Cover Photo: IBM TS3500 Tape Library David

Hobby, USA, aka Strobist, www.strobist.com

Foreword by the author Kurt Gerecke

Dear customers, dear business partners, dear readers,

embark on an exciting journey through time through the history of IBM storage. In front IBM invented the first external storage unit in the form of a roller conveyor for 58 years the IBM 726 with a storage capacity of 1.4 MB. Let yourself be fascinated by what's in this half century of inventions and products from IBM in the storage environment and has been brought onto the market. You may want to go through this book also remember your early days in IT and mentally your professional career review. But it is also for all young people who have not experienced this period of time directly fascinating reference book.

The periods described clearly reflect the holistic IBM development strategy: computers and storage systems Teme belong together as a unit! Storage technologies and storage solutions have always been in line with advanced the server technologies to provide balanced, high-performance and coordinated overall solutions To make available. This strategy spawned many new inventions and made IBM the global patent leader - for 17 years in a row.

Three years ago, the Nobel Prize in Physics went to Peter Grünberg from the KFA in Jülich and Albert Fert from the University in Paris, which discovered the GMR effect (Giant Magnetoresistance) at extremely low temperatures in 1988. We all know that without this discovery, today's hard drive capacities would never have been possible. What a lot not knowing is that it is thanks to IBM that the discovery became a product. The IBM started back in 1989 Researcher Stuart Parkin in the IBM Almaden Laboratory in California with the implementation and goal of achieving the effect even with normal ones Realizing temperatures. He researched more than 30,000 material combinations, in 1997 the first usable GMR To provide read heads for disk drives. All three researchers worked together in 1997 for GMR development awarded the Euro Physics Prize.

Maybe not everything, but everything fundamentally important was discovered and developed by IBM in the storage area. I would like In addition to GMR, give you a few more key data: In 1991, IBM introduced three pioneering technologies, thin-film layers on the disks, the first MR heads (magnetoresistive heads) and RAID5-based subsystem architectures. In 1999, IBM discovered the paramagnetic effect and in the same year, IBM launched the first microdrive Drive the size of a two euro coin. In 2000, IBM was the first company to receive the “National Medal of Technology for Innovations in Storage ”by the American government, one of the highest technological honors to date was finally awarded to individuals. All of these fascinating inventions and many more can be found in techno

logieAnnex to this storage compendium. IBM performed in the field of memory development
Unique and I am convinced that this will also be the case in the future.

Pay particular attention to the current era, especially through virtualization
in all areas of data storage. All current IBM products are described there in detail.
The compendium is therefore not just a historical reference work, but above all it informs you about everyone
current IBM memory developments.

The revision of the IBM System Storage Compendium as Update 2010 again required a lot of free time
and many weekends. So first of all I want to go back to my wife Gabriele for that
thanks for the understanding. Her patience was once again admirable.

Many thanks also to the IBM Storage product team for their valuable information. Dr.
I would like to thank Stefan Radtke for his technical implementation of the XIV storage system, Mr. Jens
Gerlach from LSI for his DS5000 input, Mr. Stefan Neff for his TS7700 contribution, Mr. Hartmut
Reason for his help with the IBM Information Archive, Jürgen Loeb for his SOFS or SPSC input and Axel
Melber for the photo editing of the exhibit collection.

I would particularly like to thank Hans W. Spengler and Heinz Graichen from the IBM Museum in
Sindelfingen for their tabular compilations, graphics and pictures of the now finished in the museum
thank you plate rotunda and the interview with me, both of which detail the RAMAC 350 access
set out. I would also like to thank both of them for the plate profiles, especially Hans
W. Spengler has put together ten years of work and was able to integrate this new compendium.

All other colleagues who helped me with this and who are not mentioned by name now apply the same
big thank you.

Many thanks also to Mr. Ivano Rodella from IBM System Storage Marketing, who made the agency work possible,

as well as the sponsors who financially supported the printing of this compendium.

I am most pleased that I am Messrs. Werner Bauer and Klemens Poschke, all long-time colleagues and

Friends, could win to work on this compendium. Werner has the topic of the system

managed storage (system managed storage) in the mainframe environment. Klemens has it

Documented the history of Tivoli Storage Manager (TSM), which is celebrating its 20th anniversary this year.

I would also like to sincerely thank Edelgard Schittko, who simply added this at very short notice

agreed to work on the topic of BRMS in the System i environment. The compendium now reflects one

holistic documentation of the most important storage hardware and software products.

I would also like to do a little advertising! If you are in the Stuttgart area, forget it

Not you, the IBM Museum, the house of the history of IBM data processing, Bahnhofstrasse 43 in 71063

To visit Sindelfingen. There you can "live" the history of information technology from 1890 to

Experience the year 2000. Also the first disk product from IBM, the RAMAC 350 as part of the former computer

stems IBM RAMAC 305, you can see in operation there. A special highlight is the newly built one

Plate rotunda. Accompanied by a slide show you can go on a journey through time and the fascinating

Get to know products and inventions in the field of plates from 1956 to the year 2000.

Register in time for an appointment via me, Tel. 01702207066, or Mr. Hans W.

Spengler from the IBM Museum, Tel. 07031271378.

I hope you enjoy reading and studying the IBM System Storage Compendium

and hope you have many opportunities to successfully use this information for your purposes.

With best regards

Yours Kurt Gerecke

A few words of personal reference ...

The idea for a historical documentation of Tivoli Storage Manager Soft goods were created in conversation with my friend and colleague Kurt Gerecke at the CeBIT 2009, when he extended an extended edition of this book for the 100th year IBM Germany 2010 was planned. I am happy with a TSM Chapters contribute to the success of the book because of the story of Tivoli Storage Manager and its naming processes closely related to history IBM's storage technology. In this respect, this chapter completes Kurt's Gerecke and colleagues wrote the hardware part, but also documents the tradition of the successful common a lot of work between people from different areas of IBM.

This compilation about the history of the IBM Tivoli Storage Manager is also a reflection animals of the last 20 years of my own professional history at IBM, since I have been dealing with this since 1990 Deal with the topic as an architectural consultant. When it comes to collecting the material and preparing the content, always happy to be reminded of many personable people that I talk to customers about this topic, Got to know business partners and as colleagues.

The chapter is deliberately limited to the subject areas in which TSM and its predecessors are part of IBM data protection / archive solutions. I purposely have information about company information and accompanying products of the backup market to include the historical development and To describe market importance holistically.

20 years of working on this topic allows me to use a large amount of resources, who helped provide the following information, documents and historical images. On behalf of all contributions I would like to thank the following colleagues especially for the overseas Background information, documents, stories and screenshots:

Jim Smith (San José), Frank Albert Müller (Mainz), Chris Zaremba (Endicott), Edward Gretz (West Chester), Norm Pass (Almaden), Barry Fruchtman (Tucson), Dr. HansJoachim Renger (Böblingen), Claire Rankin (San José), Mike Welkley (San José), Paul Bradshaw (Almaden).

Special thanks go to my colleagues Wolfgang Hitzler (Ehningen) and Dr. HansJoachim Renger (Böblingen), who sacrificed their time to validate the content of the text, and my wife Doris, who took care of the subject has adopted Chinese on a grammatical / stylistic level.

I would also like to thank the editors of the magazine "Computerwoche" (Sandra Holleber) and the Online magazine "Monitor" (Dipl.Ing. Rüdiger Maier) for approval, article extracts from their publications to be allowed to use the Internet.

If, despite intensive research, content errors have crept in, I apologize and to inform me about it (Klemens.Poschke@de.ibm.com).

I hope you enjoy reading and studying the 20-year TSM success story and

Best regards,

Your Klemens Poschke

General:

The history of IBM storage technology and storage systems is almost certainly one of the most fascinating events in our history over the past 58 years. Exciting from the beginning to the present day, this brochure is intended to provide an overview and, above all, to develop an understanding of how inventive mankind can be when it comes to crossing technological boundaries in order to open up new possibilities for storing data. The multitude of products and inventions in this area make it necessary to organize the history a bit and to divide it into contemporary history. Since this endeavor has never been done before, the author allows himself to divide the storage events from 1952 to the present day into technological epochs, to then examine in detail the individual epochs from the product and product development side. After the product eras, the chapter "IBM Storage Software" follows, which represents the most important software developments in the storage environment. At the end of this book in the “Technology Appendix” section there is a technology section that describes the development of storage-based technologies and recording methods, including upcoming technologies.

Development of storage systems from 1952 to 2010

The individual epochs always describe a time phase that was characterized by special products that came onto the market in this time phase. In many cases, the products were actively sold beyond the defined epoch and were mostly withdrawn from marketing only in the subsequent epoch or later. The individual products are described in the epoch in which they were launched in such a way that they represent the entire life cycle.

The Storage Compendium does not claim to be complete, but describes the most important products in the respective epoch.

Price development of magnetic disk storage

The speed of technological development in magnetic disk storage from 1956 to 2010 can be seen in particular from the drop in prices for a stored megabyte. The process was something like this:

year	1956	1964	1975	1987	1991
Euro / MB	12,000	8,000	70	12	9

year	1997	2000	2006	2009
Euro / MB	1	0.25	0.015	0.0054 (FC disk) 0.0025 (SATA)

If you add the reduced space and energy requirements as well as the lower maintenance costs per megabyte, the trend in cost development is rather more favorable.

Capacity development of magnetic disk storage

While the price decline for the stored megabyte continued, there was an over dimensionally high price decline in the past 10 years, which has reached a range of 40 to 70% per year, especially in the past years. At the same time, the recording density was increased to such an extent by technological innovations that it could withstand the fall in prices for the stored megabytes and take them into account.

The increase in recording density was as follows:

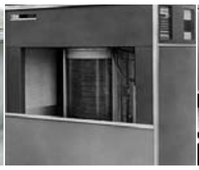
year	product	Recording density (million bits / qcm)
1956	RAMAC 350	0.00031
1961	1311	0.008
1964	2311	0.017
1965	2314	0.034
1970	3330	0.12
1973	3340	0.26
1975	3350	0.48
1979	3370	1.2
1980	3375	1.51
1980	3380	1.9
1985	3380E	2.99
1987	3380K	5.47
1991	33903	13.95
1994	33909	41.6
1996	RAMAC 3	129.2
1999	ESS	427.5
2004	ESS	2.9 billion bits / qcm
2006	DS8000	5.8 billion bits / qcm
2009	DS8000FC plates DS8000SATA plates	8.7 billion bits / qcm 19.3 billion bits / qcm

In 2003, with the 146GB drives available for ESS, the magic limit of 1 billion bits on the square centimeter was exceeded. With the current availability of 1 TB SATAL drives, a recording density of over 19 billion bits / qcm can be achieved. The big leaps have occurred especially in the last 10 years (see also technology appendix).

1952 to 1961

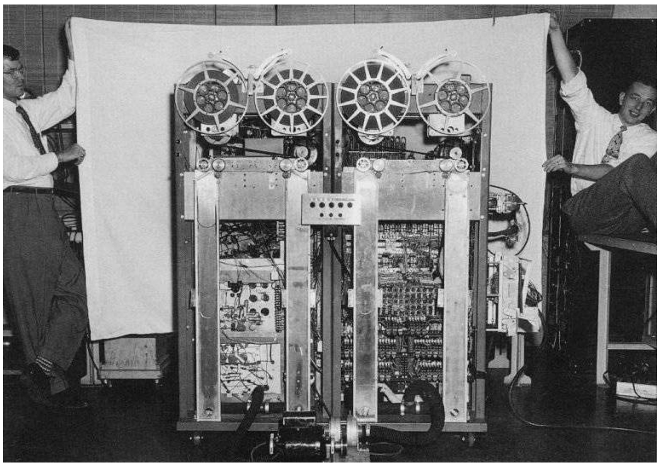
The early days of electromagnetic storage







The early days of electromagnetic storage



Recording of the first prototype of the IBM 726 magnetic tape system in 1951 with two tape drives per unit

For the first time in memory history, external storage with magnetic recording from IBM **1952**
offered: the **IBM magnetic tape unit IBM 726** with a storage capacity of 1.4 MB on a 12 inch roller tape (reel diameter) with a length of 720 meters. The unit had **two integrated tape drives.** However, this proved to be extremely impractical because neither drive was still available in the event of a fault or during maintenance work. Just a year later, **1953,** followed

the IBM 727 with one drive per unit, which offered a capacity of 4 MB on a standardized roller belt with a belt length of approx. 740 meters. Until 1956, part of the control was housed in what was then the IBM 701 computer unit. The first dedicated tape control units were then released **1956.**

Medium roller conveyor for IBM 726

IBM model 726 magnetic tape unit, launched in 1952

1956 that came **IBM RAMAC 305 system** on the market. This system had one for the first time **Disk storage unit type IBM 350.** The IBM 305 system was announced under the registered trademark 'RAMAC' (Random Access Method of Accounting and Control). The computer of the overall system controlled both the positioning of the read / write head on the desired disk surface and disk track and the transfer of the data in both directions (write and read).

The recording density of the RAMAC 350 disk was 100 bits per inch (40 bits per cm) and the disk stack rotated at 1200 revolutions per minute (1200 RPM). The average search time was 600 ms. The disk stack consisted of 51 disks with a diameter of 24 inches and offered a maximum capacity of 5 MB (5 million characters). The RAMAC 350 can be viewed in operation at the IBM Museum in Sindel fingen.

IBM 350 RAMAC plate, 51 plates in a 24 inch diameter stack. 1956: 5 MB model (5 million 'characters'), 1958: 10 MB model (10 million 'characters'), withdrawal from sales in 1960

IBM RAMAC 350 disk stack

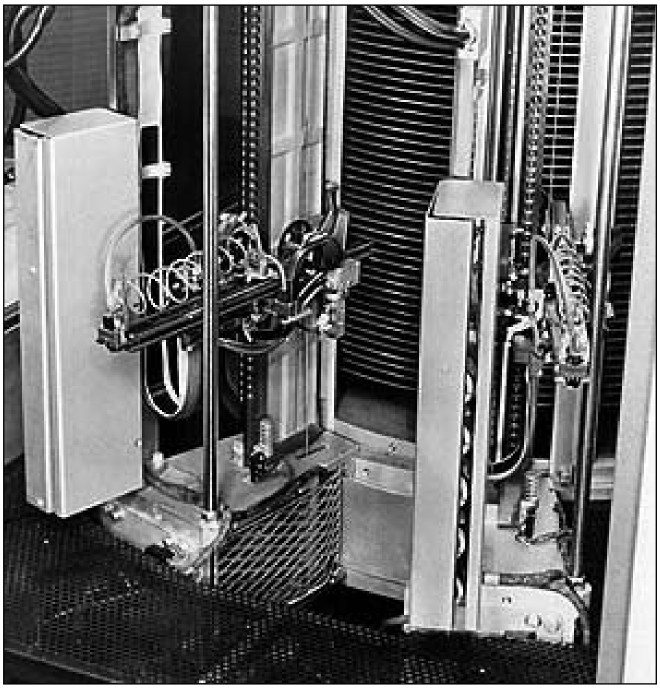
At that time there were no access combs, just an access arm that was controlled into the interstices of the plate stack with a cable pull mechanism. The exact positioning was carried out with a rack and pawl and was electronically controlled by compressed air. The heads hovered over the plate at a distance of 800 micro inches. In 1958, with the announcement of Model 2, the possibility was created to represent double the capacity (10 MB). In this model, two access stations were used, so that one station could read or write and the second could search at the same time in order to prepare the read / write head for the processing of the next sentence.

In terms of sales technology, the same marketing approach was already used at that time with the RAMAC, which can still be found today, but in other sizes. How do you make it clear to the buyer and user what 5 MB is? At that time, 5 MB was beyond normal imagination. For this reason, 5 MB were defined in most specification documents so that this corresponded to 60,000 - 80,000 punched cards, or 2,000 feet of tape length (the tapes were already 4 years earlier) or 940 pages with printed paper.

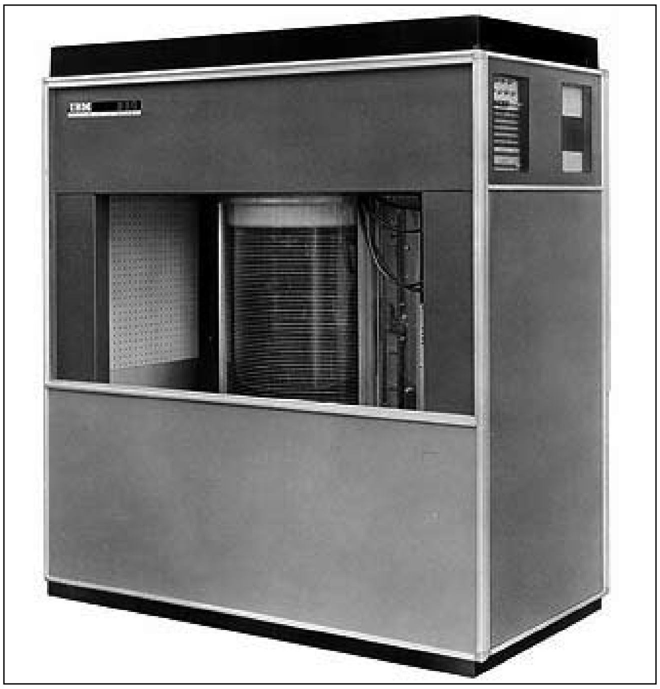
In addition to the RAMAC plate stack and the associated access station, the unit 350 contained an electronic and pneumatic control for the access station and an air compressor. The lubrication of the plate bearings was carried out with an oil vapor lubrication.

The unit 350 had a length of 152.4 cm (60 inches), a height of 172.72 cm (68 inches) and a depth of
73.66 cm (29 inches) and weighed about a ton.

The entire system, consisting of the computer, disk unit, card reader, card punch, printer and console unit, cost $ 185,000 and was also offered under leasing at a monthly rate of $ 3,200. In the period from 1956 to 1960, 1,500 systems were installed worldwide. In 1960 all models were withdrawn from sales and production ceased in 1961.



IBM RAMAC 350 disk with two access stations



IBM RAMAC 350 unit

1952 to 1961

The early days of electromagnetic storage

First delivery of the IBM RAMAC 305 in Brazil in 1956

1957 was the **'Data Synchronizer'** introduced on the IBM System 709, the forerunner of today's transmission channels.

1959 IBM introduced a new program package, the input / output control system **IOCS** for magnetic tapes and magnetic disks (**Input / output control system)** a. These were macro instructions for creating, retrieving and processing files, and for correcting errors. IOCS also controlled data block formation and deblocking. IOCS remained essential for later systems.

1958 IBM used the technology of automatic program interruption for the first time with the systems IBM 7070 and 7090 (**Program interrupt).** Under certain conditions indicated by input / output units, the processor automatically branched to routines which handled the 'interrupt'.

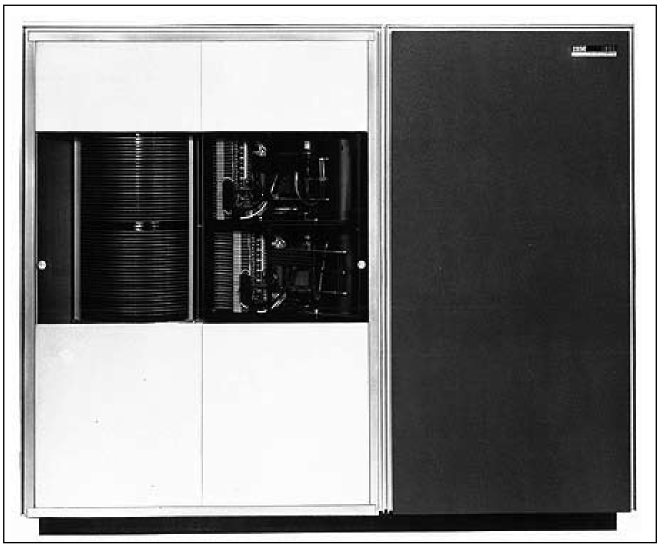
1958 was the **IBM 729 tape unit** launched as successor to the predecessor IBM 727. In addition to higher data rates and higher storage density, the IBM 729 tape family offered the first write control. With the introduction of the IBM 729, it was automatically checked when writing to magnetic tape whether the information just written was valid. For this purpose, a check register was read in and checked point by point.

In October 1960 was the **IBM 1405 disk storage** announced that was operated on the systems IBM 1401, 1410 and 1460 (not on System 7000). This disk storage quadrupled the capacity of a disk stack compared to the RAMAC from 1956. Both the number of tracks per

'Inch' and the number of bits in a track were doubled, so that four times the capacity of the disk stack was achieved. The plate stacks were available in two versions, one with 25 plates in a stack and one with 50 plates. The smaller unit had a capacity of 10 MB and the larger one of 20 MB. The recording density was 220 bits per inch, which corresponds to 40 tracks per inch. The hovering height between head and plate was 650 micro inches and the plates rotated at 1,800 revolutions per minute. The data rate was 17.5 kilobytes per second. The IBM 1405 disk storage was mainly used for the IBM 1401 computer.

A year later, in **June 1961,** was the **IBM 1301 disk unit** announced. It became a year later, in the 3rd quarter
1962, delivered to customers. The development of this plate unit contributed greatly to the further development of future plate systems and is as **milestone** to look at throughout the record history. The 1301 plate system paved the way for future technologies to be used in two areas: the **Air Bearing Slider** Technology and **separate read / write heads for each plate surface** with the option of electronically selecting all the heads physically set on a data cylinder for the relevant disk surface and then one after the other

IBM 1405 disk storage, announcement October 10, 1960, withdrawal June 30, 1970



IBM 1301 disk storage device, announced in June 1961, all models withdrawn in October 1970

IBM 7340 Hyper Tape Drives in the center of the picture, on the left the IBM 7640 control unit, on the right the IBM 729 tape units

to edit. This laid the cornerstone of the cylinder principle for disk storage and that **'Plate cylinder'**
born!

In addition to higher capacity and performance, the machine offered a much more flexible and higher transmission bandwidth thanks to the principle of having a read / write head per disk surface, and was adapted to the new IBM 7000 series of computers (7070, 7094, 7080 and 7090) . The capacity increase for the first RAMAC was a factor of 13. The plates rotated at 1,800 revolutions per minute. By reducing the flight distance between head and disk to 250 micro inches, 50 tracks per inch and up to 520 bits per inch could be recorded. The Model 1 of the IBM 1301 had a disk module with a capacity of 28 MB, the Model 2 worked with two modules and a capacity of 56 MB. Connected to the IBM 7631 control unit, up to 10 modules (or up to five 1301 units) could be used, they offered a maximum capacity of up to 280 MB for the 7000 series. The prices at that time were also interesting: The Model 1 could be leased for $ 2,100 a month or for a sum of

$ 115,500 can be bought. The Model 2 with two modules was priced at $ 3,500 per month and was priced at $ 185,500.

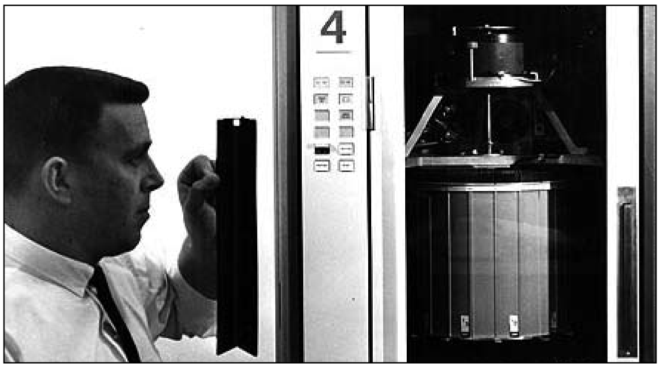
The **Hyper Tape Drives IBM 7340** reflected the latest tape technology when they **1961** have been launched on the market. Together with the control unit **IBM 7640** they operated the IBM computers 7074, 7080 and 7090. The Hyper Tapes achieved much faster reading and writing speeds compared to the previous tape units. Connected to the IBM 7090 computer, they delivered double the transfer rates compared to the IBM 729 tape system, the further development of the IBM 727.

At the **April 7, 1964** was with the **System / 360** the **Magnetic stripe memory IBM 2321** announced. At that time, with its total storage capacity from 400 MB to 1,000 plastic strips with a magnetizable surface, it was considered a mass storage device. Each strip contained 100 data tracks, had a capacity of 0.2 MB and was written in so-called

'IBM 2321 Data Cell Drives' kept. 200 such strips could be accommodated in one data cell. The data cells located on a rotating carousel were positioned by rotation so that the memory strip sought could be gripped and passed around the read / write drum. This meant that the addressed data track could be read or written. The elaborate construction and the associated high maintenance expenditure prevented a wide spread.

1952 to 1961

The early days of electromagnetic storage



Principle and mode of operation of the IBM 2321 magnetic stripe memory

IBM 2321, total capacity of 400 MB, expansion levels of 40 or 80 MB, access times from 95 to 600 ms

Commentary on the beginning of electromagnetic storage

In addition to the invention of the first external storage units with the IBM 726 roller conveyor and the IBM RAMAC 350 disk unit, the top milestone was the introduction of the IBM 1301 disk unit.

The IBM 1301 was the first disk system to work with 'Air Beam Ring Sliders' and to use an access comb that controlled every disk surface in the disk stack with dedicated read / write heads. With the IBM 1301, the cylinder architecture was introduced in the plate stack, which has retained its status to this day.

Also noteworthy is the fact that even back then people were looking for ways to store small files that were not used as often on a cheaper mass storage device. The IBM 2321 magnetic stripe memory was introduced for this purpose.

The era was also shaped by the fact that with new system announcements, the associated new disk units, which were adapted to the performance of the new system, were always part of the system announcement.

The beginning epoch laid many foundations for the subsequent epochs.

1962 to 1974

The era of the removable plates and the 'Winchester' period

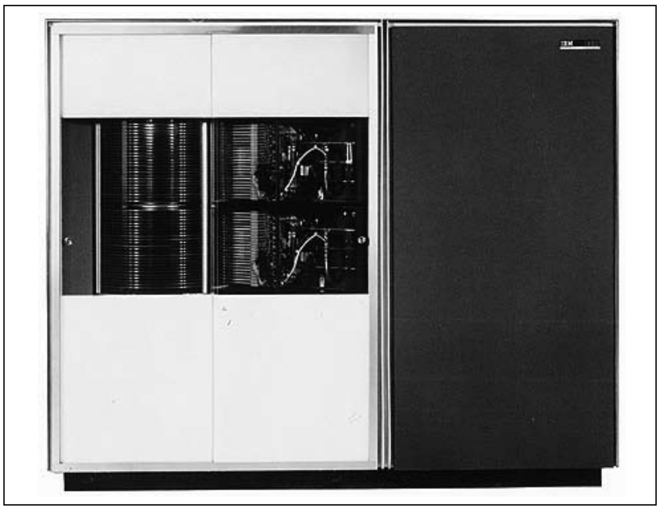
1962 to 1974

The era of the removable plates and the 'Winchester' period


IBM 1311: Announcement on October 11, 1962, withdrawal of all models in 1971

The **between 1962 and 1964** introduced **Magnetic disk storage IBM 1311, IBM 2311 and IBM 2314** to connect to the 360 systems worked with interchangeable magnetic plate stacks. The new **Disk control unit IBM 2841** was **Micro-program controlled for the first time.** This was necessary because several types of storage had to be controlled, including the newly used CKD format. IBM developed the recording format CKD (Count, Key, Data, freely translated: sector identification, key, data), which as an optimized version ECKD remained valid beyond the year 2000 and is still used in the mainframe (z / OS) environment. The 'E' stands for 'Extended'. The IBM 3330 product describes how data is accessed using the CKD method.

The first model of the **Magnetic disk storage IBM 1311** with removable plates was in October **1962 with the IBM 1440 system** announced. The removable disk stack consisted of six 14-inch disks, weighed 10 pounds, and had a storage capacity of 2 MB. Compared to the IBM 1301, the linear storage density doubled to 1,025 bits per inch. This was achieved by reducing the distance between the heads to the plate surfaces by a factor of 2 compared to the IBM 1301 to 125 MicroInches. The disks rotated at 1,500 RPMs and offered an average access time of 150 ms.



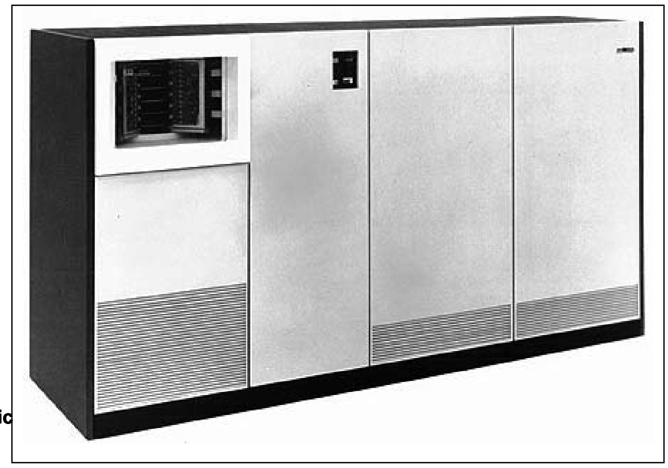
The IBM 1302 hard drive system was announced on September 23, 1963 and withdrawn from sales on February 9, 1965

IBM 2401 magnetic tape unit

1963 became another **Successor to the IBM 1301** announced with a permanently installed panels. The **IBM 1302** looked optically the same as the IBM 1301, but offered 500 data tracks on the disk surface (1301 had 250 tracks per disk surface). As with the 1301, 20 disks were used in the stack and the capacity increased from 28 MB to 58.5 MB. The Model 1 of the 1302 had a capacity of 117 MB and the Model 2 of 234 MB. The data transfer rate was increased from 90 KB / s to 180 KB / s.

1964 came **especially for the System / 360 computer family** the new **Magnetic tape system IBM 2401** onto the market, which was able to record with 9 tracks on the roller conveyor at the same time and was much more reliable than the predecessors IBM 726, 727 and 729 This was achieved by introducing a so-called CRC (Cyclic Redundancy Check), which enabled the first automatic error correction (Automatic Error Capture & Correction) and provided the basis for the later ECCs (Error Correction Codes) in tape systems.

In April **1965** - one year after the announcement of the System / 360 processor - that was **IBM 2314 disk system** announced as 'IBM Direct Access Storage Facility'. This was the first plate system in which the controller and the strand with plate stacks could be ordered and assembled in one configuration. With the 2314, a new disk stack was introduced with twice the number of disk surfaces, which represented a storage capacity of 29.2 MB. The overall system had a maximum capacity of 233.4 MB. Compared to the 1311, the 2314 was a factor of 4 cheaper in terms of the stored MB. The data rate was 310 KB per second. In January 1969, two new versions of the 2314 were launched, the A1 and A2 models



IBM 2305 Fixed Head Storage, known as 'Zeus Storage', launched 1971, withdrawal from sales in January 1980

offered a 20% faster access time, which ranged from 60 to 75 ms. The capacity remained the same. In December 1970 the plate family 2314 was supplemented by a model B1, which offered 50% higher capacity.

in the **January 1970** became the record system **'IBM 2305 Fixed Head Storage'** announced that was developed under the development name 'Zeus' and as **IBM Zeus** became better known than under the official type name. The first deliveries were made in 1971. The IBM 2305 was ideally suited for large database applications and for batch processing and was used for this type of application in large System / 360 computers of the models 85 and 195 and later in the System / 370 models 155 and 165. With small storage capacities of 5.4 MB and

10.8 MB, but with an average access time of
2.5 ms or 5 ms **the system was the fastest of its time.** The data rate was 3 MB / s and thus 10 times higher than previous disk systems.

IBM 2314 production line in Mainz 1969

IBM 2314 family of disk systems, model 1 announced on April 22, 1965, models A1 and A2 on January 10 1969, model B1 on December 14, 1970, withdrawal from sales in 1978

1970 to 1972: The virtual memory, which was mapped on magnetic disks, created a new quality in the relationship between processors and magnetic disk memories. Online applications quickly gained in importance. Like the new IBM / 370 processors, subsystems carried control units **IBM 3830** and **Magnetic disk units IBM 3330** the increasing performance and reliability requirements. New block multiplex channels ensured significantly higher data transmission rates and contributed to a significantly higher overall performance. Here are some details:

The IBM 3830 disk control unit was microprogram-controlled, but for the first time with a loadable microprogram memory that was loaded from a floppy disk. The microprogram made it possible to implement more functions than in the previous models. This included the detection and elimination of bit errors, repeated repetition of channel commands if necessary before transferring error handling to the computer and the collection of information about errors that occurred for the system log (**LogRec).** These records provided the operators with information about the technical condition of the data carriers and provided the technical field service with instructions for preventive maintenance.

In terms of the sensitivity of magnetic recording to electrical and mechanical interference, bit errors had to be expected from the start. With additional bits for parity checking, IBM developed increasingly sophisticated methods to generate redundancy bits that were attached to the stored data. Both **Recording formats CKD and ECKD** this 'redundancy' recording takes place in the gaps between the data records. With the help of appropriate programs in computers and control units, most bit errors during transmission can be corrected without any significant time delay, and data that has been falsified by technical influences can practically be excluded. The procedures require the definition of so-called 'physical records', a uniform length in bytes, for each file. Such sentences were called blocks. Whole blocks were always transferred and checked during input / output operations. In most cases it was useful to combine several 'logical records' in one block. They consisted of an order term, which was followed in 'fields', which were defined according to location and length, the associated information and information. In administrative applications such. B. Articles, customers or personnel number the order terms. The fields then contained alphabetic and numerical information. The number of logical records in the physical record was called the blocking factor.

IBM 3330, capacity 800 MB, 2-8 drives, access time 30 ms, data rate 806 KB / s, announcement model 1 on June 30, 1970, model 2 on October 4, 1972, model 11 on July 17, 1973, withdrawal from sales on December 20, 1983

According to previous hydraulic access mechanisms, the magnetic disk units used **IBM 3330** for the first time for the 'Search', the **horizontal setting of the access comb** with the read / write heads, a linear motor. As with loudspeakers, a magnetic field generated by variable current flow in a coil counteracts the static field of a permanent magnet. Hence the English name **'Voice Coil Motor' (** Loudspeaker coils motor). The position of the read / write heads at certain times resulted from the strength of the magnetic field generated by the coil. This technology required that one surface of a plate stack was provided with servo tracks that were unchangeable by the user. The (read only) reading head above provided the electronic control with the information required for targeted positioning within the scope of a control loop. At the time, this technique reduced the average duration of a search operation by half: instead of 60 only 30 ms.

IBM 3340, maximum capacity up to 4,800 MB, Winchester removable disks model 35: 35 MB, model 70: 70 MB, up to 64 drives (models 158, 168), data rate 885 KB / second, access time 25 ms

The following was important for an improved process of direct access to data: The IBM 3830 control unit was the first to use the information from the servo plate to recognize the rotational or angular position of the magnetic disk units connected to it (**RPS - Rotational Position Sensing).** With the IBM 3830/3330 disk system, routine access was as follows: According to information from the data carrier table of contents, the so-called

VTOCs (Volume Table Of Contents), and the indexes to the files, the computer generated a channel command using the access method control programs. This contained the disk address from which to read or write, by unit, cylinder, track and sector. The concentrated traces on the plate surfaces were divided into sectors. The cylinder was the sum of the superimposed tracks on a stack of disks, of which, with a given horizontal setting of an access comb, read / write heads could also be read or written on top of one another. The control unit converted the channel command into specific commands for the units. The units carried out the search command to set the access comb on the correct cylinder independently and reported the completion of the search operation to the control unit. The control unit then activated the correct read / write head, analyzed its position in relation to the sectors and began to transfer the data via the

Channel of the computer as soon as the beginning of the sector specified in the channel command was reached. The channel was only occupied during data transmission. Multiple disk units could overlap search commands. The rotation waiting time (at that time around 17 to 20 ms) did not burden the channel either. This method allowed higher system-effective data rates of the magnetic disk system, which, with favorable configuration and application profiles, was not very far below the transmission capacity of the units.

1973: The new one that can be directly connected to medium-sized computer systems IBM / 370, models 115, 125 and 135 **Magnetic disk system IBM 3340** used that as a removable disk

IBM 3348 data module. The interchangeable 3348 data modules offered 35 or 70 MB capacity. One module contained the stack of disks and a horizontally movable access comb with two read / write heads per data surface in a closed, dust-proof housing. The module was placed on the shaft of the drive motor in the drive. After the automatic opening of a lateral blind, electrical plug connections connected the access comb to the drive and a mechanical connection connected the access comb to the linear motor. The reason for this construction: If interchangeable plate stacks were used as data carriers, this meant extreme demands on the manufacturing tolerances, especially for access combs, because every stack had to fit any drive and even minimal horizontal deviations of the heads above the tracks led to reading errors. This set limits on horizontal track density and reduced the security and reliability of the systems. A maximum of four modules with a maximum capacity of 280 MB could be operated via a direct connection to the IBM / 370 computer.

1962 to 1974

The era of the removable plates and the 'Winchester' period

Construction of the IBM 3340 Winchester removable plate stack

The IBM 3340 with the models A2, B1, B2 and C2 were announced on March 13, 1973. Models B1 and C2 were withdrawn from sales on December 20, 1983 and models A2 and B2 on May 1, 1984.

The fact that the data access module of the IBM 3340 magnetic disk memory always wrote and read the same access comb enabled higher track densities with increased reliability. The overall design resulted in shorter search times and somewhat higher transmission performance compared to the IBM 3330.

With the new ones **Magnetic tape units IBM 3420, models 3, 5 and 7,** who worked in 9-track technology with a recording density of 1,600 bytes per inch, a data rate - depending on the model - of 120, 200 and 320 kilobytes per second could be achieved. These models were introduced in 1970.

The new magnetic tape units **IBM 3420, models 4, 6 and 8,** the **1973** have been introduced on the market, the process of 'group coded recording' increased the recording density of

1,600 to 6,250 bytes per inch and the maximum transfer rate - depending on the model - was increased to 470, 780 and 1,250 kilobytes per second. At the same time, reliability increased by a factor of 7 (compared to models 3, 5 and 7).

To make operation easier, IBM announced **1971** parallel to the 3420 magnetic tape system **IBM 3410 magnetic tape system** which had the shape of the desk and for the first time allowed the roller belts to be inserted horizontally from above. The capacities and technology were comparable to the 3420, but the read / write speed was significantly lower. The 3410 was therefore offered - in addition to the 3420 high-performance magnetic tape system - as a cost-effective alternative for users of the IBM System / 370 and IBM System / 360.

1974: In the course of the increased transition from the direct processing started with the system / 370 with inexpensive terminals for program development and problem solving (VM / 370, APL), the number of files for users of large systems grew dramatically. In contrast to the database applications (IMS, CICS) that work with a few large files and sophisticated backup procedures, it became more and more difficult, especially the many small ones, irregularly

IBM 3420 magnetic tape unit, 1970 models 3, 5 and 7, 1973 models 4, 6 and 8

IBM 3410 magnetic tape system with horizontal tape insertion option secure and archive used stocks under manual management. In order to solve this task and to enable more cost-effective storage of data under system control, IBM introduced new hardware and software.

The mass storage **MSS (Mass Storage System) IBM 3850** depicted the contents of a data carrier of the IBM 3330 magnetic disk storage on two special cylindrical magnetic tape cartridges. You can use it as

first virtual 'disk storage' designate, which at the same time represented an automatic tape archive. All the active cartridges were stored in a cupboard with shelves opposite each other in honeycomb cells. When requested by the programs, a sensible design automatically moved a so-called picker to the correct cartridge. In contrast to later mechanical grippers, the picker held the data carrier electromagnetically and brought it to a read / write station, from which the desired data was transferred to certain magnetic disk drives (staging devices) for processing or vice versa. After the operation, the same device returned the cartridges to their storage location. The operator was only concerned with the cartridges, which the mass storage device controlled according to program specifications via corresponding stations or which it requested from outside. The mass storage could hold up to 236 GB in the access of the processors. The technical development of the magnetic storage - disk, tape - did not allow a meaningful successor, while the associated software has remained up to date in 2006.

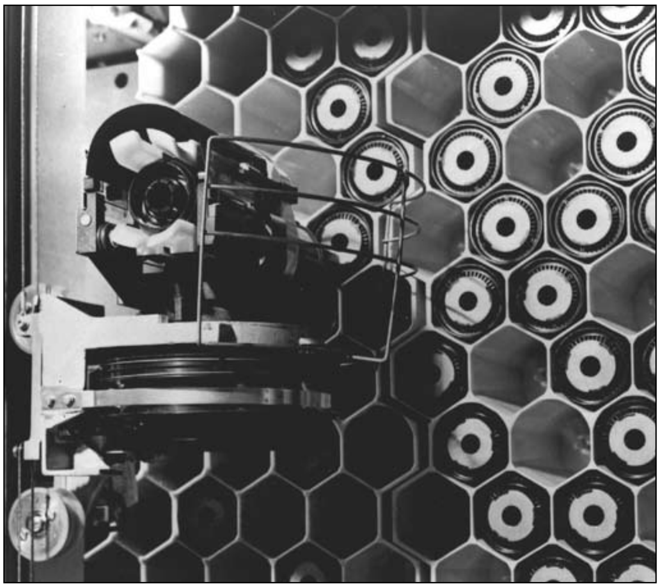
IBM MSS 3850 in the data center in the background on the right

ILM, Information Lifecycle Management, that is on everyone's lips today was already lived intensively in 1974! At that time, the HSM (Hierarchical Storage Manager), as the administrator of the storage hierarchy, was a program that, in connection with the mass storage device, automatically backed up the files assigned to it in accordance with the specified control criteria and restored them if necessary. Another function that was already available at the time was to move files from expensive disk storage space to cheaper storage (originally to the MSS) during times when they were not needed. In order to make optimal use of the storage space, the HSM was able to compress files to be archived by omitting empty areas on the data carriers, and it was also able to compress data using binary arithmetic methods (compression, compaction).

IBM Mass Storage System 3850 with 'Honeycomb Racks' 1974

Adapted to a wide variety of configurations with different external memories, the HSM remained a permanent part of later packages of service programs for the area of external memories, e.g. B. the IBM DF products (Data Facility), which aimed at the system-managed storage. For example, the DF products were introduced as program products in the main frame environment, starting with DFP, in 1989. The DFDSS and DFHSM followed shortly afterwards, all of which were then combined under the DFSMS (Data Facility System Managed Storage) program package in 1992 and were fully integrated into the MVS operating system a year later. To this day, HSM still plays a very important role under different names.

The era of the removable plates and the 'Winchester' period



IBM 3850 robot with double gripper (standard) in action

IBM 3850 tape cartridges with 19.5 meter tape length and a capacity of 50 MB each in size comparison to the 3330 plate stack with 100 MB capacity

The MSS 3850 was available in eight model variants. The smallest model stored up to 706 cartridges, had a capacity of 35.3 GB and could be expanded to the largest variant with up to 4,720 cartridges and a capacity of 236 GB.

Commentary on the epoch of removable plates and the 'Winchester' period

The strong increase in computer performance due to new systems made it necessary to introduce many new things on the memory side in order to meet the new performance requirements of these computers.

With the IBM 2841 and IBM 3830, the first micro-program-controlled control units and the CKD format that is still valid today (today ECKD at zSeries) were introduced. The newly introduced block multiplexing channels ensured higher transmission bandwidths, which made a significant contribution to improving the overall performance of the computer and memory peripherals.

Smaller form factors and smaller flight heights between the head and the surface of the plate have been introduced to improve performance through faster speeds, data rates and capacitive capabilities. Almost in parallel, the performance of the tape systems was adjusted and the data rates were increased by new possibilities of multi-track technology.

Like the early days, the era was marked by new system announcements, in which the appropriate plate and tape system was always part of the announcement. Storage announcements that were not detached from this did not take place.

The use of the removable disks used at that time was very popular with the end users because these portable disk modules could also be used for data exchange between different locations. The 3420 models newly announced in 1970 and 1973 were used exclusively for backup purposes. It was only in the subsequent epoch, in which fixed panels were used again and the principle of interchangeable panels was abandoned, that the 3420 experienced a real heyday. The roller conveyor was then the only medium that could be used as a removable data medium.

As the capacitive requirements were increasing and new mass storage devices with higher capacities were required in addition to online storage, one of the milestones of this era was certainly the IBM 3850 Mass Storage System, the world's first nearline robot system with tape cartridges, which, like the removable disks at that time, were addressed as online storage devices . The fact that ILM was introduced for the first time via an HSM functionality has to be emphasized. This HSM functionality laid the foundation for the future DF products (data facility) in the mainframe environment.

The era of permanently installed panels with external control units

The era of permanently installed panels with external control units

IBM 3350, capacity up to 2,536 MB, 317 MB per LW, 8 drives per 3350 line, access time 25 ms, data rate 1,198 KB / s, left rear: display unit for microfiches, right rear: control console of the system / 370 model 168

IBM 3310 Direct Access Storage, disk system for IBM 4331 computers

1975: With the **Magnetic disk storage IBM 3350** the IBM went back to permanently installed plates. This was a crucial step towards higher reliability and availability as well as higher recording density. Compared to the IBM 3340 type working with data modules, there were no longer any potential defects in this design, which accounted for 70% of the necessary technical interventions. In fact, the IBM 3350 technical field service recorded an average of less than one intervention per unit per year. The IBM 3350 and the successor products continue the strand concept started with the IBM 3330. Up to three auxiliary units could be connected to a head unit (master unit) with two drives. So one strand consisted of eight drives. The head unit basically had two inputs / outputs to the IBM 3830 control unit. The two connections could lead to different 3830 units. If a channel, controller, or controls in the head unit failed, operation on the second connection to each drive continued, albeit with less power. In this new concept, you can **first entry into concepts of fault tolerance**

detect.

Because with the magnetic disk system IBM 3350 with permanently installed disks, the magnetic tape reel is now the only removable storage medium for data backup, offline

Archiving and data medium exchange remained, the IBM 3350 triggered a significant increase in sales of magnetic tape units, especially for the fast models 6 and 8 of the magnetic tape system 3420. From then on, magnetic tapes were an indispensable component of larger and also medium-sized data processing systems.

In 1979, IBM introduced another disk system, the **IBM 3310 Direct Access Storage,** which offered high performance in a compact design at a very reasonable price and was used in connection with IBM 4331, 4341, 4321, 4361 and small 4381 computers. Each disk stack had a total capacity of 64.5 MB and the entire disk string 258 MB. Up to four disk strings could be operated on the above computer models and offered a total capacity of 1,032 MB.

1979 to 1983: IBM 3370 and 3375 were newly developed magnetic disk units on the new one **Control unit IBM 3880**

for medium-sized systems at that time. They worked with fixed blocks of 512 bytes for the sake of simplification and acceleration of program sequences. The user defined longer block lengths with a multiple of 512. The first FBA (Fixed Block Architecture) plate was born.

At that time, IBM 3380 disks and 3880 control units filled entire data center halls

The **new IBM 3880 control unit** introduced a new control principle, which in 1985 resulted in the first integration of buffer memories, initially from 8 to 16 MB. These small buffers were designed for paging and swapping pages and initially had an impact on the performance of older configurations without extended memory. The use of larger buffer memories then led to the introduction of new caching algorithms. The data to be written were temporarily stored in the control unit on semiconductor memory cards and duplicated on a current-independent write memory. The control unit then reported the completion of the operation to the computer, although the data had not yet been written to the disks. This operation was called 'DASD Fast Write' at the time, because this principle has significantly improved the response times between the disk systems and the computer. This 'caching' principle for disk subsystems, which was introduced at the time, was then quickly developed and is still valid today.

1981: IBM 3380 were **Magnetic disk units,** which were optimized as a new development for the 3880 control units and continued the CKD format. The IBM 3380 transferred at 3 MB per second and used special capabilities of the block multiplex channels (data streaming). The new run works came just in time, because at that point it had become clear how necessary it was, given the close relationship between processors and disk storage, to match the performance of the two.

1981 the standard models AA4 and B04 of the IBM 3380 were announced. **1985** the extended models AD4, BD4, AE4 and BE4 followed. **1987** the fast AJ4 and BJ4 and the fast, large-capacity models AK4 and BK4 were announced. Likewise took place **1987** the announcement of the 3380 CJ2, a model in which the control unit was integrated in the plate housing.

IBM 3370, 1 disk stack per unit with 2 access arms and 2 addresses, model 1: 571 MB, model 2: 730 MB, each per unit, up to 4 units with 2,285 - 2,920 MB capacity, access 20 ms, data rate 1.86 MB / s

The era of permanently installed panels with external control units

The time of the 3380 plate brought great progress in response time behavior, throughput and in the high capacities for the then operated IBM computers 3031, 3032, 3033, 3042 and / 370. Depending on the model, the average search times ranged between 12 and 16 ms. Two drives were always housed in one housing unit. Capacities of 1.26 GB (JModels) and

3.78 GB (KModels) shown. This meant a total capacity of 20.16 GB for a fully expanded 4-way line of J models and a total capacity of up to 60.5 GB was achieved for the KM models.

The plates were built in 14 inch format. In order to lift the important plate stacks out of the housing units, specially designed lifting trucks were used by the technical field service.

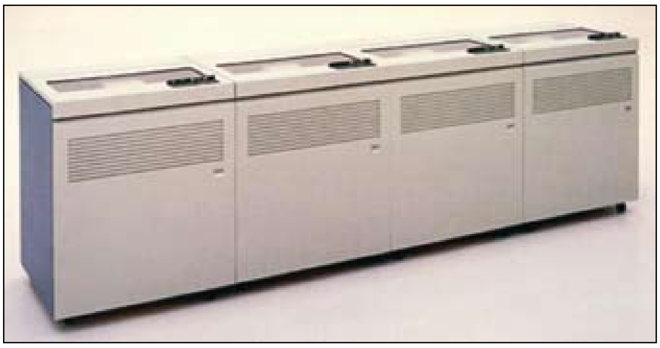
The function was implemented in the 3880/3380 subsystem for path redundancy reasons **'Dynamic Path Reconnect'** introduced, which in the event of a path error ensured that the data transfer to each HDA of the subsystem could take place on an alternative path. This multi-path architecture in conjunction with the new mainframe channel subsystem ensured 30% higher throughput in the storage subsystem.

Physical specification of 3380 disk subsystems: models

	FROM	D	E	CJ2	J	K
Actuators per unit	4th	4th	4th	2nd	4th	4th
Cylinder per device	885	885	1770	885	885	2655
Tracks per device	13275	13275	26550	13275	13275	39825
GB per unit (box)	2,520	2,520	5,041	1,260	2,520	7,562

IBM led on the magnetic tape subsystem side **1983** the last roller conveyor system with **the IBM 3430** one that was used primarily with the IBM 43xx and IBM 303x Series processors, but also with the Midrange System / 38 at that time. The compact design allowed users to set up on a much smaller area compared to the IBM 3420. The roller belts were inserted horizontally from above, as with the IBM 3410 (1971).

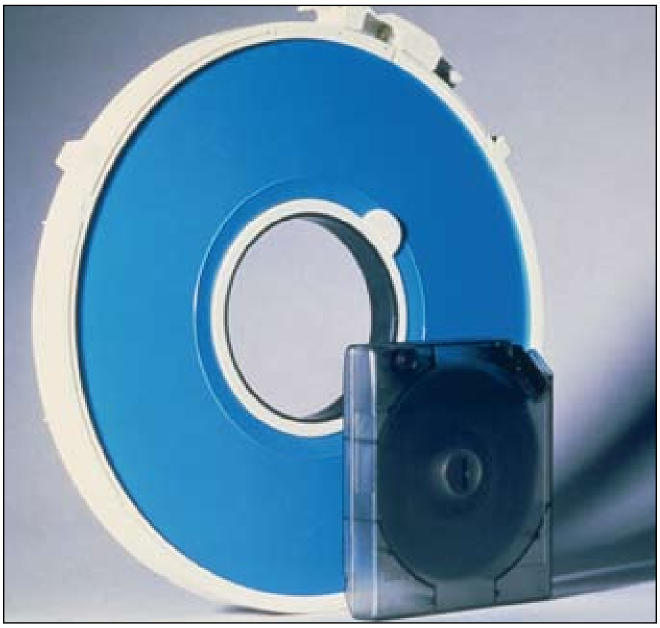
Due to the new IBM 3480 cassette technology, which was announced a year later as a replacement for the roller belts, 3430 units were not sold in large quantities, but would nevertheless remain on the market for quite a long time in order to meet the data carrier exchange requirement that continued in the following years Roller tapes was made.



IBM 3430 magnetic tape subsystem, last IBM roll, market launch in 1983

1984:

The dynamic achieved in the performance of 3880/3380 disk subsystems and / 370 computers also required a performance adjustment on the backup side, **the magnetic tape systems.** With the **IBM 3480** the change from the previously used roller conveyor systems to the cassette technology was initiated. The new IBM 3480 magnetic tape system transferred at 3 MB per second, just as fast as the IBM 3380 disk storage, thus closing the gap in overall system behavior. The 10.5-inch rolls of 3420 technology that had been used until then were replaced by palm-sized, rectangular cassettes. With a recording density of 38,000 bytes per inch, the new cartridges recorded twice as much data as conventional reels, i.e. up to 200 MB. With the addition of a storage chute (stacker), several cassettes could be automatically preloaded one after the other. This made it much easier to back up large amounts of data.



IBM 3480 magnetic tape system

1984: Change from classic roller conveyor to 3480 cassette technology

1989:

With the **IBM 3390 disk technology** the form factor used with the IBM 3380 has been reduced from 14 inches to 10.8 inches. Initially, the 3390 technology was still produced with the conventional 'brown coating', where iron oxide particles (rust particles), which were distributed as homogeneously as possible in a synthetic resin mass for uniform magnetization, were used as magnetization carriers. The introduction of a new coating, its development **1989** was initiated and the **1991** used, allowed the use of a new read / write technology. The new heads - like their predecessors - wrote inductively, creating magnetic fields in the coating of the plates with the help of a tiny gap in the metal core of miniaturized electromagnets. Up to this point the electromagnets were also used for reading. For the number of coil turns you had to find a compromise between reading and writing to enable both processes. Now they could only be optimized for the writing process, because reading takes place in the 3390 technology via a second component, a reading film. The film was very small and consisted of thin layers of different metals such as iron, chrome and cobalt. A short time later, the film was switched to nickel and iron. If the direction of magnetization reversed within the track on the disk, the electrical resistance changed abruptly, especially in the middle layer of the head above. This is known as a magnetoresistive effect, which is appropriately amplified and evaluated as a signal. This new technology ensured safe reading with the highest recording density, because even the smallest stray fields could be differentiated from information units. Simply put, two specialists were introduced, an inductive head optimized for writing and a specialist for the reading process (detailed description in the technology appendix). Both specialists together formed the read / write element, which was positioned by the access mechanism.

Size comparison of magnetic tape systems IBM 3420 versus IBM 3480

The new read / write head technology of the IBM 3390 required a new coating using what is known as thin-film technology. The surfaces of glass plates made of a specially hardened glass were coated with an alloy and then subjected to a baking process. The magneto resistive read / write technology in connection with thin film coatings on the plates was used for the first time in the 3390 series with the models 3 and 9. The aluminum plate, which had been used since the introduction of magnetic disks and which was coated with finely ground, soft magnetic material in a synthetic resin as a binder, had thus become obsolete.

Compared to the IBM 3380 series, the 3390 series delivered an improvement in access time of 28% and the data transfer rate was increased by 40% because 4.5 MB / s could be transferred instead of 3 MB / s. This enabled the channel speeds of the then new 3090 computer to be used. It was also possible to transfer from the computer to the subsystems via four simultaneous paths (four path data transfer). The 3390 disks and also the 3380 disks were connected to the new IBM 3990 control unit, which was equipped with larger disk buffer sizes (cache) from 16 to 64 MB (3990 models 3).

The era of permanently installed panels with external control units

1989: IBM 3390-1 / 2, 1991: IBM 3390-3, 1994: IBM 3390-9, rear IBM 3380-K line with 60 GB capacity, front 3390-3 line with 90 GB capacity, right IBM 3990 control units

A fully expanded 3390 Model 3 line with 1 x A38 and 2 x B3C had a total capacity of 90 GB. The large capacity model 9, which came on the market in 1994, tripled the capacity even more. However, the 9 models were used only to a limited extent because the disks were slower due to their high capacity, and were used for applications where access times and data were sufficient.

Physical specification of the 3390 disk subsystem models

	Capacity / actuator	Capacity / HDA	Number of HDAs	Capacity / unit
A14	0.946	1.89	2nd	3.78
A18	0.946	1.89	4th	7.56
B14	0.946	1.89	2nd	3.78
B18	0.946	1.89	4th	7.56
B1C	0.946	1.89	6	11.35
A24	1.89	3.78	2nd	7.56
A28	1.89	3.78	4th	15.1
B24	1.89	3.78	2nd	7.56
B28	1.89	3.78	4th	15.1
B2C	1.89	3.78	6	22.7
A34	2,838	5.67	2nd	11.34
A38	2,838	5.67	4th	22.68
B34	2,838	5.67	2nd	11.34
B38	2,838	5.67	4th	22.68
B3C	2,838	5.67	6	34.02

1991:

This year, IBM took the first step towards smaller form factors of magnetic disk drives in subsystems. The offered **IBM 9340 disk systems** and IBM 0662, for the first time on the basis of disk diameters with 5.25 or 3.5 inches, consisted of a frame in which the control unit was installed together with several disk drives. A standard 9309 frame accommodated up to eight IBM 9345 plate racks. Each drive contained two drives, each with 1 or 1.5 GB of storage capacity. They were 5¼ inch drives of the latest technology with a data rate of

4.4 MB / s, average rotation waiting time of only 5.6 ms and in proven CountKeyData (CKD) storage form. The IBM 9341 control module was available for requests up to 24 GB. It was connected to the 9221 or 9370 processor series as a 2-path subsystem without a cache.

Larger memory requirements and even higher requirements were covered with the standard frame IBM 9343 with built-in control unit with 4-way functionality and with 9345 bays. The capacity was gradually increased to 48 GB. The IBM 9341/9343 allowed the uninterrupted addition of disk trays. High performance without cache was especially achieved in database applications with high random access. Significant increases in corresponding applications were later made possible by a model-dependent 32MB to 64MB cache.

1989/1991:

Parallel to the 3390 plate development series **1989** on the tape side of the successor to the 3480 tape system, the **IBM 3490,** available and 2 years later the extended version in the form of the **IBM 3490E.** The 3490 drive technology used 18 tracks, the 3490E technology 36 tracks.

Priming process worked, the IDRC (Improved Data Recording Capability), which allowed compression factors of up to 3: 1. The new 3490 systems could be controlled with ESCON channels, as was the case with the 3390 plates. This laid the foundation stone for ESCON, which was to replace the BMPX (Block Multiplex) cabling that had been used until then.

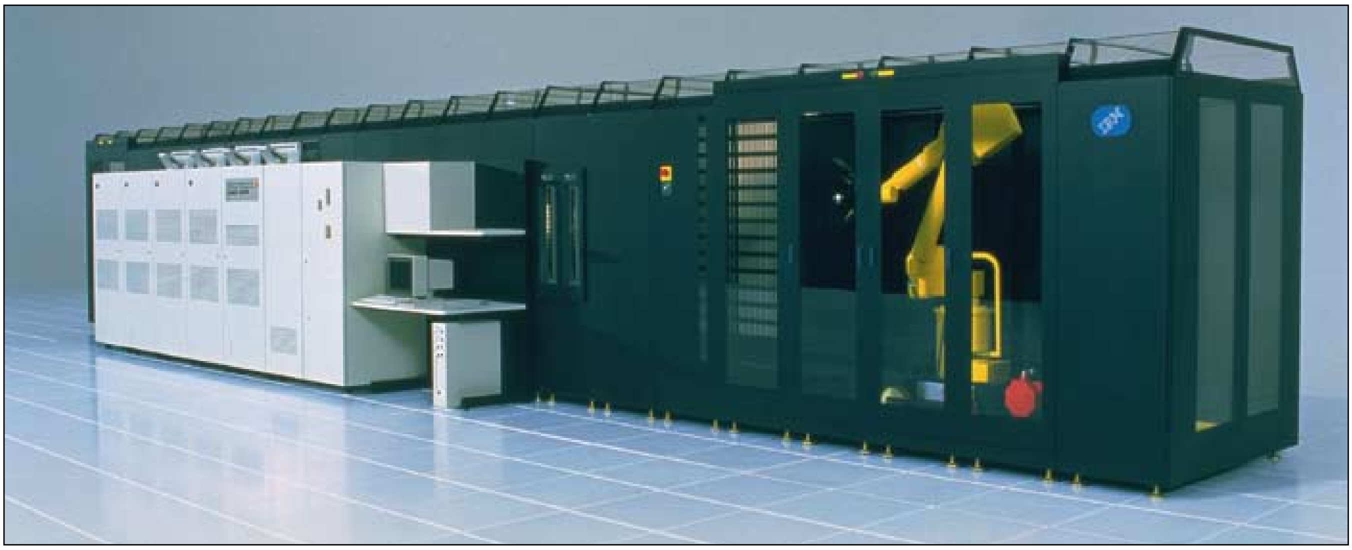
For the first time, the 3490 technology used very powerful control units with a dynamic buffer memory of 2 or 8 MB to control the tape subsystem. Likewise, when recording with a new comm

The achieved recording density was 3490 with IBM 38,000 BPI with cassette capacities of 600 and 1,200 MB, with 3490E, double the number of tracks resulted in 78,000 BPI with cassette capacities of 1,200 and 2,400 MB.

IBM 3490 magnetic tape system

1975 to 1993

The era of permanently installed panels with external control units



IBM 3495 tape archive with 5,660 to 18,920 cartridges (13.5 to 45.5 TB), height 2.38 meters, depth 2.46 meters, length depending on the model 13.4, 18.3, 23.2 or 28.0 meters, 4 to a maximum of 64 IBM 3490 drives

1992: IBM announced its own automatic after several years of cooperation in the band archive area with the companies Grau and Haushahn **Magnetic tape archive IBM 3495** for large processors that were supported under the MVS operating system. With an IBM 3490 cassette capacity from

2.4 GB, users were able to use the library capacities of Expand 13.5 TB to 45.5 TB.

For comparison: The mass storage device from 1974 allowed a maximum of about 0.24 TB.

The 3495 robot system used at that time was heavy and was not particularly accepted by the market except for a few customers. IBM was often asked why a robot weighing several tons was necessary to transport a small 3490 cassette.

1993: Then came the automatic one and a half years later **Magnetic tape archive IBM 3494** in a much more compact design and with an adequate robot system, which supported the MVS platform as well as AS / 400 and RISC6000 systems.

The 3494 library system was functionally expanded in the following years and is still an active library system that is mainly used in the mainframe environment. It was also the first IBM TapeLibrary to mix 3490 and 3490E drives and cartridges, and also to support the successor to the 3490, the IBM 3590 drive technology. With 3590 drives, which became available as successors of the 3490E in 1995, channel data rates of 9 MB / s could be achieved with the first Model B. The 3494 was also a library that could and can operate the mainframe and open systems at the same time. The E and H models of the 3590 were later supported. The library manager of the 3494 was continuously expanded developed and adapted to the VirtualTapeServerSystems VTS with their further developed functionalities that can still be operated with the library today. With the availability of the first Jaguar 3592 drives in September 2003 (see era of multi-platform systems and SAN), IBM developed an intelligent integration of these small drives into the 3494 library. In the exact size of a Magstar 3590 drive, so-called 'cradles' housing units were built, where two Jaguar drives can be accommodated in front of their part. The cradles are open to the rear. There are two power supplies. The two Jaguar drives are connected to both power supplies to ensure power supply from two independent sources. The IBM 3494 will still be operated in some data centers in 2010. In 2005, IBM published the 'SOD' (Statement of Direction) to continue to support future Jaguar drive technologies. The new high-end TS1130 drive can also be installed in the library. Marketing will be withdrawn in autumn 2009. IBM plans to maintain Library 3494 in IBM maintenance through the end of 2015.

1993 IBM 3494

IBM 3494 housing unit D22 with 'Cradles' and TS1120 (Jaguar) drives

The era of permanently installed panels with external control units

Commentary on the era of the permanently installed panels with external control units

The requirements regarding high availability became ever higher. That is why the company switched from removable plates to permanently installed plates. Fixed panels were many times more reliable. In addition to availability, the requirements for investment protection have also increased. To take this into account, IBM established the principle of changing control units and disk storage throughout the entire era. This looked like new plate units could be connected to the existing control units, then came a new, more powerful control unit on which the old plates were operated. Then new plates became available that could be connected to the existing control unit, and then a new control unit came. This change principle was so positively received by the market

On the control unit side, the IBM 3880 was the first to introduce cache and corresponding caching algorithms in order to keep pace with the performance development of the computer systems. The most important function here was the 'DASD Fast Write'. This was also maintained and expanded with the new 3990 control unit. Function after function was added. The functional expansions continued until the last model of the external control units, the 3990 Model 6, until the mid-1990s.

At the beginning of the 90s, the 'brown' plate production was changed and thin film coatings were introduced. This led to significantly higher reliability, because it was possible to work with grain structures and iron oxide particles (rust) were no longer required as data carriers. The newly developed magnetoresistive head technology (MR head technology) was used (see also technology appendix).

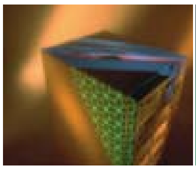
The big technological change came in 1984 in the tape sector. The roller belt was replaced by a small cassette with the 3480 belt technology. This was primarily necessary to match the belt area in terms of performance to the plates and the overall system.

At the beginning of the 1990s, the first automated tape archives came with the products IBM 3495 and IBM 3494 with robots and gripper systems that could transport the tape cartridges.

Overall, the very long-lasting era of the permanently installed panels with external control units was an era that was characterized by constant further developments with regard to data security and investment protection and ensured that the storage kept pace with the newly developed computer systems. In addition to data security, the top priority was the overall performance of the network of computers and storage systems.

1994 to 1998

The era of RAID systems



1994 to 1998

The era of RAID systems

Disk systems with RAID architectures

In 1994, IBM offered a completely new concept for external magnetic disk systems under the registered trademark, which was used again for the first time since 1956 **'RAMAC'** at. This time the letter order RAMAC stood for ' **R** AID **A** rchitecture with **M** ultimate level **A** daptive **C.** ache '.

What was the background and what did this mean specifically?

1. ***The processing speed of comparable process***
Between 1964 and 1994, sensors had grown about 400 times faster than the transfer rate of the disk access mechanisms and about 30 times faster than the transfer rate of the common channels. The fastest way was to increase the capacity per access mechanism - by a factor of around 6,000 - twice as fast as the processing speed of the processors. The demands on the availability of data in international and in-house networks increased constantly and because technical failure of components could never be completely ruled out, the demand for fault-tolerant systems grew.

2nd ***Parallel to the development of ever more powerful magnetic***
The industry had included disk storage for large and medium-sized systems - including IBM - for the fast growing market of single-user computers (PC, laptop, notebook) in the dimensions small, but at a high technical level, powerful and reliable disk drives. At that time it was foreseeable that the capacities of these drives would increase rapidly. In addition, these small disk drives could be manufactured much more cost-effectively than the large disk files, which at that time were already known as SLEDs (Single Large Expensive Disks).

3rd ***The University of Berkeley in California already described in 1987***
(the study was carried out on behalf of IBM) in a document entitled "A Study for Redundant Requirements of Cost-Effective Disk Storage" (A Case for Redundant Arrays of Inexpensive Disks), how to interconnect the disk drives in question into a single addressable disk from the operating system's point of view to use them for larger systems and achieve either higher transfer rates or higher data availability or both. In this paper, 5 RAID levels were defined, which differed in the balance between transmission performance and the degree of fault tolerance.

RAID1 describes the double storage of data on two disks with completely identical content (mirroring, mirro ring). If one drive fails, the system accesses the other.

At **RAID2** the data is copied byte by byte to several disks (multiple mirroring). An error code is saved on another disk, which can be used to reconstruct lost data.

At **RAID3** the individual bytes are also stored on several disks - albeit alternately - and on a separate disk so-called parity bits. If a disk fails, its content can be restored using the disk that has remained intact and the parity bits.

RAID4 differs from RAID3 in that the data is divided into whole blocks of several kilobytes instead of individual bytes.

At **RAID5** the system creates parity bits for blocks of several kilobytes, which are distributed across all disks so that they are always on a different disk than the data from which they were generated. The method offers a high level of security with relatively quick access, because the distribution enables parallel write updates. That is why this RAID level was most widespread.

For the sake of completeness, RAID levels added later are mentioned:

RAID6 offers additional parity bits (two parity schemes) compared to RAID5 and thus even more security (up to two disks can fail within an array), but at the expense of the performance of an array.

RAID7 also logs all write accesses.

RAID0 distributes the data in small packets over several disks and enables fast access, but due to the lack of redundancy there is no increased security. Finally there is RAID10, also known as level 1 + 0. It works like RAID0, but with two sets of disks of equal size, the second group containing the exact mirror image of the content of the first group. This solution requires many drives.

In the meantime (until today, in 2010) entirely new RAID levels have developed from these basic RAID levels, which are beyond the standard levels, are sometimes nonsensical, but in some cases also represent an optimization of the basic levels (such as RAID 1E, 1E0 , 1.5, Matrix RAID, RAID15, 51, 53, 45, RAID5E, 5EE and RAID5DG / RAID ADG).

4th ***A RAID advisory board was established in 1992 (RAID Advisory Board) from 40 user and manufacturer companies, including IBM. This made RAID a kind of industry standard in the 1990s. The meaning of the original name RAID has been slightly corrected: 'Redundant Array of Inexpensive Disks' have become 'Independent Units' (Redundant Array of Independent Disks).***

IBM has had a patent since 1978 for the restoration of data stored on a failed unit from a group of drives. Like the rest of the industry, IBM endeavored to effectively address the challenges arising from the points outlined above. The increasing changes in the market with its constantly growing number of providers, more and more established hardware and software platforms, dynamic innovations and the ever shorter product cycles made it necessary to significantly change and adapt the IBM business policy. So far IBM produced almost exclusively for its own needs for distribution through its own organization to end users and was only occasionally a subcontractor of other manufacturers (industry term for such businesses is OEM, Original Equipment Manufacturer). One example was contracts concluded in the 80s and 90s for delivering slightly modified IBM disk storage to Siemens. IBM always allowed third-party products of all kinds within their configurations and software platforms, without appearing as suppliers in the environments of other manufacturers. On the networking of data processing products and DV platforms from various manufacturers, which is widespread among DVB users - heterogeneous client / server environments

- IBM responded with offers that supported all of its own and third-party interfaces relevant to the respective products. The storage business was further promoted through the acquisition of appropriate business partners in addition to the actually blue sales. It worked with more and more business partners who successfully tried to sell IBM hardware and software products in their sales organization. This trend has significantly increased to this day.

As already described at the beginning of this chapter, IBM made **1994** as the first company on that **RAID5 based magnetic disk system** available, the former **RAMAC 1 (IBM 9391).** RAMAC 1 was announced in June 1994 and was available from September 1994. A served as the basis for a RAMAC storage unit **IBM 9392 slot (** Drawer), which was integrated into a frame with several other inserts. Each of these bays contained four 3.5-inch disk drives with 2 GB capacity each (Allicat drives), two power supplies, two cooling fans, a RISC microprocessor as controller, a battery-supported buffer memory (non-volatile) and a rechargeable battery. Each slot represented a self-contained RAID5 system. Up to 16 of these RAID5 slots could be integrated into the frame. RAMAC came in two versions. The **RAMAC array** only contained the RAID5 bays and was connected to the powerful 3990 control unit at that time. The **RAMAC subsystem** included the higher-level controller function in the same housing in which the plug-in units were integrated and was installed on the side of the frame. In the full expansion stage, both systems delivered a usable disk capacity, RAID5 secured, of 90 GB.

RAMAC 2 and **RAMAC 3** followed because of the rapid progress in the capacitive possibilities of the 3.5-inch drives at intervals of only one year. The UltrastarXP drives were used for RAMAC 2 and the Ultrastar 2XPL drives for RAMAC 3. The computational capacity doubled for each subsystem.

With 16 bays, RAMAC 2 offered a usable capacity of 180 GB and RAMAC 3 of 360 GB.

With RAMAC 3 in the version of the RAMAC array for connection to 3990 control units, a new model of the control unit was introduced at the same time, the 3990 model 6. This control unit was a double control unit that was integrated into a housing and on which two RAMAC 3 arrays are operated could. Two RAMAC 3 arrays on the 39906 thus delivered a total capacity of up to 720 GB usable as a 'double subsystem'.

The RAID5RAMAC series was very successful on the market, but could not be continued over the years because the cost factor in production could not withstand the price drop of the gigabytes in the offered subsystems. Each slot as an independent system and up to 16 slots under its own control unit, thus two cache levels: this was very complex to produce.

This was one of the main reasons why **1996** a **Cooperation with the company StorageTek** regarding the distribution and further development of the available 'Iceberg' plate system was closed and the IBM put two plate subsystems from STK on the market. The 'Ice berg' was named by IBM **IBM RVA (RAMAC Virtual Array)** marketed, the large capacity system from STK as **IBM RSA (RAMAC Scalable Array).** The RAMAC Virtual Array was announced in July 1996 with immediate availability. The more powerful RVA in the turbo version became available almost a year later, in April 1997.

The RAMAC Virtual Array **IBM RVA** IBM 9393 met the requirements of **RAID6** and offered a very smart solution for creating and growing files. Previously, users had to reserve sufficient space for growing files. This meant keeping a constant supply of empty tracks (allocated tracks) or unused capacity. If the system ran out of capacity when running a program for a growing file, the operation was aborted abnormally. The system automatically searched for free space at the RVA. Reservation (allocation) was no longer necessary and cancellations due to lack of storage space were excluded. This special algorithm was referred to as a logically structured file algorithm (Log Structured File).

left: IBM RAMAC subsystem with integrated control unit right: IBM RAMAC array for connection to 3990 control units

1996 - 1998 IBM RAMAC Virtual Array RVA, IBM RAMAC Virtual Array 1996, IBM RAMAC Virtual Array Turbo 1997, IBM RAMAC Virtual Array Turbo 2 1998, capacities from 840 GB to 1,680 GB arrays 13 + 2 + 1 RAID6 or 2 x (5+ 2 + 1), array capacity up to 420 GB, minimum configuration: 2 arrays, maximum configuration: 4 arrays

If there were any traces to be written on the disks, the control unit sought the space on any disks in the available arrays. The obsolete traces were simply released for overwriting in a space reclamation process. The whole thing was managed via pointer tables and the pointers of a current changed file were simply adapted accordingly. This new LogStructured File Architecture, as the first virtual disk architecture on the market, also allowed single files or volumes to be copied in a matter of seconds without having to reserve additional space for the copy. Only the changes in the form of new tracks or track updates again required new space. The system created corresponding copy pointers, that matched the original pointers on a single track basis. This function was called 'SnapShot'. Thanks to this principle, the RVA could be used much more in terms of disk allocation compared to conventional disk systems. With the RAMAC Virtual Array, IBM SSA disks were used in the backend and the system could be configured with 4.5GB, 9GB and later with 18GBSSA disks.

IBM 7133 disk subsystem, SSA architecture with SCSI host connections, model T40 'Deskside' Entry, model D40 'Rackmountable Disk System' disk type

	Capacities from / to
18.2 GB	72.8 - 291.2 GB
36.4 GB	145.6 - 582.4 GB
72.8 GB	291.2 GB - 1.16 TB
146.8 GB	587.2 GB - 2.34 TB

The second product from the StorageTek series, the **IBM RSA** (RAMAC scalable array) corresponded to the conventional subsystem architectures, but could already **1996** can be expanded to 1.4 TB in relatively small scalable steps. The system was, however, capacitively overhauled by the RVA in the Turbo and Turbo2 versions at relatively short notice and showed only very little market penetration.

Disk systems for open systems

While IBM mainly concentrated on the RAMAC series for magnetic disk subsystem development in the mainframe area, new architecture developments for disk subsystems in the OpenSystems area were carried out in parallel. Until the end of the 1990s, there was no disk system that could serve servers from the OpenSystems area (AIX and other UNIX derivatives and the Intel platforms) and the mainframe with its MVS operating system at the same time. On the OpenSystems side, IBM went different ways. Starting in the 1990s, a new, OpenSystem-compliant plate architecture was developed in the IBM laboratory in Hursley in England **SSA** should be used for many years. SSA stood for **Serial storage architecture,** an architecture that complies with the limitations of the SCSI disks previously connected to OpenSystems servers

systems (Small Computer Systems Interface) eliminated. At SSA, a loop technology was used for the connection of disk drives, which allowed two write and two read I / Os to be placed in the loop at the same time using newly developed adapters, and all drives in the loop were operated with equal rights. The well-known Arbi tration, as was the case with SCSI-based drive lines, was eliminated. With a SCSI drive train, one I / O always had to be dealt with before the next one could start. In addition, drives that were connected at the very end of the SCSI line were disadvantaged in terms of operation. An SCSI line had the limitation that a maximum of 15 drives could be connected. The new SSA architecture overcame these limitations **1992**

was first used in subsystems.

The best known system was the IBM **7133 Serial Storage Architecture Disk Subsystem,** that could be connected to RISC systems / 6000, power servers and power stations and was always equipped with the latest 3½ inch drives. The IBM 7133 subsystem was announced in July 1995 with immediate availability. It was one of the most successful disk systems in the AIX environment and was sold directly via the AIX sales channel. During the life cycle of the 7133, over 42,000 panel systems were put into production worldwide, a unique sales success.

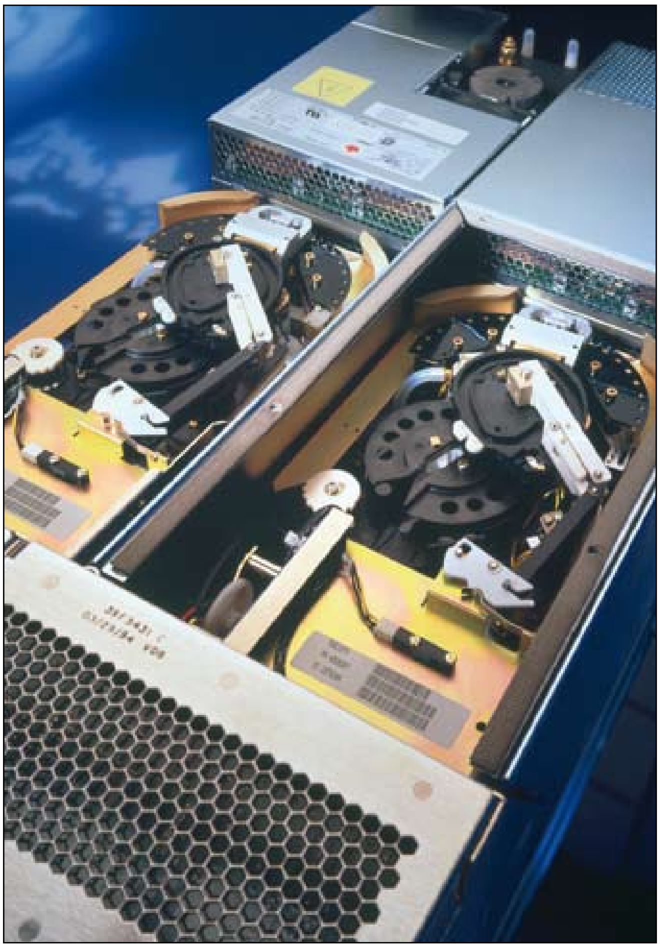
The last systems in SSA architecture (IBM 7133 models T40 and D40) remained until **2003** in active IBM sales and at that time could connect to existing FibreChannel infrastructures (SANs, Storage Area Networks) via gateways.

1999 IBM announced its own self-developed disk subsystem under the name for the mainframe **IBM ESS (Enterprise Storage Server)** , but more under the development name **'Shark'** became known and ultimately replaced the RVA and RSA in IBM sales. Sto rageTek then again sold the RVA itself under the name STK Shared Virtual Array. The ESS was not only a disk system for the mainframe, but could also be connected to all common OpenSystems platforms. This created the prerequisite to operate both the mainframe and the Open SystemsServer with a disk subsystem, as had been possible on the tape archive side for years. Over the following years, the ESS partially replaced the IBM 7133 Serial Disk System in the OpenSystems area. At the same time, FibreChannel-based plate systems for the Open Systems platforms were introduced.

Belt systems

In addition to disk subsystem development in the 1990s, IBM introduced one **new tape drive technology** as a replacement for the 3490 technology previously used

IBM 3590, the one under the name **Magstar** got known. The first drive of the Magstar 3590 was the Model B, which recorded a native capacity of 10 GB on a ½ inch cartridge. The 3590B drive was announced in April 1995 and was available in September of the same year. This new recording was the first



A look inside the IBM 3590 Magstar drives

Magnetoresistive read / write heads with the 'Track Following Servo' principle were used in tape technology, where the reading elements constantly control how well the bits were written by the write head. The cassettes were filled with 128 tracks and the data rate of the Model B of the 3590 was a fascinating 9 MB per second (native).

The drives were available as a standalone version, i.e. permanently installed in a housing with appropriate autoloaders, but they were also used in the current IBM 3494 tape library. Fully expanded, the IBM 3494 with 3590B drives had a capacity of 180 TB (native). The drives themselves had two SCSI interfaces and could either be connected directly to OpenSystems servers or via a control unit to ESCON-based mainframe servers.

The Magstar development series then became **1999** with the new **E-model** expanded, which had twice the number of tracks with 256 tracks and a capacity of 20 GB on the Kas

IBM 3590 Magstar Tape Subsystem 1995: IBM 3590-B with 9 MB / S data rate 1999: IBM 3590-E with 14 MB / S data rate 2002: IBM 3590-H with 14 MB / S data rate Cassette capacities (native): Model

	normal cassette	long cassette
3590-B	*10 GB*	*20 GB*
3590-E	*20 GB*	*40 GB*
3590-H	*30 GB*	*60 GB*

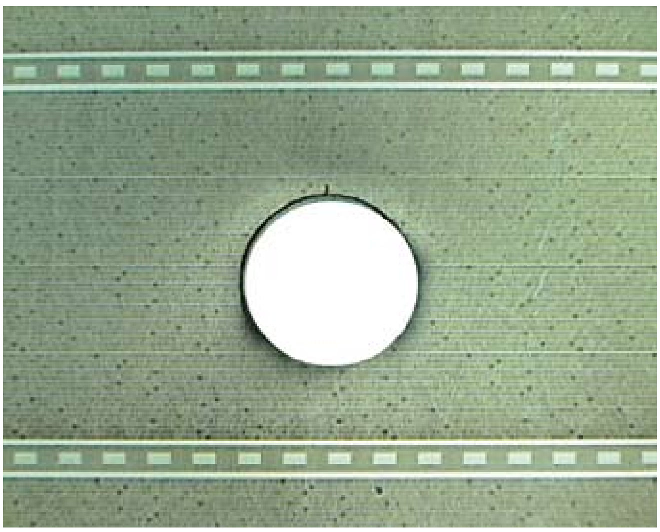
set saved (native without compression). The data rate was increased from 9 MB / s to 14 MB / s. At the same time, IBM launched a new cartridge, the **3590 'Extended Length Cartridge',** which with the double tape length (described with 3590E) also represented twice the capacity, namely 40 GB. This increased the 3494 library capacities to 374 TB using the normal cartridge and up to 748 TB using the long cartridge. The Magstar development series was available from May 1999.

The Model E of the Magstar series had a powerful controller, which in addition to recording data also blocks **Parity blocks** recorded in order to enable a much higher error recovery. (When 16 tracks were recorded at the same time, after 7 data blocks in the eighth and after another 7 data blocks in the

16. Track recorded a parity block and 'striped' over the logical track set of eight tracks - RAID5 on tape.)

A test with the Magstar E model, which was carried out in June 1999 and in July 2000 in the IBM laboratory in Tucson, Arizona, shows what Magstar can do in terms of recovery. All key consultants such as Gartner Group and IDC were invited to the experiment. With the E model, a cassette with 256 tracks was fully written and a clean Read Verify was obtained. Then the cartridge was unloaded from the drive. A 1.25 mm hole was then punched into the center of the tape. This hole reflected approximately 100,000 bits, which were then missing. The cassette with the hole was reloaded into the drive and the read operation was carried out successfully without errors. All missing data bits could be 'recovered' by the ECC (Error Correction Code) via the bit block distribution and the parity bit blocks.

The drive had two SCSI interfaces. From May 2000, the Magstar drives were also available with two FibreChannel interfaces. This enabled the drives to be connected directly to a FibreChannel-based SAN network. The corresponding controllers came onto the market with a delay, which also enabled the drives to be connected to FICON-based mainframe computers.



Hole punch test with Magstar 3590 tapes in 1999 and 2000 in the IBM laboratory in Tucson

The era of RAID systems

The last model of the 3590 Magstar series came in 2002 with the 3590 Model H, which recorded 384 tracks on the same cassettes and thus increased the cassette capacities by 50%. Model H was announced on June 25, 2002 and was already available on June 28, 2002. The data rate of the 3590H remained at the level of the predecessor 3590E, at 14 MB per second. The 3494 library cassette capacities now reached 1 petabyte native capacity in full configuration (without compression).

In September **1996** was created by IBM **new cassette tape system** was introduced, which achieved extremely fast access times in addition to high streaming speed. The
IBM Magstar MP 3570 offered a data transfer rate of 15 MB per second and access times of just a few seconds. This was made possible with a special cassette design, where the basic position of the tape in a two-spool design was in the center position. For example, when loading the tape, it could be spooled both to the left and to the right in order to quickly access a file. This new concept, however, only enabled small cassette capacities of 5 MB and later 10 MB, which was not widely accepted by the market as was originally thought. Therefore, the life cycle of the Magstar MP was short and limited to about 3 years.

With the availability of the first Magstar models **3590-B 1995**
there was one in the mainframe environment **new problem,**
because the now high cassette capacities of 10 GB native and the high data rates of 9 MB / s native were only used to a small extent. A worldwide study showed that the degree of utilization in the mainframe environment was around 200300 MB per cartridge and the data rates were on average 23 MB / s and for many backup jobs even less than 1 MB / s. Around

IBM Magstar MP 3570 tape subsystem

to fix this problem, IBM started **Tape virtualization solutions** for the mainframe area in order to make maximum use of the physical resource behind it, both in terms of the capacity of the cartridges and the data rate of drives. The emerged from an initial software-based volume stacking solution

3494 Virtual Tape Server VTS, the one in June **1997** with the first **Model B16** came onto the market. The tape volumes that had to be backed up were buffered on a disk buffer and indexed. As soon as so many volumes were written to the disk buffer that they could fill a 3590 cartridge, a copy of the entire collection of disks was written to the fast drives (premigration) and the cartridges were filled to the maximum (stacked volume principle). The original remained in the plate buffer. This has maximized the utilization of cartridge capacities and drive speeds. Recalls of volumes that were still in the disk buffer were carried out at disk speed, and many jobs could be started in parallel during backup, since many virtual drives were available. IBM implemented a so-called 'outboard' solution, which was completely transparent to the host and the MVS operating system and showed the same transparency with regard to the tape management system and the catalog entry. 3490E drives with cartridge capacities of 400 or 800 MB were emulated and the host did not register that he had no physical drives in the access.

The first virtual tape server **IBM 3494 B16** was then integrated directly into the 3494 tape archive as a housing unit. The B16 housing housed the RS / 6000 server, a RAID 5 secured disk cache with SSA disks, which could be expanded up to 72 GB, as well as the ESCON connections and 400 cassette slots required for the host. Later, one worked with external units, as the further development of this concept shows. In today's world, virtual tape solutions have become indispensable because it is the only way to make maximum use of tape and tape library resources in this way.

The principle of integrating the VTS (Virtual Tape Server) directly into the 3494 tape archive turned out to be extremely unfavorable. In addition, the throughput of the B16 at that time was very small and only reached 68 MB / s because it was operated without compression. That is why IBM terminated a year later, in June **1998,** with an availability in August 1998, a new one **external model B18** which could be operated as a separate unit next to the 3494 library. This new unit offered compression, which was carried out directly in the host adapters, then based on ESCON. This had the advantage that all the data in the disk buffer was already compressed and the utilization rate in the disk buffer was increased by an average of a factor of 3. The use of SSA drives with 9 and 18 GB capacity increased the size of the disk buffer to 288 GB and 576 GB. Later, with 36GBSSA files, disk buffers of up to 5.2 TB (usable) became possible. As a result, the dwell time of the volumes in the disk buffer was increased many times and recalls were carried out directly from the disk buffer as virtual mounts at disk speed. The B18 units at that time, depending on the expansion stage, was still available for the old model B18 in a limited form.

These new extensions, as **'Advanced Function'** designated, allowed over a **APM (Advanced Policy Management)** new functionalities at the IBM VTS. They are still in active use today and form the **Interlocking with SMS (System Managed Storage)** as an integral part of the z / OS operating system.

The functionality **'Logical Volume Pool Assignment'** offers the possibility to manage logical volumes individually per number range in different physical cassette pools. This made the IBM VTS multi-client capable. With **'Selective Dual Copy'** two copies of a logical volume can be created in two different physical cartridge pools. The function **'Peer to Peer Copy Control'** addresses a PtPVTS complex and allows you to control which logical volumes are mirrored in a PtP complex in Immediate Copy Mode and which logical volumes are mirrored in Deferred Copy Mode.

Many other companies jumped on this business very quickly and a whole series of band virtualization solutions, HW-based, but also pure SW stacking solutions became available on the market. However, many disappeared from the market very quickly.

The model B18 was functionally expanded in the following years. Import / export options for logical volumes were possible. With the introduction of follow-up models in August

2001, of the models **B10 and B20,** In addition to the ESCON connections, FICON connections were also available and a wide range of new functions were used

IBM Virtual Tape Server VTS, 1997: Model B16, 1998: Model B18, 2001: Model B10 and B20

With **'Tape Volume Cache Management' (** this function was previously available, but host-controlled), logical volumes can be assigned to a cache management preference group. There are two preference groups available. The preference group controls how the logical volumes in the disk buffer are handled.

But the most important introduction in August 2000 was that

Principle of mirroring entire VTS systems, which is still referred to today as Peer to Peer VTS and quickly established itself in the market, especially among banks and insurance customers. A PeertoPeerVTS complex offers this **synchrone mirror possibility** and asynchronous mirroring. Another feature that **2004** was introduced (**Selective Peer to Peer Copy),** Individual volumes or entire storage groups can also be managed both mirrored and non-mirrored in a VTSPeerto Peer mirror complex.

The VTS story, which began in 1997, was a unique success story for IBM, especially in the mainframe environment, and IBM decided to continue this success story with new systems and new functionalities over the following years.

1994 to 1998

The era of RAID systems

Linear tape open LTO Ultrium

In the **first half of the 90s** had one in the OpenSystems environment, especially in the area of small and medium-sized businesses and in smaller companies **Tape technology** that dominated this market: **DLT (** Digital Linear Tape), a cheap alternative to high-end ½ inch technology that was affordable for small businesses and had a market penetration of 90% by mid-1995. DLT was the manufacturer and supplier of Quantum. The smaller part of the market went to HelicalScan based technologies and the QIC (Quarter Inch Cartridge) and MiniQIC technologies.

With the **Merger of the companies IBM, HP and Seagate** in a new tape development committee, the so-called **LTO (Linear Tape Open)** Panel, an initiative was launched to counter the dominant and dominant Quantum DLT with an alternative. The LTO Development Board started its work in 1996.

The specifications of LTO were from **1997 to 1999** developed to as **first tape technology standard** to come onto the market and stop the then rampant growth of different technologies and media. LTO is an official ISO standard, both in terms of technology and medium, which allows LTO media to be processed on any LTO drive, regardless of which manufacturer, regardless of which manufacturer. Today, over 30 companies have acquired LTO manufacturing licenses. In addition, there are the OEM agreements that the development companies IBM, HP and Seagate had entered into. LTO therefore had a great opportunity to prevail over DLT in the market in the long term.

If you look at the market shares today, 2010, this calculation worked. Today DLT has almost disappeared from the market.

As the LTO Group **1997** When the standards for standardization were first developed, she first assessed all existing tape technologies for their plus and minus points in order to incorporate the good in LTO and omit the negative. No more wild growth, compatibility was required! Whether Sony AIT or 4 mm, 8 mm helical scan technology from Exabyte, Quantums DLT to Super DLT, Philipps NCTP, Quarter Inch Cartridge or Mini Quarter Inch Cartridge: all technologies that were not compatible and therefore had difficulties, continue to advance to assert itself permanently on the market!

Due to the high reliability, the recording technology was chosen for the IBM Magstar linear serpentine recording, which describes the tape at a data rate of 15 MB / s with 384 tracks (LTO1) and 35 MB / s with 512 tracks (LTO2). Eight tracks were always written at the same time. At these track densities, the cartridges reached native capacities of 100 GB (LTO1) and 200 GB (LTO2). With this high wheel density, it was necessary to ensure that the read / write elements were guided as precisely as possible.

For this purpose, a new, **time-based tracking system** used that worked with five prescribed servo tapes, each with eight servo areas. The AIT (Advanced Intelligent Tape) technology from Sony adopted the memory chip integrated in the cassette, but with a larger storage capacity (4 KB).

LTO1 drive with LTO1 Ultrium cartridge

Cartridge loading on the IBM LTO1 drive

IBM autoloader and library offering with LTO drives, late 2000, left IBM 3584 Library, bottom right IBM 3583 Library, above it the Autolaoder IBM 3581 and the LTO stand alone drive IBM 3580

There is recorded what z. B. at Magstar, the tape is stored in the VolumeControlRegion region: Manufacturers specified ReadOnlyInformation, the table of contents of the tape and information on how many read / write errors occur on the tape in order to move the cartridge to 'Read Only' status for a certain threshold value for security reasons put. There was also space for user-specific information. In addition to being saved on the chip, for security reasons, all information was also stored in the tape's VolumeVolumeControl region.

IBM developed to match the LTO drive development **a new library concept,** that works with a gripper system (until today) that could optimally handle the features of the standardized LTO cassettes. The robot system with this special gripper was **Late 2000** with the **new IBM 3584 Library** realized. The IBM 3584, which at that time could only be operated on Open Systems, was the perfect counterpart to the IBM 3494 Library, but with much more modern options and a robot speed that is still unmatched today.

With the introduction of LTO tape technology on the market, IBM changed its previously pursued concept of producing everything itself. It was impossible to survive as a single manufacturer in the already competitive Open Systems market, which represented all types of solutions and built them in terms of production technology. That is why IBM signed an OEM agreement with ADIC back in 1999 that looked like ADIC received LTOL drives from IBM and in return IBM the small libraries from ADIC, which were then marketed under the IBM logo with IBM LTOL drives.

IBM 3583 Ultrium Scalable Tape Library with the door open

So ADIC provided the 'midrange libraries' at that time, the **IBM 3583** and later the **IBM 3582 (** with ADIC Skalar 100 and Skalar 24). The only difference to the ADIC products was that IBM integrated its own FibreChannel-based control unit in both libraries, in order to give the libraries the possibility to create logical partitions as sub-library raries and to directly install FibreChannelLTO drives. ADIC used SCSI drives, which could then be connected to SANs via a gateway. FC drives were only used and installed at ADIC in early 2005.

The current autoloaders, midrange libraries and the special features of the 3584 tape archive are described in more detail below.

IBM 3582 Display - mini library with up to 23 LTO cartridges

Commentary on the era of RAID systems

Compared to the previous era, where the same control units and panels remained up-to-date for years, the era of RAID systems was very hectic with very fast development cycles. RAMAC 1, 2 and 3 came in succession at intervals of only one year.

In this era there was enormous price pressure, which was caused by new competitors and players in the record environment.

RAMAC 3 was the last Rolls Royce that IBM built, then the costs ran away and the products were far too expensive compared to the competition. This led to the merger of StorageTek and IBM and the IBM distribution of the RAMAC virtual array (formerly STK Iceberg). Suddenly, IBM sales had to sell a product they had always distanced themselves from. A massive rethink was announced in storage sales. Log Structured File, the 'Garbage Collection - which can't be anything', which was previously negatively commented by IBM, had to be displayed positively. The fraternization with STK in the plate environment had to be digested first.

In the OpenSystems area, the successful introduction of the SSA architecture and the successful sale of the IBM 7133 disk system compensated for the slump in the mainframe environment.

The years of development in the tape environment were just as hectic. In addition to Magstar 3590, Magstar MP and the development of the Virtual Tape Server (VTS), the LTO (Linear Tape Open) tape technology was developed and specified together with HewlettPackard and Seagate. In autumn 1999 the standard of LTO1 technology was ready. A prestigious race between IBM, HP and Seagate then began. Everyone wanted to be the first to launch LTO1 drives.

Seagate was the first company to announce its LTO1 drives in the spring of 2000. The Seagate drives became the last available. IBM announced in September 2000 with immediate availability. HP followed a month later with the announcement and availability.

The epoch was marked by sales struggles by many competitors in the record and tape environment and was therefore so hectic because of enormous price and prestige struggles.

1999 to 2005

The era of multiplatform systems and FibreChannel SAN and NAS

The era of multiplatform systems and FibreChannel SAN and NAS

1999 - 2006: IBM SAN, switches and directors from Brocade, McData (CNT) and Cisco

SAN, NAS, iSCSI

Up until 1999, the era of RAID systems was designed in such a way that disk systems were available for the mainframe area, while the OpenSystems area used its own, then directly connected disk systems. Only tape libraries were able to support both worlds. Data growth during this period was dramatic, often over 60% a year for data center operations. The IP networks and the associated infrastructure, at that time Ethernet on a 10 Mbit basis, hardly allowed large amounts of data to be moved back and forth over IP infrastructures. The call for

separate storage network infrastructures got louder and louder to counter the flood of data. In 1998 Brocade launched the first 8-port FibreChannelSwitches with 1GbitPorts. A **new infrastructure, the SAN (Storage Area Network)** based on FibreChannel, should experience a real heyday in the coming years.

The beginnings of **first SAN infrastructures** However, took place 2 years earlier. You came already **1996**

in **High performance data centers** of the universities and were made available by the Ancor company.

1998 the first 8PortSwitch was used by Brocade in the commercial data center environment at IT Austria in Vienna and in the following years many companies started to implement SAN infrastructures in the most varied of company areas. Not only in the Open Systems area, fiber optics experienced a new heyday, the mainframe environments also changed, beginning in 2001, from the previous ESCON environment (Enterprise Storage CONnection) with the previously used ESCON directors to fiber optic networks, in order to bundle combined ESCON packages via the fiber optic transfer. The whole thing is now called FICON (Fiber CON nection). As a result, the high transmission rates of the single-mode glass fiber with corresponding multiplexing methods could also be used in the mainframe area.

Already in the year **2003** were directors on the market that could map both FibreChannel and FICONPorts. Unfortunately, until now no FCPort can be used as a FICONPort or vice versa and it will certainly take some time until SAN-based infrastructures from both the OpenSystems environment and the mainframes can be used simultaneously. The Fiber Channel SANs evolved rapidly. **2002** the first came **Switches with 2 Gbit ports** on the market. in the

Year 2006 all manufacturers were in the process of converting from 2Gbit to 4Gbit technology, and by the end of 2006 the changeover had come to the point that 'end to end' SAN implementations on a 4Gbit basis could be implemented. Both the host adapters of all server manufacturers and the host ports of all storage providers were converted to 4Gbit technology. Today, in 2008, this game is only repeated on an 8Gbit technology basis. The first 8Gbit-capable directors have been on the market since the beginning of 2008.

In the early days of this era **SAN implementations are very expensive.** Not every company, especially in the middle class sector, could afford SANs financially to take the new bandwidth possibility into account. Added to this was the high price policy of the glass fiber manufacturers, which made the monomode fiber required for multiplexing very expensive over the early years. The ServerAdapter manufacturers joined this high price policy. As a result, the developers of multiplexing processes on fiber optics were no longer able to identify any major sales potential and therefore turned to new multiplexing processes based on Ethernet. So it happens that IP infrastructures based on Ethernet today offer much larger transmission bandwidths (2Gbit and 10GbitEthernet) compared to FibreChannel networks.

The high price policy also led to the arrival of a new technology, NAS (Networking Attached Storage), in addition to the SAN. NAS technology in the form of file server devices with built-in disks was used, as were gateways, which in principle were nothing more than a file server, but had the option of using disk capacities within a SAN for file serving purposes.

Today SANs are affordable even for very small businesses. The drop in prices from 2003 to 2006 is in a range of almost 96%. This drop in prices was largely achieved by the new bandwidth options of the IP networks based on Ethernet and certainly, especially today (10Gbit Ethernet), questions the need for SANs when it comes to bottlenecks in the data transmission option. Despite this, mainly due to the management options, it can be assumed that SAN networks from 4 Gbit to 8 Gbit (or 12 Gbit) will be further developed, which remain compatible with the existing networks (autosensing).

In addition to the possibilities and advantages of SANs, IBM also offered solutions and products in the sector from 2000 to 2004 **Networking Attached Storage (NAS)** at. NASSystems are preconfigured file servers. To date, they consist of one or more internal servers with pre-configured disk capacity and are connected to the LAN via Ethernet, where they make their disk storage available as a file server or as an HTTPServer. NAS servers are specially developed for file sharing. The advantage of NAS compared to classic file servers was the low installation and maintenance effort.

The IBM NAS models 200 and 300 were solutions especially for customers who wanted to consolidate storage in a Windows environment. All IBM NAS appliances were delivered with a preconfigured Microsoft Windows operating system. This Windows was specially adapted for file serving. Routines that were unnecessary have been removed from the operating system to increase performance and reliability. As a result, the IBM NAS models offered very good performance in a Windows environment (CIFS). The IBM NAS models also performed well in a mixed Windows and UNIX (NFS) environment.

Such file servers were not intended for classic block I / O applications such as databases, because these applications directly access a formatted disk space without an overlying file system. Solutions such as Direct Attached Storage (DAS), Storage Area Networks (SAN) and iSCSI products were more suitable for this.

The era of multiplatform systems and FibreChannel SAN and NAS

IBM NAS Gateway 500, available February 2004, pSeries-based, up to 224 TB capacity in the SAN

In order to use the storage resources in the SAN for file serving, IBM, at that time the only vendor at the time, offered the option of using existing IP network structures **NAS gateway** connect to the SAN. The NAS gateway was a dedicated file server. This was connected to the SAN via Fiber Channel and was able to B. an Enterprise Storage Server ESS in the SAN optimally for file serving. By using a SAN storage pool via gateways, file servers could be consolidated into a SAN. Using the storage in the SAN also enabled server-independent scaling in terms of capacity. The **NAS Gateway 300G**

was later replaced in 2004 by the much more powerful Gateway 500.

iSCSI is a joint development of the companies **IBM and Cisco.** The SCSI protocol is transmitted via TCP / IP. iSCSI thus follows a similar approach to SAN, with the difference that a TCP / IP connection replaces the SCSI cable for iSCSI, and at that time was a cheaper alternative to SANs in areas with low or medium performance requirements. The advantage was that existing ones IP networks could be used directly and there was no need to set up a separate fiber optic network in the form of a SAN. The implementation of iSCSI solutions compared to SANs was much easier and did not require the high IT skill that was required when setting up a Storage Area Network (SAN). The IBM models **IP Storage 200i (Models 200 and 225)**

In contrast to the NAS and NAS gateways used Linux as the operating system (kernel) and offered a capacity of 108 GB to 1.74 TB.

However, the iSCSI solutions did not really prevail on the market, especially since a short time later the enormous drop in prices for single-mode fiber began, which made the implementation of SANs affordable even for small and medium-sized businesses.

Plate systems

In July 1999, the first IBM multi-platform disk system, the Enterprise Storage Server ESS, was announced. Under its development name **'Shark'** however, the system was much more widespread than under the term ESS. The type designation was 2105 and the **1999** announced **Models** were those **E10 and E20** with a capacity of 420 GB to 11.2 TB. Disk drives of 9 GB, 18 GB and 36 GB were used as SSA disks, which were connected to the computer via four so-called device adapter pairs using SSALoop technology. In the beginning, the plate types could not be mixed, but later it was possible to install different plates in the arrays. Cache sizes of up to 6 GB were configurable and a stream of independent write memory (NVS non-volatile storage) of 384 MB was available. At the beginning, the arrays could only be operated under RAID5. The 3380 and 3390 track formats were available for the mainframe; corresponding LUNs (Logical Unit Number) were emulated on Open Systems (UNIX, Windows NT and AS / 400).

From the predecessor of the ESS, the Versatile Storage Server, a short run of a few months, the **IBM Data Path Optimizer** adopted for AIX and Windows NT as an integral part of the machine's multi-path capability.

Functionally, the machine was still weak on the chest at the beginning and there were only a few copy services available as functions, which did not exactly simplify the market launch. The mainframe functions came later **'Concurrent Copy'** and **XRC,** 'eXtended Remote Copy' as an asynchronous mirror option for entire systems for S / 390 servers (with consistent data on the secondary side) and again the function is delayed **PPRC 'Peer to Peer Remote Copy'** as synchronous mirroring for both the mainframe and the OpenSystems area. **FlashCopy (instant or point-in-time copy)** was implemented on the machine so that copies could be made in seconds with which one could work immediately.

Almost 2 years after the market launch, the machine then offered a much greater variety of functions in the mainframe environment compared to the competitor systems on the market. The function **PAV (Parallel Access Volume)** allows multiple I / O operations on the same logical volume in the mainframe environment. **Multiple allegiance** allows parallel write and read access to the same logical volume from different S / 390 servers. **I / O priority queuing** uses the prioritization information of the OS / 390 workload manager in Goal Mode to optimally control the sequence of the individual I / Os and to optimally handle workloads with different importance according to their prioritization.

The system came with new management tools, the so-called **'StorWatch'** Family, to the market that allowed the machines to be administered via the web. The machine was standard with the **ESS Specialist** delivered, which was mainly necessary for the configuration and for later modifications. As a separate, additional commercially available product, the **ESS Expert** provided that provided detailed information about used capacities, performance, response times, cache hit ratios and more, in order to be able to perform an optimal tuning on the machine for the different workloads.

Enterprise Storage Server ESS (Shark), 1999: models E10 / E20 to 11 TB, 2001: models F10 / F20 to 35 TB, 2002: models 800 and 800 Turbo II to 56 TB, April 2004: entry model 750 to 5 TB (Gross capacities)

In May **2001** came the **Successor models F10 and F20,** which for the first time offered native FICON connections to work with FibreChannel in the mainframe environment. The previously used ESCON and FC adapters had also been improved and offered 100 MB / s throughput per port in FICON and Fiber Channel Speed. They provided distances of up to 10 km (native) and up to 100 km via the corresponding SANFabric components. The new models could be equipped with up to 16 FICON connections and an additional cache option of 24 GB, i.e. a maximum of 32 GB. This improved the performance of the previous models by approx. 40%.

In terms of functionality, the new models also made the PPRC and FlashCopy functions available to the AS / 400 and iSeries system platforms. in the **March 2002** was also the **FICON support** available for the Transaction Processing Facility (TPF) application used in airline booking systems.

The change from ESCON to FICON infrastructures brought enormous infrastructure improvements and infrastructure savings because four ESCON channels with 18 MB / s could usually be replaced by a FICON channel with 100 MB / s. However, the penetration continues to this day.

In August 2002 came the last series of the **ESS with the 800 and 800 Turbo II models** to the market, which is then replaced by a **Entry level 750** in the **Spring 2004** was added. The new models were more powerful and used two 2, 4 or 6WaySMP computers with up to 64 GB cache.

This enabled approx. 42% more transactions with cache standard workloads. The new models could be mixed drives with 18 GB, 36 GB, 73 GB and 146 GB and offered a gross capacity of 582 GB to 56 TB. The drives with 18 GB, 36 GB and 73 GB were available with rotation speeds of either

10,000 revolutions / minute and 15,000 revolutions / minute, while the large-capacity 146GB drive was only available at 10,000 revolutions / minute.

Optionally, the arrays could be configured with RAID0, 1, 5 and 10 and all disks in an array were controlled with a striping process (all drives work simultaneously if a track is written or read in small pieces in parallel on all drives of the array), to ensure optimal read and write performance. SSA160 adapters were used for the array connection, which achieved a total throughput of 320 MB / s with eight simultaneous I / Os in a double loop. The overall throughput (internal bandwidth) of the TurboII model was included

3.2 GB / s. The maximum cache size was up to 64 GB.

The RAID management was completely relocated to the adapter cards and with sequential write workloads the machine was able to switch from RAID5 to RAID3 mode, in which up to 64 simultaneous physical I / Os were possible on the disks.

Right from the start, all machines were designed for 2 Gbit host connection technology, i.e. 2 Gbit FICON and 2 Gbit Fiber Channel. In addition, the machine could still be configured with ESCON and UltraSCSI adapters.

The **Models 800 and 800 Turbo II** were the basis for a complete **new remote copy procedure** to develop, which is still unique in its performance to this day and is used to a significant extent in the successor products of the DS6000 and DS8000. This new RemoteCopy procedure was made available to the new ESS models in April 2004 and, through a bidirectional implementation via a Fiber Channel connection, enabled a synchronous mirror load (PPRC) of up to 16,000 I / Os per second (ESS), which was later found in the new DS8000 up to **23,000 I / Os** per second - and all this with a single FibreChannelLink. With this new implementation extremely short response times can be achieved, even with synchronous mirroring over long distances. With the 800 models of the ESS, response times of 2 ms were achieved with synchronous mirroring of over 75 km.

The management options for the new models have also been expanded. For the first time, open APIs were made available as Industry Standard Interface (SNIASMIS). The machines could be automated with scripting tools via the CLI (Command Line Interface). A web-based graphical user interface (GUI) was also available for easy administration.

RS / 6000 based POWER5 processors were used for all ESS models. The EModels used an H50 as a 4WaySMP version with a clock rate of 340 MHz. The FM models used an H70 as 4WaySMP at 540 MHz. The last models in the 800 series used an H80 calculator. The entry-level model 750 used an H80 as 2x2Way with 600 MHz, the models 800 an H80 as 2x4Way with 600 MHz and the most powerful ESS 800 Turbo 2 an H80 as 2x6Way with 750 MHz. PCI buses were used for the internal transmission in the machines.

In January 2006, IBM withdrew all ESS models from sales effective April 2006.

Parallel to the ESS, IBM gave the go-ahead in 1999 to work on server-based storage architectures that laid the basis for today's first server-based storage architecture and are now available with the product of the DS8000. This new architecture development took place parallel to the further development of the ESS and was pursued with the greatest secrecy.

The successful, SSA-based IBM 7133 disk system, which was mainly used in RS / 6000 and pSeries environments, also had the opportunity in 2002 to be used in Fiber Channel SANs. With the adapter 7140, the disk system could be connected to SANs and provided a new function via this adapter, 'InstantCopy'. In addition to a RAID1 mirror, another copy could be created, which could then be used for asynchronous backup or test purposes. Since the SSA technology used here was also used in the Enterprise Storage Server ESS (Shark), the combination of 7133 and the adapter 7140 was also referred to as the 'Hammershark'. The proximity to the ESS (Shark) was also underlined by an upgrade option, in which 7133 capacity could be used to expand a Shark.

After the market launch of the ESS as a multi-platform enterprise disk system, IBM quickly realized that even entry-level configurations of a Shark were too expensive for medium-sized businesses. Therefore, was already in

Year 2000 with the company **Compaq** a **OEM sales contract** signed, which allowed IBM to sell the current Compaq plate system with FibreChannel connections. The MSS Modular Storage Server product supported heterogeneous server platforms in the WindowsNT and UNIX environment and was connected to FC servers via hubs and / or switches.

However, the time of the MSS Modular Storage Server was very short and only a few systems were installed. The alliance with Compaq ended with the takeover of Compaq by Hewlett Packard HP.

IBM then resorted to one that had existed since the early 1990s **OEM alliance** with the **LSI company** back, which allowed IBM to sell plate products from LSI under the IBM logo, especially in the Intel environment (NetFinity) via the IBM xSeries channel. The LSI products IBM 7135110 RAIDiant Array were launched in 1993, 1995 the IBM 7135210 RAIDiant Array, 1998 the IBM 3526 FC RAID Controller, 2000 the IBM 3552 FAStT500 and the IBM 3542 FAStT200 and in 2001 the IBM 1742 FAStT700.

in the **Year 2001** This alliance between IBM and LSI was expanded to include all storage sales, an alliance that was to become very successful and is highly topical up to the present day. The first product that IBM sold through this alliance was an entry plate product, the **FAStT200,** at that time also known as the IBM 3542, which with up to two FibreChan nelPorts with the basic unit (10 built-in hard disks) and two expansion units EXP500 (additional 15 disks per EXP500) had a maximum capacity of up to

2.1 TB offered. The system was an inexpensive, permanent disk solution for decentralized departments and work groups with up to four servers. As an entry system, the system already offered a whole range of security options at the time, such as dual active RAID controllers, RAID 0, 1, 10, 3 and 5, redundant power supply units and battery as well as the possibility of exchanging all components as 'hot swaps'. In the system, disks with 18 GB, 36 GB and 73 GB could be used in mixed operation.

The abbreviation **'FAStT'** stands for **'Fiber Array Storage Technology'.** The original name was supposed to be FAST - for Fiber Array Storage Server - but was already taken when the LSI products were announced and could not be used. That is why FAStT was chosen with the intention that the small 't' should not be pronounced. Despite everything, the pronunciation as 'fast_T' prevailed across the board.

	FAStT900	FAStT700	FAStT600 Turbo	FAStT600	FAStT200	FAStT200
Host interface	2 Gbps FC	2 Gbps FC	2 Gbps FC	1 Gbps FC	2 Gbps FC	2 Gbps FC
SAN attachments (max)	4 FC-SW	4 FC-SW	4 FC-SW	4 FC-SW	4 FC-SW	4 FC-SW
Direct attachments (max)	8 FC-AL	8 FC-AL	4 FC-AL	4 FC-AL	2 FC-AL	4 FC-AL
Redundant drive channels	Four 2 Gb FC	Four 2 Gb FC	Four 2 Gb FC	Four 2 Gb FC	Four 1 Gb FC	Four 2 Gb FC
Drive types supported	FC, SATA	FC	FC, SATA	FC, SATA	FC	FC, SATA
Max drives	224	224	112	56	66	56
Max physical capacity with FC	32 TB	32 TB	16.4 TB	8.2 TB	9.6 TB	-
Max physical capacity with SATA	56 TB	-	28 TB	28 TB	-	14 TB
XOR technology	ASIC	ASIC	Integrated	Integrated	Integrated	Integrated
Subsystem cache	2 GB	2 GB	2 GB	512 MB	256 MB	512 MB

IBM FAStT disk subsystem family, specification summary

The era of multiplatform systems and FibreChannel SAN and NAS

The FAStTPlattensysteme series was developed in **October 2001** the more powerful products **FAStT500 and FAStT700** expanded, which offered higher capacities with up to four FibreChannel connections. In the high performance area came in **February 2003** the **FAStT900** also to round off the series in the upper performance range. in the **April 2003** was the FAStT500 by the **FAStT600 and FAStT600 Turbo** replaced. In addition to the plate expansion units EXP500 and EXP700 (EXP stands for expansion), which were equipped with FibreChannel plates, IBM announced in **October 2003** the EXP100 expansion unit, which was equipped with cheaper SATA disks. The matching controller was created in **May 2004** with the **FAStT100** announced. **SATA disks**

(Serial Advanced Technology Attached) are inexpensive IDE disks with a serial interface, which so far have only been used in the home PC area. With the exception of the FAStT200, which had been available for some time, all systems were now equipped with 2 Gbit FibreChannel, both on the host side and in the connection of the plate loops to the respective FAStT controller. Over 90% of the sales of this LSI plate subsystem series was handled by IBM or IBM business partners and was a huge success in the FibreChannel environment, which continues to this day.

Depending on the strength of the controller, a large number of new functionalities were introduced in the FAStTPlate systems, which were then used as standard in the upper models and are still used in the systems to this day. These functions are described below for better understanding.

DCE and DVE: Dynamic Capacity Expansion and Dynamic Volume Expansion enable allocated storage areas to be increased during operation. Within the FAST system, physical drives are bound to individual array groups. The drives can be distributed over various EXP expansion units. The selection of the drives can be left to the system (then the best performing configuration is always selected) or manually defined by the administrator. A RAID level is defined for the created ArrayGroup, which is valid for the entire ArrayGroup. Different volumes can be defined within this array group. These volumes are assigned to the individual servers. In the event that the selected plate combination (ArrayGroup) more capacity or higher performance is required, there is the option to add one or more physical drives to this group (DCE). The volumes in this group use this new capacity, the disk is automatically included in the volume distribution.

If more capacity is required within an assigned volume, a capacity expansion is also possible here during operation (DVE). The prerequisite for this is that corresponding capacities are available within the array group (in this case, any undefined disk can be supplemented via DCE). These functionalities enable the user to respond flexibly to capacity requirements and to implement them in the most professional way. All functions work with Data in Place during operation and are initialized via the FAStT Storage Manager.

DRM: Dynamic RAID migration. Just like assigned capacities, assigned RAID levels are not static, but can also be changed during operation. If changes in the load profile, capacity or security requirements result in changes to the RAID level used, this can be adjusted for each RAID array during operation. RAID levels 0, 1, 1 + 0, 3 and 5 are available. You can switch from any RAID level to any other RAID level. The prerequisite is that the corresponding disk capacities (e.g. when changing from RAID5 to RAID1) are available (DCE). This is also possible with Data in Place during operation. The RAID calculation is carried out in the controller using HW functionality.

DAE: Dynamic array expansion. If additional physical capacity is required within a FAStT, there is the possibility to equip the connected expansion units with other FC drives in different capacity sizes and with different RPMs during operation (up to 14 drives per EXP700). If there are no free slots available in the connected EXPs, there is the option of connecting another EXP700 to the FAStTController during operation. In this way, individual drives or entire expansion units can be added during operation.

In addition to these standard functionalities, other optional functions are available:

- *FlashCopy: The function FlashCopy (Point-in-Time-Copy)*
  enables up to 4 virtual copies of an existing volume. This function is mainly used for creating test environments and increasing online times for backups. How does FlashCopy work with the FAStT systems?

  The FlashCopy command is issued via the GUI interface via the FAStT Storage Manager (FSM). In this step, a repository is also created via the GUI, a physical data area for the changes between the copy and T0, the time at which mirroring began. The size of the repository depends on the number of changes to the original volume. Experience shows that a repository size of 20% of the original volume makes sense. If more repository capacity is required in the case of longer copy storage or an above-average number of changes to the source, this area can be flexibly enlarged (DVE). The creation of a FlashCopy is completed within a few seconds. To continue using a FlashCopy, it is important to use a consistent database for T0, otherwise an inconsistent copy is created. The virtual copy created (after a few seconds) can e.g. B. be assigned to the backup server for data backup. If a change is made to the source (block A) while using a FlashCopy (target), the original is transferred to the repository before the change. If the target is changed, this change is also saved in the repository. If no changes have been made to the source, the source is automatically used when using the copy (target). the original is then transferred to the repository before the change. If the target is changed, this change is also saved in the repository. If no changes have been made to the source, the source is automatically used when using the copy (target). the original is then transferred to the repository before the change. If the target is changed, this change is also saved in the repository. If no changes have been made to the source, the source is automatically used when using the copy (target).

- *VolumeCopy: In addition to the function FlashCopy is for*
  FAStT systems with FSM 8.4 also have the VolumeCopy function available. Unlike FlashCopy, VolumeCopy creates a physical copy. Therefore, the same physical capacity as that of the source must be available for the creation of a VolumeCopy. It is possible to prioritize this copy as desired. This means that I / O activities to the host are preferred, the creation of the copy is only treated secondarily. VolumeCopy is a function that is particularly suitable for test environments (e.g. for performing release changes etc.) or for online backups.

- *Remote Mirroring: During the copy functions FlashCopy*
  and VolumeCopy provide data copies within a FAStT system, Remote Mirroring creates a copy of data over one SAN infrastructure from one FAStT to another. This function is available for the 700 and 900 FAStTSystems. Remote mirroring is a bidirectional, synchronous copy. Up to 32 volumes can be mirrored to a second FAStT via dedicated FibreChannel ports. Remote mirroring is used for disaster preparedness and high-performance copies.

- *Partitioning: To different storage areas*
  To separate FAStT from each other, the partitioning function is available. The delimitation means that defined servers can only access the memory area assigned to them. Access to the memory areas of 'foreign' servers is therefore not possible. This increases the security of the individual memory areas. This function is particularly valuable in Microsoft environments, since the operating systems do not offer such functionality.

FAStT RAID controller family

Drive expansion units, management software and the built-in rack

EXP100
Serial ATA technology

EXP700
2Gb FC-AL technology

Drive units

FAStT Storage Manager (FSM)

The rack

With the announcement of the first **Server-based storage architectures** With the disk products DS6000 and DS8000, IBM undertook **September 2004** a general, uniform re-branding of the plate products. DS stands for Disk System. The number behind it should reflect the performance of the respective system: the higher the number, the higher the performance.

In addition to the standardization of the plate product names, the names of the respective functionalities have also been adjusted accordingly. Comparable functions of the DS6000 and DS8000 were referred to in the FAStT systems with the same terms. This particularly affected the RemoteCopyMirror procedures.

In the DS6000 and DS8000, the synchronous mirroring of disk systems is referred to as Metro Mirroring, the asynchronous mirroring process as Global Mirroring. The newly introduced terms are to take due account of the newly developed mirror, which was developed as a platform on the ESS Model 800 and is today the most powerful mirror on the market.

The following rebrandings were carried out for the FAStT plate families:

Naming Prior to Sept 7, 2004	New naming as of Sept 7, 2004
IBM TotalStorage FAStT Storage Server	IBM TotalStorage DS4000 FAStT
	DS4000
FAStT Family	DS4000 series
FAStT Storage Manager vX.Y (example FSM v9.10)	DS4000 Storage Manager vX.y (example 9.10)
FAStT100	DS4100
FAStT600	DS4300
FAStT600 with Turbo Feature	DS4300 Turbo
FAStT700	DS4400
FAStT900	DS4500
EXP700	DS4000EXP700
EXP100	DS4000EXP100
FAStT FlashCopy	FlashCopy for DS4000
FAStT VolumeCopy	VolumeCopy for DS4000
FAStT Remote Volume Mirror (RVM)	Enhanced Remote Mirroring for DS4000
FAStT Synchronous Mirroring	Metro Mirroring for DS4000
FAStT Asynchronous Mirroring (New Feature) w / o Consistency Group	Global Copy for DS4000
FAStT Asynchronous Mirroring (New Feature) with Consistency Group	Global Mirroring for DS4000

The **now renamed DS4000 disk subsystem family** was in **May 2005** with a new one **High performance model DS4800** added, which delivered twice the performance compared to the DS4500. The DS4800 was already a 4Gbit machine when it became available and was the first 4Gbit disk system on the market. The cache sizes at the beginning could be configured with 4 GB, 8 GB and later 16 GB depending on the model. The DS4800 also had another positive feature in construction: the machine was **RoHS Compliant** and already complied with the new EU directive in 2005, which has been in force since July 1, 2006.

In 2003, a body was formed in the United States with the task of testing disk subsystems from different manufacturers for real and productive workloads and publishing these results. This body, which is firmly established today, is called **'Storage Performance Council' or SPC.**

The performance of the new DS4800 as a midrange plate system far exceeded that of all other manufacturers and achieved a top score of 42,254. This makes the DS4800 the fastest disk system in the mid-range market today.

Since the older DS4000 products, with the exception of the newly available DS4800, did not meet the RoHS directive, in May 2006 all non-RoHS-compliant products were replaced by new ones.

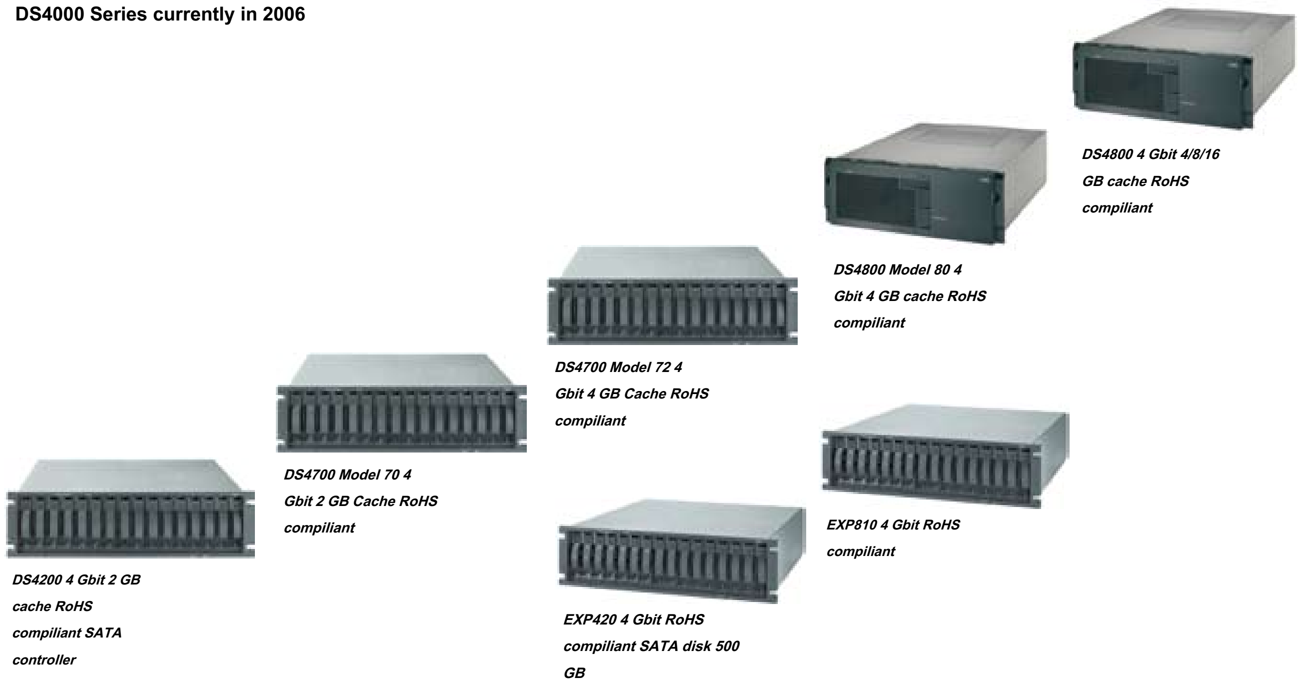
RoHS stands for 'Restriction of Hazardous Substances'
and since July 1, 2006 it is a legally controlled EU directive that specifies exactly which materials, Alloys, barium derivatives etc. may be used in new products. This new directive applies across the EU and applies to all new products in the electrical and IT industry that will be distributed from July 1, 2006.

The DS4100 as SATA controller was replaced by the new one **DS4200 SATA controller** replaced the EXP100 with 250GB and 400GBSATA disks by the EXP420 with 500GB SATA disks.

The DS4300 Dual Controller was developed by **DS4700 model 70** and the DS4300 Turbo through the **DS4700 model 72** replaced. Likewise, the EXP710 with FibreChannel disks was replaced by the EXP810 with FC disks on a 4Gbit basis.

The DS4500, formerly FAStT900, was replaced by a new entry-level model, the **Model 80 of the powerful DS4800,**
replaced. With this entry-level model 80, the internal transmission bandwidth was slightly reduced and the processors of the controllers were provided with a 100 MHz clock (in comparison to a 133 MHz clock). The Model 80 of the DS4800 thus closes the performance gap between the high-performance models of the DS4800 and the midrange Model 72 of the DS4700.

DS4000 Series currently in 2006



Current models of the IBM DS4000 disk family in 2006

DS4000 extensions 2007 and 2008

In 2007, IBM launched the **DS4000 disk family** a whole bunch of extensions. in the **May 2007** were made for the DS4000 **750 GB SATA disks** and **300 GB FibreChannel disks with 15,000 revolutions** introduced in the minute. This means that a DS4800 scales up to 68 TB or 168 TB storage capacity (depending on the use of the disk type). In October 2007, many functional expansions for all DS4000 systems were added. RAID6 is available as a new RAID option for the DS4200 and DS4700 systems. RAID6 secures the failure of two disks in the RAID array, since two parity schemes are used (see also under RAID). The number of mirror relations has been doubled, as has the number of flash copies. With availability in February 2008, LUNs larger than 2 TB can be configured for all systems.

The **powerful DS4800** got the extension that up to 512 servers can be connected to the disk system. There are now four current models of the DS4800 available, which differ mainly in the cache size. Depending on the model, the cache sizes 4 GB, 8 GB and 16 GB are available.

The RAID levels 0, 1, 3, 5 and 10 are operated (the new RAID6 level is only available for the DS4200 and DS4700). The calculation of the RAID level or the

XOR calculation in the DS4800 takes place in one **specially developed ASIC.** This is a special processor designed specifically for this workload. This makes the DS4800 free of latency because it does not load the cache or other parts of the machine that are relevant to throughput.

View and structure of the high-performance model DS4800

The DS4800 has configured all hard drives in the EXP810 expansion unit. Additional expansion units as well as individual hard disks can be added and put into operation during operation.

All components of the DS4800 are designed redundantly,
e.g. B. Controller, power supply and cooling. The cache is also mirrored and is protected against loss by a battery for up to 72 hours. In addition, the midplane of the DS4800, which is referred to as the Interconnect Module, can be replaced during operation. So far, this is unique for plate systems in the midrange environment.

The performance of the DS4800 is also unique. The machine still holds the leading position in the benchmark of the Storage Performance Council. This is a random I / O-oriented benchmark, in which real workloads are stored. The exact test data and the description can be found at www.storageperformance.org.

The DS4000 Storage Manager is supplied free of charge with all systems for all systems. The multi-pathing software and the call home functionality can also be used at no additional cost. The functions FlashCopy, VolumeCopy and eRVM are available as premium features. These are not free of charge, but only one license per system and function is used.

In terms of capacity, performance and price, the current DS4000 series offers a tailor-made plate solution in the FibreChannel environment for every end user.

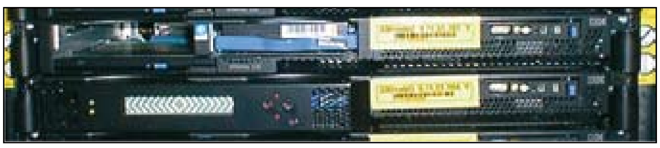
SAN virtualization

The **Epoch of multi-platform systems** was strongly influenced by the introduction of FibreChannel networks and the establishment of SANs. That was it **Storage virtualization in SAN** on one of the most current topics in the storage environment - although the concept of virtualization is not new. Storage virtualization concepts have long been in use in the mainframe (e.g. DFSMS) or in the UNIX area in the form of 'logical volume managers'. The use of storage area networks (SAN) has accelerated the development towards storage virtualization. Likewise, the complexity of today's heterogeneous infrastructures with regard to servers, storage networks and storage subsystems.

The primary approach to storage virtualization was to decouple physical storage resources from direct association with server systems. This SAN-based solution establishes a virtualization level between the server and storage systems. The primary goal is to share memory across the entire storage hardware as well as all server platforms and operating systems. Virtualization in the storage network makes it possible to integrate, upgrade, migrate, replicate and distribute storage resources regardless of the platform.

To get this tremendous flexibility within a SAN, IBM in Hursley, UK developed the **Product SAN Volume Controller, also SVC** called that in June **2003** announced and became available in September 2003.

The IBM **SAN volume controller** was developed for an application in which the capacities of several heterogeneous storage systems are combined into a single storage reservoir that can be managed from a central point. It enables changes to physical storage systems with minimal or no impact on applications running on the hosts, and minimizes downtime through scheduled or unplanned events, maintenance, and backups. In addition, the SVC increases the utilization of storage capacities



IBM SAN Volume Controller SVC

Online availability and the productivity and efficiency of administrators. In addition, to further simplify operations, it offers the option of using extended copy services across multiple systems for storage systems from many different providers. In its fourth release, the SAN Volume Controller was designed to manage even larger and diverse storage environments. With expanded support for numerous third-party storage systems, such as B. EMC, HP and HDS, the SAN Volume Controller allows the establishment of a multi-level storage environment, so that each file can be saved on the corresponding subsystem due to its importance. The **latest version of the SAN Volume Controller,** the in **July 2006** became available based on **4 Gbit technology and is RoHS compliant.**

Belt systems

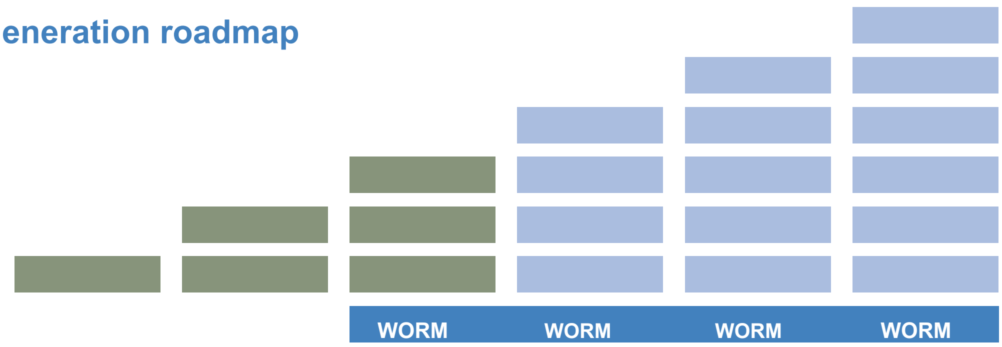
In addition to the further development of the disk subsystem and advances in the field of SANG fiber optic networks, the further development of tape technologies also remained exciting. After the introduction of the **LTO tape technology 2000 (** Ultrium 1) with 100 GB native capacity on the cassette was released in February **2003** Generation 2 (Ultrium 2) of the LTO drives with a cassette capacity of 200 GB native was introduced. In February **2005** came the **Third generation (Ultrium 3),**

again with a capacity doubling to 400 GB native per cartridge. The speed of the drives has also been massively improved. LTO2 drives already worked at 35 MB / s with 8-track technology. At **LTO3** was converted to 16-track technology and a data rate of 80 MB / s.

Read / write backward compatibility was maintained as planned. For example, LTO3 drives can still process generation 1 cartridges for reading and generation 2 cartridges can both write and read.

With the introduction of the LTO3 drive generation in February **2005** In addition to the previously used rewritable cassettes, so-called **WORM cartridges (Write Once Read Many)** introduced that no longer offer the possibility of changing or deleting stored data due to multiple labeling as overwrite protection. The WORM identification is stored on the left side of the cassette in a built-in transponder chip, which is read out with radio frequency. The label is also in the built-in memory chip

Ultrium LTO six-generation roadmap



Native Capacity Native Transfer Rate	Generation 1 100 GB up to 20 MB / s	Generation 2 200 GB up to 40 MB / s	Generation 3 400 GB up to 80 MB / s	Generation 4 800 GB up to 120 MB / s	Generation 5 1.6 TB up to 180 MB / s	Generation 6 3.2 TB up to 270 MB / s

Standardized LTO roadmap up to generation 6

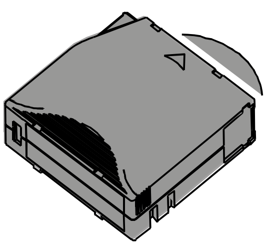
Cassette and in the tape volume control region at the beginning of the tape. In order to be absolutely tamper-proof, the manufacture of these WORM cartridges is still marked on the prescribed servo tapes: a quadruple label that ensures that LTO3WORM cartridges can no longer be overwritten. LTO WORM cassettes are marked in two colors. One half is black and the other is white.

With the availability of LTO3, the official roadmap was expanded by two more generations with LTO5 and LTO6. It can be assumed that a new LTO generation according to this roadmap will be available on the market every 2 to 2 ½ years.

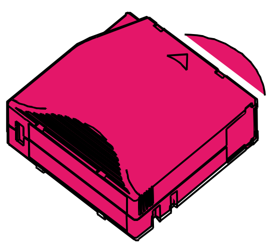
The somewhat idiosyncratic choice of colors of the standardized LTO cassettes comes about due to an important requirement. Color blind people must be able to visually distinguish the cassettes.

In 1999 the market share was still 90% with Quantum DLT technology, since then the picture has rotated 180 degrees and LTO technology dominates today with over 80%.

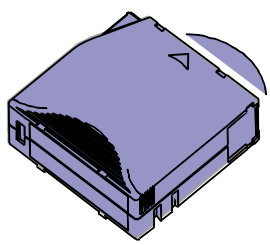
The LTO mass market with cheap LTO drives in half-high construction was previously only covered by HewlettPackard HP. With an announcement in October 2006, IBM also enters this segment and offers **mid-height LTO3 drives** in the low price range. The drives are equipped with an Ultra 160 SCSI interface and work with a data transfer rate of 60 MB / second (native).



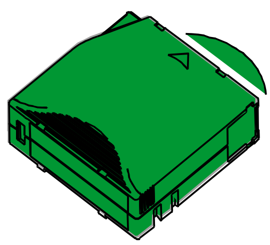
LTO1 cartridge, 100 GB capacity (native)



LTO2 cartridge, 200 GB capacity (native)



LTO3 cartridge, 400 GB capacity (native)



LTO4 cartridge, 800 GB capacity (native)

in the **April 2007** IBM announced as the first manufacturer with immediate availability **LTO4 drives and cartridges** at.

The IBM TS1040 drive is a full-height LTO4 drive and is intended for installation in IBM libraries. The drive operates at a data rate of up to 120 MB / s (native) and the cartridge capacity of the new LTO4 cartridges is 800 GB (native).

The drive interface offers a connection via a 4 Gbit FibreChannel, SCSI or 3 Gbps Dual Port SAS. The TS1120 (3592) technology has adopted the surface control guiding mechanism for optimized belt guidance. The drive also offers predictive error handling SARS (Statistical Analysis and Reporting System), an improved ECC (Error Correction Code) and digital speed matching. An internal buffer memory of 256 MB (LTO3 had only 128 MB) ensures that the data stream does not stop. Overall, the drive adapts to 6 speeds of the server (30, 48, 66, 84 103 and 120 MB / s) and thus reduces the start / stop activities. Adaptive compression of data is used for optimal use of the band capacities. WORM cartridges can be used in addition to the standard cartridges and the drives are the first LTO generation to be encrypted. LTO4 from IBM is characterized by very low energy consumption with power management function, including sleeping mode when not in use (maximum power consumption 27 watts). IBM relies on high quality, e.g. B. the use of metal, and ver avoids plastic (less interference). These features lead to a more stable operation compared to the old LTO generations. Drive / access statistics can be kept and errors predicted if necessary, so that drives can be replaced preventively before an error occurs. Many innovations in the IBM System Storage TS1120 enterprise tape drive have been integrated into the IBM LTO4 drives.

IBM TS1040 LTO4 drive without case

The differences in encryption technology are in 'KeyHandling' and are described in detail in the 'Encryption' chapter.

The IBM TS1040 drives can be installed in the autoloader TS3100 and in the IBM libraries TS3200, TS3310 and TS3500. In addition to the TS1040 drives for the libraries, IBM announced in **April 2007** also **Stand alone LTO4 drives** which can also be installed in a 19 inch rack. The 'stand alone' drives are marketed under the term TS2340 and have the same specifications as TS1040. Again, these are full-height drives. The TS2340 have a connection via 3 Gbps dual port SAS or SCSI and an LED display for administration.

LTO4 drives can no longer process LTO1 cartridges. You can read out LTO2 cassettes and set and read LTO3Kas in LTO3 mode. With the LTO4 cartridges, an LTO4 drive offers the maximum capacity (800 GB on the cartridge) and the high data rate of up to 120 MB / second.

In November 2006, IBM announced the LTO4 series **IBM TS2240** as an external LTO4 drive with half height (Half High) as a further option to FullHigh drives. The half-height LTO4 drive is available as a 'stand alone' drive or 'rack mounted'. In contrast to LTO3, the half-height LTO4 drives achieve the same data rates (120 MB / s) as the full-height drives. In mid-2006, IBM took a new one for the production of the half-height LTO4 drives **Manufacturing plant in Singapore** in operation. Many elements from the production of the full-height drives were taken over there, so that it is ensured that the half-height drives meet an extremely high standard in terms of quality and safety.

in the **February 2008** IBM announced that the **mid-height LTO4 drives** in the **TS3100 autoloader** and in the **TS3200 Library** can be integrated. Instead of a very high drive, two half-height drives can be installed. The TS3100 autoloader can hold up to two half-height LTO4 drives and the TS3200 library up to four. The TS3200 also supports mixed operation of very high and half-height drives (one very high and two half-height drives).

The **Magstar-3590** Tape development as a 1/2-inch format ran parallel to the LTO development, but the sophisticated recording process with parity information was retained. The Magstar development series took place in **June 2002**

with the **Generation 3, models 3590-H,** their graduation. The reading backwards compatibility of the previous generation, i.e. for cassettes that were written with the B or E model, was ensured.

With the new HModel, 384 tracks were written on the cassette in an even more compact form. For the Extended Length Cartridge this meant a capacity of 60 GB uncompressed and 180 GB compressed (3: 1). The data rate of 14 MB / s was maintained. Magstar thus achieved the same track density in 2002 as LTO Generation 1 in the year

2000.

With the IBM from **September 2003** newly announced

Compact drive 3592 was a **new 'tape age'**

initiated, the dimension of which is still only known to a few! But if you take a closer look at the integrated technology and the new coating type of the 3592 cassettes, you can compare this development step with 1984. In 1984, IBM initiated the switch from roller conveyor to cassette technology. With the 3592 technology, which was developed under the development name 'Jaguar', there were technological possibilities for tape cartridges and the use of tape drives, which until then could only be dreamed of.

The drive itself has next to the new one **Flat lap head technology,** the new, customized guiding system with roller bearing and the time-controlled tracking system with servo tapes (all three elements are also integrated in the LTO2 and LTO3 drive) **PRML encoding for the first time**

(Partial Response Maximum Likehood - see also technology appendix) implemented, which allows bit mapping from 1: 1 to tape.

The **PRML encoding** had already made its way onto the record in 1995, but it was previously not possible to use this method on tape. The basic requirement for using PRML is the generation of extremely well-magnetized bits on the data carrier. The coating of the 3592 cassette in connection with the FlatLap heads with 10 times higher induction strength makes it possible to produce such high quality bits. The 3592 cassette is also the first medium to work with strong induction, without negative phenomena such as. B. to get less capacity. This means that you can store approx. 50% more in the data track and thus realize 50% more capacity on cassettes with the same tape length and the same number of tracks.

The FlatLap head technology improves the read / write signal between the head and the tape, because it enables much higher induction. In addition, a vacuum is created in front of and behind the head row, which sucks up dust particles that - for whatever reason - collect accordingly (**'Self Cleaning Machines').** This head technology was fully used for the first time in LTO2 and was implemented in a more extensive form in the new 3592 drive. As a result, the quality of the bits generated is much higher and the ability to read the data again is quite second to none in the industry. It also ensures that there is almost no loss of momentum, even with long-term storage of tapes.

The 'Surface Control Guiding System' enables the read / write elements to be guided very precisely via servo tapes. The repositioning of the elements on the servo tracks is carried out via a time control, which is carried out via prescribed analog signals (analog tracks) in the servo bands and a time measurement. As a result, the surface of the tape is used for fine guidance and not the edge area of the tape where servo tracks were applied in classic recordings. This avoids the sources of errors that can occur outside due to the belt tension. FlatLap technology and the new guiding system were implemented for the first time in the LTO2 and 3592 products. PRML encoding is only realized in the 3592 because the new coating of the 3592Kas is a prerequisite for this new encoding method.

With a data rate of 40MB / s and incredibly fast winding and FastSearch times, the IBM 3592 drive turned out to be the fastest and most reliable drive on the world market. Another special feature of Jaguar technology are new, functional possibilities that could be implemented due to the 3592 coating, because a 3592 medium has no restrictions in the number of users. With the function **'Virtual Back Hitch'** Many reset times for file retrieval and all start / stop positioning times for individual file transfers that cannot be processed sequentially can be eliminated on a broad basis.

Traditionally, a drive works in such a way that it is always written to tape via the buffer memory. The drive thus receives data in the buffer memory and, via the so-called FlushBuffer command, the instruction to write out the previously received data on tape. Once this has been done, the FlushBuffer command is stored in the track as a tapemark. If there is no more data to be written, the drive stops. If something comes back into the buffer memory with the corresponding FlushBuffer command, the command cannot be executed immediately because the drive has stopped. The drive must now reset to be at the streaming speed at which the FlushBuffer command can be executed on the tape mark that was last written. This reset is called backhitch. A backhitch is therefore not a 'good thing' because it takes time and stresses the tape (especially outdoors). With Virtual Backhitch in the 3592 such positioning processes are limited to the minimum level!

How does Virtual Backhitch work now: If the Control Unit determines, due to the buffer utilization, that it is no longer possible to work in streaming mode, the drive writes a so-called RABF mark in the current track. RABF stands for Recursive Accumulative Backhitchless

Flush. The RABF stamp simply refers to a track set that lies in front of you and is actually not intended for writing. In this new track set, the drive simply continues to stream, ie if something drips into the buffer memory, a copy is made on tape in the track set that is not yet intended for writing. Since streaming simply continues, regardless of whether data is coming or not, there can of course be real 'holes' between the actually written data. The whole process could also be called **'Nonvolatile caching on tape'** designate, since the tape and the buffer are written in parallel. The tape represents the current-independent NVS write memory and the buffer memory represents the cache. If the cache breaks for any reason, any drive can still process the data written to tape, since the reference to the track set is ensured via the RABF label.

If the buffer memory is now 50% full or if the drive comes to the end of the tape, the drive turns over and writes back. If the buffer memory is full, the entire buffer memory is now written out sequentially into the correct data track set. Then the reference to the RABF brand is resolved, as if nothing had happened. Virtual backhitch or nonvolatile caching on tape is an ingenious tape functionality because repositioning times are saved and additional tape stress is avoided.

The **3592 drive** itself has **unprecedented shock and vibration properties** and can still write and read cleanly even with high vibrations. This was achieved by a self-calibrating suspension of the drive in the housing.

in the **November 2005** the second generation of Jaguar drives became available. The **3592 generation 2,** the marketing side is referred to as the TS1120 drive (TS stands for tape system), was and is the first tape drive to use **two 4 Gbit FibreChannel interfaces** is controlled and thereby realizes a data rate of 260 MB / s on a compressed basis. The 8-track technology previously used was converted to 16-track technology and the induction of the heads was doubled again, so that even 'better' bits can be written on the same medium. The 60GB cassette designed with a short tape length becomes a 100GB cassette when written with the second generation and the 300GB cassette becomes a 500GB cassette. Jaguar 2 writes 896 tracks on the 3592 cartridges with an incredible data rate of 100 MB / s. Due to the extremely well-generated stray fields of bits and tapemarks, the streaming speed during high-speed search was increased to 10 ms. This means that the tape moves forward at 36 km an hour when a file is requested. This high speed is unique in the market and, despite the high capacitance, allows average access times of 27 seconds (100 GB cassette) and 46 seconds (500 GB cassette). Electricity consumption was also reduced to 46 watts. Comparable drives today require twice the power.

The unique function of the virtual backhitch was further optimized by using a 512 MB buffer memory (Generation 1 had a buffer memory) with 128 MB and two built-in control units.

In comparison to LTO, Jaguar technology also offers WORM (Write Once Read Many) cartridges in addition to rewritable cartridges. **3592 WORM cartridges** are high priced compared to LTO3WORM cassettes, but in addition to all WORM security, they offer the advantage that **Updating always possible** is, even if faulty tracks or track areas should occur in the area where updating is to be carried out. This is ensured by the fact that not only the WORM marking is applied to the 3592 cassettes on the servo tapes, but also so-called **Append flags,** which are used for the error correction. LTO3WORM cassettes work without app end flags, which means that if there are faulty areas when the cassette is updated, the LTOWORM cassette simply cannot be rewritten.

IBM announced one at the end of October 2006 **new 700 GB cartridge** for the TS1120 drives. This means that three cartridges are now available (overwritable and as WORM): the 100GB cartridge with 120 meter tape length, the 500GB cartridge with 609 meter tape length and the new 700GB cartridge with 825 meter tape length.

With the launch of LTO in the year **2000** IBM announced a new tape archive, the IBM 3584, for enterprise-wide use in the OpenSystems area. In the course of the following years, the 3584 library was significantly expanded. Today she puts it **strategic library platform of IBM**

since **June 2005** In addition to the OpenSystems platforms, the **zSeries (** Mainframe) is supported as a server platform. They turn the 3584 into a tape archive that can be used company-wide for all computer platforms and that outperforms all comparable archive products on the market due to its high performance and flexibility.

Top-class IBM TS1120 (3592 Generation 2) tape drive

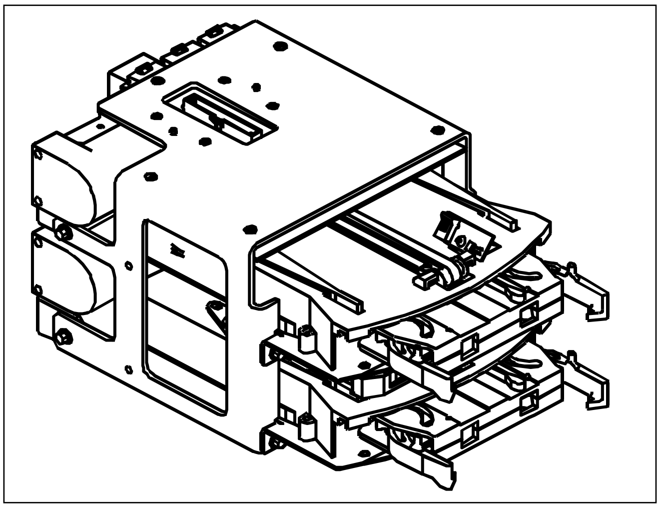
IBM 3592 WORM cartridges in platinum gray color and WORM labels

IBM 3584 basic unit with input / output station

At CeBIT in Hanover **the 3584 always ensured a lot of people.** Many cannot believe that it is technically possible to operate a robot at this incredibly fast speed. At

a robot service time of less than 3 seconds In a 2-frame configuration (fetching the cassette, loading it into a tape drive, unloading the cassette from another tape drive and returning it to the compartment), it is almost impossible to see the moving cassettes with the naked eye. It really couldn't be faster!

This speed is possible because of itself **a newly developed gripper (IBM patent)** hooks the notches into the LTO and 3592 cassettes, pulls the cassettes into a shaft and the gripper remains hooked until the cassettes are returned to a drive or to a parking space. The cassette cannot be lost even at the highest speeds or fastest rotations. All robot movements are servo controlled to take these speeds into account.



IBM 3584 double gripper

LTO drives and cartridges as well as 3592 drives and cartridges can be operated in the 3584 archive. Mixed operation in separate frames is also possible.

As introduced earlier in the older 3494 library, the system was implemented in the **Spring 2005** so expanded that with it

two robot systems can be operated, whereby the second robot is always operated in active mode. Numbers of well over 1,000 mounts are achieved per hour. Furthermore, the 3584 can be expanded with two robot systems with frames almost without interruption (max. 60 seconds, whereby no robot command can be lost).

Since May 2005 can the **3584 Archive also on zSeries server** be connected. This applies only to the operation of 3592 drives and was realized by the fact that on the 3584 **now several Library Managers 3953 L05 and J70 ESCON / FICON controllers** can be operated. In contrast to the 3494 library, the library manager and J70 controller are no longer integrated in the library frames, but are installed in external frames 3953 F05 and connected to the library. This offers a much higher configuration flexibility. The 3584 supports up to four library managers and thus the connection of up to eight virtual tape servers (3494B10 / B20).

With VSS mirroring (Peer to Peer VTS) a 3494 library can now be on the primary side and a 3584 library on the secondary side or vice versa, ie in VTSPeertoPeer mirror operation the mixed operation of both libraries is possible.

There are almost no limitations for operating native drives on zSeries systems, since up to 64 x J70 controllers with up to 192 drives can be configured in the 3584.

The 3584 archive offers significant advantages in every respect compared to conventional libraries. A **Multipath architecture** With direct connection of LTO and 3592 Fiber Channel drives, the library can be partitioned logically (up to 192 partitions). The options are also available **'Control Path Failover' and 'Data Path Failover'**

to disposal. With the function **ALMS (Advanced Library Management System)** the logical partitions can be dynamically enlarged and reduced, regardless of where the drives or cartridges are located in the library.

The **WWN (World Wide Name)** The address is assigned to the drive sled and allows a drive replacement without having to 'reboot'.

By connecting the 3584 to the zSeriesServer, IBM clearly signaled that the 3584 with its unique robot system represents the strategic library platform for future developments. The 3494 Library also benefits from this further development, as it receives all the extensions that are developed at the Library Manager. In addition to the 3584, the 3494 will also integrate all planned new generations of the 3592 drive (SOD from May 2005).

in the **May 2006** was the **IBM 3584** in the course of the frame changes **4 Gbit FibreChannel technology in TS3500 (TS stands for tape system)** renamed and corresponding new frames are available. The IBM 3584 or TS3500 library is produced by IBM itself. With the changeover to 4Gbit technology, the **RoHS compliance** of the TS3500 Library product announced. The previously used were also used **J70 control units,** that were required to connect TS1120 drives to the mainframe are more powerful thanks to new RoHS-compliant **C06 control units** replaced with 4GbitFICON connections. RoHS stands for 'Restriction of Hazardous Substances' and has been a legally controlled EU directive since July 1, 2006, which specifies exactly which materials, alloys, barium derivatives, etc. may be used in new products. This new directive applies across the EU and applies to all new products in the electrical and IT industry that will be redistributed from July 1, 2006.

In the middle and lower system segment, IBM did not produce its own libraries, but bought the 'naked' Libra ries through OEM contracts and then equipped them with IBM LTO drives.

IBM 3583 models L18, L36 and L72, capacity with LTO1 cartridges up to 7.2 TB (native), ADIC sold the product as Scalar 100 under the ADIC logo

With the announcement of the first LTO drives in the year **2000** were too **OEM libraries** and Autoloader, which were further distributed as an IBM logo product after the conversion / expansion. So were **1999** an autoloader from ADIC **as IBM 3581** Product and a medium-sized archive system also from ADIC **as IBM 3583** introduced. The autoloader was equipped with an LTOSCSIL drive (LVD or HVD) and could manage up to eight LTO cartridges. The archive system 3583 could be equipped with six LTO drives and, depending on the model, manage 18 cartridges, 36 and 72 cartridges. In the beginning, only SCSI drives were installed. In order to be able to connect the library to a FibreChannel network, up to six SCSI drives were daisy chained via an integrated FCAL gateway.

This FibreChannel connection of the IBM 3583 caused addressing problems with certain HostFibreChannel adapters. For this reason, a separate IBM controller was installed in the 3583 in May 2003. Thanks to a multi-path architecture, it was possible to directly operate FC drives in the library and to partition the library. Up to three logical partitions were possible in the 3583.

in the **May 2004** this midrange library portfolio was supplemented by a mini library, the **IBM 3582.** The library is an OEM product from ADIC. It is equipped with the IBM controller for connecting FCLTO drives and offers the option of being operated with two logical partitions.

1 to 2 LTO drives were installed in the MiniLibrary and up to 23 cartridges could be managed. Nine cassettes had built-in spaces, the other 14 spaces were covered by two removable 7-cassette magazines.

in the **May 2006** the ADIC OEM autoloader was replaced by a new autoloader from **BDT IBM 3581** replaced, which could also manage up to eight cassettes, but was constructed in a flat design and could therefore be integrated into a standard 19-inch rack in a very space-saving manner.

In the midrange library series, too, the changeover to the new EURoHS directive, effective from July 1, 2006, had to be carried out by June 30, 2006.

The Library 3583 was already in the **October 2005** by a **new library TS3310** replaced. The TS3310 also comes from ADIC, had integrated prepared 4Gbit technology and was up to the availability period **RoHS Compliant.**

In the beginning, the library could only be configured with the basic unit, an expansion unit and up to six LTO3 drives and 122 cassette slots, in the **May 2006**

then the extensions up to the maximum expansion with up to four expansion units and thus up to 18 LTO3 drives and up to 398 cassette slots were announced. With this new library, IBM in turn uses its own FC controller to connect the drives. This makes the library partitionable and up to 18 logical partitions can be operated. The assignment of drives and storage spaces to logical partitions is much more flexible than with the predecessor 3583.

IBM 3581 models L28 and F28, autoloader from BDT

IBM 3582 mini library with up to 23 LTO cartridges. ADIC sold the product as Scalar 24 under the ADIC logo

The new IBM Midrange Library IBM TS3310 with up to 18 LTO3 drives and a maximum of 398 cassette slots, here with an expansion frame. ADIC markets the product as the ADIC 500i under the ADIC logo

In the course of the entire RoHS conversion, both the autoloader from BDT IBM 3581 and the MiniLibrary from ADIC IBM 3582 were replaced by new products from BDT. The 3581 was built by the **IBM TS3100** replaced, a new autoloader with a 4GbitFibreLTO3 drive and up to 22 cassette slots. The MiniLibrary 3582 was developed by the **IBM TS3200** replaced that up to two **4 Gbit LTO3** Drives and can hold up to 44 cassettes. The TS3200 has the same partitioning option as its predecessor.

This means that all library products and tape drives that IBM sells are **consistently on 4 Gbit technology** converted and correspond **the EU RoHS** Directive that came into force on July 1, 2006 for new products.

Tape virtualization in the open systems environment

So far, tape virtualization was only useful in the mainframe environment (see VTS Virtual Tape Server). With the increase in speed and data rate, it became more and more difficult to take full advantage of the speed of such drives. An LTO3 drive now works at 80 MB / s, a Jaguar TS1120 drive even at 100 MB / s (native). In order to make maximum use of these high drive speeds, disk buffers are increasingly being used as a buffer. In order to establish a suitable, sensible backup concept, you have to switch between the classic backup via the IP network (**LAN backup)**

and the **LAN-free backup,** a distinction is made between where the data is transferred to the tape drives via a Fiber Channel network (SAN).

If the backup is traditionally run over the LAN, the most sensible type of virtualization is the IBM **TSM (Tivoli Storage Manager)** to be backed up to a disk buffer via the TSM server and migrated from the disk buffer to tape under the control of the TSM. The TSM has had this functionality for many years. In this way, the high speeds of the drives can be used. The advantage of this solution is automation, because you can use all TSM policy rules accordingly.

In the LANfree area, virtualization makes sense if many LANfree clients are operated, because each LANfree client requires a dedicated tape drive. If a data center runs the backup with many LANfree clients, you need the same number of physical drives, and this can be expensive. Virtual tape archives make sense here because you have many virtual drives available. In order to do justice to this fact, IBM announced in **October 2005** the **virtual tape system TS7510** for open systems environments.

IBM TS3200

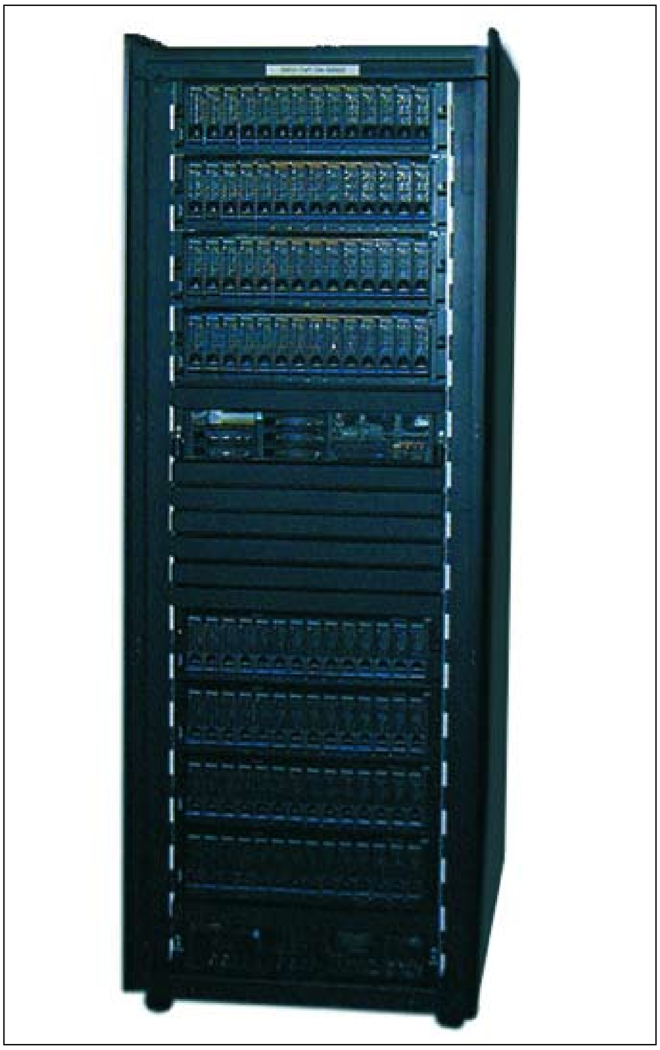
Tape virtualization for Open Systems (Virtual Tape Library VTL)

The virtual tape archive **TS7510** consists of several important components and is server-based. XSeries computers with a Linux kernel are used as servers. The machine can be equipped with one (single) or two xSeries (dual server configuration). The dual server configuration offers the possibility of 'active failover and failback'. The servers used represent the virtual unit and emulate virtual tape drives. Up to 512 virtual drives are emulated per computer unit. With a dual server configuration, up to

1,024 virtual drives emulated. LTO2, LTO3 and 3592 (Jaguar 1) drives are shown. Up to 128 virtual libraries and up to 8,192 virtual volumes are available, i.e. 64 virtual libraries and

4,096 virtual volumes.

It is written to a disk buffer. For this purpose, two plate controllers and corresponding plate expansion modules based on the DS4000 are installed in the housing. The first frame is expanded with a second frame if larger capacities are required. The minimum configuration allows 5 TB of usable disk space and can be expanded to the maximum configuration of up to 46 TB. A total of eight 2GbitPorts are available, four for disk access and four for the host and the physical tape drives. The machine can be mirrored via two 1Gbit Ethernet ports via the IP network (remote replication). Compression and encryption can be used as additional options.



Vitae tape library IBM TS7510

1999 to 2005

The era of multiplatform systems and FibreChannel SAN and NAS

Archiving solutions

Archiving Information is a very old topic, one thinks of the cave paintings of the Chau vet grotto in the Vallon Pont d'Arc in the south of France, which are estimated to be around 31,000 years old - probably the first transmission of information from human hands.

Since then, the techniques for providing information to posterity have improved. First, information was carved in stone and burned in clay until the Egyptians discovered the papyrus - the predecessor of today's paper - about 4,000 years before our time, in order to hold information on it using a simple technique. As technology progressed, it was now possible to store more information and leave it for posterity. Let us just think of the Old Testament, which spanned a period of approx.

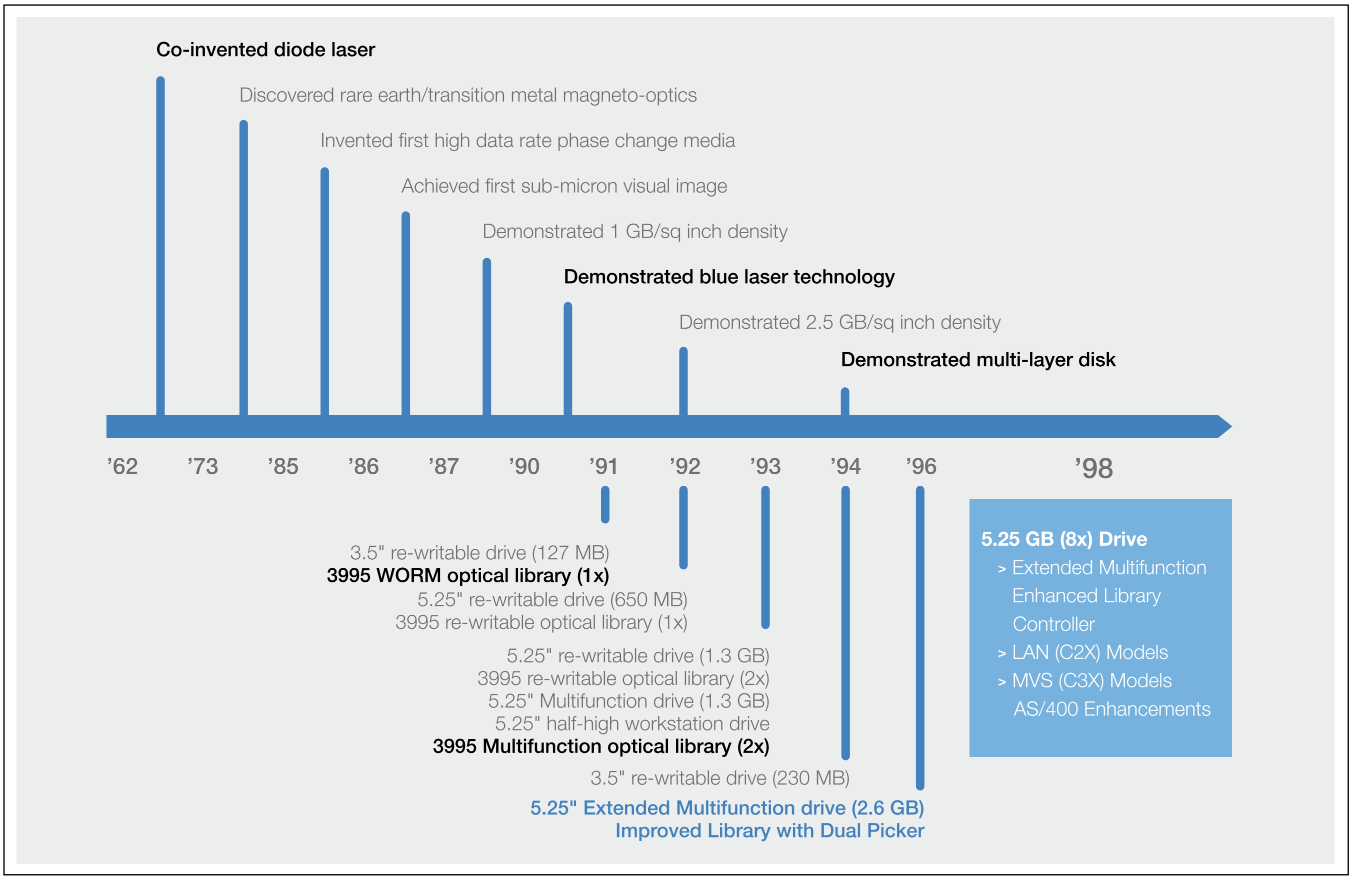
2,000 years before the birth of Christ and consists of a total of twelve books.

With the invention of buck printing by Johannes Gutenberg from Mainz in the middle of the 15th century, it was then also possible to simply reproduce information, which has contributed to many interesting traditions and of course also to the growth of information.

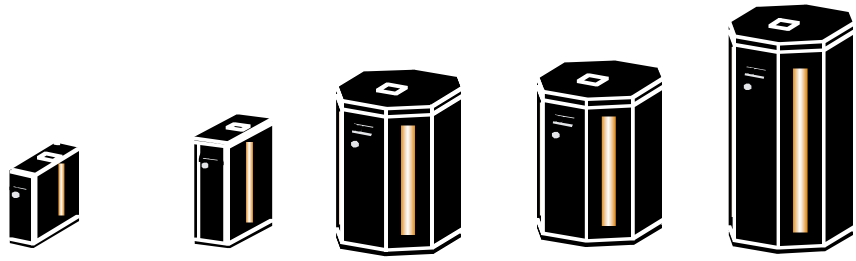
Already **1982** IBM and Sony founded a development alliance with the aim of **optical technologies** to further develop together. In the IT environment at the beginning of the 1990s, there were optical archiving options, so-called 'juke' boxes with optical disks and corresponding read / write devices, which then also became the archiving units for long-term archiving. Three different optical media were established, **the WORM disk (Write Once Read Many),** the **magneto-optical disc MO**

and the 'ScheinWORM' plate, the so-called **CCW (Continous Composite WORM) board,** a magneto-optical disc that is given a label during manufacture to ensure that the medium is not accidentally overwritten.

Already **1992** began to standardize the WORM, MO and CCW formats developed for the IT industry, based on the RoteLaser technology. An ISO standardization body was formed, and even the German DIN formed the NI23 working group, a committee of the German DIN for the standardization of optical data carriers, which worked directly on the international ISO and gave its corresponding position. The ISO standard, adopted as a 1 x standard, reflected a 650GB disk in all three formats with which



Development of optical technologies in the 90s



MODEL	CX0	CX2	CX4	CX6	CX8
Cartridges	20	52	104	156	258
Drives	1 or 2	2	2 or 4	4 or 6	4 or 6
Capacity	104 GB	270 GB	540 GB	810 GB	1.34 TB
Attachments: X = 6 X = 2 X = 4 X = 3 X = 1 (Expansion)	 SCSI LAN AS/400 -- --	 SCSI LAN AS/400 S/390 S/390	 SCSI LAN AS/400 S/390 --	 SCSI LAN AS/400 S/390 S/390	 SCSI LAN AS/400 S/390 S/390
Model Upgrades	--	--	From	To	--

IBM 3995 optical archiving system with the individual model variants with 8 x standard

2 x standard came the 1.3GB, with the 4 x standard the
2.6GB disk. In 1998 the last standard was adopted as the 8 x standard, which was still based on the red laser. The 8 x standard offered on all three formats
5.2 GB capacity per disk.

In order to ensure backward compatibility, the standard was designed in such a way that media of the 1 x standard could be read and written by read / write devices of the 2 x standard and that even the 4 x standard was still able to process media of the 1 x Reading standards. The break came with the 8 x standard, since it was no longer intended to be able to process media of the 1 x standard on read / write devices of the 8 x standard.

In the mid-1990s, the WORM plate - recognized by legislators for many archiving requirements - was specified with a shelf life of 300 years. Many companies switched to long-term archiving on these optical JukeBoxes. IBM offered that in this area **IBM 3995 optical archive system** at.

After the adoption of the 8 x standard, nothing happened regarding a meaningful standardization for many years. This was due to the fact that a 16 x standard was technically not feasible, since the red laser showed a scatter in this area that did not allow 4 pits to be accommodated on a pit produced according to the 8 x standard by optimizing the laser frequency. The technological change to blue laser technology (also known as 'blue ray') also contributed to the fact that a still achievable standard in red laser technology as the 12 x standard was no longer pursued.

Here is another important note that shows how deeply IBM was involved in the development of optical technologies. Today's **blue laser is an IBM patent,** which resulted from a joint development by IBM and Sony.

in the **Year 2000** there was a massive change in long-term archiving in Germany. The **GDPdU** Directive was adopted and in the year **2001 verified again.** This guideline abandoned the requirement to use optical WORM media for certain data to be stored. This cleared the way for the development of new solution concepts that are not necessarily based on optical technologies.

If the archiving is subject to legal provisions and requirements, we also speak of revision-proof archiving. Legislators recognize digital archiving as revision-proof in many areas, but at the same time make demands on the type of archiving. In Germany there are e.g. B. the **Principles for data access and the verifiability of digital documents (GDPdU).** Many laws and regulations place general requirements on revision-proof archiving. The retention periods of archived data are usually prescribed. Furthermore, it must be ensured that the data in the digital archive cannot be changed or deleted - not so easy in a world with two to three new computer viruses per day. Some regulations also require copies of the original data to be made in separate rooms. Revision security naturally also means that the integrity of the data is demonstrably guaranteed at all times. It must therefore also be demonstrated by means of protocols that the data correspond to the original. The technology used for archiving is not required by almost any law or regulation today.

In some cases, e.g. B. in the medical and pharmaceutical field, data must be archived for 30 years and longer. One question that arises from this is: Is there a digital archive that will still exist in 30 years? A modern answer to this is that the information has to be transferred from time to time to new systems and technologies. Transfer of information and information processing systems to new systems and technologies - hereinafter also referred to as migration - is the only way today to be able to read the data even in 30 years.

Due to the changed conditions for long-term archiving, IBM announced that **DR450 system** in the **October 2003**

at. The system was built with standard components and made available as a long-term archiving solution for the OpenSystems area. The solution consisted of two pSeriesp615 servers, which were combined in a highly available HACMP cluster (AIX operating system). The IBM Tivoli Storage Manager for Data Retention was installed on these servers, which ensured that the archived files and documents were not deleted or modified within the retention period. The data was saved on SATA disks (see also technology appendix) of a FAStT600 with EXP100 expansion units. Capacities from 3.5 TB to 56 TB could be configured. The magnetic disk system was connected to the servers via a redundant, FibreChannel-based SAN.

Already in **Year 2005** came the **Further development of the DR450 with the DR550** on the market, which was converted to RoHS compliance in 2006 with all of its components.

The IBM DR550 system today consists of the IBM standard software and hardware components AIX, SSAM, pSeries p52A, the SAN components 2005B16, Disk System DS4700 and disk expansion units EXP810. Optionally, you can order a highly available DR550 system. The system also includes the cluster software HACMP. The advantage of this concept is obvious: The user receives a system whose components have been tried and tested on the market for a long time. The core component of DR550 is the System Storage Archive Manager, a derivative of the Tivoli Storage Manager for Data Retention, which prevents any change or deletion of information and thus enables revision-proof storage.

In the IBM System Storage ™ product family **DR550 there are three models:**

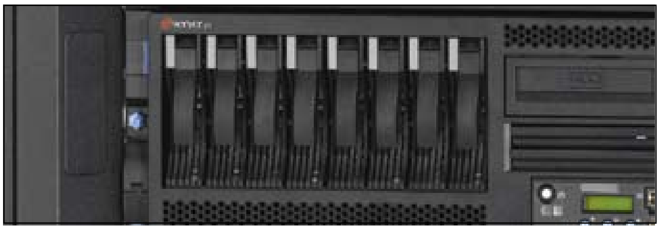
- ***The DR550 Express system is an entry-level solution that consists of an IBM pSeries Model 52A with internal SCSI hard disks configured as RAID5. A monitor kit with keyboard and mouse is supplied with the DR550 Express Model, as well as a SAN Switch 2005-B16, which is intended for the connection of tape or a disk expansion unit. A DR550 Express system is available with an entry capacity of 1.1 TB (gross). The system can be expanded by 4 TB or 8 TB (gross) by connecting a DS4700-RAID5 system. A corresponding built-in cupboard can optionally be ordered.***

- ***The DR550 single node system consists of a single cabinet in which an IBM pSeries p52A, a SAN Switch 2005-B16, a Disk System DS4700 and optionally one or more EXP810 are installed and configured. A single-node system can be ordered with a hard disk capacity of 8 TB or 16 TB and expanded up to 112 TB (gross).***

- ***The DR550 dual node system consists of the same Components like the single-node system, with the difference that all components are designed redundantly (double). In other words, two IBM p52A and two SAN switches 2005-B16 are installed in the built-in cabinet and configured redundantly. A dual-node system can be ordered with a hard disk capacity of 8 TB or 16 TB and expanded to 112 TB (gross). It is a highly available system.***

The information is stored on hard drives within the DR550 - optionally also encrypted - and allows fast access times. The performance from the user point of view can be increased with the multi-object transaction, especially if many objects are read or written at once. Several objects are saved or read within one transaction.

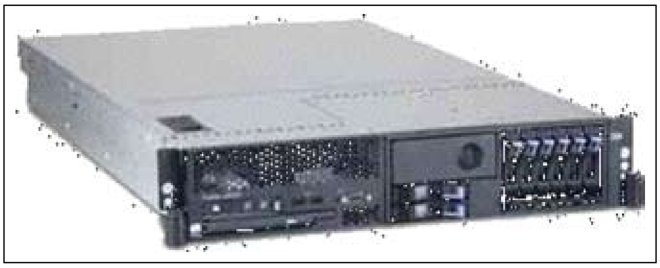
Multi-tenancy can also be realized with SSAM. This allows the logical separation of storage areas for different clients and also the reporting of used storage capacity. This ensures that a client can only access the data that is stored in its partition.



DR550 Express

The archiving system IBM DR550 is connected to the application via the Tivoli Storage Manager for Data Retention API (Application Programming Interface). Typical applications for archiving data on a DR550 system are document management systems such as B. IBM Content Manager or Enterprise Content Management Systems, such as. B. Open text Lifelink Enterprise Archive Server. The TSM API is freely available to the user. All applications that have implemented the TSM API - regardless of which platform this application is operated on - can archive and read data objects on the DR550.

IBM DR550 dual-node system IBM



DR550 File System Gateway

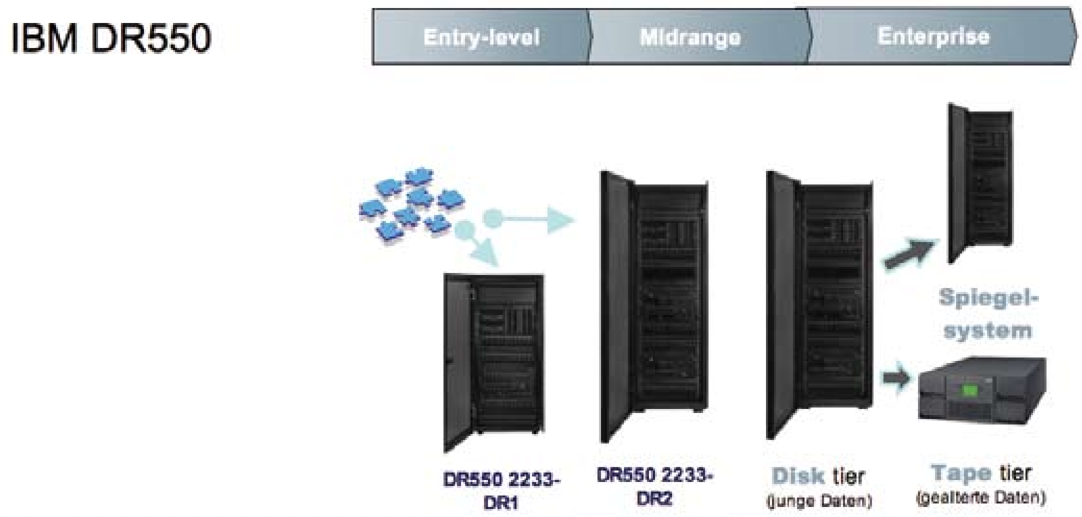
in the **May 2007** IBM announced for the DR550 **File system gateway** at. A file system can be set up via the DR550 gateway in which the data is protected from being overwritten. The gateway is delivered preconfigured. A CIFS or NFS file system is issued externally. The gateways can be 'clustered' to build a HA-capable (high availability) solution.

in the **August 2007** IBM provided the **DR550 750 GB SATA disks** to disposal. With these new large plates, a DR550 can accommodate up to **168 TB capacity** get extended.

since **August 2007** stand with the DR550 **Complete solution packages** for the **Email archiving** to disposal. These complete solutions for Lotus Domino and Microsoft Exchange address the requirements of small and medium-sized companies. The one-stop solution package contains coordinated, scalable software and hardware components from IBM and can be implemented quickly. Based on company-specific retention and access profiles, the solution enables secure management and archiving of emails, including file attachments, throughout their life cycle. Business requirements as well as national or international compliance regulations are taken into account, as is the wish of many companies to optimize storage requirements and reduce administration costs. The complete package for email archiving for the SMB area (Small and Medium Business) can be expanded at any time to a comprehensive archiving solution that contains all unstructured company information such as B. digitized corres pondez, office documents, faxes, presentations, audio and video files etc. can be managed. A connection to SAP for documents and data archiving is also possible. In this way, the complete package combines state-of-the-art technology with rapid ROI as well as maximum investment and future security.

The complete package consists of the user-friendly and powerful email archiving solution IBM Common Store for Lotus Domino and Microsoft Exchange as well as the IBM Content Manager as the basic repository. In addition, there is an IBM System x3650 server with Intel Xeon QuadcoreProcessors, which was specially developed for demanding corporate tasks such as Enterprise Content Management (ECM), Virtualization, Enterprise Resource Planning (ERP) or database applications. The IBM System Storage DR550 serves as the storage component, offering powerful functionality for storing relevant documents in accordance with legal retention periods and regulations on magnetic storage media. The DR550 supports non-erasable and non-rewritable data storage.

in the **February 2008** IBM did that **DR550 solution as a machine** with two models DR1 and DR2 in the **Version 4.5**
available. This clearly shows that IBM is investing heavily in the information retention segment. Through the extended use of the System Storage Archive Manager (SSAM) for policy-based information retention operation, the DR550 enables transparent and automated movement of archived data between different storage classes. This can save costs without endangering the security of the archived data. The now award-winning system is now available as a machine. Two models are available: The DR1 consists of a 25U rack, is pre-integrated and is particularly suitable for medium-sized customers. The DR2 was specially developed for large companies and is housed in a larger 36U unit rack.



Kapazität:	0.88 – 48TB	8 – 224TB	Erweiterbar in 8TB Schritte
Tapeunterstützung:	Ja	Ja	Kapazitätserweiterung und Backup
Zweites Serversystem:	Nein	Ja, als Option	
Spiegelsystem:	Nein	Ja, als Option	

The physical connection of the DR550 to the server systems takes place via Ethernet interfaces and is based on the TCPIP protocol. An Ethernet interface is used as standard. Optionally, 2 or more Ethernet interfaces can also be connected, which allows the data throughput to be scaled.

The DR550 also offers the connection of other external storage technologies such as. B. WORM tape or optical storage. In general, it is recommended that an external device have a NativeWORM functionality when connected to a DR550 system.

The TSM API in conjunction with SSAM offers the application various options for controlling the retention time. An application can use the event-based retention rule to delete data objects using an event. Of course, only on the condition that the configurable minimum retention time for the data object has already expired. SSAM uses the chronological retention rule to delete the data after a specified retention period. The retention rules are defined in so-called management classes, the application then only assigns an object to one management class, which means that the corresponding retention time is assigned to the object. With the additional deletion protection, another option of the TSM API, the predefined retention rule for objects can be overridden. This can prevent an object from being deleted after the normal retention period.

Tape drives are connected to the DR550 via the SAN. The connection of WORM tape has two decisive advantages for the user:

1. Copies of the data can be written on WORM tape
disasters and are safely stored in another fire compartment.
When the data on the hard disk in the DR550 'age' and

2nd
thus the accesses become less frequent, they can be outsourced to WORM tape. The tapes require less power, maintenance and replacement and are therefore much cheaper than hard drives.

The era of multiplatform systems and FibreChannel SAN and NAS

IBM 3996 optical archives - 960 GB to 10.4 TB

In the autumn **2003** Finally, a new standard based on the blue laser became available on the market. The new 1 x standard reflected a 30 GB disk in the traditional formats WORM, MO and CCW. The Plasmon company, which was also very active in the standardization of the new blue laser technology, was the first company to offer this year

2004 JukeBoxes with the new optical plates. **In autumn 2005** came in **OEM contract between Plasmon and IBM** which allows IBM to market these JukeBoxes under the IBM logo. In the beginning, sales for connecting to iSeriesServer had been limited since **June 2006** the JukeBoxes can also be connected to pSeries-based servers. With the models 32, 80 and 174, the **IBM 3996** Capacities from 960 GB to 5.2 TB. in the **August 2007** IBM announced the use of the new for the 3996 optical archive systems **60 GB optical disks** in the formats WORM, MO and CCW. The new plate generation reflects the 2 x standard, based on the blue laser technology. The IBM 3996 thus scales to a capacity of up to 10.4 TB. The capacitive possibilities are therefore significantly below the possibilities of a DR550.

It remains to be seen whether - based on the blue laser technology - a classic optical 4 x standard will be adopted, since other solution options for long-term archiving have now been established on the market.

Added to this is the fact that the standard of the first holographic plate with a capacity of 150 GB in XY format was also adopted in 2003, also based on blue laser technology, the 2 x standard with a 500 GB plate in early 2005 and a second in late 2005 x Standard as additional standard for the 'consumer' industry in the form of a 300GB disk. Holographic CDs and CDROMs can be produced due to the poly mer coating (plastic) used much cheaper than z. B. classic DVDs or WORM, MO or CCW disks used in the IT environment.

New NAS products

The **Year 2005** was characterized by another surprise. On April 4, 2005, a close cooperation and alliance between **IBM** and the company **Network appliance** announced. This alliance enables IBM to distribute all NetApp products as IBM logo products. Since IBM was previously only limitedly active in the field of NAS (Network Attached Storage), this partnership enables IBM to offer a broad portfolio of storage solutions in the field of 'file serving'. In the beginning, only the small NetApp devices were available under the IBM logo, since the end of 2005 the devices of medium performance class, since June 2006 corresponding gateways and since August 2006 the high-performance devices and thus almost the entire product range of NetApp.

IBM Nseries N5000 and Nseries N7000

The **IBM Nseries** offers answers to the questions of simplifying the storage infrastructures and central data management, backing up the data and restoring them in the most professional way.

With the N series, IBM heralds the convergence of the storage worlds. NAS stood for 'simplicity' and a wide range of functions, arguments such as high performance, scaling and reliability were mentioned on the subject of SAN. These differences are now eliminated with the Nseries.

There are generally two model types available. Systems with internal disk drives (filers) and systems that use external SANDisk resources, the so-called NAS gateways.

The scaling of the two system series ranges from entry to midrange to enterprise storage capacities and connection options. An outstanding feature is that all systems use the same Data ONTAP operating system.

The Nseries products offer a great opportunity to connect servers in the network with their specific access protocols. These include the NASFileI / OP protocols (CIFS, NFS) and the BlockI / OP protocols iSCSI and FCP.

The Nseries offers enormous flexibility with regard to the expansion stages and the possibility to put together various disk technologies for the respective solution. Fiber Channel Diskdrives and SATA Diskdrives can be mixed.

An Nseries, equipped with FibreChannel Disks, can be used for a MissionCritical, HighPerformance and transaction oriented environment. Nseries, equipped with SATA disk drives, can be the ideal choice for customers who are looking for a disk-to-disk backup option or who are interested in archiving.

All Nseries systems use a single operating system with an enormous variety of additional SW options, starting with storage and system management, through inbound and outbound copy services to a complete D / R solution with integrated backup functions.

IBM N3700

Since the subject of 'legally compliant archiving' is becoming increasingly important, the Nseries WORM offers functions. This means that data can be saved as 'not erasable / changeable' in order to take the relevant guidelines into account. Certain memory areas in the system or the entire series can simply be defined as 'WORM' areas. There is broad support from the well-known ApplicationManagementSW manufacturers.

The Nseries offers software functions that make it easier for the system administrator to manage his Microsoft Exchange, Microsoft SQL, IBM DB2 and Oracle databases. With 255 SnapShots (PointinTimeKopien) applications can be easily restored in case of errors. A patented RAIDDP algorithm ensures high data availability and protects against data loss if two disks in a RAID group fail. This is particularly useful when using SATA disks.

The entry-level system **N3700** offers a capacity of up to 16 TB, the middle systems **N5200** up to 84 TB (up to 168 LUNs) and **N5500** up to 168 TB (up to 336 LUNs). 72 GB, 144 GB and 300 GB disks can be configured. In the middle systems there are two

2.8GHzXeon processors used. The so-called 'head' of the filer (the term has become established for the controller of the appliance solution) is also available as a gateway in the medium-sized systems and uses the available disk capacity in a SAN.

The high performance systems Nseries 7000 consist of the **N7600,** which capacitively scales up to 420 TB (up to 672 LUNs), and the **N7800,** which can be expanded to up to 504 TB (up to 672 LUNs). Four 2.6GHzAMD Opteron processors are used with the N7600 and eight 2.8GHzAMD Opteron processors with the N7800. In both high-performance systems, 500GB SATA disks can be installed in addition to the other disks. The 'heads' of the high-performance systems are also available as gateways.

To round off the scalability of the Nseries product portfolio between N5000 and N7000, IBM announced this in November 2006 **Model N5600** as the new most powerful model in the N5000 series with availability from December 8, 2006. Based on a 64-bit architecture, the N5600 offers capacities of up to 252 TB and 30 - 40% higher performance compared to the N5500 and thus closes the gap between the N5500 and N7600.

To complement the N5600 appliance, IBM announced in **May 2007** the NAS **Gateway N5600** at. At the same time, with availability in **June 2007,** the Nseries series will be the new models **N5300** and the NAS **Gateway N5300**

added. The N5300 integrates a 64-bit machine and capacitively scales up to 126 TB. The 64Bit machine makes the gateway ideal for 4GbitSANs. in the

August 2007 there are two new entry systems, the entry system **N3300,** that can be expanded to a capacity of 24 TB, and the **N3600,** that can be expanded to up to 69 TB capacity. The old N5500 and the N5500 gateway were in the **October 2007** withdrawn from sales.

in the **October 2007** is a new management software for the Nseries products **Virtual File Manager VFM** made available. This software product was developed by Brocade and offers the possibility of file virtualization.

in the **February 2008** IBM announced one **new generation of nseries** for large data center operations.

The next generation of the N7000series is available both as an appliance and as a gateway. The new models
N7700 and **N7900** enable greater scalability for large-scale data center operations. The

new systems offer scalability with up to 1,176 TB
Capacity support for memory-intensive requirements in data centers. The N7000series enables IT operators to consolidate SAN and NAS storage requirements on a single system.

New disk drives are available for the EntrySystems. For the N3300 and N3600 SATA or SAS drives in the controller are now also supported. The N3300 has expanded scalability up to 68 TB and the N3600 up to 104 TB. In addition, **the following models Capacity increased:** N5300 (from 168 to 336 TB), N5600 (from 252 to 504 TB), N7600 (from 420 to 840 TB) and N7800 (from 504 to 1,008 TB). The SnapManager for Office SharePointServer from Microsoft is now available on all systems of the Nseries.

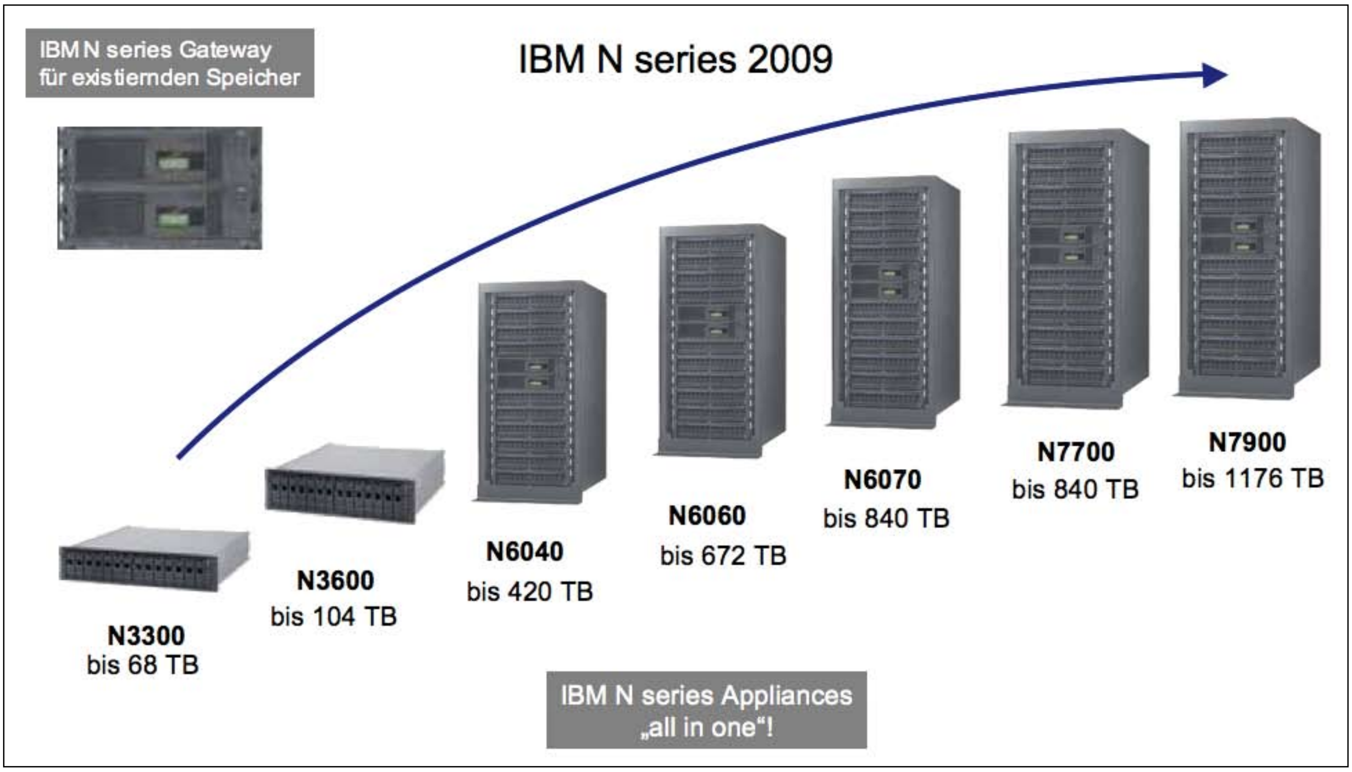
In August 2008 IBM announces the new Nseries systems **N6040** and **N6070** at. The N6040 offers capacities of up to 420 TB and the N6070 of up to 840 TB. The N6000 series is launched in **February 2009** through the model **N6060**
with a capacity of up to 672 TB. in the

As of October 2008, the N5200, N5300 and N5600 systems will be withdrawn from Marketing in October 2008. Compared to the N5000 series, the N6000 series offers 1.4 to 2 times higher performance because they are based on modern hardware and have more cache. The N6000 systems also have the option of working with newer operating systems based on Data ONTAP.

Nseries supports a variety of servers and operating systems via FCP, CIFS, NFS and iSCSI. The connected disk arrays can be configured in RAID4 or RAIDDP (Double Parity), which corresponds to the RAID6 level.

With the Nseries family, IBM provides tailor-made disk storage solutions in filers and NAS environments.

There are currently 40 different software functions available for the IBM Nseries. The most important are described in the appendix.



SnapShot

SnapShots are readable, consistent instant copies (point-in-time copy) of a database, more precisely a file or a LUN. Up to 255 SnapShots can be created per file / LUN. The special thing is that they are space-saving, which means that they do not require any additional disk space when they are created. This is done by storing only the reference to the data (pointer). SnapShots can be triggered manually or processed automatically. B. PointinTime copies are created every 5 minutes. Due to the high number of SnapShots (255), a very granular and timely backup / recovery is possible. Since SnapShots are stored in a special directory, it is also very easy for the user to restore their deleted files themselves (using drag & drop).

SnapRestore

SnapRestore uses the up-to-date SnapShot copies to restore a file, a file system or a LUN. If SnapShots are only readable copies and must be copied into the active file system in order to restore the data, SnapRestore can easily be used to put the active file system onto a SnapShot copy. This means that no data is copied around for data recovery. Such a restoring can be done very quickly with a single command! A 300 GB database can be restored with this technology in about 2 to 3 minutes.

SnapManager

SnapManager was specially developed for databases (Oracle and MS SQL) and mail applications (MS Exchange). Based on SnapRestore, the SnapManager offers a highly integrative SW package for the system administrator. Instead of complex scripts, the user can set up simple backup tasks via GUI, initiate restoring processes or create database clones for test purposes with a 'click'. It is even possible to easily restore individual mailboxes (with calendar entries, contacts, attachments) for the MS Exchange environment after an error.

SnapMirror

If SnapShot and SnapRestore are backup / recovery functions, a function has been developed with SnapMirror that provides a D / R function for mirroring the data from one Nseries to another. SnapMirror reflects Data - to put it simply - synchronously or asynchronously via SAN or LAN on a Disaster Recovery Site. If the productive side fails, the data is available despite the failure. When the productive side is operational again, the data is copied back from the DR side.

SnapVault and SnapLock

SnapVault and SnapLock are solutions for backing up and archiving data. SnapVault allows you to backup one series to another (or many to one). By using cheap SATA disks, the backup system can be backed up quickly and inexpensively. Since the disk is restored, the data can also be restored quickly. SnapLock offers the option of using all or part of the disks as a WORM area (Write Once Read Many). Data that are stored here can only be deleted and / or changed after the retention date has expired.

Open System SnapVault (OSSV)

OSSV behaves like SnapVault with the difference that the OSSV (Open System SnapVault) software server can back up and restore. For this, the OSSV SW must be installed on the server (Open, ie Unix and Windows). The servers then back up to an Nseries as a backup system.

Cluster failover

Cluster Failover (CFO) is a standard component for every DualController (A20 or G20) of the Nseries. If one controller fails, the second controller automatically takes over the server I / Os and data access is retained.

MetroCluster

If you expand the cluster failover function (CFO) across building boundaries, you get the MetroCluster.

The MetroCluster replicates the data from controller 1 of the primary data center to controller 2 of the secondary data center and guarantees data consistency and availability. If the primary site fails, MetroCluster allows failover with a single command from the administrator. Unlike SnapMirror, MetroCluster is not a DR, but a business continuity solution. Based on the design, in which the data is already in sync on both sides and with immediate access, an immediate continuation of the business processes is possible. This takeover is automated to such an extent that the

Administrator only has to issue a single command. Stretch MetroCluster offers protection of up to 300 m between two Nseries systems. Fabric MetroCluster also offers protection of up to 100 km with SAN switches.

A-SIS (Advanced Single Instance Storage)

DataDeduplication can be used to reduce the amount of data stored on a backup or archive system. At Nseries, this function is called ASIS (Advanced Single Instance Storage Deduplication). Identical data is only saved once, which can save up to 70% in storage capacity for certain applications. At system level, the Nseries analyzes identical blocks in the background (for the application transparent) and saves them only once. ASIS is ideal for archives or backup systems.

SMBR (Single Mailbox Recovery for Exchange)

In contrast to Lotus Domino, in which each mailbox is saved separately as a separate database (and thus represents an easily recoverable unit), Exchange merges various mailboxes and stores them in .edb and .stm files. These files become very large over time. This makes restoring more difficult because you have to work with very large .edb and .stmFiles if a single mailbox needs to be restored.

The solution for this is SMBR, a software that searches these large files to restore the desired mailbox (including attachments, folders, etc.). SMBR is based on the SnapShots and does not have to run on the production server. A 'content analysis' is an option of the SMBR: with it the content of emails, attachments etc. can be analyzed and logged.

SnapDrive

SnapDrive is the software that makes it possible to manage volumes from the application server, enlarge them, trigger consistent SnapShots and clone LUNs (without FlexClone feature).

If the volumes for a database need to be increased in a classic storage subsystem, the server administrator must then initiate a process at the latest to request more storage space.

With SnapDrive this is now very easy: From the GUI of the application server you select the appropriate volume and enter the new larger capacity. The volume is enlarged in the background (i.e. without interruption). SnapDrive communicates with the Nseries to perform these actions. SnapDrive saves time and costs and helps to simplify processes, as these can be done via the GUI. In addition, SnapDrive can also create new drives and 'mount' them (ie make them available to the servers). SnapDrive supports MSCS (Micro soft Cluster) and VSS and should always be used for iSCSI and FC. This function supports Microsoft, Windows and Unix / Linux derivatives.

SyncMirror

SyncMirror is a standard feature where each I / O is actually written to two separate disk pools. This corresponds to a RAID1 within a system. SyncMirror can best be compared to LVM (Logical Volume Manager) on AIX.

RAID-DP

Conventional single parity RAID technologies offer protection in the event of a disk drive failure. No other error in the data reconstruction time is expected to occur. If an error occurs in this phase, it can lead to data loss. This is particularly the case with the ever-growing SATA drives, which can now hold 1 TB of data volume. RAID DP (Double Parity) secures the failure of two DiskDrives. RAID DP writes 16 stripes, which are calculated beforehand (data and parity), to the DiskBackend at the same time.

Save drives take effect if a disk fails. RAIDDP thus corresponds to RAID6.

FlexVol

FlexVol is a standard SW that was developed for LUN management. FlexVol uses the entire disk storage pool of the Nseries to form the individual LUNs. All available 'spindles' in the backend are used without I / O bottlenecks due to dedicated disk allocations.

FlexShare

FlexShare prioritizes the service for important workloads by assigning priority levels from 1 to 5 for each volume. So z. B. the I / O of a database can be used preferentially over file services. FlexShare is a standard feature of the Nseries.

FlexClone

FlexClone enables multiple clones of FlexVols to be created (e.g. for test purposes, QA, DWH etc.). FlexClones do not consume any additional storage space at the time of creation. They are based on SnapShots and can be created and changed in seconds. They have the same performance as FlexVols including all the properties of a FlexVol (dynamic zooming in and out). One restriction has to be considered: 'Parent FlexVol' and Base SnapShot cannot be deleted as long as a dependent FlexClone volume exists. With 'Split' the FlexClone volume can be separated from its 'Parent Volume' and continue to exist independently. This requires free space in the unit to copy the 'shared blocks'.

Commentary on the era of multi-platform systems and FibreChannel SAN and NAS

The speed and hectic pace increased even more in this epoch compared to the previous epoch. Added to this were new complexities due to new FibreChannel networks (SANs) and the possibilities in the networking attached storage environment. Many new products in the SAN and NAS environment made storage more versatile, but it also made it more complex and difficult.

In this era, the storage systems detached themselves from their previous dependency on servers. Multiplatform systems such as the ESS were able to serve any server platform. In addition to these multi-platform systems, FibreChannel systems were built in parallel for the OpenSystem platforms, which could be connected to the emerging and increasingly widespread SANs. SANs and SAN strategies dominated the market more and more. That still applies today! The zSeries also switched from ESCON to FICON in order to transport ESCON protocol packages via the FibreChannel. SAN virtualization was the focus topic that IBM answered with the SAN Volume Controller. SAN management, above all central management has become one of the main requirements of data centers that operate SANs. For this purpose, IBM developed the program package IBM TPC (Total Productivity Center). After 1GbitSANs, 2Gbit and 4GbitSANs were used at intervals of around three years. 4GbitFibreChannelSAN 'End to End' solutions have been possible since 2006.

In the NAS environment, IBM mainly brought solutions in the form of gateways, which, however, could not prevail throughout the epoch. IBM had too little focus on NAS capabilities and focused mainly on SANs. At the end of the era, in 2005, there was an alliance between the companies IBM and Network Appli ance, which were very successful in the field of NAS and are still successful today. This alliance enables IBM to also offer solutions developed in the field of NAS and to expand the overall storage portfolio.

On the tape development side, in addition to the introduction of LTO2 and LTO3, one of the milestones was the introduction of the sensational Jaguar tape drive technology with the 3592 and later TS1120 drive. Jaguar is the 'vehicle' to break new ground in the tape environment. Jaguar opens up technical possibilities that we may not even be able to imagine today (see the technology appendix, chapter Magnetic tape).

In addition to optical archives, new and better solutions were created for long-term archiving because legislators no longer prescribed which medium must be used. This made better and more meaningful solutions possible. IBM developed the DR450 (Data Retention) solution, which was replaced by the DR550 and is now used successfully in long-term archiving.

The products available at the end of this multifaceted epoch will certainly accompany the subsequent epoch with corresponding further developments and offer solutions for the storage world that has become complex due to SAN and NAS. In particular, the areas of the DS4000 disk system, the San Volume Controller, the entire tape and tape archive area, tape virtualization solutions and the area of long-term archiving will be rapidly advanced with the DR550.

ynchronous copy with Metro Mirror

(up to 300 km)

Asynchronous c

(no distanc

The era of server-based storage systems and storage virtualization

A

B

The era of server-based storage systems and storage virtualization

What will the new era of server-based storage architecture and storage virtualization bring us? What can we expect technologically in the next few years?

The author allows himself to incorporate a personal and quite realistic estimate of the near future into this storage compendium.

Server-based disk systems

The approaches of server-based storage architectures can already be seen in the previous epoch in the products DR550 (pSeriesServer) for long-term archiving and the TS7500 family (xSeriesServer with Linux Kernel) for tape virtualization in the OpenSystems area. SAN virtualization with the SAN Volume Controller must also be included.

However, server-based storage architectures mean much more. Systems based on this are not only machines on which data is stored, but also servers that can run applications nen. Today, if servers are consolidated at the server level and storage units at the storage level, server-based storage architectures allow vertical consolidation of servers and storage. Applications can be moved from the server level to the storage level. This is of particular advantage for applications that produce many I / Os between a server unit and a storage unit. These I / Os can all be saved if the application is carried out by a storage unit ('No I / Ois the best I / O').

The first product in this server-based storage architecture is the DS8000 disk system **October 12, 2004** was announced by IBM. The announcement was undoubtedly much too early and it took almost a whole year until the machine had the degree of stability to be used productively. Today, the machine has unprecedented stability and performance that no other comparable system on the market delivers. The little brother, the DS6000, was announced on October 12 with the DS8000.

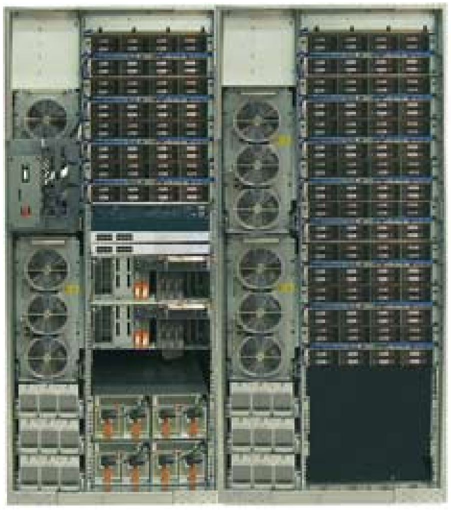
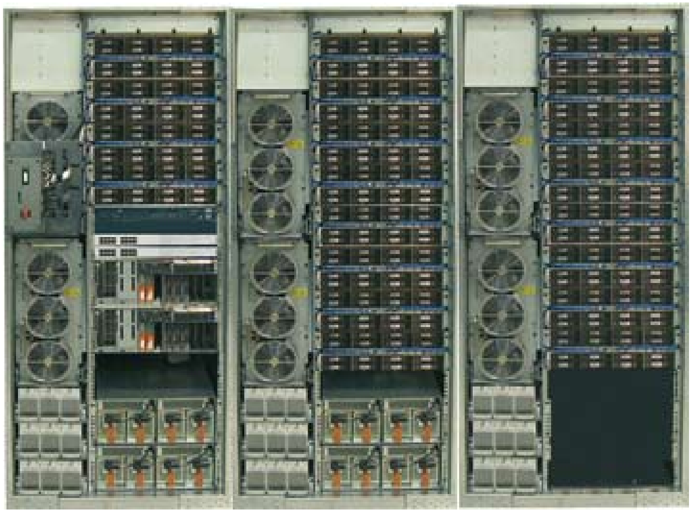
IBM DS8100 2-way system

IBM DS8300 4-way system

The DS8000 is currently the most powerful disk system on the market. With 3.4 million I / Os per second, it is unbeatable in comparison to other disk systems. However, IBM did not compete with the DS8000 storage system in order to achieve any top positions on any ranking lists, but to make data storage cheaper with less, more powerful hardware. It's about economy and cutting-edge technology helps to make the Total Cost of Ownership (TCO) cheaper. TCO consist of hardware acquisition costs and operating costs. Both are reduced with the use of DS8000. DS8000 leads to lower acquisition costs and significantly lower SAN infrastructure costs because fewer components are required. At the same time, the management and tuning effort is reduced, because the hardware is more powerful and requires less monitoring and optimization. With architecture principles such as a strip-like design and self-learning caching algorithms that automatically adapt to changing workloads, the DS8000 ensures that the hardware is used optimally.

The following table provides an overview of the scalability of the DS8000 storage system. The systems were designed with a view to almost linear scalability. This is shown in the table below. The entry into the DS8000 world is provided by the 2-way system DS8100, which scales up to 116 TB on the basis of 300 GB drives. If the capacity requirements increase, all the components required to increase performance, such as the number of processors, cache, FC / FICON adapter and disk adapter, increase in the system. This also applies to future extensions. If the capacity requirements exceed 192 GB, 8-way and 12-way systems up to the 32-way system will be feasible in the future without changing the architecture. With a tripled capacity of over 500 TB, the system scales almost linearly with because then all essential components for the provision of services such as cache, processors, FC / FICON adapters and disk adapters are tripled when upgrading from 4-way to 12-way system. The future systems are part of a long-term roadmap. The 8-way and 12-way systems will be realized as soon as there is a need for this extremely high performance on the market.

	2-way	**4-way**	**8-way**	**12-way**
Server processors	2-way POWER5	4-way POWER5	8-way POWER5	12-way POWER5
Cache	16 to 128 GB	32 to 256 GB	64 to 512 GB	96 to 768 GB
FICON (2 Gb / s) (4 ports per adapter)	8 to 64	8 to 128	up to 256	up to 384
FibreChannel (2 Gb / s) (2 ports per adapter) 8 to 64		8 to 128	up to 256	up to 384
ESCON (2 Gb / s) (4 ports per adapter)	4 to 32	8 to 64	up to 128	up to 192
Device ports	8 to 32	8 to 64	8 to 128	8 to 192
Drives 73 GB (15K RPM), 146 GB, (15K RPM) 146 GB (15K RPM), 300 GB (10K RPM)	16 to 384	16 to 640	up to 1792	up to 1792
Physical capacity	1.2 to 115 TB	1.2 to 192 TB	up to 538 TB	up to 538 TB
Number of frames	1 to 2	1 to 3	2 to 8	2 to 8

IBM DS8100 interior construction

IBM DS8300 interior construction

The DS8100 consists of a base unit and a maximum of one expansion unit. Up to two expansion units can be connected to the DS8300. **Since October 2006** Up to four expansion units can be connected to the DS8000 Turbo models. The DS8000 Turbo thus offers the option of operating over 300 TB FC disks and over 500 TB FATA disks (maximum

1,024 disk drives).

The **Storage systems DS8000** are the first storage systems on the market with real StorageLPARs. As with a server, logical LPARs can be formed on a physical unit with their own processor, bandwidth, HBAs, DiskAdapter, drives and microcode. The LPARs are robustly isolated from one another. Even a crash in one LPAR does not affect the second LPAR. The LPAR concept is therefore particularly well suited for ensuring a certain service level for certain areas of application. Production and testing can be run on one machine without the test activities affecting production. zSeries and Open Systems Workload can also be operated on the same machine without hesitation, since each LPAR is equipped with its own resources.

As of today, the DS8300 offers two LPARs, each of which receives 50% of the resources. The division will soon become more flexible and sub-processor allocations will be possible.

In the future it will be next to the StorageLPARs too **Application LPARs** give. Since the DS8000 has integrated a pSeries, all options of a server can be used for data storage. In particular, it is intended to run applications with a high affinity for storage (many I / Os) directly on the storage system and to use the internal bandwidths of the DS8000 for optimal performance.

A key feature of the DS8000 are the **superior copy services,** that enable uninterrupted data center operation. These include the functions synchronous and asynchronous copying (Metro Mirror / Copy, Global Mirror / Copy) and PointinTime Copy (FlashCopy). The mirror functions of the DS8000 are unique in the market today due to their performance (distance, number of I / Os via one line, number of simultaneous mirrors) and their functionality (consistency groups, suspend / resume option, reversing the direction of mirroring, three mirrors, etc.). There are currently **no more powerful remote copy implementation** on the market.

While DS8000 clearly addresses the high-end memory market based on the POWER5 architecture, there is still huge market potential with lower requirements in terms of performance and scalability. IBM addresses this market potential with a modular storage system based on PowerPC processor technology. What is important for this market segment is a compact, modular design that enables high performance in the smallest space, makes no special demands on the data center infrastructure and can be perfectly integrated into the existing IT landscape with servers in 19-inch racks.

For IBM, the DS6000 represents a breakthrough in space requirements. While the previous entry into the enterprise world with the ESS 750 did not mean without at least one square meter of floor space plus additional service space and a stable double floor, which had to allow a load capacity of over a ton was possible, the DS6000 reached completely new dimensions. The scale comparison below shows the great technological progress. The entry into the enterprise world is now possible with less than 5% of the volume of previous storage systems. This goes hand in hand with reduced requirements in terms of service space, data center infrastructure and electricity requirements.

This corresponds to these design requirements **Storage system DS6000.** The system is fully designed for the mass market and its specific requirements. The system is based on modular 3U racks that can be installed in existing 19-inch racks without compromising on performance. The storage system scales from 288 gigabytes to 38.4 TB.

The logical structure of the DS6000 corresponds to that of the DS8000. The difference is in the platform on which the system is operated. HA and DACode are essentially identical because they are self-sufficient units. The functionality of the storage system is also identical. The big difference is the processor platform. Since the PowerPC processor has no partitioning options, this component is not required. In addition, the processor is not controlled by AIX, but by Linux.

IBM ESS 750 with 5 TB capacity

IBM DS6000 with 4.8 TB capacity

The era of server-based storage systems and storage virtualization

View of the IBM DS6800 disk system

Significant changes also affect the abstraction layer, which isolates all hardware specifications from the functional level. DS6000 is therefore a pure storage system that cannot be used for future application purposes.

The machine basically consists of 5 components, so-called 'Customer Replaceable Units' (CRUs). These are drives, power supply, controller, battery and a 'LightPathDiagnoseModul'. Whenever one of these parts is defective, the customer will receive a spare part that can be used in a few simple steps.

The same applies to the microcode. Whenever a new version is available, the customer will be informed and can download the latest microcode from the Internet and import it without interruption.

The error analysis is supported by light path diagnostics, SNMP messages and a GUI. The fault analysis is fully supported by the system. Interactive help functions simplify repairs. For simple administration and quick implementation, the storage system has an Express Configuration Wizard, which greatly simplifies the initial configuration or reconfiguration.

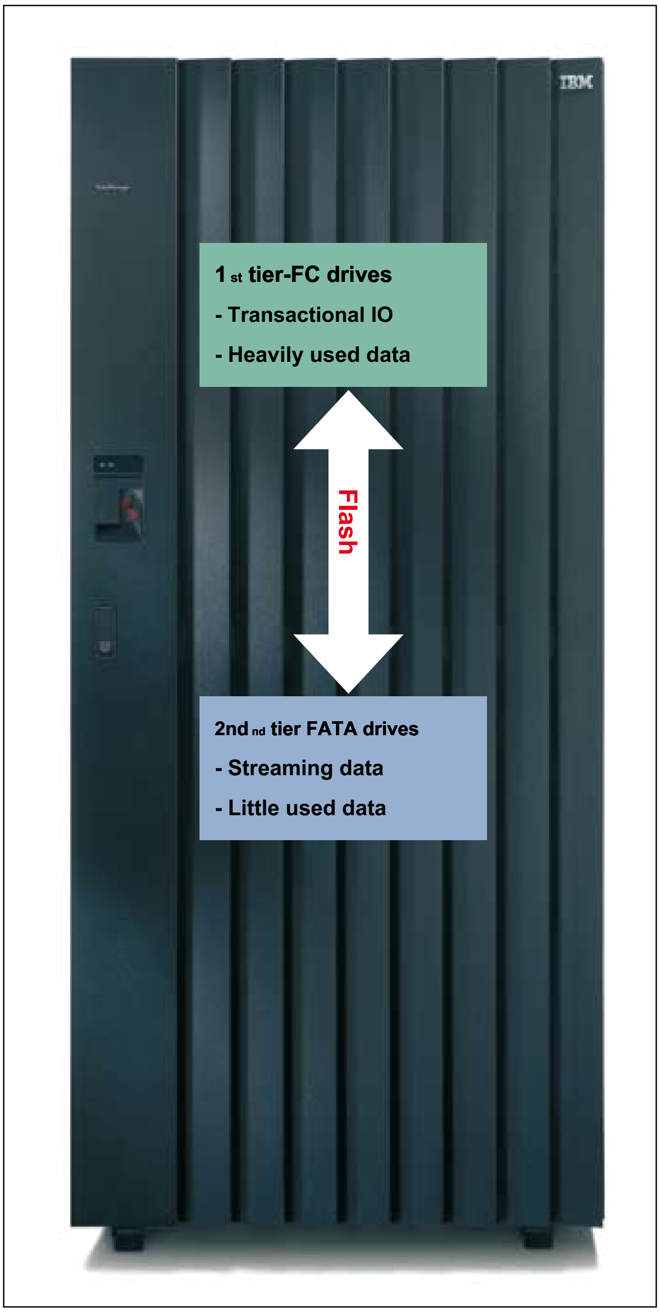
The **DS6000 and DS8000** are the first plate systems from IBM that **with a four-year warranty** were announced. Due to this long guarantee, you can already see that this new architecture will last for a very long time.

On August 22, 2006 with availability in September 2006, IBM announced the latest enhancements for the DS6000 and DS8000. In addition to the selection of FibreChannel plates, both plate systems can now also be used with cheaper FibreATA (**FATA) plates** be equipped. The plates have a capacity of **500 GB** per drive. This offers the possibility for the plate systems, within the system with a **Two-tier storage hierarchy** to work.

So they can **FC plates** for classic transaction processing and high-traffic data and the cheaper but slower ones **FATA disks**

for sequential data processing and little used data.

A fully expanded DS8000 with FATA plates offers a gross capacity of 320 TB and a small DS6000 up to 64 TB in gross capacity.



New models have been announced for the DS8000, the **POWER P5 + processor technology** have integrated and thus deliver 15% higher performance. The I / OCages were on **4 Gbit FibreChannel and FICON ports** switched. This makes the DS8000 the first disk system on the market that can work with 4GbitFICONPorts on the zSeries host. The DS8000 systems with P5 + processors are available as **DS8000 turbo systems** designated.

A new facility that as **'Synergy Feature'** , which will be available in November 2006, will significantly improve the overall performance of the DS8000 Turbo, DB 2 and AIX components and achieve significantly more effective communication between the computer and the storage system for DB2 applications. This is made possible by targeted 'end to end' priority control.

Part of the announcement of August 22, 2006 also included the option to work with a third mirror in remote copy processes for DS8000 turbo systems. This solution is called **'3 Site Disaster Recovery Solution'**

designated. This ensures constant access to the data even in the event of a disaster. Until now, this was only possible via RPQ (Request for Price Quoting).

The **triple remote mirroring** consists of the components 'Metro Mirror' and 'Global Mirror' (formerly PPRC XD - Peer to Peer Remote Copy eXtended Distance - and FlashCopy). Location A is mirrored with the DS8000Turbo synchronously with location B. Location B then mirrors asynchronously to location C.

If location A fails, the application is redirected to location B and Global Mirror continues to run without interruption between locations B and C. When location A comes online again, location B to A is resynchronized on an incremental basis. After resynchronization is complete, the application can be moved back from B to A.

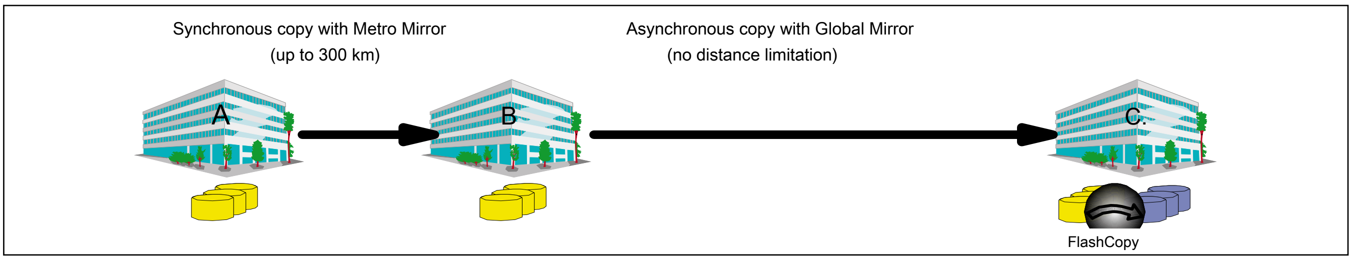
If location B fails, incremental resynchronization procedures establish the connection between A and C and thus continue to ensure uninterrupted mirror operation. If location B is available again, an incremental procedure from A to B is used to resynchronize. After resynchronization is completed, locations B and C are reconnected via Global Mirror, thus restoring the initial situation.

If location C fails, application A is not affected and the synchronous mirroring between A and B is retained. If location C is available again, Global Mirror is rebuilt incrementally and the starting situation is restored.

Parallel Access Volume (PAV): The function of writing and reading from multiple computers and applications to the same disk address simultaneously was introduced in the z / OS environment with the Enterprise Storage Server (see also ESS) and further optimized for the DS8000. The work load manager (WLM) in z / OS controlled the dynamic assignment and reordering of the necessary 'alias' addresses, which is referred to as dynamic PAV. In November 2006, IBM announced for the DS8000 **Hyper PAV**

at. The function was changed so that z / OS in combination with the DS8000 can take an alias address from a pool for each individual I / O without the need for coordination between z / OS systems. This means that you can react immediately to changing workloads. With the same workload, HyperPAV only needs about half of alias addresses and, compared to dynamic PAV with the same number of alias addresses, offers the possibility of processing 50% more I / Os.

On August 22, 2006 IBM was launched as a new software product **TotalStorage Productivity Center (TPC)** announced for replication. TPC for replication enables central management of the CopyServices functions Metro Mirror, Global Mirror and FlashCopy for the products DS6000, DS8000, DS8000 Turbo, SAN Volume Controller SVC and ESS Model 800.



In 2007, IBM provided significant extensions for the DS8000 disk systems.

in the **February 2007** IBM announced March availability for all DS8000 systems in addition to the previous disk types
300 GB FibreChannel disks with 15,000 revolutions
at. The DS8000 was the first system on the market that could be equipped with these fast, large-capacity plates. In May 2008, significant capacity expansions were added. Previously, a 4WaySystem with up to 640 plates (with 2Way up to 384 plates) could be equipped in a configuration of three units, since June 2007 it was possible to use a 4WaySystem DS8300

with up to 1,024 plates to configure. For this, four additional expansion units are connected to the base unit, in total a configuration with five housing units. When using the 500GBFATA disks, the maximum capacity increased up to 512 TB per system.

since **November 2007** further performance-optimizing features are available to the DS8000. The function **Storage pool striping with Rotate Extends** reflects a new default algorithm that distributes new logical disks in 1GB increments across the backend, thus optimizing performance without special tuning. **AMP (Adaptive Multistream Prefetching)** provides a new caching technology from IBM Research that performs self-optimizing prefetching in relation to workloads. AMP dynamically decides what and when is loaded into the cache. This can dramatically improve the throughput of sequential and batch workloads, which means that runtimes can be significantly reduced. AMP improves the read throughput from a RAID5 array by almost a factor of two! AMP can prevent hot spot situations when very high demands are placed on sequential workloads.

IBM z / OS Global Mirror Multiple Reader offers a significantly higher throughput at RemoteSpie in the zSeries environment.

In addition to the new performance features, new functional expansions were also available for the machines in November 2007. **Space efficient FlashCopy** can significantly reduce costs by reducing the required disk capacity for copies. This means that electricity and climate requirements can be reduced at the same time. Flash copies only require disk space if changes are made to the original. **Dynamic volume expansion**

Simplifies management by increasing the size of logical drives online as data grows. In addition to these extensions, the TPC (Total Productivity Center) will be renamed as standard when new DS8000 systems are delivered **SSPC (System Storage Productivity Center)** supplied and thus offers uniform access to the management of IBM and other storage systems. This provides the IT user with a uniform management interface and a common console for DS8000 systems and the SVCs (San Volume Controller).

At the **February 26, 2008** Together with the announcement of the new mainframe system z10, new extensions for the DS8000 disk system were announced, especially in the z / OS environment. With the function **z / OS Metro / Global Mirror Incremental Resync** With triple remote mirroring, there is no need to create a full copy for resynchronization after a hyper swap situation. The extension **Extended Distance FICON** Reduces the need for channel extenders with z / OS Global Mirror (double remote mirroring) and z / OS Metro / Global Mirror configurations (triple remote mirroring) due to a higher parallelism of the reading operations. The new possibility of **z / OS Basic Hyper Swap** allows the configuration of disk replication services via the graphical user interface (GUI) of the DS8000 disk systems. This option can also be used for DS6000 and ESS systems where the GUI is also available. The z / OS Basic Hyper Swap option can also be used without GDPS (Geographically Dispersed Parallel Sysplex). z / OS Basic Hyper Swap is a unique feature that automatically fails over to the remote disk system even without GDPS when the primary site becomes unavailable. However, this is only possible with the IBM systems DS8000, DS6000 and ESS. Disk systems from other manufacturers are not supported.

With the **Release 4.0** gets the DS8300 in **August 2008** the possibility with **variable LPARs** and work with more flexibility. Previously, only 50% / 50% allocation was possible, now a variable allocation of 25% / 75% or 75% / 25% of the system resources is supported.

Different versions of the microcode can be operated within the separate system resources,
e.g. B. separate production systems or test systems and that within a physical storage system. Variable LPARs are provided within the DS8300 at no additional cost from Microcode R.4.0. With R.4 the following divisions can now be made: 50/50% split (factory default), 75/25% split and 25/75% split. The division can be made for the system resource processors, NVS and cache. Host adapter slots, device adapter slots and storage enclosure slots remain in the 50/50% split.

With the **Release 4.1** in the **September 2008** the DS8000 family receives the following extensions:

Support of **Raid6** as extended plate protection, especially suitable for large plates, e.g. B. with 1 TB SATA disks. With FC disks, RAID5 is still preferred due to the high disk performance required and the fast architecture-related restoration of the RAID system in the event of a disk failure. With the integration of **450 GB FC 15K disks** On the one hand, the possible total capacity expands and thus increases the efficiency of the system. A new IBM service

Disk drive above, solid state disk below

' **Secure data overwrite** " guarantees the secure deletion of sensitive data on disks. This service is performed by a trained IBM technician.

With the **Release 4.2** in the **February 2009** further fundamental extensions come into play for the DS8000 systems. The DS8000 systems can now be used in addition to FC plates and SATA plates **SSDs (Solid State Disks)** be equipped. Solid state disks of 73 GB and 146 GB are supported. Factory installation is provided as standard; an RPQ must be provided for field installation. This RPQ is used to ensure optimal performance.

DS8000 is the first IBM subsystem to support encryption of data in conjunction with special DiskDrives (**Encryption)** on the plates. This ensures that the data cannot be read even if the DiskDrives are lost. In contrast to other solutions, encryption on DiskDriveLevel enables encryption without loss of performance. No changes need to be made in applications or the storage infrastructure. The keys are managed in the same way as tape encryption with the software product TKLM, formerly EKM (see also under Tape Encryption).

With the introduction of high-capacity **1 TB SATA disks**
The DS8000 disk system is also suitable for the use of DataDeDuplication technologies, e.g. B. in connection with the Diligent products (see under DataDeDuplication IBM ProtecTIER).

With the **Release 4.3** in the **July 2009** supports the DS8000 ' **Thin provisioning** " without impairing performance. The thin provisioning function can be ordered optionally and can be used for the entire capacity. Thin provisioning can increase the use of storage capacity and thus make a significant contribution to the economic use of the storage used. Storage administration costs can be reduced through automated storage provisioning.

A new feature that ' **Quick initialization feature** " enables quick setup of copy services.

zHPF Multi-Track Support for DS8000

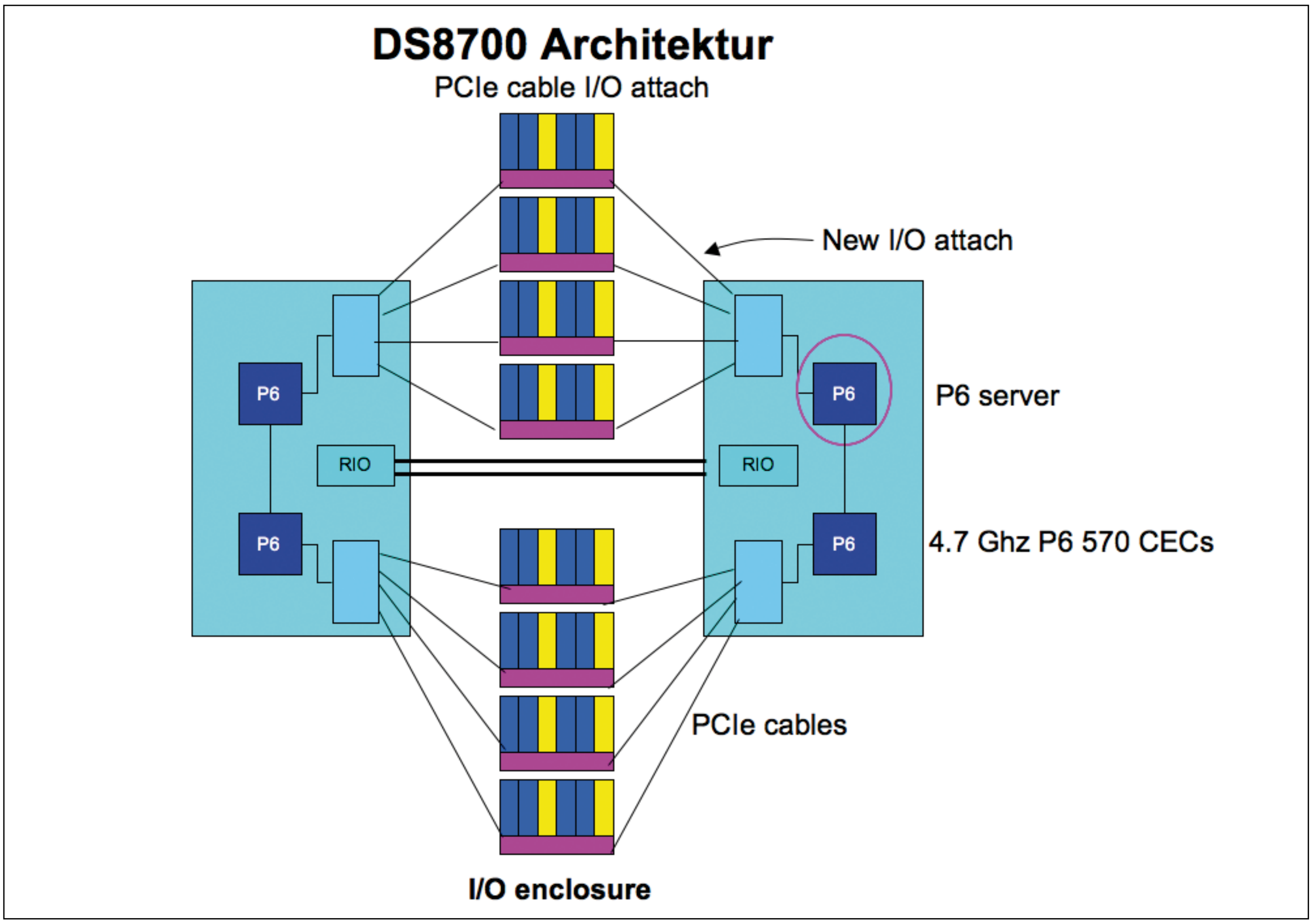
This unique feature of the DS8000 in conjunction with zSeries leads to significantly improved I / O performance. This performance improvement is achieved through an optimized I / O protocol.

IBM DS8700 disk system

in the **October 2009** IBM announces the new machine **DS8700** as the successor to the disk systems DS8100 and DS8300. The DS8700 is based on Power6 technology and works with **P6 processors with 4.7 GHz.** The performance of the machine is about 2.5 times faster than the DS8300. The RioG Loop is used exclusively for P6 'Server to Server' communication, while the disk arrays are controlled via PCIE (Express) cards. This architecture change gives the machine an optimal back-end connection and higher back-end bandwidth, which pays off especially when using fast SSDs (Solid State Disks).

Like the previous systems, the machine is equipped with the new optimized cache algorithms: **SARC (** Simplified Adaptive Replacement Cache) as the basic algorithm combined with **AMP (** Adaptive Multistream Prefetching) to optimize read operations from the disks to the cache and **IWC (** Intelligent Write Caching) to optimize the write workload. The response time behavior of the DS8700 should currently far exceed all available disk systems on the market.

Connected with the DS8700 announcement is an outlook **Roadmap until 2013.** This roadmap contains functions for optimal use of SSDs by analyzing the I / O behavior and dynamic relocation of extents, which are ideal for storage on SSDs. This new algorithm gives the machine the ability to make optimal use of storage tiers and to store the data on SSDs, FC disks and SATA disks according to I / O behavior.



IBM XIV Storage

Today, almost every company is affected by the increasing flood of digitally stored data. According to the IDC study “The Diverse and Exploding Digital Universe”, the installed storage volume will grow by 600% in the next four years alone. Storage system administration is already posing considerable problems for many companies. However, the expected high growth does not come from the classic data center area, but occurs in the area of unstructured data. While in classic data center operation, so-called 'scale up' plate systems take account of the requirements for ever faster response times (e.g. DS8000 with fast FibreChannel plates and SSDs), there are no cost-effective 'scale out' systems in the area of data growth, which, despite growing capacity, offer consistently high performance. Such a requirement can only be met with an architecture that allows standard components such as cost-effective SATA disks and PC-based processors to be integrated in such a way that high reliability and high performance are guaranteed, even if the system grows capacitively.

The former chief developer of the EMCSymmetrix systems accepted this new 'scale out' requirement **Moshe Yanai**
to found the company XIV to develop a new type of storage system that takes these new aspects into account from the very beginning. XIV was founded in 2002 and the first systems were already in production in 2005. At the **12/31/2007** XIV was aquired by IBM. By the end of 2008, more than 100 systems had been in productive use and since October 2008 the second generation systems have been available as IBM products.

IBM XIV Storage System

The creation of XIV Storage is something special! The first one sees with the new XIV architecture **Disk subsystem socialism** The Light of the World! It is not human, so socialism works perfectly! Whenever the machine has something to do, be it writing something on the disk, reading from the disk or reorganizing itself internally, everyone has to work on it. The integrated PCs, all disk drives without exception, always process the pending workload at the same time.

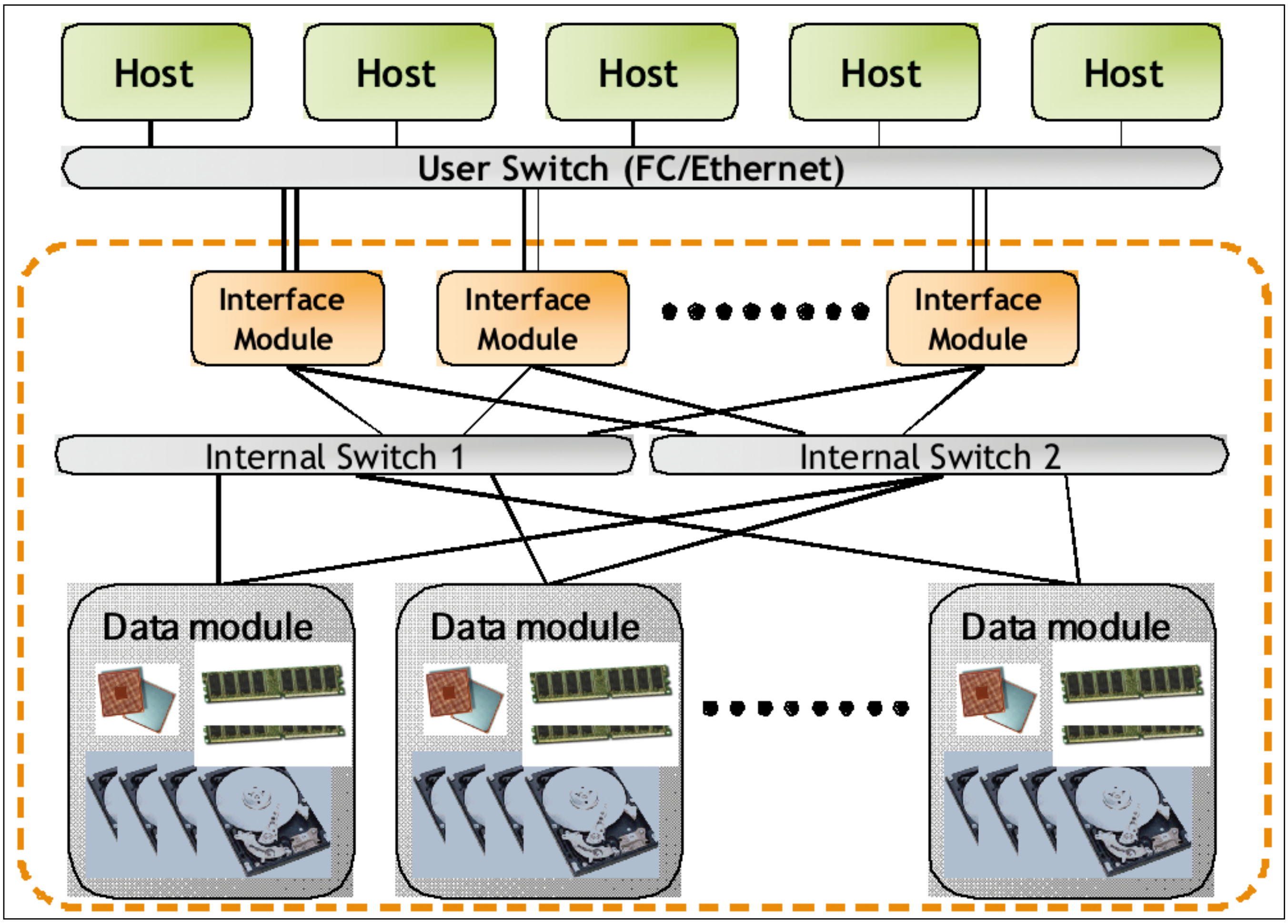
Z. B. each plate always the same fill level and the same workload. If the system grows in capacity, more 'workers', i.e. more disk drives, are available and increase the performance and throughput of the overall system. This is perfect socialism!

XIV architecture

A key feature of the IBM XIV systems is the high degree of parallelism thanks to a modular GRID architecture. Unlike traditional systems, there are not only two controllers that have to process all I / Os, but several independent modules. These modules consist of standard servers with an optimized Linux operating system. The high degree of parallelism is now made possible by the fact that a patented load balancing process distributes all I / Os across all system components: modules, switches, caches and disks.

XIV standard hardware

An IBM XIV Storage System consists of several modules, each module with its own CPUs, its own cache memory and its own (local) disks. Some modules also have FibreChannel and iSCSI interfaces, which are used for



XIV GRID architecture with distributed memory, caches and CPUs

serve redundant connection of the hosts. These modules are also referred to as interface modules. Optimized Linux serves as the operating system on the modules (servers with the latest generation of Intel CPUs), which is, however, transparent to the user. The modules communicate via two Gigabit Ethernet switches. The high performance is achieved by the parallelization of all I / Os, which always extends across all modules.

A fully expanded XIVRack consists of:

- ***15 modules with each***
  - ***Intel quad-core CPUs***
  - ***8 GB RAM (cache)***
  - ***4x Gigabit Ethernet for internal communication***
  - ***12 disk drives (1 TB SATA II)***
- ***6 interface modules with a total of***
  - ***24 x 4 GB FC adapter***
  - ***6 x 1GB iSCSI adapter***
- ***2 Gigabit Ethernet switches for internal communication***
- ***3 uninterruptible power supply units***

When using 1 TB disks, the gross capacity is 15 * 12 * 1 TB = 180 TB. Due to the high I / O parallelism, slower disks with SATAII technology can be used compared to FibreChannelDisks, without the system performance being slower than with traditional high-end storage systems with FibreChannelDisks. The advantage of using the standard components for the modules and the Intercon nect is that new technology developments can be used quickly and without new developments. For example, if faster modules in the form of standard servers or new adapters are available, they can be replaced immediately. The same applies to the Inter connect switches. Each module is currently connected via 4Gigabit Ethernet connections.

Data distribution and load balancing

To enable the most granular and even distribution of data and I / O operations possible, the system always distributes each volume created by the administrator to 16,384 (or a multiple thereof) primary and just as many secondary partitions, each with a size of 1 MB:

- ***The partitions are named after a 'pseudo-random' Algorithm distributed across all disks (and thus across all modules). This basically applies to all volumes, which always ensures optimal performance for all volumes.***

- ***A partition contains either a primary or secondary copy of the data. Primary and secondary data are always stored in different modules, i.e. also on different discs.***

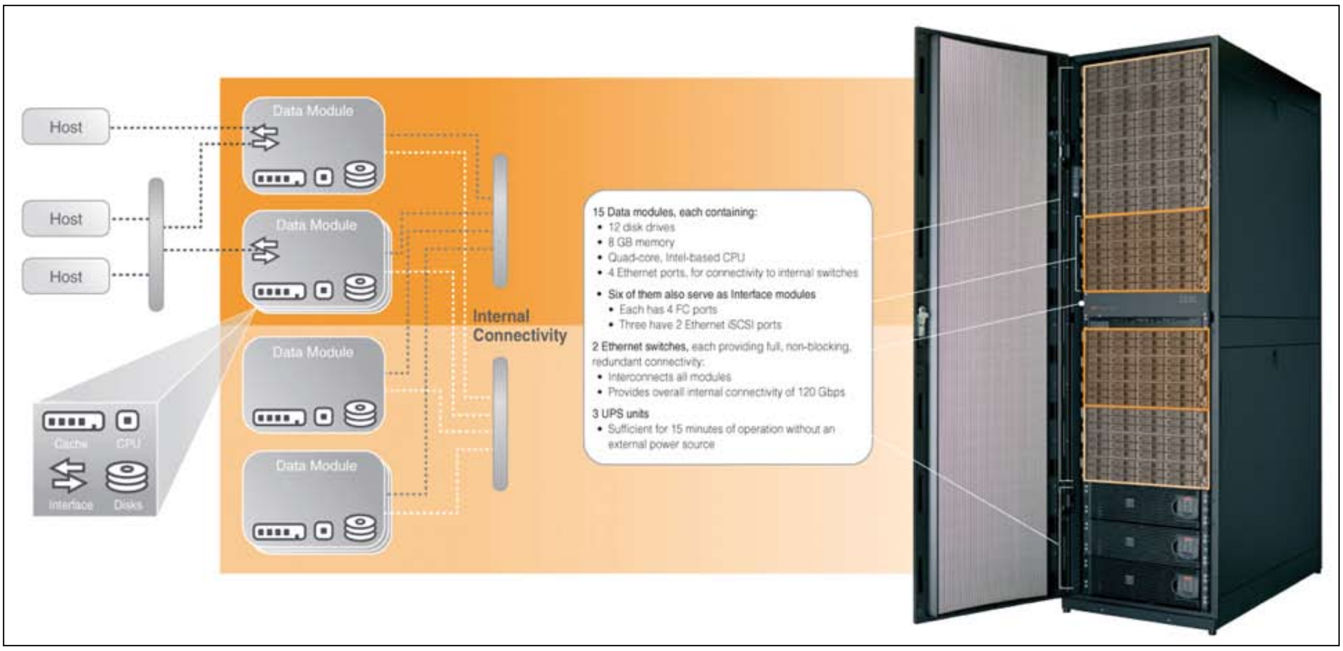
The distribution process works autonomously and dynamically. Falls z. If, for example, a disc is removed or is removed from the system, all (primary or secondary) 1 MB partitions on it are immediately copied or redistributed from all other discs. As a result, this recovery works extremely quickly. Incidentally, only partitions that actually contain data are copied. In practice, the data of a 1 TB SATAIIDisk is redundant again within a few minutes. The maximum time for making a 1TB disc is 30 minutes. It usually takes several hours, sometimes days, to recover a failed disk from a RAID5 or RAID6 array. In addition, the perfor system, because additional XOR operations are necessary when accessing data from this volume.

Since primary and secondary partitions are always "mirrored" across modules, even several disks within a module or even a complete module can fail, since even then an intact data copy is still available in the system. In these cases, too, the system immediately begins to create the missing copies in the background, always while maintaining the same distribution and thus even utilization of all system components.

The module that describes a primary partition for an IO operation sends the same data block for the secondary copy to another module in parallel to the local IO. The IO operation is only confirmed to the host when both modules have the data blocks in the cache. The data blocks are written to the respective local caches in the background.

Comparison to known RAID methods

The data distribution method shown does not correspond to any of the usual RAID1 to RAID10 mechanisms. In terms of properties, it is most similar to a RAID 10 method in which the data blocks are striped and mirrored. If you configured 180 disks in a RAID10 array, the performance of both disk arrangements would be identical at the start of operation. In practice, however, this “tie” quickly evaporates:



- ***When expanding capacity or changing volume***

  Sizes are no longer guaranteed to be equally distributed with RAID arrays. Manual reworking of the storage and host system (rebuild of striping etc.) is labor intensive and often does not take place. Sometimes entire arrays or volumes have to be moved for such operations. Even distribution of all I / Os across all system components

-

  Traditional systems require precise knowledge of the architecture and all virtualization layers. This work can often only be carried out by specialists. Outsourcing such activities to users is not possible.

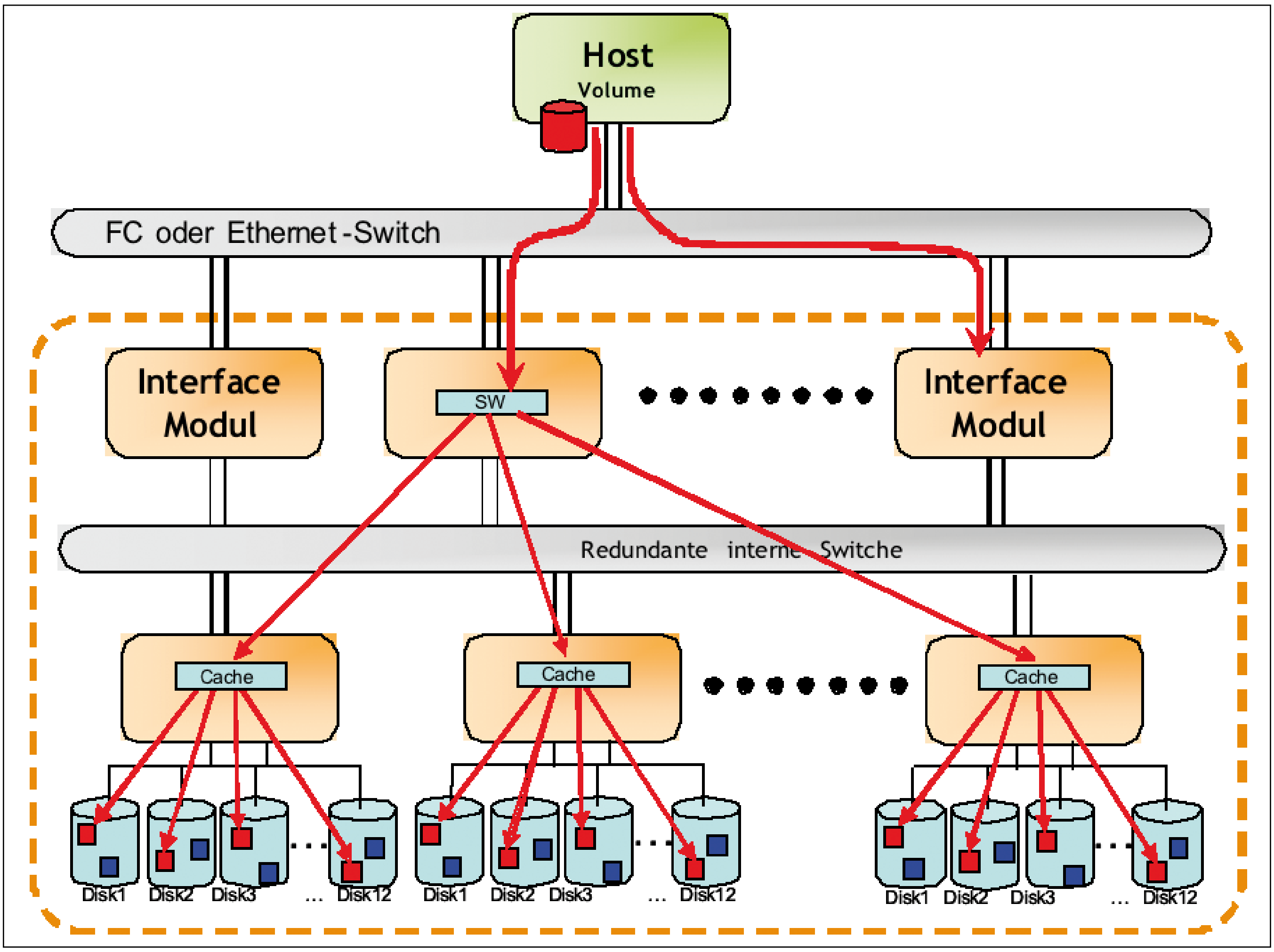
As already mentioned, such measures are not necessary (and also not possible) with XIV systems because the distribution algorithm works autonomously and dynamically. As soon as the system receives additional modules, the system begins to redistribute the 1MB partitions in order to achieve an even distribution across all system resources. Even if modules or disks are removed, a redistribution begins in order to achieve an even distribution and

Restoring complete data redundancy. In contrast to RAID5 or RAID6 methods, however, the recovery works much faster because all the remaining disks are involved in the necessary IO operations and, moreover, access to the data during this time does not require any additional XOR operations.

These properties mean that the XIV system, despite slower SATAII technology, has higher performance and higher availability than classic high-end RAID systems with fiber channel technology.

Distributed caches

The distributed architecture also provides advantages over DualController systems in terms of caches. The independent caches are only responsible for the 12 disks in the modules. This ensures minimal latency and maximum bandwidth. The full bandwidth of two dedicated PCIX buses is available for writing data from the cache to the disks.



Distribution of the 1 MB partitions of a volume across all disks in the system.

The entire CPU performance of a module is only available for this local cache, which enables three things:

- ***Aggressive prefetching of***
  Data blocks using intelligent algorithms. Very small data blocks in
- ***the cache. Traditional systems***
  work with 32 or 64 KB data blocks due to simpler cache management. The XIV system can work with 4 KB blocks due to the high local CPU performance, which results in a very efficient cache usage.

- ***A cache synchronization is not necessary because everyone***
  Cache is only responsible for its local data or disks.

A distributed cache structure also has advantages with regard to the availability required for high-end systems: Even if an entire module fails, only a fifteenth of the cache capacity and performance is lost in a single-rack system.

All caches are secured by redundant uninterruptible power supply units (UPS). The system has a total of three such UPS units, which are permanently located between the public grid and the modules, so that these short-term failures or voltage peaks are hidden. The system only starts to shut down in an orderly manner if there is a failure of more than 30 seconds. A write-through procedure common to other systems in the event of a controller failure or a power failure is therefore not necessary with XIV.

Storage capacity

As already described, each module has 12 local disks. When using 1 TB SATAII disks, the gross capacity for a full one-rack system is:

K (gross) = 12 disks * 15 modules * 1 TB = 180 TB

As described in the chapter Data distribution and load balancing, all data is always distributed across all disks, redundantly. There are no dedicated spare discs that do nothing in normal operation. Nevertheless, as much capacity must of course be kept in “reserve” as the disk should be able to fail. For these reserve purposes

the 12 disks of a complete module plus three additional ones are taken into account. The system also requires approx. 3.5 TB for internal purposes. This results in the following for the net capacity:

K (net) = (15 * 12 - 12 - 3) * 1TB / 2 - 3.5 TB = 79.5 TB

If more or less than 15 modules are used, this will reduce or increase the net capacity accordingly.

XIV functions Thin

Provisioning

With the help of thin provisioning, storage capacities can be used more effectively. In a traditional system, the storage space allocated for a volume is allocated immediately and is therefore no longer available to other applications, even if the storage space is not used at all. In practice, this leads to very poor utilization. Users want to reserve as much storage space from the outset as will be necessary in the foreseeable future, because otherwise manual rework will be necessary later. In contrast, with the XIV system, a volume can in principle be of any size. If, for example, it is known for a new application that the data volume will be 10 TB in three years, a volume of the appropriate size can be created immediately. The actually used storage space only corresponds to that of the data actually saved. If the physical storage space of the system (or pools) is eventually exhausted, new modules can simply be added. The distribution algorithm immediately redistributes the data and the capacity is expanded. No rework is necessary here either.

Incidentally, this over-provisioning is not valid system-wide, but only within a storage pool. A pool represents an administrative domain within which quota management is possible.

What happens if this over-provisioning occurs for many volumes, they grow over time and the physical capacity limits are reached? As in any storage system, capacity utilization is monitored. If configurable capacity limits are exceeded, the administrator is notified (SNMP, SMS or EMail).

If this does not respond, the system starts deleting snapshots when the workload is critical. The system prioritizes by importance (configurable when taking snapshots) and by age. If these measures are no longer sufficient, the system behavior can be configured for two reactions: either all volumes are locked or set to 'readonly'. In any case, data integrity is preserved.

In practice, thin provisioning improves storage utilization by 20–30%.

Snapshots

XIV systems are equipped with extremely fast and easy-to-use Snapsho mechanisms. The described distribution algorithm of the data presupposes that distribution tables contain the information as to which data blocks are stored on which disks. If a snapshot is now created for a volume or a group of volumes, only this meta information is saved. This is a pure memory operation, which always takes about 150 ms, regardless of the size of a volume (this also applies to deleting snapshots). The storage space used for a snapshot is also zero at this time. Only when data blocks are changed does a 'redirectonwrite' ensure that the modified data block is written to another location.

In practice, XIV snapshots therefore do not lead to any loss in performance. Up to 16,000 snapshots can be created per system.

Volume copies

Of course, volume copies can also be created. Like snapshots, volume copies can be used immediately after they are initiated. The physical copying runs in the background until all data blocks have been copied.

Writable snapshots

A snapshot can be assigned to any host. This is often helpful for further processing of the data, e.g. B. Creating backups or as a data source for a DataWarehouse. By default, snapshots are only read-only. If necessary, snapshots can also be marked as writeable with one hand. This is often helpful for tests or other scenarios in which the data must be accessed in writing.

Snapshots of snapshots

Sometimes it is desirable to have another logical copy of a snapshot. XIV systems can create multiple copies of a snapshot using the same mechanism. So a snapshot could e.g. B. be marked as writable, while the second copy is used for backup.

Consistency groups

For some applications, it is important to create consistent snapshots when the application uses multiple volumes. This applies e.g. B. to databases if this stores their log files on separate volumes. In these cases, several volumes can be grouped into consistency groups. When creating a snapshot, it is then ensured that the same contains consistent data, ie that the time stamp of the different volumes is identical. Volumes in a consistency group must all belong to the same StoragePool.

Automatic snapshot when mirroring is removed

XIV systems automatically create a so-called 'lastcon sistencysnapshot' on the target system before a mirroring process is removed. The reason for this is to ensure a consistent state in the event that the source system fails or the connection is interrupted during the synchronization process. In this case it is possible that the target system is in an inconsistent state. The automatic snapshot can then be used to return to a defined, consistent state. These automatic snapshots have a deletion priority of zero, which means that they are not automatically deleted if a pool is over-provisioned.

Remote mirroring / disaster recovery

For disaster-resistant installations, it is possible to create distant mirrors on another XIV system. This can be implemented either via dedicated iSCSI or FibreChannelPorts.

Two options for synchronous mirroring are currently supported:

Best effort:
If the communication path or the target system fails, work can continue on the primary system. As soon as the link is available again, the blocks changed during this time will be re-synchronized.

Mandatory:
If the communication path or the target system fails, write access is no longer possible on the primary system. This type of implementation requires special care when planning the entire SAN. All components (fabrics, links, switches etc.) should be designed redundantly.

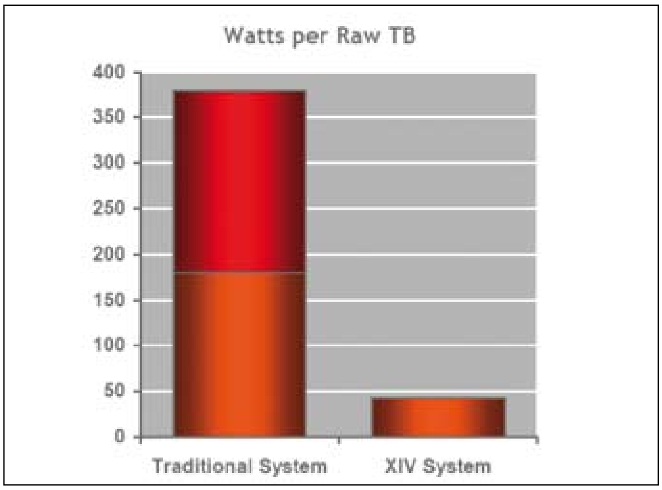
Mirrors can be created per volume and multiple target volumes can be defined.

Hosts connected to a target system can see the volumes, but they are in 'readonly' mode. If the secondary target system becomes the primary system in the event of a disaster, the volumes can be written to. The volumes of the formerly primary system now become secondary volumes and are only readable for this period.

power consumption

Thanks to the GRID architecture and the described I / O load balancing mechanisms, discs with high storage density and SATAII technology can be used even for very high performance requirements. In order to achieve a corresponding performance, traditional systems have to use FCDisks, which are not only much more expensive, but above all consume more energy because firstly they rotate faster and secondly more disks are required for the same capacity.

A fully expanded XIVRack with 180 disks and a gross capacity of 180 TB has an energy consumption of 7.7 kW. This corresponds to a consumption of 43 W per TB of data volume. Traditional storage systems with 146GB FC disks have a consumption of 180-380 W per TB.



Energy consumption of traditional systems with FC disks vs. XIV

Other features of the XIV architecture also help to reduce energy consumption:

Physical capacity is only "consumed" when the application is actually writing data, but not when a volume is created. Thin provisioning thus allows over-provisioning to be carried out, which means up to 30% less unused capacity.

By evenly distributing the data of a volume across all disks in the system, there are no unused “remnants” that often arise when different RAID arrays are configured for different purposes or volumes have to be distributed across different arrays.

management

The simple management of XIV systems is their great strength. Due to the described dynamic and autonomous distribution of the data to all volumes, planning for this is completely eliminated. With traditional systems, extensive planning and preparatory work is often required to ensure optimal performance and availability for the various applications. Furthermore, the settings would have to be adjusted when the system changes, which in turn is labor-intensive. All of these things are completely eliminated with XIV systems. The same applies to the placement of snapshots. No special precautions need to be taken here either.

User interfaces

There are three ways to manage XIV systems:

A graphical user interface (GUI). This can be operated intuitively, the complete administration of a XIV system can be learned in a few hours. Experienced storage administrators often get along without any preparation. A demo for graphical user interfaces can be found on the XIV website: http://www.xivstorage.com/demo/GUI_Demo.html

A command line interface (CLI). All administrative tasks can be entered here from the command line or automated with scripts.

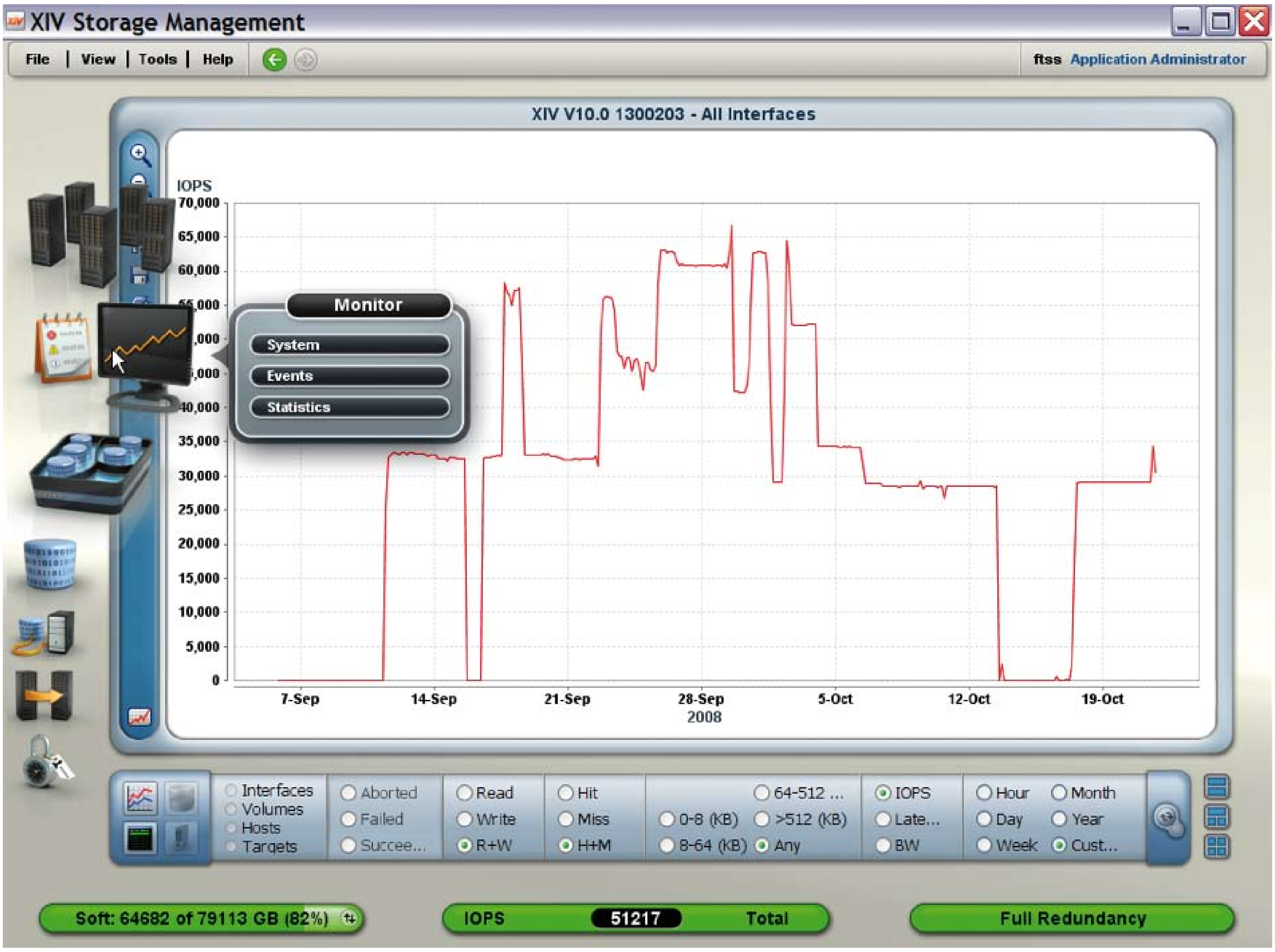
SMIS-compliant storage management tools: Storage tools, such as the Tivoli Productivity Center, can be used to administer XIV systems.

Statistical data

In addition to the regular administration tasks, statistical data can also be called up via the user interfaces mentioned. The system saves the performance data of a year. For example, I / Os per second or throughput numbers in MB / s can be called up from any time window within the last 12 months and examined more closely. When outputting z. B. be filtered on:

- ***Interfaces***
- ***Hosts***
- ***Volumes***
- ***Targets***
- ***Cache hit / miss rates***
- ***Read, Write or Read / Write***

and much more. Of course, this data can also be exported to other tools for further processing via the CLI.



View statistical data via the GUI

Interoperability

XIV systems can be operated on all common open systems that support iSCSI and / or FibreChannel storage systems. These include e.g. B. AIX, Solaris, HPUX, Linux and the various Windows variants. Common cluster solutions also support the XIV support systems.

Release XIV storage system software version 10.2

In November 2009, IBM announced release 10.2 for XIV, which introduced the following enhancements:

Asynchronous mirroring of XIV volumes and consistency groups and support for the Thin Provisioning Space Reclamation feature of the Veritas Storage Foundation.

DS3000 panel entry system

in the **February 2007** IBM surprised the market with the announcement of an entry plate system, the **DS3000.** Like the DS4000 series, the system is built by LSI and sold as an IBM logo product. In the beginning, the system was only available through the IBM System x sales channel. When the first performance benchmarks showed the respectable performance class in which the machine is scaled, the system was integrated in all of IBM's storage sales in April 2007.

The DS3000 family is generally characterized by an excellent price-performance ratio and offers an extremely affordable entry-level price. There are three models available.

The DS3200 is characterized by 6 SAS host connections. The DS3300 is characterized by a maximum of 4 x 1 Gbit Ethernet iSCSI host connections. The DS3400 is characterized by a maximum of 4 FibreChannelPorts, which can work at speeds of 4, 2 and 1 Gbps. The plate and backend technology is SAS in all three models with the possibility of connecting SATA plates. Initially, only 300GBSAS (Serial Attached SCSI) disks were available with a maximum expandable capacity of up to 14.4 TB. 750GBSATA disks have also been integrated since October 2007. The systems are scaled to a maximum capacity of 35.2 TB. The mixed operation of SAS plates and SATA plates was ensured.

In general, the models are characterized by a low height of only 2U and a scalability of up to 48 hard drives. Up to 12 hard disks can be accommodated in the head unit. The controller and thus the entire system can be expanded using the EXP3000 expansion units (up to 3 additional units).

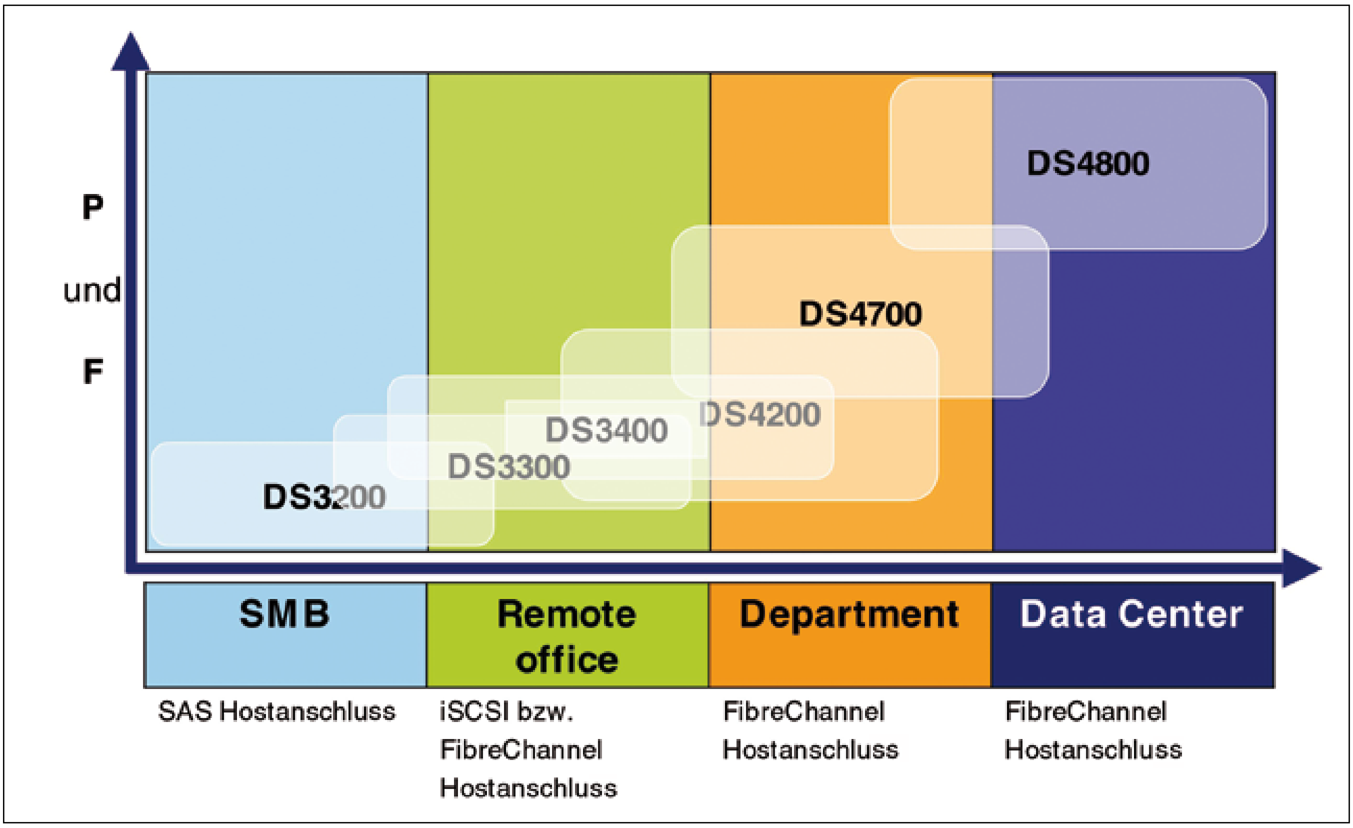
On the operating system side, the DS3000 family has a smaller portfolio of connectable operating systems. The Windows, Linux and Novell systems are offered as standard here. Depending on the model, the machine also offers the option of connecting VMware, SVC (San Volume Controller) or AIX.

Rear and front view of the DS3000

Each model has an Ethernet access for management. In addition, the machine can also be administered via host connectivity. With 512 MB RAM per controller, there is sufficient memory available for caching. This memory can be expanded with an additional 512 MB per controller. This results in a maximum cache size of 2 GB per system.

Despite the "entry system", the DS3000 does not do without redundant components. The DS3000 also has redundant fans and power supplies in a single controller variant.

Like the DS4000 Manager, the DS3000 Storage Manager is also intuitive to use and easy to install. The initial setup of the machine takes place in 6 dialogue-guided steps and should be completed within 45 minutes, ie after 45 minutes at the latest



Positioning of the entry systems DS3000 to the existing DS4000 systems (P = performance, F = functionality)

first volumes announced to the connected servers. The DS3000 Storage Manager supports RAID levels 0, 1, 3, 5 and 10.

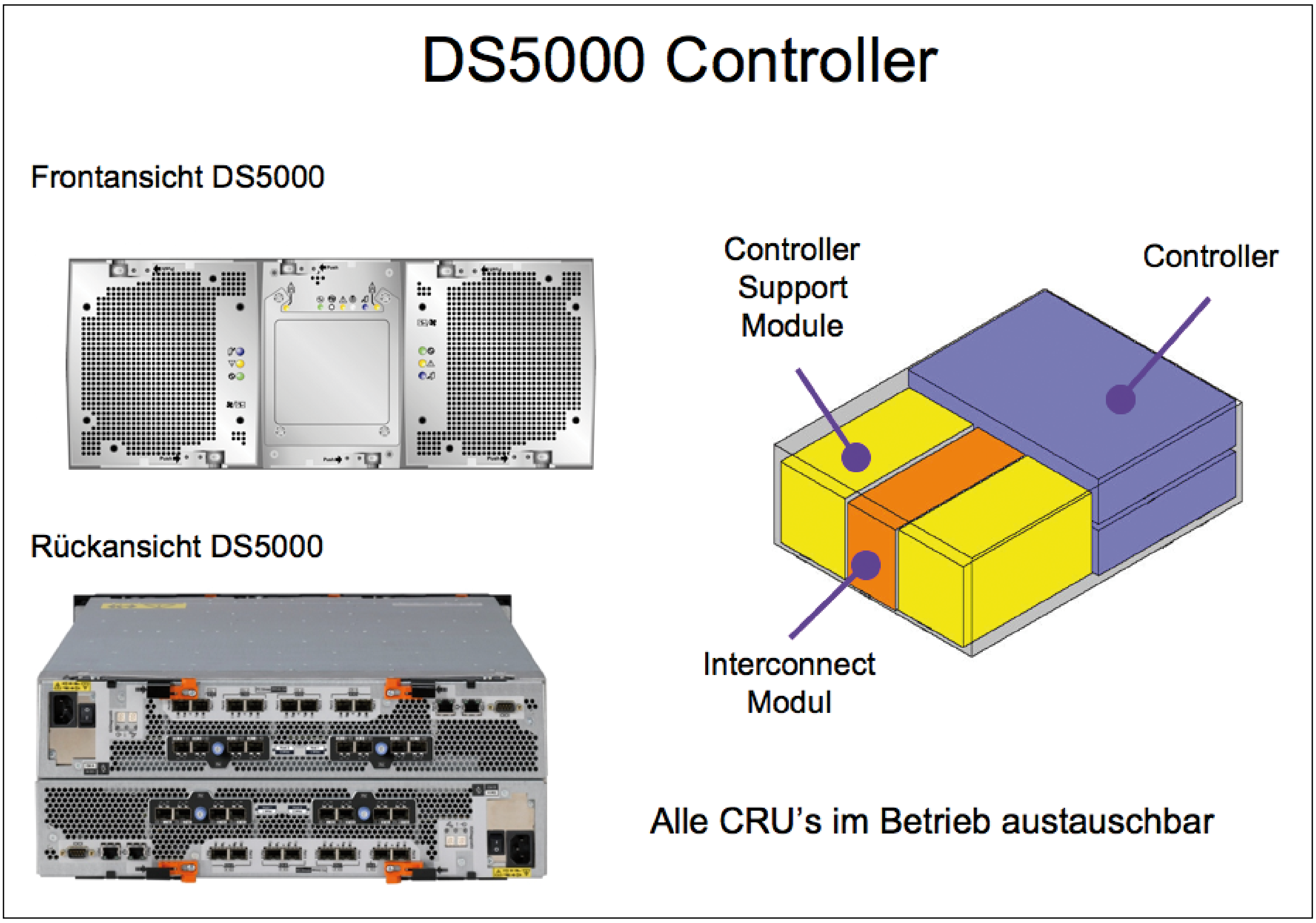
The recovery guru and automatic notifications proactively receive notifications of possible errors or exceeding of threshold values, even before a crash situation. The notification takes place via SNMP and / or by email.

in the **February 2008** IBM announced for all three DS3000 models 1,000 GB (**1 TB) SATA disks** at. This makes the EntrySystem DS3000 the first disk system from IBM to use disks with a capacity of 1 TB. With these disks, each expansion unit can be configured with 12 TB disk space. The entire system, consisting of a head unit and three expansion units, delivers one with the 1 TB SATA disks **maximum capacity of 48 TB** out.

A new model 32T is available for the DS3300, which has integrated a 48 V-based DC power supply in the controller unit especially for Telco environments (iSCSI DCPowered Controller). This means that the DS330032T can be used in telco environments without any problems.

DS5000 disk systems

The **August 2008** was all about announcing the new midrange **' Flagship ' DS5000,** that should round off the product portfolio. Taking over the tried-and-tested hardware design with the interchangeable midplane of the DS4800, the system has been equipped with several other functions. In the DS5000 family (DS5100 and DS5300), which consists of two products, the host interface cards (HIC for short) can be modularly exchanged. The user has the option of connecting via 1 GBit iSCSI, optionally also via 4 Gbit or 8 Gbit FC.



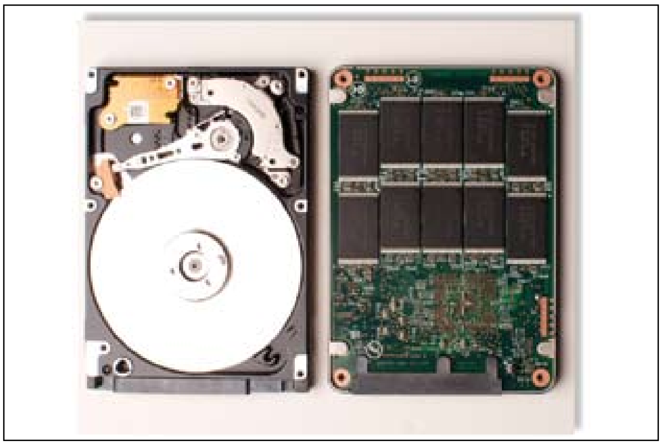
The capacitive scalability has also been increased from 224 disks on the DS4800 to 448 disks. As a result, the number of DiskBackend channels has been doubled to sixteen 4GbitFC channels.

When designing the storage controller, attention was once again paid to linear scalability right down to the last disk. The RAID calculation is carried out via a separate ASIC. These two features ensure excellent transaction and data throughput performance. Both are impressively demonstrated by the SPC1 and SPC2 tests (www.storageperformance.org).

An innovation compared to all other midrange products in the IBM portfolio is a change in the cache protection. While in all previous models the cache was backed up by a backup battery for up to 72 hours in the event of a power failure, the cache content in the DS5000 is maintained with the help of FlashMemory modules receive. In this case, the controller itself is operated by a battery in the event of a power failure until the cache content has been moved to the permanent memory. Then the DS5000 will shutdown properly.

Dedicated channels are now available for cache mirroring on the DS5000 in order to be able to use the entire internal bandwidth of the system without restrictions.

Functionally, the DS5000 has been expanded by RAID6 compared to the predecessor DS4800. The RAID process used is P + Q RAID6, an industry standard that was developed together with Intel.



Conventional plate SSD drive

in the **Year 2009** The DS5000 family was then expanded step by step:

- *Full Disk Encryption Drive (FDE) : This is to an FC disk that is able to encrypt data on it using a key on the disks. Encryption is activated using another key held on the storage system. This ensures that the data cannot be read when an FDE disk is removed and that the data on it is therefore protected against unauthorized access. A function that increases data security, especially when storing personal or customer-related data.*

- *System I connection : System I was also able to over a VIOS can be connected, but with the announcement in autumn 2009 there was the possibility to connect System I native to a DS5000. This is a feature that many System I users have been waiting for. This feature in combination with the simultaneous possibility of supporting up to 64 GB cache ensures the response and service time that is so important for System I applications.*

- *Solid State Disk Drives (SSD) : Announced in autumn 2009 added the possibility to operate up to 20 SSD drives in a DS5000. In contrast to a conventional disk (figure), the SSD has no rotating or moving components. This eliminates the rotational latency, so that a significant improvement can be achieved in the area of transactions. The log areas of a database can be viewed here as a typical area of application. At the time of the announcement, the SSD was available in the size of 73 GB.*

- *High Density Enclosure EXP5060 : In addition to the 16 plates expansion unit EXP5000 and EXP810 with the announcement of the EXP5060 a 60 SATA 4U expansion unit is available. Spread over 5 drawers, this offers a high capacity (60 1TB SATA disks) and very little space. To ensure the required performance, it is possible for the EXP5060 to use trunking from several FCBackend channels of the DS5000 for the first time. The EXP5060 is primarily intended to address the HPC (high performance computing) and archiving market.*

Announcement DS5020 - the end of the DS4000 era

Mid 2009 The task was to find a successor to the DS4700. At the announcement date of August 25, 2009, the **DS5020** presented.

Like its predecessor, the DS5020, the DS5020 can be expanded up to 112 disks, which are connected via several 4Gbit channels. Standard FibreChannel, FDE FibreChannel and SATA plates are available as disc types. All plate types can be operated side by side in the same expansion unit.

In order to meet the requirements of virtualized environments in particular, special attention was paid to balanced performance during controller development. On the one hand, there is excellent transaction performance that makes maximum expansion seem sensible, and on the other hand, the system delivers very impressive throughput results. Both values were documented in corresponding SPC1 and SPC2 tests. The data cache secured with flash memory can be expanded up to 4 GB.

Another special feature is the possibility to use different protocols in the front end, i.e. after the server. Here are 4/8 Gbit FibreChannel and also 1 Gbit iSCSI available. Both protocols can be operated in parallel on one controller and enable a corresponding 'SAN tiering'.

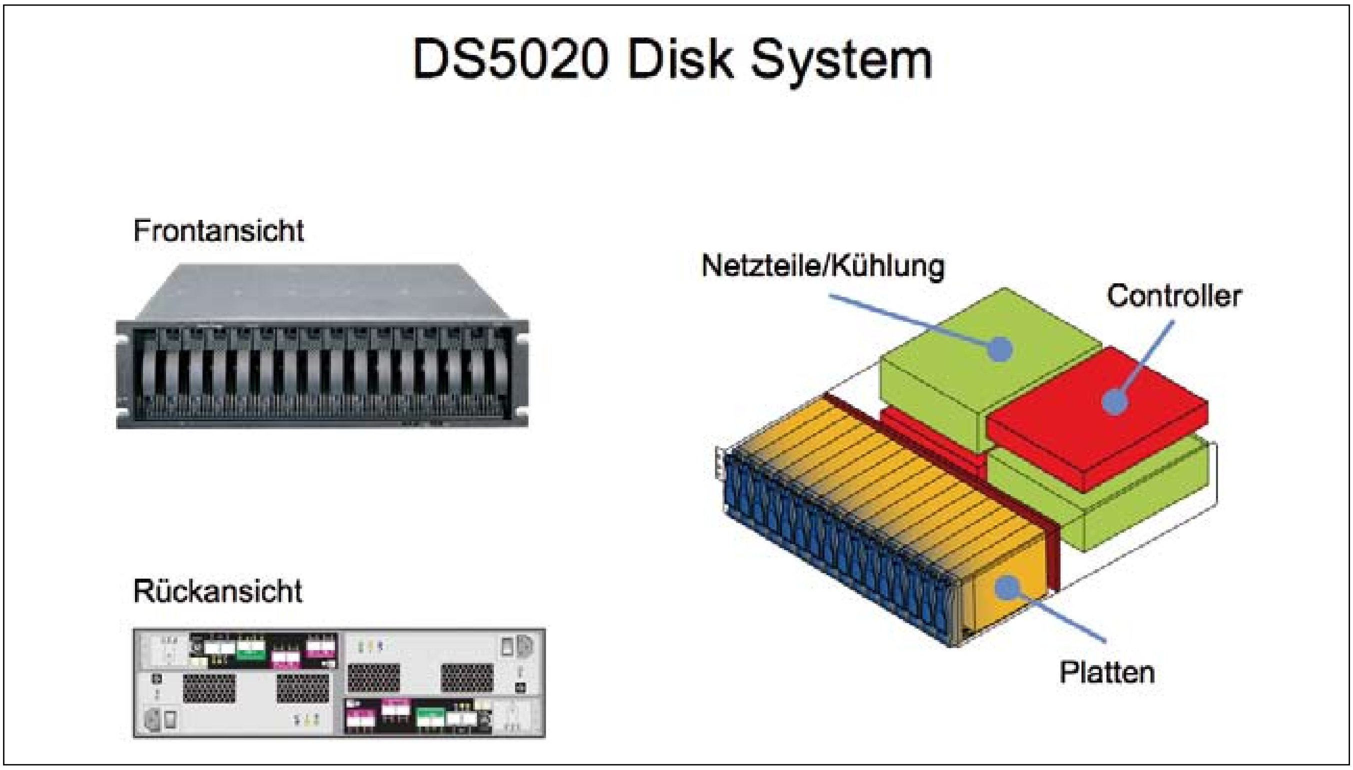
Like all members of the DS4000 / DS5000 family, the DS5020 supports a variety of additional functions such as:

- ***Storage partitions***
- ***FlashCopy***
- ***VolumeCopy***
- ***RemoteVolume Mirroring.***
- ***Disk Encrypton etc.***

Data from existing DS4000 systems can be transferred directly to new DS5000 systems, e.g. B. by Remote VolumeMirro ring or in some constellations also by physical migration of the disk groups.

After the announcement of the DS5020, the DS4700 will remain available until the end of 2009. Then the chapter of the DS4000 storage systems is closed.

Both announcements (DS5000 and DS5020) accompanied a large number of solutions to ensure that the integration of the applications and their operating system platforms was so important for operations.

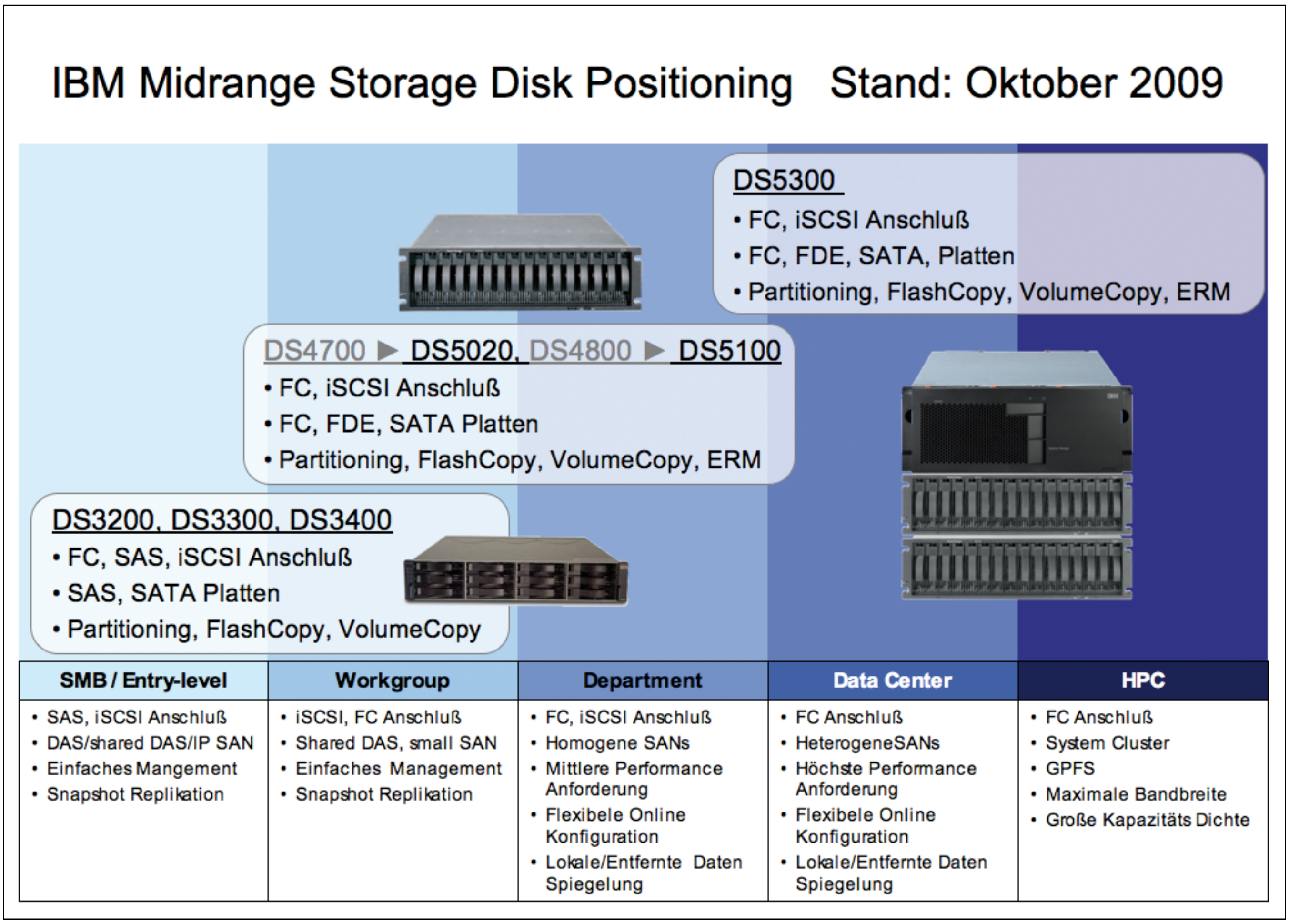


The era of server-based storage systems and storage virtualization

Here are some examples:

- ***Integration in VMware vCenter via a plug-in***
- ***VSS / VDS support***
- ***SMI-S integration***
- ***Oracle Enterprise Manager PlugIn***
- ***Solution offering for de-duplication solutions (ProtecTier)***
- ***TSM FlashCopy Manager***
- ***Backup integration such as B. with FastBack***
- ***Integration into the System Storage Productivity Center (SSPC)***

Complete integration into the IBM virtualization solution SVC (SAN Volume Controller) is also available for the entire system family.



Virtualization

SAN Volume Controller SVC - Virtualization of disk systems in the SAN environment

The first devices of the SAN Volume Controller SVC were available in September 2003 and are described in the era of multiplatform systems and the FibreChannel SAN and NAS. The design of the SVC and the associated software has been continuously developed over the years and today's product plan of the SVC provides for many new hardware and software enhancements over the next few years. It is therefore necessary to better understand the concept and the SVC itself.

The administration of storage capacities continues to represent a high proportion of IT costs. Virtualized storage landscapes offer a possible solution in that they can manage storage systems of different types and manufacturers together and flexibly assign them to the servers. Storage virtualization with the SAN Volume Controller meets these challenges with a scalable architecture. In recent years, many companies have put a lot of effort into developing from a decentralized storage landscape with direct attached storage (DAS) to a central storage solution with a storage area network (SAN). Storage area networks do not exist across the board, but in many cases they already form the standard. In the current phase, it is a matter of largely virtualizing both server environments and storage systems. The currently existing SAN solutions are often quite rigid environments and mostly form individual solutions for platforms from certain manufacturers and specialist areas. As a result, changes are complex and usually interrupt the application operation. If an IT operator has installed different storage systems - possibly from several manufacturers - he needs different administration tools. The hardware-based data mirroring between storage systems of different designs is also not possible, the same applies to storage systems from different manufacturers. The currently existing SAN solutions are often quite rigid environments and mostly form individual solutions for platforms from certain manufacturers and specialist areas. As a result, changes are complex and usually interrupt the application operation. If an IT operator has installed different storage systems - possibly from several manufacturers - he needs different administration tools. The hardware-based data mirroring between storage systems of different designs is also not possible, the same applies to storage systems from different manufacturers. The currently existing SAN solutions are often quite rigid environments and mostly form individual solutions for platforms from certain manufacturers and specialist areas. As a result, changes are complex and usually interrupt the application operation. If an IT operator has installed different storage systems - possibly from several manufacturers - he needs different administration tools. The hardware-based data mirroring between storage systems of different designs is also not possible, the same applies to storage systems from different manufacturers. If an IT operator has installed different storage systems - possibly from several manufacturers - he needs different administration tools. The hardware-based data mirroring between storage systems of different designs is also not possible, the same applies to storage systems from different manufacturers. If an IT operator has installed different storage systems - possibly from several manuf

Administration work is very time-consuming because platform knowledge and the associated maintenance have to be kept equally varied. Today's storage continues to be burdensome made poor use of their installed storage capacities. Worldwide customer surveys have shown that effective use is around 50%. A lot of effort was also required to replace storage systems as part of business continuity measures, and the application operation was mostly interrupted for the entire migration period. The IBM SAN Volume Controller (SVC) makes this much easier by: **Migrations during operation** enables.

Storage virtualization with IBM SVC offers the advantage over traditional storage solutions that **Storage space allocated largely independent of the manufacturer** can be. This results in easier administration, since the virtual user interface can be configured independently of the storage systems installed. The basic idea is to be able to allocate disk storage quickly and flexibly wherever there is a need. So far, this has stood in the way of servers connected to the SAN only having access to the memory areas assigned to them. A virtualized storage solution like the SVC resolves this rigid pattern, since the servers here only have access to virtual storage areas. The SVCSoftware decides according to the corresponding specifications where the data is physically stored. The solution offers interfaces to all common operating systems, hardware platforms and storage systems and enables management via a central console.

Some important hardware functions remain in the storage system and are independent of the SVC. The systems are set up before connecting to the SVC as in the past. RAID arrays and then the LUNs (Logical Units) are formed. This takes place in the course of the installation using the administration software supplied by the manufacturer. The LUNs defined in this way are then assigned to the SVC, which converts them into so-called 'Managed Disks' (MD). An MD has a fixed size and can no longer be changed. Then one or more MDs are grouped into so-called 'Managed Disk Groups'. For the connected servers, virtual LUNs are defined in the SVC, which the respective server accesses. The definition also defines which managed disk group the virtual LUN is assigned to. This determines

Easy data migration

The SVC carries out data migration of LUNs between different storage systems during ongoing operations. If an application whose data is stored on a high-performance memory, for example, loses priority for the company - for example due to high operating costs for the specialist department - its data can be migrated online to cheaper low-performance memory. This is done by changing the assignment of a virtual LUN to the Managed Disk Group.

In addition to data migration for storage system exchange, maintenance work can now also take place during working hours. Memory contents of the subsystems can be moved to other, free areas of another subsystem during operation. All data transfers take place under the control of the storage administrator and must be initiated by him. The SVC does not offer rule-based data shifts due to access frequencies. This can be done by analysis tools that are embedded in the IBM Total Storage Productivity Center. It offers complete SAN management including storage space and access analysis and can trigger rule-based copying processes in a hierarchical storage system (HSM). All IBM storage products, as well as tape libraries, can be integrated.

Overall, the SVC provides one **complete appliance solution** that includes hardware, software and management console. Each of the two IBM System xServers belonging to a so-called node pair is with a 2-processor system **with 2 x 2.4 GHz Intel processors** has 8 GB cache and four 4Gbit / sFibreChannelPorts (2145 8G4). These new processors have been available since May 2007 and increase the performance from 160,000 I / Os to up to 276,000 I / Os. All administrations can be carried out via the management console and the necessary analysis data can be read out in the event of an error. Included are also two components for uninterruptible, battery-backed power supply (UPS).

The SVC is implemented like other devices that are newly integrated into the SAN. Adjustments in SAN zoning and the assignment of data paths should not mean anything new to an experienced SAN administrator. For data acquisition in the virtual level, the administrator first executes an image mode with each individual LUN. This means that the previous LUNs of a server are assigned to the SVC and from there passed on to the respective server as a virtual LUN. In the next step, these LUNs can be assigned to another 'Managed Disk Group'. As a result, the data is shifted to the MDs assigned to this group. This is done during operation and is transparent to the respective server.

SVC redundancy

The redundant structure of the SVC ensures the highest possible availability of the storage environment. The Linux cluster solution consists of at least two IBM System x servers. All critical elements are designed in duplicate, which largely minimizes the risk of system failures. A Linux kernel runs in each of the servers belonging to the cluster, which was adapted by IBM to virtualization needs and cannot be modified externally. Necessary changes are treated like a firmware update that the customer or a service provider carries out. The virtualization software runs like this optional copy service routine under the control of this Linux kernel.

SAN Volume Controller SVC: 3 nodes in the cluster

A highlight of the SAN Volume Controller is the **high performance.** The SVC meets all scaling requirements by expanding a cluster to up to eight nodes today. If a cluster with a node pair encounters performance or capacity limits, an additional node pair can be added during operation (up to a maximum of four node pairs per cluster). Each node of the IBM SVC has an 8 GB main memory that is used almost entirely as a cache. At **up to eight nodes per cluster** this results in a **Cache size up to 64 GB.** The use of the cache in the IBM SVC brings performance advantages over traditional storage solutions, since all memory access runs against the cache of the SVC. As a result, the performance of the connected storage systems loses importance because a large part of the I / Os are served by the SVC's upstream cache.

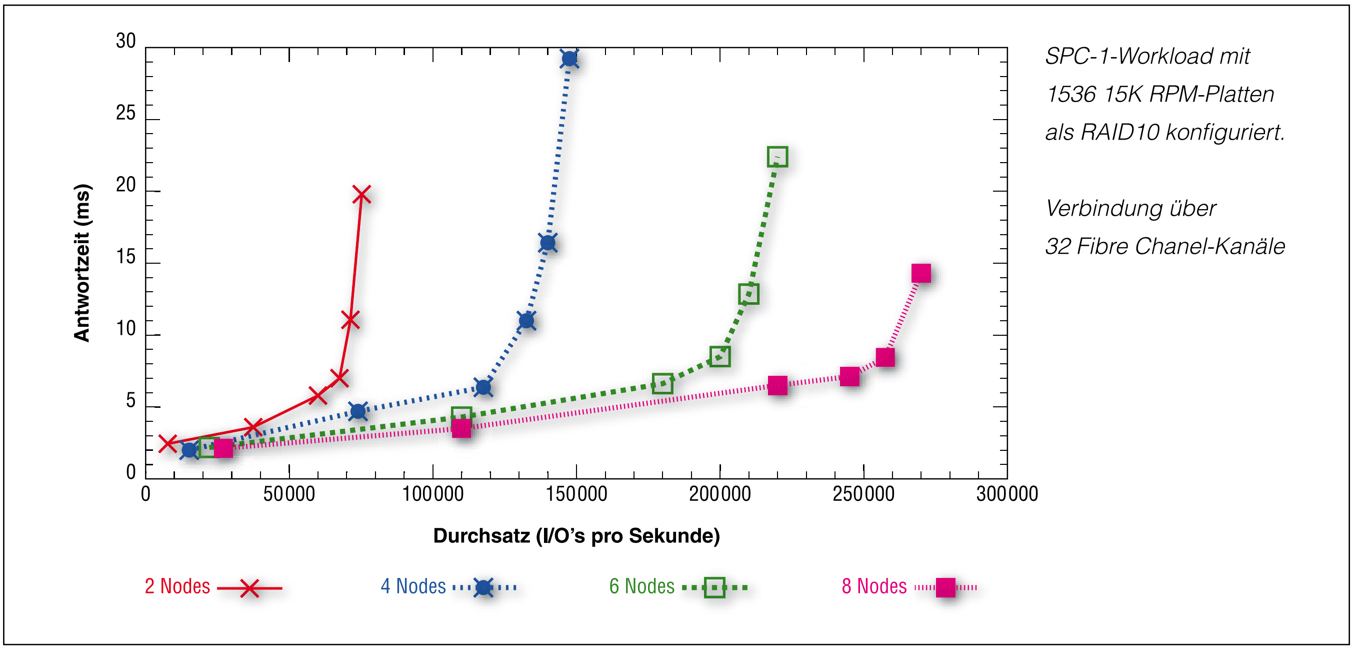
The IBM SAN Volume Controller thus fulfills all performance requirements of marketable systems. This is underlined by the SPC benchmark test (Storage Performance Council) of the manufacturer's association of the same name, which all providers of storage systems can carry out voluntarily.

SVC cross-location reflections

Hardware-based data mirroring (synchronous or asynchronous) as a typical function for fulfilling business continuity requirements is normally not possible across manufacturers. The SVC performs data mirroring in the virtualization layer. In a virtualized storage environment, data mirroring is now also possible between systems of different designs and manufacturers. In the event of a disaster, you can quickly switch from primary to secondary storage. This usually requires intervention by the administrator, which can be supported or automated by previously created scripts.

SVC mirroring options

Overall, at the SVC, all mirroring activities between distributed locations mostly take place via fiber optics. For long distances, SAN routers can be used, which convert the FC protocol into an IP protocol on one side and back into an FC protocol on the other side. Alternatively, some manufacturers also offer DWDM (Dense Wave Division Multiplexing) or CWDM technology (Coarse Wave Division Multiplexing). With these devices it is possible to bundle various protocols (e.g. IP, FC, FICON, ESCON) to the minimum number of physical connections (usually dark fiber).



This graphic impressively describes the almost linear scaling of an SVC cluster from 2 nodes to 4, 6 to 8 nodes. In these tests, 1536 15K RPM drives in RAID10 were required to saturate an 8 node SVC. So that keeps SVC the record in the SPC-1 benchmark. The benchmark results can be found at www.storageperformance.org be retrieved.

SVC copy functions (FlashCopies)

FlashCopies are suitable for backup support of databases and file systems. A FlashCopy is executed as a PointinTime Copy and, when activated, creates a hard drive that is an exact copy of the source hard drive at the start.

With the software release **SVC version 4.2,** that in **May 2007** was made available, the new function came
Multi Target FlashCopy which allows 16 copies of a hard disk to be created, ie to map 16 targets from a source.

With the **Software release SVC version 4.2.1 from October 2007** IBM announced the function **of the Incremental FlashCopy** and the **Cascaded FlashCopy** at. Incremental means that an existing FlashCopy can be updated to the current status. Then only the changes to the original hard disk that have been made up to that point are copied. Cascaded means that you can also make additional copies of the copy. This is especially important if you want to save an incremental FlashCopy before continuing. In total, up to 16 copies of an original hard disk can be created. Consistency groups are supported in all modes.

Version SVC Version 4.2.1 also offers the option of CopyService capacities of 256 TB per I / O group
1,024 TB, and now opens up the possibility of a maximum capacity of **8 PetaByte (** previously it was only 2 PB) to map using SVCCluster.

in the **May 2008** IBM announces the **SVC Release 4.3** with availability in June 2008. Release 4.3 contains many enhancements and new functions, which are described below.

Space Efficient Virtual Disk (SEV)

Space Efficient Vdisk is nothing more than ”Thin Provisioning” or ”Storage Overallocation”.

Users can create virtual disks that have different real and virtual capacities. Real capacity defines how much physical capacity is actually used by a Vdisk. The virtual capacity defines how large the virtual disk appears to the connected servers.

An SEV differs only minimally from normal discs and can be enlarged or reduced like a normal disc. Like other NON SEV discs, it is created from any pool. The conversion between SEV and NON SEV is done via Vdisk Mirroring. SEVs help to save space from allocated but not used capacity. The saving potential depends on the operating system environment and the application (in Windows environments often over 50%). SEVs also improve the degree of utilization of the physical disks (in addition to the potential deriving from the pooling concept of the disk and spare capacity). SEV is the basis for Space Efficient FlashCopy.

Space Efficient FlashCopy (also known as SnapShot)

Space Efficient FlashCopy works like Standard Flash Copy with the difference that the FlashCopy Target is a Space Efficient Vdisk. This means that only as much space is allocated for the FlashCopy as is currently required.

Virtual disk mirroring

Vdisk Mirroring can be thought of as an internal LVM (Logical Volume Manager) in an AIX environment. Vdisk Mirroring represents a new Vdisk configuration, in which a 'mirrored Vdisk' is written simultaneously on two different MdiskSets, which are normally on two different disk systems. The SVC keeps both copies in sync. In the event of failure of a spoke chersystems has no influence on the availability of a mirrored Vdisk. The mirror is automatically resynchronized after availability. Mirror copies can be split and the discs can be in disk groups with different extent sizes. One or both copies can be Space Efficient. Vdisk Mirroring can be used together with SVC Copy Services.

SVC clusters are usually set up in a data center. With Vdisk Mirroring it is possible to build a 'Stretched SVC Cluster' (RPQ required) by distributing the node pairs to two data centers. The SVC node pairs are connected via Longwave FC in the switch. Distances of up to 10 km are possible.

FlashCopy Target Volumes increased to 256

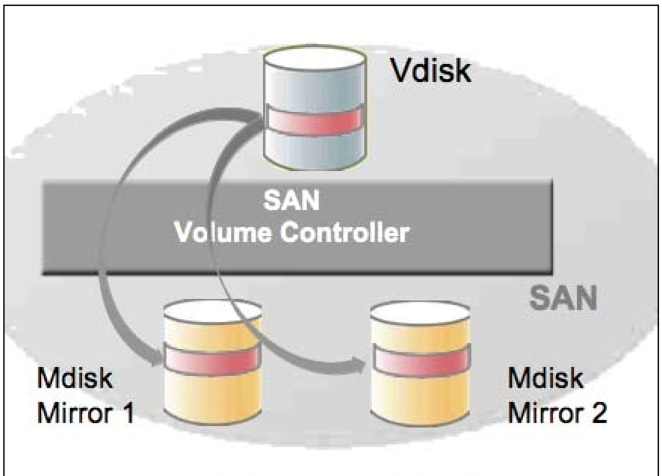
There are only two limits to the capacity of FlashCopy: Up to 256 target volumes can be created from a source volume and the maximum capacity of all mirror actions must not exceed 1 PB. There are no longer any restrictions regarding the mixing of copy actions.

Increased scalability of virtual disks

The scalability in the number of virtual disks has been doubled and increased to 8,192 virtual disks.

SVC Entry Edition

At the same time as Release 4.3, IBM is also making a cost-effective entry-level solution, the SVC Entry Edition, available. These devices are based on an x3250 server with an Intel Xeon E3110 3.0 GHz 6 MB L2 cache dual core processor. Like the current model 8G4, the entry devices have 8 GB cache and four 4 Gbps FibreChannel connections. The SVC Entry Edition is about 40% cheaper and achieves about 60% of the throughput and performance of the 8G4 model.



IBM SVC Virtual Disk Mirroring

New high-performance SVCs with Solid State Disks (SSDs) with SVC Release 5.1

in the **October 2009** IBM announces the availability of November 2009 **SVC Release 5.1** which, in addition to functional expansions, includes new, more powerful processor hardware. The new SVC Model 2145 CF8 is based on the IBM System x3550 M2Server with Intel Core i7 quadcore

2.4 GHz processors and performs almost twice as the predecessor 2145 8G4. The new machines are equipped with the triple cache size of 24 GB and an I / O group offers 48 GB as cache memory. This means that four pairs of nodes in the cluster network result in a cache size of 192 GB. Each SVC node also offers 8 x 8 Gbps FC ports. This means that in the maximum configuration with four pairs of nodes in the cluster, 32 x 8 Gbps FCPorts are available. Due to the faster processors and the tripling of the cache, the SVC gains an unusually high performance. 120,000–140,000 IOPS are possible per I / O group, i.e. per node, i.e. 480,000–540000 IOPS when fully expanded.

The new SVC engines can with **built-in SSDs** (Solid State Disks). Up to four SSDs with a capacity of 146 GB can be built into one engine, i.e. up to eight SSDs per I / O group and thus an SSD gross capacity of 1,168 GB per I / O group. Only 600 GB are used because the majority of the SSDs are mirrored. With a maximum configuration with four pairs of nodes in the cluster, up to 2.4 TB of usable SSD capacity is available.

iSCSI support

The new SVC devices can also be equipped with built-in ethernet ports and thus directly connected to iSCSI. The existing LAN ports of the SVC are used for this.

If one of the SVC's nodes fails, the corresponding IP address is transferred to the other nodes. It is therefore not necessary to install MPIO software for iSCSI on the host.

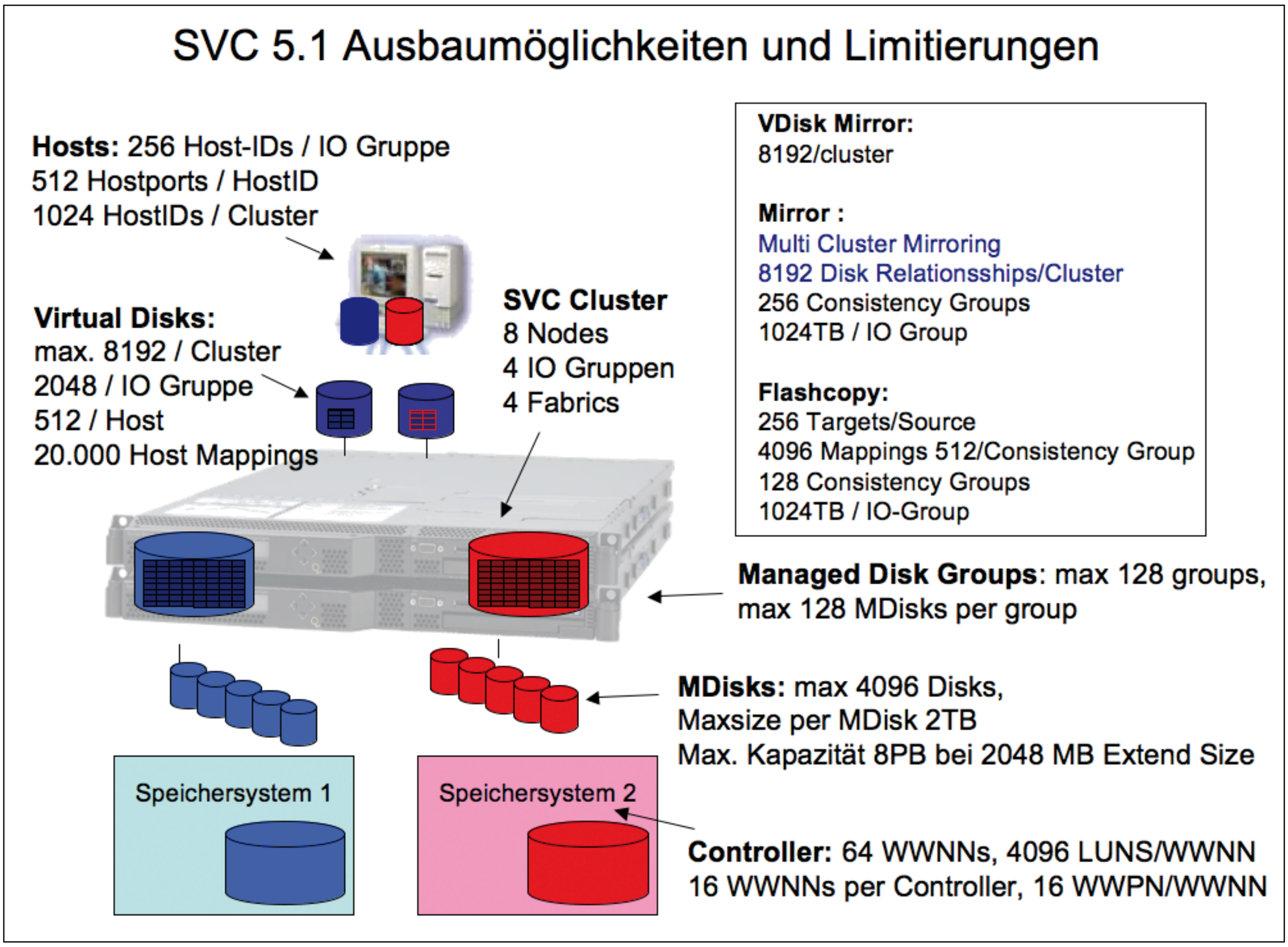
New function

With **Release 5.1** new functions are available to the SVC. **Reverse FlashCopy** allows FlashCopy targets to be used as a restore point for the source without breaking the FlashCopy relationship. Multiple targets and multiple 'roll back points' are supported.

Multiple cluster mirroring allows you to build a triple reflection. Four SVC clusters form a mirror group. Each disk can be mirrored from any cluster to another cluster.

Installation basis December 2009

By December 2009, over 17,000 SVC nodes had been delivered worldwide. These nodes are in productive use in more than 5,000 SVC systems.



Backup concepts

Regardless of the implementation, the importance of backing up business-critical data is undisputed. In addition to the classic three business factors of people, land and capital, another factor has been imposed in recent years. The information factor. Information is the lifeline of every company and IT is there to secure and support the existence and further development of this lifeline.

A backup is a copy of data in a local or remote location depending on the implementation. Such a data copy protects against total loss in the event of human errors, hardware errors, virus infections and total loss of a location. A well-thought-out backup concept is personal insurance for the continued ability to work and, in extreme cases, for the continued existence of an entire company.

Plate technology vs. Tape technology in the backup environment

One disk offers direct access and advantages for single file restore and file system restore. There is also the option of working with “multiple streams” on the disk in order to reduce the backup time window. But records also have disadvantages! They offer no advantage for large file backups or full image backups (here the tape is usually much faster). Plates cannot be moved, that is, they cannot be used as moving data carriers. They require electricity and cost more than tape solutions if the capacities are high. The fundamental question also arises, which disk types do you use for backup? Should it be ATA, SATA, FATA, SAS or FC disks? What is the right price-performance ratio in relation to reliability and security?

Tapes, on the other hand, are sequential 'streaming devices' and today, in terms of data rate, are usually faster than discs. That is why it is usually more efficient to save large files or large databases directly onto native tapes. Tapes are 'on demand', so they scale in terms of performance and capacity by adding either additional drives or additional cartridges to a tape library.

Tapes are exchangeable data carriers, which means that they can be outsourced to security rooms or exchanged between different locations. Tapes do not require electricity once they have been written on and are much more durable (LTO as ECMA standard up to 30 years). In addition to rewritable cartridges, WORM tapes (write once read many) are available in almost all tape technologies, which can be used particularly in the area of long-term archiving.

What are the alternatives for a secure backup today?

The classic and traditional backup begins with a tape drive connected locally to the server. If there are several servers, this becomes maintenance-intensive very quickly and therefore only scales to a limited extent. The next step would be a central, LAN-based data backup e.g. B. via TSM and a corresponding backup server, in which disk or tape systems are connected directly to the server. This means that multiple servers can use a common tape backup infrastructure (e.g. a tape library). With larger data volumes, e.g. For example, for DBs, a SAN-based backup is available that keeps the high data transfer volumes away from the local network (LAN) and thus realizes advantages in terms of speed and access.

Despite the immense further developments in the field of tapes, shorter and shorter access times with higher write throughputs and capacities, the data is still not permanently accessible (online).

This can only be achieved through a disk storage infrastructure. Against the background of the drop in prices for hard disk drives, especially for the inexpensive SATA disks (Serial ATA), there are many possibilities. One possibility is to completely change the backup concept from tape to disk. This ensures high availability and speed, but it is difficult or impossible to physically move the data to a remote location (mine, bunker, vault, bank). Depending on the sensitivity of the data, this is recommended or is partly part of legal requirements. In addition, there are very high energy costs and additional migration costs for a pure disco solution, which quickly nullify the advantage of the supposedly cheap hardware investment. To take advantage of both

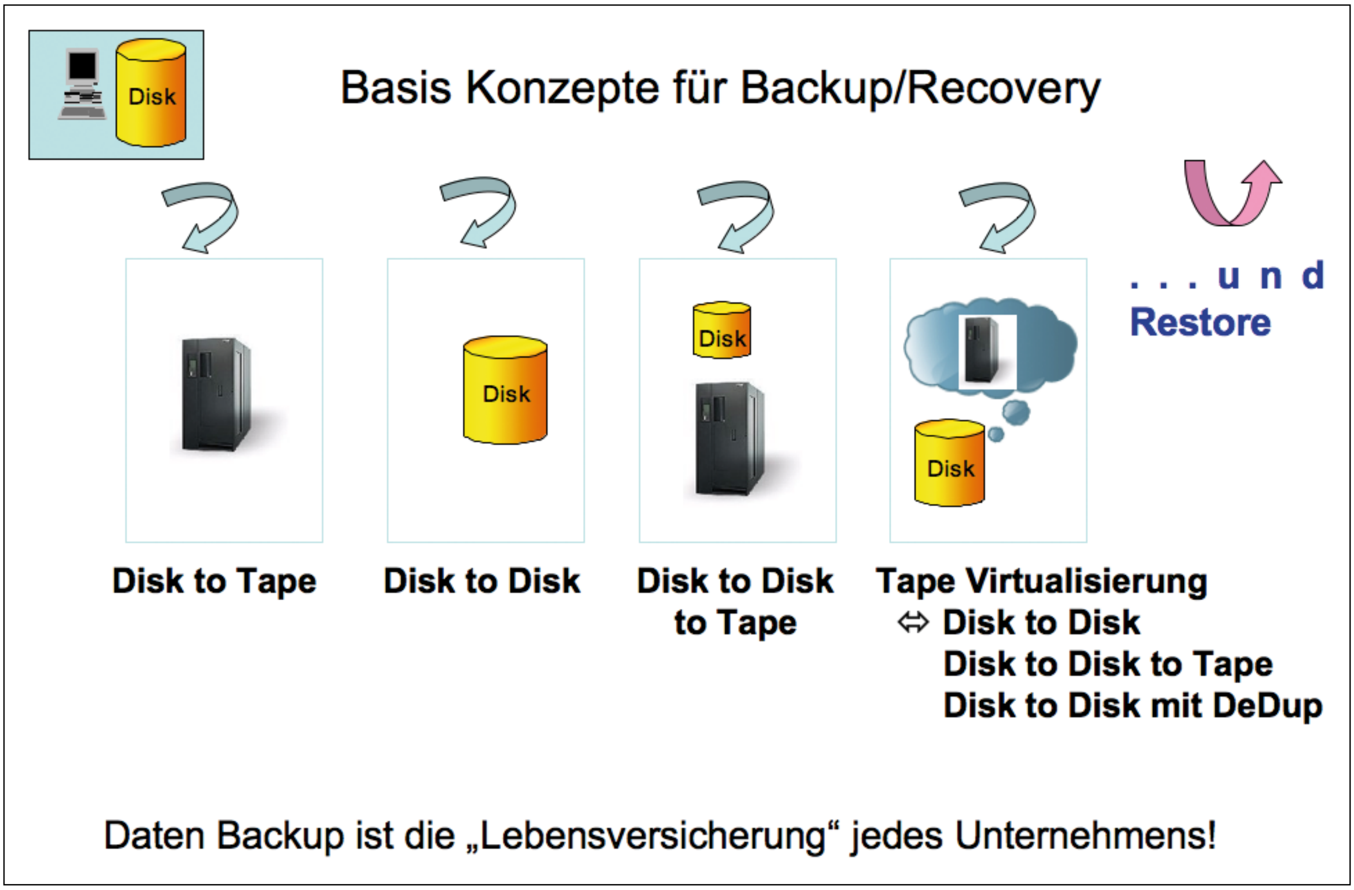
The era of server-based storage systems and storage virtualization

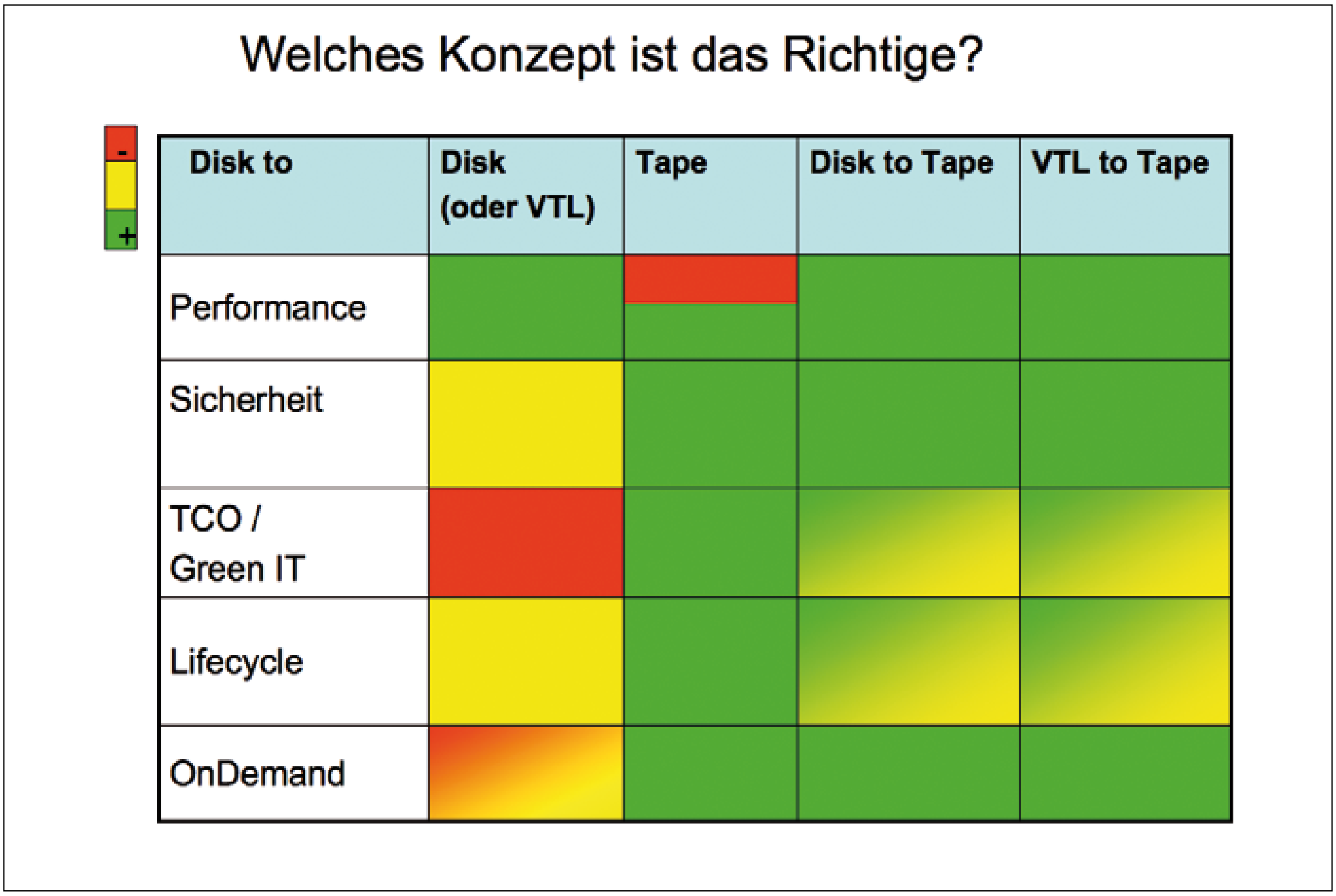
To be able to use technologies, the combination of disk and tape is becoming increasingly popular for fast, highly available and yet secure backup. This is known throughout the industry as DisktoDisktoTapeBackup. In this concept, the disks are used as backup buffers (staging area) in order to keep recent backups online for a certain time and then to transfer them to tape after a set period. This also applies to virtualization solutions from Tape: Here tape drives are emulated on disk, which means that disks appear like a tape drive. Integrated tape virtualization solutions combine hardware and software into an easy-to-use virtual tape library (VTL). Only the option of migrating to physical tapes offers an operational, cost-effective and efficient environment.

When working with a combination of disk and tape gear, the question arises as to how large the upstream disk system must be in terms of capacity, how the tape library behind it must be configured and how to implement the migration from disk to tape in a sensible way! Individual sizings of the existing environment and the required requirements are necessary for a sensible solution!

There are VTLs on the market that do not have a tape connection. Of course, this is not the question! However, such conceptions are strongly discouraged, as it leads to a technological one-way street, from which you can only get out with a lot of effort and costs. VTLs without tape connection with very cheap ATA or SATA disks usually have to be replaced completely after three years for security reasons to avoid disk failures and data loss. If very large capacities are affected here, there is usually a real explosion of costs!

When using VTLs and disk buffers, more effective use of the disk buffer can be saved by de-duplicating processes, which means that a large amount of disk space that would otherwise be required can be saved because only data blocks are saved once and duplicates are avoided. The saving is directly dependent on the de-duplication factor achieved. There are different de-duplication processes from different providers. The most important requirement for a de-duplication process is to ensure 100% data integrity in addition to scalable performance. Only procedures with 100% data integrity ensure a secure





Restore option available, because backup is made because of the restore!

The de-duplication factors achieved depend on many factors: which backup software is in use, how is the backup carried out (e.g. always incrementally or a full backup every week etc.) and how long should the retention period be, where the data in Disk buffers remain stored (see also under Data De-duplication).

Which concept offers the best effectiveness in terms of performance, security, costs, durability and easy expandability depends on the existing infrastructure and whether the data is backed up traditionally via LAN or LANfree via SAN or both. In most cases, combination solutions of disk and tape are available to use the strengths of both technologies for an effective backup and restore.

It is important that DataBackup is safe and reliable and that data recovery is guaranteed in any case. This applies to unforeseen errors such as B. Disaster (natural disasters, floods, technical errors, disk crash), but also for human errors (accidental deletion of data, incorrect editing of files) and application errors (virus, software errors). So it is

Among other things, it is important that data copies are created at different locations, on different media and with different tools or software and by different responsible persons. Snaps in the same unit alone do not offer any protection, neither do mirroring with the same software!

The demands on the level of security and the recovery time (recovery time objective) also play a decisive role in the selection of the components and the backup strategy.

Virtualization of tape systems

New server-based tape virtualization for System z

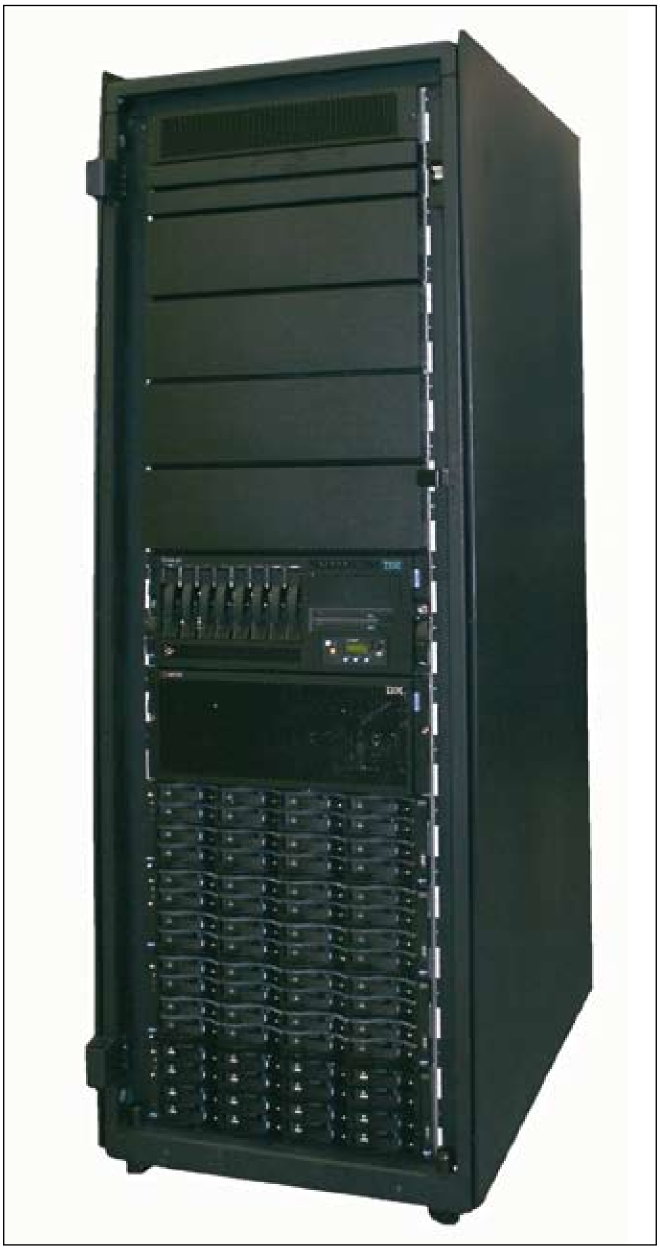
At the **August 29, 2006** IBM announced a new virtualization solution for the z / OS environment. The **TS7700** will be available as the successor to the IBM 3494 B10 and B20 Virtual Tape Server at the end of September 2006. This was particularly necessary because the old units B18, B10 and B20 did not comply with the RoHS EU directive, which came into force on July 1. The new TS7700 series for tape virtualization in the mainframe environment forms one **new, long-term basis** through a modular and scalable server architecture. One speaks of one **'Distributed Node Architecture'.**

The TS7700 consists of the 3592 F05 housing, the TS7740 distributed node architecture with different nodes (see below) that perform the corresponding functions, including two active 'I / ODrawer', the corresponding FICON connections to the host, Fiber Channel connections to the plates and Ethernet connections to the Library Manager, the TS7740 cache controller as a RAID control unit and corresponding RAID-based disk modules that reflect the TVC (tape volume cache) as a disk buffer.

RAID arrays with 146 GB disks are used for the disk buffer capacities. The RAID controller works with 16 integrated disks, a TS7740 cache extension also contains 16 disks. In addition to the RAID controller, up to three disk expansion units can be integrated into the housing unit. This means that up to 64 plates can be installed in the basic unit. The total capacity is 6 TB (native) and 18 TB (compressed).

Nodes are functional tasks which run as an application on a pSeriesServer consisting of two DualCore 64 bit, 1.9 GHz processors. The vNode (v stands for Virtual) provides the host with virtual drives and is responsible for the host communication. The hNode (h stands for Hierarchical Storage Management) is responsible for the administration of the logical volumes in the disk buffer. It also manages the migration from disk to tape and is responsible for the entire management of the physical tape resources, including the tape recovery process. The combination of vNode and hNode is called gNode (g stands for general) and reflek the virtualization unit of the system. This structure enables simple future scaling in terms of capacities and performance by coupling several gNodes or even just vNodes in the cluster.

The announcement includes a 'SOD' (Statement of Direction) after a cluster with two gNodes will be available. This ensures increased availability, increased performance as well as more virtual devices within the cluster and enables uninterrupted code updates. With later extensions, vNodes and hNodes can run on different systems and thus scale horizontally within the cluster. In the future it is conceivable that the vNodes will then perform different tasks, e.g. B. Virtualization for Open Systems or data archiving.



IBM TS7700 Virtual Tape Server for zSeries

Two clusters can be in one today **Dual cluster grid** (or also Multi Cluster Grid) and serve to build appropriate disaster recovery solutions. The way they work is comparable to the Peer to Peer VTS with the B10 and B20 models (synchronous and asynchronous mirroring). The newly announced TS7700 Grid Communication Feature is required for this.

Since August 2007, three clusters can be configured in one grid. The architecture itself allows up to eight clusters in a grid in the long term and will be announced in a later announcement.

The difference in mirroring is that the TS7700 uses dedicated Ethernet connections (IP). The mirroring is carried out with two 1Gbit Ethernet connections; it is considerably more powerful than Peer to Peer VTS with FICON. This lowers costs, uses open standards and simplifies the infrastructure.

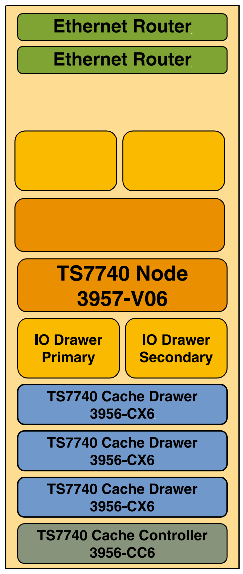
In addition to the 6TB disk buffer (18 TB compressed), the unit offers a maximum throughput of 600 MB / s (peak), almost double that of a 3494B20. Up to four 4GbitFICON connections are available in this basic unit. The TS3500 Library (3584 Library) is supported by connecting the IBM 3953 Library Manager Model L05. The unit supported drives in emulation mode of the 3592 J1A in the TS1120 backend. 'Native' support has also been available since February 2007. Up to 16 TS1120 physical tape drives can be operated on one TS7700.

As with the old VTS models, 3490E drives are mapped to the z / OS host. Functionally, the new unit offers functions that were available on the 3494B20. These are mainly the functions of the APM (Advanced Policy Management) such as 'Volume Pooling', 'Dual Copy', 'Logical Volume Sizes' and 'Multiple Reclamation Policies' (description of the functions under 3494B20), which are used in TS7700 as Outboard Policy Management be designated.

TS7700 Aufbau

- **System pSeries Server**
  POWER5 Prozessor (2 Dual Core CPUs)
- **Zwei I/O Expansion Drawer**
- **RAID Disk Controller und drei Disk Expansion Drawer**
- **Redundante Network Router**
  Stellen interne Netzwerkverbindungen zum Library Manager und zur Konfiguration des Disk Controllers zur Verfügung

  Stellen geschütztes NAT Interface für den Zugriff auf die Dienste des Management Interfaces, die auf dem Controller laufen, zur Verfügung
- **Platz für einen zweiten 3957 Controller und I/O Drawer (Statement of Direction)**



IBM 3952 Frame Modell F05

The TS7740 node emulates many 3490E drives to the host, which are visible to the host like physically existing drives. The TS7740 is therefore completely transparent to the operating system, the tape management system and the catalog entry. The emulated drives are called virtual drives. The TS7700 supports up to 256 virtual drives and up to 100,000 virtual volumes and this will increase accordingly later through the formation of multi-node clusters. 512 virtual drives are available in a dual cluster grid (mirroring).

The TS7740 cache is written to the available disk buffer (the same is read from the disk buffer for recalls). If a volume is stored in the disk buffer, it is referred to as a virtual volume. Using an algorithm called premigration, copies of the virtual volumes are migrated to the physical tape drives and written to 3592 physical cartridges, with a time delay. The copy of the virtual volume that is now stored on a physical cartridge is called a logical volume. Many logical volumes on a tape cartridge are called 'stacked volumes'. A stacked volume is therefore a physical cartridge that contains many logical volumes.

The premigration process makes full use of the drive speeds of the physical drives. The capacity of the cartridges is also used to the maximum, because there are many logical volumes on one cartridge get saved. If the host is recalled to a logical volume, the recall is served from the disk buffer as long as the original of the virtual volume is still available there. If the virtual volume in the disk buffer is no longer available, the physical cartridge that contains the copy of the logical volume is loaded into a tape drive.

The requested logical volume is then reloaded into the disk buffer and the recall is served from the disk buffer.

If the disk buffer space (TS7740 cache) becomes scarce due to a large write workload, the virtual volumes that were least used (LRU Least Recently Used Algorithm) are released for overwriting in the disk buffer after their copy has already been 'premigrated' to tape has been.

The reclamation process of the physical cartridges takes place in the background from drive to drive without loading the disk buffer. There is no reloading of the still valid logical volumes on the cartridge into the disk buffer, but these are written directly from drive to drive on a new cartridge. The old cassette then automatically goes back into the existing scratch pool.

The TS7700 forms the new basis of a mainframe-creating virtual tape server. Many new functions, improvements in performance and capacity and larger configurations are planned for the next few months, which will make the TS7700 an absolute tape virtualization specialist in the mainframe environment that cannot be compared to any other solution. This was already the case with the B10 and B20 models, which were the only systems that had a connection to the SMS (System Managed Storage) via the APM functionality as part of the z / OS operating system. The SMS construct names, such as Data Class, Storage Class and Management Class, as well as the StorageGroup names including the VolSer assignment are automatically transferred to the Library Manager Data Base. No other system on the market offers this.

For the 3494B10 and B20 models as well as for the new TS7700, even stronger integration into the SMS of the z / OS operating system is planned in the long term.

Grid reflections with TS7700 systems

With a multi-cluster grid mirroring, it is necessary to define how each storage group of each cluster in the grid is to be treated. Setting a policy on the LM defines the consistency point for each defined site in the subsystem. The policies become an array of consistency point values, with each element of the array representing the policy for a particular site. If e.g. For example, if you want site A to have a copy at the time of unloading and site B to have a copy later, the array looks like this: site A has RD and site B has RD.

The replication policies can be different on the LMs. The policy settings on each LM dictate what actions are taken when a volume is mounted on a virtual disk belonging to that LM. **Determine consistency policies** the site that receives the data first (TVC selection), which sites receive a copy of the data and at what time.

Consistency policies are configured via the management class. Policies are assigned to a management class name. Policies are set via the ETL Specialist on the Library Manager.

A new definition is that of **'Copy Consistency Points',**
which determines when a target volume on a TS7700 is consistent with the source volume. Both **Consistency Policy Options** are set for each site consistency points. There are three definitions available:

RUN (R) - This site requires a valid copy of the logical volume before Device End is sent to the RUN command (Rewind Unload) by the host (largely corresponds to the Immediate Mode with the PtP VTS).

Deferred (D) - this site needs a valid copy later, after the job has ended (similar to deferred mode for PtP VTS).

No copy (N) - This site should not receive a copy of volumes in this management class.

This new option enables extremely flexible mirror variants to be implemented in MultiClusterGrid configurations.

TS7700 extensions 2007

When announcing the TS7700, IBM stated that with four additional releases in 2007, a variety of new functional enhancements would be integrated into the TS7700 systems that go far beyond the scope of the previous B20 system. A new release became available every quarter, R1.1 in the 1st, R1.2 in the 2nd, R1.3 in the 3rd and R1.4 in the 4th quarter of 2007. The available functions are in AF-based functions (Advanced Functions) and non-AF-based functions.

physical cassette pools are created. The TS7700 then saves a duplicate of a logical volume on a second cassette to protect against a media defect. **'Cross-site replication'** is used in a grid configuration to create a copy in a second site. Depending on the setting, this is done on a synchronous or asyn chronical basis. Unmirrored volumes can also be managed in a grid configuration. **'Logical Virtual Volume Sizes'** is used to select volume sizes that exceed the standard volume sizes of 400/800 MB. You can choose between 1,000, 2,000 and 4,000 MB.

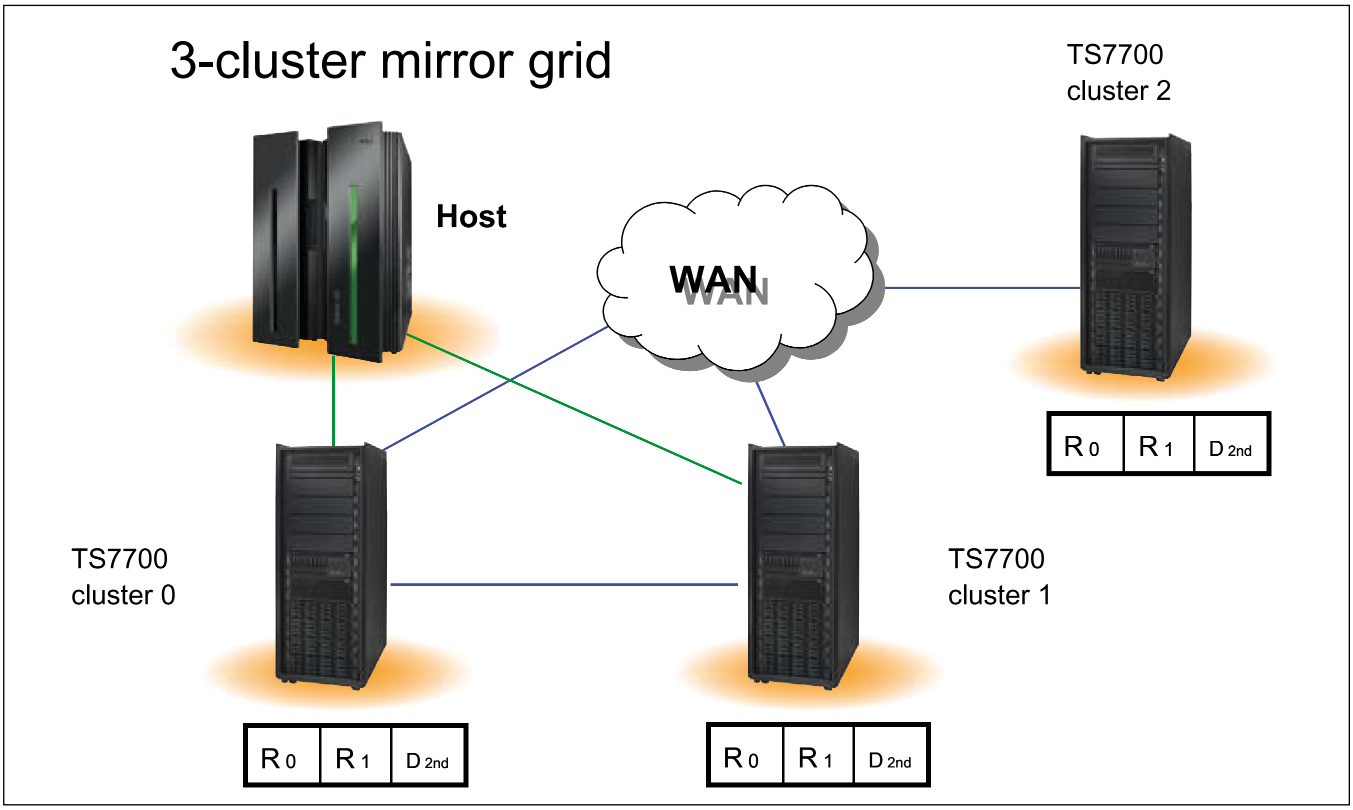
Advanced Functions (AF)

The 'Logical Volume Pooling' functionality enables you to manage logical volumes individually per number range in different physical cassette pools. The TS7700 is now multi-client capable.

With **Tape Volume Cache Management** logical volumes can be assigned to a cache management preference group. There are two preference groups available. The preference group controls how the logical volumes are treated in the disk buffer and influences the retention time of virtual volumes in the tape volume cache (disk buffer). With **Dual Selective Copy** can have two copies of a logical volume in two different

Secure data erasure with encryption shredding reflects the return of the function already known from the VTS for the fully automated 'destruction' of databases on empty cartridges (after complaint). This is a brilliant trick when TS1120 Tape Encryption is used: 'only' the key on the cartridge is deleted, not the entire tape data structure! At a **3-site grid** a TS7700 Grid is expanded so that up to 3 clusters are available for even better business continuity support (triple remote mirroring). Now you can have two TS7700 in one 3SiteGrid

eg mirrored locally synchronously and then both systems asynchronously at a great distance over WAN into a third TS7700.



Three-sided TS7700 mirror grid, clusters 0 and 1 are mirrored synchronously and clusters 0 and 1 asynchronously over a WAN link on cluster 2

Copy export (Standalone systems in R1.3, grid systems in R1.4) allow an entire copy of the data of a TS7700 to be exported, e.g. B. for disaster recovery purposes, while the original data remains in the TS7700 and is available for the production site.

Copy Export is a new TS7700 function to support the transfer of data (on stacked tapes) to a regional / national security zone for disaster recovery very purposes. The Dual Selective Copy function can be used to create a copy of volumes within a TS7700, which is stored in a separate storage pool and thus separate cartridges. Copy Export then exports a copy of selected data (log. Volumes), but the primary copy remains in the productive TS7700.

The 'Copy Exported physical volumes' are still managed by the productive TS7700, ie the volumes are still kept in the TS7700 database. The TS7700 also manages the 'Offsite' Volume Reclamation Process (recycling of the outsourced cartridges). Information about the exported volumes can be initiated via a host console request and a BVIR request.

A 'Copy Export Recovery' is carried out by the TS7700 Management Interface under the Service and Troubleshoot menu or via UserId and Password with 'rolebased' access. The TS7700 is 'offline' to the hosts during recovery. The VolSer of the volume from which the database is to be restored is required.

There is also the customer option to clean up existing data records in a TS7700 for the copy export recovery process (mfg cleanup). During the recovery process, the only panel displayed in the management interface is the status panel for copy export recovery. When the recovery is complete, the TS7700 can be taken 'online' to the connected hosts again - as soon as the TCDB and TMS have also been restored.

Additional functions (non-AF)

Autonomic ownership takeover

Ownership means that every volume in a grid configuration has an owner (cluster). To 'mount' a volume on a specific cluster, the cluster must be the owner. If it is not, the TS7740 requests ownership transfer from the other cluster site on this cluster site. In the event that the other site does not respond, the TSSC (Total Storage Service Console) checks what is wrong with the TS7740 on the other site. If the TS7740 is actually no longer available, the Autonomic Ownership Takeover is initiated. Ownership transfer can always be initiated manually.

Tape TS1120 encryption support

The TS7740 supports encryption on the connected TS1120 tape drives (from Release 1.2). The virtual drives have nothing to do with it. Encryption is managed by the existing storage pools (up to 32 pools) on the TS7740. The host uses the existing storage constructs of the storage groups and the management class to determine which logical volumes go to which pool of physical 3592 cartridges. If a pool is defined where encryption has to be used, it is treated accordingly. It can be defined for each pool which 'Public / Private' key pairs are used to store encrypted keys on the cassette that have been processed with encryption.

Host console request

This option allows an MVS console operator to perform queries and simple problem analysis of the TS7700 status without having access to the Web Interface (MI).

Enhanced Automated Read Only Recovery (in the backend)

This is the return of the function already known in the previous VTS for the fully automated recovery of error-prone 3592 back-end cartridges, also using the second copy in a grid system.

Remote Read and Write Pipelining

This function brings significant performance improvements when reading and writing to the 'remote' cache of a TS7700 in the grid.

FICON

The previous 128 logical channels per FICON channel, which could become a bottleneck in very complex configurations with many LPARs on the System z Site, have been expanded to 256 logical channels. This should no longer result in configuration bottlenecks.

The TS7700 is a tape virtualization product for the mainframe environment, for which due to its new architecture both continuous expansions and the integration of completely new developments are planned in the coming years.

Dynamic link load balancing

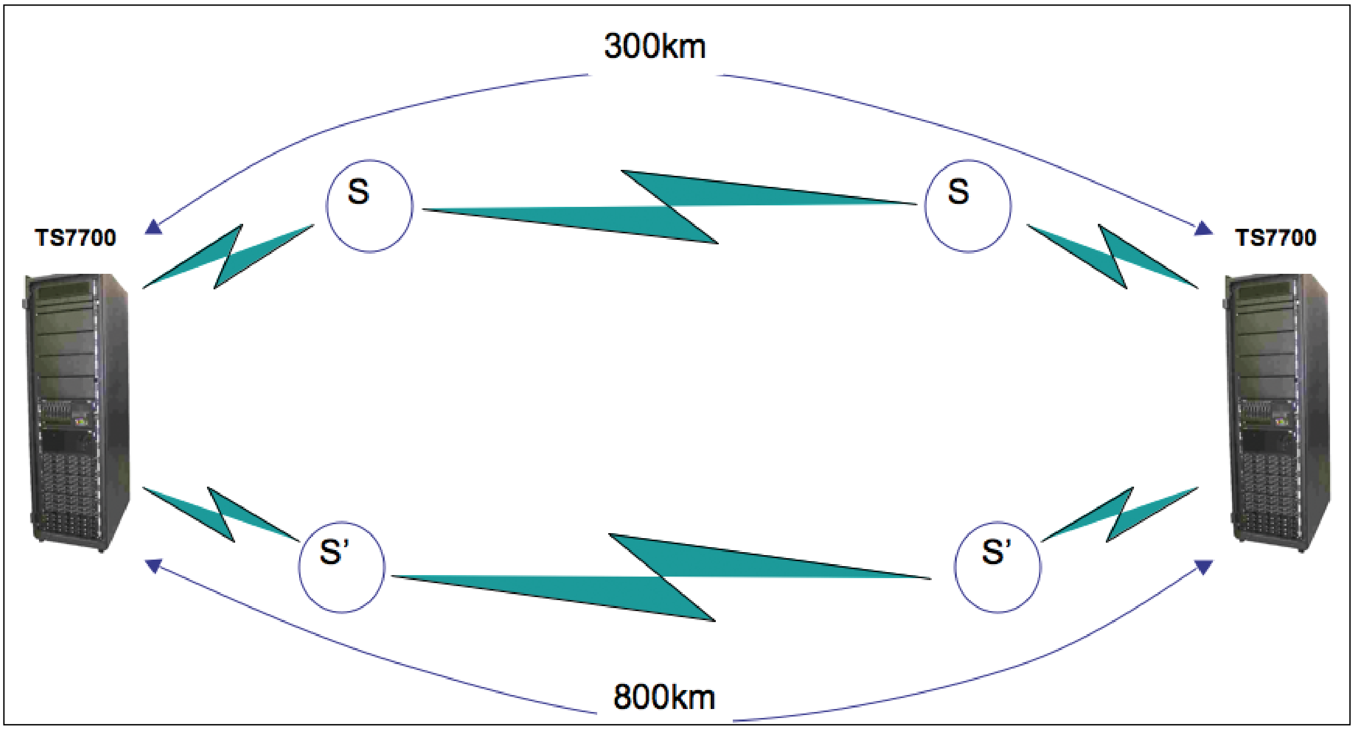
The one in **Spring 2008** announced **Release 1.4a** In addition to functional extensions, it also offers the option of dynamic link load balancing, which ensures in a mirror grid that the available channel connections are balanced with respect to the workload so that a balanced grid performance is reflected. This essentially affects the runtime differences in the network due to the different lengths of the grid connections. In order to compensate for these differences accordingly, an automatic “check” is carried out every 5 minutes, the current load is assessed and new CopyJobs are distributed according to the available service resources.

TS7700 Library Manager integration

in the **Fall 2008** IBM announces the TS7700 **Release 1.5** at. One of the main components of this R1.5 extension is that the functions of the external library manager have been integrated into the TS7700 itself. This external library manager was still a holdover of the IBM 3494 Tape Libary and has the task of managing the logical and physical volume inventory for the host and managing the associated advanced functions. This was necessary because the TS3500 tape library is a pure SCSI medium changer tape library. Since the Library Manager platform was still based on OS / 2, porting 1: 1 to the TS7700 was out of the question. However, the developers were able to use another existing IBM product: The IBM eRMM was included in the TS7700 microcode to manage the physical inventory.

TS7720 - new model

This simplification then allows the first appearance of a mainframe virtual tape system that is purely disk-based; This means that there is no longer any physical tape or tape library in the backend. This system (the TS7720) has the same structure and the same com



Dynamic Link Load Balancing for TS7700 mirror grid configurations

Components like a TS7740 (even the Ucode is the same) with the exception of the internal disk buffer. The TS7740's Fiber Channel disk cache in the TS7720 is equipped with a SATA disk system based on the IBM DS4700. 1TBSATA drives are used which are operated in the machine under RAID6. Two configurations with 40 TB or 70 TB native capacity are available (usable cache of 120 TB or 210 TB with 1: 3 compression). The TS7720 offers 2/4 FICON connections, offers 256 virtual drives and manages up to 1 million logical volumes. With the new model TS7720, grid configurations can be built up like with the TS7740. The number of virtual drives and the disk buffer capacities increase accordingly.

Additional functionality of the R1.5 release was the support of the TS1130 (Jaguar Generation 3) tape drives on the TS7740 as well as larger expansion levels of the internal TS7740 disk buffer up to 14 TB.

New opportunities with TS7700 in the mainframe environment

in the **October 2009** IBM announces the TS7700 family

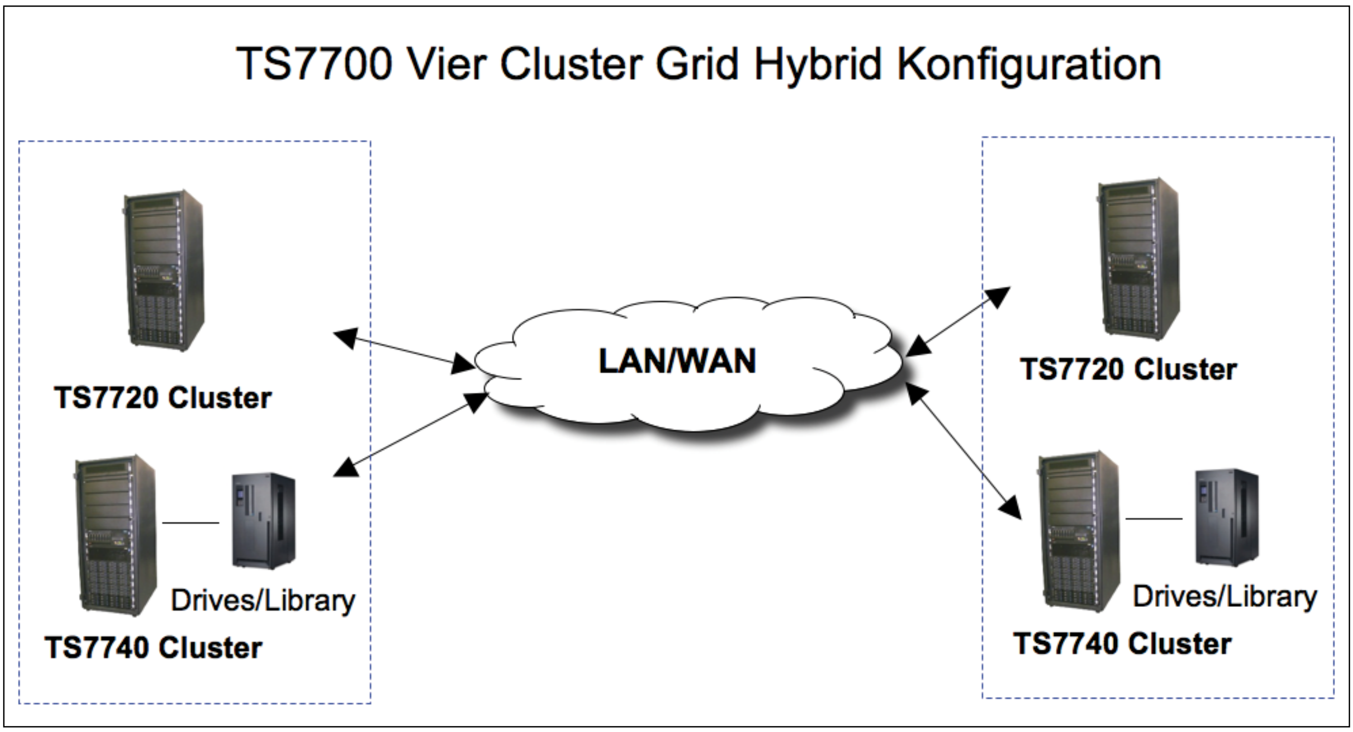
Release 1.6 at. The new release contains many new features that enable grid computing of TS7700 clusters and the associated new high-availability solutions.

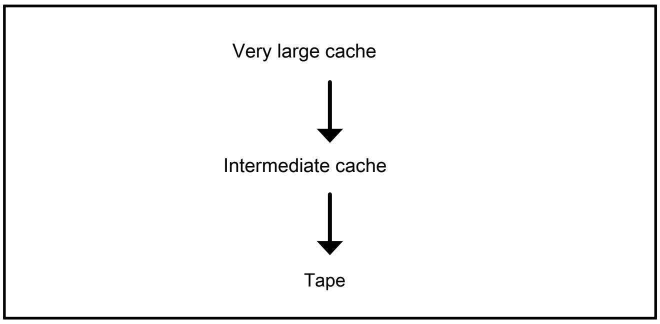
4 cluster grid and hybrid grid configurations

The R1.5, published a year earlier, laid the foundations for using the subsequent lease R1.6 to develop the unique scaling options of a TS7700 environment. This essentially required two important extensions:

- ***Increase in the maximum number of TS7700 clusters***
  a grid environment (formerly 3 to 4) Hybrid support of TS7740 and
- ***TS7720 within one***
  Grid network

This release forms the basis of backup and storage architectures that were previously not possible. So 4 location concepts can now be realized or e.g. B. a 2-location concept, in which each location is now equipped with 2 systems with high availability even in the event of a fault. A so-called 'hub and spoke' architecture is also possible, in which 3 systems serve as I / O recording systems, which then mirror their content together to form a TS7740 as a disaster recovery system.





These new architectural approaches are made even more interesting by mixing TS7720 and TS7740 systems (hybrid configurations). This allows z. B. Costs can be saved because a TS7720 without backend tape is cheaper to buy. However, the more important aspect is the 2-dimensional scalability, which allows scaling in the direction of higher I / OP performance requirements or satisfying the requirement for larger internal TS7700 caches - inexpensively. So z. For example, a customer who already operates a TS7740 2ClusterGrid solution and needs more throughput in the front end can simply expand the grid by adding 2 x TS7720 (in the sense of intelligent I / OvNodes with their own cache). Using the known methods, the data can then either be replicated to the other TS7720 or directly to the remote TS7740. With a 2-copy strategy, the possible throughput can be increased by approx. 75%. Without these new options, a completely new, second TS77402 cluster grid system with backend tape would otherwise have to be purchased and administered.

The second scaling option of a hybrid solution is that the larger internal disk storage of the TS7720 can serve as a larger but cheaper cache (SATA) of the TS7700 GridSystem.

It has its charm! So now z. B. three TS7720 can be connected with a TS7740 including tape connection. The TS7740 therefore has a huge upstream plate buffer. With three TS7720s, this would be up to 3 x 70 TB, i.e. up to 210 TB native, compressed up to 630 TB (3: 1) of upstream usable disk buffer capacity.

To optimize administration and performance in hybrids and multiple cluster environments, two important functions have been integrated in R1.6.

The ' **Cluster family function** " allows the operator to tell the TS7700 systems in which geographic positions they are in relation to each other. This enables the replication paths to be optimized. So z. For example, in a 4-cluster grid environment in 2 locations, first make a copy between the locations and only then within the locations make a copy of the respective system within the family. This procedure saves the line capacity of the replication route between the locations.

The new function has been implemented so that the disk memories of the TS7720 clusters do not overflow in a hybrid TS7700 environment **'Automatic Volume Removal Policy** ' introduced. It automatically ensures that the oldest logical volume of a TS7720 is removed from the cache (if a copy exists on another TS7740 system) and thus new space is released when required. This intelligent management allows the disk buffers of the TS7720 to appear as an additional cache in a hybrid grid environment and make optimal use of them in order to keep even larger amounts of data in immediate access, and relieves the backend of recalls from the tape.

Existing 2 or 3wayGrids can be expanded to a 4wayGrid with availability in December 2009. The merging of two 2way grids into one 4way grid is to be supported from 2010.

TS7700 Logical WORM Support

With regard to functional enhancements, release R1.6 provides logical WORM support. This function makes it possible to manage the logical volumes as WORM media on the corresponding applications on the host. The host sees the TS7700 as a logical WORM 'compliant' library. Host software extensions support the management of logical WORM volumes via the DataClass constructs. Support is available for both TS7720 and TS7740. Volumes that have already been written cannot be converted into logical WORM volumes!

New server-based tape virtualization for open systems

IBM launched the virtualization solution back in October 2005 **TS7510** A virtual tape library for the Open Systems area is available, which, however, differs only slightly in functionality from the other virtual tape libraries on the market.

in the **April 2007** IBM announced the new one **Virtual tape library TS7520 (VTL)** as the successor to the previous TS7510. In addition to higher capacity and performance, the TS7520 offers a completely new range of functions that has been integrated into the machine specifically for the requirements in the OpenSystem environment. Before the new TS7520 product is described in detail, it is important that **Trend regarding VTLs in open systems** To illuminate the environment.

After hard disks have become cheaper in recent years, especially due to the introduction of new technologies such as SATA, disk-based backup solutions are increasingly being propagated. Many market observers and customers often compare disk solutions with old tape technologies and thus find advantages in disk solutions such as faster backup and restore or more stable and reliable operation than with older tape technologies. But the reliability with pure disk solutions did not want to adjust as many wanted! Especially when using the inexpensive SATAL drives there are always disk drive errors up to data loss, even under RAID protection.

Tape is an independent medium and offers the advantages of very fast reading and writing of large amounts of data, long storage security and also does not require any electricity (**Tape is 'cool').** In addition, tape can also be used as a removable data medium. Disk is also an independent medium and offers the advantage of quick direct access with short response times and the possibility of parallelism. Optimal tailor-made solutions are therefore available in the combination of both technologies.

Nowadays, disk backup solutions are mainly used as disk buffers for high-capacity and fast tape drives. Backup data is stored on these disk systems for a short time (one to several days) and migrated or copied to tape promptly or later. Disk backup systems are also used to optimize the restore process, but only useful where disks have advantages over tape, i.e. for small files and file systems, but not for databases or image backups.

Due to the climate change, the related discussion about the reduction of CO_2 (GreenIT) and the increasing energy costs, disk backup systems should only be used with care and only where it is really necessary or where disk offers real advantages over tape. This is especially true for small and medium-sized companies. Disk backup systems entice with supposedly low costs, but tend to have electricity and climate costs that are significantly higher than the investment costs over a period of six years. In addition, there are costs for data migration that arise due to the short life cycle of disk units.

In addition to the long-standing possibility that **TSM (Tivoli Storage Manager)** so that a **LANBackup on a disk buffer** is running and the migration to tape takes place later, IBM now also offers a new VTL (Virtual Tape Library) in the form of the product **TS7520** at. The main advantage of the TS7520 is in the LANfree area, because many LANfree clients can back up directly to the TS7520 without the need for many physical tape drives (a LANfree Client always requires a direct tape drive, be it physical or virtual!). In principle, the TS7520 provides the disk buffer and emulation software that emulates tape drives. However, the data is written directly to the disk buffer.

The migration from the TS7520 disk buffer to tape can be carried out by the backup software (e.g. TSM) or directly - but with corresponding performance restrictions - by the TS7520 itself.

First of all, VTLs are nothing more than DiskBackup solutions. However, VTLs present themselves as tape libraries with tape drives. Thus, VTLs offer advantages if the backup application does not support a native DiskBackup or the license costs for such a DiskBackup are too high.

VTLs can also be used sensibly where there are many slow LANfree backups, because otherwise the tape drives required can be saved. However, VTLs are usually also integrated solutions that keep users free of any administrative tasks that would otherwise be necessary with native disk systems.

IBM Virtual Tape Library TS7520 for open system environments

A major advantage of VTL systems is the ability to compress the data and thus make better use of the disk capacity. In most cases, however, the performance of the systems drops when compression is used. Another advantage is that the data can be migrated or copied directly from the VTL to tape and this happens transparently for the BackupServer.

The IBM Tape Virtualization Engine TS7520 provides one **integrated solution of hardware and software** provides band virtualization for OpenSystemsServer via physical FibreChannel (FC) and iSCSI connections. By **System x-based servers (1-4)** tape drives are emulated. The TS7520 emulation supports the LTO2, LTO3 and LTO4 drive formats as well as the 'High End' technologies 3592 and TS1120. The **Disk buffer**

is reflected by 500GB and / or 750GBSATA disks. The plates are in one slot as **RAID5 array**

pictured. The 16 disks in the slot reflect a 6 + P and a 7 + P RAID5 array (P = parity disk) as well as a hot spare disk, which acts as a rebuild disk for both arrays in the event of a fault. The capacities can with

500 GB disks up to 884 TB and with **750 GB disks up to 1.3 PB (** native capacities). The smallest capacitive entry size starts at 6.5 TB. Depending on the type of plate, extensions can be made in 6.5TB steps or 9.75TB steps.

The TS7520 offers one **unique performance,**
by enabling up to 4,800 MB / s throughput (4,800 MB / s when reading from the disk buffer and up to
4,500 MB / s when writing to the disk buffer). In terms of both performance and capacitance, there is currently no other VTL system on the market that offers these possibilities.

The TS7520 can be configured so that up to four servers can be 'clustered'. Up to 512 virtual libraries, up to 4,096 virtual tape drives and up to 256,000 virtual cartridges are emulated. Up to 32 FC connections based on 4Gbit Fiber Channel can be configured for the hosts. Up to 16 FC connections, also based on 4GbitFibre, are available for the tape connection.

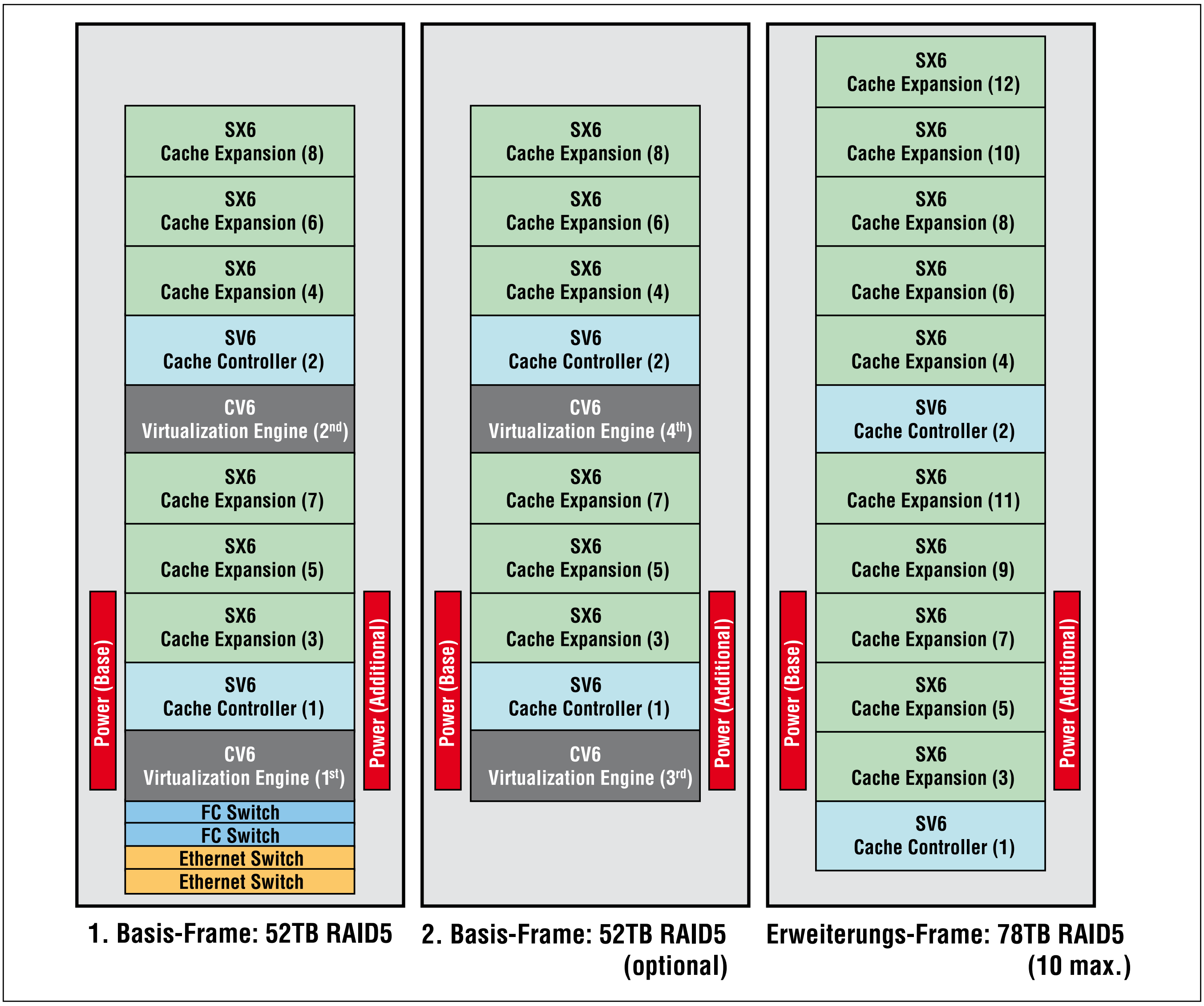
In addition to the high flexibility, performance and capacity, the TS7520 offers a range of functions that no other VTL on the market currently outperforms. This functionality will be discussed in detail later.

With a maximum configuration of the TS7520, up to **12 housing units** needed. It is important to note that even with small configurations, two plate controllers (SV6) are always used due to safety and performance requirements (Enterprise Edition). Both the disk controllers and the virtualization engines (CV6) are connected to each other via an integrated SAN based on 4GbitFibreChannel. This highly redundant structure is also evident in the power supply. Each housing unit is equipped with two power supplies.

The entry-level version is the **TS7520 Limited Edition** to disposal. This affordable entry-level option is available for entry-level configurations with a virtualization engine (CV6) and only one disk controller (SV6) and supports disk capacities of up to 29.25 TB (maximum). The limited edition does not contain the full range of functions as standard. Tape caching is e.g. B. a chargeable feature.

The **Enterprise Edition** supports every type of TS7520 configuration up to maximum expansion and offers tape caching as standard.

The IBM TS7520 was developed as a complete appliance and all individual components were coordinated. The complete system was tested with all functions and subjected to extensive performance measurements.



Setup of a TS7520 in a configuration with three housing units

For IT operators, the IBM TS7520 presents itself as one or more 'virtual' tape libraries and you do not have to worry about the connection or configuration of the disk storage system. The IBM TS7520 VTL is to be regarded as a single unit and the support is carried out accordingly, i.e. no individual components, but always the entire appliance are viewed and serviced (e.g. for firmware updates). This means that only one support team takes care of the IBM TS7520. A central point takes customer inquiries. A preliminary analysis of the problem by the customer, as is the case with implementations that were put together from individual components, is not necessary.

With **up to 48 x 4 Gbit FC connections** the IBM TS7520 offers the highest number of connectivity options.

EndtoEnd4GbitSupport means that the machine realizes a balanced performance for the front end, disk cache and tape backend in its entire connection to the hosts as well as to the disk and tape backend. Up to 256 physical backend tape drives (IBM TS1120 J1A / E05 Drives and / or LTO2, LTO3 and LTO4 Drives) and up to 32 physical backend tape libraries (IBM TS3500 Library, IBM TS3310 Library, IBM TS3200 Library, IBM 3582 Library, IBM 3583 Library and IBM 3494 Library) supported.

While some other VTL providers charge extra license costs for the number of emulated libraries, drives and volumes, the TS7520 always includes everything.

The IBM TS7520 offers an independent call home functionality to the IBM service team. Comparable VTLs only offer an 'email' call home functionality, so that errors are only reported internally, without notifying the manufacturer.

As a major unique selling point, the IBM TS7520 is the only VTL system on the market to offer one **Multipath device driver for FC failover and load balancing.** The machine can be easily integrated into a DualFabric. All VTLs available on the market offer DualFabricSupport, but only the IBM TS7520 also offers failover and load balancing. This is guaranteed by the multipath driver.

Without a multipath device driver, the individual virtual tape drives can be laboriously and manually distributed via the front-end FCPorts of the VTL, but load balancing cannot be achieved because the backup application, e.g. B. TSM, receives no information from the distribution of the virtual tape drives over several FC ports. All backup applications distribute the workload to the available tape drives according to 'round robin' and / or 'last recently used' algorithms. However, in connection with VTLs, this means that individual FC links are overloaded and at the same time some links are without load.

Especially that **Load balancing** is extremely important for a balanced and high performance with VTLs. Load balancing has long been a standard function in disk systems. This is not the case with VTLs that are actually a DiskSystem or offer DiskBackup, although most VTLs offer several FCPorts in the frontend. The **IBM Multipath Device Driver** for tapes, the workload distributes the available FC links, so that the individual links are one **balanced workload distribution**

to have. In addition, the IBM Multipath Device Driver enables automatic failover if a link or a fabric fails or is only offline for maintenance work.

Specific **Disk caching algorithms** accelerate the recall of tape when the logical volume is no longer in the disk buffer. If VTLs are used, the backup software writes the individual files to a disk buffer. In case of migration from disk to tape, the logical volume is saved in a 1: 1 copy on a cartridge. In the event of a restore of a single file that is no longer in the disk buffer, many VTLs have to reload the entire logical volume into the disk buffer in order to serve the HostRequest from there. Depending on the emulated cassette (remember: an LTO4 cassette with a capacity of 800 GB), this can take a long time, often up to two hours or more! The **Caching algorithm of the TS7520** works much smarter in such cases! Instead of reloading the entire logical volume into the disk buffer, the file to be restored is read directly from tape (without reloading) and made available to the host. This saves valuable time during restore activities.

There is the option of multiple TS7520 VTLs **Mirror asynchronously via IP connections.** After the 'demount' of a cassette, at a certain time or at intervals, a copy of the logical volume is transferred to another TS7520 via IP. Before the copied volumes created can be accessed, these volumes must be released by an administrative intervention so that they can be actively worked with. IP replication cannot be done concurrently with tape caching.

TS7530 - new model

in the **May 2008** IBM announces the more powerful IBM Tape Virtualization Engine **TS7530** which corresponds to the functional scope of the TS7520, but in addition to higher performance also offers the possibility of RAID6 and the use of 1TBSATA disks and a configuration level of up to

1.7 PB (petabyte) allowed. The smallest capacitive entry size starts at 6.5 TB. Depending on the type of plate, extensions can be made in 6.5 TB steps, in 9.75 TB steps and in 13 TB steps.

The TS7520 and TS7530 support the **Encryption options for backend tape drives (** TS1120 or LTO4). With this HWEncryption, which is integrated in these drives, no performance impairments are to be expected. For tape drives that do not use encryption technology, the TS7520 also offers the option of encrypting the data using software encryption methods before it is written to the physical tape.

In addition to FC connections, the TS7520 also enables iSCSI connections when customers use or want to use iSCSI.

About the functionality of the **Hosted backups** it is possible to install the backup software directly on the TS7520. However, this is only recommended if the backup server does not have too high performance requirements.

in the **July 2009** IBM withdraws all models of the TS7500 family from Marketing. One of the main reasons is likely to be the lack of data de-duplication. Through the acquisition of Diligent in April 2008 and the rapid implementation of this technology in IBM hardware, IBM relies on the ProtecTIER VTL, which is a unique data de

Duplication method offers and thus requires significantly less physical disk space to map the same capacity (see under IBM ProtecTIER VTL with Hyperfactor).

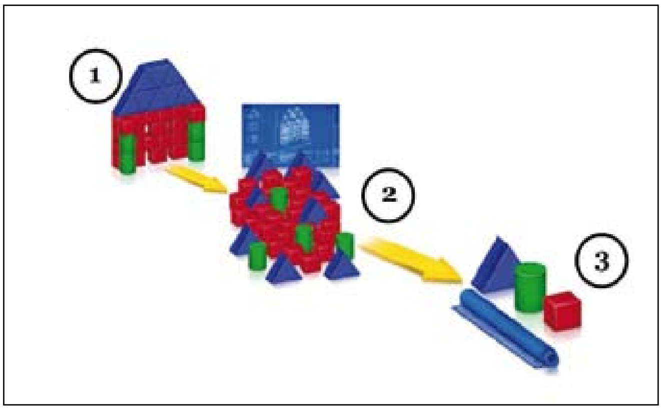
Data de-duplication

DeDuplication became an often discussed topic in 2007. DeDuplication is **nothing new.** It is a **mathematical algorithm,** the block performs bit comparisons. The only goal is to avoid duplicates. There are many different approaches and procedural options. Such procedures can be established in the operating system. An example of this is the z / OS with the Hyper PAV function. Such methods are also possible via software tools, such as B. Analysis Tools, ILM, TPC or Common Store or using comparison algorithms in SW and HW. Today's de-duplication processes are mainly used when backing up to virtual tape libraries (VTLs) because most of the duplicates are created during backup and archiving.

IBM launched the first professional de-duplication process back in 2006 for the product **IBM Common Store** a. Common Store makes these comparisons for email archiving on a mail basis and / or attachment basis and ensures that an email and / or an attachment is only archived once. DeDuplication is a feature of Common Store and is only available in conjunction with the IBM Content Manager as a backend repository.

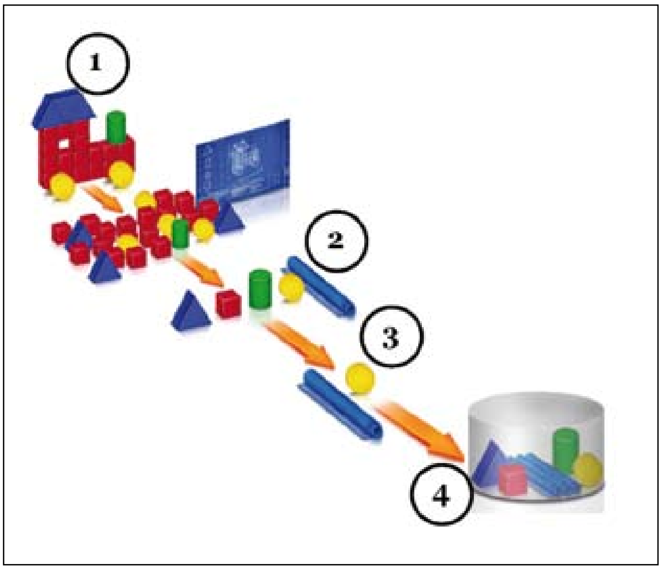
The DeDuplication approach is particularly popular with friends **Virtual tape libraries (VTLs).** The mathematical comparison algorithm runs on the VTL and carries out bit-block comparisons. The savings in disk space can easily reach a factor of 10 - 20. Different approaches are available: via software, via microcode with hardware or a combination of software and hardware. A basic distinction is made between two methods. At the **Inline process** the comparisons are carried out before the BitBlock is saved on disk. This poses the problem of scalability and performance. If too many TB are processed with DeDuplication, the VTLs quickly decrease in performance because the computer only executes the DeDup algorithm. The general recommendation of the providers is not to process more than 15 - 20 TB with DeDup. The other

To better understand de-duplication, let's look at the following simple example:



Our file to be saved is the house shown here (1). It is broken down into its components (2). It consists of red cubes, green cylinders and blue roofs. De-duplication now saves each component only once. The three building blocks are accompanied by assembly instructions on how to rebuild the house and how many building blocks are required for each element (3).

Our second file to be saved is shown in the form of a locomotive (1):



Here we find the same components that were included in the first file: the red cube, the green cylinder and the blue roof. Another additional element is added, the yellow wheel. All four components are accompanied by assembly instructions (2). Since the red cube, the green cylinder and the blue roof have already been saved, we only have to save the yellow wheel and the assembly instructions (3). For both files, the 'house' and the 'locomotive', only four components and two assembly instructions were saved. The critical thing about this procedure is: If assembly instructions are lost, there is no possibility of reconstruction!

Process is that **Post-processing process,** the algorithm taking place downstream. You write the blocks on disk first without DeDup and then carry out the comparisons. Additional storage space is then required.

DeDuplication processes are used today by **Companies IBM (Diligent), EMC 2 (Data Domain), Quantum, FalconStor, Sepaton** and **Network appliance** offered. DeDuplication is also available for the IBM Nseries in the form of the ASIS (Advanced Single Instance Storage) function.

IBM has been dealing with data de-duplication for several years. The 'Bit Block Mathematics' project was launched in January 2004 in the IBM Almaden laboratory in California. The group consisted of microcode specialists and mathematicians. The goal was to develop a powerful and highly scalable algorithm that can be integrated into both hardware and software solutions. In autumn 2008, a de-duplication process will be integrated as a post-processing process in the backup software TSM (Tivoli Storage Manager) and will be available with TSM Release 6.1.

In April 2008, IBM acquired the Israeli company Diligent to use DILIGENT technology called ProtecTIER to address the challenges of data growth and the desire for longer disk-based data storage within data backup. ProtecTIER is a software solution that has won awards for the enterprise market and has been tried and tested for five years, and which combines the technologies of virtual tape library (VTL) and data de-duplication (DDD). The DDD engine within the ProtecTIERVTL solution is called Hyper factor.

IBM data de-duplication with ProtecTIER and Hyperfactor

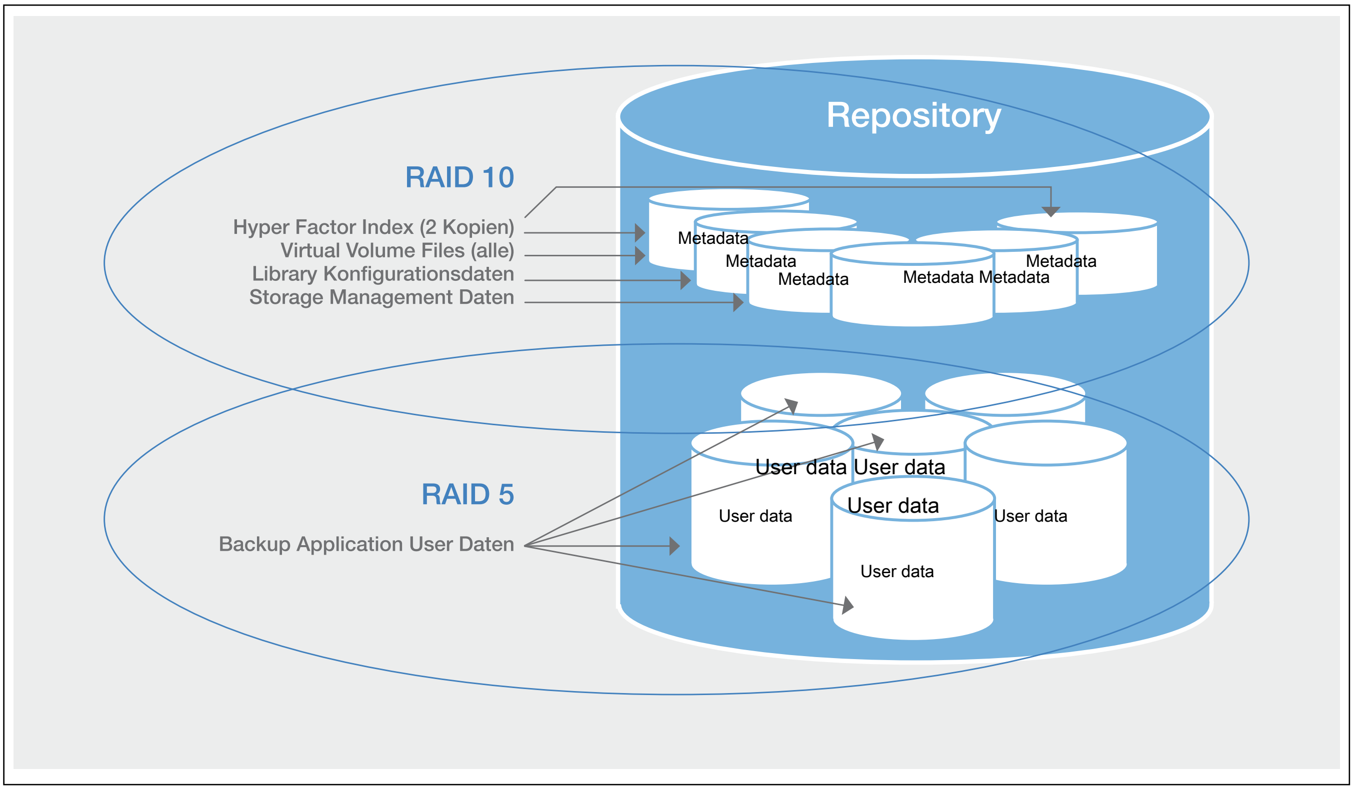
After the takeover of Diligent, IBM established a product plan to implement the ProtectVTL solution with Hyperfactor in suitable IBM hardware products. Already in August 2008 the gateway solutions, i.e. x-series-based servers, were announced, the disk systems, which serve as backup repositories, are connected upstream and work with the ProtecTIER and Hyperfactor algorithm. In February 2009, IBM announced the corresponding solutions as appliances and in July 2009 announced the direct replication option for the gate ways.

HyperFactor as data de-duplication is a mathematical algorithm that filters out blocks that have already been saved using variable BitBlock comparisons so that they are not saved twice. Hyperfactor therefore avoids storing duplicates. The saving is directly dependent on the de-duplication factor achieved, which can go up to a factor of 25. The highest de-duplication factors are achieved especially with backup processes that contain many full backups and have a relatively long retention period. With TSM, the factors are much smaller because TSM already works on an 'incremental' basis and only writes the changes made in the backup.

ProtecTIER Gateway TS7650G

After the ProtecTier gateway is installed between the servers and the disk systems on which the backup is to be made, the hyperfactor algorithms are used to filter out redundant data. This requires an index that can be used to determine which data blocks are already stored in the repository, ie on the disk systems. The algorithm analyzes the incoming data stream and uses the index to identify similarities (agnostic method). If no similarities are found, the corresponding block is saved directly in the repository. If there are similarities, the corresponding blocks are read in by the repository and compared with the new data. The identical data is filtered out and only the remaining difference is saved again in the repository. This method ensures 100% data integrity. The index has a fixed size of 4 GB and is kept permanently in the main memory of the gateway (xSeries).

The ratio of stored data to the index (ratio of repository to index) reflects the efficiency of the index. The number of 250,000: 1 in the hyperfactor process means that ProtecTier with a fixed memory index of 4 GB manages 250,000 times, ie 1 petabyte of data



Storage on the ProtecTIER repository

can. The de-duplication factor has not yet been taken into account. If the factor achieved were 25, 25 PetaByte of data could be managed. This is the great advantage of this solution, because the managed capacities can be scaled to this high dimension without loss of performance (in comparison: hash-based methods usually only achieve a ratio of about 400: 1. This means that the requirement for the main memory size in the If the computer is significantly higher and the index can no longer be kept completely in the main memory after a certain amount of data, the performance immediately drops in. The Protect solution is therefore much better suited for very high capacities and high performance compared to hash-based methods a real 'enterprise' solution).

The ProtecTier repository contains both the backup data and the associated metadata. The meta data contains 2 copies of the Hyperfactor Index, all virtual volume files, the configuration data of the virtual libraries and storage management data (pointer tables, reference counter etc.).

The data blocks themselves are stored in the RAID5 repository, while the metadata is saved based on RAID10, i.e. twice, for security reasons.

The repository can be kept on a single RAID array or distributed over many RAID arrays. LUNs that are used for the ProtecTier repository must not come from an array group that is used by other applications. It is recommended that a LUN be created for the repository so that it covers the entire RAID group.

IBM **Hyperfactor** can lead to 25: 1 data reduction and scales down to 1 PB native disk capacity. If a de-duplication factor of 25: 1 is reached, 25 PB of data can be managed with a native 1PB disk capacity. The gateway is available as a single note gateway or in a cluster configuration with two nodes. The performance in a cluster configuration is up to

1,000 MB / s, but is directly dependent on the underlying disk infrastructure and disk type (FC disks or SATA, FATA disks).

Hyperfactor's high performance comes from the fact that the algorithm works on a dynamic block basis and identified small blocks that are smaller than 8 KB are not considered and are immediately saved as 'new' in the repository. If one were to treat these small blocks as the large blocks, scalability in this high performance class would not be possible.

ProtecTier uses virtual LTO drives (LTO3) as tape drive emulation. In addition to LTO, DLT (DLT7000) emulation is also possible. Up to 256 virtual drives per node and 512 drives per cluster are supported. A cluster configuration can map up to 16 virtual libraries and emulate up to 500,000 virtual cartridges.

The HyperfactorDeDuplication process is currently the only one on the market **100% data integrity** for sure. All current IBM disk systems are supported in the backend, i.e. behind the gateway. IBM non systems from HDS, EMC and HP can also be operated.

ProtecTIER new processors for the TS7650G gateway

In the first quarter of 2009, IBM introduced new processors for the TS7650G gateways. The x3850 M2 MT7233 processor is used. This is a 4 x 6 core processor with 2.6 GHz and 32 GB RAM with two integrated RAID1-operated 146 GB SAS disk drives, 2 Emulex Dual Port FC adapters for the host connection and two Qlogic Dual Port FC HBAs for connecting the Backend Plate Repositories. Dual port Gigabit Ethernet adapters are available for replication. The new processor increases gateway performance by approximately 30%. This is made possible by the 4 x 6 core processor, which can now process 24 data streams simultaneously with de-duplication.

TS7650G disk backend as repository

In the back end of the TS7650G gateway, the IBM disk systems DS3400, the DS4000 and DS5000 models, the DS8000, XIV Storage, SVC and Nseries are supported. The non-IBM disk systems HDS AMS1000, EMC CX and HP EVA can also be operated.

ProtecTIER replication

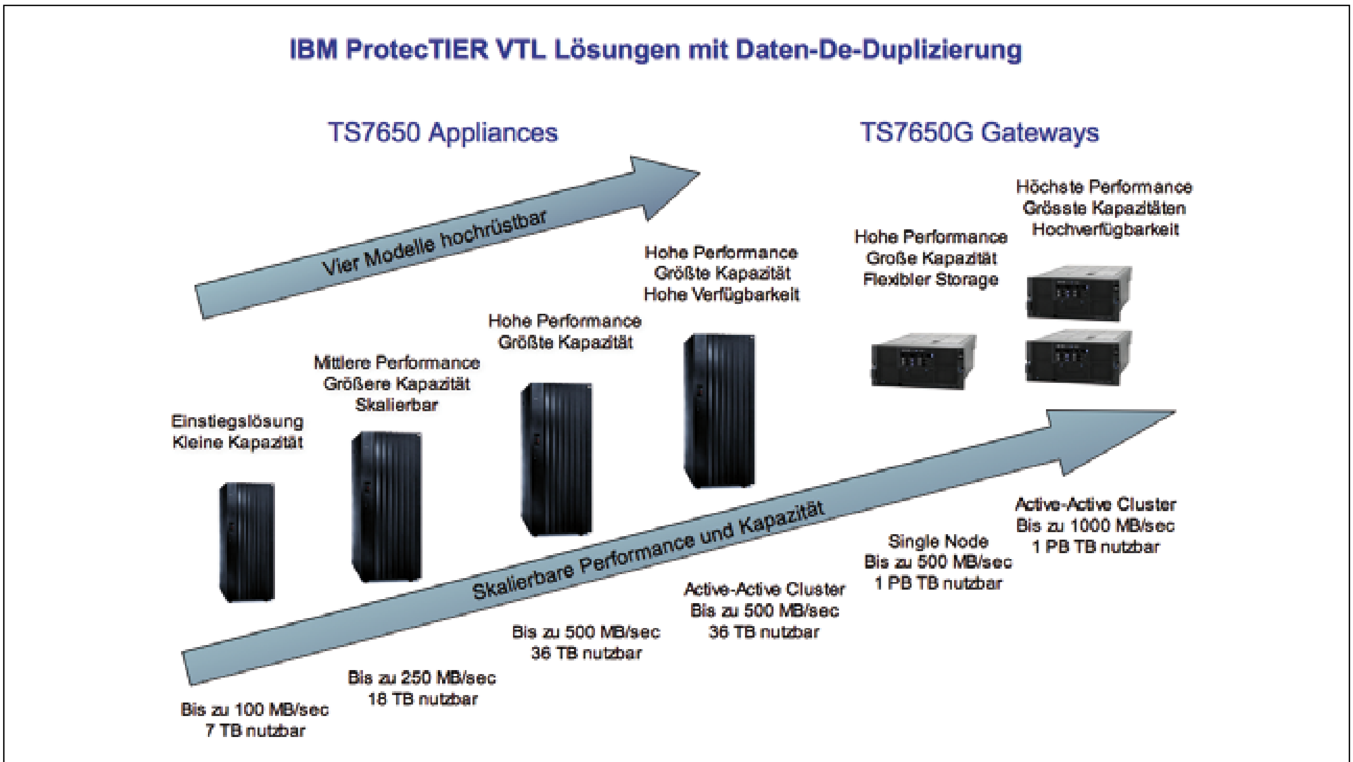
The gateway solution has had the option of IP-based replication since September 2009 (availability). This means that virtual tapes can be copied directly to another location in order to avoid the need for physical tape transports. Since only the changed blocks are transferred during replication, the required bandwidth of the IP connections is limited. The requirement for a third, 'secure' copy has existed in many IT environments for a long time, but so far this has not been technically possible because the corresponding line capacities were not available or were not affordable. Nothing stands in the way of this now and such concepts can be implemented without the need for huge transmission bandwidths.

The replication can be operated policy-based. The replication can be carried out continuously and immediately (immediate mode). However, replication can also be carried out in a planned time window (replication window). The policies can administer both individual cartridges and a pool of many cartridges. There are two operating modes: The VisibilityControl mode imports or exports the cassettes, which are controlled via the backup application (CheckOut). Cassettes that are exported are replicated in the TargetVTL. In BasicDR mode (disaster recovery), replication runs continuously and transparently with the backup application, and the cartridges are available on both the primary and remote sides.

Two EthernetIP ports are available for each node. Older systems can be retrofitted with a second Ethernet card. The special feature of this replication function is the automatic failover and failback. If the primary page fails, the disaster / recovery page can be made 'primary' at the push of a button and operation continues. When the primary page is available again, a failback is performed and the data is replicated back to the primary page.

Multipathing and control path failover

The TS7650G gateway and the TS7650 appliances together with the TS7500 family are currently the only VTL systems on the market to offer a multipath device driver for FC failover and load balancing. This means that the VTL solutions can be operated redundantly in SAN fabric infrastructures. If adapters, paths or entire SAN switches fail, operation continues uninterrupted because the 'commands' for the virtual library and drives can be automatically transferred via the remaining adapters, paths and SANS switches (control path failover and data path failover).



In addition, the data path failover offers load balancing on the redundant paths from the server to the virtual drives. This enables an even utilization of the physical FibreChannelLinks and the FCPorts.

ProtecTIER TS7650 appliances

In addition to the gateway solution, appliance solutions in the sense of 'allinone' are also available.

With the **Appliances** IBM primarily addresses medium-sized customers, who now also have the opportunity to cost-effectively establish an enterprise solution for DeDupVTLs.

IBM has checked the HyperfactorDeDuplication algorithm according to all the rules of the art to ensure that 100% data integrity is actually guaranteed. External companies and 'hackers' were also commissioned for this. No vulnerabilities could be found. It can therefore be assumed that IBM data de-duplication will adapt over time to the entire IBM storage portfolio and all operating system platforms. ProtecTier is already planned for the main frame environment (System z) in 2010.

TS1130 tape drive ' Enterprise '- great

IBM TS1130 tape drive ' Enterprise '- Great - the breakthrough in tape technology

in the **July 2008** IBM announces the new tape drive with availability in September 2008 **TS1130** and thus leads one **new technology base** for tape development!

The technological highlight of the TS1130 drive is reflected in the new read / write elements. For the first time, a tape drive technology is used in addition to the inductive write heads as a read head **GMR read head**

(Giant Magneto Resistance). Everyone was talking about GMR in 2007 when Peter Grünberg and Albert Fert were awarded the Nobel Prize in Physics for discovering the GMR effect. Without this discovery, the high capacity of hard drives would not have been possible. The implementation of this effect into a product was made possible by the IBM researcher Stuart Parkin, who needed seven years to do this and in late 1997 integrated the first GMR reading head into a disk drive. Now this technology is also finding its way into the tape sector!

GMR read heads are able to **smallest stray fields of bits** read out and at a speed that is not possible with any other technology. It is thanks to the GMR head that the TS1130 **'High speed search** '- **speed** from previously 10 m / s (for TS1120) to an incredible 12.4 m / s, i.e. from 36 km / h to 45 km / h. In high-speed search, the tape marks on the tape are read out during fast-forwarding and this is now much faster with the GMR head, so that the access speed to a file to be read out is considerably increased.

In addition to the highlights that the previous TS1120 technology had integrated, many additional facilities have been improved to ensure even greater availability and reliability.

The previous TS1120 drives are also based on the new TS1130 drives **upgradable,** which will please every TS1120 user. The media, i.e. the cassettes, remain the same, ie no media change.

3592 cassettes	Tape length	Cassette type	Gene. 1 3592	Gene. 2 TS1120	Gene. 3 TS1130
YY	200 m	Read Write	60 GB	100 GB	128 GB
YES	609 m	Read Write	300 GB	500 GB	640 GB
JB	825 m	Read Write		700 GB	1,000 GB
JR	200 m	WORM	60 GB	100 GB	128 GB
JW	609 m	WORM	300 GB	500 GB	640 GB
JX	825 m	WORM		700 GB	1,000 GB

IBM 3592 tape cartridge formats

The TS1130 drive currently operates at the highest data rate (160 MB / s native) and only needs one hour to back up compressed data (2.3: 1) with up to 1,296 GB. The maximum capacity per cassette is 1 terabyte (JB and JX cassettes). 1,152 tracks are recorded on the cassette (288 per data tape x 4 = 1,152 tracks).

The format factor of the drive and the cartridges enables the drive to be used in the IBM 3494 and IBM TS3500, TS3400 tape libraries or in standalone rack solutions. Company-wide connection options are available: FibreChannel 4 Gb / s via switched fabric with two ports for OpenSystemsServer and FICON or ESCON support (2 Gb) via A60 u J70 and (4 Gb) via C06 for servers with zOS.

Tape formats and compatibility

The 1st generation of 3592J1A drives wrote in the Gen1 format 300 GB on the JAKassette and 60 GB on the JJ cassette. The second drive generation TS1120 (3592E05) writes in the Gen2 format 500 GB on the JAKassette and 100 GB on the JJ cassette. The TS1120 drive can also write in Gen1 format 300 GB on the JA and 60 GB on the JJ cartridge. Since the availability of the JB cartridge, the TS1120 drive in Gen2 format has written 700 GB to the JB cartridge.

The third drive generation TS1130 (3592E06 / EU6) writes 1,000 GB in Gen3 format and 700 GB in Gen2 format on the JB cartridge; 640 GB in Gen3 format and 500 GB in Gen2 format on the JA cassette and 128 GB in Gen3 format and 100 GB in Gen2 format on the JJ cassette.

As with the predecessor TS1120, the ability to write to IBM 3592 cartridges in such a way that only the instance that knows and has the associated encryption key can read them again.

With the throughput / capacity features and the fast winding and FastSearch times, the IBM TS1130 drive is unique in the market. A backup window can be reduced to 80% by using TS1130 drives with the same number of drives (e.g. compared to Generation 1).

The drive is equipped with two control processors. This leads to enormous performance and reliability. In order to be able to use the enormous throughput of 160 MB / s, the TS1130 drives are equipped with two 4GbitFC interfaces. With the 3592 drives, IBM was already able to claim the world record in terms of winding speed ('High Speed Search' of 8 m / s.). With the new TS1130 drives, IBM surpasses its own record and winds at 12.4 m / s. This corresponds to 45 km / hour - breathtaking for a tape drive! Without GMR technology that could never have been achieved!

The buffer memory has been doubled to 1 GB compared to its predecessor (TS1120). This achieves a significant increase in performance with regard to virtual backhitch, since the drive remains 'streaming' for a longer time. A SkipSync feature was introduced to optimize the processing of large files, which results in an up to 80% performance improvement for these worloads. Despite the higher capacitive use of the cartridges (up to 1 TB), the TS1130 drives is much faster in all areas than its predecessor.

Overview of IBM 3592 / TS1120 / TS1130 drive specifications

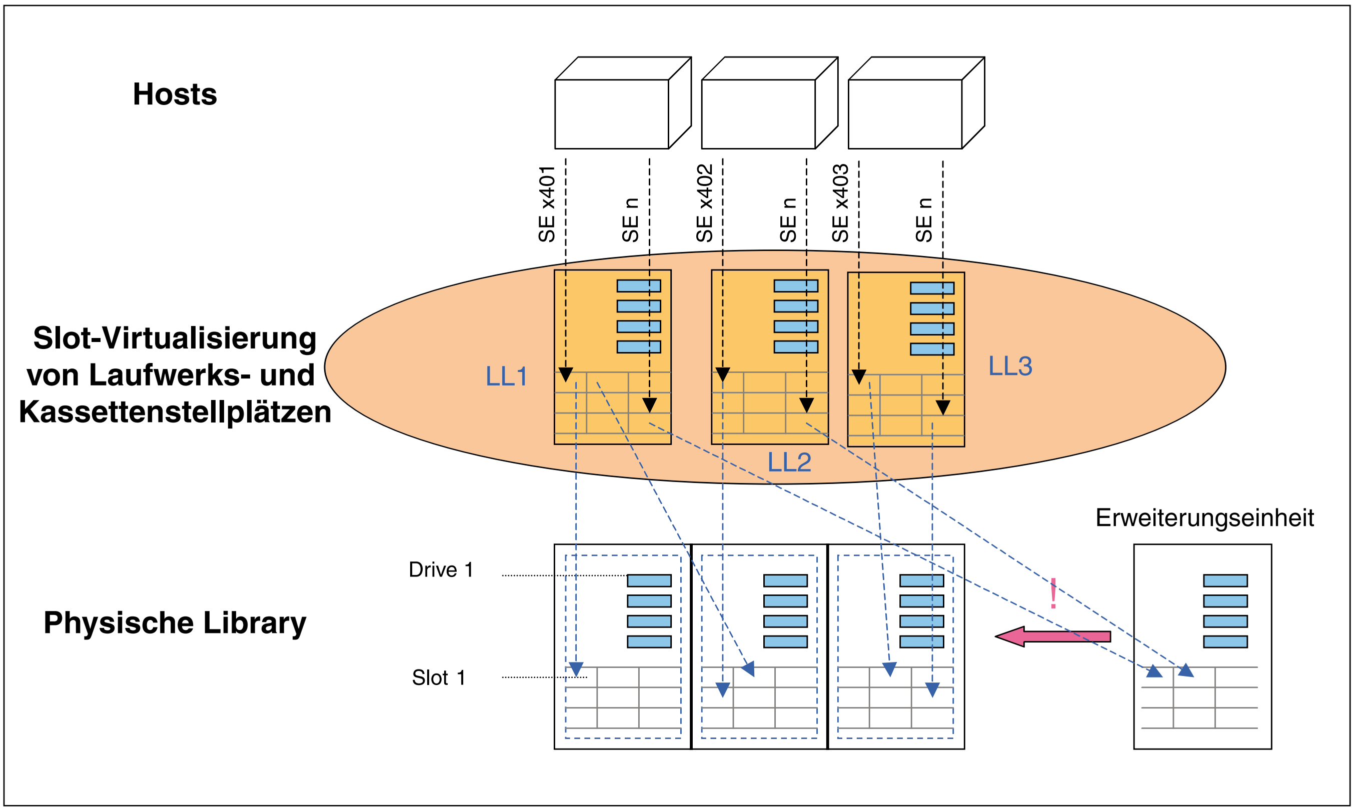
Drive feature	IBM 3592	IBM TS1120	IBM TS1130
Max. Cassette capacity	300 GB	700 GB	1,000 GB
Data rate (native)	40 MB / s	104 MB / s	160 MB / s
System z connection	2 Gb FICON / ESCON	4 Gb FICON / ESCON	4 Gb FICON / ESCON
Open system connection	2 Gbps FibreChannel	4 Gbps FibreChannel	4 Gbps FibreChannel
Virtual backhitch	Yes	Yes	Yes (extended)
Buffer memory	128 MB	512 MB	1 GB
Speed matching	6 speeds	6 speeds	6 speeds
Load thread times	16 seconds	13 seconds	13 seconds
Average access time including loading	60 seconds	47 seconds	38 seconds
Encryption	No	Yes (no extra charge)	Yes (no extra charge)
Power consumption	42 watts	42 watts	46 watts
Standby	27 watts	27 watts	17 watts

Virtualization of tape libraries

TS3500 (3584) library virtualization with ALMS (Advanced Library Management System)

The TS3500 with ALMS (Advanced Tape Library Management System) offers unique functionality. ALMS has been available for the TS3500 since April 2006, but was only really used in 2007. The physical library is virtualized with ALMS, ie the hardware is decoupled from the hosts. Virtual drives and virtual cassette slots are displayed for the hosts. This enables unprecedented flexibility because logical partitions as logical libraries can be dynamically changed during operation, i.e. reduced or enlarged.

The hosts, i.e. the backup applications, see virtual drives and virtual cassette slots completely transparently as if they were physically present. ALMS moves a virtual layer between the library hardware and the hosts, which allows hardware expansions to be carried out during operation, logical partitions to be reconfigured, reduced and enlarged, whether it concerns the drives or the slots. ALMS is a unique functionality that is only available with the TS3500 Library. ALMS offers four important functionalities for the TS3500 Library.



Storage slot and drive virtualization with ALMS

1.) Dynamic partitioning

With ALMS partitions can be created or changed during operation. The physical position in the library is no longer relevant. This applies to cartridges and drive bays.

2.) Drive sharing

With ALMS drives can be assigned to different backup servers and brought online for the backup server that needs them.

3.) Virtual I / O

ALMS offers a larger logical I / OStation. The 'Virtual I / O' function enables the backup server to display a virtual I / O station of up to 255 slots. This simplifies cassette management when checking in and out.

4.) Overall location / underallocation

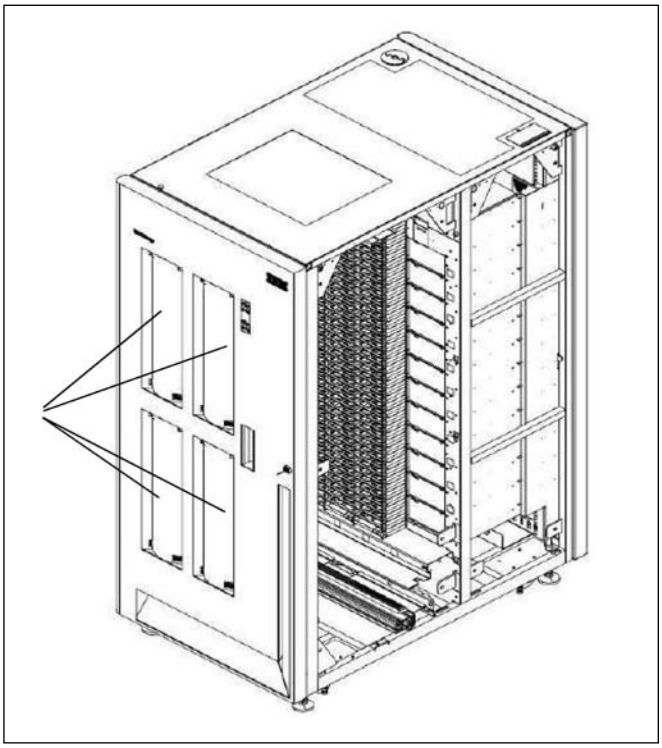
With ALMS a flexible representation of the library size is possible and thus over and underallocation. Overalloca tion reduces the administration effort for a library extension. Underallocation saves license costs, e.g. B. in the Legato environment.

TS3500 Library Hardware Extensions 2007

In addition to the Virtual I / O option of the ALMS, there has been **June 2006** the possibility of using DFrames **additional physical I / O stations** to provide. Up to four I / O stations, each with up to 16 I / O slots, can be integrated in a DFrame door.

The TS3500 supports three DFrames equipped in this way within the library. Together with the I / O station in the LFrame with 32 I / OSlots, three DFrames with up to 64 I / OSlots, the TS3500 offers the possibility **224 physical cartridges** to be stored in the library as 'bulk' or to be removed from the library as 'bulk'.

Stand for the TS3500 **flexible model conversions** to disposal. So old frames such as B. L22 and D22 with TS1120 drives can be converted into the new L23 and D23 frames. The same applies to the old LTO frames L52 and D52, which can be converted into the new LTO frames L53 and D53.



TS3500 Library 4 I / O Station D-Frame

With the new frames, TS1120 frames (L23, D23) can now be converted into LTO frames (L53, D53). The new LTO frames (L53, D53) can also be converted into new TS1120 frames (L23, D23). The TS3500 thus offers a very flexible option for converting existing frames to newly desired target frames.

The TS3500 Library offers with the functionality of the **Data gatherings** the possibility to get information about the different things that happen in the library.

The **Drive statistics** contains information about the last mounts of each drive (not possible with LTO1), the

Port statistical contains FibreChannelPort information of the last mounts (not possible with LTO1) that **Mount history**

shows the mount statistics of the last 100 cartridges that have been de-mounted and the **Library Statistics** provides information about the effective utilization of the logical libraries, ie it becomes clear whether some logical libraries are used above average. This allows the partitions to be changed to improve library performance. This is only possible with the new frames. In addition, a lot of information like **Residency Max Time**

(maximum mount time of a cassette), **Residency Avg Time**

(average mount time of a cassette), **Total mounts**

(Number of total mounts), **Mounts Max Time**

(Maximum times of mounts), **Mounts Avg Time (** By cutting times of the mounts), **Ejects Total (** Number of cassette issues), **Ejects Max Time (** Maximum times of cassette issues), **Ejects Avg Time (** Average times of cassette output) and **Total inputs (** Number of cassette inserts that are fed to the library via the I / OStation) in the last hour.

IBM TS3500 - fastest tape archive on the market

IBM TS3500 Library - New Extensions and HD (High Density) Technology

in the **July and August 2008** IBM announces massive enhancements to the TS3500 (3584) high end library.

With the on **July 15, 2008** announced new **Release 8A**
the TS3500 tape library supports the new TS1130 drives. Mixed operation with the previous drives is also guaranteed. ALMS (Advanced Library Management System) is a prerequisite for using the new TS1130 drives. In addition to supporting the TS1130 drives, the new Release 8A offers the option of SSL functionality (Secure Socket Layer). This ensures that when the library is administrated via the web, no one else can “hack in between” and thus guarantees absolute administration security. With regard to standardization, IPv6 support was also made available (an important requirement in the USA) and the standardized SMI interface was integrated.

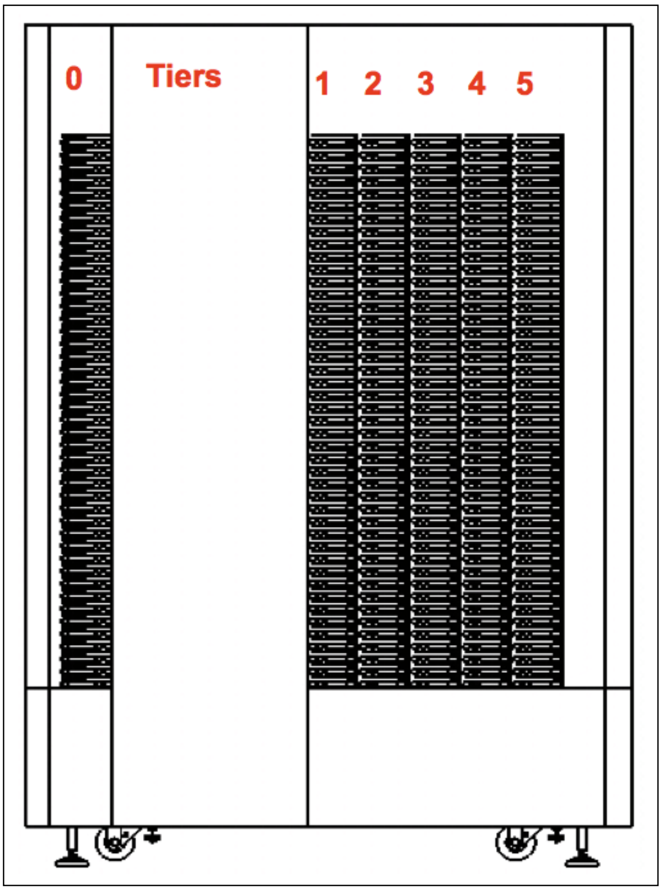
A new tool is that **TSR,** the 'Tape System Reporter'. This allows it with Windows based systems (ie
e.g. To get and collect current status queries from the library (even over a longer period). All updates and innovations are available via a microcode update.

Now stand for all Sx4 housing units **LEDs** available as lighting (picture on the right). This means that an observer can see more precisely what the robot and the cassettes are doing, and can visually follow the operation in the library. It is interesting that this requirement was made by many TS3500 users!

With the on **8/26/08** announced new **Release 8B**
are completely new, so-called **High Density (HD) frames,** available that allow over three times the amount of physical cartridges in a frame to store. The new S24 frame is used for the 3592 cassettes and the new S54 frame for the LTO cassettes. ALMS (Advanced Library Management System) and a new gripper are required to operate the new frames in the TS3500 Library. The cassettes are placed in a row in several rows of slots. If the gripper, which has been adapted to this new handling, grabs a cassette from the front row, the cassettes behind it slide forward by one slot and vice versa! An ingenious and IBM patented construction to accommodate many cassettes in the smallest space and to ensure full automation.

IBM TS3500 - LED lighting

In the S24 frame for 3592 cassettes (1/2 inch), which are slightly larger than LTO cassettes, 5 tier levels have been established, ie on the back of the frame there are 3 additional rows of slots behind each slot row. An S24 frame can store up to 1,000 cassettes. There are even 6 tier levels available for the smaller LTO cassettes, which means that 4 additional rows of slots are located behind each row of slots on the back. An S54 frame can therefore store up to 1,320 cassettes.



IBM TS3500 - HD frame S54 for LTO cartridges

The hardware in the form of HD frames is one thing, but now it is important to manage these HD frames effectively. New functions are available for the TS3500 Library.

Here is the example of the S54 frame (see illustration) with up to 5 animal levels **HD slot management** described. All tiers and thus all HD slots are completely transparent to the hosts. Three key elements are responsible for HD slot management. The function **Floating home cell** optimizes the cassette allocation in the slots of tier levels 0 and 1.

In addition, the tier 0 slots are considered **Cartridge cache** treated, ie all cassettes that are in Tier 0 slots are subjected to an LRUCaching algorithm (Least Recently Used). If the cassettes are not used for a longer period of time (compared to other cassettes of tier 0), they are moved from tier 0 to slots of tier 1. If they are not used there in comparison to other tier 1 cassettes, they are moved to tier 2 slots. Depending on the filling level of the HD slots, this goes into the last tier. Cartridge caching ensures that cartridges that are frequently required are always in the robot's direct access, i.e. in tier levels 0 and 1.

If several HD frames are now in operation within a library, the function ensures **Tier load balancing** to ensure that all HD slots of the cassette fill level are used evenly across all HD frames (balanced slot utilization across all HD frames). This ensures that the cassettes are located in the slots of the smaller animal levels as far as possible and that all HD frames are evenly optimized with regard to their animal use.

Over a **SCSI command option** The cartridge caching algorithm can be bypassed for certain Kas. The option allows you to always store selected cartridges in the cache, i.e. in tier 0 or not. The command controls the targeted prestaging and destaging of selected cartridges in and out of the cache (tier 0).

For the new HD frames, there is a " **Capacity on demand** "Option (CoD) available. By default, the new S24 and S54 frames are delivered so that 3 vertical rows of animals are activated. Tier 0 on the front and the first two Tiers 1 and 2 on the back of the frame. With this activation, an S24 Frames can use 600 slots and an S54 Frames 660 slots.

With an 'On Demand' activation, the slots that have not yet been used can be activated, ie tiers 3 and 4 are activated with the S24 frame and tiers 3, 4 and 5 with the S54 frame. With this activation, all slots in the frame can be activated be used and thus the maximum number of cassettes and thus the highest capacity can be mapped.

Model conversions of the new frames, i.e. the conversion from S24 to S54 or from S54 to S24, have been available since then February 20, 2009 available.

Mount times HD frames

The robot mount times were determined for the HD frames in a 6 frame configuration of the TS3500 under production conditions at a customer (see graphic 'Robot mount times HD frames').

Tier	"Mount" Zeiten (Sekunden)
0 (Cache)	5.93
1	6.69
2	8.78
3	19.69
4	20.81

Robot mount times HD frames

Access to cassettes in tier 0 is the fastest because the robot's gripper is oriented in its home position towards tier 0. Access to Tier 1 is 760 ms slower, as the gripper now has to turn 180 degrees to point towards Tier 1.

Tier 2 tapes require an additional 2 seconds. These occur because the robot must first remove the cassette in tier 1. Then the cassette in tier 2 slides forward into tier 1. The robot then picks up this cassette with its second gripper (dual gripper standard). The large time span is between tier 2 and 3. The 10.9 seconds come about because the robot first has to remove the tier 1 and 2 cartridges and move them to another location. This is done in one operation due to the two grippers. In the meantime, the cassettes from tier 3 and 4 have slid into tier 1 and 2. The robot can then pick it up there in a second action.

Tape library virtualization with the Enterprise Removable Media Manager (eRMM, iRMM)

The Enterprise Removable Media Manager forms a virtualization layer between data backup applications and the tape library hardware. By decoupling data backup applications and tape library hardware, the greatest possible flexibility and scalability of the tape library infrastructure is achieved and the range of functions of libraries with SCSI media changers such as. B. IBM TS3500, supplemented by expanded media management. eRMM (since

August 2007 as a program package iRMM - integrated Removable Media Manager - available) offers the more efficient use of drives and libraries through dynamic sharing and pooling of the tape drives. It allows **central management of tapes and tape drives,** that are used by different applications, thus offering the greatest possible flexibility and scalability of the tape library infrastructure. The result is shorter backup runtimes thanks to flexible tape device assignment and significant resource savings in the tape library infrastructure.

In modern environments with storage networks, servers and tape libraries are connected to one another via a storage network. This means that the technical requirements for the efficient sharing of tape resources between different backups and archiving servers are only met to a limited extent. There is no abstraction layer for tapes, which is comparable to a volume manager (SAN volume controller) or a file system (SAN file system). This means that tape resources can only be shared to a very limited extent by several applications. The backup applications usually require exclusive access to some of the resources. Until now there was no centralized management of tape resources, which enable the prerequisite for cross-application sharing. With eRMM (iRMM), IBM makes a solution available that exactly meets this requirement. This program product was developed in the IBM location in Mainz. iRMM is now also an integral part of the TS7700 Virtual Tape Server systems in a modified form and replaces the library manager previously required in System z environments.

TS3400 Mini-Library with TS1120 drives for mainframe and open systems

TS3400 mini library for high-end tape technology

in the **February 2007** IBM announced a small library **TS3400** that works like the large TS3500 with the TS1120 and TS1130 high-end drives. The library was specially designed as an entry-level solution for the TS1120 high-end tape drives (Jaguar technology). This small library can be operated both in the OpenSystems environment and in the mainframe environment (via the TS1120 FICON controller since August 2007). The TS3400 offers a native capacity of 12.6 TB (TS1120) and 18 TB (TS1130).

Can in the TS3400 **1 to 2 TS1120 and TS1130 FC tape drives** with 4GB / sDualPortFibreChannel connections. All cartridges processed by a TS1120 and TS1130 drive can be used (100, 500 and 700GB cartridges with TS1120 and 128, 640 and 1,000GB cartridges with TS1130). The library supports both rewritable and WORM cartridges and offers direct encryption and encryption key management support, an encryption technology that was made available for the TS1120 drives in autumn 2006 and is also integrated into the TS1130 drives. The library has 2 removable interchangeable cassette magazines. Each magazine can hold up to 9 cassettes. The lower magazine can be configured so that 3 slots in the magazine serve as I / OStation. The upper magazine can be configured with two cleaning cassette slots.

The TS3400 offers a partitioning of up to 2 logical libraries. Each logical library contains a drive and a magazine and can work in SequentialMode (Autoloader) or RandomMode (Library). The TS3400 is via the local operator panel or remotely via a web GUI admi nistrier and manageable. The library can be operated 'Stand Alone' or 'Rack Mounted' (5U). The storage capacity is up to 12.6 TB (up to 37.8 TB with 3: 1 compression) when using the 700 GB cartridges and up to 18 TB when using the 1,000 GB cartridges (up to 54 TB with 3: 1 compression).

The TS3400 is a highly redundant library and is equipped with redundant power supply, 'hot swappable' drives and power supplies, a double power connector as well as control path and data path failover options. The library can be connected to IBM systems i, p, x and zLinux as well as to z / OS via TS1120 controller. In addition to the IBM systems, the operating system platforms from HP, Sun and Microsoft Windows are supported.

This enables IBM to operate large TS3500 libraries with TS1120 and TS1130 drives in a central data center, while branch offices can work with the same technology in the form of a mini library. This enables data media to be exchanged in the form of 3592 cartridges between the data center and branch offices. In order to guarantee the highest security aspects, the TS1120 and TS1130 drives offer the possibility to write to the cartridges in encrypted form (see also under Tape Encryption).

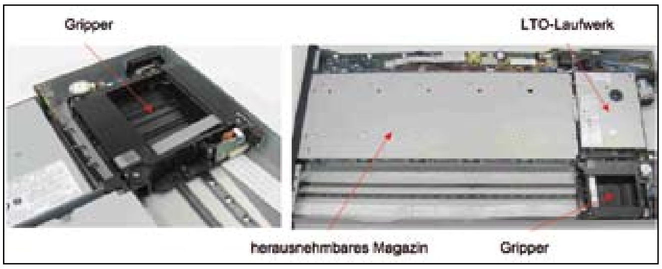
With the availability of the TS3400, the IBM Library portfolio expands in such a way that small to very large solutions with TS1120 and TS1130 technology as well as LTO technology are possible.

TS2900 autoloader with HD technology (High Density)

IBM announces in **August 2008 (** together with the HDFrame announcement of the TS3500) a new entry-level library solution

TS2900 offers capacities from 3.6 TB to 7.2 TB (native) and can be equipped with a half-height LTO3 drive or a half-height LTO4 drive. The older LTO technologies LTO1 and LTO2 are not supported in this entry-level solution.

The new autoloader has been available for the entry level since September 2008. It is only 1U high and therefore looks almost like a thin “pizza slice”. Because of this low overall height, only half-high LTO drives can be integrated. The TS2900 can be used with an HH LTO 4 SAS or an HH LTO 3 SAS drive



configure and offers space for up to nine LTO cassettes, which are made available to the autoloader in an interchangeable magazine. One slot of the magazine serves as an I / OStation. Despite the small size, the autoloader offers considerable capacity. With LTO4 you can achieve up to

14.4 TB (with 2: 1 compression). This small MiniLibrary offers direct access to every cassette and can be administered via the web.

The TS2900 MiniLibrary is an ingenious construction in the smallest space, because in the construction of 1U height a half-high LTO3 or LTO4 drive has been accommodated, a gripper system that has direct access to all nine cartridges in the magazine, power supply, all electronics including the blower as well a removable magazine with ten cassette spaces, of which only nine can be filled with cassettes.

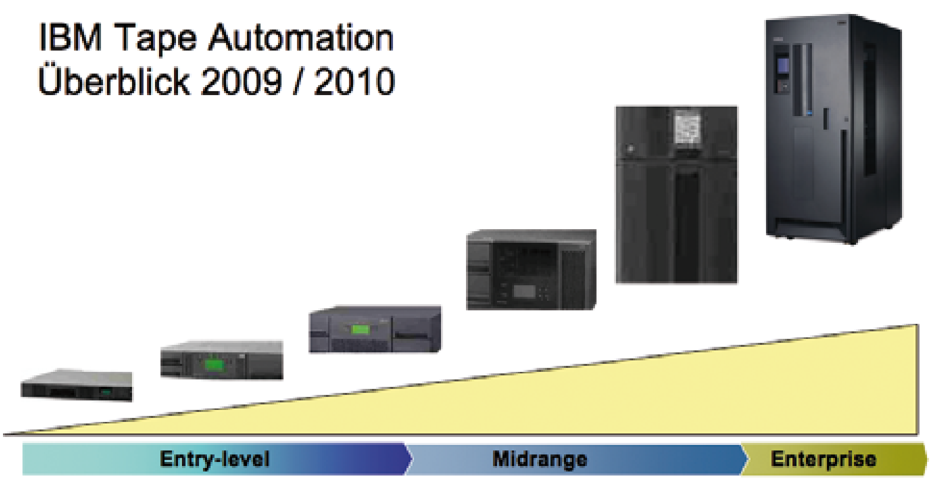
It is only possible to accommodate nine cassettes at this low height and to have them directly accessible to the gripper

through the new patented **HD technology (high density)** from IBM possible.

The magazine that can be removed is equipped with this HD technology. It actually holds ten cassettes, always two in five adjacent rows of two cassettes each and may only be equipped with nine cassettes - this means that a storage space in the magazine is always free and is used as a 'Cartride Caching'Slot.

On the left in the picture is the gripper system, which moves back and forth in the front on the right. If you have to access a cassette that is stored in one of the five slots at the back, the robot takes out the first cassette and places it in the slot where there is only one cassette! The cassette behind it automatically slides into place and is then picked up by the robot in a second step and then brought to the drive that is located on the far right of the magazine. This means that the gripper is always able to bring each cartridge to the drive on request. It is simply fascinating how ideas can be used to implement completely new things!

The graphic below shows an overview of IBM's complete library offering in 2009/2010, from the Mini Library TS2900 to the High End Library TS3500. TS2900 and TS3400 are IBM proprietary developments and are made by



Library	TS2900	TS3100	TS3200	TS3400	TS3310	TS3500
LTO 3	3.6 TB	8.8 TB	17.6 TB		bis 159 TB	bis 8 PB
LTO 4	7.2 TB	17.6 TB	35.2 TB		bis 318 TB	bis 16 PB
TS1120				12.6 TB		über 10 PB
TS1130				18.0 TB		über 15 PB

Kapazitätsangaben sind "native" ohne Kompression produced by NEC as an OEM product for IBM. The entry-level solutions TS3100 and TS3200 come from BDT and are sold as IBM logo products. As a midrange library for LTO drives, IBM has the TS3310 in its portfolio, a library that is built by Quantum / Adic and sold by OEM as an IBM logo product. Quan tum itself sells this library as the Scalar 500i itself. The difference in the product is that the IBM TS3310 supports end-to-end path failover. In the high-end area, IBM uses the TS3500, which IBM itself produces and constantly develops. The TS3500 is currently the fastest library on the market and offers many unique features and advantages over comparable libraries.

File Virtualization and File Area Networks (FAN)

Global Parallel File System (GPFS) and Scale Out File Services (SOFS)

In November 2007, using proven IBM modules such as GPFS, BladeCenter and Storage, IBM presented a highly scalable and high-performance file system with integrated ILM functionalities. The system is able to grow linearly with the requirements for storage space and throughput, and in addition to extremely high availability, it is easy to administer. In spring 2010 the IBM PLC solution will be used as the successor to SOFS. SPSC stands for Smart Private Storage Cloud.

IBM SPSC - Smart Private Storage Cloud (formerly SOFS Scale Out File Services) solution overview

The IBM PLC solution was developed to provide simple and high-performance unstructured data in customer environments and to eliminate known bottlenecks (administration, growth, throughput, e.g. in web applications) in current network-attached storage solutions - NAS for short.

Proven IBM hardware components - IBM Blade Center with IBM Storage - serve as the basis for the IBM PLC solution

Products - and the IBM General Parallel File System - GPFS for short, which has been used very successfully in high-performance computing environments since 1996. IBM GPFS provides a stable grid platform with proven scaling options and integrated high availability functions.

Through expansions with proven modules (Redhat Linux, Samba), SPSC has a highly modern and an open system platform. This grid-based solution concept provides a stable scale-out platform for future storage requirements. By focusing the PLC development on optimized management and reduced complexity requirements, IBM already integrates many functionalities into the SPSC standard and supports them with processes. These functionalities include trend analysis, data migration processes and information lifecycle management processes.

Just like the management functionalities such as monitoring, administration and operation integrated in the IBM PLC solution, interfaces allow simple integration into IT monitoring and guarantee low administration costs. To increase backup / restore times, IBM Tivoli Storage Manager - TSM for short - was integrated into the solution. Using the SPSC mechanisms significantly reduces the scan times, and the same applies to the restore.

IBM SPSC offers:

- ***a global namespace for all directories (File virtualization)***
- ***simple management and administration***
- ***High availability mechanisms of a cluster with grid Mechanisms***
- ***dynamic expansion / reduction without operational break***
- ***integrated information lifecycle functionalities with automatic automated rules at file level, in combination with TSM HSM an integration down to tape flexible coupling of application servers***
- ***through CIFS, NFS, GPFS Cross Cluster Mount very large storage volumes in one***
- ***system (> 10 peta-Byte)***
- ***integrated data migration mechanisms***

2006 to 2010

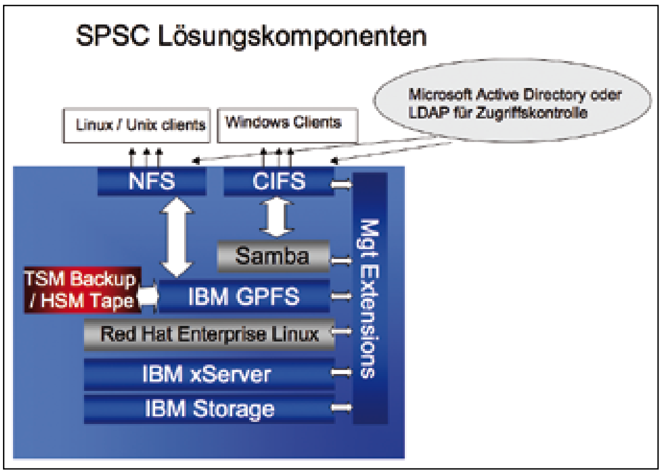
The era of server-based storage systems and storage virtualization

Solution components

In the picture 'PLCC solution components' the schematic structure of the solution is shown in a complete overview. The modules HSM, TSM are options that can be integrated into the solution on request. The PLCC solution always consists of IBM components and extensions of GNU components.

IBM solution components:

- ***IBM hardware (BladeCenter, Storage)***
- ***IBM software (GPFS, TSM, HSM, SPSC license material) including subscription***
- ***Additional software - Redhat Linux, CIFS, Samba .. including subscription, if necessary IBM***
- ***implementation service***

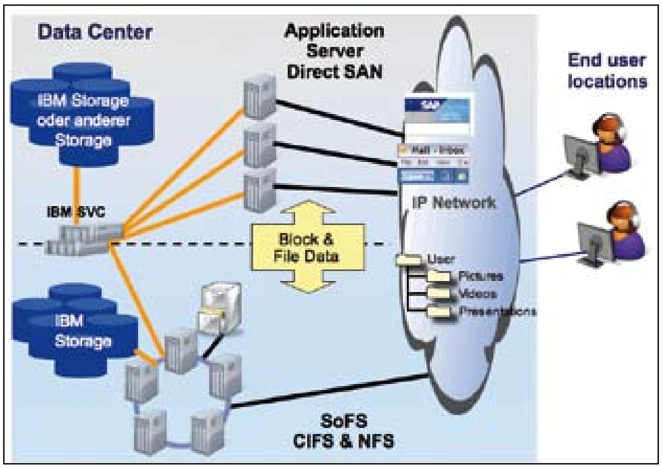


IBM server and storage hardware

A system is put together for every requirement based on IBM HW modules. The configuration is determined from the required performance (throughput requirements, memory requirements, availability and expansion requirements). By using the IBM SAN Volume Controller there is the option to integrate existing disk systems from other manufacturers (investment protection) and to take advantage of block virtualization.

IBM software

GPFS builds and provides the runtime environment of the solution. Optionally, the solution can be expanded with TSM and / or with efficient backup and recovery mechanisms and / or hierarchical storage management.



IBM Implementation Service

With the IBM Implementation Service, the PLC solution is integrated into existing IT environments:

- ***Configuration and connection planning***
- ***Installation of all IBM HW and SW components***
- ***Representing the installation of the open source products***
- ***Commissioning of the SPSC and support at Integration into the IT infrastructure Transfer of the SPSC solution to***
- ***the operating team***
- ***Documentation of the implementation***
- ***Project management***

IBM Basis Support Service

The service includes solution-oriented support with a central call entry for all questions and problems with the PLC solution. This support structure includes the problem solving of the open source components and ensures maximum solution availability.

- ***Central customer care center for error messages***
- ***Provision of information on error corrections, Program corrections for known errors Updates of the***
- ***SPSC code and the associated documentation***
- ***Providing support for the current and that previous release of the SPSC code***

IBM Premium Service (includes basic support service)

The Premium Service extends the basic support by taking on basic operating tasks such as level monitoring, change and active problem solving and maintenance of the system with software updates, optimizations and technical administration of the solution.

The IBM SPSC Premium Service provides operational services based on 7 days x 24 hours and includes the following activities of IBM:

- ***Central customer care center for all customer inquiries***
- ***Regular monitoring of the solution and active operation rating***
- ***Provision of patches / updates and import into the SPSC environment after consultation with the customer Solution***
- ***support for problems (problem owner ship including the integration of necessary functions, e.g. B. HW maintenance)***
- ***Regular support for operational optimization and use***
- ***IBM project manager for all operational inquiries***
- ***Regular status report (usage parameters, Utilization, throughput)***
- ***Optimization of the SPSC solution based on the operational experience***
- ***Planning changes (introducing new customer projects performance parameters)***
- ***Strategic development of the solution for planned Projects***

SPSC standard functionalities

SPSC provides a scalable platform for all file service requirements. Due to the very modular

The PLCC solution easily adapts to the needs and requirements. In most cases, changes to the configuration and system components can be made online.

Global namespace for all directories

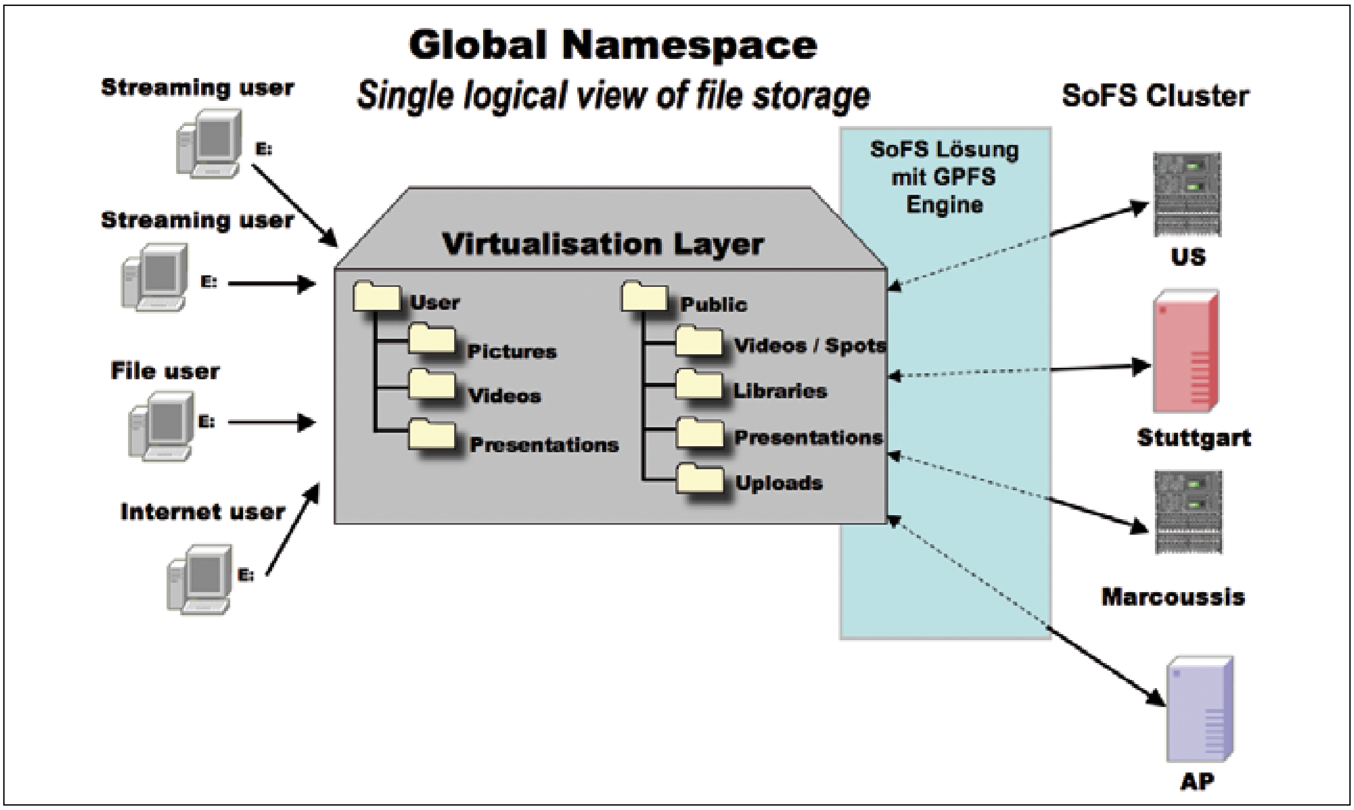
By using the GPFS file system, SPSC enables the representation of a transparent file system within the solution (file system virtualization).

This approach allows data to be stored in different storage classes without changing their representation to the user. It is irrelevant whether this is a PLC cluster or several PLC clusters are connected together.

High availability mechanisms of a grid

A PLC cluster can consist of up to 26 blades with an additional quorum server. The Quorum Server allows an automated switchover in the event of failure / unavailability of a cluster part.

The basis for this is the splitting of a PLCC cluster into two BladeCenters in two data centers (<20 km away). In this cluster structure, all nodes are active and serve incoming requests. In the event of a data center failure, the remaining part of the cluster takes over



Total amount of requests and receives the service for the user. These mechanisms are also used when mirroring the file systems.

Mirror mechanisms at the file system level

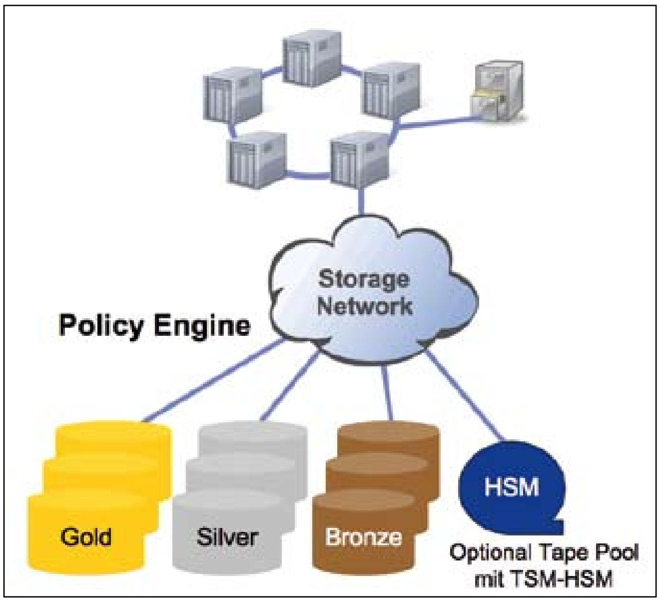
When mirroring files to increase data security, mirroring can be set up using policies based on file shares. The mirror can be defined for each (and / or) in the same share, in a different share or in a different cluster.

In SPSC, the defined files are mirrored by writing the files in parallel under the control of GPFS. This allows the write processes to be processed in parallel, especially with multiple mirrors in several clusters. A coupling of the plate systems is not necessary.

Dynamic expansion without business interruption

The PLCC solution allows the solution to be expanded / reduced under full online operation.

Changes to the PLC solution take place in independent steps. The distribution of the data is the sole responsibility of SPSC.



Expansion of storage capacity (example):

1. Basic configuration of the new storage components (LUNs)

2. Connection of the storage components to the BladeCenter

3. Transfer of resources to SPSC in the form of additional LUNs

4. The LUNs shares are assigned using the SPSC administration interface.

5. SPSC automatically strips the data via the extended share.

SPSC integrated ILM functionalities

With the integrated ILM functions (Information Lifecycle Management), IBM SPSC offers a platform for the efficient and automated handling of unstructured data. The definition rules at the file level allow automated relocation processes to be initiated on different storage classes and, in combination with IBM TSM (Tivoli Storage Manager) and HSM (Hierarchical Storage Manager), they are swapped out to tapes, whereby the file entries are displayed transparently in the file system for the user.

As already described, several different storage classes can be defined in SPSC.

In order to optimize the use of the storage classes, SPSC already has ILM extensions (Information Lifecycle) in its basic functionality. Based on policies at share level, it can be defined on which type of disk storage data should be kept. You can optionally migrate to tape:

- ***Definition of policies for files at share level***
- ***Automatic migration of the files to the defined ones Policies***
- ***Transparent access for the user (including the tape use)***
- ***Optional: When using TSM HSM - Tivoli Storage Manager Hierarchical Storage Manager - a tape drive can be integrated during migration.***

Flexible connection of application systems

In addition to the CIFS / NFS protocols, a high-performance connection to Linux application servers using GPFS Cross Cluster Mount is available. The coupling is carried out using the IBM GPFS protocol, which is used very successfully in HPC environments. To optimize data throughput, IBM SPSC supports the direct coupling of GPFS-based application servers. This coupling allows a maximum data throughput due to the protocol implemented in the GPFS.

Integrated data migration mechanisms

Based on the described procedures for online expansion of the storage, data migration can be carried out very easily and efficiently. The data migration can be carried out with minimal system loads without additional tools.

The necessary steps are:

1. ***Provision of the LUNs in the SPSC***
2. ***Set up the migration relationship between the storage areas to be replaced***
3. ***Migration of data to the new LUNs***
4. ***Dissolve the migration relationship in the storage area***
5. ***Remove the old LUNs from the SPSC cluster***

Expansion potential for SPSC

With SPSC, a scalable NAS platform is available. The scaling can take place under several history points:

- ***Throughput when accessing the files: additional blades***
- ***Volume: additional storage***
- ***Incorrect operation / data integrity: asynchronous mirror***
- ***Higher availability: SPSC Cross Cluster Mount***

All of these extensions are available within the solution, regardless of which variant and in which environment.

ILM processes with integrated outsourcing on belts

Since filers are usually filled with a large amount of data from inactive files (60–70% of the data because it is no longer required), they can be transferred to tapes at low cost. On the basis of policies, outsourcing takes place automatically via storage classes down to tape cartridges.

- ***The shift first takes place within the filer Different disk types The policies work***
- ***at the file level***
- ***According to rules, the data is automated to TSM-HSM transferred and marked for the Windows user as an offline file.***

- ***When the file is opened, it is automated from the TSM-HSM environment restored.***

Optional extensions of asynchronous mirror

With the asynchronous mirror, data can be copied online with a time delay in order to have an additional database against unwanted data deletion or in the event of a disaster. Data can be retrieved at short notice (within the synchronization period or in the case of an operation of the asynchronous cluster as an emergency system). The asynchronous mirror is a PLC cluster that runs in normal operation as part of the production cluster. The mirroring rules are defined as standard in SPSC per policy.

Long-distance coupling of SPSC clusters through grid processes

With SPSC, several variants of coupling company-wide solutions are available.

- ***Coupling of two independent SPSC clusters under a global namespace. Each of the clusters holds part of the user data.***

- ***Alternatively, an SPSC cluster can be used as an asynchronous mirror location in which the operating data are available with a time delay. If necessary, this location can be activated as an emergency location.***

Encryption for tape

In August 2006 IBM announced the encryption technology and an associated key management for the **TS1120 tape drives (Jaguar 2)** at. This means that 3592 cassettes can be written in such a way that only the instance that knows and has the associated encryption key can read them out again. Misuse of data on the cassettes can thus be excluded. If cassettes fall into the wrong hands for any reason, they cannot be read out.

Encryption options on tape have so far been available in different forms. With specially developed encryption software on a server, with specially developed encryption applications or external encryption devices, it was possible to record encrypted on tape. The tape drive itself had nothing to do with the actual encryption. Such solutions have so far only ever been used in certain areas. Implementation on a company-wide basis, regardless of the different operating system platforms and also across company boundaries, has so far not been sensible. Added to this are the many types of encryption such as Mars, RC5, Serpent, Twofish, DES or AES, to name just a few, which make the whole thing correspondingly complex. To do encryption, an encryption algorithm is required on the one hand and a so-called data key, on the other hand, which secures the encrypted data against unauthorized access. The encrypted data can only be decrypted using the data key. The different encryption methods are described below for better understanding.

Symmetric encryption: In order to prevent a third party from spying on sent data, cryptographic methods are generally used. With symmetric encryption, the data is encrypted or decrypted using a secret key. The key must be known to both the sender and the recipient and must be exchanged in person for this purpose.

Examples of well-known symmetrical encryption algorithms are the Data Encryption Standard (DES), which was developed by IBM in the early 1970s and works with a key length of 56 bits, and the International Data Encryption Algorithm (IDEA), which was developed by the Swiss Lai and Massey and published in 1990 and with a key length of 128 bits is significantly more secure than the DES. The general disadvantage of this algorithm is the direct exchange of the secret keys, which makes it difficult to use it in a customer-dealer relationship. The advantage is the relatively low computing power required.

Asymmetric encryption: Another encryption method is the so-called asymmetrical encryption. It is based on the use of a key pair that belongs together, one key being used for encryption and one for decryption. With the public key method, one of the keys is now published and can be used by any sender to encrypt a message to the recipient. Only the recipient who has the second, private key can then decrypt the message.

A representative of the asymmetrical encryption algorithms is the RSA algorithm (RSA Data Security Inc.), named after its developers Ron Rivest, Adi Shamir, Leonard Adleman, which was developed and patented in the USA in 1977 and for export in a limited encryption depth (40 Bit) is available.

Asymmetric encryption can also be used to solve the authentication problem. For this purpose, the public keys of sender and receiver are made known to each other. The sender encrypts the message first with his own, private and then with the recipient's public key. After receiving the message, the recipient decrypts the message first with his private and then with the sender's public key. However, this last step is only successful if the message really came from the designated sender, otherwise the public key used is not suitable.

Hybrid processes: The disadvantage of asymmetrical encryption is the high computing effort. This is why a combination of symmetrical and asymmetrical encryption is often used. A message is first symmetrically encrypted by the sender using a special secret key (session key). This key is then asymmetrically encrypted and transmitted with the recipient's public key. The recipient can now asymmetrically decrypt the session key and thus the actual message symmetrically with his private key. Since the asymmetrical encryption is only used for the encryption of the symmetric key, the computing effort for the asymmetrical encryption remains relatively low.

IBM is the first company which offers direct encryption technology for tape drives. To be able to work with the TS1120 drives with the Encryption option, a hardware upgrade and a microcode update must be carried out on installed TS1120 drives. TS1120 drives, which have been delivered since September 2006, are equipped with encryption capability as standard. IBM opted for AES encryption (Advanced Encryption Standard), which is available with 128, 192 and 256 bit key lengths and uses the Rijndael algorithm as a symmetric cryptography method. AES is the successor to DES (Data Encryption Standard with 56 bit key length). AES is recognized by the U.S. agency NIST (National Institute of Standards and Technology) and the U.S. Department of Commerce and is recognized to be non-crackable for years to come. The TS1120 tape drive works with the secure AES256 bit encryption algorithm. This is the first time that the drive itself has been able to perform the encryption.

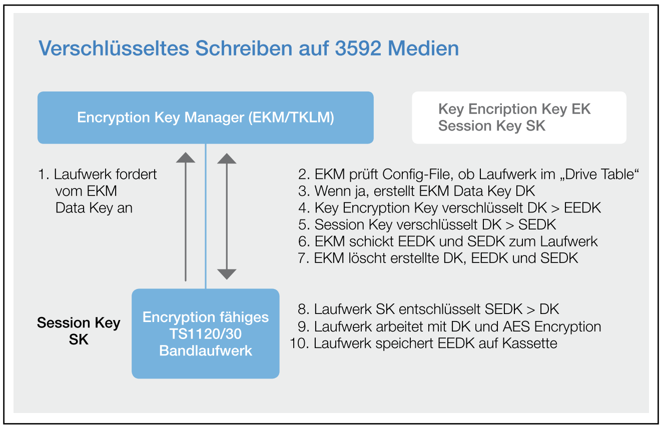
The encryption method is one side. The data keys that protect against unauthorized access and with which the encrypted data can be decrypted are just as important. In the mainframe environment, where encryption with appropriate crypto processors has been used for years, forms the heart of data key management **EKM Encryption Key Manager** in the z / OS operating system. The EKM used in z / OS offers absolutely non-manipulable and one hundred percent secure central data key management. This

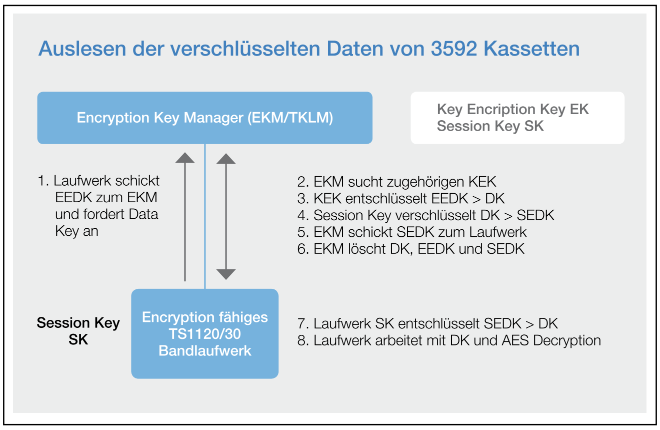
EKM has now been ported to the Java platform and is therefore available for all operating system platforms. The new EKM delivers the keys to the TS1120 drive, can run on a wide variety of system platforms and supports centralized encryption key management that can be used throughout the company.

The **EKM** stands for the operating systems z / OS 1.6 or 1.7, AIX 5.2 or later, I5 / OS 5.2 or later, HPUX 11.0, 11i, and 11.23Pl, Sun Solaris 8, 9 and 10, Linux - System z, System p and Intel, Red Hat Enterprise Linux 4 (REHL 4), SuSE Linux Enterprise Server 9 (SLES 9), Windows 2000 and Windows 2003 are available.

If a TS1120 tape drive with encryption now has to be recorded, the drive requests the data key from the EKM Encryption Key Manager. The EKM then checks whether the drive is approved and is known in the 'Drive Table' of the EKM. If not, the drive must be added to the configuration file in the EKM drive table. The EKM then creates the requested data key, which is referred to below as the 'data key'.

Since the data keys cannot be transferred unencrypted over a network to the tape drive for security reasons, the EKM works internally with two different encryption methods for the data key. One algorithm is the KEK Key Encryption Key, which runs exclusively in the EKM and with which only the EKM knows what to do. The other algorithm is the SK Session Key, which is used in both the EKM and the tape drive. If a session is established between the EKM and the tape drive and the data key is encrypted with the session key SK and transferred to the tape drive, the tape drive decrypts the original data key again and then records the data with the AES256 bit encryption. A new SK session key is always used for every newly established session between EKM and tape drive. The data key encrypted with the SK is called SEDK (Session Encryption Data Key). In addition to the SEDK, the EKM also transfers the data key encrypted via the KEK (Key Encryption Key) to the drive, which is referred to as the EEDK (Externally Encryption Data Key) and with which the drive cannot do anything, because only the EKM can use the KEK knows and can apply. The drive therefore has no way of transferring the transferred EEDK which is called EEDK (Externally Encryption Data Key) and with which the drive can do absolutely nothing, because only the EKM knows and can use the KEK. So the drive has no way to transfer the EEDK which is called EEDK (Externally Encryption Data Key) and with which the drive can do absolutely nothing, because only the EKM knows and can use the KEK. The drive therefore has no way of transferring the transferred EEDK





decrypt to get to the actual data key. The drive can only decrypt the data key using the transferred SEDK. Once the drive has completed the encrypted recording, the tape volume control region of the EEDK is stored at the end of the tape, with which the tape drive does not know what to do. The EEDK is also stored in the memory chip of the cassette. If this cassette should now be read out unauthorized on an encryption-capable TS1120 drive, this is not possible because the drive cannot do anything with the recorded EEDK. Unauthorized reading is not possible.

The drive requests the data key from the EKM to read out and transfers the recorded EEDK to the EKM. The EKM searches for the KEK Key Encryption Key belonging to the EEDK and decrypts the EEDK. The original data key is therefore available again. The data key is now encrypted again using a new SK session key and transferred to the drive. The drive receives the new SEDK Session Encryption Data Key and decrypts the SEDK to get the original data key. The encrypted recorded data can then be decrypted and read out on the cassette.

So there is no way to read the encrypted data without the original data key. The sophistication of this implementation lies in the fact that the drive performs symmetrical encryption using a data key, while the key management above is implemented asynchronously. The encrypted data key (EEDK) stored on the tape cannot be used by TS1120 drives, but only by the EKM (Encryption Key Manager) for the restoration of the original data key. The drive must transfer the saved encrypted data key to the EKM. The EKM restores the original data key and sends it back to the drive in an encrypted form (session key). The drive is now able Create the original data key and thus read out the data decrypted. The combination of synchronous and asynchronous technology make this encryption absolutely 'watertight'.

If cartridges that have been recorded with encryption are transported to another location, for example to a business partner, TS1120 drives with encryption capability and an EKM must be installed there. If the location now wants to read the data, a TS1120 drive sends the EEDK stored on the cassette to the local EKM, which has a private key to decrypt the EEDK stored on the cassette.

Up to two encrypted EEDKs with two different private keys can be stored on the 3592 cassette. The first key belongs to the EKM that generated the key, and the second key can belong to another EKM, which can convert the second EEDK into the original key based on the private key available there. It makes sense to work with two keys on the cassette if the data carrier is exchanged in encrypted form between two different locations. The drive always sends the first EEDK to the EKM. If the EKM cannot do anything with it, the second EEDK is automatically transferred to the existing EKM.

With the announcement of the LTO4 drives in **April 2007**
was for **LTO4** also **Encryption** announced. LTO4 drives, like TS1120 drives, also work with AES 256-bit encryption technology. The difference to TS1120 is that no encrypted keys (EEDK) are stored on the LTO4 cassette, but only the key identifier, i.e. the key fob. Key management and key storage is carried out via the EKM. If an encrypted cassette is to be read, the LTO4 drive transmits the identifier to the EKM, which is then able to assign the correct original key to the identifier.

The EKM was held in **May 2007** with the **EKM Release 2** so expanded that it can perform key handling for TS1120, TS1130 and LTO4. This means that both technologies can be managed with the same EKM.

The new **TS1130** The tape drive, which will be available as the successor to the TS1120 in September 2008, is equipped with the same encryption technology. The tape cartridges remain the same, which means that the new TS1130 drive uses the same cartridges. The storage of the keys on the cassettes is identical.

TKLM - Tivoli Key Lifecycle Manager

in the **Spring 2008** is available as new key management software from TKLM (Tivoli Key Lifecycle Manager), which has the same range of functions as the EKM and also provides key handling for encryption on disk (disk) and has an easy-to-use interface. The TKLM can be obtained as an official IBM program package.

IBM offers three encryption implementation options on tape:

- ***The application selects the data using encryption Tape must be written and makes the keys available to the TSM. The TSM Tivoli Storage Manager then manages the encryption key database as a backup server.***
- ***In the mainframe environment, System Utilities des DFSMS in the z / OS operating system can be selected via new data class parameters, which is to be processed with encryption. Similarly, policies for encryption can also be set up in the 'AIX Tape Device Driver'. In both cases, the new Enterprise Key Manager (EKM) manages the encryption key database. The third option goes through the library itself. Both the***
- ***3494 and the TS3500 library establish corresponding policy rules as to which 'VolSers' or which logical libraries are to be processed with encryption. The encryption key database is managed by the new Enterprise Key Manager (EKM).***

All three options allow very flexible, company-wide and tape-based encryption solutions with central key management and are unique in their solution form.

Encryption on tape secures the data written on tape cartridges against misuse and offers one
new security standard on tape, that protects the written data against unauthorized access.

Archiving

Archiving requirements have changed in recent years. In addition to the classic storage of data based on legal or internal company regulations, there are additional requirement profiles such as increased access speed and automatic outsourcing of data that is no longer used for operational purposes to archive systems. Furthermore, archive solutions are now in demand that contain hierarchical storage management and thus offer the possibility of automatically managing archived data according to ILM regulations (Information Lifecycle Management) on different storage systems. Due to the ever increasing amounts of data in the archiving area, very powerful and scalable archive systems are in demand today. which must also ensure the requirements for high availability and data security. This applies to the archiving of classic documents and content management systems as well as to applications based on filesystems. Modern archive solutions must also meet the migration requirements with regard to the data as well as the system itself when new technologies are available.

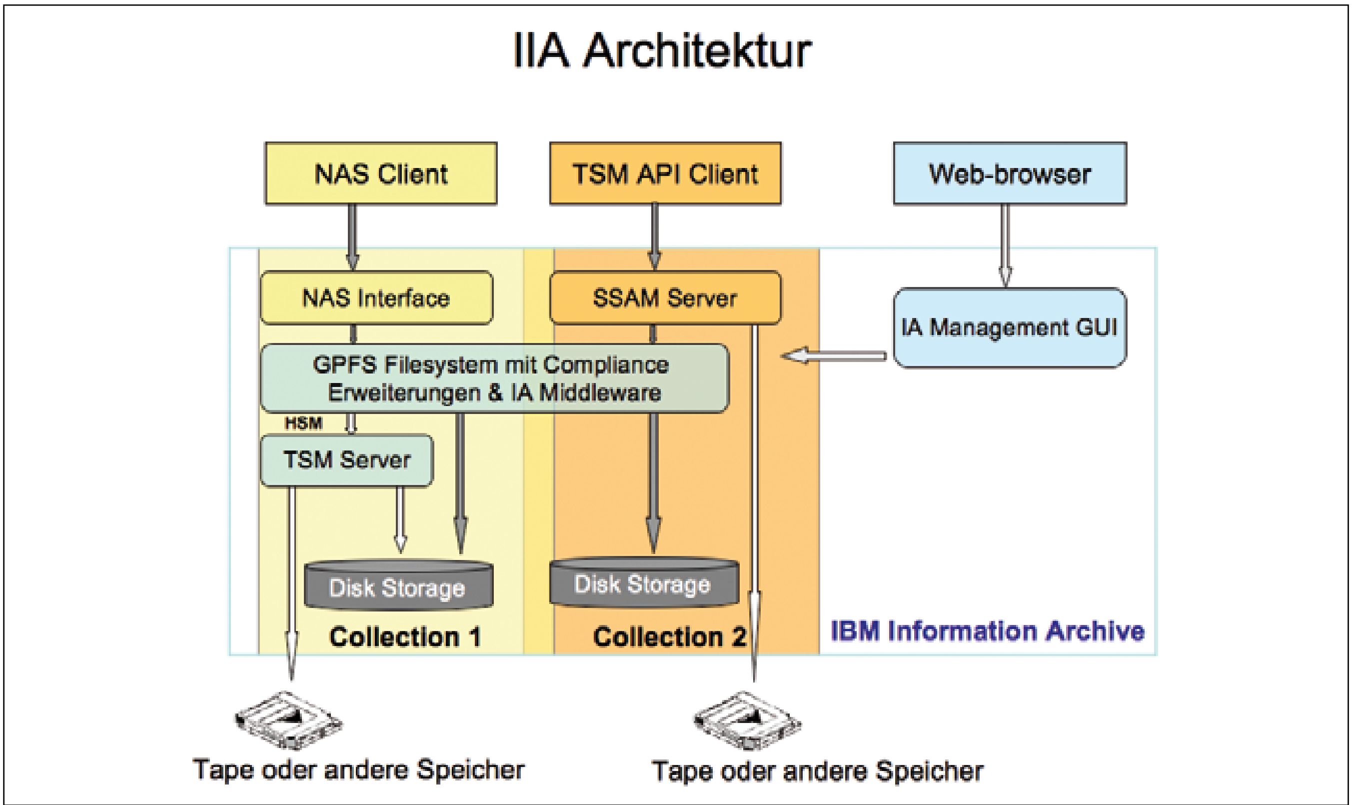
IBM Information Archive (IIA)

IBM announces in **November 2009** the new archiving system **'IBM Information Archive',** IIA for short, as the successor to the previous DR550 system.

The new system addresses the changed or expanded requirements for archive systems. IIA is a universal solution for the archiving of structured and unstructured data, which both fulfills the legal retention requirements and generally takes into account the long-term requirement of data storage free of any flaws. The system is characterized by the fact that it offers a high degree of flexibility in data storage, because applications to be archived are made possible by different standardized access methods. The high performance and scalability are special features of the system.

architecture

The system consists of up to three computer nodes (referred to as 'collection' in the IIA terminology), which are internally connected to the disk arrays by a FibreChannel network and provide each other with the integrated GPFS (General Parallel File System) backup and thus high availability guarantee. Applications can pass through nodes on the one hand via the NAS interface



the NFS or HTTP interfaces or via the classic interface SSAM System Storage Archive Manager (TSM API). FTP and CIFS will also be supported at a later date. The nodes with regard to the interfaces to the applications to be addressed are freely selectable. Both interfaces offer the possibility of other infrastructures such as B. Connect tape.

The ability to operate several archiving instances at the same time within one system increases the performance of the overall system considerably. In contrast to the DR550, the database of the SSAM server is DB2-based (through the integration of TSM 6.1). This increases the number of data objects to be saved by a factor of 3 per collection. The SSAM server ensures data migration on connected external storage devices. Within a collection for file connection (node with NAS interface), data stored on external storage such as

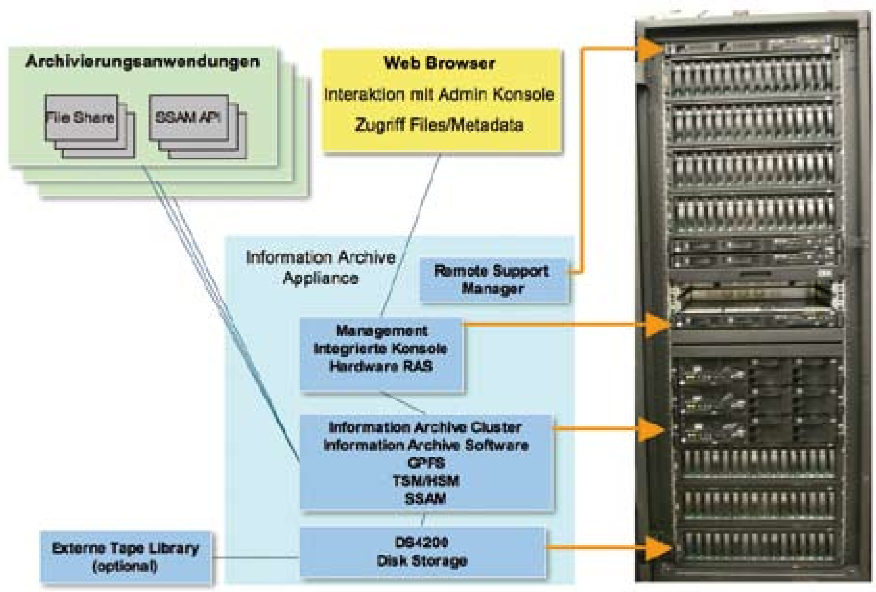
e.g. B. to be outsourced to tape, migrated through the internal TSM server using its HSM functionality.

The management and administration of the system is carried out via a uniform user interface.

technical structure

The computer nodes (GPFS nodes) consist of an IBM System x Server with two quad-core processors, 24 GB memory and a Linux operating system. Up to three nodes can be configured within a cluster. An IBM System x with a quad-core processor and 4 GB memory is used to manage the system. The disk arrays consist of 1TBSATA disks and are operated with RAID6 on dual redundant Active / Active disk control units. Each control unit is equipped with a 2 GB cache and offers up to 2 x 4 Gbit FCPorts for server connection and up to 2 x 4 Gbit FCPorts for remote mirroring. Computer nodes and disk control units are connected to each other by an 8 Gbit FCSAN based on 24-port switches, with operation initially based on 4 Gb. Each disk controller has access to every GPFS node and vice versa. The maximum configuration consists of two housing units and represents a gross capacity of 304 TB. 209 TB can be used. In the first rack, the disk capacity is up to 112 TB gross (77 TB usable). Up to 2 x 96 TB gross (2 x 66 TB usable) can be configured in the second rack. The connection of tape drives and tape libraries is fully supported.

The new architecture of the IBM Information Archive will offer scalable expansion options in the future by expanding the system with additional computer nodes and additional disk capacity.



SAN and IP networking

Brocade had in the **January 2007** McData acquired and announced a roadmap for the next 18-24 months. The architecture of the acquired McData SAN directors IBM 2027140/256 and the architecture of the Brocade 2109 M48 SAN director are mapped in a new architecture in a new product, the 'Converged' SAN director, with 8 Gbps. The announcement of the **IBM SAN768B** took place in **February 2008.** The 2027140/256 native can then be connected to the new 8Gbps director and used again.

IBM director SAN768B (Brocade DCX) with up to 896 ports

The IBM director SAN768B is being built by Brocade (Brocade DCX) and provides a platform that adapts to upcoming IT development stages through high-performance IT and future FCoE technologies (Fiber Channel over Ethernet). It is the first 'director-like' product to support transfer rates of 8 gigabits per second (Gbps) via fiber channel connections, which in the future will almost double the speed of data transfer. The director supports FibreChannel, Fiber Channel over Ethernet (FCoE), Data Center Ethernet (DCE), Gigabit Ethernet and the iSCSI protocol.

The SAN768B is the first product in the industry to support real interoperability between directors of the bType (Bro cade) and those of the mType (McDATA) in an 8 or 10Gbps FCSAN infrastructure. For users of the mType and large customers who operate their mainframes via FICON channels, this interoperability offers a road map for the expansion of the existing FICON structure. The new platform reduces complexity, failure risk and energy consumption with a special scalability of up to 768 FibreChannelPorts (external user ports) based on 8 Gbps over two domains.

In a large configuration with two domains (picture above, picture right), the director offers up to 896 ports. Of these, 768 ports are used as 'user ports for external connection'. The remaining 128 ports are routed to the backplane of the two directors via so-called ICL Inter Chassis Links and are referred to as ICL ports. The transmission bandwidth between the two domains (SAN868B chassis) can be increased up to 1 Tbit / s.

The adaptive networking services features in the SAN768B (Brocade DCX) allow the dynamic allocation of distributed resources in the fabric at the moment when the requirements of the virtual server or networked storage occur. Should bottlenecks occur (or be foreseeable), the bandwidths in the fabric can be adjusted automatically and according to the predefined service levels. This helps to automatically provide operations with a higher priority the required resources.

The **Year 2008** is the year of the introduction of 8Gbps technology in storage area networks that make it possible
8 Gbps end-to-end solutions to implement. In 2009 there were new opportunities to connect SAN infrastructures to one another via IP networks, and the magic word is **FcoE,** FibreChannel over Ethernet.

FCoE - FibreChannel over Ethernet

In addition to FibreChannel and iSCSI, FCoE is a new standard with the goal of standardizing data traffic in the corporate network and connecting FibreChannel SAN infrastructures with each other via Ethernet. LAN uses Ethernet and the TCP / IP protocol for data transfer. SANs communicate via FibreChannel. Here, servers use for the different network types

Separate interface modules, the network interface cards (NICs) for Ethernet and FC, as well as separate switch infra structures. FCoE transfers the data in a FCSAN via the local productive network via Ethernet. This means that existing SAN systems can continue to be used and integrated into other systems via Ethernet, which in the long term will reduce operating costs through a consolidated connection. Instead of maintaining two parallel networks, all data traffic should be able to be processed and administered via one network. FCoE supports a consistent management model that is common in FibreChannelSANs. This new converged technology enables both fiber channel data and IP data to be transmitted via the same transport mechanism.

So that fiber channel packets can now be sent via an Ethernet transport protocol, various manufacturers of storage solutions and network components have come together and developed a proposal for the FCoE standard. FCoE specifies how FC data can be transported losslessly over IP. This 'lossless Ethernet' consists of two protocols: first the protocol for the transport of data, called FCoE, and secondly the protocol for controlling the FCoE connections, called FIP (FCoE Initialization Protocol). FIP takes care of the login and logout of devices in an FCoENetwork. FCoE packs the FCFrame into an Ethernet header and sends it via Ethernet, but now no 'traditional Ethernet' is used, but an improved version,

What about iSCSI? iSCSI sends SCSI commands via Ethernet and TCP / IP and does not require expensive FC networks. Classic FibreChannel loses against iSCSI especially in terms of costs. iSCSI uses the existing infrastructure, while IT requires completely new components for a FCSAN. Now FCoE is coming as a new technology and has all the properties that iSCSI has, including the costs. This raises the question of how iSCSI will continue if FCoE prevails as a technology.

Brocade acquired Foundry Networks in July 2009 and has made a pretty good connection. Brocade builds hardware and software for the storage landscape, Foundry is on the move in the classic corporate network. In the future, Brocade will be able to expand in the FC and in the ether net world, thereby fulfilling the convergence that FCoE is striving for.

Cisco acquired Nuova in April 2009, a company that has been dealing with FCoE for a long time.

FCoE is now backed by Brocade, Cisco, Nuova, EMC, Emulex, IBM, Intel, QLogic and Sun. The specification for 'FibreChannel over Ethernet' are that **T11 committee** of the US National Standards Institute (ANSI) to establish a common standard. The FCoE standard has been adopted since June 2009. It will take some time before FCoE can be used professionally in the IT environment! All necessary protocol standards are expected to be completed by mid or late 2010.

SAN & Converged Networking

Following the announcement of the IBM SAN768B in February 2008, the integration of the McData technology that Brocade acquired in January 2007 was complete in a new 8 Gbps Director product. Since then, a great deal has been added in the collaboration between Brocade and IBM in the area of storage networking: after expanding the range of solutions with 8 Gbit / s HBAs from Brocade, a manufacturer-wide server-to-storage connectivity is available for the first time - with great advantages in the continuous management of the environment . In the area of network convergence, IBM now offers with the IBM® Converged Switch B32 a top-of-rack solution for users who use FibreChannel over Ethernet (FCoE) and want to gain experience with the new technology. Finally, the SAN384B is another core and edge switching platform for medium to large SAN environments, with which further design options in the backbone area are possible.

' End to end '- support

The 8 Gbps FibreChannel host bus adapters enable a high degree of robustness and reliability to be built for the IBM System x product family. Based on the

The high performance of the Brocade DCX Backbone product family means that the 8 Gbit / s HBAs are integrating important adapter functions such as virtual channels for the first time in the industry.

Using the 'Adaptive Network Services', the HBAs offer a real Quality of Service (QoS) in the SAN from virtual servers to storage arrays and support functions for data authentication that enable faster and more secure communication between virtual servers and storage.

The 8 Gbit / s HBAs were designed to increase the I / O performance on the server to 500,000 input / output operations per second (IOPS) per port. The 8 Gbit / s HBAs thus achieve twice the throughput compared to comparable products. Due to the tight integration, the HBAs are extremely easy to install and integrate into existing SAN environments.

Expanded Fabric Backbone Options with the SAN384B (Brocade DCX-4S)

The IBM SAN384B is an entry-level version of the already available IBM SAN768B backbone (Brocade DCX). The new switching platform, which can be used both in the core and in the edge, is designed for the consolidation of servers, storage area networks (SAN) and data centers. Since the complete DCX functionality is available, the SAN384B can reduce the costs for the infrastructure and administration.

The IBM SAN384B scales in four blade bays to a total of 198 ports with full 8 Gbit / s and is therefore suitable for operators of medium-sized networks. Large storage network operators can also use the IBM SAN384B as a network edge. The IBM SAN384B is equipped with the latest Brocade Fabric OS (FOS) version 6.2, which offers a new VirtualFabric option that enables the logical partitioning of a physical SAN into logical fabrics.

The IBM SAN384B optionally supports the encryption of data on hard drives and tape drives (Data at Rest) in connection with different key management systems. The convergence in the network area ensures the IBM SAN384B with its multi-protocol architecture by the support of emerging

Standards such as converged enhanced Ethernet (CEE) and FibreChannel over Ethernet (FCoE). These protocols are made possible by simply adding appropriate blades at a time desired by the user.

The adaptive networking services features - as in the SAN768B - allow dynamic allocation of distributed resources in the fabric at the moment when the requirements of the virtual server or the networked memory occur. Should bottlenecks arise (or be predictable), the bandwidths in the fabric can be adjusted automatically and according to the predefined service levels. This helps to automatically provide operations with a higher priority the required resources.

FibreChannel over Ethernet (FCoE) products

Even if the development is not yet complete - the first standards for FibreChannel over Ethernet (FCoE) are there. With the IBM Converged Switch B32 and the associated 10 Gb Converged Network Adapters (CNA for IBM System x), a solution is available that already allows the use of this new technology. FCoE should actually be called FCoCEE, since a new Ethernet protocol was developed for lossless transport with converged enhanced Ethernet (CEE).

The IBM Converged Switch B32 is a multi-protocol capable, Layer 2, TopoftheRack FCoE switch with 24 x 10 Gigabit Ethernet (GbE) CEE ports and eight FC ports with 8 Gbit / s. It enables the consolidated input / output (I / O) of memory and network ports on the server side. The new converged network adapters (CNAs) offer a performance of up to 500,000 I / Os per second (IOPS) and are available in two versions: the single port brocade CNA 1010 and the dual port brocade CNA 1020. The new brocade adapters are part of the range the first CNAs with SingleASIC, which offer classic Ethernet based on the TCP / IP protocol, converged enhanced Ethernet and storage network functionality via a single 10 Gb / s FCoE link from the servers via the SAN to the LAN.

In combination, the IBM Converged Switch B32 and the 1010/1020 CNAs enable end-to-end FCoE connectivity from the server to the storage. The solutions can be integrated into existing FC environments and can be administered with a comprehensive management platform, the Data Center Fabric Manager (DCFM).

Green IT

The topic of energy efficiency is increasingly becoming the focus. Some data centers can no longer accommodate servers or storage units today because it is technically impossible.

According to an IDC survey from 2007, topics relating to energy consumption and energy supply are becoming increasingly important. Rising electricity prices are burdening the available budgets. In many data centers, the topic of energy efficiency has become a hot topic. 'Electroless' data carriers such as tape cassettes or optical disks will experience a renaissance in the next few years. According to a Gartner study from 2006 and 2008, more than 70% of all data is saved on tape. If all of the data that is stored on tape cassettes in productive use were moved to disk systems (disk), around 40 new nuclear power plants would have to be built for the power supply. The Internet and the increasing use of the Internet also increase the energy problem immensely:

in the **May 2007** IBM announced the project **'Big Green'** at. IBM has been investing since this announcement **$ 1 billion** per year in the project to improve 'green' technologies and services and works on a roadmap to deal with the IT energy crisis that has occurred for data centers. To this end, IBM has built a team of more than 850 energy specialists. The added value for IT companies often results from the fact that energy costs for a typical data center can be halved. In comparison, this corresponds to a reduction in emissions that approximately 1,300 cars produce. In addition to the hardware and software, IBM offers a variety of corresponding services for data center operations. Part of these services include energy efficiency analyzes and / or 'EnergySelfAssessments'.

The exemplary activities of IBM in the 'Big Green' project and the associated long-term commitments led to IBM becoming the 'Top Green IT Company' of 2008 by both IDG (International Data Group) and in February 2008 was chosen and awarded by the computer world.

General virtualization

Virtualization in any form will shape the era of server-based storage architectures. In addition to tape virtualization and disk virtualization solutions in the SAN, virtualization will expand to far more areas. The **IBM** represents the worldwide **'Virtualize Everything' strategy.** This is the only way to make optimal use of heterogeneous servers and storage units and a wide variety of infrastructures and to enable flexible data management based on regulations. Automation processes in today's different infrastructures will only be possible through virtualization in all areas. Server-based storage architectures considerably simplify the overall virtualization approach.

Effective virtualization solutions achieve the maximum possible use of the available physical resource, which means that existing storage capacities are used much more. Thus, virtualization makes a significant positive contribution to dealing with energy problems in the IT environment.

New infrastructures and bandwidths

Today's Fiber Channel networks and IP infrastructures are certain to be further developed over the next few years. Today's 4GbitSANs and 8GbitSANs are heading in the direction **12 Gbit SAN** Technology further. IP networks today with up to 2Gbit and 10GbitEthernet could in the next few years **100 Gbit Ethernet technology**

be replaced. The IP infrastructure, especially Ethernet, will be of particular importance in the next few years. It is planned to provide every private household in Germany with a 100 Mbit Ethernet connection within the next five years. This gives private households new, huge transmission bandwidths. Ethernet infrastructures are also becoming a key factor in the private environment.

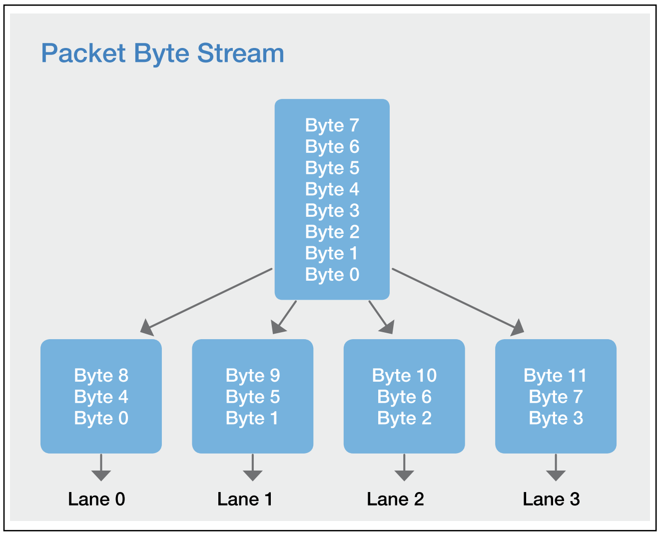
Just as the first SAN infrastructures were used two years earlier in the high-performance data centers of the universities before they were used in the commercial data center environment, similar trends are developing today for a new transmission technology, the

InfiniBand referred to as. **InfiniBand** stands for **'Infinite Bandwidth'.** InfiniBand is a new transport protocol that can connect server units, storage units and entire networks of data centers. The standard has been driven by the InfiniBand Trade Association since 1999 after Compaq, Dell, HP, IBM, Intel, Microsoft and SUN announced that they would adapt their future I / O developments to InfiniBand technology. The standard has been available since autumn 2005.

InfiniBand enables today's LAN, SAN, NAS and IPC (Interprocessor Communication) infrastructures in one **universal network** to unite. In addition, separate InfiniBand networks could be operated via iSCSI in addition to existing SAN and IP infrastructures.

Copper cables (distance limitation at 17 m) and FibreOptic cables (up to 10 km) can be used for the transport protocol. The FibreOptic cables for InfiniBand are manufactured differently from today's FibreChannel connections and are far less sensitive. A kink in the cable causes almost no damping and therefore no loss. If you were to do this with today's FibreChannel, the cables would have to be completely replaced. This feature makes InfiniBand Fiber Optic ideal for use in internal machine cabling.

For external use, it is important to advance the FibreOptic solutions both in the support of the serial interfaces and in the distance options.



4, 8 or 12 'physical lanes' can be switched to an InfiniBand 'physical link', the packet transmission is 'bytemultiplexed' via the physical lanes.

InfiniBand represents an extremely high scalable link interface SLI. It is possible to dynamically change and dynamically adjust the number of physical lanes - depending on throughput requirements and data traffic.

InfiniBand offers a highly scalable connection technology with 'nX' multiple lane pairs with a data transfer rate of 2.5 Gbit / s (Single Data Rate SDR), 5 Gbit / s (Double Data Rate DDR) or 10 Gbit / s (Quad Data Rate QDR) in every direction. The table below shows the maximum bandwidth for each of the options:

	SDR	GDR	QDR
1X	2.5 Gb / s	5.0 Gb / s	10.0 Gb / s
4X	10.0 Gb / s	20.0 Gb / s	40.0 Gb / s
8X	20.0 Gb / s	40.0 Gb / s	80.0 Gb / s
12X	30.0 Gb / s	60.0 Gb / s	120.0 Gb / s

InfiniBand has another significant advantage. The Infini BandUpperLevel protocols such as SDP, SRP and iSER use the method of Remote Direct Memory Access RDMA to write data directly into application-related buffer areas of the computer and read them out at the same time.

This reduces overhead data transfer by a third and frees up CPU capacity that can be used differently.

After the first InfiniBand networks have been established in some of the university's high-performance data centers, it can be assumed - analogous to the FibreChannel - that the first InfiniBand networks and infrastructures will be implemented in 'commercial' data centers within the next 2-3 years. Accelerating is also the fact that InfiniBand is cheaper compared to FibreChannel.

New basic technologies

The experts in the storage area agree that

Plastic is the storage medium of the future can be. Experiments with organic polymers and polycrystalline chalcogenide have been going on for several years. Organic polymers are suitable for optical data storage. With the new one **blue laser technology** interference fields can be generated on the plastic that map a large number of bits. This technology is also known as holographic storage and allows much larger capacities than with conventional optical methods. A blue laser is used as the 'data beam'. Part of the laser is sent through a light modulator using mirror processes so that it receives a different light diffraction. Then both beams come together again on the polymer surface. The different light diffraction creates an interference field in the polymer, which is known today as a hologram (see technology appendix under HVD Holographic Versatile Disk, PCM Phase Change Memories and NanoTechnologies).

Since no one can yet estimate how plastics and modified plastics will behave as they age, IBM is working on sensible aging processes for plastics in order to get more precise information. Today it is assumed that a modified plastic should have a life cycle of approximately 100 years. However, this statement still has to be substantiated and confirmed by appropriate aging processes.

Commentary on the era of server-based storage architectures with new infrastructure options

Server-based storage architectures such as DS8000, SVC, new archiving solutions, the new tape virtualization units ProtecTIER and TS7700, TS3500 tape library virtualization and file virtualization via file area networks offer unprecedented flexibility and scaling in performance, capacity and functionality. They will significantly determine the era ahead because consolidations will be possible on a completely new basis. These new architectures make it possible to connect many servers together via clusters. This marks the first beginnings of grid computing in a commercial environment.

New bandwidths in data transmission and new networking technologies are giving rise to new approaches for the optimal design of data centers and their infrastructure. Virtualization solutions in a wide variety of data center areas will lead to a significant simplification of data center processes and greatly simplify administration.

Standards are forcing IT providers to move closer together, and the end user is finally getting more compatibility between the different systems than was previously the case. Virtualization plays a key role here.

New storage technologies will come onto the market in good time to meet the demands of ever-increasing capacities. Plastics as the storage medium of the future will significantly shape this era. Power-independent flash memories in the form of SSDs (solid state disks), phase change memories and storage class memories can also replace today's expensive FC disks if they are produced accordingly cheaper in the next few years.

The pace of development and product cycles will increase even further. It will be increasingly difficult to keep an overview of all options in the storage environment that can be actively used, and advice will be more important than before. Hardware alone plays a minor role.

The era will surely be the most innovative and will easily produce things that we don't even dare to think about today. If you had told somebody in the 90s what is available today, you would surely have been stamped with a fan.

What will be in ten years? It will be an exciting epoch and the author asks himself whether his statements will be correct.

Maybe our imagination is still far behind the things that lie ahead ...

software

IBM storage software

The system-managed memory in the mainframe environment

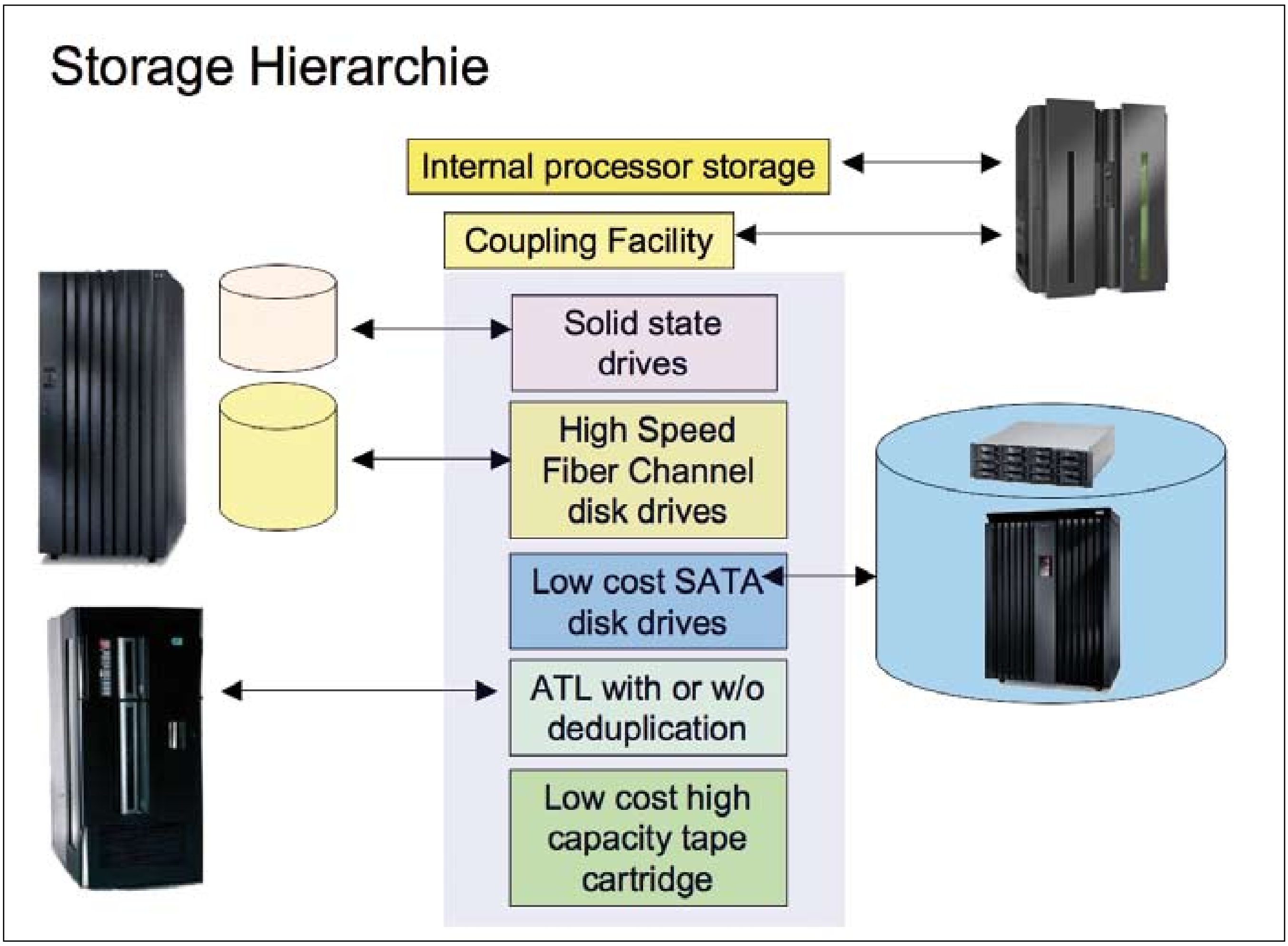
As the name implies, the external storage environment is managed by the system, more precisely by the connected server and its operating system. In this context, managed means that the operating system automatically manages all activities in connection with external storage according to a predefined set of rules - policy-based. Other or similar terms for system-managed storage - System Managed Storage - are Information Lifecycle Management (ILM) or Scaleout File Services (SoFS). Such terms can be found in IT landscapes that are primarily equipped with Unix and / or Intel-based servers. System managed storage, or SMS for short, was massively shaped by the approaches in the mainframe and host environment in the 1980s.

The aim of SMS was to quickly adapt to the rapid growth in storage at that time without increasing the number of storage administrators. SMS behaves - almost - completely elastic to the growing external storage volume, scales excellently and manages a 400GB disk-based storage environment in 1990 just as well as a 4,000,000 GB disk configuration today.

Storage hierarchy

In addition to comparing the logical requirements for a file with the physically existing configuration, an automatic storage management solution is expected to meaningfully store the files in the existing storage hierarchy and move them as required.

Today's disk systems can be equipped with SSDs (Solid State Disks), FibreChannel disks and SATA disks and thus map a storage hierarchy within the system. SSDs are still very expensive, but they offer extremely fast access and extremely fast response times.



The storage hierarchy for magnetic disks is determined by two opposing trends: the faster the memory access and the better and shorter the service time, the more expensive the stored GB. In other words, the stored GB becomes cheaper if the requirements for access time and speed decrease.

A fast FibreChannel disk drive also rotates
15,000 revolutions per minute or 250 revolutions per second, ie one revolution takes 4 ms. The usual capacity of a FibreChannelDisk today is between 300 GB and 650 GB with a form factor of 3.5 inches in diameter. Due to the high number of revolutions, the number of possible accesses is considerably faster compared to an inexpensive, high-capacity SATA plate with 7,200 revolutions per minute or 120 revolutions per second with 8.3 ms per revolution. This takes more than twice as long compared to a FibreChannel plate. The capacities of SATA disks today are 1 TB and 2 TB.

A common size to rank a disk's performance is the number of I / O accesses per second and per GB. This makes it clear that SATA is much weaker than a FibreChannel disk and is therefore unsuitable for highly active files. The technology for 15,000 rpm is also more demanding - and therefore more expensive - than the technology required to rotate at just 7,200 revolutions per minute.

Virtual tape configurations play a special role in the mainframe environment. In order to make the advantage of direct access to the disk accessible for backup and outsourced files, but on the other hand to use the space used economically, a technology is used that avoids the same data structures being saved multiple times (“deduplication”). Real tape drives are completely dispensed with and the backup files and the migrated data sets are saved on disks. Such a system simulates tape drives compared to the connected z / OSServers.

Automatic tape libraries and virtual tape servers, however, continue to use real tape drives and cartridges, on which the files are ultimately stored after the written tape data has passed through an upstream disk buffer. Such a disk buffer can be used to emulate virtual tape drives that simulate many tape drives per virtual tape server. The actual backend consists of only a few real cassette drives, which describe high-capacity cassettes. These real cartridge drives are not visible in a virtual tape server for the z / OS operating system.

Emergence

Storage management software

DFSMS / MVS today sets the standard for storage management software, without which storage growth in the mainframe environment would not have been administrable in recent years.

The beginnings of DFSMS / MVS go back to the early 1980s. At that time, customers and corresponding user organizations, including IBM, recognized the need for fully automated storage management:

- ***Rapid storage growth***
- ***The TCO for storage has already escalated over the years***
  Early 1980
- ***"Storage keeps getting cheaper, but also harder***
  to manage ", Datamation, December 15, 1985 GUIDE / SHARE asked
- ***IBM from 1983–1985 to***
  to offer a "policy based storage management" and defines
  "requirements for futures of storage management", GUIDE 1983

- ***"Without improvements in methods, or the emergence of***
  improved automated managers, the storage management forces will have to grow exponentially as the data grows ", GUIDE Paper, November 15, 1985

- ***Internet, e-business, b-2-b interactions, life sciences, WEB***
  2.0, etc. continue to exacerbate the demand for storage capacity that needs to be managed.

The system-managed memory in the mainframe environment - Werner Bauer

Automatic storage and data management was implemented with DFSMS / MVS and is described in the following sections.

Data facility storage management subsystem, DFSMS / MVS

In 1989, IBM implemented MVS / DPF 3.0, the second stage of a strategy that emerged under the internal term “Jupiter Strategy” in the early 1980s. This 2nd stage is what we understand as SMS today in the host environment.

The Jupiter strategy was then planned in three stages:

1. ***Jupiter Stage I is what is in the DFSMS / MVS or z / OS Interactive Storage Management Facility (ISMF) is understood. ISMF is a panel-based front end to the external storage and the associated repository, in which the external storage is recorded in a special, very effective and efficient database. ISMF was the first step and was announced in 1987 and launched on the market.***

2nd ***Jupiter Stage II was announced in 1989 as MVS / DFP 3.0 and generally available. In addition to the Data Facility Product (DFP), the components Data Set Services (DSS), the Hierarchical Storage Manager (HSM) and the Removable Media Manager (RMM)***

integral part of this Jupiter stage II. That is the state of SMS, as it is known today in the mainframe environment. It takes into account the realization of the idea of automatically managing the memory by the operating system according to a predefined set of rules. This level has been refined and supplemented over the years and is now an integral part of the operating system software under z / OS.

3rd ***Jupiter Stage III was planned as the last stage. At this stage the complete separation of logical data and physical storage should be implemented. The external memory should be standardized to a block-level architecture and the logical data should be automatically placed as specified by rules and service levels. If a rule - or the required service level - can no longer be adhered to, the system itself ensures that the affected blocks are automatically relocated in the memory hierarchy so that they cannot be used, so that the required service level can be maintained again.***

At the beginning of the 1990s, IBM decided to no longer implement this third stage as part of the Jupiter strategy. It was assumed that a corresponding investment should be made for Unix-based servers.

functionality

Files in z / OS

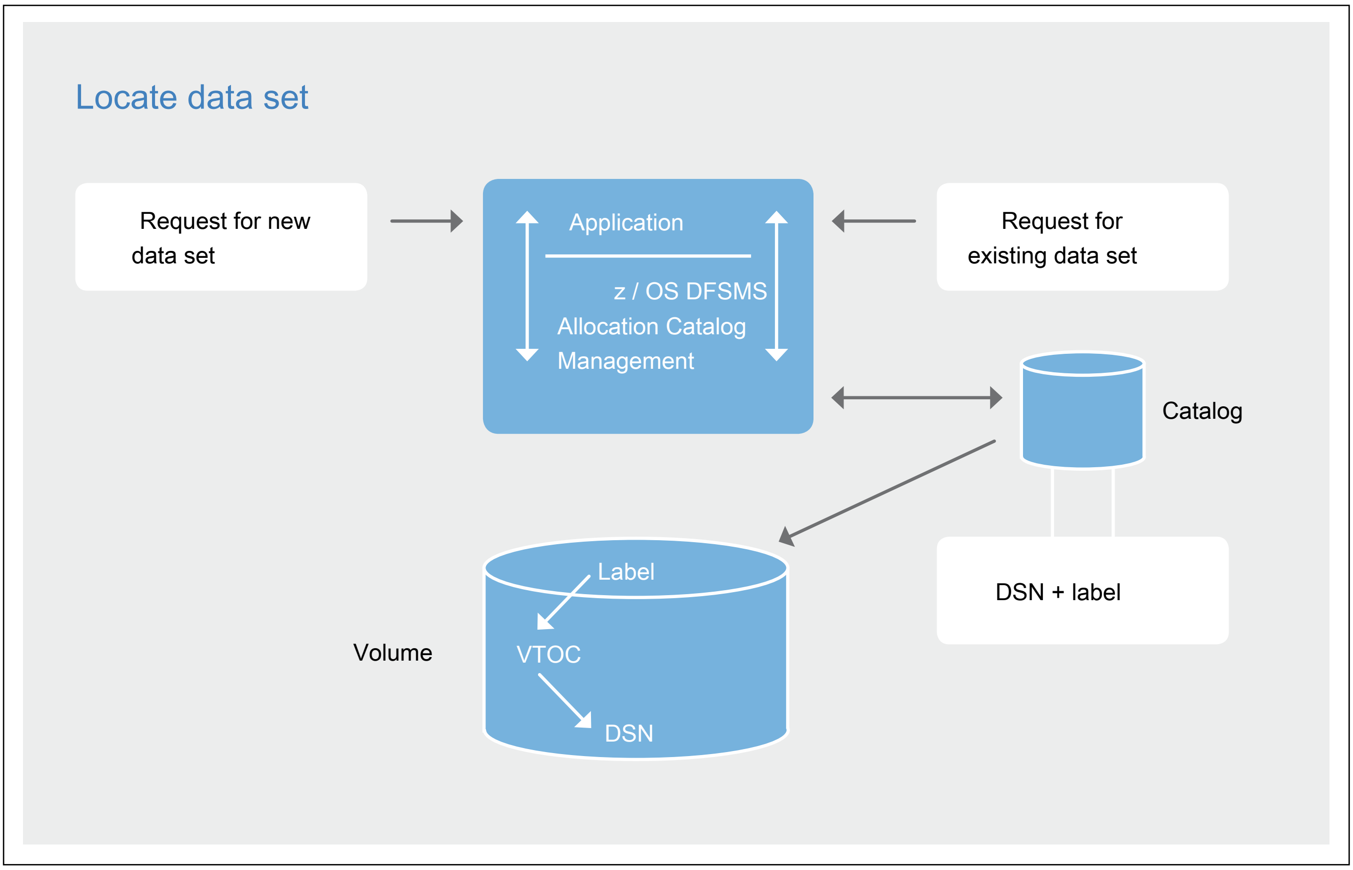
Before we take a closer look at SMS and its components, we want to describe the principle of the file in z / OS and how a file is created and found.

Each file - comparable to a "File" in Unix or Windows - is assigned a unique name or more precisely a file name, a "Data Set Name" (DSN) in z / OS. A DSN can be up to 44 characters long and consists of individual "qualifiers". These qualifiers are between 1 and 8 characters long and consist of alphanumeric characters. As a rule, there are no regulations on how such a name should be pronounced. An example: "SAP.PROD.KDNR1984.INDEX" has the 1st qualifier with SAP, which is also known as a high level qualifier (HLQ), a 2nd qualifier PROD, the 3rd qualifier KDNR1984 and the so-called last level qualifier (LLQ) with INDEX. The qualifiers are separated by a dot.

Such a name can also be broken down and checked by the system. For example, a rule can be defined that reads <WHEN LLQ = & variable_llq THEN action1>. Under & variable_llq, for example, INDEX, VER *, GROUP1 is defined as a possible LLQ list, whereby this LLQ would also be selected with the first 3 characters VER. It would be irrelevant how the other qualifiers are structured. So a file with DSN = TEST. D30M09.VERZ1 would accept the selection as valid because VER * is also in the LLQ list.

With corresponding AND links to other rules, the selection can be more selective, such as <WHEN (LLQ = & variable_llq AND HLQ = & variable_hlq) THEN action2> with SAP *, PROD as a list in the variable & variable_hlq. Then all DSN with SAP in the first 3 digits of the HLQ and with the LLQ as defined in the variable & variable_llq fall into the selection.

A file must always be cataloged in an SMS environment. A catalog is a directory in which each individual file is listed by name and where this file can be found. This is done via a volume label or the so-called volume serial number, VOLSER. The VOLSER is also unique in an SMS system. With hundreds of thousands or millions of files in a large system, you will not only be using a single catalog, but via HLQ to certain catalogs



approach. In our example, all files that identify SAP as HLQ are cataloged in the U.SAP catalog. Since the catalogs must also be found via a DSN, there is a higher-level instance called the “Master Catalog”. There is only one active master catalog in an SMS system. Its content is stored resident in the z / OS, since it has to be pulled up for each “Data Set Locate” in order to attract the associated [user] catalog. This is done via the HLQ and is referred to as the ALIAS relationship. SAP is entered as HLQ in the master catalog and the reference to the user catalog U.SAP. Whenever a DSN appears that identifies SAP as HLQ, this file is defined or searched for in the U.SAP catalog.

Since the catalog only refers to a volume on which the file is or is stored, each volume also has a fixed structure so that the file can ultimately be found at the volume level or entered accordingly. The “Volume Label” is always on the same place of each volume, Cylinder 1, Track 1, Record 1. The label also has a pointer to the table of contents of the volume, the so-called Volume Table Of Content or VTOC. The VTOC is roughly comparable to a file system in Unix or Windows. The VTOC contains all DSNs that have a home on this volume and the start and end address of this file on the volume.

A file or DSN is therefore a unique name within an SMS configuration and the associated file must always be cataloged. The request to create a new file goes through the SMS - more precisely through the DFSMS, allocation and catalog management - and looks for a suitable place in the storage hierarchy, depending on the required service level that is specified for this new file. A request for an existing file only goes through allocation and catalog management in order to localize the existing file and make it accessible to the application.

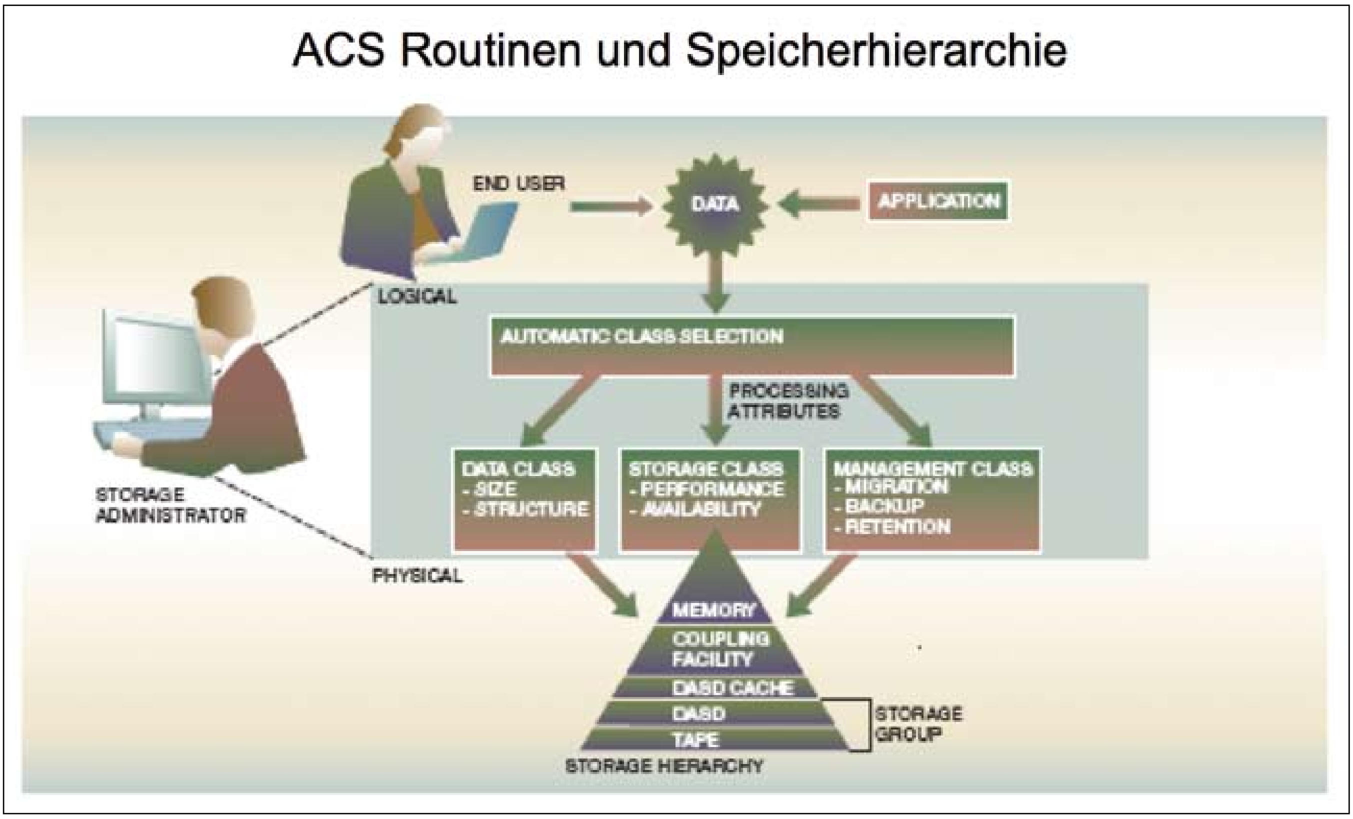
Overview of policy-based storage management

Policy-based storage management takes over the tasks of the storage administrator that was previously carried out manually. DFSMS allows the logical view of the data to be separated from the physical memory configuration. Logically here is synonymous with Service Level Agreements (SLA) and physically means where the data is actually stored. Logical view in SMS terminology means: data class, storage class and management class. The storage group ultimately specifies the physical storage space.

- ***The data class contains information about the storage required and the data attributes made for a file. The storage class***
- ***contains information on the required Performance and response time and the level of availability for the affected files. This includes placement on a physical volume with fast access time or that a file is automatically distributed to several volumes using the "sequential data striping" attribute. The management class contains information on the number and***
- ***Retention of backups for a file. Here you can also define how long the file itself remains before it is automatically deleted. The storage group defines the potentially physical***
- ***Storage space for the affected file. A storage group is a group of volumes in disk storage subsystems or a group of tape cartridges and tape drives. The affected volumes are managed collectively by DFSMS.***

The decisive factor is the assignment of these attributes, which are defined in these classes or "constructs". This task is performed by so-called “Automatic Class Selection Routines” (ACS routines). Based on criteria such as DSN, job name, size of the file and approx. Another 25 criteria, appropriate classes are assigned to each new file. There is such an ACS routine for each class.

The overview ”ACS routines and storage hierarchy” shows the role of the ACS routines. Based on these logical classes, the system searches for the optimal storage space for the affected file in the existing storage hierarchy.



The selection of the space in the storage hierarchy is called "Volume Selection". DFSMS goes through a number of steps and creates a list of potential volumes, a so-called “candidate disk volumes” list, which can be used to save the new file there. This volume list is based on the selection of the corresponding storage groups as assigned in the Storage Group ACS routine.

The so-called “primary candidate” list contains volumes that meet all the criteria specified in the corresponding storage class and data class. There is also a "secondary candidate volumes" list, which contains all the volumes that do not meet all criteria, but were also assigned in the Storage Group ACS routine.

If this “primary candidate” list is not empty, it is passed on to the z / OS component System Resource Manager (SRM). SRM then selects one or more volumes from this list. Criteria for SRM are e.g. B. the shortest waiting time for an I / O on the concerned Volume or the volume is not yet in use for the current job. If the primary list is empty, DFSMS tries to select one or more volumes directly from the secondary list without going through the SRM. The most important criterion is the available free space.

After one or more volumes have been selected, the file is created on the volume concerned (volume allocation) and a corresponding entry in the catalog is created.

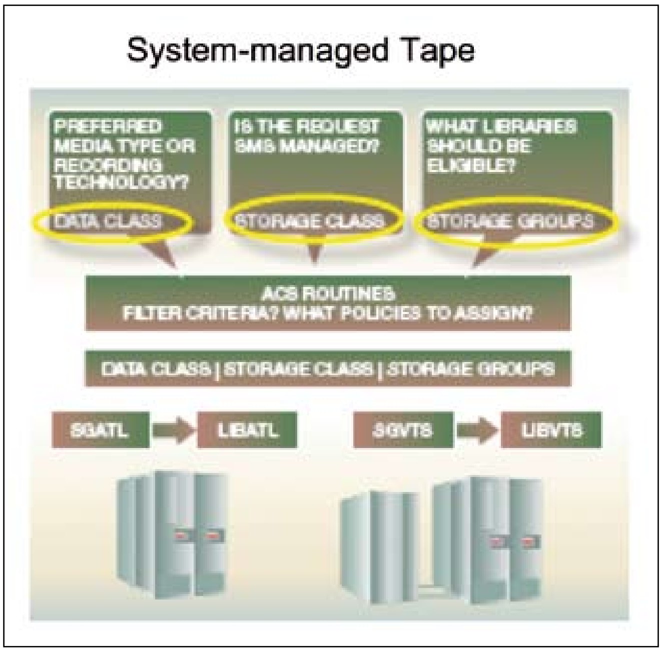
If the allocation is not successful, the volume selection is repeated until either volumes are found to create the file there, or the process terminates with a corresponding message if no eligible volume can be found.

SMS for tape

Policy-based storage management for tape libraries
DFSMS not only manages the files in a disk storage subsystem, but also the data and units in a tape library. This applies to automatic tape libraries with robot access to the cassettes and storage spaces or to so-called “manual tape libraries” with manual loading of the cassettes into a tape drive. In this approach, DFSMS groups the cartridges and the associated cartridge drives in a corresponding tape storage group.

System-managed tape manages the tape libraries, the tape drives and the tape cartridges themselves, but not the individual files that are on these tape cartridges. This function is guaranteed by DFSMSrmm (Removable Media Manager) by keeping a directory of which file is on which cartridge. The cartridges also have a label that is similar to the disk volumes.

The “System Managed Tape” illustration shows how a specific tape library is assigned for a specific file on tape using the ACS routines.



The policy-based storage management for files on tapes can also be outsourced to the affected libraries themselves via an "outboard" management. DFSMS assigns the constructs Data Class, Storage Class and Management Class from a tape file to the tape subsystem and transfers these construct names to the library. The form of the "Constructs" is defined in the library itself and then executed accordingly - in contrast to the central approach of DFSMS in the host.

DFSMSrmm manages the actual files on the cassettes. Here too, attributes from the management class are used accordingly, such as: B. Lifetime and automatic deletion of files if the files have not been referenced for a defined period of time.

Components and functions

Components
The figure "z / OS components for centralized storage and data management" shows the system components in z / OS that are responsible for the allocation, i.e. the placement in the storage hierarchy. It also shows the components that are used for space management and availability.

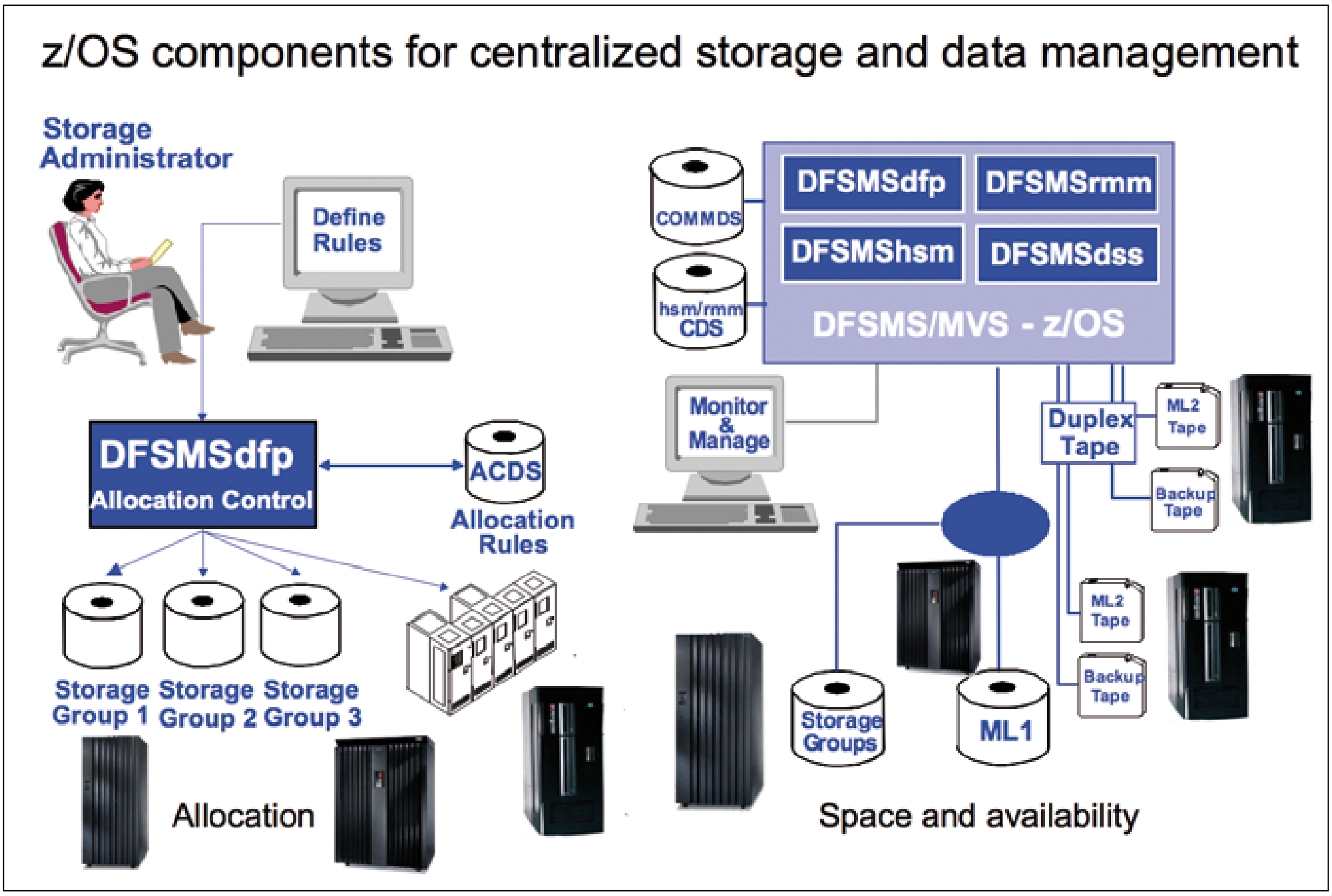
In addition to the DFSMSdfp, the z / OS SRM is also used for "data set allocation" when a file is newly created.

All four components are linked to each other for space management and availability management. Each component has to fulfill a specific order within SMS:

1. DFSMSdfp - creates the interface to the policies and instructs the other components
2. DFSMSdss - is a kind of data mover and moves almost all files

3. DFSMShsm - is responsible for space and availability management. DFSMShsm determines which file is copied for backup purposes or creates free space by swapping out and migrating inactive files.

4. DFSMSrmm - is responsible for the recording of the cassette contents or for the files on tape.



DFSMS works with two repositories, the Active Control Data Set (ACDS) and the Communication Data Set (COMMDS). DFSMShsm and DFSMSrmm each have their own repositories, which are supplemented and compared with the DFSMS files via DFSMSdfp.

Functions

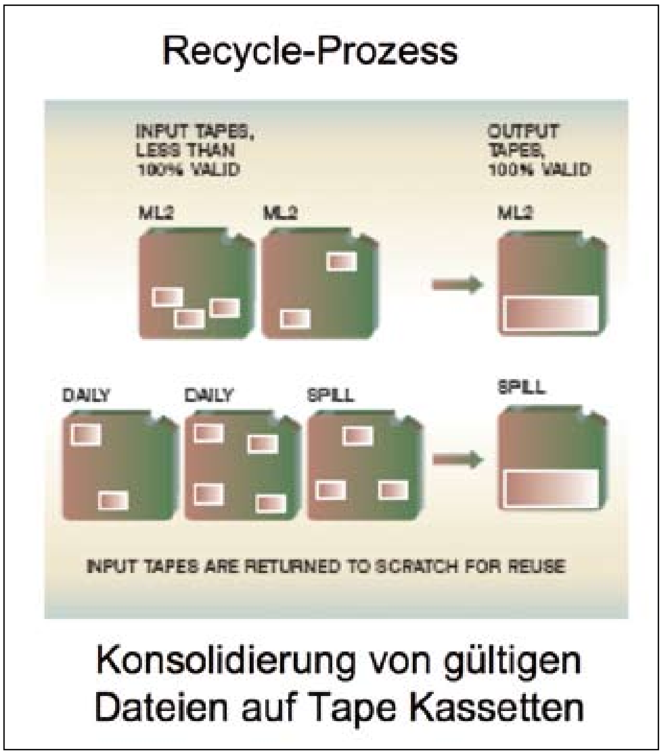
Space management

Using the SMS constructs and here especially the management class, service level objectives are specified with regard to the available space on disk storage subsystems or the volumes in such a subsystem. The goal is to prevent applications from aborting because of a problem with insufficient space.

DFSMShsm is the DFSMS component that monitors the policies for space management and takes necessary actions when the policies are no longer adhered to. In addition to the management class, information from the storage groups is also used to determine the space on all managed volumes in the disk storage subsystems at FileLevel. The criteria are so-called thresholds at volume level and at storage group level, which must be observed in order to always have enough free space for new files. This migration process is periodic and inactive files are identified as candidates for moving to another class in the DFSMShsm storage hierarchy.

Even if a file is moved from a volume to a disk / tape storage hierarchy managed by DFSMShsm, the file remains cataloged. If such a file is referenced by the application, DFSMShsm automatically performs a recall and brings the file back to an active application volume. This process is completely transparent and unnoticed for the application concerned or the user.

DFSMShsm also automatically deletes files whose lifespan has expired according to their assigned management class (deleting expired data sets). However, temporary files will also be deleted or there will be space that is not used with data within a file if this is required as an attribute by the associated management class.



Availability management

Availability Management manages the backup copies of active files that are on active disk storage volumes, also known as L0 volumes. The goal is that a backup copy can be used if the original file is faulty or damaged - or even an entire volume is no longer readable. DFSMShsm uses the policy information in the management class and information from the storage group to automatically and periodically create backup copies of the original files. These backups can also be initiated using appropriate commands.

Consolidate valid files on tape cartridges

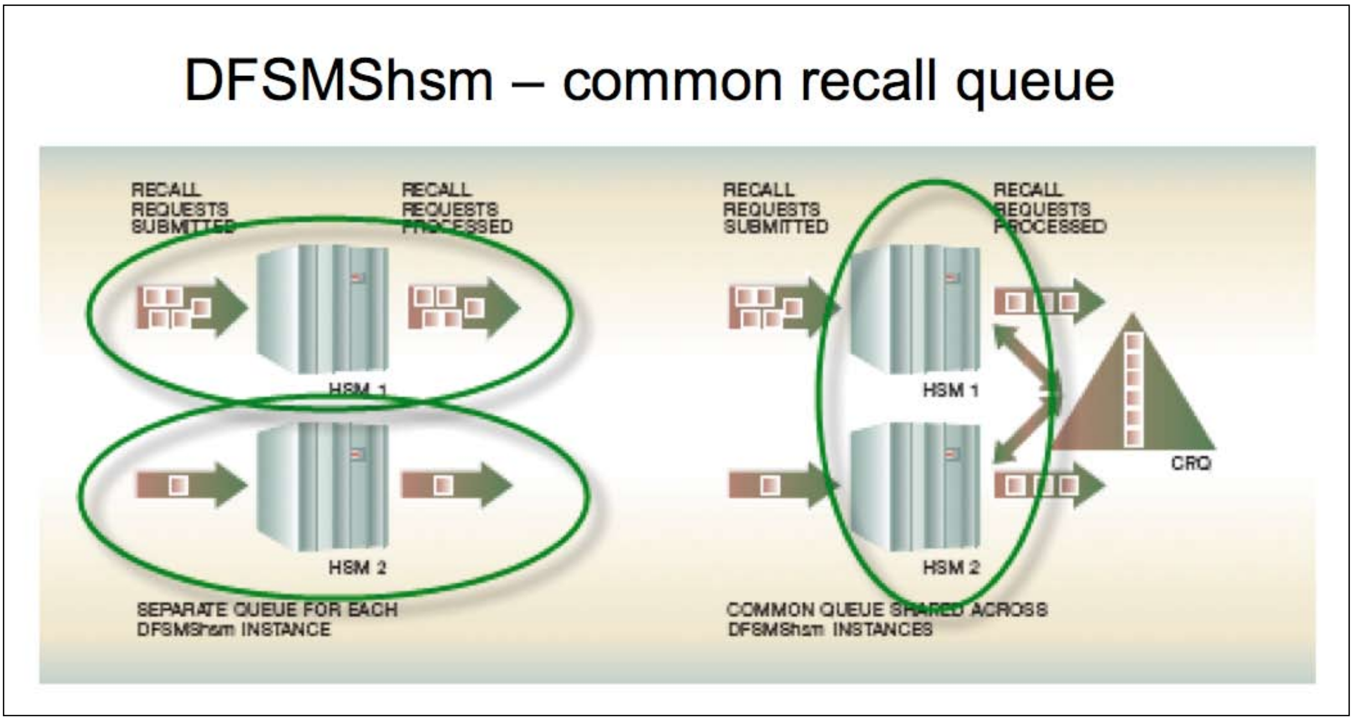
The management class stipulates policies on how and when backup copies are to be made or when original files are to be outsourced. The management class also specifies when a file should be deleted or how many backup versions of an original file should be kept and for how long. Hundreds of files about space management or availability management have often been written on a tape cassette and can be deleted at different times.

The illustration "Consolidation of valid files on tape cassettes" shows how a "Reclamation" or "Recycle" process removes backups that are no longer valid or files that are no longer needed and have been removed.

New tape cassettes are created, which are then rewritten with valid files. The thus released tape cassettes are returned to a scratch pool and are thus available again for new tape requests.

DFSMShsm Common Recall Queue

The decisive factor for the performance of DFSMShsm is the intelligent use of functions and services that z / OS and Parallel Sysplex provide. The performance of DFSMShsm is particularly critical when many files have to be recalled from the DFSMShsm-managed repository (using the Recall function of DFSMShsm). If an application or user references a file that has been outsourced to the DFSMShsm-controlled storage hierarchy as part of space management, the application must wait until DFSMShsm has recalled the file to the ApplicationVolume (L0) (Recall). If this waiting time is minimized and the throughput in the system is increased, this contributes to a more efficient use of the entire system. DFSMShsm therefore optimizes this recall challenge and uses the coupling facility as a repository for a common queue of recall requests. Communication from the DFSMShsm systems within a sysplex takes place via XFC technology, in which the coupling facility (CrossSystem Coupling Facility) practically represents a fast and shared semiconductor memory.



A typical parallel sysplex configuration consists of a DFSMShsm instance in every z / OS image within such a sysplex. The recall requirements can be very different in the various z / OS images. For example, the HSM1 in the figure "DFSMShsm - common recall queue" has to process significantly more recall requests than the HSM2 in the second system. Each DFSMShsm processes the requirements in its local queue and knows nothing about the requirements in the other system.

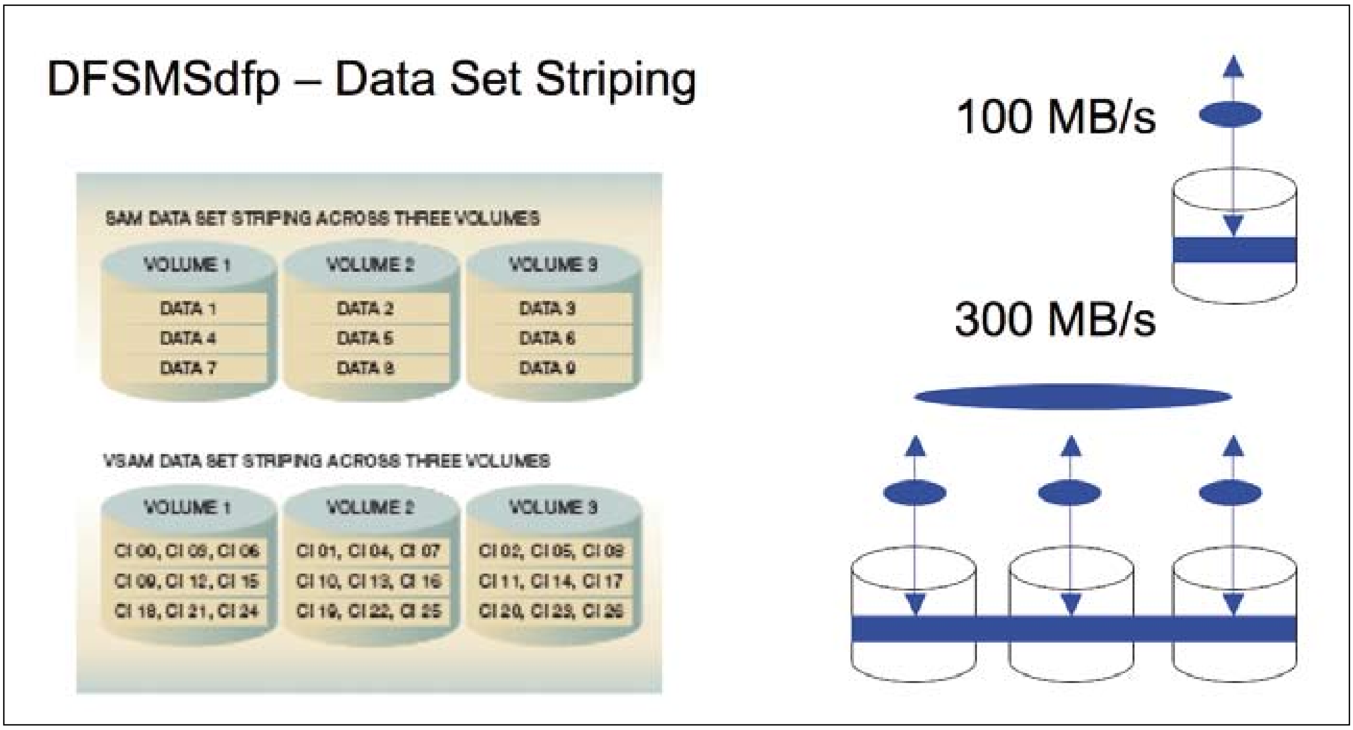
ParallelSysplex-wide load balancing became possible with the introduction of a common recall queue.

In addition to other positive side effects, it is obvious that this can improve overall throughput. Side effects concern a better consideration of priorities, more efficient use of tape resources, e.g. For example, all requests are sorted by tape cassette, and greater flexibility with a Sysplex configuration.

DFSMSdfp and data striping

Another example of a variety of functions and services provided by DFSMS is data striping. Instead of placing a file on a single volume, a file is spread over several volumes or “striped”. The resulting improvement in the combined input / output contributes to the fact that the “single file” throughput can be increased accordingly. This can significantly reduce the dwell time of applications and jobs with high demands on the throughput of sequential I / Os. This also applies to a certain extent to random I / Os (Random I / Os) if the application can issue and process I / Os in parallel. This is the rule for database subsystems such as DB2 or IMSDB.

The figure ”DFSMSdfp - Data Set Striping“ illustrates this striping effect for the sequential throughput.



The right side of the figure symbolizes the effect of data striping. In the case of the file on a volume, the maximum sequential throughput of the access path is 100 MB / s. If an attribute is given for this file via the storage class, which says that the sequential throughput must be 300 MB / s, then DFSMS would automatically ensure that this file is distributed to three volumes when this file is created, if each volume can be used at the speed of 100 MB / s. DFSMS or z / OS knows exactly the possibilities of each individual volume and uses this information for the automatic selection of the volumes if a corresponding new file is to be created.

The z / OS system with its integrated DFSMS automatically reacts to required service levels and carries out corresponding actions regarding data distribution and optimization of the memory configuration found. The system-managed memory is the highest level of perfection in automatic data management and data storage in the z / OS environment.

Summary

DFSMS or system-managed storage is a software-based solution for z / OS. It has been continuously improved since the announcement in 1989 and automatically takes advantage of new opportunities in modern servers and storage technology. This chapter gives just a few examples of how DFSMS helps to ensure and continuously improve the increasing requirements for availability and performance in modern data centers.

DFSMS also serves as a model for similar developments in the OpenSystems environments with Unix and Intel-based operating systems.

Tivoli Storage Manager - the success story of the first 20 years - Klemens Poschke

How it all started… 1983–1993

When you consider that large parts of the architecture used today in Tivoli Storage Manager were designed 20 years ago, the performance of the developers at that time cannot be overestimated. The architecture elements “forever incre mental”, the use of a transactional meta database and the “tiered storage” concept shape the overall success of the TSM solution particularly sustainably. These components are still effective differentiation criteria from other offers on the market today.

From the early 1980s, the original ideas for a data backup solution came from developments in Almaden by Mark Brown, Robert M. Rees and Michael D. Penner on the VM / 370 platform. They developed an application for internal use called CMSback that allowed single files from CMSMinidisks [1]

to incrementally save and restore in the VM. At that time, CMSMinidisks could only be saved and reconstructed as a dump volume using DDR (Data Dump Restore).

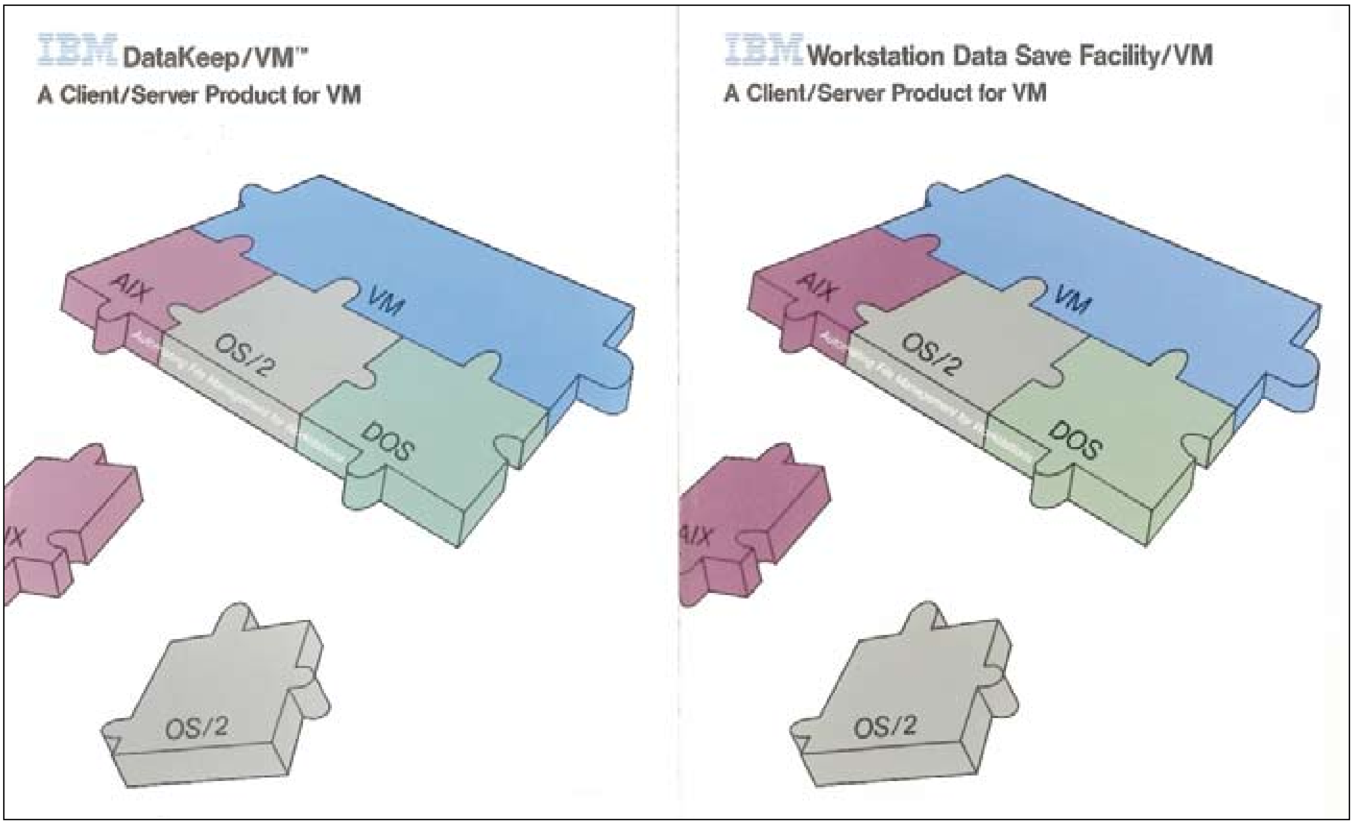
Bob Rees describes the application components of CMSback as follows:

“In CMSback we explored a crawling model where:

- *a central process crawled user 'file systems' (VM minidisks) looking for files that have been changed or deleted a framework for*
- *supporting tape drives together with disk*
- *a restore GUI that could communicate with a central server to allow users to restore their own files. ”*

As a result, an internal project with names was created **AMORE (A** grubs **M** ulti **O** perating system backup and **RE** store), which took up a number of ideas from CMSback, RDBMS research projects and other internal tools. These initiatives ultimately led to a formal project in the late 1980s **Datakeep / VM** - a joint development between the Almaden and Endicott laboratories. At the same time, there was another development on this topic in Tucson / Boulder called ESMS.

[1] CMS = Cambridge Monitoring System - the realization of a virtualized single-user system in the mainframe operating system VM, in which a user can work interactively in an independent system environment and "minidisks" are available as storage resources that he can use with any data or programs or can share with others - analogous to the possibilities that a PC / laptop user or a VMware user also has today.



Announcement brochures DataKeep / VM (name discarded) and WDSF / VM

In the first pilot projects, the DataKeep / VM solution proved to be more stable than ESMS and led to the announcement of the WDSF / VMSoftware product (Workstation DataSave Facility Release 1) on November 1, 1990. The change in name from “DataKeep / VM” to “WDSF / VM” was necessary due to reservations by the IBM legal department shortly before the announcement, even though the entire product documentation had already been printed.

WDSF / VM Release 1 was IBM's first Cliet / Server application for backing up PC / workstation data from the MSDOS, OS / 2 and AIX operating systems. The product was designed according to the principles of transaction-based operation, rule-controlled data storage, a purely incremental data backup cycle for file system data and the use of hierarchical storage pools for the combined use of different storage technologies (DisktoDisk or DisktoDisktoTape).

The development team has also set a monument in the depths of the TSM server. If you issue the undocumented "show developers" command in the Admin Commandline of the TSM server, you will get the protagonists listed (see screenshot below).

The metadata management was implemented in the WDSF / VMServer component in a database that used the basics of the ARIES architecture. ARIES = Algorithms for Recovery and Isolation Exploiting Semantics was among others by IBMFellow Dr. Mohan from Almaden was developed. The principles of the "writeahead logging based recovery method" defined in ARIES are still the basis of TSM and a large number of other IBM applications (DB2, MQSeries, Domino etc.) and the basis of relational database systems in general.

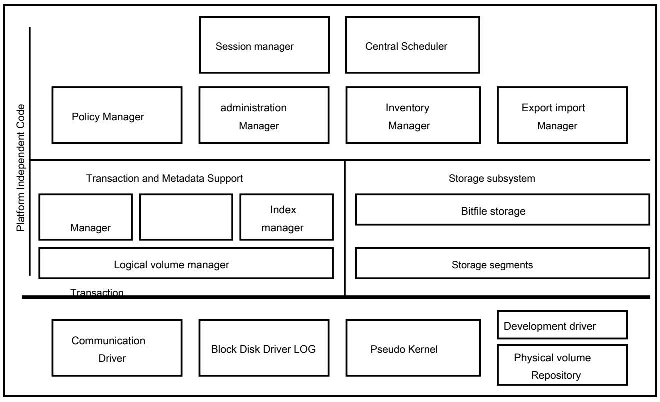
The implementation was based on a hierarchical database architecture (B + Tree), which fulfilled the claim to be executable on all operating systems of the planned server platforms, which DB2 did not fulfill at the time. Only now - almost 20 years later - has the managed metadata volume in our customers' TSM servers reached sizes that can no longer be processed scalably with the previous database architecture. That is why TSM in its current version 6.1 has been switched to DB2 as database technology.

The concept also provided foresight to be able to integrate the functions of data backup into general systems management solutions, e.g. by providing a comprehensive management console and interfaces for creating reports, collecting accounting information and forwarding error conditions.

The developers set a memorial - “show developers” command in the TSM server

The main design principles of the new data backup software were:

- *Enable the server to preserve the physical locality that data has at its source even when clients deliver this data over long periods of time.*
- *Estimate the target workload and develop the system to accommodate them.*
- *Deploy the system in many computing platforms.*
- *Provide lights-out, unattended operation with unattended recovery from failures. Provide*
- *continuous operation*
- *Accommodate the peculiarities of a wide variety of storage devices*
- *Manage the storage system through user controlled policies*
- *Minimize the periods of time in which an entity in the system cannot be retrieved*
- *Use additional temporary storage space to gain concurrency and to increase the availability of user data Source: "Applying*
- *Database Technology in the ADSM Mass Storage System "by Luis-Felipe Cabrera, Robert Rees and Wayne Hineman IBM Almaden Research Center 1995.*



ADSM server: Principal Software Components

Here is another excerpt from the same publication with a structural representation of the application components and a description of the interaction of the modules in "incremental backup":

*For incremental backup, the session begins by the client requesting from the server the corresponding policies from the **Policy Manager.** Policies are stored in system catalogs administered by the **Index manager.** The client then builds the candidate list of files for backup and requests the server to send, from the*

***Inventory manager,** the latest information for each of the candidate files. The Inventory Manager stores all its information in catalogs administered by the Index Manager. The client then sends to the server, optionally compressing them, only those user files that have changed or have been created since the last incremental backup. The server constructs appropriate entries in the Inventory Manager and appropriate bitfile using the*

Bitfile storage** component. Bitfiles are stored using the **Storage segment manager.

At the end of sending all the user files the client also sends the server the list of user files that have been deleted since the last incremental backup allowing the server to mark them and, eventually, to expire them from the system. The client then commits the transaction and disconnects from the server […]

[…] Database technology has been applied before to distributed systems, to file systems, operating systems, to message queuing systems and to network I / O subsystems.

***To our knowledge we are the first ones to apply trans actions to a mass storage system that administers a storage hierarchy. The fundamental difference with all other systems is the need to support high degrees of con currency for a variety of transactions in the presence of devices with enormous latency times.** A second difference is that our server can replicate its metadata with up to three copies. No other backup system or file system we know of has this function […].*

Extract from the WDSF / VM notice of announcement on November 1, 1990:

Workstation Data Save Facility / VM turns VM host into a server

WDSF / VM Release 1 is a client-server licensed program for backing up, archiving, and restoring disk data files. It enables a VM system to act as a server for workstations and personal computers, and moves the responsibilities of back-up archive media management to VM, while providing client applications to several workstation-based environments including MS-DOS, OS / 2 Extended, AIX / RT and AIX version 3.

It consists of a server program that runs under VM and one of the above environments, and while the archive function is userinitiated, the back-up function may be automated through scheduling. Backups may be performed from either end-user workstations or from workstation-based servers. Incremental backups may be run as background tasks under OS / 2 and AIX, and when files are backed to VM-attached disks, older and larger files are automatically moved to tape. Archival functions enable the use of host storage resources as a off-line storage pool and WDSF / VM uses TCP / IP, NetBIOS or 3270 data stream support for communications.

There is an optional facility for compressing data in the workstation client program. It operates on any System 370 or System 390 processor and requires a 1,600bpi or 6,250bpi tape drive, a 3480 or 3490 magnetic tape subsystem, 10 cylinders of 3380 disk space and 6Mb of virtual storage.

[…] WDSF / VM will be available March 29, and AIX 3 Client, September 27, 1991

The first WDSF / VM version included the client support of the then standard IBM PC / workstation platforms OS / 2, AIX / RT, AIX V3 and MSDOS with the connection via the LAN communication protocols TCP / IP, NetBIOS and 3270 LU2.0. In addition, the protocols SNA LU 6.2, IPX / SPX, CLIOS and Named Pipes were added. This variety of protocols was supported until TSM version 4.1 in 2001. Today TSM only supports TCP / IP, SNMP for distributed environments and shared memory and named pipe protocol for client / server solutions on the same physical system.

A first enhancement was the announcement of WDSF / VM Release 1 Enhanced in November 1991 with the availability of additional clients on SunOS and Apple MacOS. This was the beginning of the era of far-reaching support for other OpenSystem operating system platforms. In the heyday, the client support list had around 40 system environments.

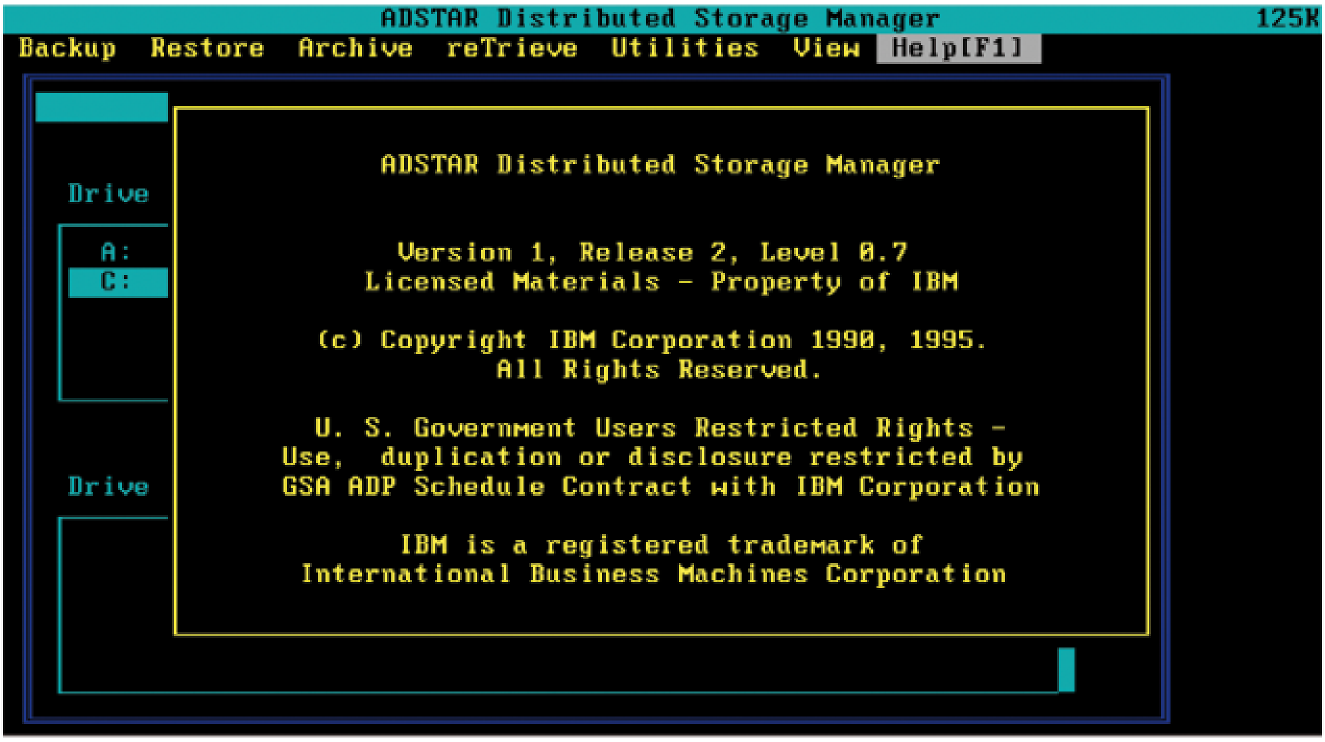
Workstation Data Save Facility / VM Release 1 Enhanced

Workstation Data Save Facility / VM is enhanced to provide support for Sun Microsystems Inc's SunOS Version 4.0.1 and Version 4.1.1, and Apple Computer Inc MacOS Version 6.0.7 and System 7 clients.

AIX Version 3 for RS / 6000 client will be available September 27, while WDSF / VM Programmer's Reference, SunOS Client, and MacOS Client will ship June 26, 1992.

The general trend towards distributed systems was the reason for the next step that Tivoli Storage Manager took in its product history. In 1992, Chip Coy and Bill Fischofer's application in Endicott was ported to MVS and renamed to DFDSM for the VM and MVS platforms - the new data protection module for workstations had thus become a member of the S / 370 DFxxx storage software family.

In parallel, Norm Pass (RINGMASTER project) and Mike Kaczmarski and their teams in California drove porting to AIX and OS / 2 as servers. The new DFDSM product was also shown at the COMDEX in Las Vegas in 1992, albeit with a somewhat questionable focus, as the developer Barry Fruchtman describes as a trade fair participant.



ADSM V1.2 MS-DOS client initial screen

***Barry Fruchtman:** "One item that may be interesting is that IBM Storage made an appearance at COMDEX in 1992 in Las Vegas with a demonstration of the DFDSM. We demonstrated, AIX, OS / 2, and DOS clients backing up and restoring to VM, MVS, AIX and OS / 2 servers; **The OS / 2 clients were running a video at the same time that they were in a loop backing up and restoring.** For the few people which looked at the DFDSM sign and stopped to watch, it was an amazing demo, given the capabilities of the time. It was located in **the multime dia area** of COMDEX, however, so there was a mismatch with the intended audience ... ".*

The anecdote documents with a wink that at that time the IBM trade fair planners subordinated the importance of workstation data backup to the fact that a multimedia video was running on the OS / 2PC at the same time.

As can be seen from the eyewitness report, the first internal porting of the DFDSMServer component to the AIX and OS / 2 platforms already existed in 1992, but only a year later became available as commercial products under the new name ADSM V1.2.

During this time there were other developments and products that used VM and MVS as a server platform for data services of the open system workstations:

- ***WLFS / VM - Workstation LAN File Services (announcement December 1992 - later renamed to LAN File Services / ESA due to availability on MVS) - an application that connected an OS / 2 LAN server to the VM server via an S / 370 BMPX-Bus / Tag adapter card and transparent CMS minidisks made available as drive resources for the OS / 2 LAN server. This was done with the idea of synergistically combining the reliability of mainframe data storage with the new interactive and graphic possibilities of the workstations.***

- ***LANRES / VM and LANRES / MVS (LAN Resource Extension and Services 1992) - the analog solution to WLFS / VM for Novell Netware file server with connection either via a BMPX channel card, TCP / IP or an SNA LU6.2 communication link. Not only file services were supported, but also bi-directional printing.***

At the time, all of these solutions were seen in the computer press as an attempt to prevent the impending demise of the mainframe computer as a central application server by using it as an open system server resource.

The fact that the use of the mainframe as a server still exists today for a large number of business-critical and cross-platform applications (e.g. SAP, Domino, ContentManager etc.) and services (such as TSM) would have made the "industry experts" at that time sustainable in their worldview shocked.


COMPUTERWOCHE

Newsletter | Suchen

Home | Management | Hardware | Software | Netzwerke | Mittelstand | Secur

Big Blue sucht Lebenselexier fuer Mainframe-Saurier

Mit Lanres sollen Mainframes als Server im LAN ueberleben

22.01.1993 Drucken | Empfehlen | PDF | Merken | ShareThis

Waren Funktionen wie Druckersteuerung und zentrale LAN-Verwaltung frueher nur in Lanres/VM verfuegbar, so bietet das MVS-Release 1.2 diese Features nun auch.

Nach Meinung von Branchenkennern sind die neuen Versionen ein typisches Beispiel dafuer, wie Big Blue versuche, den Mainframe in Zeiten der allgemeinen Migration zu PC-Systemen am Leben zu erhalten.

"Einst primaer als leistungstarke Anwendungsrechner gedacht, werden Mainframes nun immer mehr als Server eingesetzt", kommentiert Gerry Machi, Vice President von Novells Netware Group, die Entwicklung.

CW article (excerpt) from January 22, 1993 announcing LANRES (Source: http://www.computerwoche.de/heftarchiv/1993/4/1125617/)

Departure into the open system world

ADSM V1.1 - July 29, 1993 and ADSM V1.2 November 16, 1993

The name DFDSM, which is common in mainframe circles, could not be used for the target market of the Open Systems, because during this time the polarization between supporters of mainframe systems and Open Distributed Systems was already taking on class-like dimensions and the product success was feared.

Shortly after the availability of DFDSM, IBM decided internally to port the server component to PC / workstation operating systems and to change the name to ADSM (ADSTAR Distributed Storage Manager).


COMPANY NEWS; New I.B.M. Name For Storage Unit

The International Business Machines Corporation said it had changed the name of its data-storage products business to Adstar, which stands for advanced storage and retrieval. The unit, based in San Jose, Calif., produces disk and tape drives, optical storage devices, control devices and storage-management software. I.B.M.'s revenue from storage devices last year was more than $8 billion.

"The new name is another step in establishing our identity as an independent business operation within the framework of the I.B.M. corporation," Ray S. Abu Zayyad, an I.B.M. vice president and general manager of Adstar, said. "We plan to continue I.B.M.'s long tradition of offering world-class storage solutions, and we want Adstar to be synonymous with such solutions," he said.

AdStaR announcement - New York Times 1992 (Source: http://www.nytimes.com/1992/03/20/business/company-news-newibm-name-for-storage-unit.html)

ADSTAR

As described in the article, AdStaR is the acronym of " **AD** vanced **ST** orage **A** nd **R** etrieval "and was introduced as a brand name for the IBM Storage Systems Division (IBM ADSTAR Storage Systems Division). The reason for this renaming was the mind games of the then IBM management to spin off IBM divisions as independent parts of the company and individualize them with their own trademarks (see also MagStaR). However, these plans were not subsequently implemented, but the ADSM trademark remained until 1999.

The "Safeman" logo was designed by the former IBM graphic designer Mike Welkley, who describes the story of its creation as follows:

Here'sa little history on how the Safeman came about ... I was working on the ADSM GUI back in 1993. Periodically we would bring in prospective customers to test the product and the software team asked if I would create a mouse pad that could be given to the customers as a complimentary gift. Given a blank canvas to create whatever I wanted for the design, I decided to create a cartoon character that represented ADSM Storage.

Brand name and identification -
The original of the first ADSM "Safeman" logo

The mouse pad with the ADSM Safeman character on it, for whatever reason, never got printed. However, the ADSM Marketing team somehow got hold of the mouse pad design and wanted to print the character on buttons as a tradeshow giveaway. This was the first use of the character.

The ADSM V1.1 announcement consisted of two independent components, the storage management functions for data backup / archiving and the data access section **THE (** Data Access Services), which actually had nothing to do with data backup or storage. Here is an extract from the letter of announcement:

Announcement Letter Number 293-382 dated July 29, 1993 US - Last Revised on July 29, 1993

Brief Description of Announcement, Charges, and Availability

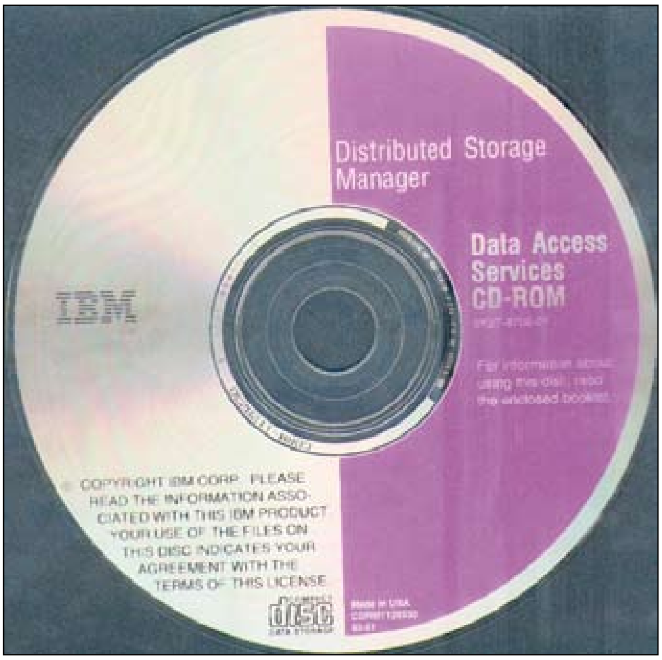
*IBM announces the availability of ADSTAR (TM) Distributed Storage Manager Version 1 Release 1 *. The ADSTAR Distributed Storage Manager provides storage management and data access capabilities. The storage management capabilities provide an automated, highly-reliable, high-performance, network-based backup and archive product for workstations and local area network (LAN) file servers. It consists of an MVS- or VM-based backup / archive server and backup / archive clients for DOS, OS / 2 (R), AIX (R) for RISC System / 6000 (R), Microsoft Windows, Apple Macintosh, SPARC / Solaris, HP-UX, and Novell NetWare systems. The data access services (DAS) provide Distributed FileManager for OS / 2 Version 2 and the Record Level Input / Output (RLIO) API. Distributed FileManager provides local or remote record level access for OS / 2 Version 2 applications.*

ADSTAR Distributed Storage Manager Version 1 Release 1 is the successor product of IBM Workstation Data Save Facility / VM (5684122) and supports existing WDSF / VM backup / archive clients.

The "Data Access Services (DAS)" were a combination of the "Distributed File Manager (DFM)" for OS / 2 V2 and the access API RLIO (Record Level I / O). It was therefore an application middleware which, according to the specifications of the "IBM System Application Architecture (SAA)", could have different application parts communicated with each other on different operating systems, provided the required interface compatibility was available (DDM / LU6.2 2).

In version ADSM V1.1 there were a number of applications that were tested for data backup with ADSM:

- ***DB2 / 6000 provides an online backup of log and data to a remote MVS or VM system using ADSTAR Distributed Storage Manager.***
- ***IBM LAN File Services / ESA [formerly WLFS / VM d. A] Customers using LAN File Services / ESA to distribute data objects will now be able to protect their data with ADSTAR Distributed Storage Manager.***
- ***IBM DFSMS / VM uses ADSTAR Distributed Storage Manager to provide migration support for its SMS-managed storage hierarchy.***



Data Access Services CD - additional component for ADSM V1.1

- ***The Norton Backup for Windows. Windows V3.1 users will now be able to backup their data to ADSTAR Distributed Storage Manager using the popular Norton Backup for Windows interface.***
- ***Additionally, IBM has tested the ADSTAR Distributed Storage Manager backup / archive clients with the following database products for support of offline database backup and archive***
  - ***IBM DB2 OS / 2 Version 1***
  - ***IBM DB2 AIX / 6000 version 1***
  - ***SYBASE 4.5 and 4.9 (except for raw partition support)***
  - ***INFORMIX-OnLine for AIX / 6000***
  - ***Borland dBASE IV 1.5 for DOS***
  - ***Borland PARADOX 4.0 for DOS***
  - ***Borland PARADOX 1.0 for Windows***
  - ***Microsoft ACCESS 1.0 for Windows***
  - ***Microsoft FOXPRO 2.5 for Windows***
  - ***INGRES Intelligent Database for AIX / 6000***
  - ***ORACLE for IBM***

2nd IBM definition DDM:

"Distributed Data Management DDM" is a function of the operating system that allows an application program or user on one system to use database files stored on remote systems. The systems must be connected by a communications network, and the remote systems must also be using DDM.... Remote systems on which data can currently [1993 dA] be read or written are: CICS / MVS (TM), CICS / VSE (TM), OS / 400 (R) and 4680 Store Systems. "

Announcement Letter Number 293-686 dated November 16, 1993
US - Last Revised on November 16, 1993

Brief Description of Announcement, Charges, and Availability

Two additions to the ADSTAR (TM) Distributed Storage Manager family of products, ADSTAR Distributed Storage Manager/2 Version 1.2 and ADSTAR Distributed Storage Manager/6000 Version 1.2 are being announced today. They will provide the same storage management functions on AIX/6000 (TM)- and OS/2 (R)-based systems as the original ADSTAR Distributed Storage Manager Version 1.1 did for VM- and MVS-based systems.

ADSTAR Distributed Storage Manager is a client/server storage management product that provides administrator controlled, highly automated, centrally scheduled, network-based back-up and archive functions for workstations and LAN fileservers. It backs up data from clients running on OS/2, NetWare (1), Windows (2), DOS, Macintosh (3), and UNIX (4) platforms to a back-up-server running on AIX/6000, OS/2, VM or MVS.

Extract from the ADSM V1.2 announcement with OS / 2 and AIX as server platforms

In November 1993, ADSM servers on OS / 2 and AIX were announced with ADSM V1.2 (ADSM / 2 and ADSM / 6000). This was the beginning of the era of OpenSystem platforms for ADSM.

The backup market was characterized by a strong polarization between desktop / PC-based backup solutions (DOS, OS / 2, Novell NetWare, Apple Macintosh) and corresponding solutions for UNIX workstations. What was common to all of these solutions was that they could support magnetic tape as a backup medium because of the advantages as an inexpensive mass storage device for long-term storage.

The following lists are an extract from a market overview from the PCMagazine of March 29, 1994, pp.227272 (source: http://stason.org/TULARC/pc/storage2/ index.html). The lists only contain solutions in which the backup server itself was operated on OpenSystem platforms, in parallel there were also offers that used the mainframe as a server, such as “Harbor” from New Era or “FDR / Upstream” from Innovation Data Processing:

- ***Arcada Software - Storage Exec. (NT)***
- ***Avail (NT)***
- ***Cheyenne Software - ArcServe (Netware)***
- ***Conner Storage Systems - Backup Exec (Netware)***
- ***Emerald Systems - Xpress Librarian***
- ***Fortunet - NSure NLM / AllNet***
- ***Hewlett-Packard - OmniBack II (NT)***
- ***Legato - NetWorker (Netware)***
- ***Mountain Network Solutions - FileSafe***
- ***NovaStor (Netware)***
- ***Palindrome - Backup Director and Network Archivist (Netware, OS / 2, Windows)***

Below is the list of UNIX solutions, whereby the importance of magnetic tape support was not as serious as that of desktops / PCs, because UNIX already knew tape support through the system functions tar, cpio, dd and dump.

- ***APUnix - FarTool***
- ***Cheyenne - ArcServe***
- ***Dallastone - D-Tools***
- ***Delta MicroSystems (PDC) - BudTool***
- ***Epoch Systems - Enterprise Backup***
- ***Hewlett-Packard - OmniBack II***
- ***Legato - networker***
- ***Network imaging systems***
- ***Open Vision - AXXion Netbackup 2.0 software***
- ***Software Moguls - SM-arch***
- ***Spectra Logic - Alexandria***
- ***Unitree software - Unitree***

ADSM V1.2 broke into these markets, a company-wide backup solution on OpenSystem platforms for desktops and UNIX workstations, which had already proven itself as a server platform under the mainframe operating systems VM and MVS and brought in all of IBM's experience in managing large amounts of data. This was particularly important for those customers who, on the one hand, used the new "democratic" organization of PCs and workstations in distributed processing, and on the other hand did not want to do without the high security and availability standards of mainframe technology for central data management.

ADSM V1.1 - manuals from the "black books collection"

Subsequently, one of the top design principles of WDSF / VM, ADSM and TSM proved itself for ADSM - the use of "common code". It did not have to be redeveloped for every client / server platform, but the reuse of code proved to be cost-saving. The process reduces the effort required to update and maintain the developers and speeds up the development cycles considerably - a principle that has made hardware and software developments economical until today.

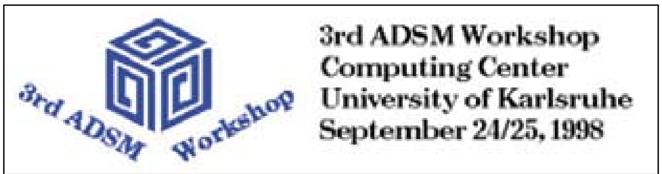
The establishment of Internet forum and user groups

The abbreviation ADSM is still a widespread synonym for Tivoli Storage Manager and you can always discover areas in the TSM environment in which this old product name is still used.

So z. B. in the well-known ADSML Internet UserForum (www.adsm.org), which was founded at the beginning of the 90s primarily from the large user group of the universities and is still maintained on the Internet as a knowledge database and communication platform for all those interested in ADSM / TSM. During this time (1994), user groups also began to form in Germany, such as the GSE working group "Storage Management" (www.gsenet. De), initially led by Chairman FranzXaver Maier and since 2003 actively lived by his successor Gerd Becker (IBM business partner Empalis). Here in Germany, interested customers meet twice a year on three consecutive days for lectures and exchanges of experience on the topics of mainframe storage, open system storage and TSM.

The first ADSM Large European Sites (ALES) workshop was held in Karlsruhe in 1994, initiated by the Universities of Karlsruhe and Heidelberg, namely Prof. Dr. Gerhard Schneider, Klaus Dilper and Rolf Bogus, an exchange of experiences between large European ADSM users from the university, scientific and commercial customer base. As a result, the biennial TSM Symposium (Karlsruhe 1994–1998, Oxford 1999–2007 - initiator Sheelagh Treweek, Königswinter / Petersberg 2009 - initiator Claus Kalle, University of Cologne) with more than 200 international participants and speakers from the Participants, from business partners, software manufacturers and active personal support from the TSM laboratories (including Paul Bradshaw, Andy Raibeck and from 1999 Dave Cannon).

- ***March 1994 - 1st ALES workshop Karlsruhe***
- ***April 1996 - 2nd ADSM Workshop Karlsruhe***
- ***September 1998 - 3rd ADSM Workshop Karlsruhe***



- ***September 1999 - ADSM Oxford Symposium:***
  Meeting the challenge
- ***September 2001 - TSM Symposium Oxford:***
  Managing The Impossible
- ***September 2003 - TSM Symposium Oxford:***
  A New Perspective
- ***September 2005 - TSM Symposium Oxford:***
  Facing the future
- ***September 2007 - TSM Symposium Oxford:***
  Preparing the path

- ***September 2009 - TSM Symposium Cologne / Petersberg:***
  Experience The Challenge

ADSM server on other open system platforms - V2.1 1995

COMPUTERWOCHE

Home | Management | Hardware | Software | Netzwerke | Mittelstand

Big Blue oeffnet sich jetzt bei Backup-Software

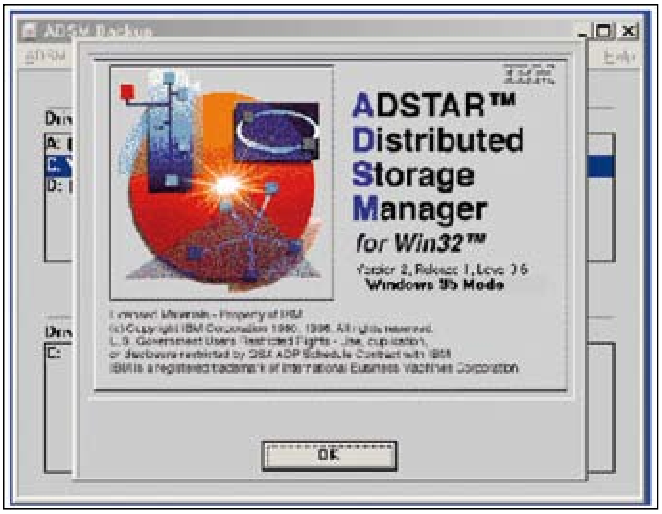
31.03.1995

ARMONK (pi) - Die IBM macht mit der Oeffnung fuer Fremdsysteme nun auch im Backup-Bereich ernst. Mit den zwei neuen Tools "ADSM for HP-UX" und "ADSM for Solaris" kommen ab sofort die Unix-Systeme von Hewlett-Packard und Sunsoft als Backup-Server in Frage. Bisher unterstuetzten die Tools der ADSM-Reihe auf der Server-Seite lediglich die Palette der IBM-Betriebssysteme. Als Clients koennen Solaris, AIX, HP-UX, Ultrix, Sinix, SCO-Unix, das System-V.4.-Unix sowie OS/2, DOS, MacOS, Windows und Netware fungieren.

CW article (excerpt) from March 31, 1995 to announce ADSM V2.1 (Source: http://www.computerwoche.de/heftarchiv/1995/13/1113172/)

Through the use of "common code" ADSM V2.1 was expanded in the course of one year to November 1996 by the server platform OS / 400, VSE and by a number of client and application platforms.

- ***Hewlett-Packard HP-UX 10.01***
- ***SunOS and Sun Solaris***
- ***Windows NT and Windows 95***
- ***Apple Macintosh Windows 32-bit client***
- ***NCR UNIX SVR4 3.0 and NCR EWS-UX / V***
- ***Bull DPX / 2 and Bull AIX***
- ***Digital UNIX***
- ***Siemens Nixdorf SINIX 386/486 and SINIX RISC***
- ***SCO UNIX386 and SCO Open Desktop***
- ***Sequent PTX***
- ***Silicon Graphics IRIX***
- ***Pyramid Nile (via SINIX RISC)***
- ***Open Edition MVS***
- ***AS / 400 API client for BRMS***
- ***Lotus Notes 4.1***
- ***HSM client for AIX and Solaris 2.5***
- ***Cray UNICOS and Fujitsu UXP as a service offering***



ADSM V2.1 Win 95 Client GUI and logo

The "Lotus Notes Backup Agent" was the first regular product expansion (in contrast to optional service offers such as Backint / ADSM) for online application backup in ADSM. It opened up the possibility of creating consistent online backups and also using the "incremental forever" procedure for Notes documents. This procedure had the advantage for Notes users that they could also save and reconstruct individual mails or attachments, but suffered from the disadvantage of poor throughput and heavy workload for the ADSM database, because every saved document also generated a metadata entry. The later database-oriented architecture of Lotus Domino from V5 no longer had this “single object backup” interface and the ADSMAgent was adapted accordingly.

Other functional innovations of ADSM V2.1 included:

- *the provision of the "Disaster Recovery Manager" for the automation of disaster recovery precautionary measures and a structured procedure for status tracking of the outsourcing tapes (offsite vault management)*
- *the automated generation of ADSM database Backups and storage pool backups (copypool)*
- *Extension of the scheduling process with pre- and post-processing options and support of admin commands in the scheduler HSM clients for the capacity management of*
- *UNIX file systems (HSM client for AIX V3.2 / V4.1 and Solaris 2.5)*
- *Extensions of the ADSM API client with the functions "Partial Object Retrieve" and the provision of an OS / 400 API client*
- *Announcing the availability of an Oracle / EBU client for late 1996*

During this time, the first complementary products with ADSA connections appeared on the market, some of which were resold by IBM under reseller contracts or where there were development partnerships with the manufacturers.

- *IBM Personally Safe'n Sound - IBM data backup Solution for OS / 2 servers for connecting local storage resources or the optional connection to the ADSM server.*
- *DataTools "SQL BackTrack" - November 1994 (data tools was taken over by BMC in 1997), a backup software for various relational databases (DB2, Oracle, Sybase, Informix etc.) with the unique feature of the OBSI interface, which allows all interface-compatible data backup applications to be used as a backend without the Having to replace the backup code in the application.*
- *Avail Systems "Netspace for Netware" (Avail was Acquired in 1995 from Wang Laboratories Inc., later sold by Kodak) - an HSM solution for Novell Netware servers.*
- *ETI back home! - Backup integration for tandem guar dian high availability solutions*
- *Core Data Inc. "Remoteworx" (Core Data was 1999 acquired by Sterling) - a data protection solution for desktop and laptop users with Windows95 and WindowsNT as the operating system*
- *Storage Solution Specialist Inc. SSSI - "ABC OpenVMS Client" for ADSM*

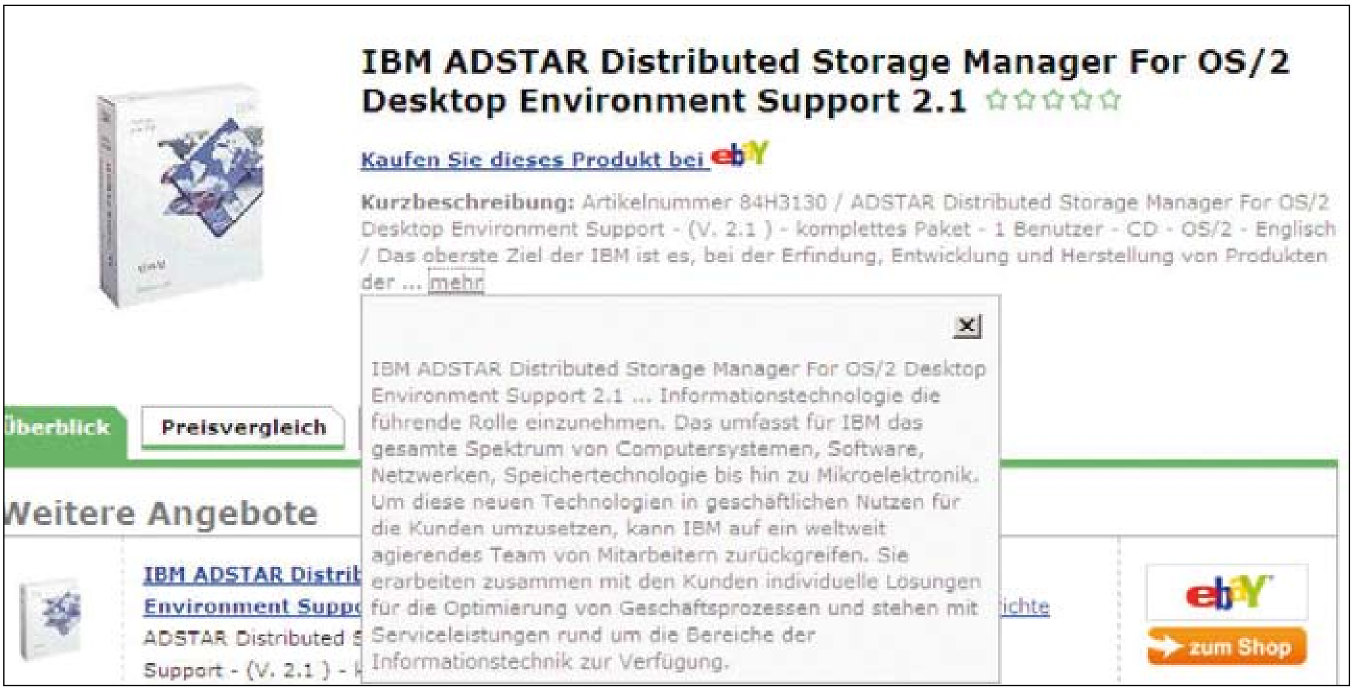
- ***Seagate "Backup Exec"*** *(formerly Conner Storage Systems - Seagate was acquired by Veritas in 2000) - data protection software for small and medium-sized environments with an ADSM interface for using the hierarchical ADSM storage pools as an alternative to directly connected tape drives.*

- ***EMASS / gray*** *- Connection interface for AML / S, AML / E, AML / 2 tape libraries via the ADSM External Library Management API.*
- ***Sterling SAMS Vantage DP*** *- Reporting tool for ADSM environments (Sterling was acquired by CA in 2000)*

It is noteworthy that with appropriate research on the Internet you can still find advertisements for ADSM V2.1 today, as the figure "Result from an Internet search 2009 for the keyword ADSTAR" shows. However, it must be doubted whether the product is really still available under this source.

SAP R / 3 integration - a success story from the IBM Laboratory in Böblingen

The early connection to the "Backint interface" developed by SAP was particularly important for the success of ADSM. The SAP developers considered this interface necessary because the integrated backup functions of the Oracle database were not sufficient at the time to create consistent online backups and the requirements provided for a close link between backup and the SAP-specific management tools. Due to its geographical proximity to the SAP development in Walldorf, the IBM Laboratory in Böblingen had a strategic advantage for this implementation, because continuous agreements and tests with the application manufacturer are necessary to maintain such an integration.



Result of an internet search 2009 for the keyword "ADSTAR"

Here is a brief tabular Backint / ADMP product history of the early days:

- ***The idea for Backint / ADSM came up in summer 1993 due to numerous customer inquiries.***
- ***September 1993 Meeting with SAP - Backint interface***
- ***Market launch as Service Offering 1994 - V1: call the ADSM functions through commandline - dmsc calls, AIX, Oracle***

- ***First customer installation April 1994 - approx. 10 customers until the end of 1994***
- ***Extensions 1995 and 1996 - parallel sessions, Porting to other operating systems (DEC, HP, SNI, SUN), first customer in the USA on HP-UX***
- ***Backint / ADSM V2 - API interface to ADSM as a replacement the 1996 CmdLine implementation and Windows / NT Support***

- ***Extensions 1997-1999 - multithreading, parallel / alternate Paths and Servers Around 3000***
- ***customers worldwide at the end of 1999***
- ***Year 2000 Transfer to Tivoli and announcement as a member the TSM product family as Tivoli Data Protection for SAP R / 3, Administration Tools***
- ***V3 - Object-oriented implementation with modular Structure, "common code" for Oracle and DB2, support for DB2, Administration Assistant***

Media collection (3.5 "diskette, 8 mm tape cassette, CD, manual) for Backint / ADSM, TDP for SAP R / 3 and TDP for hardware from different years

Atypical for the other ADMP product structure was the fact that the Backint / ADMS offer for SAP R / 3 was a key solution, but until 1999 it was only offered as a service offering. The reason for this was possibly the fact that SAP R / 3 was initially only successful on the European market. It is thanks to the creativity of the developers and the market commitment of the development team that this module has become the most important and economically most successful ADSM / TSM components overall. In 2000 the module became a formal IBM Tivoli product and the Böblingen department an official TSM development location.

October 1997 ADSM V3.1 - Extensions for company-wide use

ADSM V3.1 start screen

Version V3.1 of ADSM brought the greatest amount of substantial functional innovations in the TSM history:

- ***ADSM server on Windows / NT and HP-UX V3.1***
- ***HSM client for Solaris V2.5.1 with Veritas file system and SGI IRIX with XFS file system (integration via the X / Open DMAPI interface published in 1997) Provision of the backup GUI***
- ***also for UNIX plat shape based on the "motif" application architecture, graphical editor for the options file, "estimate" function, progress bar for backup / restore, search and filter options***

- *Point-in-time restore option for clients*
- *"No-query-restore" protocol in which the ADSM server and not the client does the optimized compilation of the data to be restored*
- *"Small file aggregation" and "server file aggregation" - the Summary of small files in transaction groups and storage in logical aggregates "Restartable Restores" - restart on the last one*
- *"Uncommitted transaction"*
- *The replacement of the generic admin GUIs by the web Admin client, which led to user interface standardization Web client as*
- *an alternative to the generic client GUI - thus the availability of a GUI for Netware clients and the support of central help desks SQL SELECT queries against the database*
- *(SQL'95 Command set) and ODBC interface*
- *Extension of the ADSM scheduler to define one-time action plans or executing operating system commands*
- *"Migdelay" and "migcontinue" for predictable migration of data on chained storage pools*
- *"Space trigger" for the automatic expansion of DB and log volumes of the ADSM database Storage of*
- *scripts and macros in the ADSM Database and provision of the "run" command to execute macros*
- *"Client option sets" for unification and central Client Configuration Management Server-to-Server*
- *Communication - Deploy the "ADSM Configuration Managers" for uniform script, policy, schedule management between multiple ADSM servers and the option of defining "virtual volumes" on a remote ADSM server (peer-to-peer disaster protection scenarios)*
- *"Overflow storage pool" and "single-drive reclamation" For configurations with only one magnetic tape drive, event logging,*
- *reporting and monitoring with interfaces Among other things, to SNMP managers, Tivoli NetView and Netview / MVS*
- *ADSMConnect agents for:*
  - *DB2 for AIX, HP / UX, SINIX, Solaris, OS / 2, Windows NT*
  - *Informix for AIX, HP / UX, Solaris*
  - *Lotus Notes for OS / 2 and AIX*
  - *OnDemand for AIX*
  - *Oracle 7 (EBU) and Oracle 8 (RMAN) for AIX and Solaris*
  - *SAP with Oracle for AIX, Digital UNIX, HP / UX, NCR UNIX, SINIX, Solaris, Windows NT based on the Backint interface*
  - *SAP with ADABAS-D for AIX, HP / UX, SINIX, Solaris, Windows NT based on the Adint interface (service offering) MS-Exchange*
  -

With the ADSM V3.1 product portfolio, IBM catapulted the breadth of the functional innovations and the provision of future-oriented technologies and processes into the premier class of data management applications, which the market and the analysts also rewarded accordingly.

Synergy through new combinations of IBM storage hardware and ADSM

Excerpt from an internal laudation (1997) for "Outstanding Complications" in the IBM Storage Systems Division:

… A major new release of ADSM (Version 3) was made generally available in October 1997, with Research contributions that include a Web administrative interface, support for SQL queries of its internal database, point-in-time restore, and improved archive file management . ***In addition to its sales as a separate product, ADSM is a key building block in SSD's Network Storage Manager and Virtual Tape Server offerings, which are integrated hardware and software solutions.***

The interesting news from this award is that the recently launched “Seascape Storage Building Block” initiative by the IBM Storage Systems Division (the reuse of proven IBM components as a development basis for complex subsystems) has also been extended to the ADSM software. This initiative subsequently resulted in successful combination products with ADSM such as the IBM 3494 Virtual Tape Server (VTS), the IBM 3466 Network Storage Manager (NSM) and the IBM 3466 WebCache Manager (WCM). In parallel, there were collaborations with the IBM software group in the form of the integration of ADSM with IBM ContentManager, Content Manager OnDemand and IBM CommonStore as a complete ECM solution on Open Systems.

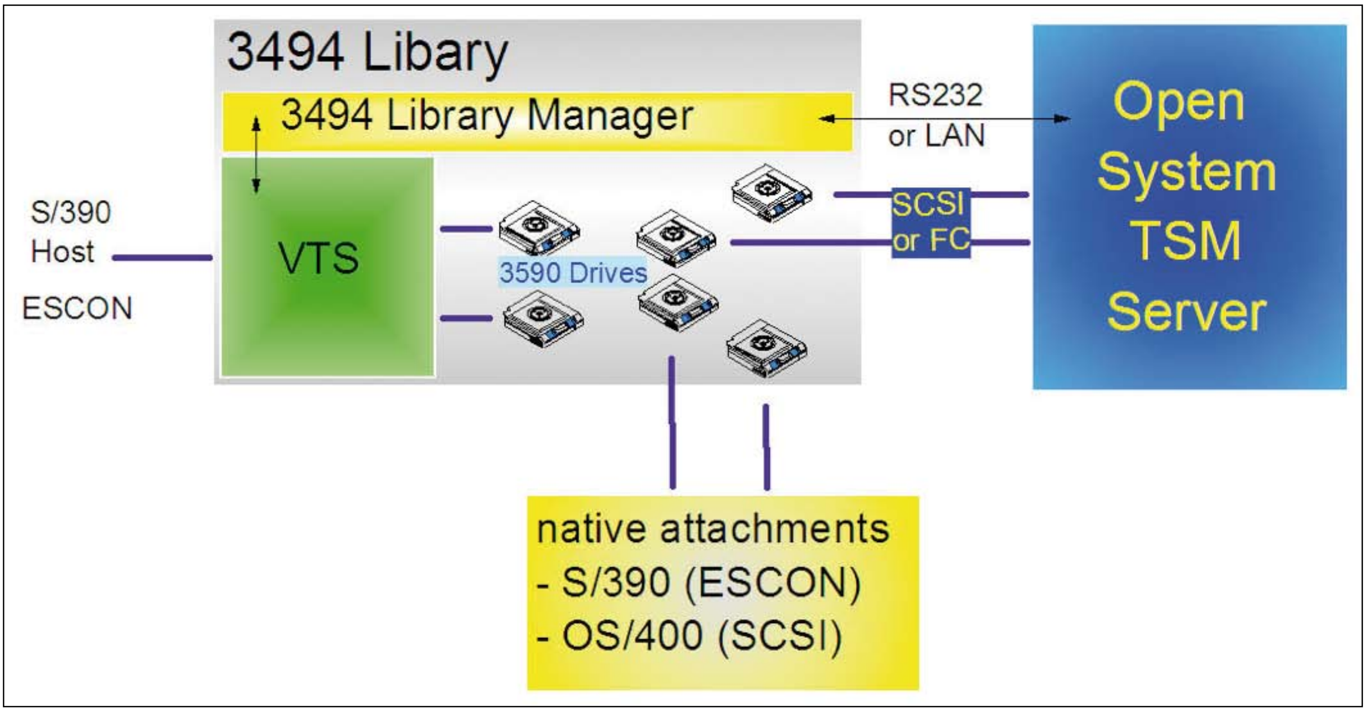
• *IBM 3494 Virtual Tape Server (the hardware-specific cations are described in the hardware chapter "The era of RAID systems") - the VTS was the implementation of a large number of virtual tape drives and virtual tape cartridges and was based on a modified AIX ADSM server / HSM client solution in the B16 Control unit implemented. Outwardly, 3490 drives and the 3490 virtual volumes (typically 800 MB) were represented and, after filling, were migrated to physical drives of the IBM Library 3494 using the ADSM / HSM functions. With the "Open System Attachment Feature", ADSM servers on open system platforms were also able to use the VTS functions. However, this solution was used less frequently because in the open system environment the ADSM server only shows its full strengths for managing the storage media. if he is the sole manager of all storage units. This reservation also applies to those Open System VTLs that have implemented transparent migration to physical drives.*

IBM 3466 Network Storage Manager

could. The first Model A00 had integrated SCSI adapters to which you could connect external tape libraries, Model A01 had a built-in library (either DLT or Magstar MP 3570) and Model B01 was delivered with IBM 3494/3590. Successor models with changed technology components had the designations C00, C10, C20, C30.

• *IBM 3466 Network Storage Manager (April 1998) - the "ADSM-in-the-box" solution - today one would call it an "appliance" - was a pre-configured ADSM server based on an RS6000 / H50 with disk pool on SSA disks, which can either be used as a JBOD (up to 814 GB) or pre-configured in RAID5 mode (up to 610 GB) and process the first data backups in a short time after the hardware installation*

• *IBM 3466 WebCache Manager (September 1998) - these The solution was a fascinating modification of the NSM, which was designed for the special task as an Internet file transfer resource. File transfer services via web servers were already in great demand at this time, but the bandwidth of the public networks to the web servers was often limited (ISDN speed) and online disk storage was expensive.*



VTS with Open System Attachment Feature and native connections to the 3494

The WebCache Manager - like the NSM - consisted of an RS6000 / H50 computer with ADSM server on AIX, an internal SSA disk storage whose file system was managed by an AIX / HSM client and an IBM 3494/3590 tape library.

In the HSM implementation of ADSM, stub files are used as a proxy for the physically migrated files. The stub files are usually the size of a block of the file system (4 k). In the WCM, this block size was changed to a value of 256 kB and larger. If a file transfer of a very large file via the Internet (ISDN) was requested, the transfer could be started immediately and the HSM agent started to load the rest of the file from tape in the background. Reloading was quicker when the stub file was transferred to the user, which gave the user the illusion of having immediate access to any file, however large.

The economic advantage of this solution is described in an extract from an article from MONITOR Online 2/99 by the customer Deutsche Telekom AG (source: http: //www.monitor. Co.at/index.cfm/storyid/101):

… The new IBM 3466 Web Cache Manager is part of the Seascape storage architecture developed by IBM and represents a combination of leading storage technologies. A powerful storage server, disk and tape storage media as well as the corresponding storage management software were integrated into an overall solution. With a unique combination of hierarchically used tape and disk storage media, the product offers up to two terabytes a multiple of the storage capacity of conventional web proxies and can also cache very large pages and files with 10 MB and more. Since the pages and files are kept locally, the Web Cache Manager makes an important contribution to improving response times on the Internet […].

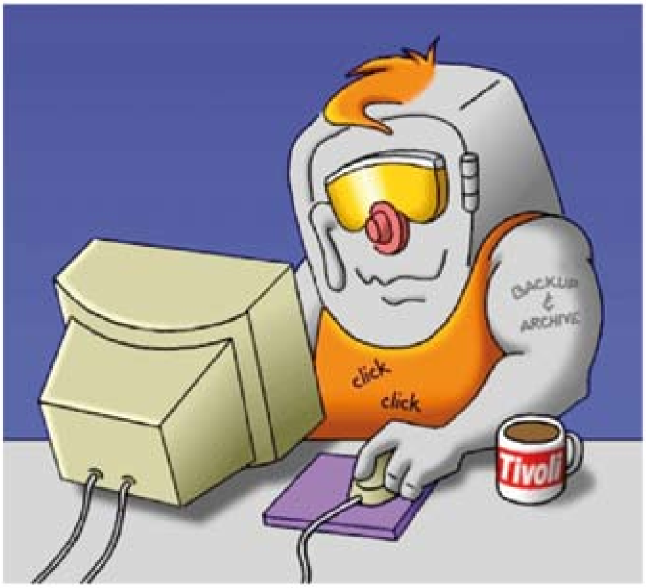
[…] Our current tests have shown that we can save more than half of the web traffic in bytes with the Web Cache Manager. In contrast to other web proxies, the Web Cache Manager works with a hierarchy of two storage media, which enables significant cost reductions on the hardware side. SSA (Serial Storage Architecture) disks and Magstar tapes are used. This choice of data storage offers ISPs the best possible price-performance ratio and the largest cache currently available. Tape storage is five to ten times cheaper than disk storage, making it the most cost-effective way to store larger web objects with a megabyte or more

- because around 50 percent of web traffic now consists of content of this magnitude ...

- ***IBM ContentManager, ContentManager OnDemand, Commonstore, Admira - From the mid-1990s, the efficient media management functions of ADSM were also used as a backend for the IBM archive applications on Open Systems. The ADSM-API client was used for this, which is the basis for all application integrations. During this time, optical media with WORM functions on the ADSM server were preferably used as back-end storage for ECM archive solutions (e.g. IBM 3995).***

1999–2002: ADSM becomes Tivoli ADSM and eventually Tivoli Storage Manager

In March 1996, IBM took over the systems management developer Tivoli Software Inc. This takeover had consequences for the further product home of ADSM within the IBM. From January to October 1999 sales, marketing and development of the product were transferred to the responsibility of the IBM Tivoli division and the product was first renamed to IBM Tivoli ADSM (January 1999). In September 1999 the product was released as Tivoli Storage Manager Version 3.7 under a new name and the Safeman logo changed.



ADSM Safeman in Tivoli look

Extract from the press release regarding the transfer of marketing and sales responsibility for ADSM from the IBM Storage Systems Division to IBM Tivoli Systems:

Tivoli Systems Announces Next-Generation Applications Management Storage Product Line

Customers Gain ADSM Enhanced Benefits for Storage Management

SAN JOSE, Calif. - Dec. 2, 1998 - Tivoli Systems Inc. and IBM's Storage Systems Division (SSD), both part of the IBM company, announced today the transfer of ADSTAR Distributed Storage Manager (ADSM) to Tivoli. Responsibilities for the industry leading cross-platform application and data storage management solution will be transferred to Tivoli, effective January 1, 1999.

With the transfer, customers will gain a broader, more tightly integrated application and data management solution for their enterprise storage management needs. The move is also designed to significantly increase IBM's share of the growing distributed storage management market by combining the strength and technology of two of IBM's most successful divisions. ADSM also wants to provide customers with next-generation application-centric storage management solutions, supporting more platforms and operating systems than any other product available on the market today.

Today's announcement underscores the importance of this new Tivoli business segment. The company plans to further enhance its enterprise solutions with significant investments in integrated storage solutions spanning mainframes to notebooks. Plans include additional investment in cross-platform storage area networks (SANs) and application-centric solutions.

The versions V3.7 (September 1999) and V4.1 (July 2000) followed quickly. The "crooked" version number 3.7 was mainly caused by the name change from Tivoli ADSM to Tivoli Storage Manager; these versions also brought some important functional enhancements to version 3.1. Particularly worth mentioning is the entry into the world of storage area networks through the provision of tape library sharing, LANfree data transfer and the SAN data sharing options obtained through the acquisition of the Sanergy product from Mercury Computer Systems.

Tivoli Storage Manager V3.7 - new name and new logo up to V4.2

- ***Improvements in TSM server V3.7:***
  - ***No-Log Database Load & Audit***
  - ***Snapshot database backup***
  - ***SAN tape library sharing between TSM server system AIX, Windows 2000 and Solaris using SAN Data Gateways for SCSI Libraries Secure Web Admin Proxy***
  -
  - *Introduction of the backup sets (Instant Archive and Rapid Restore)*
  - ***Windows 2000 with cluster support***

- ***Improvements in the TSM clients V3.7:***
  - ***Subfile backup for Windows clients***
  - ***Multi-session client (multithreaded backup)***
  - ***Windows 2000 support with the support of Cluster, System Objects, Active Directory, Journaling, Encryption, DFS***
  - ***Linux clients on Redhat, Suse, Caldera and TurboLinux***
  - ***LAN-free client transfer for the TSM API client (Application LAN-free) Client***
  - ***Encryption DES 56-bit***

- ***New products and name changes:***
  - ***Renaming of the ADSMConnect Agents to Tivoli Data Protection for ...***
  - ***Renaming of the TSM / HSM UNIX clients to Tivoli Space manager***
  - ***Tivoli Data Protection for IBM ESS and EMC Symmetrix for Oracle and SAP / Oracle connection to Tivoli***
  - ***Decision Support for Storage Management Analysis (12/99) - Reporting and analysis of TSM server operations***
  - ***Tivoli SANergy file sharing (04/00) - LAN-free-to-disk Backup technology as a result of the acquisition of the "Shared Storage Unit" from Mercury Computer Systems***

  - ***Tivoli Removable Media Manager (07/00)***
  - ***Tivoli Data Protection for Workgroups (TDPfW) - Bare Machine Recovery for Windows NT, Windows 2000 and Netware through development partnership (June 1999) with STAC ™ software using the STAC / Replica ™ module.***
  - ***Tivoli Plus modules for TSM and TDPfW for integration into the Tivoli framework (reporting, monitoring, software distribution)***

The turn from the 20th to the 21st century was marked in the storage technology by the introduction of the FibreChannel Storage Area Networks for the open systems with big consequences regarding the "socialization" of resources and the basis for the storage virtualization techniques common today. The deliberately conservative approach of the TSM developers in integrating this technology, which is new for Open Systems, was accordingly critically commented on by the press accompanying the market.

COMPUTERWOCHE

Newsletter

Home Management Hardware Software Netzwerke Mittelstand Se

Adstar in neuer Version angekündigt

Tivoli läßt sich mit einer eigenen SAN-Strategie noch etwas Zeit

17.09.1999 Drucken | Empfehlen | PDF | Merken | ShareThis

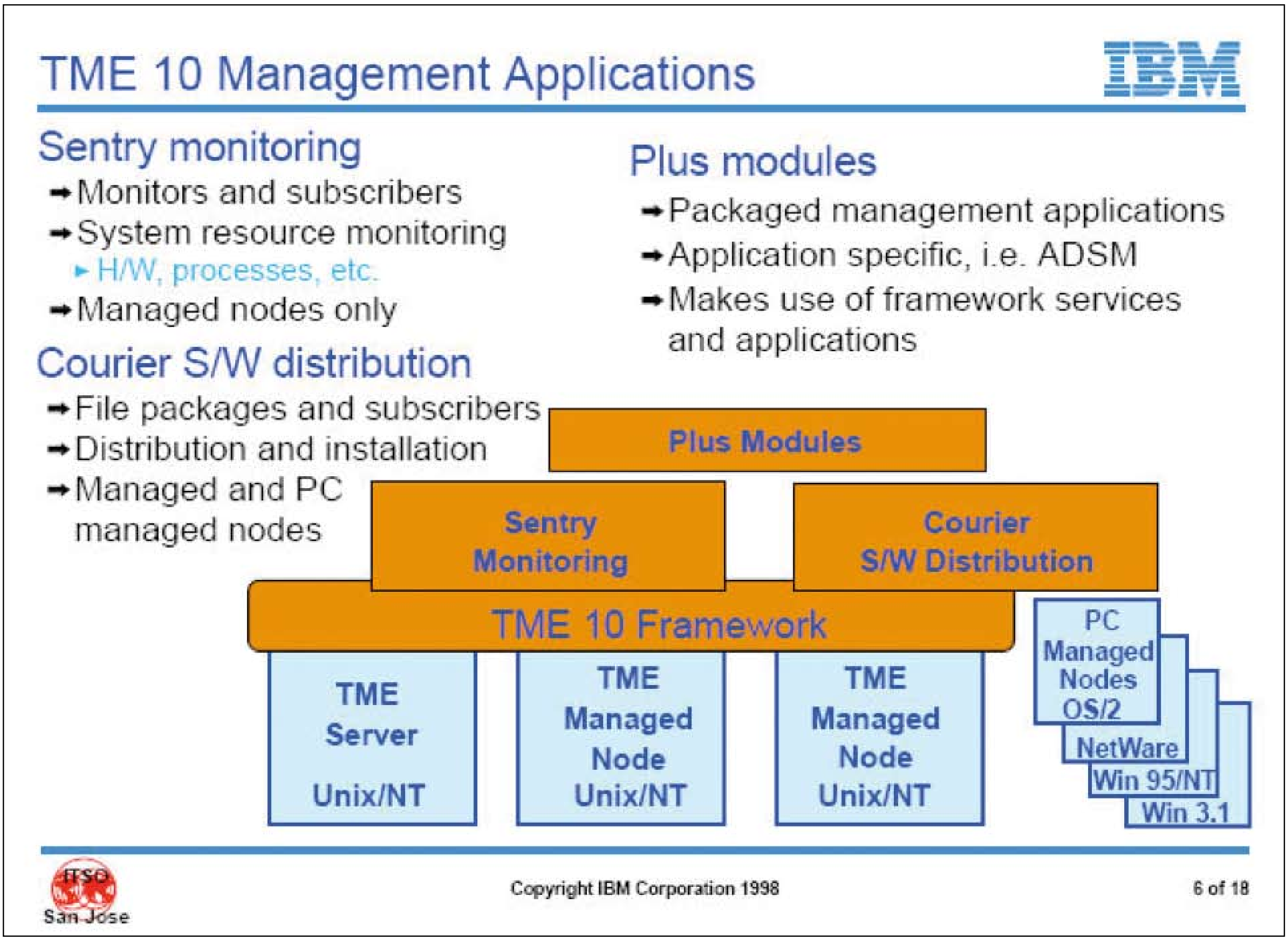
MÜNCHEN (jha) - Dem boomenden Speichermarkt kann Tivoli Systems Inc. nicht widerstehen. Die IBM-Tochter bemüht sich um eine schlüssige Strategie zur Verwaltung von Storage Area Networks (SANs). Als erster Schritt wird dazu ab September die Version 3.7 des "Tivoli Storage Manager" auf den Markt gebracht. Eine komplette SAN-Lösung soll jedoch erst in einem Jahr folgen.

Kräftige Rückendeckung für den Eintritt in den Speicherverwaltungs-Markt hat Tivoli von der Konzernmutter bekommen. Seit dem 1. Januar 1999 verantwortet die IBM-Tochter alle Aktivitäten in diesem Bereich, als Beigabe erhielt der System-Management-Spezialist das Produkt "Adstar Storage Management" (ADSM). Nach dem Motto "Alter Wein in neuen Schläuchen" wurde das Tool in "Tivoli Storage Manager" umbenannt, erste Neuerungen beinhaltet die für den September angekündigte Version 3.7.

Storage Manager als eigenes Paket

Obwohl sich der Storage Manager in das Verwaltungs-Framework einbinden läßt, wird das einstmalige IBM-Produkt, das früher als Beigabe zu Bandlaufwerken vertrieben wurde, auch als eigenständiges Paket verkauft. Ob integriert oder separat - eine schlüssige Antwort auf die Herausforderungen des SAN-Marktes ist der Storage Manager noch nicht. Mit einer Lösung für diese Speicherverbände nimmt sich Tivoli noch etwas Zeit. Ungefähr in einem Jahr, so Dale Head, als Marketing-Manager zuständig für die Speicherverwaltungslösungen, erscheint der "Tivoli SAN Manager", mit dem sich beispielsweise Server-freie Backups vornehmen lassen.

CW article (excerpts) from September 17, 1999 on the TSM V3.7 announcement (source: http://www.computerwoche.de/heftarchiv/1999/37/1088572/)



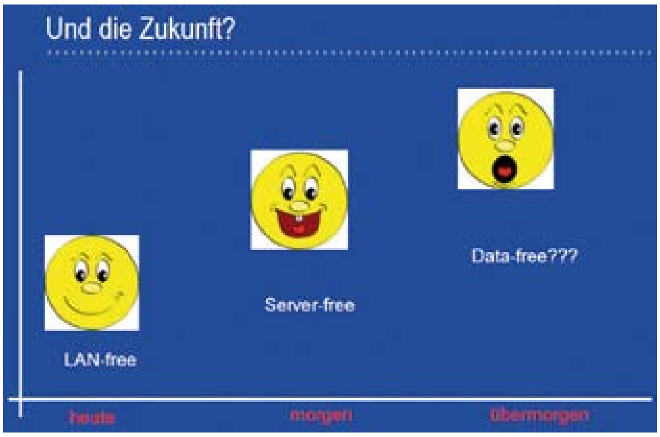
Integration of ADSM in the Tivoli TME10 framework via PLUS modules

The swipe of the article about "ADSM as an add-on to tape drives" was indicative of the zeitgeist, since the editor obviously did not rate complete infrastructure solutions as a desirable goal for companies. However, with these combinations of hardware, software and services, IBM was steadily gaining market share. These solutions are coordinated and tested and offer customers one-stop support in the event of a fault.

The cautious approach to implementing SANTechniques in TSM such as library sharing and LAN free backup proved to be justified because the path to reliable SAN environments was bumpy due to the lack of compatibility of infrastructure components and protocols. In addition, they were required to manage and monitor storage area networks Management functions only rudimentary available. The necessary standardization (Bluefin, SMIS) was initially hampered rather than accelerated by competing standardization consortia (FibreChannel Association, SNIA etc.).

For TSM, the new possibilities of the storage area networks meant a quantum leap in the implementation of distributed backup instances and disaster protection solutions. However, some SAN advertising messages also included salvation promises that this technology could not deliver:

"Your LAN backup takes 8 hours today, with LAN-free it will take 2 hours and with Server-free you will be finished after 20 minutes." (Original sound from a marketing event of that time).



Not really serious final slide of a TSM4.1 lecture

The LANfree and Serverfree backup methods were the dominant topics at that time. As is common with fashion topics, publications have often dealt with the possibilities of using the new technologies in a striking rather than precise manner.

The most frequently advertised server-free approach was the implementation of the SCSI3 T10 extended copy / 3rd party copy standard developed in 1996 (also standardized under SNIA 143R1).

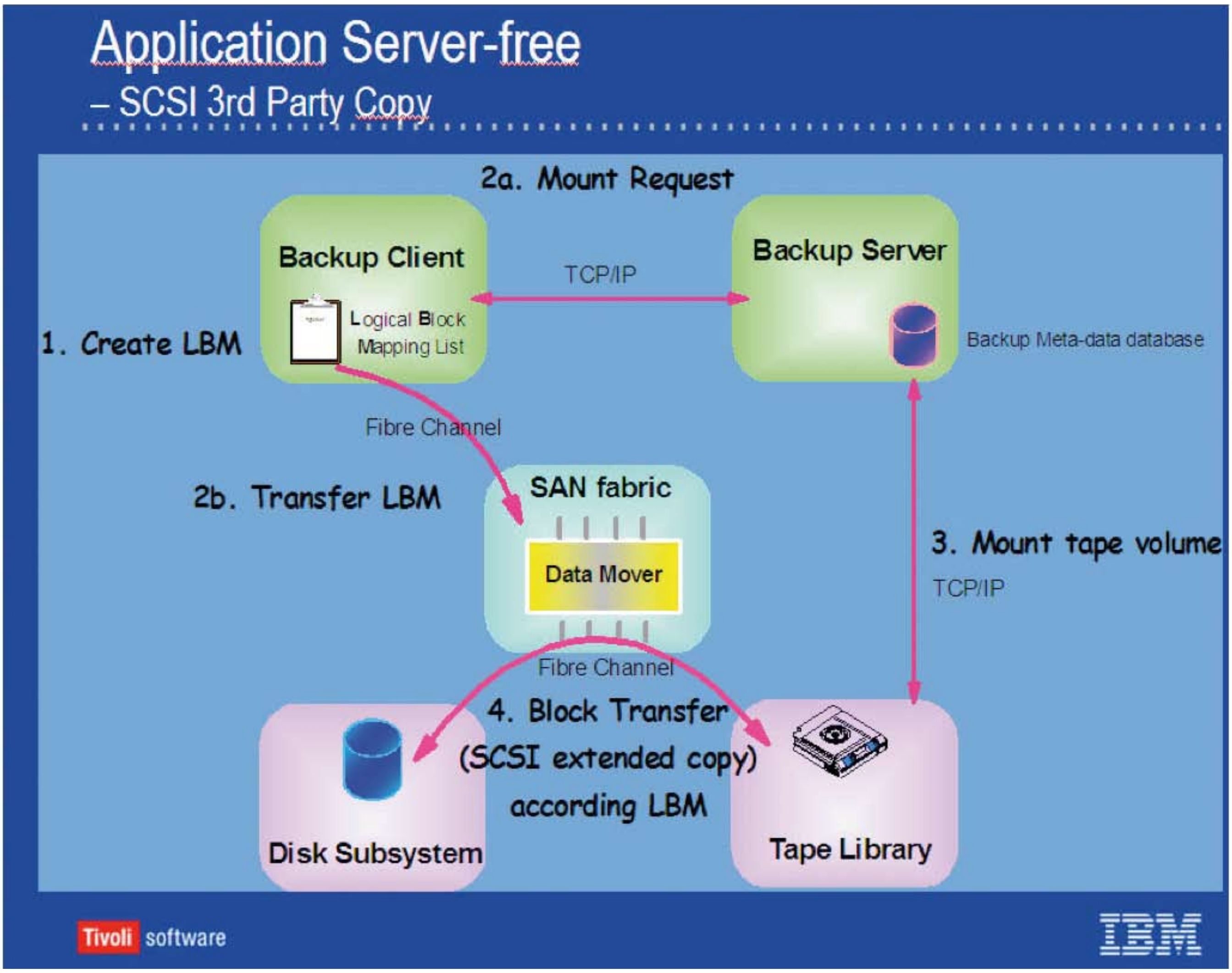
It was defined here how a “data mover” instance could be mapped within SAN infrastructure components (e.g. SAN data gateways expanded with microcode functions), which was able to copy data in the SAN directly from disk to tape without the Having to take a "time-consuming" detour via the application server.

This solution was only feasible by adhering to a number of framework conditions and required a chain of T10 compatible and tested hardware and software components. The number of manufacturers for "data mover" instances (SAN data gateways) was limited with the manufacturers Pathlight ™ and Crossroads ™. It was also not possible to save or reconstruct individual files with this method because the copying process was a volume block level operation. The necessary opening of the file system manufacturers did not take place, which would have been necessary in order to be able to map the logical files to the physical blocks using standardized API interfaces (LBM = Logical Block Mapping List). The LBMs were not needed for volume dumps. With the Celestra ™ product, Legato ™ attempted to implement this itself using reengineering processes, but then failed in the medium term due to the necessary maintenance work due to poorly published changes by the file system manufacturers. Other approaches such as the Vertex initiative from Veritas ™ tried to solve the problem on the basis of their own proprietary volume manager or file system.

The topic SCSI3 T10 occupied the IT world for several years and was obviously so important for the market presence that TSM with the version V5.1. made a corresponding implementation available for Windows 2000 client, IBM SAN Data Gateway 2108 (Pathlight) and TSM server on Windows 2000 (the TSM demo highlight at CeBIT 2002).

Today, “Application Serverfree” solutions are defined by other technical options, whereby the principle of “workload offloading” is still the heart of the process, such as: For example, the "Copy Backup" procedure DisktoDisk via the snapshot functions of operating systems, file systems, disk subsystems or proxy solutions such as VMware Consolidated Backup.

Critical for all online backup solutions that work via replacement instances (server-free and copy backup) is the necessary application integration to ensure data consistency during backup. This must either be done using script-based methods or via the operating system or application interfaces provided, e.g. B. Microsoft Volume Shadowcopy Services or SAP "Splitint".



Server-free process flow with SCSI 3rd party copy (T10 standard)

TSM V4.2 was released in June 2001 with the innovations:

- ***The expansion of LAN-free support for others Platforms and the range of TDP clients***
- ***Journal-based backup for Windows clients***
- ***zLinux client, USS client (zOS Unix System Services)***
- ***Tivoli Data Protection for NDMP (filer-to-tape)***
- ***Tivoli Data Protection for Websphere Application Server***
- ***TSM Management Console as MMC plugin***
- ***Unicode support for Windows NT and 2000***

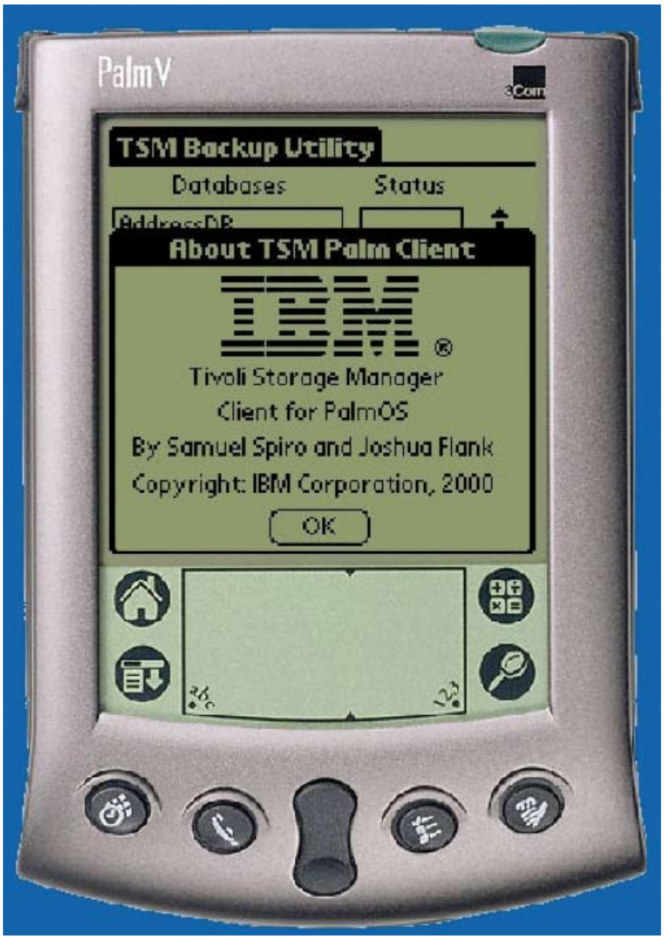
The years 2000 to 2003 were characterized by a growing number of development partnerships and reseller agreements between IBM Tivoli and independent software houses that had developed complementary software to TSM. Here are the most important partnerships and products:

- ***"The Kernel Group" (TKG) - The "Bare Metal Restore" - TKG's product was resold by IBM as "Tivoli Bare Metal Recovery Manager" and "Tivoli Disaster Recovery for Servers" from 2000 onwards. It was software that implemented the requirements of operating system reconstruction and was based on the saved data of the TSM server. The operating systems AIX, HP-UX, Solaris, Windows NT and later Windows 2000 and Novell Netware were supported. Veritas bought TKG in January 2002, and Veritas terminated TSM support in June 2003.***

- ***Cristie PC-BaX and later Cristie Bare Machine Re covery (CBMR) June 2003 - The partnership and the product, which closed the functional gap for operating system reconstruction in the TSM portfolio and is still an important building block as CBMR and TBMR (Tivoli Bare Machine Recovery - reconstruction of operating system solely by using the data stored in the TSM)***

Holistic data backup strategy in the TSM architecture is for the operating systems Windows 2003/2008, Linux, Solaris and HP-UX.

- *St. Bernhard Open File Manager (OFM) - The additional Solution that was necessary to be able to consistently back up file systems (Windows, Netware) or applications from a file system perspective (Exchange, Groupwise etc.).*
- *Gresham Computing plc. Enterprise DistribuTape (EDT) - formerly Open Microsystems (taken over by Gresham 1998) - The connection of ACSLS tape libraries (StorageTek and others) to the TSM External Library Management (ELM) interface. The reseller contract was signed in June 2001.*
- *Open Technology Group (OTG) - DiskXtender - The Windows 2000 HSM solution for TSM. The cooperation between IBM and OTG began in 2000 and ended in December 2004 because Legato acquired OTG in June 2002 and has since been bought by EMC.*
- *Monitoring and reporting solutions for TSM environment exercises - Servergraph Professional, JamoDat TSMManager, STORServer Manager*



TSM client for PalmOS in V4.2
(Prototype - was never an official product)

2002–2008 TSM V5.x - the era of server consolidation, virtualization and NAS

As of version V5.1, the license structure of the TSM has been changed. For this purpose, the previously used environment licenses (desktop environment, workstation environment, extended device support module) were replaced in favor of a processor-based price model. To do this, two basic license models were defined for a TSM environment:

- *Tivoli Storage Manager with tape library up to 2 tape drives and 40 cartridges*
- *TSM Extended Edition included for larger configurations TSM Disaster Recovery Manager and NDMP Support as a replacement for the paid modules Extended Device Support, TDP for NDMP and DRM.*

The versions V5.1 (April 2002), V5.1.5 (October 2002) and V5.2 (October 2003) extended the TSM functionality by the following focal points:

TSM V5.1 - April 2002

- *Server-free data movement based on the SCSI3 T10 proto- Collaboration, Windows 2000 client full-volume backup / restore*
- *Simultaneous Writes to Copy Storage Pools Write swap copies directly with the backup multithreaded restore -*
- *restoration over several parallel lele sessions, provided the backup was also multithreaded on different backup volumes Windows 2000 Image Backup and Open*
- *File Backup with Help from the TSM Logical Volume Snapshot Agent (LVSA) "MOVE*
- *NODEDATA " to selectively move Backup data of a client from one storage pool to another*
- *CRC (Cyclic Redundancy Check) - checksum check for TSM client and TSM server to ensure consistent data transfer over networks (LAN, WAN and SAN)*

- *Image backup for filesystem and raw devices from AIX, HP-UX and Solaris*
- *Reorganization of the TSM Application Clients in:*
  - *TSM for Databases (Oracle, MS-SQL, Informix)*
  - *TSM for Mail (Exchange, Domino)*
  - *TSM for ERP (SAP / Oracle and SAP / DB2)*
  - *TSM for Application Servers (Websphere)*
  - *TSM for Hardware (Copy Backup for Oracle, DB2 and SAP / Oracle, SAP / DB2)*

TSM V5.1.5 - October 2002

- *Full Linux / x86 support including TSM server, LAN-free, NDMP and Domino client*
- *Server-to-server import / export - the serious association Simplification of system migration and division processes of the TSM database for improved horizontal scaling Exchange Single Mailbox*
- *Restore via script integration into the Xmerge utility from Microsoft*
- *TSM-Server on OS / 400 PASE*

TSM V5.2 - October 2003 and V5.2.2 - December 2003

- *TSM server on zLinux*
- *TSM zOS LAN-free*
- *Flashback functions for TSM for ERP / DB2 together with TSM for Hardware*
- *Improved firewall support (server initiated sessions, single port communication)*
- *SAN device discovery for Windows 2000 TSM servers*
- *NAS environments: File-level restore through NDMP table of Contents (TOC) and EMC Celerra-Support*
- *Mixed-Media Libraries*
- *TSM for System Backup and Recovery (operating system reconstruction for AIX based on the IBM service solution Sysback6000)*
- *Extension of the retention time for archived data from 9,999 days (approx. 27 years) to 30,000 days (approx. 83 years) with no dependency on the year 2038 problem - also called "The Friday 13th Bug" - (http: //www.2038bug. com / index.html) .*
- *Tivoli Storage Manager for Data Retention - the software WORM variant of the TSM for use with revision-proof archiving*
- *TSM support for EMC Centera and NetApp Snaplock (TSM for DR)*

The TSM logo of versions V5.1 to V5.3

Tivoli Storage Manager for Data Retention and DR450 / 550 (October 2003)

With TSM for Data Retention, a special variant of the TSM has been provided for revision-proof archiving of data. This software is particularly important for the IBM storage landscape because it is the software basis for the archive backend solutions DR450, DR550 and IBM Integrated Archive (October 2009).

Hardware details on the DR450 / DR550 and a description of the legal framework aspects are described in detail in the HW chapter "The era of multiplatform systems and the FibreChannel SAN and NAS".

The reason for the development of such archive backend solutions was, on the one hand, the poor future prospects of optical storage media in general, which had previously been the key technology for legally compliant long-term archiving, and, on the other hand, the stricter legal regulations that, after a few cases of falsification of balance sheets, particularly in the United States were (among other things in response to the ENRON / Worldcom bankruptcy).

The fascination of TSM for Data Retention lies in the simple function that you can operate rewritable data carriers (magnetic disk, magnetic tape) in SoftwareWORM mode (writeoncereadmany) and additionally use new physical WORM media such as WORM tapes (IBM 3592, LTO3 and LTO4) . WORM means preventing overwriting and ensuring deletion only after the specified blocking periods have expired.

The idea came from the archive function that TSM has always offered in parallel with the backup. In "Archive" mode, TSM does not overwrite data, but rewrites each object, even if the object with all its content and attributes has already been saved. Objects of the same name differ by a clear object ID and by the date / time of storage and are also clearly accessible again. The basic mode of operation of "ArchiveMode" thus corresponds to the first requirement of revision-proof archiving, preventing "overwriting". The second cardinal requirement is to ensure deletion operations only under the control of the controlling archive application, which specifies and controls the blocking periods. For this purpose, an additional "retention" procedure was implemented in the TSM, which enables the traditional expiration control of the TSM ("chronological retention") to be replaced by a multi-stage expiration control controlled by the archive application ("event-based retention" and "deletion hold / release")). The deletion protection in the TSM for Data Retention is additionally supplemented by measures that prevent the TSM administrator from deleting archive data carriers or changing the binding of retention rules to objects.

The interface for data transfer to TSM for Data Retention is either the ArchiveClient or the archive function of the TSMAPI interface as of V5.2.2, which over the following years was used in a variety of DMS and archive applications within IBM (IBM ContentManager, IBM Common store, IBM Filenet and IBM Optim) and in solutions from independent software houses (including OpenText, SER, Saperion, d.velop, Easy, Ceyoniq, Waters, PBS and many others).

Logo for the Tivoli interface certification of external applications

In 2005, “Tivoli Storage Manager for Data Retention” was renamed “IBM System Storage Archive Manager”, mainly because of the changed licensing procedure (TBL license based on the managed archive data stock).

TSM V5.3 became available in December 2004 with the following important innovations:

- *Introduction of the Integrated Service Console as a result*
  a general user interface consolidation for Tivoli products for graphical administration extensions for day-to-day definitions in
- ***TSM***
  Scheduler
- ***Multiple processes for complaints and migration***
- ***Throughput improvements in the TSM file pool***
- ***Group collocation for improved use of high-capacity***
  quotative tape cassettes
- ***Automatic SAN device detection and reconfiguration***
- ***Generic support for ACSLS libraries without***
  external software
- ***Segmentation of NDMP backups, NDMP dumps***
  of snapshots or quota trees on NetApp NAS New encryption
- ***method for clients AES 128-bit***
- ***Introduction of proxy node support for sub-***
  support of virtualized system environments Windows Open File
- ***Support and online Image Support***
  Support for TSM clients in VMware guest systems via the TSM
- ***Logical Volume Snapshot Agent (LVSA)***
  and support of the Linux B / A client for the backup of ESX / GSX
- ***Administration Assistant for TSM for mySAP***
- ***TSM HSM for Windows - as a result a company cooperation***
  tion with Intercope GmbH in Hamburg (OpenStore for File Server OS4FS)

With version TSM 5.4 (January 2006), the promised maintenance intervals per release changed from 3 years to 5 years, which is very helpful for TSM customers who plan to upgrade in the long term. The most important functional innovations:

- ***Support of TS1120 encryption and the implementation***
  Encryption Manager in the TSM for the encryption mode "application managed"
- ***Implementation of the NDMP copy function for creation***
  copies for NDMP dump tapes
- ***NDMP "Filer-to-Server" mode - NAS Filer sends its***
  Volume dumps via LAN to the TSM server, where the data can be saved in TSM format at all storage hierarchy levels, including the creation of copypool tapes.
- ***"Data Shredding" function for disk pools - the multiple***
  Overwrite deleted data with random HEX values
- ***Backupset improvements (point-in-time, combination***
  with image backup, single file restoration via TOC, multiple backup sets on the same medium) "Active Datapool" - the possibility of
- ***copying the last one***
  Create or store full backups on a separate storage pool
- ***TSM for Copy Services - Use of Microsoft Volume***
  Shadowcopy services for the consistent snapshot backup of Exchange and MS-SQL including the "instant restore" function when using IBM disk storage (IBM DS6000 / DS8000) or IBM SAN Volume Controller.

TSM V5.5 was announced in November 2007, this was the last release using the proprietary database. The most important innovations:

- ***Encryption key management by the TSM server,***
  This also means that application clients (TSM for… clients) can send their data encrypted over networks without the application having to support the management of the keys. Improvements in sequential disk pools (including multi-
-
  threaded read in the same volume) Import and export
- ***processes can be restarted***
- ***VMware integration " VCB integration module" in***
  the Windows B / A client to support VMware Consolidated Backup VCB
- ***TSM for Advanced Copy Services - Extensions in the***
  Device support and usage for DB2 Snapshots NetApp
- ***"Snapmirror to tape" support as alternative***
  tive protocol for NDMP with improved scaling

TSM B / A Client Logo V5.4 and V5.5

2005-2008 Solutions for end users, small and medium-sized companies and branches of large customers

Tivoli Continuous Data Protection CDP (2005) was the product implementation of the "Vitalfile" tool, which has been available internally at IBM since 2004, with the target group of laptop users in the company. Interesting especially for those users who only have sporadic access to their company's security services.

CDP introduced a revolutionary innovation in data backup, the backup of changed files directly when saving (continuous data protection). The target medium for the data copy is first local storage on the laptop like

e.g. B. another logical drive of the partitioned hard disk or locally connected external memory (USB). You can also write a second copy to a remote network drive or to the TSM, with CDP trying to transfer data until a valid network connection is established.

At the end of 2009, CDP was renamed TSM Fastback for Workstations and expanded to include central configuration functions.

TSM Express was announced in March 2006 as the entry-level solution for TSM in small system environments. The modifications to the standard TSM were:

- ***Data backup only on disk***
- ***Optional connection of magnetic tape for generation of outsourced copies via TSM backup sets for file data and Exchange / SQL Very easy installation and operation***
-
- ***Integrated reporting***
- ***Support only for Windows environments with two specific application agents for Exchange and MS-SQL***
- ***Migration option to the standard TSM***

TSM Fastback comes from the company acquisition of the American / Israeli company FilesX Inc. by IBM in April 2008. TSM Fastback replaces TSM Express and provides completely new functions for securing small system environments - the exclusive backup of file servers or application servers using snapshot processes. By implementing a snapshot filter driver on the client, it is possible to implement all data backups via snapshots, which - in the tradition of the TSM - are purely block-incremental. The main features:

- ***Disk-only backup on the local fastback backup server***
- ***Online snapshots for file servers and application servers through the use of provided consistency functions***
- ***Windows and Linux clients (from early 2010)***
- ***The ability to save any saved snapshot either Restore as a complete volume or make the snapshot available as a network drive to selectively restore individual files. Item-level recovery / single mailbox recovery from Microsoft***
-

Exchange objects

- ***Restore the operating system partition from the Snapshots by using a loading CD replication of selective snapshots***
- ***from many remote ones Fastback servers to a central Fastback server over WAN lines***
- ***Connection of the central Fastback server to the central one TSM server for creating swap tapes.***

April 2009 –TSM V6.1 with new database architecture and deduplication

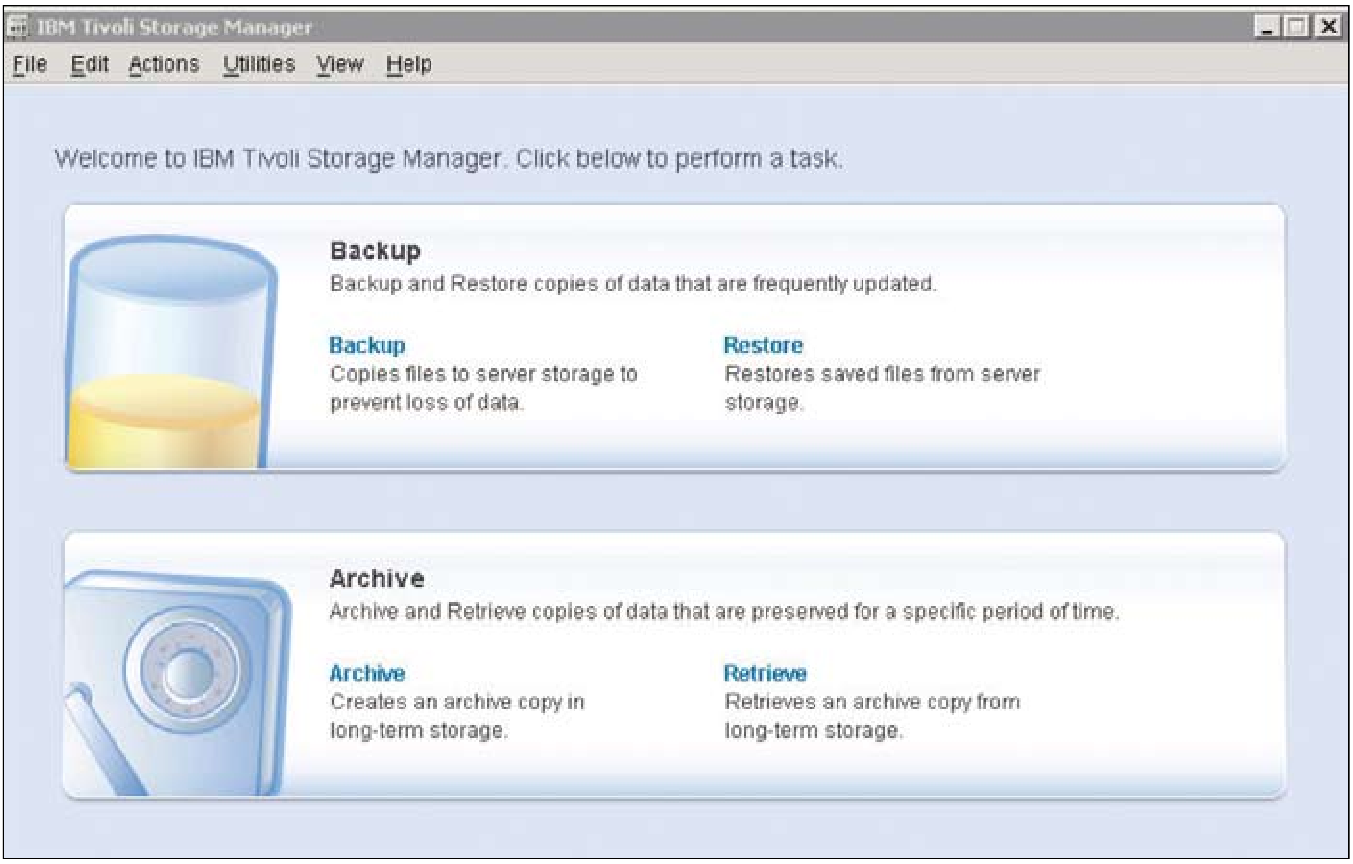
The database architecture based on B + Tree technology had existed for almost 20 years. The increasing backup volumes and, above all, the increasing number of backup objects made an architecture change necessary in order to scale better and adapt online availability to the increased requirements. Therefore V6.1 introduced DB2 as database technology for the management of TSM metadata.

The **DB2** On the one hand, technology meets the increased scalability requirements of modern environments, on the other hand enables the use of new methods such as online consistency checking, online reorganization and comprehensive index generation and opens up additional degrees of freedom for the TSM architecture for future extensions.

The port to the new database technology was planned and implemented in close cooperation with the DB2 laboratory in Toronto and the developers were able to port all specific functions that were technologically linked to the original database into the new DB2 environment.

Deduplication is in the opinion of self-proclaimed industry experts "... the biggest thing in backup since the incremental." (W. Curtis Preston in February 2009 in "TSM
6.1, dedupe & DB2 ”at www.backupcentral.com). Basically, deduplication is only a new method of data compression with the property of breaking down data objects into blocks and allowing comparative operations to be carried out over these blocks with the aim of detecting redundancies. Redundant blocks are referenced in the metadata and then deleted. The amount of redundant blocks deleted is then the measure of the memory saving (= compression factor = DedupRatio).

TSM V6.1 optionally provides deduplication functions for its disk pools and works asynchronously, ie the backup data is first saved unmodified, then analyzed and deduplicated in subsequent processes. It is planned to extend the procedure in subsequent releases of TSM V6.1 to the TSM clients, which will distribute the calculation load and, above all, reduce the amount of data that the clients need. B. must be sent over public networks. The process is extremely efficient because the comparison operations take place on the central TSM backup server and the data volume of all clients is used as the basis for the redundancy detection.



TSM V6.1 new layout of the B / A client

Summary and closing word:

Here is a brief summary of the historical development of Tivoli Storage Manager from 1990 to 2009:

- *Nov. 1, 1990 - WDSF / VM R1 with clients on OS / 2, DOS and AIX*
- *November 1991 - WDSF / VM R1 Enhanced - Sun, Apple clients*
- *May 1992 - DFDSM / VM and DFDSM / MVS*
- *July 29, 1993 and November 16, 1993 - IBM ADSTAR Distributed Storage Manager V1.1 and V1.2 with OS / 2 and AIX as server platforms*
- *Spring 1995 - ADSM V2.1 - OS / 400, VSE, Sun and HP as platforms for ADSM servers and a variety of supported client platforms, DRM and copypool*
- *October 1997 - ADSM V3.1 - ADSM server on Windows NT, server-to-server communication, Web Admin GUI, Web GUI for clients, SQL queries and integration in systems management applications*
- *January 1999 - IBM ADSM Marketing moves to Tivoli, new name Tivoli ADSM*
- *September 1999 - Tivoli Storage Manager (TSM) V3.7 replaces Tivoli ADSM, LAN-free, tape library sharing*
- *July 2000 - TSM V4.1 - subfile backup, Linux clients and Windows 2000 support*
- *June 2001 - TSM V4.2 - Journal-based backup, LAN-free backup for the backup archive client, NDMP and SANergy*
- *March 2002 - TSM V5.1 - Windows server-free backup and online Image Backup / Open File Backup for Windows 2000 with LVSA*
- *October 2002 - TSM V5.1.5 - Linux / x86 and Windows 2000 as server, server-to-server import / export, TSM on OS / 400 PASE*
- *July 2003 - TSM V5.2 - TSM server on zLinux, TSM / zOS LAN-free, TSM for Data Retention, NDMP TOC*
- *December 2004 - TSM V5.3 - Introduction of ISC as Admin console, AES128-bit, proxy node support, group collocation*
- *January 2006 - TSM V5.4 - NDMP filer-to-server, Data shredding, active data pool, TSM for copy services*
- *November 2007 - TSM V5.5 - VMware Consolidated Backup*
- *July 2008 - TSM Fastback*
- *April 2009 - TSM V6.1 - DB2 as a database, Storage pool deduplication*

The TSM version and function list of the last 20 years is also a reflection of the technological development in IT. Data processing has developed in comparable phases, from mainframe-centric IT to the era in which distributed processing on as many heterogeneous systems as possible seemed to be the principle, to centralized structures on virtualized hardware.

The requirements for which TSM was originally designed have not changed over this period: data backup is a necessary evil that causes costs and requires constant maintenance, but to which there is no alternative. Therefore, I am sure that Tivoli Storage Manager will continue to be used and will be successful. In November 2010, IBM celebrates Tivoli Storage Manager

20th birthday, plus my best wishes for a long and successful product life.

1. BRMS overview

The "Backup, Recovery and Media Services for IBM i" (BRMS) license program enables the implementation of a backup strategy for IBM POWER systems with the IBM i operating system (formerly i5 / OS).

Complex scenarios can be mapped fully automatically, such as: B. the online backup of LotusSer vern (Domino) in an i environment. This means that BRMS offers more options than when using the backup commands from the native operating system IBM i.

The current license program 5761BR1 "Backup, Recovery and Media Services for IBM i" for version i 6.1 consists of three chargeable features: * **BASE, Network Feature, Advanced Feature,** which are presented in Figure 1 and are explained in detail below.

Useful links: IBM i InfoCenter - BRMS http://publib.boulder.ibm.com/infocenter/iseries/v6r1m0/ index.jsp? topic = / rzai8 / rzai8overview.htm

BRMS HomePage http://www03.ibm.com/systems/i/support/brms

2. BRMS - current features

Most IBM i customers using BRMS have "only" the * BASE feature licensed from 5761BR1. With this feature * **BASE** are four extensive modules

Backup, recovery, media management and **Tape library support** available, see picture 2.

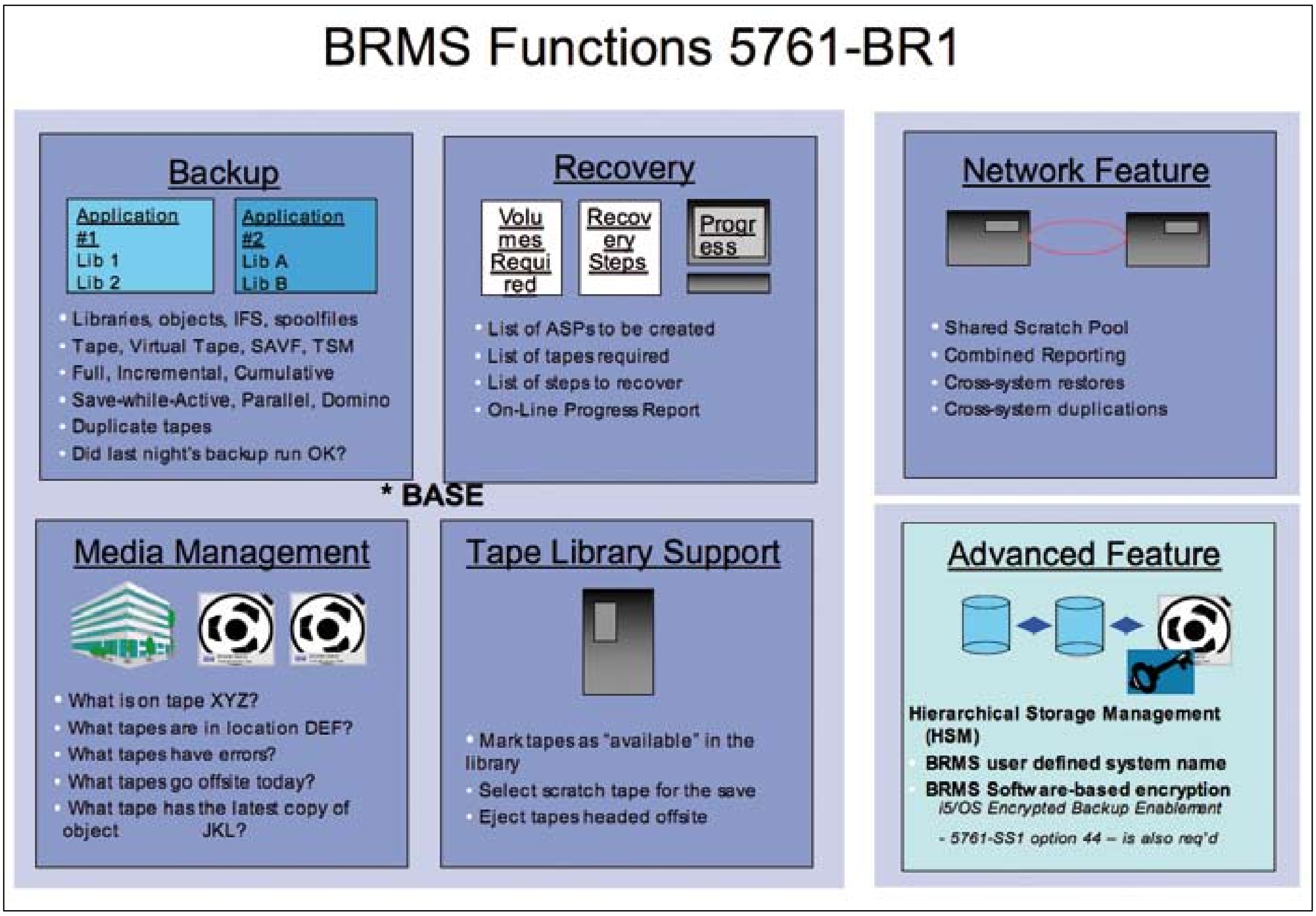
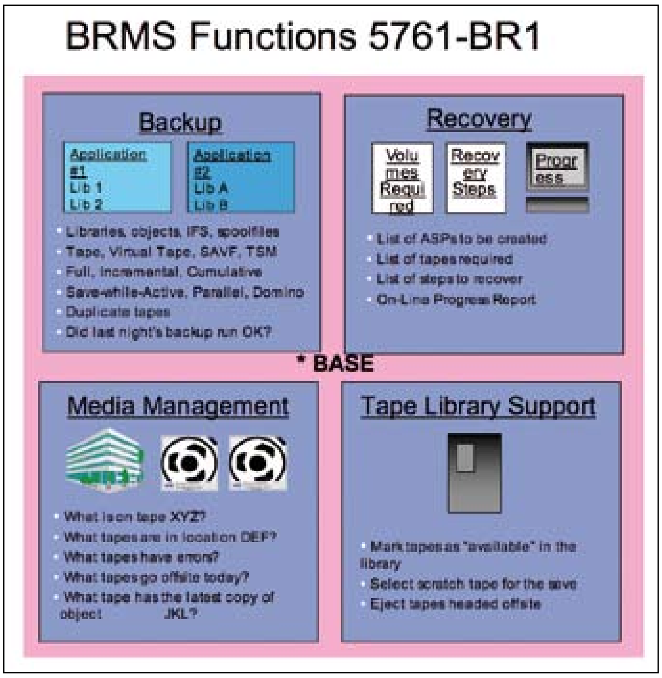


Image 1



picture 2

Further details on the four modules mentioned in Figure 2 **Backup, recovery, media management** and **Tape library support** of the * BASE feature see:

http://www.redbooks.ibm.com/redbooks/pdfs/sg244840.pdf

http://publib.boulder.ibm.com/infocenter/iseries/v6r1m0/ topic / rzai8 / sc415345.pdf

With i 6.1, three different interfaces (interfaces) are available for BRMS for implementation and management:

- ***traditionally "green screen" (5250 interface), see*** *picture 3*
- ***BRMS PlugIn in the System i Navigator, see*** *Image 4*
- ***BRMS PlugIn in the IBM Systems Director Navigator for IBM i (formerly i5 / OS), see Image 5***

The use of the 5250 interface (**Picture 3)** has two advantages:

- ***a distinctive menu structure, useful for BRMS "Beginner"***
- ***a variety of BRMS commands that can easily be Have the customer's existing programs integrated, which makes sense for the BRMS "experts"***

picture 3

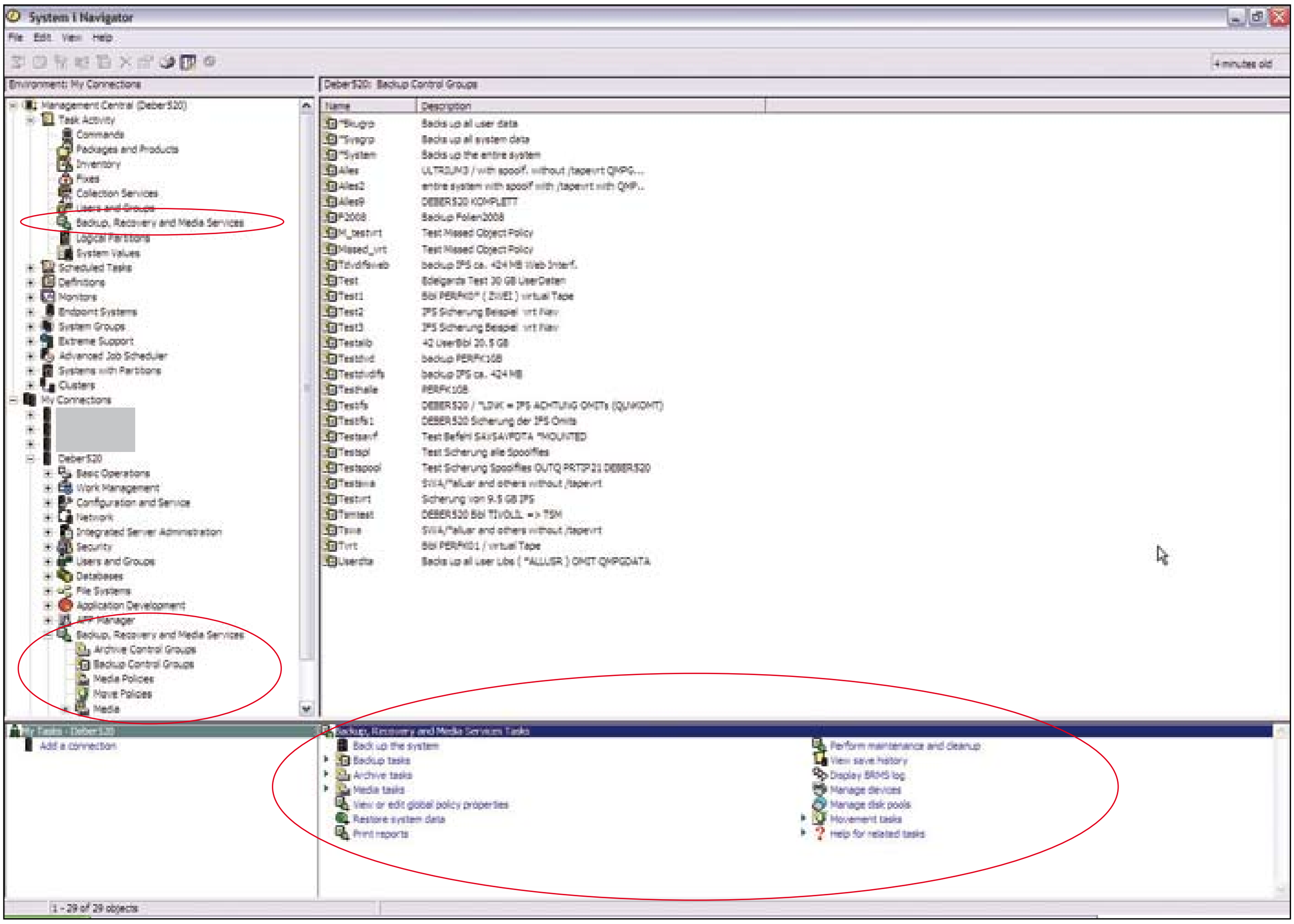


Image 4

The use of the two in **Image 4** and **Image 5** The presented GUI interfaces for BRMS have the advantage that GUI-oriented users and / or users who are not very familiar with the IBM i operating system can also work with BRMS.

GUI = Graphical User Interface

in the **Image 4** are a few examples of so-called "backup control groups". They are definitions for automatic backups. The scheduling for the backups is not visible here, but is available in the background.

Image 4 was created with a System i Navigator installation on a laptop with English Windows XP. System i Navigator is also available in German.

The examples for the backup control groups are from a relatively small IBM i partition (name "DEBER520") of a POWER system with approx. 700 GB disk capacity. Of course, BRMS is also used by IBM i customers in the use of larger partitions with several TB disk capacity.

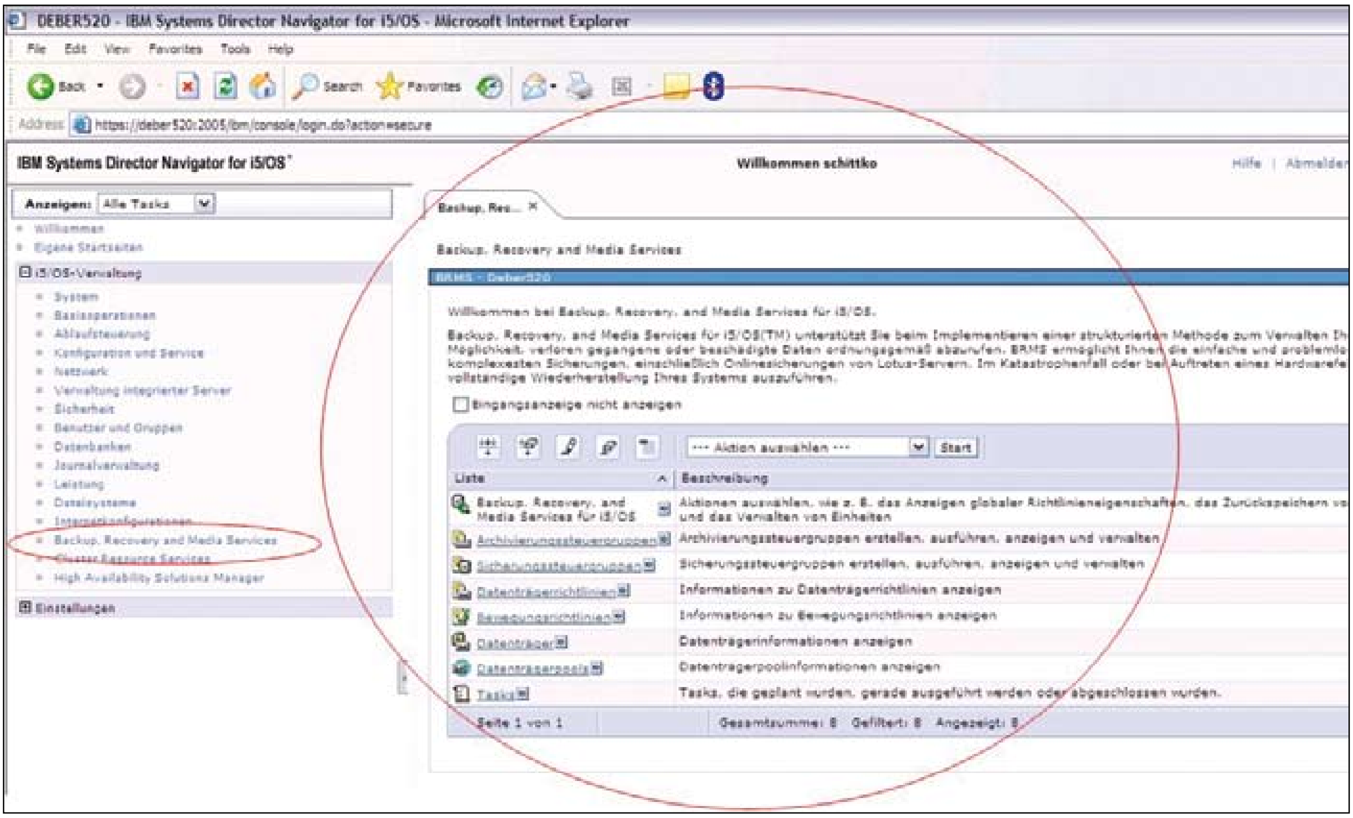


Image 5

New with IBM i 6.1 is the web browser interface "IBM Systems Director Navigator for i" (former name "... for i5 / OS"). With the previous release IBM i 5.4 (formerly i5 / OS V5R4) this interface was not yet available.

In this web browser interface with i 6.1 there is also a BRMS plug-in, see **Image 5.** Details on **BRMS plug-in** in "IBM Systems Director Navigator for i" see Chapter 12.1 in the red book "IBM Systems Director Navigator for i (SG24778900)": http://www.redbooks.ibm.com/redbooks/pdfs/sg247789.pdf

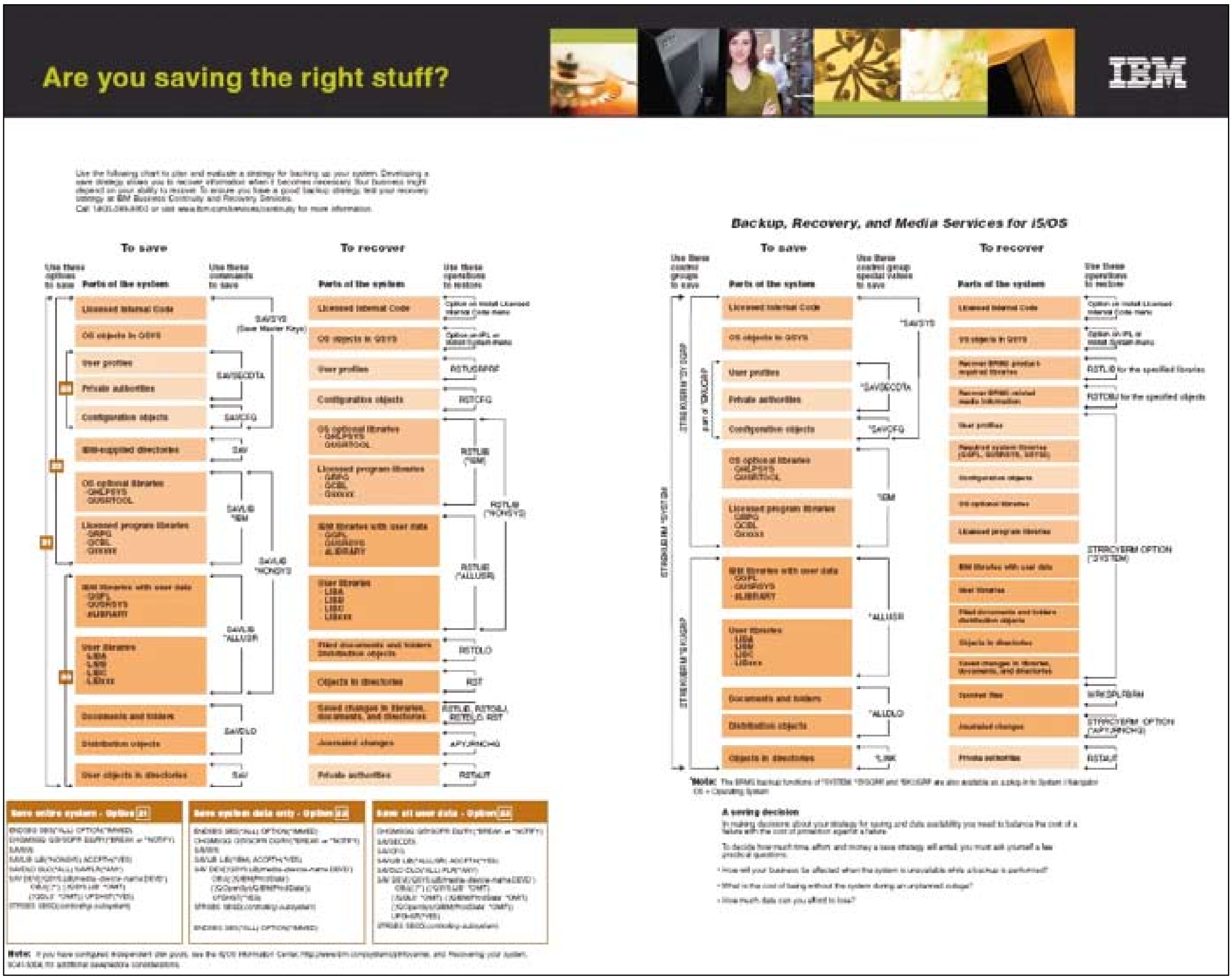


Image 6

In the above **Image 6** you can see a comparison between the options (commands, menus, backup control group special values) for saving (to save) and saving back (to recover) in a native IBM i operating system environment and a BRMS environment.

The picture is taken from the poster "Are you saving the right stuff?" (IBM Form Number G325632805, http://publibfp.dhe.ibm. Com / epubs / pdf / 32563285.pdf).

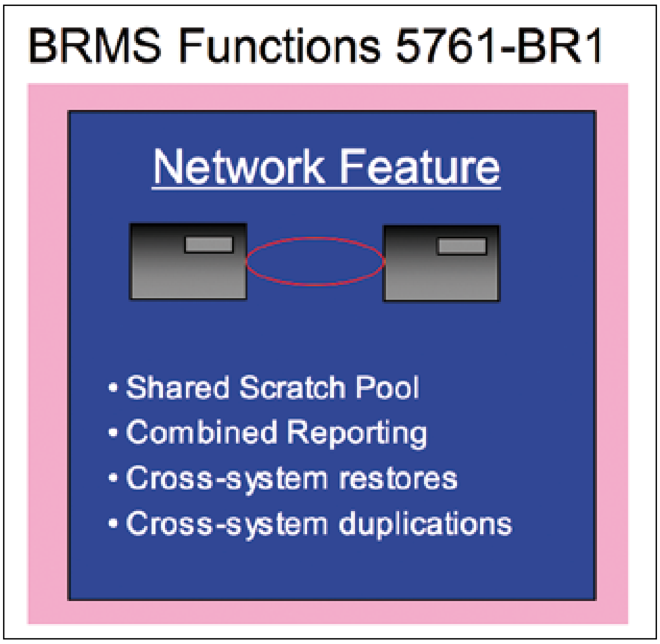


Image 7

The licensed BRMS **Network feature** makes it possible - see **Picture 7** to use a common tape library with several shared partitions or POWER systems with IBM i and BRMS, with a common tape pool and a certain number of tape drives. Each partition / system must be connected to the tape library.

As a result, fewer cartridges and fewer tape drives are generally required than with dedicated allocation. In addition, system-wide restores (restores) or duplications can be carried out. B. is common in two systems, where one is the production system and the other the backup system. With partitioned systems, the BRMS Network can also be set up between the partitions. Here it is useful to use the virtual Ethernet available in all partitioned POWER systems as an internal TCP / IP interface for the BRMS network.

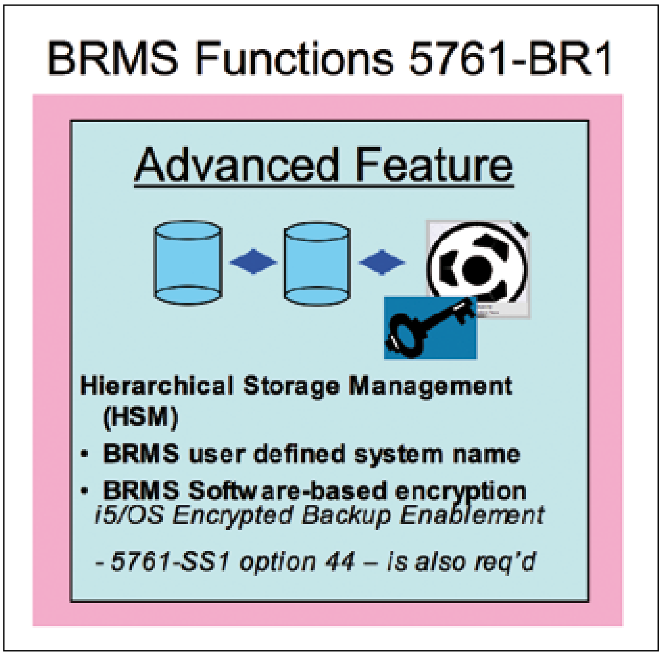


Image 8

The licensed BRMS **Advanced feature** - please refer **Image 8** - In addition to HSM (Hierarchical Storage Management) since March 21, 2008 in version 5761BR1 with i 6.1 there are two more options, on the one hand to assign an individual BRMS system name ("BRMS user defined system name") and on the other hand the software-based encryption (" **BRMS software-based encryption** ").

While HSM was hardly widespread in the BRMS and thus the previous BRMS Advanced Feature, the use of the two new functions in the BRMS Advanced Feature with i 6.1 will certainly increase. So z. For example, the option of assigning an individual BRMS system name in connection with high availability scenarios is required.

Note: The so-called "hardware-based encryption", also referred to as " **Tape drive encryption** "Has been supported as" Library Managed Encryption "in the BRMS since October 2006, see http://www03.ibm.com/systems/i/support/ brms / tapeEncrypt.html

The BRMS features described above also include support for so-called "special environments", listed below with further links:

- ***BRMS Application Client to an IBM TSM Server,***
  see Chapter 9 in
  http://www.redbooks.ibm.com/redbooks/pdfs/sg247031.pdf BRMS and
- ***FlashCopy, see Chapter 15 in***
  http://www.redbooks.ibm.com/redbooks/pdfs/sg247103.pdf BRMS and
- ***WORM tapes, see Chapter 6 in***
  http://publib.boulder.ibm.com/infocenter/iseries/v6r1m0/ topic / rzai8 / sc415345.pdf BRMS and SAP, see Chapter 23 in
-
  http://www.redbooks.ibm.com/redbooks/pdfs/sg247166.pdf BRMS and Domino,
- ***see among others***
  http://www-03.ibm.com/systems/i/support/brms/domino.html

New with IBM i 6.1.1 (available since October 23, 2009) on POWER6 or POWER6 + systems or corresponding POWER blades there is one **NPIV** Support (NPIV = N_Port ID Virtualization) for certain tape libraries (tape robots), see Informational APAR II14526:

http://www912.ibm.com/n_dir/nas4apar.nsf/c79815e083182f ec862564c00079d117 / 02327e067599839e86257659003c6d 66? OpenDocument & Highlight = 2, NPIV

These tape libraries with their LTO4 drives are fully supported in the BRMS, although they can only be accessed via a virtual fiber connection from the IBM i 6.1.1 partition.

3. BRMS - history

BRMS was first announced on September 1, 1992 (EMEA announcement letter ZP920587) under the former name "Backup, Recovery and Media Services / 400" with license program number 5798RYT with general availability on January 15, 1993. The "license program" BRMS had the "exotic" license program number 5798RYT in 1992/1993 because at that time it was still being marketed as a result of an alliance with the company MerschBacher & Associates, Inc. and was not a "full" license program, but "only" an LPO (Licensed Program Offering).

On May 3rd, 1994 BRMS became a "full" license program 5763 **BR** 1 (Version 3, EMEA Announcement Letter ZP940348) announced with general availability on
November 25, 1994 and thus integrated into the normal licensed program procedures from IBM Rochester for (at that time) AS / 400 systems with OS / 400 operating system. From this point on, BRMS has the "normal" license program designation 57xx **BR** Get 1.

With EVERY operating system version of OS / 400 or i5 / OS or IBM i, a new version of BRMS was released in 1994 with various improvements and functional expansions. Here is a link to an overview of all operating system versions of OS / 400 or i5 / OS or IBM i since 1988. The period from 1994 onwards is important for BRMS:
http://www947.ibm.com/systems/support/i/planning / upgrade / suptschedule.html

The exact features of the current BRMS with IBM i 6.1 (5761BR1) can be found in the manual:
http://publib.boulder.ibm.com/infocenter/iseries/v6r1m0/ topic / rzai8 / sc415345.pdf

4. BRMS - summary and outlook

BRMS is a powerful tool for IBM i environments to implement complex backup requirements and to use the six different modules (backup, recovery, media management, tape library support, network feature, advanced feature) for various restores including the complete restoration of the entire system from a backup to be "armed" as well as to take into account special requirements from customer-specific implementations (e.g. Domino, SAP, FlashCopy).

In the case of a complete restoration of the entire system, the so-called “recovery report” is used, a manual that is created automatically after the last backup about the restore steps to be carried out.

The (*) announcements for POWER7 and IBM i 7.1 planned for 2010 and the extensions planned with them, such as the expansion of the "IBM i Network Upgrade" that is already possible with IBM i 6.1.1, also for a scratch installation of a system with IBM i, will also be available BRMS have an impact.

One can count on gratifying innovations for BRMS in 2010.

Already existing links (*):
www.ibm.com/systems/power/hardware/sod.html
www.ibm.com/systems/power/hardware/sod2.html

(*) Disclaimer: All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Any reliance on these Statements of Direction is at the relying party's sole risk and will not create liability or obligation for IBM.

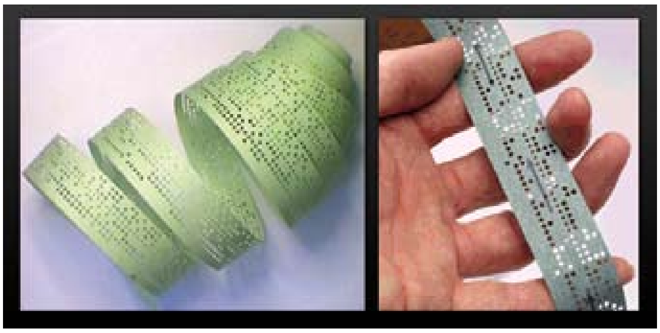
Technology appendix

The predecessor of today's mass storage

Punch cards

Punch cards as mass storage existed long before the computers. Hollerith's census machine already used punch cards in 1891. Punched cards represent binary data in which a certain position on the card was punched (or not punched). With the advent of computers in the 1950s, this technique was mastered very well: the cards could be sorted, counted and evaluated in every conceivable form.

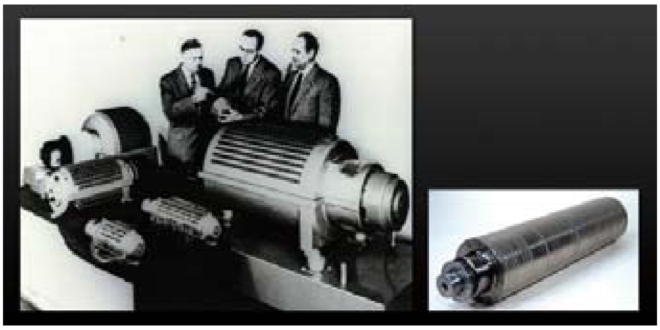
The cards were made using punch cards. With the ready-coded cards, the computer was fed via a reader and the desired program and processing data were fed to it. Instead of individual cards, perforated strips were later used, which allowed a higher mechanical processing speed.



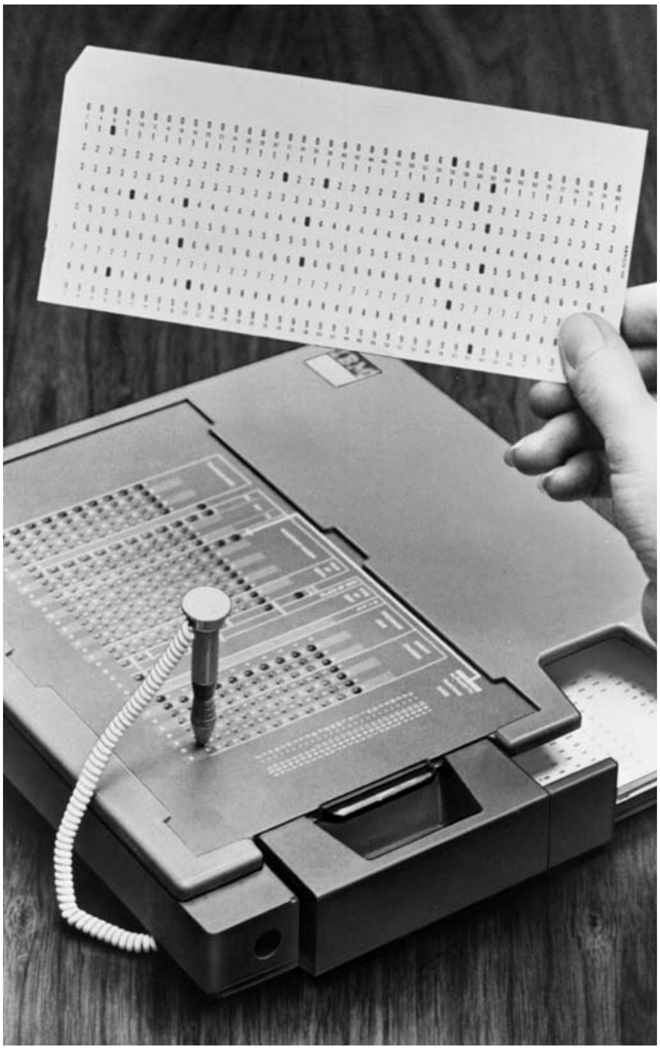
punched tape

Drum storage

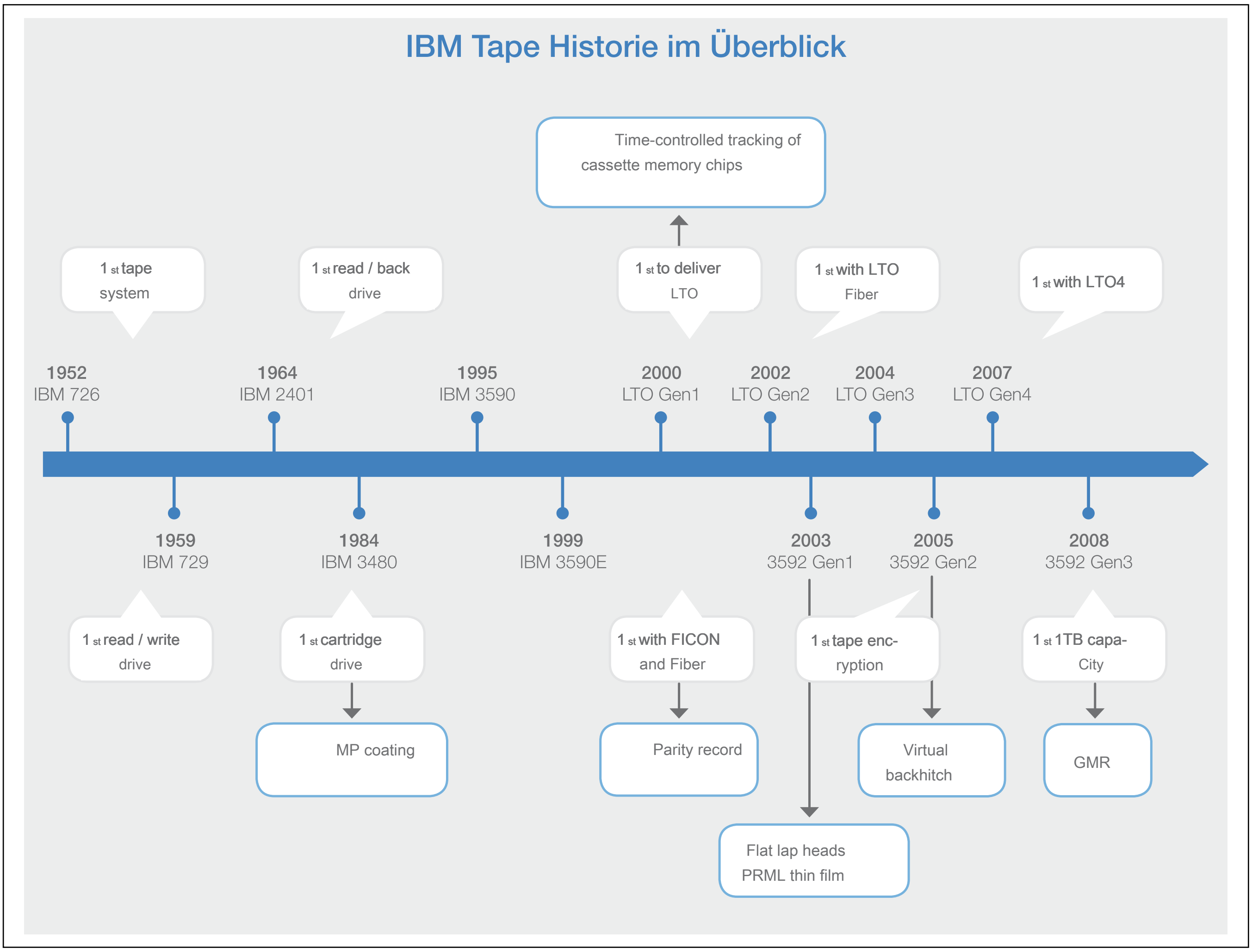
The first electrical mass storage devices that were used in connection with computers were the so-called drum storage devices in 1947/48. A coating of a magnetic material was applied to a solid drum. This could be magnetized inductively by a current or generated an induced current when read out. The smaller dimensions of the drum storage also reduced the size of the computers at that time considerably. From 1950 onwards, drum storage devices were occasionally used to load and save programs or data.



Drum storage 1947



Punch cards



Magnetic tape

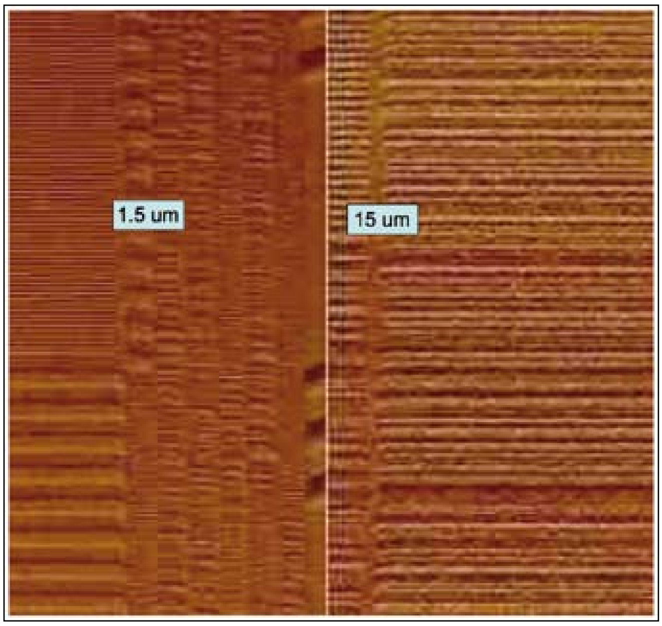
From 1952 the magnetic tape (IBM 726, IBM 727) replaced the punch cards as mass storage. It was a 12.4 mm wide plastic tape with one side covered with a magnetizable coating. The data record was that the ferrite strips contained on the tape were magnetized or not, thus representing the information in the binary system required for the computer. So-called parity bits could later be accommodated on an additional tape track, which ensured high data security. In addition to a high processing speed, the advantage of the magnetic tapes was above all the high storage capacity. However, data on a tape could only be stored and read again sequentially (i.e. one after the other).

still used today. In the future, the magnetic tape is likely to experience a renaissance due to the energy efficiency problem and the "Green IT" discussions. Magnetic tapes do not need electricity. Tape is "cool"!

As a medium, tape has the longest history compared to all other storage media. Today, tape as a medium offers the highest development potential in terms of capacity and performance in the next few years compared to all existing technologies such as disk, optical memory, RAM, flash memory or PCMs. This was achieved with the consistent technical development of tape (see picture "IBM tape history at a glance"). MP coating (metal particles) was already introduced with the first cassette technology, the IBM 3480. With the 3590, parity recording was introduced in 1999, wherever the associated parity block was also recorded after seven data blocks, which is a much higher error correction enabled. With the introduction of LTO in 2000, something very new came about: the time-controlled tracking via servo tapes, which allows the read / write heads to be guided precisely over the track set - even with very thin tracks. This led to a capacity and performance increase of a factor of 20 in just 10 years (from 2000 to 2010). With the introduction of 3592 technology (Jaguar) in 2003 for the enterprise sector, technologies from disk technology were adopted in the tape drives. A new head technique is used which, in conjunction with the first thin film coating, allows recording on the medium itself with significantly higher induction on the heads and thus reflecting much more stable bits on the medium. PRML was first introduced on the plates in 1991 and has entered tape technology with the 3592 (see also PRML in the technology appendix). New functions such as Virtual Backhitch come into play that keep the tape in the drive “streaming”, even if only small amounts of data are transferred to the drive.

The year 2008, however, makes a big story in the history of tapes. The Jaguar 3 drive, now called the TS1130, is the first to use GMR read heads (GiantMagnetoResistance). GMR read heads were first used on the 1997 disks and without this technology, our current hard disk capacities would never have been possible. Now the basis of GMR has also been introduced for tape, a basis that creates completely new capacitive options. The introduction of thin film coating and highly inductive write heads allows extremely thin traces to be recorded. This is done in such a way that a relatively thick track is generated in a highly inductive manner from the beginning of the tape to the end of the tape. When writing back from the end of the tape to the beginning of the tape, part of the written track is overwritten again. With this overwrite technique, which is also known as "shingling", it is possible to create wafer-thin traces. The problem now is reading out. How can the small magnetic stray fields be read out again on such a thin track? The solution for this is the GMR reading heads that have now been introduced. GMR read heads can easily pick up and read out the smallest stray fields even at extremely high speeds. This creates a new basis for tape development. GMR read heads can easily pick up and read out the smallest stray fields even at extremely high speeds. This creates a new basis for tape development. GMR read heads can easily pick up and read out the smallest stray fields even at extremely high speeds. This creates a new basis for tape development.

In May 2006, the use of GMR reading heads in a prototype based on 3592 Jaguar technology and a recording density of 1 gigabit per cm were presented in the IBM laboratory in Almaden, California 2nd reached. The graphic "GMR use with tape" shows the tracks created on the tape with GMR with a track width of 1.5 µm compared to an LTO3 track with 15 µm.



Almaden: May 16, 2006, GMR use with tape 1 gigabit per cm^2, 8 TB per cartridge

Magnetic disks

Magnetic disks are the successors of the voluminous drum storage and are considered the direct forerunners of today's hard drives. On September 13, 1956, IBM presented the first magnetic disk called RAMAC 350 (Random Access Method of Accounting and Control) as part of the RAMAC 305 system and with a capacity of 5 MB to the public. This capacity was distributed over 51 disks, each 24 inches (60 cm) in diameter. The information was scanned by a floating write / read head (see under RAMAC 305).

Development of the IBM System 305 and the first magnetic disk storage IBM 350

At the beginning of the 1950s, the IBM employee Reynold B. Johnson was given the task of designing a computer system that was able to process every business transaction when it occurs with an output of around 10,000 cases per day (without batch processing, as is the case) punch cards and punched strips used until then). Reynold B. Johnson put together a unique team of developers and experts to start this difficult task. In order to have the best possible environment for this task, which should generate as many innovations as possible, he started the project in 1952 in San Jose, in sunny California. At that time the project was called "source recording". Johnson's team included AJ Critchlow, WA Goddard,

JW Haanstra, HP Luhn, J. Ratinow and CD Stevens. Just one year later, in 1953, Johnson made the decision that the data storage device should be a file system using magnetic storage disks.

At that time, the development work for the IBM 726 magnetic tape unit and the Reynold team had just been completed

B. Johnson quickly realized that magnetic tape units could work very quickly in the sequential range, but were slow when it came to direct access. The access times were in the range of minutes and therefore the tapes were not suitable for immediate access in very specific applications. A new storage system that could meet these requirements had to be developed. The solution consisted of 51 aluminum plates coated on both sides, which were mounted on a vertical shaft and rotated at a speed of 1200 revolutions per minute. One was content with a pair of read / write heads that were positioned two-dimensionally by means of a steel cable drive according to the content of an address register. This process took a maximum of 800 ms. This fulfilled the requirement to process around 10,000 transactions a day, each immediately upon arrival of the case. In 1956 the time came that the new RAMAC 305 computer system with the associated RAMAC 350 disk memory could be brought onto the market.

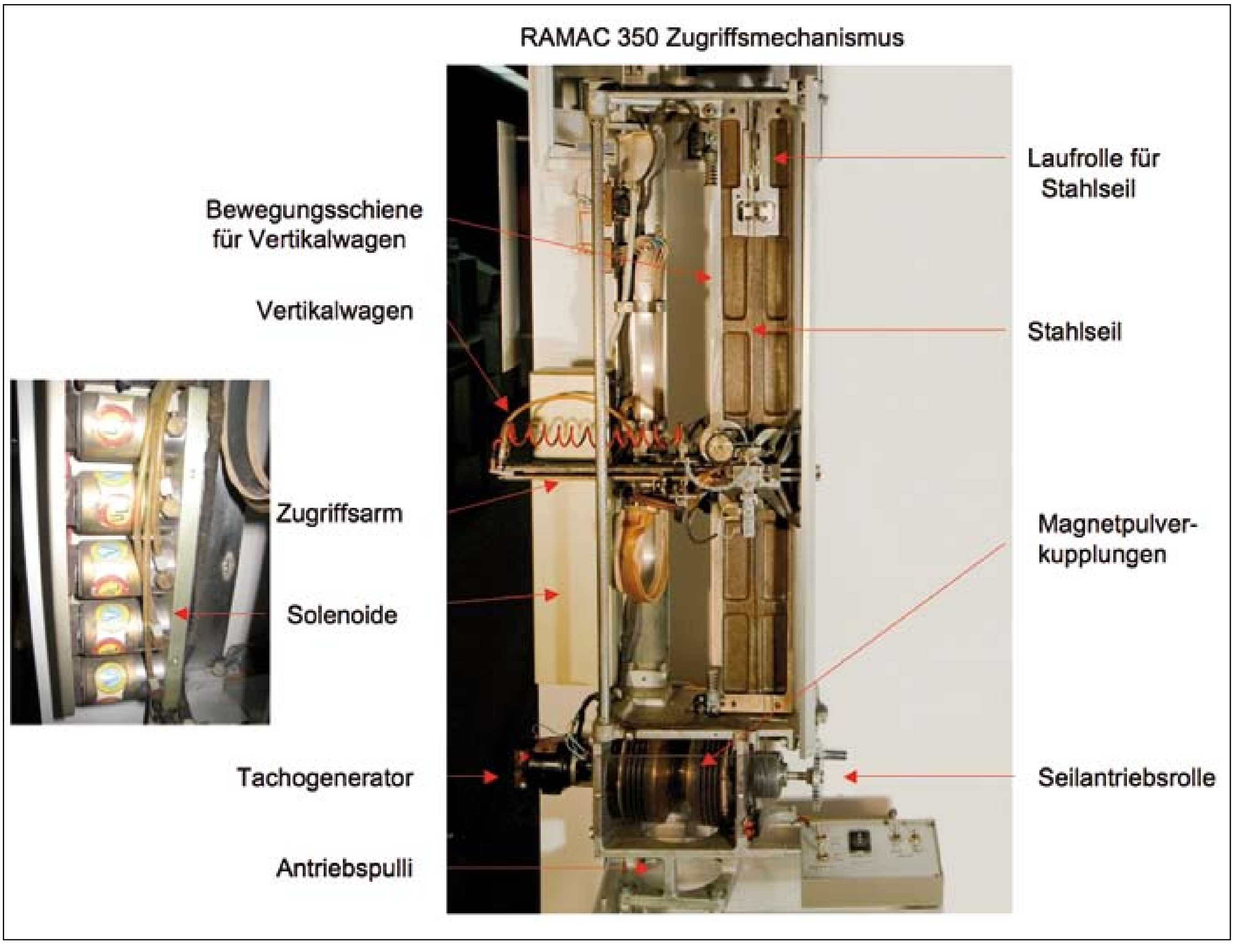
RAMAC 350 as part of the RAMAC 305 computer system

Reynold B. Johnson, nicknamed "Rey", was 90 until he retired in 1971

US patents pending, almost exclusively patents related to plate development.

With the RAMAC 350, the first magnetic disk memory saw the light of day and further developments of the disk memory followed without interruption for decades with the constant goal of achieving capacity increases, shorter access times, higher data rates and lower costs per stored megabyte. This course of development was made possible by the **American Society of Mechanical Engineers** in **1984** recognized by using the RAMAC disk storage for **"International Historic Mechanical Engineering Landmark" (** Landmark stands for Milestone), in recognition of its ongoing impact on storage technology.

Reynold B. Johnson



How the IBM 350 (1956) and IBM 355 (1957) Magnetic Disk Storage Works

How does the RAMAC 350 magnetic plate work and how is it positioned? The processor uses the search command to provide the address to be controlled. This address is set in address relays, then a start signal ensures access control. Starting from the actual position of the vertical carriage, the target position must be approached by activating one of the two magnetic powder clutches (see figure "RAMAC access mechanism").

At the very bottom, the drive motor hangs with a V-shaped drive sweater, a steel cone with a rubber surface. This cone drives two magnetic powder clutches, ie one runs to the right, the other to the left. The two clutches run on a shaft. If neither of the two clutches is activated, the output cable drive pulley stops (bottom right). If one of the clutches is activated, the vertical carriage that moves the access arm (center) moves up or down, depending on which coupling has been activated.

The address relays represent the target position as the target position. The contact from the vertical carriage to one of 50 contacts (figure "RAMAC 350 vertical positioning", copper-colored contact strip) shows the vertical carriage actual position. An ohmic measuring bridge determines the size of the difference between the actual and target position and generates a corresponding start signal. The vertical carriage is now either moving up or down.

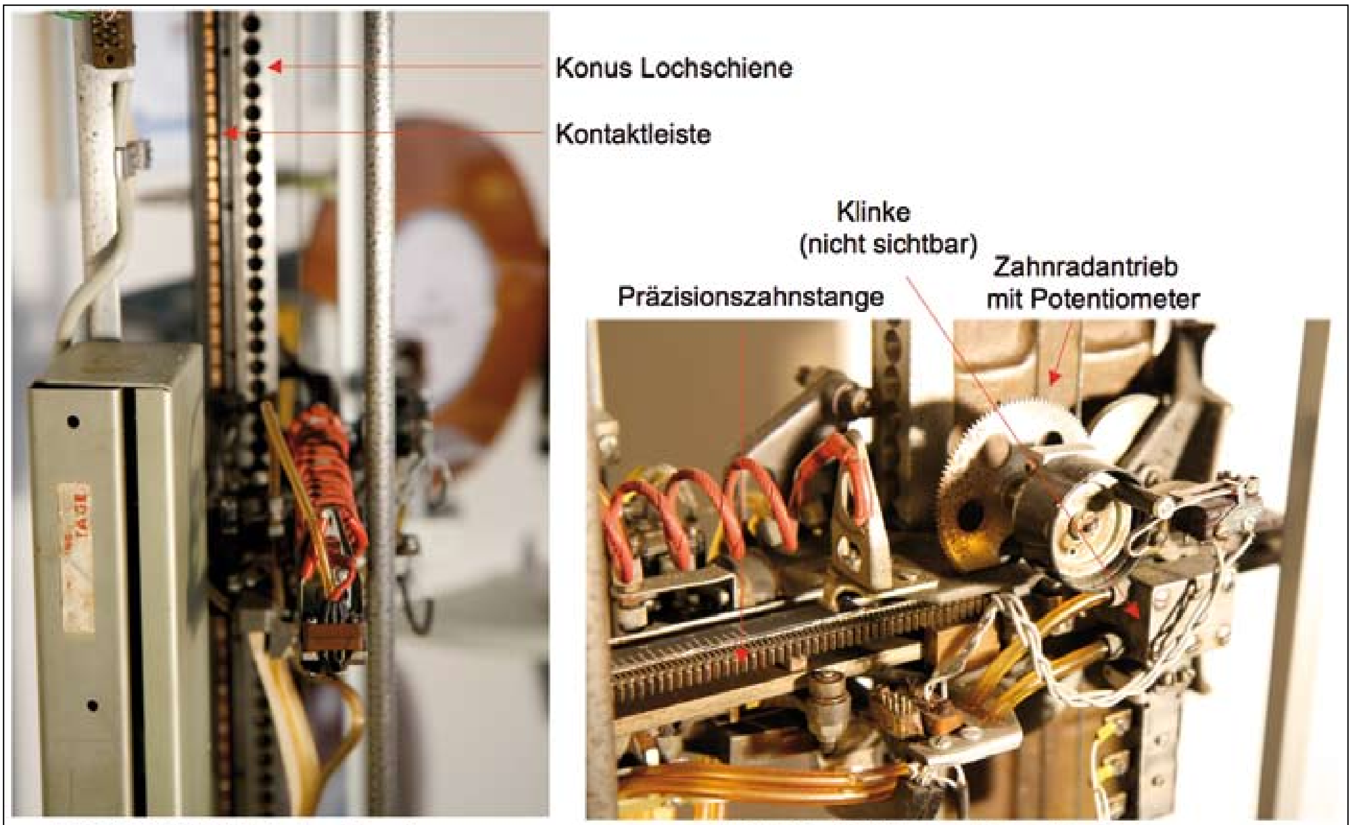
As a result, the actual resistance in the measuring bridge changes continuously until it reaches zero (bridge equality). If this is the case, the driving clutch is de-energized (either one or the other) and the movement is stopped. This completes the vertical movement. In order to maintain this position mechanically precise, a precision cone is driven by compressed air into one of the cone holes for locking (Figure "RAMAC 350 vertical positioning", perforated rail on the right next to the copper-colored contact strip).

The vertical carriage is thereby firmly locked on the V-shaped movement rail.

In order to control the forces that occur during the stopping process, a tachometer generator runs (figure "RAMAC 350 access mechanism" below left), which shows the speed in the form of the voltage output. The closer the vertical carriage comes to its target position, the less coupling current is available and the movement slows down.

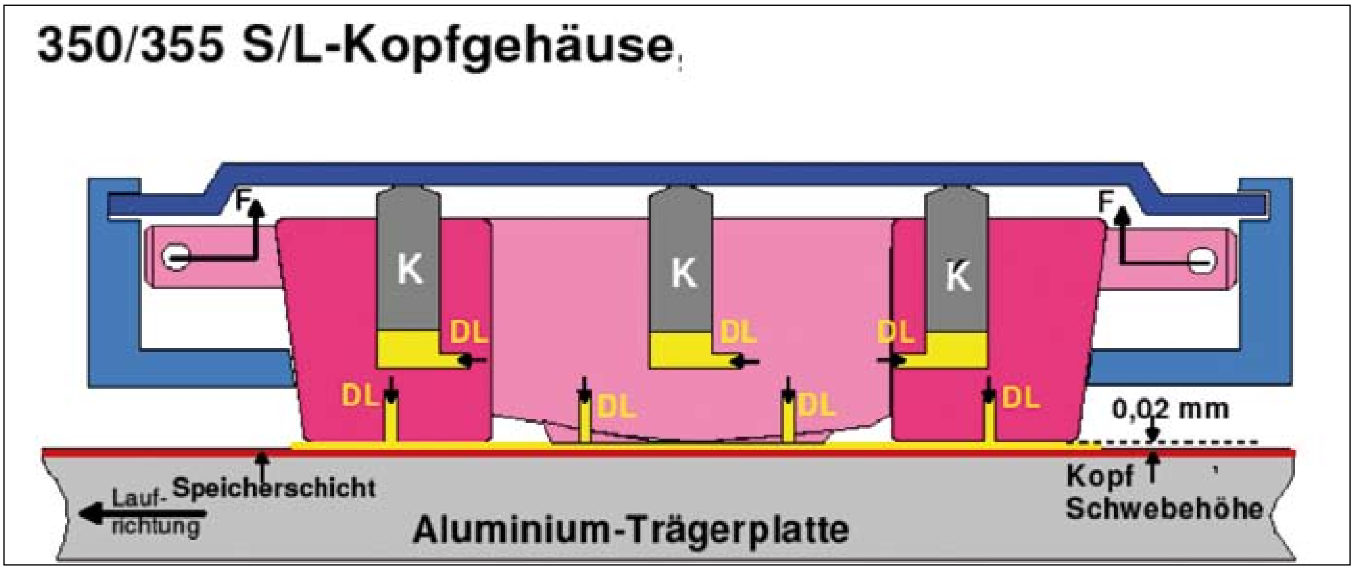
Once the search position has been reached, ie vertical locking by the precision cone in the perforated rail has taken place, the control signal for the phase of the horizontal movement of the access arm is started (Fig. "RAMAC 350 horizontal positioning").

The horizontal positioning is similar to the vertical positioning. The position of the access arm is accompanied by the potentiometer (gear drive) visible at the top right and supplies the target actual difference for the further sequence control via the measuring bridge. When the target position (track address) is reached, the precise track position is locked onto the precision rack by means of a pneumatically operated pawl. This completes the mechanical positioning in both access axes. An "End of Access" signal controls the pneumatic extension of the S / L head to the level of suspension. The precision rack contains 100 notches where the pawl can snap into place because the plate reflects a total of 100 tracks.



RAMAC 350 Vertikalpositionierung

RAMAC 350 Horizontalpositionierung



How did the S / L head really hover?

The head is in the non-activated state in the retracted position due to spring force (see Figure 350/355 S / L head housing, F left and right). In this position, the head is at a considerable distance from the top surfaces. When the charging valve is activated, compressed air (see picture "DL") flows into the three cylinder chambers of the pistons K. This moves the head down towards the plate surface (storage surface). The controlled compressed air emerges simultaneously from six holes arranged in a hexagon (only four visible in the picture) with a diameter of only 0.1 mm. As a result, the head cannot be pressed deeper because the air cushion under the six holes can only be raised to a height of 0.02 mm (due to the design). This floating state of the head at this height is continuously maintained, even if there are slight fluctuations in the plates with a diameter of 24 inches (approx. 60 cm) in the outer area (Bernoullisches law). The S / L head is now ready to read or write.

The solenoids (see figure “RAMAC 350 access mechanism”, small picture on the left, 5 pieces) use compressed air to control the locking and releasing of the carriage locking bolt during vertical positioning and the compressed air locking arm on the precision rack at the Hori

zonal positioning. They also control the compressed air supply to the 350/355 head housing to extend the heads to the plate at a floating distance. Solenoids are tubular, cylindrical metal coils that act like a bar magnet when current flows through them. In the case of the RAMAC 350/355, solenoids work as fast needle valves for compressed air, which are opened by a magnetic field and closed by spring force.

Note: There are often irritations as to whether the RAMAC 350 plates had 50, 51 or 52 plates! The RAMAC 350 used 50 plates, i.e. 100 plate surfaces, for recording. The stack itself consisted of 51 plates. The two outer surfaces of the two outer plates (at the top and at the bottom) were not used for the recording. In the RAMAC 355, which came on the market a year later, in 1957, the disk stack comprised 52 disks, 50 for recording. The outer plates only served to stabilize and protect the 50 inner plates that were used for recording.

The compressed air-controlled floating head only learned to fly in 1961 and was introduced with the product IBM 1301 (see also under IBM 1301). “Runways” were introduced in the outer area of the plates and the head was designed so aerodynamically that it reached its flying height at a revolution of almost 1,800 revolutions per minute, which at that time was 0.0635 mm.

Note: In the appendix to the chapter on magnetic disks you will find an interview that the author conducted with the IBM RAMAC 350/355 specialists Hans W. Spengler and Heinz Graichen from the IBM Museum in Sindelfingen. Here the IBM RAMAC 350/355 access is explained in all technical details.

Information technology still dominates magnetic disk storage, although optical storage in the form of rewritable compact discs has also conquered a significant share of the mass storage market. Magnetic disk storage all work according to the same basic principle: A round disk serves as storage medium, on which there is a layer of hard magnetic material (previously different ferrites, now thin film). The plate is divided into concentric tracks. A movable magnetic head is moved radially over this plate in order to magnetize the needle-shaped ferrites on the individual tracks so that binary data are imaged. It is also able to quickly switch from one track to the other by moving the drive arm. The tracks are divided into sectors.

In principle, magnetic disk storage devices are designed for random access. This means that the medium is not - as e.g. B. with tape drives - must be searched sequentially from the beginning to find a specific location (file). The heads of the magnetic disks - comparable to the tonearm of a turntable and the selection of a certain piece of music - can jump directly to the point at which the desired file is created.

Floppy disks

Alan Shugart, who worked for IBM in the late 1960s, was credited with inventing the 8-inch diskette in 1971. The floppy disk consists of a plastic film which is provided with a non-oriented magnetic layer. Data is recorded either on one side (SS) or on both sides (DS). For protection and better handling, the disc is in a rectangular plastic cover, which is lined with a sliding and cleaning fleece. The cover has openings for the working cone (via which the disc is driven), the index hole and the read / write head. In addition, the case has a recess for setting a write protection. Depending on the system, write protection is set by covering or leaving this recess. The read / write head touches the disk surface when writing and reading, otherwise the head is raised. In 1971, IBM introduced the first 8-inch diskette drive to the public. This is the first time that random access has been achieved with a movable data carrier, because the read / write head can be positioned at any point on the data carrier.

Floppy disks

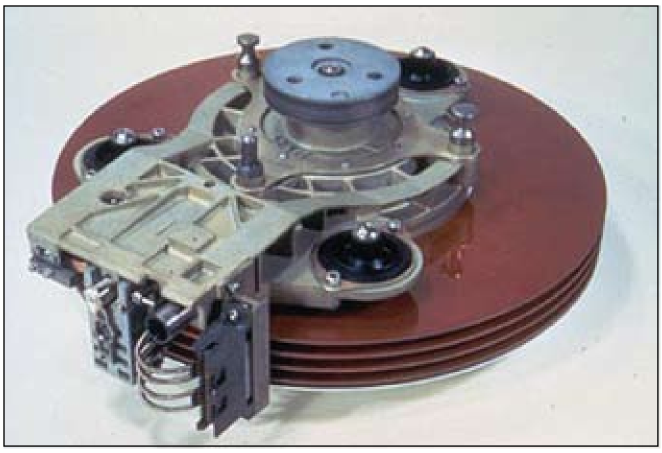
Alan Shugart founded his own company, Shugart Associates, in 1973 to develop and manufacture disk drives. He had long wanted to integrate processors and floppy drives as parts in complete computer systems. In 1976 it was Alan Shugart's company that launched the first

5.25-inch floppy disk drive (customer was Wang Laboratories) on the market. In 1978, ten companies were already producing 5.25-inch drives. In 1980 Sony launched the
3.5-inch diskette to the public. This disk format with the increased capacity of 1.44 MB is still used today in very old PCs. The miniaturization of the floppy disk drives meant that a combination of 5.25-inch and 3.5-inch floppy drives could be accommodated in a half-height housing. Various 3-inch (Amdisk) and 2-inch formats as well as the LS120 diskette (Panasonic, even backwards compatible with the 1.44 MB diskette) and the Zip diskette (Iomega, 1994, 100 MB - 750 MB) could no longer prevail as the standard.

Hard drives

The first approaches for mobile hard drives were seen in 1973. IBM introduced the "Winchester 3340" hard drive in 1973. Capacity: 35 MB. The development name "Winchester" combines the storage capacity of around 30 MB per module and access time of 30 ms with the American rifle type "Winchester 3030". This plate module was easy to transport because it was smaller than the previous IBM 3330 and only weighed about half (see also under IBM 3340).

How do hard drives work now? Actually to this day, the same principle. In an airtight sealed housing (almost airtight, because there is a certain air exchange takes place) several magnetic plates rotating one above the other are mounted. With newer hard drives these are - for

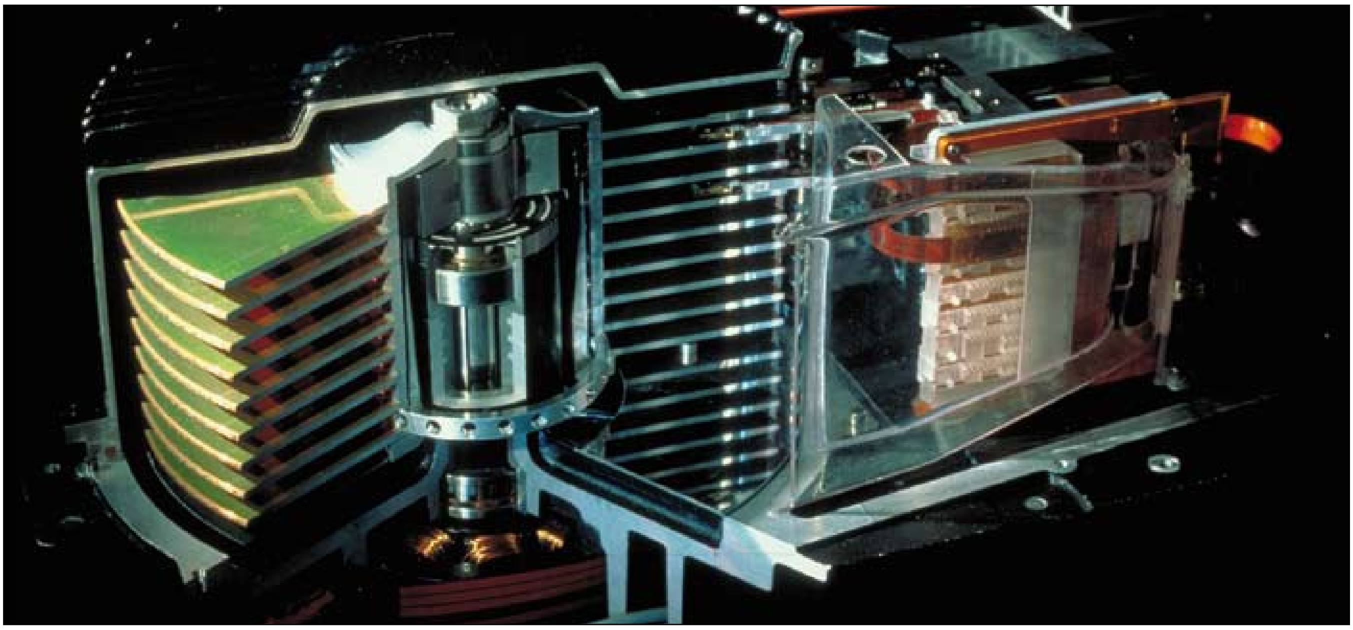


1973: IBM 3340 Winchester removable disk

Reduction of the overall height - only a few magnetic disks left. In 1977 Shugart launched the first inexpensive drive (14 inches, 30 MB). The further development led to smaller plates. Seagate built the first 5.25-inch hard drive in 1979. SCSI came in 1981 and Seagate's ST506 interface in 1982, from which IDE, EIDE, ATA and ATAPI developed. The Seagate ST506 drive, after which the interface was named, (like the RAMAC drive from 1956) had a capacity of 5 MB.

In 1996, Seagate presented the Cheetah series, the first hard drives with 10,000 rpm. In 1998 the SeagateBarracuda series had a maximum capacity of 50 GB. Just two years later, it was almost 200 GB. This far exceeded the usual increase of 60% in one year or the doubling in 18 months. Between 1957 and 1990 the rate of increase was still around 25% a year. The areal density on the hard disk disks rose from 2,000 bits / inch 2nd in 1957 to over 1 Gbit / inch 2nd in the years 1995 to 1997.

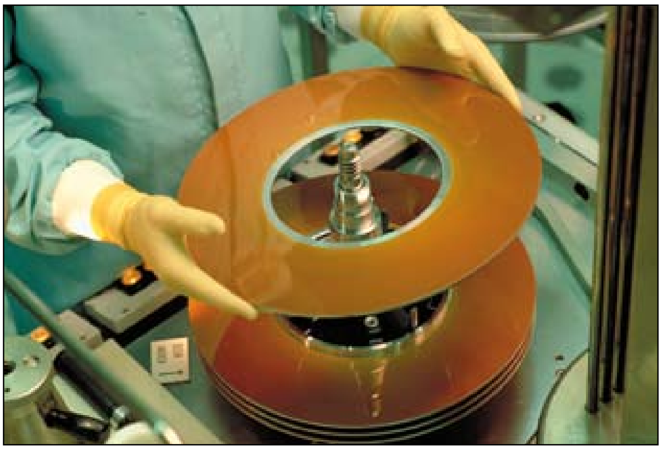
1990: IBM 3390 assembly of the access arms



On average: IBM 3390 drive

The data transfer rates of the hard drives were at least ten times higher than with floppy disks, because the disks rotated at up to 20 times the speed of a floppy disk, depending on the type, between 3,500 and 15,000 rpm. Each plate has at least two read and write heads that scan the plates on the top or bottom. However, there are also hard drives that have several sets of read / write heads, e.g. B. hard drives in high-performance computers or SCSI hard drives with four R / W heads, which reduces access time. Rotating the magnetic disks creates an air cushion on which the read / write heads then move. These are mechanically connected to the other heads via the read / write arm (comb), so that a lane change is carried out for all disks at the same time.

Cylinder introduced. This includes all tracks that have the same track number. Each record is divided into tracks, which are arranged in concentric circles around the center (similar to a record). Tracks are numbered from 0 to N, where N is the total number of tracks minus 1. The outer track has the number 0, the next one the number 1. Each track takes up a constant number of sectors, which divide the track into individual sections of the same size. The structure is similar to that of a floppy disk. Each sector contains 512 bytes of data (Fixed Block Architecture FBA) and at the same time represents the smallest unit of access. Each hard disk has significantly more concentric tracks than a floppy disk,



1987: IBM 3380 production at the IBM plant in Berlin

1990: IBM 3390 assembly

MFM (Modified Frequency Modulation)

MFM became the new method of transferring data via the read / write head to the surface of the hard disk. MFM hard drives were the standard for PC hard drives until the mid-1980s. Hard disks with a capacity of 20 MB were very common at first, later 40 MB. Consider: At the same time, a fully executable MSWord 2.0 also fit on a 360KDis chain. Current motherboards no longer have connections for this type of hard disk, but corresponding 8 and 16-bit ISA controllers can still be purchased at flea markets.

RLL (Run Length Limited)

In principle, the structure is identical to that of the MFM hard disk, only the storage capacities were larger. This resulted from the improved surface of the plates. The control of the drives also improved a lot. This made 26 sectors possible per track, which meant a significant increase in storage density. The only thing left of RLL plates today is the type of recording process. Otherwise they are - like the MFM hard disks - outdated.

Development of storage density per product from 1956 to 2000

year	product	Bits / inch (bpi)	Tracks / inch (tpi)	area density (MBit / inch 2)	Area Density (Mbit / cm 2)	Head technique
1956	350	100	20th	0.002	0.00031	Floating head
1957	355	100	20th	0.002	0.00031	
1961	1405	220	40	0.009	0.00140	
1962	1301	520	50	0.026	0.00400	Flight head
1963	1311	1,025	50	0.051	0.00790	
1964	2311	1 100	100	0.110	0.01700	
1966	2314	2,200	100	0.220	0.03400	
1971	3330	4,400	192	0.780	0.12100	
1973	3340	5 636	300	1,690	0.26200	
1976	3350	6 425	478	3,070	0.47600	
1979	3370	12 134	635	7,800	1,20000	Thin film head
1981	3380-D	15 200	840	12,000	1.86000	
1985	3380-E		1,400	19,290	2.99000	
1987	3380-K	15 240	2 083	36,000	5.47000	
1989	3390-1 / 2	27 483	2,235	60.510	9.38000	
1991	3390-3	29 718		90,000	13.50000	DFpl + MR
1991	9340/9345	24 130	4,450	111,000	16.00000	
1993	3390-9		2 921	270,000	41.85000	
1994	RAMAC 1			260,000	40.30000	
1995	RAMAC 2			544,000	84.32000	
1996	RAMAC 3			829,000	129,20000	
1999	2105 ESS			2,758,000	427,50000	GMR read head
2000	2105 ESS			5,516,000	855.00000	

DFpl + MR = thin film plate and magnetoresistive read / write head, GMR = Giant magnetoresistive read head

IBM Travelstar 60 GB 7,200 RPM Serial ATA drive

IBM Ultrastar 146 GB drive with 10,000 RPM

Hard drives in the first IBM PCs

The first interface standard for hard drives was developed by Seagate in 1980: ST506. This was used in a 5 MB hard drive with the same model name. The ST412 standard was created just one year later - also by Seagate. This 10 MB hard disk was installed in the popular IBM XT (5160). These two early standards differed significantly from their successors, IDE / ATA and SCSI: they had no intelligence of their own. The complete control was carried out by a complex controller, to which they were connected with two cables. For the control of the plate mechanics, the controller was connected to the plate via a 34-wire cable, and for the data transport with a 20-wire cable. The whole thing was still very error-prone and slow, Because the data processed by the read / write heads first had to be transported to the controller before an error could be determined. In addition, these early hard drive standards were complex to install. Many complicated parameters (e.g. the “interleave factor”) and the entire geometry of the hard disk had to be set by hand because there was no BIOS in which these values had been preset. And plug & play (i.e. the automatic recognition of hard drive geometry) as we know it today only came in the mid-1990s. the “interleave factor”) and the entire geometry of the hard disk had to be set by hand, because there was no BIOS in which these values had been preset. And plug & play (i.e. the automatic recognition of hard drive geometry) as we know it today only came in the mid-1990s. the “interleave factor”) and the entire geometry of the hard disk had to be set by hand, because there was no BIOS in which these values had been preset. And plug & play (i.e. the automatic recognition of hard drive geometry) as we know it today only came in the mid-1990s.

The successor to the MFM / RLL hard drives was the IDE hard drive (Integrated Drive Electronics) from IBM. But also these are no longer used today because the capacity was limited to 504 MB net. The types of hard drives used today are EIDE (Enhanced Integrated Drive Electronics), SCSI (Small Computer Systems Interface) and currently SATA (Serial Advanced Technology Attachment), a further development of ATA (Advanced Technology Attachment) that was only defined in 2001.

The difference with these newer standards was mainly in the controller area. SCSI hard drives were a bit faster at the time, but also much more expensive. The advantage of SCSI was its flexibility, because pretty much all types of devices (floppy, streamer, CDROM, hard disk, scanner, etc.) can be connected to a SCSI controller, up to seven pieces per controller (further development then 16 per channel). An EIDE controller, on the other hand, could only manage two devices because this interface was actually designed exclusively for operation with hard disks. Nevertheless, there are now CD drives or streamers for EIDE as well. Modern motherboards often already contained two integrated EIDE controllers. This enabled four corresponding devices to be operated, which was completely sufficient for most users. Unlike all other systems, SATA transmits the data serially, ie bit by bit. Nevertheless, transfer rates of up to 600 MB per second are achieved. SATA is compatible with the ATAPI standard and the various ATA predecessors. Another great advantage: there is almost no configuration and devices can be plugged in and unplugged when the computer is switched on (HotPlugging).

In the 80s to 1991 **large brown plates built.** The IBM 3380 technology with 14-inch disks, which was later replaced by 10.5-inch disks, the 3390 technology, constantly dealt with the question of how much capacity can be managed under an actuator, i.e. under a read / write head, without it queues occur and the performance of the drive is impaired. Due to this problem, 2 read / write heads were usually brought up on the access. One head was responsible for the interior and the other for the exterior of the plate.

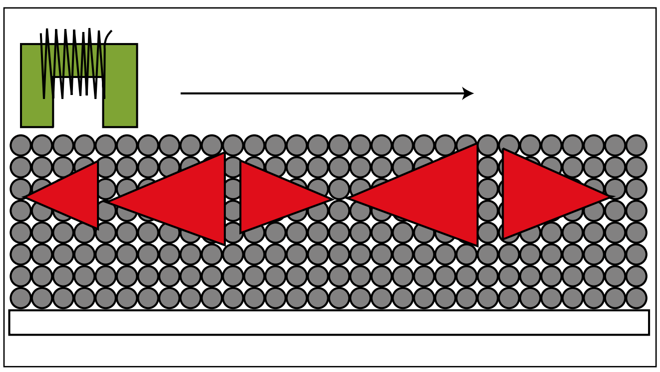
The **inductive recording** has been **until 1991** tracked and deployed. Let us imagine such a recording read / write head as a cut transformer with a corresponding copper coil. The writing process now actually requires few copper turns when using as thick a wire as possible so that as much current as possible can be transported and a well-magnetized bit is generated on the magnetic layer. Reading a bit, however, requires exactly the opposite, namely as many copper turns as possible with the thinnest possible copper wire, so that the stray field of an information unit causes the highest possible inductive voltage, which can then be tapped as a clear read signal. Due to the conflicting requirements regarding the number of turns and the thickness of the copper wire, the head was designed that both processes were possible. This **Compromise** but created a direct dependency on the speed between head and disk. If you turned too fast, the writing process became difficult, if you turned too slow, the speed of rotation was not sufficient to generate a correspondingly high inductive voltage, which could then be tapped as a clean read signal. The most complex manufacturing process at that time was that **Manufacture of the plate** self.

Iron oxide particles, i.e. rust particles, were used as the magnetization carrier and the art was to distribute these rust particles as homogeneously as possible in a synthetic resin mass in order to be able to magnetize them to some extent evenly. The manufacturing process of the heads was easier in comparison.

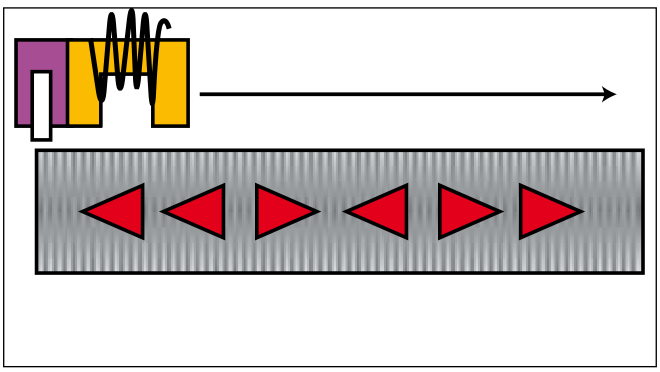
These brown plates caused enormous problems at the beginning of the 1990s (1990-1993), especially with models 1, 2 and 3 of the 3390 family, because after many years of use, the iron oxide particles worked their way out of the synthetic resin mass and accumulated on the heads . The accumulation of these rust particles then produced the famous flight effect, that is, the distance between the heads and the surface of the plate increased until a height was reached that no longer allowed writing or reading. In a unique campaign that took over two years, IBM then replaced all of the world's installed disk stacks.

With inductive recording technology **linear access mechanisms** used, which were designed very complex by the mechanics to position the head, which was responsible for both writing and reading, exactly over a track.

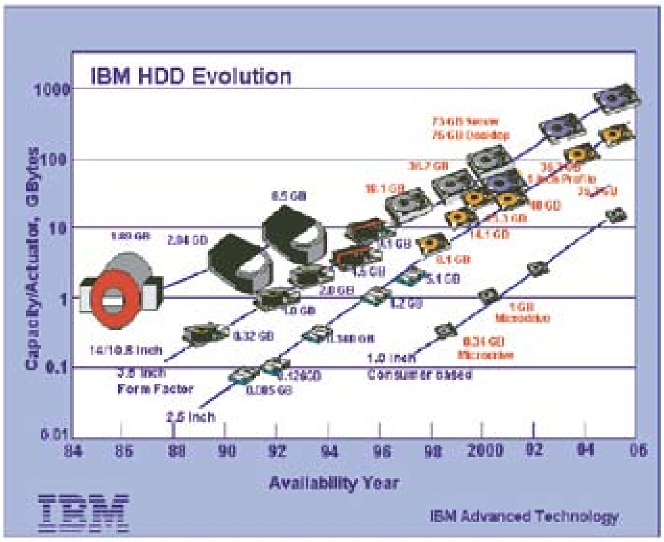
1991 IBM said goodbye to the inductive recording used at that time and was the first company to lead
MR technology for the head technique (**magnetoresistive recording)** for the products 3390 model 3 and 9 as well as for the 9340 system with its 5.25 inch plates. Today, every manufacturer works with this technology that IBM developed. The MR head technology had so much potential for further development that this technology has accompanied us to this day (2008). The difference to inductive recording is that you can no longer use one



Principle of inductive recording technology



Principle of magnetor-resistive recording technology



1991:
RAID-Architekturen
Dünnfilmbeschichtung
MR-Aufzeichnung

1995 – 1997:
PRML Encoding

1999:
Micro Drives
Paramagnetischer Effekt

2000:
USA National Medal of Technology for Innovations in Storage

Read / write head works in which the number and thickness of the copper coils must be designed so that both writing and reading is possible. The MR head still uses an inductive head for the writing process, but the number and thickness of the copper coils are now designed so that an optimal writing process can take place. This head no longer has anything to do with reading. This enables much faster rotation speeds. The reading process takes advantage of the "magnetoresistive effect from physics" that metals or alloys can change their resistance if they come under the influence of a magnetic field. The work is usually done with a nickel iron film that is connected to low current. If this slider is in the area of influence of a magnetic field, the alignment of the electrons changes. If you amplify this accordingly, you get a clear, unambiguous read signal. With this method, even the smallest stray fields of individual information units can be tapped. This is how you create one **Specialist for the writing process and a specialist for the reading process.**

In parallel, as **second important element** changed the plate making. The lamination and vacuum processes were now so far that thin-film coatings (alloys) could be used on the plate and a metal film was used instead of iron oxide. The metal film is sprayed onto a specially hardened glass plate under vacuum (this process is called cathode sputtering). The plate then runs directly from the vacuum into an oven to be "baked" (this process is called sputtering). The heating process (sputtering) creates "grains" in the thin film, i.e. grain structures, and the individual grains represent the magnetization carriers of the plate. The production of such thin film plates has been retained from the principle to the present day and is described in more detail later in the part of the AFC technology. With the combination of MR head technology and thin-film plates, it was possible to increase the recording density from then since 1991, when it was introduced - 16 million bits in the square centimeter to over 5 billion bits in the square centimeter. This corresponds to an increase of over a factor of 300 and represents a unique technological evolution, as described in no other technological segment has ever been achieved in terms of the growth factor.

The **third driving factor in 1991** was the introduction of new subsystem architectures, the RAID-based systems.

The RAID architectures ushered in the step of changing from large panel sizes to smaller form factors. Although the small disks did not have the comparable reliability factor as the large brown disks, the so-called SLEDs (Single Large Expensive Disk), as they were called at the end of their lifespan, they also did not need this, because drive failures in an array are protected by RAID were (see also under RAID).

In the year **1990** IBM integrated a new encoding method into the 5.25 and 3.5 inch drives manufactured at that time

as PRML encoding is called and is still in use today. **Partial Response / Maximum Likelyhood (PRML)** is a writing procedure to write data on magnetic data carriers as space-saving as possible. For this purpose, this data is encoded beforehand. This trick gives you a higher data density on the surface of a hard drive. This happens completely unnoticed by the user through the controller of the hard disk. In contrast to older methods such as MFM (Modified Frequency Modulation), no individual magnetic field reversals are identified in the analog signal stream when the data carrier is read and their sequence is interpreted as a sequence of data and synchronization bits. Instead, the analog signal is broken down into sections, these are processed ("partial response") and the result is compared with specified patterns in order to find the most similar ("maximum likely hood"). Each of the given patterns stands for one

certain bit sequence. PRML allows a 30 - 40% higher data density than MFM. PRML is still used today in modern hard drives, but is increasingly being replaced by EPRML (Enhanced), which enables 20 to 70% higher data densities due to its improved algorithms and signal processing.

PRML was used for the first time in 1990 for plates of the IBM RS / 6000 and AS / 400 systems, further developed and then also used for 3.5-inch drives of all IBM product series, which were then primarily manufactured, as well as in many subsequent products. The process was developed in the IBM laboratory in Rüschlikon near Zurich.

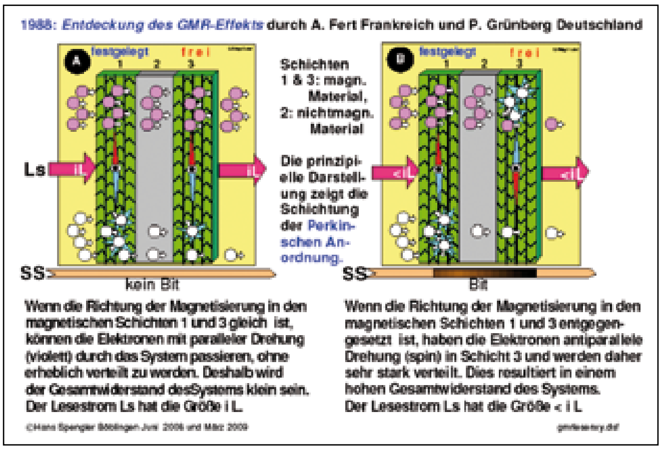
An example: While around 800–900 magnetized grains were conventionally required to represent a stable information unit, the PRML process has reduced the number of grains to 200–250 over the years, ie four times the capacity.

The capacitive effects of PRML were seen above all from 1995 to 1997, when the capacities of the 3.5 inch drives produced at that time tripled within 12 to 15 months.

1997 was developed as a further development of the MR heads **GMR technology,** the giant magnetoresistance reading head technology was introduced, which above all made it possible to read out the smallest stray fields. The **GMR effect** discovered independently of one another in the late 1980s **Peter Grünberg** at the KFA in Jülich and **Albert Fert** at the University of Paris South. Extremely low temperatures and strong magnetic fields were generated in both discoveries

Introduction of PRML channels in IBM plates

plate	GB capacity	Diameter inches	Data rate MB / S	Record density Mb / sq "	Head type plate type	PRML
Redwing 0681 Feb. 1990	0.857	5¼	3rd	63	MIG thin film	1st gen.
Corsair 0663 Sep 1991	1.0 / 1.2	3½	3rd	132	MR thin film	1st gen.
Allicat 0664 Nov. 1992	2.0	3½	5.2	264	MR thin film	2nd gen
Spitfire 0662 March 1993	1.0	3½	5.2	354	MR thin film	2nd gen
Starfire series 0667/0668 Nov. 1993	1.0 - 5.2	3½	12.2	564	MR thin film	3rd gen



effect needed. Some time after this work started **Stuart Parkin** and other researchers at **AlmadenResearch Center by IBM** in San Jose, California, with an attempt to further develop the effect at normal temperature and with less exotic material compositions for mass production. It took them many years to do this and researched more than 30,000 different multilayer material combinations. In 1997 the three researchers named together received the **Europhysics Prize**

for GMR development. Peter Grünberg and Albert Fert were also **2007** with the **Nobel Price for physics** excellent.

The GMR effect (giant magnetoresistance) is observed in structures that consist of alternating magnetic and non-magnetic thin layers with a thickness of a few nanometers. The effect has the effect that the electrical resistance of the structure depends on the mutual orientation of the magnetization of the magnetic layers, and it is significantly higher when magnetizing in opposite directions than when magnetizing in the same direction.

It is a **quantum mechanical effect,** which can be explained by the spin dependence of the scattering of electrons at interfaces. Electrons that can spread well in one of the two ferromagnetic layers because their spin is favorably oriented are strongly scattered in the second ferromagnetic layer if it is magnetized in the opposite direction. However, they pass through the second layer much more easily if the magnetization has the same direction as in the first layer.

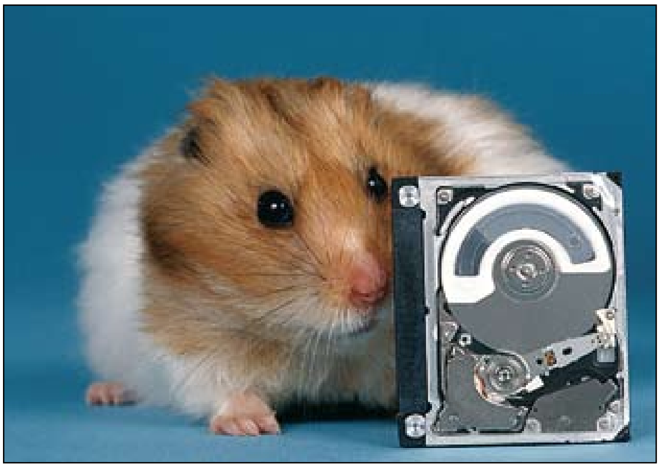
If two layers of a ferromagnetic material are separated by a thin non-magnetic layer, the magnetizations are oriented in opposite directions at certain thicknesses of the intermediate layer. Even small external magnetic fields are sufficient to return this antiferromagnetic order back to the ferromagnetic order.

In conjunction with the GMR effect, variations in the external magnetic field in suitable structures therefore cause large changes in the electrical resistance of the structure.

The possibilities of using the effect in a sensor for a magnetic field (and thus as a new type of reading head in a hard disk) were quickly discovered by an IBM research team led by Stuart Parkin, who showed that the effect can also be found in polycrystalline layers occurs.

IBM introduced in **December 1997** the first commercial drive to use this effect. GMR is still in use today. In addition to use in hard drives, the GMR effect is also used in magnetic field sensors in the automotive and automation industries.

1999 could be the first **Micro drives** - Back then with a capacity of 340 MB - to be launched on the commercial market. The capacities of the MicroDrives developed rapidly and today, in 2008, capacities of 6 GB and 8 GB are already being offered.



IBM Micro-Drive

By luck, the IBM developers discovered in the second half of the year **1999** the possibility of a new recording technique. One spoke and still speaks of the **Discovery of the paramagnetic effect.** Already on October 4, 1999, the IBM laboratory in Almaden was able to demonstrate a hard disk laboratory version using this new technology. In a press release on October 4, 1999, the recording density was 35.3 Gbits / inch 2nd

spoken - these are **over 5 billion bits to the square centimeter.** This initiated the physical limits that had previously been seen for MR technology. With this technology, four times higher capacities can be realized in the same room!

To produce the magnetic coating of a hard disk today, a complex alloy is sprayed onto a glass pane surface under vacuum (cathode atomization), which is then "baked / fired" (sputtering). This creates approx. **80-100 nanometer magnetic grains (** MR technology). To store one bit, an inductive write head aligns the magnetic orientation of a few hundred such particles. In order to increase storage densities, attempts have been made to reduce the number of grains that form a bit. However, this had the serious disadvantage that the signal-to-noise ratio deteriorated. Therefore, attempts were made to reduce the size of the grains and to maintain the number that make up a bit. However, if the grain size comes below a certain limit, the particles lose their ferromagnetic properties (paramagnetic limit). This effect can be partially offset by using "harder" magnetic materials with a higher coercive force. Problem here is the fact

With the discovery of the **paramagnetic effect** and the paramagnetic recording option became one
Grain size of approximately 25-30 nanometers possible without having to switch to harder materials. For the recording density, this meant a quadrupling of the capacities that could be imaged in the same space.

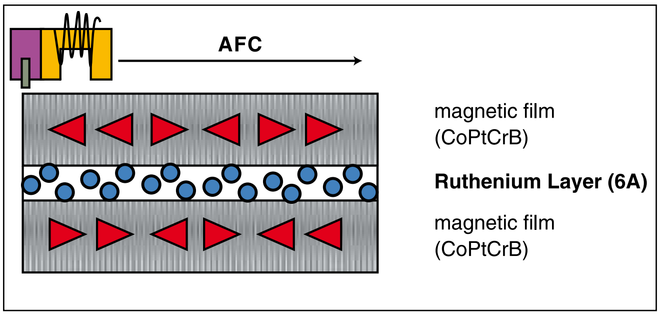
A special head is connected upstream of the MR head, which generates a vertical bit in the metal grain structure. The horizontal data bit is written immediately afterwards. **Stabilize both bits** with certain alloys and allow the horizontal data bit to be made much smaller and to work with significantly higher rotational speeds and data rates on the plate.

Mass production of this technology would have become very complex and costly. That is why the developers took action **other ways.** This led to IBM's mass production of the so-called **Pixie dust technology**

started in May 2001, which at that time shook up the entire trade press, because much higher capacities could be realized at even faster turning speeds.

in the **Year 2000** IBM was the first company to receive the **National Medal of Technology for Innovations in Storage** by the American government, an award that has so far been given only to individuals. The reason for the award: the mass production of MicroDrives and the discovery of the paramagnetic effect. It is one of the highest technological honors.

The AFC recording technology (Pixie Dust) allows for
average grain sizes of 4-8 nanometers to achieve a much higher thermal long-term stabilization in the thin film without having to put up with a higher coercive field strength of the material. That means: using significantly smaller grains as magnetization carriers for softer materials. For this purpose, boron is added to the classic metal film made of cobalt, platinum and chrome that has been used up to now.



Principle of AFC (AntiFerromagnetically Coupled) recording technology

This technology is called **AFC (AntiFerromagnetically Coupled)** or, shaped by the IBM researchers in Almaden,

"Pixie Dust". AFC media consist of three layers. There is only one between two magnetic layers consisting of a complex cobalt-platinum-chromium-boron alloy **three atomic layers (6 angstroms) thick layer made of the non-magnetic metal ruthenium.** With exactly this thickness, the ruthenium layer ensures that the magnetization orientations of the other two magnetic layers **coupled anti-parallel** are. This means that the lower magnetic layer is always magnetized the other way round than the upper layer. This leads to a

enormous stabilization of the two bits and allows considerably higher write densities (up to 100 gigabits per square inch and more - approx. 15-20 billion bits per square centimeter) with significantly higher rotational speeds and data rates. Because of the miraculous effect of the ruthenium layer, the IBM researchers in Almaden christened this new technique "Pixie Dust", ie fairy dust, based on Peter Pan's fairy tale.

Another **advantage** lies in the fact that you can continue to work with today's MR write / read head technology. The miniaturization of these heads is already so far today that write / read gap widths of 1–2 nanometers are possible as laboratory versions. This represents another potential to drive recording density.

With the new AFC technology, disk recording technologies will surely be able to go into capacitive areas of more than 1 TB in the years 2010-2011.

Because of the heat generated in the drives, you have to find out more about it **Packaging,** especially think about cooling these plates. **Water-cooled racks** could be an alternative. There are also possibilities of "old technology" in development, such as that used for the 3380 plates, but in a correspondingly miniaturized form. It works with two heads on the arm, one controls the inside of the plate, the other the outside. This is one way of compensating for the additional capacity under the actuator in terms of performance without having to massively increase the rotational speed of the plate stack. So there are opportunities to advance capacity in the next few years.

In the year **2003** founded **IBM and Hitachi form a joint development and manufacturing alliance** for disk drives. In this alliance, IBM is responsible for the further development of plate technology and Hitachi for mass production. This alliance has proven itself to this day and it can be assumed that it will continue for many years and will be extended to other technological areas. Hitachi is now the third largest plate supplier in the world. The second largest production volume is covered by Western Digital. The largest is still the Seagate company, which produces about 40% of the world's total volume. Today, in 2010, two form factors are being produced: 2.5 inch and 3.5 inch drives.

Within the next two years at the latest - all panel manufacturers agree - only one form factor will generally be used because no manufacturer can afford to produce two different form factors in the event of price drops of around 40% per year. The

3.5-inch drives, which were previously mainly installed in disk subsystem architectures, will be replaced by 2.5-inch drives.

The prerequisite for this strategy is the production of powerful 2.5-inch drives. It emerges that three different techniques are then used within the 2.5-inch design: the **AFC technology,** a new technique with **Perpendicular recording**

and **IDE** Disks known as ATA, SATA, or FATA disks. This approach makes the world easier.

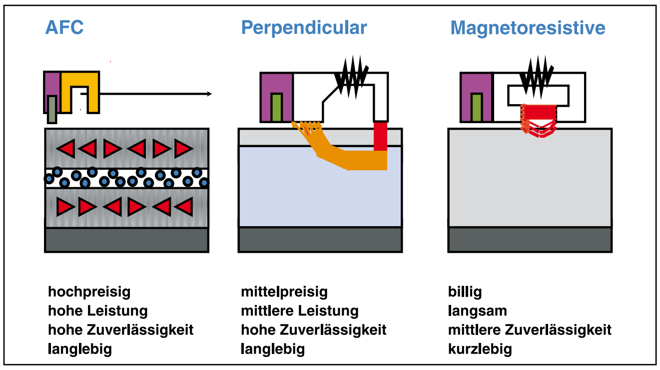
Just one more form factor, which means: AFC as a high-priced, high-performance drive with extreme reliability, drives with perpendicular recording as a medium-priced drive with good performance and comparable reliability as AFC and the cheap SATA / FATA disks, which correspond in structure to IDE disks.

Perpendicular recording

In perpendicular recording, the magnetic moments, which, together with the logical writing methods used such as PRML, each represent a logical bit, are not parallel to the direction of rotation of the data carrier, but perpendicular to it. This leads to a potentially much higher data density, which means that more data can be stored on the same surface. This was first established in 1976 by Professor Shunichi Iwasaki at the Tohoku Institute of Technology in Japan.

The disadvantage is that the smaller white areas also require a shorter distance between the read / write head and the magnetic surface so that the data can still be written and read. Therefore, this recording technique is technically more difficult to implement. The counterpart to perpendicular recording is the longitinal recording used up to now. Here, however, the so-called superparamagneic effect occurs if the magnetic particles are too close together (too high data density), ie the individual bits lose their magnetic direction and begin to “tilt”, which leads to data loss. 3.5-inch hard drives thus have a maximum capacity of currently 1 TB, 2.5-inch hard drives currently have a capacity of 320 GB. This corresponds to an average data density of 132 Gbit per square inch or 20,

The hard disk capacities can be increased by a maximum of ten times with this new recording technology. In addition, the significantly higher data density results in an increase in the read / write speed, since the read head reads more data per revolution and thus the data rate increases with the same number of revolutions. Seagate has been delivering records with perpendicular recording since 2006.



Comparison of recording methods

HAMR (Heat Assisted Magnetic Recording)

HAMR (Heat Assisted Magnetic Recording) describes a process that allows even higher data densities for hard drives than the new process with perpendicular recording, which has already reached market maturity. This means that hard disks with volumes in the mid single-digit terabyte range should be possible in the future. With the help of HAMR, data can also be written on materials with a weakly pronounced superparamagnetic effect. The so-called effect occurs with today's hard drives if you increase the data density too much. This means that the domains into which data is written lose their ferromagnetic properties and become superpara-magnetic, destroying the bits.

With HAMR, the domain into which data is to be written is first locally heated by a laser in order to keep the magnetic field required for a writing process as small as possible and to enable writing despite a weakly pronounced superparamagnetic effect. However, since the protective lubricant of the hard disk evaporates when heated, the hard disk manufacturer Seagate wants to store supplies in nanometer-thin carbon tubes, which the latter should spray onto the affected domains if necessary.

There is already speculation as to what storage capacity can be achieved using this technology. For example, on January 4, 2007, itwire.com reported that Seagate's capacity of 37.5 TB may already exist in 2010, today. That would be 37.5 times as much as Hitachi's largest hard drive with around 1 TB. However, it is also reported that this technology could increase the number of HeadCrashes and thus the quality of the hard drive would suffer. However, this speculation contradicts Seagate's Techworld statement that 3TB hard drives can be expected today in 2010.

Plate rotunda in the IBM Museum Sindelfingen

The house for the history of IBM data processing in Sindelfingen, Bahnhofstrasse 43

Today, two unique systems can be viewed in the IBM Museum, the IBM RAMAC 355 storage unit, which was launched in 1957, and the legendary IBM 1401 EDP system with IBM 1311 disk storage, which was announced in October 1959 and last year the 50th anniversary celebrated.

Particularly noteworthy is the recently completed disk rotunda, which shows the history of IBM's magnetic disks from 1956 to the year 2000 and presents many exhibits of the most diverse technologies during this period. In the IBM Museum you can take a fascinating journey through time through the world of IBM magnetic disk games issue. The absolute "highlight" of the rotunda can be seen on the far left in the plate rotunda under plexiglass: the RAMAC 355's access mechanism, which can be demonstrated functionally. For this purpose, the compressor required for this is located directly behind the rotunda wall (not visible in the picture).

IBM plate rotunda in the IBM Museum Sindelfingen

The initiators and builders of the plate rotunda in the IBM Museum were Messrs. Hans W. Spengler and Heinz Graichen. It took over ten years for this work to be completed and released for inspection. The data collection alone took many years, which was primarily driven by Hans W. Spengler. For this purpose, he created a disk profile with all technical details for each IBM magnetic disk from 1956 to the year 2000 and fortunately made these profiles available for the update of the IBM System Storage Compendium 2010. You can find these profiles on the following pages.

Hans W. Spengler

The author of the compendium contributed to the fact that the exhibits collected at the IBM plant in Mainz were transferred to the IBM Museum a few years ago so that the construction of the plate rotunda could begin. You will also find an interview afterwards, in which the author asks the then specialists Heinz Graichen and Hans Spengler about the IBM RAMAC 350/355 disk storage device, how the access unit with the read / write head did not work in principle, but actually.

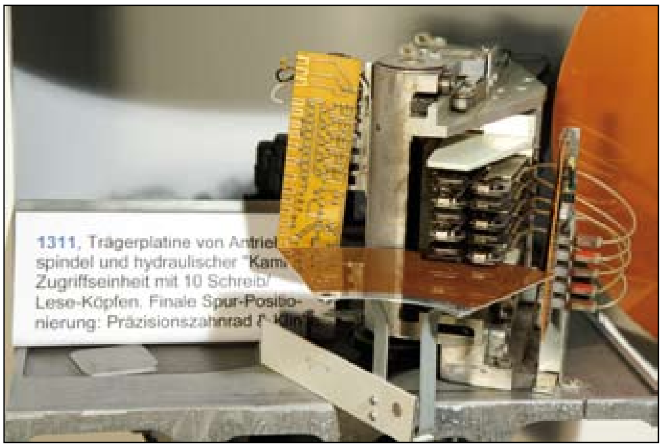
Heinz Graichen

Magnetic disk exhibit collection at the disk rotunda of the IBM Museum in Sindelfingen

The beginning of electromagnetic storage

Partial rotunda RAMAC 350/355 access and IBM 1301 representation from 1956-1961

IBM RAMAC 350/355 plate with 24 inch diameter years 1956-1957

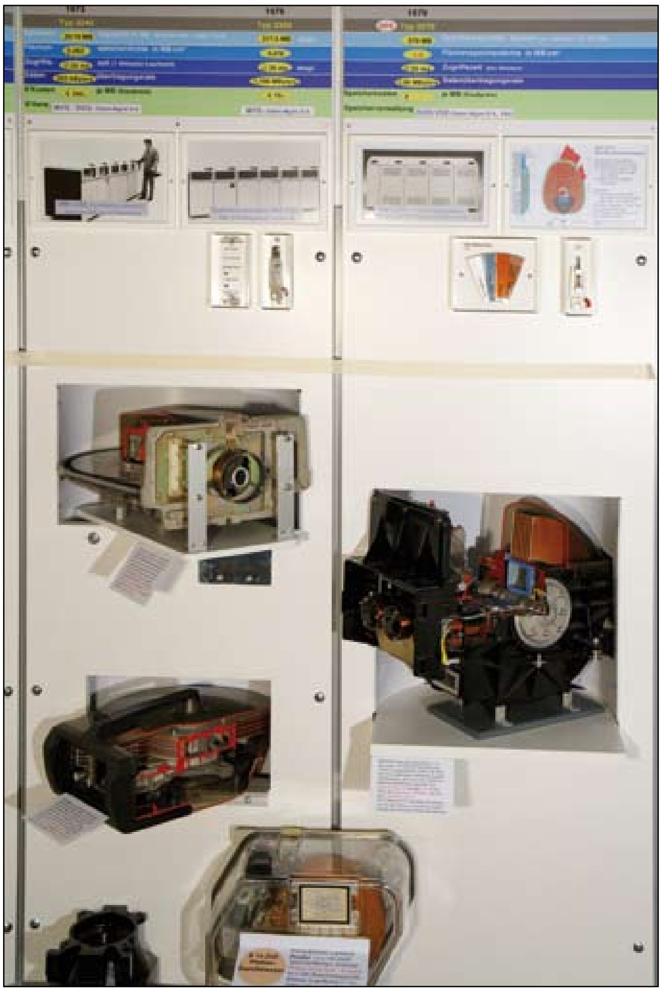


IBM 1311 access 1962

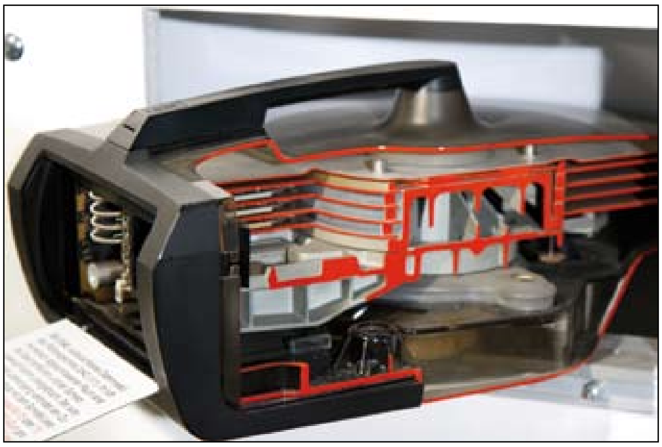
Epoch of removable plates

Partial rotunda IBM 2314 open (see below) and IBM 3330 representation 1965-1972

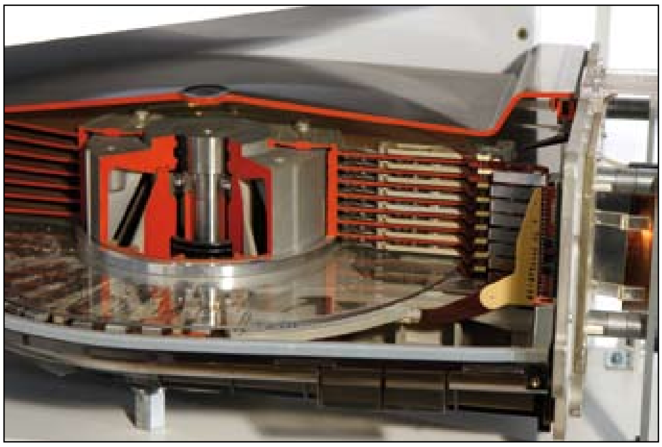
Epoch of removable plates (Winchester period) and epoch of permanently installed plates with external control units



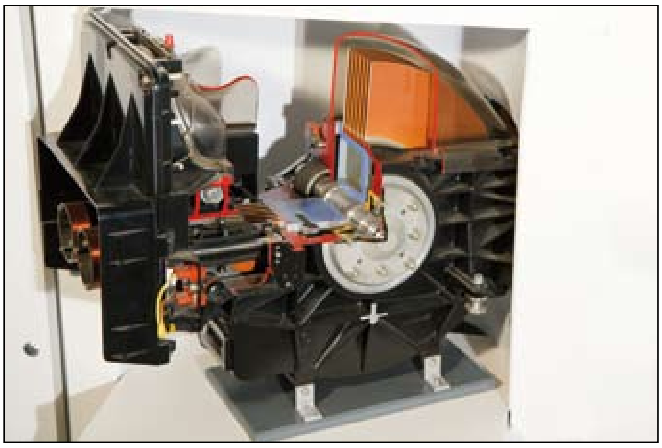
Partial rotunda IBM 3340, 3350 and 3370 illustration 1973-1979



1973 IBM 3340 drive data modules



IBM 3350 hard drive 1975

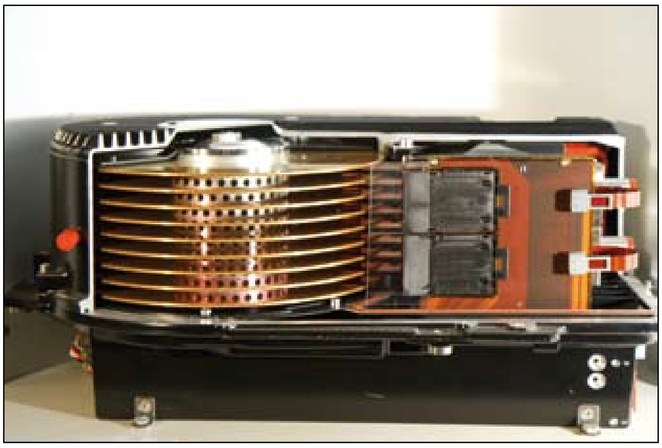


IBM 3370 drive 1979

Period of the permanently installed panels with external control units

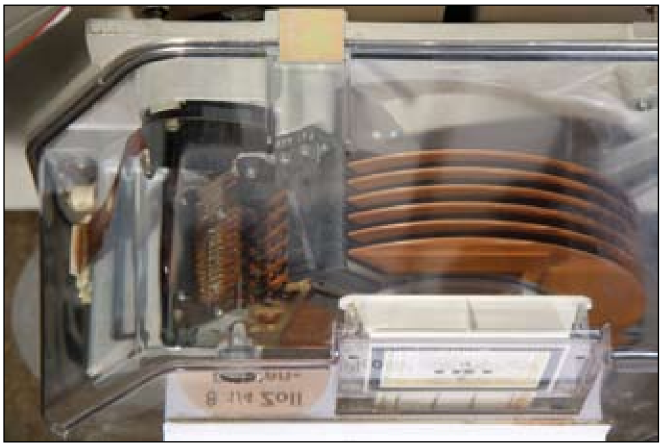
Partial rotunda IBM 3380 and IBM 3390 representation from 1981-1994

IBM 3380 drive 1981



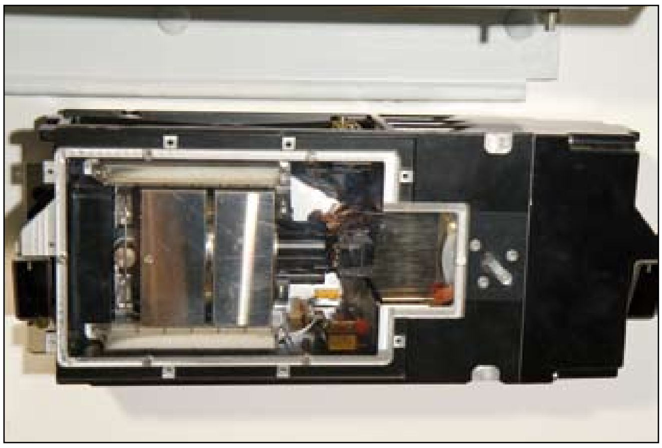
IBM 3390 drive 1989

PC XT hard drive 5 ¼ inch 1980



IBM Piccolo drive Hursley 1978

drive with PRML 1990



IBM RAMAC 9345 drive bay 1991 IBM 0681 - 5 ¼ inch

The era of RAID systems and the beginnings of the era of multi-platform systems

Partial rotunda IBM RAMAC and IBM ESS representation from 1994-2000

IBM RAMAC 1, 2, 3 plates 1994-1998

IBM ESS 9 GB SSA drives 1999-2000

IBM Travelstar 1996-1998 as a mobile plate in laptops

Zur Geschichte der IBM Plattenspeicher: Profile ausgewählter IBM DASD's

Das *Haus zur Geschichte der IBM DV* nahm 2008 die neue Ausstellungseinheit, genannt *„DASD-Rotunde"* in Betrieb. Damit können Interessierten Einblicke in die Entwicklung der Magneto-motorischen Direktzugriffspeicher von IBM ab 1950 bis 1999 gegeben werden, die i.d. Form sonst nicht zugänglich sind. Der Terminus **D**irect **A**ccess **S**torage **D**evice wurde erst lange nach der Produktverfügbarkeit in 1956 benutzt in den 70er Jahren. Ab dieser Zeit konnten Computersysteme ohne DASD's nicht mehr effizient betrieben werden.

Auch im *Haus zur Geschichte der IBM DV* ist Stellfläche knapp. Dies und insbesondere die strikte Fokusierung auf die Kernstücke wie *Speicherplatten, Schreib/Lese-Köpfe und Zugriffsmechanismen* je als echte Exponate, ermöglichten es, zusätzlich erläuternde Bilder und Graphiken unterzubringen. Leitzeilen über die ganze Rotunde hinweg geben dem Betrachter Orientierung über Schlüßel-Kenngrößen.

Das Konzept der Gesamtgestaltung erlaubt es dem Besucher, an einer konventionellen Führung in freier Präsentation teilzunehmen oder einer semi-automatischen PowerPoint-Präsentation zu folgen oder auch sich durch Eigenanschauung mit dem Thema zu befassen.

Um den außerordentlichen Umfang des Stoffes auf das Wesentliche zu verdichten und trotzdem authentisch zu bleiben, war es erforderlich, eine geeignete Informationsbasis zu erarbeiten, die Kenndaten der großen, mittleren und kleinen Laufwerke / Speichereinheiten abzubilden. Für jedes Laufwerk/Speichereinheit wurde ein Profil erarbeiet. Selten genutzte Einheiten wurden nicht berücksichtigt. Die Profile bildeten die Fakten-Basis der *DASD-Rotunde*, ihre Gliederung folgt den zeitlichen Dekaden seit 1956.

Kleinsysteme hatten - beginnend mit IBM 1130 in 1965 - auch integrierte DASD's. In den 70er Jahren folgten solche in gesteigertem Umfang bei Branchen-Subsystemen, dann der Serien S/3, S/32, S/34, S/38 und S/36, dem branchenunabhängigen Subsystem 8100 sowie den Prozeß-Steuerungssystemen S/7 und S/1. Dies bewog uns, auch deren intergrierte Laufwerke exemplarisch einzubeziehen, denn große Volumen dieser Laufwerkkategorie zusammen mit Technologieentwicklung brachte fallende Preise pro gespeichertem MB. Dies führte zur schrittweisen Ablösung der jeweils größeren Speicherplattendurchmesser.

Herrn Kurt Gerecke's Wunsch, die über Jahre hinweg erarbeiteten Profile in dieses, *„sein Speicherkompendium"* aufnehmen zu dürfen, ehrt uns, wir erfüllen diesen Wunsch sehr gerne.

Profile von IBM DASD's
nach Ankündigungsjahr / Dekade

Jahr	DASD	Computer-System(e)	Kategorie Klein < Mittel = Gross >
1956	IBM 350	RAMAC 305	<
1957	IBM 355	RAMAC 650	=
1960	IBM 1405	RAMAC 1401	=
1961	IBM 1301	70XX / 1410 / 1460	>
1962	IBM 1311	1440/ 1401/ 1410/ 1620	=
1964	IBM 2311	S / 360	= >
1965	IBM 2314	S / 360	>
1969	IBM 5444/45	S / 3	<
1970	IBM 3330	S / 370	>
1973	IBM 3340	S / 370	=
1975	IBM 3350	S / 370	>
1978	IBM 62PC	S/ 38	<
1979	IBM 3370	S / 370	=
1979	IBM 3310	S / 370	<
1980	IBM 3380	S / 370	>
1987	IBM 9332	S/36 / AS/400 / 9370	< =
1989	IBM 3390	S / 370 / 390	>
1990	IBM 0681	RS/6000-M930	=
1991	IBM 9340 mit Einschub 09345 à 2 LW	S / 390	< =
1992	IBM 0662	Workstations, Network-Server	<
1994	IBM 9394 RAMAC-neu an S/390 mit Drawer 9392 mit dem Laufwerk Allicat. Mit dem Laufwerk Ultrastar XP in 1995 und mit dem Laufwerk Ultrastar 2XP in 1996		= >
1999	IBM 2105 ESS	ES/9000	>>

Die Profile der oben nach Ankündigungsjahr gelisteten DASD's sind auf den folgenden Seiten enthalten.

Der Betrachtungszeitraum der IBM DASD-Entwicklung - über fünf Jahrzehnte hinweg - enthält eine Fülle von Erfindungen, Entwicklungen und Innovationen aus einer ganzen Reihe von Wissenschaft- und Technik-Bereichen. Diese DASD-Profile mögen den Leser unterstützen, die große Bandbreite dieser Entwicklungen zu erkennen.

z GdEDZS.tmd

Einheit **Plattenspeichereinheit 350** der
Random Access Memory ACcounting Machine 305 (RAMAC); angekündigt 1956
Entwicklungslabor: San Jose unter dem Namen „Source Recording Project"
Produktion: San Jose und Sindelfingen

Oben im Hintergrund IBM 350, 225 cm breit, 82 cm tief und 188 cm hoch; Gewicht 790 kg. Vorne die 305 System-Steuerkonsole

Modell 1 mit 5 Millionen BCD-kodiert. Zeichen Speicherkapazität.
2 mit 10 Millionen BCD-kodiert. Zeichen Speicherkapazität.
Plattenspeicherturm mit vertikal rotierender Welle. Für beide Modelle betrug die Datenrate 8,8 Kilozeichen pro Sekunde.
Direktanschluß an den Prozessor, dieser steuert alle Abläufe*.

Konzeption *Bei Testläufen der in Entwicklung befindlichen IBM 701 Magnetbandeinheit 726 zeigt sich die Überlegenheit des Mediums Magnetband im Hinblick auf Speicherkapazität und Schreib/ Lesegeschwindigkeit. Jedoch wird auch offenbar, daß sich die Zugriffszeit bis in den Minutenbereich hinein bewegen kann, d.h Magnetband für Sofortzugriffe bei ganz bestimmten Anwendungen* ***nicht das geeignete Speichermedium ist.***

Ein neuer Speicher, der diese Anforderungen erfüllen konnte, musste entwickelt werden. *Die Lösung* bestand aus *50 beidseitig beschichteten Aluminuimplatten* montiert auf einer vertikalen Welle, die mit *1200* Umdrehungen pro Minute rotiert; auch Plattenturm genannt.

Man begnügte sich mit einem Paar Schreib/Leseköpfe, das mittels eines Stahlseilantriebes zweidimensional positioniert wurde gemäß dem Inhalt eines Adreßregisters. Dieser Vorgang dauerte max. 800 ms. Damit konnte die Forderung erfüllt werden, etwa 10.000 Transaktionen pro Tag je ***sofort beim Eintreffen des Geschäftvorfalles*** zu bearbeiten.

Technische Daten

Speicherplatte
- Plattenturm..................*mit 100 nutzbaren Speicherflächen, fest eingebaut*
- Trägermaterial..............*Aluminium, Dicke 1/10 Zoll =* **2,54 mm**
- Speicherschicht...........*1,2 Millizoll* = **0,0305 mm** Dicke aus *FeO^3*
- Plattendurchmesser.....*24 Zoll =* **609,5 mm**
- Speicherspuren...........*100 konzentrische Spuren pro Oberfläche, jede Spur war in 10 adressierbare Sektoren je 100 BCD-Stellen eingeteilt.*
- Spurendichte...............*20 Spuren pro Zoll* **= 1,27 mm Spurabstand.**
- Speich. Pl.-Drehzahl...*1200 Upm*

Zugriffs-mechanismus
(Prinzip-Skizze siehe unter 1405-Profil)
- *Vertikalwagen mit Zugriffsarm, letzterer ist als Gabel ausgebildet die das Schreib/Lese-Kopfpaar trägt. Der Vertikal-Wagen positioniert sich entsprechend der adressierten Speicherfläche, danach fährt der Zugriffsarm die adressierte Speicherspur an. Den Antrieb besorgen ein Motor, Magnetpulverkupplungen und ein bi-direktional laufendes Stahlseil. Die finale, pneumatische Positionierung des Vertikalwagens erfolgt mittels Präzisionskonus, die des Zugriffsarmes mittels Präzisionszahnstange und pneumatischer Klinkenverriegelung. Alle Zugriffsabläufe sind elektronisch gesteuert.*
- *Zugriffszeit max. 800 ms*
- *1957 gab es optional eine 2. Zugriffsstation, um während des Schreibens /Lesens einer Station, gleichzeitig mit der 2. Station Suchen der nächsten Speicher-Adresse zu ermöglichen.*

Schreib/Lese-kopf
- *Schwebehöhe über der Plattenoberfläche = 0, 02032 mm, diese wurde pneumatisch erzielt und konstant gehalten durch Druckluft. Diesen „Kopf-Arbeitszustand" bewirkte - nach der Positionierung - ein Steuerventil durch das Signal „Kopf laden" für das S/L-Kopfpaar.*
- *Kopfdurchmesser 5/8 Zoll =* **16 mm** , *Kopfhöhe* **4,5 mm.**
- *der Schreib/Lesekopf enthält eine Schreib/Lesespule und eine Löschspule und determiniert durch die Jochdimensionen sowohl die Spurbreite als auch die Spur-Bitdichte.*

Aufzeichnung auf der Spur
- ***N****on* ***R****eturn to* ***Z****ero IBM (NRZI), d.h. ständiger Schreibstromfluß, umgeschaltet wird nur die Stromrichtung pro zu schreibendem Bit.*
- *Die lineare Bitdichte betrug 100 Bit/Zoll bzw. rund* 4 Bit/mm *mit resultierende Bitzeit 12 µs*
- *Die Entkopplung von Magnettrommel-Hauptspeicher und Plattenspeicher besorgte ein 100 BCD-Stellen großer Kernspeicher als Pufferspeicher.*

* Alle die Einheit 350 betreffenden Steuerungsabläufe beim Suchen, Lesen oder Schreiben wurden durch die Prozessoreinheit IBM 305 trotz der Zykluszeit von 20 ms geleistet, damit konnte das Ziel der Direktverarbeitung von etwa 10.000 Geschäftsvorfällen pro Tag erreicht werden.

Quellen

- IBM 305 Manual of Operation 1956
- IBM Journal of R & D, Vol 1, No. 1, January 1957
- IBM 305 Elektronische Simultanabrechnungsanlage, Form 74073, 4-1959
- Ch. Bashe a.o.: IBM's Early Computers, the MIT Press 1986

 350Profil.tmd

IBM DASD-Profile

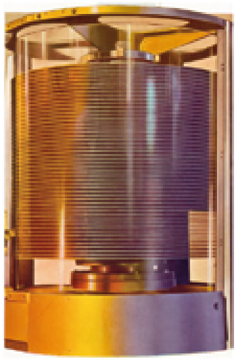
Einheit — **Plattenspeichereinheit 355,** angekündigt als DASD von System „650 RAMAC" in 1957.
Hier steht RAMAC für „**R**andom **A**ccess **M**emory **A**ccounting and **C**ontrol".
Entwicklungslabor: San Jose, für die Steuereinheit und den Pufferspeicher: Endicott.
Produktion: San Jose und Sindelfingen.

Die Magnetplattenspeichereinheit IBM **355**, 158 cm breit, 77 cm tief und 181 cm hoch; Gewicht 947 kg.

Modell — 1 mit 6 Millionen Stellen Speicherkapazität
2 mit 12 Millionen Stellen Speicherkapazität
Beide Modell nutzen „2 von 5 Bit"-Kodierung, die Datenrate beider betrug 10,5 Kilocharakter pro Sekunde.

Konzeption — Der mechanische Aufbau von Plattenturm, Zugriffstation und Doppelkopf incl. Kopfschwebeprinzip ist gleich dem von IBM 350, ebenso die Plattendrehzahl und die mittleren Zugriffzeiten. *Im Gegensatz zu 305/350,* konnten an die Steuereinheit IBM 652 und den erforderlichen Pufferspeicher IBM 653 *vier Einheiten 355* angeschlossen werden. Dies ermöglichte Direktzugriff auf 4 x 12 = 48 Millionen numerischen Zeichen. Die Transaktionsleistung 650/355 wurde determiniert durch die Plattenspeicherleistung (Zugriffszeit + Schreib/Lese-Zeit), doch initiierte der Prozessor lediglich über die Steuereinheit diese Vorgänge, die Durchführung oblag der Steuereinheit, unabhängig vom Prozessor. Letzterer erhielt z.B. beim Lesen die Daten und die abschließende *„Fertig und Korrekt-Meldung"* der 355.
Das EDV System 650 war ab 1956 mit Magnetbandeinheiten verfügbar. Damit konnte die *Datensicherung* von 355-Einheiten elegant durch Kopieren auf Magnetband erfolgen.

Technische Daten



Speicherplatte — *Plattenturm mit 100 nutzbaren Speicherflächen, fest eingebaut (siehe rechts)*
-Trägermaterial — *Alu, Dicke 1/10 Zoll* = **2,54 mm**
-Speicherschicht — *1,2 Millizoll Dicke* = **0,0305 mm** *aus* FeO^3
-Plattendurchm. — *24 Zoll =* **609,5 mm**
-Speich.-Spuren — *100 konzentrische Spuren pro Speicheroberfläche, Eine Speicherspur= 60 numerische Worte à 10 Stellen plus zugehörigem Vorzeichen.*
-Speicherdichte — *20 Spuren pro Zoll* = **1,27 mm Spurabstand**
-Plattendrehzahl — *1200 Upm*

Zugriffsmechanismus (Prinzip-Skizze siehe unter 1405-Profil)
- *Vertikalwagen mit Zugriffsarm, letzterer ist als Gabel ausgebildet die das Schreib/Lesekopf-Paar trägt. Der Vertikal-Wagen positioniert sich entsprechend der adressierten Speicherfläche, danach fährt der Zugriffsarm die adressierte Speicherspur an. Den Antrieb besorgen ein Motor, Magnetpulverkupplungen und ein bi-direktional laufendes Stahlseil. Die finale Positionierung des Vertikalwagens erfolgt pneumatisch mittels Präzisionskonus, die des Zugriffsarmes mittels Präzisionszahnstange und pneumatischer Klinkenverriegelung. Alle Zugriffsabläufe sind elektronisch gesteuert.*
- *Zugriffszeit max. 800 ms*
- *Optional konnten bis zu max. 3 Zugrifsstationen eingebaut werden. So konnte z.B. während des Schreibens einer Station, gleichzeitig mit der 2. Station Lesen vorbereitet und mit der 3. Station Suchen des nächsten Datensatzes (Speicher-Adresse) vorbereitet werden.*

Schreib/Lesekopf
- *Schwebehöhe über der Plattenoberfläche* = **0, 02032 mm** *, diese wurde pneumatisch erzielt und konstant gehalten durch Druckluft. Diesen „Kopf-Arbeitszustand" bewirkte -nach der Positionierung - ein Steuerventil durch das Signal „Kopf laden" für das Schreib/Lese-Kopfpaar.*
- *Kopfdurchmesser 5/8 Zoll* = **16 mm**, *Kopfhöhe* **4,5 mm**
- *der Schreib/Lesekopf enthält eine Schreib/Lesespule und eine Löschspule und determiniert durch die Jochdimensionen sowohl Spurbreite als auch die Bitdichte auf der Spur.*

Aufzeichnung auf der Spur
- *Non Return to Zero IBM (NRZI), d.h. ständiger Schreibstromfluß, umgeschaltet wird nur die Stromrichtung pro zu schreibendem Bit.*
- *Die lineare Bitdichte betrug 100 Bit pro Zoll bzw. rund* 4 Bit pro mm. *Die resultierende Bitzeit: 10 µs*
- *Entkopplung von Magnettrommel-Hauptspeicher und Plattenspeicher erfolgt durch einen Pufferspeicher mit 60 numerischen Worten (à 10 Stellen & Vorzeichen), technisch ausgeführt als Kernspeicher in 2 von 5-Kodierung, d.h. mit voller Gültigkeitsprüfung.*

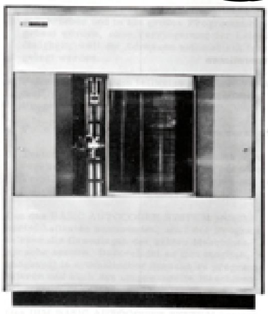
Quellen

- IBM 355 CE Manual of Operation, Form 724-56200, 1957
- IBM 650 RAMAC, Manual of Operation, Form 22-6270, 1956
- Ch. Bashe a.o.: IBM's Early Computers, the MIT Press 1986
- IBM 355 CE Reference Manual Form 724-56210, 1957
- IBM 650 RAMAC, Faltprospekt, Form 74074, 1959

355Profil.tmd

Einheit Plattenspeichereinheit 1405, angekündigt Okt. 1960
als DASD des Systems „1401 RAMAC".
Entwicklungslabor San Jose,
Produktion: San Jose und Sindelfingen

Modell Plattenspeicher mit vertikal rotierendem Speicherplattenturm.
Direktanschluß an den Prozessor, letzterer steuert alle Abläufe.
1 mit 10 Millionen BCD-kodierter Stellen Speicherkapazität
2 mit 20 Millionen BCD-kodierter Stellen Speicherkapazität
Datenrate: beide Modelle 22,5 Kilo-BCD-Zeichen pro Sekunde.



IBM **1405**
Höhe 220 cm
Breite 197 cm, Tiefe 84 cm
Gewicht 920 kg

Konzeptionelles

Konzeptioneller Aufbau und Struktur der IBM 1405 folgen dem Vakuumröhren-Vorläufer IBM 350/355, doch ist die Schreib/Lese- und Steuerungstechnik voll transistorisiert mittels der IBM Standard Modular System Technologie (SMS-Technik). Die Kopfschwebehöhe wurde verkleinert, dadurch erhöht sich die Speicherdichte, außerdem wurde die Spurendichte verdoppelt auf 200/Zoll.

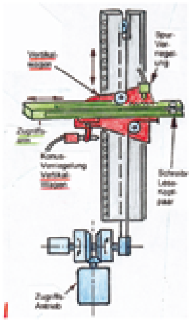
Jede Speicherspur ist eingeteilt in 5 Sektoren mit je 200 alphanumerischen Stellen. Die 1405-Adressierungskomponenten sind *Plattenoberfläche* (1 von 100), *Speicherspur* (1 von 200) und *Sektor* (1 von 5), d.h. ein Datensatz mit 200 alphanumerischen Stellen kann direkt adressiert werden.

Technische Daten

Speicherplatte

Plattenturm.......................*mit 100 nutzbaren Speicherflächen, fest eingebaut.*
-Trägermaterial.................*Aluminium, Dicke 1/10 Zoll* = **2,54 mm**
- Speicherschicht..............*Dicke 0,9 Millizoll* = **0,023 mm** *aus* FeO^3
- Durchmesser..................*24 Zoll* = **609,5 mm**
- Speicherspuren..............*100/200 konzentr. Spuren pro Oberfläche, Modell-abhängig*
- Speichersp.-Dichte..........*40/80 Spuren pro Zoll, Modell-abhängig*
- Speicherpl.-Drehzahl.......*1200 Upm*

Zugriffsmechanismus



- *Vertikalwagen mit Zugriffsarm, letzterer ist als Gabel ausgebildet, die das Schreib/Lesekopf-Paar trägt. Der Vertikal-Wagen positioniert sich entsprechend der adressierten Speicherfläche, danach fährt der Zugriffsarm die adressierte Speicherspur an. Wagen- und Zugriffsarm-Antrieb besorgt ein Motor, Magnetpulverkupplungen und ein bi-direktional laufendes Stahlseil. Die finale, pneumatische Positionierung des Vertikalwagens erfolgt mittels Präzisionskonus, die des Zugriffsarms mittels Präzisionszahnstange und pneumatischer Klinkenverriegelung. Alle Zugriffsabläufe sind elektronisch gesteuert.*
- *Zugriffszeit max. 800 ms*
- *Die 1405-Standardausführung hatte eine Zugriffsstation, doch konnte optional noch eine zweite eingebaut werden. Dies erlaubte z.B. auf Station 1 zu Lesen/Schreiben und gleichzeitig auf Station 2 eine Suchoperation zur Vorbereitung der Bearbeitung des nächsten Datensatzes durchzuführen.*

Schreib/Lesekopf

Der Schreib/Lesekopf enthält eine Schreib/Lesespule und eine Löschspule, deren Jochbreiten determinierten die Spurbreite und Bitdichte auf der Spur.
- *Schwebehöhe über der Plattenoberfläche = 0, 0165 mm , diese wird pneumatisch erzielt und konstant gehalten durch den Effekt des Bernoully-Gesetzes mittels Druckluft. Der Schwebzustand wird nach der S/L-Kopf-Positionierung auf Platte und Spur eingenommen, ausgelöst durch das elektronische „head load signal".*
- *Kopfdurchmesser), 5/8 Zoll* = **16 mm**, *Kopfhöhe 4,5 mm*

Spur-Aufzeichnung

- ***N**on **R**eturn to **Z**ero IBM (NRZI), d.h. ständiger Schreibstromfluß, doch wird die Stromrichtung pro zu schreibendem Bit umgeschaltet und damit der Magnetfluß in der Speicherspur.*
- *Die lineare Bitdichte beträgt 220 Bit pro Zoll.*
- *Schreiben und Lesen erfolgt direkt „aus dem" oder „in den" Hauptspeicher, technisch ausgeführt als Magnetkernspeicher mit 11,4 μs Zykluszeit pro BCD-kodierte Stelle.*
- *Die Bitzeit in der Speicherspur beträgt ~ 5μs.*
- *Beim Lesen steuern zwei Oszillatoren die Zeitgebung.*

Quellen

- IBM 1405 Manual of Operation, 1961
- IBM 1405 Manual of Instruction. 1961
- IBM Journal of R & D 25th Anniversary, 5-1981
- Ch. Bashe a.o.: IBM's Early Computers, the MIT Press 1986

 Prof1405.tmd

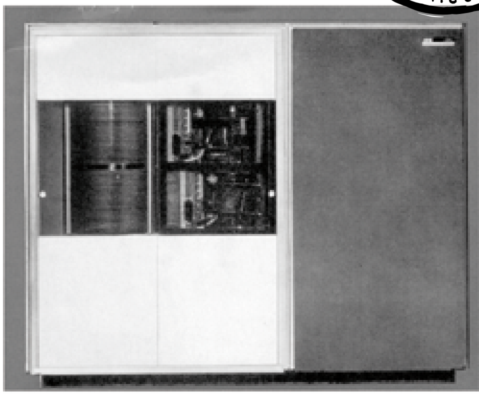
IBM DASD Profile

Einheit **Plattenspeichereinheit 1301**
Ankündigt am 2. Juni 1962, Entwicklung und Produktion in San Jose California.

Anschluß an Systeme bei Ankündigung:
IBM 7000 System Serie und System IBM 1410

Modell 1 Einheit mit einem Modul, Speicherkapazität 28 Mio Zeichen (an IBM 70xx) bezw. in gepacktem Format= 43,3 Mio Ziffern.

2 Einheit mit zwei Modulen, Speicherkapazität 56 Mio Zeichen (an IBM 70xx) bezw. in gepacktem Format = 86,6 Mio Ziffern.



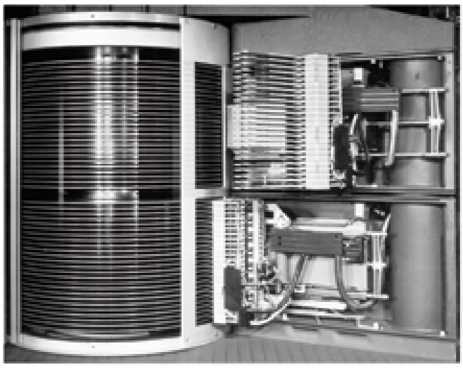
Plattenspeichereinheit IBM 1301, links die beiden Plattenmodule übereinander, rechts davon die Zugriffskämme, ganz rechts verdeckt Hydraulik und Steuer-Elektronik

Generelles
Vorbemerkung: Bei 1301 treibt eine gemeinsame Welle zwei Plattensätze an. Ein Satz wird *Modul* genannt. Jedes Modul hat seinen eigenen Kamm-Zugriffsmechanismus, der eine horizontale Bewegung *aller Köpfe gleichzeitig* zur Spur/Zylinder-Selektion ausführt. Dieses neue *Zylinderkonzept* ermöglicht Zugriff i*nnerhalb von Zylindern* durch elektronisches Schalten.

Leistungsparameter
Die **Speicherkapazität** an Systemen beträgt bei
- *IBM 1410* = 25 MC [Millionen alphanum. Zeichen (BCD-Kode)] und 2 Module x max. **5** Einheiten = **250 MC**
- *IBM 7070* = 28 MC [Millionen numer. Zeichen (2 out of 5 Kode)] x 2 Module x max. **5** Einheiten = **280 MC**

Die große Speicherkapazität und deren Verwaltung forderten die Erweiterung von IBM *IOCS* (Input-Output-Control System) zur Unterstützung der Anwender-Systemprogrammierer.
Der *lineare* Kammzugriff pro Modul erfolgte in *50 - 165 ms, „im Zylinder“ in 34 ms*
Datenraten sind System-abhängig: *90 kC/s an 7000er Systemen, 75 kC/s an System 1410.*

Technische Daten
- Speicherflächen — ***Pro Modul:*** *20 +1 Zeitgeber +1 Format + 6 Reserve*
- Plattendrehzahl — *1790 Upm*, ***34 ms*** */ Umdrehung*
- Plattendurchmesser — *24 Zoll* = **600 mm**
- Speicherschicht-Art — *Partikular beschichtet mit epoxyharz-gebundenem* Fe_2O_3
- Speicherschichtdicke — *0,001 Zoll* = **0,0254 mm**
- S/L-Kopftechnologie — Kopfform und Luftführung ermöglichen ***erstmals*** den *„aerodynamisch fliegenden S/L-Kopf“, ausgeführt in Mu-Metall-Technik.*
- *Kopf-Flughöhe* — *250 µZoll* = **0,00635 mm**
- *Aufzeichnungsspuren* — *50 pro Zoll* = **1,96 pro mm**
- *Zeitgeber* — *Schreiben: Zeitgeberspur auf Platte*
  Lesen: 625 kHz Quarzoszillator,
  (625000 Bits/s, eine „Bit-Zellenzeit“ =1,6µs)
- *Schreibverfahren* — *NRZI (Non return to zero IBM)*
- Flächenspeicherdichte *0,026 MBit/Zoll*^{2} = **0,004 MBit/qcm**



Im Bild ist das untere S/L-Kopf-Modul in Arbeitsstellung, das obere Modul ist mit ausgefahrenem S/L-Kopfsatz abgebildet.

Steuereinheit: IBM 7631 *Model 1 =* Anschluß an IBM 1410
2 = Anschluß an IBM 7070/7074/7080/7090
3 = Anschluß „selektiv geschaltet“ zu IBM 70xx oder 1410 (shared)
4 = Anschluß „selektiv geschaltet“ zu zwei IBM 70xx Systemen

1963/4: Anschluß an IBM 1440 & 1460: max 5 Einheiten pro System mit total 100 MC Speicherkapazität.

Quellen

1. General Information IBM 1301 with Series 7000/1410 Systems, 1961, 1962
2. 1301 CE Manual of Instruction, Form 227-5582, 1962
3. 1301 Disk Storage Models 11, 12, 21 and 22 System Reference Library Form A24-3157, 1963
4. Charly Bashe and others: IBM's Early Computers, MIT Press, 1986

Einheit

Plattenspeichereinheit IBM 1311 mit geöffneter Abdeckung zur Auswechselung des Plattenstapels.

Breite	110 cm
Höhe	103 cm
Tiefe	61 cm
Gewicht	180 kg

Plattenstapel Typ IBM 1316 mit 10 Speicheroberflächen.

Durchmesser	37 cm
Höhe	10 cm
Gewicht	4,25kg

Plattenspeichereinheit 1311

Generelles

IBM bot bis dahin großvolumige Plattenspeicher mit 24 Zoll Plattendurchmesser an, deren preisliche Schwelle lag entsprechend hoch. Mit dem System IBM 1440 und dessen Plattenspeicher IBM 1311 - angekündigt 10-1962 - gelang es IBM, auch mit kleineren Systemen *Direktverarbeitung* anbieten zu können.

Das Neue und besondere an der Plattenspeichereinheit 1311 war ein durch den Operator schnell auswechselbarer Plattenstapel 1316 mit 14 Zoll Plattendurchmesser mit 2 Millionen alphanumerischen Zeichen Speicherkapazität. Die Auswechselbarkeit ermöglichte unbegrenzte off-line Speicherkapazität im Stapel-Regal.

Kunden, die bereits mit 1401, 1410 und 1620 Systemen arbeiteten, waren für bestimmte Anwendungen ebenfalls am neuen Speicher interessiert. Dies führte, zu den folgenden Modellen - denn eine Standard-Schnittstelle gab es - *noch* - nicht:

Modell *1*	Haupteinheit mit Steuerung und Stromversorgung für weitere Modelle 2 zum Anschluß an System *1440*
Modell *2*	Nebeneinheit zum Anschluß an Haupteinheiten, maximal vier.
Modell *3*	Haupteinheit zum Anschluß an Systeme *1620* und *1710*
Modell *4*	Haupteinheit zum Anschluß an System *1401*
Modell *5*	Haupteinheit zum Anschluß an die Kanäle der Systeme *1410/7010*

Leistungsparameter und Adressierungskomponenten

Der *auswechselbare Plattenstapel 1316* wird über eine vertikale Präzisionskonus-Verbindung angetrieben, enthält 10 Speicheroberflächen und rotiert mit 1500 Umdrehungen pro Minute in staubgeschützem Gehäuse. Die Datenrate beträgt 69 Kilozeichen pro Sekunde. Ein Trägerkamm mit 20 Schreib/Lese-Köpfen - einer pro Speicheroberfläche - bewegt sich programmgesteuert, hydraulisch angetrieben, horizontal zu *einer* von insgesamt 100 Speicherspuren und wird dort per Zahnrad und Klinke final positioniert. Zugriffszeit im Mittel 150 Millisekunden.

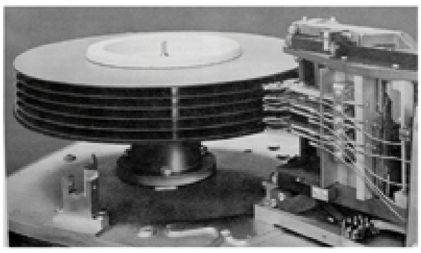
Die Spurorganisation arbeitet mit 20 Sektoren mit je 100 alphanumerischen Zeichen pro Spur. Die Adressierung für System 1401 - mit 1x1311 Haupteinheit und max. vier Nebeneinheiten erfolgt durch: Einheit (**1 - 5**), Plattenoberfläche (**oben - unten**), Spur (**001-100**) und Sektor (**1-20**).

Einsatzbereiche

Für Anwendungen mit kleineren bis mittleren Datenvolumen auf kleineren bis mittleren Systemen. Die Speichereinheit 1311 selbst generierte Anwendungswachstum in kleineren EDV-Anlagen, da Kunden für bestimmte Aufgaben weniger Lochkarten benötigten wegen direkten Zugriffs auf bereits gespeicherte Daten. Darüber hinaus konnten Anwendungsprogramme schnell und bequem direkt vom Plattenspeicher geladen werden. Systemsoftware *Input-Output Control System* unterstützte Systemprogrammierung und Speicherverwaltung ab 1963.

Technische Daten

- Speicherflächen	*10 pro Plattenstapel*
- Plattendrehzahl & Durchm.	*1500 Upm*, *14 Zoll* = **350 mm**
- Speicherschicht	*Partikular beschichtet mit Epoxyharz-gebundenem* $Fe^{2}O^{3}$
- Speicherschichtdicke	*0,001 Zoll* = **0,0254 mm**
- S/L-Kopftechnologie	*Aerodyn.fliegender S/L-Kopf, Flugkörper Stahl, S/L-Element Mu-Metall*
- Kopf-Flughöhe	*125 µZoll* = **0,0318 mm**
- Aufzeichnungsspuren	*100 pro Speicherfläche*
- Zeitgeber	***Quarz*** *699,3 kHz*, *pro Bitzelle= 1,43 µs*
- Schreibverfahren	*NRZI (Non return to zero IBM)*
- Flächenspeicherdichte	*0,051 MBit/Zoll*^{2} = **0,008 MBit/cm²**



Plattenstapel in **Arbeitsposition** innerhalb der Einheit 1311, rechts davon der Schreib/Lesekopf-Kamm (Die Verkleidung und Staubsicherung wurde zur Photoaufnahme entfernt).

Quellen

- SRL, IBM 1311, Modell 3, Form A26-5650, 1962
- Reference Manual IBM 1311 OEM Information, Form A26-5713, 1963
- Reference Manual IBM 1316 Disk Pack Requirements, Form A26-5747, 1966
- W. Fink: Informationen über Prinzipielles in der Magnetkopf- und Plattentechnik, 1968
- Charly Bashe and others: IBM's Early Computers, MIT Press, 1986

pro1311g.tmd

IBM DASD Profile

Einheit **Plattenspeichereinheit 2311**
Ankündigung: April 1964

Entwicklungslabor: *San Jose,*
Produktion: *San Jose, Sindelfingen und/* [illegible]

Generelles
Für Anwendungen mit mittleren Speichervolumen wurde die Plattenspeichereinheit 2311 geschaffen zum Betrieb an kleinen bis größeren S/360 Prozessoren. 2311 hat ein Laufwerk für den auswechselbaren Plattenstapel 1316 als Speichermedium.
Modell 1 wird mit der Steuereinheit 2841 betrieben, letztere ist über einen Selektorkanal mit dem Prozessor verbunden. Die Speicherkapazität pro Plattenstapel beträgt **7,25 MB.**
Modell 2 ist für das S/360-Modell 20 mit *Direktanschluß* eingerichtet, die Speicherkapazität pro Plattenstapel beträgt dabei **5 MB.**

Splattenspeicher-einheit IBM 2311. Rechts oben: Zugangsklappe zum Plattenstapel, links oben die Bedienerkonsole.

Leistung:	**Speicherkapazität**	**Datenrate**	**Zugriffszeit**
Modell **1:**	**7,25 MB**	**156 kB/s**	**85 ms** im Mittel, im Zylinder **0 - 25 ms**
Modell **2:**	**5 MB**	**156 kB/s**	**85 ms** im Mittel, im Zylinder **0 - 25 ms**

Steuereinheit: Ist IBM **2841-** Modell 1**,** anschließbar max. 8 Laufwerke 2311, Mikroprogrammsteuerung. Formate- und Zylinder variabel nutzbar. Der Systemprogrammierer entscheidet anwendungsabhängig.

Betriebssystem-Unterstützung
Systembetrieb mit dem Betriebssystem DOS oder OS ermöglichten den Betrieb vieler Anwendungen, deren Datenvolumen mit 2311 bewältigt werden konnten.

Technische Daten

Speicherflächen	*10 pro Stapel (gleicher Plattenstapel wie bei 1311) ,*
- Plattendrehzahl	***2400 Upm,** d.h. **25 ms** / Umdrehung, durchschn. Latenzzeit 12,5 ms*
- Plattendurchmesser	***14 Zoll / 356 mm,** Trägermaterial AlMg5 diamantgedreht*
- Speicherschicht-Art	*Partikular beschichtet (ca. 30%) mit Epoxyd-Phenolharz - gebundenem* Fe_2O_3
- Speicherschicht-Dicke	***ca 5 µm***
S/L-Kopftechnologie	
- Glaskeramikflugkörper	*Aerodynamisch fliegend, Kurvenradius mit 6,35 m ø, 2 Stabilisierungslöcher*
- S/L-Element	*Ferrit, mit Epoxyharz eingeklebt.*
- Spaltbreite	***2,5 - 3 µm***
- Kopf-Flughöhe	***3 - 5 µm** (innere/äußere Spur)*
Speicherung	
- Aufzeichnungsspuren	***203** pro Speicherfläche* (bei gleichem Plattenstapel 1316 = ***Verdoppelung***)
- Zeitgeber	***Quarz 2,5 MHz**, Quarzoszillator, 2 Zeitgeberzyklen pro Bitzelle, Bitzellenzeit = 400 ns; beim Lesen: Variable Frequency Oscillator ,*
- Schreibverfahren	***NRZI** (Non return to zero IBM), **Double Frequency**, Schreibstrom **70 mA***
- Lesen (Detektion)	***Peak,** Lesesignalgröße ~ **2 mV***
- Lineare Bitdichte	***1,1 kBit pro Zoll** =* **44 Bits pro mm**
- Spurendichte	***100** Spuren pro Zoll* = **3,9** Spuren pro mm
- Flächenspeicherdichte	***0,026 MBit/Zoll²*** = **0,004 MBit/cm²**
Zugriffsmechanismus	*Hydraulisch bewegter Zugriffskamm, dieser trägt 10 S/L-Köpfe und positioniert diese horizontal auf einer der 203 Spuren .Die finale Spurposition wird fixiert mittels Präzisionszahnrad und hydraulisch betätigter Klinke durch elektronische Steuerung.*

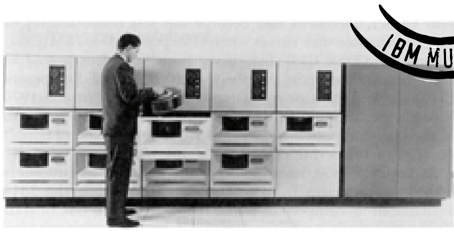
Quellen

1. CE Reference Manual IBM 2311, 1966
2. CE Manual of Instruction IBM 2311, 1966
3. S/360 DASD-Handbuch zur Datenorganisation auf DASD's, Form 79 966, 1966
4. IBM WS AP Unterlagen W. Fink, Jan. 1968
5. IBM Journal of R&D, 25th Anniversary Edition, 9-1981
6. Charly Bashe and others: IBM's Early Computers, MIT Press, 1986

.

 prof2311.tmd

Einheit *Direct Access Storage Facility 2314*
(Speicheranlage) angekündigt im April 1965
für auswechselbare Plattenstapel IBM 2316.
Entwicklungslabor San Jose
Produktion in San Jose und Mainz



Anschluß An *S /360* Prozessormodelle 40 und größer.
Das erste Plattensystem, bei dem ***Steuereinheit und ein Strang mit Laufwerken - bis zu max. 8 - in einer Konfiguration*** bestellt und aufgebaut werden konnte (2313 = Einheit mit 4 Laufwerken, 2312 = Einheit mit 1 LW (als Reserveeinheit) und einer integrierten Steuereinheit (File Control Unit). Die 2314-Laufwerke nutzen den *auswechselbaren,* neuen Plattenstapel *IBM 2316* mit 22 Speicheroberflächen und ***29,17 MB Speicherkapazität.*** Die in 2314 integrierte Steuereinheit ist ähnlich 2841. Für den wahlweisen Betrieb an zwei Kanälen ist optional ein **Kanalumschalter** in der Steuereinheit **IBM 2844** möglich.

.Leistungsparameter

- Die **Speicherkapazität** pro Anlage **2314** wird durch max. 8 Laufwerke gebildet, sie beträgt 8 mal 29,17 MB = **233 MB.** Alle 8 Laufwerke werden auch als Strang bezeichnet.
- Die **Datenrate beträgt *312 kB / s***

Zugriffsmechanismus Jedes Laufwerk hat einen Zugriffskamm mit je einem S/L-Kopf pro Plattenoberfläche. Insgesamt werden 22 S/L-Köpfe über den Zugriffskamm hydraulisch positioniert.
Datenzugriffszeiten: Mittlere=***75 ms,*** *wenn „im Zylinder"=**0- 25 ms;** (erreicht durch „3-Geschwindigkeits-Actuator), die finale Position wird fixiert durch Präzisionszahnrad und Klinke.*

Technische Daten

- Plattenanzahl:	***11*** *pro Stapel =* ***22 Speicheroberflächen*** *pro Stapel (20 Datenspeicherflächen, 1 Indexfläche und 1 Servofläche).*
- Plattendrehzahl:	***2400 UpM****, d.h.* ***25 ms*** */ Umdrehung, durchschnittliche Latenzzeit 12,5 ms; Plattengeschwindigkeit am Außendurchmesser etwa 160 km / h.*
- Plattendurchmesser:	***14 Zoll*** *(***356 mm***)*
- Speicherschicht-Art:	*Partikular beschichtet (ca. 30%) mit Epoxyd-Phenolharz-gebundenem* Fe_2O_3
- Plattenträgermaterial:	***AlMg5,*** *diamantgedreht, Oberflächengüte 3,2 Mikrozoll*
- Speicherschichtdicke:	***1 Millizoll*** *(=***2,5 Mikrometer***)*
- S/L-Kopftechnologie:	*Kopfform sorgt für aerodynamisch fliegenden Schreib/Lesekopf" aus Mu-Metall mit 2 Stabilisatorbohrungen (Kurvenradius der Flugfläche 250 Zoll =* **6,35 m** *besorgt keilförmige, stabile Luftführung)*
- Kopf-Flughöhe:	***0,8 -1 Millizoll*** **= (2 - 2,5 µm)**
- Aufzeichnungsspuren:	***100 tpi*** *(d.h. Spurdistanz =* **0,254 mm***)*
- Zeitgeber:	*Quarz* ***4 MHz,*** *2 Zeitgeberzyklen pro Bitzelle, Bit-Zellenzeit =* ***400 ns;*** *beim Lesen: Variable Frequency Oscillator.*
- Schreibverfahren:	***NRZI*** *- Double Frequency, Schreibstrom* ***30 mA***
- Detektion:	***Peak***
- S/L-Spaltbreite:	***2,3 bis 3*** *Mikrometer*
- Lineare Bitdichte:	***2,2 kBit pro Zoll*** *(* **88 Bits pro mm***)*
- Flächenspeicherdichte:	***0,22 MBit /Quadratzoll*** *(***34 kBit/cm^2***)*

Fazit Die 2314-Marktakzeptanz war weit besser als erwartet, auch durch das Betriebssystem OS/360, das es ermöglichte, ein System über ausgedehnte Zeitabschnitte ohne Operator-Eingriff zu fahren. Mit 2314 war Zugriff zu großen Datenbanken möglich geworden. Das Betriebssystem konnte vollautomatisch überleiten von einer Anwendung zur nächsten. OS/360 half mit, IBM 2314 zum profitabelsten Speicherprodukt seiner Zeit zu machen und zusätzlich beizutragen zum Erfolg der Systemfamilie IBM S/360.

Quellen

1. General Information IBM 2314, Form ( )
2. 2314 System Reference Library, Form ( )
3. 2314/2844 FE Theory of Operation, Form SY-3671, 1966
4. Charly Bashe and others: IBM's Early Computers, MIT Press, 1986
5. IBM Journal of R&D, 25th Anniversary Edition, 9-1981
6. IBM WS AP Unterlagen W. Fink Jan. 1968

 DASD2314.tmd

IBM **DASD-Profile**

Einheit Plattenspeichereinheit 5444, angekündigt 1969 für das System /3 Modell 15, als das *kleinste IBM System mit DASD's seiner Zeit.* Ab 2/1971 wurde die DASD-Kapazität ergänzt mit IBM 5445.

Steuereinheit: Direktanschluß an die Systemeinheit von S/3

	5444*	5445**
Speicherkapazität:	*2,5/ 7,5 MB* pro LW	*20 MB* pro LW
Datenrate	*199 KB/s*	*312 KB/s*

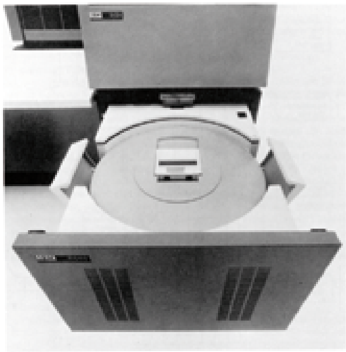


Bild oben: Magnetplatteneinheit 5444 mit Direktanschluß a.d. ZRE, Blick auf ausgezogene Schublade auf die **Wechselplattenkassette.** Daruter befindet sich die **Festplattenspindel**

Technische Daten

Aufzeichnungsdichte-
- Lineare Bitdichte...*1,1 KBit/Zoll =44 Bit/mm*.....*2,2 KBit/Zoll = 88 Bit/mm*
- Spurendichte...................................*100/Zoll = 3,95/mm*
- Flächenspeicherdichte *0,11 MBit/Zoll² = 0,017 MB/cm².*
  0,22 MBit/Zoll² = 0,034 kBit/cm²

Schlüßelparameter
- Kopf/Platte-Flughöhe......... *4 µm*...............................*2,5 µm*
- Kopf-Luftspalt in..................*ca. 3 µm*.........................*2,3-3,0 µm*
- Speicherschicht..................*ca. 5 µm*............................*2,5 µm*

Platte-
- Durchmesser...*14 Zoll*
- Technik..*"Braune Platte"*
- Drehzahl ....*1500 UpM*.....................*2400 UpM*
- Latenzzeit...
- fest/auswechselbar..........*1 fest, 1 ausw'bar* ....*1 ausw'bar*
- Speicheroberflächen / Spindel...*4*.................................... *22*
  (die untere ist fest),
  (die obere in auswechselbarr Kartusche).

Aktuator
- Zugriffsgeometrie.............*Linearbewegung bei allen Modellen*
- Köpfe-Anzahl............................*4*...*22*.....
- Positionierung.................................... ..*per Hydraulik.*
- Finale Position.............*per hydr. gest. Klinke & Präzisionszahnrad*

Bild oben: Magnetplatteneinheit 5445 mit Schublade für den Standard- Wechselplattenplattenstapel 2516 (in der Mitte).

Weitere S/3-DASD-Kapazitätszeigerung.
Die zweite Erhöhung der S/3-DASD-Kapazität auf 82 MB erfolgte 1975 bei S/3-M12 durch die Anschlußoption IBM 3340 (siehe auch DASD-Profil IBM 3340) über die Einheit 5447, eine kombinierte Plattenspeicher-und Steuereinheit in einem Gerät.

* DASD 5444 nutzte die 2311-DASD-Technologie in modifizierter Form.
** DASD 5445 nutzte die 2314-DASD- Technologie in modifizierter Form.

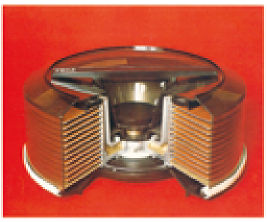
Quellen:
- Das IBM System /3 Modell 15, Form F12-1596, 9-1974, und Herr D. Christoph
- IBM Journal of R&D, 25th Anniversary Issue, 1981
- IBM Rochester: A Half Century of Innovation, 2006 (1956 - 2006)

ASD54445.tmd

Einheit **Speicheranlage 3330**
angekündigt 1970

DASD-Speicheranlage ist modular aufgebaut (Bild rechts), kann in maximaler Ausrüstung **800 MB** speichern auf Plattenstapeln. Betrieb an S/370 - Großsystemen.

Plattenstapel Type **3336**-001 mit 100 MB Speicherkapazität als Schnittmodell (Bild unten)



x-----------------------x-----------------------x-------------------x-------------x----------x

bis zu max. 8 LW insgesamt.	2 LW **3330**	2 LW **3330**	2 LW mit Steuereinheit **3330**	**3830-1**

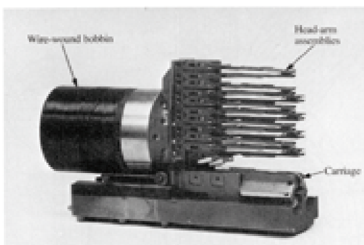
Konzeption der Speicheranlage
- Auswechselbare Plattenstapel (Type 3336-001) als **Speichermedium** mit **100 MB** Kapazität
- in Schubfächer eingebaute Laufwerke, pro Laufwerk Antriebsspindel, Zugriffsarm und Schreib/Leseköpfe in Kammanordnung
- Anschlußkonzept für max. **8 Laufwerke** „pro Laufwerkstrang"
- Ein 3330-Subsystem mit Maximalausrüstung bildet **800 MB** ab.
- Die **Datenübertragungsrate** beträgt **806 kB/Sekunde** für alle Modelle.
- **Unbegrenzte** off-line Speicherkapazität.

Speichereinheit
- 3333-001 Haupteinheit, versorgt Laufwerke der angeschl. Nebeneinheiten und die Stromversorgung des Stranges. Enthält 2 Laufwerke.
- 3333-002 Haupteinheit, versorgt zwei weitere Stränge (bei Verdoppelung der Speicherkapazität)
- 3330-001 Nebeneinheit mit 2 Laufwerken
- 3330-002 Nebeneinheit mit Einzelaufwerk
- 3330 ist erstes IBM DASD-Laufwerk mit Voice Coil Aktuator (siehe Abb. unten)

Steuereinheit **3830-1: Standard, 3830-2** mit Zweistrang-Steuerung
- 3830-3
- Automatische Positionsbestimmung (RPS) und Türöffnung Standard
- **Befehlswiederholung** (Fehlerkorrektur ist verbessert dadurch, daß Kanal-und Steuereinheit ohne Programmeingriff Befehle wiederholen können).
- Erweiterte **Fehlererkennung** und Korrektur.
- **2-Kanalschalter:** erlaubt den Anschluß der Steuereinheit an 2 Kanäle
- **Positionsbestimmung:** ermöglicht Kanalfreigabe während Suchoperationen
- Anschliessbar ab S/370- **ab /370-125** aufwärts

Programmunterstützung Betriebssysteme OS oder VS1/2 oder MVS oder VM/370

Weitere technische Daten
- *Zugriffszeit durchschnittlich* ***30 ms***
- *Latenzzeit durchschnittlich* ***8,4 ms***
- *Drehzahl* ***3600 Upm***
- *Plattendurchmesser* ***14" = 350 mm,*** *partikular beschichtet, Schichtstärke* ***1,27*** *Tausendstel Millimeter*
- *S/L-Kopf* ***Ferrit*** *in Keramikflugkörper eingeklebt*
- *Flughöhe* ***1,25*** *Tausendstel Millimeter*
- *Schreibverfahren* ***mfm***
- *Lineare Bitdichte* ***4400 bpi / 176 Bits pro mm***
- *Spurendichte* ***192 tpi / 7,6 tpmm***
- *Flächenspeicherdichte* ***780 kBit/ Zoll²*** */* ***121 kBit/qcm***



Quellen
- Datenblatt Vertrieb-D, Form O12-0229-1
- DPCE Folder ohne Formnummer
- Überblick über das IBM System /370, Form GA12-1040-5, Seite 7-13
- IBM Journal of R&D, 25th Anniversary, 1981, page 678

 DASD3330.tmd

IBM **DASD-Profile**

Einheit **Plattenspeichereinheit 3340,** angekündigt 1973 als Teil einer modular angelegten Speicheranlage für Mittlere Systeme in Winchester* - Technik mit auswechselbarem **Datenmodul**, d.h. der Plattenstapel wird *zusammen* mit dem Schreib/Lesekopfkamm durch den Operator gewechselt. * Entwicklungsname.

2 Einheiten IBM 3340, rechts die Kopfeinheit mit eingebauter Steuereinheit
Quelle: DPCE Prod. Folder IBM 3340

Speicher-anlage-konzeption Eine 3340-Speicheranlage setzt sich aus bis zu 4 Einheiten 3340 zusammen und damit aus max. 8 Laufwerken an einem Strang. 3 Modelle sind verfügbar: A02 mit zwei Laufwerken und der Steuereinheit für den ganzen Strang (Kopfeinheit), B02 mit zwei Laufwerken und B01 mit nur einem Laufwerk.

Ein Datenmodul IBM 3348 **speichert** als Modell 35=**34,94 MB**, als Modell 70=**69,89 MB.** Damit stehen pro Laufwerkstrang max. 560 MB Speicherkapazität on-line zur Verfügung und zusätzlich pro Modell 70 als Option 0,5 MB Speicherkapazität mit Festköpfen. Datenmodule werden durch den Operator anwend-ungsgemäß ausgewechselt.

Neu ist die **Schreibsperre** -ein Hebel am Daten-modul- der unbeabsichtigtes Schreiben verhindert.

Die Luft im hermetisch abgeschlossenen Daten-modul bildet über Mikrofilter einen geschlossenen Kreislauf zur Vermeidung von Kontaminierung.
Die **Datenrate** beträgt **885 kB/s** bei 3000 UpM, die mittlere **Zugriffszeit 25 ms**

Datenmodul IBM 3348, der Griff von links oben zur Mitte und dahinter der mechanische Anschluß zum Aktua-tor der S/L-Kopfpositionierung.

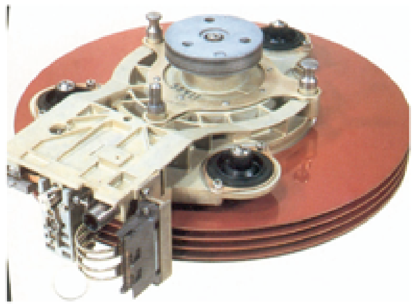
Die mechanische Kopplung von Plattenantrieb und Zugriffs-kamm erfolgt vollautomatisch nach drücken des Startknopfes..

Steuereinheit i.d. Kopfeinheit **von 3340**
- Fehlererkennung durch die Steuereinheit, diese stellt Korrekturdaten bereit bei Lesefehlern von bis zu drei aufeinander folgenden Bitfehlern in einem Datensatz (Error Correction Coding = ECC)
- **R**otational **P**osition **S**ensing = **RPS**, ermöglicht die Freigabe der Steuerung während Suchoperationen
- Anschliessbar an S/370 **- 115, 125, 135** und **145**

*Anmerkung: Der Entwicklungsname ver-bindet die angestrebte Speicherkapazität von ~30 MB/Modul in 30 ms Zugriff mit dem amer. Gewehrtyp „Winchester 30-30"

Programmunterstützung: DOS / VS

Technische Daten
- *Zugriffszeit durchschnittlich* ***25 ms***
- *Drehzahl* ***3000 Upm****, Latenzzeit* ***10,1 ms***
- ***Platten****durchmesser* ***14 Zoll / 350 mm***
- *Trägerplattendicke* ***1,9 mm***
- *Speicherschichtstärke* ***1,04*** *Tausendstel Millimeter partikular beschichtet,*
- ****S/L-Kopf*** *und Flugkörper aus* ***Ferrit****, erstmals Doppelkopf (Weghalbierung), Auflagekraft nur 10 g; Schmierung, Landebahn, ohne Lademechanismus!*
- *Kopfflughöhe (nur noch)* ***18*** *Mikrozoll =* ***0,457*** *µm*
- *Schreibverfahren: mfm*
- *Detektion: delta*
- *Lineare Bitdichte* ***5636*** *bpi (Bit/Zoll) /* ***221,88*** *Bit/mm*
- *Spurdichte:* ***300*** *tpi (tracks/inch) /* ***11,81*** *Spuren/mm*
- *Finale Position: per Servospur*
- *Flächenspeicherdichte:* ***1,69 MBit/Zoll***^{2}, ***0,262*** *MBit/cm*^{2}



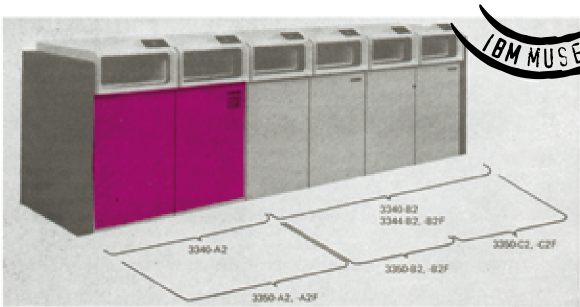
Das Bild zeigt das Datenmodul 3348 ohne hermetisch dichte Kunststoffver-kleidung und „auf dem Kopf" stehend photographiert.

Quellen
- DP CE Product Folder 1973
- DV Information 13-1973, IBM Deutschland
- Pugh, Johnson and Palmer: IBM's 360 and Early 370 Systems, MIT Press, 1991

 DASD3340.tmd

Einheit Plattenspeichereinheit 3350 angekündigt 1976 als Teil einer modular angelegten Speicheranlage für Großsysteme.

Eine Speicheranlage als Subsystem mit mit hoher Speicherkapazität und Datenrate für Kanalbetrieb über die Speichersteuereinheit IBM 3830.



System-konzeption
- Fest eingebautes und versiegeltes **Speichermedium** bestehend aus Platten, Spindel, Zugriffsarm und Schreib/Leseköpfen, genannt Head Disk Assembley (HDA).
- Pro Laufwerk werden max. 317.498.850 Bytes gespeichert = 317,5 MB.
- Alle Modelle haben eine Datenübertragungsrate von 1,198 MB/Sekunde.
- Laufwerke sind über Kopfeinheit und zwei E/A-Gänge mit Steuereinheit 3830 verbunden, bei Ausfall eines Pfades wurde so - mit kleinerer Leistung - der Betrieb aufrecht erhalten.
- Ein Laufwerkstrang mit max. 8 Laufwerken bildet 2,54 GB Speicherkapazität ab.

Speicher-einheiten-Modelle
- Jedes Modell hat zwei Laufwerke, diese sind im oberen Teil der Einheit sichtbar.
- dabei enthalten die Modelle **A**2, **A**2F, **C**2 und **C**2F die **Steuereinheit** integriert.
- die Modelle B2 und B2F sind nur Laufwerke, Modelle A2**F**, B2**F** und C2**F** besitzen einen **F**estkopfspeicher mit max. **1,144MB** Gesamtkapazität pro Laufwerk.

Steuereinheit 3830
- Steuert einen Laufwerksstrang mit max. 8 Laufwerken und 2,54 GB Speicherkapazität
- Besitzt Einrichtung zur Auswahl des **Betriebsmodus:** Originalmodus oder Verträglichkeitsmodus mit 3340 -11.
- Erste Stufe „Fehlertoleranz“ durch zwei Datenpfade von Kopfeinheit zu Steuereinheit.
- Befehlswiederholung. Fehlerkorrektur wird dadurch verbessert, daß Kanal-und Steuereinheitbefehle ohne Eingriff der Steuerprogramme automatisch wiederholt werden.
- Fehlerkorrektur von Daten bis zu 4 Bytes.
- Error Correction Coding (ECC).
- Strangschalter: ermöglich einen 3350-Strang umzuschalten auf 2 CU's oder RPS's.
- **R**otational **P**osition **S**ensing = RPS, ermöglicht Kanalfreigabe während Suchoperationen.
- Anschliessbar ab **S/370-** Modell 135 und aufwärts.

Programmunterstützung durch die Betriebssysteme OS/VS1, OS/VS2, DOS/VS und VM /370.

Technische Daten
- *Zugriffszeit durchschnittlich 25 ms*
- *Drehzahl 3600 Upm, Latenzzeit 10,1 ms*
- *Plattendurchmesser 14“ = 350 mm, partikular beschichtet, Speicherschichtstärke= 1,04 Tausendstel Millimeter Trägerplattendicke 1,9 mm*

- *S/L-Kopf: Ferrit Flughöhe 18 Mikrozoll = 0,46 µm, finale Kopfposition durch Servospur*
- *Schreibverfahren: mfm*
- *Detektion: delta*

- *Lineare Bitdichte 6425 bpi (Bit/Zoll) / 476 Bit/mm*
- *Flächenspeicherdichte: 3,07 MBit/Quadratzoll / 476 kBit/qcm*

Quellen

- DPCE Folder (ohne Formnummer) 1976
- Überblick über das IBM System /370, Form GA12-1040-5, Seite 7-18, 1976
- IBM R&D Journal, 25th Anniversary Issue, 1981
- INA 10-1983: „Anzahl 3350 installiert bei *Aachener-Münchener*: 126 Einheiten mit 40 Milliarden Bytes, dann erfolgte Umstellung auf 3380“

 DASD3350.tmd

IBM **DASD-Profile**

Einheit DASD-Einheit 62 PC* für den integrierten Einbau in räumlich kleinere Systeme, erstmals eigesetzt im Datenbanksystem S/38 1978 und später weiterer Produkte. 62PC wurde konzipiert für *kompakte Baugröße* und *hohe Zuverlässigkeit* im Entwicklungslabor Hursley.
*PC steht für Piccolo. Es handelt sich um eine Weiterentwicklung der Laufwerke Spartan und Delphi.

Systemfamilie S/38 und andere Kleinsysteme

Steuereinheit Bei S/38 Direktanschluss a.d. Systemeinheit 5381, ebenso bei Anderen.

Speicherkapazität Pro 62PC-Einheit = 3,4 GB. System /38 konnte max. 6 Einheiten = 20,4 GB.haben.

Datenrate 1,031 MB/Sek.

Technische Daten

Aufzeichnungsdichte
- Lineare Bitdichte.................*7535 bpi / 336 Bit pro mm...........*
- Spurendichte.......................*457 tpi / 18 Spuren pro mm...*
- Flächenspeicherdichte........*1,69 Mbit/Zoll² / 0,2619 Mbit/cm²...*

Schlüßelparameter
- Kopf/Platte-Flughöhe in.......*Kleiner 18 Mikrozoll (0,457 µm)...*
- Kopf-Luftspalt in.................*1,8 µm...*
- Speicherschicht in...............*1,04 Tausendstel mm...................*

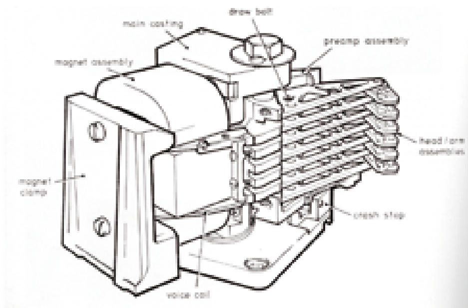
Platte-
- Durchmesser......................*8 ¼ Zoll(Datenzone ~35 mm breit*
- Technik..............................*"Braune Platte", Winchesterkopf*
- Kopfschutz bei Start/Stop... *Schmierschicht für v < 50%*
- Substrat-Dicke ....*1,9 mm* ....
- Drehzahl............................*2964*............................
- fest/auswechselbar.............*fest* ....
- Speicheroberfl./Spindel......*12*...............................

Luftträger & magnetisches Element
- Trägertyp............................*Flugkörper*................
- Oberflächenkontur..............Keine Angabe.............
- Gleitermaterial....................Keine Angabe.............
- Kernmaterial ....*Ferrit*..........................

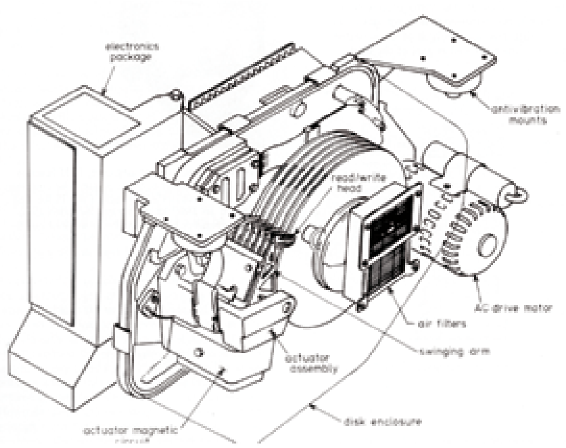
R/W-Elektronik
- Datenrate.............................*889 KB/Sek*..............
- Encoding..............................Keine Angabe............
- Detektion.............................*delta*..........................
- Zeitgebung...........................Keine Angabe............

Aktuator
- Zugriffsgeometrie.................*Rotary Aktuator*.........
- Köpfe-Anzahl.......................*12*..............................
- Positionierung......................*Tauchspule*...............
- Finale Position.....................*Servospur***...............
- Aktuator/Spindel (max. Anzahl)..*1*................................
- Durchschnittliche Suchzeit...*27 ms* ....

** Die auf 18 tpi gesteigerte Spurendichte forderte eine Spur-Ansteuerung- und Haltung mittels *hybridem* „Position Error Signal" (PES). Dieses nutzt sowohl eine Servospur als auch unveränderbare Datenspuren-Teile zur Kompensation von Schwankungen durch Temperatur, Luftfeuchtigkeit und Viberationen.



DASD **62PC**. Die Zeichnung stellt den Aktuator um 90º nach rechts gekippt dar (bezogen auf die Einbauposition, siehe auch Bild unten). Kopfträgerarme tragen auch Schreib/Lese-Verstärker zwecks kürzester Verbindungen.



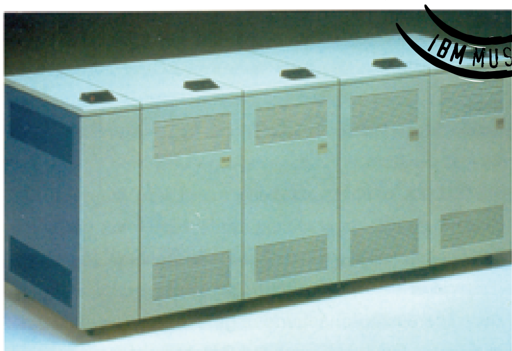
DASD **62PC.** Gesamtüberblick der Plattenspeichereinheit-Komponenten. Die hermetisch schließende Plexiglashaube ist entfernt, den Luftausgleich ermöglichen die „air filters".

Quellen:
- IBM Disk Storage Technology, Februar 1980, GA 26-1665,
- IBM Journal of R&D, 25th Anniversary Issue, 1981, page 685
- IBM Rochester: „A Half Century of Innovation", 2006 (1956 -2006)

 DASD62PC.tmd

Einheit **Plattenspeichereinheit 3370,** angekündigt 1979 für mittlere Systeme. 1983 erweitert auf die Modellreihe 2 mit größerer Spurendichte.

Anschluss Modellabhängig an S /370- 4331, 4341, 4361 und S /38 über die Kopfeinheit eines 3370-Stranges mit max. 4 Einheiten. Die Kopfeinheit (A) enthält die Steuereinheit.



Im Bild oben links die Kopfeinheit 3370 A1 mit der Speichersteuereinheit und einem Laufwerk. Rechts davon drei weitere 3380 B1 Einheiten mit je 1 LW.

Allgemeines Jede Einheit enthält ein Head-Disk-Assembly (HDA) als fest eingebautes und versiegeltes Aggregat aus Spindel und zwei Zugriffsmechanismen. Diese greifen unabhängig voneinander je zur Hälfte der Speicheroberflächen zu, sind mit zwei Schreib/Leseköpfen pro Speicherfläche versehen, so daß sich die Zugriffswege halbieren.

Modellreihe	1 (1979)	2 (1983)
• Speicherkapazität pro Laufwerk	571 MB	730 MB
pro Zugr. Mechanism.	285 MB	365 MB
pro Subsystem (Strang)	2,285 GB	2,920 GB
• Fixed Block Architektur (FBA) mit *512 Bytes/Block* bezw. Vielfachen davon.		
Anzahl der Blöcke je Zugriffsmechanismus	558.000	712.752
• Mittlere Zugriffzeit	20 ms	19 ms
• Mittlere Latenzzeit	10,1 ms	10,1 ms
• Datenübertragungsrate	1,86 MB/Sekunde	1,86 MB/Sekunde
• Rotational Position Sensing (RPS)	ja	ja

- Befehlswiederholung (automatische Fehlerkorrektur, innerhalb von Kanal-und Steuereinheit ohne Eingriff durch Programme).
- Daten-Fehlerkorrektur mittels error-corection code (ECC) trotz der hohen Datenrate durch Erzeugung der ECC-Bytes in „serial by byte“ Modus (nicht wie bisher „serial by bit“), mit Hilfe µP-gesteuerter „table look ups“ zur Herstellung der Ersatz-und Korrekturmuster. Dies war nur durch die FBA möglich.
- Automatische Fehlerstatistik durch EREP (Environmental Error Recording)

Betriebssystem-Unterstützung 3370 Modelle A01, A02 und B01, B02 werden durch **DOS/VSE**, **VM/SP** und **SSX** unterstützt, die Modelle A11, A12 und B11, B12 durch das S/38 Betriebssystem CPF.

Techn. Daten
- *Spindeldrehzahl beider Modelle* ***2964 Upm***
- *Die finale Spurposition wird durch Servospur erreicht und gehalten.*

- *Partikular beschichtete, „braune Platten“, Schichtstärke 1,04 Tausendstel Millimeter*
- *S/L-Kopf in* ***Dünnfilm-Technologie:*** *erstmals werden S/L-Köpfe monolythisch mit hoher Genauigkeitsreproduktion hergestellt und ermöglichen einen Jochluftspalt von* ***1/1000 mm.*** *Die resultierende Bitdichte beträgt erstmals 12134 bpi, nahezu Verdoppelung zu 3350.*

- *Schreibverfahren*	***2,7***
- *Detektion*	***delta clip***
- *Zeitgebung*	***VFO***
- Lineare Bitdichte	***12134 bpi*** *bzw.* ***4777 Bits pro cm***
- *Spurendichte*	***35 tpi*** *bzw.* ***25 Spuren pro mm***
- *Aufzeichnungsdichte*	***7,8 Mbit/ Zoll^2*** *bzw.* ***1,2 MBit/qcm***

Quellen
a) DPCE Folder G-7902-716
b) Marketinginformation 38/1983
c) IBM Journal od R&D, 25th Anniversary
d) 1:0 für Dünnfilmköpfe: INA 7-1981
e) IBM Disk Storage Technology, Form GA 26-1665

 DASD3370.tmd

IBM DASD-Profile

Einheit Die Plattenspeichereinheit 3310 wurde 1979 für das System 4331 angekündigt. 3310 war unter dem Arbeitsnamen „Swallow" die letzte DASD-Entwicklung des IBM Entwicklungslabors Hursley UK. Der Vorläufer war das Laufwerk „Piccolo" das u.a. im System IBM /32 als integrierte („embedded) Disk File" zum Einsatz kam. 3310 hat Truhenform.

Modelle

A1 Einfach-Plattenlaufwerk als Haupteinheit
A2 Doppel-Plattenlaufwerk als Haupteinheit
B1 wie A1, jedoch Sklave von A1
B2 wie A2, jedoch Sklave von A2

- Jedes Laufwerk hat eine Speicherkapazität von 64,5 MB.
- Alle Modelle wurden via *Direktanschluß* (FTA= file-tape-adapter) an mittlere oder kleine EDV-Syteme angeschlossen. Ein LW-Strang konnte mit max. 4 Laufwerken die Speicherkapazität von 258 MB abbilden (siehe Photo unten).
- Aufzeichnungsmodus: Festblock, Größe 512 Bytes, 32 Blöcke pro Speicherspur.

Technische Daten:

Platte-
- Durchmesser.............................*8¼ Zoll*....................
- Technik....................................*"Braune Platte"*.........
- Substrat-Dicke..........................*0,075 Zoll (1,9 mm)*
- Drehzahl...................................*3125 Upm*.................
- fest/auswechselbar....................*fest*..........................
- Speicheroberflächen/Spindel.....*11*.............................

Schlüßelparameter
- Kopf/Platte-Flughöhe in µZoll....*13*.............................
- Kopf-Luftspalt in µZoll................*40*.............................
- Speicherschicht in µZoll ....*25*.............................

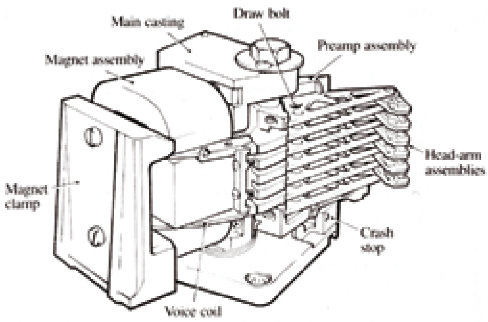
Aufzeichnungsdichte
- Flächenspeicherdichte . ...*3,87Mbit/Zoll² (0,6 Mbit/cm²)*
- Spurendichte (tpi) ....*450/Zoll (17,72/mm)*
- Lineare Bitdichte (bpi) ....*8530 Bit/Zoll (336 Bit/mm)*

Flugkörper & magnetisches Element
- Trägertyp..................................*Flugkörper*...............
- Oberflächenkontur.....................*flache Keiform*..........
- Gleitermaterial...........................*Ferrit* ....
- Kernmaterial..............................*Ferrit*........................

R/W-Elektronik
- Datenrate...................................*1,031 MB/Sek*..........
- Encoding....................................*mfm*.........................
- Detektion....................................*delta*........................
- Zeitgebung................................*vfo*..........................

Aktuator
- Zugriffsgeometrie.......................*Rotary movement* (RVCA)
- Köpfe-Anzahl.............................*Ein S/L-Kopf/Speicherfläche*
- Positionierung............................*Elektromagnetischer Aktuator*
- Finale Position...........................*Servofläche+ Sektorinformation*
- Aktuator.....................................*Einer pro Spindel*
- Durchschnittliche Suchzeit.........*27 ms*
- Umdrehungszeit.........................*19,2 ms*
- Latenzzeit ....*9,6 ms*
- Automatic Rotational Position Sensing (*RPS*)

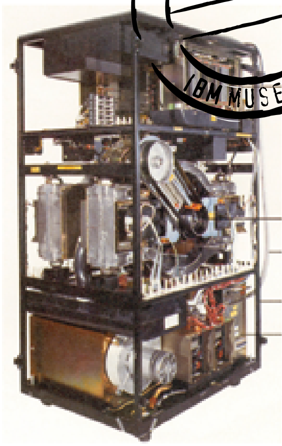
Zwei Einheiten 3310, eine davon als Haupteinheit, die zweite als „Sklave".



Actuator des Laufwerks „Picollo", dargestellt 90º nach rechts gekippt.

Quellen:
- IBM Deutschland Vertrieb, Produkt-Information Form GT 12-2580
- DP CE Informationfolder, Form G-7902-714
- Journal of R&D, 25th Anniv. 1981, Page 678;

DASD3310.tmd



Speichereinheit 3380 mit zwei HDA's (Mitte) und deren Versorgungseinheiten (darunter und darüber).

Einheit **Plattenspeichereinheit 3380,** angekündigt 1981 für S/370 Großsysteme als Modell D, als Modell E in 1985 und als Modell K in 1987; jeweils als DASD's großer Speicherkapazität und Datenrate sowie kurzer Zugriffszeit.

Subsystem-konzeption

- Ein Subsystem besteht aus Speichersteuereinheit (SPSTE) 3880 mit 2 *Direktoren* und je daran angeschlossenen Speichereinheiten 3380 mit eingebauten Steuereinheiten. Die neue Steuereinheit kann ab 1985 optional mit 8 oder 16 MB großen Pufferspeichern ausgerüstet werden, um das Ein/Auslagern von Seiten des VS-paging / swapping - zu beschleunigen.
- Das schon bekannte **CKD-Format** wird fortgesetzt.
- Jede 3380 enthält zwei fest eingebaute Laufwerke, genannt HDA (head disk assembly) mit je 15 Speicherflächen. Jedes Laufwerk hat zwei unabhängige Zugriffsmechanismen, die je zu einer Hälfte der Speicheroberflächen zugreifen.Die adressierte Spur aktiviert einen der S/L-Doppelköpfe, dadurch wird der Zugriffsweg halbiert.
- 2 Steuereinheiten in 3380 Modellen **AA4, AD4 und AE4** ermöglichen „*Dynamische Pfad-Selektion*“, d.h. jedes Laufwerk kann dynamisch über 2 Pfade erreicht werden.
- Flexibilität und Zugriffssicherheit -auch an 2 Prozessoren- ermöglicht der *Kanalschalter*, der jeden Direktor einem von 2 oder 4 oder 8 Kanälen *selektiv* zuschalten kann.

Modellreihe 3380	**D** (1981)	**E** (1985)	**K** (1987)
- Sp'Kapazität/Einheit::	630 MB	1260 MB	1890 MB
- Mittlere Zugriffszeit :	16 ms	17 ms	16 ms
- Mittlere Latenzzeit:	8,3 ms	8,7 ms	8,3 ms
- Strangkapazität:	10,8 GB	20,16 GB	30,24 GB

- Datenrate bei allen Modellen...................................... 3 MB/s ....

Technische Daten

Plattentechnik:

- Spindeldrehzahl aller Modelle .... *3600 Upm*
- Tragerschicht-Dicke für alle Modelle beträgt............ *3,8 mm*
- Partikulare Plattenbeschichtungsstärke.....*bei allen Modellen kleiner als 1/3000 mm*

S/L-Kopftechnik

- alle Modelle verwenden................................ *Dünnfilm-Kopftechnik*
- S/L-Kopf-Flughöhe für alle Modelle...................... *< 1/3000 mm*
- Lin. Bitdichte bei allen Modellen..............*15200 bpi*, bzw. *5984 Bits pro cm*

	D	E	K
- Spurendichte	*801 tpi* bzw. *31,5 /mm*	*1260 tpi* bzw. *49,6 /mm*	*2139 tpi* bzw. *64,2 /mm*
- Flächenspeicher-dichte	*12 MBit / Zoll²* *1,86 Mbit/qcm*	*19 MBit / Zoll²* *2,99 Mbit/qcm*	*38 MBit / Zoll²* *5,47 Mbit/qcm*

Elektronik

- Schreibverfahren für alle Modelle................................. 2,7
- Detektion bei allen Modellen.......................................*delta clip*
- Zeitgebung bei allen Modellen...................................... *VFO*

Softwareunterstützung: MVS/XA-Datenmanagement E/A., MVS/ 370, VM/System Product, VM/HYPO, VM/XA, VSE/SP, VSE/AF und TPF2.

Quellen:

- IBM 3380 DASD General Information, Storage Subsystem Library, Form GC26-4193, 1981
- IBM Journal of R&D, 25th Anniversary Issue 1981
- IBM 3380 DASD, Reference Summary, Form GX26-1674, 1984/1986/1987
- IBM Report 4-1986: „IBM Werk Mainz entwickelt Flughöhenmesser für S/L-Köpfe“.
- IBM Mainz, Wattrodt-Präsentation 1-1986
- IBM Deutschland Geschäftsbericht 1986
- IBM Report 1-1988: 100 000 Plattenlaufwerk der Familie 338X ausgeliefert
- G. Sandner Papier „Daten über IBM DASD's“, 5-2005

 DASD3380a.tmd

IBM **DASD-Profile**

Einheit Magnetplatteneinschub 9332 zum Gehäuse IBM 9309 angekündigt 1987 als kleine, wirtschaftliche Einheit im unteren und mittleren Systembereich. Zuverlässigkeit, Ausbaufähigkeit und Festblockarchitektur sind weitere Merkmale des Magnetplatteneinheit-Konzeptes.

Systemfamilien: 9370, 6150, S/38, S/36

Modelle	Modell 220	Modell 40x	Modell 60x
Kapazität	200 MB	400 MB	600 MB
	Freistehendes Modell mit einem Zugriffsmechanismus	Beide Modelle haben 2 Zugriffsmechanismen, jeder greift unabhängig vom andern die halbe Speicherkapazität zu. Beide Modelle haben 2,5 MB/s Datenübertragungsrate	

Anschluß Direktanschluß, 9332 enthält die Steuereinheit und das/die Laufwerk(e). Außerdem können mehrere 9332 an einen Datenpfad angeschlossen werden. Die Modelle 400 und 600 werden an ein System angeschlossen. Die Modelle 402 und 602 können über Umschalter an ein zweites System S/370 angeschlossen werden.

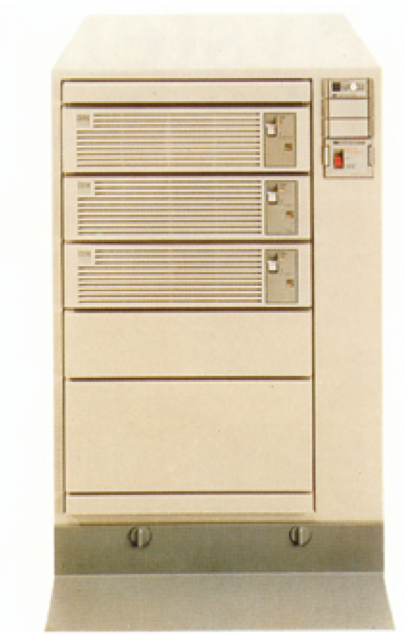
Aufzeichnungsmodus 9332 zeichnet in Festblockarchitektur auf, d.h. die Spuren werden in feste Blöcke mit je 512 Byte eingeteilt. Dies hat den Vorteil, daß bei entsprechender Satzorganisation die nominale Laufwerkskapazität weitgehend zur Datenspeicherung selbst benutzt werden kann. Speicherplatz für die Daten-Identifizierung wird nicht benötigt.

Softw. Unterstützung
- Für Prozessoren 9373, 9375, 9377: VSE 2.1.6; VM/IS ab Release 4, VM/SP ab Rel. 1; IX/370, DPPX/370
- Prozessor 9406: OS/400

Technische Daten

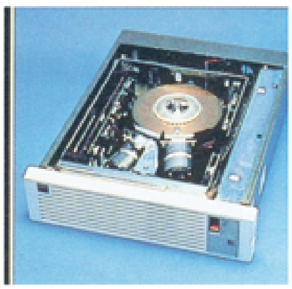
Speicherplatte...................*Festplatte, 8¼ Zoll Durchmesser*
- Technik............................*“Braune Platte“*
- Drehzahl...........................keine Angabe
- Latenzzeit.........................*9,6 ms*

Aktuator
- Zugriffsgeometrie.............*Linear*...................................
- Köpfe-Anzahl....................keine Angabe.........................
- Positionierung...................*Voice Coil Motor*...................
- Aktuator............................*zwei*......................................
- Durchschn. Suchzeit........*min. 3,2 ms, max. 19,5 ms*



Drei Magnetplatteneinschübe IBM 9332 Modell 400 oder 600 im Gehäuse IBM 9309. Breite 48,3 cm. Höhe 80 cm, Tiefe 56,1 cm.

Einschub (Drawer)
- der Einheit 9332 im Bild rechts in geöffnetem Zustand.
- Vor dem 8¼ Zoll-Plattenstapel die beiden Linear-Zugriffsstationen für je eine Hälfte der Kapazität (bei den Modellen 40x und 60x.

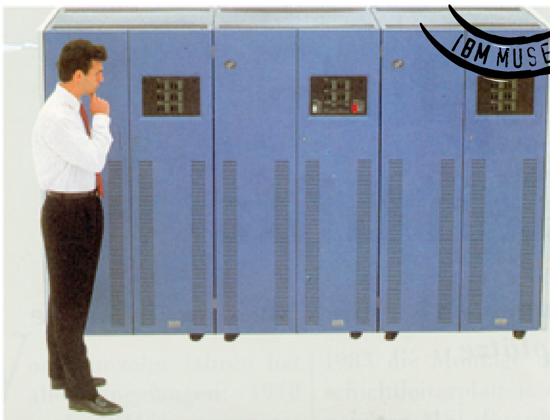


Quellen:
- IBM D, Marketing ’87, 23.06.1987, Nr. 19
- IBM D, Produkt Hardware, Form GT12-3434-2 (1990)
- IBM Rochester, Half a Century of Innovations, 2006

DASD9332.tmd

Einheit

Plattenspeichereinheit 3390, angekündigt 1989 für S/390 Großsysteme als DASD großer Speicherkapazität, hoher Datenrate, kurzer Zugriffszeit und neuer Vorkehrungen gegen Ausfälle.



3390-3 Mod. A38 und 2 x B3C, 90 GB (ohne Steuereinheit).

Subsystem-konzeption

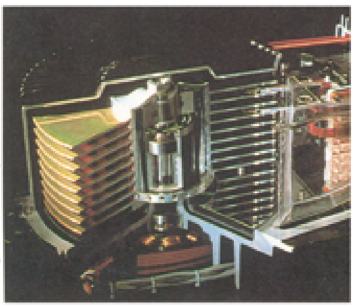
- Ein Subsystem besteht aus Speicher-Steuereinheit **3990** mit einem Laufwerksstrang in Haupt (A)- und Nebeneinheiten (B). Modell **1** an 2-8 Kanälen mit *2 Pfaden*, Modell **2** an 4-16 Kanälen und *4 Pfaden*; d.h. *2-4 Datentransfers können simultan stattfinden.*
- Fest eingebaute **Head Disk Assembleys** (HDA's) mit je **zwei Zugriffsmechanismen** mit Schreib/Leseköpfen.
- Jeder Zugriffsmechanismus, d.h. die „halbe HDA-Speicherkapazität" ist adressierbar und wird ***Volume*** genannt.
- Alle Modelle arbeiten mit der Datenübertragungsrate **4,2 MB/Sekunde**
- Ab 1991 konnte die Steuereinheit zur Durchsatzsteigerung optional mit dynamischem Cache von **16- 64 MB** + **4MB Non-volatilem** Pufferspeicher ausgerüstet werden.
- Neue automatische Diagnoseverfahren zur Wartung im laufenden Betrieb.

Modellreihe	3390 (1989)	3390-3 (1991)	3390-9* (1992)
HDA-Kapazität	**1,89 GB**	**5,67 GB**	**17 GB**
Kapazität/ LW (Aktuator)	**0,946 GB**	**2,84 GB**	**8,5 GB**
Kapazität/Einheit	**7,56 GB**	**22,68 GB**	**68 GB**
Kapaz./ Strang	**30,26 GB**	**90,8 GB**	**272,5 GB**
Mittlere Zugriffszeit	**je nach Modellgruppe zwischen 9,5 und 12,5 ms**		**16 ms**
Mittlere Latenzzeit	**7,1 ms**	**7,1 ms**	**8 ms**

*Hochkapazitätsmodell, für Anwendungen mit hohen Kapazitäts- jedoch geringeren Zugriffs-Anforderungen

Technische Daten

	3390	3390-3	3390-9
Plattentechnik:			
- Plattendurchmesser		*- 10,8 Zoll / 274 mm alle Modelle -*	
- Spindeldrehzahl,		*- 4260 Upm alle Modelle -*	
- Trägerplatte	***Aluminium 3 mm***	***Al-Mg, 1,9 mm***	*wie Modell 3*
- Plattenbeschichtung	*braune Platte*	***Dünnfilm-Vakuumbeschichtung aus CoCrPt** mit **0,05 µm** Stärke*	*wie Modell 3*
S/L-Technik			
- Schreib/Lese-Element	*induktiv/induktiv*	***induktiv/magneto-resistiv***	*wie Modell 3*
- Flughöhe		*- < 1/3000 mm alle Modelle -*	
- Lineare Bitdichte	***27.483 bpi***	***29.718 bpi***	
- Spurendichte	***2235/Zoll** - 88/mm*	***2972/Zoll** - 117/mm*	
- Flächenspeicherdichte	***60,5 MBit/Zoll²*** *9,38 MBit/cm²*	***90 MBit/Zoll²*** *13,5 MBit/cm²*	



Ein 3390-HDA-Schnittmodell, links der Plattensatz, ganz rechts übereinander die beiden Aktuatoren mit den Schreib/Lesekopfen.

Software-Unterstützung MVS/370, MVS/XA, MVS/ESA-**Data Mgmt.-**
I/O -**Data FacilityStorage Management Subsystem**
VM/XA SP, VM/SP HPO, TPF

Quellen

- Storage Subsystem Library, „Using IBM 3390 DASD in a MVS Environment"
- DPCE Folder without form number
- IBM 3390 DASD General Information, Storage Subsystem Library, Form
- IBM System Guideline Overview
- IBM Report 8/'89, Seite 33

 DASD3390.tmd

IBM DASD-Profile

Einheit	**DASD-Laufwerk 0681**, 1990
Anwendungsfeld	**IBM und OEM-Produkte,** im SCSI-Einschubeinheit für IBM RS/6000 und für OEM-Produkte.
Speicherkapazität **Datentransferrate**	0,857 GB pro Laufwerk. Im SCSI-Einschub max. 4 x 0,857 = 3,4 GB 3 MB/s
Modularität	Für das System RS/6000 Modell 930 (Rackaufbau) wurde ein SCSI Einschub (Drawer) für 4 Laufwerke 0681 geschaffen. Einschub-Bauhöhe 6,7 Zoll. Siehe auch RS/6000-M930 Abbildung unten.

Technische Daten

Aufzeichnungsdichte
- Lineare Bitdichte........................Keine Angabe
- Spurendichte..............................Keine Angabe
- Flächenspeicherdichte...............*63 MBit/Zoll²*

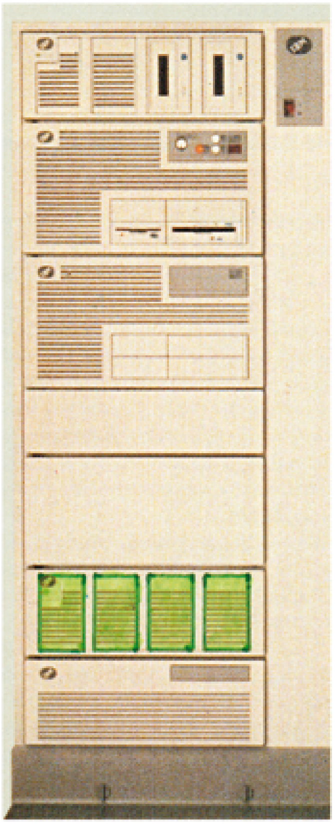
Platte
- Durchmesser..............................*5¼ Zoll*
- Speicheroberflächen..................*14 x 2*
- Technik.......................................*MIG Dünnfilmplatte*
- Substrat-Dicke............................Keine Angabe
- Drehzahl.....................................*7000 Upm*.
- Durchschn. Latenzzeit................*6 ms*

Flugkörper & magn. Element
- Trägertyp....................................*Flugkörper mit MR-Kopf*
- Oberflächenkontur......................*Flache Keilform*
- Kernmaterial...............................*Dünnfilm*

R/W-Elektronik
- Schreiben/Lesen.........................*Induktiv / Magneto-resistiv*
- Elektronik....................................*VLSI-Technologie*
- Detektion.....................................*Partial Response Maximum Likelihood*
....(*PRML erstmals in Produktanwendung* mit ex-
....ternem Analogfilter und adaptivem
....Cosinus Equalizer)
- Rohdatenrate...............................*24 Mbit / s*

Aktuator-
- Zugriffsgeometrie.......................*Linear*
- Zugriffskamm-Führung...............*Zwei Keramik-Präzisionsrundstäbe mit 2 x 4 Mini-Präzisionskugellagern*
- Daten-Schreib/Leseköpfe...........*27*
- Positionierung............................*Tauchspulenmotor*
- Finale Position...........................*Servospeicherfläche*
- Mittlere Zugriffzeit......................*17,2 ms*

Frontalansicht des damaligen Spitzenmodells RS/6000-M 930 bestückt mit einem Drawer mit 4 Laufwerken 0681 (unten grün markiert).

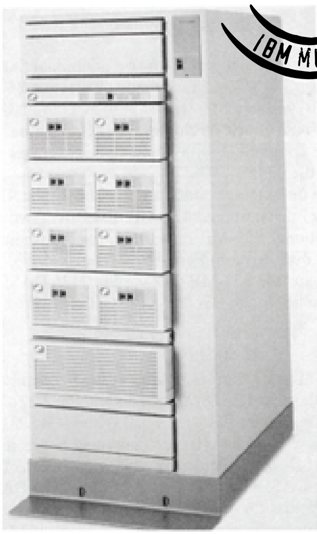


Quellen
- RISC System /6000 Technologie, Form SA23-2619, 1990
- IBM Hursley Park, Special Edition of „Developments", 2-1990
- IBM D, Marketing, „Neu von IBM", 2-1990
- IBM Report, Neue Produkte, 2-1990

Hans Spengler Böblingen, 5-2006, 9-2009 DASD0681.tmd

Type DASD Subsystem 9340, angekündigt 1-1991 für mittlere und kleine Systeme in ES/9000 Architektur in Gestell-Bauweise (Rack).

9340 leitet mit seinem DASD-Einschub IBM 9345 den Übergang auf 5¼-Zoll-Laufwerke in dieser DASD-Kategorie ein. 9345 hat zwei Laufwerke mit je 1 GB, d.h. 2 GB insgesamt bei Modell 1 ***oder*** 3 GB bei Modell 2. Die Datenübertragungsrate beträgt 4,4 MBit/s. Einschübe können jetzt im Betrieb entfernt / hinzugefügt werden!



Subsystem 8 Einschüben 9345 und darunter die Steuereinheit Modell 9341, alles im Rahmen 9309.
Maße/Gewicht:
breit: 65 cm, tief: 94 cm,
hoch: 158 cm, Gewicht 800 kg

Subsystem-konzeption Für Anforderungen von 2 bis 24 GB Speicherkapazität wird 9345 bis zu 8 mal in den Standardrahmen 9309 eingefügt, darunter befindet sich die Steuereinheit 9341. System anschluß an 2 BMPX-Kanäle von ES/9000-9221. **Zwei Datenpfade.**

Für Anforderungen von 4 bis 48 GB ist in einem weitren Standardrahmen die Steuereinheit 9343 erforderlich, deren Systemanschluß an bis zu 4 parallele Kanäle erfolgen kann, dabei ist ESCON eine Option. Bis zu **vier Datenpfade** sind möglich (ES/9000, 3090, 43xx, 937x). 9340 DASD's nutzen die CKD-Datenarchitektur für die Informationsspeicherung und ab 5-1992 a) die ECKD-Archtitektur und b) weitere Leistungssteigerung modellabhängig und optional mittels 32-64 MB Cache; c) Bei ES/9000-ESCON XDF und DFSMS/MVS kann die Entfernung zwischen Prozessor und Subsystem jetzt bis zu 43 km betragen.

Betriebsssystem-Unterstützung VSE/ESA und VM/ESA,

Technische Daten Einschub 9345, ***Maße:*** *hoch: 17,2 cm, breit: 21,2 cm, tief. 76,8 cm,* ***Gewicht:*** *27 kg*

Plattentechnik
- Durchmesser......................*5¼ Zoll*= 130 mm........................
- Speicherschicht ....***Vakuum-Dünnfilmbeschichtung***
- Drehzahl............................*7000 UpM*......................................
- Speicherflächen/Spindel.....*18*...
- Mittlere Latenzzeit..............*5,6 ms*..

Aufzeichnungsdichte-
- Lineare Bitdichte ....*44500 bpi* = 1752 Bit/mm...........
- Spurendichte ....*2413 tpi* = 95/mm........................
- Flächenspeicherdichte ...... *111 MBit/Zoll²* = 16 MBit/cm².....

Flugkörper & magnetische Elemente
- Trägertyp............................*Flugkörper*..................................
- Oberflächenkontur...............*Flache Keilform*.........................
- Schreib/Lese-Element.........***Induktiv schreiben / Magneto-resistiv Lesen***
- Schreib-Luftspalt.................keine Angaben
- Joch-Breite..........................keine Angaben

Aktuator
- Zugriffsgeometrie................*Linearbewegung*........................
- Köpfe-Anzahl......................*18* (*16 für Datenspeicherung*).......
- Positionierung.....................***Tauchspulenmotor***.....................
- Finale Position....................***Servospeicherfläche*** ....
- Aktuator/Spindel.................*1*...
- MittlereZugriffszeit.............*11 ms*...

Quellen

IBM D Marketing INFO 9-1991 und 5-1992;
Flyer „Speicherperipherie GF12-1670-6, 6-1994 sowie Unterlagen aus der Sammlung des Verfassers

 Prof9340.tmd

IBM **DASD-Profile**

Einheit — **DASD-Laufwerk** 0662, 1993

Anwendungsfeld — IBM **und** OEM-**Prosdukte,**
Einschub-Einheit, 0662 setzte neue Maßstäbe für größtmögliche Datenmenge auf kleinstem Raum in den folgenden Anwendungsfeldern: Technische/Kommerzielle Workstations, Netz-werk-Servern, Massenspeicher-Arrays, PC's im oberen Leistungsbereich.

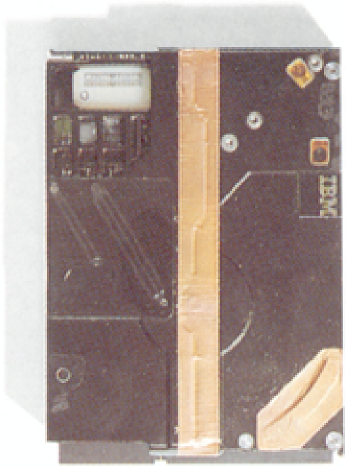
Speicherkapazität — 1,05 GB pro 0662-Laufwerk, formatierte Kapazität mit 512 Bytes pro Sektor, 512 KB Datenpuffer, mehrfachsegmentiert

Modelle
- S12, fast 10 MB/s, SCSI-2
- S1D, fast 10 MB/s, SCSI-2
- SW1, fast and wide 20 MB/s SCSI-2
- SWD, fast and wide 20 MB/s SCSI-2

Alle Modelle mit den Abmessungen 25,4 mm hoch, 101,6 mm breit und 146 mm tief; Gewicht 0,46 kg. (siehe auch Abbildung unten)

Zuverlässigkeit — MeanTime Between Failure = 800 Kilostunden (Industry Leading Reliability), ECC on the fly; Predictive failure analysis,
Produktion: *IBM Werk Mainz und IBM ADSTAR San Josè California*

Technische Daten



Das Photo zeigt die Einschub-Einheit IBM 0662. Ganz unten die elektrische Steckverbindung. Darüber das herme-tisch abgeschlossene Laufwerk.

Aufzeichnungsdichte
- Lineare Bitdichte.................*86360 Bit/Zoll² = 3400 Bit/mm²*
- Spurendichte.......................*4064/Zoll = 160/mm*
- Flächenspeicherdichte........*354 MBit/Zoll² = 0,5487 MBit/mm²* .

Platte- -
- Durchmesser.......................*3,5 Zoll = 95 mm* ....
- Technik..............................*Dünnfilm*..............................
- Substrat-Dicke.....................Keine Angabe........................
- Drehzahl.............................*5400 UpM*............................
- fest/auswechselbar.............*fest*..
- Speicheroberfl./Spindel........*6*...
- Durchschn. Latenzzeit..........*5,56 ms*..................................

Flugkörper & magnetisches Element
- Trägertyp.............................*Flugkörper mit MR-Kopf*.......
- Oberflächenkontur...............*Flache Keilfom*......................
- Gleitermaterial.....................keine Angabe...........................
- Kernmaterial ....*Dünnfilm*.................................
- Flugkörper/Kern-Verbindg...*deposited*...............................

R/W-Elektronik
- Datenrate.............................*Modellabh. 40/48 Mbit/s*, *Read ahead caching, Write cache supported*
- Datenpfadbreite...................*Modellabh. 1* oder *2 Byte* ....
- Detektion.............................*PRDF-Kanal* (partial response maximum likelihood with digital filter)
- Zeitgebung......................... *Quarzoszillatoren*..................

Aktuator
- Zugriffsgeometrie................*Drehbewegung* ....
- Daten-S/L-Köpfe.................*5*..
- Positionierung.....................*Rotary Voice Coil Motor*........
- Finale Position....................*Servospeicherplatte*.............
- Aktuator/Spindel.................*1*..
- Mittlere Zugriffszeit.............*9,0 ms*...................................

Quellen

- OEM Storage Products from ADSTAR, Folder Form G511-1787-01, 1996
- K. Wattrodt-Referat IBM Werk Mainz 1996
- Flyer „Speicherperipherie GF12-1670-6, 6-1994
- IBM Report 3-1993: „*Der Speicherstar* „

 DASD0662.tmd

Einheit RAMAC Array 9394, angekündigt 1994. RAMAC steht jetzt für „Raid Architecture with Multilevel Adadaptive Cache“, eine eine völlig neue DASD-Speichersystem-Konzeption, die hohe Leistung bei ständiger Datenverfügbarkeit ermöglicht.

Die innovative Architektur mit Parallelverarbeitung, mehrstufiger Datenpufferung und Unterstützung verschiedener Datenformate setzen einen neuen Meilenstein: *9394 ist gleichzeitig die Steuereinheit des flexibelsten aller bisherigen IBM Speichersysteme.*

Parallel arbeitende DASD-Einschübe ermöglichen bis zu 64 parallele Schreib/Lese-zugriffe. Dies in Verbindung mit mehrstufiger Datenpufferung (Cache-Kapazität von 64 MB bis 2,04 GB) resultiert in hoher Leistung bei höchster Datenverfügbarkeit.

Das RAMAC-Speichersystem ist fehlertolerant durch redundante Auslegung aller kritischen Komponenten. Selbst bei Ausfall eines Netzteils, eines Gebläses, eines Datenpfades oder DASD-Laufwerkes wird für den Anwender weder die Datenverfügbarkeit vermindert, noch tritt Datenverlust ein.

Speicher-bausteine

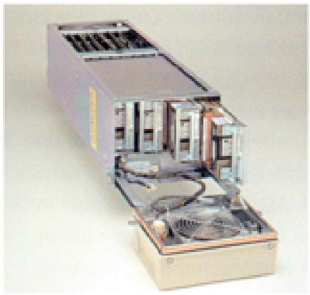
1. Der Grundbaustein ist die Einschubeinheit 9395 als unabhängige, fehlertolerante Einheit in der RAID 5 Array. Die Einschubeinheit wurde ab 1994 mit Laufwerken vom Typ ***Allicat*** mit *2 GB* Speicherkapazität bestückt, ab 1995 konnte es der Typ ***Ultrastar XP*** mit *4 GB* sein und ab 1996 der Typ ***Ultrastar 2XP*** mit *9 GB* Speicherkapazität.

2. Jeweils 4 Laufwerke bilden einen Einschub (Drawer) Typ 9395 (Bild rechts).

3. Ein Gestell (Frame) kann max. 16 Einschübe fassen. Daraus ergibt sich pro Gestell die max. Speicherkapazität von *128 GB / 256 GB / 288 GB.*

2-faches Netzteil

13 Mikroprozessoren, 32 Bit, RISC, CMOS-IV Technik



Einschub-Cache **32 MB.**
Nicht flüchtiger Laufwerk-Cache **512 KB**, ab 2XP **1,024 MB**

Das 3,5 Zoll Laufwerk ganz rechts ist halb herausgezogen aus der Normal-position.

2-fache Kühlung

Batterie

2-fache SCSI-Pfade

4. Ein oder mehrere Gestelle bilden ein Gesamt-DASD-Speichersystem.

Kanalanschluss: Parallel / ESCON

BS-Unterstützung: MVS/ESA/SP/XA, VM/ESA/SP/XA/XA HPO, VSE/ESA/SP, AIX/ESA/370, TPF.

Technische Daten	**Alicat** (1994)	**Ultrastar** (1995)	**Ultrastar 2XP** (1996)
Speicherkapazität	*2 GB*	*4 GB*	*9 GB*
Datenrate im Mittel	*5,2 MB/s*	*12,6 MB/s*	*15,4 MB/s*
Zugriffzeit im Mittel	*9,2 ms*	*8 ms*	*8,5 ms*
Speicherplatte			
- Durchmesser	*3,5 Zoll, alle Modelle*		
- Beschichtung	*Dünnfilm, alle Modelle*		
- Flächen/Spindel	*9*	*alle Modelle*	
- Drehzahl	*5400 UpM*	*7200 Upm*	*7200 Upm*
- Latenzzeit	*4,7 ms*	*4,2 ms*	*4,2 ms*
Aufzeichnungsdichte			
- Lineare Bitdichte	*81913 bpi*	keine Angaben	
- Spurendichte	*3168 tpi*	keine Angaben	
	125 tpmm	keine Angaben	
- Flächenspeicherdichte	*260 Mbit/Zoll²*	*544 Mbit/Zoll²*	*839 Mbit/Zoll²*
	40,4 Mbit/cm²	***84,3*** Mbit/cm²	***129*** Mbit/cm²
- Aufzeichnungsart	*RAID 5, alle Modelle*		
Magnetisches Element			
Flugkörper	***Induktives Schreiben, MR-Lese-Element***		
R/W-Elektronik			
Detektion	***Partial Response Maximum Likelihod***		
	PRML 2nd Generation...	***3rd Generation***	

IBM 9394 RAMAC Subsystem mit integrierter Steuereinheit. Breit 75 cm, tief 98 cm, hoch 158 cm, Gewicht 754 kg

Quellen

- IBM Produktinformation RAMAC Array System, Form GT2-5147, 1994
- IBM DIrekt Katalog S/390, 1995/96
- K. Gerecke: Technologie-Folien „The RAMAC Disk Evolution“, 1994

IBM **DASD-Profile**

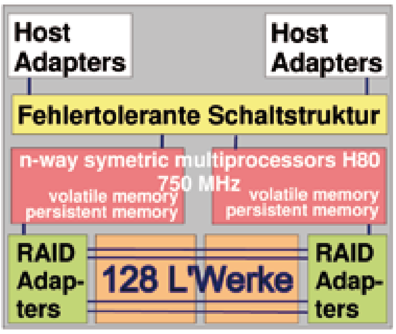
Einheit Enterprise Storage Server 2105 Modell E, 1999*, Multiplattform-fähig für die Funktionen
- Konsolidierte Unternehmensdaten auf DASD's abzubilden und unternehmensweit/platt-form-unabhängig für Geschäftsprozesse bereitzustellen/managen.
- Im Umfeld S/390, RS/6000, AS/400 und PC's, d.h. deren Betriebssysteme sowie UNIX-LINUX und WINDOWS NT, Daten abzugeben / entgegenzunehmen.
- Speicherung bis 11 TB* in RAID-Modus und Verfügbarkeit / Sicherheit mit konfigurierbaren Cache-Kapazitäten bis 6 GB (später* bis 64 MB. Laufwerkzugriff in 7,5 ms auf SAA-Laufwerke 3,5 Zoll mit 9 GB und später* /18 /36 GB pro Laufwerk. Kurze Latenzzeiten durch Plattendrehzahl 10 000 / 15 000 Upm entsprechend Laufwerkmodell.
- Noch effizienteres & komfortableres Speichermanagement durch die neuen Werkzeuge StorWatch, ESS Specialist und Expert; graphische Benutzeroberfläche, fernsteuerbar via Internet.
- Systemkommunikationsmodus „Jeder mit jedem". Leistungsbereich 18-100 MB/s (Ficon) für Entfernungen im Nahbereich bis 10 km und bis zu 100 km im **S**torage **A**rea **N**etwork.

* ESS Modell F (2001) mit 35 TB; ESS Modell 800 (2003) mit 56 TB.
Die 2105-Dimensionen: 138 cm breit, 91 cm tief, 192 cm hoch. Gewicht 1.174 kg.

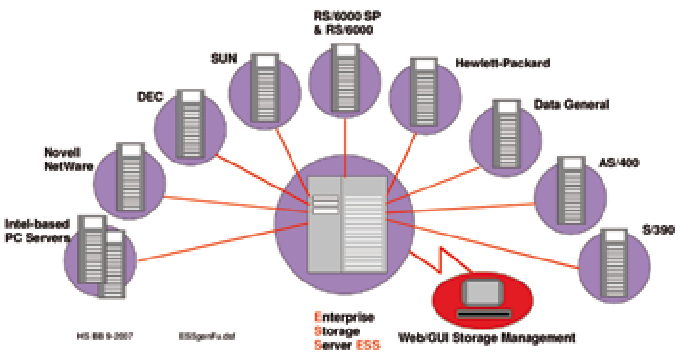
System-struktur Das Bild rechts zeigt die Struktur der 2105 mit den Hauptkomponenten. Die Datenpfade der 128 Laufwerke werden über zwei RAID-Adaptergruppen von einem n*-Wege RS/6000 Prozessor gesteuert, ebenso die fehlertolerante Schaltstruktur zu den beiden Hostadaptergruppen, *n= Modellabhängig, das Topmodell H80 hat 2 x 6 Wege bei 750 Mhz.

Intere Systemkommunikation erfolgt über ein PCI Switching Bus System zum Cluster Processor Complex (2 x 6), Cluster Cache und Non Volatile Storage zu den Device-Adaptern.

Externe Systemkommunkation mit anderen Computersystemen über ESCON, FICON, FibreChannel und SCSI (siehe Übersicht rechts Mitte).



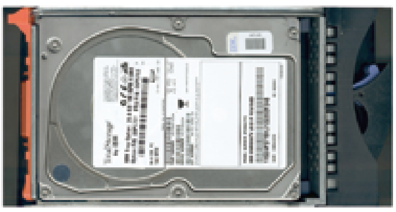
HS BB, 9-2009 Shark1.dsf



Technologie der 2105-Laufwerke
Herausragendes Merkmal des *Einschub-Laufwerkes* ist die Speicherkapazität von *9 GB* (spätere Kapazität wie oben genannt). Es nutzt erstmals die *Giant Magneto Resistive* (GMR) Lesemethode. So konnte die Flächenspeicherdichte von *427,5 MBit/cm²* erzielt werden bei LW-Dimension wie im Bild rechts unten genannt.

Der erst 1988 endeckte GMR-Effekt (bei *Tieftemperatur* und *komplexer* Schichtung) wurde von IBM Forscher *S. Perkin* in 10 jähriger Arbeit seines Teams technisch realisierbar entwickelt, um bei ***Normaltemperatur* und *wenigen Schichten*** durch *Sputtering*** im Nanometerbereich einen hochempfindlichen Lesesensor zu erhalten.

** Sputtering: Eine Material-Auftragstechnik für Nanometer-Partikel, die Kathodenzerstäubung und Beschichtung im Vakuum nutzt, um reproduzierbare Ergebnisse bei hoher Genauigkeit sicherzustellen.



Ein Laufwerk von 2003 mit 36 GB Speicherkapazität. Links elektrische Steckverbindungen, rechts der *snap-in* Mechanismus für einfache Auswechselung durch den Operator. 21,3cm breit, 12,2 hoch, 2,8 cm tief

Quellen
- IBM Journal of Research & Developm., V42, Nb 1, January 1998
- IBM zBulletin 4-2001
- The Royal Swedish Academie of Sciences: „The Nobel Prize in Physics 2007"

 DASD2105.tmd

Interview by the author Kurt Gerecke

with RAMAC experts Heinz Graichen and Hans W. Spengler

Kurt Gerecke:

Why was the invention of the RAMAC 305/350 product so important?

Hans Spengler:

The first product RAMAC 305 with the RAMAC 350 magnetic disk storage was followed without interruption for decades by further developments of disk storage with the aim of achieving increased capacity, shorter access times, higher data rates and lower costs per megabyte stored. The RAMAC 350 started this development. The American Society of Mechanical Engineers recognized this development in 1984 by declaring RAMAC disk storage to be the “International Historic Mechanical Engineering Landmark *” in recognition of its ongoing influence on disk storage technology (* milestone).

Kurt Gerecke:

How did the access mechanism of RAMAC 350 (1956) and 355 (1957) disks work? A single pair of read / write heads served all 100 memory areas. How do you imagine the positioning of the read / write head pair?

Heinz Graichen:

Positioning means covering a path. The 100 storage surfaces are applied to 50 carrier plates and their stack height is approximately 53 cm. 100 storage tracks are defined on each storage surface. The outermost track (00) is about 12.7 cm from the innermost (99) (maximum travel = 78.4 cm). The constructors solved the "positioning" of the S / L head pair by means of a trolley, which first brings an access arm vertically to the desired plate and locks it firmly in this position. Then, in a further phase, the access arm is moved horizontally to the desired track and even

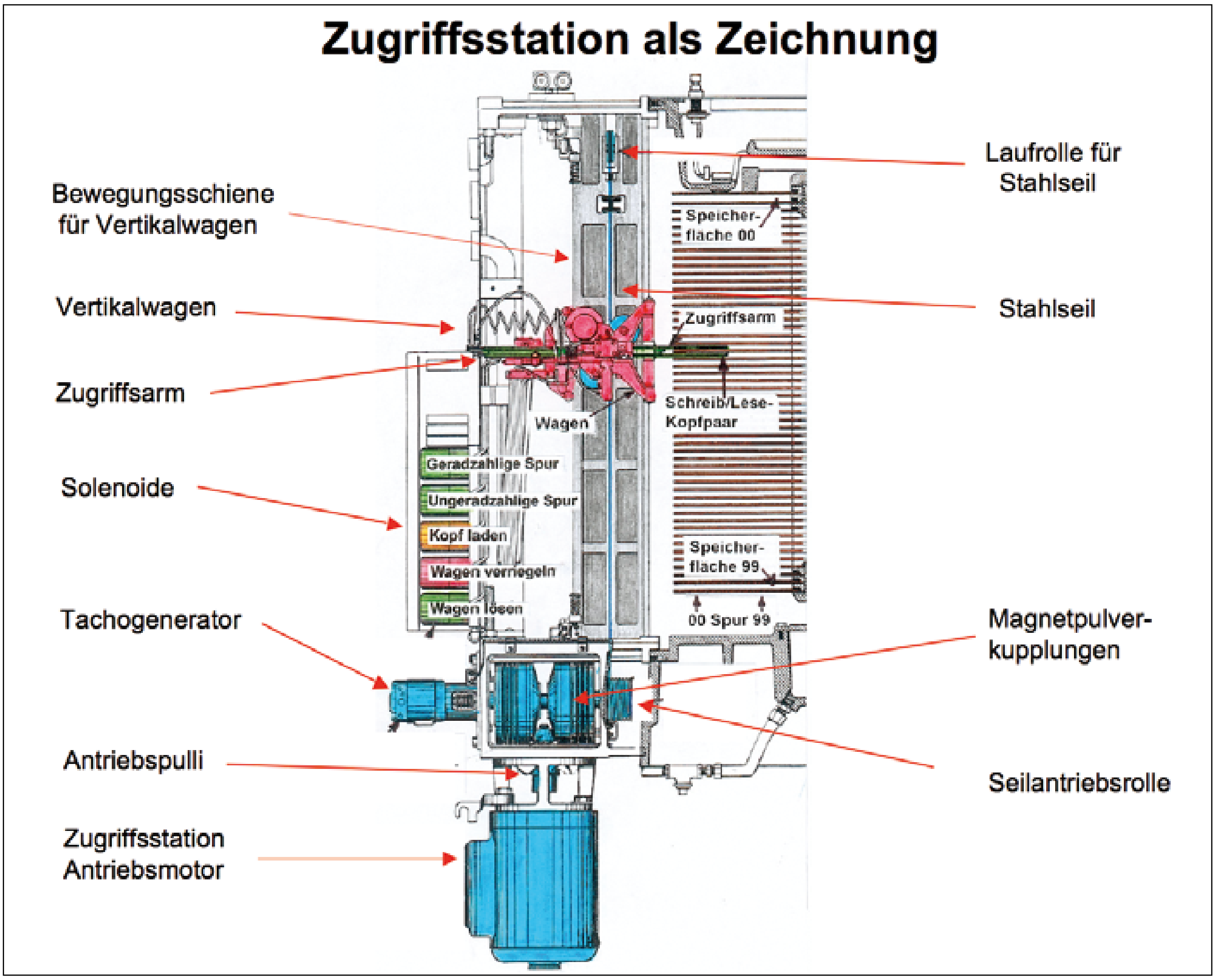


Fig. 1

if firmly locked. The graphic "Access station as drawing" (Fig. 1) shows the carriage colored in red and the access arm in green.

Hans Spengler:

This is how colleague Graichen described the way, but this has yet to be done. A refined steel cable drive with a cable drive roller (see Fig. 1 below), a roller (see Fig. 1 at the top) and two other rollers on the trolley itself cause both the up and down of the trolley and the in and out of the access arm. To do this, the rope drive roller must be able to run forwards and backwards. This creates a magnetic powder double clutch for one or the opposite direction of rotation. Both halves run on a common shaft that drives a speedometer generator (see Fig. 1 left) and the cable drive pulley (see Fig. 1 right). Both coupling halves are conical, so that the rubber-coated cone of the access station drive motor can drive the two halves in opposite directions.

one of the two halves. The access station drive motor runs continuously.

Kurt Gerecke:

So far the mechanics are clear. But how does the "Search" computer instruction enable the access station to automatically carry out and implement corresponding movements using the information on the disk surface and track?

Heinz Graichen:

Based on the previous explanation, we consider the drawing "flow chart" (Fig. 2). At the bottom left we find the cable drive pulley, the speedometer generator and, as a “black box”, the electronic control stages for both drive coupling halves. The control voltage entering the box can be positive or negative. If it is positive, the carriage is moved in one direction to the desired plate. If it is negative, the carriage is moved in the other direction to the desired plate ("up / down"). This also applies to the direction of movement of the access arm ("in / out").

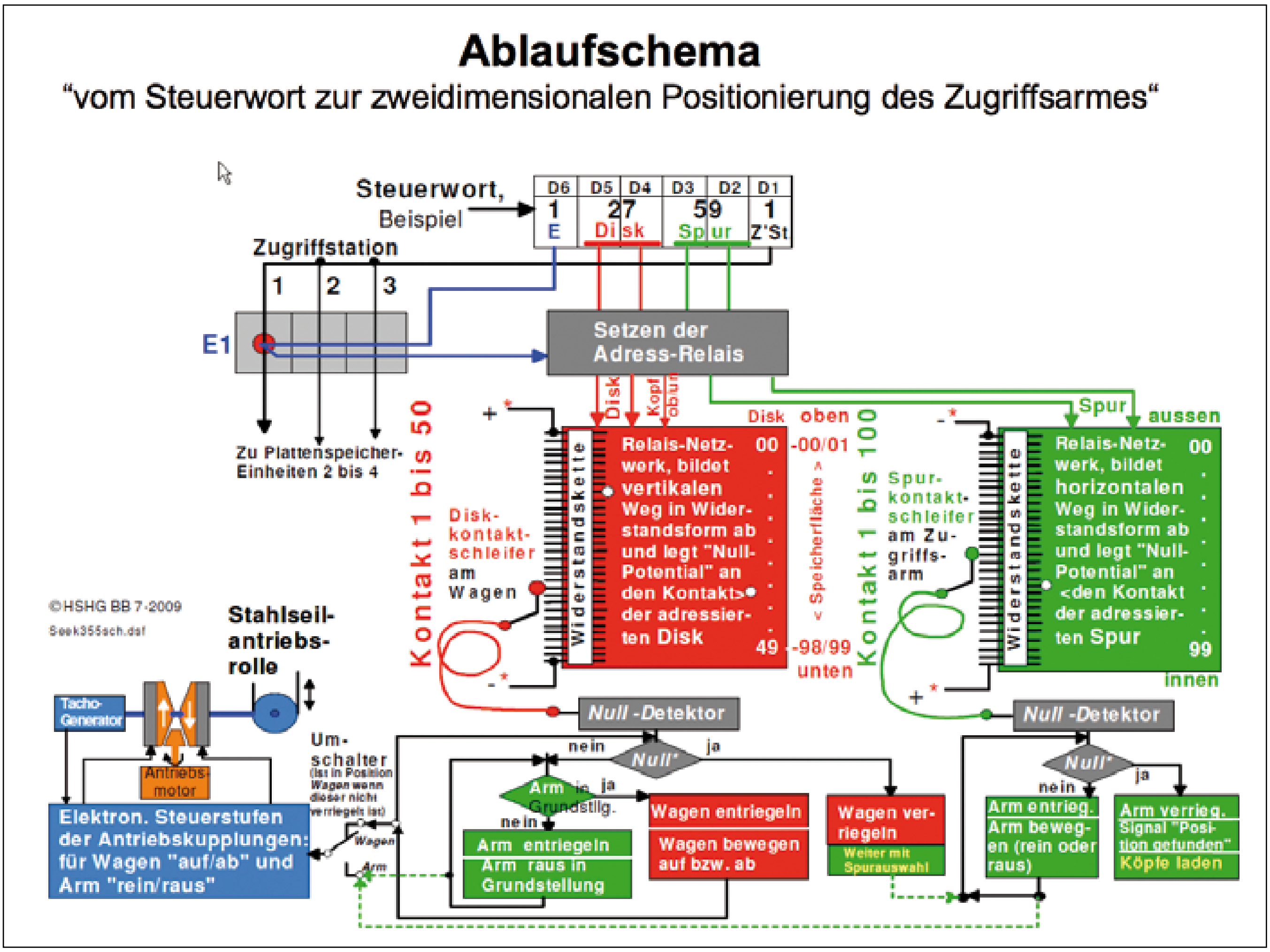


Fig. 2

We already know the control principle of the drive clutches, but not yet the conversion of the addressed storage disk (one of 50) and the addressed storage track (one of 100) into corresponding processes and movement paths.

Hans Spengler:

For the implementation of the control principle in movement paths, further elements are required, which Heinz Graichchen and I have summarized from extensive circuit diagrams into a simplified diagram (see Fig. 2 "Flow diagram"). The essentials are an address relay set, two resistance chains, each with a relay contact chain, a zero detector and a switching logic that controls the processes, including the drive clutch control stages.

The address relay set: saves the disk side (0099), the relay contact chains provide zero potential for the target address and control the later S / L head selection. Relays also save the track address (0099). The address relays are "2 of 5 coded" for the purpose of validity checking.

Resistance chains: Each movement step is represented by a resistance. Connected in series, each path of movement is represented in terms of resistance. A non-grounded DC voltage (150 V) is applied to each resistor chain. Each chain has a sliding contact, which decreases the actual position in the form of a tension and its polarity. A chain is used for the disk and a chain for track display. The address relay contact chain applies zero potential to the target position.

Zero detector: If the actual position and target position are present as voltage difference and polarity at the zero detector, this delivers the signal "Not zero", ie movement towards the target must be initiated. We consider the example of wagon movement. The latter may only be moved when the “arm is in the basic position” if it is “fully extended”. Then the car is first unlocked and then moved until the "zero signal" stops the movement, locks the car and then leads to the lane selection. The same applies analogously to the latter. The "zero signal" stops the arm movement, the arm is locked, the signal "position found" is generated and the read / write heads are loaded.

Kurt Gerecke:

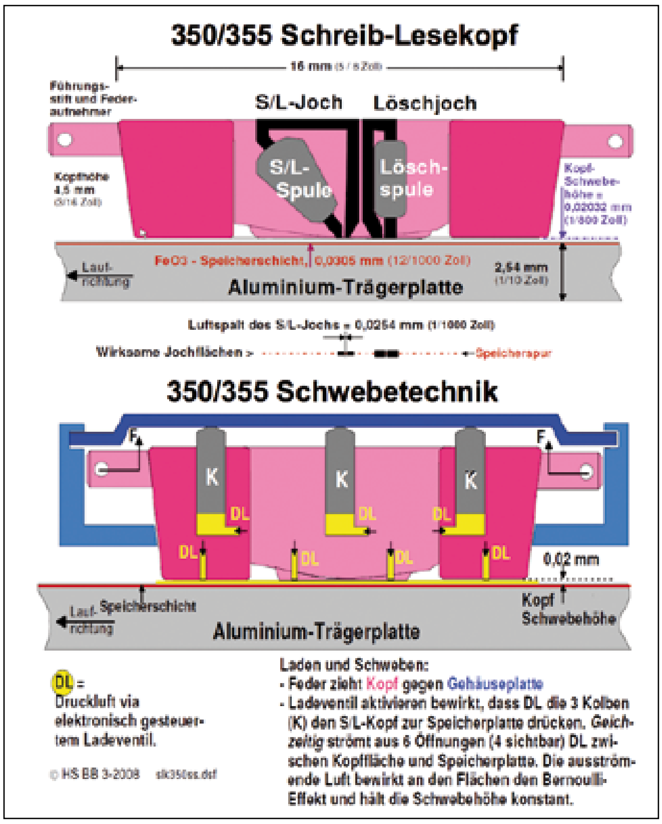
Why hasn't the resistance chain for disk and arm been designed uniformly?

Heinz Graichen:

For design reasons, it was not possible at the time to design the resistance chain for disk and arm uniformly. For the arm, a special potentiometer with 10 precise taps was used together with a measuring bridge to simulate tracks 00 to 99 using a resistance chain. However, this did not change the principle applied. The potentiometer is attached to the carriage and is actuated by the rack on the access arm and a gear wheel also attached to the carriage. The rack of the arm is locked for odd-numbered tracks by an "odd detent". The even detent locks the even-numbered tracks. This doubles the tooth size, the stability and thus the reliability.

Kurt Gerecke:

Now we have reached both the addressed memory area and the addressed memory track. The access arm carries the pair of read / write heads. How is this “pioneer read / write head” structured?



Hans Spengler:

The head structure can be seen in the picture "350/355 read / write head": the read / write coil with yoke (black) is on the left and the erase coil with yoke is on the right. The yoke surfaces are projected downwards against the storage plate and show the air gap size of 0.0254 mm and the applied principle "erase wide and read narrow".

The entire magnetic element is guided through a conical hole in the access arm frame (see figure "350/355 suspension technology" in light blue). A hairpin spring engages in both holes of the head guide pins (F) and presses the head against the cover plate until its guide pins stop there. This state is known as "unloading the head". In order to prepare for reading or writing, the control electronics generate the "Head Load" signal (see Fig. 2 "Flow diagram" at the bottom right). This excites the "head load solenoid". This allows compressed air (DL) both to escape from the six 0.1 mm hexagonal holes in the sliding surface of the element and to act simultaneously on three cylinders with pistons arranged at a 120 degree angle. This causes the magnetic element to move downwards.

Summary:

CPU (305):

- ***In the instruction stream of the central processing unit, the Search command decoded. Transfer of the search command (control***
- ***word) and the Address components to the control unit 652 and further to the 350/355.***

- ***The central processing unit ends the introduction of Search instruction and continues in the instruction stream.***

350/355:

- ***starts the sequence of the required sequences.***
- ***Activate zero potential at the address-equivalent contact the resistance chain "Disk". Corresponding to zero detector output a)***
- ***non-zero = arm unlock and move to the basic position. Then unlock the carriage and move it vertically towards the target plate, b) if " Disk zero result is reached, lock the cart and transfer to start " Track Search ". Track search, with the access arm horizontal, as long as the***
- ***Zero detector " Non-zero ", drive. If " Disk zero "is reached, lock the arm. The signal " Generate position "and load heads. The access station for a read or write***
- ***operation prepared. Then the upper or lower S / L head is electronically activated by the held address relays.***

- ***The movement paths determine the access time. mini- Times occur when changing lanes on the same memory surface (11 ms). Vertical movements of the car require more time. Maximum time is from " top inside down inside "(800 ms). There are two intermediate stops due to the transitions horizontal / vertical and then vertical / horizontal.***

Optical storage technologies

CD-ROM (Compact Disc Read Only Memory)

In 1983 Philips and Sony introduced the CDROM, a kind of offshoot of the music CD. The medium is also structured similarly to the music CD. The data is saved during the production of the disk and the data can only be read (analogy: ROM). In contrast to magnetic disks, recording takes place in a single, spiral track - just like a record. Holes (pits) are burned into this pre-embossed, reflective layer during the manufacture of the master plate using a laser. Any number of copies can then be made from the master plate.

The copy is scanned by the laser beam, which is modulated by digital information due to the different structure of the storage area. The track density is up to 16,000 tracks / inch. The ISO 9660 format has established itself as the recording standard (transfer rate: 1.2 MBit / s, capacity: approx. 600 MB). The CDROM is mainly used to distribute larger amounts of data, as a photo CD and more recently as a video carrier. The first CDROM application to be delivered on CDROM was the Microsoft Book shelf in 1987. In 1990 the Commodore CDTV was launched, a computer based on the Amiga 500 with an integrated CDROM drive. In 1993 and 1994 NEC specified the clock with triple (triple speed) or quadruple CDROM drive (quad speed).

The structure of the writable CDR is more complex than that of the CDROM. The base layer is made of polycarbonate, followed by a light-sensitive organic substance that is translucent. Then there is a reflective gold layer and finally a protective lacquer layer. With increased laser energy, the organic material can be discolored or melted, giving it a different reflection property. The disk can then be read like a CDROM.

WORM (Write Once Read Many)

WORM plates can be written by the user, but only once (analogy: PROM). With 5.25 inch disks, storage capacities of well over 1 GB are possible (plasmon today offers 60GB disks based on blue laser technology in the IT environment). WORM can be used to archive all types of data (backup medium). The disk works like a magnetic disk drive and can be addressed in exactly the same way; the driver software ensures that the most recent version is always addressed if a file is saved several times (older versions can be read using special programs) -> Saving a file chronology. When writing, the plate structure is permanently changed by high laser energy. When reading, this change is sensed and detected with low laser energy.

When the bubbles are generated, the laser beam heats a polymer layer that lies under a thin metal film. A bubble is formed that permanently deforms the metal film. When scanning with low laser energy, the changed scatter can be evaluated.

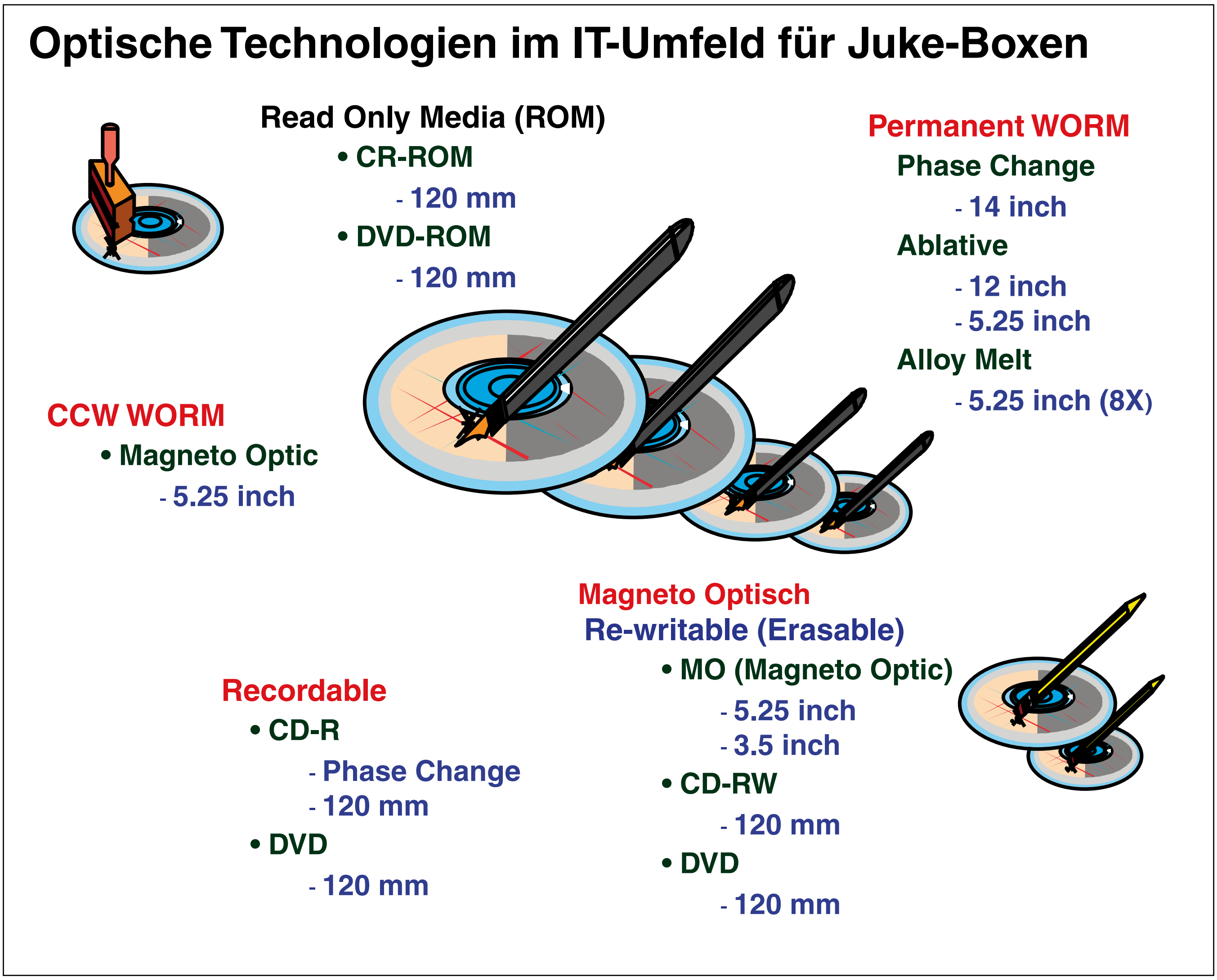
When the pit is created, the laser beam burns through an opaque layer that lies over a reflection layer (pit is created). When reading, the resulting light-dark zones are evaluated.

At the end of the 1990s, the durability of data on WORM plates was estimated to be around 300 years.

120 mm optical disk in the IT environment for long-term archiving (WORM, MO and CCW)

MOD (Magneto Optical Disc)

MOD is related to WORM, a modern storage medium for computers on a magneto-optical basis. It allows digital recording and deletion via laser. The MOD contains up to one billion micromagnets in order. These are premagnetized in a certain direction. At the points heated with a laser beam, these micromas can be reversed by an applied magnetic field. When scanning, the laser beam is circularly thrown back by the magnets, which are now polarized differently, either clockwise or anti-clockwise (magneto-optical Kerr effect). The change in polarization can be read using appropriate optics. The MOD is less sensitive to light than the CDR, can be written to more quickly, is insensitive to moisture,



CCW (Continuous Composite WORM) is based on the MO technique. The plate is given a label during production so that it cannot be overwritten. In addition, software in the drive protects the disk from being overwritten. CCW is therefore a “bill” WORM plate that is based on MOD technology.

In the IT environment, the technologies were standardized on a 120 mm (5.25 inch) basis as MO, CCW and WORM plates in the early 1990s. The 120mm technology is still used in the IT environment for archiving purposes. Optical plates based on 90mm (3.25 inches) are primarily used in the consumer environment.

UDO (Ultra Density Optical Disk)

In 2003 and 2004, the 120mm technology, which had previously worked with a red laser, was replaced by a new standard, UDO technology (Ultra Density Optical). The UDO technology is not backwards compatible with the old 120mm technology and works with a blue-violet laser with a 405nm wavelength. The plate formats WORM, MO and CCW have been retained. UDO offers a "Phase Change" WORM and "Phase Change" rewritable medium for use. The first plates on this new laser base as the new 1x standard offered a capacity of 30 GB. The second generation as a 2x standard, which became available in 2006, offered twice the capacity, i.e. 60 GB. The UDO technology as the new 120mm standard will be further developed according to a roadmap that has been anchored in the UDO standard.

UDO roadmap

	Generation 1	Generation 2	Generation 3
Capacity	30 GB	60 GB	120 GB
Transfer rate	8 MB / s	12 MB / s	18 MB / s
RPM	2,000	3,000	3,600
Avg Seek Time	25 ms	25 ms	25 ms
Numerical aperture	0.7	0.7	0.85
Media layers	1	2nd	2nd
Encoding	1.7	1.7	ML
Sector size	8 KB	8 KB	8 KB
SCSI transfer rate	80 MB / s	80 MB / s	80 MB / s
Load time	5 s	5 s	5 s
Unload time	3 s	3 s	3 s
MSBF	750,000	750,000	750,000

DVD (Digital Versatile Disk)

DVD stands for "Digital Versatile Disk" (formerly "Digital Video Disk"). The medium is as large as a normal CDROM, but a much higher storage density is used. There are four different media. The simple single-sided DVD can store 4.7 GB on one layer. But there are also two-layer DVDs. The information is stored on two superimposed layers, one of which is transparent.

The right layer is controlled by focusing the laser differently. This makes 8.5 GB possible. And then there is the whole thing on two sides. This enables 17 GB of data on a single DVD. The first drives came on the market in the 90s and can read a layered, single-sided DVD. Unfortunately there were still few DVD titles with videos at this time. The videos were encoded in MPEG2, which results in very good quality when played back. The first burners for single-sided, single-layer DVDs had already been introduced in the 1990s; the Pioneer burner was offered in autumn 1997 for a little over DM 10,000. It was recorded at approx. 1 - 2 MB / s and could save a maximum of 3.9 GB.

The readers can also read normal CDs, but mostly not CDRs, i.e. the writable CDs. This is because a laser with a shorter wavelength is used, which can no longer read the self-burned CDs correctly. For this purpose, Sony had developed a drive with two laser diodes, which can then be used to read the CDRs again.

HD DVD

The HD DVD (High Density Digital Versatile Disk, previously: Advanced Optical Disk, short: AOD) is a data carrier format and was traded from 2005 to February 2008 alongside the Bluray Disc and VMD as a possible successor format to the DVD.

The HD DVD was specified by the Advanced Optical Disc Consortium (AOD), which included NEC, Microsoft, Toshiba, Intel, IBM and HewlettPackard. Afterwards, these companies merged to form the HD DVD Promotion Group in order to make the HD DVD better known. In November 2003, the DVDForum had adopted the HD DVD as the HD successor to the DVD in accordance with the HDDVD specifications for ReadOnlyDiscs. The Advanced Access Content System (AACS) copy protection from the area of digital rights management (DRM) was provided for the HD DVD.

In March 2006, Toshiba's HDXA1, the first HDDVD player, was launched in Japan. In August 2006, Elephants Dream, the first HD DVD in German, was released.

In February 2008, Toshiba announced that the development, manufacture and marketing of the technology would cease at the end of March 2008.

The HD DVD was based on a blue-violet laser with a wavelength of 405 nm. The thickness of the carrier layer is with 0.6 mm identical to that of the DVD. In contrast, the numerical aperture (NA) is 0.65 compared to 0.6 for the DVD and 0.85 for the Bluray disc.

The HD DVD had one layer and a storage capacity of:

- ***15 GB (for HD DVD-ROMs - pressed media) or***
- ***15 GB (for HD DVD-R / RWs - once and rewritable bare media) or***
- ***20 GB (with HD DVD-RAM - rewritable media with optional sector access)***

and with two layers a storage capacity of

- ***30 GB (for HD DVD-ROMs - pressed media).***

In addition, a three-layer version was approved by the DVDForum in August 2007, with space for 17 GB per shift and thus having a total capacity of 51 GB. The definitive standard for films on HD DVD initially included the 15GB and the 30GB variant, whereby the two-layer 30GB variant was almost always used for films.

In January 2008, Time Warner announced that its Warner Bros. and New Line Cinema studios would not release any more films for the HD DVD in the future, but would only use the Bluray disc. This has been described by some media as a decision of the format war. Over the next few days, other providers followed, such as the large European film distribution company Constantin Film on January 10, the US porn provider Digital Playground on January 12 and the German Senator Film. The US providers Universal Studios and Paramount Pictures, however, denied rumors of abandoning their contracts with HD DVD, they are considered the last remaining large providers. After Toshiba, of which a large part of the HD DVD players came, had drastically reduced the prices of the devices in the USA on January 15, some media spoke of a “sellout”. On the other hand, the pressure on the HD DVD continued to increase: at the beginning of 2008, MediaSaturnHolding stores took the HDDVD player in payment when buying a Bluray player. The format suffered further setbacks in February 2008. The USE electronic chain Best Buy announced on February 12, 2008 that it would only concentrate on the Bluray format in the future. A day earlier, the largest US online video rental company Netflix announced that it would remove the HD DVD from its range by the end of the year. On February 15, 2008, WalMart, the largest retailer in the USA, announced that it would sell out HDDVD inventory and would only want to rely on Bluray in the future. On February 19, 2008, Toshiba issued a press release announcing that the development, Production and sales of the HD DVD and corresponding devices will no longer be pursued and will be discontinued at the end of March 2008. As a result, the same day, the Universal Studios announced that they would switch from the HD DVD to the Bluray format.

Blu-ray disc

The Bluray Disc (abbreviated BD or more rarely BRD) is a digital optical storage medium. In addition to HD DVD and VMD, it was advertised as a possible successor to the DVD. After Toshiba announced in February 2008 that production and further development of HDDVD technology - including the devices - would be discontinued in March 2008, the Bluray Disc was the winner in the format war. The name Bluray is of English origin. Blue ray literally means “blue light beam”, which refers to the blue laser used. The deliberate deviation from the orthographically correct spelling "blue ray disc" aims to encourage registration of the printout as a brand.

The Bluray Disc specifications were approved on February 19, 2002 by the nine companies of the Bluray Group, Panasonic, Pioneer, Philips, Sony, Thomson, LG Electronics, Hitachi, Sharp and Samsung; This group also joined Dell and HewlettPackard at the end of January 2004 and Apple in mid-March 2005. However, HewlettPackard resigned from the Bluray consortium in 2005 after some suggestions for improvement were rejected and switched to the HDDVD camp.

The new PhaseChange technology should enable a double transfer rate (data transfer rate) of 9.0 MB / s (72 Mbit / s) when writing instead of the originally maximum specified simple 4.5 MB / s (36 Mbit / s). An important part of the specification is copy protection in the form of a unique identification number. Bluray discs are also particularly suitable for Full HD videos, which, thanks to the high resolution, offer better quality than conventional systems such as PAL and NTSC, but also require more storage space accordingly. BDS players generally support the high-definition AVCHD format, which was first introduced by Pana sonic and Sony.

The Bluray Disc is available in three versions: as a readable BDROM (comparable to DVDROM), as a write-once variant BDR (comparable with DVD ± R) and as a rewritable BDRE (comparable with DVD ± RW). By using a blue-violet laser and its short wavelength (405 nanometers), higher data densities can be achieved than with a DVD. 25 GB of data can be stored on a BDROM per data layer, so that a capacity of up to 50 GB results with two-layer media.

Another innovation compared to the DVD is the reduced distance of the laser to the data carrier and the shorter wavelength (and therefore different color) of the laser beam. Furthermore, the protective layer on the laser side is significantly smaller at 0.1 mm compared to 0.6 mm on the DVD. Due to the resulting greater susceptibility to scratches and dust, it was initially planned to offer the Bluray Disc only in one cartridge. Instead, a coating called "Durabis" was developed, which makes the use of a cartridge no longer necessary.

A four-layer version of the Bluray Disc, which is said to have a capacity of 100 GB on one side, was presented by TDK. TDK has managed to accommodate 200 GB on a six-layer disc. The capacity of a layer was increased to 33.3 GB. However, these are primarily feasibility studies. Writable media are also planned from the outset, as the Bluray Disc is primarily intended for entertainment electronics and serves as a storage medium for high-resolution films. The rewritable Bluray Disc works with the phase change technique.

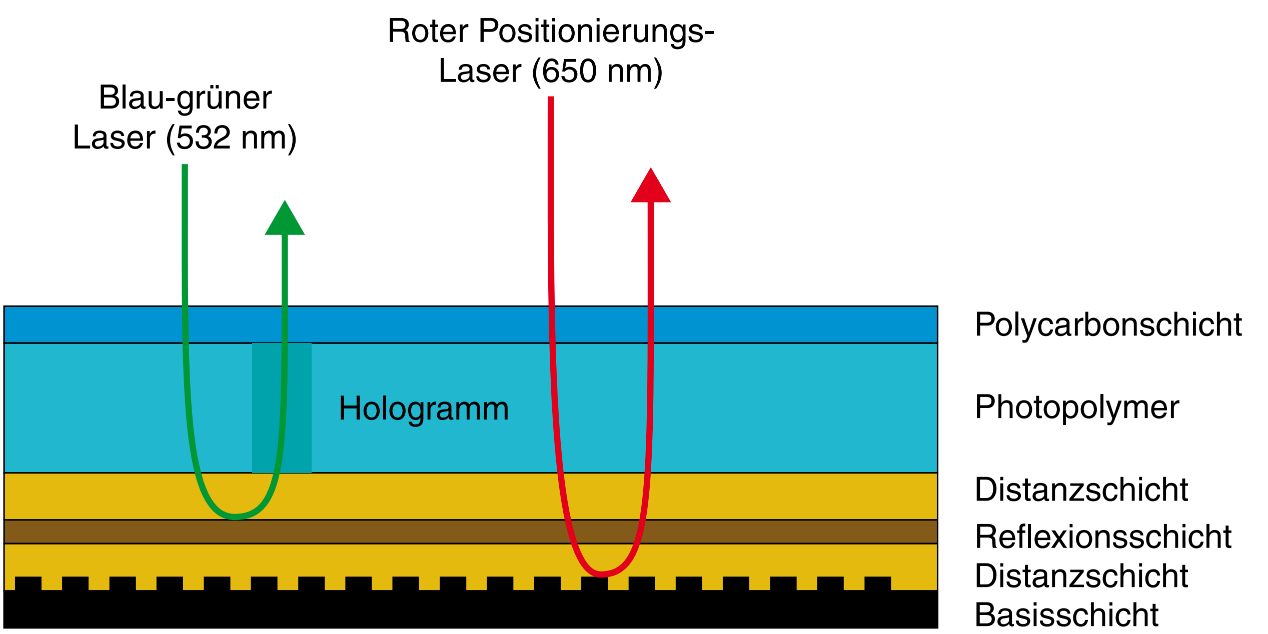
Because of the smaller distance between the medium and laser optics and the thinner protective layer, a lens with a more favorable numerical aperture can be used, which can focus the beam more efficiently. Thus, who reduces the typing errors and greater scatter and it is possible to combine a Bluray disc made of metal or other stable, opaque materials to build with a thin transparent support layer, which can be operated at significantly higher speeds than a disk made of polycarbonate, which then results in higher transmission rates. In addition, the smaller wavelength of the laser beam compared to the DVD allows a much higher data density and thus an increased storage capacity.

Verbatim presented a "Bluray Disc Rewriteable" at IFA 2005. According to the manufacturer, this medium should allow a storage capacity of 25 gigabytes (i.e. approx. 135 minutes of video in MPEG2 HD quality). In addition, the BDRE should be provided with a particularly impact and scratch-resistant protective layer. In addition to the immediate availability of 25GB blanks, TDK has also announced 50GB discs for October 2006. TDK is one of the founding members of the Bluray Disc Association (BDA) and was instrumental in the development.

Holographic Versatile Disc (HVD)

The holographic versatile disc, also known as HVD, is already traded in specialist circles as the successor to today's HDDVD and Bluray technology. HVDs have a capacity of up to 3.9 TB, which is many times more than the largest 200 GB Bluray disk available today. The transfer rate is also many times higher and reaches 1 Gbit / s compared to 36 or 72 Mbit / s with the Bluray plate. In addition, the HVD is far from exhausted and new drives with even higher rotational speeds are conceivable. The specification for the HVD was approved in December 2004 by the TC44 Committee of Ecma International. The "HVD Alliance" was founded in February 2005 to drive the development of HVD. The Alliance now includes 19 companies: Alps Electric Corporation, CMC Magnetics, Dainippon Ink and Chemicals, EMTEC International, Fuji Photo Film, Konica Minolta Holdings, LiteOn Technology, Mitsubishi Kagaku, Nippon Kayaku, Nippon Paint, Optware Corporation, Pulstec Industrial, Shibaura Mechatronics, Software Archi tects, Suruga Seiki, Targray Technology, Teijin Chemicals, Toagosei Company and Tokiwa Optical Corporation.

Two superimposed lasers with different wavelengths, a blue-green write / read laser with 532 nm and a red positioning laser with 650 nm, are reflected on two different layers in the plate and generate an interference field in the photopolymer layer, a so-called hologram. In the schematic drawing, the two lasers are not superimposed to simplify Darge to better illustrate the principle of different reflection.



Prinzip der HVD (Holographic Versatile Disc)

When reading, the blue-green laser reads the hologram of the data encoded as a laser interference pattern in the polymer layer in the upper area of the medium, while the red laser serves to read out auxiliary information from a CD-like aluminum layer in the lower area. The auxiliary information is used for exact positioning where you are on the disk (comparable to sector, head and segment information of a hard disk). A dichroic mirror layer (reflection layer), which is located between the polymer layer and the aluminum layer (base layer), allows the blue-green laser to reflect while the red laser is passing through.

Polymers, i.e. plastics, are long chain molecules that consist of the same building blocks. They have the advantage over crystals that they can be modified almost indefinitely. Polymers are extremely sensitive to light, highly transparent and insensitive to temperature fluctuations. They do not change their performance even after frequent reading and are ideal for laser processing. After exposure to laser light, the photosensitive molecules in the polymer change their orientation. In the exposed areas, the material deflects light more than in the unexposed areas. The “holographic” information is contained in this change in the molecular order, which is targeted as an interference field. Since the data is not recorded and read as individual bits, but as whole bit patterns, extremely many bits can be called up at the same time with a single “flash” laser. Depending on the type and thickness of polymer, this can be several hundred thousand bits. The transfer rates are gigantic. A feature film that now fits on a DVD could be read in about 30 seconds. The large storage capacity is due to the fact that the holograms are not only written on the surface of a storage material, but in the entire volume.

Particularly sensitive data could also be encrypted using a special mask that is placed between the storage material. In contrast to information (e.g. on chips) that is encrypted using software, the hardware encryption of holograms cannot be broken in principle. For the same reasons a hologram is also difficult to fake. The shelf life of information in holographic memories is now estimated to be around 50 years.

Flash memory - short story

The development of the flash memory known today took place relatively late in the 90s. The driving force was digital photography. After all, it is practically impossible to install a CD burner in a small digital camera, which burns the pictures taken directly onto the CD. The floppy disk was also unsuitable because it simply saved too little data too slowly. The microdrives developed by IBM at the end of the 1990s, which were an alternative to flash memory cards for a few years, consumed a lot of power in use, so that the battery of a digital camera often had to be charged. In the long run, the microdrives were therefore not a sensible alternative in digital photography.

The first flash memory was launched by Sandisk in 1994. At that time it included 4MB memory. In the meantime, this form of data backup has gone many technological steps further, since you could hardly do anything with 4 MB storage capacity. Nowadays, USB sticks based on this technology are available with several gigabytes.

The term flash memory came about in the Toshiba laboratories, where researcher Shoji Ariizumi suggested this name because deleting the data blocks reminded him of a flash light (in English: "Flash").

Flash memory of a compact flash memory card (photo: SANDISK)

Description and functionality

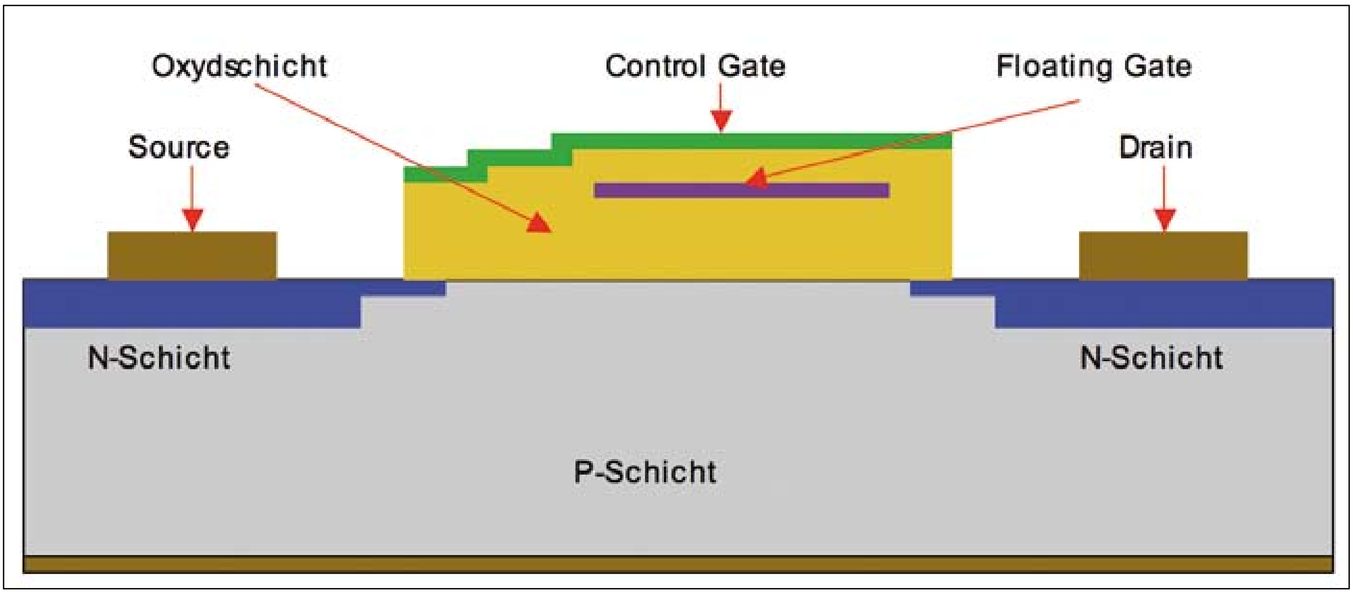
Flash memories are based on semiconductor technology and are found in all memory cards that are common today. They are used in hybrid hard disks or flash hard disks in personal computers, but also in memory cards for work computers or in fast buffers of mainframes; but also above all in SIMMs, PC Cards and PCMCIA cards, in digital cameras and MP3 players, in USB sticks, camcorders and printers, notebooks and handhelds. There are special flash modules for note books, laptops, PDAs, printers and other peripheral devices such as digital cameras. The different versions have a protected designation and spelling: Examples are the CompactFlash card (CF), the Smart Media Card, Memory Sticks, the Multi Media Card (MMC), the SD card and others with storage capacities of several gigabytes (GB). Flash memories can be read in special card readers that are connected to the personal computer (PC) via an interface cable such as USB. In addition, the flash memories can also be operated via PCCard adapters.

In addition to a very compact design, the basic requirements for a memory card are high storage capacities, rewritability, the use of power-independent memory modules and the lowest possible susceptibility to interference in mobile use. These requirements were and are primarily defined by digital photography.

In contrast to hard disks, flash memories no longer have any moving parts integrated, which means that mechanical components of a hard disk are excluded and are taken over by semiconductors, i.e. electrical components, with regard to their functionality. These must be able to store data permanently and non-volatile, i.e. independent of electricity. Even if the power supply is cut off, the information must not be lost.

A flash cell consists of a modified MOSFET transistor (metal oxide field effect transistor), which is equipped with an additional electrode, the so-called “floating gate”. This electrode is placed between the control electrode (control gate) and the active transistor channel and surrounded with a thin semi-permeable separating layer. If free electrons are now "pressed" through the barrier, usually an oxide layer, onto the floating gate using a higher voltage (write voltage), the transistor behavior changes.

Each transistor can either conduct or block an electrical charge in the form of electrons. This corresponds exactly to the binary information states of a bit, ie "0" and "1". The transistor reveals the stored information when a voltage is applied and depending on whether current is flowing or not, it has the logical information "0" or "1".



Construction of a flash memory cell with floating gate

The floating gate is completely electronically isolated from all other components by an oxide layer, so that the stored charge, i.e. the information on the data cell below, is really stored in it and cannot be lost over a long period of time. Under normal circumstances, this isolation would not allow transmission because the isolation would block it. The flash memory therefore uses the so-called tunnel effect from quantum mechanics, which makes it possible to transfer charges even through a non-conductive material. Depending on the amount of electrons absorbed, the floating gate becomes an electronic throttle valve for the transistor function. When reading out a transistor memory cell, the resulting current is evaluated as an indicator for logic "0" (many electrons in the floating gate) and logic "1" (few or almost no electrons in the floating gate). A charged "floating gate" thus reflects "0" and an uncharged one is rated as "1". A different assignment would be possible, but was never used. Each individual bit is thus transmitted via the "floating gate". In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process makes it possible to change this state in the other direction. A charged "floating gate" thus reflects "0" and an uncharged one is rated as "1". A different assignment would be possible, but was never used. Each individual bit is thus transmitted via the "floating gate". In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process makes it possible to change this state in the other direction. A charged "floating gate" thus reflects "0" and an uncharged one is rated as "1". A different assignment would be possible, but was never used. Each individual bit is thus transmitted via the "floating gate". In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process makes it possible to change this state in the other direction. In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process makes it possible to change this state in the other direction. In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process make

The schematic diagram "Structure of a flash memory cell with a floating gate" shows how the floating gate forms a charge trap in which the electrical charge is stored. It lies in an oxide layer below the control gate and normally prevents the charge from flowing out to the N and PS layers (N = heavily negatively doped electrodes drain and source, P = heavily positively doped substrate). The charge on the floating gate forms a conductive channel between drain and source via its electrical field, via which the charge is read out. The deletion takes place block by block. By applying a negative quenching voltage, the charge carriers are driven out of the floating gate.

Further considerations

Flash memories are current-independent and therefore fundamentally differ from current-dependent (volatile) RAM modules (Random Access Memory), because the information stored in the main RAM of a computer is only kept as long as the computer is running. If the computer is without power, RAM modules lose the ability to save data.

A classic bipolar transistor requires current for several functions. On the one hand, electrons are released from the surface of the transistor to change the switching state. Information is saved or deleted in this way. On the other hand, the current in the form of "refresh" pulses is required to stabilize the state reached.

MOSFETs, on the other hand, are able to use a thin oxide layer to build up a small capacitor that transports charge carriers between the silicon layers via a built-up electrical field. This creates tension until a limit is reached. The result is a thin conductive channel that remains stable without a "refresh" pulse. MOSFETs therefore only need current to charge or discharge the capacitor, while in the other case current must be supplied continuously in order to keep the conductivity in a stable state and to maintain the stored information.

NAND architecture

The name NAND, which describes the NAND flash memory, actually comes from the statement logic. Here, a gate is described that represents a specific AND operation (Not AND = NAND). Circulated in the world of electronics, it means that the memory cells on a flash memory are connected in series, that is, they are arranged in series on different data lines. The metal oxide semiconductor field transistors are connected to each other via a NAND gate. The great advantage of this architecture is that only a small number of data tracks are required, since one track contains several memory cells. In comparison to the NOR architecture, space savings of well over 50% can be achieved.

The access times of the NAND architecture are relatively high, since you cannot access individual memory cells, but can only read out the entire line. So you have to read the data in blocks. You use the page and block-oriented method of working. A page consists of 512 or 2,048 bytes, which in turn are combined into blocks that are between 16 and 128 KB.

As already mentioned, you can only rewrite a memory area, i.e. a page, if you delete it beforehand. Since the data cells are connected in series, you have to delete the entire block in which the page is located.

NANDFlashs are used in areas where a relatively large amount of memory is required and the FlashModule should be small and very compact. The slower access times compared to the NOR architecture are somewhat disadvantageous.

NAND memories or SLC flashes (single level cell) have largely prevailed over NOR memories due to their high storage density and cost-effective production method.

NOR architecture

The so-called NOR memory exists as a counterpart to the NAND memory. As the name suggests, this type of memory consists of NORGattern (Not OR) instead of NAND. This also results in the fact that the memory cells of a NORFlash are connected in parallel to one another, which enables parallel and therefore random access. So you don't have to go block by block when reading.

The parallel NORFlash has a lower storage density and is comparatively expensive, but executes polling processes much faster. NOR memories are also known as MLC memories (Multi Level Cell) and can be addressed flexibly due to their “grid” structure. Due to their design, they have relatively little storage space compared to NAND memories, since the parallelism simply takes up more space.

As a general rule

NAND is used in large data stores in commercial systems because it can be used to build more compact, faster and cheaper storage systems. In the NOR memory, the memory cells can be read out in parallel, which is why such memories are particularly fast for z. B. very suitable as a program or code memory. NAND memory cells can be packed more densely. However, many storage cells are connected in series. Therefore, the optional read access is slower. Erase and write are faster with NAND memories when working in blocks. The storage densities for NAND and NORT technologies for one-bit cells differ. Usually NAND uses single-bit cells and NOR multi-bit cells. However, both techniques are now also available in multi-bit versions with two, three or four bits in a cell. In general, multi-bit cells are slower and more prone to failure than single-bit cells. Multi-bit cells have the same cell size and the bits in the cell are defined by different voltage levels. This explains the cost reduction of MLC memories compared to SLC memory (single level cell). However, there are always difficulties in precisely reading out the small voltage differences. For this reason, SLCFlash in NAND technology should be used in the current state of the art in professional IT operations for reliability and speed reasons. Multi-bit cells have the same cell size and the bits in the cell are defined by different voltage levels. This explains the cost reduction of MLC memories compared to SLC memory (single level cell). However, there are always difficulties in precisely reading out the small voltage differences. For this reason, SLCFlash in NAND technology should be used in the current state of the art in professional IT operations for reliability and speed reasons. Multi-bit cells have the same cell size and the bits in the cell are defined by different voltage levels. This explains the cost reduction of MLC memories compared to SLC memory (single level cell). However, there are always difficulties in precisely reading out the small voltage differences. For this reason, SLCFlash in NAND technology should be used in the current state of the art in professional IT operations for reliability and speed reasons.

lifespan

Due to their mode of operation, FLASH memories cannot be kept indefinitely. Every write and erase process causes a certain amount of wear in the oxide layers of the cells, so that the manufacturers of flash memory cards specify a kind of minimum durability for their products that is based on the number of erase cycles performed.

With NAND (SLC) this value is 100,000 cycles, with NOR (MLC) only 10,000 cycles. A generous security buffer is always included in the calculation, which means that NAND memory cards can usually still function after a million or more erase cycles.

A kind of defect management additionally increases the lifespan of flash memory cards. As soon as the wear and tear within a block affects the integrity of the data, this block is flagged as defective and the data stored there is transferred to a reserve block. This block management also takes care of an even distribution of write accesses right from the start, so that the blocks are used as evenly and uniformly as possible.

control

The control that performs defect management has a significant influence on the performance of a flash memory card. In addition to the maintenance of the flash cells, it primarily deals with the communication of the memory card with the end device and is therefore largely responsible for the data transfer rates that are achieved when reading from and writing to a flash cell. The access times are around 100 µs, the reading speed is 25 MB / s and goes up to over 50 MB / s.

In addition to the speed, the reliability of the data communication and the protocol that ensures that the memory card logs in and out of the operating system smoothly are the responsibility of the controller. Controllers are used for this control.

With most memory cards available today, the controller is integrated in the data carrier unit, but in some models it is also used externally in order to achieve an even more compact design. The disadvantage of the external variant is that external controllers often have to manage more than one card model and are not optimally tailored to the requirements of each one. Loss of performance and compatibility problems can then result.

outlook

It is now assumed that there is a period of four to five quarters between two flash generations, in which the capacity can be increased by at least 50% or in many cases doubled. This can continue similarly in the next generations, but the curve could flatten out considerably! For this reason, new and different flash memory technologies are already being worked on. IDC expects the use of flash memory in mobile PCs to increase from almost 0 to 20% in 2011.

RAM - Random Access Memory

RAM is short for Random Access Memory. RAM is also referred to as main memory or main memory and is a type of memory whose memory cells can be addressed directly via their memory addresses. In this context, one speaks of “optional”, which refers to “random” and has nothing to do with “random” or “random access”. In this context one speaks of "random access memory" or "direct access memory".

RAM allows access to each individual memory cell, both reading and writing. RAM memory always requires power. If the power supply is switched off, the data in RAM are lost.

RAM is used as working memory in computer and microcontroller systems. Programs and data from external storage media and hard disks are loaded into this working memory. For fast processing, the processor can access it, process it and then write it back into the main memory.

Short story

The origin of the term goes back to the early days of computers, in which all data was stored in sequentially readable forms of storage such as punch cards, magnetic drum memories and magnetic tapes, which were loaded into fast calculation registers for processing. In order to have intermediate results available faster than these block-oriented devices, delay lines were temporarily used for intermediate values until the ferrite core memories were then introduced. These writable memories already had the same form of matrix access as today's RAMs. At that time, the fast memory types were all writable and the main innovation was the random access to the magnetic core memory and the RAM modules that were then placed on semiconductor memory.

functionality

RAMs are semiconductor memories and there are basically two types of RAM: SRAM as static RAM and DRAM as dynamic RAM. In the case of static RAMs, the information is stored in feedback circuits (so-called flip-flops), in the case of dynamic RAMs in capacitors whose charge is periodically refreshed. While of the recharge, the processor (CPU) has no access to the DRAM. This is why computers with DRAMs often work slower than those with SRAMs. However, the memory capacity of the DRAMs is significantly higher than that of SRAMs.

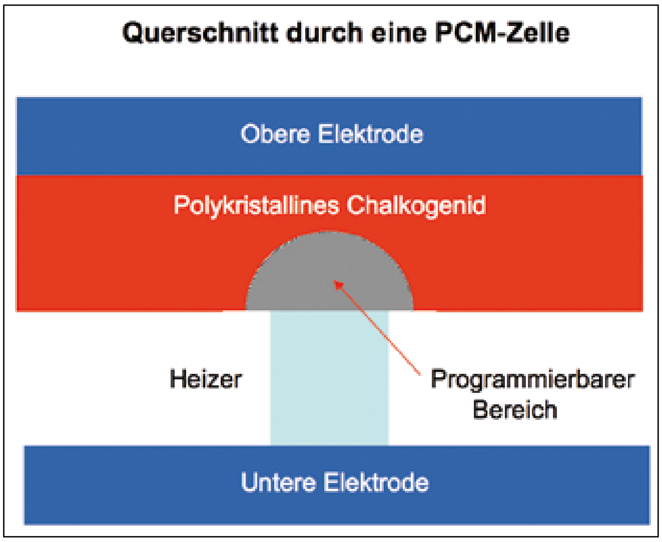
SRAM is static. This means that the memory content is stored using flip-flops and is thus retained after the memory content has been called up. As a result, the power consumption is very high, but this leads to fast work within the storage. Due to its high price and high power consumption, SRAM is only used as a cache or buffer memory with low capacities (e.g. L1, L2 and L3 cache of large computer systems). Its content is volatile, ie the stored information is lost when the operating voltage is switched off. In combination with a backup battery, a special form of non-volatile memory NVRAM can be realized from the static RAM,

DRAM is simpler and slower in comparison. For a long time, only this one type of memory with its various variants was known in the computer area for the main memory.

A DRAM memory cell consists of a transistor and a capacitor, which is the actual memory element. The information is stored in the form of the state of charge of a capacitor. This very simple construction makes the memory cell very small, but the capacitor discharges quickly due to the leakage currents that occur, and the information content is lost. Therefore, the memory cells must be refreshed regularly. DRAM is much cheaper than SRAM. That is why DRAM is mainly used for the main memory of a PC.

Phase Change Memories (PCM) - phase change memory

Already in the 1920s it was observed that the electrical conductivity changes due to a structural change on a chalcogenide. The semiconductor properties of crystalline and amorphous chalcogenides were researched in the 1950s. In the 1960s, reversible phase-changing materials started on their elek



trical and then look at optical properties.

Chalcogenides are chemical compounds from one or more chalcogen elements such as oxygen, sulfur, selenium and tellurium, which form solids with an ionic or covalent character with other binding partners (such as arsenic, germanium, phosphorus, lead, aluminum and other). The solids mostly appear crystalline.

In 1968 phase change memories were considered for the first time as a possible storage medium. At this point, however, the technology was not yet ready to keep up with other storage media such as DRAM and SRAM.

However, the chalcogenides were specifically researched further into optical storage and found their market with the CD / DVD RW. A point on a CD is heated with a laser so that it changes its state (amorphous to crystalline and back again).

Only when materials were discovered in the course of these developments that came to interesting regions in terms of writing times and flow, did the Phase Change RAM development pick up speed again.

With phase change memory, the chalcogenide is stimulated with a current pulse to change the phase in contrast to the optical application. Changing the state changes the electrical resistance. A distinction can be made between two or more states and more can be saved in one cell.

PCM - How It Works

Inside a phase change memory is a tiny amount of a semiconductor alloy that can quickly switch between an ordered crystalline phase with a lower electrical resistance and an unordered amorphous phase with a much higher electrical resistance. Since the maintenance of both phase states does not require a power supply, the phase change memory is not volatile.

The phase state of the material is determined by the amplitude and duration of the electrical current surge that is used to heat the material. If the temperature of the alloy is increased to slightly above the melting point, the atoms excited by the supply of energy are distributed in a random arrangement. If the power supply is then abruptly interrupted, the atoms solidify in this arbitrarily arranged, amorphous phase. If the power supply is reduced over a longer period of time - approximately 10 nanoseconds - the atoms have enough time to return to the preferred, well-ordered structure of the crystalline phase.

PCM versus flash

Technologically, PCM connects where NANDFlash reaches the limits of feasibility. Many experts believe that flash technology will face major problems in the near future due to its limited scalability.

In IBM's Almaden research laboratories, in collaboration with Macronix and Qimoda, new storage technologies were developed that can replace FlashMemories. The "Phase Change Memory" technology (PCM) enables significantly higher speeds for smaller sizes. This improves the properties of PCMs as they grow smaller. Less electricity is required because less material has to be heated and the speed increases due to the small distances. All advantages that can be achieved with the shrinking of PCM cells. However, as the cells become smaller, the electrical resistance also shrinks, making it more difficult to distinguish between the states. It is therefore important to find a good balance here.

"Together we have now succeeded in developing a new material for phase change memory that offers high performance even with extreme miniaturization," says Rolf Allenspach, head of the nanophysics group at the IBM research laboratory in Rüschlikon near Zurich.

NANDFlash is currently still on the rise, but the technology is still expensive, the capacities for use in computers alone are too low, and they run out after only 100,000 read and write cycles. This relatively fast material fatigue is not a problem for consumer applications, but it disqualifies the modules for use as main and buffer storage in network or storage systems. Hybrid solutions, which combine a conventional hard disk with flash modules, are also increasingly entering the market to enable faster data access.

Flash memory technology is likely to become a victim of miniaturization in the future. New production processes in chips make smaller and smaller traces possible, but according to current research, flash memory has ended at 45 nanometers. Litter losses prevent data from being saved without power.

The PCM modules, on the other hand, are intended to enable production processes of up to 20 nanometers and even less. The theoretical limit for the scalability of PCMs is less than 5 nanometers. The developers use a germanium alloy for their storage as new material.

PCM - small and universally applicable

Together with high performance and reliability, PCM could therefore become the basic technology for universal memory in mobile applications, because phase change memories, like flash, are not volatile. This means that they remember the information even when no power is being supplied. In addition, the PCM prototype from IBM has very advantageous technical values: It can be written 500 times faster and requires only half the energy.

In addition, significantly more read and write cycles are possible and the components with a cross-section of three times 20 nanometers are considerably smaller than the smallest flash memories that can be produced today. “As soon as the scalability of Flash finally reaches its limits, a new technology like PCM will prevail,” says Allenspach, from the IBM laboratory in Zurich. This could soon become possible.

Today's PCMs are about 30 times faster than Flash and range from 100 to 200 nanometers. However, if you consider the rapid development in terms of write and read speed, PCM is also suitable as a replacement for the RAM that is common today. With PCM as a replacement for RAM, the previous booting of computers would no longer be necessary, however, security problems could arise since the PCMs are not volatile (one could read out the confidential data in RAM such as passwords etc.). In addition, PCM has much less wear and tear, such as occurs with flash due to the wear of the oxide layer on the floating gate. PCM therefore offers a long data shelf life and not too high energy consumption during operation.

Millipede nano technology

The “Millipede” project was started in the IBM basic research laboratory in Rüschlikon in the mid-1990s. The aim of the project was to use nano-technology and nano-storage architectures for data storage and to integrate them into storage subsystems. The project manager in Rüschli kon was the IBM employee Peter Vettinger and the technological driving force was the IBM and Nobel Prize winner Gerd Binnig. The development of Millipede brought IBM many new patents in the field of nano-technology and nano-mechanics and gave many new insights into polymer structures and the possibility of using plastics as a data carrier. Millipede development is therefore of crucial importance for future technologies.

Basic principle

The basic principle is simple and is comparable to that of the earlier punch card, but now with structure sizes in the range of nanometers. Another crucial sub

schied: using the technology used, bits can be deleted and overwritten. Tiny levers with a fine tip made of silicon also melt tiny recesses into a polymer medium to write bits. The same tips can also be used to detect these depressions, i.e. to read out the bits again. To do this, bring the tip close to the polymer film and heat it. If the tip is immersed in a bit crater, the heat exchange between it and the storage medium increases, which reduces the electrical resistance of the lever, also called the cantilever. In order to overwrite a bit like that, the tip of the crater is used to create new depressions, the edges of which overlap the old depression and thus push the polymer material towards the crater.

Because the holes are so extremely small, they can be placed very close together and such fantastic data densities can be achieved: With revolutionary nanotechnology, IBM scientists in the Rüschlikon research laboratory have managed to penetrate the millionth of a millimeter range. A record density of 1 terabit per square inch could be exceeded when storing data - which corresponds to the content of 25 DVDs on the surface of a stamp. The Terabit density was achieved with a single silicon tip, the depressions with a diameter

of approximately 10 nanometers. In order to increase the data rate, i.e. the write and read speed, not only a tip is used, but a whole matrix of levers that work in parallel. The new version has more than 4,000 such tips, which are arranged in a small square with a side length of 6.4 millimeters. These dimensions enable a complete, high-capacity storage system to be packaged in the smallest standardized format for flash memory.

A cantilever in **Millipede** writes and reads from a cell that is around 100 by 100 micrometers in size. For example, while the read and write head and the storage medium move on hard disks, only the medium moves on the Millipede memory. Two tiny coils, which are placed between magnets, drive the movement of the plate: The microscanner can be positioned with an accuracy of up to 2 nanometers. The position can be precisely determined at any time from the overlapping area of strip-shaped sensors.

Cantilever array

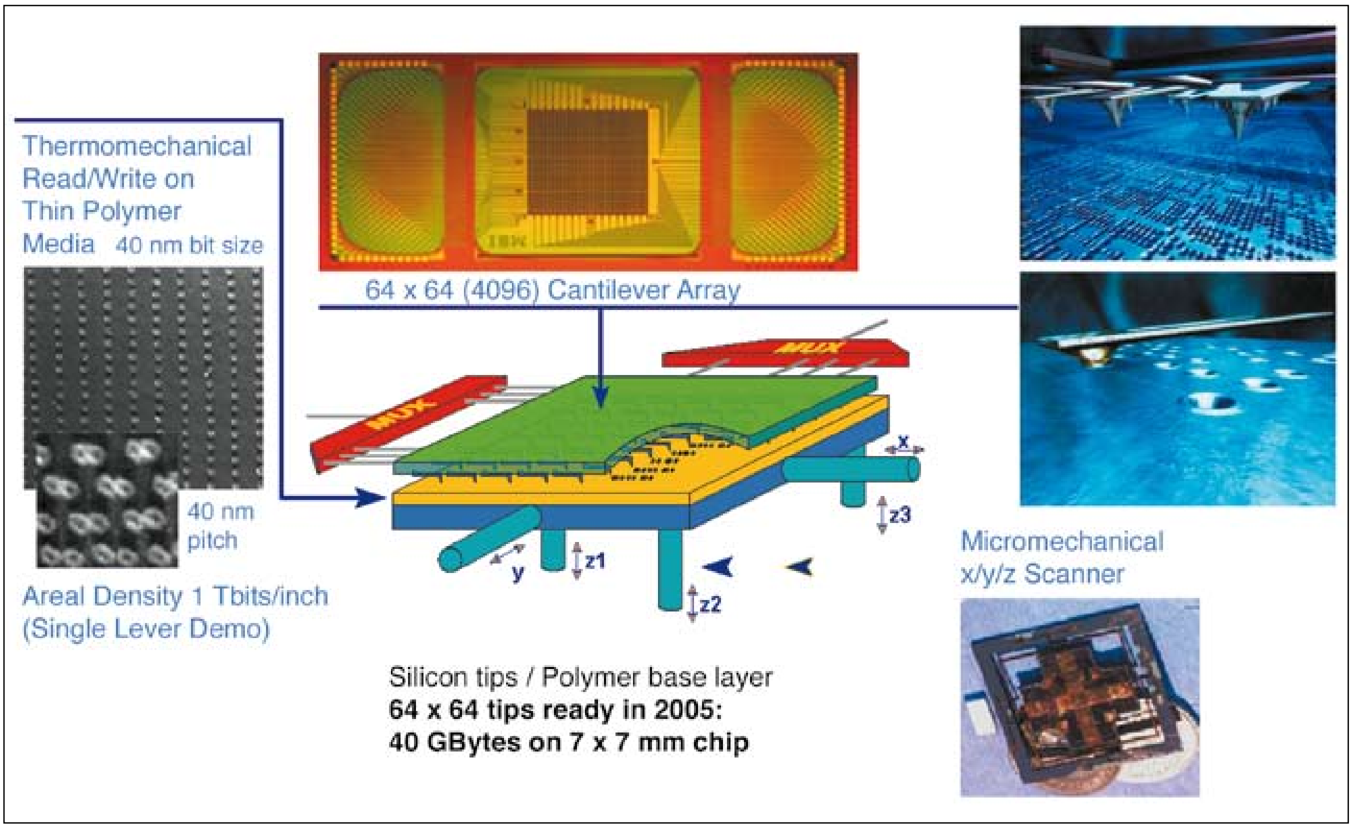
The core of Millipede technology is a two-dimensional arrangement of v-shaped silicon spring tongues (cantilever) that are 70 micrometers (thousandths of a millimeter) long. At the end

Each "tongue" has a sensor for reading and a resistor above the tip for writing. The tip is almost 1 micron long and the radius is only a few nanometers. The cells with the cantilevers are arranged on a cantilever array chip. The chip is 7 x 14 mm. At the center is the array, which currently consists of a total of 4,096 (64 x 64) cantilevers etched out of the silicon. The actual data carrier consists of a polymer film just a few nanometers thick on a silicon substrate. Controlled individually via multiplexers, the heads read, write or delete the desired bit. Up to 100,000 write and overwrite cycles have been successfully tested so far. Although mechanical engineering is used in the design,

Microscanner

The movement of the storage medium relative to the cantilever array is realized using the silicon-based x / y micro scanner. The scanner consists of approx.

6.8 x 6.8 mm² scan table that carries the polymer medium and two electromagnetic triggers. The scanner chip is mounted on the silicon plate, which serves as the mechanical base of the system. The distance between his



The surface and the surface of moving parts of the scanner is approx. 20 µm. The scan table can be moved by 120 µm in the x and y directions. Each trigger consists of two permanent magnets built into the silicon plate and a small coil located between the magnets. In order to extinguish the vibrations from the outside, a so-called pivot is used, which is coupled to the triggers.

If the tip "falls" into a recess, the reading sensor cools down slightly due to the shorter distance to the substrate, but this leads to a measurable change in the resistance. In order to overwrite data, the tip places depressions in the surface. Their outer edges overlap the old depressions and thus delete the old data. More than 100,000 write and overwrite cycles have proven that the concept is suitable for a rewritable memory type.

Position determination

The positioning information is provided by four thermal sensors. These sensors are located directly above the scan table, on the cantilever array. The sensors have thermally insulated heaters. Each sensor is positioned over an edge of the scan table and is heated by electricity. Part of this heat is conducted through the air to the scan table, which now serves as a cooler. Moving the scan table causes a change in the efficiency of this cooling system, which changes the temperature of the electrical resistance from the heater.

A sophisticated design ensures that the tips are precisely leveled over the storage medium and dampens vibrations and impacts from the outside. Time division multiplexing electronics, as used in a similar way in memory chips (DRAM), enables the addressing of each individual tip in parallel operation. Electromagnetic actuation moves the substrate with the storage medium on its surface very precisely in the x and y direction, so that each tip can read and write in its storage field with a side length of 100 micrometers. The short distances make a significant contribution to low energy consumption.

For the functions of the device, ie reading, writing, erasing and overwriting, the tips are brought into contact with the polymer film on the silicon substrate.

Writing technology

Bits are written by heating the resistor integrated in the cantilever to typically 400 degrees Celsius. The tip, which is also heated as a result, softens the polymer, sinks in and leaves a depression of a few nanometers. For reading, the cantilever's reading sensor is heated without softening the polymer film.

In order to get the data into and out of the memory faster, a complete matrix arrangement of levers processes the medium simultaneously. However, it turned out that it is extremely difficult to manufacture mechanics and electronics in one piece on one chip. So the scientists decided to build it in two parts:

1. ***The matrix of the micro tips is provided with tiny contact pins, which look like the knobs of Lego bricks under the electron microscope.***
2. ***These studs are then contacted with their counterparts on the electronic board.***

The prototypes shown have recently demonstrated the technical feasibility of a product, for example in terms of storage density, performance and reliability. While the storage technologies used today are gradually reaching fundamental limits, the nanomechanical approach has enormous development potential for a storage density that is a thousand times higher. This nanomechanical data carrier generates almost no heat, consumes little electricity and is completely shock-resistant.

Millipede project phases

Chips were developed in three generations throughout the project. The last chip completed in 2007 works with 128 x 128 tips, has a capacitive storage capacity of 160 Gbytes and a data rate of 120 Mbit / s. IBM is currently considering how and how MillipedeChips can be used sensibly. So it remains exciting to see how nano-technologies and nano-mechanics will find their way into the world of data storage.

Revolutionary storage technology from IBM research is getting closer to being realized

"Racetrack Memory" opens up new dimensions in storage density, speed and reliability

April 11, 2008: IBM researchers report in the renowned science journal Science about a breakthrough in the research of a new storage technology, which has become known under the name "Racetrack" (German: race track). The technology could create a new class of storage that combines the high performance of flash storage with the large capacity and low cost of hard disk storage.

In the current issue of Science, IBM fellow Stuart Parkin and his colleagues from the IBM Almaden Research Center, USA, describe an important step in the development of memory chips based on the "race track process". Within the next ten years, the technology could be used and open up new dimensions of storage capacities, enable extremely fast switching on of electronic devices and be much cheaper, more energy efficient and more robust than today's systems.

It was thanks to Stuart Parkin that the GMR effect (plate technology, see GMR) was converted into a product in the form of a GMR reading head, and it took seven years to do so. It is thanks to him and his team that the high hard disk capacities were possible today. The author does not agree that Stuart Parkin was disregarded when the Nobel Prize for Physics was awarded in 2007, especially since the Nobel Prize could have been awarded to three people. With Racetrack Memories he will write real history in the memory area! The author is convinced of that!

Universal storage of the future

The revolutionary concept is based on the storage of information in the form of tiny, oppositely magnetized regions (domains) in a nanowire. In the conventional hard disk used as a storage medium, the medium and a read / write head are moved to read, write or erase data. It is different with the "racetrack process": Here the magnetic domains are shifted towards the central reading and writing units, which are attached in the middle of the nanowire - and at extremely high speed. The stored data bits seem to "race" through the data conductor, hence the name "Racetrack".

Stuart Parkin, IBM Fellow and one of the largest memory researchers and developers at IBM

Since a single 'racetrack' is only a few nanometers in size and can store between 10 and 100 bits, the technology allows extremely high storage densities. Compared to flash memories, a 'racetrack memory' could record 100 times more data in the same area. That would correspond to around 500,000 music tracks or 3,500 films. Due to the minimal power consumption of a 'Racetrack memory', an MP3 device could also be operated with a single battery for weeks.

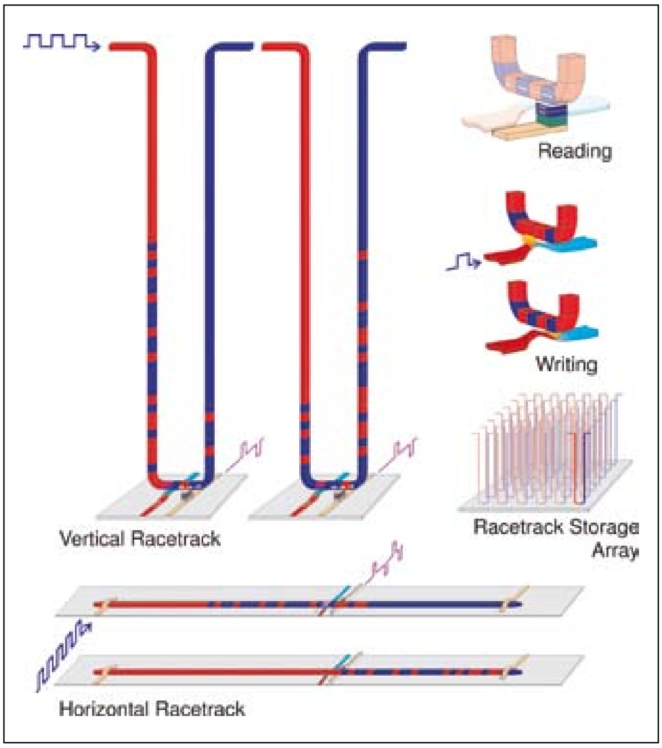
In addition, the process does not require any moving parts, which means that there are almost no signs of wear or tear. This makes the 'Racetrack storage' more resistant than all existing storage technologies and gives it an almost unlimited lifespan.

“The connection between the research of new physical phenomena that occur at the molecular and atomic level and nano-engineering - the ability to selectively structure materials from individual atoms, since a single“ racetrack ”is only a few nanometers in size and between 10 and 100 Can store bits, the technology allows extremely high storage densities. Compared to flash memories, a “racetrack memory” could record 100 times more data on the same area. That would correspond to around 500,000 music tracks or 3,500 films. Due to the minimal power consumption of a “Racetrack memory”, an MP3 device could also be operated with a single battery for weeks.

In addition, the process does not require any moving parts, which means that there are almost no signs of wear or tear. This makes the "Racetrack storage" more resistant than all existing storage technologies and gives it an almost unlimited lifespan.

“The connection between the research of new physical phenomena that occur at the molecular and atomic level and nano-engineering - the ability to create material structures from individual atoms and molecules - always poses new and exciting challenges,” explains Stuart Parkin. He continues:

"Exploiting the potential of our" racetrack technology "could lead to fascinating new applications - applications that are still unimaginable today."



Schematic representation of racetrack memory storage

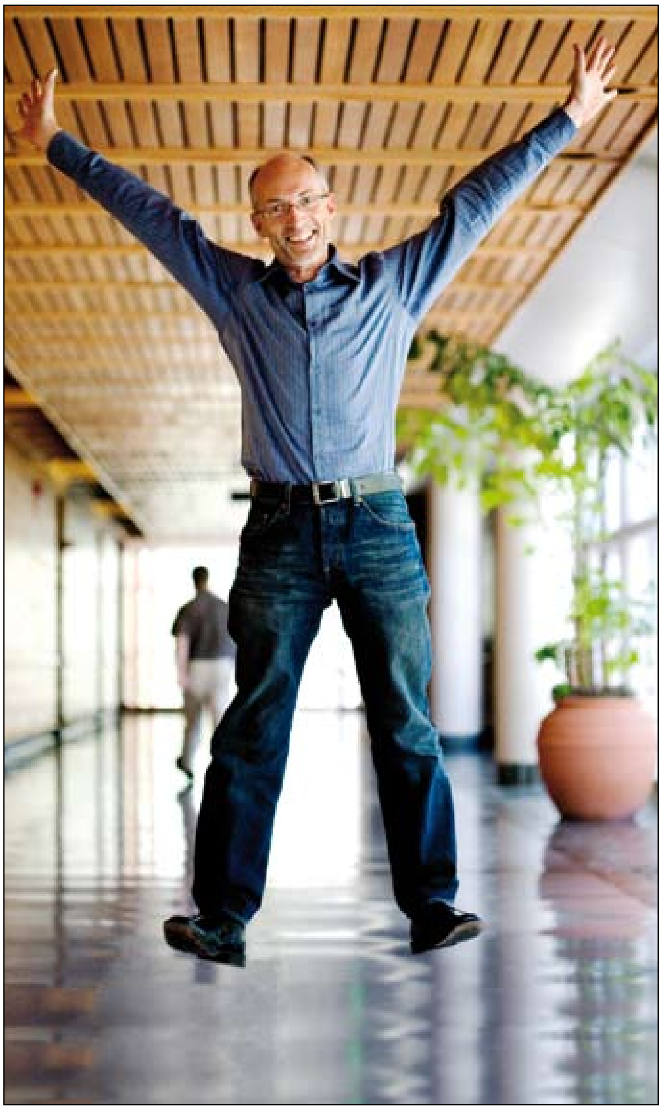
Radical technical innovations that started in basic research have already opened up new markets and fields of application several times in the history of IBM. These include the memory chip, the relational database or the hard disk.

The Racetrack Method

For around 50 years, scientists have been trying to store digital information through so-called magnetic domain walls - the interfaces between two oppositely magnetized regions in a magnetic material. Until now, domain walls could only be manipulated using costly, complex and energy-intensive processes.

In their recently published work “CurrentControlled Magnetic DomainWall Nanowire Shift Register”, Parkin and his teammates demonstrate a new, very efficient method. Here, the domain walls are moved by a spin polarized current, which can rotate the magnetization in the wall. The reason for this is the spin transfer effect, which considerably simplifies the storage technology because the current flows directly through the domain walls and no additional magnetic fields have to be generated.

In the second work "Magnetic DomainWall Racetrack Memory", the researchers summarize the basic principles of "Racetrack Technology". This is based on the storage of data in tiny magnetic regions within a data conductor. The researchers use small nanowires as data conductors, which are magnetically described with information. These are attached lengthwise horizontally or in the form of a loop vertically on a silicon surface. When a voltage is applied, neither the wires nor the silicon move, but only the magnetic domain walls in which the information is stored.



Stuart Parkin is right to be happy, he did it and Racetrack can change the world

Each magnetic domain that represents a bit has a "head", a magnetic north pole, and a "tail", a magnetic south pole. The domain walls that form at the interfaces alternate between "HeadtoHead" and "TailtoTail" configurations in the data conductor. The distances between two successive domain walls are equal to the bit length and are determined by markings along the nanowire, so-called pinning centers.

In their experiment, the researchers used nanowires made of permalloy (an FeNi alloy) and showed how they can create, move and read out domain walls by using current pulses of precisely controlled duration in the nanosecond range. The write and shift cycle only takes a few 10 nanoseconds.

The goal is to arrange many thousands of these nanowire memories, which can store between 10 and 100 bits, densely on one surface. The vertical arrangement of nano wires also opens up new types of three-dimensional architecture - a paradigm shift compared to the current two-dimensional chip and memory technologies. The exploitation of the third dimension creates new scope for performance increases and cost reductions that are no longer only achieved through miniaturization - as has been the case until now.

IUPAP Magnetism Award and Louis Neel Medal

At the **July 27, 2009** receives IBM Fellow Dr. Stuart Parkin one of the highest honors in the field of magnetism research. He will receive the "International Union of Pure and Applied Physics Magnetism Award (at the" International Conference on Magnetism "in Karlsruhe **IUPAP)** and the **Louis Neel medal** for his fundamental research and development results in the field of nano-magnetism. The IUPAP Magnetism Committee only awards this award every three years. Most of all, he was honored for his latest research on Racetrack Memories. The Louis Neel Medal goes back to the scientist Louis Eugene Felix Neel, who received the Nobel Prize in Physics in 1970 for the discovery of antiferromagnetism.

Nanotechnology - key technology of the 21st century

Nanotechnology is a new technology that uses functions on an extremely small scale. It focuses on structures and processes in dimensions under 100 nanometers - about 400 times thinner than a human hair. In the nanometer size range, many fundamental processes in biology, chemistry and physics take place, which can be controlled to an unprecedented degree and open up new perspectives in many areas. These include highly developed functional materials, nanoelectronics, information and communication technology, sensors, life sciences and energy technology. Nanotech applications could contribute to the more efficient use of solar energy or to new types of water treatment.

Through research at the ETH Zurich (Swiss Federal Institute of Technology) and the IBM research laboratory in Rüschlikon ZRL (Zurich Research Laboratory), Zurich has repeatedly provided decisive impulses in quantum mechanics and nano research. These include, for example, the groundbreaking concepts of quantum mechanics by ETH physicist and Nobel laureate Wolfgang Pauli or the development of the scanning tunneling microscope (RTM) by Gerd Binnig and Heinrich Rohrer at the IBM research laboratory in Rüschlikon. For this, the researchers received the Nobel Prize in Physics in 1986. This device gave the first glimpse into the world of atoms. Shortly thereafter, it became possible to manipulate individual atoms, which opened the door to nanotechnology widely has been. The RTM is generally viewed as the instrument that opened the door to the nanocosmos. The atomic force microscope (RKM), which is closely related to the RTM, was invented in 1986 by Gerd Binnig.

The scanning tunneling microscope (RTM) is called STM (Scanning Tunneling Microscope), the atomic force microscope (RKM) is called AFM (Atomic Force Microscope).

IBM and ETH Zurich found a new joint nanotech research center

The new Nanotech research center from IBM and ETH Zurich covers around 6,000 square meters. The two institutions share a long tradition of scientific cooperation. At a joint media conference in Zurich, Ralph Eichler, President of ETH Zurich, and John E. Kelly, Senior Vice President and Director of Research at IBM, announced the planned collaboration. The central element of this collaboration is a new building that is being built on the premises of the ZRL in Rüschlikon. The foundation stone will be laid in spring 2009.

The new research center on the IBM campus in Rüschlikon will enable nano research at the highest level. The "Nanoscale Exploratory Technology Laboratory" is part of an important strategic partnership in nanotechnology with the Swiss Federal Institute of Technology Zurich (ETHZ). Research activities will start in 2011.

IBM research location Rüschlikon: new laboratory for nano-technology

The research focus of the two institutions ranges from basic research to applied research. Research is carried out jointly in various areas, such as carbon-based materials, NanoPhotonics, Spintronics, nanowires and tribology.

New breakthrough in nanotechnology - IBM researchers measure the charge state of individual atoms with the atomic force microscope

In collaboration with scientists from the universities of Regensburg and Utrecht, IBM researchers achieved a breakthrough in the field of nanotechnology in June 2009. For the first time, they were able to measure the charge state of individual atoms directly using atomic force microscopy (RKM). The precision with which they were able to distinguish between uncharged or positively or negatively charged atoms was a single electron charge with a spatial resolution that was accurate to the nanometer. This opens up new possibilities for the research of nanostructures and building blocks on an atomic and molecular scale in application areas such as molecular electronics, catalysis or photovoltaics.

For their experiments, the researchers used a combination of a scanning tunnel (RTM) and an atomic force microscope (RKM), operated in ultra-high vacuum and at low temperatures (5 Kelvin) in order to achieve the stability required for the measurements.

An RKM uses an atomically fine tip attached to a vibrating cantilever to measure the forces that occur between this tip and the atoms on the substrate. In the present work, the researchers used a so-called qPlus force sensor, in which the tip is attached to a prong of a tuning fork, as found in mechanical watch movements, while the other prong is fixed. The tuning fork is mechanically excited and vibrates with an amplitude of 0.02 nanometers. This corresponds to only about a tenth of the diameter of an atom. Becomes

If the RKM tip is placed very close to the sample, for example over a single atom, the resonance frequency of the tuning fork changes due to the forces that occur between the sample and tip.

With this method and under extremely stable conditions, the IBM researchers were now able to measure the minimal differences in the force that exist between the tip and individual, differently charged atoms. The force difference between a neutral gold atom and a gold atom with an additional electron is only about 11 picoNewtons, measured with a minimum distance between tip and sample of about half a nanometer. The measurement accuracy of these experiments is in the range of 1 picoNewton, which corresponds to the gravitational force that two people exert on each other at a distance of more than half a kilometer. The researchers also determined how the force changed with the voltage applied between the tip and the sample. This allowed the distinction

This breakthrough is another important advance in the field of nano research. In contrast to the RTM, which relies on electrically conductive samples, the RKM can also be used for non-conductive samples. In molecular electronics, in which the use of molecules as functional building blocks in future circuits and processors will be researched, non-conductive carrier substances (substrates) are required. Therefore, atomic force microscopy would be preferred in such experiments.

"The atomic force microscope with an accuracy of one electron charge is ideally suited to investigate the charge transfer in molecular complexes, which could provide us with valuable, new knowledge and physical foundations and one day could also lead to new components in information technology," explains Gerhard Meyer, who leads the research in the area of scanning tunnels and atomic force microscopy at the IBM research laboratory in Zurich. Computer devices on the molecular scale have the potential to be orders of magnitude smaller, faster and also more energy efficient than today's transistors and memory devices.

For future experiments, the researchers can imagine making connections from individual metallic atoms and molecules, charging them with electrons, and measuring their distribution directly with the RKM. IBM researcher Leo Gross also points out the importance of her work for other areas: “Charge status and charge distribution are critical factors in catalysis and in the conversion of light into electrical energy. Mapping the charge distribution on an atomic scale could lead to a better understanding of the basic processes in these areas. ”

The latest results follow a series of fundamental scientific breakthroughs that IBM researchers have achieved in recent years: In 2008, scientists at the IBM Almaden Research Center in California were able to use a qPlusRKM to measure the force required to create a force for the first time Moving atom on a surface. This paved the way for the current experiments. In 2007, Gerhard Meyer's team at the IBM Research Laboratory in Zurich demonstrated a molecule that could be switched between two states in a controlled manner without changing its external shape. As early as 2004, the same group had successfully carried out an experiment in which they could manipulate the charge state of a single gold atom with an RTM. By applying voltage pulses to the RTM tip, the researchers were able to apply an additional electron to a single atom that was on an insulating substrate. This negatively charged atom remained stable until a corresponding voltage pulse with the opposite sign occurred using the RTM tip. The researchers also used this method in their current experiments to charge the individual atoms.

IBM researchers are showing the inner structure of molecules with atomic resolution for the first time

IBM researchers in the IBM ZRL Rüschlikon succeeded in August 2009, for the first time atomically resolving the complete chemical structure of a molecule using an atomic force microscope. This enabled the scientists to make a breakthrough in the field of nanosciences, which could also open up new opportunities for research into new types of electronic components on the atomic and molecular scale.

In recent years, astounding progress has been made in characterizing nanostructures on the atomic scale using atomic force microscopy. Up to now, however, science has been prevented from looking at the internal structure of a molecule with atomic resolution.

In the edition of the science journal Science dated
On August 28, 2009, IBM researchers Leo Gross, Fabian Mohn, Nikolaj Moll and Gerhard Meyer and Peter Liljeroth from the University of Utrecht reported how they were able to use an RKM to map the complete chemical structure of pentacene molecules ($C22H14$) with atomic resolution The experiments were carried out in an ultra-high vacuum and at very low temperatures (5 Kelvin or -268 ° C). The images obtained are somewhat similar to X-rays, which allow a look inside the human body. The “RKM with X-ray vision” from the IBM researchers can look through the electron cloud that envelops the molecule and map the atomic backbone of a pentacene.

This publication follows just two months after another work by the same group, published on June 12, 2009 in Science. The charge measurement of individual atoms using RKM was described in this. "Scanning probe technologies offer incomparable potential for building prototypes of complex functional structures on the atomic scale and thus tailoring and examining their electronic and chemical properties in a targeted manner," explains Gerhard Meyer, who is responsible for research in the field of scanning tunnel microscopy and atomic force microscopy at IBM in Rüschlikon directs.

The results of the two research projects open up new possibilities for studying the charge distribution in specific molecules or molecular networks. Knowledge of these processes is generally very important for the development of electronic components on the atomic and molecular scale. Such components could enable faster, more powerful and more energy-efficient processors and memory chips in the future if the limits of today's chip technologies are pushed to the limit.

The RKM uses a tiny, very sharp metal tip to measure the minimum forces that occur when this tip is brought very close to a sample, such as a molecule. With a point-by-point measurement of these forces, an image of the surface can be created. The pentacene molecules that the IBM researchers used in this work consist of 22 carbon atoms and 14 hydrogen atoms and are only 1.4 nanometers (millionths of a millimeter) long. The distances between the carbon atoms arranged in hexagons are around ten times smaller. These hexagonal carbon rings are surprisingly clearly resolved on the RKM images and even the positions of the hydrogen atoms can be clearly identified.

IBM researcher Leo Gross emphasizes: "The decisive factors for the resolution were an atomically sharp tip with a defined structure and a very high stability of the overall system." In order for the chemical structure of a molecule to be visible, the tip must be extremely close - less than one Nanometer - to introduce the molecule. Forces occur only in this area, which are largely determined by chemical interaction. In order to achieve this, the researchers had to improve the sensitivity of the tip and overcome a major hurdle: similar to two magnets lying next to each other that attract or repel each other, the molecule shifts or attaches itself to the tip if it comes too close. If this happens, no further measurements can be carried out.

Both problems could be solved by choosing a suitable atom or molecule at the RKM tip. “In order to sharpen our tip, we specifically added atoms and molecules to the tip using manipulation techniques. Our measurements with differently prepared tips clearly show that the foremost atom or molecule of the tip has a decisive influence on the resolution, ”states Leo Gross. A carbon monoxide molecule (CO) at the top brought about the breakthrough and, at a distance of about half a nanometer from the pentazene, ensured optimal resolution of the individual atomic positions and their chemical compounds.

The scientists also managed to create a complete three-dimensional, topographical representation of the forces over the molecule. “The data collection took more than 20 hours. This placed the highest demands on the thermal and mechanical stability of our system in order to ensure that the tip and the molecule remained unchanged throughout the entire time, ”describes Fabian Mohn, PhD student at IBM Research in Rüschlikon.

Theoretical calculations carried out by IBM researcher Nikolaj Moll confirmed the experimental results and provided information on the exact nature of the imaging mechanism. Nikolaj Moll explains: "The calculations show that the so-called Pauli repulsion between the C molecule at the tip and the pentacene is responsible for the atomic contrast." This repulsive force is due to a quantum mechanical effect that prevents two identical ones from being formed Too close electrons.

For readers who want to know more about the two experiments and research work, the scientific work of L. Gross, F. Mohn, N. Moll, P. Liljeroth and G. Meyer with the title “The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy ”, published in Science, Vol. 325, No. 5944, pp. 1110-1114 (August 28, 2009) is recommended.

The author is convinced that nanotechnologies will find their way into memory technology to a large extent in the next few years and enable memory chips that go beyond the limits of today's chip technologies!

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

CCW (Continuous Composite WORM) is based on the MO technique. The plate is given a label during production so that it cannot be overwritten. In addition, software in the drive protects the disk from being overwritten. CCW is therefore a “bill” WORM plate that is based on MOD technology.

In the IT environment, the technologies were standardized on a 120 mm (5.25 inch) basis as MO, CCW and WORM plates in the early 1990s. The 120mm technology is still used in the IT environment for archiving purposes. Optical plates based on 90mm (3.25 inches) are primarily used in the consumer environment.

UDO (Ultra Density Optical Disk)

In 2003 and 2004, the 120mm technology, which had previously worked with a red laser, was replaced by a new standard, UDO technology (Ultra Density Optical). The UDO technology is not backwards compatible with the old 120mm technology and works with a blue-violet laser with a 405nm wavelength. The plate formats WORM, MO and CCW have been retained. UDO offers a "Phase Change" WORM and "Phase Change" rewritable medium for use. The first plates on this new laser base as the new 1x standard offered a capacity of 30 GB. The second generation as a 2x standard, which became available in 2006, offered twice the capacity, i.e. 60 GB. The UDO technology as the new 120mm standard will be further developed according to a roadmap that has been anchored in the UDO standard.

UDO roadmap

	Generation 1	Generation 2	Generation 3
Capacity	30 GB	60 GB	120 GB
Transfer rate	8 MB / s	12 MB / s	18 MB / s
RPM	2,000	3,000	3,600
Avg Seek Time	25 ms	25 ms	25 ms
Numerical aperture	0.7	0.7	0.85
Media layers	1	2nd	2nd
Encoding	1.7	1.7	ML
Sector size	8 KB	8 KB	8 KB
SCSI transfer rate	80 MB / s	80 MB / s	80 MB / s
Load time	5 s	5 s	5 s
Unload time	3 s	3 s	3 s
MSBF	750,000	750,000	750,000

DVD (Digital Versatile Disk)

DVD stands for "Digital Versatile Disk" (formerly "Digital Video Disk"). The medium is as large as a normal CDROM, but a much higher storage density is used. There are four different media. The simple single-sided DVD can store 4.7 GB on one layer. But there are also two-layer DVDs. The information is stored on two superimposed layers, one of which is transparent.

The right layer is controlled by focusing the laser differently. This makes 8.5 GB possible. And then there is the whole thing on two sides. This enables 17 GB of data on a single DVD. The first drives came on the market in the 90s and can read a layered, single-sided DVD. Unfortunately there were still few DVD titles with videos at this time. The videos were encoded in MPEG2, which results in very good quality when played back. The first burners for single-sided, single-layer DVDs had already been introduced in the 1990s; the Pioneer burner was offered in autumn 1997 for a little over DM 10,000. It was recorded at approx. 1 - 2 MB / s and could save a maximum of 3.9 GB.

The readers can also read normal CDs, but mostly not CDRs, i.e. the writable CDs. This is because a laser with a shorter wavelength is used, which can no longer read the self-burned CDs correctly. For this purpose, Sony had developed a drive with two laser diodes, which can then be used to read the CDRs again.

HD DVD

The HD DVD (High Density Digital Versatile Disk, previously: Advanced Optical Disk, short: AOD) is a data carrier format and was traded from 2005 to February 2008 alongside the Bluray Disc and VMD as a possible successor format to the DVD.

The HD DVD was specified by the Advanced Optical Disc Consortium (AOD), which included NEC, Microsoft, Toshiba, Intel, IBM and HewlettPackard. Afterwards, these companies merged to form the HD DVD Promotion Group in order to make the HD DVD better known. In November 2003, the DVDForum had adopted the HD DVD as the HD successor to the DVD in accordance with the HDDVD specifications for ReadOnlyDiscs. The Advanced Access Content System (AACS) copy protection from the area of digital rights management (DRM) was provided for the HD DVD.

In March 2006, Toshiba's HDXA1, the first HDDVD player, was launched in Japan. In August 2006, Elephants Dream, the first HD DVD in German, was released.

In February 2008, Toshiba announced that the development, manufacture and marketing of the technology would cease at the end of March 2008.

The HD DVD was based on a blue-violet laser with a wavelength of 405 nm. The thickness of the carrier layer is with 0.6 mm identical to that of the DVD. In contrast, the numerical aperture (NA) is 0.65 compared to 0.6 for the DVD and 0.85 for the Bluray disc.

The HD DVD had one layer and a storage capacity of:

- ***15 GB (for HD DVD-ROMs - pressed media) or***
- ***15 GB (for HD DVD-R / RWs - once and rewritable bare media) or***
- ***20 GB (with HD DVD-RAM - rewritable media with optional sector access)***

and with two layers a storage capacity of

- ***30 GB (for HD DVD-ROMs - pressed media).***

In addition, a three-layer version was approved by the DVDForum in August 2007, with space for 17 GB per shift and thus having a total capacity of 51 GB. The definitive standard for films on HD DVD initially included the 15GB and the 30GB variant, whereby the two-layer 30GB variant was almost always used for films.

In January 2008, Time Warner announced that its Warner Bros. and New Line Cinema studios would not release any more films for the HD DVD in the future, but would only use the Bluray disc. This has been described by some media as a decision of the format war. Over the next few days, other providers followed, such as the large European film distribution company Constantin Film on January 10, the US porn provider Digital Playground on January 12 and the German Senator Film. The US providers Universal Studios and Paramount Pictures, however, denied rumors of abandoning their contracts with HD DVD, they are considered the last remaining large providers. After Toshiba, of which a large part of the HD DVD players came, had drastically reduced the prices of the devices in the USA on January 15, some media spoke of a “sellout”. On the other hand, the pressure on the HD DVD continued to increase: at the beginning of 2008, MediaSaturnHolding stores took the HDDVD player in payment when buying a Bluray player. The format suffered further setbacks in February 2008. The USE electronic chain Best Buy announced on February 12, 2008 that it would only concentrate on the Bluray format in the future. A day earlier, the largest US online video rental company Netflix announced that it would remove the HD DVD from its range by the end of the year. On February 15, 2008, WalMart, the largest retailer in the USA, announced that it would sell out HDDVD inventory and would only want to rely on Bluray in the future. On February 19, 2008, Toshiba issued a press release announcing that the development, Production and sales of the HD DVD and corresponding devices will no longer be pursued and will be discontinued at the end of March 2008. As a result, the same day, the Universal Studios announced that they would switch from the HD DVD to the Bluray format.

Blu-ray disc

The Bluray Disc (abbreviated BD or more rarely BRD) is a digital optical storage medium. In addition to HD DVD and VMD, it was advertised as a possible successor to the DVD. After Toshiba announced in February 2008 that production and further development of HDDVD technology - including the devices - would be discontinued in March 2008, the Bluray Disc was the winner in the format war. The name Bluray is of English origin. Blue ray literally means “blue light beam”, which refers to the blue laser used. The deliberate deviation from the orthographically correct spelling "blue ray disc" aims to encourage registration of the printout as a brand.

The Bluray Disc specifications were approved on February 19, 2002 by the nine companies of the Bluray Group, Panasonic, Pioneer, Philips, Sony, Thomson, LG Electronics, Hitachi, Sharp and Samsung; This group also joined Dell and HewlettPackard at the end of January 2004 and Apple in mid-March 2005. However, HewlettPackard resigned from the Bluray consortium in 2005 after some suggestions for improvement were rejected and switched to the HDDVD camp.

The Bluray Disc is available in three versions: as a readable BDROM (comparable to DVDROM), as a write-once variant BDR (comparable with DVD ± R) and as a rewritable BDRE (comparable with DVD ± RW). By using a blue-violet laser and its short wavelength (405 nanometers), higher data densities can be achieved than with a DVD. 25 GB of data can be stored on a BDROM per data layer, so that a capacity of up to 50 GB results with two-layer media.

A four-layer version of the Bluray Disc, which is said to have a capacity of 100 GB on one side, was presented by TDK. TDK has managed to accommodate 200 GB on a six-layer disc. The capacity of a layer was increased to 33.3 GB. However, these are primarily feasibility studies. Writable media are also planned from the outset, as the Bluray Disc is primarily intended for entertainment electronics and serves as a storage medium for high-resolution films. The rewritable Bluray Disc works with the phase change technique.

The new PhaseChange technology should enable a double transfer rate (data transfer rate) of 9.0 MB / s (72 Mbit / s) when writing instead of the originally maximum specified simple 4.5 MB / s (36 Mbit / s). An important part of the specification is copy protection in the form of a unique identification number. Bluray discs are also particularly suitable for Full HD videos, which, thanks to the high resolution, offer better quality than conventional systems such as PAL and NTSC, but also require more storage space accordingly. BDS players generally support the high-definition AVCHD format, which was first introduced by Pana sonic and Sony.

Another innovation compared to the DVD is the reduced distance of the laser to the data carrier and the shorter wavelength (and therefore different color) of the laser beam. Furthermore, the protective layer on the laser side is significantly smaller at 0.1 mm compared to 0.6 mm on the DVD. Due to the resulting greater susceptibility to scratches and dust, it was initially planned to offer the Bluray Disc only in one cartridge. Instead, a coating called "Durabis" was developed, which makes the use of a cartridge no longer necessary.

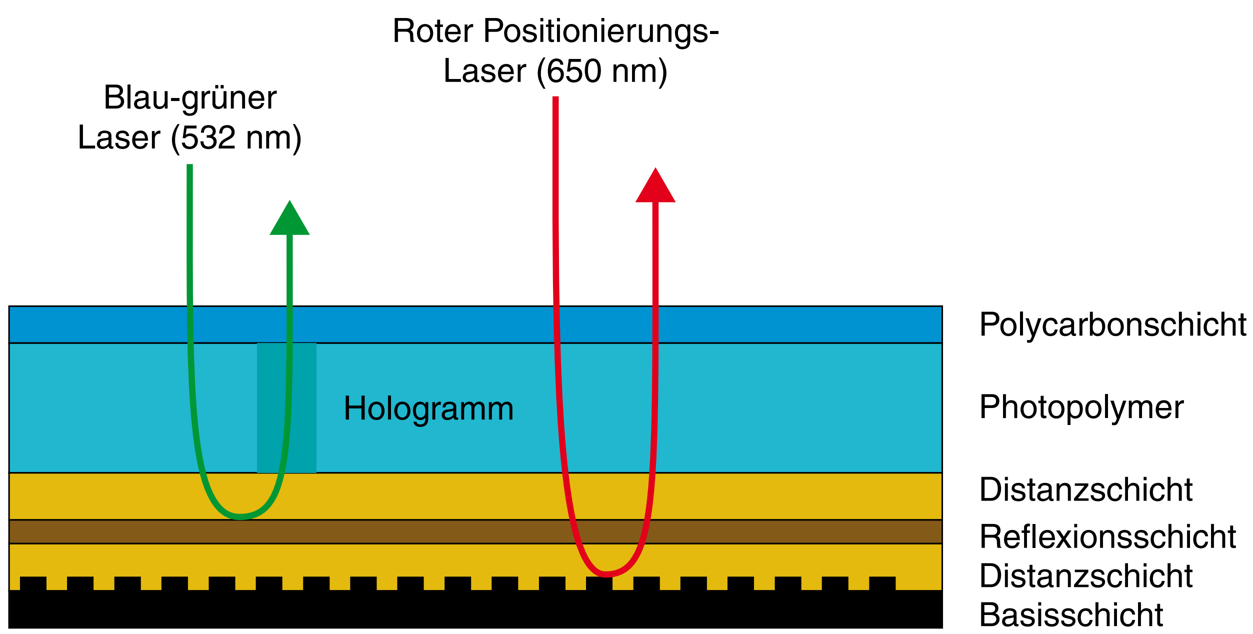
Because of the smaller distance between the medium and laser optics and the thinner protective layer, a lens with a more favorable numerical aperture can be used, which can focus the beam more efficiently. Thus, who reduces the typing errors and greater scatter and it is possible to combine a Bluray disc made of metal or other stable, opaque materials to build with a thin transparent support layer, which can be operated at significantly higher speeds than a disk made of polycarbonate, which then results in higher transmission rates. In addition, the smaller wavelength of the laser beam compared to the DVD allows a much higher data density and thus an increased storage capacity.

Verbatim presented a "Bluray Disc Rewriteable" at IFA 2005. According to the manufacturer, this medium should allow a storage capacity of 25 gigabytes (i.e. approx. 135 minutes of video in MPEG2 HD quality). In addition, the BDRE should be provided with a particularly impact and scratch-resistant protective layer. In addition to the immediate availability of 25GB blanks, TDK has also announced 50GB discs for October 2006. TDK is one of the founding members of the Bluray Disc Association (BDA) and was instrumental in the development.

Holographic Versatile Disc (HVD)

The holographic versatile disc, also known as HVD, is already traded in specialist circles as the successor to today's HDDVD and Bluray technology. HVDs have a capacity of up to 3.9 TB, which is many times more than the largest 200 GB Bluray disk available today. The transfer rate is also many times higher and reaches 1 Gbit / s compared to 36 or 72 Mbit / s with the Bluray plate. In addition, the HVD is far from exhausted and new drives with even higher rotational speeds are conceivable. The specification for the HVD was approved in December 2004 by the TC44 Committee of Ecma International. The "HVD Alliance" was founded in February 2005 to drive the development of HVD. The Alliance now includes 19 companies: Alps Electric Corporation, CMC Magnetics, Dainippon Ink and Chemicals, EMTEC International, Fuji Photo Film, Konica Minolta Holdings, LiteOn Technology, Mitsubishi Kagaku, Nippon Kayaku, Nippon Paint, Optware Corporation, Pulstec Industrial, Shibaura Mechatronics, Software Archi tects, Suruga Seiki, Targray Technology, Teijin Chemicals, Toagosei Company and Tokiwa Optical Corporation.

Two superimposed lasers with different wavelengths, a blue-green write / read laser with 532 nm and a red positioning laser with 650 nm, are reflected on two different layers in the plate and generate an interference field in the photopolymer layer, a so-called hologram. In the schematic drawing, the two lasers are not superimposed to simplify Darge to better illustrate the principle of different reflection.



Prinzip der HVD (Holographic Versatile Disc)

When reading, the blue-green laser reads the hologram of the data encoded as a laser interference pattern in the polymer layer in the upper area of the medium, while the red laser serves to read out auxiliary information from a CD-like aluminum layer in the lower area. The auxiliary information is used for exact positioning where you are on the disk (comparable to sector, head and segment information of a hard disk). A dichroic mirror layer (reflection layer), which is located between the polymer layer and the aluminum layer (base layer), allows the blue-green laser to reflect while the red laser is passing through.

Polymers, i.e. plastics, are long chain molecules that consist of the same building blocks. They have the advantage over crystals that they can be modified almost indefinitely. Polymers are extremely sensitive to light, highly transparent and insensitive to temperature fluctuations. They do not change their performance even after frequent reading and are ideal for laser processing. After exposure to laser light, the photosensitive molecules in the polymer change their orientation. In the exposed areas, the material deflects light more than in the unexposed areas. The “holographic” information is contained in this change in the molecular order, which is targeted as an interference field. Since the data is not recorded and read as individual bits, but as whole bit patterns, extremely many bits can be called up at the same time with a single “flash” laser. Depending on the type and thickness of polymer, this can be several hundred thousand bits. The transfer rates are gigantic. A feature film that now fits on a DVD could be read in about 30 seconds. The large storage capacity is due to the fact that the holograms are not only written on the surface of a storage material, but in the entire volume.

Particularly sensitive data could also be encrypted using a special mask that is placed between the storage material. In contrast to information (e.g. on chips) that is encrypted using software, the hardware encryption of holograms cannot be broken in principle. For the same reasons a hologram is also difficult to fake. The shelf life of information in holographic memories is now estimated to be around 50 years.

Flash memory - short story

The development of the flash memory known today took place relatively late in the 90s. The driving force was digital photography. After all, it is practically impossible to install a CD burner in a small digital camera, which burns the pictures taken directly onto the CD. The floppy disk was also unsuitable because it simply saved too little data too slowly. The microdrives developed by IBM at the end of the 1990s, which were an alternative to flash memory cards for a few years, consumed a lot of power in use, so that the battery of a digital camera often had to be charged. In the long run, the microdrives were therefore not a sensible alternative in digital photography.

The first flash memory was launched by Sandisk in 1994. At that time it included 4MB memory. In the meantime, this form of data backup has gone many technological steps further, since you could hardly do anything with 4 MB storage capacity. Nowadays, USB sticks based on this technology are available with several gigabytes.

The term flash memory came about in the Toshiba laboratories, where researcher Shoji Ariizumi suggested this name because deleting the data blocks reminded him of a flash light (in English: "Flash").

Flash memory of a compact flash memory card (photo: SANDISK)

Description and functionality

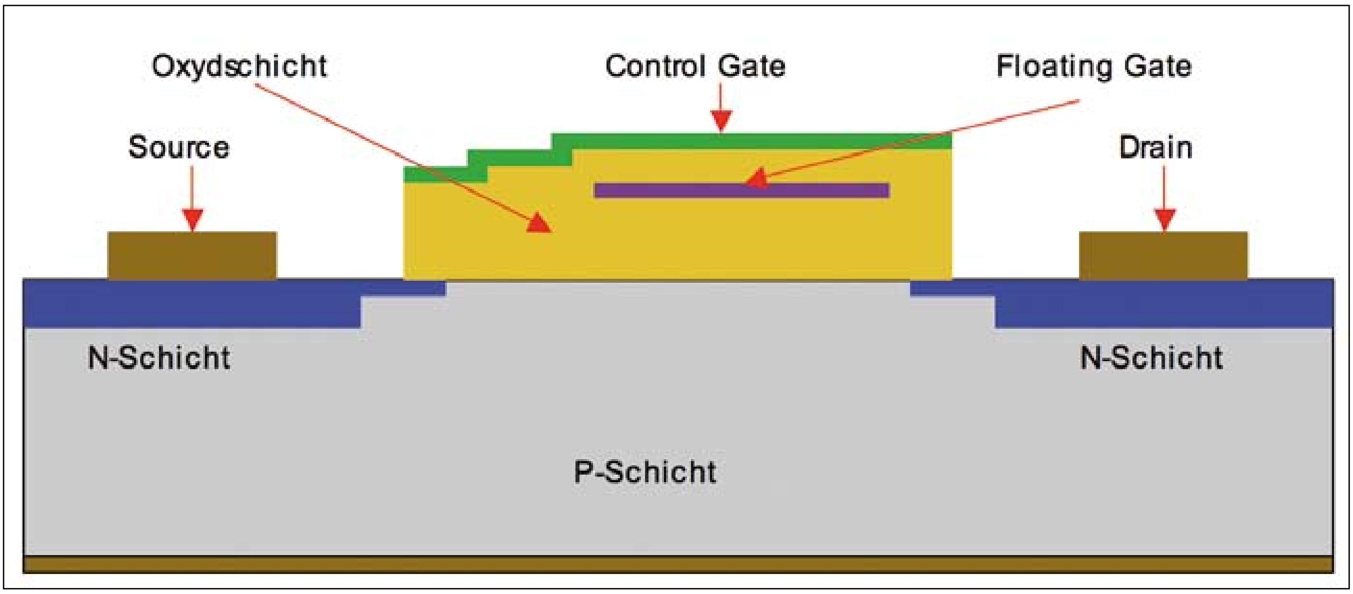
Flash memories are based on semiconductor technology and are found in all memory cards that are common today. They are used in hybrid hard disks or flash hard disks in personal computers, but also in memory cards for work computers or in fast buffers of mainframes; but also above all in SIMMs, PC Cards and PCMCIA cards, in digital cameras and MP3 players, in USB sticks, camcorders and printers, notebooks and handhelds. There are special flash modules for note books, laptops, PDAs, printers and other peripheral devices such as digital cameras. The different versions have a protected designation and spelling: Examples are the CompactFlash card (CF), the Smart Media Card, Memory Sticks, the Multi Media Card (MMC), the SD card and others with storage capacities of several gigabytes (GB). Flash memories can be read in special card readers that are connected to the personal computer (PC) via an interface cable such as USB. In addition, the flash memories can also be operated via PCCard adapters.

In addition to a very compact design, the basic requirements for a memory card are high storage capacities, rewritability, the use of power-independent memory modules and the lowest possible susceptibility to interference in mobile use. These requirements were and are primarily defined by digital photography.

In contrast to hard disks, flash memories no longer have any moving parts integrated, which means that mechanical components of a hard disk are excluded and are taken over by semiconductors, i.e. electrical components, with regard to their functionality. These must be able to store data permanently and non-volatile, i.e. independent of electricity. Even if the power supply is cut off, the information must not be lost.

A flash cell consists of a modified MOSFET transistor (metal oxide field effect transistor), which is equipped with an additional electrode, the so-called “floating gate”. This electrode is placed between the control electrode (control gate) and the active transistor channel and surrounded with a thin semi-permeable separating layer. If free electrons are now "pressed" through the barrier, usually an oxide layer, onto the floating gate using a higher voltage (write voltage), the transistor behavior changes.

Each transistor can either conduct or block an electrical charge in the form of electrons. This corresponds exactly to the binary information states of a bit, ie "0" and "1". The transistor reveals the stored information when a voltage is applied and depending on whether current is flowing or not, it has the logical information "0" or "1".



Construction of a flash memory cell with floating gate

The floating gate is completely electronically isolated from all other components by an oxide layer, so that the stored charge, i.e. the information on the data cell below, is really stored in it and cannot be lost over a long period of time. Under normal circumstances, this isolation would not allow transmission because the isolation would block it. The flash memory therefore uses the so-called tunnel effect from quantum mechanics, which makes it possible to transfer charges even through a non-conductive material. Depending on the amount of electrons absorbed, the floating gate becomes an electronic throttle valve for the transistor function. When reading out a transistor memory cell, the resulting current is evaluated as an indicator for logic "0" (many electrons in the floating gate) and logic "1" (few or almost no electrons in the floating gate). A charged "floating gate" thus reflects "0" and an uncharged one is rated as "1". A different assignment would be possible, but was never used. Each individual bit is thus transmitted via the "floating gate". In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process makes it possible to change this state in the other direction. A charged "floating gate" thus reflects "0" and an uncharged one is rated as "1". A different assignment would be possible, but was never used. Each individual bit is thus transmitted via the "floating gate". In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process makes it possible to change this state in the other direction. A charged "floating gate" thus reflects "0" and an uncharged one is rated as "1". A different assignment would be possible, but was never used. Each individual bit is thus transmitted via the "floating gate". In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process makes it possible to change this state in the other direction. In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process makes it possible to change this state in the other direction. In order to rewrite a data cell that already contains information, it is necessary to delete it beforehand. The reason for this is that the bits in a flash memory can only be changed from 1 to 0. Only a deletion process make

The schematic diagram "Structure of a flash memory cell with a floating gate" shows how the floating gate forms a charge trap in which the electrical charge is stored. It lies in an oxide layer below the control gate and normally prevents the charge from flowing out to the N and PS layers (N = heavily negatively doped electrodes drain and source, P = heavily positively doped substrate). The charge on the floating gate forms a conductive channel between drain and source via its electrical field, via which the charge is read out. The deletion takes place block by block. By applying a negative quenching voltage, the charge carriers are driven out of the floating gate.

Further considerations

Flash memories are current-independent and therefore fundamentally differ from current-dependent (volatile) RAM modules (Random Access Memory), because the information stored in the main RAM of a computer is only kept as long as the computer is running. If the computer is without power, RAM modules lose the ability to save data.

A classic bipolar transistor requires current for several functions. On the one hand, electrons are released from the surface of the transistor to change the switching state. Information is saved or deleted in this way. On the other hand, the current in the form of "refresh" pulses is required to stabilize the state reached.

MOSFETs, on the other hand, are able to use a thin oxide layer to build up a small capacitor that transports charge carriers between the silicon layers via a built-up electrical field. This creates tension until a limit is reached. The result is a thin conductive channel that remains stable without a "refresh" pulse. MOSFETs therefore only need current to charge or discharge the capacitor, while in the other case current must be supplied continuously in order to keep the conductivity in a stable state and to maintain the stored information.

NAND architecture

The name NAND, which describes the NAND flash memory, actually comes from the statement logic. Here, a gate is described that represents a specific AND operation (Not AND = NAND). Circulated in the world of electronics, it means that the memory cells on a flash memory are connected in series, that is, they are arranged in series on different data lines. The metal oxide semiconductor field transistors are connected to each other via a NAND gate. The great advantage of this architecture is that only a small number of data tracks are required, since one track contains several memory cells. In comparison to the NOR architecture, space savings of well over 50% can be achieved.

The access times of the NAND architecture are relatively high, since you cannot access individual memory cells, but can only read out the entire line. So you have to read the data in blocks. You use the page and block-oriented method of working. A page consists of 512 or 2,048 bytes, which in turn are combined into blocks that are between 16 and 128 KB.

As already mentioned, you can only rewrite a memory area, i.e. a page, if you delete it beforehand. Since the data cells are connected in series, you have to delete the entire block in which the page is located.

NANDFlashs are used in areas where a relatively large amount of memory is required and the FlashModule should be small and very compact. The slower access times compared to the NOR architecture are somewhat disadvantageous.

NAND memories or SLC flashes (single level cell) have largely prevailed over NOR memories due to their high storage density and cost-effective production method.

NOR architecture

The so-called NOR memory exists as a counterpart to the NAND memory. As the name suggests, this type of memory consists of NORGattern (Not OR) instead of NAND. This also results in the fact that the memory cells of a NORFlash are connected in parallel to one another, which enables parallel and therefore random access. So you don't have to go block by block when reading.

The parallel NORFlash has a lower storage density and is comparatively expensive, but executes polling processes much faster. NOR memories are also known as MLC memories (Multi Level Cell) and can be addressed flexibly due to their “grid” structure. Due to their design, they have relatively little storage space compared to NAND memories, since the parallelism simply takes up more space.

As a general rule

NAND is used in large data stores in commercial systems because it can be used to build more compact, faster and cheaper storage systems. In the NOR memory, the memory cells can be read out in parallel, which is why such memories are particularly fast for z. B. very suitable as a program or code memory. NAND memory cells can be packed more densely. However, many storage cells are connected in series. Therefore, the optional read access is slower. Erase and write are faster with NAND memories when working in blocks. The storage densities for NAND and NORT technologies for one-bit cells differ. Usually NAND uses single-bit cells and NOR multi-bit cells. However, both techniques are now also available in multi-bit versions with two, three or four bits in a cell. In general, multi-bit cells are slower and more prone to failure than single-bit cells. Multi-bit cells have the same cell size and the bits in the cell are defined by different voltage levels. This explains the cost reduction of MLC memories compared to SLC memory (single level cell). However, there are always difficulties in precisely reading out the small voltage differences. For this reason, SLCFlash in NAND technology should be used in the current state of the art in professional IT operations for reliability and speed reasons. Multi-bit cells have the same cell size and the bits in the cell are defined by different voltage levels. This explains the cost reduction of MLC memories compared to SLC memory (single level cell). However, there are always difficulties in precisely reading out the small voltage differences. For this reason, SLCFlash in NAND technology should be used in the current state of the art in professional IT operations for reliability and speed reasons. Multi-bit cells have the same cell size and the bits in the cell are defined by different voltage levels. This explains the cost reduction of MLC memories compared to SLC memory (single level cell). However, there are always difficulties in precisely reading out the small voltage differences. For this reason, SLCFlash in NAND technology should be used in the current state of the art in professional IT operations for reliability and speed reasons.

lifespan

Due to their mode of operation, FLASH memories cannot be kept indefinitely. Every write and erase process causes a certain amount of wear in the oxide layers of the cells, so that the manufacturers of flash memory cards specify a kind of minimum durability for their products that is based on the number of erase cycles performed.

With NAND (SLC) this value is 100,000 cycles, with NOR (MLC) only 10,000 cycles. A generous security buffer is always included in the calculation, which means that NAND memory cards can usually still function after a million or more erase cycles.

A kind of defect management additionally increases the lifespan of flash memory cards. As soon as the wear and tear within a block affects the integrity of the data, this block is flagged as defective and the data stored there is transferred to a reserve block. This block management also takes care of an even distribution of write accesses right from the start, so that the blocks are used as evenly and uniformly as possible.

control

The control that performs defect management has a significant influence on the performance of a flash memory card. In addition to the maintenance of the flash cells, it primarily deals with the communication of the memory card with the end device and is therefore largely responsible for the data transfer rates that are achieved when reading from and writing to a flash cell. The access times are around 100 µs, the reading speed is 25 MB / s and goes up to over 50 MB / s.

In addition to the speed, the reliability of the data communication and the protocol that ensures that the memory card logs in and out of the operating system smoothly are the responsibility of the controller. Controllers are used for this control.

With most memory cards available today, the controller is integrated in the data carrier unit, but in some models it is also used externally in order to achieve an even more compact design. The disadvantage of the external variant is that external controllers often have to manage more than one card model and are not optimally tailored to the requirements of each one. Loss of performance and compatibility problems can then result.

outlook

It is now assumed that there is a period of four to five quarters between two flash generations, in which the capacity can be increased by at least 50% or in many cases doubled. This can continue similarly in the next generations, but the curve could flatten out considerably! For this reason, new and different flash memory technologies are already being worked on. IDC expects the use of flash memory in mobile PCs to increase from almost 0 to 20% in 2011.

RAM - Random Access Memory

RAM is short for Random Access Memory. RAM is also referred to as main memory or main memory and is a type of memory whose memory cells can be addressed directly via their memory addresses. In this context, one speaks of “optional”, which refers to “random” and has nothing to do with “random” or “random access”. In this context one speaks of "random access memory" or "direct access memory".

RAM allows access to each individual memory cell, both reading and writing. RAM memory always requires power. If the power supply is switched off, the data in RAM are lost.

RAM is used as working memory in computer and microcontroller systems. Programs and data from external storage media and hard disks are loaded into this working memory. For fast processing, the processor can access it, process it and then write it back into the main memory.

Short story

The origin of the term goes back to the early days of computers, in which all data was stored in sequentially readable forms of storage such as punch cards, magnetic drum memories and magnetic tapes, which were loaded into fast calculation registers for processing. In order to have intermediate results available faster than these block-oriented devices, delay lines were temporarily used for intermediate values until the ferrite core memories were then introduced. These writable memories already had the same form of matrix access as today's RAMs. At that time, the fast memory types were all writable and the main innovation was the random access to the magnetic core memory and the RAM modules that were then placed on semiconductor memory.

functionality

RAMs are semiconductor memories and there are basically two types of RAM: SRAM as static RAM and DRAM as dynamic RAM. In the case of static RAMs, the information is stored in feedback circuits (so-called flip-flops), in the case of dynamic RAMs in capacitors whose charge is periodically refreshed. While of the recharge, the processor (CPU) has no access to the DRAM. This is why computers with DRAMs often work slower than those with SRAMs. However, the memory capacity of the DRAMs is significantly higher than that of SRAMs.

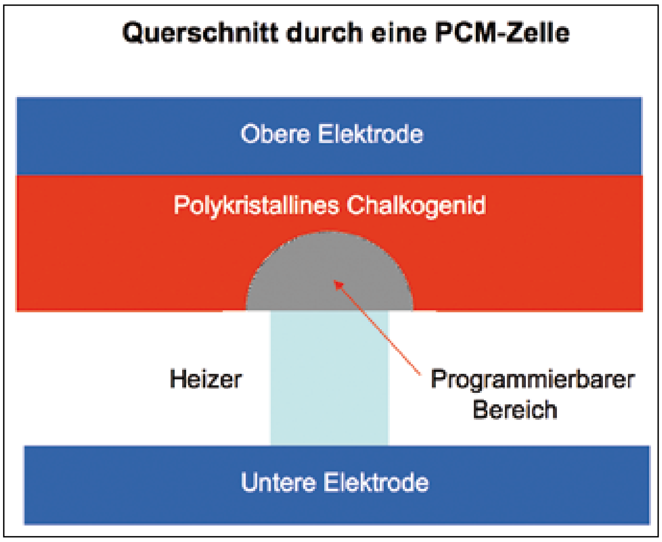
SRAM is static. This means that the memory content is stored using flip-flops and is thus retained after the memory content has been called up. As a result, the power consumption is very high, but this leads to fast work within the storage. Due to its high price and high power consumption, SRAM is only used as a cache or buffer memory with low capacities (e.g. L1, L2 and L3 cache of large computer systems). Its content is volatile, ie the stored information is lost when the operating voltage is switched off. In combination with a backup battery, a special form of non-volatile memory NVRAM can be realized from the static RAM,

DRAM is simpler and slower in comparison. For a long time, only this one type of memory with its various variants was known in the computer area for the main memory.

A DRAM memory cell consists of a transistor and a capacitor, which is the actual memory element. The information is stored in the form of the state of charge of a capacitor. This very simple construction makes the memory cell very small, but the capacitor discharges quickly due to the leakage currents that occur, and the information content is lost. Therefore, the memory cells must be refreshed regularly. DRAM is much cheaper than SRAM. That is why DRAM is mainly used for the main memory of a PC.

Phase Change Memories (PCM) - phase change memory

Already in the 1920s it was observed that the electrical conductivity changes due to a structural change on a chalcogenide. The semiconductor properties of crystalline and amorphous chalcogenides were researched in the 1950s. In the 1960s, reversible phase-changing materials started on their elek



trical and then look at optical properties.

Chalcogenides are chemical compounds from one or more chalcogen elements such as oxygen, sulfur, selenium and tellurium, which form solids with an ionic or covalent character with other binding partners (such as arsenic, germanium, phosphorus, lead, aluminum and other). The solids mostly appear crystalline.

In 1968 phase change memories were considered for the first time as a possible storage medium. At this point, however, the technology was not yet ready to keep up with other storage media such as DRAM and SRAM.

However, the chalcogenides were specifically researched further into optical storage and found their market with the CD / DVD RW. A point on a CD is heated with a laser so that it changes its state (amorphous to crystalline and back again).

Only when materials were discovered in the course of these developments that came to interesting regions in terms of writing times and flow, did the Phase Change RAM development pick up speed again.

With phase change memory, the chalcogenide is stimulated with a current pulse to change the phase in contrast to the optical application. Changing the state changes the electrical resistance. A distinction can be made between two or more states and more can be saved in one cell.

PCM - How It Works

Inside a phase change memory is a tiny amount of a semiconductor alloy that can quickly switch between an ordered crystalline phase with a lower electrical resistance and an unordered amorphous phase with a much higher electrical resistance. Since the maintenance of both phase states does not require a power supply, the phase change memory is not volatile.

The phase state of the material is determined by the amplitude and duration of the electrical current surge that is used to heat the material. If the temperature of the alloy is increased to slightly above the melting point, the atoms excited by the supply of energy are distributed in a random arrangement. If the power supply is then abruptly interrupted, the atoms solidify in this arbitrarily arranged, amorphous phase. If the power supply is reduced over a longer period of time - approximately 10 nanoseconds - the atoms have enough time to return to the preferred, well-ordered structure of the crystalline phase.

PCM versus flash

Technologically, PCM connects where NANDFlash reaches the limits of feasibility. Many experts believe that flash technology will face major problems in the near future due to its limited scalability.

In IBM's Almaden research laboratories, in collaboration with Macronix and Qimoda, new storage technologies were developed that can replace FlashMemories. The "Phase Change Memory" technology (PCM) enables significantly higher speeds for smaller sizes. This improves the properties of PCMs as they grow smaller. Less electricity is required because less material has to be heated and the speed increases due to the small distances. All advantages that can be achieved with the shrinking of PCM cells. However, as the cells become smaller, the electrical resistance also shrinks, making it more difficult to distinguish between the states. It is therefore important to find a good balance here.

"Together we have now succeeded in developing a new material for phase change memory that offers high performance even with extreme miniaturization," says Rolf Allenspach, head of the nanophysics group at the IBM research laboratory in Rüschlikon near Zurich.

NANDFlash is currently still on the rise, but the technology is still expensive, the capacities for use in computers alone are too low, and they run out after only 100,000 read and write cycles. This relatively fast material fatigue is not a problem for consumer applications, but it disqualifies the modules for use as main and buffer storage in network or storage systems. Hybrid solutions, which combine a conventional hard disk with flash modules, are also increasingly entering the market to enable faster data access.

Flash memory technology is likely to become a victim of miniaturization in the future. New production processes in chips make smaller and smaller traces possible, but according to current research, flash memory has ended at 45 nanometers. Litter losses prevent data from being saved without power.

The PCM modules, on the other hand, are intended to enable production processes of up to 20 nanometers and even less. The theoretical limit for the scalability of PCMs is less than 5 nanometers. The developers use a germanium alloy for their storage as new material.

PCM - small and universally applicable

Together with high performance and reliability, PCM could therefore become the basic technology for universal memory in mobile applications, because phase change memories, like flash, are not volatile. This means that they remember the information even when no power is being supplied. In addition, the PCM prototype from IBM has very advantageous technical values: It can be written 500 times faster and requires only half the energy.

In addition, significantly more read and write cycles are possible and the components with a cross-section of three times 20 nanometers are considerably smaller than the smallest flash memories that can be produced today. “As soon as the scalability of Flash finally reaches its limits, a new technology like PCM will prevail,” says Allenspach, from the IBM laboratory in Zurich. This could soon become possible.

Today's PCMs are about 30 times faster than Flash and range from 100 to 200 nanometers. However, if you consider the rapid development in terms of write and read speed, PCM is also suitable as a replacement for the RAM that is common today. With PCM as a replacement for RAM, the previous booting of computers would no longer be necessary, however, security problems could arise since the PCMs are not volatile (one could read out the confidential data in RAM such as passwords etc.). In addition, PCM has much less wear and tear, such as occurs with flash due to the wear of the oxide layer on the floating gate. PCM therefore offers a long data shelf life and not too high energy consumption during operation.

Millipede nano technology

The “Millipede” project was started in the IBM basic research laboratory in Rüschlikon in the mid-1990s. The aim of the project was to use nano-technology and nano-storage architectures for data storage and to integrate them into storage subsystems. The project manager in Rüschli kon was the IBM employee Peter Vettinger and the technological driving force was the IBM and Nobel Prize winner Gerd Binnig. The development of Millipede brought IBM many new patents in the field of nano-technology and nano-mechanics and gave many new insights into polymer structures and the possibility of using plastics as a data carrier. Millipede development is therefore of crucial importance for future technologies.

Basic principle

The basic principle is simple and is comparable to that of the earlier punch card, but now with structure sizes in the range of nanometers. Another crucial sub

schied: using the technology used, bits can be deleted and overwritten. Tiny levers with a fine tip made of silicon also melt tiny recesses into a polymer medium to write bits. The same tips can also be used to detect these depressions, i.e. to read out the bits again. To do this, bring the tip close to the polymer film and heat it. If the tip is immersed in a bit crater, the heat exchange between it and the storage medium increases, which reduces the electrical resistance of the lever, also called the cantilever. In order to overwrite a bit like that, the tip of the crater is used to create new depressions, the edges of which overlap the old depression and thus push the polymer material towards the crater.

Because the holes are so extremely small, they can be placed very close together and such fantastic data densities can be achieved: With revolutionary nanotechnology, IBM scientists in the Rüschlikon research laboratory have managed to penetrate the millionth of a millimeter range. A record density of 1 terabit per square inch could be exceeded when storing data - which corresponds to the content of 25 DVDs on the surface of a stamp. The Terabit density was achieved with a single silicon tip, the depressions with a diameter

of approximately 10 nanometers. In order to increase the data rate, i.e. the write and read speed, not only a tip is used, but a whole matrix of levers that work in parallel. The new version has more than 4,000 such tips, which are arranged in a small square with a side length of 6.4 millimeters. These dimensions enable a complete, high-capacity storage system to be packaged in the smallest standardized format for flash memory.

A cantilever in **Millipede** writes and reads from a cell that is around 100 by 100 micrometers in size. For example, while the read and write head and the storage medium move on hard disks, only the medium moves on the Millipede memory. Two tiny coils, which are placed between magnets, drive the movement of the plate: The microscanner can be positioned with an accuracy of up to 2 nanometers. The position can be precisely determined at any time from the overlapping area of strip-shaped sensors.

Cantilever array

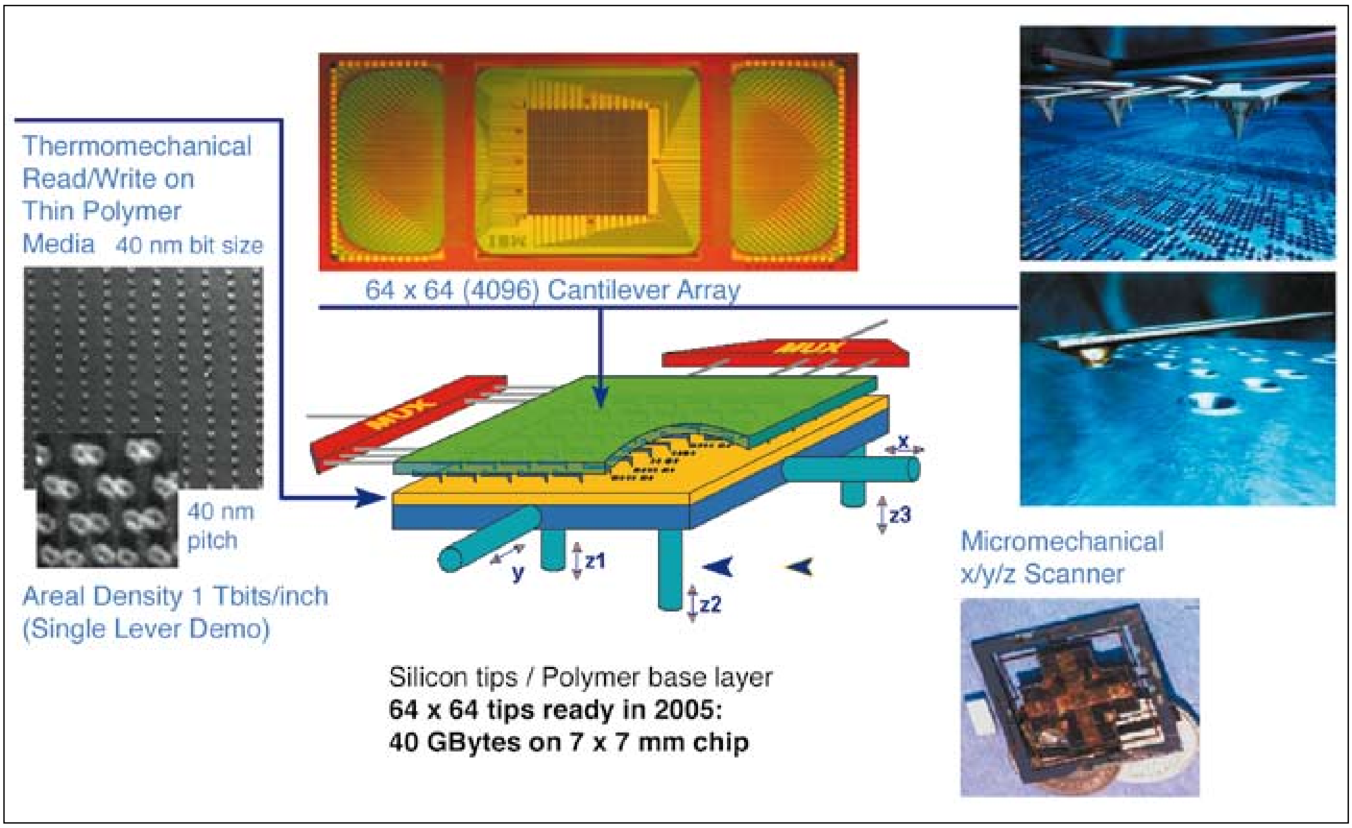
The core of Millipede technology is a two-dimensional arrangement of v-shaped silicon spring tongues (cantilever) that are 70 micrometers (thousandths of a millimeter) long. At the end

Each "tongue" has a sensor for reading and a resistor above the tip for writing. The tip is almost 1 micron long and the radius is only a few nanometers. The cells with the cantilevers are arranged on a cantilever array chip. The chip is 7 x 14 mm. At the center is the array, which currently consists of a total of 4,096 (64 x 64) cantilevers etched out of the silicon. The actual data carrier consists of a polymer film just a few nanometers thick on a silicon substrate. Controlled individually via multiplexers, the heads read, write or delete the desired bit. Up to 100,000 write and overwrite cycles have been successfully tested so far. Although mechanical engineering is used in the design,

Microscanner

The movement of the storage medium relative to the cantilever array is realized using the silicon-based x / y micro scanner. The scanner consists of approx.

6.8 x 6.8 mm² scan table that carries the polymer medium and two electromagnetic triggers. The scanner chip is mounted on the silicon plate, which serves as the mechanical base of the system. The distance between his



The surface and the surface of moving parts of the scanner is approx. 20 µm. The scan table can be moved by 120 µm in the x and y directions. Each trigger consists of two permanent magnets built into the silicon plate and a small coil located between the magnets. In order to extinguish the vibrations from the outside, a so-called pivot is used, which is coupled to the triggers.

If the tip "falls" into a recess, the reading sensor cools down slightly due to the shorter distance to the substrate, but this leads to a measurable change in the resistance. In order to overwrite data, the tip places depressions in the surface. Their outer edges overlap the old depressions and thus delete the old data. More than 100,000 write and overwrite cycles have proven that the concept is suitable for a rewritable memory type.

Position determination

The positioning information is provided by four thermal sensors. These sensors are located directly above the scan table, on the cantilever array. The sensors have thermally insulated heaters. Each sensor is positioned over an edge of the scan table and is heated by electricity. Part of this heat is conducted through the air to the scan table, which now serves as a cooler. Moving the scan table causes a change in the efficiency of this cooling system, which changes the temperature of the electrical resistance from the heater.

A sophisticated design ensures that the tips are precisely leveled over the storage medium and dampens vibrations and impacts from the outside. Time division multiplexing electronics, as used in a similar way in memory chips (DRAM), enables the addressing of each individual tip in parallel operation. Electromagnetic actuation moves the substrate with the storage medium on its surface very precisely in the x and y direction, so that each tip can read and write in its storage field with a side length of 100 micrometers. The short distances make a significant contribution to low energy consumption.

In order to get the data into and out of the memory faster, a complete matrix arrangement of levers processes the medium simultaneously. However, it turned out that it is extremely difficult to manufacture mechanics and electronics in one piece on one chip. So the scientists decided to build it in two parts:

1. ***The matrix of the micro tips is provided with tiny contact pins, which look like the knobs of Lego bricks under the electron microscope.***
2. ***These studs are then contacted with their counterparts on the electronic board.***

The prototypes shown have recently demonstrated the technical feasibility of a product, for example in terms of storage density, performance and reliability. While the storage technologies used today are gradually reaching fundamental limits, the nanomechanical approach has enormous development potential for a storage density that is a thousand times higher. This nanomechanical data carrier generates almost no heat, consumes little electricity and is completely shock-resistant.

For the functions of the device, ie reading, writing, erasing and overwriting, the tips are brought into contact with the polymer film on the silicon substrate.

Writing technology

Bits are written by heating the resistor integrated in the cantilever to typically 400 degrees Celsius. The tip, which is also heated as a result, softens the polymer, sinks in and leaves a depression of a few nanometers. For reading, the cantilever's reading sensor is heated without softening the polymer film.

Millipede project phases

Chips were developed in three generations throughout the project. The last chip completed in 2007 works with 128 x 128 tips, has a capacitive storage capacity of 160 Gbytes and a data rate of 120 Mbit / s. IBM is currently considering how and how MillipedeChips can be used sensibly. So it remains exciting to see how nano-technologies and nano-mechanics will find their way into the world of data storage.

Revolutionary storage technology from IBM research is getting closer to being realized

"Racetrack Memory" opens up new dimensions in storage density, speed and reliability

April 11, 2008: IBM researchers report in the renowned science journal Science about a breakthrough in the research of a new storage technology, which has become known under the name "Racetrack" (German: race track). The technology could create a new class of storage that combines the high performance of flash storage with the large capacity and low cost of hard disk storage.

In the current issue of Science, IBM fellow Stuart Parkin and his colleagues from the IBM Almaden Research Center, USA, describe an important step in the development of memory chips based on the "race track process". Within the next ten years, the technology could be used and open up new dimensions of storage capacities, enable extremely fast switching on of electronic devices and be much cheaper, more energy efficient and more robust than today's systems.

It was thanks to Stuart Parkin that the GMR effect (plate technology, see GMR) was converted into a product in the form of a GMR reading head, and it took seven years to do so. It is thanks to him and his team that the high hard disk capacities were possible today. The author does not agree that Stuart Parkin was disregarded when the Nobel Prize for Physics was awarded in 2007, especially since the Nobel Prize could have been awarded to three people. With Racetrack Memories he will write real history in the memory area! The author is convinced of that!

Universal storage of the future

The revolutionary concept is based on the storage of information in the form of tiny, oppositely magnetized regions (domains) in a nanowire. In the conventional hard disk used as a storage medium, the medium and a read / write head are moved to read, write or erase data. It is different with the "racetrack process": Here the magnetic domains are shifted towards the central reading and writing units, which are attached in the middle of the nanowire - and at extremely high speed. The stored data bits seem to "race" through the data conductor, hence the name "Racetrack".

Stuart Parkin, IBM Fellow and one of the largest memory researchers and developers at IBM

Since a single 'racetrack' is only a few nanometers in size and can store between 10 and 100 bits, the technology allows extremely high storage densities. Compared to flash memories, a 'racetrack memory' could record 100 times more data in the same area. That would correspond to around 500,000 music tracks or 3,500 films. Due to the minimal power consumption of a 'Racetrack memory', an MP3 device could also be operated with a single battery for weeks.

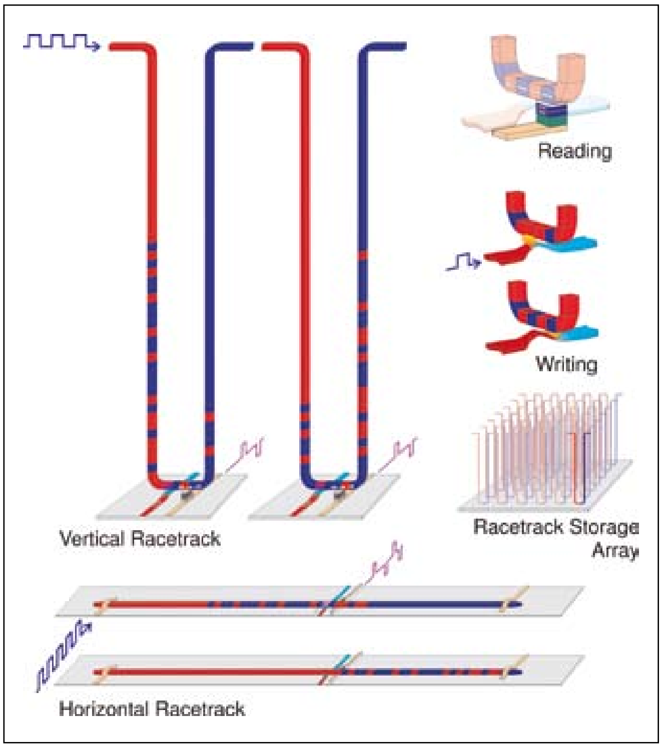
In addition, the process does not require any moving parts, which means that there are almost no signs of wear or tear. This makes the 'Racetrack storage' more resistant than all existing storage technologies and gives it an almost unlimited lifespan.

“The connection between the research of new physical phenomena that occur at the molecular and atomic level and nano-engineering - the ability to selectively structure materials from individual atoms, since a single“ racetrack ”is only a few nanometers in size and between 10 and 100 Can store bits, the technology allows extremely high storage densities. Compared to flash memories, a “racetrack memory” could record 100 times more data on the same area. That would correspond to around 500,000 music tracks or 3,500 films. Due to the minimal power consumption of a “Racetrack memory”, an MP3 device could also be operated with a single battery for weeks.

In addition, the process does not require any moving parts, which means that there are almost no signs of wear or tear. This makes the "Racetrack storage" more resistant than all existing storage technologies and gives it an almost unlimited lifespan.

“The connection between the research of new physical phenomena that occur at the molecular and atomic level and nano-engineering - the ability to create material structures from individual atoms and molecules - always poses new and exciting challenges,” explains Stuart Parkin. He continues:

"Exploiting the potential of our" racetrack technology "could lead to fascinating new applications - applications that are still unimaginable today."



Schematic representation of racetrack memory storage

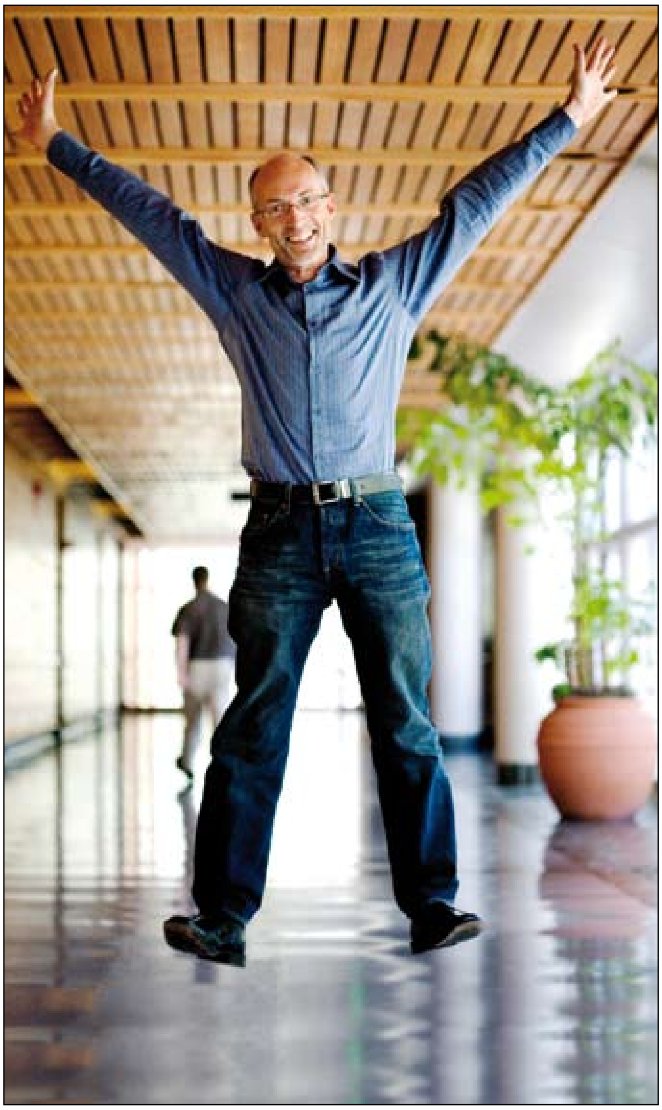
Radical technical innovations that started in basic research have already opened up new markets and fields of application several times in the history of IBM. These include the memory chip, the relational database or the hard disk.

The Racetrack Method

For around 50 years, scientists have been trying to store digital information through so-called magnetic domain walls - the interfaces between two oppositely magnetized regions in a magnetic material. Until now, domain walls could only be manipulated using costly, complex and energy-intensive processes.

In their recently published work “CurrentControlled Magnetic DomainWall Nanowire Shift Register”, Parkin and his teammates demonstrate a new, very efficient method. Here, the domain walls are moved by a spin polarized current, which can rotate the magnetization in the wall. The reason for this is the spin transfer effect, which considerably simplifies the storage technology because the current flows directly through the domain walls and no additional magnetic fields have to be generated.

In the second work "Magnetic DomainWall Racetrack Memory", the researchers summarize the basic principles of "Racetrack Technology". This is based on the storage of data in tiny magnetic regions within a data conductor. The researchers use small nanowires as data conductors, which are magnetically described with information. These are attached lengthwise horizontally or in the form of a loop vertically on a silicon surface. When a voltage is applied, neither the wires nor the silicon move, but only the magnetic domain walls in which the information is stored.



Stuart Parkin is right to be happy, he did it and Racetrack can change the world

Each magnetic domain that represents a bit has a "head", a magnetic north pole, and a "tail", a magnetic south pole. The domain walls that form at the interfaces alternate between "HeadtoHead" and "TailtoTail" configurations in the data conductor. The distances between two successive domain walls are equal to the bit length and are determined by markings along the nanowire, so-called pinning centers.

In their experiment, the researchers used nanowires made of permalloy (an FeNi alloy) and showed how they can create, move and read out domain walls by using current pulses of precisely controlled duration in the nanosecond range. The write and shift cycle only takes a few 10 nanoseconds.

The goal is to arrange many thousands of these nanowire memories, which can store between 10 and 100 bits, densely on one surface. The vertical arrangement of nano wires also opens up new types of three-dimensional architecture - a paradigm shift compared to the current two-dimensional chip and memory technologies. The exploitation of the third dimension creates new scope for performance increases and cost reductions that are no longer only achieved through miniaturization - as has been the case until now.

IUPAP Magnetism Award and Louis Neel Medal

At the **July 27, 2009** receives IBM Fellow Dr. Stuart Parkin one of the highest honors in the field of magnetism research. He will receive the "International Union of Pure and Applied Physics Magnetism Award (at the" International Conference on Magnetism "in Karlsruhe **IUPAP)** and the **Louis Neel medal** for his fundamental research and development results in the field of nano-magnetism. The IUPAP Magnetism Committee only awards this award every three years. Most of all, he was honored for his latest research on Racetrack Memories. The Louis Neel Medal goes back to the scientist Louis Eugene Felix Neel, who received the Nobel Prize in Physics in 1970 for the discovery of antiferromagnetism.

Nanotechnology - key technology of the 21st century

Nanotechnology is a new technology that uses functions on an extremely small scale. It focuses on structures and processes in dimensions under 100 nanometers - about 400 times thinner than a human hair. In the nanometer size range, many fundamental processes in biology, chemistry and physics take place, which can be controlled to an unprecedented degree and open up new perspectives in many areas. These include highly developed functional materials, nanoelectronics, information and communication technology, sensors, life sciences and energy technology. Nanotech applications could contribute to the more efficient use of solar energy or to new types of water treatment.

Through research at the ETH Zurich (Swiss Federal Institute of Technology) and the IBM research laboratory in Rüschlikon ZRL (Zurich Research Laboratory), Zurich has repeatedly provided decisive impulses in quantum mechanics and nano research. These include, for example, the groundbreaking concepts of quantum mechanics by ETH physicist and Nobel laureate Wolfgang Pauli or the development of the scanning tunneling microscope (RTM) by Gerd Binnig and Heinrich Rohrer at the IBM research laboratory in Rüschlikon. For this, the researchers received the Nobel Prize in Physics in 1986. This device gave the first glimpse into the world of atoms. Shortly thereafter, it became possible to manipulate individual atoms, which opened the door to nanotechnology widely has been. The RTM is generally viewed as the instrument that opened the door to the nanocosmos. The atomic force microscope (RKM), which is closely related to the RTM, was invented in 1986 by Gerd Binnig.

The scanning tunneling microscope (RTM) is called STM (Scanning Tunneling Microscope), the atomic force microscope (RKM) is called AFM (Atomic Force Microscope).

IBM and ETH Zurich found a new joint nanotech research center

The new Nanotech research center from IBM and ETH Zurich covers around 6,000 square meters. The two institutions share a long tradition of scientific cooperation. At a joint media conference in Zurich, Ralph Eichler, President of ETH Zurich, and John E. Kelly, Senior Vice President and Director of Research at IBM, announced the planned collaboration. The central element of this collaboration is a new building that is being built on the premises of the ZRL in Rüschlikon. The foundation stone will be laid in spring 2009.

The new research center on the IBM campus in Rüschlikon will enable nano research at the highest level. The "Nanoscale Exploratory Technology Laboratory" is part of an important strategic partnership in nanotechnology with the Swiss Federal Institute of Technology Zurich (ETHZ). Research activities will start in 2011.

IBM research location Rüschlikon: new laboratory for nano-technology

The research focus of the two institutions ranges from basic research to applied research. Research is carried out jointly in various areas, such as carbon-based materials, NanoPhotonics, Spintronics, nanowires and tribology.

New breakthrough in nanotechnology - IBM researchers measure the charge state of individual atoms with the atomic force microscope

In collaboration with scientists from the universities of Regensburg and Utrecht, IBM researchers achieved a breakthrough in the field of nanotechnology in June 2009. For the first time, they were able to measure the charge state of individual atoms directly using atomic force microscopy (RKM). The precision with which they were able to distinguish between uncharged or positively or negatively charged atoms was a single electron charge with a spatial resolution that was accurate to the nanometer. This opens up new possibilities for the research of nanostructures and building blocks on an atomic and molecular scale in application areas such as molecular electronics, catalysis or photovoltaics.

For their experiments, the researchers used a combination of a scanning tunnel (RTM) and an atomic force microscope (RKM), operated in ultra-high vacuum and at low temperatures (5 Kelvin) in order to achieve the stability required for the measurements.

An RKM uses an atomically fine tip attached to a vibrating cantilever to measure the forces that occur between this tip and the atoms on the substrate. In the present work, the researchers used a so-called qPlus force sensor, in which the tip is attached to a prong of a tuning fork, as found in mechanical watch movements, while the other prong is fixed. The tuning fork is mechanically excited and vibrates with an amplitude of 0.02 nanometers. This corresponds to only about a tenth of the diameter of an atom. Becomes

If the RKM tip is placed very close to the sample, for example over a single atom, the resonance frequency of the tuning fork changes due to the forces that occur between the sample and tip.

With this method and under extremely stable conditions, the IBM researchers were now able to measure the minimal differences in the force that exist between the tip and individual, differently charged atoms. The force difference between a neutral gold atom and a gold atom with an additional electron is only about 11 picoNewtons, measured with a minimum distance between tip and sample of about half a nanometer. The measurement accuracy of these experiments is in the range of 1 picoNewton, which corresponds to the gravitational force that two people exert on each other at a distance of more than half a kilometer. The researchers also determined how the force changed with the voltage applied between the tip and the sample. This allowed the distinction

This breakthrough is another important advance in the field of nano research. In contrast to the RTM, which relies on electrically conductive samples, the RKM can also be used for non-conductive samples. In molecular electronics, in which the use of molecules as functional building blocks in future circuits and processors will be researched, non-conductive carrier substances (substrates) are required. Therefore, atomic force microscopy would be preferred in such experiments.

"The atomic force microscope with an accuracy of one electron charge is ideally suited to investigate the charge transfer in molecular complexes, which could provide us with valuable, new knowledge and physical foundations and one day could also lead to new components in information technology," explains Gerhard Meyer, who leads the research in the area of scanning tunnels and atomic force microscopy at the IBM research laboratory in Zurich. Computer devices on the molecular scale have the potential to be orders of magnitude smaller, faster and also more energy efficient than today's transistors and memory devices.

For future experiments, the researchers can imagine making connections from individual metallic atoms and molecules, charging them with electrons, and measuring their distribution directly with the RKM. IBM researcher Leo Gross also points out the importance of her work for other areas: “Charge status and charge distribution are critical factors in catalysis and in the conversion of light into electrical energy. Mapping the charge distribution on an atomic scale could lead to a better understanding of the basic processes in these areas. ”

The latest results follow a series of fundamental scientific breakthroughs that IBM researchers have achieved in recent years: In 2008, scientists at the IBM Almaden Research Center in California were able to use a qPlusRKM to measure the force required to create a force for the first time Moving atom on a surface. This paved the way for the current experiments. In 2007, Gerhard Meyer's team at the IBM Research Laboratory in Zurich demonstrated a molecule that could be switched between two states in a controlled manner without changing its external shape. As early as 2004, the same group had successfully carried out an experiment in which they could manipulate the charge state of a single gold atom with an RTM. By applying voltage pulses to the RTM tip, the researchers were able to apply an additional electron to a single atom that was on an insulating substrate. This negatively charged atom remained stable until a corresponding voltage pulse with the opposite sign occurred using the RTM tip. The researchers also used this method in their current experiments to charge the individual atoms.

IBM researchers are showing the inner structure of molecules with atomic resolution for the first time

IBM researchers in the IBM ZRL Rüschlikon succeeded in August 2009, for the first time atomically resolving the complete chemical structure of a molecule using an atomic force microscope. This enabled the scientists to make a breakthrough in the field of nanosciences, which could also open up new opportunities for research into new types of electronic components on the atomic and molecular scale.

In recent years, astounding progress has been made in characterizing nanostructures on the atomic scale using atomic force microscopy. Up to now, however, science has been prevented from looking at the internal structure of a molecule with atomic resolution.

In the edition of the science journal Science dated
On August 28, 2009, IBM researchers Leo Gross, Fabian Mohn, Nikolaj Moll and Gerhard Meyer and Peter Liljeroth from the University of Utrecht reported how they were able to use an RKM to map the complete chemical structure of pentacene molecules (C22H14) with atomic resolution The experiments were carried out in an ultra-high vacuum and at very low temperatures (5 Kelvin or -268 ° C). The images obtained are somewhat similar to X-rays, which allow a look inside the human body. The “RKM with X-ray vision” from the IBM researchers can look through the electron cloud that envelops the molecule and map the atomic backbone of a pentacene.

This publication follows just two months after another work by the same group, published on June 12, 2009 in Science. The charge measurement of individual atoms using RKM was described in this. "Scanning probe technologies offer incomparable potential for building prototypes of complex functional structures on the atomic scale and thus tailoring and examining their electronic and chemical properties in a targeted manner," explains Gerhard Meyer, who is responsible for research in the field of scanning tunnel microscopy and atomic force microscopy at IBM in Rüschlikon directs.

The results of the two research projects open up new possibilities for studying the charge distribution in specific molecules or molecular networks. Knowledge of these processes is generally very important for the development of electronic components on the atomic and molecular scale. Such components could enable faster, more powerful and more energy-efficient processors and memory chips in the future if the limits of today's chip technologies are pushed to the limit.

The RKM uses a tiny, very sharp metal tip to measure the minimum forces that occur when this tip is brought very close to a sample, such as a molecule. With a point-by-point measurement of these forces, an image of the surface can be created. The pentacene molecules that the IBM researchers used in this work consist of 22 carbon atoms and 14 hydrogen atoms and are only 1.4 nanometers (millionths of a millimeter) long. The distances between the carbon atoms arranged in hexagons are around ten times smaller. These hexagonal carbon rings are surprisingly clearly resolved on the RKM images and even the positions of the hydrogen atoms can be clearly identified.

IBM researcher Leo Gross emphasizes: "The decisive factors for the resolution were an atomically sharp tip with a defined structure and a very high stability of the overall system." In order for the chemical structure of a molecule to be visible, the tip must be extremely close - less than one Nanometer - to introduce the molecule. Forces occur only in this area, which are largely determined by chemical interaction. In order to achieve this, the researchers had to improve the sensitivity of the tip and overcome a major hurdle: similar to two magnets lying next to each other that attract or repel each other, the molecule shifts or attaches itself to the tip if it comes too close. If this happens, no further measurements can be carried out.

Both problems could be solved by choosing a suitable atom or molecule at the RKM tip. “In order to sharpen our tip, we specifically added atoms and molecules to the tip using manipulation techniques. Our measurements with differently prepared tips clearly show that the foremost atom or molecule of the tip has a decisive influence on the resolution, ”states Leo Gross. A carbon monoxide molecule (CO) at the top brought about the breakthrough and, at a distance of about half a nanometer from the pentazene, ensured optimal resolution of the individual atomic positions and their chemical compounds.

The scientists also managed to create a complete three-dimensional, topographical representation of the forces over the molecule. “The data collection took more than 20 hours. This placed the highest demands on the thermal and mechanical stability of our system in order to ensure that the tip and the molecule remained unchanged throughout the entire time, ”describes Fabian Mohn, PhD student at IBM Research in Rüschlikon.

Theoretical calculations carried out by IBM researcher Nikolaj Moll confirmed the experimental results and provided information on the exact nature of the imaging mechanism. Nikolaj Moll explains: "The calculations show that the so-called Pauli repulsion between the C molecule at the tip and the pentacene is responsible for the atomic contrast." This repulsive force is due to a quantum mechanical effect that prevents two identical ones from being formed Too close electrons.

For readers who want to know more about the two experiments and research work, the scientific work of L. Gross, F. Mohn, N. Moll, P. Liljeroth and G. Meyer with the title “The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy ”, published in Science, Vol. 325, No. 5944, pp. 1110-1114 (August 28, 2009) is recommended.

The author is convinced that nanotechnologies will find their way into memory technology to a large extent in the next few years and enable memory chips that go beyond the limits of today's chip technologies!

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

Accelerating Your Success™

AVNET®
technology solutions

Wie sich einzelne Technologien zu einer zukunftsträchtigen Arbeitsumgebung optimal kombinieren lassen, lässt sich mit dem Avnet Mobile Solution Center nicht nur theoretisch erklären, sondern in Hands-On Trainings und im Demonstrationsbetrieb live miterleben.

Wir bringen vorkonfigurierte Systemlösungen in zwei rollbaren Flight-Cases zu Ihnen oder Ihren Kunden und zeigen bequem vor Ort das IBM Low End und Midrange Produktportfolio zu den Themen Virtualisierung, E-Mail-Archivierung, BladeCenter und Backup & Recovery mit TSM FastBack.

Nutzen Sie die Vorteile:

- Lead- und Demand-Generierung bei Kundenveranstaltungen
- Abstrakte Themen in Hands-On-Trainings zum Anfassen
- Aufbau von Technologiekompetenz
- Erweiterung der technischen sowie vertrieblichen Ressourcen
- Erhöhung Ihres Mehrwerts bei Kunden durch die Adressierung von IT-Wachstumsthemen

Einsatzmöglichkeiten:
Endkundenveranstaltungen, Machbarkeitsstudien, Performance-Analysen, Benchmarks, Datenportierungslösungen, Produktschulungen, Hands-On Trainings

Oft kopiert und nie erreicht.
Nutzen Sie außerdem unser mobile Prometric Testcenter um im Rahmen von Zertifizierungsschulungen vor Ort die Aus- und Weiterbildung Ihrer Mitarbeiter voranzutreiben.

Jetzt testen und überzeugen lassen.

Individuelle Beratung durch unsere Spezialisten:
+49 (0)711-51881-0 oder amsc@avnet.com

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:


LÖWEN-
STARK!
Storage-Lösungen von Azlan.
Dem besten Partner auf der Spur? Sie haben ihn gefunden!
Der Storage-Markt wächst kontinuierlich! Wachsen Sie mit!
Azlan, einer der führenden Value Added Distributoren Europas, gibt Ihnen dazu alles an die Hand, was Sie benötigen!
Und das sind 6 gute Gründe, die nicht nur im Storage-Umfeld für Azlan sprechen:
• lösungsorientiert
• professionell
• innovativ
• kundennah
• individuell
• erfahren
Sprechen Sie uns an! Azlan. Ihr Value Added Distributor für Storage
Ihr Azlan Team ist gerne für Sie da. Telefon 089 4700-3003. Wir machen Sie stark: www.azlan-ibm.de
Treffen Sie die richtige Entscheidung.
Mit den richtigen Partnern.
Mit Value Add. Von Azlan.
IBM Premier Business Partner
Azlan
A Trademark of Tech Data

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

becom | Computacenter Services & Solutions

Erster zu zweit. Das freut Dritte.

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

Kundenreferenz von FRITZ & MACZIOL

IBM XIV Storage System –
das technologisch führende und innovative Speichersystem

EURONICS

Die EURONICS Deutschland eG ist der zweitgrößte Marktteilnehmer der Branche für Consumer Electronics in Deutschland. Für SAP und VMware benötigte die Verbundgruppe eine hochverfügbare Produktionsumgebung. Stetige Zuwächse im Datenbereich mussten abgedeckt, der hohe Wartungsaufwand der Storagesysteme unterschiedlicher Hersteller sollte reduziert werden. Vornehmliches Ziel war die Automatisierung der Failover-Szenarien speziell im VMware-Bereich. Eine hochverfügbare, RZ-übergreifende Storagelösung für VMware sowie die Homogenisierung der Storage-Plattform für AIX, Windows, Linux, Novell und VMware mit zentraler Administration gehörten ebenso zu den gewünschten Ergebnissen wie die Konsolidierung im Bereich Windows, Linux und Novell. Die von **FRITZ & MACZIOL** realisierte Lösung zeichnete sich durch zwei Kriterien aus: Als erster IBM Business Partner im deutschen Markt hat **FRITZ & MACZIOL** das neue IBM XIV Storage System erfolgreich implementiert. Und im unternehmenseigenen Demo-Center konnte im Vorfeld bezüglich Performance und Ausfallszenarien die Leistungsfähigkeit der XIV getestet werden. EURONICS verfügt nun an zwei Standorten über jeweils ein System p570, drei x3850, zwei stretched SVC Cluster und zwei IBM XIV Systeme. Die Daten sind mittels SVC in beide Rechenzentren gespiegelt. Mit Hilfe von stretched SVC Clustern wurde eine hochverfügbare VMware-Umgebung mit automatischem Failover aufgebaut. Die Datensicherung erfolgt an einem dritten Standort. Das IBM XIV Storage System ist als CoD (Capacity on Demand) System ausgelegt, damit Speicheranforderungen „on the fly" provisioniert werden können. Kürzere Batch-Fenster und eine vereinfachte Administration sorgen nicht nur für höhere Anwenderzufriedenheit, sondern auch für Kosteneinsparungen.

Accredited for IBM Dynamic Infrastructure Specialty Elite

IBM Premier Business Partner

Das Unternehmen FRITZ & MACZIOL

Als erstes autorisiertes IBM XIV Storage Solution Center in Deutschland ist **FRITZ & MACZIOL** Vorreiter für innovative Storage-Technologien, damit unterstreichen wir einmal mehr unsere führende Position in IBM Storage Lösungsbereichen. So dürfen wir die ersten zwei XIV Installationen im deutschen Markt offiziell zu unseren Kundenreferenzen zählen: Euronics Deutschland eG und Deutsche Vermögensberatung AG. Wir präsentieren live in unserem Demo-Center (auch für Remote Demo verfügbar) das revolutionäre Speichersystem.

Mit rund 700 Mitarbeitern an 15 Standorten hat sich **FRITZ & MACZIOL** mit einem ganzheitlichen Portfolio aus Hardware, Software und Services aus einer Hand sowohl in Deutschland als auch der Schweiz positioniert. Mit der FuM Asia Inc. und Servicebüros in verschiedenen Ländern ist das Unternehmen auch international vertreten. **FRITZ & MACZIOL** vertreibt Software- und Systemlösungen für öffentliche Auftraggeber, den Mittelstand und Großunternehmen.

Ansprechpartner:
Robert Schreiner
Phone +49 89 4567 8587
e-mail rschreiner@fum.de

Hörvelsinger Weg 17 · 89081 Ulm · Tel. +49 731 1551-0 · www.fum.de

Deutschland Berlin · Düsseldorf · Frankfurt · Hamburg · Hannover · Karlsruhe · München · Nürnberg · Reilingen · Stuttgart · **Schweiz** Neuheim/Zug · Zürich · **Asien** Manila/Philippinen **Servicebüros: England** Gloucester · **Polen** Warschau · **Saudi Arabien** Riyadh · **Spanien** Murcia · **USA** Alpharetta
FRITZ & MACZIOL : INFOMA®, Part of Imtech N.V., Gouda, Netherlands

FRITZ & MACZIOL
Software, Systeme und Dienstleistungen

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

LSI is the technology partner powering the IBM System Storage™ DS3000 & DS5000 series.

The superior technology provides mixed workload capabilities exceeding today's demands in transaction and bandwidth intensive applications.

The DS3000 & DS5000 series storage arrays offer relentless data security including self encrypting drives and are ideal for key application environments including VMware, Oracle, Exchange and Hyper-V.

LSI Logic GmbH Orleansstrasse 4
81669 München
www.lsi.com

Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:

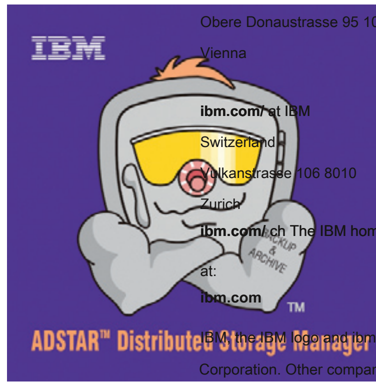
Many thanks also to the sponsors of this IBM System Storage Compendium, who supported the print run:


SVA

IBM Deutschland GmbH
IBM-Allee 1 71139 Ehningen

ibm.com/ de IBM Austria
Obere Donaustrasse 95 1020
Vienna

ibm.com/ at IBM
Switzerland
Vulkanstrasse 106 8010
Zurich
ibm.com/ ch The IBM homepage can be found
at:
ibm.com



xiv
Storage Reinvented

€ 49.90 - recommended retail price

IBM System Storage Compendium 2010