UNIVERSITÄT LEIPZIG

Enterprise Computing
Einführung in das Betriebssystem z/OS

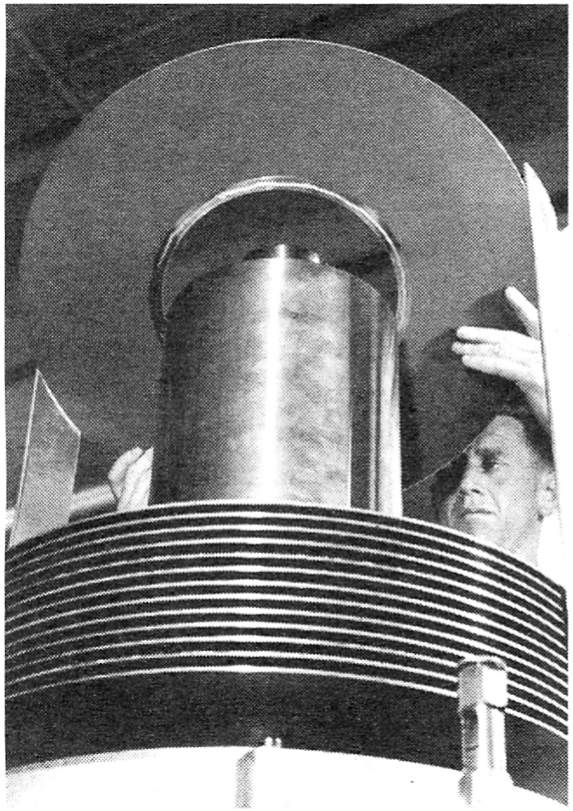
Prof. Dr. Martin Bogdan
Prof. Dr.-Ing. Wilhelm G. Spruth

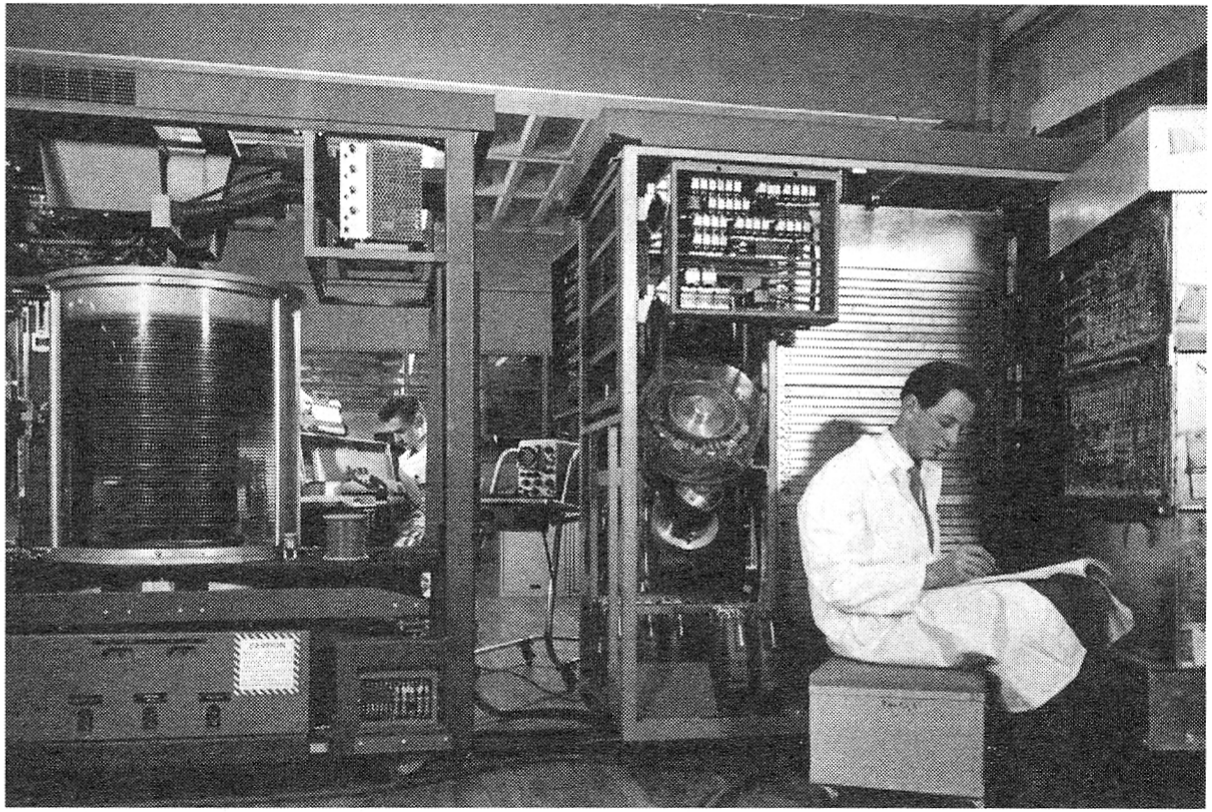
WS2012/2013

Input/Output Teil 1

Plattenspeicher Technologie

© copyright W. G. Spruth,





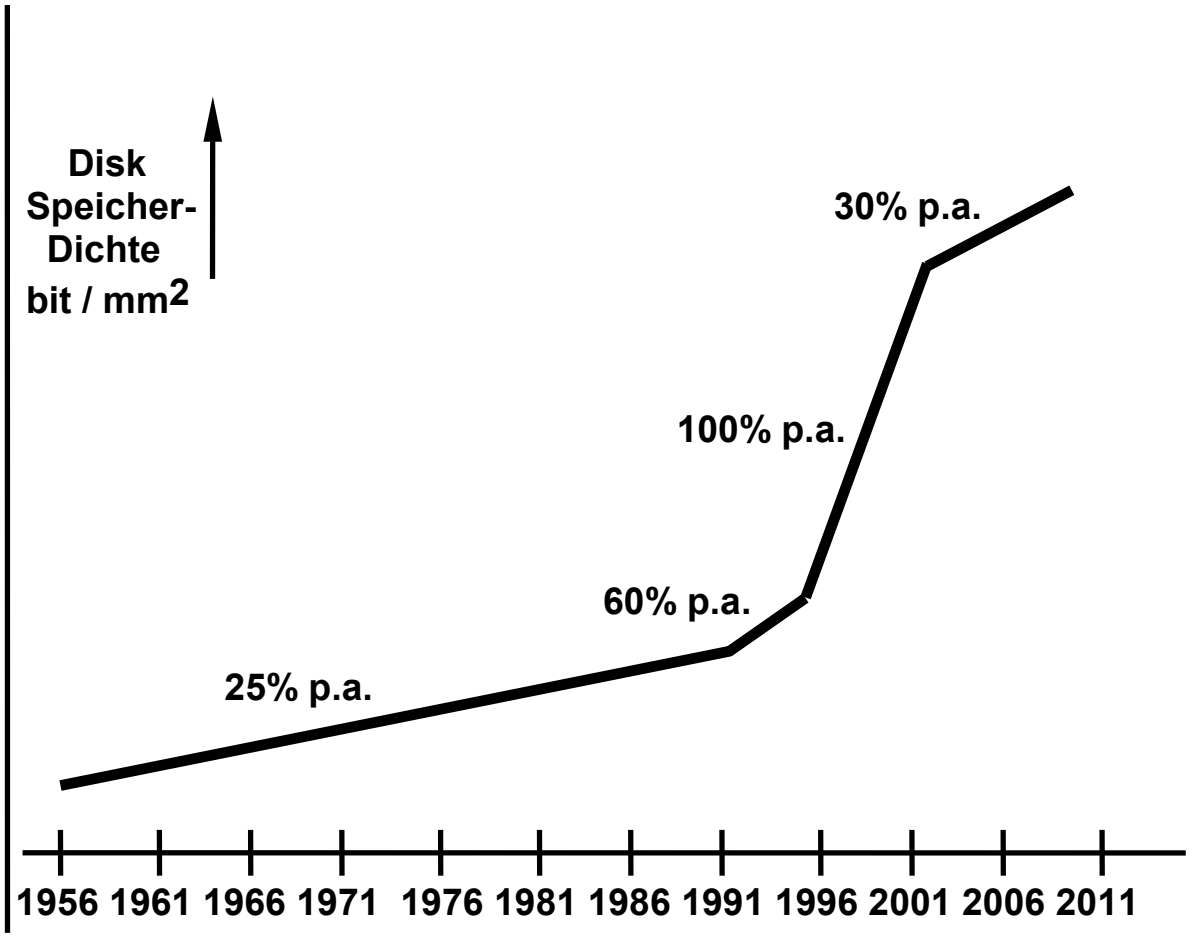
IBM 350 von 1956:

IBM 350 aus dem Jahre 1956:

Am 13. September 1956 stellte IBM das erste Plattenspeichersystem der Welt vor, die "IBM 350 Plattenspeichereinheit", die zusammen mit dem "IBM 305 RAMAC" (Random Access Method of Accounting and Control)-Rechner ausgeliefert wurde. Der Festplattenspeicher im Kleiderschrankformat wog fast eine Tonne.

Auf einundfünfzig mit Eisenoxyd-beschichteten Platten mit einem Durchmesser von 61 cm (24-Zoll) speicherte die RAMAC sechs Millionen Zeichen. Hätte man damals schon Daten in Form von Bytes abgelegt, entspräche dies einer Kapazität von fünf Megabyte. Die IBM 350 Plattenspeichereinheit verbrauchte 2,5 Kilowatt an elektrischer Energie und wurde seinerzeit auf Mietbasis für 10 000 D-Mark pro Monat vertrieben.

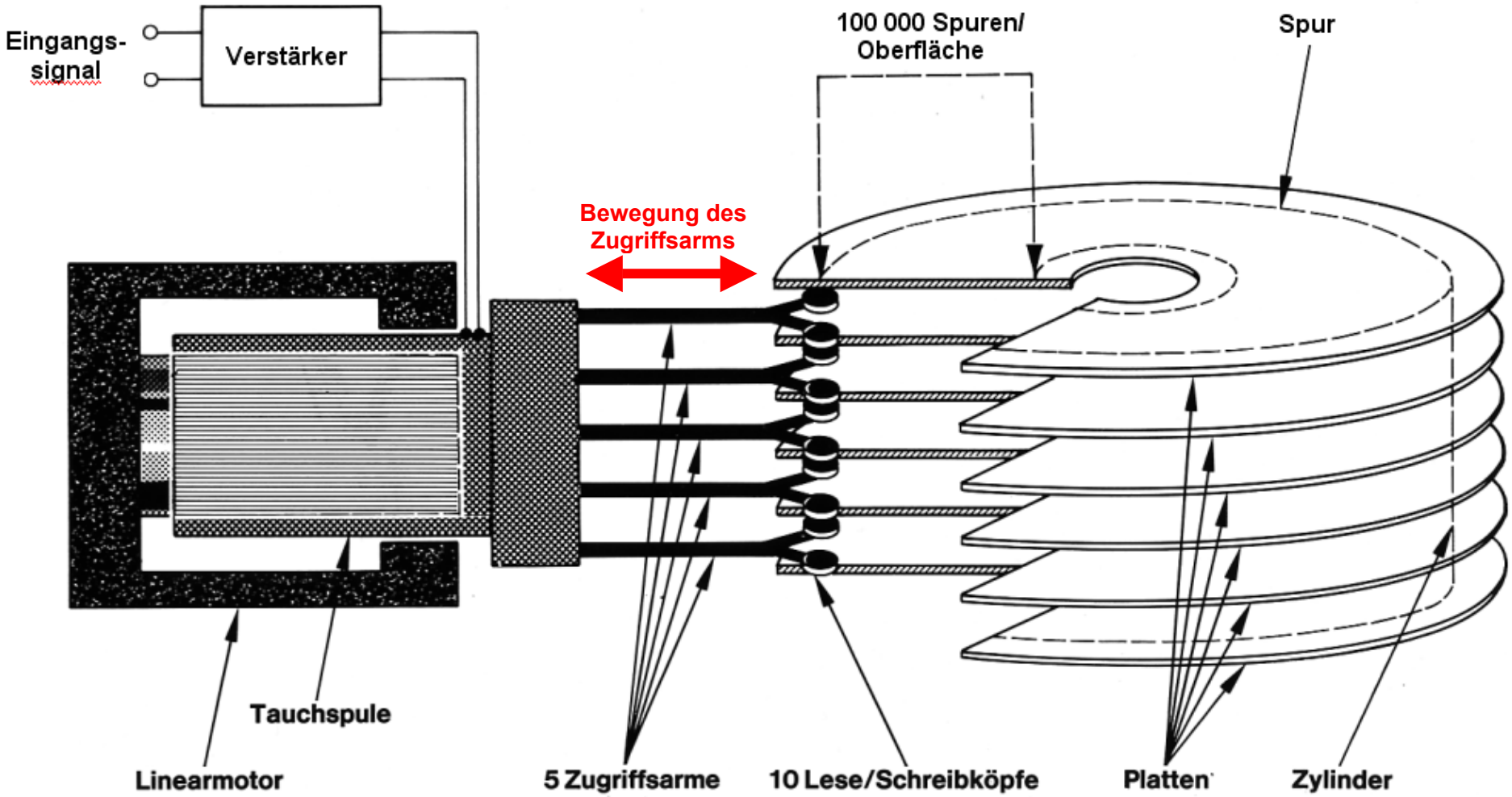
Die modernen Festplattenscheiben verfügen nur über einen Durchmesser von 2,5 Zoll (6,35 cm) oder weniger. Das derzeit stärkste aktuelle Speichersystem der IBM, das System Storage DS8870, hat mit 2,3 PByte Kapazität eine um ca. 400 Millionen mal größere Speicherkapazität als die RAMAC.



Bei den magnetischen Festplattenspeichern wird der technologische Fortschritt daran gemessen, wie viele Bit auf einem Quadratmillimeter Oberfläche gespeichert werden können.

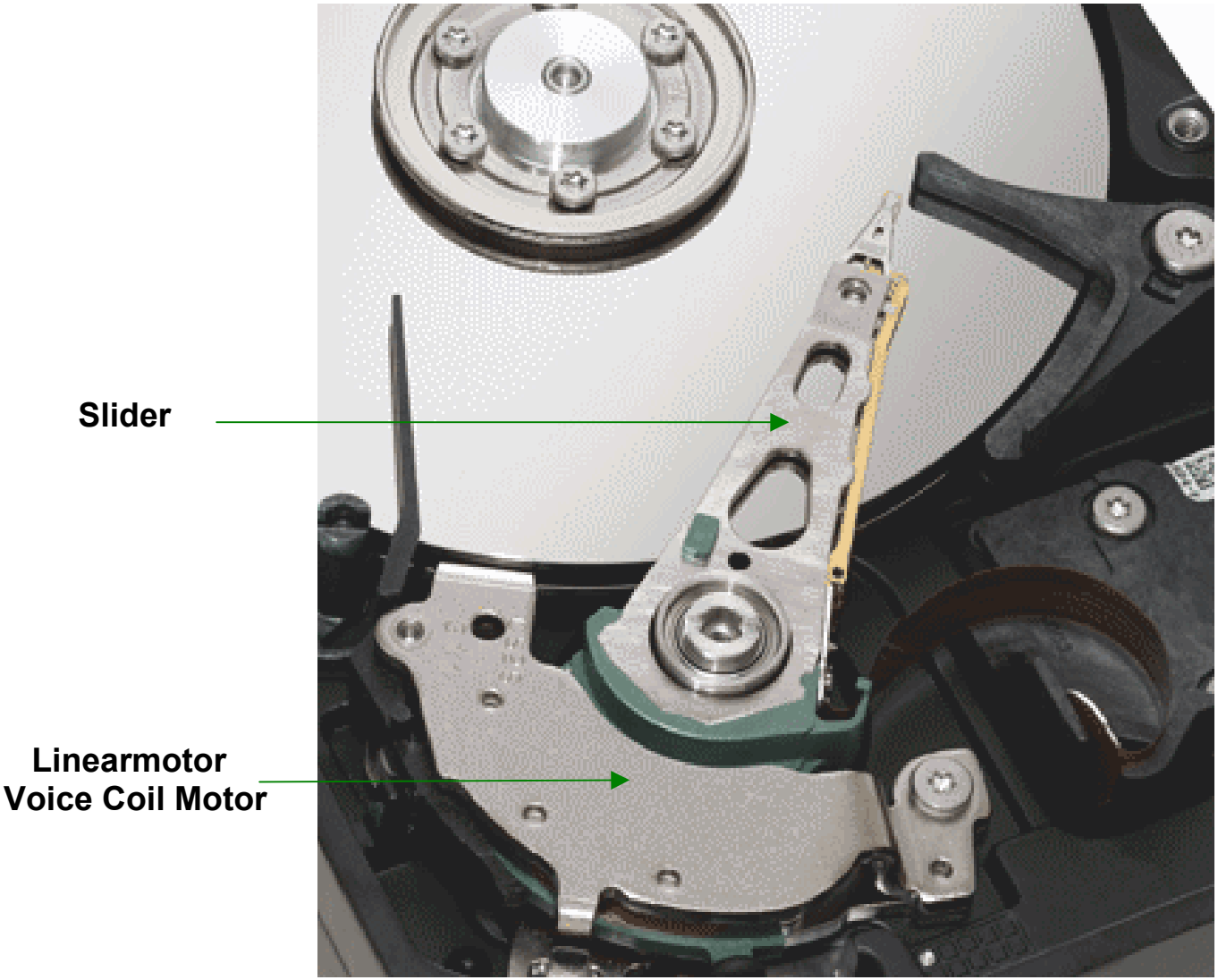
Dargestellt ist die 50 jährige Entwicklung mit dem jährlichen Wachstum der Plattenspeicherdichte. Auf der y-Achse ist die Speicherdichte mit einer logarithmischen Skala dargestellt. Das bedeutet, ein exponentielles Wachtum wird durch eine gerade Linie darsgestellt.

Über mehrere Jahrzehnte betrug das Wachstum konstant etwa 25 % pro Jahr. In den 90er Jahren beschleunigte sich das Wachstum auf etwa 100 % pro Jahr. Ab Anfang der 2000er Jahre verlangsamt sich das Wachstum wieder auf etwa 30 % pro Jahr.



An der ursprünglichen Form eines Plattenspeichers hat sich in den Jahrzehnten nicht viel geändert. Der Plattenspeicher besteht in der Regel aus mehreren Platten, von denen beide Oberflächen für die Aufzeichnung von Information genutzt wird. Die Information ist in kreisförmigen Spuren (tracks) auf der Oberfläche angeordnet. Jeder Spur ist eine Spuradresse zugeordnet. Die Spuren selbst sind in Sektoren aufgeteilt.

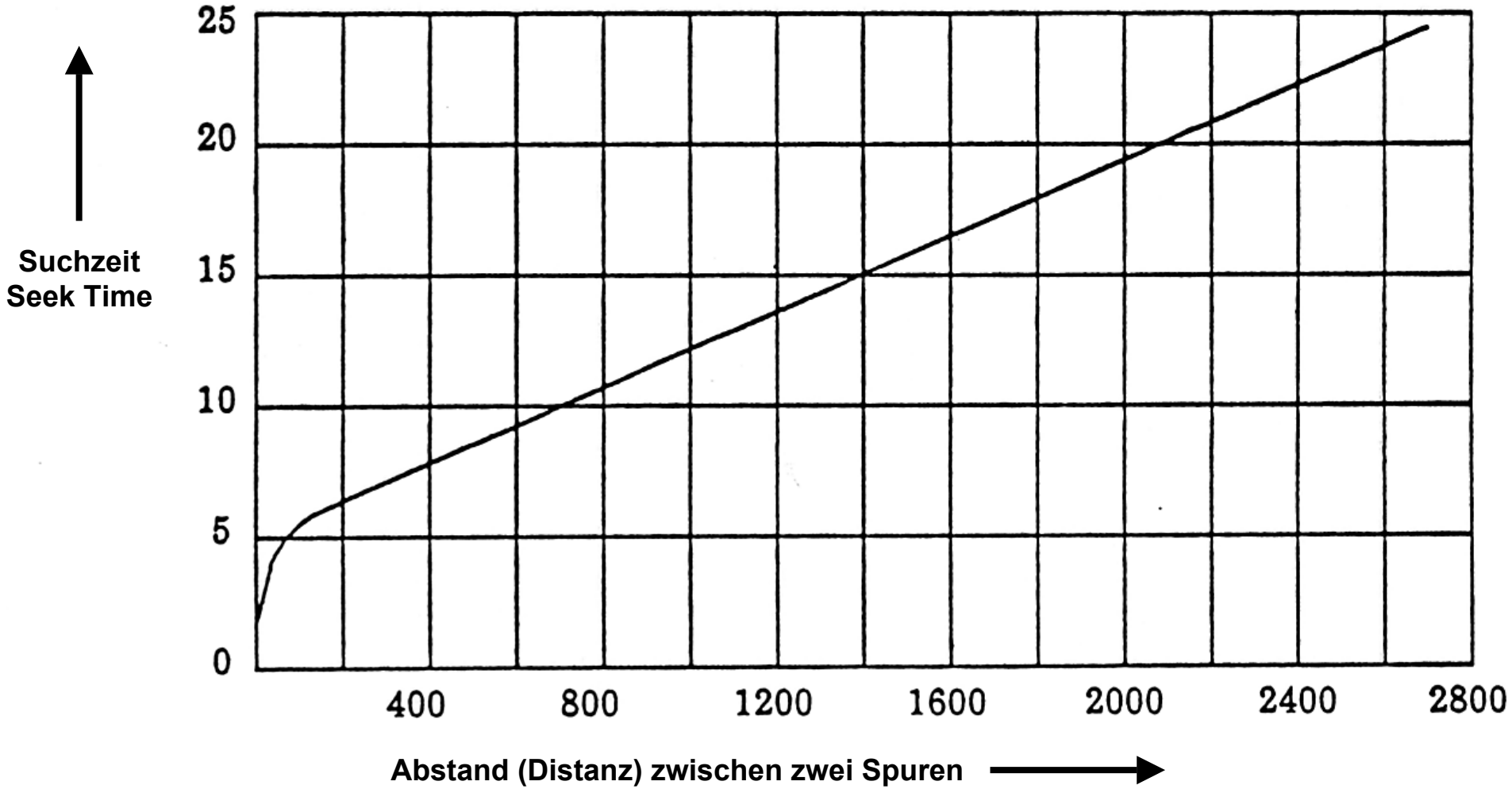
Pro Oberfläche gibt es mindestens einen Lese/Schreib-Kopf. Die Köpfe für alle Oberflächen sind auf einer gemeinsamen Zugriffsarmstruktur (slider) befestigt, welche alle Köpfe mittels eines Linearmotorsauf eine der vielen Spuren bewegt.



Abgebildet ist die Implementierung, einer Seagate Barracuda 7200.7 Festplatte (2006). Der Linearmotor bewegt sich senkrecht zur Achse des Sliders.

Eine Life Demo kann unter http://www.youtube.com/watch?v=L0nbo1VOF4M gesehen werden.

Die Zugriffsarme (slider) werden durch einen Linearmotor angetrieben. Der Linearmotor wird oft auch als Voice Coil Motor bezeichnet, weil er das gleiche Prinzip benutzt wie der Linearmotor in einem Lautsprecher, der die Membrane antreibt. Eine Ansteuerungselektronik setzt eine Spuradresse in die entsprechenden Signale für den Linearmotor um.



Zugriffscharakteristik - Profil der Suchzeiten (Seek Time)

Die Bewegung durch den Linearmotor erfolgt relativ langsam. Der Wechel von einer Spur auf eine unmittelbar benachbarte Spur kann mehrere Millisekunden betragen. Der Wechel von einer Spur auf eine entfernte Spur kann bis zu 100 Millisekunden dauern. Diese Zeit ist kritisch für das Leistungsverhalten eines Plattenspeichers; sie wird als „Suchzeit“ (seek time) bezeichnet. Hersteller von Festplatten geben eine „mittlere Zugriffszeit“ an, die aus einer Mischung von vielen benachbarten und wenigen entfernten Zugriffen ermittelt wird. Zu fragen ist, ob diese Mischung realitätsnahe ist.

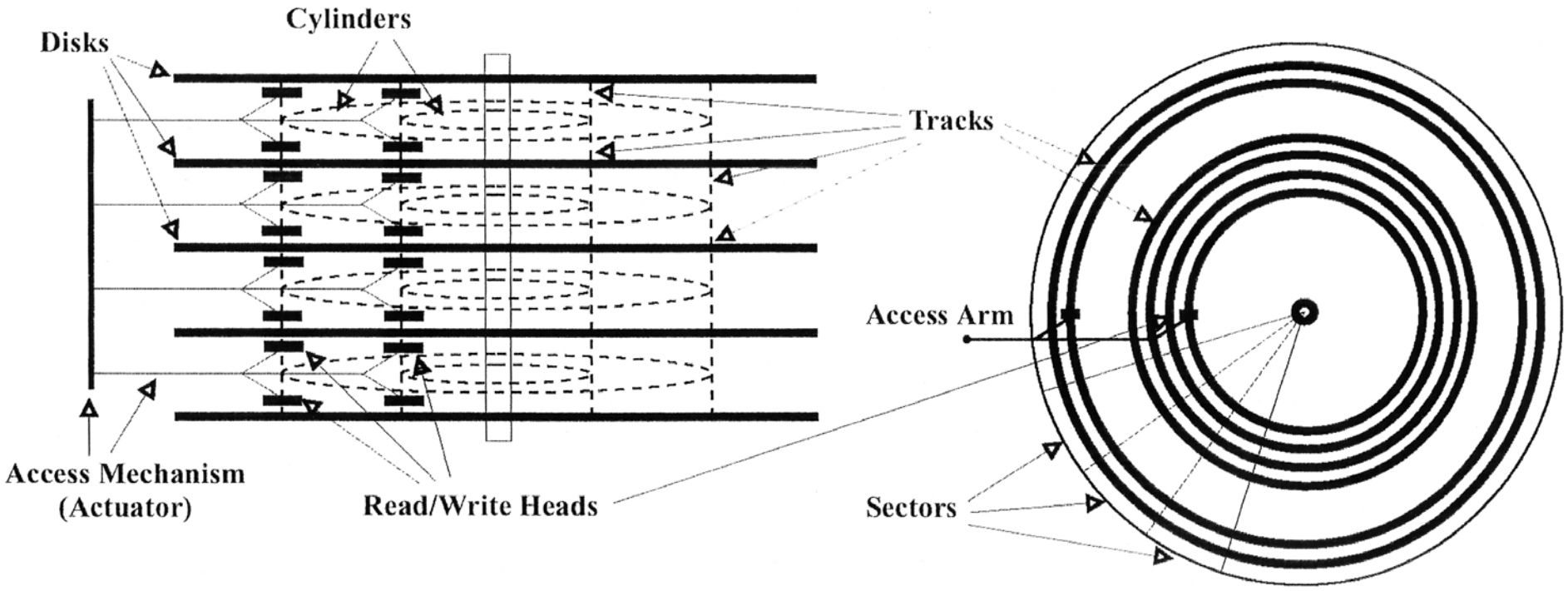
Die Speicherkapazität von Plattenspeichern ist in mehreren Jahrzehnten enorm gestiegen, von 5 MByte bei der ursprünglichen IBM RAMAC 305 bis zu mehr als einem TByte heute. Spuren werden sehr viel enger gepackt und auch die Anzahl der Bits auf einer Spur ist sehr viel größer geworden.

An der Zugriffscharakteristik hat sich aber nur wenig geändert.

Der Lese/Schreibkopf wird mit etwa 100 g beschleunigt und abgebremst (1 g = Anziehungskraft auf der Erdoberfläche). Jenseits von 100 g beginnt Metall sich zu verformen. Es ist daher in den letzten Jahrzehnten nicht möglich gewesen, die Geschwindigkeit der mechanischen Bewegung des Linearmotors deutlich zu verbessern.

Die Platte dreht sich gegenüber dem Kopf mit etwa $^2/_3$ Schallgeschwindigkeit. Es entsteht ein Luftstrom zwischen Platte und Kopf. Der Kopf hat eine Fläche von etwa 1 mm^2 · Er ist aerodynamisch wie der Flügel eines Flugzeuges geformt und erhält einen Auftrieb gegenüber der Plattenoberfläche. Der Auftrieb ist umso größer, je geringer die Entfernung Kopf – Platte ist. Eine Feder drückt den Kopf gegen die Plattenoberfläche. Es stellt sich eine Flughöhe ein, bei der Auftrieb und Federdruck im Gleichgewicht sind. Die Flughöhe beträgt etwa 50 nm.

Ein kleines Gedankenexperiment soll dies verdeutlichen. Vergrößern wir diese Struktur um einen Faktor 100 000. Das Ergebnis ist ein Flugzeug-ähnliches Gebilde mit einer Flügelweite von 100 Meter, dass mit nahezu Schallgeschwindigkeit mit einem Abstand von 5 mm über die Erdoberfläche fliegt. Jede Berührung des Kopfes mit der Plattenoberfläche bewirkt einen irreparablen Plattenspeichercrasch.



Um die Zugriffszeit zu verbessern werden häufig zwei (oder mehr) Lese/Schreibköpfe pro Plattenoberfläche eingesetzt. Dabei kann ein Kopf z.B. nur die äußeren Spuren und der andere Kopf nur die inneren Spuren abdecken. Hierdurch wird die maximale Distanz zwischen zwei Spuren halbiert.

Zusätzlich fasst man Spuren, die auf zwei oder mehr Oberflächen übereinander liegen, zu einem „Zylinder“ zusammen. Dadurch entstehen größere Speichereinheiten, die ohne eine Bewegung des Zugriffsarms adressiert werden können. Beim Wechsel von einer Spur zu einer anderen Spur des gleichen Zylinders muss nur die Elektronik der Lese/Schreibköpfe umgeschaltet werden, was im µs Bereich erfolgen kann.

Gebräuchliche Umdrehungsgeschwindigkeiten bei modernen Platten betragen:

3600 U/min	16,67 ms/Umdrehung
5400 U/min	11,11 ms/Umdrehung
7200 U/min	8,33 ms/Umdrehung
10 000 U/min	6 ms/Umdrehung
15 000 U/min	4 ms/Umdrehung

Bei diesen Umdrehungszeiten bewegt sich der Kopf gegenüber der Plattenoberfläche mit etwa $^2/_3$ Schallgeschwindigkeit. Bei einer größeren Geschwindigkeit würde der laminarer Luftstrom zwischen Plattenoberfläche und Kopf abreißen und zu Turbulenzen führen. Dann kann der geringe Abstand zur Plattenoberfläche nicht mehr eingehalten werden.

Füllt man einen Plattenspeicher mit Helium statt mit Luft, sind höhere Umdrehungszahlen möglich. Die Schallgeschwindigkeit beträgt in Luft etwa 340 m/s, in Helium etwa 980 m/s .

In der Hardware und Betriebssystem-Software eines Rechners existieren sehr weitgehende Einrichtungen, die helfen, das an sich sehr schlechte Zugriffsverhalten eines Plattenspeichers zu verbessern. Es hat auch in den letzten Jahrzehnten viele Versuche gegeben, Plattenspeicher auf Grund Ihres schlechten Zugriffsverhaltens durch eine alternative Technologie zu ersetzen. Das ist über viele Jahre nicht erfolgreich gewesen. Mit der Einführung von Solid State Drives (SSD) wird sich das in der absehbaren Zukunft wahrscheinlich ändern.

Solid State Drive

Ein Solid State Drive (SSD) ist ein Speichermedium, das wie eine herkömmliche magnetische Festplatte eingebaut und angesprochen werden kann, ohne eine rotierende Scheibe oder andere bewegliche Teile zu enthalten, da nur Halbleiterspeicherbausteine verwendet werden.

Der englische Begriff solid state in der Geschichte der Elektronik bedeutet, dass keinerlei bewegliche mechanische Teile wie Relais oder Röhren usw. verwendet werden, sondern Halbleiterbauteile, die durch die Entwicklungen in der Festkörperphysik möglich wurden. Insofern erscheint die Bezeichnung paradox, da ein eben gerade ohne bewegliche Teile auskommendes Medium als 'Drive' bzw. 'Disk' angesprochen wird. Dies geschieht in Analogie zu anderen Festplatten.

In SSD-Laufwerken werden in aller Regel Flash-Speicher Chips vom Typ "Nand" eingesetzt. Dabei wird die Information in Form von elektrischen Ladungszuständen der einzelnen Flash-Speicherzellen gespeichert. Das Speichern der Daten erfolgt nicht flüchtig, das heißt die Ladungen bleiben erhalten unabhängig davon, ob das Medium an die Stromzufuhr angeschlossen ist oder nicht. Das bedeutet, dass sich SSD-Speicher in der gleichen Art und Weise nutzen lassen wie Festplatten.

In den letzten Jahren werden zusätzlich zu (und bei mobilen Geräten an Stelle von) Plattenspeichern Solid State Drives eingesetzt.

Vorteile eines Solid State Drive sind Robustheit, schnelle Zugriffszeiten und Energieverbrauch. Nachteile sind Kapazität und Preis. In großen Mainframe Installationen findet man häufig zusätzlich zu den Plattenspeichern eine begrenzte Anzahl von SSDs für die Auslagerung von Daten für besonders performance-kritische Vorgänge.

Wenn sich SSDs als Alternative zu magnetischen rotierenden Plattenspeichern durchsetzen ist es denkbar, diese anders als herkömmliche Plattenspeicher zu nutzen. Dies könnte zu signifikanten Änderungen in der Betriebssystem Struktur führen.

Die Zukunft des Plattenspeichers

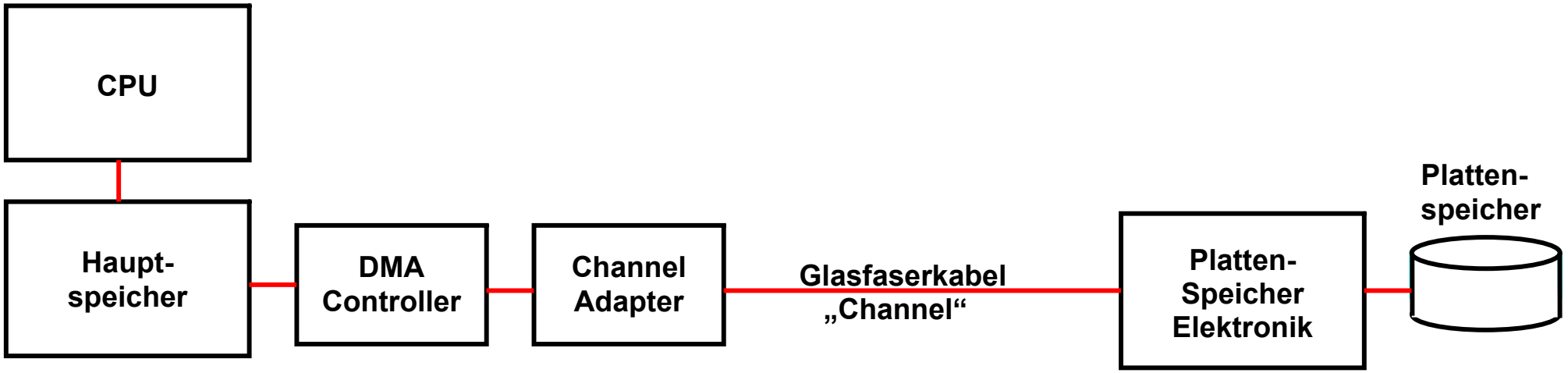
Plattenspeicher behaupten ihre dominierende Rolle gegenüber SSDs, weil der Preisunterschied pro Byte Speicherkapazität etwa einen Faktor 10 beträgt.

Seit Jahrzehnten wird angenommen, dass SSDs in der Zukunft kostenmäßig gleichziehen werden.

Dennoch ist mit einem Zeithorizont von Jahrzehnten anzunehmen, dass man dann den magnetischen Plattenspeicher als genauso archaisch betrachten wird, wie man heute die Lochkarte sieht.

Panta rhei (griechisch ,πάντα ῥεῖ **, „Alles fließt“) sagte schon der griechische Philosoph Heraklit.**

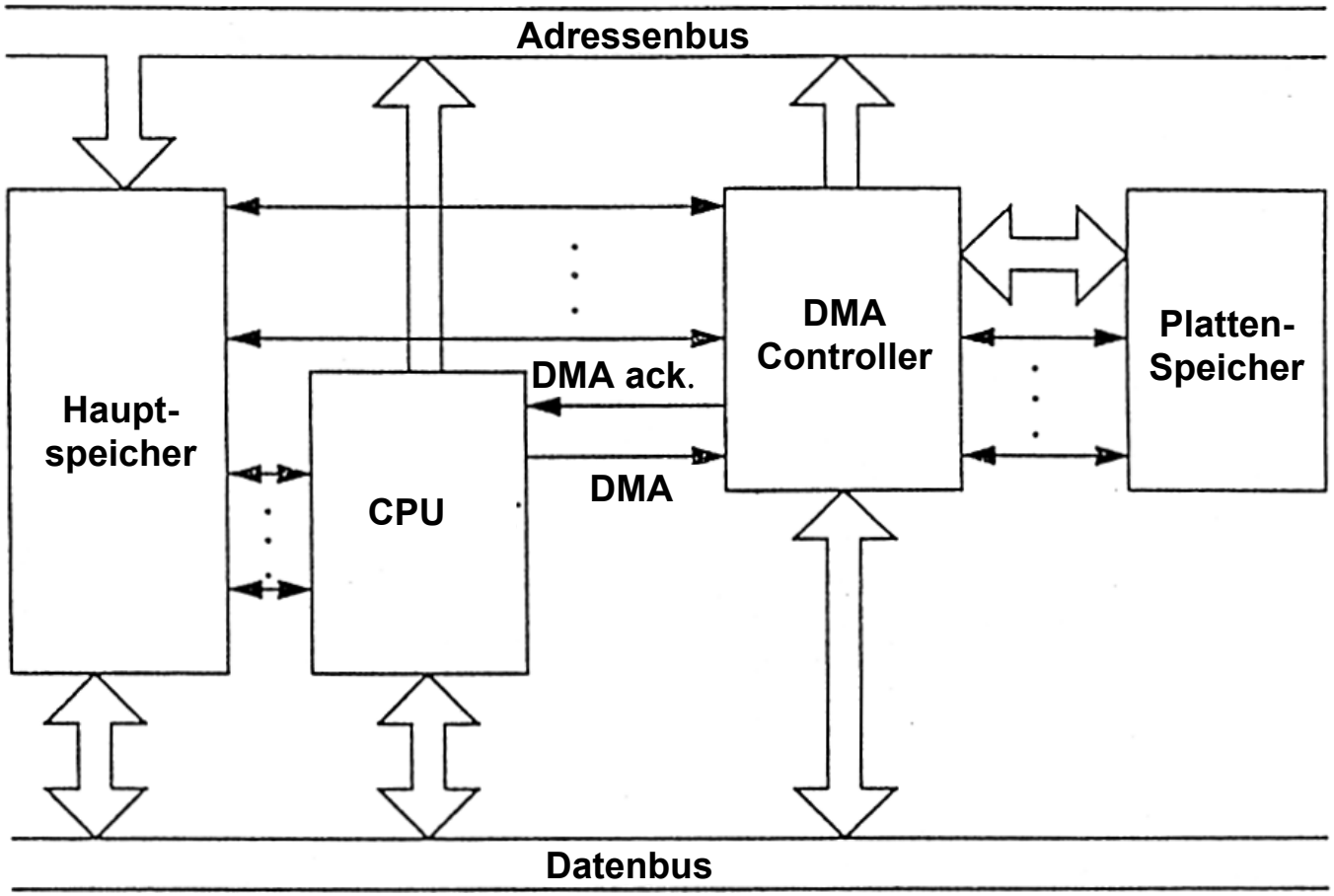
Direct Memory Access



Plattenspeicher werden mittels eines „Direct Memory Access“ (DMA) Controllers an den Hauptspeicher angeschlossen. In einem z196 Mainframe ist dies eine Fanout Card innerhalb eines Books. Der DMA Controller (die Fanout Card) ist wiederum mit einer Channel Adapter Card in einem I/O Cage verbunden.

Die Channel Adapter Card stellt mittels eines Glasfaserkabels die Verbindung zur Elektronik eines Plattenspeichers her. Das Glasfaserkabel wird als Channel, Ficon Channel oder Channel Path bezeichnet (die Begriffe sind weitgehend austauschbar). Ein Mainframe kann mehrere 100 Glasfaseranschlüsse aufweisen. Die Distanz Recher – Plattenspeicher kann viele Kilometer betragen.

Plattenspeicher werden in getrennten Gehäusen, sog. „Enterprise Storage Servern“ (ESS) untergebracht. Ein ESS kann Hunderte von Plattenspeichern enthalten. Für die Ansteuerung enthält ein ESS mehrere Spezialrechner, deren Firmware von außen nicht zugreifbar ist.

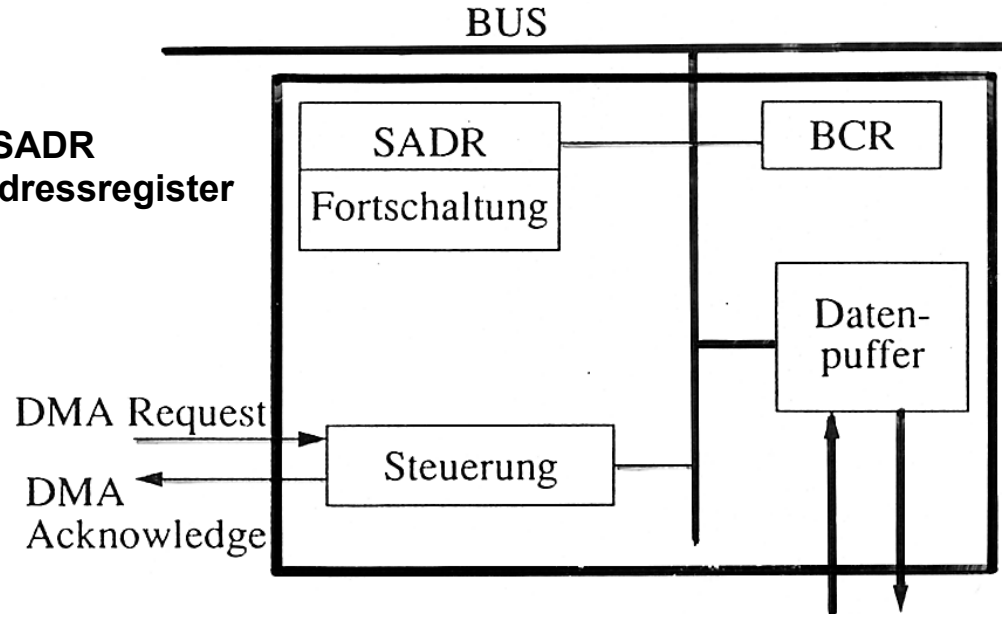


DMA Steuerung

Der Processor (CPU) und der DMA (Direct Memory Access) Controller greifen mit Adress-, Daten- und Steuerleitungen gleichzeitig und parallel auf den Hauptspeicher zu. Jeder Hauptspeicherzyklus wird entweder von der CPU oder von dem DMA Controller genutzt.

Die Memory Bus Adapter (MBA) in dem z9 Rechner, bzw. Host Channel Adapter (HCA) in den z10/z196/zEC12 Rechnern werden gemeinsam auch als Fanout Card bezeichnet, und enthalten einen DMA Controller.

SADR
Speicheradressregister



BCR
Bytecountregister
(Length Count)

DMA Steuerung (Direct Memory Access)

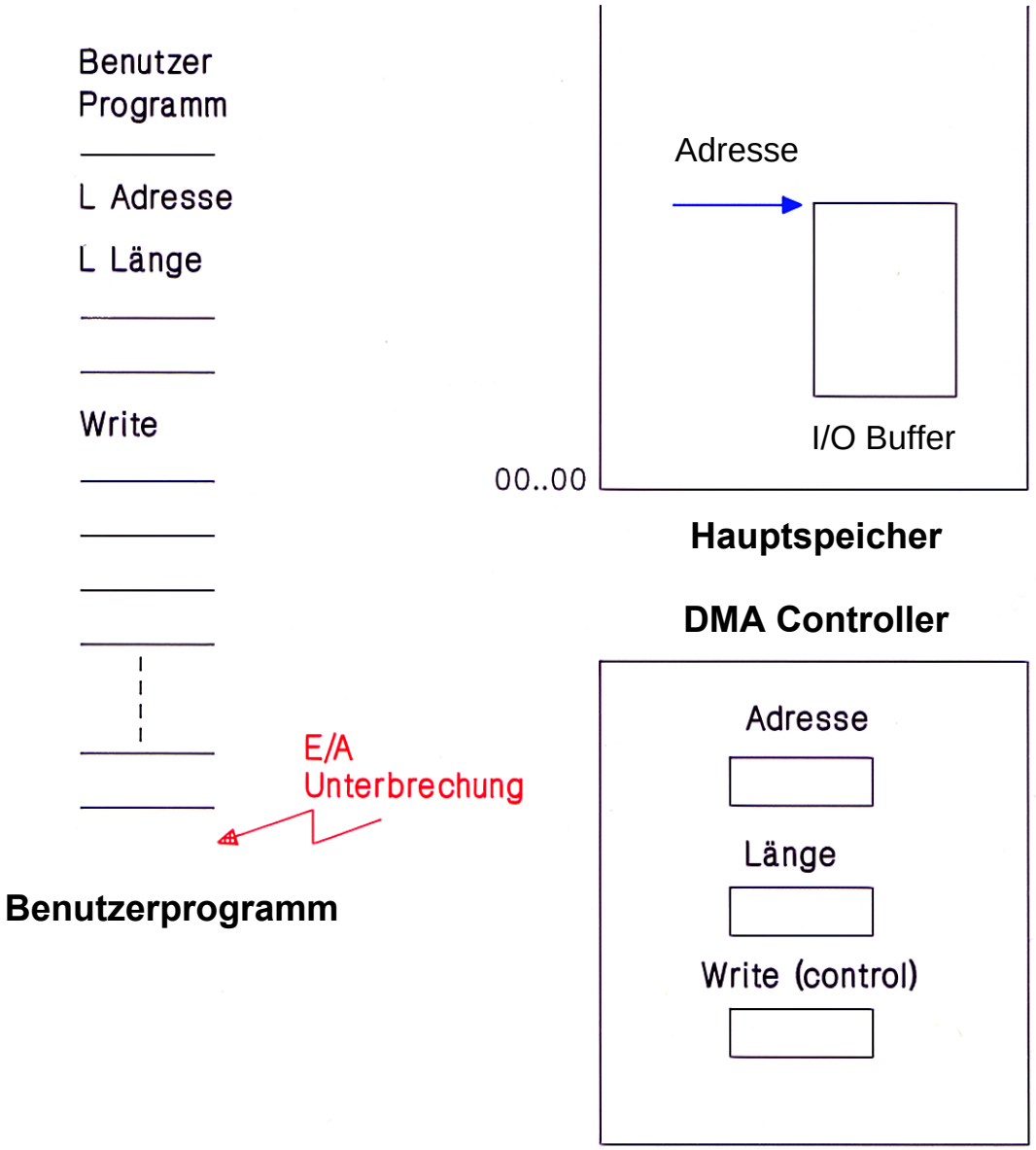
Aufgaben der DMA-Steuerung:

- **Adressieren des Hauptspeichers durch Adressfortschaltung**
- **Adressieren der Geräteschnittstelle**
- **Steuerung der Buszugriffe für Lesen oder Schreiben**
- **Zählen der übertragenen Bytes**
- **Rückmelden an CPU**

Das SADR enthält die Hauptspeicheradresse des zu übertragenenden Bytes. Für die Übertragung eines größeren Datenblocks wird das SADR automatisch hochgezählt. Das BCR enthält die Länge des zu übertragenden Datenblocks.

Für einen Übertragungsvorgang werden DMA-Steuerung und I/O Controller (I/O adapter card) von der CPU initialisiert, z.B. Laden der Steuer(Control)- und Adressregister.

Unterbrechungsgesteuerte Ein/Ausgabe



Die zu übertragenden Daten befinden sich in einem als I/O Buffer bezeichnetem Bereich des Hauptspeichers. Das Anwendungs-programm definiert die Adresse und Länge des I/O Buffers,, und führt z.B. einen WRITE Befehl aus.

Der DMA Controller übernimmt Adresse und Länge in in seine SADR und BCR Register und setzt sein Control Register auf Write.

Danach liest er automatisch und unabhängig von der CPU die Daten aus dem I/O Puffer und überträgt sie zum Plattenspeicher. Nach der Übertragung eines jeden Bytes wird der Length Count in dem BCR Register (Byte Count Register) heruntergezählt und das SADR Register wird inkrementiert. Die CPU arbeitet während dieser Zeit unabhängig weiter.

Wenn der Wert in dem Längenregister den Wert 0 erreicht, ist die I/O Operation beendet. Dies wird der CPU über eine I/O Unterbrechung mitgeteilt.