

# **Enterprise Computing**

## **Einführung in das Betriebssystem z/OS**

**Prof. Dr. Martin Bogdan**  
**Dr. rer. nat. Paul Herrmann**  
**Prof. Dr.-Ing. Wilhelm G. Spruth**

**WS 2009/2010**

**Teil 5**

**Ein/Ausgabe Verarbeitung**

# Literatur

**zSeries Connectivity Handbook  
Redbook, May 2003  
SG24-5444-03**

## **Ein/Ausgabe Performance**

**Das Leistungsverhalten in großen kommerziellen C/S Systemen wird in der Regel weniger durch die CPU Geschwindigkeit und mehr durch die Leistungsfähigkeit der Speicherverwaltung und des I/O Systems bestimmt.**

**Es ist allerdings sehr schwierig das I/O Leistungsverhalten zu charakterisieren.**

**Eine Messgröße ist die gesamte maximale I/O Datenrate. Eine Angabe hierüber enthält das Februar 1996 Heft der Zeitschrift „Manufacturing Systems“. Hiernach konnte das S/390 I/O Subsystem 1,000 bis 20,000+ MByte/Minute übertragen. Sehr große UNIX Systeme konnten 2 bis 100 Mbyte/Minute übertragen.**

**A z10 Book hat eine maximale I/O Datenrate von 96 GByte/s pro Book. Für ein 4 Book System wäre das eine theoretische maximale I/O Datenrate von 384 GByte/s, die allerdings in der Praxis nie erreicht werden wird.**

# Moderne Plattenspeicher-Anschlussarten

Seit über 15 Jahren entwickeln sich zwei unterschiedliche Plattenspeicher-Anschlussarten ATA (IDE) und SCSI mit unterschiedlichen Einsatzschwerpunkten nebeneinander her. Daran wird sich auch in Zukunft nichts ändern. Während früher parallele Anschlussarten benutzt wurden, werden moderne Plattenspeicher heute fast ausschließlich über ein serielles Protokoll angeschlossen. Die wichtigsten Protokolle sind:

- **SATA (serial ATA)**                      Nachfolger für parallel ATA bzw. IDE
- **SAS (serial attached SCSI)**        Nachfolger für parallel SCSI, Kupferkabel
- **FC SCSI**                                Fibre Channel SCSI, Glasfaserkabel
- **iSCSI**                                    Internet SCSI (Ethernet Protokoll)

IDE ist die Abkürzung für „Integrated Drive Electronics“. ATA (AT Bus Adapter) ist eine andere Bezeichnung für IDE. SCSI ist die Abkürzung für „Small Computer System Interface“.

**SATA (serial ATA) dominiert beim PC und anderen Arbeitsplatzrechnern. Die verschiedenen SCSI Arten dominieren bei Servern. Serial attached SCSI und SATA benutzen die gleichen Kupferkabeltypen. Fibre Channel SCSI (FC SCSI) benutzt in der Regel Glasfaserkabel, und kann sowohl für Punkt-zu-Punkt Verbindungen als auch als FC-AL Version (Fibre channel arbitrated Loop) eingesetzt werden. Beim Anschluss einer größeren Anzahl von Plattenspeichern an einen Server ist der Fibre Channel das dominierende Protokoll.**

**Seagate z.B. bietet die Barracuda Familie von Plattenspeichern mit der SATA und der SAS Schnittstelle an; Speicherkapazität bis zu 2 TByte. Für „Mission Critical Applications“ ist die Seagate Cheetah Familie von Plattenspeichern mit einer 4-Gbit/s Fibre Channel SCSI Interface verfügbar; Speicherkapazität bis zu 600 GByte. 2008 waren die Werte 1,5 GByte und 400 GByte. Cheetah Plattenspeicher sind für Anwendungen vorgesehen, „**where system availability and reliability is of utmost importance**“.**

**Letzteres ist vor allem bei Mainframes gegeben. Dafür wird die deutlich geringere Speicherkapazität in Kauf genommen. In anderen Worten: Plattenspeicher in Ihrem PC haben pro Einheit in der Regel eine deutlich höhere Speicherkapazität als die Plattenspeicher eines Mainframes.**

# **FATA**

## **Fibre Channel ATA**

**In der Praxis verwenden SCSI Platten bessere mechanische und elektronische Komponenten als ATA Platten. Dies bewirkt**

- **schnellere Zugriffszeiten**
- **höhere Zuverlässigkeit**
- **deutlich höhere Kosten**

**Aus Zuverlässigkeitsgründen verwenden SCSI Platten selten die neueste Plattenspeichertechnologie. Dies ist einer der Gründe, warum für einen PC Platten mit einer höheren Speicherkapazität erhältlich sind als dies im Mainframe Bereich der Fall ist.**

**FATA-Laufwerke sind ATA-Plattenlaufwerke für einen Fibre Channel (FC) Anschluss. Sie arbeiten mit den mechanischen Komponenten von ATA-Festplatten, jedoch mit vorgeschalteter FC-Schnittstelle. In anderen Worten, es sind SATA Platten, bei denen die elektrische serielle ATA Schnittstelle durch eine Fibre Channel Schnittstelle ersetzt wurde. Die FATA-Technologie hat den Vorteil, dass sie in einer Mischung mit anderen FC-Laufwerken betrieben werden können.**

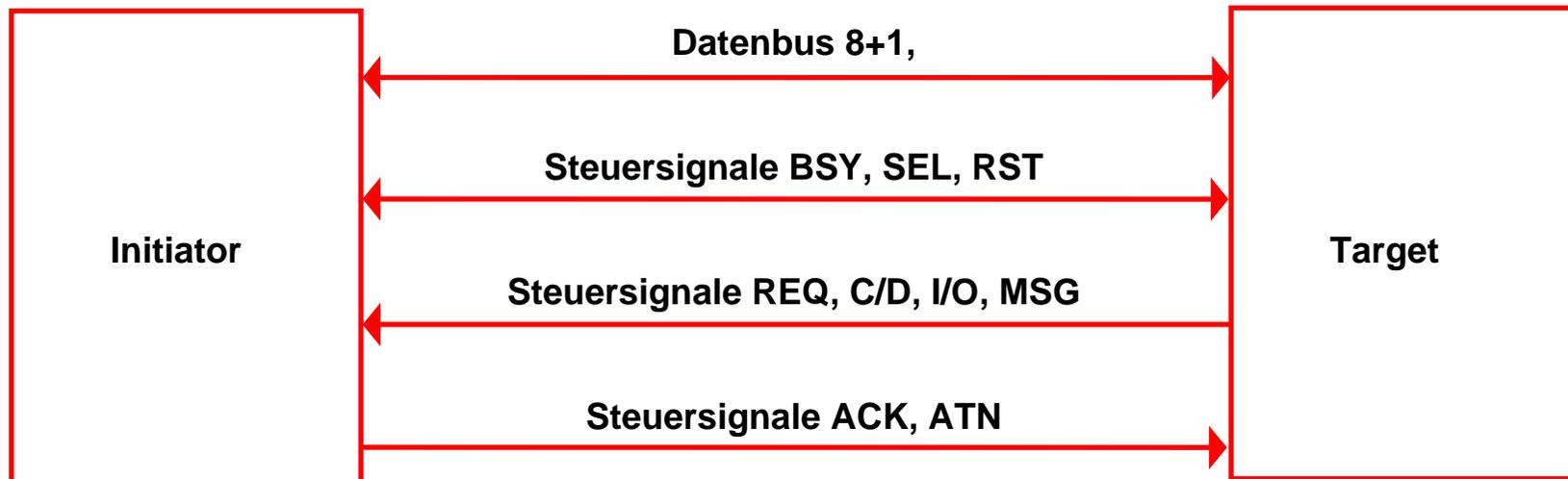
**FATA Platten werden auch als Nearline-Platten bezeichnet. Sie werden im Großrechnerbereich dann eingesetzt, wenn Zuverlässigkeit und Zugriffszeit weniger wichtig sind, z.B. um Bilddateien (Images) zu archivieren.**

# Historische Entwicklung des Peripherie-Busses

S/370 Block Multiplex Channel



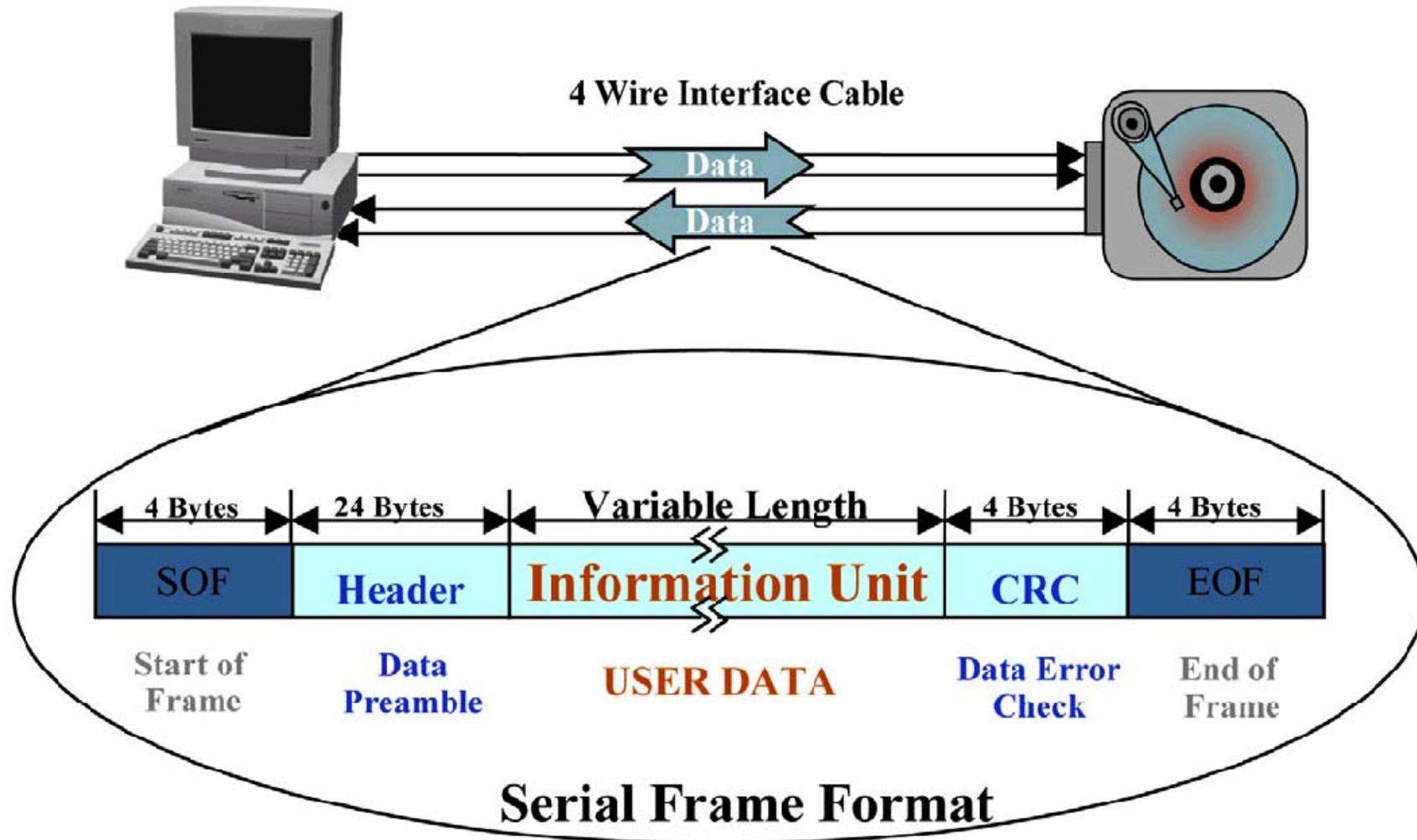
Die ursprüngliche S/360-Kanal-Schnittstelle wurde schon frühzeitig unter dem Namen *OEMI (Original Equipment Manufactures Interface)* veröffentlicht. Zahlreiche Hersteller haben E/A-Geräte mit dieser Schnittstelle hergestellt. OEMI wurde mit einigen Abänderungen unter dem Namen SCSI standardisiert. Dabei wurde die Spezifikation der SCSI-Kabel gegenüber den OEMI-Kabeln geändert, weil bei dem OEMI-Kabel bezüglich Abschirmung und Steckverbinder ein sehr hoher, kostenintensiver Hardware-Aufwand betrieben wird.



## OEMI und SCSI BUS

Die logische OEMI und SCSI Interface bestand aus einem 8Bit (plus Parity) Datenbus und einer Reihe von teils unidirektionalen, teils bidirektionalen Steuersignalen. Die OEMI Schnittstelle wird heute von IBM als „Parallel Channel“ bezeichnet. Später entwickelten sich beide Schnittstellen unabhängig voneinander weiter. Der ursprünglich 8 Bit breite Datenbus der SCSI-1 Schnittstelle wurde auf 16 und später 32 Bit verbreitet. Anschließend entstanden serielle Versionen, die bei IBM zu den Glasfaser-gestützten ESCON und dann der FICON Kanälen führten. Bei SCSI entstanden die Serial SCSI und die FC SCSI Versionen.

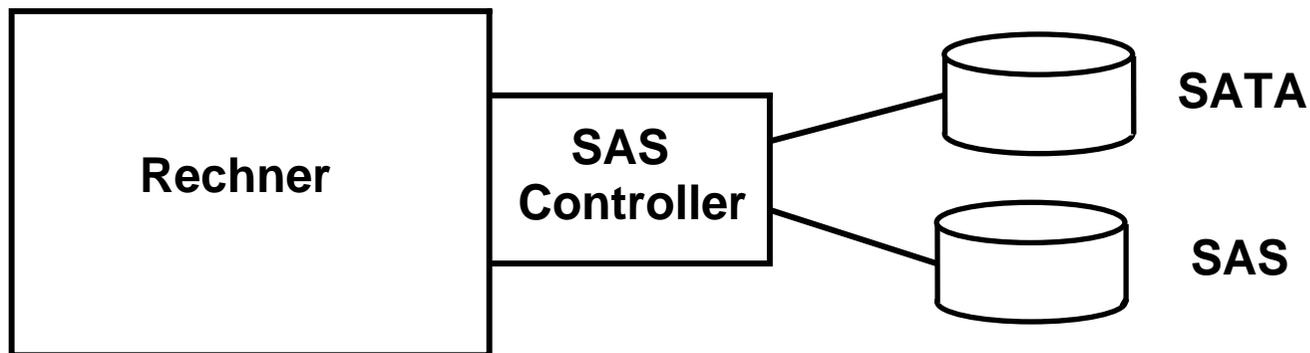
Wegen der höheren Anforderungen im Großrechnerbereich hatten die IBM Kanäle immer einen deutlich höheren Funktionsumfang als die SCSI Schnittstellen. Auch heute kann der mit einem FICON Netzwerk erzielbare Durchsatz durch ein SCSI Netzwerk nicht erreicht werden.

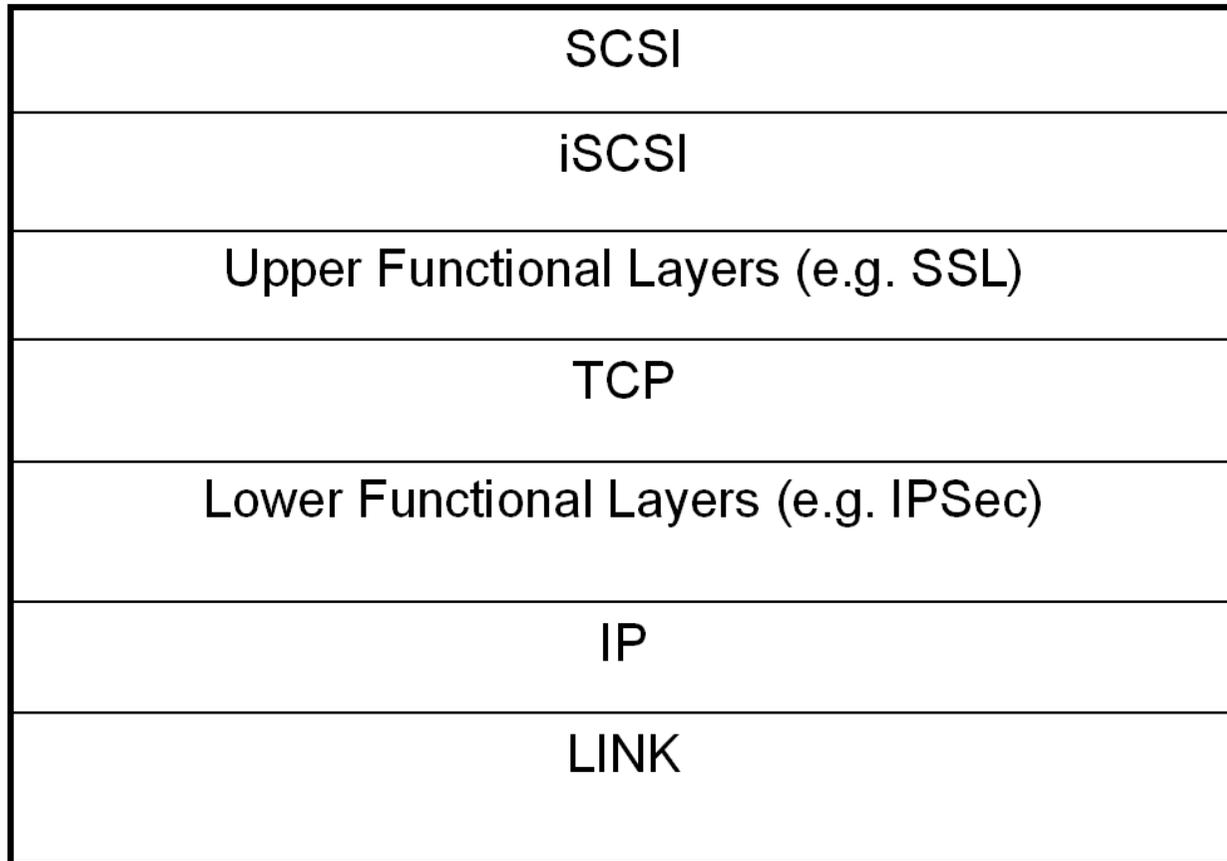


Daten werden bei SAS (serial attached SCSI) und Fibre Channel SCSI seriell im Serial-Frame-Format übertragen. Fibre Channel Protocol (FCP) ist das Schnittstellenprotokoll von SCSI über den Fibre Channel.

**SATA-(Serial-ATA) Endgeräte lassen sich an SAS-(serial attached SCSI) Controller anschließen. Der Grund liegt darin, dass SAS neben dem Serial SCSI Protocol (SSP) und dem SCSI Management Protocol (SMP) ein drittes Protokoll beherrscht, das SATA Tunneling Protocol (STP). Es verpackt die SATA-Befehle in eigene Datenpakete. Außerdem erkennt der Host Bus Adapter (HBA) in der Initialisierungsphase anhand von definierten Signalen, welche Sprache das Target-Gerät spricht und benutzt dann das entsprechende Protokoll.**

**Jeder RAID-System-Hersteller bietet auch Fibre Channel - to - SAS RAID-Controller an**





**iSCSI**

# iSCSI

**Das iSCSI Protokoll (internet SCSI) erlaubt es Klienten (als Initiators bezeichnet), SCSI commands (CDBs) an SCSI Plattenspeicher oder SCSI Bänder zu senden. Dies ist ein populäres Storage Area Network (SAN) Protokoll. Es erlaubt eine Konsolidierung von Speichergeräten in einem Rechenzentrum. Gleichzeitig behalten Server (z.B. Database und Web server) die Illusion von lokal angeschlossenen Speichergeräten. Im Gegensatz zu Fibre Channel, welches Spezial-Verkabelung erfordert, kann iSCSI über große Distanzen mit einer existierenden Netzwerk Infrastruktur betrieben werden.**

**iSCSI verwendet TCP/IP (typisch TCP ports 860 und 3260). Es erlaubt zwei Hosts, die über ein IP Netzwerk verbunden sind, SCSI Commands auszuhandeln und zu benutzen. iSCSI implementiert hiermit ein Storage Area Network (SAN). Im Gegensatz zu anderen SAN Protokollen erfordert iSCSI keine eigene Verkabelung; es kann über eine existierende Switching und IP Infrastruktur (z.B. Ethernet) betrieben werden. Deshalb wird iSCSI oft als eine low-Cost Alternative zu Fibre Channel gesehen, welches eine eigene Infrastruktur erfordert.**

**In der Praxis wird iSCSI fast ausschließlich benutzt, um Server mit Storage Arrays zu verbinden.**

# **Fibre Channel (FC)**

**Fibre Channel ist für serielle, kontinuierliche Hochgeschwindigkeitsübertragung großer Datenmengen konzipiert worden. Die erreichten Datenübertragungsraten liegen heute bei 4 Gbit/s und 8 Gbit/s, was im Vollduplex-Betrieb für Datentransferraten von 800 MB/s ausreicht. Als Übertragungsmedium findet man gelegentlich Kupferkabel (hauptsächlich innerhalb von Storage-Systemen; überbrückt bis zu 30 m), meistens aber Glasfaserkabel. Letzteres erfolgt meist zur Verbindung von Rechnern mit Storage-Systemen oder aber von Storage-Systemen untereinander. Hierbei werden Entfernungen bis zu 10 km überbrückt. Der Zugriff auf die Festplatten erfolgt blockbasiert.**

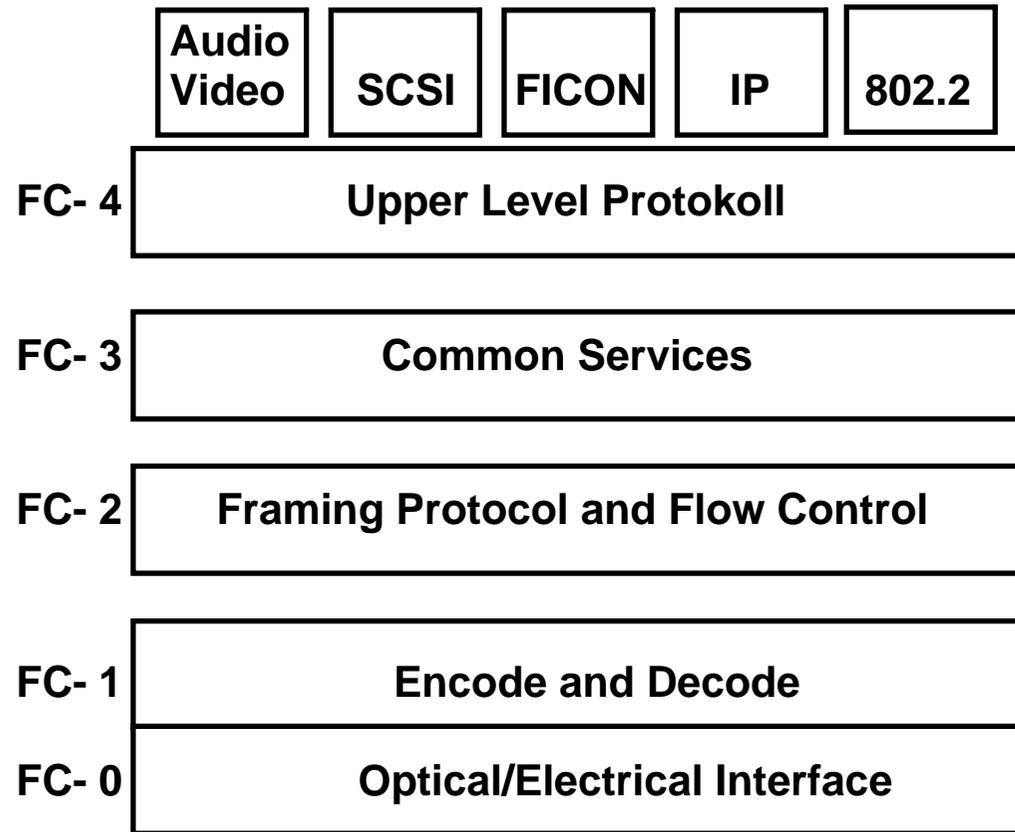
**Es können generell zwei Arten von Fibre-Channel-Implementierungen unterschieden werden, die Switched Fabric, die meist als Fibre Channel, oder kurz FC-SW, bezeichnet wird und die Arbitrated Loop, kurz als FC-AL bekannt.**

## **Zwei Arten des Fibre Channel**

**Bei der Fibre Channel-Switched Fabric werden Punkt-zu-Punkt-Verbindungen (Point To Point) zwischen den Endgeräten geschaltet. Beim Fibre Channel-Arbitrated Loop handelt es sich um einen logischen Bus, bei dem sich alle Endgeräte die gemeinsame Datenübertragungsrate teilen.**

**Bei der Fibre Channel-Switched Fabric handelt es sich um die leistungsfähigste und ausfallsicherste Implementierung von Fibre Channel. In den meisten Fällen ist Switched Fabric gemeint, wenn nur von Fibre Channel gesprochen wird. Im Zentrum der Switched Fabric steht der Fibre Channel Switch oder der Director. Über dieses Gerät werden alle anderen Geräte miteinander verbunden, so dass es über den Fibre Channel Switch möglich wird, direkte Punkt-zu-Punkt-Verbindungen zwischen je zwei beliebigen angeschlossenen Geräte zu schalten.**

**FC-AL erlaubt es, bis zu 127 Geräte an einem logischen Bus zu betreiben. Dabei teilen sich alle Geräte die verfügbare Datenübertragungsrate (bis 4 GBit/s). Die Verkabelung kann sternförmig über einen Fibre Channel Hub erfolgen. Es ist auch möglich, die Geräte in einer Schleife (Loop) hintereinander zu schalten (Daisy Chain), da viele Fibre-Channel-Geräte über zwei Ein- bzw. Ausgänge verfügen. Dies ist z.B. beim IBM DSS 8000 Enterprise Storage Server der Fall.**



## **Fibre Channel Architektur**

# **Fibre Channel Architektur**

**Die Fibre Channel Architektur verwendet ein Schichtenmodell vergleichbar mit (aber unabhängig von) den TCP/IP oder OSI Schichtenmodellen. Die unterste Schicht verwendet in den meisten Fällen optische Kabel. Wichtig ist besonders die oberste Schicht FC – 4.**

**Hierüber ist es möglich, unterschiedliche Protokolle zu betreiben. FC SCSI ist eine serielle Form des SCSI Protokolls, die über Fibre Channel Verbindungen erfolgt. (Das als „Serial SCSI“ bezeichnete Protokoll benutzt keinen Fibre Channel). FICON ist das universell von Mainframes eingesetzte Protokoll, um Rechner miteinander und mit I/O Geräten zu verbinden.**

# Fibre Channel Schichtenmodell

Fibre Channel ist ein Schichten-Protokoll und besteht aus 5 Lagen:

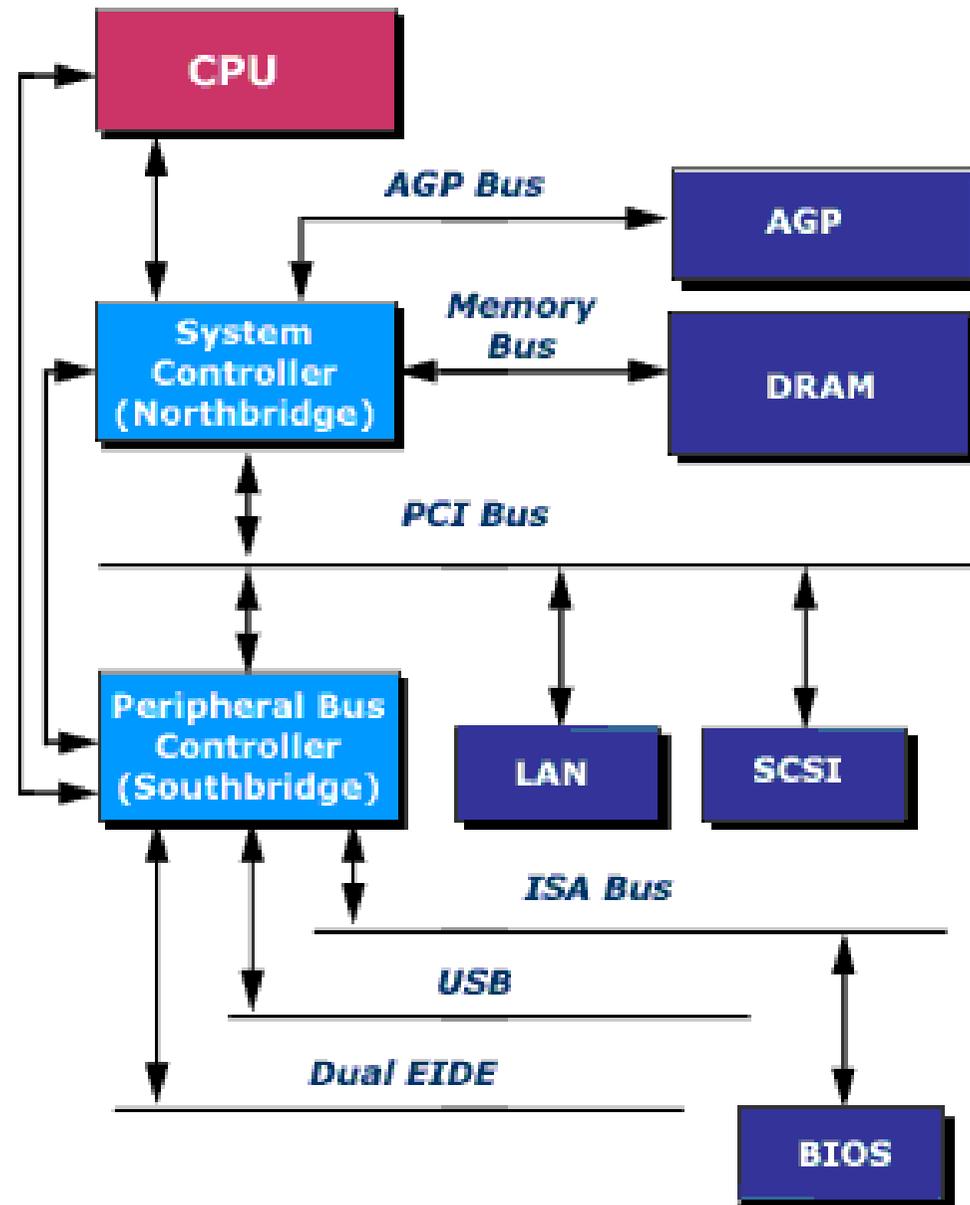
- **-FC0** Die physical layer beschreibt Kabel Fiber Optics, Konnektoren, Pinouts usw.
- **-FC1** Die data link layer implementiert das the 8b/10b Encoding und Decoding der Signale.
- **-FC2** Die network layer, definiert durch den FC-PI-2 standard, ist das Kern Element des Fibre Channel.
- **-FC3** Die common services layer, ist für Erweiterungen vorgesehen, und könnte in Zukunft Funktionen wie Encryption oder RAID implementieren.
- **-FC4** Die Protocol Mapping layer. In dieser Schicht werden Protokolle wie FC SCSI oder FICON abgebildet.

# **Fibre Channel Datenraten**

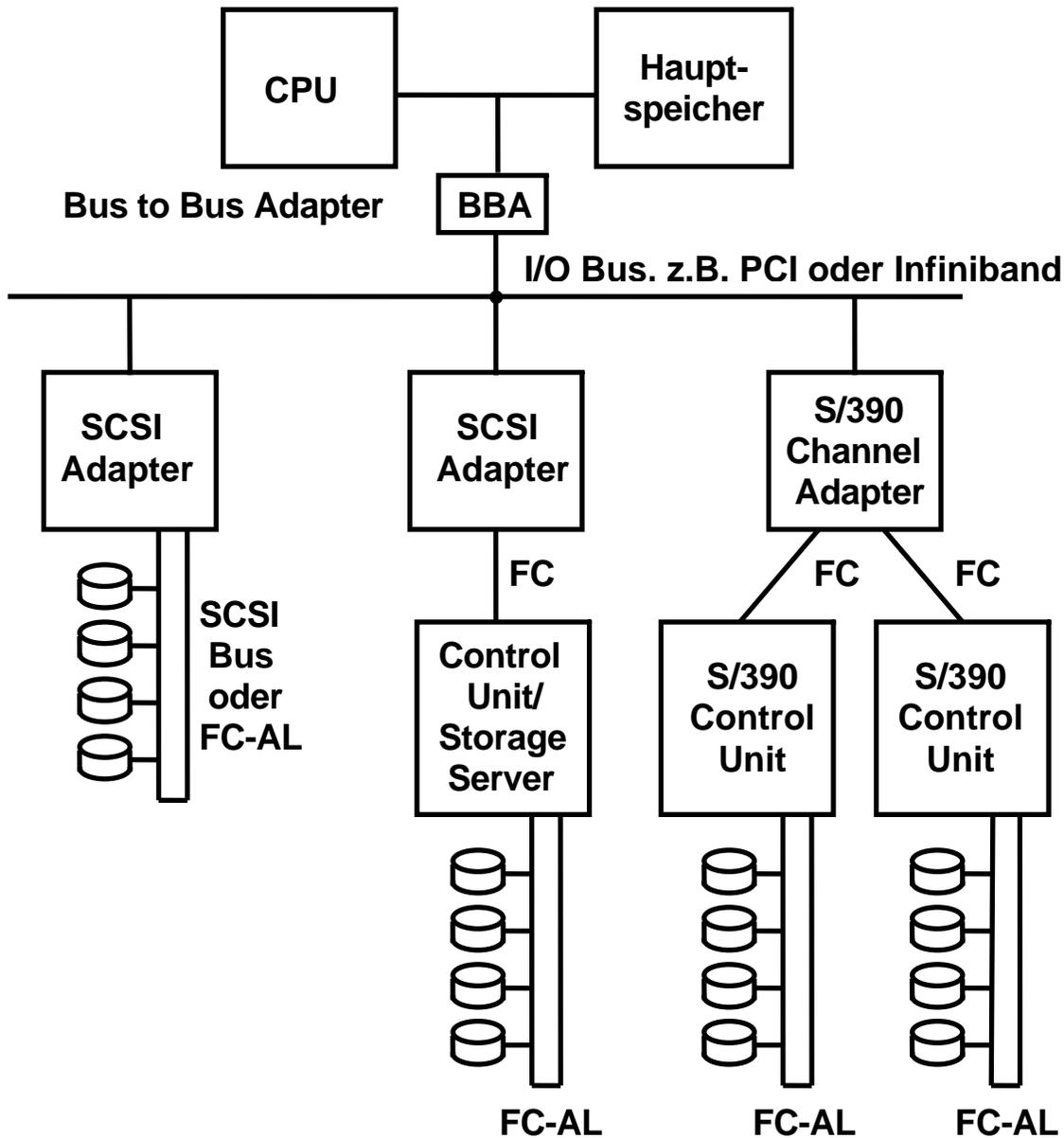
**Fibre Channel Produkte mit 1 Gbit/s, 2 Gbit/s und 4 Gbit/s sind verfügbar. Die 4 Gigabit/s Fibre Channel Version (4GFC) ist der derzeitige Standard. An einem 8 Gbit/s Standard wird gearbeitet. Ein 10 Gbit/s Standard wurde verabschiedet, wird aber nur eingesetzt, um Switche miteinander zu verbinden.**

**Fibre Channel, oder FC, ist eine Gigabit/s Netzwerk Technologie, die für Storage Area Networks (SANs) eingesetzt wird. Im Prinzip könnte es als ein general-Purpose Netzwerk für die Übertragung von ATM, IP und anderen Protokollen eingesetzt werden; in der Praxis wurde es fast nur für die Verbindung von Servern mit Plattenspeichern eingesetzt, hauptsächlich mit FC SCSI oder FICON Verbindungen.**

**Fibre Channel Protocol (FCP) ist das Schnittstellenprotokoll von SCSI über den Fibre Channel.**



**I/O Bus Struktur des Personal Computer**

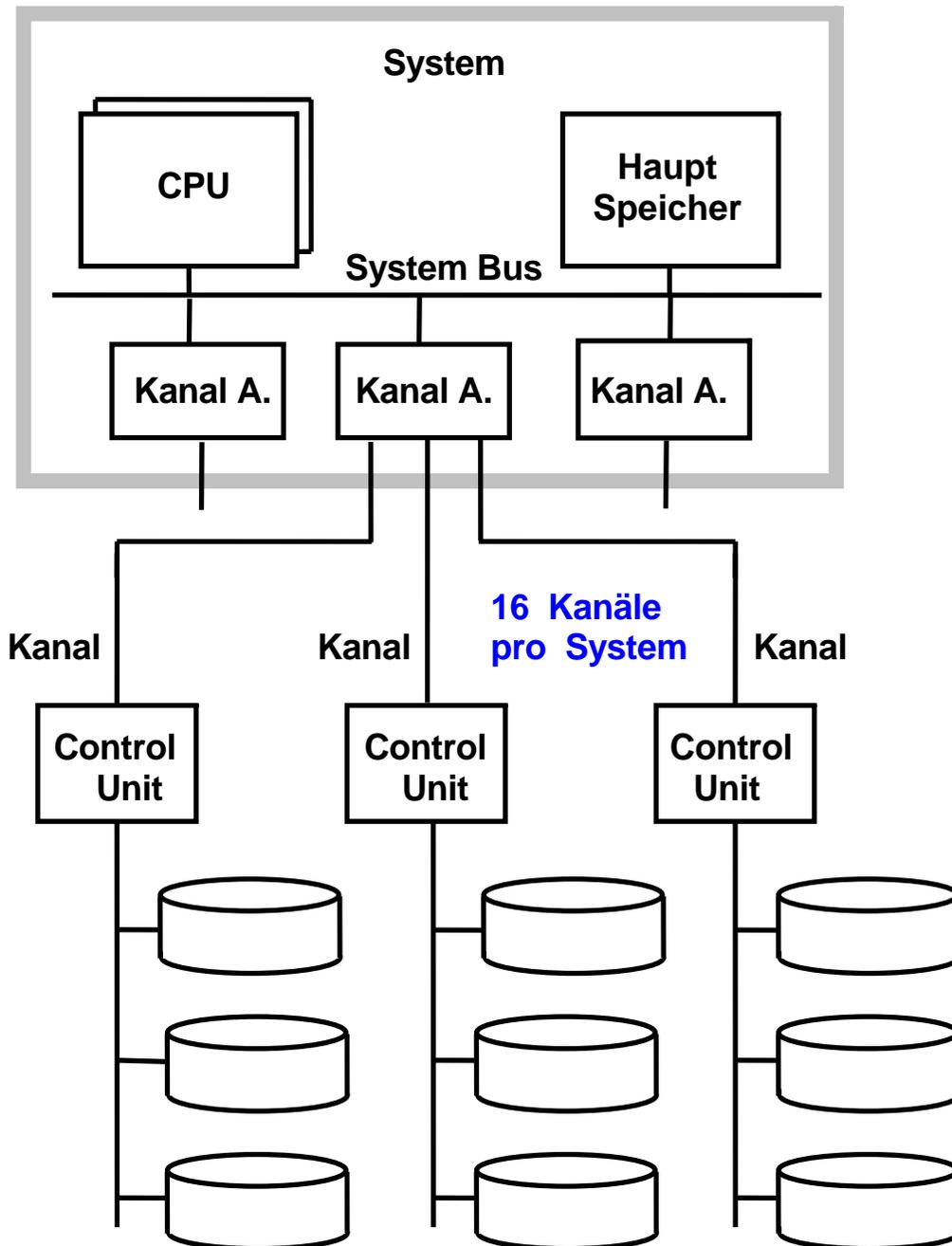


## Plattenspeicher Anschluss Alternativen

Ein SCSI Adapter kann eine oder mehrere SCSI Platten mit dem I/O Bus eines Rechners verbinden. Bei größeren Mengen an Plattenspeichern sind diese über eine Control Unit oder einen Storage Server mit dem SCSI Adapter verbunden.

Mainframes ersetzen den SCSI Adapter durch einen Channel Adapter. Plattenspeicher sind grundsätzlich über Control Units mit dem Channel Adapter verbunden.

FC            Fibre Channel  
 FC-AL       Fibre Channel Arbitrated Loop



## S/360 I/O Konfiguration

Dargestellt ist die ursprüngliche S/360 Konfiguration.

Plattenspeicher sind über Control Units, Kanal-Verbindungskabel und Kanal-Adapter mit dem Hauptspeicher des Systems verbunden. Die Verbindungskabel des „Parallel Channels“ waren bis zu 400 Fuß ( 130 m ) lang. Die Kanal-Adapter konnten mittels DMA direkt auf den Hauptspeicher des Systems zugreifen.

Die Control Unit führte I/O Befehle (Channel Command Words, CCWs) aus, die vom Kanal-Adapter aus dem Hauptspeicher ausgelesen und zwecks ausführung der Control Unit übergeben wurden.

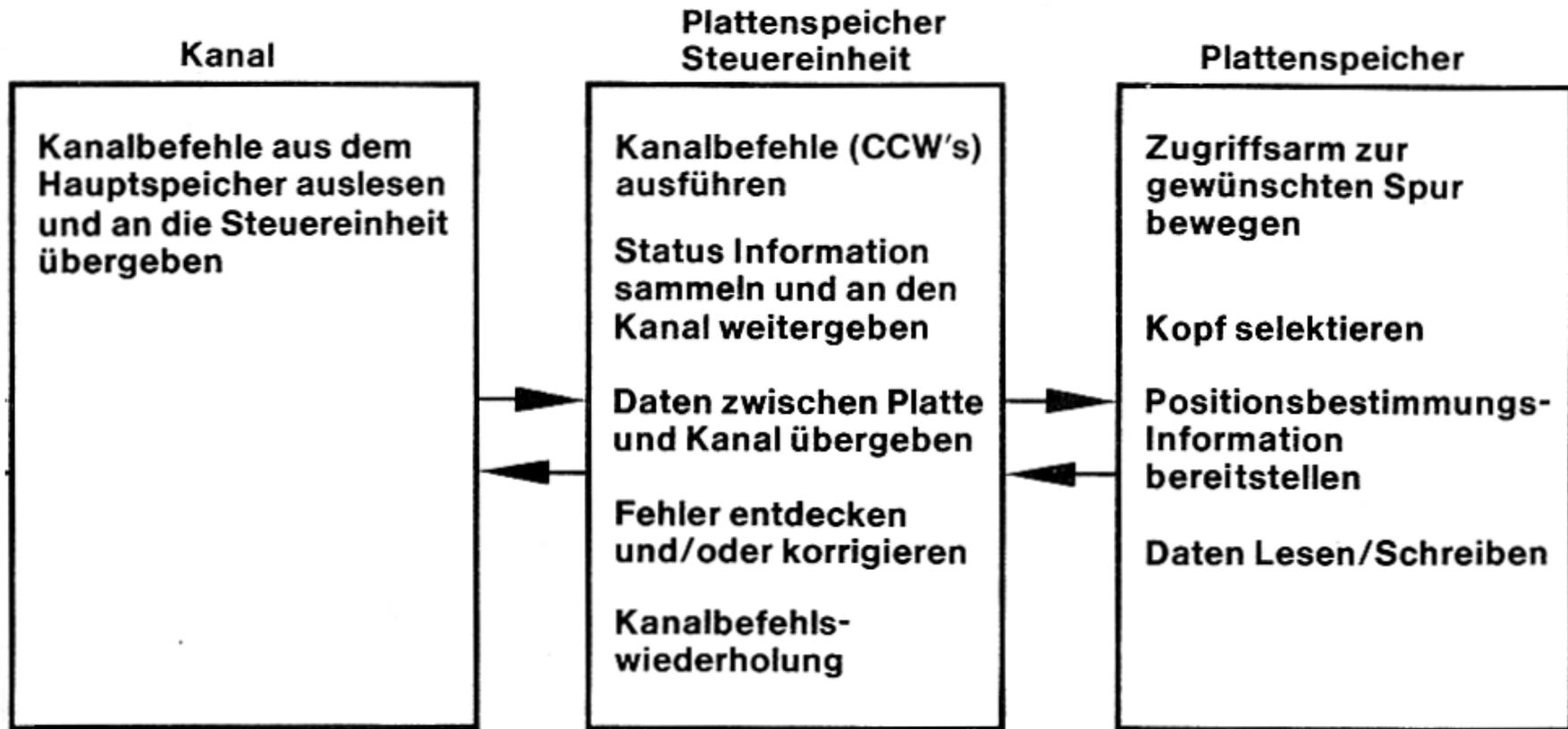
# **Kanal und Control Unit**

**deutsche Begriffe: Steuereinheit bzw. Plattenspeicher Steuereinheit**

**Kanal und Control Unit sind zwei physische Einheiten, die über ein Kanal-Kabel miteinander verbunden sind, und eine logische Einheit darstellen. Die Aufteilung ist erforderlich, weil aus Platzgründen die control Units in einer gewissen Entfernung voneinander und vom Rechner aufgestellt werden müssen.**

**In manchen Fällen ist es möglich (Beispiel OSA Adapter), die Control Unit im gleichen Gehäuse wie die CPUs unterzubringen. In diesem Fall werden Control Unit und Kanal Adapter als eine einzige Baugruppe implementiert.**

**Das Konzept der Control Unit entstand 1953, zeitgleich mit der Einführung des IBM 702 Rechners.**



## Aufgaben der Plattenspeicher-Steuereinheit

Dargestellt ist die Aufgabenaufteilung zwischen Kanal-Adapter, Steuereinheit (Control Unit) und der Plattenspeicher-Elektronik.

Der Kanal (die Kanal-Adapter Karte) ist nur dazu da, um stellvertretend für die Steuereinheit per DMA Daten und Kanalbefehle aus dem Hauptspeicher auszulesen und an die entfernte Steuereinheit weiter zu geben.

# Beispiel I/O Scheduler

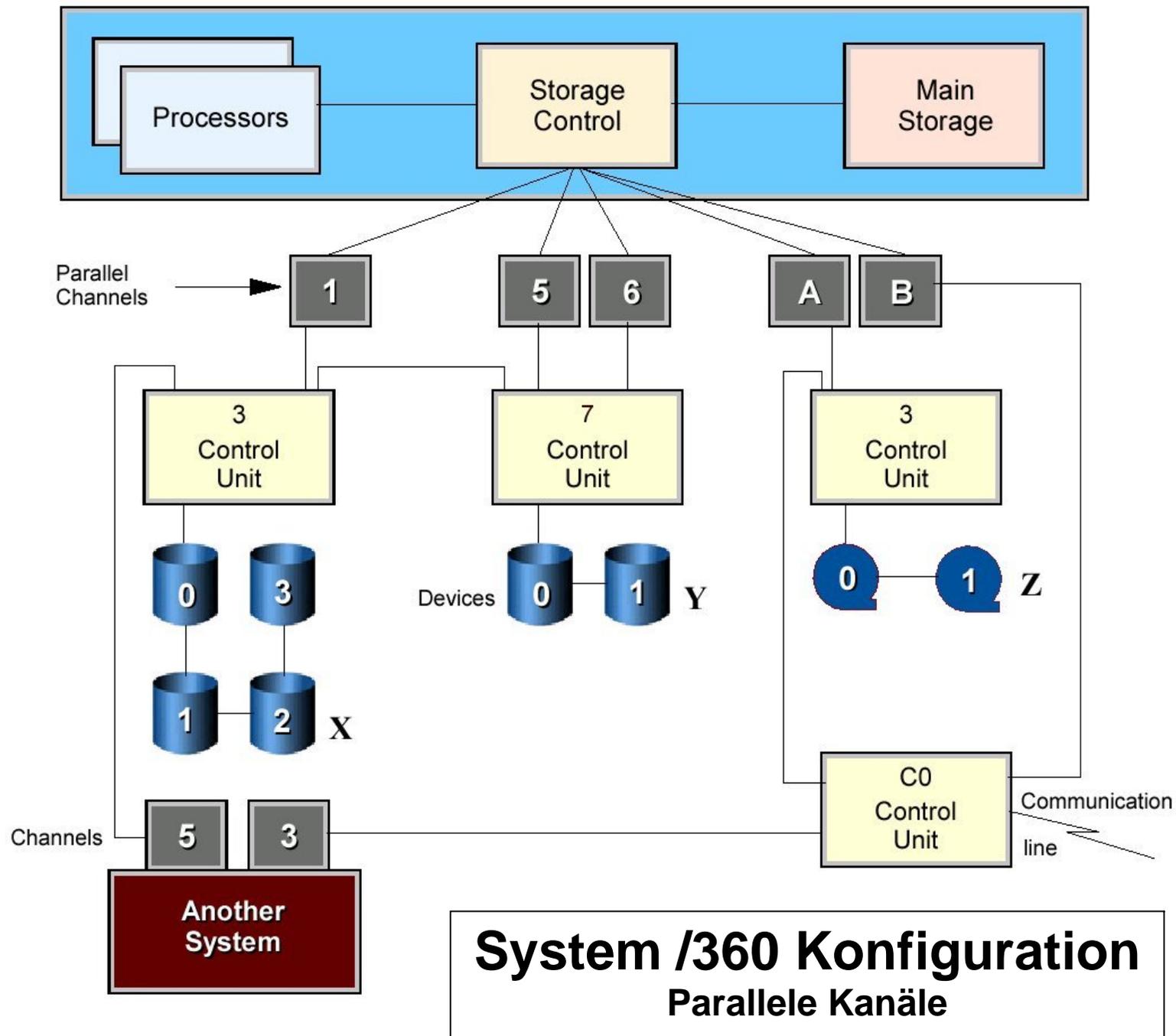
Ein I/O Scheduler bestimmt die Reihenfolge, in der der Zugriffsarm einer DASD mehrere anstehenden I/O Anforderungen abarbeitet. Hierbei werden Algorithmen wie Elevator Disk Arm, Shortest Seek First, usw. eingesetzt.

Unter Linux und zLinux stehen vier I/O Scheduler zur Verfügung: noop, anticipatory, deadline und cfq. Diese sind \_Bestandteile des I/O Drivers.

Ändern lässt sich der Scheduler im laufenden Betrieb pro device dynamisch via:  
`sysX:~ # echo "noop" > /sys/devices/css0/0.0.0004/0.0.2b2b/block/dasda/queue/scheduler.`

Bei einem Mainframe ist dies die Aufgabe der Control Unit (z.B. 3990), und/oder des Enterprise Storage Subsystems. Mit derartigen Verarbeitungsaufgaben wird die CPU nicht belastet; es fehlt auch die entsprechende CCW Funktionalität.

Ohne ein Auslagern derartiger Funktionen wäre ein Mainframe nie in der Lage, den erforderlichen I/O Durchsatz zu erbringen.



Die obige Abbildung zeigt eine mögliche S/360 Konfiguration.

Jeder Kanal, jede Control Unit und jedes I/O Gerät hat eine Adresse in Form einer 24 Bit hexidezimalen Ziffer. Der mit X markierte Plattenspeicher hat die Adresse 132:

1	3	2
---	---	---

Kanal Nummer    Control Unit Nummer    Device Nummer

Betrachten wir den mit Y makierte Plattenspeicher. Auf ihn kann mit den Adressen 171, 571, oder 671 zugegriffen werden, da er mit den Kanälen 1, 5 und 6 verbunden ist. Mehrfache Pfade zu einem I/O Gerät verbessern Performance und Zuverlässigkeit. Wenn eine Anwendung auf Plattenspeicher 171 zugreifen möchte, versucht sie es zunächst über Kanal 1, dann über Kanal 5 und schließlich über Kanal 6.

Die Abbildung mit der System /360 Konfiguration zeigt ein zweites S/360 System („Another System“), welches mit 2 Kanälen mit den Control Units des ersten Systems verbunden ist. Diese gleichzeitige gemeindame Nutzung (sharing) von I/O Devices ist in allen Mainframe Installationen üblich. Magnetband Z hat die Adresse A31 für das erste System, aber Adresse 331 Für das zweite System. Die gemeinsame Nutzung von Plattenspeichern durch zwei Systeme ist kein einfaches Thema. Es existieren Hardware und Software Funktionen die verhindern, dass die gleichen Daten auf einem Plattenspeicher von zwei Systemen zum gleichen Zeitpunkt abgeändert werden.

# **Serielle und parallele Kanäle**

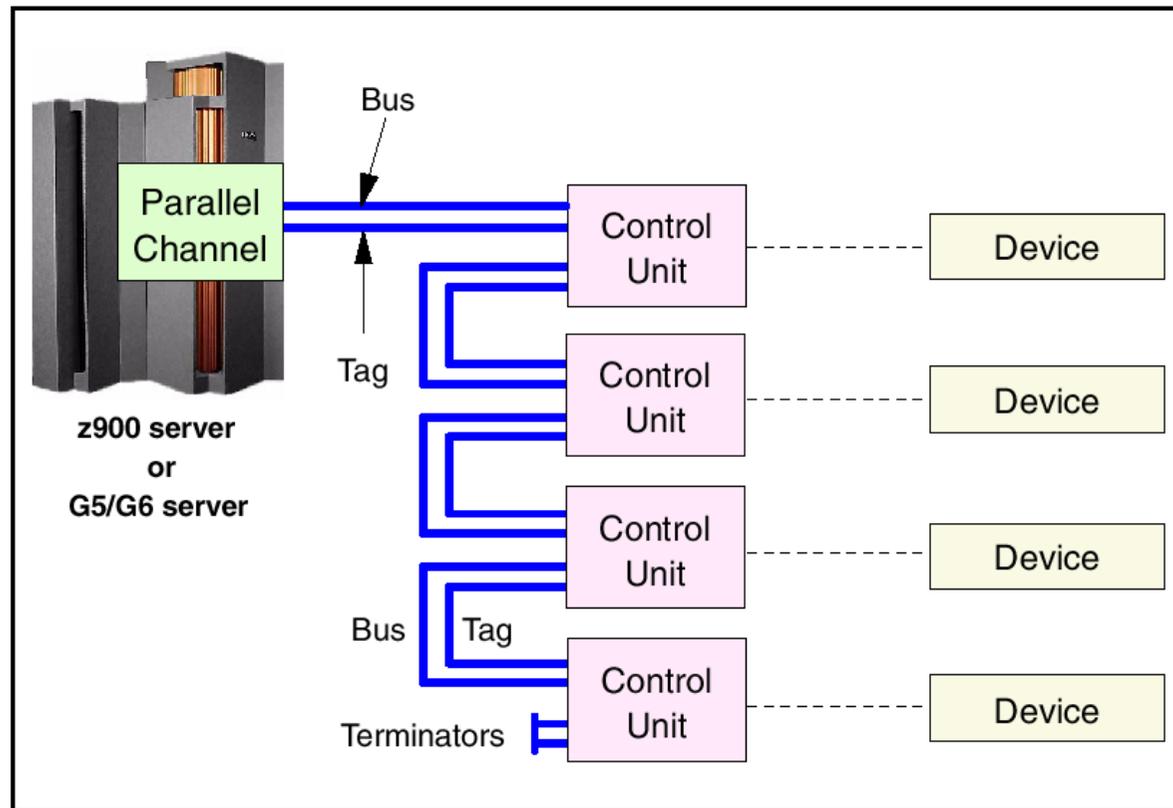
**Der Parallel Channel dominierte die Mainframe I/O Konfigurationen bis zum Anfang dieses Jahrhunderts. Die z10 Rechner unterstützen den Parallel Channel nicht mehr.**

**Serielle Channel haben den älteren parallel Channel abgelöst. Serielle Channel verwenden Glasfaser Kabel und eine Punkt-zu-Punkt Verbindung. Es können Entfernungen bis zu 100 km überbrückt werden.**

**Es existieren zwei Serial Channel Typen:**

- Der (ältere) „ESCON Channel erlaubt Datenraten von 17 MByte/s. Eingeführt 1990**
- Eine FICON Verbindung erlaubt Datenraten von 100, 200 oder 400 MByte/s. Eingeführt 1997 für Magnetbänder, 2001 für Plattenspeicher.**

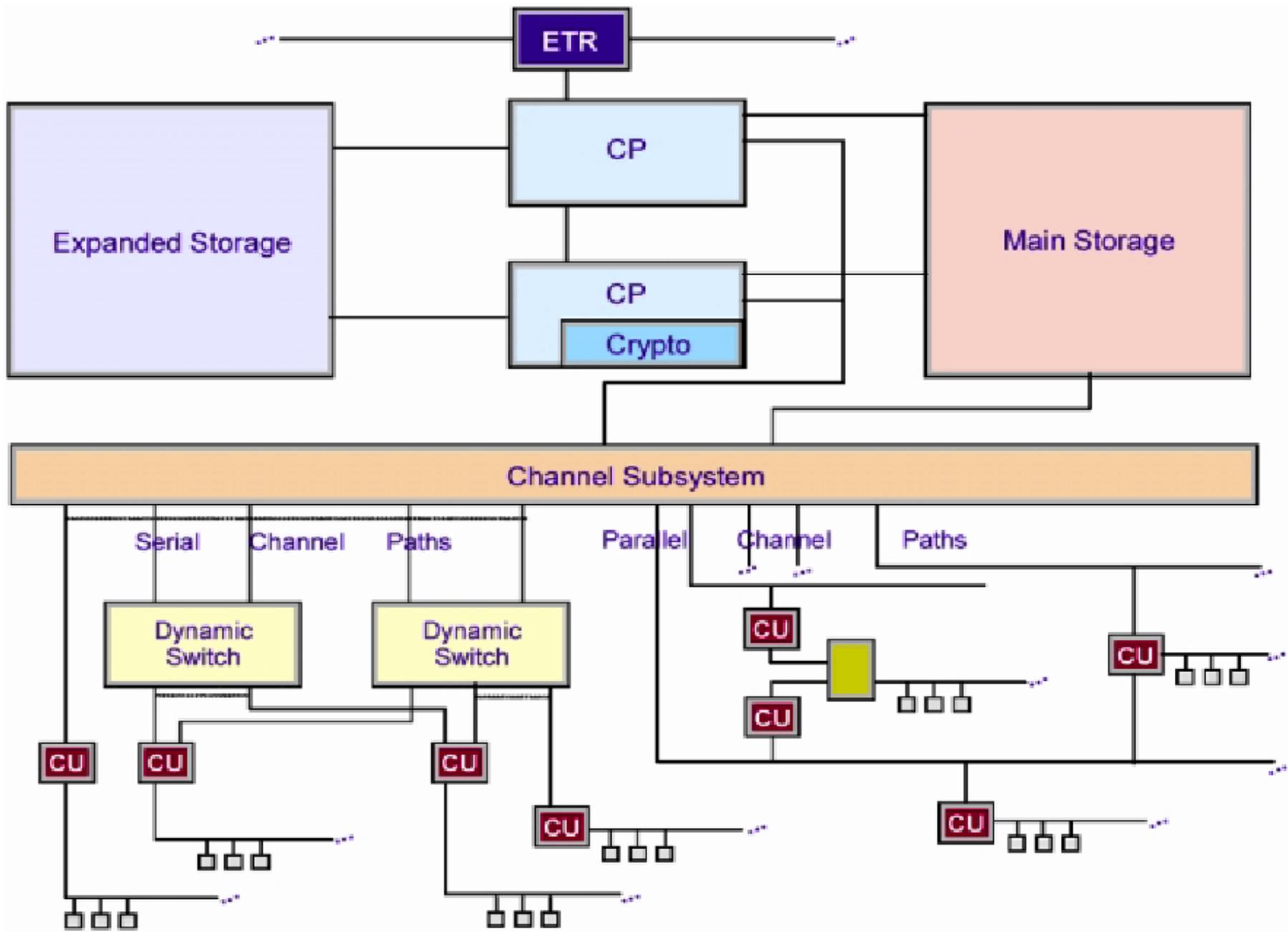
**Serielle Kanäle verwenden Glasfaser- am Stelle von Kupferverbindungen. Sie können deshalb wesentlich größere Entfernungen überbrücken. Außerdem verfügen sie über eine erweiterte I/O Adressierung.**



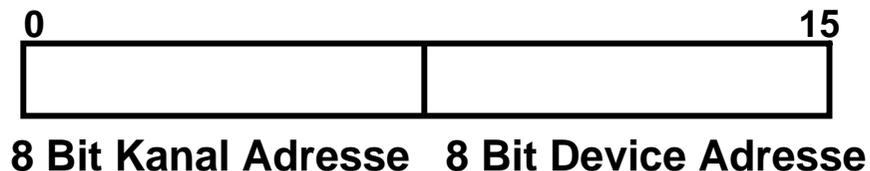
Der S/360 „Parallel“ Channel wird auch als Parallel I/O Interface, bus and tag channel, OEMI channel und als block multiplexor channel bezeichnet. Die Parallel I/O Interface hat Ähnlichkeit mit der parallelen SCSI Interface.

Ein Parallel Channel hat eine Datenrate von max. 4.5MByte/s. und überbrückt Distanzen bis zu 130 m. Der Parallel Channel benutzt zwei Kupfer Kabel: *bus* und *tag*. Ein bus Kabel überträgt Information (ein Byte in jeder Richtung). Daten auf dem tag Kabel definieren die Bedeutung der Information auf dem bus Kabel.

Beim Parallel Channel können mehrere Control Units an einen Kanal Adapter in einer „Daisy Chain“ Konfiguration angeschlossen werden. Eine Control Unit imt einer Parallel I/O Interface hat deshalb zwei Anschluss-Stecker, je einen für die einkommende und für die ausgehende Verbindung.



# Änderungen gegenüber der ursprünglichen S/360 I/O Architektur



Beim Übergang von der S/360 auf die S/370 Architektur (1980) wurde die maximale Anzahl der Kanäle von 16 auf 256 erhöht. I/O devices werden über eine 16 Bit Device Address angesprochen. Diese Adresse besteht aus zwei Feldern. Kanäle haben die Adressen 0 .. 255, und I/O Geräte (Devices) haben die Adressen 0 .. 255.

Beim Übergang auf die S/370 XA Architektur (1983) beschloss man, dass die Kanäle (Channels) in Zukunft Kanalpfade (Channel Path) heißen und durch einen 8 Bit CHPID (Channel Path Identifier) gekennzeichnet werden. In der Umgangssprache werden Channel Path nach wie vor als Kanäle bezeichnet, und auch Experten kennen den Unterschied nicht.

## Mehrfache Pfade zu einem I/O Gerät

System z I/O Geräte werden über Steuereinheiten an das Kanal Subsystem angeschlossen.

Steuereinheiten können über mehr als einen Kanalpfad an das Kanalsubsystem angeschlossen werden und I/O Geräte können an mehr als eine Steuereinheit angeschlossen werden.

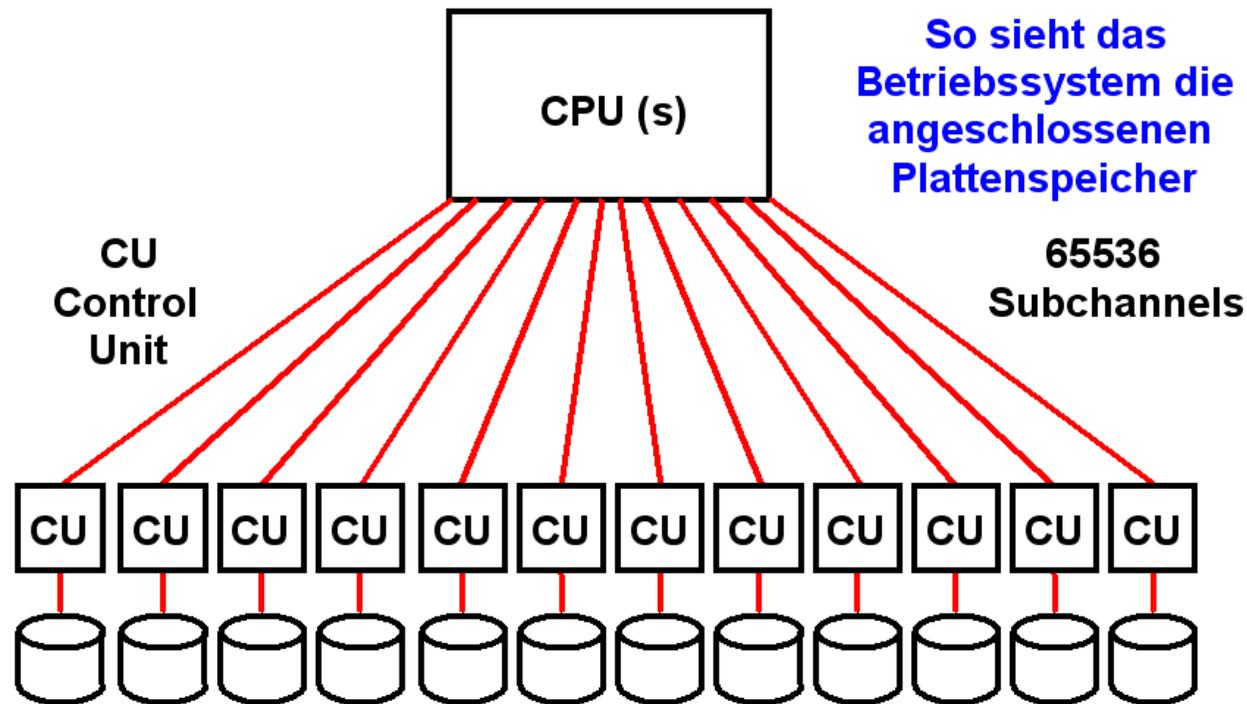
Eine z/OS Instanz (auf einem System kann mehr als eine z/OS Instanz laufen) kann auf ein beliebiges I/O Gerät über bis zu 8 unterschiedliche Kanalpfade zugreifen (und umgekehrt).

Der Zugriffsweg kann dynamisch geändert werden (DPS - Dynamic Path Selection). Eine I/O Operation muss nicht auf dem gleichen Weg abgeschlossen werden, auf dem sie gestartet wurde. Hiermit kann erreicht werden, dass sequentielle und zufallsbedingte Zugriffe sich nicht gegesätzlich beeinträchtigen.

Z.B. angenommen zwei Plattenspeicher, die an die gleiche Steuereinheit angeschlossen sind. Ein Plattenspeicher überträgt einen großen Block sequentieller Daten, während der zweite Plattenspeicher gleichzeitig viele kurze Datenpakete mit einem zufallsbedingten Zugriffsmuster überträgt.

In anderen Worten, es gibt mehrere Wege, auf denen Daten (und Steuerinformation) zwischen Platte und CPU übertragen werden können.

Diese dynamische Weg-Steuerung ist rechenaufwendig. Deswegen belastet man mit dieser Aufgabe nicht das Betriebssystem, sondern überträgt sie einem getrennten I/O Rechner, dem **Channel Subsystem**. Eine Beispiel-Konfiguration ist in der folgenden Abbildung wiedergegeben.



## Vereinfachte I/O Konfiguration aus Sicht des z/OS Betriebssystems

Plattenspeicher können während der Ausführung einer I/O Operation über unterschiedliche Channel Path und Control Unit Adressen angesprochen werden. Das **Channel Subsystem** ist für diese Ansteuerungsoptimierung verantwortlich.

Das Betriebssystem braucht aber eine eindeutige I/O Device Identifizierung. zSeries und S/390 Rechner arbeiten mit einer vereinfachten und standardisierten Sicht der angeschlossenen I/O Struktur (virtuelles I/O Subsystem). Deshalb wird jeder Plattenspeicher über eine 16 Bit (0 .. 65 535) **Subchannel ID** angesprochen. Aus der Sicht des Betriebssystems ist jedes I/O Device über einen dedizierten Subchannel mit dem Rechner verbunden. Die I/O Ansteuerung des Betriebssystem Kernels kennt die Einzelheiten der I/O Konfiguration nicht.

Es ist die Aufgabe des Channel Subsystems, die Subchannel ID in eine physische Adresse umzusetzen.

# Channel Subsystem

## Subchannel

Die CPU und das Betriebssystem interessiert nicht, auf welchem Weg die Daten zum/vom Plattenspeicher übertragen werden.

Es spricht jeden Plattenspeicher (und jedes andere I/O Gerät) über eine logische Adresse an. Diese logische Adresse wird als **Subchannel** bezeichnet. Soweit es das Betriebssystem betrifft, ist jedes einzelne I/O Device über einen eigenen Subchannel mit den CPUs des Rechners verbunden (über eine eigene Control Unit).

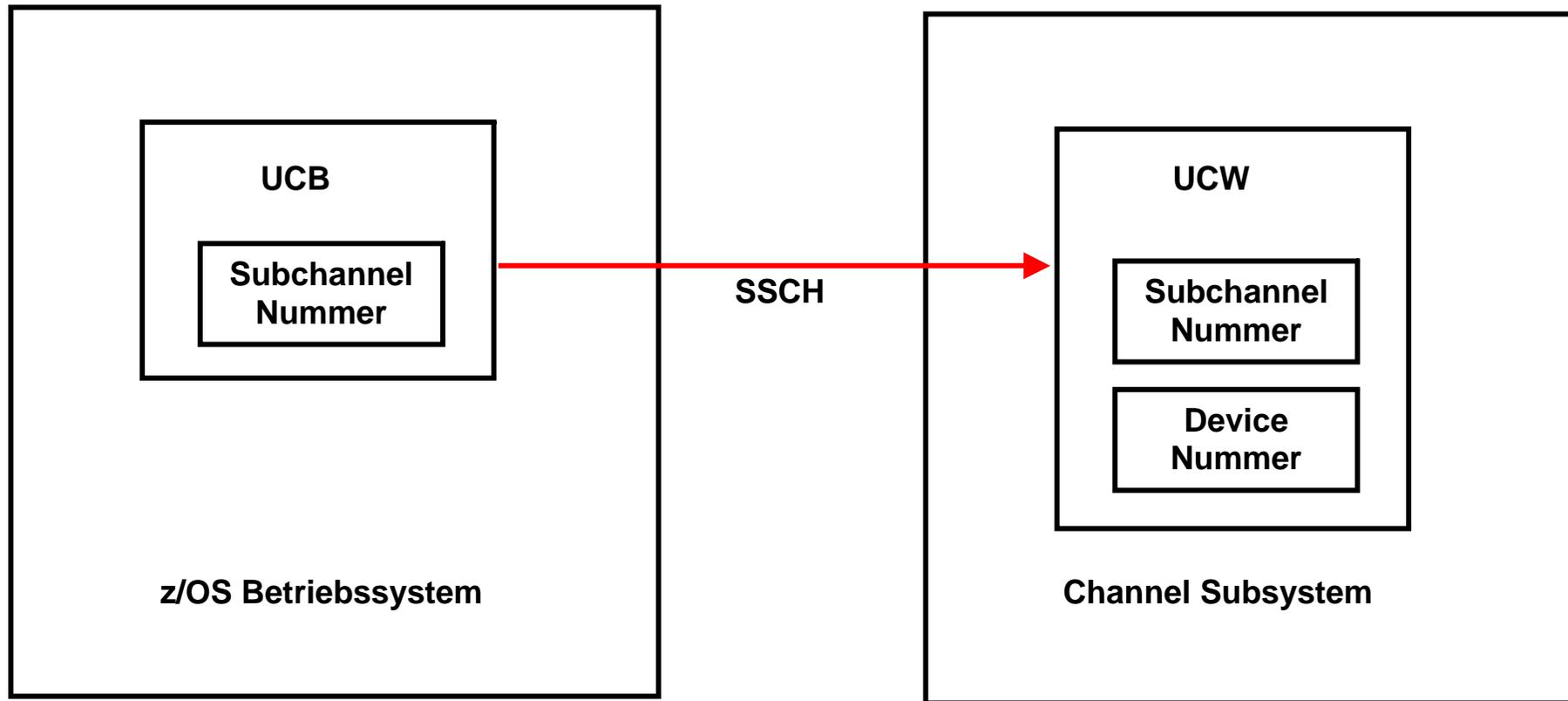
Bitte nicht die Begriffe „Channel Subsystem“ und „Subchannel“ miteinander verwechseln.

Jeder Subchannel wird durch eine 16-Bit Subchannel Nummer repräsentiert.. Es sind also bis zu 65 536 Subchannels (und angeschlossene I/O Devices) pro Channel Subsystem möglich.

Für das Betriebssystem wird jeder Subchannel durch einen Bereich im Hauptspeicher dargestellt, der als UCB (Unit Control Block) bezeichnet wird. Es existiert je ein UCB pro Subchannel. Subchannels und UCBs werden beim Booten (Initial Program Load, IPL) von z/OS initiiert. Der UCB repräsentiert ein I/O Device eindeutig gegenüber einem Anwendungsprogramm und dem Betriebssystem. Er enthält die für die Ausführung einer I/O Operation erforderliche Information. Hierzu enthält er die Subchannel Nummer, die Adresse des Kanalprogramms (in der Form von CCWs), sowie weitere Daten.

Der z/OS Kernel (Supervisor, Basic control Program) führt einen Maschinenbefehl (Start Subchannel, SSCH) aus, der das Channel Subsystem auffordert, eine I/O Operation auszuführen. Damit wird die CPU von der Aufgabe entlastet, die einzelnen Schritte der I/O Operation durchzuführen. Datenverarbeitung und I/O-Verarbeitung werden gleichzeitig und unabhängig voneinander ausgeführt.

Das Channel Subsystem selektiert einen von mehreren möglichen physischen Verbindungswegen von/zum I/O Device. Es testet die Verfügbarkeit des Verbindungsweges und übernimmt prioritäts-Steuerungen.



## Aufgabenaufteilung CPU – Channel Subsystem

Die Ausführung des SSCH (Start Subchannel) Maschinenbefehls bewirkt, dass der angesprochene UCB einschließlich der Subchannel Nummer an das Channel Subsystem übergeben wird.

Das Channel Subsystem unterhält für jeden Subchannel ein dazugehöriges UCW (Unit Control Word). Das UCW enthält die zu der Subchannel Nummer (logische Adresse) dazugehörige Device Nummer (Physische Adresse).

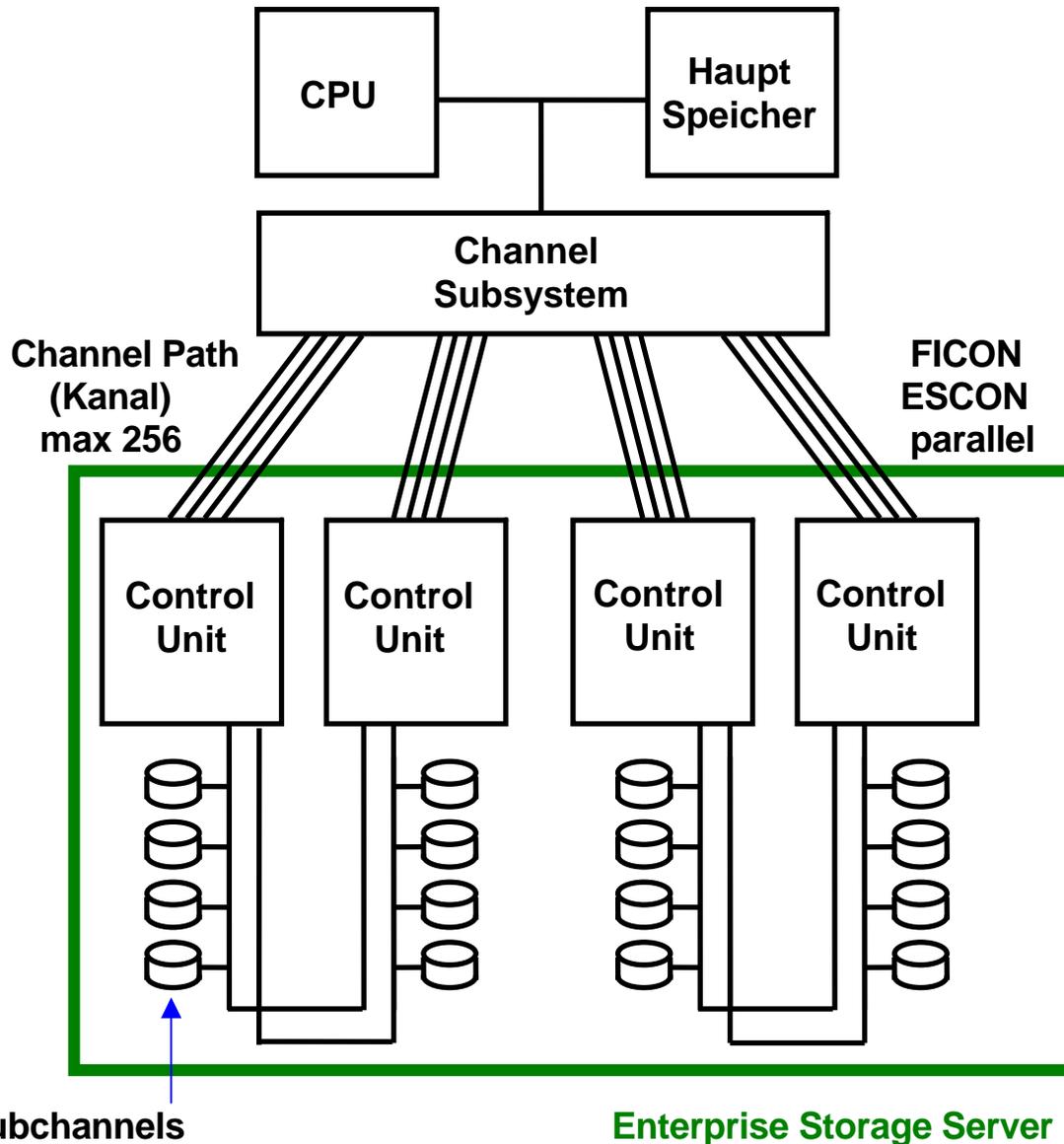
An dieser Stelle sei vorwiegend genommen, dass neben der Subchannel Nummer und der Device Nummer das Channel Subsystem noch eine dritte Adresse benutzt, den Device Identifier. Hierauf wird später im Abschnitt I/O Konfiguration eingegangen.

## Subchannel und Channel Path

Als **Channel Path** wird die physikalische und logische Verbindung zwischen einem System und einer Control Unit bezeichnet. Während einer I/O Operation kann der Datentransfer über einen Subchannel auf unterschiedliche Channel Path abgebildet werden. Ein Channelpath wird durch eine 8 Bit CHPID gekennzeichnet.

(Die Begriffe „Channel Path“ und „Channel“ (Kanal) sind weitestgehend identisch.)

Jedem angeschlossenen Gerät (z.B. Plattenspeicher) ist ein **Subchannel** zugeordnet. Jeder aktive Subchannel wird durch einen Speicherblock innerhalb des Hauptspeichers (UCB) dargestellt und durch eine 16 Bit Adresse gekennzeichnet.

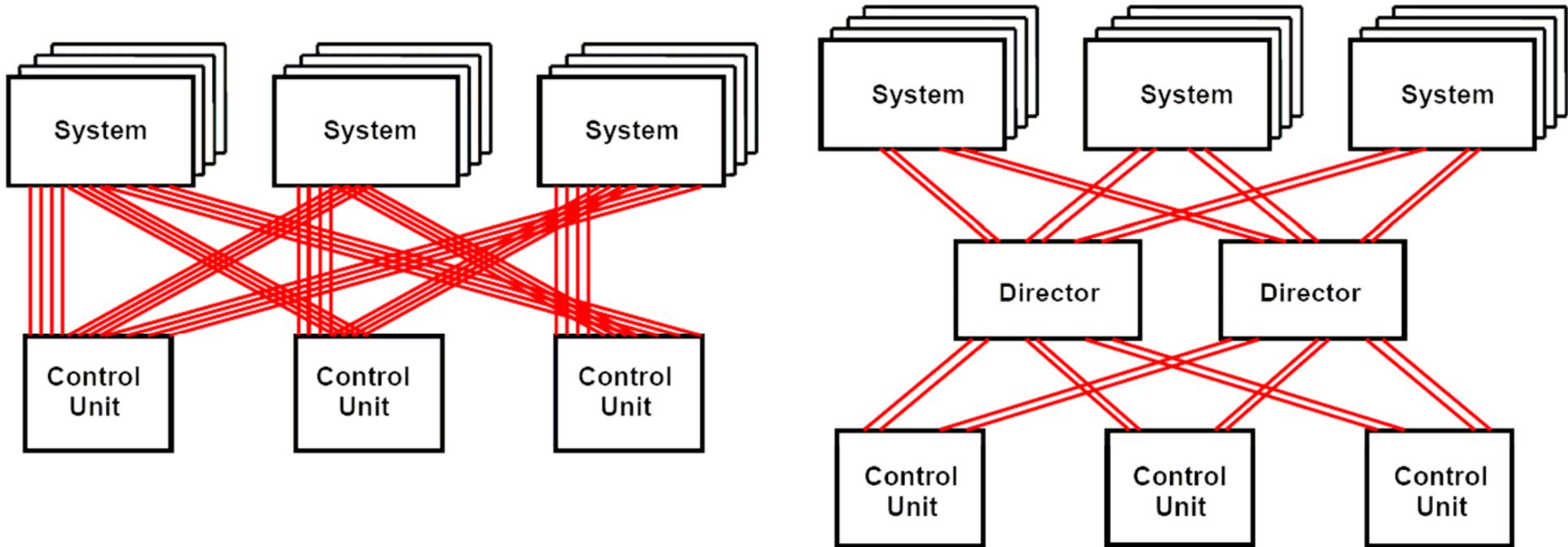


## System z Plattenspeicher Anschluss

Ein System kann über mehrere (bis zu 8) Kanäle mit einer Control Unit verbunden werden, und ein I/O Gerät kann an mehr als eine Control Unit angeschlossen werden.

Das Channel Subsystem bildet die logische I/O Konfiguration wie sie das Betriebssystem sieht, auf die physische Konfiguration ab.

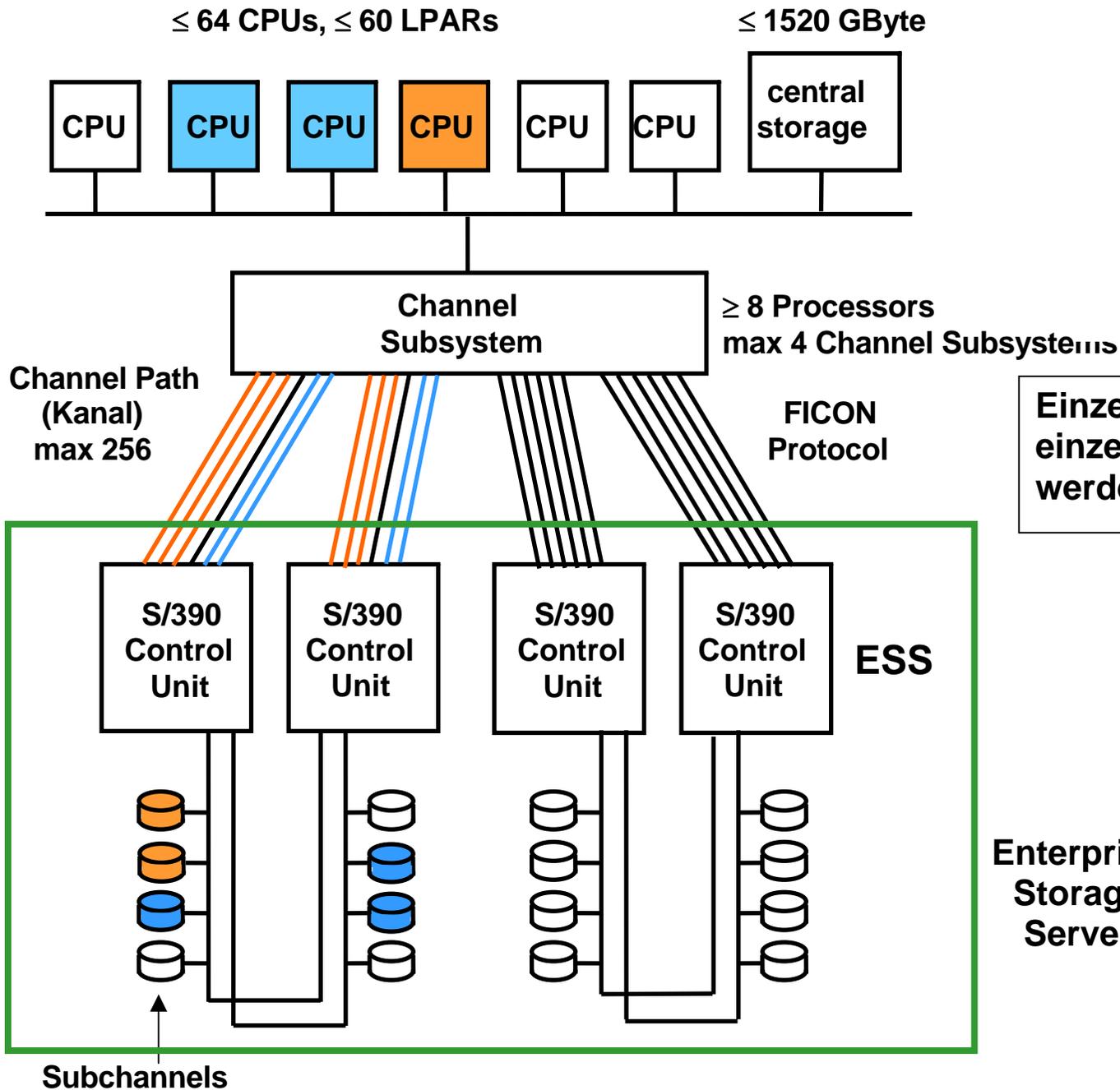
Heute werden mehrere Control Units und angeschlossene Plattenspeicher zu einem physischen „Enterprise Storage System“ zusammengefasst.



**Control Units sind in der Regel über mehr als einen Channel (Glasfaserkabel) mit dem Rechner (System) verbunden. Es können viele Control Units an ein System angeschlossen werden.**

**Vor allem in Installationen mit mehr als einem System kann das zu sehr unübersichtlichen Kabelverbindungen führen. Tausende oder zehntausende von Glasfaserkabeln sind keine Seltenheit.**

**Um die Anzahl der Kabel zu verringern, schaltet man häufig zwischen Channel Subsystem und Control Unit noch einen (oder viele) Switch(e), die dynamisch Verbindungen zwischen den einzelnen Systemen und Control Units schalten. Diese Switche werden als „Director“ bezeichnet.**



## z10 Disk Storage Attachment

Einzelne Channel Path können einzelnen CPUs fest zugeordnet werden.

Enterprise Storage Server

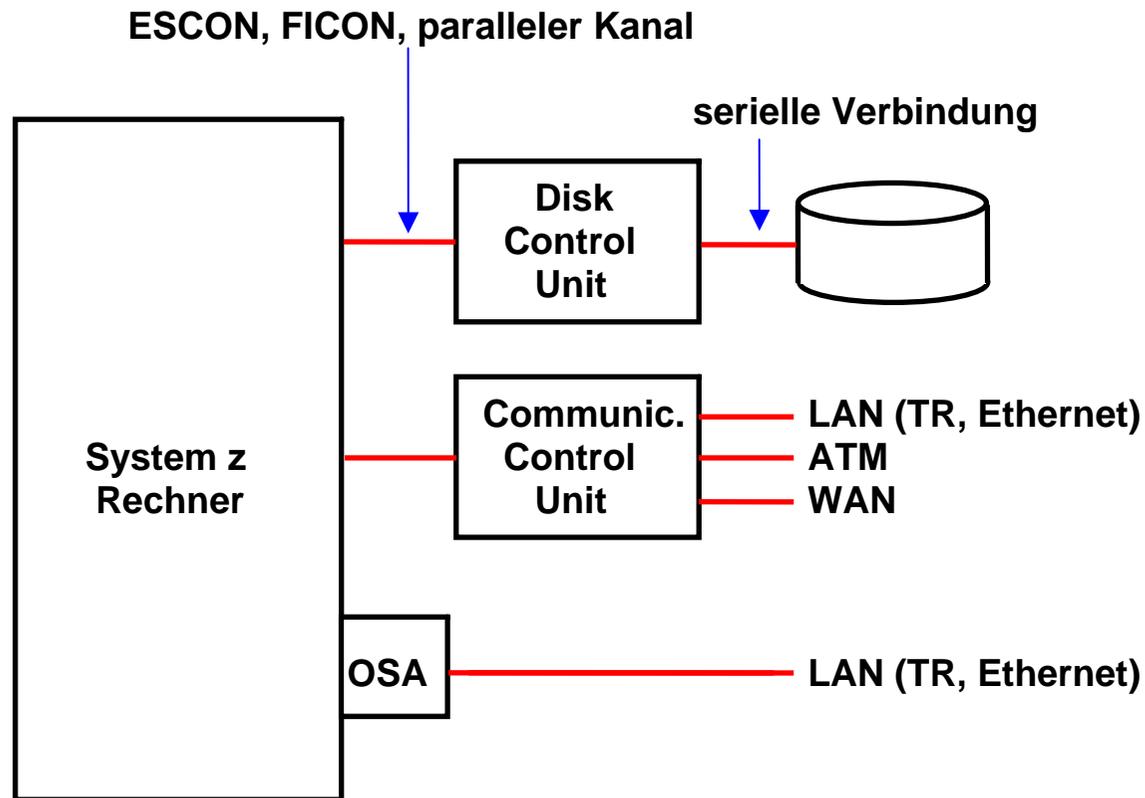
## **Logical Channel Subsystem (LCSS).**

**Ein Channel Subsystem kann bis zu 65,000 I/O Devices (vor allem Plattenspeicher) anschließen. Es gibt heute Situationen, wo dies nicht ausreicht.**

**Dies führte zur Einführung der Logical Channel Subsystems (LCSS) Technologie.**

**Ein System z kann bis zu vier Logical Channel Subsysteme aufweisen. Jedes LCSS kann 256 Channels und 65,000 I/O Devices unterstützen. Damit kann ein einzelner System z Rechner über 1,000 Channels und über 250 000 I/O Devices anschließen.**

**Die LCSS Architektur ist nach oben geöffnet. Es ist denkbar, dass zukünftige System z Implementierungen bis zu 256 logical Channel Subsystem aufweisen können, jedes mit 65,000 I/O Devices, oder bis zu 16 Millionen I/O Devices (Plattenspeicher) insgesamt.**



Einige Steuereinheiten können in den System z Rechner integriert werden. Das wichtigste Beispiel ist der OSA Adapter für den Anschluß von LAN's.

### System z I/O-Konfiguration

I/O Geräte werden grundsätzlich über Steuereinheiten (Control Units) angeschlossen. Steuereinheiten sind meistens in getrennten Boxen untergebracht, und über Glasfaser (ESCON, FICON) an den S/390 oder System z Rechner angeschlossen.

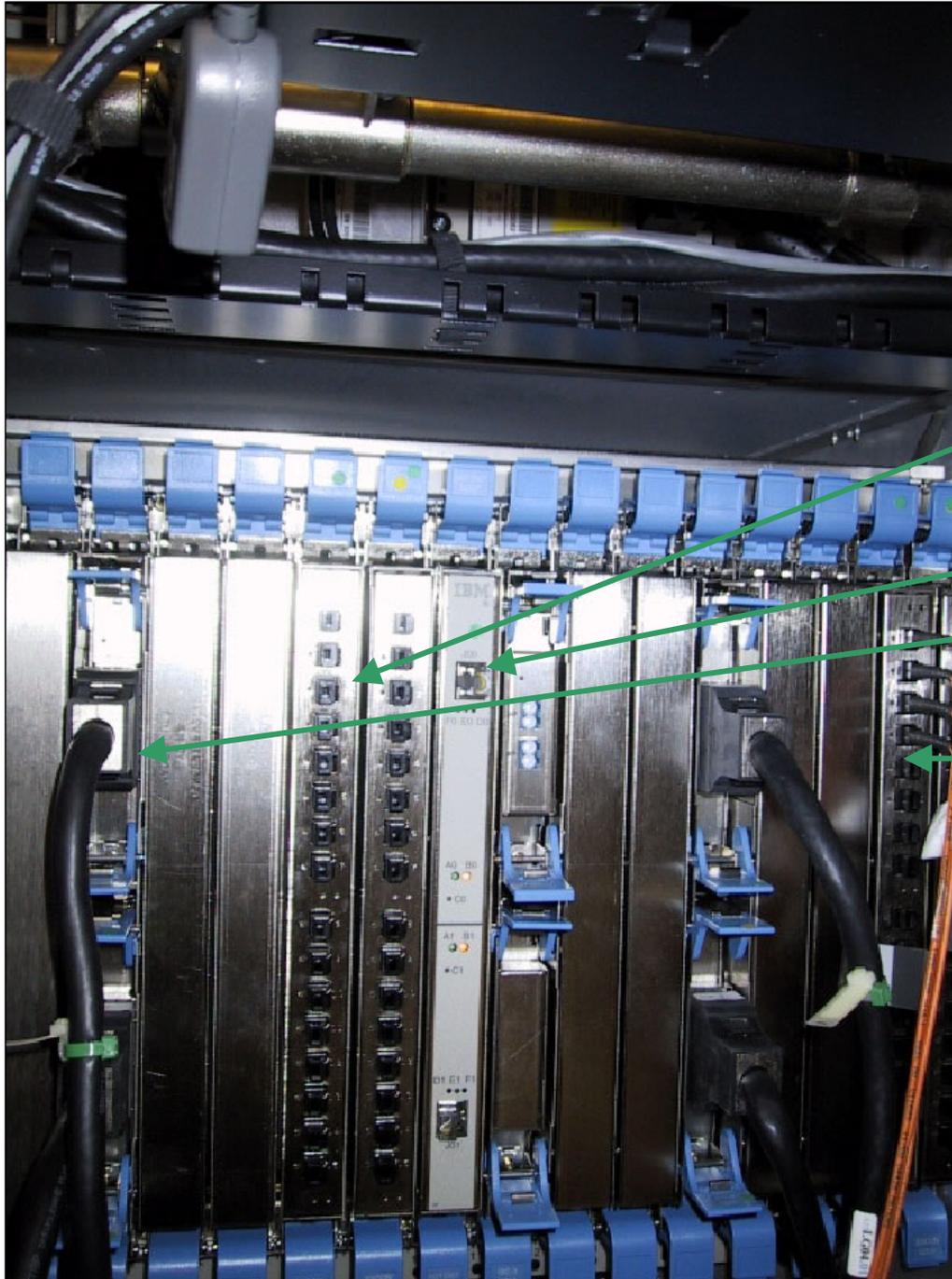
Es existieren viele unterschiedliche Typen von Steuereinheiten. Die wichtigsten schließen externe Speicher (Platten, Magnetbänder Archivspeicher) und Kommunikationsleitungen an.

Es existieren Steuereinheiten für viele weiteren Gerätetypen. Beispiele sind Belegleser für Schecks oder Druckstraßen für die Erstellung von Rentenbescheiden.

## **OSA Adapter**

**Der Open Systems Adapter (OSA) ermöglicht über mehrere Kanalanschlüsse einen direkten Weg in Token Ring, ATM oder Gigabit Ethernet Netzwerke.**

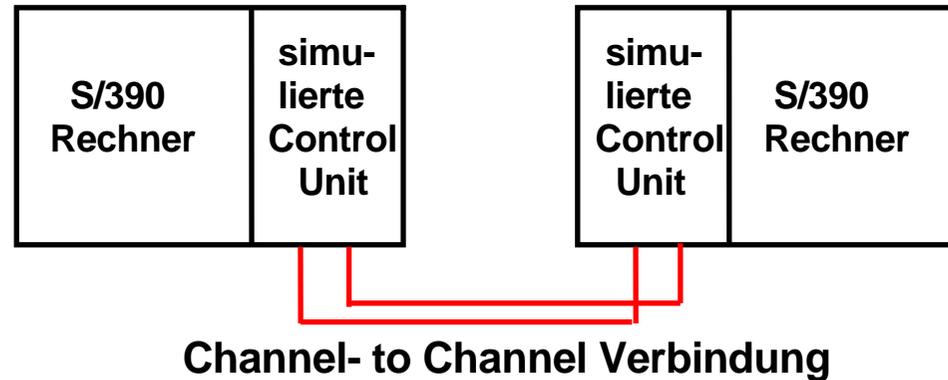
**Die heute verfügbare Variante auf den z/Series Maschinen ist der OSA Express Adapter. Er verfügt über eine höhere Bandbreite, außerdem können die Kanalverbindungen lastabhängig den jeweiligen LPARs dynamisch zugeordnet werden.**



The picture shows a System z I/O cage.

The two cards with numerous small black fiber channel connectors are ESCON cards. The next card to the right is an OSA card, with two RJ-45 connectors. The cards connected to the large black cables are inter-system coupling (ISC) cards for coupling links. To the far right of the image is another ESCON card with several fiber optic cables connected to it.

## Channel- to Channel Verbindung (CTC) Cross-System Coupling Facility (XCF)



Die Channel- to Channel Verbindung wird durch eine Hardware Einrichtung eines zSeries Systems verwirklicht, die dieses System gegenüber einem anderen zSeries System wie eine I/O Einheit erscheinen lässt.

Für eine Full Duplex Verbindung werden normalerweise zwei CTC Verbindungen eingesetzt.

Die Cross-System Coupling Facility (XCF) ist eine Komponente des z/OS Betriebssystems. Sie verwendet das CTC Protokoll und die CTC Hardware. Sie stellt die Coupling Services bereit, mit denen OS/390 Systeme innerhalb eines Sysplex miteinander kommunizieren.