

1. S/390 Hardware

1.1 Einführung

Für die Entwicklung der S/390-Familie ist es interessant, den historischen Hintergrund dieser Produkt-Linie zu betrachten. Am 7. April 1964 kündigte IBM die S/360-Rechnerfamilie an. Die Ziffer "360" bezieht sich dabei auf alle Punkte eines Kompasses, um die universelle Anwendbarkeit, den weitgefächerten Bereich von Performance und Preis des Produkts sowie die Entwicklungsrichtungen des Unternehmens IBM zu demonstrieren.

Die Existenz der S/360-Architektur ist den drei genialen Wissenschaftlern

Gene Amdahl,
Gerry Blaauw,
Fred Brooks

und der damaligen IBM-Entwicklungs-Abteilung unter der Leitung von B.O. Evans zu verdanken. Ursprünglich war geplant, an der S/360-Architektur sechs Jahre keinerlei Veränderungen vorzunehmen. Dagegen begann das IBM-Management schon sechs Monate nach der S/360-Einführung mit der Planung, die Monolithic Circuit (MLC)-Technologie in der Architektur einzusetzen. Gegen Ende 1965 war die Entwicklung einer neuen Familie mit dem Namen "New Systems" (NS) abgeschlossen. Diese NS-Architekturen implementierten die MLC-Technologie, und die Extended S/360-Architektur erhielt den Namen S/370. Letztere war aufwärtskompatibel zum System/360, d.h. der gesamte für die S/360-Architektur entwickelte Code war auch auf den S/370-Rechnern ohne Änderung lauffähig. Während der Entwicklung der S/370-Architektur ergab sich die Notwendigkeit, den für die Anwender-Programme verfügbaren Teil des Hauptspeichers zu erweitern. Die Konsequenz daraus führte zur virtuellen Speichertechnik, die im August 1972 mit den S/370-Modellen 158 und 168 veröffentlicht wurde. Die Modelle 158 und 168 implementierten ausserdem Multiprozessor-Konfigurationen der S/370-Rechner. Weitere Hauptspeicher-Unterstützung erhielt die S/370 Extended Architecture (370-XA), die 1981 von IBM angekündigt und 1983 vertrieben wurde.

Die nächste Verbesserung in der Rechnerarchitektur wurde 1988 mit der Einführung der Enterprise System Architecture/370 (ESA/370) vorgenommen. Diese Architektur verbesserte die virtuelle Speicher-Adressierung durch zusätzliche Register. Letztere erlauben den Zugriff zu einer anderen Form des virtuellen Speichers, genannt "Data Spaces". Die Data Spaces ermöglichen, mehr Daten im Haupt- und Expanded-Speicher abzulegen, wobei die Ein-/Ausgabe reduziert und der Durchsatz verbessert wird.

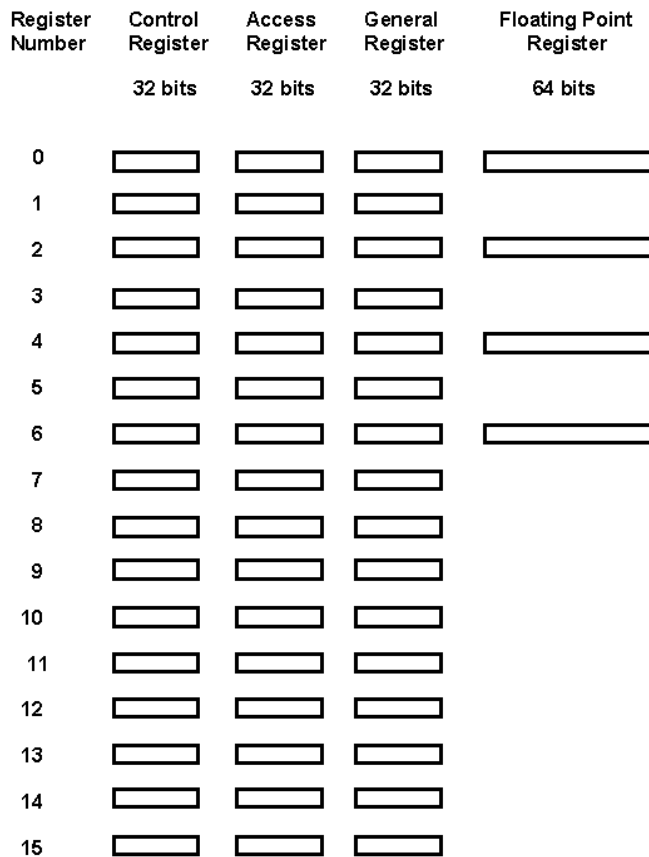
Im September 1990 brachte IBM die Enterprise System Architecture /390 und die ES/9000 S/390-Rechnerfamilie auf den Markt. Die ESA/390-Architektur implementiert ESCON und Sysplex. 1994 kündigte IBM eine Reihe von Erweiterungen der S/390-Familie einschliesslich der ES/9000-Produktlinie an. Dazu gehörten speziell skalierbare S/390-Parallelrechner in einem Parallel Sysplex Environment. Mit dieser Ankündigung setzte sich IBM zwei neue Rechnerarchitektur-Entwicklungslinien: Die Complementary Metal Oxide Semiconductor (CMOS) Technologie wird für die Bausteine sehr grosser Rechner eingeführt und ergänzt die bipolare Technologie; Rechnerarchitekturen mit speziellen Anwendungs-Umgebungen, die über die General Purpose-Umgebungen hinausgehen, sind in der Entwicklung.

1.2 S/390-Architektur

Die S/390-Architektur verfügt als 32 Bit-Rechner über unterschiedliche 32 Bit-Register, die sich in Control-, Access- und Mehrzweck-Register gliedern. Zusätzlich stehen noch 64 Bit-Gleitkomma-Register zur Verfügung. Das Registermodell ist in der Abbildung 3 dargestellt. Ende November 2000 wurde von der Firma IBM die S/390 als 64 Bit-Architektur angekündigt. Damit reagierte sie auf die avisierten 64 Bit-Architekturen von Intel, Compac (Dec), Sun und Motorola. Der Unterschied einiger 64 Bit-Architekturen zur /390-Architektur besteht darin, dass die sogenannte X-Serie von IBM mit der momentan eingesetzten 32 Bit-Architektur weitestgehend kompatibel ist. Die 32 Bit-Register werden auf 64 Bit erweitert, so dass auf einem Rechner sowohl 32 Bit- als auch 64 Bit- Programme bearbeitet werden können. Dagegen ist z.B. die 64 Bit-Architektur (Itanium) von Intel im Prinzip inkompatibel zum derzeitigen Pentium-Prozessor.

Bei der Entwicklung der /390-Architektur hat man bewusst auf Mechanismen verzichtet, die bis zur Einführung von RISC-Architekturen in den meisten Prozessoren implementiert waren. Dazu gehört u.a. der sogenannte Stack. IBM benutzt beim Aufruf eines Unterprogramms anstatt des Stack einige dedizierte Register, die als Save Area bezeichnet werden. Prinzipiell könnten als Aufruf-Adresse des Unterprogramms und für die Rückkehr-Adresse zum Hauptprogramm beliebige Register benutzt werden. Zu einem guten Programmierstil gehört es aber, die Linkage Conventions zu benutzen. Diese empfehlen dafür die Register 13, 14 und 15. Das Register 13

enthält den Pointer auf einen Bereich im Hauptspeicher (Save Area). In dem Register 14 wird der Inhalt des Befehlszählers (Return-Adresse) und im Register 15 die aufrufende Adresse abgespeichert. Dieser Mechanismus wurde schon in den 60iger Jahren in der /360-Architektur erfolgreich implementiert. Er stellt sich heute noch als sehr modern heraus und wird in ähnlicher Form in RISC-Architekturen verwendet.



S/ 390 Control-, Access, General, and Floating-Point Register

2 adjacent General or Floating Point Registers may be coupled as a double register

Abbildung 3: S/390-Registermodell

Eine Eigentümlichkeit der /390-Architektur stellt die Codierung der alpha-numerischen Zeichen sowie diverser Sonderzeichen dar. Im Unterschied zum 7 Bit-ASCII-Code wird der /390 spezifische EBCDIC-Code durch 8 Bit dargestellt. Abbildung 4 zeigt die entsprechenden Code-Tabellen.

0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111	↑
NUL	DLE	DS		SP	&	-						{	}	\	0	000
SOH	DC1	SOS				/		a	j	~		A	J		1	000
STX	DC2	FS	SYN					b	k	s		B	K	S	2	001
ETX	DC3	WUS	IR					c	l	t		C	L	T	3	001
SEL	RES	BYP	PP					d	m	u		D	M	U	4	010
HT	NL	LF	TRN					e	n	v		E	N	V	5	010
RNL	BS	ETB	NBS					f	o	w		F	O	W	6	011
DEL	POC	ESC	EOT					g	p	x		G	P	X	7	011
GE	CAN	SA	SBS					h	q	y		H	Q	Y	8	100
SPS	EM		IT					i	r	z		I	R	Z	9	100
RPT	UBS	SM	REF	‡	!	!	:									101
VT	CU1	CSP	CU3	.	\$,	#									101
FF	IFS	MFA	DC4	<	*	%	@					^		h		110
CR	IGS	ENQ	NAK	()	-	'									110
SO	IRS	ACK		+	i	>	=					~				111
SI	IUS	BEL	SUB		~	?	"								EO	111

bits 1,2,3			ASCII				bits 4,5,6,7		
000	001	010	011	100	101	110	111	↓	
NUL	DLE	SP	0	@	P	`	p	0000	
SOH	DC1	!	1	A	Q	a	q	0001	
STX	DC2	"	2	B	R	b	r	0010	
ETX	DC3	#	3	C	S	c	s	0011	
EOT	DC4	\$	4	D	T	d	t	0100	
ENQ	NAK	%	5	E	U	e	u	0101	
ACK	SYN	&	6	F	V	f	v	0110	
BEL	ETB	'	7	G	W	g	w	0111	
BS	CAN	(8	H	X	h	x	1000	
HT	EM)	9	I	Y	i	y	1001	
LF	SUB	*	:	J	Z	j	z	1010	
VT	ESC	+	;	K	[k	{	1011	
FF	FS	,	<	L	\	l		1100	
CR	GS	-	=	M]	m	}	1101	
SO	RS	.	>	N	^	n	~	1110	
SI	US	/	?	O	_	o	DEL	1111	

EBCDIC and ASCII Schemes

NOTE: The seven ASCII bits are numbered so as to enable direct comparison with EBCDIC. For example, the letter J is

1001010 in ASCII
11010001 in EBCDIC

The control character mnemonics are defined below. Note that SP is the space, or blank, character.

Abbildung 4: EBCDIC- und ASCII-Code-Tabellen

Bezüglich der Darstellung von Gleitkommazahlen weisen die S/390-Architekturen eine weitere Besonderheit auf: Sie unterstützen neben dem IEEE- noch einen spezifischen S/390-Gleitkommastandard. Da historisch der /390-Standard eher implementiert war, liegen gewaltige Datenmengen in diesem Standard weltweit auf Magnetbändern, deren Verarbeitung sich auf Architekturen ohne den /390-Gleitkommastandard schwierig gestaltet. Das betrifft in erster Linie Daten aus naturwissenschaftlich-technischen Forschungszentren. Neben dieser Besonderheit bietet die /390-Architektur die Möglichkeit, mit Dezimalzahlen zu operieren. Diese Dezimalarithmetik wird ausser von der /390- auch in identischer Form von der AS/400-, Vax- und BS/2000-Architektur benutzt. Die Methode, wie z.B. eine Ziffer + 00123 auf dem Speicherplatz 2500 abgelegt wird, ist in der Abbildung 5 dargestellt.

Gepackte dezimale Zeichenfolge auf Hauptspeicheradresse 02500

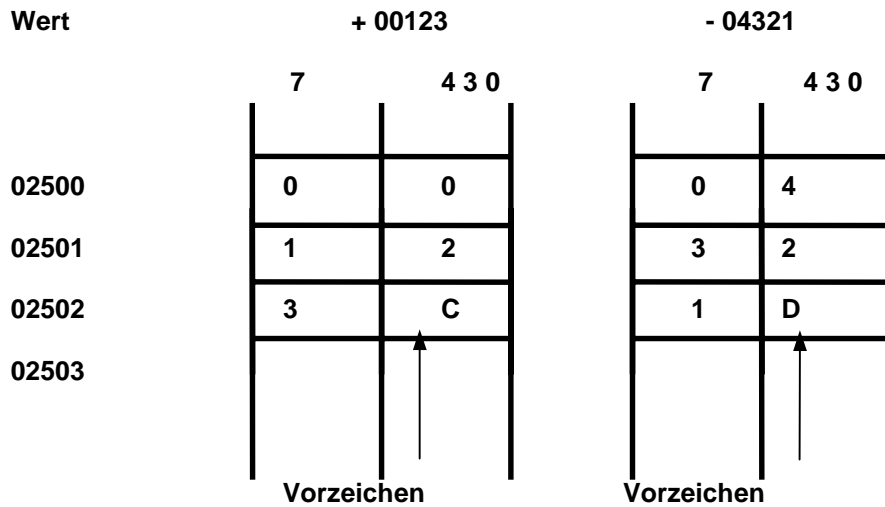


Abbildung 5: Dezimalarithmetik

Die S/390-Architektur benutzt wie die meisten Unix-Rechner beim Laden eines Registers das Big Endian-Format. Das bedeutet, dass ein Wort (4 Byte) beim Lesen aus dem Hauptspeicher in ein Register (4 Byte) in derselben Reihenfolge, in der die einzelnen Bytes im Hauptspeicher adressiert sind, auch in den Register-Adressen abgespeichert werden. Ein Beispiel zeigt die Abbildung 6, in der das Wort 12345678, das in vier aufeinanderfolgenden Hauptspeicher-Adressen 4710, 4711, 4712, 4713 abgespeichert ist, in ein Register gelesen wird. Durch den Ladebefehl wird zuerst das Byte, das auf der Hauptspeicher-Adresse 4710 steht, in die werthöchste Byte-Position des Registers, anschliessend das Byte, das auf der Adresse 4711 abgelegt ist, in die zweithöchste Byte-Position des Registers gelesen usw. Dieses Endian-Problem stellt sich nicht nur bei der Datenübertragung zwischen einem PC (Pentium: Little Endian) und einem /390-Rechner sondern auch bei der Kommunikation zwischen einem PC und einer HP-, Sun- oder IBM-AIX-Architektur.

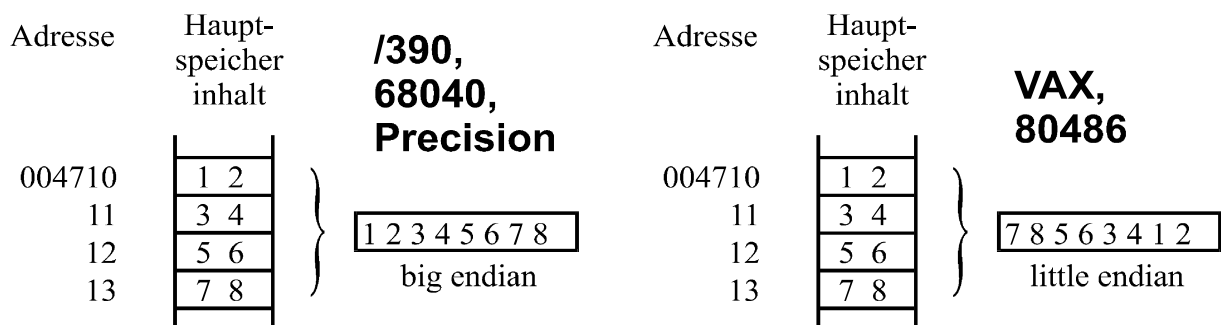


Abbildung 6: Byte Ordering

Die S/390-Architektur verfügt wie andere Rechner auch über einen Speicherschutz. Dieser ist im Normalfall so implementiert, dass der Hauptspeicher in Blöcke aufgeteilt und jedem dieser Blöcke ein Speicherschutzschlüssel zugeordnet wird. Neben dieser allgemein üblichen Methode zeichnet sich aber der Speicherschutz der IBM /390-Architektur noch durch einen zusätzlichen Hardware-Mechanismus aus, der in anderen Architekturen nicht bzw. erst in Ansätzen (Intel) implementiert ist. Dieser Hardware-Speicherschutz verwendet im Program Status Word (PSW) einen 4 Bit-Speicherschutzschlüssel (Abbildung 7). Bei jedem Speicherzugriff wird aus einem Schnellspeicher dieser Speicherschutzschlüssel ausgelesen und mit dem 4 Bit-Feld im PSW verglichen. Nur wenn dieser Vergleich positiv ausfällt (4 Bit-Felder sind identisch), erfolgt der Zugriff. Der Hardware-Speicherschutz benutzt das Konzept des sogenannten License Internal Code (LIC). Dahinter verbarg sich in der Vergangenheit spezieller Micro-Code, der von einer separaten, primitiveren Architektur ausgeführt wurde. Inzwischen sind diese zusätzlichen Funktionen, die insgesamt die Zuverlässigkeit und Sicherheit der S/390-Rechner erhöhen, in der Maschinensprache realisiert, da in diesem Fall die gesamte Entwicklungs-Umgebung

zur Verfügung steht. Da der LIC vor jedem normalen Benutzer-Zugriff geschützt sein muss, wird der Code-Bereich ausserhalb des Adressbereichs der einzelnen Maschinenbefehle im Hauptspeicher untergebracht. Ähnlich verhält es sich mit dem Code, der das Initial Program Load (IPL) ausführt. Der Unterschied zum LIC besteht u.a. darin, dass der IPL-Code nicht im Hauptspeicher sondern auf einem Plattenbereich, der für den Anwender nicht zugänglich ist, untergebracht wird.

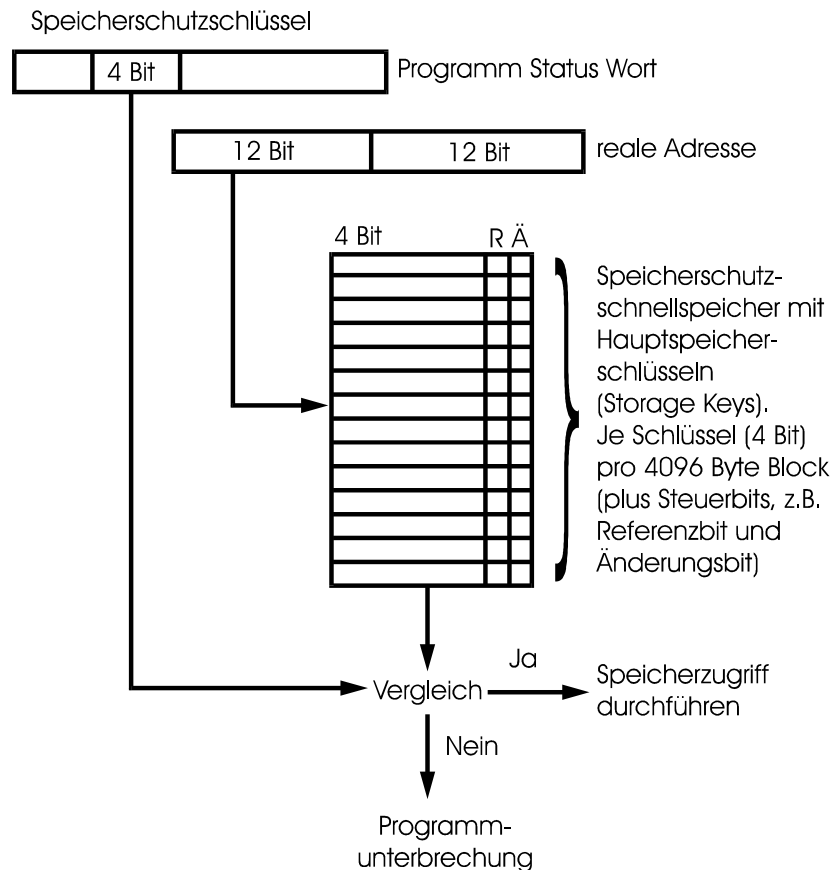


Abbildung 7: Hardware-Speicherschutz

Eine sehr wesentliche Eigenschaft der S/390-Rechner bezieht sich auf die Ein-/Ausgabe-Leistung. Letztere hat einen besonderen Stellenwert bei der Transaktionsverarbeitung. In diesem Fall wird das Leistungsverhalten eines Rechners nicht vorrangig durch die Leistungsfähigkeit der CPU bestimmt sondern dadurch, wie schnell der Zugriff auf die gespeicherten Daten erfolgen kann. Es ist allerdings sehr schwierig, das E/A Leistungsverhalten zu charakterisieren. Eine Meßgröße ist die gesamte maximale E/A Datenrate. Eine Angabe hierüber enthält das Februar 1996 Heft der Zeitschrift „Manufacturing Systems“. Hiernach kann das S/390 E/A Subsystem 1.000 bis 20.000 MByte/Minute übertragen. Sehr große UNIX Systeme können 2 bis 100 MByte/Minute übertragen. Beispiele solcher Transaktionen stehen im Zusammenhang mit Geldautomaten, Kontoauszugsdrucker und diversen Abrechnungsplätzen im Einzelhandel, bei denen es sich um den bargeldlosen Geschäftsverkehr handelt. E-Business und e-Commerc sind komplexere Anwendungen, die leistungsfähige Client/Server-Systeme voraussetzen. Den Kern solcher Systeme bildet der Server, der primär auf sehr hohe Ein-/Ausgabe-Leistungen optimiert ist. Historisch gewachsen sind in diesem Anwendungsbereich die /390-Architekturen. Die bereits laufenden Server und deren Neuinstallationen in Banken, Versicherungen und Versandhäusern werden zum überwiegenden Teil von /390-Architekturen realisiert.

G5/6 Systemstruktur: High Performance Peripherie-Anbindung

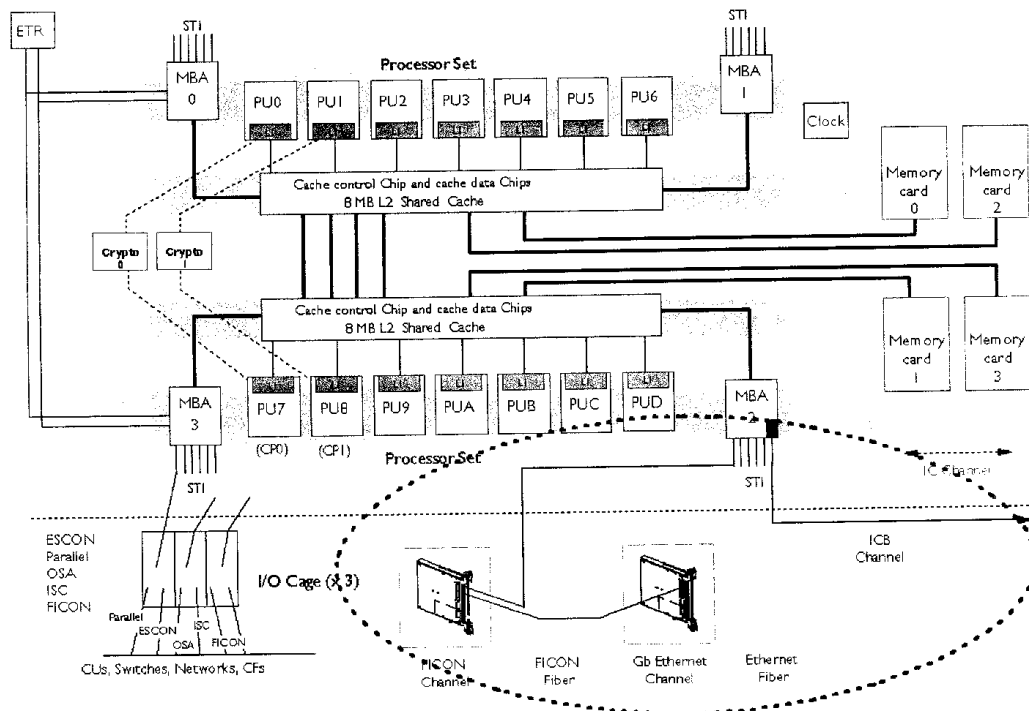


Abbildung 8:/390-Systemstruktur

In der Abbildung 8 ist die /390-Systemstruktur der Modelle G5/G6 dargestellt. Jeweils 2 Prozessor-Sets zu je 7 CPUs haben Zugriff auf einen Set-eigenen L2-Cache. Dazu kommen noch zwei Krypto-Prozessoren, die bei Internet-Zugriffen die Verschlüsselung in der Secure Socket Layer sehr zeiteffektiv ausführen. 4 Ein-/Ausgabe-Prozessoren übernehmen den Datenverkehr mit den Plattenspeichern. Diese Verbindung erfolgt mittels ESCON-Switche bzw. Glasfaser-Leitungen.

S/390-Rechner, deren Marktanteile in der ersten Hälfte der 90iger Jahren geringer ausfielen, seit 1995 aber wieder ansteigende Tendenz zeigen, werden von mehreren Herstellern angeboten. Die Firma IBM ist als /390-Anbieter der Marktführer mit einem Anteil von ca. 50-60%. Die Installationen umfassen die Serien G5 und G6, die Multiprise-Serie sowie die P/390-Implementierung. Letztere integriert als normaler PC zwei CPUs. Die /390-Funktionalität wird von einem CPU-Board übernommen, während die zweite CPU als Intel-Pentium-Architektur nur für die Ein-/Ausgabe-Funktion zuständig ist. Auf dem /390-Prozessor-Komplex läuft ein vollwertiges OS/390-Betriebssystem einschließlich der DB2-Datenbank sowie CICS-Transaktionsmonitor. Der Pentium-Prozessor implementiert das OS/2-Betriebssystem. Aus der Abbildung 9 ist der funktionale Ablauf zwischen den beiden Betriebssystemen des P/390-Servers dargestellt. Es sind zwei Prozessoren erkennbar: Die linke Seite bildet den S/390- und die rechte Seite den Server-Prozessor (Pentium). Beide Teile haben ihren eigenen Speicher (S/390: 128 MByte, Pentium 32 MByte).

Den Rest des /390-Marktes teilen sich die Firmen Amdahl (Millenium-Server), Hitachi/Compaq (Skyline-Serie), Fujitsu und Siemens.

Functional Flow

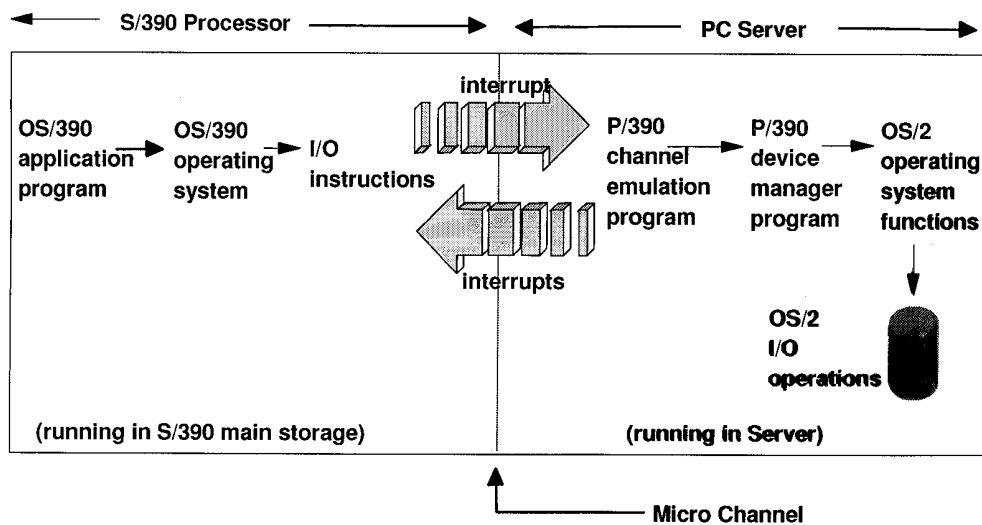
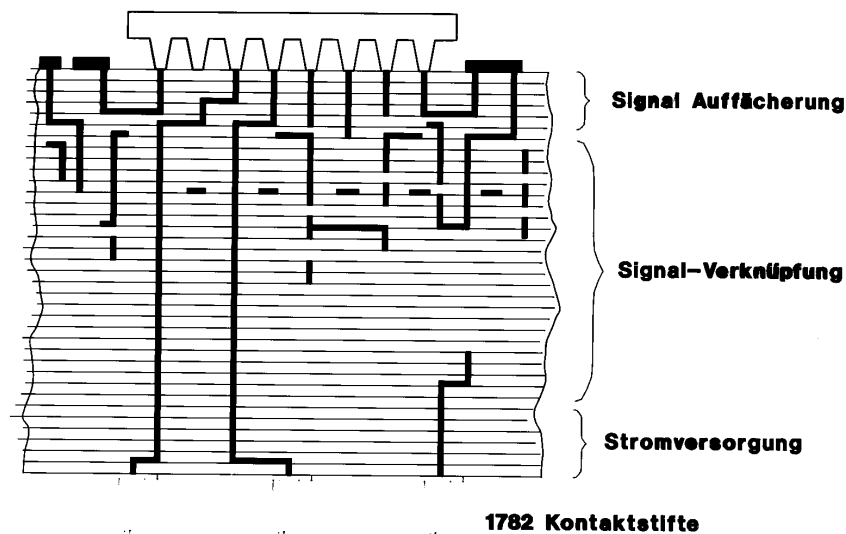


Abbildung 9: Funktionsweise der P/390-Architektur

1.3 S/390-Technologie

Die /390-Architektur einschliesslich ihrer Vorgängerinnen ist seit ihrer Einführung vorzugsweise in der bipolaren Technologie implementiert worden. Der Grund dafür ergab sich aus der Tatsache, dass die bipolare der CMOS-Technologie bezüglich Geschwindigkeit, d.h. Schaltzeiten der Transistoren in der kombinatorischen und seriellen Logik, überlegen war. Sie hatte aber den Nachteil, dass Rechner dieser Technologie relativ viel Energie verbrauchten. Um die Wärmeenergie abzuführen, war Wasserkühlung notwendig. Das hatte wiederum zur Folge, dass diese Rechner geometrisch gewaltige Dimensionen annahmen. Ab der zweiten Hälfte der 80iger Jahre änderte sich das Geschwindigkeitsverhältnis der beiden Technologien. In dieser Zeit wurde die erste S/390-Architektur in CMOS-Technologie auf den Markt gebracht. Der CMOS-Mikroprozessor war ein Produkt des IBM-Entwicklungs-Labors in Böblingen. Mitte der 90iger Jahre war die CMOS-Technologie soweit fortgeschritten, dass die Rechnerarchitekturen auf CMOS-Basis im gleichen Geschwindigkeitsbereich lagen wie die mit bipolarer Technologie. Inzwischen liegt das CPU-Performance-Verhältnis eines /390-Rechners und einer Workstation von HP, Sun, SG oder eines leistungsfähigen PCs nahe der Zahl eins.

Der wesentliche Unterschied zwischen der /390-Architektur und anderen Rechnerarchitekturen besteht in der Benutzung der Packaging-Technologie. Die CPU-Chips der /390-Rechner sind auf einem Multilagen-Keramik-Modul untergebracht. Der Keramik-Träger nimmt ausser den 12 bzw. 16 CPU-Chips weitere für L2-Caches, Vektor- und Crypto-Verarbeitung sowie Ein-/Ausgabe auf. In der Abbildung 10 ist ein Querschnitt durch einen solchen Keramik-Modul dargestellt. Ausser IBM hatte Intel die Multilagen-Keramik-Technologie im Pentium Pro implementiert. Aus Kostengründen wurde diese Technologie in den folgenden Pentium-Architekturen nicht mehr verwendet. Diese Entscheidung hatte besonders auf die Performance von Servern negative Auswirkungen, weil den L2-Cache bei der Multilagen.-Keramik-Technologie wesentlich geringere Zugriffszeiten infolge der Entfernung zum L1 auszeichnet.



**ESA/9000 TCM Mehrschichtkeramik
63 Schichten; über 5 Meter Leiterbahnen pro cm²**

Abbildung 10: Querschnitt durch TCM-Multilagen-Keramik

Die Thermal Conduction Module (TCM)-Technologie ist heute die am weitesten verbreitete, da sie als einzige in der Lage ist, auf relativ einfache Weise sehr große Wärmemengen zu transportieren. Der Nachteil liegt in den hohen Herstellungskosten dieser Module. Ihr Vorteil besteht dagegen in der hohen Zuverlässigkeit. Nach Aussagen des Herstellers sind in der gesamten Zeit des Betriebes der TC-Module keinerlei Ausfälle registriert worden.

Eine weitere Schwierigkeit wird von der TCM-Technologie beseitigt. Die einzelnen Verbindungen der Kontakte müssen von den Chips zu den Pins nach außen geführt werden. Bei einem einzelnen Chip ist das noch mit einigem geometrischen Aufwand möglich. Befinden sich aber mehrere Chips auf einem Keramik-Modul, dann müssen die Verbindungen zwischen den verschiedenen Chips extrem kurz ausfallen. Als Folge davon werden keine gedruckten Schaltungen mehr benutzt, vielmehr sind für derartig komplizierte Verdrahtungsmuster spezielle mehrlagige Keramik-Module entwickelt worden (Abbildung 10). Als Substrat für die Mehrlagen-Module verwendet man Aluminiumoxid (Al_2O_3). Innerhalb der verschiedenen Schichten, deren Anzahl bei der IBM-TCM-Technologie bis zu 88 betragen kann, entstehen sehr komplexe Verdrahtungsmuster. Die senkrechten Verbindungen zwischen den verschiedenen Schichten bestehen aus leitenden Bohrungen, die wiederum innerhalb einer Schicht in horizontalen Leiterbahnen weitergeführt werden und an einer Bohrung zu einer darunter- oder darüberliegenden Schicht enden usw. Die Abbildung 11 zeigt ein ESA/9000 Mehrschichtkeramik TC-Modul, der neben dem keramischen Chip-Träger die Mechanik für die Energieableitung mittels Stempel und Kühlplatte enthält.

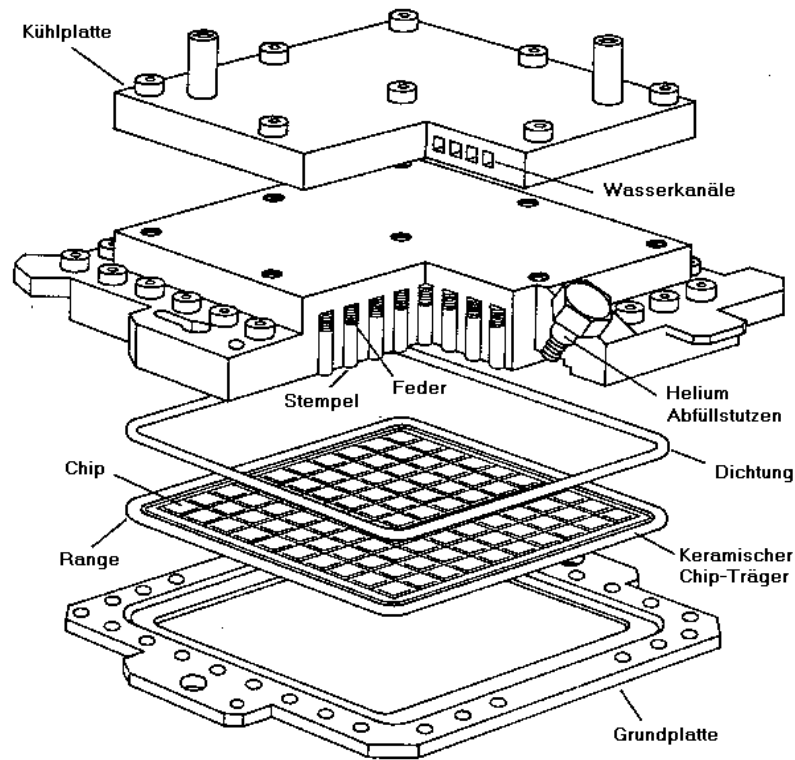


Abbildung 11: ESA/9000 Mehrschichtkeramik TC-Modul

Der momentan modernste G6-Prozessor der /390-Architektur integriert insgesamt 12 CPUs (Abbildung 12), die über einen Main Store Controller auf einen gemeinsamen Hauptspeicher (32 Gbyte) zugreifen. Je 6 CPUs sind 8 MByte Cache zugeordnet. In der von IBM neu angekündigten X-Serie der /390-Architektur wird die Anzahl der CPUs auf 16 erhöht. Weiterhin sind in der X-Serie die beiden Caches des G6-Modells wieder in einem einzigen Cache integriert.

IBM S/390 9672 G6 Server Processor MCM

- **S/390 Microprocessor**
 - ▶ Processor Chip (PU)
 - 25 million transistors include
 - 256 KB On-Chip Cache
 - Floating Point Processor
- **L2 Cache**
 - ▶ L2 Cache Chip on MCM (SD)
 - 121 million transistors (2 x G5)
 - ▶ 2MB per chip (1 MB for G5)
 - ▶ 16 MB (2 x 8 MB) Shared
- **Memory Bus Adapter (MBA)**
 - ▶ 24 STI I/O Attach
- Up to 12 CPs
- Up to 7 SAPs
- Up to 11 ICFs
- 2 Crypto Element Chips (CE)

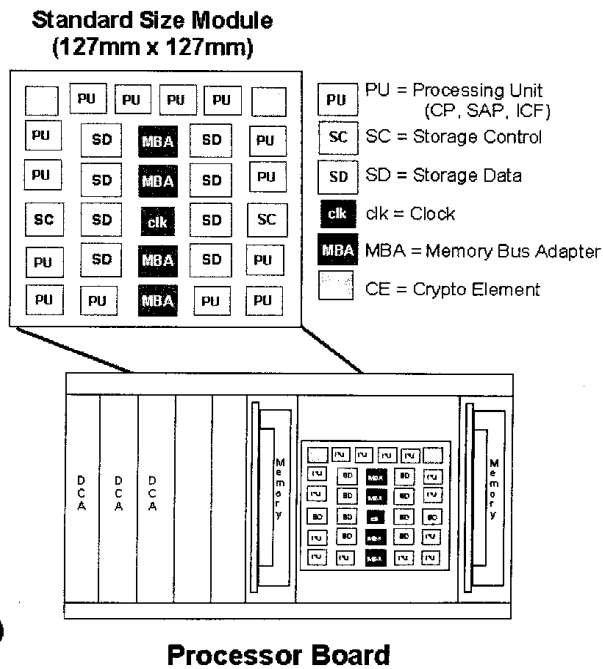


Abbildung 12: G6-Processor Board

Die Firma IBM verwendet als eine der ersten die Silicon-On-Insulator (SOI)-Technologie in der Chip-Produktion.

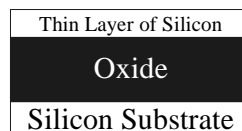


Abbildung 13: SOI-Schicht

Die SOI-Technologie bezieht sich auf eine dünne Silizium-Schicht, die auf dem Silizium-Oxid platziert wird (Abbildung 13). Die Idee besteht darin, dass die SOI-Schicht die Kapazität eines Schalters, der als Transistor implementiert ist, reduziert. Die kleinere Kapazität äußert sich in einer verkürzten Schaltzeit des Transistors. In einem MOS-Schalter können Ladungen in einem Bereich zwischen dem dotierten Silizium und dem Silizium-Substrat gespeichert werden (Abbildung 14). Diese Grenz-Kapazität (junction capacitance) wird durch die Silizium (Glas)-Isolationsschicht eliminiert.

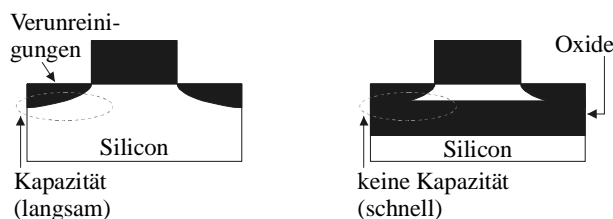


Abbildung 14: SOI mit Verunreinigungen