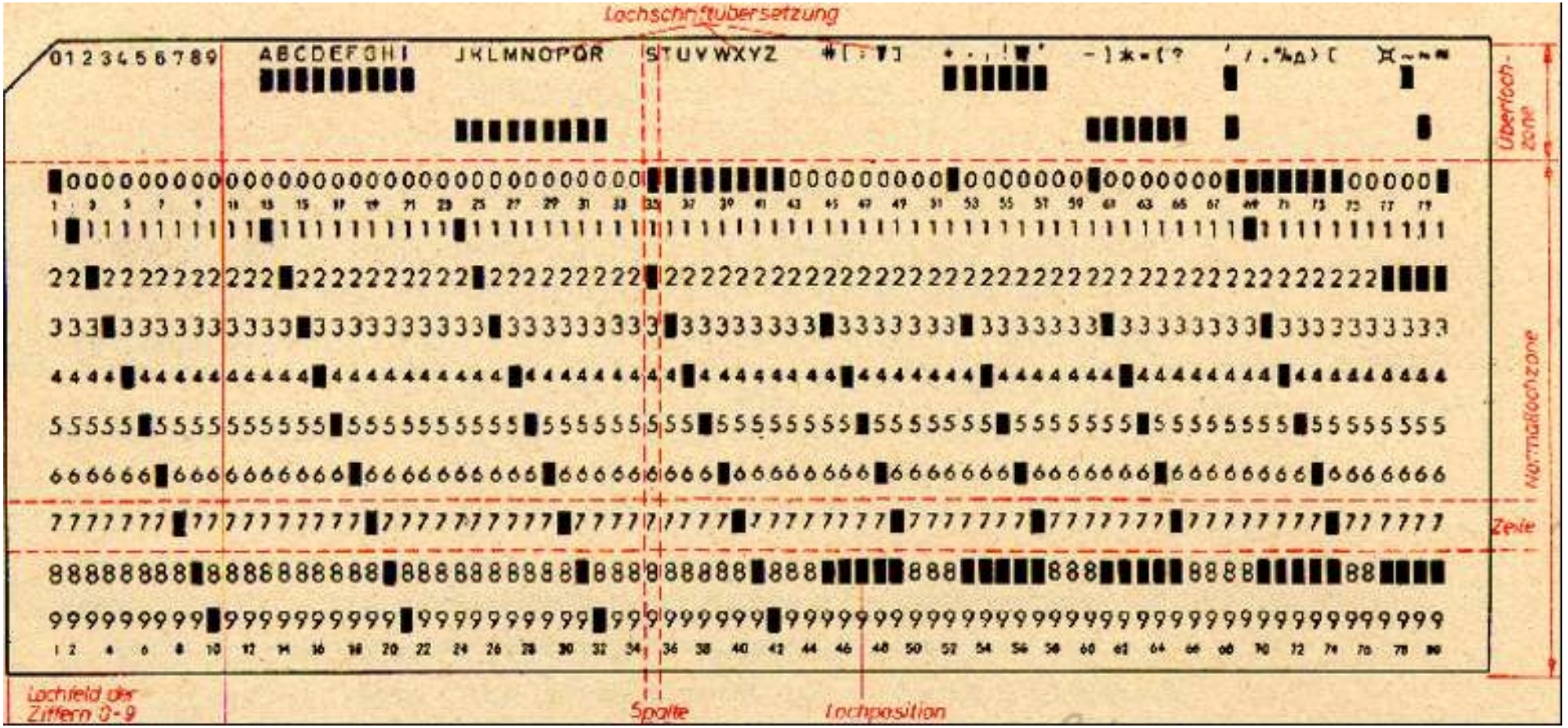
Einführung in z/OS und OS/390

Dr. rer. nat. Paul Herrmannn
Prof. Dr.-Ing. Wilhelm G. Spruth

WS 2006/2007

Teil 2

zSeries Hardware



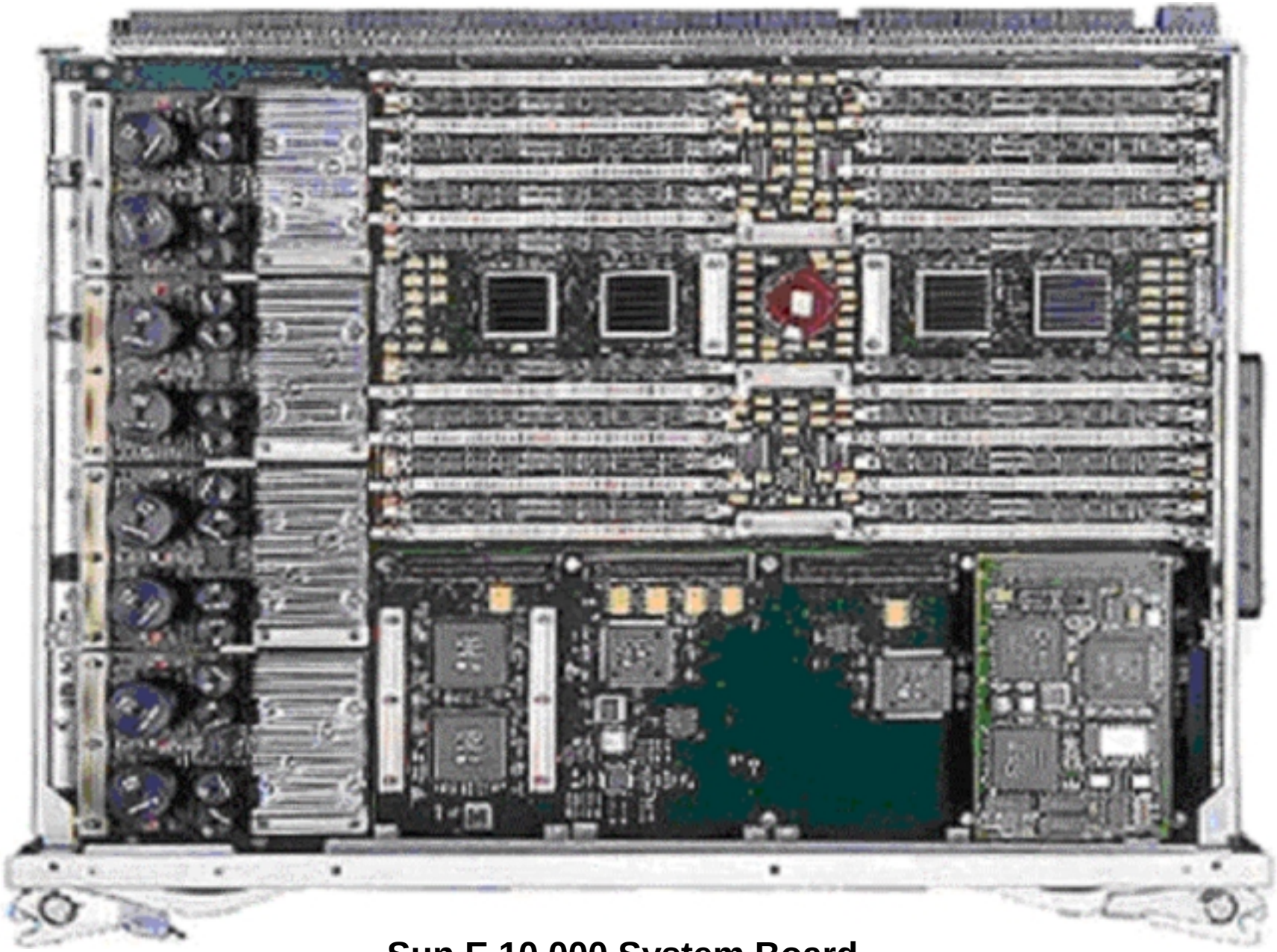
IBM Lochkarte


IBM 026 Card Punch

Ursprünglicher S/360 Rechner (Modell 62)

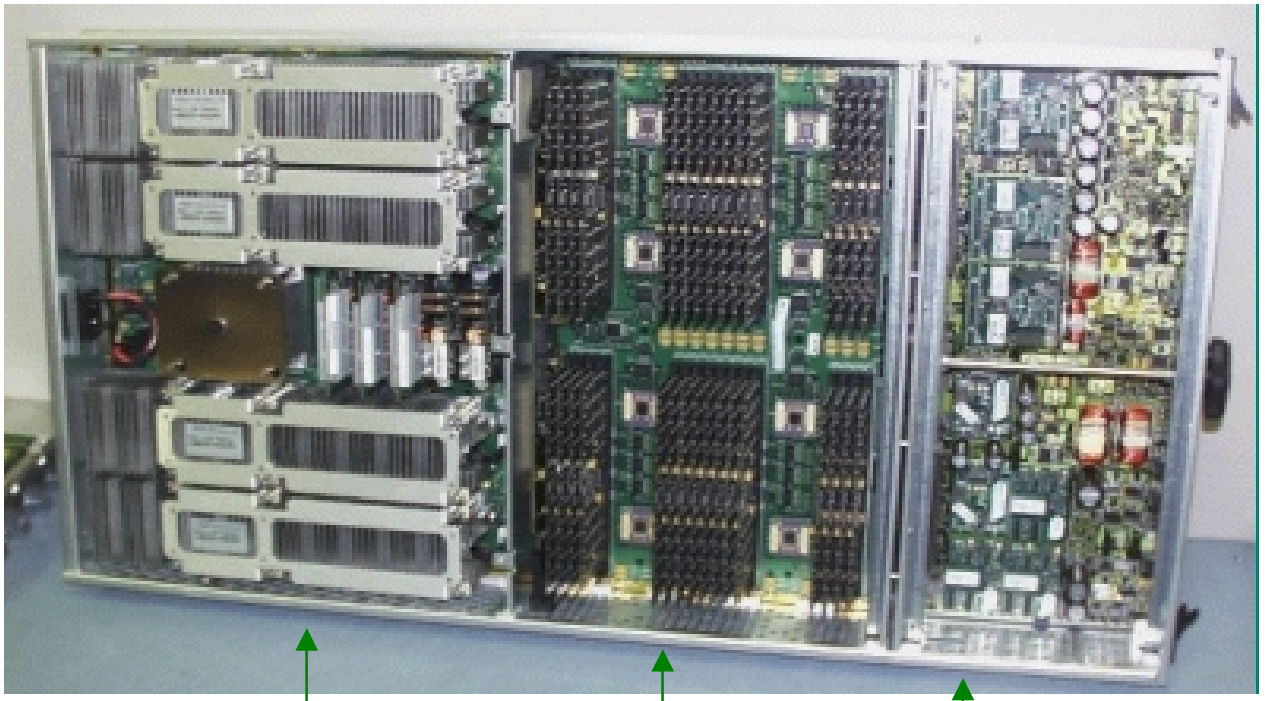
mit Bandeinheit und Plattenspeicher, 1966

IBM 3278 „Green Screen“ Bildschirm



Sun E 10 000 System Board

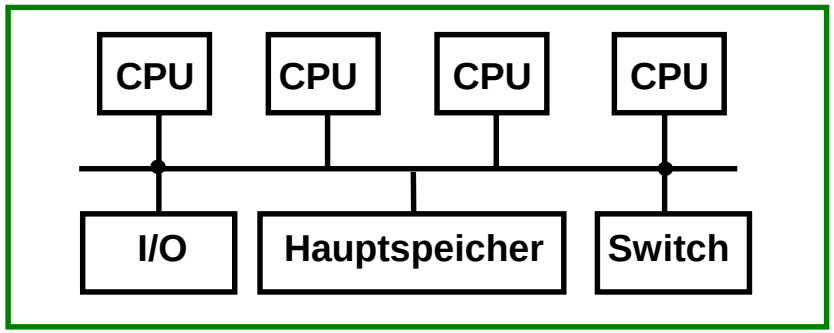
HP Superdome Cell Board

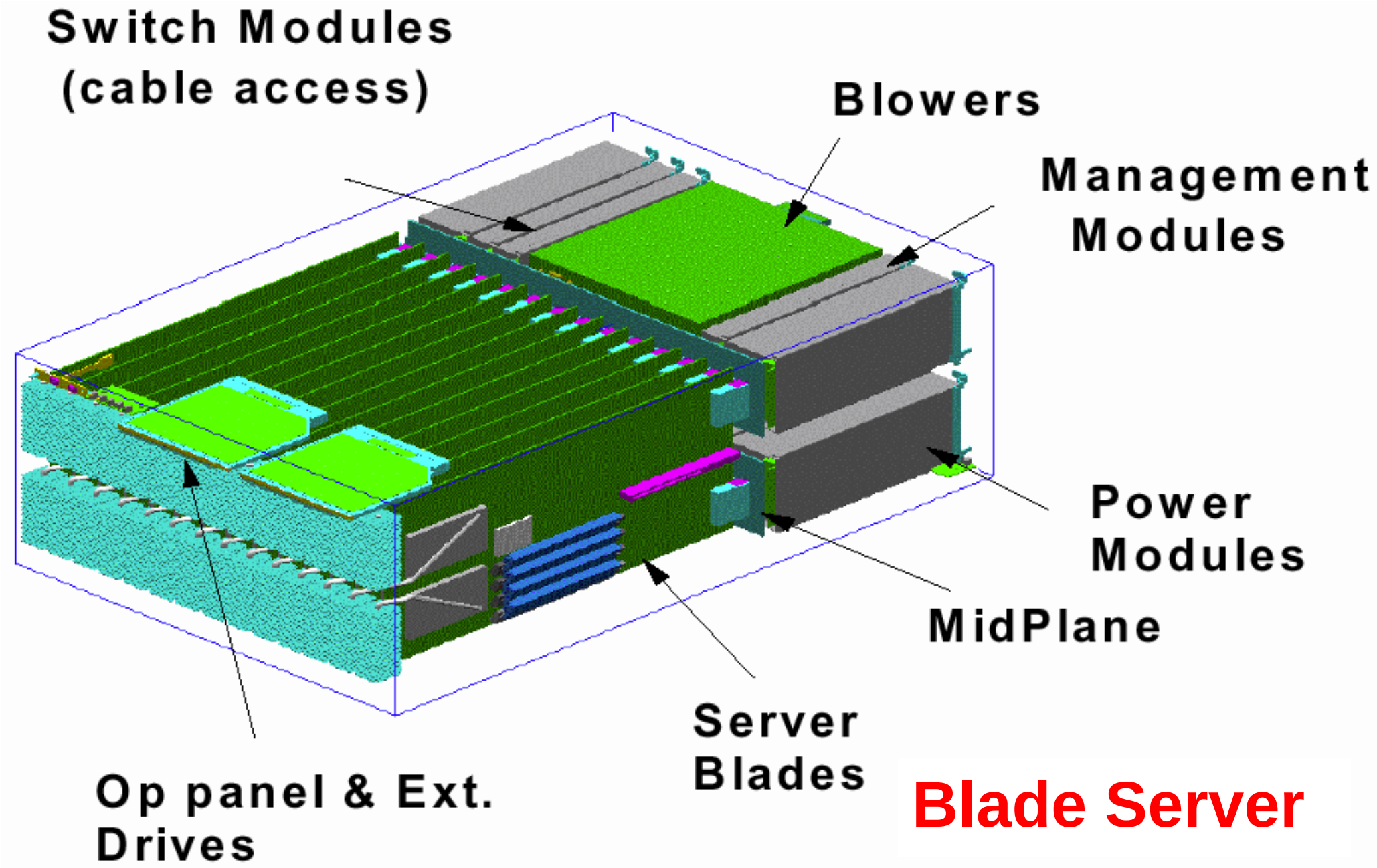


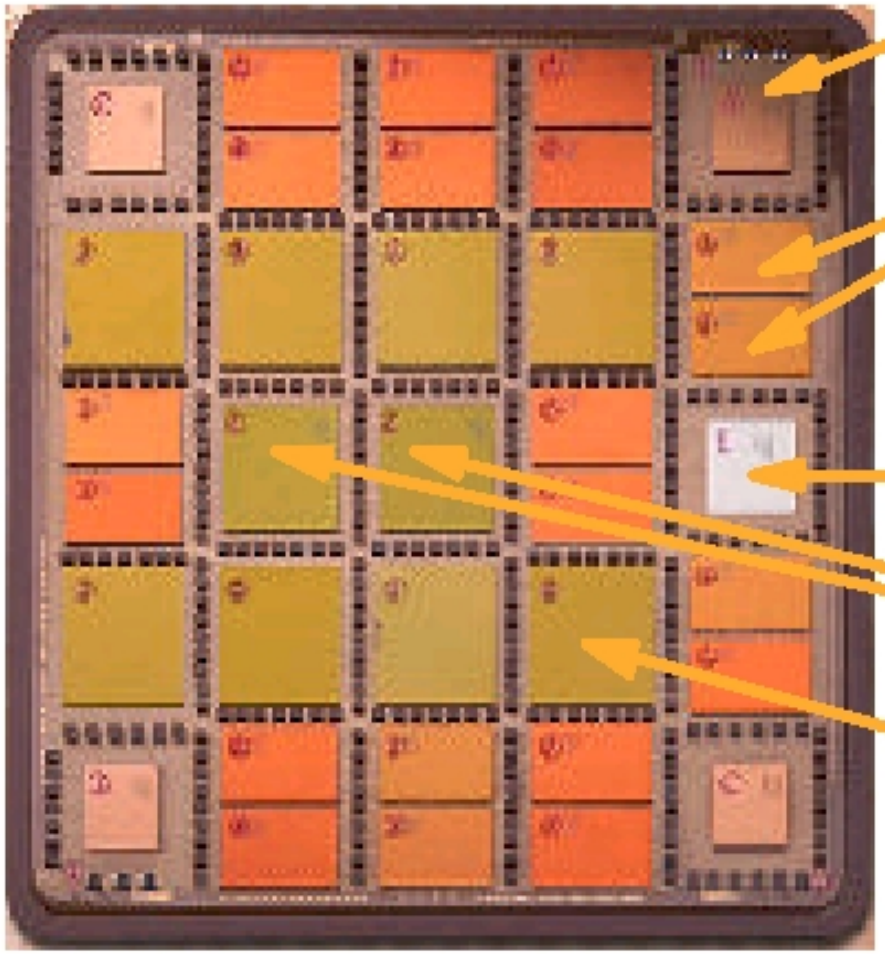
4 Itanium 2 CPUs
1,5 GHz

32 Gbyte
Hauptspeicher

E/A Bus
Anschlüsse
12 PCI




Switch Modules
(cable access)
Blowers
Management
Modules
Power
Modules
MidPlane
Server
Blades
Op panel & Ext.
Drives
Blade Server



(4) Memory Bus Adapters

(20) Dual Processors:
CPU, System Assist,
Internal Coupling,
Integrated Linux, Spares

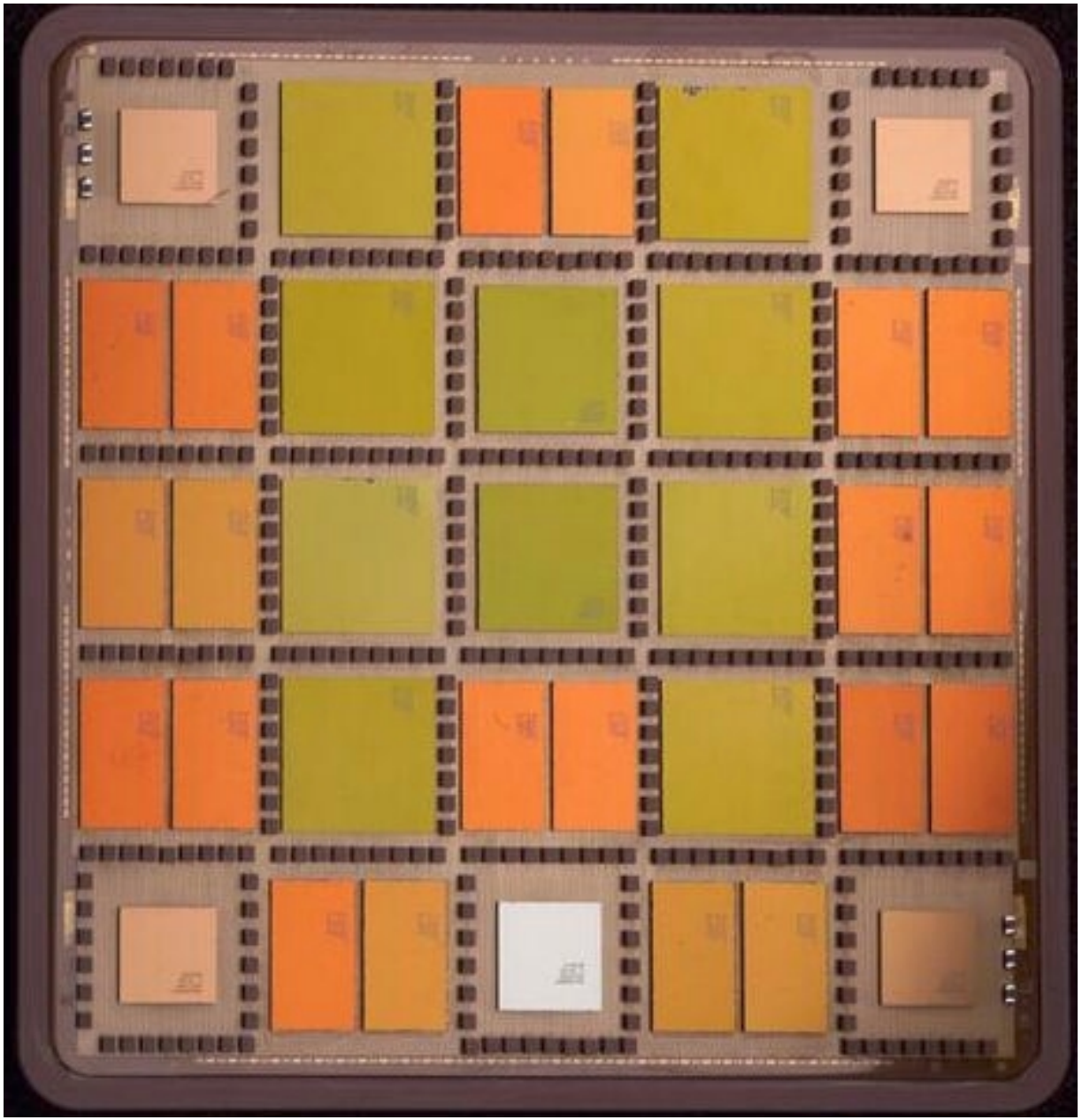
(1) Clock Chip

(2) L2 Storage Control

(8) CMOS L2 Chips

z900 Multi Chip Modul

16 CPU Chip SMP, Shared L2 Cache mit Error Correction (ECC),



Z900 Multichip Module (MCM)

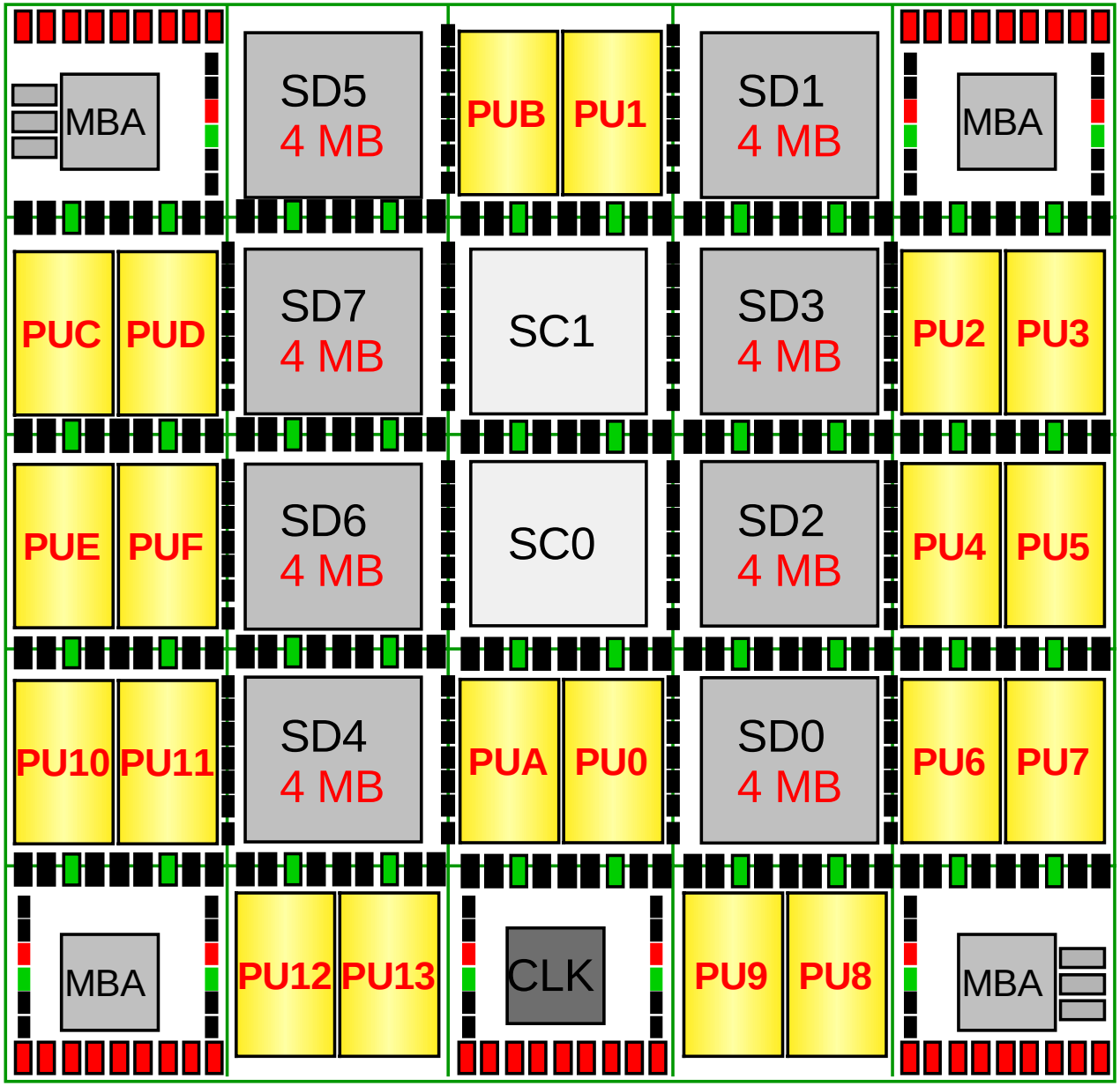
127,5 x 127,5 mm Modul

35 Chips, 2,5 Milliarden Transistoren

234 Million Transistoren L2 chip

101 Glas-Keramik + 6 Dünnfilm Verdrahtungslagen

1 km Drahtlänge insgesamt



Z900 Multichip Module (MCM)

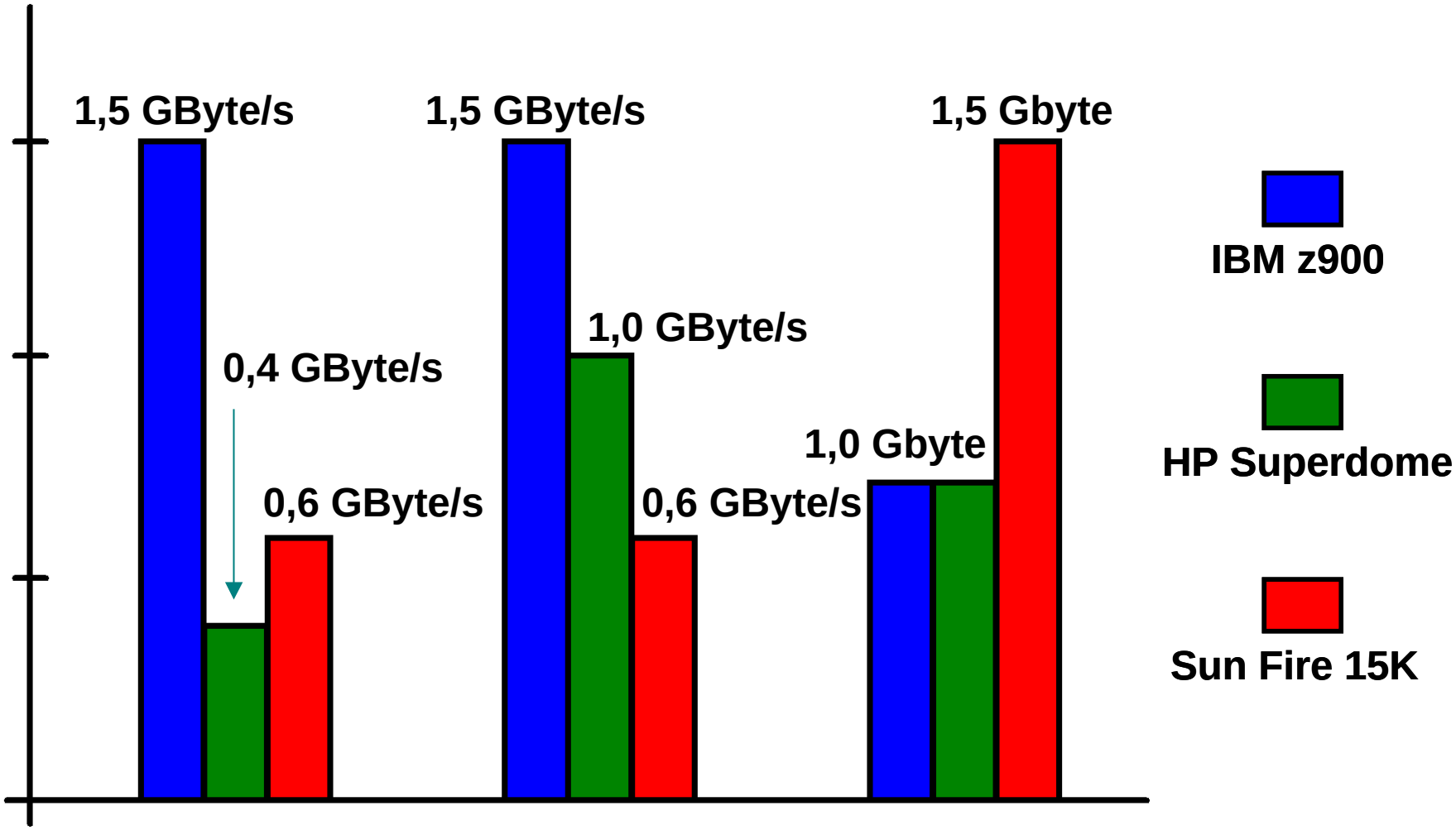
**20 Mikroprozessor Chips,
16 CPU, 3 Channel Subsystem,
1 Spare**

**8 L2 Cache Chips, je 4 Mbyte,
je 234 Mill. Transistoren**

2 Storage Control Chips

4 Memory Bus Adapter Chips

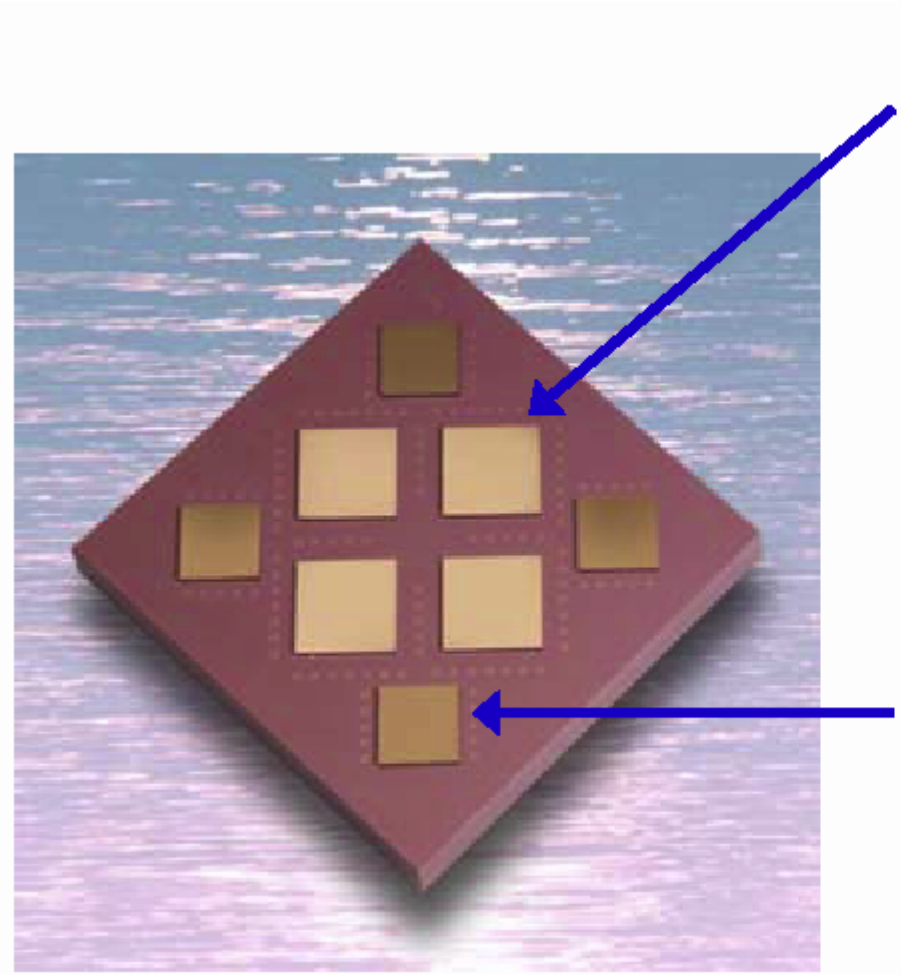
1 Clock Chip



System interne Daten Transfer Raten

Simultaneous Multithreading (SMT)

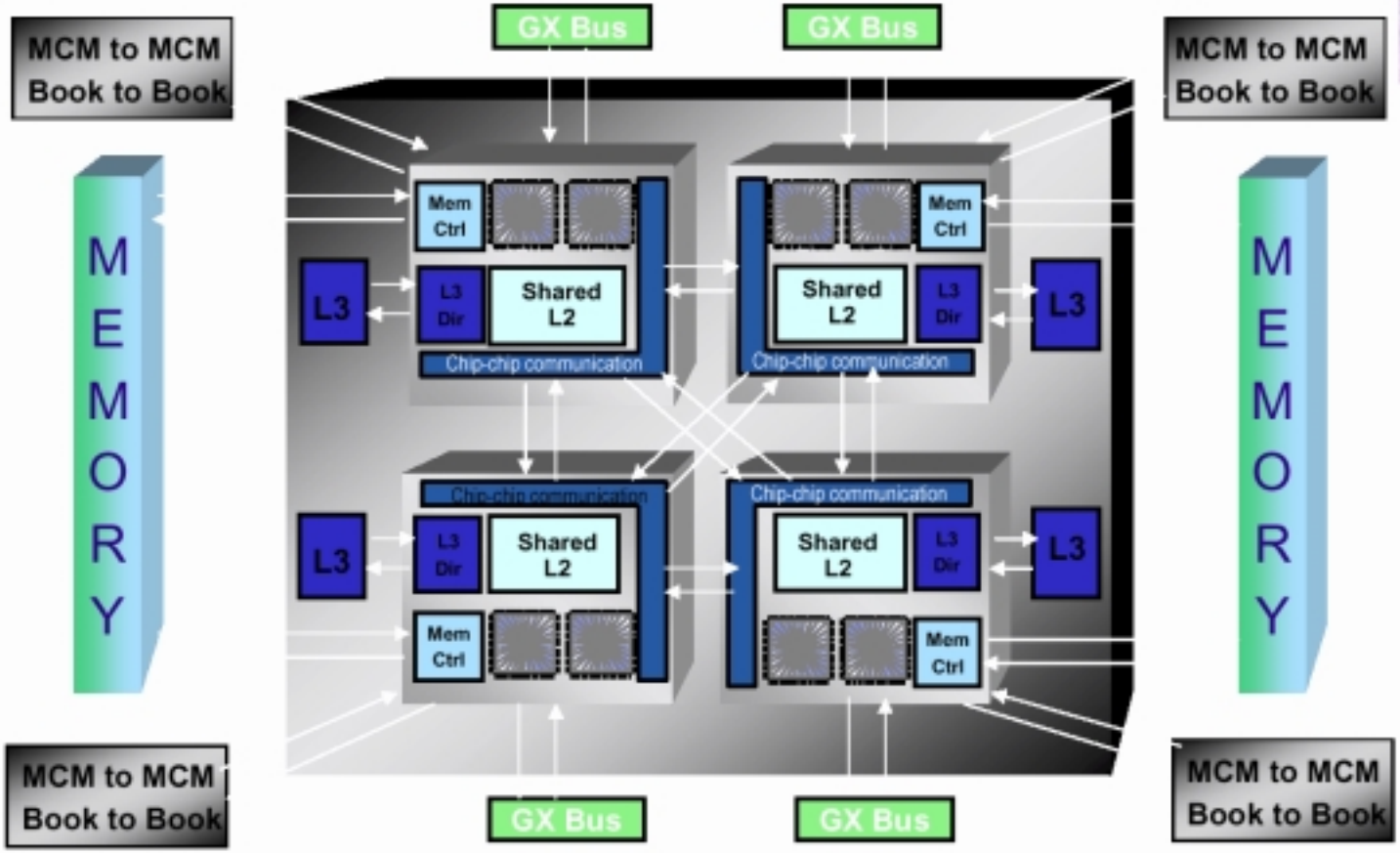
The POWER5 processor is the first to introduce simultaneous multithreading (SMT) technology into the IBM Power Architecture. There are two independent physical processors on a POWER5 chip. SMT provides two independent threads of instruction execution for each physical processor. Each thread appears to the operating system as a logical processor.

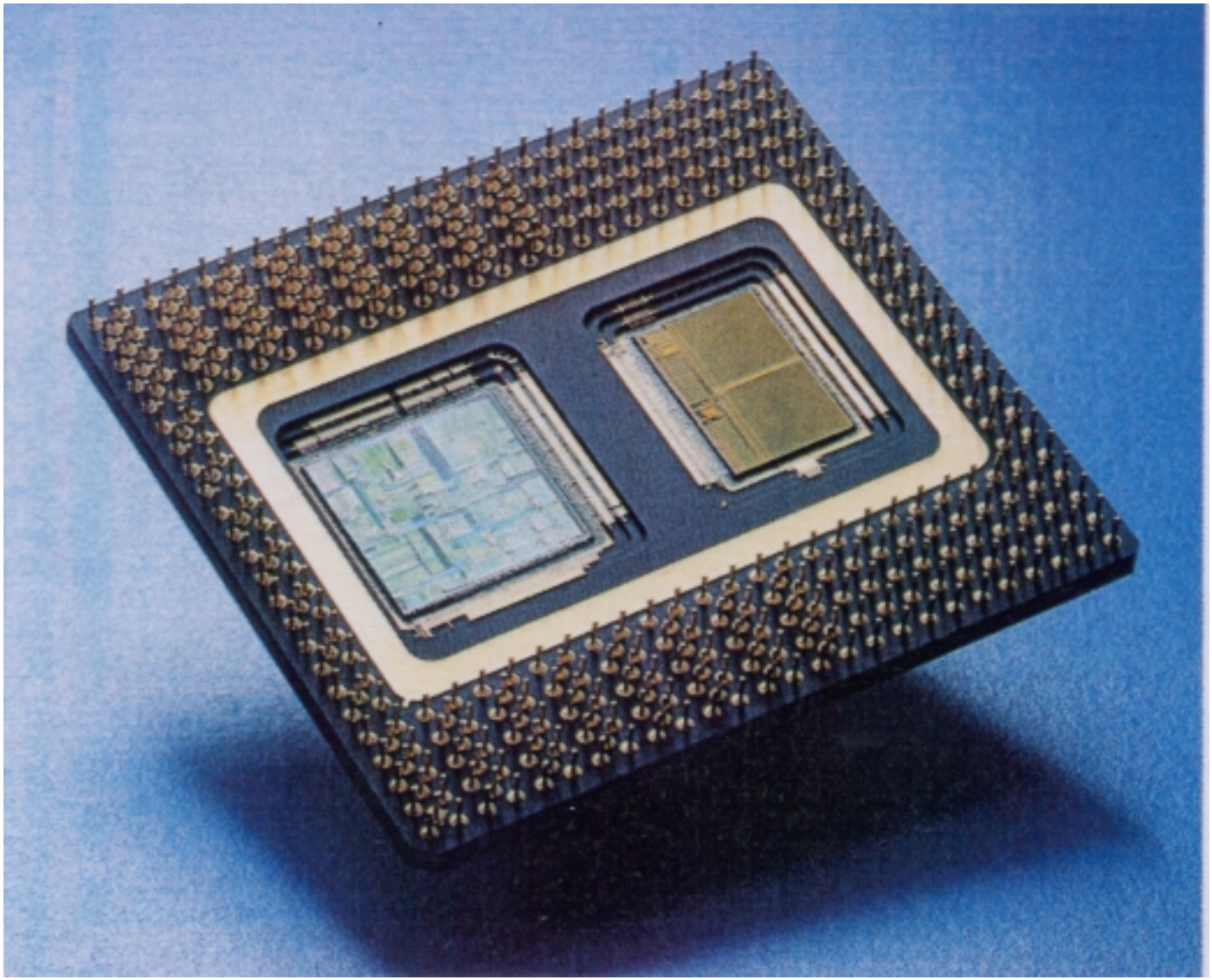


- POWER5 chip
  - 4 chips (8 cores) per MCM
  - 2 cores per chip
  - 4 logical processor with SMT
  - Up to 1.9 MB shared L2 cache per chip
  - Integrated L3 cache & Memory controller
- L3 cache
  - Up to 36 MB
  - Up to 144 MB per MCM

POWER5 Multi-chip Module (MCM)

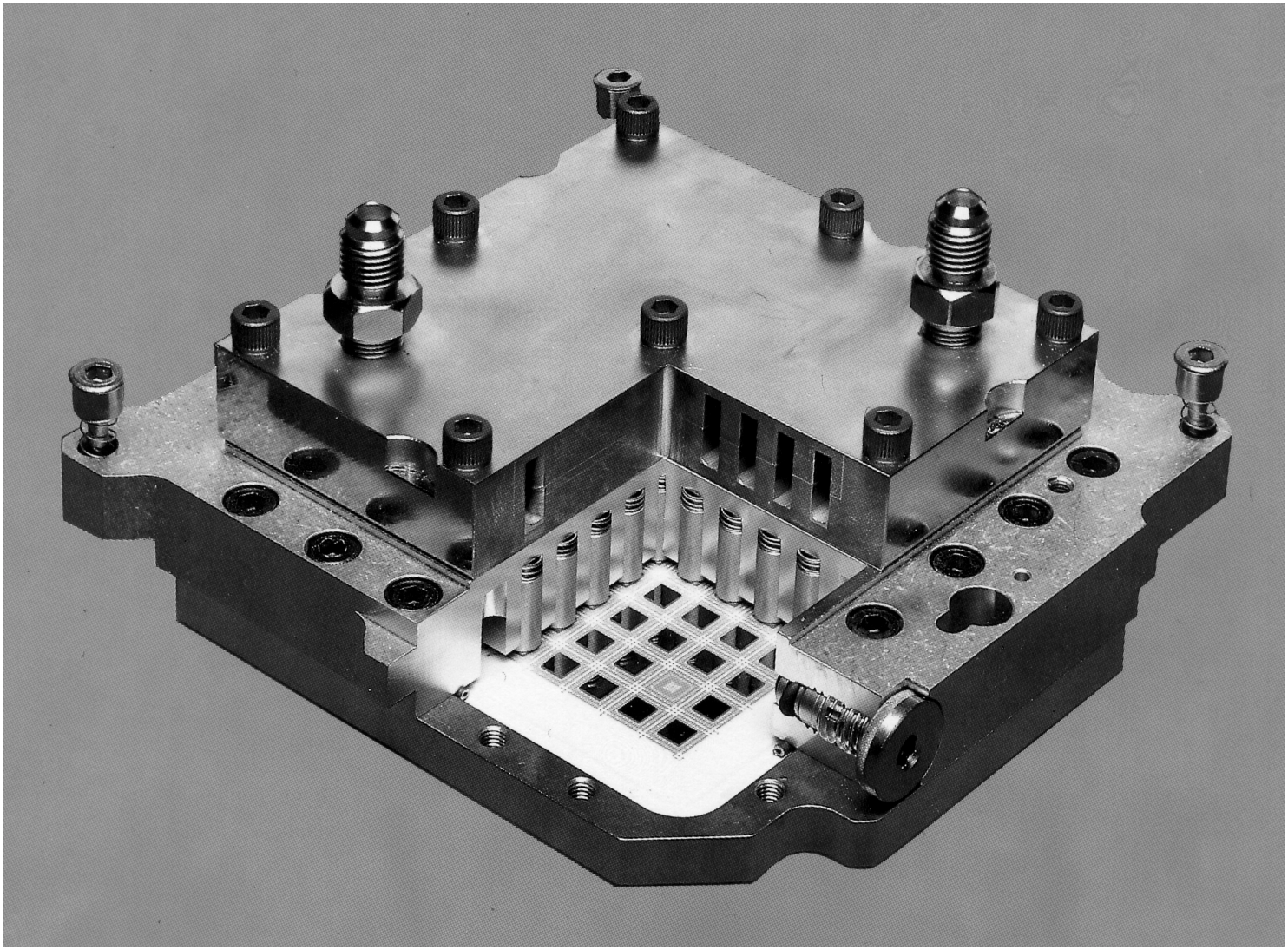
(Multichip Module in einem 64-way server)



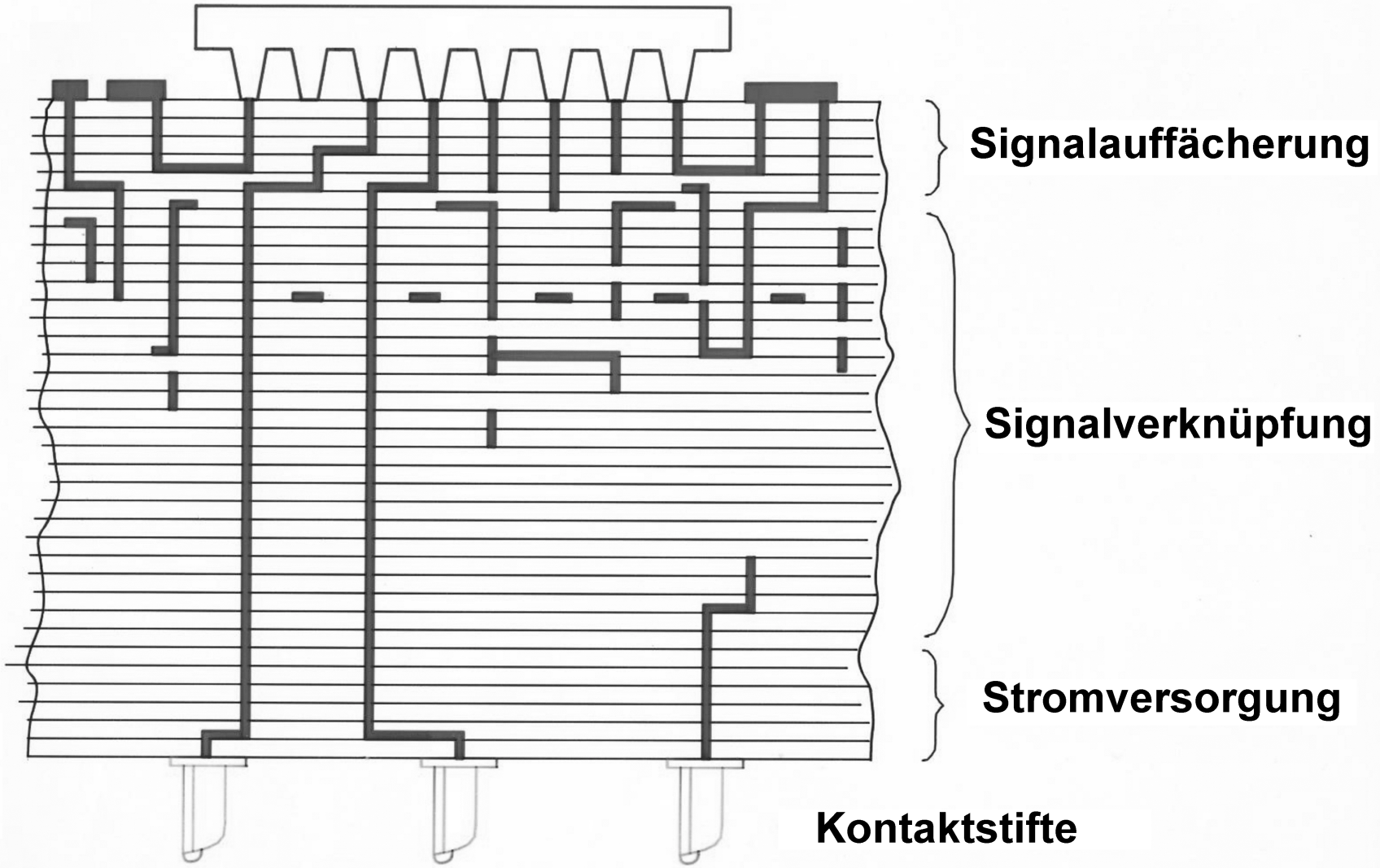


Pentium Pro

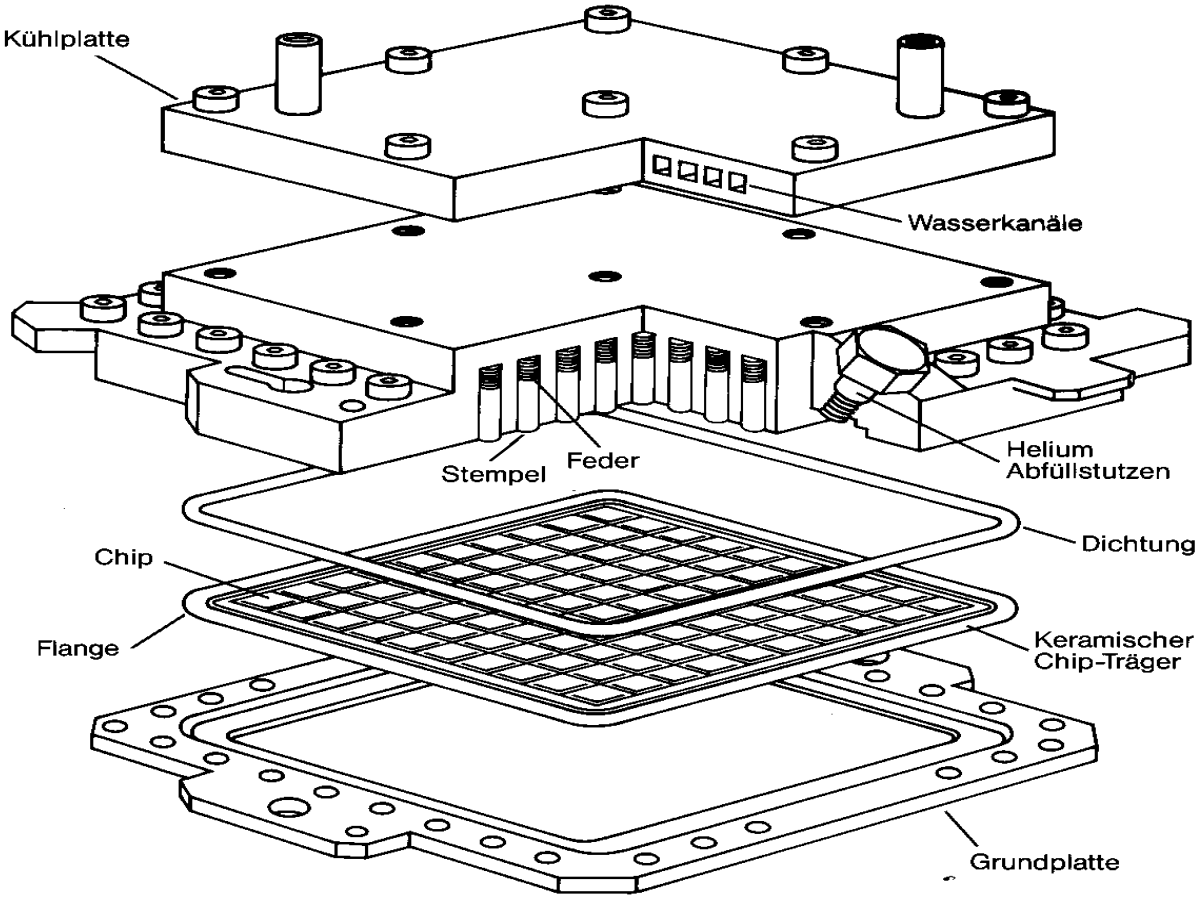
387 Pin Multi Layer Ceramic (MLC) Multi Chip Carrier (MCM) Module



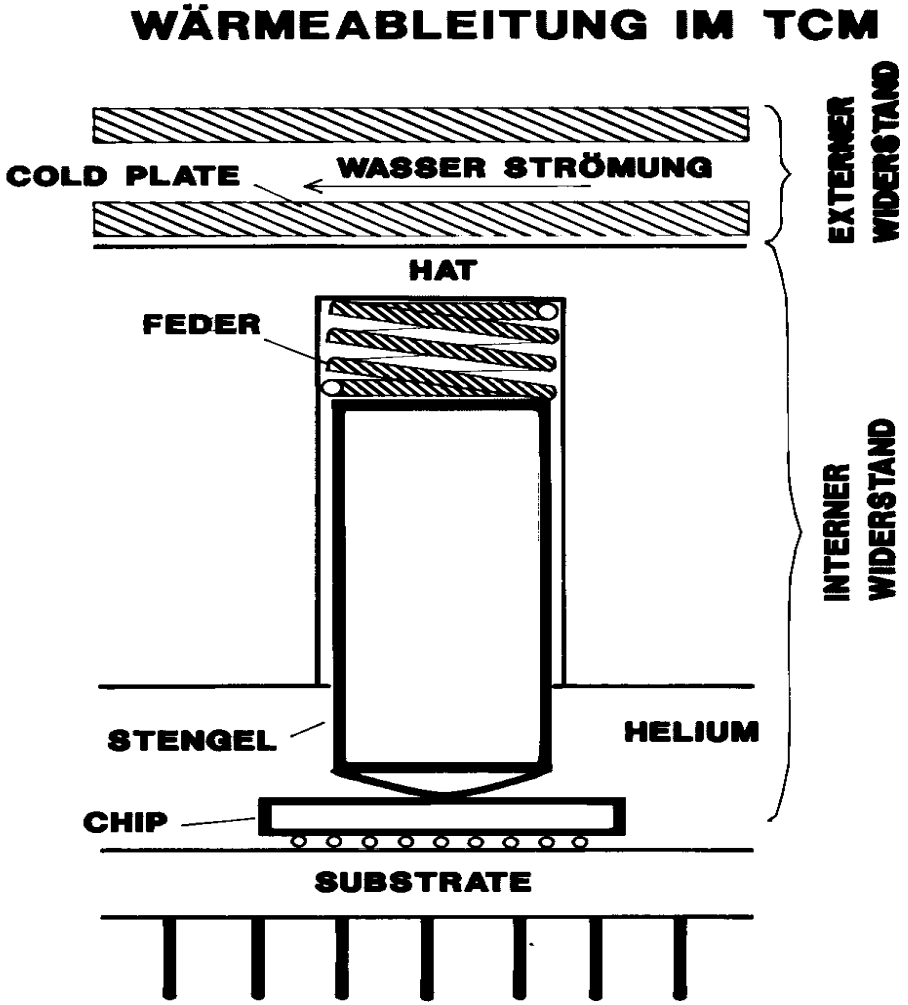
Thermal Conduction Module (TCM)



101 + 6 Schichten, über 4000 Kontaktstifte
1 km Verdrahtungslämge

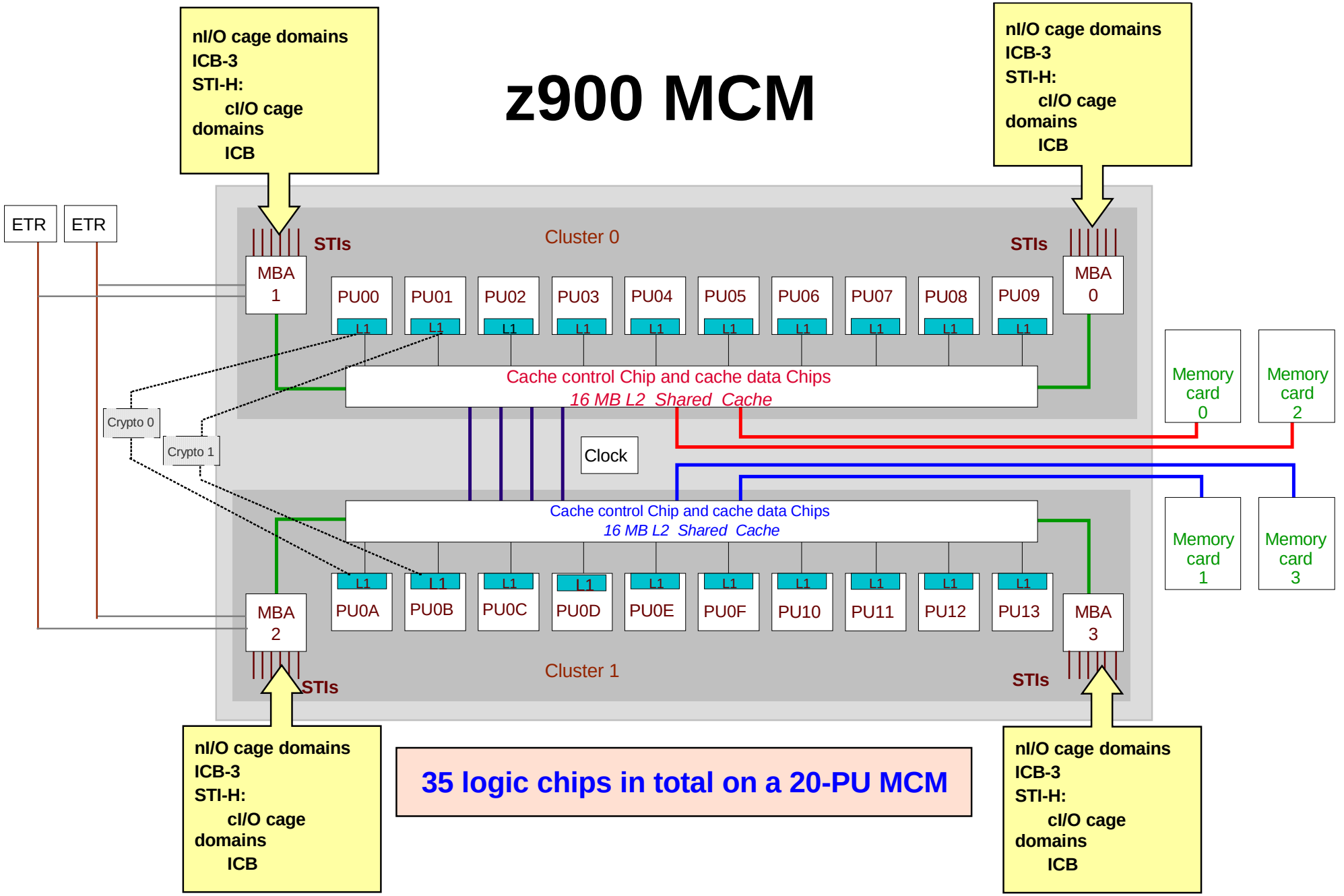


Aufbau eines „Thermal Conduction Module“

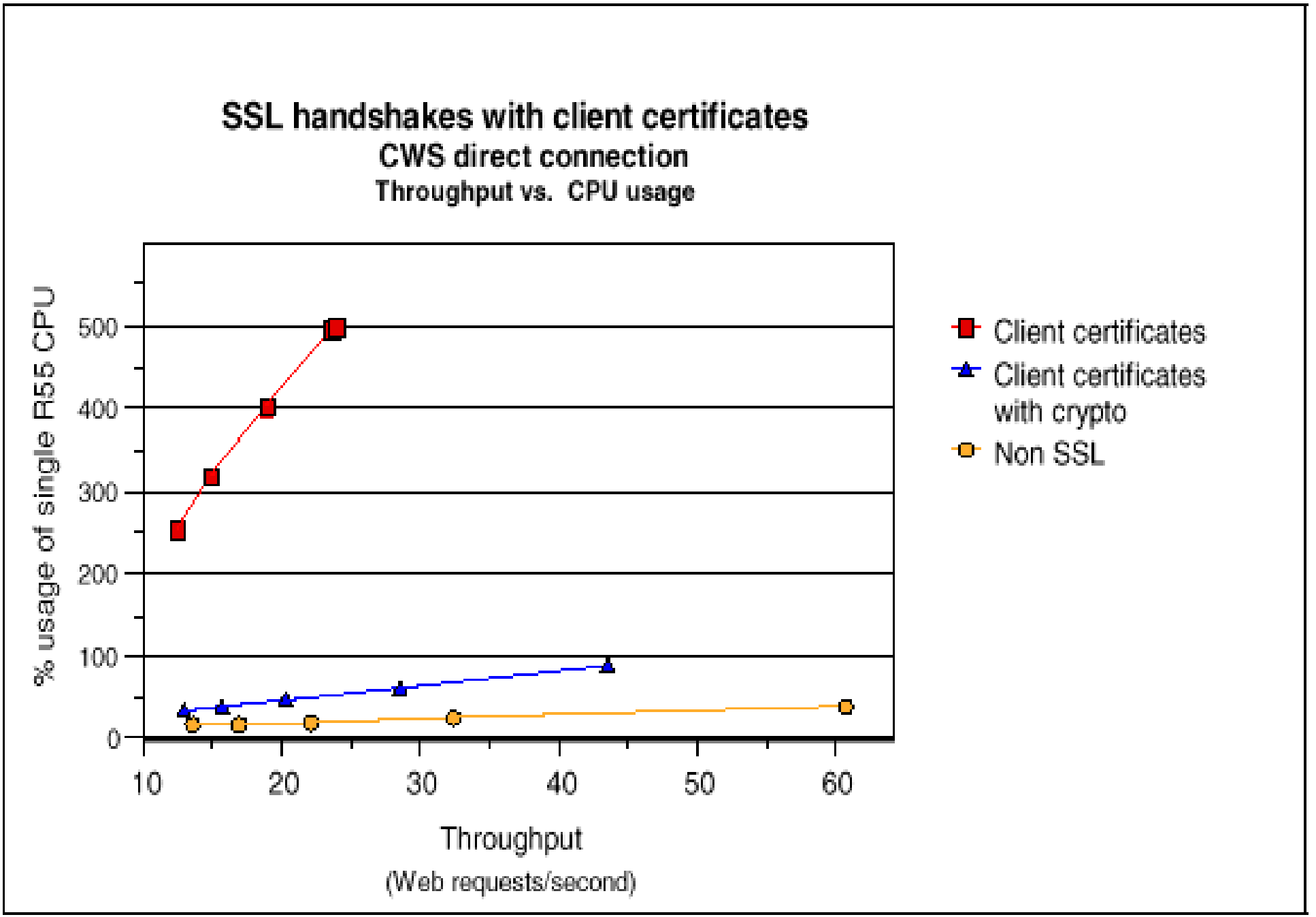


TCM Wärmeübergang

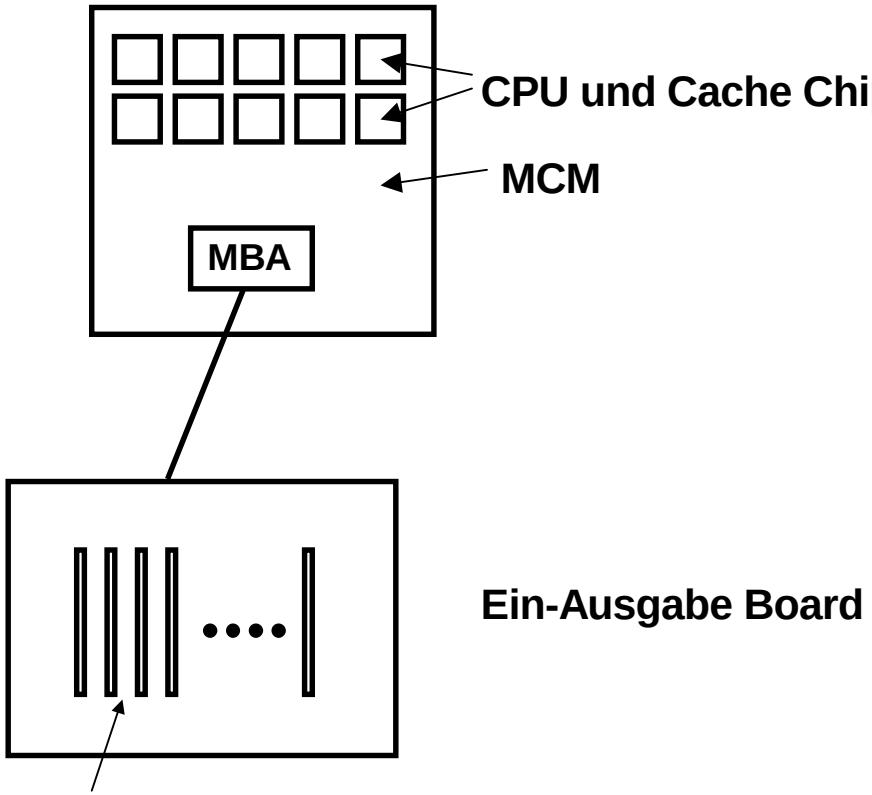
z900 MCM



35 logic chips in total on a 20-PU MCM


SSL handshakes with client certificates
CWS direct connection
Throughput vs. CPU usage
% usage of single R55 CPU
500
400
300
200
100
0
10
20
30
40
50
60
Throughput
(Web requests/second)
Client certificates
Client certificates with crypto
Non SSL

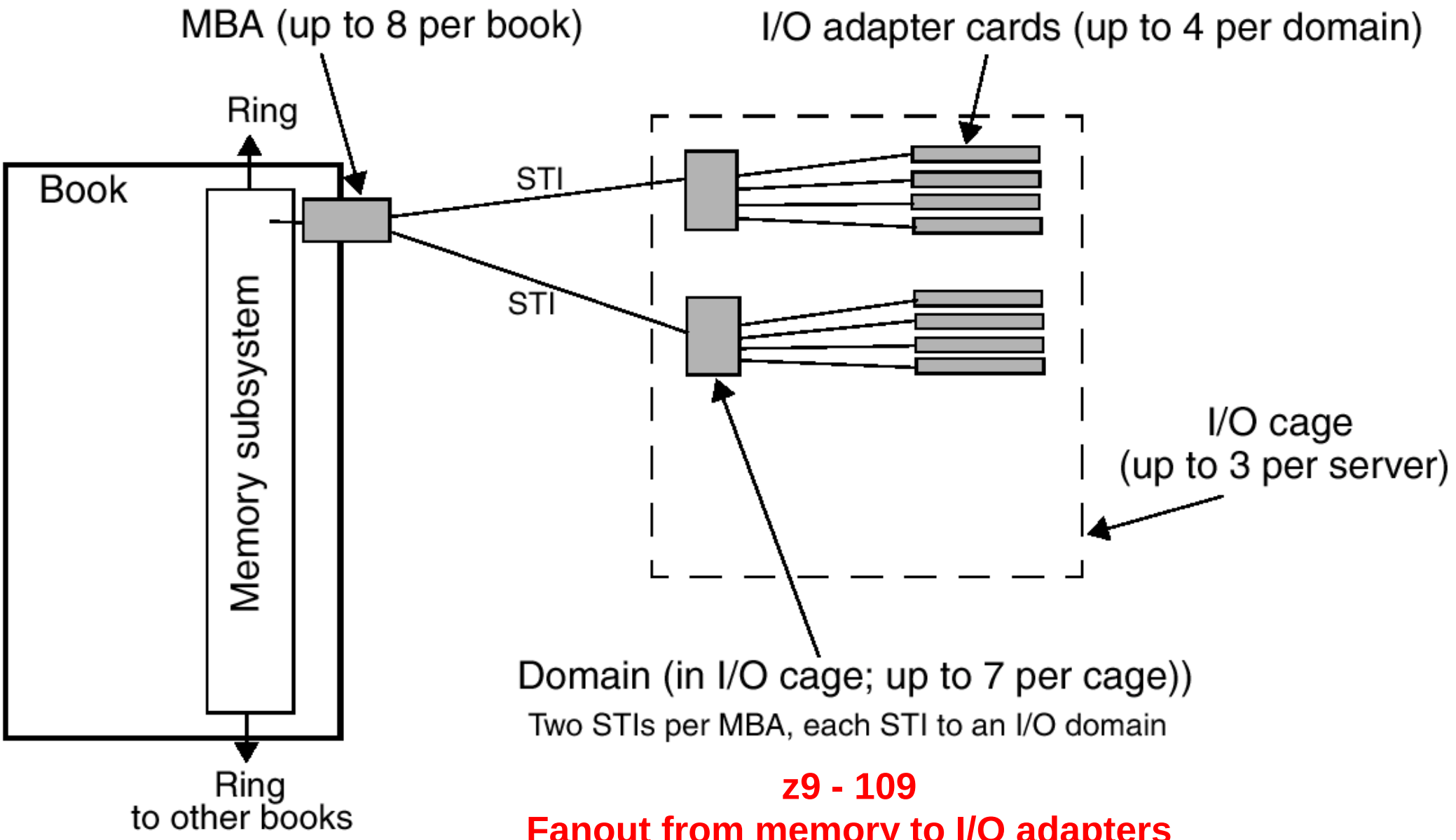
Ein-/Ausgabe Board



Auf dem Multichip Module sitzen neben den CPU- und Cache Chips vier Memory Bus Adapter (MBA) Chips, die eine ähnliche Rolle wie das Southbridge Chip in einem PC übernehmen. Aus jedem MBA werden sechs STI Busse herausgeführt, vergleichbar mit den PCI Bussen in einem PC. Die STI Busse münden in Ein-/Ausgabe Boards, die über STI Slots für die Aufnahme der Ein-/Ausgabeadapter Karten verfügen.

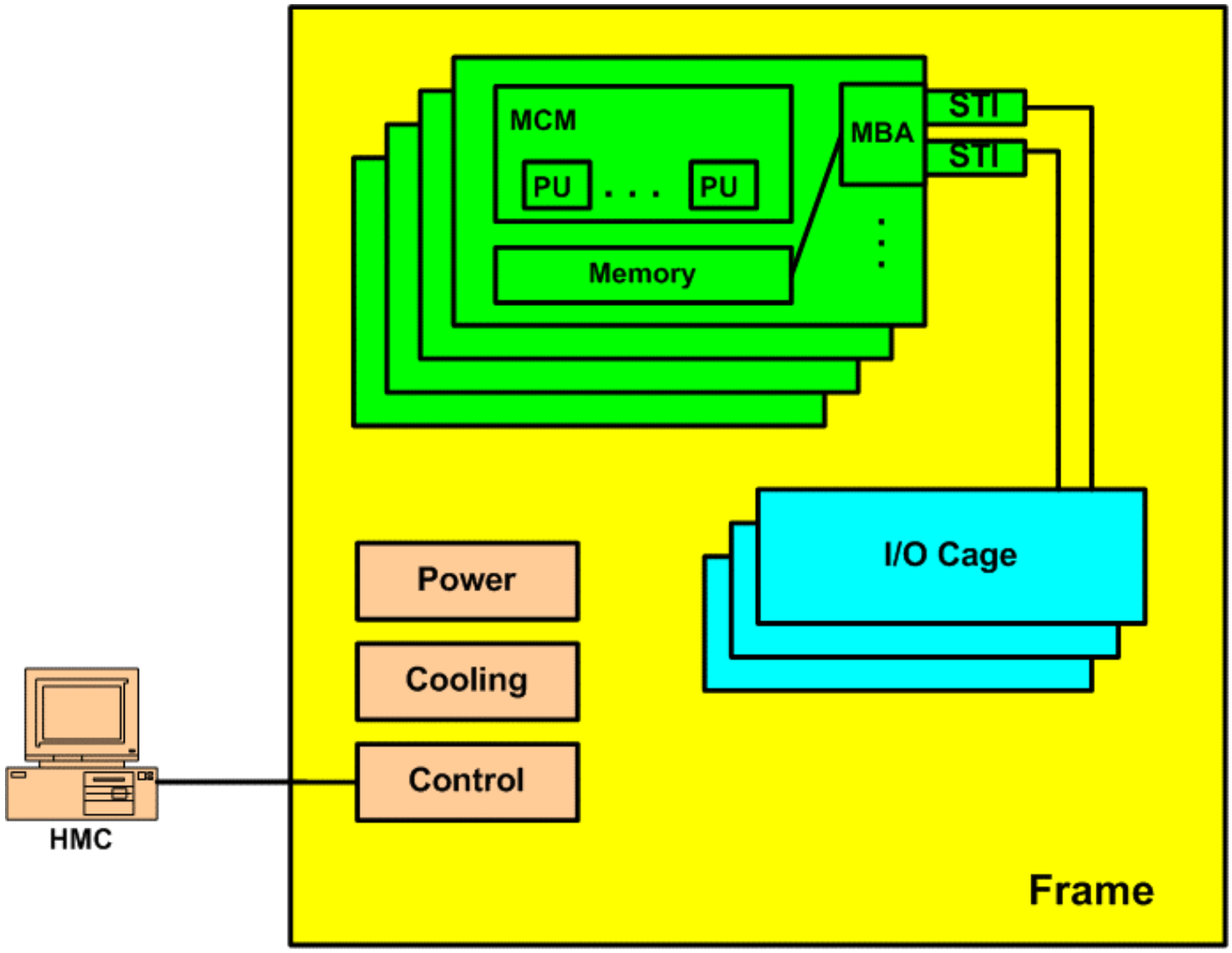
Die wichtigsten E/A Adapter Karten Typen haben Anschlüsse für

ESCON Kanal
FICON Kanal
OSA Adapter für Ethernet, ATM, ...

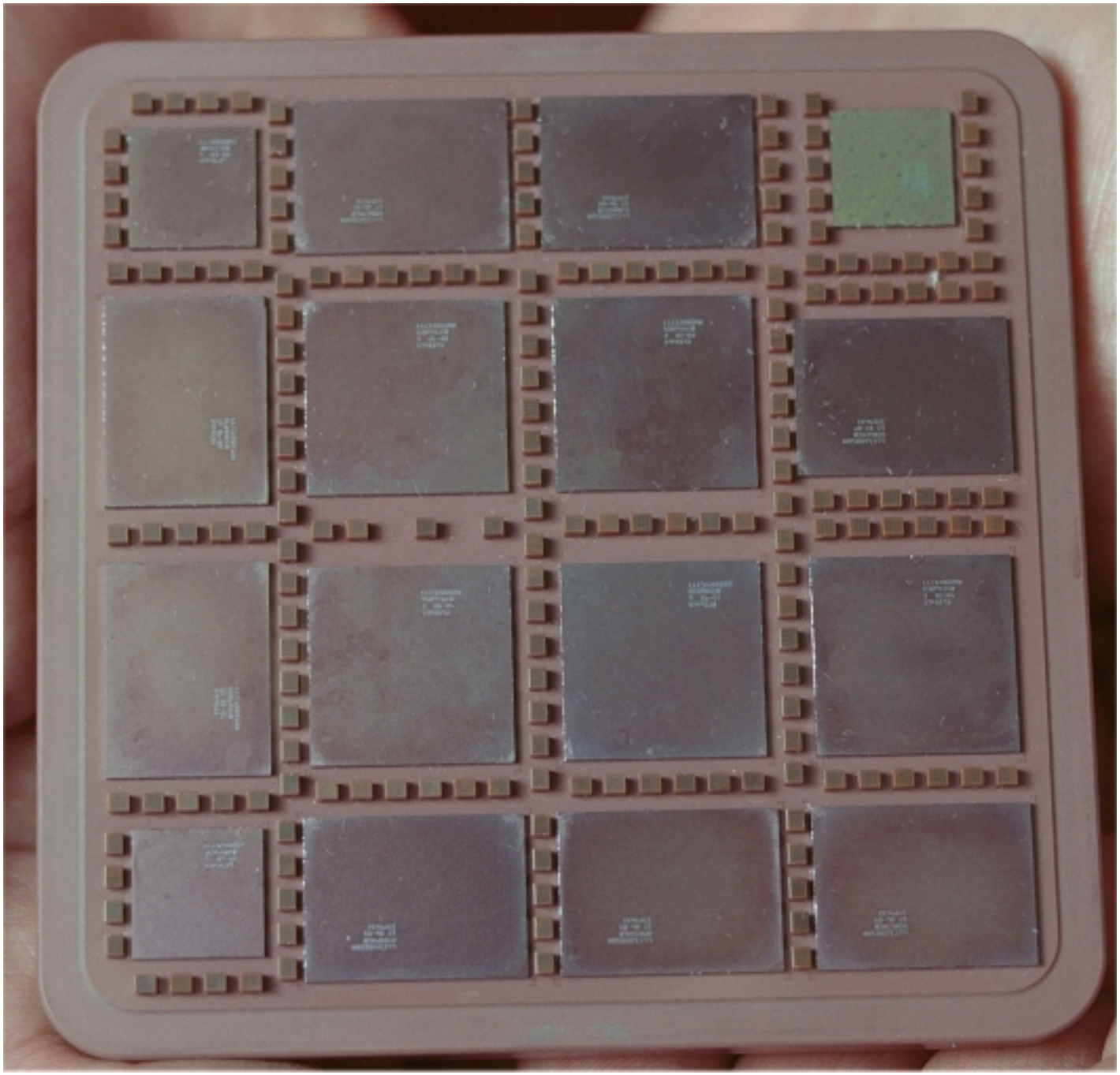


z9 - 109

Fanout from memory to I/O adapters



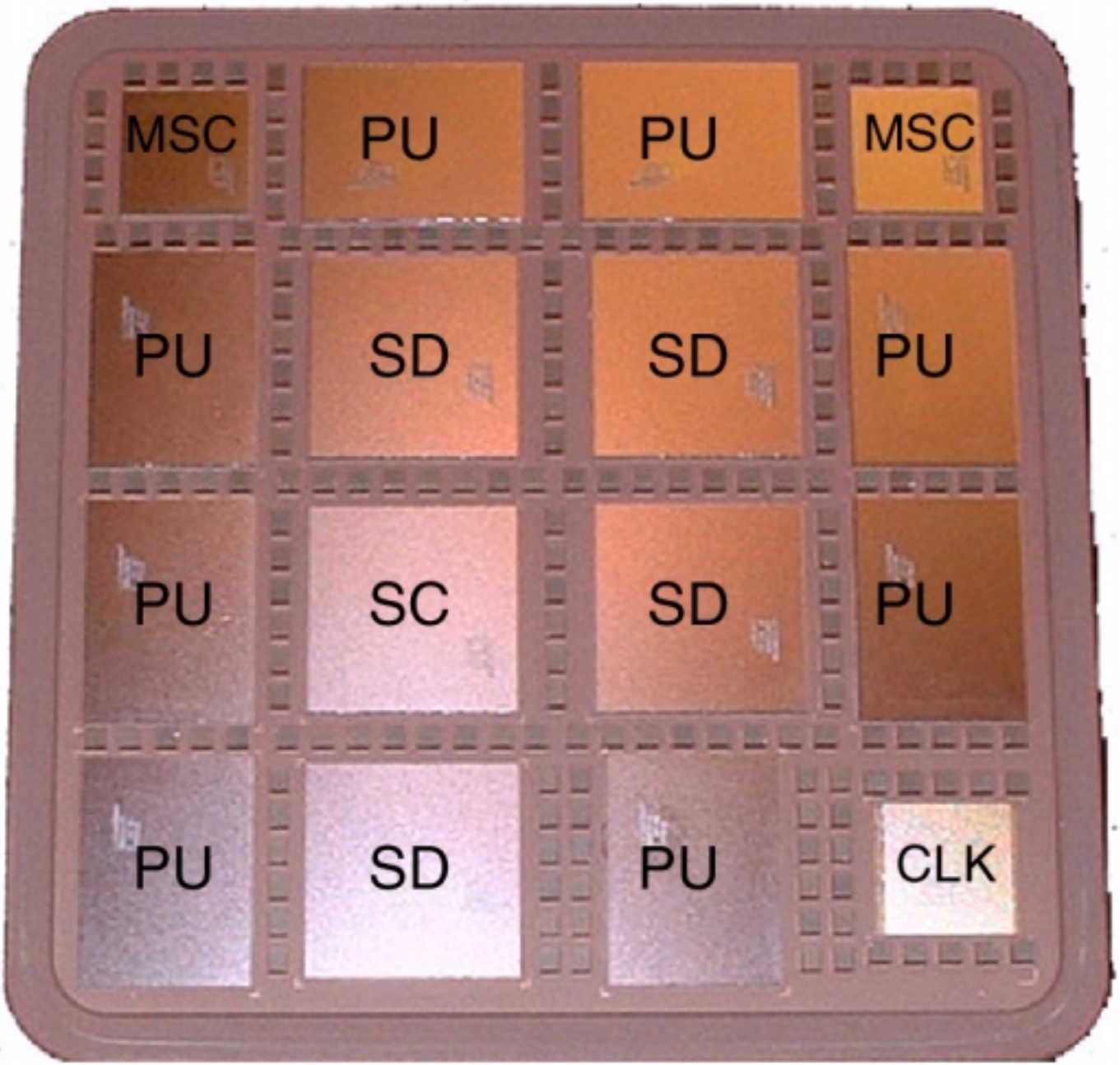
MCM = Multi Chip Module
STI = Self Timed Interface
HMC = Hardware Management Console
PU = Processing Unit
MBA = Memory Bus Adapter

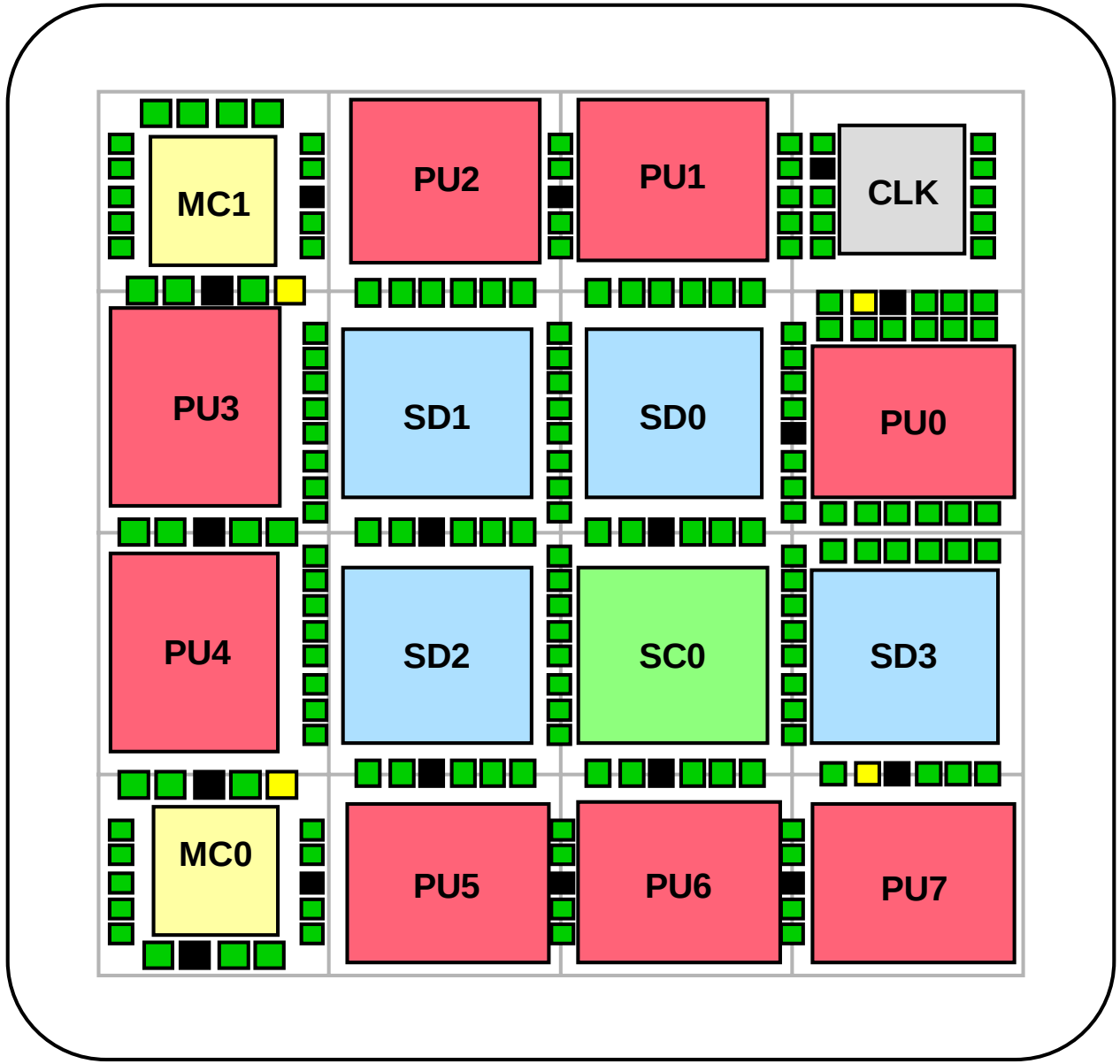


z990 Multichip Module (MCM)

SOI technology

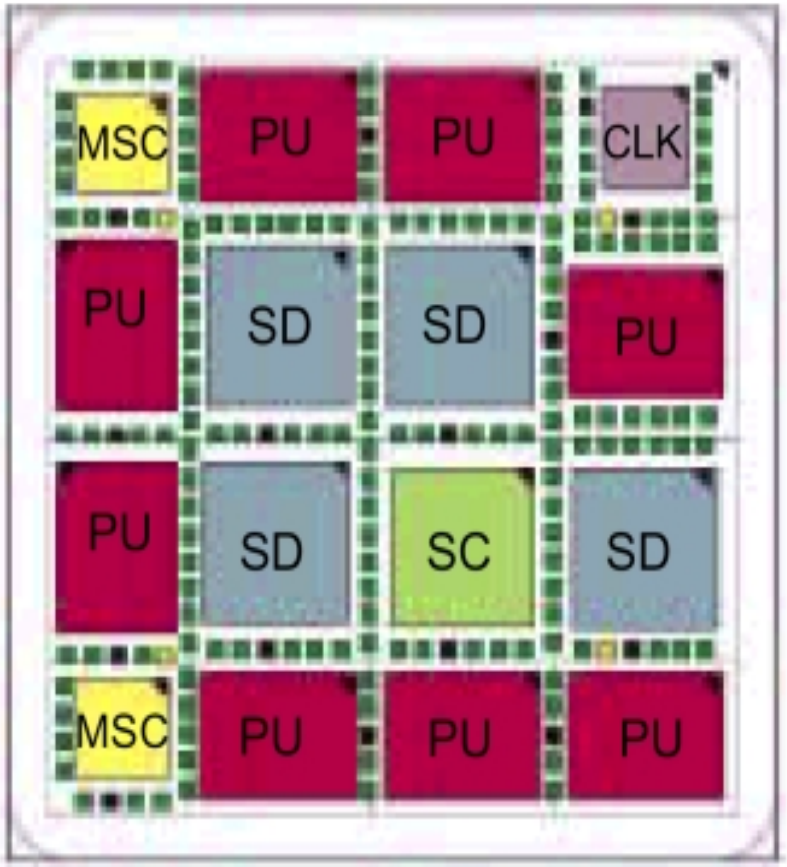
configuration of one to four books. Each book comprises a MCM And memory cards up to 64 GByte per book


MSC
PU
PU
MSC
PU
SD
SD
PU
PU
SC
SD
PU
PU
SD
PU
CLK


MC1
PU2
PU1
CLK
PU3
SD1
SD0
PU0
PU4
SD2
SC0
SD3
MC0
PU5
PU6
PU7

z990 12 Processor Unit (PU) MCM

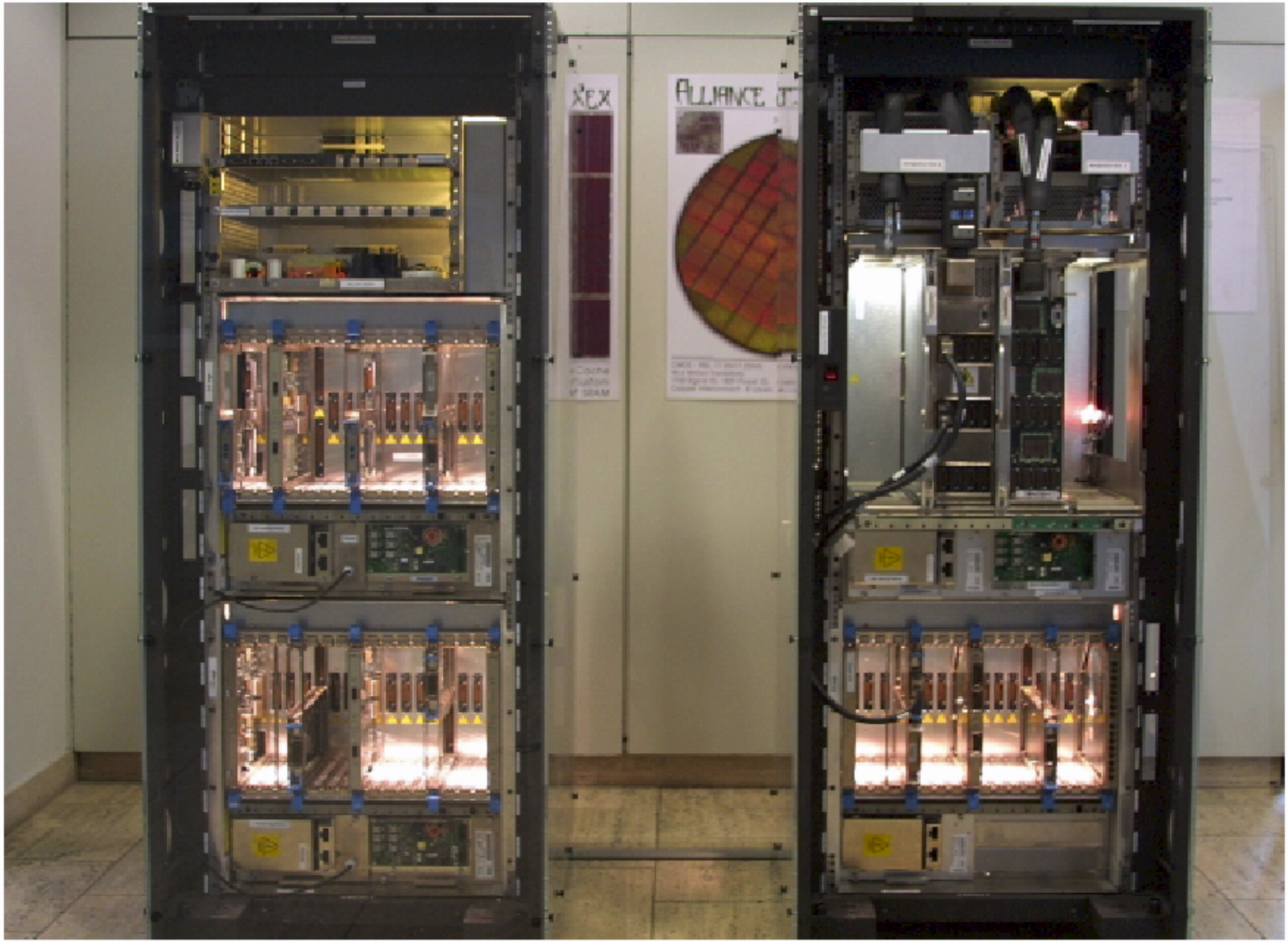
- Advanced 93mm x 93mm MCM
  - 16 chip sites, 185 capacitors
  - 100 Glass Ceramic layers
  - 1.3 Volts, 713 Watts
  - 46% smaller than z900

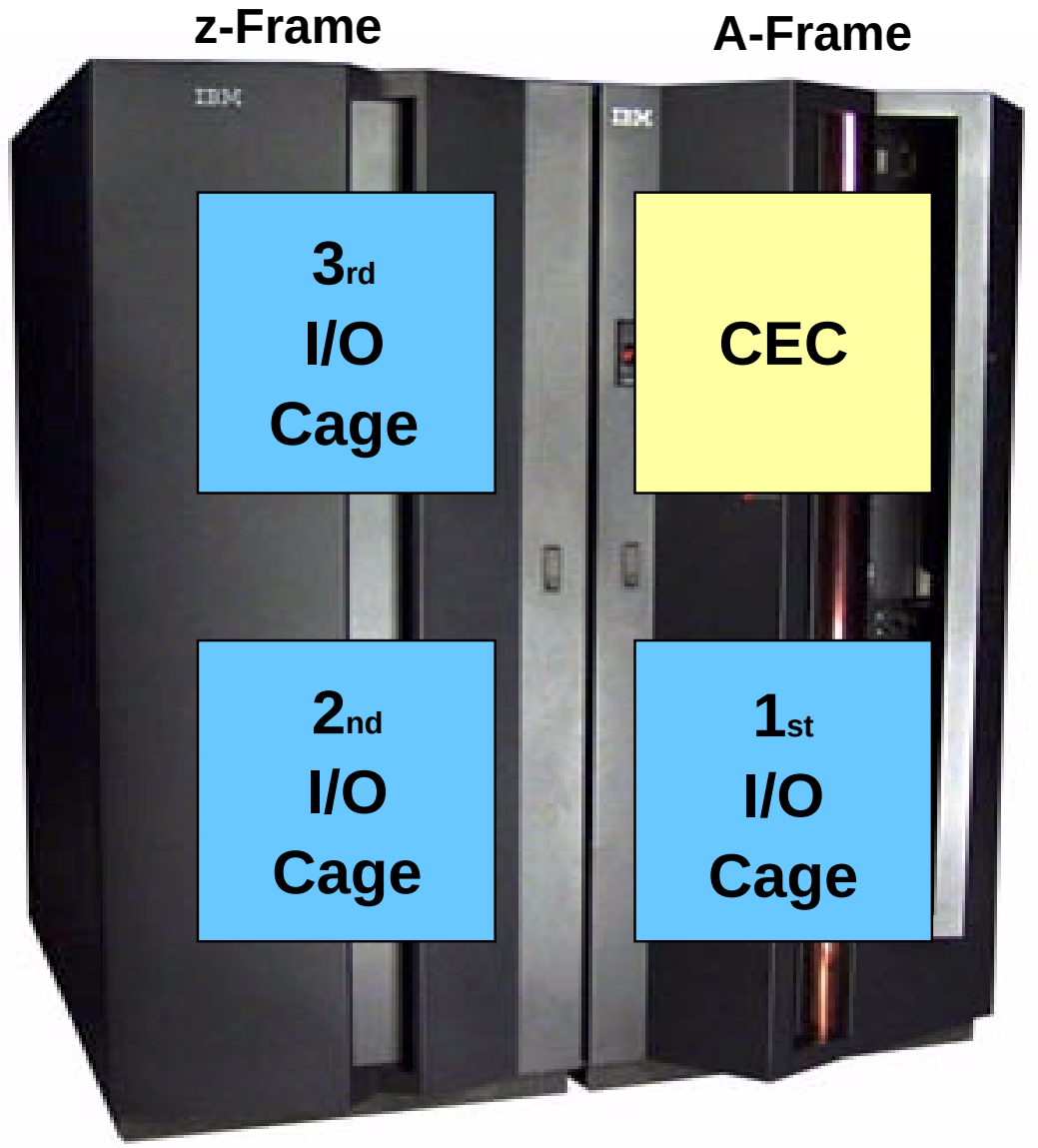


- CMOS 9SG - SOI chip technology
  - PU, SC, SD and MSC chips
  - Copper interconnections, 8 copper layers
- 4 Dual PU, 4 Single PU Chips per MCM
  - 14.1 mm x 18.9 mm
  - 122 million transistors/Chip
  - L1 cache/PU
    - 256 KB I-cache
    - 256 KB D-cache
- 4 System Data (SD) cache chips per MCM
  - 521 million transistors/chip
  - 8 MB L2 cache per chip
  - Single 32 MB L2 cache per MCM
- 1 Storage Control (SC) chip
  - L2 cache crosspoint switch
  - L2 access rings to/from other MCMs
  - L2 access to/from MBAs (off MCM)
- 2 Storage Control (MSC) chips
  - Memory cards (L3) interface to L2
- 1 Clock (CLK-ETR) chip
  - CMOS 8SF, 7 copper layers
  - Clock and ETR Receiver

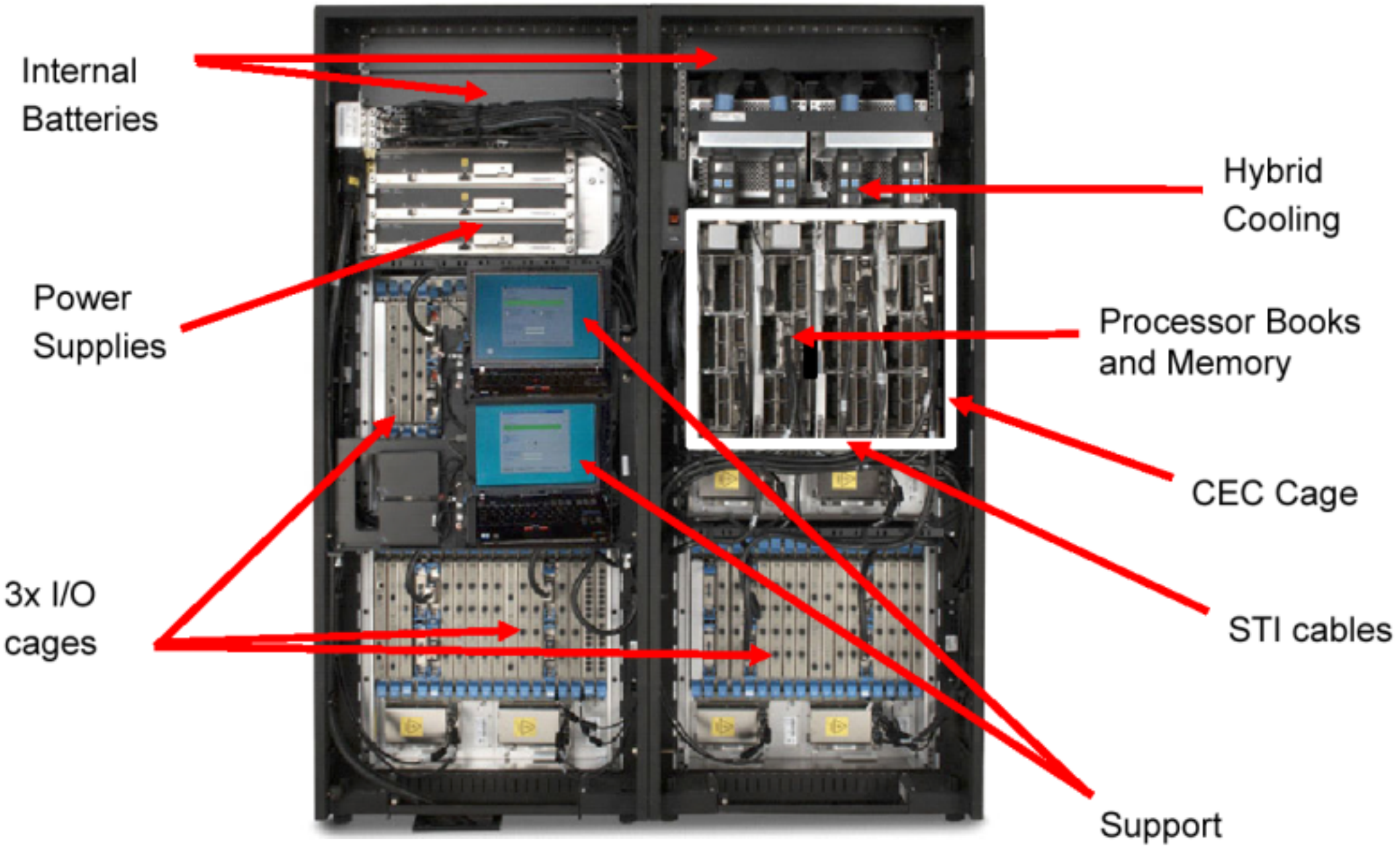
- CMOS 9S-SOI chip Technology
  - PU, SC, SD and MSC chips
  - Copper interconnections, 8 copper layers
  - 8 PU chips/MCM
    - 14.1 mm x 18.9 mm
    - 122 million transistors/PU
    - L1 cache/PU
      - 256 KB I-cache
      - 256 KB D-cache
    - 0.83 ns Cycle Time
  - 4 System Data (SD) cache chips/MCM
    - 17.5 mm x 17.5mm
    - World's densest chip
    - L2 cache per Book
      - 521 million transistors/chip
      - 32 MB
  - One Storage Control (SC) chip
    - 17.3mm x 17.3mm
    - 98 million transistors
    - Densest I/Os
    - 3692 Power Signal I/Os
    - L2 cache crosspoint switch
    - L2 access rings to/from other MCMs
    - L2 access to/from MBAs (off MCM)
  - Two Storage Control (MSC) chips
  - Memory cards (L3) interface to L2
  - One Clock (CLK) chip - CMOS 8S
    - Clock and ETR Receiver


ALLIANCE


ALLIANCE

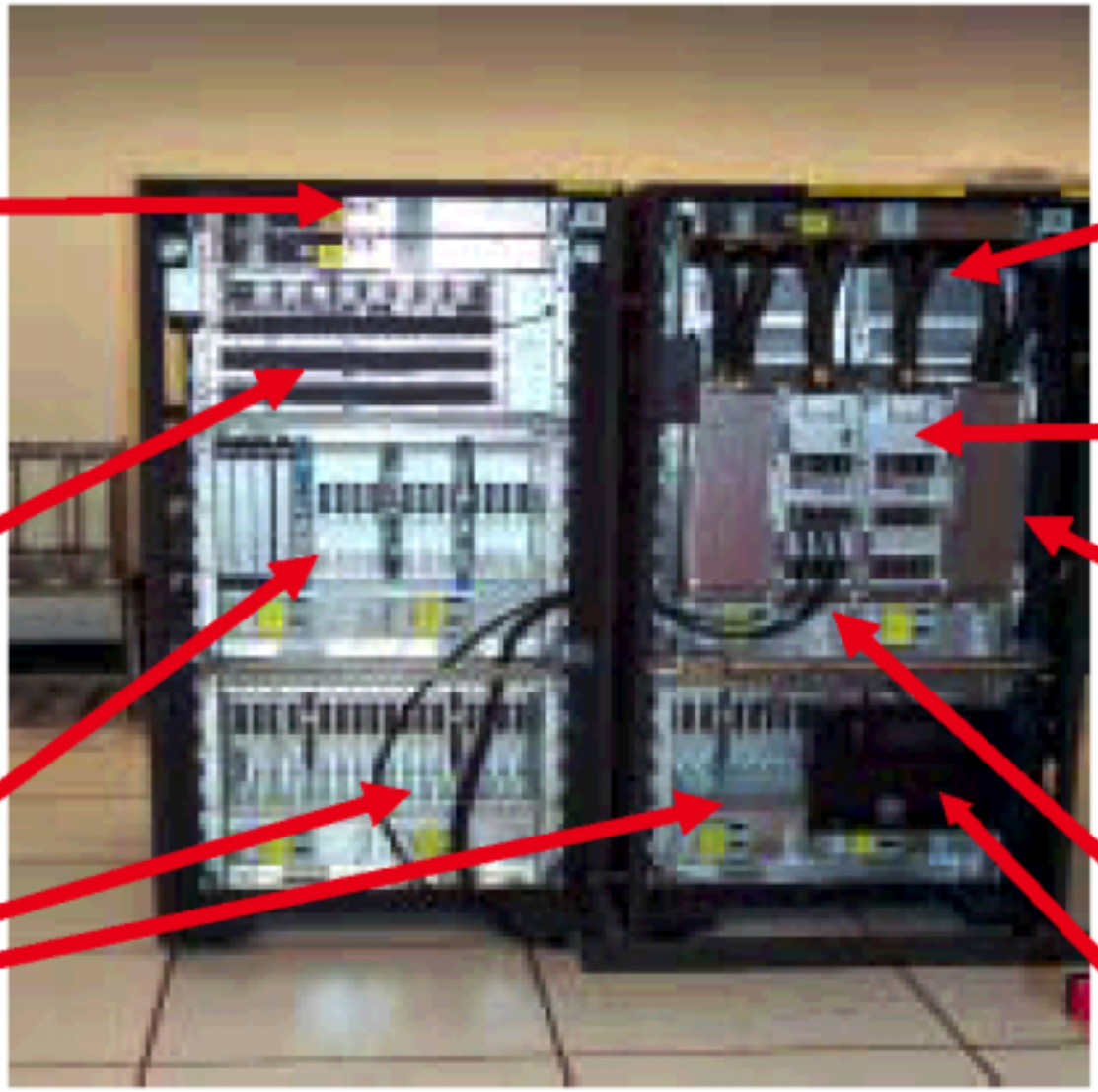


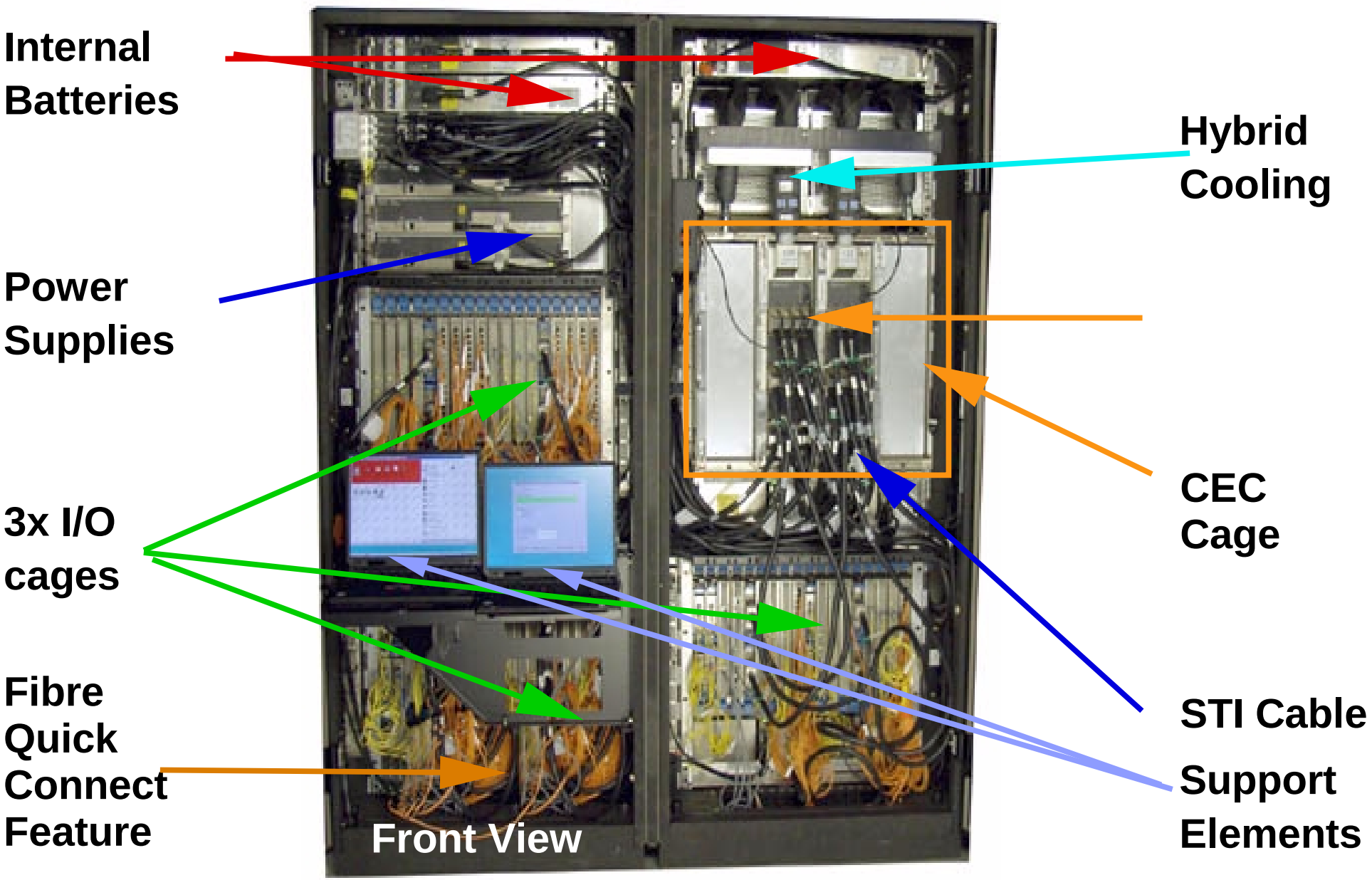
Front View

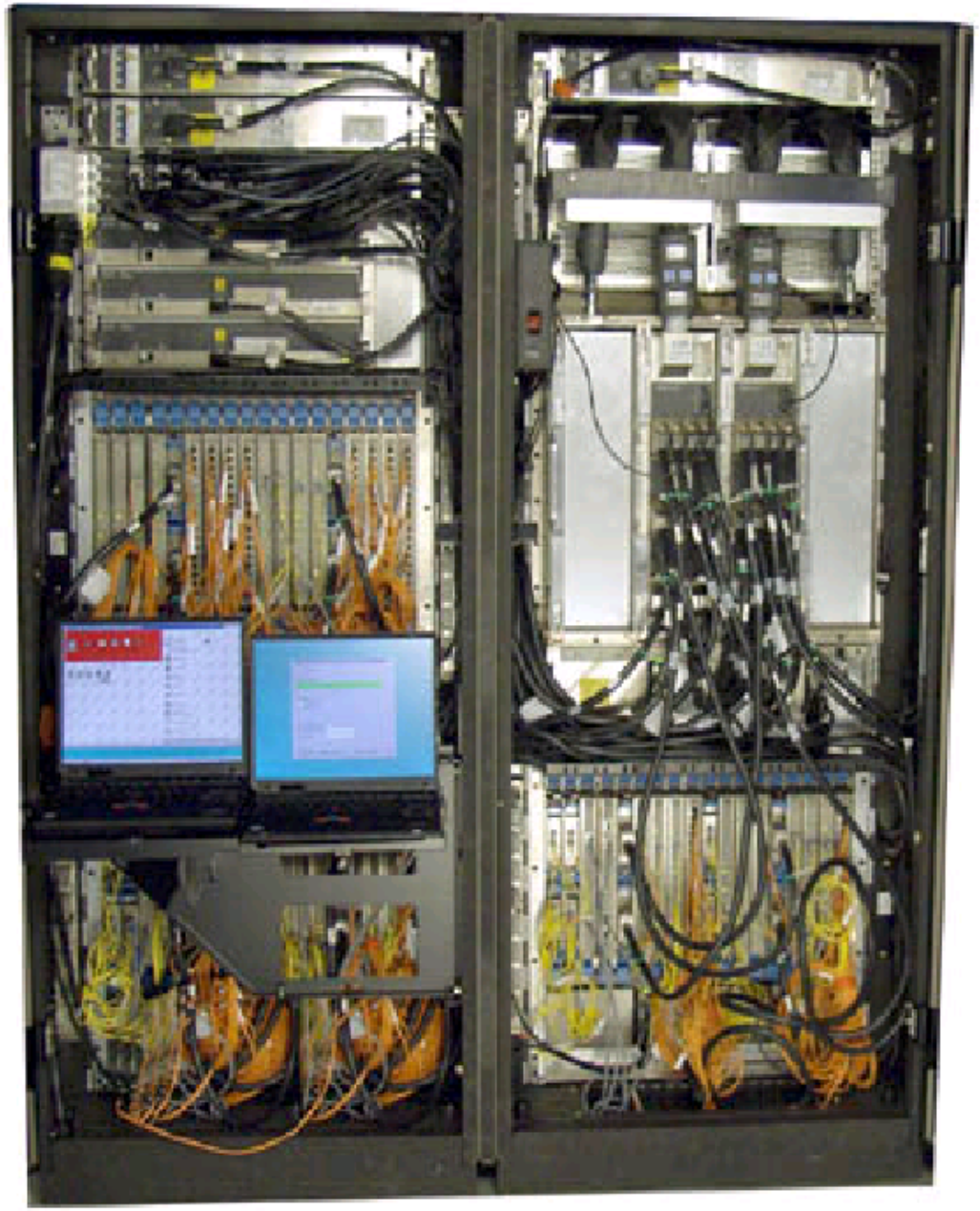


z9-109 Server – Ansicht von vorne

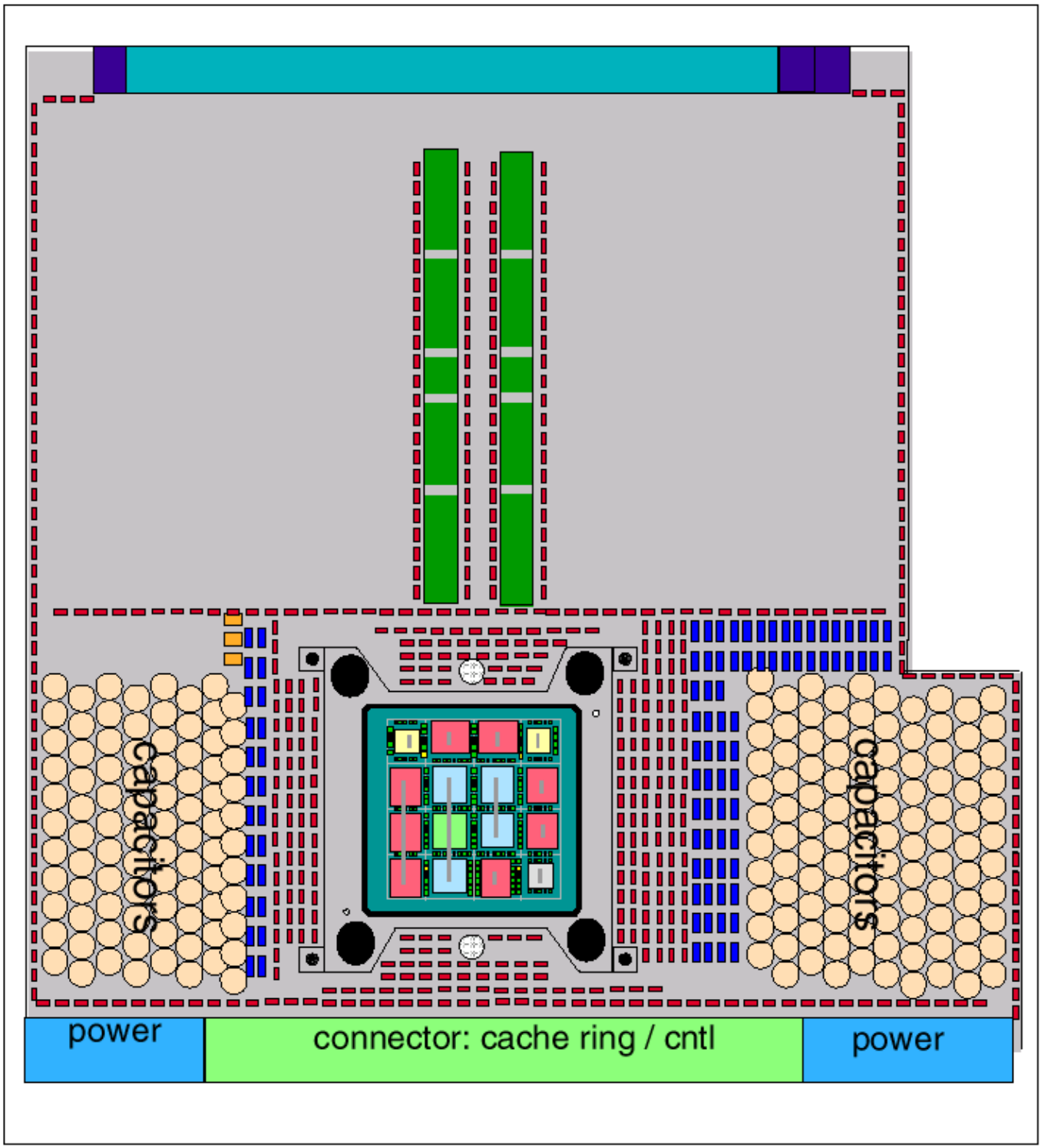
z990 System


Internal Batteries
Cooling Units
Processor Book
Power Supply
CEC Cage
IO cages
STI cables
Service Element


Internal Batteries
Power Supplies
3x I/O cages
Fibre Quick Connect Feature
Hybrid Cooling
CEC Cage
STI Cable
Support Elements
Front View

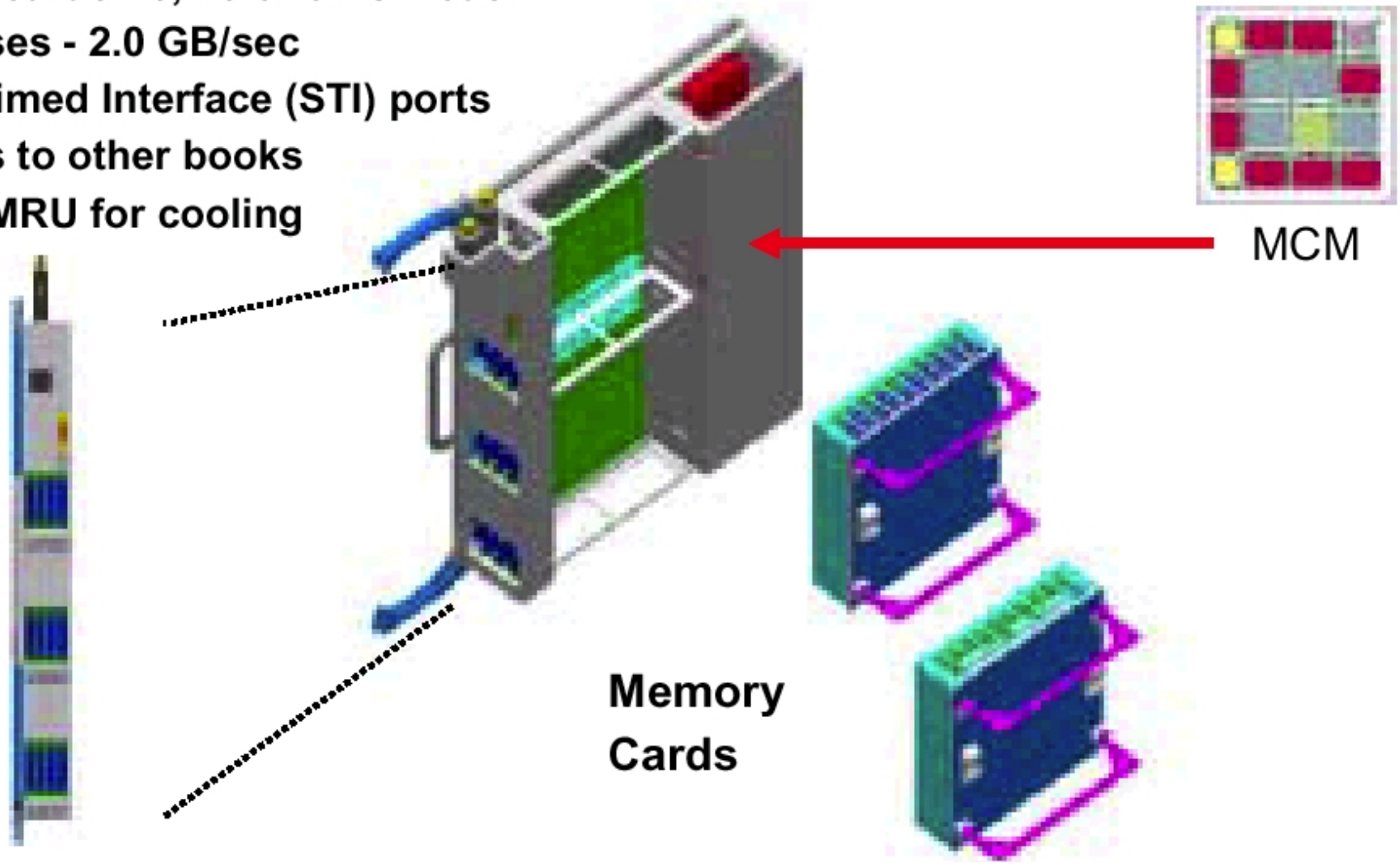



D46
D47
eserver

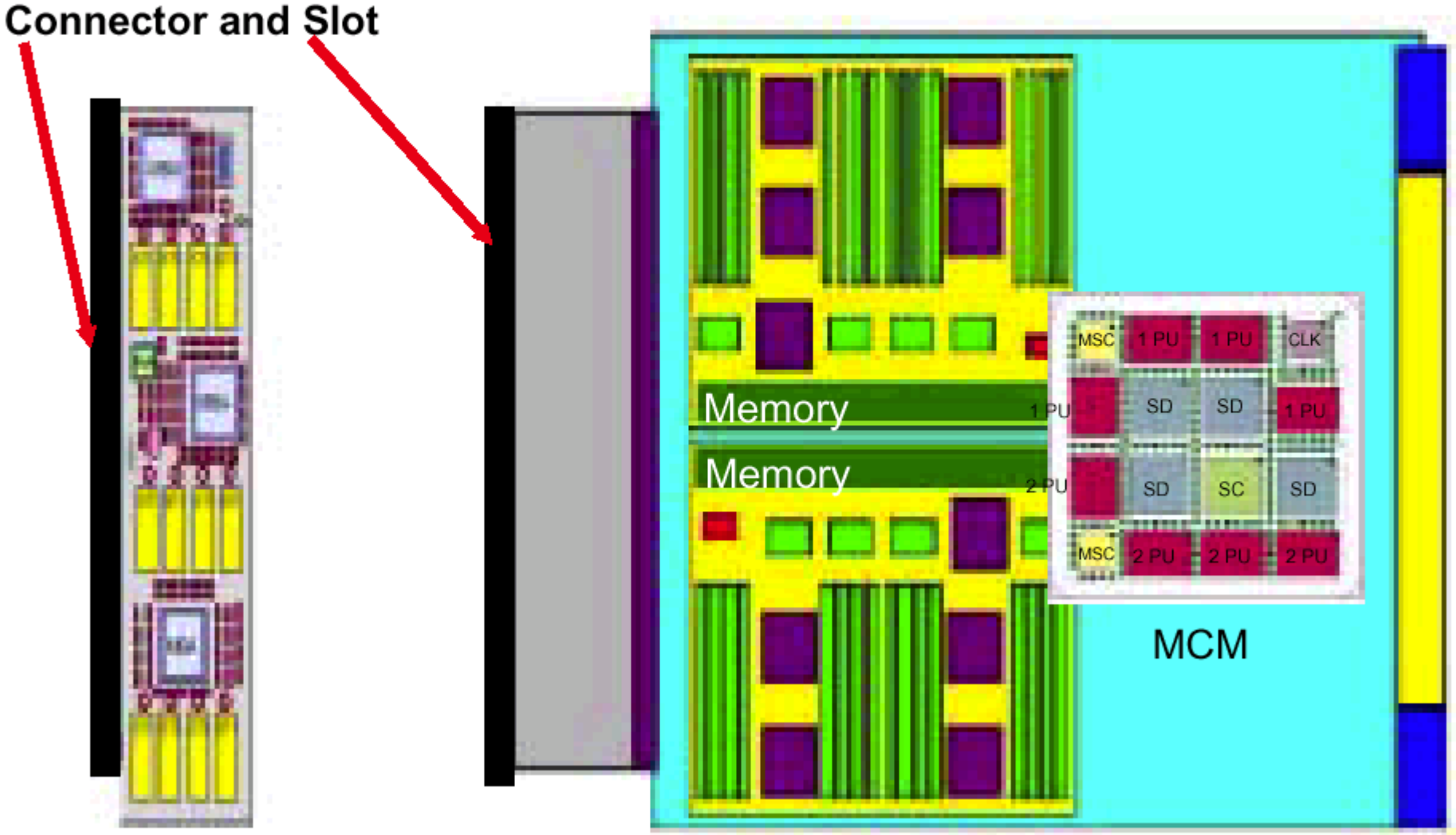

capacitors
capacitors
power
connector: cache ring / cntl
power

z990 Processor Book

- Multichip module with 12 processor units
- Two memory cards - 8, 16 or 32 GB each
- Channel busses - 2.0 GB/sec
  - 12 Self-Timed Interface (STI) ports
- Interconnects to other books
- Connects to MRU for cooling

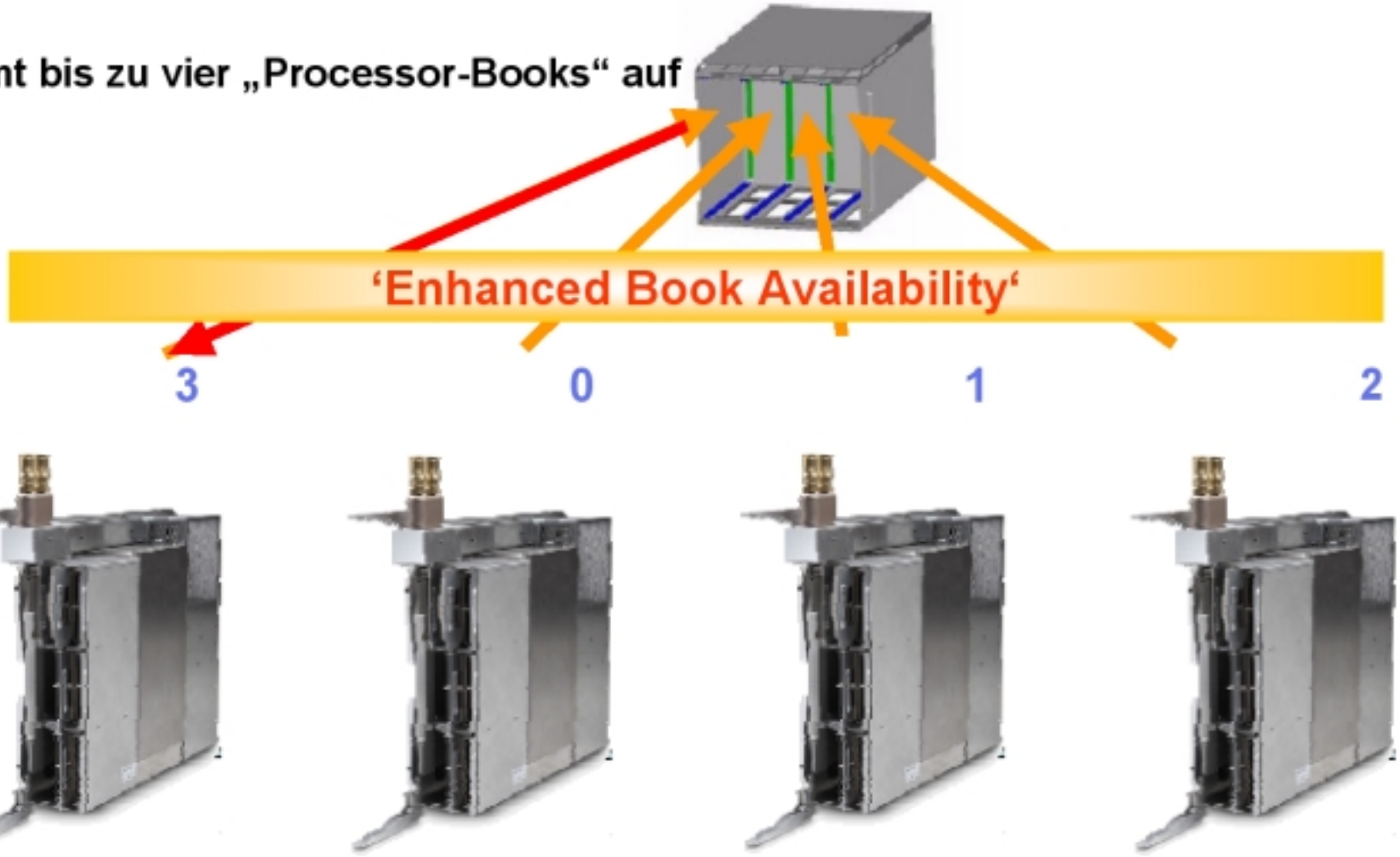


z990 Processor Book Riser Card
Memory Bus Adapters (MBAs) and STI Ports

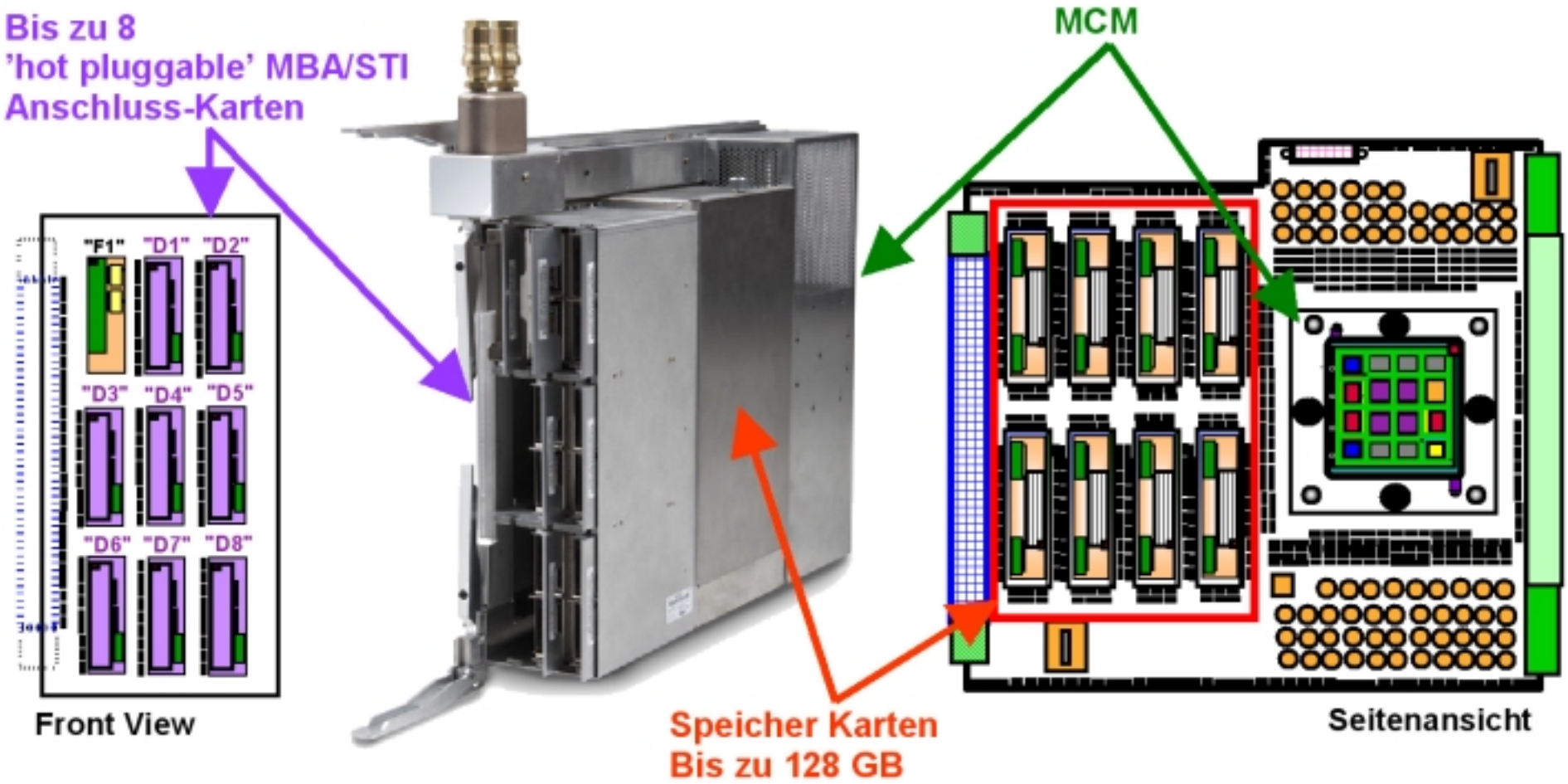


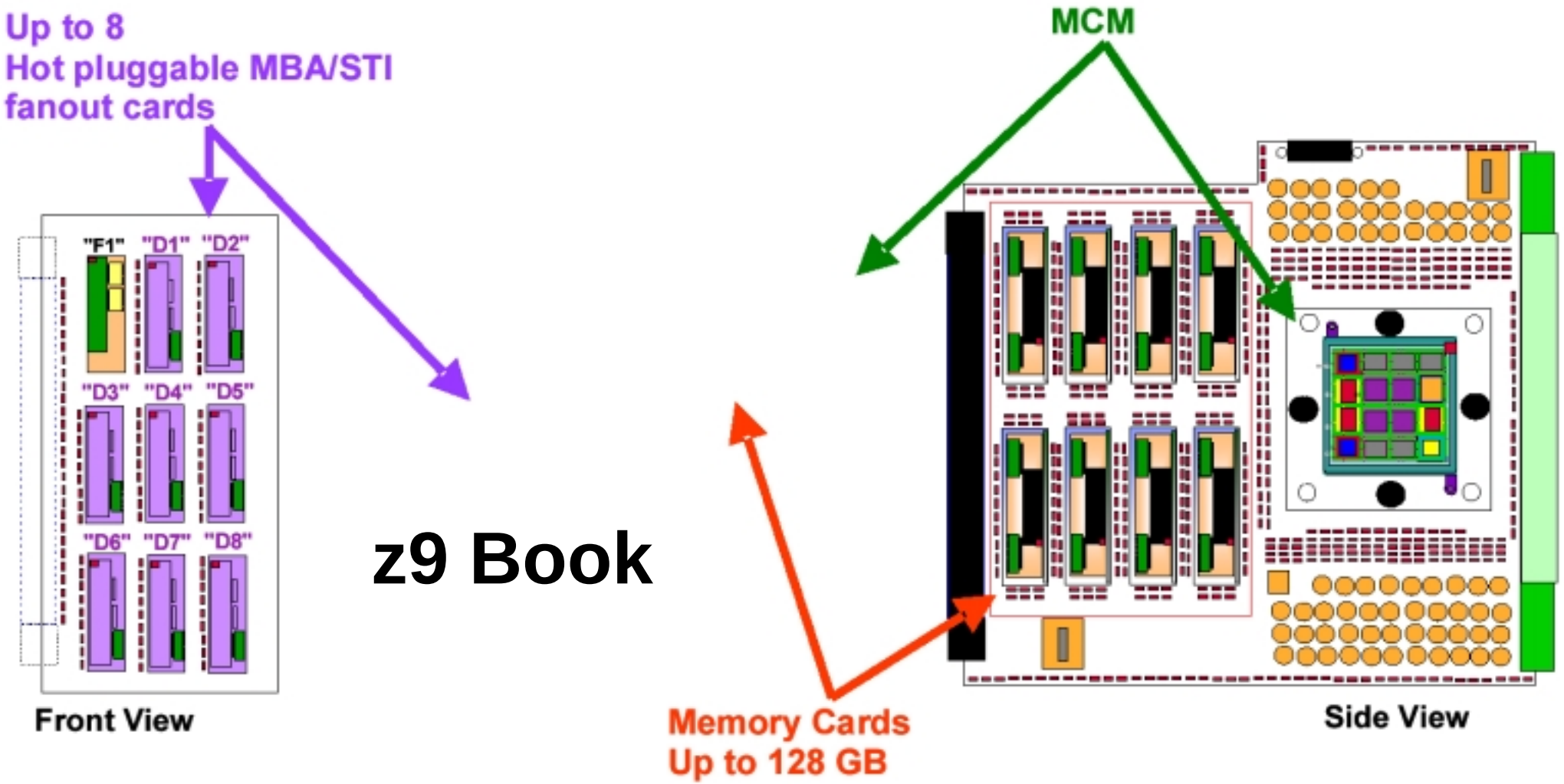
z9-109 'Central Electronic Complex': PUs, Speicher, I/O-Anschlüsse...

- Nimmt bis zu vier „Processor-Books“ auf

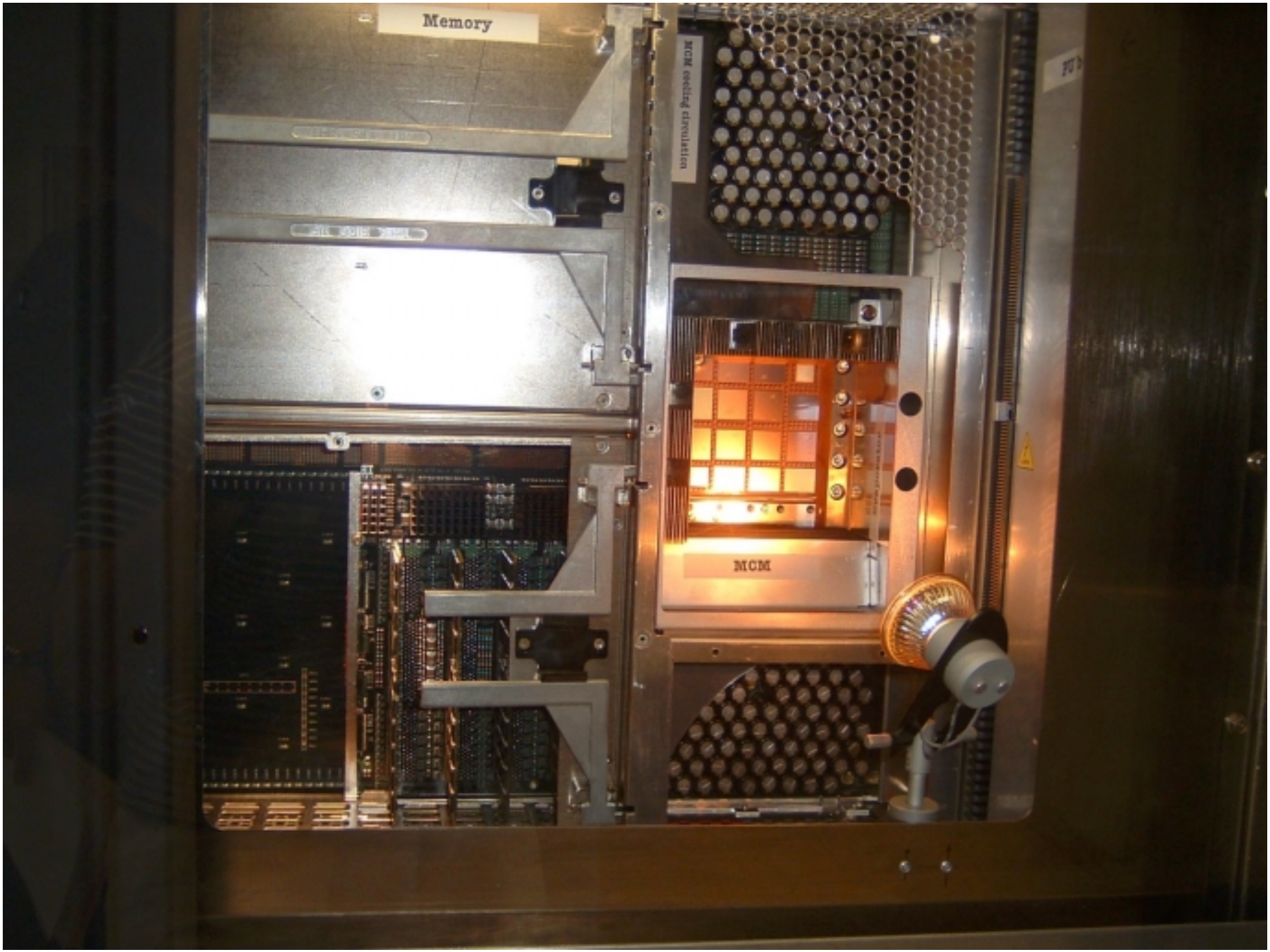


z9-109 Prozessor Book Layout

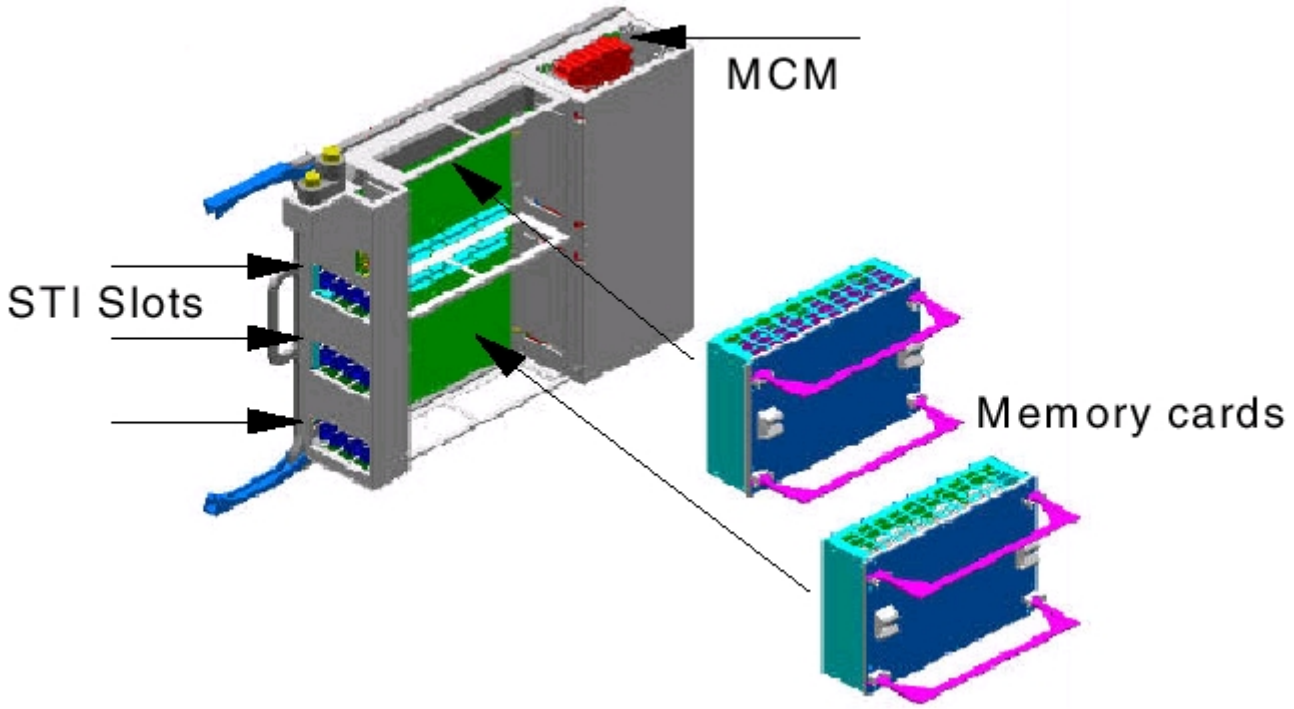


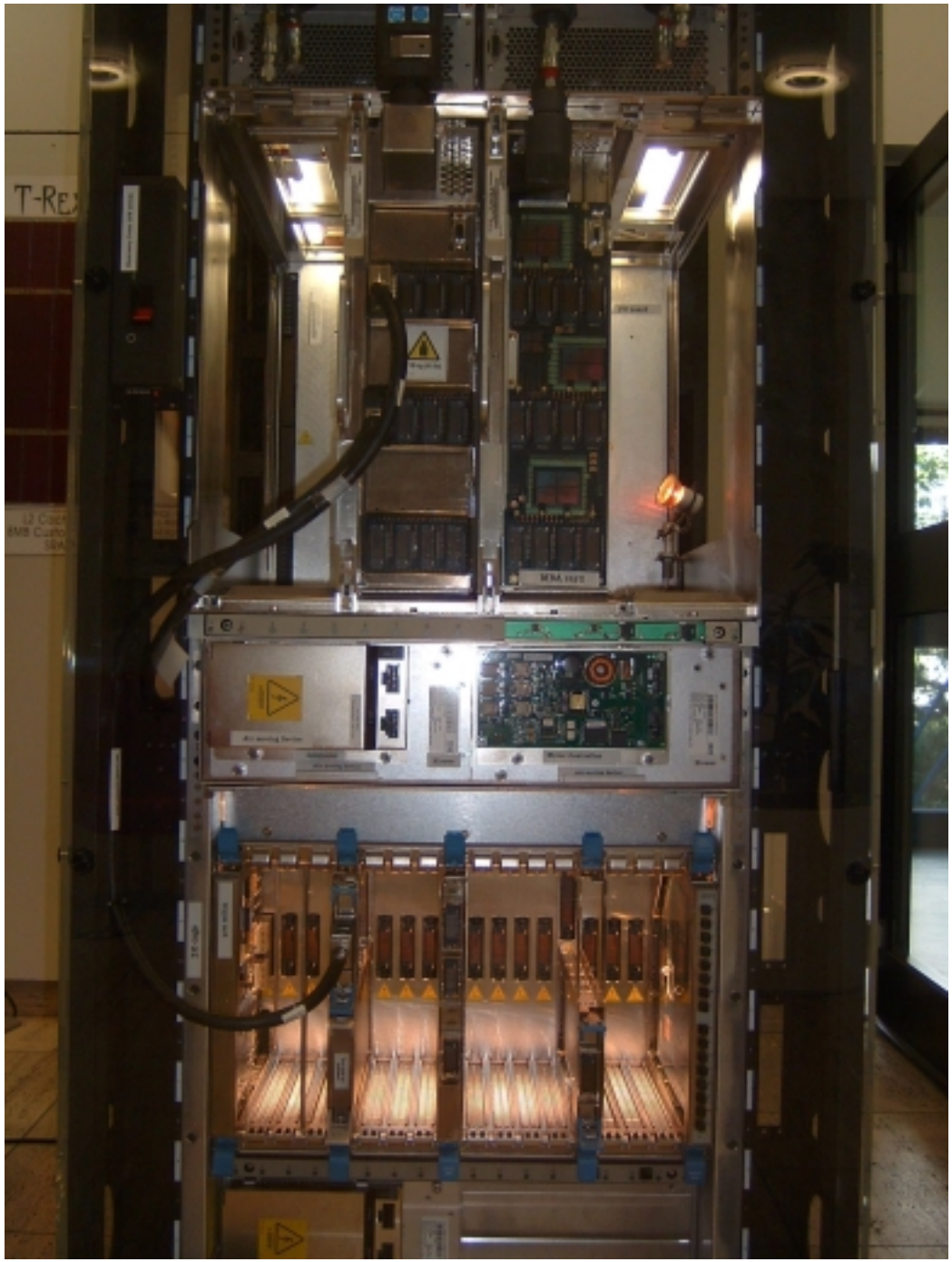


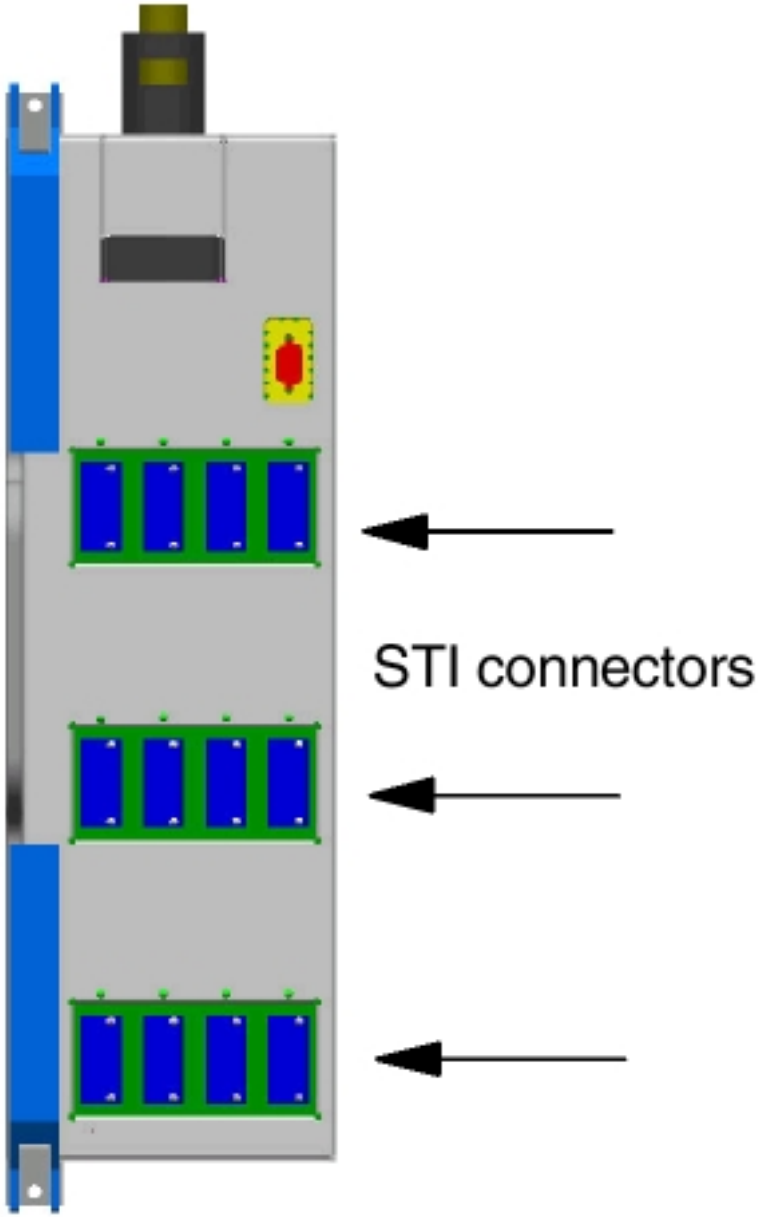
Connectivity from the CPU/memory area to the I/O area is established via eight MBA/STI hot pluggable cards at the front of the book. Each of the eight MBA/STI cards provides two 2.7 GByte/sec bidirectional interfaces to the I/O cage realized by the big black cables, i.e.,the aggregate external bandwidth is more than 43 GB/sec.


Memory
MCM cooling circulation
MCM

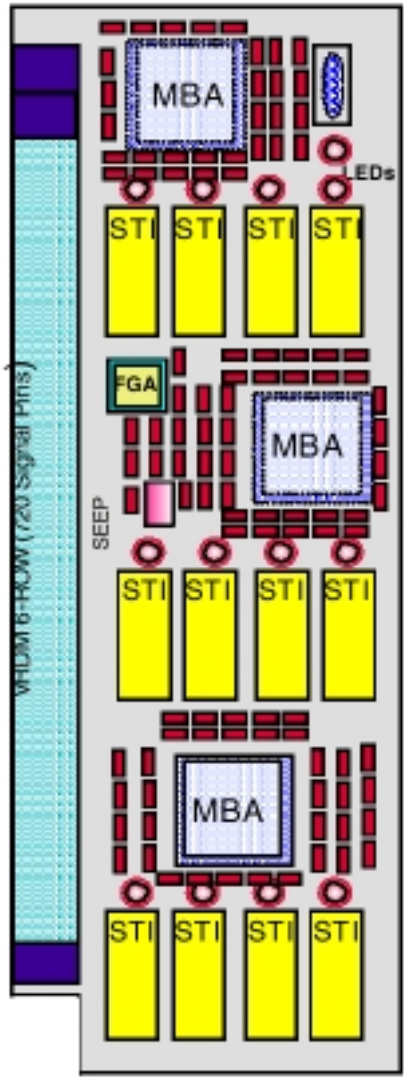
MCM


MCM
STI Slots
Memory cards

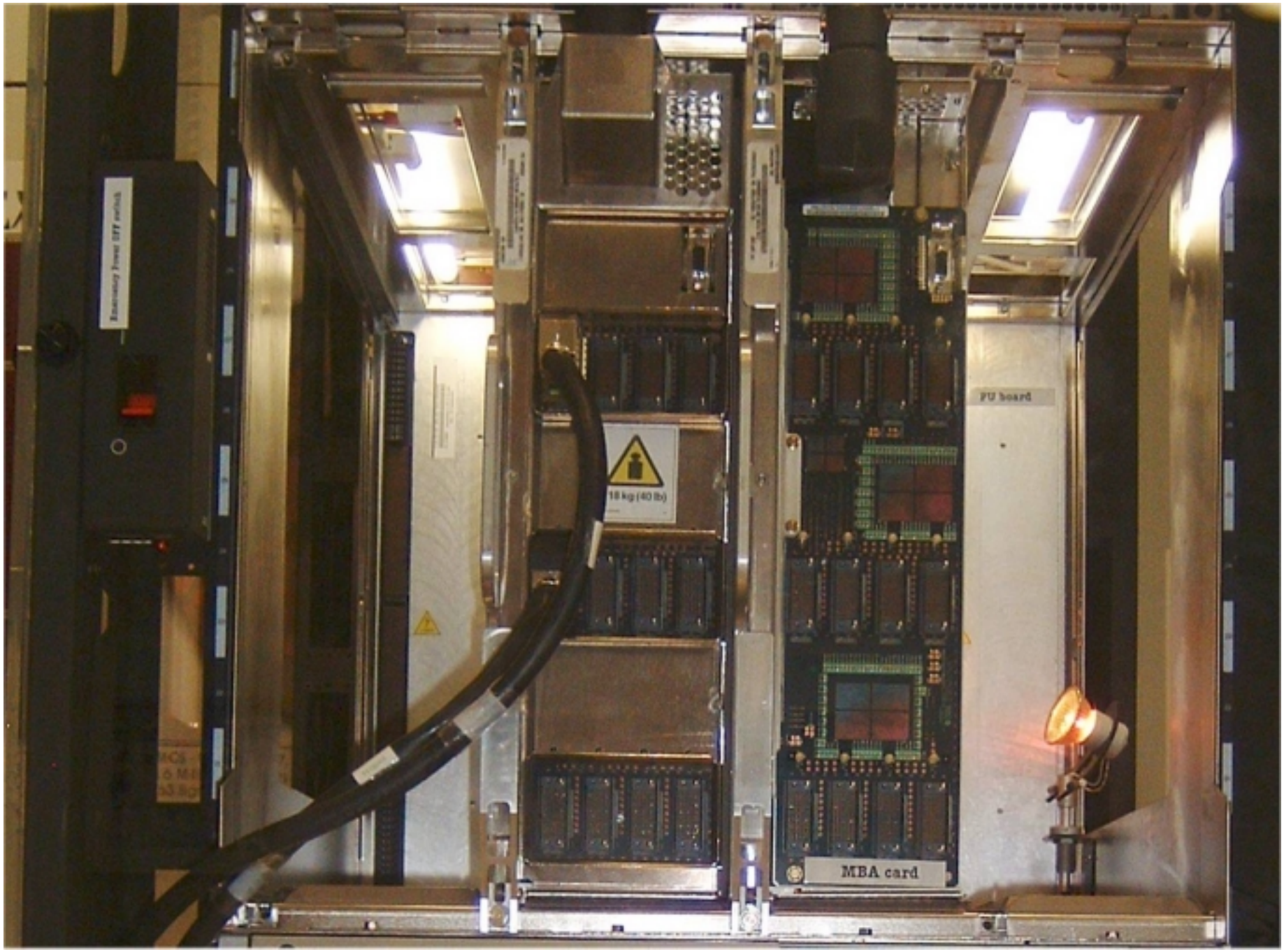


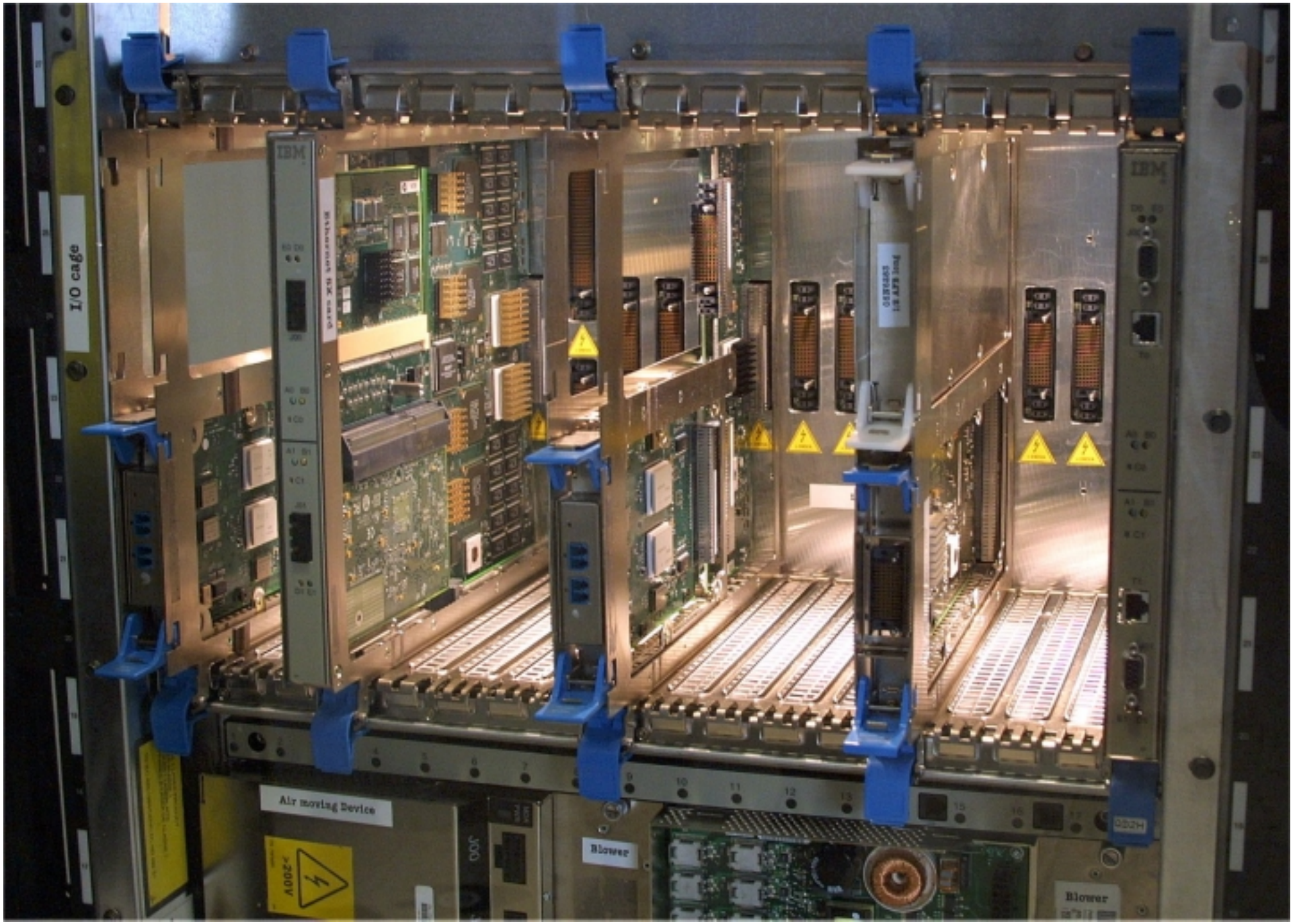


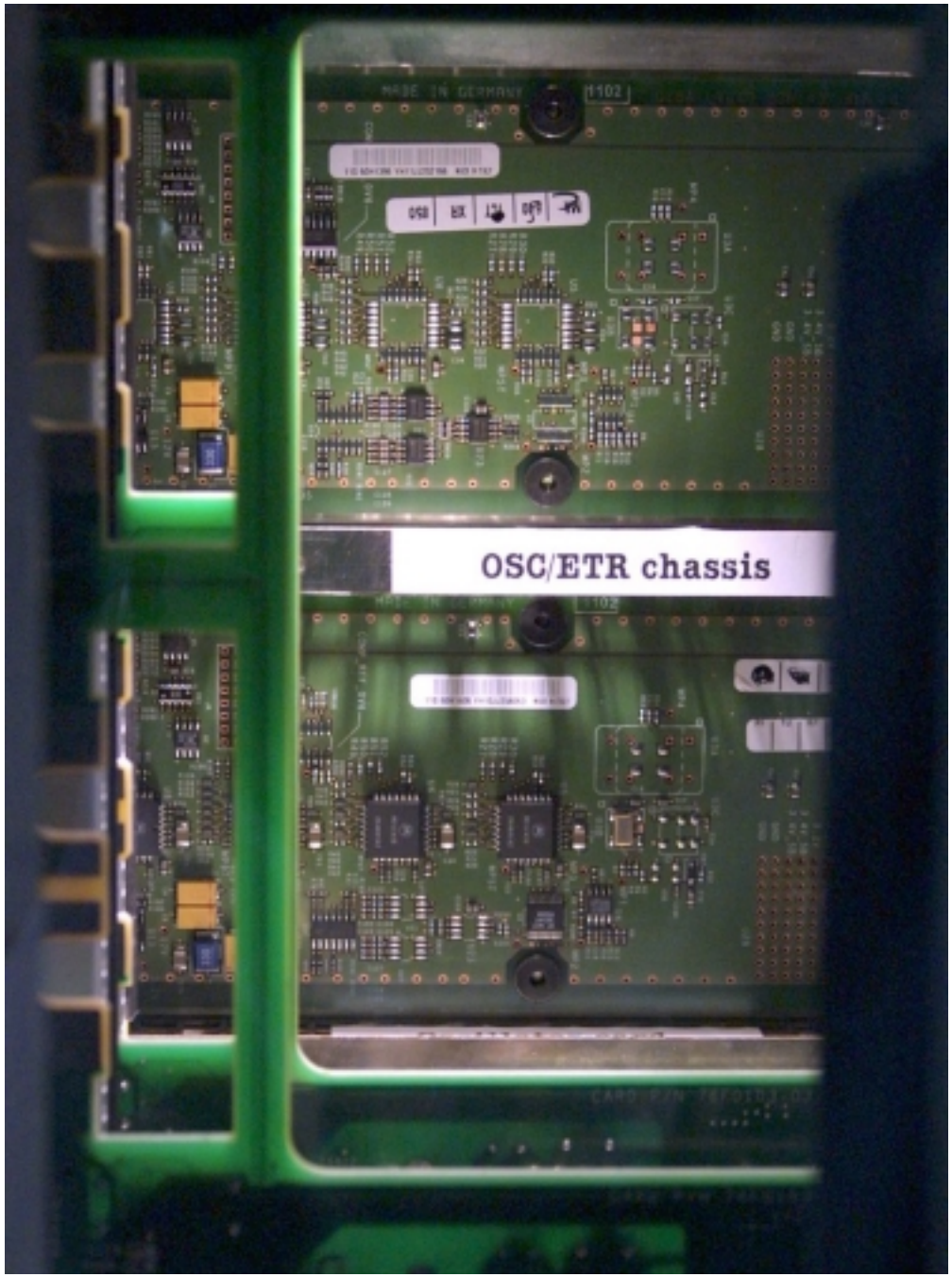
Book front view

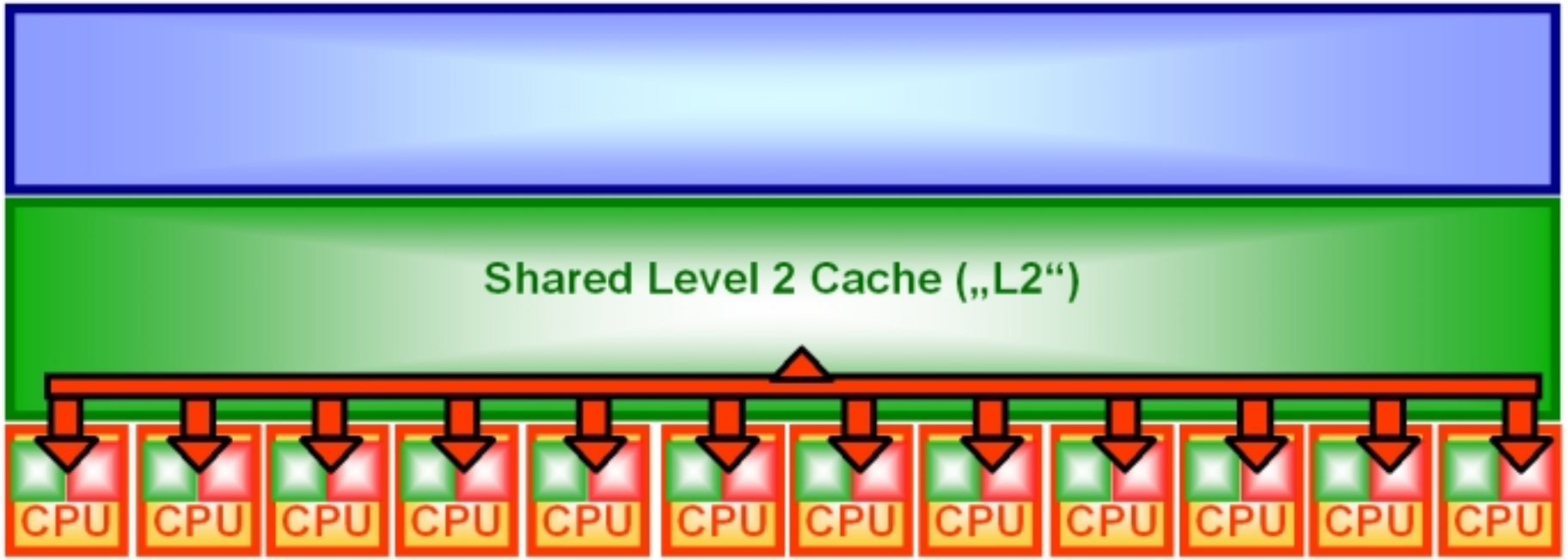


MBA card


PU board
MBA card

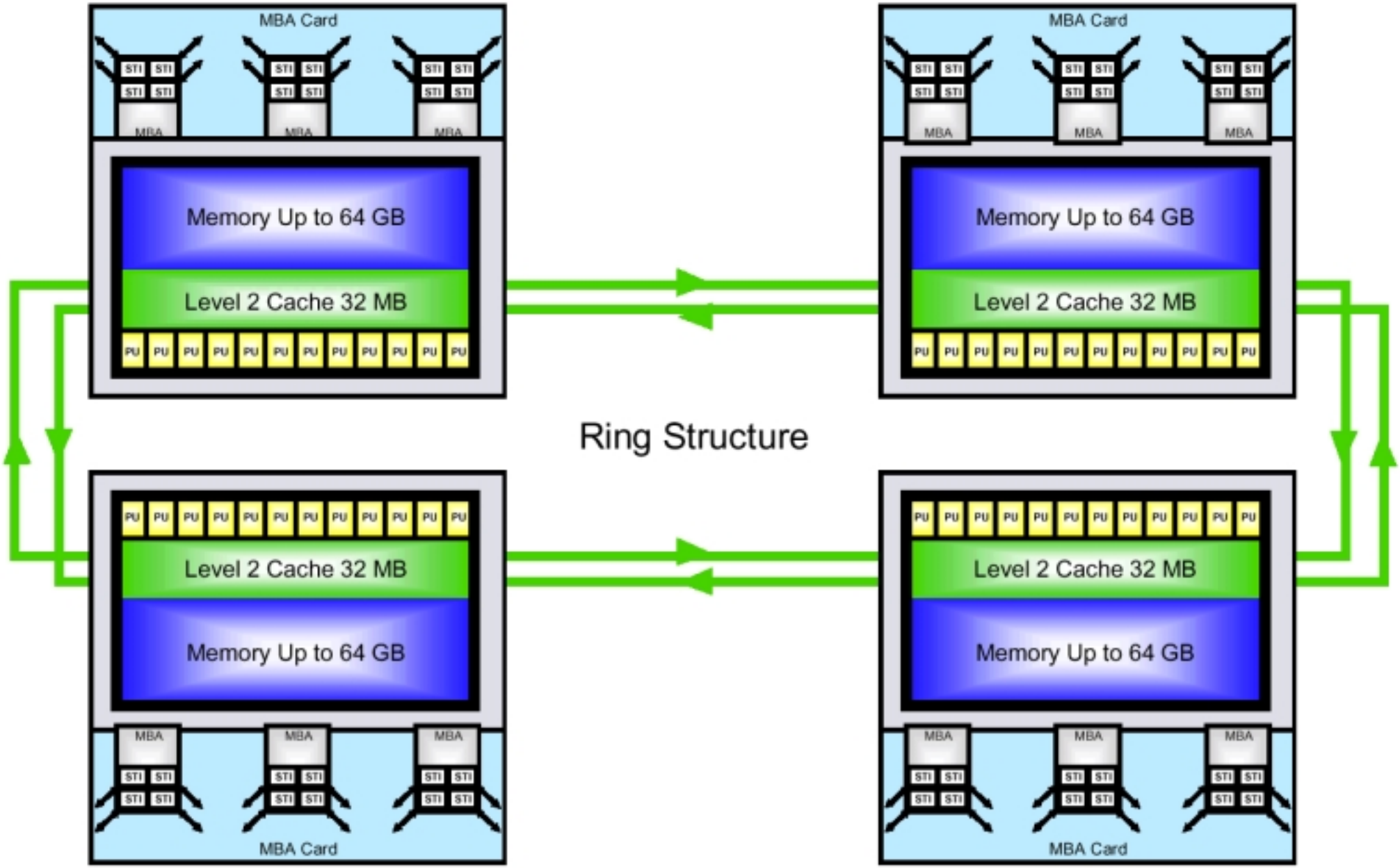

I/O cage
IBM
Ethernet SX card
Air moving Device
Blower
Blower

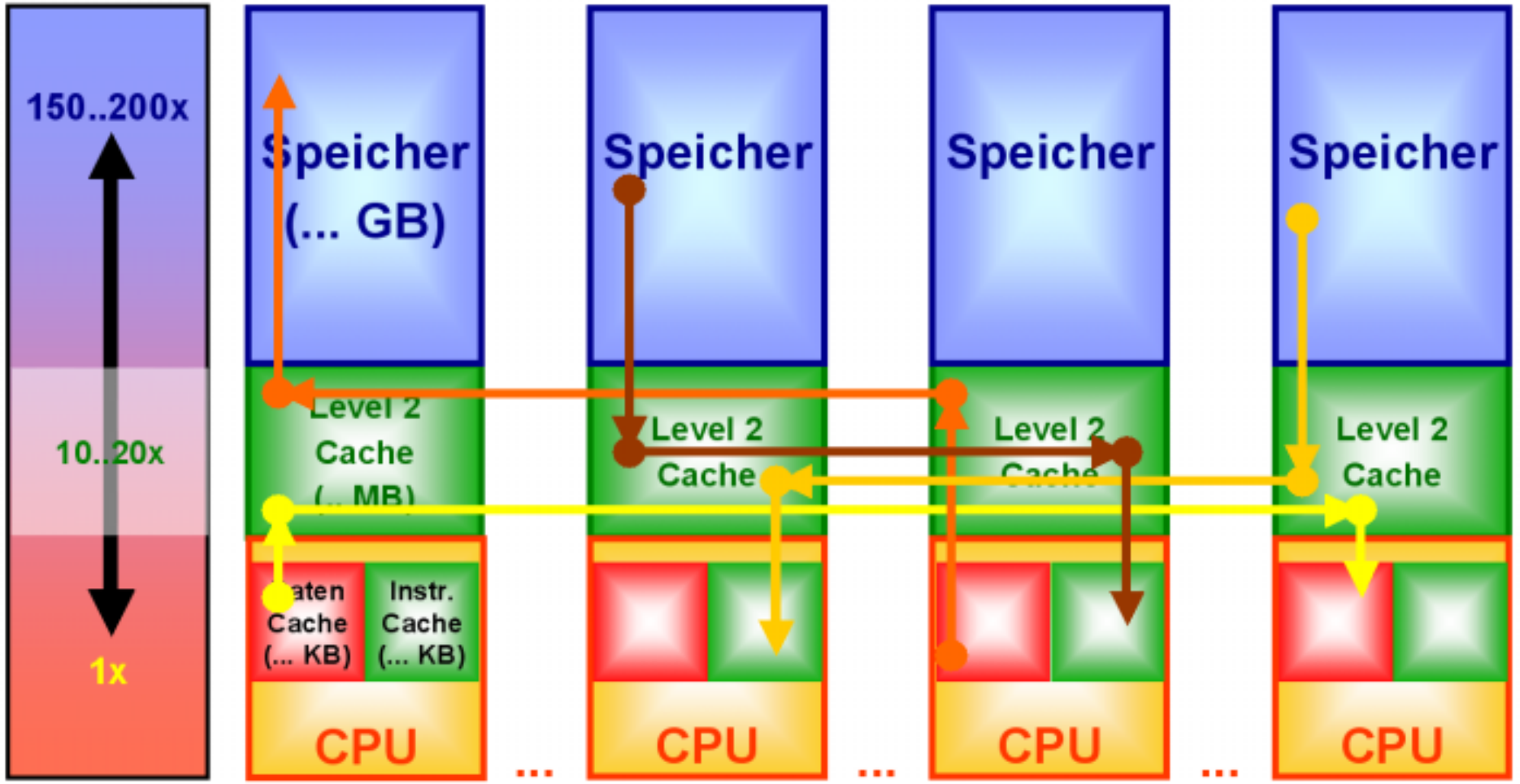

OSC/ETR chassis



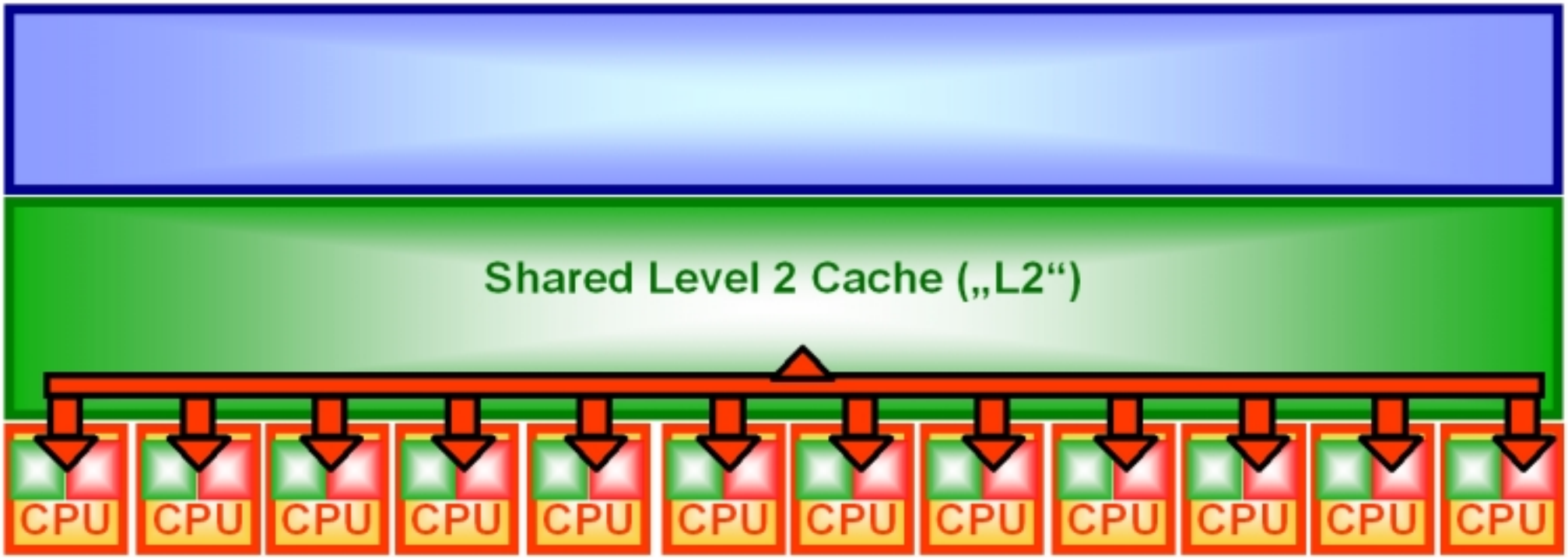
zSeries - Zentraler L2 Switch mit gleichzeitigem Zugriff aller Prozessoren
CICS, DB2 und IMS Sperrverwaltung, Signifikante Leistungsverbesserung

z990 Model D32 (GA2)



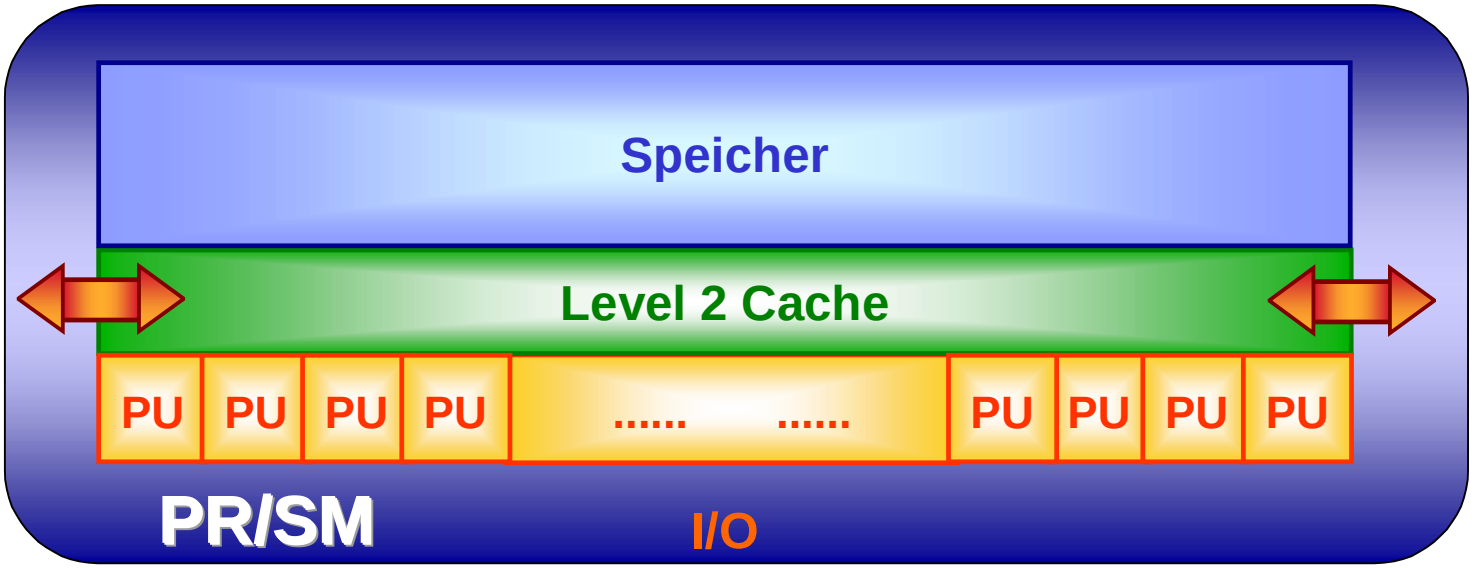


Sun und HP Cluster – Kommunikation zwischen den Prozessoren

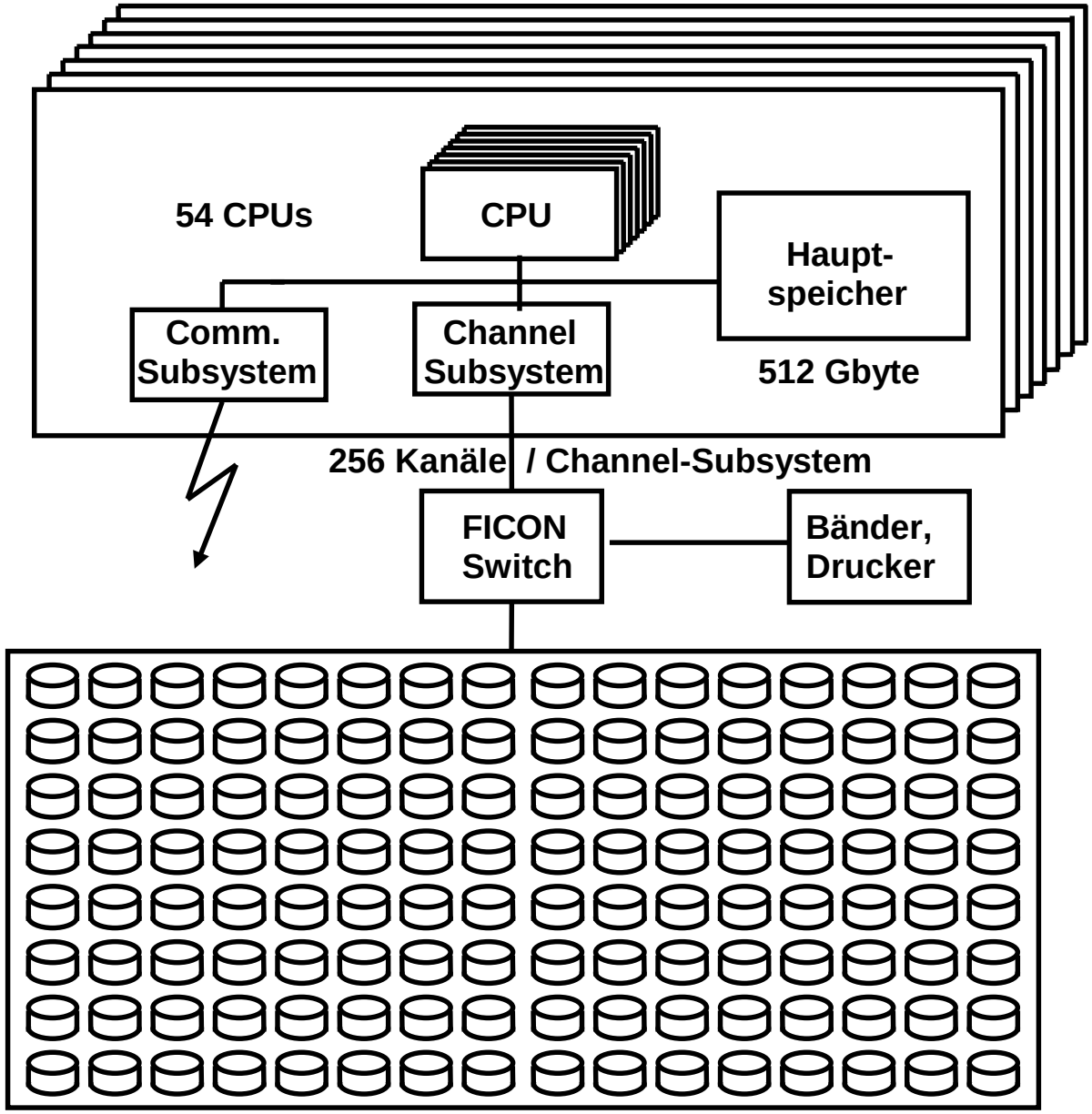


zSeries - Zentraler Switch mit gleichzeitigem Zugriff aller Prozessoren

zSeries und System z9: Multi-Book-Strukturen (Logische Sicht)



- Betriebsart: „Logical Partition Mode“ (Processor Resource/Systems Manager)
- Ein Pool von physischen Ressourcen (CPU's, Speicher, I/O) in modularer Implementierung (1/2/3/4 Knoten/'books')
- Nutzung durch virtuelle Server: Bis zu 60 LPARs ...100+... (VM)
- Multiple Kanal Subsysteme: Bis zu 4 x 256 „E/A-Kanäle“ (Channel Path IDs)

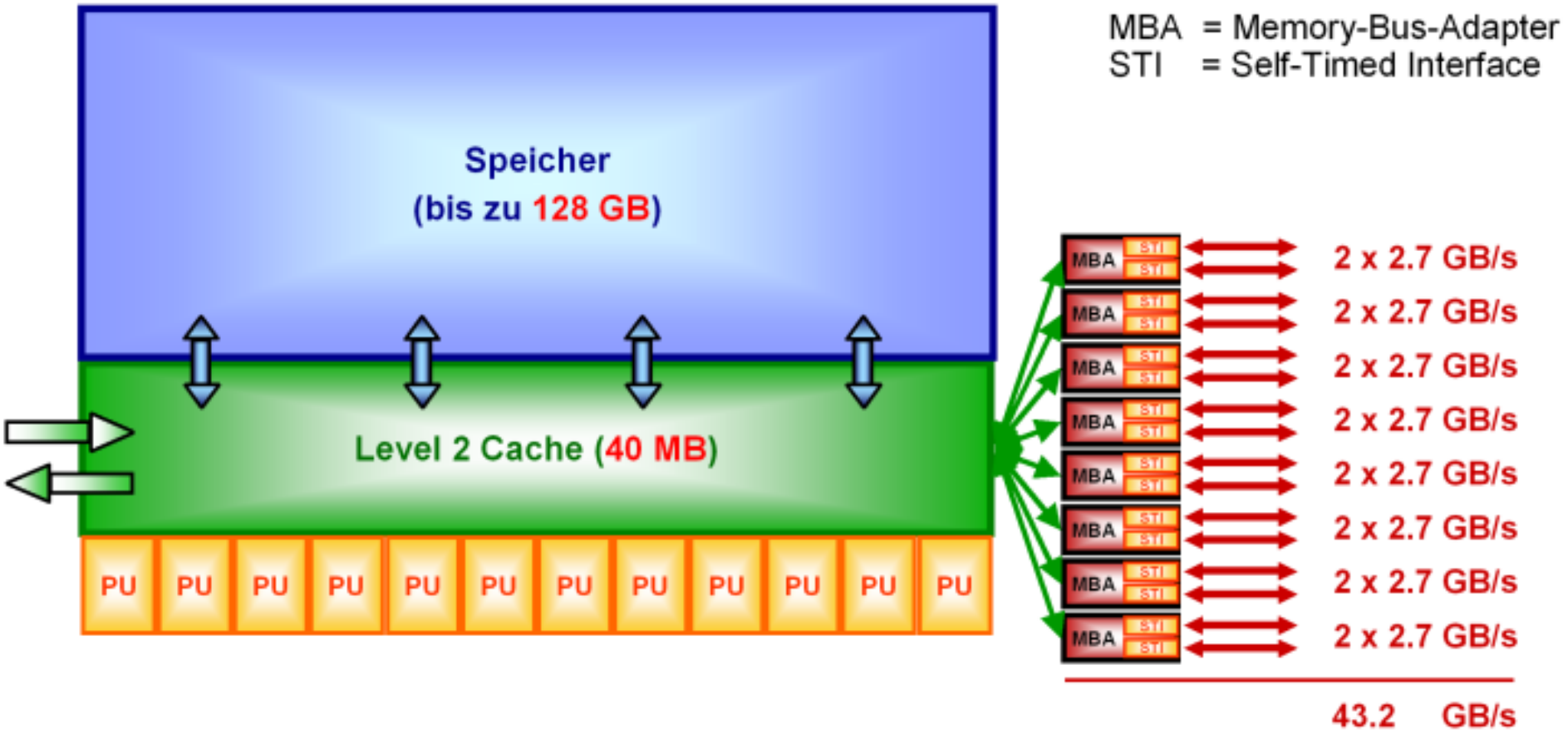


zSeries (S/390) Großsystem-konfiguration

128 000 Plattenspeicher (Devices)

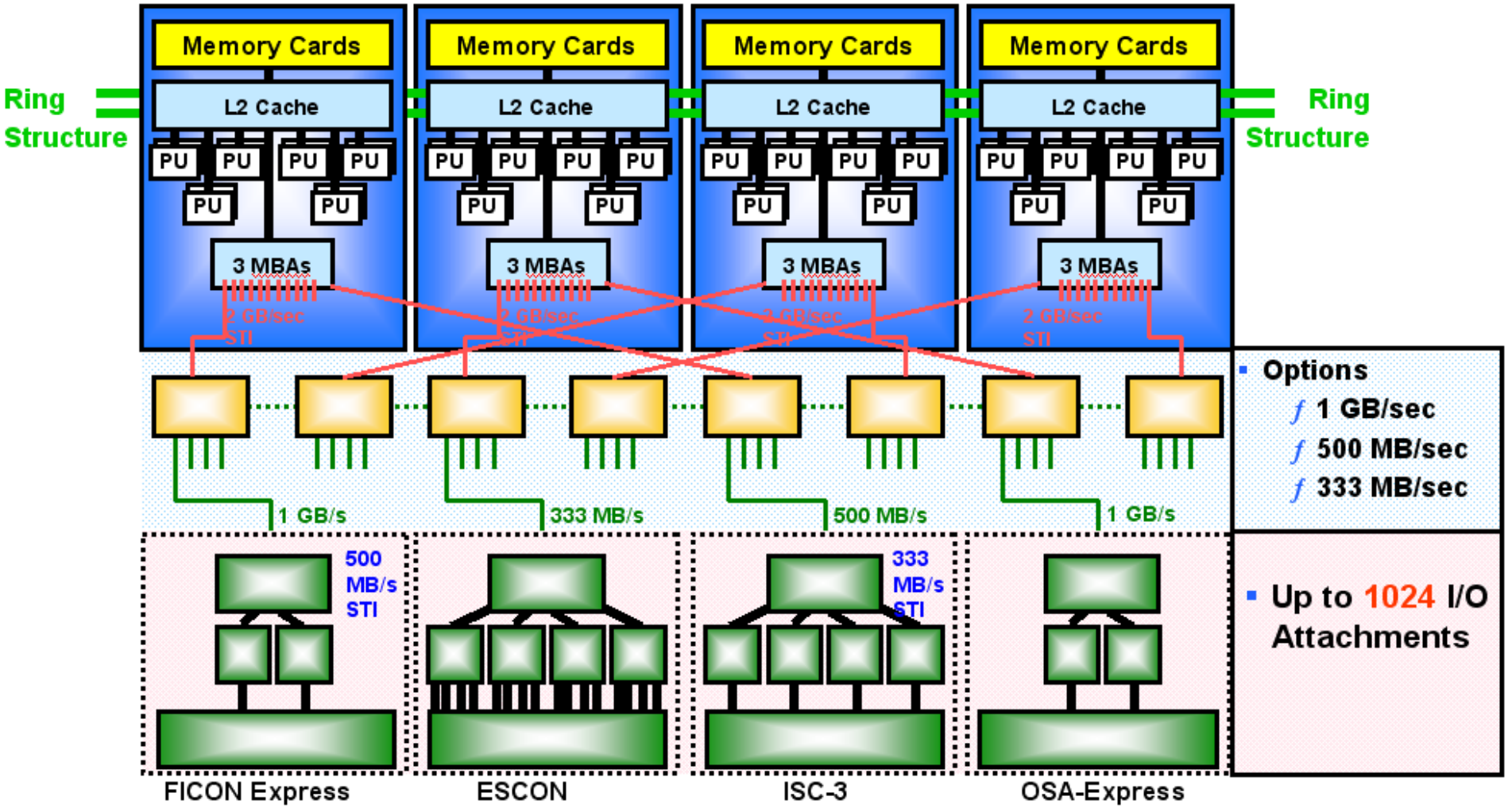
Logische Plattenspeicher (logical Volumes)

10 - 1000 TeraByte Plattenspeicherplatz



Gesamte externe z9 Sysembandbreite: 4 x 43,2 GByte/s.

Im Gegensatz zu allen anderen Architekturen kommunizieren die E/A Geräte mit dem L2 Cache und nicht mit dem Hauptspeicher. Die zSeries Ingenieure waren in der Lage, die resultierenden Cache Koheränz Probleme zu lösen.


Memory Cards
Memory Cards
Memory Cards
Memory Cards
Ring Structure
Ring Structure
L2 Cache
L2 Cache
L2 Cache
L2 Cache
PU
3 MBAs
2 GB/sec STI
1 GB/s
333 MB/s
500 MB/s
1 GB/s
500 MB/s STI
333 MB/s STI
FICON Express
ESCON
ISC-3
OSA-Express
Options
1 GB/sec
500 MB/sec
333 MB/sec
Up to 1024 I/O Attachments

Hercules

Hercules ist ein open Source Software Paket. Es emuliert die System/390 und zSeries Architektur.

Hercules läuft auf einem regulären, leistunggsfähigen PC unter Linux, Windows 98, Windows NT und Windows 2000.

Einzelheiten:

http://www.conmicro.cx/hercules/

http://neuro2.med.uni-magdeburg.de/~markgraf/mvs/doc/hercfaq.html