

IBM System z10 for Hochschule Esslingen January 9, 2009

by

Klaus-Dieter Mueller

IBM Executive Briefing Center Europe

kmuller@de.ibm.com



Smart. Cool. Affordable.

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

AlphaBlox*	GDPS*	RACF*	Tivoli*
APPN*	HiperSockets	Redbooks*	Tivoli Storage Manager
CICS*	HyperSwap	Resource Link	TotalStorage*
CICS/VSE*	IBM*	RETAIIN*	VSE/ESA
Cool Blue	IBM eServer	REXX	VTAM*
DB2*	IBM logo*	RMF	WebSphere*
DFSMS	IMS	S/390*	xSeries*
DFSMShsm	Language Environment*	Scalable Architecture for Financial Reporting	z9*
DFSMSrmm	Lotus*	Sysplex Timer*	z10
DirMaint	Multiprise*	Systems Director Active Energy Manager	z10 BC
DRDA*	MVS	System/370	z10 EC
DS6000	OMEGAMON*	System p*	z/Architecture*
DS8000	Parallel Sysplex*	System Storage	z/OS*
ECKD	Performance Toolkit for VM	System x*	z/VM*
ESCON*	PowerPC*	System z	z/VSE
FICON*	PR/SM	System z9*	zSeries*
FlashCopy*	Processor Resource/Systems Manager	System z10	

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

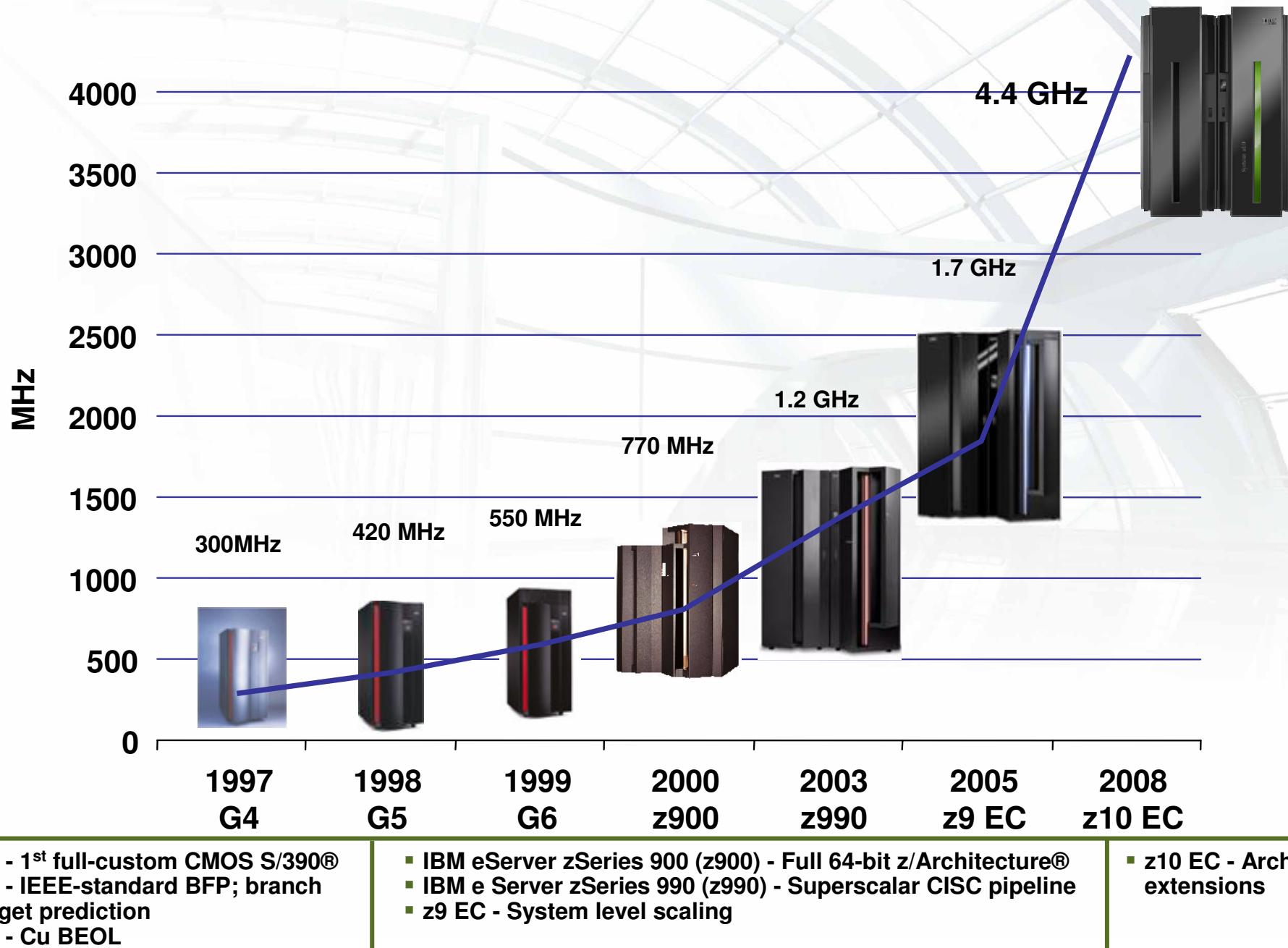
Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

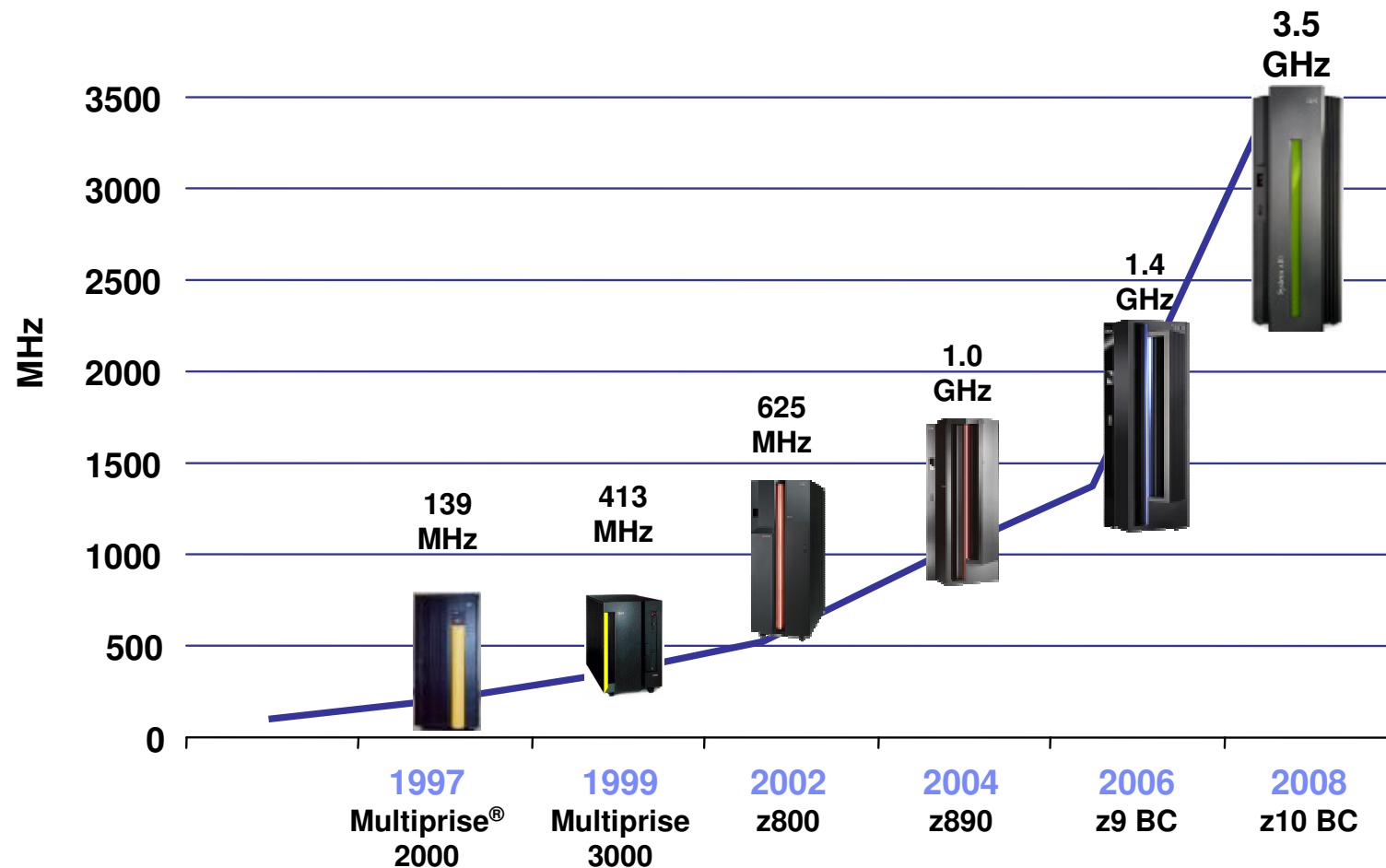
Von 'System/360' (S/360) zu ESA/390 und z-/Architektur

- 1964 S/360
 - CISC, 24bit Adressierung, 'Real Storage', Uniprozessoren
 - Amdahl, G.M., Blaauw, G.A., and Brooks, F.P.: *Architecture of the IBM System/360*
 - IBM Journal of Research and Development 8,2 (April 1964)
- 1971 S/370
 - 'Virtual Storage', Multiprozessor-Unterstützung, ...
- 1981 S/370 XA (Extended Architecture)
 - 31bit Adressierung (2GB), 'Expanded Storage' (>2GB), 'Channel Subsystem'
- 1988 ESA/370
 - ESA = Enterprise Systems Architecture, Logische Partitionierung
 - Ausbau der Speicher-Zugriffsmethoden: Mehr als ein 'address space'
- 1990 ESA/390
 - 'ESCON' (Enterprise Systems Connection Architecture) Glasfasertechnologie ...
 - Datenkompression, Kryptographie, LPAR Erweiterungen
- 1994 Parallel Sysplex, Übergang von Bipolar zu CMOS Technologie
 - 'Coupling Facility', Cluster von bis zu 32 x 16-way MultiProzessoren
 - 'FICON' (Fiber Channel Connectivity), Ausbau der Glasfasertechnologie
- 2000 z-/Architektur (64-bit)
 - Hardware 2000-2008: zSeries z900/z800 & z990/z890, System z z9EC/BC, z10EC/BC, ...

IBM z10 EC Continues the CMOS Mainframe Heritage



IBM z10 BC continues the CMOS Mainframe heritage



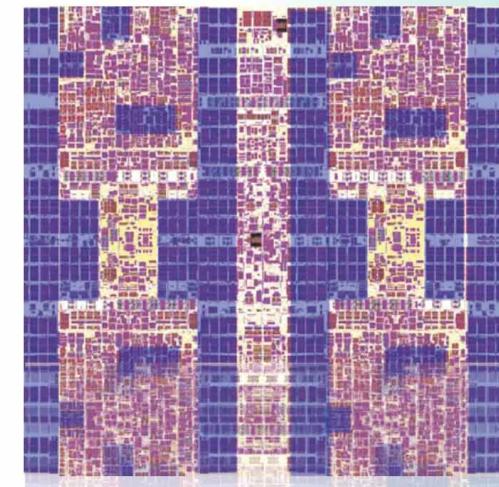
- Multiprise 2000 - 1st full-custom Mid-range CMOS S/390
- Multiprise 3000 – Internal disk, IFL introduced on midrange

- IBM eServer zSeries 800 (z800) - Full 64-bit z/Architecture®
- IBM eServer zSeries 890 (z890) - Superscalar CISC pipeline
- z9 BC - System level scaling

- z10 BC - Architectural extensions
- Higher frequency CPU

Making high performance a reality

- **New Enterprise Quad Core z10 EC processor chip**
 - 4.4 GHz - additional throughput means improved price/performance
 - Cache rich environment optimized for data serving
 - 50+ instructions added to improve compiled code efficiency
 - Support for 1MB page frames
- **Hardware accelerators on the chip**
 - Hardware data compression
 - Cryptographic functions
 - Hardware Decimal Floating point
- **CPU intensive workloads get performance improvements from new core pipeline design**



**Enterprise Quad Core
z10 EC processor chip**

z Architecture Elements

Mainstorage

- Byte-wise addressable
- 64bit
- shared by all CPUs

I/O

- old: parallel (copperbased), 4.5MB/sec
- 1990': serial (fiber optics), 17MB/sec
- 1999: FiCON (fiber optics), 100+MB/sec
- 2006: FiCON Express 4, 500MB/sec

ESCON & FiCON 'Directors'

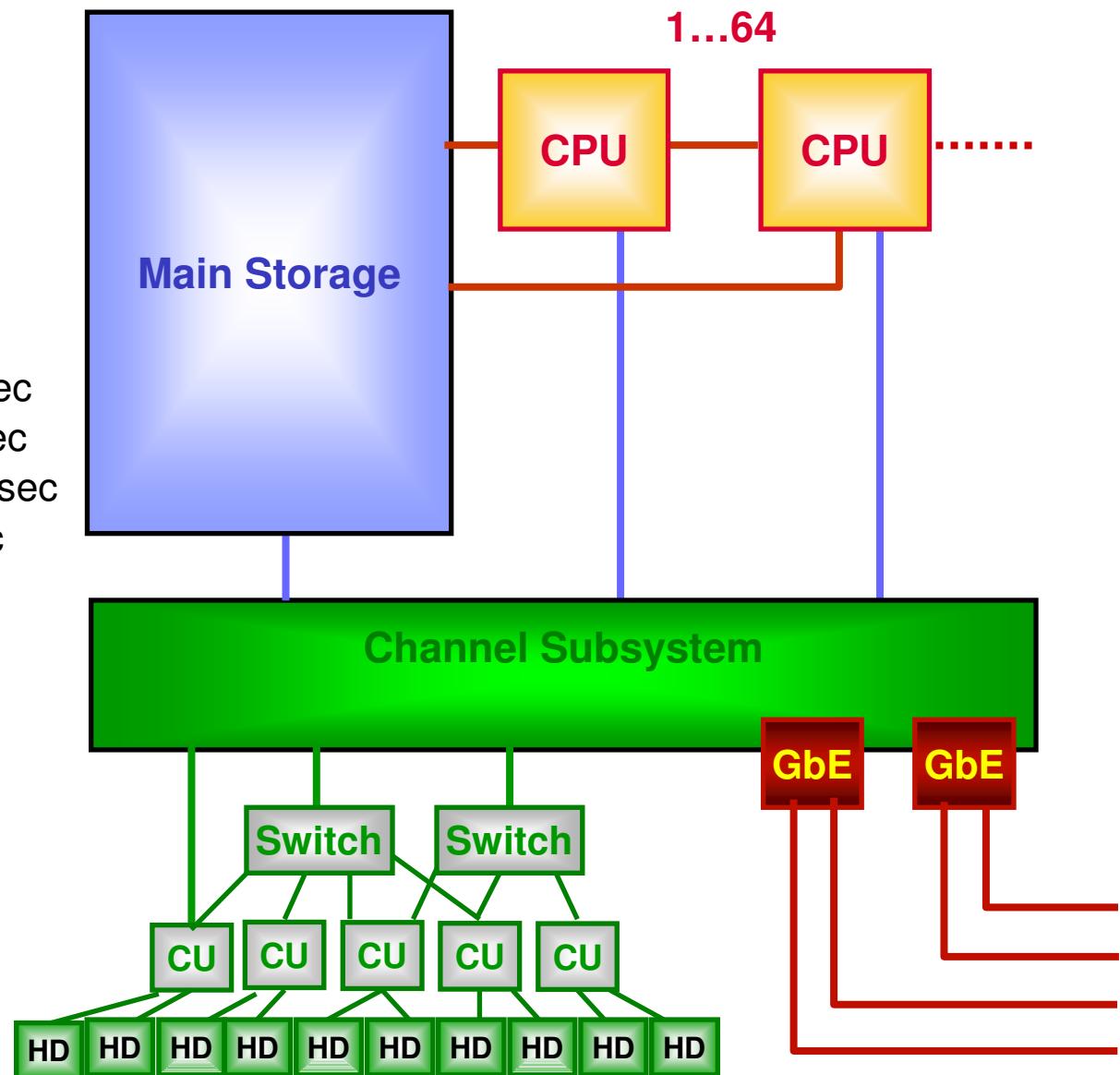
- 'switches'

Control Units (CU)

- Control units

Devices

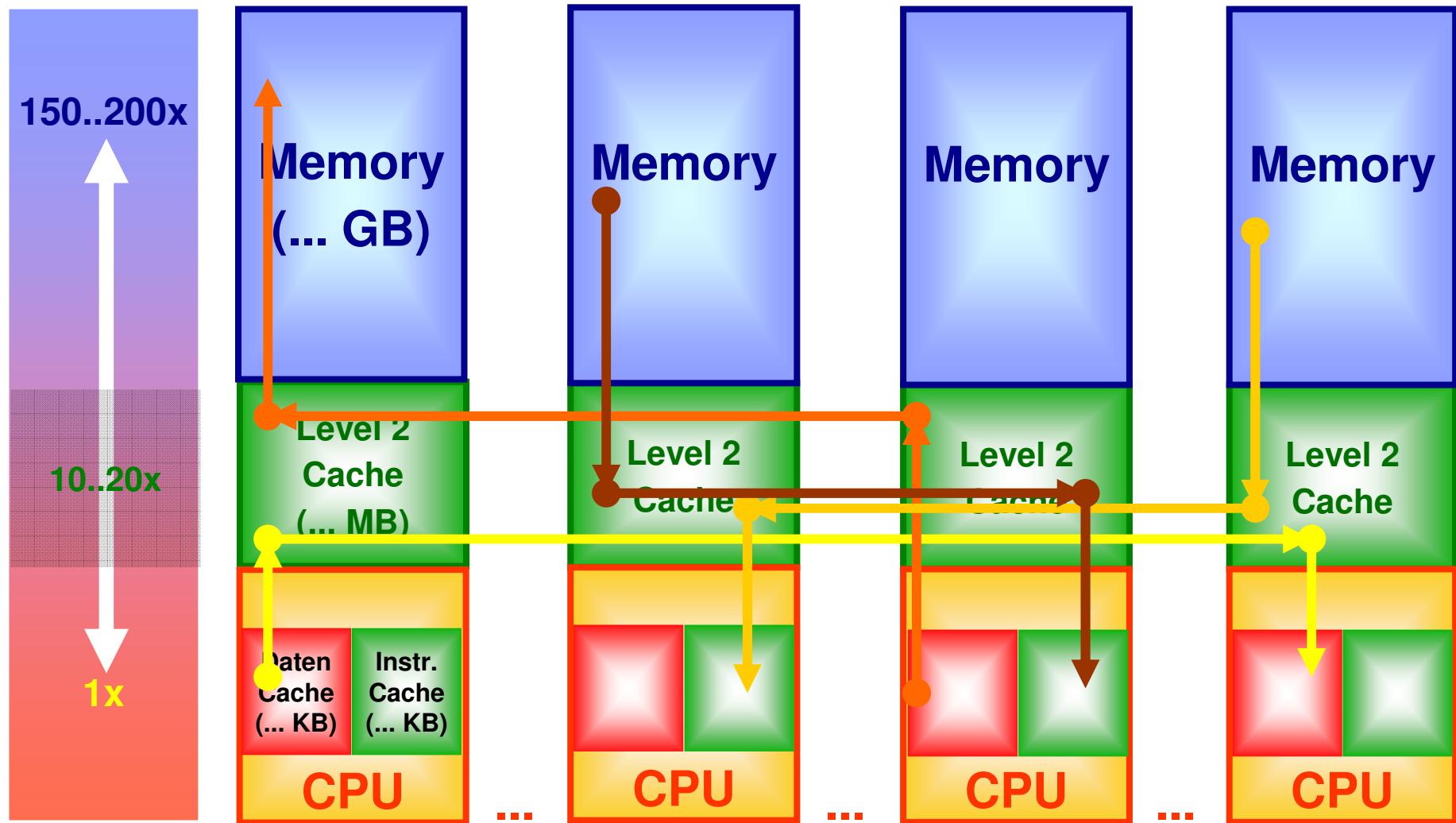
- Harddrives(HD), Tapes, Printer, ...



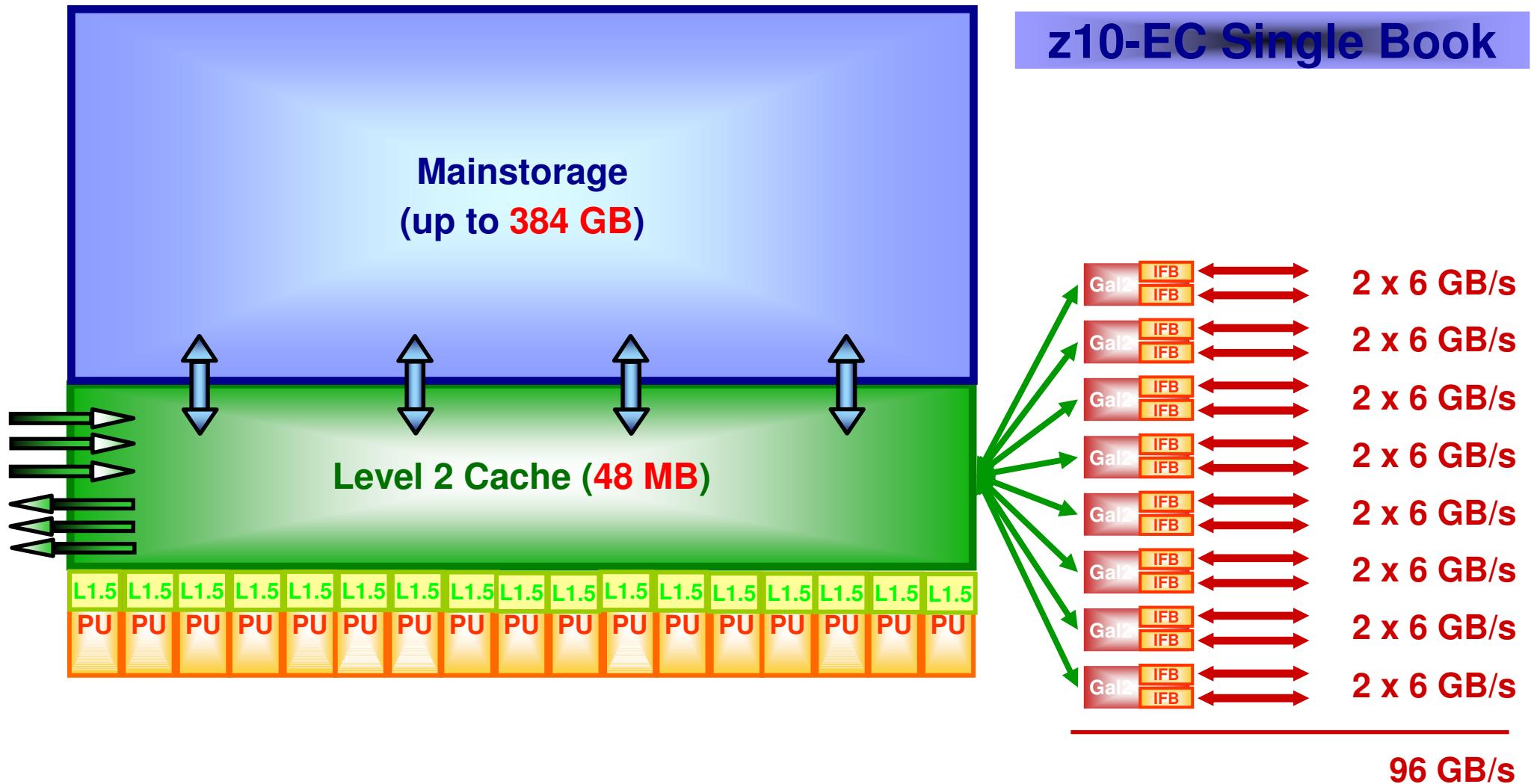
Network (GbE, 10GbE, ...)

Scalability: System-Structures optimized for data

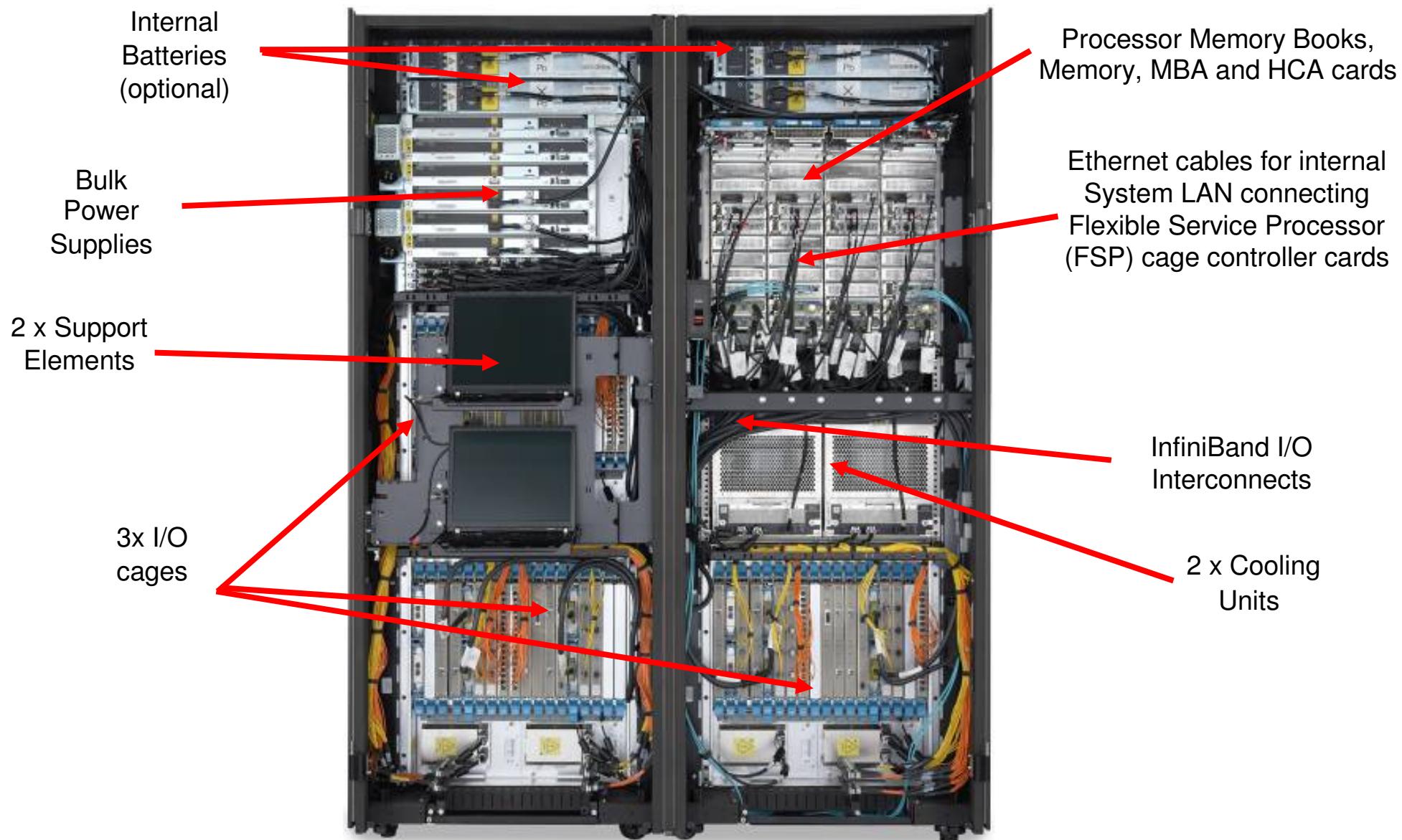
The key problem of current microprocessor-systems:
Memory access does not scale with CPU-cyclette !



z10-EC Systemstructure:

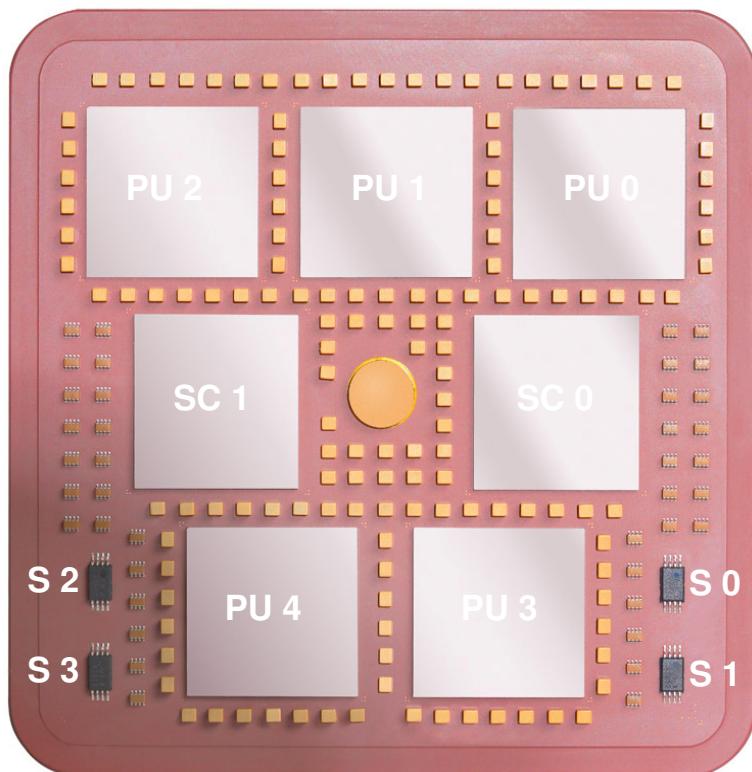


z10 EC – Under the covers (Model E56 or E64)

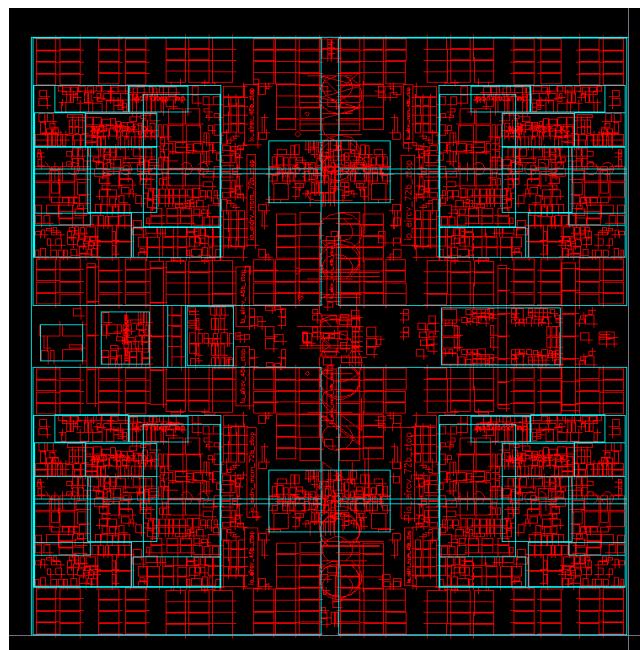
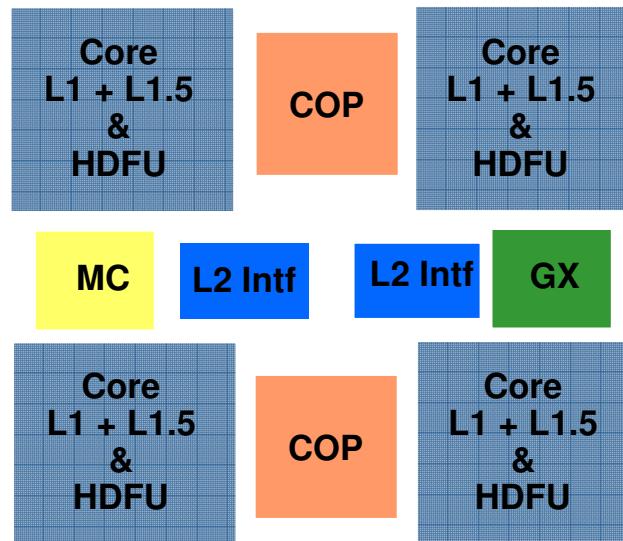


z10 EC Multi-Chip Module (MCM)

- 96mm x 96mm MCM
 - 103 Glass Ceramic layers
 - 7 chip sites
 - 7356 LGA connections
 - 17 and 20 way MCMs
- CMOS 11s chip Technology (65 nm)
 - 5 PU chips/MCM – Each up to 4 cores
 - One memory control (MC) per PU chip
 - 21.97 mm x 21.17 mm
 - 994 million transistors/PU chip
 - L1 cache/PU core
 - 64 KB I-cache
 - 128 KB D-cache
 - L1.5 cache/PU core
 - 3 MB
 - 4.4 GHz
 - 0.23 ns Cycle Time
 - 6 km of wire
 - 2 Storage Control (SC) chip
 - 21.11 mm x 21.71 mm
 - 1.6 billion transistors/chip
 - L2 Cache 24 MB per SC chip (48 MB/Book)
 - L2 access to/from other MCMs
 - 3 km of wire
 - 4 EEPROM (S) chips
 - 2 x active and 2 x redundant
 - Product data for MCM, chips and other engineering information
 - Clock Functions – distributed across PU and SC chips
 - Master Time-of-Day (TOD) and 9037 (ETR) functions are on the SC

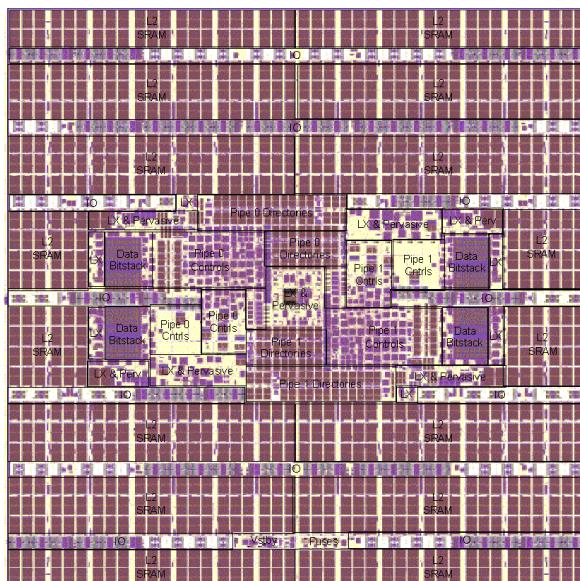
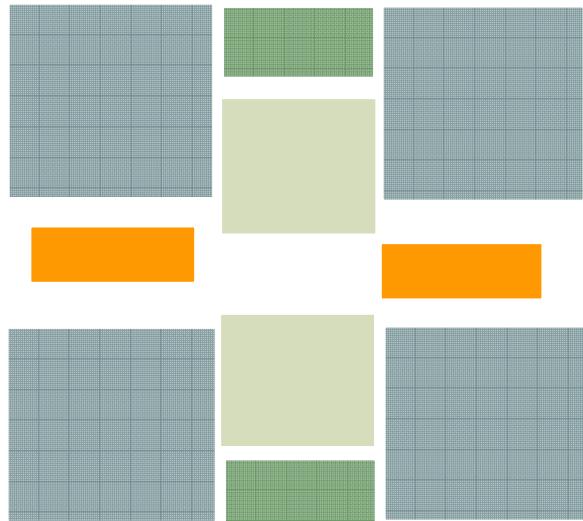


z10 EC – Enterprise Quad Core z10 PU Chip



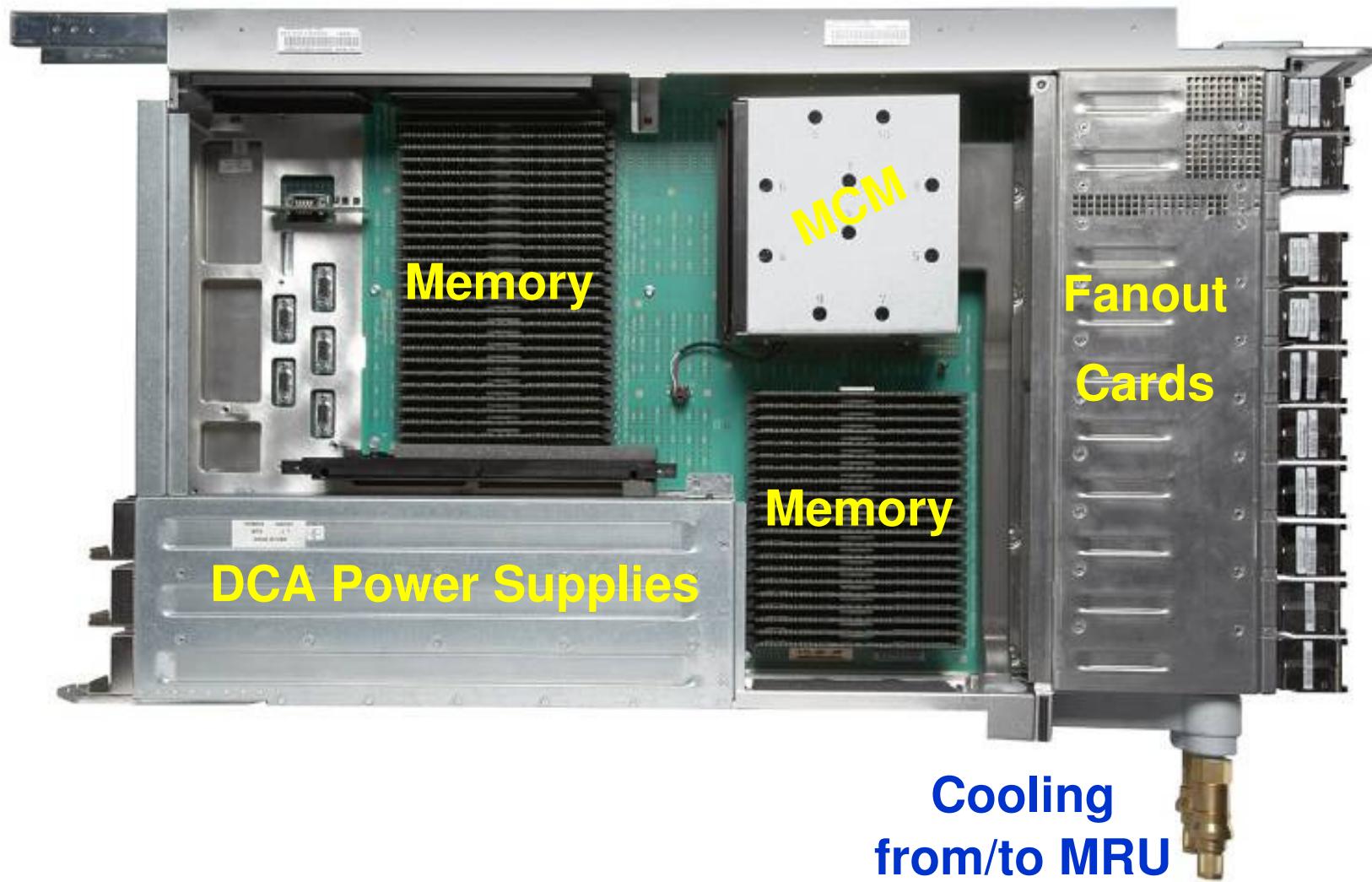
- Up to Four cores per PU
 - 4..4 GHz
 - L1 cache/PU core
 - 64 KB I-cache
 - 128 KB D-cache
 - 3MB L1.5 cache/PU core
 - Each core with its own Hardware Decimal Floating Point Unit (HDFU)
- Two Co-processors (COP)
 - Accelerator engines
 - Data compression
 - Cryptographic functions
 - Includes 16KB cache
 - Shared by two cores
- L2 Cache interface
 - Shared by all four cores
 - Even/odd line (256B) split
- I/O Bus Controller (GX)
 - Interface to Host Channel Adapter (HCA)
 - Compatible with System z9 MBA
- Memory Controller (MC)
 - Interface to controller on memory DIMMs

z10 EC SC Hub Chip

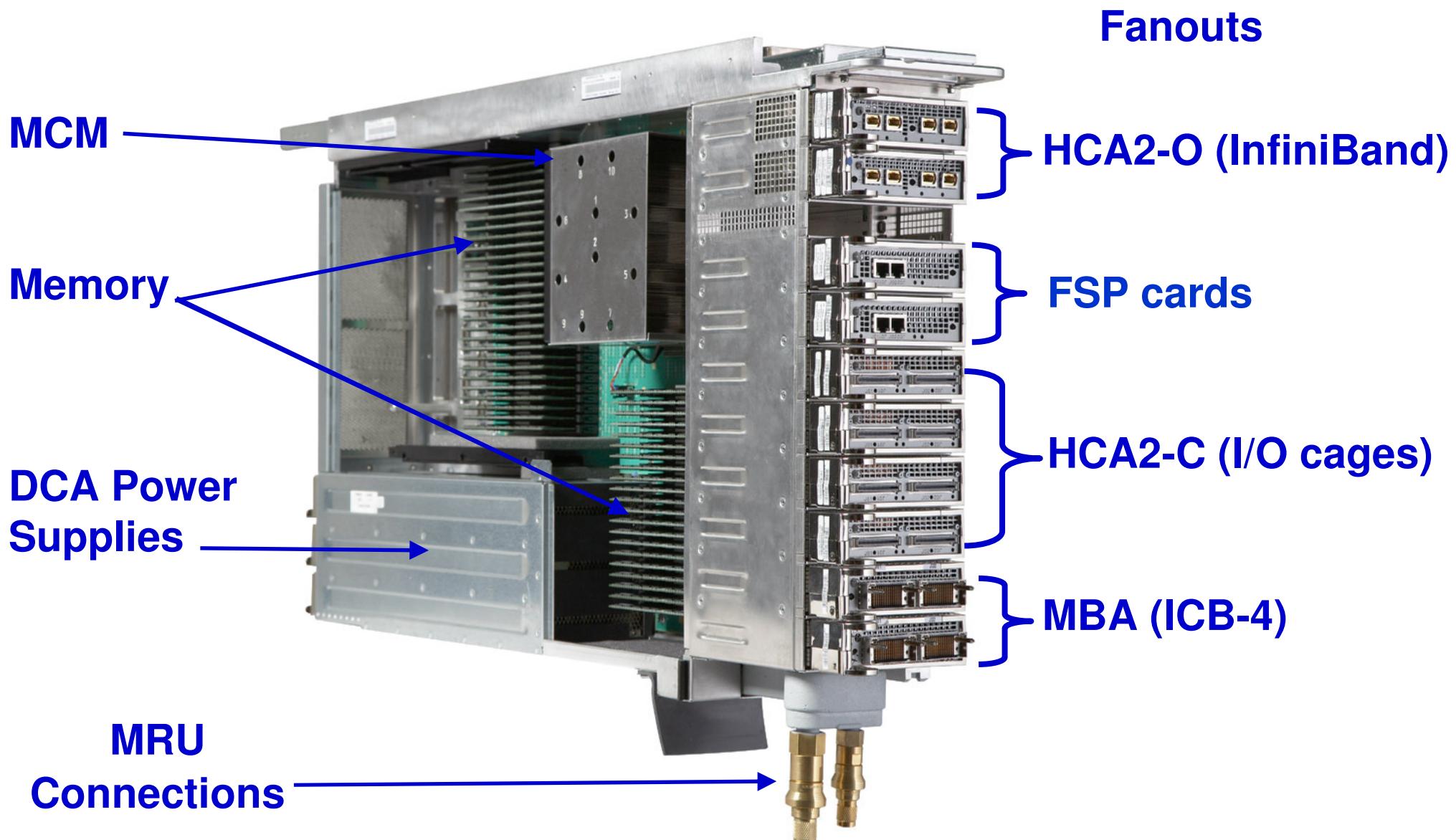


- Connects multiple z10 PU chips
 - 48 GB/Sec bandwidth per processor
- Shared Level 2 cache
 - 24MB SRAM Cache
 - Extended directory
 - Partial-inclusive discipline
 - Hub chips can be paired
 - 48MB shared cache
- Low-latency SMP coherence fabric
 - Robust SMP scaling
 - Strongly-ordered architecture
- Multiple hub chips/pairs allow further SMP scaling

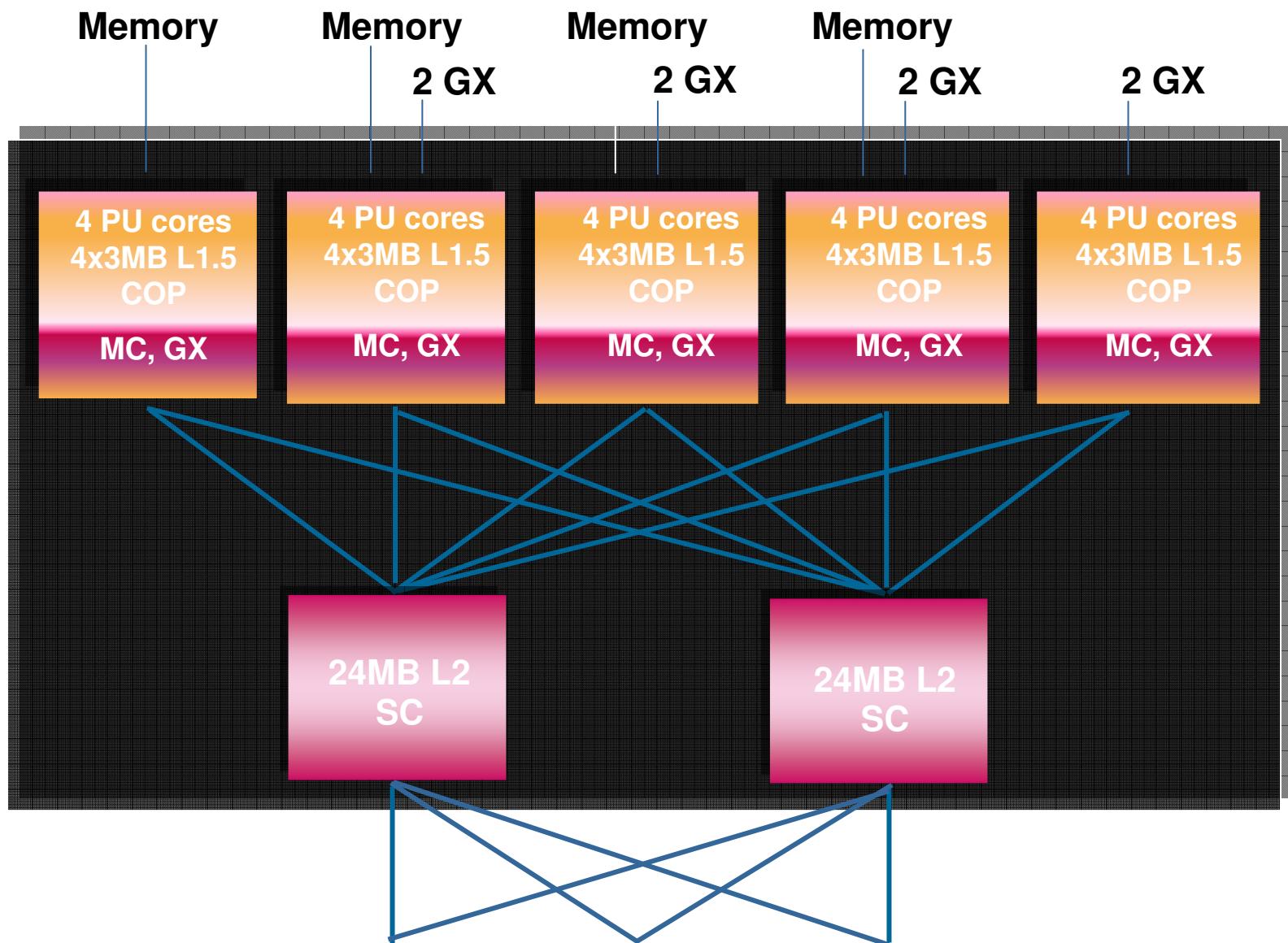
z10 EC Book Layout



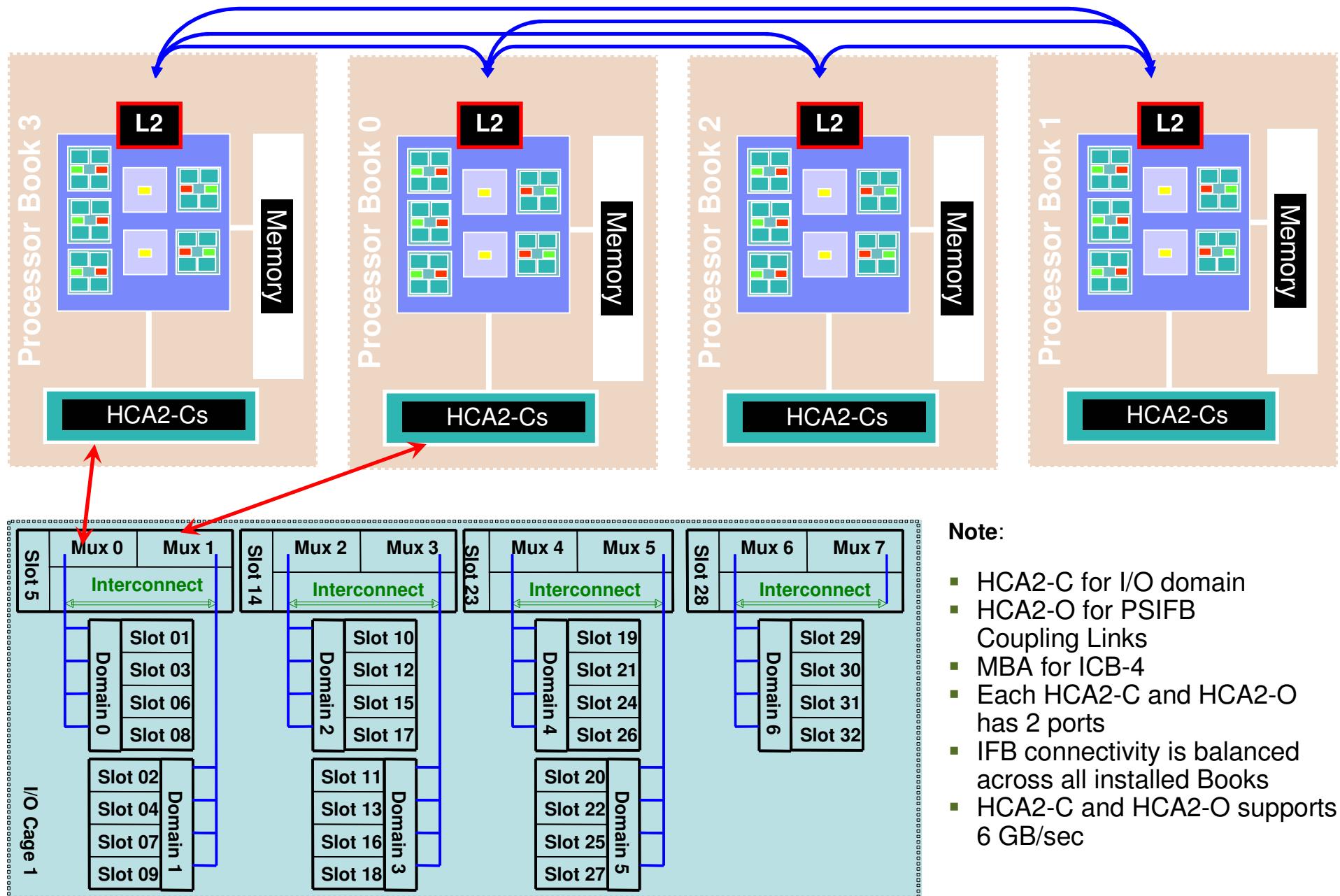
z10 EC Book Layout – Under the covers



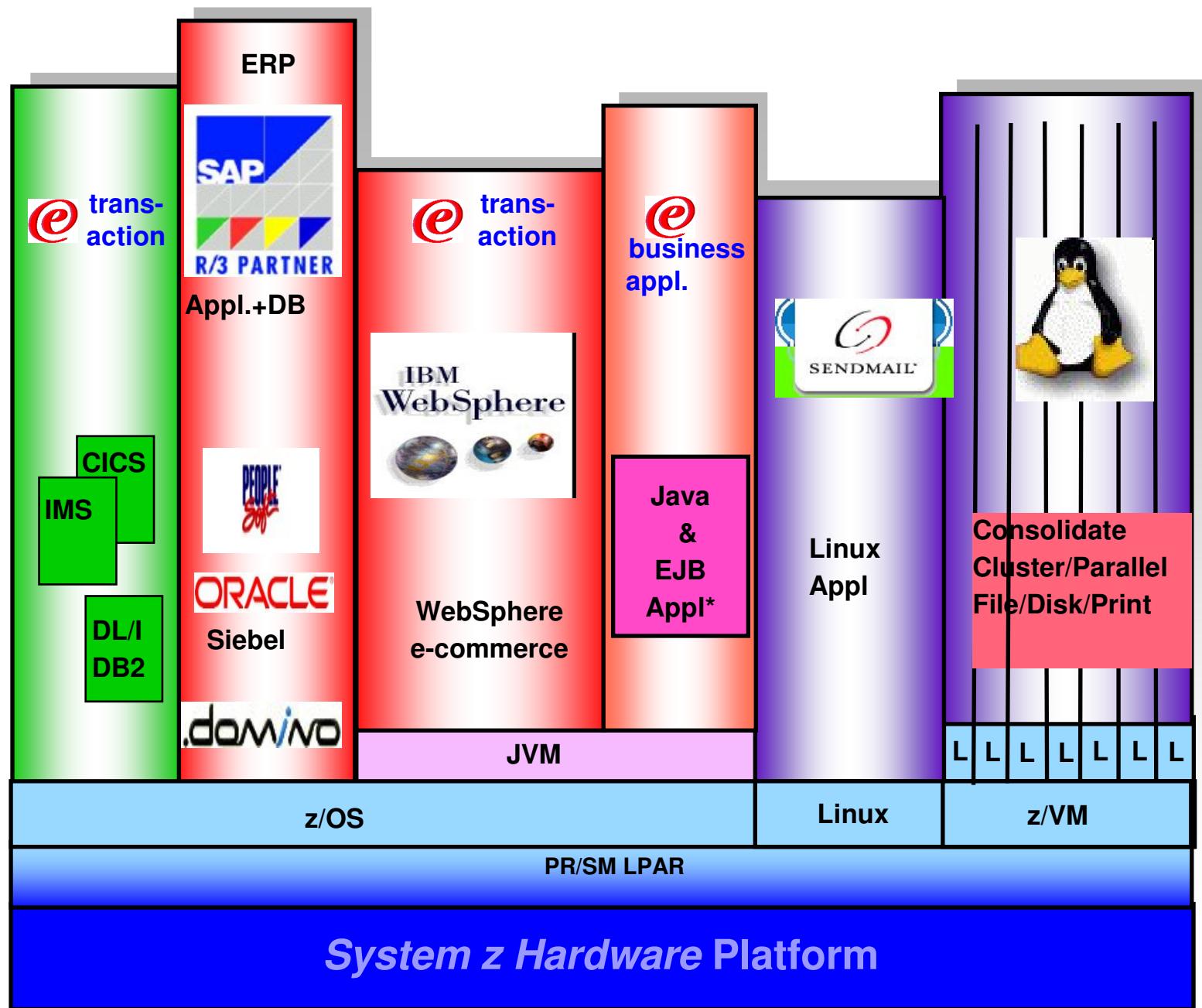
20 PU MCM Structure



z10 EC – Inter Book and I/O Communications – Models E54/E64



Mainframe / System z Server Environments

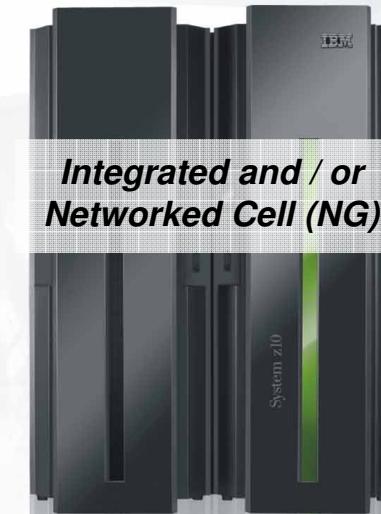
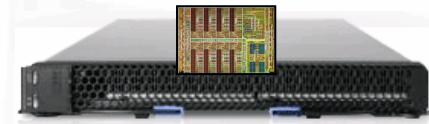


System z and Cell Broadband Engine – The Vision

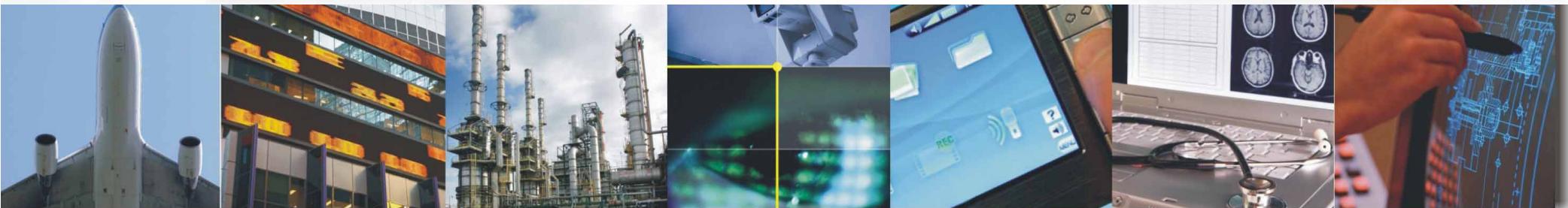
A ‘Marriage’ of Two Technologies that Perfectly Complement Each Other



z today



z tomorrow



Aerospace and
Defense

Financial
Services Sector

Chemicals and
Petroleum

Digital Video
Surveillance

Digital
Media

Information
Based Medicine

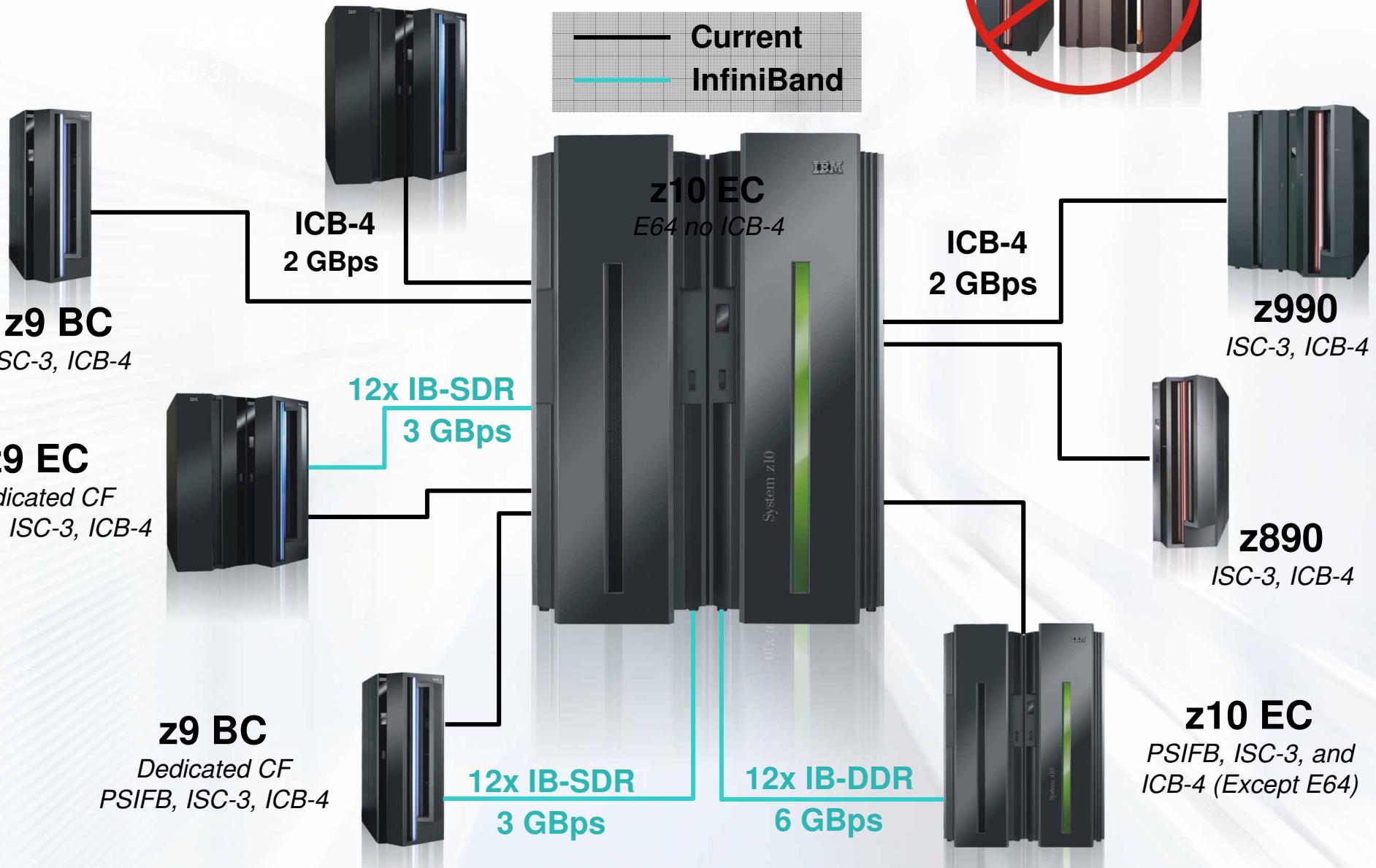
Electronic Design
Automation

z10 EC Parallel Sysplex coexistence and coupling connectivity

z800, z900
Not supported!



— Current
— InfiniBand

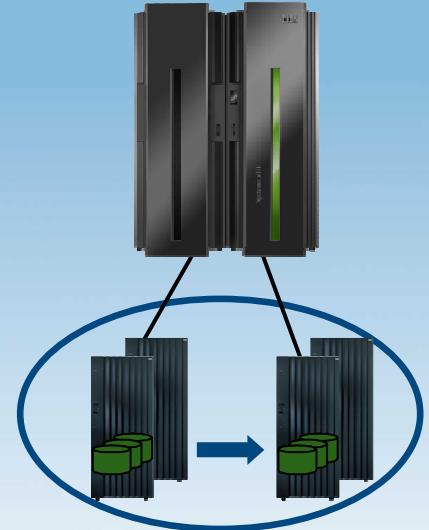


The right level of business continuity protection for your businessGDPS family of offerings

Continuous Availability of Data within a Data Center

Single Data Center
Applications remain active

Near-continuous availability to data

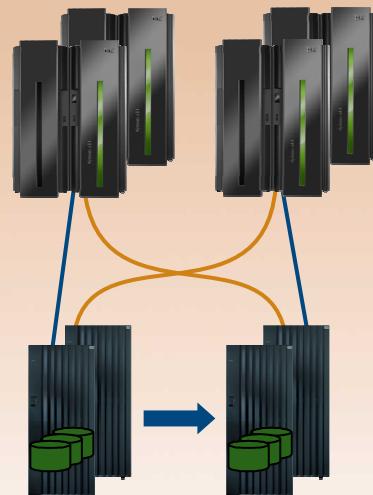


GDPS®/PPRC
HyperSwap Manager

Continuous Availability / Disaster Recovery Metropolitan Region

Two Data Centers
Systems remain active

Automated D/R across site or storage failure
No data loss

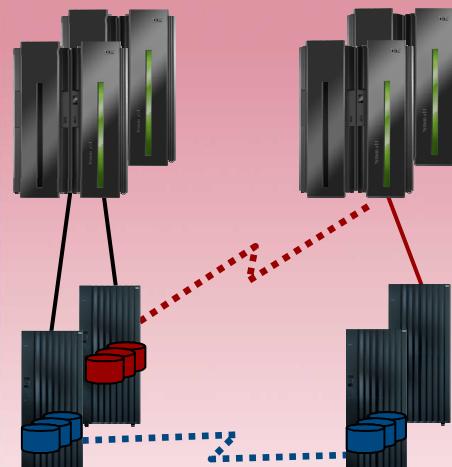


GDPS/ PPRC
HyperSwap Manager
GDPS/PPRC

Disaster Recovery at Extended Distance

Two Data Centers

Automated Disaster Recovery
“seconds” of Data Loss

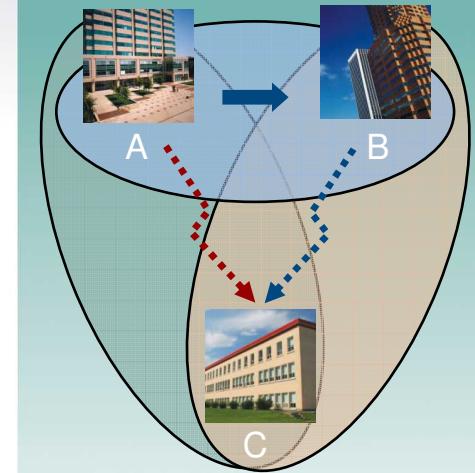


GDPS/GM
GDPS/XRC

Continuous Availability Regionally and Disaster Recovery Extended Distance

Three Data Centers

Data availability
No data loss
Extended distances



GDPS/MGM
GDPS/MzGM

Operating systems

z/OS

- Providing intelligent dispatching on z10 EC for performance
- Up to 64-way support
- Simplified capacity provisioning on z10 EC
- New high availability disk solution with simplified management
- Enabling extreme storage volume scaling
- Facilitating new zIIP exploitation

Linux on System z

- Large Page Support improves performance
- Linux CPU Node Affinity is designed to avoid cache pollution
- Software support for extended CP Assist instructions AES & SHA

z/TPF

- Support for 64+ processors
- Workload charge pricing
- Exploit encryption technology



z/VSE™

- Interoperability with Linux on System z
- Exploit encryption technology
- MWLC pricing with sub-capacity option

z/VM

- Consolidation of many virtual images in a single LPAR
- Enhanced management functions for virtual images
- Larger workloads with more scalability

Protecting your investment in IBM technology

- Designed to protect your investment by offering upgrades from z9 EC and z990 to the z10 EC
- Full upgradeability within the System z10 family
 - Upgrade to Model E64 will require a planned outage
- Temporary or permanent growth when you need it
 - New provisioning architecture



z10 EC Overview



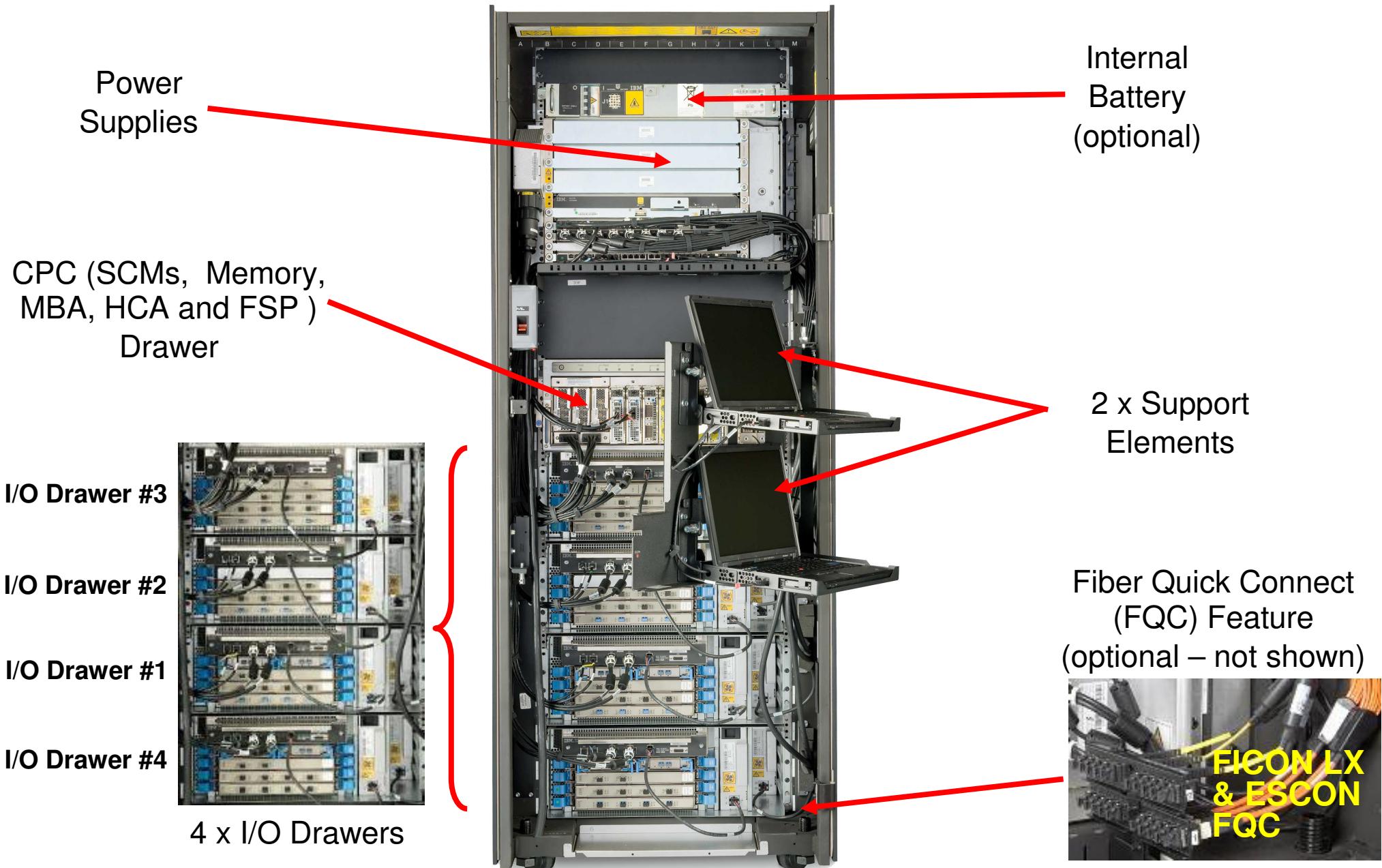
- **Machine Type**
 - 2097
- **5 Models**
 - E12, E26, E40, E56 and E64
- **Processor Units (PUs)**
 - 17 (17 and 20 for Model E64) PUs per book
 - Up to 11 SAPs per system, standard
 - 2 spares designated per system
 - Dependant on the H/W model - up to 12, 26, 40, 56 or 64 PUs available for characterization
 - Central Processors (CPs), Integrated Facility for Linux (IFLs), Internal Coupling Facility (ICFs), System z10 Application Assist Processors (zAAPs), System z10 Integrated Information Processor (zIIP), optional - additional System Assist Processors (SAPs)
- **Memory**
 - System Minimum of 16 GB
 - Up to 384 GB per book
 - Up to 1.5 TB GB for System
 - Fixed HSA, standard
 - 16/32/48/64 GB increments
- **I/O**
 - Up to 48 I/O Interconnects per System @ 6 GBps each
 - Up to 4 Logical Channel Subsystems (LCSSs)
- **ETR Feature, standard**

z10 BC Overview

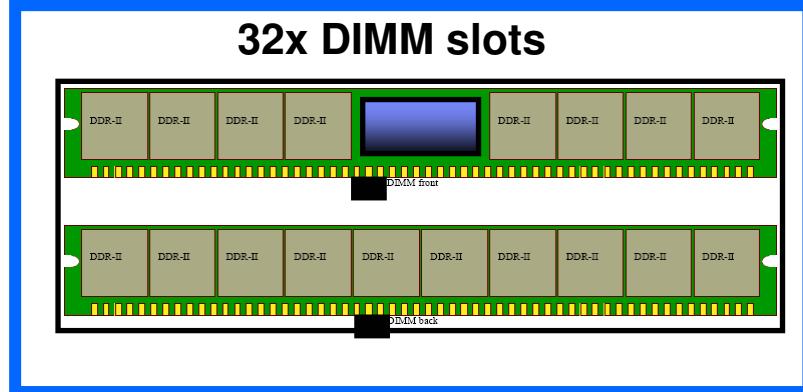


- **Machine Type**
 - 2098
- **Single Model – E10**
 - Single frame, air cooled
 - Non-raised floor option available
- **Processor Units (PUs)**
 - 12 PU cores per System
 - 2 SAPs, standard
 - Zero spares when all PUs characterized
 - Up to 10 PUs available for characterization
 - Central Processors (CPs), Integrated Facility for Linux (IFLs), Internal Coupling Facility (ICFs), System z10 Application Assist Processors (zAAPs), System z10 Integrated Information Processor (zIIP), optional - additional System Assist Processors (SAPs)
- **Memory**
 - System Minimum of 4 GB
 - Up to 128 GB for System, including HSA (up to 256 GB, June 30, 2009)
 - 8 GB Fixed HSA, standard
 - Up to 120 GB for customer use (up to 248 GB, June 30, 2009)
 - 4, 8 and 32 GB increments (32 GB increment, June 30, 2009)
- **I/O**
 - Up to 12 I/O Interconnects per System @ 6 GBps each
 - 2 Logical Channel Subsystems (LCSSs)
 - Fiber Quick Connect for ESCON and FICON LX
 - New OSA-Express3 Features
 - ETR feature, standard

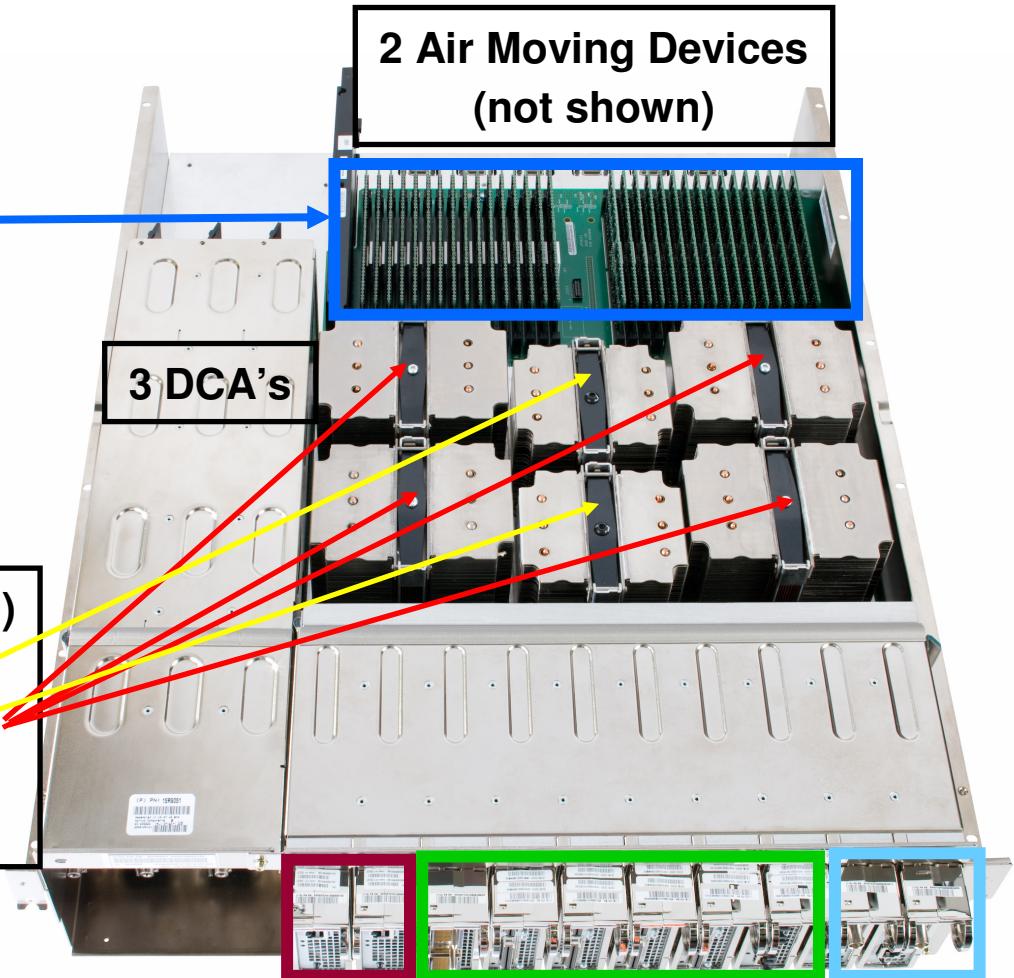
z10 BC – Under the covers Front View



z10 BC CPC and Memory Drawer Layout



**2 Air Moving Devices
(not shown)**



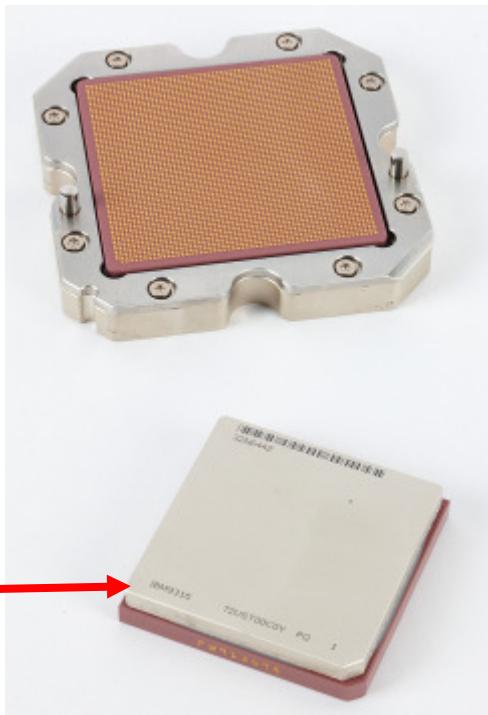
- Two new types of Single Chip Modules (SCMs)
- Processor – PU (4 SCM's x 3 cores = 12 PU's)
- System Controller – SC (2)

2 Flexible Support Processor (FSP) card slots providing support for the Service Network subsystem (hot swappable)

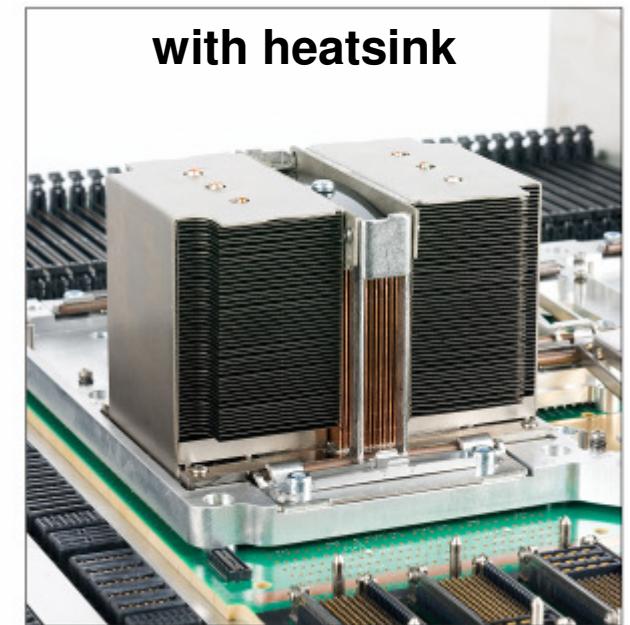
6 fanout card slots providing support for the I/O subsystem and/or coupling

2 card slots for the oscillator/ETR function (standard) – dynamic switchover support

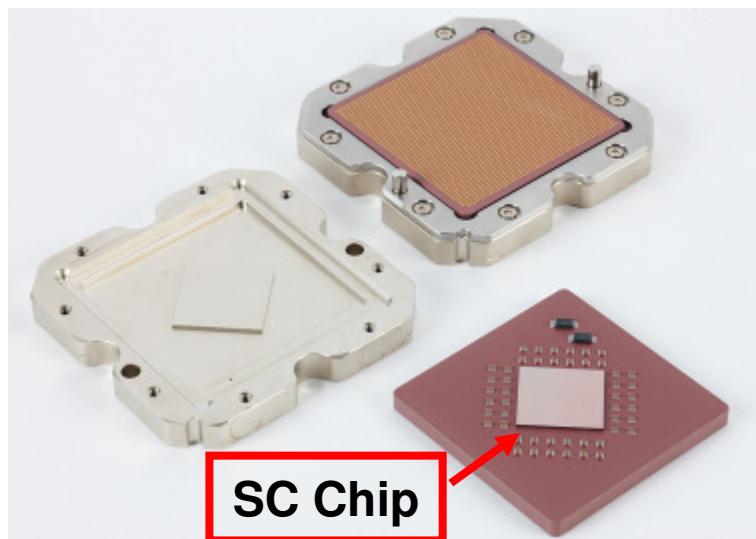
z10 BC PU/SC SCM Components



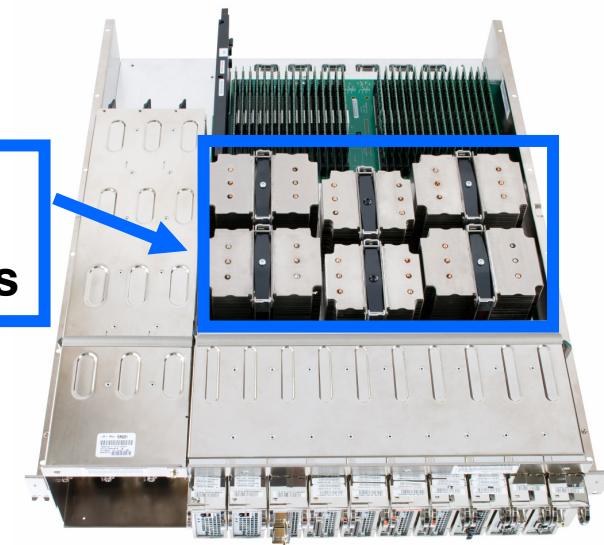
Assembled SCM
with heatsink



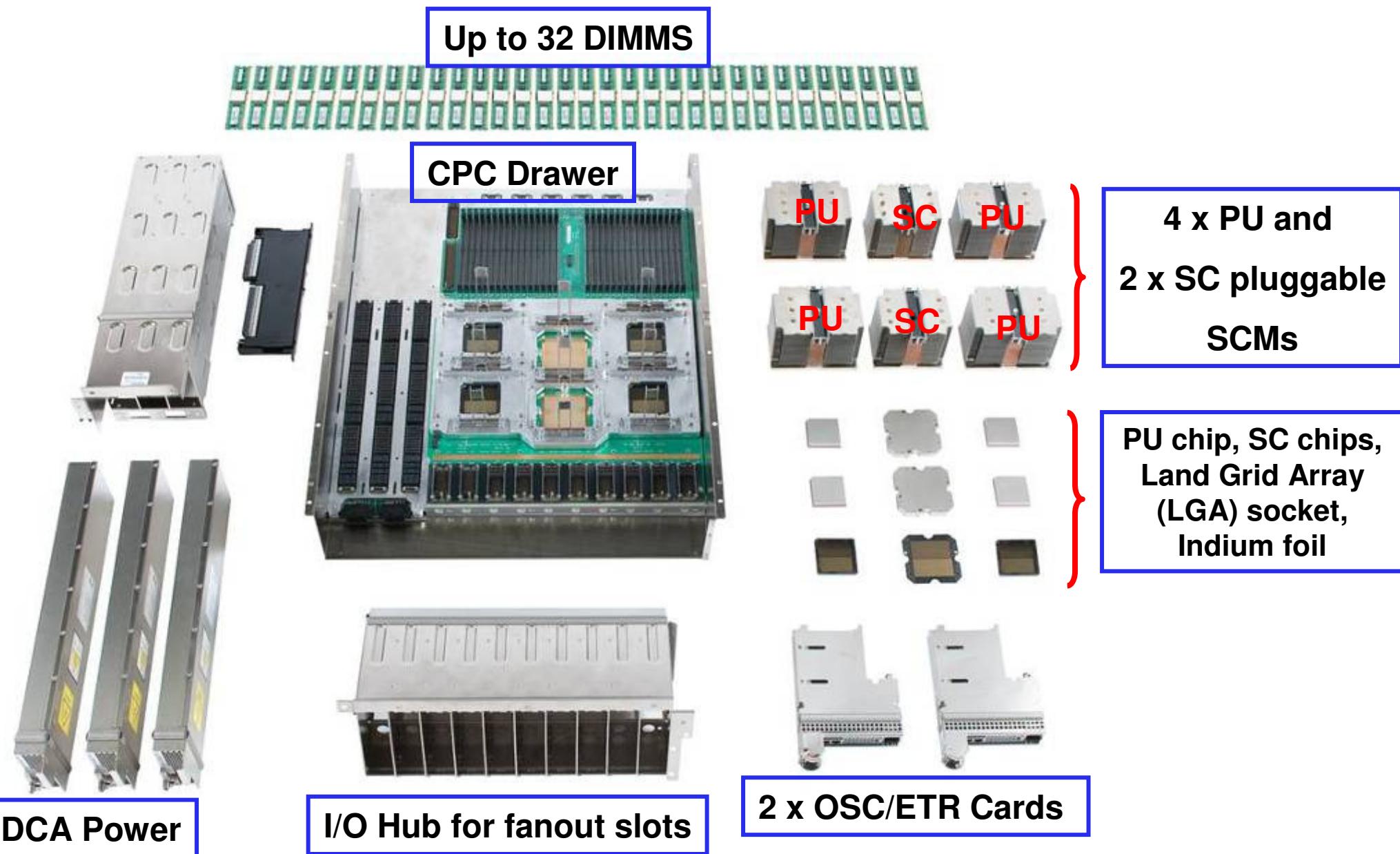
SCM Components



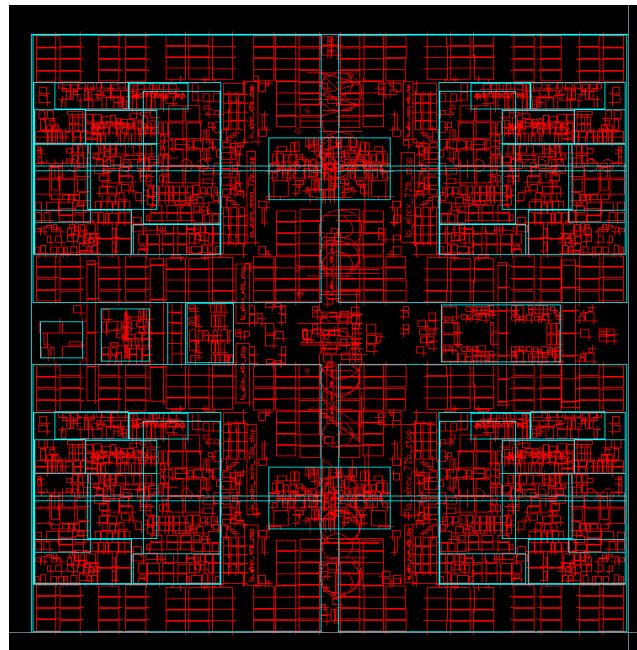
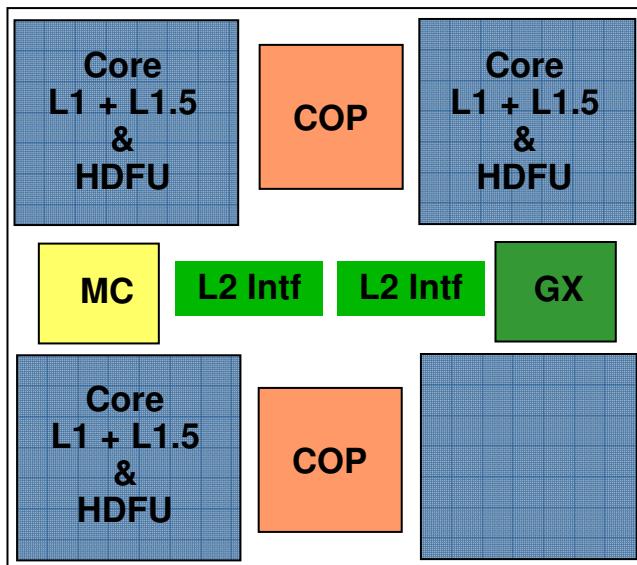
CPC Drawer with
4 PU SCMs, 2 SC SCMs



z10 BC CPC Drawer Components



z10 BC – Enterprise Quad Core z10 PU Chip

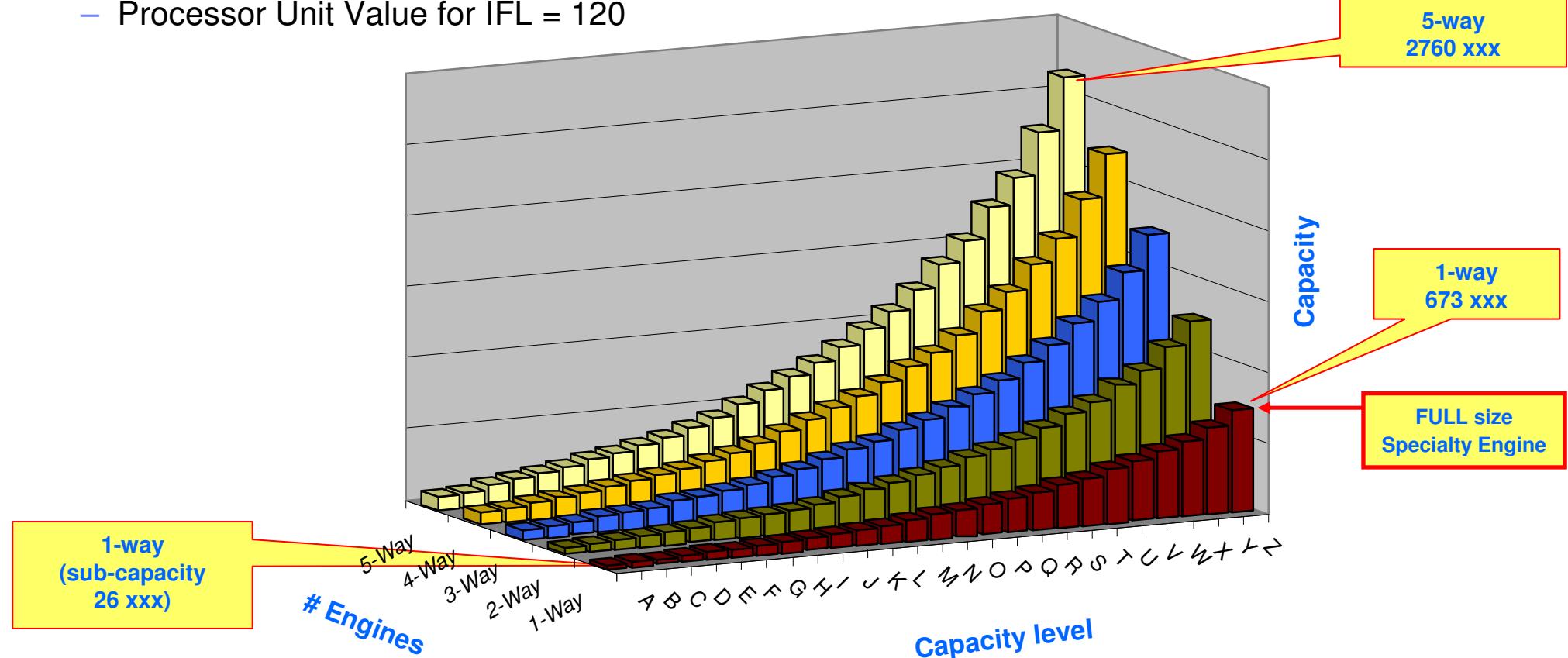


- **Three active cores per PU**
 - 3.5 GHz
 - 0.286 ns cycle time
 - L1 cache/PU core
 - 64 KB I-cache
 - 128 KB D-cache
 - 3 MB L1.5 cache/PU core
 - Each core with its own Hardware Decimal Floating Point Unit (HDFU)
- **Two Co-processors (COP)**
 - Accelerator engines
 - Data compression
 - Cryptographic functions
 - Includes 16 KB cache
 - Shared by two cores
- **L2 Cache interface**
 - Shared by all cores
 - Even/odd line (256B) split
- **I/O Bus Controller (GX)**
 - Interface to fanout
 - Compatible with System z9 MBA
- **Memory Controller (MC)**
 - Interface to controller on memory DIMMs

z10 BC Sub-capacity Processor Granularity

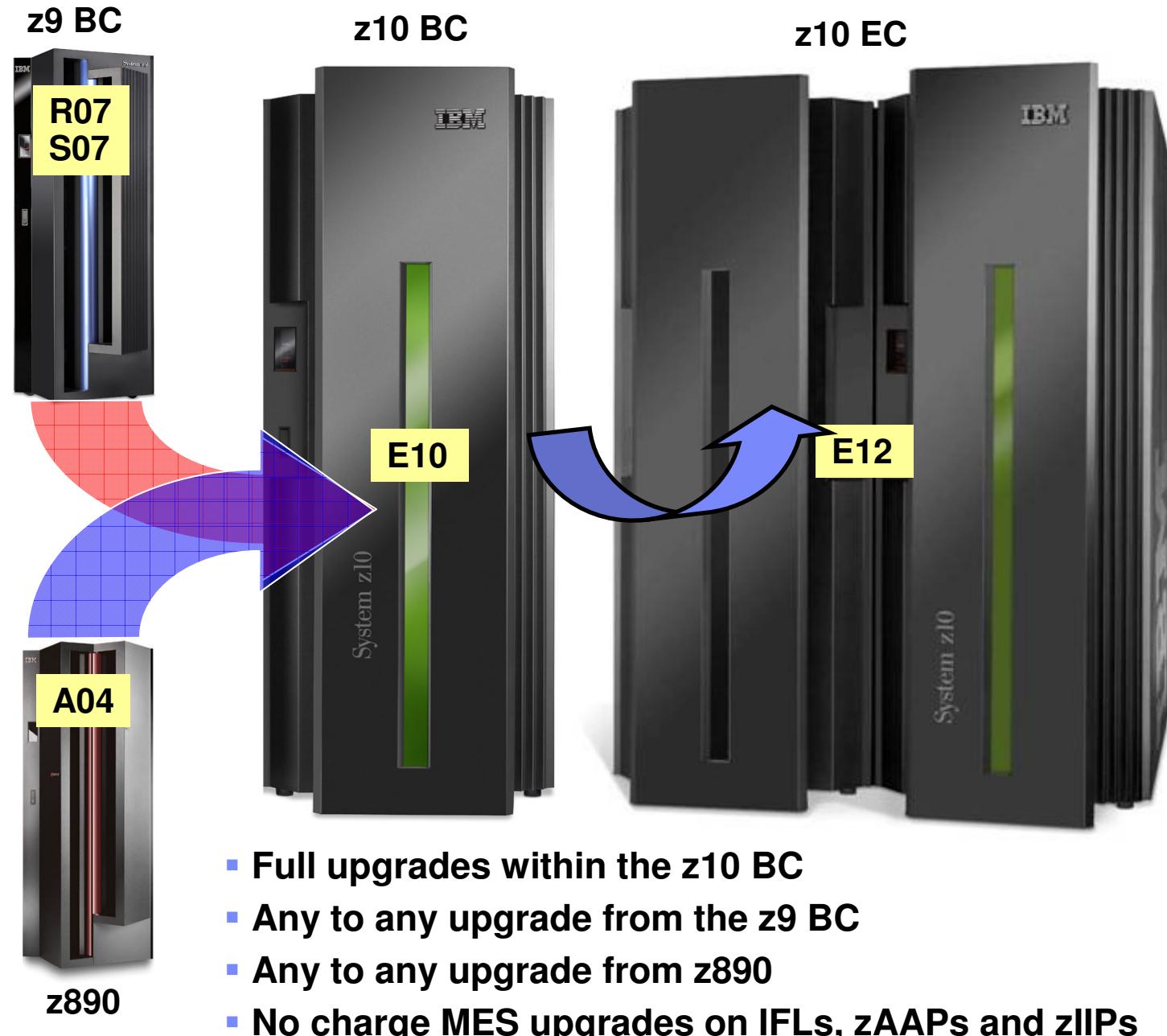
- The z10 BC has 26 CP capacity levels ($26 \times 5 = 130$)
 - Up to 5 CPs at any capacity level
 - All CPs must be the same capacity level
- The one for one entitlement to purchase one zAAP and/or one zIIP for each CP purchased is the same for CPs of any speed.
 - All specialty engines run at full speed
 - Processor Unit Value for IFL = 120

Number of z10 BC CPs	Base Ratio	Ratio z9BC to z10BC
1 CP	z9 BC Z01	1.40
2 CPs	z9 BC Z02	1.36
3 CPs	z9 BC Z03	1.30
4 CPs	z9 BC Z04	1.28
5 CPs	Z9 BC Z04	1.54



z10 BC Upgrade Paths

- Can enable dynamic and flexible capacity growth for mainframe servers
- Temporary capacity upgrade available through On/Off Capacity on Demand
- Temporary, nondisruptive addition of CP processors, IFLs, ICFs, zAAPs or zIIPs
- New options for reconfiguring specialty engines if the business demands it
- New options for changing On/Off CoD configurations
- Subcapacity CBU engines



I/O Subsystem – internal host bus interconnect speeds



InfiniBand I/O Bus
z10
2008



STI
z9
2005



STI
z990/z890
2003



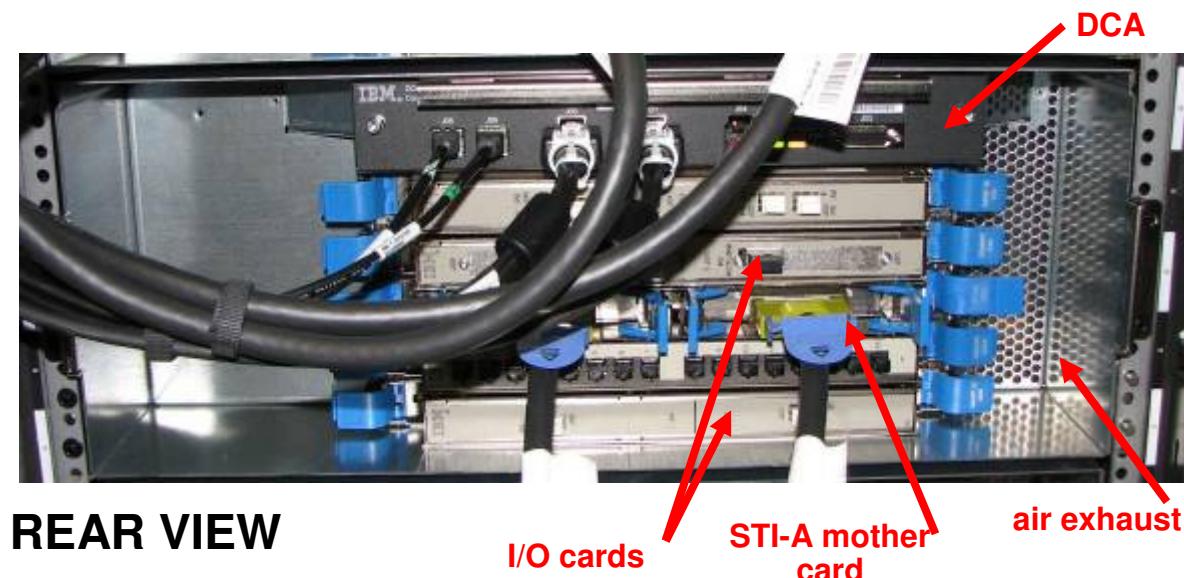
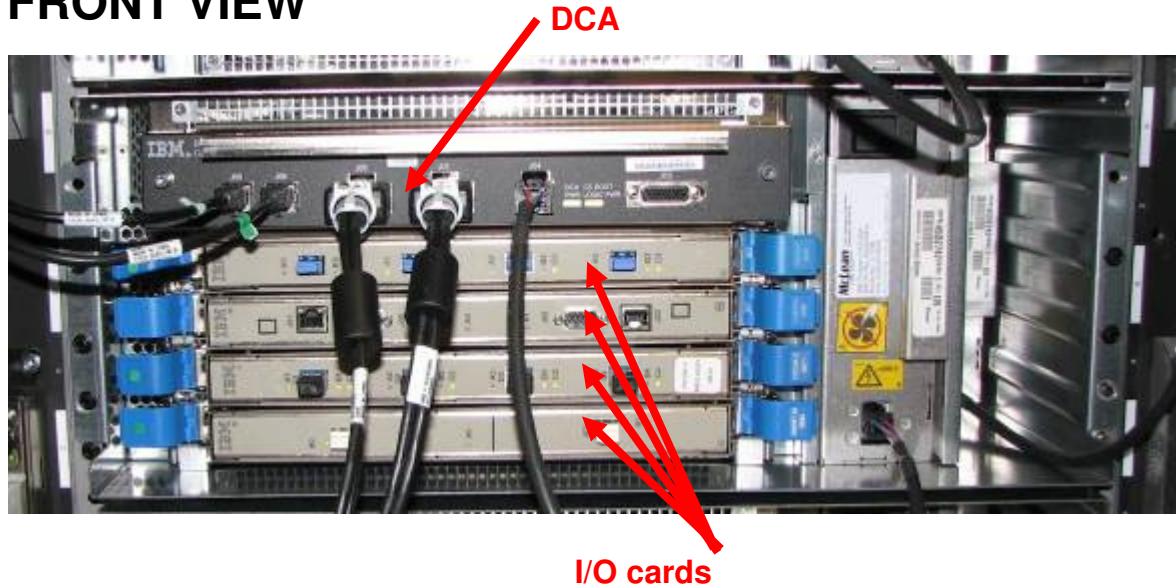
STI
z900/z800
200x



STI: Self-Timed Interconnect

z10 BC I/O Drawer

FRONT VIEW



REAR VIEW

- The new I/O drawer together with logic board are field replaceable units (FRUs)
- Drawer can be removed without affecting system input power or power to any other unit
- Up to 4 I/O drawers supported
 - Up to 8 I/O cards in each drawer
 - 4 in front and 4 in rear
- I/O cards are horizontal; Very important that cables are routed to the side or else concurrent replacement of I/O cards may not be possible
- Concurrent add
- Concurrent replacement, repair for Systems with >1 I/O drawer
- No support for I/O cage from prior Servers

z10 CoD Offerings

- **On-line Permanent Upgrade**
 - Permanent upgrade performed by customer (previously referred to Customer Initiated Upgrade - CIU)
- **Capacity Backup (CBU)**
 - For disaster recovery
 - Concurrently add CPs, IFLs, ICFs, zAAPs, zIIPs, SAPs
 - Pre-paid
- **Capacity for Planned Event (CPE)**
 - To replace capacity for short term lost within the enterprise due to a planned event such as a facility upgrade or system relocation
 - Predefined capacity for a fixed period of time (3 days)
 - Pre-paid
- **On/Off Capacity on Demand (On/Off CoD)**
 - Production Capacity
 - Supported through software offering – Capacity Provisioning Manager (CPM)
 - Payment:
 - Post-paid or Pre-paid by purchase of capacity tokens
 - Post-paid with unlimited capacity usage
 - On/Off CoD records and capacity tokens configured on Resource Link
- **Customer Initiated Upgrade (CIU)**
 - Process/tool for ordering temporary and permanent upgrades via Resource Link
 - Permanent upgrade support:
 - Un-assignment of currently active capacity
 - Reactivation of unassigned capacity
 - Purchase of all PU types physically available but not already characterized
 - Purchase of installed but not owned memory

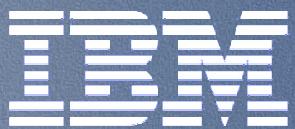
z10 BC System Power

- **z10 BC maximum configuration calculated AC input power (Statistical Maximum)**
 - All systems should draw less power than this
 - Typical systems will draw less power than this

	1 I/O Drawer	2 I/O Drawers	3 I/O Drawers	4 I/O Drawers
normal room (<28 degC)	3.686 kW	4.542 kW	5.308 kW	6.253 kW
warm room (>=28 degC)	4.339 kW	5.315 kW	6.291 kW	7.266 kW

- **30 Amp plug capacity (208 VAC)**
 - 5.5 kW single phase or unbalanced 3 phase
 - Supports up to 2 I/O drawers
 - 8.9 kW balanced 3 phase
 - Supports all system configurations – have balanced 3 phase feature
 - Plug 2 additional BPR's per side

Always refer to the z10 BC IMPP (GC28-6875) for detailed planning information



Noch Fragen?

010101010101010101010101

Danke !