ITSO Poughkeepsie

Redbooks IBM

zEnterprise.
A New Dimension in Computing

IBM zEnterprise: Technical Overview

Redbooks IBM

Trademarks

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

AIX*	IBM*	RACF*	z9*	z/VM*
BladeCenter*	IBM eServer	System x*	z10 BC	z/VSE
DataPower*	IBM (logo)*	System z*	z10 EC	
DB2*	Parallel Sysplex*	System z9*	zEnterprise	
FICON*	POWER7	System z10	z/OS*	
GDPS*	PR/SM	WebSphere*	zSeries*	
Geographically Dispersed Parallel Sysplex				

* 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 there from.
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.
InfiniBand is a trademark and service mark of the InfiniBand Trade Association.
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.

Redbooks IBM

IBM zEnterprise Technical Overview

- *z196 introduction*
- *z196 Key dates*
- *Product positioning*
- *z196 Highlights*
- *Model structure*
- *Granularity & scalability*
- *Memory configurations*
- *HSA considerations*
- *Flexible and Plan ahead memory*
- *z196 Functions & Features*

Redbooks IBM

IBM zEnterprise System naming guidelines

Description	Full name – first use	Name after first use
Family name	IBM System z®	System z
System name	IBM zEnterprise™ System	zEnterprise System or zEnterprise
Name on the outside of the server	zEnterprise	N/A
CPC name	IBM zEnterprise 196 (z196)	zEnterprise 196 or z196
Model numbers	M15, M32, M49, M66, M80	N/A
Hybrid infrastructure name	IBM zEnterprise BladeCenter®Extension (zBX) - Model 002	zEnterprise BladeCenter Extension or zBX
Name on the outside of the zBX	z BladeCenter Extension (note that there is a space between z and BladeCenter)	Do not use this on documents – not approved form except on the physical zBX
Management Firmware (ensemble management)	IBM zEnterprise Unified Resource Manager	Unified Resource Manager or zManager
Bundles	Two suites of tiered functionality – Manage suite and Automate suite	Manage, Automate
System x blade and WebSphere DataPower Appliance	IBM System x® blade* WebSphere® DataPower® Appliance*	Comment: Use the Statement of Direction footnote.
Linux running on a System x	Linux® on System x	Do not use other forms like Linux /x86. They are not correct.
Thus : z196 + zBX + Unified Resource Manager = zEnterprise System		

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

Redbooks IBM

IBM zEnterprise 196 (z196) Key Dates

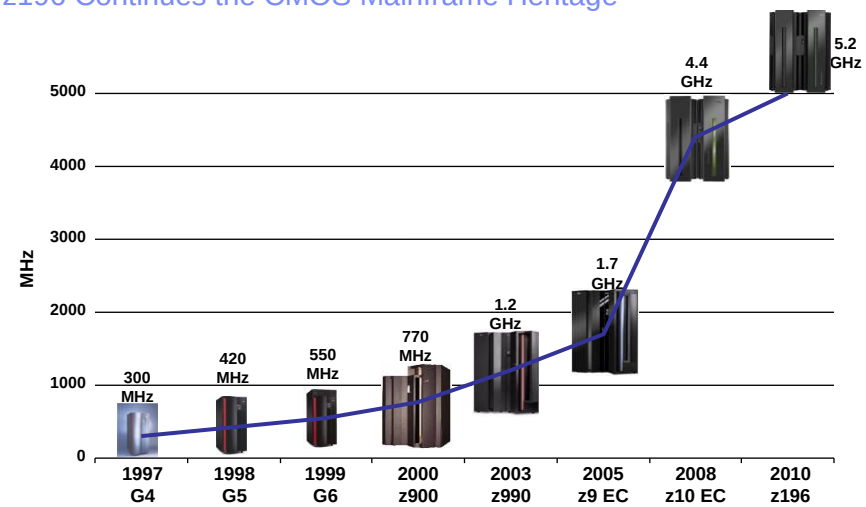
- **Planned Availability Dates:**
  - **September 10, 2010**
    - Features and functions for z196
    - z196 Models M15, M32, M49, M60 and M80
    - Manage Suite for zEnterprise
    - Water Cooling for z196
    - IBM System z9® Enterprise Class (z9® EC) upgrades to air-cooled z196
    - z9 EC upgrades to water-cooled z196
    - IBM System z10™ Enterprise Class (z10 EC™) upgrades to air-cooled z196
    - z10™ EC upgrades to water-cooled z196
    - 3-in-1 Bolt Down Kit for new build z196
    - System z discovery and autoconfiguration (zDAC)
  - **November 19, 2010**
    - Manage suite enhancement functions for zEnterprise
    - Automate suite for zEnterprise
    - IBM Smart Analytics Optimizer for DB2® for z/OS® for z196 in blade quantities of 7, 14 and 28
    - POWER7® blade for z196 in a zBX
  - **December 17, 2010**
    - IBM Smart Analytics Optimizer for DB2 for z/OS for z196 in blade quantities of 42 and 56
    - Model 002 MES feature upgrades within model
  - **December 31, 2010**
    - 3-in-1 Bolt Down Kit Mes for z196
    - MES features for Models M15, M32, M49, M66 and M80
    - Model conversions for z196

Note: At GA, Ethernet switch (#0070) is not available in the following countries: Belize, Brazil, Egypt, Haiti, Moldavia, Nicaragua, Oman, Qatar, and Venezuela. For alternatives, refer to the Ethernet LAN switch support section of Installation Manual - Physical Planning (IMPP), offering ID GC28-6897.

5

Redbooks IBM

z196 Continues the CMOS Mainframe Heritage



- **G4** – 1st full-custom CMOS S/390®
- **G5** – IEEE-standard BFP; branch target prediction
- **G6** – Copper Technology (Cu BEOL)

- **z900** – Full 64-bit z/Architecture
- **z990** – Superscalar CISC pipeline
- **z9 EC** – System level scaling

- **z10 EC** – Architectural extensions
- **z196** – Additional Architectural extensions and new cache structure

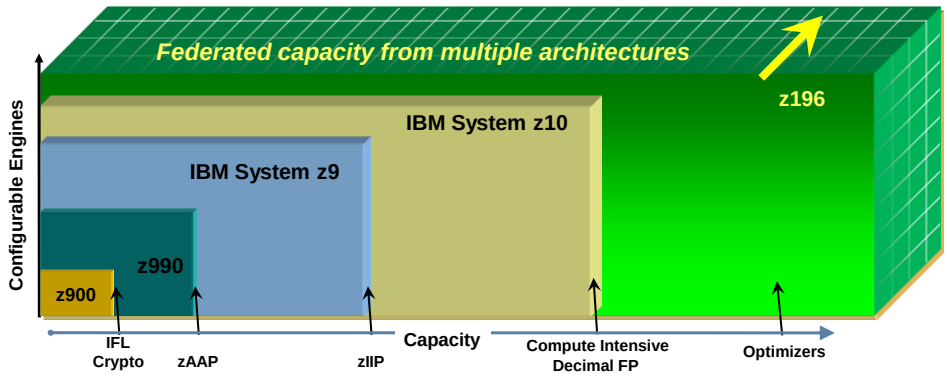
6

Redbooks IBM

z196 performance and scalability

Think Inside the box! Think System z Qualities of Service!

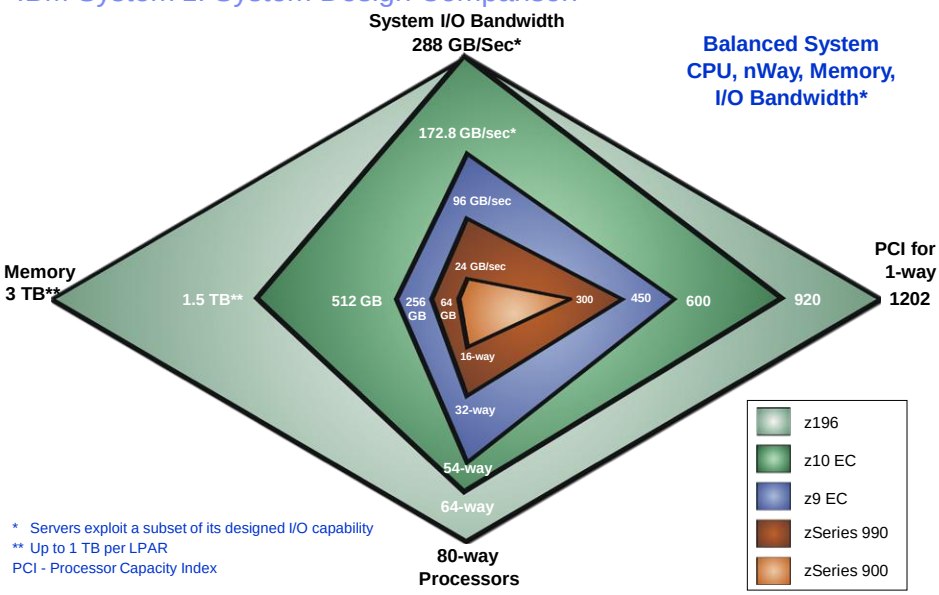
- **zEnterprise ensembles – Multiple nodes**
  - Node – z196 with or without zBX
- **zEnterprise Unified Resource Manager**
- **Multiple architectures**
  - z/Architecture®
  - Power Architecture® - POWER7
  - X-Architecture®
- **z196 ITR**
  - Largest z196 model 1.6x compared to z10 EC E64
  - Equivalent n-ways 1.4x compared to z10 EC
  - With compiler optimization additional 30% additional for some CPU intensive work



7

Redbooks IBM

IBM System z: System Design Comparison

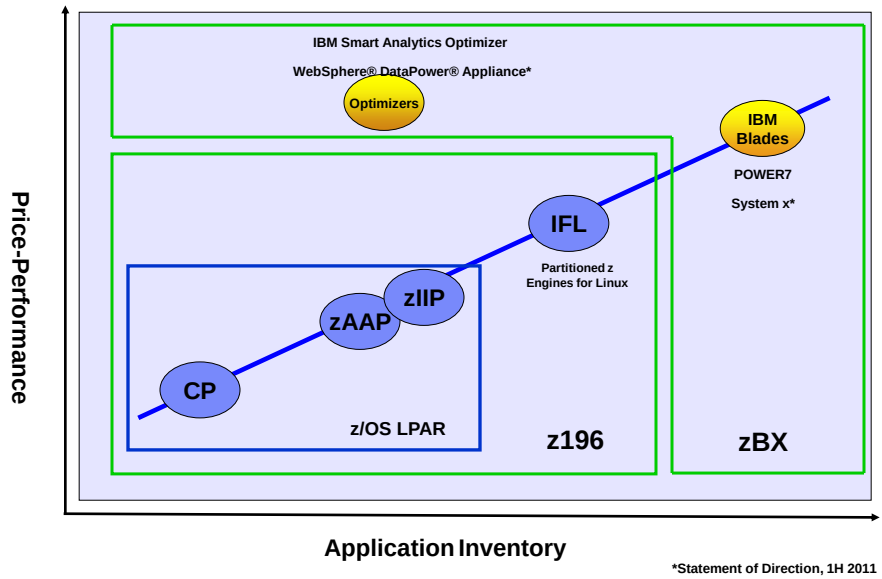


* Servers exploit a subset of its designed I/O capability
** Up to 1 TB per LPAR
PCI - Processor Capacity Index

8

Redbooks IBM

System z "Specialty Engine" Evolution to the zEnterprise Ensemble



*Statement of Direction, 1H 2011

Redbooks IBM

IBM System z family

IBM System z10 EC (2097)

- Announce 2/08 – Server w/ up to 77 PU cores
- 5 models – Up to 64-way
- Granular Offerings for up to 12 CPs
- PU (Engine) Characterization
  - CP, SAP, IFL, ICF, zAAP, zIIP
- On Demand Capabilities
  - CoD, CIU, CBU, On/Off CoD, CPE
- Memory – up to 1.5 TB for Server and up to 1 TB per LPAR
  - 16 GB Fixed HSA
- Channels
  - Four LCSSs
  - 2 Subchannel Sets
  - MIDAW facility
  - 63.75 subchannels
  - Up to 1024 ESCON® channels
  - Up to 336 FICON® channels
  - FICON Express2, 4 and 8
  - zHPF
  - OSA 10 GbE, GbE, 1000BASE-T
  - InfiniBand Coupling Links
- Configurable Crypto Express3
- Parallel Sysplex® clustering
- HiperSockets™ – up to 16
- Up to 60 logical partitions
- Enhanced Availability
- Operating Systems
  - z/OS, z/VM, z/VSE, TPF, z/TPF, Linux on System z

IBM System z10 BC (2098)

- Announced 10/08 – Server w/ 12 cores
- Single model – Up to 5-way CPs
- High levels of Granularity available
  - 130 Capacity Indicators
- PU (Engine) Characterization
  - CP, SAP, IFL, ICF, zAAP, zIIP
- On Demand Capabilities
  - CoD, CIU, CBU, On/Off CoD, CPE
- Memory – up to 256 GB for Server
  - 8 GB Fixed HSA
- Channels
  - Two LCSSs
  - 2 Subchannel Sets
  - MIDAW facility
  - 63.75 subchannels
  - Up to 480 ESCON channels
  - Up to 128 FICON channels
  - FICON Express2, 4 and 8
  - zHPF
  - OSA 10 GbE, GbE, 1000BASE-T
  - InfiniBand Coupling Links
- Configurable Crypto Express3
- Parallel Sysplex clustering
- HiperSockets – up to 16
- Up to 30 logical partitions
- Enhanced Availability
- Operating Systems
  - z/OS, z/OS.e, z/VM, z/VSE, TPF, z/TPF, Linux on System z

IBM zEnterprise z196 (2817)

- Announce 7/10 – Server w/ up to 96 PU cores
- 5 models – Up to 80-way
- Granular Offerings for up to 15 CPs
- PU (Engine) Characterization
  - CP, SAP, IFL, ICF, zAAP, zIIP
- On Demand Capabilities
  - CoD, CIU, CBU, On/Off CoD, CPE
- Memory – up to 3 TB for Server and up to 1 TB per LPAR
  - 16 GB Fixed HSA
- Channels
  - Four LCSSs
  - 3 Subchannel Sets
  - MIDAW facility
  - 63.75 subchannels
  - Up to 240 ESCON channels
  - Up to 288 FICON channels
  - FICON Express4 and 8
  - zHPF
  - OSA 10 GbE, GbE, 1000BASE-T
  - InfiniBand Coupling Links
- Configurable Crypto Express3
- Parallel Sysplex clustering
- HiperSockets – up to 32
- Up to 60 logical partitions
- Enhanced Availability
- Unified Resource Manager
- Operating Systems
  - z/OS, z/VM, z/VSE, z/TPF, Linux on System z

IBM zEnterprise Blade Extension (2458)

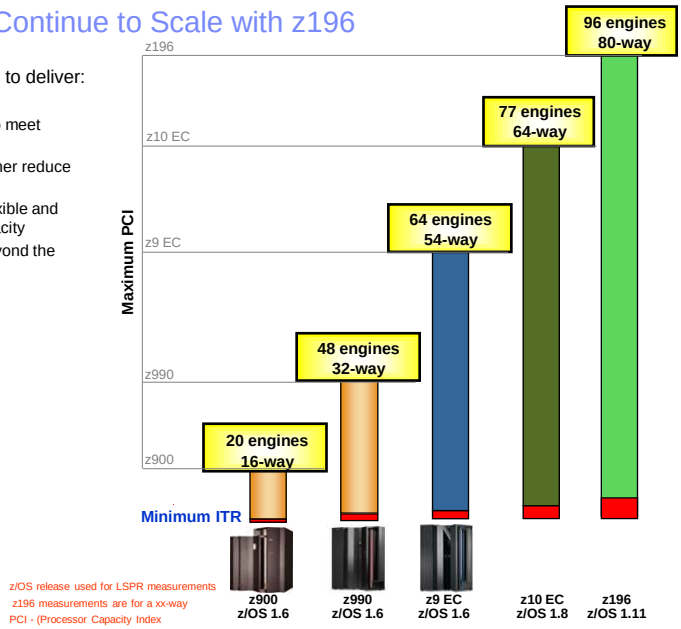
- Announce 7/10
- 2 models
- zBX Racks with:
  - BladeCenter Chassis
  - N + 1 components
  - Blades
  - Top of Rack Switches
  - 8 Gb FC Switches
  - Power Units
  - Advance Management Modules
- Model 002 for z196
- 0 to 112 Blades
  - IBM Smart Analytics Optimizer Solution
  - POWER7 Blades
  - System x Blades*
  - WebSphere DataPower Appliance*

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

Redbooks IBM

System z Servers Continue to Scale with z196

- Each new range continues to deliver:
  - New function
  - Unprecedented capacity to meet consolidation needs
  - Improved efficiency to further reduce energy consumption
  - Continues to delivering flexible and simplified on demand capacity
  - A mainframe that goes beyond the traditional paradigm



z/OS release used for LSPR measurements
z196 measurements are for a xx-way
PCI - (Processor Capacity Index

Redbooks IBM

z196 GHz Leadership

- **System z continues to grow frequency by**
  - Leveraging 45nm technology
  - Advancing low latency pipeline design
  - Leveraging high speed yet power-efficient circuit designs
- **z196 achieves industry leadership 5.2 GHz operation**
- **High-frequency processor design is enabled by**
  - Innovative power-efficient circuit designs
  - IBM's SOI silicon technology
  - Leadership packaging technology including glass-ceramic MCMs
  - Leadership cooling technology, with choice of MRU or chilled-water cooling
- **GHz is especially important for CPU-intensive applications**
- **System performance is not linear with frequency**
  - Need to use LSPR + System z capacity planning tools for real client / workload sizing
- **GHz is not the only dimension that matters**
  - Must balance frequency, pipeline efficiency, energy efficiency, cache/memory design, I/O design

Redbooks IBM

z196 Architecture

- **Continues line of upward-compatible mainframe processors**
- **Rich CISC Instruction Set Architecture (ISA)**
  - 984 instructions (762 implemented entirely in hardware)
  - 24, 31, and 64-bit addressing modes
  - Multiple address spaces robust inter-process security
  - Multiple arithmetic formats
  - Industry-leading virtualization support
    - High-performance logical partitioning via PR/SM
    - Fine-grained virtualization via z/VM scales to 1000's of images
  - Precise, model-independent definition of hardware/software interface
- **Architectural extensions for IBM z196**
  - 100+ new instructions added to improve compiled code efficiency
  - Decimal floating point quantum exceptions
  - New crypto functions and modes
  - Virtual architectural level
  - Non-quiescing SSKE

Redbooks IBM

z196 Highlights

Machine Type
- 2817

5 Models
- M15, M32, M49, M66 and M80

Processor Units (PUs)
- 20 (24 for M80) PU cores per book
- Up to 14 SAPs per system, standard
- 2 spares designated per system
- Dependent on the H/W model - up to 15,32,49,66 or 80 PUs available for characterization
  - Central Processors (CPs), Integrated Facility for Linux (IFLs), Internal Coupling Facility (ICFs), System z Application Assist Processors (zAAPs), System z Integrated Information Processor (zIIP), optional - additional System Assist Processors (SAPs)
- Sub-capacity available for up to 15 CPs
  - 3 sub-capacity points

Memory
- System Minimum of 32 GB
- Up to 768 GB per book
- Up to 3 TB for System and up to 1 TB per LPAR
  - Fixed HSA, standard
  - 32/64/96/112/128/256 GB increments

I/O
- Up to 48 I/O Interconnects per System @ 6 GBps each
- Up to 4 Logical Channel Subsystems (LCSSs)
- Up to 3 Subchannel Sets per LCSS

- **STP - optional** (No ETR)

Redbooks IBM

z196 - PU Distribution

z196																	
H/w Model	Total CPUs	Total SAPs	1st Book - LG06				2nd Book - LG15				3rd Book - LG10				4th Book - LG01			
			PUs	Avail	SAPs	Spare	PUs	Avail	SAPs	Spare	PUs	Avail	SAPs	Spare	PUs	Avail	SAPs	Spare
M15	15	3	20	15	3	2	-	-	-	-	-	-	-	-	-	-	-	-
M32	32	6	20	16	3	1	20	16	3	1	-	-	-	-	-	-	-	-
M49	49	9	20	16	3	1	20	16	3	1	20	17	3	0	-	-	-	-
M66	66	12	20	16	3	1	20	16	3	1	20	17	3	0	20	17	3	0
M80	80	14	24	20	3	1	24	20	3	1	24	20	4	0	24	20	4	0

Redbooks IBM

z196 PU Characterization

Model	Books/ PUs	CPs	IFLs uIFLs	zAAPs	zIIPs	ICFs	SAPs Std	Optional SAPs	Std. Spares
M15	1/20	0-15	0-15 0-14	0-7	0-7	0-15	3	0-4	2
M32	2/40	0-32	0-32 0-31	0-16	0-16	0-16	6	0-10	2
M49	3/60	0-49	0-49 0-48	0-24	0-24	0-16	9	0-15	2
M66	4/80	0-66	0-66 0-65	0-33	0-33	0-16	12	0-20	2
M80	4/96	0-80	0-80 0-79	0-40	0-40	0-16	14	0-18	2

- z196 Models M15 to M66 use books each with a 20 core MCM (two 4-core and four 3-core PU chips)
- Concurrent Book Add is available to upgrade from model to model (except to the M80)
- z196 Model M80 has four books each with a 24 core MCM (six 4-core PU chips)
- Disruptive upgrade to z196 Model M80 is done by book replacement

Notes: 1. At least one CP, IFL, or ICF must be purchased in every machine
2. One zAAP **and** one zIIP may be purchased for each CP purchased even if CP capacity is "banked".
3. "uIFL" stands for Unassigned IFL

Redbooks IBM

z196 Concurrent PU core Conversions

- **Must order (characterize one PU core as) a CP, an ICF or an IFL**
- **Concurrent model upgrade (book add) is supported – M15 to M32 to M49 to M66**
  - Upgrades to Model M80 are disruptive
- **Concurrent processor upgrade is supported if PU cores are available**
  - Add CP, IFL, unassigned IFL, ICF, zAAP, zIIP or optional SAP

From/To->	CP	IFL	Unassigned IFL	ICF	zAAP	zIIP	Optional SAP
CP	x	Yes	Yes	Yes	Yes	Yes	Yes
IFL	Yes	x	Yes	Yes	Yes	Yes	Yes
Unassigned IFL	Yes	Yes	x	Yes	Yes	Yes	Yes
ICF	Yes	Yes	Yes	x	Yes	Yes	Yes
zAAP	Yes	Yes	Yes	Yes	x	Yes	Yes
zIIP	Yes	Yes	Yes	Yes	Yes	x	Yes
Optional SAP	Yes	Yes	Yes	Yes	Yes	Yes	x

YES

Exceptions: Disruptive if ALL current PU cores are converted to different types May require individual LPAR disruption if dedicated PU cores are converted.

17

Redbooks IBM

Model Summary I/O Memory Matrix

Model	Processors	Customer Memory[1]	HCA[2]	I/O Cages[3]	I/O Drawers
M15	1 to 15	32 to 752GB	16	0 to 2	0-6
M32	1 to 32	32 to 1520GB	32	0 to 2	0-6
M49	1 to 49	32 to 2288GB	40	0 to 2	0-6
M66	1 to 66	32 to 3056GB	48	0 to 2	0-6
M80	1 to 80	32 to 3056GB	48	0 to 2	0-6

1.An additional 16GB is delivered and reserved for HSA. 1 additional memory channel will be dedicated for RAIM. Every fifth dimm configured is reserved for this purpose.

2.Total of internal (to I/O cage) and external IFB (coupling) ports. With the addition of the third and fourth books air flow concerns require that the number of hub cards plugged be reduced to increase cooling around the MCM.

3.With RPQ 8P2506 up to three I/O cages can be ordered

4.Offerings are the same for both MRU/air and MWU.

18

Redbooks IBM

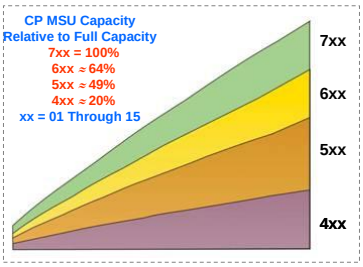
CP Granularity on z196

CP	Speed	Model Capacity Indicator	Kneecap ID	Capacity Marker Features	Related Pricing Features
1877	CP4	4xx where xx = 01 to 15	54	7202-7216 where xx = 01 to14	1751-1765 where xx = 01 to 15
1878	CP5	5xx where xx = 01 to 15	55	7217-7231 where xx = 01 to 14	1766-1780 where xx = 01 to 15
1879	CP6	6xx where xx = 01 to 15	56	7232-7246 where xx = 01 to 14	1781-1795 where xx = 01 to 15
1880	CP7	7xx where xx = 00 to 80	57	7247-7327 where xx = 00 to 80	1796-1876 where xx = 00 to 80

Redbooks IBM

Model Capacity Identifier

Hardware Model	Model Capacity Identifier			
	Full Capacity	Sub-Capacity		
M15	700 –715	601-615	501-515	401-415
M32	700 –732	601-615	501-515	401-415
M49	700 –749	601-615	501-515	401-415
M66	700 –766	601-615	501-515	401-415
M80	700 –780	601-615	501-515	401-415



- The CP Model is the critical value by which the Independent Software Venders (ISV) sets the price for software billing
- Support Element code will pass Model Capacity Indicators for both permanent and temporary capacity
- The indicator represents the number of active CPs rather than total physical PUs delivered (purchased)
- There is only one set of Model Capacity Indicators for both MRU and MWU models

Redbooks IBM

z196 – Hardware System Area considerations

- Fixed HSA of 16 GB provided as standard
- The HSA has been designed to eliminate planning for HSA. Preplanning for HSA expansion for configurations will be eliminated as HCD/IOCP will, via the IOCDS process, always reserve:
  - 4 Logical Channel Subsystems (LCSS), pre-defined
  - 60 Logical Partitions (LPARS), pre-defined
  - Subchannel set 0 with 63.75k devices
  - Subchannel set 1 with 64K-1 devices
  - Subchannel set 2 with 64K-1 devices
  - Dynamic I/O Reconfiguration - always enabled by default
  - Concurrent Patch - always enabled by default
  - Add/Change the number of logical CP, IFL, ICF, zAAP, zIIP, processors per partition and add SAPs to the configuration
  - Dynamic LPAR PU assignment optimization CPs, ICFs, IFLs, zAAPs, zIIPs, SAPs
  - Dynamically Add/Remove Crypto (no LPAR deactivation required)

Redbooks IBM

z196 HSA Comparisons

- **HSA significantly larger than pre-z10 Servers**
- **Fixed 16 GB HSA and does not affect customer purchased memory**
- **Size of HSA on prior Servers (dependant on defined configuration)**
  - Multiprise® 2000 — From 12 MB up to 40 MB
  - 9672 G4 — From 48 up to 64 MB
  - Multiprise 3000 — From 38 MB up to 136 MB
  - 9672 G5/G6 — From 64 MB up to 192 MB
  - z800 — From 160 MB up to 256 MB
  - z900 — From 288 MB up to 512 MB
  - z890 — From 768 MB up to 1.9 GB
  - z990 — From 1 GB MB up to 2 GB
  - z9 BC — From 896 MB up to 2.7 GB
  - z9 EC — From 1.2 GB up to 4.2 GB
  - z10 BC — 8 GB – Fixed
  - z10 EC — 16 GB – Fixed
  - z196 — 16 GB – Fixed
- **HSA Estimator on Resource Link not relevant**

Redbooks IBM

zEnterprise z196 Functions and Features (GA Driver 86 – August, 2010)

Five hardware models
Quad core PU chip
Up to 80 processors configurable as CPs, zAAPs, zIIPs, IFLs, ICFs, or optional SAPs
Increased capacity processors
Out of order instruction execution
Over 100 new and enhanced instructions
Improved processor cache design
Up to 15 subcapacity CPs at capacity settings 4, 5, or 6
Up to 3 TB of Redundant Array of Independent Memory (RAIM)
Unified Resource Manager suites
Cryptographic enhancements
On Demand enhancements
Energy efficiencies

2 New OSA CHPIDs – OSX and OSM
Three subchannel sets per LCSS
8 slot, 2 domain I/O drawer
Concurrent I/O drawer add, remove, replace
FICON discovery and autoconfiguration
Doubled HiperSockets to 32
Physical Coupling Links increased to 80
Doubled Coupling CHPIDs to 128
CFCC Level 17
Optional water cooling
Optional High Voltage DC power
Optional overhead I/O cable exit
STP enhancements
zBX-002 with IBM Smart Analytics Optimizer, IBM Blades

Redbooks IBM

IBM zEnterprise System Structure & Design

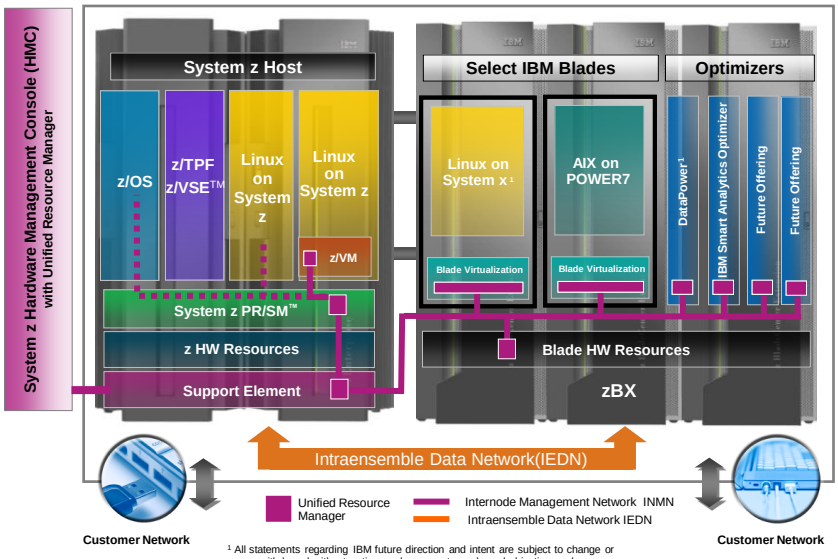
- *System design overview*
- *Processor Book*
- *CPC components*
- *z196 Quad Core PU Chip*
- *PU core functional units*
- *Out-of-order execution*
- *Compression and Cryptography*
- *Memory Design*
- *Book and HCA Plugging*

Redbooks IBM

z196 System Design Overview

Redbooks IBM

zEnterprise – System Design



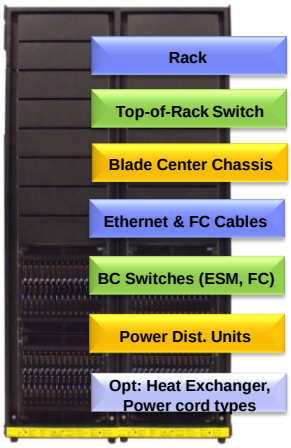
¹ All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.

Redbooks IBM

z196 and zBX Hardware Components

z196

zBX Infrastructure



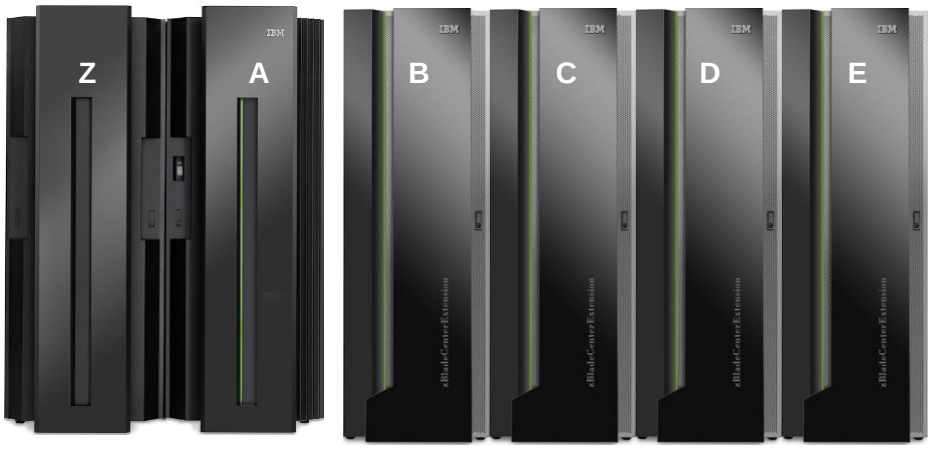
- **New System z Machine Types:**
  - CPC = 2817
  - zBX = 2458 Model 002
  - Customer supplied IBM System x* and POWER7 Blades
- **Key points:**
  1. CPC Availability and Integrity
  2. Independent Lifecycles
     - Asynchronously Upgradeable
     - Enhancements not tied to CPC HW
  3. Data availability not tied to single CPC
     - Shareable between CPCs
  4. Growth
  5. System z service

*All statements regarding IBM's plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these Statements of General Direction is at the relying party's sole risk and will not create liability or obligation for IBM.

27

Redbooks IBM

IBM zEnterprise System – Frame Layout



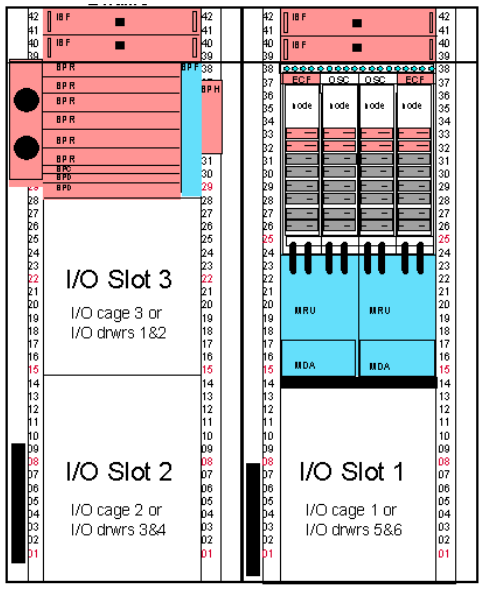
IBM zEnterprise 196 (z196) | **IBM zEnterprise BladeCenter Extension (zBX)**

IBM zEnterprise Unified Resource Manager (zManager)

28

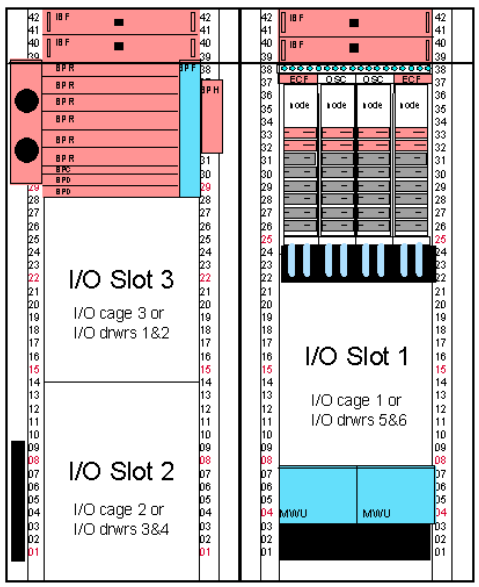
Redbooks IBM

z196 Frame Layouts – Air Cooled Front View



Redbooks IBM

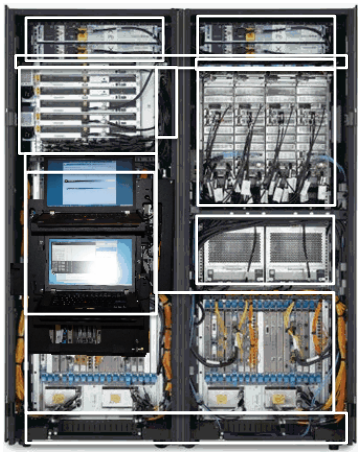
z196 Frame Layouts – Water Cooled Front View



Redbooks IBM

IBM zEnterprize 196: Front View - Air Cooled

- ***Internal Battery Feature**:*
  - ***A-Frame**:* Each 2U mounted on the front only
  - ***Z-Frame**:* Each 2U, Internal Batteries Feature (4x total - back to back)
- ***Frame Brake**:* top hat removal for height reduction feature
- ***VIPAL Board**:* A 2U board that from the front side houses:
  - The PPS/FSP and Oscillator cards
- ***Processors Book**:* 1 to 4 Processor Books
- ***MRU**:* The 1 or 2, side-by-side, 7U hybrid refrigeration Modular Refrigeration Units
- ***BPA**:* The 8U Bulk Power Assembly (back to back) which houses up to 3 BPR's, 1 BPC, 1 BPD, 1 BPH, 1 BPF and Line Cord. (3 BPR´s are shown: standard + two for 3 phase feature)
- ***BPH-A:*** Hub which provides internal communication within the system elements
- ***I/O Cages and/or Drawers**:* 14U (Cargo Cage) or 5U (Cayuga Drawer) which contain the I/O, STI-MP cards and DCAs
- ***SE Tray**:* Houses the primary and alternate Support Elements
- ***Tailgate Asm.:*** For possible fiber trunk system



Redbooks IBM

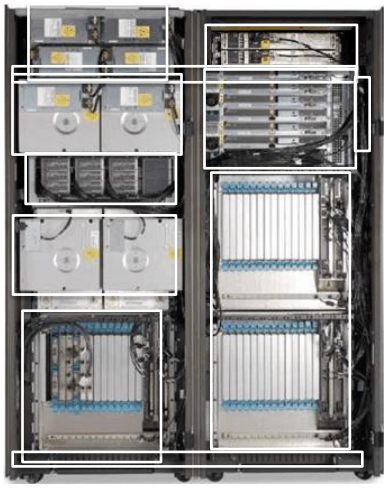
IBM zEnterprize 196: Front View - Water Cooled

- ***Internal Battery Feature**:*
  - ***A-Frame**:* Each 2U mounted on the front only
  - ***Z-Frame**:* Each 2U, Internal Batteries Feature (4x total - back to back)
- ***Frame Brake**:* top hat removal for height reduction feature
- ***VIPAL Board**:* A 2U board that from the front side houses:
  - The PPS/FSP and Oscillator cards
- ***Processors Book**:* 1 to 4 Processor Books
- ***Water Circulation Units**:* 2, side-by-side redundant water pumps used to maintain MCM cooling through a closed internal loop.
- ***BPA**:* The 8U Bulk Power Assembly (back to back) which houses up to 3 BPR's, 1 BPC, 1 BPD, 1 BPH, 1 BPF and Line-Cord. (3 BPR´s are shown: standard + two for 3 phase feature)
- ***BPH-A:*** Hub which provides internal communication within the system elements
- ***I/O Cages and/or Drawers**:* 14U (Cargo Cage) or 5U (Cayuga Drawer) which contain the I/O, STI-MP cards and DCAs
- ***SE Tray**:* Houses the primary and alternate Support Elements
- ***Tailgate Asm.:*** For possible fiber trunk system

Redbooks IBM

IBM zEnterprize 196: Rear View - Air Cooled

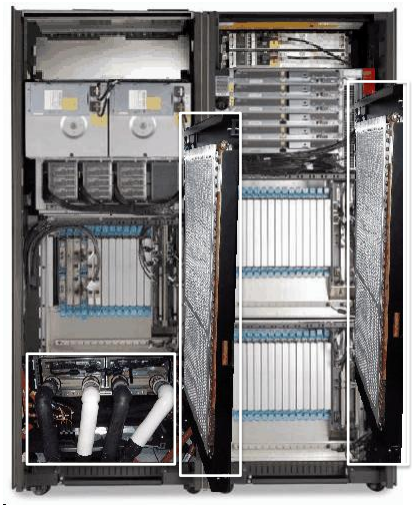
- ***Internal Battery Feature:*** Each 2U, mounted on the front and rear,.1 or 2 per side if feature is installed.
- ***Frame Break****:* top hat removal for height reduction feature
- ***BPA****:* The 8U Bulk Power Assembly (back to back) which houses up to 6 BPR's, 1 BPC, 2 BPD's, 1 BPF and 2 Line Cords
- ***BPH-B:*** Hub which provides internal communication within the system elements
- ***I/O cages and/or drawers****:* 14U (Cage) or 5U (Drawer) which contain the I/O, STI-MP cards and DCAs
- ***Backup Blowers****:* The 1 or 2 blowers which run when necessary to keep the MCM at optimal temperature
- ***CEC Primary Blowers****:* The 2, side-by-side, blowers providing cooling for the processor book components
- ***Book DCA's****:* 3 per book provide power to the books N+1
- ***MRU Blowers****:* The 1 or 2, side-by-side blowers which provide cooling for the 7U Modular Refrigeration Units
- ***Tailgate Asm.*** *:* For possible fiber trunk system



Redbooks IBM

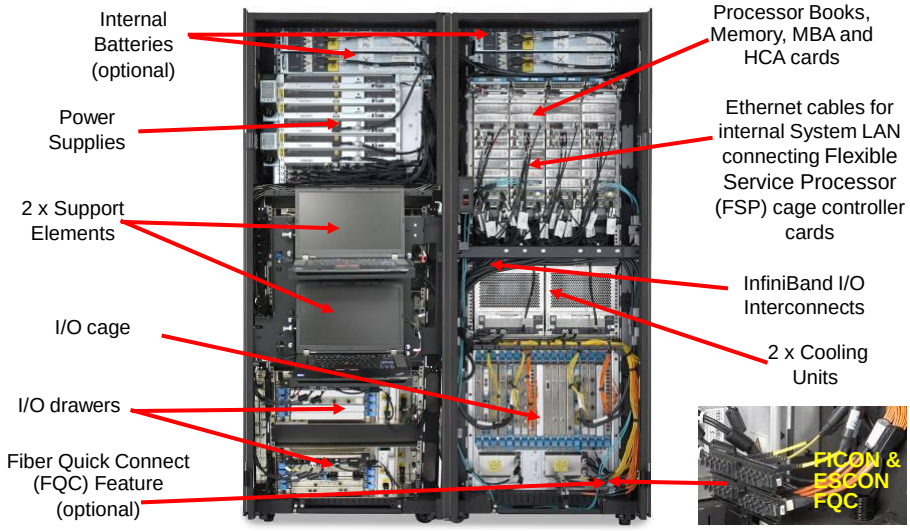
IBM zEnterprize 196: Rear View - Water Cooled

- ***Heat Exchanger Units:*** Cool the hot air blown out the rear of the frames and cool the circulating water in backup mode.
- ***Internal Battery Feature:*** Each 2U, mounted on the front and rear,.1 or 2 per side if feature is installed.
- ***Frame Break****:* top hat removal for height reduction feature
- ***BPA****:* The 8U Bulk Power Assembly (back to back) which houses up to 6 BPR's, 1 BPC, 2 BPD's, 1 BPF and 2 Line Cords
- ***BPH-B:*** Hub which provides internal communication within the system elements
- ***I/O cages and/or drawers****:* 14U (Cage) or 5U (Drawer) which contain the I/O, STI-MP cards and DCAs
- ***CEC Primary Blowers****:* The 2, side-by-side, blowers providing cooling for the processor book components
- ***Book DCA's****:* 3 per book provide power to the books N+1
- ***Water Circulation Units****:* 2, side-by-side redundant water pumps used to maintain MCM cooling through a closed internal loop.
- ***Tailgate Asm.*** *:* For possible fiber trunk system



Redbooks IBM

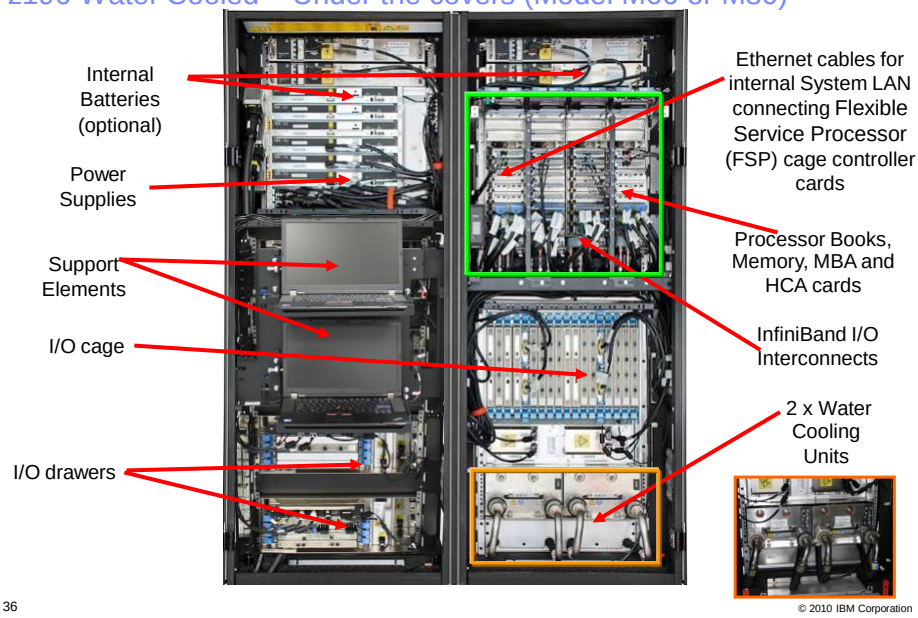
Z196 Air Cooled – Under the covers (Model M66 or M80)



35

Redbooks IBM

z196 Water Cooled – Under the covers (Model M66 or M80)

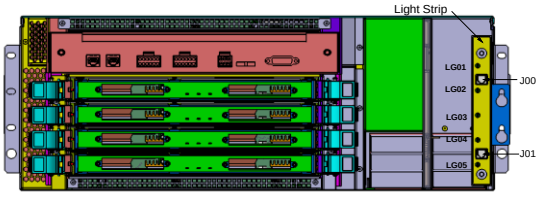


36

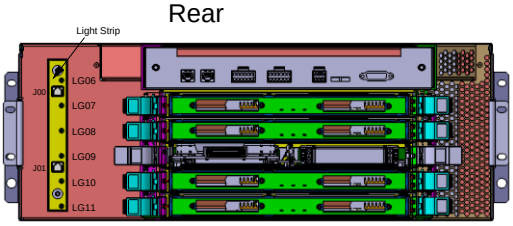
Redbooks IBM

z196 I/O Drawer

Front



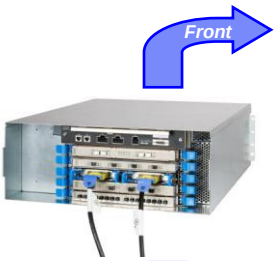
Rear



- The I/O drawer together with logic board and stiffener are FRUs. It can be removed without affecting system input power or power to any other unit. Removing this drawer requires the assistance of the lift tool.
- The blowers are driven by a motor drive card located within the DCA's. There are no external motor drive fru's. Each blower is driven by one DCA motor drive card.
- The FSP's are located in the DCA's.
- The cage has light strips front and rear, each driven by both DCA/FSPs. The PSC24V card is supported in location LG11 of the A01 drawer. This supports two PSC boxes.
- The front and rear lightstrips have a new design

Redbooks IBM

z196 I/O Drawer

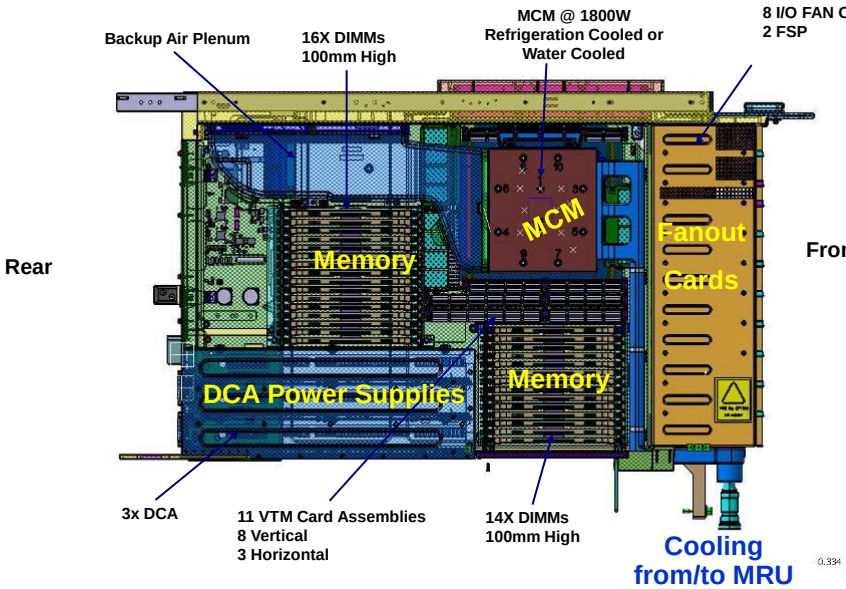


Redbooks IBM

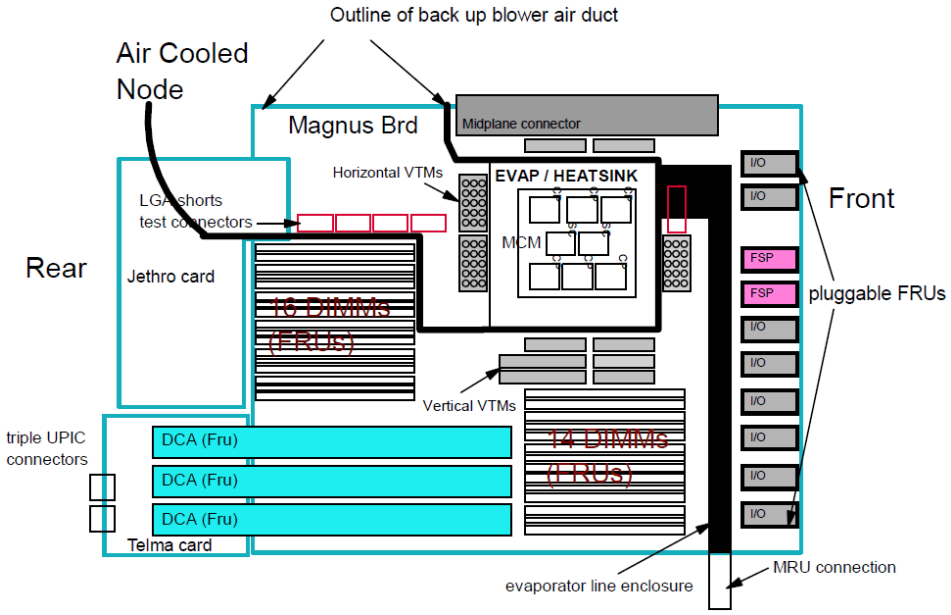
z196 Processor Book

Redbooks IBM

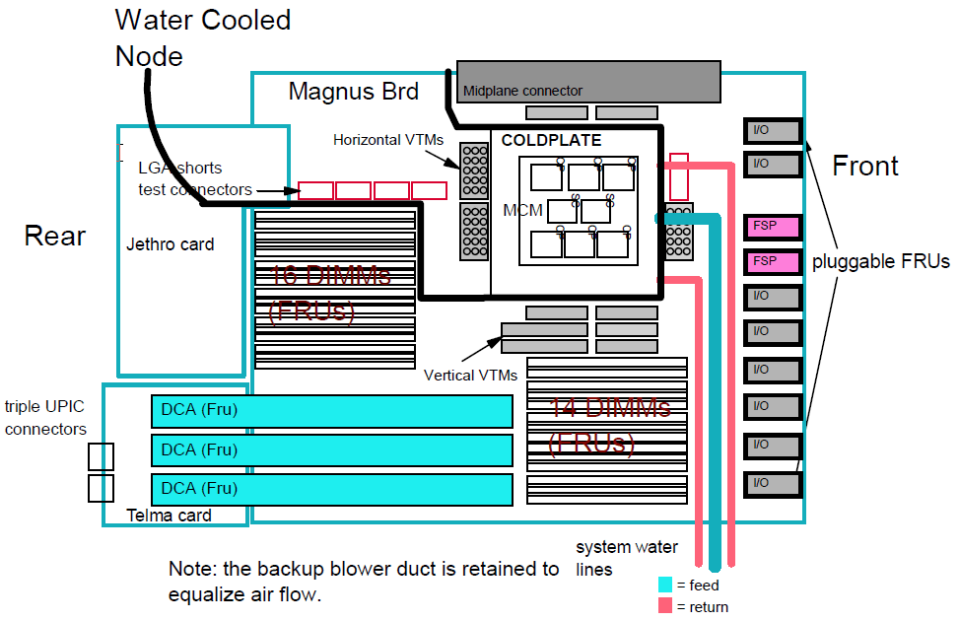
z196 Book Layout



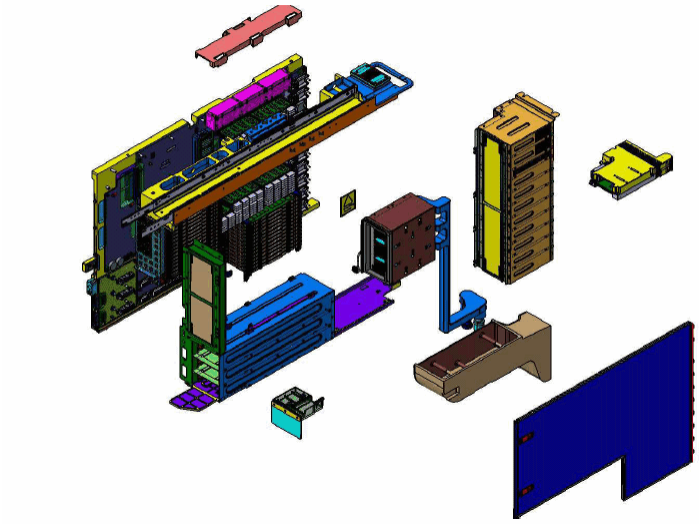
Redbooks IBM



Redbooks IBM

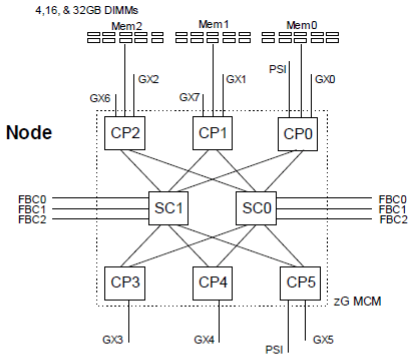


Processor Book – Exploded View



z196 Processor Attributes

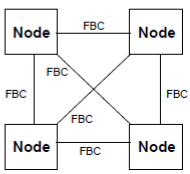
- **4 book System**
  - Fully connected topology
  - 96 Processing Units
  - 12 Memory Controllers
  - Up to 24 I/O Hub and 48 ports
- **up to 3TB Memory capacity**
  - 5-channel Memory RAIM protection
  - DIMM bus CRC error retry
  - 4-level cache hierarchy
  - eDRAM (embedded) caches
  - L3, L4 and SP Key Cache
- **Concurrent Maintenance**
  - Dynamic Book add/repair
- **Support for Galaxy2, IO Hubs**



Chip Info
CP @ 512 sqmm, 1150 $\sqrt{o}$'s, ~250W
SC @ 478 sqmm, 1800 $\sqrt{o}$'s, ~85W

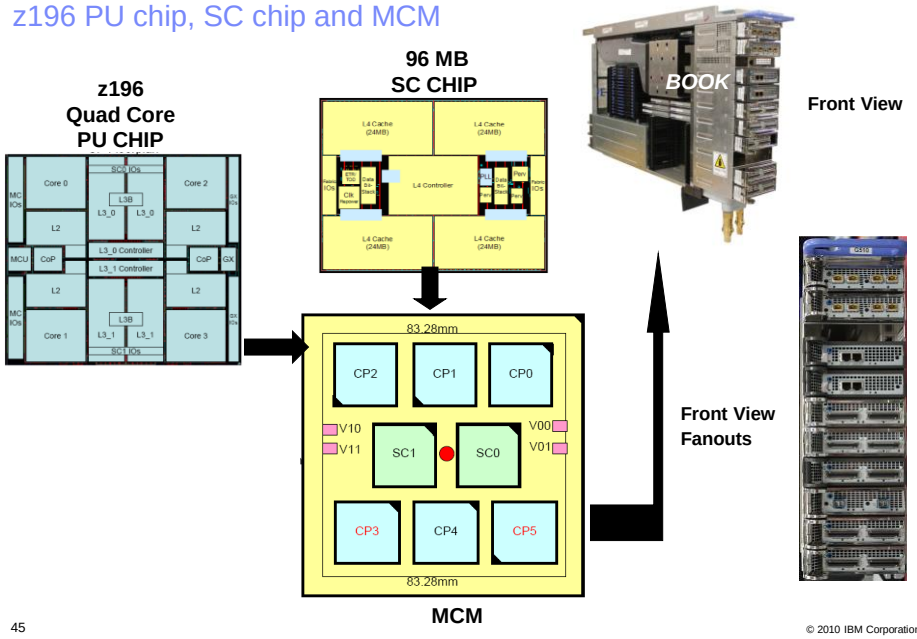
Buses

CP/SC bus	8B@2:1 per dir, EI3
CP/GX bus	4B@~2:1 per dir, Async, EI3
CP/Mem bus	8B@4.8Gb/s per dir, Async, Diff
SC/SC (FBC)	8B@2:1 per dir, EI3



Redbooks IBM

z196 PU chip, SC chip and MCM



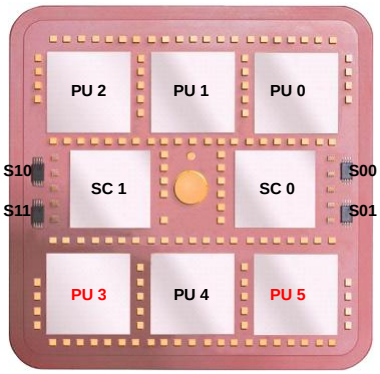
Redbooks IBM

z196 CPC Components

Redbooks IBM

z196 Multi-Chip Module (MCM) Packaging

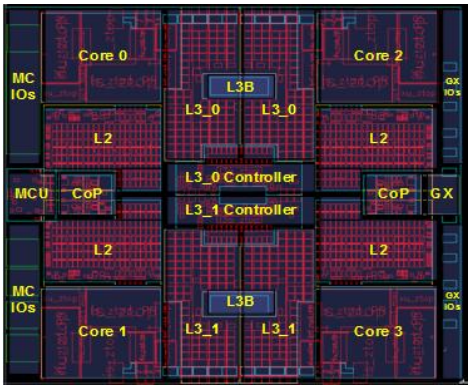
- **96mm x 96mm MCM**
  - 103 Glass Ceramic layers
  - 8 chip sites
  - 7356 LGA connections
  - 20 and 24 way MCMs
  - Maximum power used by MCM is 1800W



- **CMOS 12s chip Technology**
  - PU, SC, S chips, 45 nm
  - 6 PU chips/MCM – Each up to 4 cores
    - One memory control (MC) per PU chip
    - 23.498 mm x 21.797 mm
    - 1.4 billion transistors/PU chip
    - L1 cache/PU core
      - 64 KB I-cache
      - 128 KB D-cache
    - L2 cache/PU core
      - 1.5 MB
    - L3 cache shared by 4 PUs per chip
      - 24 MB
    - 5.2 GHz
  - 2 Storage Control (SC) chip
    - 24.427 mm x 19.604 mm
    - 1.5 billion transistors/SC chip
    - L4 Cache 96 MB per SC chip (192 MB/Book)
    - L4 access to/from other MCMs
  - 4 SEEPROM (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) function is on the SC

Redbooks IBM

z196 Quad Core PU Chip Detail



- **12S0 45nm SOI Technology**
  - 13 layers of metal
  - 3.5 km wire
- **1.4 Billion Transistors**
- **Chip Area – 512.3mm²**
  - 23.5mm x 21.8mm
  - 8093 Power C4's
  - 1134 signal C4's

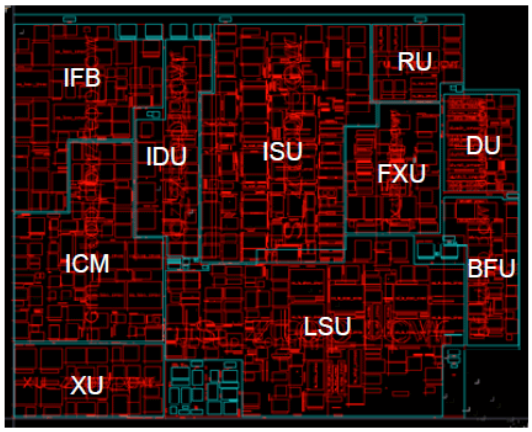
- **Up to Four active cores per chip**
  - 5.2 GHz
  - L1 cache/ core
    - 64 KB I-cache
    - 128 KB D-cache
  - 1.5 MB private L2 cache/ core
- **Two Co-processors (COP)**
  - **Crypto & compression accelerators**
  - Includes 16KB cache
  - Shared by two cores
- **24MB eDRAM L3 Cache**
  - Shared by all four cores
- **Interface to SC chip / L4 cache**
  - 41.6 GB/sec to each of 2 SCs
- **I/O Bus Controller (GX)**
  - Interface to Host Channel Adapter (HCA)
- **Memory Controller (MC)**
  - Interface to controller on memory DIMMs
  - Supports RAIM design

Redbooks IBM

PU Floor plan

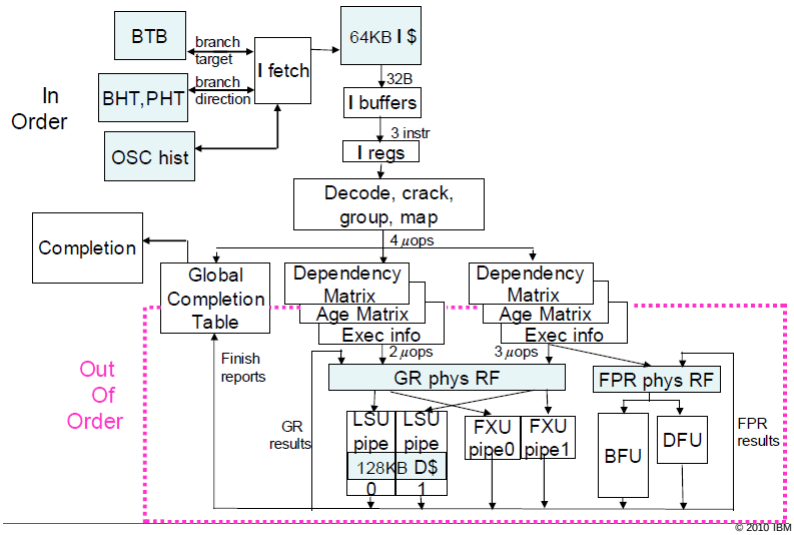
- **IFB** - Instruction fetch and branch (I Buffers)
- **ICM** - Instruction cache & merge (64K I Cache)
- **IDU** - Instruction Decode
- **ISU** - Issue (OOO)
- **XU** - TLB, DAT(Address Translation, I/O Interrupt ctrl)
- **LSU** - Load/Store (128K D Cache)
- **FXU** - Fixed Point (Ex.)
- **BFU** - Binary/ floating point (Ex.)
- **DU** (DFU) - Decimal and Floating point Ops. (Ex.),
- **RU** – Checkpoint and retry controls

Core Floorplan



Redbooks IBM

Processor Structure



Redbooks IBM

z196 PU core

- **Each core is a superscalar, out of order processor with these characteristics:**
  - Six execution units
    - 2 fixed point (integer), 2 load/store, 1 binary floating point, 1 decimal floating point
  - Up to three instructions decoded per cycle (vs. 2 in z10)
  - 211 complex instructions cracked into multiple internal operations
    - 246 of the most complex z/Architecture instructions are implemented via millicode
  - Up to five instructions/operations executed per cycle (vs. 2 in z10)
  - Execution can occur out of (program) order
    - Memory address generation and memory accesses can occur out of (program) order
    - Special circuitry to make execution and memory accesses appear in order to software
  - Each core has 3 private caches
    - 64KB 1st level cache for instructions, 128KB 1st level cache of data
    - 1.5MB L2 cache containing both instructions and data

Redbooks IBM

z196 New CP Architecture

- **New trunc and OR inexactness Binary Floating Point rounding mode**
- **New Decimal Floating Point quantum exception**
  - Eliminates need for test data group for every operation
- **Virtual Architecture Level**
  - Allows the z/VM Live Guest Relocation Facility to make a z196 behave architecturally like a z10 system
  - Facilitates moving work transparently between z196 and z10 systems for backup and capacity reasons
- **On Non-quiescing SSKE:**
  - Significant performance improvement for systems with large number CPUs (typically 30+)
  - Improves SMP scaling for OS images
  - Up to 10% performance increase when exploited by the operating system
  - Exploited by all z/OS 1.10 and above (with PTF for 1.10 and 1.11)
- **Other minor architecture**
  - RRBM, Fast-BCR-Serialization Facility, Fetch-Store-Access Exception Indicator, CMPSC Enhancement Facility

Redbooks IBM

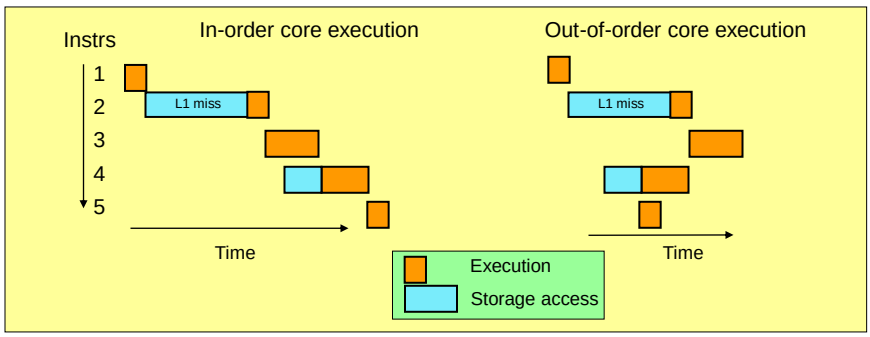
z196 New Instruction Set Architecture

- **Re-compiled code/apps get further performance gains through 100+ new instructions**
- **High-Word Facility (30 new instructions)**
  - Independent addressing to high word of 64-bit GPRs
  - Effectively provides compiler/ software with 16 additional 32-bit registers
- **Interlocked-Access Facility (12 new instructions)**
  - Interlocked (atomic) load, value update and store operation in a single instruction
  - Immediate exploitation by Java
- **Load/Store-on-Condition Facility (6 new instructions)**
  - Load or store conditionally executed based on condition code
  - Dramatic improvement in certain codes with highly unpredictable branches
- **Distinct-Operands Facility (22 new instructions)**
  - Independent specification of result register (different than either source register)
  - Reduces register value copying
- **Population-Count Facility (1 new instruction)**
  - Hardware implementation of bit counting ~5x faster than prior software implementations
- **Integer to/from Floating point converts (21 new instructions)**

Redbooks IBM

z196 Out of Order (OOO) Value

- **OOO yields significant performance benefit for compute intensive apps through**
  - Re-ordering instruction execution
    Later (younger) instructions can execute ahead of an older stalled instruction
  - Re-ordering storage accesses and parallel storage accesses
- **OOO maintains good performance growth for traditional apps**



Redbooks IBM

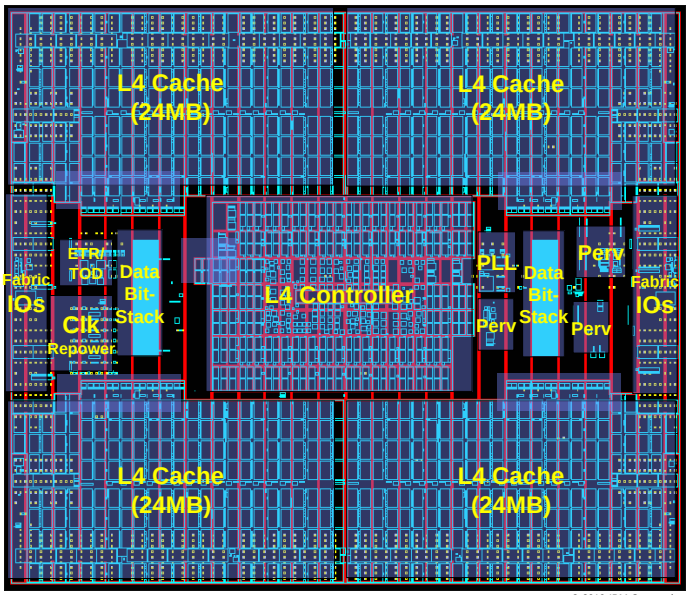
z196 Out of Order Detail

- **Out of order yields significant performance benefit through**
  - Re-ordering instruction execution
    - Instructions stall in a pipeline because they are waiting for results from a previous instruction or the execution resource they require is busy
    - In an in-order core, this stalled instruction stalls all later instructions in the code stream
    - In an out-of-order core, later instructions are allowed to execute ahead of the stalled instruction
  - Re-ordering storage accesses
    - Instructions which access storage can stall because they are waiting on results needed to compute storage address
    - In an in-order core, later instructions are stalled
    - In an out-of-order core, later storage-accessing instructions which can compute their storage address are allowed to execute
  - Hiding storage access latency
    - Many instructions access data from storage
    - Storage accesses can miss the L1 and require 10 to 500 additional cycles to retrieve the storage data
    - In an in-order core, later instructions in the code stream are stalled
    - In an out-of-order core, later instructions which are not dependent on this storage data are allowed to execute

Redbooks IBM

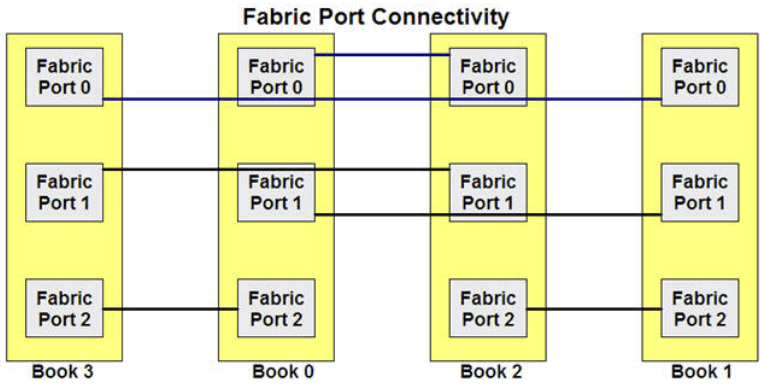
z196 SC Chip Detail

- **12S0 45nm SOI Technology**
  - 13 layers of metal
- **Chip Area – 478.8mm^2**
  - 24.4mm x 19.6mm
  - 7100 Power C4's
  - 1819 signal C4's
- **1.5 Billion Transistors**
  - 1 Billion cells for eDRAM
- **eDRAM Shared L4 Cache**
  - 96 MB per SC chip
  - 192 MB per Book
- **6 CP chip interfaces**
- **3 Fabric interfaces**
- **2 clock domains**
- **5 unique chip voltage supplies**



Redbooks IBM

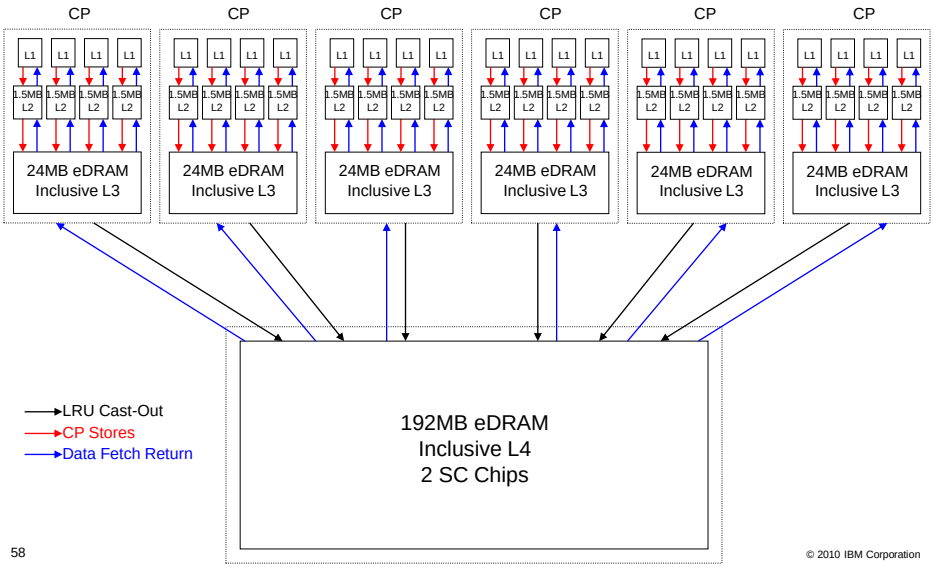
Fabric Port Connectivity



The fabric ports provide book to book communications
Part of the SC chips

Redbooks IBM

z196 Book Level Cache Hierarchy



Redbooks IBM

z196 PU Allocation

- For models M15, M32, M49, M66, **(**20 PU Books) the processor units on the MCM in each book are implemented with a mix of four PU chips with three active cores (PU0, PU1, PU2, PU4) and two PU chips with four active cores (PU3, PU5), 20 active cores per MCM.
- For the model M80, each processor chip has four active cores, resulting in 24 active-cores per MCM. This means that there are 96 active cores on model M80
- SPARES - Book 1 and 3 get assigned one spare each on the high chip, core 1, then core 0. If not available then look at the next highest chip, core 1, then core 0
- SAPs - Spread across books and high chips. Start with high chip high core, then next highest chip high core. This prevents all the SAPs from being assigned on one chip.
- CP's - Fill chip and spill into next chip on low order book first before spilling over into next book. Exp: 3 book system: 1st CP assigned to 20, then 21, 22, 23, then 24, 25, 26, 27, Then 28, 29 etc..The CPs are assigned from low to high clustered by 4
- CF (Coupling) - Fill high chip on high book
- IFL (Linux) - Fill high chip on high book
- zAAP (IFA) - Attempts are made to align these close the CP's
- zIIP - Attempts are made to align these close the CP's

Redbooks IBM

Assign resources to processors

- LPAR can Address 80 CPs.with CP numbering from 0 to 95
- For performance reasons, CPs for a logical partition are grouped together as much as possible.
- Having all CPs grouped in as few books as possible limits memory and cache interference to a minimum

	0PU0x (x=0-3)		0PU1x (x=0-3)		0PU2x (x=0-3)		0PU3x (x=0-3)		0PU4x (x=0-3)		0PU5x (x=0-3)	
BOOK0	00	01	04	05	08	09	0C	0D	10	11	14	15
	02	03	06	07	0A	0B	0E	0F	12	13	16	17

	1PU0x (x=0-3)		1PU1x (x=0-3)		1PU2x (x=0-3)		1PU3x (x=0-3)		1PU4x (x=0-3)		1PU5x (x=0-3)	
BOOK1	20	21	24	25	28	29	2C	2D	30	31	34	35
	22	23	26	27	2A	2B	2E	2F	32	33	36	37

	2PU0x (x=0-3)		2PU1x (x=0-3)		2PU2x (x=0-3)		2PU3x (x=0-3)		2PU4x (x=0-3)		2PU5x (x=0-3)	
BOOK2	40	41	44	45	48	49	4C	4D	50	51	54	55
	42	43	46	47	4A	4B	4E	4F	52	53	56	57

	3PU0x (x=0-3)		3PU1x (x=0-3)		3PU2x (x=0-3)		3PU3x (x=0-3)		3PU4x (x=0-3)		3PU5x (x=0-3)	
BOOK3	60	61	64	65	68	69	6C	6D	70	71	74	75
	62	63	66	67	6A	6B	6E	6F	72	73	76	77

Note: The hexadecimal physical PU numbers in the boxes are in the form that one would obtain from the PU Configuration Register [R-unit MCR8B(57:63), where 57:58 is the book#, 59:61 is the chip#, and 62:63 is the processor# (also known as the core#)].

Redbooks IBM

z196 – M49 / 716 + 6 ICFs + 4 zAAPs + 2 IFLs and 4 zIIPs

	Total CPUs	Total SAPs	Ratio	Node 1 PUs	Avail	SAPs	Spares	Node 3 PUs	Avail	SAPs	Spares	Node 2 PUs	Avail	SAPs	Spares	Node 0 PUs	Avail	SAPs	Spares
m15	15	3	5.0	20	15	3	2												
m32	32	6	5.3	20	16	3	1	20	16	3	1								
m49	49	9	5.4	20	16	3	1	20	16	3	1	20	17	3	0				
m66	66	12	5.5	20	16	3	1	20	16	3	1	20	17	3	0	20	17	3	0
m80	80	14	5.7	24	20	3	1	24	20	3	1	24	20	4	0	24	20	4	0

	PU Chip 0		PU Chip 1		PU Chip 2		PU Chip 3		PU Chip 4		PU Chip 5	
Book 0 LG01	PU00	PU01	PU04	PU05	PU08	PU09	PU0C	PU0D	PU10	PU11	PU14	PU15
	PU02	PU03	PU06	PU07	PU0A	PU0B	PU0E	PU0F	PU12	PU13	PU16	PU17
Book 1 LG06	CP1		CP4	CP5		CP7	CP10	CP11	CP13		CP15	Spare
	CP2	CP3	CP6		CP8	CP9	SAP3	CP12	CP14	SAP2	CP16	SAP1
Book 2 LG10	zAA1	zAA2	zII2	PU45		PU49	PU4C	PU4D	PU50	PU51	PU54	PU55
	zII1		PU46		PU4A	PU4B	PU4E	SAP6		SAP5	PU56	SAP4
Book 3 LG15	PU60	PU61	PU64	PU65	PU68	PU69	PU6C	PU6D		IFL1	ICF2	Spare
		PU63	PU66		PU6A		PU6E	SAP9	IFL2	SAP8	ICF2	SAP7

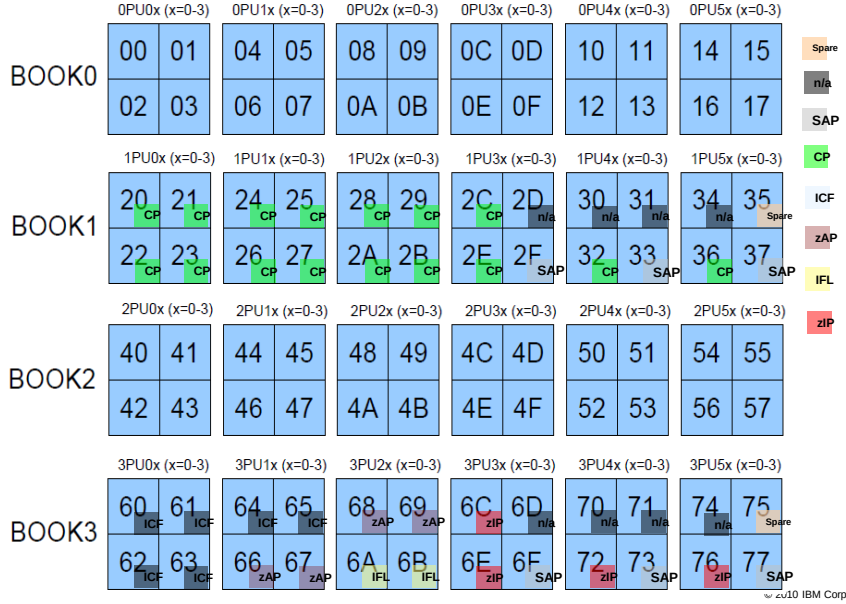
Redbooks IBM

z196 PU arrangement

	0PU0x (x=0-3)		0PU1x (x=0-3)		0PU2x (x=0-3)		0PU3x (x=0-3)		0PU4x (x=0-3)		0PU5x (x=0-3)	
BOOK0	00	01	04	05	08	09	0C	0D	10	11	14	15
	02	03	06	07	0A	0B	0E	0F	12	13	16	17
	1PU0x (x=0-3)		1PU1x (x=0-3)		1PU2x (x=0-3)		1PU3x (x=0-3)		1PU4x (x=0-3)		1PU5x (x=0-3)	
BOOK1	20	21	24	25	28	29	2C	2D	30	31	34	35
	22	23	26	27	2A	2B	2E	2F	32	33	36	37
	2PU0x (x=0-3)		2PU1x (x=0-3)		2PU2x (x=0-3)		2PU3x (x=0-3)		2PU4x (x=0-3)		2PU5x (x=0-3)	
BOOK2	40	41	44	45	48	49	4C	4D	50	51	54	55
	42	43	46	47	4A	4B	4E	4F	52	53	56	57
	3PU0x (x=0-3)		3PU1x (x=0-3)		3PU2x (x=0-3)		3PU3x (x=0-3)		3PU4x (x=0-3)		3PU5x (x=0-3)	
BOOK3	60	61	64	65	68	69	6C	6D	70	71	74	75
	62	63	66	67	6A	6B	6E	6F	72	73	76	77

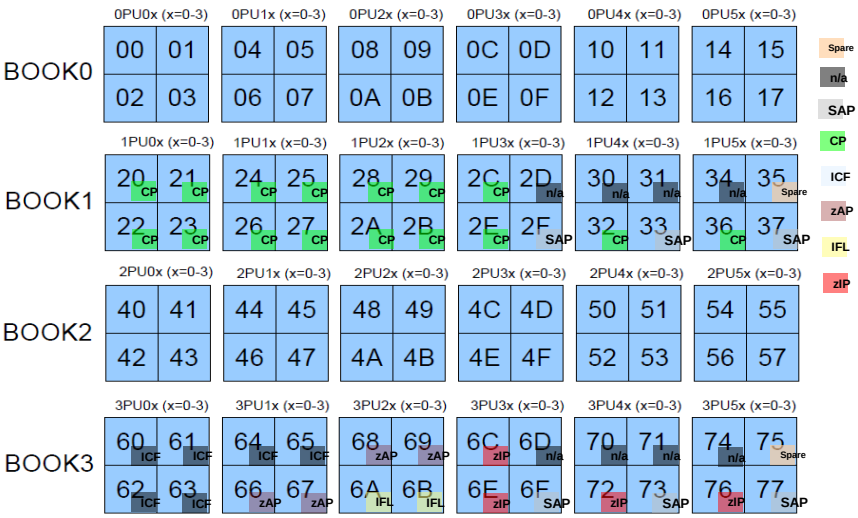
Redbooks IBM

z196 – M32 / 716 + 6 ICFs + 4 zAAPs + 2 IFLs and 4 zIIPs



Redbooks IBM

z196 – M32 / 716 + 6 ICFs + 4 zAAPs + 2 IFLs and 4 zIIPs



Redbooks IBM

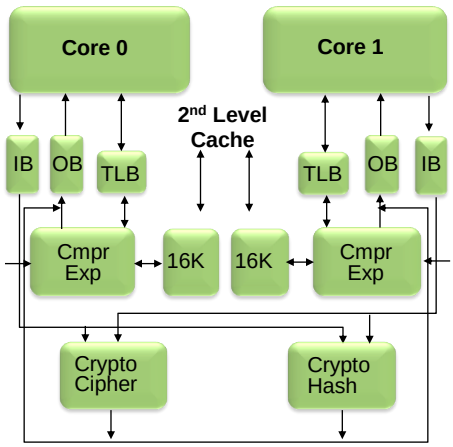
PU assignment considerations

- LPAR will always remap the PU assignments when the images activate and can swap PU's with a spare within the same book.
- LPAR can optimize PU's every 50 sec's
- WLM can changes weights, capping and park/unpark PU's every 60 sec's.
- CPM (Capacity Provisioning Manager) can configure new processors online via OOCoD.
- Change default time is set at 15 minutes, however this is a PARM in CPM and can be made shorter or longer by the customer. The recommended range is a value between 5 & 15 minutes.
- After a CBR storage assignments are not moved back into the CBR'd book (Same as z10).

Redbooks IBM

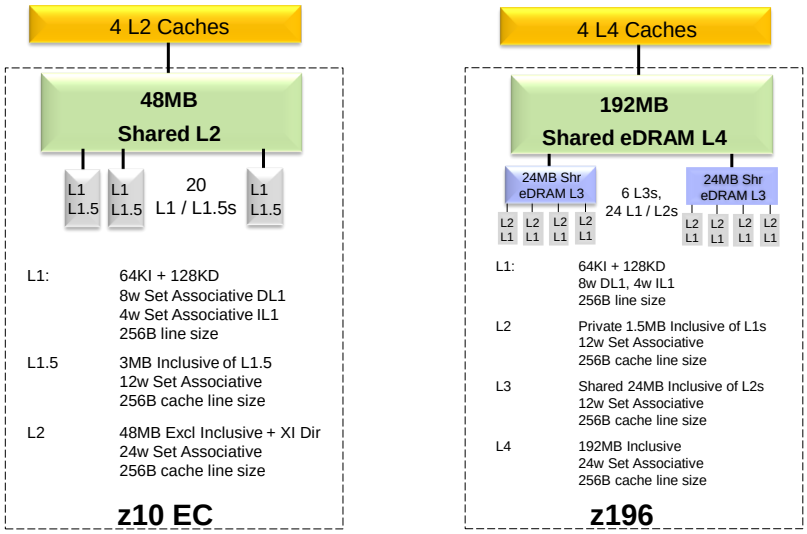
z196 Compression and Cryptography Accelerator

- **Data compression engine**
  - Static dictionary compression and expansion
  - Dictionary size up to 64KB (8K entries)
    - Local 16KB cache per core for dictionary data
- **CP Assist for Cryptographic Function (CPACF)**
  - Enhancements for new NIST standard
  - Complemented prior ECB and CBC symmetric cipher modes with XTS, OFB, CTR, CFB, CMAC and CCM
  - New primitives (128b Galois Field multiply) for GCM
- **Accelerator unit shared by 2 cores**
  - Independent compression engines
  - Shared cryptography engines



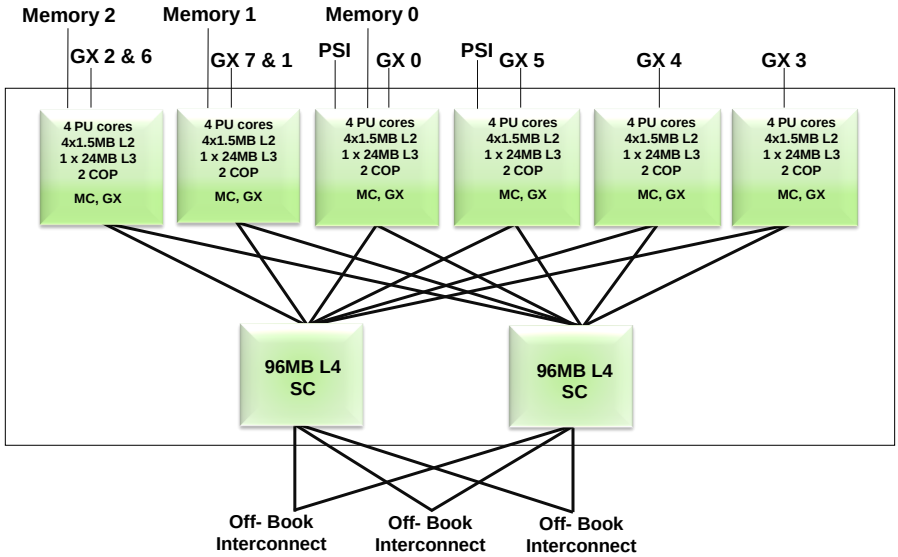
Redbooks IBM

System z Cache Topology – z10 EC vs. z196 Comparison



Redbooks IBM

z196 24 PU MCM Structure



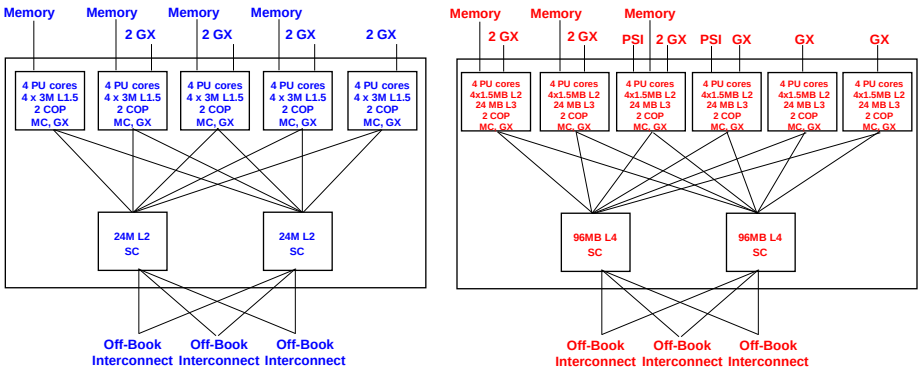
Redbooks IBM

z10 EC vs. z196 CPC Structure

	z10 EC	z196
SMP Configuration	4 books, 77 PU cores	4 books, 96 PU cores
Topology	Fully Connected	Fully Connected
Max Memory	1.5TB	3.0TB
Cache Levels	L1, L1.5, L2	L1, L2, L3, L4

77 PU core System

96 PU core System



Redbooks IBM

z10 EC MCM vs. z196 MCM Comparison

z10 EC MCM

- **MCM**
  - 96mm x 96mm in size
  - 5 PU chips per MCM
    - Quad core chips with 3 or 4 active cores
    - PU Chip size 21.97 mm x 21.17 mm
    - 4.4 GHz
    - Superscalar, In order execution
    - L1: 64K I /128K D private/core
    - L1.5: 3M I+D private/core
  - 2 SC chips per MCM
    - L2: 2 x 24 M = 48 M L2 per book
    - SC Chip size 21.11 mm x 21.71 mm
  - Power 1800 Watts

z196 MCM

- **MCM**
  - 96mm x 96mm in size
  - 6 PU chips per MCM
    - Quad core chips with 3 or 4 active cores
    - PU Chip size 23.5 mm x 21.8 mm
    - 5.2 GHz
    - Superscalar, OOO execution
    - L1: 64K I / 128K D private/core
    - L2: 1.5M I+D private/core
    - L3: 24MB/chip - shared
  - 2 SC chips per MCM
    - L4: 2 x 96 MB = 192 MB L4 per book
    - SC Chip size 24.4 mm x 219.6 mm
  - Power 1800 Watts

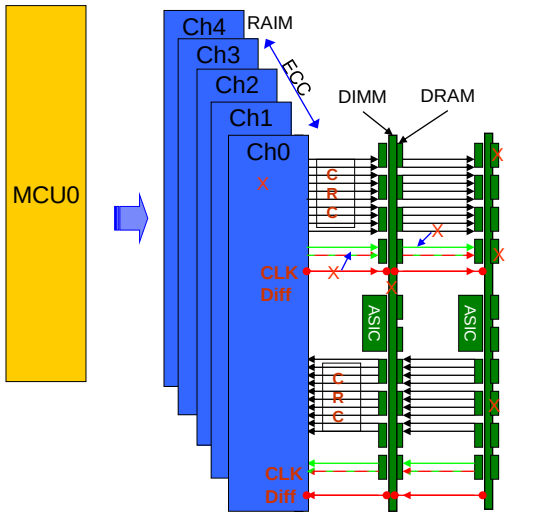
Redbooks IBM

z196 Memory Design

- **System z10 EC memory design:**
  - Four Memory Controllers (MCUs) organized in two pairs, each MCU with four channels
  - DIMM technology is Nova x4, 16 to 48 DIMMs per book, plugged in groups of 8
  - 8 DIMMs (4 or 8 GB) per feature – 32 or 64 GB physical memory per feature
    Equals 32 or 64 GB for HSA and customer purchase per feature
  - 64 to 384 GB physical memory per book = 64 to 384 GB for use (HSA and customer)
- **z196 memory design:**
  - Three MCUs, each with five channels. The fifth channel in each z196 MCU is required to implement memory as a Redundant Array of Independent Memory (RAIM). This technology adds significant error detection and correction capabilities. Bit, lane, DRAM, DIMM, socket, and complete memory channel failures can be detected and corrected, including many types of multiple failures.
  - DIMM technology is SuperNova x81, 10 to 30 DIMMs per book, plugged in groups of 5
    5 DIMMs (4, 16 or 32 GB) per feature – 20, 80 or 160 GB physical RAIM per feature
    Equals 16, 64 or 128 GB for use per feature. RAIM takes 20%. (There is no non-RAIM option.)
  - 40 to 960 GB RAIM memory per book = 32 to 768 GB of memory for use
    (Minimum RAIM for the M15 is 60 GB = 48 GB = 16 GB HSA plus 32 GB customer memory)
- **For both z196 and z10**
  - The Hardware System Area (HSA) is 16 GB fixed, outside customer memory
  - In some cases, offering granularity can prevent purchase of all available memory in a book

Redbooks IBM

z196 RAIM Memory Controller Overview



2- Deep Cascade
Using Quad High DIMMs

Layers of Memory Recovery

ECC
- Powerful 90B/64B Reed Solomon code

DRAM Failure
- Marking technology; no half sparing needed
- 2 DRAM can be marked
- Call for replacement on third DRAM

Lane Failure
- CRC with Retry
- Data – lane sparing
- CLK – RAIM with lane sparing

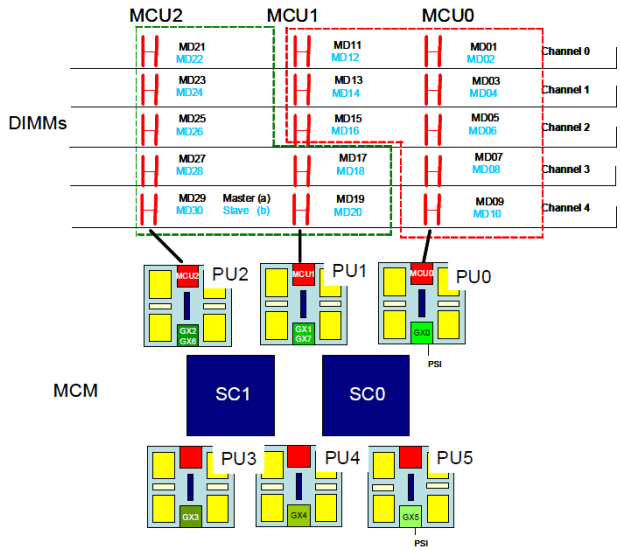
DIMM Failure (discrete components, VTT Reg.)
- CRC with Retry
- Data – lane sparing
- CLK – RAIM with lane sparing

DIMM Controller ASIC Failure
- RAIM Recovery

Channel Failure
- RAIM Recovery

Redbooks IBM

SuperNOVA Main Memory Sub-system Overview



Redbooks IBM

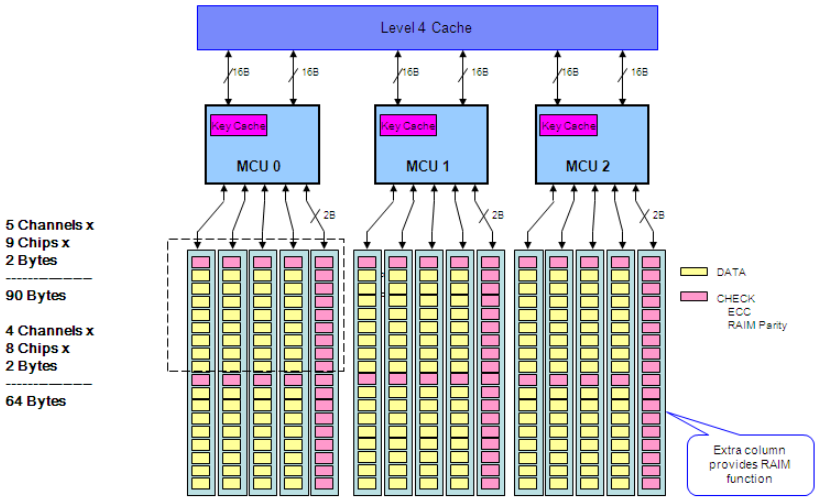
SuperNOVA Main Memory Sub-system Overview…..

- 3 MCUs / Book, 2-deep cascade, 5-channels
  - Minimum plugged 2 MCUs with 1 cascade - 10 DIMMs
- Differential Memory Interface (DMI)
  - 800Mhz bus speed, CRC, Retry, and dynamic lane sparing
- "RAIM" redundancy implementation
  - 5th. channel added for check bits, can survive catastrophic DIMM failure
  - Order process impacted by 20% more memory
- Up to 30 DIMMs per Book
  - 4/16/32 GB DDR3
  - max 0.75TB/Book
  - 3TB/system
  - Min plug 10
  - Increment 5
- No mixing of DIMM sizes on a Book

Redbooks IBM

z196 RAIM Memory Structure

RAIM Memory Structure



Redbooks IBM

z196 Purchase Memory Offerings

- **Purchase Memory – Memory available for assignment to LPARs**
- **Hardware System Area – Standard 16 GB outside customer memory for system use**
- **Standard Memory – Provides minimum physical memory required to hold base purchase memory plus 16 GB HSA**
- **Flexible Memory – Provides additional physical memory needed to support activation base customer memory and HSA on a multiple book z196 with one book out of service.**
- **Plan Ahead Memory – Provides additional physical memory needed for a concurrent upgrade (LIC CC change only) to a preplanned target customer memory**

Redbooks IBM

z196 Memory Offerings and Features

Model	Standard Memory GB	Flexible Memory GB
M15	32 - 704	NA
M32	32 - 1520	32 - 704
M49	32 - 2288	32 - 1520
M66	32 - 3056	32 - 2288
M80	32 - 3056	32 - 2288

RAIM Physical Memory Math:

20 GB RAIM = 16 GB usable
80 GB RAIM = 64 GB usable
160 GB RAIM = 128 GB usable
Orderable = Usable less 16 GB HSA

- **Memory Granularity for ordering:**
  - 32 GB: Std - 32 to 256; Flex - 32 to 256
  - 64 GB: Std - 320 to 512; Flex - 320 to 512
  - 96 GB: Std - 608 to 896; Flex - 608 to 896
  - 112 GB: Std - 1008; Flex - 1008
  - 128 GB: Std - 1136 to 1520; Flex - 1136 to 1520
  - 256 GB: Std - 1776 to 3056; Flex – 1776 to 2288
- **16 GB separate fixed HSA standard**
- **Maximum Physical Memory:**
  960 GB RAIM per book, 3.75 TB RAIM per system
  Redundant Array of Independent Memory (RAIM) design detects and recovers from DRAM, socket, memory channel or DIMM failures. Extra physical memory required for RAIM is not available for HSA or purchase.
  - Up to 30 DIMMs per book (no mixing of DIMM sizes)
  - 40 GB RAIM minimum physical memory in each book
  - Physical Memory Increments and Features:
    20 GB RAIM – Five 4 GB DIMMs (FC #1640 or #1641)
    Preferred if can fulfill purchase memory
    80 GB RAIM – Five 16 GB DIMMs (FC #1616 or #1617)
    Used where necessary
    160 GB RAIM – Five 32 GB DIMMs (FC #1632 or #6133)
    Used where necessary
- **For Plan Ahead or Flexible Memory, if required, 16 GB "Pre-planned Memory" features (FC # 1996) are added to the configuration.**

Redbooks IBM

z196 Flexible Memory

- **Provides additional physical memory needed to support activation all customer purchased memory and HSA on a multiple book z196 with one book down for**
  - Scheduled concurrent book upgrade (e.g. memory)
  - Scheduled concurrent maintenance
  - Concurrent repair of a book "fenced" during Activation (POR)
  - Note: All of the above can be done without Flexible Memory; but, all customer purchased memory will not be available for use in most cases. Some work may have to be shut down or not restarted.
- **Offered on M32, M49, M66 and M80 in:**
  - 32 GB increments from 32 GB to 256 GB
  - 64 GB increments from 320 GB to 512 GB
  - 96 GB increments from 608 GB to 896 GB (M32 limit 704 GB)
  - 112 GB increment to 1008 GB
  - 128 GB increments from 1136 GB to 1520 GB (M49 limit 1520 GB)
  - 256 GB increments from 1776 GB to 2288 GB
- **Selected by checking the "Flexible" box when configuring memory**
- **Additional physical memory, if required, is added to the configuration and priced as "Plan Ahead Memory"**

Redbooks IBM

z196 Standard and Flexible Memory Offerings

Increment	GB, Notes	Growth %	Increment	GB, Notes	Growth %	Increment	GB, Notes	Growth %
32 GB	32	100%	96 GB	608	16%	256 GB	1776	17%
	64	50%		704[1]	13%		2032	13%
	96	33%		800	12%		2288[3]	11%
	128	25%		896	12%		2544	10%
	160	20%					2800	9%
	192	17%	112 GB	1008	13%		3056[4]	NA
	224	14%						
	256	25%						
64 GB	320	20%	128 GB	1136	11%			
	384	17%		1264	10%			
	448	14%		1392	9%			
	512	19%		1520[2]	17%			

Notes – Memory Maximums:

1. M15 Standard, M32 Flexible = 704
2. M32 Standard, M49 Flexible, (z10 EC Standard)= 1520
3. M49 Standard, M66 and M80 Flexible = 2288
4. M66 and M80 Standard = 3056

Redbooks IBM

z196 Plan Ahead Memory

- **Provides the capability for concurrent memory upgrades without exploitation of Enhanced Book Availability, Licensed Internal Code upgrades**
  - Memory cards are pre-installed to support target Plan Ahead capacity
  - Available on all z196 models
  - Can be ordered with standard memory on any z196 model (Standard plus Plan Ahead will NOT be Flexible in most cases.)
  - Can be ordered with Flexible memory on a multiple book z196 model
- **Pre-planned memory features are chargeable**
  - Charge to install **ADDITIONAL** memory hardware needed to enable the selected plan ahead target. (Additional means larger than the hardware needed for base memory)
  - FC #1996 – One feature for each 16 GB (20 GB RAIM) of additional hardware added
- **Pre-planned memory activation is chargeable**
  - Subsequent memory upgrade orders will use Plan Ahead Memory first
  - Charged when Plan Ahead Memory is enabled by concurrent LIC upgrade
  - Add FC #1901, Delete FC #1996 – For each 16 GB of memory activated for use
- **Note: Plan Ahead Memory is NOT temporary, On Demand memory**
  Why? Because Memory LIC downgrade is disruptive.

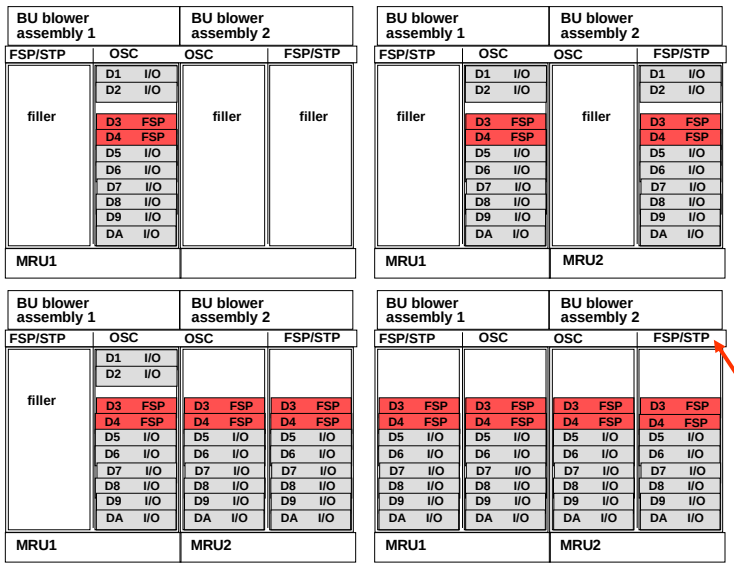
Redbooks IBM

z196 – Three M49 new-build memory examples

- **M49 with 160 GB, no plan ahead**
  - 1640 4 GB Memory DIMM(5/feature) 11 = 220 GB RAIM, 176 GB memory
  - 1901 16 GB Memory Capacity Incr 8 = Pricing feature (128 GB memory, first 32 GB is N/C)
  - 2433 160 GB Memory 1 = Customer use memory (HSA is separate)
  - 160 GB customer memory +16 GB HSA = 176 GB
  - Book placement – 80, 80, 60 RAIM = 64, 64, 48 GB = *Two books: 96 GB + HSA worst case*
- **M49 with 160 GB, Flexible**
  - 1640 4 GB Memory DIMM(5/feature) 17 = 340 GB RAIM, 272 GB memory
  - 1901 16 GB Memory Capacity Incr 8 = Pricing feature (128 GB memory, first 32 GB is N/C)
  - 1996 Preplanned Memory 6 = Pricing feature (96 GB Plan Ahead memory)
  - 2433 160 GB Memory 1 = Customer use memory (HSA is separate)
  - 160 GB customer memory +16 GB HSA + 96 GB Plan Ahead, not active = 272 GB
  - Book placement – 120, 120, 100 RAIM = 96, 96, 80 GB = *Two books: 160 GB + HSA worst case*
- **M49 with 160 GB, Plan Ahead to 224 GB**
  - 1640 4 GB Memory DIMM(5/feature) 15 = 300 GB RAIM, 240 GB memory
  - 1901 16 GB Memory Capacity Incr 8 = Pricing feature (128 GB memory, first 32 GB is N/C)
  - 1996 Preplanned Memory 4 = Pricing feature (64 GB Plan Ahead memory)
  - 2433 160 GB Memory 1 = Customer use memory (HSA is separate)
  - 160 GB customer memory +16 GB HSA + 64 GB Plan Ahead, not active = 240 GB
  - Book placement – 100, 100, 100 RAIM = 80, 80, 80 GB = *Two books = 128 GB + HSA with 16 GB not usable because 144 GB is not a valid memory configuration.*

Redbooks IBM

z196 – Book and HCA plugging order-front view



- **I/O fanouts**
  - M15 – up to 8
  - M32 – up to 16
  - M49 – up to 20
  - M66 – up to 24
  - M80 – up to 24
- **Cooling Design**
  - HCA locations D1 and D2 are NOT used in the 2nd and 3rd Book for Model M49
  - HCA locations D1 and D2 are NOT used in any Book for Models M66 or M80

Redbooks IBM

IBM zEnterprise I/O Architecture

- *z196 I/O Infrastructure*
- *Physical Lay-out*
- *Fanout Cards*
- *z196 PCHID numbering*
- *I/O drawer structure*
- *HCA & MBA plugging*
- *Redundant I/O Interface*
- *z196 new build I/O offerings*
- *z196 I/O Connectivity*
- *zDAC*

Redbooks IBM

Glossary for System z I/O

Acronym	Full name	Comments
AID	Adapter identification	HCA fanout has AID instead of a PCHID
CIB	Coupling using InfiniBand	CHPID type z196, z10, System z9
HCA	Host Channel Adapter	Path for communication
MBA	Memory Bus Adapter	Path for communication
PSIFB	Parallel Sysplex using InfiniBand	InfiniBand Coupling Links
IFB-MP	InfiniBand Multiplexer	I/O drawer/cage intra-connection
STI-MP	Self-Timed Interconnect Multiplexer	I/O drawer/cage intra-connection

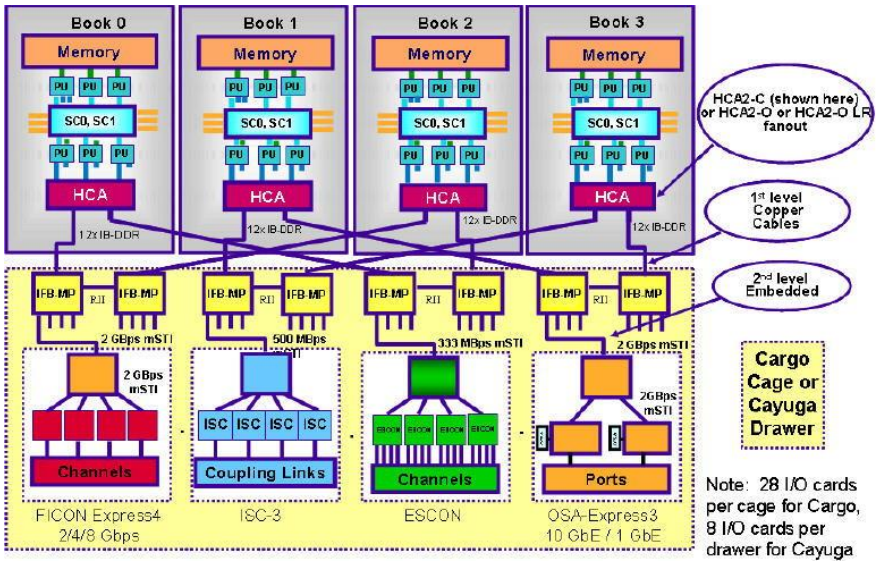
Type	System z10	System z196
HCA1-O fanout	NA	NA
HCA2-C fanout	Copper - I/O drawer/cage - 12x IB-DDR	Copper - I/O drawer/cage - 12x IB-DDR
HCA2-O fanout	Optical - Coupling 12x IB-DDR	Optical - Coupling 12x IB-DDR
HCA2-O LR fanout	Optical - Coupling 1x IB-DDR	Optical - Coupling 1x IB-DDR
MBA fanout	Coupling (ICB-4)	N/A
IFB-MP	I/O drawer/cage intra-connection	I/O drawer/cage intra-connection
STI-MP	NA	NA

Redbooks IBM

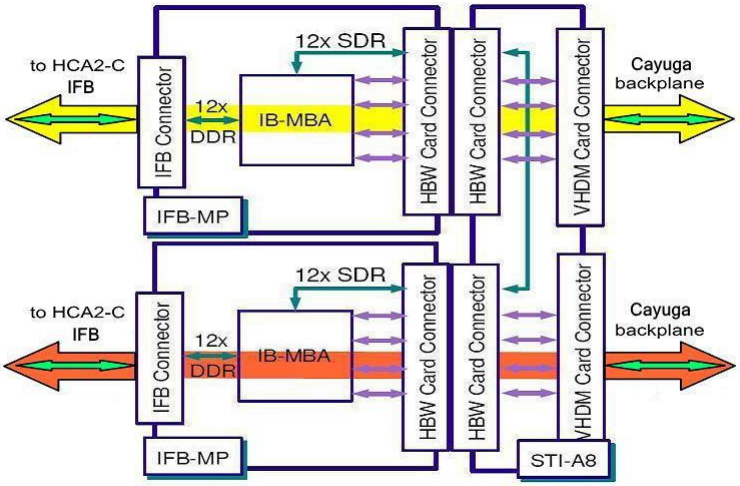
z196 I/O Infrastructure

Redbooks IBM

z196 I/O Infrastructure

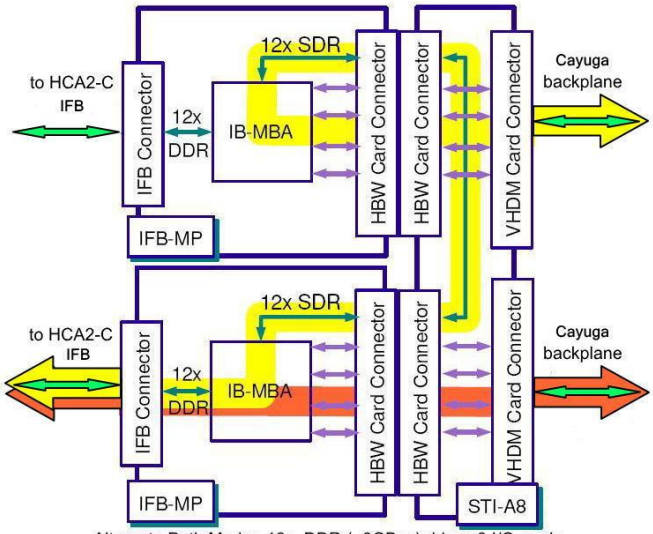


I/O Interconnect Failover Redundancy



Normal Operation : 12x DDR (~6 GBps), drives 4 I/O cards

I/O Failing Path

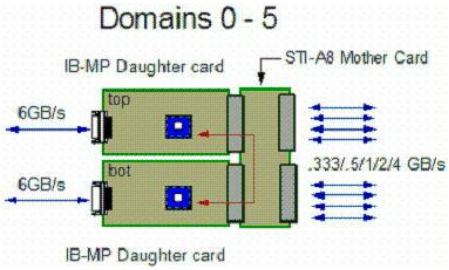


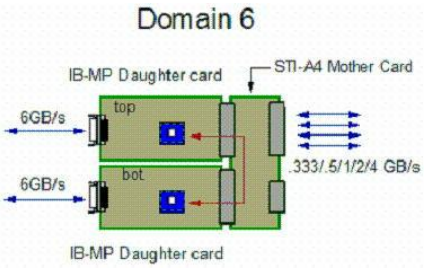
Alternate Path Mode: 12x DDR (~6GBps) drives 8 I/O cards
4 of which are connected via 12x SDR (~3GBps) chaining link

IMPORTANT NOTE: Failing paths are persistent through a POR/IML.

Redbooks IBM

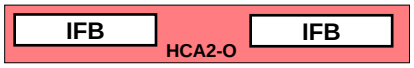
Cargo I/O Cage and the Seventh Domain



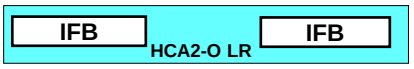


Redbooks IBM

Connectivity for Coupling and I/O



Up to 16 CHPIDs – across 2 ports



Up to 16 CHPIDs – across 2 ports

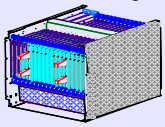
- **Up to 8 fanout cards per book**
  - Up to 16 ports per book
    - 48 Port System Maximum
- **Fanout cards – Two-port InfiniBand host channel adapters dedicated to function**
  - **HCA2-C fanout – I/O Interconnect**
    - Supports FICON, ESCON, OSA, ISC-3 and Crypto Express3 cards in I/O drawer and I/O cage domains. Always plugged in pairs.
  - **HCA2-O fanout – 12x InfiniBand coupling links**
    - CHPID type – CIB for Coupling
      - Fiber optic external coupling link – 150 m
  - **HCA2-O LR fanout – 1x InfiniBand coupling links – Long Range**
    - CHPID type – CIB for Coupling
    - Fiber optic external coupling link – 10 km (Unrepeated), 100 km repeated

Redbooks IBM

z196 I/O Cages and Drawers

z196 I/O infrastructure will support I/O cages (z10 EC) and I/O drawers (z10 BC)

Current I/O cage



28 I/O Slots – Supports all current I/O

Current I/O drawer

8 I/O Slots – Supports all current I/O

I/O cage and Drawer considerations

- Can't order cages or drawers, have to order I/O and/or Crypto features, eConfig will deliver the correct mix of I/O drawers and I/O cages
- There is NO Plan Ahead option available due to the concurrent nature of the I/O drawers
- The I/O drawer can be concurrently added or removed (non-disruptively)
- I/O cage additions and removals are disruptive
- Air cooled models will have a max of 2 x PSC, Water cooled models will have a max of 1 x PSC

Consolidation and Pre-Planning

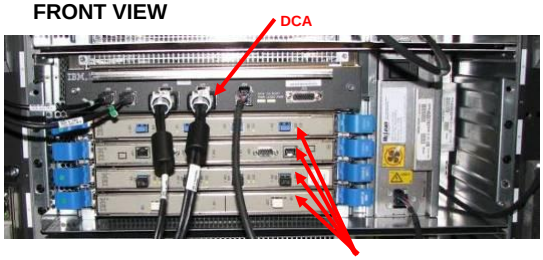
- For customers with large I/O requirements, focus on consolidating to fewer then 72 slots, reducing to 64 slots or less would be ideal as it would leave room for future I/O expansion.

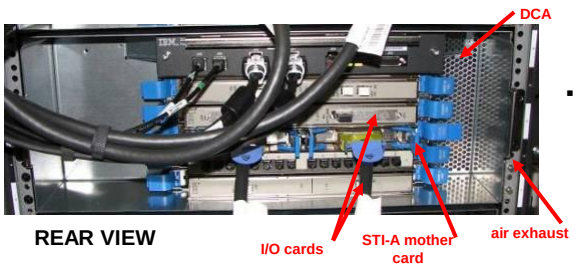
91

Redbooks IBM

z196 I/O Drawer

FRONT VIEW





REAR VIEW

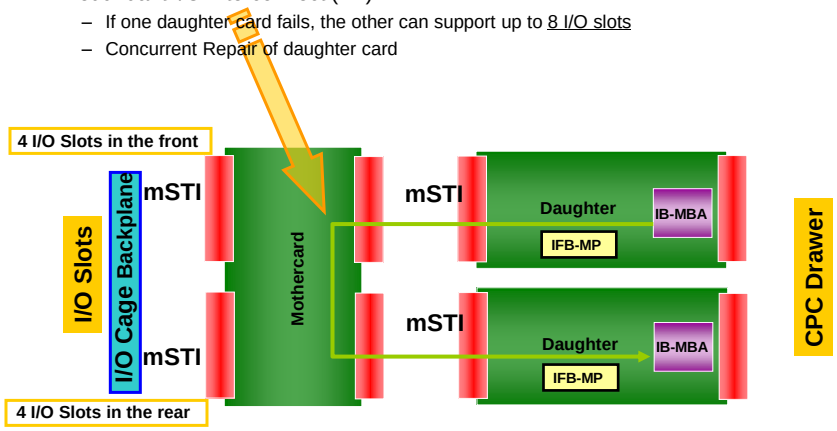
- **Introduced with z10 BC**
  - Up to 8 I/O cards in each drawer – 4 in front and 4 in rear
- **Concurrent drawer add, repair and replacement for Systems with >1 I/O drawer**
- **Drawer can be removed without affecting system input power or power to any other unit**
- **Drawers are favored on z196**
  - Up to 32 I/O cards use 1 to 4 drawers
  - For 33 to 72 I/O cards use 1 or 2 z10 I/O cages plus up to 2 drawers
- **I/O cards in drawers are horizontal; Very important that cables are routed to the side or else concurrent replacement of I/O cards or drawers may not be possible**

92

Redbooks IBM

STI-A8 Mother Card and HCA2-C Daughter Cards

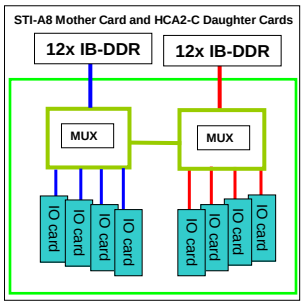
- Each daughter card supports 4 I/O Slots
- Redundant I/O Interconnect (RII)
  - If one daughter card fails, the other can support up to 8 I/O slots
  - Concurrent Repair of daughter card

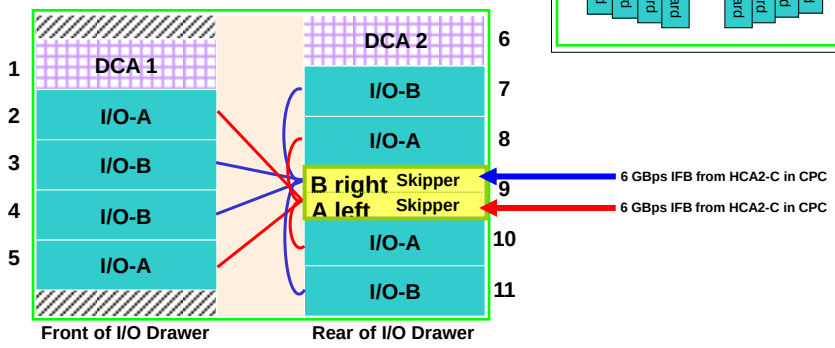


Redbooks IBM

z196 I/O Drawer – HCA and MBA plugging

- 2 I/O Domains per Drawer
- 4 I/O Cards per Domain
  - Domain A and B shown below
- Redundant I/O Interconnect (RII)





Redbooks IBM

z196 – Z frame front and rear views with I/O drawers

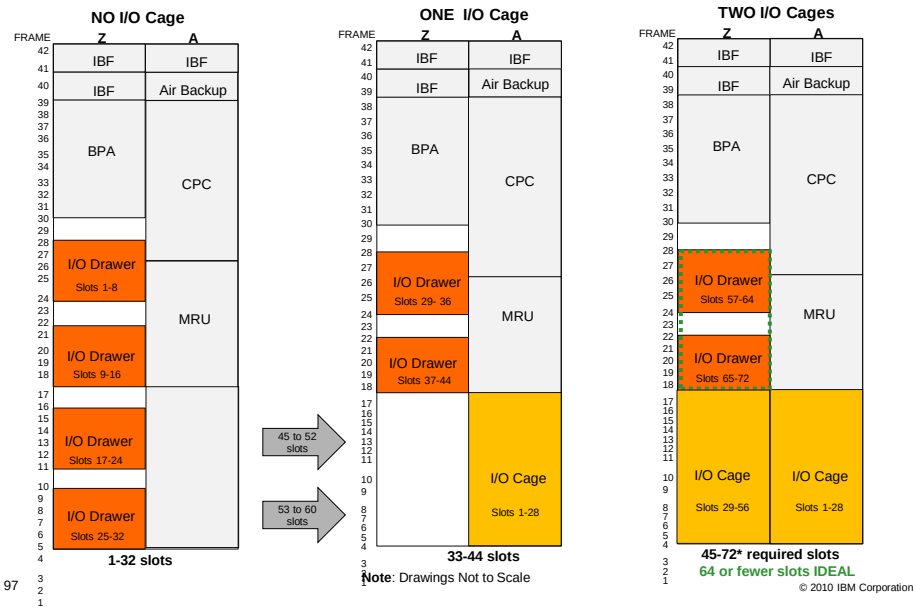
Redbooks IBM

Consolidation and pre-planning required for z196

- **For z196, IBM will begin the migration to the new I/O infrastructure by shipping I/O drawers (i.e. drawer available on the z10 BC) along with the current I/O cage available in z10 EC**
  - Can't order cages or drawers, have to order I/O and/or Crypto features and eConfig will deliver the correct mix of I/O drawers and I/O cages
  - There is NO Plan Ahead option available at due to the concurrent nature of the I/O drawers
  - The I/O drawer can be concurrently added or removed (non-disruptively)
- **Consolidation and Pre-Planning**
  - The total number of I/O slots available are 72 slots (84 is the total number of slots available today with 3 I/O cages).
  - For customers with large I/O requirements, we need them to consolidate their I/O slots to less than 72 slots
  - 64 or less slots would be ideal because that leaves an empty location for future expansion
- **Customers should consolidate/eliminate the following: ESCON, ISC-3, OSA-Express2 (2 ports), FICON Express4**
  - ESCON → Optica Prizm solution
  - ISC-3 → InfiniBand
  - OSA-Express2 (2 ports) → OSA-Express3 (4 ports)
  - FICON Express4 → FICON Express8 (i.e. channel consolidation)

Redbooks IBM

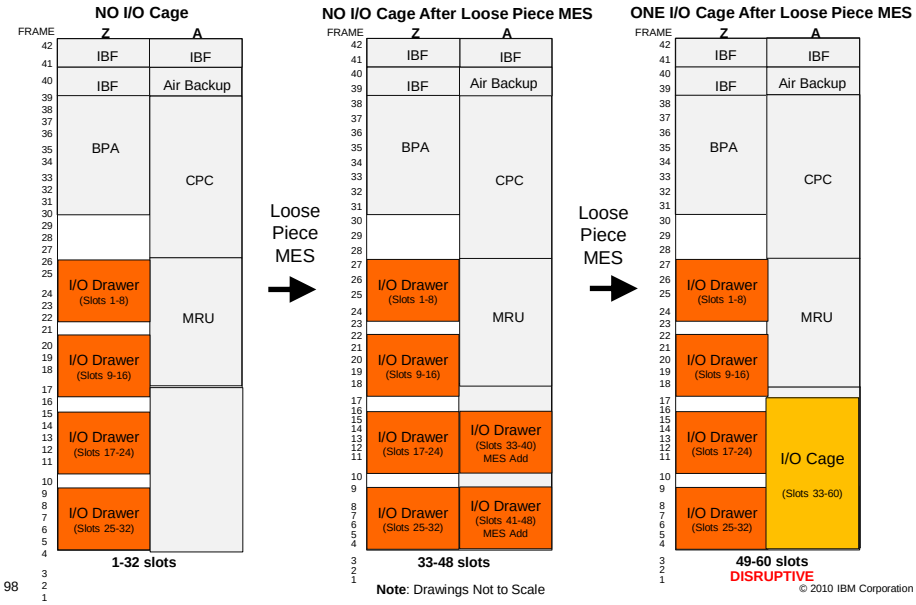
z196 New Build I/O offerings



Note: Drawings Not to Scale

Redbooks IBM

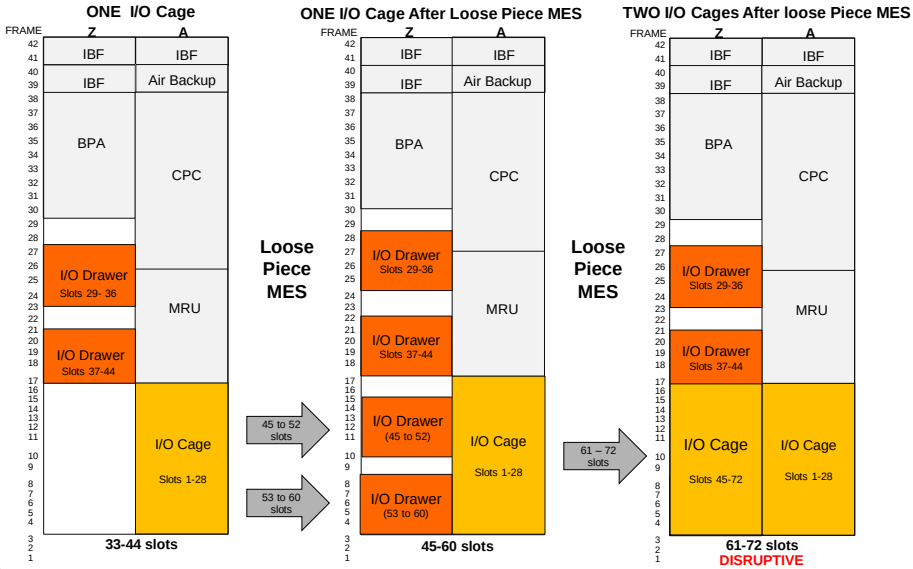
Post-z196 GA MES from no I/O cage configuration



Note: Drawings Not to Scale

Redbooks IBM

MES from one I/O cage configuration



Redbooks IBM

PCHID Assignement

PCHID Ranges							
I/O Card Slot	Cage/Drawer Type	A01B (Air cooled) A08B (Water cooled)	A08B (Air cooled) A15B (Water cooled)	Z01B	Z08B	Z15B	Z22B
LG01	Cargo	100 - 10F		300 - 30F		500 - 50F	
	Cayuga	DCA 1	DCA 1	DCA 1	DCA 1	DCA 1	DCA 1
LG02	Cargo	110 - 11F		310 - 31F		510 - 51F	
	Cayuga	100 - 10F	180 - 18F	300 - 30F	380 - 38F	500 - 50F	580 - 58F
LG03	Cargo	120 - 12F		320 - 32F		520 - 52F	
	Cayuga	110 - 11F	190 - 19F	310 - 31F	390 - 39F	510 - 51F	590 - 59F
LG04	Cargo	130 - 13F		330 - 33F		530 - 53F	
	Cayuga	120 - 12F	1A0 - 1AF	320 - 32F	3A0 - 3AF	520 - 52F	5A0 - 5AF
LG05	Cargo	IFB MP		IFB MP		IFB MP	
	Cayuga	130 - 13F	1B0 - 1BF	330 - 33F	3B0 - 3BF	530 - 53F	5B0 - 5BF
LG06	Cargo	140 - 14F		340 - 34F		540 - 54F	
	Cayuga	DCA 2	DCA 2	DCA 2	DCA 2	DCA 2	DCA 2
LG07	Cargo	150 - 15F		350 - 35F		550 - 55F	
	Cayuga	140 - 14F	1C0 - 1CF	340 - 34F	3C0 - 3CF	540 - 54F	5C0 - 5CF
LG08	Cargo	160 - 16F		360 - 36F		560 - 56F	
	Cayuga	150 - 15F	1D0 - 1DF	350 - 35F	3D0 - 3DF	550 - 55F	5D0 - 5DF
LG09	Cargo	170 - 17F		370 - 37F		570 - 57F	
	Cayuga	IFB MP	IFB MP	IFB MP	IFB MP	IFB MP	IFB MP
LG10	Cargo	180 - 18F		380 - 38F		580 - 58F	
	Cayuga	160 - 16F	1E0 - 1EF	360 - 36F	3E0 - 3EF	560 - 56F	5E0 - 5EF

Redbooks IBM

PCHID Assignment

PCHID Ranges							
I/O Card Slot	Cage/Drawer Type	A01B (Air cooled) A08B (Water cooled)	A08B (Air cooled) A15B (Water cooled)	Z01B	Z08B	Z15B	Z22B
LG11	Cargo	190 - 19F		390 - 39F		590 - 59F	
	Cayuga	170 - 17F	1F0 - 1FF	370 - 37F	3F0 - 3FF	570 - 57F	5F0 - 5FF
LG12	Cargo	1A0 - 1AF		3A0 - 3AF		5A0 - 5AF	
	Cayuga	X	X	X	X	X	X
LG13	Cargo	1B0 - 1BF		3B0 - 3BF		5B0 - 5BF	
	Cayuga	X	X	X	X	X	X
LG14	Cargo	IFB MP		IFB MP		IFB MP	
	Cayuga	X	X	X	X	X	X
LG15	Cargo	1C0 - 1CF		3C0 - 3CF		5C0 - 5CF	
	Cayuga	X	X	X	X	X	X
LG16	Cargo	1D0 - 1DF		3D0 - 3DF		5D0 - 5DF	
	Cayuga	X	X	X	X	X	X
LG17	Cargo	1E0 - 1EF		3E0 - 3EF		5E0 - 5EF	
	Cayuga	X	X	X	X	X	X
LG18	Cargo	1F0 - 1FF		3F0 - 3FF		5F0 - 5FF	
	Cayuga	X	X	X	X	X	X
LG19	Cargo	200 - 20F		400 - 40F		600 - 60F	
	Cayuga	X	X	X	X	X	X
LG20	Cargo	210 - 21F		410 - 41F		610 - 61F	
	Cayuga	X	X	X	X	X	X

101

Redbooks IBM

PCHID Assignment

PCHID Ranges							
I/O Card Slot	Cage/Drawer Type	A01B (Air cooled) A08B (Water cooled)	A08B (Air cooled) A15B (Water cooled)	Z01B	Z08B	Z15B	Z22B
LG21	Cargo	220 - 22F		420 - 42F		620 - 62F	
	Cayuga	X	X	X	X	X	X
LG22	Cargo	230 - 23F		430 - 43F		630 - 63F	
	Cayuga	X	X	X	X	X	X
LG23	Cargo	IFB MP		IFB MP		IFB MP	
	Cayuga	X	X	X	X	X	X
LG24	Cargo	240 - 24F		440 - 44F		640 - 64F	
	Cayuga	X	X	X	X	X	X
LG25	Cargo	250 - 25F		450 - 45F		650 - 65F	
	Cayuga	X	X	X	X	X	X
LG26	Cargo	260 - 26F		460 - 46F		660 - 66F	
	Cayuga	X	X	X	X	X	X
LG27	Cargo	270 - 27F		470 - 47F		670 - 67F	
	Cayuga	X	X	X	X	X	X
LG28	Cargo	IFB MP		IFB MP		IFB MP	
	Cayuga	X	X	X	X	X	X
LG29	Cargo	280 - 28F		480 - 48F		680 - 68F	
	Cayuga	X	X	X	X	X	X
LG30	Cargo	290 - 29F		490 - 49F		690 - 69F	
	Cayuga	X	X	X	X	X	X

102

Redbooks IBM

PCHIDS

Domain to cage position, IFB MP slot and I/O card slot relation				
Domain	I/O Cage Position (In installation order)	IFB MP card position	I/O card slots in the domain	PCHID ranges per domain
A	Z22B	Z22BD109	02 05 08 10	580 - 58F 5B0 - 5BF 5D0 - 5DF 5E0 - 5EF
B	Z22B	Z22BD209	03 04 07 11	590 - 59F 5A0 - 5AF 5C0 - 5CF 5F0 - 5FF
A	Z15B	Z15BD109	02 05 08 10	500 - 50F 530 - 53F 5D0 - 5DF 5E0 - 5EF
B	Z15B	Z15BD209	03 04 07 11	510 - 51F 520 - 52F 540 - 54F 570 - 57F
A	Z08B	Z08BD109	02 05 08 10	380 - 38F 3B0 - 3BF 3D0 - 3DF 3E0 - 3EF
B	Z08B	Z08BD209	03 04 07 11	390 - 39F 3A0 - 3AF 3C0 - 3CF 3F0 - 3FF

Redbooks IBM

PCHIDS

Domain to cage position, IFB MP slot and I/O card slot relation				
Domain	I/O Cage Position (In installation order)	IFB MP card position	I/O card slots in the domain	PCHID ranges per domain
A	Z01B	Z01BD109	02 05 08 10	300 - 30F 330 - 33F 350 - 35F 360 - 36F
B	Z01B	Z01BD209	03 04 07 11	310 - 31F 320 - 32F 340 - 34F 370 - 37F
A	A08B (Air cooled) A15B (Water cooled)	A08BD109 (Air cooled) A15BD109 (Water cooled)	02 05 08 10	180 - 18F 1B0 - 1BF 1D0 - 1DF 1E0 - 1EF
B	A08B (Air cooled) A15B (Water cooled)	A08BD209 (Air cooled) A15BD209 (Water cooled)	03 04 07 11	190 - 19F 1A0 - 1AF 1C0 - 1CF 1F0 - 1FF
A	A01B (Air cooled) A08B (Water cooled)	A01BD109 (Air cooled) A08BD109 (Water cooled)	02 05 08 10	100 - 10F 130 - 13F 150 - 15F 160 - 16F
B	A01B (Air cooled) A08B (Water cooled)	A01BD209 (Air cooled) A08BD209 (Water cooled)	03 04 07 11	110 - 11F 120 - 12F 140 - 14F 170 - 17F

Redbooks IBM

PCHID

Cayuga I/O Drawer PCHID Numbering						
I/O Slot	Z22B	Z15B	Z08B	Z01B	A08B (Air cooled) A15B (Water cooled)	A01B (Air cooled) A08B (Water cooled)
2	580 - 58F	500 - 50F	380 - 38F	300 - 30F	180 - 18F	100 - 10F
3	590 - 59F	510 - 51F	390 - 39F	310 - 31F	190 - 19F	110 - 11F
4	5A0 - 5AF	520 - 52F	3A0 - 3AF	320 - 32F	1A0 - 1AF	120 - 12F
5	5B0 - 5BF	530 - 53F	3B0 - 3BF	330 - 33F	1B0 - 1BF	130 - 13F
7	5C0 - 5CF	540 - 54F	3C0 - 3CF	340 - 34F	1C0 - 1CF	140 - 14F
8	5D0 - 5DF	550 - 55F	3D0 - 3DF	350 - 35F	1D0 - 1DF	150 - 15F
10	5E0 - 5EF	560 - 56F	3E0 - 3EF	360 - 36F	1E0 - 1EF	160 - 16F
11	5F0 - 5FF	570 - 57F	3F0 - 3FF	370 - 37F	1F0 - 1FF	170 - 17F

Redbooks IBM

PCHIDS

Cargo I/O Cage PCHID Numbering			
I/O Slot	PCHIDs for A01B (Air cooled) or A08B (Water cooled)	PCHIDs for Z01B	PCHIDs for Z15B (Only available through RPQ)
1	100 - 10F	300 - 30F	500 - 50F
2	110 - 11F	310 - 31F	510 - 51F
3	120 - 12F	320 - 32F	520 - 52F
4	130 - 13F	330 - 33F	530 - 53F
6	140 - 14F	340 - 34F	540 - 54F
7	150 - 15F	350 - 35F	550 - 55F
8	160 - 16F	360 - 36F	560 - 56F
9	170 - 17F	370 - 37F	570 - 57F
10	180 - 18F	380 - 38F	580 - 58F
11	190 - 19F	390 - 39F	590 - 59F
12	1A0 - 1AF	3A0 - 3AF	5A0 - 5AF
13	1B0 - 1BF	3B0 - 3BF	5B0 - 5BF
15	1C0 - 1CF	3C0 - 3CF	5C0 - 5CF
16	1D0 - 1DF	3D0 - 3DF	5D0 - 5DF
17	1E0 - 1EF	3E0 - 3EF	5E0 - 5EF
18	1F0 - 1FF	3F0 - 3FF	5F0 - 5FF
19	200 - 20F	400 - 40F	600 - 60F
20	210 - 21F	410 - 41F	610 - 61F
21	220 - 22F	420 - 42F	620 - 62F
22	230 - 23F	430 - 43F	630 - 63F
24	240 - 24F	440 - 44F	640 - 64F
25	250 - 25F	450 - 45F	650 - 65F
26	260 - 26F	460 - 46F	660 - 66F
27	270 - 27F	470 - 47F	670 - 67F
29	280 - 28F	480 - 48F	680 - 68F
30	290 - 29F	490 - 49F	690 - 69F
31	2A0 - 2AF	4A0 - 4AF	6A0 - 6AF
32	2B0 - 2BF	4B0 - 4BF	6B0 - 6BF

Redbooks IBM

z196 I/O Connectivity

Redbooks IBM

Channel Type and Crypto Overview

Supported Channel Types

- **I/O Channels**
  - FICON Express8
  - FICON Express4 (Carry forward on upgrade)
  - ESCON – Migrate Away (240 or fewer)
- **Networking**
  - OSA-Express3
    - 10 Gigabit Ethernet LR and SR
    - Gigabit Ethernet LX and SX
    - 1000BASE-T Ethernet
  - OSA-Express2 (Carry forward on upgrade)
    - 1000BASE-T Ethernet
    - Gigabit Ethernet LX and SX
  - HiperSockets (Define only)
- **Coupling Links**
  - InfiniBand Coupling Links
    - 12x InfiniBand
    - 1x InfiniBand
  - ISC-3 – Migrate Away (Peer mode only)
  - IC (Define only)
- **Crypto**
  - **Crypto Express3**
    - Configurable Coprocessor / Accelerator

Non-Supported Channel Types

- **I/O Channels**
  - ⊗ **FICON (before FICON Express4)**
  - ⊗ **FCV – ESCD Model 5 Bridge Card**
- **Networking**
  - ⊗ **OSA-Express2 10 GbE LR**
  - ⊗ **OSA-Express (pre OSA-Express2)**
- **Coupling Links**
  - ⊗ **ICB-4 and earlier ICB**
- **Crypto**
  - ⊗ **Crypto Express2 and earlier**
- **ETR**
  - ⊗ **Sysplex Timer (ETR) Attachment**

I/O discontinuations and recommendations

- **More than 240 ESCON channels, more than 72 I/O cards**
  - Migrate and consolidate to FICON Express8, consider OEM FICON conversion technology if required
  - Eliminate parallel control units or consider OEM FICON conversion technology
- **FICON Express2 – Migrate and consolidate to FICON Express8**
- **FICON Express – Ends support for FICON Conversion with ESCON Director Model 5 Bridge Card**
  - Migrate and consolidate to FICON Express8, consider OEM FICON conversion technology if required
- **OSA-Express2 10 GbE – Migrate to OSA-Express3 10 GbE**
- **Crypto Express2 – Migrate to Crypto Express3**
- **Parallel Sysplex discontinuations**
  - Sysplex with z990, z890 – Limit z196 sysplex coexistence to z9 and z10
  - Sysplex Timer 9037 (ETR) – Migrate to STP and ensure only z9, z10, and zCPC in the CTN
  - Dynamic ICF Expansion – Change production CF to dedicated ICFs (shared ICFs test only)
  - ICB-4 – Migrate to InfiniBand 12x links
  - Note: ISC-3 is supported but consider consolidation to 12x InfiniBand or migration to 1x InfiniBand

ESCON Future Support

- **ESCON channels to be phased out (HW Announcement, April 28, 2009)***
  - System z10 EC and System z10 BC will be the last server to support greater than 240 ESCON channels
  - Currently, 1024 channels are supported on System z10 EC and 480 channels on the System z10 BC
- **FICON bridge has been phased out and 9032-005 directors are approaching End Of Service**
- **Customer Options**
  - Run unsupported
  - Replace control units with FICON interfaces and FICON channels
  - Utilize Optica PRIZM solution, FICON to ESCON converter
- **Similar approach was utilized to phase out of Parallel Channels (i.e. Bus and Tag)**
  - Optica was IBM's strategic partner utilizing many ESCON to parallel converters (which still exist in the field).

*All statements regarding IBM's plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these Statements of General Direction is at the relying party's sole risk and will not create liability or obligation for IBM.

Redbooks IBM

OSA-Express3

- **Double density of ports compared to OSA-Express2**
  - Reduced CHPIDs to manage
  - Reduced I/O slots
  - Reduced I/O cages or I/O drawers
  - Up to 96 LAN ports versus 48
- **Designed to reduce the minimum round-trip networking time between z196 systems (reduced latency)**
  - Designed to improve round trip at the TCP/IP application layer
    - OSA-Express3 10 GbE
      - **45%** improvement compared to the OSA-Express2 10 GbE
    - OSA-Express3 GbE
      - **45%** improvement compared to the OSA-Express2 GbE
  - Designed to improve throughput (mixed inbound/outbound)
    - OSA-Express3 10 GbE
      - 1.0 GBytes/ps @ 1492 MTU
      - 1.1 GBytes/ps @ 8992 MTU
      - 3-4 times the throughput of OSA-Express2 10 GbE
      - 0.90 of Ethernet line speed sending outbound 1506-byte frames
      - 1.25 of Ethernet line speed sending outbound 4048-byte frames

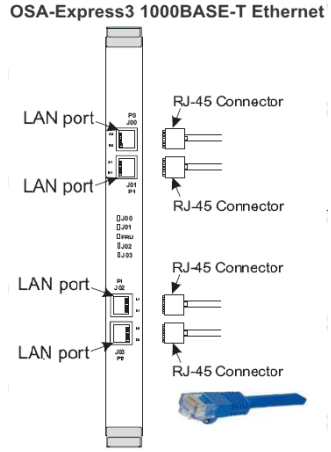
The above statements are based on OSA-Express3 performance measurements performed in a test environment on a System z10 EC and do not represent actual field measurements. Results may vary.

Redbooks IBM

OSA-Express3 – 1000BASE-T

- 1000BASE-T Ethernet
- Two and Four ports per feature options
  - RJ45, Cat 5 UTP, up to 100 meters (328 feet)
  - Two ports* per PCI-E adaptor/CHPID
  - CHPIDs Supported
    - OSC (OSA-Integrated Console)
    - OSD (TCPIP and Layer 2)
      - OS PTF required to use 2nd port
    - OSE (Non-QDIO TCPIP and SNA/APPN®)
    - OSN (OSA-Express for NCP)
- New microprocessor and hardware data router
  - Large send for IPv4 traffic
  - Checksum offload
  - Concurrent LIC update
  - Auto-negotiation 10/100/1000
  - Large send packet construction, inspection and routing performed in hardware

FC 3367
OSA-Express3 1000BASE-T Ethernet



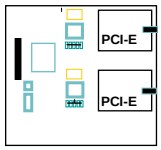
Notes:
For CHPID type OSC, Port 0 is only used.
- **Port 1 not 'visible' to OS**

For CHPID type OSD, OSE, both ports on each PCI-E adaptor are used
Each PCI-E adaptor can be defined as the same or different CHPID type

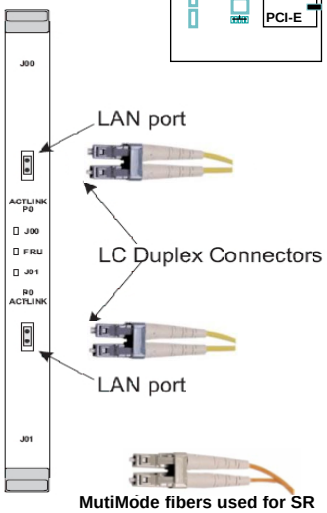
*** NOTE: To use 2-Ports per PCI-E adaptor, the following is required – z/OS V1.9+, z/VM V5.2+, z/VSE V4.1+, zTPF 1.1 PUT 4 with APARs.**
If this support isn't installed, only port zero on a PCI-E adaptor is 'visible' to the Operating System.

Redbooks IBM

OSA Express3 10GbE LR and SR – FC 3370/71

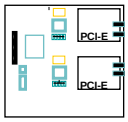


- FC3370 LR / 3371 SR
  - two ports per card
  - Maximum number of OSA Express3 10GbE cards is 24
    48 ports per system
  - Small form factor connector (LC Duplex)
    LR = Single Mode 9 micron fiber
    SR = Multimode 50 or 62.5 micron fiber
  - New small form factor LC Duplex connector
    replaces 1 port / SC duplex
  - Unrepeated distance: 10 Km (6.2 miles)
    Data rate: 10 Gigabits per second
    Operating Mode: Full Duplex
    Data rate: 10 Gigabits per second
  - Operating Mode: Full Duplex
  - CHPID type OSD (QDIO) (QDIO TCPIP and Layer 2)
  - PCHID's xx0 / xx1



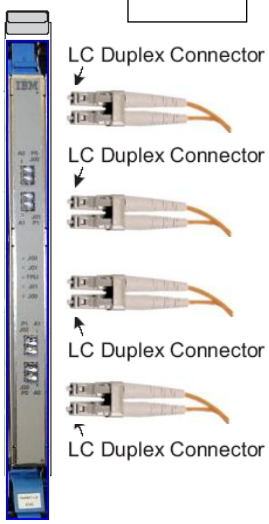
Redbooks IBM

OSA-Express3 - 4P Gigabit Ethernet SX



FC3363

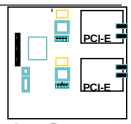
- **Four ports per card**
- **Maximum number of OSA Express3 GbE SX cards is 24**
  - **96 ports per system**
- **Uses multimode fiber**
- **Small form factor connector (LC Duplex)**
- **Unrepeated distances:**
  - **50 micron fiber at 500 MHz-Km: 550m (1804 feet)**
  - **62.5 micron fiber at 200 MHz-Km: 275m (902 feet)**
  - **62.5 micron fiber at 160MHz-Km: 220m (722 feet)**
- **Data rate: 1 Gigabit per second**
- **Operating Mode: Full Duplex**
- **2 x PCHID's (xx0 / xx1)**
  - **Two ports per PCHID/CHPID**
  - **Each PCHID will support port 0 and port 1 and supports QDIO mode only (CHPID type OSD), and is designed to deliver line speed - 1 Gbps in each direction**
  - **Each port supports attachment to a one Gigabit per second (Gbps) Ethernet Local Area Network (LAN).**



Note: For the operating system to recognize all four ports on an OSA Express3 Gigabit Ethernet feature, a new release and/or PTF is required. If software updates are not applied, only two of the four ports will be visible by the operating system.

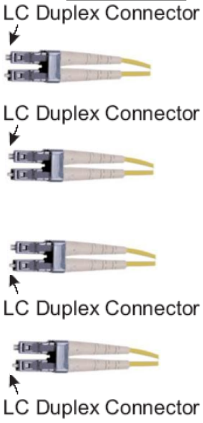
Redbooks IBM

OSA-Express3 - 4P Gigabit Ethernet LX



FC3362

- Four ports per card
- Maximum number of OSA Express3 GbE SX cards is 24
  - 96 ports per system
- 9 micron single mode fiber
- small form factor connector (LC Duplex)
- Unrepeated distance: 5 Km (3.1 miles)
- Data rate: 1 Gigabit per second
- Operating Mode: Full Duplex
- 2 x PCHID's (xx0 / xx1)
  - two ports per PCHID/CHPID
  - Each PCHID will support port 0 and port 1 and supports QDIO mode only (CHPID type OSD), and is designed to deliver line speed - 1 Gbps in each direction
  - Each port supports attachment to a one Gigabit per second (Gbps) Ethernet Local Area Network (LAN).



*** NOTE: To use 2-Ports per PCI-E adaptor, the following is required – z/OS V1.9+, z/VM V5.2+, z/VSE V4.1+, zTPF 1.1 PUT 4 with APARs.**
If this support isn't installed, only port zero on a PCI-E adaptor is 'visible' to the Operating System.

Redbooks IBM

z196 – New OSA-Express3 CHPIDs types OSX and OSM

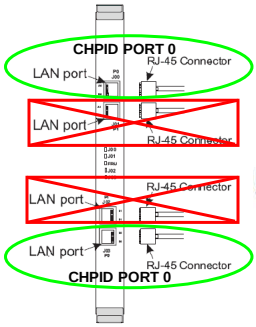
- **Two new OSA-Express3 CHPID types to support new types of z196 networks**
- **A z196 System can have up to 6 types of OSA-Express3 CHPID's**
  - External (customer managed) networks
    - Defined as OSC,OSD,OSE, & OSN
    - Existing customer provided and managed OSA ports used for access to the current customer external networks - no changes
  - Intranode management network (INMN)
    - Defined as CHPID type OSM for Unified Resource Manager
      - When the PCIe adaptor on 1000BASE-T is defined as CHPID type OSM, the second port cannot be used for anything else
    - OSA-Express3 1000BASE-T configured as CHPID type OSM for z196 Node management of the z196 Ensembles
    - OSA connection via the Bulk Power Hub (BPH) on the z196 to the Top of the Rack (TORs) switches on zBX
  - Intraensemble data networks (IEDN)
    - Defined as CHPID OSX for zBX
    - OSA-Express3 10 GbE configured as CHPID type OSX for data between z196 and zBX
- **Functions Supported:**
  - Dynamic I/O support
  - HCD
  - CP Query capabilities
  - Ensemble Management for these new channel paths and their related subchannels.
- **z/VM 5.4 – CHPID types OSX and OSM cannot be varied online**

Redbooks IBM

z196 CHPID Types OSX and OSM – Usage

OSM (INMN)

FC 3367
OSA-Express3 1000BASE-T Ethernet



OSA-Express3 1GbE
2 CHPIDS
2 PORTS/CHPID

CCIN 57AC FC3367

CAT 6 ETH CABLE

MUST USE CHPID PORT 0

Supports IOCP CHPID types:
OSC, OSD, OSE, OSN, and OSM (ONLY 1000BASE-T).
PCHID = xx0 & xx1

OSM IOCDS EXAMPLE:

```
• CHPID PCHID=191,PATH=(CSS(0,1,2,3),23),TYPE=OSM,,SHARED,
• CNTLUNIT CUNUMBR=0910,PATH=((CSS(0),23)),UNIT=OSM
• IODEVICE ADDRESS=(0910,15),CUNUMBR=(0910),UNIT=OSA,UNITADD=00,      *
        MODEL=M,DYNAMIC=YES,LOCANY=YES
```

OSX (IEDN)

OSA-Express3 10 GbE
2 CHPIDS
1 PORT/CHPID

IEDN Distances
•MM (Short Range Optics)
50 micron at 2000 MHz-km: 300 meters (984')
50 micron at 500 MHz-km: 82 meters (269')
62.5 micron at 200 MHz-km: 33 meters (108')
•SM (Long Range Optics) 10 km (6.2 miles)

CCIN 57A3 FC3370 (LR)
Single Mode 9 micron LC duplex

CCIN 57AD FC3371 (SR)
Multi Mode 50/62.5 micron LC duplex

Supports IOCP CHPID types:
OSD and OSX (ONLY 10 GbE).
PCHID = xx0 & xx1

OSX IOCDS EXAMPLE:

```
•CHPID PCHID=5E1,PATH=(CSS(0,1,2,3),2F),TYPE=OSX,SHARED,  ....
•CNTLUNIT CUNUMBR=09F0,PATH=((CSS(0),2F)),UNIT=OSX
• IODEVICE ADDRESS=(09F0,15),CUNUMBR=(09F0),UNIT=OSA,UNITADD=00,      *
        MODEL=X,DYNAMIC=YES,LOCANY=YES
```


Redbooks IBM

OSA-Express3 GbE & 1000BASE-T CHPID usage

Feature Type	CHPID Type	CHPID Use	Number of available ports	Software requirements
1000BASE-T 2 or 4 ports	OSC	3270 data streams	One port per PCI-E available for use	No unique software requirements
1000BASE-T 2* or 4* ports - - - - - - - - - - - - Gigabit Ethernet 2* or 4* ports	OSD	QDIO L2 & L3	Two ports per PCI-E available for use	Requires software exploitation
1000BASE-T 2* or 4* ports	OSE	SNA/APPN/HPR TCP/IP passthru	Two ports per PCI-E available for use	No unique software requirements
1000BASE-T 2 or 4 ports	OSN	NCP & CCL	Not applicable LPAR-to-LPAR only	No unique software requirements

Note: OSA-Express3 10 GbE has one port per PCI-E and one port per CHPID. It supports CHPID type OSD exclusively. Therefore it is not covered in the table above.

*** NOTE: To use 2-Ports per PCI-E adaptor, the following is required – z/OS V1.8+, z/VM V5.2+, z/VSE V4.1+, zTPF 1.1 PUT 4 with APARs. If this support isn't installed, only port zero on a PCI-E adaptor is 'visible' to the Operating System**

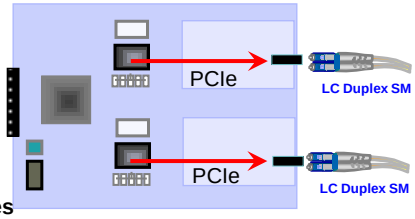
Redbooks IBM

OSA-Express3 'fiber'

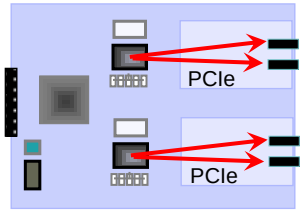
- **Double the port density of OSA-Express2**
- **Reduced latency & improved throughput**
  - Ethernet hardware data router
- **Improved throughput – standard & jumbo frames**
  - New microprocessor
  - New PCI adapter
- **CHPID types**
  - 10 Gigabit Ethernet – OSD, OSX
  - Gigabit Ethernet – OSD and OSN

    CHPID type OSN is for LPAR-to-LPAR communication. Does not use ports

	OSA-Express2	OSA-Express3
Microprocessor	500 MHz – 10 GbE 448 MHz – 1 GbE	667 MHz
PCI bus	PCI-X	PCIe G1



10 GbE LR #3370,
10 GbE SR #3371



CHPID shared by two ports
GbE LX #3362, GbE SX #3363

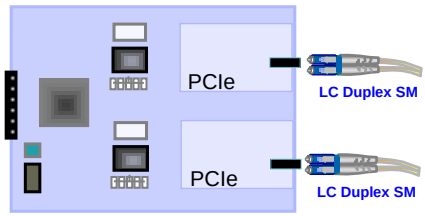
Redbooks IBM

OSA-Express3 – 10 GbE

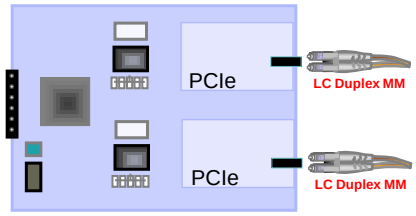
- **10 Gigabit Ethernet LR (Long Reach) and SR (Short Reach)**
  - One port per PCI-E adaptor
  - Two ports per feature
  - Small form factor connector (LC Duplex)

    LR = Single Mode 9 micron fiber
    SR = Multimode 50 or 62.5 micron fiber
  - Two CHPIDs, one port each

    Type OSD (QDIO TCPIP and Layer 2)
    Type OSX (for IEDN)
- **New Microprocessor and hardware data router**
  - Large send packet construction, inspection and routing preformed in hardware instead of firmware
  - Large send for IPv4 traffic
  - Checksum offload
  - Concurrent LIC update
  - Designed to improve performance for standard (1492 byte) and jumbo frames (8992 byte)
- **Up to 45% reduction in latency compared to OSA-Express2 10 GbE**



10 GbE – LR, 2 ports

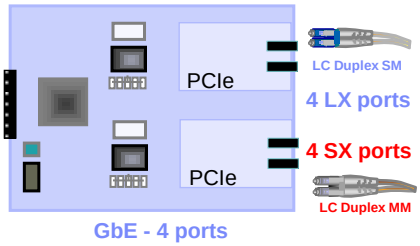


10 GbE – SR, 2 ports

Redbooks IBM

OSA-Express3 GbE – 4 ports feature

- **Gigabit Ethernet LX and SX**
  - Four ports per feature options
  - Two ports* per PCI-E adaptor/CHPID
    - OS PTF required to use 2nd port
  - CHPIDs support
    - OSD (QDIO TCPIP and Layer 2)
    - OSN (OSA-Express for NCP)
  - Small form factor connector (LC Duplex)
- **New microprocessor and hardware data router**
  - Large send packet construction, inspection and routing preformed in hardware instead of firmware
  - Large send for IPv4 traffic
  - Checksum offload
  - Concurrent LIC update
- **Up to 45% reduction in latency compared to OSA-Express2 GbE**

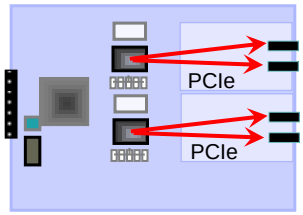


*** NOTE:** To use 2-Ports per PCI-E adaptor, the following is required – z/OS V1.9+, z/VM V5.4+, z/VSE V4.1+, zTPF 1.1 PUT 4 with APARs. If this support isn't installed, only port zero on a PCI-E adaptor is 'visible' to the Operating System

Redbooks IBM

OSA-Express3 'copper'

- **Auto-negotiation to 10, 100, 1000 Mbps**
- **Double the port density of OSA-Express2**
- **Reduced latency & improved throughput**
  - Ethernet hardware data router
- **Improved throughput – standard & jumbo frames**
  - New microprocessor
  - New PCI adapter
- **Usage of ports**
  - OSC, OSD, OSE can exploit four ports
  - OSM, Port 0 only
  - CHPID type OSN is for LPAR-to-LPAR communication. Does not use ports



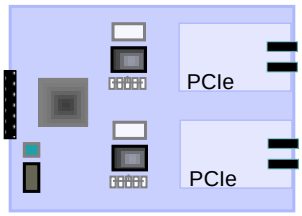
CHPID shared by two ports
1000BASE-T # 3367

	OSA-Express2	OSA-Express3
Microprocessor	448 MHz	667 MHz
PCI bus	PCI-X	PCIe G1

Mode	CHPID	Description
OSA-ICC	OSC	TN3270E, non-SNA DFT, IPL CPCs, and LPARs, OS system console operations
QDIO	OSD	TCP/IP traffic when Layer 3, Protocol-independent when Layer 2
Non-QDIO	OSE	TCP/IP and/or SNA/APPN/HPR traffic
OSA for NCP (LP-to-LP)	OSN	NCPs running under IBM Communication Controller for Linux (CDLC)
Unified Resource Manager	OSM	Connectivity to intranode management network (INMN) from z196 to zBX

Redbooks IBM

OSA-Express3 1000BASE-T, OSA-ICC



FC 3367 1000BASE-T, 4 ports

(FC 3369 – 2 ports)

- **1000BASE-T Ethernet**
- **Four ports per feature options**
  - RJ45, Cat 5 UTP, up to 100 meters (328 feet)
  - Two ports* per PCI-E adapter/CHPID
  - CHPIDs Supported
    - OSC (OSA-Integrated Console)
    - OSD (TCPIP Layer 2 and Layer 3)
      - OS PTF required to use 2nd port
    - OSE (Non-QDIO TCPIP and SNA/APPN®)
    - OSN (OSA-Express for NCP)
    - OSX (for INMN)
- **New microprocessor and hardware data router**
  - Large send for IPv4 traffic
  - Checksum offload
  - Concurrent LIC update
  - Auto-negotiation 10/100/1000
  - Packet construction, inspection and routing performed in hardware

Notes:

- For CHPID type OSC, OSD, OSE, both ports on each PCI-E adaptor are used
- CHPID type OSX only uses port-0 of a PCHID. Port-1 can not be used
- Each PCI-E adaptor can be defined as the same or different CHPID type

*** NOTE:** For OSD CHPID type to use 2-Ports per PCI-E adaptor, the following is required – z/OS V1.9+, z/VM V5.4+, z/VSE V4.1+, zTPF 1.1 PUT 4 with APARs. If this support isn't installed, only port zero on a PCI-E adaptor is 'visible' to the Operating System

Redbooks IBM

OSA-Express3 – Data router

- **Consists of IBM 'System on Chip' (SOC) ASIC in 90nm technology and companion high performance FPGA**
- **SOC contains:**
  - Dual PowerPC embedded microprocessors
  - 20Mb embedded DRAM
  - Hundreds of thousands of random latches
  - Hundreds of interface signals
- **Hardware-based Ethernet virtualization offload engine**
  - Performs packet construction, inspection, and routing
    Previously accomplished in Licensed Internal Code
  - Functions as a bus master on the self-timed interconnect (STI) host connection to fetch and store packets within System z host memory as required
- **Extensive error detection and correction policy throughout the design**

Redbooks IBM

Open Systems Adapter performance

- **OSA processor becomes more efficient as data traffic increases**
- **Window size**
  - TCP window determines amount of data that the sender can transmit to receiver without needing an acknowledgment from the receiver
  - Faster and longer networks require larger windows to keep data flowing smoothly
- **Blocking**
  - Performance is affected by the amount of data blocked together for transfer between OSA and TCP
- **Frame size**
  - Larger frames perform better
  - Larger frames reduce host and OSA processing costs
  - Size of frame depends on LAN type, MTU setting, size of data sent
- **Measure: throughput, transaction response time, server utilization**
  - Bulk data transfer and interactive transactions
- **QDIO and jumbo frames (8992 byte MTUs) yield the highest streams**

Redbooks IBM

OSA-Express3 family of LAN adapters

Feature	Feature Name	Ports	CHPIDs	Connectors
3362	GbE LX	4	OSD L2/L3, OSN	LC Duplex
3363	GbE SX	4	OSD L2/L3, OSN	LC Duplex
3367	1000BASE-T	4	OSC, OSD L2/L3, OSE, OSM, OSN	RJ-45
3370	10 GbE LR	2	OSD L2/L3, OSX	LC Duplex
3371	10 GbE SR	2	OSD L2/L3, OSX	LC Duplex

Redbooks IBM

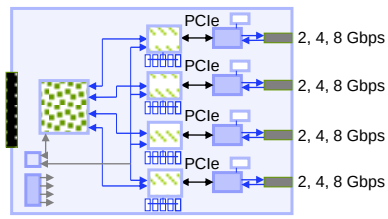
Summary – CHPID type controls operation

CHPID type	Purpose / Traffic	Operating Systems
OSD **All OSA features** **z196, z10, z9, zSeries**	Supports Queue Direct Input/Output (QDIO) architecture TCP/IP traffic when Layer 3 (uses IP address) Protocol-independent when Layer 2 (uses MAC address)	z/OS, z/VM z/VSE, z/TPF Linux on System z
OSE **1000BASE-T** **z196, z10, z9, zSeries**	Non-QDIO; for SNA/APPN/HPR traffic and TCP/IP "passthru" traffic	z/OS, z/VM z/VSE
OSC **1000BASE-T** **z196, z10, z9, z990, z890**	OSA-Integrated Console Controller (OSA-ICC) Supports TN3270E, non-SNA DFT to IPL CPCs & LPs	z/OS, z/VM z/VSE
OSM **1000BASE-T** **z196 exclusive**	OSA-Express for Unified Resource Manager Connectivity to intranode management network (INMN) from z196 to Unified Resource Manager functions	z/OS, z/VM Linux on System z*
OSN **GbE, 1000BASE-T** **z196, z10, z9 exclusive**	OSA-Express for NCP Appears to OS as a device supporting CDLC protocol Enables Network Control Program (NCP) channel-related functions Provides LP-to-LP connectivity OS to IBM Communication Controller for Linux (CCL)	z/OS, z/VM z/VSE, z/TPF Linux on System z
OSX **10 GbE** **z196 exclusive**	OSA-Express for zBX Connectivity and access control to intraensemble data network (IEDN) from z196 to zBX	z/OS, z/VM Linux on System z*

***IBM is working with its Linux distribution partners to include support in future Linux on System z distribution releases.**

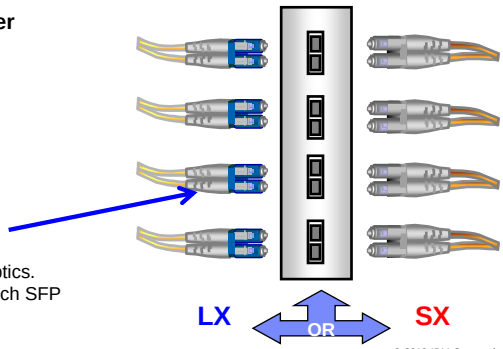
Redbooks IBM

FICON Express8



- **Auto-negotiate to 2, 4, or 8 Gbps**
- **Connector – LC Duplex**
- **10KM LX – 9 micron single mode fiber**
  - Unrepeated distance - 10 kilometers (6.2 miles)
  - Receiving device must also be LX
- **SX – 50 or 62.5 micron multimode fiber**
  - Variable with link data rate and fiber type
  - Receiving device must also be SX
- **4 channels of LX or SX (no mix)**

3325 – 10KM LX, # 3326 – SX

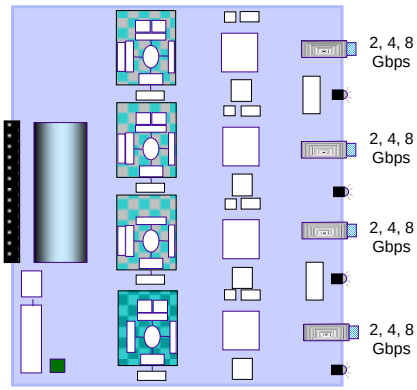


Small Form Factor Pluggable (SFP) optics.
Concurrent repair/replace action for each SFP

Redbooks IBM

FICON Express8 10 Km LX

- **FICON Express8**
  - Supported on System z10 EC and z10 BC
  - Can be shared among LPARs, and defined as spanned
  - High Performance FICON for System z (zHPF) for FICON Express8
  - Same Software Requirements as for FICON Express4
- **8 Gbps with Auto-negotiate capability (2, 4, or 8 Gbps)**
  - 1 Gbps NOT supported
- **LX – 9 micron single mode fiber**
  - Unrepeated distance - 10 kilometers (6.2 miles)
  - Receiving device must also be LX
- **Small Form Factor pluggable optics for concurrent repair/replace**
- **Personalize as:**
  - FC
    Native FICON
    Channel-To-Channel (CTC)
    - z/OS, z/VM, z/VSE, z/TPF, TPF, Linux on System z
  - FCP (Fibre Channel Protocol)
    Support of SCSI devices
    - z/VM, z/VSE, Linux on System z



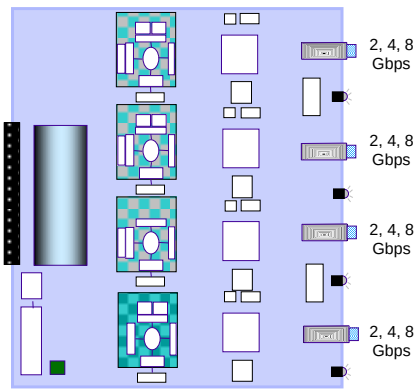
10 Km LX
FC 3325

LC Duplex

Redbooks IBM

FICON Express8 SX

- **FICON Express8**
  - Supported on System z10 EC and z10 BC
  - Can be shared among LPARs, and defined as spanned
  - High Performance FICON for System z (zHPF) for FICON Express8
  - Same Software Requirements as for FICON Express4
- **8 Gbps with Auto-negotiate capability (2, 4, or 8 Gbps)**
  - 1 Gbps NOT supported
- **SX - 50 or 62.5 micron multimode fiber**
  - Unrepeated distance (20 to 500 meters) varies with speed and fiber type
  - Receiving device must also be SX
- **Small Form Factor pluggable optics for concurrent repair/replace**
- **Personalize as:**
  - FC
    Native FICON
    Channel-To-Channel (CTC)
    - z/OS, z/VM, z/VSE, z/TPF, TPF, Linux on System z
  - FCP (Fiber Channel Protocol)
    Support of SCSI devices
    - z/VM, z/VSE, Linux on System z



SX FC 3326

LC Duplex

Redbooks IBM

Buffer credits for FICON Express8

- 40 buffer credits available per channel
- Conforms to Fibre Channel standard flow control
  - Avoids overwriting the buffer
  - 40 credits is 80 KB of buffer at the receiver
  - Sufficient to hold the maximum data payload on the fiber for 20 km (round trip) of data. Anything less will not fill the buffer
- Receiver side of port should contain sufficient buffer credits to ensure channel is fully utilized
  - Receiver must contain = > 4 buffer credits per km of fiber distance at 8 Gbps
- Buffer credits are directly related to the distance supported
  - 8 Gbps needs about 4 credits/km...

    10 km = about 40 credits. 100 km = about 400 credits
  - 4 Gbps needs about 2 credits/km...

    10 km = about 20 credits. 100 km = about 200 credits
- In the short block case, the concern is response time, not data bandwidth. Where the buffer-to-buffer credits (BtoB credits) becomes an issue is when large blocks per I/O are being transferred, so link bandwidth becomes the factor of interest. In this case, the frames will be almost always full (2K).

Redbooks IBM

z/OS Discovery and Auto-Configuration (zDAC) for z196

- **Provides automatic discovery for FICON disk and tape control units**
- **Reduces level of IT skill and time required to configure new I/O devices**
- **Ensures system (host) and control unit definitions are compatible with each other**
- **Automatically discovers storage devices accessible to the system but not currently configured and proposes host definition values**
  - For those discovered control units, explore for defined logical control units and devices
  - Compare discovered logical control units and devices against those configured previously
  - Add missing logical control units and devices to the configuration, proposing control unit and device numbers, and proposing paths to reach them

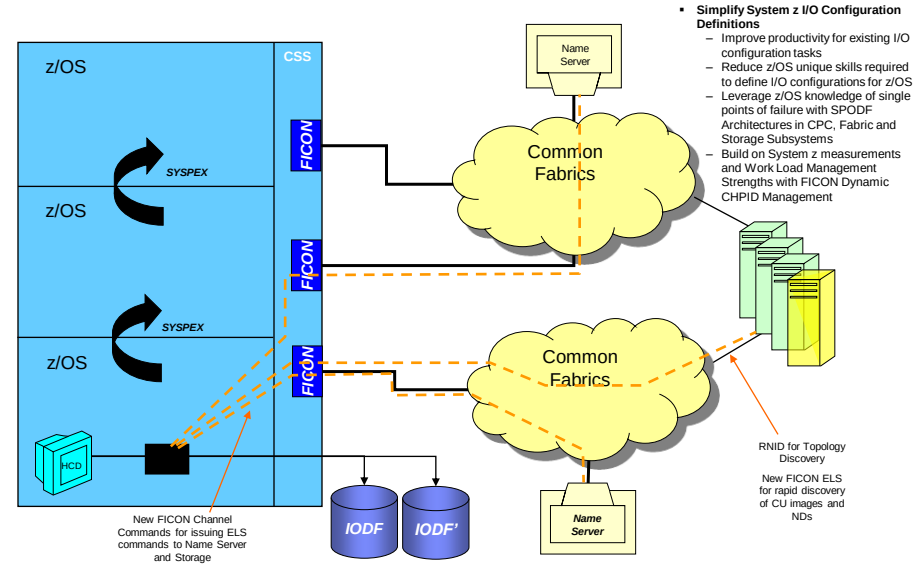
    Channel paths chosen using algorithm to minimize single points of failure
- **Integrated into existing System z host configuration tools (HCD & HCM)**
- **Requirements**:
  - IBM zEnterprise™ 196 (z196) server with FICON Express8 or FICON Express4
  - Switch/director attached fabric (no direct attachment)
  - z/OS V1.12 (at least 1 LPAR for Dynamic I/O capability)
  - First exploiter is IBM System Storage DS8700

    DS8000® licensed machine code level 6.5.15.xx (bundle version 75.15.xx.xx), or later.

    see IBM PSP buckets
  - Suggested: FICON DCM (z/OS dynamic channel management) to help manage performance

Redbooks IBM

z/OS Discovery and Auto-Configuration (zDAC)



- **Simplify System z I/O Configuration Definitions**
  - Improve productivity for existing I/O configuration tasks
  - Reduce z/OS unique skills required to define I/O configurations for z/OS
  - Leverage z/OS knowledge of single points of failure with SPODF Architectures in CPC, Fabric and Storage Subsystems
  - Build on System z measurements and Work Load Management Strengths with FICON Dynamic CHPID Management

Redbooks IBM

zDAC Implementation

- zDAC is invoked through HCD or HCM
- Provides automatic discovery for FICON DASD and TAPE Control Units
- Reduces level of IT skills and time required to configure new I/O devices
- Ensures system (host) and Control Unit definitions are compatible with each other
- Automatically discovers storage devices accessible to the system but not currently configured and proposes host definition values
- This discovery capabilities are delivered via new options on HCD and HCM

What can zDAC do?

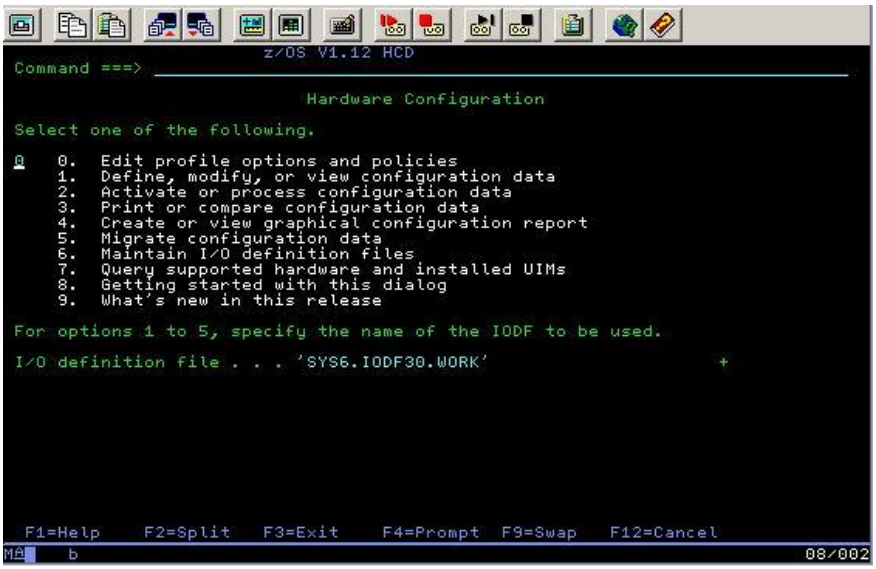
- Display a list of discovered DASD and TAPE controllers
  - Indicates whether the controller is new -- ie. no existing devices/LCUs defined in the target IIODF
- Discover new devices for existing Logical Control Units on a controller
- Discover new devices and new Logical Control Units on a controller, and propose new paths for each new Logical Control Unit found
- New devices and Control Units are displayed for either inclusion or exclusion

Discovery and Auto configuration Policy

- Through HCD or HCM the users establishes a policy for the discovery operation which can:
  - Limit the scope of the discovery
  - Limit the proposal information
  - Indicate the desired number of paths to discovered Logical Control Units
  - Indicate the method used for device and Control unit numbering

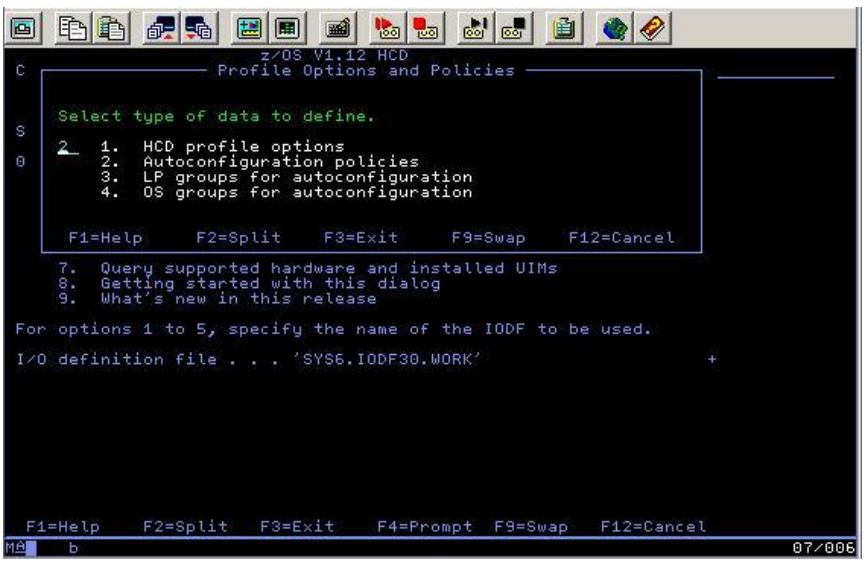
Redbooks IBM

HCD Options Panel



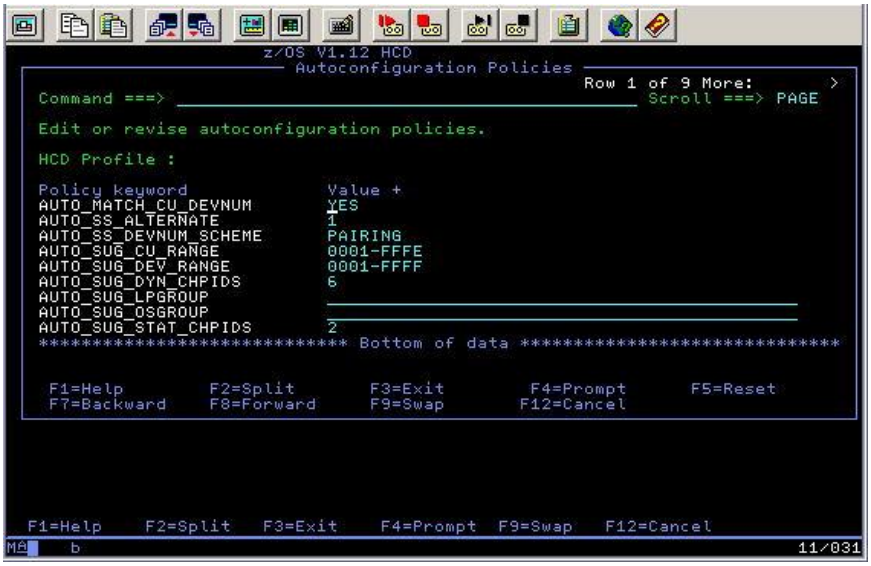
Redbooks IBM

Auto configuration Options



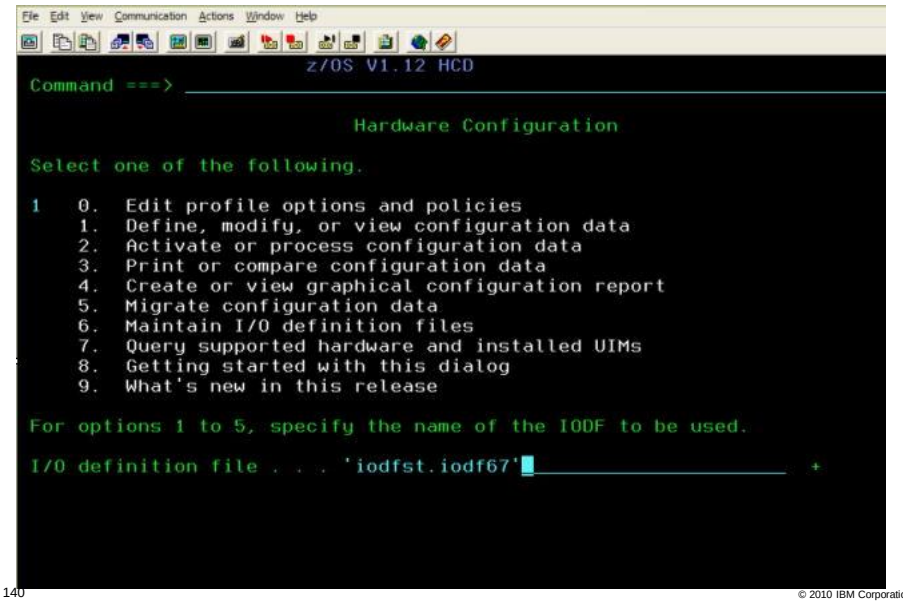
Redbooks IBM

HCD Autoconfiguration Policies



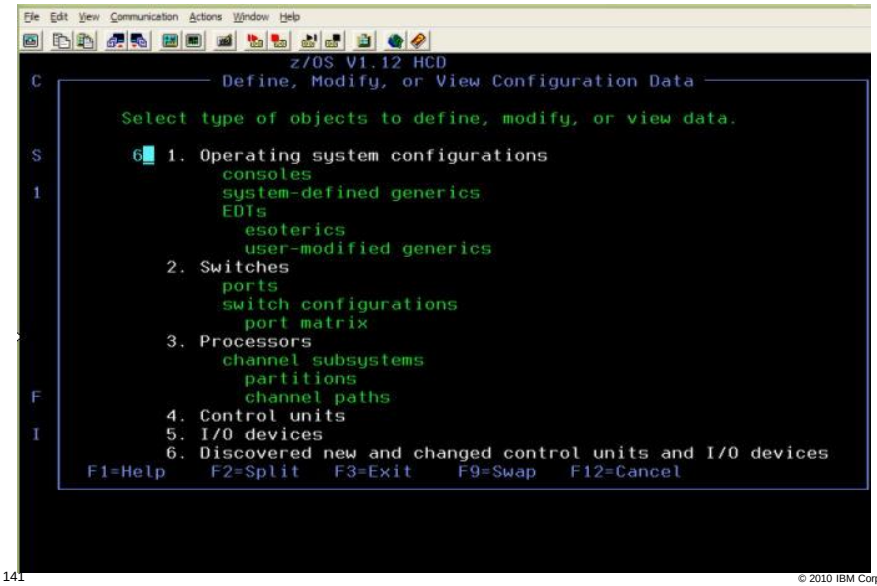
Redbooks IBM

HCD



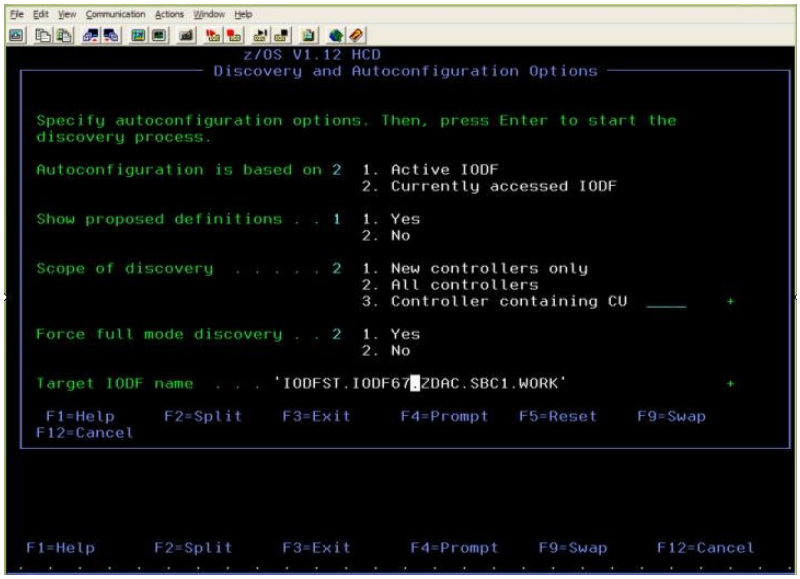
Redbooks IBM

Using zDAC



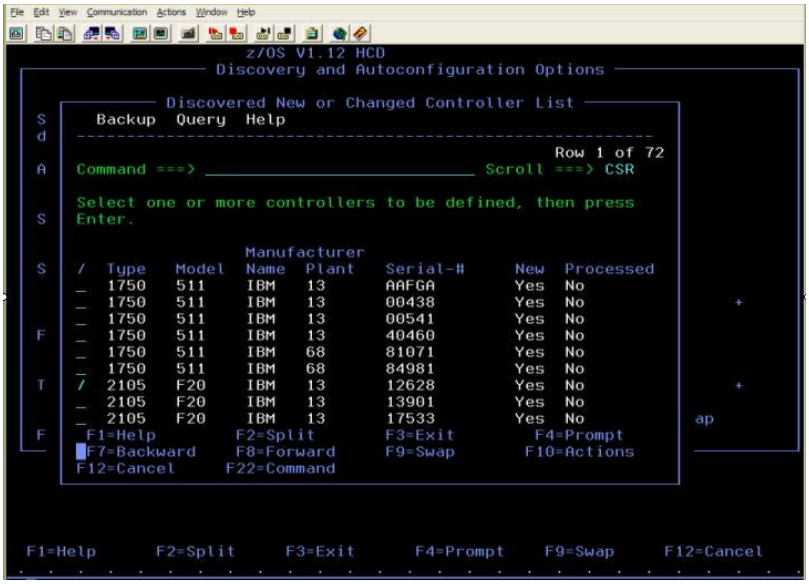
Redbooks IBM

zDAC Options



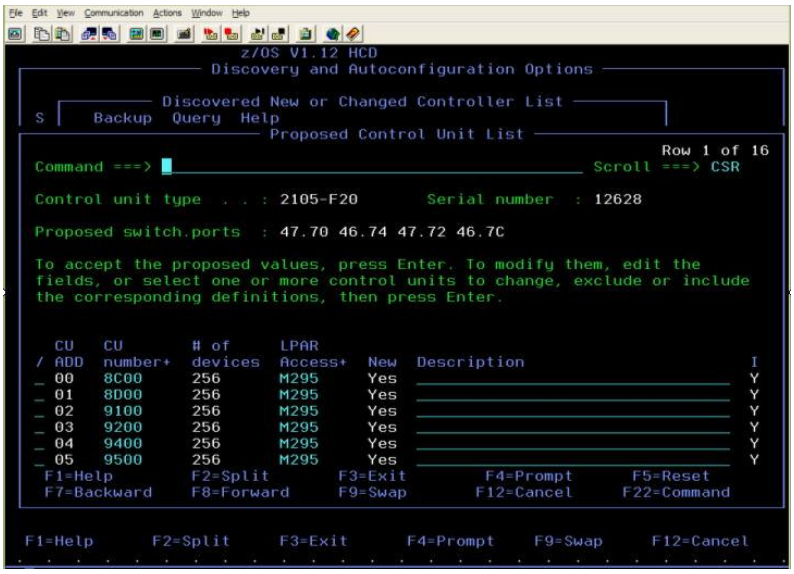
Redbooks IBM

Discovered Control Units



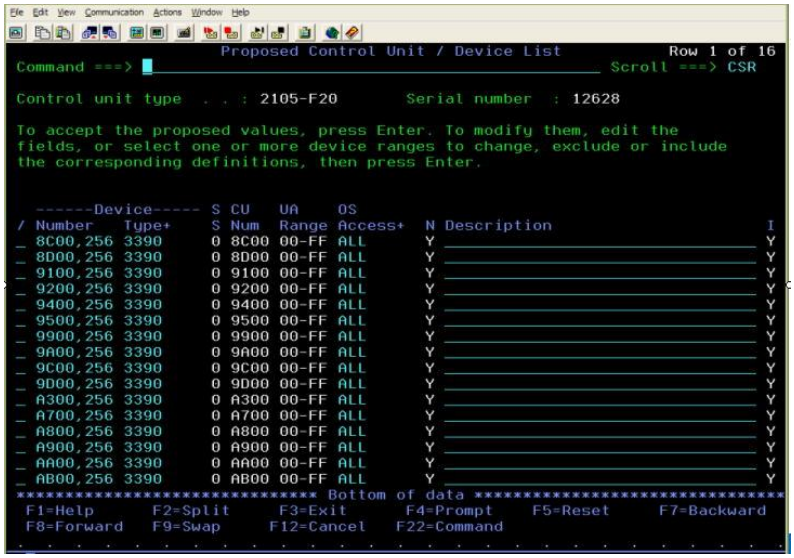
Redbooks IBM

Proposed List



Redbooks IBM

Accepting Configurations



Redbooks IBM

zDAC Hardware and Software Dependencies

- zDAC uses new capabilities in the 2817 processor for Fabric discovery
- Explores FICON channel attached to switches
- Dynamic I/O Enabled Partitions
- Up to date controller microcode
  - Tested with IBM 2107 Controller level R12p.9b090910b - Bundle 64.30.87.0
- Up to date switch microcode
  - Brocade Firmware Version: V6.2.0e
  - McData Firmware Version : 09.09.00
- All target systems must be at z/OS R1.12 or higher

Redbooks IBM

Extension to zHPF multitrack support - removing the 64k byte data transfer limit

- **Performance improvement for large block transfers**
  - Multi-track support, fully supporting 256 tracks of data
    Today limited to 64k data transfers in a single operation
- **Raising limit on the amount of data that can be transferred in a single operation, allowing the channel to operate at rates that are designed to fully exploit the bandwidth of FICON Express channel**
- **Customers will see higher throughputs at lower levels of channel utilization allowing more work to be done by fewer channels**
- **CHPID type: FC**
- **Features: FICON Express8, FICON Express4**
- **z/OS V1.12**
- **z/OS V1.10 and V1.11 with PTFs**
- **Transparent to control unit supporting zHPF**

Redbooks IBM

WWPN assignment to physical FCP channel on z196

- **Adds support to allow assignment of WWPNs to physical FCP ports**
  - Currently LIC assigns worldwide port names (WWPNs) to virtual FCP ports
- **Allows WWPN Prediction Tool to calculate WWPNs for virtual and physical ports ahead of system installation time**
- **Customer value - Can keep their SAN configuration after replacement of an FCP Channel card. Improved system installation. Simplifies FCP channel card replacement procedures**
- **CHPID type: FCP**
- **Features: FICON Express8, FICON Express4**
- **Transparent to operating systems**
- **WWPN tool available for download from Resource Link**

URL:https://www-304.ibm.com/servers/resourcelink/hom03010.nsf/pages/wwpnMain?OpenDocument&pathID=

Redbooks IBM

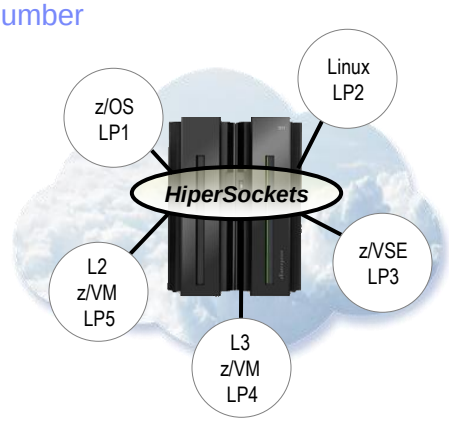
Three subchannel sets per LCSS for z196

- **Addition of a subchannel set of 64K devices to the existing two subchannel sets**
- **Value**
  - Extends the amount of addressable storage for the largest System z customers
  - Provides a means to provide consistent device address definitions to help simplify addressing scheme for congruous devices
    - Can utilize same device number in different subchannel sets
- **The first subchannel set (SS 0) allow definitions of any type of device as is allowed today, (i.e. bases, aliases, secondaries, and those other than disk that do not implement the concept of associated aliases or secondaries)**
- **Second and third subchannel sets (SS1 and SS2) now be designated for use for disk alias devices (of both primary and secondary devices) and/or Metro Mirror secondary devices only**
- **CHPID types: ESCON - CNC, FICON - FC (both native FICON and zHPF paths)**
- **Features: ESCON, FICON Express8, FICON Express4**
- **z/OS V1.12**
- **z/OS V1.10 and V1.11 with PTFs**

Redbooks IBM

z196 HiperSockets – doubled the number

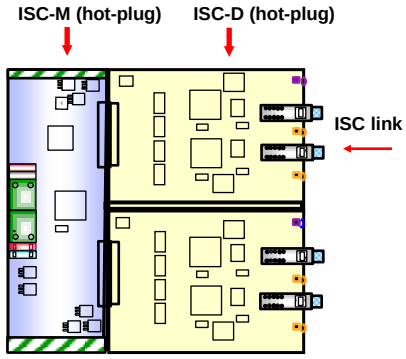
- **High-speed "intraserver" network**
- **Independent, integrated, virtual LANs**
- **Communication path – system memory**
- **Communication across LPARs**
  - Single LPAR - connect up to **32 HiperSockets**
  - 4096 communication queues
- **Support for multiple LCSS's & spanned channels**
- **Virtual LAN (IEEE 802.1q) support**
- **HiperSockets Network Concentrator**
- **Broadcast support for IPv4 packets**
- **IPv6**
- **HiperSockets Network Traffic Analyzer (HS NTA)**
- **No physical cabling or external connections required**
- **z/OS V1.12, z/OS V1.10 and V1.11 with PTFs**
- **z/VM V5.4 and later with PTFs**



Redbooks IBM

ISC-3 coupling links

- **InterSystem Channel-3 (ISC-3)**
  - ISC-3 links ordered in increments of one
  - Activated links balanced across features
- **Peer mode only – 2 Gbps**
  - #0217 (ISC-M), #0218 (ISC-D / ISC link)
  - Activate link - #0219
  - Four links per ISC-M
    Two links per ISC-D
  - Supports 9µ single mode fiber
- **Up to 48 links per machine**



ISC-3

LC Duplex connector

Redbooks IBM

Coupling Link Options

- **12x InfiniBand for high speed communication at medium distance**
  - New CHPID – CIB (Coupling using InfiniBand)
  - New 50 micron OM3 (2000 MHz-km) multimode fiber with MPO connectors
  - Up to 150m
- **1x InfiniBand (SDR or DDR)**
  - Unrepeated distanced up to 10 km
  - Repeated distances at up to 100 km when attached to a Dense Wavelength Division Multiplexer (DWDM) qualified by System z or when point-to-point with another z196 or z10
  - Speed may be auto-negotiated if the attached DWDM is capable of operating at SDR or DDR
    Supports SDR at 2.5 Gbps when connected to a DWDM capable of SDR speed
    Supports DDR at 5 Gbps when connected to a DWDM capable of DDR speed
  - 9 micron single mode fiber optic cables with LC Duplex connectors.
- **ISC-3 for extended distance over fiber optic cabling**
  - No change to current cabling
- **Internal Coupling channels (IC)**

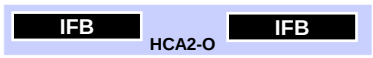
Note: The InfiniBand link data rates of 6 GBps, 3 GBps, 2.5 Gbps, or 5 Gbps do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. With InfiniBand coupling links, while the link data rate may be higher than that of ICB (12x InfiniBand (SDR or DDR) or ISC-3 (1x InfiniBand (SDR or DDR), the service times of coupling operations are greater, and the actual throughput may be less than with ICB links or ISC-3 links.

Redbooks IBM

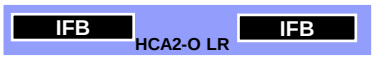
InfiniBand coupling links

Type	Speed	Distance	Fanout	Cabling
12x InfiniBand	6 or 3 GBps	150 meters	HCA2-O	50µ MM (OM3) fiber
1x InfiniBand	5 or 2.5 Gbps	10 km	HCA2-O LR	9µ SM fiber

Up to 16 CHPIDs – across 2 ports



Up to 16 CHPIDs – across 2 ports



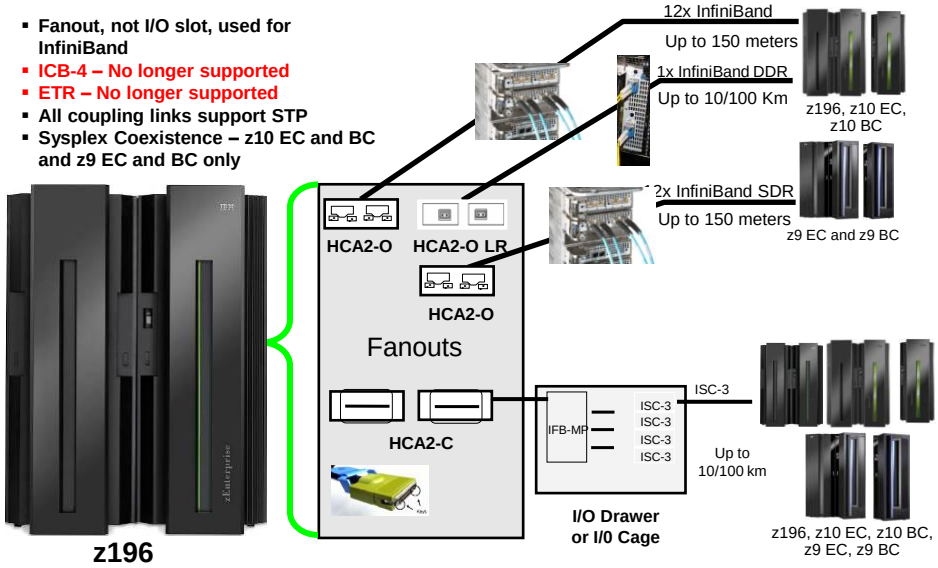
- **Ports exit from the front of a book**
  Does not use I/O card slots
- **12x InfiniBand – z196, z10, z9**
  - DDR at 6 GBps
    - z196 and z10
  - SDR at 3 GBps
    - z196 & z10 to z9
    - First addition to z9 is disruptive
    - z9 to z9 connection not supported
- **1x InfiniBand – z196 and z10 (not z9)**
  - DDR at 5 Gbps
  - SDR at 2.5 Gbps (if DWDM requires)

DDR = double date rate, SDR = single data rate

Redbooks IBM

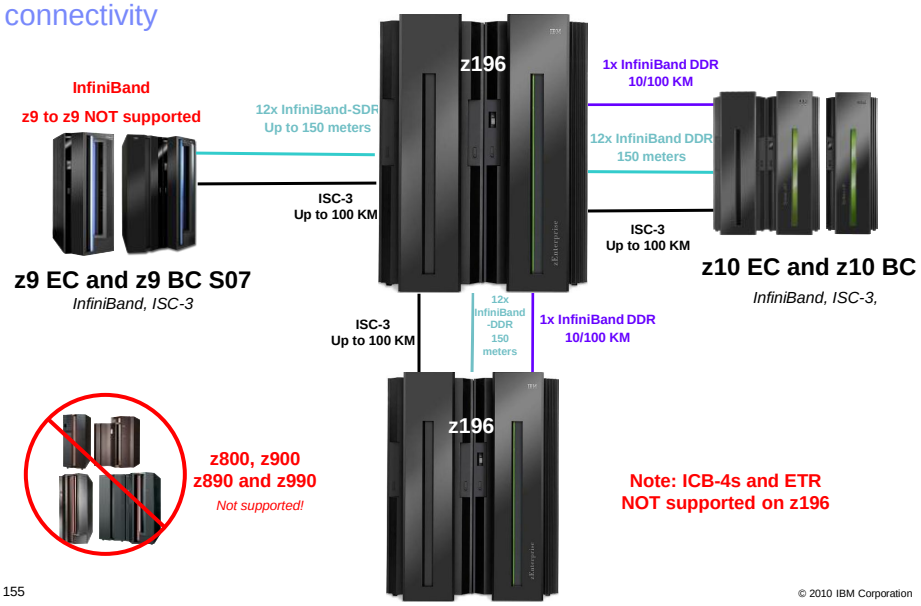
Coupling Links

- **Fanout, not I/O slot, used for InfiniBand**
- **ICB-4 – No longer supported**
- **ETR – No longer supported**
- **All coupling links support STP**
- **Sysplex Coexistence – z10 EC and BC and z9 EC and BC only**



Redbooks IBM

z196 Parallel Sysplex coexistence of Servers/CFs and coupling connectivity



Redbooks IBM

System z InfiniBand HCA Adapter Options – Summary

- **Host Channel Adapter Optical Fanouts – 2 ports per fanout**
  - z196, up to 16 HCA2-O fanouts
- **12X InfiniBand-SDR (3 GBps)**
  - z196 to z9
  - Uses 12 lanes for a total link rate of 3 GBps and is a point-to-point connection, maximum length of 150 meters
  - New 50 micron OM3 (2000 MHz-km) multimode fiber with MPO connectors
- **12X InfiniBand-DDR (6 GBps)**
  - z196 to z196
  - z196 to z10
  - Uses 12 lanes for a total link rate of 6 GBps and is a point-to-point connection, maximum length of 150 meters
  - New 50 micron OM3 (2000 MHz-km) multimode fiber with MPO connectors
- **1X InfiniBand-DDR LR (5 GBps)**
  - z196 to z196
  - z196 to z10
  - Uses one lane for a total link rate of 5 Gbps and supports an unrepeated distance of 10 km
  - Repeated distances at up to 100 km when attached to a Dense Wavelength Division Multiplexer (DWDM) qualified by z196
  - Speed may be auto-negotiated if the attached DWDM is capable of operating at SDR or DDR
    - Supports SDR at 2.5 Gbps when connected to a DWDM capable of SDR speed
    - Supports DDR at 5 Gbps when connected to a DWDM capable of DDR speed
  - 9 micron single mode fiber optic cables with LC Duplex connectors (same cable as ISC3)

Note: The InfiniBand link data rates of 6 GBps, 3 GBps, 2.5 Gbps, or 5 Gbps do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. With InfiniBand coupling links, while the link data rate may be higher than that of ICB (12x InfiniBand (SDR or DDR) or ISC-3 (1x InfiniBand (SDR or DDR), the service times of coupling operations are greater, and the actual throughput may be less than with ICB links or ISC-3 links.

Redbooks IBM

System z – CF Link Connectivity (Peer Mode only)

Connectivity Options	z196 ISC-3	z196 1x InfiniBand	z196 12x InfiniBand
z196 / z10 / z9 ISC-3 (RPQ 8P2197 – 20 km)	1 Gbps	N/A	N/A
z196 / z10 / z9 ISC-3 (10/100 km)	2 Gbps	N/A	N/A
z9 EC or z9 BC with 12x InfiniBandSDR (150 m)	N/A	N/A	3 GBps
z196 / z10 1x-InfiniBand DDR (10/100 km)	N/A	5 Gbps	N/A
z196 /z10 12x InfiniBand DDR (150 m)	N/A	N/A	6 GBps

- Theoretical maximum rates shown
- 9672 / z900 / z800 / z990 / z890 Connectivity to z196 Not Supported
- z196 does not support ICB connectivity
- 1x InfiniBand supports single data rate (SDR) at 2.5 Gbps when connected to a DWDM capable of SDR speed and double data rate (DDR) at 5 Gbps when connected to a DWDM capable of DDR speed or when point-to-point with another z196 or z10
- System z9 does NOT support 1x InfiniBand DDR or SDR InfiniBand Coupling Links

Note: The InfiniBand link data rates of 6 GBps, 3 GBps, 2.5 Gbps, or 5 Gbps do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. With InfiniBand coupling links, while the link data rate may be higher than that of ICB (12x InfiniBand (SDR or DDR) or ISC-3 (1x InfiniBand (SDR or DDR), the service times of coupling operations are greater, and the actual throughput may be less than with ICB links or ISC-3 links.

Redbooks IBM

z196 coupling link summary

Type	Description	Use	Link data rate	Distance	z196 Maximum	z196 Max links	z196 Max CHPIDs
IC (ICP)	Internal Coupling Channel	Internal communication	Internal speeds	NA	32	NA	128 CHPIDs
InfiniBand (CIB)	12x InfiniBand	z196 & z10 z196 & z10 to z9	6 GBps 3 GBps*	150 meters (492 feet)	32	32	
InfiniBand (CIB)	1x InfiniBand	z196 & z10	5 Gbps or 2.5 Gbps**	10 km unrepeated (6.2 miles) 100 km repeated	32		
ISC-3 (CFP)	InterSystem Channel-3	z196, z10, z9	2 Gbps	10 km unrepeated (6.2 miles) 100 km repeated	48	48	

- **Maximum of 16 InfiniBand fanouts are allowed, 2 links per fanout**
- **Maximum of 128 coupling CHPIDs (ICP, CIB, CFP) per server**
  Each InfiniBand link supports definition of multiple CIB CHPIDs, up to 16 per fanout

* z196 & z10 negotiate to 3 GBps when connected to a z9
** May negotiate to 2.5 Gbps when connected to a DWDM

Note: ICB-4 is not supported on z196

Note: The InfiniBand link data rates of 6 GBps, 3 GBps, 2.5 Gbps, or 5 Gbps do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload.

Redbooks IBM

Maximum Coupling Links and CHPIDs

Server	1x InfiniBand	12x InfiniBand	IC	ICB-4	ICB-3	ISC-3	Max**** External Links	Max***** Coupling CHPIDs
z196	32* M15 - 16	32* M15 - 16	32	N/A	N/A	48	80	128
z10 EC	32** E12 - 16	32* E12 - 16	32	16** (32/RPQ)	N/A	48	64	64
z10 BC	12***	12***	32	12***	N/A	48	64	64
z9 EC	N/A	16 S08 - 12	32	16	16	48	64	64
z9 BC	N/A	12	32	16	16	48	64	64

* Maximum of 32 InfiniBand links of all types on System z196.
** Maximum of 32 InfiniBand links of all types + ICB4 links on System z10 EC. ICB-4 not supported on Model E64
*** Maximum of 12 InfiniBand links of all types + ICB4 links on System z10 BC.
**** Maximum external links is the maximum total number of physical link ports (Does not include IC)
***** Maximum coupling CHPIDs defined in IOCDS includes IC and multiple CHPIDs defined on InfiniBand physical links.

Redbooks IBM

Fibre Channel Physical Interface (FC-PI-4) standard (Rev. 8.00)

- **Applies to FICON Express8 (2, 4, 8 Gbps), FICON Express4 (1, 2, 4 Gbps) FICON Express2 (1, 2 Gbps), and FICON Express (1, 2 Gbps) features**
- **CHPID types FC (FICON, zHPF, CTC) and FCP (Fibre Channel Protocol)**
- **Unrepeated distances in kilometers (km), meters (m), and feet (ft)**

	1 Gbps		2 Gbps		4 Gbps		8 Gbps		10 Gbps ISLs	
Fiber Core (µ) Light source	Distance meters feet	* Link loss budget	Distance meters feet	* Link loss budget	Distance meters feet	* Link loss budget	Distance meters feet	* Link loss budget	Distance meters feet	* Link loss budget
9µ SM LX laser	10 km 6.2 miles	7.8 dB	10 km 6.2 miles	7.8 dB	10 km 6.2 miles	7.8 dB	10 km 6.2 miles	6.4 dB	10 km 6.2 miles	6.0 dB
9µ SM LX laser	4 km # 2.5 miles	4.8 dB #	4 km # 2.5 miles	4.8 dB #	4 km # 2.5 miles	4.8 dB #	N/A	N/A	N/A	N/A
50µ MM OM3 SX laser	860 m 2822 ft	4.62 dB	500 m 1640 ft	3.31 dB	380 m 1247 ft	2.88 dB	150 m 492 ft	2.04 dB	300 m 984 ft	2.6 dB
50µ MM OM2 SX laser	500 m 1640 ft	3.85 dB	300 m 984 ft	2.62 dB	150 m 492 ft	2.06dB	50 m 164 ft	1.68 dB	82 m 269 ft	2.3 dB
62.5µ MM OM1 SX laser	300 m 984 ft	3.00 dB	150 m 492 ft	2.10 dB	70 m 230 ft	1.78 dB	21 m 69 ft	1.58 dB	33 m 108 ft	2.4 dB

OM1 62.5 micron at 200 MHz-km, OM2 50 micron at 500 MHz-km, OM3 50 micron at 2000 MHz-km
Inter-Switch Links (ISLs) is the link between two FICON directors; FICON features do not operate at 10 Gbps
* The link loss budget is the channel insertion loss as defined by the standard.
\# This distance and dB budget applies to FICON Express4 4KM LX features

Redbooks IBM

Coupling links using InfiniBand Trade Association standard

- **12x Multi-fiber Push-On (MPO) connector**
- **24-fiber cable with Duplex 12x MPO connectors**
  - 12 fibers for transmit and 12 fibers for receive

		3 GBps (SDR)		6 GBps (DDR)	
Fiber Core (µ) (Light source)	Fiber Bandwidth @ wavelength	Unrepeated distance	Optical passive loss	Unrepeated distance	Optical passive loss
50µ MM (SX laser)	2000 MHz-km @ 850 nm	150 meters 492 feet	2.06 dB	150 meters 492 feet	2.06 dB

1. 12x InfiniBand SDR links operating at 3 GBps (2.5 Gbps per lane) are used to connect a System z10 to a System z9 or z196 to z10
2. 12x InfiniBand DDR links operating at 6 GBps (5.0 Gbps per lane) are used to connect System z10 servers or z196 to z10

Redbooks IBM

Coupling links using InfiniBand Trade Association standard

- **1x InfiniBand**
- **LC Duplex connector**
- **One pair of fiber (1x) - one fiber for transmit and one fiber for receive**

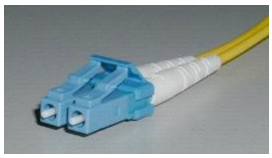
		2.5 Gbps (SDR)		5 Gbps (DDR)	
Fiber Core (µ) (Light source)	@ wavelength	Unrepeated distance	Optical passive loss	Unrepeated distance	Optical passive loss
9µ SM LR laser	@ 1310 nm	10 km 6.2 miles	5.66 dB	10 km 6.2 miles	5.66 dB

- 1x InfiniBand operating at 2.5 Gbps or 5 Gbps is exclusive to System z10 and z196 servers
- 1x InfiniBand is for use within a building only. It is not intended to go outside the building. All attachments to an outside cable plant (including public "dark fiber") are supported only through a patch panel or Wavelength Division Multiplexer (WDM) product.

Redbooks IBM

1x InfiniBand 9/125 micrometer single mode fiber optic cabling

LC Duplex connector



- **Cables available from:**
  - IBM Global Technology Services (GTS)
  - Your preferred cable provider
- Fiber core – 9µ single mode
- Light source – LR laser @ wavelength: @ 1310 nm

Note: the fiber optic cabling is the same
as used with ISC-3, FICON LX, 10 GbE LR, and GbE LX

LC Duplex harness

Redbooks IBM

Supported 12x InfiniBand DDR cable lengths
OM3 50/125 micrometer multimode fiber optic cabling

- **Cables available from:**
  - IBM Global Technology Services (GTS)
  - Anixter www.anixter.com/
  - Computer Crafts Inc. www.computer-crafts.com/
  - Tyco www.tycoelectronics.com/
  - Fujikura www.fujikura.com/
- **Fiber core – 50u multimode**
- **Light source – SX laser**
- **Fiber bandwidth @ wavelength: 2000 MHz-km @ 850 nm**
- **IBM cable part numbers highly recommended**

Item Description	Cable IBM P/N	Cable Length Meters	Cable Length Feet	Connector Type
Duplex 24-fiber cable Assembly	41V2466	10.0 m	32.8 f	MPO - MPO
Duplex 24-fiber cable Assembly	15R8844	13.0 m	42.7 f	MPO - MPO
Duplex 24-fiber cable Assembly	15R8845	15.0 m	49.2 f	MPO - MPO
Duplex 24-fiber cable Assembly	41V2467	20.0 m	65.6 f	MPO - MPO
Duplex 24-fiber cable Assembly	41V2468	40.0 m	131.2 f	MPO - MPO
Duplex 24-fiber cable Assembly	41V2469	80.0 m	262.4 f	MPO - MPO
Duplex 24-fiber cable Assembly	41V2470	120.0 m	393.7 f	MPO - MPO
Duplex 24-fiber cable Assembly	41V2471	150.0 m	492.1 f	MPO - MPO
Duplex 24-fiber cable Assembly	42V2083	Custom	N/A	MPO - MPO

Redbooks IBM

z196 I/O connectivity summary

Features	Avail	Maximum # of features	Maximum connections	Increments per feature	Purchase increments
ESCON	Yes	16	240 channels	15 active	4 channels
FICON **FICON Express8** **FICON Express4**	 Yes CF	72*	288 channels	4 channels	4 channels
ISC-3	Yes	12	48 links	4 links	1 link
OSA **OSA-Express3** **OSA-Express2**** **GbE** **1000BASE-T**	 Yes CF CF	24	96 ports	 2 ports – 10 GbE 4 ports – GbE 4 ports – 1000BASE-T	1 feature
Crypto **Crypto Express3**	 Yes	8	16 PCI adapters	2 PCI adapters	2/1*** feature

* The standard z196 offering is limited to 72 features. Customers who have a good business need to have >72 features, on special approval, RPQ 8P2506 will be made available.
** OSA-Express2 10 GbE LR is not supported as a carry forward (CF)
*** Two Crypto features on initial order; one feature thereafter

Redbooks IBM

z196 I/O connectivity summary

Features	Minimum # of features	Maximum # of features	Maximum connections	Increments per feature	Purchase increments
16-port ESCON	0 (1)	16	240 channels	16 channels 1 reserved as a spare	4 channels
FICON Express8	0 (1)	72*	228 channels	4 channels	4 channels
FICON Express4	0 (1)	72*	228 channels	4 channels	4 channels
ISC-3	0 (1)	12	48 links (2)	4 links	1 link
HCA2-O LR (1x)	0 (1)	16	32 links (2)	2 links	2 links
HCA2-O (12x)	0 (1)	16	32 links (2)	2 links	2 links
OSA-Express3	0	24	48/96 ports	2 or 4	2 ports/ 4 ports
OSA-Express2**	0	24	48 ports	1 or 2	1 ports / 2 port
Crypto Express3	0	8	16 PCI-X adapters	2 PCI-X adapters	2 PCI-X adapters (3)

1. Minimum of one I/O feature (ESCON, FICON) or Coupling Link (InfiniBand, ISC-3) required.
2. The maximum number of InfiniBand features is 16 or 32 links. There is a maximum of 128 coupling link CHPIDs per server (ICs, ISC-3 links, and IFBs)
3. Initial order of Crypto Express3 is 4 PCI-X adapters (two features). Each PCI-X adapter can be configured as a coprocessor or an accelerator.

* The standard z196 offering is limited to 72 features. Customers who have a good business need to have >72 features, on special approval, RPQ 8P2506 will be made available.
** Available only when carried forward on an upgrade from z9 or z10. OSA-Express2 10 GbE LR is not supported as a carry forward (CF)

Redbooks IBM

z196 Capacity on Demand

- *CoD Offerings*
- *CoD provisioning*
- *CoD Enhancements*
- *CoD Auto renwal*
- *System z CoD summary*

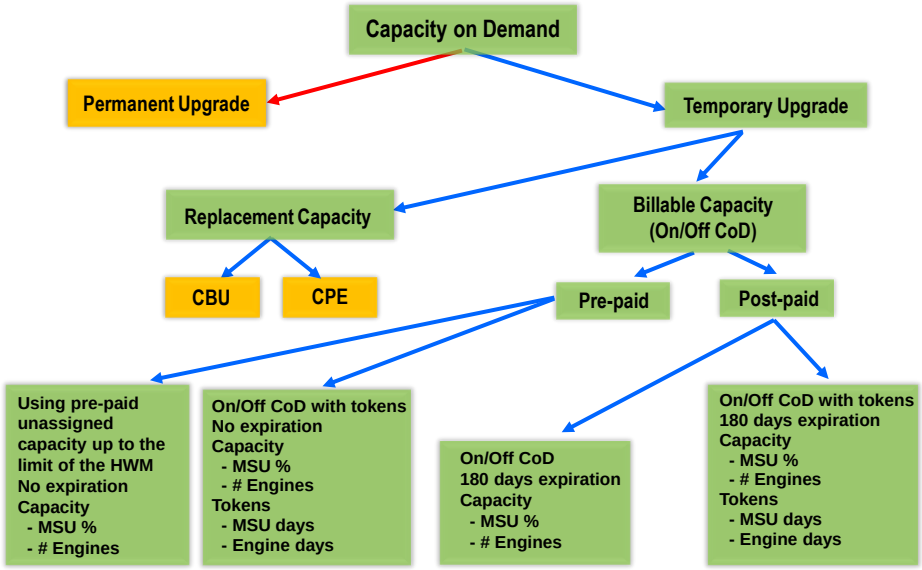
Redbooks IBM

IBM System z Capacity on Demand Offerings (review)

- 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 (three 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

Redbooks IBM

z196 – Basics of CoD



Redbooks IBM

z196 CoD Resources, Time elements and Tokens Summary

Resources	CBU	CPE	On/Off CoD	Remarks
CP	CP	CP	up to 100% more MSU CP capacity	
Specialty	zIIP, zAAP, ICF, IFL, SAP			
Time Elements	CBU	CPE	On/Off CoD	Remarks
Test Duration	10 days	NA	NA	
Real activations	90 days	3 days	Post-paid – Unlimited Prepaid - Limited	
Grace Period	2 days	N/A	One hour	Auto deactivation upon end of grace period
Expiration Date	1-5 years	No Expire	180 days	Auto deactivation upon expiration
Tokens	CBU	CPE	On/Off CoD	Remarks
Number of test	5,10 or 15	0	1	On/Off CoD tests are managed via a separate record
Number of real activation	1	1	Post-paid – Unlimited Prepaid - Limited	

Limiting On/Off CoD capacity via resource tokens

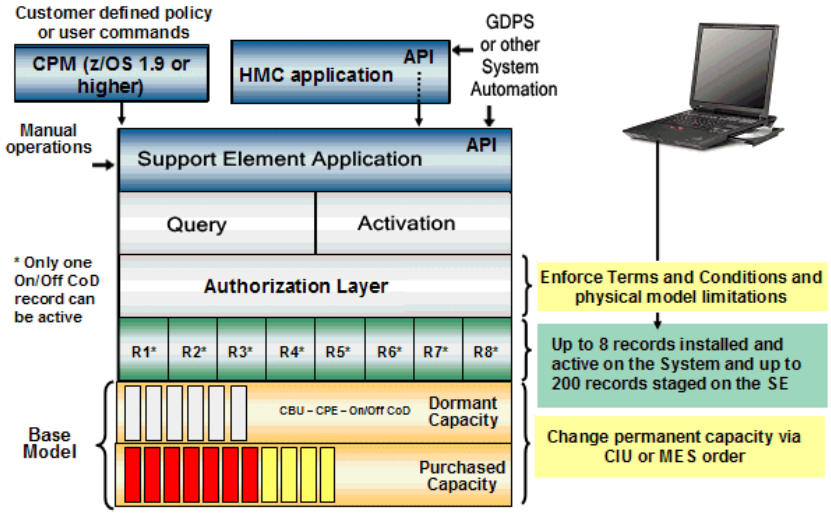
- Helps bound the hardware costs associated with using On/Off CoD
  - Allows PO (Purchase Order) customers to assign tokens to limit their financial obligation
- Optional **Prepaid** or **Post-paid**
  - When **prepaid**, you are billed for the total amount of resource tokens contained within the On/Off CoD record when the record is downloaded
  - When **post-paid**, the total billing against the On/Off Cod record is limited by the total amount of resource tokens contained within the record
- For **CP capacity**, a resource token represents an amount of processing capacity that will result in 1 MSU of SW cost for 1 day - **an MSU-day**
- For **specialty engines**, a resource token represents activation of 1 engine of that type for 1 day – an IFL-day, a zIIP-day or a zAAP-day **(specialty engine-day)**
- Tokens are decremented at the end of each 24 hour period. If for any engine type, there are insufficient tokens for the next 24 hour period the entire record is deactivated
  - Customers notified when token pool falls below 5 days remaining (again at 24 hours)
  - Customers can order replenishment records to add tokens

z196 – Running Production workload with CBU Test

- Customers may now execute production workload during a CBU test provided:
  - *An amount of System z production workload Capacity equivalent to the CBU Upgrade is shut down or otherwise made unusable by the Customer for the duration test*
  - *The appropriate contracts are in place. All new CBU contract documents contain these new CBU Test terms*
  - *Existing CBU customers will need to execute IBM Customer Agreement Amendment for IBM System z Capacity Backup Upgrade Tests*
  - *Applies to any System z capable of CBU*

Redbooks IBM

z196 Capacity on Demand Provisioning Architecture



Redbooks IBM

z196 Capacity on Demand Enhancements

z10	z196
Separate orders for purchase of unassigned engines	Unassigned engine purchase via CIU
On/Off CoD records must be replenished manually	Auto replenishment of On/Off CoD records
CoD records staged on machine deliver	Manufacturing install of up to 4 CoD records with system ship.
No On/Off CoD administrative test	On/Off CoD Administrative tests

Redbooks IBM

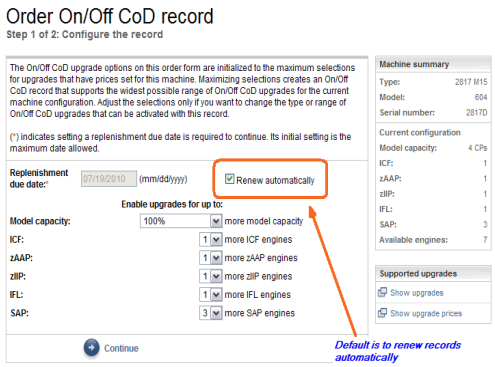
Capacity on Demand – Auto renewal of On/Off CoD records

- **Auto-renewal of On/Off CoD records**
  - With System z10, IBM introduced an expiration date on On/Off CoD records
    - New On/Off CoD records gave considerable flexibility, but increased IBM's exposure
    - Date was introduced to limit IBM's asset exposure
    - All On/Off CoD records expire within 180 days of download
    - When records expire, any active capacity is automatically deactivated on the SE
  - Expiration of records has met with universal customer complaints
  - For z196: Resource Link will monitor all installed On/Off CoD records
    - Process will only support installed records
    - Records staged on the SE or on RETAIN and not installed will not be auto-renewed
    - Auto-renewal will only be available to machines that can connect to RETAIN – not available to RPQ customers.
  - Every 90 days, Resource Link will generate a replenishment record for each installed record that will move the expiration date out 180 days
    - Machine must have connected to back to IBM in past 14 days with updated VPD / Billing history
    - Record must be "enabled" for auto-renewal
    - IBM Reserves right to disable auto-renewal if customer does not meet contracts terms and conditions
  - Next time the machine connects to RETAIN, replenishment record is pushed to the machine and installed

Redbooks IBM

Auto renewal of On/Off CoD records

- **Allow for the automatic renewal of installed records - this function is intended for use with On/Off CoD, to automatically update the expiration date for "customers in good standing".**
  - Customer in good standing = consistent transmittal of VPD and regular and timely payment of bills. This would be managed at the Resource Link level.
- **The request is to auto-replenish the expiration date of installed On/Off CoD records, there is no intent to replenish tokens automatically, only the expiration date.**
- **Staged records are not subject to automatic replenishment.**
- **Participation in this program is a customer option.**
- **Not available for records on RPQ machines**
- **Resource Link will "disable" auto replenishment on a per machine profile basis**
  - Customer does not pay bills
  - Customer drops off IBM Maintenance



Redbooks IBM

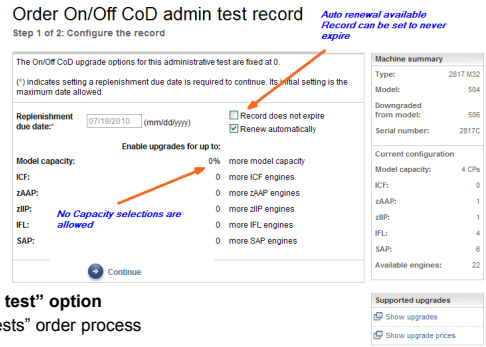
Capacity on Demand – On/Off CoD Administrative Test

- **On/Off CoD Administrative Test**
  - IBM Introduced On/Off CoD administrative tests on the System z9
    - On/Off CoD records that could be activated, but did not activate any capacity
    - Records allowed for testing of processes / procedures without any resulting HW or SW charges
  - With new Capacity on Demand architecture on System z10, administrative test function was dropped.
    - New reusable record was thought to eliminate need for administrative test capability
    - New record structure did not readily support administrative tests
  - For z196, Resource Link reintroduces the "Order Administrative On/Off CoD test" option
    - Standard order flow, including approval steps – "tests" order process
    - Default will be for 180 day expiring record
    - Customer optionally may select renewal
    - Customer optionally may select "record does not expire" (unique to admin test)
  - Resource Link generates On/Off CoD LICCC request
    - All capacity levels set to zero
    - If non-expiring, no expiration date set
    - "Training record" bit set – instructs SE to not allow any capacity activations
  - SE changes to support admin test
    - Training bit results in no engine upgrade choices (0% records)
    - Record can be "activated" without actually changing activation levels

Redbooks IBM

On/Off CoD Administrative Test

1. **Customers use for operations training**
   - Allows multiple tests to train a number of personnel
   - Allows customers to do periodic retraining
2. **API testing**
   - Customers using APIs needed way to test their code without running up charges



- **Resource Link provides "Order On/Off CoD admin test" option**
  - Standard order flow, including approval steps – "tests" order process
  - Default will be for 180 day expiring record
  - Customer optionally may select auto-replenishment
  - Customer optionally may select "record does not expire" (unique to admin test)
- **Resource Link generates On/Off CoD LICCC request**
  - All capacity levels set to zero and will not allow activation of any capacity
  - If non-expiring, no expiration date set

Redbooks IBM

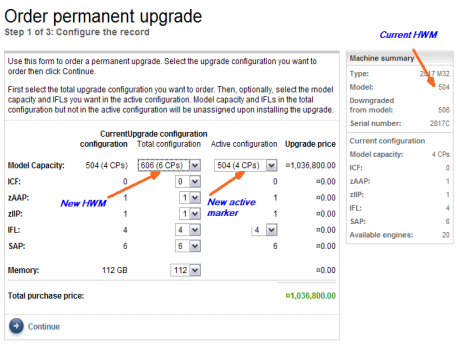
Capacity on Demand – additional Enhancements

- **Other CoD Enhancements**
  - Purchase of unassigned CP or IFL capacity
    New ordering option on Resource Link
    Has been available via eConfig (Standard MES process) before, but not via CIU
    Allows customer to create an upgrade that move the HWM without moving the active capacity marker

Redbooks IBM

Purchase of permanent unassigned engines

- **Purchase of unassigned CP or IFL capacity**
  - New ordering option on Resource Link
  - Has been available via eConfig (Standard MES process) before, but not via CIU
  - Allows customer to create an upgrade that moves the HWM without moving the active capacity marker
- **On z10, via CIU, you can: add active engines; activate already purchased, but inactive engines; deactivate engines**
  - All transactions must be separate and distinct
  - Each CIU order takes 2 hours
  - VPD must be transmitted and processed between transactions
  - Net: to purchase and deactivate and engine takes approx: 4-6 hours
- **Deactivation of engines is a "priced" feature**
- **Customers: "I can do this via sales, why can't I do this via CIU?"**



- **On permanent engine purchased, Resource Link will provide the customer with the options to adjust the active capacity mark**
- **Customers will be billed for purchase of new engines**
- **No charges made for deactivation of the capacity**
- **HW Billing request will contain necessary Feature Codes to indicate new capacity level**

Redbooks IBM

Pre-Installed Records

- With z10, the manufacturing process puts temporary records that have been ordered with the system into the staging area, but not automatically install them.
- The record installation (which is a pre-req for being able to activate the record) is not performed until the Service Representative at the customer site does it (guided by install instructions) during the initial install of the machine.
  - This step is sometimes not executed during initial machine install and then causes problems because customers cannot find their records in the installed area.
- With z196, system will come with records pre-installed
  - Limitations: Manufacturing will only install 1st 4 records

Redbooks IBM

System z Capacity Resource Availability by Server

Capacity on Demand	Server
Capacity BackUp	z196, z10 EC, z9 EC, z9 BC, z990, z890, z900, z800
Capacity Upgrade on Demand	z196, z10 EC, z9 EC, z9 BC, z990, z890, z900, z800
Customer Initiated Upgrade	z196, z10 EC, z9 EC, z9 BC, z990, z890, z900, z800
On/Off Capacity on Demand	z196, z10 EC, z9 EC, z9 BC, z990, z890
Capacity for Planned Event	z196, z10 EC, z10 BC

	Capacity BackUp	Capacity Upgrade on Demand	Customer Initiated Upgrade – for ordering	On/Off Capacity on Demand	Capacity for Planned Event
z900	Yes (CP only)	Yes (CP, I/O, IFL, ICF, Memory)	Yes (CP, IFL, ICF, Memory)	No	No
z800	Yes (CP only)	Yes (CP, I/O, IFL, ICF)	Yes (CP, IFL, ICF)	No	No
z890/z990	Yes (CP only)	Yes (CP, I/O, IFL, ICF, zAAP, Memory*)	Yes (CP, IFL, ICF, zAAP, Memory*)	Yes (CP, IFL, ICF, zAAP)	No
z9	Yes (CP, ICF, IFL zAAP, zIIP)	Yes (CP, I/O, IFL, ICF, zAAP, zIIP, Memory*)	Yes (CP, IFL, ICF, zAAP, zIIP, Memory*)	Yes (CP, IFL, ICF, zAAP, zIIP)	No
z10/ z196	Yes (CP, ICF, IFL zAAP, zIIP, SAP)	Yes (CP, I/O, IFL, ICF, zAAP, zIIP, SAP, Memory*)	Yes (CP, IFL, ICF, zAAP, zIIP, SAP, Memory*)	Yes (CP, IFL, ICF, zAAP, zIIP, SAP)	Yes (CP, IFL, ICF, zAAP, zIIP SAP)

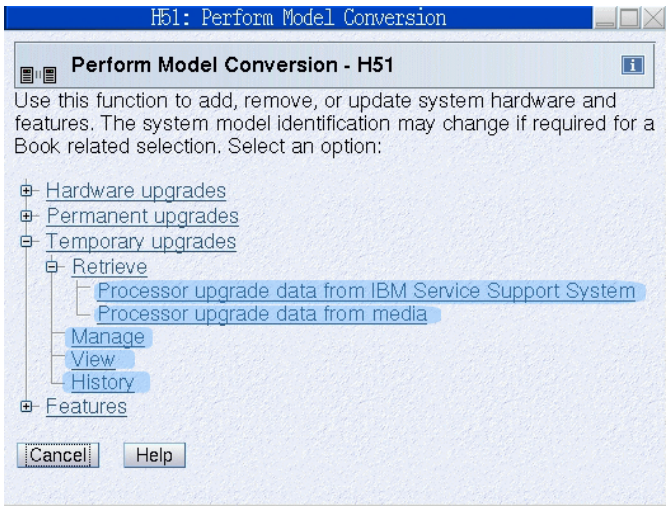
* Not supported for some memory upgrades

Note: Upgrades are non-disruptive only where there is sufficient hardware resource available and provided pre-planning has been done

CBU characteristics on z196

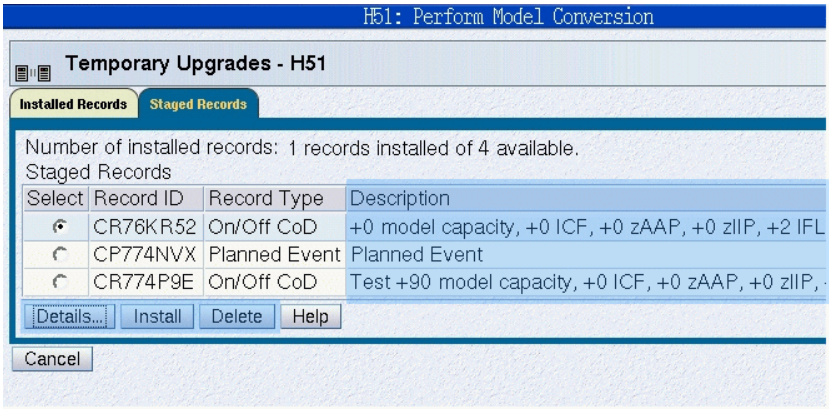
- There is no requirement for a CBU password
- CBU upgrades can be from a sub-capacity model to another sub capacity
- Multiple CBU records can be installed and staged at the same time
- The sum of resources of installed CBU records may exceed physical capacity
- Resource activation of an installed CBU record can be incremental manner
- When a CBU record is deactivated, the record stays in the installed "slot“
- It is possible to replenish a CBU record
- The activation period for a test is 10 days with a 2 day grace period
- A real activation is set to 90 days with 2 days grace period
- When the CBU record has been active for it's allotted number of days, the CPU no longer gets kneecapped
- After the 2 day grace period is over, system will automatically deactivate CBU resources
- The CBU record has an expiration date
- More then 1 CBU record can be ordered

Perform Model Convertion



Redbooks IBM

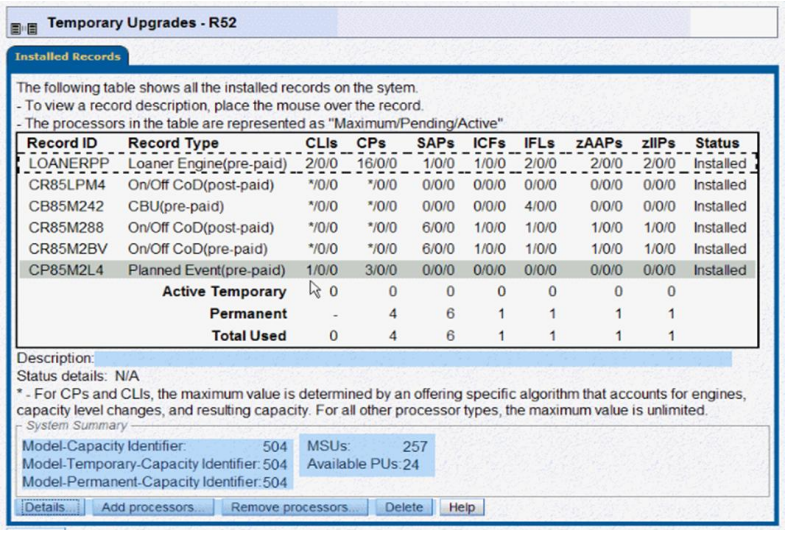
Managed Panel



Note: When a new machine is ordered with the CBU feature up to a total of 4 TERs will be installed. Any additional TERs will be staged

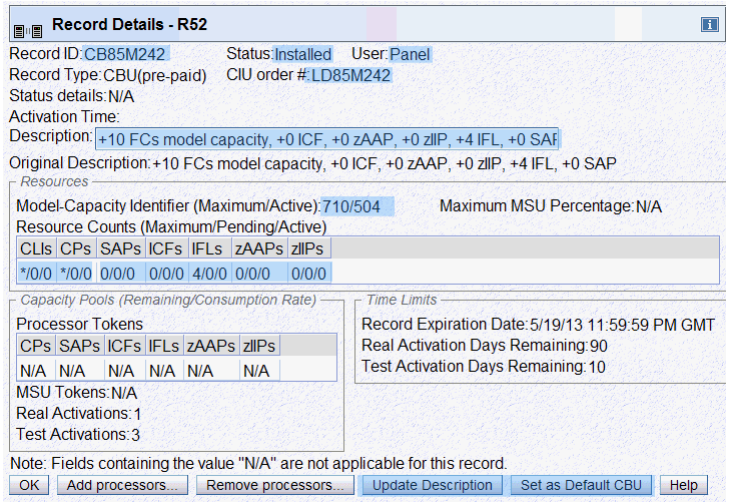
Redbooks IBM

Temporary Upgrade



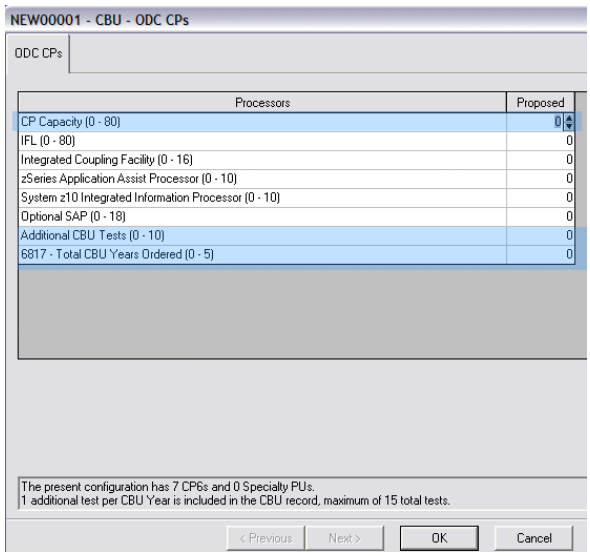
Redbooks IBM

Record Details



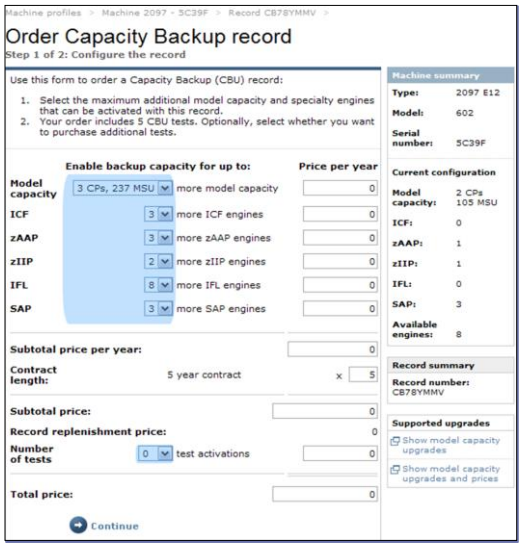
Redbooks IBM

eConfig CBU Request



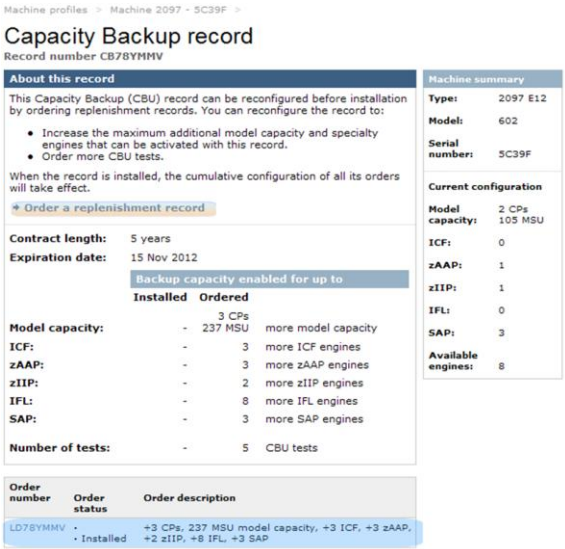
Redbooks IBM

Order Capacity Backup Record



Redbooks IBM

Capacity Backup Record



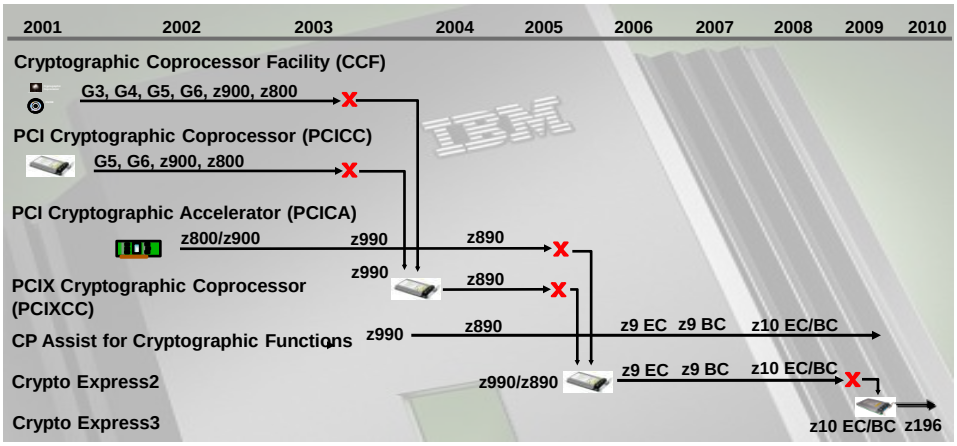
Redbooks IBM

z196 Cryptography

- *System z Crypto history*
- *Protect key CPACF*
- *Crypto Express 3*
- *z196 new cryptographic features*
  z196 cryptographic enhancements
- *TKE*
- *Certification for system z*

Redbooks IBM

System z Crypto History



- **Cryptographic Coprocessor Facility – Supports "Secure key" cryptographic processing**
- **PCICC Feature – Supports "Secure key" cryptographic processing**
- **PCICA Feature – Supports "Clear key" SSL acceleration**
- **PCIXCC Feature – Supports "Secure key" cryptographic processing**
- **CP Assist for Cryptographic Function allows limited "Clear key" crypto functions from any CP/IFL**
  - NOT equivalent to CCF on older machines in function or Crypto Express2 capability
- **Crypto Express2 – Combines function and performance of PCICA and PCICC**
- **Crypto Express3 – PCI-e Interface, additional processing capacity with improved RAS**

Redbooks IBM

CP Assist for Cryptographic Function

- The CP Assist for Cryptographic Function (CPACF) is available on every Processor Unit defined as a Central Processor (CP) and or Integrated Facility for Linux (IFL).
- The CPACF provides a set of symmetric cryptographic functions that enhance the encryption/decryption performance of clear-key operations.
- Cryptographic keys must be protected by the application system, as these keys are provided in the clear-key form to the CPACF.
- CPACF must be explicitly enabled, using a no-charge enablement feature (#3863), except for SHA-1, SHA-224, SHA-256, SHA-384 and SHA-512, which are shipped enabled with each server.
- The CPACF function is supported by z/OS, z/VM, z/VSE and Linux on System z.

Redbooks IBM

CP Assist for Cryptographic Function (CPACF)

- **For data privacy and confidentially**
  - Data Encryption Standard (DES)
  - Triple Data Encryption Standard (TDES)
  - Advanced Encryption Standard (AES) for 128-bit, 192-bit and 256-bit keys
- **For data integrity**
  - Secure Hash Algorithms
  - SHA-1: 160 bit
  - SHA-2: 224, 256, 384 and 512 bit
- **For key generation**
  - Pseudo Random Number Generation (PRNG)R
  - Random Number Generation Long (RNGL) -- 8 bytes to 8096 bytes
  - Random Number Generation Long (RNG) with up to 4096-bit key RSA support
- For message authentication
  - Single-key MAC
  - Double-key MAC
- For Personal Identification Number (PIN) processing
  - PIN generation, verification and translation functions

Protected key CPACF – a blending clear key and secure key cryptography

Clear versus Secure Keys
The security of encryption relies upon keeping the value of the key a secret. A secure key is simply a key that has been encrypted under another key, usually the master key. A clear key is a key that has not been encrypted under another key and, therefore has no additional protection within the cryptographic environment.

The CPACF enhancement is designed to:

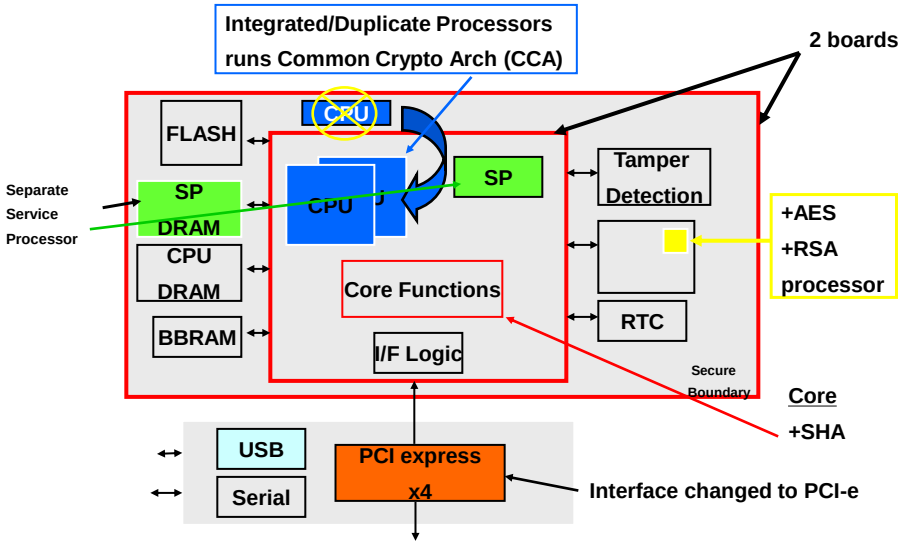
- Help facilitate the continued privacy of key material when used by the CPACF for high performance data encryption.
- Provide additional security for cryptographic keys.
- Leverage the unique z/Architecture and helps to ensure that key material is not visible to applications or operating systems when used for encryption operations.
- Provide significant throughput for large volumes of data and low latency for small blocks of data.
- Enhance the information management tool, IBM Encryption Tool for IMS and DB2 Databases, by improving performance for protected key applications.

Crypto Express3 – Hardware Design

- **Crypto Express3**
  - One or two Coprocessor features. One Coprocessor feature for z10 BC only
  - Each processor can be defined as a Cryptographic Coprocessor or an Accelerator
  - A minimum of two features must be ordered
- **Integrated and duplicated Processors into field-programmable gate array (FPGA) to support Common Cryptographic Architecture (CCA)**
- **Specialized hardware to perform DES, TDES, AES, RSA, SHA1 and SHA-2 cryptographic operations**
  - **SHA-2**
    SHA-2 (256-bit) hardware based on FIPS PUB 180-2 Secure Hash Standard
    SHA-256 is intended to provide 128 bits of security against collision attacks
  - **RSA**
    Two 2048-bit RSA engines are designed to provide improved performance for symmetric and asymmetric operations
- **Separate Service Processor**
- **PCI-express (PCI-e)**
- **Designed to provides a state-of–the art tamper sensing and responding, programmable hardware to protect the cryptographic keys and sensitive custom applications**
- **The tamper-resistant hardware security module, which is contained within the Crypto Express3, is designed to meet the FIPS 140-2 Level 4 security requirements for hardware security modules**

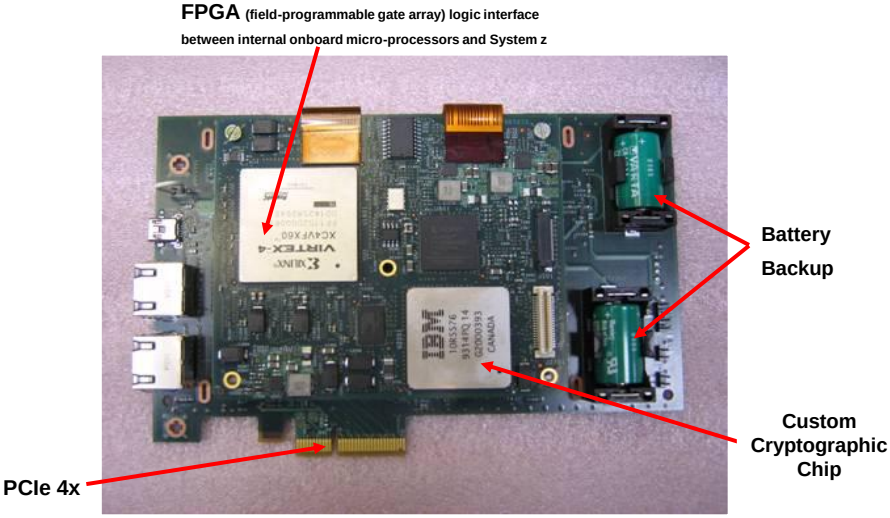
Redbooks IBM

Crypto Express3 (logical view)



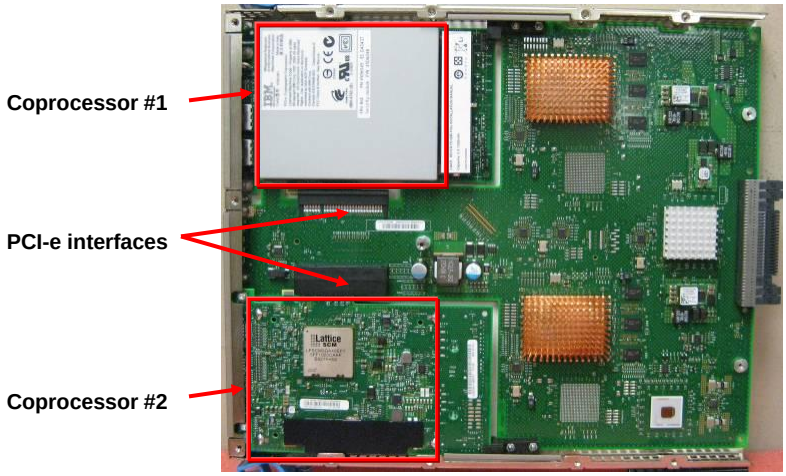
Redbooks IBM

Crypto Express3 Internal View of single Coprocessor



Redbooks IBM

Crypto Express3 2-P Physical View



Redbooks IBM

Crypto Express3 feature highlights

- Dynamic power management to maximize RSA performance while keeping within temperature limits of the tamper-responding package
- Virtualization: all logical partitions (LPARs) in all Logical Channel Subsystems (LCSSs) have access to the Crypto Express3 feature, up to 32 LPARs per feature
- Designed for improved reliability, availability and serviceability (RAS)
- Secure code loading that enables the updating of functionality while installed in application systems
- Executes its cryptographic functions asynchronously to a Central Processor (CP) operation

Continued

Redbooks IBM

Crypto Express3 features highlights

- **Lock-step-checking of dual CPUs for enhanced error detection and fault isolation of cryptographic operations performed by coprocessor when a PCI-E adapter is defined as a coprocessor**
- **Dynamic addition / configuration of cryptographic features to logical partitions without an outage**
- **Updated cryptographic algorithms used in loading the Licensed Internal Code (LIC) with the TKE workstation to keep in step with current recommendations for cryptographic strength**
- **Support for smart card applications using Europay, MasterCard Visa specifications**
- **Health Monitoring of mesh, temperature, voltage, soft tamper and low battery. Monitoring is internal to the card. Cannot be accessed or exploited by the customer**

Redbooks IBM

z196 New and exclusive cryptographic capabilities

- **Elliptic Curve Cryptography Digital Signature Algorithm (ECC)**, an emerging public key algorithm expected eventually to replace RSA cryptography in many applications. ECC is capable of providing digital signature functions and key agreement functions. The new CCA functions provide ECC key generation and key management and provide digital signature generation and verification functions compliance with the ECDSA method described in **ANSI X9.62 "Public Key Cryptography for the Financial Services Industry: The Elliptic Curve Digital Signature Algorithm (ECDSA)"**. ECC uses keys that are shorter than RSA keys for equivalent strength-per-key-bit; RSA is impractical at key lengths with strength-per-key-bit equivalent to AES-192 and AES-256. So the strength-per-key-bit is substantially greater in an algorithm that uses elliptic curves.
- **ANSI X9.8 PIN security** which facilitates compliance with the processing requirements defined in the new version of the **ANSI X9.8 and ISO 9564 PIN Security Standards** and provides added security for transactions that require Personal Identification Numbers (PIN).
- **Enhanced Common Cryptographic Architecture (CCA),** a Common Cryptographic Architecture (CCA) key token wrapping method using Cipher Block Chaining (CBC) mode in combination with other techniques to satisfy the key bundle compliance requirements in standards including **ANSI X9.24-1** and the recently published **Payment Card Industry Hardware Security Module (PCI HSM) standard**.
- **Secure Keyed-Hash Message Authentication Code (HMAC),** a method for computing a message authentication code using a secret key and a secure hash function. It is defined in the standard **FIPS 198, "The Keyed-Hash Message Authentication Code"**. The new CCA functions support HMAC using SHA-1, SHA-224, SHA-256, SHA-384, and SHA-512 hash algorithms. The HMAC keys are variable-length and are securely encrypted so that their values are protected.
- **Modulus Exponent (ME) and Chinese Remainder Theorem (CRT),** RSA encryption and decryption with key lengths greater than 2048-bits and up to 4096-bits.

Redbooks IBM

z196 GA Cryptographic features

Elliptic Curve Cryptography (ECC)

- **ECC provides more "security per bit" than other public-key algorithms.**
  - Ideal for resource-constrained systems
  - RSA cryptography is impractical at key lengths with strength equivalent to AES-192 and AES-256.
- **Elliptic Curve Key generation and digital signatures**
  - Only clear and internal EC keys will be supported in z196 GA1
  - Keys will be protected using a new ECC master key (256-bit AES key)
  - CCA verbs modified: PKB, PKG, DSG, DSV, KTC, PKX, CFQ, KYT
  - Hardware will be used as appropriate to accelerate ECC performance

NIST guidelines for public key sizes for AES

ECC KEY SIZE (Bits)	RSA KEY SIZE (Bits)	KEY SIZE RATIO	AES KEY SIZE (Bits)
163	1024	1 : 6	
256	3072	1 : 12	128
384	7680	1 : 20	192
512	15 360	1 : 30	256

Supplied by NIST to ANSI X9F1

Redbooks IBM

z196 Cryptographic enhancements

- **CP Assist for Cryptographic Function (CPACF) Message-Security Assist 4 enhancements**
  - New Instructions
    - **KMF** – Cipher message with Cipher Feedback (CFB) Mode
    - **KMCTR** – Cipher Message with Counter
    - **KMO** – Cipher Message Output Feedback (OFB) mode
  - Enhanced Instruction
    - **KIMD** – Extension for GHASH
- **Trusted Key Entry 7 Workstation (FC #0841)**
  - Includes the same cryptographic hardware as Crypto Express3
  - Required for z196, supports System z10, System z9, and eServer zSeries 990 and 890
  - Requires Trusted Key Entry 7.0 Licensed Internal Code (FC #0860)
  - Requires use of USB flash memory for TKE removable media functions
  - Supports only the USB Smart Card reader (FC #0885) introduced on System z10
- **Trusted Key Entry 7.0 Licensed Internal Code (FC #0860)**
  - Requires Trusted Key Entry 7 Workstation (FC #0841)
  - Supports USB flash memory for TKE removable media functions
  - Supports enhanced migration of configuration data and key material using a secure protocol
  - Supports Elliptic Curve Cryptographic master keys
  - Supports upload of TKE Audit Records
  - Requires 6-digit PIN support for TKE Smart Cards (Supports existing cards with 4-digit PIN)
  - Strengthens Passphrase requirements for logon

Redbooks IBM

Trusted Key Entry Usability Enhancements

TKE Domain Grouping

- Grouping of multiple domains and configuration attributes on a Crypto Express3 feature to substantially reduce the time required to load new master keys.
- Example: A TKE user can load a master key part to all domains with one command.

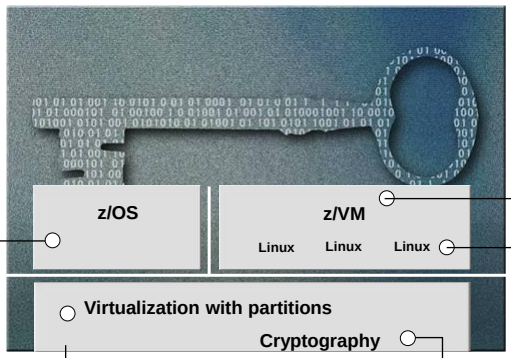
Crypto Migration Wizard

- Allows users to migrate Crypto Express3 and Crypto Express2 configuration data and to significantly reduce migration task duration.
- Collect configuration data from a Crypto Express3 and Crypto Express2 coprocessor and migrate the data to a different Crypto Express coprocessor.
- The target Crypto Express coprocessor must have the same or greater capabilities than the source Crypto Express coprocessor

Redbooks IBM

Certifications for System z

The Common Criteria program developed by NIST and NSA establishes an organizational and technical framework to evaluate the trustworthiness of IT Products and protection profiles



- **z/VM**
- **Common Criteria EAL4+**
  - z/VM 5.4 + RACF with CAPP and LSPP – designed for compliance
  - Plan to have z/VM 6.1 evaluated during 2011
- **Linux on System z**
- **Common Criteria**
  - Novell SUSE SLES9 and 10 certified at EAL4+ with CAPP
  - RHEL5 EAL4+ with CAPP and LSPP

- **Common Criteria EAL4+**
  - with CAPP and LSPP
  - z/OS 1.10
  - z/OS 1.11 June 2010
  - z/OS 1.12 under evaluation
- **IdenTrust™ certification** for z/OS as a Digital Certificate Authority (PKI Services)

z9 EC, z9 BC, z10 EC, z10 BC

z196 - Pending

- **Common Criteria** EAL5 for Logical partitions
- **FIPS 140-2 level 4** for Crypto Express2

See: www.**ibm.com**/security/standards/st_evaluations.shtml

z196 RAS

- *RAS Strategy*
- *Preventing Outages*
- *Improving Reliability*
- *Memory RAS*

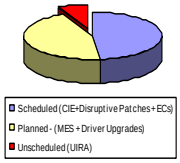
zEnterprise System RAS

- **The zEnterprise System continues to reduce customer down time by focusing on all sources of outages: unscheduled outages, scheduled outages and planned outages. Power and cooling requirements were reduced while still managing reliability.**
  - Major new Memory design for preventing outages
  - Introducing new IO drawers technologies designed for concurrent service
  - Introducing System z management to the mixed computing environment
  - Delivering Green functions and RAS together

Redbooks IBM

System z overall RAS Strategy
…..Continuing our RAS focus helps avoid outages

Sources of Outages
Pre z9
-Hrs/Year/Syst-



Impact of Outage

	Prior Servers	z9 EC	Z10 EC	z196
Unscheduled Outages	✓	✓	✓	✓
Scheduled Outages	✓	✓	✓	✓
Planned Outages		✓	✓	✓
Preplanning requirements			✓	✓
Power & Thermal Management				✓

Increased Focus over time

Temperature = Silicon Reliability Worst Enemy
Wearout = Mechanical Components Reliability Worst Enemy.

Redbooks IBM

Preventing All Outages

- **Unscheduled Outages**
  - ✓ Advanced Memory RAIM (Redundant Array of Independent Memory) design
    - ✓ Enhanced Reed-Solomon code (ECC) – 90B/64B
    - ✓ Protection against Channel/DIMM failures
    - ✓ Chip marking for fast DRAM replacements
  - ✓ Mirrored Key cache
  - ✓ Improved chip packaging
  - ✓ Improved condensation management
  - ✓ Integrated TCP/IP checksum generation/checking
  - ✓ Integrated EPO switch cover (protecting the switch during repair actions)
  - ✓ Continued focus on Firmware

Redbooks IBM

Preventing All Outages

- **Scheduled Outages**
  - ✓Double memory data bus lane sparing (reducing repair actions)
  - ✓Single memory clock bus sparing
  - ✓Field Repair of interface between processor chip and cache chip and between cache chips (fabric bus)
  - ✓Fast bitline delete on L3/L4 cache (largest caches)
  - ✓Power distribution using N+2 Voltage Transformation Modules (VTM)
  - ✓Redundant (N+2) Humidity Sensors
  - ✓Redundant (N+2) Altitude Sensors
  - ✓Unified Resource Manager for zBX

Redbooks IBM

Preventing All Outages

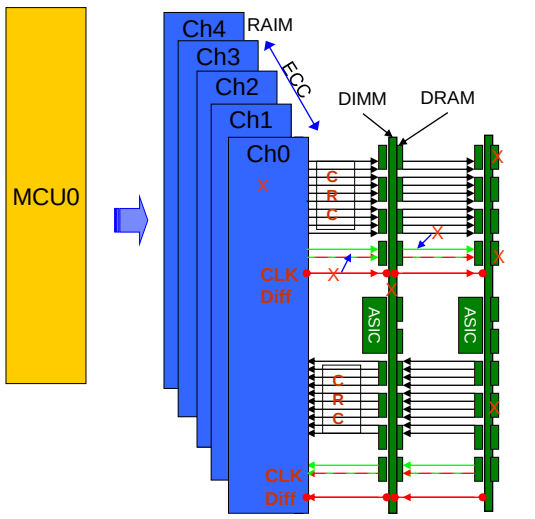
- **Planned Outage**
  - ✓Concurrent zBX attachment and control enablement
- **Preplanning**
  - ✓Fully concurrent MRU firmware upgrade (no cycle steering)

Improving Reliability

Power & Thermal Optimization and Management

- ✓ Static Power Save Mode under z/OS control
- ✓ Smart blower management by sensing altitude and humidity
- ✓ Enhanced evaporator design to minimize temperature variations
- ✓ MCM cooling with N+1 design feature
- ✓ MCM backup cooling with Heat Exchanger
- ✓ Available air to water heat exchanger for customer water cooling

z196 – Memory Overview



2- Deep Cascade
Using Quad High DIMMs

Layers of Memory Recovery

ECC
- Powerful 90B / 64B Reed Solomon code

DRAM Failure
- Marking technology; no half sparing needed
- 2 DRAM can be marked
- Call for replacement on third DRAM

Lane Failure
- CRC with Retry
- Data – lane sparing
- CLK – RAIM with lane sparing

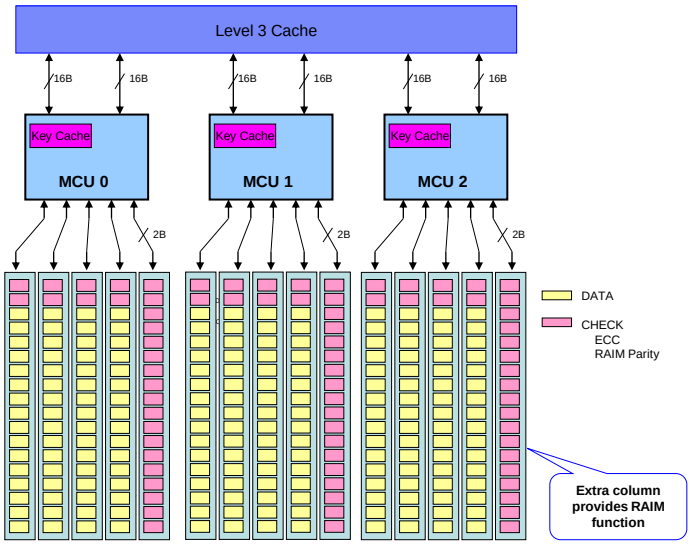
DIMM Failure (discrete components, VTT Reg.)
- CRC with Retry
- Data – lane sparing
- CLK – RAIM with lane sparing

DIMM Controller ASIC Failure
- RAIM Recovery

Channel Failure
- RAIM Recovery

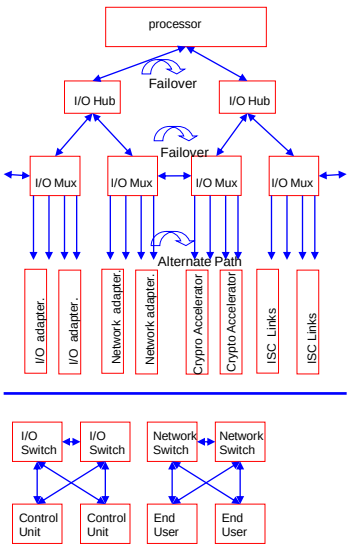
Redbooks IBM

z196 Redundant Array of Independent Memory (RAIM) Structure



Redbooks IBM

z196 RAS Design of
……..Fully Redundant I/O Subsystem – of existing IO cage and drawers



Fully Redundant I/O Design

- SAP / CP sparing
- SAP Reassignment
- I/O Reset & Failover
- I/O Mux Reset / Failover
- Redundant I/O Adapter
- Redundant I/O interconnect
- Redundant Network Adapters
- Redundant ISC links
- Redundant Crypto processors
- I/O Switched Fabric
- Network Switched/Router Fabric
- High Availability Plugging Rules
- I/O and coupling fanout rebalancing on CBA
- Channel Initiated Retry
- High Data Integrity Infrastructure
- I/O Alternate Path
- Network Alternate Path
- Virtualization Technology

Summary

IBM zEnterprise 196 - System Summary

PROCESSOR/MEMORY
1.3x Faster Processor Unit (PU)
1.6x Faster CPU Intensive
Larger capacity 80 customer PUs
Flexible 45 CP Subcapacity Settings
Up to 3 TB memory • 768GB per book • 32/64/96/112/128GB increments • 16GB fixed HSA
RAIM Memory RAS Improvement

I/O
8.0 GBps InfiniBand HCA to I/O interconnect
Concurrent Add delete I/O drawers
zDAC Discovery and Automatic Configuration
FICON enhancements

MACHINE INFO
Machine Type: 2817
5 Models: M15, M32, M49, M66 and M80

HETEROGENEOUS PLATFORM MANAGEMENT
Integrated Performance Management for heterogeneous architecture
Application Server Blades
Data Power Integration

ENVIRONMENTALS
Optional Water Cooling
Optional High Voltage DC Input
Optional Overhead I/O Cabling
Static Power Savings
New Capacity on Demand architecture and enhancements

HMC / SE
Ensemble Management
UFD Memory Media
Driver 86 at GA

Dank u
Dutch

Merci
French

Спасибо
Russian

Gracias
Spanish

شكراً
Arabic

감사합니다
Korean

Tack så mycket
Swedish

धन्यवाद
Hindi

תודה רבה
Hebrew

Obrigado
Brazilian Portuguese

谢谢
Chinese

Thank You

Dankon
Esperanto

ありがとうございます
Japanese

Trugarez
Breton

Danke
German

Tak
Danish

Grazie
Italian

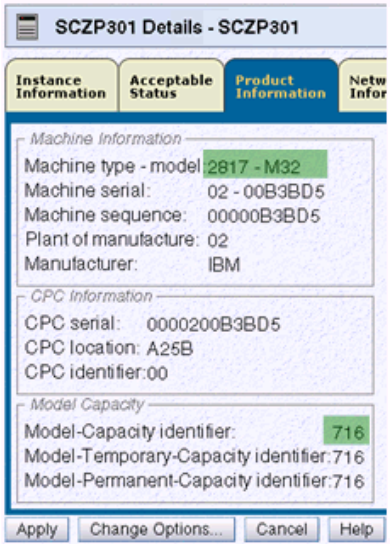
நன்றி
Tamil

děkuji
Czech

ขอบคุณ
Thai

go raibh maith agat
Gaelic

Hardware and Software model




```
D M=CPU
IEE174I 16.36.24 DISPLAY M 907
PROCESSOR STATUS
ID  CPU                     SERIAL
00  +                        013BD52817
01  +                        013BD52817
02  -
03  -
04  -I
05  -I
06  -A
07  -A

CPC ND = 002817.M32.IBM.02.0000000B3BD5
CPC SI = 2817.716.IBM.02.00000000000B3BD5
         Model: M32
CPC ID = 00
CPC NAME = SCZP301
```

Redbooks IBM

Coupling Facility Limits by CFCC Levels

Coupling facility Limit	Coupling Facility Code Level										
	Level 17	Level 16	Level 15	Level 14	Level 13	Level 12	Level 11	Level 10	Level 9 SL >2.00	Level 9 (SL< 2.00)	Level 8
Max number of CPs	16	16	16	16	16	16	16	16	16	12	12
Storage increment	512KB	512KB	512 KB	256 KB	256 KB	256 KB	256 KB	256 KB	256 KB	256 KB	256 KB
Structure ID limit	2047	1023	1023	1023	1023	1023	1023	1023	1023	1023	1023
Retry buffer limit	2039	2039	2039	1799	1799	1559	1559	1559	1559	519	519
Facility information	64	64	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes
Maximum list element characteristic	4	4	4	4	4	4	4	4	4	4	4
Maximum lock t.e. characteristic	5	5	5	5	5	5	5	5	5	5	5
User identifier limit	64	32	32	32	32	32	32	32	32	32	32
Maximum d.a. element characteristic	4	4	4	4	4	4	4	4	4	4	4
Local cache identifier limit	255	255	255	255	255	255	255	255	255	255	255
Storage class limit	127	63	63	63	63	63	63	63	63	63	63
Castout class limit	2048	1024	1024	1024	1024	1024	1024	1024	1024	1024	1024

221

Redbooks IBM

CPC support for Coupling Facility code levels

CPC models	Coupling faciliity code level						
	Level 17	Level 16	Level 15	Level 14	Level 13	Level 12	Level 11
2817 zEnterprise 196	EC N29796 (Ver 2.11.0) MCL 000	NA	NA	NA	NA	NA	NA
2097 z10 EC	NA	EC N10964 (Ver 2.10.1) MCL015	EC F85900 (Ver 2.10.0) MCL006	NA	NA	NA	NA
2096 z9 BC	NA	NA	EC G40953 (Ver 2.9.2) MCL 001	EC J99670 (Ver 2.9.1) MCL 005	NA	NA	NA
2094 z9 EC	NA	NA	EC G40953 (Ver 2.9.2) MCL 001	EC J99670 (Ver 2.9.0) MCL 005	NA	NA	NA
2084 z990	NA	NA	NA	EC J13481 (Ver 1.8.2) MCL 018	EC J13481 (Ver 1.8.2) MCL 003	EC J12555 (Ver 1.8.0) MCL 009	NA
2086 z890	NA	NA	NA	EC J13481 (Ver 1.8.2) MCL 018	EC J13481 (Ver 1.8.2) MCL 003	NA	NA
2066 z800	NA	NA	NA	NA	EC J11207 (Ver 1.7.3) MCL 029	EC J11207 (Ver 1.7.3) MCL 022	NA

222