System z10 I/O Subsystem

Connie K. Beuselinck
IBM Corporation
Poughkeepsie, NY U.S.A
conniek@us.ibm.com

IBM Systems

Trademarks

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

390	HiperSockets	Parallel Sysplex*	Tivoli *
ACF/VTAM *	HPR	PR/SM	TotalStorage *
APPN *	IBM*	RACF *	VM/ESA *
CICS *	IBM logo*	Redbooks	VSE/ESA
DB2 *	IBM System z	Resource Link	VTAM *
ESCON *	IBM System z9	RMF	WebSphere *
eServer	IBM System z10	System Storage	z9
GDPS*	IMS	Sysplex Timer*	z10
FICON*	Infoprint	System z9	z/Architecture*
		System z10	z/OS*
			zSeries*
			z/VM *
			z/VSE

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies:

Linux is a registered trademark of Linus Torvalds

Penguin (Tux) compliments of Larry Ewing

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries

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

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

* 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.

IBM Systems

Performance data disclaimers

The performance data contained herein was obtained in a controlled environment based on the use of specific data. Actual results that may be obtained in other operating environments may vary significantly. These values do not constitute a guarantee of performance.

Product data is accurate as of initial publication and is subject to change without notice.

No part of this presentation may be reproduced or transmitted in any form without written permission from IBM Corporation.

References in this document to IBM products, programs, or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM program product in this document is not intended to state or imply that only IBM's program product may be used. Any functionally equivalent program may be used instead. Future plans and announcements are subject to change.

The information provided in this document has not been submitted to any formal IBM test and is distributed "As Is" basis without any warranty either express or implied. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into their operating environment.

While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

The mentioning of 3rd party products does not constitute a warranty or endorsement of any kind.

IBM Systems

InfiniBand coupling links

- Note: The InfiniBand link data rate of 6 GBps or 3 GBps does 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, the service times of coupling operations are greater, and the actual throughput may be less than with ICB links.

- Refer to the *Coupling Facility Configuration Options* whitepaper for a more specific explanation of when to continue using the current ICB technology versus migrating to InfiniBand coupling links.

- The *Coupling Facility Configuration Options* whitepaper is available at:
  - ► http://www.ibm.com/systems/z/advantages/pso/whitepaper.html

- Refer also to *Getting Started with InfiniBand on System z10 and System z9* (SG24-7539)
  - ► http://www.redbooks.ibm.com/

IBM Systems

Agenda

- *System z10 I/O Subsystem*
- **Coupling**
  - *IC, ICB-4, ISC-3*
  - *InfiniBand on System z10, System z9*
- *Storage*
  - *ESCON, FICON, FCP*
- *Networking*
  - *HiperSockets*
  - *OSA-Express3*

Glossary

Acronym	Full Name	Use
CCW	Channel Control Word	An architecture supporting I/O communications
CHPID	Channel Path Identifier	Identifier for channel path type in I/O Control Program
CRH	Channel Request Handler	Internal host bus for communication between I/O & memory
CSS	Channel Subsystem	A group of resources for use by channels
CTC	Channel-To-Channel	Protocol for host-to-host communication; ESCON, FICON
ESCON	Enterprise Systems Connection	Channel architecture used by switches, directors, storage, printers
EREP	Environmental Record Editing and Printing Program	Processes SYS1.LOGREC datasets
FCP	Fibre Channel Protocol	Fixed format protocol to communicate with SCSI devices
FICON	Fibre Connection	Channel architecture used by switches, directors, storage, printers
IFB	InfiniBand	Refer to next slide
ISL	InterSwitch Link	Data path between Directors to minimize cross-site fiber optic cabling
km	kilometer	Distance measurement identified with fiber optic cabing
LCSS	Logical Channel Subsystem	Architecture allowing more than one physical channel subsystem

New

Glossary

Acronym	Full Name	Use
LUN	Logical Unit Number	Associated with fixed block / SCSI devices
LX	Long wavelength	Light source designed for use with single mode fiber
MBps	MegaByptes per second	Information transfer at one million bytes per second
MBA	Memory Bus Adapter	Building block of Central Electronic Complex and MCM
MCM	MultiChip Module	"Brick" containing Processor Units
MM	Multimode	Short form used in graphics; refers to multimode fiber
MPO	Multi-fiber Push-On	Connector used with 12x IB-DDR fiber optic cables; packaged with 12 fibers for transmit and 12 fibers for receive
PCI	Peripheral Component Interconnect	Local bus standard used with Crypto, FICON, OSA hardware
PCI-X	PCI Extended	Faster than PCI
RAS	Reliability, Availability, Serviceability	
SAN	Storage Area Network	
SCSI	Small Computer System Interface	Fixed block architecture; communication between servers & devices
SM	Single mode	Short form used in graphics; refers to single mode fiber
STI	Self-Timed Interconnect	System z internal host bus for communication between I/O & memory
SX	Short wavelength	Light source designed for use with multimode fiber

Glossary

Term	Description
Gbps	**Gigabits per second; Information transfer at one billion bits per second**
GBps	**GigaBytes per second; Information transfer at one billion bytes per second (1,024 megabytes)**
1x	**One "lane", one pair of fibers**
12x	**12 "lanes", 12 pairs of fiber**
SDR	**Single Data Rate – 2.5 Gbps per "lane" (0.25 GBps)**
DDR	**Double Data Rate – 5 Gbps per "lane" (0.5 GBps)**
12x IB-SDR	**12 "lanes" (pairs) for a total link data rate of 3 GBps, 150 meters point-to-point Used with OM3, 2000 MHz-k 50 micron multimode fiber with MPO connectors**
12x IB-DDR	**12 "lanes" (pairs) for a total link data rate of 6 GBps, 150 meters point-to-point Used with OM3, 2000 MHz-k 50 micron multimode fiber with MPO connectors**
1x IB-DDR LR* (SOD)	**One "lane" (one pair), 5 Gbps link data rate, unrepeated distance of 10 km Used with 9 micron single mode fiber with LC Duplex connector**

* Statement of Direction (SOD)
All statements regarding IBM's future direction and intent are subject to change
or withdrawal without notice, and represent goals and objectives only.

IBM Systems

Glossary

Acronym	Full name	Comments
AID	**Adapter identification**	**HCA fanout has an AID instead of PCHID; used to assign a CHPID to a port using HCD/IOCP**
CIB	**Coupling using InfiniBand**	**CHPID type System 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 cage intra-connection**
STI-MP	**Self-Timed Interconnect Multiplexer**	**I/O cage intra-connection**

Type	System z9	System z10 EC
HCA1-O fanout	Optical - Coupling 12x IB-SDR	NA
HCA2-C fanout	NA	Copper - I/O Cage - 12x IB-DDR
HCA2-O fanout	NA	Optical - Coupling 12x IB-DDR
MBA fanout	I/O cage connection to CEC or Coupling (ICB-4)	Coupling (ICB-4)
IFB-MP	NA	I/O cage intra-connection
STI-MP	I/O cage intra-connection	NA

IBM Systems

IBM System z10 EC

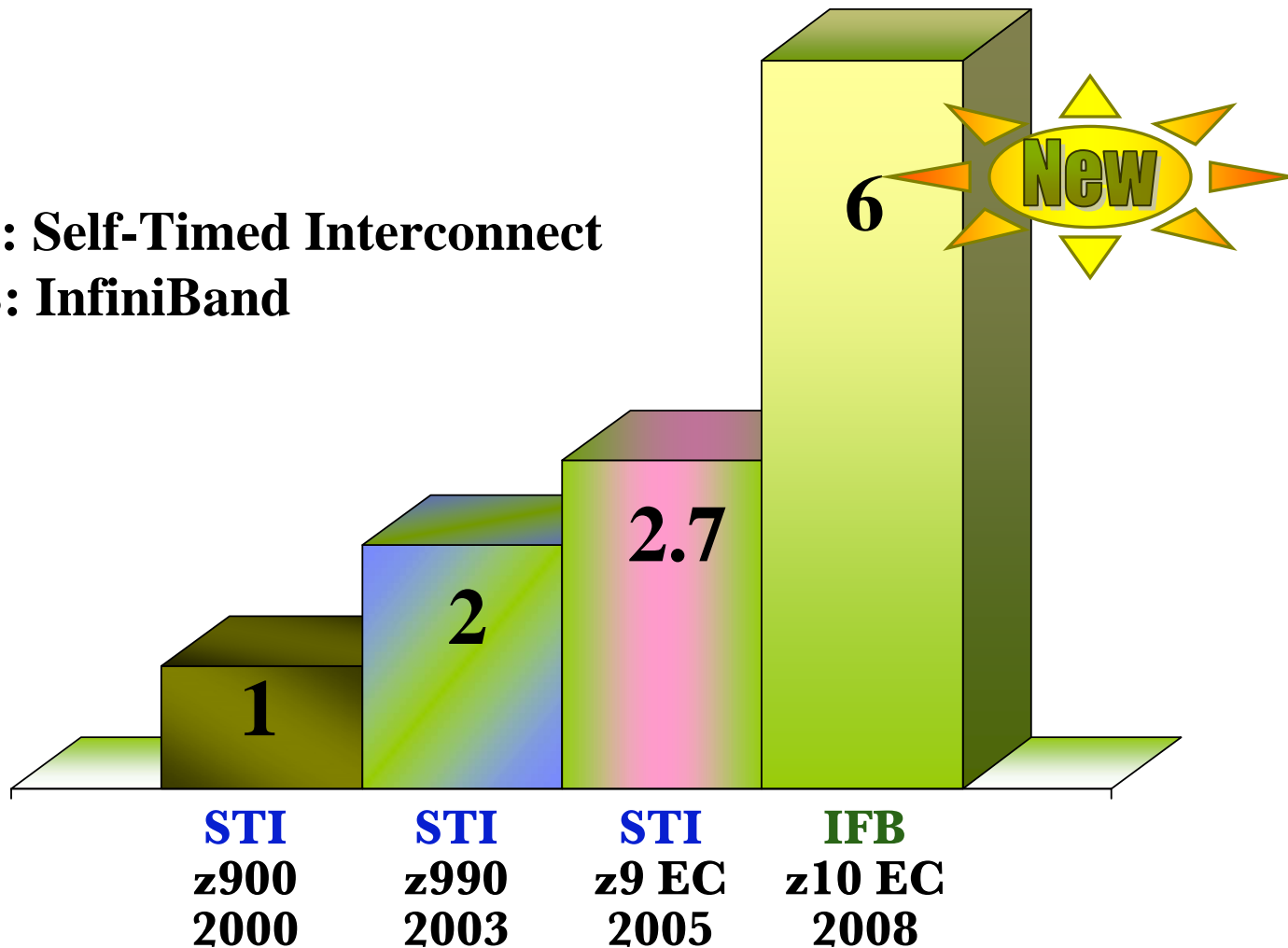**Coupling**

Storage

Network

IBM Systems
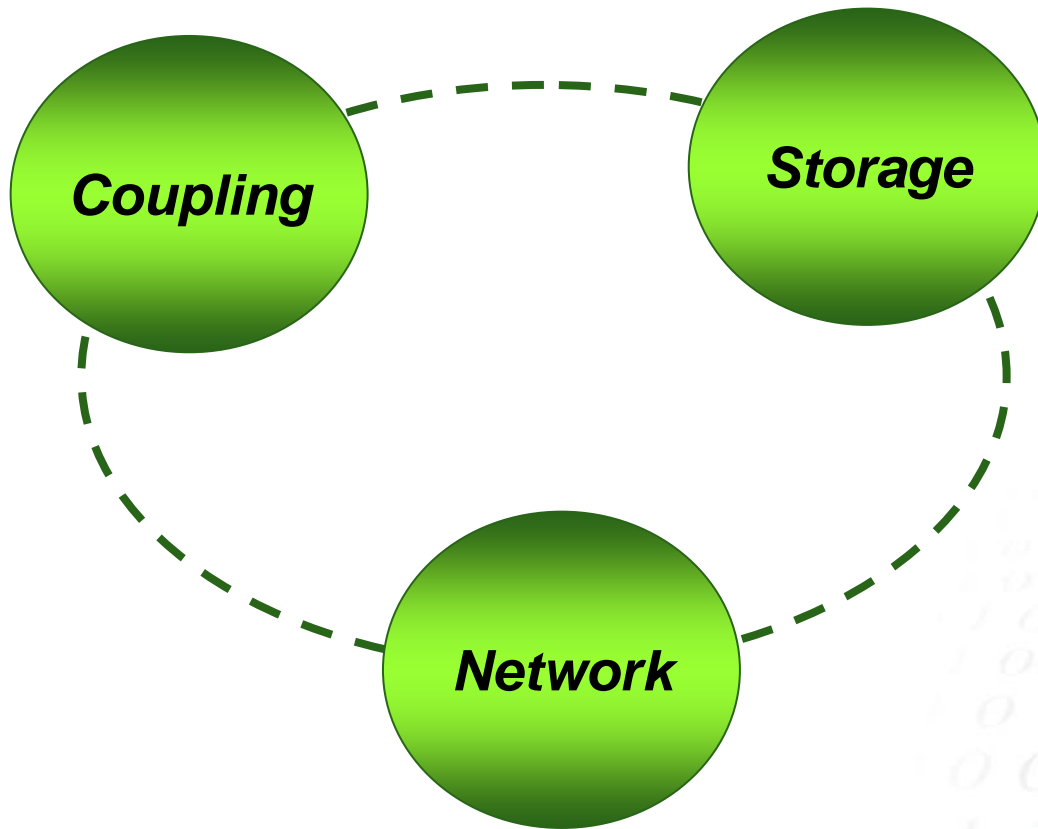
Host Bus Evolution

The Future Runs on System z

IBM Systems

I/O subsystem host bus interface GigaBytes per second (GBps)

STI: Self-Timed Interconnect
IFB: InfiniBand

New

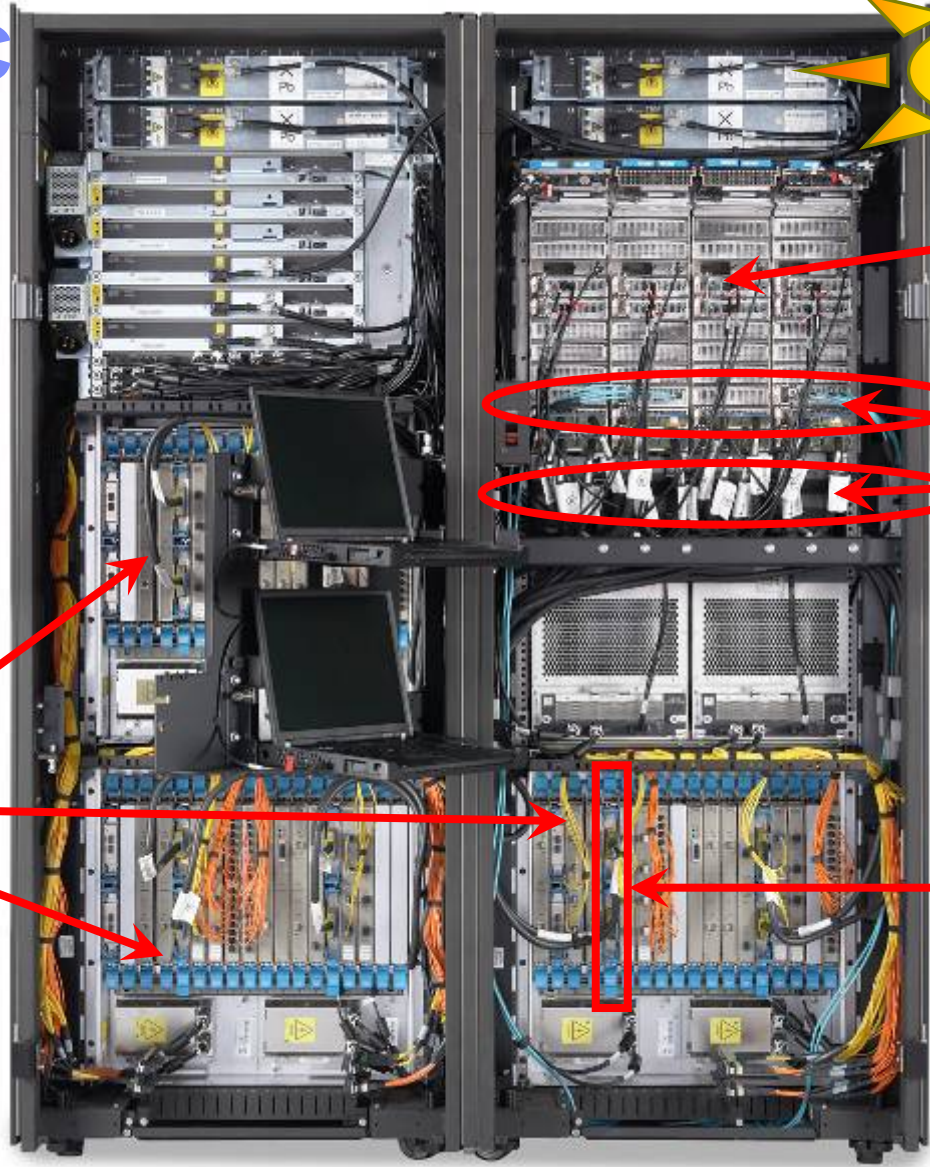| STI	STI	STI	IFB
1	2	2.7	6
z900	z990	z9 EC	z10 EC
2000	2003	2005	2008

© 2008 IBM Corporation

IBM Systems

System z10 EC

New

Books

12x IB-DDR optical copper cables

I/O cages

IFB-MP

IBM Systems

z10 EC   Book Under the covers

MCM

Memory

DCA Power Supplies

MCM

HCA2-O (InfiniBand)

FSP cards (SE LAN)

HCA2-C (I/O cages)

New

MBA (ICB-4)
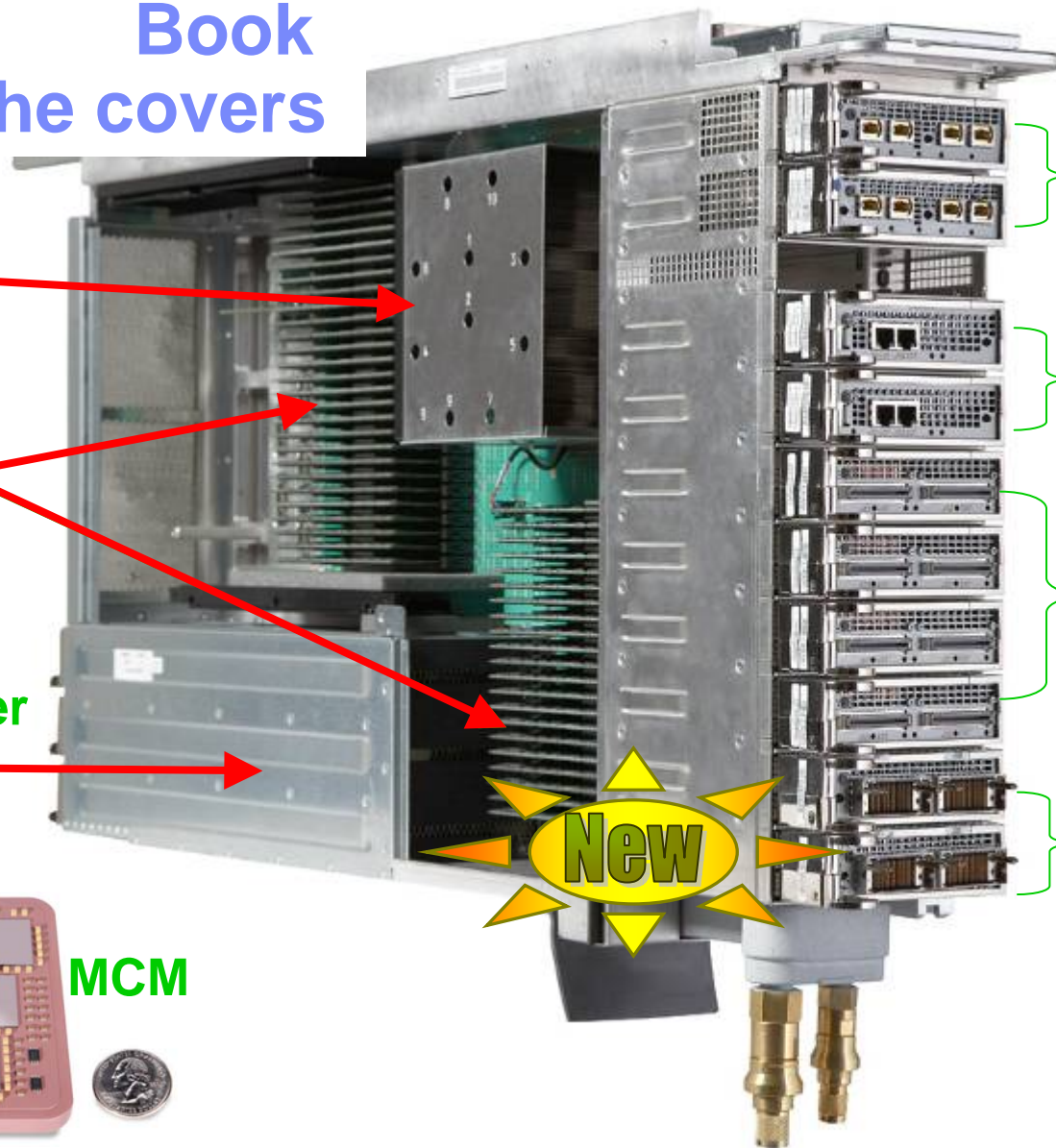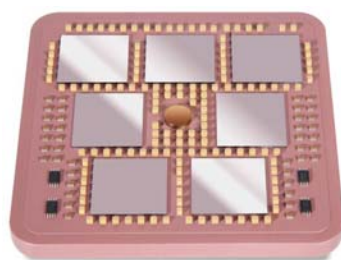
IBM Systems

Host Channel Adapter 2 (HCA2) fanouts

New

HCA2-C

HCA2-O

- *12x IB-DDR interfaces, 6 GBps*

  ➤ *HCA2-C (copper) = to I/O cages*

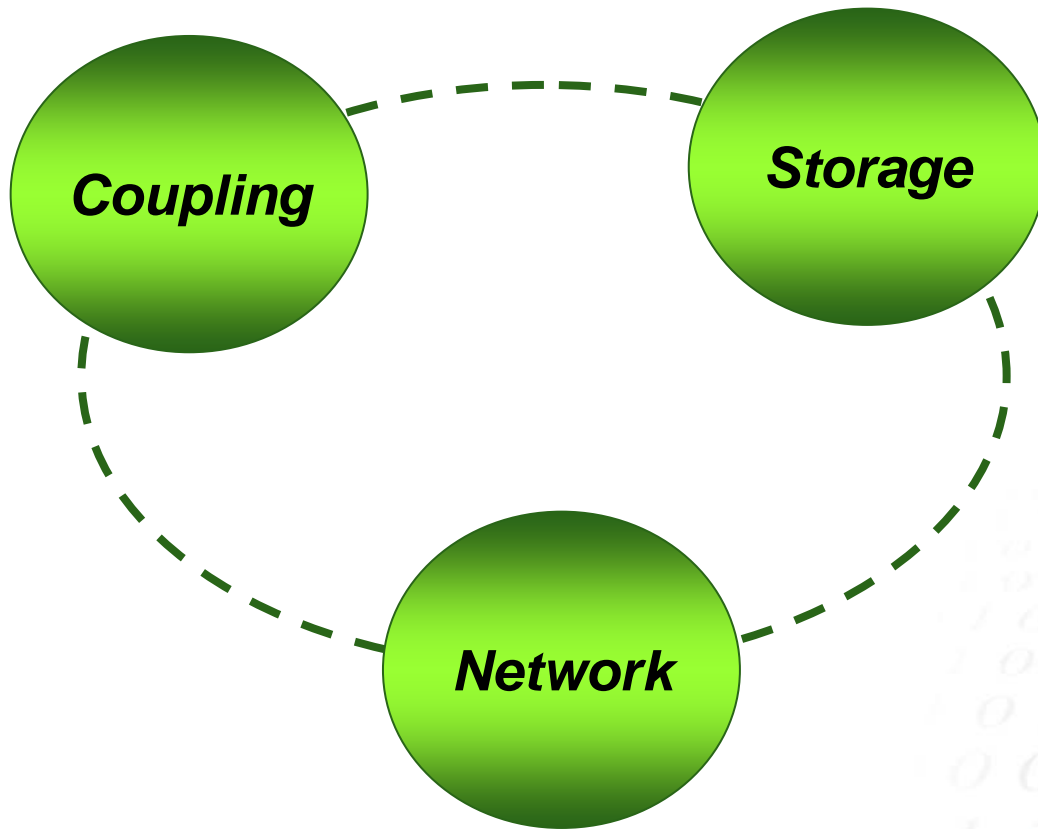 ➤ *HCA2-O (optical) = External coupling*
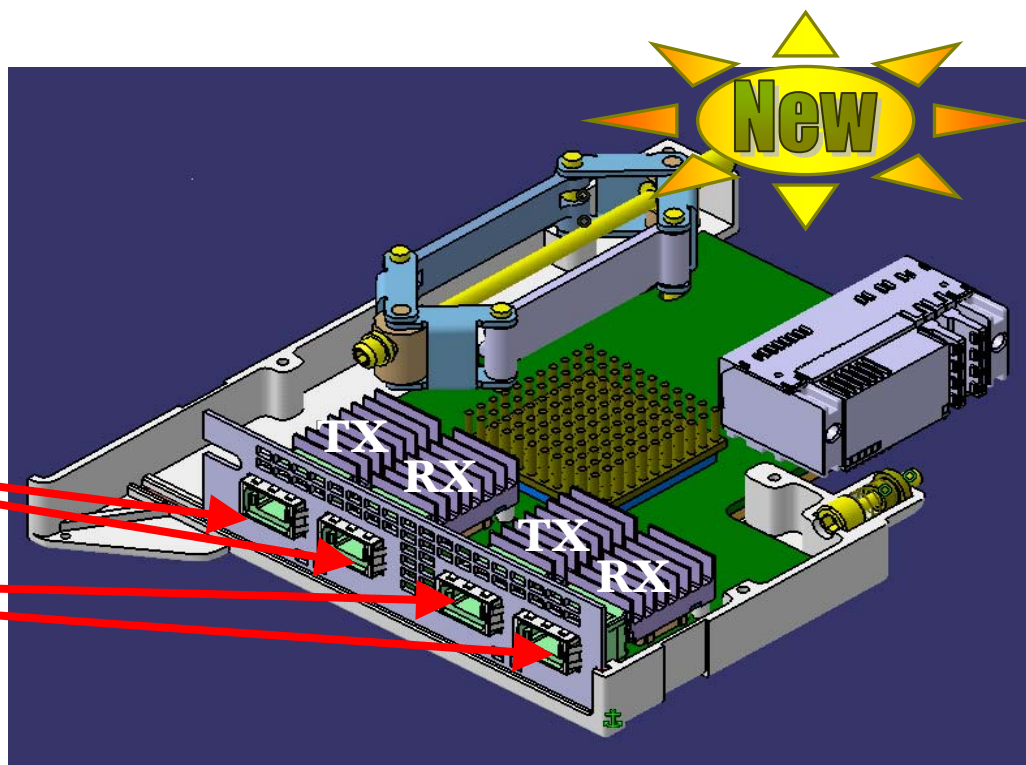
IBM Systems

z10 EC I/O infrastructure

Book 0 — Memory — HCA (8x) — 12x IB-DDR (16x)

Book 1 — Memory — HCA (8x) — 12x IB-DDR (16x)

Book 2 — Memory — HCA (8x) — 12x IB-DDR (16x)

Book 3 — Memory — HCA (8x) — 12x IB-DDR (16x)

HCA fanout or MBA fanout for ICB-4

IFB-MP RII IFB-MP — 2 GBps — Channels — FICON Express4 1/2/4 Gbps

IFB-MP RII IFB-MP — 500 MBps — Coupling Links — ISC-3

IFB-MP RII IFB-MP — 333 MBps — Channels — ESCON

IFB-MP RII IFB-MP — 2 GBps — Ports — OSA-Express3 10 GbE LR

I/O cage (3x)

28 features per I/O cage

IBM System z10 EC

Coupling

Storage

Network

IBM Systems

HCA2-O fanout

Port 0

Port 1

New

TX
RX
TX
RX
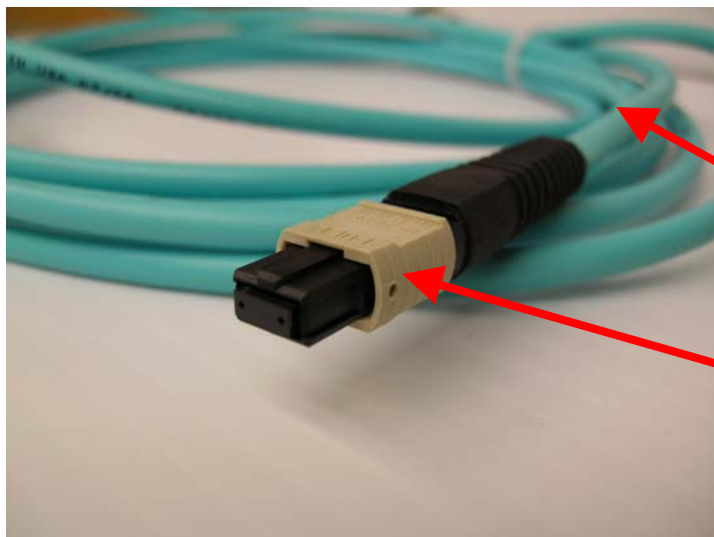
OM3 cable (aqua cladding)
2000 MHz-km, 50 micron multimode fiber

MPO connector
150 meters point-to-point

New

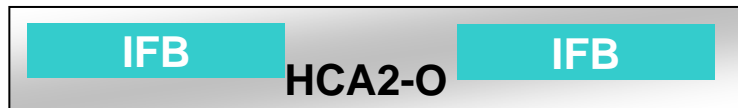*HCA2-O fanout*

IBM Systems

12x IB-DDR fiber optic cabling

- 12x IB-DDR jumper cables - two MPO connectors on each side

  ► **One labeled TX for transmitter**

  ► **One labeled RX for receiver**

  ► **Ports on HCA1-O and HCA2-O fanout cards are also labeled TX and RX**

    ● **Connect the MPO connector labeled TX to the fanout card port labeled TX**

    ● **Connect the MPO connector labeled RX to the fanout card port labeled RX**

- **Use of patch panels**

  ► **Acceptable as long as the link loss budget does not exceed 2.06 dB.**

  ► **A minimum of two fiber-to-fiber connections can generally be supported with the link loss budget for InfiniBand coupling links.**
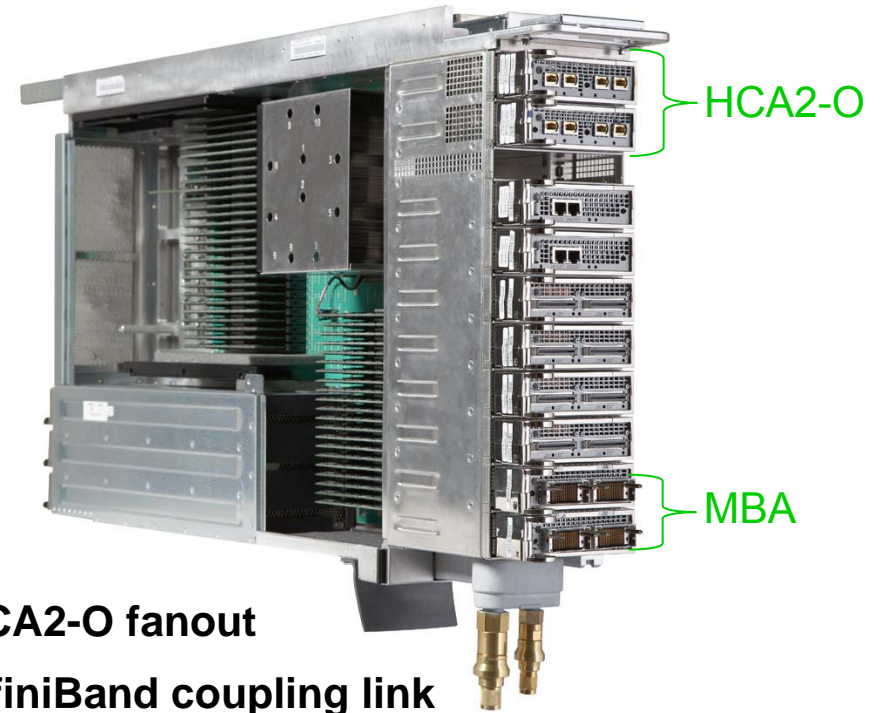
IBM Systems

Coupling on z10 EC

Type	Distance
ICB-4	10 meters
InfiniBand	150 meters

Up to 16 CHPIDs – across 2 ports

IFB	HCA2-O	IFB

AID

2 CHPIDs – 1 per port

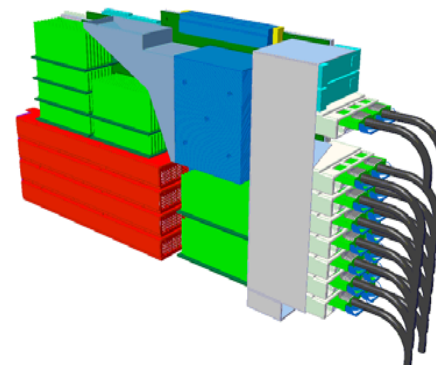| STI	MBA	STI

HCA2-O

MBA

- **HCA2-O fanout**
- **InfiniBand coupling link**
  - ► **New CHPID type CIB**
    - ● **Fiber optic external coupling link**
      - – **Cabling is customer-supplied**

- **MBA fanout (not available on Model E64)**
- **ICB-4 - still uses an STI**
  - ► **Same CHPID type CBP**
  - ► **New cable and connector**
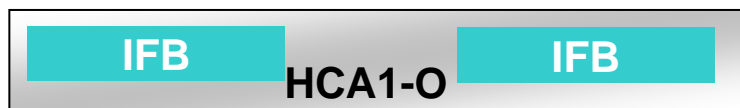   - ● **Cable available as feature**

IBM Systems

Coupling on z9 EC, z9 BC

New

Type	Distance
ICB-4	10 meters
InfiniBand	150 meters

Up to 16 CHPIDs – across 2 ports

IFB   HCA1-O   IFB

AID

- **HCA1-O fanout**
- **InfiniBand coupling link**
  - ► **New CHPID type CIB**
    - ● **Fiber optic external coupling link**
      - – **Cabling is customer-supplied**

2 CHPIDs – 1 per port

STI   MBA   STI

- **MBA fanout**
- **ICB-4 - still uses an STI**
  - ► **Same CHPID type CBP**
  - ► **New cable and connector for connection to z10**
    - ● **Cable available as feature**

IBM Systems

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

Fiber Core (μ) (Light source)	Fiber Bandwidth @ wavelength	3 GB / sec		6 GB / sec	
		Unrepeated distance	Link loss budget*	Unrepeated distance	Link loss budget*
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 IB-SDR links operating at 3 GBps (2.5 Gbps per lane) are used to connect a System z10 to a z9 EC or a z9 BC Coupling Facility.
2. 12x IB-DDR links operating at 6 GBps (5.0 Gbps per lane) are used to connect System z10 EC servers.

* The link loss budget is the total link optical passive loss (fiber, connectors, and unallocated link margin) as defined by the standard for fiber specified above.

IBM Systems

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

- Cables available from:
  - ► **IBM Global Technology Services (GTS)**
  - ► **Anixter**
- Fiber core – 50u multimode
- Light source – SX laser
- Fiber bandwidth @ wavelength: 2000 MHz-km @ 850 nm
- Only IBM cable part numbers accepted

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

IBM Systems

Announced 05 May

HCA-O fanouts for InfiniBand coupling links

Planned availability
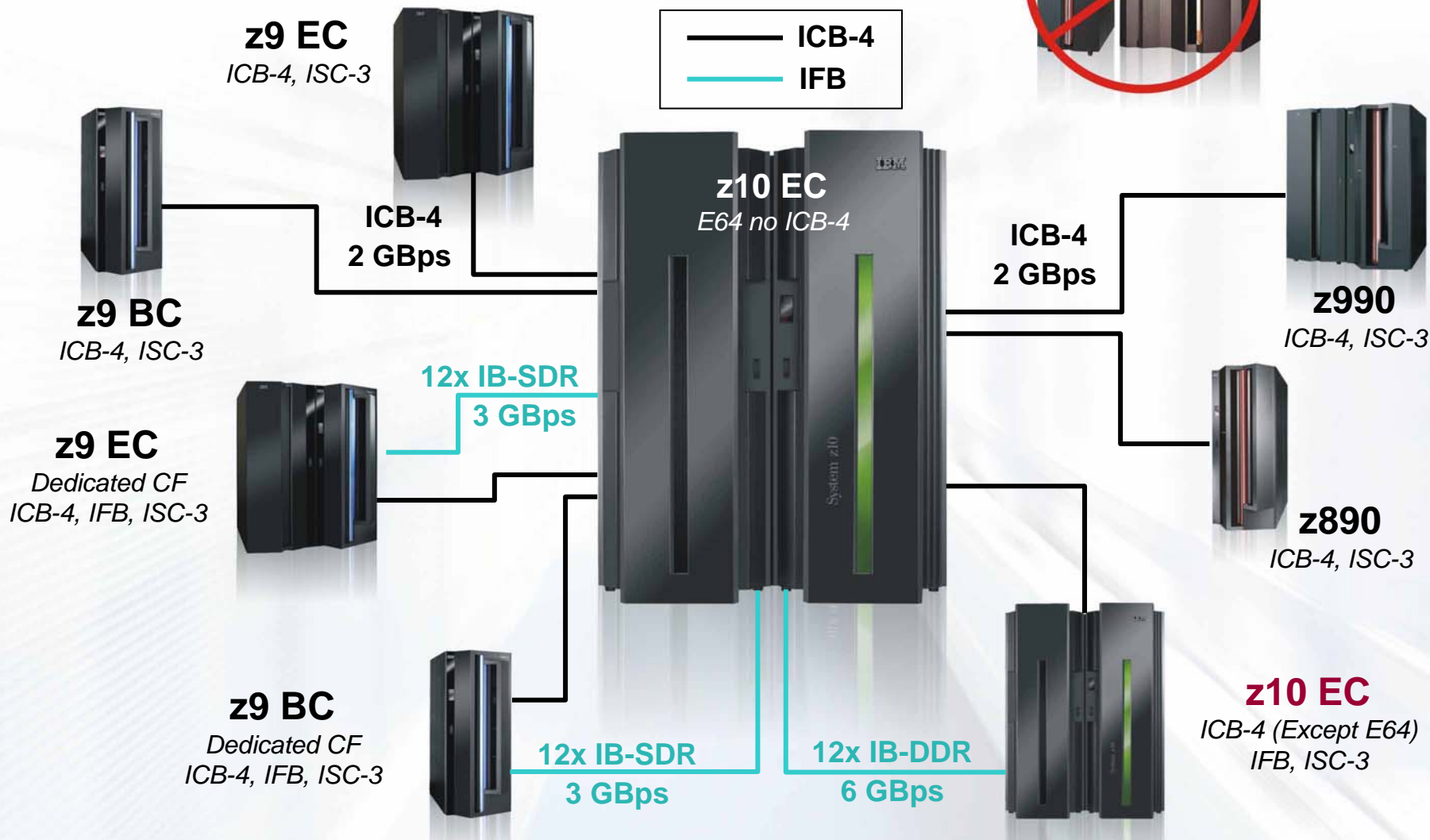## May 30, 2008

**System z9 EC dedicated CFs
and
z9 BC Model S07 dedicated CFs**

System z10

IBM Systems

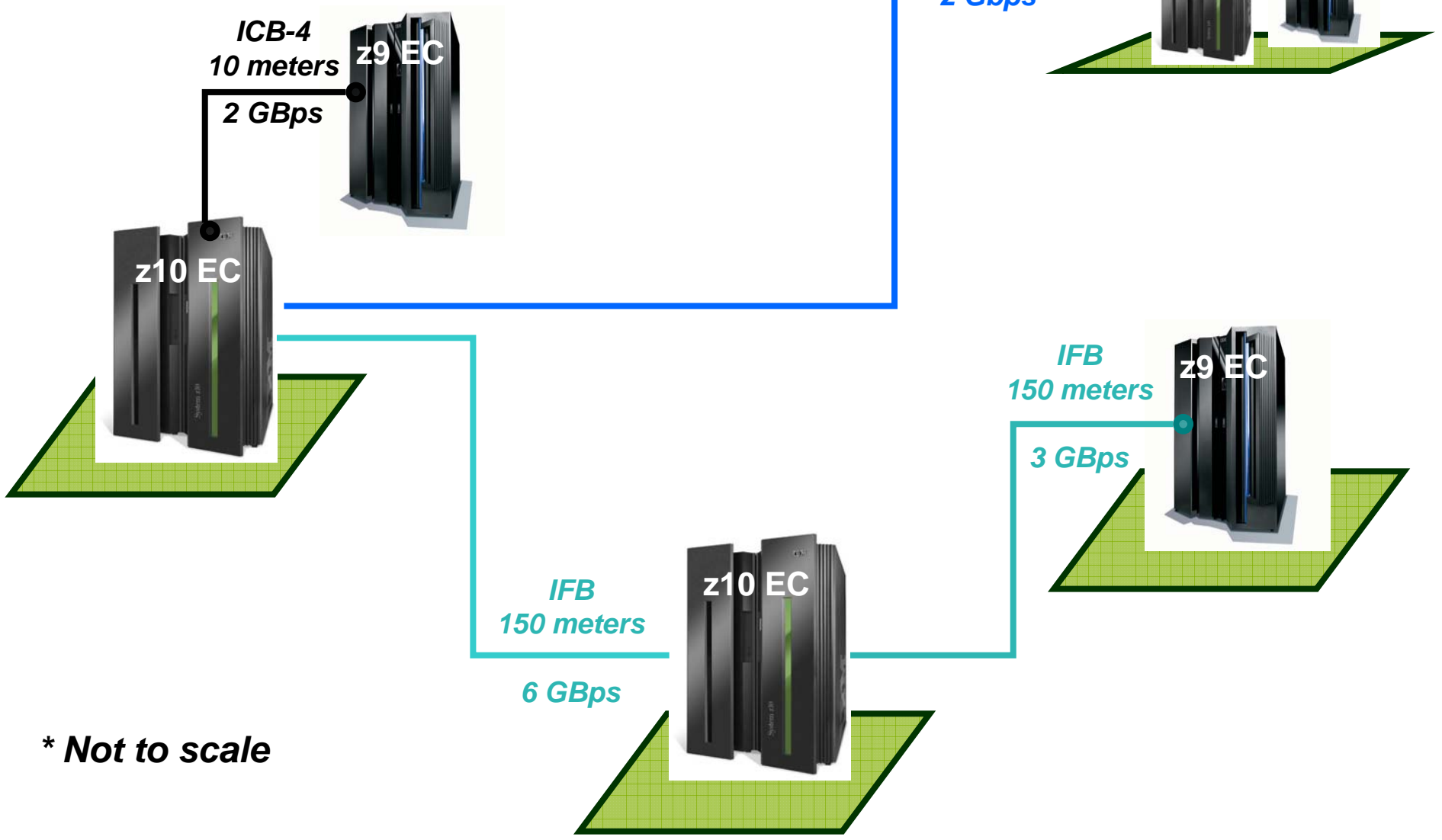# System z Parallel Sysplex coexistence

z800, z900
Not supported!

z9 EC
ICB-4, ISC-3

ICB-4	
IFB	

ICB-4
2 GBps

z9 BC
ICB-4, ISC-3

z10 EC
E64 no ICB-4

ICB-4
2 GBps

z990
ICB-4, ISC-3

12x IB-SDR
3 GBps

z9 EC
Dedicated CF
ICB-4, IFB, ISC-3

z890
ICB-4, ISC-3

z9 BC
Dedicated CF
ICB-4, IFB, ISC-3

12x IB-SDR
3 GBps

12x IB-DDR
6 GBps

z10 EC
ICB-4 (Except E64)
IFB, ISC-3

IBM Systems

1x IB-DDR Statement of Direction
One "lane" (one pair), 5 Gbps link data rate, unrepeated distance of 10 km
Used with 9 micron single mode fiber with LC Duplex connector

- **One fiber pair (1x – transmit fiber and receive fiber)**
- **Double data rate (DDR) – 5 gigabits per second (Gbps) for one lane**
- **Unrepeated distance of 10 km (6.2 miles)**
- **Expected repeated distance of 100 km (62 miles)**
- **Complement and/or replace ISC-3 (2 Gbps link data rate**
- **Same cabling as ISC-3, single mode fiber, 9 micron (yellow cladding)**
- **Same connecter as ISC-3, LC Duplex**
- **Exclusive to System z10 EC**
- **Will be an additional fanout, complementing HCA2-O fanout**
- **Server time protocol (STP) synchronization signals can flow over 1x IB-DDR**

IBM Systems

Coupling links distances*

ISC-3
10 km

2 Gbps

z10 EC   z9 BC

ICB-4
10 meters   z9 EC

2 GBps

z10 EC

IFB
150 meters   z9 EC

3 GBps

IFB
150 meters   z10 EC

6 GBps

* Not to scale

z10 EC Coupling link options

Type	Description	Use	Link data rate	Distance	z10 EC Maximum	Combined Maximum	z10 EC Maximum
IC	Internal Coupling channel	Internal communication	Internal speeds	NA	32	NA	
ICB-4[a]	Copper connection between OS and CF	z10 EC, z9 EC, z9 BC z990, z890	2 GBps	10 meters [b] (33 feet)	16 [c]	32 links	64 CHPIDs
IFB[d]	12x IB-DDR	z10 EC to z10 EC z10 EC to z9 CF	6 GBps 3 GBps[e]	150 meters [b] (492 feet)	32 [d f]		
ISC-3	Fiber connection between OS and CF	z10 EC, z9 EC, z9 BC z990, z890	2 Gbps	10 km (6.2 miles) unrepeated 100 km repeated	48	NA	

For each MBA fanout installed for ICB-4's, the number of HCA fanouts is reduced by one

[a] ICB-4 is not supported on z10 EC Model E64. See RPQ 8P2334
[b] 3 meters (10 feet) reserved for internal routing and strain relief
[c] ICB-4 and IFB in combination cannot exceed 32 links.
[d] Each HCA2-O fanout supports definition of up to 16 CHPID type CIB
[e] z10 EC negotiates to 3 GBps (12x IB-SDR) when connected to a System z9 Dedicated CF
[f] 16 IFB links on z10 EC Model E12

IBM Systems

System z – Supported Coupling Links

System	IFB** (2Q2008)	ICB-4	ICB-3	ISC-3	IC	Max # Links
z10 EC	32*	16* Except E64	N/A	48	32	64
z9 Dedicated CF	16	16	16	48	32	64
Any z9	SOD**	16	16	48	32	64
z990	N/A	16	16	48	32	64
z890	N/A	8	16	48	32	64

* Maximum of 32 IFB + ICB-4 links on System z10 EC. ICB-4 not supported on Model E64.

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

IBM Systems

InfiniBand coupling links software requirements

- **z/OS V1.7 with PTFs.**

- **z/VM V5.3 to define, modify, and delete an InfiniBand coupling link, CHPID type CIB, when z/VM is the controlling LPAR for dynamic I/O.**

IBM Systems

Guidance - InfiniBand coupling links

- **Refer to the *Coupling Facility Configuration Options* whitepaper for a more specific explanation of when to continue using the current ICB technology versus migrating to InfiniBand coupling links.**

- **The *Coupling Facility Configuration Options* whitepaper is available at:**
  - ► http://www.ibm.com/systems/z/advantages/pso/whitepaper.html

- **Refer also to *Getting Started with InfiniBand on System z10 and System z9* (SG24-7539)**
  - ► http://www.redbooks.ibm.com/

- **Note: The InfiniBand link data rate of 6 GBps or 3 GBps does 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, the service times of coupling operations are greater, and the actual throughput may be less than with ICB links.**

IBM Systems

IBM System z10 EC

I/O cage connectivity

IBM Systems
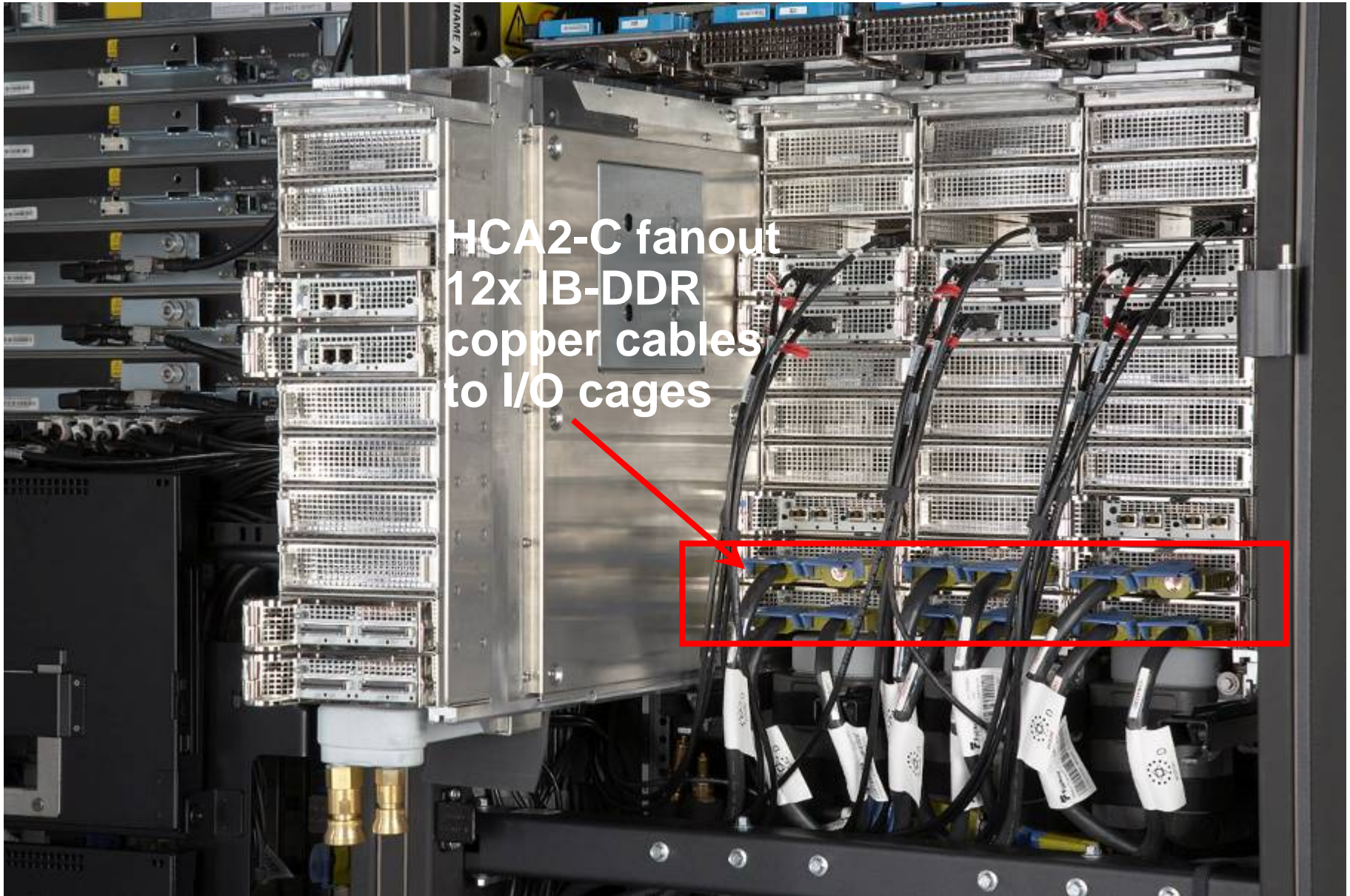
HCA2-C fanout
12x IB-DDR
copper cables
to I/O cages

IBM Systems

System z10 EC - 8 STIs in I/O cage Top view
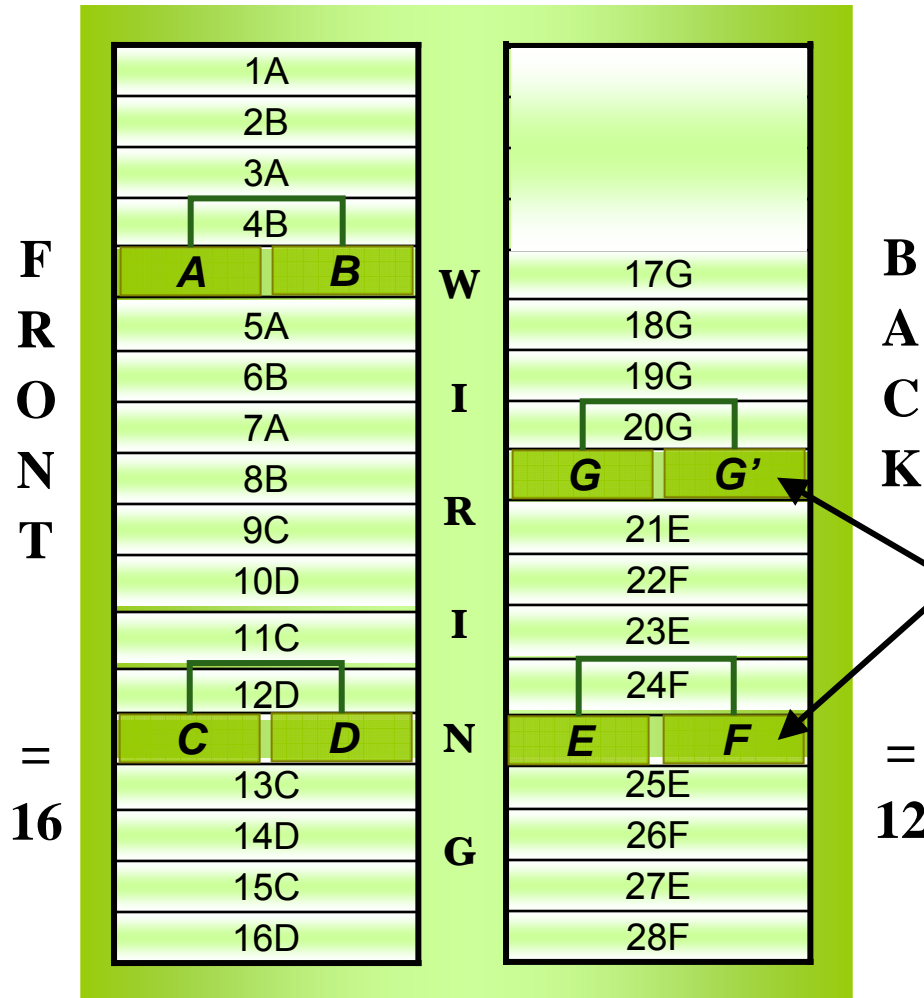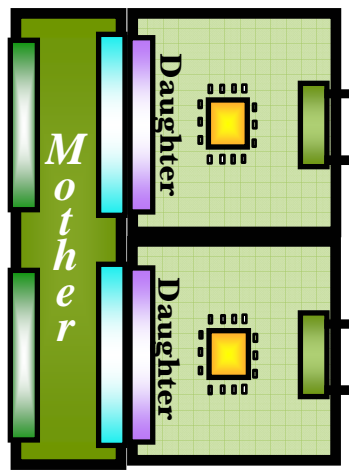
- ❑ I/O cage
  - ➢ 8 STIs
    - ▪ G' is for redundancy
- ❑ mSTI speeds for downstream channels
  - ➢ 333 MBps
    - ▪ ESCON
    - ▪ FICON Express
  - ➢ 500 MBps
    - ▪ ISC-3
  - ➢ 1 GBps
    - ▪ FICON Express2
    - ▪ OSA-Express2
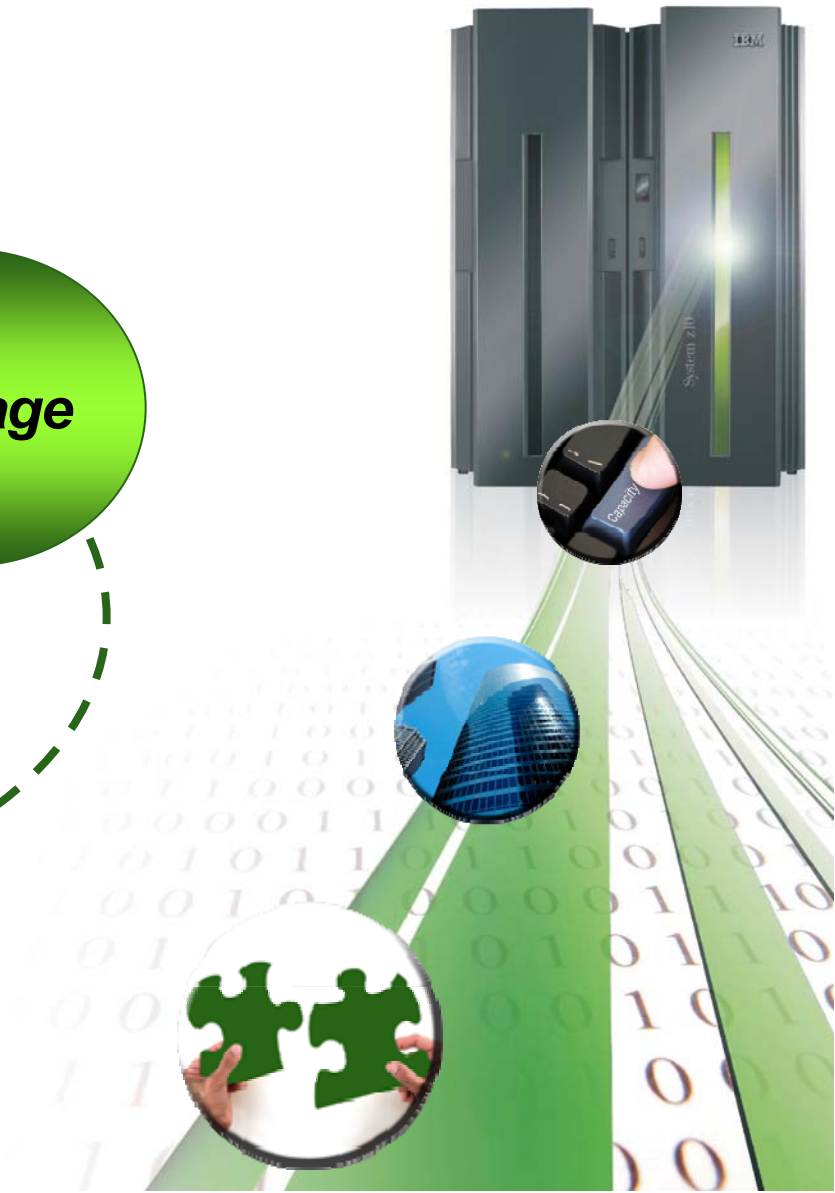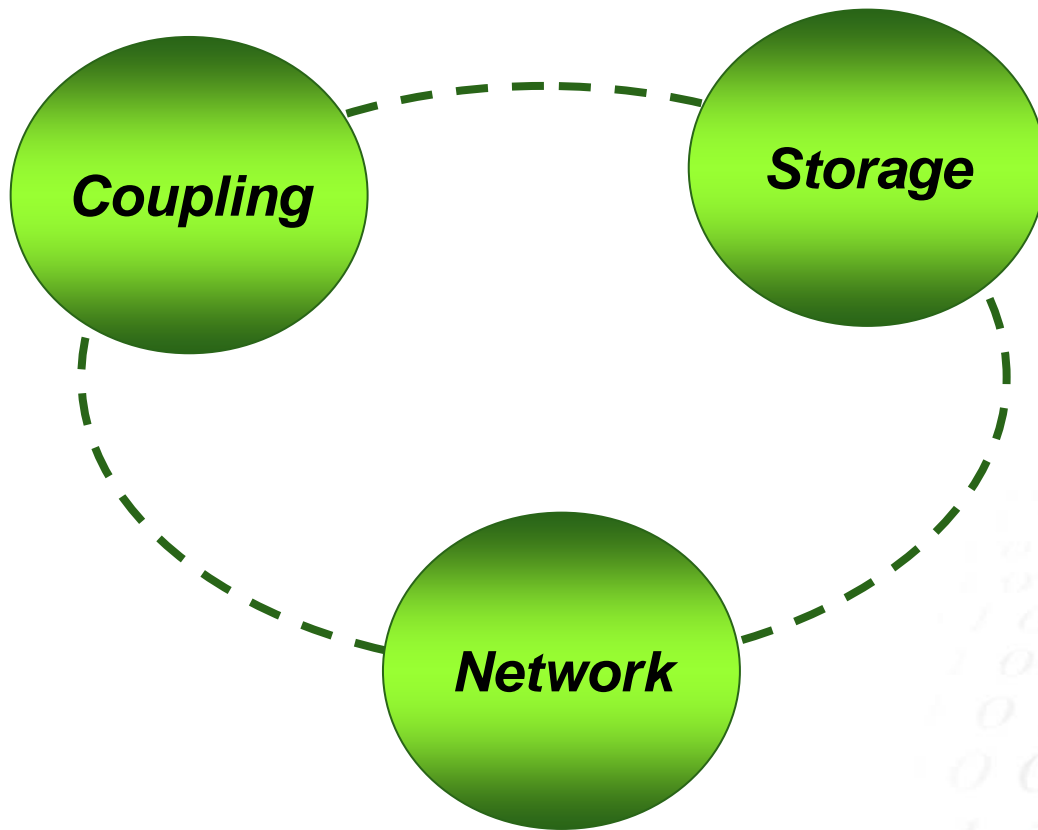 - ➢ 2 GBps
    - ▪ FICON Express4
    - ▪ **OSA-Expresss3**

New

FRONT = 16		WIRING		BACK = 12
1A				
2B				
3A				
4B				
A	*B*		17G	
5A			18G	
6B			19G	
7A			20G	
8B			*G* *G'*	
9C			21E	
10D			22F	
11C			23E	
12D			24F	
C *D*			*E* *F*	
13C			25E	
14D			26F	
15C			27E	
16D			28F	
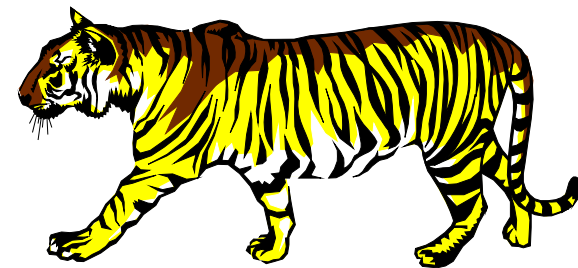
IFB-MPs

IFB-MP with Redundant I/O interconnect
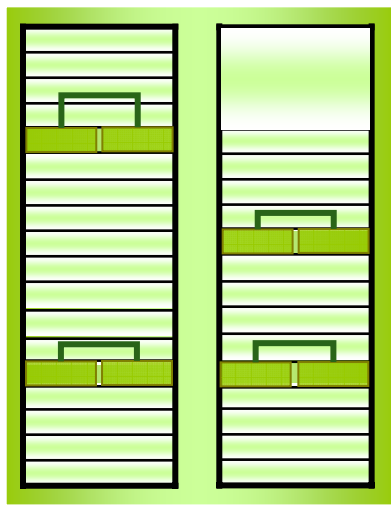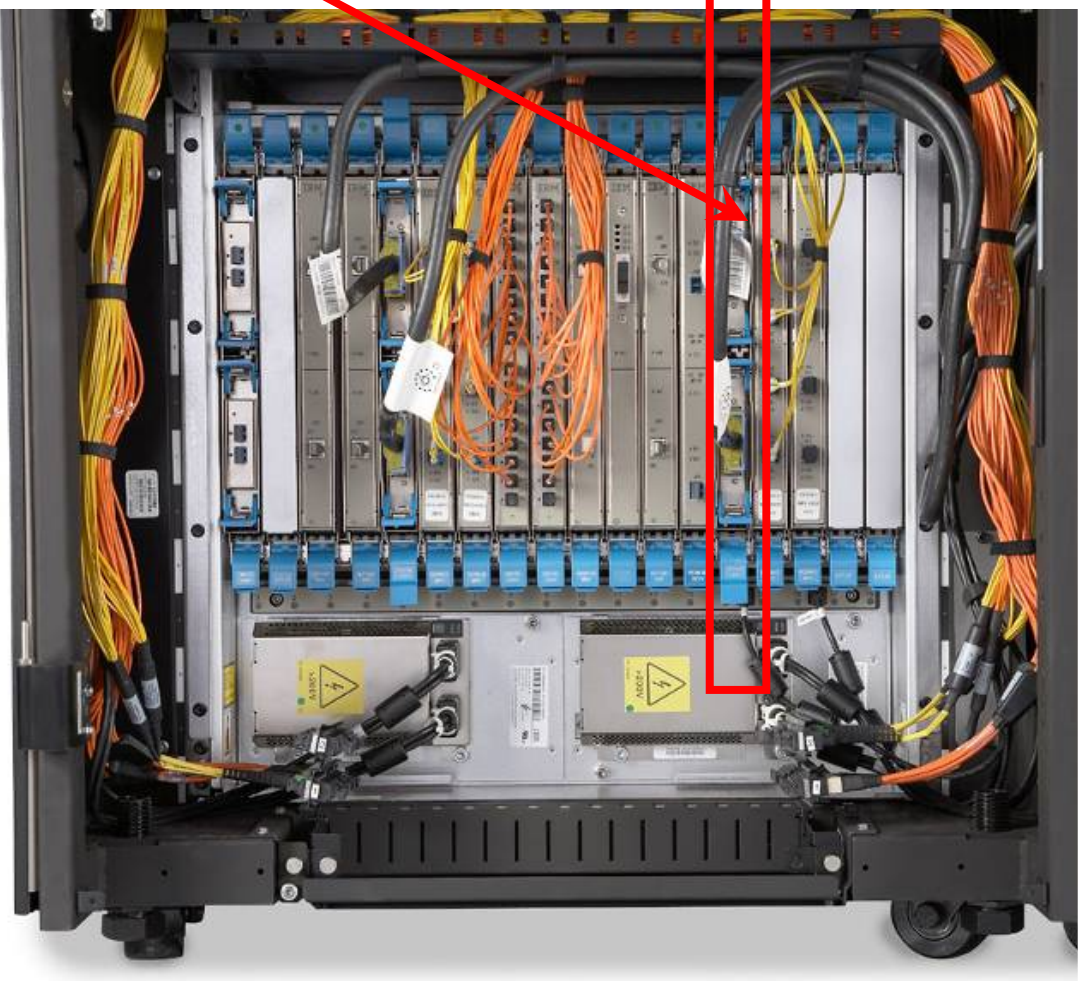
IFB-MP(multiplexer)

IBM System z10 EC

Coupling

Storage

Network

IBM Systems

System z
Storage Area Network (SAN)

Strict requirements isolation and RAS

ESCON, 17 MBps link data rate

FICON Bridge (FCV)

FICON CTC

Native FICON (FC)

1, 2, 4 Gbps link data rates

ISLs

SCSI

FICON/ FCP Intermix

FCP
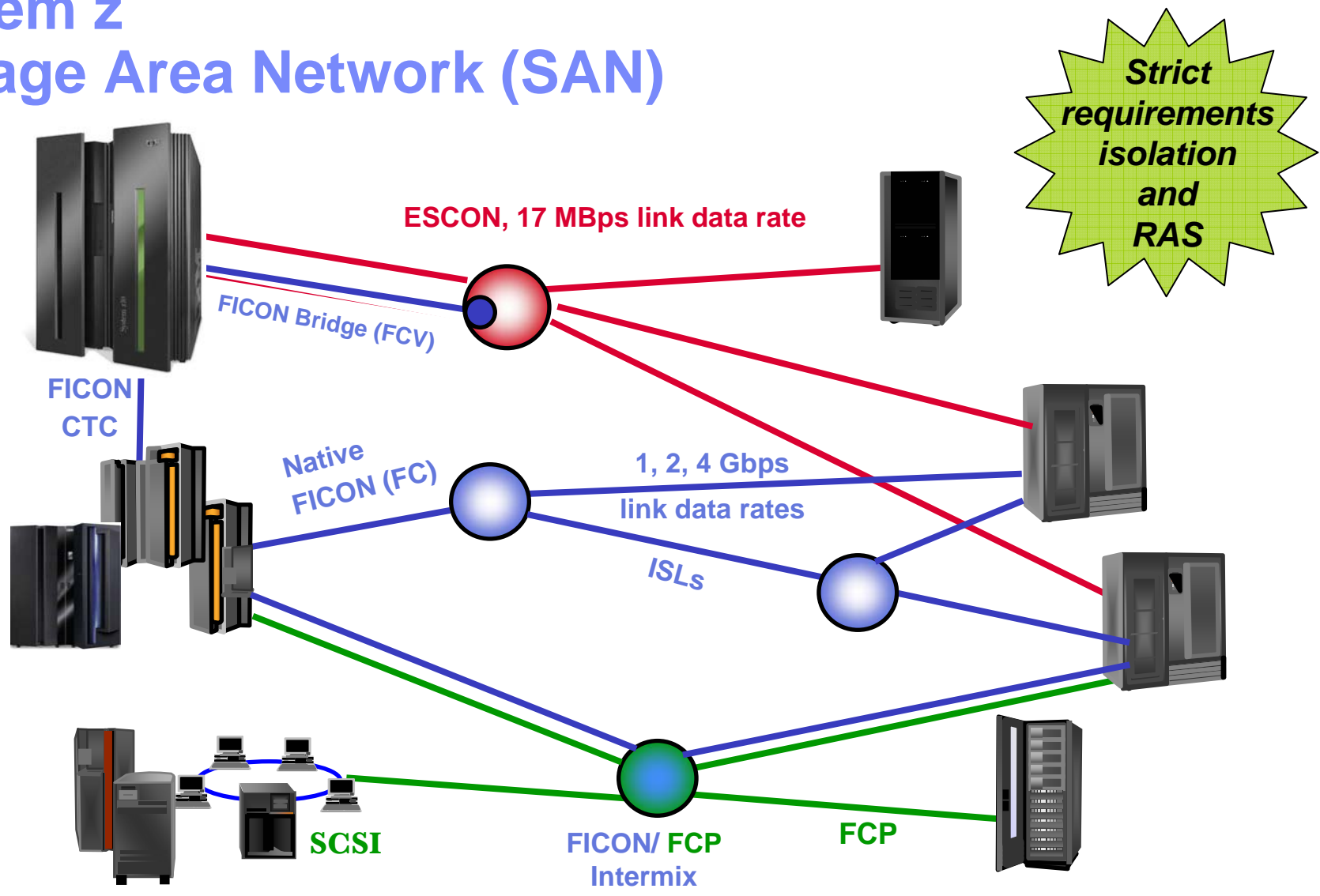
16-port ESCON

MT-RJ
connector

Increment
4 channels

Minimum
0

Maximum
1024

15
active
LED

Fiber
62.5µ MM

3 km
Distance
1.86 miles

17 MBps

IBM Systems

ESCON channel plugging

IBM Systems

FICON Express4

Minimum 0

Increment 4 channels

Maximum 336

4
SX, 4KM LX
10KM LX

Fiber
50,62.5µ MM
9µ SM

1, 2, 4 Gbps

Distance variable

LC Duplex MM

LC Duplex SM
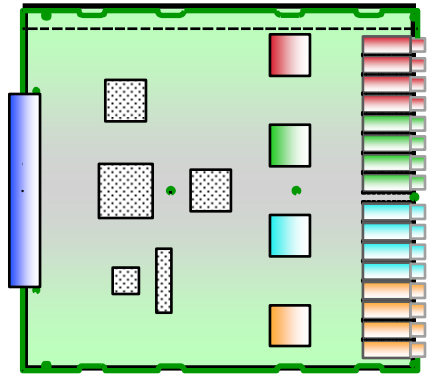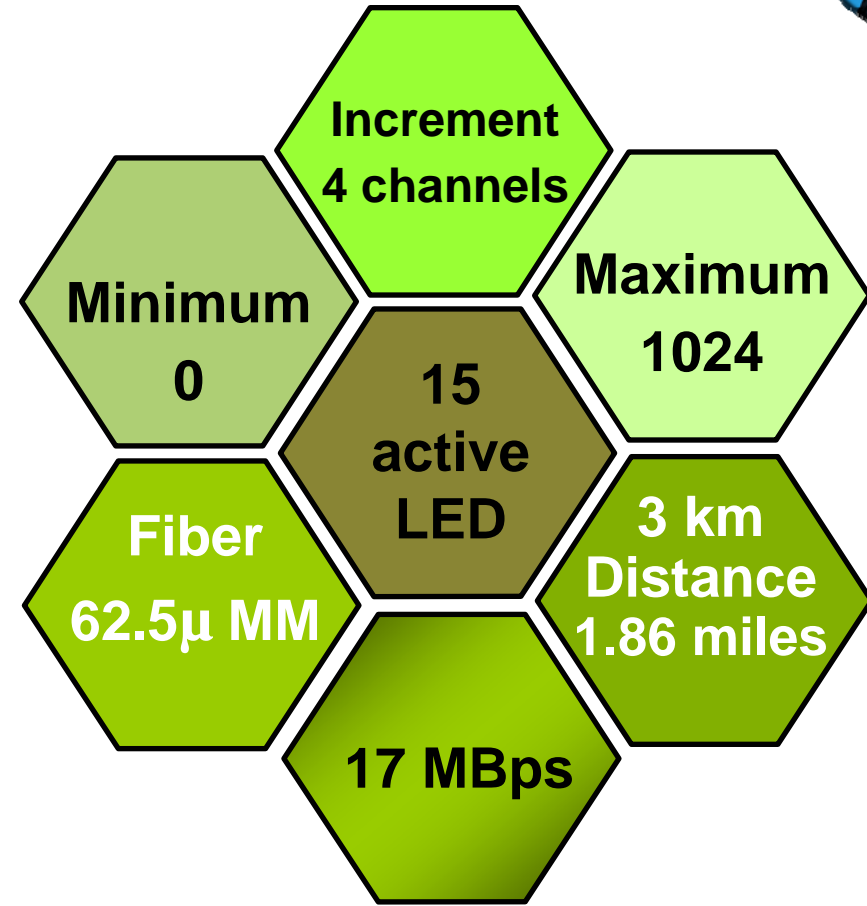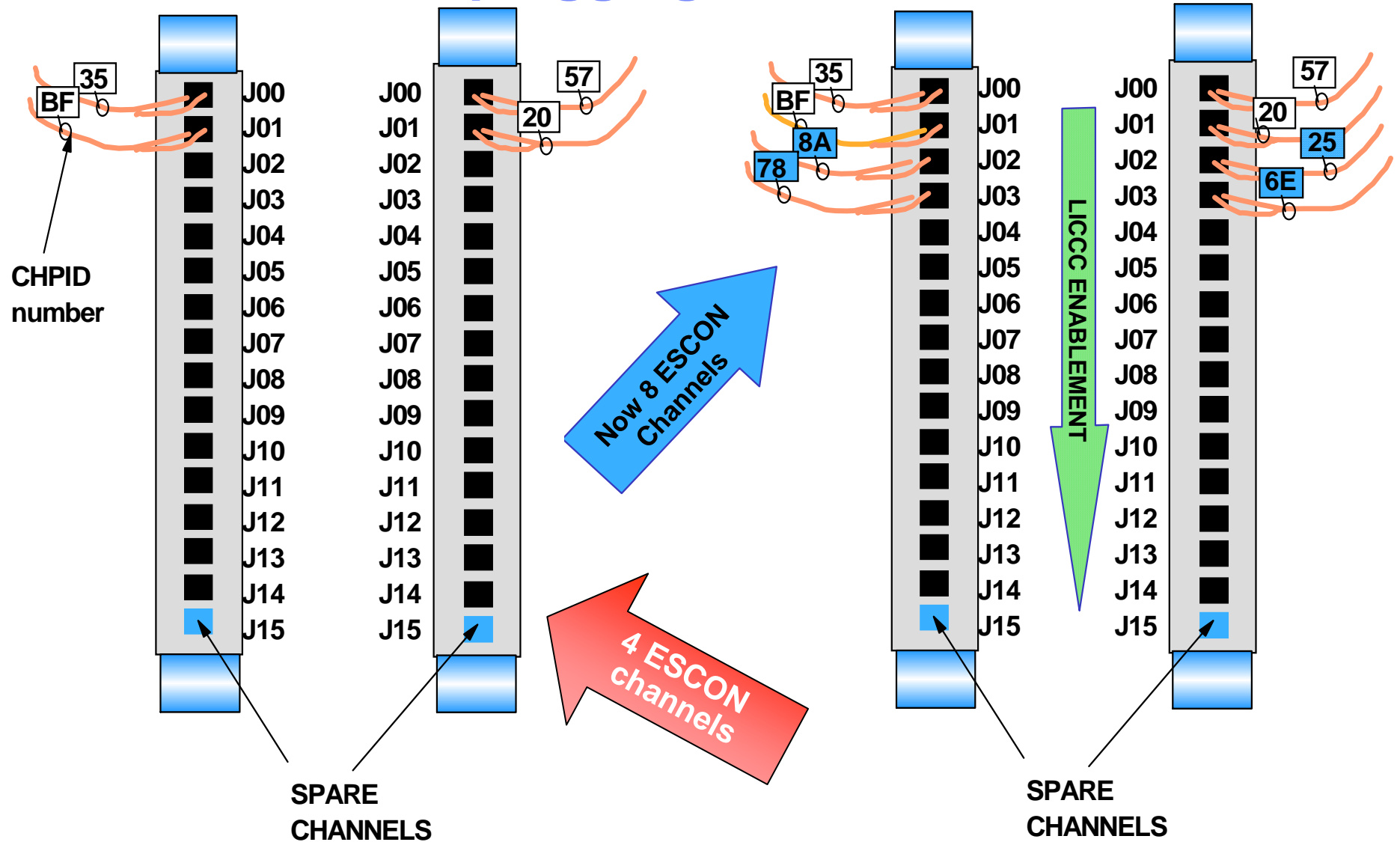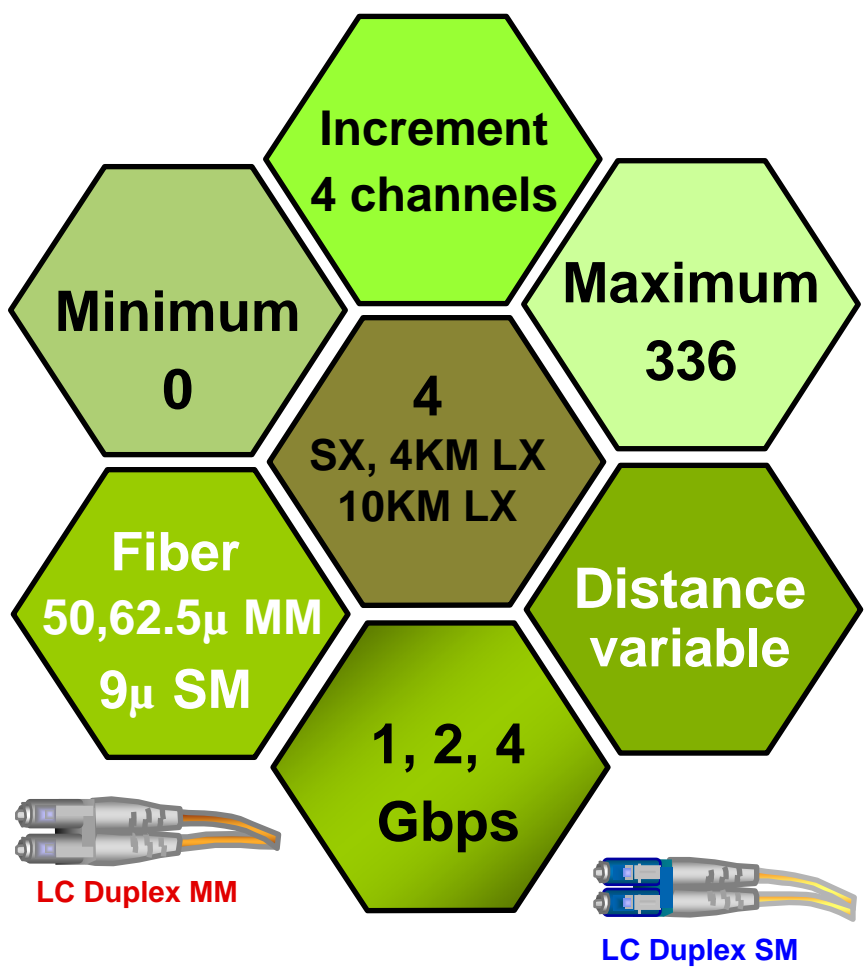
- **Personalize as:**
  - FC (Fibre Channel)
    - Native FICON
    - Channel-To-Channel (CTC)
      - z/OS, z/VM, z/VSE, z/TPF, Linux on System z
  - FCP (Fibre Channel Protocol)
    - Support of SCSI devices
      - z/VM, z/VSE, Linux on System z
  - FCV not supported

IBM Systems

FICON Express4

Example – FICON Express4 10KM LX

- **All channels same type**
  1. **4KM LX**
  2. **10KM LX**
  3. **SX**
- **Auto-negotiate**
  - ► **1, 2, or 4 Gbps**

1, 2, 4 Gbps

1, 2, 4 Gbps

1, 2, 4 Gbps

1, 2, 4 Gbps

Small form factor pluggable (SFP) optics
Concurrent repair/replace for each SFP

IBM Systems

FICON Express4 LX optics - Unrepeated distances

9 micron single mode fiber

4 kilometers (2.5 miles)

4KM LX | 4KM LX | 4KM LX | 4KM LX → **4KM LX**

9 micron single mode fiber

4 kilometers (2.5 miles)

10KM LX | 10KM LX | 10KM LX | 10KM LX → **4KM LX**

9 micron single mode fiber

10 kilometers (6.2 miles)

10KM LX | 10KM LX | 10KM LX | 10KM LX → **10KM LX**

IBM Systems
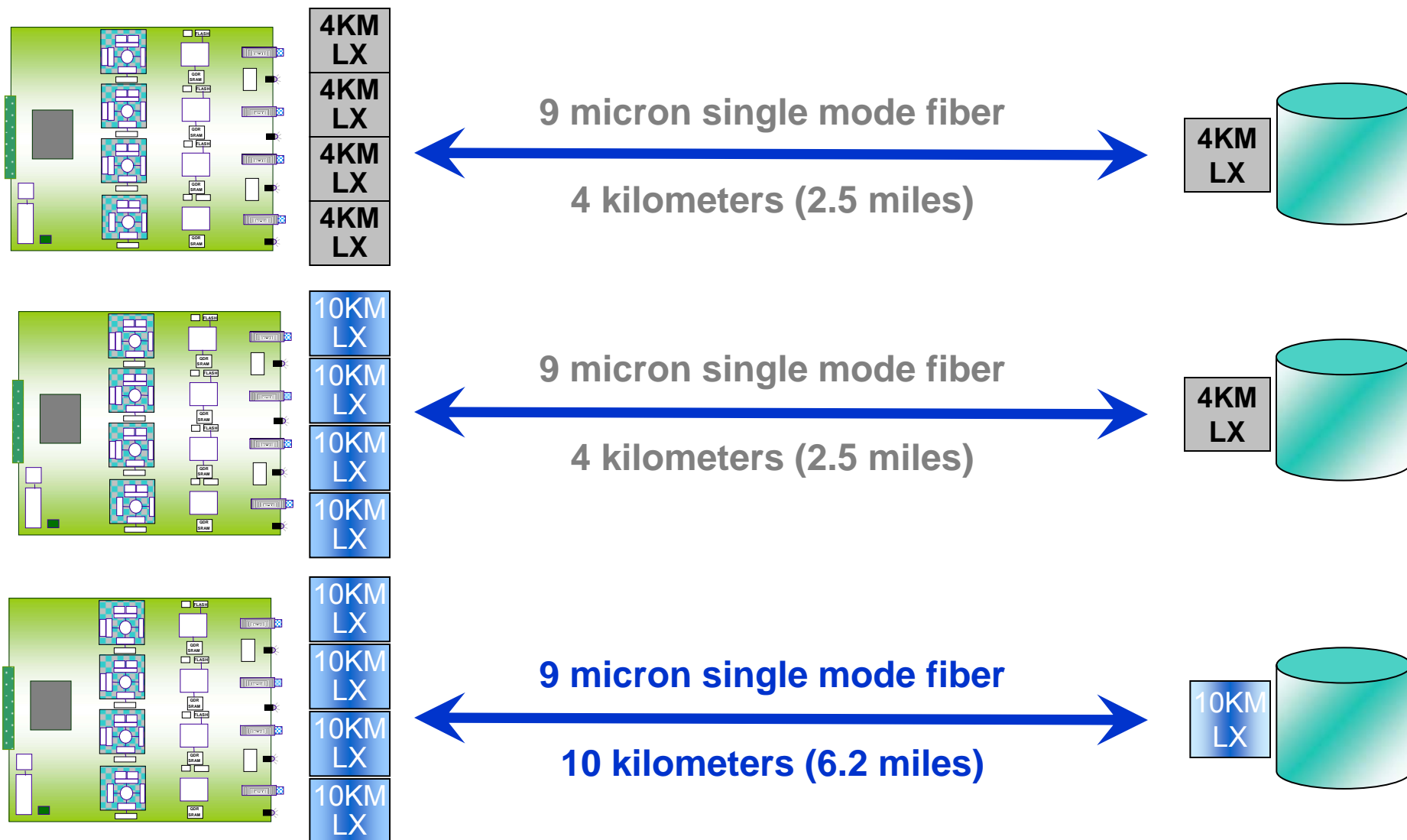
FICON features that can be "carried forward"

1 Gbps, 2 channels

Small form factor - LC Duplex

LX – configured as **FC**, **FCP**, **FCV**

SX – configured as **FC**, **FCP**
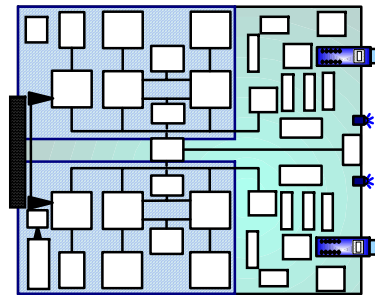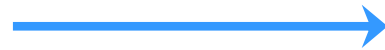
Was available on:

z990, z900

FICON Express

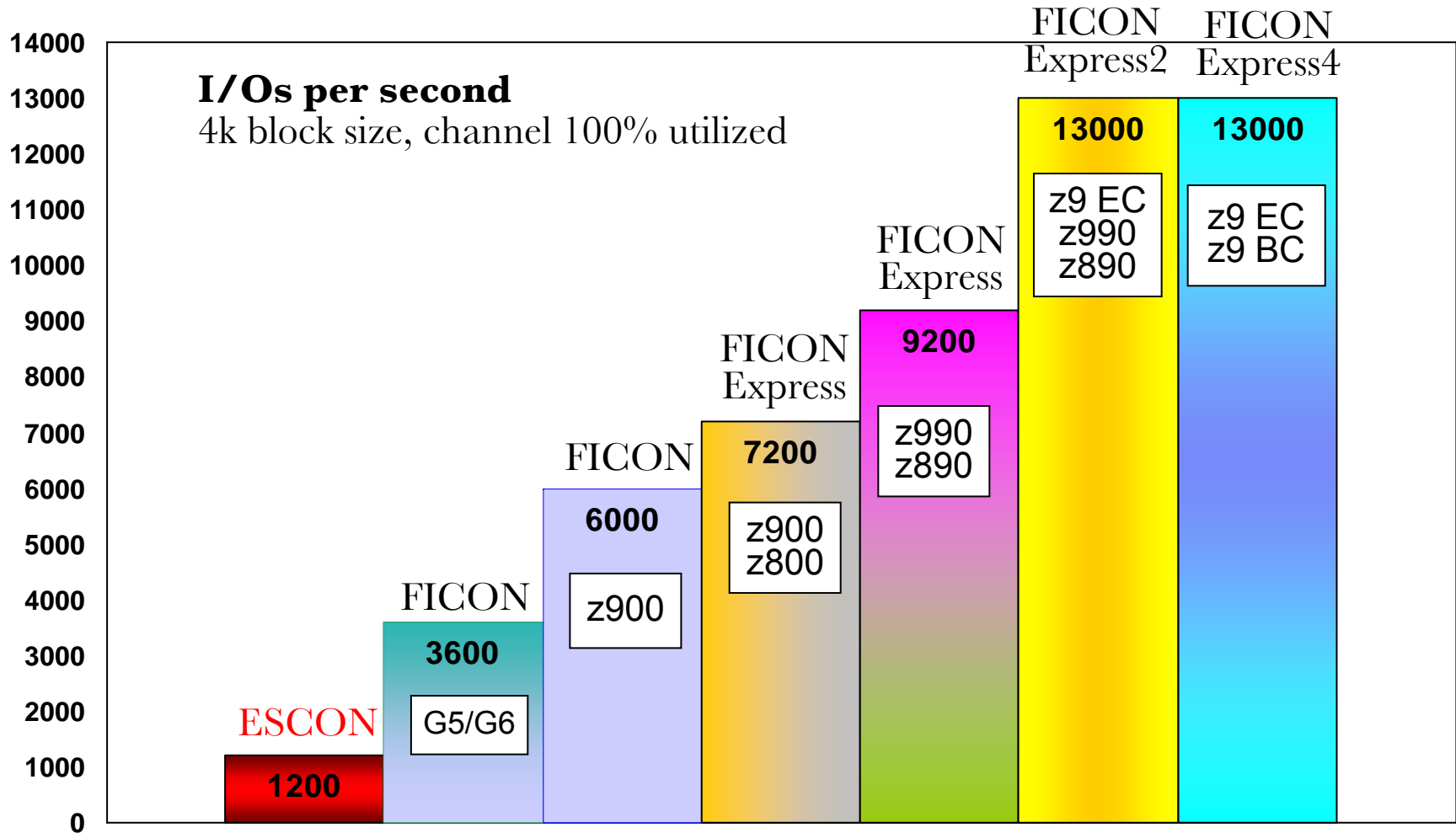**2 Gbps, 4 channels**

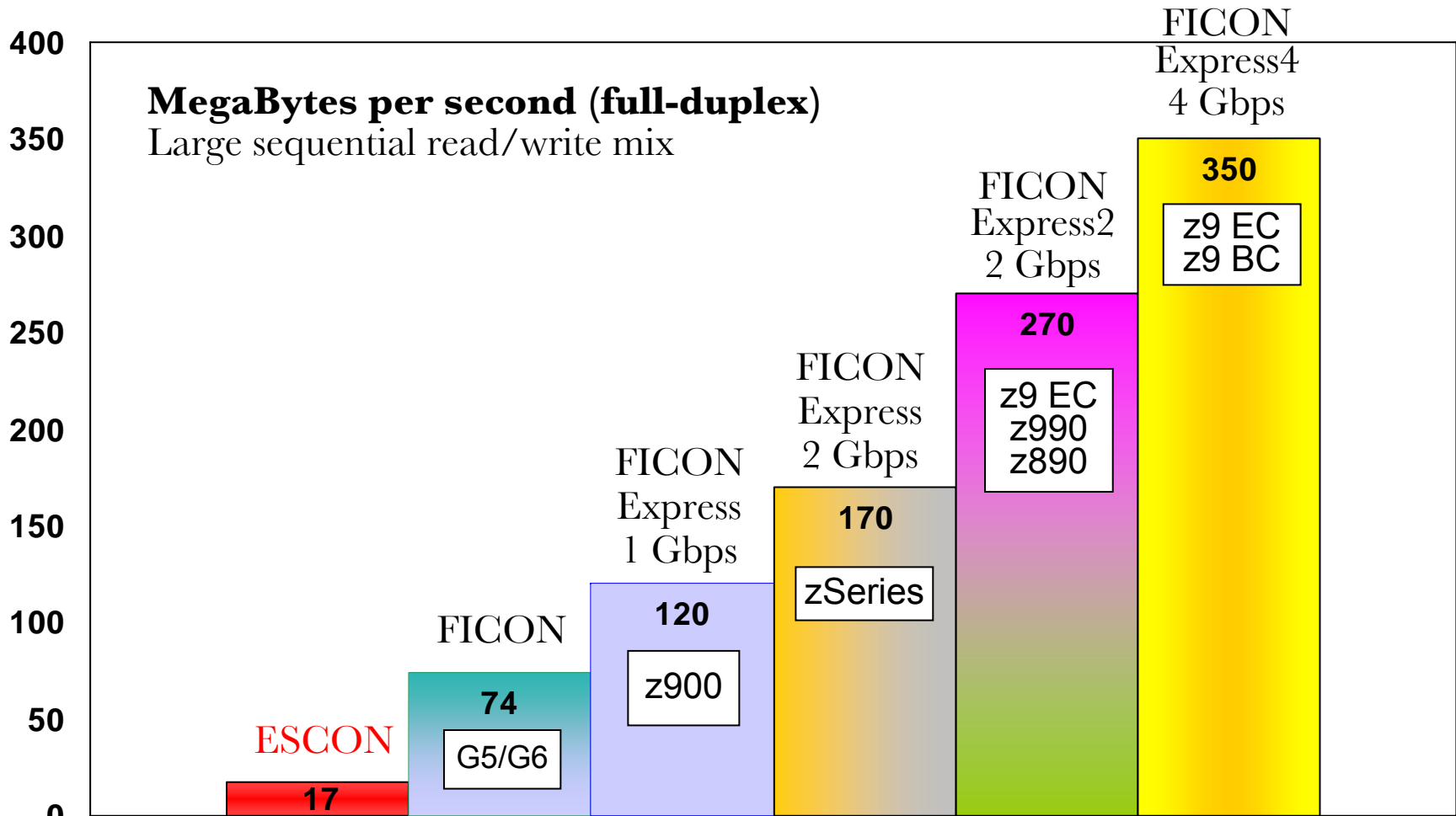LX & SX – configured as **FC**, **FCP**
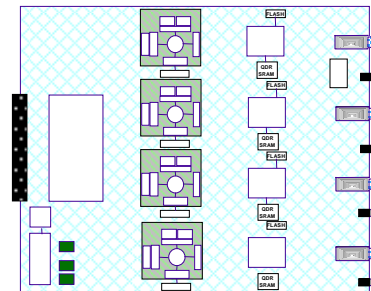
Was available on:
z9 EC, z990
FICON Express2

IBM Systems

FICON performance – Start I/Os

I/Os per second
4k block size, channel 100% utilized

Chart data (I/Os per second):
- ESCON — G5/G6 — 1200
- FICON — G5/G6 — 3600
- FICON — z900 — 6000
- FICON Express — z900, z800 — 7200
- FICON Express — z990, z890 — 9200
- FICON Express2 — z9 EC, z990, z890 — 13000
- FICON Express4 — z9 EC, z9 BC — 13000

FICON channel performance
Full duplex read/write mix data transfers

MegaBytes per second (full-duplex)
Large sequential read/write mix

Channel	Value
ESCON	17
FICON (G5/G6)	74
FICON Express 1 Gbps (z900)	120
FICON Express 2 Gbps (zSeries)	170
FICON Express2 2 Gbps (z9 EC, z990, z890)	270
FICON Express4 4 Gbps (z9 EC, z9 BC)	350

*This performance data was measured in a controlled environment running an I/O driver program under z/OS . The actual throughput or performance 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.

IBM Systems

FCP performance - Reads/writes/mix
An *80% increase on z10 EC compared to z9 EC*

I/Os per second - 4k block size - Channel 100% utilized

FICON Express4
4 Gbps

57000

New

31500

15750

z9 EC
z9 BC

z9 EC
z9 BC

z10 EC

60000
50000
40000
30000
20000
10000
0

Apr 2006   **Apr 2007**   **Feb 2008**
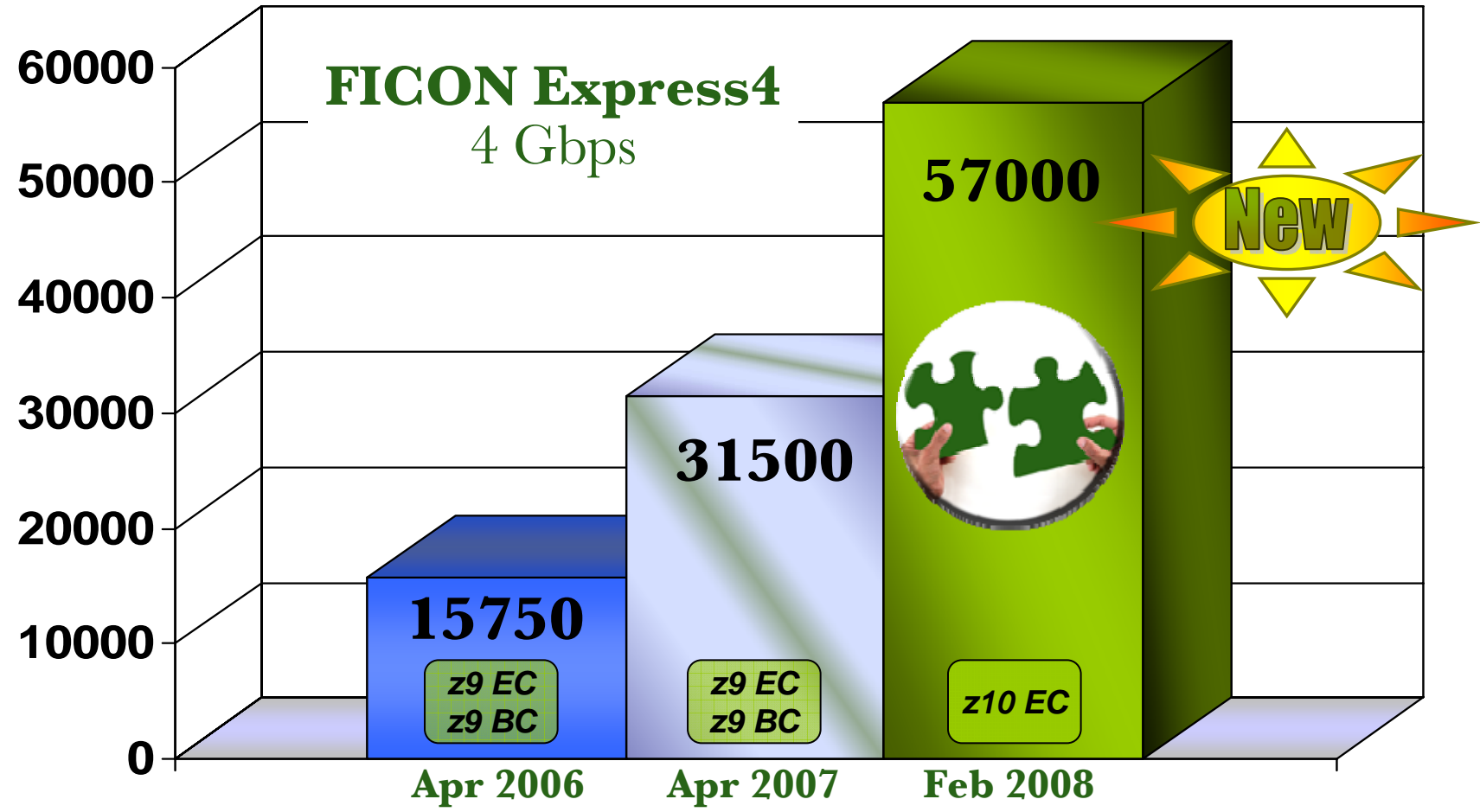
IBM Systems

System z10 EC + *Extended Distance FICON* + z/OS Global Mirror + DS8000 = Enables protection at less cost

- Optimize the FICON pacing to increase the number of commands in flight
  - **IBM System Storage DS8000 series Licensed Machine Code (LMC) level 5.3.1xx.xx (bundle version 63.1.xx.xx), or later**
  - **Available 07 March 2008**
- Enables communication over greater distances without substantial reduction to effective data rate
- Supports increased link utilization
- Can significantly reduce the cost of remote mirroring over FICON for z/OS Global Mirror (XRC) solution
  - ► **Eliminates need for more expensive 3rd party protocol-specific channel extender products**

zGM primary controller

z/OS Global Mirror

System Data Mover

SDM

Sec

zGM Secondary controller

Channel extension devices

IBM Systems

Extended distance FICON

New
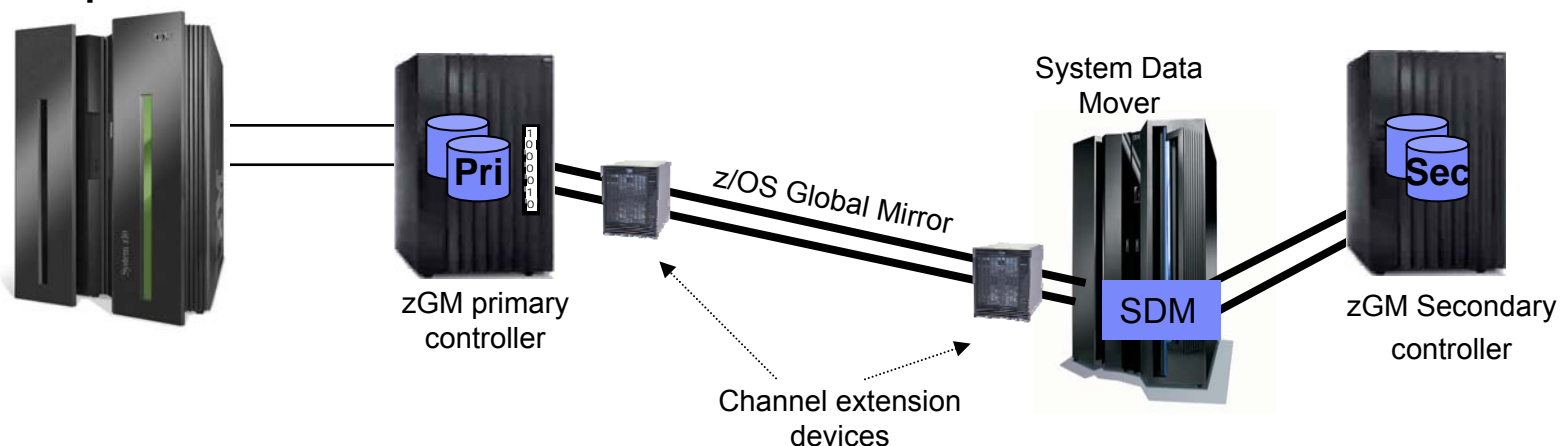
Servers, operating systems, features

- ► System z10 EC

- ► Transparent to operating systems

- ► FICON Express4, FICON Express2 (CHPID type FC – native FICON)

Fabric support

- ► Require sufficient buffer credits on FICON directors to support the link data rate at extended distance

IBM Systems

Fiber Quick Connect (FQC)

- **Available for ESCON 62.5 multimode fiber optic cabling**
- **Facilitates adds, moves/changes, deletes**
- **Harnesses factory-installed to support maximum-configured cages**
- **Reduces fiber installation time by up to <u>80%</u>**
- **Significant fiber optic cable reduction**
  - ▸ 144 ESCON cables (288 fibers) versus 2 trunks (144 fibers in a trunk)

Front View   Front View

Fiber cabling solutions

MTRJ

1

MTP connectors

MTP coupler bracket

2

FQC Harness

MTP connectors

3

72 channel trunk cable

4

SC-DC panel-mount box (rear view)

SC-DC connectors

FTS Solution

IBM Systems

Fiber Quick Connect (FQC) in System z

Replace individual cables

with trunk cables

and fiber harnesses

© 2008 IBM Corporation

IBM Systems

FQC for FICON LX channels

- **FQC - factory installation of Fiber Transport System (FTS) fiber harnesses**

- **Connection to all ESCON and FICON LX channels in all I/O cages**

- **Help minimize disruptions**

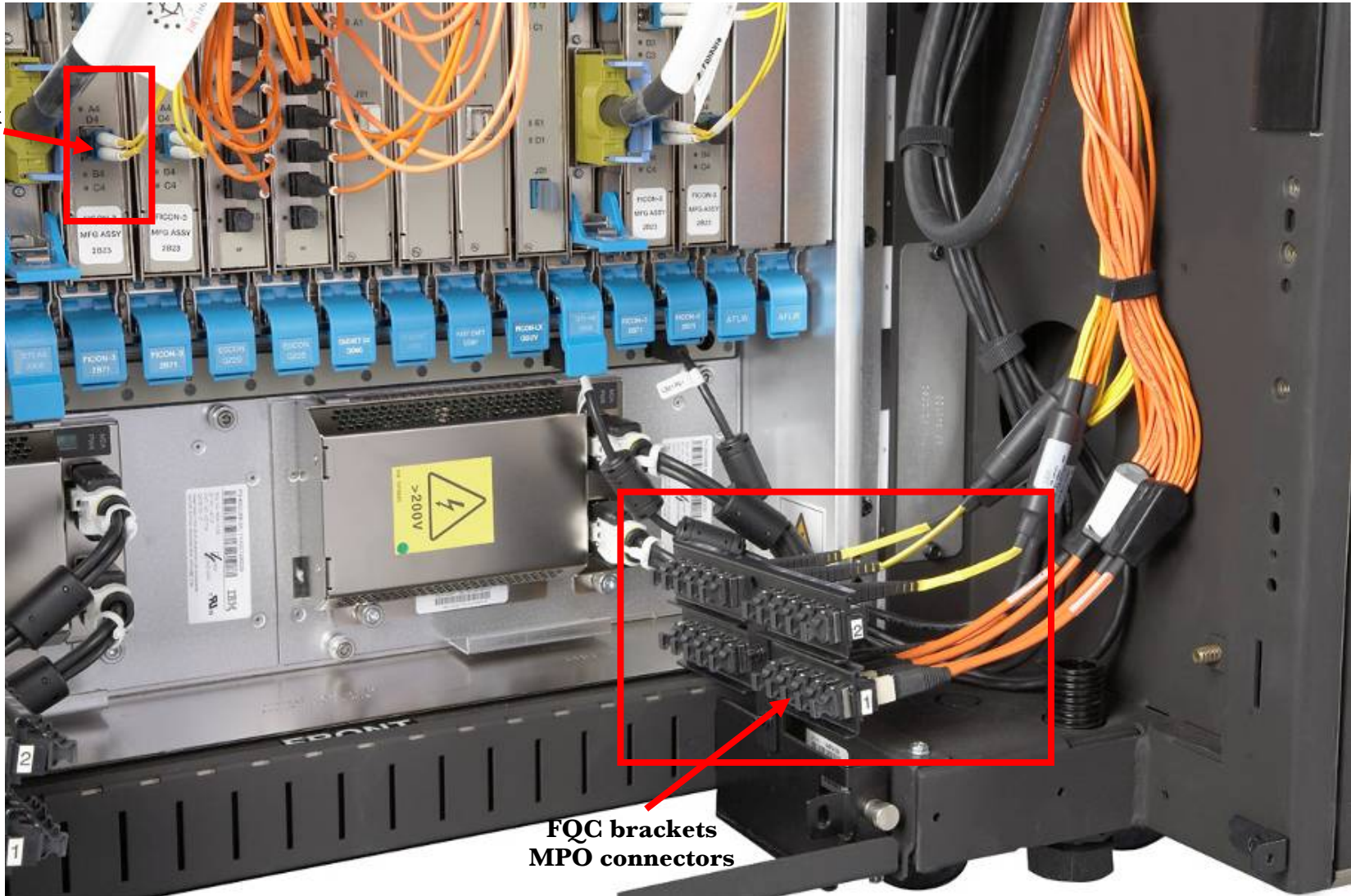- **Isolate fiber cabling activities away from the active system**
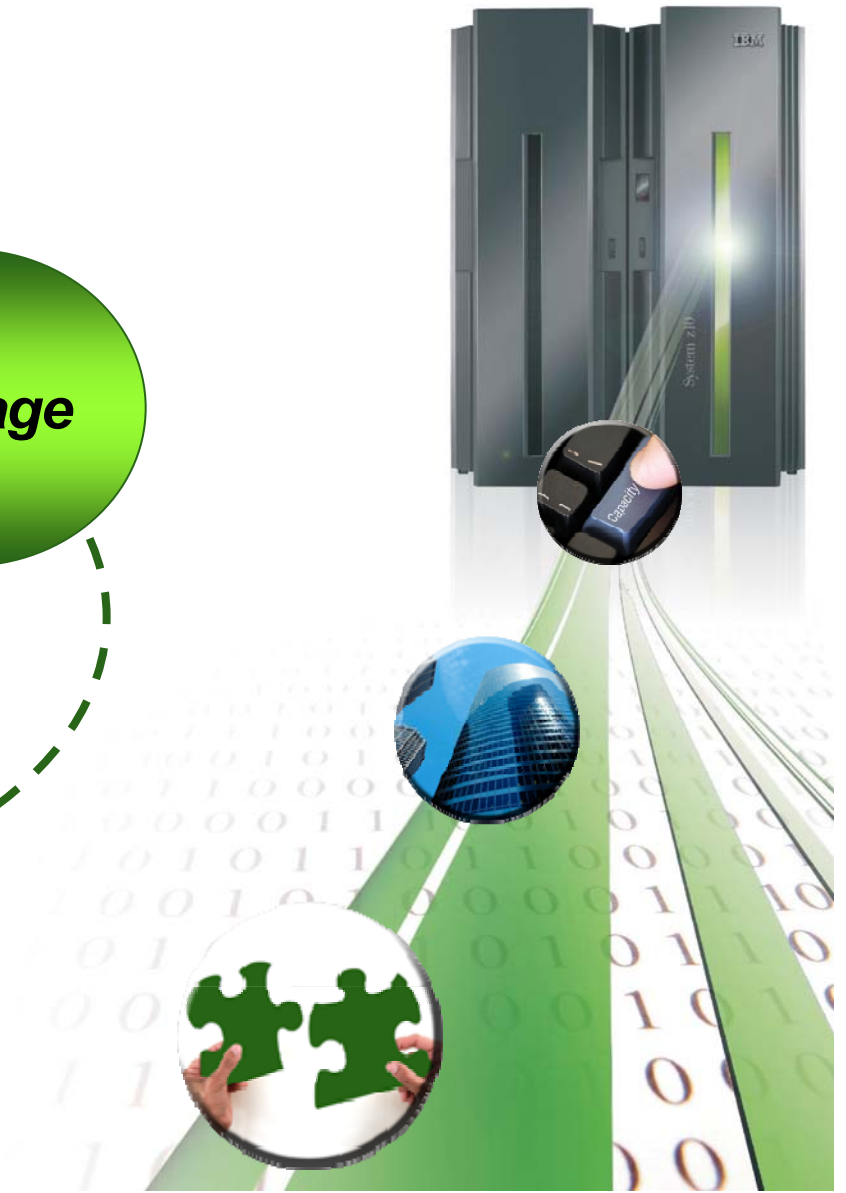
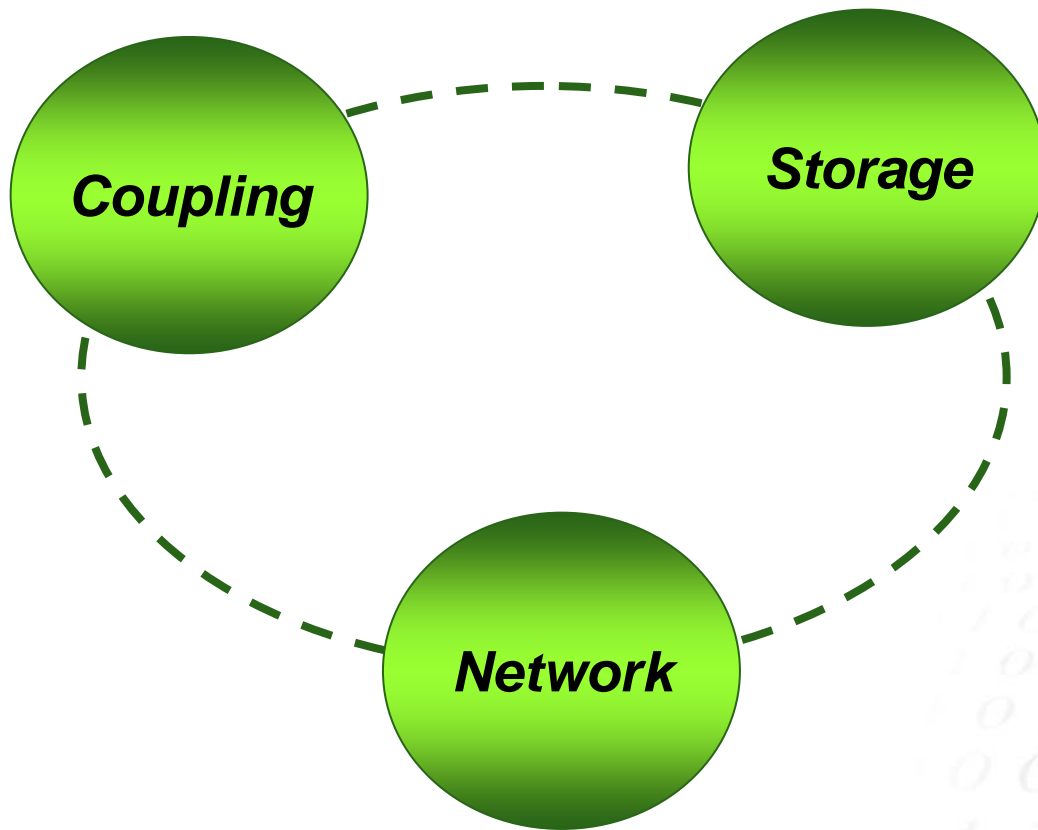LC Duplex
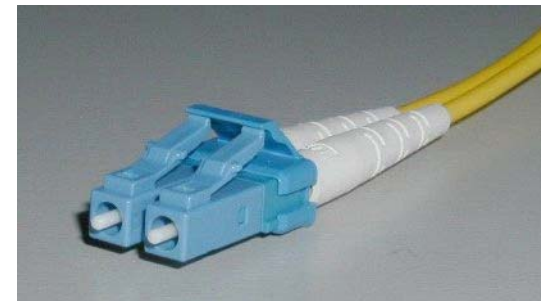
System z10 EC Cabling in frames

IBM Systems

Fiber Quick Connect (FQC) harnesses

LC Duplex connector

FQC brackets
MPO connectors

IBM Systems

IBM System z10 EC

**Coupling**

Storage

Network

IBM Systems

System z HiperSockets

- **High-speed "intraserver" network**

- **Independent, integrated, virtual LANs**

- **Communication path – system memory**

- **Communication across LPARs**

  - ► **Single LPAR - connect up to 16 HiperSockets**

  - ► **One HiperSocket shared by up to 30 LPARs**

- **Support for multiple LCSS's & spanned channels**

- **Virtual LAN (IEEE 802.1q) support**

- **Broadcast support for IPv4 packets**

- **No physical cabling or external connections required**

z/OS LP1

Linux LP2

HiperSockets

z/VSE LP3

New

L2 z/VM LP5

L3 z/VM LP4

IBM Systems
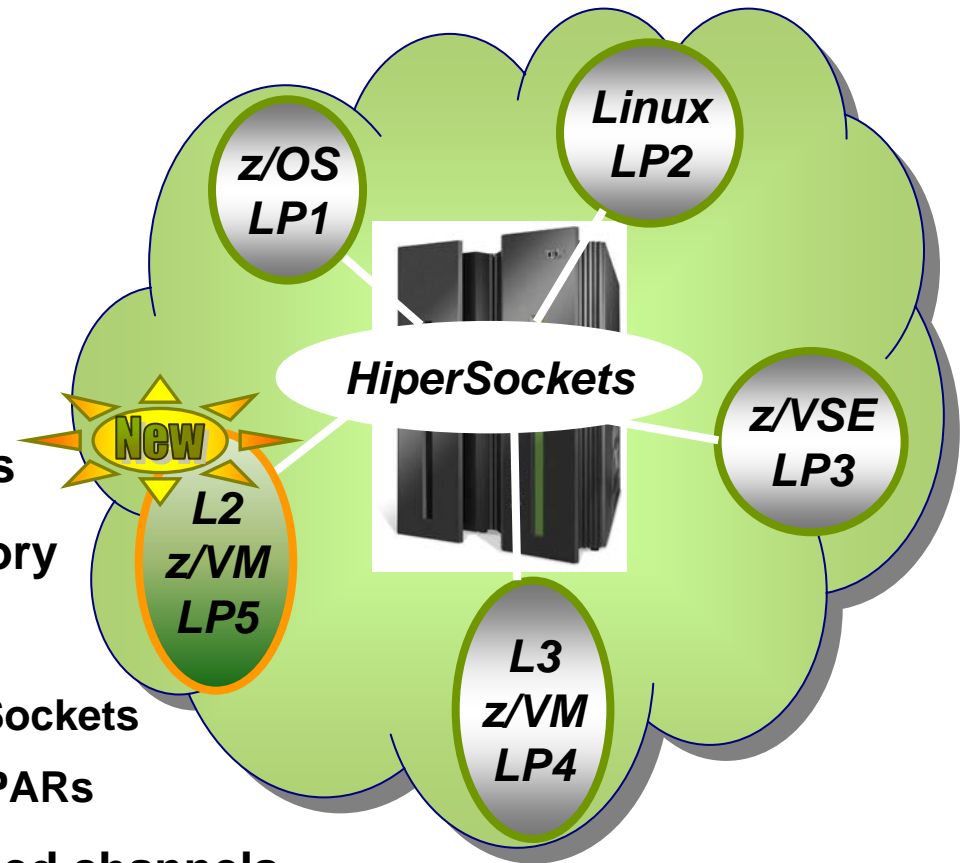
z10 EC HiperSockets Enhancements

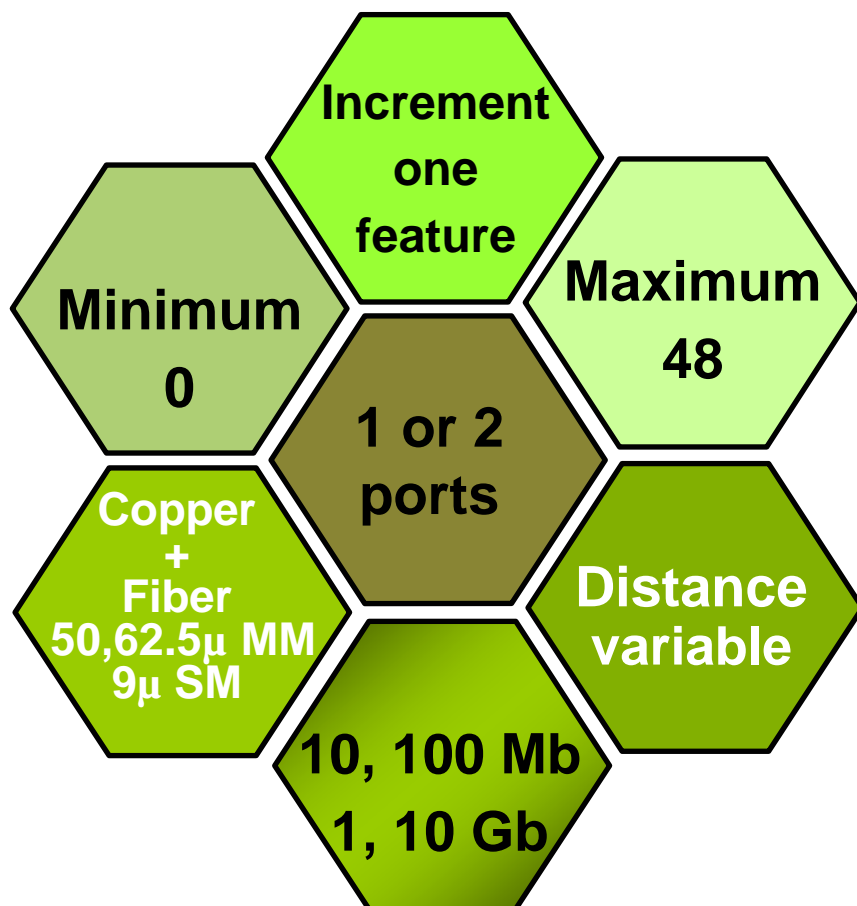- **HiperSockets Multiple Write Facility**
  - ► **Performance improvements**
    - ● **For the streaming of bulk data over a HiperSockets link between LPARs**
    - ● **Allows receiving LPARs to process a much larger amount of data per I/O interrupt**
      - – **z/OS V1.10***
        - ♦ **Transparent to software in receiving LPARs**
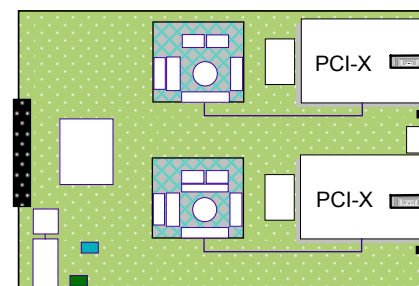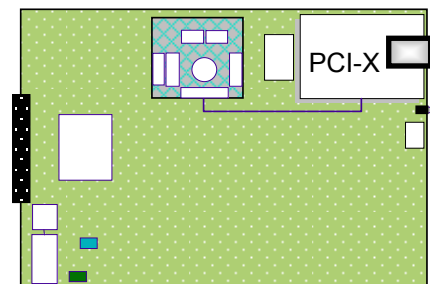   - – **z/VM 5.2 – Guest support**

- **HiperSockets Layer 2 support**
  - ► **Hosting of new workloads**
  - ► **Host non-IP protocols (IPX, NetBIOS, SNA)**
  - ► **Bridge from and into distributed switched fabrics**
  - ► **Supports broadcast, unicast, or multicast**
  - ► **VLANs: In Layer 2 the same rules apply as for Layer 3 VLAN handling.**
    - ● **Linux on System z**
  - ► **Layer 3 applications cannot communicate with Layer 2 applications**
  - ► **z/VM 5.2 or higher – Guest support**

IBM Systems

Open Systems Adapter2

Increment one feature

Minimum 0

Maximum 48

1 or 2 ports

Copper + Fiber 50,62.5µ MM 9µ SM

Distance variable

10, 100 Mb 1, 10 Gb

- **OSA-Express2 10 GbE**
  - ▶ **10 Gbps**
    - • **LR – 9µ fiber**
    - • **SC Duplex connector**

- **1000BASE-T Ethernet**
  - ▶ **10, 100, 1000 Mbps**
    - • **Category 5 copper, RJ-45**
- **GbE**
  - ▶ **1 Gbps**
    - • **SX – 50, 62.5µ fiber, LC Duplex**
    - • **LX – 9µ fiber, LC Duplex**

IBM Systems

Announced 05 May on System z10

✓**OSA-Express3 GbE**
✓**Availability of OSA-Express3 10 GbE**
May 30, 2008

IBM Systems

Open Systems Adapter3

Increment one feature

Minimum 0

Maximum 96

2 or 4 ports

Fiber 50,62.5µ MM 9µ SM
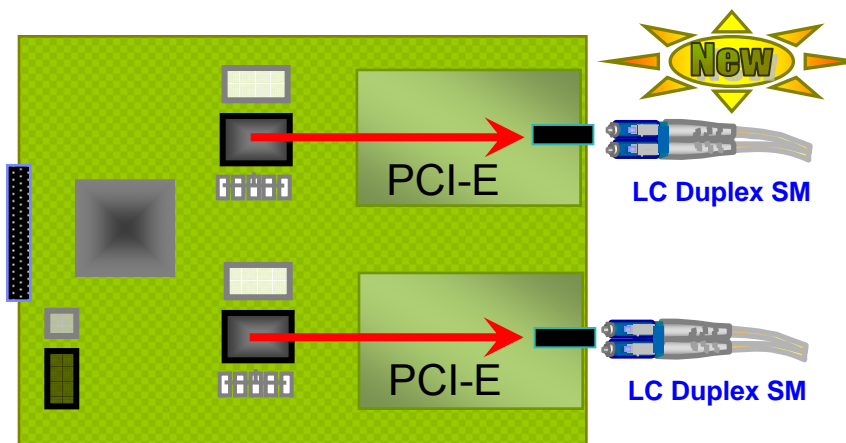
Distance variable

1 Gbps 10 Gbps

IBM Systems
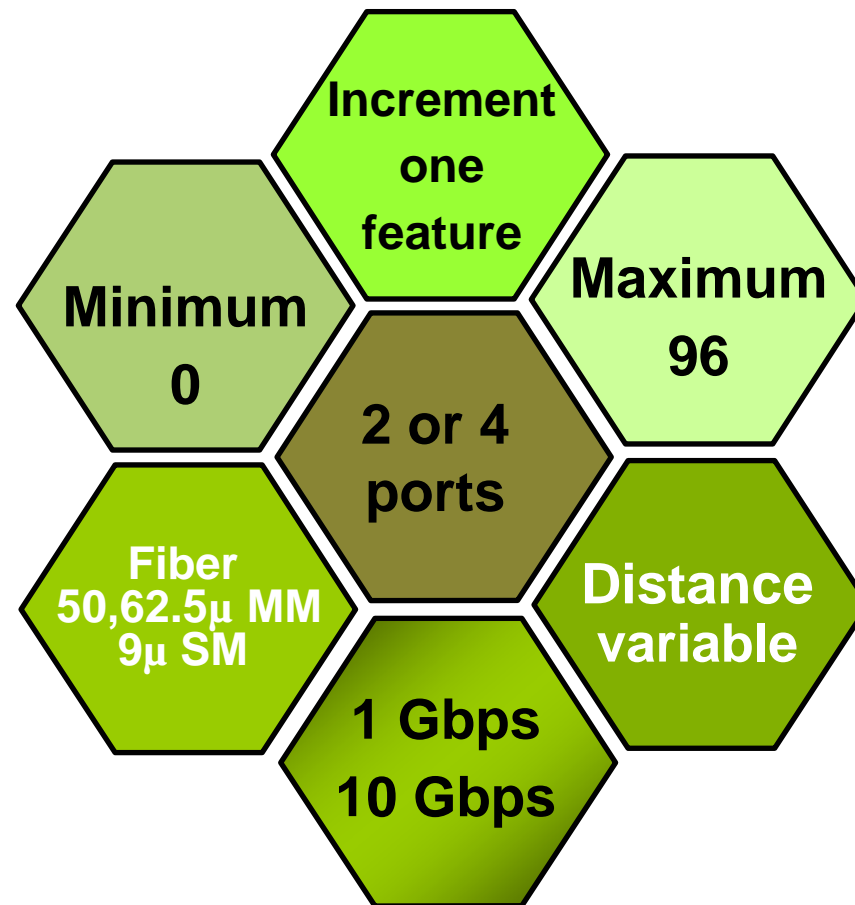
OSA-Express3
May 30, 2008

- **Double the port density**
  - ► **CHPID type OSD (QDIO)**
  - ► **Two CHPIDs per feature**
- **Reduced latency**
  - ► **Hardware data router**
- **Improved throughput – standard & jumbo frames**
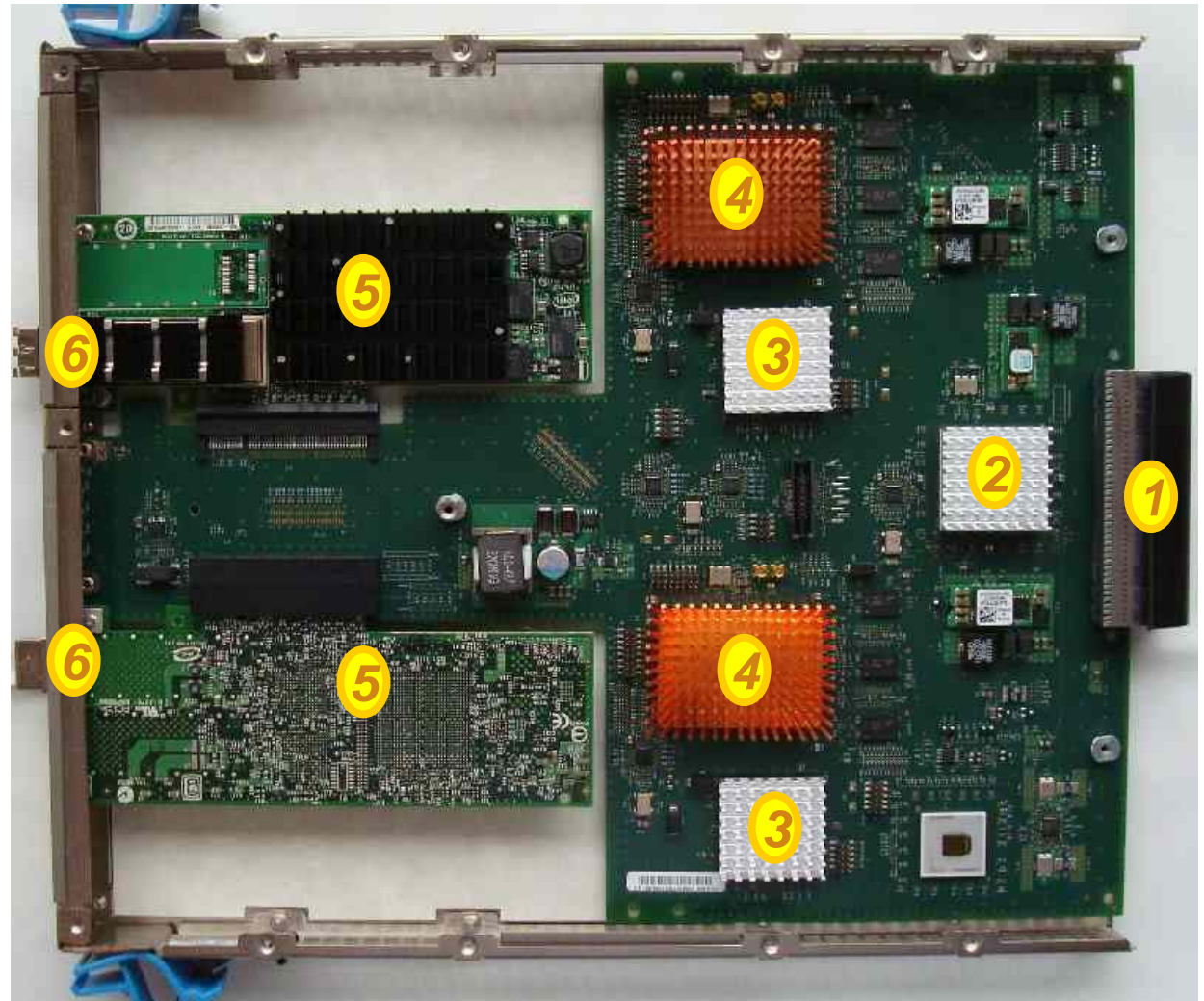 - ► **New microprocessor**
  - ► **New PCI adapter**

PCI-E
PCI-E
LC Duplex SM
LC Duplex SM

10 GbE LR (#3370)

PCI-E
PCI-E

GbE SX (#3363) GbE LX (#3362)

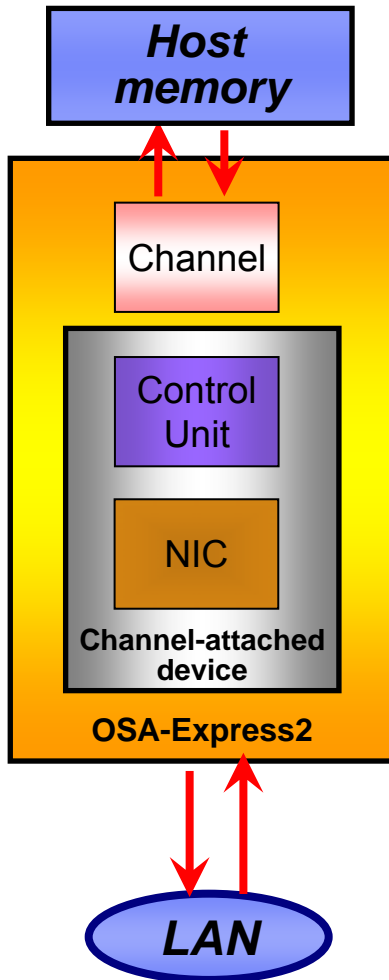|	OSA-Express2 10 GbE GbE	OSA-Express3
Microprocessor	500 MHz 448 MHz	660 MHz
PCI bus	PCI-X	PCI-E G1

IBM Systems

OSA-Express3 10 Gigabit Ethernet

1. mSTI – 2 GBps
2. IBM multiplexer
3. FPGA
4. IBM ASIC
5. PCI-E adapter
6. LC Duplex connector with LR transceiver supporting 9μ single mode fiber

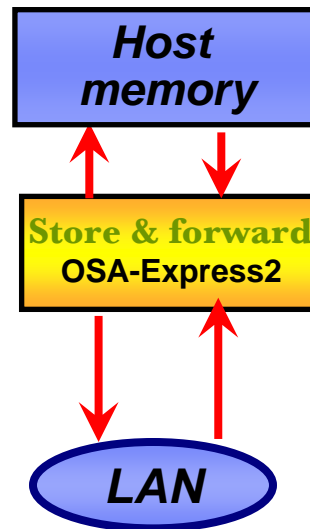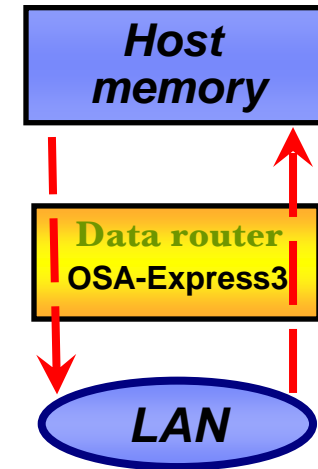IBM Systems

Open Systems Adapter data paths

non-QDIO/LCS

Host memory

Channel

Control Unit

NIC

Channel-attached device

OSA-Express2

LAN

QDIO

Host memory

Store & forward
OSA-Express2

LAN

QDIO

Host memory

Data router
OSA-Express3

LAN

- **OSA-Express3**
- **Direct memory access**
  - ► Hardware data router
  - ► Eliminates "store and forward" delays

IBM Systems

OSA-Express3 performance benchmarks

- **OSA-Express3 round-trip networking time up to a 45% reduction in latency**
  - **10 GbE**
    - **Minimum round-trip time at the TCP/IP application layer 40% less than with OSA-Express2 10 Gigabit Ethernet**
  - **GbE**
    - **45% less than with OSA-Express2 Gigabit Ethernet**

- **OSA-Express3 10 GbE throughput up to a 4x improvement**
  - **Mix of inbound and outbound TCP streams traffic**
    - **1 gigabytes per second (1 GBps, i.e., 1,000,000,000 bytes per second) when the packet maximum transmission unit (MTU) size is 1492 bytes**
    - **1.1 GBps when MTU size is 8992-byte**

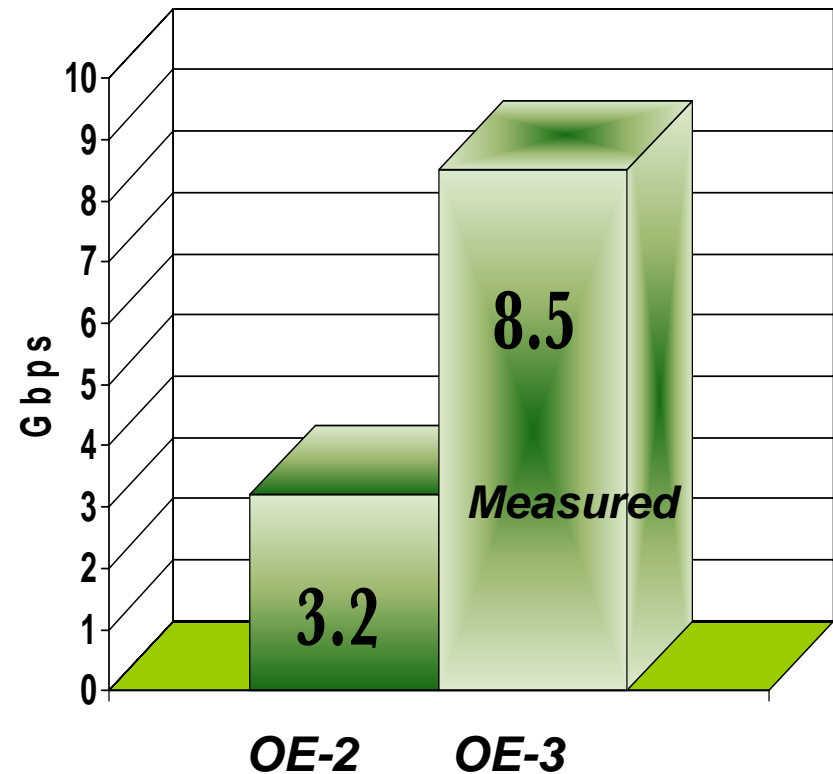**Improvements from 3x to 4x that of OSA-Express2 10 GbE on z9 EC system**
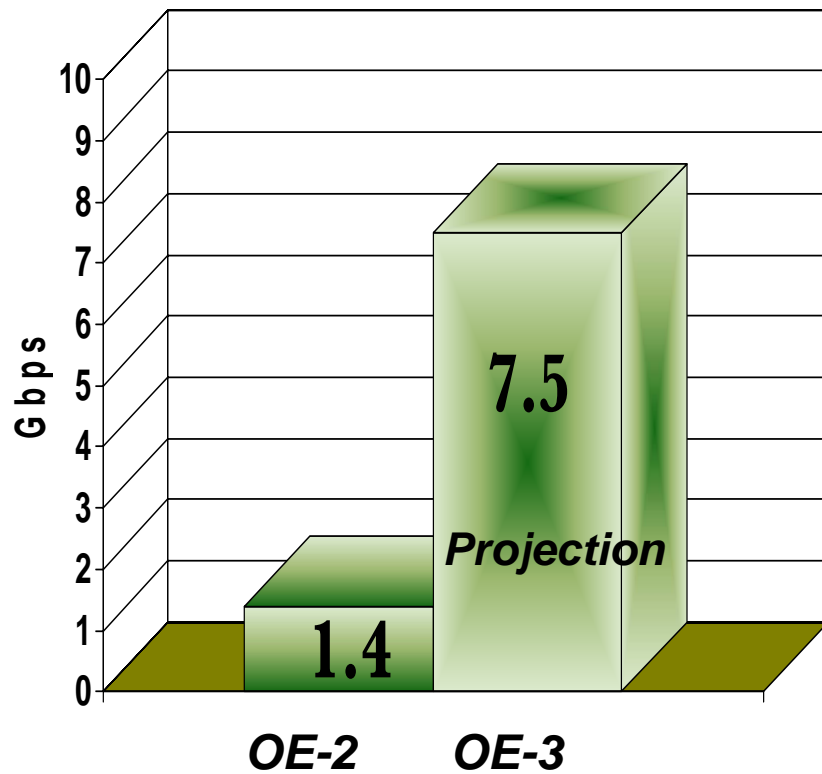
IBM Systems

*** Projections – performance ***
OSA-Express3 10 Gigabit Ethernet

New

Full-duplex data transfers
Mixed direction streams workload

Standard frames
1492 bytes

Jumbo frames
8992 bytes

Standard frames chart (Gbps):
- OE-2: 1.4
- OE-3: 7.5 *Projection*

Jumbo frames chart (Gbps):
- OE-2: 3.2
- OE-3: 8.5 *Measured*

© 2008 IBM Corporation

OSA-Express3 software requirements

- **Exploitation of 4 ports on Gigabit Ethernet**
  - ► **z/OS V1.9 with PTFs (planned to be available third quarter 2008).**
  - ► **z/VM V5.2 with PTFs (planned to be available July 2008).**
  - ► **z/VSE V4.1 with PTFs (planned to be available July 2008).**
  - ► **z/TPF 1.1 PUT 4 with APARs (planned to be available June 2008).**
  - ► **Linux on System z distributions**

- **10 Gigabit Ethernet and use of 2 of the 4 ports on Gigabit Ethernet**
  - ► **/OS V1.7.**
  - ► **z/VM V5.2.**
  - ► **z/VSE V3.1 with PTF.**
  - ► **TPF V4.1 at PUT 13 with PTF.**
  - ► **z/TPF 1.1.**
  - ► **Linux on System z distributions:**
    - ● **Novell SUSE SLES 9 and SLES 10.**
    - ● **Red Hat RHEL 4 and RHEL 5.**

IBM Systems

OSA-Express3 4-port GbE features - Benefits

- Reduced latency (up to 45% reduction) and increased throughput (up to 4x) for applications (software changes are not required to exploit the two ports on OSA-Express3 10 GbE or to exploit two of the four ports on OSA-Express3 GbE SX or LX)

- More physical connectivity to service the network and less resources:

    ► Fewer CHPIDs to define and manage

    ► Reduction in the number of required I/O slots

    ► May reduce the number of I/O cages

    ► Double the port density of OSA-Express2

    ► Satisfies requirement for more than 48 LAN ports (now up to 96 ports)

- 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" to the operating system.

IBM Systems

OSA-Express3 GbE support of OSA for NCP Preview

- **Connectivity between System z operating systems and IBM Communication Controller for Linux (CCL). CCL**

- **Allows you to keep your business data and applications on the mainframe operating systems while moving NCP functions to Linux on System z.**

- **CCL helps to simplify their network infrastructure while supporting traditional Systems Network Architecture (SNA) functions such as SNA Network Interconnect (SNI).**

- **Communication Controller for Linux on System z**

  - ► **Program Number 5724-J38**

  - ► **Helps improve network availability by replacing token-ring networks and ESCON channels with an Ethernet network**

IBM System z10 EC
Solutions

Scalability

Capacity

Coupling

Storage

Network

IBM Systems
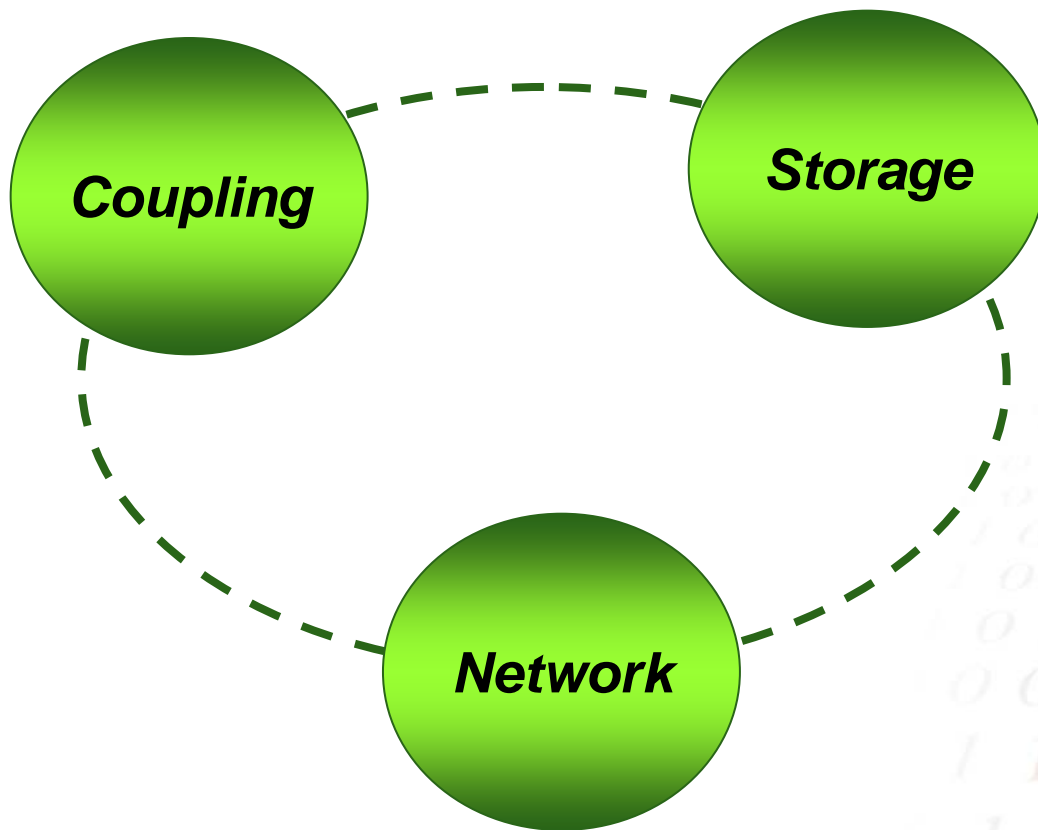
System z
Thanks You!

IBM Systems

z10 EC Connectivity summary

Maximum of 1024 CHPIDs, three I/O cages, 84 I/O slots (28 I/O slots per I/O cage) *

Features	Minimum # of features	Maximum # of features	Maximum connections	Increments per feature	Purchase increments
16-port ESCON	0 [1]	69*	1024 channels	16 channels 1 reserved as a spare	4 channels
FICON Express4	0 [1]	84*	336 channels	4 channels	4 channels
FICON Express2**	0 [1]	84*	336 channels	4 channels	4 channels
FICON Express**	0 [1]	60	120 channels	2 channels	2 channels
ICB-4	0 [1]	8	16 links [2] [3] [4]	2 links	1 link
HCA2-O	0 [1]	16	32 links [3] [4]	2 links	2 links
ISC-3	0 [1]	12	48 links [2]	4 links	1 link
OSA-Express3	0	24	48 ports	2 ports - 10 GbE	2 ports
OSA-Express2	0	24	48 ports	2 or 1 (10 GbE has 1)	2 ports / 1 port
Crypto Express2	0	8	16 PCI-X adapters	2 PCI-X adapters	2 PCI-X adapters [5]

1. Minimum of one I/O feature (ESCON, FICON) or one Coupling Link (ICB-4 , IFB, ISC-3) required.
2. Maximum number of Coupling Links combined (ICB-4s, IFBs, and active ISC-3 links) cannot exceed 64 per server.
3. ICB-4 and HCA2-O (InfiniBand) are not included in the maximum feature count for I/O slots but are included in the CHPID count.
4. Total number of ICB-4 + IFB ports cannot exceed 32.
5. Initial order of Crypto Express2 is 4 PCI-X adapters (two features). Each PCI-X adapter can be configured as a coprocessor or an accelerator.

* Model E12 has 1 book: Supports maximum of three I/O cages and 64 I/O slots.

** Available only when carried forward on an upgrade from z990 or z9 EC.

Publications

- **SA24-7172** **FICON I/O Interface Physical Layer**

- **GA23-0367** **Planning for Fiber Optic Links (Resource Link)**

  http://www.ibm.com/servers/resourcelink
  After logging in select "library" on the left
  Select the server
  Scroll down to find the publication

- **SG24-5444** **System z Connectivity Handbook (*Redbook*)**

- **SG24-6497** **FICON Implementation Guide (*Redbook*)**

On the Internet

- **IBM Redbooks**
  - www.redbooks.ibm.com

- **Network connectivity home page**
  - www.ibm.com/systems/z/networkiing

- **I/O connectivity home page**
  - www.ibm.com/systems/z/connectivity
    - Go to this location for a list of FICON/FCP supported devices

- **Announcement Letters, sales manual**
  - www.ibm.com/common/ssi/OIX.wss

IBM Systems