IBM TRAINING

Session G03

IBM System z9 Business Class: I/O Structure and Connectivity

Harv Emery, IBM Americas System z Advanced Technical Support

IBM
SYSTEM z9 AND zSERIES EXPO
October 9 - 13, 2006

Orlando, FL

RETURN TO INDEX

Trademarks

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

APPN*
CICS*
DB2*
DB2 Connect
DirMaint
e-business logo*
ECKD
Enterprise Storage Server*
ESCON*
FICON*
GDPS*
Geographically Dispersed Parallel Sysplex
HiperSockets
HyperSwap
IBM*
IBM eServer
IBM e(logo)server*
IBM logo*
IMS
Language Environment*
MQSeries*
Multiprise*
NetView*
On demand business logo
OS/390*
Parallel Sysplex*
PR/SM
Processor Resource/Systems Manager
RACF*
Resource Link
RMF
S/390*
Sysplex Timer*
System z9
TotalStorage*
Virtualization Engine
VM/ESA*
VSE/ESA
VTAM*
WebSphere*
z/Architecture
z/OS*
z/VM*
z/VSE
zSeries*

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

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

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

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.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

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.

Agenda

- I/O Introduction and New Function
- I/O Infrastructure
- LCSSs and PCHIDs
- Channels
- Networking
- Parallel Sysplex® Support
- HMC and SE
- Server Time Protocol
- Backup Charts

IBM System z9™
I/O Introduction and New Function

z9-109 I/O Enhancements (2005)

- **I/O Enhancements**
  - Up to 16 x 2.7GB STIs per book* with Redundant I/O Interconnect
  - Up to 80% increase in I/O bandwidth per book
  - Multiple Subchannel Sets (MSS)
    - Second subchannel set definable per LCSS
    - Up to 63.75K subchannels for Set-0
  - Modified Indirect-Data-Address Word (MIDAW) facility
  - Storage Area Network (SAN) enhancements:
    - N_Port ID Virtualization (NPIV) for switched FCP
    - Linux® program directed re-IPL
    - Remote Node ID Reporting for FICON FC
  - FICON® Express2 – 64 Open Exchanges
  - Up to 28 FICON Express2/Express, Crypto Express2, OSA-Express2/Express features per I/O cage
    - Up to 336 FICON Express2 channels per machine (Up to 256 for the Model S08)
- **Networking enhancements**
  - HiperSockets™ IPv6
  - OSA-Express2 1000BASE-T Ethernet
  - OSA-Express2 OSN (OSA for NCP support)
  - GARP VLAN management (GRVP)

*z9-109 exploits only a subset of its designed I/O capability

System z9 I/O Enhancements – May 2006

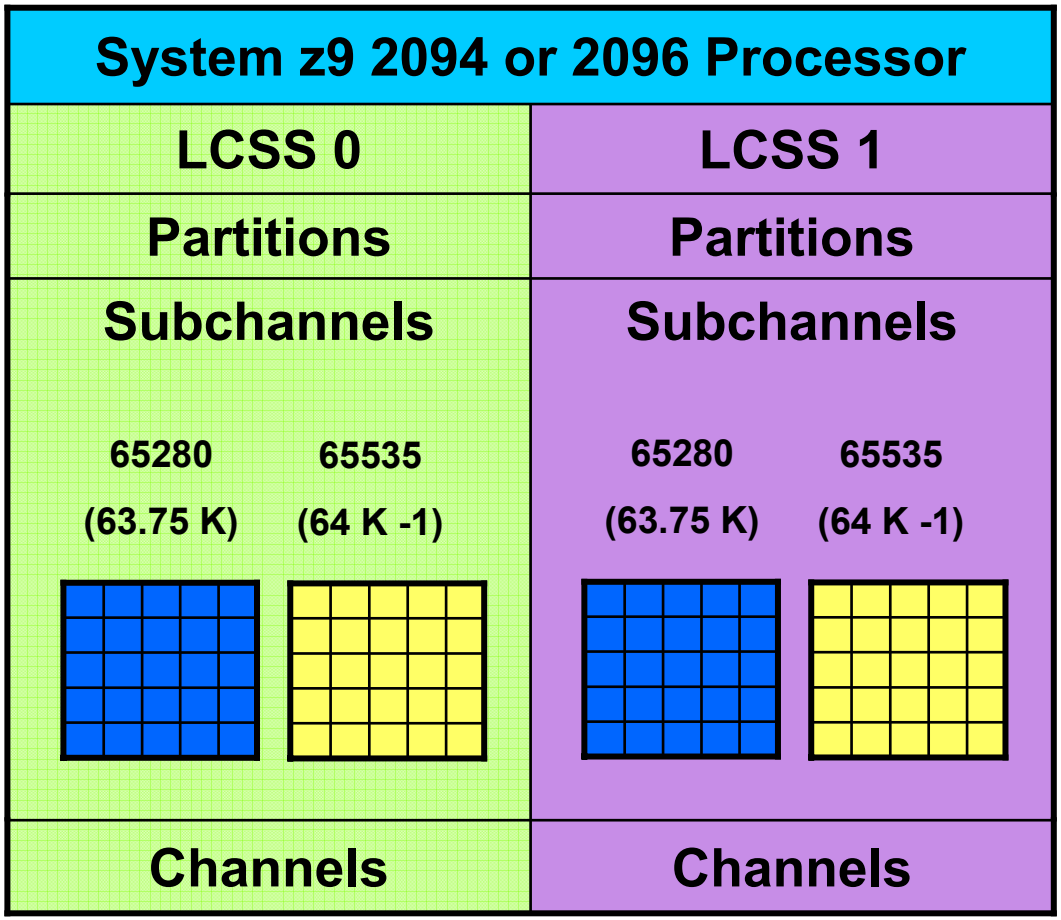
- **FICON Express4**
  - 1, 2, 4 Gbit/sec auto-negotiated
  - Small form factor pluggable optics
  - 2 GB/sec bandwidth between card and memory
- **OSA-Express Enhancement**
  - Layer2/Layer3 transport mode port sharing
- **OSA Express2 1000Base-T Ethernet**
  - Announced for z990 and z890

Note: OSN function not included on z990 or z890

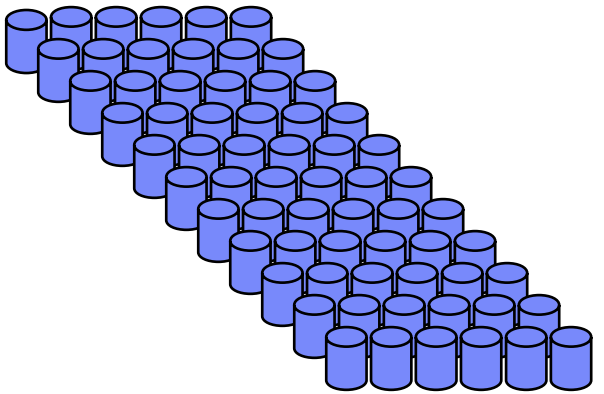
z9 BC I/O Overview

- **I/O Enhancements compared to z890**
  - Up to 28 - FICON Express; FICON Express2; FICON Express4
    - 4 channels/feature FICON/FCP
    - 1, 2, 4 Gbps auto-negotiated 4Gbps for FICON Express4
  - Modified Indirect Data Address Word (MIDAW) facility
  - Multiple (2) Subchannel sets (MSS)
    - Increase to 63.75K Subchannels for Set-0
  - Up to 16 x 2.7GB STI's
    (7 STIs max for the single I/O cage. Remaining STIs for ICB-4s)
- **Storage Area Networks (SANs) enhancements**
  - N_Port ID Virtualization
  - Program Directed re-IPL
  - FICON Link Incident Reporting
- **Networking enhancements**
  - HiperSockets IPv6
  - OSA-Express2 1000BASE-T Ethernet
  - OSA-Express2 OSN (OSA for NCP support)
  - GARP VLAN management (GRVP)

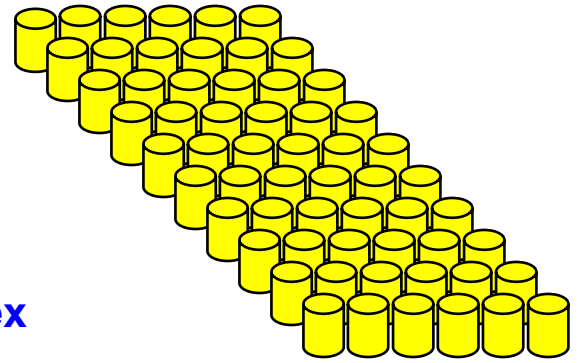
System z9 MSS Capabilities



Subchannel per device



Base Devices



Alias Devices

- **More base volumes, more data in an LPAR or sysplex**
- **More alias volumes, better performance**
- **Synergy with z/OS® large volume capability**

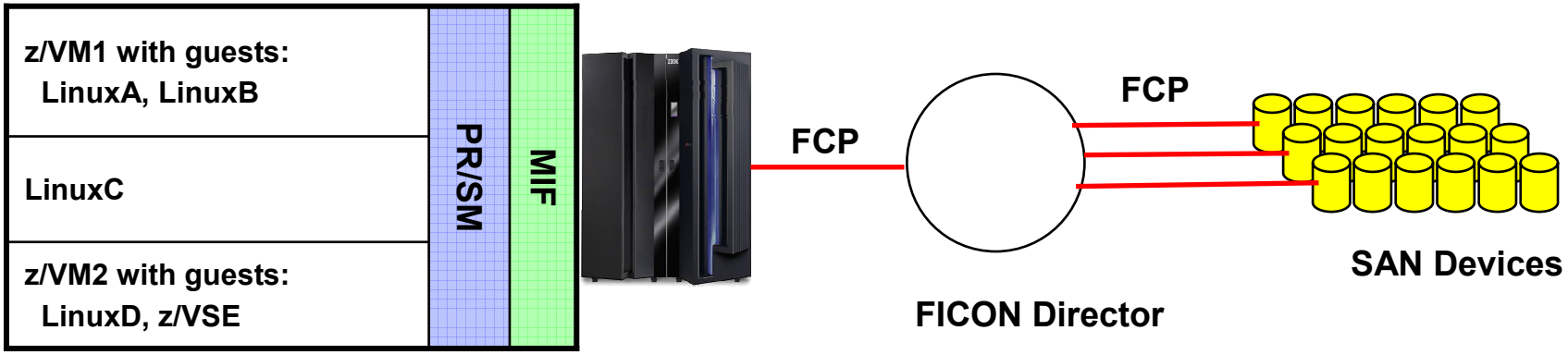
System z9 Modified Indirect-Data-Address Words (MIDAWs)

- **Modified-Indirect-Data-Address Word facility:**
  - The MIDAW facility is a new CCW-indirect-data-address word facility being added to z/Architecture™ to coexist with the current IDAW facility
  - Both MIDAW and IDAW facilities offer, for FICON and ESCON® channels, alternatives to using CCW data chaining in channel programs
  - Both facilities are designed to reduce channel, director, and control unit overhead by reducing the number of CCWs and frames processed
  - The MIDAW facility is usable in certain case where the IDAW facility is not because it does not have IDAW boundary and data length restrictions
- **Designed to:**
  - Be compatible with existing disk control units
  - Decrease response time for exploiting I/O.
  - Increase the number of I/O operations per second that can be processed and thus move more data per second, especially on faster FICON channels
- **Applications that may benefit include:**
  - DB2®, VSAM, Partitioned Data Set Extended (PDSE), Hierarchical File System (HFS), z/OS File System (ZFS), and other datasets exploiting striping and compression

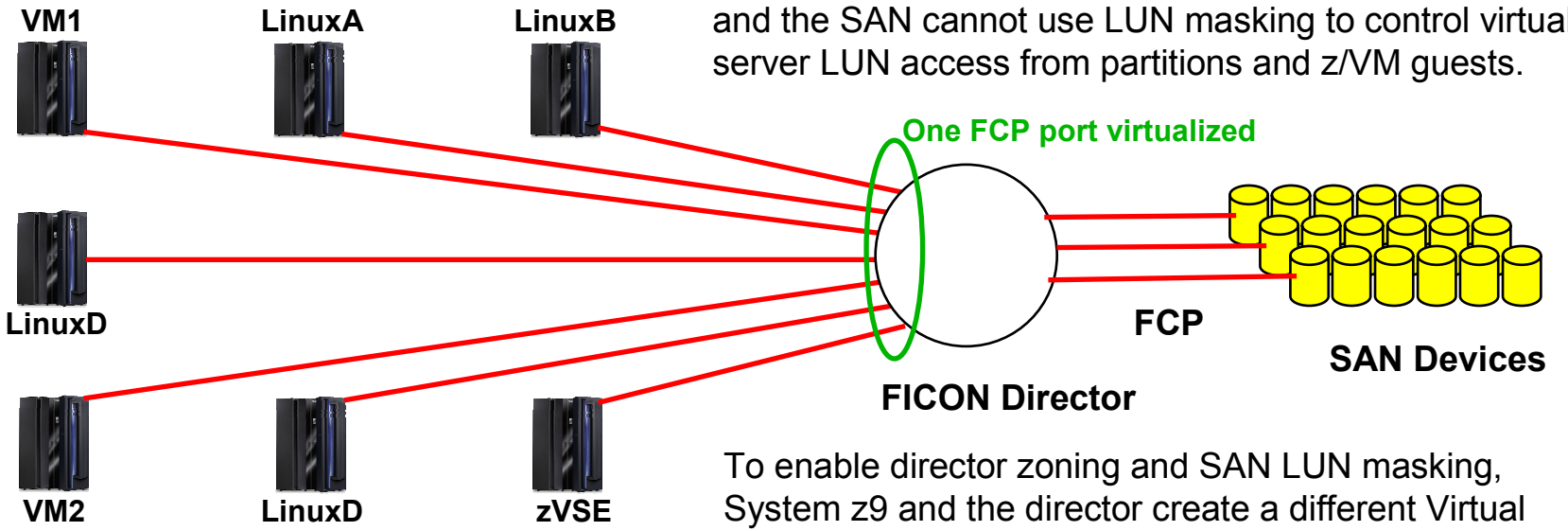
System z9 N_Port ID Virtualization (NPIV)

- NPIV provides the capability for an FCP channel to define itself as multiple virtual FCP channels, each with its own unique virtual Fibre Channel port name and FC N_Port ID.
- By assigning different virtual N_Port IDs to different operating system images (running in logical partitions or as z/VM guests in virtual machines), operating systems can use the virtual FCP channels as if they were using dedicated physical FCP Channels
- The I/O transactions of each operating system image are separately identified, managed, and transmitted, and are processed by the fibre channel and SAN infrastructure as if each image had its own unique physical N_Port. Each virtual FCP Channel port identifier (WWPN) can be used in device-level access control (LUN masking) in storage controllers and in switch-level zoning.
- **Designed to permit multiple operating system images concurrently to access SAN attached devices (LUNs) via a single, shared FCP Channel.**
  - Designed to work with all FICON features supported on System z9
  - May reduce number of FCP channels needed
  - May simplify infrastructure by reducing ports and cabling
  - May improve channel utilization
- **NPIV is an industry standard solution** which supersedes FCP LUN Access Control
  - **Requires Fibre Channel Director support in the FCP channel's entry director only**
  - **IBM is working with Linux on System z9 distribution partners to provide NPIV support**

NPIV Physical and Virtual Views



With only one N_Port ID, the director cannot use zoning and the SAN cannot use LUN masking to control virtual server LUN access from partitions and z/VM guests.

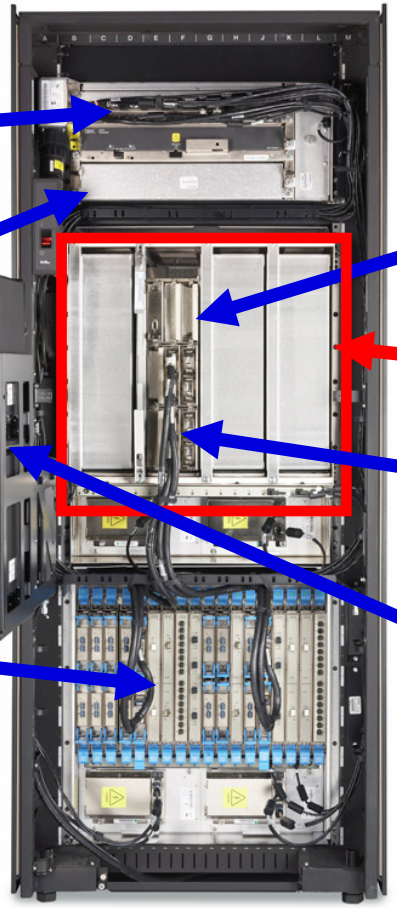


To enable director zoning and SAN LUN masking, System z9 and the director create a different Virtual N_Port ID for each virtual server.

System z9 I/O Infrastructure

System z9 BC – Under the covers

A Frame



Internal Batteries (optional)

Power Supplies

I/O Cage

Book 0: Processors, Memory and Channel Busses (STIs)

CEC Cage

STI Cables

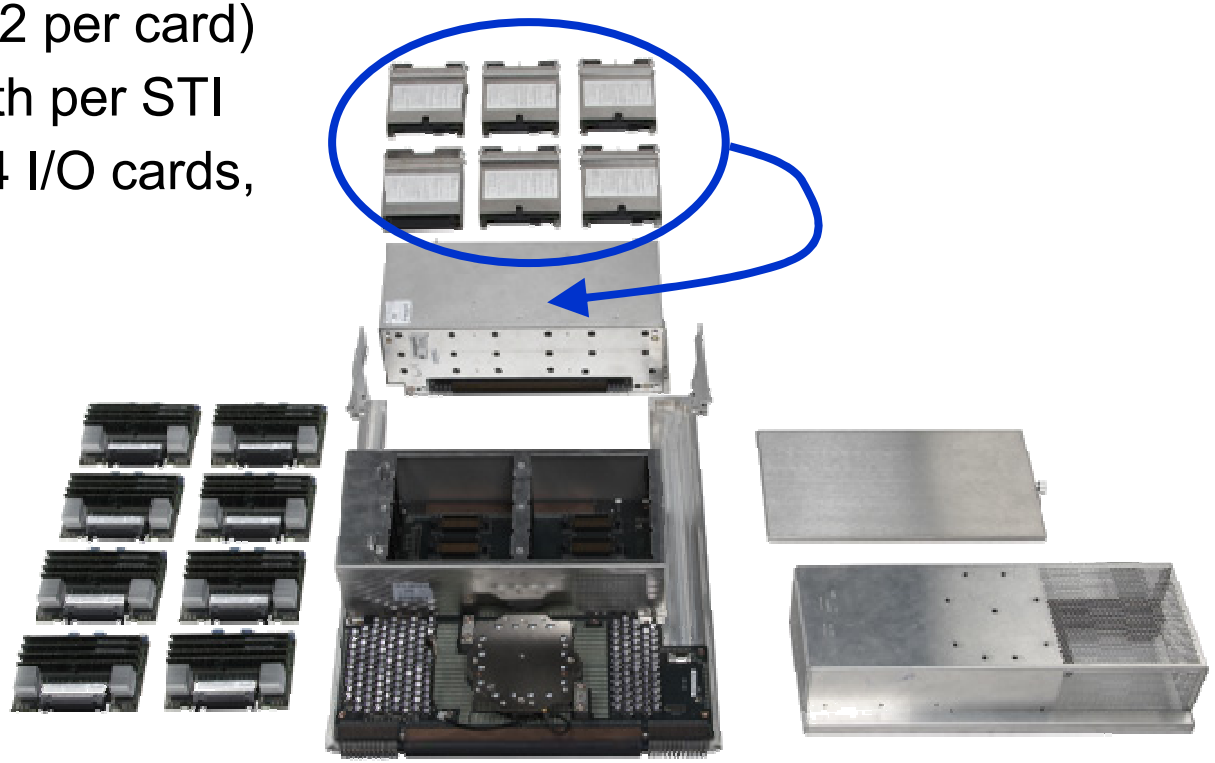
Support Elements (gate with Laptops swung open)

Fiber Quick Connect Feature (optional)

Front View

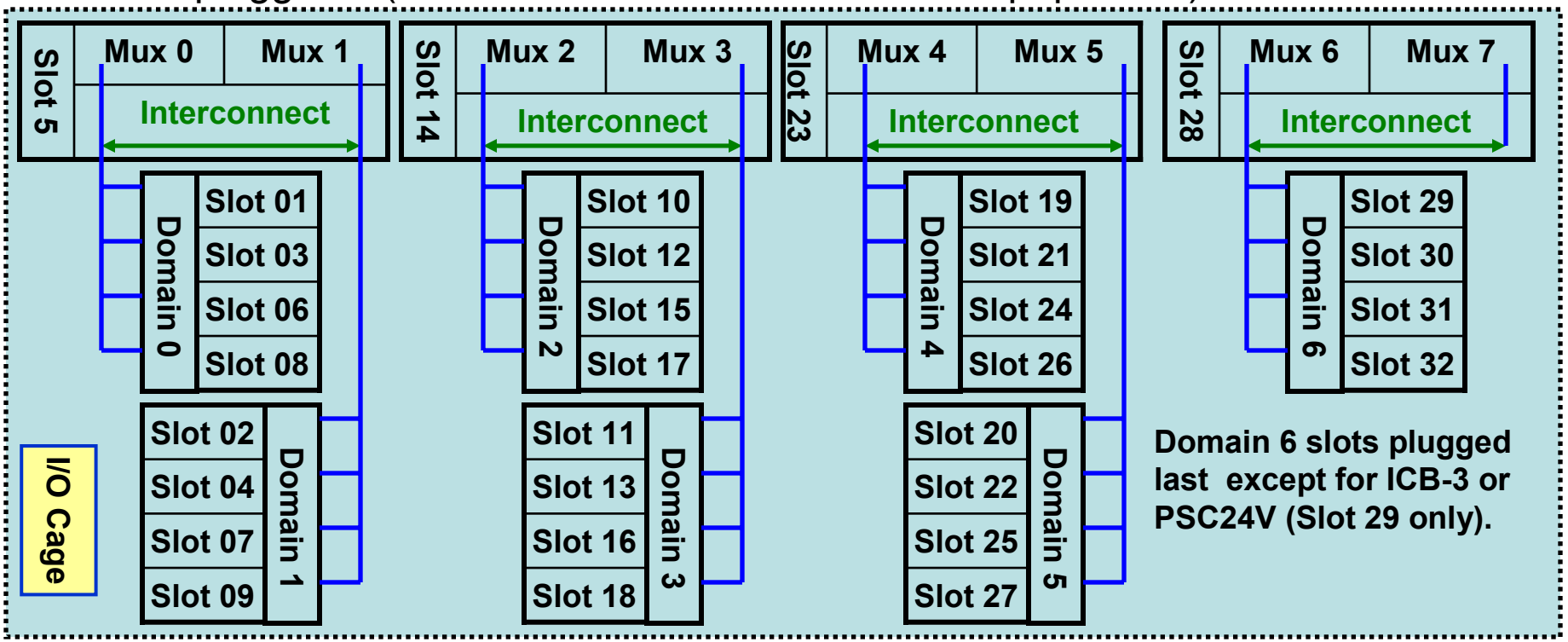
System z9 BC MBA Cards (Channel Bus Cards)

- **System supports 2, 4, 6, or 8 MBA cards, pairs for availability**
  - Hot Plug
  - Hot Repair
- **Each card supports two STIs (channel buses)**
  - 4, 8, 12, or 16 STIs (2 per card)
  - 2.7 GB/sec bandwidth per STI
  - Each STI supports: 4 I/O cards, one ICB4 or 2 ICB3

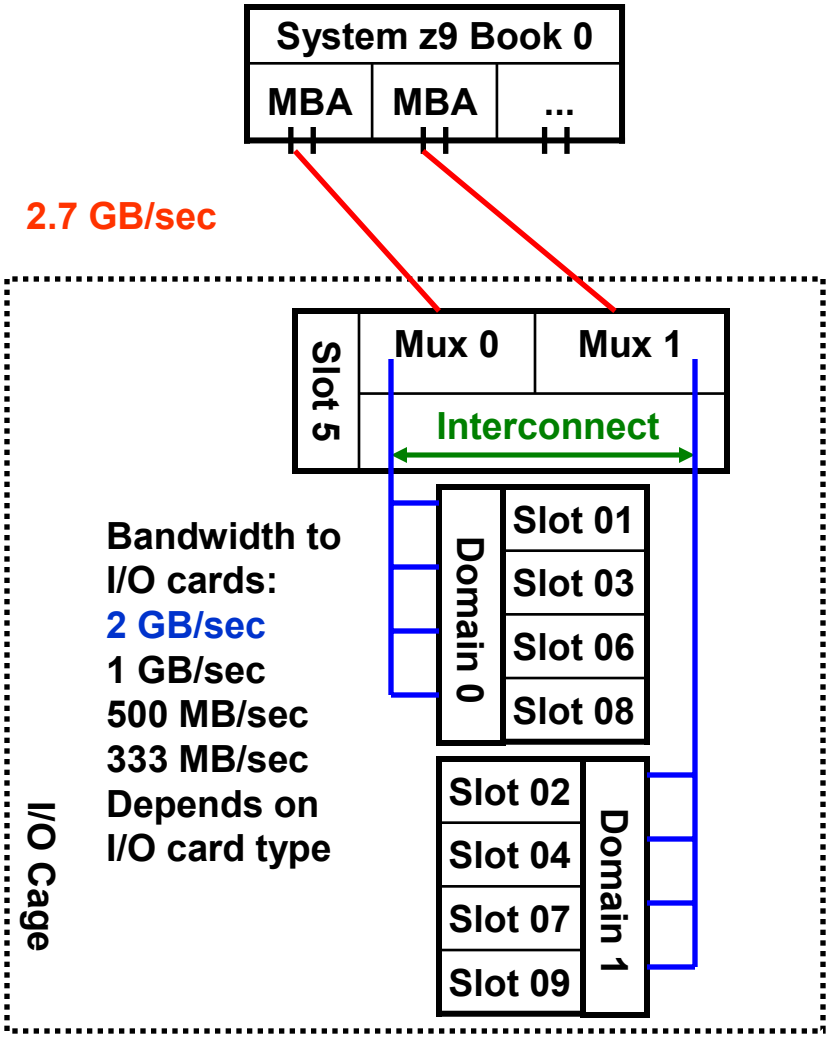


System z9 BC I/O cage

- **Same cage technology as System z9 EC**
- **Increased power compared to z890 –** Supports any combination of I/O cards
- **New STI multiplexers with Redundant I/O Interconnect**
  - Up to four pairs supporting Domains 0/1, 2/3, 4/5, and 6
  - Up to 2 GB/sec bandwidth per I/O card in each domain
  - Hot pluggable (Airflow cards used if domains not populated)



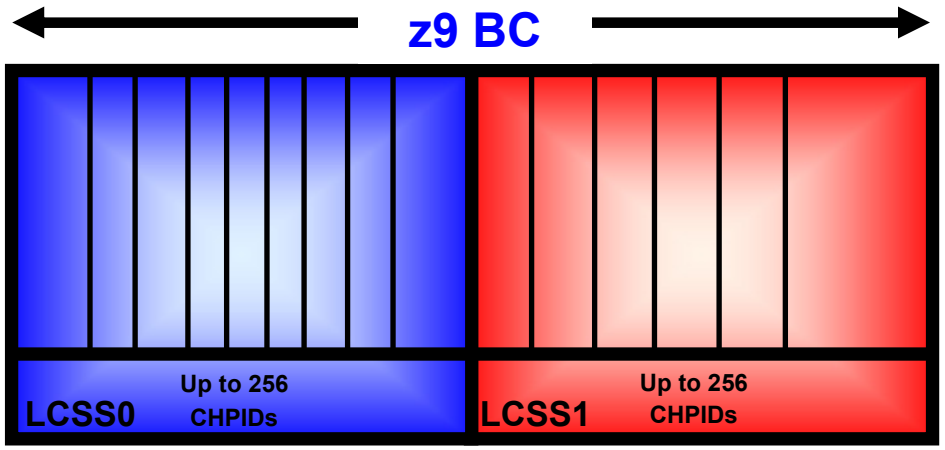
System z9 Redundant I/O Interconnect – One book



- **Normal Operation**
  - Mux 0 STI supports only Domain 0
  - Mux 1 STI supports only Domain 1
  - STIs from different MBAs
- **Redundant operation – one book**
  - Using the interconnect Mux 0 STI **OR** Mux 1 STI supports **both** domains
  - Supports System z9:
    - **Concurrent MBA repair**
    - **Concurrent STI cable repair**

System z9 LCSSs and PCHIDs

z9 BC Logical Channel Subsystems (LCSSs)



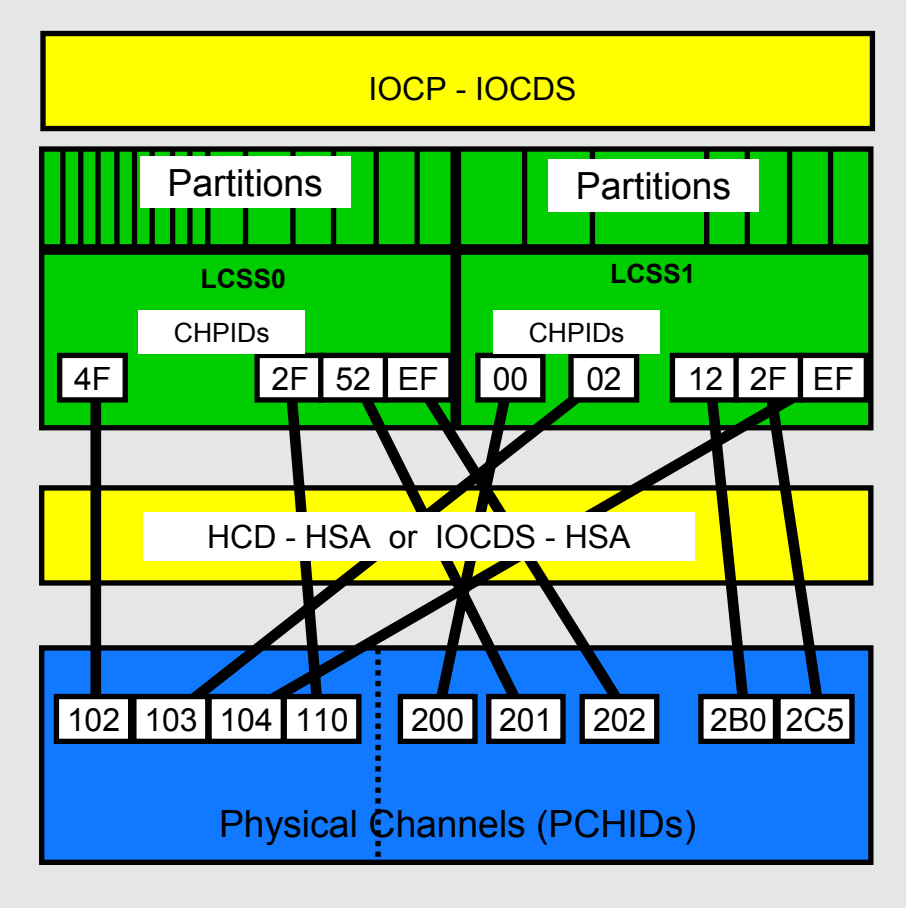
- **One or two Logical Channel Subsystems (LCSSs) on z9 BC, up to four on z9 EC**
  - Up to 15 LPARs per LCSS (**15 total on z9 BC Model R07**)
  - Up to 256 channels per LCSS
- Multiple LCSSs enable
  - **Up to 60 Logical Partitions per z9 EC (Requires four LCSSs)**
  - **Up to 30 Logical Partitions per z9 BC S07 (Requires two LCSSs)**
  - **Up to 420 external data channels on z9 BC Model S07, 240 on R07, 1024 on z9 EC**
- An LPAR can access channels ONLY in its assigned LCSS
- Some channels may be assigned to multiple LCSSs - "**Spanned Channels**"
  - ICP, IQD, FC, FCP, OSE, OSD, OSC, OSN, CBP, CFP
  - But not ESCON or FICON Conversion

System z9 compared to z990, z890, z900, z800

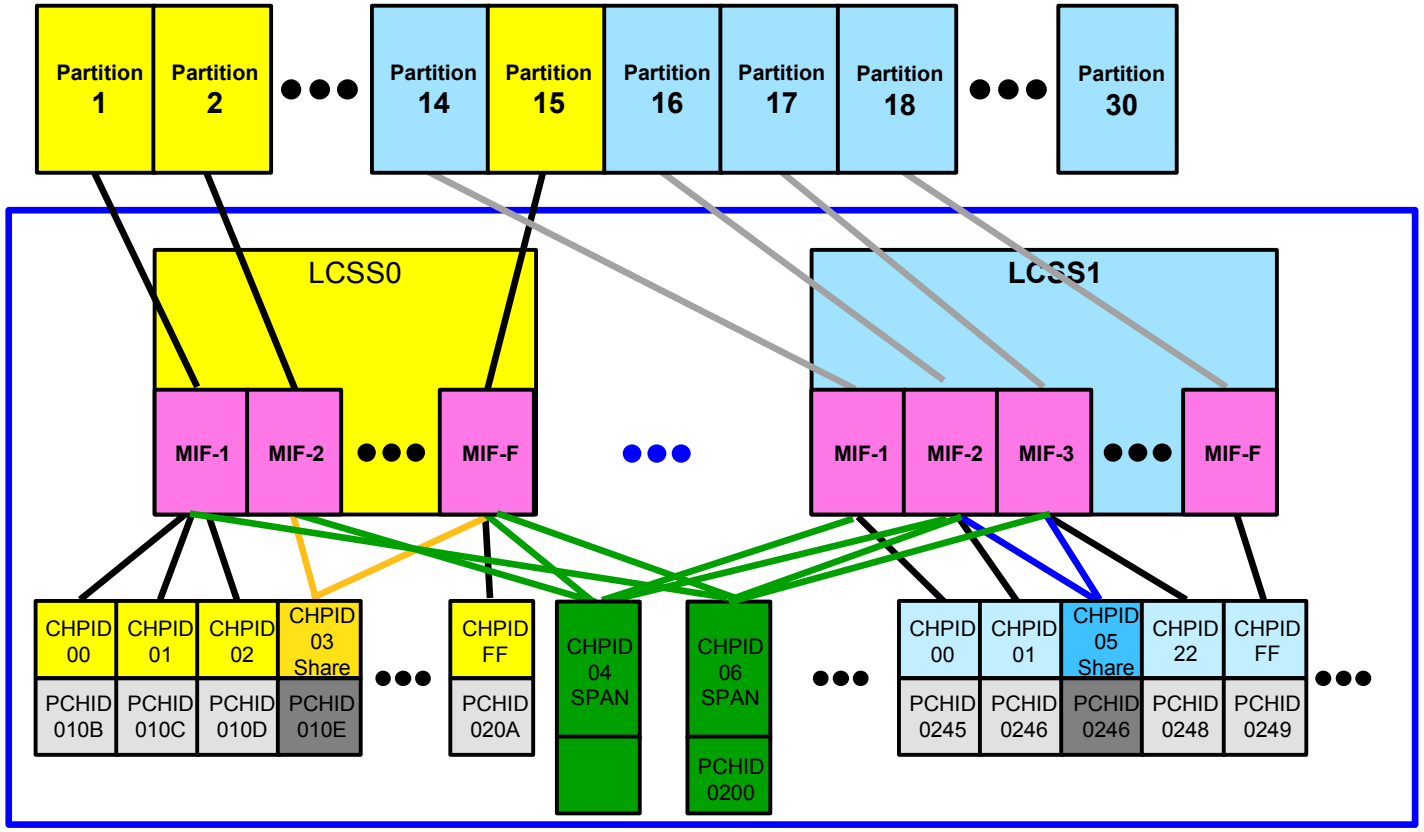
zSeries Processor Type	MCSS (LCSSs)	Logical Partitions	Subchan Sets	Subchannels Per LCSS (Per LPAR)	Base Devices (per CSS)	SPANNED Channels
2064 z900	No (1 CSS)	15	1 per LCSS	63 K	64512	N/A
2066 z800	No (1 CSS)	15	1 per LCSS	63 K	64512	N/A
2084 z990	Yes 4 x LCSS	30 (15 / LCSS)	1 per LCSS	63 K	64512	Yes
2086 z890	Yes 2 x LCSS	30 15 - 110 (15 / LCSS)	1 per LCSS	63 K	64512	Yes
2094 z9 EC	**Yes 4 x LCSS**	**60 (15 / LCSS)**	**1 or 2 per LCSS**	**65280 + 65535**	**65280 (Set 0)**	**Yes**
2096 z9 BC	**Yes 2 x LCSS**	**30 – S07 15 – R07 (15 / LCSS)**	**1 or 2 per LCSS**	**65280 + 65535**	**65280 (Set 0)**	**Yes**

System z9 Physical Channel IDs (PCHIDs)

- Channel ports have no CHPID numbers pre-assigned as on z900/800 and earlier machines
- CHPID numbers must be assigned to channel ports (PCHIDs) as part of the definition process
  - HCD/IOCP manual process, and/or
  - Updated CHPID Mapping Tool
- CHPID assignment
  - Define the channel to an LCSS(s)
  - Associate the CHPID number to a physical channel port location (PCHID)
  - CHPID numbers are still 00 - FF and must be unique within an LCSS
- Physical channel location, known as the PCHID, is assigned by manufacturing and reported by eConfig in the PCHID report
- Except for ESCON sparing, a PCHID relates directly to a jack location on a channel card in a specific I/O slot, in a specific I/O cage
  - Other exceptions
    - ICs and HiperSockets - no PCHID
    - ICB-4 - assigned to CEC cage



z9 BC Two LCSSs and External Spanned Channels



- CHPID 04 Spanned Internal HiperSockets (IQD) or Internal Coupling Link (ICP)
- CHPID 06 Spanned external channel (FICON, OSA, or External Coupling Link) Supported: FC, FCP, OSE, OSD, OSC, OSN, CBP, CFP
  - **Not supported: ESCON, FICON Conversion (FCV)**

System z9 Channels

System z9 Channel Type and Crypto Overview

- **FICON/FCP**
  - **FICON Express4**
  - FICON Express
  - FICON Express2
- **Networking**
  - OSA-Express2
    - Gigabit Ethernet LX and SX
    - 10 Gigabit Ethernet LR
    - 1000BASE-T Ethernet (#3366)
  - OSA-Express
    - Gigabit Ethernet LX and SX
    - 1000BASE-T Ethernet (#1366)
    - Fast Ethernet
  - HiperSockets
- **Coupling Links**
  - ISC-3 (Peer mode only)
  - ICB-3, ICB-4
  - IC
- **ESCON**

- **Crypto**
  - Crypto Express2
    - Configurable Coprocessor or Accelerator

- **Channel types not supported:**
  - OSA-Express Token-Ring (SOD Oct 2004)
  - FICON (pre-FICON Express)
  - Internal Disk / Emulated I/O
  - ICB-2 (SOD 2003)
  - ISC-3 Links in Compatibility Mode (SOD April 2004)
  - Parallel (use ESCON Converter)
  - OSA-Express ATM 155
  - OSA-2
  - PCIXCC
  - PCICA

Note: Only ICB cables orderable.
All other cables have to be sourced separately.
Blue Items Only Available on MES Upgrades

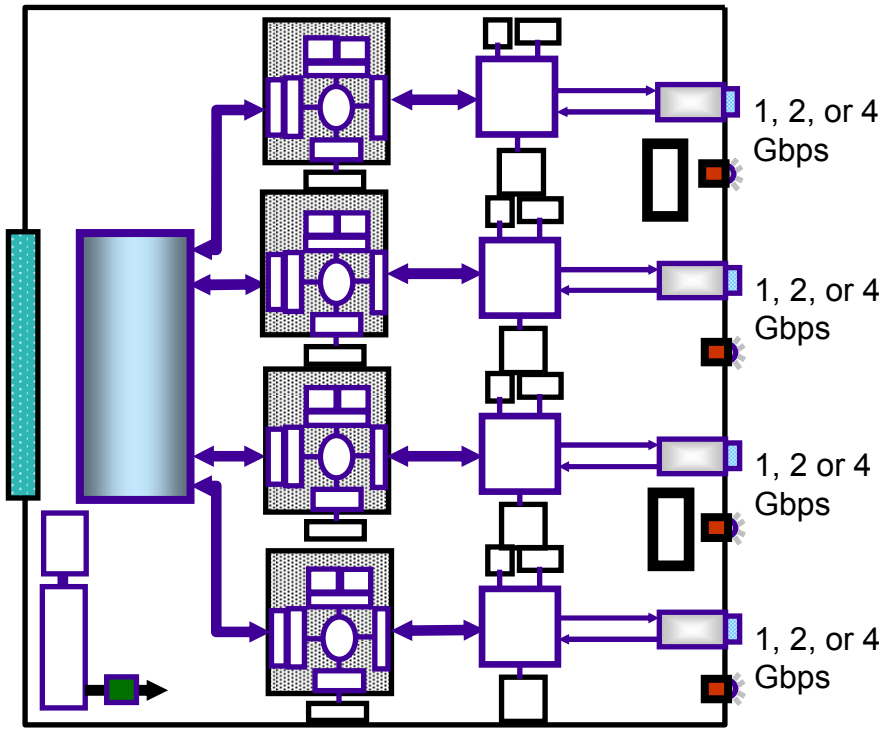
System z9 BC I/O connectivity overview

- **FICON Express4, FICON Express2, FICON Express**
  - Up to 28 features / 112 channels on S07 (FICON Express4 + FICON Express2)
  - Up to 16 features / 64 channels on R07 (FICON Express4 + FICON Express2)
  - FICON Express required for FCV support
- **ESCON Channels, 16-port cards: up to 15 channels per card with spare**
  - Up to 420 channels, 28 cards (S07) or 240 channels, 16 cards (R07)
- **OSA-Express2, OSA-Express**
  - Up to 24 cards, 48 ports S07, 16 cards, 32 ports R07 except 12 cards, 24 ports for A01
  - Gigabit Ethernet, 10 Gigabit Ethernet, 1000BASE-T Ethernet, Fast Ethernet
- **HiperSockets internal LANs, up to 16**
  - Nothing to order created by I/O definition
- **Coupling Links, up to 64 total**
  - IC (up to 32), ICB-3 (up to 16), ICB-4 (up to 16, 12 on R07 A01), ISC-3 (up to 48 links)
- **Crypto Express2**
  - Up to 8 features, 16 coprocessors (If ordered, minimum order is 2 features)
  - Default Secure Coprocessor or configurable as “Public key” Accelerator

System z9 BC FICON Express4

- **Supports all the function of the FICON Express2 feature plus:**
  - 4 Gbps with Auto-negotiate capability (1, 2, or 4 Gbps)
  - Can be shared among LPARs, and defined as spanned
  - Small Form Factor Pluggable (SFP) optics for Service / Repair
    - Concurrent repair/replace action for each SFP
- **Ordering**
  - Two or Four port increments
  - Intermix is not supported on a single card
  - All ports must be of the same type, either LX or SX.
    - LX Feature Code 332 has 4 LX 10 km SFPs
    - SX Feature Code 3322 has 4 SX SFPs
    - LX Feature Code 3323 has 2 LX 4 km SFPs
    - LX Feature Code 3324 has 4 LX 4 km SFPs

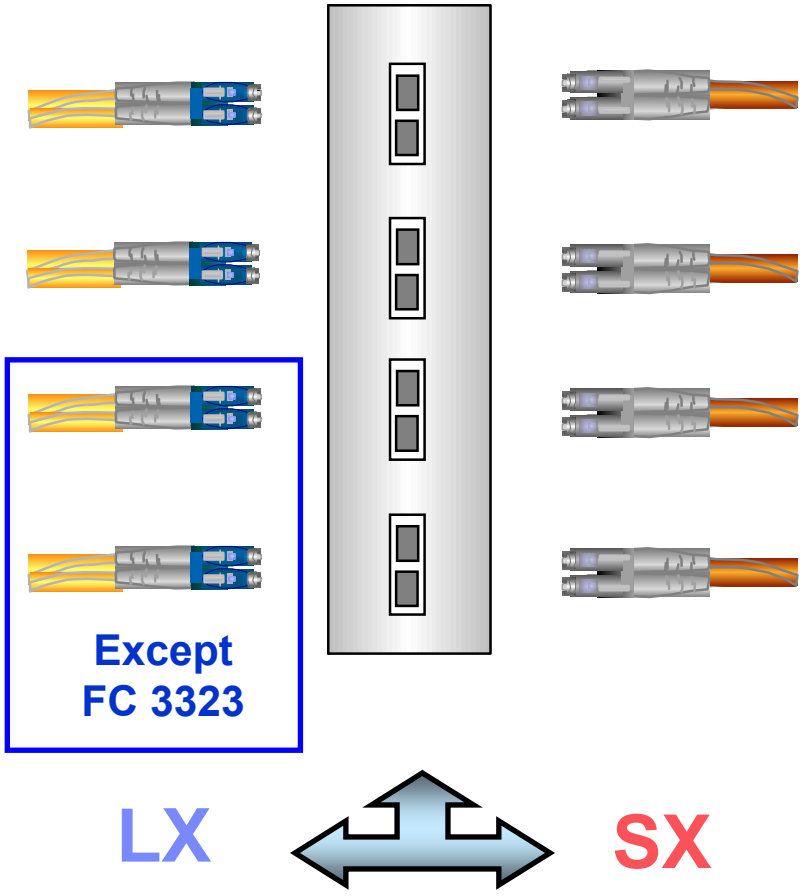
FICON Exprees4 4 Port card shown



FC 3321 FICON Express4 10 km LX
FC 3322 FICON Express4 SX
FC 3323 FICON Express4-2C 4km LX (2 ports)
Note: FC 3323 is NOT supported on z9 EC
FC 3324 FICON Express4 4km LX

System z9 BC FICON Express4

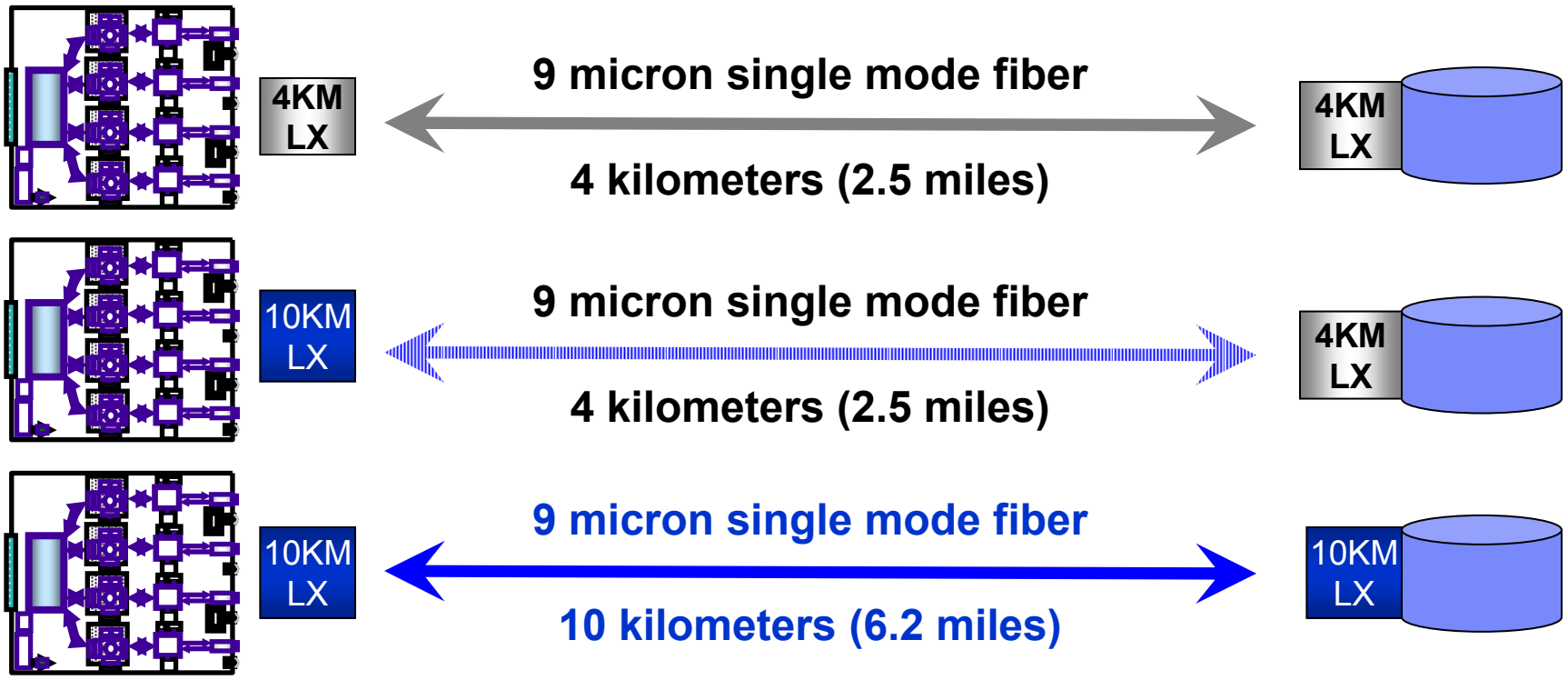
- **Three LX features**
  - 9 micron single mode fiber
  - Feature unrepeated distance
    - Up to 4 kilometers (2.5 miles) OR
    - Up to 10 km (6.2 miles)
  - Receiving device must also be LX
- **One SX feature**
  - 50 or 62.5 micron multimode fiber
  - Unrepeated distance varies with speed and fiber type
  - Receiving device must also be SX
- **Connectors - LC Duplex**
- **4 channels of LX or SX or 2 of LX**
- **Exploits up to 2 GB/sec bandwidth to STI**
- **Two CHPID types can be configured**
  - FC - native FICON and CTC
  - FCP - communication with SCSI devices
    - half-track block sizes



System z9 FICON Express4 Unrepeated Distances

IBM System z9 4Gb FICON/FCP connectivity:

- The ANSI Fibre Channel Physical Interface (FC-PI-2) standard 10 kilometer (km) and 4 kilometer transceivers when using 9 micron single mode fiber optic cabling
- IBM supports FC-PI-2 variants
- IBM supports interoperability of 10 km transceivers with 4 km transceivers provided the unrepeated distance between a 10 km transceiver and a 4 km transceiver does not exceed 4 kilometers (2.5 miles)



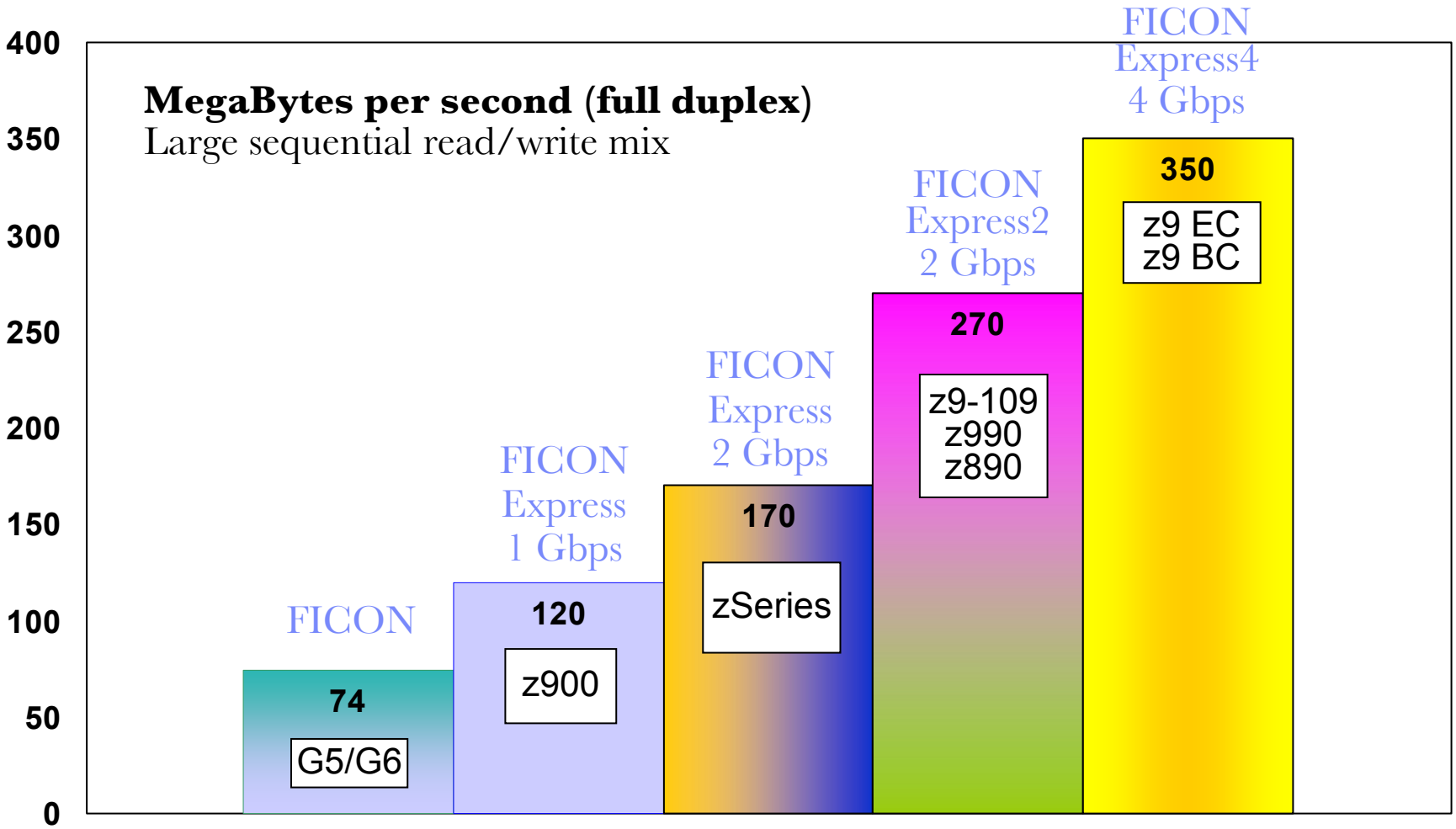
FICON Express4 – Designed to improve performance

- **FICON channel performance[1]**
  - Up to a 25% improvement in FICON channel throughput when processing a mix of read and write large data transfers
  - Up to a 65% improvement in FICON channel throughput when processing all read and or all write large data transfers
- **Application performance – DB2 with the System z9 MIDAW facility, FICON Express4 and the DS8000[1,2]**
  - Internal IBM DB2 Table Scan tests[2] with EF datasets with **FICON Express4 on the z9 EC with the MIDAW facility compared to FICON Express2 not using the MIDAW facility on z9 EC**
    - Greater than 220% improvement in I/O throughput (from 84 to 270 READ MB/sec)
- **FCP channel performance for z/VM and Linux environments[3]**
  - Up to 50% improvement in FCP channel throughput when processing a mix of read and write large data transfers (Up to 525 MB/sec)
  - Up to 100% improvement in in FCP channel throughput when processing all read or all write large data transfers (Up to 392 MB/sec)

1. Large sequential data transfers on z9 EC with FICON Express4 operating at 4 Gb/sec (running z/OS V1.7) when compared to FICON Express2 on z9-109 (running z/OS V1.6).
2. Results of internal DB2 table scan tests with the z9 EC, the MIDAW facility, FICON Express4 operating at 4 Gb/sec (running z/OS V1.7), and the DS8000 compared to z9-109, and FICON Express2 operating at 2 Gb/sec (running z/OS V1.6).
3. Large data transfers on z9 EC with FICON Express 4 operating at 4 Gb/sec (running SUSE Linux SLES 9 SP3) when compared to FICON Express2 operating at 2 Gb/sec on z9-109.

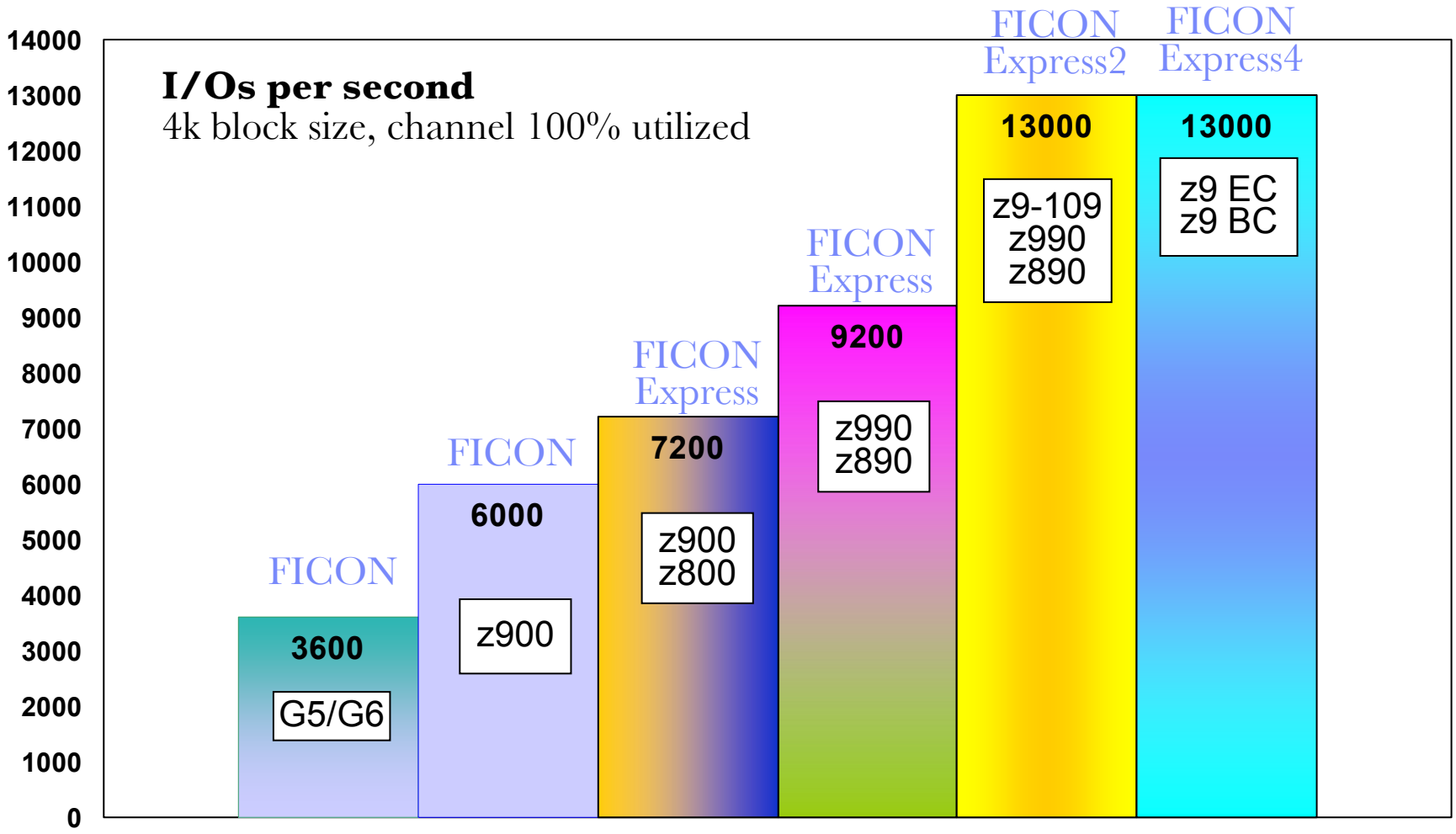
Results were obtained in a controlled environment. No assurance can be given that an individual user will achieve equivalent throughput or performance.

FICON Comparison: Full duplex data transfers



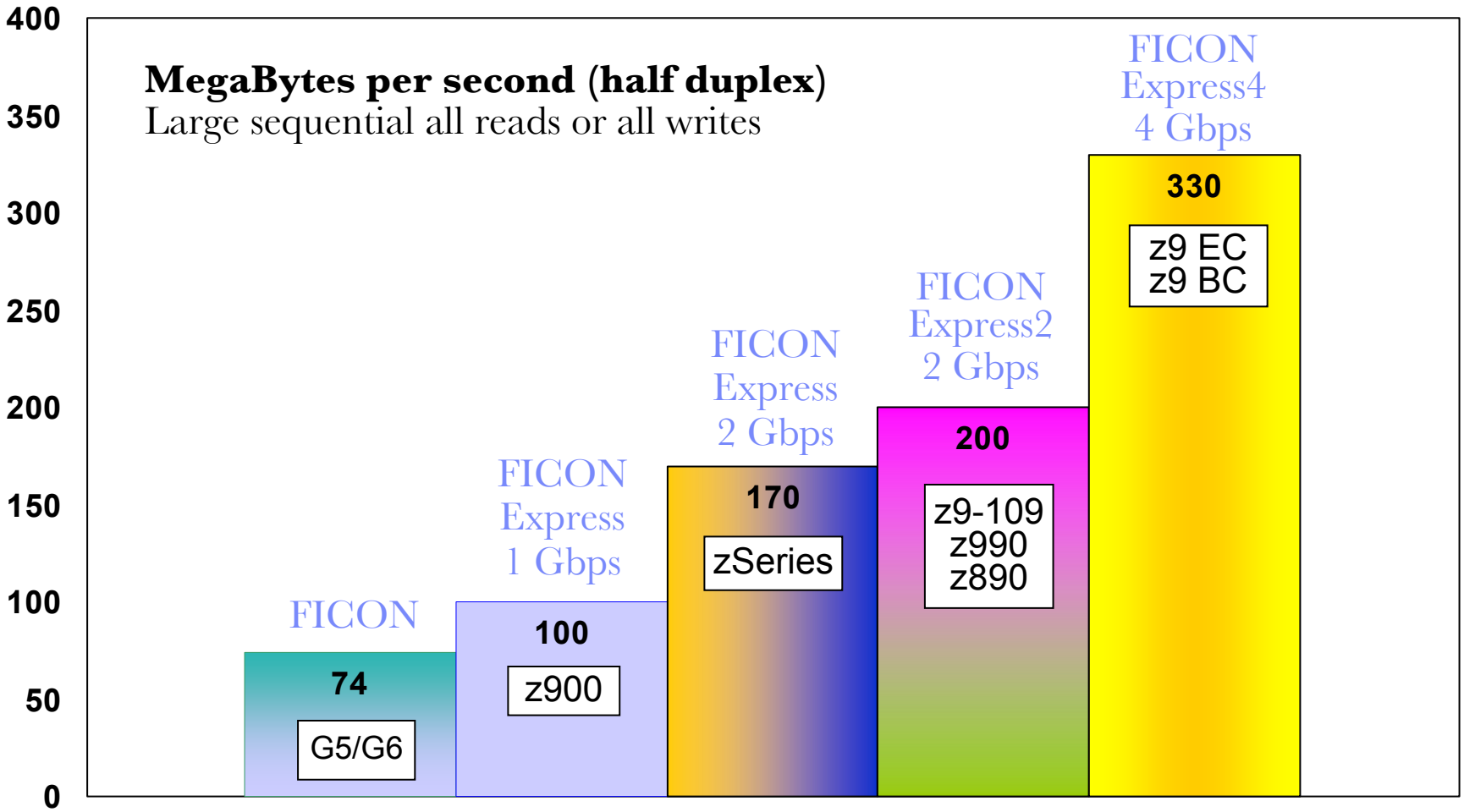
*This performance data was measured in a controlled environment running an I/O driver program under z/OS 1.7. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed.

FICON Comparison: Start I/Os per Second



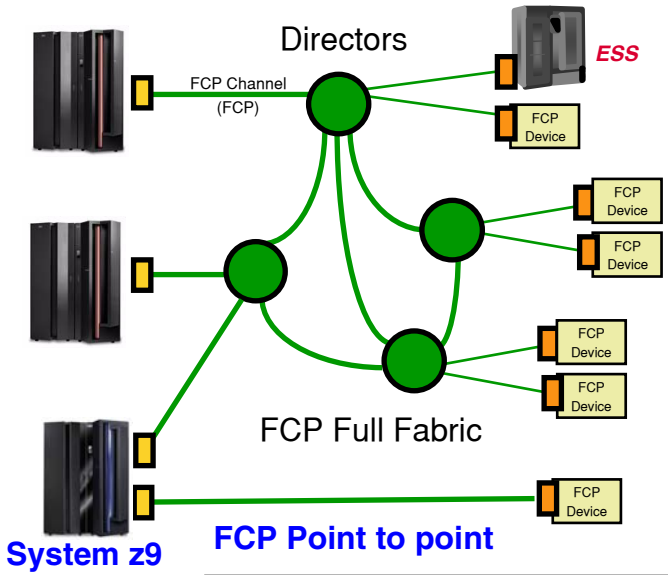
*This performance data was measured in a controlled environment running an I/O driver program under z/OS 1.7. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed.

FICON Comparison – Half duplex data transfers

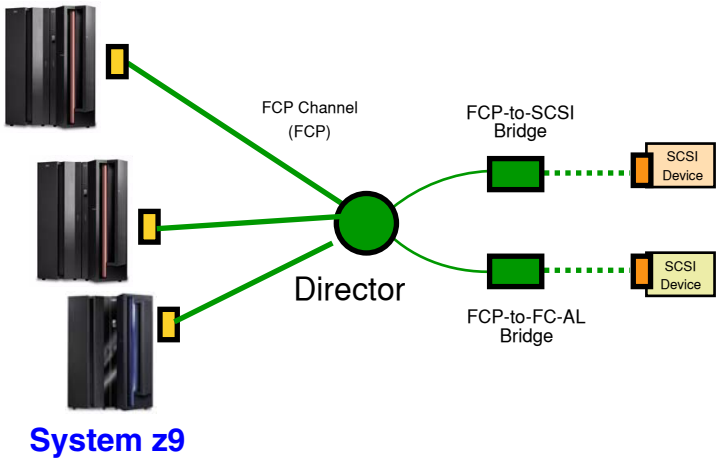


*This performance data was measured in a controlled environment running an I/O driver program under z/OS 1.7. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed.

System z9 FCP Attachment Options for z/VM and Linux



- FCP Full Fabric Connectivity
  - Homogeneous, single vendor fabric
  - Fibre channel directors, switches
  - **NPIV support**
  - **64 Open Exchanges**
- **FCP point to point – Designed to support all FICON features supported on System z9**
  - **Direct attachment to FCP CU port**
  - NPIV support not applicable



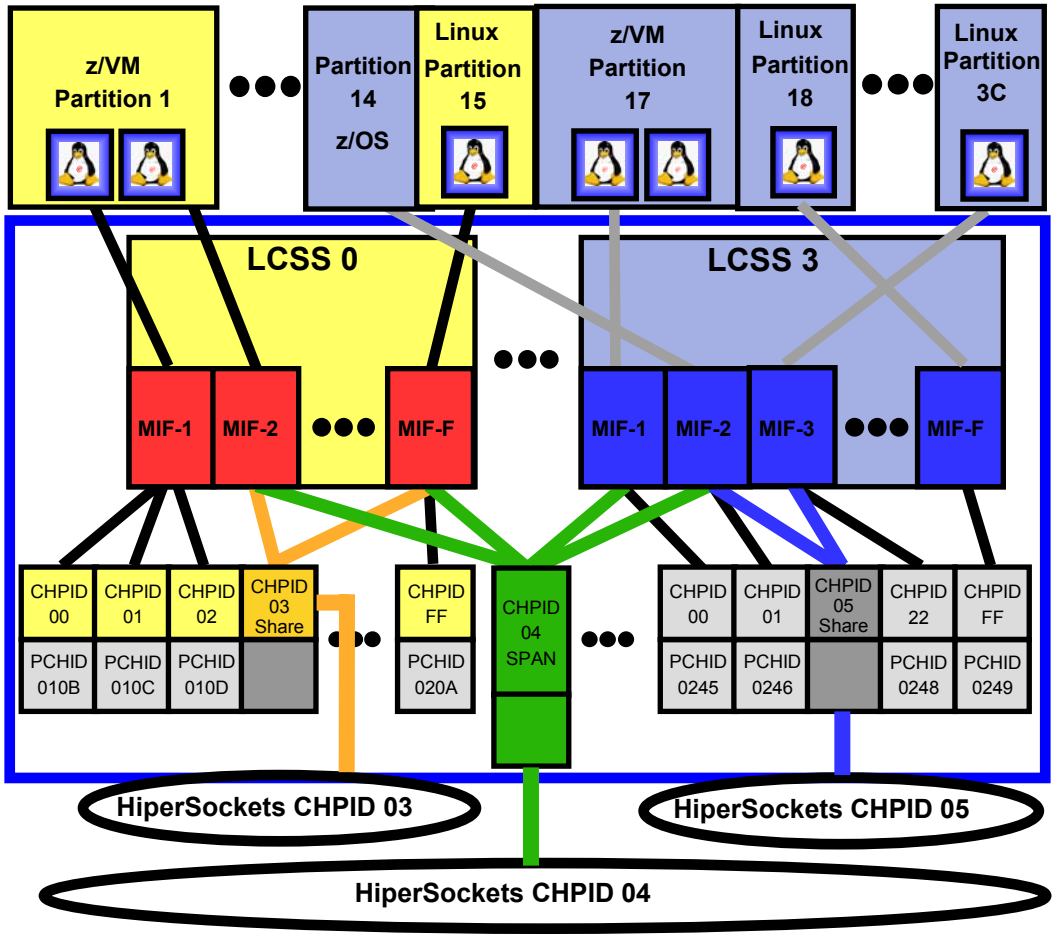
- FCP switched to SCSI Bridge
  - FCP-to-SCSI Bridges
  - FCP-to-FC-AL bridge

Supported devices: ***www.ibm.com/servers/eserver/zseries/connectivity/#fcp***

System z9 Networking

HiperSockets IPv6

- **Internet Protocol Version 6 (IPv6)**
- **More unique IP addresses**
  - Expands the IP address space
    - From 32 bits to 128 bits
  - Follow-on to IPv4
- **Operating system support:**
  - z/OS V1.7
  - z/VM V5.2: Statement of Direction – System and guest support

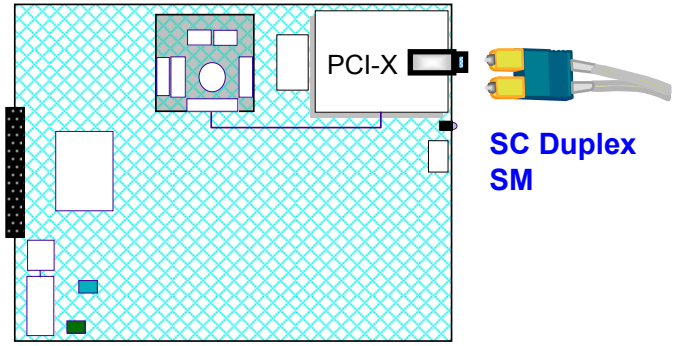


Very High Speed Interconnection between programs running z/OS, z/VM or Linux on System z9

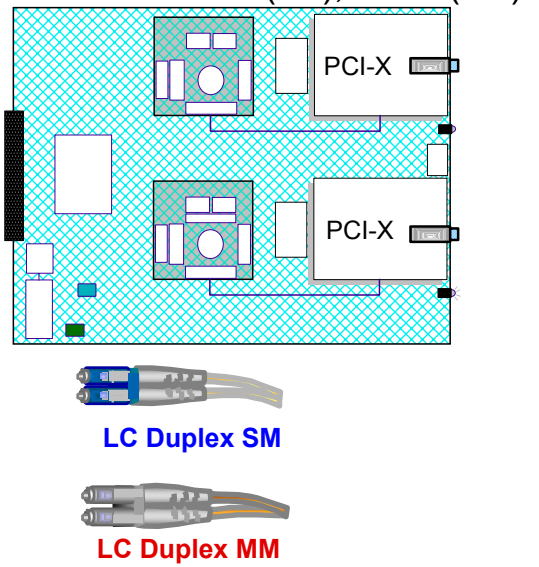
System z9 OSA-Express2 GbE and 10 GbE

- **10 Gigabit Ethernet LR (long reach)**
  - One port per feature
  - CHPID type OSD (QDIO)
  - 9 micron single mode fiber, **SC Duplex connector**
- **Gigabit Ethernet features, 2 ports per feature**
  - **CHPID types OSD (QDIO), OSN (OSA for NCP)**
  - **Designed to achieve line speed - 1 Gbps in each direction**
  - Gigabit Ethernet LX (Long wavelength)
    - 9 micron single mode fiber, LC Duplex connector
  - Gigabit Ethernet SX (Short wavelength)
    - 50 or 62.5 micron multimode fiber, LC Duplex connector
- **OSA-Express2 GbE and 10 GbE support**
  - **Large send -** offloading TCP segmentation
  - **Concurrent LIC update** to minimize network traffic disruption
  - **640 TCP/IP stacks -** improved virtualization
  - **Layer 2 support -** protocol-independent packet switching

10 Gigabit Ethernet
Feature 3368



Gigabit Ethernet
Features 3364 (LX), 3365 (SX)

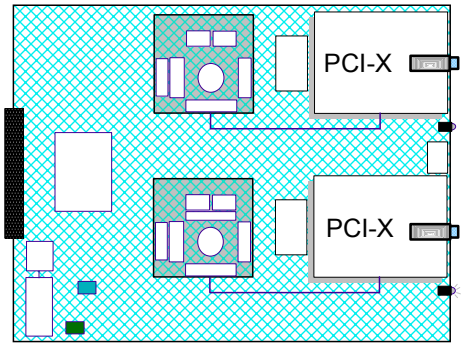


System z9 OSA-Express2 1000BASE-T Ethernet

- **New to OSA-Express2 family**
- **Supports auto-negotiation to 10, 100, 1000 Mbps over Category 5 copper**
- **Capable of achieving line speed**
  - Actual throughput is dependent upon environment
- **Supports:**
  - **Large send -** offloading TCP segmentation
  - **Concurrent LIC update** to minimize network traffic disruption
  - **640 TCP/IP stacks -** improved virtualization
  - **Layer 2 support -** protocol-independent packet switching

Mode	CHPID	Description
OSA-ICC	OSC	3270 data streams
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 NCP	OSN	Open System Adapter for NCP

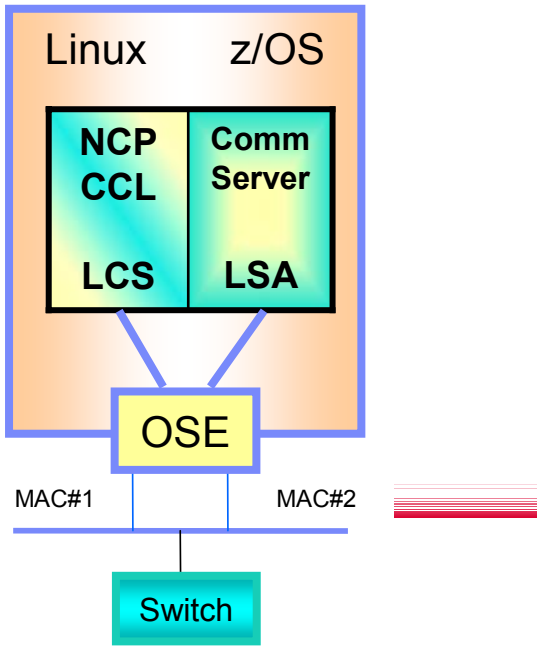
1000BASE-T Ethernet #3366



OSA-Express2 OSN (Open System Adapter for NCP)

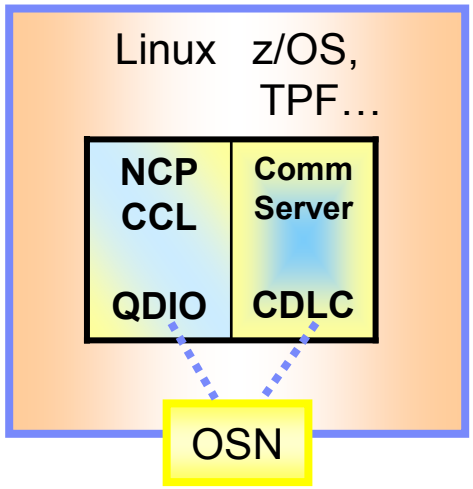
- **Supporting Channel Data Link Control (CDLC)**
  - Channel protocol for communication from a host operating system to 374x
    - Communications Controller for Linux on System z9 (CCL)
      - Emulates 374x control unit

CCL V1R1



- ✓CCL
- ✓LAN connectivity – external communications
- ✓Traditional connectivity
- ✓OSA and non-QDIO (CHPID type OSE)
- ✓Traffic flows on the LAN

OSA for NCP



- ✓Traffic flows LPAR-to-LPAR
- ✓OSA-Express2 (1000BASE-T, GbE) is the bridge
  - Internal communications – CCL to CDLC
  - SNA PU type 5 and PU type 2.1
- ✓QDIO extensions for CHPID type OSN
- ✓Unlike-to-unlike with “glue logic”

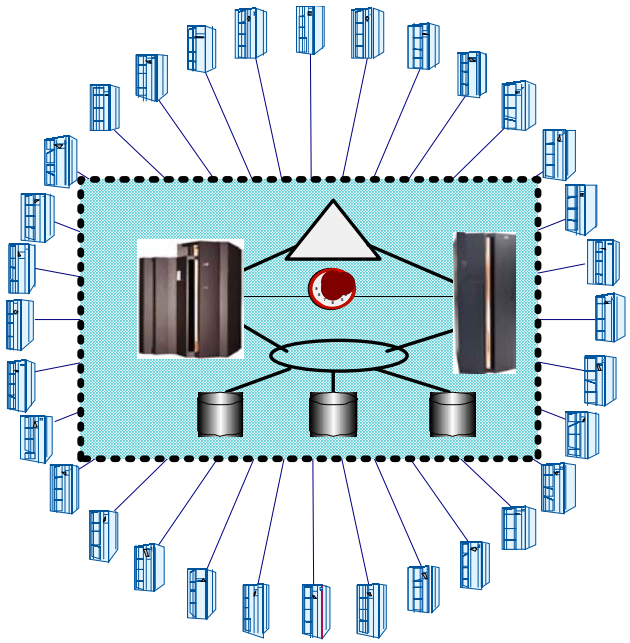
OSA Layer 2/Layer 3 and LPAR-to-LPAR enhancement

- **Currently OSA port sharing is only supported when using the same transport mode (either Layer 2 or Layer 3). That means, if one LPAR is using a given OSA-Express port for Layer 2 traffic, then all LPARs sharing that same port can only use it for Layer 2 traffic (obviously, the same with Layer 3). This restriction applies for traffic going LPAR-to-LPAR and out on the LAN.**
- **With this new enhancement, restriction is removed.**
- **OSA port sharing between virtual switches can now communicate whether the transport mode is the same (Layer 2 to Layer 2) or different (Layer 2 to Layer 3).**
  - This enhancement is designed to allow mixing of Layer 2 and Layer 3 traffic, potentially helping reduce the total cost of networking.
  - Previously, Layer 2 and Layer 3 TCP/IP connections through the same OSA port (CHPID) were unable to communicate with each other LPAR-to-LPAR using the Multiple Image Facility (MIF).
- **This enhancement is designed to facilitate a migration from Layer 3 to Layer 2 and to continue to allow LAN administrators to configure and manage their mainframe network topology using the same techniques as their non-mainframe topology.**

System z9 Parallel Sysplex Support

System z9 CFCC Level 14

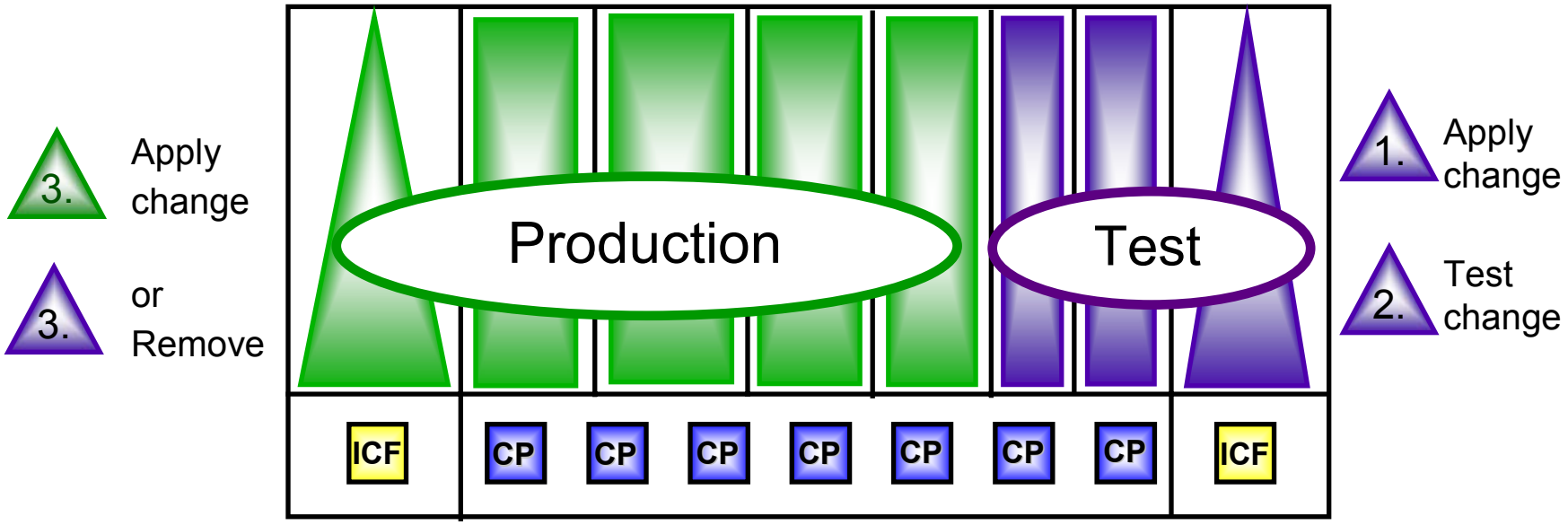
- **Function and Potential Benefit**
  - Contains improvements to the CF dispatcher and internal serialization mechanisms designed to better manage coupled workloads
- **Requirements and Support**
  - z/OS 1.4 and higher
    - Optional APAR fix OA08742 to allow sysplex connectors to request structure allocation in a Level 14 Coupling Facility
  - z/VM 4.4 and higher for virtual CF support
- **CF Storage Sizing with CFCC level 14**
  - May increase storage requirements
  - Use CFSIZER tool to determine: www.**ibm.com**/servers/eserver/zseries/cfsizer/



1 to 32 Systems

System z9 continues to meet the requirements for advanced clustering

System z9 concurrent CFCC changes

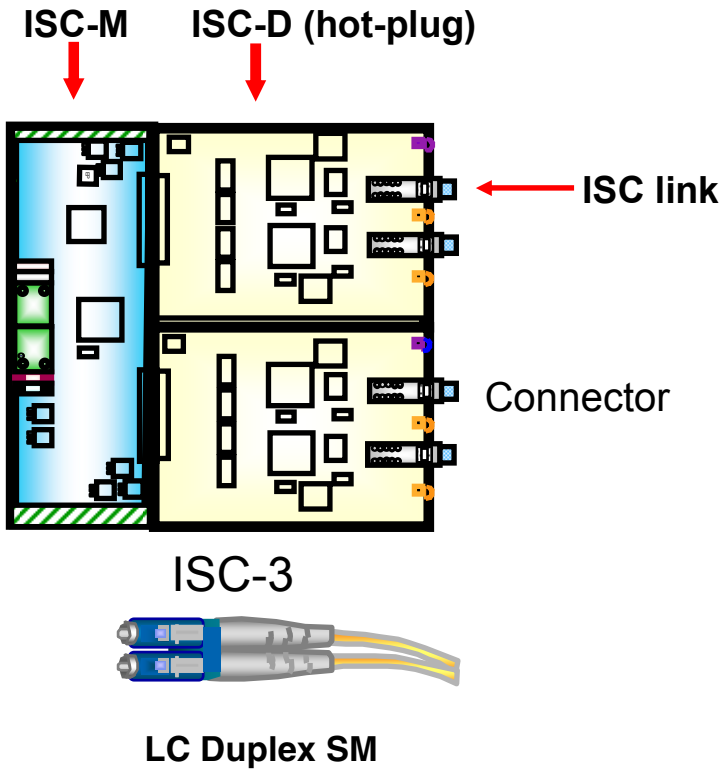


- **Apply previously disruptive CFCC changes with little disruption**
  - Disruption occurs one CFCC LPAR at a time to activate or remove a change
  - Allows rolling CFCC maintenance across CF LPARs
    - Similar to rolling z/OS maintenance across OS images
  - Helps reduce the requirement to isolate test CFs from production OS/CF images

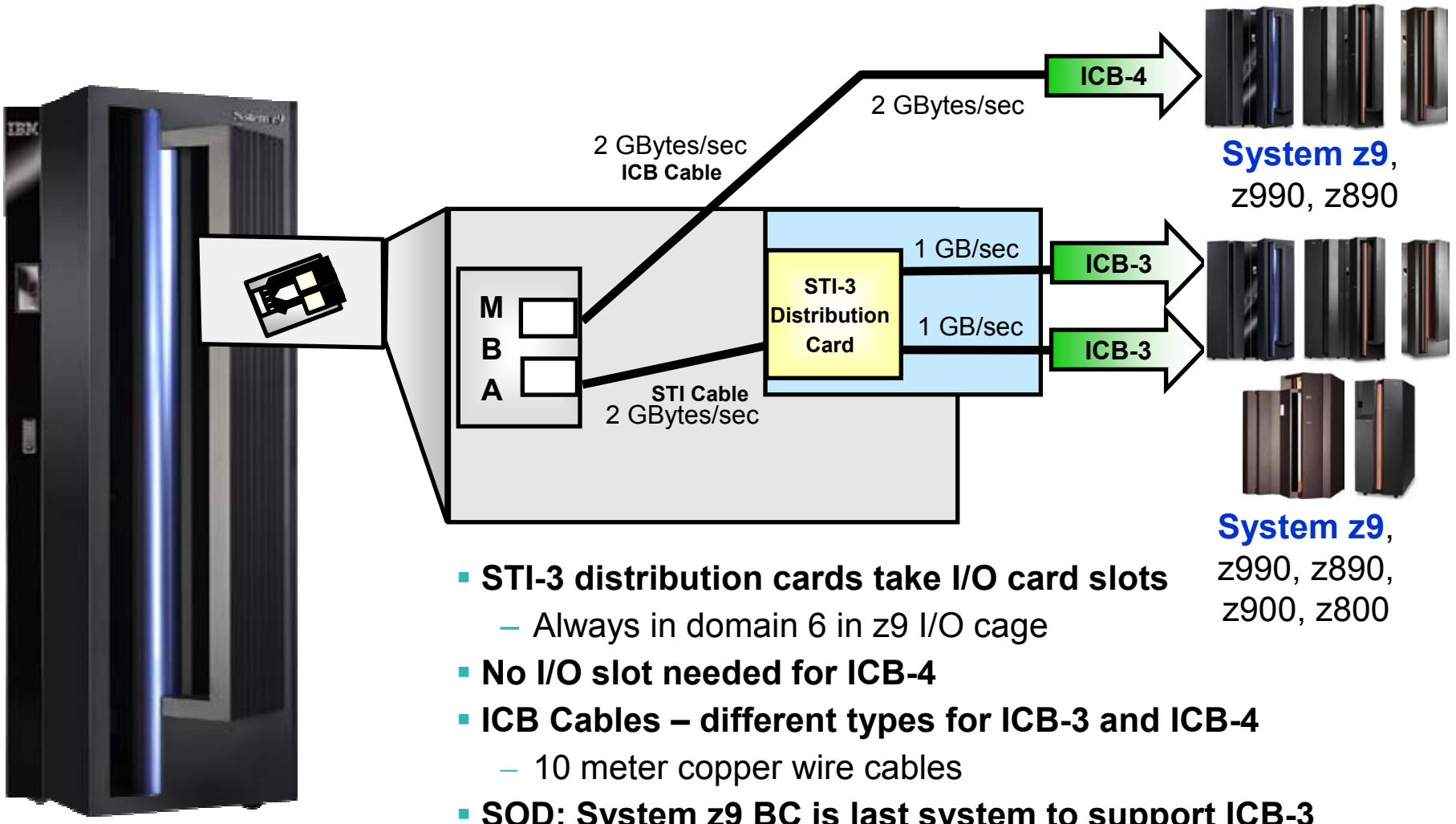
Note: Some very infrequent CFCC related changes may still require power-on-reset

System z9 Fiber Optic Coupling Links

- **InterSystem Channel-3**
  - 3rd generation fiber optic coupling link
  - ISC-3 links ordered in increments of one (two for 8P2197 “long distance” links)
  - Activated links balanced across features
- **Peer mode only**
  - Connects to System z9 and zSeries® only
  - **No connectivity to any 9672 or 9674**
- **Distance supported**
  - Unrepeated up to 10 km (20 km RPQ 8P2197)
  - Repeated up to 100 km
- **ISC-3 feature codes: 0217(ISC-Mother card), 0218 (ISC-D / ISC link)**
  - Active link - FC 0219 (2 Gbit/second)
  - Two links – RPQ 8P2197 (1 Gbit/second)
  - Four ports per ISC-M (two ports per ISC-D)
  - Supports 9 micron single mode fiber
- **Up to 48 ISC-3 links per System z9 server**



System z9 ICB Coupling Links



- **STI-3 distribution cards take I/O card slots**
  - Always in domain 6 in z9 I/O cage
- **No I/O slot needed for ICB-4**
- **ICB Cables – different types for ICB-3 and ICB-4**
  - 10 meter copper wire cables
- **SOD: System z9 BC is last system to support ICB-3**

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

System z9 HMC and SE

System z9 HMC configuration

HMC orderable on 2094:

Closed desktop, browser like UI

HMC #0081 – Dual Ethernet

9.117.59.1

192.101.150.2

Primary SE LAN
Always Ethernet

Secondary SE LAN
Always Ethernet

192.101.150.154

192.101.150.155

9.117.59.154

9.117.59.155

Other HMCs OK:

HMC #0077, #0078, #0079, #0080

HMC #0076 conversion of #0075

HMC #0075

Note: **Ethernet attach standard.**
Only HMCs #0076, #0077, and #0079 have dual Ethernet support

2094 SE LAN attach feature

#3065 – Dual Ethernet

HMC/SE LAN Components

#0089 – 10/100, 16-port Ethernet switch

Note: Offered on orders including an HMC

2094 has no internal switch

For z990/890, z900/800:

Migrate to Ethernet

Order HMC #0081

Convert #0075->#0076

Order Ethernet SE #3063

Order Ethernet TKE

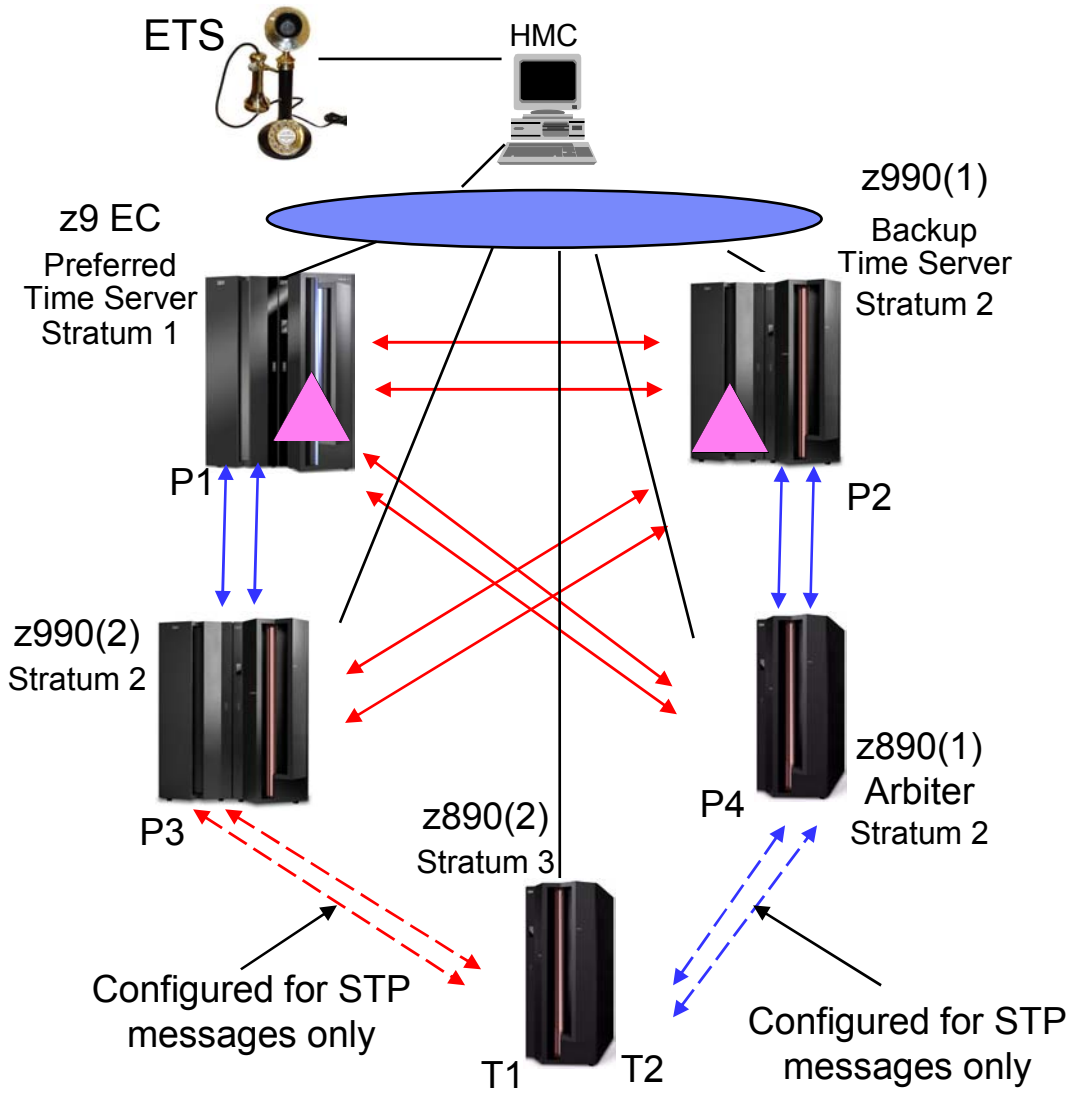
Use SE and HMC Ethernet available on older equipment

Server Time Protocol (PREVIEW)

Prerequisites

- **Hardware**
  - System z9, z990 or z890 server
    - Concurrent MCLs on Driver 55 (z990, z890) to install STP Licensed Internal Code (LIC)
    - Concurrent MCL to enable STP
  - System z9 HMC
    - New HMC
    - Upgrade of existing HMC
- **Operating System**
  - z/OS 1.7 or higher
    - STP code in z/OS 1.7 (default=disabled)
    - PTFs for STP support
    - PTF to enable STP
  - Toleration PTFs for z/OS 1.4, 1.5, 1.6 in Mixed Timing Network

STP-only Timing Network example

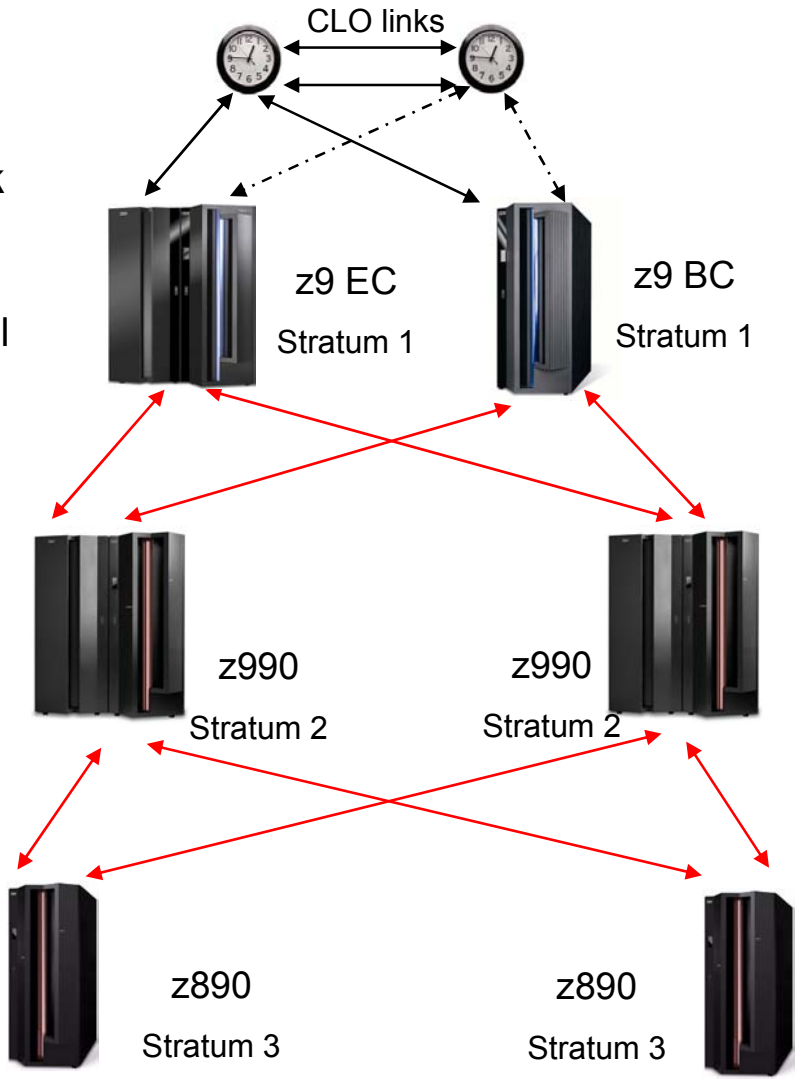


- **Parallel Sysplex**
  - P1, P2, P3, P4, (2) ICFs
- **If Servers configured with ICFs selected as Preferred and Backup Time Servers, connectivity in place for CF messaging can be used for STP message communication**
  - Any STP capable server can be selected as Preferred or Backup
- **If CTN scope is greater than Sysplex scope, additional ISC or ICB links may need to be configured for STP**
  - T1, T2 on z890(2) not in P1-P4 Parallel Sysplex but in same CTN as P1-P4

ISC-3 links
ICB-3 links

Mixed Timing Network

- **Allows co-existence of ETR and STP networks**
- **9037 ETRs are the "root" for both networks**
- **Need at least one server in Mixed Timing Network capable of:**
  - Time coordination using ETR (attach to 9037) and time coordination using STP message-based protocol
- **Any STP capable server (z9, z990, z890) attached to 9037 can be a Current Time Server (Stratum 1) for the STP network**
  - As long as one server in the CTN is still connected to the ETR, there is no Current Time Server (CTS) configured. Configuring a CTS removes connectivity of all Servers to the ETR
  - Backup time server can be Stratum 2 or 3
    - Stratum 2 uses Stratum 1 as clock source
    - Stratum 3 uses Stratum 2 as clock source
- **9037 console continues to be used for all timing related functions of the Mixed Timing Network**
- **Hardware Management Console (HMC) must be used for Coordinated Timing network configuration**



Maximum Stratum level for STP is 3. Stratum 3 not recommended

NOTE: Example may not be a typical Parallel Sysplex configuration

Implementation Assistance Program (IAP)

- **Program to accelerate customer adoption of STP with IBM's assistance.**
- **You will receive:**
  - technical planning assistance
  - review of migration plans to an STP network
  - technical support during the implementation of STP
- **Potential candidates for IAP**
  - Require a multi-site sysplex across more than 40 km

    NOTE: using STP at a distance greater than 40km is not a requirement for IAP.
  - New Parallel Sysplex implementation – single or multi site

    Note: a SOD was included in the z9 EC announcement that IBM plans to withdraw the 9037-002 from marketing during 2006
  - Latest Parallel Sysplex technology
- **IAP on z990, z890 and System z9 has started**
- **Contact Angelo Corridori (angeloc@us.ibm.com) if interested or have questions**

System z9 Connectivity Backup Charts

System z9 and zSeries Coupling Link Connectivity

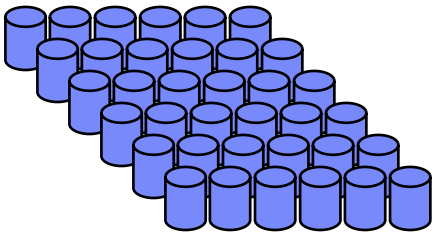
To: / From:	System z9 ISC-3	System z9 ICB-3	System z9 ICB-4
z900/z800 ISC-3	2 Gb/sec* Peer Mode	No	No
z9/z990/z800 ISC-3	2 Gb/sec* Peer Mode	No	No
z900/z800 ICB-3	No	1 GB/sec Peer Mode	No
z9/z990/z890 ICB-3	No	1 GB/sec Peer Mode Recommendation: Use ICB-4	No
z9/z990/z890 ICB-4	No	No	2 GB/sec Peer Mode
z900/z990 ICB-2	Impossible!		

***RPQ 8P2197** links – Run at 1 Gb/sec in Peer mode and support 20 km unrepeated instead of 10 km.
Coupling Connectivity to 9672 or 9674 fiber optic links is not supported
Coupling Connectivity to compatibility mode (sender/receiver) links on any machine not supported

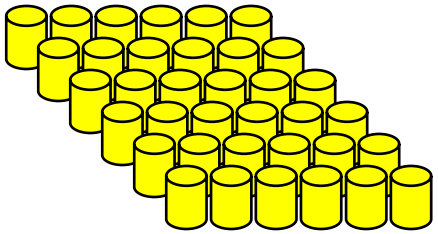
Single Subchannel Set per LCSS (per Partition)

z890 2086 Processor	
LCSS 0	**LCSS 1**
Partitions	**Partitions**
Subchannels **64512** **(63 K)**	**Subchannels** **64512** **(63 K)**
Channels	**Channels**

Subchannel per device

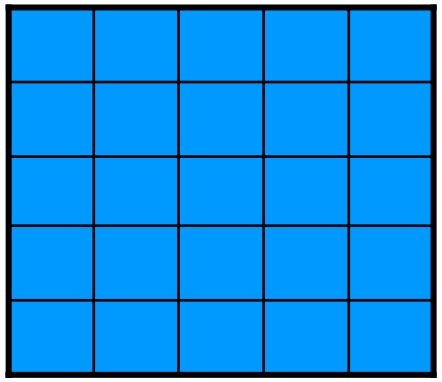


Base Devices

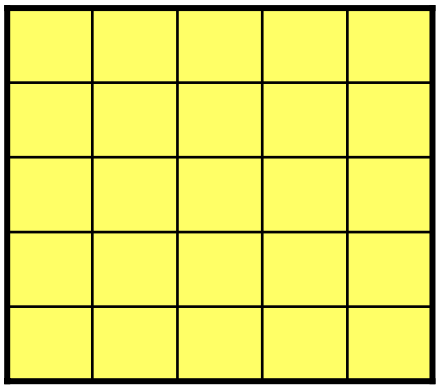


Alias Devices

System z9™ Multiple Subchannel Sets per LCSS (MSS)



- **Set 0 – Up to 65,280 subchannels**
  - Always present if the LCSS is defined
  - Any device for allocation
  - Reference in JCL
  - UCB Services
  - Messages
  - Commands

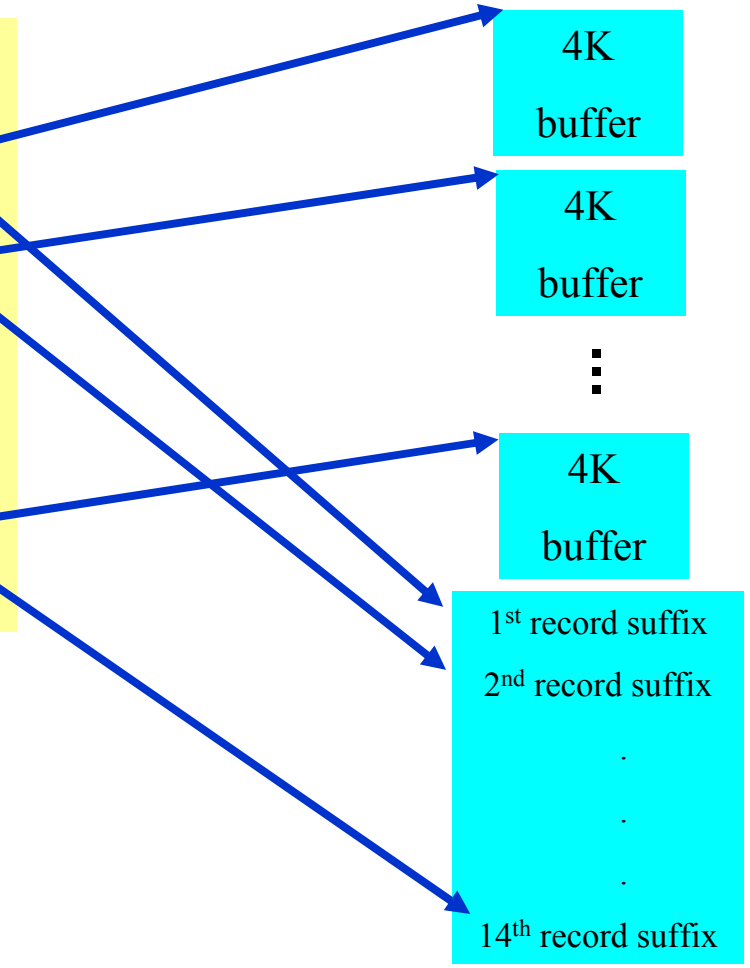


- **Set 1 – Up to 65,535 subchannels**
  - Optionally defined
  - PAV alias devices only
  - Display Commands
  - Limited messages

- **Designed to be compatible with existing storage CUs that support PAV**

Without the MIDAW facility – 29 CCWs with data chaining Reading 14 disk records with 4K data, 32 byte suffix

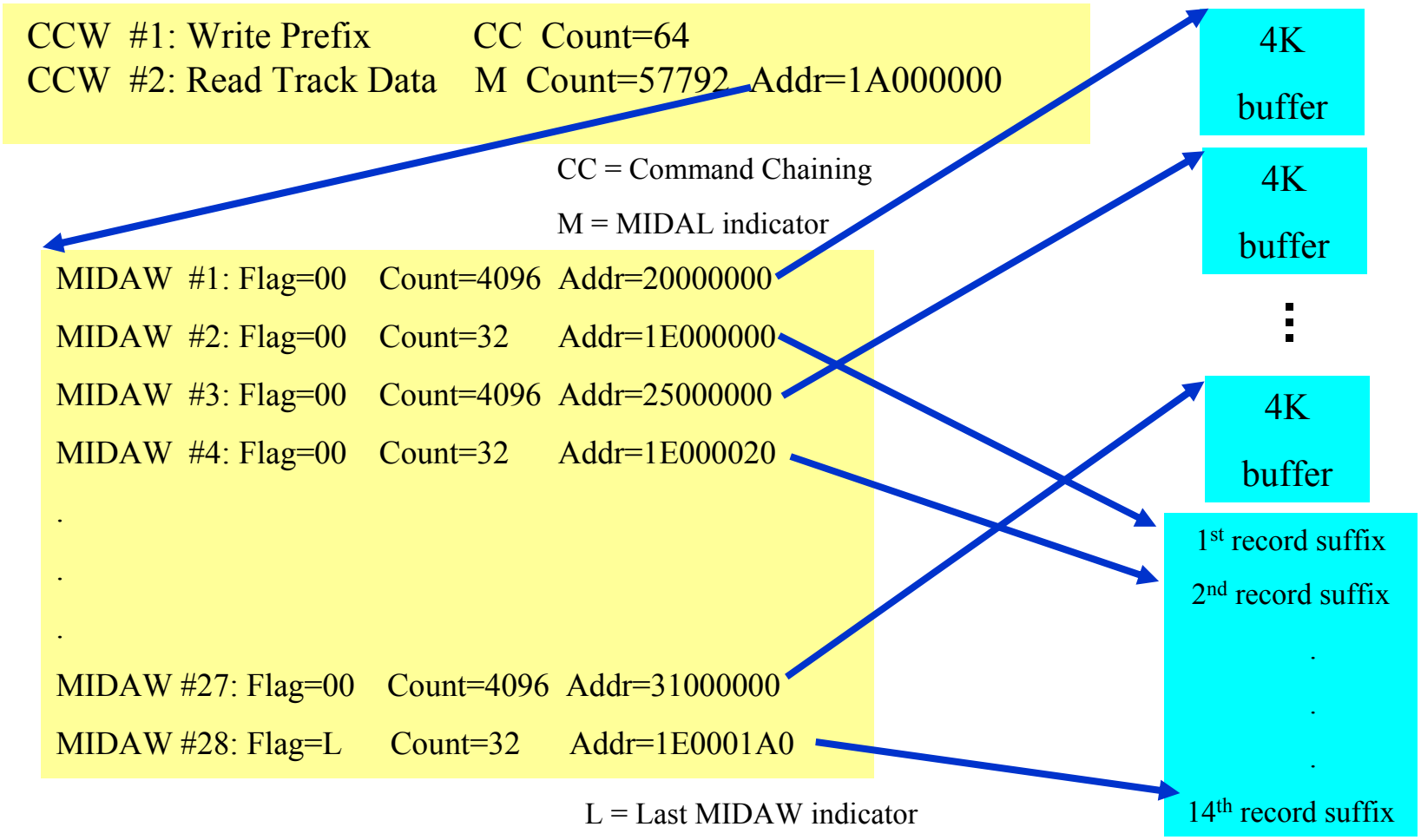
```
CCW  #1: Write Prefix          CC  Count=64
CCW  #2: Read Track Data  CD  Count=4096  Addr=20000000
CCW  #3: Read Track Data  CD  Count=32    Addr=1E000000
CCW  #4: Read Track Data  CD  Count=4096  Addr=25000000
CCW  #5: Read Track Data  CD  Count=32    Addr=1E000020
            .
            .
            .
CCW #28: Read Track Data  CD  Count=4096 Addr=31000000
CCW #29: Read Track Data      Count=32    Addr=1E0001A0
```



CC = Command Chaining

CD = Data Chaining

System z9 using the MIDAW facility – 2 CCWs Reading 14 disk records with 4K data, 32 byte suffix



zEnd

System z9

RETURN TO INDEX