Washington Systems Center - Storage

IBM

Accelerate with IBM Storage:
Copy Services Manager Update and zHyperLink Update

Randy Blea, CSM Software Architect and Craig Gordon, DS8000 Subject Matter Expert

Accelerate with IBM Storage Webinars

The Free IBM Storage Technical Webinar Series Continues in 2019...

Washington Systems Center – Storage experts cover a variety of technical topics.

Audience: Clients who have or are considering acquiring IBM Storage solutions. Business Partners and IBMers are also welcome.

To automatically receive announcements of upcoming Accelerate with IBM Storage webinars, Clients, Business Partners and IBMers are welcome to send an email request to accelerate-join@hursley.ibm.com.

Located in the Accelerate with IBM Storage Blog:
https://www.ibm.com/developerworks/mydeveloperworks/blogs/accelerate/?lang=en

Also, check out the WSC YouTube Channel here:
https://www.youtube.com/channel/UCNuks0go01_ZrVVF1jgOD6Q

2019 Upcoming Webinars:

April 18 - IBM DS8880 R8.5 GUI Demo
  Register Here: https://ibm.webex.com/ibm/onstage/g.php?MTID=ea2d6470284e858931dedd393edf2cc58
April 23 - Copy Services Manager and zHyperLink Update
  Register Here: https://ibm.webex.com/ibm/onstage/g.php?MTID=e4ad33b2fc4451f8fb205ab6cad2f125c
April 25 – Data Reduction Pools: Overview and Best Practices
  Register Here: https://ibm.webex.com/ibm/onstage/g.php?MTID=e647561ca5cd145596bb1770fa1bd5fe7
May 2 - TS7760 Best Practices (A View from the Field)
  Register Here: https://ibm.webex.com/ibm/onstage/g.php?MTID=e553c6faed4058a414f133151d5743f71

Copy Services Manager Update

What's New in IBM Copy Services Manager V6.2.5 - eGA'd 3/22/19

Allow session operators to only see the sessions that they have access to

Ability to remove Practice relationships after disaster recovery testing

CSM CLI on z/OS SAF support for Facility Class Management

New Dual Control support for added security

Allow Session Operators to only view the sessions they can manage

- Session operators in the past could only issue commands to sessions they were authorized to manage however, they had monitor authority to view all other sessions

- With CSM 6.2.5, a session operator will no longer have the ability to view sessions they have not be authorized to manage

Facility Class support for CSM CLI running on z/OS

- In previous releases, the CSM CLI while running on z/OS required a user name and password to either be specified on the command line on in an encrypted file

- Customers running the CSM CLI had to update the CSM environment every time they changed their password due to the security rules in their environment

- Customers requested the ability to use a Facility Class like other z/OS applications to authorize the CSM CLI

- In CSM 6.2.5, when running on z/OS, a Facility Class can now be defined such that the user currently logged into the OMVS shell, or specified in a BPXBATCH job are validated through the Facility Class instead of having to enter a password

Remove Practice relationships in Practice sessions

- CSM Practice sessions allow customers to Flash a consistent copy at the remote site in order to practice their disaster recovery

- In previous release once the Flash was issued, if No Copy was specified the relationships remained on the hardware for the life of the session

- When using Extent Space Efficient volumes for practice volumes, No Copy is the recommended setting so that the practice copy does not consume backend space

- However, if the relationship remains on the hardware for the life of the session it may eventually consume that space

- In CSM 6.2.5, CSM has a new Terminate HxIx command that allows customers to remove the practice relationship when they have completed their testing

Dual Control Support for increased security

- In 6.2.5, CSM has added the ability to enable Dual Control support across the CSM server

- When Dual Control is enabled on a server, all actions that could have negative impacts on the replication environment, require two users with proper authority to authorize the action

- This helps both in the prevention of malicious acts as well as provides a mechanism for 'double checking' actions against sessions.

- The following actions will require two users when Dual Control is enabled

  - Run commands against sessions
  - Modifying properties of a session
  - Adding/Removing copy sets from a session
  - Modifying/Deleting/Enabling/Disabling/Running Scheduled Tasks
  - User administration actions (adding/removing/etc)
  - Setting up Active/Standby server support
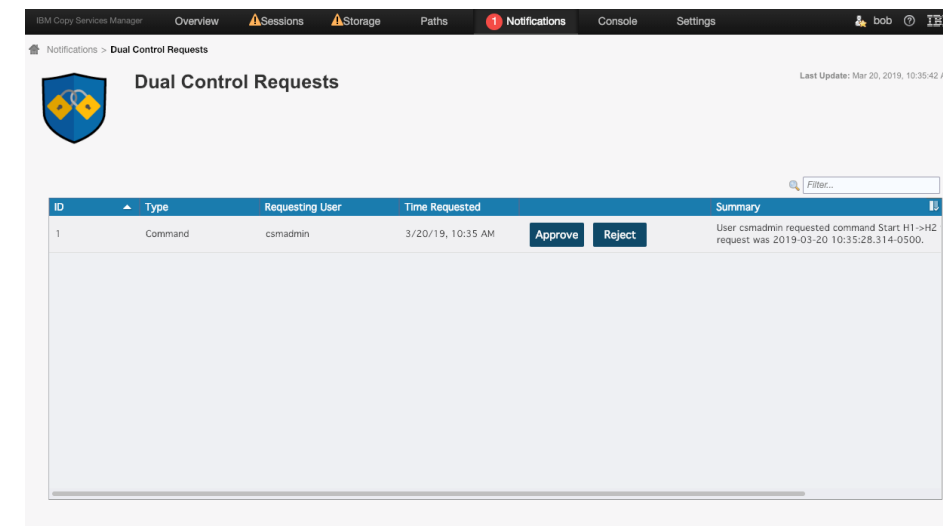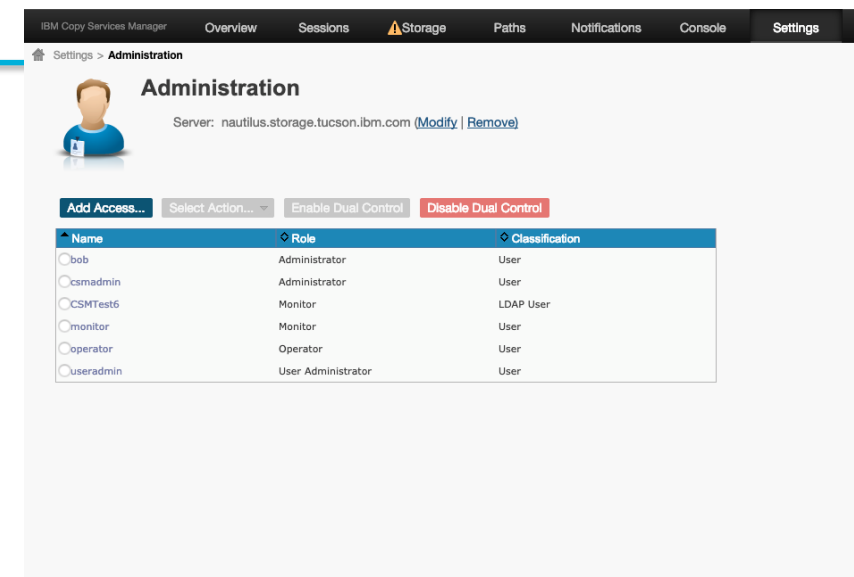
zHyperLink Update

What is zHyperLink?

zHyperLink Express is a direct connect short distance IBM Z I/O feature designed to work in conjunction with a FICON or High Performance FICON SAN infrastructure

- New card that resides in the z14 PCIe I/O drawer

- Connects directly into the I/O enclosure of the DS8880. A new host adapter card is not required

- 150 meter maximum distance
  - Metro mirror secondary must also be within 150 meters to support writes

- FICON connectivity to each storage subsystem still required
  - You may continue to define up to 8 FICON paths per logical control unit
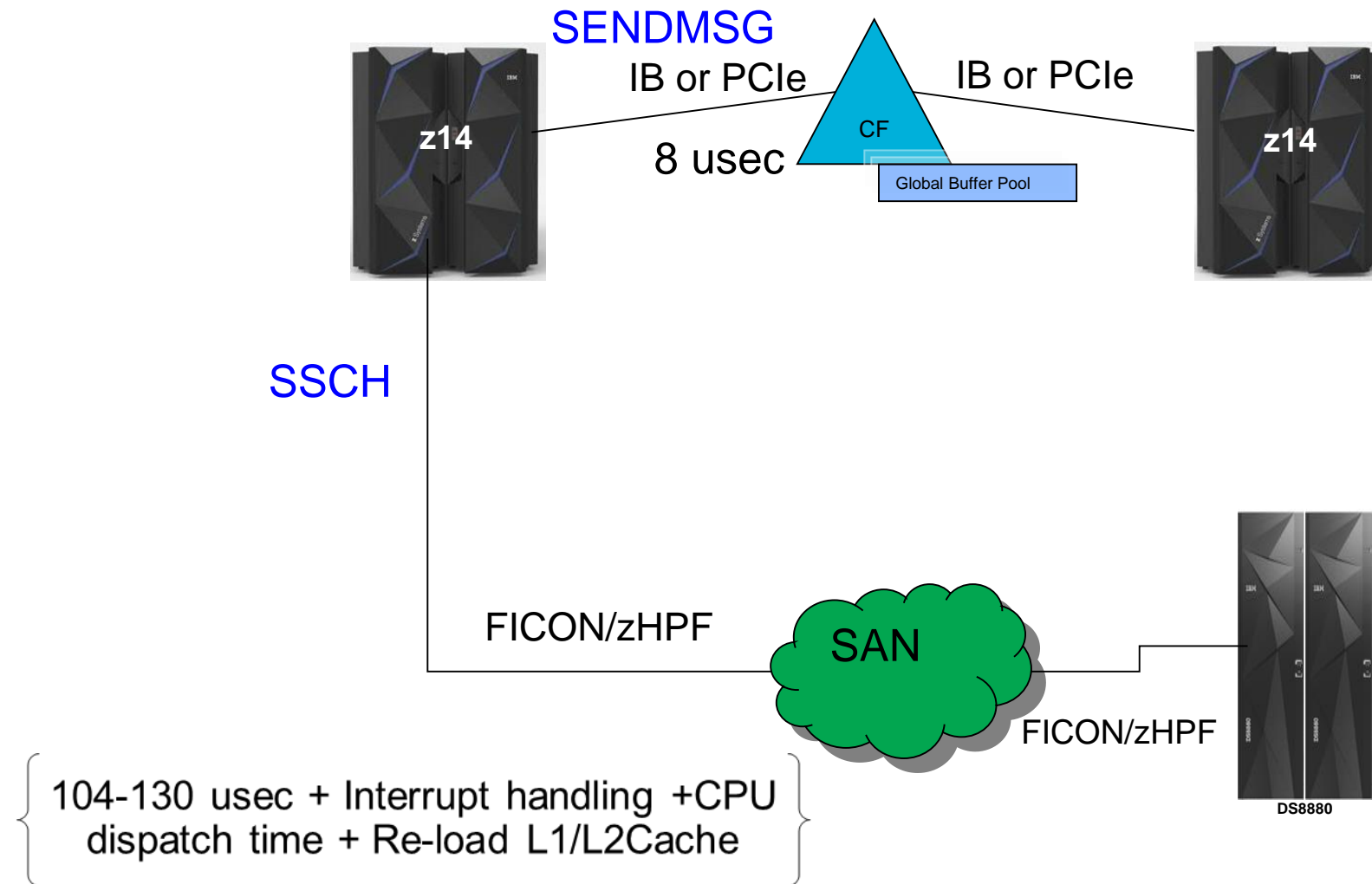
IBM z14

Up to 10x reduction In latency

IBM DS8880

Today's I/O

SENDMSG

IB or PCIe   IB or PCIe

z14   CF   z14

8 usec

Global Buffer Pool

SSCH

Typically over 80% cache hit ratio on random reads.

100% cache hit on writes.

FICON/zHPF   SAN

FICON/zHPF

DS8880

104-130 usec + Interrupt handling +CPU dispatch time + Re-load L1/L2Cache

With zHyperLink

SENDMSG

IB or PCIe

IB or PCIe

z14

CF

z14

8 usec

Global Buffer Pool

SSCH

zHyperLink™

Typically over 80% cache hit ratio on random reads.

100% cache hit on writes.

FICON/zHPF

SAN

FICON/zHPF

DS8880

104-130 usec + Interrupt handling +CPU dispatch time + Re-load L1/L2Cache

vs 20 usec
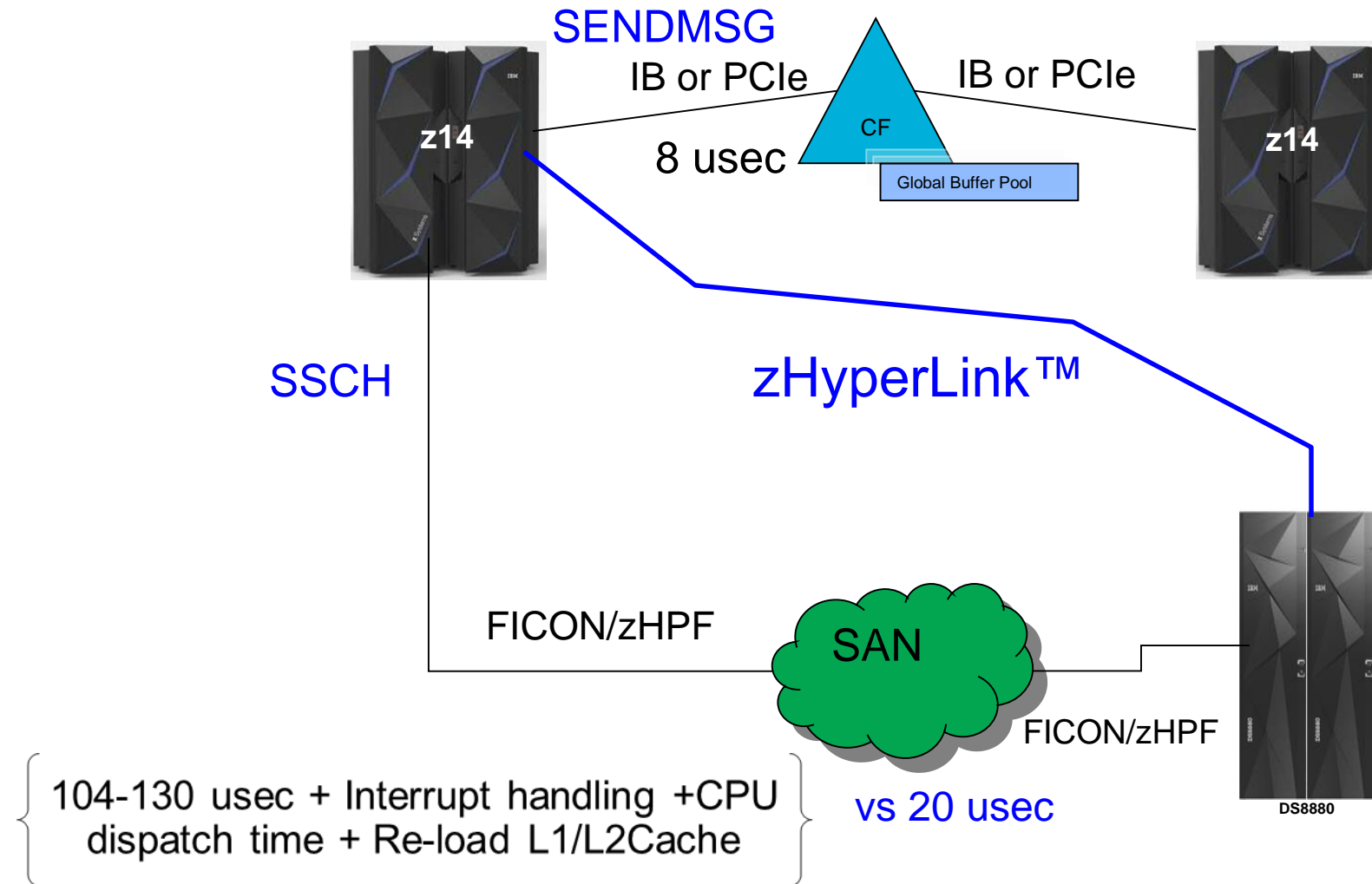
zHyperLink Infrastructure at a Glance

- **z14 zHyperLink Express Adapter**
  - Two ports per adapter
  - Maximum of 16 adapters (32 ports)
  - Function ID Type = HYL
  - Up to 127 Virtual Functions (VFs) per port
  - Point to point connection using PCIe Gen3
  - Maximum distance: 150 meters

- **DS8880 zHyperLink Ports**
  - Two ports per enclosure
  - Maximum ports
    - Up to 16 ports on DS8888
    - Up to 12 ports on DS8886
    - Up to 4 ports on DS8884
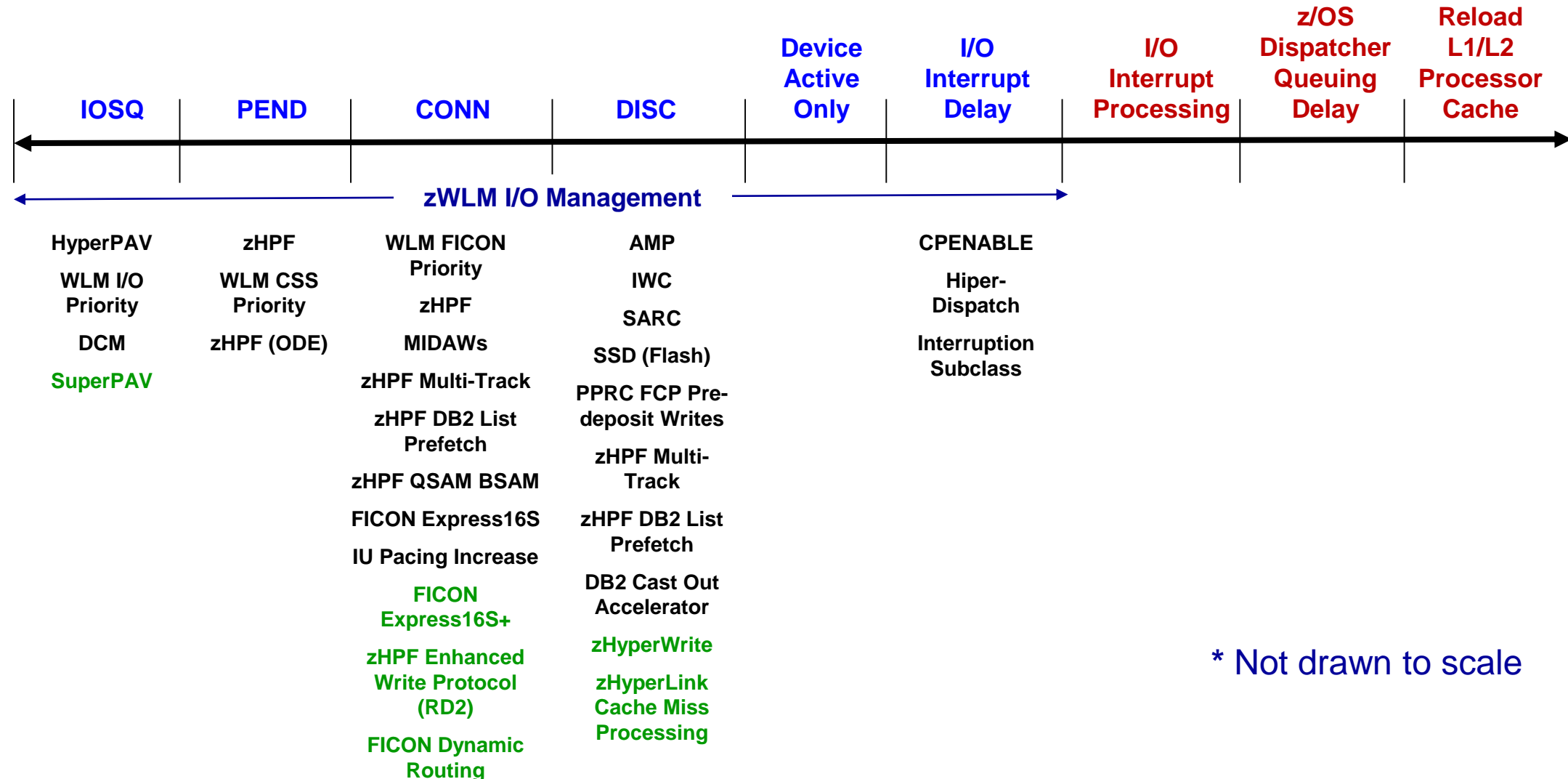 - Point to point connection using PCIe Gen3
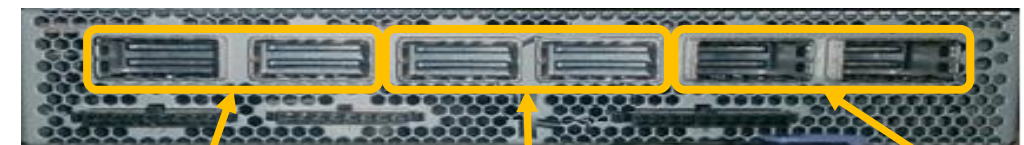
DS8880 internal PCIe Fabric

zHyperLink ports

HPFE Gen2

z/OS I/O Latency Improvement Technologies

Components of I/O Latency*

IOSQ	PEND	CONN	DISC	Device Active Only	I/O Interrupt Delay	I/O Interrupt Processing	z/OS Dispatcher Queuing Delay	Reload L1/L2 Processor Cache

zWLM I/O Management

IOSQ	PEND	CONN	DISC		I/O Interrupt Delay
HyperPAV	zHPF	WLM FICON Priority	AMP		CPENABLE
WLM I/O Priority	WLM CSS Priority	zHPF	IWC		Hiper-Dispatch
DCM	zHPF (ODE)	MIDAWs	SARC		Interruption Subclass
SuperPAV		zHPF Multi-Track	SSD (Flash)		
		zHPF DB2 List Prefetch	PPRC FCP Pre-deposit Writes		
		zHPF QSAM BSAM	zHPF Multi-Track		
		FICON Express16S	zHPF DB2 List Prefetch		
		IU Pacing Increase	DB2 Cast Out Accelerator		
		FICON Express16S+	zHyperWrite		
		zHPF Enhanced Write Protocol (RD2)	zHyperLink Cache Miss Processing		
		FICON Dynamic Routing			

* Not drawn to scale

zHyperLink Expectations

- zHyperLink Express improves application response time, cutting I/O sensitive workload response time by up to 50% without requiring application changes.

  - Client Value:
    - Lower cost and less risky than application changes to get lower response time
    - Enables new business logic in existing application at same or lower response time.

- zHyperLink Express is designed for up to 10X lower read latency than High Performance FICON

  - Client Value:
    - Higher availability through Faster Recovery operations
    - Higher availability through improved Peak IO capacity
    - Basis for faster middleware with better scaling
      - Middleware must use zHyperLink or exploit software that uses zHyperLink
    - Lower Latency and better throughput for FICON storage systems

Note: Workload response time estimate is based on IBM internal measurements and projections that assume 75% or more of the workload response time is associated with read DASD I/O and the storage system random read cache hit ratio is above 80%. Projections of read I/O latency are based on I/O service times and CPU queueing delays from IBM internal measurements. The actual performance that any user will experience may vary.

As an exploiter, how do I know if zHyperLink is right for me?

- Good candidates

  - Synchronous (blocking) reads where the data is likely to be in cache
  - Synchronous writes that follow a log write pattern. If synchronous mirroring is used (metro mirror):
    - The exploiter must also support zHyperwrite (software mirroring)
    - The metro mirror secondary device must be within a 150 meters of the processor and support zHyperlink writes

- Bad candidates

  - Synchronous reads where data is not likely to be in cache
  - Random writes
  - Lazy/delayed writes – non-blocking
  - Read-ahead/pre-staging of data – non-blocking
  - I/O requests that do not meet the requirements on the next page such as very large sequential writes

- Note: Exploiter = DB2, VSAM

IBM zHyperLink Exploitation

- Currently Supported

  - DB2 V11 and V12 Read Support (Batch and Transactions)
    - Synchronous DB2 database reads – 4K CI/page sizes only
    - For DB2, think of zHyperLink as part of the storage hierarchy in terms of latency that is between the DB2 buffer pools and DS8K cache
  - DB2 Active Log Support
    - zHyperLink Write exploitation; supports metro mirror with HyperSwap storage replication and zHyperWrite

  - Ability to turn zHyperLinks on and off at the Dataset level
    - Via Storage Class policy or via VARY SMS command

- Future Support

  - VSAM Read Support (Batch and Transactions)
    - All VSAM record type data sets (i.e. KSDS, ESDS, RRDS, VRRDS) support zHyperLink, along with all types of access, except NSR Sequential.

Software Requirements

- z/OS - Support for z/OS 2.2 and up

- Fix Category Information

  - Used to group/associate PTFs for a particular category of software fixes (e.g., support a particular hardware function).

  - Fix Category: IBM.Function.zHyperLink

    - The following are the z/OS 2.3 APAR numbers:

```
FMID     APAR      Comments
======= ======= ======= ============================
HDZ2230 OA53199      DFSMS (Media Mgr, Dev. Support)
        OA50681      DFSMS (Media Mgr, Dev. Support)
        OA53287      DFSMS (Catalog)
        OA53110      DFSMS (CMM)
        OA52329      DFSMS (LISTDATA)
        OA52876      VSAM RLS zHyperlink Exp. (open)
        OA52941      VSAM zHyperlink Exp.  (open)
        OA52790      SMS zHyperlink Exp.  (open)
        OA54822      SMS/CMM/BAM
        OA54824      Media Manager
        OA54825      ISMF
        OA54826      Naviquest
        OA54872      AMS
```

I/O Requirements

- Exploitation is via Media Manager interface
  - Used for DB2, VSAM, system logger etc.
  - No EXCP or EXCPVR support

- Read and update writes
  - No format writes, no keys
  - Update writes must be log type writes - always go forward in data set and may wrap around

- Fixed sized records
  - Data buffers must be 16 byte aligned
  - DS8000 record sizes currently supported
    - Reads: 512 byte multiples up to 4K, plus any of the preceding record sizes plus 32 bytes (extended format) – single record per I/O only
    - Writes: 4K, 4K+32 – multiple records per I/O

zHyperLink Enablement Requirements for Writes (1 of 2)

- The following must be done to enable zHyperlink writes for a DB2 log:

  - The DB2 log data set must be defined in a storage class with zHyperLink writes enabled

  - DB2 log writes must be enabled for zHyperLink writes in DB2 ZPARM

  - The device must be enabled for zHyperLink writes. This requires all of the following to be true:

    - Storage controller must be enabled for zHyperLink writes
    - System must be enabled for zHyperLink writes (IECIOSxx parmlib or SETIOS command)
    - zHyperLinks must be available

zHyperLink Enablement Requirements for Writes (2 of 2)

- If device is a primary device that is in a metro-mirror (PPRC) relationship, the following must also be true:
  - Secondary storage controller(s) must be enabled for zHyperLink writes
  - Primary and secondary storage controllers must support zHyperwrite
  - zHyperLinks must be available to the secondary storage controller(s). This implies the secondary must be within 150 meters of the processor.
  - System must be enabled for zHyperwrite (IECIOSxx or SETIOS)
  - The PPRC volume state for each metro-mirror secondary must be duplex. zHyperlink writes are not supported if the volume state is copy pending or suspended.
  - The primary and secondary devices must be in the Hyperswap configuration. Note that hyperswap does not need to be currently enabled.
    - For example, z/OS is unable to test a particular secondary device for connectivity during its 5 minute polling loop. This will disable the ability to perform a hyperswap, but the hyperswap configuration is still loaded and zHyperlink writes can still be done to other secondary devices.

Additional Factors Affecting zHyperLink Usage

- Even if all of the requirements on the preceding pages have been met, the following can prevent zHyperLink from being used for a particular I/O request:

  - There are no available links (PFIDs). Only 4 requests can be active on a link at a time and if the number if exceeded on all links, zHyperlink is not used for that I/O request.

  - The device is reserved to another system or a reserve is pending from the current system

  - A device or subsystem long busy conditions exists or in some other state that prevents asynchronous I/O from starting or completing

  - z/OS is performing some type of recovery for the device. E.g., verifying the device connected at the other end of a FICON link after an error occurs or the link becomes operational

DB2 zHyperLink Write Exploitation

- Improved throughput for high insert workloads
  - Record inserts cause keys to be inserted, which may require an index page to be split
  - Index splits cause log writes to occur, so log latency is important
- Improved throughput for commit limited workloads
  - The combination of Faster IO link time (5X) for write I/Os and elimination of dispatcher delays can substantially improve Commit duration
- Reduced lock hold times
  - Enables higher transaction rates through existing applications.
  - Combined benefits with zHyperLink Read support which also reduces lock hold time

- Notes:
  - DB2 log writes are eligible to run on a ZIIP
  - zHyperWrite technology is a pre-requisite to zHyperLink to mitigate synchronous replication delays

VSAM zHyperLink™ Exploitation

- When zHyperLink™ is enabled on a LPAR and for individual data sets:

  - All VSAM (physical) read I/O will request zHyperLink™ support.
  - Includes all record type VSAM data sets (KSDS, ESDS, RRDS, VRRDS, LDS).
  - Includes all types of VSAM read access (NSR,LSR,GSR,RLS), except NSR Sequential.
  - Currently only CISIZEs less than or equal to 4096 are supported.
  - User buffers must be on a quad-word boundary.
  - If the physical read I/O meets all the requirements for a zHyperlink request, then the task (TCB or SRB) issuing the I/O will not suspend or expect an interrupt.
  - Depending on system conditions, a "syncio" request can have a much improved elapse time, allowing VSAM applications to run faster with less latency. Ideally the request will spin for <20 us (times are dependent on the length of the links to the disk subsystem.

Enabling zHyperLink at the z/OS System Level

- New keyword in IECIOSxx and SETIOS to enable or disable zHyperLink

  - Separately disable reads vs writes

  - Default is NONE

  - This is a very broad control – enables <u>all</u> zHyperLink read or write activity on this z/OS system (LPAR)

IECIOSxx: ZHYPERLINK OPER=[ALL|READ|WRITE|NONE]

SETIOS ZHYPERLINK,OPER=[ALL|READ|WRITE|NONE]

ZHL can be used as an abbreviation for ZHYPERLINK

D IOS,ZHYPERLINK
IOS634I hh.mm.ss IOS SYSTEM OPTION
ZHYPERLINK IS {ENABLED FOR READ AND WRITE OPERATIONS}
{ENABLED FOR [READ|WRITE] OPERATIONS}
{DISABLED}
{NOT SUPPORTED BY THE PROCESSOR}

Enabling a Db2 Database or Log for zHyperLink

- Defined in the Db2 ZPARM. Can be changed dynamically while DB2 is active.

```
DSNTIPA3                 INSTALL DB2 - DATA PARAMETERS PANEL
 ===>


 Check parameters and reenter to change:
  1 PERMANENT UNIT NAME    ==> 3390_____   Device type for MVS catalog and
                                            partitioned data sets

  2 TEMPORARY UNIT NAME    ==> SYSDA_____   Device type for temporary data sets
  3 DB2 zHyperLinks SCOPE  ==> DISABLE__    Scope of zHyperLinks I/O connection
                                            (DISABLE, ENABLE, DATABASE, ACTIVELOG)


                                    ---------------- SMS ------------------
                                                                      ASS
                                                                      ---
  4 C                                                                   __
  5 N                                                                   __
  6 V                                                                   __
    A
```

Default: DISABLE
Data sharing scope: Member scope. It is recommended that all members use the same setting.
Online changeable: Yes

DISABLE
    Db2 does not use zHyperLink for any I/O requests.
ENABLE
    All eligible I/O requests that Db2 makes request the zHyperLink protocol.
DATABASE
    Db2 requests the zHyperLink protocol only for database synchronous read I/Os.
ACTIVELOG
    Db2 requests the zHyperLink protocol only for active log write I/Os.

VSAM zHyperLink™ System Parameters

- To enable VSAM zHyperLink™ Exploitation on each LPAR:

  SYS1.PARMLIB(IGDSMSxx):   **VSAM_zHyperLink(<u>NO</u>/YES)**

  Where:

  NO – Do not request zHyperLink™ exploitation for VSAM reads on this lpar (default).

  YES – Request zHyperLink™ exploitation for VSAM reads on this lpar.

Storage Class/Data Set Level Granularity

- Storage class definition is updated to allow clients to specify whether zHyperLink should be enabled or disabled

  - zHyperLink read and write capability can be enabled separately
  - Storage class attribute is examined on an I/O by I/O basis
  - Activating a new SMS configuration with a changed storage class definition will immediately affect all SMS data sets using the storage class

- VARY SMS command can be used to update the zHyperLink capability for a data set on the fly

  - Good for testing zHyperLink performance impact on a data set
  - Not hardened – will not be retained across an IPL
  - Can be used for non-SMS data sets and override storage class for SMS data sets

zHyperLink™ STORCLAS Enablement

- To enable zHyperLink™ Exploitation for data sets via STORCLAS:

```
SCDS Name . . . . . : D50.RLS.SCDS
Storage Class Name  : SXPXXS02

To ALTER Storage Class, Specify:

   Guaranteed Space  . . . . . . . . . . . . Y         (Y or N)
   Guaranteed Synchronous Write  . . .   N         (Y or N)
   Multi-Tiered SG . . . . . . . . . . . . . . .              (Y, N, or blank)
   Parallel Access Volume Capability. .   N        (R, P, S, or N)
   CF Cache Set Name . . . . . . . . . . .  CS2      (up to 8 chars or blank)
   CF Direct Weight  . . . . . . . . . . . .   1         (1 to 11 or blank)
   CF Sequential Weight  . . . . . . . . .   1         (1 to 11 or blank)
   CF Lock Set Name  . . . . . . . . . . .                (up to 8 chars or blank)
   Disconnect Sphere at CLOSE  . . . .  Y         (Y or N)
   zHyperLink Eligible for Read  . . . . N         (Y or N)
   zHyperLink Eligible for Write . . . . . . N         (Y or N)

Use ENTER to Perform Verification; Use UP Command to View previous Page;
Use HELP Command for Help; Use END Command to Save and Exit; CANCEL to Exit.
```

VSAM zHyperLink™ STORCLAS Enablement

- Displaying VSAM zHyperLink™ Exploitation in the STORCLAS:

CDS Name : ACTIVE

Enter Line Operators below:

LINE OPERATOR ---(1)----	STORCLAS NAME --(2)---	MULTI- TIERED SG --(19)---	PARALLEL ACCESS VOL ---(20)---	OAM SUBLVL -(21)-	CF LOCK SET NAME --(22)--	DISC SPHERE -(23)-	**ZHLINK** **READ** **-(24)-**	ZHLINK WRITE -(25)-
	ACCTDSN4	---	NOPREF	-	--------	NO	**YES**	NO

NO is the default.

Changing zHyperLink Eligibility of a Data Set

- V SMS,DSNAME(*datasetname*),ZHLREAD={YES|NO|DEFAULT}

  - Change the zHyperlink read eligibility for a data set or reset it to its default based on the storage class (SMS managed data sets). If DEFAULT is specified for a non-SMS data set, zHyperlink read eligibility is disabled.

- V SMS,DSNAME(*datasetname*),ZHLWRITE={YES|NO|DEFAULT}

  - Change the zHyperlink write eligibility for a data set or reset it to its default based on the storage class (SMS managed data sets). If DEFAULT is specified for a non-SMS data set, zHyperlink write eligibility is disabled.

- V SMS,DSNAME(*datasetname)*,DELETE

  - Reset the attributes to the storage class values (SMS managed data sets) or disable zHyperlink access (non-SMS data sets)

- Notes:

  - Vary applies to all current and future opens for this data set

  - The data set must be a VSAM data set

  - D SMS,DSNAME(dsn | ALL),ZHL can be used to display zHyperLink overrides by data set

z/OS zHyperLink Related Display Commands

- z/OS has display commands at the device, control unit, and link (PFID) level

- What do we mean by PFID?
  - Each zHyperLink port (link) is defined in the IBM z I/O configuration as a PCI "Function"
    - Each PCI function has a PCI function identifier or PFID, a physical channel id that represents the location of the zHyperLink card in the IBM Z, a port number, and access lists used to control which LPARs have access to the PFID
    - Up to 127 PFIDs may be defined for a single zHyperLink port

  - For each LPAR, define up to 4 PFIDs per zHyperLink port
    - The storage controller can handle up to 4 simultaneous requests on a single link
    - IBM Z and z/OS used the PFID to represent a particular I/O request on a link

  - FICON connectivity to each storage system is still required
    - Used for initialization, for I/Os that are not eligible for zHyperLink, and for failback when zHyperlink request fails (e.g., cache miss, busy condition)
    - You may continue to define up to 8 FICON CHPIDs per logical control unit (LCU)
    - Association between zHyperLinks and storage subsystem is discovered – no I/O configuration change required to tie a zHyperLink port with a DS8000

Display Device - D M=DEV(devno)

```
D M=DEV(8400)
IEE174I 16.10.51 DISPLAY M 652
DEVICE 08400    STATUS=ONLINE
CHP                       3D   4D   3E   4E   3F   4F
ENTRY LINK ADDRESS       17B7 1888 17A5 1897 17C9 18A3
DEST LINK ADDRESS        1750 1753 17D4 17D1 1752 17D3
PATH ONLINE              Y    Y    Y    Y    N    N
CHP PHYSICALLY ONLINE Y    Y    Y    Y    Y    Y
PATH OPERATIONAL         Y    Y    Y    Y    Y    Y
MANAGED                  N    N    N    N    N    N
CU NUMBER                8400 8400 8400 8400 8400 8400
INTERFACE ID            0501 0510 0401 0413 0320 0101
MAXIMUM MANAGED CHPID(S) ALLOWED:   0
DESTINATION CU LOGICAL ADDRESS = 50
SCP CU ND        = 002107.961.IBM.75.0000000BMN41.0502
SCP TOKEN NED    = 002107.900.IBM.75.0000000BMN41.5000
SCP DEVICE NED   = 002107.900.IBM.75.0000000BMN41.5000
WWNN             = world-wide-node-name
HYPERPAV ALIASES CONFIGURED = 32
ZHYPERLINKS AVAILABLE = 8
FUNCTIONS ENABLED = MIDAW, ZHPF, XPAV, ZHYPERLINK
```

Display Device - D M=DEV(devno)

- Interface ids
  - The interface ids are useful for determining whether CHPIDs share the same host adapter card (the first 3 digits represent the card)
  - If a D M=DEV(devno) command is issued and the particular path was never online to any device in the LCU, but other paths were online, the interface id will be displayed as "...."
  - If a D M=DEV(devno,(chpid)) command was issued in the above scenario, the interface id line would not be displayed
- WWNN
  - The WWNN will be displayed if the WWNN was read when a path first came online to the LCU
  - The WWNN is used by z/OS to tie the devices to the zHyperLink PFIDs connected to the DS8000
- zHyperlinks available
  - If there are no zHyperlinks available, this line will not be displayed regardless of whether the device is enabled to use zHyperlinks.
  - A link is considered available (usable) if the PFID is online, the port is operational, and z/OS has allocated the PFID
  - Links where the PFID is offline, the port is not operational, or where z/OS failed to allocate the link are not included
- Functions enabled
  - If the device is enabled for zHyperlink reads and writes, then ZHYPERLINK is displayed
  - If the device is enabled for zHyperlink reads only, then ZHYPERLINK(R) is displayed
  - If the device is enabled for zHyperlink writes only, then ZHYPERLINK(W) is displayed

Display Device zHyperLink Capability - D M=DEV(devno),ZHYPERLINK

- New option to display the zHyperLink capabilities and why a device is not enabled for zHyperLink
  - New message id: IEE587I
  - Not valid for PAV alias devices
  - ZHL can be used as an abbreviation for ZHYPERLINK
- If disabled entirely or disabled just for reads or writes, display reasons why it is disabled
  - Processor doesn't support synch I/O
  - Control unit doesn't support synch I/O
  - z/OS zHyperLink option turned off
  - No hyperlinks available
  - z/OS zHyperwrite option turned off
  - Secondaries (for synchronous mirroring) don't support synch I/O
- If a client doesn't see any zHyperLink activity in RMF, first issue this command to ensure that the device is at least enabled for zHyperLink

zHyperLink Disablement Reasons

- Processor reasons
  - `PROCESSOR DOES NOT SUPPORT ZHYPERLINK`

- z/OS reasons (IECIOSxx, SETIOS)
  - `ZHYPERLINK IS DISABLED FOR THE SYSTEM`
  - `ZHYPERLINK {READS | WRITES} ARE DISABLED FOR THE SYSTEM`
  - `HYPERWRITE IS DISABLED FOR THE SYSTEM` (Writes with metro mirror only)

- Control unit reasons
  - `CONTROL UNIT DOES NOT SUPPORT ZHYPERLINK`
  - `CONTROL UNIT DOES NOT SUPPORT ZHYPERLINK {READS | WRITES}`
  - `CONTROL UNIT WWNN COULD NOT BE RETRIEVED`
  - `CONTROL UNIT DOES NOT SUPPORT HYPERWRITE` (Writes with metro mirror only)

- zHyperlink reasons
  - `THERE ARE NO ZHYPERLINKS AVAILABLE`

- Device reasons (writes with metro mirror only)
  - `PPRC PRIMARY DEVICE IS NOT HYPERSWAP MANAGED`
    - `GDPS and CSM only load a hyperswap configuration when all volume states are duplex. A device will be in this state when the PPRC relationship is established until it goes duplex.`
    - `If a PPRC session is suspended, GDPS and CSM purge the hyperswap configuration.  This will also cause the device to be disabled for writes for this reason`
  - `PPRC SECONDARY DEVICE MISSING FROM THE HYPERSWAP CONFIGURATION`
    - `E.g., first leg of multi-target metro mirror relationship was loaded as part of a hyperswap configuration)`
  - `THERE ARE NO ZHYPERLINKS AVAILABLE FOR SECONDARY devno`
  - `CONTROL UNIT WWNN COULD NOT BE RETRIEVED FOR SECONDARY devno`
  - `CONTROL UNIT DOES NOT SUPPORT ZHYPERLINK WRITES FOR SECONDARY devno`
  - `CONTROL UNIT DOES NOT SUPPORT HYPERWRITE FOR SECONDARY devno`

D M=DEV(devno),ZHYPERLINK - Examples

Fully enabled

```
IEE587I 16.10.51 DISPLAY M 652
DEVICE 08400    STATUS=ONLINE
DEVICE IS ENABLED FOR ZHYPERLINK
READ AND WRITE OPERATIONS ARE ENABLED
ZHYPERLINKS AVAILABLE = 12
```

Enabled for reads only

```
IEE587I 16.10.51 DISPLAY M 652
DEVICE 08400    STATUS=ONLINE
DEVICE IS ENABLED FOR ZHYPERLINK
READ OPERATIONS ARE ENABLED
ZHYPERLINKS AVAILABLE = 12
WRITE OPERATIONS ARE DISABLED FOR THE FOLLOWING REASON(S):
  ZHYPERLINK WRITES ARE DISABLED FOR THE SYSTEM
  HYPERWRITE IS DISABLED FOR THE SYSTEM
```

D M=DEV(devno),ZHYPERLINK - Examples

- If zHyperLink I/O writes are disabled and this is a display for a primary device, we will include a list of secondary devices and their states, if conditions associated with the secondary devices also caused writes to be disabled.

```
IEE587I 16.10.51 DISPLAY M 652
DEVICE 08400    STATUS=ONLINE
DEVICE IS ENABLED FOR ZHYPERLINK
READ OPERATIONS ARE ENABLED
ZHYPERLINKS AVAILABLE = 12
WRITE OPERATIONS ARE DISABLED FOR THE FOLLOWING REASON(S):
  ZHYPERLINK WRITES ARE DISABLED FOR THE SYSTEM
  CONTROL UNIT DOES NOT SUPPORT ZHYPERLINK WRITES
  ZHYPERWRITE IS DISABLED FOR THE SYSTEM
  THERE ARE NO ZHYPERLINKS AVAILABLE FOR SECONDARY 10980
  CONTROL UNIT DOES NOT SUPPORT ZHYPERLINK WRITES FOR SECONDARY 10980
```

The first 3 lines are for the primary (the device specified in the command) or the system as a whole

The last 2 lines are for the secondary. Information for each secondary that is preventing writes from being enabled is shown.

Display Control Unit (LCU) – D M=CU(cunum)

```
D M=CU(700)
IEE174I 19.16.24 DISPLAY M 715
CONTROL UNIT 0700
CHP                       43   44   45   20   65   75
ENTRY LINK ADDRESS     B355 B353 B157 B051 ..   ..
DEST LINK ADDRESS      B37C B37D B17C B06B ..   ..
CHP PHYSICALLY ONLINE Y    Y    Y    Y    Y    Y

...
ZHPF - CU INTERFACE    Y    Y    Y    Y    Y    Y
INTERFACE ID           0501 0510 0401 0413 0320 0101
MAXIMUM MANAGED CHPID(S) ALLOWED = 0
DESTINATION CU LOGICAL ADDRESS = 07
CU ND             = 002107.000.IBM.PK.000000000003.0001
CU NED            = 002107.F20.IBM.PK.000000000003.0701
TOKEN NED         = 002107.900.IBM.PK.000000000003.0700
WWNN              = world-wide-node-name
FUNCTIONS ENABLED = ZHPF, ZHYPERLINK
DEFINED DEVICES
  00700,00710-00717
DEFINED PAV ALIASES
  00718,00719,10718
ZHYPERLINKS
  PFID     PCHID Port LinkId S/W St   Port St
  pfid     pchid  p    link  sssssss  pppppppp
  pfid     pchid  p    link  sssssss  pppppppp
```

Logical control unit or LSS level display

See next page for details

Only online PFIDs are shown. PFIDs are usable if software (S/W) state is "Alloc" and port state is "Oper" or "Degraded"

D M=CU Changes

- Interface ids
  - The interface ids are useful for determining whether CHPIDs share the same host adapter card (the first 3 digits represent the card)
  - Interface id will be displayed as "...." if the path was never online to any device.
- WWNN
  - The WWNN will be displayed if the WWNN was read when a path first came online to the LCU
- Functions enabled
  - If the device is enabled for zHyperLink reads and writes, then ZHYPERLINK is displayed
  - If the device is enabled for zHyperLink reads only, then ZHYPERLINK(R) is displayed
  - If the device is enabled for zHyperLink writes only, then ZHYPERLINK(W) is displayed

D M=CU Changes

- zHyperLink list
  - Only <u>online</u> PFIDs are shown regardless of whether the link is usable. Offline (standby) PFIDs are not included.
  - Processor side information includes PFID, PCHID and port
  - Control unit side information includes CU link id (similar to interface id)
  - Software states:
    - Alloc – The PFID is allocated and usable
    - Error – The PFID is in permanent error or port is not operational
    - RcvPend – Recovery is pending for the PFID
    - AllocEr – An error occurred attempting to allocate the PFID
    - DealEr – An error occurred attempting to deallocate the PFID
    - IntvReq – Intervention is required to make the PFID usable
  - Port states
    - Only high level states are shown. Issue D PCIE,PFID=pfid to get more details for not operational conditions
    - None – The state of the port is unknown
    - Oper – The port is operational
    - NotOper – The port is not operational
    - Degraded – The port is operational but running in degraded state (port state qualifier is 02 or 03)

Display PCI Functions – D PCIE

- Use these commands to display the state of the link if the D M=DEV or D M=CU display does not show the expected number of usable links

```
D PCIE
IQP022I hh.mm.ss DISPLAY PCIE
PCIE      0012 ACTIVE
PFID     DEVICE TYPE NAME        STATUS  ASID  JOBNAME  CHID  VFN  PN
00000001 8GB zHyperLink          ALLC    0013  IOSAS    034E  0001 1
00000002 8GB zHyperLink          ALLC    0013  IOSAS    034E  0002 2
00000099 Hardware Accelerator    ALLC    0011  FPGHWAM  0524  0002
```

- If Status is ALLC, then the PFID is usable
- If the PFID is online (CNFG), use D PCIE,PFID=pfid to get more details

```
D PCIE,PFID=1
IQP024I hh.mm.ss DISPLAY PCIE
PCIE      0012 ACTIVE
PFID     DEVICE TYPE NAME        STATUS  ASID  JOBNAME  CHID  VFN  PN
00000001 zHyperLink 8GB          ALLC    0013  IOSAS    034E  0001 1
 CLIENT ASIDS: None
 CU WWNN: world-wide-node-name  CU Link Id: linkid
 S/W State: swstate
 Port State: portstate
 CU Node Descriptor: devtype.model.mfg.plant.seqno
```

- Will show whether the port is operational, degraded, or not operational

Performance Statistics

z/OS collects performance statistics at multiple levels. Some of these show up in RMF reports and others appear only in SMF records

- Device level

- zHyperLink PFID level

- DASD subsystem link level

- Storage class and data set level

RMF Device Activity Report

```
                          D I R E C T   A C C E S S   D E V I C E   A C T I V I T Y

           z/OS V2R3              SYSTEM ID SYSF           DATE 04/04/2016           INTERVAL 05.04.737
                                  RPT VERSION V2R3 RMF      TIME 15.09.29            CYCLE 1.000 SECONDS

TOTAL SAMPLES =    304   IODF = 00   CR-DATE: 04/01/2016   CR-TIME: 07.22.21    ACT: ACTIVATE
                                          DEVICE  AVG   AVG   AVG   AVG   AVG   AVG   AVG   AVG    %
STORAGE   DEV   DEVICE   NUMBER  VOLUME PAV  LCU  ACTIVITY RESP IOSQ  CMR   DB    INT  PEND  DISC CONN   DEV
 GROUP    NUM    TYPE    OF CYL  SERIAL          RATE      TIME TIME  DLY   DLY   DLY  TIME  TIME TIME   CONN...
          02180 33909     10017  SYST10  1  001C  1.024S .384 .000  .128  .000  .128  .256 .000 .128   0.00...
          02181 33909     10017  SYST11  1  001C  0.073S .112 .000  .031  .000  .000  .000 .000 .000   0.00...
                          LCU            001C     1.097  .379 .000  .118  .000  .128  .256 .000 .128   0.00...
```

This part of the report continues to show only FICON and zHPF activity.
"S" means synch I/O requests were performed for this device, so synchronous I/O device activity
report should be examined.

RMF Device Activity Report

```
                          S Y N C H R O N O U S   I/O   D E V I C E   A C T I V I T Y

          z/OS V2R3                     SYSTEM ID SYSF              DATE 04/04/2016            INTERVAL 05.04.737
                                        RPT VERSION V2R3 RMF        TIME 15.09.29             CYCLE 1.000 SECONDS


 TOTAL SAMPLES =    304... CR-DATE: 04/01/2016   CR-TIME: 07.22.21    ACT: ACTIVATE
                              DEVICE ACTIVITY RATE -   -- AVG RESP TIME --   AVG SYNCH I/O  %          %      %      %
 STORAGE   DEV   ... LCU    -- SYNCH I/O --     ASYNCH   -SYNCH I/O - ASYNCH   TRANSFER RATE   REQ        LINK CACHE --REJECTS--
 GROUP     NUM   ...          READ    WRITE       I/O    READ  WRITE    I/O    READ    WRITE  SUCCESS   BUSY  MISS  READ WRITE
           02180 ... 001C    0.345   0.404     1.024   0.002  0.002   0.384   1.078   0.999    99.85   0.11  0.00  0.00 0.00
           02181 ... 001C    0.702   0.491     0.073   0.001  0.001   0.112   0.500   0.456    99.94   0.00  0.01  0.00 0.00
           LCU 001C          1.047   0.895     1.097   0.001  0.001   0.379   1.578   1.455    99.84   0.11  0.01  0.00 0.00
```

- Report shows comparison of synch and asynch activity rate (I/Os per second) and response time (ms)
- Transfer rate is in MB/sec
- Req Success – Percentage of zHyperLink requests that were successful
- Link Busy – Percentage of zHyperLink requests that failed because all available links were busy
- Cache Miss – Percentage of zHyperLink requests that failed because of a read or write cache miss
- Rejects – Percentage of zHyperLink requests that failed for other reasons (e.g., timeouts)
- LCU numbers for asynch activity only reflect those devices that had zHyperLink activity

- Read and write success, link busy and cache miss are combined into a single value in the report. However, the RMF SMF 74.1 record breaks these down into read and write values. Note that most likely a client won't have the DB2 log on the same volume as another data set using zHyperlink read.

RMF PCIE Activity Report

Shows both PFID and link (CEC) level activity

▼ Synchronous I/O Link Activity

Function ID ⇅	Function CHID ⇅	Port ID ⇅	Serial Number ⇅	Type-Model ⇅	Total Request Rate ⇅	Total Request Rate (CPC) ⇅	Successful Request % ⇅	Successful Request % (CPC) ⇅	Read Transfer Rate ⇅
0101	013C	1	0000000CF811	002017-932	0.1	0.57	0.083	44.64	1.67
0112	013C	1	0000000CF811	002017-932	0.035	0.23	0.033	17.86	0.33
0109	013C	2	0000000CC887	002017-932	0.075	0.27	0.067	35.71	0.033

Read and write transfer rates are in MB/sec for link

Read Transfer Rate (CPC) ⇅	Read Transfer Ratio ⇅	Read Transfer Ratio (CPC) ⇅	Write Transfer Rate ⇅	Write Transfer Rate (CPC) ⇅	Write Transfer Ratio ⇅	Write Transfer Ratio (CPC) ⇅	Time Busy % ⇅	Time Busy % (CPC) ⇅
0.625	16.67	0.57	8.93	1.33	13.33	7.14	0.83	0.588
0.125	3.33	0.14	1.79	0.33	3.33	1.79	0.167	0.188
0.0125	0.33	0.014	0.79	0.033	0.33	0.179	0.33	0.235

Read and write transfer ratios are in MB per I/O request

Busy time percentage

RMF ESS Synchronous I/O Link Statistics

```
                         E S S   S Y N C H R O N O U S   I / O   L I N K   S T A T I S T I C S


           z/OS V2R2                    SYSTEM ID MVSC              START 10/15/2018-15.06.00  INTERVAL 000.00.59
                                     CONVERTED TO z/OS V2R4 RMF          END   10/15/2018-15.07.00  CYCLE 0.200 SECONDS
SERIAL NUMBER  00000RA321        TYPE-MODEL  002107-985    CDATE 10/15/2018     CTIME 15.06.00    CINT 00.01.00
                                     ----CACHE READ OPERATIONS----   ----CACHE WRITE OPERATIONS---   -----NVS WRITE OPERATIONS--
--
SIID  ------LINK TYPE------      OPS    BYTES    RTIME   %SUCC      OPS    BYTES    RTIME   %SUCC      OPS    BYTES    RTIME
%SUCC
                                 /SEC    /OP     /OP             /SEC    /OP     /OP             /SEC    /OP     /OP
0080  Optical PCIe  GEN3  8    59391    4.1K   0.019   99.8   9230.0   49.2K   0.125   93.2   9228.8   45.9K   0.138 93.2
0180  Optical PCIe  GEN3  8    58522    4.1K   0.018   99.8   9466.4   49.2K   0.127   92.7   9467.6   45.6K   0.139 92.7
0181  Optical PCIe  GEN1  0    NO DATA TO REPORT OR ZERO
0380  Optical PCIe  GEN3  8    62416    4.1K   0.019   99.7   10992    49.2K   0.140   92.5   10986    45.5K   0.148 92.5
0580  Optical PCIe  GEN3  8    NO DATA TO REPORT OR ZERO
0680  Optical PCIe  GEN3  8    61317    4.1K   0.018   99.7   11630    49.2K   0.149   93.0   11635    45.8K   0.149 93.0
```

VSAM zHyperLink™ SMF Support

- The following new SMF 42-6 fields are available for zHyperlink reads:

  - S42SNERD   Eligible Synchronous I/O Reads
  - S42SNERH   Eligible Synchronous I/O Read Hits
  - S42SNRDT   No of sync_io read attempts.
  - S42SNROS   No of sync_io read successes.
  - S42SNSEQ   No of sync_io sequential operation
  - S42SNCND   No of sync_io cache candidates.
  - S42SNHTS   No of sync_io cache hits
  - S42SNRMS   No of sync_io read misses.
  - S42SNMXR   Max sync_io read response
  - S42SNRDU   Avg sync_io Read Time in usec
  - S42SNCONC   No of concurrent synch I/Os per req

- The zBNA and CP3KEXTR programs can be used to analyze zHyperLink performance. The latest versions that must be used are zBNA version 1.8.2 and CP3KEXTR version 3.75.

IBM Z Batch Network Analyzer

- IBM Z Batch Network Analyzer
  - A no charge, "as is" tool to analyze batch windows
  - Available to Customers, Business Partners and IBMers
  - PC based, and provides graphical and text reports
    - Including Gantt charts and support for Alternate Processors
- Available from IBM Z WSC CPS Tools team
  - IBMers: http://w3.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5126
  - BP's: https://www.ibm.com/partnerworld/wps/servlet/mem/ContentHandler/tech_PRS5133
  - Customers: http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5132
- zHyperLink Support
  - Top Data Set Candidate List for zHyperLink
  - Able to filter the data by time
  - Ability to estimate the benefit of zHyperLink activities
  - Generate customer reports with text and graphs to show zHyperLink benefit
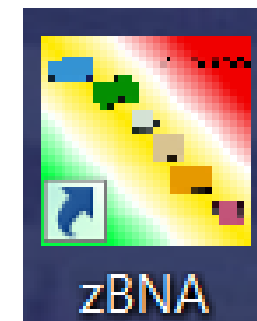
IBM Z Batch Network Analyzer (zBNA) Tool for zHyperLink is Available

https://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5132

or

Google:

IBM ZBNA DOWNLOAD

Reference Material

- [Getting Started with IBM zHyperLink for z/OS (redpaper)](#)
- z/OS MVS Initialization and Tuning Guide (SA23-1379)
  - IECIOSxx parmlib definition
- z/OS MVS System Commands (SA38-0666)
- z/OS MVS System Management Facilities (SMF) (SA38-0667)
  - SMF record formats including ones from RMF
- z/OS MVS System Messages (SA38-067 and SA38-0676) – IOS and display related messages
- z/OS MVS Diagnosis: Tools and Service Aids (GA32—0905) – System and GTF trace formats
- z/OS RMF Report Analyis (SC34-2664)

Questions