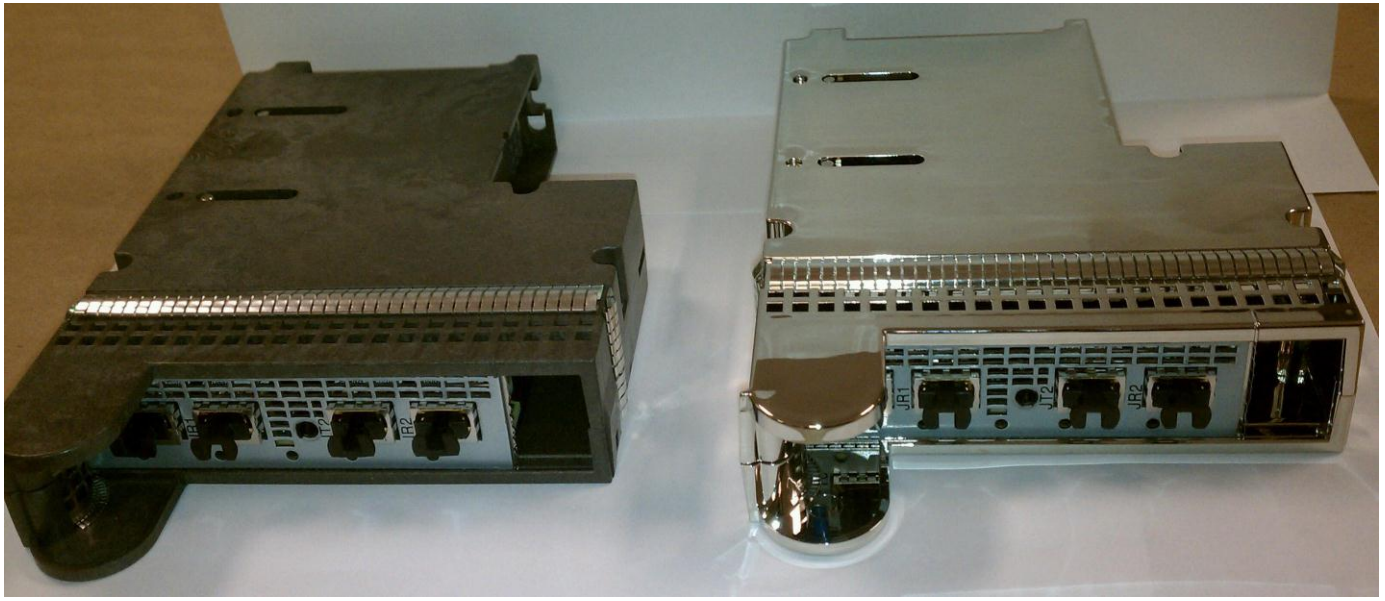
Migrating from ICB4 to Infiniband and Beyond



George Handera

HanderaG@Aetna.com

© 2012 Aetna

Migrating from ICB4 to Infiniband and Beyond

Trademarks

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

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

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

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

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

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

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

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

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

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

Migrating from ICB4 to Infiniband and Beyond

Agenda

- **Introduction/Background**
- **Implementation considerations z10 to z196**
  - **Consolidation**
  - **Performance**
- **Hardware feature comparison HCA2 to HCA3**
- **Experiences and observations**
- **z196 general performance**

Migrating from ICB4 to Infiniband and Beyond

Introduction/Background

George Handera - SE Capacity

Aetna has 2 datacenters - 25 miles apart

- 14 - 2817
- Each datacenter has a pair of external CF CECs
- One datacenter has 1 SYSPLEX over 8 CECs - heavy Datasharing
- Second datacenter 6 SYSPLEXES over 6 CECs - Datasharing load is light

Migrating from ICB4 to Infiniband and Beyond

SYSID	Mon	Day	SH	Hour	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L15P	L2LP	L2RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Ded/Shr
AE83	JUN	1	P	TOTAL	6.87	15.4	3.0	3.9	88	4.4	69.5	23.9	0.1	6.6	0.73	464.7	4.40	Ded
AE83	JUN	2	P	TOTAL	7.27	14.9	3.2	4.1	93	4.4	67.2	25.8	0.1	6.9	0.78	612.8	4.40	Ded
AE83	JUN	3	P	TOTAL	6.43	12.5	2.9	3.6	93	3.8	67.3	25.5	0.1	7.1	0.79	627.6	4.40	Ded
AE83	JUN	4	P	TOTAL	7.14	14.8	3.2	4.0	91	4.4	67.5	25.6	0.1	6.8	0.77	553.4	4.40	Ded
	Average				6.93	14.4	3.1	3.9	91	4.3	67.9	25.2	0.1	6.9	0.77	564.6		Ded
AE83	MAY	25	P	TOTAL	7.66	17.6	3.2	4.5	87	5.2	68.6	25.4	0.1	6.0	0.70	641.6	4.40	Shr
AE83	MAY	26	P	TOTAL	7.36	16.1	3.1	4.3	87	4.9	69.0	24.7	0.1	6.3	0.72	621.1	4.40	Shr
AE83	MAY	27	P	TOTAL	7.31	14.5	3.1	4.2	88	4.8	68.5	25.0	0.1	6.4	0.73	395.6	4.40	Shr
	Average				7.44	16.0	3.1	4.3	87	5.0	68.7	25.0	0.1	6.2	0.72	552.8		Shr
Dedicated / Shared					0.93	0.90	0.98	0.90	1.04	0.86	0.99	1.01	0.98	1.11	1.07	1.02		
					1.07	Relative ITR Capacity Ratio Of Dedicated Vs Shared												

CPI – Cycles per Instruction

Prb State - % Problem State

Est Instr Cmplx CPI – Estimated Instruction Complexity CPI (infinite L1)

Est Finite CPI – Estimated CPI from Finite cache/memory

Est SCPL1M – Estimated Sourcing Cycles per Level 1 Miss

L1MP – Level 1 Miss %

L15P – % sourced from Level 2 cache

L2LP – % sourced from Level 2 Local cache (on same book)

L2RP – % sourced from Level 2 Remote cache (on different book)

MEMP - % sourced from Memory

Rel Nest Intensity – Reflects distribution and latency of sourcing from shared caches and memory

LPARCPU - APPL% (GCPs, zAAPs, zIIPs) captured and uncaptured

Eff GHz – Effective gigahertz for GCPs, cycles per nanosecond

Migrating from ICB4 to Infiniband and Beyond

Implementation considerations z10 to z196

- ICB4 Links no longer supported on the z196 processor. Infiniband becomes the Link of choice.
  - HCA2 fanout cards were initially available and provided acceptable performance. Expectation was some degradation when converting from ICB4 technology.
  - HCA3 fanout cards were available Sept 10 2011 with significant performance improvements.
  - HCA3 fanout cards are compatible with HCA2.
- STP must be implemented before introducing a z196 processor.

Migrating from ICB4 to Infiniband and Beyond

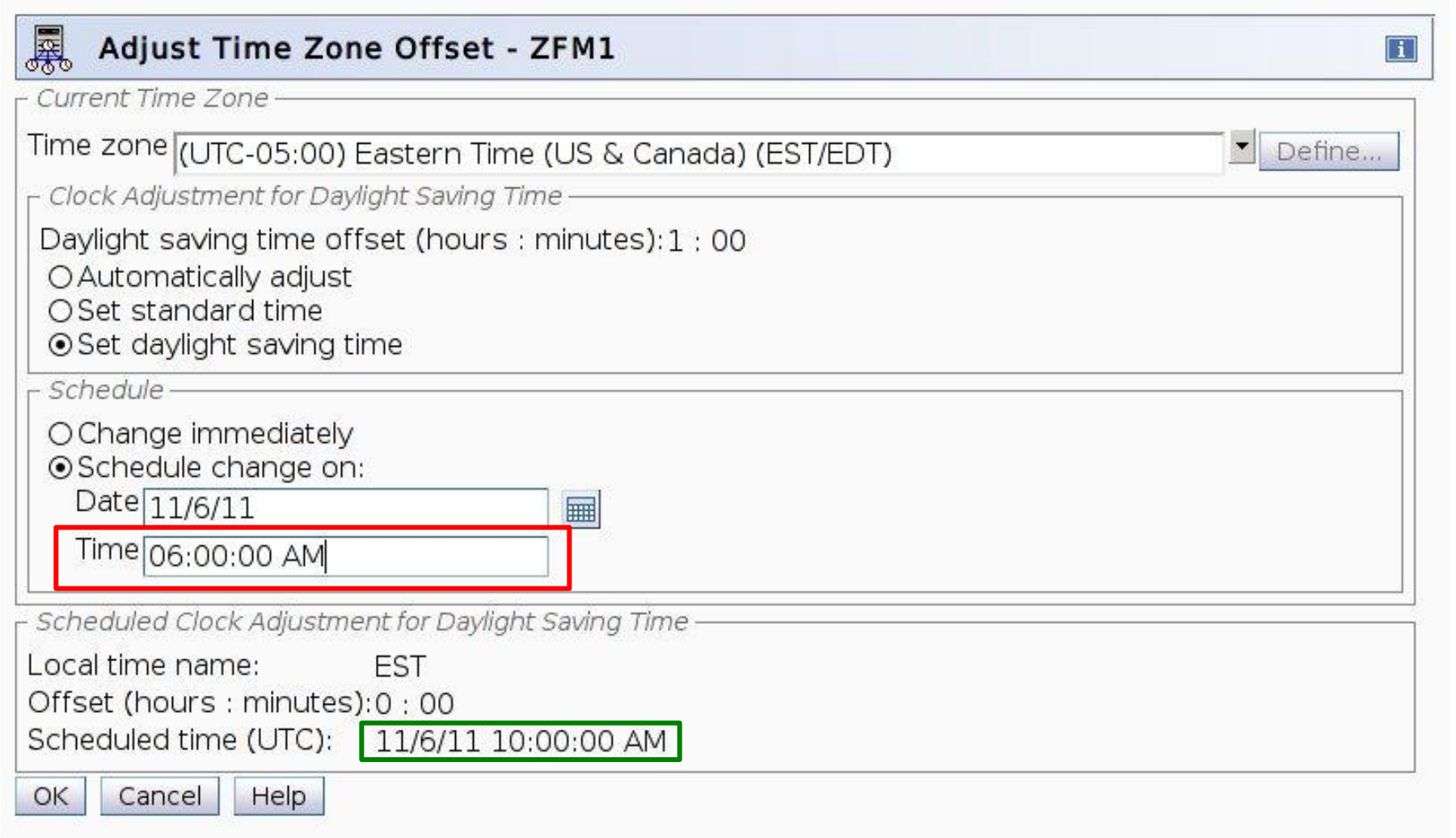
STP Configuration/Comments

- Time Source
  - We use a pair of HMC's as our NTP Servers
  - HMC's use a PPS server as their time source
- External CF CPCs are used for the PTS/CTS and BTS roles
  - This ensures that a maximum stratum level of 2 (CF links)
- Role player CPC's (PTS/BTS/Arbiter) must be connected (we use timing only links on the z/OS CPC's
- **STP uses UTC times only**, adjustments for zones or daylight savings time are artificial. If you schedule a DST adjustment for 10AM the user clocks adjusts but the UTC time show 10:00:01 one second later

NOTE: Driver 93 provides additional protection extending the inability to POR a CTS server to the other role players

Migrating from ICB4 to Infiniband and Beyond

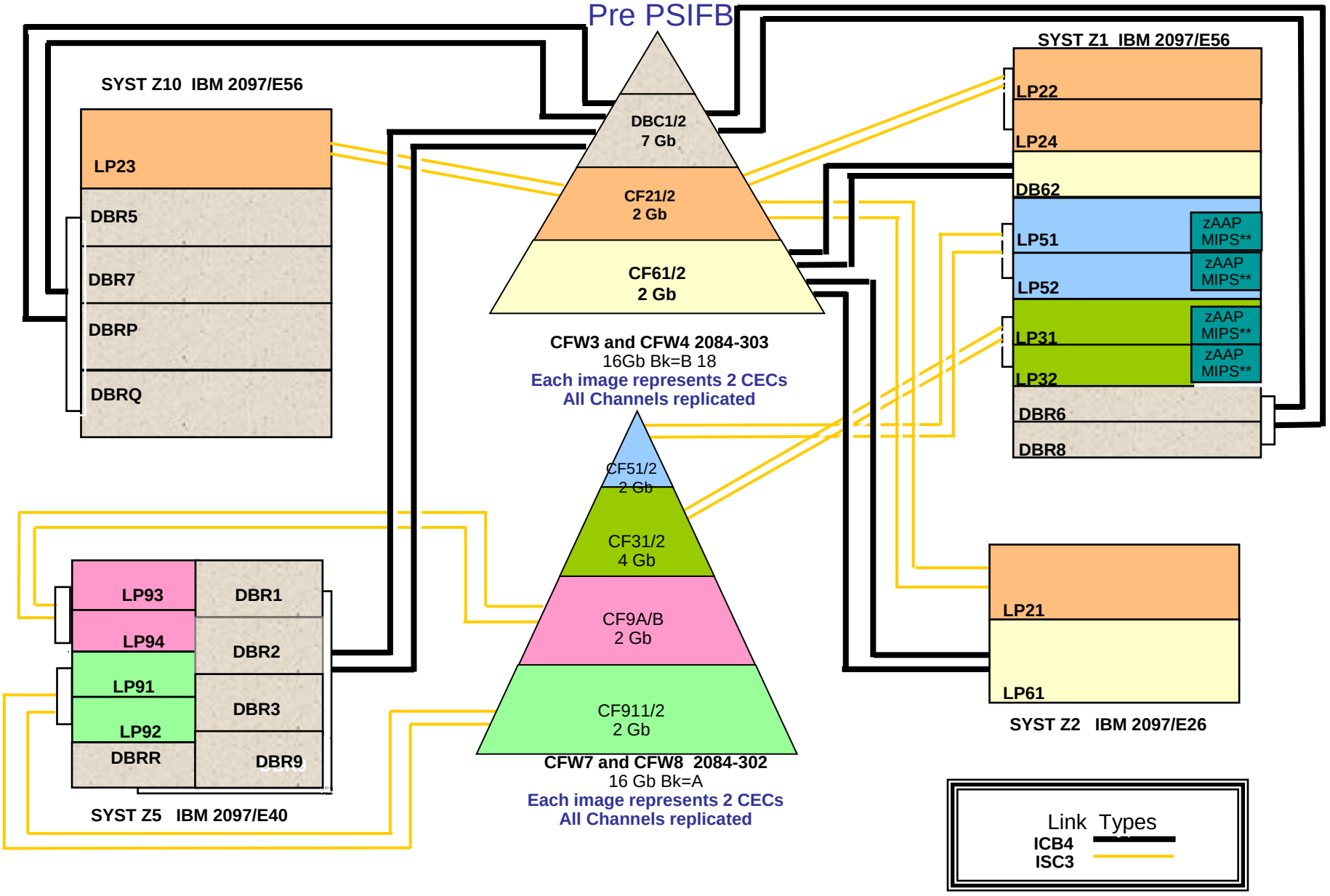
STP



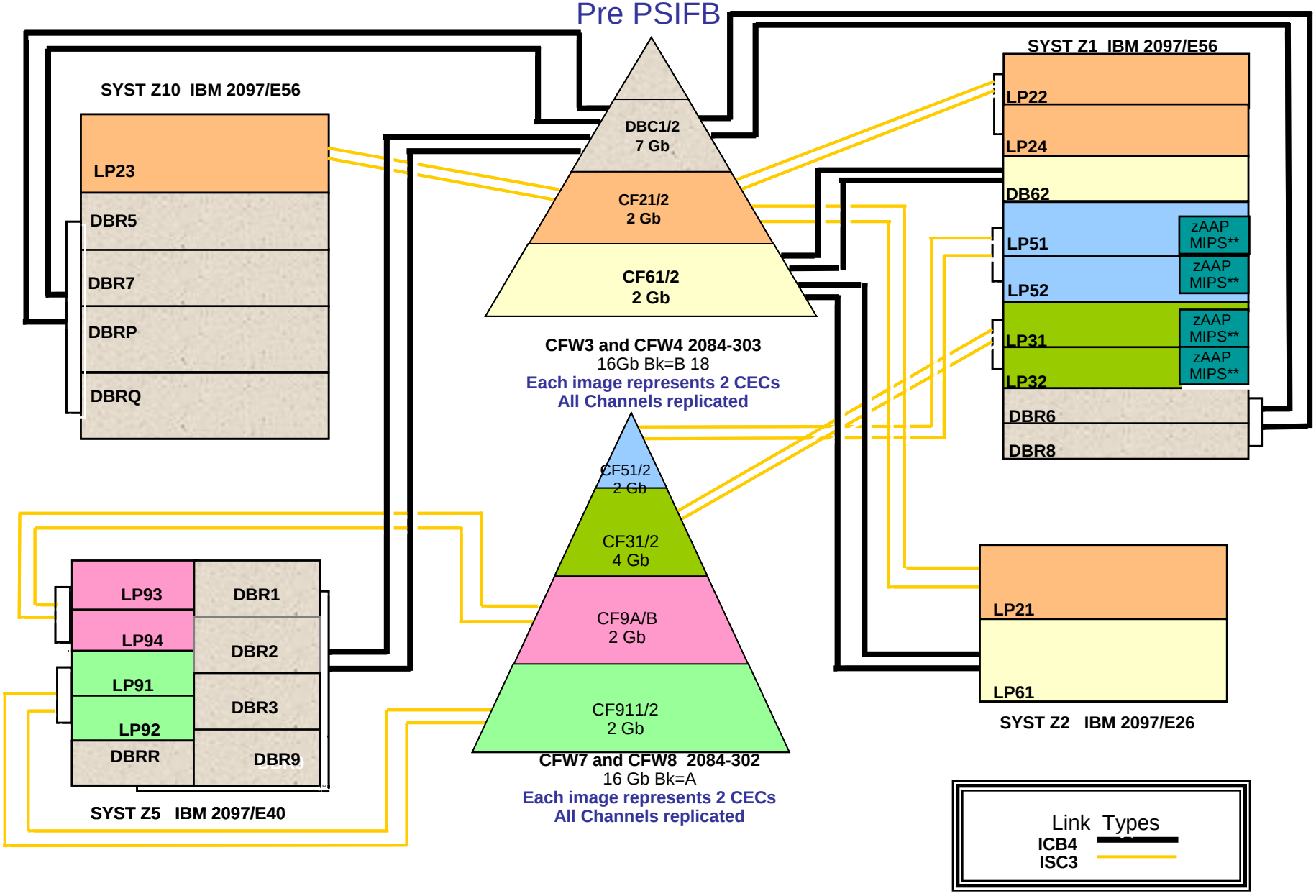
INFINIBAND

Consolidation and Performance

Windsor Computer Center
CPU Configuration
Pre PSIFB



Windsor Computer Center
CPU Configuration
Pre PSIFB



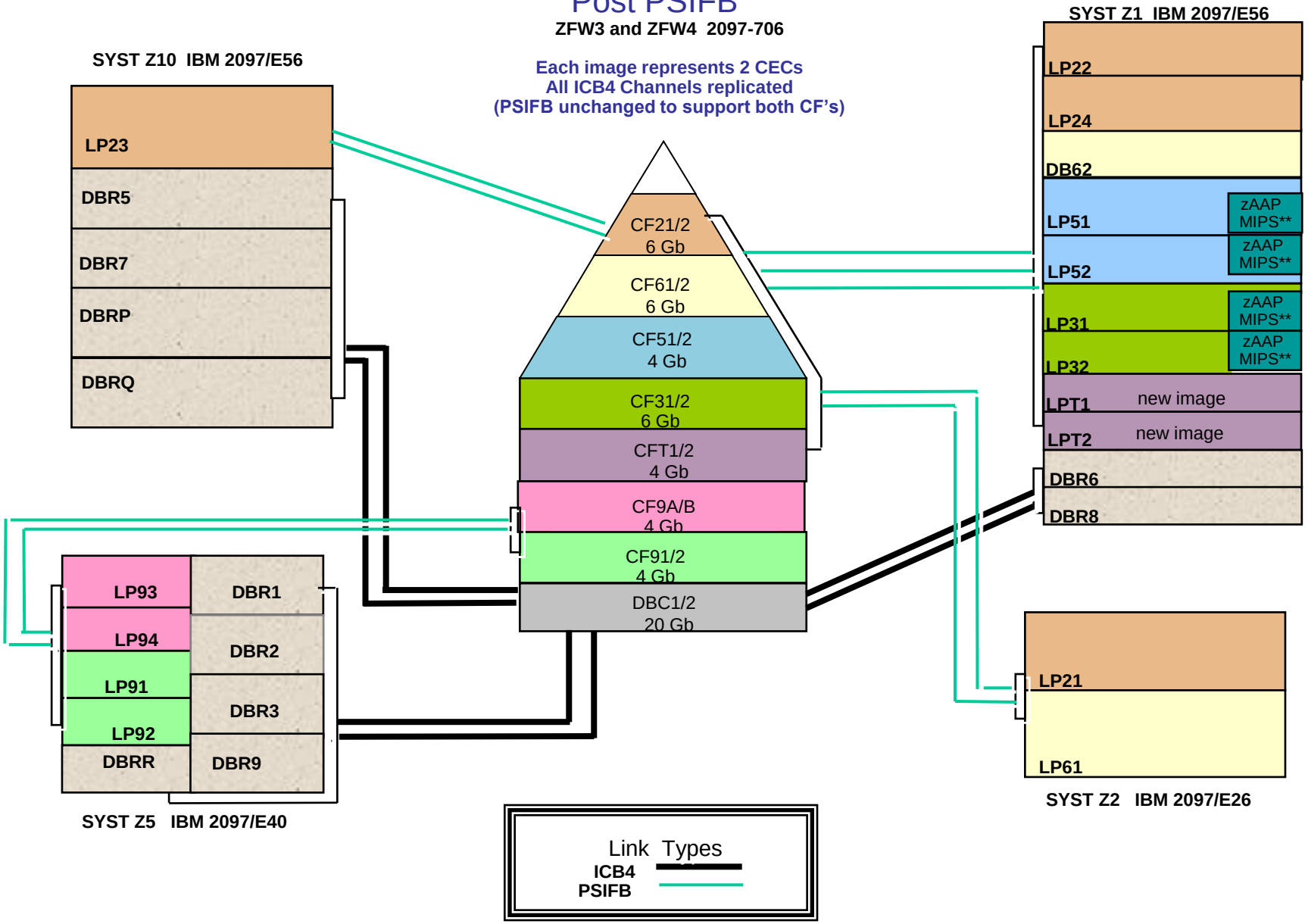
Windsor Computer Center

CPU Configuration

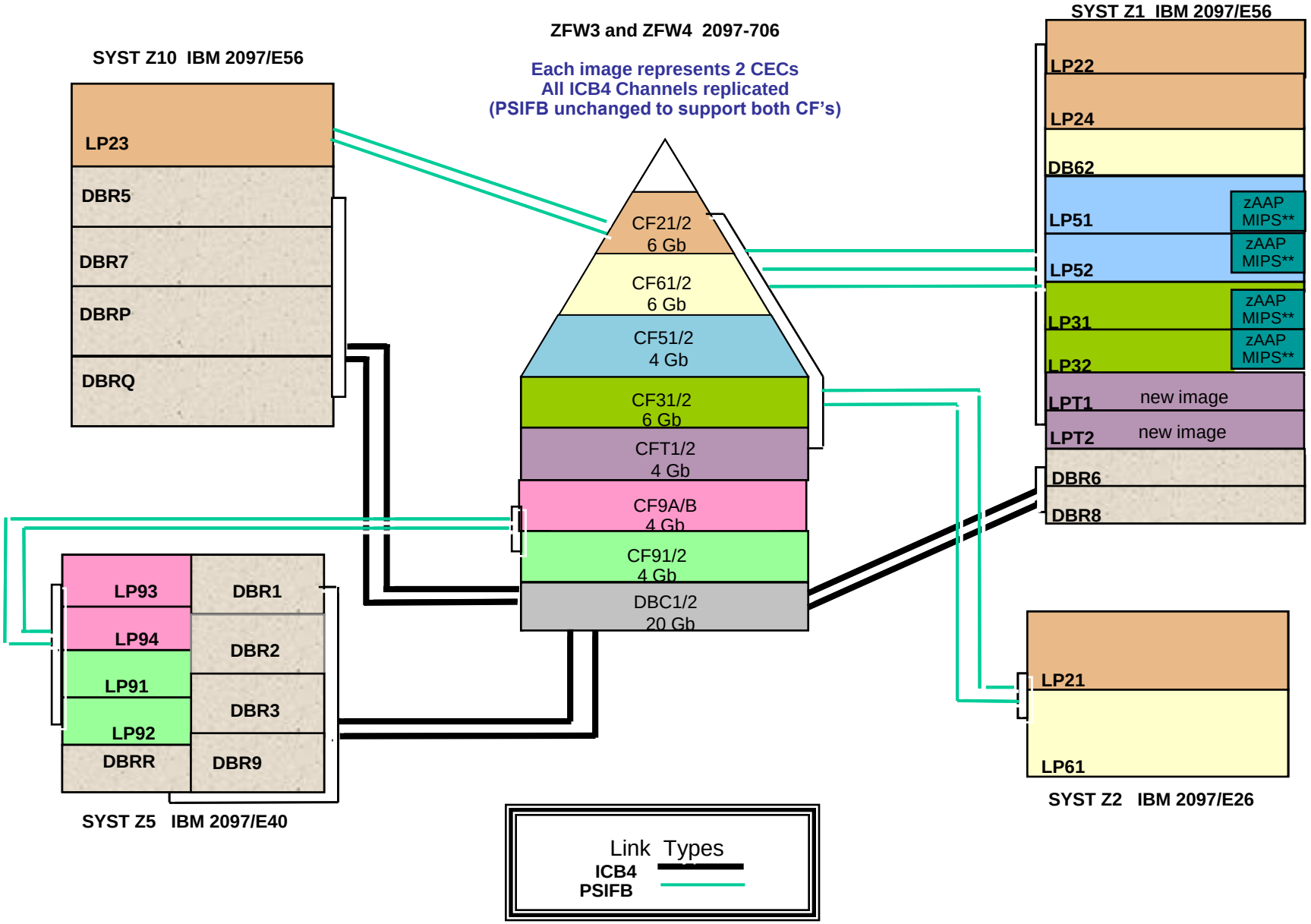
Post PSIFB

ZFW3 and ZFW4 2097-706

Each image represents 2 CECs
All ICB4 Channels replicated
(PSIFB unchanged to support both CF's)

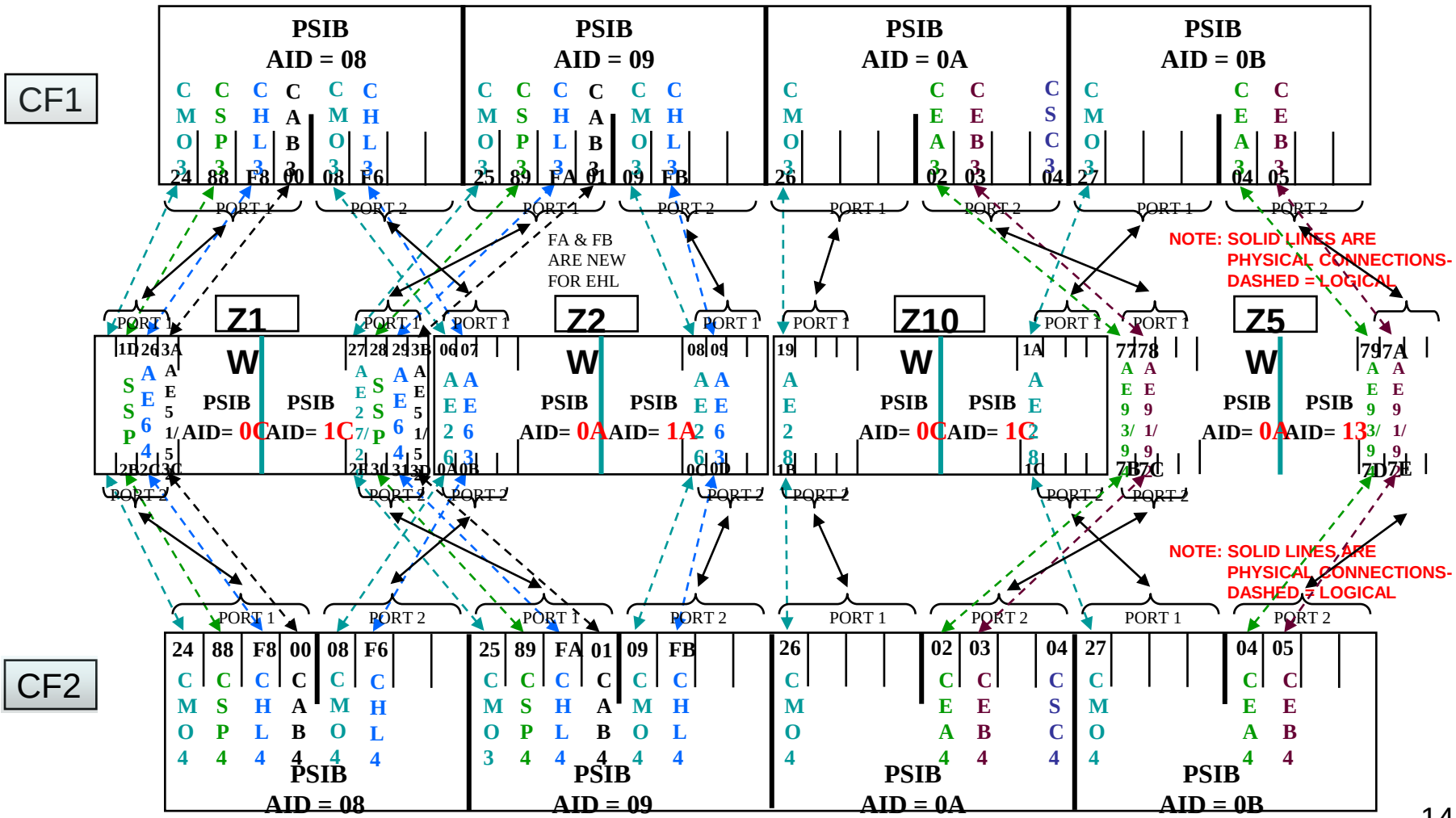


Windsor Computer Center
CPU Configuration
Post PSIFB



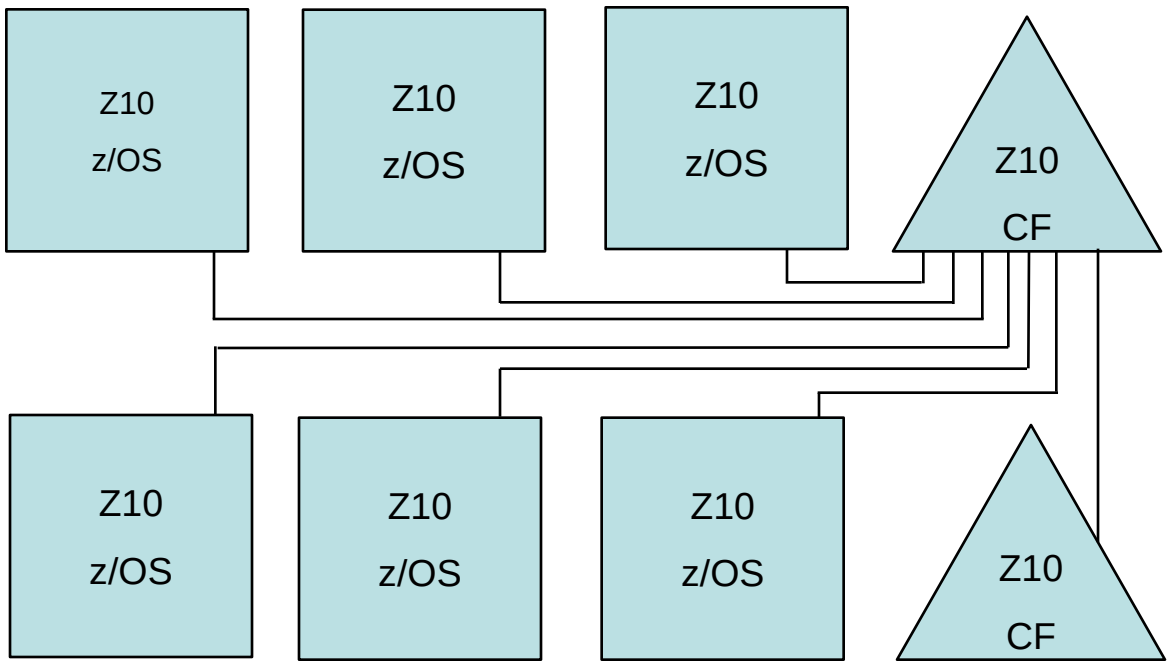
Migrating from ICB4 to Infiniband and Beyond

PSIFB Connectivity



Migrating from ICB4 to Infiniband and Beyond

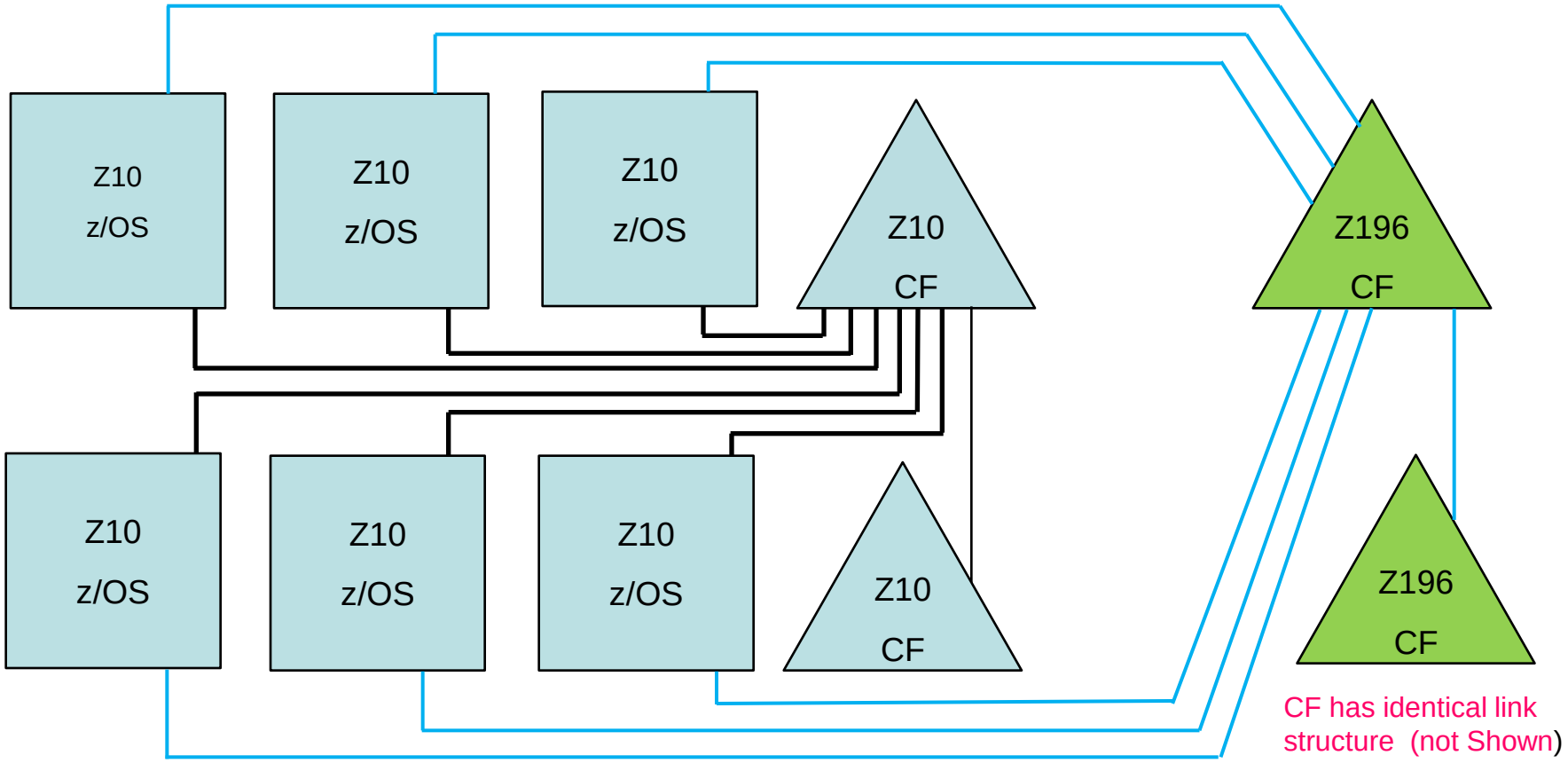
Pre Z196 - Z10 All ICB4 (1 SYSPLEX)



To prepare for roll of the floor the Z10 z/OS CPCs received HCA2-O PSIFB cards

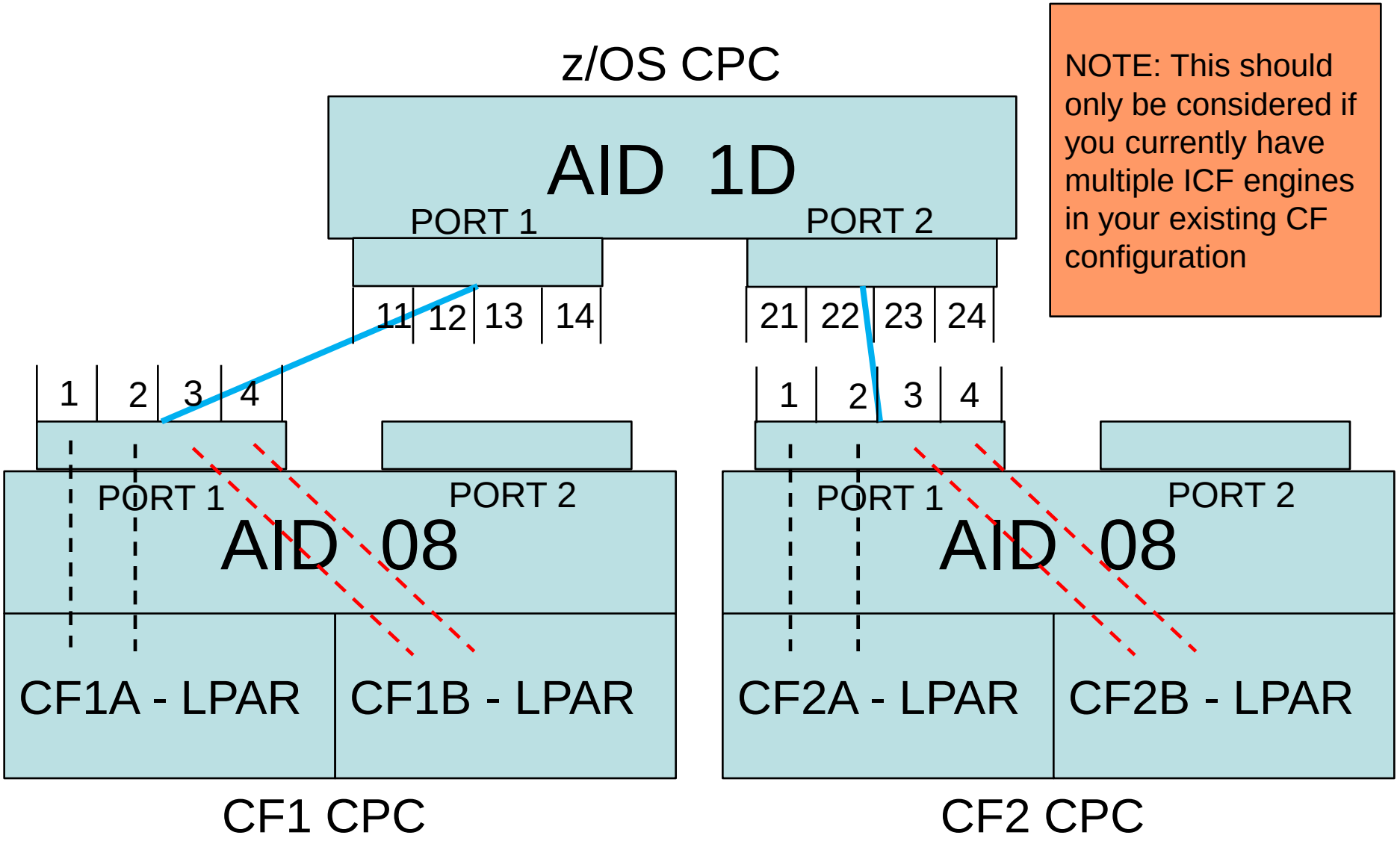
Migrating from ICB4 to Infiniband and Beyond

Z196 and PSIFB connectivity introduced to the SYSPLEX



New Z196 CF CPCs installed and PSIFB connectivity was implemented in addition to the Z10 ICB4 links.

Migrating from ICB4 to Infiniband and Beyond



NOTE: This should only be considered if you currently have multiple ICF engines in your existing CF configuration

One Cable is used to supply connectivity to 2 CF LPARs on the same CPC

Migrating from ICB4 to Infiniband and Beyond

Switching from ICB4 to PSIFB

- Alter the CFRM Policy to include the new Coupling Facility LPARs
- Place Z10 CF's in Maintmode
  SETXCF START,MAINTMODE,CFNAME=(CF1,CF2)
- Move the structures to the new Z196 CF's
  SETXCF START,REALLOCATE
- Verify structures moved
  D XCF,CF,CFNAME=(CF1,CF2)
- Observe
- To backoff
  SETXCF STOP,MAINTMODE,CFNAME=(CF1,CF2)
  SETXCF START,MAINTMODE,CFNAME=(CF1A,CF1B,CF2A,CF2B)
  SETXCF START,REALLOCATE

Migrating from ICB4 to Infiniband and Beyond

Results/Recommendations

- INFINIBAND performed better than our ICB4 environment / rec: at least 28 Subchannel Buffers for each LPAR. With ICB4 links I had bursts of activity that would overrun my subchannel buffers.
- CF Processor Utilization decreased dramatically due to a reduction in the MP effect and the increased cycle rate on the z196
- We went from 2 CF only CPCs each having an LPAR with 5 dedicated engines, to a configuration with 2 LPARs on each CPC…
- Each CPC had a 2 way and 3 way dedicated engine configuration.
- The 3 way is targeted with our “Loved ones” (SYNC)
  The 2 Way gets the ASYNC traffic.

Note: CFs only know they have a request. They do not know if the request is SYNC or ASYNC, but the z/OS LPARs do. Sharing SYNC and ASYNC requests tends to increase the ASYNC service times.

Migrating from ICB4 to Infiniband and Beyond

Samples: 60 Systems: 10 Date: 04/20/11 Time: 16.01.00 Range: 60 Sec

CF Policy: POLICY2 Activated at: 04/19/11 01.02.01

----- Coupling Facility ------					------ Processor -------				Request	- Storage --	
Name	Type	Model	Lvl	Dyn	Util%	Def	Shr Wgt	Effect	Rate	Size	Avail
CFM1	2097	E12	16	OFF	0.0	5	0	5.0		30G	30G
CFM2	2097	E12	16	OFF	0.0	5	0	5.0		30G	30G
CF1A	2817	M15	17	OFF	12.3	3	0	2.8	92662	30G	22G
CF1B	2817	M15	17	OFF	8.4	2	0	1.9	18312	22G	21G
CF2A	2817	M15	17	OFF	24.9	3	0	2.8	85490	30G	21G
CF2B	2817	M15	17	OFF	13.9	2	0	1.9	30850	22G	21G

Migrating from ICB4 to Infiniband and Beyond

Lock Structure Comparison - Peak Hour

CF Structure Name=DSNDB3G_LOCK1 Hour of Day=14

Week of Year	Day of Month	Reqs - Synch	Req Time - Synch	Reqs - Asynch	Req Time - Asynch	Defined Processors Utilization	Requests Changed from Synch to Asynch	AVESYTM	Reqs - Asynch	AVEASYTM	Requests Completed - Total
16	14	184834753	2292.5972	38804	3.418064	13.22 %	4469	.000012403	38804	.000088085	200046152
17	20	200053247	1941.6895	601	0.245176	8.64 %	351	.000009706	601	.000407947	212059129

Day	Tot Sync Requests	Tot Sync Time Seconds	Ave Sync Time
14	184,834,753	2292	12.4
20	200,053,247	1941	9.6

- Total and SYNC request rate increased
- Total SYNC service time decreased
- Average SYNC service time dramatically reduced

Migrating from ICB4 to Infiniband and Beyond

Summary of 12 hour Weekday Activity for SYNC and ASYNC Activity

Sysplex Name=AEPLEX04 Time Zone=1

Obs	Year of Century	Week of Year	Day of Month	_TYPE_	_FREQ_	Reqs - Synch	Req Time - Synch	Reqs - Asynch
1	11	16	13	0	4184	5403693742	79379.144	1646294832
2	11	17	19	0	4207	5885592760	69484.639	1665926063
ZONE						11289286502	148863.78	3312220895

Obs	Req Time - Asynch	Requests Changed from Synch to Asynch	Requests Completed - Total	AVESYTM	AVEASYTM
1	91693.984	2753375	7173079565	.000014690	.000055697
2	79962.040	841730	7791578304	.000011806	.000047999
ZONE	171656.02	3595105	14964657869		

Migrating from ICB4 to Infiniband and Beyond

Intro to HCA3

- Machine needs the GA2 - Driver 93
- Compatible with HCA2 Fanout card
- Runs Multiple protocols - HCA2 and IFB3
- Two card types HCA3-12x and HCA3-LR
- Comments
  - LR fanout cards may be defined with 7 or 32 subchannels. 32 should only be used for distance
  - LR fanout cards have 4 Ports, and use standard Ficon
  - The HCA3 - 12x cards use the same PSIFB cables as the older HCA2 12x cards.

Migrating from ICB4 to Infiniband and Beyond

HCA3 Test Results

- HCA2 Mode - CF on 2817
- HCA2 Mode - CF on 2097
- HCA3 MIX of LR and 12X Channel definitions
- HCA3 12X
- HCA3 LR comparison (a CF utilization story)
- Duplex

Activity driver requests were consistent for all tests

All activity was completion based (zero think time)

Migrating from ICB4 to Infiniband and Beyond

Test Environment

```
Samples: 1800    System: AE92  Date: 09/21/11  Time: 20.30.00  Range: 1800  Sec

Partition:   AE92        2817 Model 734
CPC Capacity:    3134    Weight % of Max: ****   4h Avg:  626   Group: Z15WCAP
Image Capacity:  3130    WLM Capping %:    0.0   4h Max: 1532   Limit:  3130

Partition  --- MSU ---  Cap  Proc    Logical Util %   - Physical Util % -
             Def   Act  Def   Num    Effect   Total   LPAR  Effect  Total

*CP                       38.0                        0.2    98.1   98.3
AE91           0  1525  NO   17.0      97.3    97.4    0.0    48.7   48.7
AE92           0  1523  NO   17.0      97.1    97.2    0.0    48.6   48.6
AE93           0    21  NO    2.0      10.9    11.1    0.0     0.6    0.7
AE94           0     9  NO    2.0       4.9     5.0    0.0     0.3    0.3
PHYSICAL                                                0.1            0.1

*IFL                       2.0                         0.8     0.7    1.5
AEVM                  NO   1.0       0.3     0.4     0.0     0.3    0.4
AEV2                  NO   1.0       0.4     0.5     0.1     0.4    0.5
PHYSICAL                                                0.7            0.7
```

Two 17 way 2817 processors were used to run a batch and CICS/DB2/MQ workload

Migrating from ICB4 to Infiniband and Beyond

HCA2 Performance - CF on 2817

Samples: 1800 Systems: 2 Date: 09/21/11 Time: 15.00.00 Range: 1800 Sec

CF Policy: POLICY2 Activated at: 09/14/11 17.40.19

----- Coupling Facility ------					------ Processor -------				Request	- Storage --	
Name	Type	Model	Lvl	Dyn	Util%	Def	Shr Wgt	Effect	Rate	Size	Avail
CFC4	2817	M49	17	OFF	39.6	3	0	3.0	362K	20G	18G
CFC5	2817	M49	17	OFF	0.0	2	0	2.0		20G	19G

Samples: 1800 Systems: 2 Date: 09/21/11 Time: 15.00.00 Range: 1800 Sec

CF: CFC4	Type	ST	System	CF	--- Sync ---		--------- Async --------			
				Util	Rate	Avg	Rate	Avg	Chng	Del
Structure Name				%		Serv		Serv	%	%
DSNDBMG_GBP20	CACHE	A	*ALL	18.5	16846	18	3771	141	0.0	0.0
	CACHE		AE91		8165	17	1456	161	0.0	0.0
	CACHE		AE92		8682	19	2315	129	0.0	0.0
DSNDBMG_GBP21	CACHE	A	*ALL	6.0	9652	15	1077	111	0.0	0.0
	CACHE		AE91		4880	15	380.3	130	0.0	0.0
	CACHE		AE92		4772	15	696.8	100	0.0	0.0
DSNDBMG_GBP22	CACHE	A	*ALL	6.5	10449	15	1182	110	0.0	0.0
	CACHE		AE91		5301	15	428.4	128	0.0	0.0
	CACHE		AE92		5147	15	753.5	100	0.0	0.0
DSNDBMG_LOCK1	LOCK	A	*ALL	69.0	319K	11	477.1	55	0.1	0.1
	LOCK		AE91		147K	11	129.4	59	0.1	0.1
	LOCK		AE92		172K	11	347.7	53	0.0	0.1

Migrating from ICB4 to Infiniband and Beyond

HCA2 Performance - CF on 2097

Samples: 1800 Systems: 2 Date: 09/21/11 Time: 20.30.00 Range: 1800 Sec

CF Policy: POLICY2 Activated at: 09/14/11 17.40.19

----- Coupling Facility ------					------ Processor -------				Request	- Storage --	
Name	Type	Model	Lvl	Dyn	Util%	Def	Shr Wgt	Effect	Rate	Size	Avail
CFE1	2097	E40	16	OFF	54.4	3	0	3.0	360K	7841M	6396M
CFE2	2097	E40	16	OFF	0.0	2	0	2.0	3.7	7841M	7633M

Samples: 1800 Systems: 2 Date: 09/21/11 Time: 20.30.00 Range: 1800 Sec

CF: CFE1	Type	ST	System	CF	--- Sync ---		--------- Async --------			
Structure Name				Util %	Rate	Avg Serv	Rate	Avg Serv	Chng %	Del %
DSNDBMG_GBP20	CACHE	A	*ALL	21.8	16304	22	4264	139	0.0	0.0
	CACHE		AE91		7964	21	1652	155	0.0	0.0
	CACHE		AE92		8341	23	2612	129	0.0	0.0
DSNDBMG_GBP21	CACHE	A	*ALL	6.6	9159	18	1284	109	0.0	0.0
	CACHE		AE91		4633	19	579.3	117	0.0	0.0
	CACHE		AE92		4526	18	704.3	103	0.0	0.0
DSNDBMG_GBP22	CACHE	A	*ALL	6.8	9303	18	1368	105	0.0	0.0
	CACHE		AE91		4731	19	619.6	112	0.0	0.0
	CACHE		AE92		4572	18	748.1	100	0.0	0.0
DSNDBMG_LOCK1	LOCK	A	*ALL	64.8	315K	13	1494	62	0.1	0.1
	LOCK		AE91		150K	13	756.5	62	0.1	0.1
	LOCK		AE92		165K	13	737.8	62	0.2	0.2

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - LR and 12X IFB3

Samples: 1800 Systems: 2 Date: 09/26/11 Time: 12.00.00 Range: 1800 Sec

CF Policy: POLICY2 Activated at: 09/14/11 17.40.19

----- Coupling Facility ------					------ Processor -------					Request	- Storage --	
Name	Type	Model	Lvl	Dyn	Util%	Def	Shr	Wgt	Effect	Rate	Size	Avail
CFC4	2817	M49	17	OFF	35.3	3	0		3.0	351K	20G	18G
CFC5	2817	M49	17	OFF	0.0	2	0		2.0		20G	19G

Samples: 1800 Systems: 2 Date: 09/26/11 Time: 12.00.00 Range: 1800 Sec

CF: CFC4	Type	ST	System	CF	--- Sync ---		--------- Async --------			
				Util	Rate	Avg	Rate	Avg	Chng	Del
Structure Name				%		Serv		Serv	%	%
DSNDBMG_GBP20	CACHE	A	*ALL	19.0	22792	17	3187	141	0.0	0.0
	CACHE		AE91		11396	17	1589	142	0.0	0.0
	CACHE		AE92		11395	17	1597	141	0.0	0.0
DSNDBMG_GBP21	CACHE	A	*ALL	6.5	9597	16	991.0	128	0.0	0.0
	CACHE		AE91		4800	16	520.2	128	0.0	0.0
	CACHE		AE92		4797	16	470.8	128	0.0	0.0
DSNDBMG_GBP22	CACHE	A	*ALL	6.7	9781	16	1032	127	0.0	0.0
	CACHE		AE91		4887	16	530.3	126	0.0	0.0
	CACHE		AE92		4894	16	501.2	127	0.0	0.0
DSNDBMG_LOCK1	LOCK	A	*ALL	67.8	303K	9	625.0	49	0.0	0.0
	LOCK		AE91		144K	9	302.8	49	0.0	0.0
	LOCK		AE92		159K	9	322.2	49	0.0	0.0

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - LR and 12X IFB3

```
                RMF Coupling Facility - Subchannels and Paths

 Details for System      : AE91
 Coupling Facility       : CFC4

 Subchannels  Generated  :  70
              In Use     :  70
              Max        : 156

 Path IDs    : C1    C3    C5    C7    14    15    0D    0E
       Types : CIB   CIB   CIB   CIB   CIB   CIB   CIB   CIB

 Press Enter to return to the Report panel.
```

- CHPIDs are used in the sequence shown
- When all subchannels are used the on a CHPID, the next CHPID is used. ***TYPE MATTERS!!***

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - 12X IFB3

Samples: 1800 Systems: 2 Date: 09/26/11 Time: 13.00.00 Range: 1800 Sec

CF Policy: POLICY2 Activated at: 09/14/11 17.40.19

----- Coupling Facility ------					------ Processor -------					Request	- Storage --	
Name	Type	Model	Lvl	Dyn	Util%	Def	Shr	Wgt	Effect	Rate	Size	Avail
CFC4	2817	M49	17	OFF	33.7	3	0		3.0	399K	20G	18G
CFC5	2817	M49	17	OFF	0.0	2	0		2.0		20G	19G

Samples: 1800 Systems: 2 Date: 09/26/11 Time: 13.00.00 Range: 1800 Sec

CF: CFC4	Type	ST	System	CF	--- Sync ---		--------- Async --------			
				Util	Rate	Avg	Rate	Avg	Chng	Del
Structure Name				%		Serv		Serv	%	%
DSNDBMG_GBP20	CACHE	A	*ALL	21.1	30053	8	2827	96	0.0	0.0
	CACHE		AE91		15002	8	1398	96	0.0	0.0
	CACHE		AE92		15051	8	1428	95	0.0	0.0
DSNDBMG_GBP21	CACHE	A	*ALL	6.7	11872	7	728.0	89	0.0	0.0
	CACHE		AE91		5943	7	328.7	92	0.0	0.0
	CACHE		AE92		5929	7	399.3	87	0.0	0.0
DSNDBMG_GBP22	CACHE	A	*ALL	6.7	11516	7	727.3	89	0.0	0.0
	CACHE		AE91		5750	7	331.1	91	0.0	0.0
	CACHE		AE92		5767	8	396.3	86	0.0	0.0
DSNDBMG_LOCK1	LOCK	A	*ALL	65.5	342K	5	88.5	43	0.0	0.0
	LOCK		AE91		162K	5	50.6	43	0.0	0.0
	LOCK		AE92		180K	5	37.9	42	0.0	0.0

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - 12X IFB3 - 12 Hour Summary

Sysplex Name=AEPLEX04 Time Zone=1

Obs	Year of Century	Week of Year	Day of Month	_TYPE_	_FREQ_	Reqs - Synch	Req Time - Synch	Reqs - Asynch
1	11	47	15	0	4321	6463998487	76191.654	2411328715
2	11	48	21	0	4296	6003154530	32898.516	2246209316
ZONE						12467153017	109090.17	4657538031

Obs	Req Time - Asynch	Requests Changed from Synch to Asynch	Requests Completed - Total	AVESYTM	AVEASYTM
1	103515.26	1339223	9136090593	.000011787	.000042929
2	78676.353	1109901	8527457968	.000005480	.000035026

Day	Tot Sync Requests	Tot Sync Time Seconds	Ave Sync Time
15	6,463,998,487	76,191	11.8
21	6,003,154,530	32,892	5.5

Saved almost an hour of CPU per hour - resource and MSU's

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - LR

```
Samples: 1800    Systems: 2    Date: 09/26/11  Time: 14.30.00  Range: 1800  Sec

CF Policy: POLICY2     Activated at: 09/14/11 17.40.19

----- Coupling Facility ------  ------ Processor -------  Request  - Storage --
Name      Type  Model Lvl  Dyn  Util% Def Shr Wgt Effect  Rate     Size   Avail

CFC4      2817   M49  17   OFF  36.3    3   0       3.0    332K     20G    18G
CFC5      2817   M49  17   OFF   0.0    2   0       2.0           20G    19G
```

```
Samples: 1800    Systems: 2    Date: 09/26/11  Time: 14.30.00  Range: 1800  Sec

CF: CFC4          Type  ST System   CF     --- Sync ---    --------- Async --------
                                     Util   Rate    Avg     Rate    Avg    Chng   Del
Structure Name                        %             Serv            Serv    %      %

DSNDBMG_GBP20     CACHE A  *ALL      18.9   20416     23     3953     171    0.0    0.0
                  CACHE    AE91             10222     23     1935     172    0.0    0.0
                  CACHE    AE92             10194     23     2018     169    0.0    0.0
DSNDBMG_GBP21     CACHE A  *ALL       6.1    8112     20     1256     145    0.0    0.0
                  CACHE    AE91              4065     20    636.9     146    0.0    0.0
                  CACHE    AE92              4047     21    619.5     144    0.0    0.0
DSNDBMG_GBP22     CACHE A  *ALL       6.3    8202     21     1277     152    0.0    0.0
                  CACHE    AE91              4105     21    630.0     154    0.0    0.0
                  CACHE    AE92              4096     21    646.6     151    0.0    0.0
DSNDBMG_LOCK1     LOCK  A  *ALL      68.7    287K     12     1258      57    0.0    0.0
                  LOCK     AE91              138K     12    624.6      57    0.0    0.0
                  LOCK     AE92              149K     12    633.3      57    0.0    0.0
```

Migrating from ICB4 to Infiniband and Beyond

Duplex Test Statement

PLEASE NOTE

All Duplex testing was performed with ICP (memory to memory) links. Results are provided to demonstrate the impact of implementing SYSTEM MANAGED DUPLEX.

SYSTEM MANAGED DUPLEX is avoided at our shop and due to the number of physical cards had available for testing I made the choice to use ICP links.

External CF's allow the DB2 lock structure to be recovered from IRLM virtual storage, eliminating the need for duplexing the structure.

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - HCA3 IFB3 - Duplex (primary)

```
Samples: 1799    Systems: 2    Date: 10/06/11  Time: 12.00.00  Range: 1800  Sec

CF Policy: POLICY2     Activated at: 09/14/11 17.40.19

----- Coupling Facility ------  ------ Processor -------  Request  - Storage --
Name      Type  Model Lvl  Dyn  Util% Def Shr Wgt Effect  Rate     Size   Avail

CFC4      2817    M49  17  OFF  57.2    4   0      4.0     231K     20G    18G
CFC5      2817    M49  17  OFF  40.4    4   0      4.0     203K     20G    18G
```

Duplex ⇧ ⇩ Simplex

```
Samples: 1800    Systems: 2    Date: 09/26/11  Time: 13.00.00  Range: 1800  Sec

CF Policy: POLICY2     Activated at: 09/14/11 17.40.19

----- Coupling Facility ------  ------ Processor -------  Request  - Storage --
Name      Type  Model Lvl  Dyn  Util% Def Shr Wgt Effect  Rate     Size   Avail

CFC4      2817    M49  17  OFF  33.7    3   0      3.0     399K     20G    18G
CFC5      2817    M49  17  OFF   0.0    2   0      2.0              20G    19G
```

- Utilization 57.2% of **4** engines to do 231K primary duplex requests
- Only 33.7% utilization on **3** engines to do 399K simplex requests

IMPORTANT - ALL CF to CF duplex Communication was performed on ICP Links providing better performance than Physical links

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - HCA3 IFB3 - Duplex (primary)

CF: CFC4 Structure Name	Type	ST	System	CF Util %	--- Sync --- Rate	Avg Serv	--------- Async --------- Rate	Avg Serv	Chng %	Del %
DSNDBMG_GBP20	CACHE	AP	*ALL	7.9	17685	12	2208	116	0.7	0.7
	CACHE		AE91		8869	12	1079	117	0.5	0.5
	CACHE		AE92		8816	12	1129	115	0.9	0.9
DSNDBMG_GBP21	CACHE	AP	*ALL	1.9	2615	16	532.8	118	0.3	0.3
	CACHE		AE91		1070	13	223.3	127	0.2	0.2
	CACHE		AE92		1546	19	309.5	112	0.5	0.5
DSNDBMG_GBP22	CACHE	AP	*ALL	2.7	7223	11	659.2	106	0.3	0.3
	CACHE		AE91		3636	11	326.7	108	0.2	0.2
	CACHE		AE92		3587	11	332.6	105	0.5	0.5
DSNDBMG_LOCK1	LOCK	AP	*ALL	87.5	75521	33	124K	51	0.2	0.2
	LOCK		AE91		36931	33	59346	51	0.1	0.1
	LOCK		AE92		38589	33	64432	51	0.3	0.3

Note: Service time on software managed duplex is still within the Sync threshold of 26 Mics, but the critical lock structure has largely turn Async due to system managed duplex

Migrating from ICB4 to Infiniband and Beyond

Flashback to HCA3 Performance - 12X IFB3 - SIMPLEX

Samples: 1800 Systems: 2 Date: 09/26/11 Time: 13.00.00 Range: 1800 Sec

CF: CFC4 Structure Name	Type	ST	System	CF Util %	Sync Rate	Sync Avg Serv	Async Rate	Async Avg Serv	Async Chng %	Async Del %
DSNDBMG_GBP20	CACHE	A	*ALL	21.1	30053	8	2827	96	0.0	0.0
	CACHE		AE91		15002	8	1398	96	0.0	0.0
	CACHE		AE92		15051	8	1428	95	0.0	0.0
DSNDBMG_GBP21	CACHE	A	*ALL	6.7	11872	7	728.0	89	0.0	0.0
	CACHE		AE91		5943	7	328.7	92	0.0	0.0
	CACHE		AE92		5929	7	399.3	87	0.0	0.0
DSNDBMG_GBP22	CACHE	A	*ALL	6.7	11516	7	727.3	89	0.0	0.0
	CACHE		AE91		5750	7	331.1	91	0.0	0.0
	CACHE		AE92		5767	8	396.3	86	0.0	0.0
DSNDBMG_LOCK1	LOCK	A	*ALL	65.5	342K	5	88.5	43	0.0	0.0
	LOCK		AE91		162K	5	50.6	43	0.0	0.0
	LOCK		AE92		180K	5	37.9	42	0.0	0.0

- Duplex sync/async rate was 75K/124K total of 200K
- Duplex sync service time was 33 mics

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - HCA3 IFB3 - Duplex (secondary)

----- Coupling Facility ------					------ Processor -------				Request	- Storage --	
Name	Type	Model	Lvl	Dyn	Util%	Def	Shr	Wgt	Effect	Rate	Size	Avail
CFC5	2817	M49	17	OFF	40.4	4		0	4.0		203K	20G

Samples: 1799 Systems: 2 Date: 10/06/11 Time: 12.00.00 Range: 1800 Sec

CF: CFC5	Type	ST	System	CF Util %	--- Sync --- Rate	Avg Serv	--------- Async -------- Rate	Avg Serv	Chng %	Del %
Structure Name										
DSNDBMG_GBP20	CACHE	AS	*ALL	3.6	0.1	8	2280	65	0.0	0.3
	CACHE		AE91		0.0	0	1135	66	0.0	0.3
	CACHE		AE92		0.1	8	1145	65	0.0	0.4
DSNDBMG_GBP21	CACHE	AS	*ALL	1.0	0.1	8	482.0	70	0.0	0.1
	CACHE		AE91		0.0	0	222.6	68	0.0	0.1
	CACHE		AE92		0.1	8	259.4	71	0.0	0.2
DSNDBMG_GBP22	CACHE	AS	*ALL	1.2	0.1	8	621.4	67	0.0	0.1
	CACHE		AE91		0.0	0	328.9	68	0.0	0.1
	CACHE		AE92		0.1	8	292.5	67	0.0	0.2
DSNDBMG_LOCK1	LOCK	AS	*ALL	94.2	75371	30	124K	48	0.2	0.1
	LOCK		AE91		36867	30	59348	48	0.1	0.0
	LOCK		AE92		38504	30	64434	48	0.3	0.1

Migrating from ICB4 to Infiniband and Beyond

Analyze Channel Information - PCHID0700

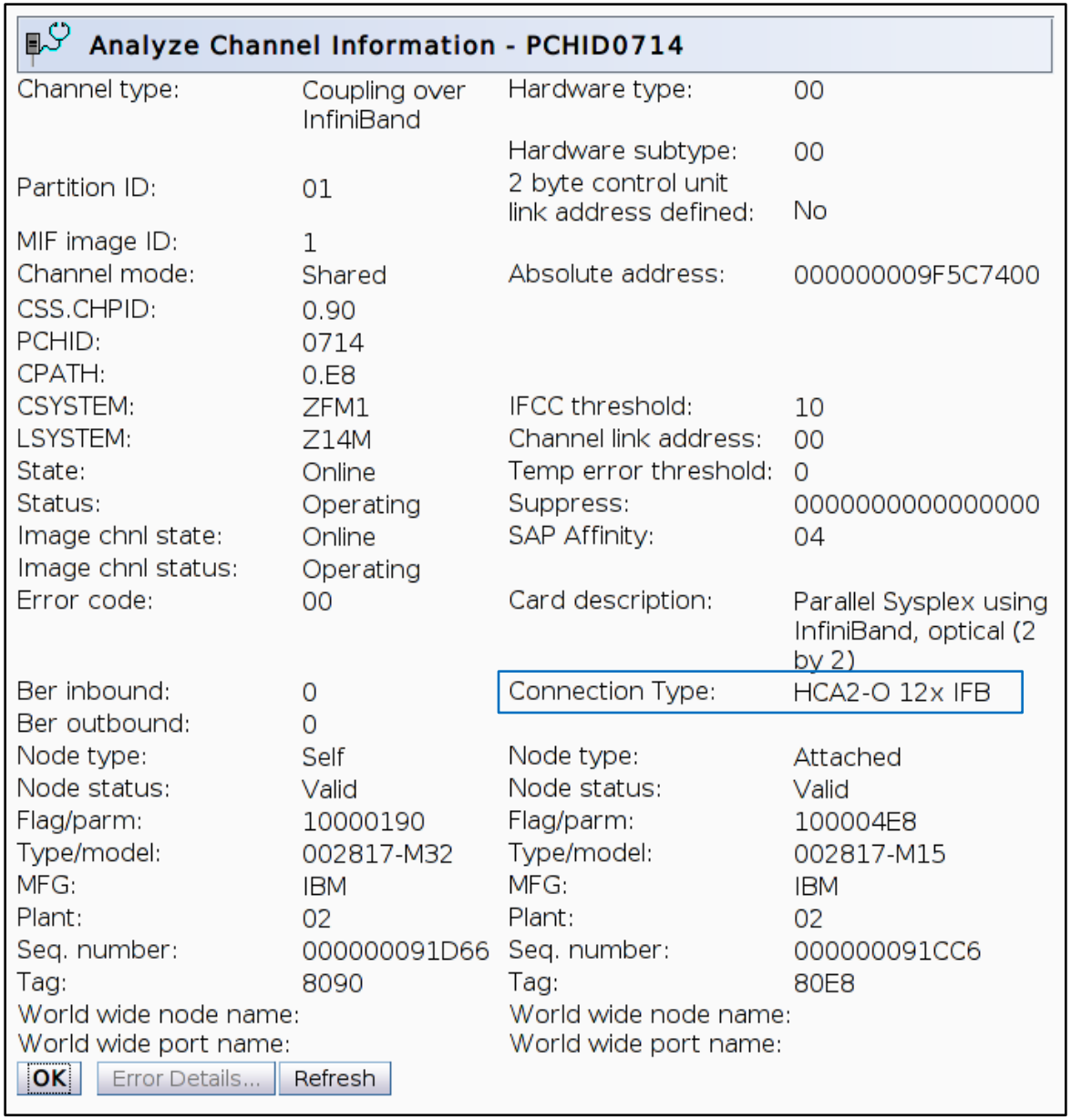
Channel type:	Coupling over InfiniBand	Hardware type:	00
		Hardware subtype:	00
Partition ID:	01	2 byte control unit link address defined:	No
MIF image ID:	1		
Channel mode:	Shared	Absolute address:	000000009F5C2400
CSS.CHPID:	0.05		
PCHID:	0700		
CPATH:	0.10		
CSYSTEM:	Z4M	IFCC threshold:	10
LSYSTEM:	Z7M	Channel link address:	00
State:	Online	Temp error threshold:	0
Status:	Operating	Suppress:	0000000000000000
Image chnl state:	Online	SAP Affinity:	08
Image chnl status:	Operating		
Error code:	00	Card description:	Parallel Sysplex using InfiniBand, optical (2 by 2)
Ber inbound:	0		
Ber outbound:	0		
Node type:	Self	Node type:	Attached
Node status:	Valid	Node status:	Valid
Flag/parm:	10000105	Flag/parm:	10000410
Type/model:	002817-M49	Type/model:	002817-M49
MFG:	IBM	MFG:	IBM
Plant:	02	Plant:	02
Seq. number:	000000091D96	Seq. number:	000000091D36
Tag:	8005	Tag:	8010
World wide node name:		World wide node name:	
World wide port name:		World wide port name:	

OK | Error Details... | Refresh

Driver 86 version

no card detail

Migrating from ICB4 to Infiniband and Beyond



Analyze Channel Information - PCHID0714

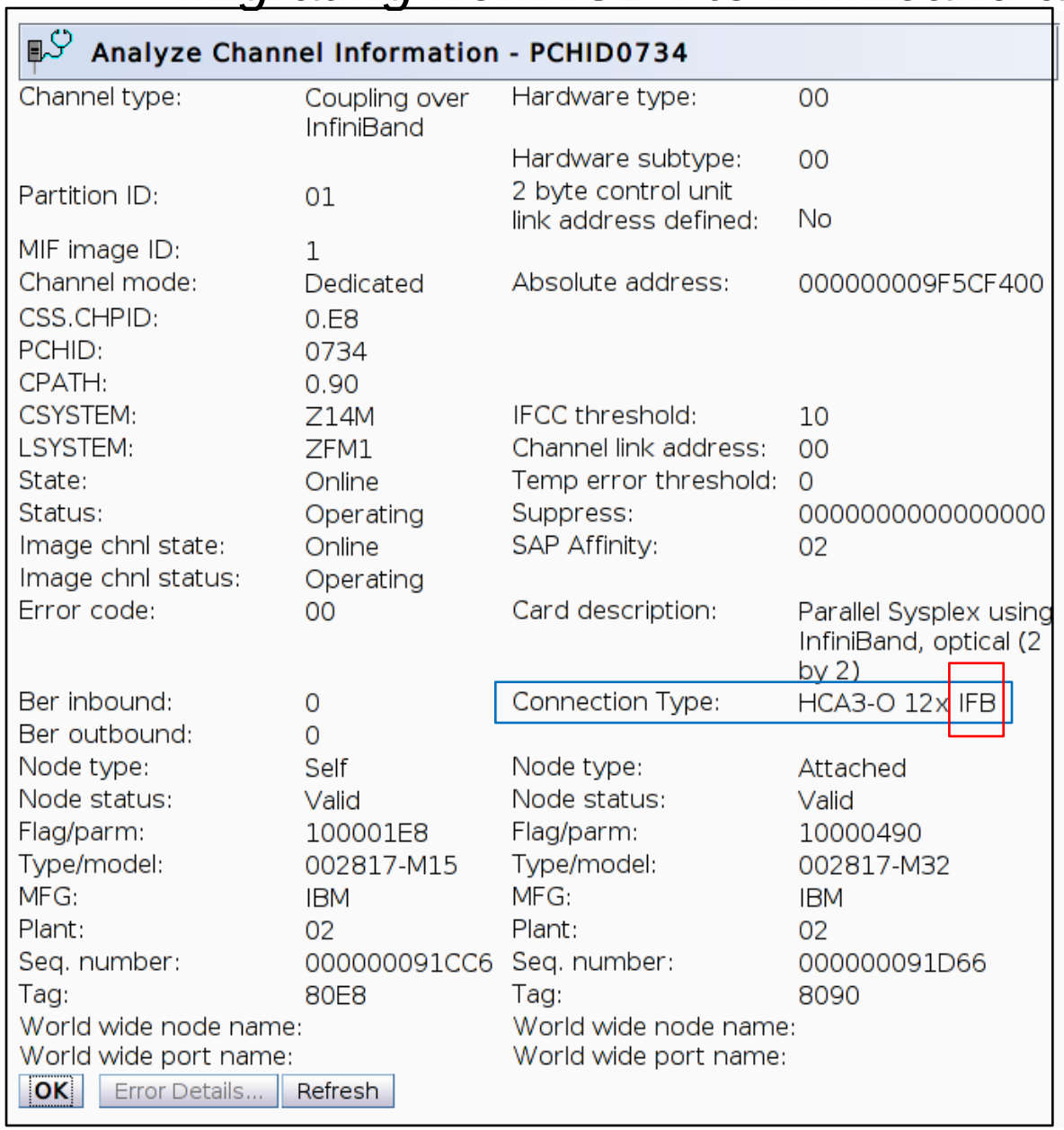
Channel type:	Coupling over InfiniBand	Hardware type:	00
		Hardware subtype:	00
Partition ID:	01	2 byte control unit link address defined:	No
MIF image ID:	1		
Channel mode:	Shared	Absolute address:	000000009F5C7400
CSS.CHPID:	0.90		
PCHID:	0714		
CPATH:	0.E8		
CSYSTEM:	ZFM1	IFCC threshold:	10
LSYSTEM:	Z14M	Channel link address:	00
State:	Online	Temp error threshold:	0
Status:	Operating	Suppress:	0000000000000000
Image chnl state:	Online	SAP Affinity:	04
Image chnl status:	Operating		
Error code:	00	Card description:	Parallel Sysplex using InfiniBand, optical (2 by 2)
Ber inbound:	0	Connection Type:	HCA2-O 12x IFB
Ber outbound:	0		
Node type:	Self	Node type:	Attached
Node status:	Valid	Node status:	Valid
Flag/parm:	10000190	Flag/parm:	100004E8
Type/model:	002817-M32	Type/model:	002817-M15
MFG:	IBM	MFG:	IBM
Plant:	02	Plant:	02
Seq. number:	000000091D66	Seq. number:	000000091CC6
Tag:	8090	Tag:	80E8
World wide node name:		World wide node name:	
World wide port name:		World wide port name:	

OK | Error Details... | Refresh

Driver 93 version

Connection type added

Migrating from ICB4 to Infiniband and Beyond



Analyze Channel Information - PCHID0734

Channel type:	Coupling over InfiniBand	Hardware type:	00
		Hardware subtype:	00
Partition ID:	01	2 byte control unit link address defined:	No
MIF image ID:	1		
Channel mode:	Dedicated	Absolute address:	000000009F5CF400
CSS.CHPID:	0.E8		
PCHID:	0734		
CPATH:	0.90		
CSYSTEM:	Z14M	IFCC threshold:	10
LSYSTEM:	ZFM1	Channel link address:	00
State:	Online	Temp error threshold:	0
Status:	Operating	Suppress:	0000000000000000
Image chnl state:	Online	SAP Affinity:	02
Image chnl status:	Operating		
Error code:	00	Card description:	Parallel Sysplex using InfiniBand, optical (2 by 2)
Ber inbound:	0	Connection Type:	HCA3-O 12x IFB
Ber outbound:	0		
Node type:	Self	Node type:	Attached
Node status:	Valid	Node status:	Valid
Flag/parm:	100001E8	Flag/parm:	10000490
Type/model:	002817-M15	Type/model:	002817-M32
MFG:	IBM	MFG:	IBM
Plant:	02	Plant:	02
Seq. number:	000000091CC6	Seq. number:	000000091D66
Tag:	80E8	Tag:	8090
World wide node name:		World wide node name:	
World wide port name:		World wide port name:	

OK | Error Details... | Refresh

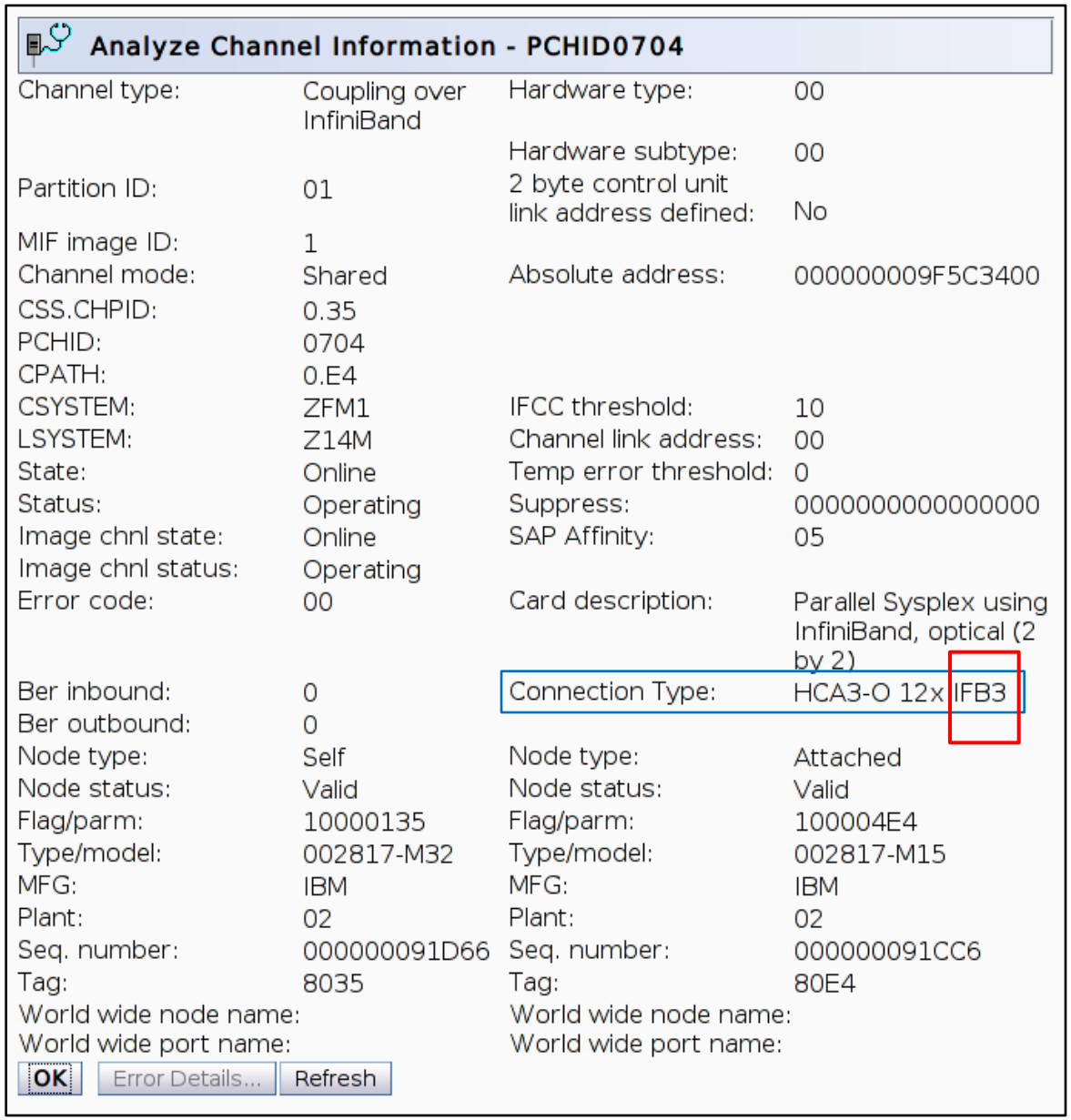
Driver 93 version

Connection type added

HCA3 12x when connected to an HCA2 12x

IFB indicates the HCA3 card is operating in HCA2 mode

Migrating from ICB4 to Infiniband and Beyond



Analyze Channel Information - PCHID0704

Channel type:	Coupling over InfiniBand	Hardware type:	00
		Hardware subtype:	00
Partition ID:	01	2 byte control unit link address defined:	No
MIF image ID:	1		
Channel mode:	Shared	Absolute address:	000000009F5C3400
CSS.CHPID:	0.35		
PCHID:	0704		
CPATH:	0.E4		
CSYSTEM:	ZFM1	IFCC threshold:	10
LSYSTEM:	Z14M	Channel link address:	00
State:	Online	Temp error threshold:	0
Status:	Operating	Suppress:	0000000000000000
Image chnl state:	Online	SAP Affinity:	05
Image chnl status:	Operating		
Error code:	00	Card description:	Parallel Sysplex using InfiniBand, optical (2 by 2)
Ber inbound:	0	Connection Type:	HCA3-O 12x IFB3
Ber outbound:	0		
Node type:	Self	Node type:	Attached
Node status:	Valid	Node status:	Valid
Flag/parm:	10000135	Flag/parm:	100004E4
Type/model:	002817-M32	Type/model:	002817-M15
MFG:	IBM	MFG:	IBM
Plant:	02	Plant:	02
Seq. number:	000000091D66	Seq. number:	000000091CC6
Tag:	8035	Tag:	80E4
World wide node name:		World wide node name:	
World wide port name:		World wide port name:	

OK | Error Details... | Refresh

Driver 93 version

Connection type added

HCA3 12x when connected to an HCA3 12x

IFB3 indicates the HCA3 card is operating in the desired mode

Migrating from ICB4 to Infiniband and Beyond

For more detailed information on Infiniband… look out for the soon to be released **Redbook**

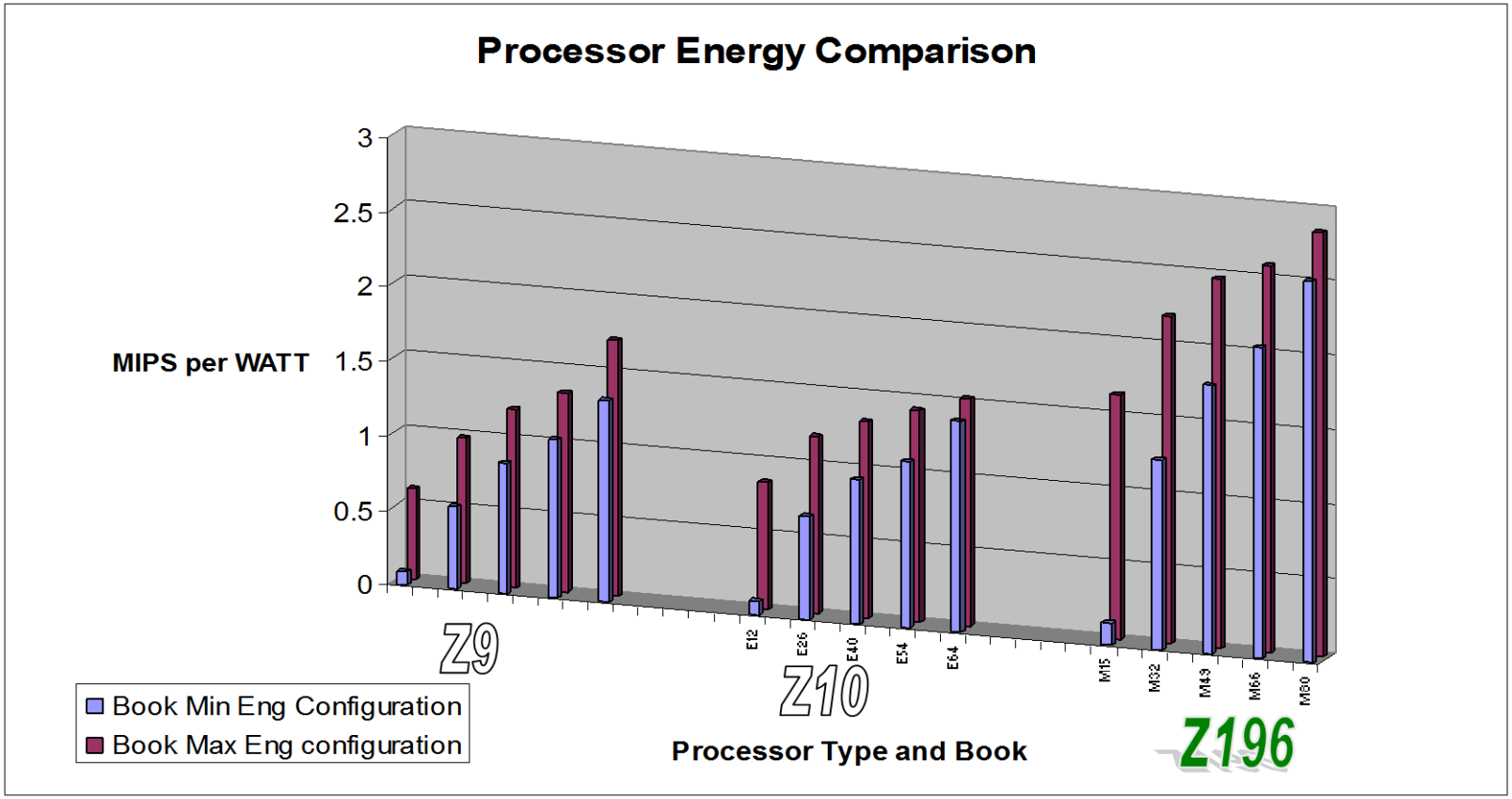
Infiniband Coupling Links on System Z

Questions?

Migrating from ICB4 to Infiniband and Beyond

Appendix A: z196 Performance

Migrating from ICB4 to Infiniband and Beyond



- These numbers were generated by the power estimation tool.
- Power requirements and MIPS per Book were generated using a minimum and maximum processors per book.

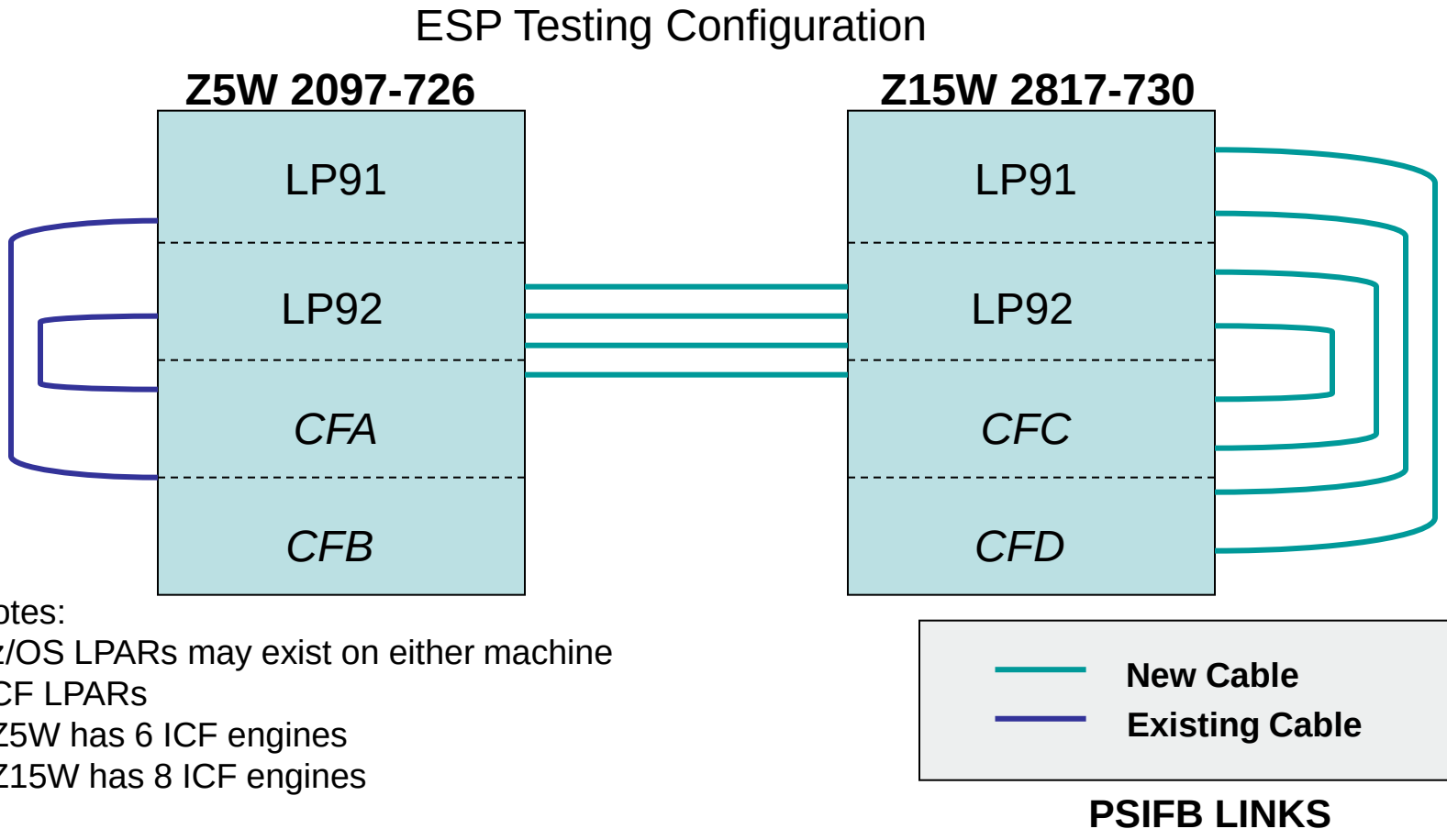
Migrating from ICB4 to Infiniband and Beyond

Hiperdispatch

Hiperdispatch characteristics are an evolving science

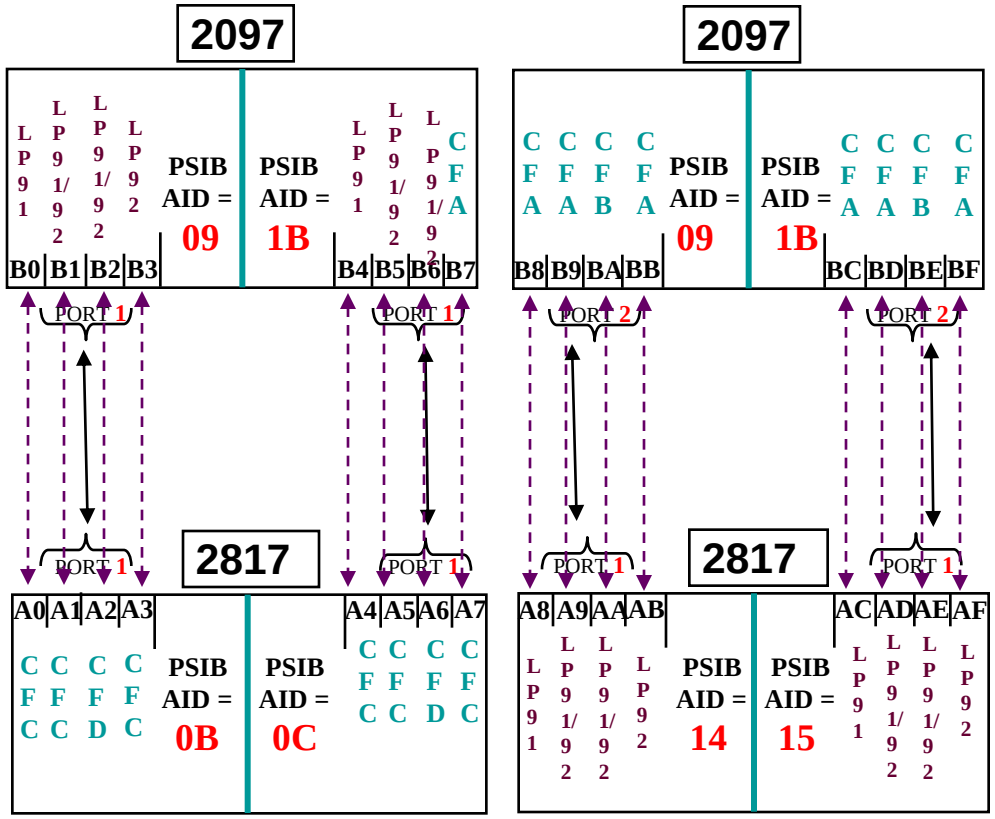
- The Hipervisor continues to align physical processors vertically to a subset of logical processors (as introduced with the z10)
- WLM's management of the Affinity nodes has evolved.
  - Node affinity scope has changed from same book to same core (sharing level 3 cache influences this change)
  - WLM now places work on a weighted basis on the affinity nodes. 4 core chips/nodes are assigned more work that a 3 core chip.

Migrating from ICB4 to Infiniband and Beyond



Notes:
- z/OS LPARs may exist on either machine
- CF LPARs
- Z5W has 6 ICF engines
- Z15W has 8 ICF engines

Migrating from ICB4 to Infiniband and Beyond



ESP - **Z5W TO Z15W** INFINIBAND CONNECTS

Migrating from ICB4 to Infiniband and Beyond

CICS TRANSACTION KEY

Tran Name	Threadsafe	Description
GHDL	**No**	**MQ 100 - .25 MS WAITS + 16 Million instructions per POST**
GHC1	**YES**	**1100 fetches per tran out of cache**
GHC2	**NO**	**1100 fetches per tran out of cache**
GHI0	**NO**	**1000-1100 fetches per tran prefetch I/O**
GHI1	**YES**	**1000-1100 fetches per tran prefetch I/O**

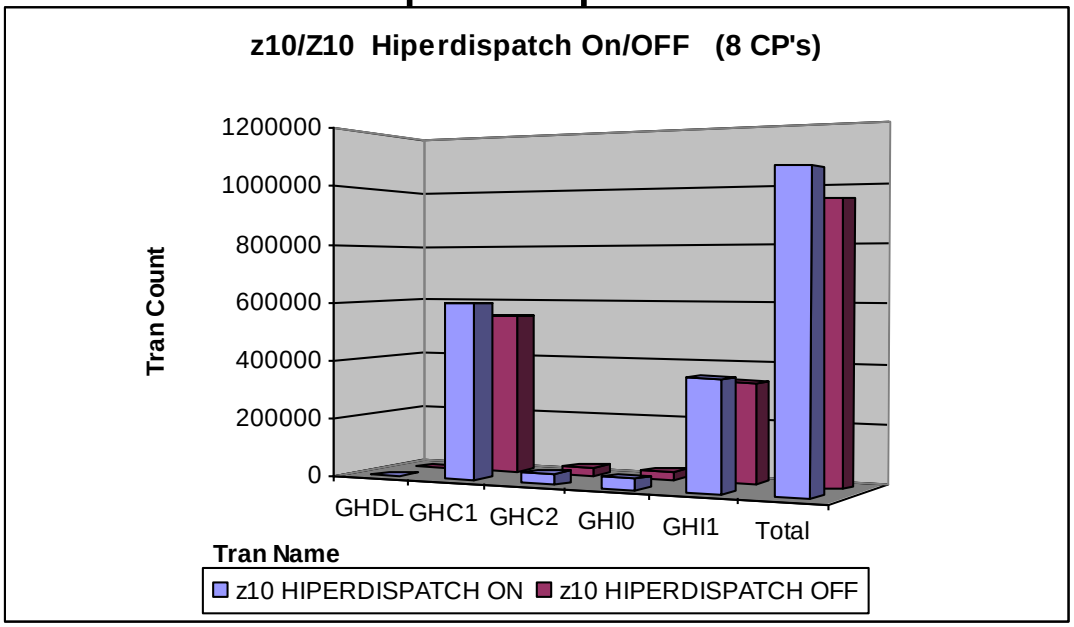
Transaction Driver - completion based initiation

Reads MQ Queue for transaction specifications for:

- Duration
- Transaction name
- Transaction concurrency

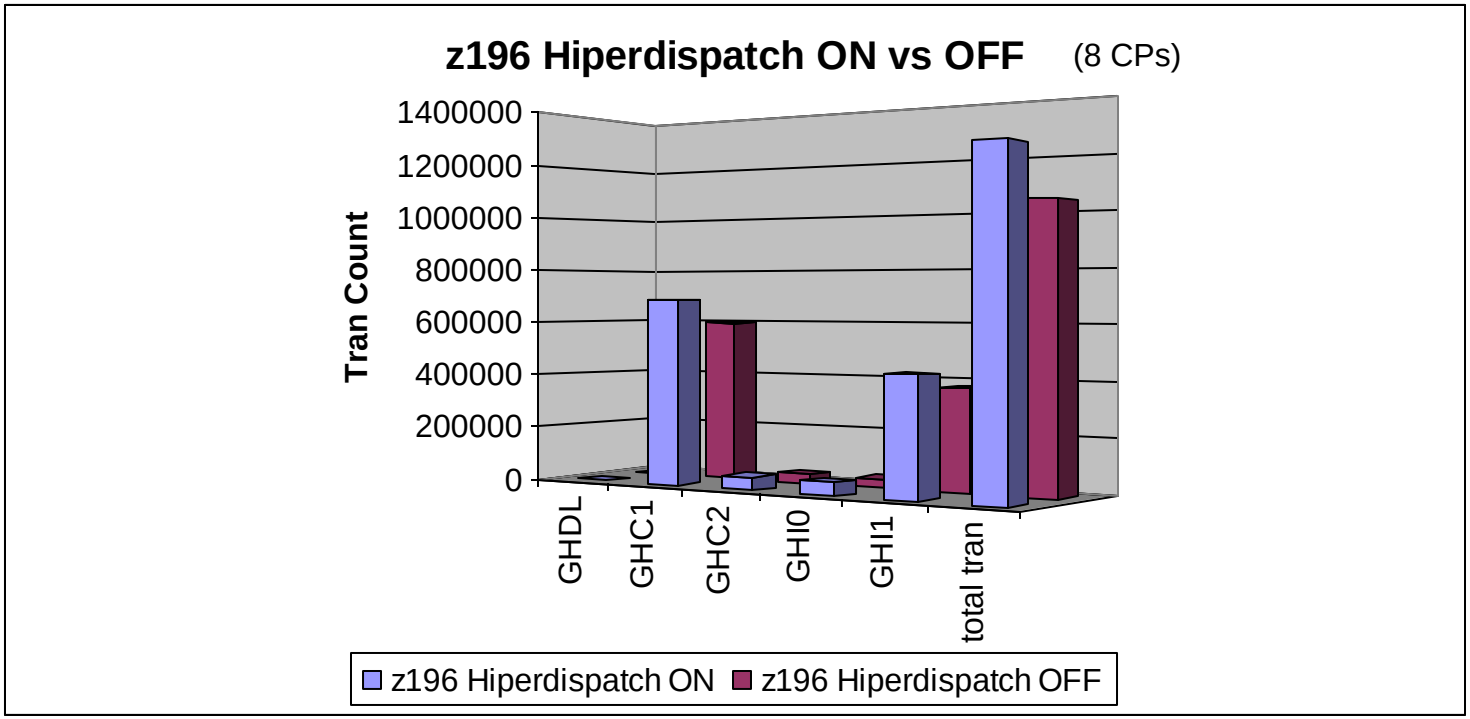
Migrating from ICB4 to Infiniband and Beyond

Hiperdispatch



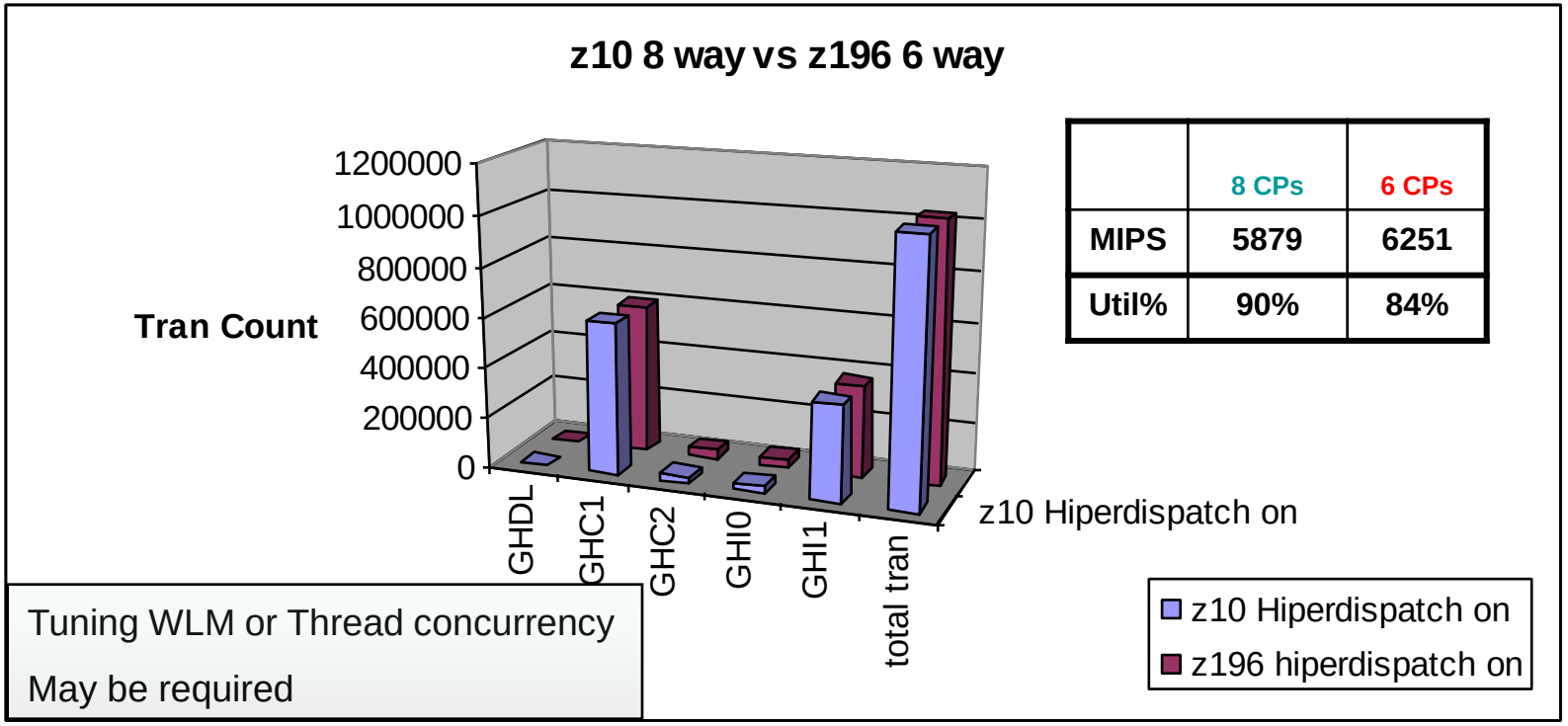
Tran	GHDL	GHC1	GHC2	GHI0	GHI1	total tran
z10 Hiperdispatch ON	**160**	**598667**	**36133**	**35100**	**367649**	**1037709**
z10 Hiperdispatch OFF	**144**	**551831**	**27873**	**27667**	**338888**	**946403**
Tran delta	**16**	**46836**	**8260**	**7433**	**28761**	**91306**
% delta	**11.11%**	**8.49%**	**29.63%**	**26.87%**	**8.49%**	**9.65%**

Migrating from ICB4 to Infiniband and Beyond



Tran	GHDL	GHC1	GHC2	GHI0	GHI1	total tran
z196 Hiperdispatch ON	180	688194	48813	47881	433296	1218364
z196 Hiperdispatch OFF	168	596079	34169	34001	373456	1037873
Tran delta	12	92115	14644	13880	59840	180491
% delta	7.14%	15.45%	42.86%	40.82%	16.02%	17.39%

Migrating from ICB4 to Infiniband and Beyond



	8 CPs	6 CPs
MIPS	5879	6251
Util%	90%	84%

Tuning WLM or Thread concurrency

May be required

	GHDL	GHC1	GHC2	GHI0	GHI1	total tran
z196 Hiperdispatch ON	172	604553	26864	26853	383130	1041572
z10 Hiperdispatch ON	160	598667	36133	35100	367649	1037709
Tran delta	12	5886	-9269	-8247	15481	3863
% delta	7.50%	0.98%	-25.65%	-23.50%	4.21%	0.37%

Migrating from ICB4 to Infiniband and Beyond

2817 Model Characteristics by Book

	1st Book			2nd Book			3rd Book			4th Book			
Model	**Avail CPs**	**SAPs**	**Spares**	**Avail CPs**	**SAPs**	**Spares**	**Avail CPs**	**SAPs**	**Spares**	**Avail CPs**	**SAPs**	**Spares**	**Max Memory for Model**
M15	**15**	**3**	**2**										**752 GB**
M32	**16**	**3**	**1**	**16**	**3**	**1**							**1520 GB**
M49	**16**	**3**	**1**	**16**	**3**	**1**	**17**	**3**	**0**				**2288GB**
M66	**16**	**3**	**1**	**16**	**3**	**1**	**17**	**3**	**0**	**17**	**3**	**0**	**3056GB**
M80	**20**	**3**	**1**	**20**	**3**	**1**	**20**	**4**	**0**	**20**	**4**	**0**	**3056GB**

Migrating from ICB4 to Infiniband and Beyond

Hiperdispatch

- The z196 processor rewards your ability to hold a dispatch!
- MIPS are nice but point of dispatch must also be considered
- Chatty workloads (CICS QR for example) vs Batch or CICS Threadsafe lose ground to workloads that hold that their dispatch.
- When the point of dispatch is reduced and the MIPS per engine is increased the capability of the QR throughput increases from an engine perspective. BUT the competition for that engine also increases
- Machine upgrades often result in more MIPS being delivered by a reduced number of engines.
- Tuning WLM goals or managing concurrency may be required!

Migrating from ICB4 to Infiniband and Beyond

SYSID	Mon	Day	SH	Hour	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L2P	L3P	L4LP	L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	CICS
PAR1	SEP	1	P	16	7.0	31.8	2.8	4.2	50	8.5	69.3	19.0	7.6	3.4	0.7	0.46	1459.4	5.2	QR
PAR2	SEP	1	P	16	4.7	23.4	3.0	1.7	24	7.0	89.2	5.6	4.5	0.1	0.7	0.20	1546.2	5.2	Threadsafe

Benchmark Description:

Comprises of CICS transactions and some Batch…

- All Batch is heavy Update and running on both LPARs
- The CICS transactions are cloned pairs. One group is left to run in QR mode and the other is marked threadsafe in the CICS PPT definition. This test Focused all the Quasi-Reentrant transactions in one LPAR and all the Threadsafe transactions in the other LPAR. Transaction concurrency was establish in order to drive the LPARs to 90%+ utilization levels.

Threadsafe Vs QR Results

CICS 110s

- Increase of 52% of transactions
- Decrease of 42% in CPU per Transaction
- Decrease of average response time by 67% (3.0x)

RMF 72s – CICS Storage Class

- Ended Transactions up 2.4x
- Response Time down 3.6x

SMF 113s – LPAR

- CPI down 1.48x from 7.0 to 4.7

  L1MP down 1.5% from 8.5% to 7.0%

  L2P up 19.9% from 69.3% to 89.2%

CICS Threadsafe is an option that may help you reduce CPU cost for applicable transactions by reducing switches between different TCB types

CPU MF example to supplement CICS and RMF performance metrics

As a secondary data source to understand why performance may have changed

These numbers come from a synthetic Benchmark and do not represent a production workload

Migrating from ICB4 to Infiniband and Beyond

SYSID	Mon	Day	SH	Hour	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L15P	L2LP	L2RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Ded/Shr
AE83	JUN	1	P	TOTAL	6.87	15.4	3.0	3.9	88	4.4	69.5	23.9	0.1	6.6	0.73	464.7	4.40	Ded
AE83	JUN	2	P	TOTAL	7.27	14.9	3.2	4.1	93	4.4	67.2	25.8	0.1	6.9	0.78	612.8	4.40	Ded
AE83	JUN	3	P	TOTAL	6.43	12.5	2.9	3.6	93	3.8	67.3	25.5	0.1	7.1	0.79	627.6	4.40	Ded
AE83	JUN	4	P	TOTAL	7.14	14.8	3.2	4.0	91	4.4	67.5	25.6	0.1	6.8	0.77	553.4	4.40	Ded
	Average				6.93	14.4	3.1	3.9	91	4.3	67.9	25.2	0.1	6.9	0.77	564.6		Ded
AE83	MAY	25	P	TOTAL	7.66	17.6	3.2	4.5	87	5.2	68.6	25.4	0.1	6.0	0.70	641.6	4.40	Shr
AE83	MAY	26	P	TOTAL	7.36	16.1	3.1	4.3	87	4.9	69.0	24.7	0.1	6.3	0.72	621.1	4.40	Shr
AE83	MAY	27	P	TOTAL	7.31	14.5	3.1	4.2	88	4.8	68.5	25.0	0.1	6.4	0.73	395.6	4.40	Shr
	Average				7.44	16.0	3.1	4.3	87	5.0	68.7	25.0	0.1	6.2	0.72	552.8		Shr
Dedicated / Shared					0.93	0.90	0.98	0.90	1.04	0.86	0.99	1.01	0.98	1.11	1.07	1.02		
					1.07	Relative ITR Capacity Ratio Of Dedicated Vs Shared												

CPI – Cycles per Instruction

Prb State - % Problem State

Est Instr Cmplx CPI – Estimated Instruction Complexity CPI (infinite L1)

Est Finite CPI – Estimated CPI from Finite cache/memory

Est SCPL1M – Estimated Sourcing Cycles per Level 1 Miss

L1MP – Level 1 Miss %

L15P – % sourced from Level 2 cache

L2LP – % sourced from Level 2 Local cache (on same book)

L2RP – % sourced from Level 2 Remote cache (on different book)

MEMP - % sourced from Memory

Rel Nest Intensity – Reflects distribution and latency of sourcing from shared caches and memory

LPARCPU - APPL% (GCPs, zAAPs, zIIPs) captured and uncaptured

Eff GHz – Effective gigahertz for GCPs, cycles per nanosecond

Migrating from ICB4 to Infiniband and Beyond

HIS Field Definitions

CPI	NUM	8	6.1	CYCLES*PER*INSTRUCTION
EFFGHZ	NUM	8	6.1	EFFECTIVE*GIGAHERTZ*CYCLES*PER NANO
ESTFINCP	NUM	8	6.1	ESTIMATED*CPI FROM*FINITE*CACHE/MEM
ESTICCPI	NUM	8	6.1	ESTIMATED*INSTRUCTION*COMPLEXITY*CPI
ESTSCP1M	NUM	8	6.1	ESTIMATED*SOURCING*CYCLES*PER L1 MISS
LPARBUSY	NUM	8	6.1	LPARCPU*PERCENT*CAPTURED AND*UNCAPTURED
L1MP	NUM	8	6.1	LEVEL*1*MISS*PERCENT
L15P	NUM	8	6.1	PERCENT*SOURCED*FROM*L1.5*CACHE
L2LP	NUM	8	6.1	PERCENT*SOURCED*FROM*L2*SAME BOOK
L2P	NUM	8	6.1	PERCENT*SOURCED*FROM*L2*CACHE
L2RP	NUM	8	6.1	PERCENT*SOURCED*FROM*L2*DIFFEERNT*BOOK
L3P	NUM	8	6.1	PERCENT*SOURCED*FROM*L3*SAME CHIP CACHE
L4LP	NUM	8	6.1	PERCENT*SOURCED*FROM*L4*SAME BOOK
L4RP	NUM	8	6.1	PERCENT*SOURCED*FROM*L4*DIFFERENT*BOOK
MEMP	NUM	8	6.1	PERCENT*SOURCED*FROM*MEMORY
MIPSEXEC	NUM	8		EXECUTED*MIPS
PRBSTATE	NUM	8	6.1	PERCENT*PROBLEM*STATE
PTEPCTMI	NUM	8	6.1	PAGETABLE*ENTRY*PCT OF TLB*MISSES
RNI	NUM	8	6.1	RELATIVE*NEST*INTENSITY
TLB1CYCL	NUM	8	6.1	CYCLES*PER*TLB*MISS
TLB1MISS	NUM	8	6.1	TLB*CPU MISS*PERCENT OF*TOTAL CPU

Migrating from ICB4 to Infiniband and Beyond

zIIP consumption in DB2 V10 was a pleasant finding as they were not observed in DB2 V9. The DBM1 address space shows zIIP consumption attributed to an Enclave that appears to be classified under the MSTR address space.

Samples: 1799 System: AE92 Date: 01/05/11 Time: 08.03.00 Range: 1800 Sec

		Service	--- Time on CP % ---			----- EAppl % -----		
Jobname	CX	Class	Total	AAP	IIP	CP	AAP	IIP
T8D1EASM	BO	ONLISO1	230.7	0.0	0.0	230.7		0.0
T8D1EASP	BO	ONLISO1	229.8	0.0	0.0	229.8		0.0
T8D1EASO	BO	ONLISO1	228.7	0.0	0.0	228.7		0.0
T8D1EASN	BO	ONLISO1	228.4	0.0	0.0	228.4		0.0
T8D1EASQ	BO	ONLISO1	227.7	0.0	0.0	227.7		0.0
T8D1EASR	BO	ONLISO1	227.6	0.0	0.0	227.6		0.0
DBPBDBM1	S	SYSSTC	18.2	0.0	0.0	18.3		25.9
DBUBP921	B	BATISO03	8.9	0.0	0.0	8.9		0.0
DBUBP925	B	BATISO03	8.9	0.0	0.0	8.9		0.0

Migrating from ICB4 to Infiniband and Beyond

Appendix B: RMF Additional Reports

Migrating from ICB4 to Infiniband and Beyond

HCA2 Performance - CF on 2817 - T1

Samples: 1800 Systems: 2 Date: 09/21/11 Time: 15.00.00 Range: 1800 Sec

CF Name	System	Subchannel Delay %	Subchannel Busy %	-- Paths -- Avail	-- Paths -- Delay %	-- Sync --- Rate	-- Sync --- Avg Serv	------- Async ------- Rate	Async Avg Serv	Async Chng %	Async Del %
CFC4	AE91	0.0	8.3	4	0.0	166K	11	2395	145	0.0	0.0
	AE92	0.0	10.1	4	0.0	190K	12	4113	112	0.0	0.0
CFC5	AE91			4							
	AE92			4							
CFE1	AE91			4							
	AE92			4							
CFE2	AE91			4							
	AE92			4							
CFE3	AE91			4							
	AE92			4							
CFE4	AE91			4							
	AE92			4							
SEB3	AE91	0.0	0.6	2	0.0	12.1	11	2504	32	0.0	0.0
	AE92	0.0	0.6	2	0.0	12.3	12	2514	32	0.0	0.0
SEB4	AE91	0.0	0.2	2	0.0	224.9	56	72.0	284	0.0	0.0
	AE92	0.0	0.4	2	0.0	293.4	67	117.0	304	0.0	0.0

Migrating from ICB4 to Infiniband and Beyond

HCA2 Performance - CF on 2097

Samples: 1800 Systems: 2 Date: 09/21/11 Time: 20.30.00 Range: 1800 Sec

CF Name	System	Subchannel Delay %	Subchannel Busy %	-- Paths -- Avail	-- Paths -- Delay %	-- Sync --- Rate	-- Sync --- Avg Serv	------- Async ------- Rate	------- Async ------- Avg Serv	------- Async ------- Chng %	------- Async ------- Del %
CFC4	AE91			4							
	AE92			4							
CFC5	AE91			4							
	AE92			4							
CFE1	AE91	0.0	10.3	4	0.0	168K	14	3607	122	0.0	0.0
	AE92	0.0	11.4	4	0.0	183K	14	4802	110	0.0	0.0
CFE2	AE91			4							
	AE92			4							
CFE3	AE91			4							
	AE92			4							
SEB3	AE91	0.0	0.6	2	0.0	12.0	11	2473	31	0.0	0.0
	AE92	0.0	0.6	2	0.0	12.2	12	2473	31	0.0	0.0
SEB4	AE91	0.0	0.4	2	0.0	251.4	127	38.6	711	0.0	0.0
	AE92	0.0	0.5	2	0.0	319.4	90	75.3	581	0.0	0.0

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - LR and 12X IFB3

Samples: 1800 Systems: 2 Date: 09/26/11 Time: 12.00.00 Range: 1800 Sec

CF Name	System	Subchannel Delay %	Subchannel Busy %	-- Paths -- Avail	-- Paths -- Delay %	-- Sync --- Rate	-- Sync --- Avg Serv	------- Async ------- Rate	------- Async ------- Avg Serv	------- Async ------- Chng %	------- Async ------- Del %
CFC4	AE91	0.0	3.0	8	0.0	165K	10	2943	127	0.0	0.0
	AE92	0.0	3.2	8	0.0	180K	10	2891	126	0.0	0.0
CFC5	AE91			8							
	AE92			8							
SEB3	AE91	0.0	1.3	2	0.0	11.8	11	1584	118	0.0	0.0
	AE92	0.0	1.5	2	0.0	11.9	12	1584	131	0.0	0.0
SEB4	AE91	0.0	0.8	2	0.0	74.8	92	98.7	1073	0.0	0.0
	AE92	0.0	1.2	2	0.0	108.2	90	121.0	1305	0.0	0.0

Migrating from ICB4 to Infiniband and Beyond

HCA3 Performance - 12X IFB3

Samples: 1800 Systems: 2 Date: 09/26/11 Time: 13.00.00 Range: 1800 Sec

CF Name	System	Subchannel Delay %	Subchannel Busy %	-- Paths -- Avail	-- Paths -- Delay %	-- Sync --- Rate	-- Sync --- Avg Serv	------- Async ------- Rate	------- Async ------- Avg Serv	------- Async ------- Chng %	------- Async ------- Del %
CFC4	AE91	0.0	4.4	4	0.0	189K	5	2109	94	0.0	0.0
	AE92	0.0	4.8	4	0.0	206K	5	2262	91	0.0	0.0
CFC5	AE91			4							
	AE92			4							
SEB3	AE91	0.0	0.7	2	0.0	11.5	12	2424	39	0.0	0.0
	AE92	0.0	0.7	2	0.0	11.6	12	2424	38	0.0	0.0
SEB4	AE91	0.0	0.3	2	0.0	58.3	101	55.7	542	0.0	0.0
	AE92	0.0	0.6	2	0.0	106.6	104	92.7	818	0.0	0.0