#SHAREorg

SHARE
Technology · Connections · Results

Migrating from z10 ICBs to z196 Infiniband- a Detailed Performance Study and User Experience

Meral Temel
Garanti Technology

15 March 2012 10593

Agenda

- Who is GT ?
- GT-Parallel Sysplex Configuration
- General Information About PSIFB
- General Information About CF Request
  - Life Cycle Of CF Requests
  - How Does Sync/Async Heuristic Algorithm Work ?
- Effect Of Upgrade To Performance Items
- RMF Mon III,I panels That Are Used and Key Fields
- More Information & Backup Slides

Who is GT ?

GARANTİ TECHNOLOGY

SHARE

- A wholly-owned subsidiary of Garanti Bank, the second largest private bank in Turkey owned by Doğuş Group and BBVA.
- One of the largest private internal IT service providers in Turkey
- Most up-to-date IT infrastructure
- Tightly integrated and fully in-house developed, custom-fit IT solutions
- Uninterrupted transaction capability and infrastructure security
- Well-reputed as a company of “firsts”
- Visionary and continuous investment in technology since 90’s

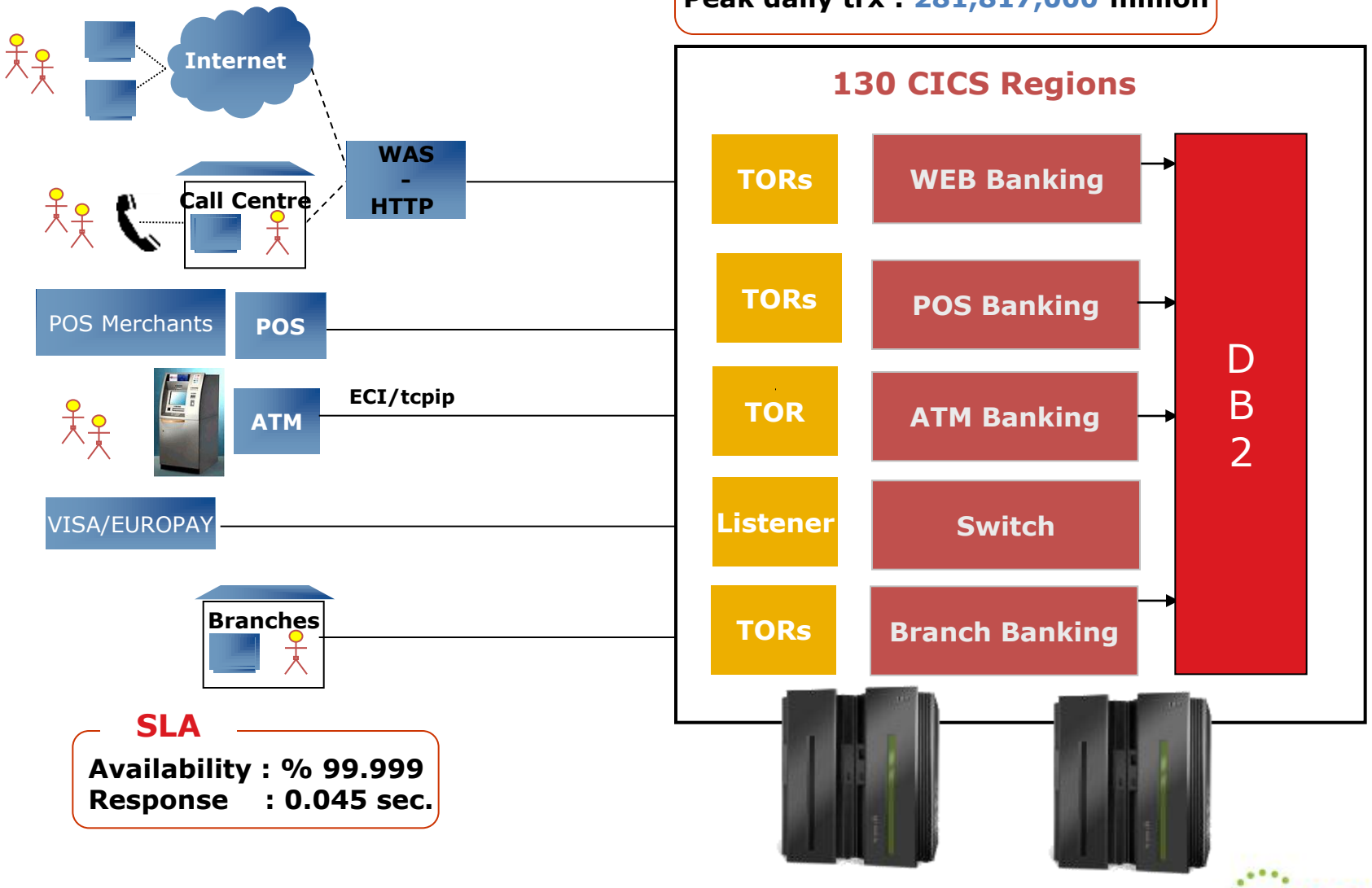
Garanti

- Fast decision making and strong communication from top to down
- Centralized management reporting systems, enable management to take timely actions
- Advanced CRM applications
- Paperless banking

Our Customers

GT- CICS Configuration –TORs & AORs

Average daily trx : 205 million
Peak daily trx : 281,817,000 million



SLA

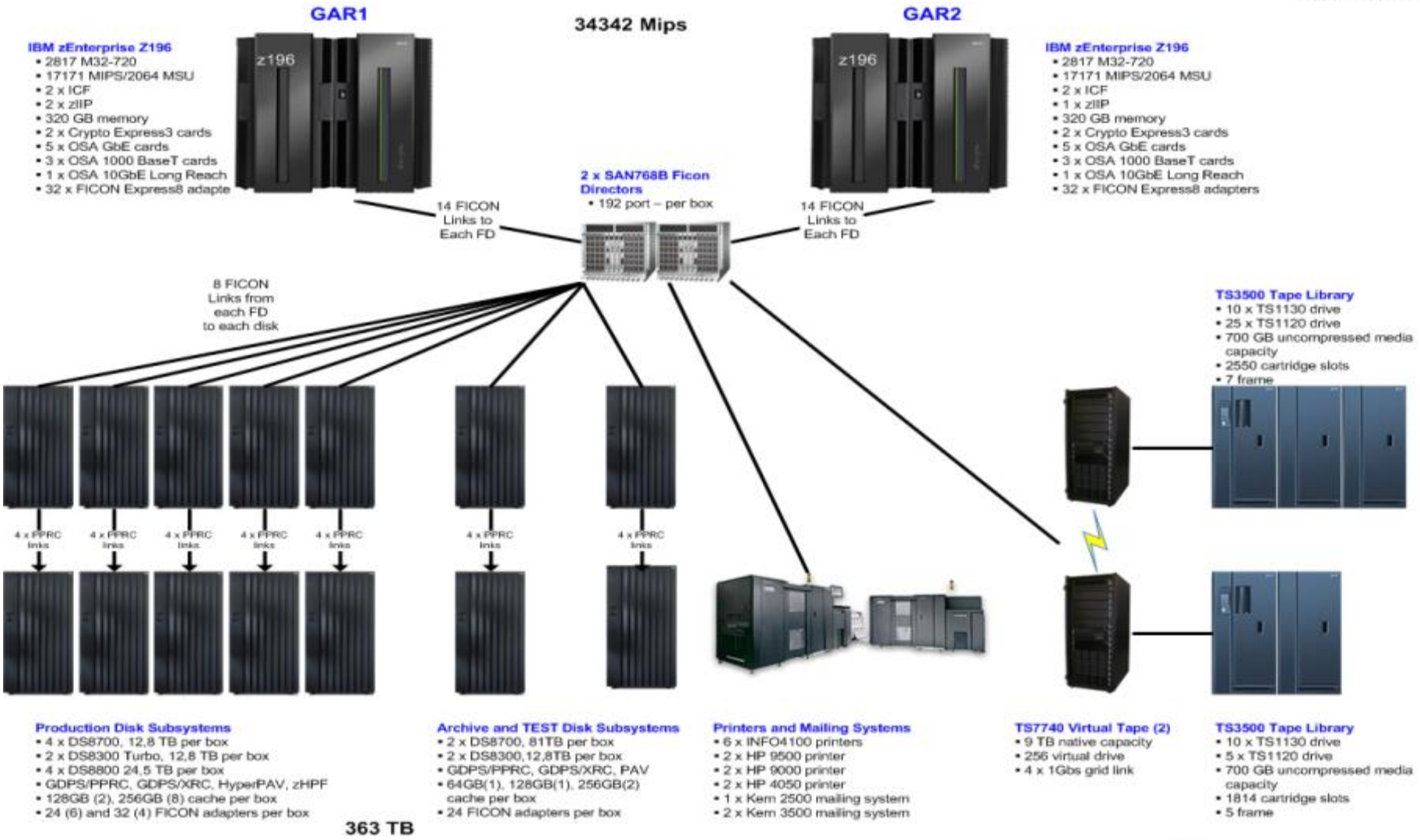
Availability : % 99.999
Response : 0.045 sec.

GT Is A Member Of …

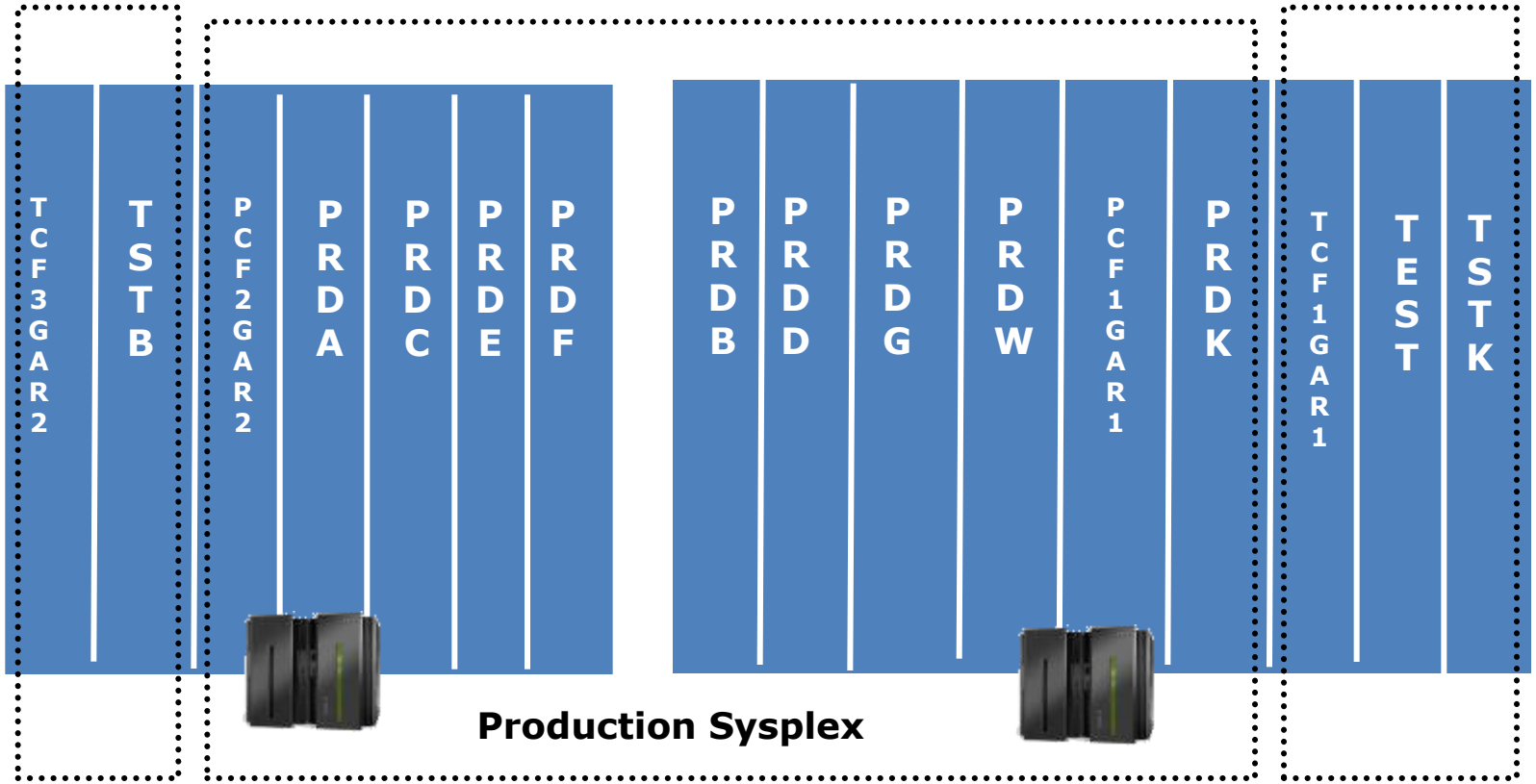
- SHARE
- CMG
- GDPS Design Council
- zBLC

GT Parallel Sysplex Configuration

GT-Mainframe Configuration

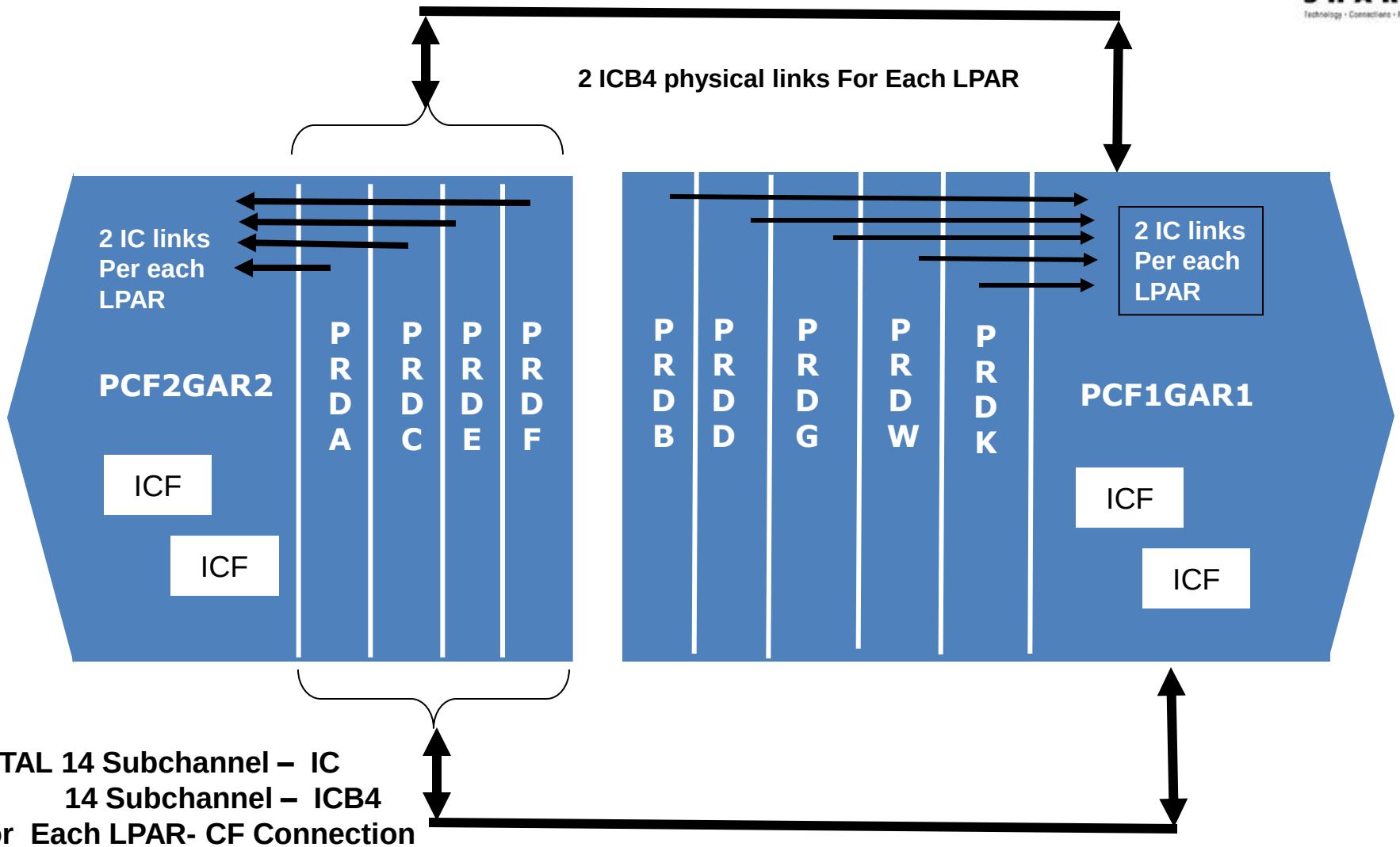


GT Parallel Sysplex Configuration - LPARS

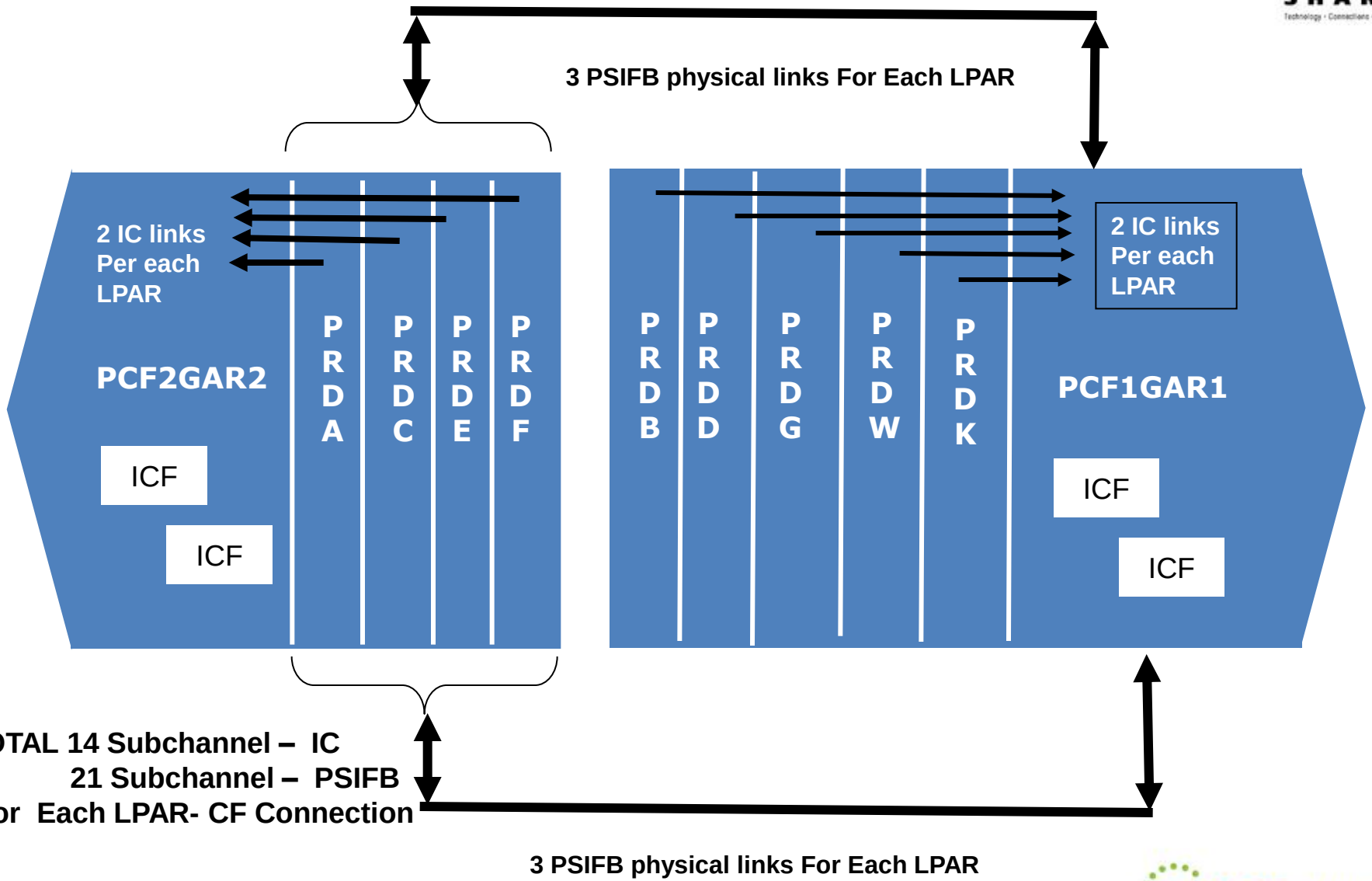


z/OS 1.12
DB2 V9
CICS TS 3.2

GT Parallel Sysplex Configuration - ICFs & CF Links z10



GT Parallel Sysplex Configuration - ICFs & CF Links z196



z196 Infiniband Adapters & Chpids

GAR1

6 Infiniband Adapters
12 Physical Ports/Links

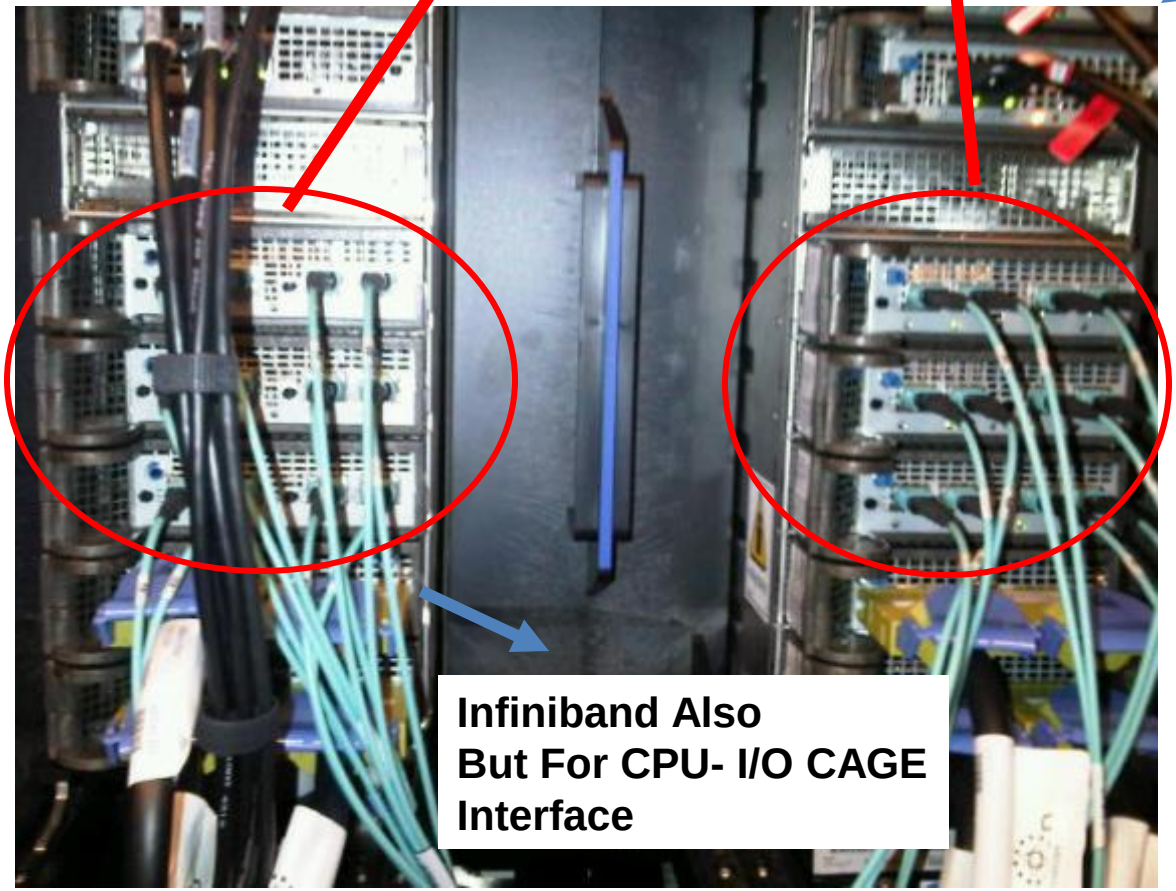
GAR2

6 Infiniband Adapters
12 Physical Ports/Links

z196 Infiniband Adapters & Chpids

3 PSFIB Adapters + 3 PSFIB Adapters

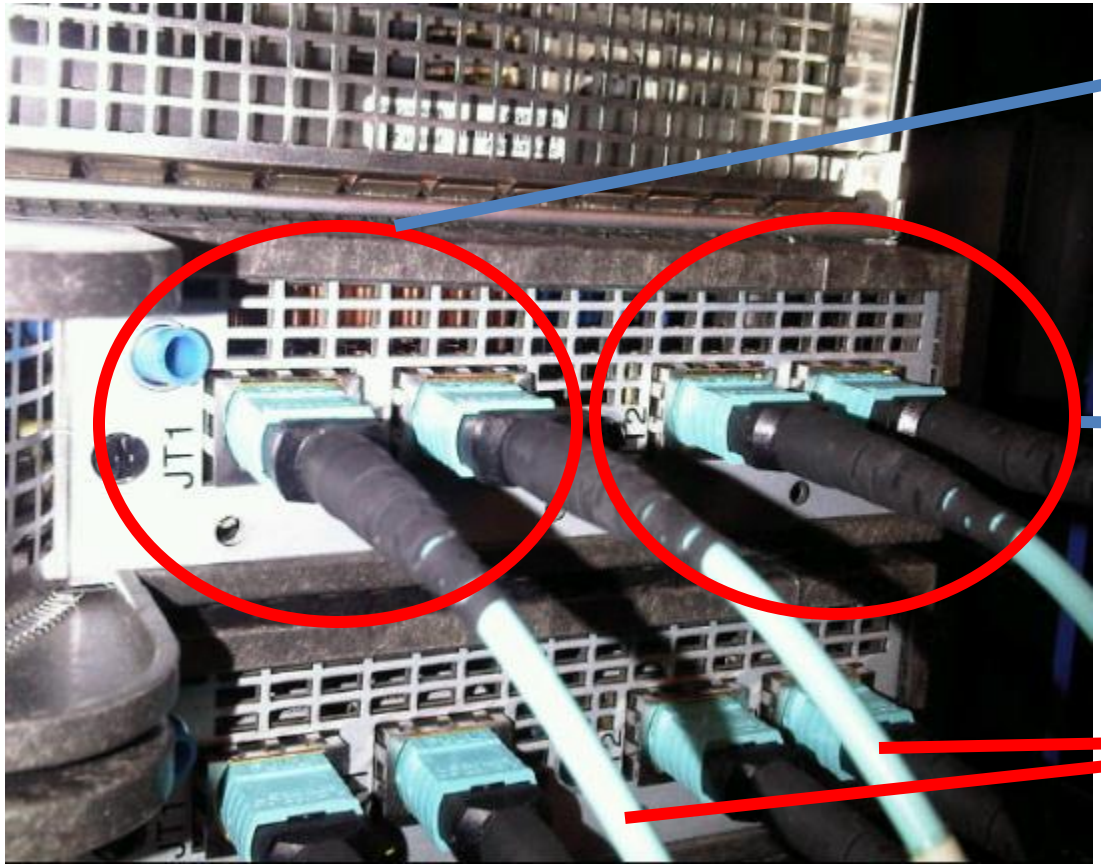
Book 1 Book 3



Infiniband Also
But For CPU- I/O CAGE
Interface

z196 Infiniband Links- CHPIDs

IN ONE INFINIBAND ADAPTER CARD ,THERE ARE 2 PORTS



PORT1
ONE PHYSICAL CONNECTION(LINK)
ONE CABLE (Transmit/Receive)

PORT2
ONE PHYSICAL CONNECTION(LINK)
ONE CABLE (Transmit/Receive)

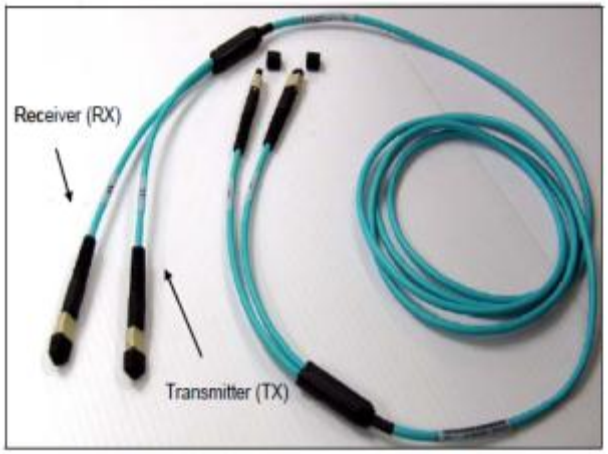
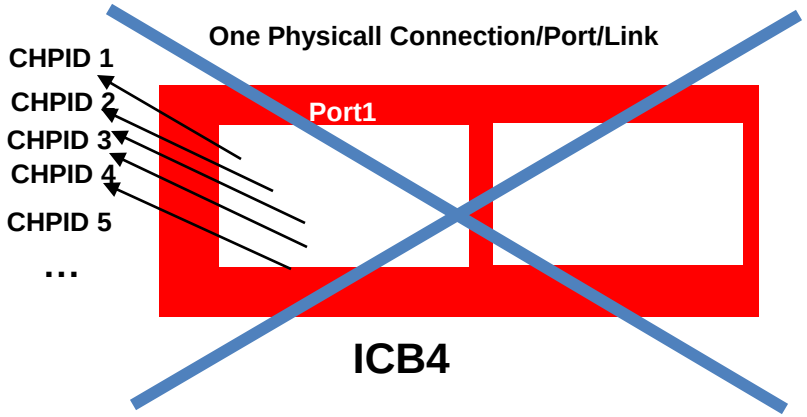
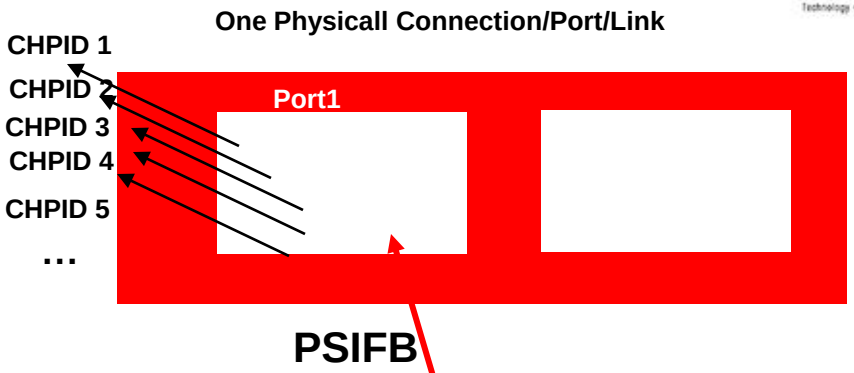


Figure 2-10 Optical InfiniBand cable, including TX and RX labels

z10 ICB4 & z196 Infiniband Links- CHPIDs Relationship

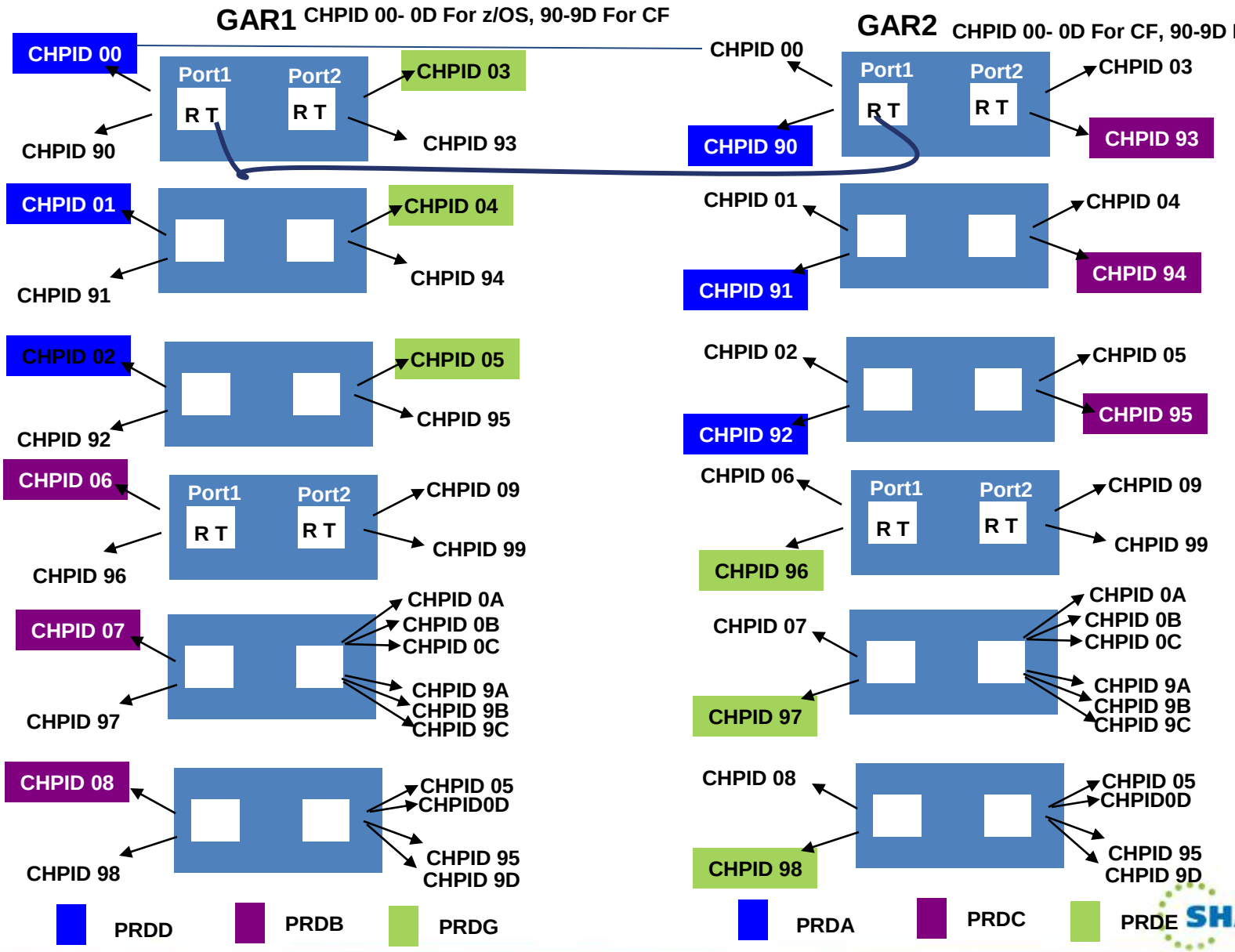




ONLY ONE CHPID FOR EACH PHYSICALL CONNECTION/PORT/LINK

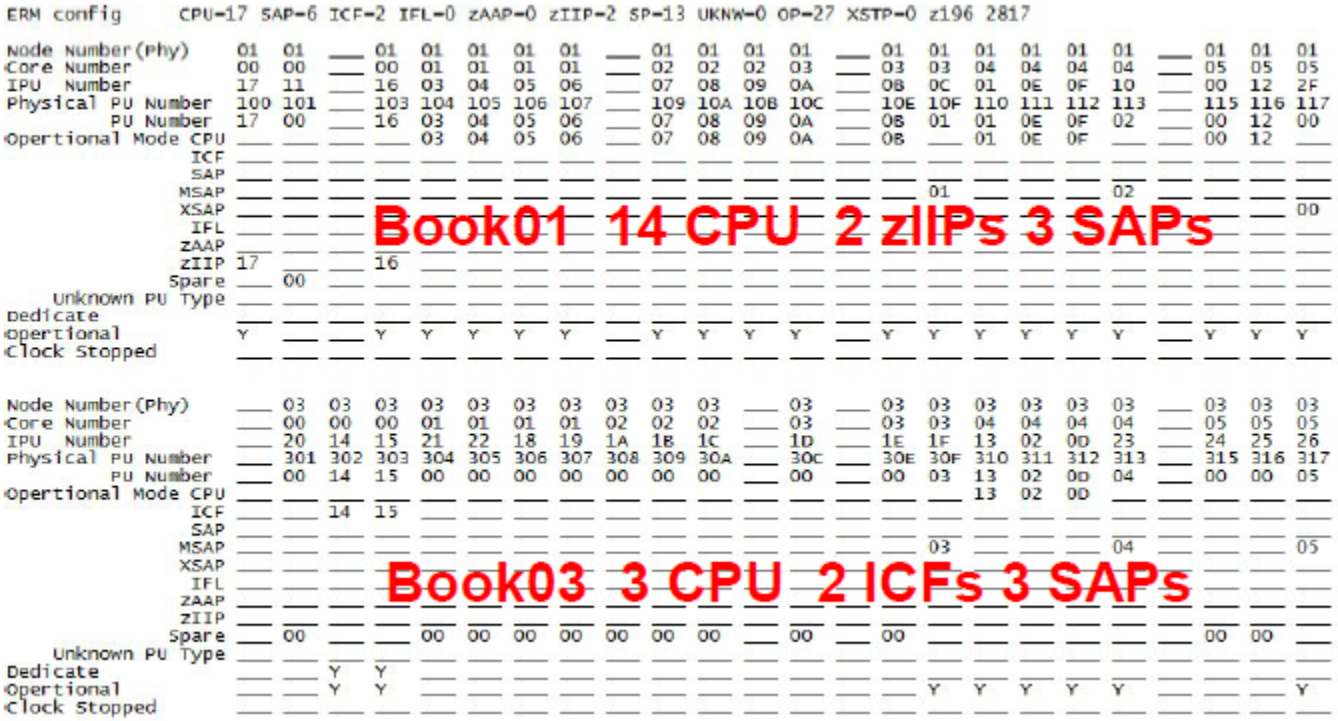
- More subchannels per physical link BY HAVING A CHANCE TO DEFINE MORE THAN ONE CHPID FOR SAME PHYSICALL CONNECTION!
- Subchannel Limit For One CHPID Is Still 7
- Up To 16 CHPIDs Across The Two Ports Of Single InfiniBand Coupling HCA (Adapter)

z196 Infiniband Cables & CHPIDs



z196 Where Are My ICFs ?

GAR1



z196 Where Are My ICFs ?

GAR2

ERM config CPU=17 SAP=6 ICF=2 IFL=0 zAAP=0 zIIP=1 SP=14 UKNW=0 OP=26 XSTP=0 z196 2817

Node Number(Phy)		01	01	01	01	01	01	01	01		01	01	01		01	01	01	01	01	01	01		01	01
Core Number		00	00	00	01	01	01	01	02		02	02	03		03	03	04	04	04	04	05		05	05
IPU Number		0A	16	11	01	04	05	06	07		08	09	02		0B	0C	0D	0E	0F	10	00		12	2F
Physical PU Number		101	102	103	104	105	106	107	108		10A	10B	10C		10E	10F	110	111	112	113	114		116	117
PU Number		0A	16	00	01	04	05	06	07		08	09	02		0B	01	0D	0E	0F	02	00		12	00
Opertional Mode CPU		0A			01	04	05	06	07		08	09	02		0B		0D	0E	0F		00		12	
ICF																								
SAP																								
MSAP																01				02				
XSAP																								00
IFL																								
zAAP																								
zIIP			16																					
Spare				00																				
Unknown PU Type																								
Dedicate																								
Opertional		Y	Y		Y	Y	Y	Y	Y		Y	Y	Y		Y	Y	Y	Y	Y	Y	Y		Y	Y
Clock Stopped																								

Book01 15 CPU 1 zIIP 3 SAPs

Node Number(Phy)	03	03	03		03	03	03	03		03	03	03	03		03	03	03	03	03	03	03		03	03
Core Number	00	00	00		01	01	01	01		02	02	02	03		03	03	04	04	04	04	05		05	05
IPU Number	24	14	15		25	17	18	19		1A	1B	1C	1D		1E	1F	20	21	22	23	13		03	26
Physical PU Number	300	301	302		304	305	306	307		309	30A	30B	30C		30E	30F	310	311	312	313	314		316	317
PU Number	00	14	15		00	00	00	00		00	00	00	00		00	03	00	00	00	04	13		03	05
Opertional Mode CPU																					13		03	
ICF		14	15																					
SAP																								
MSAP																03				04				05
XSAP																								
IFL																								
zAAP																								
zIIP																								
Spare	00				00	00	00	00		00	00	00	00		00		00	00	00					
Unknown PU Type																								
Dedicate		Y	Y																					
Opertional		Y	Y													Y				Y	Y		Y	Y
Clock Stopped																								

Book03 2 CPU 2 ICFs 3 SAPs

Book Configuration – GAR1 CEC- After Upgrades

```
ERM config   CPU=20 SAP=6 ICF=2 IFL=0 ZAAP=0 zIIP=2 SP=10 UKNW=0 OP=30 XSTP=0 z196 zHE 2817
Node Number(Phy)     01  01  ____ 01  01  01  01  01  ____ 01  01  01  01  ____ 01  01  01  01  01  01  ____ 01  01  01
Core Number          00  00  ____ 00  01  01  01  01  ____ 02  02  02  03  ____ 03  03  04  04  04  04  ____ 05  05  05
IPU  Number          1E  13  ____ 15  0D  00  06  04  ____ 07  08  09  03  ____ 0B  0C  01  0E  0F  10  ____ 02  12  2F
Physical PU Number   100 101 ____ 103 104 105 106 107 ____ 109 10A 10B 10C ____ 10E 10F 110 111 112 113 ____ 115 116 117
        PU Number    00  13  ____ 15  0D  00  06  04  ____ 07  08  09  03  ____ 0B  01  01  0E  0F  02  ____ 02  12  00
Opertional Mode CPU  ____ 13  ____ 15  0D  00  06  04  ____ 07  08  09  03  ____ 0B  ____ 01  0E  0F  ____ ____ 02  12  ____
                ICF  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
                SAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               MSAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 01   ____ ____ ____ 02   ____ ____ ____
               XSAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 00
                IFL  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               zAAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               zIIP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
              Spare  00   ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
    Unknown PU Type  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
Dedicate             ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
Opertional           ____ Y    ____ Y    Y    Y    Y    Y    ____ Y    Y    Y    Y    ____ Y    Y    Y    Y    Y    Y    ____ Y    Y    Y
Clock Stopped        ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

Node Number(Phy)     ____ 03  03  03  03  03  03  03  03  03  03  ____ 03  ____ 03  03  03  03  03  03  ____ 03  03  03
Core Number          ____ 00  00  00  01  01  01  01  02  02  02  ____ 03  ____ 03  03  04  04  04  04  ____ 05  05  05
IPU  Number          ____ 18  19  22  24  16  17  21  1B  20  25  ____ 11  ____ 1C  1F  1D  1A  14  23  ____ 05  0A  26
Physical PU Number   ____ 301 302 303 304 305 306 307 308 309 30A ____ 30C ____ 30E 30F 310 311 312 313 ____ 315 316 317
        PU Number    ____ 18  19  00  00  16  17  00  00  00  00  ____ 11  ____ 00  03  00  00  14  04  ____ 05  0A  05
Opertional Mode CPU  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 11  ____ ____ ____ ____ ____ 14  ____ ____ 05  0A  ____
                ICF  ____ ____ ____ ____ ____ 16  17  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
                SAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 04   ____ ____ ____ 05
               MSAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 03   ____ ____ ____ ____ ____ ____
               XSAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
                IFL  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               zAAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               zIIP  ____ 18  19  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
              Spare  ____ ____ ____ 00  00  ____ ____ 00  00  00  00  ____ ____ ____ 00  ____ 00  00  ____ ____ ____ ____ ____ ____
    Unknown PU Type  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
Dedicate             ____ ____ ____ ____ ____ Y    Y    ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
Opertional           ____ Y    Y    ____ ____ Y    Y    ____ ____ ____ ____ ____ Y    ____ ____ Y    ____ ____ Y    Y    ____ Y    Y    Y
Clock Stopped        ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

Number of CPU  = 20
Number of SAP  = 6
        XSAP   = Node Number=01 Physical PU Number=117
Number of CF   = 2
Number of IFL  = 0
Number of zAAP = 0
Number of zIIP = 2
Number of Spare = 10
```

Book Configuration – GAR2 CEC – After Upgrades

```
ERM config   CPU=20 SAP=6 ICF=2 IFL=0 ZAAP=0 zIIP=2 SP=10 UKNW=0 OP=30 XSTP=0 z196 zHE 2817
Node Number(Phy)     01  01  ____ 01  01  01  01  01  ____ 01  01  01  01  ____ 01  01  01  01  01  01  ____ 01  01  01
Core Number          00  00  ____ 00  01  01  01  01  ____ 02  02  02  03  ____ 03  03  04  04  04  04  ____ 05  05  05
IPU  Number          1E  13  ____ 15  0D  00  06  04  ____ 07  08  09  03  ____ 0B  0C  01  0E  0F  10  ____ 02  12  2F
Physical PU Number   100 101 ____ 103 104 105 106 107 ____ 109 10A 10B 10C ____ 10E 10F 110 111 112 113 ____ 115 116 117
        PU Number    00  13  ____ 15  0D  00  06  04  ____ 07  08  09  03  ____ 0B  01  01  0E  0F  02  ____ 02  12  00
Opertional Mode CPU  ____ 13  ____ 15  0D  00  06  04  ____ 07  08  09  03  ____ 0B  ____ 01  0E  0F  ____ ____ 02  12  ____
                ICF  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
                SAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               MSAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 01   ____ ____ ____ 02   ____ ____ ____
               XSAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 00
                IFL  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               zAAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               zIIP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
              Spare  00   ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
    Unknown PU Type  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
Dedicate             ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
Opertional           ____ Y   ____ Y   Y   Y   Y   Y   ____ Y   Y   Y   Y   ____ Y   Y   Y   Y   Y   Y   ____ Y   Y   Y
Clock Stopped        ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

Node Number(Phy)     ____ 03  03  03  03  03  03  03  03  03  03  ____ 03  ____ 03  03  03  03  03  03  ____ 03  03  03
Core Number          ____ 00  00  00  01  01  01  01  02  02  02  ____ 03  ____ 03  03  04  04  04  04  ____ 05  05  05
IPU  Number          ____ 18  19  22  24  16  17  21  1B  20  25  ____ 11  ____ 1C  1F  1D  1A  14  23  ____ 05  0A  26
Physical PU Number   ____ 301 302 303 304 305 306 307 308 309 30A ____ 30C ____ 30E 30F 310 311 312 313 ____ 315 316 317
        PU Number    ____ 18  19  00  00  16  17  00  00  00  00  ____ 11  ____ 00  03  00  00  14  04  ____ 05  0A  05
Opertional Mode CPU  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 11  ____ ____ ____ ____ ____ 14  ____ ____ 05  0A  ____
                ICF  ____ ____ ____ ____ ____ 16  17  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
                SAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 04   ____ ____ ____ 05
               MSAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ 03   ____ ____ ____ ____ ____ ____ ____ ____
               XSAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
                IFL  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               zAAP  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
               zIIP  ____ 18  19  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
              Spare  ____ ____ ____ 00  00  ____ ____ 00  00  00  00  ____ ____ ____ 00  ____ 00  00  ____ ____ ____ ____ ____ ____
    Unknown PU Type  ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
Dedicate             ____ ____ ____ ____ ____ Y   Y   ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____
Opertional           ____ Y   Y   ____ ____ Y   Y   ____ ____ ____ ____ ____ Y   ____ ____ Y   ____ ____ ____ Y   Y   ____ Y   Y   Y
Clock Stopped        ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

Number of CPU  = 20
Number of SAP  = 6
      XSAP  = Node Number=01 Physical PU Number=117
Number of CF   = 2
Number of IFL  = 0
Number of zAAP = 0
Number of zIIP = 2
Number of Spare = 10
```

GT Parallel Sysplex Configuration - Structures

DSNPD01_GBP0
DSNPD01_GBP1
DSNPD01_GBP16K0
DSNPD01_GBP16K1
DSNPD01_GBP2
DSNPD01_GBP21
DSNPD01_GBP22
DSNPD01_GBP23
DSNPD01_GBP24
DSNPD01_GBP31
DSNPD01_GBP32
DSNPD01_GBP32K
DSNPD01_GBP33
DSNPD01_GBP34
DSNPD01_GBP5
DSNPD01_GBP6
DSNPD01_GBP7
DSNPD01_GBP8K0
DSNPD01_LOCK1
DSNPD01_SCA

DSNPDRM_GBP0
DSNPDRM_GBP1
DSNPDRM_GBP2
DSNPDRM_GBP8K0
DSNPDRM_LOCK1
DSNPDRM_SCA

PQS1APPLSTR
PQS1CSQ_ADMIN
PQS1FFMCSTR
PQS1FFMDSTR
PQS1LOGOSTR
PQS1OLASTR
PQS1OTPSTR
PQS1SMSSTR
PQS1SYSPSTR
PQS1UTLSTR

EZBEPORT
EZBEPORT0111
EZBEPORT0113
ISTGENERIC

DFHNCLS_PRODNC1
DFHXQLS_PRODTSQ1
LOG_DFHLOG_WUI
LOG_DFHSHUNT_WUI

IXCSIG1
IXCSIG11
IXCSIG2
IXCSIG21
IXCSIG3
IXCSIG31
IXCSIG4
IXCSIG5
IXCSIG6
IXCSIG7

CKPT1
RLS_APL1
RLS_APL2
RRSSTR1
SYSARC_HSMPP_RCL
SYSIGGCAS_ECS
SYSTEM_OPERLOG
SYSZWLM_0E162817
SYSZWLM_0E262817
HSA_LOG
HZS_HEALTHCHKLOG
IBMBDG
IGWLOCK00
ISGLOCK

TOPSTR1

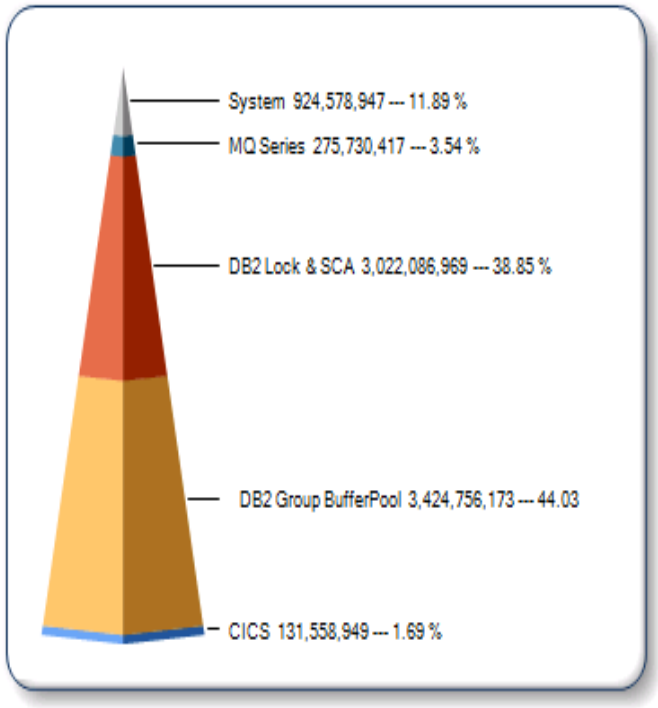
GT-CF Configuration – z196

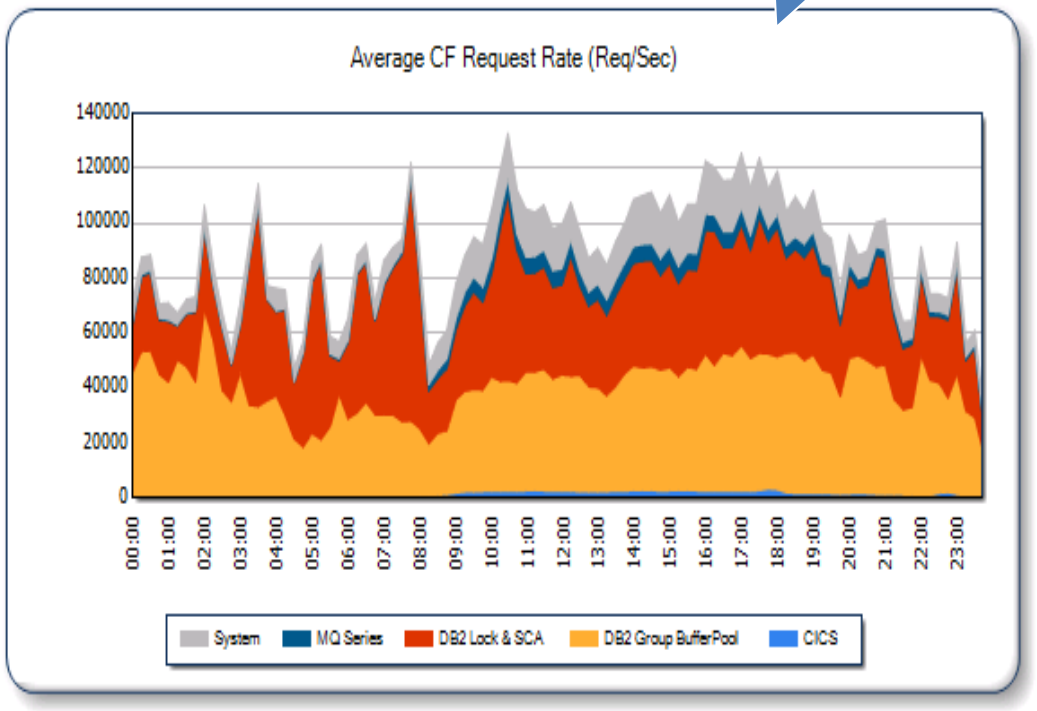
- 179 Structures Defined In CFRM Policy
- 27205 MB Storage In Each CF – Today 46 GB

RMF MON I

AVERAGE CF REQUEST RATE / WORKLOAD

20-06-2011





There are now customers processing more than 500,000 requests per second in each CF, and benchmarks in IBM have driven nearly 1,500,000 requests a second to a single CF.

General Information About PSIFB

Infiniband Technology - Drivings Of Innovation (WHY?)

Problem

Efficient Systems Must Provide Balance Between

- **CPU Performance**
- **Memory Bandwidth**
- **I/O Capabilities**

Semiconductor Technology Evolves Much Faster Than I/O Interconnect Speed.

New Technology is needed to keep up with the speed of processors

Solution

In 1999 Two Competing I/O Standards called

- **Future I/O (Developed By Compaq.IBM,HP)**
- **NextGeneration I/O (Developed By Intel,Microsoft,Sun)**

Merged Into Unified I/O Standard Called INFINIBAND

http://www.infinibandta.org/itinfo/IL

InfiniBand offers a powerful interconnect architecture that by its nature is better able to scale with increasing processor speeds. Up to 120 Gbps

Infiniband Technology - Advantages

- ✓ Superior Performance – Up To 120 Gbps
- ✓ Reduced Complexity
- ✓ Highest Interconnect Efficiency
- ✓ Reliable & Stable Connection

- ❑ First Used As Connection Between Books & I/O Cage,starting with z10.
- ❑ With z10 and supported by z9 also, it started to be used as CF link.

BUT !

PSIFB Configuration Disadvantages

In Every Document –There Is One NOTE

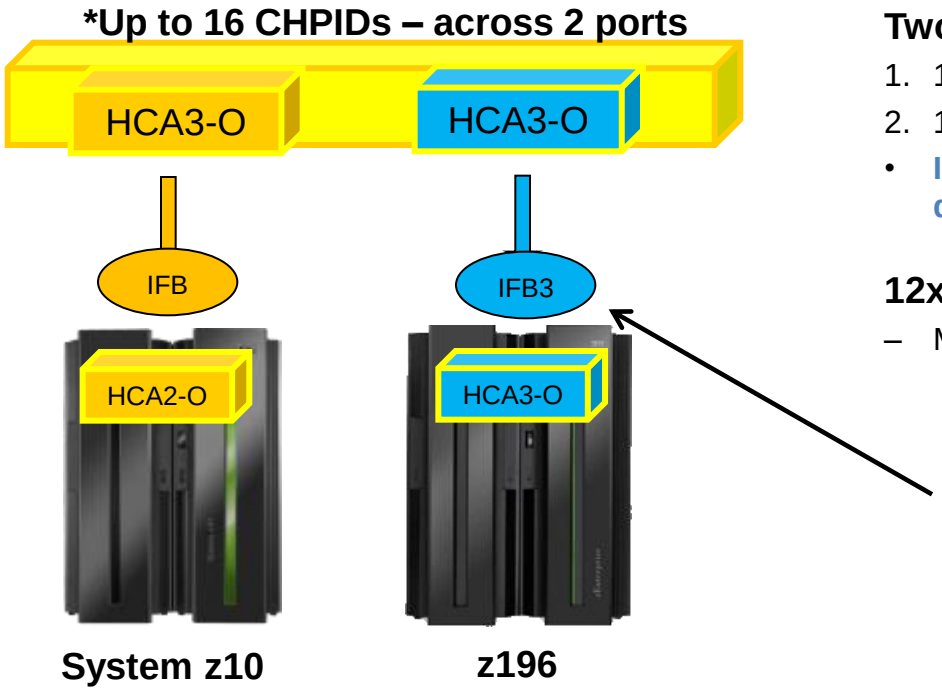
Note: **The InfiniBand link data rates of 6 GBps, 3 GBps, 2.5 Gbps, or 5 Gbps do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. With InfiniBand coupling links, while the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

Note: The InfiniBand link data rate of 6 GBps or 3 GBps does not represent the performance of the link. The actual performance depends on many factors, such as latency through the adapters, cable lengths, and the type of workload. With InfiniBand coupling links, while the link data rate might be higher than that of ICB links, the service times of coupling operations are greater.

So As Expected IBM Created A New Adapter & Protocol

New PSIFB Protocol & Infiniband Fanout Cards

New 12x InfiniBand fanout cards, exclusive to z196 and z114



Attachment to System z9 HCA1 not supported

Two protocols (IFB & IFB3)

1. 12x IFB = HCA3-O to HCA2-O
2. 12x IFB3 = HCA3-O to HCA3-O (see below)

- **Improved service times, 12x IFB3 service times are designed to be 40% faster than 12x IFB**

12x IFB3 protocol activation requirements

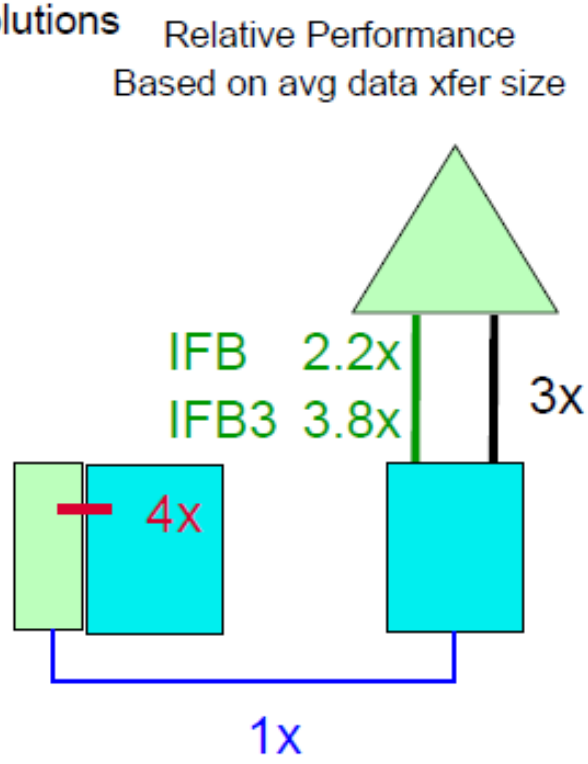
– Maximum of **four** CHPIDs per HCA3-O port

- **If more than four CHIPDs are defined per port, links will run at normal 12x IFB service times**
- **IFB3 protocol activated as long as 4 CHPIDs or less are defined. No configuration settings required.**
- Performance considerations may reduce the number of CHPIDs per port

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

Infiniband As Coupling Link Choice - PSIFB

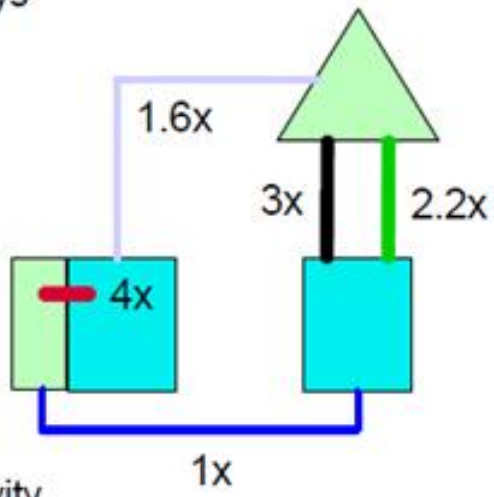
- ISC (Inter-System Channel)
  - Fiber optics
  - I/O Adapter card
  - 10km and longer distances with qualified WDM solutions
- ICB (Integrated Cluster Bus)
  - Copper cable plugs close to memory bus
  - 10 meter max length
  - Not available on z196
- IC (Internal Coupling Channel)
  - Microcode - no external connection
  - Only between partitions on same processor
- 12X IFB and 12X IFB3 (intro z196 GA2)
  - 150 meter max distance optical cabling
  - Supports multiple CHPIDs per physical link
  - Multiple CF partitions can share physical link



PSIFB Configuration Advantages

PSIFB Configuration Advantages

- Pure Capacity
  - 1 12x PSIFB replaces 1 ICB4
  - 1 12x PSIFB replaces 4 ISC3s
- Eliminating subchannel and path delays
  - Often >2 ICB4s configured not for capacity but for extra subchannels/paths to eliminate delays
  - 2 12x PSIFB links with multiple CHPIDs can replace >2 ICB4s in this case
- Multiple sysplexes sharing hardware
  - Production, development, test sysplexes may share hardware – each needs own ICB4 or ISC3 links
  - 2 PSIFB links with multiple CHPIDs can replace >2 ICB4s or ISC3s in this case
- Multiple CHPID recommendations
  - Max 16 per HCA (2 ports per HCA)
    - Use up to all 16 for lightly loaded connectivity
    - Limit to use up to 8 per HCA for heavy loads



Life Of A CF Request

CF Syncronous Request Flow-1

REQUESTED AS

1. Formats a Request To A Structure by Issuing Related XES Macros

IXLCACHE/IXLLSTE/IXLLSTM

1a. Pass Buffers For Read/Write Request

1b. Builds a Parameter List

Common Area SQA

MOB | MCB | MRB

Busy Subchannel

Free Subchannel

CF REQUEST WORK ELEMENT

Data Area Addresses
Connector Information
Structure Information
CF Request Obtained From Requester's IXLCACHE Parameter List
Requester's MODE Information

Subchannel#/Device, CHPID information

Subchannel

Subchannel#

Common Area XES Modules

2. Branch To XES Service Module

XES SERVICE Module

3. Builds a CF REquest Work Element
Verifies callers parameters
Obtains Real Storage Addresses for Data Areas
Establish Central Storage Binds for pageable storage to keep the pages in central storage while HW is referencing it.

4. Branch To XES Module & Passes CF Request Work Element

XES Module

5. Executes CF Request Work Element

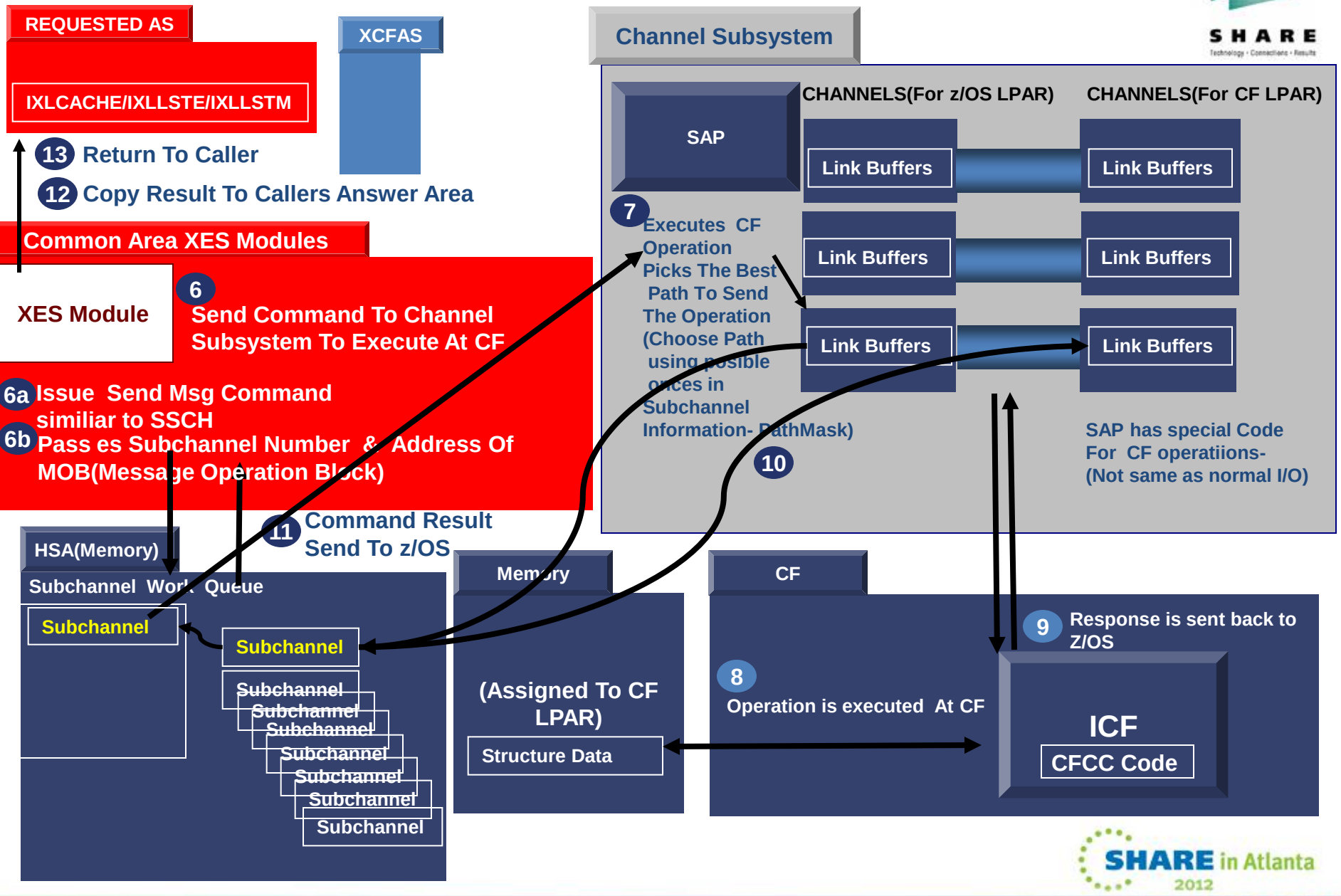
5a. Determine Which CF To Target CF Request

5b. Obtain Subchannel For That CF

5c. Determine if Request Should be processed ASYNC or SYNC based on Heuristic & Requestors Specified MODE

SASH Table-Dynamic Decisions

CF Syncronous Request Flow-2



CF Asyncronous Request Flow -1

REQUESTED AS

1. Formats a Request To A Structure by Issuing Related XES Macros
   - IXLCACHE/IXLLSTE/IXLLSTM
- 1a Pass Buffers For Read/Write Request
- 1b Builds a Parameter List

Common Area SQA

MOB | MCB | MRB

- Busy Subchannel
- Free Subchannel

CF REQUEST WORK ELEMENT

Data Area Addresses
Connector Information
Structure Information
CF Request Obtained From Requester's IXLCACHE Parameter List
Requester's MODE Information

Subchannel#/Device, CHPID information

Subchannel

Subchannel#

Common Area XES Modules

2. Branch To XES Service Module

XES SERVICE Module

3. Builds a CF REquest Work Element
   Verifies callers parameters
   Obtains Real Storage Addresses for Data Areas
   Establish Central Storage Binds for pageable storage to keep the pages in central storage while HW is referencing it.

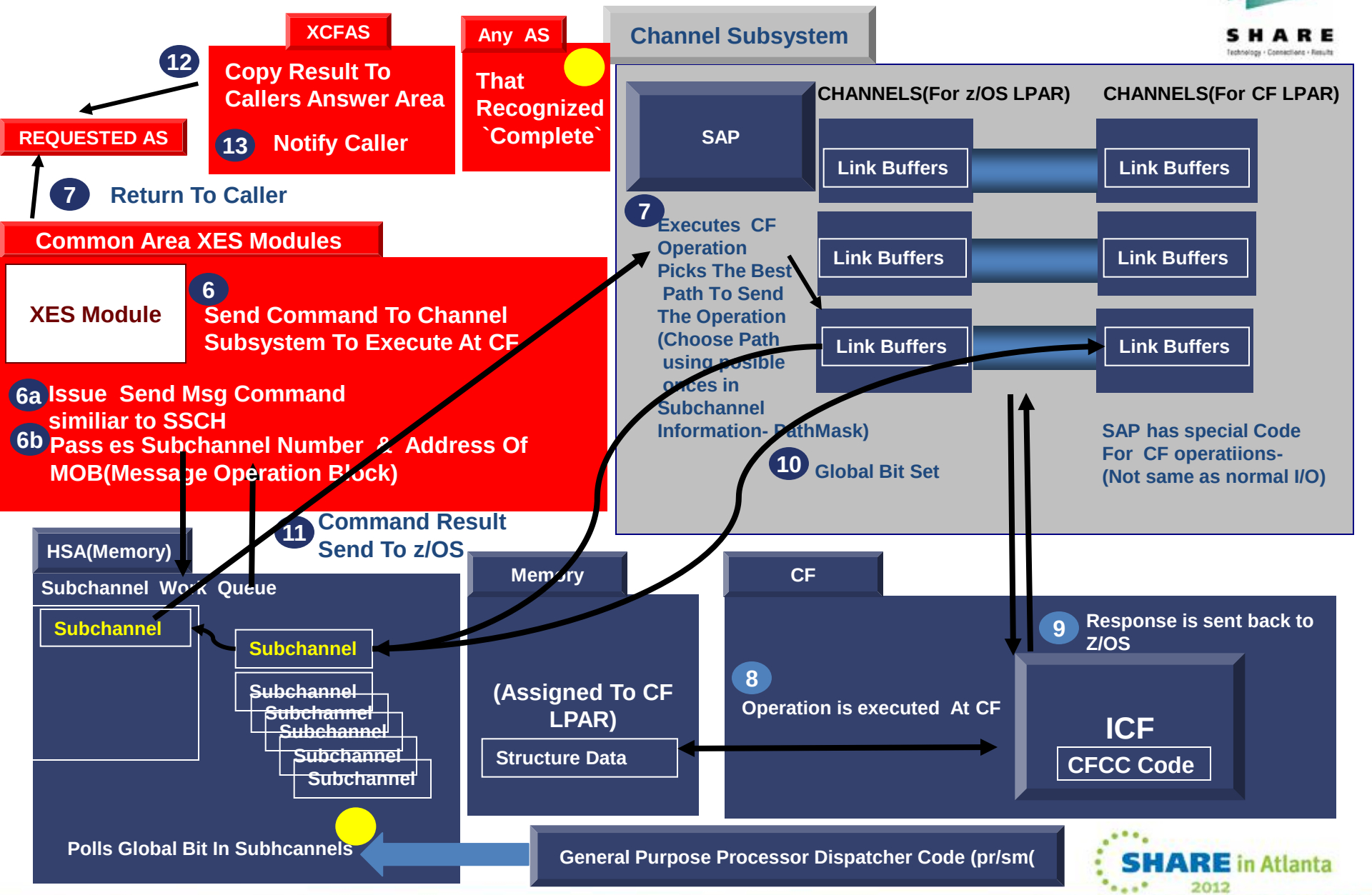
4. Branch To XES Module & Passes CF Request Work Element

XES Module

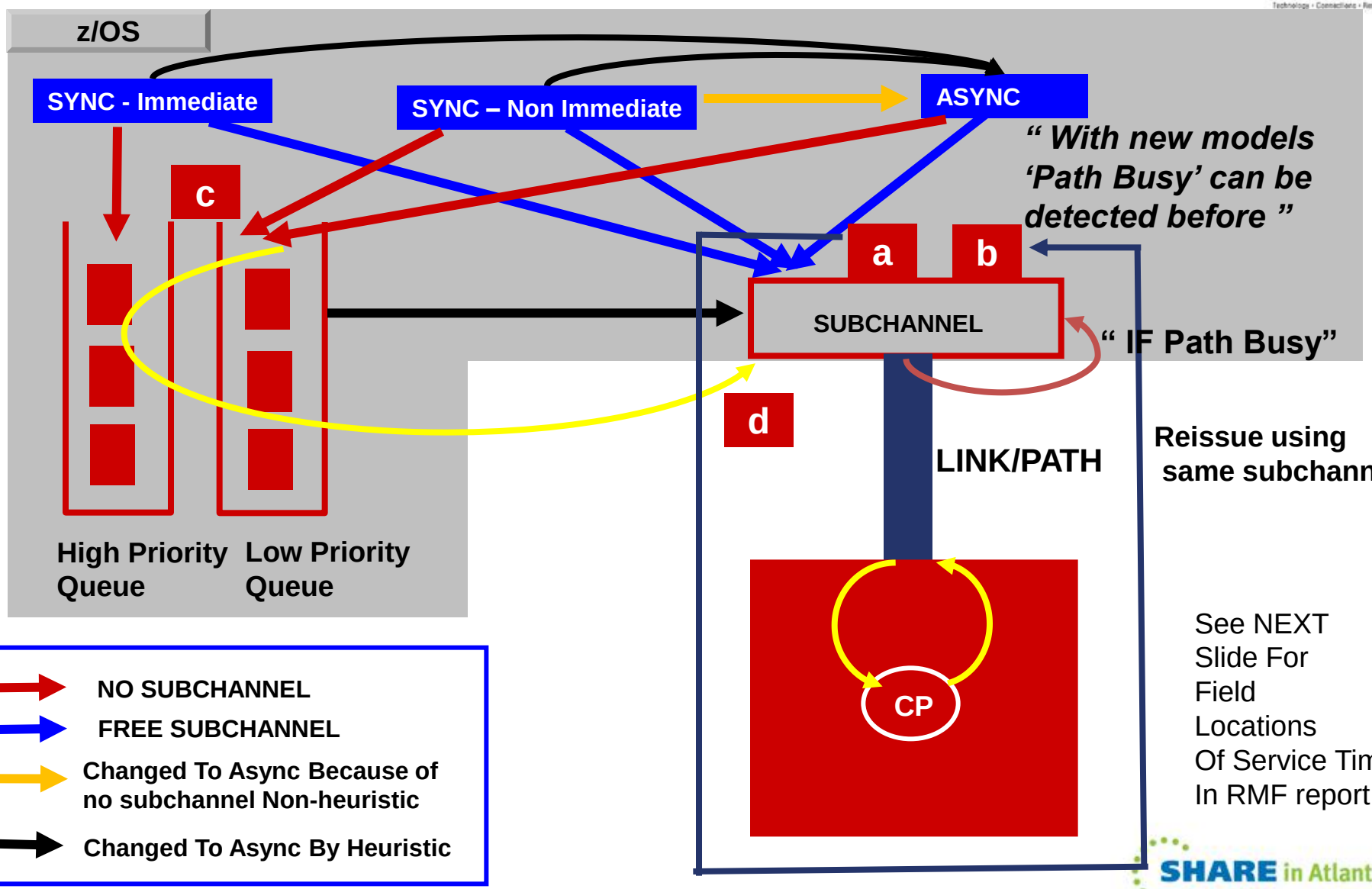
5. Executes CF Request Work Element
- 5a Determine Which CF To Target CF Request
- 5b Obtain Subchannel For That CF
- 5c Determine if Request Should be processed ASYNC or SYNC based on Heuristic & Requestors Specified MODE

SASH (Sync/Async Heuristic)Table-Dynamic Decisions

CF Asyncronous Request Flow -2



CF Request Types & Cases



CF Request Types & Cases – RMF Report

b – a

d – c

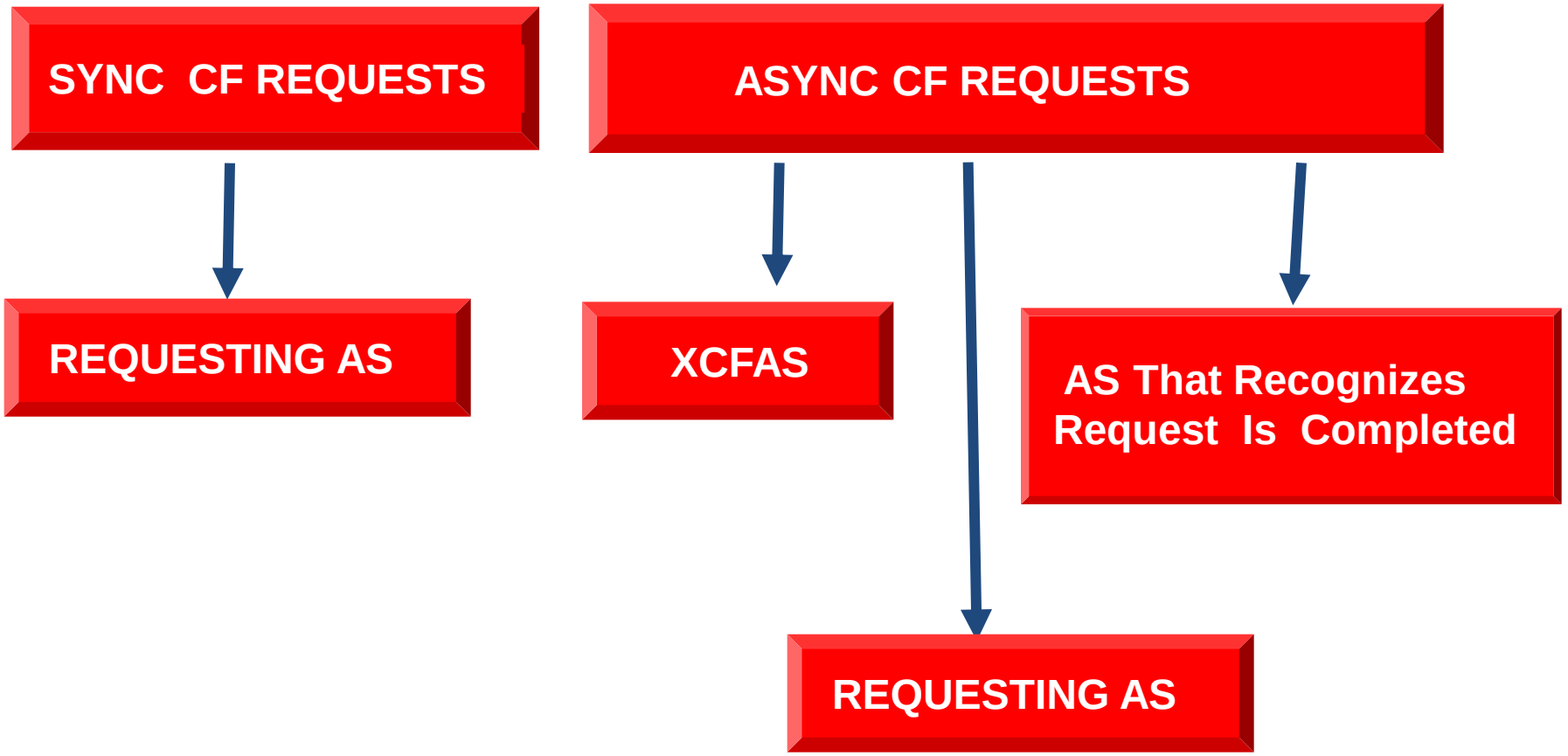
Service Time is calculated as

Delayed Time is calculated as

SUBCHANNEL ACTIVITY

SYSTEM NAME	Ö REQ TOTAL AVG/SEC	CF TYPE	LINKS GEN	LINKS USE	PTH BUSY	REQUESTS	Ö REQ	SERVICE TIME(MIC) AVG	SERVICE TIME(MIC) STD_DEV	DELAYED REQUESTS	Ö REQ	% OF REQ	AVG TIME(MIC) /DEL	AVG TIME(MIC) STD_DEV	AVG TIME(MIC) /ALL
PRDA	2391K	CIB	3	3	0	SYNC	1579K	13.0	4.7	LIST/CACHE	358	0.0	42.5	33.6	0.0
	2656.3	SUBCH	42	21		ASYNC	788047	111.8	161.0	LOCK	0	0.0	0.0	0.0	0.0
						CHANGED	0	INCLUDED IN ASYNC		TOTAL	358	0.0			
						UNSUCC	0	0.0	0.0						
PRDB	5869K	ICP	4	4	1935	SYNC	4234K	3.9	34.9	LIST/CACHE	415	0.0	789.0	601.7	0.1
	6521.0	SUBCH	56	28		ASYNC	1589K	41.7	395.3	LOCK	13	0.0	207.0	159.9	0.0
						CHANGED	413	INCLUDED IN ASYNC		TOTAL	428	0.0			
						UNSUCC	0	0.0	0.0						
PRDC	6364K	CIB	3	3	0	SYNC	4671K	12.9	4.0	LIST/CACHE	3052	0.1	707.5	573.0	0.8
	7071.1	SUBCH	42	21		ASYNC	1645K	72.8	88.7	LOCK	60	0.0	115.5	125.4	0.0
						CHANGED	2492	INCLUDED IN ASYNC		TOTAL	3112	0.0			
						UNSUCC	0	0.0	0.0						
PRDD	11892K	ICP	4	4	2718	SYNC	9162K	4.1	31.9	LIST/CACHE	582	0.0	961.6	1544	0.1
	13213	SUBCH	56	28		ASYNC	2757K	36.0	475.6	LOCK	86	0.0	393.3	1096	0.0
						CHANGED	627	INCLUDED IN ASYNC		TOTAL	668	0.0			

CPU COST OF CF REQUESTS



Sync/Async Conversion

Sync/Async Conversion

NON-HEURISTIC

- Subchannel Busy Condition
- Path Busy Condition
- Serialized List or Lock Contention

HEURISTIC

Introduced with z/OS v1r2...

- CF Link Technology
- Types Of Workload – Variable Workload Amount
- Range Of CF Utilization,Shared CP or not,...
- Actual Observed Sync Request Service Time
- Amount Of Data That Needs To Be Transfered
- Other items that effect CF response ex:Distance
- Moving Weighted Averages Of Actual CF Requests
- Every 1 of N Request not converted and send as Sync

How To Display sync/async Conversion Threshold Value

With z/OS V1R11 (APAR OA28603 for z/OS v1r8 and above)

```
PRDA D XCF,C
IXC357I  15.21.03  DISPLAY XCF 494
SYSTEM PRDA DATA
   INTERVAL    OPNOTIFY     MAXMSG    CLEANUP       RETRY   CLASSLEN
        165         165       2000         15          10        956

   SSUM ACTION  SSUM INTERVAL  SSUM LIMIT     WEIGHT  MEMSTALLTIME
       ISOLATE             0          60        100            NO

   CFSTRHANGTIME
              NO

   PARMLIB USER INTERVAL:      85
   DERIVED SPIN INTERVAL:     165
   PARMLIB USER OPNOTIFY:      87

   MAX SUPPORTED CFLEVEL: 17

   MAX SUPPORTED SYSTEM-MANAGED PROCESS LEVEL: 17

   SIMPLEX SYNC/ASYNC THRESHOLD:                26
   DUPLEX SYNC/ASYNC THRESHOLD:                 26
   SIMPLEX LOCK SYNC/ASYNC THRESHOLD:           26
   DUPLEX LOCK SYNC/ASYNC THRESHOLD:            26

   CF REQUEST TIME ORDERING FUNCTION: INSTALLED

   SYSTEM STATUS DETECTION PARTITIONING PROTOCOL ELIGIBILITY:
     SYSTEM CANNOT TARGET OTHER SYSTEMS.
       REASON: SYSPLEX COUPLE DATA SET NOT FORMATTED FOR THE PROTOCOL
     SYSTEM IS NOT ELIGIBLE TO BE TARGETED BY OTHER SYSTEMS.
       REASON: SYSPLEX COUPLE DATA SET NOT FORMATTED FOR THE PROTOCOL
```

Related To Heuristic Decision

How To Display sync/async Conversion Threshold Value

```
PRDE D XCF,C
IXC357I  15.22.57  DISPLAY XCF 859
SYSTEM PRDE DATA
   INTERVAL    OPNOTIFY      MAXMSG     CLEANUP       RETRY    CLASSLEN
        165         165        2000          15          10         956

   SSUM ACTION  SSUM INTERVAL  SSUM LIMIT     WEIGHT  MEMSTALLTIME
       ISOLATE             0          60          1            NO

   CFSTRHANGTIME
              NO

   PARMLIB USER INTERVAL:      85
   DERIVED SPIN INTERVAL:     165
   PARMLIB USER OPNOTIFY:      87

   MAX SUPPORTED CFLEVEL: 17

   MAX SUPPORTED SYSTEM-MANAGED PROCESS LEVEL: 17

   SIMPLEX SYNC/ASYNC THRESHOLD:                26
   DUPLEX SYNC/ASYNC THRESHOLD:                 26
   SIMPLEX LOCK SYNC/ASYNC THRESHOLD:           26
   DUPLEX LOCK SYNC/ASYNC THRESHOLD:            27

   CF REQUEST TIME ORDERING FUNCTION: INSTALLED

   SYSTEM STATUS DETECTION PARTITIONING PROTOCOL ELIGIBILITY:
     SYSTEM CANNOT TARGET OTHER SYSTEMS.
       REASON: SYSPLEX COUPLE DATA SET NOT FORMATTED FOR THE PROTOCOL
     SYSTEM IS NOT ELIGIBLE TO BE TARGETED BY OTHER SYSTEMS.
       REASON: SYSPLEX COUPLE DATA SET NOT FORMATTED FOR THE PROTOCOL
```

Related To Heuristic Decision

Performance Differences

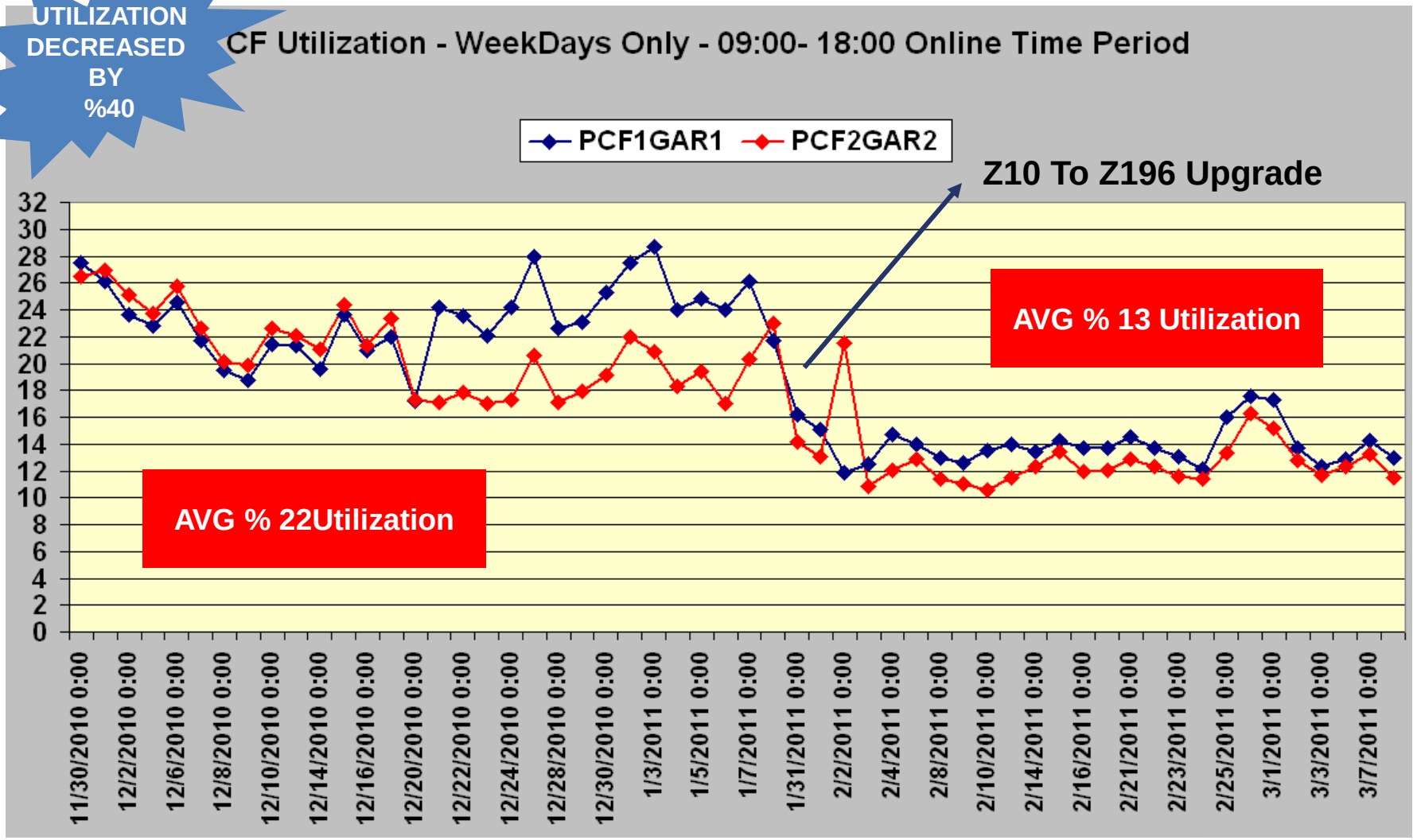
IBM - Sync Service Times For Different CF Link Types

	ISC3	1x IB	12x IFB	ICB4	IC
z10					
Lock	20-30	14-18	11-15	8-12	3-8
List/Cache (4k)	25-40	18-25	15-20	10-16	6-10
z196					
Lock	20-30	14-17	10-14	NA	2-8
List/Cache (4k)	25-40	16-25	14-18	NA	4-9

IBM POK CF Performance Group

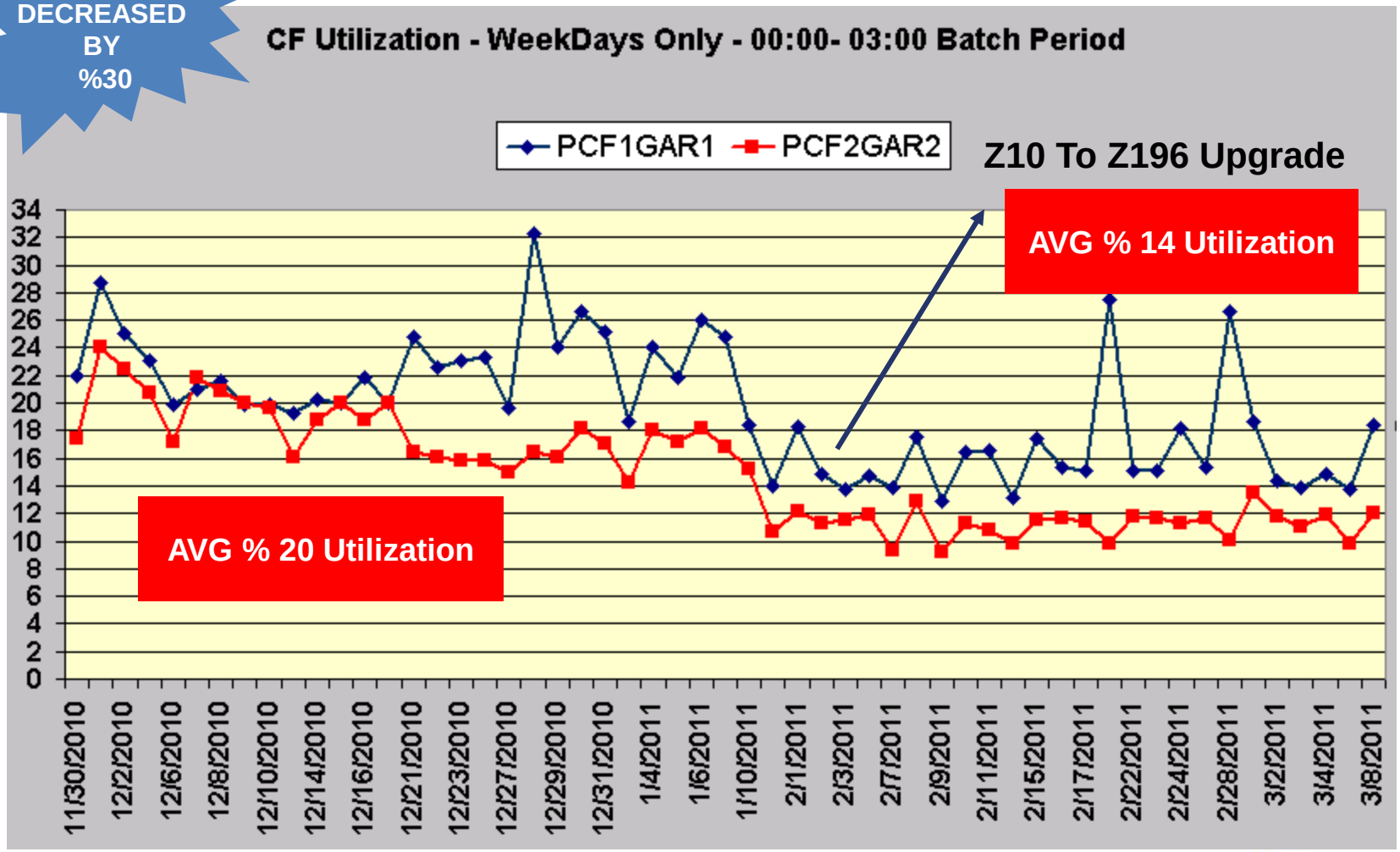
Z10 ICFs To z196 ICFs Utilization Change

UTILIZATION DECREASED BY %40



Z10 ICFs To z196 ICFs Utilization Change

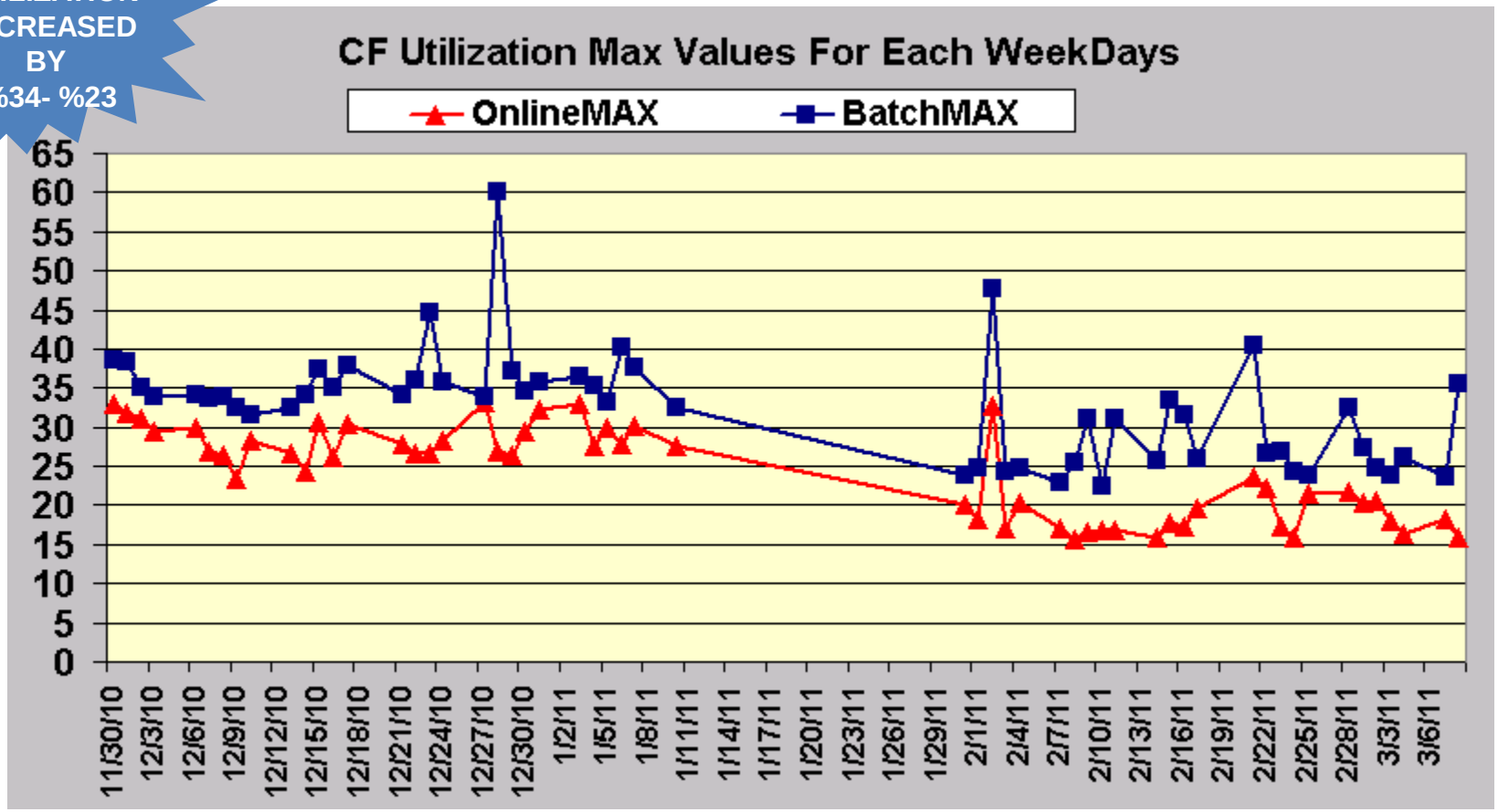
UTILIZATION DECREASED BY %30



Z10 ICFs To z196 ICFs MAX Utilization Change

Av[illegible]es May Not Be Our Concern –It is needed to check the max usages

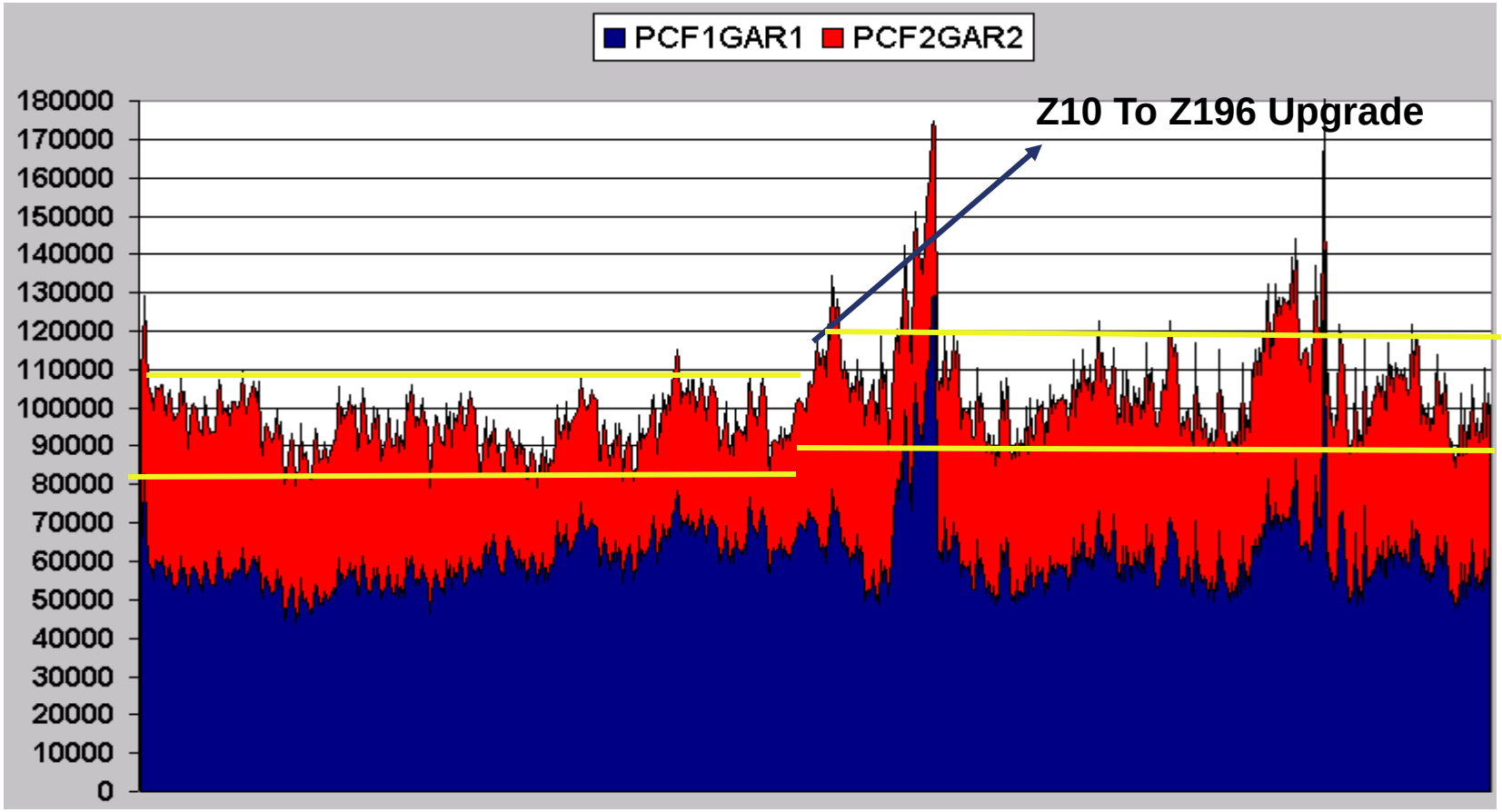
UTILIZATION DECREASED BY %34- %23



Online Time 8 -12 14-18 29 To 19
Batch Period 0-8 & 18-24 36 To 28

Z10 & z196 CF Request Rates Online

Between 90.000 – 110.000 During 14:00 – 17:00 For Each Week-Day

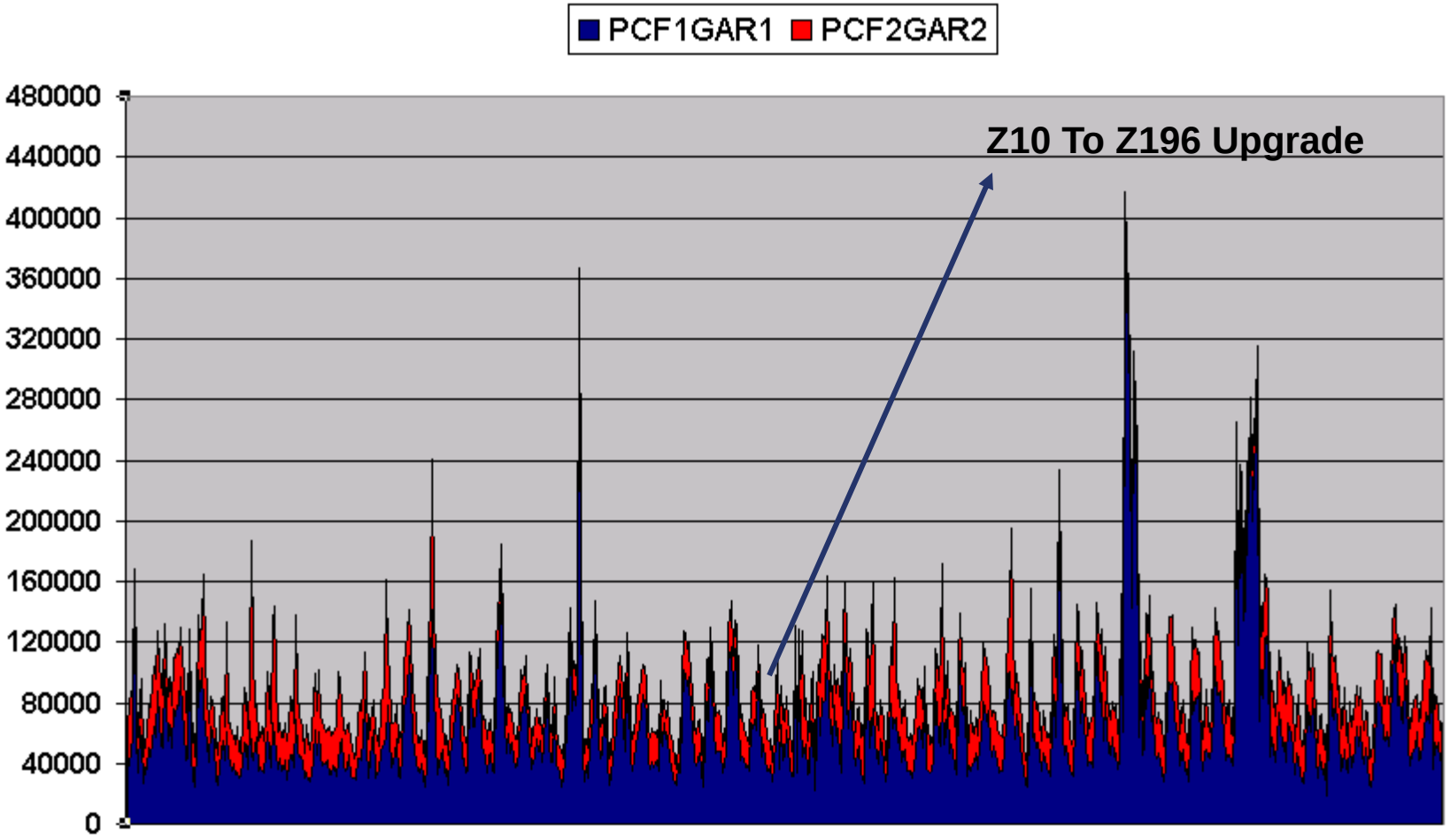


30-11-2010 **7-3-2011**

Average 96000 To 107300 CF Request Rate increased by %12

Z10 & z196 CF Request Rates Batch

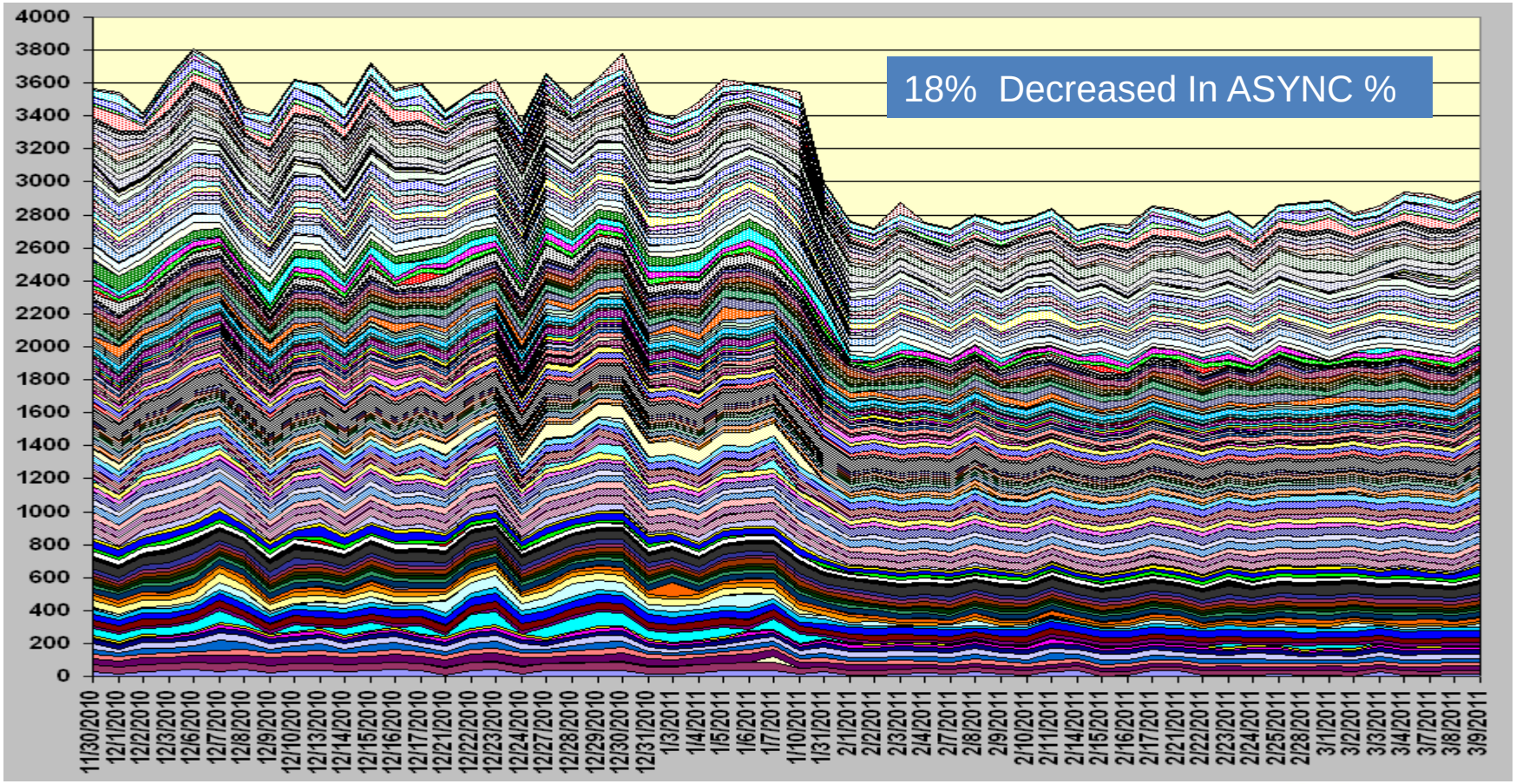
Range is wider than online 80.000 – 120.000 During 00:00 – 03:00 For Each Week-Day



Asyncronous Request %

Async Request Percentage For DB2 GBP Structures 09:00-18:00 – Weekdays Only

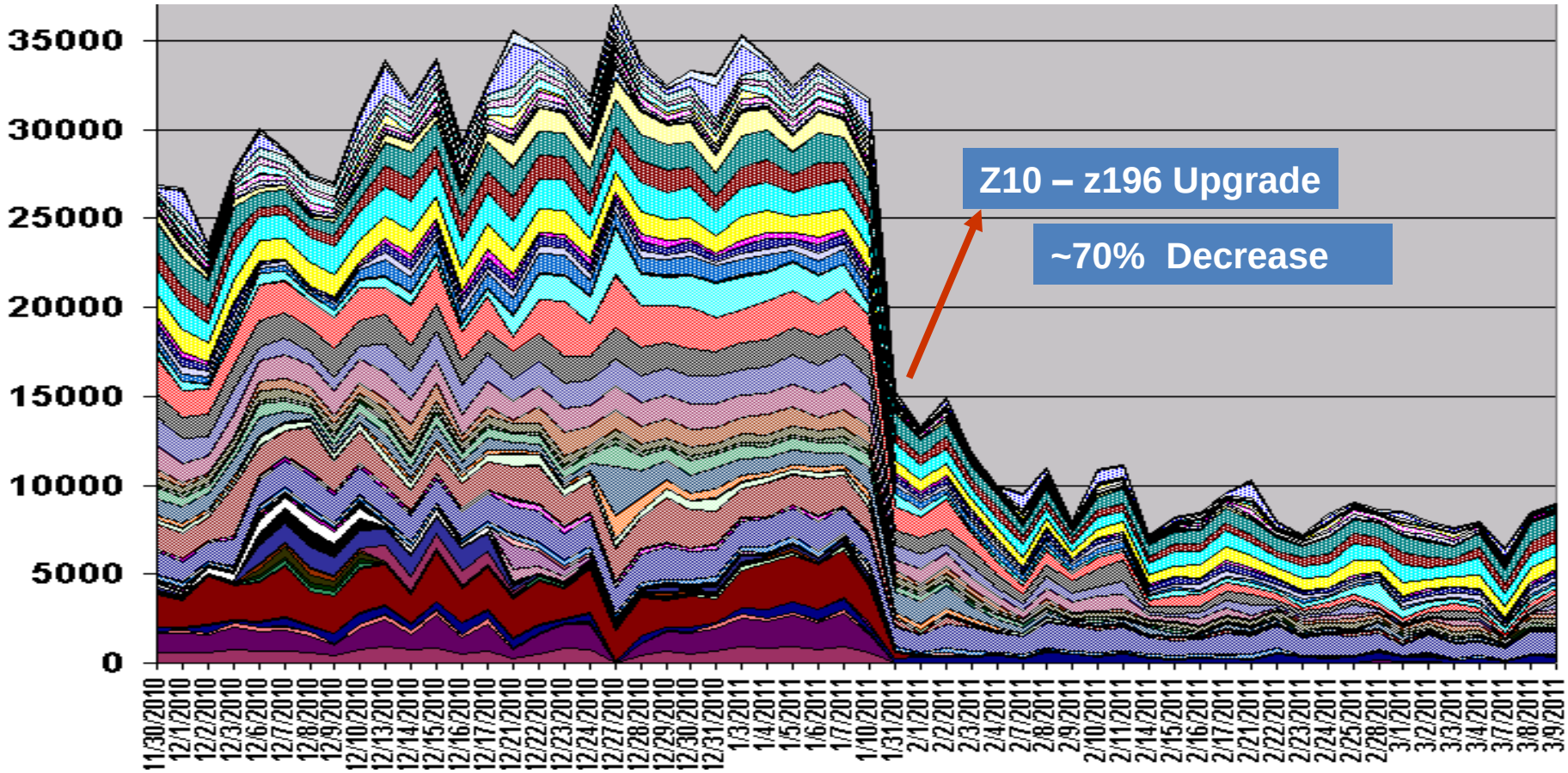
Stacked Area Graph



Number Of Asynchronous Requests Converted by Non-Heuristic

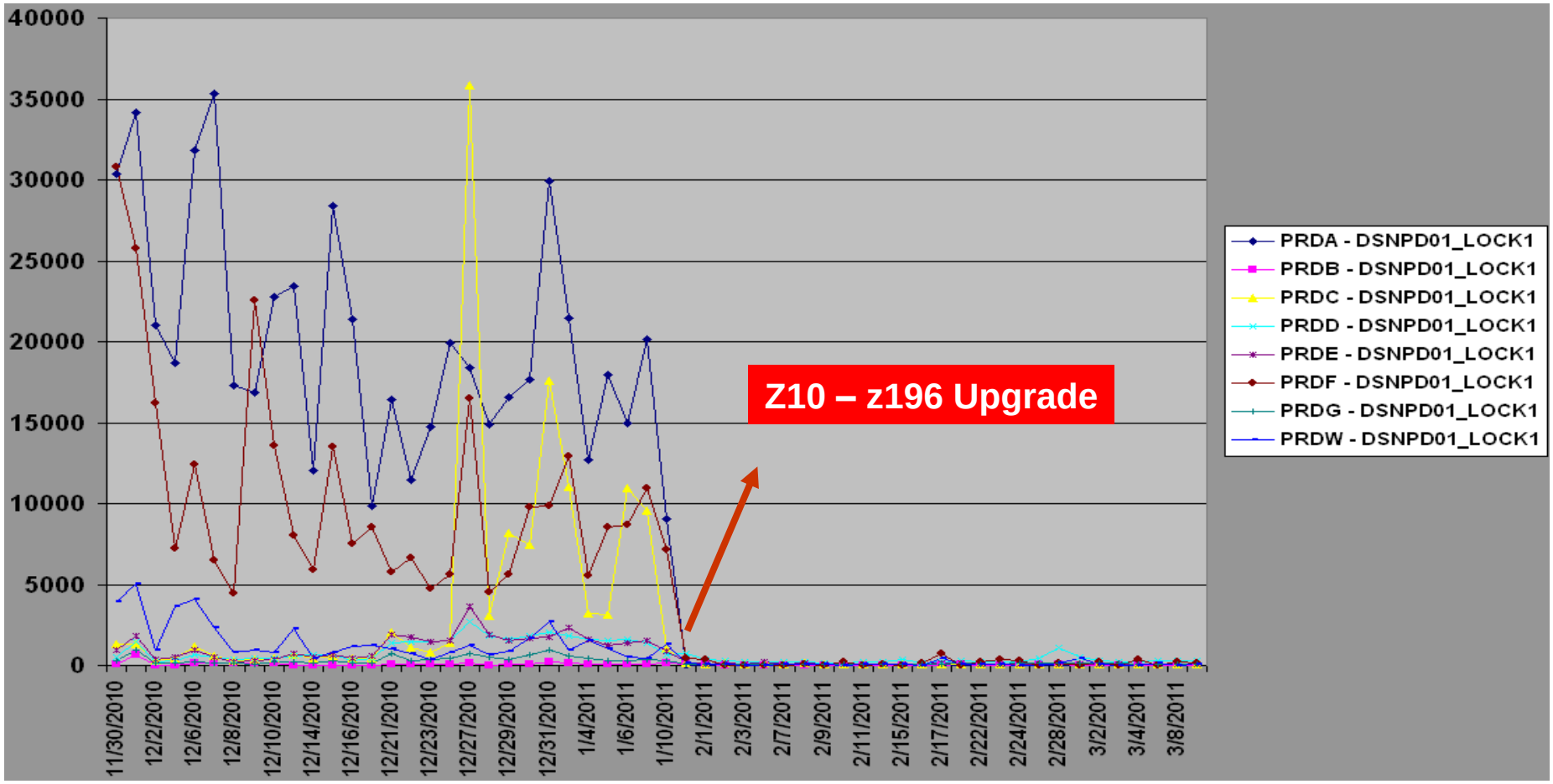
Requests Changed To async by Non-heuristic Method For DB2 GBP Structures
09:00-18:00 – Weekdays Only - Stacked Area Graph

CHANGED Request Fields in RMF Report- `No subchannel ` reason



Configuration Change Effect on # Of Async Requests For DB2 Lock Structure

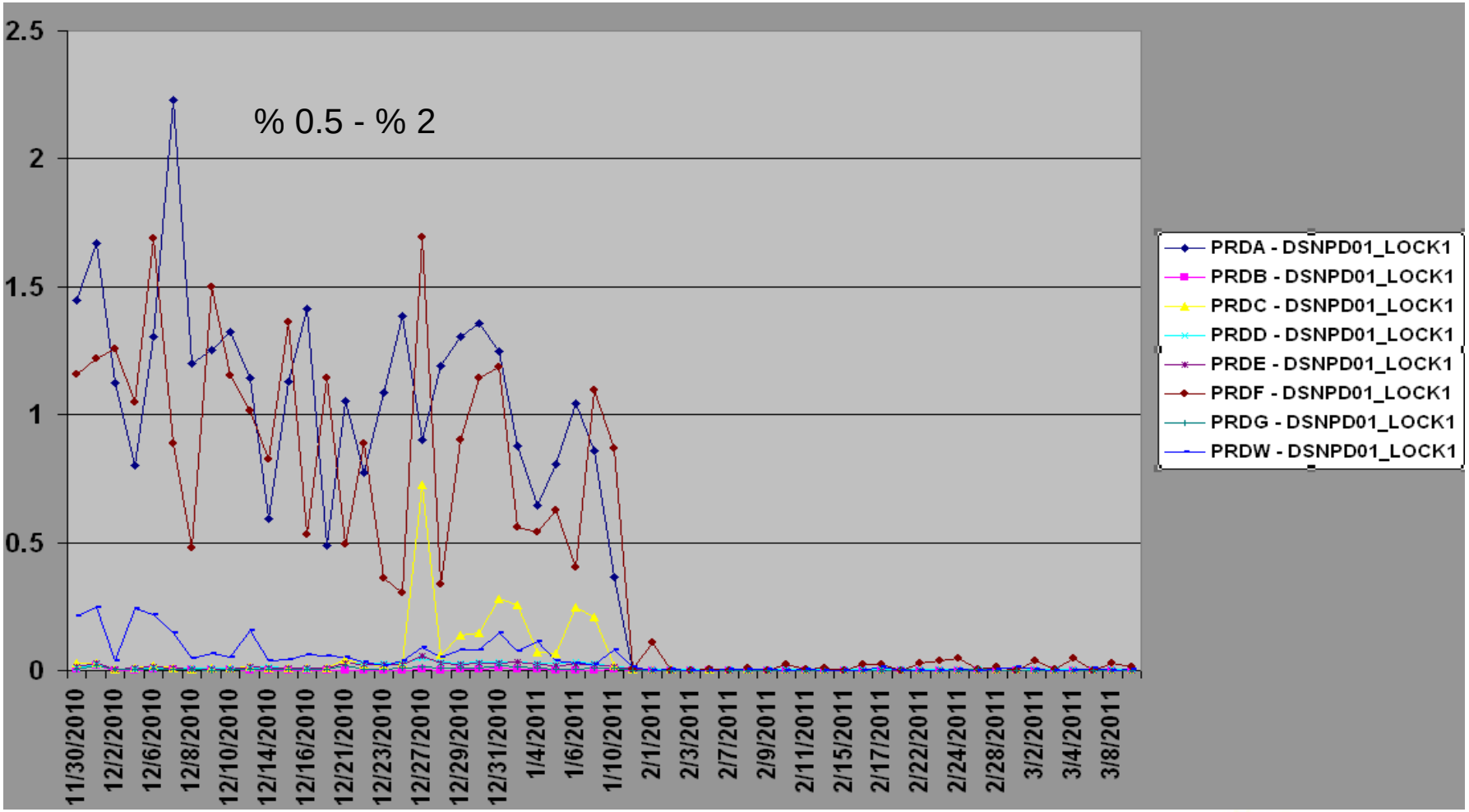
Average # Of Async Requests For DB2 Lock Structure 09:00-18:00 – Weekdays



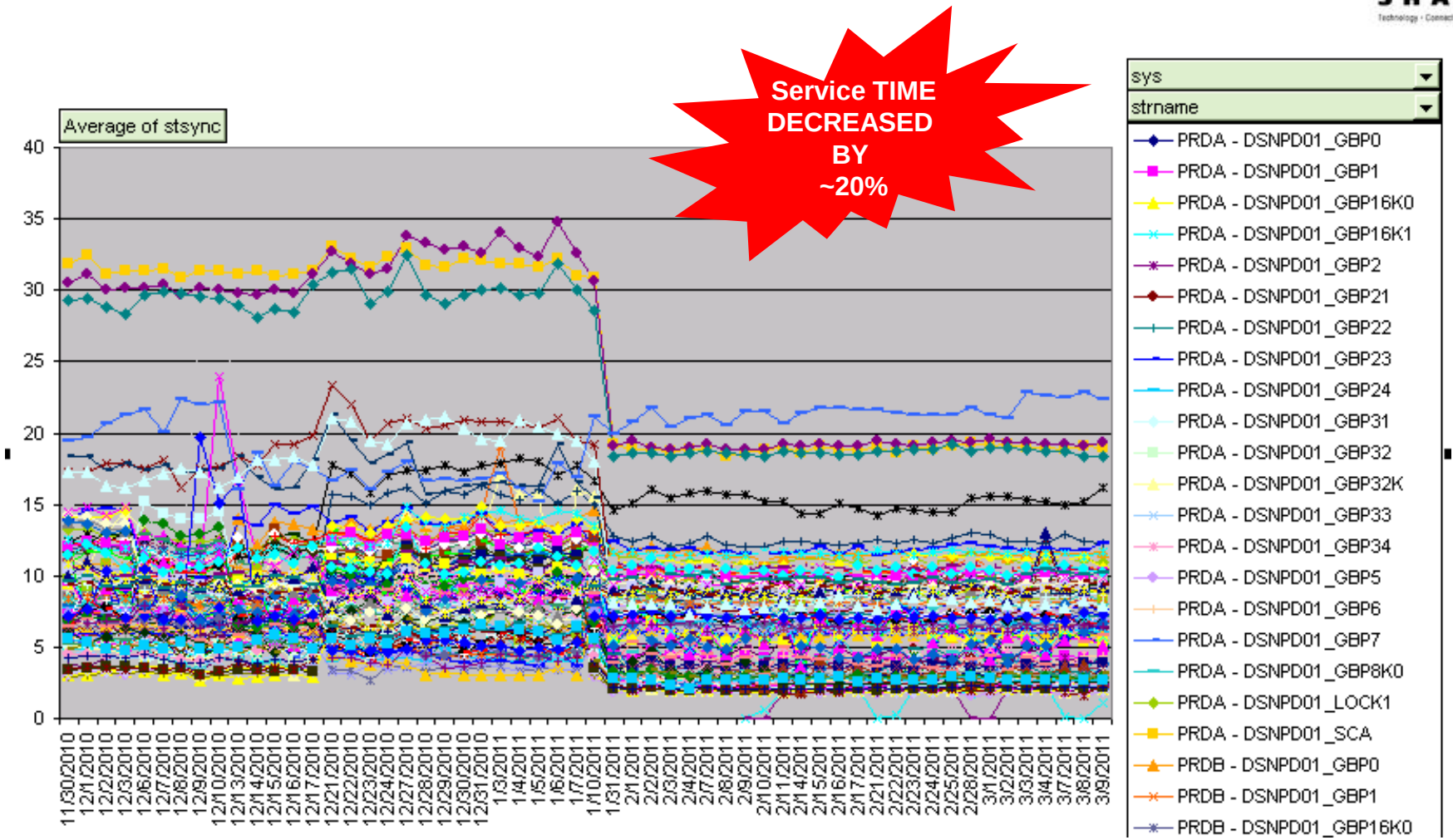
Probability of CF utilization decrease being main reason of this decrease is high

Configuration Change Effect on % Of Async Requests For DB2 Lock Structure

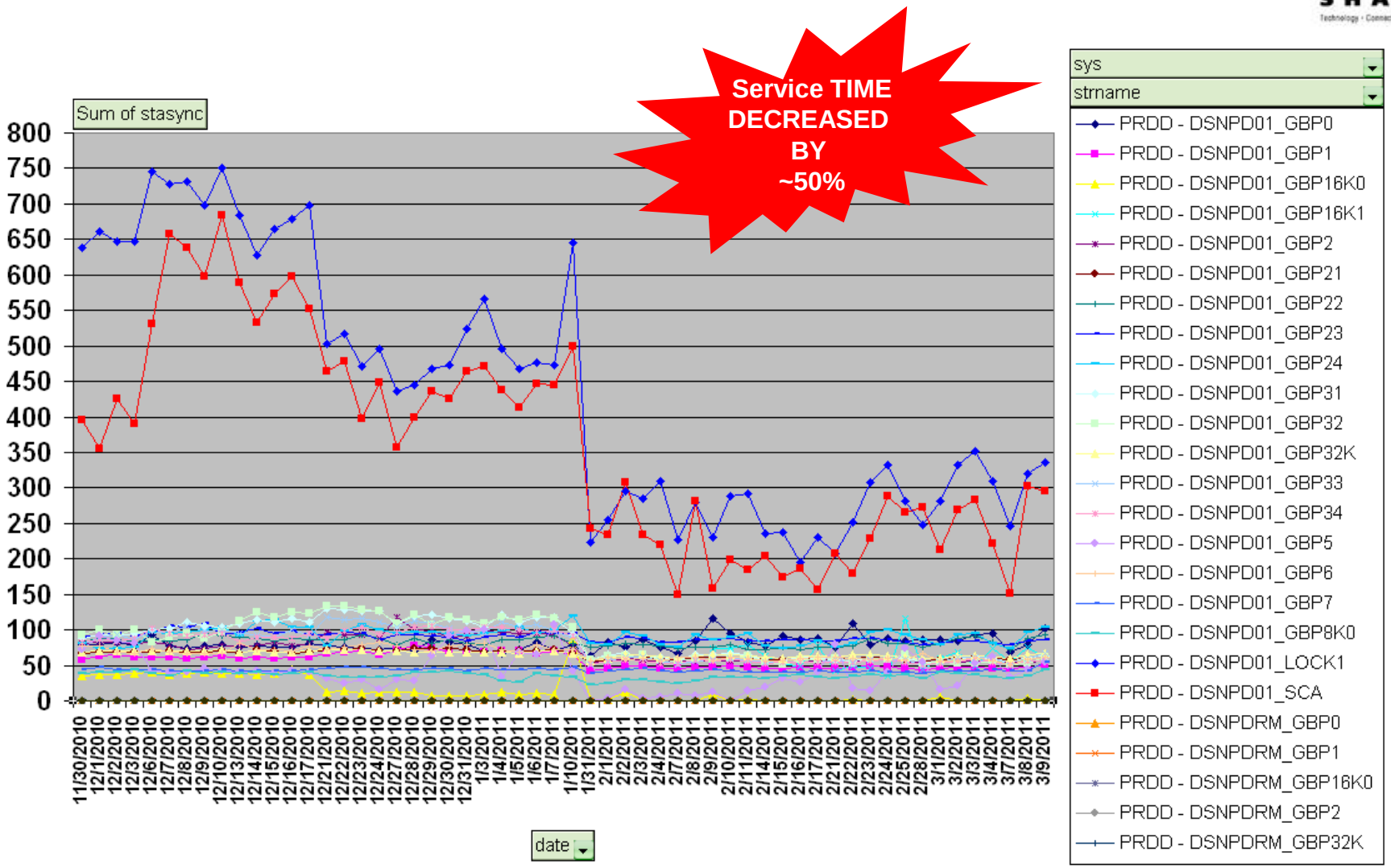
Async Request % For DB2 Lock Structures 09:00-18:00 – Weekdays



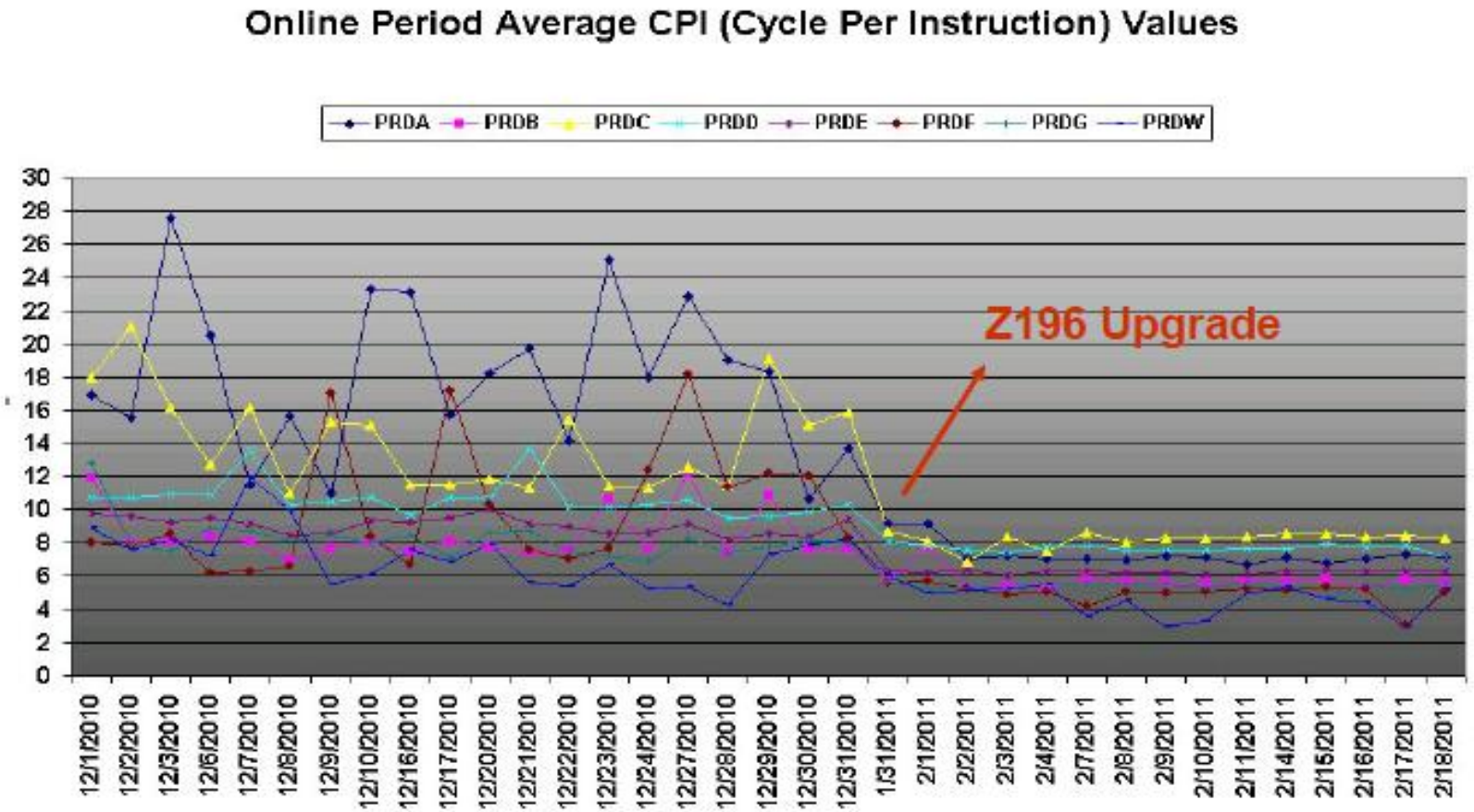
Sync Requests Daily Online Period Average Service Times



ASync Requests Daily Online Period Average Service Times



z196 Effect

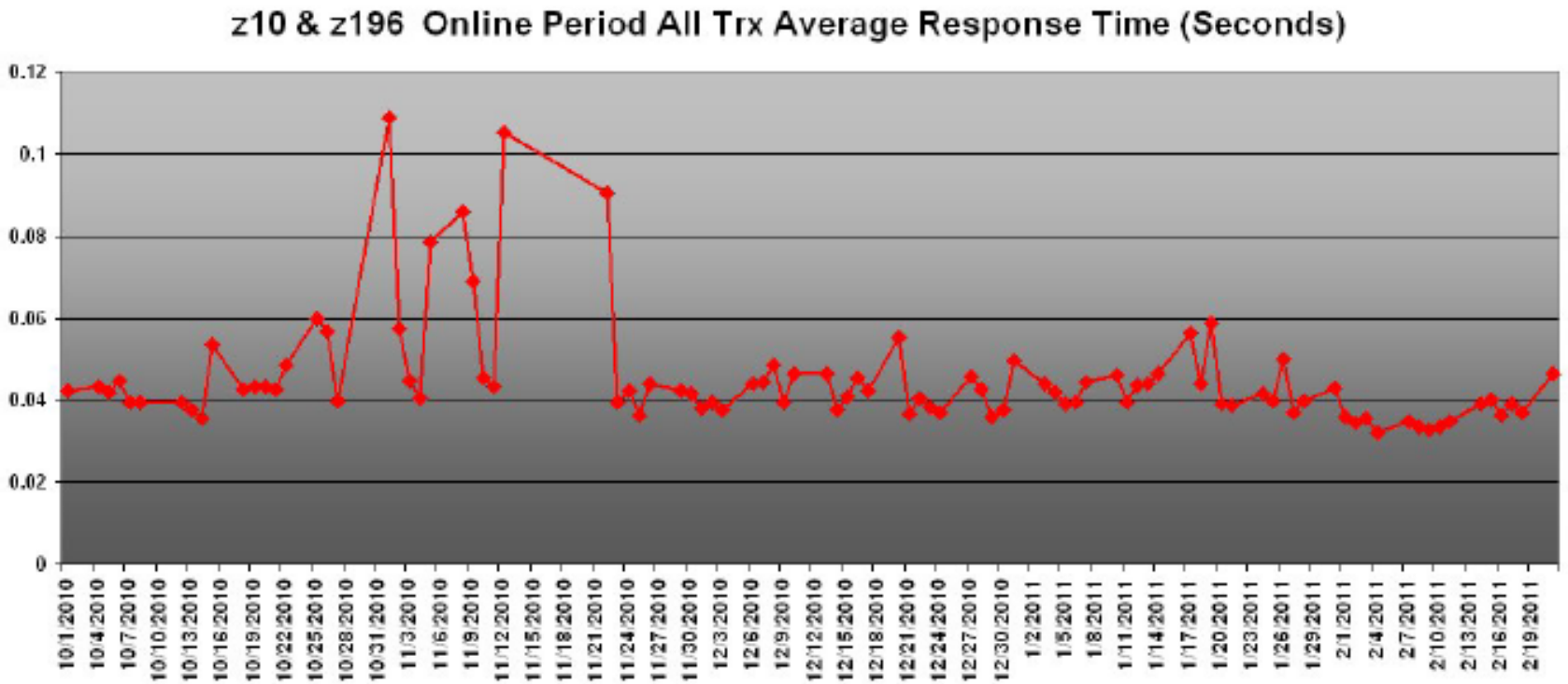

Online Period Average CPI (Cycle Per Instruction) Values
PRDA
PRDB
PRDC
PRDD
PRDE
PRDF
PRDG
PRDW
Z196 Upgrade
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
12/1/2010
12/2/2010
12/3/2010
12/6/2010
12/7/2010
12/8/2010
12/9/2010
12/10/2010
12/16/2010
12/17/2010
12/20/2010
12/21/2010
12/22/2010
12/23/2010
12/24/2010
12/27/2010
12/28/2010
12/29/2010
12/30/2010
12/31/2010
1/31/2011
2/1/2011
2/2/2011
2/3/2011
2/4/2011
2/7/2011
2/8/2011
2/9/2011
2/10/2011
2/11/2011
2/14/2011
2/15/2011
2/16/2011
2/17/2011
2/18/2011

z196 Effect

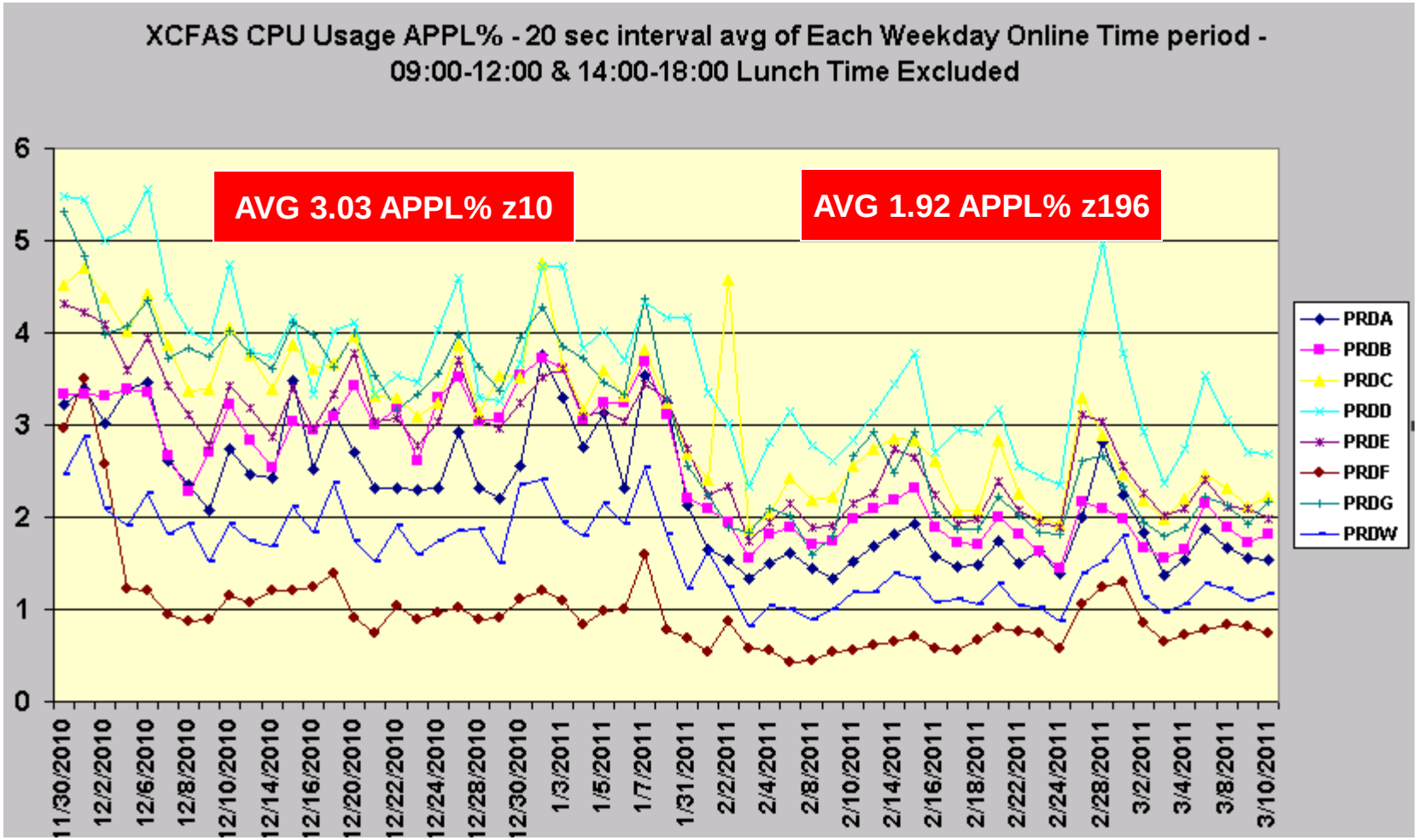
Z10 Online Time ALL Trx Avg ResponseTime 0.0042

14.5% Decrease

Z196 Online Time ALL Trx Avg ResponseTime 0.0036

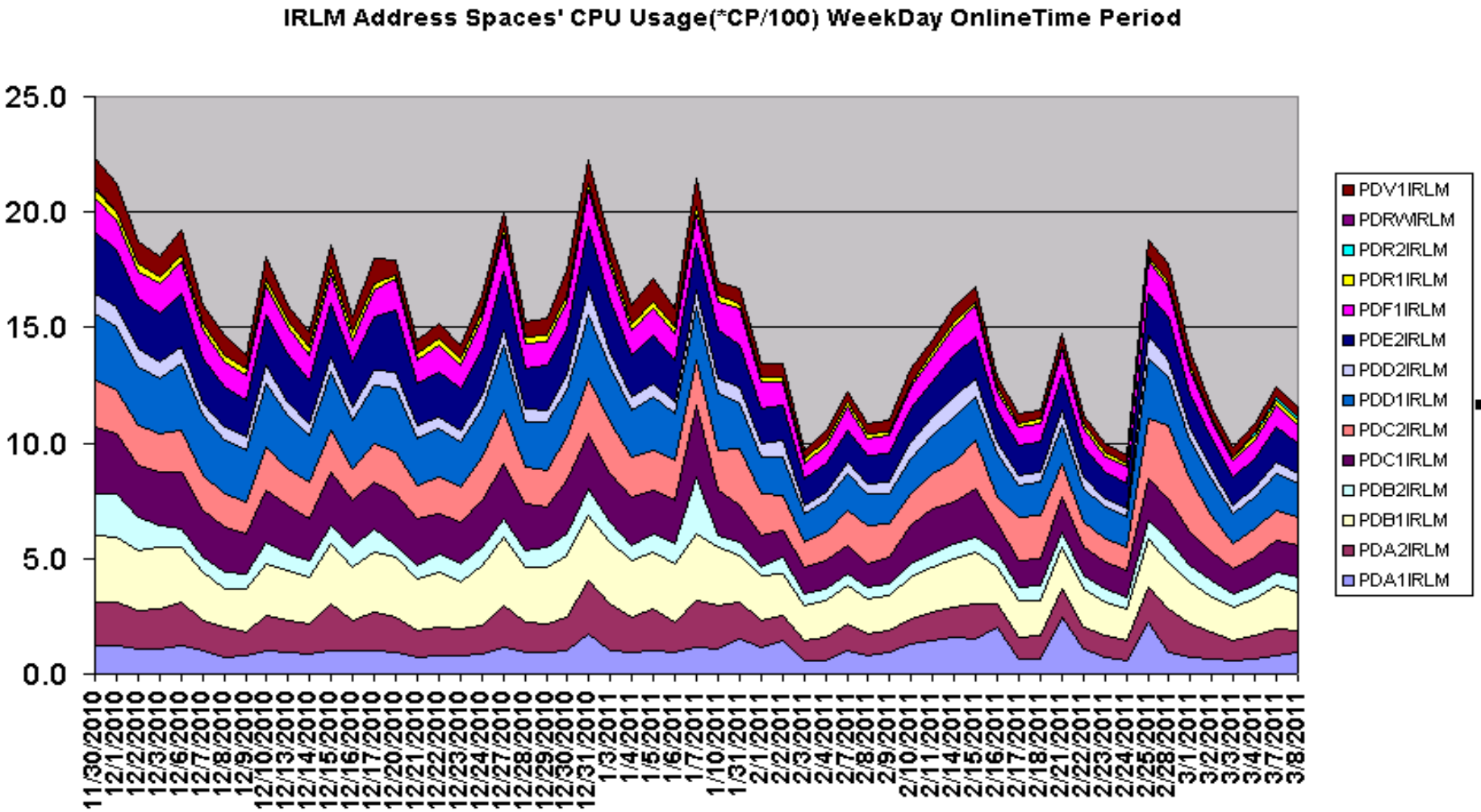


ASYNC Requests CPU USAGE CHARGED TO XCFAS Address Space



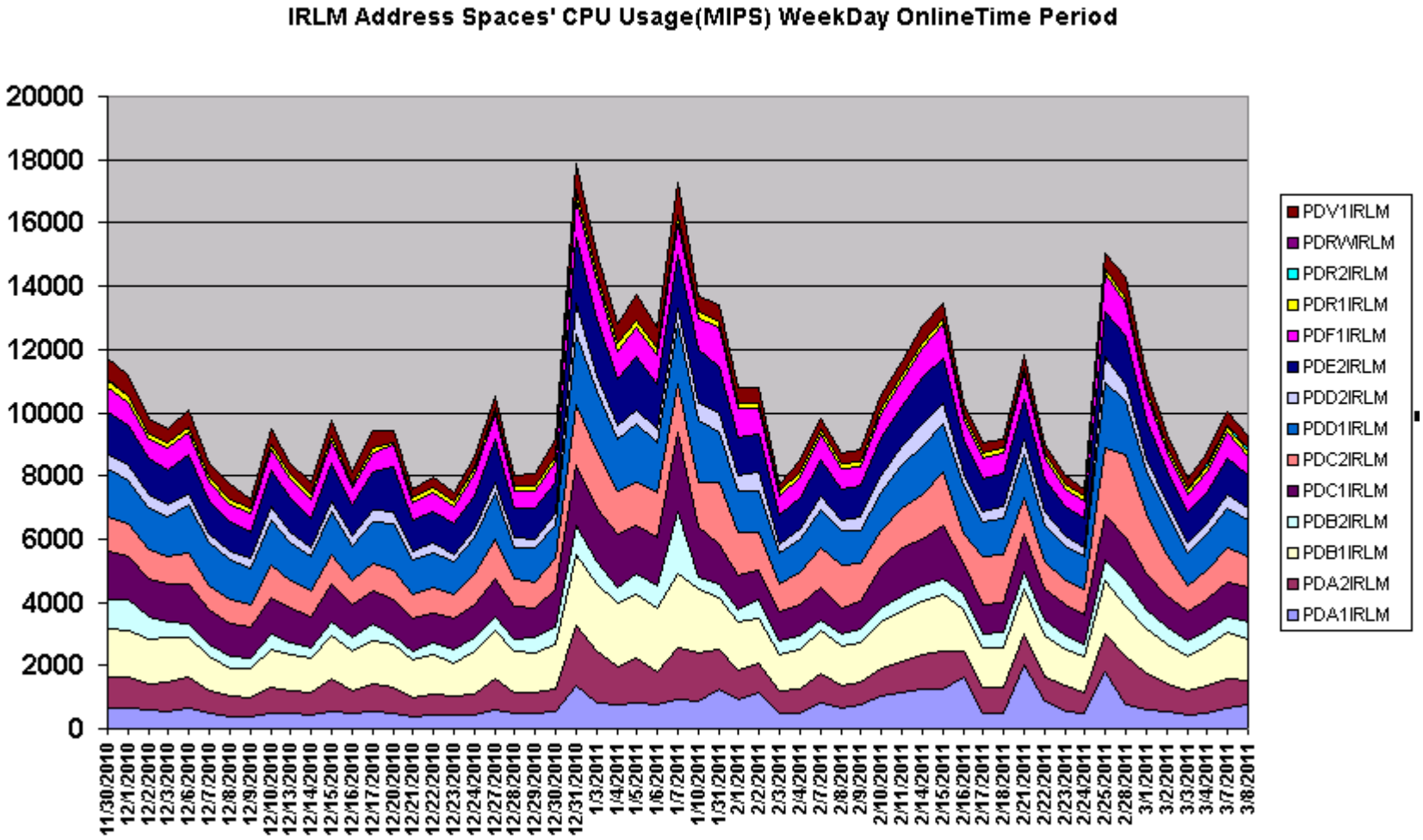
Without Normalization

SYNC Requests CPU USAGE CHARGED TO Requested AS: Sample IRLM



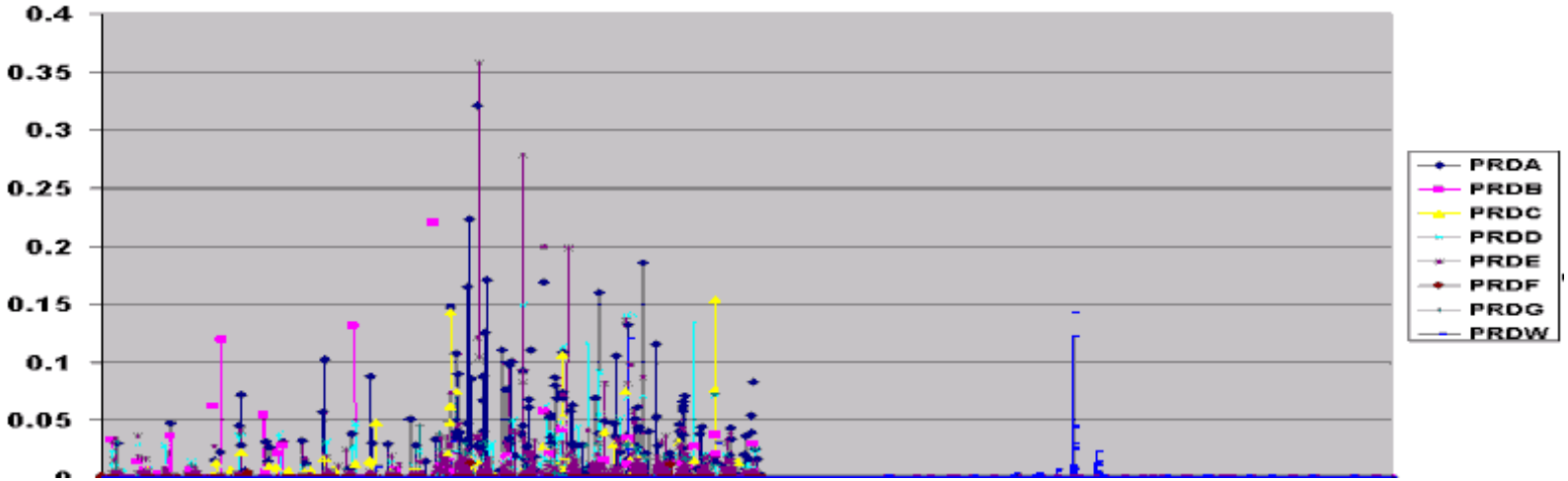
Without Normalization

SYNC Requests CPU USAGE CHARGED TO Requested AS: Sample IRLM

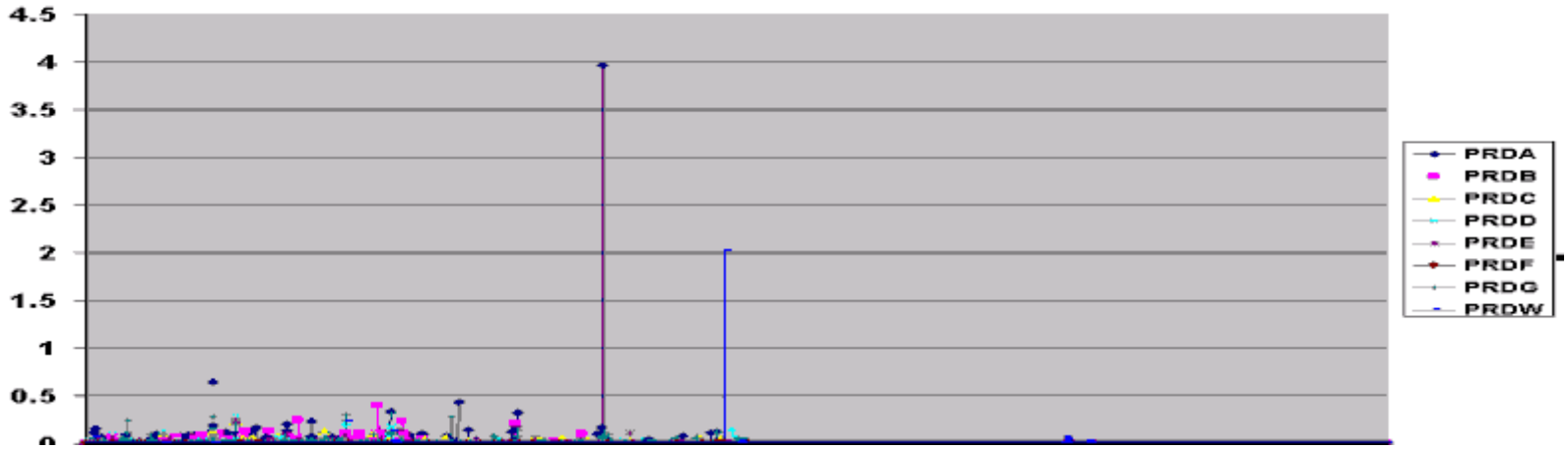


ROT Path Busy < %10 Of Total Req

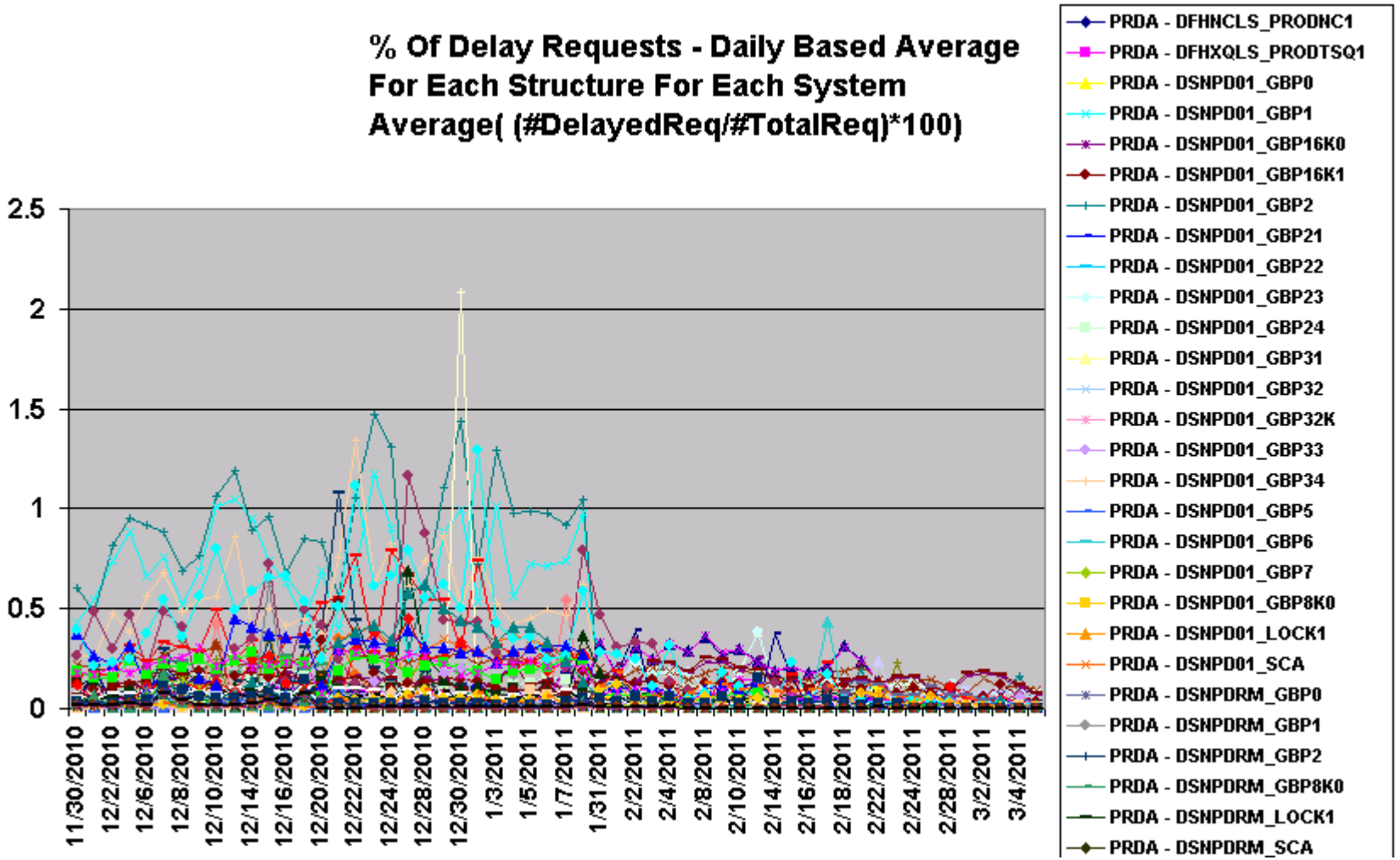
PCF1GAR1 Path Busy %



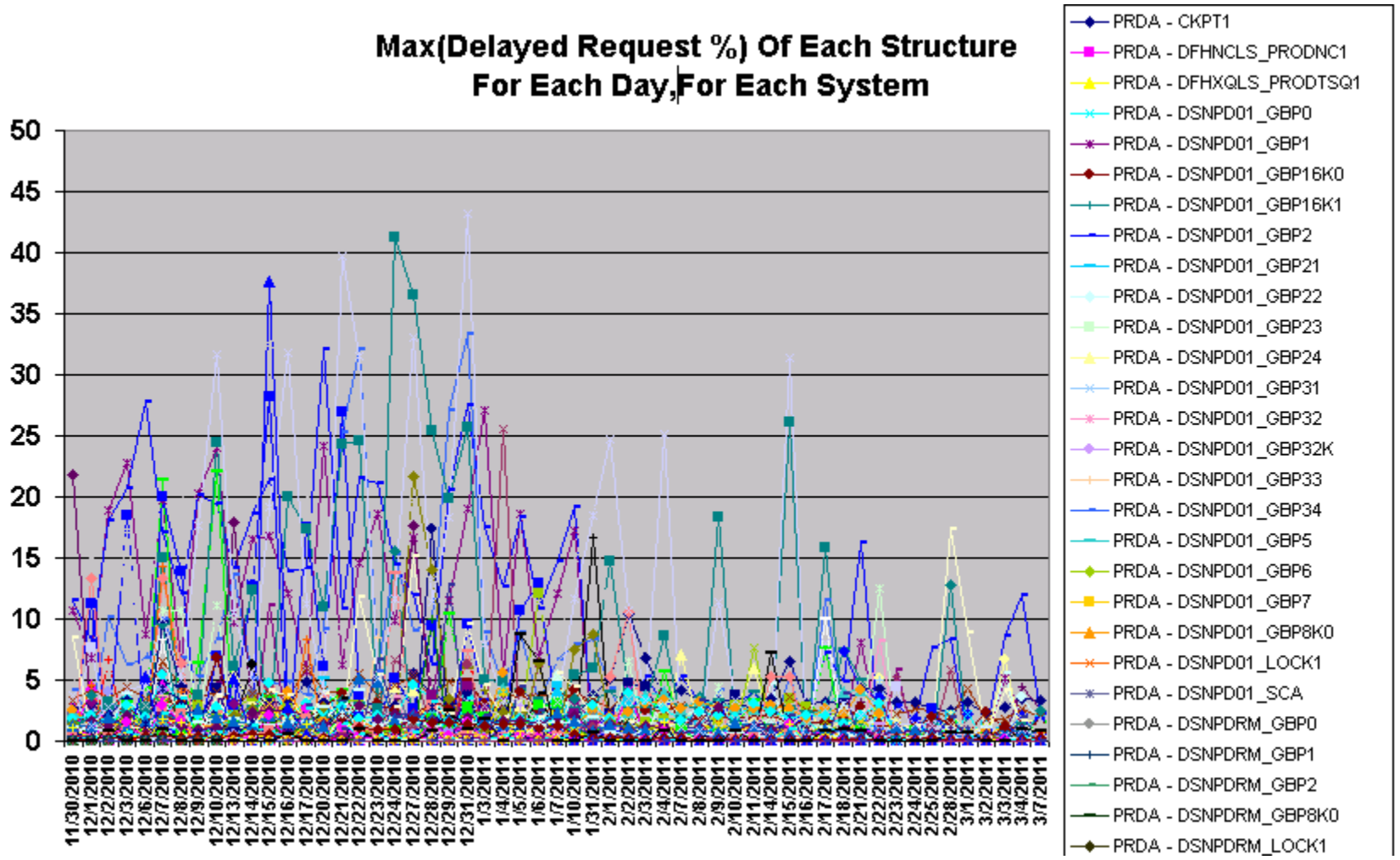
PCF2GAR2 Path Busy %



ROT Delayed Request % < %10 Of Total Req



ROT Delayed Request % < %10 Of Total Req



Host Cost (Data Sharing Cost)

Assumes 9 CF requests / MI

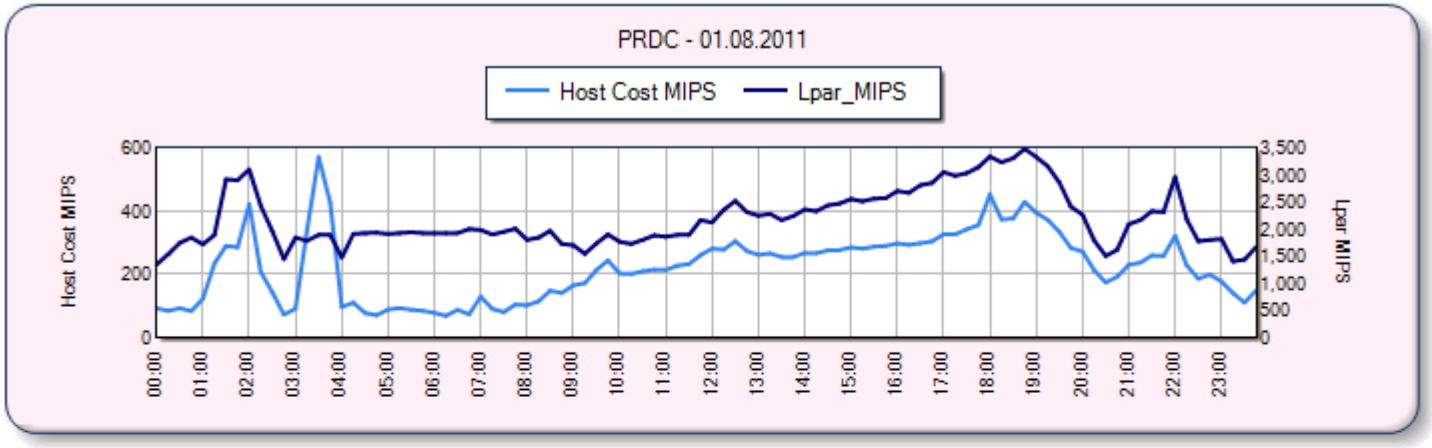
Thanks To Gary King

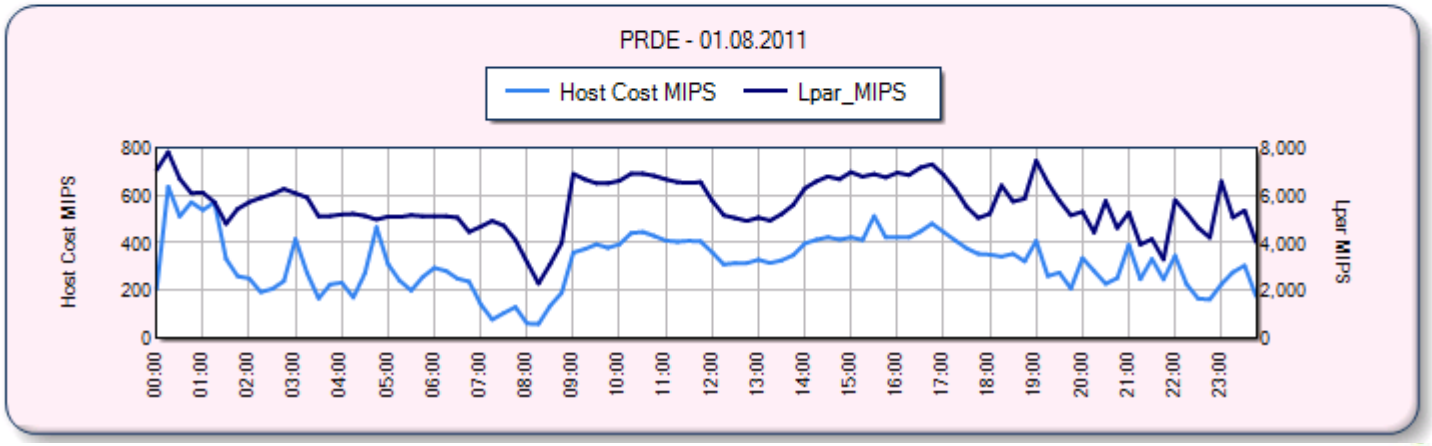
Host	z890	z990	z9 BC	z9 EC	z10 BC	z10 EC	z196
z890 ISC	13%	15%	16%	17%	19%	21%	NA
z890 ICB	9%	10%	10%	11%	12%	13%	NA
z990 ISC	13%	14%	14%	15%	17%	19%	NA
z990 ICB	9%	9%	9%	10%	12%	13%	NA
z9 BC ISC	12%	13%	14%	15%	17%	19%	23%
z9 BC PSIFB 12X	NA	NA	NA	NA	13%	14%	16%
z9 BC ICB	8%	9%	9%	10%	11%	12%	NA
z9 EC ISC	12%	13%	13%	14%	16%	18%	22%
z9 EC PSIFB 12X	NA	NA	NA	NA	13%	14%	16%
z9 EC ICB	8%	8%	8%	9%	10%	11%	NA
z10 BC ISC	12%	13%	13%	14%	16%	18%	22%
z10 BC PSIFB 12X	NA	NA	11%	12%	13%	14%	15%
z10 BC ICB	8%	8%	8%	9%	10%	11%	NA
z10 EC ISC	11%	12%	12%	13%	15%	17%	22%
z10 EC PSIFB 12X	NA	NA	10%	11%	12%	13%	15%
z10 EC ICB	7%	7%	7%	8%	9%	10%	NA
z196 ISC	NA	NA	11%	12%	14%	16%	21%
z196 PSIFB 12X	NA	NA	9%	10%	11%	12%	14%

The table does not take into consideration any extended distance effects or system managed duplexing

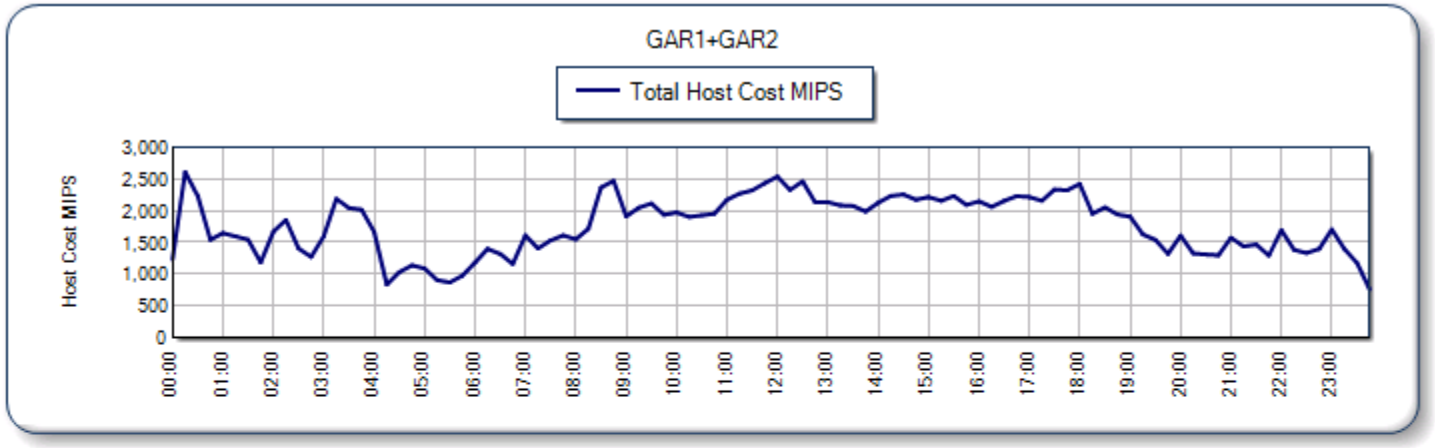
Calculating Host Cost (Data Sharing Cost)

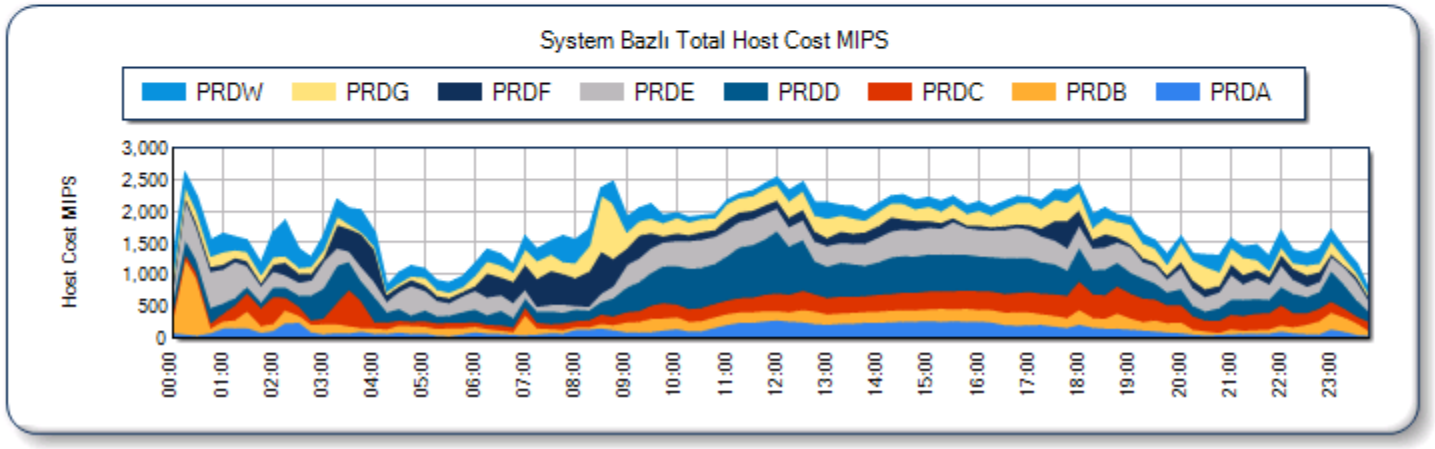
One can calculate the coupling intensity by simply summing the total req/sec of the CFs and dividing by the used MIPS of the attached systems (MIPS rating times CPU busy). Then, the values in the table would be linearly scaled. For example, if the workload was processing 4.5 CF operations per million instructions (or 4.5 CF ops/second/MIPS), then all the values in the table would be cut in half.



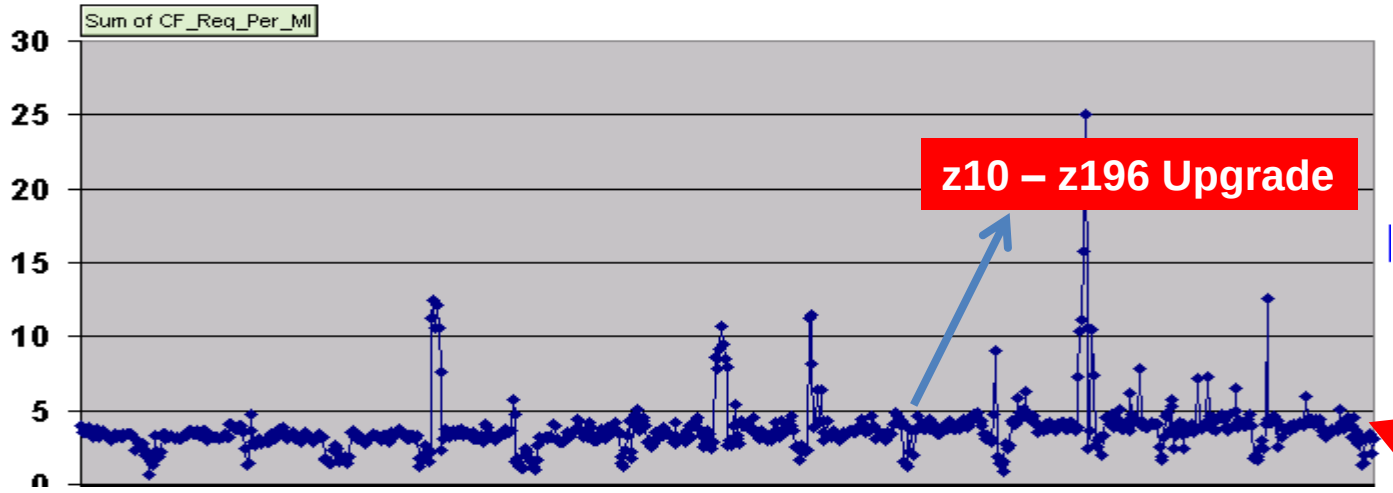


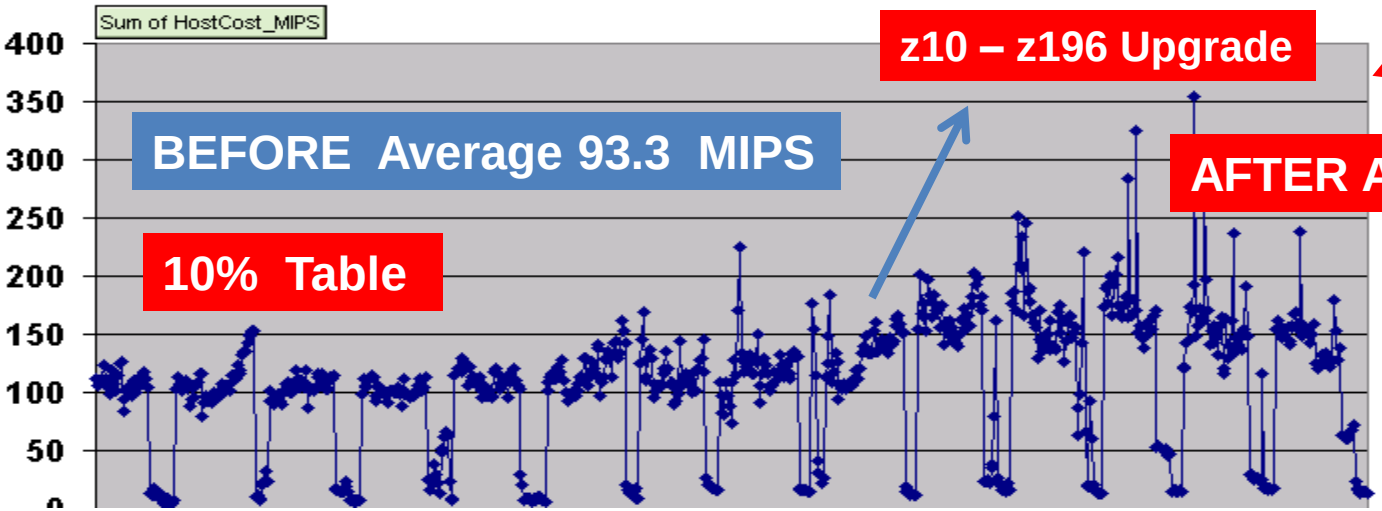
Calculating Host Cost (Data Sharing Cost)





Host Cost Before & After Upgrade – One LPAR





ACTUAL 43 % Increase In Host Cost

Host Cost With New IFB3 Protocol

Thanks To Riaz Ahmad For This Information

Host / CF	Z890	Z990	Z9 BC	Z9 EC	Z10 BC	Z10 EC	Z114	Z196
Z890 ISC	13	15	16	17	19	21	-	-
Z890 ICB	9	10	10	11	12	13	-	-
Z990 ISC	13	14	14	15	17	19	-	-
Z990 ICB	9	9	9	10	12	13	-	-
Z9 BC ISC	12	13	14	15	17	19	18	23
Z9 BC 12x IFB	-	-	-	-	13	14	13	16
Z9 BC ICB4	8	9	9	10	11	12	-	-
Z9 EC ISC	12	13	13	14	15	18	17	22
Z9 EC 12x IFB	-	-	-	-	13	14	13	16
Z9 EC ICB	8	8	9	9	10	11	-	-
Z10 BC ISC	12	13	13	14	16	18	17	22
Z10 BC 12x IFB			11	12	13	14	13	15
Z10 BC ICB	8	8	9	9	10	11	-	-
Z10 EC ISC	11	12	12	13	15	17	17	22
Z10 EC 12x IFB			10	11	12	13	12	15
Z10 EC ICB	7	7	7	8	9	10	-	-
Z114 ISC3			14	14	16	18	17	21
Z114 12x IFB			10	10	12	13	12	15
Z114 12x IFB3							10	12
Z196 ISC			11	12	14	16		21
Z196 12x IFB			9	10	11	12	11	14
Z196 12x IFB3							9	11

CF SUBCHANNEL UTILIZATION

Calculate Yourself Using SMF Records

Utilization % = (((Sync #Req * Sync service time) + (Async #Req * Async service time)) / Interval time * #Subchannels in use) * 100

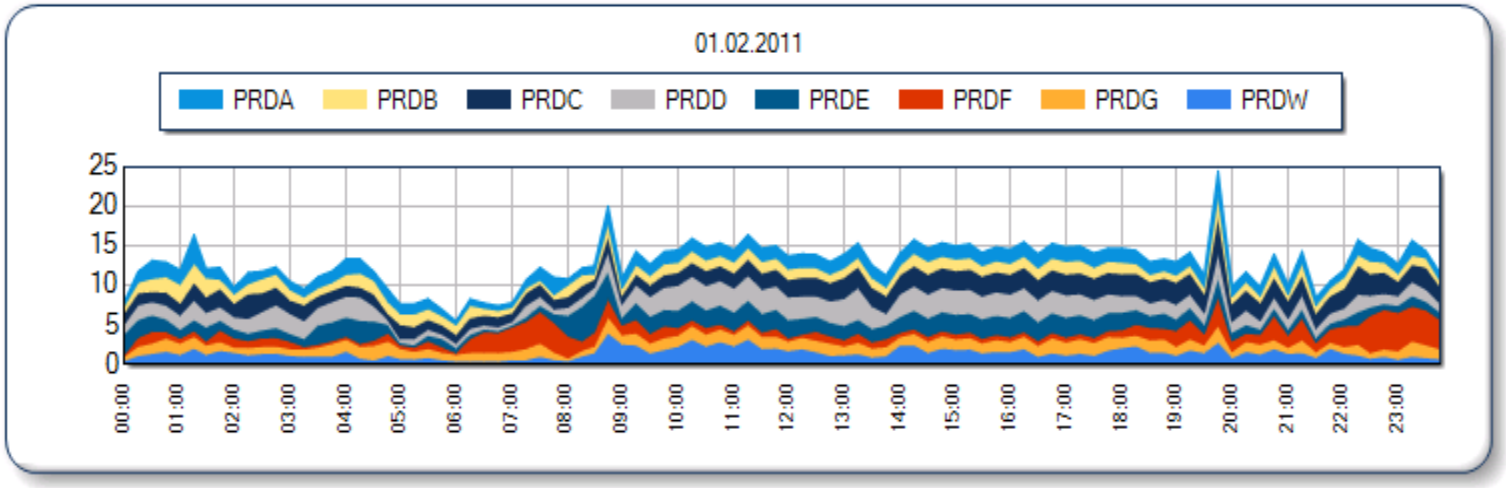
OR

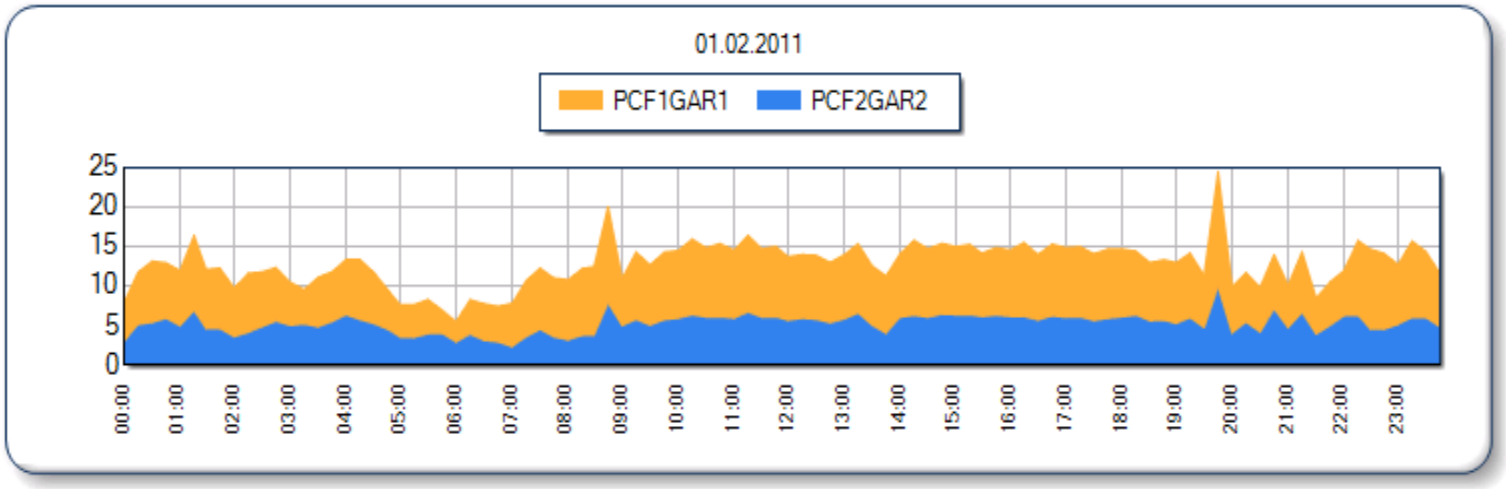
Using RMF Overview Report

OVERVIEW(REPORT)
OVW(CF1P(SUBCHBP(PCF1GAR1)))
OVW(CF2P(SUBCHBP(PCF2GAR2)))

Subchannel busy percentage	SUBCHBP	cfname	R744SSRC R744SSTM R744SARC R744SATM R744FSCU	((R744SSRC * R744SSTM) + (R744SARC * R744SATM)) *100 / Interval * R744FSCU	S

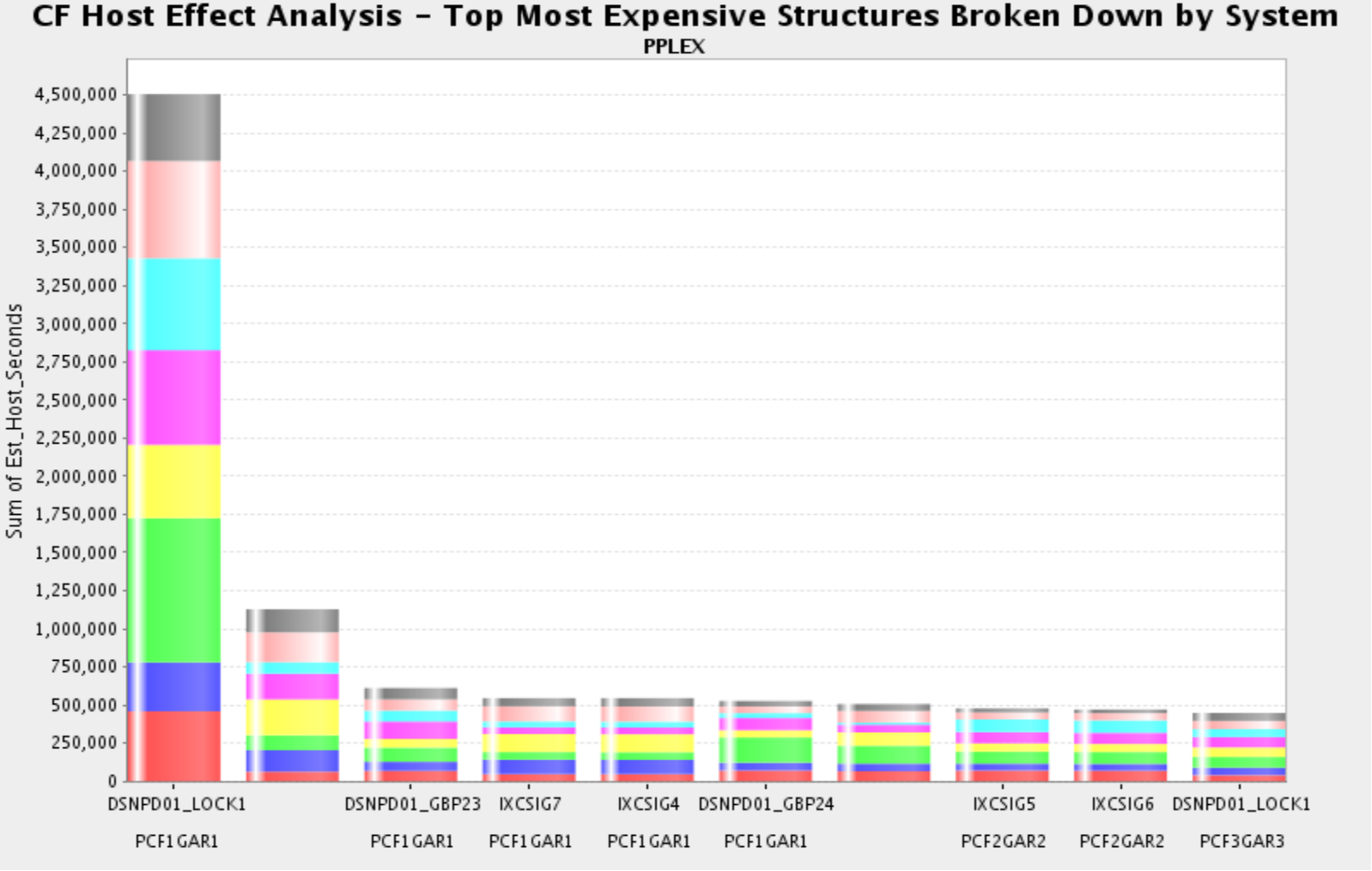
CF SUBCHANNEL UTILIZATION – ONE DAY





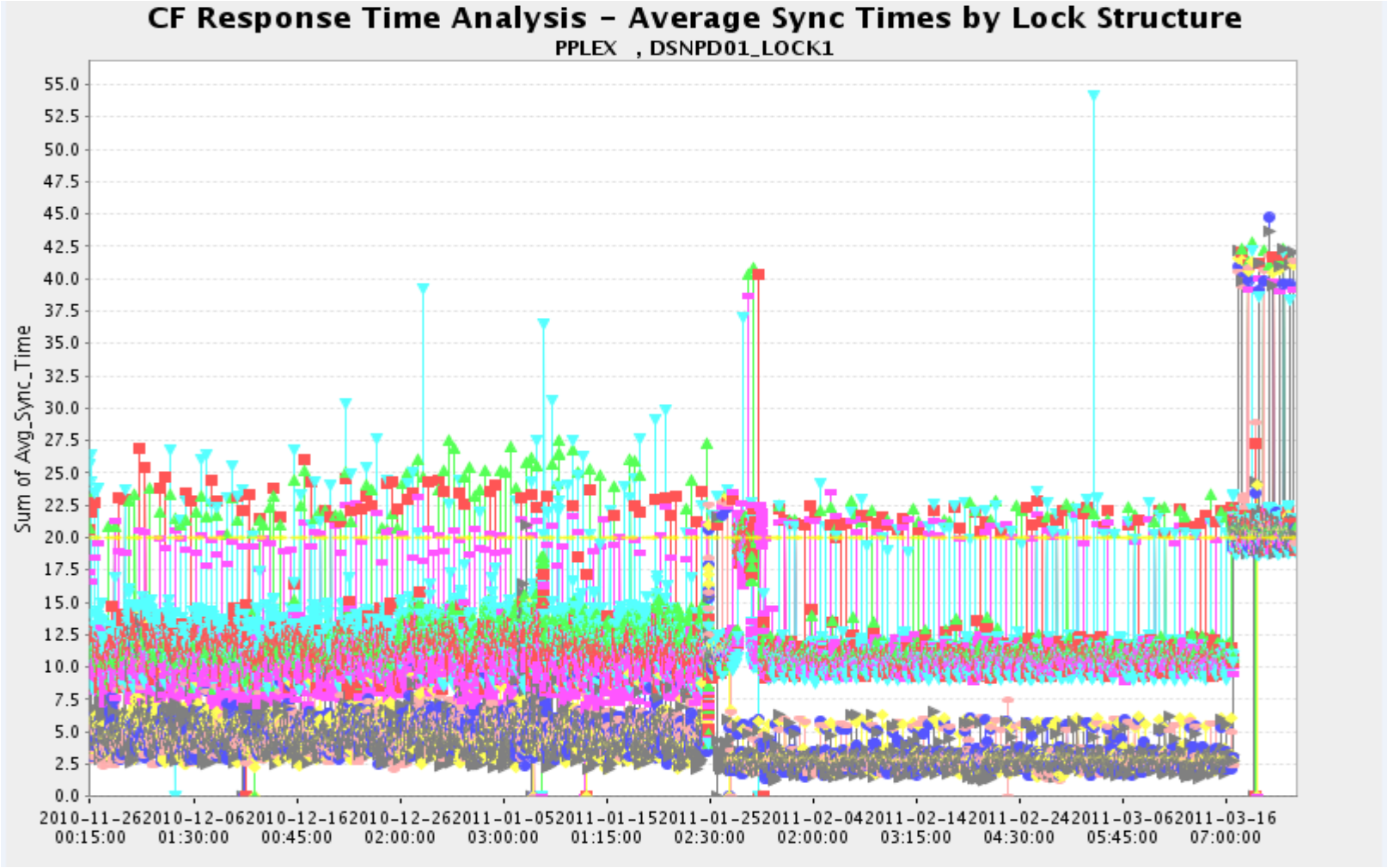
Pivotor Product of Peter Enrico

http://www.epstrategies.com/sitex/index.php



Pivotor Product of Peter Enrico

http://www.epstrategies.com/sitex/index.php



RMF Mon III,I Panels & Key Fields

Understanding Data Types In Reports

- RMF Monitor I Post Processor Report Fields
- RMF Monitor I Overview/Exception Report Fields
- RMF Monitor III Report Fields
- SMF Record Fields (RMF Related Records 70-79)

If explanation in books is not clear,

- Cross Check Related Fields in Other Types Of Data
- Google It – For APARs, Redbooks,WSC Documents
- Ask IBM – Open PMR For Information Request

Sample : Path Busy Condition

SMF Record R744FPBC = 203

Output Of RMF Post Processor Report With Statement "SYSRPTS(CF)"

SUBCHANNEL ACTIVITY

SYSTEM NAME	Ö REQ TOTAL AVG/SEC	-- CF LINKS -- TYPE	GEN	USE	PTH BUSY	REQUESTS	Ö REQ	-SERVICE TIME(MIC)- AVG	STD_DEV	
PRDA	2116K	ICP	4	4	203	SYNC	1402K	6.7	28.2	LIST/
	2351.3	SUBCH	56	28		ASYNC	691804	80.8	631.1	LOCK
						CHANGED	212	INCLUDED IN ASYNC		TOTAL
						UNSUCC	0	0.0	0.0	

Output Of RMF Post Processor Overview/Exception Report With Statements

OVERVIEW(REPORT)
OVW(PTHBUS1(PBSY(PCF1GAR1)))
OVW(PTHBUS2(PBSY(PCF2GAR2)))

R744FBC/Interval = 203/(900 secs) =0.22

R M F O V E R V I E W

z/OS V1R12 SYSTEM ID PRDA START
RPT VERSION V1R12 RMF END

NUMBER OF INTERVALS 1 TOTAL LENGTH OF INTERVALS 00.15.00

DATE MM/DD	TIME HH.MM.SS	INT HH.MM.SS	PTHBUS1	PTHBUS2
07/25	09.45.00	00.15.00	0.0	0.2

RMF Mon I Post Processor Reports – CF Reports

Postprocessor Statement – SYSRPTS(CF) - See sample JCL in backup slides

- Coupling Facility Usage Summary
- Coupling Facility Structure Activity
- Subchannel Activity

RMF Subchannel Activity Section

Delay Reason Is Related No Subchannel

Path Busy is NOT included
But IBM recognises path busy before
And reissue request using same subchannel

Service Times Do NOT include Delay Times

SUBCHANNEL ACTIVITY

SYSTEM NAME	Ö REQ TOTAL AVG/SEC	-- CF LINKS -- TYPE	GEN	USE	PTH BUSY	REQUESTS	Ö REQ	-SERVICE TIME(MIC)- AVG	STD_DEV	DELAYED REQUESTS	Ö REQ	% OF REQ	AVG TIME(MIC) /DEL	STD_DEV	/ALL
PRDA	2391K	CIB	3	3	0	SYNC	1579K	13.0	4.7	LIST/CACHE	358	0.0	42.5	33.6	0.0
	2656.3	SUBCH	42	21		ASYNC	788047	111.8	161.0	LOCK	0	0.0	0.0	0.0	0.0
						CHANGED	0	INCLUDED IN ASYNC		TOTAL	358	0.0			
						UNSUCC	0	0.0	0.0						
PRDB	5869K	ICP	4	4	1935	SYNC	4234K	3.9	34.9	LIST/CACHE	415	0.0	789.0	601.7	0.1
	6521.0	SUBCH	56	28		ASYNC	1589K	41.7	395.3	LOCK	13	0.0	207.0	159.9	0.0
						CHANGED	413	INCLUDED IN ASYNC		TOTAL	428	0.0			
						UNSUCC	0	0.0	0.0						
PRDC	6364K	CIB	3	3	0	SYNC	4671K	12.9	4.0	LIST/CACHE	3052	0.1	707.5	573.0	0.8
	7071.1	SUBCH	42	21		ASYNC	1645K	72.8	88.7	LOCK	60	0.0	115.5	125.4	0.0
						CHANGED	2492	INCLUDED IN ASYNC		TOTAL	3112	0.0			
						UNSUCC	0	0.0	0.0						
PRDD	11892K	ICP	4	4	2718	SYNC	9162K	4.1	31.9	LIST/CACHE	582	0.0	961.6	1544	0.1
	13213	SUBCH	56	28		ASYNC	2757K	36.0	475.6	LOCK	86	0.0	393.3	1096	0.0
						CHANGED	627	INCLUDED IN ASYNC		TOTAL	668	0.0			

RMF CF Usage Summary Section

COUPLING FACILITY USAGE SUMMARY

STRUCTURE SUMMARY

TYPE	STRUCTURE NAME	STATUS CHG	ALLOC SIZE	% OF CF STOR	Ö REQ	% OF ALL REQ	% OF CF UTIL	AVG REQ/ SEC	LST/DIR ENTRIES TOT/CUR	DATA ELEMENTS TOT/CUR	LOCK ENTRIES TOT/CUR	DIR REC/ DIR REC XI'S
LIST	DSNPDRM_SCA	ACTIVE	11M	0.0	2403	0.0	0.0	2.67	11K	22K	N/A	N/A
									158	458	N/A	N/A
	DSNPD01_SCA	ACTIVE	70M	0.3	36277	0.1	0.1	40.31	80K	159K	N/A	N/A
									1013	2617	N/A	N/A
	HSA_LOG	ACTIVE	14M	0.1	3	0.0	0.0	0.00	2977	9009	N/A	N/A
									153	366	N/A	N/A
	HZS_HEALTHCHKLOG	ACTIVE	15M	0.1	379	0.0	0.0	0.42	3575	32K	N/A	N/A
									2322	22K	N/A	N/A
	IBMBDG	ACTIVE	16M	0.1	971	0.0	0.0	1.08	731	25K	N/A	N/A
									56	2938	N/A	N/A
	IXCSIG1	ACTIVE	15M	0.1	164244	0.3	0.8	182.49	1537	1522	N/A	N/A

RMF CF Usage Summary Section

At the end of this section ,Summary part exists

```
                                        -------  ------  -------   -----  ------  -------
          STRUCTURE TOTALS                 13G    47.5    47169K    100     100    52410

 STORAGE SUMMARY
-------------------------------------------------------------------------------------------

                                          ALLOC      % OF CF      ------- DUMP SPACE -------
                                          SIZE       STORAGE     % IN USE   MAX % REQUESTED

 TOTAL CF STORAGE USED BY STRUCTURES      12929M       47.5
 TOTAL CF DUMP STORAGE                       49M        0.2          0.0                 0.0
 TOTAL CF STORAGE AVAILABLE               14227M       52.3
                                          -------
 TOTAL CF STORAGE SIZE                    27205M

                                          ALLOC      % ALLOCATED
                                          SIZE
TOTAL CONTROL STORAGE DEFINED             27205M       47.7
TOTAL DATA STORAGE DEFINED                    0K        0.0

PROCESSOR SUMMARY
-------------------------------------------------------------------------------------------

COUPLING FACILITY         2817        MODEL M32      CFLEVEL  17          DYNDISP OFF

AVERAGE CF UTILIZATION  (% BUSY)          14.0       LOGICAL PROCESSORS:  DEFINED   2   EFFECTIVE   2.0
                                                                          SHARED    0   AVG WEIGHT  0.0
```

RMF Structure Activity Section

STRUCTURE NAME = DSNPD01_LOCK1 TYPE = LOCK STATUS = ACTIVE

SYSTEM NAME	Ö REQ TOTAL AVG/SEC		REQUESTS Ö REQ	% OF ALL	-SERV TIME(MIC)- AVG	STD_DEV	DELAYED REQUESTS REASON	Ö REQ	% OF REQ	AVG TIME(MIC) /DEL	STD_DEV	/ALL	EXTERNAL REQUEST CONTENTIONS	
PRDA	992K	SYNC	992K	3.8	11.8	3.5	NO SCH	0	0.0	0.0	0.0	0.0	REQ TOTAL	1323K
	1102	ASYNC	0	0.0	0.0	0.0	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	11K
		CHNGD	0	0.0	INCLUDED IN ASYNC		PR CMP	0	0.0	0.0	0.0	0.0	-CONT	11K
													-FALSE CONT	5615
PRDB	3521K	SYNC	3521K	13.4	3.4	35.0	NO SCH	13	0.0	207.0	159.9	0.0	REQ TOTAL	4412K
	3912	ASYNC	13	0.0	1037.2	506.1	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	25K
		CHNGD	13	0.0	INCLUDED IN ASYNC		PR CMP	0	0.0	0.0	0.0	0.0	-CONT	25K
													-FALSE CONT	9039
PRDC	3772K	SYNC	3772K	14.4	12.2	3.0	NO SCH	60	0.0	115.5	125.4	0.0	REQ TOTAL	4052K
	4191	ASYNC	53	0.0	541.4	372.3	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	11K
		CHNGD	53	0.0	INCLUDED IN ASYNC		PR CMP	0	0.0	0.0	0.0	0.0	-CONT	11K
													-FALSE CONT	6088

RMF Structure Activity Section

STRUCTURE NAME = DSNPD01_GBP23 TYPE = CACHE STATUS = ACTIVE PRIMARY

SYSTEM NAME	Ö REQ TOTAL AVG/SEC	REQUESTS	Ö REQ	% OF ALL	-SERV TIME(MIC)- AVG	STD_DEV	DELAYED REQUESTS REASON	Ö REQ	% OF REQ	---- AVG TIME(MIC) ----- /DEL	STD_DEV	/ALL
PRDA	56964	SYNC	55K	1.6	16.2	7.5	NO SCH	0	0.0	0.0	0.0	0.0
	63.29	ASYNC	2149	0.1	214.4	215.3	PR WT	0	0.0	0.0	0.0	0.0
		CHNGD	0	0.0	INCLUDED IN ASYNC		PR CMP	0	0.0	0.0	0.0	0.0
							DUMP	0	0.0	0.0	0.0	0.0
PRDB	319K	SYNC	309K	9.0	5.4	33.0	NO SCH	138	0.0	407.3	390.7	0.2
	354.4	ASYNC	10K	0.3	109.9	361.5	PR WT	0	0.0	0.0	0.0	0.0
		CHNGD	138	0.0	INCLUDED IN ASYNC		PR CMP	0	0.0	0.0	0.0	0.0
							DUMP	0	0.0	0.0	0.0	0.0
PRDC	128K	SYNC	111K	3.2	14.5	7.9	NO SCH	7	0.0	561.4	696.3	0.0
	141.7	ASYNC	17K	0.5	119.1	164.2	PR WT	0	0.0	0.0	0.0	0.0
		CHNGD	7	0.0	INCLUDED IN ASYNC		PR CMP	0	0.0	0.0	0.0	0.0
							DUMP	0	0.0	0.0	0.0	0.0

Important SMF Fields Analized in SMF 74(4)

ASYNC*REQUESTS	R744SARC
AVERAGE*SERVICE TIME*PER ASYNC	R744SATM
SYNC*REQUESTS	R744SSRC
AVERAGE*SERVICE TIME*PER SYNC	R744SSTM
REQUESTS*CHANGED FROM*SYNC TO ASYNC	R744SSTA
REQUESTS*QUEUED	R744SQRC
REQUESTS*WAITING ON*HI PRTY QUEUE	R744SHTO
REQUESTS*WAITING ON*LO PRTY QUEUE	R744SLTO
TIMES CF*REQUEST FAILED*DUE TO PATH BUSY	R744FPBC
TOTAL REQUESTS*FROM THIS*SYSTEM	R744FTOR

Special THANKS TO .....

GEORGETTE KURDT – IBM
Very Special Thanks to Georgette for her many help

GARY KING - IBM

CHERYL WATSON – Watson & Walker

PETER ENRICO - EPS Strategies

MARIANNE HAMMER – IBM

BARBARA WEILER – IBM

REFERENCES

- z/OS 1.12 Setting Up Sysplex
- IBM z/OS Parallel Syplex Operational Senarious
- System z Parallel Sysplex Best Practices
- Coupling Facility Performance : A Real World Perspective
- z196 PR/SM Guide , z196 Technical Guide,z196 Connectivity Handbook
- Previous SHARE prezentations – Parallel Sysplex Update and many more...
- IBM WSC Papers & Flashes
- System 390 Parallel Sysplex Performance
- www.research.ibm.com
- https://www-304.ibm.com/servers/resourcelink/svc03100.nsf?Opendatabase
- IBM XES Related APARs – Really Good Information in APARs.
- www.freepatentsonline.com - For Understanding Alternatives At Least

MORE INFORMATION & BACKUP SLIDES

RMF REPORT SAMPLES

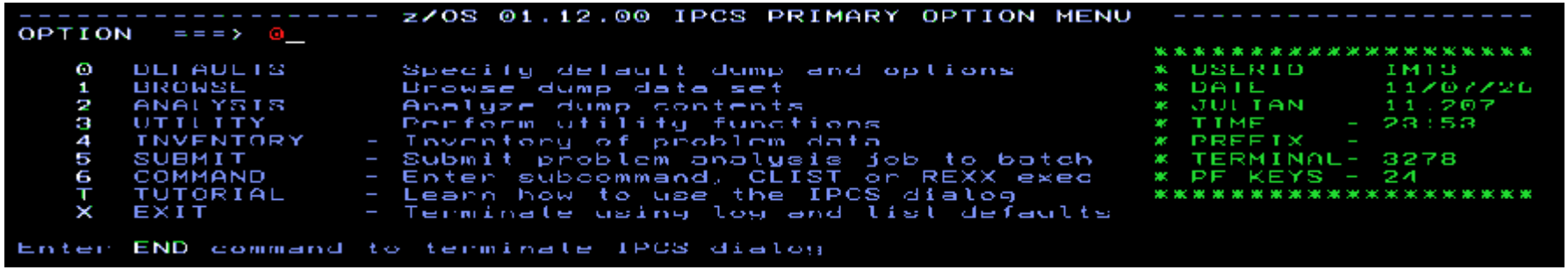
RMF Report Sample JCL * CF report

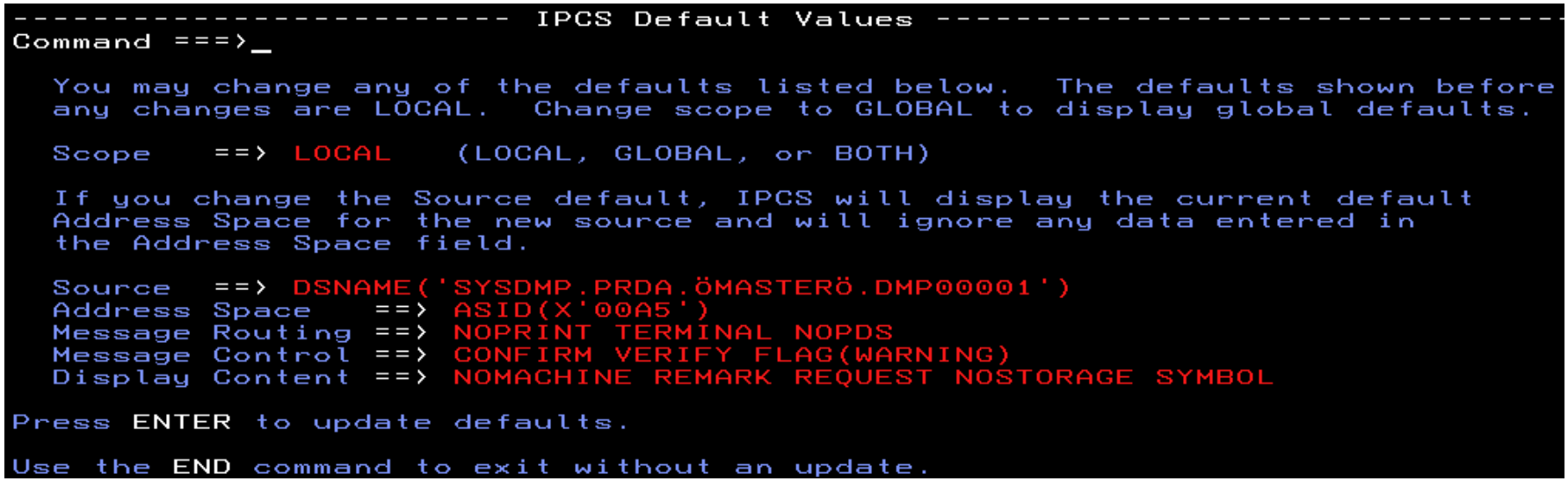
```
//SMT1RMF  JOB  MSGCLASS=X,CLASS=S,NOTIFY=&SYSUID,
//DUMPSMF  EXEC PGM=IFASMFDP,REGION=1M
//DUMPIN   DD  DSN=SYS3.SMF.PRDA.MVSARC2,DISP=SHR
//DUMPOUT  DD  DSN=&&DO,DISP=(NEW,CATLG),
//         SPACE=(CYL,(500,10)),DCB=(LRECL=137,RECFM=VBA,BLKSIZE=1693)
//SYSPRINT DD  SYSOUT=*
//SYSIN    DD  *
  INDD(DUMPIN,OPTIONS(DUMP))
  OUTDD(DUMPOUT,TYPE(70:79))
  START(1000)
  END(1015)
/*
//SORTSMF  EXEC PGM=SORT
//SORTIN   DD  DSN=&&DO,DISP=(OLD,DELETE)
//SYSOUT   DD  SYSOUT=*
//SORTOUT  DD DSN=&&SO,DISP=(,PASS),UNIT=SYSDA,
//         SPACE=(CYL,(5,10))
//SORTWK01 DD   SPACE=(TRK,200),UNIT=SYSDA
//SORTWK02 DD   SPACE=(TRK,200),UNIT=SYSDA
//SORTWK03 DD   SPACE=(TRK,200),UNIT=SYSDA
//EXITLIB  DD  DSN=SYS1.SERBLINK,DISP=SHR
//SYSIN    DD *
  SORT FIELDS=(11,4,CH,A,7,4,CH,A),EQUALS
  MODS E15=(ERBPPE15,36000,,N),E35=(ERBPPE35,3000,,N)
/*
//POSTRMF  EXEC PGM=ERBRMFPP,REGION=32M
//MFPINPUT DD  DSN=&&SO,DISP=(OLD,DELETE)
//MFPMSGDS DD SYSOUT=*
//PPRPTS   DD SYSOUT=*
//SYSIN    DD  DDNAME=SREP
//SREP     DD  *
  SYSRPTS(CF)
  RTOD(0000,2359)
  STOD(0000,2359)
  DINTV(0015)
  SYSOUT(X)
  SUMMARY(INT,TOT)
```

How To Find Out MWASDT Using IPCS

How To Find Out MWASDT Using IPCS

Select 0 ‘Defaults’ To Update Dump Dataset





How To Find Out MWASDT Using IPCS

Select 2 ‘Analysis’ And Then 6 ‘Component’

```
------------------- IPCS MVS ANALYSIS OF DUMP CONTENTS -----------------------

To display information, specify the corresponding option number.

    1   SYMPTOMS     - Symptoms                          **********************
    2   STATUS       - System environment summary        * USERID   - IMT3
    3   WORKSHEET    - System environment worksheet      * DATE     - 11/07/26
    4   SUMMARY      - Address spaces and tasks          * JULIAN   - 11.207
    5   CONTENTION   - Resource contention               * TIME     - 23:54
    6   COMPONENT    - MVS component data                * PREFIX   -
    7   TRACES       - Trace formatting                  * TERMINAL- 3278
                                                         * PF KEYS  - 24
                                                         **********************

Enter END command to terminate MVS dump analysis.
```

```
------------------- IPCS MVS DUMP COMPONENT DATA ANALYSIS --------------

To display information, specify US option nameU or enter S to the left
of the option desired.  Enter ? to the left of an option to display
help regarding the component support.

S Name       Abstract
_ NUCMAP     Nucleus CSECT Map
_ OAMDATA    OAM Control Block Analysis
_ OMVSDATA   OpenMVS analysis
_ RACFDATA   RACF control block analysis
_ RESOLVER   TCP/IP Resolver Analysis
_ RMMDATA    RMM Control Block Analysis
_ RMMPDA     RMM PDA Trace Analysis
_ RSMDATA    Real storage manager summary
_ SADMPMSG   Format SADMP console messages
_ SMSDATA    SMS control block analysis
_ SMSXDATA   SMSX Control Block Formatter
_ SRMDATA    SRM control block analysis
_ SSIDATA    Subsystem Interface analysis
_ STRDATA    Coupling Facility Structure Data
_ SUMDUMP    Format summary dump data
_ SYMDEF     Static Symbol Table Formatter
_ SYMPTOMS   Format symptoms
_ SYSTRACE   Format system trace
_ TCPIP      Tcp/Ip Dump Analysis
_ TCPIPCS    TCP/IP Analysis
_ TSODATA    TSO analysis
_ VLFDATA    Virtual Lookaside Facility data
_ VLFTRACE   Virtual Lookaside Facility trace
_ VSMDATA    VSM control block analysis
_ VTAM       Vtam Dump Analysis
_ VTAMMAP    VTAM control block analysis
_ WLMDATA    Work load manager data
S XESDATA    XES analysis
********************************* END OF LIST ********************************
```

How To Find Out MWASDT Using IPCS

Select ‘Detail’ And Then ‘S’ Command

```
------------------------- IPCS - XESDATA SUBCOMMAND ---------------------
SELECT OPTION ===>

 Select zero or more levels of detail. Default is SUMMARY reporting.
   _  SUMMARY          _  EXCEPTION      *  DETAIL

 Select zero or more report types. Default is ALL report types.
   _  CONNECTION       _  FACILITY       _  CACHE          _  LOCKMGR
   _  XESSTACK         _  LIST           _  LOCKRESOURCE   _  LOCK

 Select zero or more filtering options. Default is NO filtering.
 Filters that do not apply to a given report will be ignored.
 Additional filtering selection menus may be presented:
   _  ASID             _  HASHVALUE      _  SYSNAME        _  REQUESTORCONID
   _  RNAME            _  TARGETNAME     _  STRNAME        _  REQID
   _  SOURCENAME       _  CONNAME        _  CFNAME         _  LISTNUM
   _  JOBNAME          _  LTENTRY        _  LOCKMGRCONID

 XESDATA DETAIL

  S  = START XESDATA subcommand.
  R  = Reset all panel variables.
  END = Terminate XESDATA subcommand.
```

```
IPCS OUTPUT STREAM ------------------------------------------------ Line 0
  ********************************  TOP OF DATA  ***********************
    *****   XESDATA (CROSS-SYSTEM EXTENDED SERVICES) REPORT   *****

 Options list:
     Report(s)..............  CONNECTION
                              XESSTACK
                              LOCKRESOURCE
                              LOCKMGR
                              FACILITY
                              LIST
                              CACHE
                              LOCK
     Level(s) of detail.....  DETAIL
     Filter(s) in use.......  NONE

 Sysplex name...............  PPLEX
 System name................  PRDA

 Facility name..............  PCF1GAR1
     Structure name.........  IXCSIG1
         ASID...............  X'0006'
             Connection name..  SIGPATH_06000302
     Structure name.........  IXCSIG11
         ASID...............  X'0006'
             Connection name..  SIGPATH_06000302
```

How To Find Out MWASDT Using IPCS

Sample MWASDT 31 microseconds For PCF1GAR1

```
Queued Request Information:

   Facility Name..................... PCF1GAR1
      Low Priority Work Queue
         Number of Queued Requests... 0                  (decimal)
         Total Number of Requests.... 305193             (decimal)
         Time of Last Queued Request. 07/23/2011 22:20:04.270891
      High Priority Work Queue
         Number of Queued Requests... 0                  (decimal)
         Total Number of Requests.... 2429               (decimal)
         Time of Last Queued Request. 07/23/2011 22:12:04.078669

Moving Weighted Average Subchannel Delay Time (MWASDT) Information:

   Refresh Counter.............. 2                       (decimal)
   Refresh Limit................ 250                     (decimal)
   Queued Count................. 2777                    (decimal)
   Total Count.................. 2882                    (decimal)
   MWASDT (in microseconds)..... 31                      (decimal)

Sync/Async Heuristics Data
--------------------------

Simplex Requests:

  OpCode Acronym    Size    ReqCount ConvReqCount  Avg Svc Time
 ------------------------------------------------------------
   0301     ALST     0- 0           0            0           41
   0303     RLSC     0- 0           0            0           18
   0303     RLSC     1- 1           0            0           18
   0304     RLC      1- 1       27036            0           15
   0305     WLC      0- 0        4837            0           13
   0306     ALSU     0- 0           0            0           14
   0307     DLSU     0- 0           0            0           20
   0308     RLM      0- 0     4447777       391000           25
```

How To Find Out MWASDT Using IPCS

Sample MWASDT 6 microseconds For PCF1GAR1

```
Queued Request Information:

   Facility Name..................... PCF1GAR1
      Low Priority Work Queue
         Number of Queued Requests... 0                    (decimal)
         Total Number of Requests.... 0                    (decimal)
      High Priority Work Queue
         Number of Queued Requests... 0                    (decimal)
         Total Number of Requests.... 0                    (decimal)

   Facility Name..................... PCF2GAR2
      Low Priority Work Queue
         Number of Queued Requests... 0                    (decimal)
         Total Number of Requests.... 115497               (decimal)
         Time of Last Queued Request. 07/23/2011 21:41:39.330957
      High Priority Work Queue
         Number of Queued Requests... 0                    (decimal)
         Total Number of Requests.... 21                   (decimal)
         Time of Last Queued Request. 07/22/2011 07:43:06.658744

Moving Weighted Average Subchannel Delay Time (MWASDT) Information:

   Refresh Counter.............. 0                         (decimal)
   Refresh Limit................ 250                       (decimal)
   Queued Count................. 0                         (decimal)
   Total Count.................. 7                         (decimal)
   MWASDT (in microseconds)..... 6                         (decimal)

Sync/Async Heuristics Data
--------------------------

Simplex Requests:

  OpCode Acronym    Size    ReqCount ConvReqCount   Avg Svc Time
 ------------------------------------------------------------
   0301     ALST     0- 0           0            0             30
   0303     RLSC     0- 0           0            0             29
```

GT Structure Distribution – 2 CFs

GT Parallel Sysplex Configuration - CFs & Structures

PCF1GAR1

DSNPD01_LOCK1	DSNPD01_GBP0	RLS_APL2
DSNPD01_SCA	DSNPD01_GBP1	RRSSTR1
DSNPDRM_GBP0	DSNPD01_GBP16K0	SYSTEM_OPERLOG
DSNPDRM_GBP1	DSNPD01_GBP16K1	SYSZWLM_0E162817
DSNPDRM_GBP2	DSNPD01_GBP2	LOG_DFHLOG_WUI
DSNPDRM_GBP8K0	DSNPD01_GBP21	LOG_DFHSHUNT_WUI
DSNPDRM_LOCK1	DSNPD01_GBP22	PQS1CSQ_ADMIN
DSNPDRM_SCA	DSNPD01_GBP23	PQS1OLASTR
EZBEPORT	DSNPD01_GBP24	PQS1SMSSTR
HSA_LOG	DSNPD01_GBP31	PQS1SYSPSTR
HZS_HEALTHCHKLOG	DSNPD01_GBP32	DSNPD01_GBP5
IBMBDG	DSNPD01_GBP32K	DSNPD01_GBP6
IXCSIG1	DSNPD01_GBP33	DSNPD01_GBP7
IXCSIG11	DSNPD01_GBP34	DSNPD01_GBP8K0
IXCSIG2	IXCSIG21	IXCSIG4
IXCSIG7		

GT Parallel Sysplex Configuration - CFs & Structures

PCF2GAR2

DSNPDRM_GBP0	IXCSIG5	SYSZWLM_0E262817
DSNPDRM_GBP1	IXCSIG6	TOPSTR1
DSNPDRM_GBP2	PQS1APPLSTR	CKPT1
DSNPDRM_GBP8K0	PQS1FFMCSTR	DFHNCLS_PRODNC1
EZBEPORT0111	PQS1FFMDSTR	DFHXQLS_PRODTSQ1
EZBEPORT0113	PQS1LOGOSTR	DSNPD01_GBP0
IGWLOCK00	PQS1OTPSTR	DSNPD01_GBP1
ISGLOCK	PQS1UTLSTR	DSNPD01_GBP16K0
ISTGENERIC	RLS_APL1	DSNPD01_GBP16K1
IXCSIG3	SYSARC_HSMPP_RCL	DSNPD01_GBP2
IXCSIG31	SYSIGGCAS_ECS	DSNPD01_GBP21
DSNPD01_GBP33	DSNPD01_GBP31	DSNPD01_GBP22
DSNPD01_GBP34	DSNPD01_GBP32	DSNPD01_GBP23
DSNPD01_GBP5	DSNPD01_GBP32K	DSNPD01_GBP24
DSNPD01_GBP6	DSNPD01_GBP7	DSNPD01_GBP8K0