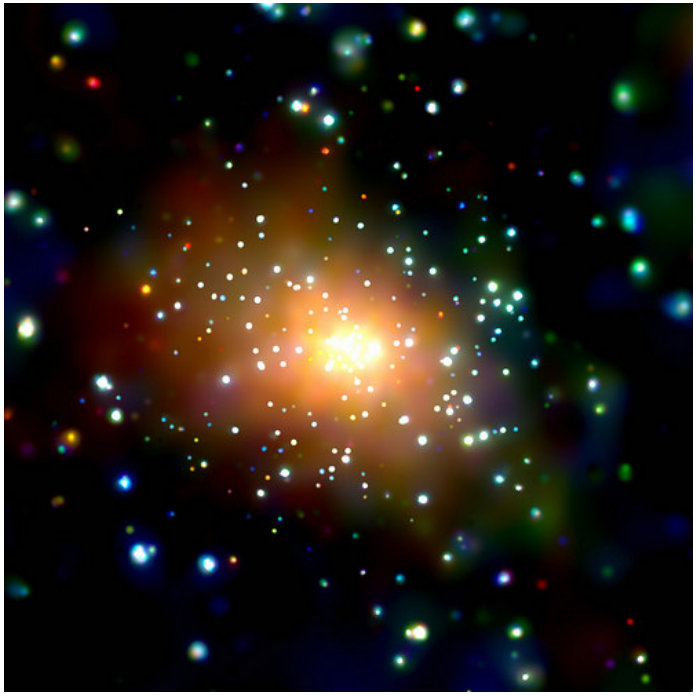
High Performance Ficon Demystified

Session 8759

Lou Ricci
lricci@us.ibm.com



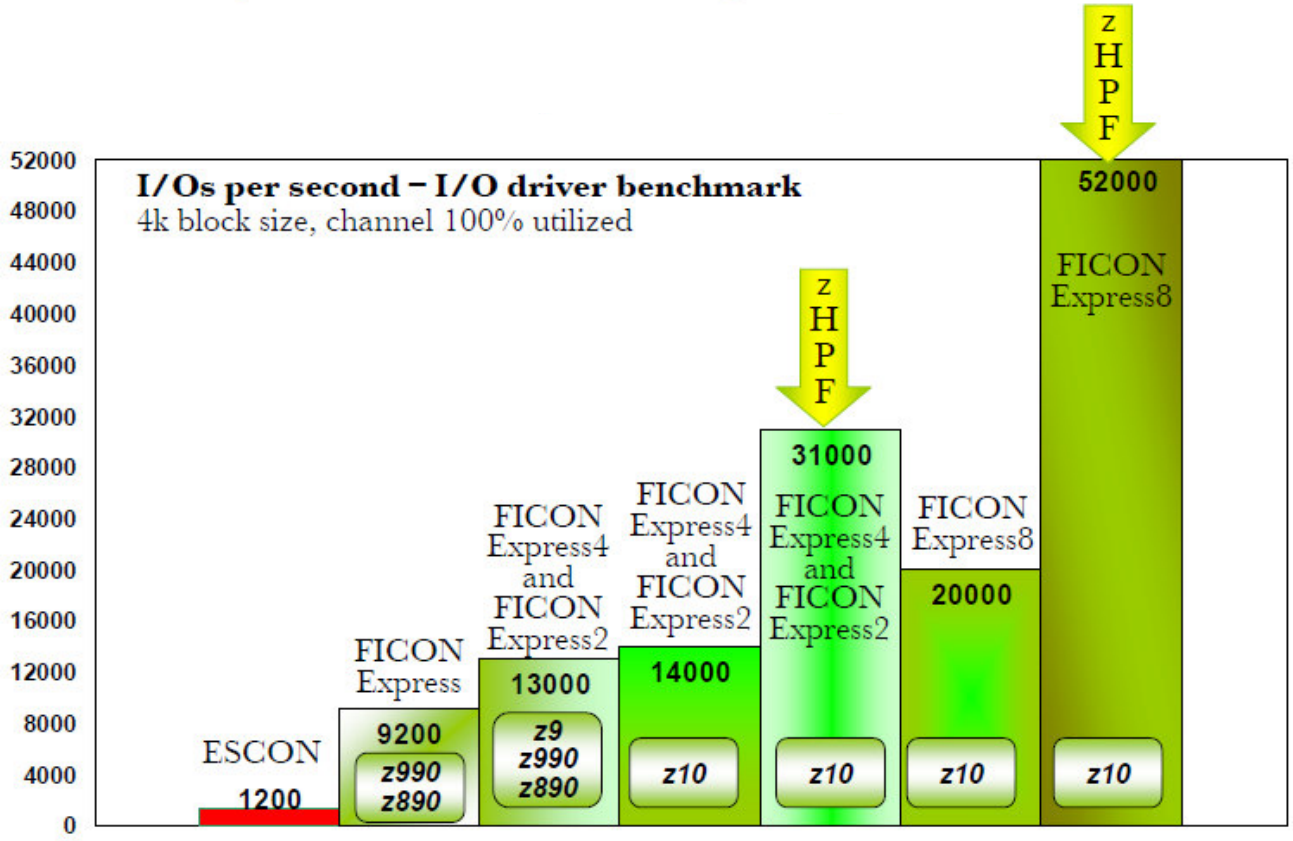
Legal Stuff

- Notice
  - IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing to: *IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.*
  - Any references in this information to non-IBM Web sites are provied for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.
- Trademarks
  - The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both: FICON® IBM® Redbooks™ System z10™ z/OS® zSeries® z10™
  - Other Company, product, or service names may be trademarks or service marks of others.

Agenda

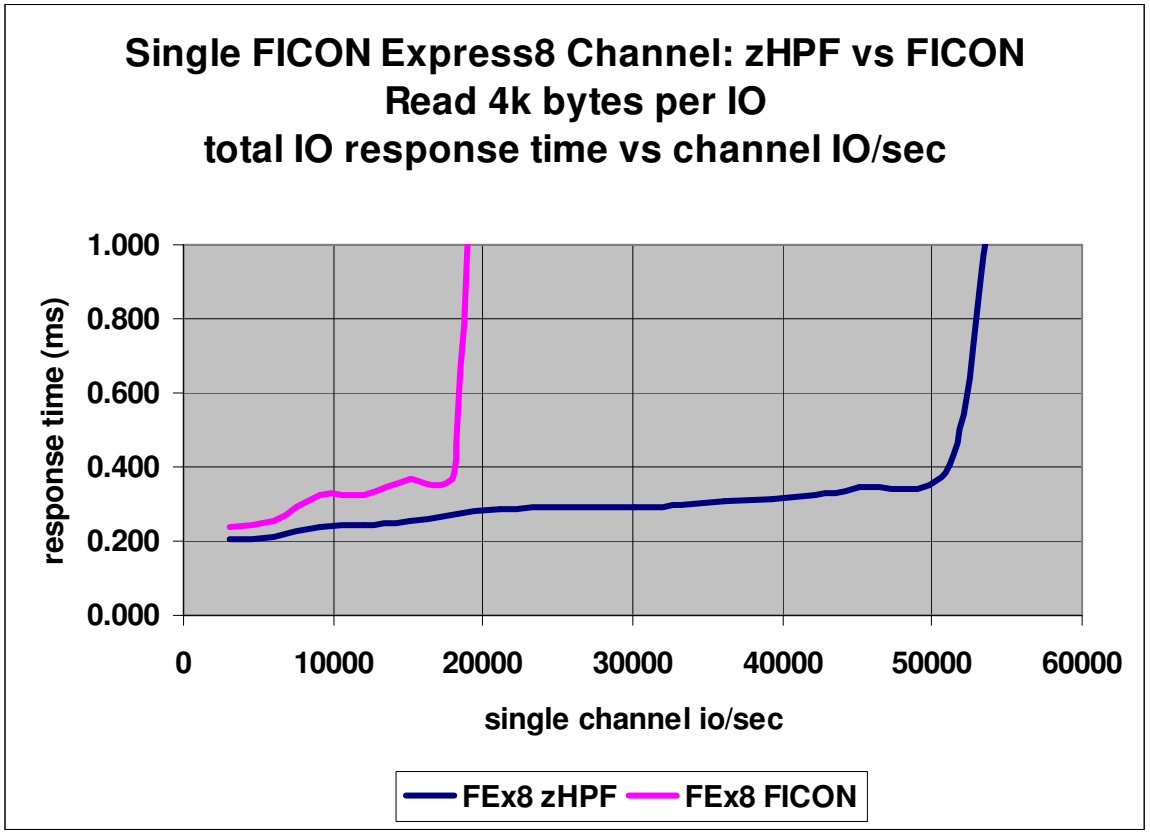
- What does zHPF Do For Me?
- How Does zHPF Do It?
- The Effect On Exchanges
- Other Improvements

More Than 2X Ficon I/Os per Second

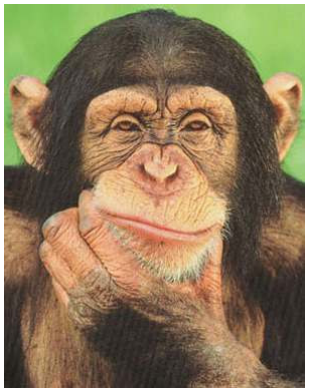


Response Time Improves Too

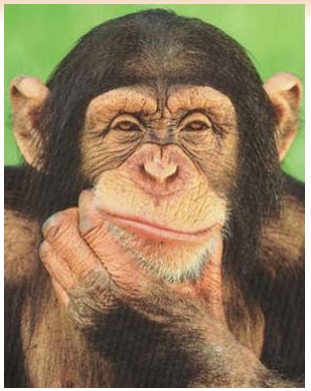
zHPF vs FICON READ 4k bytes/io



How does zHPF do it?

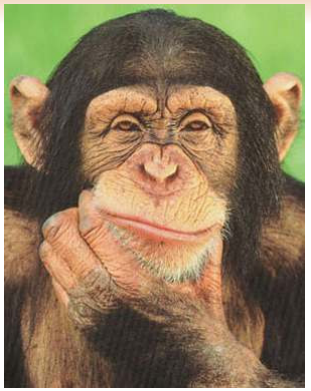


How does zHPF do it?

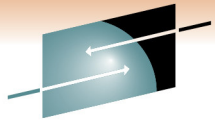


- Significantly reduced Channel and CU overhead

How does zHPF do it?



- Significantly reduced Channel and CU overhead
- Takes advantage of hardware assists in Fibre Channel interface chips
- Rides on top of an existing standard protocol called….



F. C. P.

FCP ???

- Does zHPF convert my I/O to SCSI ??????

  NO !

- FCP is a generic method to transfer commands, data, and status

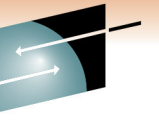
- FCP  SCSI
  - It is true however, that SCSI is the single largest user of FCP

Why FCP?

- Many HBA vendors have optimized firmware and hardware to accelerate FCP I/O

- FCP protocol has less ‘Chit Chat’

Read Comparison Summary (5 4K Reads)



SHARE
Technology · Connections · Results

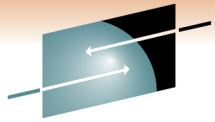
	Channel to CU in Ficon Mode	CU to Channel in Ficon Mode	Total	Channel To CU in zHPF Mode	CU to Channel in zHPF Mode	Total	% Reduction in zHPF Mode[1]
Exchanges	1	1	2	1	1	1	50
Sequences	6	6	12	1	2	3	75
Frames	6	14	20	1	10	11	45
CRC Gen / Check	5	5	10	1	1	2	80

[1]Except for exchanges, as the number of reads in a single I/O increase, the % reduction in Transport Mode increases

Let's look under the hood

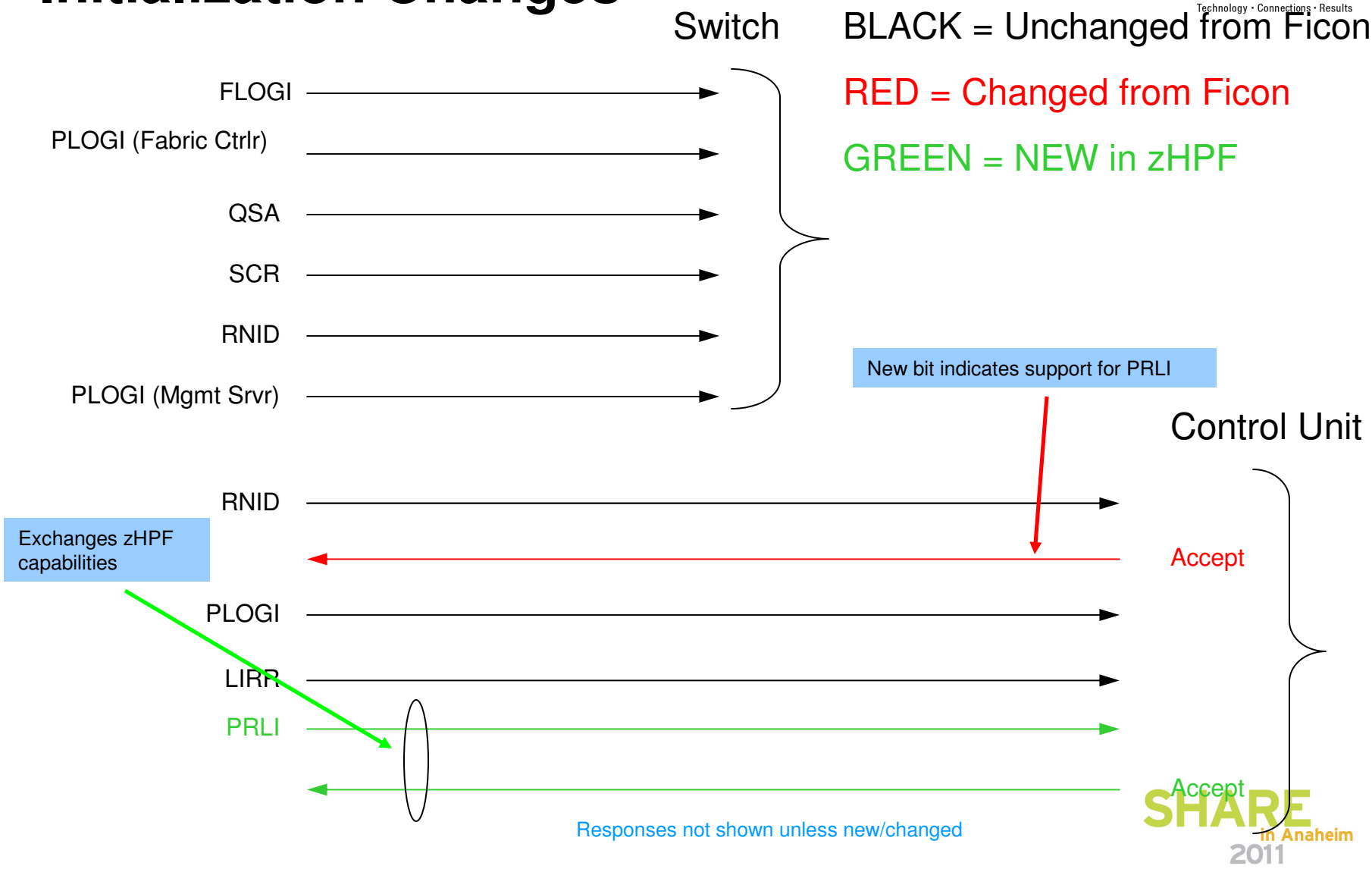
Image Credit: Flikr user aka Razz
Creative Commons License

NO I/O Definition changes



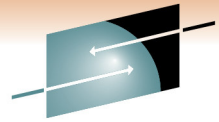
- zHPF coexists with Ficon
- Channel is STILL type=FC
- NO IOCDS changes for zHPF capable control units

Initialization Changes



Totally New I/O Structures

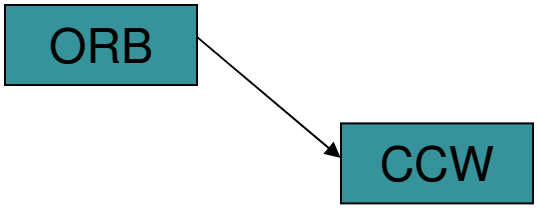
- CCWs no longer exist in zHPF (They live on happily in Ficon)
  - Replaced by Device Control Words (DCWs)
- IDAWs no longer exist in zHPF (They too are alive and well in Ficon)
  - Replaced by Transport Indirect Data Address Words (TIDAW)
- New structures added
  - Transport Control Word (TCW)
  - Transport Status Block (TSB)
  - Transport Command & Control Block (TCCB)



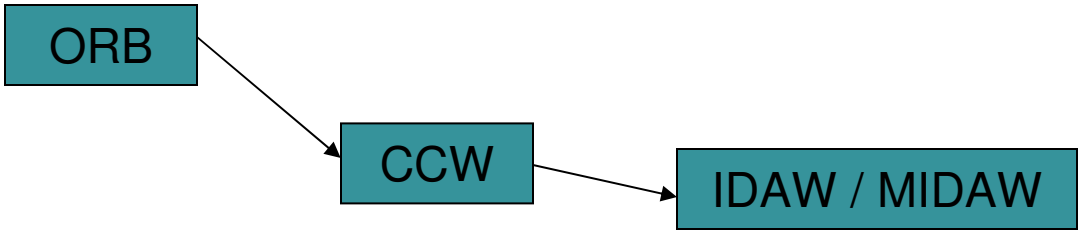
Command Mode Review

ORB

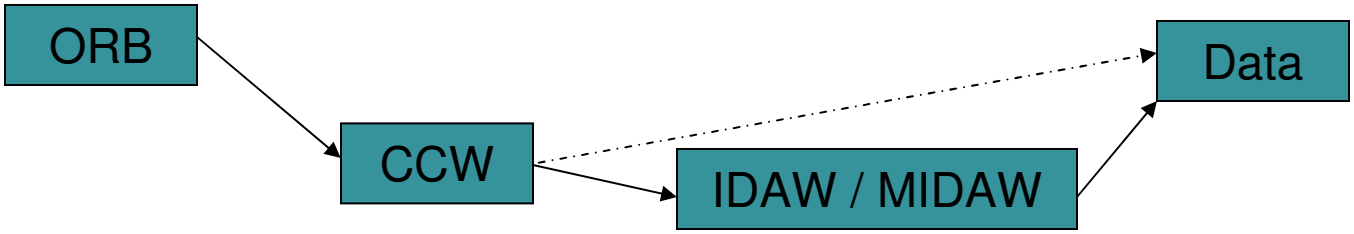
Command Mode Review



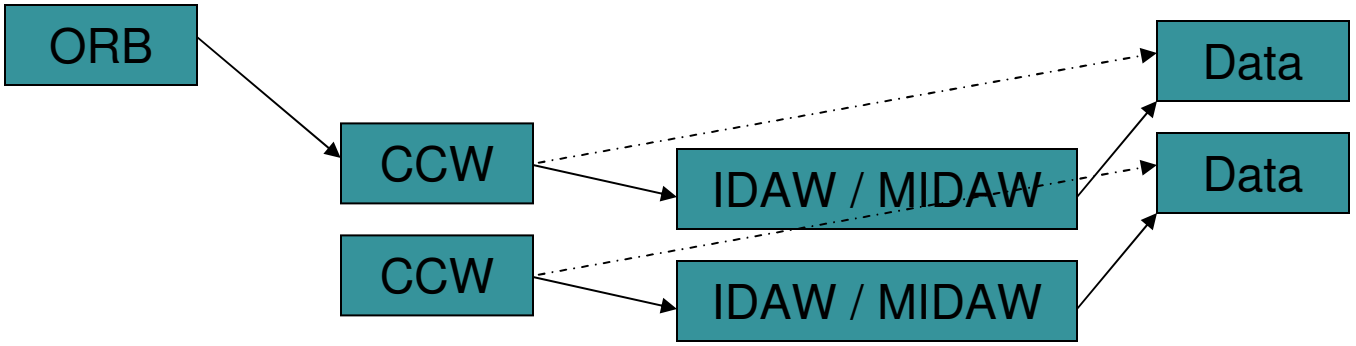
Command Mode Review



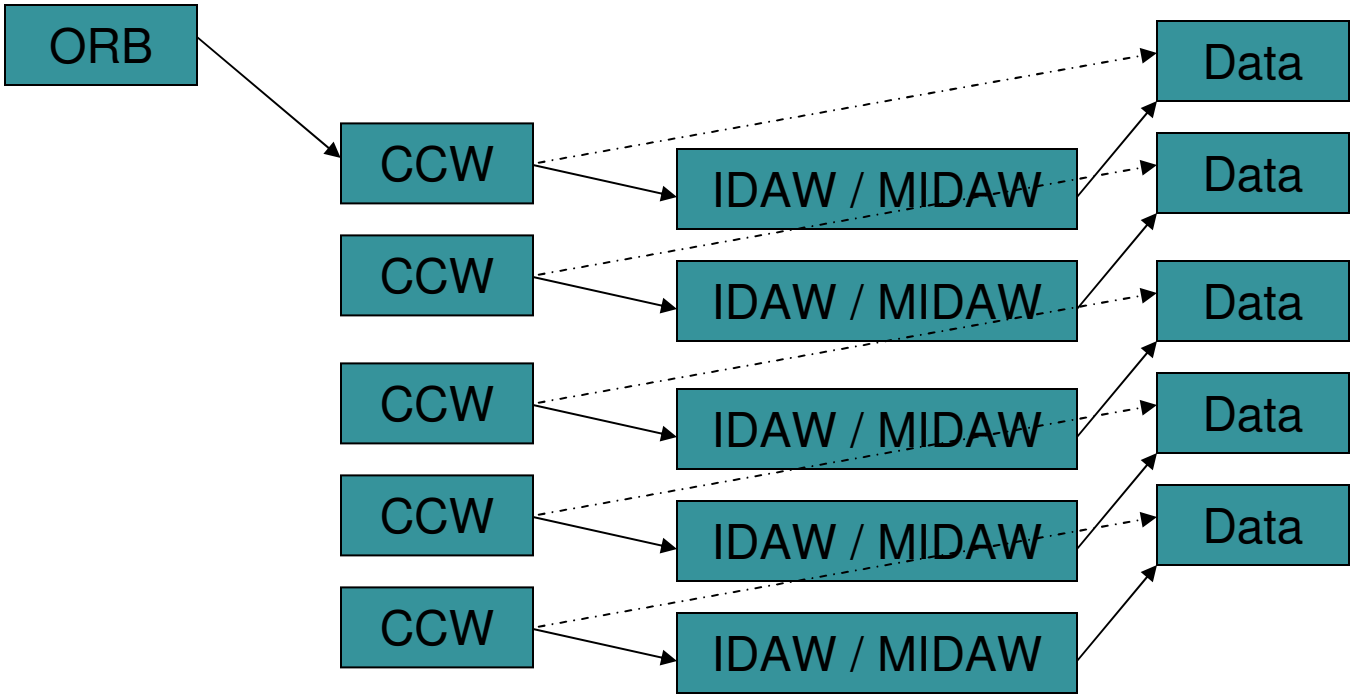
Command Mode Review



Command Mode Review

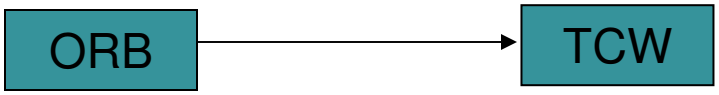


Command Mode Review

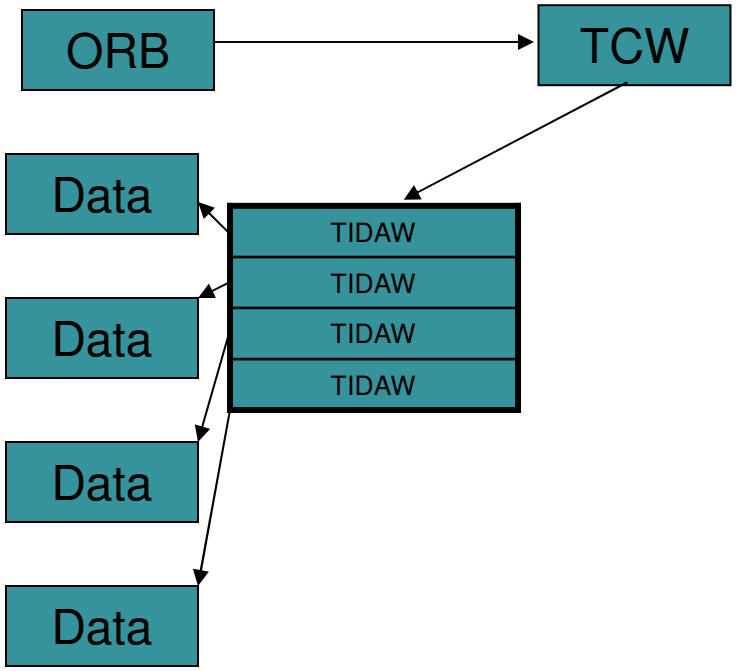


Transport Mode

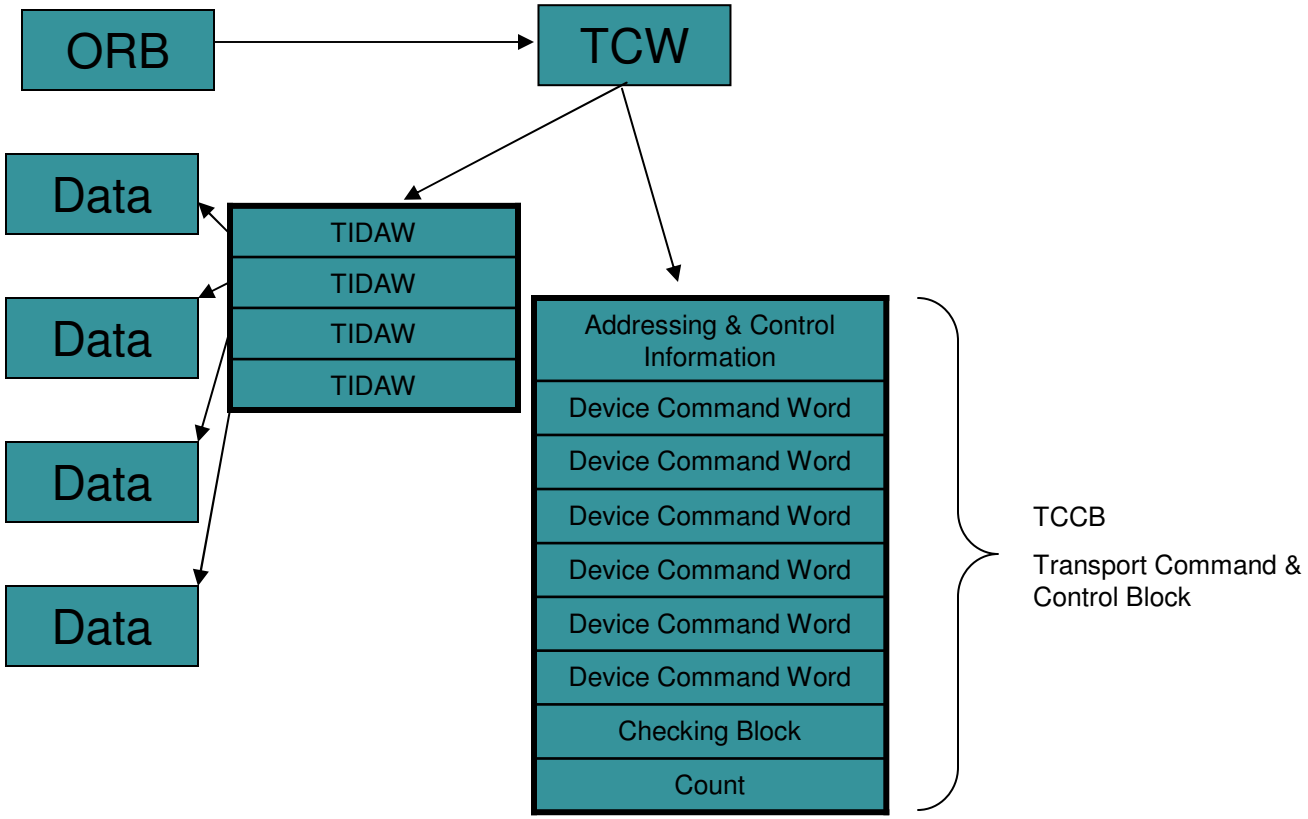
Transport Mode



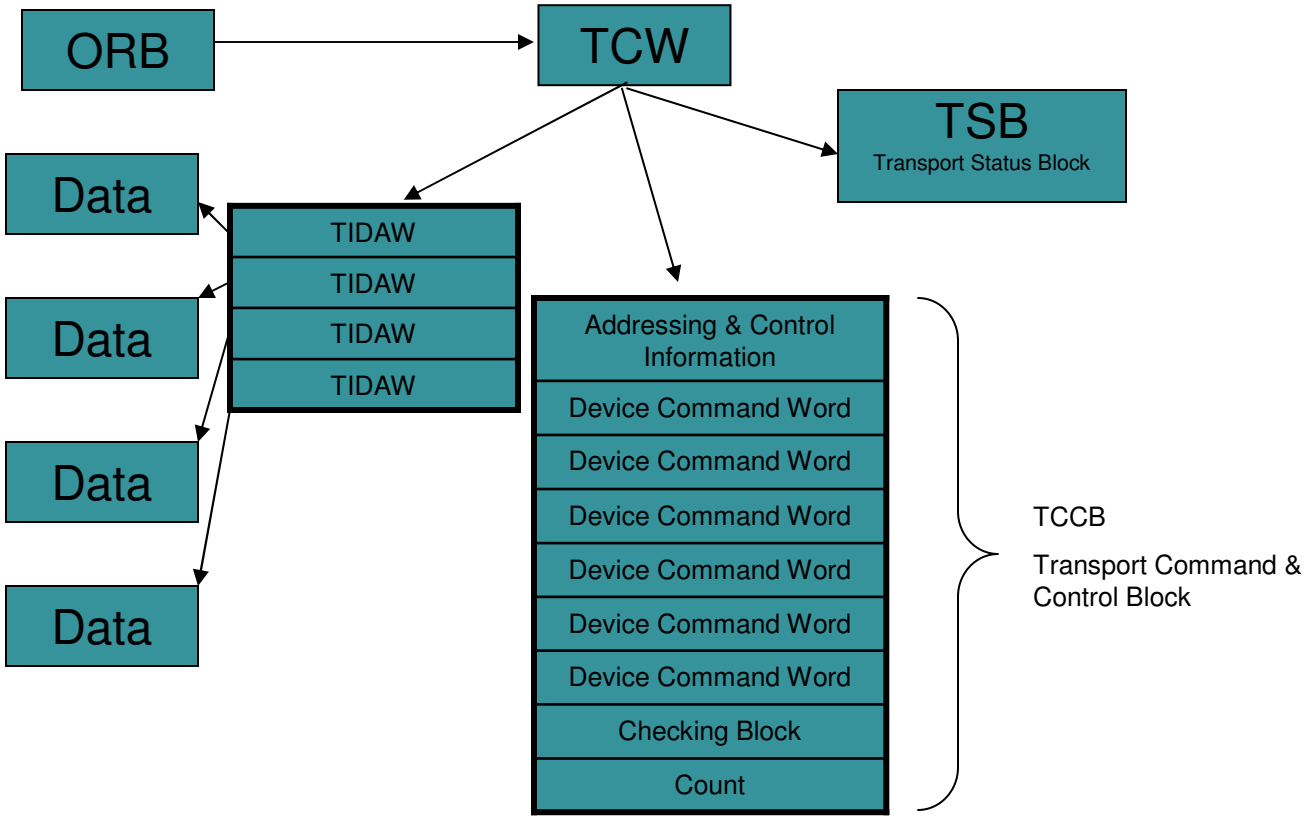
Transport Mode



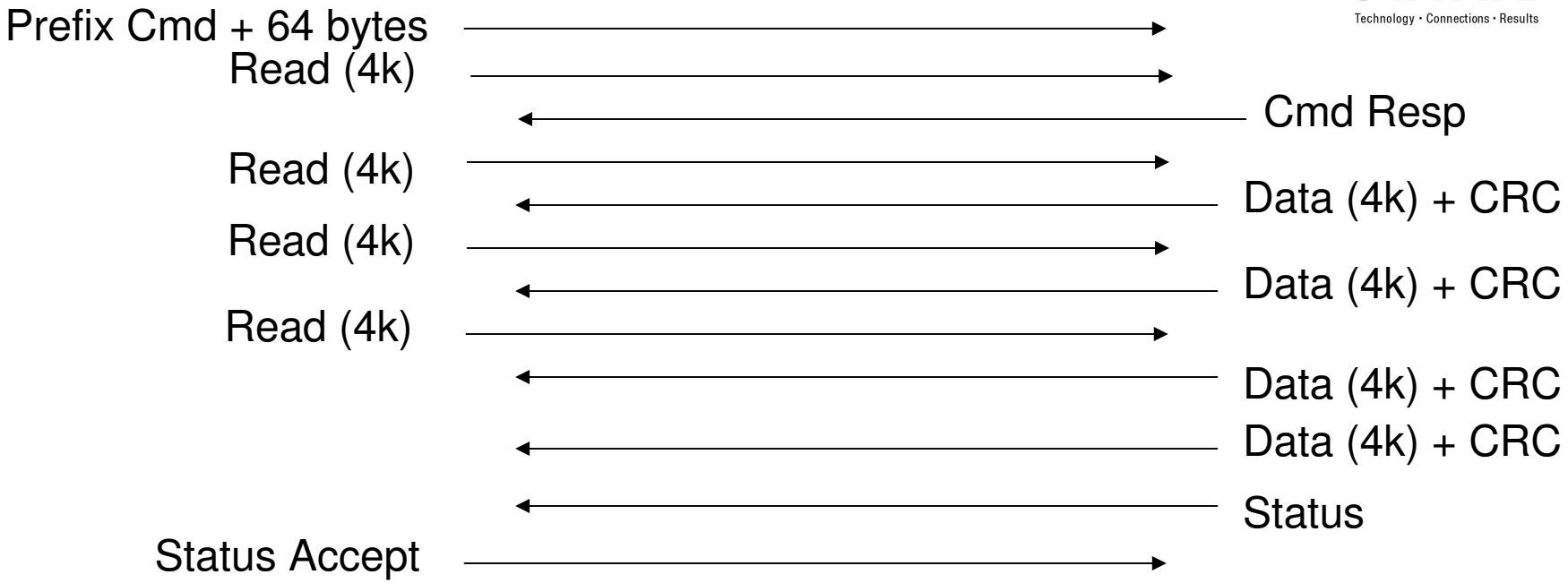
Transport Mode



Transport Mode

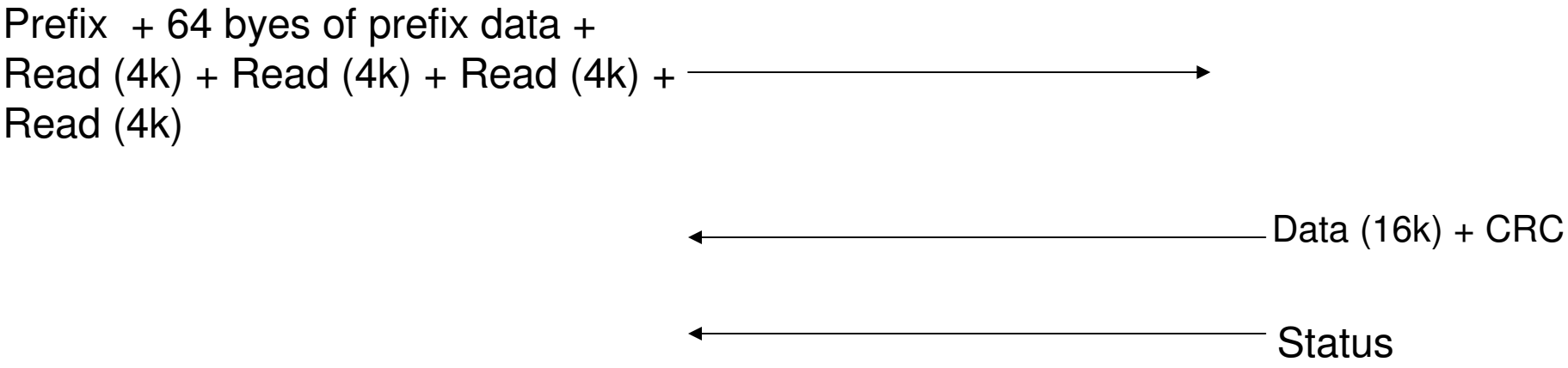


Link View of 4 Reads in Command Mode



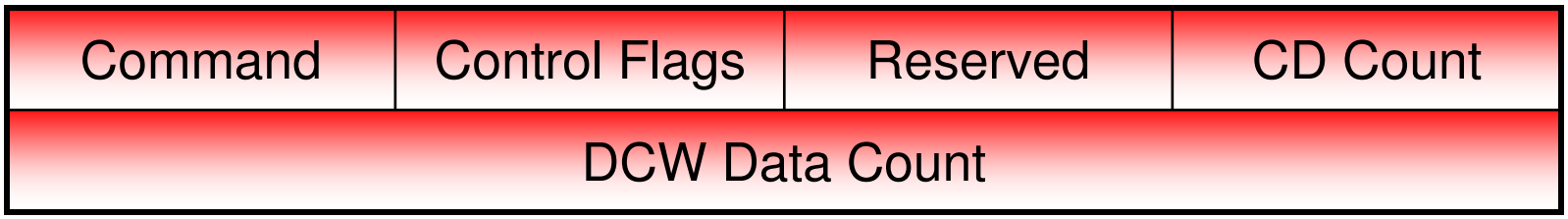
	Channel to Control Unit	Control Unit to Channel
Total Commands	5	N/A
Exchanges	2	2
Sequences	6	6
Frames	6	14
CRC Generate / Check	5	5

Link View of 4 Reads in Transport (zHPF) Mode

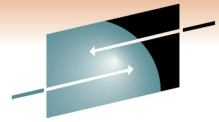


	Channel to Control Unit	Control Unit to Channel
Total Commands	5	N/A
Exchanges	1	1 (same one)
Sequences	1	2
Frames	1	10
CRC Generate / Check	1	1

Device Control Word (DCW)

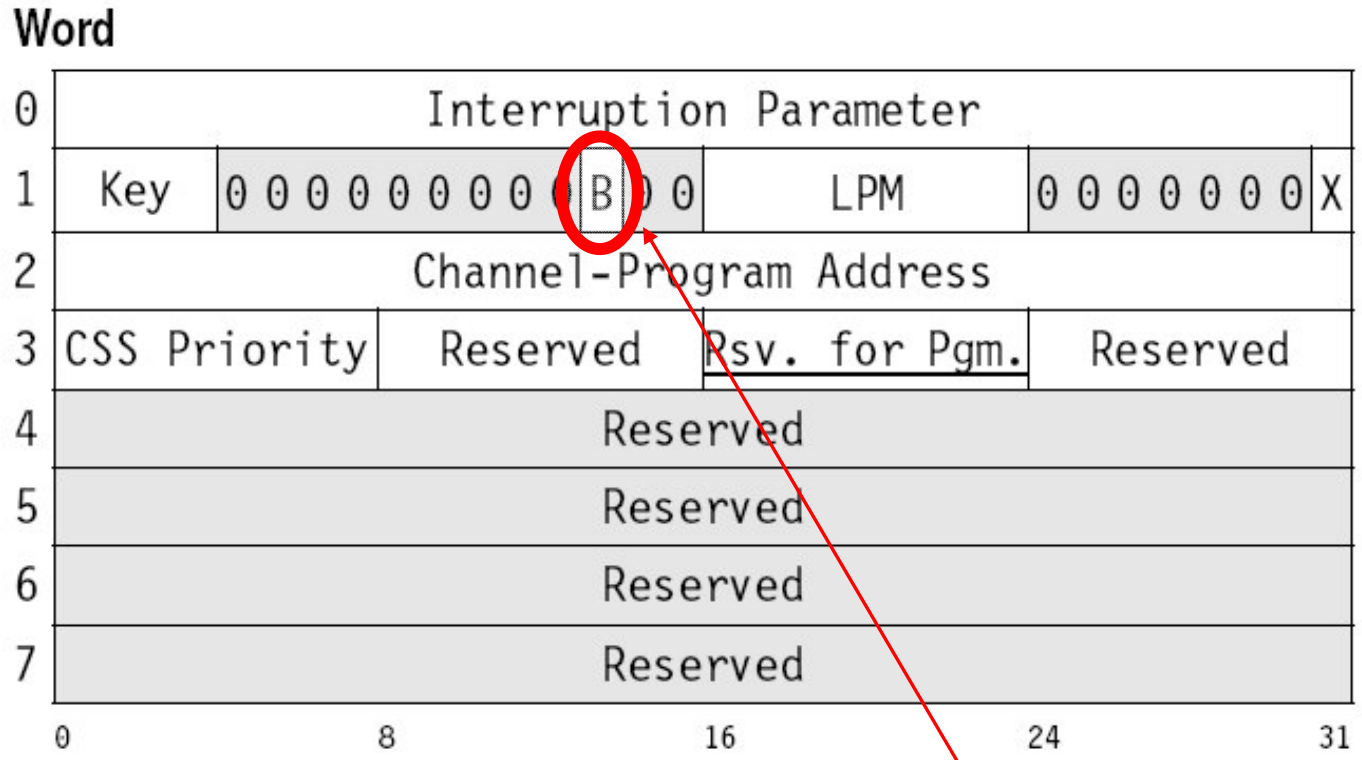


Device Control Word (DCW)



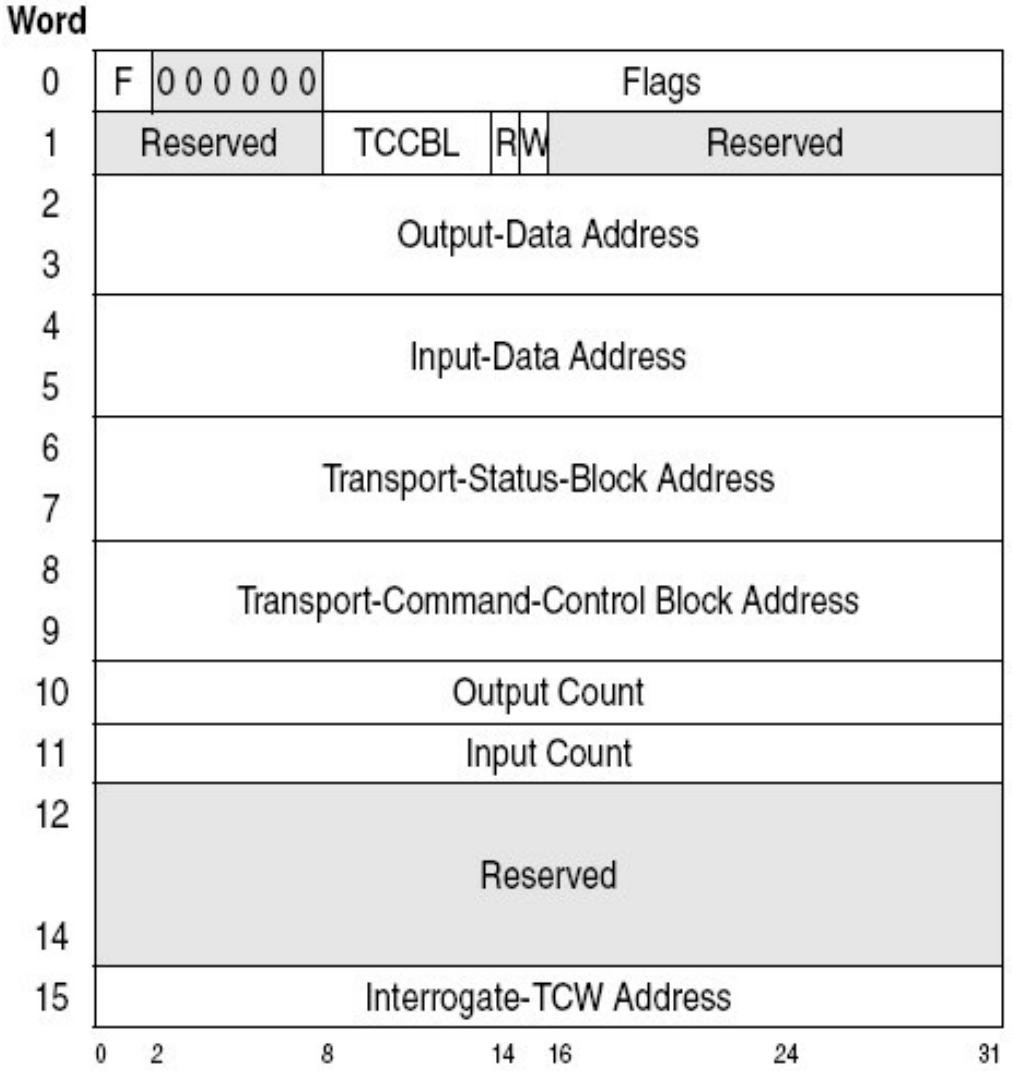
- Control Flags
  - CC (Chain Command)
    - Another command follows. If the command completes “normally” the next command is to be executed
- CD Count
  - Number of bytes that follow the DCW that contain Control Data for the DCW
- Data Count
  - Number of bytes of data to be transferred in the data phase for this DCW not including any Pad and CRC

ORB



Specifies Transport (zHPF) Mode

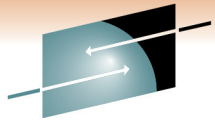
TCW (Transport Control Word)



Agenda

- What does zHPF Do For Me?
- How Does zHPF Do It?
- **The Effect On Exchanges**
- Other Improvements

How does zHPF affect EXCHANGES ?



- Little's Law states:
  - *The number of "things" in a system can be determined by multiplying the average arrival rate of those "things" by the average time each "thing" stays in the system.*

How does zHPF affect EXCHANGES ?

- Little's Law states:
  - *The number of "things" in a system can be determined by multiplying the average arrival rate of those "things" by the average time each "thing" stays in the system.*
  - Applied to zHPF:
    - The average number of Exchanges active at any given time = Average I/O rate * Average response time
    - Example: 30000 Ficon I/Os / Second on a given channel with .3ms service time[1] uses 9 Active Exchanges at any given time

[1] The amount of time the I/O is active in the channel

How does zHPF affect EXCHANGES ?

- The CU holds on to the Exchange even if the device:
  - Is reserved
  - Detects an Extent Conflict
  - Cache Miss
  - etc
- Drives requirement for higher number of possible open Exchanges

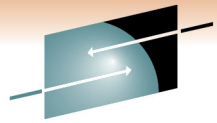
Example:

At 50,000 I/Os per Second, if 20% hit one of the above and If each of those conditions lasts for 10ms, then:

100 Exchanges are needed for Busies

20 Exchanges are needed for the rest

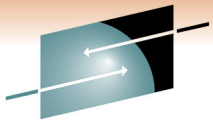
How does zHPF affect EXCHANGES ?



- CU can dynamically adjust the number of open Exchanges any one channel can open to THAT CU

- Channel maintains a Exchange count and Exchange Limit for each PHYSICAL control unit

New RMF Fields for zHPF



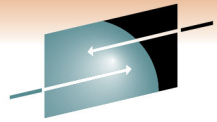
SHARE
Technology · Connections · Results

CHANNEL PATH			UTILIZATION(%)			READ(MB/SEC)		WRITE(MB/SEC)		FICON OPERATIONS			ZHPF OPERATIONS		
ID	TYPE	G SHR	PART	TOTAL	BUS	PART	TOTAL	PART	TOTAL	RATE	ACTIVE	DEFER	RATE	ACTIVE	DEFER
00	FC_S	5 Y	100.00	100.00	0.84	0.13	2.15	0.17	2.68	61.5	1.7	0.0	4.6	1.0	0.0
01	FC_S	5 Y	100.00	100.00	0.85	0.13	2.21	0.13	2.69	61.3	1.8	0.0	4.7	1.0	0.0
02	FC_S	4 Y	0.14	2.30	0.85	0.10	2.17	0.13	2.70	61.3	1.3	0.0	4.6	1.0	0.0
03	FC_S	4 Y	0.13	2.27	0.84	0.11	2.14	0.13	2.66	60.0	1.3	0.0	4.4	1.0	0.0
04	FC_S	5 Y	0.13	2.24	0.82	0.10	2.07	0.13	2.63	59.4	1.7	0.0	4.4	1.0	0.0
05	FC_S	5 Y	0.13	2.25	0.83	0.10	2.11	0.12	2.66	59.1	1.7	0.0	4.2	1.0	0.0
06	FC_S	4 Y	0.12	2.23	0.83	0.10	2.09	0.13	2.68	58.7	1.3	0.0	4.2	1.0	0.0

Where is zHPF exploited?

- All Supported Releases of z/OS
  - Media Manager Data Sets
    - DB2
    - VSAM
    - Extended Format Sequential Data Sets
    - PDSEs
    - Others

What Do I Need to Exploit zHPF?



- Z10 at Driver 76 or higher
  - Power On Reset is REQUIRED to activate zHPF
- z196

- Ficon Express-2 or above

- Control Unit that supports zHPF
  - Check with your vendor for appropriate code and/or hardware levels

- All supported releases of z/OS
  - zHPF mode has to be enabled (set ios command)

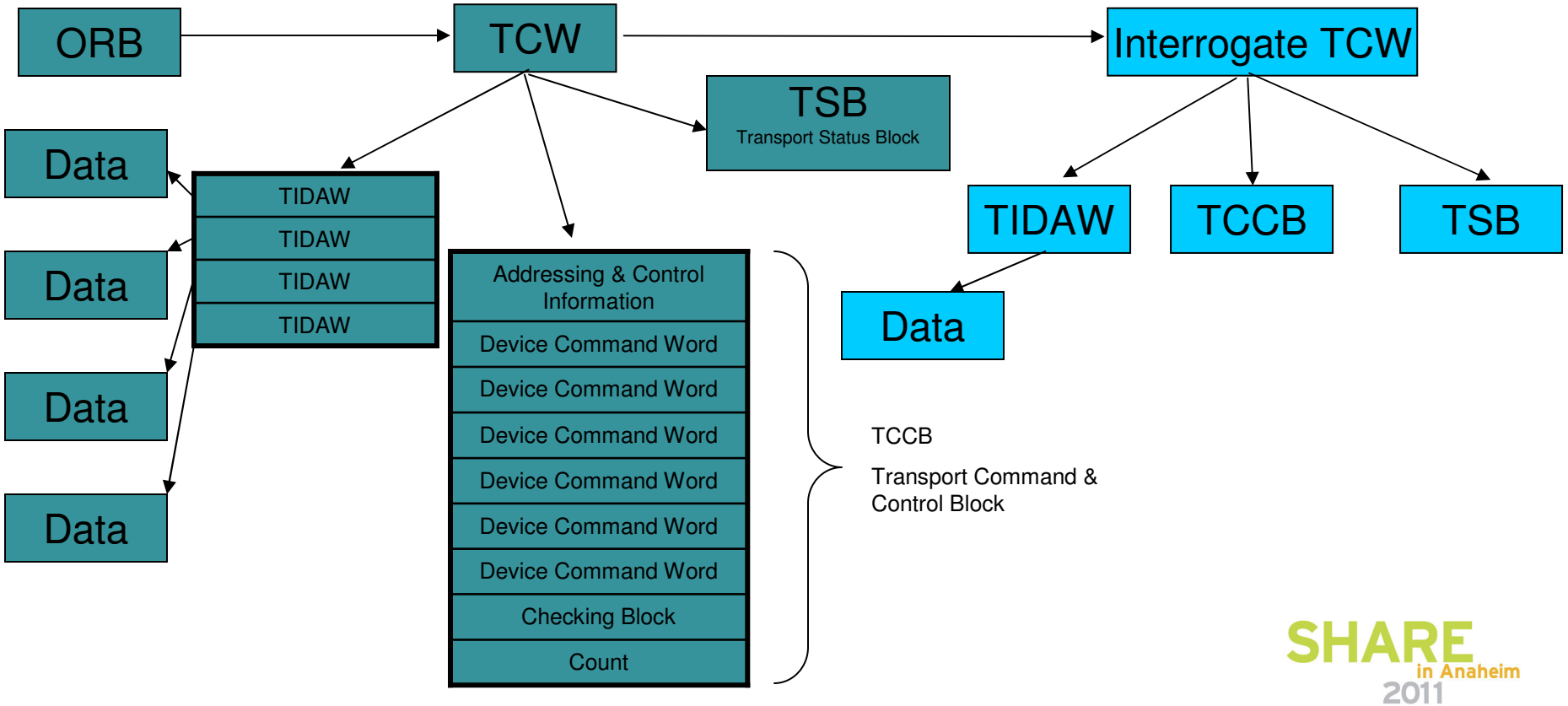
Agenda

- What does zHPF Do For Me?
- How Does zHPF Do It?
- The Effect On Exchanges
- **Other Improvements**

MIH

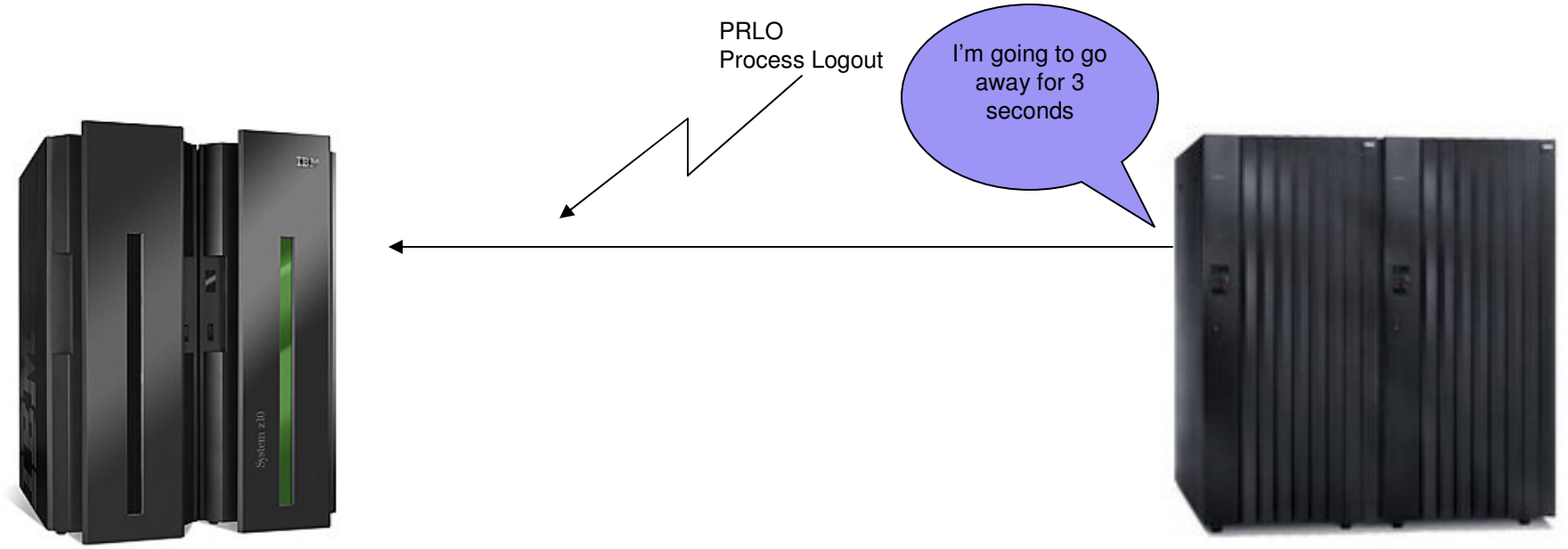
- Reduced False Missing Interrupt
  - Avoids “Go to the end of the line” penalty for MIH due to reserves
  - HPF allows the OS to interrogate the state of an existing I/O operation

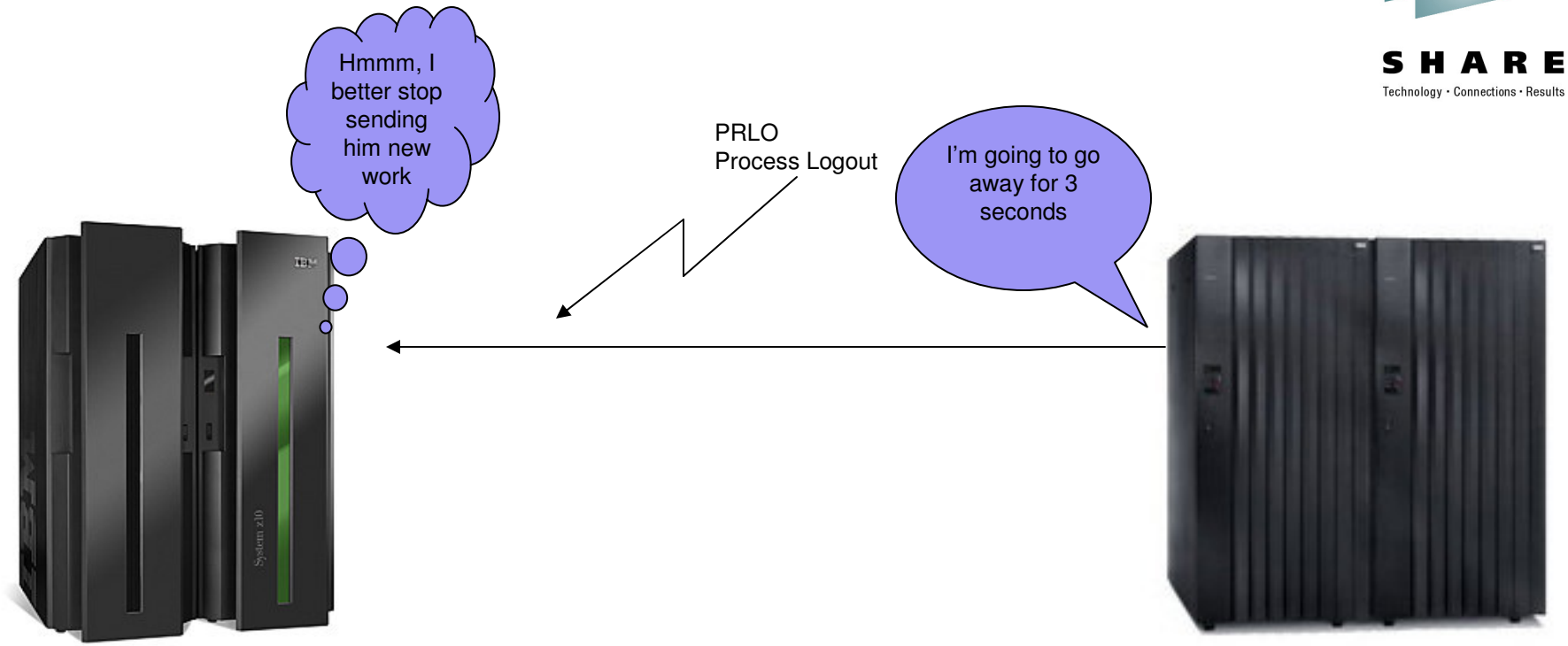
Transport Mode

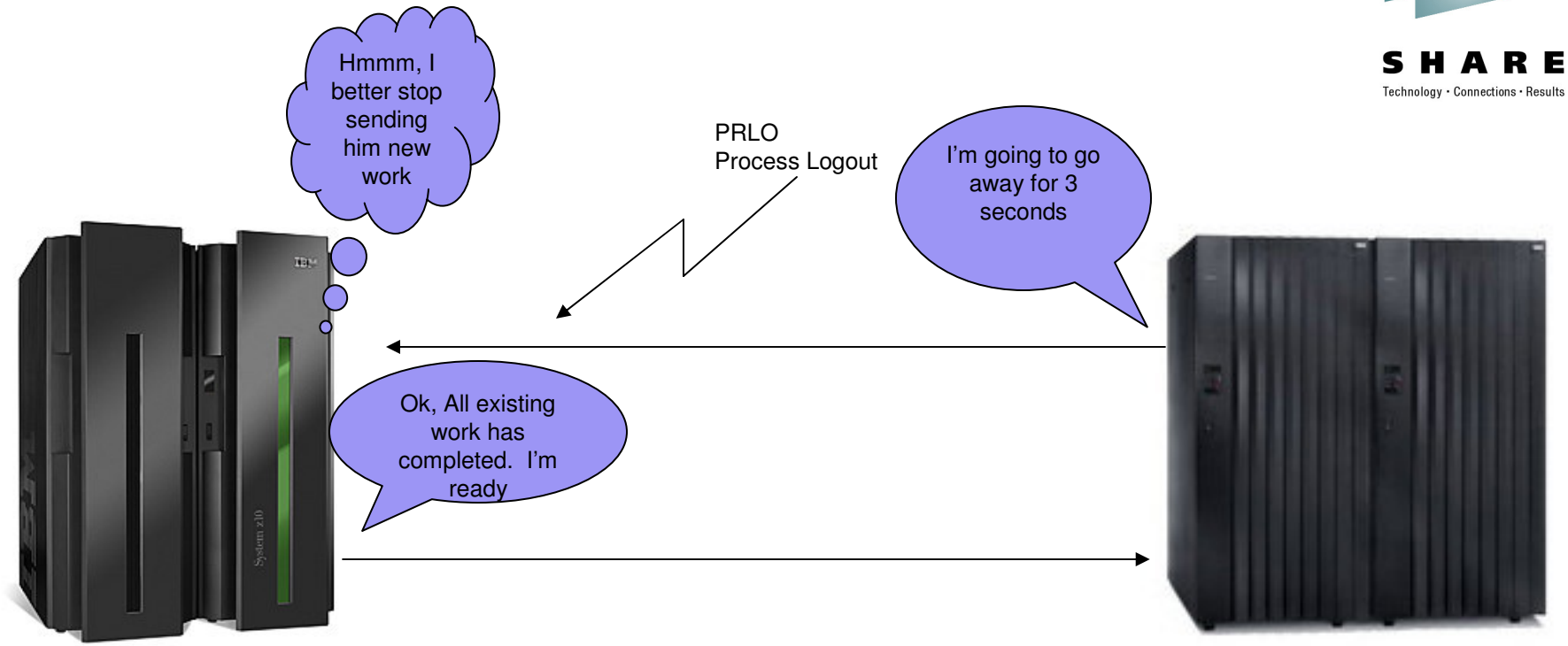


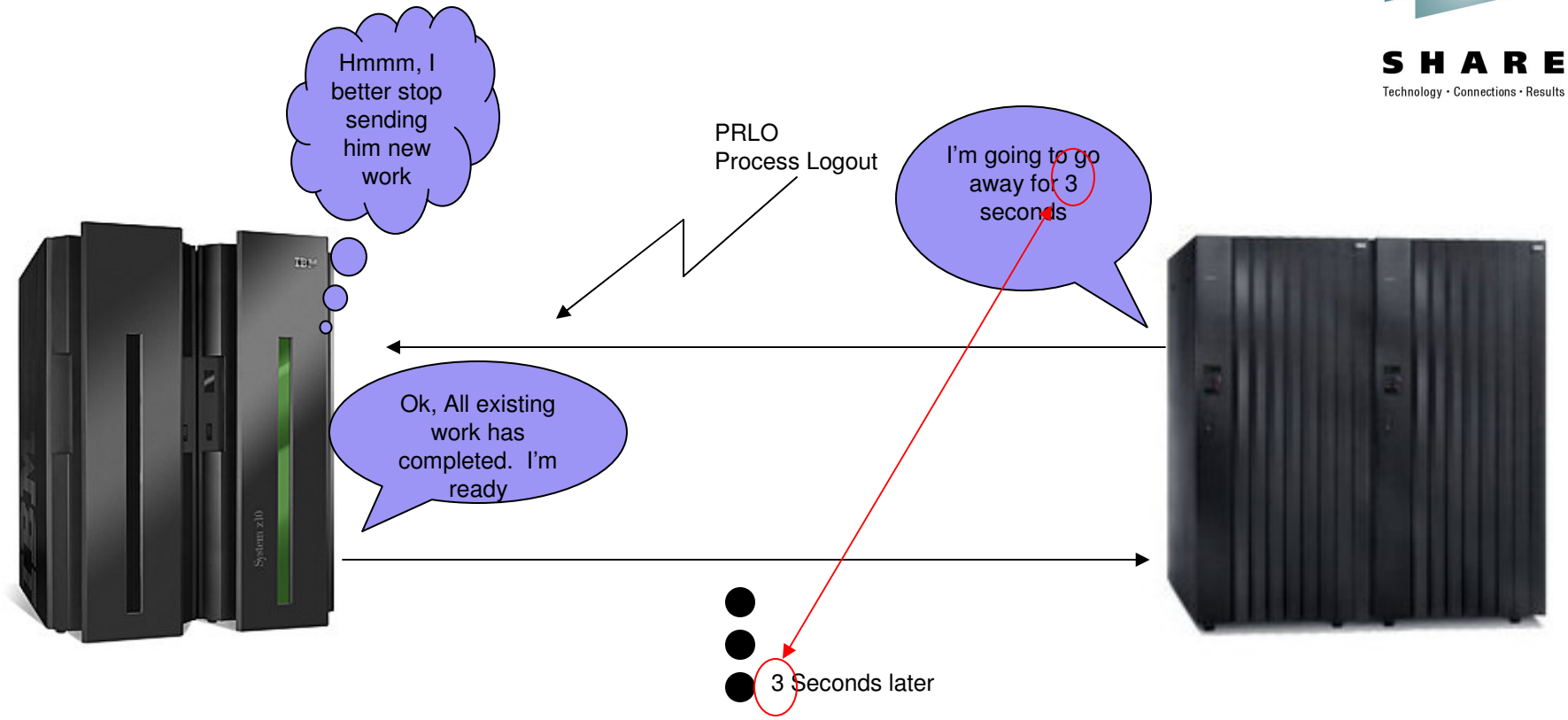
Temporary Logout

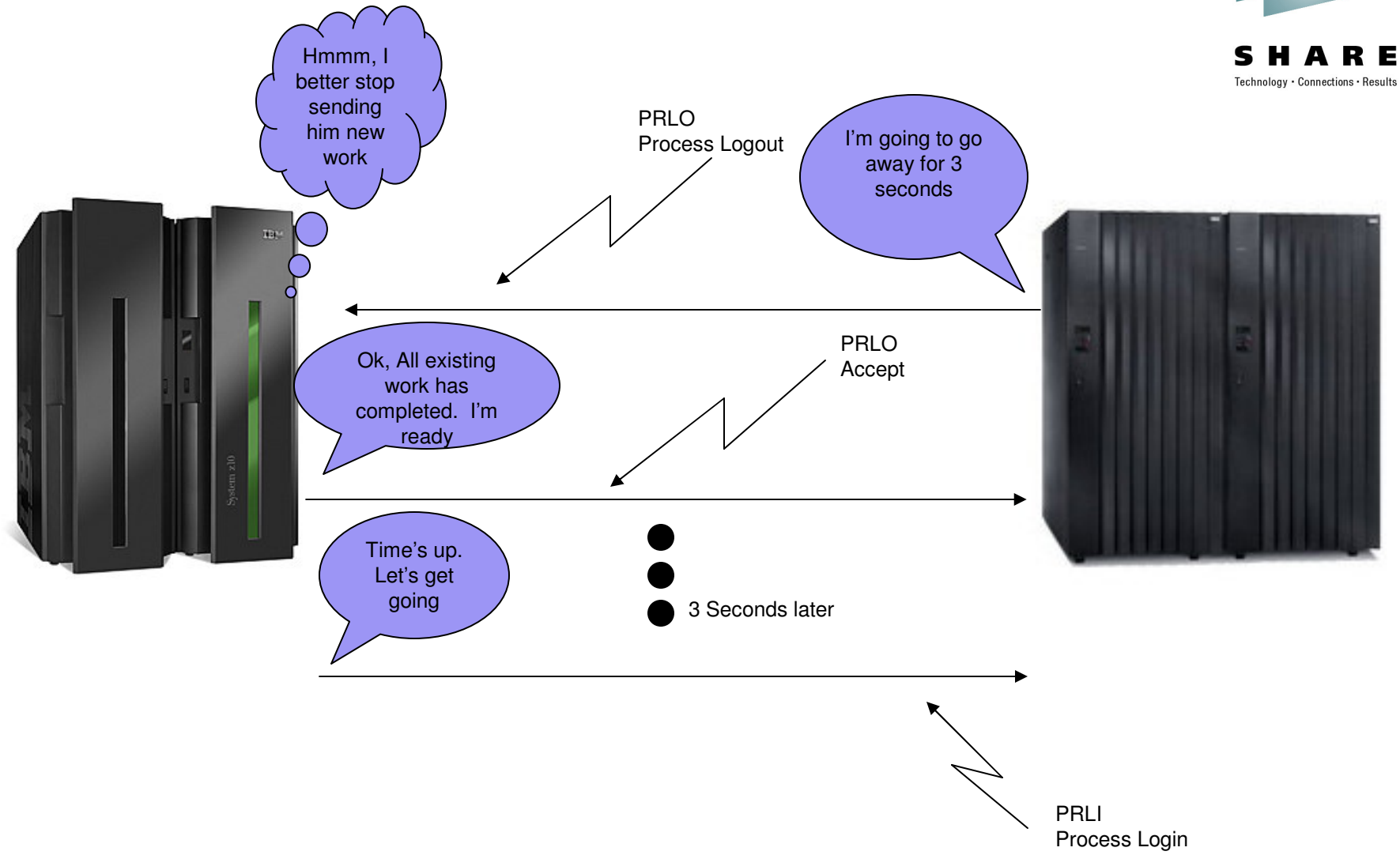
- CU firmware updates can be “cleaner” with zHPF support
- zHPF introduces a “temporary logout” concept
  - CU tells channel that it is ‘going away’

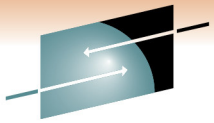

PRLO
Process Logout
I'm going to go away for 3 seconds


Hmmm, I better stop sending him new work
PRLO
Process Logout
I'm going to go away for 3 seconds


Hmmm, I better stop sending him new work
PRLO
Process Logout
I'm going to go away for 3 seconds
Ok, All existing work has completed. I'm ready


Hmmm, I better stop sending him new work
PRLO
Process Logout
I'm going to go away for 3 seconds
Ok, All existing work has completed. I'm ready
3 Seconds later


Hmmm, I better stop sending him new work
PRLO
Process Logout
I'm going to go away for 3 seconds
Ok, All existing work has completed. I'm ready
PRLO
Accept
Time's up. Let's get going
3 Seconds later
PRLI
Process Login



- Any Additional Questions ?

Thank You For Your Time And Attention

Feel free to e-mail me with any zHPF or Ficon questions