Understanding FICON Performance

David Lytle
Brocade Communications

August 4, 2010

Session Number 7345

Legal Disclaimer

- All or some of the products detailed in this presentation may still be under development and certain specifications, including but not limited to, release dates, prices, and product features, may change. The products may not function as intended and a production version of the products may never be released. Even if a production version is released, it may be materially different from the pre-release version discussed in this presentation.
- NOTHING IN THIS PRESENTATION SHALL BE DEEMED TO CREATE A WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, STATUTORY OR OTHERWISE, INCLUDING BUT NOT LIMITED TO, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NONINFRINGEMENT OF THIRD-PARTY RIGHTS WITH RESPECT TO ANY PRODUCTS AND SERVICES REFERENCED HEREIN.
- Brocade and Fabric OS are registered trademarks and the Brocade B-wing symbol, DCX, DCX-4S and SAN Health are trademarks of Brocade Communications Systems, Inc. or its subsidiaries, in the United States and/or in other countries. All other brands, products, or service names are or may be trademarks or service marks of, and are used to identify, products or services of their respective owners.

Brocade FICON Professional Certification

Brocade FICON Certification

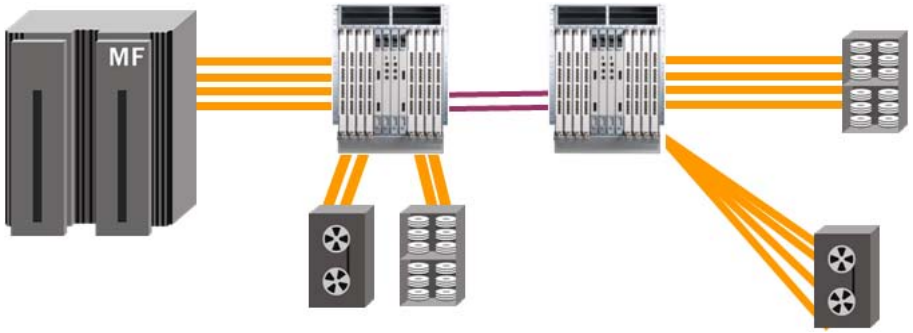
Brocade Certified Architect for FICON

- Certification for Brocade Mainframe-centric Customers – updated as of 3Q10
- For people who do or will work in FICON environments
- There is both an Brocade instructor lead 3-day class, Brocade Design and Implementation for a FICON Environment (CAF200), or a 2-day field class (FCAF200) to assist you in obtaining the knowledge to pass the certification examination
- Certification tests your ability to understand IBM System z I/O concepts, and demonstrate knowledge of Brocade FICON Director and switching fabric components.
- After the class you should be able to design, install, configure, maintain, manage, and troubleshoot common issues for Brocade FICON fabrics for local, metro area and global distance environments.
- Check the following website for complete information:
  - http://www.brocade.com/education/certification-accreditation/certified-architect-ficon/index.page

System z FICON

Some FICON Performance Considerations

Considerations for End-to-End Performance

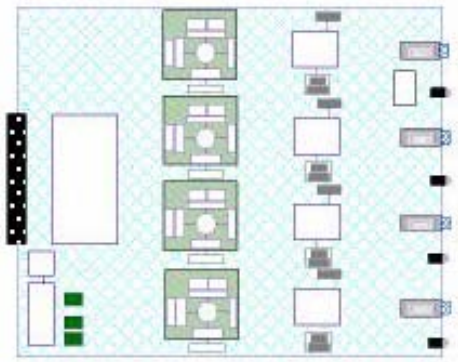


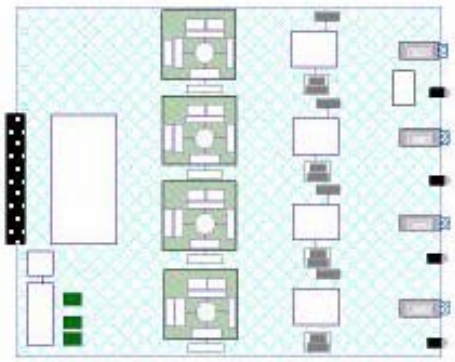
A Few Good Reasons Why Customers Should Always Deploy Switched-FICON

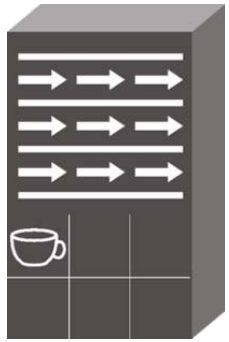
You Should Deploy Switched-FICON
Just One of the many good reasons

- System z10 and beyond will limit the number of buffer credits that are on a CHPID such that a CHPID with long wave optics and single mode cable can reach <=10km
- **Each 8G CHPID provides a maximum of 40 buffer credits**
  - FICON typically creates 1/2 to 2/3 of full size frames significantly limiting distance connectivity compared to 2G or 4G of the past
  - A port on a FICON switching device can provide 100s of buffer credits to maintain high utilization of a link across long distances
  - Tools can be utilized to determine how many buffer credits are required to be allocated to a port to keep a link busy over distance

Mainframe Channel Cards







FICON Express4

- z10, z9
- 200 Buffer Credits per port
  - 101km @ 4G full frame/port
  - 51km @ 4G half frame/port
  - 40km @ 4G 819 byte payloads

FICON Express4 provides the last native 1Gbps CHPID support

FICON Express8

- z10
- 40 Buffer Credits per port
  - 10km @ 8G full frame/port
  - 5km @ 8G half frame/port
  - 4km @ 8G 819 byte payloads

CHPIDs no longer designed to reach beyond 10km

FICON Director

- zSeries, System z
- From 8 to 1,300 Buffer Credits/port

FICON switching devices will provide BCs for long distances

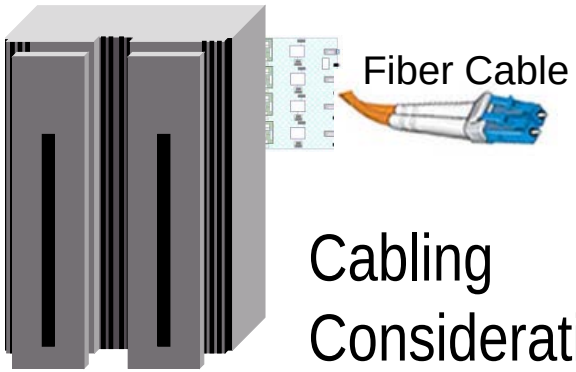
You Should Deploy Switched-FICON
Just One of the many good reasons

- Long wave, single mode connections are often deployed on CHPIDs out of the mainframe
  - Long wave (LX) optics are basically the same price as short wave (SX) optics

- Short wave, multi-mode connections are often deployed on storage ports within a data center
  - Long wave (LX) optics cost more than short wave (SX) optics on both FICON switching devices and storage devices

- But at 8Gbps and beyond the older multi-mode cables (50m) are just not up to carrying data very far
  - OM2 50m (orange cables)
  - OM3 50m (aqua cables)
  - OM4 50m (new cable type)

Multi-mode cable distance limitations



Cabling Considerations

- Long wave single mode (SM) still works well
  - 1/2/4/8/10 Gbps out to 10km with SM
- *Short wave multi-mode might be limiting!*
- 4G optics auto-negotiate back to 1, 2G
- 8G optics auto-negotiate back to 2, 4G
  - 1G storage connectivity requires 2/4G SFPs

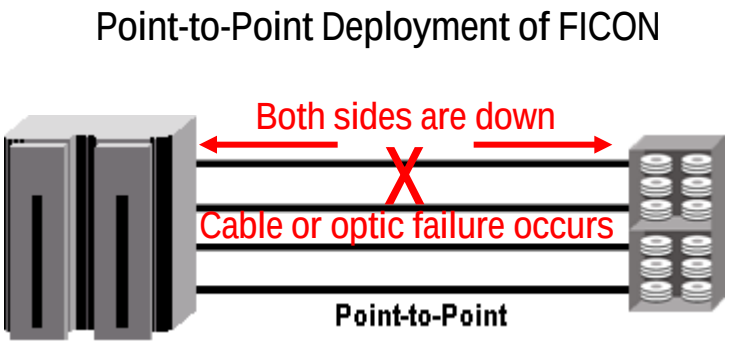
Distance with Multi-Mode Cables (meters)

Protocol (FC)	Encoding	Line Rate (Gb/sec)	OM1-62.5m (200mHz) Multi-Mode	OM2-50m (500mHz) Multi-Mode	OM3-50m (2000mHz) Multi-Mode	OM4-50m (4700mHz) Multi-Mode
1G	8b10b	1.0625	300	500	860	~
2G	8b10b	2.125	150	300	500	~
4G	8b10b	4.25	70	150	380	400
8G	**8b10b**	**8.5**	**21**	**50**	**150**	**200**
10G	64b66b	10.53	33	82	~300	~300
16G	64b66b	14.025	10.5	25	100	130

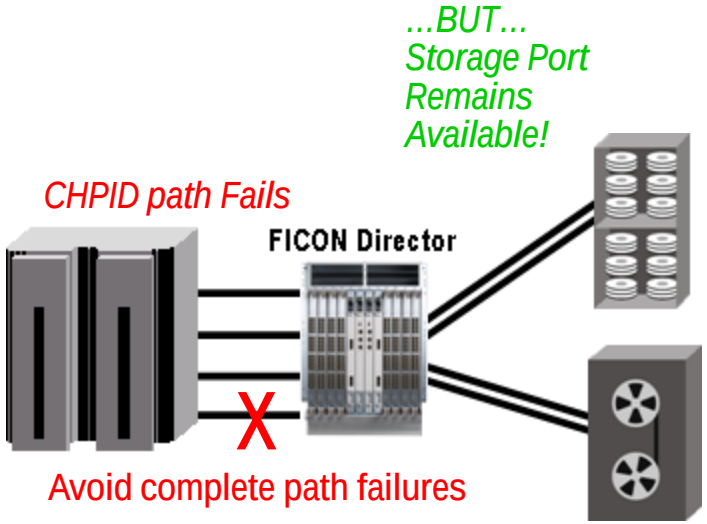
You Should Deploy Switched-FICON
Just One of the many good reasons

- Point-to-point, CHPID-to-storage port, wastes both types of connections and is the least reliable way of providing I/O connectivity for mainframe environments
  - Must deploy a CHPID for every storage port that is enabled
  - A single optic or cable failure renders two ports disabled
  - Regardless of I/O workload each DASD port requires a CHPID

Availability After A Component Failure



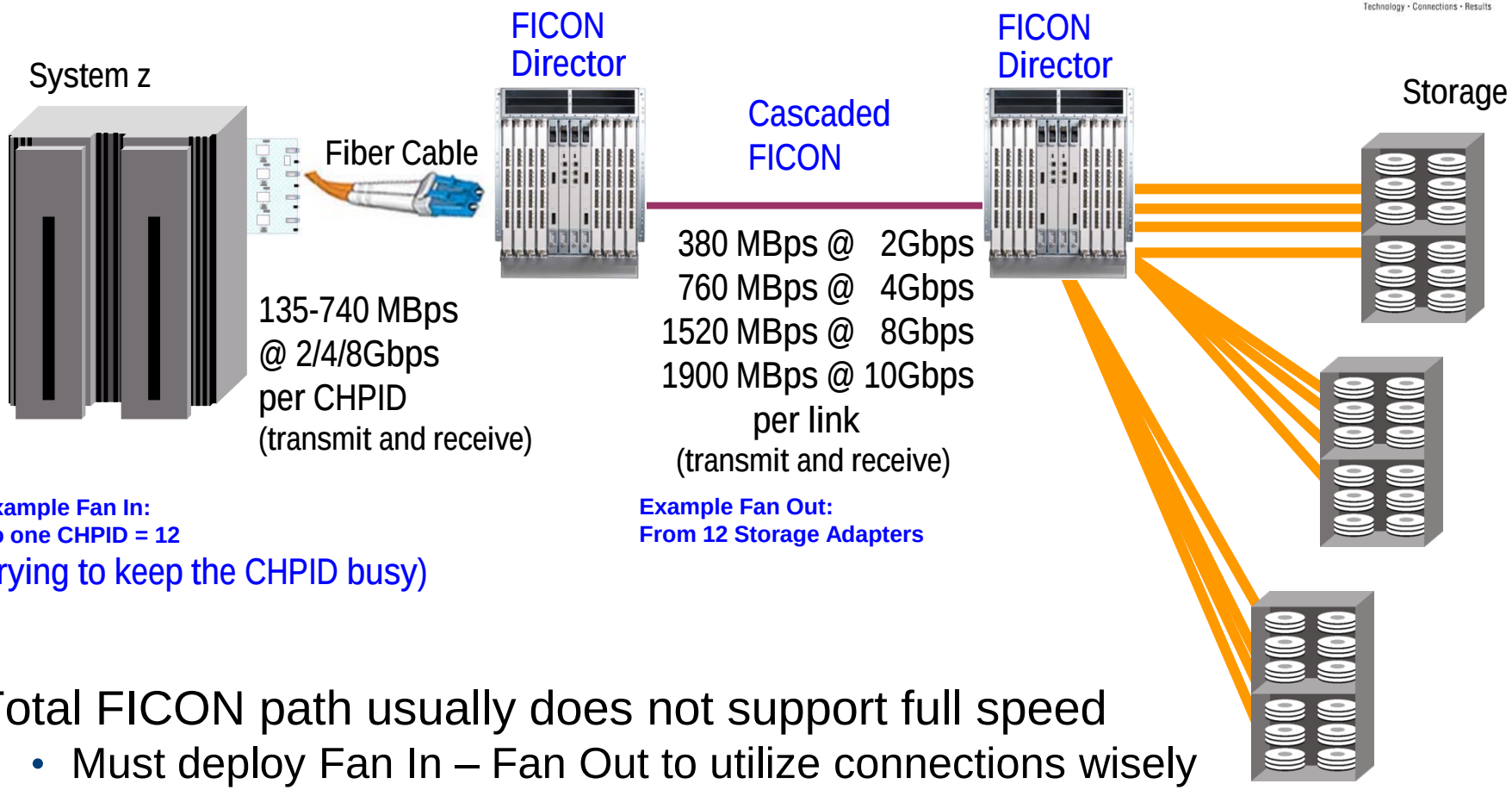
- A failure of a FE8 card …*or…* FE8 channel port …or… failure of the P-2-P cable …or… failure of the storage port optic …or… storage adapter causes:
  - FE8 port to become unavailable AND
  - Storage port becomes unavailable for everyone!
- A failure ***anywhere*** affects both the mainframe connection and the storage connection
  - Lose an SFP and lose host + storage connect
  - The WORST possible reliability and availability is provided by a P-2-P topology!
  - FC optics are most likely element to fail in the channel path – same as in a SAN
  - FC cables are second most likely failure in a fabric
- In a switched-FICON environment, only a segment is rendered unavailable:
  - The non-failing side remains available
  - If the storage has not failed, its port is still available to be used by other CHIPDs



You Should Deploy Switched-FICON
Just One of the many good reasons

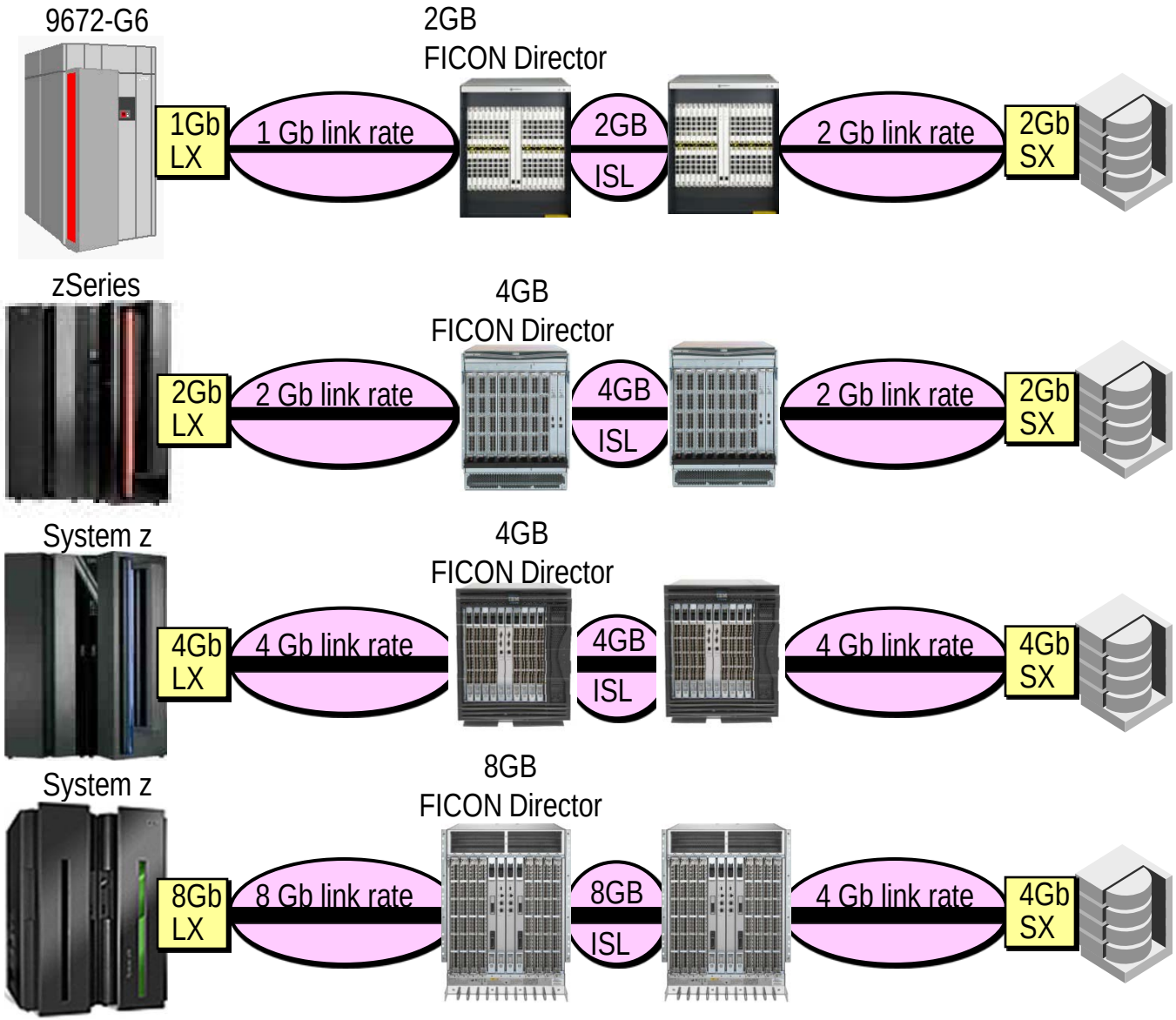
- Fan In-Fan Out was not available with switched-ESCON
  - It was a half-duplex architecture and could not take advantage of Fan In-Fan Out
  - There was never enough bandwidth available in ESCON to allow Fan In-Fan Out to provide any efficiencies anyway
- Fan In-Fan Out is the best way to maximize port utilization
  - Plenty of bandwidth is now available and it is full duplex
  - Can keep CHPIDs and storage ports at optimal utilization
- Fan In-Fan Out can only be deployed when using switched-FICON

FI-FO Overcomes System Bottlenecks



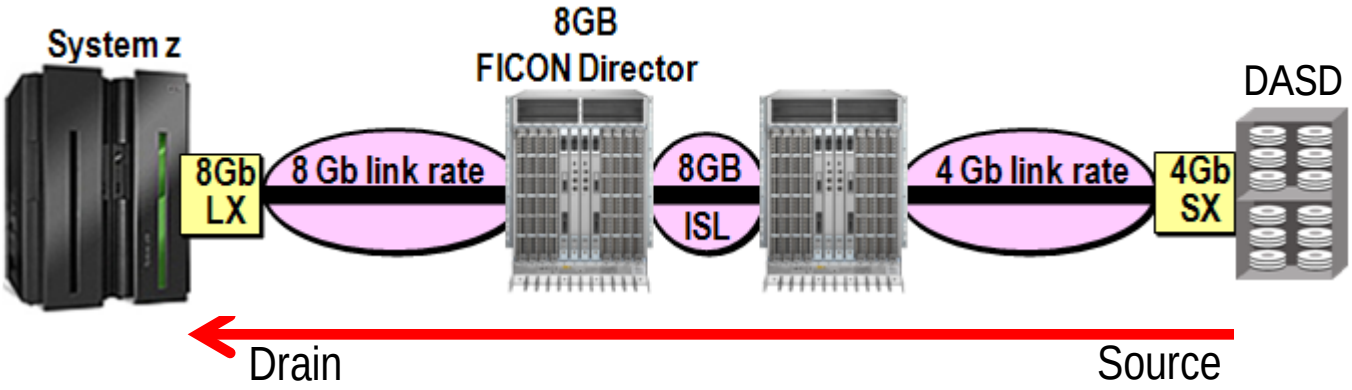
- Total FICON path usually does not support full speed
  - Must deploy Fan In – Fan Out to utilize connections wisely
    - Multiple I/O flows funneled over a single channel path

Maximum End-to-End Link Rates



- Best link rate performance is achieved when the channel, switch and control unit all operate at the same link rate
- Link rate does not guarantee that data will flow at that speed
- Take the speed of the local …AND… cascaded links into consideration.
- Cascaded Links will flow at their rated speeds even when connected with ports of lower speed

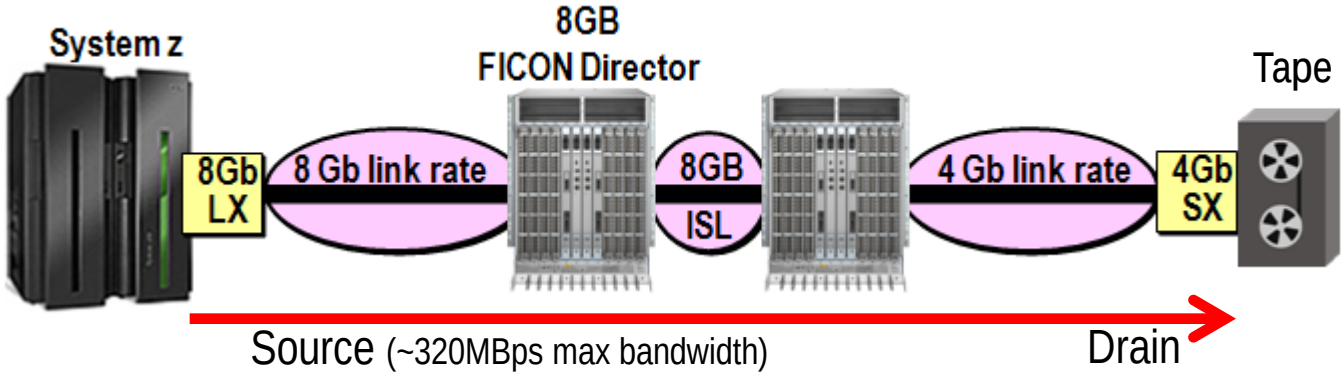
Maximum End-to-End Link Rates



- Is the above a good configuration for data flow?
- If you deployed this configuration, is there a probability of performance problems and/or slow draining devices or not?

- This is actually the ideal model!
- Most application profiles are 90% read, 10% write. So, in this case the "drain" of the pipe are the 8Gb CHPIDs and the "source" of the pipe are 4Gb storage ports.
- This represents an end-to-end network that will generally require the least amount of buffer credit pacing (assuming you implemented the correct number of ISLs)

Maximum End-to-End Link Rates



- Is the above a good configuration for data flow?
- If you deployed this configuration, is there a probability of performance problems and/or slow draining devices or not?

- For 4G tape this is OK – Tape is about 90% write and 10% read on average
- The maximum bandwidth a tape can accept and compress is about 240MBps for SUN and about 320MBps for IBM (at 2:1 compression)
- An 8G CHPID in Command Mode can do about 510MBps
- A 4G Tape channel can carry about 380MBps (400 * .95) = 380
- So a single CHPID attached to a 4G tape interface:
  - Can run a single IBM tape drive (510 / 320 = 1.594)
  - Can run two Oracle (STK) tape drives (510 / 240 = 2.125)

Directors versus Switches for FICON

- Directors are strategic devices, that solve most connectivity issues, and provide very long term investment protection and value over time in an ever changing enterprise environment
  - Excellent 5-7+ year investment
  - Non-disruptive failure repair
  - Non-disruptive and extensive scalability due to port density
  - Non-disruptive change to new and faster technology
- Switches, on the other hand, are tactical devices that solve a specific connectivity issue and provide poor investment value over time especially since technology is ever changing at a rate of about 18-24 months – best use is probably for standalone tape drives
  - Good 2+ year investment (how long have your switches lasted?)
  - Disruptive failure repair (except fans, power supplies and SFPs)
  - Limited scalability due to port density
  - Disruptive change to new and faster technology

My Story Today is:

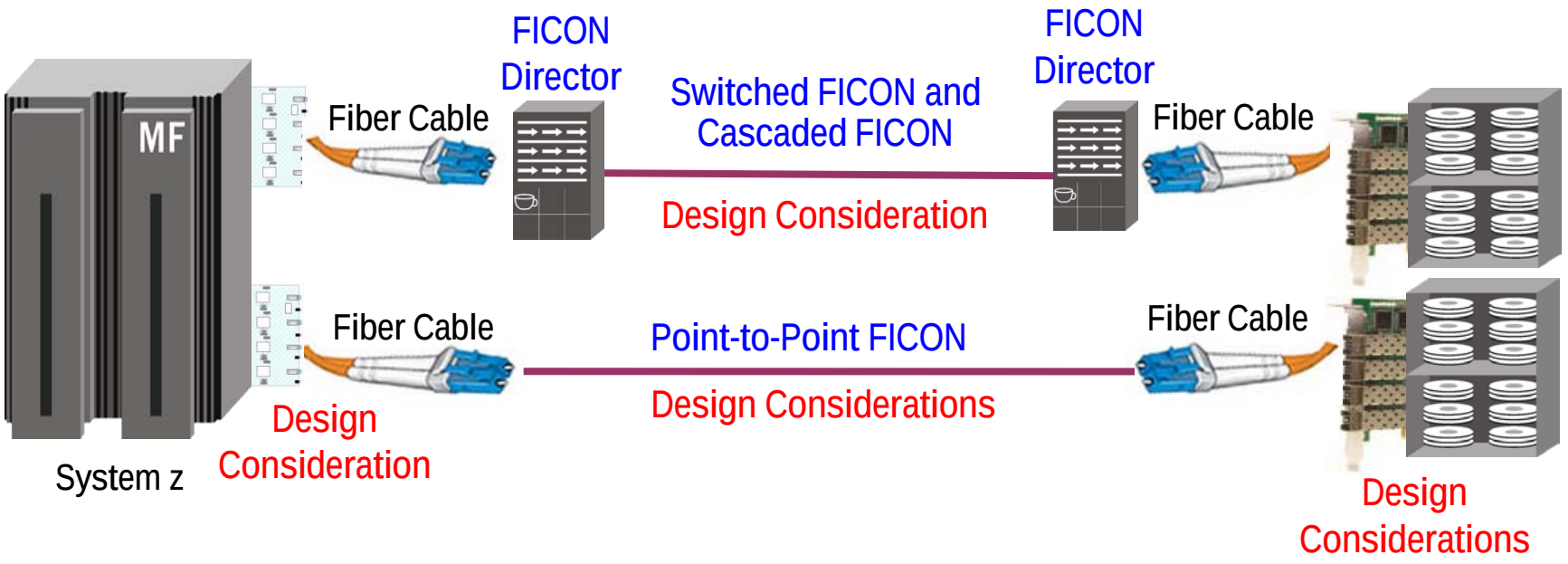
- IT IS ALL GOOD!

- BUT…

- You Have To Know *What It Really Means* In Order To Deploy FICON Technology Successfully!

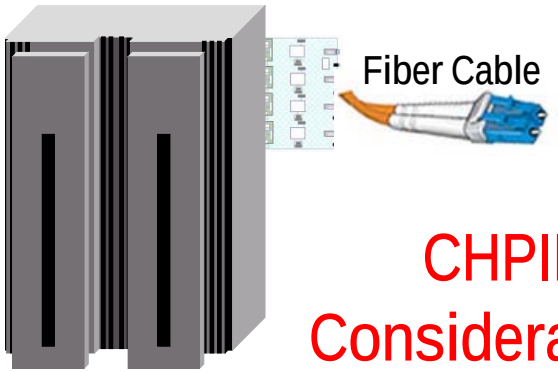
- Today's presentation is an **OVERVIEW** – I have a six hour presentation that covers these considerations in much more detail!

End-to-End FICON/FCP Connectivity

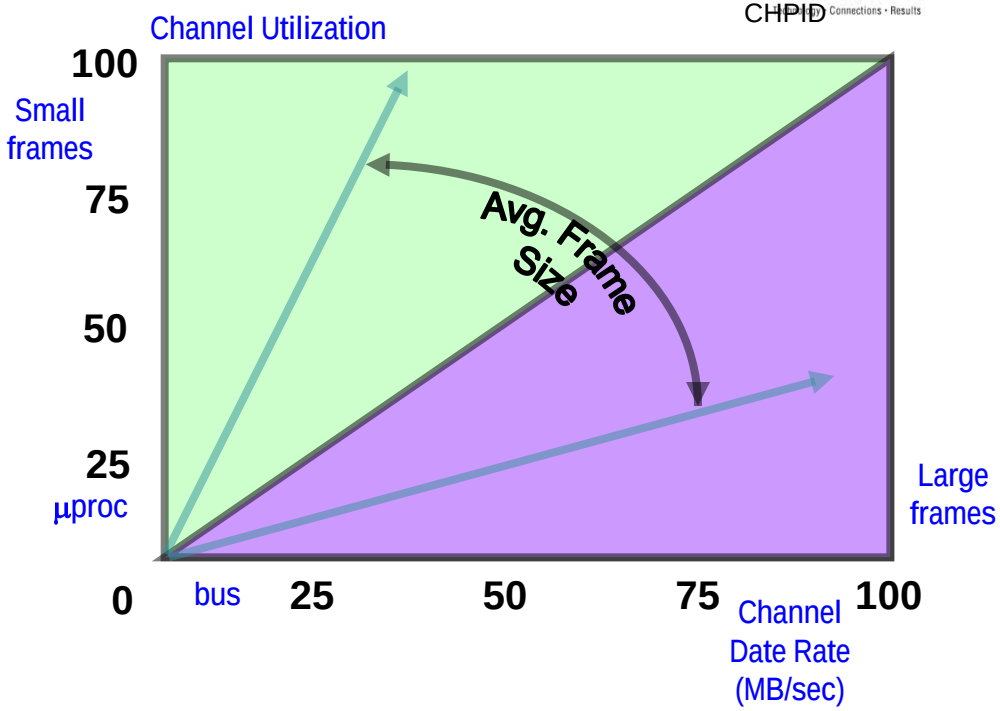


- There are a series of Design Considerations that you must understand in order to successfully meet your expectations with your FICON fabrics

End-to-End FICON/FCP Connectivity

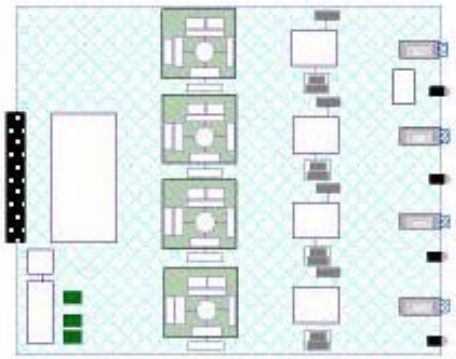


CHPID Considerations



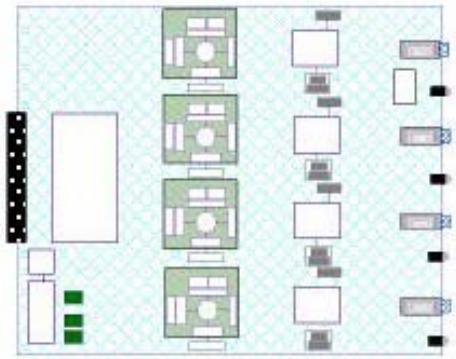
- Channel Microprocessors and PCI Bus
- Average frame size for FICON
- CHPID – LX or SX and cabling
- Buffer Credit considerations

Mainframe Channel Cards



FICON Express4

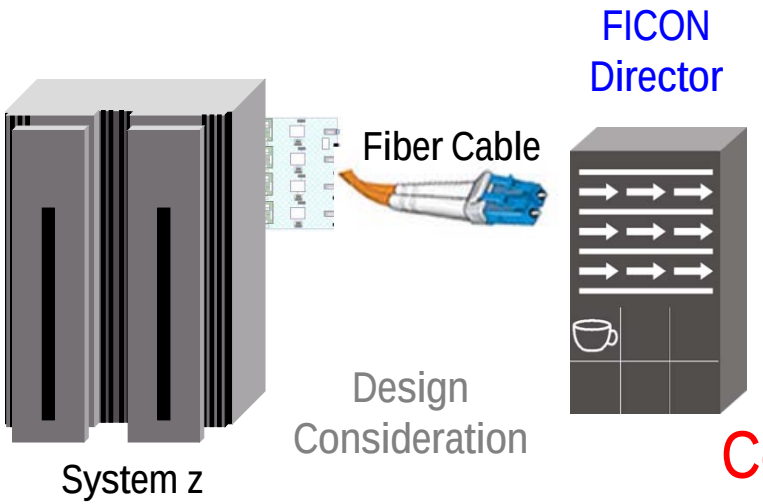
- z10, z9
- 1, 2 or 4 GBps link rate
- Cannot Perform at 4Gbps!
- Standard FICON Mode:
  <= 350MBps FD (out of 800 MBps)
- zHPF FICON Mode:
  <= 520MBps FD (out of 800 MBps)
- 200 Buffer Credits per port
  - Out to 50km with 1K frames



FICON Express8

- z10
- 2, 4 or 8 GBps link rate
- Cannot Perform at 8Gbps!
- Standard FICON Mode:
  <= 510 MBps FD (out of 1600 MBps)
- zHPF FICON Mode:
  <=740 MBps FD (out of 1600 MBps)
- 40 Buffer Credits per port
  - Out to 5km with 1K frames

FICON/FCP Switching Devices



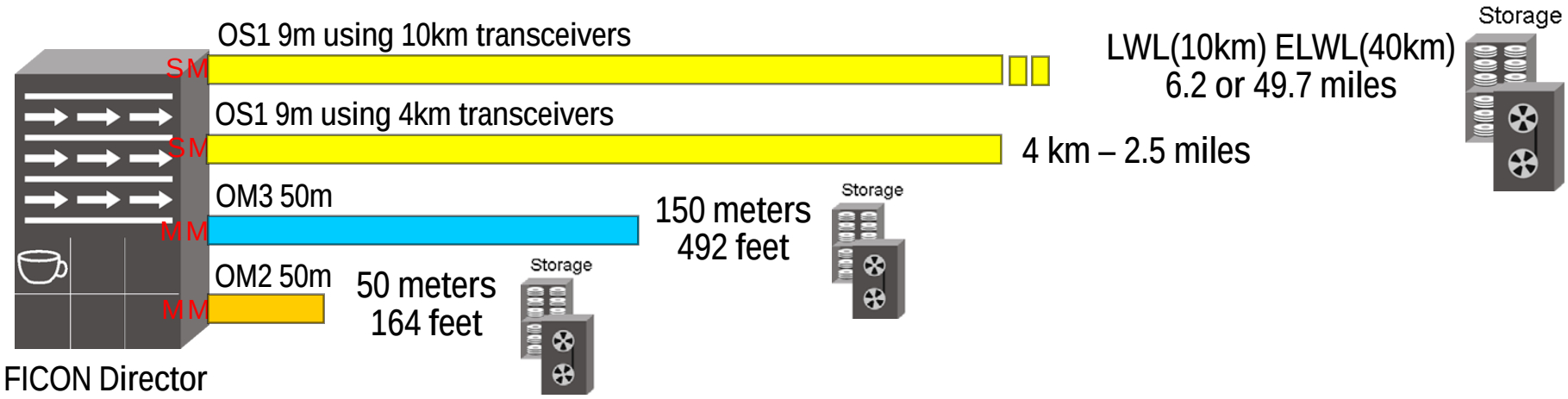
Design Consideration

FICON Director Chassis Considerations

Some of the capabilities you should be looking for in a FICON switching device to provide you with a switched-FICON fabric deployment

- Single Mode and/or Multimode cabling
- SFP Considerations
- Buffer Credits
- FICON, FCP intermix and Virtualization capabilities
- FICON connectivity and scalability alternatives
- Ease of management

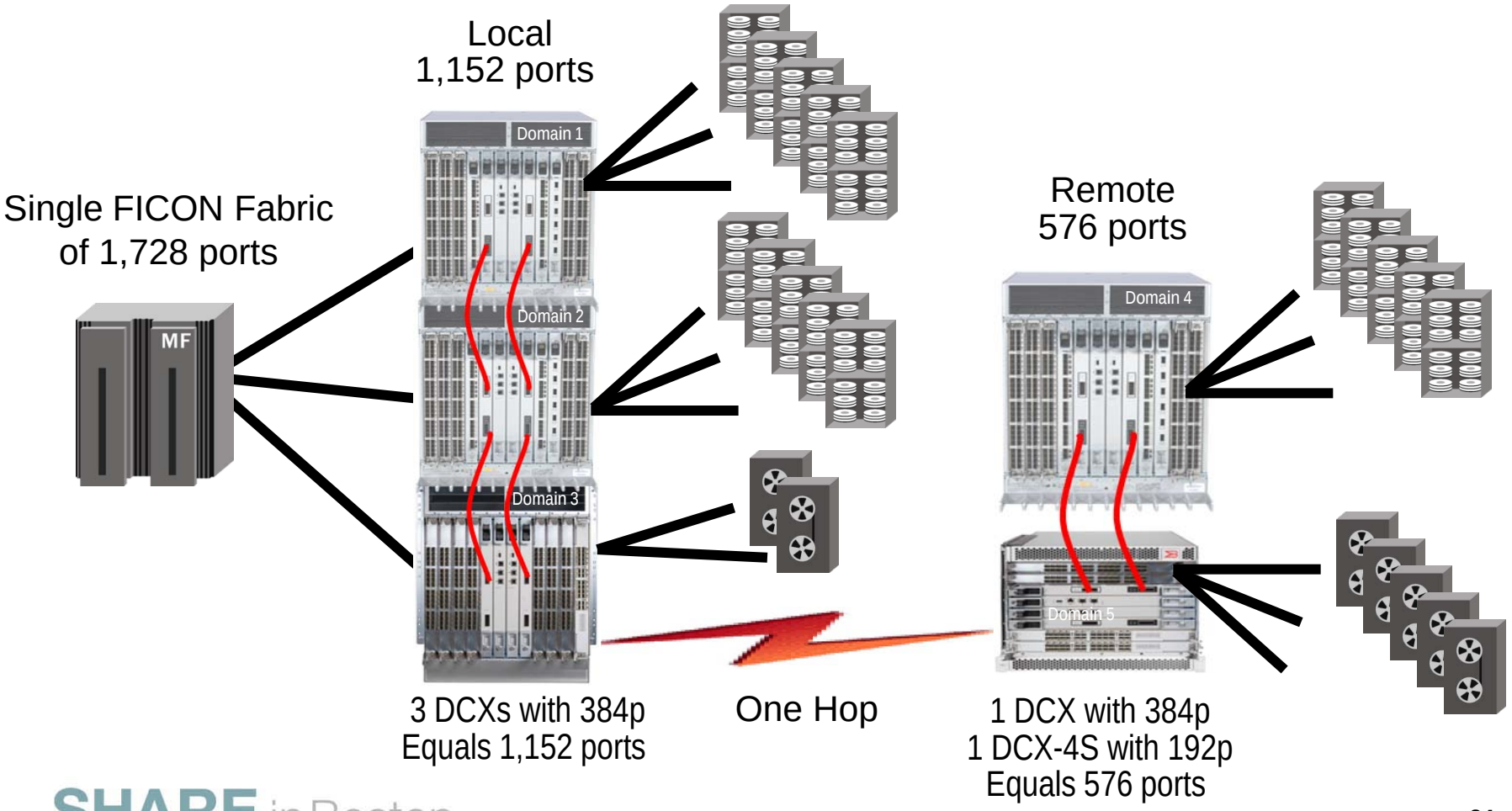
Multi-mode Cable distances at 8Gbps



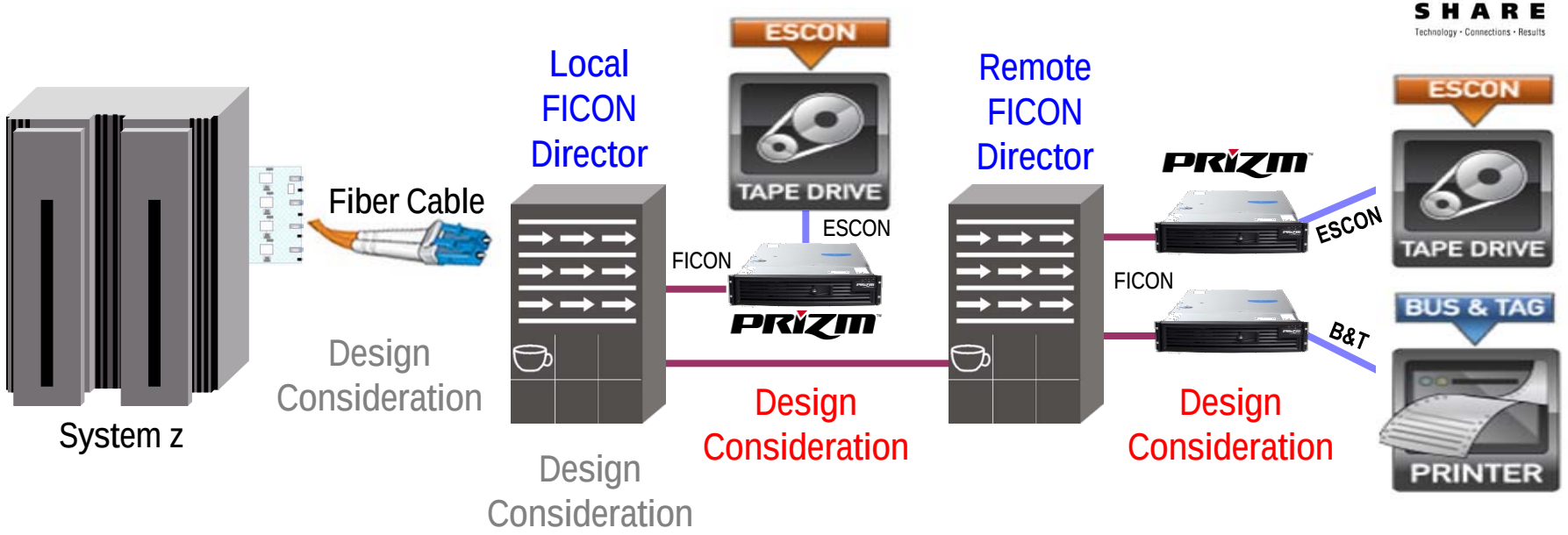
- Majority of local storage connections are still short wave
- What is your current cable distance to farthest storage attachment point?
- 4Gbps allowed it to be 1,246 feet (380m) away using OM3
  - Will you have to relocate some storage closer in at 8G?
  - Do you need to re-cable from OM2 to OM3 or up to OS1?

Inter-Chassis Links
Scalability Through Innovative Director Features

- Can have cascaded connectivity as large as 384p x 6 DCX = 2,304 ports



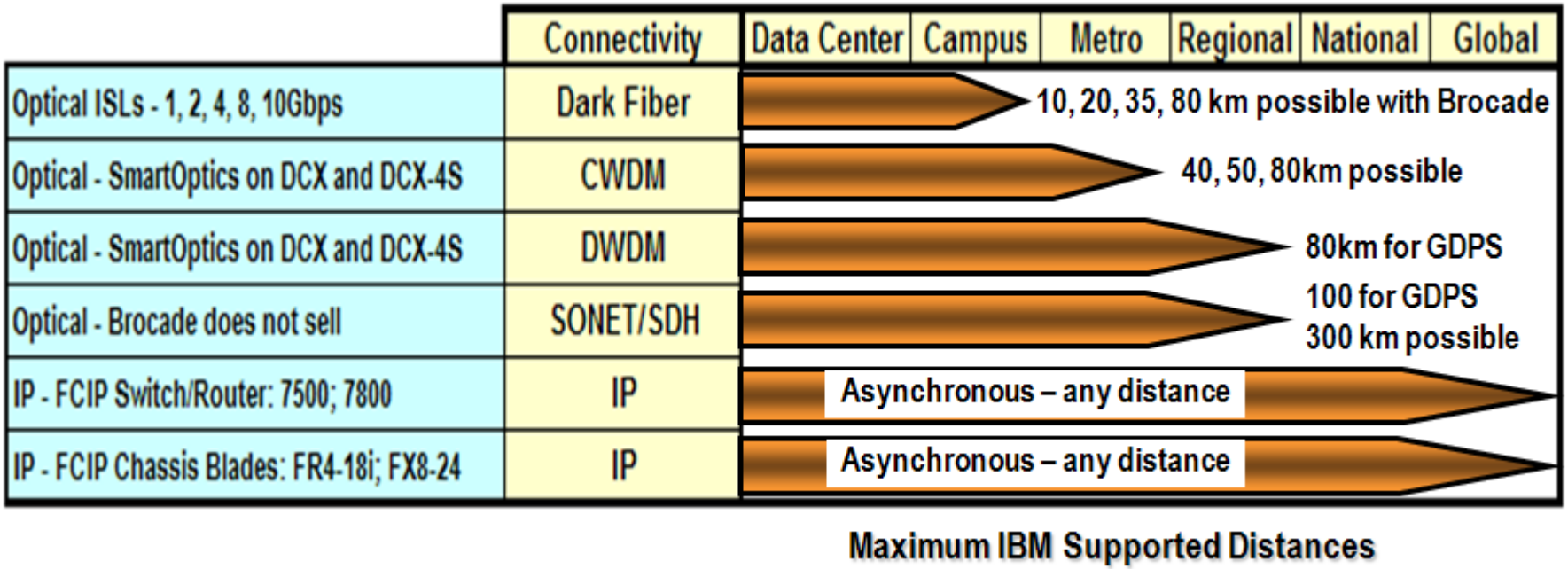
I/O Evolution and Migration



Some of you need to start converting old ESCON and even old Bus and Tag equipment to function off of FICON fabrics

Can use the Optica Technologies FICON-to-ESCON adapter box (Prizm) to address removing old ESCON Directors (9032s) while allowing ESCON legacy devices to continue to function – also converts Bus and Tag to ESCON

FICON Long Distance Connectivity



Four basic things are required for long distance connectivity:

1 – Optics or IP that will push the data where you want it to be

2 – Cabling that allows data to be pushed where you want it to be

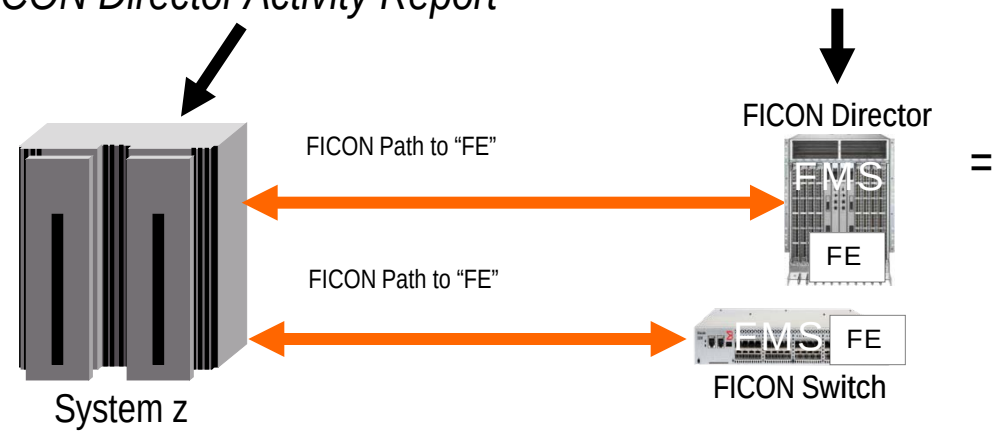
3 – Optimal buffer credits if it is an optical pathway over distance

4 – Certification from the vendor that it is supported

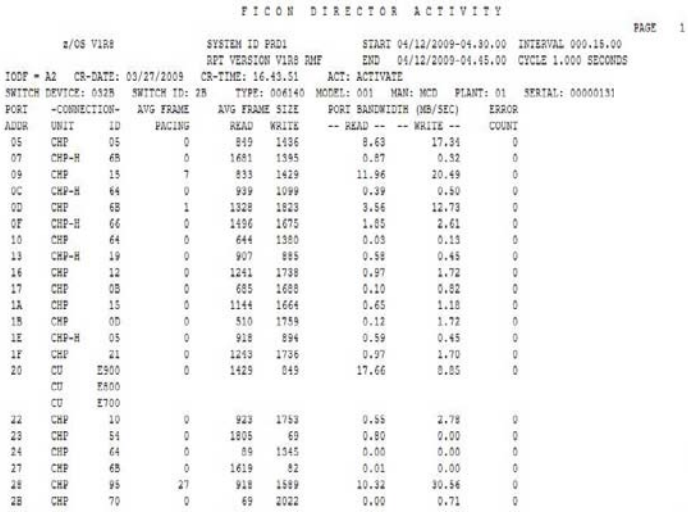
Control Unit Port for Use By RMF

Option FCD in ERBRMFxx parmlib member and STATS=YES in IECIOSnn tells RMF to produce the 74-7 *records which it uses for the FICON Director Activity Report*

FICON Management Server (FMS) license per device enables CUP to provide information to the FICON Director Activity Rpt.



FICON Director Activity Report per FICON Domain ID per Interval



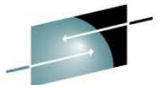
```
                                F I C O N   D I R E C T O R   A C T I V I T Y                                  PAGE    1

            z/OS V1R8                    SYSTEM ID PRD1          START 04/12/2009-04.30.00  INTERVAL 000.15.00
                                         RPT VERSION V1R8 RMF        END   04/12/2009-04.45.00  CYCLE 1.000 SECONDS
IODF = A2   CR-DATE: 03/27/2009   CR-TIME: 16.43.51     ACT: ACTIVATE
SWITCH DEVICE: 052B   SWITCH ID: 2B     TYPE: 006140  MODEL: 001   MAN: MCD   PLANT: 01   SERIAL: 00000131
PORT    -CONNECTION-   AVG FRAME    AVG FRAME SIZE     PORT BANDWIDTH (MB/SEC)      ERROR
ADDR    UNIT      ID     PACING      READ   WRITE     -- READ --  -- WRITE --      COUNT
 05     CHP       05          0       849    1436          8.63        17.34          0
 07     CHP-H     6B          0      1681    1395          0.87         0.32          0
 09     CHP       15          7       833    1429         11.96        20.49          0
 0C     CHP-H     64          0       939    1099          0.39         0.80          0
 0D     CHP       6B          1      1328    1823          3.56        12.78          0
 0F     CHP-H     66          0      1496    1675          1.85         2.61          0
 10     CHP       64          0       644    1380          0.03         0.15          0
 13     CHP-H     19          0       907     885          0.58         0.45          0
 16     CHP       12          0      1241    1738          0.97         1.72          0
 17     CHP       0B          0       685    1688          0.10         0.82          0
 1A     CHP       15          0      1144    1664          0.65         1.18          0
 1B     CHP       0D          0       510    1759          0.12         1.72          0
 1E     CHP-H     05          0       918     894          0.59         0.45          0
 1F     CHP       21          0      1243    1736          0.97         1.70          0
 20     CU     E900           0      1429     849         17.66         8.85          0
        CU     E800
        CU     E700
 22     CHP       10          0       923    1753          0.55         2.78          0
 23     CHP       54          0      1805      69          0.80         0.00          0
 24     CHP       64          0        89    1345          0.00         0.00          0
 27     CHP       6B          0      1619      82          0.01         0.00          0
 28     CHP       95         27       918    1589         10.32        30.56          0
 2B     CHP       70          0        69    2022          0.00         0.71          0
```

- FICON Management Server (FMS) is a license to enable Control Unit Port (CUP) on a FICON switching device
- CUP, when enabled, uses an internal port "FE" to establish communications to RMF running on System z
  - Even within Virtual Fabrics on a chassis, "FE" is always used for CUP

FICON Director Activity Report

FMS Key on, CUP enabled
RMF 74-7 records captured



zHPF Enabled

F I C O N D I R E C T O R A C T I V I T Y

PAGE 1

z/OS V1R8 SYSTEM ID PRD1 START 04/12/2009-04.30.00 INTERVAL 000.15.00
RPT VERSION V1R8 RMF END 04/12/2009-04.45.00 CYCLE 1.000 SECONDS

IODF = A2 CR-DATE: 03/27/2009 CR-TIME: 16.43.51 ACT: ACTIVATE
SWITCH DEVICE: 032B SWITCH ID: 2B TYPE: 006140 MODEL: 001 MAN: MCD PLANT: 01 SERIAL: 00000131

PORT ADDR	-CONNECTION- UNIT	ID	AVG FRAME PACING	AVG FRAME SIZE READ	WRITE	PORT BANDWIDTH (MB/SEC) -- READ --	-- WRITE --	ERROR COUNT
05	CHP	05	0	849	1436	8.63	17.34	0
07	CHP-H	6B	0	1681	1395	0.87	0.32	0
09	CHP	15	7	833	1429	11.96	20.49	0
0C	CHP-H	64	0	939	1099	0.39	0.50	0
0D	CHP	6B	1	1328	1823	3.56	12.73	0
0F	CHP-H	66	0	1496	1675	1.85	2.61	0
10	CHP	64	0	644	1380	0.03	0.13	0
13	CHP-H	19	0	907	885	0.58	0.45	0
16	CHP	12	0	1241	1738	0.97	1.72	0
17	CHP	0B	0	685	1688	0.10	0.82	0
1A	CHP	15	0	1144	1664	0.65	1.18	0
1B	CHP	0D	0	510	1759	0.12	1.72	0
1E	CHP-H	05	0	918	894	0.59	0.45	0
1F	CHP	21	0	1243	1736	0.97	1.70	0
20	CU	E900	0	1429	849	17.66	8.85	0
	CU	E800						
	CU	E700						
22	CHP	10	0	923	1753	0.55	2.78	0
23	CHP	54	0	1805	69	0.80	0.00	0
24	CHP	64	0	89	1345	0.00	0.00	0
27	CHP	6B	0	1619	82	0.01	0.00	0
28	CHP	95	27	918	1589	10.32	30.56	0
2B	CHP	70	0	69	2022	0.00	0.71	0

Average Frame Pacing Delay is the most important indication on this report.

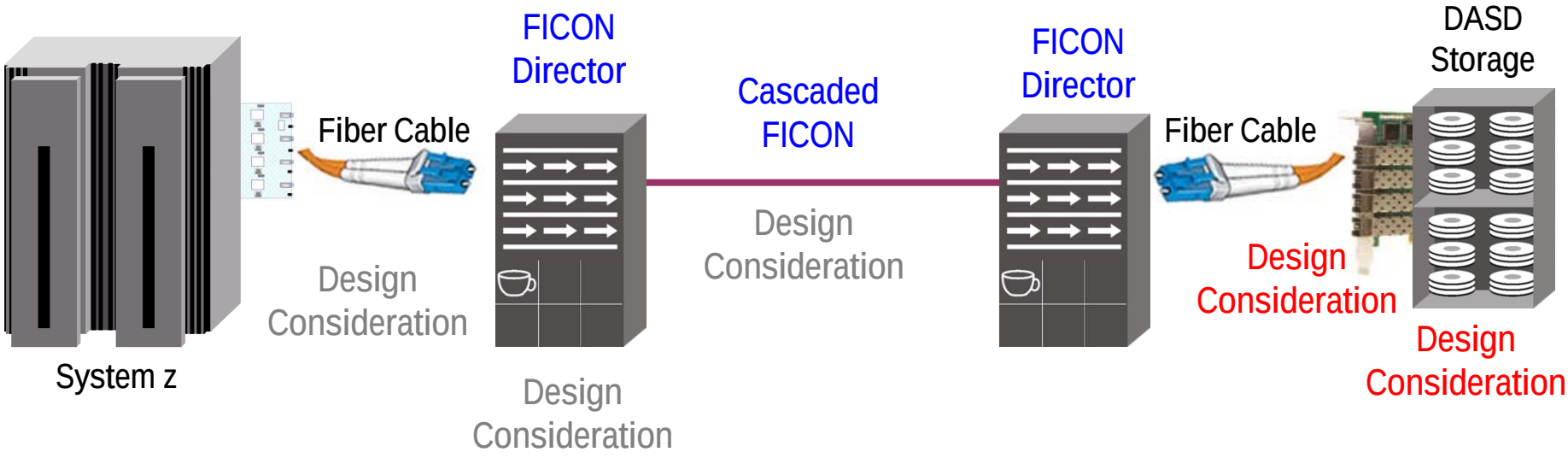
Pacing Delay indicates that during the interval a port ran out of buffer credits while a frame was waiting to be sent.

This is an indication of a performance problem.

Overall Averages: ~1116 ~1508
Note: zHPF Transport Mode results in larger frames

Command Mode FICON will probably have an average FICON frame size of 350-1100 bytes!

Connectivity with storage devices



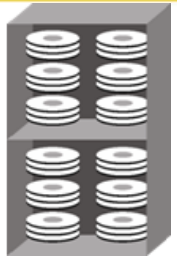
Storage adapters can be throughput constrained

- Must ask storage vendor about performance specifics
- Is zHPF supported/enabled on your DASD control units?

Busy storage arrays can equal reduced performance

- Copy services, RAID used, RPMs, etc.
- Let's look a little closer at this

Connectivity with storage devices



How fast are the Storage Adapters?

- Mostly 2 / 4Gbps today – some 8G – where are the internal bottlenecks

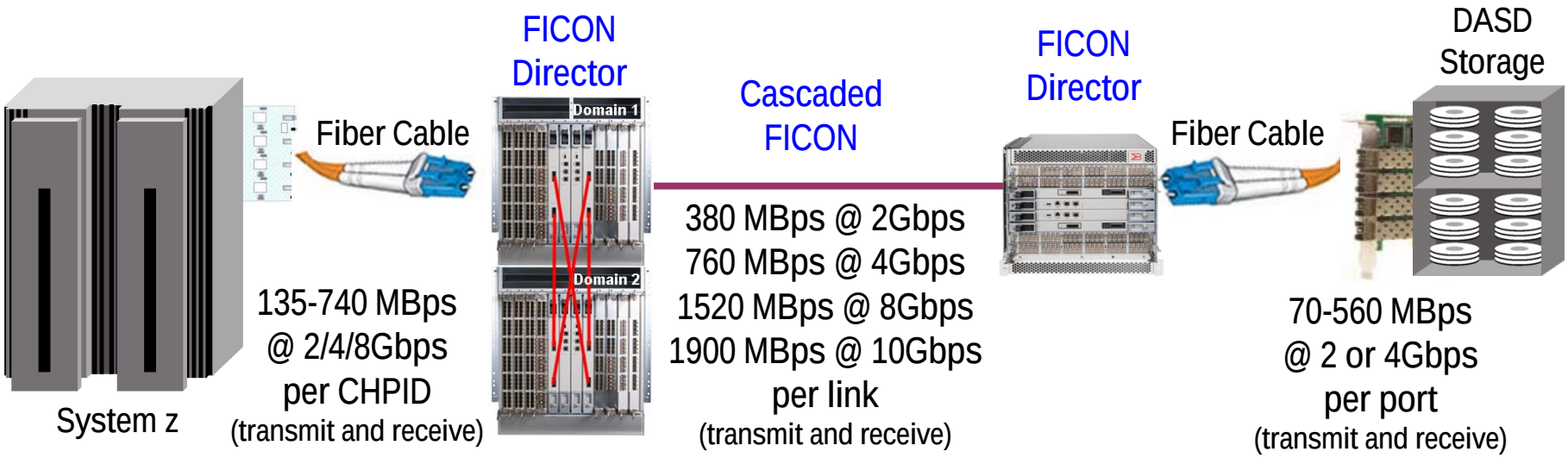
What kinds of internal bottlenecks does a DASD array have?

- 7200rpm, 10,000rpm, 15,000rpm
- What kind of volumes: 3390-3; 3390-54; EAV; XIV
- How many volumes are on a device? HiperPAV in use?
- How many HDDs in a Rank (arms to do the work)
- What Raid scheme is being used (RAID penalties)?
- Copy services? They have a higher priority than host I/O requests
- Etc.

RMF Magic from Intellimagic or Performance Associates I/O Driver

These tools perform mathematical calculations against raw RMF data to determine storage HDD utilization characteristics – use them or something like them to understand I/O metrics!

End-to-End FICON/FCP Connectivity



- In order to fully utilize the capabilities of a FICON fabric a customer needs to deploy a Fan In – Fan Out Architecture
- If you are going to deploy Linux on System z, or private cloud computing, then switched FICON flexibility is required!

FICON should never be direct attached!

The End – Thank You For Attending!

Session #7345

Please fill out your evaluation forms

Please attend the Brocade session:

Customer Deployment Examples for FICON Technologies

Thursday at 8am!