

Introduction to Storage Technologies SAN (**S**torage **A**rea **N**etworking) and

a little FICON (**F**iber **C**ONnection)

David Lytle – Brocade – dlytle@brocade.com

Mike Blair – Cisco – mblair@cisco.com (not here today)

SAN 101

Agenda for Session 10999 and 11000

Session 10999 – 09:30am – 10:30am

- **Types and Components of Storage**
- **Let's talk Fibre Channel**
- **FC Buffer Credits**
- **Fabric Routing / Virtual Fabrics / Partitioning**
- **Security / Zoning**

Session 11000 – 11:00am – 12:15pm

- **History**
- **Terminology, Connectors, Cables, and Wavelengths**
- **Addressing in FICON**
- **ESCON Status, zHPF and NPIV**
- **Buffer Credits, CUP, RMF**

Agenda for Session 10999 and 11000

Session 11000

- **Types and Components of Storage**
- Let's talk Fibre Channel
- FC Buffer Credits
- Fabric Routing / Virtual Fabrics / Partitioning
- Security / Zoning

Storage Network Components

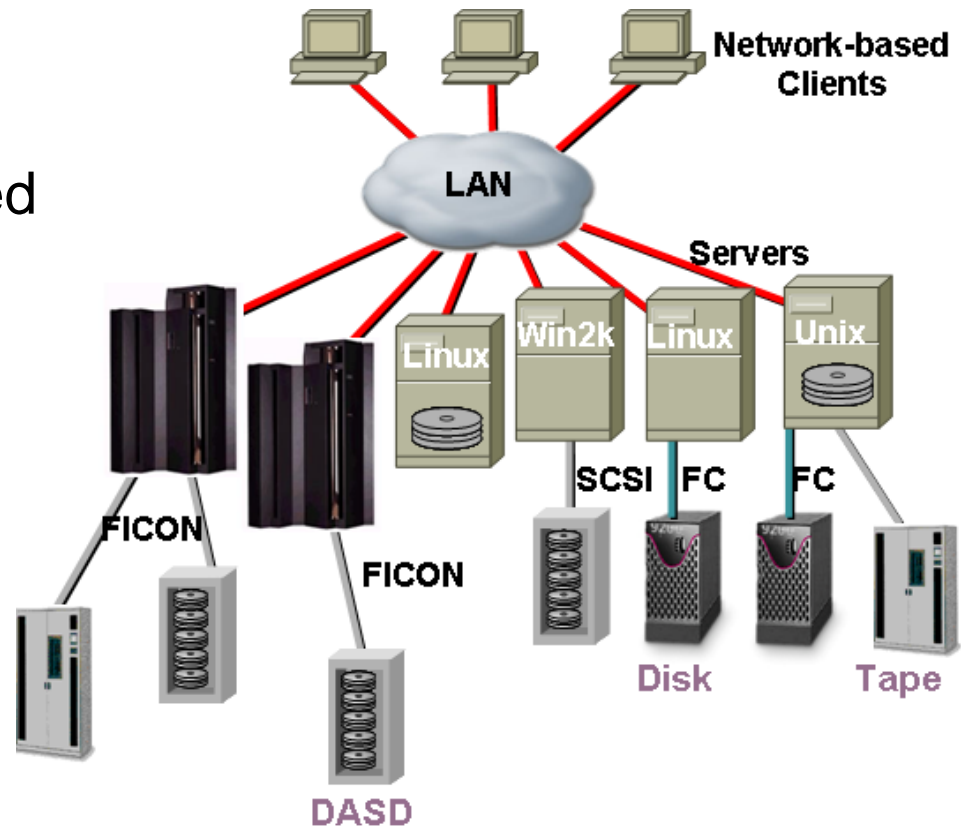
Multiple Infrastructure Vendors (eg. Brocade and Cisco)
Several components required to build a SAN

- Servers with Host Bus Adapters (HBAs)
- Mainframes with FICON Express Channels
- Storage systems
 - RAID (Redundant Array of Independent/Inexpensive Disks)
 - JBOD (Just A bunch of Disks)
 - Tape
 - VTS/VSM (Virtual Tape)
- Fibre Channel / FICON Switches or Directors
- Ethernet Switches (iSCSI / FCoE)
- SAN management software



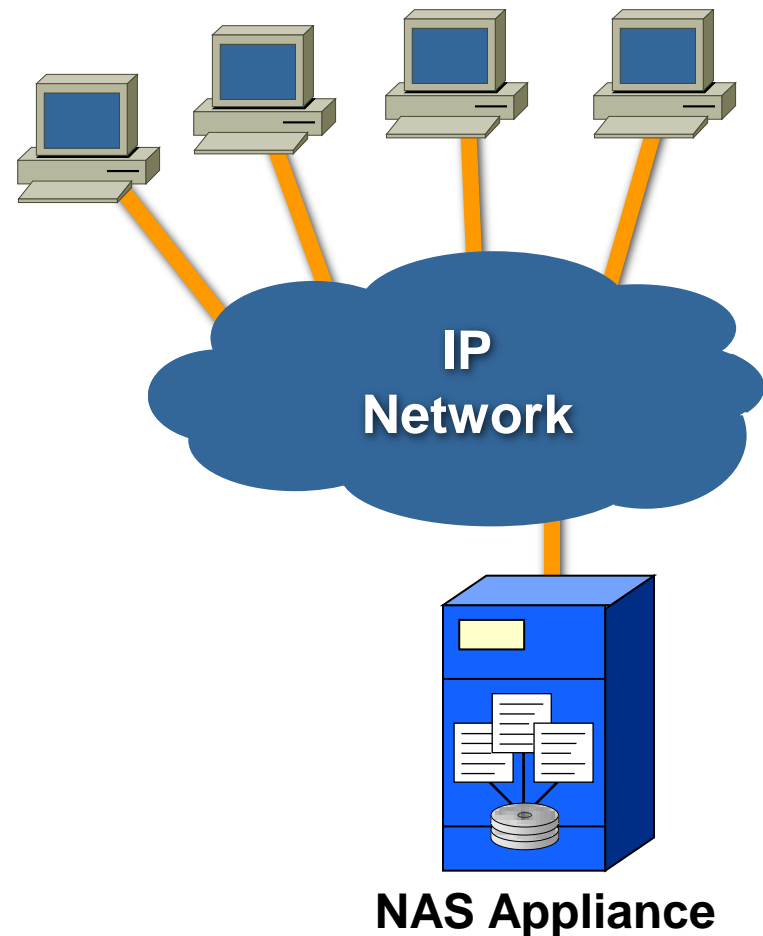
Direct Attached Storage

- Direct Attached Storage (DAS)
- Storage is captive ‘behind’ the server, limited mobility
- Limited scalability due to limited devices
- No storage sharing possible
- Costly to scale
- Management can be complex
- Often cannot take full advantage of the technology



Network Attached Storage (NAS)

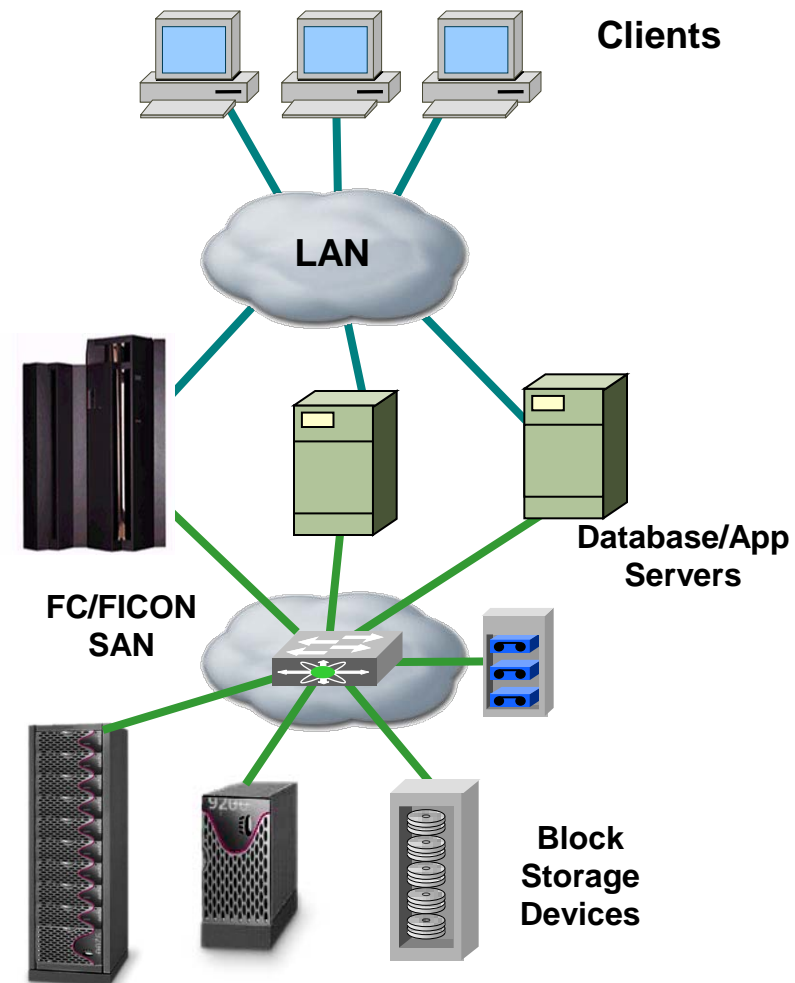
- Dedicated file server
- Optimized for file-based access to shared storage over an IP network
- Suitable for applications involving file serving/sharing
- High-performance access, data protection, and disaster recovery
- Capable of storage partitioning
- Uses network file system protocols such as NFS or CIFS



Storage Area Network (SAN)

Separation of Storage from the Server

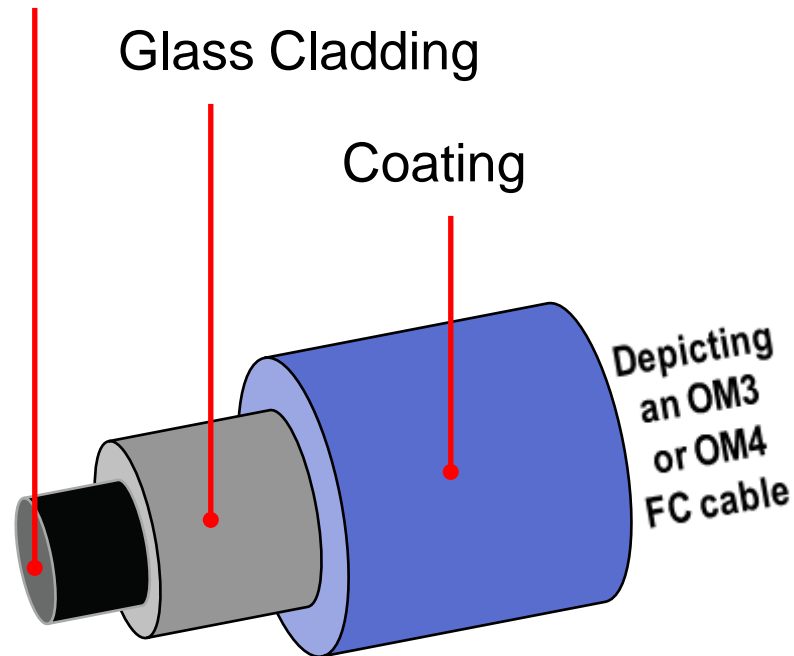
- Storage is accessed Block-level via SCSI/FICON and can be in a switched environment
- High performance interconnect providing high I/O throughput
- Lower TCO relative to direct attached storage, storage can be shared
- Have to consider Vendor Interoperability / Qualifications
- More Complex management due to size/scale



FICON or Storage Area Network (SAN)

SAN Terminology -- *Fiber Channel Link*

Glass Core

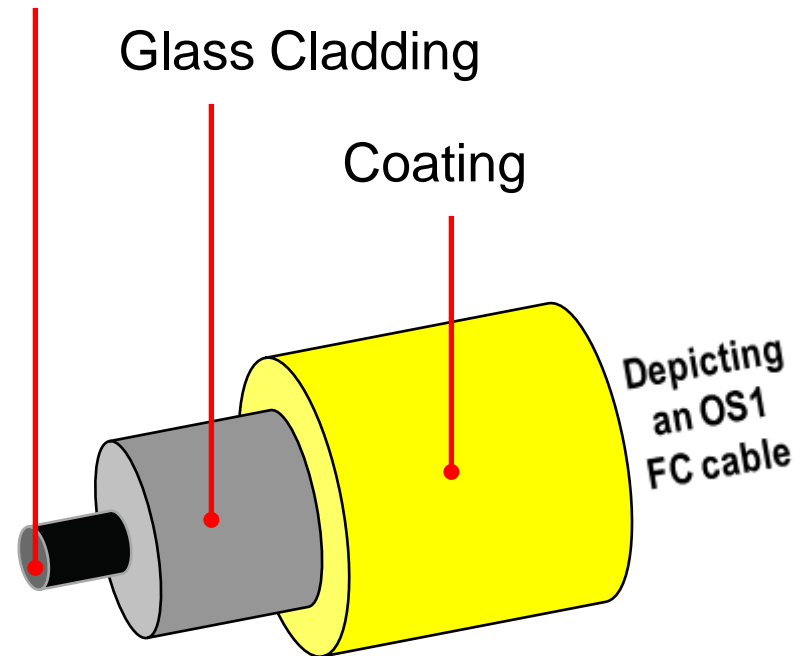


Multimode Fiber

For local distances and/or
1/2/4/8/16Gbps

MM – use with short wave optics
62.5 micron, 50 micron
(62.5m has little value today)

Glass Core



Single-mode Fiber

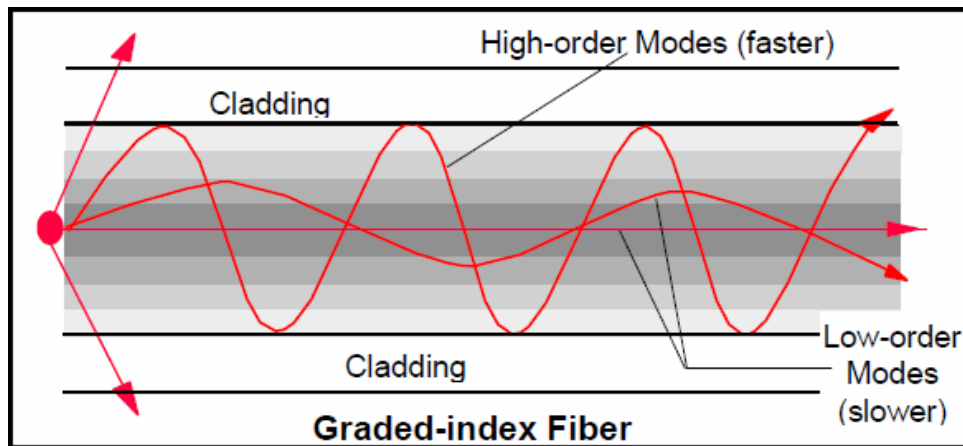
For long distances and/or
8Gbps/16Gbps link rates

SM – use with long wave optics
9 micron

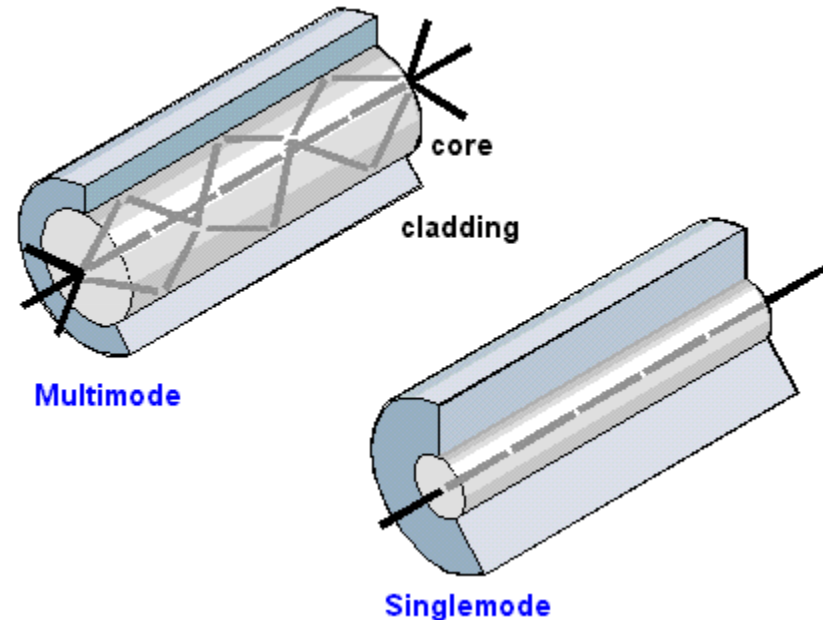
FC Storage Networking Terminology

Fiber Channel Links

- **Modal dispersion** is a distortion mechanism occurring in multimode fibers in which the signal is spread in time because the propagation velocity of the optical signal is not the same for all modes.
- Modal dispersion limits the bandwidth and distance of multimode fibers.

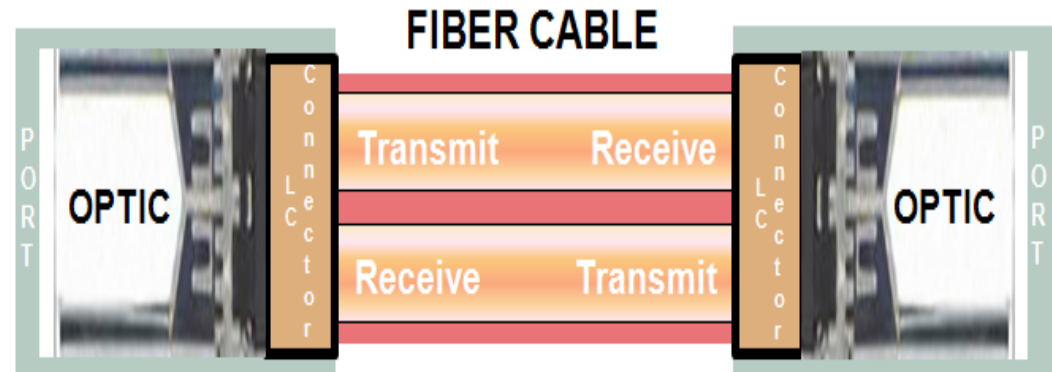


OM3 and OM4



SAN Terminology -- *Fiber Channel Link*

- **A link consists of**
 - 2 unidirectional “fibers” transmitting in opposite directions
 - May be either:
 - Optical fiber or Copper



- **Transmitters may be:**
 - Long wave laser
 - There can be multiple distances for these – ie. 4km/10km
 - Short wave laser
 - LED
 - Electrical

Switched-FCP and Switched-FICON Director Chassis's Run At:

Brocade

2Gbps
4Gbps
8Gbps
16Gbps

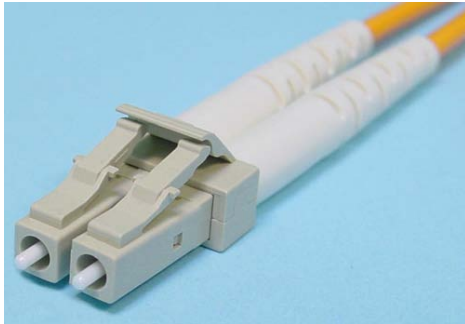
Cisco

2Gbps
4Gbps
8Gbps

SAN Terminology -- *Fiber Channel Link*



OPTICAL



Performance:
2, 4, 8, 16 Gbps

- **LC Optical Connector**
 - Standard on 2-8Gbps Switches
 - Most widely used connector

Cable Shown as Bonded Duplex

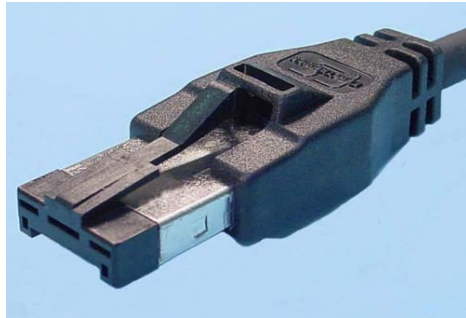


Performance:
1 Gbps

- **SC Optical Connector**
 - Standard on 1 Gbps Switches
 - Little used any longer

Cable Shown as Bonded Duplex

COPPER



Performance:
2.125 Gbps

- **HSSDC Copper Connector**
 - Smaller than older connectors
 - Easier to insert/remove



Performance:
2.5 Gbps

- **HSSDC2 Copper Connector**
 - Fits in SFP Media Slots
 - Smaller than HSSDC

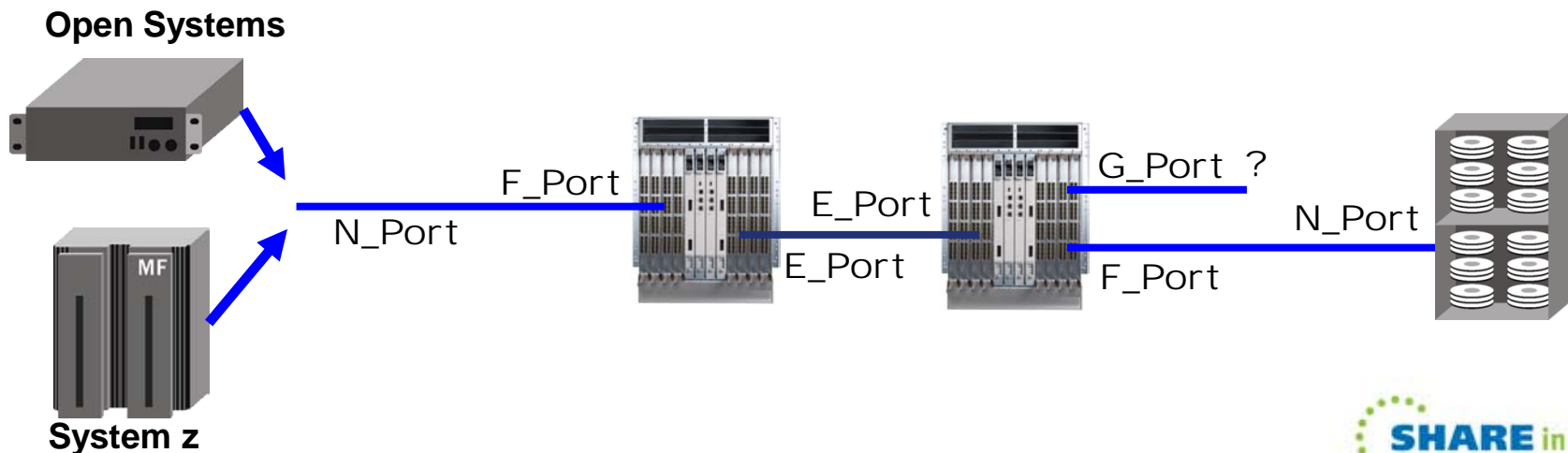
SAN Terminology -- *Fibre Channel*

Device Ports

- N_Port – Node Port, a Fabric device directly attached

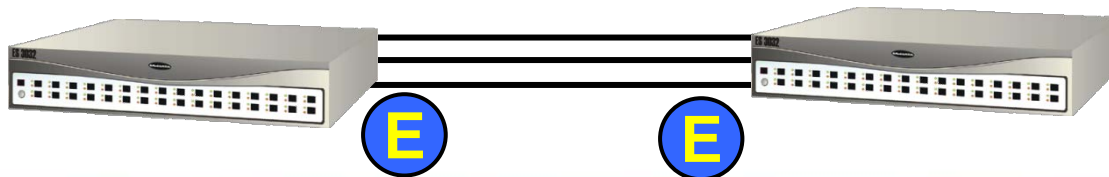
Switch Ports

- G_Port – Generic Port, a port waiting to become an E or F_Port
- F_Port – Fabric Port, a port to which an N_Port attaches
- E_Port – Expansion port used for inter-switch links



SAN Terminology -- *Fibre Channel*

- Interconnection between switches is called the Inter-Switch Link (ISL) or in FICON a Cascaded Link (uses Expansion Ports – E_Port)
 - E_Port to E_Port (aka ISL)
 - For FICON, a 10Gbps link can ONLY BE a cascaded link (ISL)
- Allows switches to be connected together to create a multi-switch Fabric
- Supports all classes of service
 - Class 1, 2, 3, and a special Class F
- The FC Standard permits consecutive frames of a sequence to be routed over different, parallel ISL links for maximum throughput



Agenda for Session 10999 and 11000

Session 10999

- Types and Components of Storage
- **Let's talk Fibre Channel**
- FC Buffer Credits
- Fabric Routing / Virtual Fabrics / Partitioning
- Security / Zoning

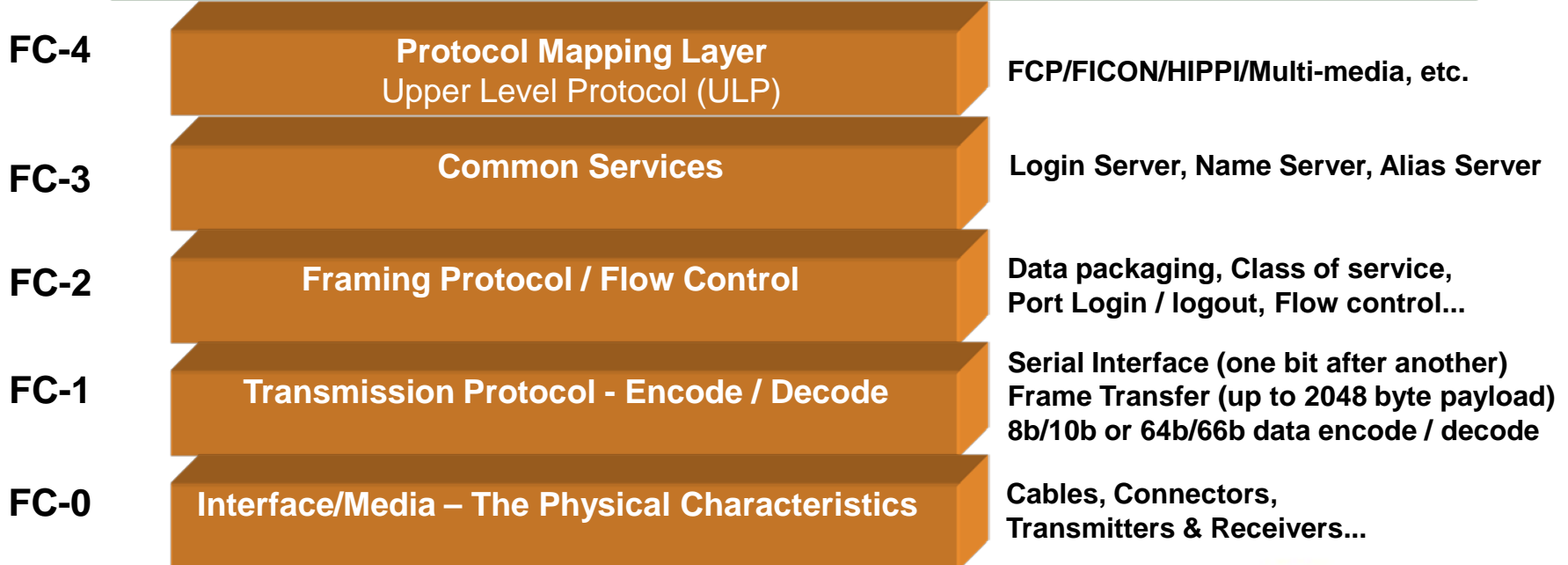
Fibre Channel Protocol

Fibre Channel (FC) provides high speed transport for Upper level (ie. FICON or SCSI) payloads

- FC is the “protocol” for a Storage Network – attributes are:
 - Highly Scale - Addressing for up to 16 million nodes
 - Various Switched Topologies
 - High Speeds - 100, 200, 400, 800 or 1600 MBps
 - 10Gb or 16Gb ISLs can be deployed
 - Segments of up to 100 Km between switches
 - Support for multiple protocols like FICON and OPEN (SAN)
 - Support for Security via Zoning and Prohibit/Allow Matrix

The Fibre Channel Protocol

- **FCP and FICON** are just a part of the upper layer (FC-4) protocol
- **They** are compatible with existing lower layers in the protocol stack



Fibre Channel Overview

- Fibre Channel is primarily utilized as a networked form of SCSI (open systems - **SAN**) or CCWs (System z - **FICON**)
 - **Actually, the lower layers of Fibre Channel are generic**
 - Able to transport multiple data types such as video on demand and Internet Protocol
 - **But, most common deployment is Fibre Channel Protocol (FCP)**
 - FCP is an upper layer protocol that provides for the transmission of SCSI commands and data over the Fibre Channel transport layers
 - **Next most common deployment is Fiber Connection (FICON)**
 - FICON is an upper layer protocol that provides for the transmission of mainframe CCW commands and data over the Fibre Channel transport layers

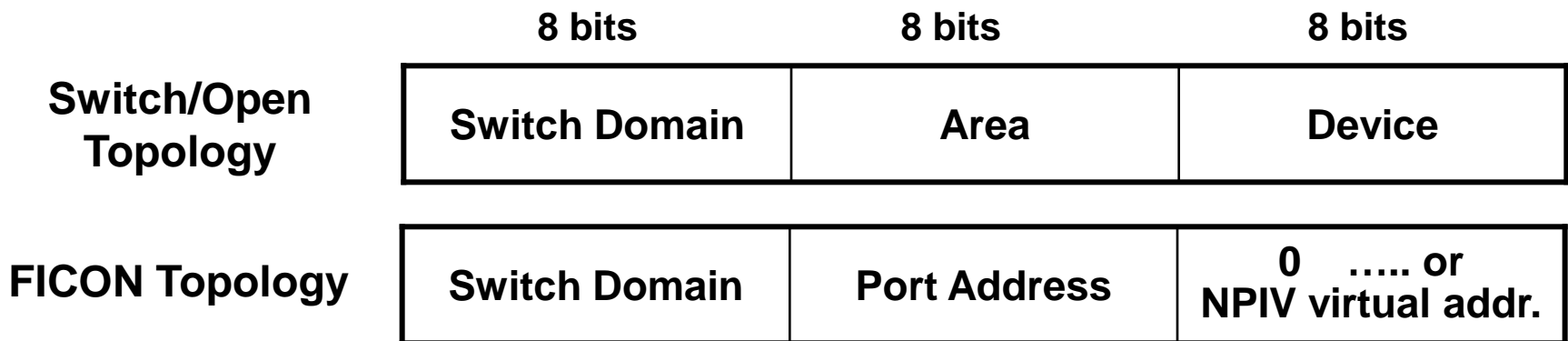
World Wide Names

- Each switch element is assigned a 64 bit WWN at time of manufacture
- Each switch port is assigned a 64 bit WWPN at the time manufacture
- During Fabric Logon (FLOGI) the switch identifies the WWN in the service parameters of the accept frame
- After FLOGI/PLOGI the WWNs and WWPNs have been mapped to Fibre Channel Identification (FCID) addressing

These Address Assignments Can then Correlate
Each Fabric Port with Switch Routing and the Fiber
Channel ID (FCID)

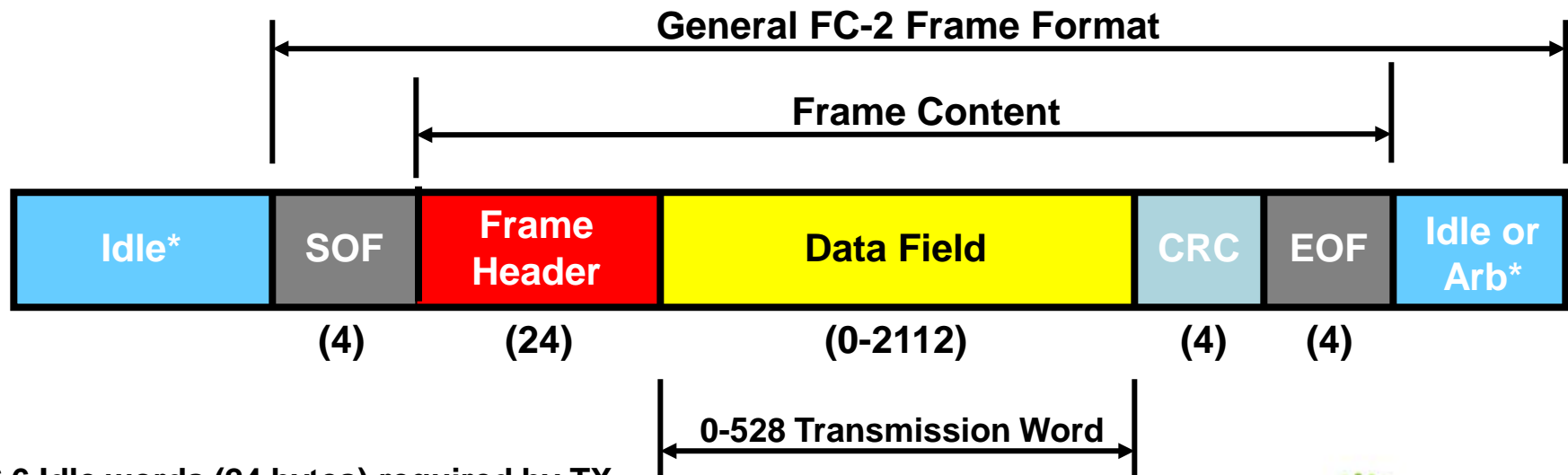
Fabric Addressing

- The 24 bit Fiber Channel Identification (FCID) address is partitioned into 3 fields:
 - Device or NPIV
 - Area
 - Domain
- This partitioning helps speed up routing
- Switch element assigns the address to N_Ports
- Address portioning is transparent to N_Ports



Fibre Channel Frame Format

- All FC-2 frames follow the general frame format as shown below
- Idle/ARB are ‘Ordered Sets’ used for synchronization and basic signaling (1, 2, 4 and 10Gbps use Idle/Idle while 8Gb and 16Gbps use Idle/Arb)
- SOF – Start-of-Frame, EOF – End-of-Frame
- 8b/10b frame encoding for 1, 2, 4 and 8Gbps frames
- 64b/66b frame encoding for 10 and 16Gbps frames



* 6 Idle words (24 bytes) required by TX
2 Idle words (8 bytes) guaranteed to RX

ISL Aggregation

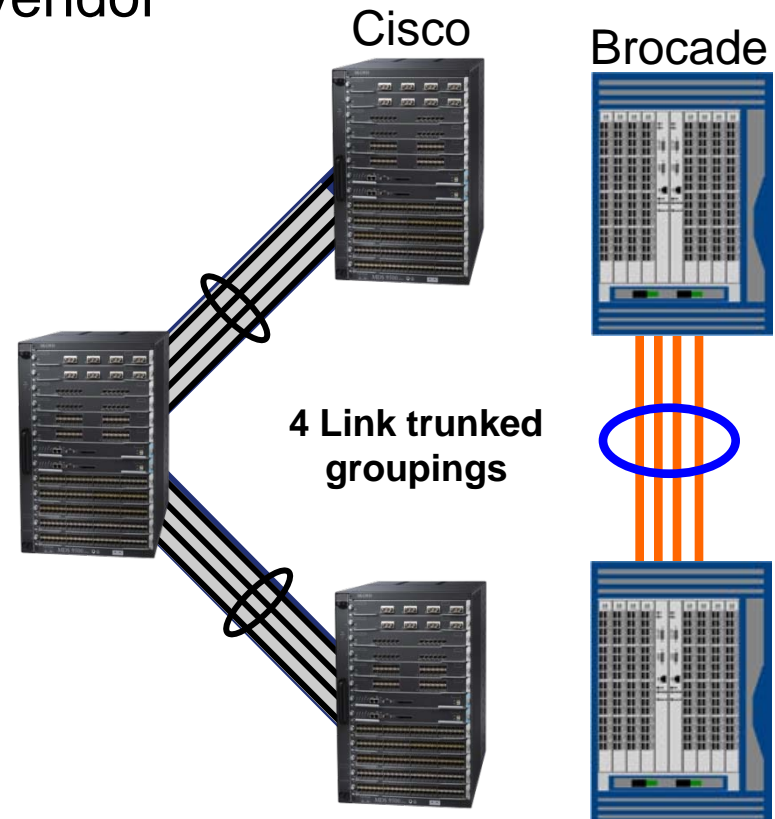
Port Aggregation Is Used to Create a Single Logical ISL from multiple Physical ISLs

Different names depending upon the vendor

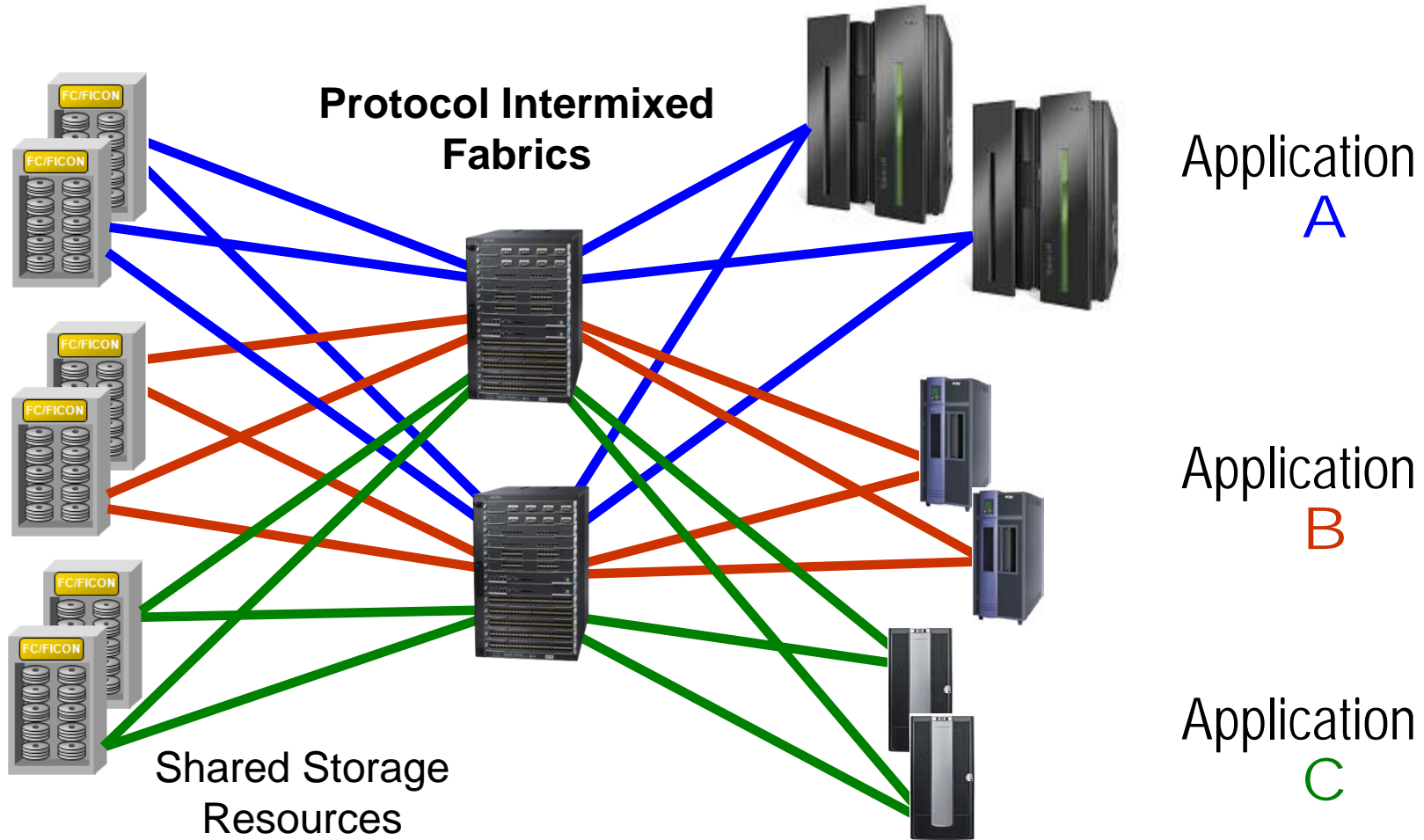
Brocade = Trunking

Cisco = Port Channel

- Increases bandwidth and availability
- Simplifies Topology
- Usually some load balancing
- Interfaces can both be added and removed in a non-disruptive manner in production environments
- Preserves FC guarantee of in-order delivery (IOD)

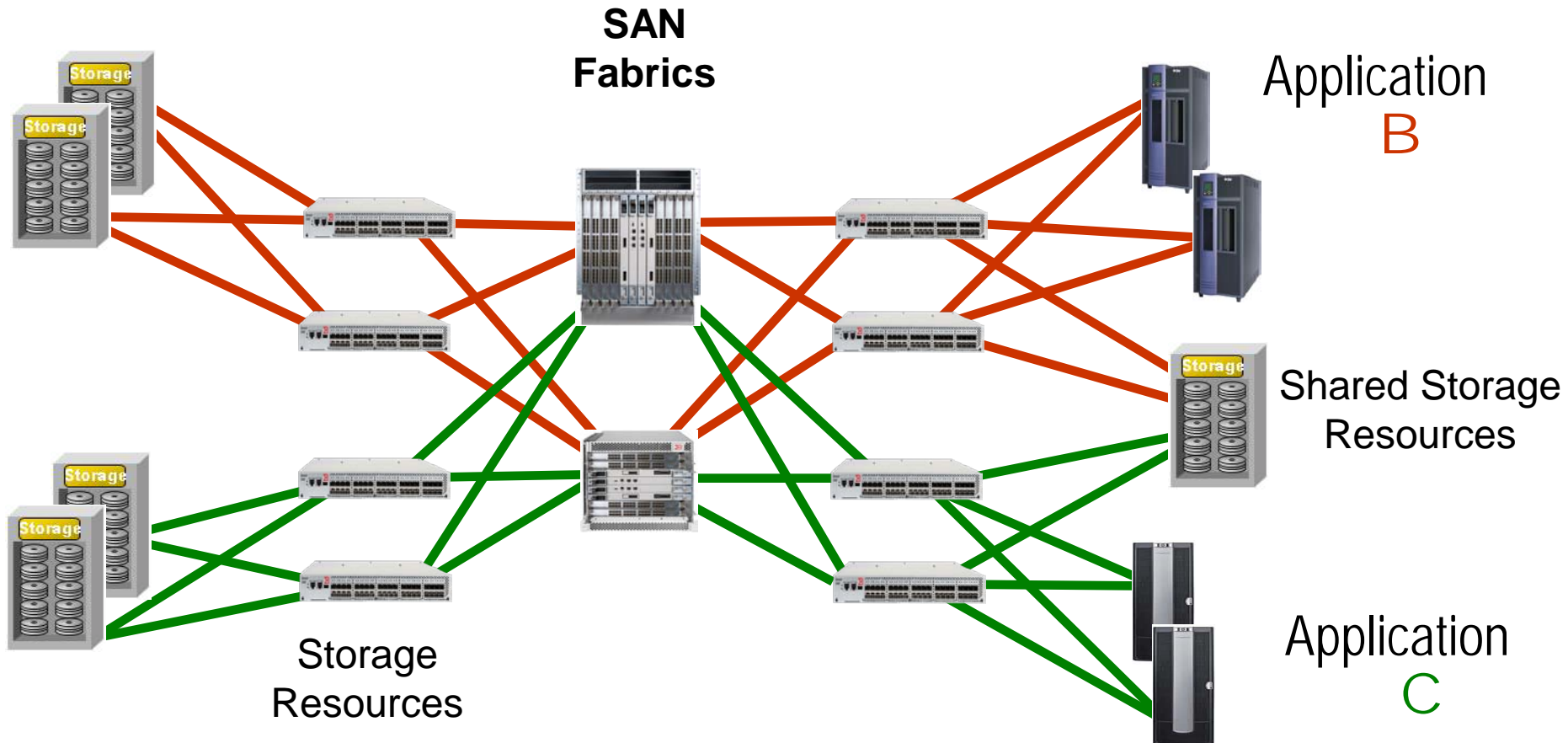


Storage Networking Topology Dual Star (non-cascaded for FICON)



- Provides an opportunity to deploy fabrics to meet five-9's of availability
- Still must insure ≤ 5 min/annual fabric downtime

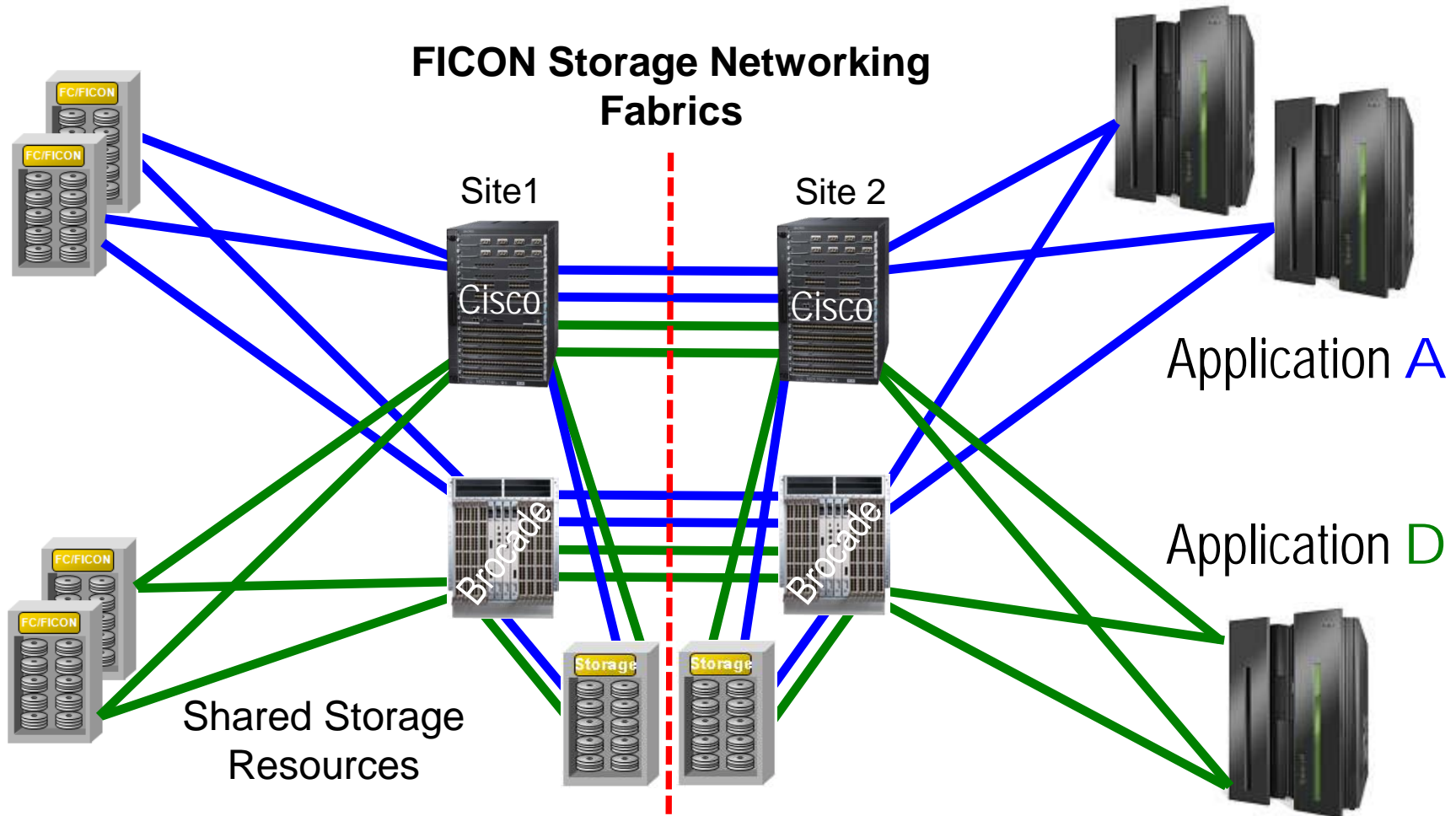
Typical SAN Topology Core-to-Edge (distributed systems only)



- Requires multi-hop so it is not currently supported for FICON due to IBM Qualification

Storage Networking Topology

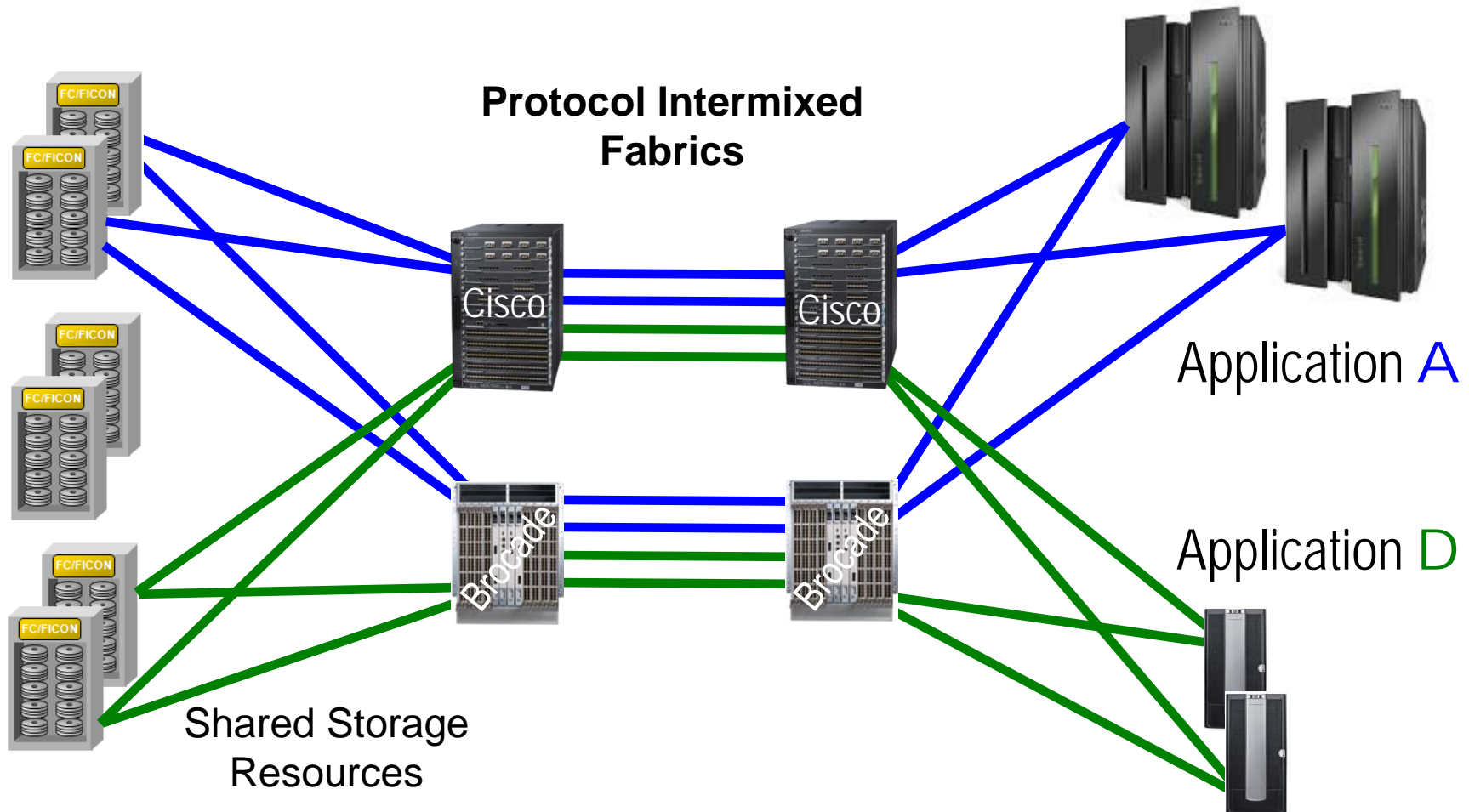
Cascaded FICON



- Provides an opportunity to deploy fabrics to meet five-9's of availability
- Only one vendor per fabric and only one hop per fabric

Storage Networking Topology

Cascaded FICON and FCP



- Provides an opportunity to deploy fabrics to meet five-9's of availability
- Only one vendor per fabric and Only one hop

Agenda for Session 10999 and 11000

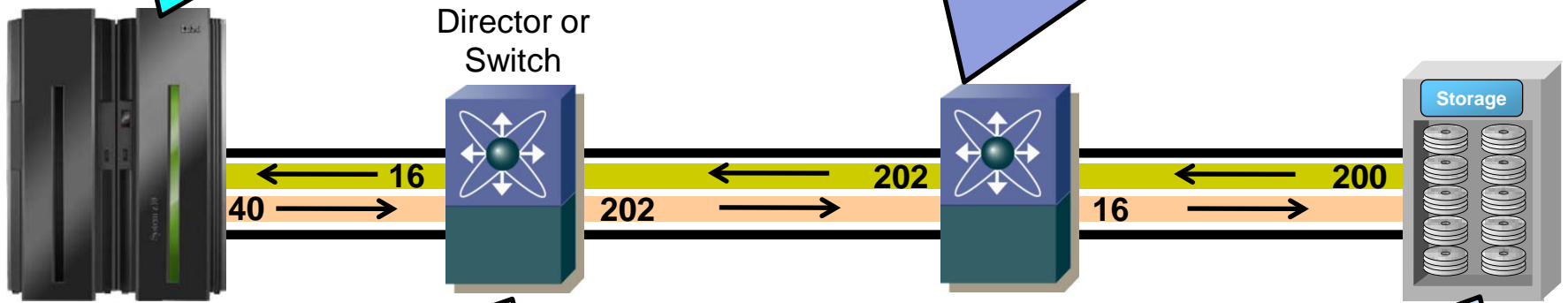
Session 10999

- Types and Components of Storage
- Let's talk Fibre Channel
- **FC Buffer Credits**
- Fabric Routing / Virtual Fabrics / Partitioning
- Security / Zoning

FC Buffer Credits and Flow Control

I can receive 40 frames (buffer credits) but after that you will have to stop sending frames until I acknowledge some or all of them.

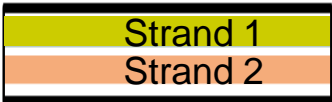
We are long distance ISL links and therefore we need extra buffer credits in order to keep the link fully utilized.



That is fine – I can only receive 16 frames myself. This is OK since we are only a few feet/meters apart anyway.

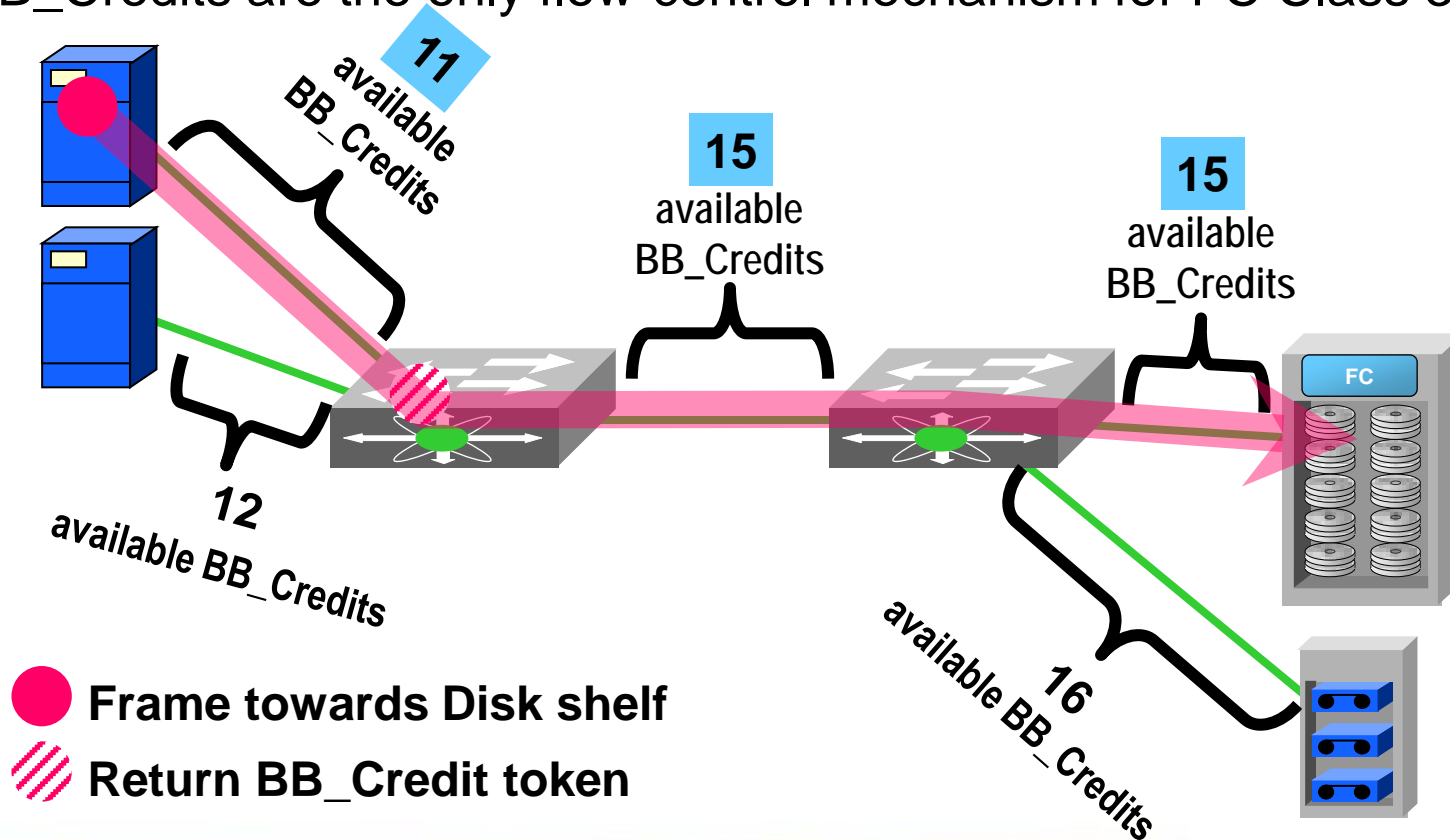
Like most modern DASD and Tape storage, I have a set number of frames that I can handle and I will let you know what the maximum BC limit will be.

Fiber Cable



FC Buffer Credits and Flow Control

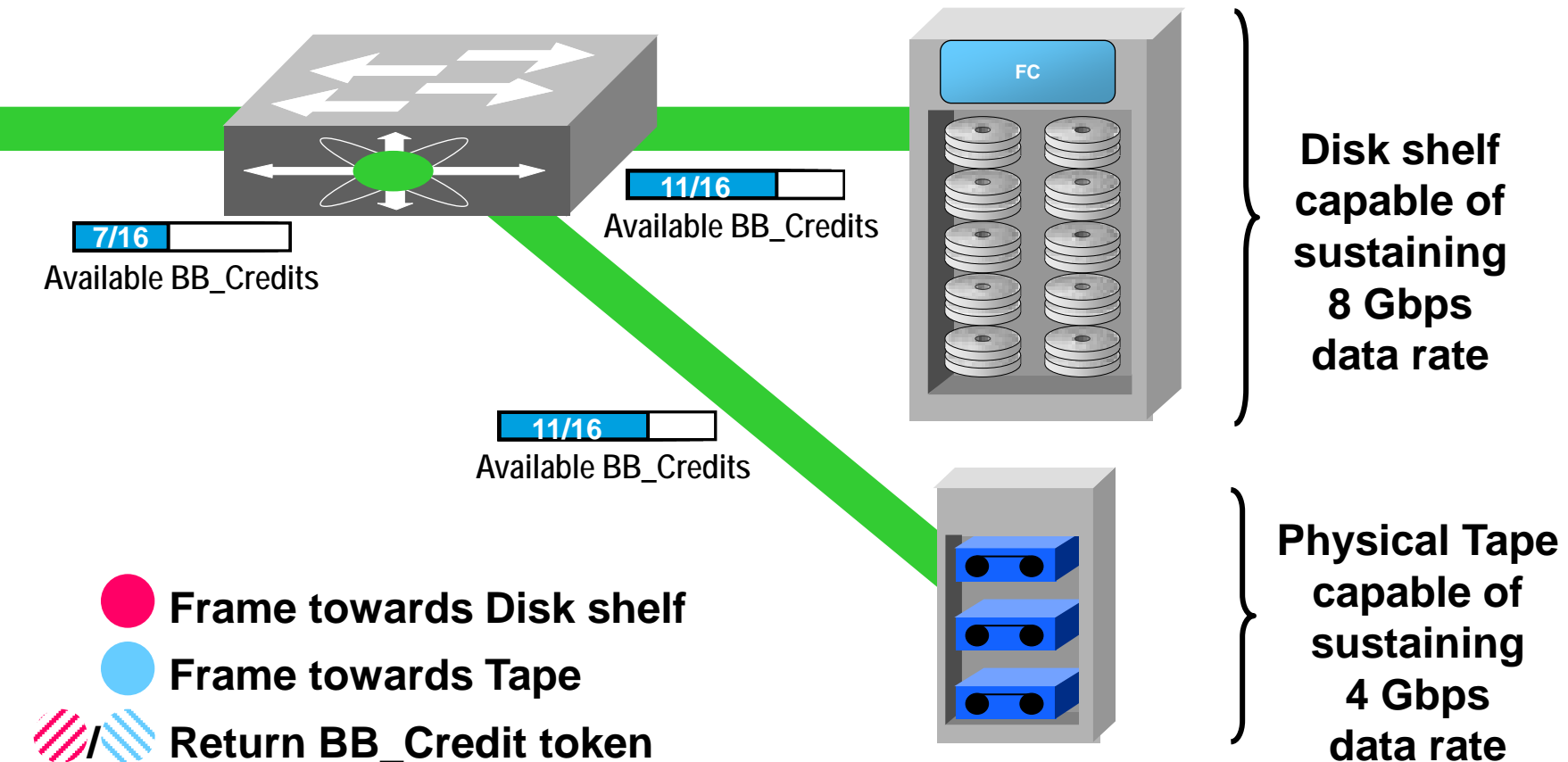
- BB_Credits are the “admission control” mechanism in FC to ensure that FC switches don’t run out of buffers (FC Switches cannot drop frames)
- For Devices operating at FC Class 3 (most devices), Buffer Credits are determined at login.
- BB_Credits are the only flow-control mechanism for FC Class 3.



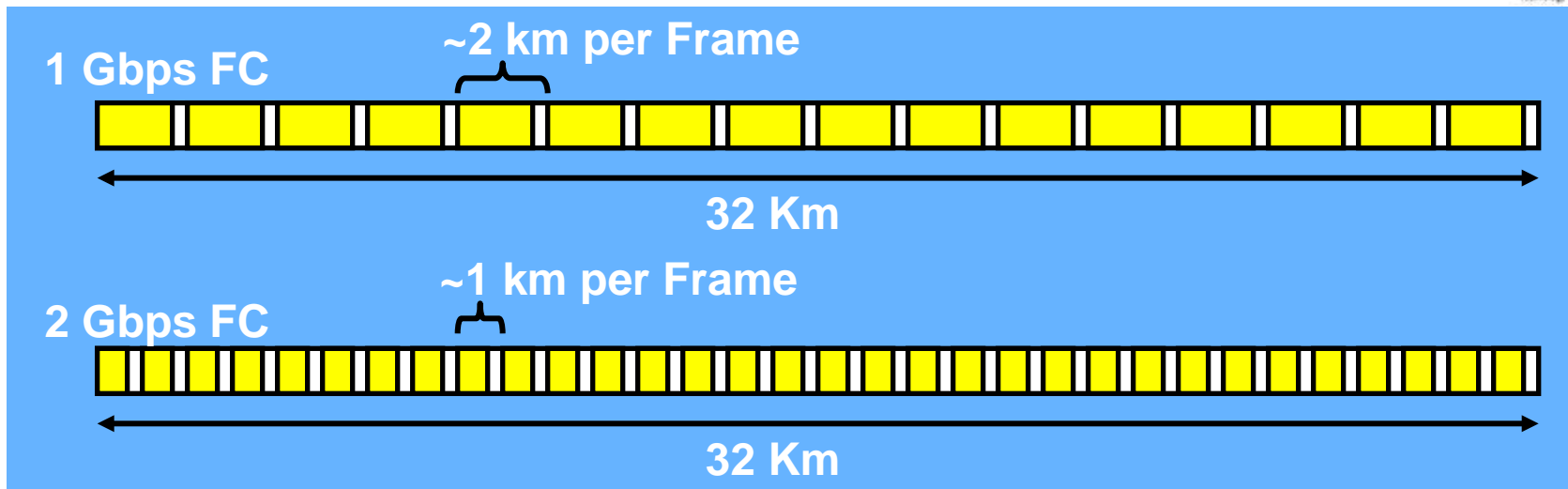
Buffer Credits (BB_Credits): Working Clean

Buffer Credits are a “Flow Control” mechanism to assure that frames are sent correctly

In an ideal FC network, there is no blocking in any device connected to the fabric. (All devices can process frames at the same rate and negotiate equal levels of BB_Credits)



FC BB_Credits and Distance



- BB_Credits are used to make sure that data frames do not overrun the receiver
- BB_Credits are used to ensure high link utilization via enough FC frames in flight
- A full (2112 byte) FC frame is approximately 1 km long @ 2 Gbps and approximately .5 km long @ 4 Gbps and approximately .25 km long @ 8 Gbps, etc.
- As distance increases, the BB_Credits needed for high link utilization increase as well
- Shorter frames will require additional BB_Credits to ensure a “full pipe” – and *for most FC data* it is smaller frames that are sent!
- Insufficient BB_Credits will throttle performance—no data will be transmitted until a frame received acknowledge is returned (R_Rdy)

Buffer Credits Required By Size of Frame and Link Speed

A distance of 20km with the link 100% utilized				2Gbps	4Gbps	8Gbps	10Gbps
SOF, Header, CRC, EOF	Payload	Total Frame Bytes	Smaller than full frame by x%	Buffer Credits Required 8b10b	Buffer Credits Required 8b10b	Buffer Credits Required 8b10b	Buffer Credits Required 64b66b
36	2112	2148	0.000%	20	40	80	117
36	2002	2038	5.138%	21	42	84	124
36	1902	1938	9.809%	22	44	88	130
36	1802	1838	14.481%	24	47	93	137
36	1702	1738	19.152%	25	49	98	145
36	1602	1638	23.823%	26	52	104	154
36	1502	1538	28.494%	28	56	111	164
36	1402	1438	33.165%	30	60	119	175
36	1302	1338	37.836%	32	64	128	188
36	1202	1238	42.507%	35	69	138	203
36	1102	1138	47.179%	38	75	150	221
36	1002	1038	51.850%	41	82	164	243
36	902	938	56.521%	46	91	182	268
36	819	855	60.398%	50	100	199	294
36	700	736	65.957%	58	116	232	342
36	600	636	70.628%	67	134	268	396
36	500	536	75.299%	80	159	318	469
36	400	436	79.970%	98	195	390	577
36	300	336	84.641%	127	254	507	748
36	200	236	89.312%	181	361	721	1065
36	100	136	93.984%	313	626	1251	1848
36	75	111	95.151%	383	766	1532	2264
36	50	86	96.319%	495	989	1978	2922

As distance link speed grows, so does the need for buffer credits!

Agenda for Session 10999 and 11000

Session 10999

- Types and Components of Storage
- Let's talk Fibre Channel
- FC Buffer Credits
- **Fabric Routing / Virtual Fabrics / Partitioning**
- Security / Zoning

Fabric Routing Across ISLs: FSPF

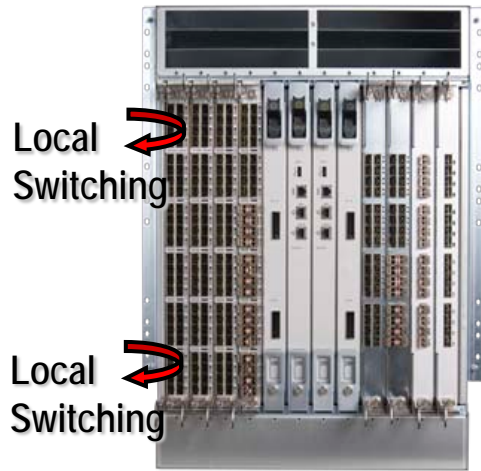


- For FSPF a domain ID identifies a single switch
 - This limits the max number of switches that can be supported in a Fabric to 239 when FSPF is supported
 - Each Cisco VSAN is treated as a separate Fabric
 - Each Brocade non-virtualized chassis is treated as a separate fabric
 - Each Brocade Virtual Fabric (DCX family) is treated as a separate fabric
 - Each Brocade Physical Partition (Mi10K) is treated as a separate fabric
- FSPF performs hop-by-hop routing
 - Each Cisco VSAN runs it's own FSPF process
 - Routing between VSAN's is done with Inter VSAN Routing (IVR)
 - Brocade will use FSPF routing if TRUNKING is not used
 - Old CNT/McDATA devices always use FSPF for initial ISL routing
- FSPF supports hierarchical path selection
 - Provides the scalable routing tables for large topologies

Deploying Virtual Fabrics / Virtual SANs

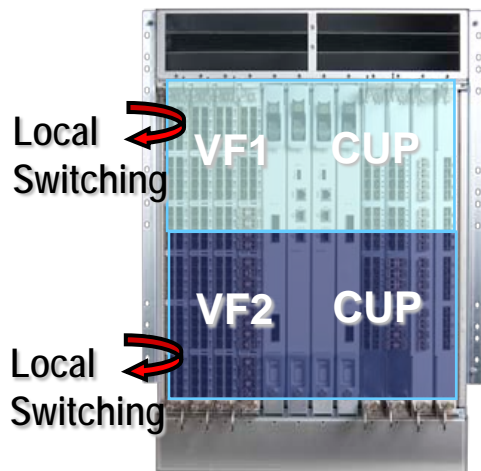
- Multiple Virtual SANs in a single hardware entity
 - Used to Converge SAN Islands
 - IETF RFC 4747
 - Although it is a Standard – Implementations are different per Vendor
- Brocade
 - Creating a Logical Switch creates a Virtual Fabric (VF)
 - Each Logical Switch will have a unique Domain ID
- Cisco
 - Virtual SANs (VSANs) are incorporated into the architecture
 - Each VSAN will have a unique Domain ID

Virtual Fabrics (VF) for DCX Family



We expect most customers to deploy FICON on a non-virtualized chassis

An infrastructure that is supporting intermixed protocol use (FICON/FCP) or hosts different customers on the same infrastructure might desire virtualization

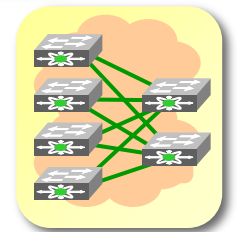
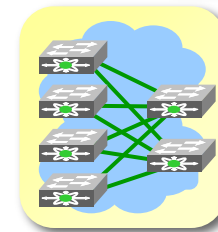
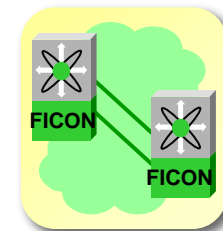
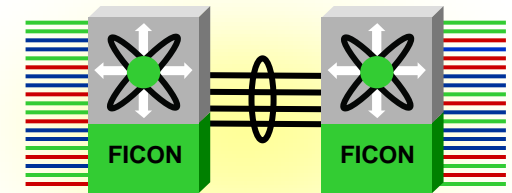


- VF is an OPTIONAL feature on all of the Brocade DCX family of FICON/SAN Directors
- But DCX 48-port blades for FICON can be used only when Virtual Fabrics are enabled on the DCX or DCX 8510-8
 - DCX-4S and the DCX 8510-4 never require the use of VF
- Creating a Logical Switch, which enables Virtual Fabric, began at FOS 6.2 and FOS 7.0 is current
- VFs create new services within each Virtual Fabric
 - Domain ID
 - Name Server, etc
- Provides additional isolation of ports within a chassis

Cisco's VSANs (Virtual SAN)

A way to Partition a Switch or SAN into a Virtual/Logical environment

- Virtual SANs created from larger cost-effective redundant physical fabric
- Reduces wasted ports of the older “island” approach
- Hardware-based isolation
- Statistics can be gathered per VSAN
- Management per VSAN
- Unique Serial Number / CUP per FICON VSAN
- Service and process level Isolation between VSANs



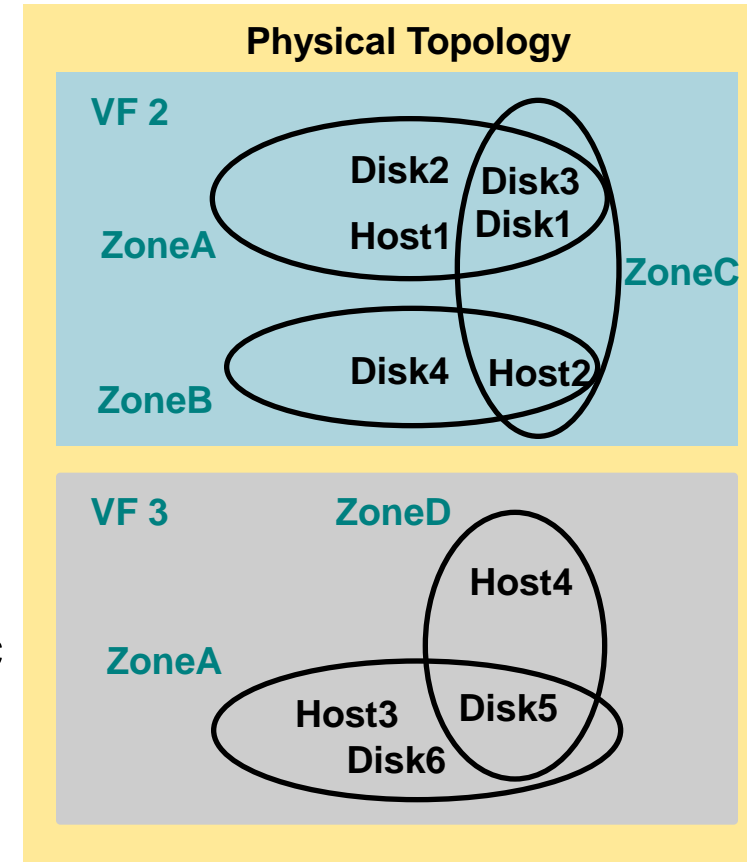
Agenda for Session 10999 and 11000

Session 10999

- Types and Components of Storage
- Let's talk Fibre Channel
- FC Buffer Credits
- Fabric Routing / Virtual Fabrics / Partitioning
- **Security / Zoning**

Zoning

- A logical grouping of fabric connected devices within a SAN (or virtual fabric)
- Zoning establishes access control
 - Devices within a zone can access each other
- Zoning increases security
 - Limiting access prevents unauthorized access
- Zone membership might be configured by:
 - Port World Wide Name (pWWN)—device
 - Fabric World Wide Name (fWWN)—fabric
 - Fibre Channel Identifier (FCID)
 - Fibre Channel Alias (FC_Alias)
 - IP address
 - Domain ID/port number
 - Interface



FICON Port Security

- Zoning
 - Able to be used with FICON (some vendors require it, others don't)
- Single Domain
 - Prohibit / Allow Matrix (ala. ESCON like port blocking/unblocking)
- Cascaded Configurations
 - Access to Cascaded resources controlled from IOCDs
 - Can also be done with Zoning but most choose not to

172.18.44.85 - Vsan 20 Port Configuration

Show Installed Ports Only Port: Prohibit:

Port Adre...	Name	Block	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D
00 (fc1/1)								X								
01 (fc1/2)								X								
02 (fc1/3)								X								
03 (fc1/4)								X								
04 (fc1/5)	mainframe port1							X								
05 (fc1/6)	mainframe port2	B	X	X	X	X	X	X	X	X	X	X	X	X	X	X
06 (fc1/7)								X								
07 (fc1/8)	disk1							X							X	
08 (fc1/9)								X								
09 (fc1/10)								X								
0A (fc1/11)								X								
0B (fc1/12)								X								
0C (fc1/13)	disk2							X		X						
0D (fc1/14)								X								

- You can block or prohibit ports, eg:
- 05 is blocked and prohibited on all
 - 04 can reach 07 and 0C
 - 07 is prohibited from 0C

SAN Sessions at SHARE this week



Monday:

Time-Session

1100 - 11000: SAN (Storage Area Networking) Overview - Part 2

1300 - 11021: Buffer-to-Buffer Credits, Exchanges, and Urban Legends

Tuesday:

Time-Session

1330 - 11003: A First Look at the Inner Workings and Hidden Mechanisms of FICON Performance

1500 - 11004: A deeper look into the Inner Workings and Hidden Mechanisms of FICON Performance

1800 - 11022: FICON Director and Channel Free-For-All

Thursday:

Time-Session

1630 - 11027: Evolution of the System z Channel

This Is The End Of Part 1

**The 2nd of this 2-part session is
Session 11000**

Monday: 11:00AM – 12:15PM

THANK YOU FOR ATTENDING!

Session 10999

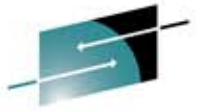
Please fill out your evaluation forms



Questions and Discussion



Session 11000



Introduction to Storage Technologies FICON (**F**iber **C**ONnection) and

a little SAN (**S**torage **A**rea **N**etworking)

David Lytle – Brocade – dlytle@brocade.com

Mike Blair – Cisco – mblair@cisco.com (not here today)

SAN 201

A decorative background graphic at the bottom of the slide. It features several overlapping, wavy bands of color in shades of blue, green, and black. A series of white lines radiate from the left side, curving towards the right. A dotted green line also curves across the bottom right area.

Agenda for Session's 10999 and 10999



1st Session 10999 – 09:30am – 10:30am

- Types and Components of Storage
- Let's talk Fibre Channel
- FC Buffer Credits
- Fabric Routing / Virtual Fabrics / Partitioning
- Security / Zoning

.... then continued in the 2nd session...

2nd Session 11000 – 11:00am – 12:15pm

- History
- Mainframe Terminology, Connectors, Cables, and Wavelengths
- Addressing in FICON
- ESCON Status, zHPF and NPIV
- Buffer Credits, CUP, RMF

Brocade, Cisco and IBM Graphics are used throughout this presentation.

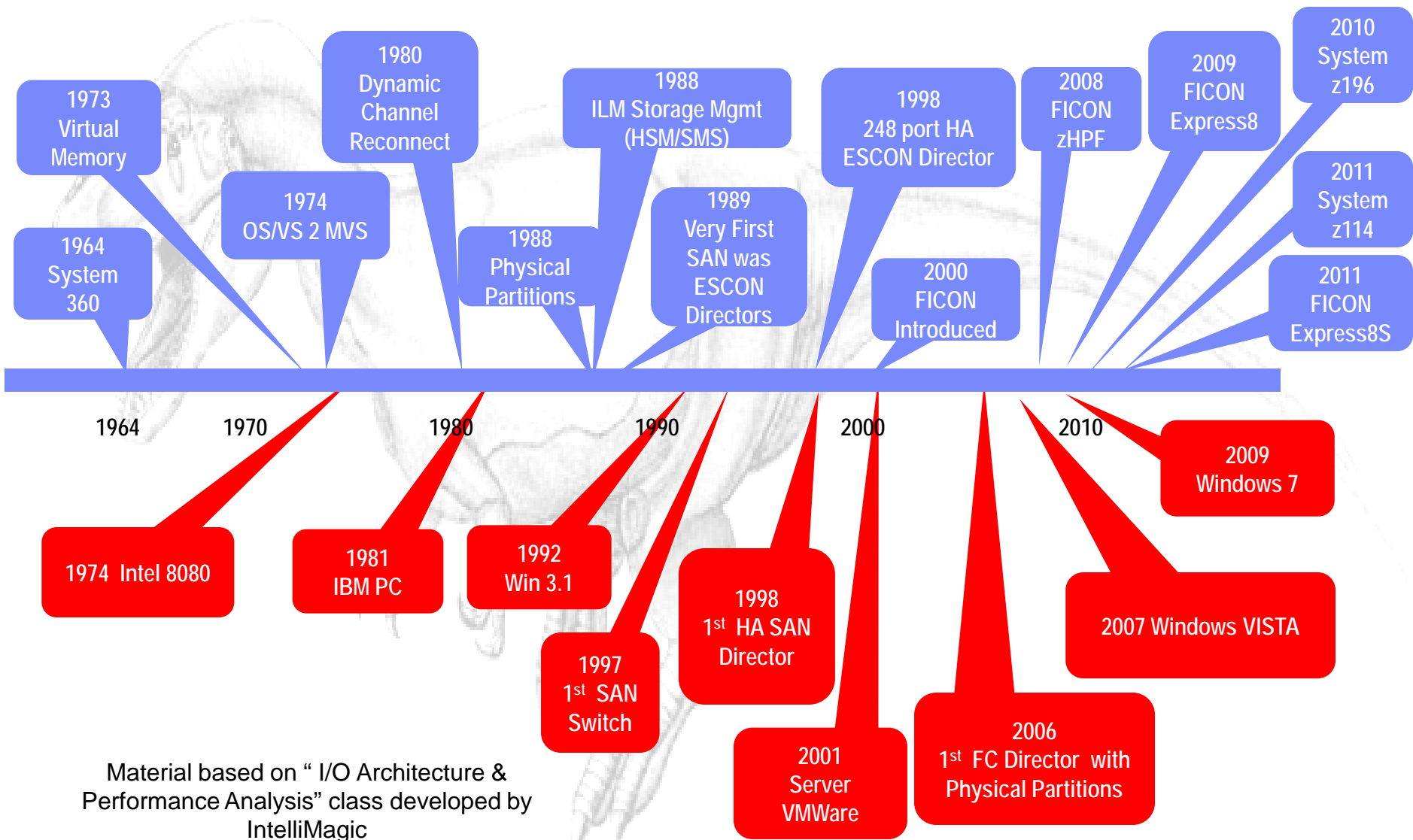


Agenda for Session 10999 and 11000

Session 11000

- **History**
- Mainframe Terminology, Connectors, Cables, and Wavelengths
- Addressing in FICON
- ESCON Status, zHPF and NPIV
- Buffer Credits, CUP, RMF

Mainframe And Open System Time Lines



Material based on " I/O Architecture & Performance Analysis" class developed by IntelliMagic

Agenda for Session 10999 and 11000

Session 11000

- History
- **Terminology, Connectors, Cables, and Wavelengths**
- Addressing in FICON
- ESCON Status, zHPF and NPIV
- Buffer Credits, CUP, RMF

SAN Terminology -- *Definitions*



- “Fibre” is the protocol/architecture used to transport frames
 - As in “the fibre channel protocol”
- “Fiber” is the glass cable used to attach connections
 - As in “the fiber cable”
- Allows up to 16 million nodes of connectivity
- Historically has used 8-bit/10-bit encoding/decoding to translate an 8-bit byte of information to 10-bit format for serial transmission within the payload of a fibre channel frame
 - But 10Gbps and 16Gbps utilize 64b66b for more efficiency
- Variable frame size with a maximum user payload of 2112 bytes.
- Supports transfers of files that are blocked at up to 128MB in size.
- Supports full duplex transmission
 - Simultaneous send and receive

SAN Terminology -- *Fiber Channel Link*



- Light wavelengths in fiber are expressed in nanometers
- Speed of light in fiber cable is $2/3^{\text{rds}}$ the speed of light in a vacuum
 - Light travels at ~5 nanoseconds per meter (3.3 ft) of distance in glass
- **Multimode fiber** is used for numerous frequencies which are all short-wave frequencies (62.5, 50 micron) of laser light.
 - Always used with short wave optics (transceivers)
 - Used for local distance connectivity (~33-1,640 feet...or...10-500 meters)
- **Single-mode fiber** has a smaller core that allows only one frequency of light (9 micron) which is long-wave laser light.
 - Always used with long wave optics (transceivers)
 - This is used for longer distance connectivity (up to 15.5 miles or 25 km)
- Optical power budgets, or link loss budgets, measured in decibels (dBs), are used to manage optical signal loss.

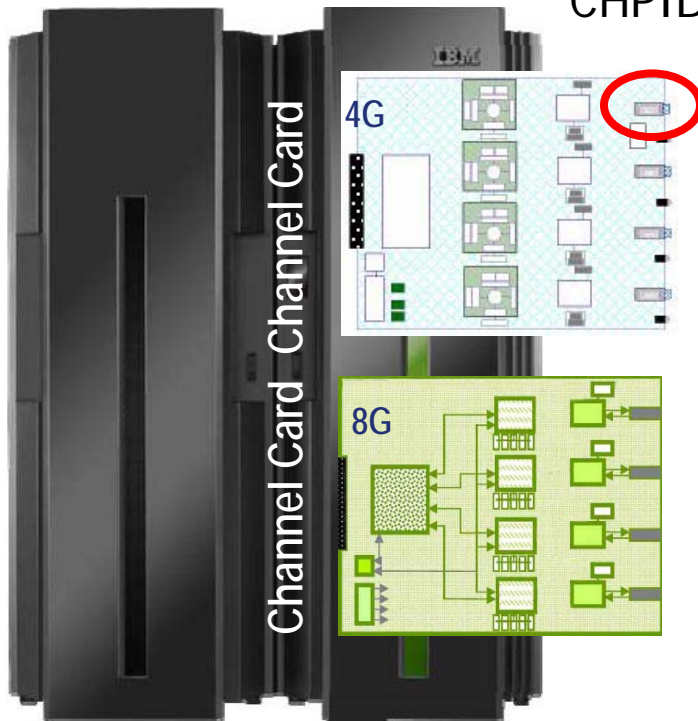
Why Customers Use Mainframes

- Organizations that run an IBM System z platform have unique business requirements.
- They need a **very powerful, highly available, robust and scalable environment** that can support high-volume transaction processing with demanding batch windows and large, critical application portfolios.
- There are tremendous opportunities (like private cloud computing) for leveraging the vast computing power of System z to help greatly simplify the computing environment while substantially reducing overall costs and power consumption.
- All of this makes System z a critically important platform for the future of large enterprise-scale computing!

Mainframe Terminology

System z

This is one
CHPID



Channel or Channel Path Identifier (CHPID)

A physical connectivity port which is embedded in the mainframe processor's channel system

A CHPID will be either long wave or short wave and depending upon the installed channel card it could be 1Gbps, 2Gbps, 4Gbps or 8Gbps referred to as FICON Express or FICON Express_X (where X is 2, 4 or 8) or FICON Express8S

Up to 336 FICON CHPIDs on z9 / z10

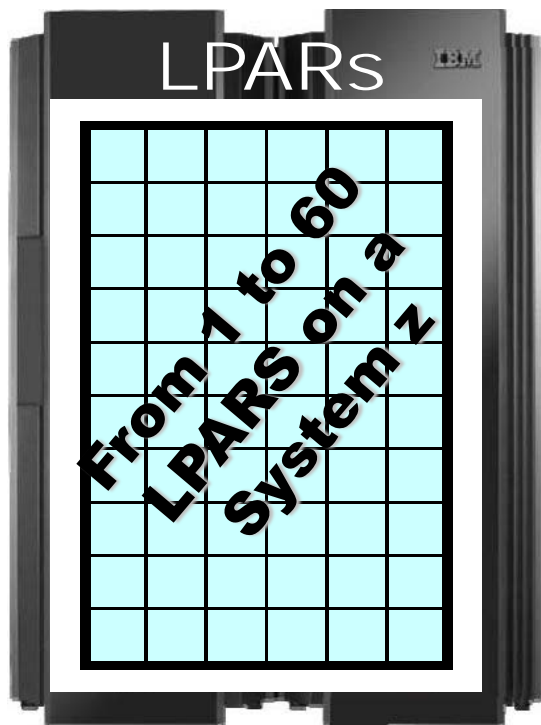
Up to 320 FICON CHPIDs on z196

Up to 128 FICON CHPIDs on z114

All CHPIDs on a channel card must be the same -- LX or SX – no mixture

Mainframe Terminology

System z Mainframe (M/F)



Each LPAR runs its own operating system and has some number of processors assigned to it as CPs and specialty engines

Logical Partitions (LPARs)

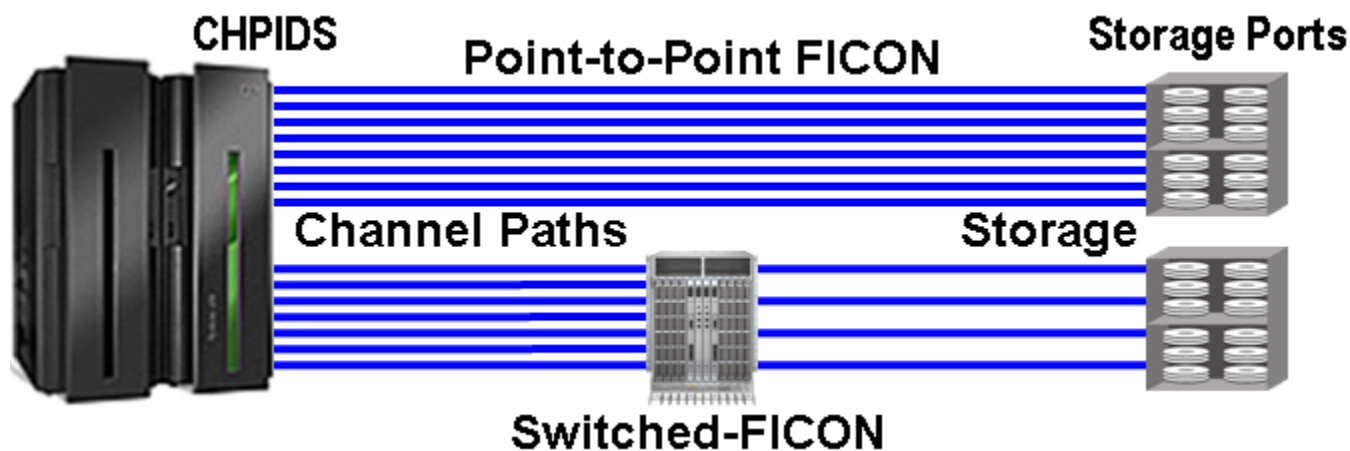
- Available since 1988 this is a way of dividing up a mainframe's capacity into **L**ogical **P**ARTitions and isolating each LPAR from every other LPAR to ensure RAS.
- System 390 and zSeries could run up to 15 partitions per physical mainframe.
- System z can run up to 60 partitions per physical mainframe.

Mainframe Terminology

Channel Path

The fiber between the channel and the storage subsystem, as well as the interface adapter on the subsystem and any interleaving directors

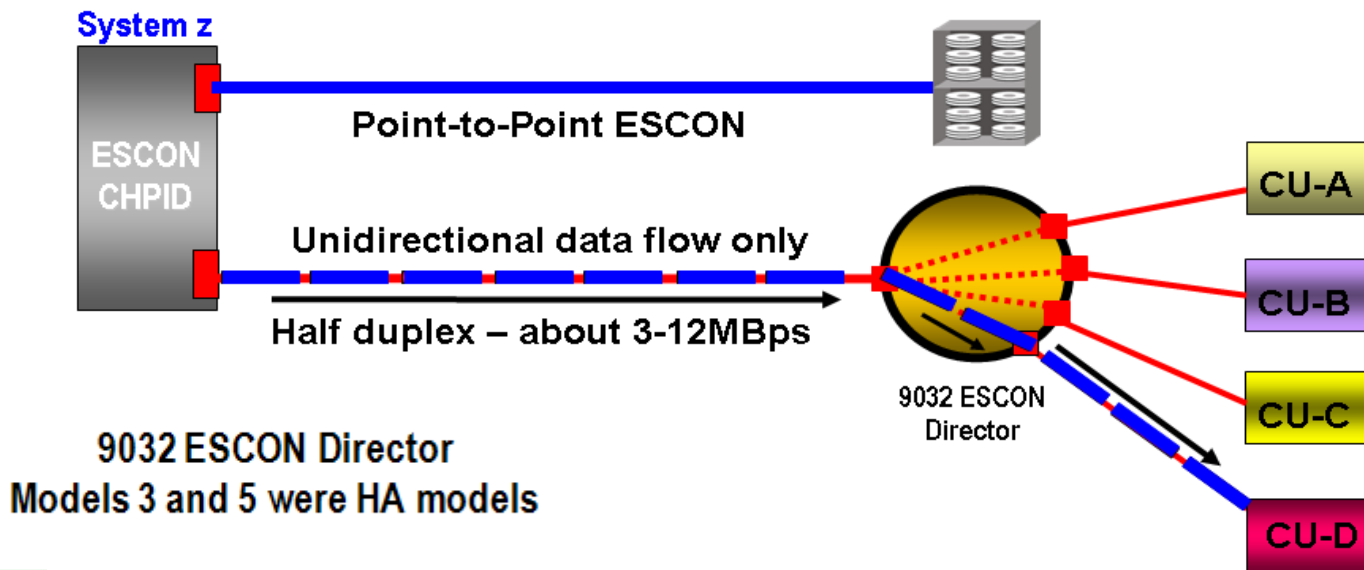
A channel path can be Point-to-Point path or it can be a switched path



Mainframe Terminology

ESCON Channel Paths – 1990

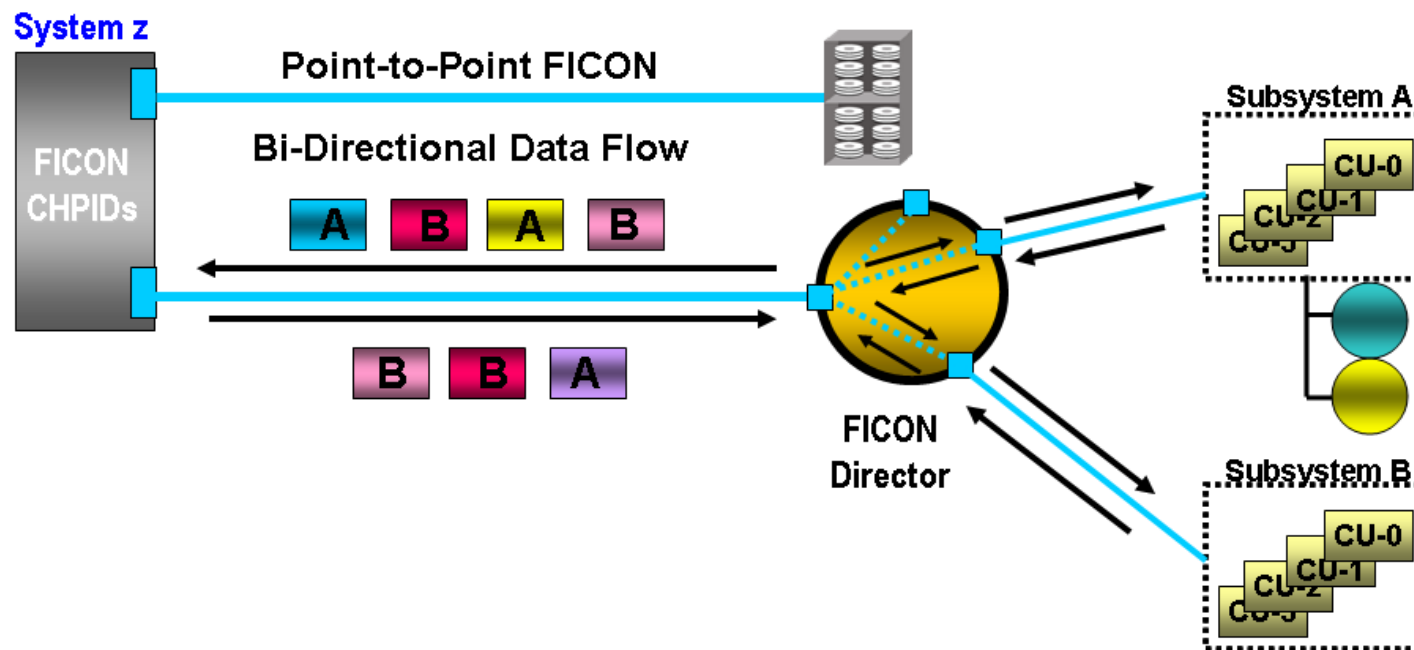
- This old protocol provides a circuit switched, unidirectional data transfer mechanism.
- Once proprietary, it did become a FC standard – SBCON
- Once a data transfer for an I/O from channel to subsystem or subsystem to channel has begun, no other I/O operations can employ the channel until that transfer has been completed.



Mainframe Terminology

FICON Channel Paths - 2000

- This newer but mature protocol provides a very robust packet switched, bi-directional data transfer mechanism.
- System z10, z196 and z114 can do up to 64 open exchanges unless zHPF is driving the I/O in which case ~600 OEs can be active.



ESCON versus FICON I/O Protocol

- **ESCON** among other things is:

- Very old (September 1989)
- Proprietary protocol / now SBCON
- Half-Duplex at ~10-14 MBps
- Maximum of about 1,200 IOps
- Short distances of ~3-9 km
- Restricted number of control units per channel - 15
- Only 1,024 devices per channel allowed
- No channel consolidation going from bus-tag to ESCON
- Lots of multi-mode optical cables used for connectivity needed to be managed

- **FICON** among other things is:

- Most Current (December 2000)
- FC Standards-based protocol
- Full-Duplex at 130-1600 MBps
- 23,000 (CCW) or 92,000 (TCW) IOps
- Long distances of Local-to-25 km
- Better use of the 15 control units than ESCON
- 16,384 devices per channel are now allowed
- Channel consolidation from 2:1 up to 16:1 when going to FICON
- Fewer single-mode and/or multi-mode cables are needed makes management easier

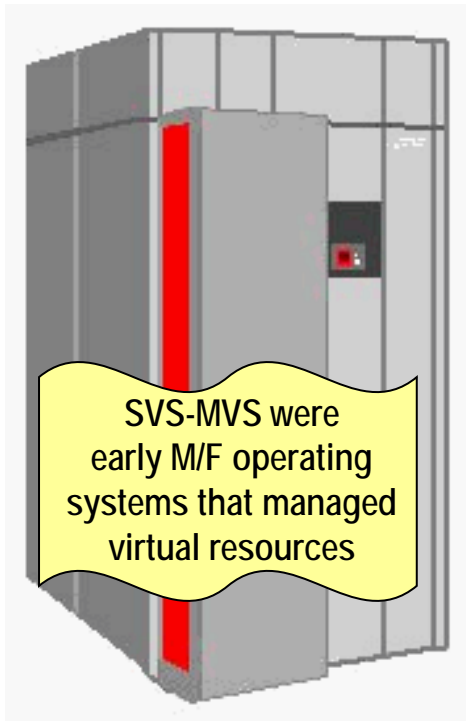
So What Is FICON (**F**iber **C**ONnection)?

- FICON has always been non-proprietary – based upon Fibre Channel Standards
- FICON was a major technical improvement over ESCON and continues to mature and get better and faster all the time
- Worldwide, there are still about ~20 to 30% of ESCON users that still need to completely migrate to FICON
 - The z114 mainframe will be the last mainframe to host ESCON CHPIDs
 - ESCON Directors will soon go to End of Support
- Many factors are pushing end users towards FICON:
 - Mainframes are getting faster and more flexible
 - DASD and Tape storage is getting faster and faster
 - Customer requirements are much wider and deeper
 - IBM support for ESCON is waning

Mainframe Terminology

1971 / 1974

S/390 Mainframe (M/F)



MVS could be run in up to 15 partitions on a M/F or could be intermixed with other operating systems running in other LPARs

Single Virtual Storage (SVS): Multiple Virtual Storage (MVS)

MVS was introduced in 1974, and it is often mentioned, even today, by old mainframer's as their primary operating system

MVS has become the de facto name for the mainframe operating system

This z/OS ancestor was a 24-bit, virtualized, batch processing-oriented operating system that managed lots of memory and DASD space for its time.

Mainframe Terminology

System z Mainframe (M/F)



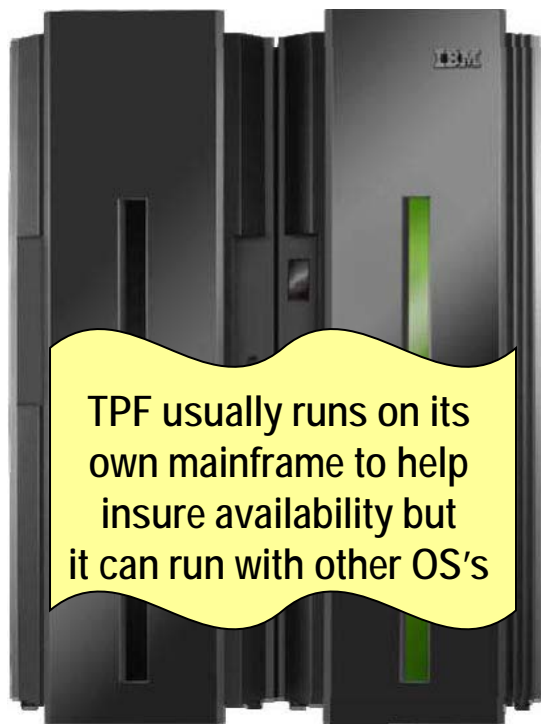
z/OS can run in up to 60 partitions on a M/F or it can be intermixed with other operating systems running in other LPARs

Z Operating System (z/OS)

- Was introduced in 2000 when the zSeries mainframes became available.
- z/OS is a 64-bit server operating system, the latest IBM mainframe operating system, combining MVS and UNIX System Services (formerly known as MVS Open Edition, or OpenMVS).
- Starting in 2010, on z196 and z114, z/OS can directly connect and manage an IBM BladeCenter with Power 7 and IBM System x blades

Mainframe Terminology

System z Mainframe (M/F)



TPF is a risk adverse system that demands the highest performance and availability levels!

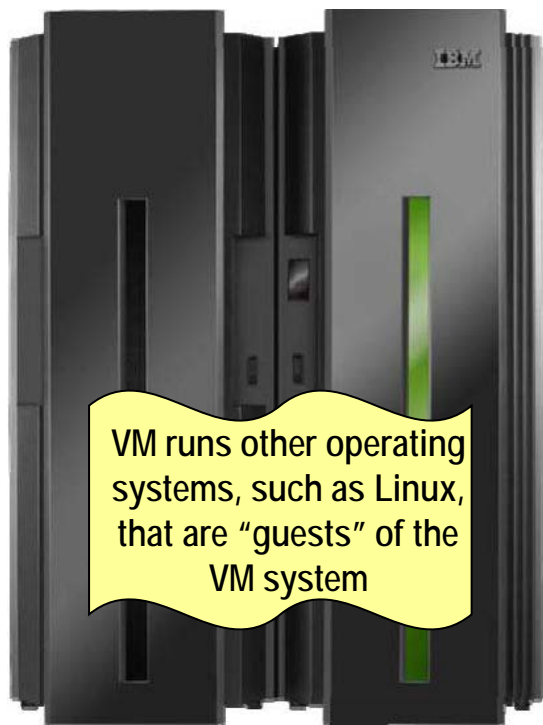
Transaction Processing Facility (TPF)

This is a mature real-time operating system that processes many requests very quickly and runs on mainframes like *zSeries* and *System z*.

Any company that needs to process very high volumes of transactions (hotels, airlines, cruise lines, etc.) often utilizes this operating system which requires very high availability.

Mainframe Terminology

System z Mainframe (M/F)



Linux on the System z is most often run as a "guest" under VM and VM actually does the I/O on behalf of Linux

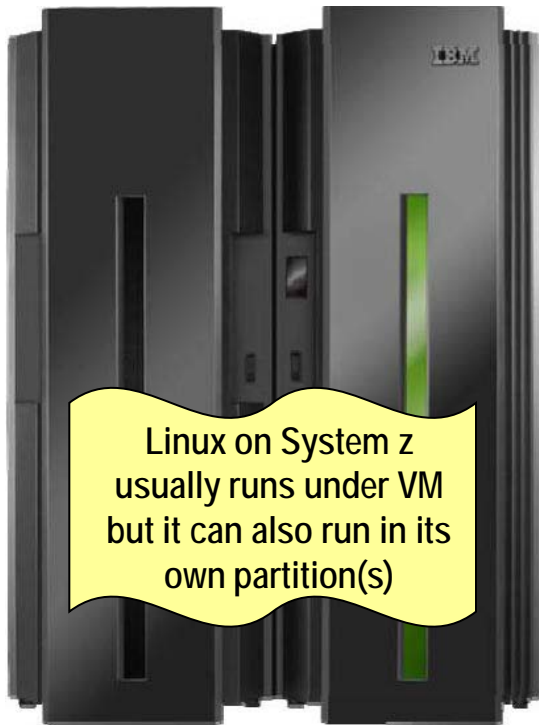
Virtual Machine (VM)

Officially called VM/ESA (Enterprise Systems Architecture), it is an operating system for mainframes that can host other operating systems, including z/OS and Linux.

Each of the guest OS's seems to have its own complete system of software and hardware resources (data storage, processor,...) but are actually sharing resources via VM services.

Mainframe Terminology

System z Mainframe (M/F)



Linux On System z (zLinux)

- Linux on System z allows a customer to leverage their highly available, reliable and scalable mainframe along with their highly available and powerful FCP and FICON infrastructure capabilities
- A Linux administrator now simply administers Linux on a “Big Server”
- Linux has been running on System z since 1999 – thirteen years ago!

FICON channels in FCP mode for use with zLinux use the Queued Direct Input/Output (QDIO) I/O for communication with the operating system.

Mainframe Terminology

Hardware Configuration Definition (HCD)

Define the mainframe's computing and I/O environment

Example FICON IOCP

ID

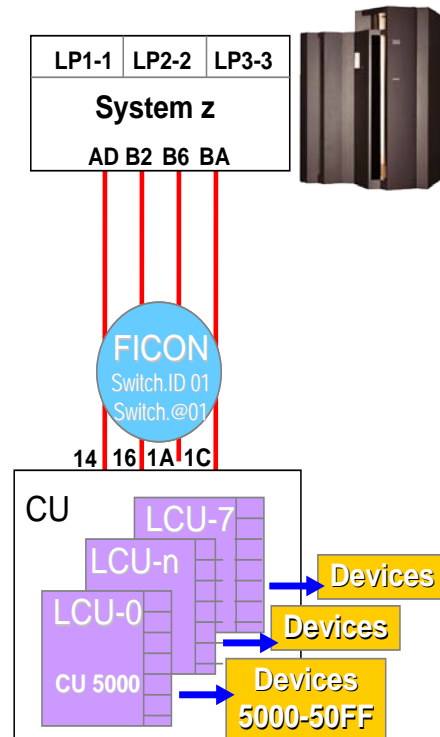
```
RESOURCE PARTITION=(LP1,1),(LP2,2),(LP3,3)
CHIPID PATH=(AD),SHARED, SWITCH=01,TYPE=FC
CHIPID PATH=(B2),SHARED, SWITCH=01,TYPE=FC
CHIPID PATH=(B6),SHARED, SWITCH=01,TYPE=FC
CHIPID PATH=(BA),SHARED, SWITCH=01,TYPE=FC
```

```
CNTLUNIT CUNUMBER=5000, PATH (AD,B2,B6,BA),
UNITADD=((00,256)),LINK=(0114,0116,011A,011C),
CUADD=0,UNIT=2105
```

*

```
CNTLUNIT CUNUMBER=5700, PATH (AD,B2,B6,BA),
UNITADD=((00,256)),LINK=(0104,0106,010A,010C),
CUADD=7,UNIT=2105
```

IODEVICE



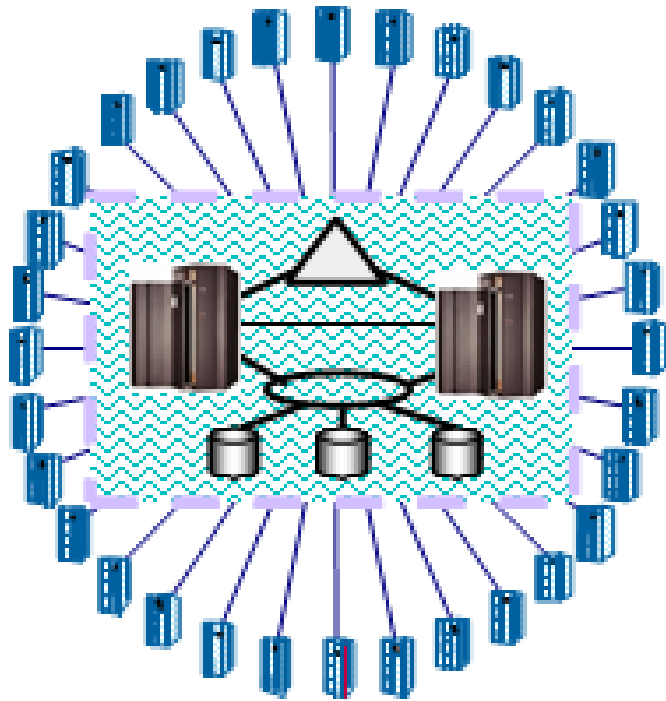
Hardware Configuration Definitions (HCD)

- HCD provides an interactive interface that allows customers to define the hardware configuration for both a processor's channel subsystem and the OS running on the processor.
- There really isn't a comparable facility in the distributed world.

HCD is an element of z/OS that provides the interactive tool which is used to define the hardware configuration for both a processor's channel subsystem and the operating system running on the processor.

Mainframe Terminology

Sysplex and Parallel Sysplex



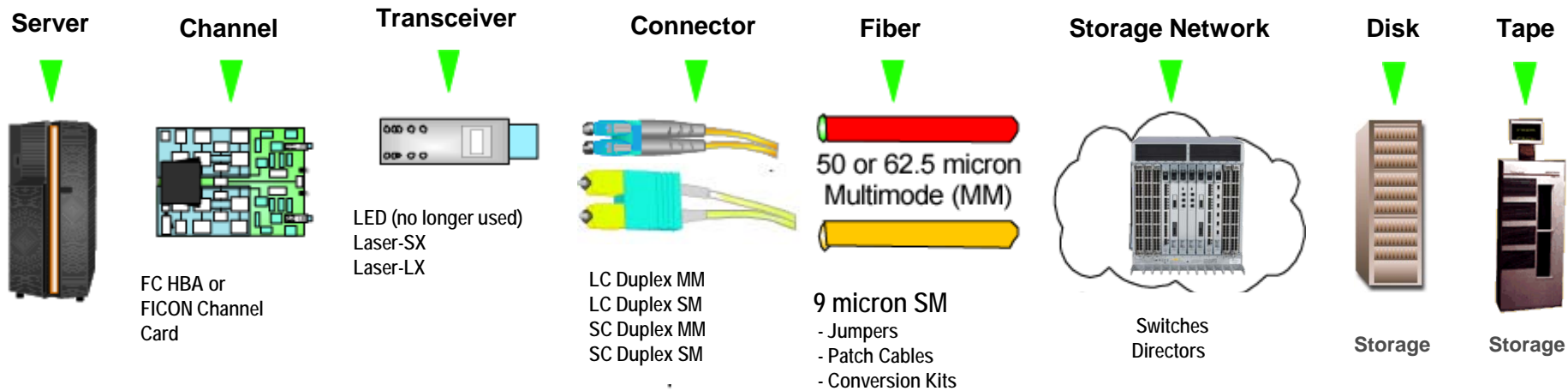
1 to 32 Mainframes

A Parallel Sysplex provides a clustered environment to provide both parallelism for application processing as well as better reliability/availability

Sysplex

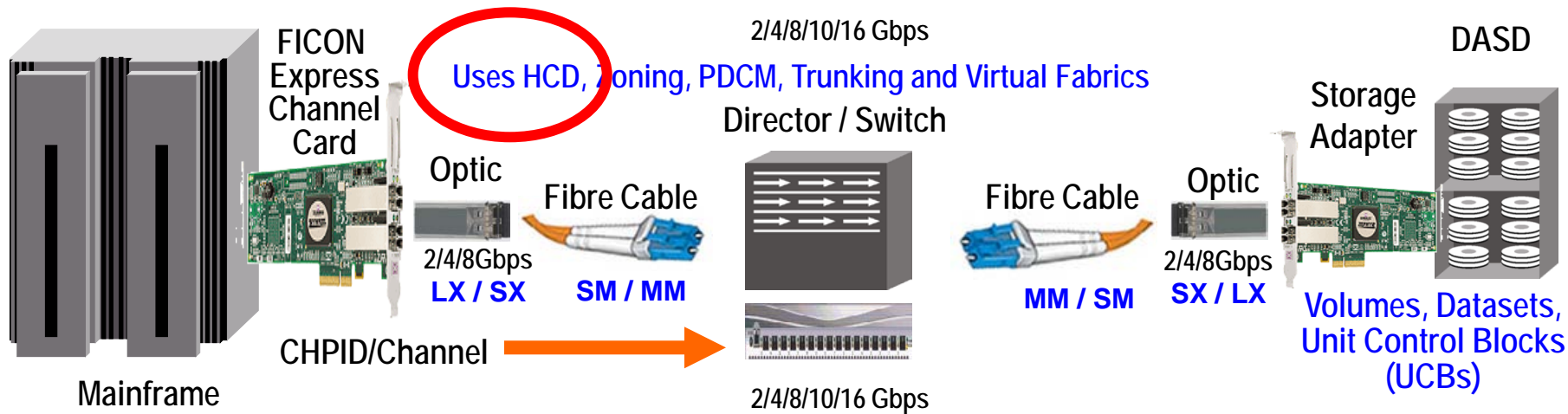
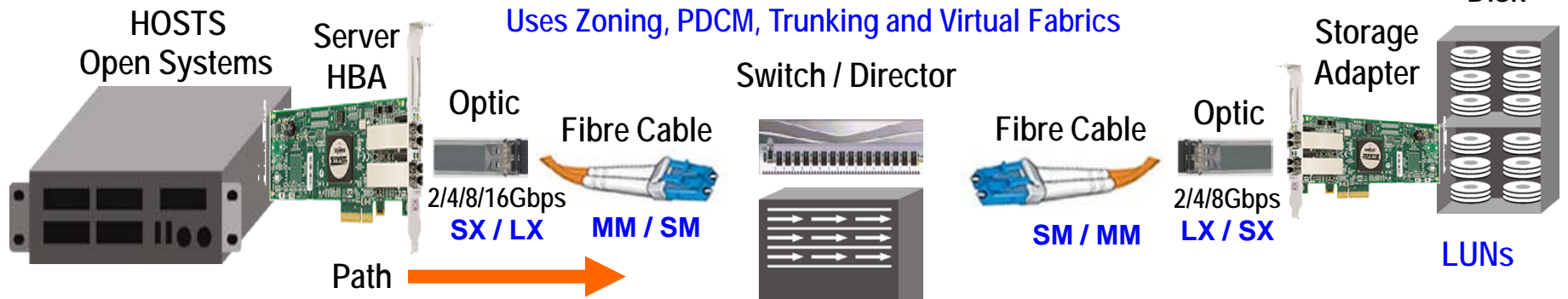
- IBM's **SY**Stems Com**P**LEX was introduced in 1990 as a platform for the MVS/ESA operating system for IBM mainframe servers. The sysplex consists of the computer or multiple computers that make up the Computer Electronics Complex (**CEC**).
- **Parallel Sysplex** is a Sysplex evolution providing a clustering architecture that has improved communication capabilities and supports more connected Central Processing Complexes (**CPCs**) and more copies of the operating system and is often deployed with a **Workload Manager capability**.

Components of a SAN Connection



This is a component summary of how everything works together in order to effectively deploy a fibre channel I/O infrastructure.

Open Systems compared to Mainframe



Multi-mode OM2
50 micron 500mHz
2/4 Gbps

Multi-mode OM3
50 micron 2000mHz
2/4/8/10/16Gbps

Multi-mode OM4
50 micron 4700mHz
2/4/8/10/16Gbps

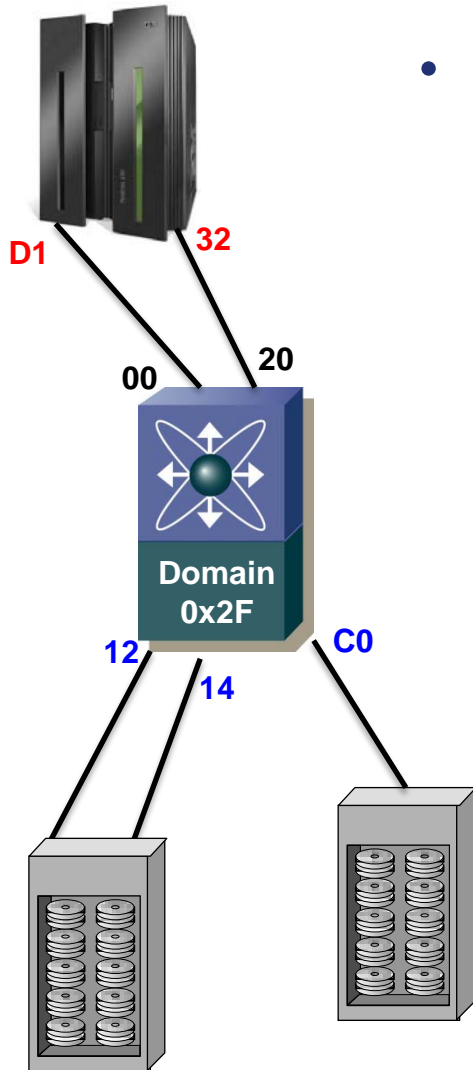
Single-mode OS1
9 micron
2/4/8/10/16Gbps ~ 10k

Agenda for Session 10999 and 11000

Session 11000

- History
- Terminology, Connectors, Cables, and Wavelengths
- **Addressing in FICON**
- ESCON Status, zHPF and NPIV
- Buffer Credits, CUP, RMF

FICON Traffic Routing – Single Switch

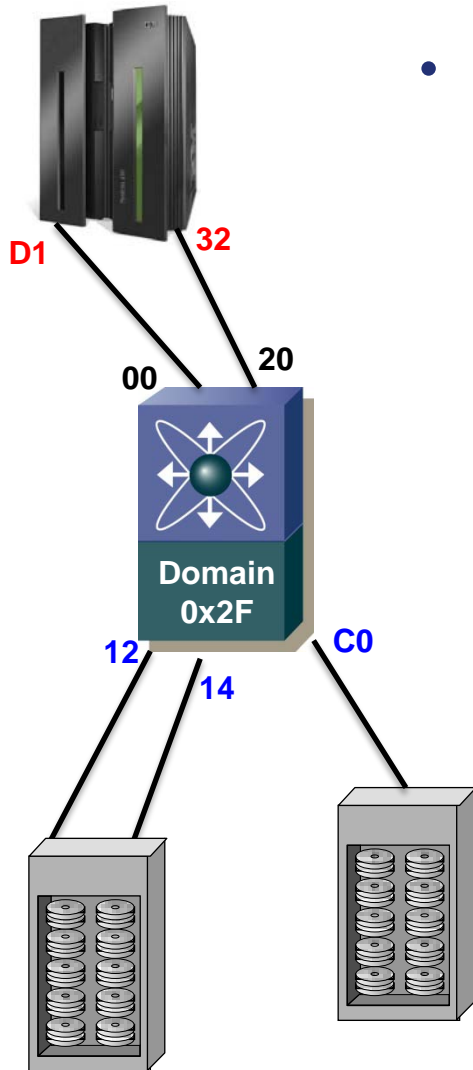


- Mainframe uses static FCID-based routing for devices
 - CHPIDs 32 and D1 are connected to a single switch
 - Switch has statically defined domain of 0x2F
 - Host IOCDS or HCD defines route to devices
 - Specifying the CHPID(s) and corresponding output port

```

CHPID PATH= (CSS (0) , D1) , SHARED , *
PARTITION= ( (LPARMVSY) , (LPARMVSX , LPARMVSA , LPARMVSB , VMLPAR*
02) ) , SWITCH=2F , PCHID=1B1 , TYPE=FC
CHPID PATH= (CSS (0) , 32) , SHARED , *
PARTITION= ( (LPARMVSY) , (LPARMVSX , LPARMVSA , LPARMVSB , VMLPAR*
02) ) , SWITCH=2F , PCHID=1B3 , TYPE=FC
    
```

FICON Traffic Routing – Single Switch



- Mainframe uses static FCID-based routing for devices
 - CHPIDs 32 and D1 are connected to a single switch
 - Switch has statically defined domain of 0x2F
 - Host IOCDs or HCD defines route to devices
 - Specifying the CHPID(s) and corresponding output port

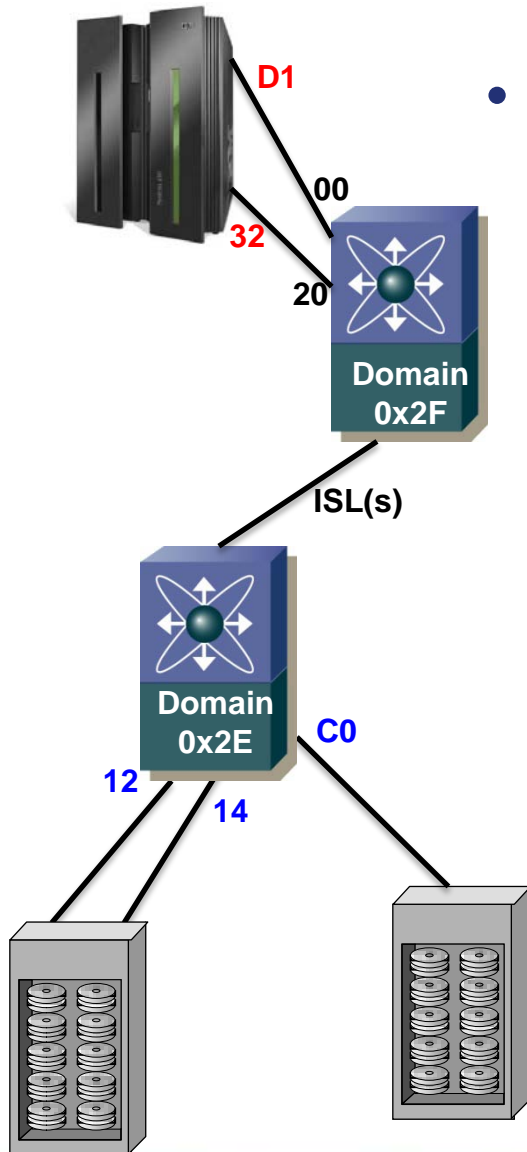
```

CHPID PATH=(CSS(0),D1),SHARED,*
PARTITION=((LPARMVSY),(LPARMVSX,LPARMVSA,LPARMVSB,VMLPAR*
02)),SWITCH=2F,PCHID=1B1,TYPE=FC
CHPID PATH=(CSS(0),32),SHARED,*
PARTITION=((LPARMVSY),(LPARMVSX,LPARMVSA,LPARMVSB,VMLPAR*
02)),SWITCH=2F,PCHID=1B3,TYPE=FC

CNTLUNIT CUNUMBR=0CC0,PATH=((CSS(0),D1,32)),*
LINK=((CSS(0),12,14)),*
UNIT=2105,CUADD=B,UNITADD=((00,032))

CNTLUNIT CUNUMBR=0C00,PATH=((CSS(0),D1)),*
LINK=((CSS(0),C0)),*
UNIT=2105,CUADD=C,UNITADD=((00,032))
  
```

FICON Traffic Routing – Cascade



- Mainframe two switch routing
 - CHPIDs 32 and D1 connected to a host-side switch (0x2F)
 - DASD is attached to a different switch (0x2E)
 - Host IOCDS or HCD defines route to devices
 - Specifying the CHPID(s) and corresponding output domain/port
- Mainframe is “blind” to the the ISL

```

CHPID PATH=(CSS(0),D1),SHARED, *
PARTITION=((LPARMVSY),(LPARMVSX,LPARMVSA,LPARMVSB,VMLPAR*
02)),SWITCH=2F,PCHID=1B1,TYPE=FC
CHPID PATH=(CSS(0),32),SHARED, *
PARTITION=((LPARMVSY),(LPARMVSX,LPARMVSA,LPARMVSB,VMLPAR*
02)),SWITCH=2F,PCHID=1B3,TYPE=FC

CNTLUNIT CUNUMBR=0CC0,PATH=((CSS(0),D1,32)), *
LINK=((CSS(0),2E12,2E14)), *
UNIT=2105,CUADD=B,UNITADD=((00,032))

CNTLUNIT CUNUMBR=0C00,PATH=((CSS(0),D1)), *
LINK=((CSS(0),2EC0)), *
UNIT=2105,CUADD=C,UNITADD=((00,032))
  
```


Agenda for Session 10999 and 11000

Session 11000

- History
- Terminology, Connectors, Cables, and Wavelengths
- Addressing in FICON
- **ESCON Status, zHPF and NPIV**
- Buffer Credits, CUP, RMF

ESCON Statement of Direction

ESCON channels will be phased out

- *It is IBM's intent for ESCON channels to be phased out!*
- 1024 total channels are supported on z10 EC while 480 total channels are supported on the z10 BC ...and...up to 960 of those channels could be ESCON channels if the customer desired it
- But only 240 ESCON channels are supported on z196 and z114
- System z196 and z114 will be the last mainframes to natively support ESCON channels

It is time to move to a FICON Infrastructure!!

...but...

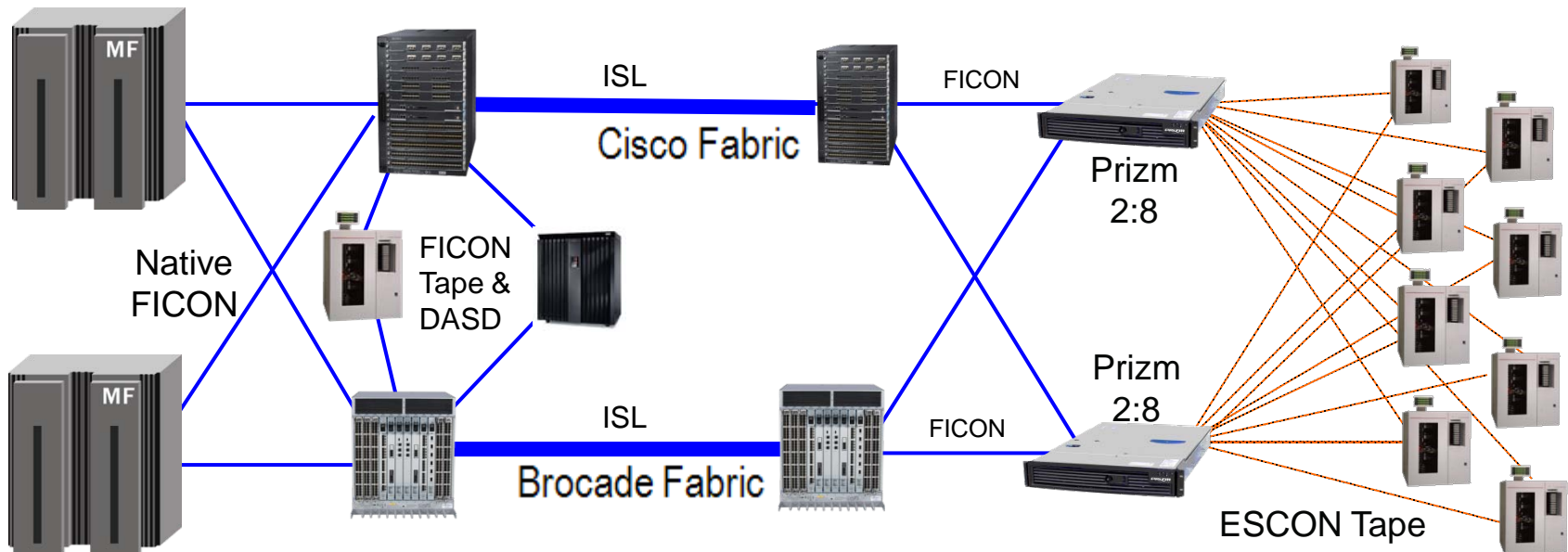
What about the old ESCON devices that you still use?



FICON to ESCON Converter



- Promotes* FICON infrastructure modernization
- Preserves* ESCON infrastructure investments
- Replaces* ESCON directors, FICON bridge, and ESCON extenders

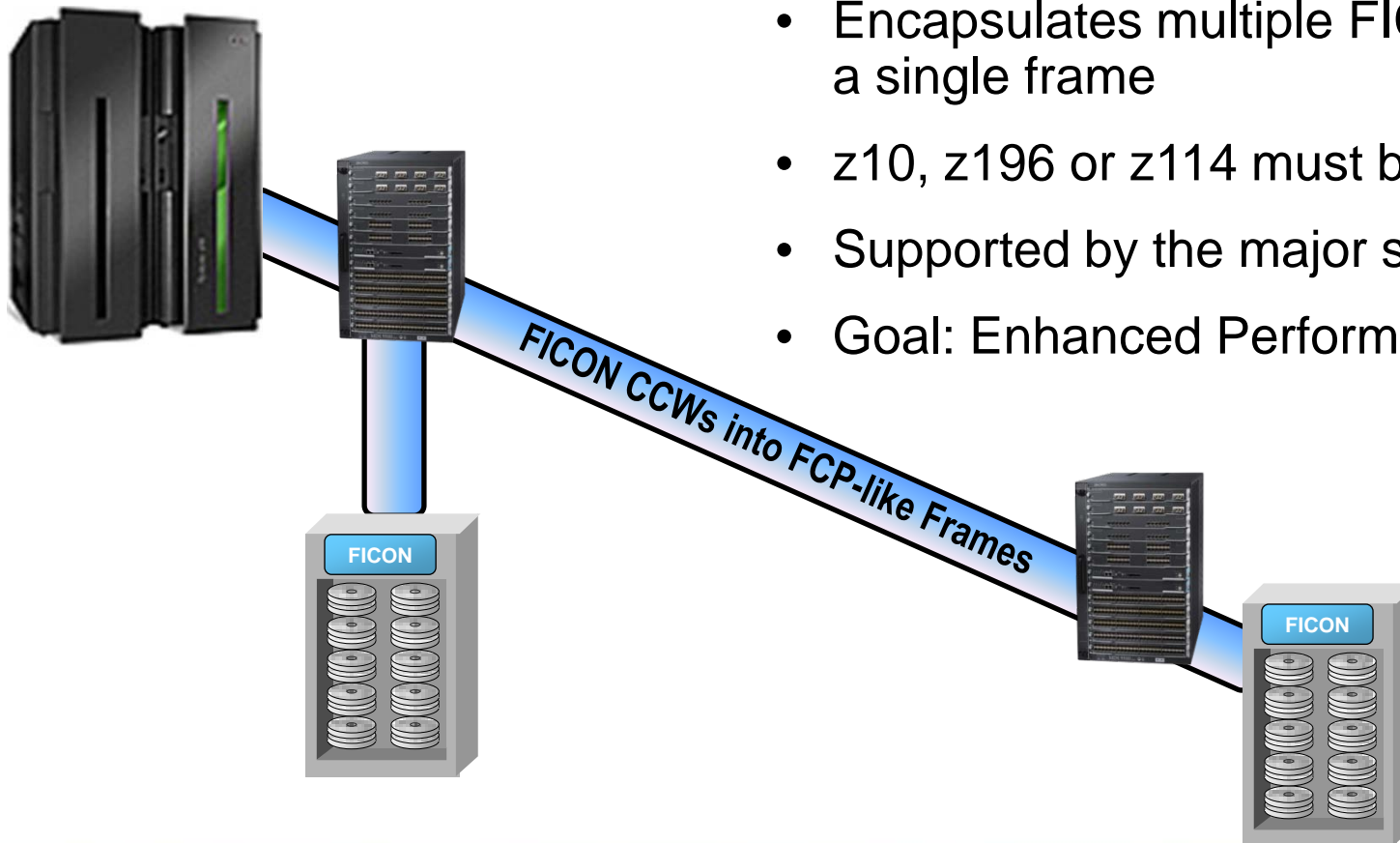


zHPF – High Performance FICON

It was first introduced in 2008

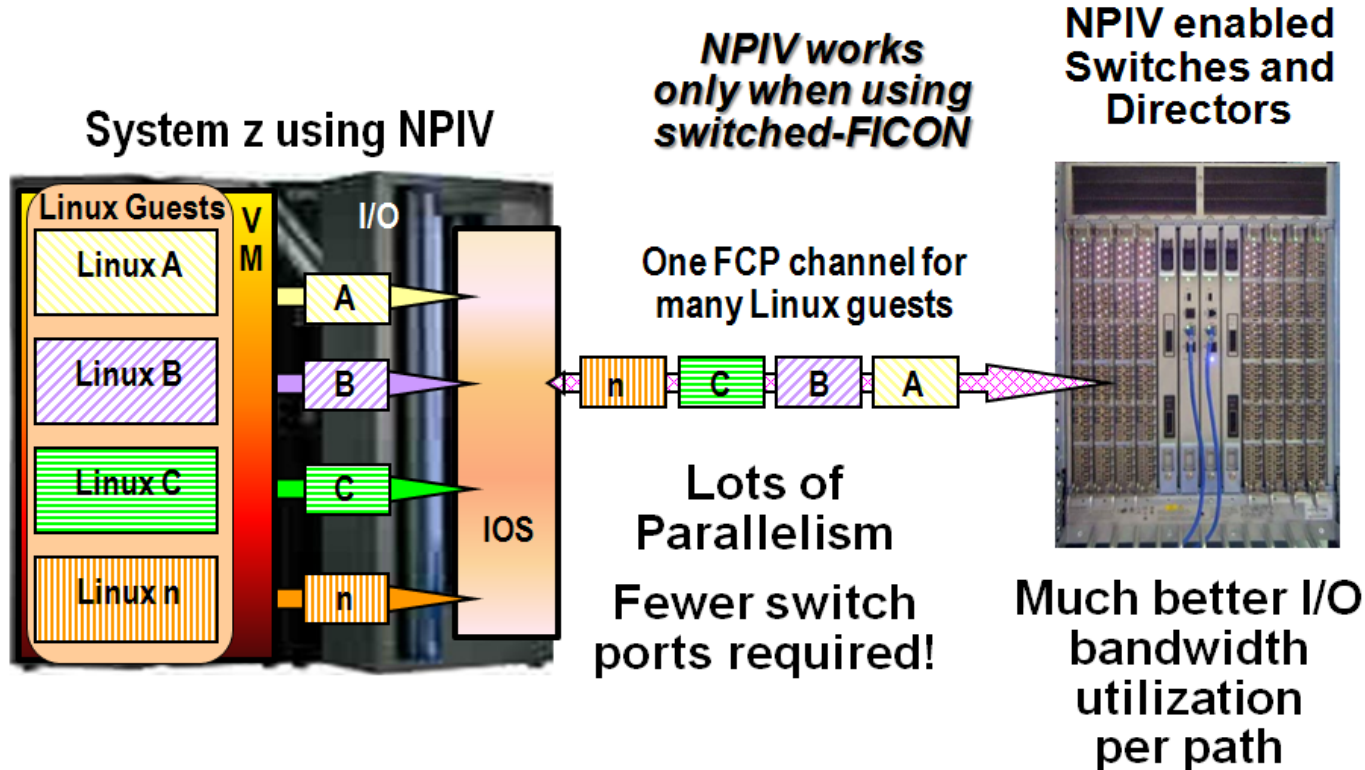
It continues to be enhanced

- Currently used for extended format data sets managed by Media Manager or by EXCP (DASD only – not tape)
- Encapsulates multiple FICON CCWs into a single frame
- z10, z196 or z114 must be the host
- Supported by the major storage vendors
- Goal: Enhanced Performance



NPIV – Node_Port ID Virtualization

- NPIV is standards based
- NPIV is used on the mainframe when Linux utilizes FCP Ports



8 Gbps Is Great For NPIV!

Agenda for Session 10999 and 11000

Session 11000

- History
- Terminology, Connectors, Cables, and Wavelengths
- Addressing in FICON
- ESCON Status, zHPF and NPIV
- **Buffer Credits, CUP, RMF**

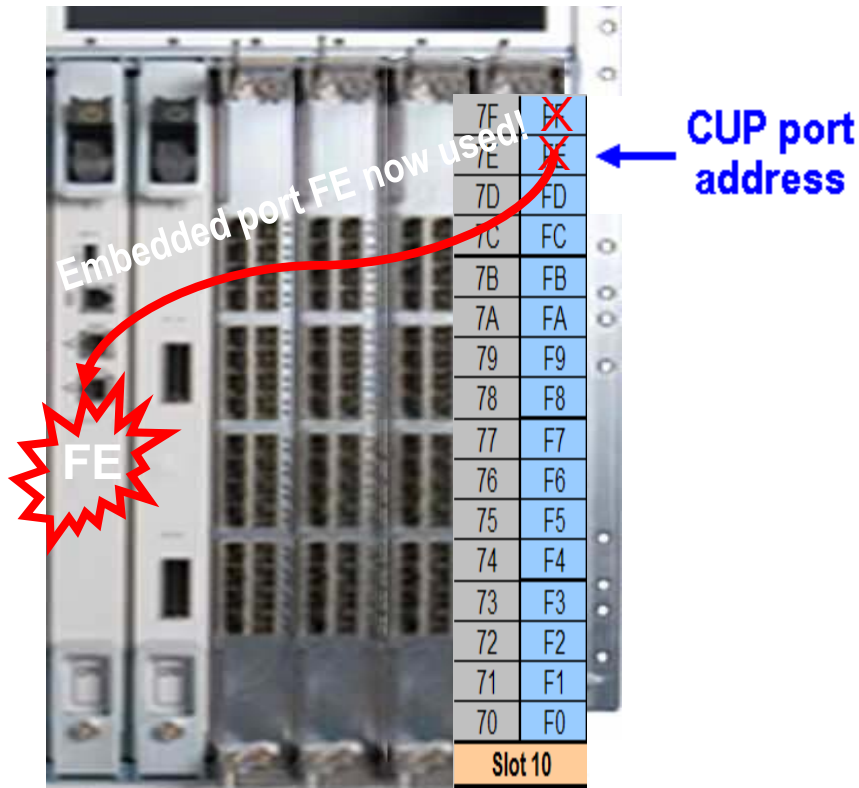
Mainframe Number of Buffer Credits - Reality

Control	76
Write	2048 2048 72
Status Accept	68

- Simple 4K write
- Will not fit into 2 buffers because of headers for FC as well as SB3

$$\text{Average} = (76+2048+2048+72+68) / 5 = 862 \text{ Bytes}$$

Mainframe Terminology



CUP should be deployed on every FICON Director! It allows you to provide the M/F environment with port performance information and buffer credit flow control information. 2 ports could be lost when CUP is used!

Control Unit Port (CUP)

- Customers install FICON Management Server (FMS) as a license on their FICON director.
- FMS enables CUP to be utilized for in-band management and FICON Director performance reporting up to the mainframe environment.
- Port address 0xFE is always the port address exclusively defined for the CUP – but it can also be a physical port which cannot be used for connectivity when CUP is enabled.

FICON Director Activity Report With Frame Delay



Using Buffer Credits is how FC does Flow Control, also called "Frame Pacing"

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

PAGE 1

I O D F = A 2 C R - D A T E : 0 3 / 2 7 / 2 0 0 9 C R - T I M E : 1 8 . 4 3 . 5 1 A C T : A C T I V A T E			S Y S T E M I D A B C D		S T A R T 0 4 / 1 2 / 2 0 0 9 - 0 4 . 3 0 . 0 0		I N T E R V A L 0 0 0 . 1 5 . 0 0	
S W I T C H D E V I C E : 0 3 2 B S W I T C H I D : 2 B			R P T V E R S I O N V 1 R 8 R M F		E N D 0 4 / 1 2 / 2 0 0 9 - 0 4 . 4 5 . 0 0		C Y C L E 1 . 0 0 0 S E C O N D S	
S W I T C H D E V I C E : 0 3 2 B S W I T C H I D : 2 B T Y P E : 0 0 6 1 4 0			M O D E L : 0 0 1 M A N : M C D P L A N T : 0 1		S E R I A L : 0 0 0 0 0 0 H I J K L M N			
PORT	-CONNECTION-	ID	A V G F R A M E P A C I N G	A V G F R A M E S I Z E	P O R T B A N D W I D T H (M B / S E C)		E R R O R C O U N T	
ADDR	UNIT			READ	WRITE	-- READ --	-- WRITE --	
05	CHP-H	05	0	849	1436	8.63	17.34	0
07	CHP	6B	1	1681	1395	50.87	10.32	0
09	CHP	15	0	833	1429	11.96	20.49	0
0C	CHP-H	64	0	939	1099	0.39	0.50	0
0D	CHP	6B	0	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	CU	C800	0	1241	738	20.97	5.72	0
	CU	CA00				70.10	3.82	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	20.80	7.30	0
24	CHP	64	0	89	1345	0.00	0.00	0
27	CHP	6B	0	1619	82	0.01	0.00	0
28	SWITCH	95	270	550	789	50.32	10.56	0
2B	CHP	70	0	69	2022	0.00	0.71	0

In the last 15 minutes



This port had a frame to send but did not have any Buffer Credits left to use to send them.

And this happened 270 times during the interval.

And this is an ISL Link! Indicators of Buffer Credit Starvation

Fabric with zHPF Enabled



SAN Sessions at SHARE this week



Monday:

Time-Session

1300 - 11021: Buffer-to-Buffer Credits, Exchanges, and Urban Legends

Tuesday:

Time-Session

1330 - 11003: A First Look at the Inner Workings and Hidden Mechanisms of FICON Performance

1500 - 11004: A deeper look into the Inner Workings and Hidden Mechanisms of FICON Performance

1800 - 11022: FICON Director and Channel Free-For-All

Thursday:

Time-Session

1630 - 11027: Evolution of the System z Channel

Visit Brocade's Mainframe Blog Page at:

<http://community.brocade.com/community/brocadeblogs/mainframe>

Also Visit Brocade's New Mainframe Communities Page at:

http://community.brocade.com/community/forums/products_and_solutions/mainframe_solutions



THANK YOU FOR ATTENDING!

Session 11000

Please fill out your evaluation forms



Questions and Discussion

