Session 9857
Introduction to Storage Technologies
SAN (Storage Area Networking)
FICON (Fiber CONnection)

David Lytle - Brocade
dlytle@brocade.com

Mike Blair - Cisco
mblair@cisco.com

SAN 101
Agenda for Session’s 9857 and 10079

1st Session 9857
- History
- Terminology, Connectors, Cables, and Wavelengths
- ESCON Status, zHPF and NPIV

2nd Session 10079
- Types and Components of Storage
- Let’s talk Fibre Channel
- FC Buffer Credits
- Fabric Routing / Virtual Fabrics / Partitioning
- Security / Zoning

Brocade, Cisco and IBM Graphics are used throughout this presentation.
Agenda for Session 9857 and 10079

Session 9857
• History
• Terminology, Connectors, Cables, and Wavelengths
• ESCON Status, zHPF and NPIV
Mainframe And Open System Time Lines

Material based on “I/O Architecture & Performance Analysis” class developed by IntelliMagic
Agenda for Session 9857 and 10079

Session 9857
- History
- Terminology, Connectors, Cables, and Wavelengths
- ESCON Status, zHPF and NPIV
SAN Terminology

• “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

• Variable frame size with a maximum user payload of 2112 bytes. Supports block transfers up to 128MB in size.

• Supports full duplex transmission (simultaneous send and receive)
SAN Terminology

- Light wavelengths in fiber are expressed in nanometers.
- The speed of light in fiber cable is 2/3 the speed of light in a vacuum:
  - Light travels at ~5nsec per meter 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).
  - This is used for local distance connectivity (~50-500 m).
- **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 10km).
- Optical power budgets, or link loss budgets, measured in decibels (dBs), are used to manage optical signal loss.
**SAN Terminology**

- **Multimode Fiber**
  - For local distances and/or 1/2/4/8Gbps
  - MM – use with short wave optics
  - 62.5 micron, 50 micron
  - (62.5m has little value today)

- **Single-mode Fiber**
  - For long distances and/or 8Gbps+ link rate
  - SM – use with long wave optics
  - 9 micron
SAN Terminology

**Fibre 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 – i.e. 4km/10km
  - Short wave laser
  - LED
  - Electrical
SAN Terminology

**OPTICAL**

- **LC Optical Connector**
  - Standard on 2-8Gbps Switches
  - Most widely used connector
  - *Cable Shown as Bonded Duplex*

- **SC Optical Connector**
  - Standard on 1 Gbps Switches
  - Little used any longer
  - *Cable Shown as Bonded Duplex*

**COPPER**

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

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

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

Open Systems

[Diagram showing connections between N_Port, F_Port, G_Port, and E_Port]
SAN Terminology

- The interconnection between switches is called the Inter-Switch Link (ISL) or in FICON a Cascaded Link
  - E_Port to E_Port (so it is also known as an E_Port)
  - 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
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

Channel Card Channel Card

8G

4G

Mainframe Terminology

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 ExpressX where X is 2, 4 or 8

Up to 336 FICON CHPIDs on z9 / z10
Up to 288 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
Logical Partitions (LPARs)

- Available since 1988 this is a way of dividing up a mainframe’s capacity into Logical PARtitions 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.

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

Diagram:
- System z
- ESCON CHPID
- Point-to-Point ESCON
- Unidirectional data flow only
- Half duplex – about 3-12MBps

9032 ESCON Director
Models 3 and 5 were HA models
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
- 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:
- Relatively new (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-100 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 (Fiber CONnection)?

- 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
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 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).
- As of 2010, on z196, z/OS will also directly connect and manage an IBM BladeCenter with Power 7 blades.
Mainframe Terminology

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.

System z Mainframe (M/F)

TPF usually runs on its own mainframe to help insure availability but it can run with other OS's.

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

System z Mainframe (M/F)

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.

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.
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 – twelve years ago!
Mainframe Terminology

Hardware Configuration Definition (HCD)
Define the mainframe’s computing and I/O environment

Example FICON IOCP

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

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

IODEVICE
```

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.
IBM's **Systems Complex (Sysplex)** 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.
Mainframe Terminology

**Geographically Dispersed Parallel Sysplex (GDPS)**

- **Parallel Sysplex**
  - Addresses Planned / Unplanned Hardware / Software Outages
  - Flexible, Non-disruptive Growth
    - Capacity beyond largest CEC
    - Scales better than SMPs
  - Dynamic Workload/Resource Management

- **GDPS**
  - Addresses Planned / Unplanned Hardware / Software Outages
  - Addresses Site Failure and Site Maintenance
  - Robust DASD/Tape Data Replication
    - Eliminates SPOF
    - Zero / Very Little Data Loss
  - Application Independent

1 to 32 Systems
What is a GDPS Solution?

• GDPS is:
  • Services offering
  • Software
  • Automation code

• Are FICON Switching Devices GDPS qualified?
  • Not automatically!
    ▪ There is a GDPS qualification beyond FICON qualification
  • Most, if not all, FICON Directors available today are GDPS qualified

• A GDPS Solution includes:
  • Two (or more) data center infrastructures
  • Plus data replication technology (PPRC, Global Mirror z/OS XRC)
  • Plus application replication technology (Parallel Sysplex)
  • Plus network infrastructure (for data replication & any-to-any connectivity)

• Note about Sysplex Protocols
  • ISC3/STP do not run across FICON Directors
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

**Open Systems**
- Hosts
- Server HBA
- Optic 2/4/8Gbps SX / LX
- Fibre Cable MM / SM
- Switch / Director
- Optic 2/4/8/10Gbps
- Uses Zoning, PDCM, Trunking and Virtual Fabrics

**Mainframe**
- CHPID/Channel
- Path
- Optic 2/4/8Gbps LX / SX
- Fibre Cable SM / MM
- Director / Switch
- Optic 2/4/8/10Gbps
- Uses HCD, Zoning, PDCM, Trunking and Virtual Fabrics

**Storage Adapter**
- Storage Adapter
- LUNs
- DASD
- Volumes, Datasets, Unit Control Blocks (UCBs)

**Optical Cables**
- Multi-mode OM2 50 micron 500mHz 2/4 Gbps
- Multi-mode OM3 50 micron 2000mHz 2/4/8/10Gbps
- Multi-mode OM4 50 micron 4700mHz 2/4/8/10Gbps
- Single-mode OS1 9 micron 2/4/8/10Gbps ~ 10k
Agenda for Session 9857 and 10079

Session 9857

- History
- Terminology, Connectors, Cables, and Wavelengths
- ESCON Status, zHPF and NPIV
ESCON Statement of Direction

ESCON channels will be phased out

- It is IBM's intent for ESCON channels to be phased out. System z10 EC and System z10 BC will be the last server to support greater than 240 ESCON channels.

- Released April 28, 2009

- Currently, 1024 channels are supported on z10 EC and 480 channels are supported in the z10 BC

- Only 240 ESCON channels are supported on z196 …AND… z196 and z114 will be the last mainframes to natively support ESCON channels
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 with Media Manager, extended format data sets
- Encapsulates multiple FICON CCWs into a single frame
- z10, z196 or z114 must be the host
- Supported by the major storage vendors
- Enhanced Performance

FICON CCWs into FCP-like Frames
NPIV – Node_Port ID Virtualization

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

System z using NPIV

8 Gbps Is Great For NPIV!
This Is The End Of Part 1

The 2\textsuperscript{nd} of this 2-part session is

Session 10079

Monday: 11:00AM – 12:00PM
SAN Sessions at SHARE this week

Monday:
Time-Session
0930 - 9857: SAN (Storage Area Networking) Overview – Part 1
1100 - 10079: SAN (Storage Area Networking) Overview - Part 2
1330 - 9671: Managing SAN for Linux on z/VM - a Nationwide Perspective
1500 - 9931: Buffer-to-Buffer Credits, Exchanges, and Urban Legends
1630 - 9934: Evolution of the System z Channel

Tuesday:
Time-Session
1100 - 9368: A First Look at the Inner Workings and Hidden Mechanisms of FICON Performance
1500 - 10079: A deeper look into the Inner Workings and Hidden Mechanisms of FICON Performance

Wednesday:
Time-Session
0800 - 9479: Planning and Implementing NPIV for System Z
0930 - 9864: zSeries FICON and FCP Fabrics - Intermixing Best Practices

Thursday:
Time-Session
0800 - 9853: FICON Over IP - Technology and Customer Use
0800 - 9899: Planning for ESCON Elimination
0930 - 9933: Customer Deployment Examples for FICON Technologies
1500 - 9316: SAN Security Overview
1630 - 10088: FICON Director and Channel Free-for-all
THANK YOU FOR ATTENDING!

Session 9857

Please fill out your evaluation forms
Session 10079

Introduction to Storage Technologies
SAN (Storage Area Networking)
FICON (Fiber CONnection)

Mike Blair - Cisco
mblair@cisco.com

David Lytle - Brocade
dlytle@brocade.com
Agenda for Session 9857 and 10079

Session 9857
- History
- Terminology, Connectors, Cables, and Wavelengths
- ESCON Status, zHPF and NPIV

Session 10079
- Types and Components of Storage
- Let’s talk Fibre Channel
- FC Buffer Credits
- Fabric Routing / Virtual Fabrics / Partitioning
- Security / Zoning
Agenda for Session 9857 and 10079

Session 10079

• Types and Components of Storage
• Let’s talk Fibre Channel
• FC Buffer Credits
• Fabric Routing / Virtual Fabrics / Partitioning
• Security / Zoning
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
Agenda for Session 9857 and 10079

Session 10079

- 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 or 800 MBps …and… 10Gb ISLs
  - 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

**FC-4**
- Protocol Mapping Layer
  - Upper Level Protocol (ULP)
  - FCP/FICON/HIPPI/Multi-media, etc.

**FC-3**
- Common Services
  - Login Server, Name Server, Alias Server

**FC-2**
- Framing Protocol / Flow Control
  - Data packaging, Class of service, Port Login / logout, Flow control...

**FC-1**
- Transmission Protocol - Encode / Decode
  - Serial Interface (one bit after another)
  - Frame Transfer (up to 2048 byte payload)
  - 8b/10b or 64b/66b data encode / decode

**FC-0**
- Interface/Media – The Physical Characteristics
  - Cables, Connectors, Transmitters & Receivers...
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

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

- The 24 bit 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

<table>
<thead>
<tr>
<th>Switch/Open Topology</th>
<th>8 bits</th>
<th>8 bits</th>
<th>8 bits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Switch Domain</td>
<td>Area</td>
<td>Device</td>
</tr>
<tr>
<td>FICON Topology</td>
<td>Switch Domain</td>
<td>Port Address</td>
<td>0 ..... or NPIV virtual addr.</td>
</tr>
</tbody>
</table>
Fibre Channel Frame Format

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

General FC-2 Frame Format

- Idles*
- SOF
- Frame Header
- Data Field
- CRC
- EOF
- Idles*

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

0-528 Transmission Word
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 <=50 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 and FCP

- 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
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)
Agenda for Session 9857 and 10079

Session 10079

- 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.
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.

Frame towards Disk shelf
Return BB_Credit token
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)

Frame towards Disk shelf
Frame towards Tape
Return BB_Credit token
FC BB_Credits and Distance

- BB_Credits are used to ensure enough FC frames in flight
- A full (2112 byte) FC frame is approximately 2 km long @ 1 Gbps and approximately 1 km long @ 2 Gbps and approximately .5 km long @ 4 Gbps
- As distance increases, the BB_Credits need to increase as well
- Shorter frames will require additional BB_Credits to ensure a “full pipe” – and **USUALLY** it is smaller frames that are sent!
- Insufficient BB_Credits will throttle performance—no data will be transmitted until R_RDY is returned
Number of Buffer Credits - Reality

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

Average = (76+2048+2048+72+68) / 5 = 862 Bytes
# Buffer Credits Required

By Size of Frame and Link Speed

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

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

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.

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!
# FICON Director Activity Report (RMF 74-7)

(requires CUP Code)

### FICON DIRECTOR ACTIVITY

<table>
<thead>
<tr>
<th>UNIT</th>
<th>ID</th>
<th>PACING</th>
<th>READ</th>
<th>AVG READ TIME</th>
<th>AVG WRITE TIME</th>
<th>DTH (MB/SEC)</th>
<th>WRITE</th>
<th>ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>05</td>
<td>CHP-H</td>
<td>47</td>
<td>0</td>
<td>433</td>
<td>1.52</td>
<td>5.81</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>06</td>
<td>CU</td>
<td></td>
<td>0</td>
<td>830</td>
<td>5.16</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>07</td>
<td>CU</td>
<td></td>
<td>0</td>
<td>140</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>08</td>
<td>CHP-H</td>
<td>45</td>
<td>0</td>
<td>400</td>
<td>1.55</td>
<td>1.00</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Additional Data

- Start: 10/03/2005-13.55.00
- End: 10/03/2005-14.00.00
- Interval: 000.05.00
- Cycle: 1.000 SECONDS

**Notes:**
- MCD
- PLANT: 01
- SERIAL: 0000013A2
Agenda for Session 9857 and 10079

Session 10079

- 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 the Fabric to 239 when FSPF is supported
    • Each Cisco VSAN is treated as a separate Fabric
    • Each Brocade Virtual Fabric is treated as a separate fabric
    • Each Brocade Physical Partition is treated as a separate fabric

• FSPF performs hop-by-hop routing
  • Each Cisco VSAN runs its 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 use FSPF for ISL routing

• FSPF supports hierarchical path selection
  • Provides the scalable routing tables in large topologies
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
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),(LPARMVX,LPARMVSA,LPARMVSB,VMLPAR*02)),SWITCH=2F,PCHID=1B1,TYPE=FC
CHPID PATH=(CSS(0),32),SHARED,*
PARTITION=((LPARMVSY),(LPARMVX,LPARMVSA,LPARMVSB,VMLPAR*02)),SWITCH=2F,PCHID=1B3,TYPE=FC
```
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),(LPARMV SX,LPARMVSA,LPARMVSB,VMLPAR* 02)),SWITCH=2F,PCHID=1B1,TYPE=FC
CHPID PATH=(CSS(0),32),SHARED, * PARTITION=((LPARMVSY),(LPARMV SX,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 are 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 ISL
FICON Traffic Routing – Cascade

- Mainframe two switch routing
  - CHPIDs 32 and D1 are 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 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
FICON Traffic Routing – Cascade

- **Mainframe two switch routing**
  - CHPIDs 32 and D1 are 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 ISL

```plaintext
CHPID PATH=(CSS(0),D1),SHARED,          *  
  PARTITION=((LPARMVSY),(LPARMVSA,LPARMVSB,VMLPAR*  
  02)),SWITCH=2F,PCHID=1B1,TYPE=FC
CHPID PATH=(CSS(0),32),SHARED,           *  
  PARTITION=((LPARMVSY),(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))
```
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
Virtual Fabrics (VF) for DCX Family

VF is an OPTIONAL feature on 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 a DCX

DCX-4S never requires VF

Creating a Logical Switch, which enables Virtual Fabric, requires FOS 6.2+

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 9857 and 10079

Session 10079

• 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

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
0930 - 9857: SAN (Storage Area Networking) Overview – Part 1
1100 - 10079: SAN (Storage Area Networking) Overview - Part 2
1330 - 9671: Managing SAN for Linux on z/VM - a Nationwide Perspective
1500 - 9931: Buffer-to-Buffer Credits, Exchanges, and Urban Legends
1630 - 9934: Evolution of the System z Channel

Tuesday:
Time-Session
1100 - 9368: A First Look at the Inner Workings and Hidden Mechanisms of FICON Performance
1500 - 10079: A deeper look into the Inner Workings and Hidden Mechanisms of FICON Performance

Wednesday:
Time-Session
0800 - 9479: Planning and Implementing NPIV for System Z
0930 - 9864: zSeries FICON and FCP Fabrics - Intermixing Best Practices

Thursday:
Time-Session
0800 - 9853: FICON Over IP - Technology and Customer Use
0800 - 9899: Planning for ESCON Elimination
0930 - 9933: Customer Deployment Examples for FICON Technologies
1500 - 9316: SAN Security Overview
1630 - 10088: FICON Director and Channel Free-for-all
Questions and Discussion
THANK YOU FOR ATTENDING!

Session 10079

Please fill out your evaluation forms