Integrating Open Systems into Mainframe Fabrics

Howard L. Johnson
Brocade

24 August 2009 (4:30pm – 5:30pm)
Session 6201

Abstract

- Focuses on the opportunities and issues to consider when intermixing FICON fabrics with open systems such as Linux on System z. Key topics include how the System z uses new techniques of N_Port_ID Virtualization (NPIV), Virtual Switches in Virtual Fabrics to support the Linux on System z under changing situations and explore these storage networking virtualization techniques in a practical application.

Agenda

- Just for Fun
  - Jargon
  - Virtual Machine Bits
- Virtualization
  - Server
  - Fabric
  - Link

Jargon and VM

JUST FOR FUN

Jargon

- Abend
  - Abnormal termination
- Bit
  - Binary digit
- CP
  - Control program
- DASD Farm
  - A huge amount of DASD managed as a single resource
- Dinosaur
  - Old minicomputers and mainframes
- Dinosaur mating
  - Big-iron mergers of the 1960's and 1970's
- Seven dwarves
  - Burroughs, Control Data, General Electric, Honeywell, NCR, RCA, and Univac
- BUNCH
  - Burroughs, Univac, NCR, Control Data, and Honeywell

Virtual Machine Bits

- Hardware and Software
  - An isolated duplicate of a real machine
  - Trapped in it's own virtual world
- History
  - 11/1961 - MIT CTSS (compatible time sharing system)
  - 12/1964 – Cambridge CP-40
  - 8/1972 - "IBM announced its first official VM product (then called VM/370) on August 2, 1972"
  - 37-years and it always sounds new ;-)
- Further Reading
  - http://www.vm.ibm.com/history/
  - http://www.vm.ibm.com/birthday.html
  - http://www.vm.ibm.com/
  - http://www.princeton.edu/~melinda/
  - http://en.wikipedia.org/wiki/VM_(operating_system)
  - http://en.wikipedia.org/wiki/Virtual_machine
  - http://www.cs.gmu.edu/cne/itcore/virtualmachine/ibm.htm

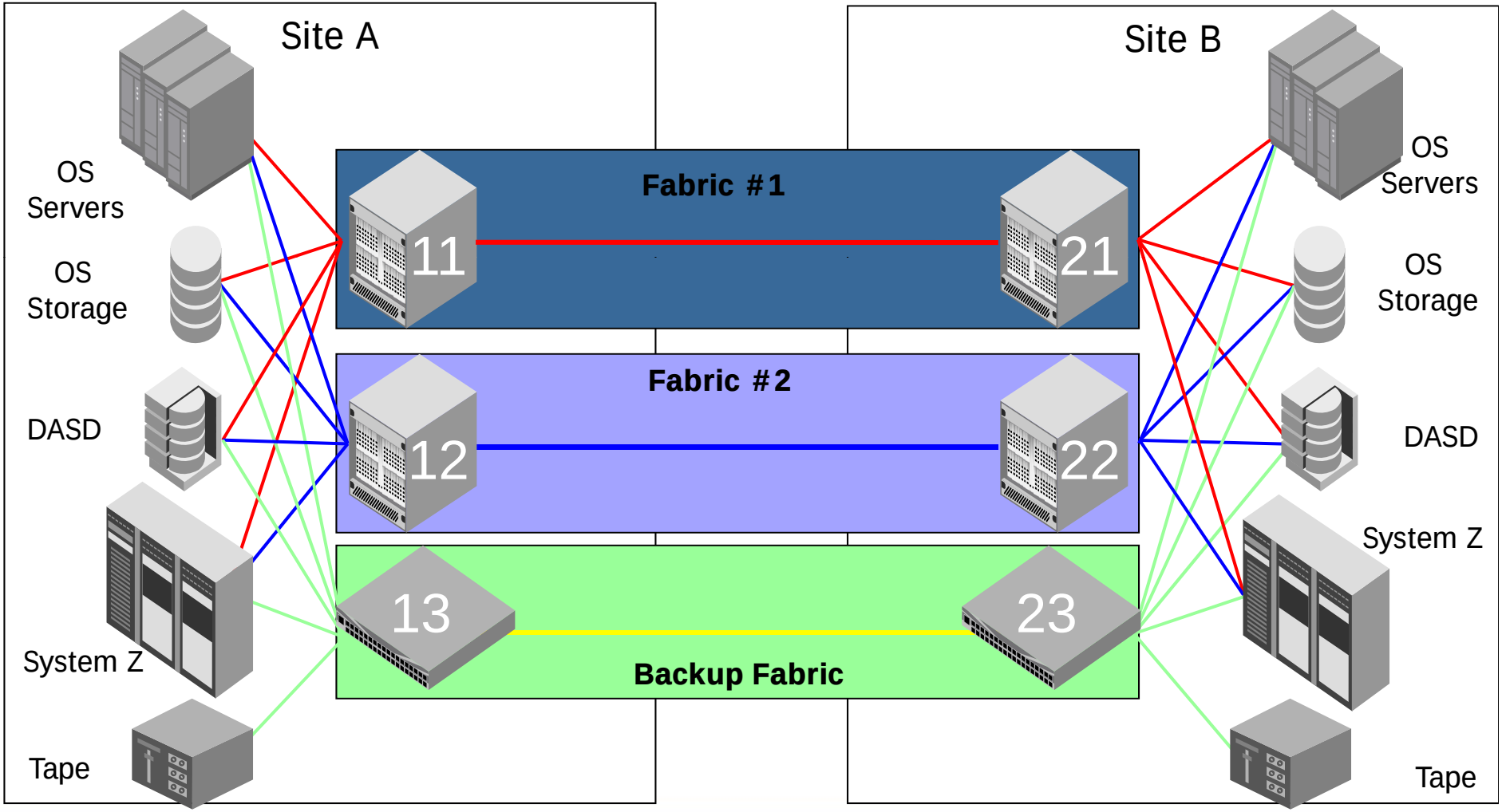
Server, Fabric, and Link Virtualization

VIRTUALIZATION

Example:
Consolidation through Virtualization

- Applications
  - Dual production fabrics
    - Four 64-port Directors
    - Cascaded across two sites
  - Backup fabric
    - Two 24-port Switches
    - Cascaded across two sites
- Devices
  - Open System Servers and Storage
  - Mainframe Server and Storage

Example:
Independent, Mixed Use Fabrics



NPIV

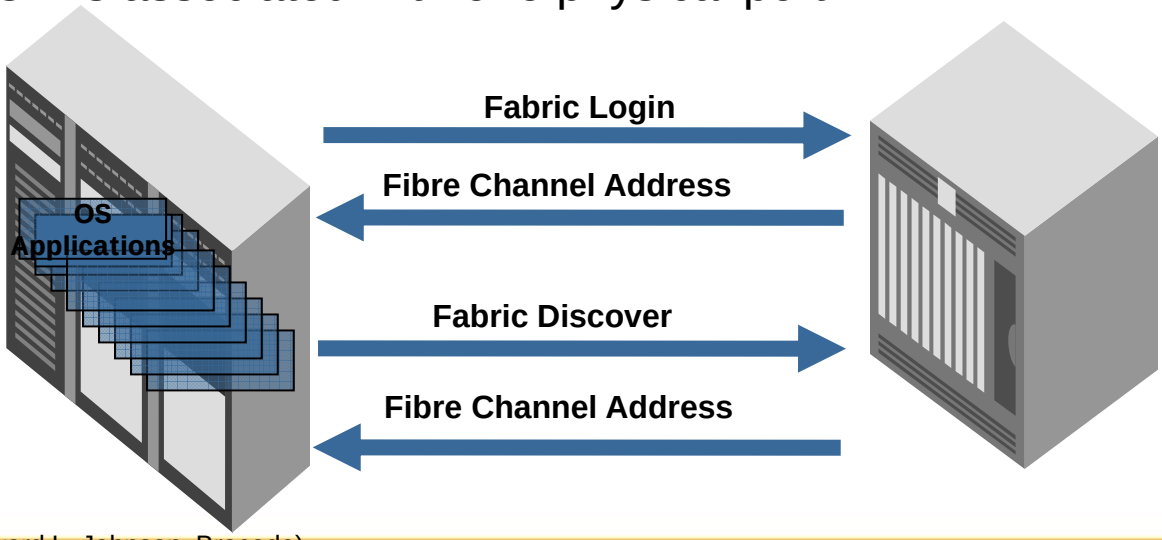
SERVER VIRTUALIZATION

zSeries Server Virtualization

- zSeries support of zLinux
  - Mainframe expanded to address open system applications
  - Linux promoted as alternative to Unix
  - Mainframe OS virtualization benefits
    - Availability, serviceability, scalability, flexibility
- Initial constraints
  - FCP requests are serialized by the OS
    - FCP header doesn't provide image addresses
    - FICON SB2 header provides additional addressing
  - Channel ports are underutilized
  - Resulting cost / performance benefit is not competitive

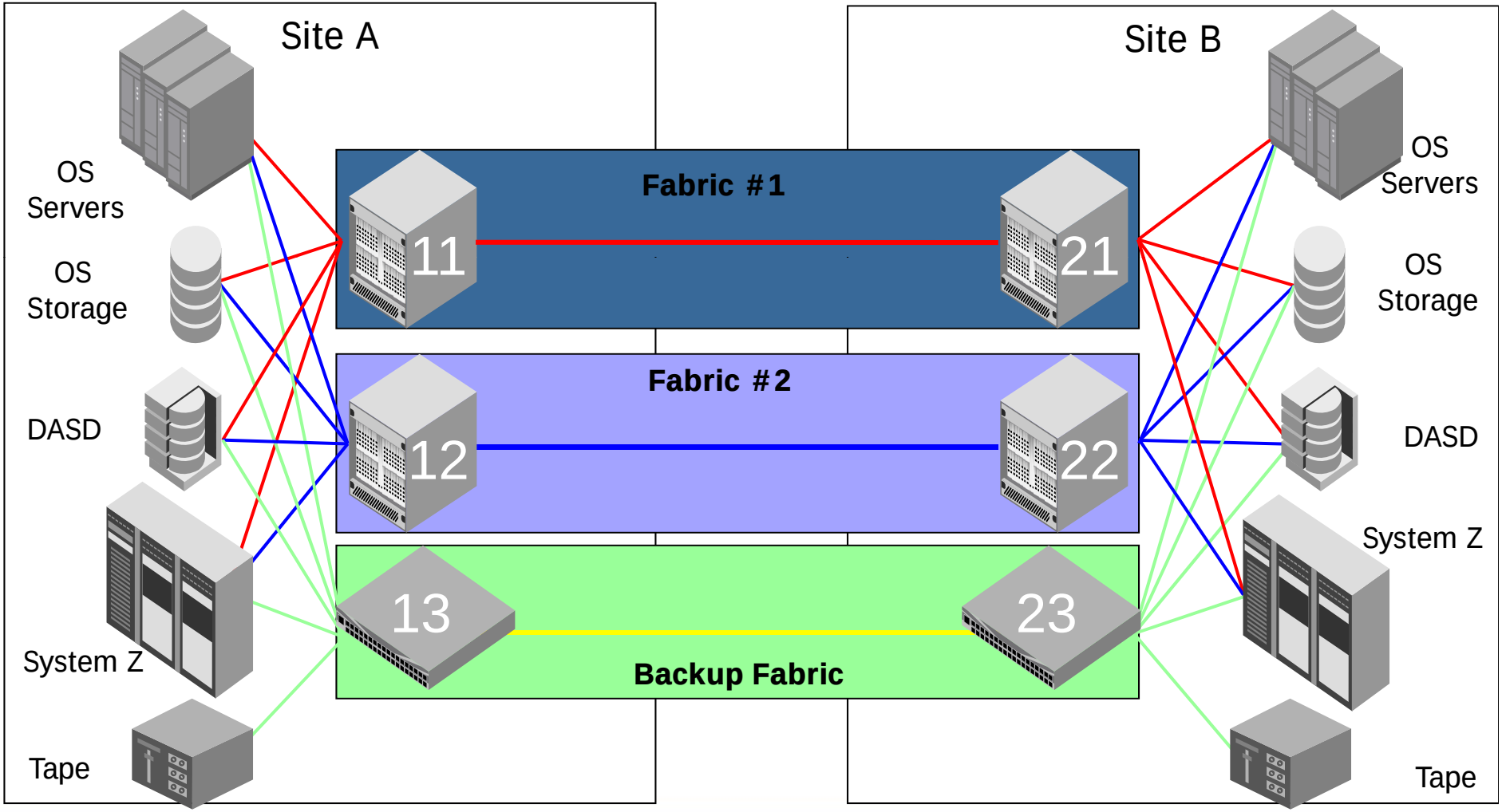
Server Consolidation
Technology

- N_Port Identifier Virtualization (NPIV)
  - N_Port becomes virtualized
    - Supports multiple images behind a single N_Port
  - N_Port requests more than one FCID
    - FLOGI provides first address
    - FDISC provides additional addresses
  - All FCID's associated with one physical port



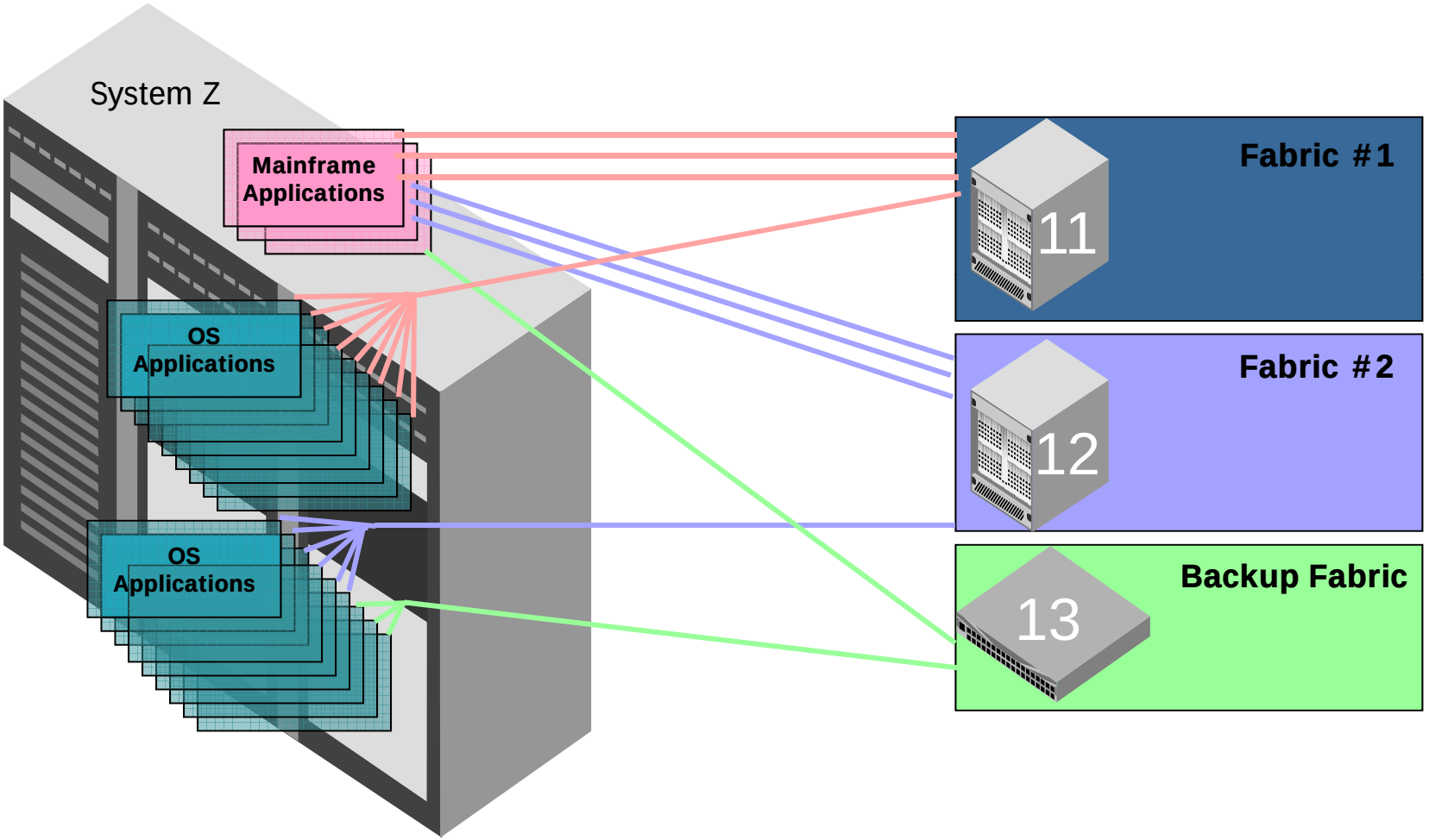
Data Center Network

Independent Cascaded Fabrics



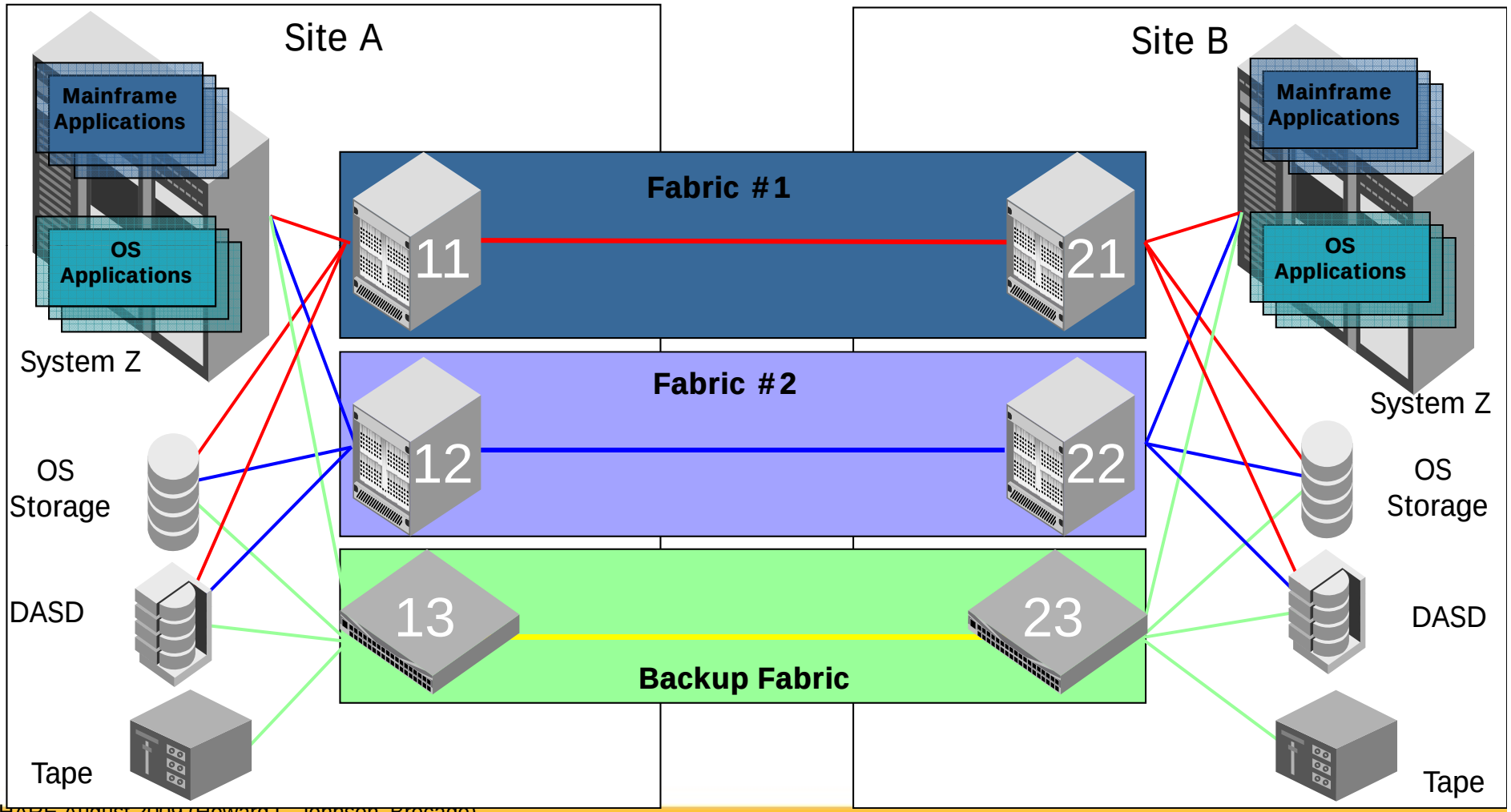
Server Consolidation

N_Port Identifier Virtualization



Consolidated Servers

Merge Open System Servers onto zSeries



Physical Switch IOCP Definition

Two-byte Link Address

- Two-byte Link Addressing

```
*
CHPID PATH=(5A),SHARED,SWITCH=11,TYPE=FC
CHPID PATH=(5B),SHARED,SWITCH=12,TYPE=FC
CHPID PATH=(5C),SHARED,SWITCH=13,TYPE=FC
*
*
*
CNTLUNIT CUNUMBR=8000,PATH=(5A),
      UNITADD=((00,256)),LINK=(2108),
      CUADD=0,UNIT=2105
CNTLUNIT CUNUMBR=8100,PATH=(5B),
      UNITADD=((00,256)),LINK=(220C),
      CUADD=1,UNIT=2105
CNTLUNIT CUNUMBR=8200,PATH=(5C),
      UNITADD=((00,256)),LINK=(2310),
      CUADD=1,UNIT=2105
*
```

Entry Switch
(ingress)

Exit Switch
(egress)

Virtual Fabrics

FABRIC VIRTUALIZATION

Data Center Fabric Consolidation

Motivating Factors

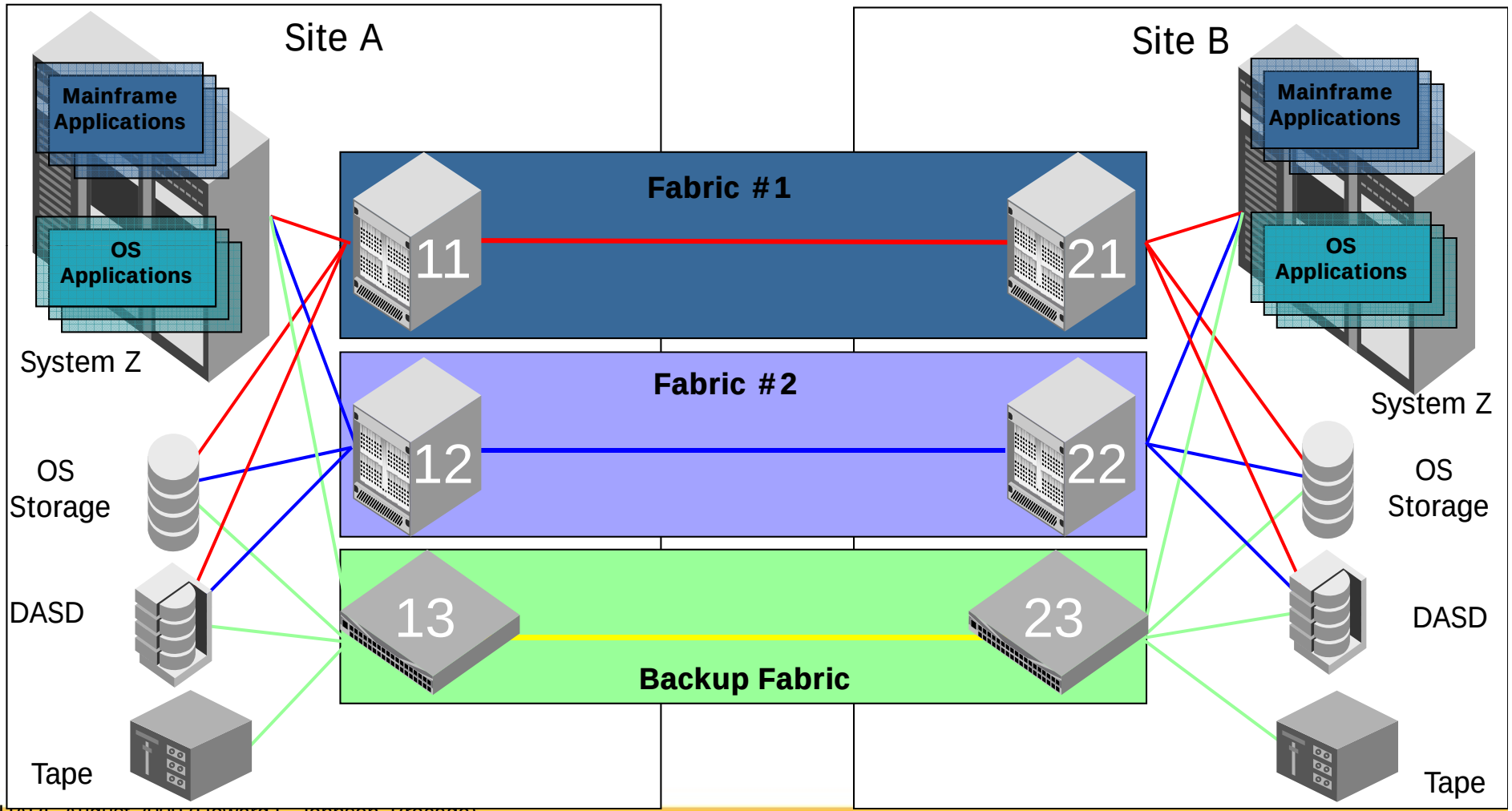
- Unorganized SAN Growth
  - Organic growth of SANs is creating large physical SAN infrastructures
  - The need to merge data centers produces larger SANs
  - Acquisition of data centers forces SAN expansion
- Controlling the growth motivates virtualization
  - Simplified management
  - Local administration
  - Access to centralized services

Fabric Consolidation
Technology

- Virtual Fabric Configuration
  - Logical Fabrics and Logical Switches
    - Utilizes frame tagging to create virtual fabrics and virtual links
  - Virtual Storage Area Networks
    - Utilizes expanded addressing to create virtual fabrics

Data Center Network

Independent Cascaded Fabrics



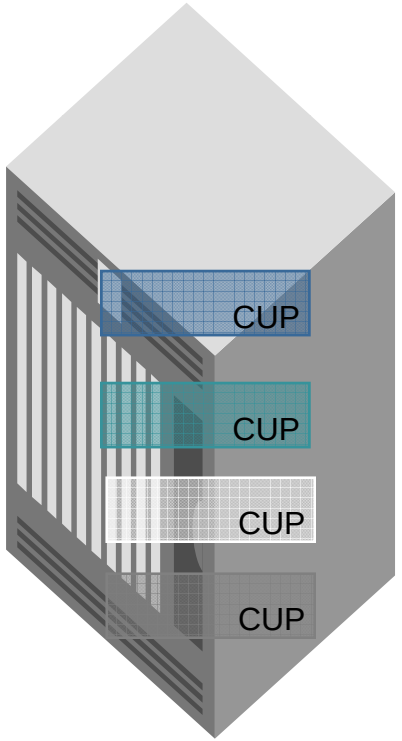
Fabric Consolidation
Large Platforms

- Larger physical building blocks
  - Fewer boxes to manage
- Logical Management
  - Virtual switch is the managed element
- Flexible port allocation
  - Allocated as needed
  - Shared between virtual switches
  - Spares efficiently managed

Virtual Fabrics

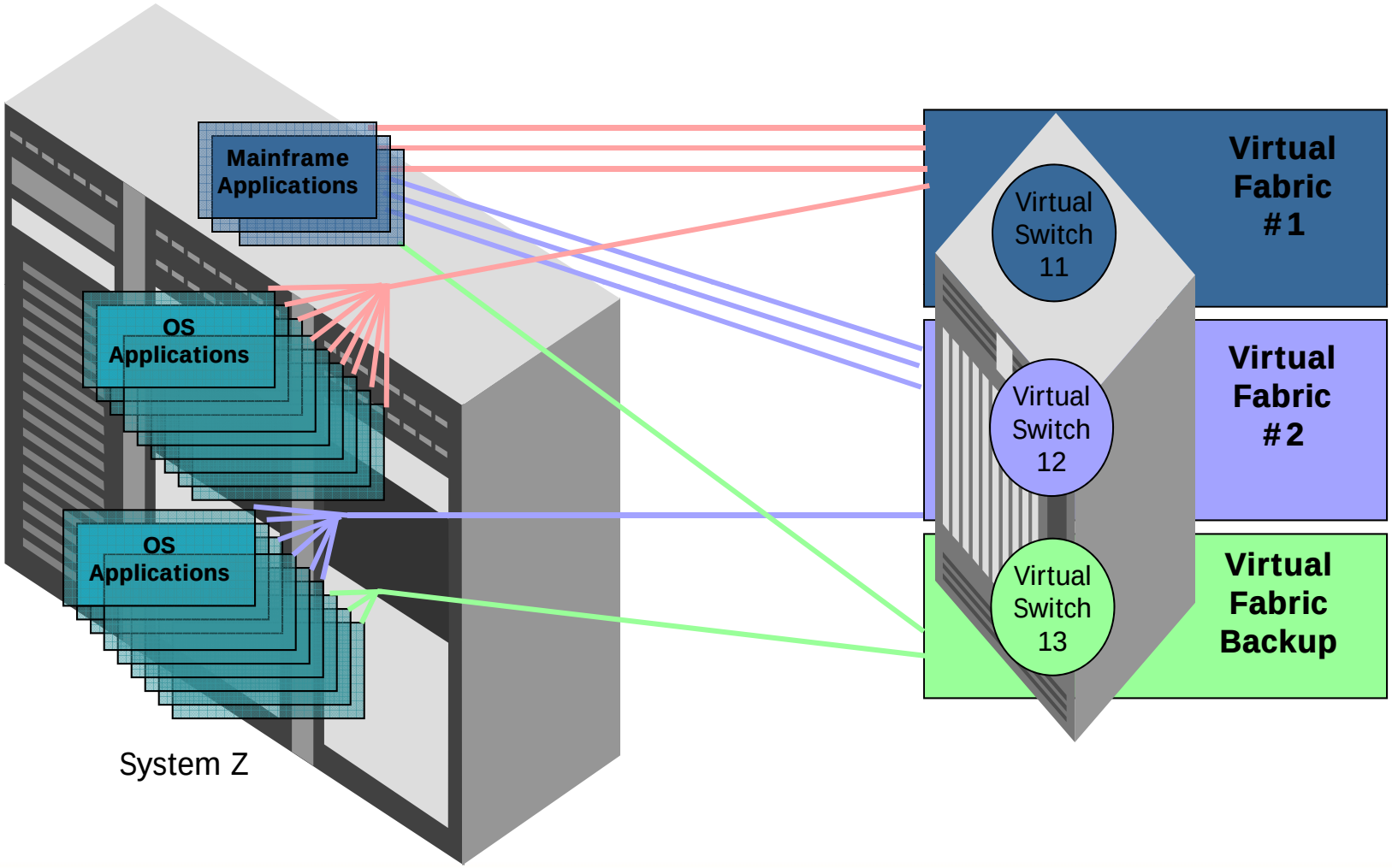
Manage SAN Growth

- Segmentation of network services
- Separation of data, control, and management
- Allocate logical fabrics at the port level
  - Administrator allocates port resources to logical fabrics
  - Administrators assigned uniquely to each logical fabric
  - Different administrative tools per logical fabric
  - Ports can be moved between logical fabrics
  - Migration path to further consolidation
  - Supports on-demand infrastructures



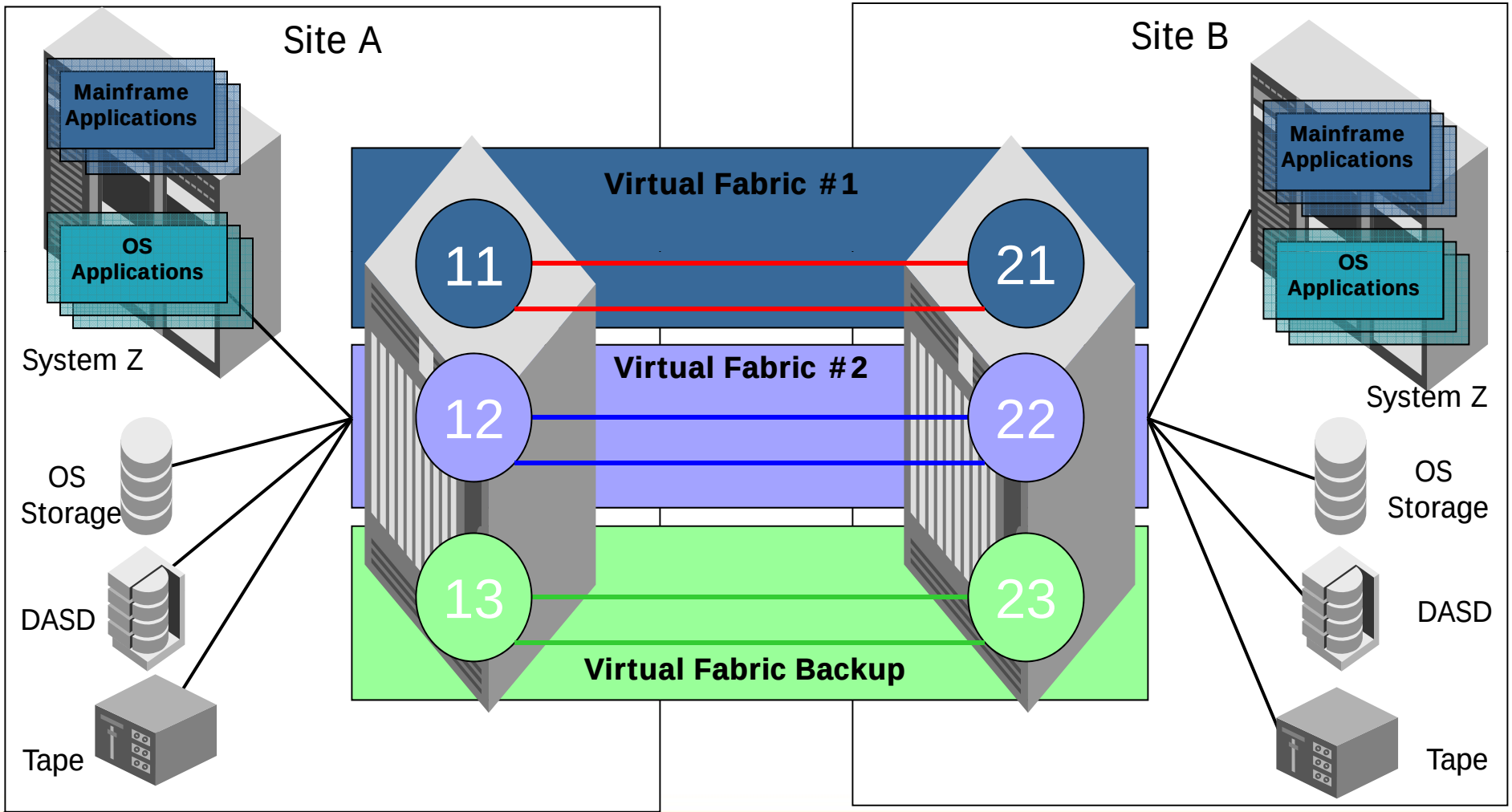
Virtualized Network

Using Virtual Fabrics



Consolidated Network

Virtual Fabrics



Virtual Switch IOCP Definition

Two-byte Link Address

- Two-byte Link Addressing

```
*
CHPID PATH=(5A),SHARED,SWITCH=11,TYPE=FC
CHPID PATH=(5B),SHARED,SWITCH=12,TYPE=FC
CHPID PATH=(5C),SHARED,SWITCH=13,TYPE=FC
*
*
*
CNTLUNIT CUNUMBR=8000,PATH=(5A),
      UNITADD=((00,256)),LINK=(2108),
      CUADD=0,UNIT=2105
CNTLUNIT CUNUMBR=8100,PATH=(5B),
      UNITADD=((00,256)),LINK=(220C),
      CUADD=1,UNIT=2105
CNTLUNIT CUNUMBR=8200,PATH=(5C),
      UNITADD=((00,256)),LINK=(2310),
      CUADD=1,UNIT=2105
*
```

Entry Switch
(ingress)

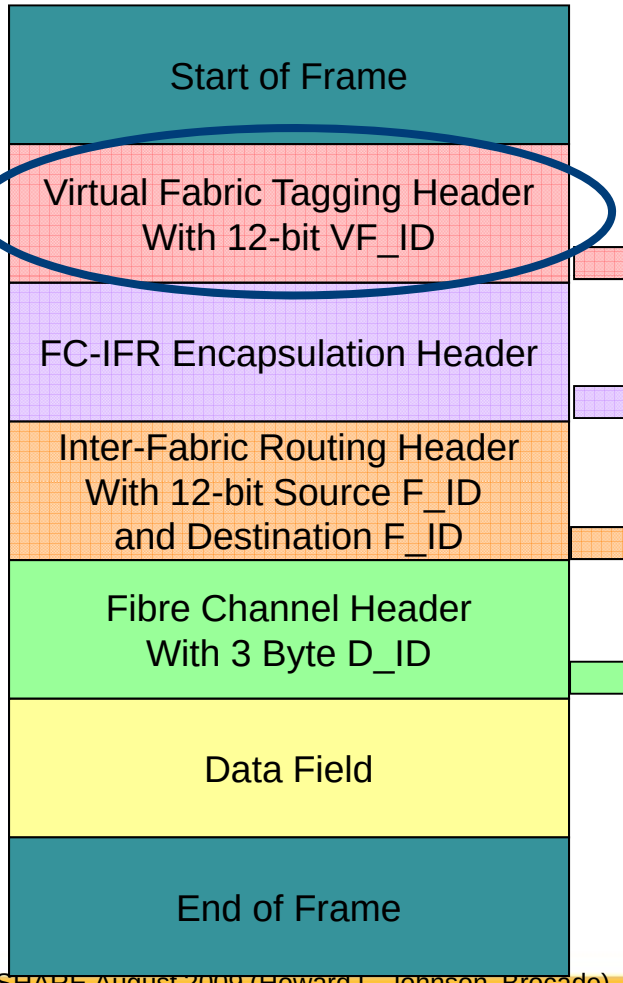
No Change!

Exit Switch
(egress)

Frame Tagging

LINK VIRTUALIZATION

Link Consolidation
Technology



- Virtual Fabric Identifier (VFID)
  - Fabric becomes virtualized
    - Supports multiple common domains on the same switch

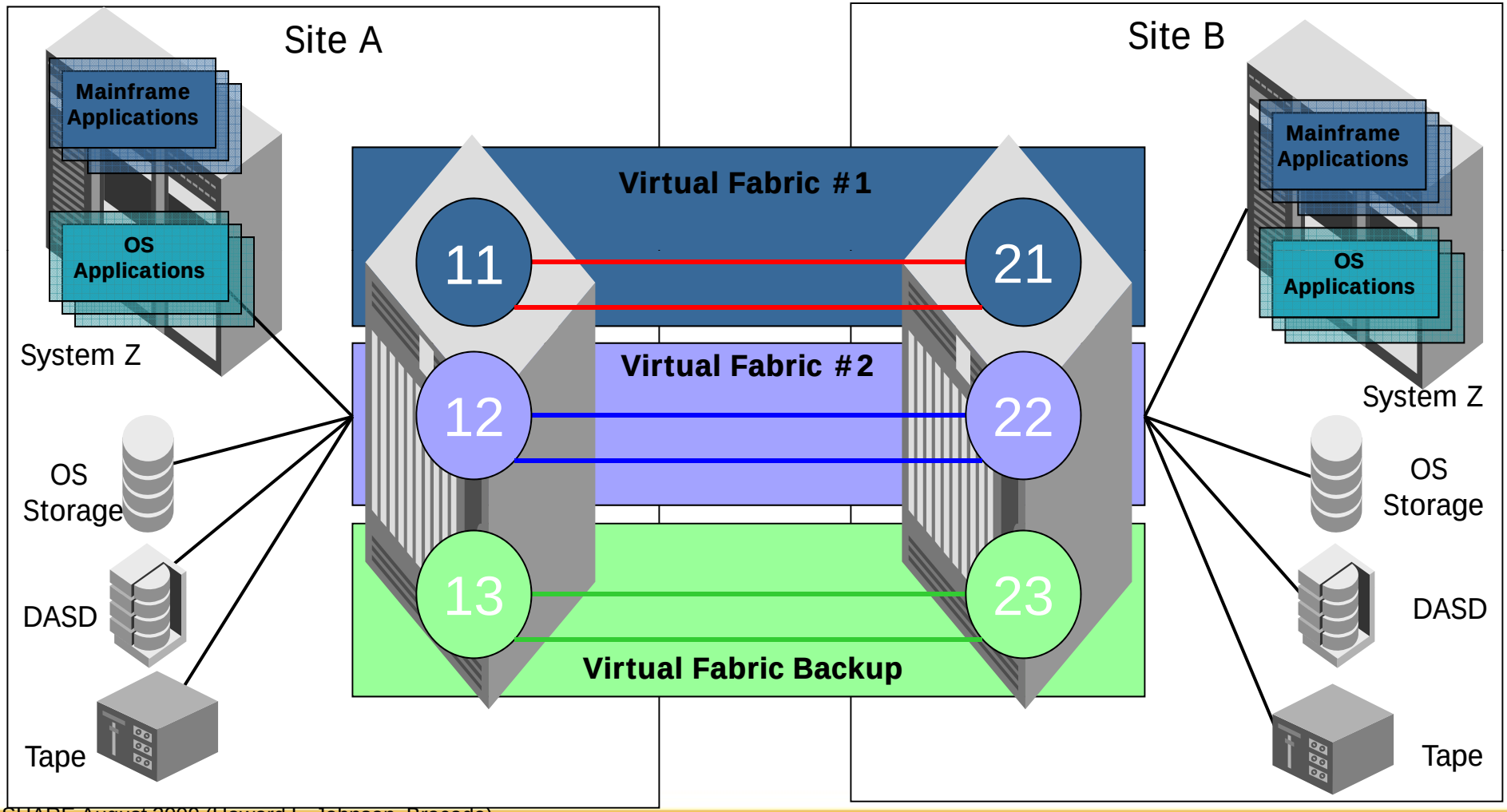
4,096 Virtual Fabric Identifiers

Encapsulation Header
- Identical to FC Header

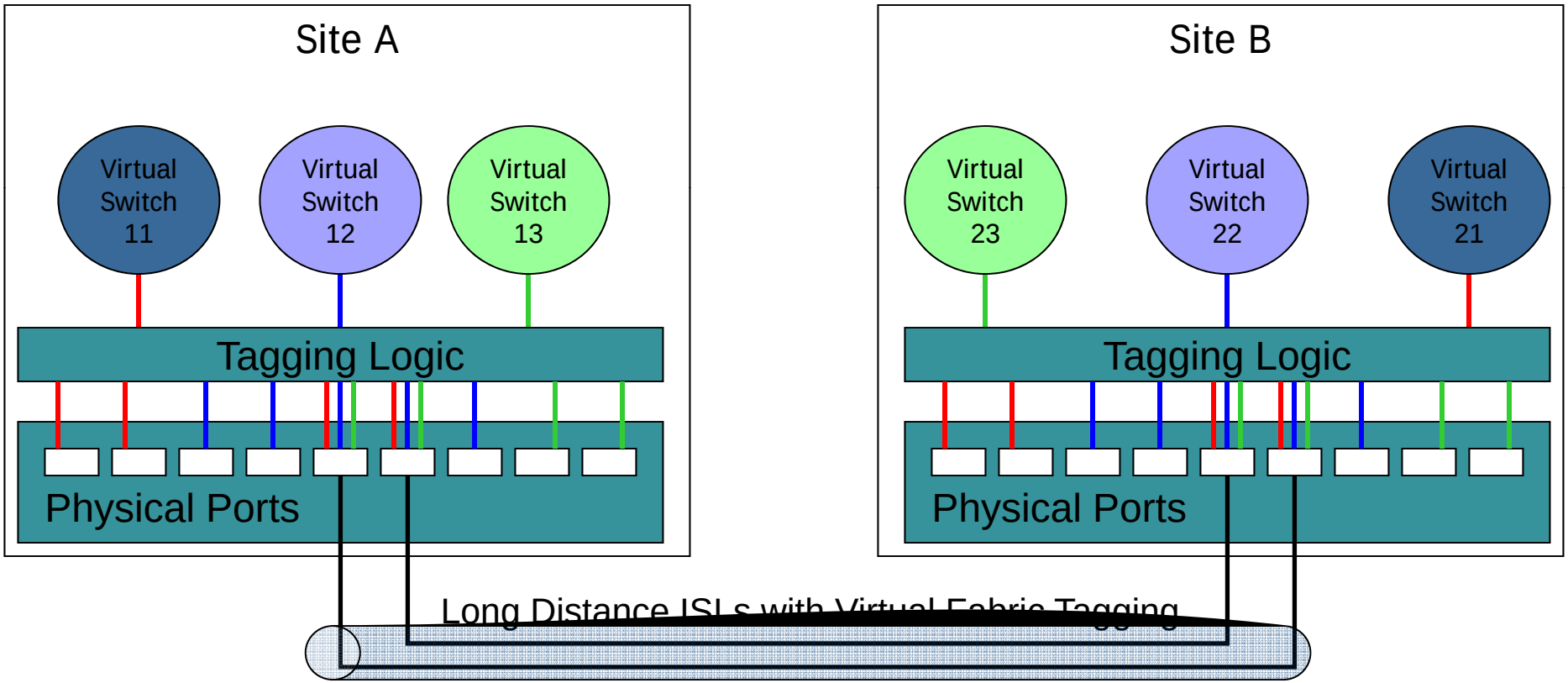
4,096 Fabric Identifiers

FC Header

Consolidated Network

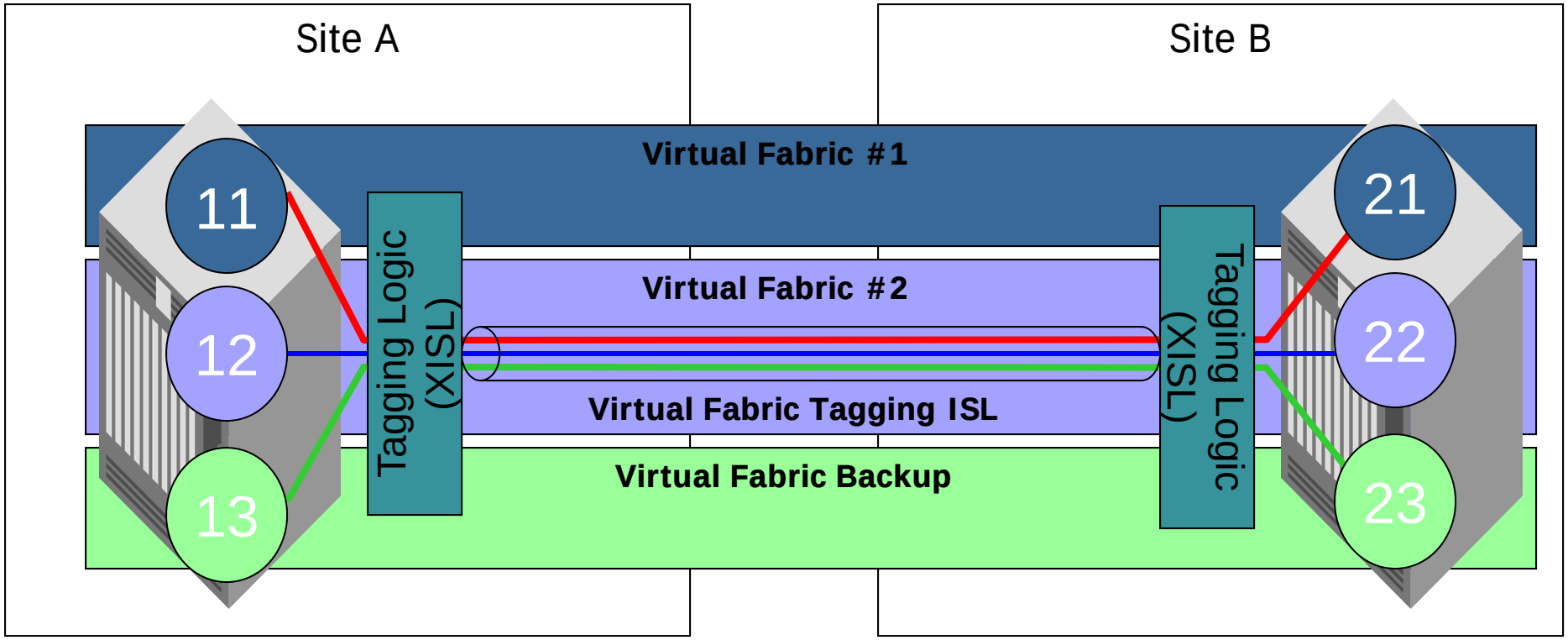


Virtual Fabric Tagging



Consolidated Data Center Network

Using Tagging ISLs



Virtual Link IOCP Definition

Two-byte Link Address

- Two-byte Link Addressing

```
*
CHPID PATH=(5A),SHARED,SWITCH=11,TYPE=FC
CHPID PATH=(5B),SHARED,SWITCH=12,TYPE=FC
CHPID PATH=(5C),SHARED,SWITCH=13,TYPE=FC
*
*
*
CNTLUNIT CUNUMBR=8000,PATH=(5A),
      UNITADD=((00,256)),LINK=(2108),
      CUADD=0,UNIT=2105
CNTLUNIT CUNUMBR=8100,PATH=(5B),
      UNITADD=((00,256)),LINK=(220C),
      CUADD=1,UNIT=2105
CNTLUNIT CUNUMBR=8200,PATH=(5C),
      UNITADD=((00,256)),LINK=(2310),
      CUADD=1,UNIT=2105
*
```

Entry Switch
(ingress)

Still
No Change!

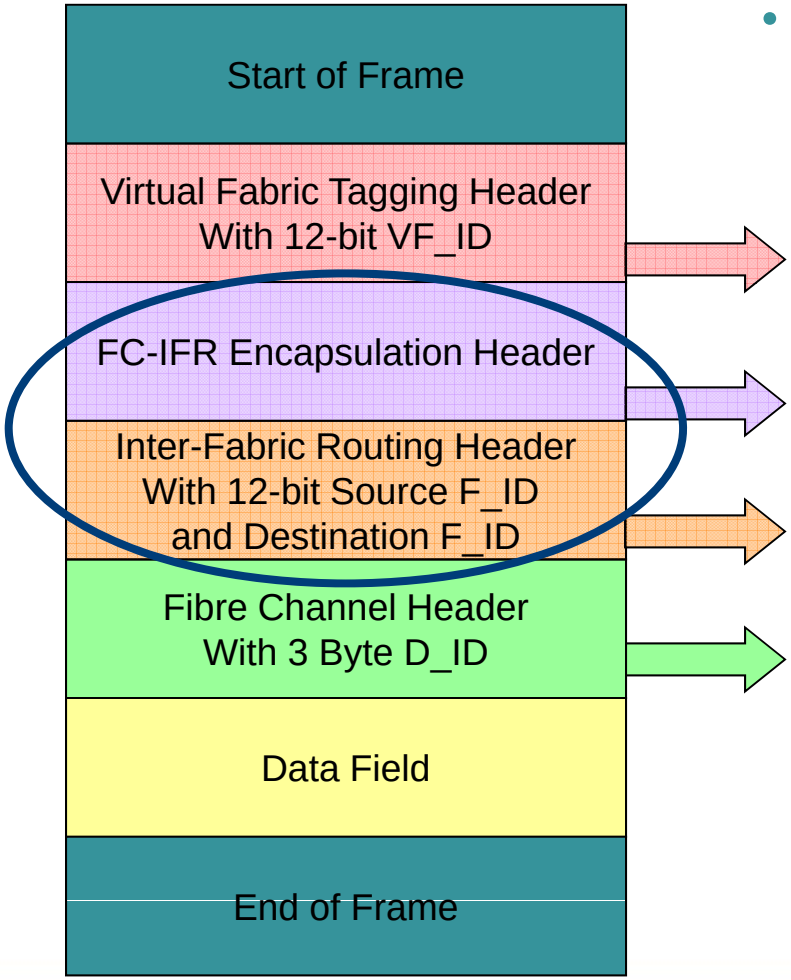
Exit Switch
(egress)

Translation and Routing

BEYOND VIRTUALIZATION

Fabric Routing Virtualization

Technology



- Inter-fabric Routing (IFR)
  - Fabric routing becomes virtualized
    - Address Translation supports multiple overlapping device addresses

4,096 Virtual Fabric Identifiers

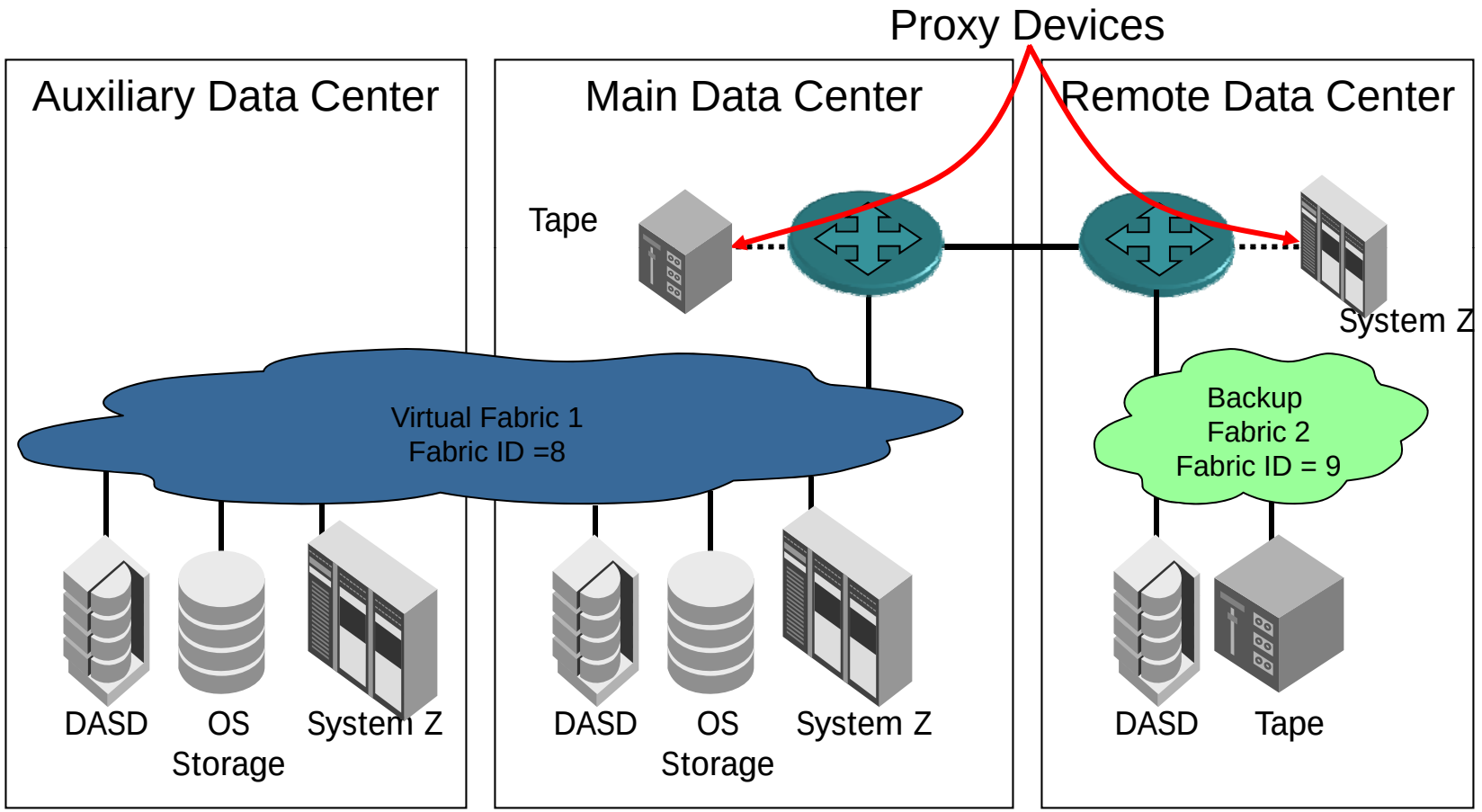
Encapsulation Header
- Identical to FC Header

4,096 Fabric Identifiers

FC Header

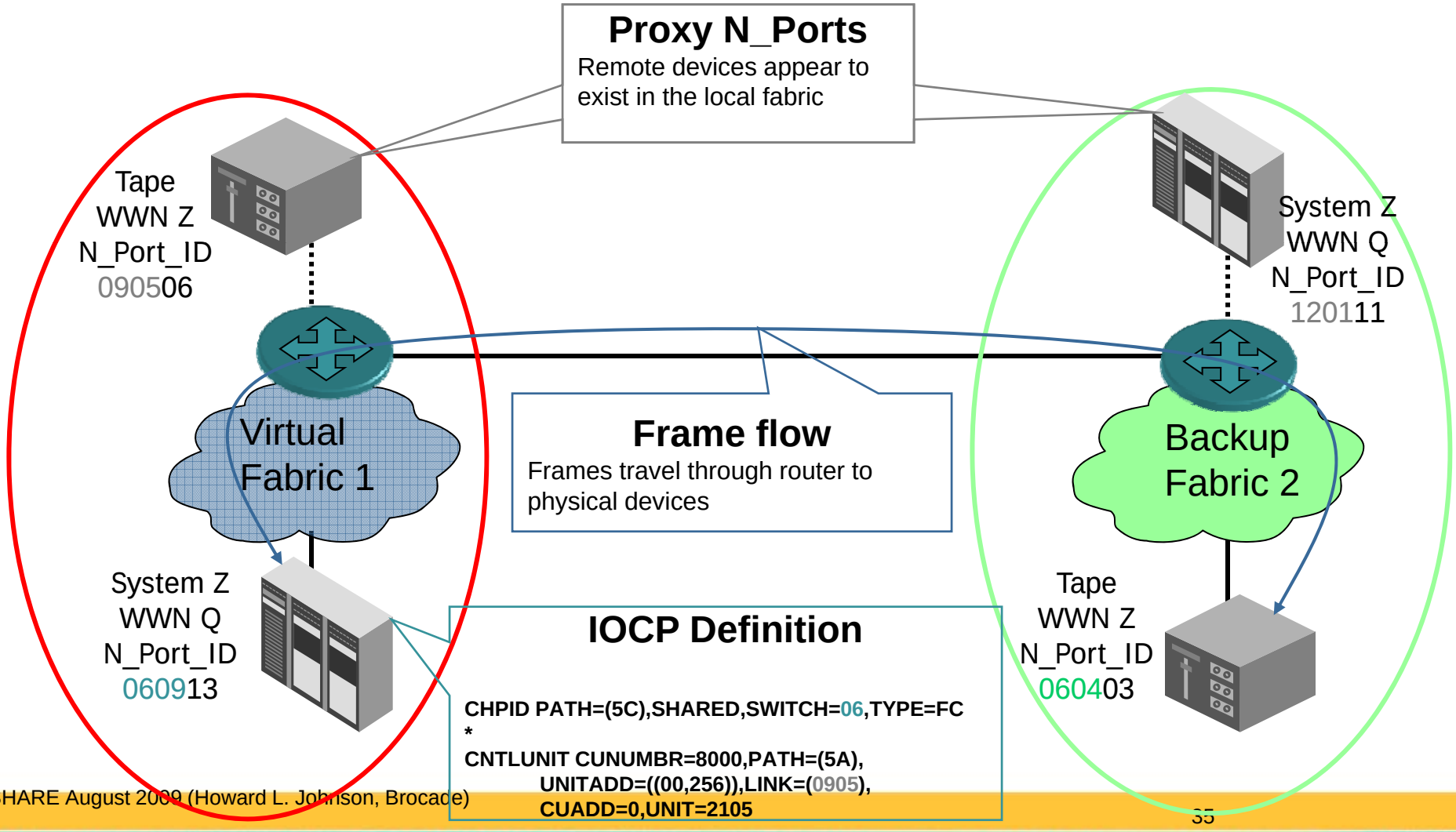
Expanding Data Centers

Accessing Remote Devices

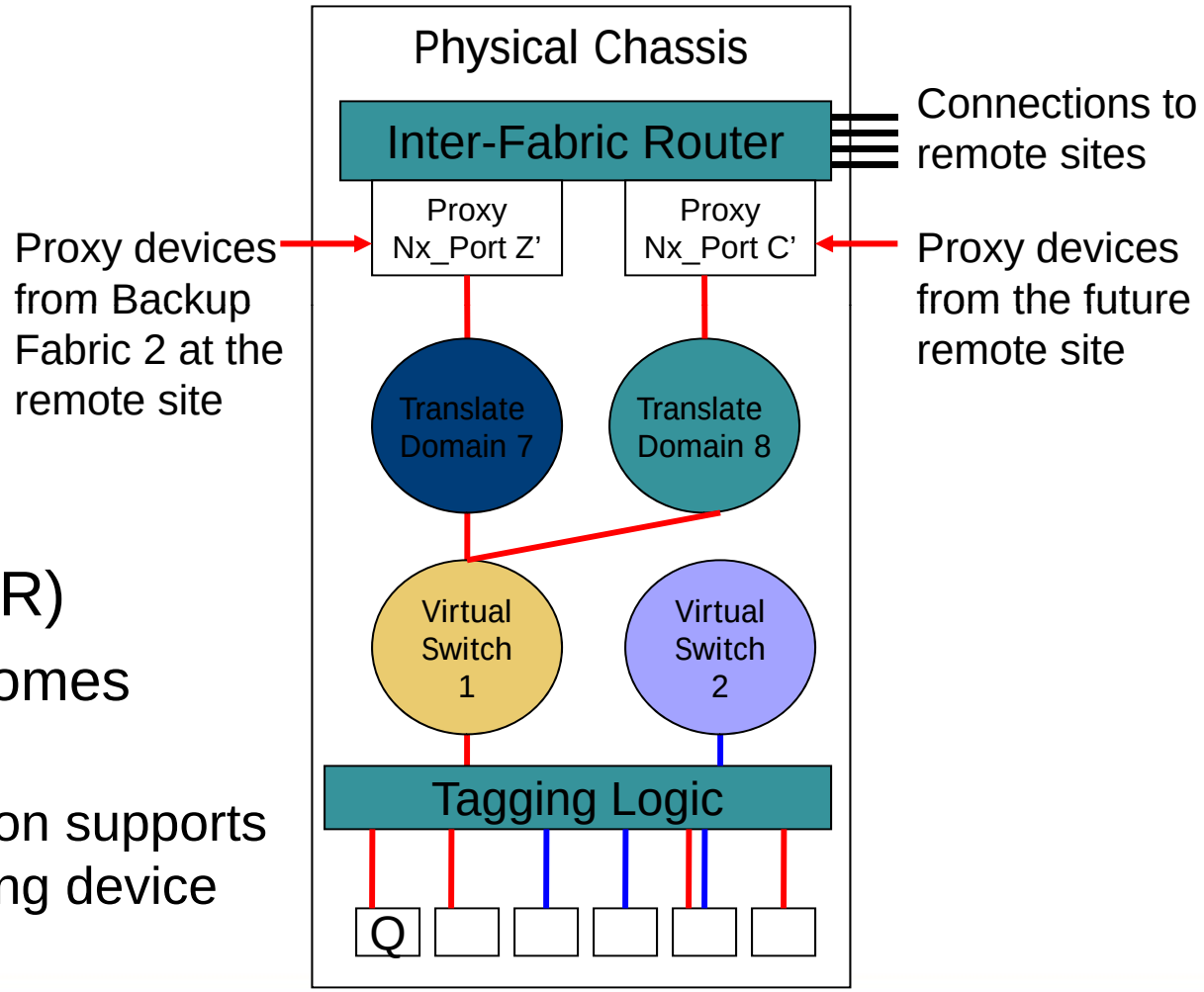


Local Representation

Proxy N_Ports



Fibre Channel Routing
Inside an IFR



Inter-fabric Router (IFR)

- Fabric routing becomes virtualized
  - Address Translation supports multiple overlapping device addresses

Summary

- Explored the fundamentals of virtualization
  - Server, fabric, link, and address/access
    - Consolidation Drives Virtualization in Storage Networks
      - *December 2006 / January 2007*
    - Fibre Channel Improves Utilization and Scalability
      - *February / March 2007*
- Examined fundamentals in theory
  - Server
    - N_Port Identifier Virtualization
  - Fabric
    - Virtual Fabric Identifier
    - Inter-fabric Routing

Other Sessions

SHAMELESS PROMOTIONS

Buffer-to-Buffer Credits, Exchanges, and Urban Legends

- Lou Ricci, IBM
- Howard L. Johnson, Brocade

- 27 August 2009
- 9:30am – 10:30am
- Session 6202

- Abstract
  - Performance in a FICON network is influenced by the underlying flow control mechanisms of Fibre Channel. In this session, we examine how Buffer-to-Buffer credits flow from the channel to the control unit. We also look at how exchanges are used in FICON applications and how they change with the introduction of zHPF. During both examinations, we explore the role of the FICON Director in managing Buffer-to-Buffer credits and exchanges over a cascaded network. Throughout the session, we debunk the various FICON "Urban Legends" featuring credits and exchanges. Take the opportunity to learn from two of the FICON industry's leading experts in channel and fabric development and join our session.

z/OS Nuggets: FICON I/O in Depth

- Lou Ricci, IBM

- 26 August 2009
- 1:30pm – 2:30pm
- Session 2214

- Abstract
- In this session, the speaker will explain what FICON is and how it is different from ESCON. Then he will cover topics that include:
  - What is a FICON exchange
  - Buffer-to-Buffer credit and how it affects a FICON fabric
- Are you curious if MIDAW will help you? The speaker, the FICON firmware team leader, will provide the answer to that question by describing the basics of the FICON architecture from the time a channel comes to life and all the way through the execution of a typical I/O request. And he will cover how a particular channel is selected to perform an I/O operation and the steps the channel goes through to pass commands and data to and from the control unit.

SAN (Storage Area Networking) Overview

- Martin Mune, Brocade
- Mike Blair, Cisco

- 26 August 2009
- 3:00pm – 4:00pm (part 1)
- 4:30pm – 5:30pm (part 2)
- Session 6210/6211

- Abstract
  - This introductory technical session will introduce SANs(Storage Area Networks) to those who are not familiar with them. It will discuss these for those with little to no background in SANs and will give the user some sense of the concepts, terms, acronyms, and challenges faced when starting to look at implementing a SAN. This is a co-sponsor with the CTEC project session 3102.

Understanding FICON Performance

- Dave Lytle, Brocade

- 24 August 2009
- 3:00pm – 4:00pm
- Session 6204

- Abstract
  - This technical session will discuss some of the limitations of a FICON infrastructure. It will discuss and explain buffer credits as well as frame pacing delay FICON issues. Lastly, a discussion on the tools and techniques for managing performance issues will also be covered.

Introduction to Virtualization: z/VM Basic Concepts and Terms

- Bill Bitner, IBM

- 27 August 2009
- 1:30pm – 2:30pm
- Session 9102

- Abstract
  - The increased interest in virtualization has resulted in many people investigating virtualization technology. Exploring a different technology has the challenge of being introduced to new terminology and concepts. With over 30 years of experience in virtualization, z/VM on System z is the premier enterprise solution. This session is designed for those who are new to z/VM, are back working on z/VM after being away from it for a few years, or are just confused about the concepts of 'virtualization'. This session describes the basics of the virtualization in z/VM, the resources it can manage, and various other facilities it offers. Terminology that will be used in other z/VM sessions this week will be introduced and explained here. Even if z/VM is not the virtualization platform you will ultimately use, understanding it is helpful as it is often the benchmark used to judge others.

REFERENCES

Articles

- zJournal
  - Consolidation Drives Virtualization in Storage Networks
    - December 2006 / January 2007
    - http://www.zjournal.com/index.cfm?section=article&aid=739
  - Fibre Channel Improves Utilization and Scalability
    - February / March 2007
    - http://www.zjournal.com/index.cfm?section=article&aid=764

About the Authors

- Howard Johnson
  - Howard is a Technology Architect at Brocade and is responsible for furthering Brocade's lead in FICON technology. His expertise encompasses Brocade's ESCON and FICON products and includes an extensive relationship with IBM's zSeries I/O development team.
  - hjohnson@brocade.com
- Scott Kipp
  - Scott develops standards, including NPIV and virtual fabrics, for Brocade. He has written several books for the Fibre Channel industry Association and speaks with storage network user groups about his latest book, Fibre Channel Advances.
  - skipp@brocade.com
- Steve Guendert
  - Steve is Brocade's Mainframe Solutions principal engineer and is an industry expert on ESCON and FICON. He has an MBA and an MS in Management Information Systems and recently completed his Ph.D.
  - sguendert@brocade.com

Articles

- IBM Journal of Research and Development
  - Sharing FCP Adapters through Virtualization
    - January / February 2007
    - http://researchweb.watson.ibm.com/journal/rd/511/srikrishnan.html

Standards
NPIV

- FC-LS
  - Describes FDISC use to allocate additional N_Port_IDs in Section 4.2.32
  - Service Parameters for FDISC are described in Section 6.6
  - NV_Ports are treated like any other port
    - Exception is they use FDISC instead of FLOGI
  - Documents the responses to NV_Port related ELSs in section 6.4.5
    - FDISC, FLOGI and FLOGO
  - http://www.t11.org/t11/docreg.nsf/ufile/06-393v6
- FC-GS-5
  - Describes Name Server queries in 5.2.5
    - Permanent Port Name and Get Permanent Port Name command
      - *Based on the N_Port ID (G_PPN_ID)*
    - The PPN shall be the F_Port Name in FC-GS-6
  - http://www.t11.org/t11/docreg.nsf/ufile/06-393v6
- FC-DA
  - Profiles the process of acquiring additional N_Port_IDs in Clause 4.9
  - http://www.t11.org/t11/docreg.nsf/ufile/04-202v2
- FC-MI-2
  - Profiles how the fabric handles NPIV requests
    - New Service Parameters are defined in Section 6.3
    - Name Server Objects in 7.3.2.2 and 7.3.2.3
  - http://www.t11.org/t11/docreg.nsf/ufile/04-109v4

Standards
Virtual Fabrics and Inter-Fabric Routing

- Virtual Fabrics
  - FC-FS-2
    - Overview of Virtual Fabrics and the Virtual Fabric Tag Header in 10.2
    - http://www.t11.org/t11/docreg.nsf/ufile/06-085v3
  - FC-LS
    - Virtual Fabrics Bit in Common Login Parameters in 6.6.2
    - Exchange Virtual Fabric Parameters in 4.2.43
    - http://www.t11.org/t11/docreg.nsf/ufile/06-393v6
  - FC-SW-4
    - Overview and Processing in 12
    - Exchange Virtual Fabric Parameters SW_ILS in 6.1.26
    - http://www.t11.org/t11/docreg.nsf/ufile/05-033v5
- Inter-Fabric Routing
  - FC-FS-2
    - Inter-Fabric Routing Extended Header in 10.3
    - http://www.t11.org/t11/docreg.nsf/ufile/06-085v3
  - FC-Inter-Fabric Routing
    - Complete definition of the protocols to initiate and manage IFRs is in progress but several pre-standard implementations are already being used
    - http://www.t11.org/t11/docreg.nsf/ufile/07-051v0
    - This draft is subject to change

SHARE

Denver

August 2009

THANK YOU!

Speaker Biography

- Howard L. Johnson
  - BROCADE
  - Technology Architect, FICON
    - 25 years technical development and management
- Technical Highlights
  - Design and Architecture
    - ESCON 9032-3/4/5
    - FICON Bridge 9032-5
    - Fibre Channel ED-6064/ES-3x32/ES-3x16
    - FICON ED-6064/ED-6140
- Contact Information
  - howard.johnson@brocade.com

THIS SLIDE INTENTIONALLY LEFT BLANK