



# **Driving Innovation Through the Information Infrastructure**

**SPRING 2011**



# **Unified Fabric Anywhere - Getting from Here to There**

J Michel Metz, Ph.D

Product Manager - FCoE

Cisco Systems

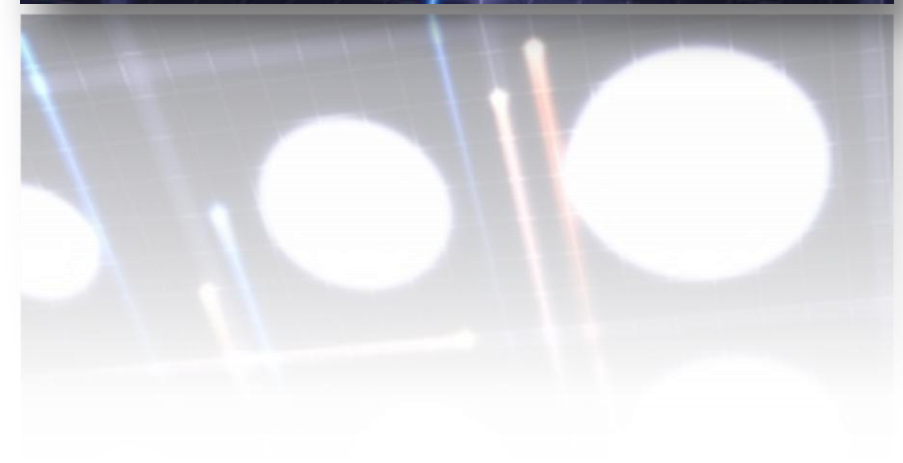
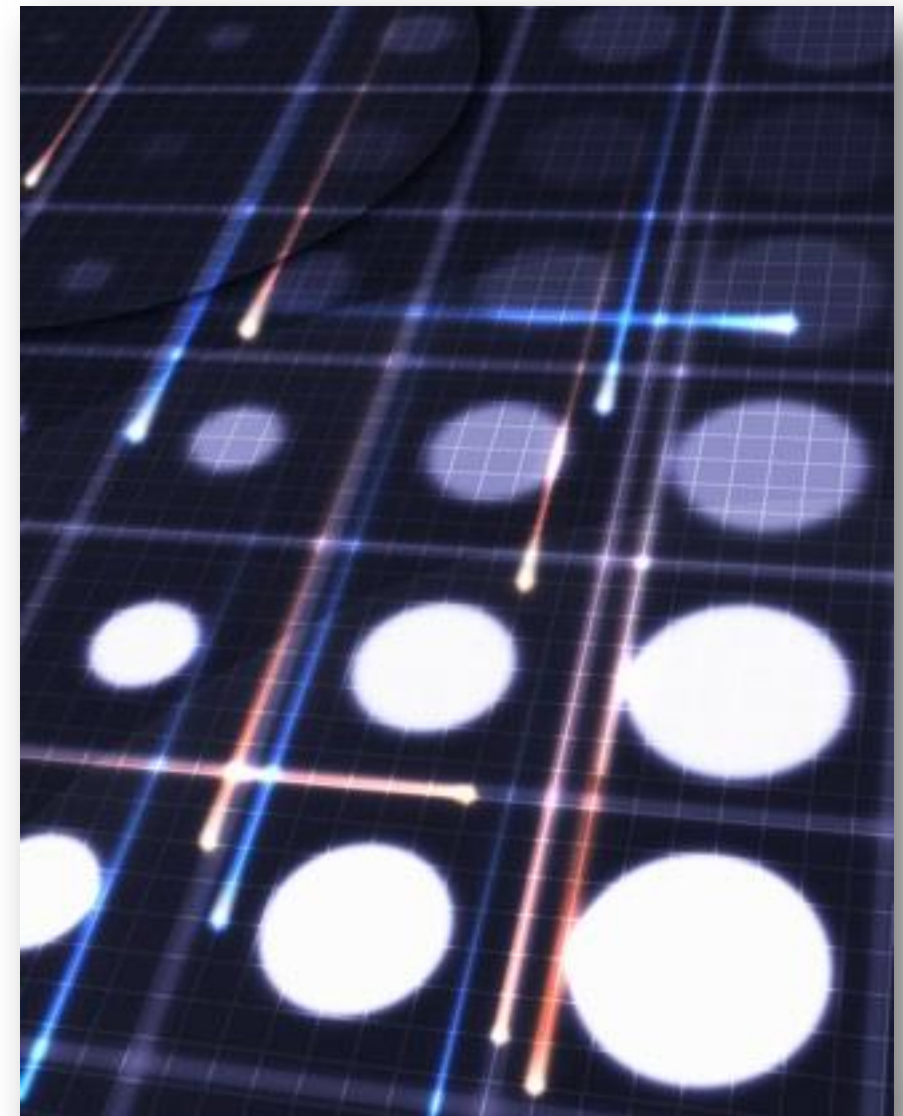
Twitter: @jmichelmetz





# Agenda

- Assumptions and Definitions
  - What we mean with what we say
- Technical Considerations
  - Expansion and Growth
  - Best Practices and Implementations
- Organizational Considerations
  - It's all about the people, people!
- Long-Term Consequences
  - What can you expect over time?





# What This Session Is

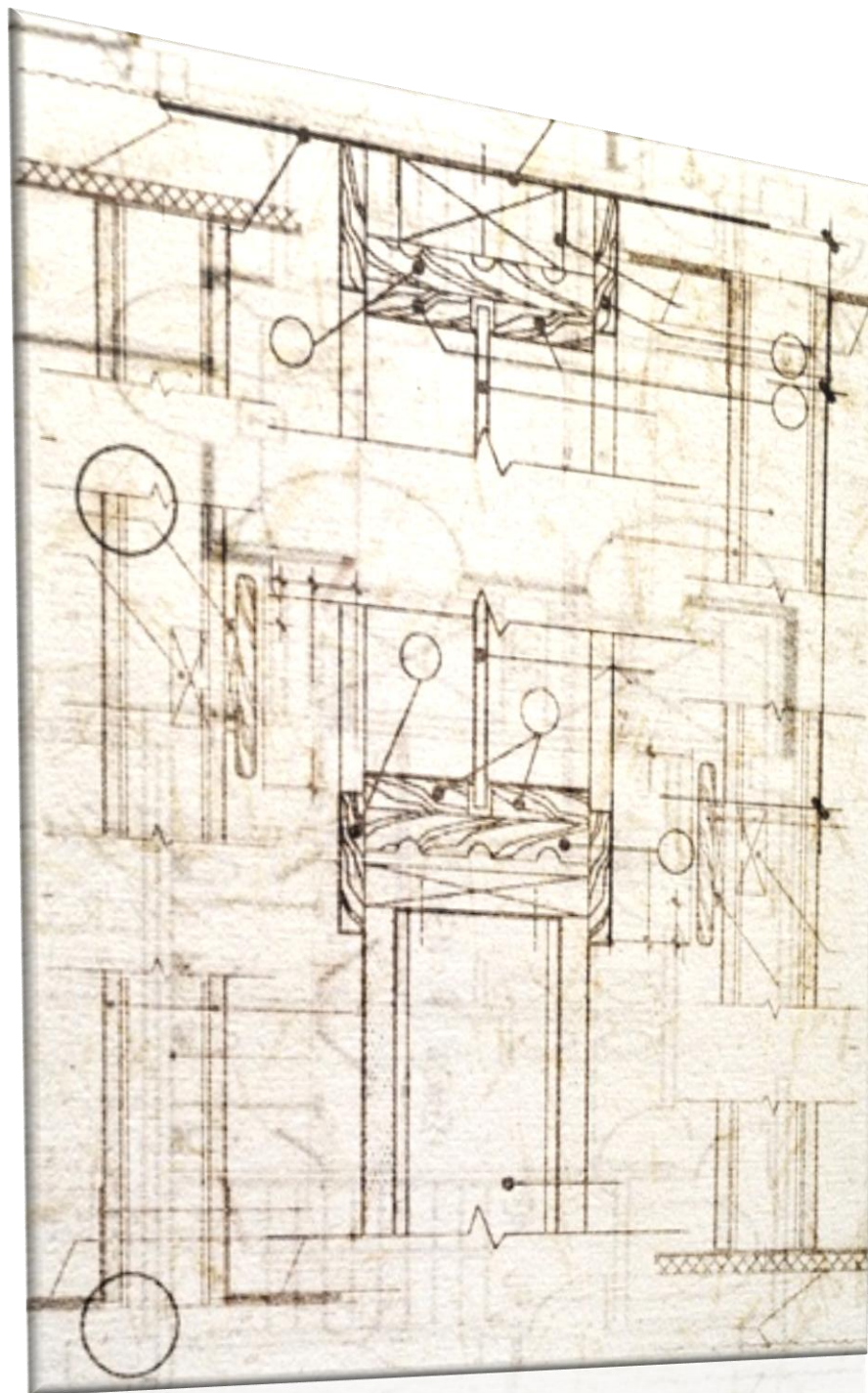


- Examination of
  - What you can do now
  - Why you might do this
  - What the terms mean
  - How to understand all the hype
- Description of things to be aware of
  - Best practices
  - Design alternatives
  - Organizational considerations
- Eliminate confusion and misdirection





# What This Session Is NOT



- An FCoE primer
- A thorough coverage of every possible deployment
- A recommendation or proscription for your data center
- A sales pitch



# Assumptions and Definitions

- You have at least *heard* of FCoE
- You are curious about the Data Center future
- You will walk away with both answers *and* questions
- You are not looking for a panacea for all the problems in the Data Center







# Facing the Same Direction

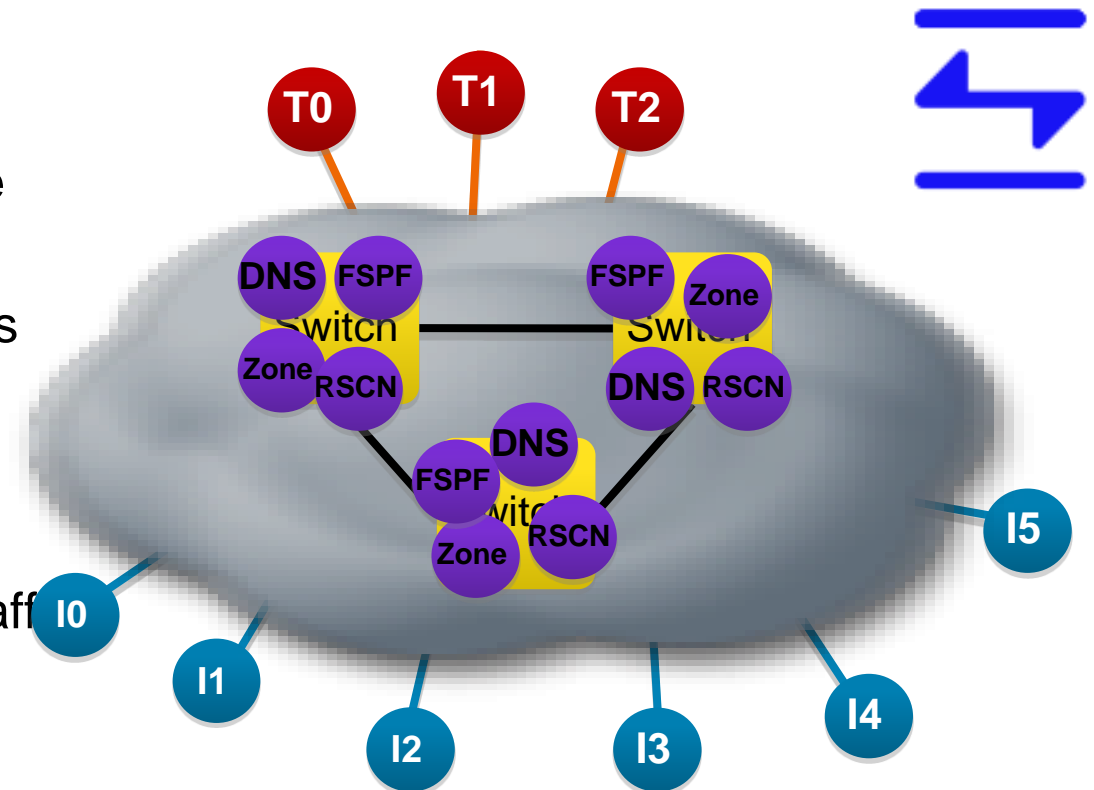
- Storage or Network? SAN or LAN?
- Your paradigm will guide your understanding
  - Ethernet is non-deterministic.
    - Flow control is destination-based
    - Relies on TCP drop-retransmission / sliding window
  - Fibre-Channel is deterministic.
    - Flow control is source-based (B2B credits)
    - Services are fabric integrated (no loop concept)
- Where you come from will guide how you see where you're going



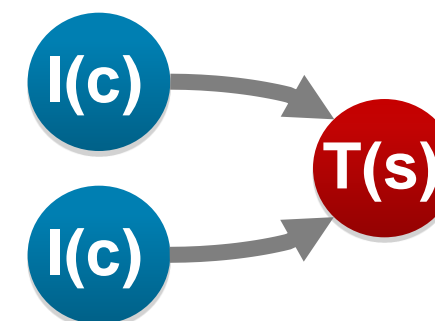


# FC: Network and Fabric

- Fibre Channel SAN
  - Transport and Services are on **same layer** in the same devices
  - Well-defined end device relationships (initiators and targets)
  - Does not tolerate packet drop – requires **lossless** transport
  - Only north-south traffic, east-west traffic mostly irrelevant
- Network designs optimized for Scale and Availability
  - High availability of network services provided through dual fabric architecture
  - Edge/Core vs Edge/Core/Edge
  - Service deployment



Fabric topology, services and traffic flows are structured



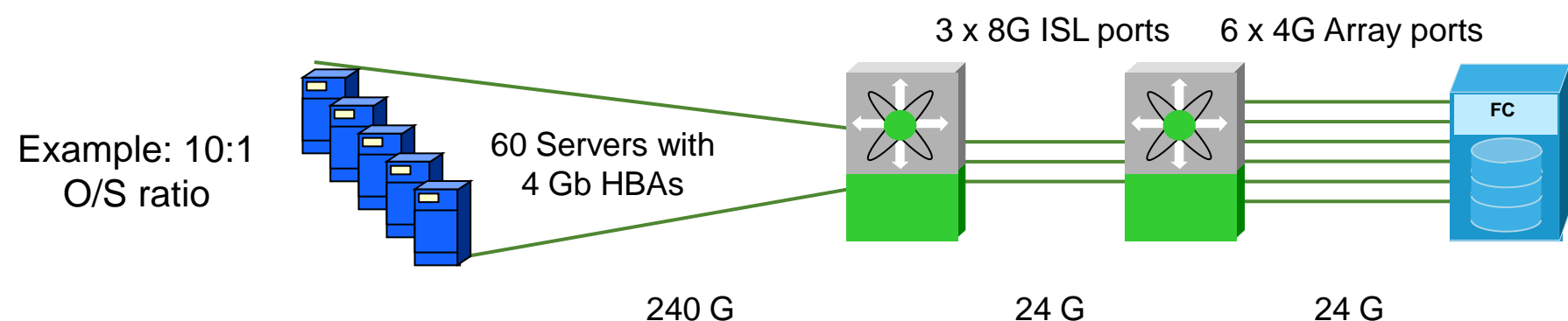
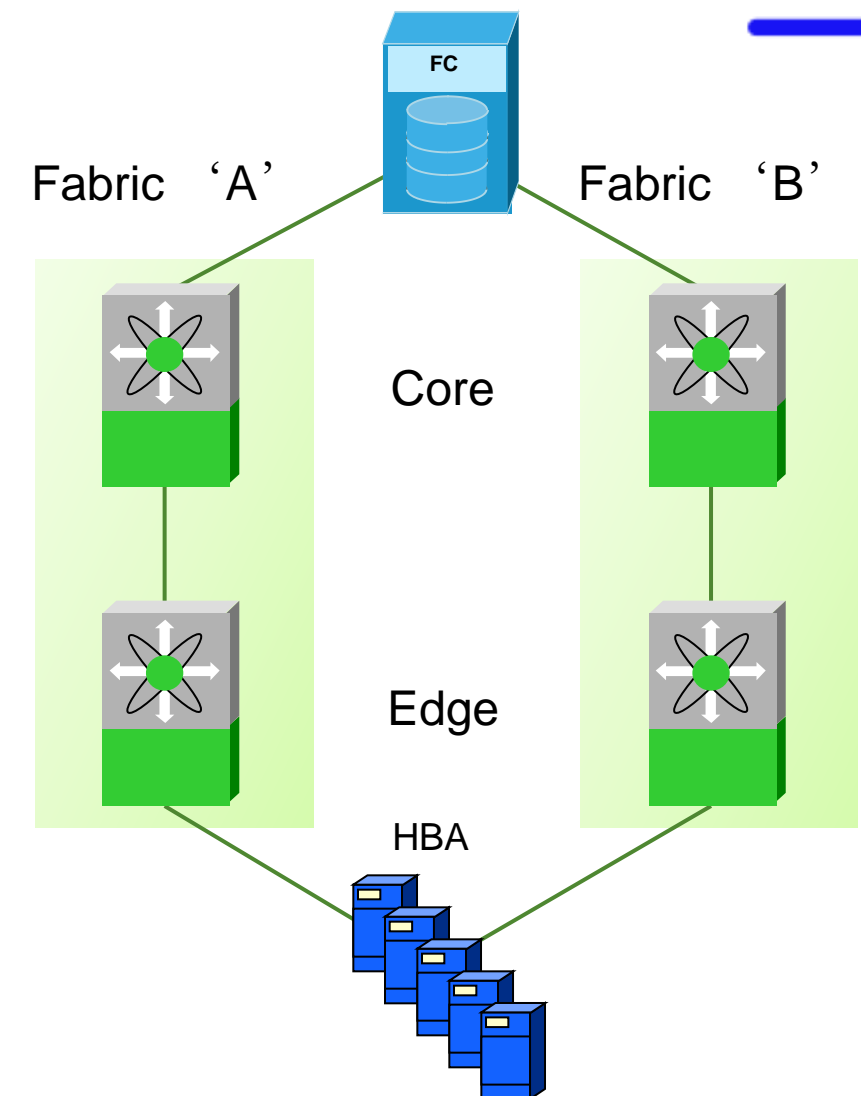
Client/Server Relationships are pre-defined





# FC: Network and Fabric

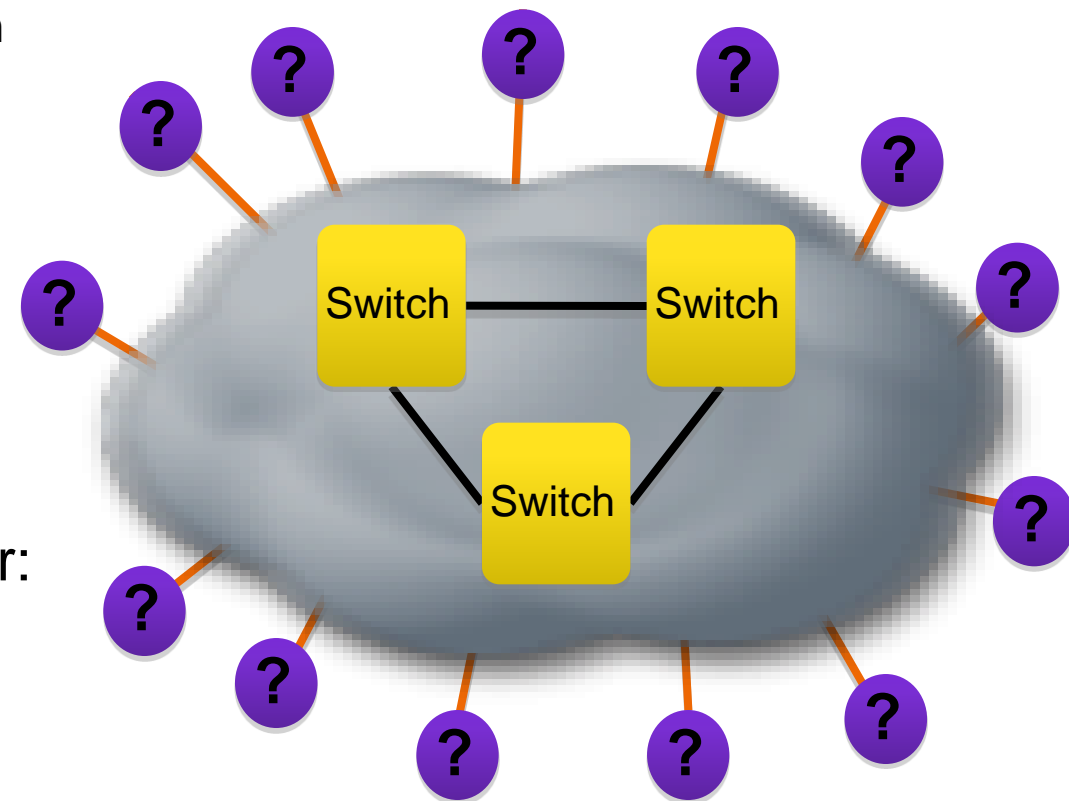
- “Edge-Core” or “Edge-Core-Edge” Topology
- Servers connect to the edge switches
- Storage devices connect to one or more core switches
- HA achieved in two physically separate, but identical, redundant SAN fabric
- Very low oversubscription in the fabric (1:1 to 12:1)
- FLOGI Scaling Considerations



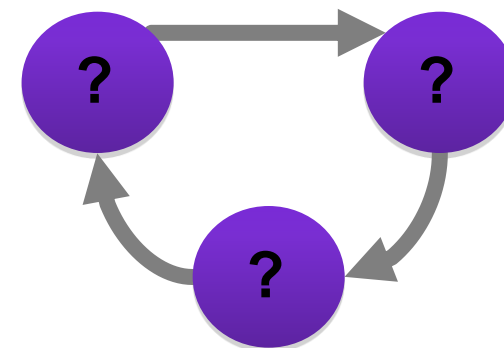


# Ethernet: Network and Fabric

- Ethernet/IP
  - Goal : provide **any-to-any** connectivity
  - Unaware of packet loss (“lossy”) – relies on Upper Layer Protocols (ULPs) for retransmission and windowing
  - Provides the transport without worrying about the services - **services provided by upper layers**
  - East-west vs. north-south traffic ratios are undefined
- Network design has been optimized for:
  - High Availability from a transport perspective by connecting nodes in mesh architectures
  - Service HA is implemented separately
  - Takes in to account control protocol interaction (STP, OSPF, EIGRP, L2/L3 boundary, etc...)



**Fabric topology and traffic flows are highly flexible**

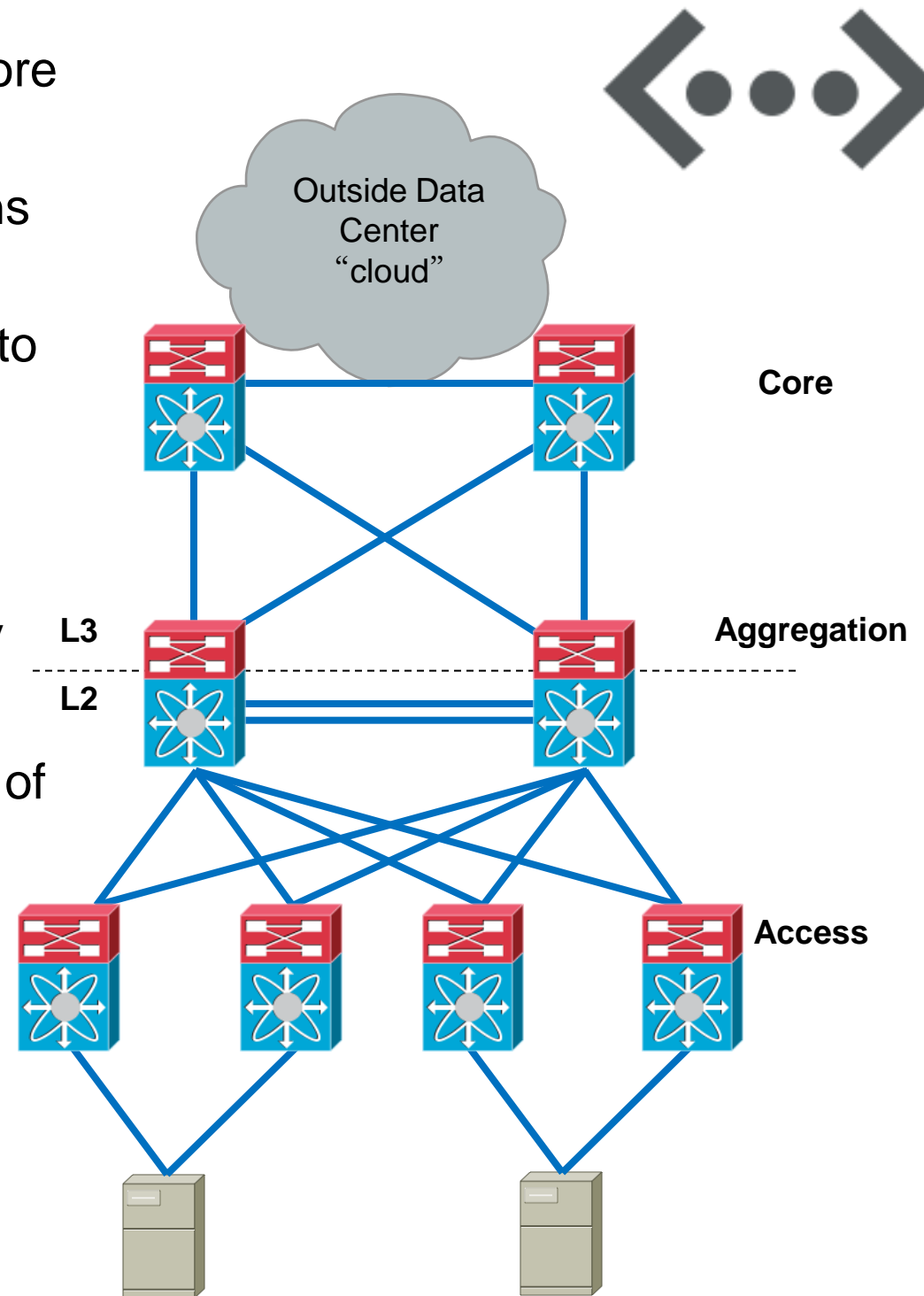


**Client/Server Relationships are not pre-defined**



# Ethernet: Network and Fabric

- Servers typically dual homed to two or more access switches
- LAN switches have redundant connections to the next layer
- Distribution and Core can be collapsed into a single box
- L2/L3 boundary typically deployed in the aggregation layer
  - Spanning tree or advanced L2 technologies used to prevent loops within the L2 boundary
  - L3 routes are summarized to the core
- Services deployed in the L2/L3 boundary of the network (load-balancing, firewall, etc)





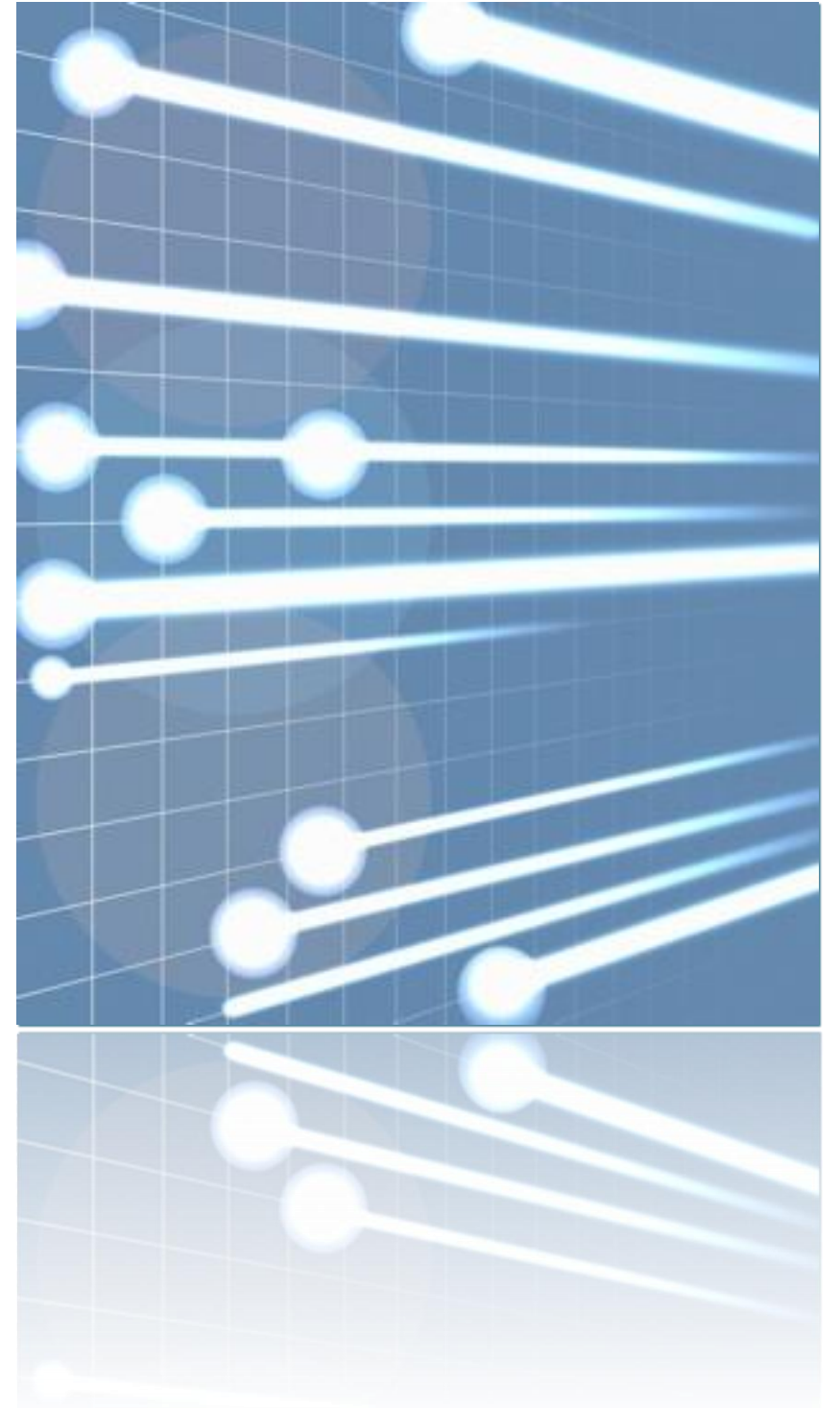
# FCoE Building Buzzwords

- **Links**

- **DCB Lossless Link** – "Data Center Bridging" Ethernet link, capable of providing any traffic types, including lossless
- **Converged Link** – a single DCB Ethernet link actively carrying **both** LAN and Storage (FC/FCoE/NAS/iSCSI) traffic simultaneously
- **Dedicated FCoE Link** -- a single DCB Ethernet link capable of carrying all traffic types but actively **dedicated** to FCoE for traffic engineering purposes

- **Networks**

- **Converged Access** – Converged switches at the access layer, dedicated switches beyond
- **Converged Network** – Converged switches beyond the access layer







# The Goals of Unified Fabric

## Unified Fabric = One Network

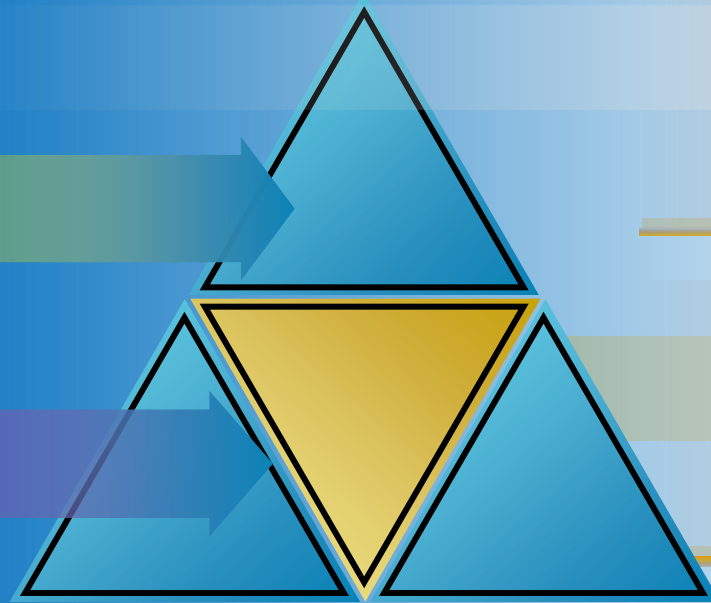
LAN + Storage (NAS + iSCSI+ FC+ FCoE)

Apply the roadmap and economies of scale of Ethernet. Extend current data center lifecycle

Provide the high performance and lossless nature of Fibre Channel with Ethernet

Wire hosts once to connect to any network - SAN, LAN, HPC. Faster rollout of new apps and services

Every host capable to mount any storage target, increasing SAN attach rate. Drive storage consolidation and improve utilization





# Why Unified Fabric?

## *Ethernet Model has Proven Benefits*

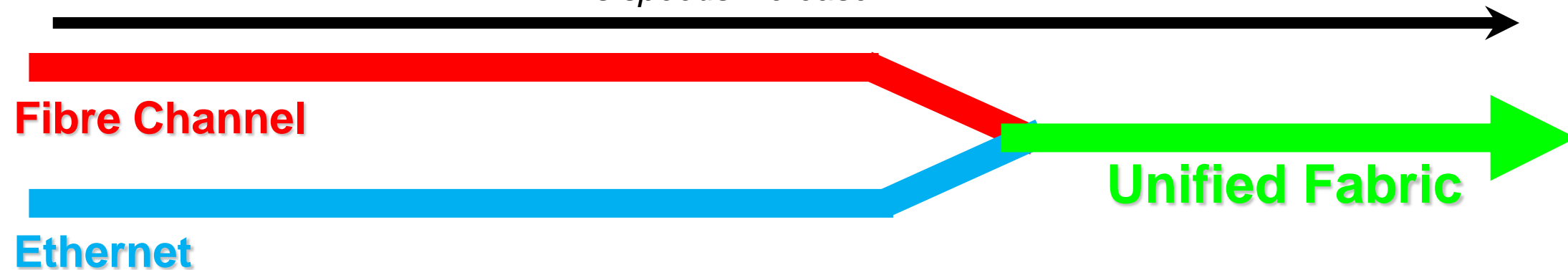
### Ethernet Economic Model

- Embedded on Motherboard
- Integrated into O/S
- Many Suppliers
- Mainstream Technology
- Widely Understood
- Interoperability by Design

### FC Economic Model

- Always a Stand-Up Card
- Specialized Drivers
- Few Suppliers
- Specialized Technology
- Special Expertise
- Interoperability by Test

*As speeds increase....*



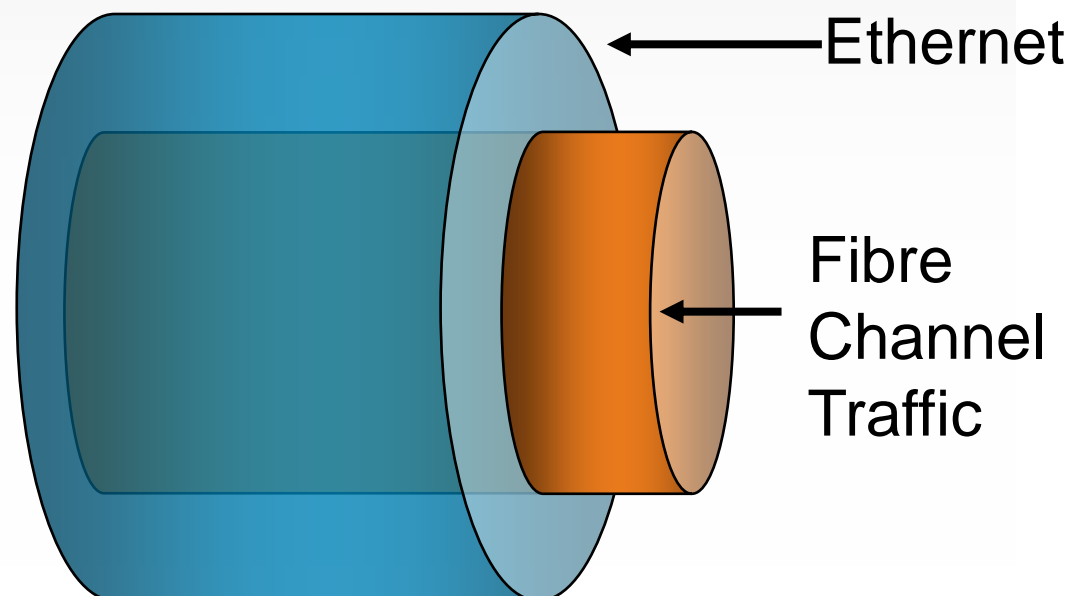




# FCoE

## *FCoE*

- Encapsulation of FC Frames over Ethernet
- Enables FC to Run on a Lossless Ethernet Network



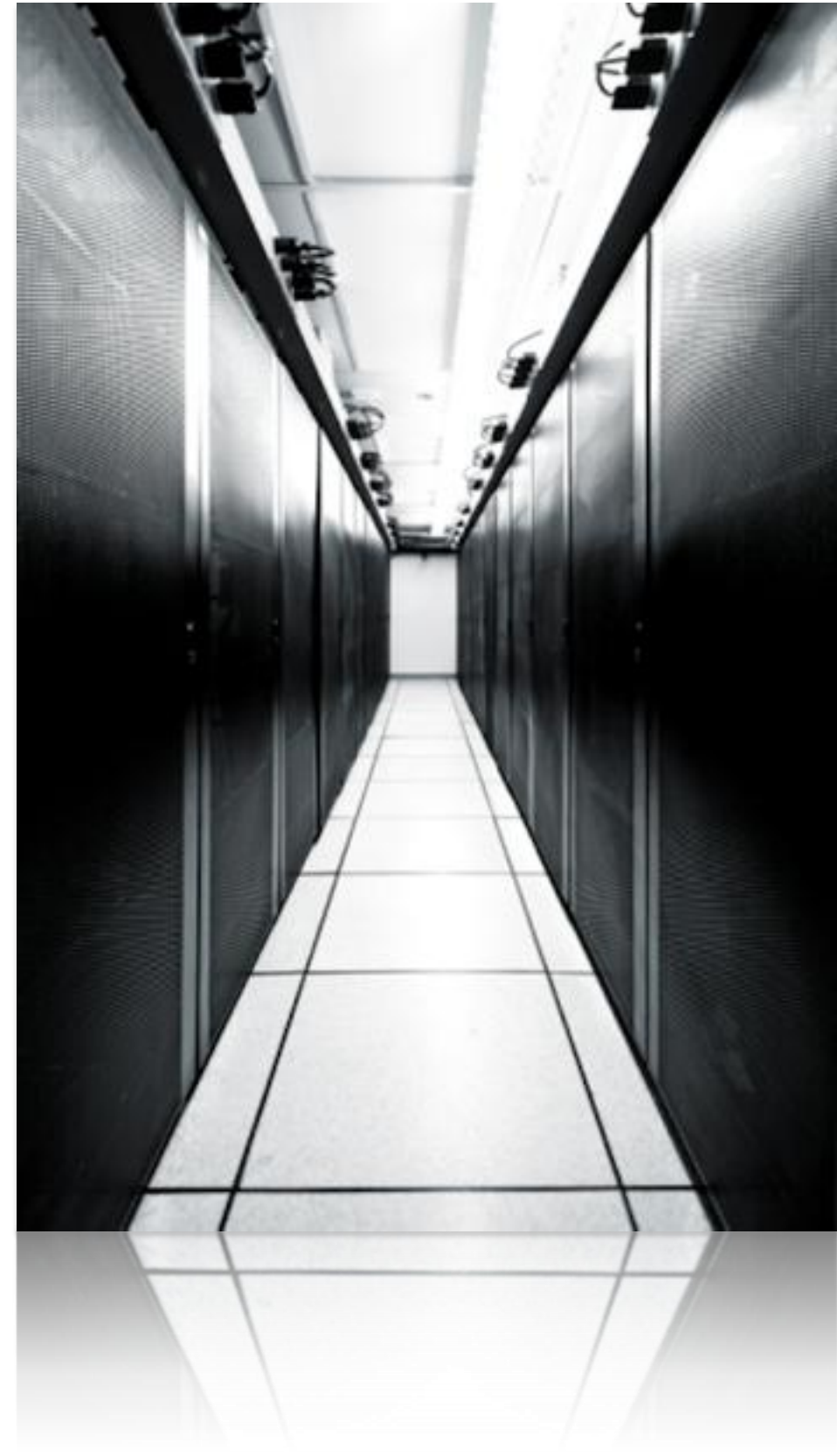
## *Benefits*

- Fewer Cables  
Both block I/O & Ethernet traffic co-exist on same cable
- Fewer adapters needed; less power
- Interoperates with existing SAN' s  
Management of SAN' s remains constant
- No Gateway



# Why FCoE?

- Because FCoE...
  - Integrates with today's Fibre Channel SANs
  - Enables "Converged Technology"
  - Is based on Ethernet, leveraging technology, investment, market presence, scaling ability
  - Invites more user choice
  - Enables FC to become more accessible





# Technical Considerations

- Expanding and Growing the Data Center
- Multitier FCoE
- Multihop FCoE





# The Big Daddy: FCF

- Fundamental piece of FCoE
- Fully functional Fibre Channel switch inside an Ethernet switch
- Has its own Domain ID
- Performs all FC logins, services, zoning
- Without this, you ain't got no FCoE



Source: BioShock's "Big Daddy"  
©2008 Take-Two Interactive





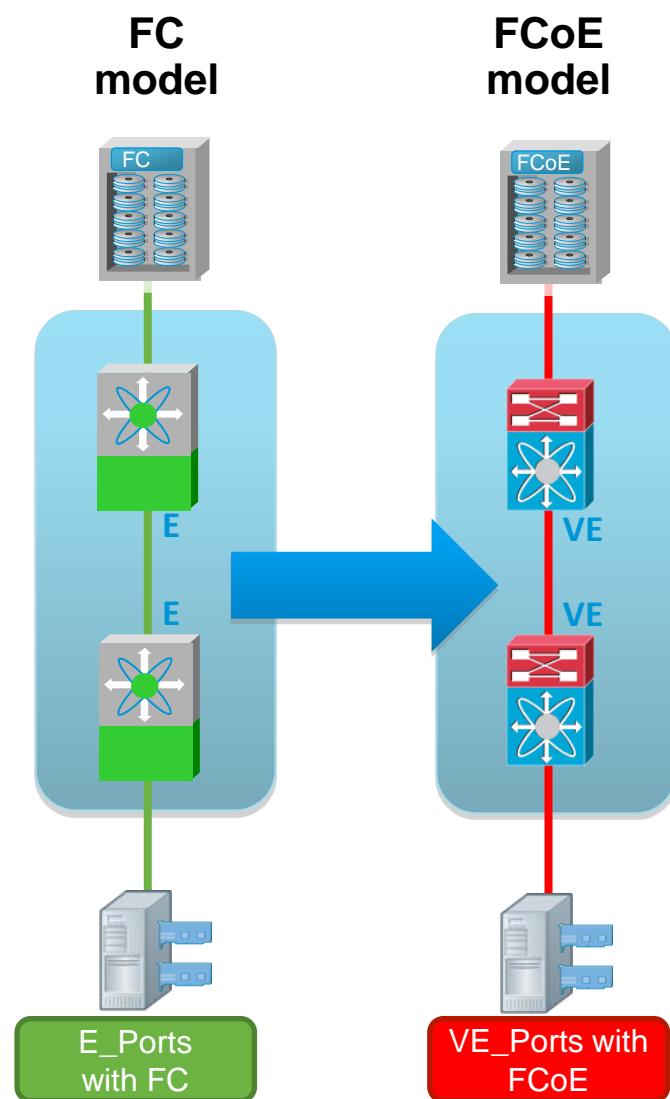
# Why Port Type Matters

- FCoE is, by definition, Fibre Channel
  - Therefore uses the same types of terms
  - The ports are called "Virtualized" in FCoE

Location	FC Name	Definition	FCoE Name
Host port that talks to switch	"N_Port"	"Node" port	"VN_Port"
Switch port that talks to host	"F_Port"	"Fabric" port	"VF_Port"
Switch port that talks to port	"E_Port"	"Extender" port	"VE_Port"

- In Fibre Channel, an "E\_Port" makes a "hop"; in FCoE, a "VE\_Port" makes a hop

# GeekSpeak: VE\_Ports



- **E\_Ports**: used to “expand” a **FC** fabric by connecting multiple Fibre Channel switches together
- **VE\_Ports**: used to “expand” a **FCoE** fabric by connecting multiple Fibre Channel over Ethernet switches together
- Creates a standards-based FCoE ISL - a FC "hop"
- Same scalability as FC





# Expansion and Growth

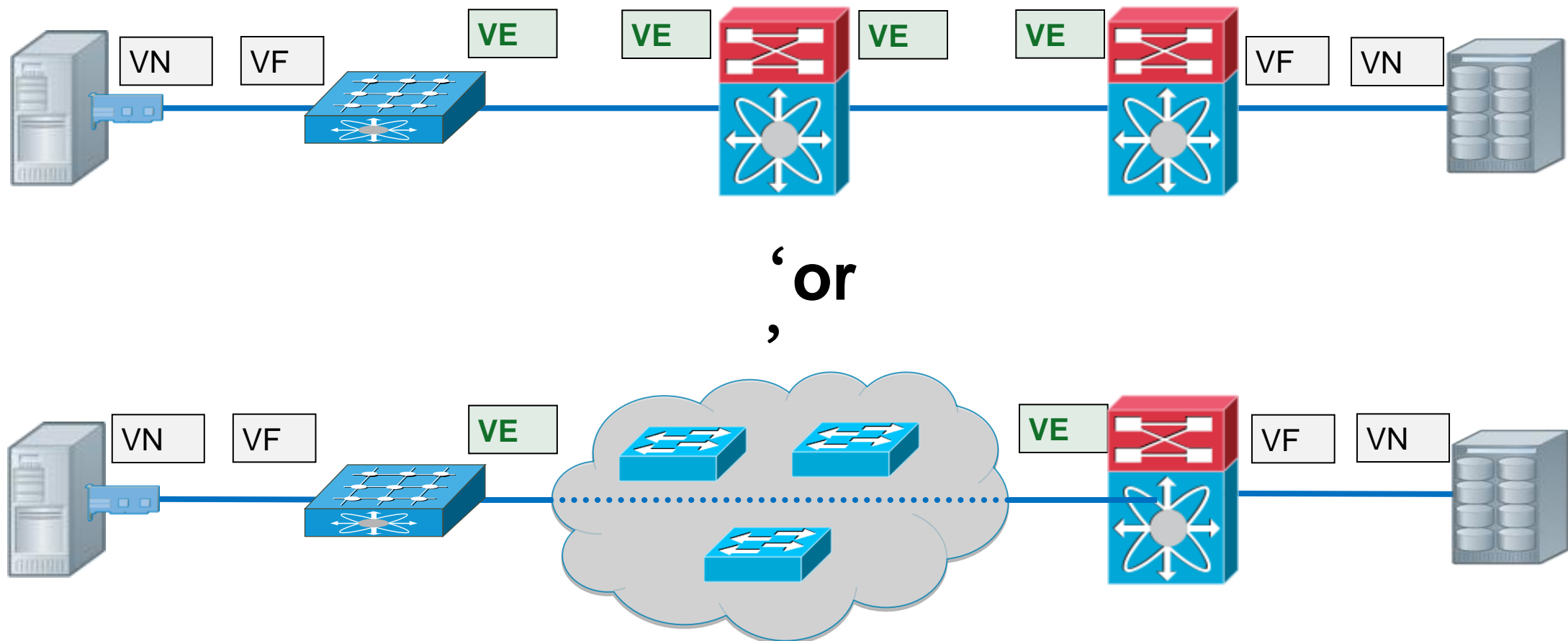
- Many ways to expand the network:
  - Additional ports, cards, switches
  - But... for Data Center-sized solutions?
  - How do you *scale*?
- Storage Questions:
  - How do you handle load-balancing?
  - How do you handle SAN A/B separation?
  - How much visibility do you have in the fabric?





# Theory vs. Practice

- What's going to make it easier in real life?
  - Do you want a 'routed' topology or a 'bridged' topology?
  - Is FCoE a layer 2 overlay or integrated topology?
  - How do these design affect operations?
  - What looks simple can be deceiving: there are wires within that cloud that need to be connected
  - More than just connectivity; you're running your business on it



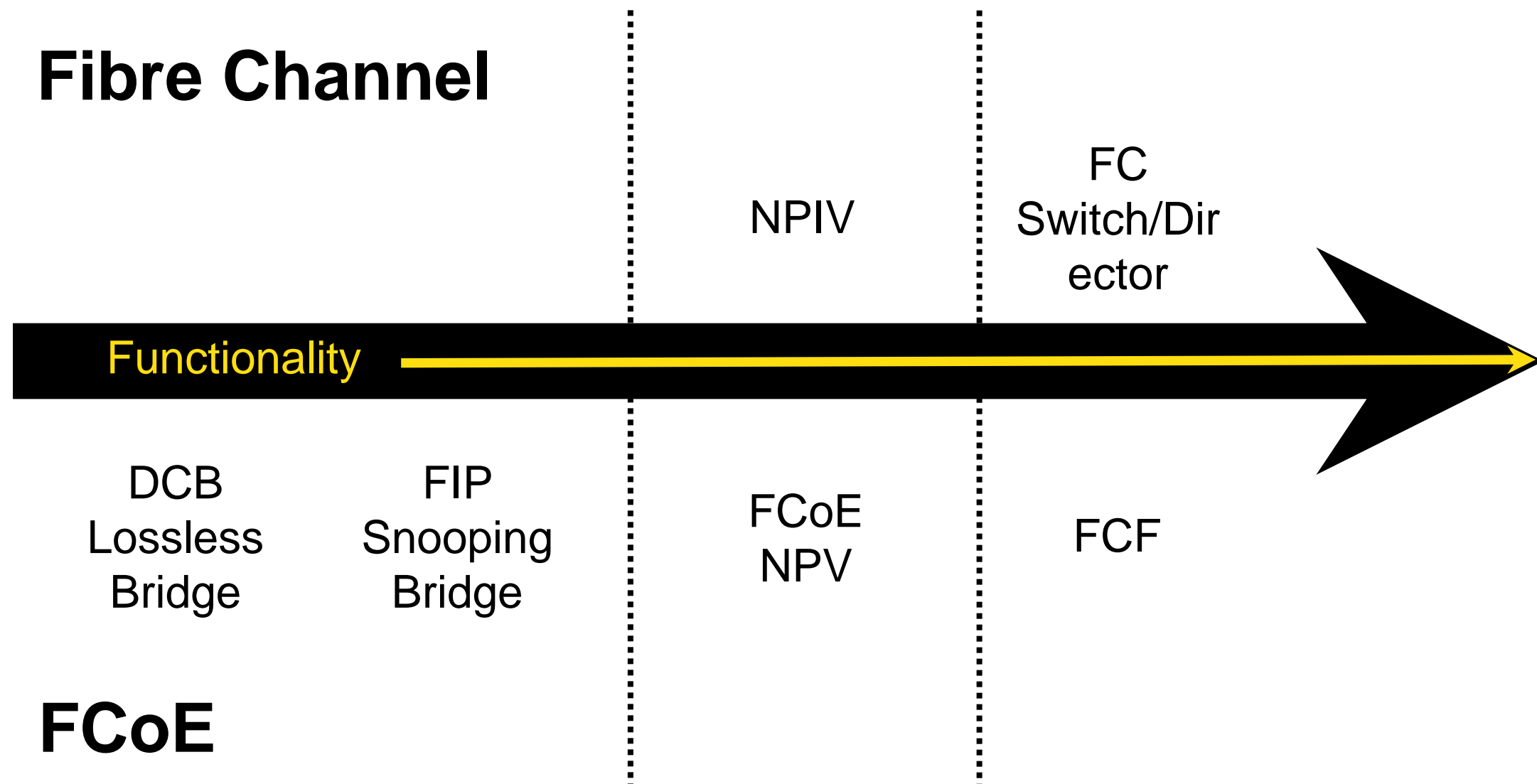




# Functionality Equivalence

- Rough comparison (not to scale)
- Increasing degrees of fabric functionality and control

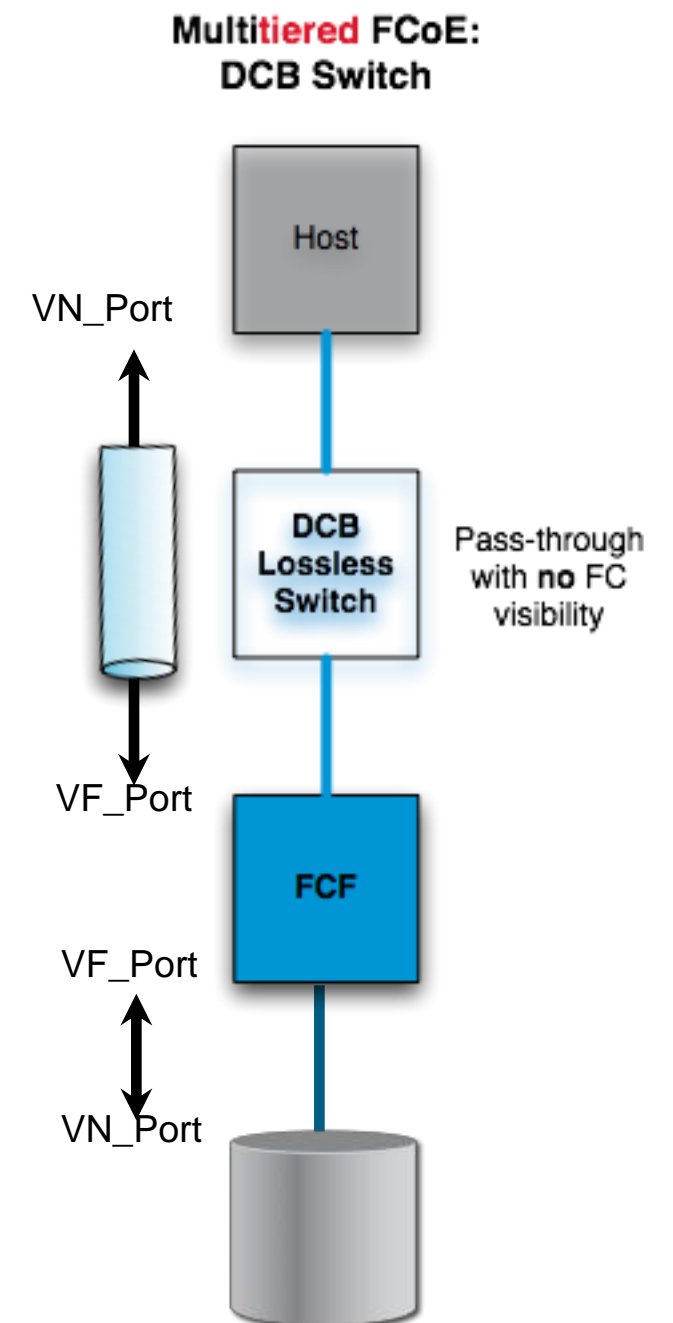
## Fibre Channel





# Multitier FCoE: Lossless Ethernet Bridge

- Paradigm: Ethernet Fabric
- What it is
  - Tunneling between through storage-unaware switch(es)
  - No knowledge of the packet type
  - No FC protocol-related activities applied to traffic between initiator and FCF

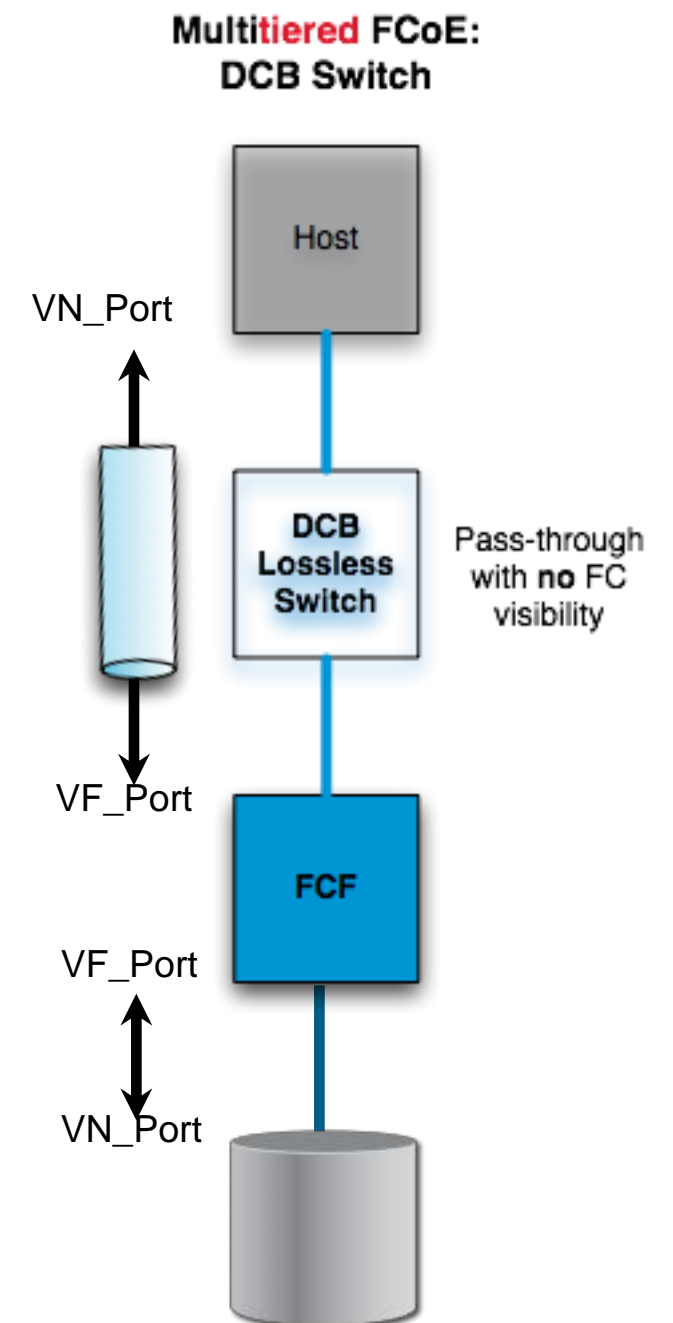






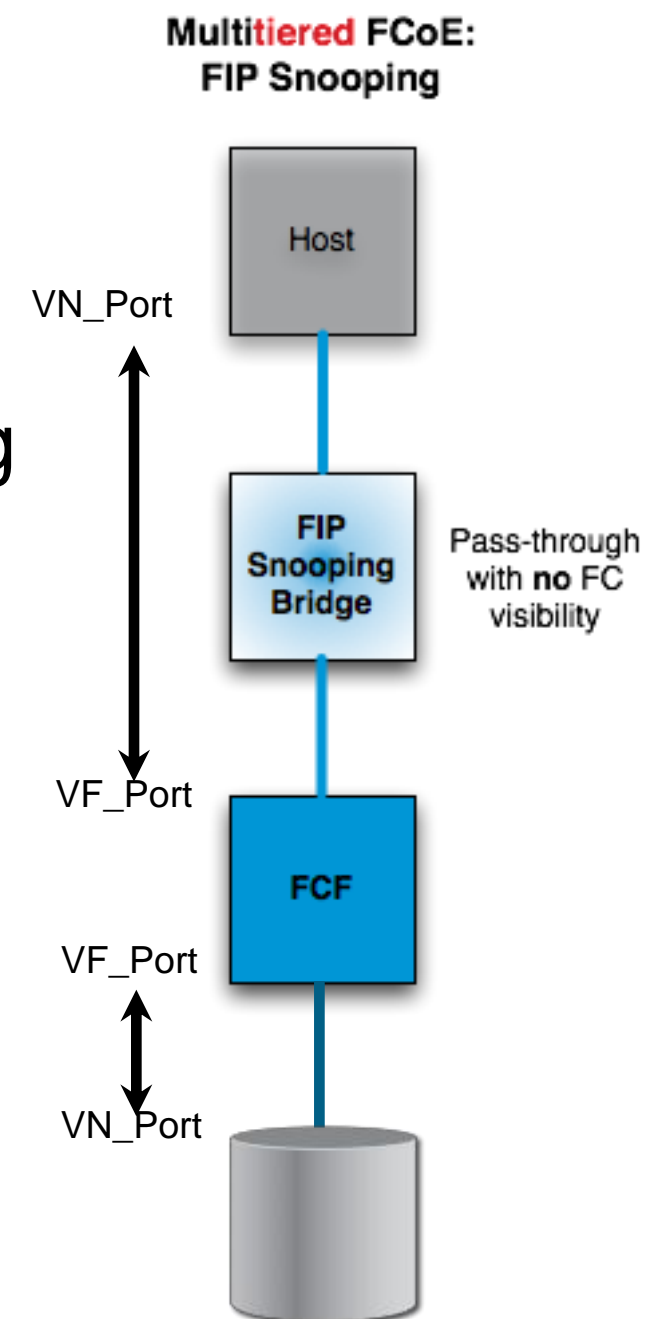
# Multitier FCoE: Lossless Ethernet Bridge

- Pros
  - Hosts get access to FCoE storage
  - Hosts get access to FC storage... maybe
  - Expands fabric size without expanding Domain IDs
- Cons
  - Not a FC model
  - No complete visibility into FC fabric
  - Troubleshooting and load balancing is difficult, especially as number of DCB switches increase
  - Cannot use FC forwarding or deterministic multipathing technology, must rely on Ethernet L2 solutions



# Multitier FCoE: FIP Snooping

- What it is
  - Used during login process; after FLOGI has no interaction with FC fabric
  - Non FCoE switch "snoops" packets to discover FC logins
  - Inspects packets and applies routing policies based on those frames
  - Uses dynamic ACLs to enforce FC rules within DCB network
  - Prevents nodes from seeing or communicating with other nodes without first going through an FCF

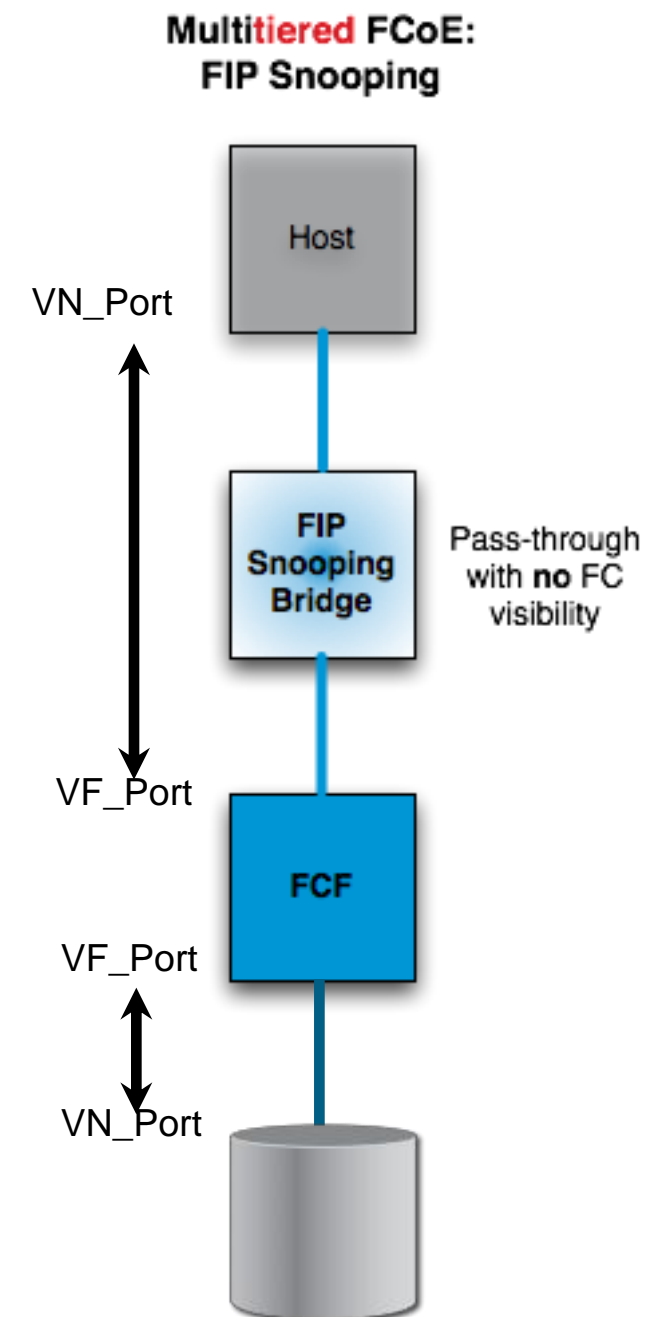






# Multitier FCoE: FIP Snooping

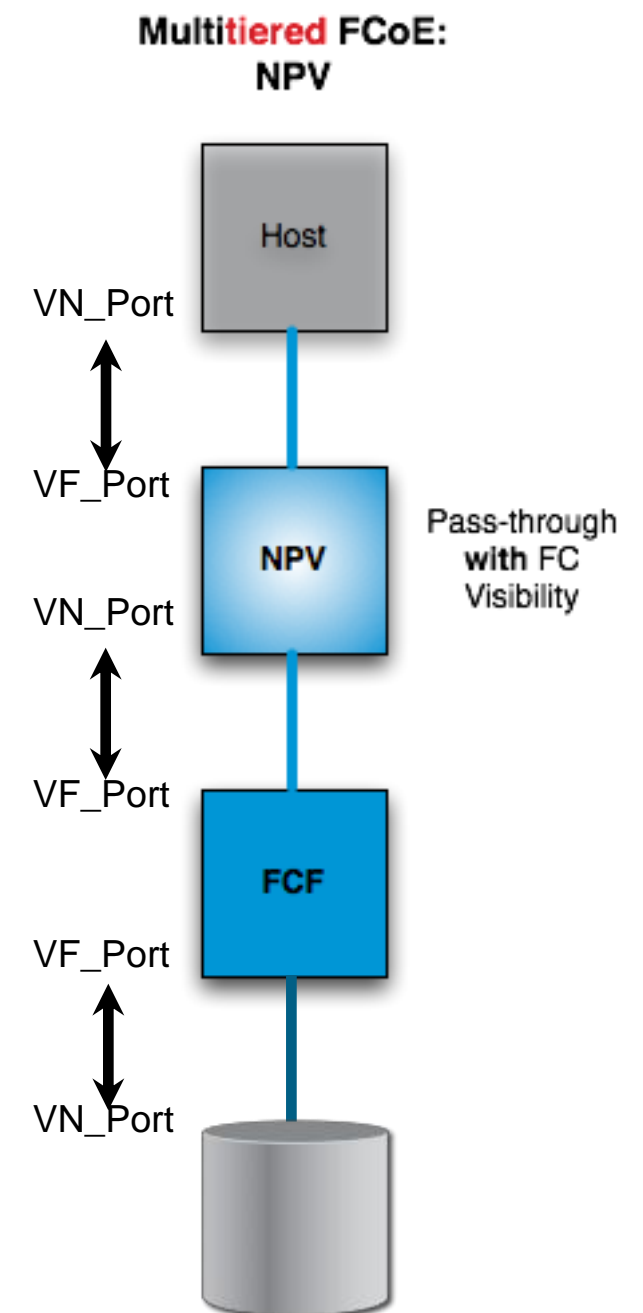
- Pros
  - Enhances FCoE security (helps prevent FCoE MAC spoofing)
  - Creates FC point-to-point links within Ethernet LAN
  - Allows auto-configuration of ACLs based on name server information in FIP frames
  - Expands fabric size without expanding Domain IDs
- Cons
  - Troubleshooting can be an issue
    - FC tools don't see FIP snooping bridges, and FIP snooping bridges don't track discovery attempts or login failures
    - FC fabric tools are unaware of CNA failures that prevent logins, must rely on CNA tools for troubleshooting
  - Potential load-balancing, SAN A/B separation issues





# Multitier FCoE - FCoE NPV

- FCoE switch using NPIV logins to an FCF
- What is it?
  - Enhanced FCoE pass through
  - FCoE switch acts like a server, performing multiple logins
  - Load balancing, traffic engineering, FCoE security, FC operational SAN model
  - Addresses FC Domain ID sprawl
  - Offers familiar management and troubleshooting to SAN admins

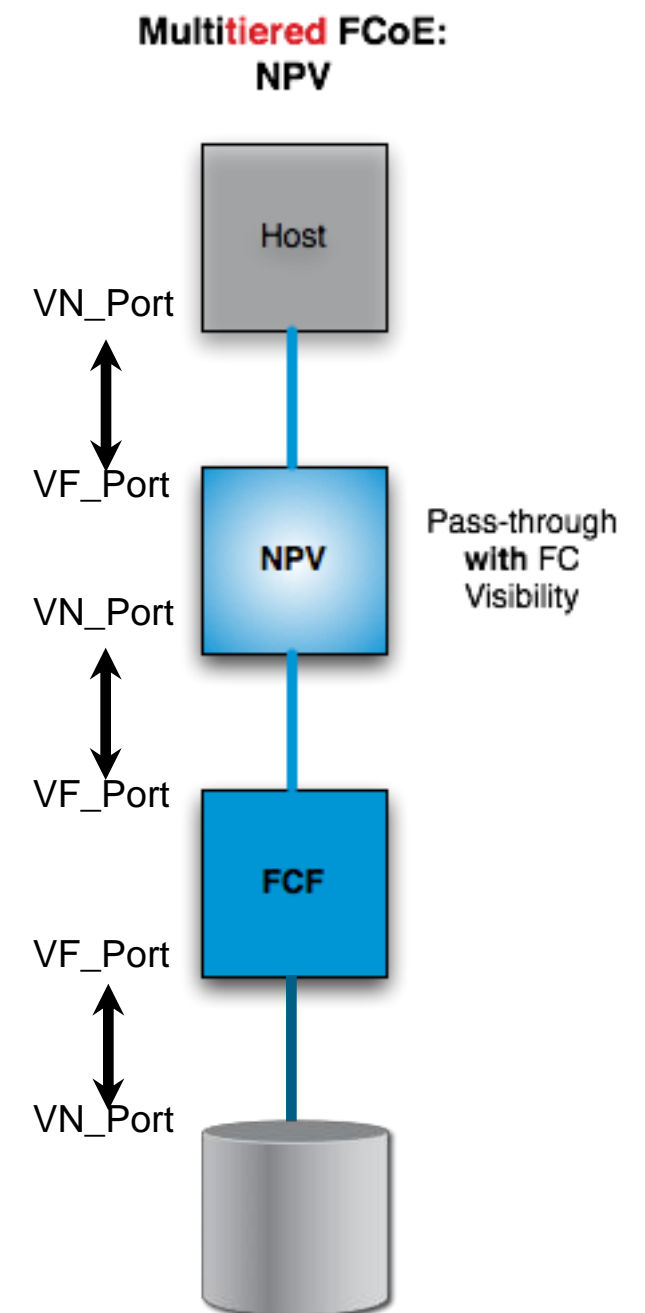






# Multitier FCoE - FCoE NPV

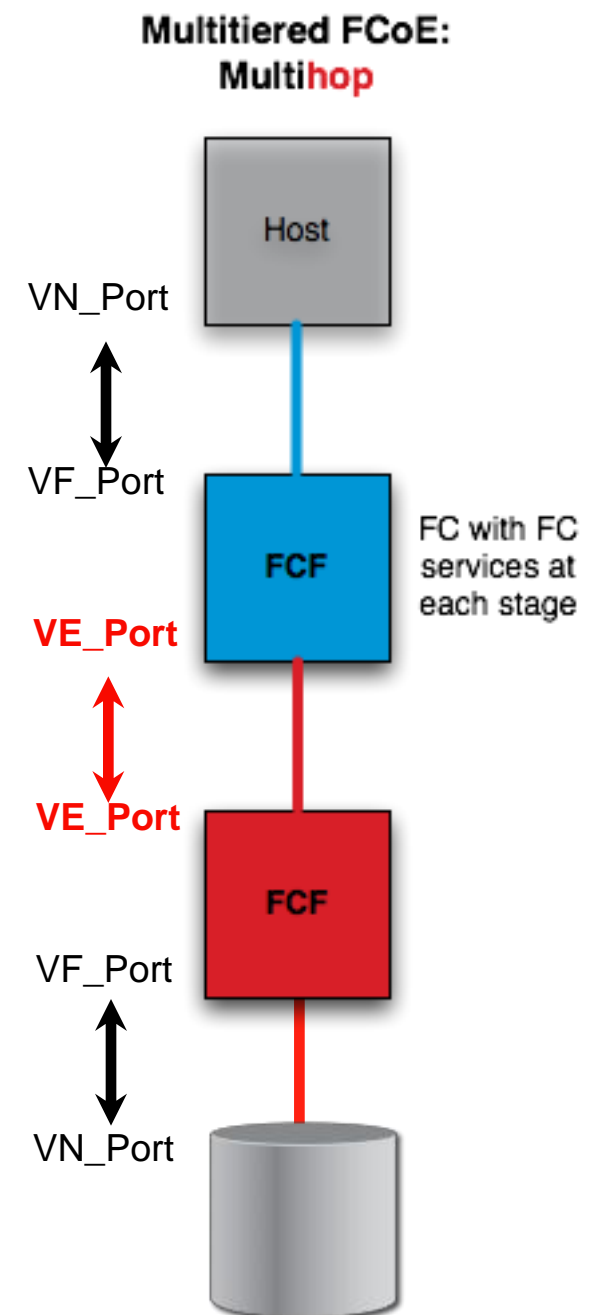
- Benefits
  - Provides same benefits as FC switches doing NPIV logins
  - Does not use a domain ID
  - Doesn't lose visibility into FC domains and zoning
- Difference from FIP-snooping
  - Makes storage traffic visible to the SAN admins for optimization, monitoring, troubleshooting
  - Limits the expectations from CNAs
  - Offers greater traffic engineering options





# Multitiering: Multihop FCoE

- Same model as FC networks today
- FCF-to-FCF, Switch-to-Switch communication
- Full, complete visibility into FC fabric at every hop
- Consistent with FC "hop" rules
- FC-BB-5 standard for FCoE Switch-to-Switch interoperability

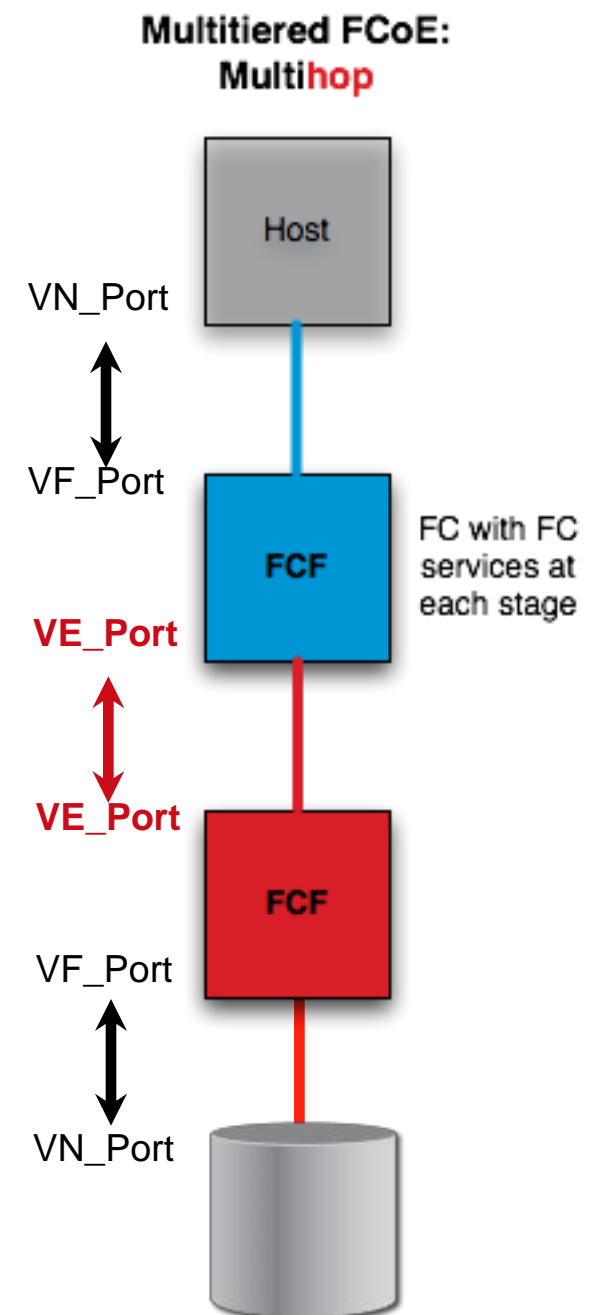






# Multitiering: Multihop FCoE

- Pros
  - Nothing further (TRILL, FabricPath, etc.) necessary for implementing “multihop” FCoE
  - Most control for SAN admins
- Cons
  - Risk of FC Domain ID sprawl





# Best Practice Considerations ... Storage Perspectives

- What design considerations do we have when extending FCoE beyond the access layer?
- SAN A/B separation
  - How does multipathing concerns affect separation?
  - Not as simple as building two separate Ethernet fabrics, which must be one contiguous fabric
  - How do you maintain a single Ethernet fabric *and* separate storage fabrics?
- High-availability
- Oversubscription for SAN and LAN

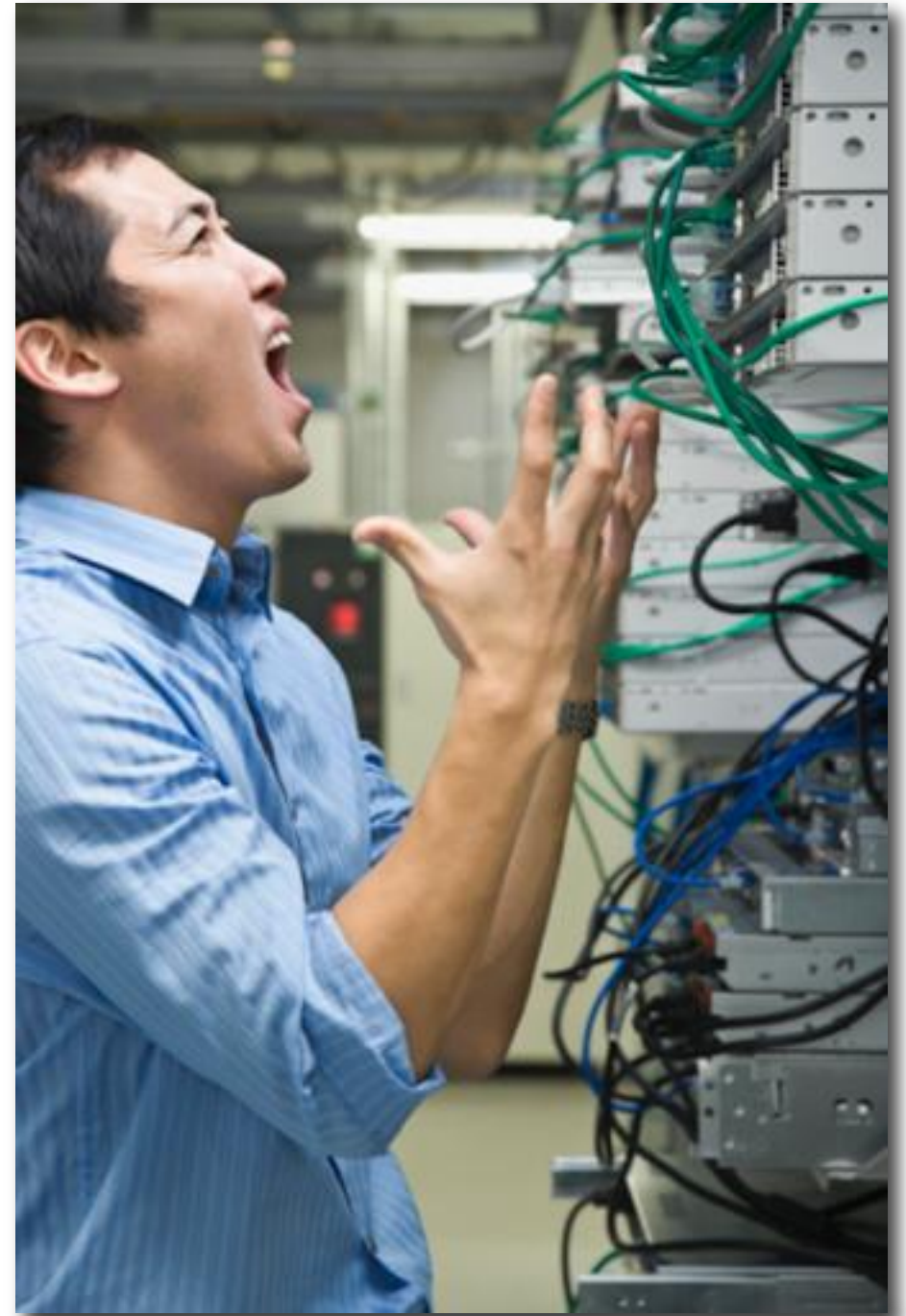






# Organizational Considerations

- You don't want to be like this guy ---->

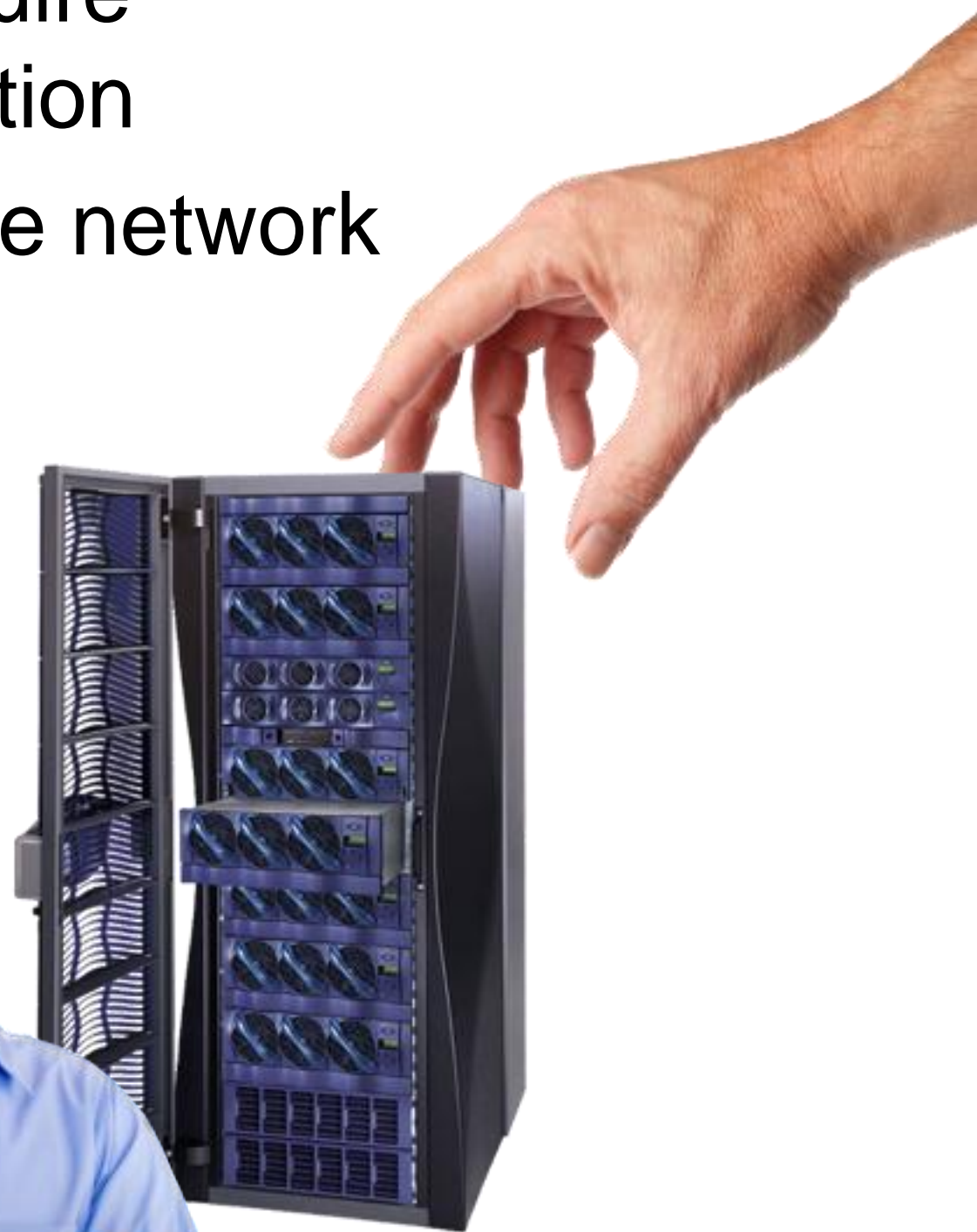






# Different “Tribe” Perspectives

- Unified Fabric will require additional communication
- Fear of takeover by the network team or storage team
- Fear of operational changes
- Consider “SLAM” team





# The Realities of Unified Fabric

- Management should stay with the relevant team
- Skills don't change
- RBAC: Role-Based Access Controls
  - Both teams still control their domains
  - Provides administrative protection to each team
  - Provides operational visibility to both teams
- Cross-functional teamwork increases







# Summary



- Storage assumptions are not the same as Ethernet assumptions
  - Make sure that storage priorities are maintained when planning
- Growth and expansion
  - Different methods of expanding Ethernet networks risk breaking storage methods
- Multitier vs. Multihop
  - Different mechanisms have different consequences to FC tools and operational models
- Organizational Considerations
  - It's all about the people, people!





# Thank You!

J Michel Metz, Ph.D  
Product Manager - FCoE  
Cisco Systems  
[jmmetz@cisco.com](mailto:jmmetz@cisco.com)