



# UM DIA NA VIDA DE UM PACOTE CEE

Marcelo M. Molinari

System Engineer - Brazil

May 2010



# CEE (Converged Enhanced Ethernet) Standards

## Making 10GbE Lossless and Spanning-Tree Free

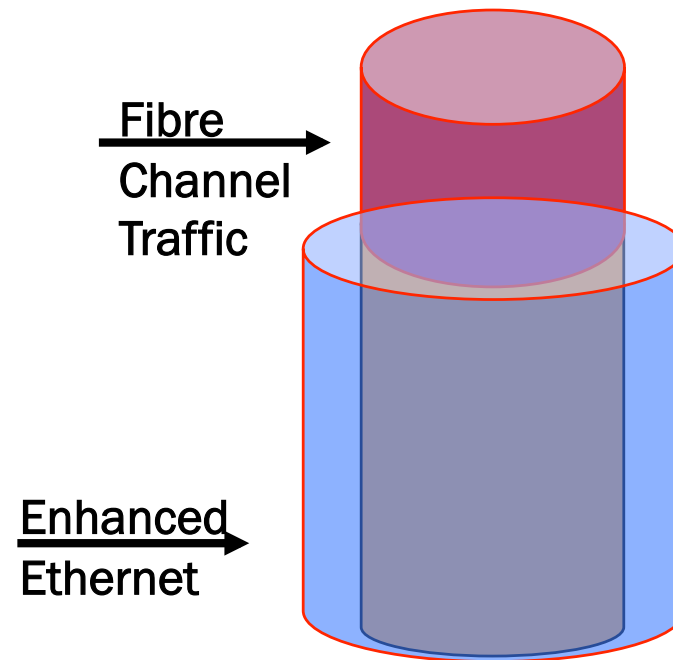
- **T11 FCoE & FIP standards**
  - FCIA approved standards
  - INCITS approved and expected to publish soon
- **IEEE draft DCB/CEE components**
  - 802.1Qbb priority-based flow control
  - 802.1Qaz enhanced transmission selection
    - DCBX capability exchange protocol
  - 802.1Qau Congestion Notification
  - Completion & approvals expected H2 of 2010
- **IETF status**
  - Transparent Interconnection of Lots of Links (TRILL)
  - Expected completion mid 2010



# FCoE: *Bird's Eye View*

FCoE straddles two worlds: Fibre Channel and Ethernet

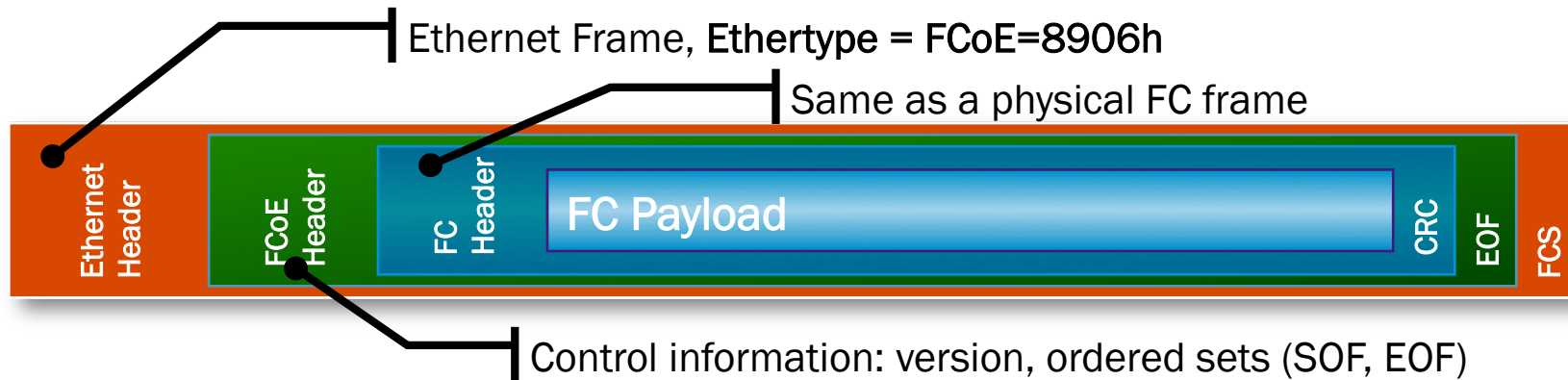
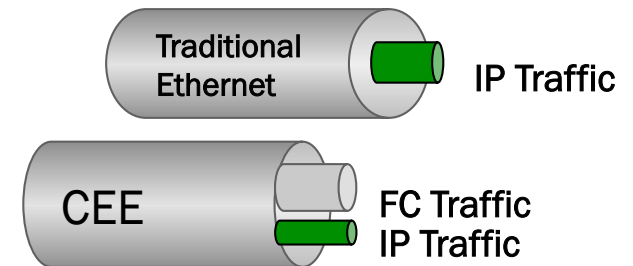
- **FC View:** FC gets a new transport in the form of lossless Ethernet
- **Ethernet view:** a new upper layer protocol, or storage application, that runs over a new lossless Ethernet



# Fibre Channel over Ethernet

## Encapsulation protocol

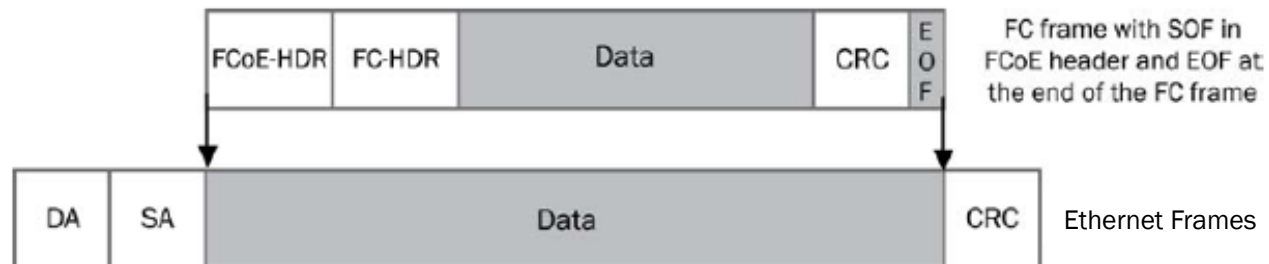
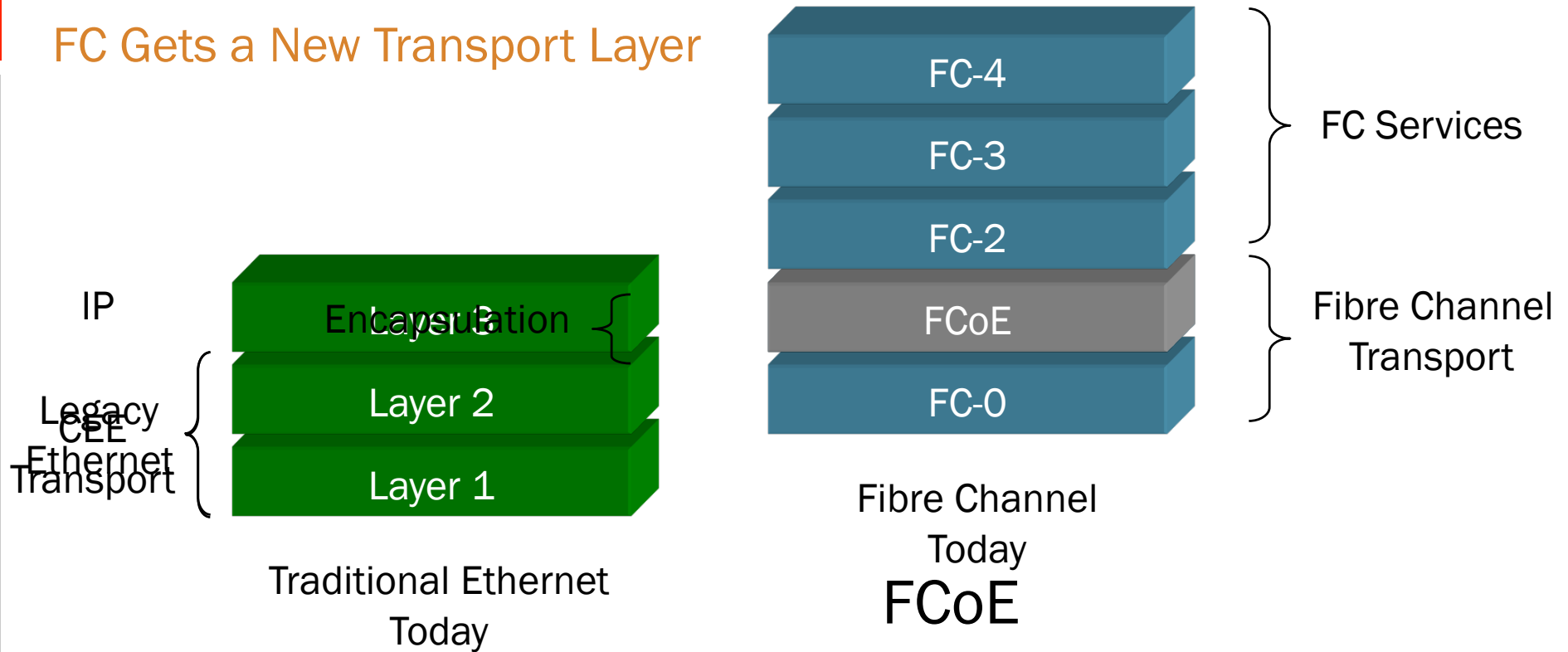
Encapsulation protocol for transporting FC over Ethernet (Lossless Ethernet: CEE)



- FC frame remains intact: FC does not change
- Ethernet needs a larger frame: Larger than 1.5 KB
- Ethernet must become lossless to carry storage data with integrity

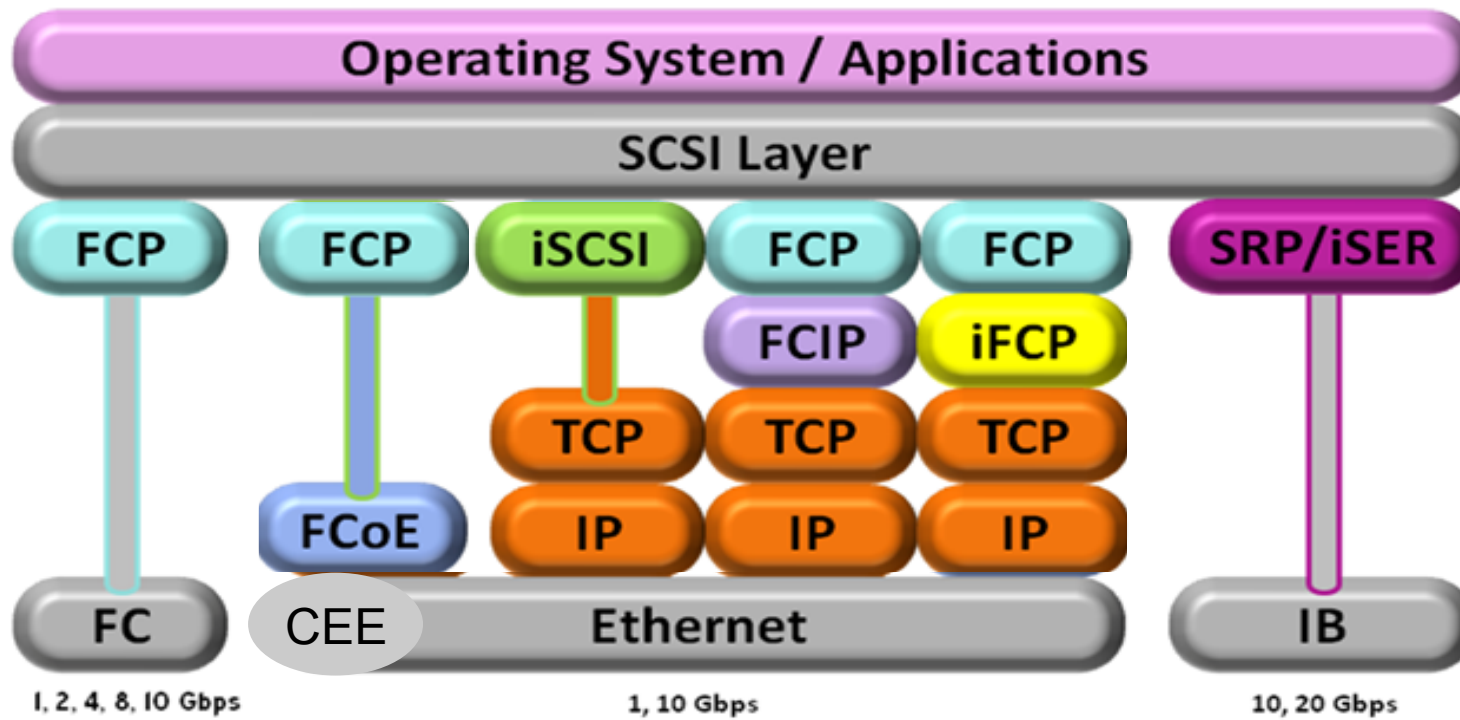
# FCoE

## FC Gets a New Transport Layer



# FCoE and Other Protocols ISO Layers

## Encapsulation v. Native Protocols

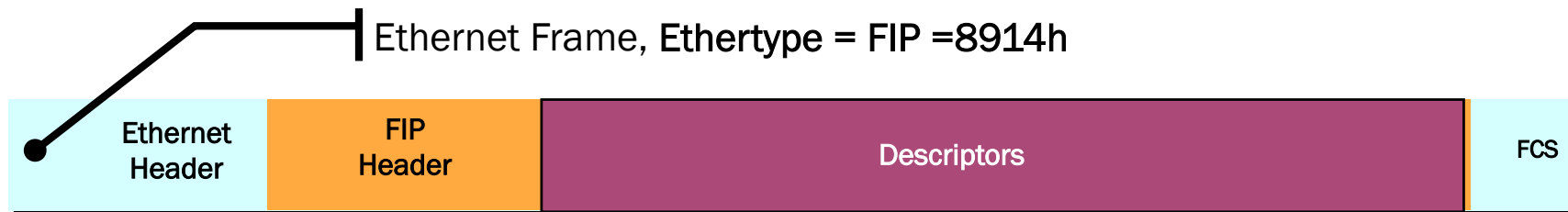


- FCoE is an encapsulation protocol with less overhead than iSCSI but more overhead than FC

# FCoE Protocol Organization

## Data Plane & Control Plane

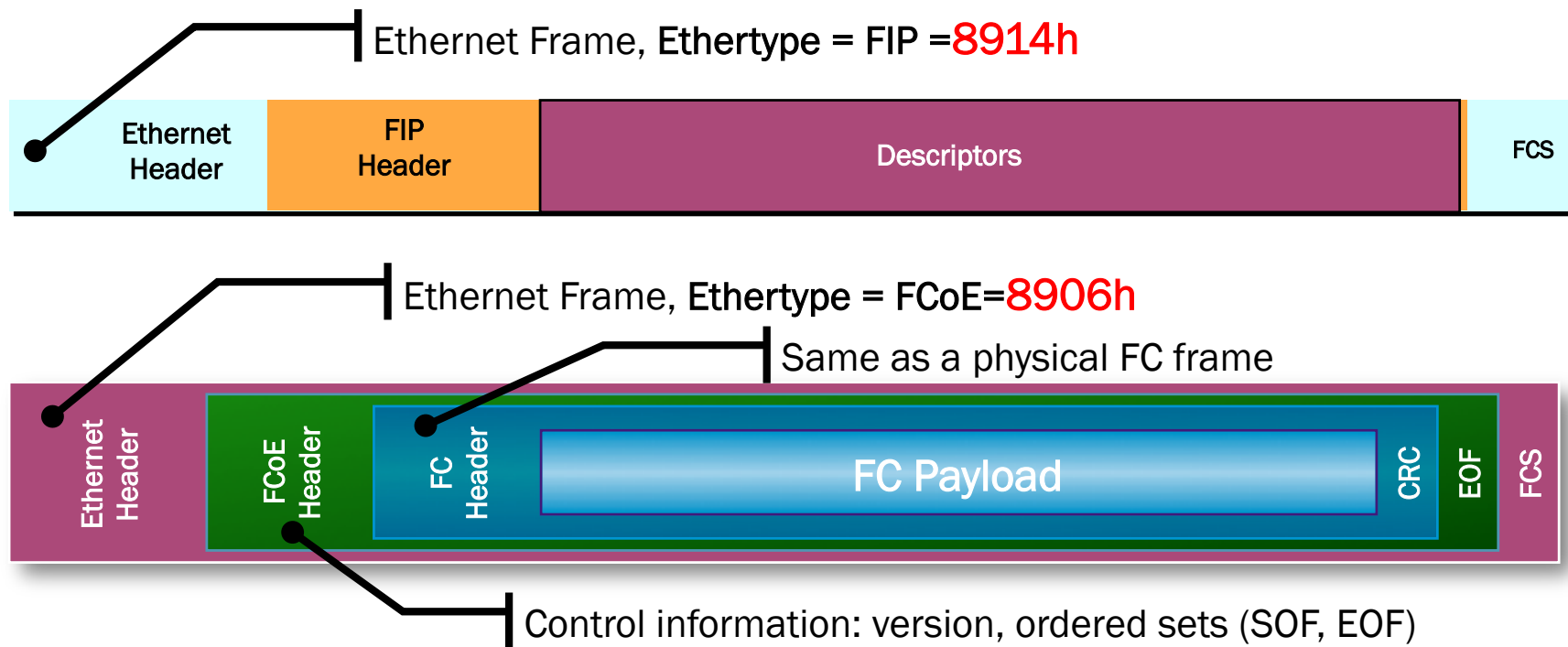
- FCoE uses two different Ethernet Packets:
  - FCoE Initialization Protocol (FIP)
  - FCoE Data Plane (FCoE)
- FIP is the control plane protocol:
  - Used for login to and logout off the FC fabric
  - Discovers FC entities connected to the Ethernet fabric
  - Associates Fibre Channel FC\_IDs and MAC addresses
  - Clear links and keep virtual links alive in an Ethernet L2 cloud
- FCoE data plane packets are used in the data transfer phase



# FIP & FCoE Frames

## Data & Control Planes

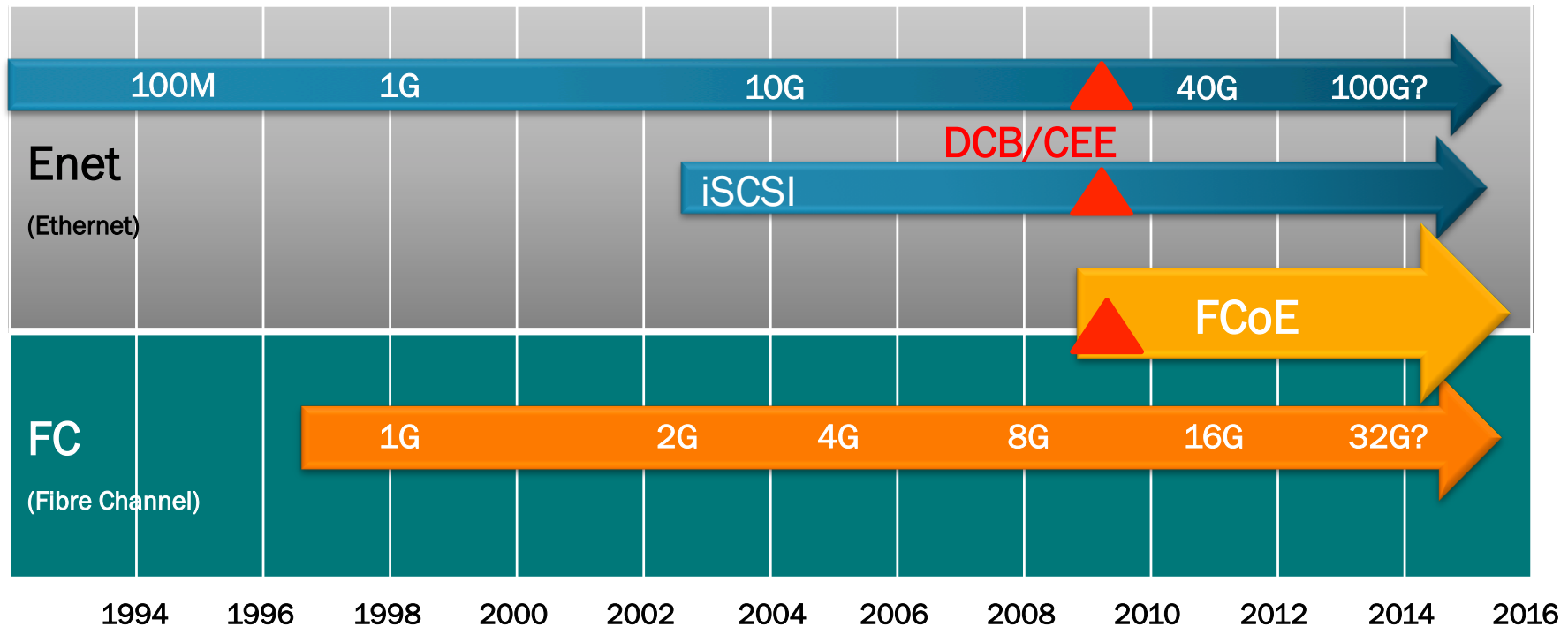
- Different Ethertype
- Only FCoE carries FC data





# Ethernet and FC Roadmaps

## Parallel Evolution & Potential for Convergence



- FC and Ethernet evolved in parallel paths with FC dominating storage SANs and Ethernet supporting IP networking
- Lossless Ethernet & FCoE open the door for server I/O consolidation



# Making Ethernet Lossless

## Converged Enhanced Ethernet (CEE)

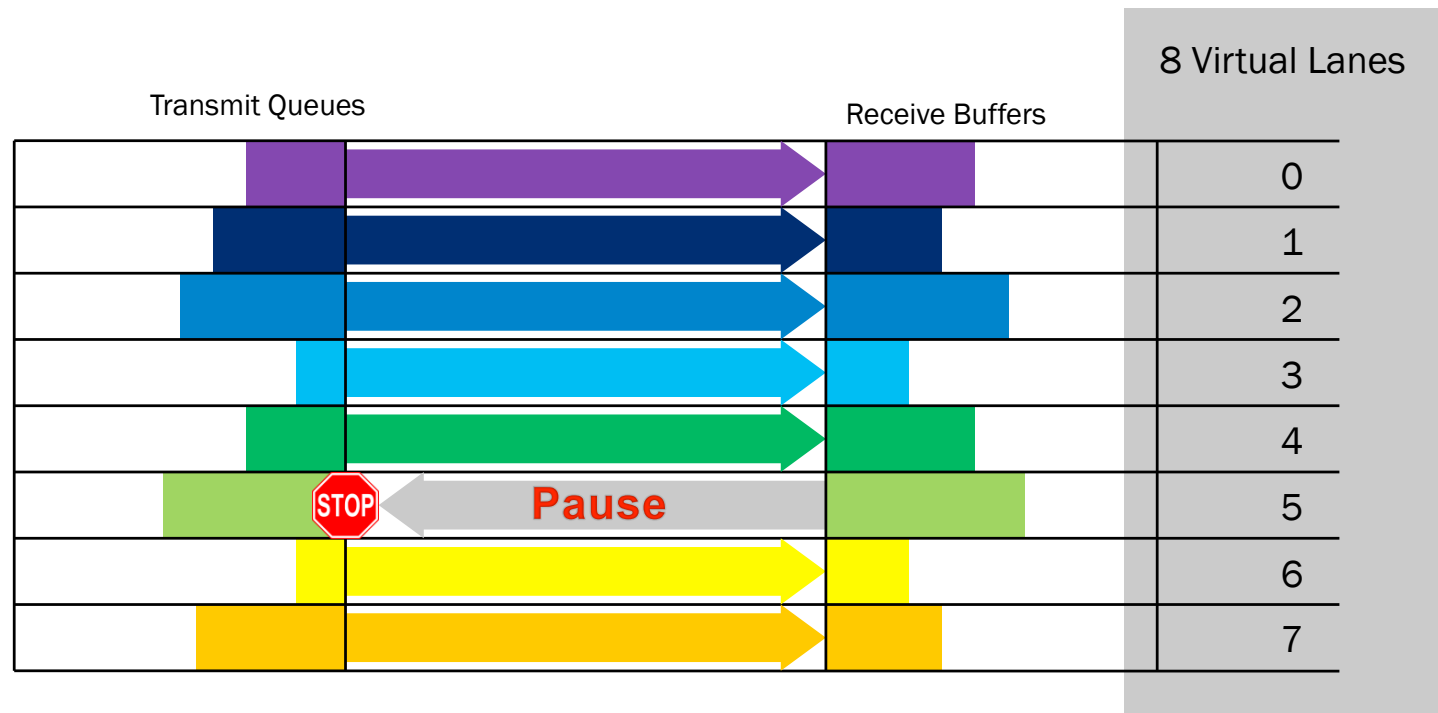
- 802.1Qbb – Priority-Based Flow Control
  - PFC: Allows identification and prioritization of traffic
- 802.1Qaz – Enhanced Transmission Selection
  - ETS: Allows grouping of different priorities and allocation of bandwidth to PFC groups
- 802.3Qau – Congestion Notification
  - CN: Signals sources of congestion



# 802.1Qbb – PFC

## Priority-based Flow Control

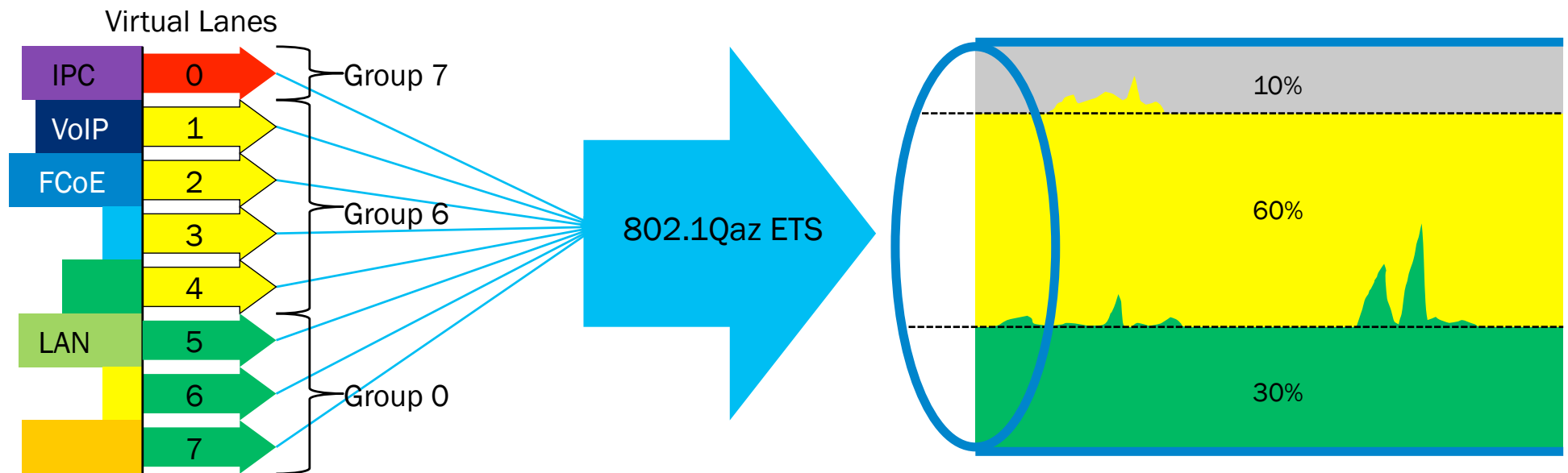
- During periods of heavy congestion
  - Ensures delivery of critical data
  - Latency sensitive traffic continues normal operation



# 802.1Qaz – ETS

## Enhanced Transmission Selection

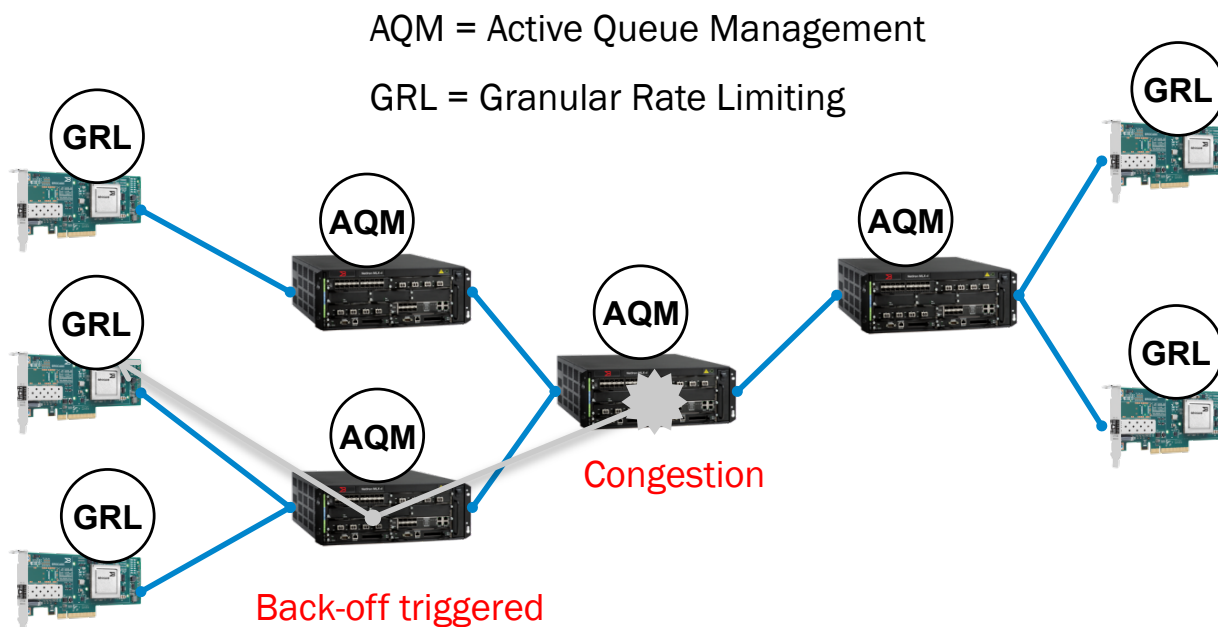
- Capability to apply differentiated treatment to different traffic within the same traffic class



# 802.1Qau - CN

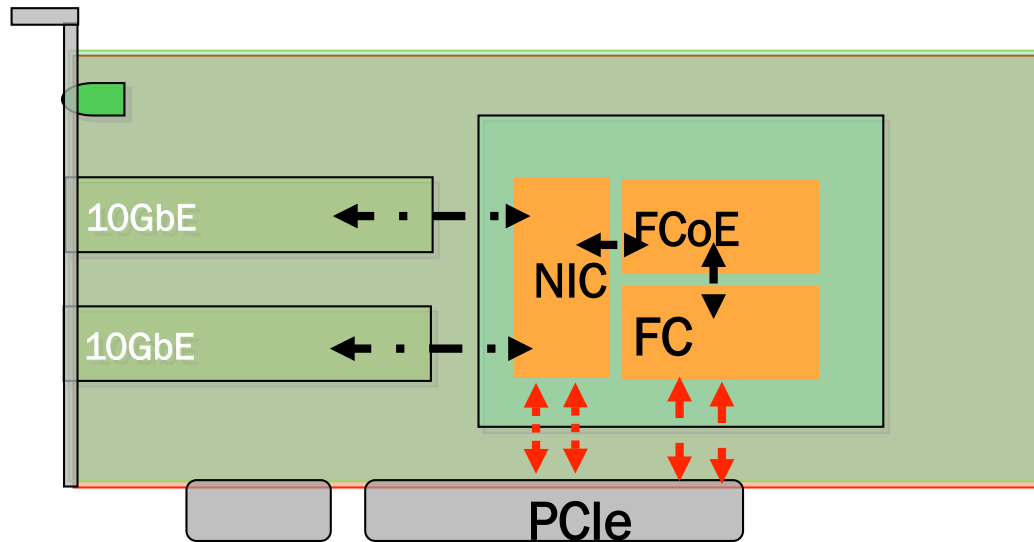
## Congestion Notification

- Signals congestion to the source allowing for lossless behavior.

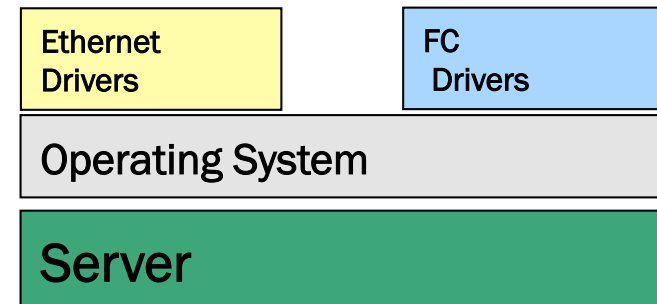
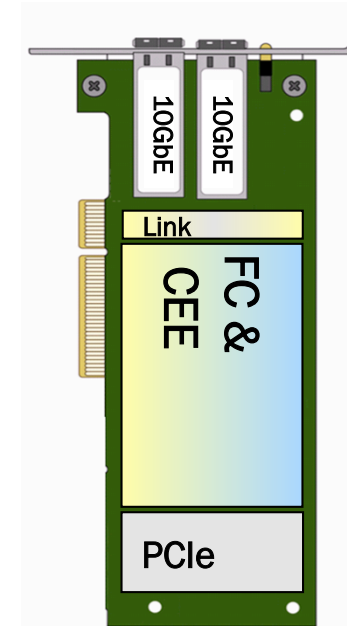


# Model of FCoE Initiator

## Converged Network Adapters (CNAs)



- One or more FC interfaces to the host
- One or more NIC interfaces to the host
- FCoE function: Not seen by the host:
  - Performs the encapsulation and de-encapsulation

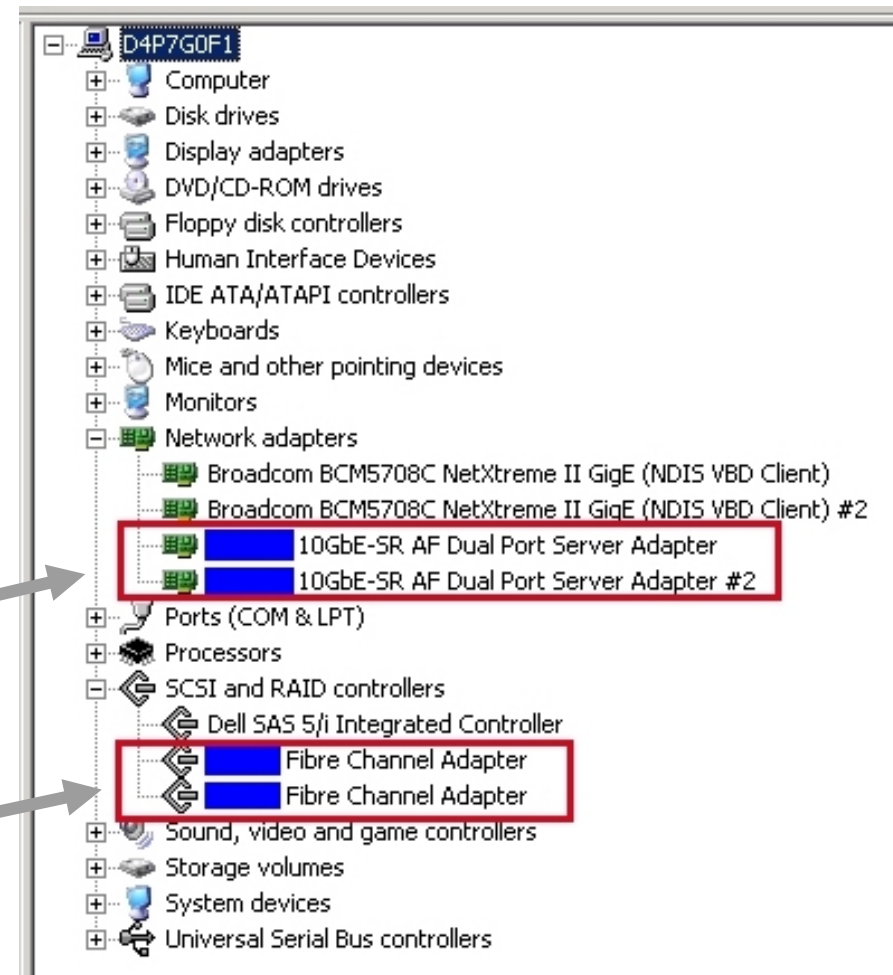


# Server OS View of CNAs

No Change in OS View

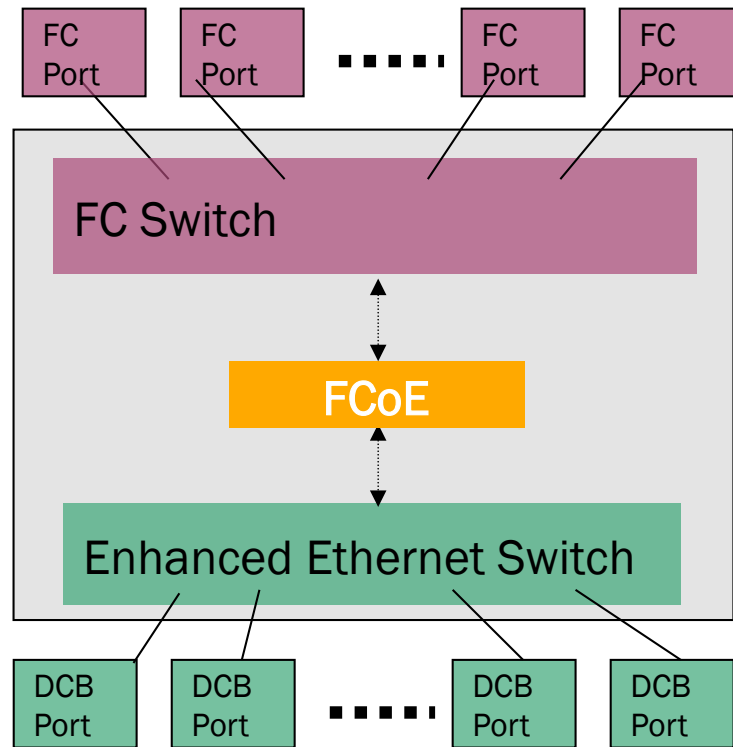
- Standard drivers
- Same management
- Operating System sees:

- Dual port 10 Gigabit Ethernet adapter
- Dual Port 4/8 G Fibre Channel HBAs



# CEE/FCoE Switch Model

## Brocade 8000 Switch



- FC Switch is the FCoE/FC forwarding entity
- Has multiple FC ports that connect to SAN
- Has multiple 10GbE ports that connect to messaging network
- FCoE entity that handles FCoE/FC encap/decap function
- Enhanced Ethernet Switch handles LAN and SAN traffic
  - Forwards FCoE traffic to FCoE entity
  - Forwards Enet traffic to Enet ports





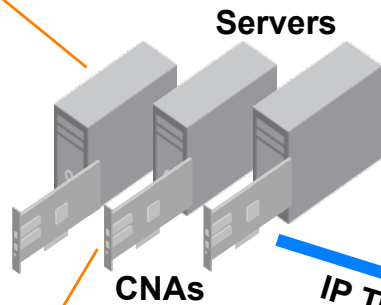
# Walk Through



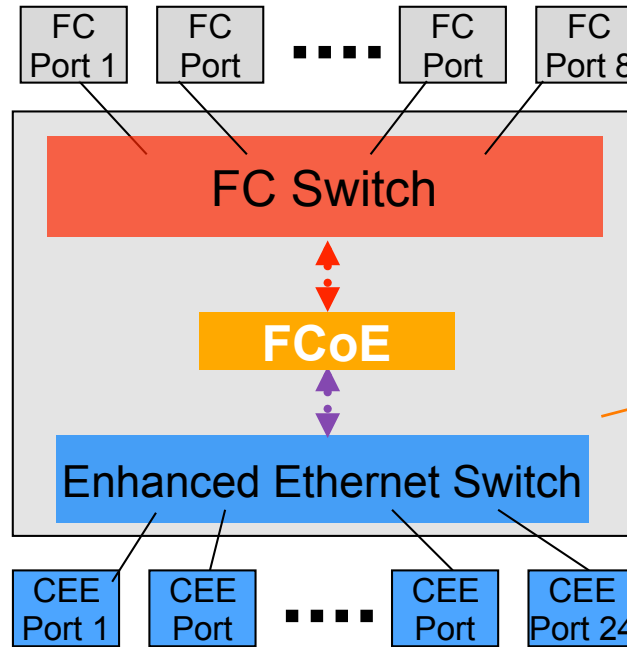
# Brocade 8000 Switch

## CEE/Ethernet Traffic

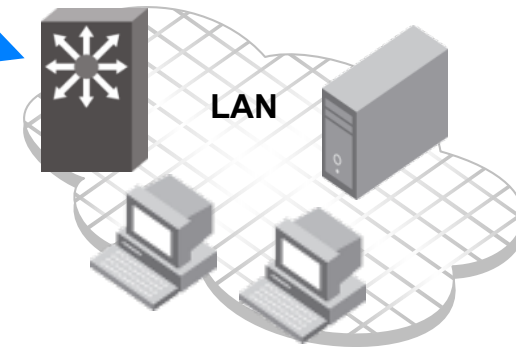
1: Client or virtual application initiates networking activity – let's say IP traffic



2: CNAs encapsulate data into Ethernet (CEE) frames and send to switch

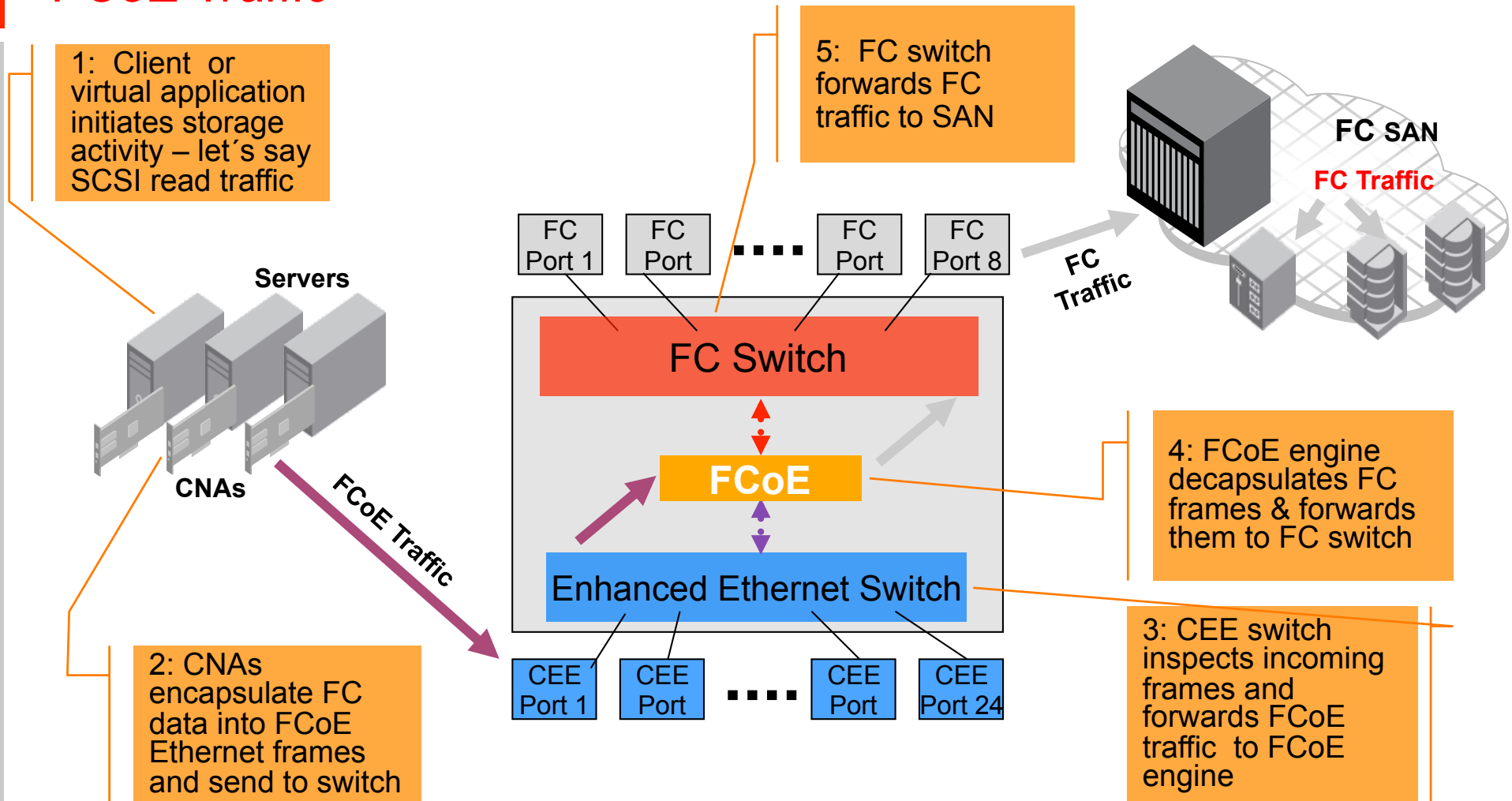


3: CEE switch inspects incoming frames and forwards CEE traffic



# Brocade 8000 Switch

## FCoE Traffic



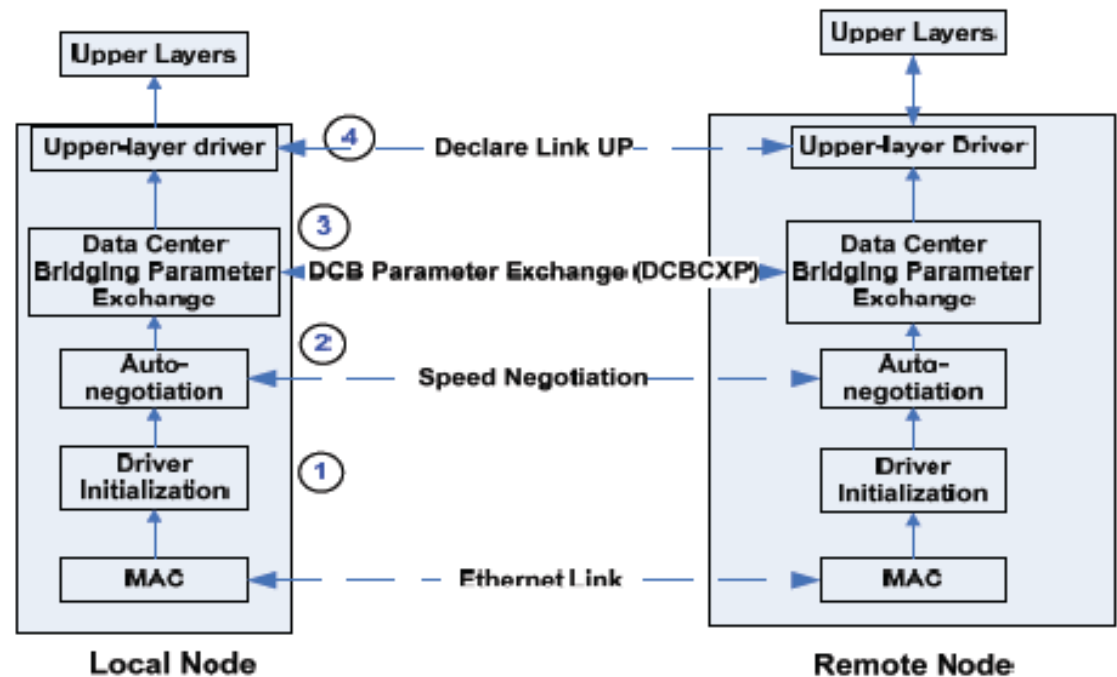
# So how does CEE gets initialized?

- Data Center Bridging Exchange (DCBX)
  - DCBX enables CNAs to receive their ETS and PFC settings from the switch port
  - DCBX is used by IEEE 802.1Qaz and IEEE 802.1Qbb
  - DCBX is based on LLDP
- This way you don't have to program ETS and PFC settings of every single CNA
- DCBX will also play a future role in CEE link negotiation between switches, but let's not get ahead of ourselves



# DCBX Overview

- DCBX allows two sides to exchange PFC and ETS attributes
- It also allows exchange of application capabilities like FCoE
- It uses LLDP as transport to exchange the above capabilities
- DCBX only works for point-to-point interfaces
- Maintains a state machine for every feature PFC, ETS, etc.



# Data Center Bridging eXchange (DCBX)

- Why DCBX?
  - Defines the limits of the CEE-capable cloud
  - Detects misconfiguration between peers
  - Can be used to configure a peer
- DCBX enhances LLDP with additional TLVs
  - PFC TLV
  - ETS TLV
  - FCoE Application TLV



# DCB & FCoE: Server I/O Consolidation



# Benefits of FCoE Server I/O Consolidation

## Lower CapEx and OpEx



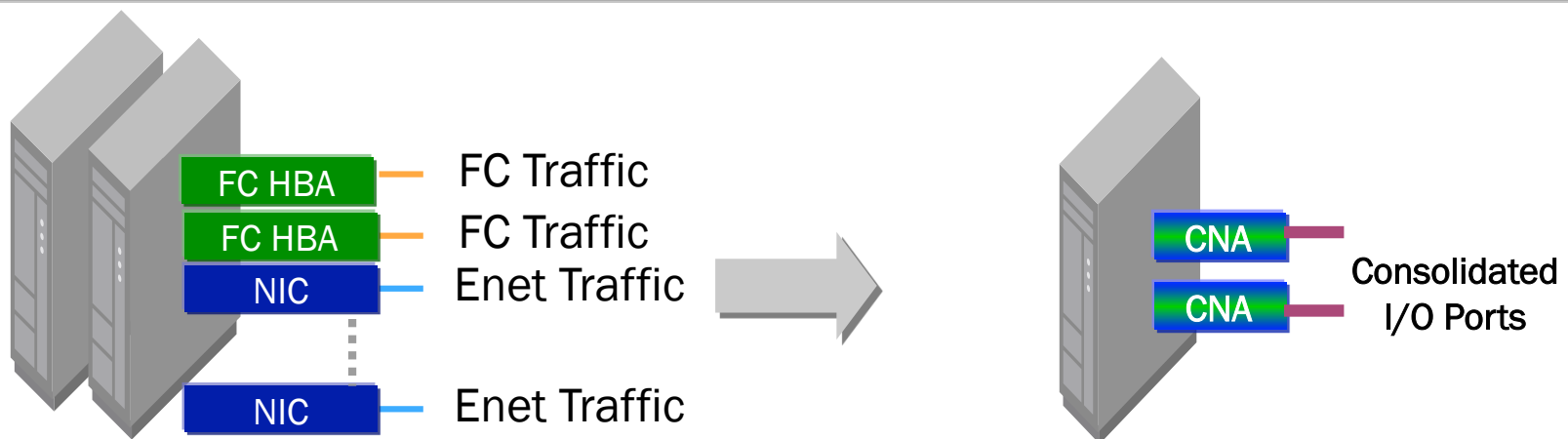
Consolidate and simplify server connectivity to LAN and SAN



Reduce server adapters, cabling, switch ports, and power usage



Lower CapEx and OpEx in Enterprise Data Centers

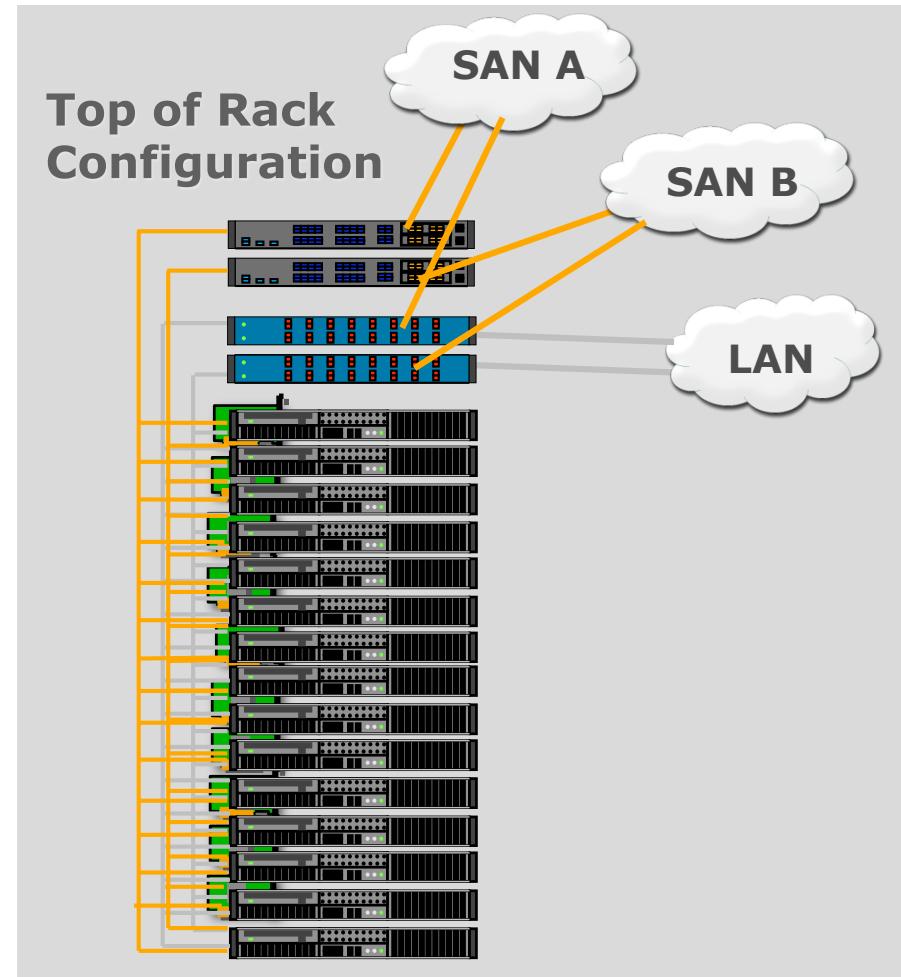




# FCoE & CEE Benefits

## Simpler Configurations & Lower Costs

- Reduce number of server adapters, ports, cables and switch ports
- Reduce power consumption
- Increase speed and utilization of links
- Simplify configurations and diagnostics



# Target Markets & Environments

## Enterprise & New Server Deployments

- Enterprises looking to implement FCoE data centers
  - Medium size to large global 1000 enterprises
  - FCoE integrates seamlessly into FC installations
- Enterprises with new data centers
  - Server I/O consolidation reduces CapEx for data centers
  - CEE/FCoE is a good fit for virtualization
  - Lossless 10 GbE or FCoE deployment
- New server installs
  - New windows and Linux environments
  - Applications running on VMs
  - Tier 3 and some tier 2 applications



# Brocade Solution & Products



# Brocade's Unified FCoE/CEE Solution

## Industry Leading Server I/O Consolidation Solution

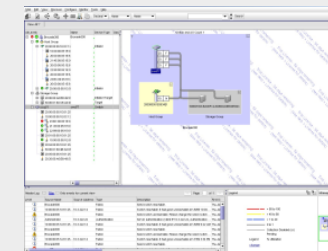
- Brocade CNAs
  - Single or dual 10 Gbps DCB/FCoE ports
  - Leading performance, broad OS support, low power draw
- Brocade 8000 Switch (ToR)
  - 1U; 24 10 Gbps CEE ports and 8 8 Gbps FC ports
  - Lossless Ethernet or FCoE deployment
- Brocade DCX<sup>®</sup> Backbones (EoR, MoR)
  - Brocade FCOE10-24 blades for server connection
  - Brocade 8 Gbps FC blades for SAN connection
- Brocade Data Center Fabric Manager (DCFM<sup>™</sup>)
  - Unified management of Brocade 8000, Brocade DCX Backbones, Brocade CNAs, and Brocade SANs

Brocade 1010/1020  
CNAs



Brocade 8000 Switch

Brocade FCOE10-24 Blade  
for DCX Backbones



Brocade  
DCFM

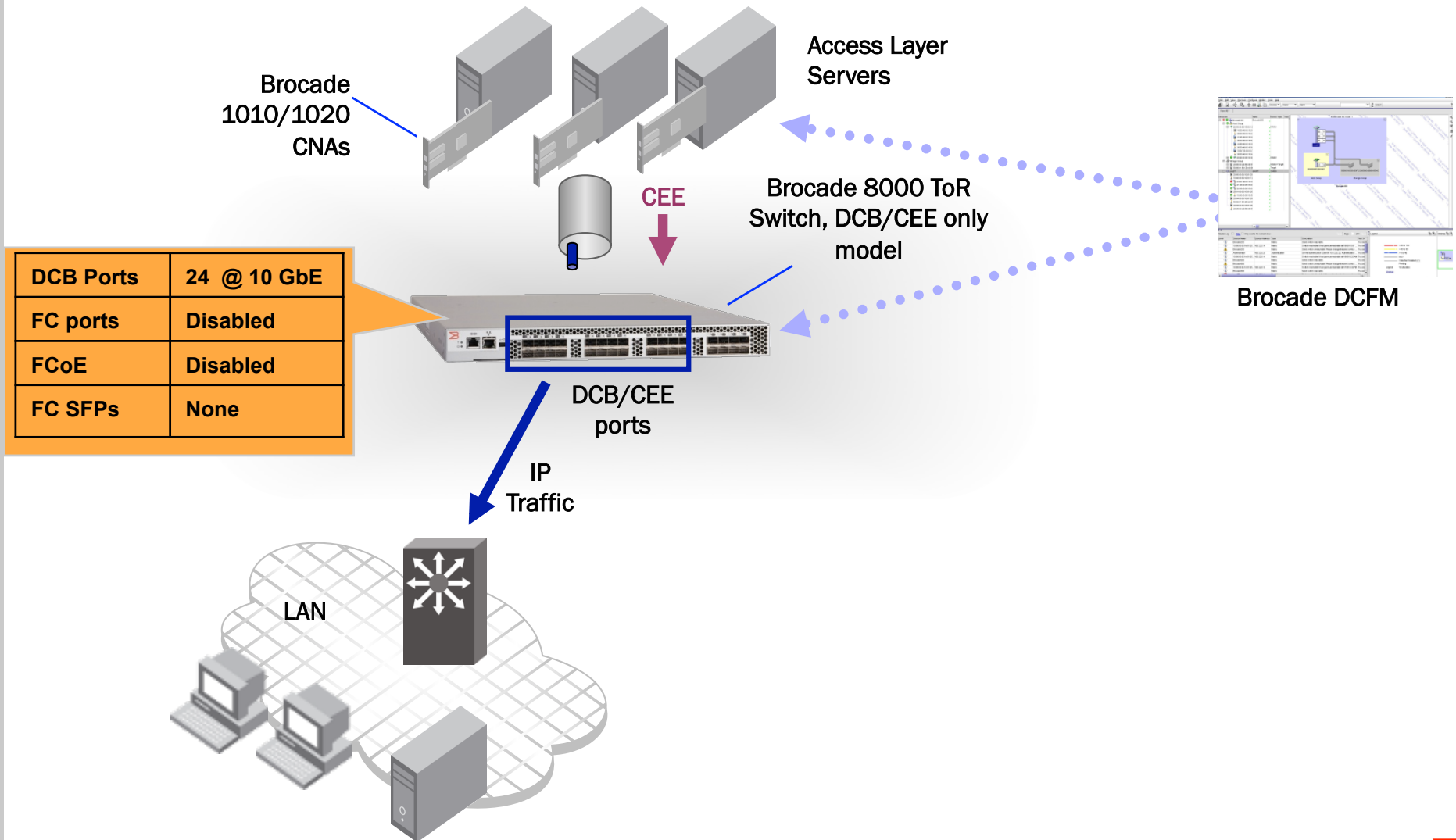
# Deployment Models





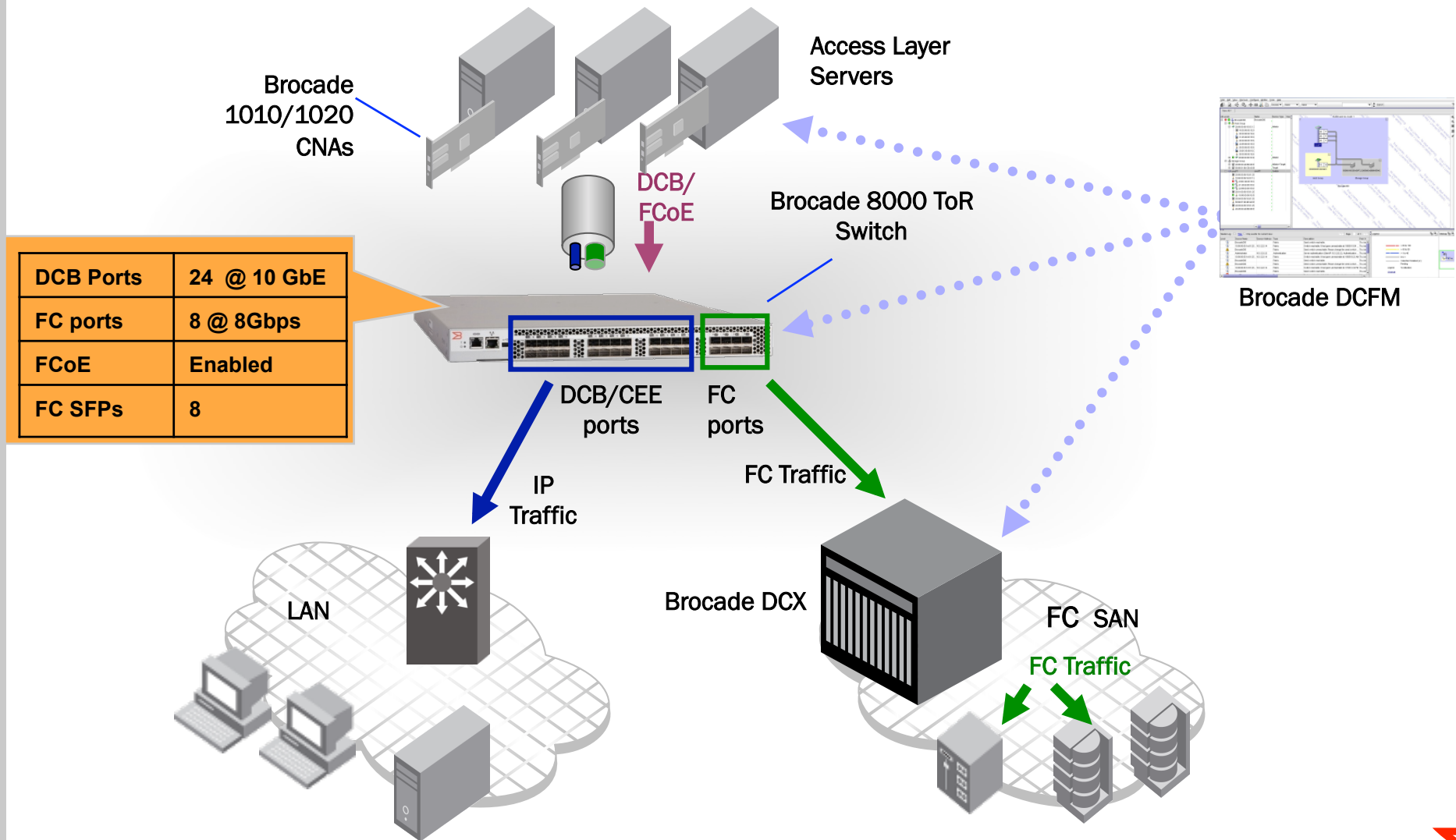
# Brocade TOR Solution

## Brocade 8000 Switch with DCB/CEE Only



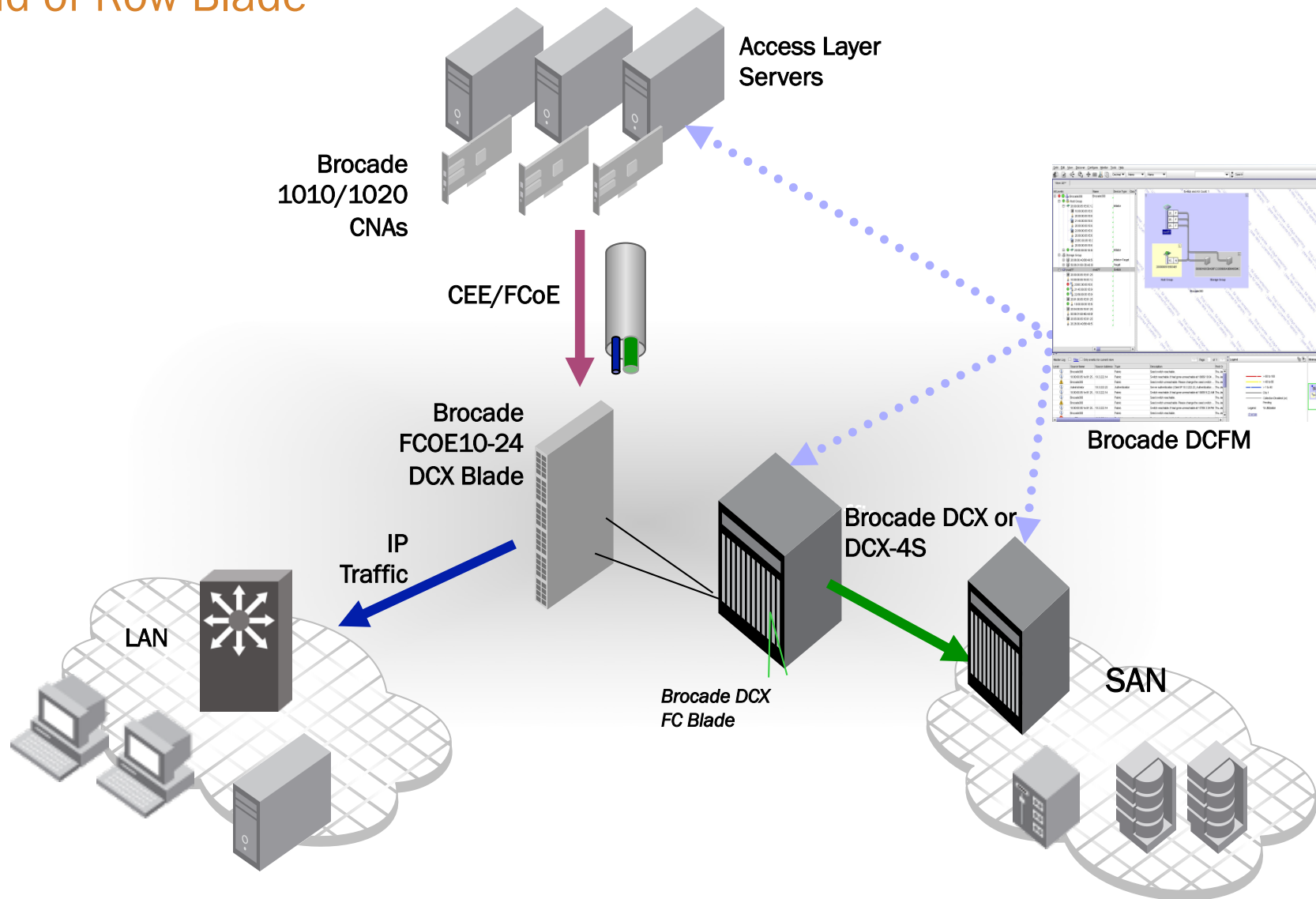
# Brocade TOR Converged Deployment

## Brocade 8000 with CEE & FCoE Enabled



# Brocade FCOE10-24 Deployment

## End of Row Blade

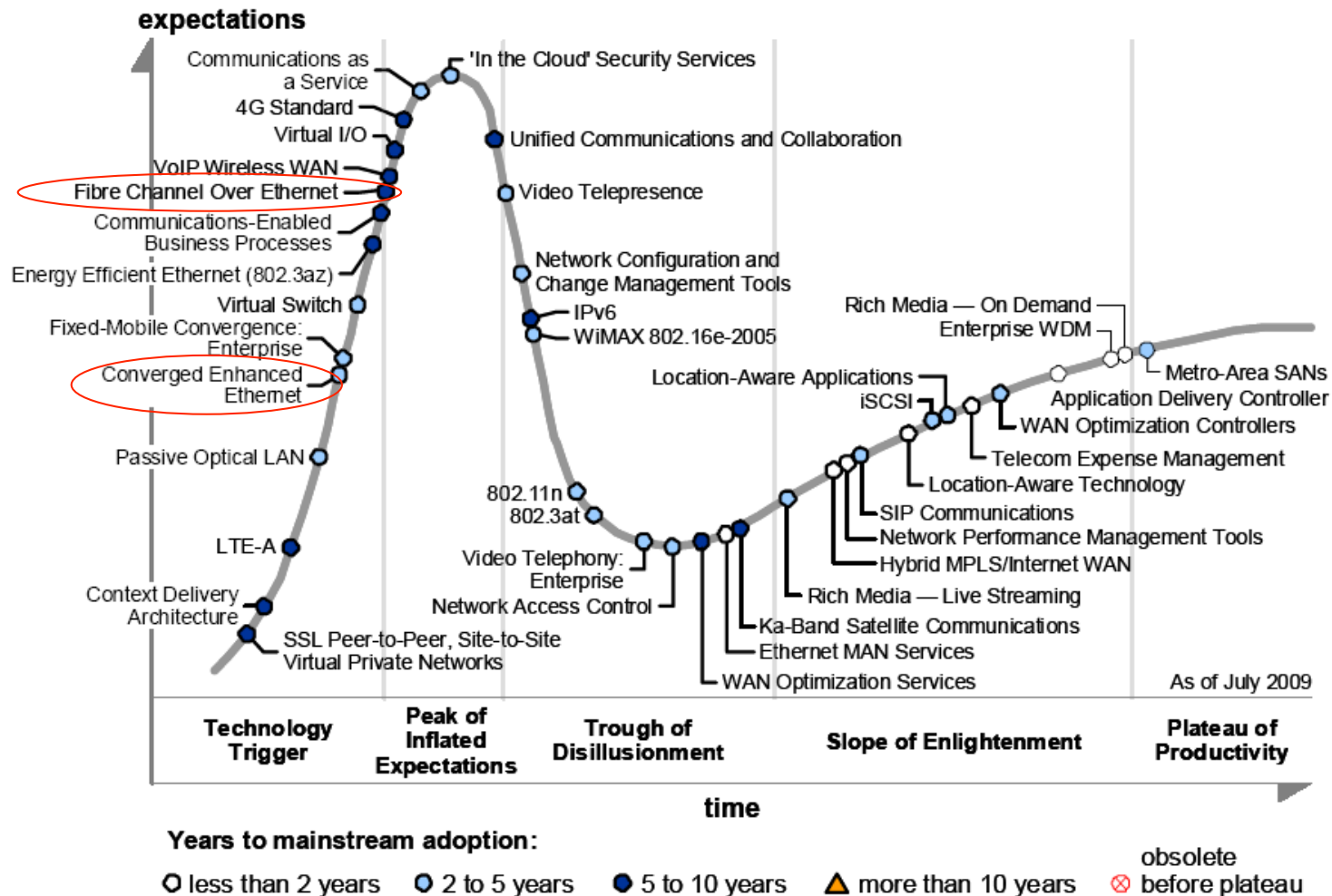




# Adoption Timelines



# Networking Technology Hype Cycle Curve - 2009



# TRILL



# TRILL

## *Transparent Interconnection of Lots of Links*

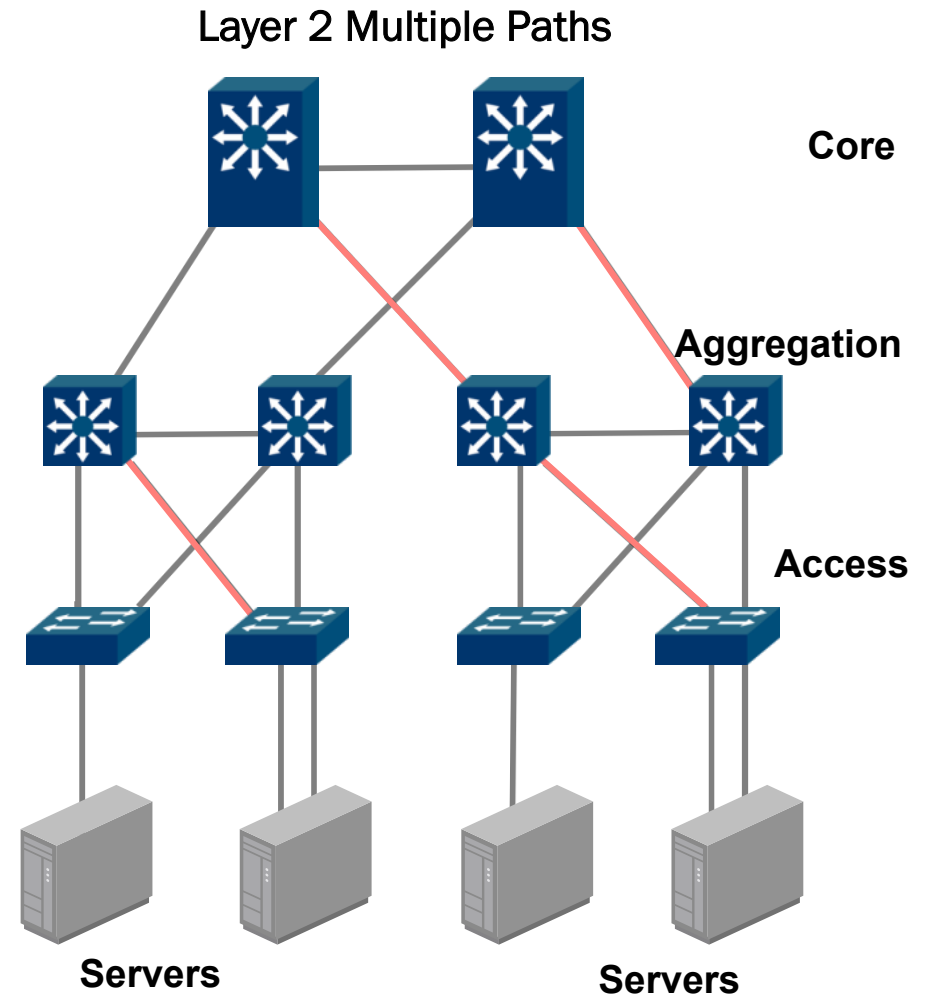
- TRILL is a new proposed data center Link Layer (L2) technology being developed by an Internet Engineering Task Force (IETF) workgroup
- *“The TRILL WG will design a solution for shortest-path frame routing in multi-hop IEEE 802.1-compliant Ethernet networks with arbitrary topologies, using an existing link-state routing protocol technology.”* - source IETF
- Development timelines
  - July 2010: Recharter or shut down the workgroup



# Data Center Networks

*Built with good intentions*

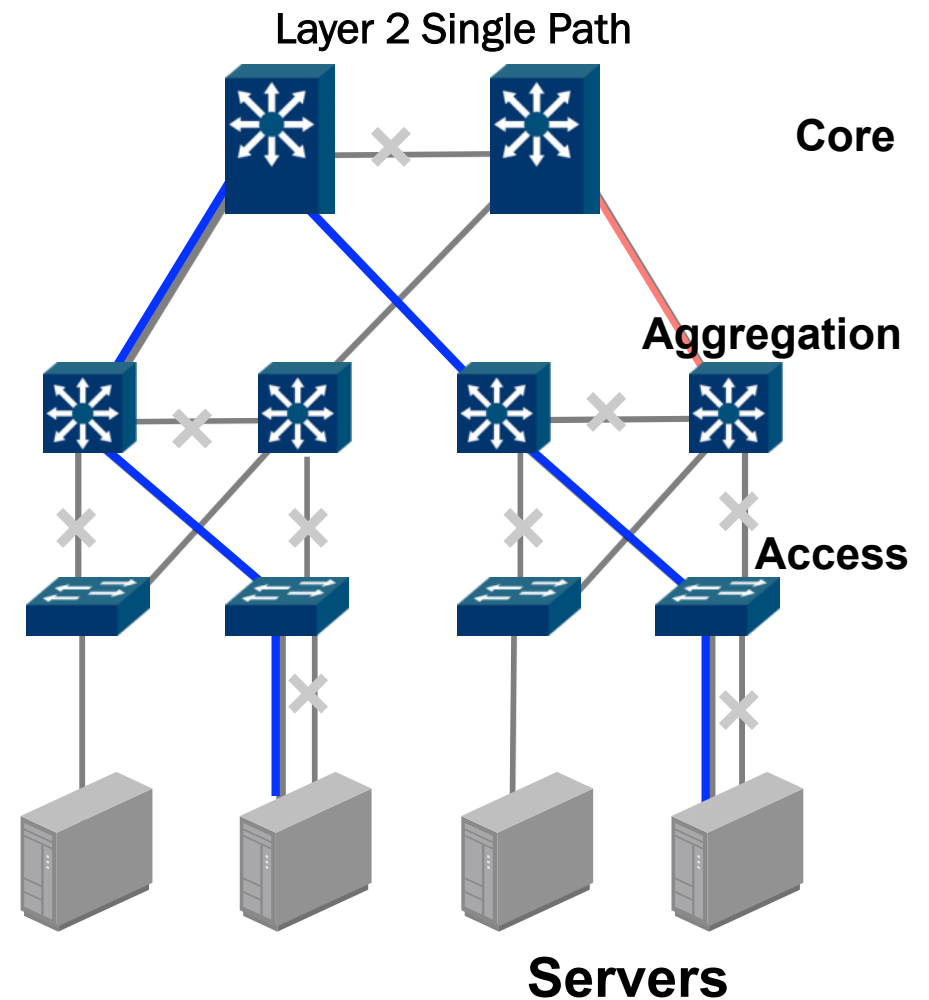
- Designed to meet business needs in an efficient and dynamic fashion while delivering performance, scalability, availability, and lower cost
- Data centers use richly connected topologies, such as fat trees
- Contain many Equal-Cost Multi-Paths (ECMP) between any given endpoints



# The STP Effect

## *Data center with Spanning Tree Protocol (STP)*

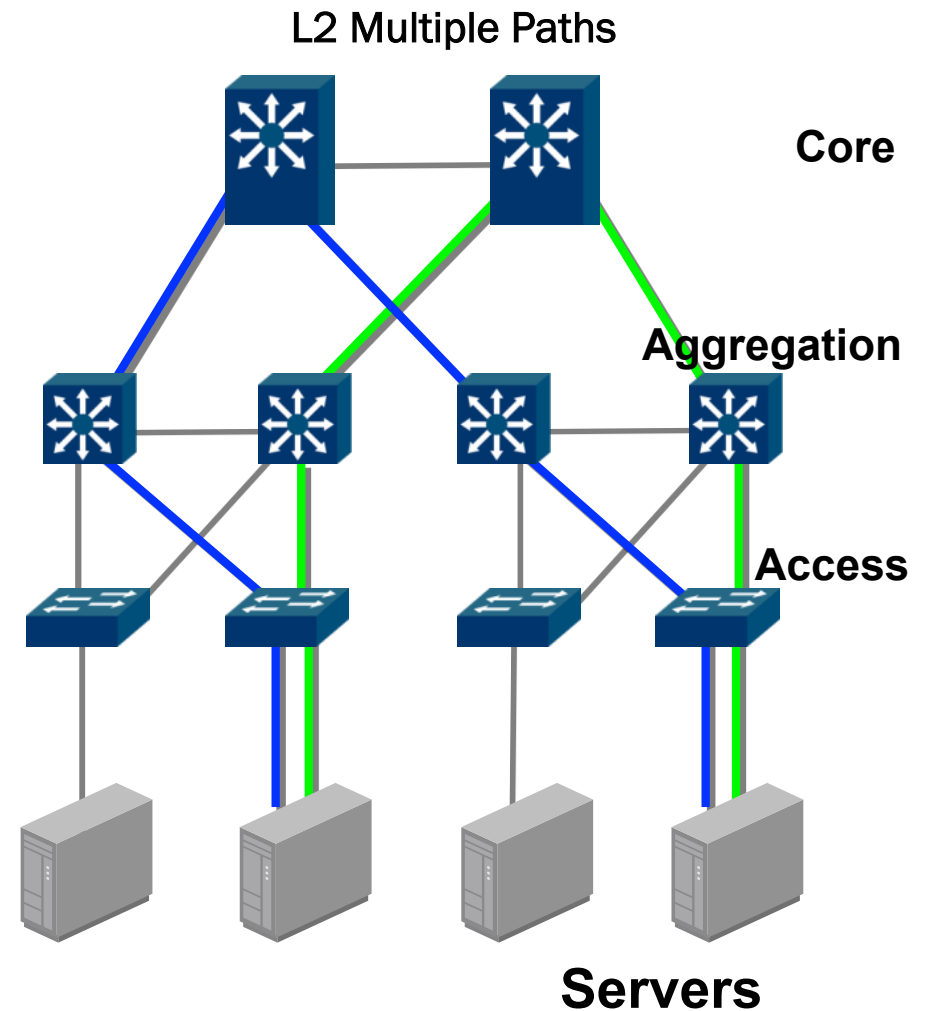
- STP is an Ethernet protocol for establishing and maintaining a single spanning tree among all the bridges on a local Ethernet segment
- In all types of spanning tree protocols (STP, MSTP, RSTP) a single tree/path is established for a given VLAN
- All alternate paths are blocked
- Inefficient use of available links reduces aggregate bandwidth
- Reacts to small topology changes with large impact on the reconfiguration of links (overreacts)
- MSTP helps, but requires additional configurations



# The TRILL Solution

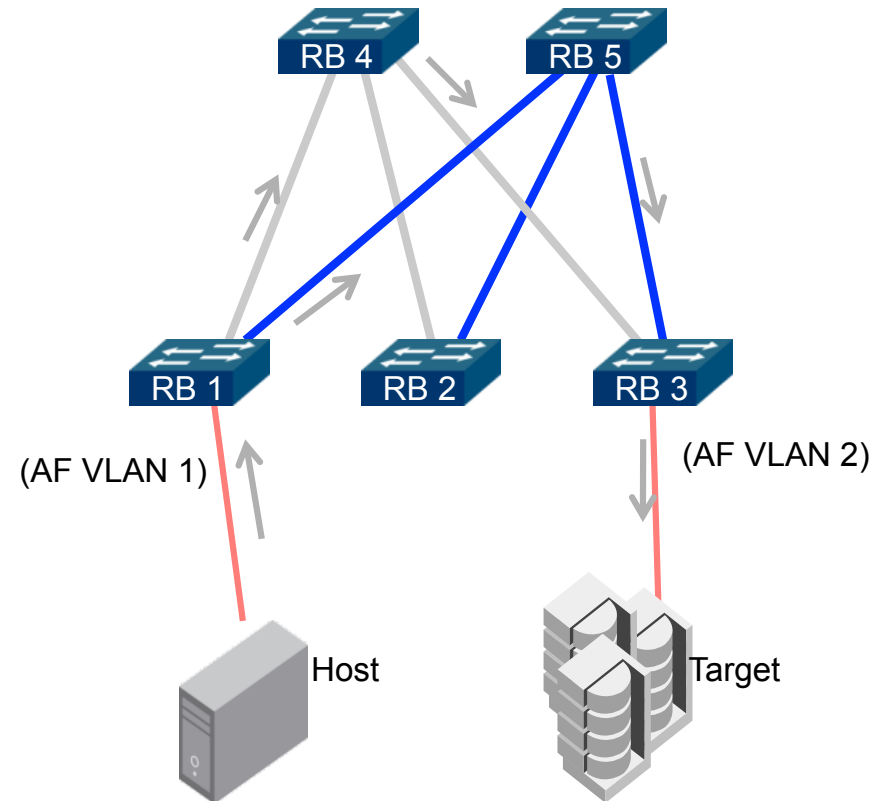
## *Layer 2 multipathing and multi hops*

- Provide Layer 2 multiple paths via load splitting among several paths
- Reclaims network bandwidth and improves utilization by spreading traffic more evenly
- Faster response to failures
- Improves efficiency by establishing shortest path through Layer 2 networks
- TRILL is backward-compatible and connects into existing infrastructures
- Delivers multiple hops by utilizing the concept of Routing Bridge (RBridge) running the IS-IS routing protocol



# TRILL Multi-Hop Concept

- Routing Bridge (RBridge): A device implementing the TRILL protocol and capable of learning VLAN configuration
- RBridges utilize multiple paths for higher throughput
- Use Layer 2 Intermediate System to Intermediate System (IS-IS) protocol
- Each VLAN has one RBridge called Appointed Forwarder (AF) that encapsulate/decapsulate frames from/to that VLAN
- Frames sent between RBridges are encapsulated, where link and TRILL headers are added





# TRILL

## *Protocol highlights*

---

Feature	Benefit
Load splitting among multiple paths	L2 multipathing: improved bandwidth and fault tolerance
Support multiple points of attachment	Redundancy
Routing loop mitigation, possibly through TTL field	Eliminates infinite loops and established hop count
Maintain compatibility with existing STP	Works with existing configurations
Minimum or no configuration requires	Faster reconfigurations
Use IS-IS to carry routing information and build shortest path tree for each VLAN	Broadcast MAC address info of devices connected to the VLAN. Established protocol

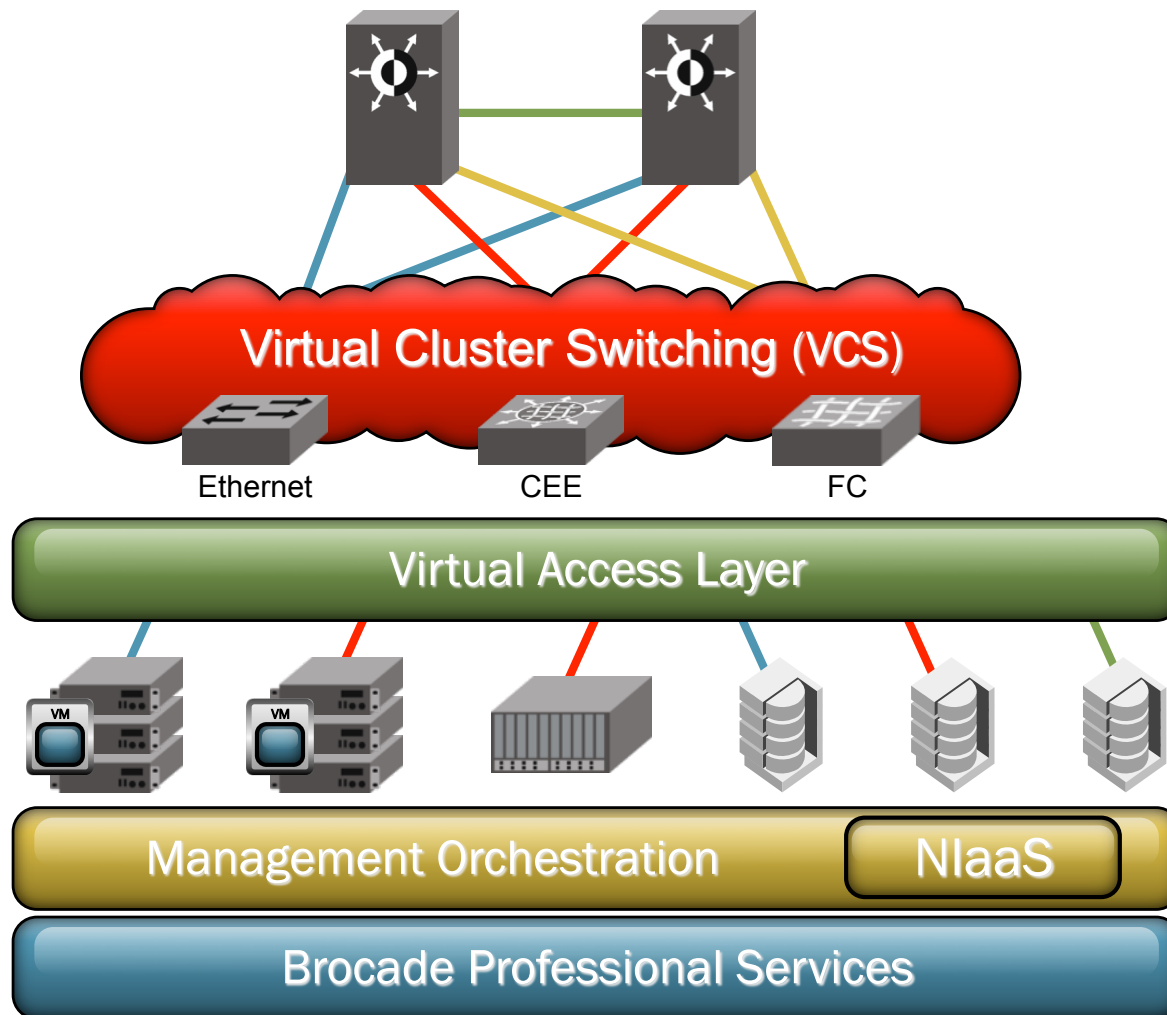
---

# The Brocade Solution

## *TRILL + Brocade enhancements*

- TRILL-ready hardware
  - Layer 2 multipathing capabilities and multiple-hop support
- Brocade Converged OS storage services on each hop
  - Built on storage-aware ASICs and protocol stack
  - FSPF running over CEE: multipathing using Fibre Channel standard routing protocol on CEE
  - Link protocol-agnostic fabric services: zoning and name server
- End-to-end troubleshooting and manageability
  - Across link layer protocols: CEE/FCoE and Fibre Channel
  - Across routing protocols: TRILL and FSPF
  - DCFM: Single management tool for all Brocade IP and SAN fabrics

# Brocade Architectural Vision



Collapse/simplify layered architecture

Map to application needs

Network innovation

Revolutionize Layer 2 connectivity (VCS)

Optimize VM performance and mobility (VAL)

Single pane of glass (NlaaS)



Thank You

