



Education

# Data Center Evolution and Network Convergence

Joseph L White, Juniper Networks

Author: Joseph L White, Juniper Networks

- ◆ The material contained in this tutorial is copyrighted by the SNIA unless otherwise noted.
- ◆ Member companies and individual members may use this material in presentations and literature under the following conditions:
  - ◆ Any slide or slides used must be reproduced in their entirety without modification
  - ◆ The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- ◆ This presentation is a project of the SNIA Education Committee.
- ◆ Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- ◆ The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

**NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.**

## ➤ Data Center Evolution and Network Convergence

- ◆ FC, FCoE, NAS, iSCSI, DCB, traditional LAN, internet/WAN, HPC, clusters, clouds, server virtualization, storage virtualization, network virtualization, and more are all colliding in your data center. Redundancy, resiliency, security, I/O consolidation, network convergence, dynamic application distribution, and thin provisioning with high levels of service are desired at all layers and all data center sizes across a broad spectrum of use cases. You worry about operational separation, buying decisions, investment protection, cost and energy savings, and evolving standards while maintaining very high levels of service and security. Is the technology evolving to a dream come true or a nightmare? If that doesn't keep you up at night nothing will.
- ◆ This tutorial will untangle, define, and illustrate the main ideas and concepts behind Data Center Evolution and Network Convergence to give context and a solid foundation for discussions with your vendors as well as for your further reading and investigation. The point of view taken for this presentation is that of the network and transport characteristics in the face of the changes taking place.

## ➤ History and Trends

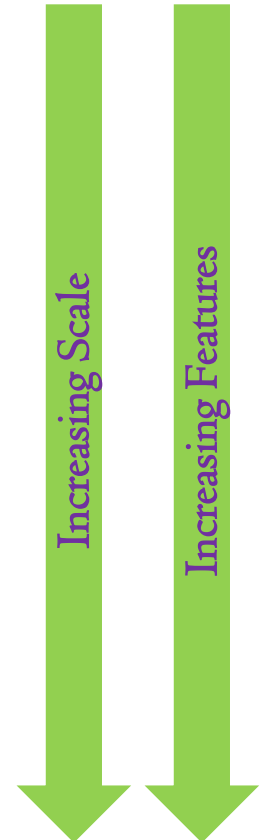
- ◆ Scaling
- ◆ Virtualization
- ◆ Data Center Consolidation

## ➤ Convergence

- ◆ I/O Consolidation
- ◆ Tier Collapse
- ◆ SAN-LAN Convergence
- ◆ Inhibitors and Accelerators

# An Abbreviated History

- ◆ **First Model**
  - ◆ isolated systems
  - ◆ direct attach storage
- ◆ **Second Model**
  - ◆ limited networking
  - ◆ direct attach storage
- ◆ **Third Model**
  - ◆ Networking Explosion
  - ◆ direct attach storage + local cluster server-server storage
- ◆ **Third and a half Model**
  - ◆ Flexible storage via NAS and network file systems
- ◆ **Fourth Model**
  - ◆ SANs for block storage attach plus fully entrenched NAS
- ◆ **Fifth Model ('current one')**
  - ◆ Server Virtualization drives first hop I/O consolidation, increased SAN attach
  - ◆ "Cloud" (pooled resources of all kinds with uniform distributed access)
- ◆ **Evolving Model**
  - ◆ Network Convergence
    - › Protocols for SAN and LAN on same infrastructure
    - › Network Scaling via virtualization and simplification (tier collapsing, distributed control planes)



## Consolidation

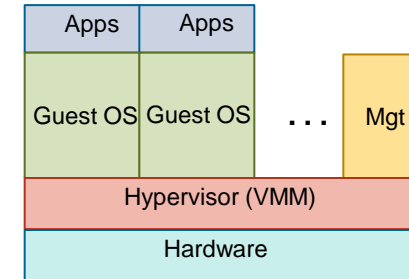
- § Mega DCs; 400K sq ft
- § 4K racks, 200K servers



DC Scale

## Server Trends

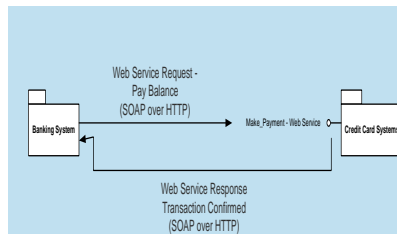
- § Multi-core (8->16 >32,...128,...)
- § Virtualization and VMs



Want Low Oversubscription

## Application Trends

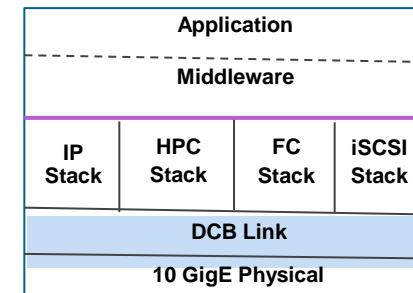
- § SOA, Web 2.0
- § MapReduce, Hadoop, Grids



Increased East-West traffic

## Interconnect Trends

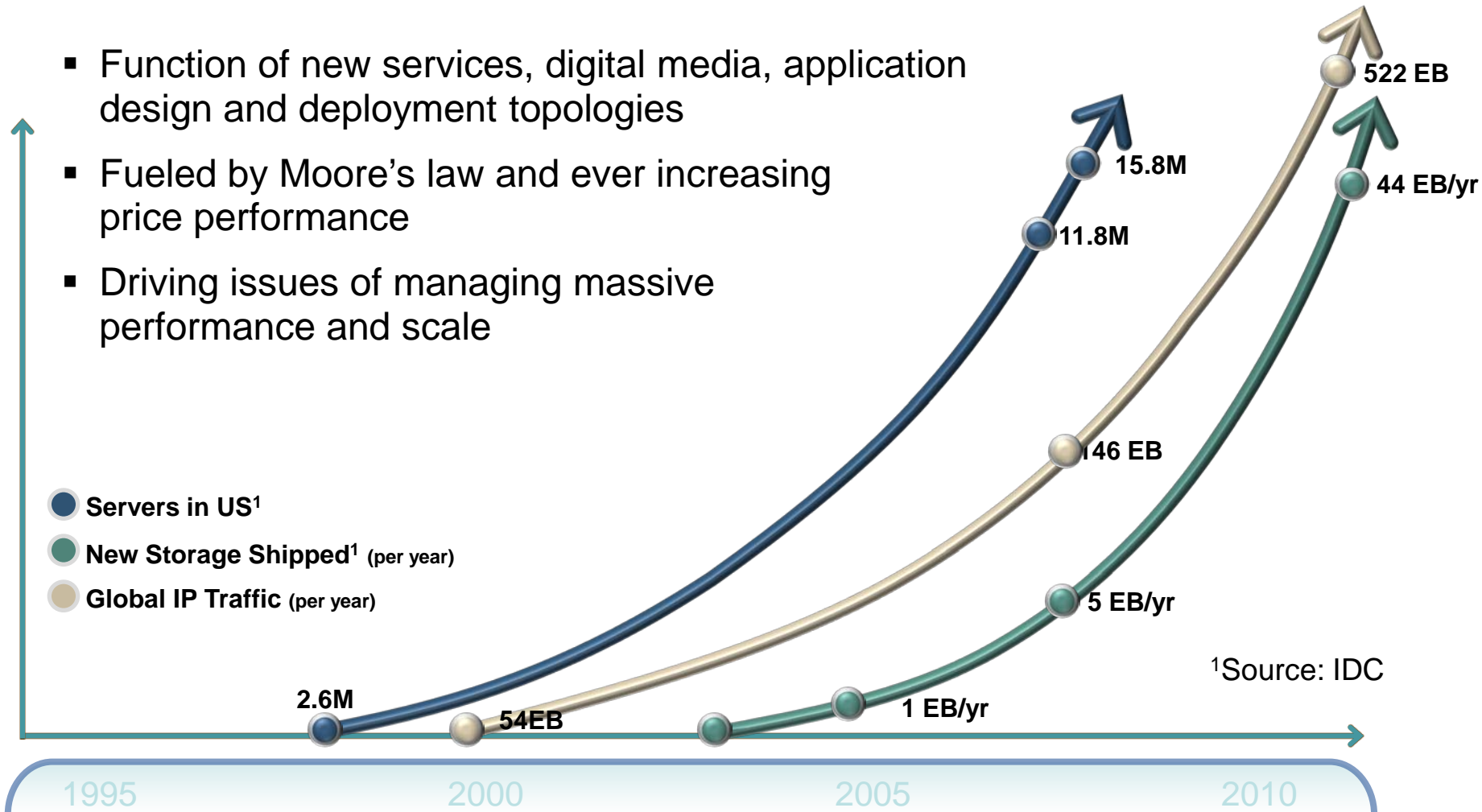
- § Convergence to 10 GE
- § Enhancements to Ethernet



Large speed increases 10/40/100 GE

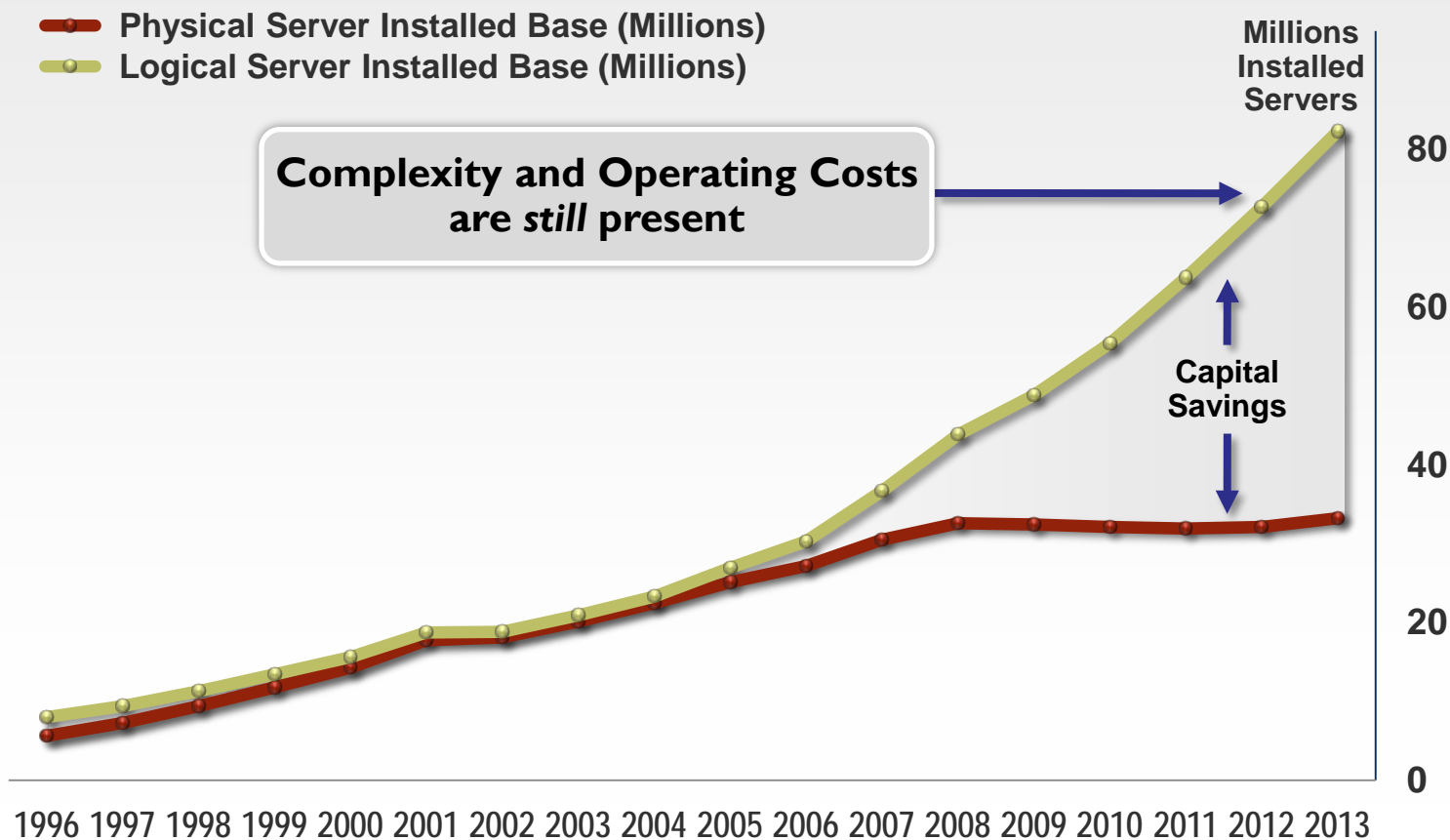
# Explosion in infrastructure

- Function of new services, digital media, application design and deployment topologies
- Fueled by Moore's law and ever increasing price performance
- Driving issues of managing massive performance and scale



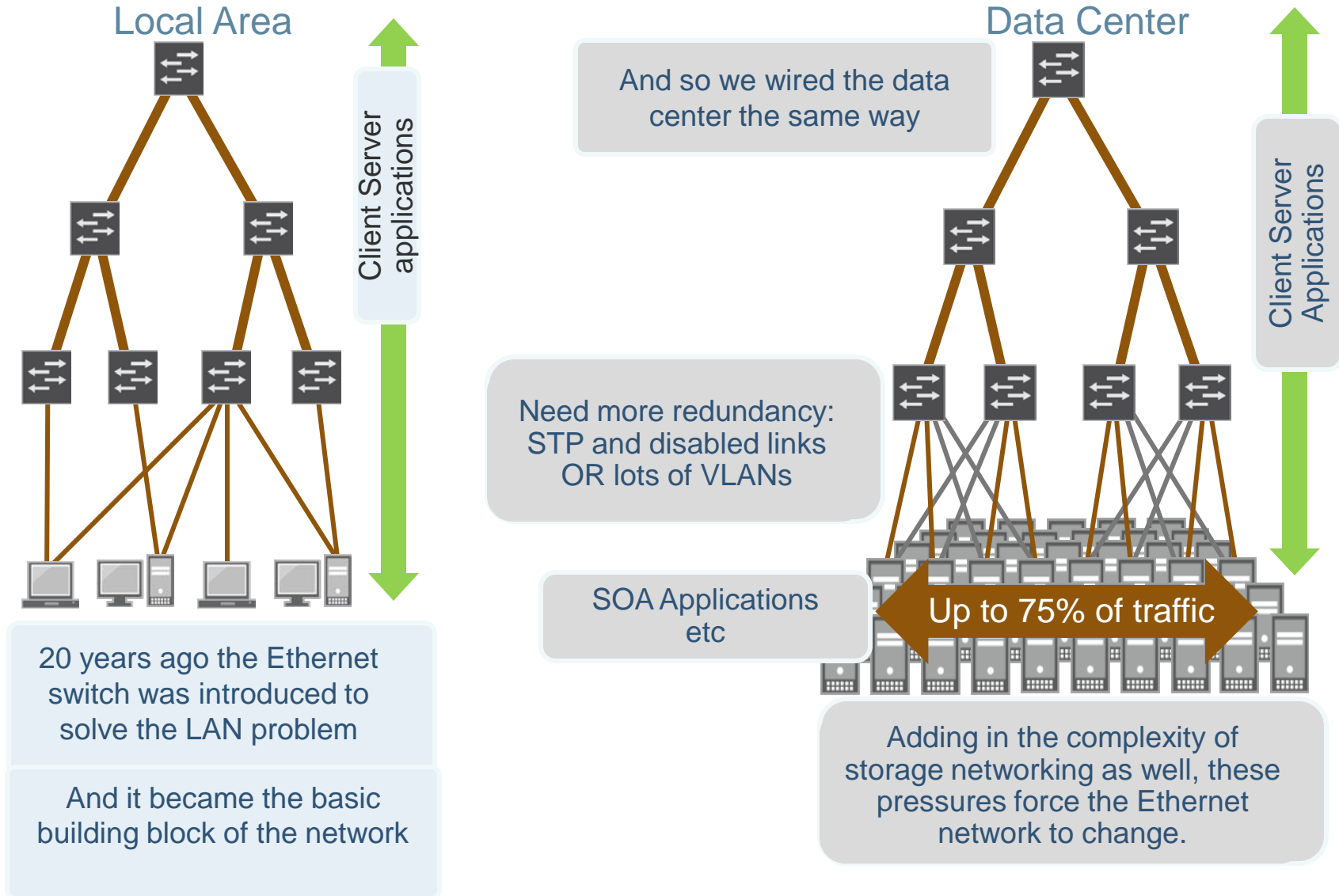
Remember: Complexity increases exponentially with scale

# Virtual to Physical Server Trend



Source: IDC

# Evolution of the Data Center LAN



## ➤ Aggregate up and Virtualize down

- ◆ many examples such as storage arrays, servers, ...
- ◆ avoid Accidental partitioning
- ◆ embrace Deliberate partitioning

## ➤ Aggregation

- ◆ Physical and Software
- ◆ Bring together and pool capacity with flexible connectivity

## ➤ Virtualization

- ◆ logical partitions of the aggregated systems to match actual need
- ◆ flexibility → fungible resources everywhere
- ◆ Utility Infrastructure with just in time & thin provisioning

**THIS IS HAPPENING TO NETWORKS AS WELL**

**... because Data Centers are always in flux**

**Application life cycle**

services introduced, updated, retired

**Load on servers and networks constantly changing**

can be unpredictable

**Resource management challenge**

- ◆ Minimize the need for excess capacity
  - > Reconfigure
  - > Reclaim/Reuse
- ◆ Adding resources is last resort

**Dynamic shared resource pools address these issues**

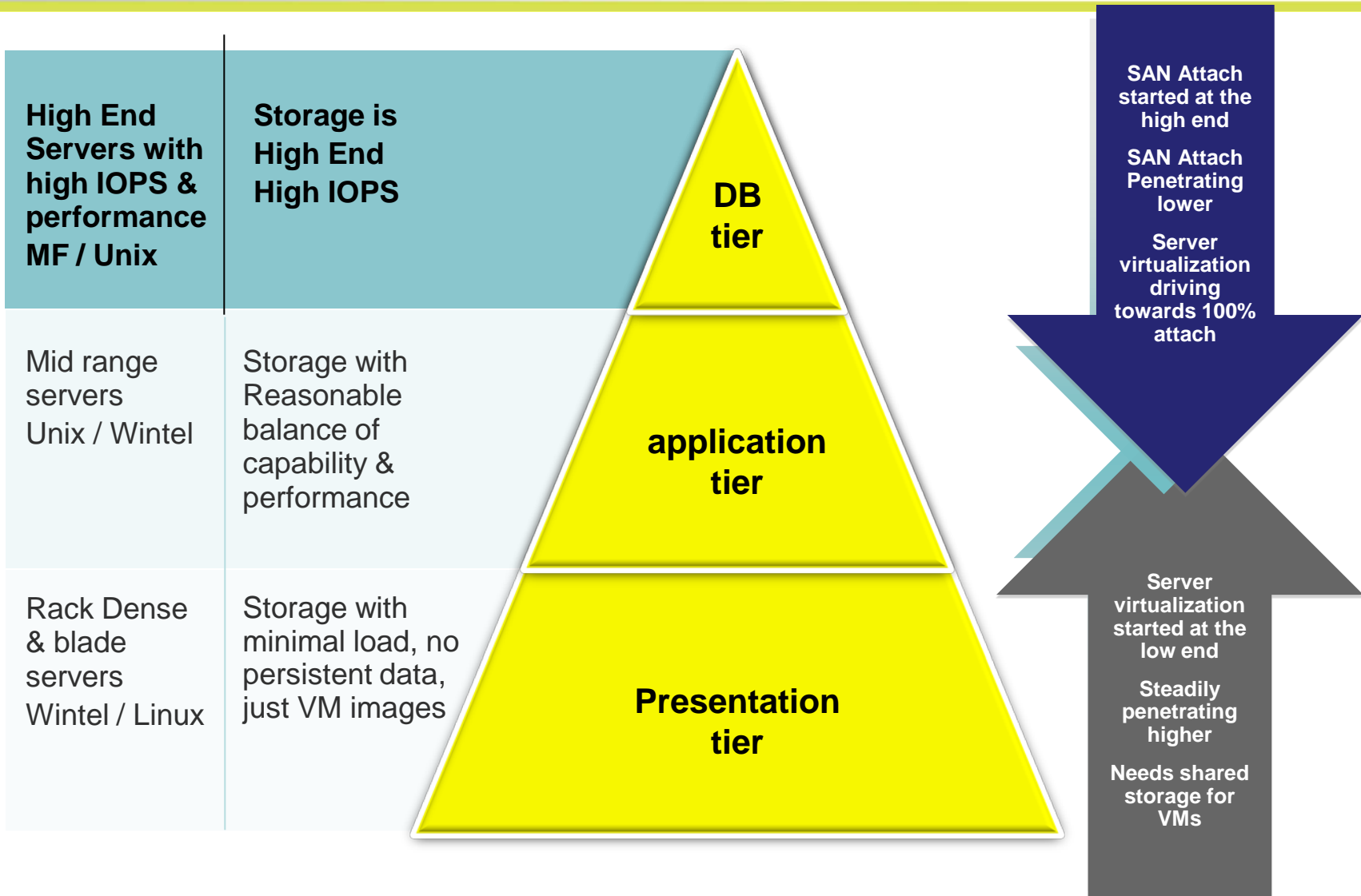
**Enabled by Virtualization + Full Connectivity Networks**

**Any servers potentially needs access to any storage**

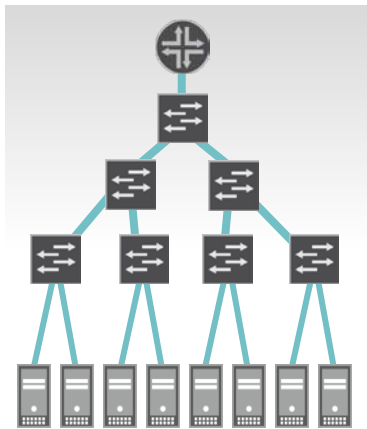
**Drives SAN attach from 20% to near 100%**

**If you don't converge you will end up connecting everything to everything anyway but across additional parallel networks.**

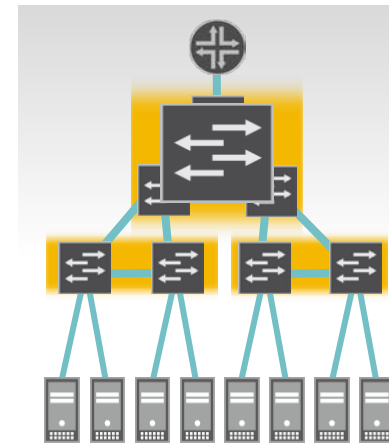
# Servers and Storage Attach



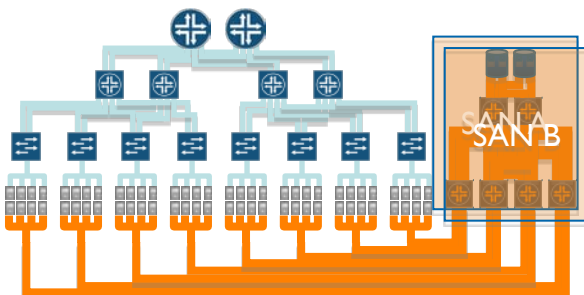
- Convergence occurs along 2 dimensions



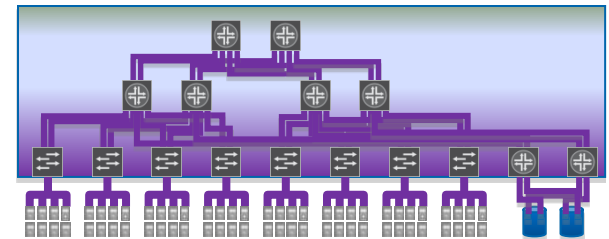
**Tier  
Collapse**



- And the Converged infrastructure will run on the collapsed infrastructure



**Infrastructure  
Convergence**



# Has Convergence already happened?

## ➤ For some aspects of Convergence: YES

- ◆ NAS Allows access to file based storage across the network
- ◆ iSCSI Allows access to block based storage across the network
- ◆ SANs have been bridged across metro and wide area networks for 10 years (FCIP, iFCP, & proprietary)
- ◆ FCoE provides an accepted protocol for FC across an Ethernet

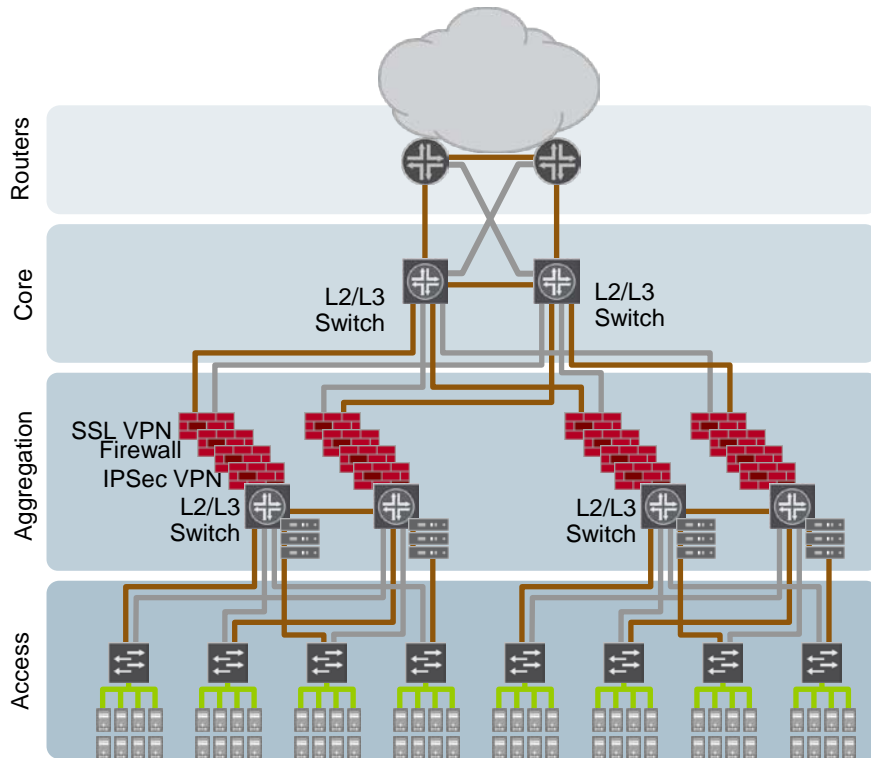
## ➤ These are good but ***not sufficient***

- ◆ Data center LANs have issues at scale
- ◆ WAN IP SAN connections do not solve the Local Data Center problem
- ◆ Operational Characteristics of FC based SANs desirable and entrenched for many applications

# What else does convergence need?

- ◆ **Transport Convergence**
  - › I/O Consolidation out of server
  - › Switch forwarding plane aggregation
    - Tier Collapsing
    - Ethernet Fabrics, not to be confused with FC Fabrics
- ◆ **Protocol Evolution**
  - › FCoE (FC-BB-5/FC-BB-6)
  - › iSCSI (especially over DCB networks)
  - › DCB for Ethernet is ‘complete’
- ◆ **Operational**
  - › Collections of physical switches acting as single logical switch
  - › Scaling of all management tools and infrastructure

***We will now explore these additional aspects of convergence***

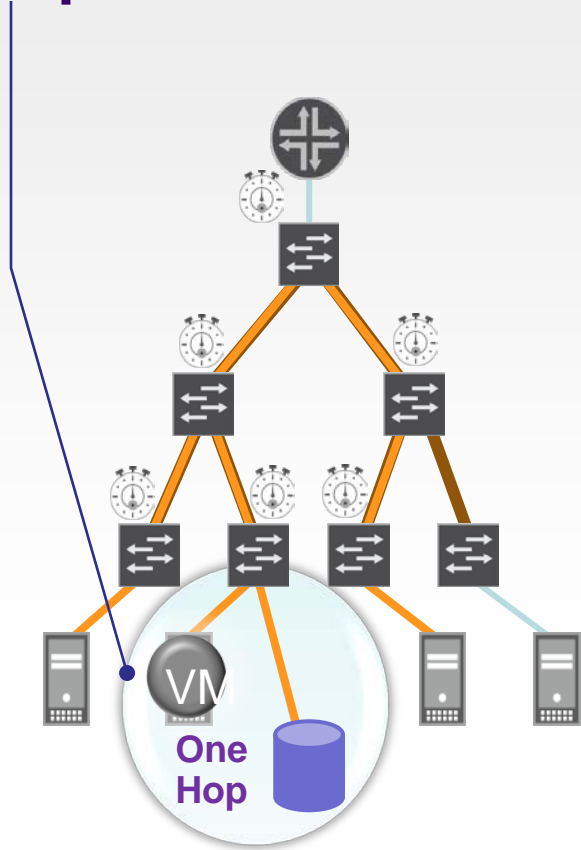


## built from Multi-Tier Trees

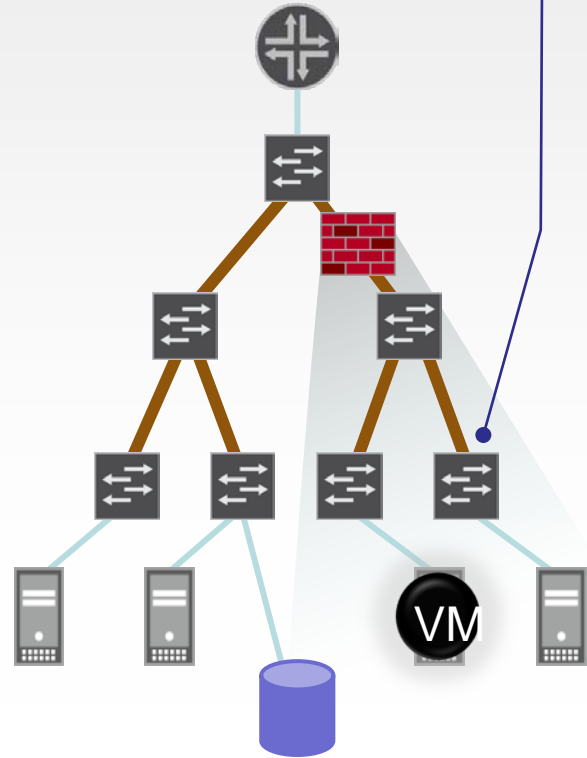
- Servers → Servers, NAS, Campus/MAN/WAN across switched network
- Multi-Tier
- 100's to many 1000's of ports
- multi-link redundancy
- 100s of meters max diameter
  - ◆ oversubscribed
  - ◆ East-West Latency can be a problem
- Ethernet carrying predominantly IP traffic
- Firewalls and security in aggregation layer
  - ◆ have to be distribute in the data path due to efficiency forced by oversubscription

# Why not Multi-Tier Trees?

## Bubbles of Optimal Performance



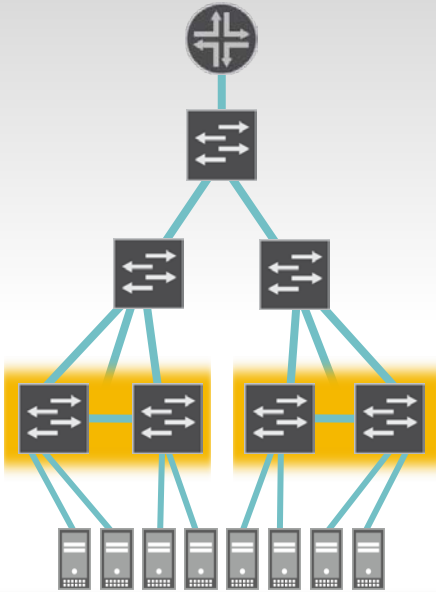
## Appliances and VLANs create Shadows of Accessibility



Location matters in a tree architecture

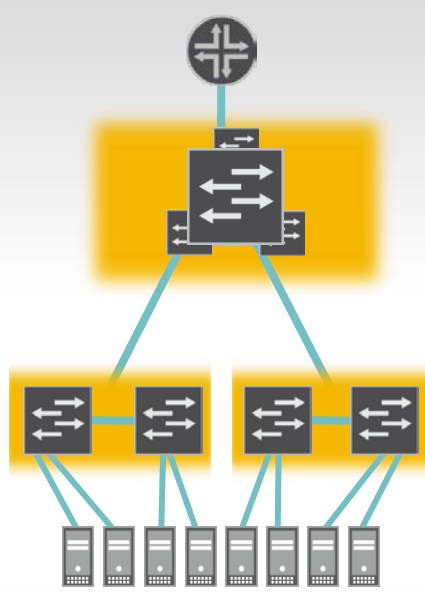
# Pruning the Trees

## Simplify the Data Center LAN by aggregation and consolidation



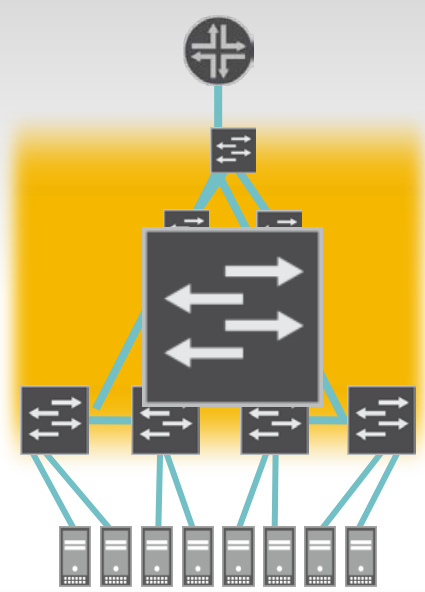
### Aggregate Switches

Multiple physical switches that operate as a single logical device for both management and traffic forwarding



### Collapse Tiers

Use an aggregated switch to do the work of multiple tiers of switches  
Allows aggregated access to services

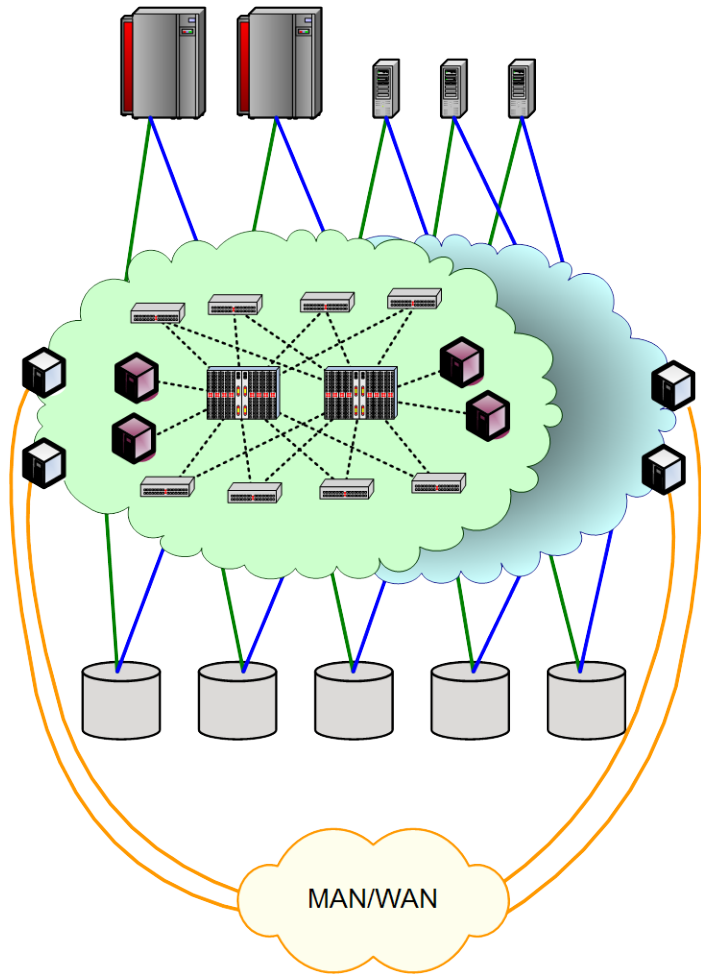


### Create A Fabric

Use both techniques at the same time to build a fabric based infrastructure

**This technology is also good for our SAN**

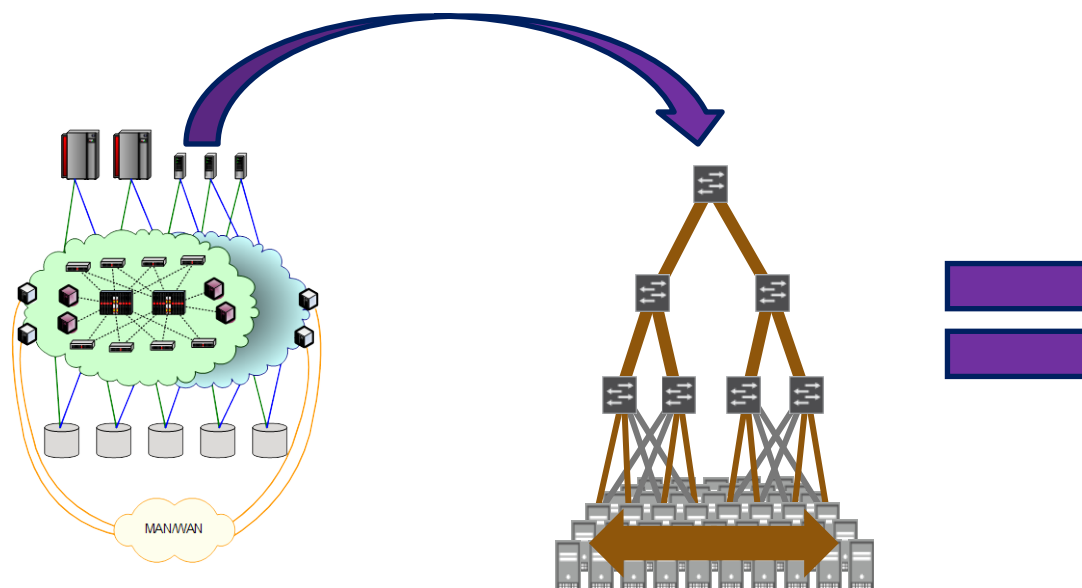
# Data Center FC SAN (today)



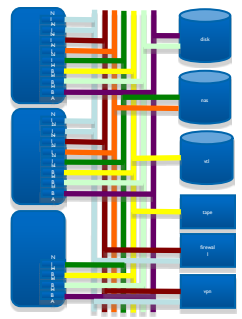
- Servers → Storage across switched network
- Core – Edge or Edge – Core – Edge
  - ◆ in effect one level of tier collapse is done
- 10's to 1000's of ports
- Full Dual Rail Redundancy
- 100s of meters max diameter
  - ◆ High Bandwidth, Low Latency
  - ◆ Lossless Links
- Fabric Services provide Discovery, Access Control, and Change Notification
- Gateways and specialized extension devices provide remote access for BC/DR
- Attached Appliances provide data services
  - ◆ Encryption, Block Virtualization

# Now Put the SAN on the LAN

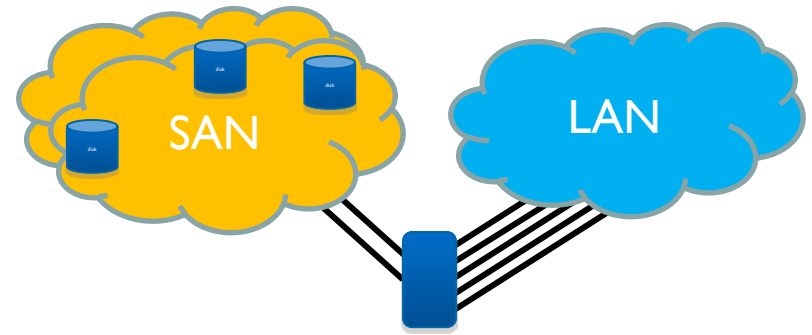
- What happens when you overlay the SAN picture ... onto the LAN picture



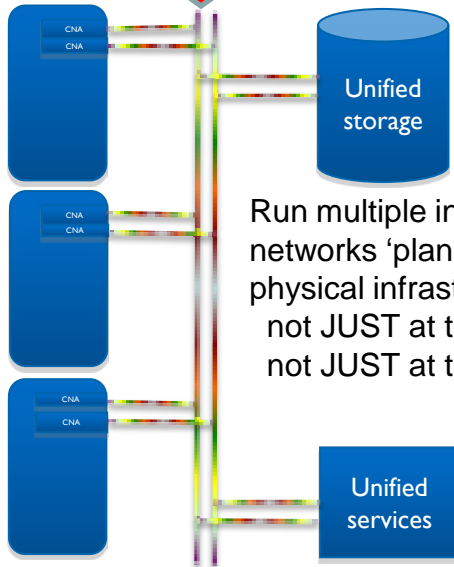
# At a high level you end up with...



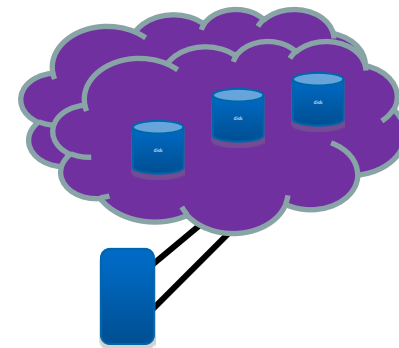
**I/O Consolidation**



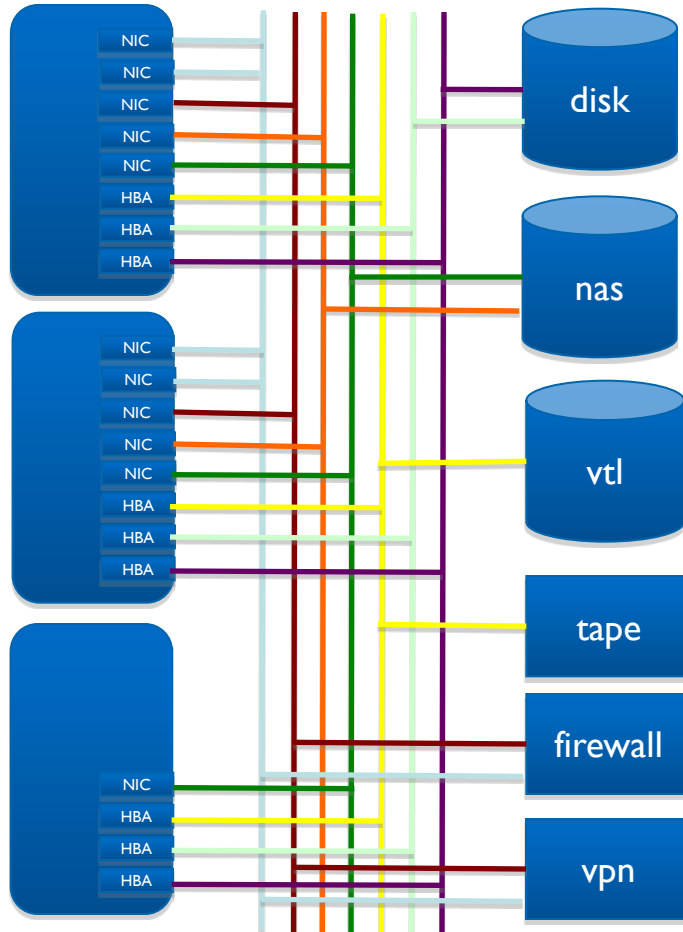
**Network Convergence**



Run multiple independent virtual networks 'planes' across a single physical infrastructure  
not JUST at the Edge  
not JUST at the SAN



# I/O Consolidation Server Point of View



## Multiple network connections

- ◆ Separate subnets
- ◆ Separate VLANs
- ◆ May be separate networks
- ◆ May be separate switches in the same network
- ◆ Each configured with different QoS settings

## Depends on the Server...

- ◆ file servers and/or NAS points to campus
- ◆ Presentation servers and intranet point to campus
- ◆ Application servers
- ◆ Database servers
- ◆ Backup Servers, etc etc

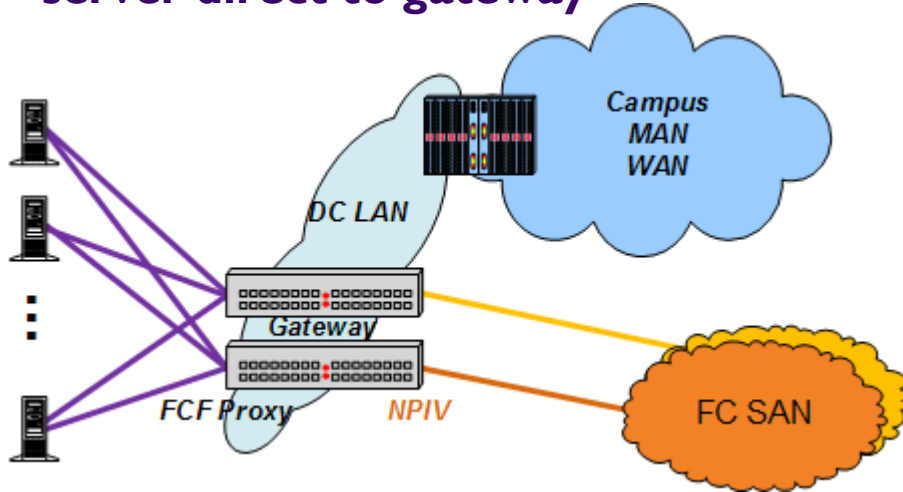
## For Example Something like this is possible:

- ◆ Campus Access (1+)
- ◆ LAN Backup NIC (1)
- ◆ Application Cluster NIC (1)
- ◆ Vmotion NIC (1)
- ◆ Presentation to application to database private NICs (1+)
- ◆ Storage Access Cards (2)

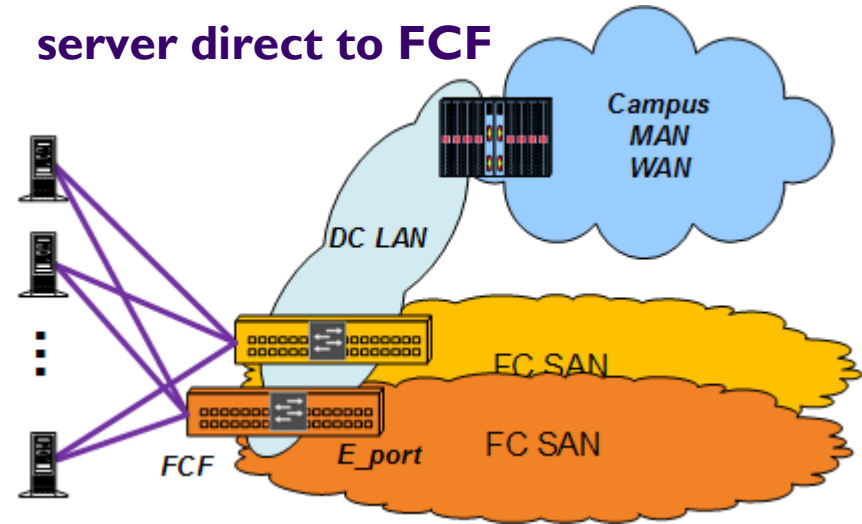
*I/O Consolidation reducing this to 2 NICs can save a bundle here*

# I/O Consolidation Deployments

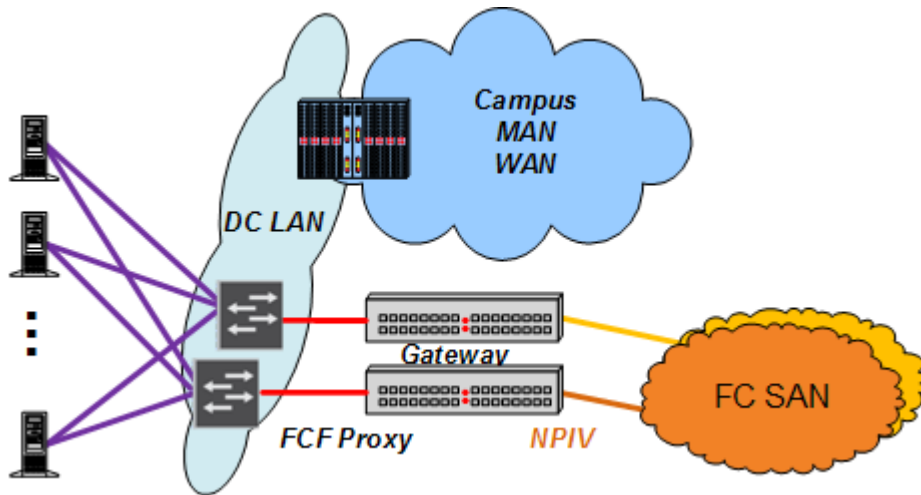
## server direct to gateway



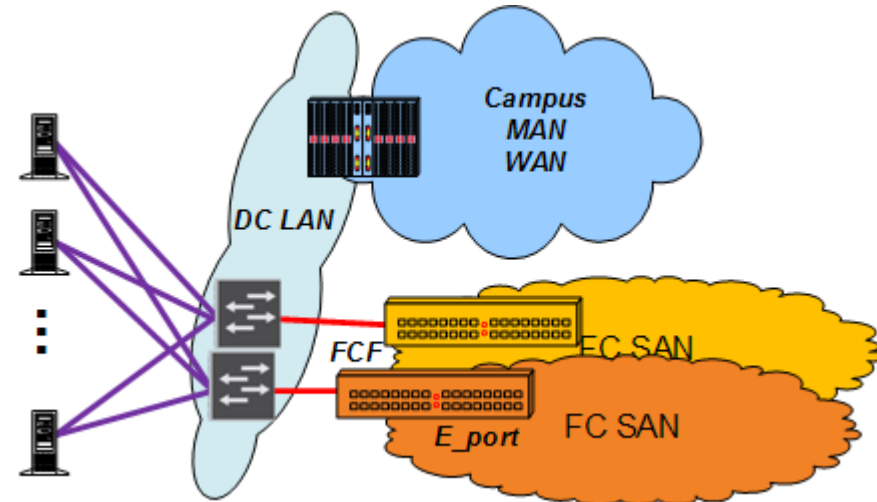
## server direct to FCF

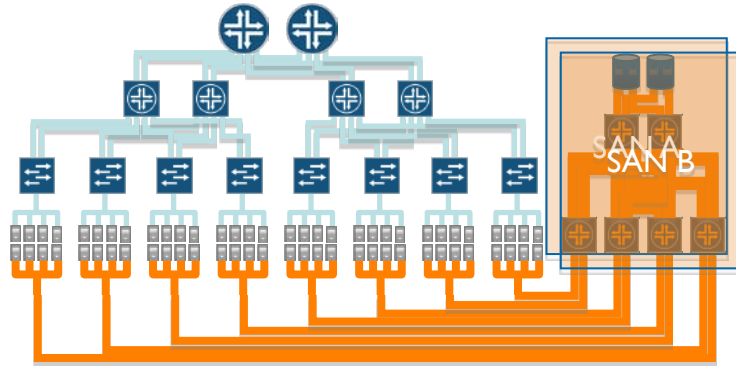


## server to L2 to gateway

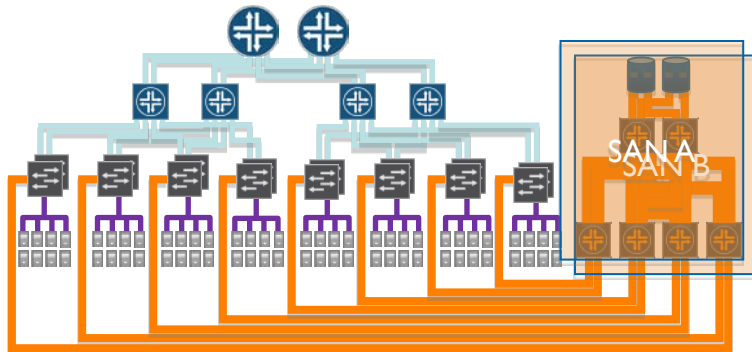


## server L2 to FCF

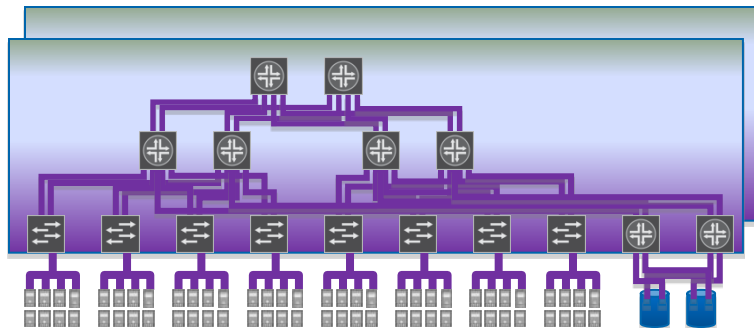




- ◆ Separate Networks
  - ◆ No overlap of operations & management
  - ◆ separate network characteristics
  - ◆ separate teams run them (typically)



- ◆ I/O Consolidation
  - ◆ overlap confined to server and 1<sup>st</sup> hop
    - ◆ team consensus needed
  - ◆ several detailed ways to accomplish
    - ◆ transit switch L2 to access SAN
    - ◆ gateway



- ◆ SAN-LAN Convergence
  - ◆ unified management and operational model
    - ◆ run by single team
    - ◆ common redundancy
    - ◆ shared bandwidth and connectivity

- results in logical overlays for forwarding on single, shared HW infrastructure
  
- **Benefits**
  - ◆ stocking of spare FRUs
  - ◆ combined operations
  - ◆ fewer stranded resources
  - ◆ better utilization
  - ◆ lower latency
  - ◆ better flexibility

# IF we use FCoE for Network Convergence

## ➤ Scaling

- ◆ FCF at the TOR hits Domain Scaling problems
- ◆ FC-BB-6 addressing this with FCF-FDF virtual domains

## ➤ Configuration of L2/L3 separation

- ◆ VLAN → Virtual Fabric mapping
- ◆ L2 network as access between server and FCF

## ➤ Full FC topology overlay onto data center network

- ◆ Multi-hop FCoE across multiple L2 networks and several FCFs
- ◆ Multiple VLAN/Virtual Fabric Configuration and Management

## ➤ Services

- ◆ services means discovery, access control, notifications, security
- ◆ iSNS exists for these but not commonly deployed

## ➤ Management

- ◆ Tools have limited support for iSCSI attach

## ➤ Performance

- ◆ to match local data center SAN requirements need DCB
- ◆ this is true even though TCP/IP is the transport

## ➤ End to End performance

- ◆ bandwidth & latency & frame processing rate
- ◆ congestion handling
- ◆ logical overlays onto physical infrastructure

## ➤ MAN and WAN extension for storage

- ◆ Under convergence the local network is Ethernet. The devices that make up the local Ethernet are not typically suited to directly support lossless distance extension.

# Inhibitors and Accelerators

***If convergence is so great why hasn't it already happened?***

***Actually it is in progress,***

***but there is a balance between various factors influencing the adoption rate***

## INHIBITORS

Existing  
Operational  
Models and  
Infrastructure

Cost of 10G/40G  
Infrastructure

Politics /  
Team Issues

Technology Maturity



## Data Center Network Convergence



Development of  
best practices for  
phased deployment  
of convergence

10GE Performance  
Cost Reductions

Convergence Solves  
Actual problems of  
scale, flexibility, and  
complexity

Server Virtualization  
Network Evolution  
New Data Center Models

## ACCELERATORS

- Please send any questions or comments on this presentation to SNIA: [tracktutorials@snia.org](mailto:tracktutorials@snia.org)

**Many thanks to the following individuals  
for their contributions to this tutorial.**

**- SNIA Education Committee**

**Joseph L White  
Simon Gordon  
Gunes Aybay  
Charles Waters  
Andy Ingram**