



Education

IP SANs - How to Consolidate and Maintain High Utilization

Dave Zenz, Dell Inc.

- The material contained in this tutorial is copyrighted by the SNIA.
- 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.

➤ IP SANs - How to Consolidate and Maintain High Utilization

This session will appeal to IT administrators and architects who are seeking a fundamental understanding of IP SANs, Consolidation, and Utilization. Additionally, detail on how Server Virtualization works together with Storage Consolidation will be reviewed.

As the amount of data being generated continues its hyper growth, even while costs and expenditures are looking to be controlled, the benefits of IP SANs can be increasingly paramount. This session will delve into how IP SAN technology can be applied in today's environment and the challenges that can be addressed.

IP SANs - How to Consolidate and Maintain High Utilization

- Or in other words...
- How to do more with less now, and in the future, by leveraging what you already have



Check out SNIA Tutorials:

**Storage: Storage
Consolidation With IP
Storage**

**Storage Management:
Introduction to Storage
Management**

➤ IP (iSCSI) SANs

- ◆ Benefits, Architecture, and Features

➤ Consolidation

- ◆ Types, Benefits, and Process

➤ Keys to consolidation via IP SANs

- ◆ Consolidation Scenarios & Server Virtualization

➤ Maintaining Utilization

- ◆ Categories and the “Be’s”

Benefits of iSCSI Based SANs

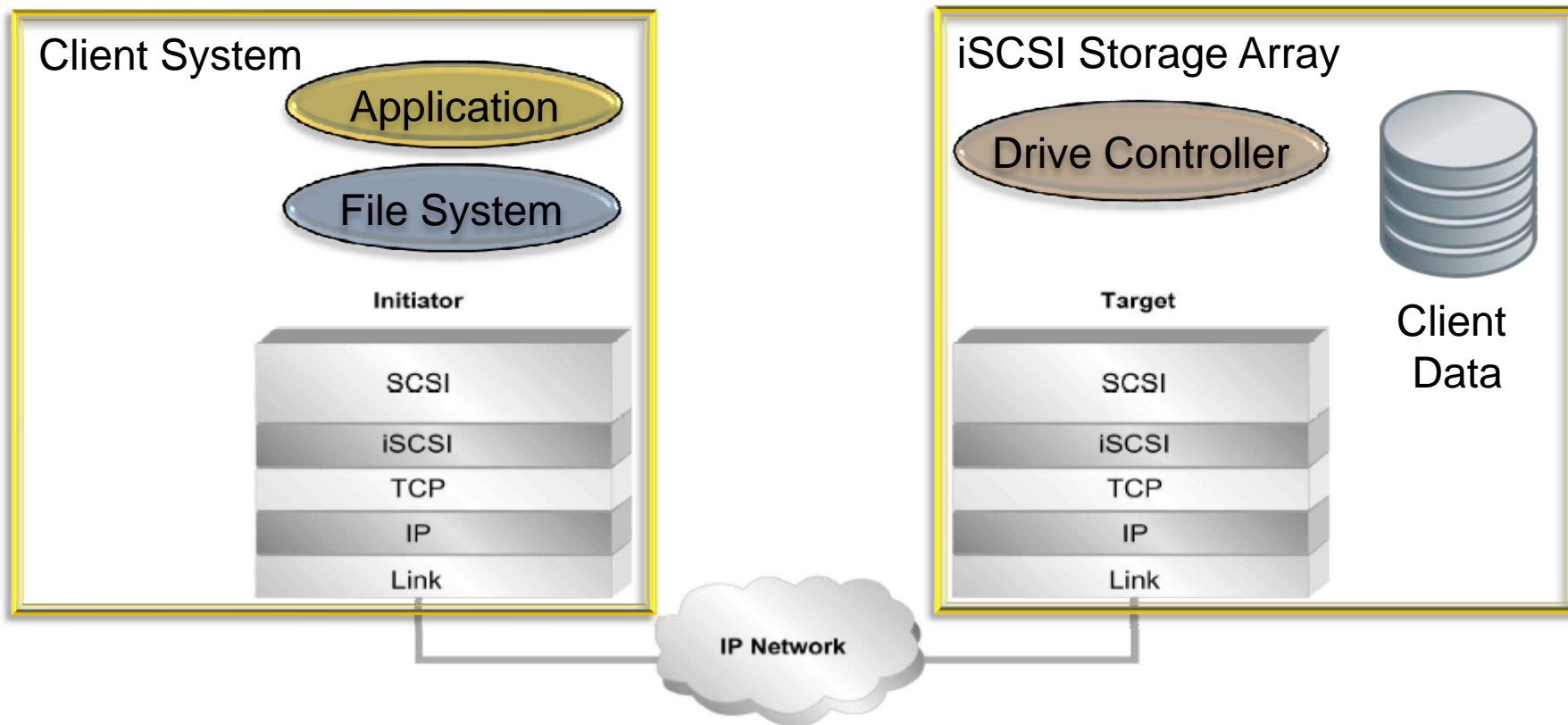
- Scalable solution cost
- Ubiquitous industry support + broad interoperability
- Internet Engineering Task Force standard
- Integrated error recovery
- *End Point Secure* – CHAP
- *Efficient* – “Jumbo” frames (9K bytes)
- *Fast* – up to 10Gbps

What is iSCSI?

- In computing, iSCSI is Internet SCSI (Small Computer System Interface), an Internet Protocol (IP)-based storage networking standard for linking data storage facilities.
- iSCSI can be used to transmit data over local area networks (LANs), wide area networks (WANs), or the Internet and can enable location-independent data storage and retrieval

Definition source: Wikipedia

iSCSI Overview



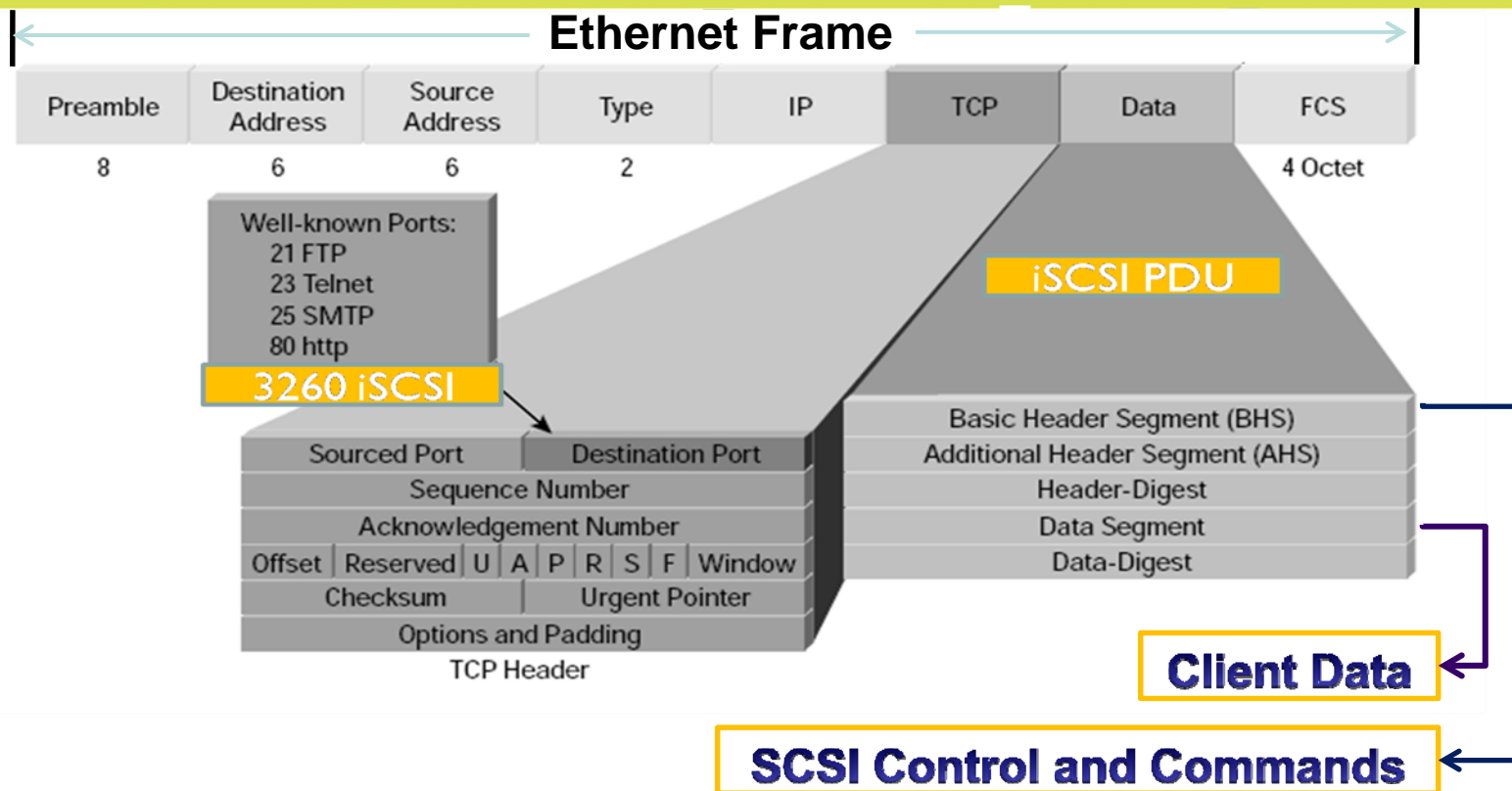
- iSCSI is a transport protocol that provides for the SCSI protocol to be carried over a TCP based IP network.

Definition source: SNIA Dictionary

Example iSCSI Interoperability Support Statements

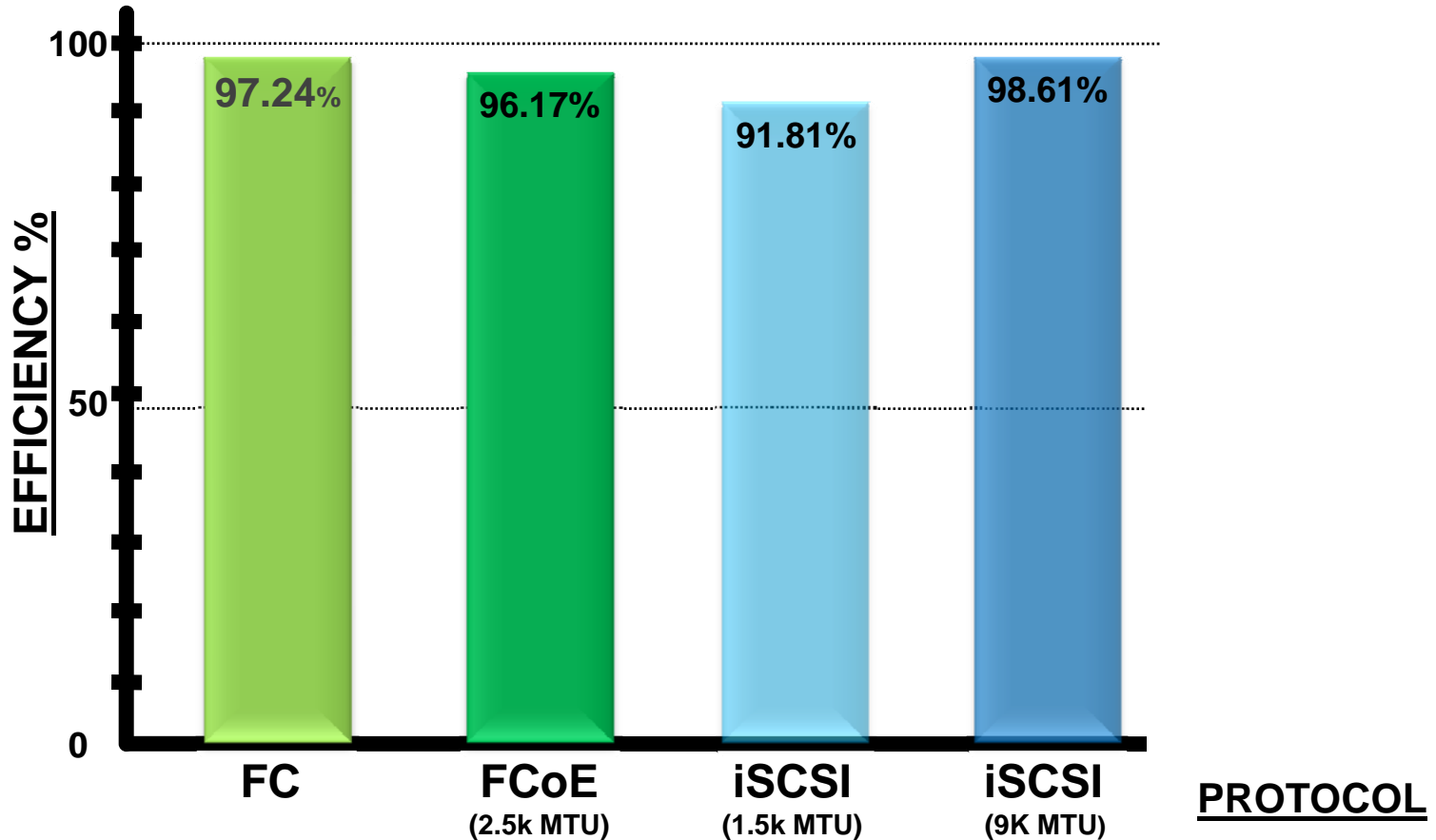
- Any industry standard 10/100/1000 Ethernet controller that conforms to IEEE 802.3ab, 802.3ac, or 802.3u is compatible with the [array].
- Any industry standard 100/1000 managed or unmanaged Ethernet switch is compatible with the [array].

iSCSI Packet and “Jumbo” Frames



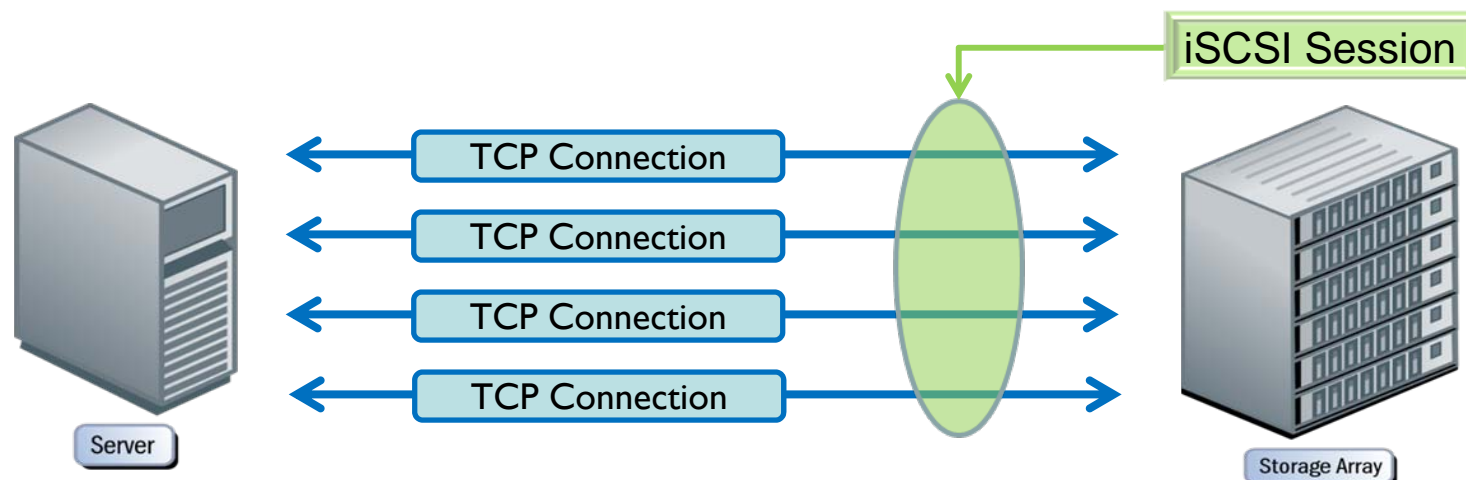
- iSCSI can leverage > 1500 byte MTU support in Ethernet devices – up to 9000 bytes
 - ◆ This is referred to as a Jumbo frame
- Jumbo frames provide greater efficiency

Protocol Efficiency



➤ Effectively the efficiency of iSCSI w/ Jumbo frames is the same as FC and FCoE

- Multiple TCP/IP Connections can be grouped into an iSCSI Session (MC/S)
 - ◆ Authentication occurs with each Connection
 - ◆ Load balancing available across Connections for > 1 aggregate performance and high availability
 - › Examples: Round Robin, Least Queue Depth, Least Path Weight
 - ◆ Lost Connections and Session can be recovered

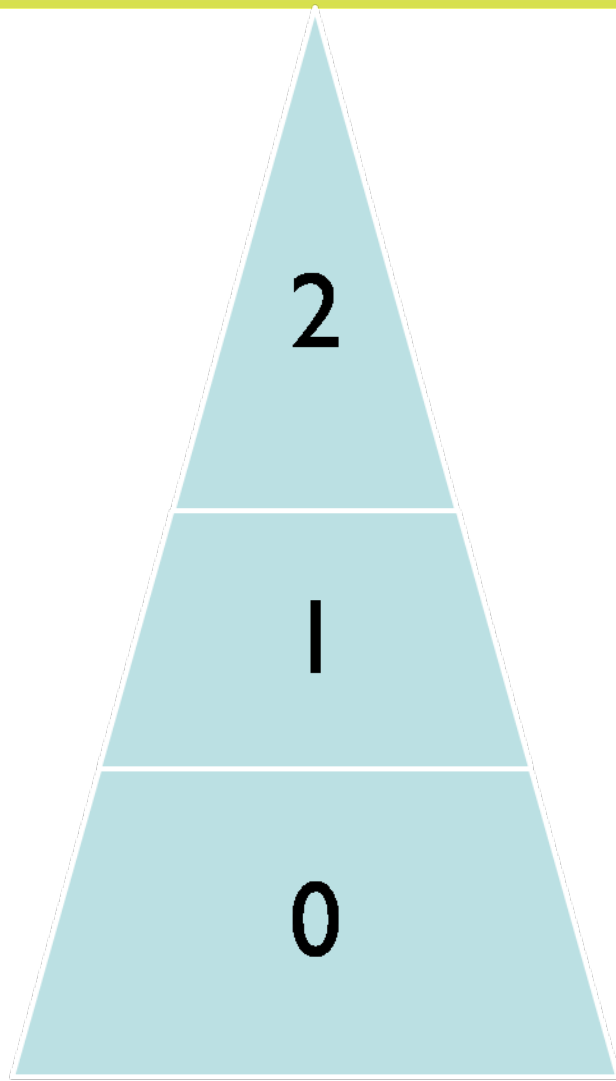


iSCSI Error Recovery Level Hierarchy

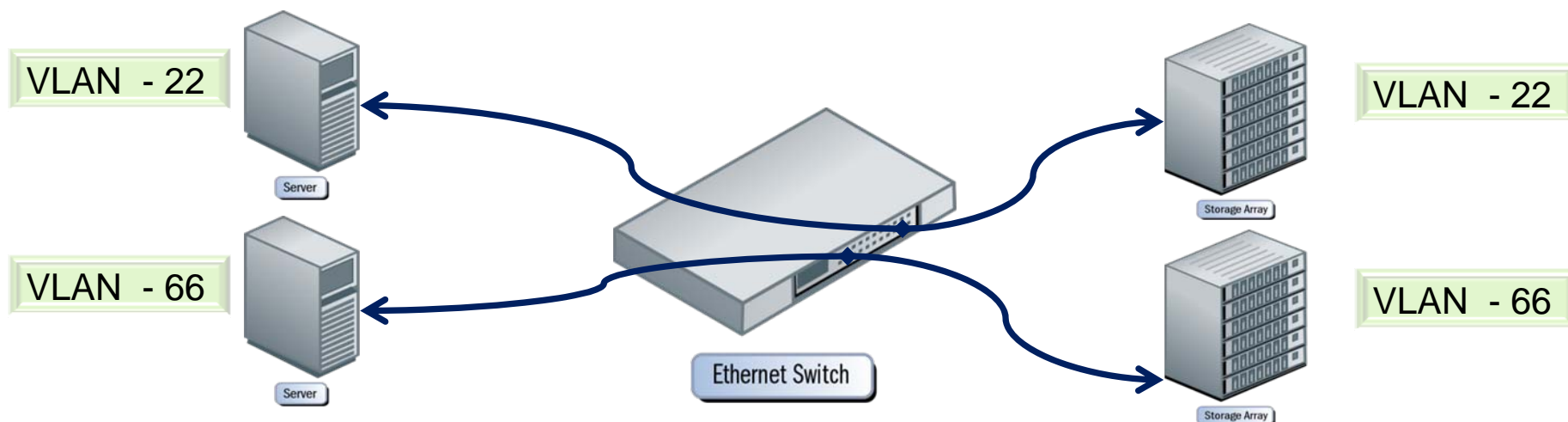
- **ERL 2 – TCP/IP Connection Error Recovery**
 - ◆ Reestablishment of Connection failures without application disruption (requires ERL 1)

- **ERL 1 – Command and Data Recovery**
 - ◆ CRC checks for iSCSI command and data errors called Header and Data Digests (requires ERL 0)

- **ERL 0 – iSCSI Session Error Recovery**
 - ◆ Reestablishment of iSCSI Session errors without application disruption



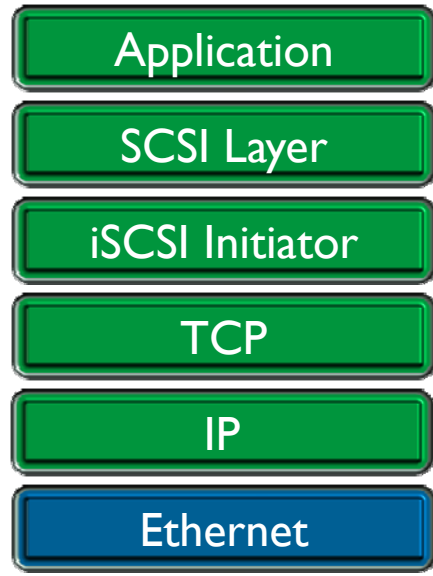
- VLANs provide a way to separate shared switched networks into private logical networks
- Each VLAN has its own broadcast domain
- Per port configuration
- Provides controlled traffic and network security



- ◆ Systems on VLAN 22 & 66 operate as if on separate physical networks

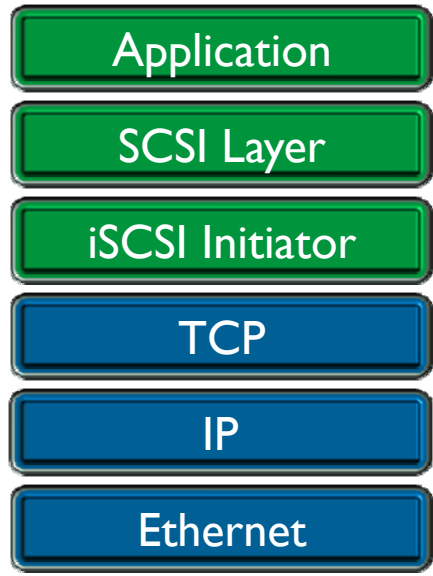
iSCSI Stack Hierarchy and Flexibility **SNIA**

Software Initiator



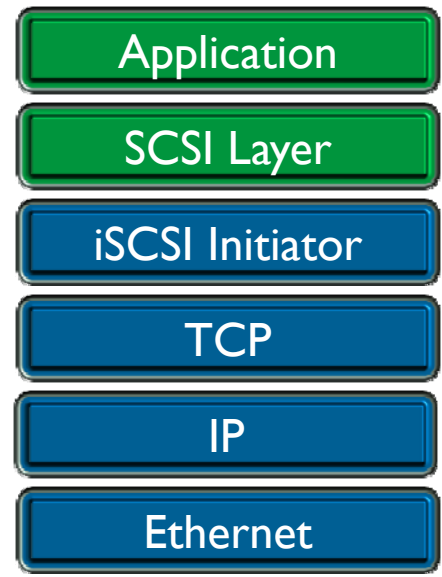
- Lowest cost (\$0)
- Highest CPU impact
- Ubiquitous driver support

TCP Offload



- TOE's increasingly offered on LOM's
- Efficiency gained is independent of iSCSI
- TOE driver support needed

iSCSI HBA

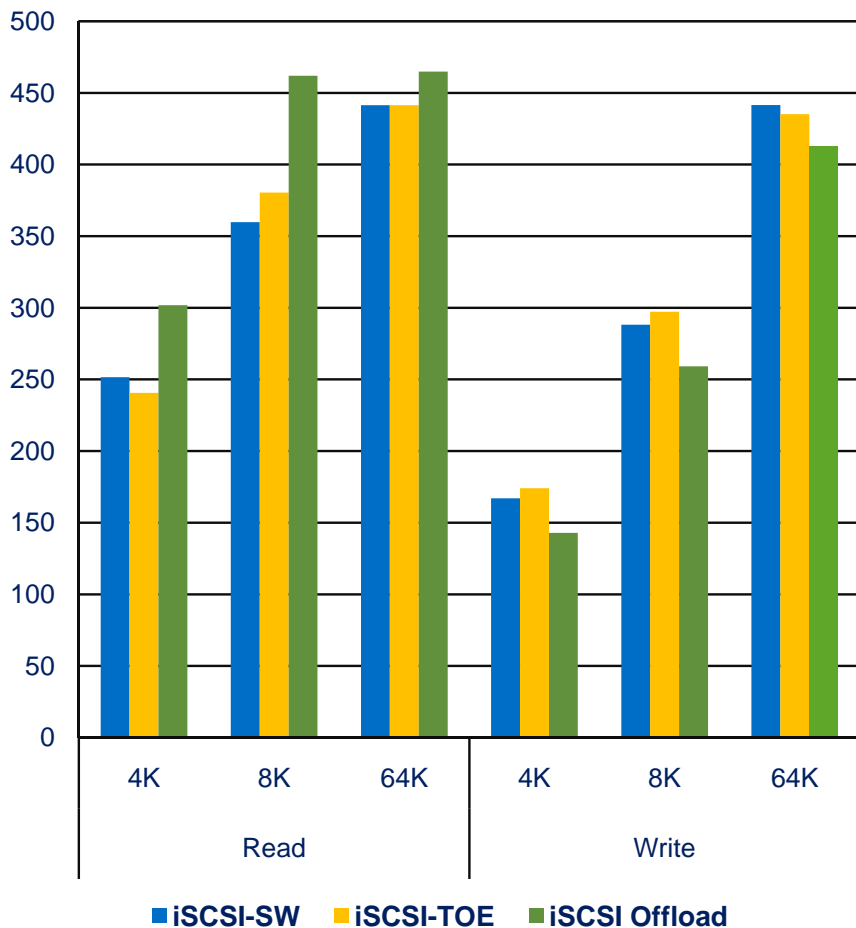


- Highest cost
- Complete iSCSI offload including digests
- HBA driver support needed

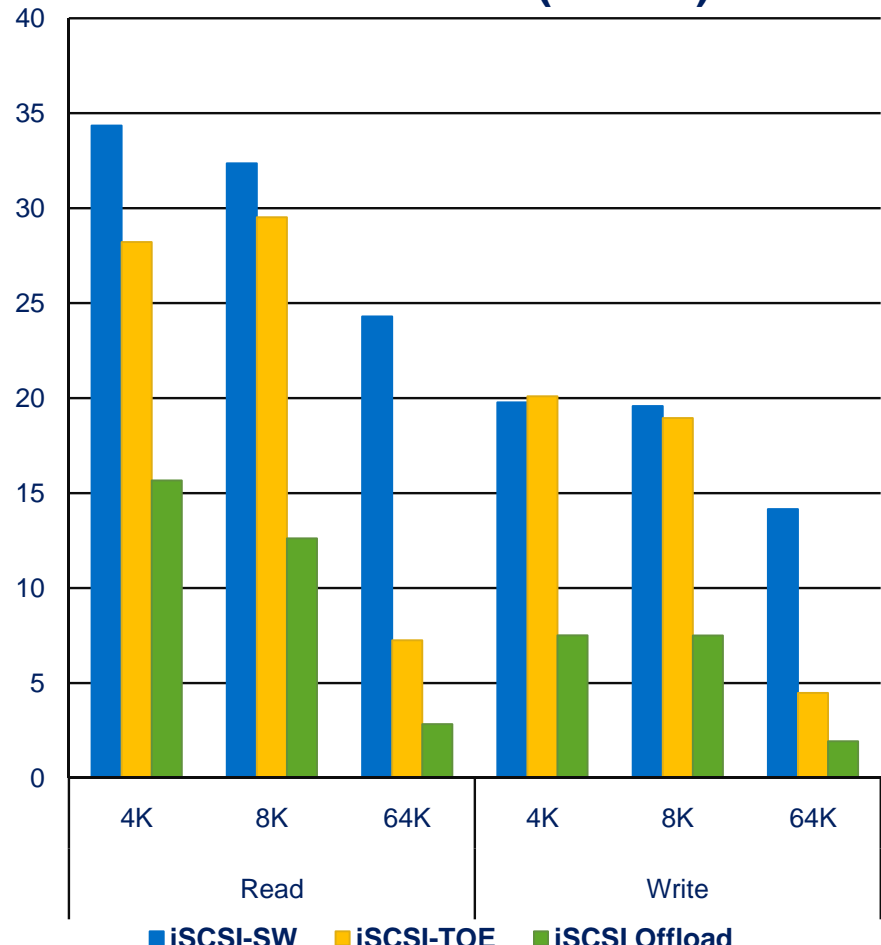


iSCSI Stack Relative Performance

Throughput (MBps)



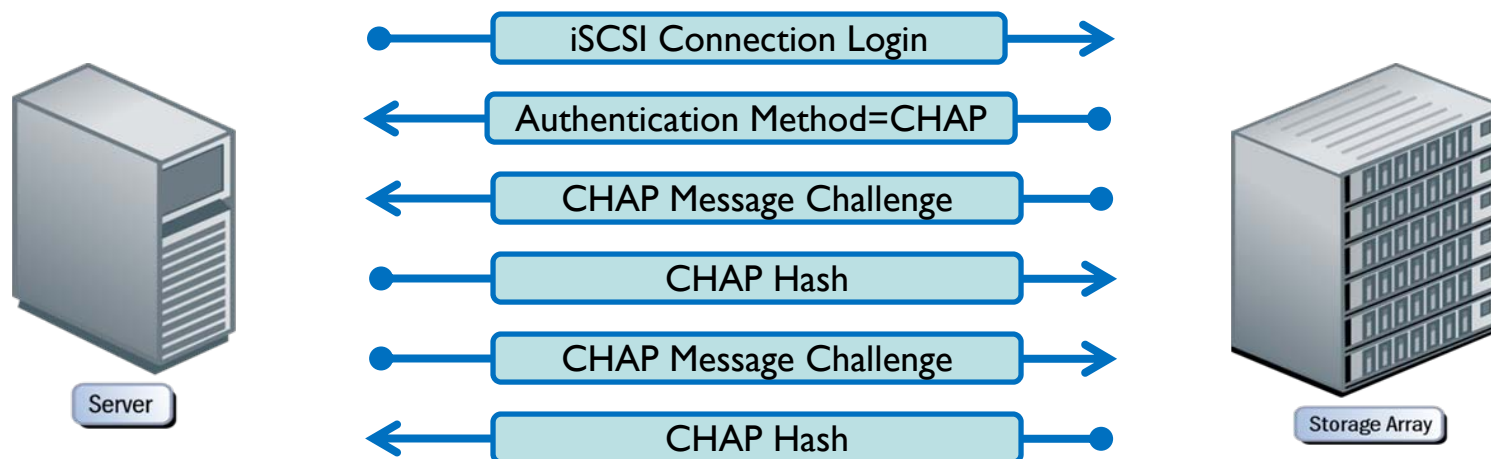
CPU Utilization (%CPU)



➤ Choose the appropriate solution for cost/performance

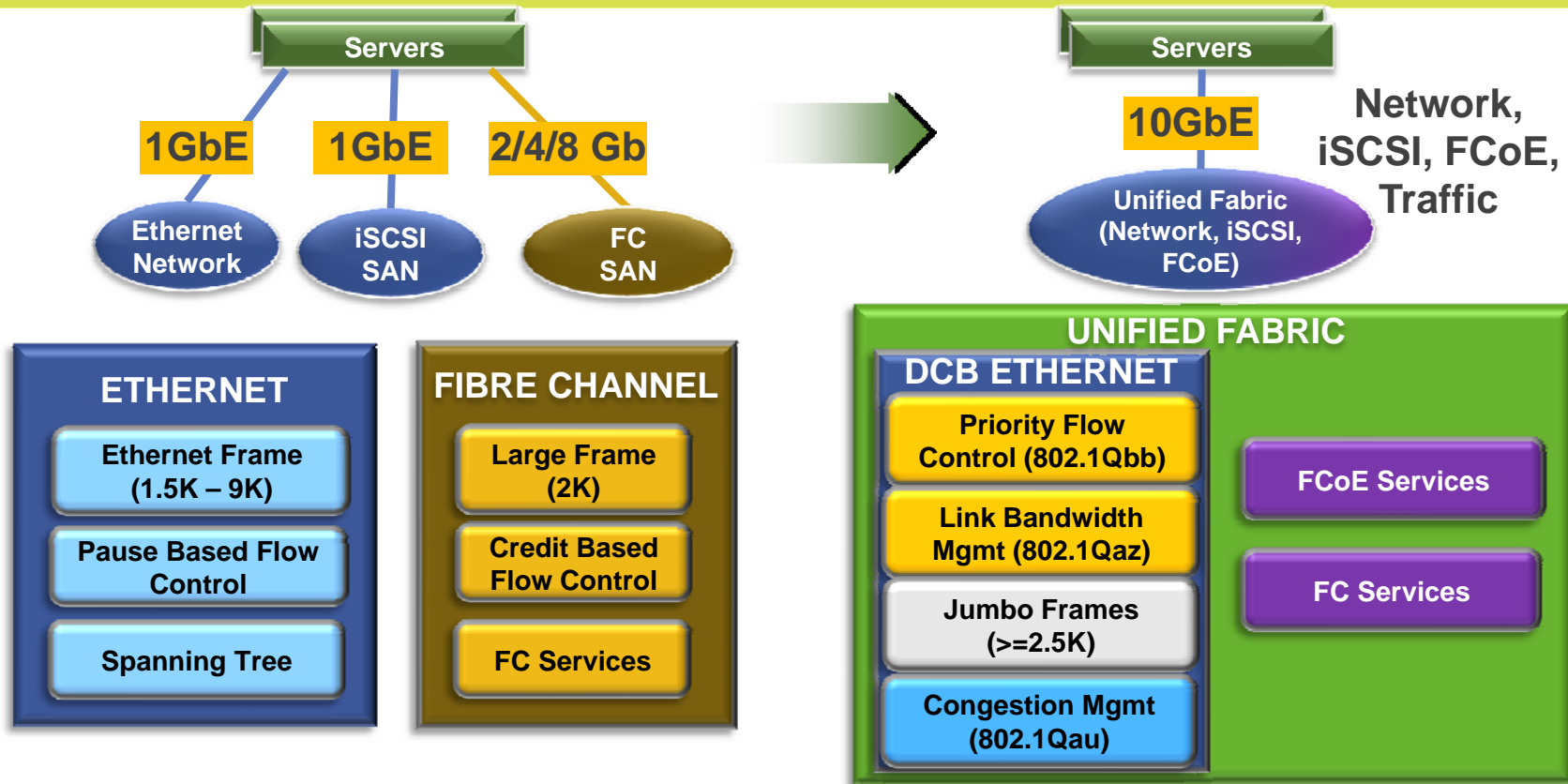
Data Source: Internal Dell paper - contact author for details.

- In-band authentication provides end-to-end trust
 - ◆ Established with each iSCSI Connection Login
- Target and Initiator models
 - ◆ Array can require the Host to authenticate
 - ◆ Host can require the Array to authenticate
 - ◆ If both are used Secrets must be different
- As with all passwords, CHAP secrets should be strong



- ◆ Mutual (Target and Initiator) Sequencing Example

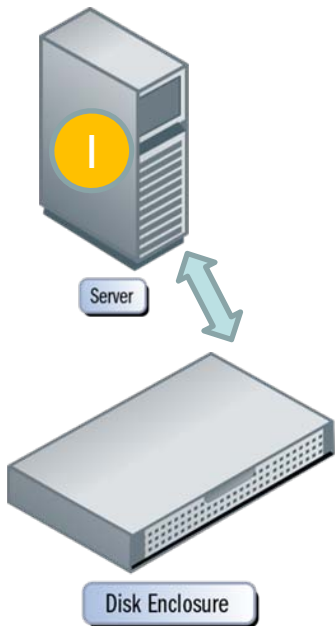
iSCSI Futures - 10Gb and DCB



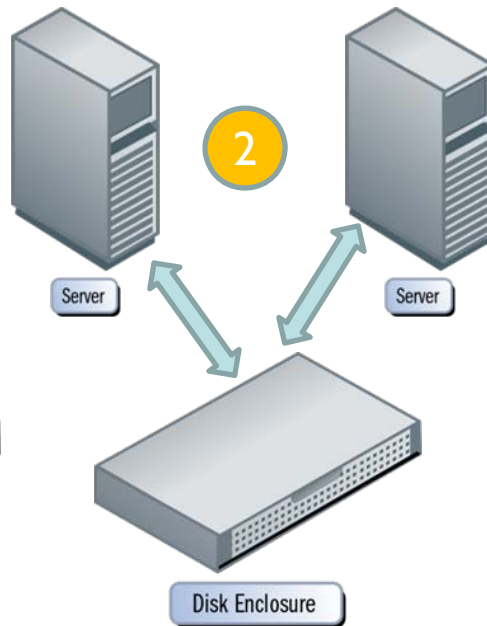
- Key value of Unified Fabric is traffic convergence
- iSCSI traffic will have end-to-end QoS with DCB
 - ◆ Utilizing Priority Flow Control and Link Bandwidth Management

Example Distributed Storage Model

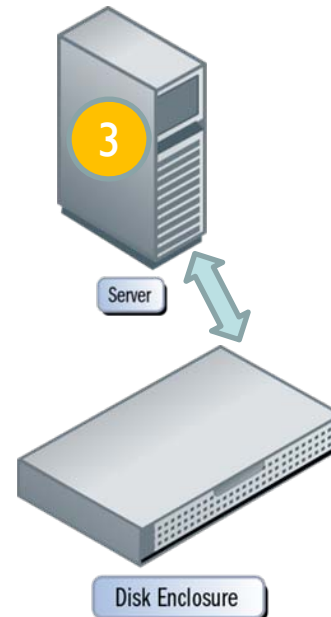
Web



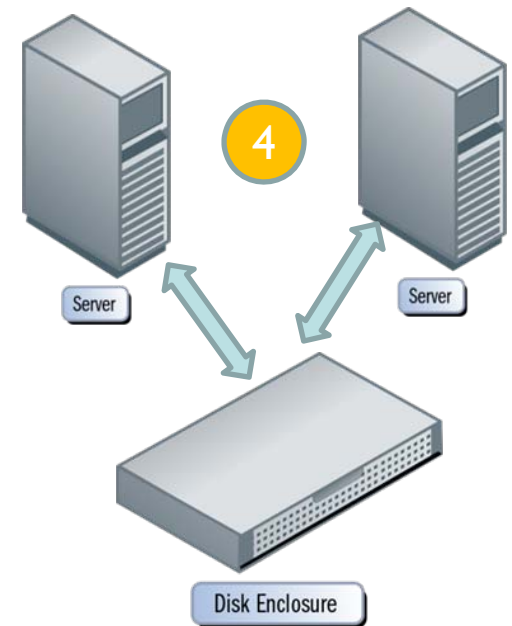
DB Nodes



File



Mail Nodes



- Example has four direct connected storage touch points
- Associated storage could be heterogeneous
 - ✦ Different RAID controllers, software, protocols, monitoring
- Significant complexity to manage

➤ **Increased Staff Productivity**

- ◆ Administrators can manage more storage capacity in a consolidated environment

➤ **Optimized Availability**

- ◆ Provide high data availability for business-critical applications

➤ **Improved Utilization and Scalability**

- ◆ Disk utilization can be increased through centralization and capacity scaled on demand

➤ **Improved Manageability**

- ◆ Administrators can gain control of the growth of data and respond to business needs more rapidly

➤ *Application*

- ◆ Combining mixed applications onto a storage system
- ◆ Key benefit
 - Cross application DR, archiving, and security procedures can be defined

➤ *Workload*

- ◆ Combining common workloads into fewer storage systems
- ◆ Key benefit
 - Improved service to users in the consolidated environment

➤ *Physical*

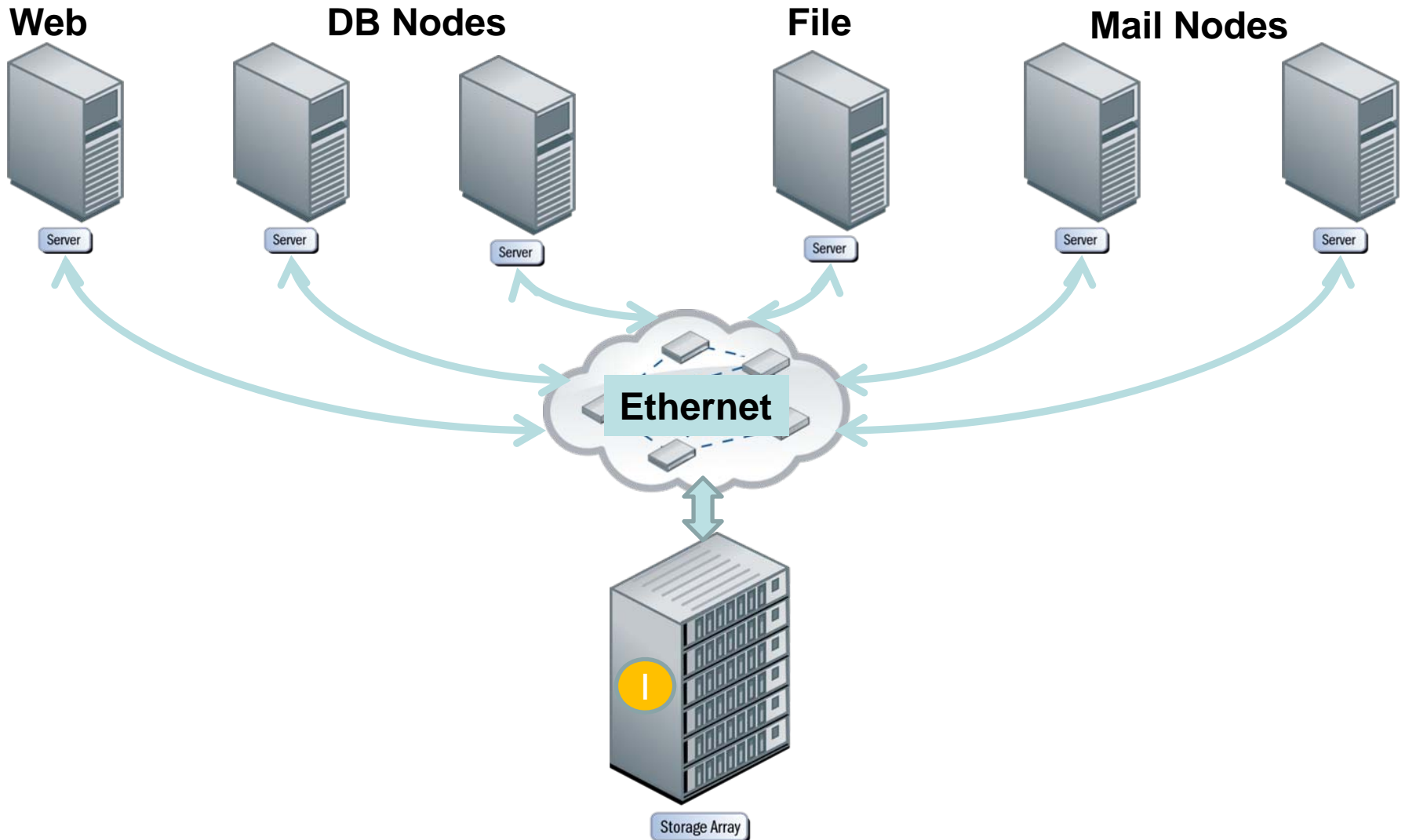
- ◆ Centralize and reduce IT storage assets
- ◆ Key benefit
 - › Improved storage management and asset control

➤ *Logical*

- ◆ Consolidation of the number of consoles used by the operations staff
- ◆ Key benefit
 - › Common processes across systems, including automation, can be defined

- Examine business requirements including mission critical applications and related SLAs
- Take inventory of your applications and their usage requirements
- Develop master plan with realistic schedules
- Prioritize and group opportunities to minimize iterations
- Plan for validation
- Consolidation needs to be a *strategic IT direction*
- Services are available to assist in the process

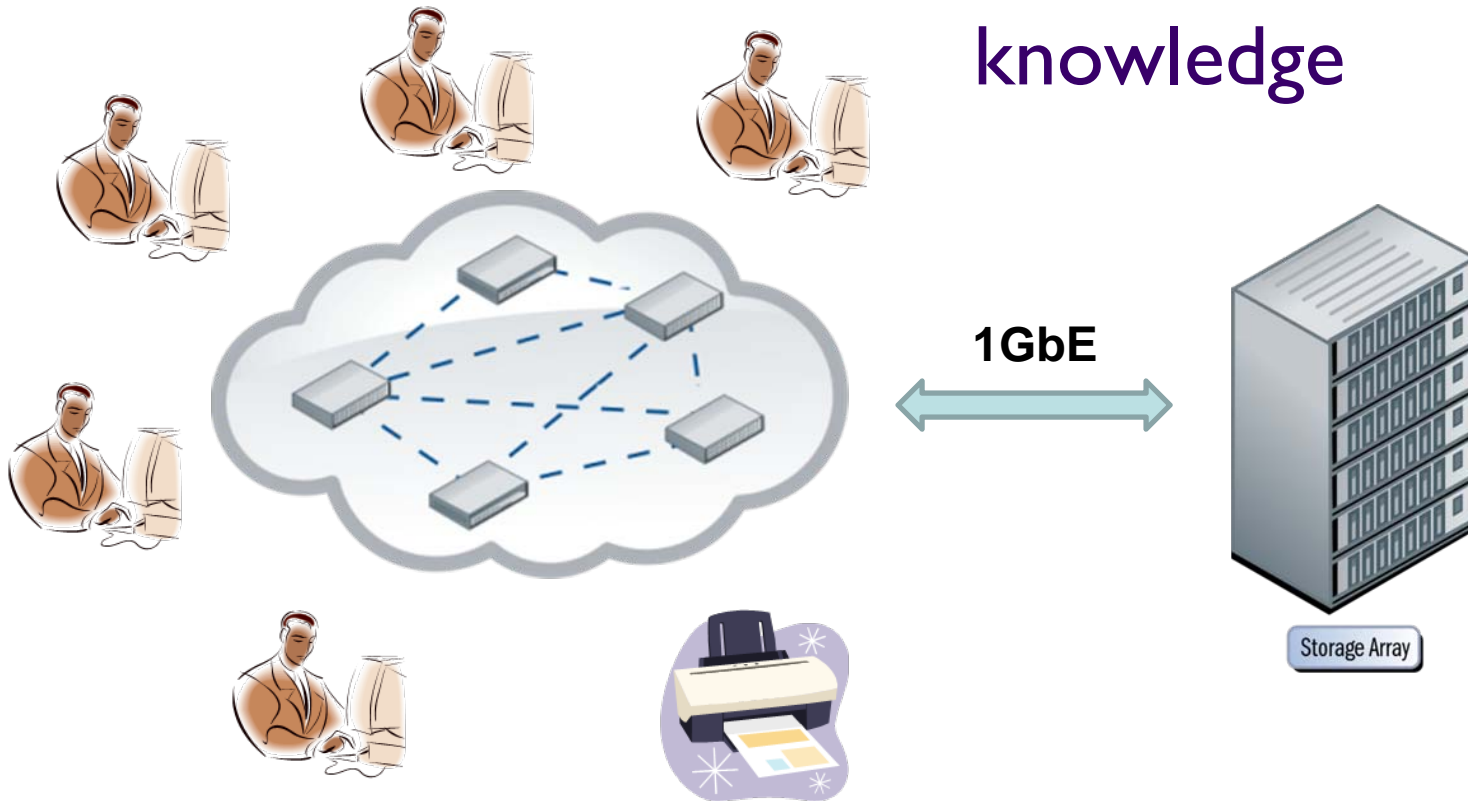
General IP SAN Consolidation



➤ Consolidated to one storage management touch point

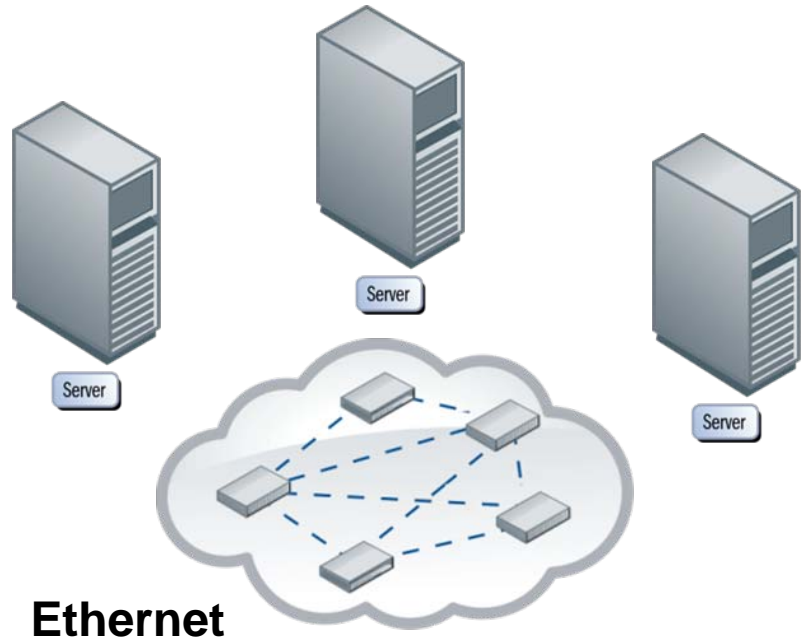
Departmental IP SAN Consolidation

- Shared files
- Print services
- User backups
- Leverage existing IP infrastructure & knowledge

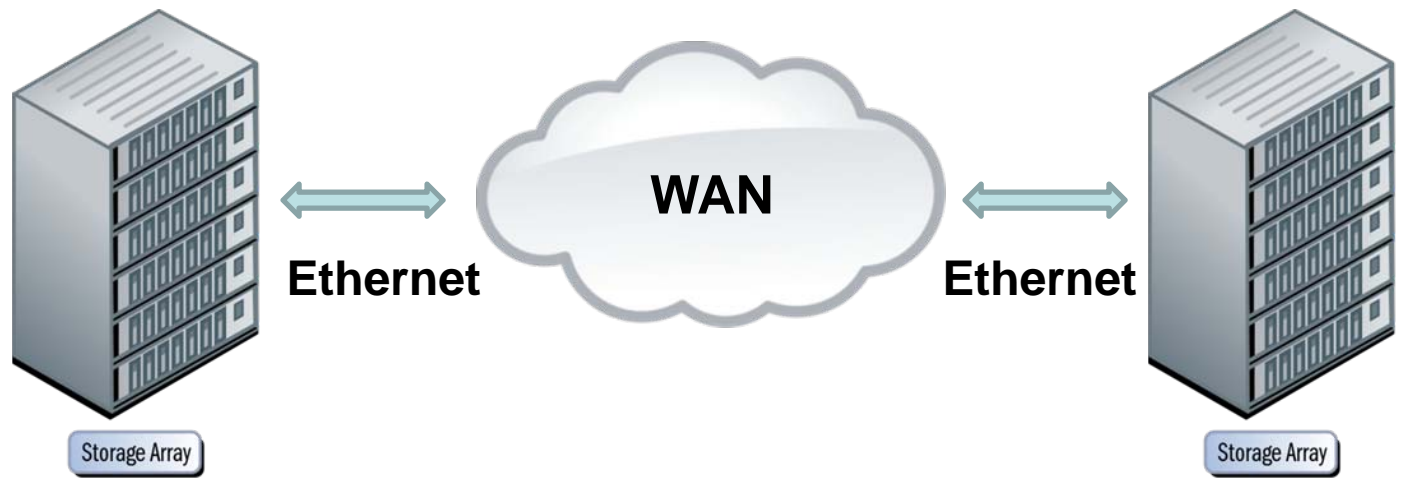


IP SAN over WAN Consolidation

- Disaster recovery
- Tiered storage
- To the Cloud
- Remote office to data center
- No FCIP gateways required



Ethernet



➤ Backup infrastructure

- ◆ Standardize backup hardware and software

➤ Blade servers

- ◆ Storage for server aggregation

➤ NAS gateway to an IP SAN

- ◆ Consolidate file storage

➤ Clusters

- ◆ Aggregate cluster nodes

➤ Updating infrastructure

- ◆ Retiring “tired” SANs

- **Server & Storage consolidation go hand-in-hand to address**
 - ◆ Capital expenditures
 - ◆ Datacenter costs
 - ◆ Power, cooling, and raised floor space
 - ◆ Backup and disaster recovery

- **IP SANs are a natural fit for server virtualization**
 - ◆ Leverage existing Ethernet processes and tools
 - ◆ Utilize existing server and switch hardware
 - ◆ Seamless movement of VM's independent of how storage is provisioned to the OS

Categories of Storage Utilization

➤ Capacity

- ◆ Under-subscribed – planning estimates too large
- ◆ Over-subscribed – hyper data growth
- ◆ Dark – can't utilize space if it's not visible

➤ Performance

- ◆ Application – host focus perspective for a solution view
- ◆ Under-utilized – workload balance out of kilter
- ◆ Over-utilized – demand outstripping architected capability

➤ Availability

- ◆ Degrees of redundancy – multi-path, clusters, & virtualization

➤ *All of these can be adjusted for within a consolidated infrastructure*

- iSCSI arrays may also have built-in capabilities that can assist with maintaining utilization
 - ◆ Thin Provisioning
 - › Avoid inefficiencies associated with over provisioning
 - ◆ Load balancing
 - › Movement of data to eliminate hot spots and performance contention
 - ◆ Tiered storage
 - › Placement of data where it makes sense for tiered application usage and data longevity

- **Be Proactive – with management**
 - ◆ This will go a long way to prevent those “oh-no” moments
- **Be Aware – of capacity and performance utilization**
 - ◆ Trending detail can be of enormous value to plan by
- **Be Tied In – with your customers**
 - ◆ Spikes in demand due to unplanned events can severely test a systems utilization
- **Be Certain – of the availability of your SAN**
 - ◆ Don’t let a path failure affect application availability
- **Be Prepared – and plan for maintenance**
 - ◆ Keep your SAN components up to date

- Remember most of all that *Consolidation* is not a one shot deal, it's an ongoing journey...
- Plan and enjoy the ride! ;-)

- Please send any questions or comments on this presentation to SNIA: tracknetworking@snia.org

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

- SNIA Education Committee

**Howard Goldstein
Joe White
Rob Peglar**

Appendix: Terms

- CHAP – Challenge Handshake Authentication Protocol
- DCB – Data Center Bridging
- ERL – Error Recovery Level
- IP SAN – iSCSI based Storage Area Network
- LOM – LAN On Motherboard
- MC/S – Multiple Connections per Session
- MTU – Maximum Transmission Unit
- TOE – TCP/IP Offload Engine
- VLAN - Virtual Local Area Network

Appendix: Protocol Efficiency Data

Ref	Protocol	MTU	Data	Total Frame Size	Efficiency
1	FC	2112	2112	2172	97.24%
2	FCoE	2112	2112	2196	96.17%
3	iSCSI over TCP/IPv4 (all frames)	1500	1412	1538	91.81%
4	iSCSI over TCP/IPv4 (all frames)	9000	8912	9038	98.61%

ENSOF = 4 DataFC = 2112
 ENHeader = 14 DataEN = 1500
 ENCRC = 4 DataFCoE = 2112
 ENEOF = 4 DataENJumbo = 9000
 ENIFG = 12
 FC SOF = 4
 FCHeader = 24
 FCCRC = 4
 FCEOF = 4
 FCIFG = 24
 FCoEHeader = 14
 FCoETrailer = 12
 TCPHeader = 20
 IPv4Header = 20 (no options)
 iSCSI Header = 48

- 1) $\text{DataFC} / (\text{FC SOF} + \text{FCHeader} + \text{DataFC} + \text{FCCRC} + \text{FCEOF} + \text{FCIFG}) = 2112 / (4 + 24 + 2112 + 4 + 4 + 24) = 0.9724$
- 2) $\text{DataFCoE} / (\text{FC SOF} + \text{FCoEHeader} + (14 + \text{FCHeader} + \text{DataFCoE} + \text{FCCRC} + 4) + \text{FCRC} + \text{FCEOF} + \text{ENIFG}) = 2112 / (4 + 14 + (14 + 24 + 2112 + 4 + 4) + 4 + 4 + 12) = 0.9617$
- 3) $(\text{DataEN} - \text{TCPHeader} - \text{IPv4Header} - \text{iSCSIHeader}) / (\text{ENSOF} + \text{ENHeader} + \text{DataEN} + \text{ENCRC} + \text{ENEOF} + \text{ENIFG}) = (1500 - 20 - 20 - 48) / (4 + 14 + 1500 + 4 + 4 + 12) = 0.9181$
- 4) $(\text{DataENJumbo} - \text{TCPHeader} - \text{IPv4Header} - \text{iSCSIHeader}) / (\text{ENSOF} + \text{ENHeader} + \text{DataENJumbo} + \text{ENCRC} + \text{ENEOF} + \text{ENIFG}) = (9000 - 20 - 20 - 48) / (4 + 14 + 9000 + 4 + 4 + 12) = 0.9861$