Deploying iSCSI in the Datacenter

SNIA™
Ethernet Storage Forum
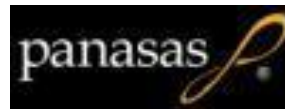
Ethernet Storage Forum Members

The SNIA Ethernet Storage Forum (ESF) focuses on educating end-users about Ethernet-connected storage networking technologies.

Housekeeping

- ◆ Questions
  - ◆ We will answer questions during the presentation.
  - ◆ Submit questions using the Questions button above the presentation.
  - ◆ Any unanswered questions will be answered at http://sniaesfblog.org/
- ◆ Votes
  - ◆ We will ask a few questions of the attendees during the presentation, and these will be indicating by a "Voting" page. Please respond by voting for one of the choices on each of these pages.
- ◆ Ratings
  - ◆ At the end of this presentation, there will be an opportunity for you to rate this presentation and provide additional feedback.

Webcast Presenters

Host

Jason Blosil, NetApp, Currently, Jason works in Solutions Marketing and is responsible promoting the advantages of iSCSI and FCoE storage. His is actively involved in partner and product marketing, as well as thought leadership activities. He is currently a co-chair of the SNIA Ethernet Storage Forum iSCSI SIG, as well as a member of the Ethernet Alliance and Marketing Committee and Solutions Committee at Blade.org.

Presenter

Dennis Martin is the founder and President of Demartek, a computer industry analyst organization with its own on-site test lab. Demartek focuses on lab validation testing and performance testing of storage and related hardware and software products. Dennis has been working in the Information Technology industry since 1980, primarily involved in software development and project management in mainframe, UNIX, and Windows environments. These include a variety of large and small end-user customers, and engineering and marketing positions for storage vendors such as StorageTek.

Agenda

- Introduction
- Basic Storage Architectures
- Marketplace
- iSCSI Technology and Definitions
- Networking Technologies
- Storage Targets
- Server Virtualization
- Best Practices
- Summary

Demartek – Who are We?

◆ Demartek – Industry Analysts with On-site Test Lab

- Approximately 90% of our research is based on running hardware and/or software in our lab
  - › The majority of our work is public and available at www.demartek.com
- Lab includes servers, networking and storage infrastructure
- We use all of the following technologies in our lab
  - › Ethernet: 1Gbe, 10GbE, WiFi (802.11 b/g/n)
  - › Fibre Channel: 2Gb, 4Gb, 8Gb
  - › FCoE: 10Gb
  - › iSCSI: 1Gb, 10Gb
  - › Unified storage
  - › SSDs (form factors: drives, PCIe, other)
  - › Virtualization (VMware, Microsoft, others)
  - › Windows, Linux, others

Poll #1 – What is Your Job Role?

- IT Administrator (server, network, storage, database, etc.)
- Software/Hardware Developer
- IT Manager, Director or CIO
- Sales or Marketing
- Press or Analyst

Poll # 2 – Size of Organization

- Less than 50
- 50-249
- 250-999
- 1000-5000
- 5000+

Introduction

- ◆ iSCSI – Internet SCSI
  - ◆ Industry standard
  - ◆ One form of a Storage Area Network (SAN)
  - ◆ Transmission of SCSI block storage commands over Ethernet networks using TCP/IP
    - › SCSI – Small Computer System Interface (not limited to "small" systems)
    - › SCSI commands are "encapsulated" inside of TCP/IP packets
  - ◆ Works over LAN and WAN

- ◆ Not surprisingly, to fully appreciate iSCSI some knowledge of both networking and storage are required

- ◆ Storage protocols require low latency and consistent response

Basic Storage Architectures

- ### DAS – Direct Attached Storage
  - Host server has private connection to storage and owns the storage
  - Used for "block" access methods
  - Distance between host CPU and storage is typically limited to within the chassis or within a rack or to an adjacent rack

- ### NAS – Network Attached Storage
  - File server shares its storage resources with clients on the network
  - Clients use network "file" access methods such as CIFS/SMB or NFS
  - The file server owns the storage but shares it with clients
    - File server advertises "shares" or "mount points" to the clients
    - File server internally uses "block" access methods behind the scenes
  - Distance between the clients on the network and the file server can be large

Basic Storage Architectures

- ◆ **SAN – Storage Area Network**
    - ◆ Provides "block" access methods over a network such as Ethernet or Fibre Channel
    - ◆ No one server owns the storage, but all servers can have access to it
        - › Some server/appliance solutions technically own the SAN storage but provide a type of pass-through access to it
    - ◆ Storage is typically carved into logical volumes and assigned to specific servers
    - ◆ Storage appears the same as DAS to the application server
   - ◆ Distance between the application server and the storage can be large

- ◆ **Unified Storage**
    - ◆ Combines NAS and SAN into one solution

Poll # 3 – Using iSCSI Today

- ◆ We use iSCSI storage today

- ◆ Planning to deploy iSCSI storage within the next 6 months

- ◆ Planning to deploy iSCSI storage in 6-12 months

- ◆ We are considering iSCSI storage for the future

- ◆ Don't know or are not in the IT department

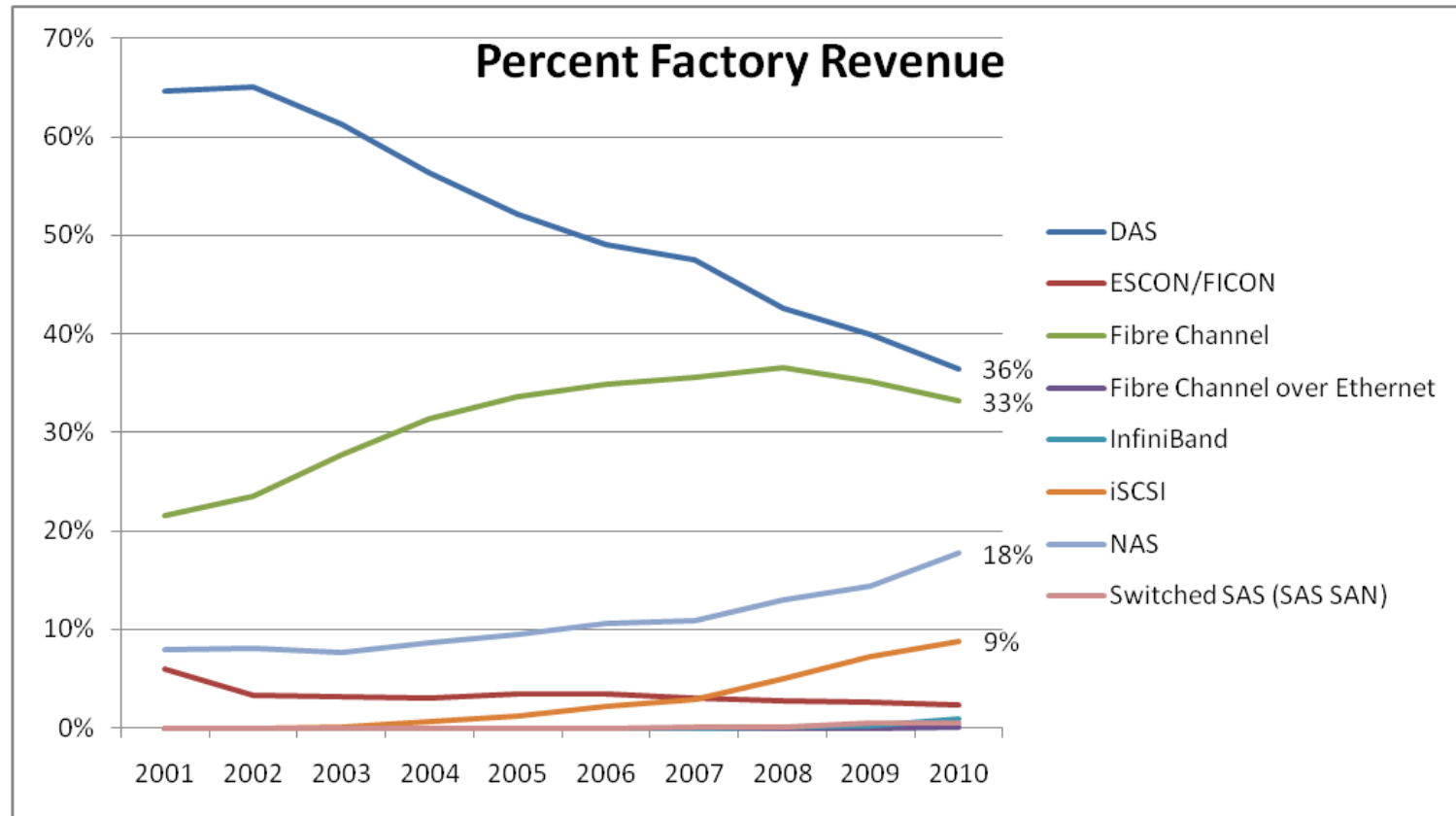## WW Server Port Shipments by Storage Interconnect Type
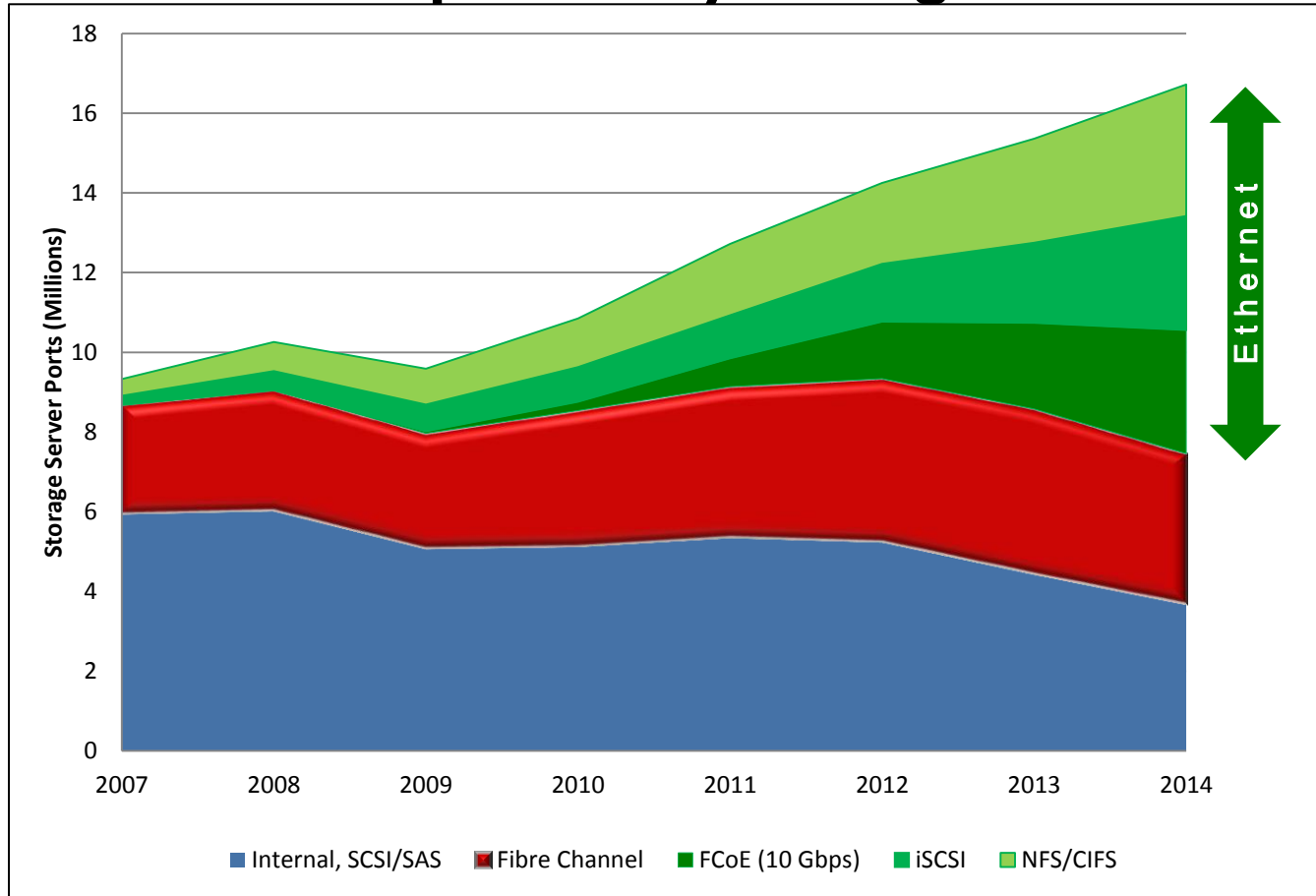
Source: IDC - Worldwide Storage Networking Infrastructure 2010-2014 Forecast, Dec. 2010

Petabytes Shipped
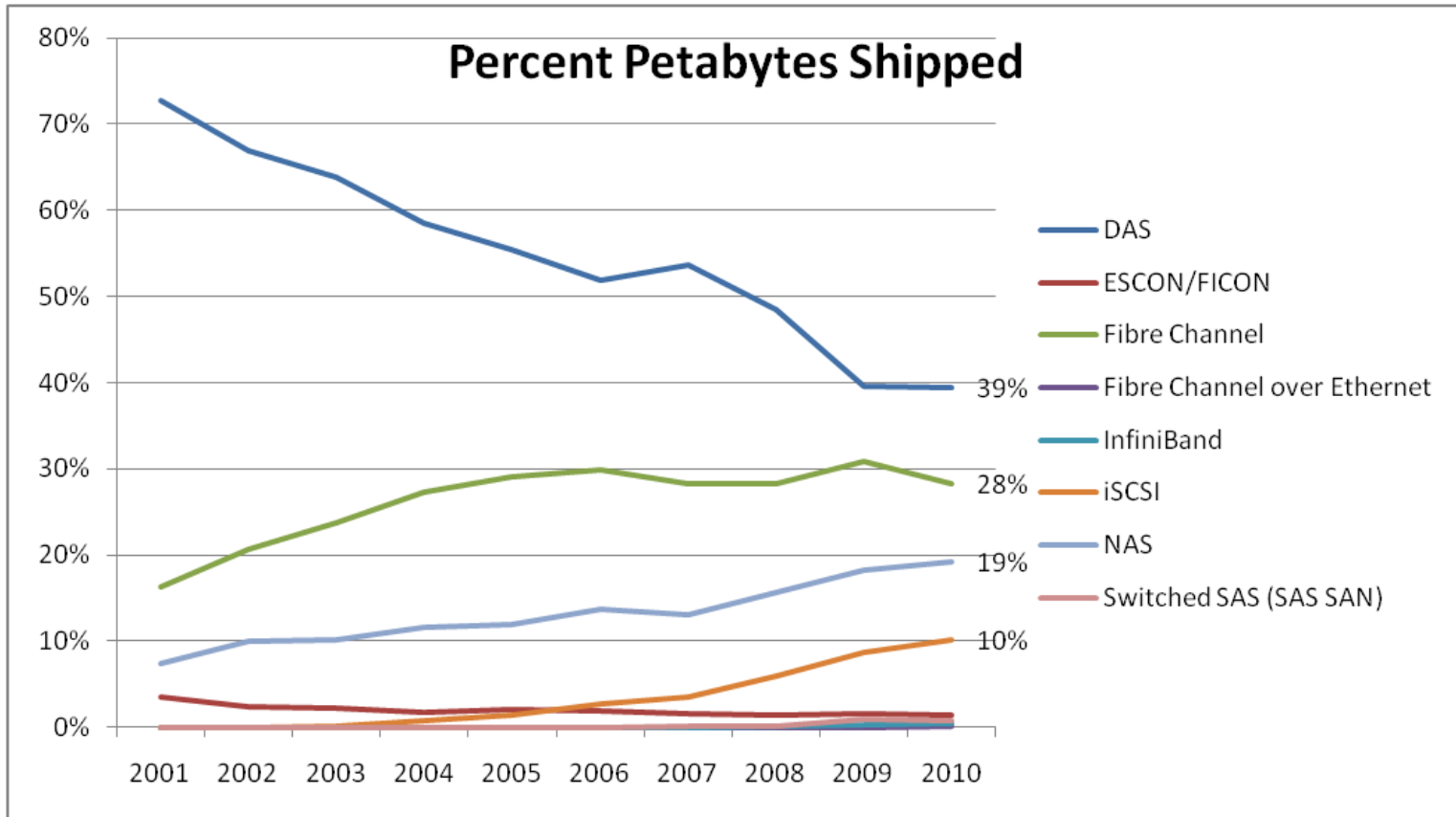
Source: IDC - Worldwide Disk Storage Systems Tracker, March 3, 2011

Factory Revenue

Source: IDC - Worldwide Disk Storage Systems Tracker, March 3, 2011

iSCSI Technology

iSCSI Initiators

- Host servers that need access to storage resources
- Follow the SCSI protocol for initiators
  - Issue commands to storage devices and wait for replies
- iSCSI initiators are included with or available for every major desktop and server operating system
- Can be implemented in software or hardware
- Can access multiple targets simultaneously
- Each initiator has a unique IQN (iSCSI Qualified Name)

iSCSI Technology

◆ iSCSI Targets

- ◆ Devices that provide storage resources for iSCSI initiators
- ◆ Follow the SCSI protocol for targets
  - › Respond to requests from initiators
- ◆ Can be implemented as dedicated hardware devices such as disk arrays or server/appliance solutions
- ◆ Can respond to multiple iSCSI initiators simultaneously
- ◆ Each target has a unique IQN (iSCSI Qualified Name)

iSCSI Technology

iSCSI sessions

- Sessions are established between iSCSI initiators and iSCSI targets
    - The iSCSI initiator performs a "logon" or "connect" with the target
- iSCSI sessions are established between a specific IP address on the initiator (host) and a specific IP address on the target (storage).
- A session can have multiple connections between the same initiator and target
    - This can be done to improve performance

iSCSI Technology

- ## iSCSI Security

  - CHAP – Challenge Handshake Authentication Protocol
    - Initiators and targets are configured with a shared secret and use a hash of that secret to authenticate communications between them
    - CHAP can be either one-way or mutual
  - IPsec – IP Security
    - Authenticates and encrypts the packets sent between the initiator and the target

- ## Performance

  - We are currently testing iSCSI configurations with various security settings to measure performance

Poll # 4 – iSCSI Security

- ◆ We use iSCSI but don't use CHAP or IPsec
- ◆ We use CHAP but not IPsec
- ◆ We use IPsec but not CHAP
- ◆ We use both CHAP and IPsec
- ◆ Don't know or are not in the IT department

iSCSI Technology

◆ Multipath I/O

- ◆ Multi-path I/O is supported by various iSCSI initiators and targets.
    - › Feature that can be enabled in Microsoft environments
- ◆ In multi-NIC environments, iSCSI sessions can be established between specific IP address pairs of initiators and targets to take advantage of multiple paths.
- ◆ Load-balance and failover policies can be specified, depending on the capabilities of the initiators and targets, including some or all of these:
    - › Round-robin, least queue depth, weighted path, failover, failback, etc.
- ◆ Target-specific MPIO drivers are usually required
    - › Microsoft: Device Specific Module (DSM)

Networking Technology

◆ Speed

- iSCSI in theory can use any speed of Ethernet, however best practice is to use either 1GbE or 10GbE
  - › We have performed some iSCSI tests using WiFi in our lab
  - › 10GbE prices are slowly dropping, making it a viable choice
  - › 10GbE is beginning to appear on server motherboards

◆ Cabling

- iSCSI (Ethernet) can use copper cabling such as
  - › 1GbE: Cat5e or Cat6
  - › 10GbE: Cat6 (up to 55m), Cat6a or Cat7
- iSCSI (Ethernet) can use fiber-optic cabling such as OM1, OM2, OM3, OM4 for 1Gbe and 10GbE

Networking Technology

♦ **Receive-Side Scaling (RSS) or Scalable I/O**

- Early TCP/IP implementations did not handle multi-core systems properly, but routed all incoming packets to one core in the system
- RSS balances incoming packet handling across cores, and requires some intelligence in the adapter
- RSS support is a recommended best practice for iSCSI

♦ **Virtual Machines**

- Some network adapters offload some virtual machine network handling, improving overall performance and lowering CPU utilization
- Not a requirement for iSCSI, but certainly helps

Networking Technologies

TCP/IP Offload

- Some network adapters have their own TCP/IP stack in hardware, offloading the main CPU
    - Some implement the Microsoft TCP Chimney offload architecture
- Can improve iSCSI performance and/or lower CPU utilization
- Available for 1GbE and 10GbE

iSCSI Offload

- Some network adapters have their own iSCSI initiator in hardware
- Can improve iSCSI performance and/or lower CPU utilization
- These adapters usually appear to the system as storage adapters (HBAs) rather than network adapters
- Available for 1GbE and 10GbE

Networking Technologies

Jumbo Frames (Jumbo Packets)

- Jumbo frames can be used to allow more data to be transferred with each Ethernet packet, reducing server overhead (protocol, checksum, etc.)
  - Standard size is ~ 1500 bytes
  - Common jumbo size is ~ 9000 bytes
  - Other sizes are ~ 4000 bytes, 14000 bytes
- In order to use jumbo frames, each physical and virtual component in the path must support jumbo frames, including the adapters, switches, routers and storage targets
- The jumbo frame size must be set to the same size for each component in the path

Performance

- We are currently testing jumbo and non-jumbo frames for performance

Networking Technologies

- **Adapter Requirements for iSCSI in the Enterprise (1Gb & 10Gb)**
  - TCP/UDP checksum offloads
  - Large Send Offload (LSO)
  - Large Receive Offload
  - Receive Side Scaling (RSS)
  - TCP Chimney Offload
  - Jumbo Frames

- **iSCSI HBA (full iSCSI/TCP/IP offload)**
  - Available for 1Gb and 10Gb

- **10GbE and Data Center Bridging (DCB)**
  - DCB provides extensions to Ethernet, but not required for iSCSI
  - We are testing the effects of DCB on iSCSI

Storage Targets

- ❖ **Storage systems are available that support**
  - ◆ 1Gb iSCSI
  - ◆ 10Gb iSCSI

- ❖ **iSCSI storage systems can scale**
  - ◆ Hundreds of TB or more per system
  - ◆ Many host ports

- ❖ **Some storage systems support multiple interfaces**
  - ◆ Fibre Channel, FCoE, iSCSI, Ethernet (NAS)

- ❖ **With 10GbE, storage performance depends less on the interface and more on the back-end disk drive (or SSD) configuration**
  - ◆ SSDs and 10GbE work well together

Server Virtualization

◆ iSCSI is supported for server virtualization

- ◆ VMware vSphere
- ◆ Microsoft Hyper-V

◆ Guests virtual machines (VMs) can access iSCSI targets

Best Practices

◆ **Separate SAN and LAN traffic**

  ◆ iSCSI SAN traffic should be separate from regular LAN traffic. This minimizes congestion and latency and improves security.

    › Either use VLANs or physically separate networks
    › With 10GbE, this separation may not be needed, depending on workloads

◆ **Use Gigabit or faster Ethernet technology for iSCSI**

  ◆ Use appropriate cabling for the rated speed and distance

◆ **Use server-class network adapters that are designed for enterprise workloads**

  ◆ In previous testing, we found that cheap, desktop NICs hung the server when running heavy iSCSI workloads

◆ **For enterprise performance and reliability, configure redundant components (adapters, switches, etc.)**

Best Practices

- **For 1Gb iSCSI**
  - Deploy enough NICs/HBAs to get the bandwidth needed for the applications

- **For 1Gb or 10Gb iSCSI**
  - Follow application best practices for storage design for RAID, disk striping, format allocation, etc.

Summary

- ◆ iSCSI uses Ethernet which is everywhere

- ◆ iSCSI is relatively easy to deploy

- ◆ iSCSI can be deployed in low-cost environments

- ◆ iSCSI can scale to meet enterprise needs

- ◆ We have published a vendor-neutral storage interface comparison reference page at
  http://www.demartek.com/Demartek_Interface_Comparison.html

- ◆ We are publishing a new iSCSI Deployment Guide in May 2011
  http://www.demartek.com/Demartek_iSCSI_Deployment_Guide.html

Questions?

Ethernet Storage Forum Members

The SNIA Ethernet Storage Forum (ESF) focuses on educating end-users about Ethernet-connected storage networking technologies.