SNIA

Education

iSCSI : A loss-less Ethernet fabric with DCB

Jason Blosil, NetApp
Gary Gumanow, Dell

SNIA Legal Notice

- 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.

Abstract

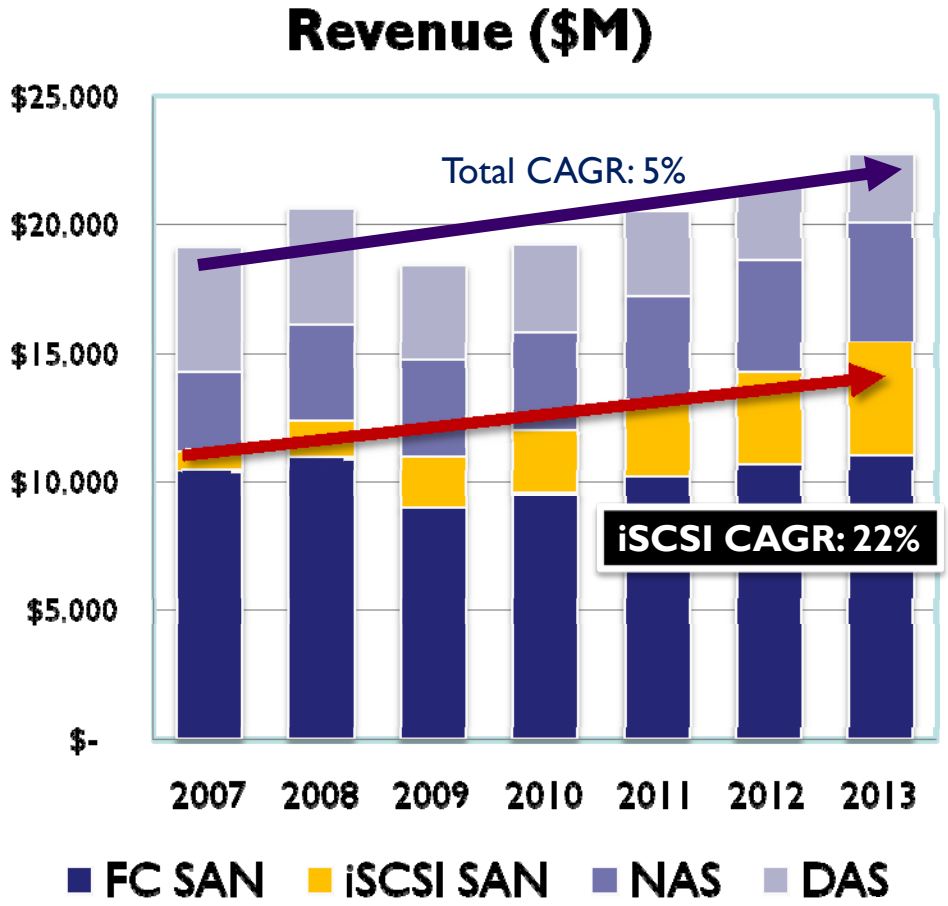
- A look at different storage technologies and a brief look at each; benefits/tradeoffs, making the case for a SAN. Then looking at making your SAN a lossless, high-performance, predictable, resource for your business. Well look at different storage protocols and how they compare to the OSI model, and the new DCB protocols. And lastly, I'll present some findings of using these technologies.

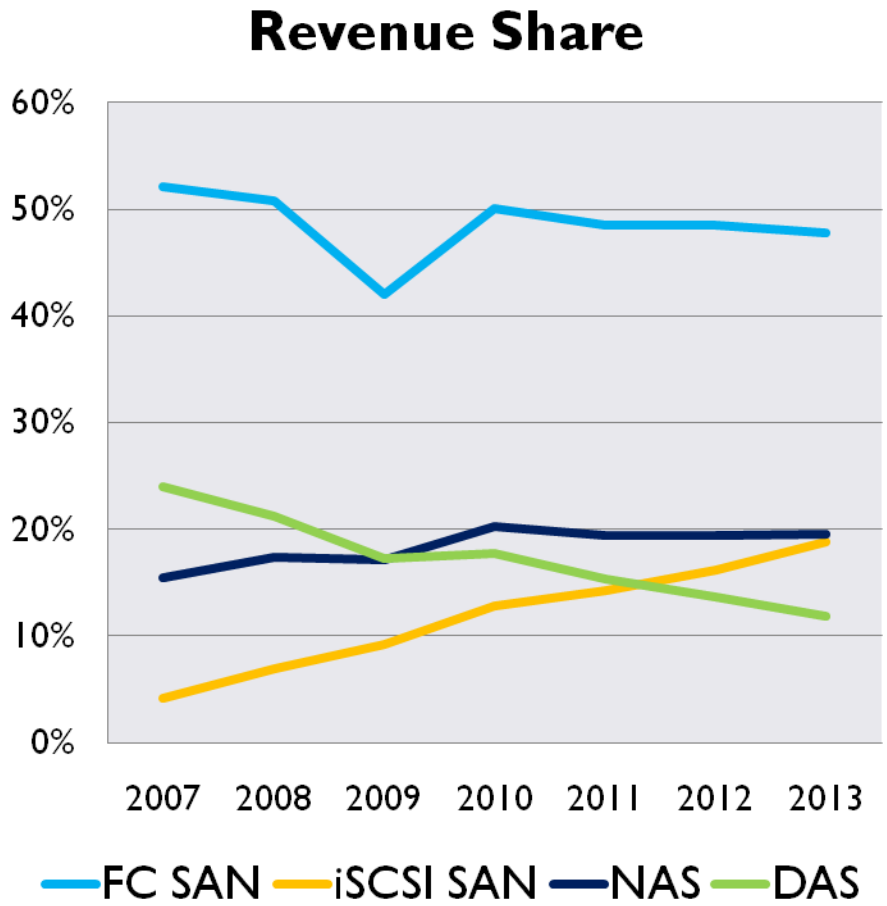
- Learning Objectives
  - A detailed look and comparison between three storage models; iSCSI, FC, and FCoE. Comparing all three to the OSI model. A discussion of the requirements for each layer of the networking stack.
  - A look at the benefits that DCB can provide iSCSI, as a lossless Ethernet fabric. A review of performance and contention issues.

What we'll talk about today…

- How is iSCSI doing today
- What is Data Center Bridging
- What's the roadmap for Ethernet
- A protocol comparison
- iSCSI over DCB performance
- Does DCB really matter for iSCSI

IDC Forecast (Dec 2009)



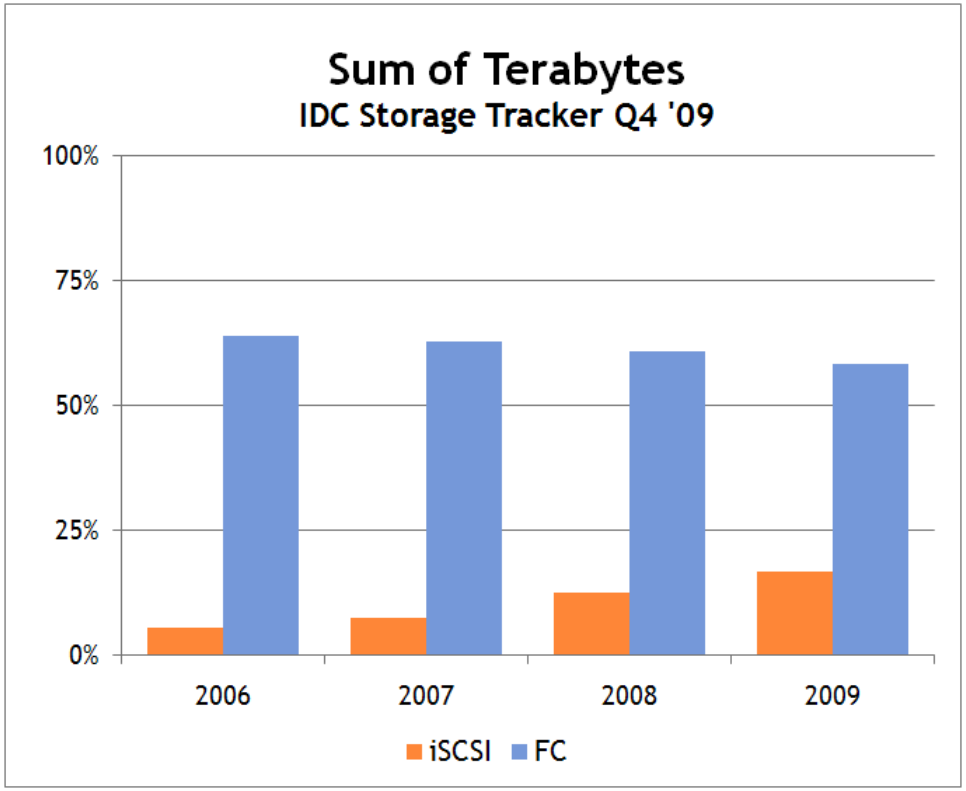


iSCSI today

- Applications of iSCSI
  - Disaster Recovery / Backup
  - Virtualization
  - Thin provisioning
  - Snapshots
  - VM integration
  - Disk de-dupe
  - Architecture choices
- SW Initiators available from all major OSes
  - MPIO support
  - Low CPU utilization without HW offload
- Architecture choice
  - Scale-out & Scale-up
- 10GbE available today

Sum of Terabytes	CAGR 2006-2009
iSCSI	43.6%
FC	-3.0%
Storage Growth	29.0%

Source: IDC Storage Tracker Q4 2009



Trends in the datacenter

- Networking
  - Data Center Bridging (DCB)
  - 10 Gigabit Ethernet
  - Top of Rack switching
- Storage
  - Increased focus on Ethernet-based storage
- Server
  - Interconnect technologies to support > 50Gbps
  - Increased density with virtualization

The need for 10GbE

Education SNIA

Virtualization

- VMs requires more network capacity
- 80% CPU increases demand on bandwidth
- 10GbE & iSCSI offers the bandwidth needed today
- 10GbE w/ DCB enables enhanced fabric virtualization by offering enhanced manageability and QoS

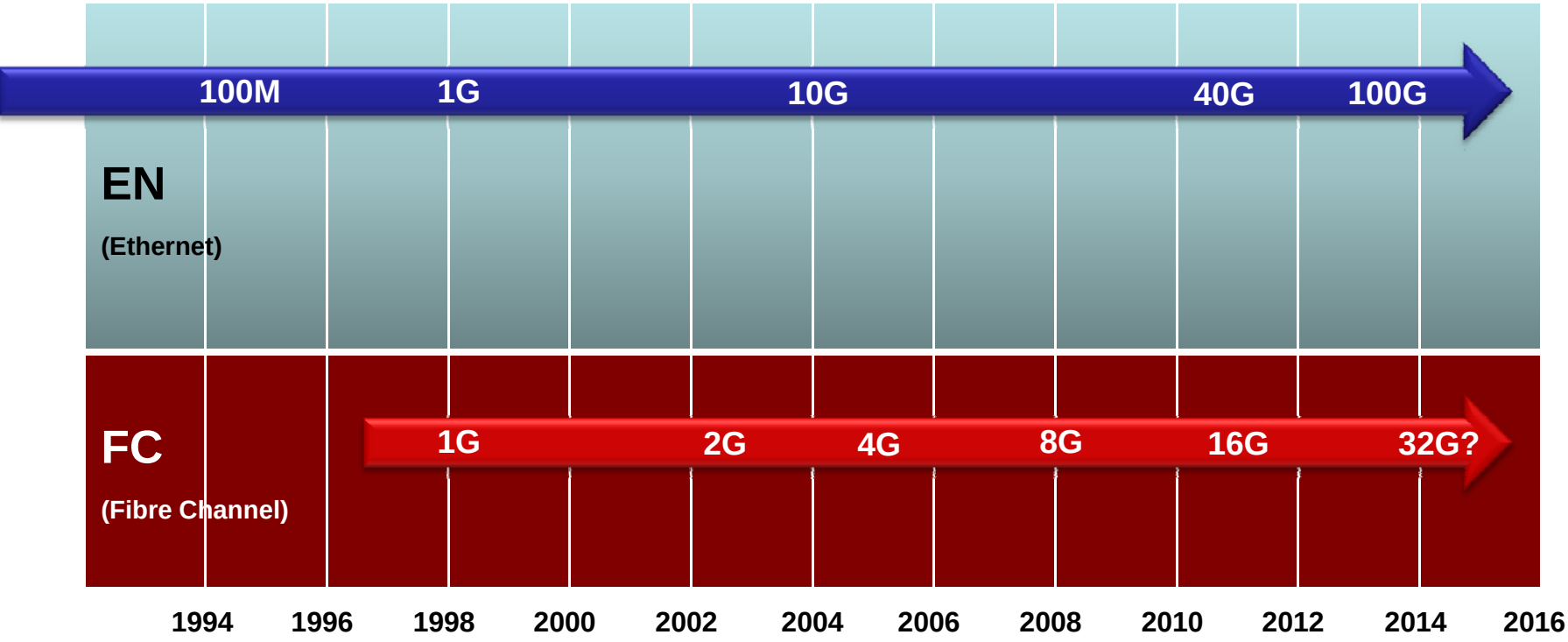
Consolidation

- Enabling server virtualization allows for increasing levels of server consolidation
- Unifying LAN & SAN consolidates network fabric
- Enables reductions in HW, power and OpEX

Application Performance

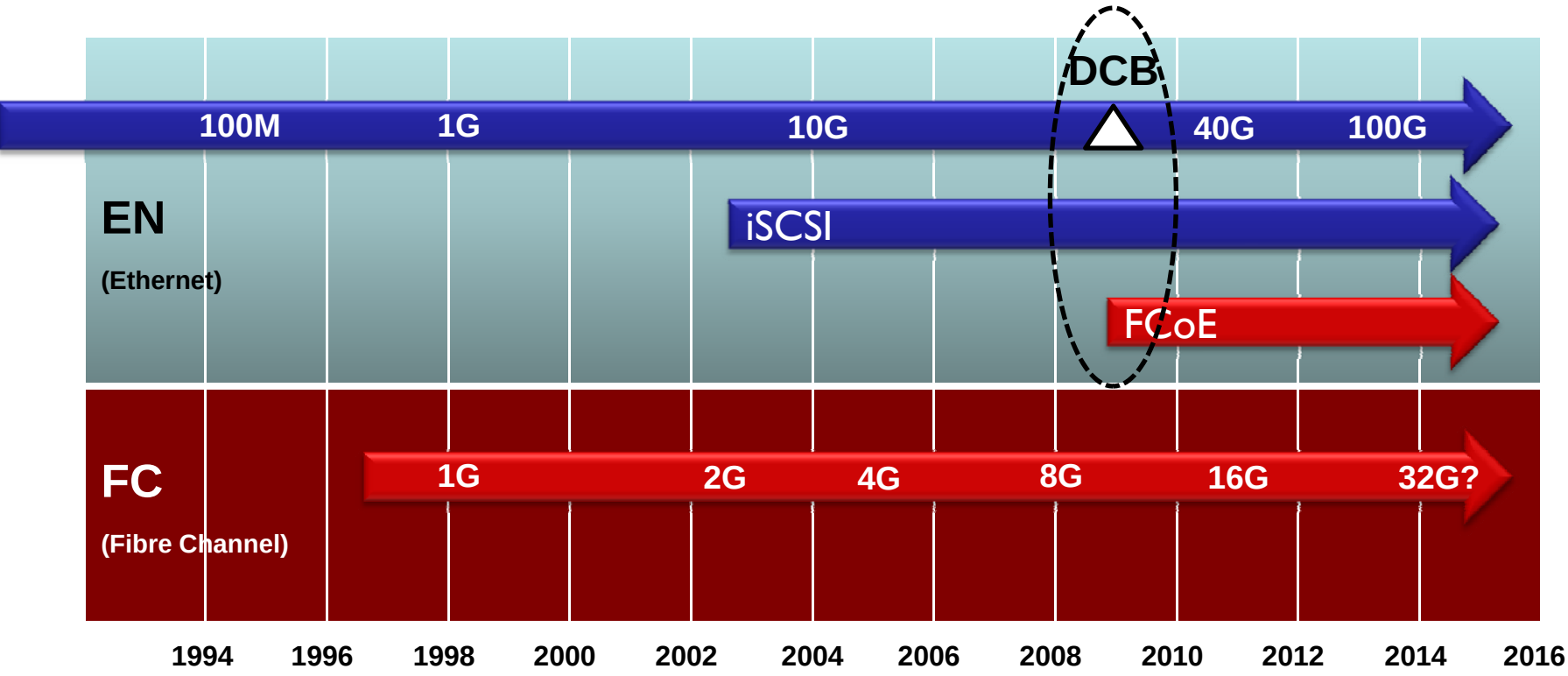
- Larger database, streaming media and other high I/O applications demand data center performance
- Enables thin client and desktop virtualization
- 10X the bandwidth reduces backup windows

Fabric Roadmaps



- Infrastructure migrates to Ethernet irrespective of storage protocols
- Convergence and virtualization are driving enhancements to Ethernet

Fabric Roadmaps

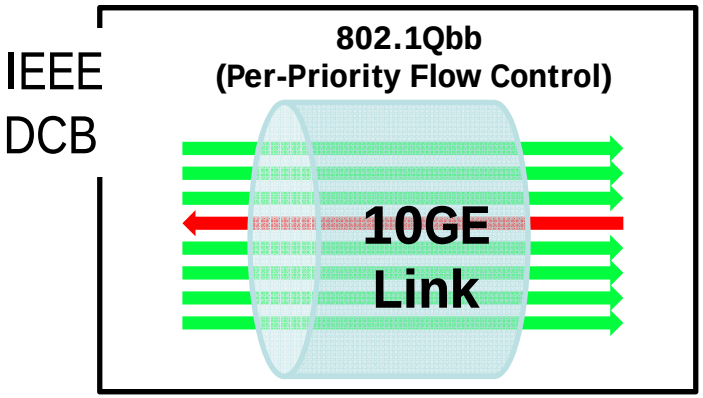


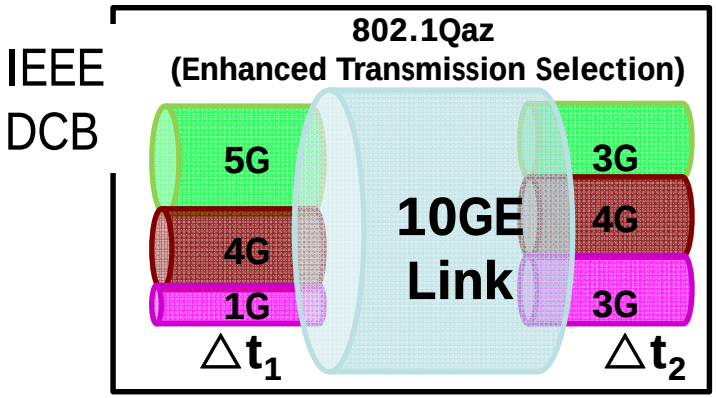
- Infrastructure migrates to Ethernet irrespective of storage protocols
- Convergence and virtualization are driving enhancements to Ethernet

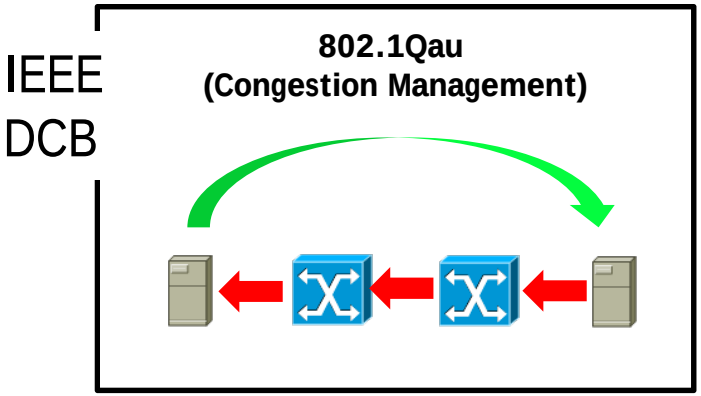
Data Center Bridging

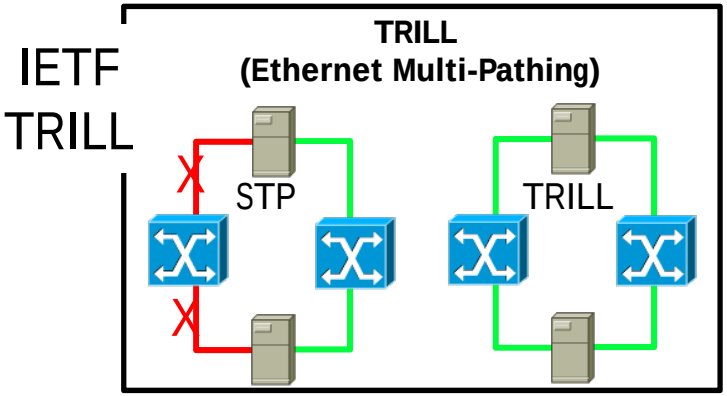
- **Enhancements to Ethernet**
  - Provided enhanced QoS support to Ethernet
- **What constitutes DCB Standards?**
  - **PFC** aka Priority based Flow Control (802.1Qbb)
  - **ETS** aka Enhanced Transmission Selection (802.1Qaz)
  - **CN** aka Congestion Notification (802.1Qau)
  - **DCBX** aka Data Center Bridging capability eXchange
    - LLDP vs. DCBX
      - LLDP: Primarily a link level information exchange protocol
      - DCBX: Neighbors can configure parameters based on info exchange and state machine

DCB Components









DCB Components

HALT an individual stream, but NOT all of them!

IEEE DCB

802.1Qbb
(Per-Priority Flow Control)

10GE Link

Allocate bandwidth based upon predetermined classes of traffic

IEEE DCB

802.1Qaz
(Enhanced Transmission Selection)

5G
4G
1G
10GE Link
3G
4G
3G
$\triangle t_1$
$\triangle t_2$

IEEE DCB

802.1Qau
(Congestion Management)

End-to-End Communication between end-points. Tells the end-point to BACK OFF!

IETF TRILL

TRILL
(Ethernet Multi-Pathing)

STP
TRILL

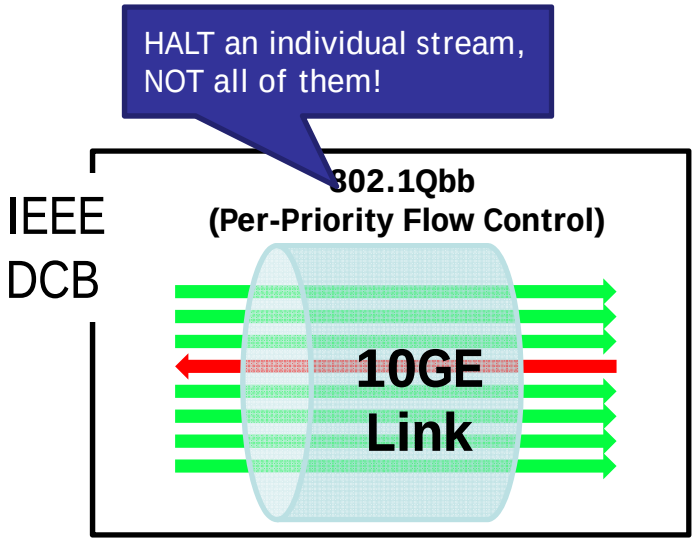
Multi-pathing for Ethernet

Replaces Spanning Tree

Is DCB relevant to iSCSI?

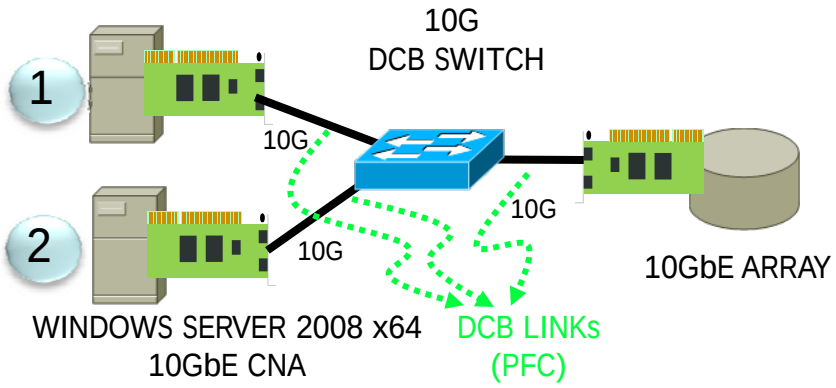
Let's only focus on 802.1Qbb (PFC) for now….

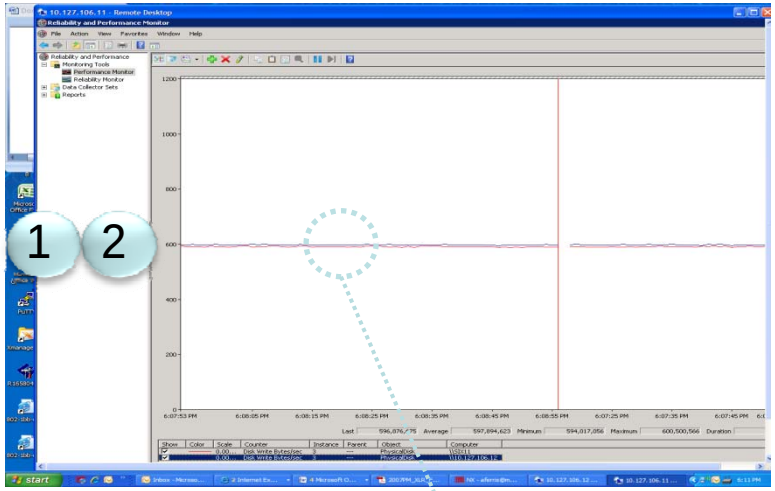
- PFC represents an enhancement to Ethernet
- PFC is often viewed as an FCoE-ONLY technology
  - FCoE requires PFC for packet delivery
  - iSCSI does not require PFC but can utilize it
- iSCSI benefits from PFC support



PRELIMINARY iSCSI DCB RESULTS

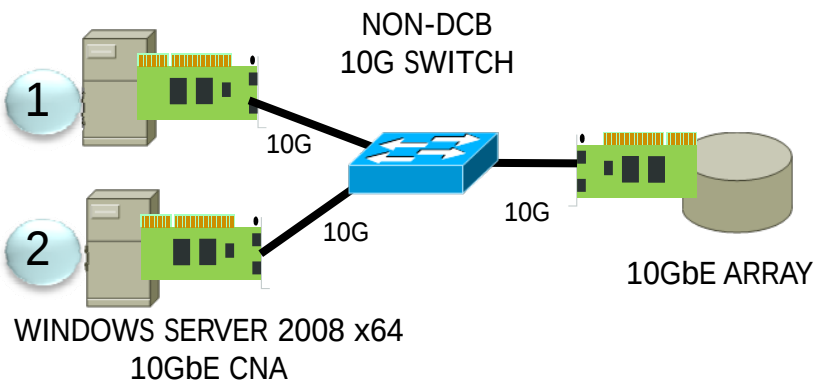
iSCSI WITH DCB

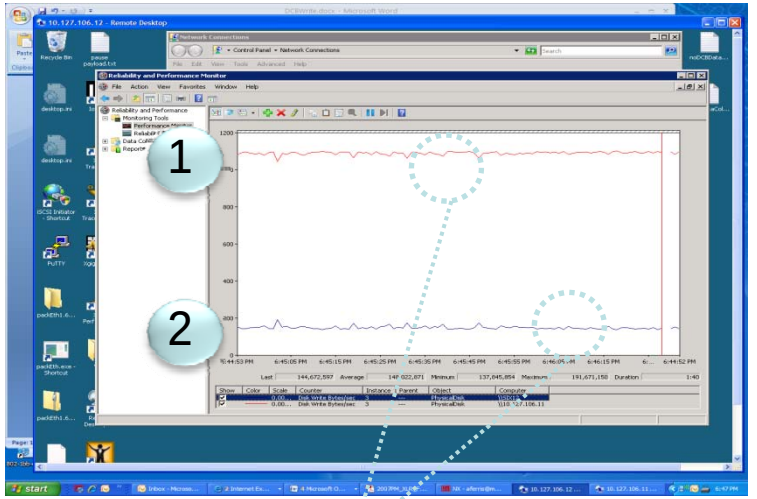




Balanced iSCSI throughput (600MB/s, 600MB/s)
Steady packet streams (no TCP burstiness)

iSCSI WITHOUT DCB



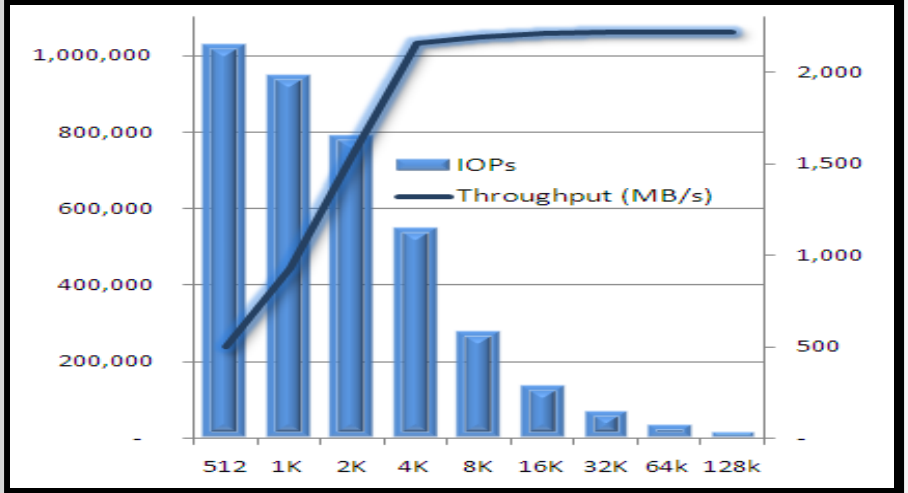


Unbalanced iSCSI throughput (1100MB/s, 100MB/s)
Typical TCP burstiness

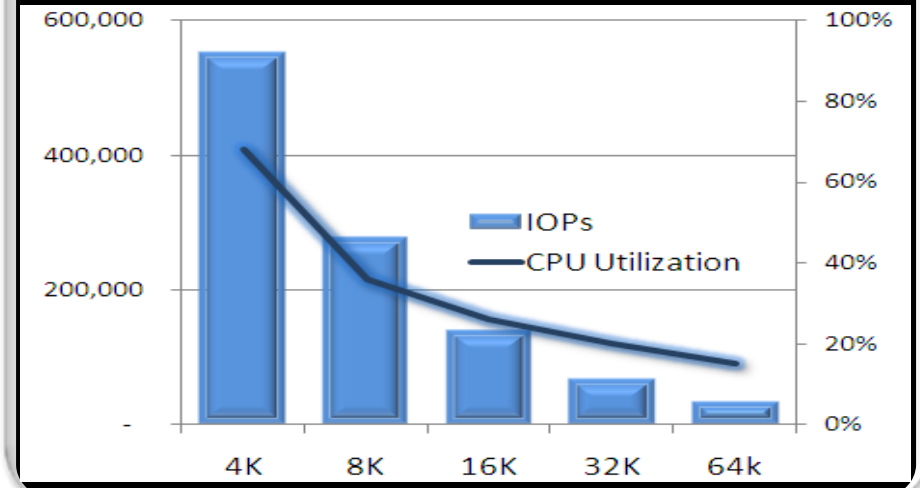
iSCSI performance at 10GbE

Education SNIA

Read/Write IOPs and Throughput Test



Read/Write IOPs and CPU Test



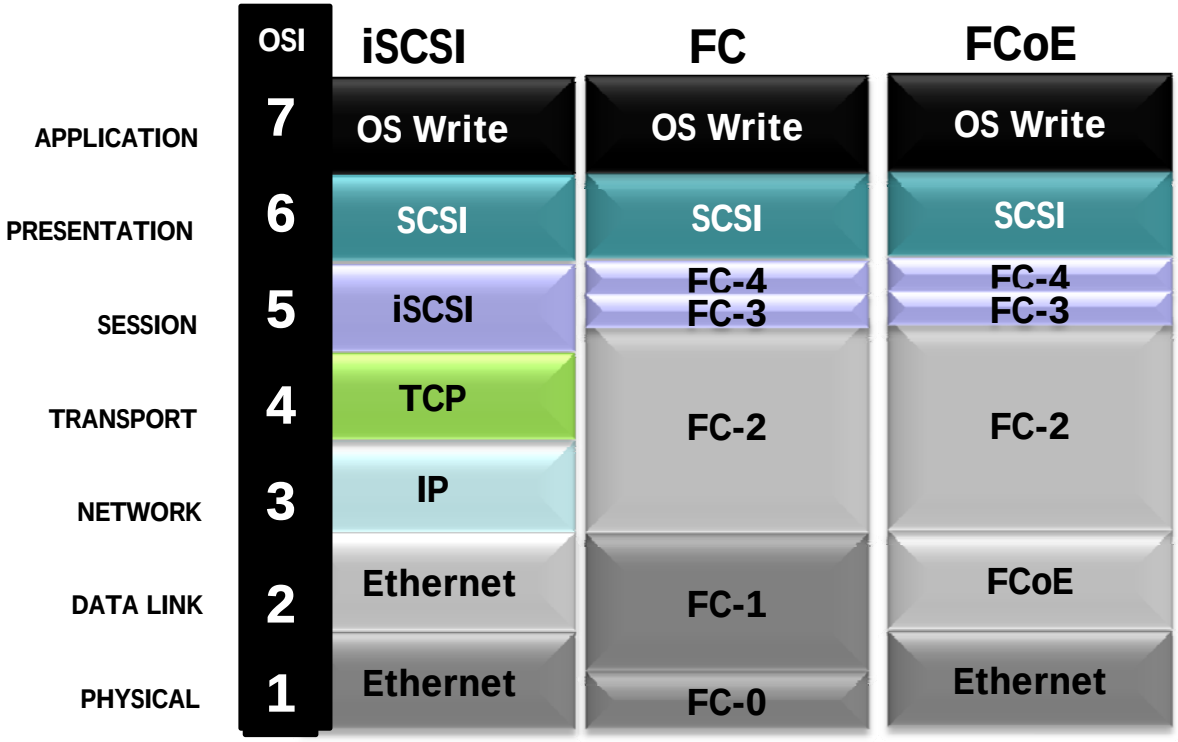
1,030,000 IOPs

- Single Port
- 10GbE line rate
- 10k IOPs per CPU point
- Performance for real world apps
- Future ready: Performance Scales

552k IOPs at 4k represents

- 3,100 Hard Disk Drives
- 400x a demanding database workload
- 1.7m Exchange mailboxes
- 9x transactions of large eTailers
- Jumbo frames: >30% CPU decrease is common for larger IO size (jumbo frames not used here)

Storage protocols



Fibre Channel is a layered protocol. It consists of 5 layers, namely:

- **FC0** The physical layer, which includes cables, fiber optics, connectors, pinouts etc.
- **FC1** The data link layer, which implements the 8b/10b encoding and decoding of signals.
- **FC2** The network layer, defined by the FC-PI-2 standard, consists of the core of Fibre Channel, and defines the main protocols.
- **FC3** The common services layer, a thin layer that could eventually implement functions like encryption or RAID.
- **FC4** The Protocol Mapping layer. Layer in which other protocols, such as SCSI, are encapsulated into an information unit for delivery to FC2.

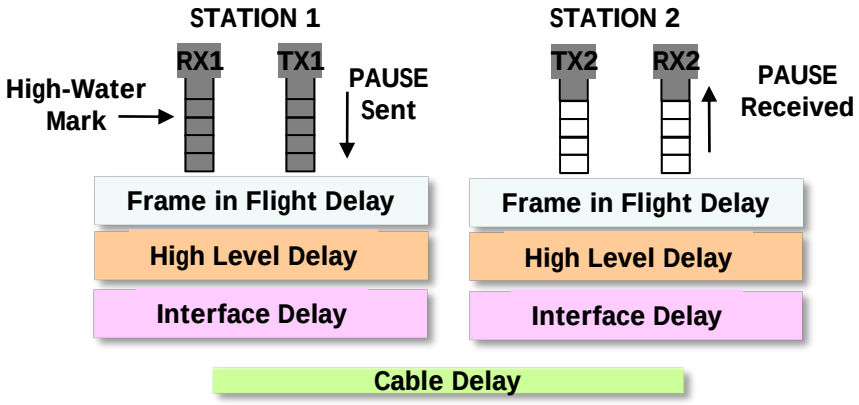
Comparison of iSCSI, FCoE, FC

Feature	iSCSI	FCoE	FCP
Uses Legacy Ethernet	Yes	No	N/A
Uses DCB Ethernet	Yes (same benefits)	Yes	N/A
REQUIRES DCB Ethernet	No	Yes (won't work without it)	N/A
Routable	Yes (IP routing)	Not Routable	Yes (FCIP, iFCP, FC-IFR)
Host Connectivity	Servers, Clients	Servers Only (no client applications)	Servers Only (no client applications)
Guaranteed Delivery	Yes (TCP)	Possible (FC Class 1/2 service)	Possible (FC Class 1/2 service)
Fabric Layer (2) Loss	rate based flows	rate based flows	credit based flows
Maturity	2003	2009	1997
Fabric Management	Ethernet Tools	FC Tools	FC Tools

FLOW Control Types

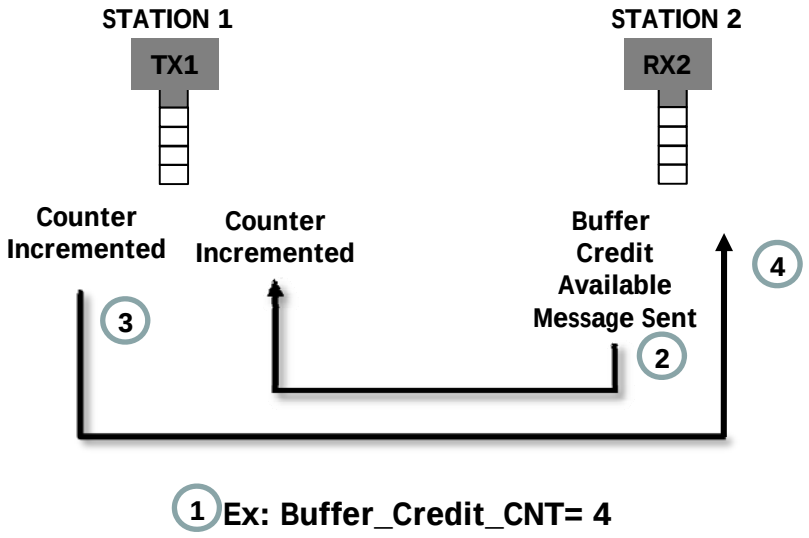
Reactive-Time/Link Dependent

PAUSE

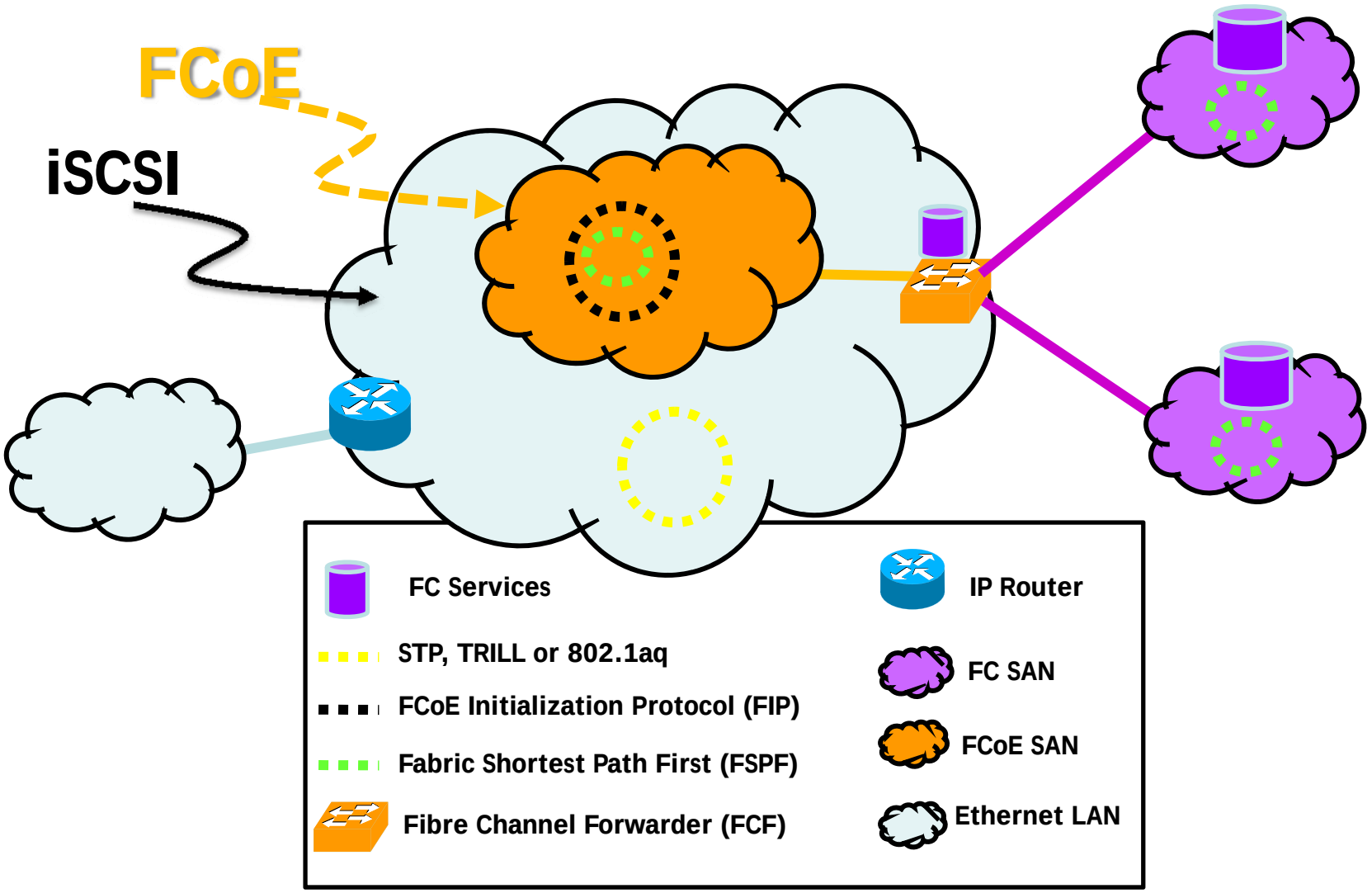


Proactive-Time/Link Independent

CREDITS



Storage Network Components



Is DCB relevant to iSCSI?

- **YES**

Improves application responsiveness

Provides QoS improvements in mixed traffic environments

Increases robust of Ethernet as a storage fabric

Improves performance of existing IP Protocols including iSCSI

Questions?

Q&A / Feedback

- 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

Jason Blosil
Gary Gumanow
Jordan Plawner