# IT-Symposium 2004

## Fibre Channel Roadmap

### Vortrags-Nr. 2C01

### 21. April 2004

**Udo Kuhr**
**Solution Consultant**

---

# CNT - Ein starkes, globales Unternehmen…

*… enstanden aus zwei führenden Storage Networking Spezialisten!*

- mehr als <u>1.200 Mitarbeiter</u>

- mehr als <u>3.000 Kunden</u> weltweit

- Kunden in <u>30+ Ländern</u> - 80% der "Fortune 500" !

- mehr als $ 360 Mio. Umsatz in 2003

- <u>weltweites</u> Vertriebs- und Support-Netzwerk

- <u>24x7</u> Call Center mit umfassenden Remote Monitoring-
  und Management-Möglichkeiten

- Global Services & Systems Integration

- Telco-Services (MAN/WAN-Bandbreite)

2
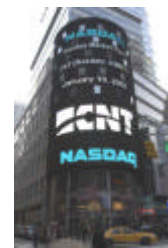
# CNT's Geschäftsbereiche

## Bereitstellung

*von komplexen, hochverfügbaren Speichernetzwerken für Business Continuity und Disaster Recovery Lösungen inklusive deren Design, Implementierung und Überwachung*

**Switching**
**Eigene Switching-Produkte für Open Systems und Mainframe- (Escon, FC, Ficon) Anbindung inklusive Managementfunktionen**

**Storage Networking für LAN/MAN/WAN**
**Hochverfügbare Speicherverbindungssysteme für die Implementierung herstellerunabhängiger Business Continuity und Disaster Recovery Lösungen**

**Optische Netzwerke**
**Skalierbare, optische Wellenlängen-Multiplexer (TDM/CWDM/DWDM) für die effektive Ausnutzung von Glasfasernetzwerken**

**Professional Services**
**Weltweite Consulting und Professional Services Dienstleitungen für komplette Speicher "end-to-end" - Lösungen**

---

# LAN / SAN / MAN / WAN
*A Unique Value Proposition*



**... langjährige Partnerschaft mit Hewlett Packard !**

## Topics

- **Overview of 4Gb/s Technology**
  - Where did it come from?
  - Who's driving it?
  - Relationship with 1Gb/s and 2Gb/s

- **Overview of 10Gb/s Technology**
  - Relationship with 1Gb/s and 2Gb/s and 4Gb/s

- **Distance Considerations at Higher Speeds**

- **Industry Support**

---

### Some Additional Material

✓ "The Resilent Enterprise – Veritas Press
✓ "Leveraging MAN and WAN Connectivity" – SNWOnline – August 2002
✓ zJournal – "zStorage" – February 2003, "Virtualization" – June 2003
✓ Industry Associations (SNIA, FCIA, IETF, T11, SCTA) and their websites
✓ SNIA Web Site and Education Committee Tutorials (www.snia.org/education)
✓ CMG Proceedings and handouts on CD's
✓ CMG MeasureIT – "Networking With Your Storage", May 2003
✓ Evaluator Group website (www.evaluatorgroup.com)
✓ "Resilient Storage Networks – Designing Flexible Scalable Data Infrastructures"



Source:Greg Schulz

## Storage Networking Update
### *The Need For Speed – 1Gb, 2Gb, 4Gb, 10Gb, and Beyond?*

**Fibre Channel Interface Speeds:**
- 4Gb for backwards compatibility with 1Gb and 2Gb devices
- 10Gb is not backward compatible with 1Gb, 2Gb ans 4Gb devices
- 10Gb for connection between switches to replace, or enhance ISL Trunks
- Servers need to support PCI-X and higher to utilize these faster speeds
- There is some *noise* about possible 8Gb for in the box, internal applications
- There is some *noise* about 20Gb, 40Gb, and beyond interfaces

**Ethernet Interfaces Speeds**
- 10/100/1000 and now 10000 (10Gb)
- 10Gb is still very expensive (e.g. $5,000+Optics per port or higher)
- Not all switches achieve 10Gb however newer generations one should support this
- 10Gb being deployed in the core backbone of networks to link switches together
- Servers need to support PCI-X and higher to utilize these faster speeds

**WAN/Metropolitan – SDH/SONET**
- Higher densities DWDM to improve distance and bandwidth utilization
- SDH/SONET OC3, OC12, OC48, OC96, OC192 and beyond

---

## 4Gb/s Technical Facts

- **FCIA Intended 4Gb/s to be Used Within Cabinets Only**
  - e.g., between internal disks in a RAID subsystem
  - 10 Gb for external interfaces to match Ethernet

- **Pushed into the "Open"**
  - QLogic (at HDD and storage subsystem)
  - Vixel for embeded switches (inside the storage cabinet)

- **FCIA Conducted Anonymous Vote in Late May**
  - Generally, switch vendors were opposed to 4Gb/s external to cabinet
  - But…majority (60%) of votes were in favor of 4G FC outside cabinet
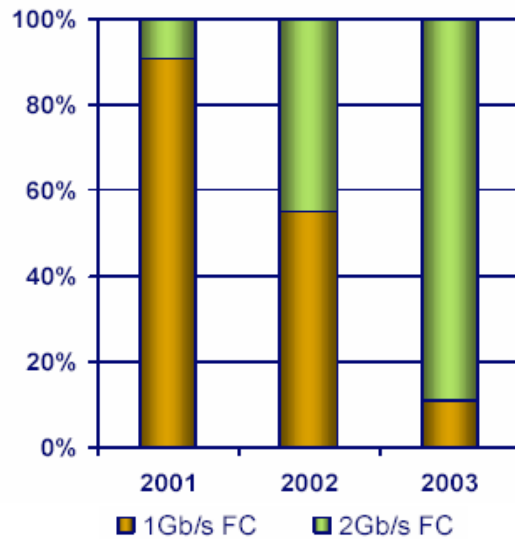
- **Cost Much Less than 10 Gb/s**
  - Slightly more expensive than 2Gb
  - 10Gb cost 3x to 5x more versus 4Gb
  - Plus, preserves customers' investment in 1Gb/2Gb infrastructure

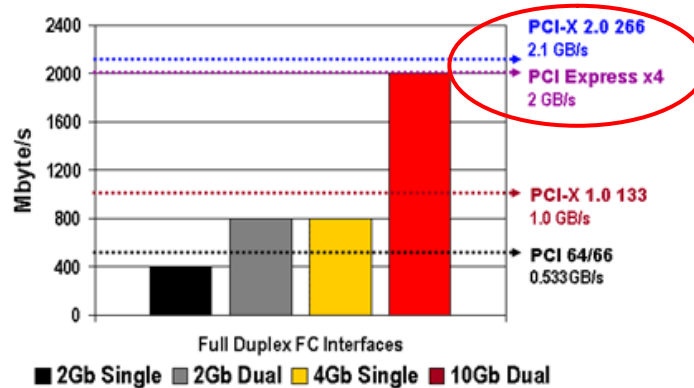## 1 Gb/s To 2 Gb/s Adoption Rate



Source: IDC, 2003

## Port Speeds versus Bus Speeds



Source: QLogic Corp.

## What About Storage Requirements?

- **2Gb/s Drives Available for Approximately a Year…**

  - Are physical disk rotation speeds fast enough yet to sustain 4 Gb/s??

- **Initial 4Gb/s FC Disk Drives Expected to Arrive 2H2004**

- **However, Disk drives will be the Key to reducing the cost of 4 Gb interfaces**

## 4 Gb/s Pro's and Con's

- **Pro's**

  - Backward switch compatinbilitity with existing 1Gb/s and 2Gb/s speed, cabling as well
    - Seamless migration to 4 Gb/s
  - Can be supported by today's PCI-X bandwidth
  - Only incrementally more expensive than 2Gb/s

- **Con's**

  - Additional speed to support for switch vendors
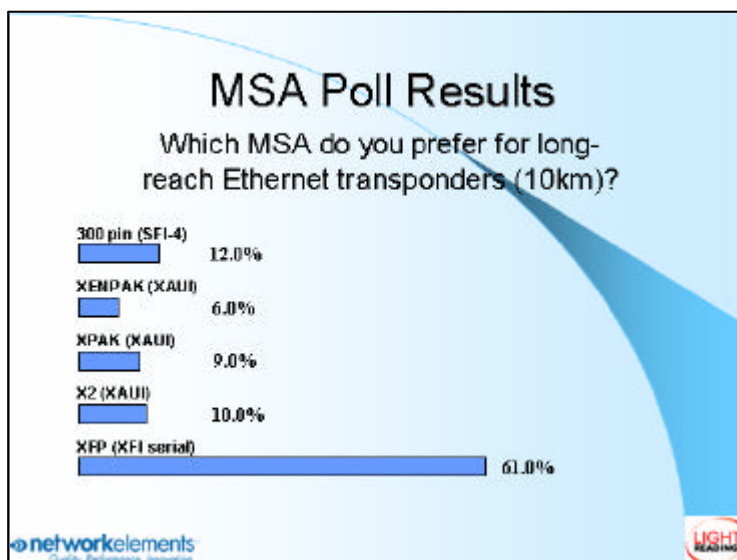
## 10Gb/s Technical Facts

- **It's Not Really 10 Gb/s !!**
  - Derived from 10 Gig-Ethernet spec
  - Four separate serial "flows" or lanes
    - vs. a single flow for 1 Gb/s, 2 Gb/s, 4 Gb/s
    - Therefore, NOT compatible with legacy components
  - Each lane runs at 3.1875 Gb/s ~or~ 12.75 Gb/s Total
    - Link actually runs at just over 10.5 Gb/s
    - But, lanes are not just muxed together in **true** 10 Gb/s
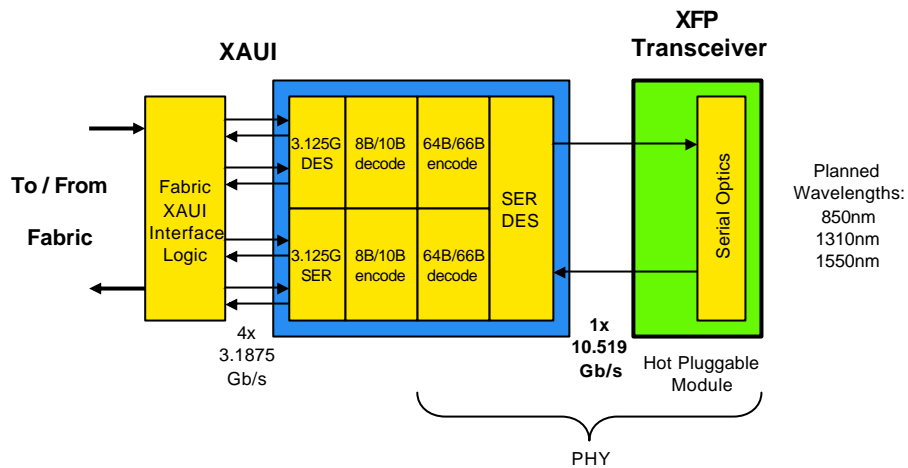- **Multiple Physical Interface Options**
  - Some support both 10Gb Gig-E and 10GFC
    - FC interface is approx. 2% faster
  - Xenpak, X2, XPAK, XFP…
    - XFP will be dominant in Fibre Channel

## XFP Will Be The Transceiver of Choice

# 10 Gb/s Serial Port Block Diagram with XFP



**XAUI**

**XFP Transceiver**

To / From

Fabric

Fabric XAUI Interface Logic

3.125G DES | 8B/10B decode | 64B/66B encode

SER DES

3.125G SER | 8B/10B encode | 64B/66B decode

Serial Optics

Planned Wavelengths:
850nm
1310nm
1550nm

4x 3.1875 Gb/s

**1x 10.519 Gb/s**

Hot Pluggable Module

PHY

ⓘ Requires at least 500 Buffer Credits for 100km of fiber.

---

# 10 Gb/s XFP Transceiver

- **Major Improvements Over Earlier 10Gb/s Designs**
- **Compact Size**
  - ~ 3.1" L x 0.9" W x 0.47" H
- **Hot Pluggable**
- **LC Type Connector**
- **Short & Longwave**
  - 850 nm
  - 1310 nm
- **Can Also Be Used for Ethernet, SONET/SDH**
  - Other uses will drive down cost

# 10 Gb/s Pro's and Con's

- **Pro's**
  - Great for ISL trunking
  - Could propel Storage-over-IP into mainstream (e.g., iSCSI)
  - Same/similar components as 10 Gb/s Ethernet
  - Maintains speed parity with Ethernet/iSCSI

- **Con's**
  - Not compatible with existing interfaces
    - 1Gb/s, 2Gb/s, 4Gb/s
  - Very long time before Devices may be able to use all 10Gb perf.
  - 10 Gb is very disruptive
  - Technology is expensive

---

# Projected Buffer Credit Quantities

- **How Many Credits for 4 Gb/s and 10 Gb/s to Fill the "Pipe" ??**

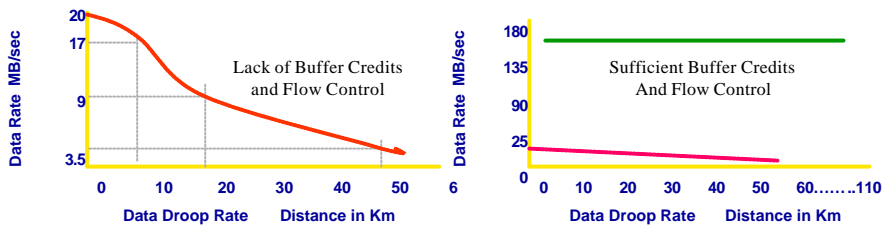| At 4.25 Gb/sec: | At 10.51875 Gb/sec: |
|---|---|
| Frame = 2,112 Bytes<br>At 4.25 Gbp/s,<br>One Frame = 4.97 µsec | Frame = 2,112 Bytes<br>At 10.52 Gbp/s,<br>One Frame = 2.01 µsec |
| Velocity in fiber approx 5 ns / meter | Velocity in fiber approx 5 ns / meter |
| ∴ **A 4 Gb/s frame fills ~1000 meters of fiber** | ∴ **A 10 Gb/s frame fills ~400 meters of fiber** |
| Round trip = 200 km so could be up to [200km ÷ 1km] or **200 full frames** | Round trip = 200 km so could be up to [200km ÷ 400m] or **500 full frames** |
| ∴ Need approximately 200 buffer credits | ∴ Need approximately 500 buffer credits |

# Storage Networking Performance
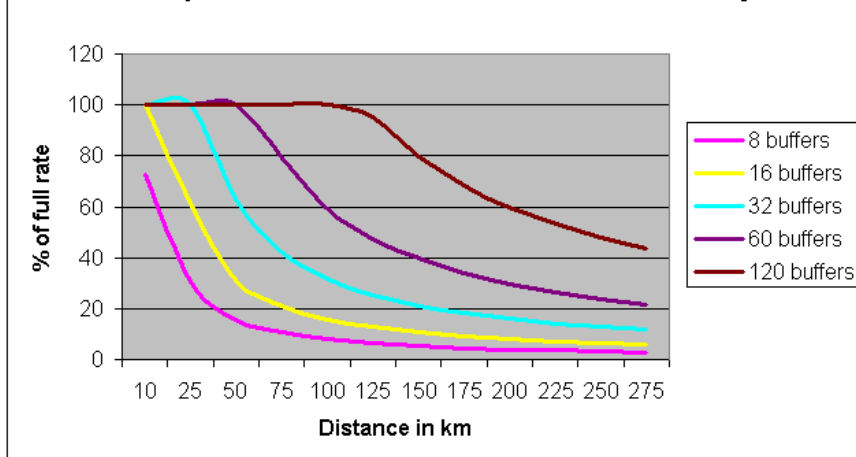## *Protocol Droop Over Distance and Flow Control*

**Droop**

- Phenomenon of performance degradation over distance
- Performance drops before theoretical limit
- Not to be confused with speed of light (5microsec/km latency)
- Function of Network Protocol vs. Physical medium
- Synchronous access can be more impacted then Async.
- Results from lack of Buffer Credits with flow control networks
  Commonly found with Storage and Deterministic Interfaces

**Left chart:** Data Rate MB/sec (20, 17, 9, 3.5) vs Distance in Km (0, 10, 20, 30, 40, 50, 6)
Lack of Buffer Credits and Flow Control
Data Droop Rate    Distance in Km

**Right chart:** Data Rate MB/sec (180, 135, 90, 25, 0) vs Distance in Km (0, 10, 20, 30, 40, 50, 60.......110)
Sufficient Buffer Credits And Flow Control
Data Droop Rate    Distance in Km

http://www.decus.de    Copyright © C N T    19

---

# Storage Networking Performance
## *Protocol Droop Over Distance and Flow Control*

**Droop in full 2Gb FC rate due to buffer difficency**

% of full rate (0, 20, 40, 60, 80, 100, 120) vs Distance in km (10, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275)

- 8 buffers
- 16 buffers
- 32 buffers
- 60 buffers
- 120 buffers

http://www.decus.de    Copyright © C N T    20
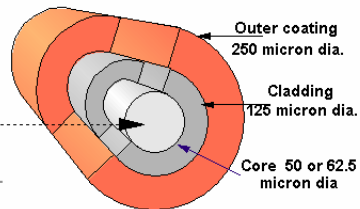
10

## Fiber Optic Cable Types



**Multimode (MM) Fiber**
"Multiple paths" for light to travel

DATA
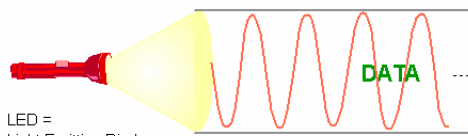
LED =
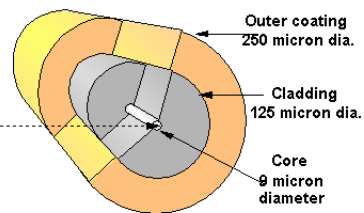Light Emitting Diode
SX Laser = Short Wavelength Laser

Outer coating
250 micron dia.

Cladding
125 micron dia.

Core 50 or 62.5
micron dia

**Single Mode (SM) Fiber**
"Single path" for light to travel

DATA

LX Laser =
Long Wavelength Laser

Outer coating
250 micron dia.

Cladding
125 micron dia.

Core
9 micron
diameter

http://www.decus.de

Copyright © C N T    21

---

## Relationship Between Speed & Distance

- **As the Link Speed (Rate) Increases, the Supported Distance and Link Budget are Reduced**
  - Assuming Multimode fiber, equivalent transceiver power, etc.

| Fiber Type | 1 Gb | 2 Gb | 4 Gb |
|---|---|---|---|
| 50 micron multimode | 500m | 300m | 150m |
| 62.5 micron multimode | 250m | 120m | 55m |

  - Speed has NO appreciable effect on distance with Single Mode ☺

http://www.decus.de

Copyright © C N T    22

# Storage Networking Performance
### *Storage Networking Enables Distance – Optical Distance*
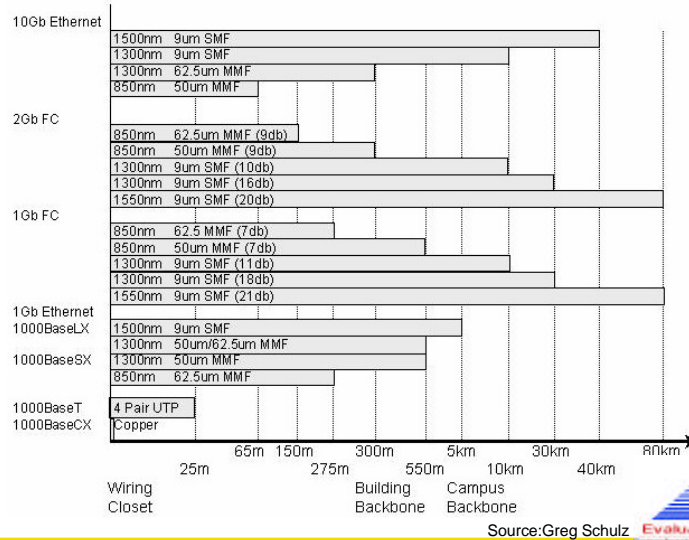


Source:Greg Schulz

---

# Storage Networking Performance

### *Storage Networking Interfaces*

- Host Bus Adapters (PCI, PCI-X, Sbus, etc)
  - PCI constrained at about 1.5-1.7Gb/sec
  - PCI-X and subsequent generations will help
  - How much of your 1Gb/2Gb interface are you using today?

- Fibre Channel (FCP, FICON, FCIP, etc)
  - 1Gb, 2Gb, and 4Gb can co-exists and are back-ward compatible
  - 10Gb will not interoperate with existing interfaces
  - 10Gb will initially be deployed for ISLs to eliminate trunking

- Ethernet (IP, NAS, iSCSI, FCIP, etc)
  - 1Gb and 10Gb today

- SONET/SDH
  - OC-1 to OC-192 (OC-768 on drawing board)

Source:Greg Schulz

# Storage Networking Performance
## *Horizontal Scaling – Scale Beyond Single Device*

Utilizing networking or fabrics of two or more switches/directors in various topologies across one or more locations to provide adequate number of ports, performance, redundancy, and meet application service requirements. First generation SANs have been based on small devices with limited vertical salacity (e.g. 8 and 16 port switches, 16 to 256 port directors). Various topologies can be implemented to meet different needs including core-fabric switch, cascade, mesh, tiered, ring, core-edge, and so on. FICON Cascade is an example of vertical scaling for DR and Resiliency.



**Dual Path Redundant SAN**

**Cascade SAN Single Path**

**Redundant Cascade SAN Dual Path**

**Redundant Tiered Dual Path SAN**

**3,000+ Port, High Performance Redundant SAN**

---
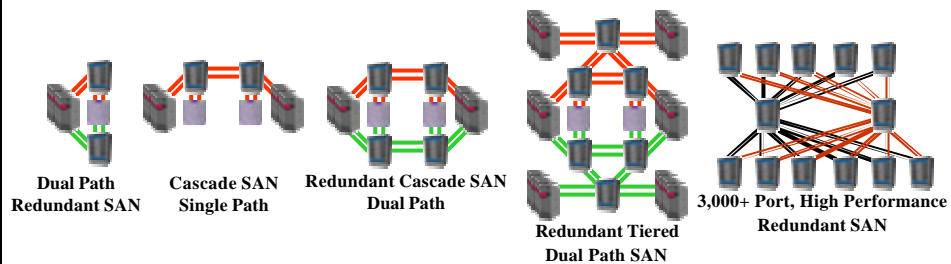
# Storage Networking Performance
## *Locality – Eliminating Congestion and Bottlenecks*

Locality describes where ports are located in a SAN with the higher locality, the better performance will be with fewer or no hops to incur increased latency, simplified security and management. In the example below on the left, all ports are 100% local however this is determined by the number of available local ports. Thus vertical scaling and locality go hand in hand. In the center example below, more ports are needed than what is available with vertical scaling so horizontal scaling is used to increase the number of ports however locality is reduced. On the right, is an example of vertical scaling to use a large port count device to improve locality to improve performance.
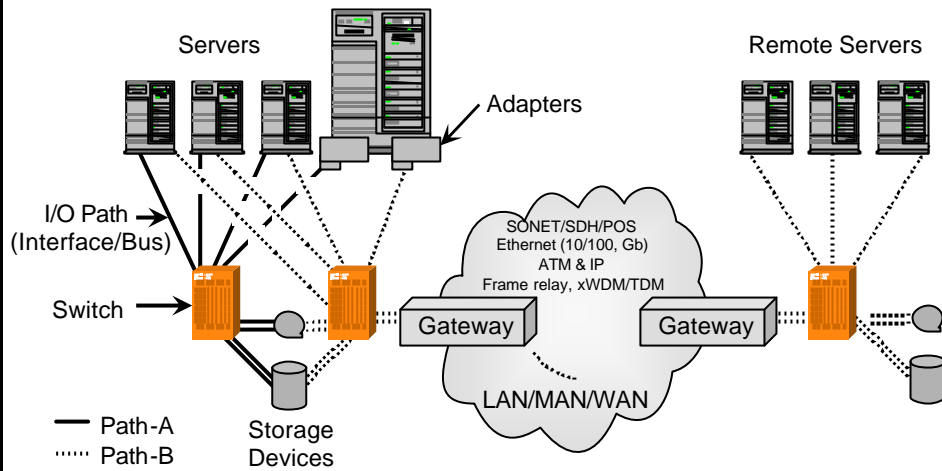


**100% Locality**

**0% Locality**

**100% Locality**

**100% Locality**

*Vertical Scaling*

*Horizontal Scaling*
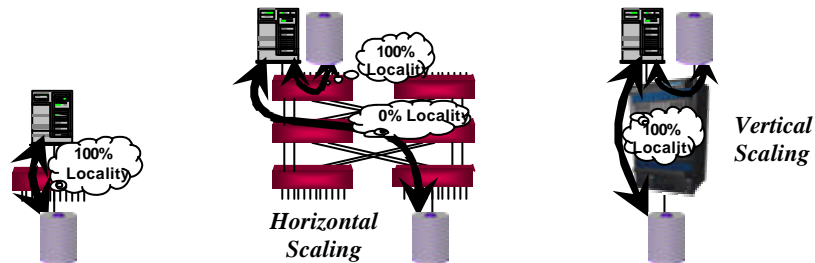
**Storage Networking Performance**
*Storage Networking Components*

**Storage Networking Performance**
*Storage Networking Enables Distance – Key To Survivability*
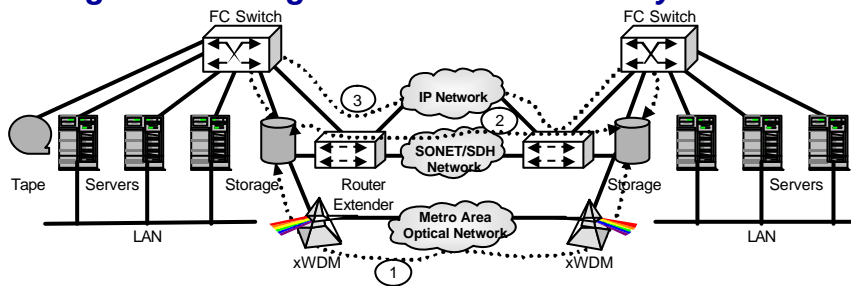
| Optical Level | SDH Equivalent | Electrical Level | Line Rate (Mbps) | Payload Rate (Mbps) | Overhead (Mbit/sec) |
|---|---|---|---|---|---|
| OC-1 | N/A | STS-1 | 51.84 | 50.112 | 1.728 |
| OC-3 | STM-1 | STS-3 | 155.5 | 150.336 | 5.184 |
| OC-9 | STM-3 | STS-9 | 466.560 | 451.008 | 15.552 |
| OC-12 | STM-4 | STS-12 | 622.080 | 601.344 | 20.736 |
| OC-18 | STM-6 | STS-18 | 933.120 | 902.016 | 31.104 |
| OC-24 | STM-8 | STS-24 | 1244.160 | 1202.688 | 41.472 |
| OC-36 | STM-12 | STS-36 | 1866.240 | 1804.032 | 62.208 |
| OC48 | STM-16 | STS-48 | 2488.320 | 2405.376 | 82.944 |
| OC96 | STM-32 | STS-96 | 4976.640 | 4810.752 | 165.888 |
| OC192 | STM-64 | STS-192 | 99 53.280 | 9621.504 | 331.776 |

# Storage Networking Update
## *Trunking To Enhance Performance*

✓ **Four Point Strategy Regarding Trunking…**

**1.** Physical Trunking
- Multiplexing Fiber Optic cables using WDM
- Cable and Fiber Management Systems (Trunk cables and connectors)
- Faster interfaces (e.g. 10Gb FC/Ethernet) logical, metro, and wide area

**2.** Load Balancing and Optimization
- Implementation (Aggregation)

**3.** Deterministic Routing (Activity Based)
- Specific which DID/SID pairs use which Isls
- Not really trunking, but may be more of what customers are looking for.

**4.** Class of Service and QoS (Note QoS is Critical for 10Gb)
- Combines #2 and #3
- Applies at native speeds as well!



http://www.decus.de

Copyright © C N T   **29**

---

# Storage Networking Update
## *Trunking To Enhance Performance*



http://www.decus.de

Copyright © C N T   **30**

**Storage Networking Industry Trends**

CNT
(Formerly INRANGE)

---

*Storage Networking Technology Evolution*

| Hubs | Switches | Directors | | Intelligent Switches And Directors |
|------|----------|-----------|--|-----------------------------------|

**HA Director**
2 Gbps ports
64-256 ports
FC & FICON

**Intelligent Director**
4. 10 Gbps ports
64-1000+ ports
FC, FICON, Ethernet (iSCSI)

**Director**
1 Gbps ports
64-128 FC ports
Core functions

**Fabric switch**
1 Gbps ports
8-16 ports
A nyto-any fabric

**FC Loop**
25-100 Mbyte/sec
126 devices
Shared loop

**Port-based storage services**
- Virtualization
- Replication
- Backup/recovery
- HSM
- Snapshot

**Non-disruptive code
load and activation
Interoperable
switches
Mainframe intermix &
CUP**

**External MAN/WAN
Port scalability
Non-blocking
RAS**

**Shared storage
access**
- Port zoning
-LUN mgmt.
LAN offload
Homogenous SAN

**Higher speed,
longer distance
storage connection**

**Management services**
- Provisioning
- Security
**Internal MAN/WAN
Heterogeneous SANs
Port partitioning**

| 1996 | 1998 | 2000 | 2002 | 2004 |
|------|------|------|------|------|

CNT

Copyright © C N T   32

16

# Storage Networking Trends
## Storage Networking Requirements and Industry Trends

- **More Scalability and Performance**
  - Port density and speeds continue to Increase
  - Quality of Service (QoS), self optimizing performance
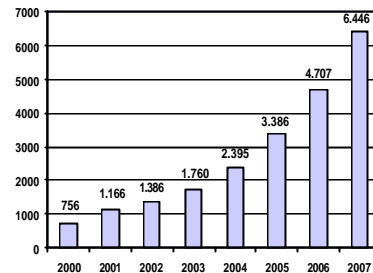  - Implement best architecture for performance and port growth requirements
- **Ultra High Availability**
  - Integrated HA for hardware, software and network
  - Fault detection, isolation, and containment
  - End to end security and data protection
  - Proactive monitoring and support
- **Multi-protocol and interfaces (Local, Metro, WAN)**
  - FC, IP, FICON, iSCSI, FCIP, GbE, SDH/SONET,IPFC, MAN

**FC Ports Forecast**



| Year | Value |
|------|-------|
| 2000 | 756 |
| 2001 | 1.166 |
| 2002 | 1.386 |
| 2003 | 1.760 |
| 2004 | 2.395 |
| 2005 | 3.386 |
| 2006 | 4.707 |
| 2007 | 6.446 |

**Over 17 million fibre channel switch and director ports forecasted to ship over next 4 years**

Source: IDC

---

# Storage Networking Trends
## Storage Networking Requirements and Industry Trends

✓ Larger, faster, intelligent switches performing "Heavy Lifting" functions

✓ Larger SAN "Islands" or distributed cores to consolidate first generation fabrics

✓ New layer of edge function moving further out into the enterpris e using IP

✓ Hardware virtualization assist in the fabric for storage virtualization functions

✓ Virtual channel interfaces (adapters) and virtual switches (logical domains)

✓ More interoperability across the board (Hardware, Software, Applications)

✓ Management Tools to address performance management, provisioning, Qos

✓ Backup improvements to speed recover with less downtime

✓ More convergence of networking interfaces and broader protocol support

✓ New interfaces and protocols including:
- IBA, iSCSI, FCIP, SMIS, FAIS, FSPF Backbone, 4Gb, 10Gb FC/GbE

## Some Additional Material

✓ "The Resilent Enterprise – Veritas Press
✓ "Leveraging MAN and WAN Connectivity" – SNWOnline – August 2002
✓ zJournal – "zStorage" – February 2003, "Virtualization" – June 2003
✓ Industry Associations (SNIA, FCIA, IETF, T11, SCTA) and their websites
✓ SNIA Web Site and Education Committee Tutorials (www.snia.org/education)
✓ CMG Proceedings and handouts on CD's
✓ CMG MeasureIT – "Networking With Your Storage", May 2003
✓ Evaluator Group website (www.evaluatorgroup.com)
✓ "Resilient Storage Networks – Designing Flexible Scalable Data Infrastructures"

**Resilient Storage Networks**

Designing Flexible Scalable
Data Infrastructures

Evaluator Group, Inc.

Source:Greg Schulz

Copyright © C N T    35

---

# Fragen ?

**Udo_Kuhr@cnt.com**

**www.cnt.com**