SNIA

STORAGE NETWORKING INDUSTRY ASSOCIATION

EDUCATION

Metropolitan and Wide Area Storage Networking

Greg Schulz, Evaluator Group

greg@evaluatorgroup.com

Paul Schoenau, CIENA Corp.

pschoena@ciena.com

Abstract

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION
EDUCATION

Metropolitan and Wide Area Storage Networking

Distance is essential to support business continuity, compliance, and consolidation. This session explains the components and provides examples of how to utilize available technologies to remove storage networking distance boundaries. This tutorial looks at different technologies and techniques including ATM, CWDM, DWDM, FCIP, Fibre Channel, Ethernet, FICON, iFCP, TCP/IP, iSCSI, SONET/SDH.

Some things you will learn in this tutorial include:

- Demystify metropolitan and wide area storage network communications
- Knowledge to aid in selection of communications services
- How to use appropriate technology for different requirements

Note: Availability of specific technology, features, and services may vary by markets and geographical area.

© 2005 Storage Networking Industry Association. All Rights Reserved.

SNIA Legal Notice

- The material in this tutorial is copyrighted by the SNIA.
- Member companies and individuals may use this material in presentations and literature under the following conditions:
  - Any slide or slides used must be reproduced without modification
  - The SNIA must be acknowledged as 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.

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Key Message

- Storage Networks can be built for a variety of applications including host to storage, storage to storage, and san to san to support diverse business needs over distance.
- Fibre Channel can extend to 100km and beyond today using dedicated fiber with CWDM/DWDM, SONET/SDH, and storage over TCP/IP.
- Metropolitan and wide area storage networks enable various distances to be spanned using different networking interfaces.
- Effective bandwidth and low latency are critical for storage applications particularly over distance.

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Agenda

- **Metropolitan storage networks overview**
  - **Applications**
  - **Networking technologies**
  - **Fibre Channel Distance Issues**
- **Transport Details**
  - **Terminology**
  - **DWDM & CWDM**
  - **SONET/SDH**
  - **TCP/IP**
- **Wrap Up**

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Terminology

- Dark Fiber (Dedicated Fiber Optics)
  - Un-lit, un-used, un-managed fiber optics
  - The physical connection used in optical networks
- TDM (Time Division Multiplexing)
  - Multiple data paths sharing bandwidth in the same time slice
- WDM (Wave Division Multiplexing)
  - Multiple unique wavelengths/lambdas/colors per fiber pair
    - Multiple network protocols per fiber (e.g. SONET/SDH, FC, IP)
    - Multiple data paths operating in parallel at the same time
  - Two technologies (CWDM & DWDM):
    - Coarse WDM – campus & metro; low cost; low density
    - Dense WDM – metro & regional; high density; long reach

© 2005 Storage Networking Industry Association. All Rights Reserved.

Terminology

- Optical Carrier (OC-x) Network
  - Fiber optic based backbone network and bandwidth allocation
  - Examples include OC-3 (155Mb), OC-12 (622Mb), OC-96 (4,976Mb)
- SONET/SDH(Synchronous Optical Network/Synchronous Digital Hierarchy)
  - Layer 1 - 2 technology – circuit/path based connection
    - High availability (150K rings in North America, 85% access)
    - Common transport for metro/regional ATM & IP
  - Clear Channel can allocate bandwidth for multi-protocol connectivity
    - GFP (Generic Framing Protocol) supports multi-vendor interop
  - Fibre Channel mapped to scaleable SONET/SDH bandwidth
    - 5MBps or less granularity with low latency, good effective bandwidth
    - Utilizes OC-x or STM-x fiber bandwidth allocation

© 2005 Storage Networking Industry Association. All Rights Reserved.

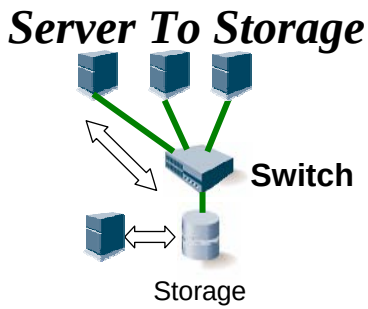
Terminology

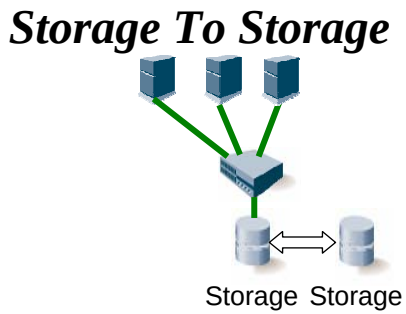
- TCP/IP – Internet Protocol
  - Layer 3 (IP) – 4 (TCP) technology – packet based routing
    - Data protocol of choice – non-deterministic, best effort transport
  - Utilizes Ethernet or some other transport locally or
    - Ethernet rides over fiber/WDM or tunneled via SONET/SDH
    - 10Gb Ethernet can be LAN or WAN using different interfaces
  - Can transport storage (FCIP, iFCP, iSCSI, NAS (NFS, CIFS))
- Effective bandwidth
  - Bandwidth that can actually be utilized vs. theoretical rate
- Wavelength Service
  - Fiber optic bandwidth provided by service provider
  - Based on a WDM lambda wavelength of light (WDM service)
  - May be protected or un-protected service

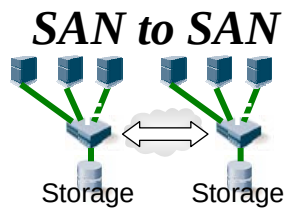
Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

SAN Implementations

- Server (host) to device (storage, disk, tape)
  - Stranded (remote) server accessing storage
  - Local server accessing remote storage
  - Clustering and failover access to storage
  - Server to NAS, and NAS to storage
- Device to device (storage to storage)
  - Local and remote mirroring
  - Synchronous and asynchronous mirroring
  - Multi-hop (cascade) mirroring to multi-sites
- SAN to SAN
  - Single SAN (Fabric), Multiple Fabrics
  - Local, Metro, and Wide Area







Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

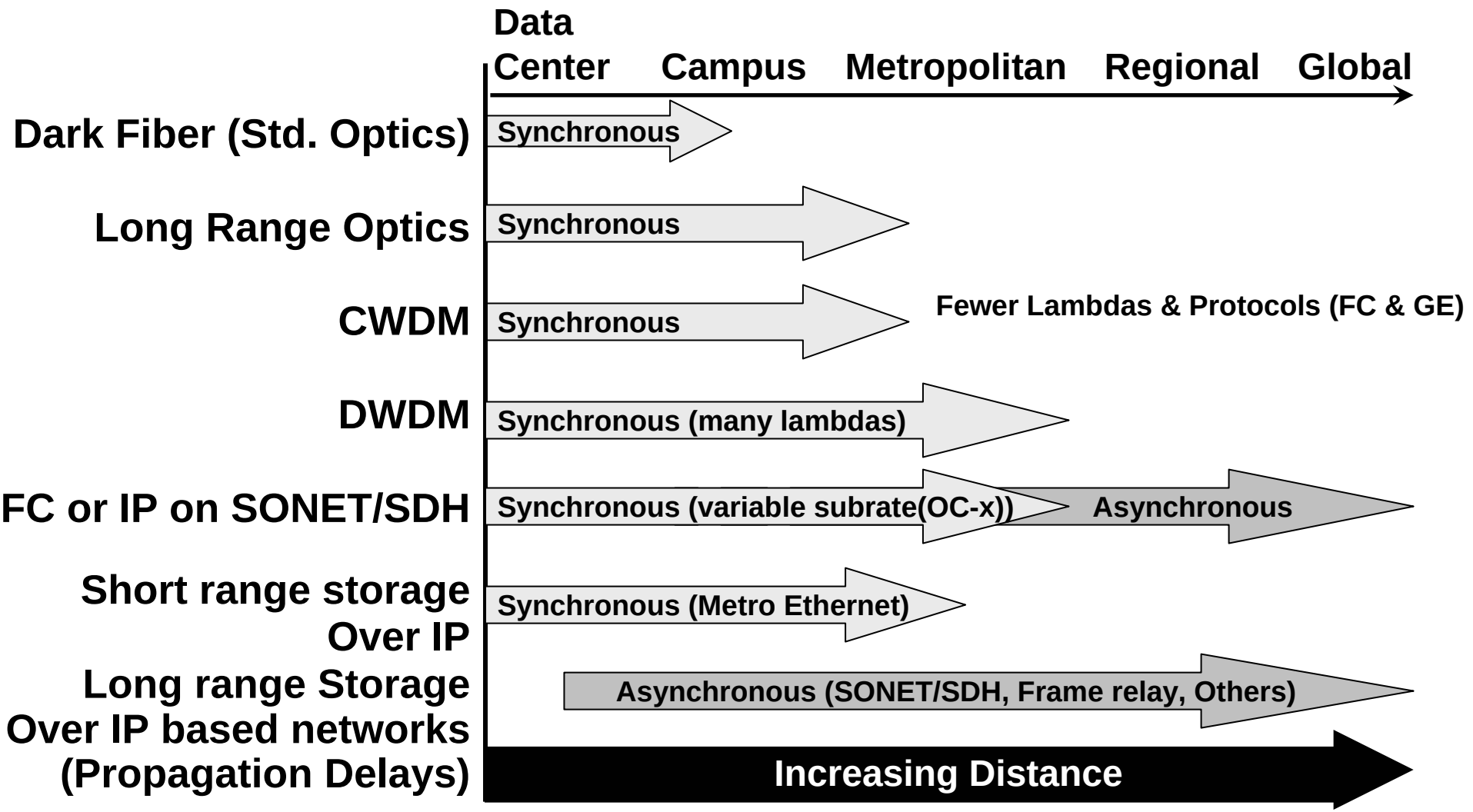
Storage Network Layers

SNIA
Storage Networking Industry Association
EDUCATION

OSI Layer	Description	FC Layer	Fibre Channel SAN	Gigabit Ethernet Network	IP Routed Network
7	Application		Filesystem, FTP	FTP, Telnet	FTP, Telnet
6	Presentation		SCSI-3, NFS	HTTP, Telnet	HTTP
5	Session		CCW FCP, IP, FICON	NFS, CIFS, iSCSI	NFS, iSCSI
4	Transport		VI, AV, AE		TCP, UDP
		FC-4	ULP		
3	Network			TCP (IP & UDP)	IP
		FC-3	Services		
		FC-2	Framing, Flow	MAC Client	LAN,WAN
2	Datalink		Flow Control	MAC Client MAC	MAN
		FC-1	Encoding, Link		
1	Physical	FC-0	Physcial	Physical	Physical

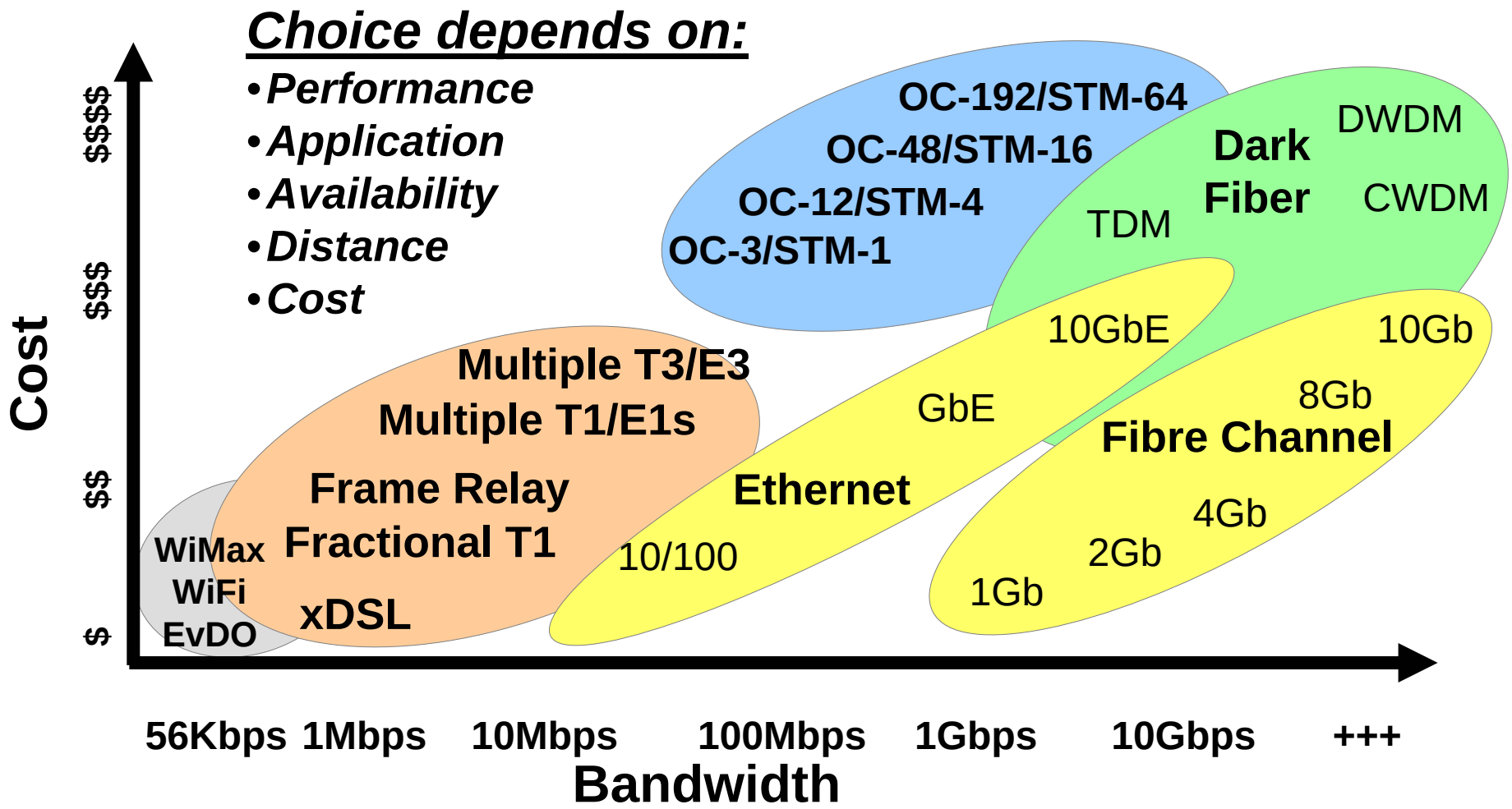
© 2005 Storage Networking Industry Association. All Rights Reserved.

Storage Over Distance



Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Storage Transport Options



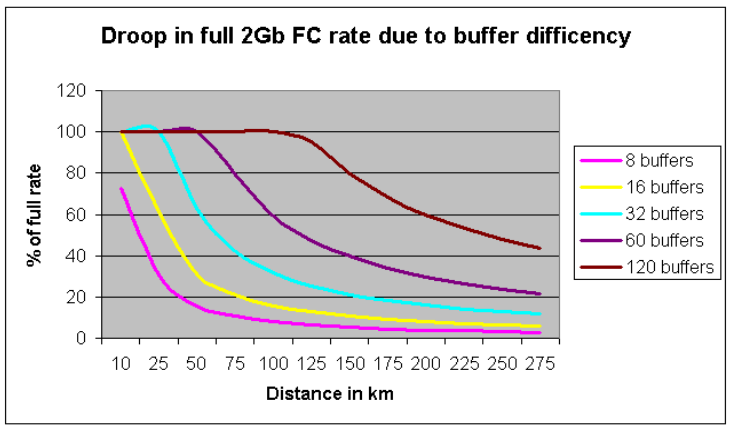
Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Fibre Channel over Distance

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION
EDUCATION

Droop

- Phenomenon of performance degradation over distance
- Performance drops before theoretical protocol bandwidth
- Not to be confused with propagation delay (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



Buffer Credit Impact Over Distance

Number of buffer credits To span various distance

	1km	2km
1Gb	½	1
2Gb	1	2
4Gb	2	4

© 2005 Storage Networking Industry Association. All Rights Reserved.

TCP/IP over Distance

Droop can also occur with TCP/IP

- The TCP window size can have an impact on performance
- A larger window size too a certain point can improve performance
- At some point a window size too large can hurt performance
  A windows size too small, or to large can result in packet drop
- A larger TCP window size can have a greater impact of packet drop
  More data too be retransmitted that can cause server overhead
- Size and tune the TCP window size appropriately
  For example factor delay times bandwidth as well as line errors
- IETF RFC 1323 adds TCP options for performance
  - Ability too scale TCP window size up too 1GB

© 2005 Storage Networking Industry Association. All Rights Reserved.

Agenda

- **Metropolitan storage networks overview**
  - **Applications**
  - **Networking technologies**
  - **Fibre Channel Distance Issues**
- **Transport Details**
  - **Terminology**
  - **DWDM/CWDM**
  - **SONET/SDH**
  - **TCP/IP**
- **Wrap Up**

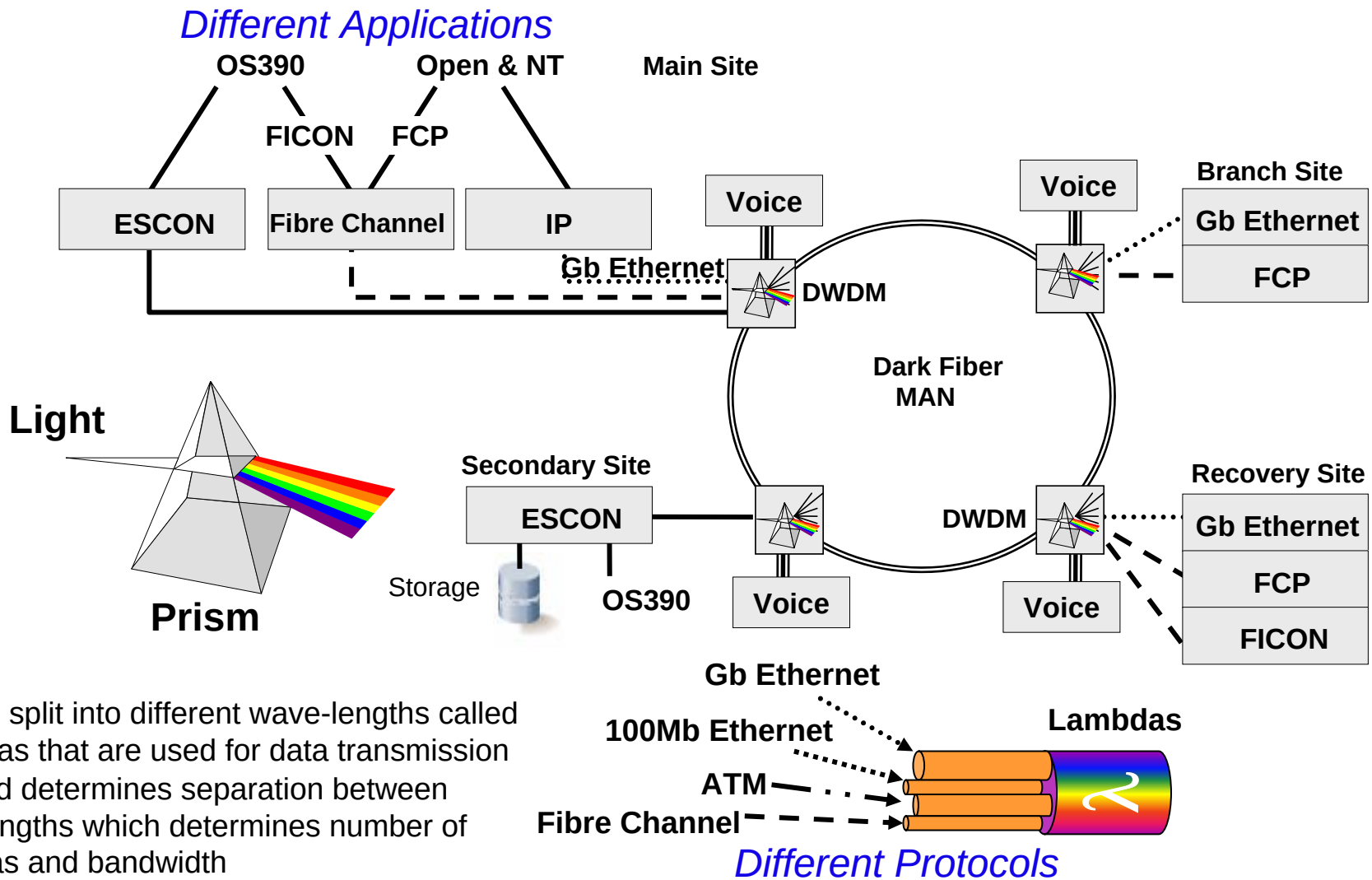
Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

DWDM (MAN) Storage Networking

- DWDM (Also CWDM) Enables and Provides:
  - Increased bandwidth and capacity between locations
    - Multiple wavelengths/networks on a single fiber
  - Increased distances
    - Optics drive storage connectivity over hundreds of km
    - Sufficient buffer credits are needed for flow control
  - Simplified cabling and infrastructure
    - Multiple physical optical network links can be consolidated
    - Common trunking of all networks over available fiber
    - Reduces total cost of ownership of fiber optic cable
  - Multiple fabrics including redundant cascades
    - Dual cascades for redundancy that span two sites
    - Protocol independent to support FICON, Fibre Channel, ESCON, ATM, Gb Ethernet, FDDI, and others all at varying speed

© 2005 Storage Networking Industry Association. All Rights Reserved.

DWDM Optical Networking



- Light is split into different wave-lengths called Lambdas that are used for data transmission
- ITU grid determines separation between wavelengths which determines number of lambdas and bandwidth

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

SONET/SDH Storage Networking

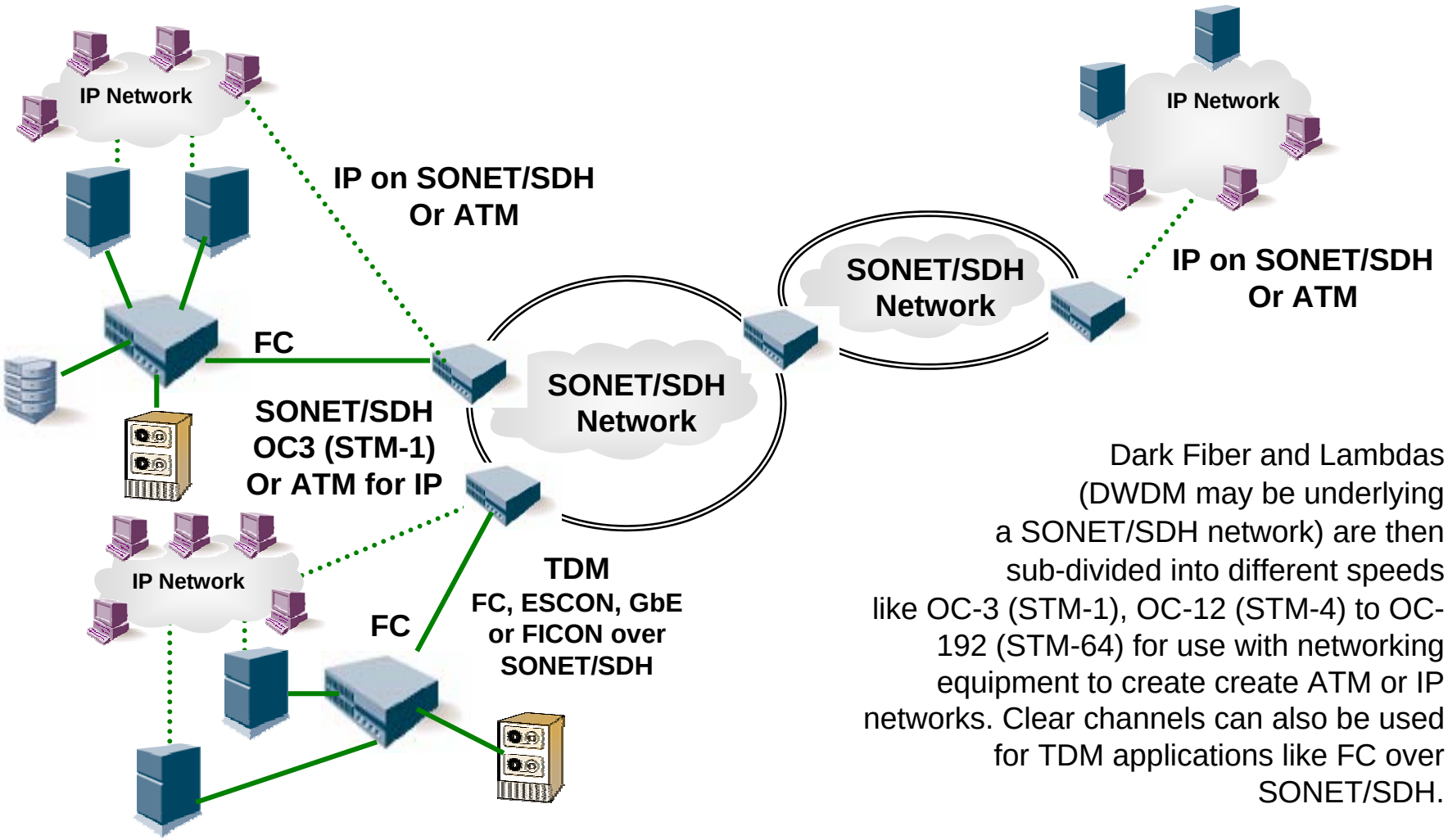
SNIA
Storage Networking Industry Association
EDUCATION

SONET/SDH Enables and Provides:

- Established with wide availability (150K rings in North America)
- Connection based with scalable bandwidth
- Path based redundancy for protection <50msec
- Long distance support from Metro to Wide area
- Good effective bandwidth with low latency
- New advances in flow control & data compression minimize WAN requirements with reduced latencies
- New standards increase interoperability and flexibility
  - Generic Framing Procedure (GFP)
  - Virtual Concatenation (VC)
  - Link Capacity Adjustment Scheme (LCAS)

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

SONET/SDH and TDM



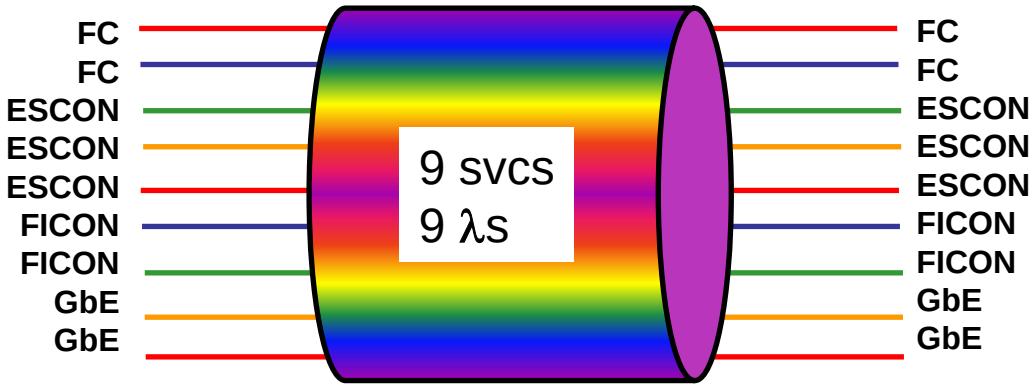
Dark Fiber and Lambdas (DWDM may be underlying a SONET/SDH network) are then sub-divided into different speeds like OC-3 (STM-1), OC-12 (STM-4) to OC-192 (STM-64) for use with networking equipment to create create ATM or IP networks. Clear channels can also be used for TDM applications like FC over SONET/SDH.

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Multi-Protocol Multi-Plexing

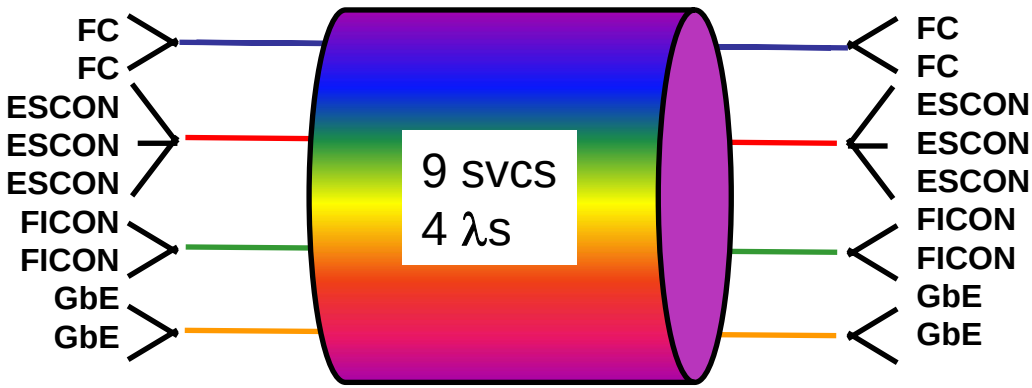
Simple DWDM

- One service channel per wavelength



Enhanced DWDM with sub-rate multiplexing

- Multiple service channels multiplexed into a single wavelength



Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Storage over IP for Distance

SNIA Storage Networking Industry Association
EDUCATION

Leverage IP for Long Haul Applications
- Asynchronous mirroring / backup / remote vaulting
- Latency can be sacrificed for distance and cost

Note: iFCP and other storage over IP (SoIP) details are covered In IP storage tutorial

FCIP (Fibre Channel over IP) & iFCP Enables:
- Fibre Channel & FICON extension over IP

iSCSI (SCSI over TCP/IP) Enables and Provides:
- Maps SCSI commands on TCP/IP (similar to FCP = SCSI on Fibre Channel)
  - Software initiators on servers, or hardware based TOEs
  - Enables IP (e.g. Ethernet) based networks to support block (SCSI) storage
  - iSCSI is an IETF standard (RFC 3720) – www.faqs.org/rfcs
  - Others include 3347-Rqmts,3721-Discovery,3722-Naming,3723-Security)
- Routers/bridges enables iSCSI servers to access FCP devices
  - Standalone devices and blades for switches
  - No changes for standard Ethernet switches

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

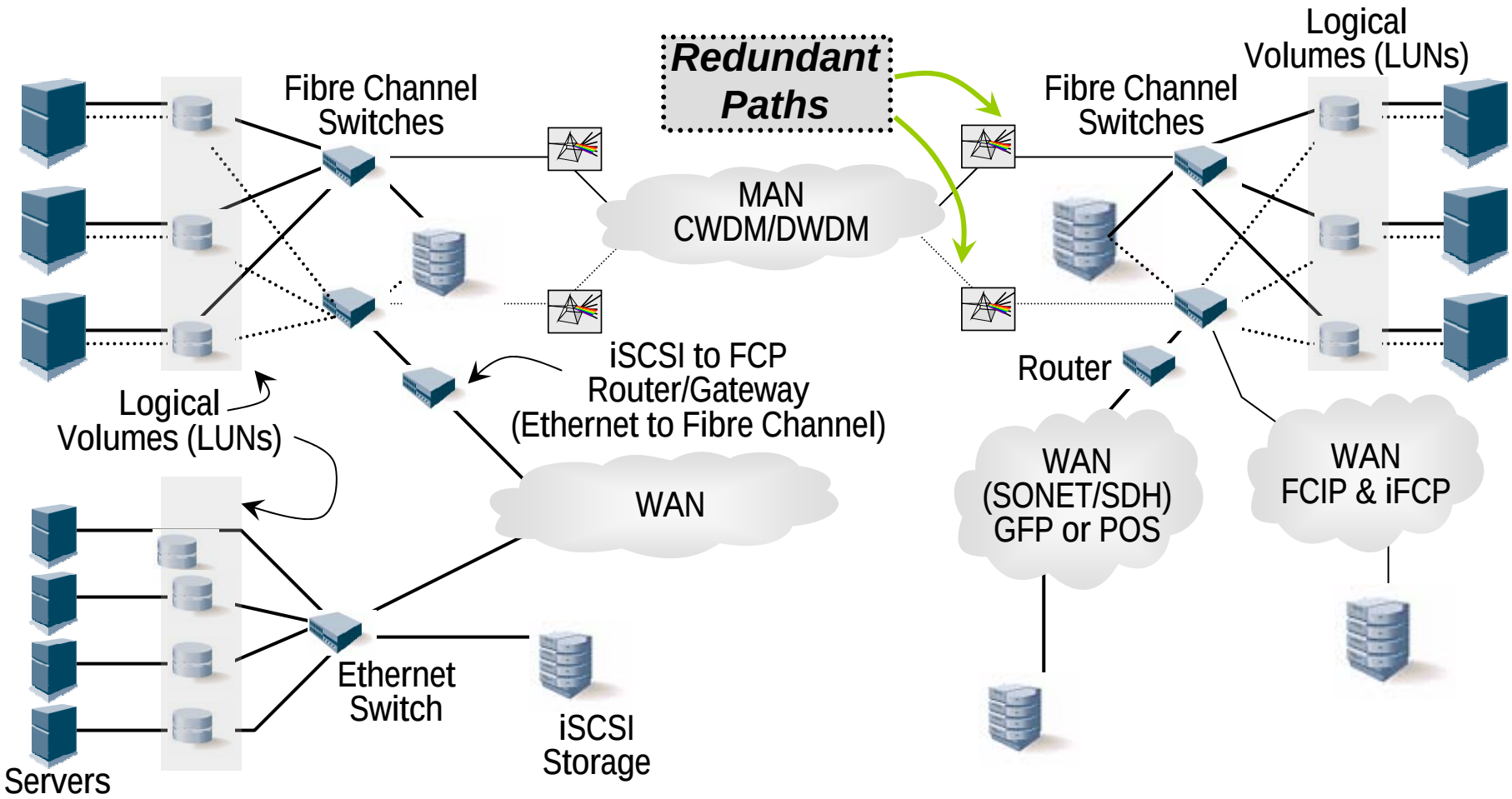
Heterogeneous Storage Networking

Multiple technologies for different needs

- Dedicated optics for LANs and local SANs
  - Ethernet (10/100/1000/10000 Mbs) including Metro Ethernet
  - Fibre Channel for FCP and FICON using long range optics
- DWDM using dedicated dark fiber and optical Lambda services
  - Support multiple network interfaces (Ethernet, FC, ESCON, etc)
  - Increased bandwidth with cost spread over more lambdas/bandwidth
- Storage over SONET/SDH using Optical Carrier (OC-x) networks
  - Fibre Channel including FICON directly over SONET/SDH - GFP
  - Ethernet at various speeds to carry IP directly over SONET/SDH
  - Variable bandwidth options available from (OC-3, OC-12, etc)
- Storage over IP based networks for extended distance
  - iSCSI, FCIP, iFCP, and NAS (NFS & CIFS)

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Heterogeneous Storage Networking



Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Agenda

- **Metropolitan storage networks overview**
  - **Applications**
  - **Networking technologies**
  - **Fibre Channel Distance Issues**
- **Transport Details**
  - **Terminology**
  - **DWDM / CWDM**
  - **SONET/SDH**
  - **TCP/IP**
- **Wrap Up**

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Some Best Practices

SNIA
Storage Networking Industry Association
EDUCATION

- The Storage Networking World Is Not Flat!
  - Fibre Channel can span distances of over 100km
  - Storage networks can span distances of 100's to 1,000s of km
- Use the right tool for the application
  - Avoid excessive layering of protocols and interfaces that result in latency
  - iSCSI for SCSI only remote access applications or FCIP for extension
  - SONET/SDH extends storage over distance for Fibre Channel & Ethernet
  - Different topologies can be used to meet specific needs and requirements
- Understand bandwidth requirements and needs
  - Wire speed differs from effective speed and protocol overhead
  - Look at low latency, and effective bandwidth vs. stated wire speeds
- Design and deploy for resiliency (High Availability) and flexibility
  - Plan for maintenance, growth, diagnostics, and management
  - Utilize redundant paths including communions paths
  - Verify that your metro and WAN communication paths are unique

© 2005 Storage Networking Industry Association. All Rights Reserved.

Additional Reading Material

SNIA
Storage Networking Industry Association
EDUCATION

- Resilient Storage Networks - Designing Flexible Scalable Data Infrastructures
  Greg Schulz – Elsevier/Digital Press Books ISBN: 1555583113
- iSCSI: The Universal Storage Connection
  John L. Hufferd - Addison-Wesley Books ISBN: 020178419X
- IP SANS: An Introduction to iSCSI, iFCP, and FCIP
  Tom Clark - Addison-Wesley Books ISBN: 0201752778
- "The Resilent Enterprise" - Veritas Press
- SNIA Dictionary – www.snia.org/dictionary
- SNIA Web Site - www.snia.org/education

© 2005 Storage Networking Industry Association. All Rights Reserved.

Q&A / Feedback

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

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

SNIA Education Committee

CIENA – Stephen Barr
Cisco – Walter Dey
CNT - Bill Bolton
CNT - Mike Naylor
Nortel – Jack Hunt
McData – Simon Gordon
Solutions Technology – David Deming
CIENA - Paul Schoenau
CNT – Brian Larsen
Evaluator Group - Greg Schulz
NeoScale - Robert Lockhart
Microsoft – SW Worth

And special thanks to:
Nancy Clay (SNIA)

© 2005 Storage Networking Industry Association. All Rights Reserved.

SNIA

STORAGE NETWORKING INDUSTRY ASSOCIATION

EDUCATION

Optional Material

Storage Network Interfaces/Protocols

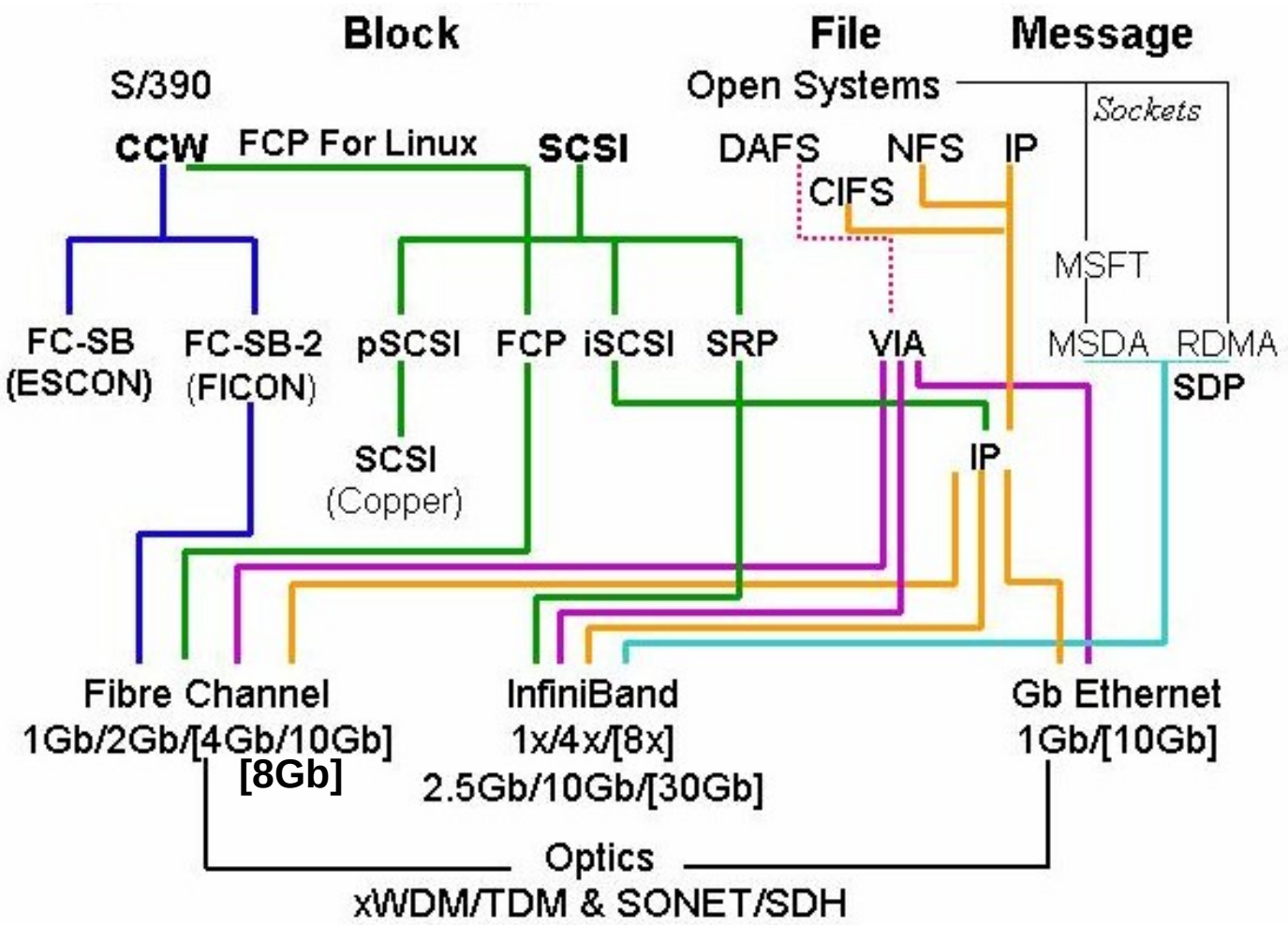
EDUCATION

Interface	Speed	Distance	Protocols	Characteristic	Usage	Comment
ESCON	17MBps	2km - 30 With Repeaters	CTC, Storage	Deterministic Channel	Mainframe Interconnect	Being replaced By FICON
Ethernet (10/100/1000 & 10000	10Mbps to 10Gbps	2km to 200+ km With Repeaters	IP, VI, IP Storage (e.g. iSCSI, FCIP)	Local Area Networks (LANs & MANs)	Voice, Video, Data	10GbE is MAN & LAN iSER & RDMA
Fibre Channel	1,2,4,,[8],10Gb	10-80km W//O Repeaters	FCP, FICON, IP, VI	Deterministic Channel	Storage	4Gb Accelerating 10Gb Emerging
InfiniBand	1x = 2.5Gbps 4x = 10 Gbps 12x = 30Gbps	Local Data Center	FCP, FICON, IP, VI	Deterministic Channel	Host Bus Connect High Speed SAN Storage	For clustering Not a PCI replacement
ATM	OC-3 (STM-1) to OC-192 (STM-64)	Variable	IP, VI, FCBB ESCON,FICON	Network	Voice, Video, Data	
SONET/SDH	OC-3 (STM-1) to OC-192 (STM-64)	Variable	ATM, IP, FCBB, ESCON, FICON	Network Backbone	Metro WAN	Storage over distance with SONET/SDH
DWDM	Over 100Tb	100-200km+	All	Multiple Networks	Local, Metro	Dedicated Fiber Optics Can be difficult to get

© 2005 Storage Networking Industry Association. All Rights Reserved.

Storage Network Interfaces/Protocols

ULPs



Note: FCIP maps Fibre Channel frames including multiple ULPs into IP packets for transmission over IP based networks

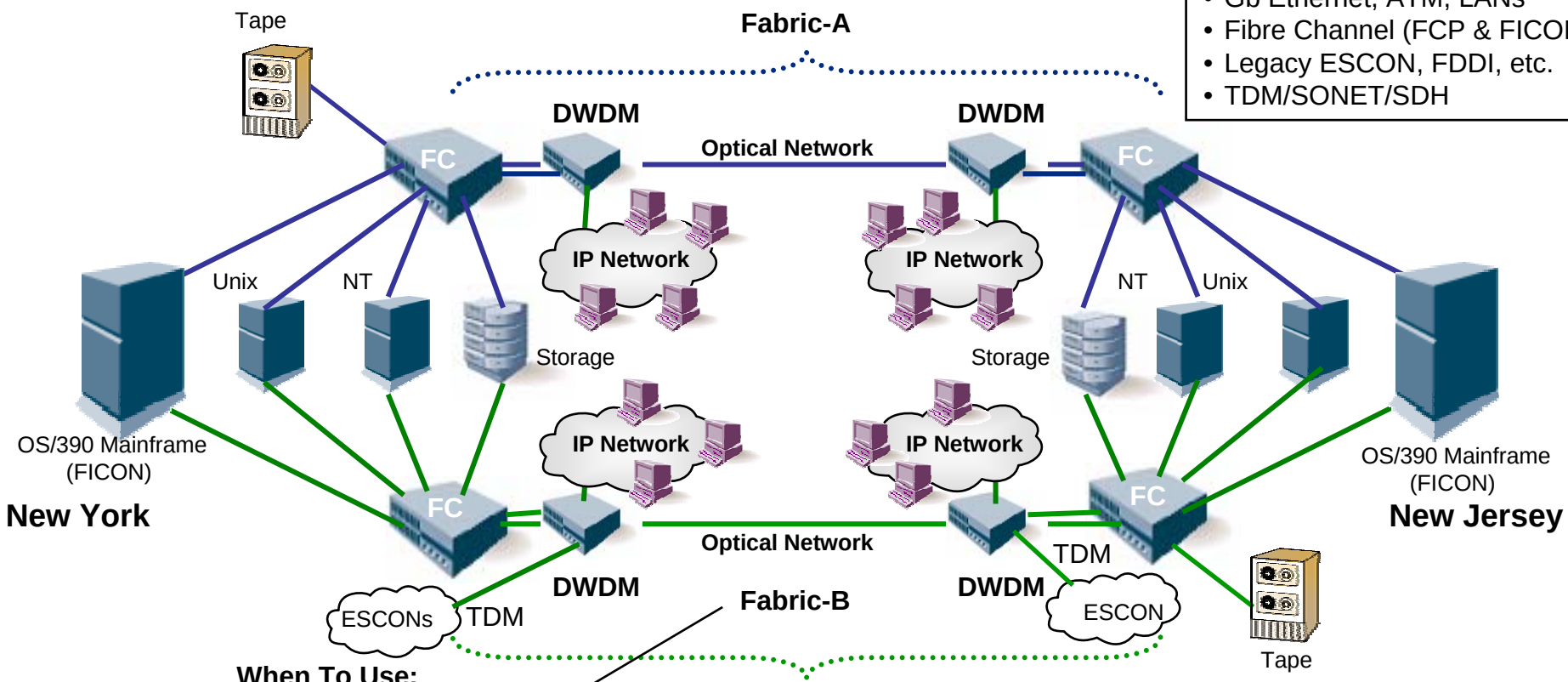
Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

DWDM (MAN) Storage Networking

SNIA
Storage Networking Industry Association
EDUCATION

DWDM Enables:

- Gb Ethernet, ATM, LANs
- Fibre Channel (FCP & FICON)
- Legacy ESCON, FDDI, etc.
- TDM/SONET/SDH



When To Use:

- Dark fiber is available
- Consolidation of fiber
- Trunking and performance
- Multiple physical networks

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

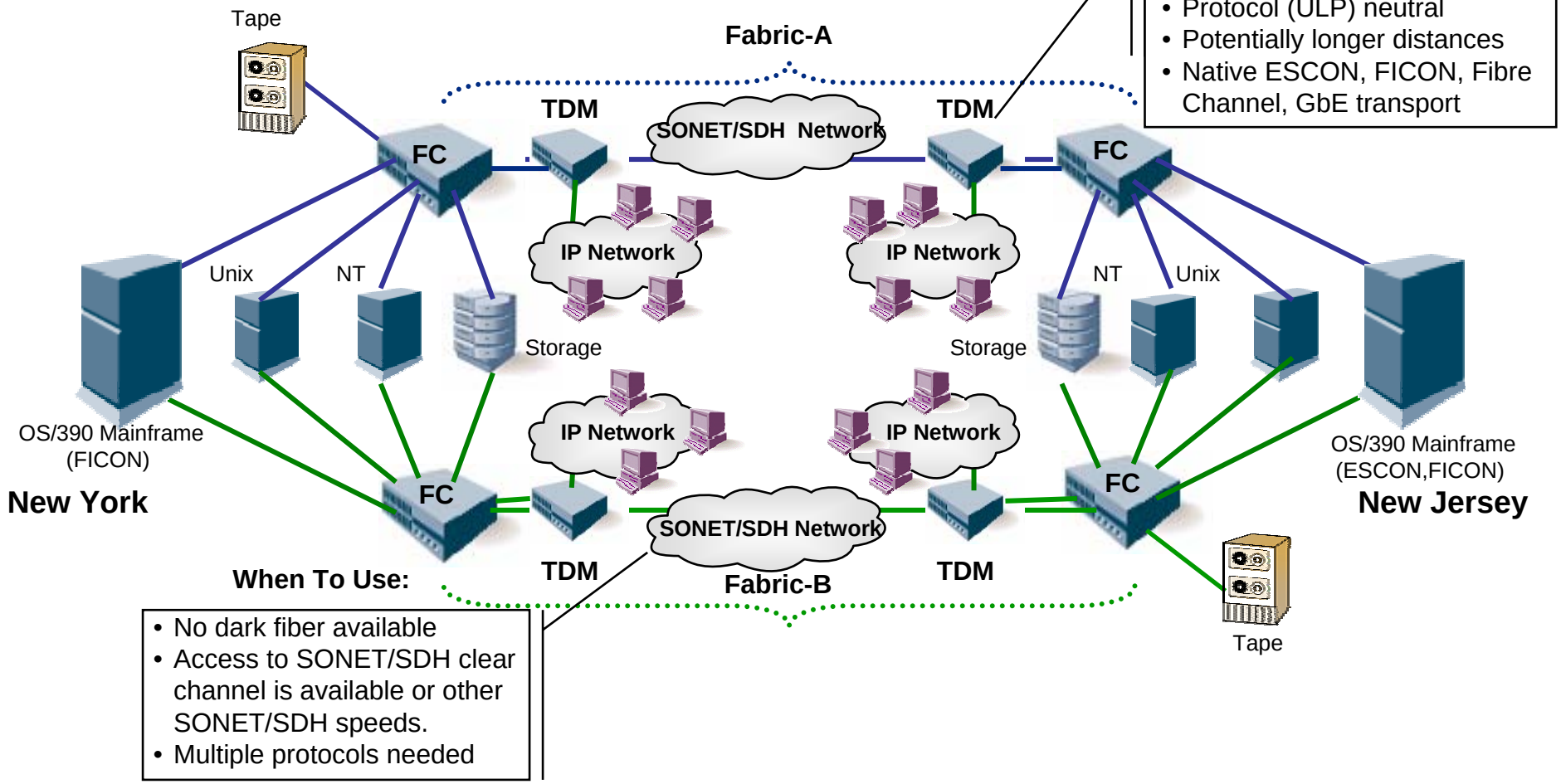
Optical Bandwidth Examples

SNIA
Storage Networking Industry Association
EDUCATION

Optical Carrier (OC) Level	SDH Equivalent	Electrical Level	Line Rate (Mbps)	Payload Rate (Mbps)
OC-3	STM-1	STS-3	155.5	150.336
OC-9	STM-3	STS-9	466.560	451.008
OC-12	STM-4	STS-12	622.080	601.344
OC-18	STM-6	STS-18	933.120	902.016
OC-24	STM-8	STS-24	1,244.160	1,202.688
OC-36	STM-12	STS-36	1,866.240	1,804.032
OC-48	STM-16	STS-48	2,488.320	2,405.376
OC-96	STM-32	STS-96	4,976.640	4,810.752
OC-192	STM-64	STS-192	9,953.280	9,621.504

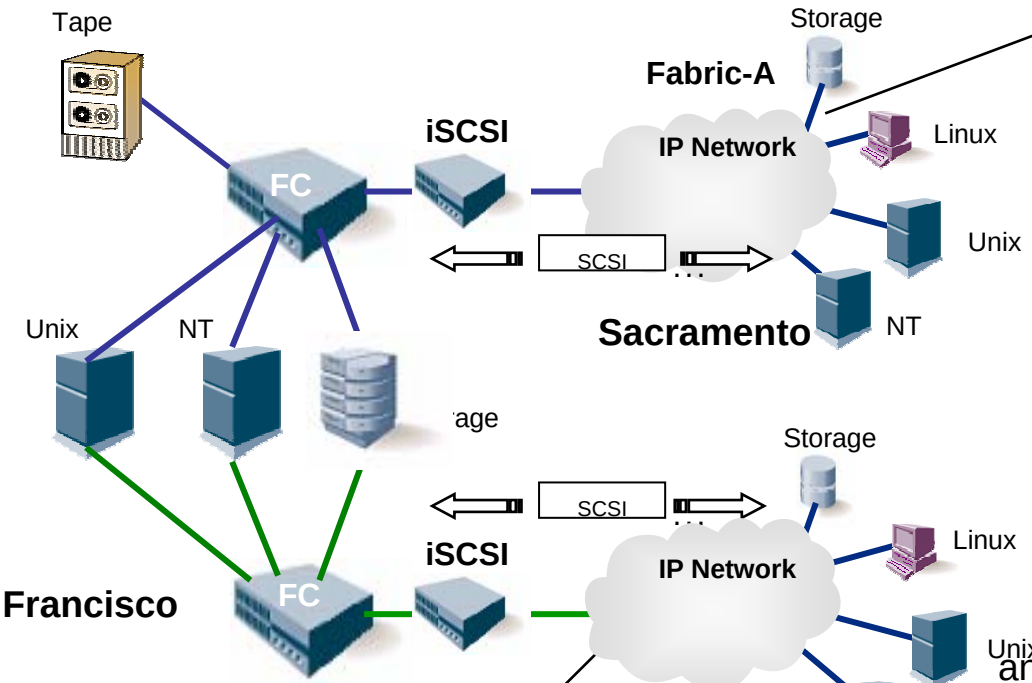
© 2005 Storage Networking Industry Association. All Rights Reserved.

SONET/SDH Storage Networking



Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

iSCSI Storage Networking



iSCSI Enables:

- IP-based SAN connected
- All SCSI SAN (FCP & iSCSI)
- Leverage IP technology
- Utilize native iSCSI devices

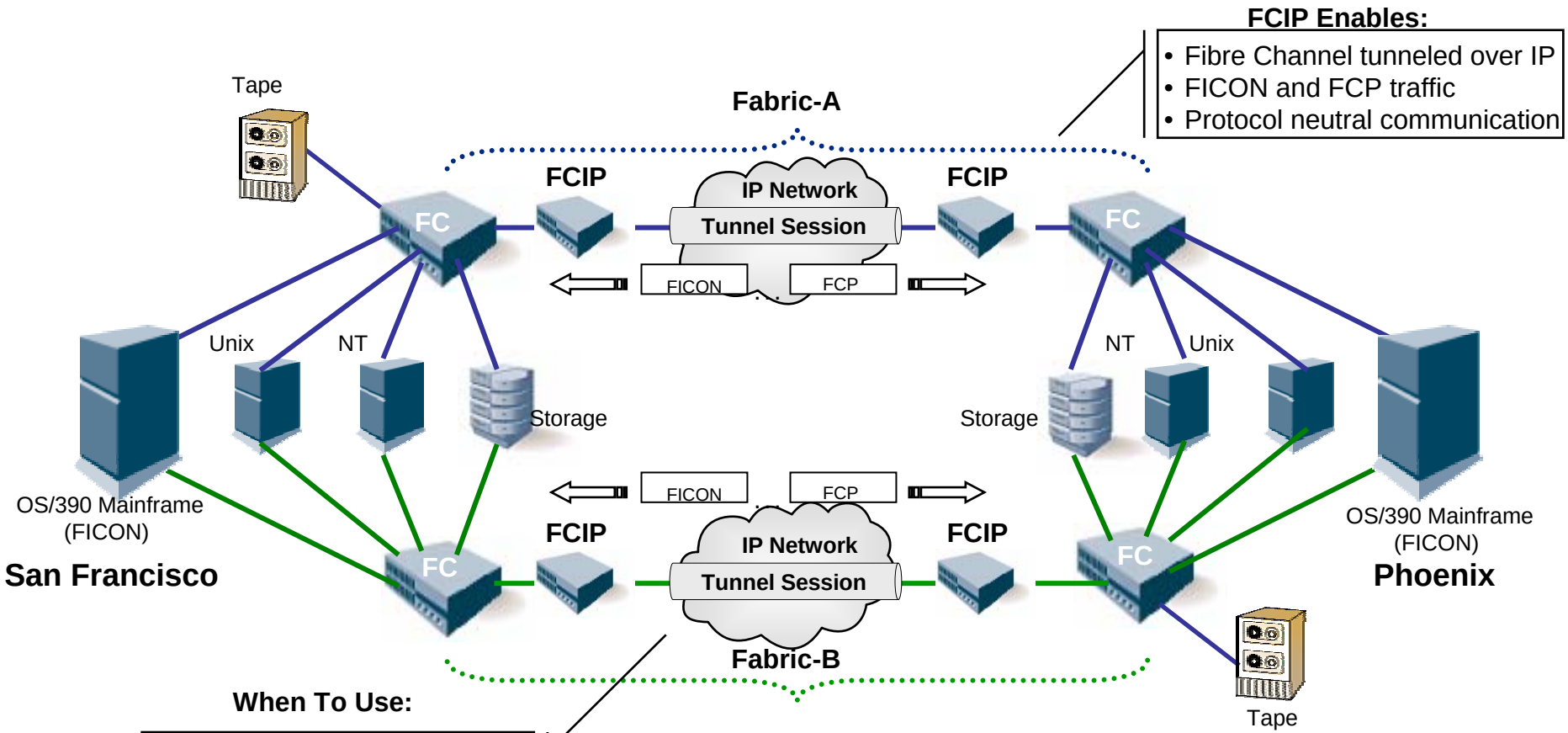
Primary storage networking at main site is based on a redundant Fibre Channel Storage network with two fabrics. Remote Sites and branch offices access the core SAN via iSCSI (SCSI over TCP/IP). The core SAN with Fibre Channel handles "heavy lifting" for high performance and shared access. In this example the remote SANs are not redundant and only attach to one of the primary core SANs. For high availability, the remote SANs should be attached to both.

When To Use:

- Workgroup, branch offices
- Low performance, low cost
- Medium distance applications
- Single protocol (SCSI)

Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

FCIP Storage Networking



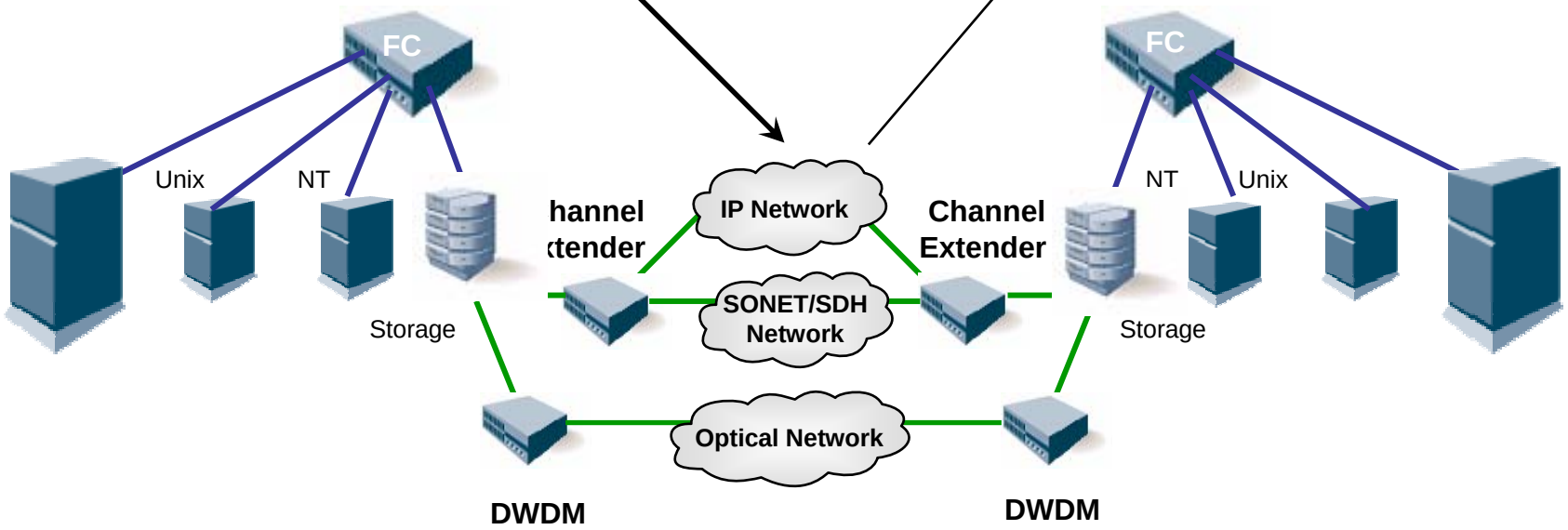
Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.

Storage to Storage Example

Tip! - While not shown, for high availability, resiliency, and flexibility, deploy redundant, divergent (unique) communication paths

Channel Extension Enables

- Remote mirroring across sites
- Compression, load balancing
- Frame passing, device emulation
- Choice of interfaces
  - ESCON of Fibre Channel
  - IP, SONET/SDH, T1, WDM, etc.



Metropolitan and Wide Area Storage Networking© 2005 Storage Networking Industry Association. All Rights Reserved.