Why Fibre Channel?

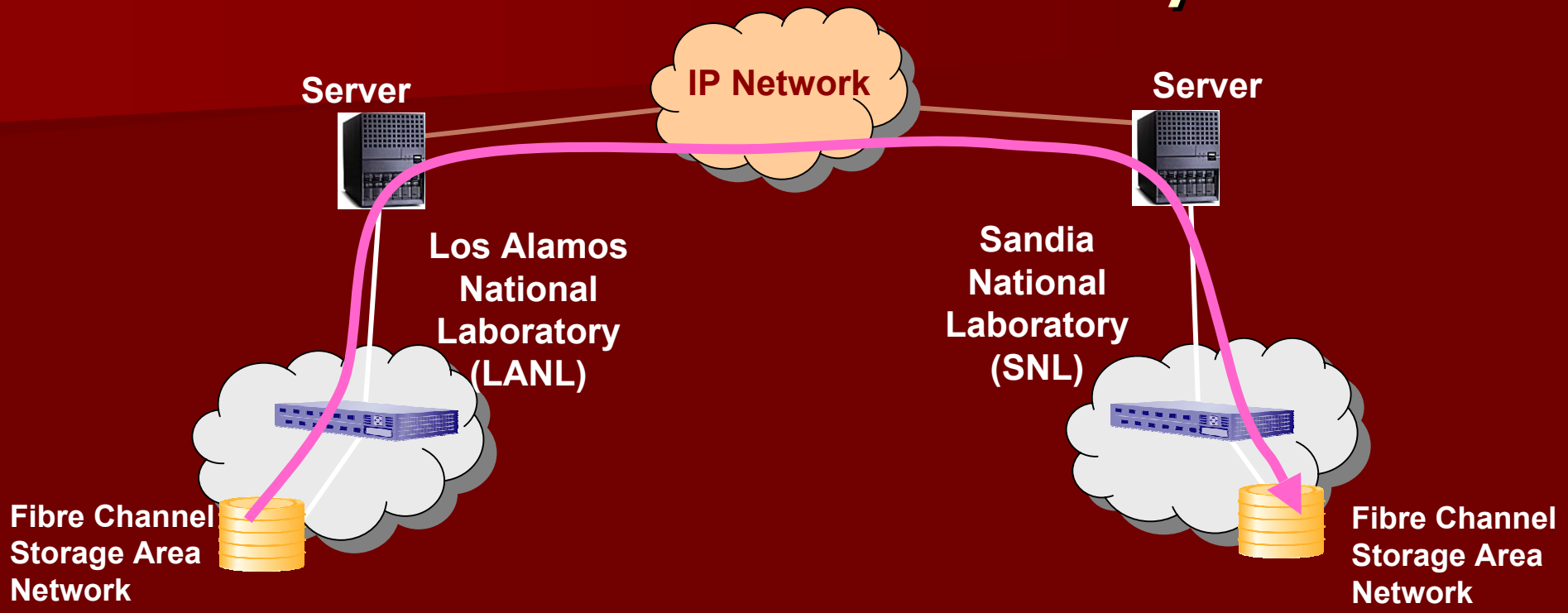
- Storage to storage transfers demands a high QoS
  - Needs "deterministic" communications
  - Generally very sensitive to latency
  - Requires in-order block delivery
- Fibre Channel is deterministic
  - Provides low latency and high bandwidth interconnect
- High performance SAN extension must maintain same QoS features over distance
  - High bandwidth
  - Low latency

How Much Bandwidth is Enough…?

- Moving 10 TB requires:
  - 2.25 hours using OC-192 (10Gb/s)
  - 9 Hours using OC-48 (2.5Gb/s)
  - 14 hours using "2G" FC (1600 Mb/s)
  - 28 hours using "1G" FC (800Mb/s)
  - 35.7 hours using OC-12 (622 Mb/s)
  - 6 days using OC-3 (155 Mb/s)
  - 20 days using T3 (45 Mb/s)
  - 1.6 years using T1 (1.5 Mb/s)

……If the pipe is fully utilized!

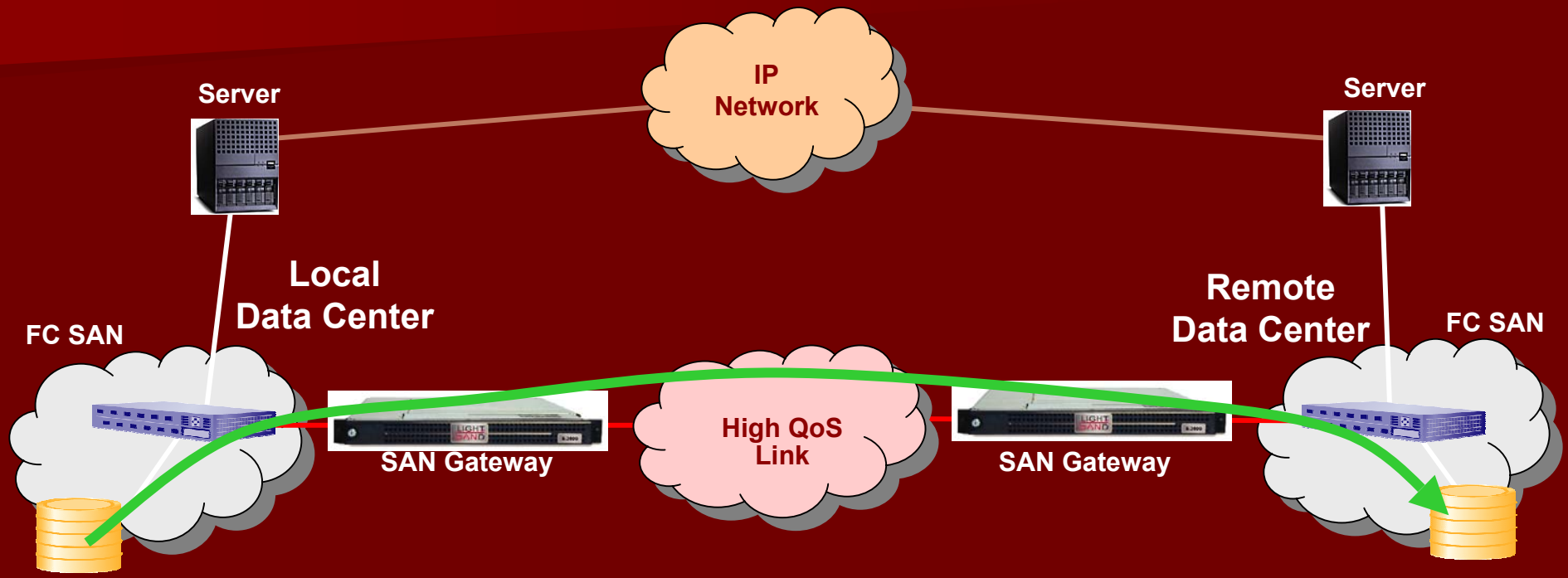
Data Movement Today – A Recent Case Study



Scientists at LANL currently dump 100GB of supercomputing data to tape and FedEx it to SNL because it is faster than trying to use the existing 155Mb/s IP WAN connection

- Actual measured throughput of 16Mb/s! (10% bandwidth utilization)

The Better Way – Directly Between Storage Systems



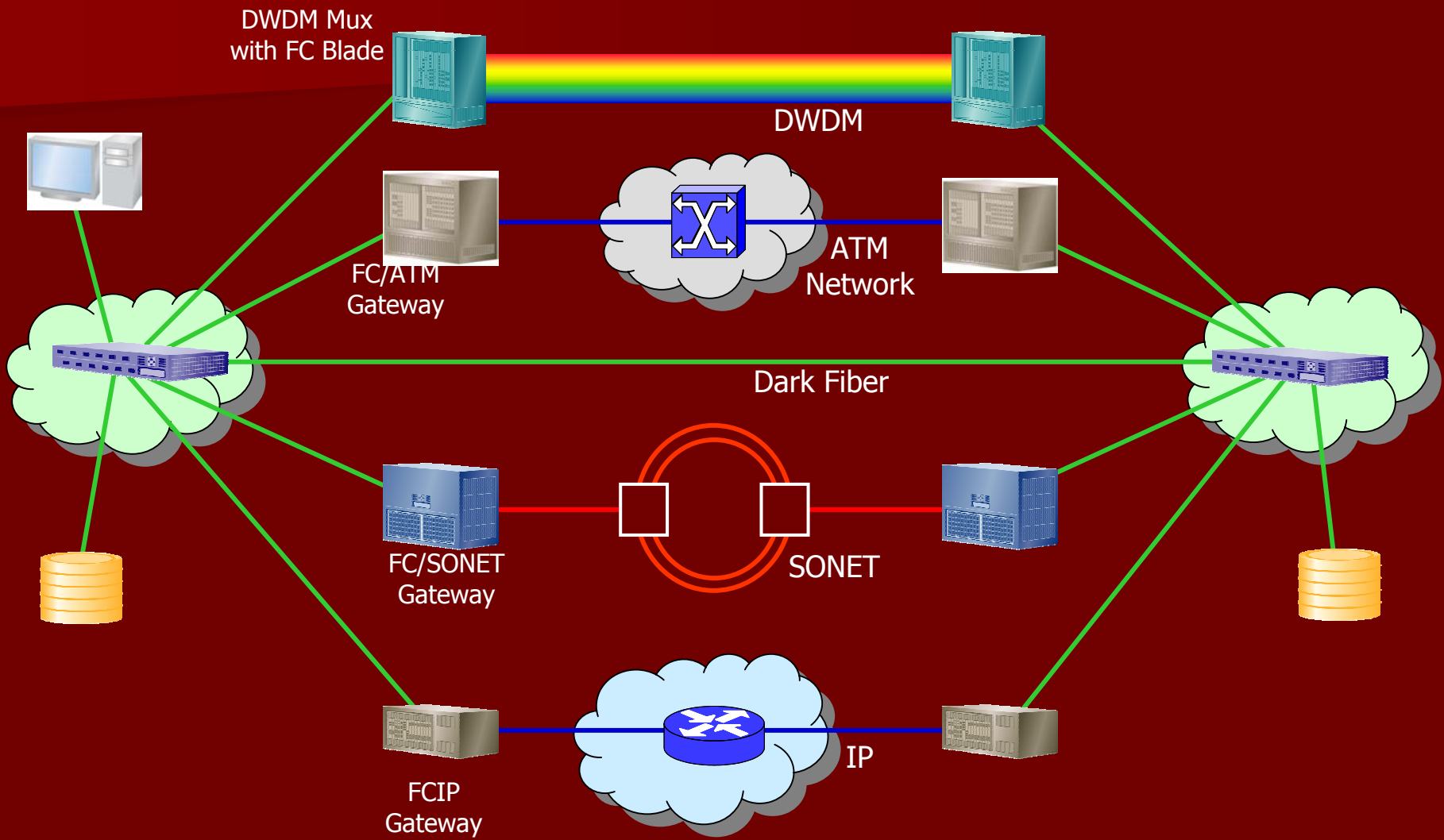
Utilizing the full bandwidth of this interconnection the data could have been moved in just over 80 minutes!

Transport Layer Extension

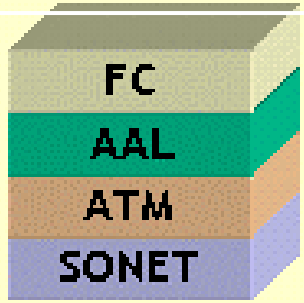
How do I move the signal?

- FC over Dark Fiber
- FC over DWDM
- FC over ATM
- FC over SONET
- FC over IP

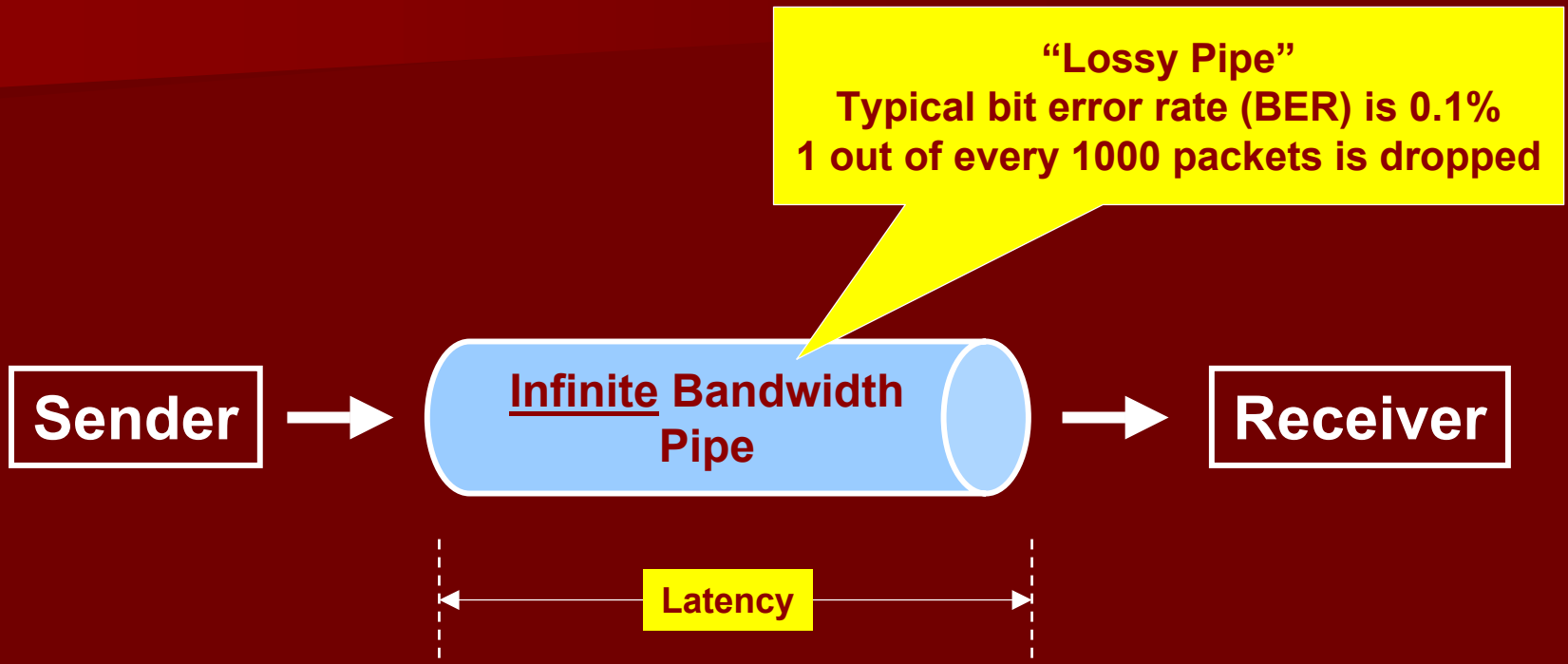
Transport Layer Extension through the MAN and WAN



FC Data Encapsulation Efficiency

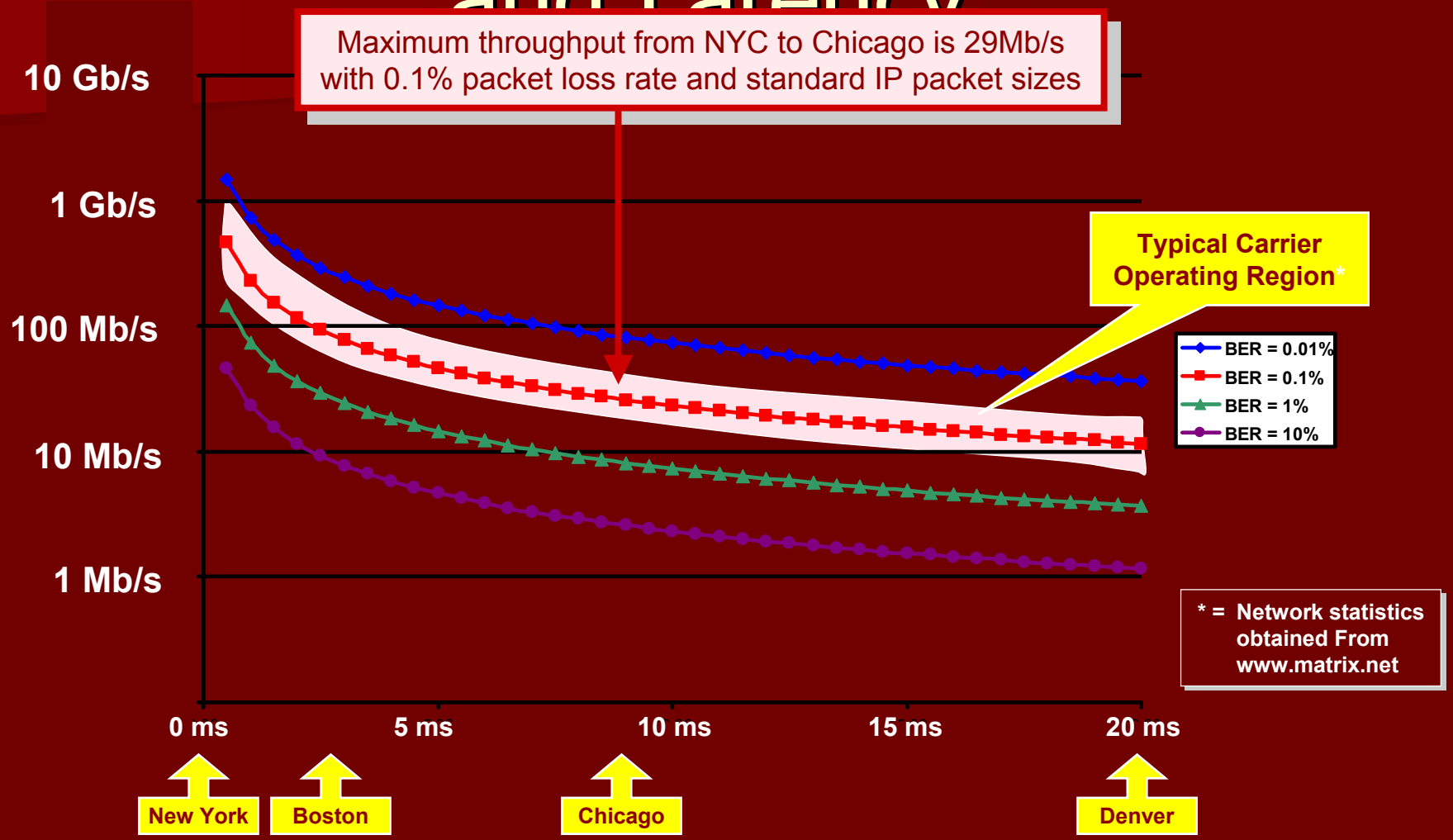
	Storage over IP	Storage over SONET	Storage over ATM
Mapping	 FCIP	 FC over SONET	 FC over ATM
Mapping Efficiency	90%	98%	80%

"Fat Pipes" Don't Guarantee High Throughput with TCP/IP



With TCP/IP over the WAN, BER and Latency affect throughput more than the bandwidth of the pipes

TCP/IP Performance vs. BER and Latency



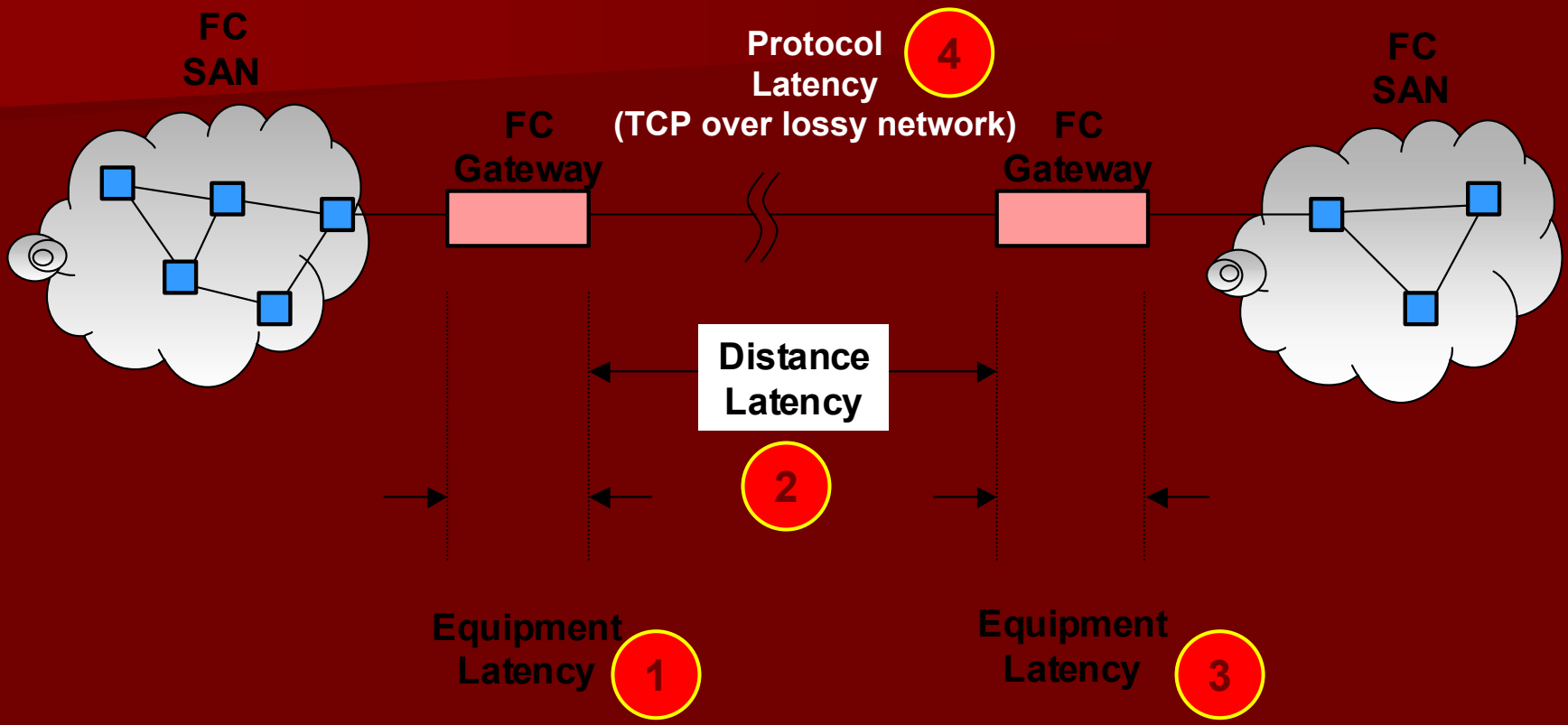
What's Wrong with Routed IP in the WAN?

- Routed IP is designed for scalability and connectivity
- Routed IP is NOT designed for performance
- TCP algorithms are designed to make the end user back-off quickly at the first sign of congestion
- Throughput is controlled by latency (geographical distance) and packet loss rate
- IP Carriers will always have packet loss
  - Sell service by peak capacity
  - Allocate equipment based on average capacity

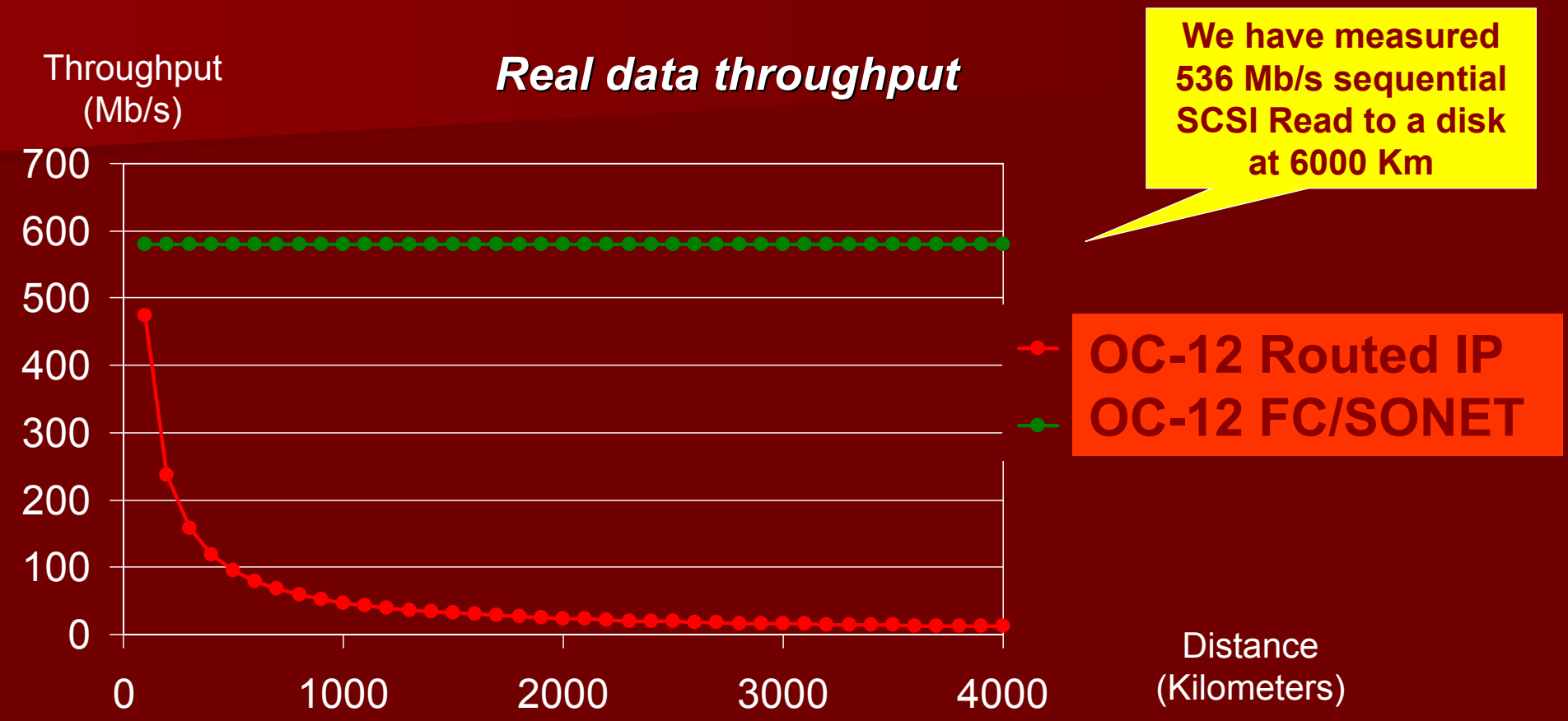
Network Layer Flow Control via Credit Buffering

- Fibre Channel maintains throughput in the data center by using flow control via buffer to buffer credits
  - Nominally FC switches provide credit buffering up to 10km distance
- Any wide bandwidth, long distance movement of FC data must couple flow control over the WAN
  - System requires end-to-end credit buffering
- WAN gateway flow Control options
  - Transparent mode - relies on the FC switches to extend the credits (good up to 130-150km) This method is used in all DWDM gateways and some SONET, FCIP devices
  - Coupled Credit mode – The gateway handles the buffer to buffer credits over the WAN. This is extensible up to great distances (1000's of km). This method is used in most FC over SONET gateways
  - FCIP gateways must provide credit buffering in conjunction with PAUSE and TCP/IP

Multiple Sources for Latency



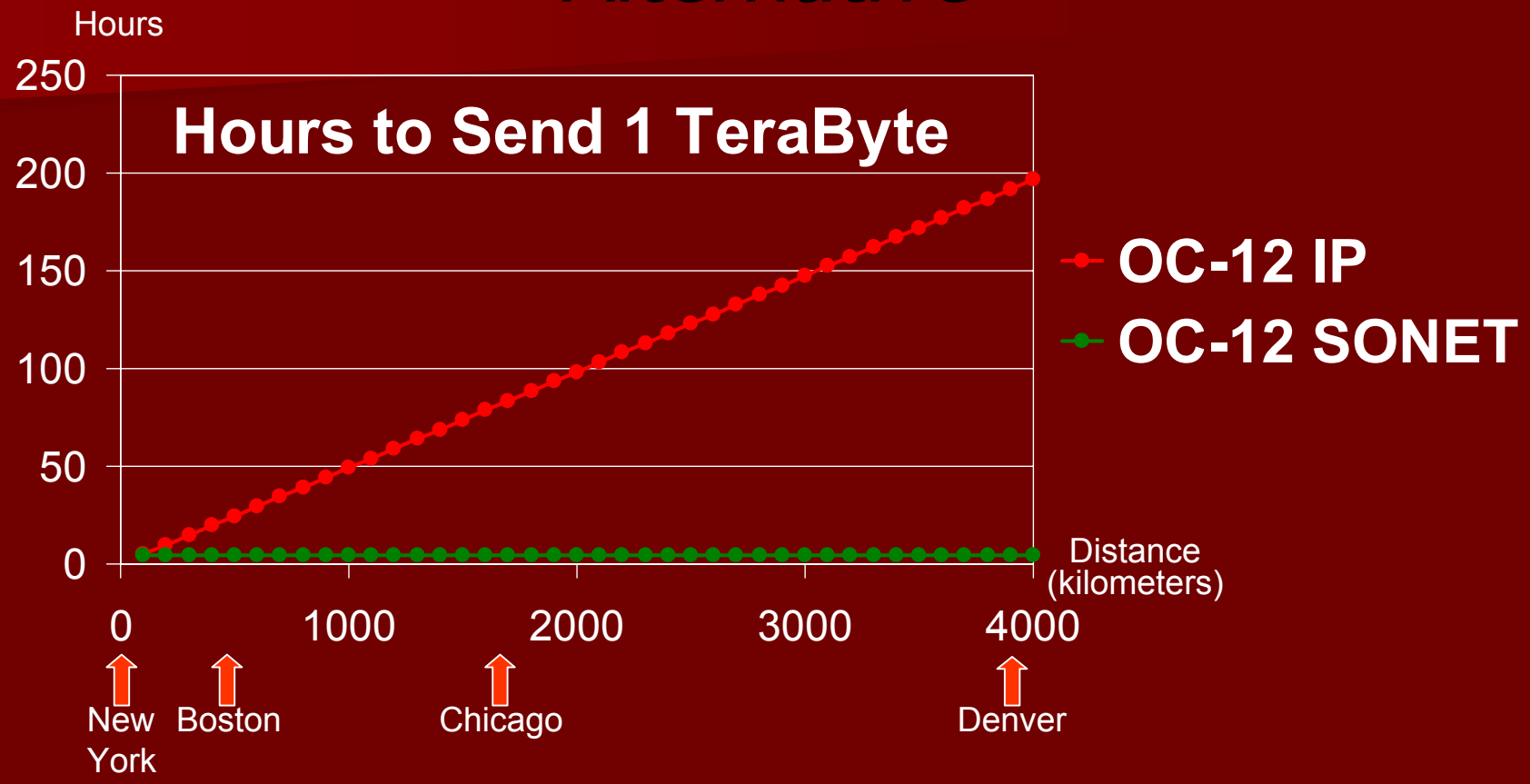
Fibre Channel Over SONET / WDM



Data re-transmission due to IP packet loss
limits actual IP throughput over

Fibre Channel over SONET

The High Efficiency, Long Distance Alternative



Data re-transmission due to IP packet loss
limits actual IP throughput over

Storage Over IP

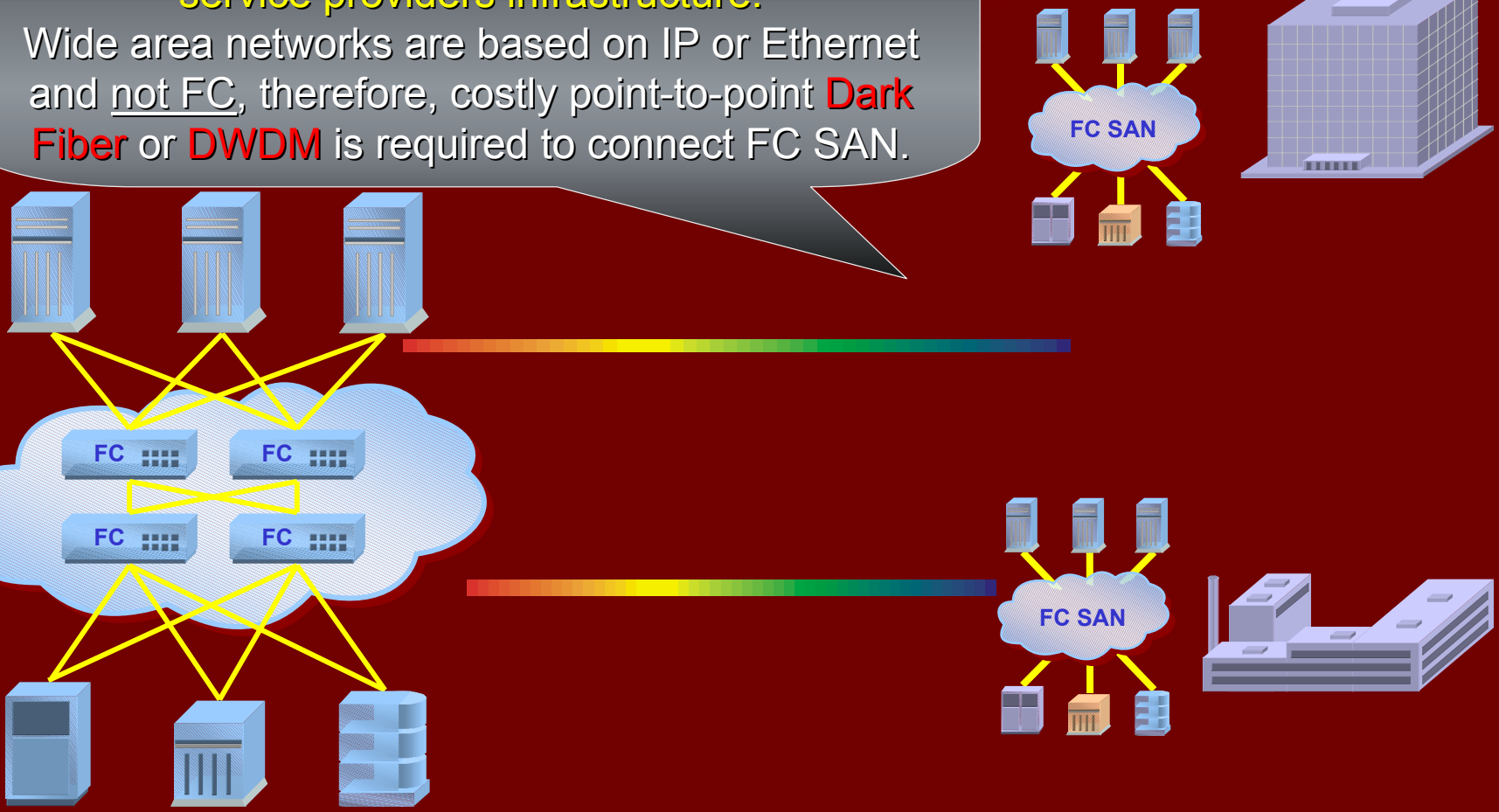
Using CNT Solutions

Joint Solution Overview

- Continuous Access XP
  - UltraNet Edge, UltraNet Storage Director
  - HP XP Disk Arrays (XP48/XP256/XP512/XP1024)
  - HP Continuous Access software
  - Cluster Extension XP and Continental Cluster
  - All inter-networking options – IP, ATM, T3, Metro Fiber optics
- StorageWorks DRM
  - Enterprise Modular Array (Enterprise Virtual Array)
  - DRM (Data Replication Manager) Software
  - All inter-networking options – IP, ATM, T3, Metro Fiber optics

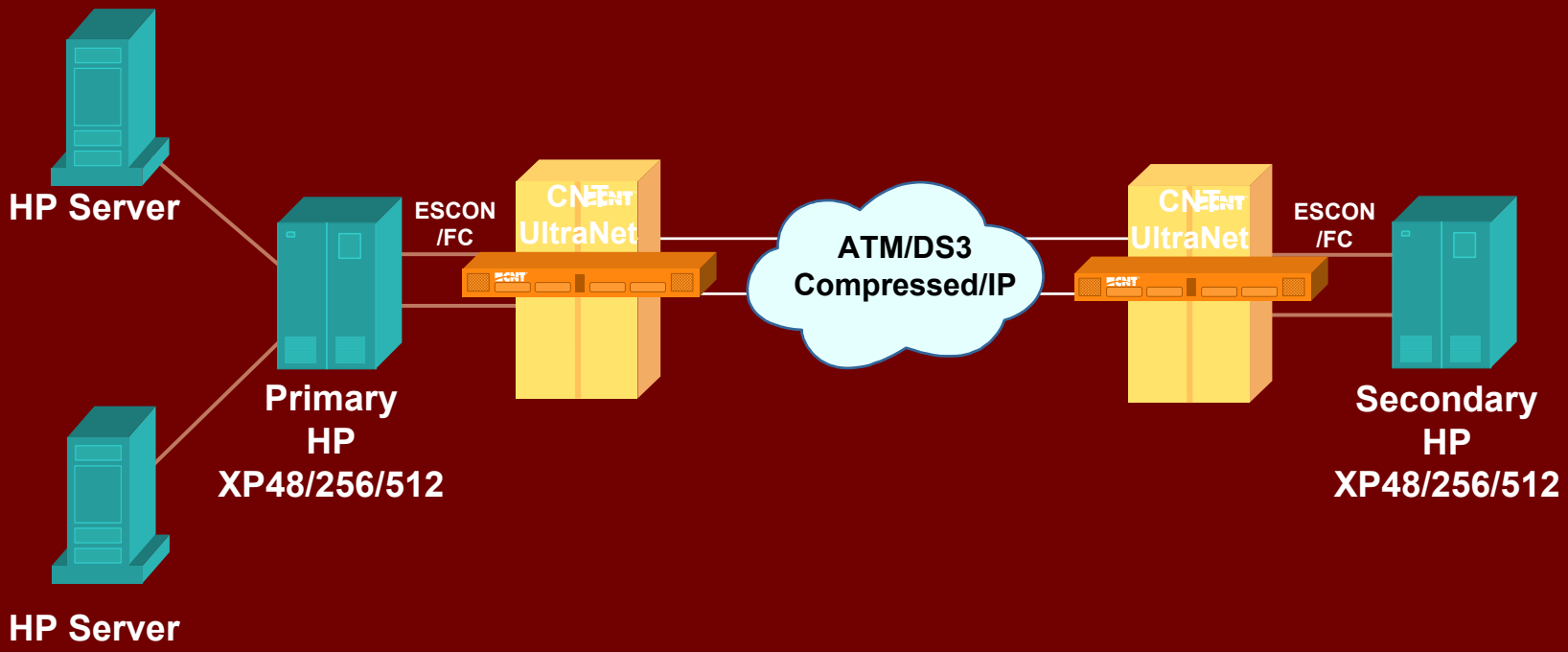
Fibre Channel SAN Issues - Extensibility

Native FC can not be transported over most service providers infrastructure.
Wide area networks are based on IP or Ethernet and not FC, therefore, costly point-to-point Dark Fiber or DWDM is required to connect FC SAN.



Fibre Channel SAN island

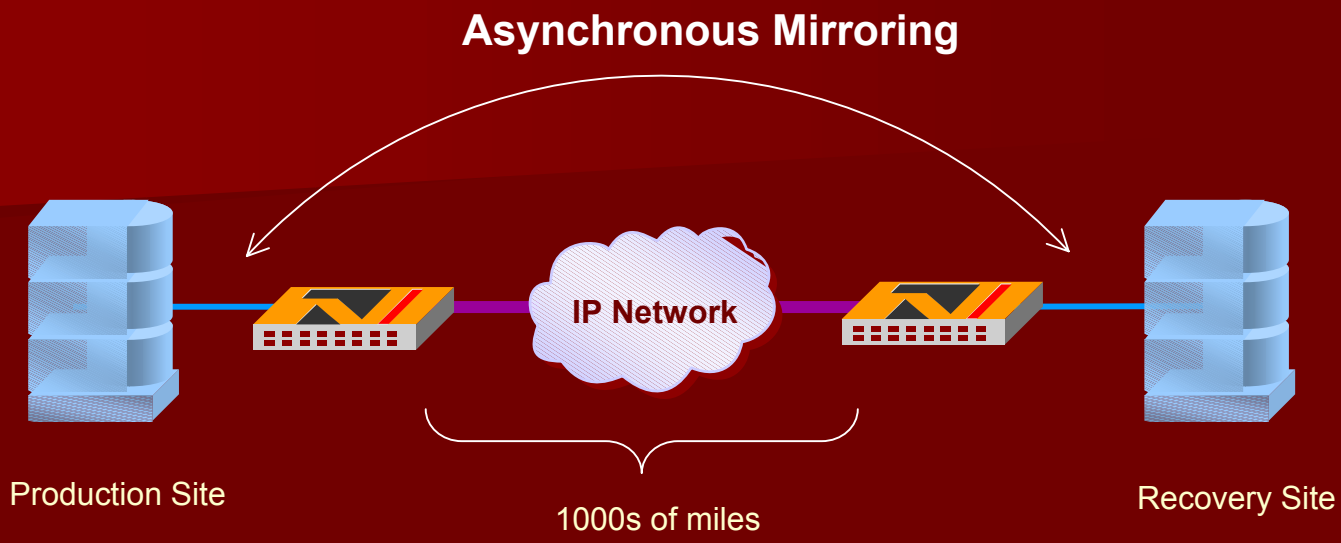
HP XP CA Remote Disk Mirroring Over WAN/IP



Storage Over IP

Using Nishan IP Storage Solutions

Asynchronous Mirroring – Going the

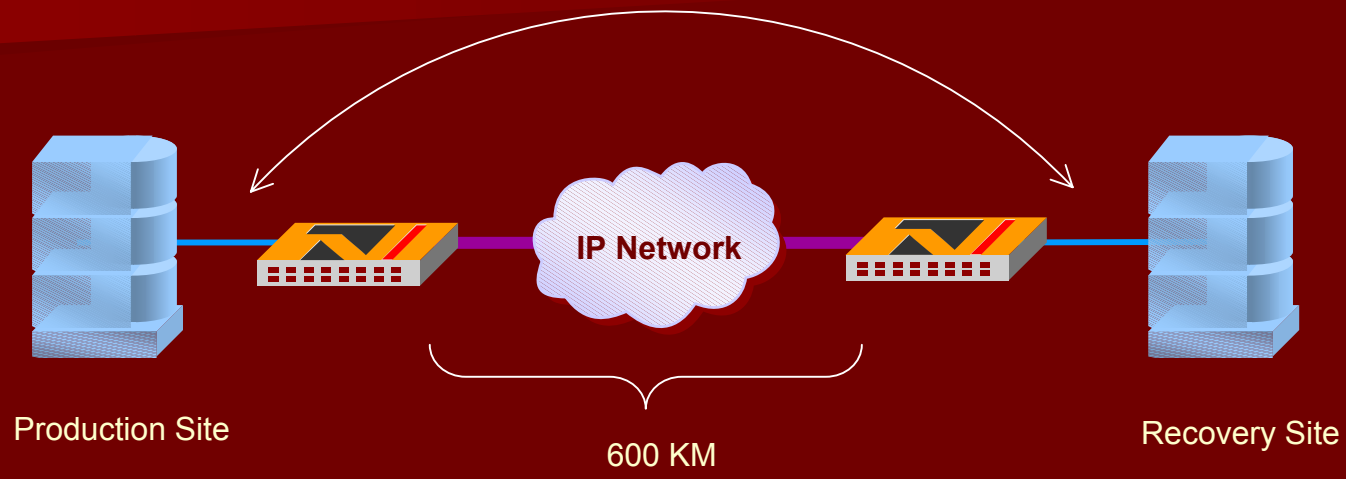


- **Customer required fast replication with 2 sites 2000 miles apart**
- **FASTWRITE technology excels at long-haul optimized performance**

Extending the Reach of Synchronous Replication

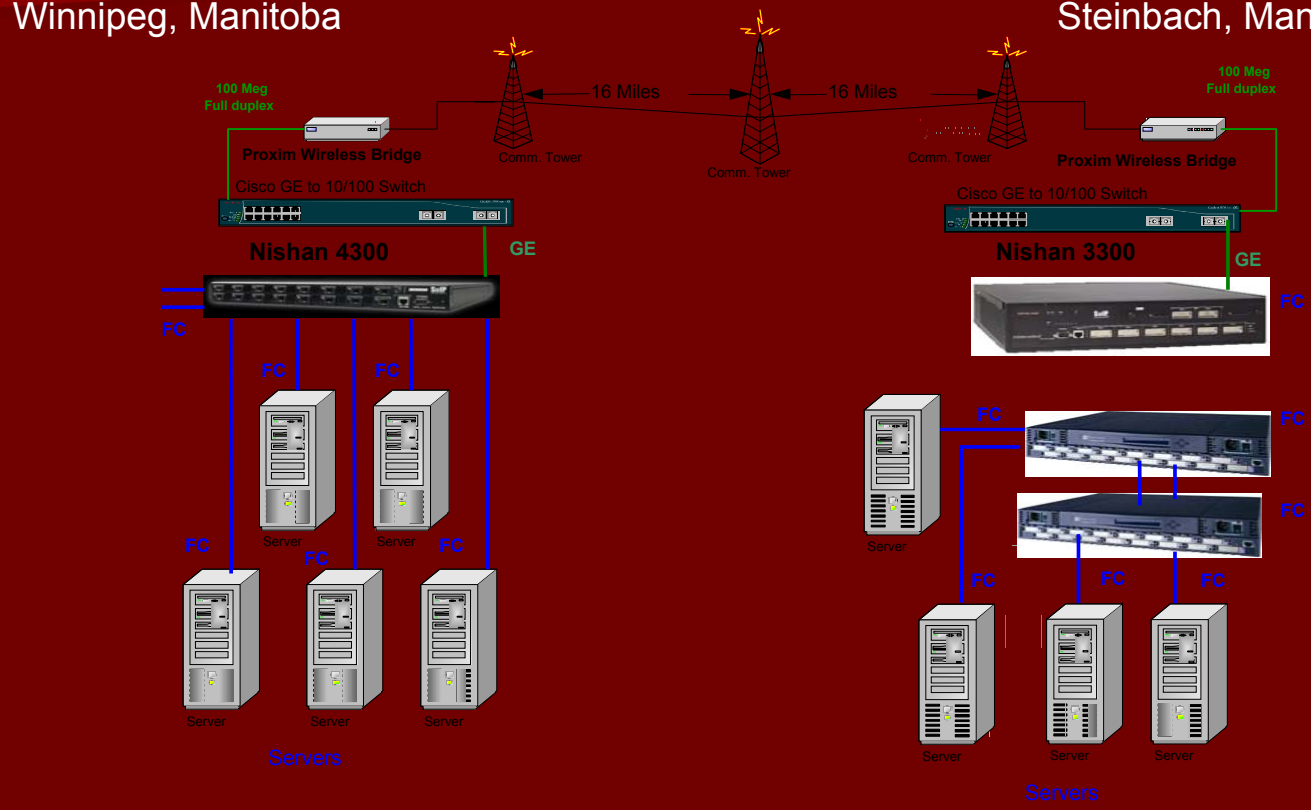
Vtesse Networks - Europe

Synchronous HDS TrueCopy

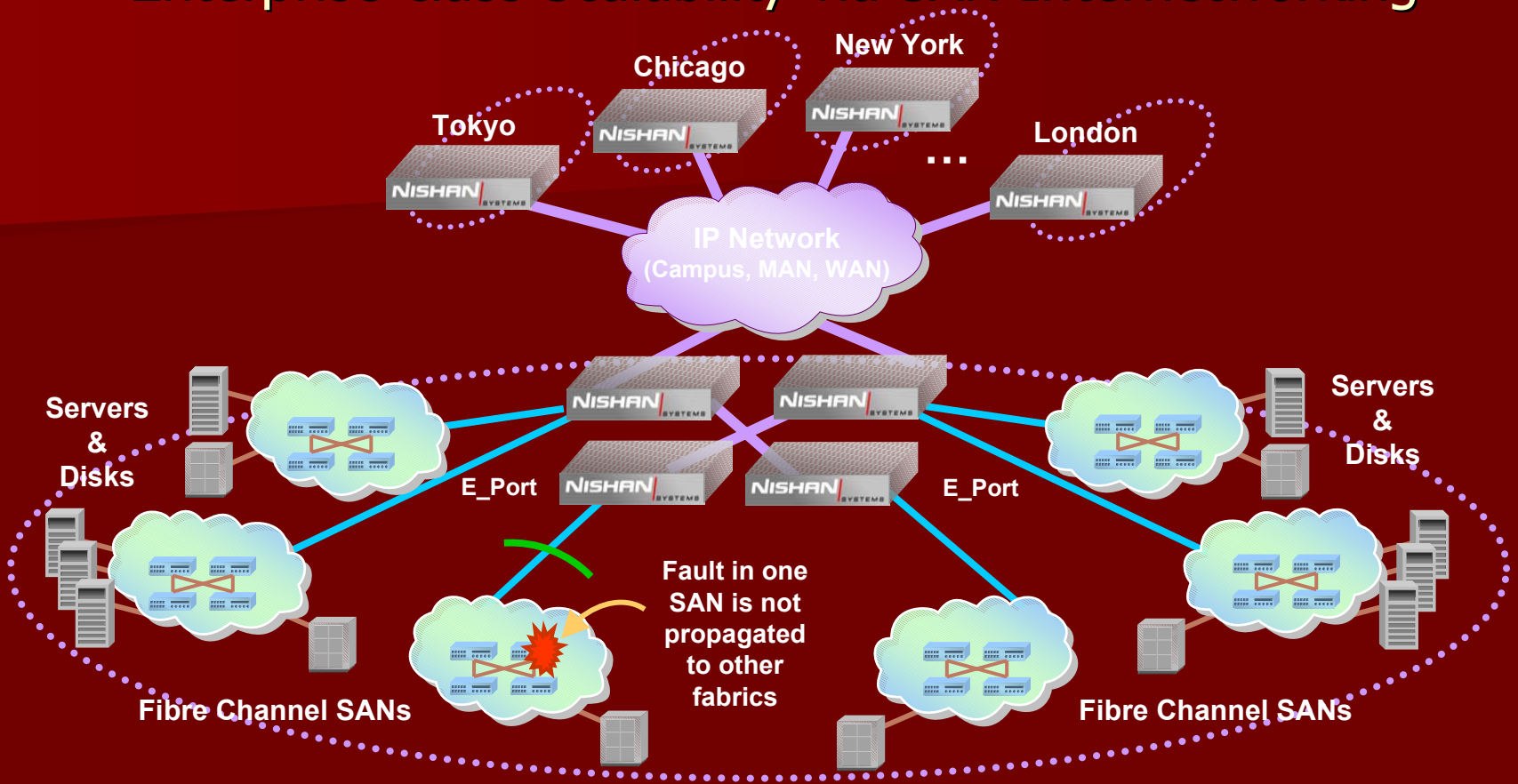


- **Nishan switches convert FC mirroring traffic to IP**
- **Synchronous data replication up to 600KM**
- **Expands the reach of Sync TrueCopy => more TC addressable market !!**

Steinbach Credit Union – Wireless IP SAN



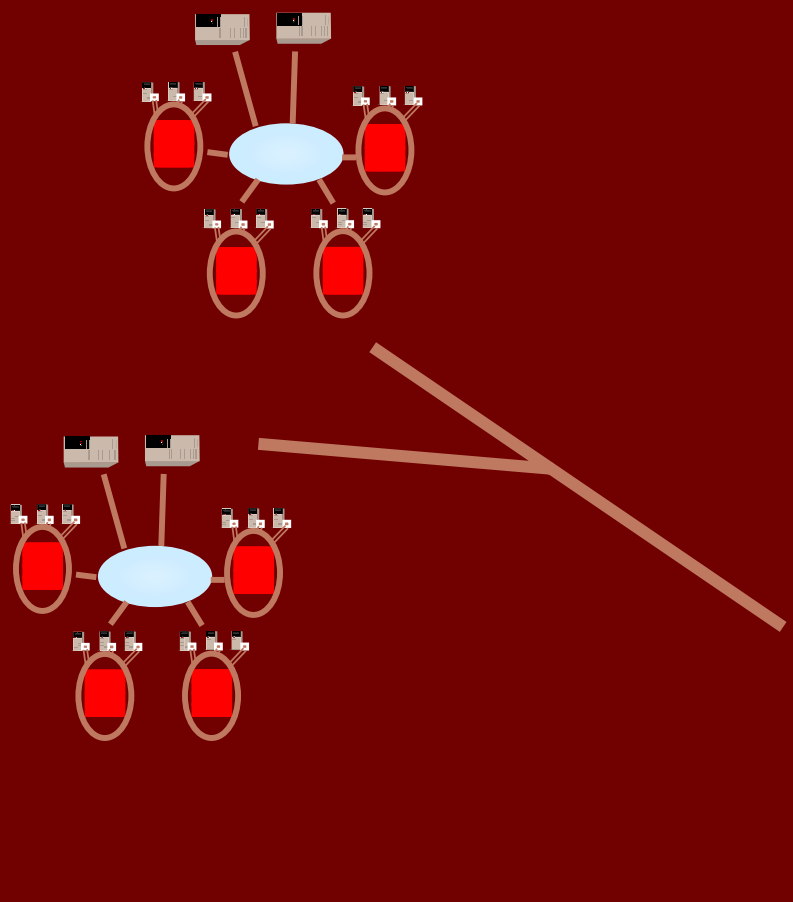
Enterprise-class Scalability via SAN Internetworking



- **Fault isolation provides router-like scalability – Enterprise-class scalability**
- **Connectivity between Heterogeneous Fibre Channel switches**
- **Extensibility across any distance: campus, metro or WAN**

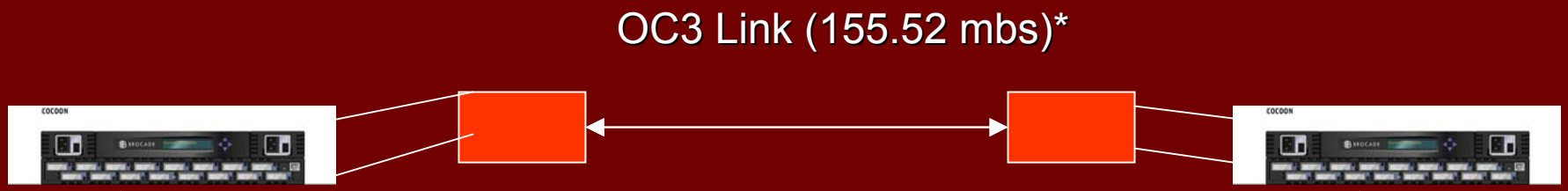
High Availability And Disaster Tolerant SAN Considerations

Storage WAN (SWAN)



- Features
  - Manage the WORLD as a single Entity
  - Wide area data sharing/migration
  - Storage Management/Reporting
  - Shadow/Remote Backup
  - Wide area DT
  - Connectivity within standard infrastructure (Network)
  - Span: The World

Fabric Extension



Extended Buffer Credits

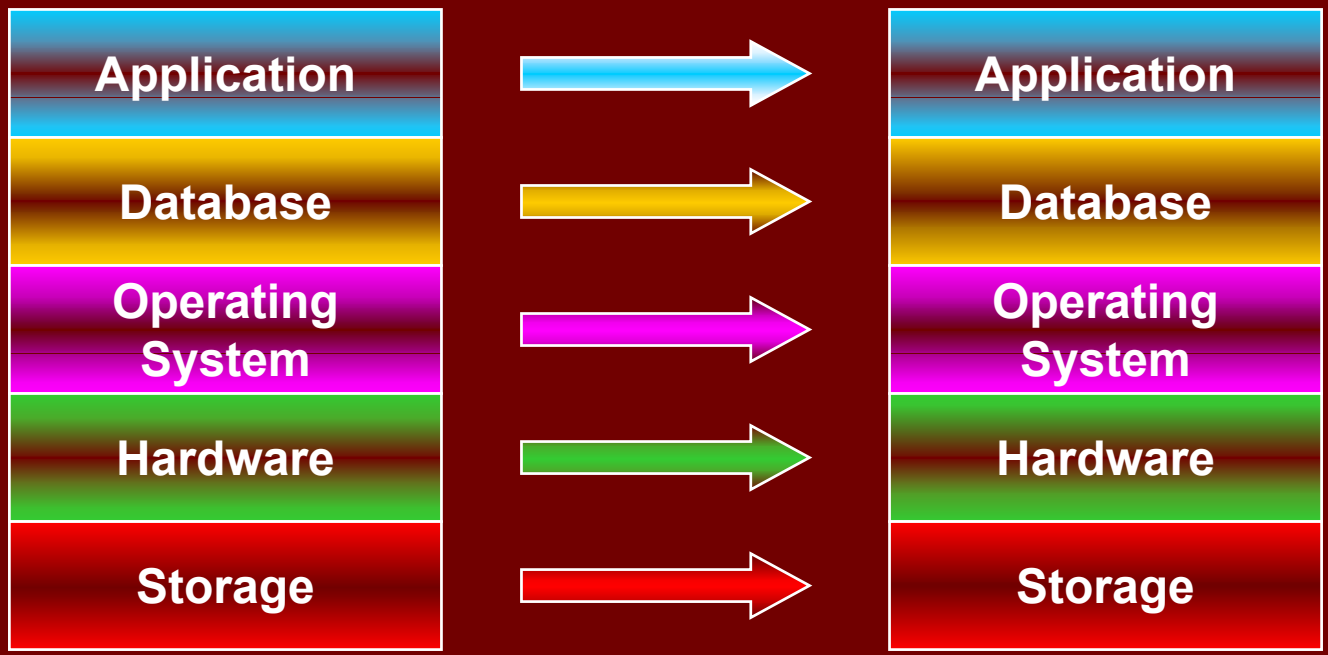
Extended Buffer Credits

*Used to specify the speed of fiber optic networks. The base rate (OC-1) is 51.84 Mbps. OC-2 runs at twice the base rate, OC-3 at three times the base rate (155.52 Mbps), etc. Planned rates are: OC-1, OC-3, OC-12 (622.08 Mpbs), OC-24 (1.244 Gbps), and OC-48 (2.488 Gbps

Data Replication

- Replication can be done at many levels

- Replication can be done at many levels



•Real Time COPY

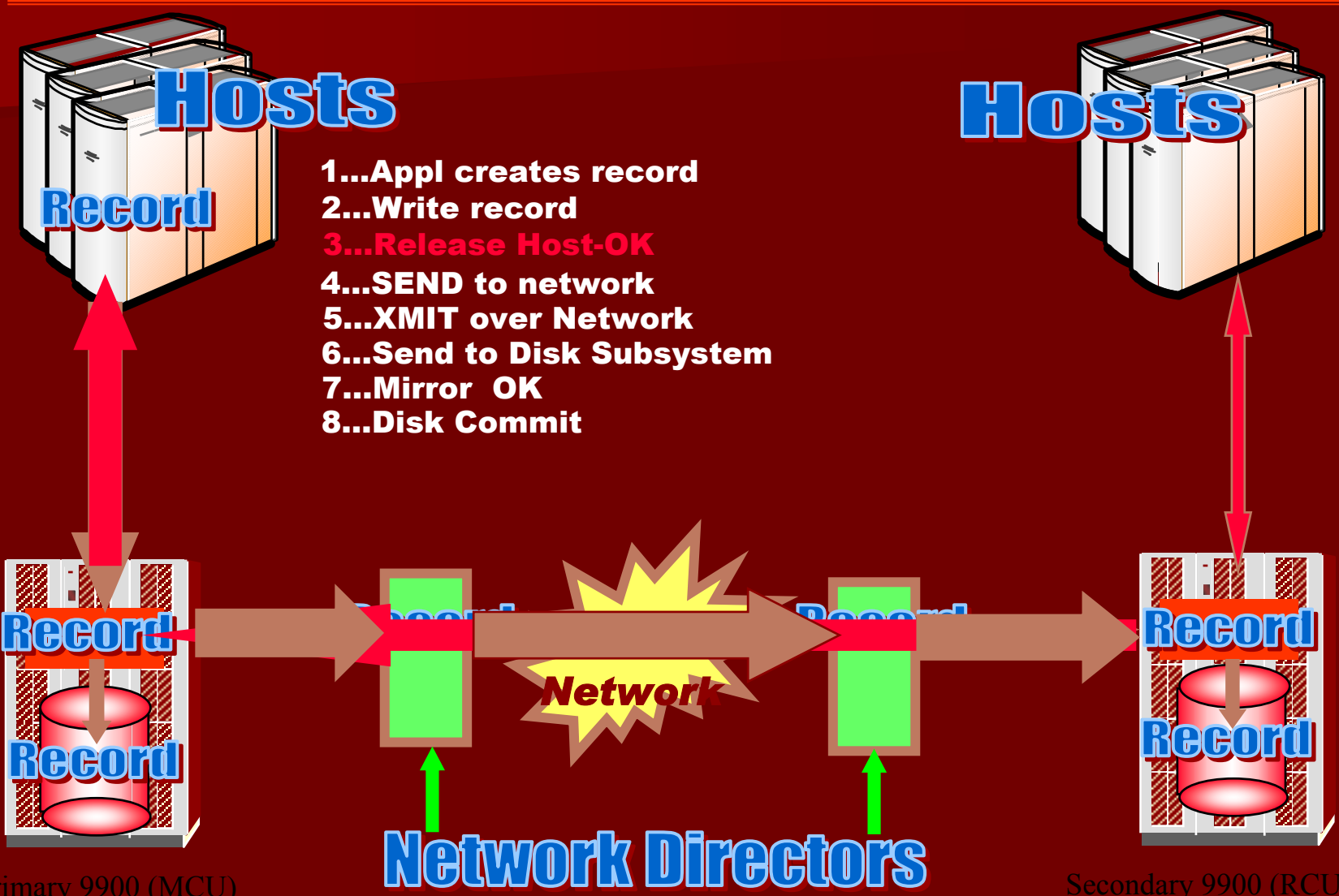
- Provide Disaster Recovery
- **NOT to maintain two identical copies**
- Provide I/O consistent copy of data

Synchronous Or Asynchronous

- If Within ***Supported*** Distance Use Synchronous Because:
  - Data is more secure
  - Best overall performance
- Asynchronous is supported
  - Cases w/low I/O rates and some potentially lost data is acceptable if links are broken

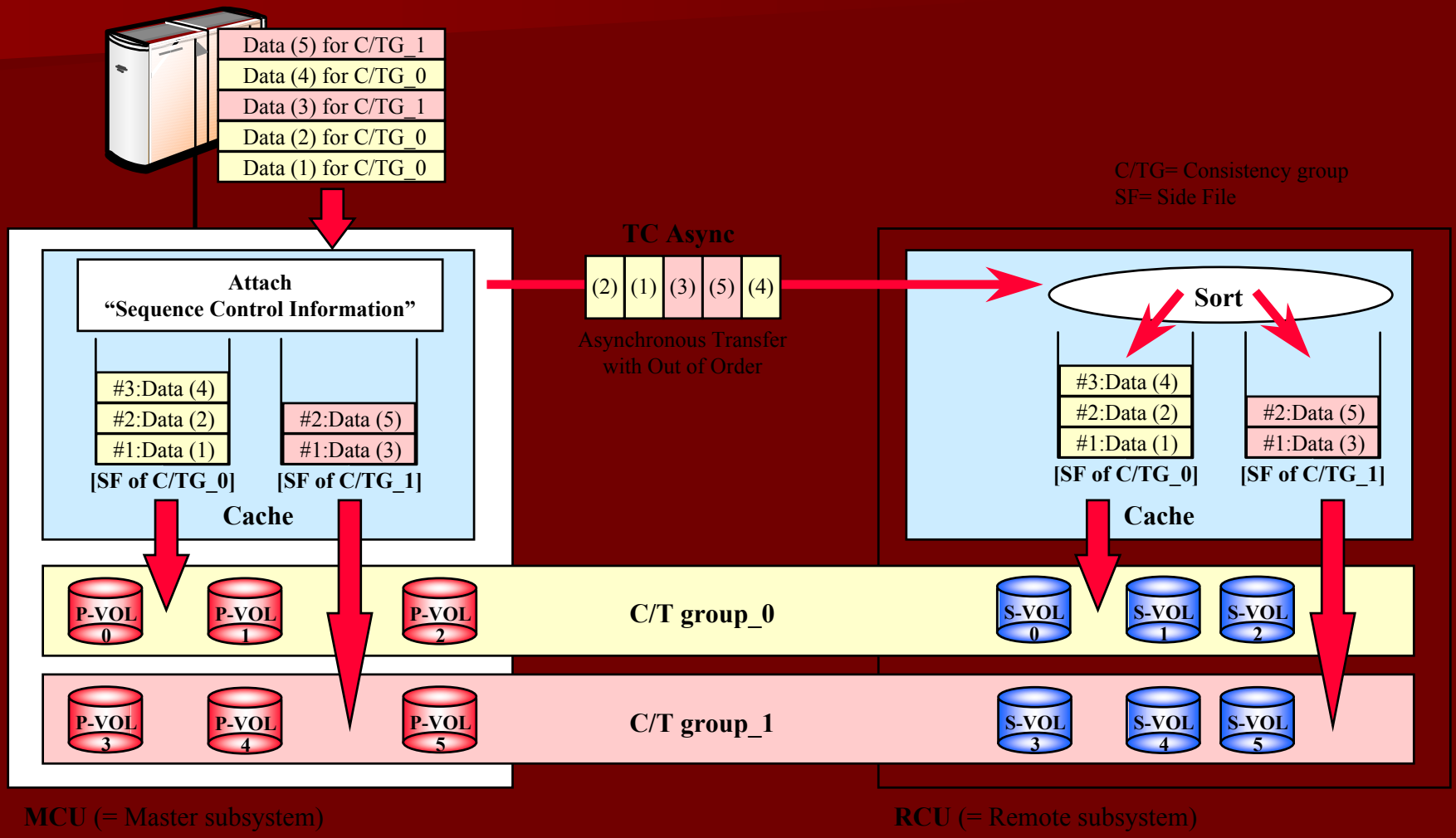
Asynchronous

Operations



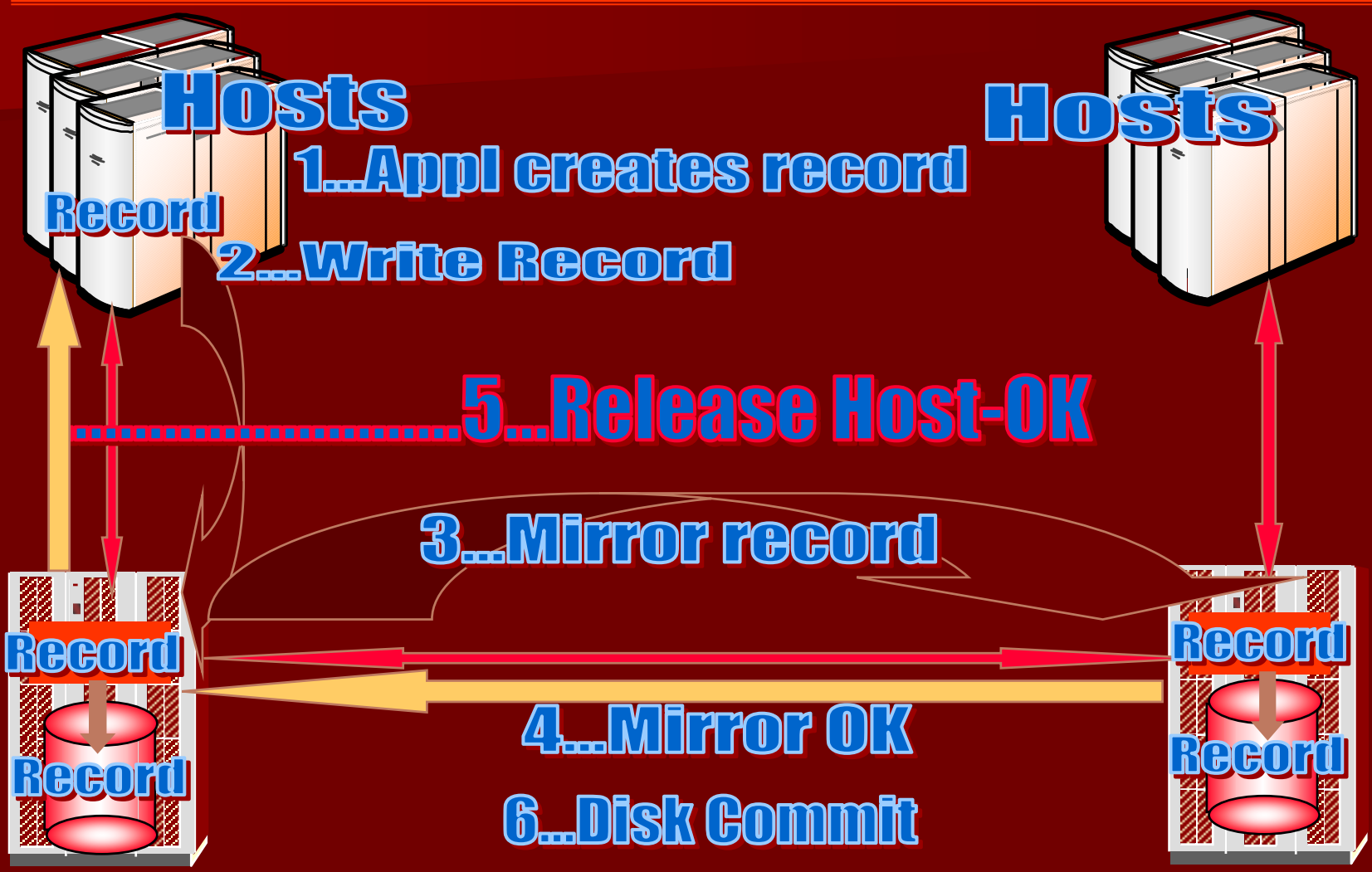
Asynchronous

Update Sequence

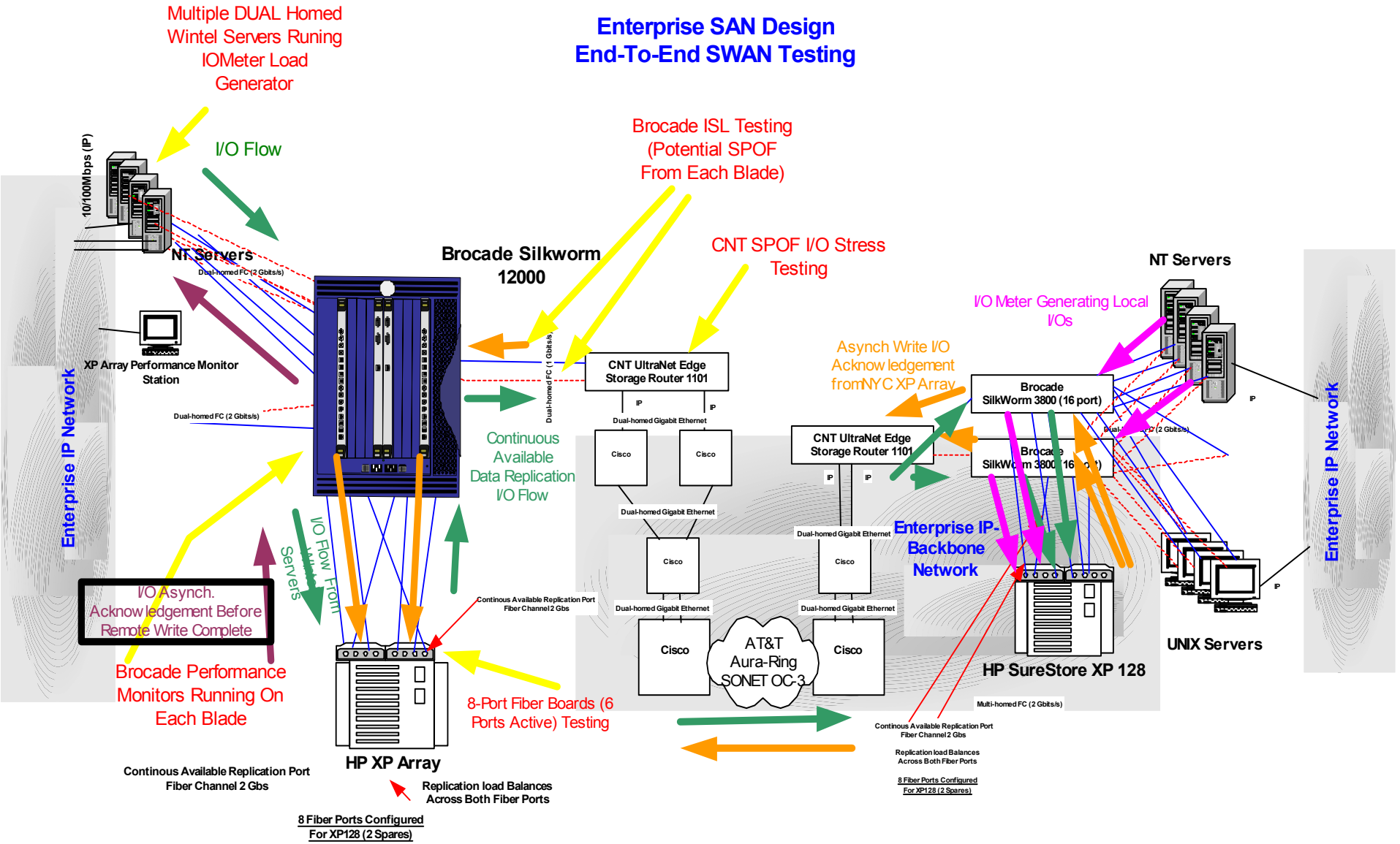


Synchronous

Operations



Enterprise SAN Design
End-To-End SWAN Testing



NOTES

1. SilkWorm 12000 comes with dual control units, multiple power-supplies and fans, and has two 16 port switches configured as separate fabrics.
2. CNT UltraNet Edge Storage Router 1101 are used exclusively for the communication between HP XP 128 SAN's.
3. Based on the design and the backbone bandwidth limitation, only Asynchronous communication between the XP 128 SAN's was recommended.
5. The design will work (in Async. mode) with existing Cisco's Enterprise IP-backbone without using QoS, Packet Prioritization, Traffic Shaping, or Layer 3 switching technology.

HP SAN's Virtualization Concepts (Differences and Considerations)

Virtualization Technology

Virtualization: Storage

- **Enables vast amount of physical capacity scattered across the enterprise to appear as a single large pool of data for presentation to application servers.**
  - **Storage pool is a single manageable entity**
  - **Virtualization can occur at any of three levels.**
  - **All levels can be used together, or independently, to maximize the benefits to customers:**
    - ***Server level*: Ideal for small entry level needs**
    - ***SAN Fabric level*: SAN-wide virtualization**
      - **increase efficiency in SAN development, management, and service**
    - ***Storage System level:* Ideal for large volumes without sacrificing performance or reliability**

Why is Virtualization Important Anyway ?

- Virtualization will:
  - Provide investment protection for non-virtualized arrays
  - Provide a Common Storage Pool that will allow for greater efficiency and not a load of stranded capacity in servers or on SAN RAID volumes/arrays
  - Data Migration that is transparent to the OS for upgrades, changes, etc to dataset for applications
  - Reduction of ownership costs

EVA Virtualization Primer

- Principles of EVA virtualization
  - "Disk blocks are disk blocks"
    - All disk blocks are equivalent*
  - Foundation object is the "Virtual Disk"
    - A SCSI LUN
      - Attributes
      - Data
    - It doesn't matter where or how data actually stored*

EVA Virtualization Primer (con't)

- Data location independence
  - Makes new benefits possible
  - BUT - to realize those benefits, we must change some old habits
    - Backend disk access pattern no longer correlated with host access pattern
    - Stop using configuration as "organizing" tool
      - Use EVM capabilities instead
    - Resist temptation to micromanage
    - Avoid temptation to utilize every last byte

Significantly Higher "Utilization" of Purchased Capacity

- **Up to twice the typical 40-50% Open Systems utilization . . . Based upon dynamic pool/LUN expansion, etc.**
- **Importance to Business/IT:**
  - **Customer minimizes purchase of unusable capacity**
  - **Just-In-Time capacity increments for application growth**
    - **Even to the point of adding one disk-at-a-time**
    - **Dynamic Pool/LUN Expansion (w/Server support)**
  - **No "droop" effect in performance**
  - **No intensive storage administration "gyrations"**

- **BOTTOM LINE:**

  Much lower "effective" price/MB . . . Easily "justifiable"

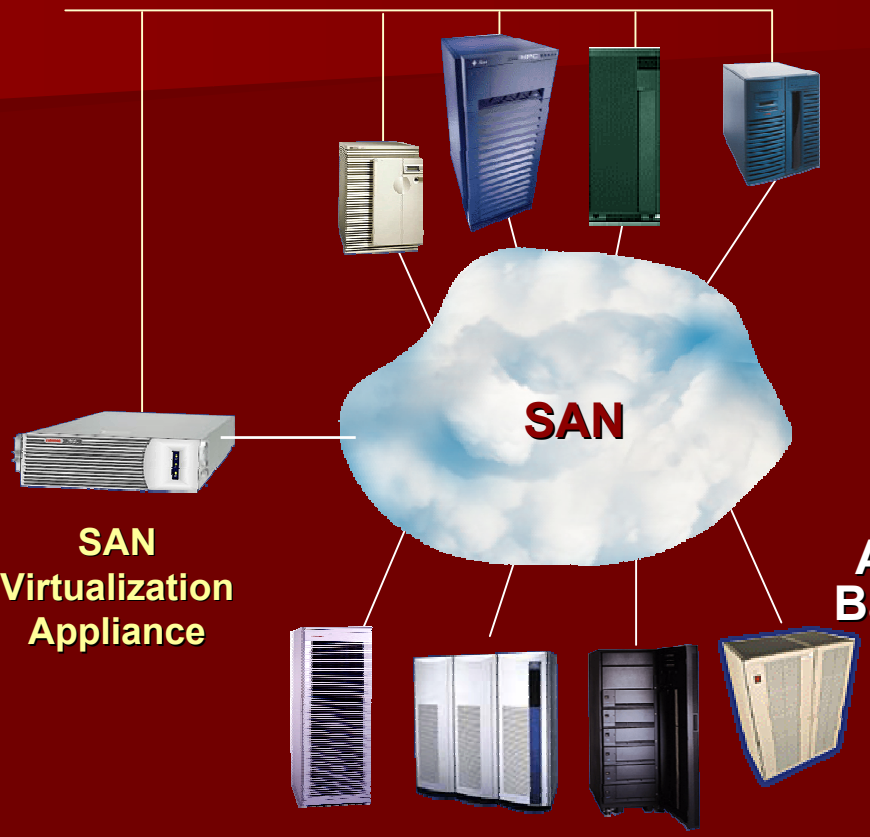
StorageWorks Enterprise Virtual Array

- HSV110 Array Controller Virtualization:
  - All raw storage is pooled
  - Virtual Disks are drawn from a pool
  - Virtual Disks managed by customer to these constraints:
    - Size range 1GB - 2TB, in 1GB increments
    - Up to 256 Virtual Disks - selectively presented to hosts
    - Each Virtual Disk can have no, medium or high redundancy
      - RAID V0, V5, V1
    - A Virtual Disk's size can be dynamically expanded, but not shrunk

What is StorageWorks Virtualization?

- Virtualization will bring new terms with it
  - Aggregation
  - Fine Grain Mapping
  - Distributed RAID
  - Capacity Free Snapshots
  - Storage Pools (Local & Global)

Strategic Storage Virtualization Views



Server Level

Local Storage Pools; Direct Access to Virtual Disks; Network Mapped Disks Served to other Clients; Local Snapshot

Fabric Level

Virtualization Within Switch

SAN Level

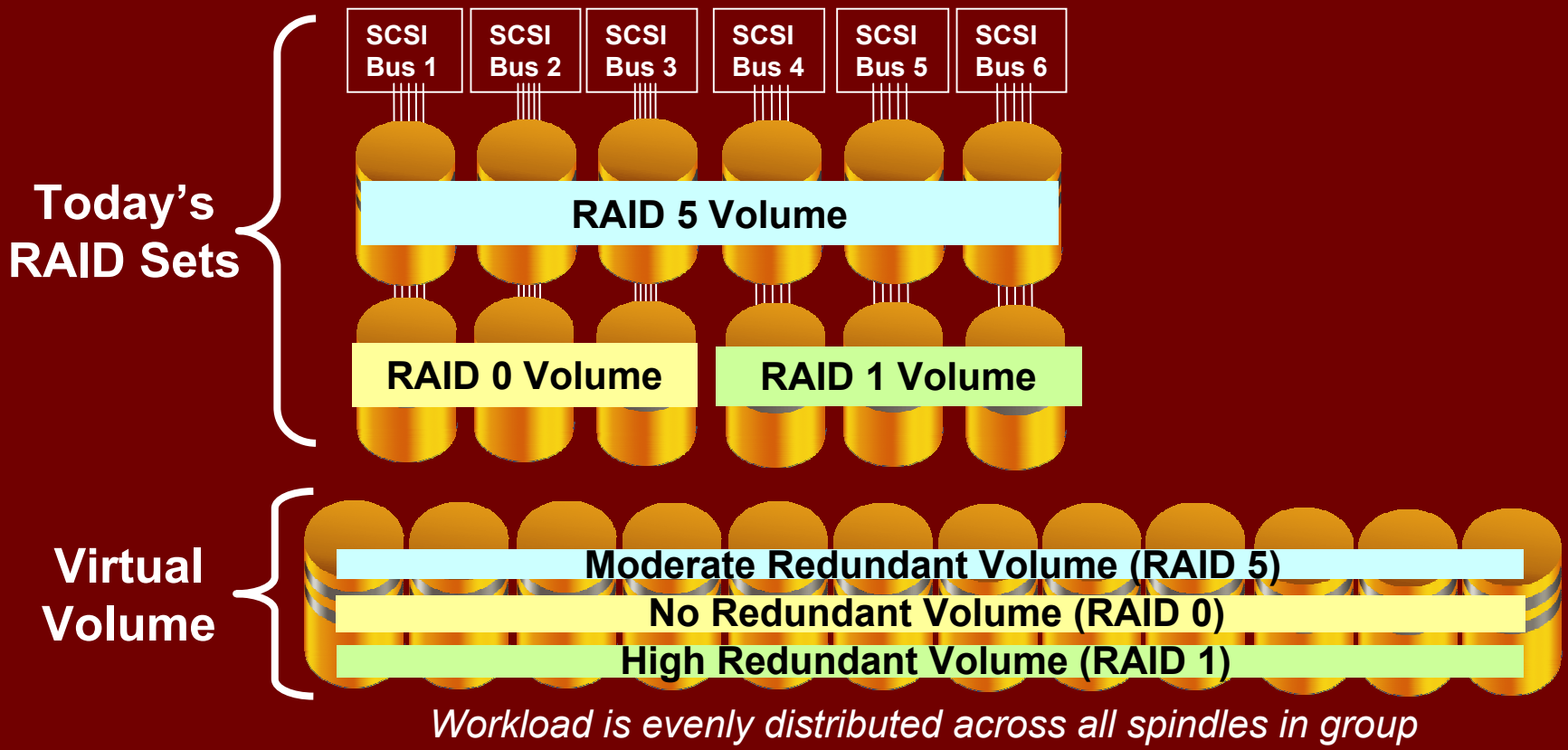
Global Storage Pools; Direct Access to Virtual Disks; Attribute Based Storage; Multiple Models for data delivery: Asymmetric & Symmetric: Builds upon RAID Volumes on SAN ;

Storage Level

Large Local Pools; Distributed RAID; Ultra High Performance

Virtualization

- Eliminate throughput bottlenecks
- Eliminate load balancing procedures for application and Data Base

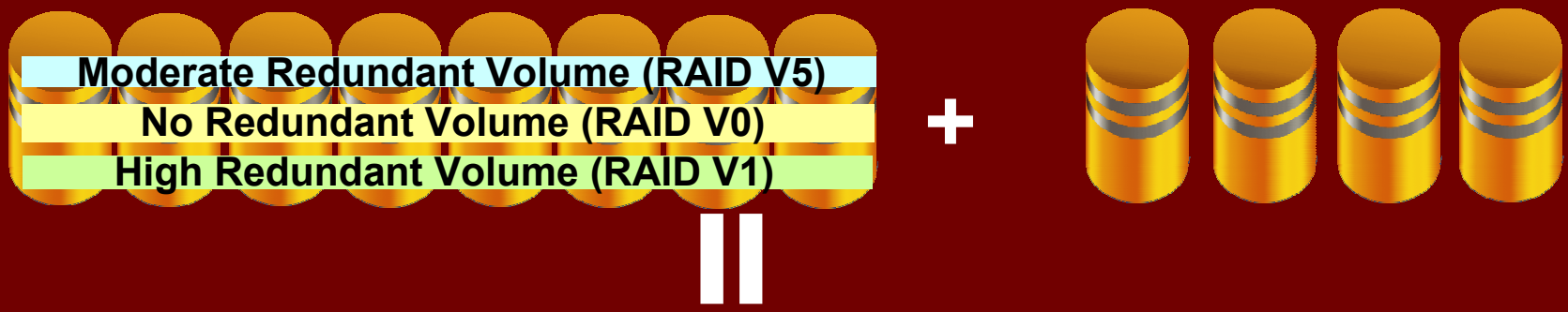


Workload is evenly distributed across all spindles in group

Virtual Storage Pools

- Dynamic pool capacity changes
- Pool capacity can be expanded by spindle(s)
- Virtual Disk blocks are automatically relocated to level spindle use

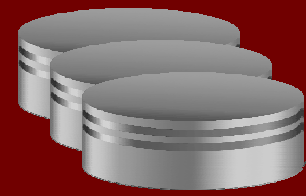
- **Disk Spindles Becoming a Throughput Bottleneck**
- **No Additional Storage Space Available**

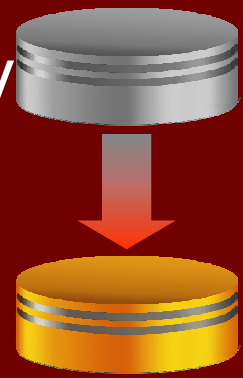


- **Disks Running at Optimum Throughput (dynamic load balancing)**
- **Additional Storage Space Available**

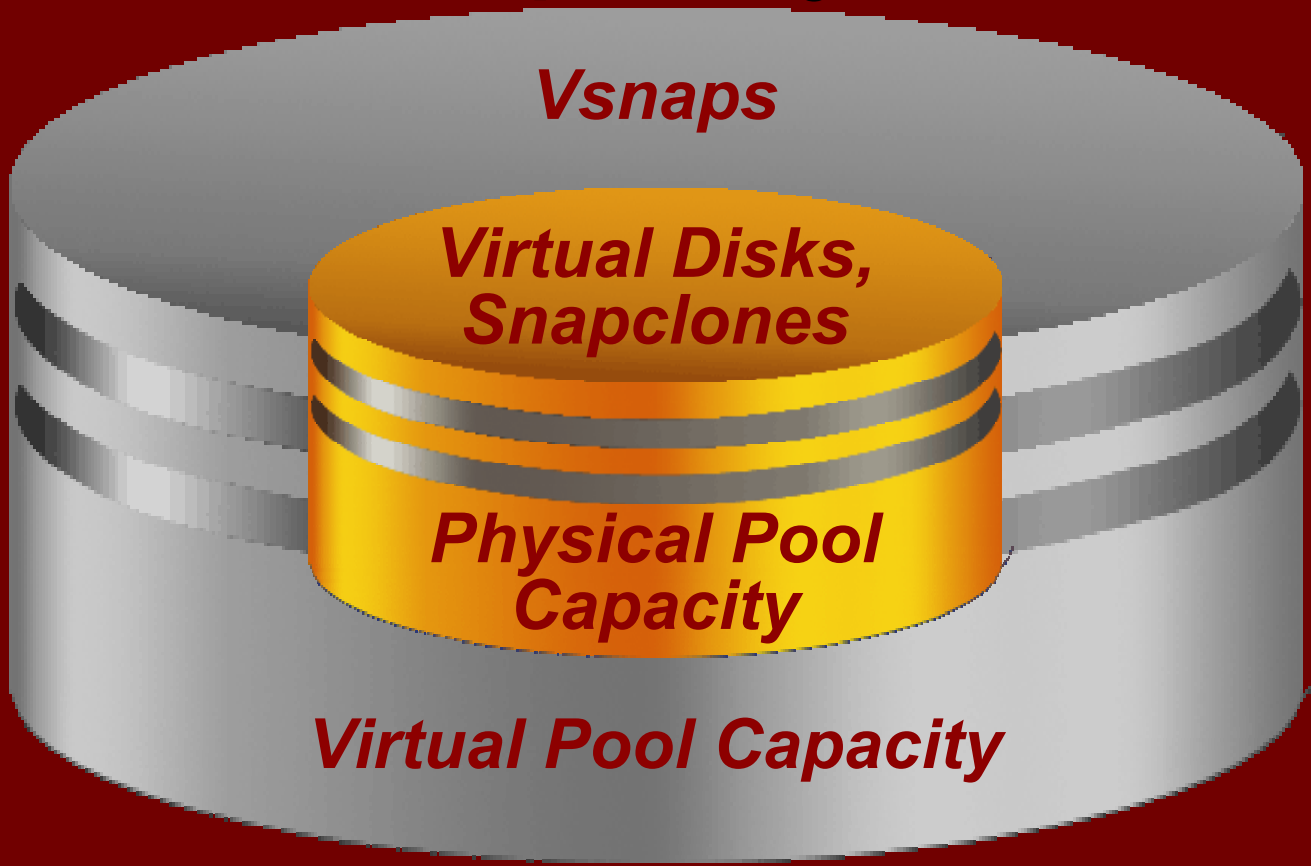
Vsnaps and Snapclones

- Double effective capacity
- Stretch virtual pool capacity for free
- Simplify data mining operations, reduce application development backlog
- Eliminate backup windows, reduce recovery time to raise service levels
- Make ad hoc point-in-time copies with minimal effort to maximize IT flexibility
- Provide administrators with powerful tools to multiply management efficiency



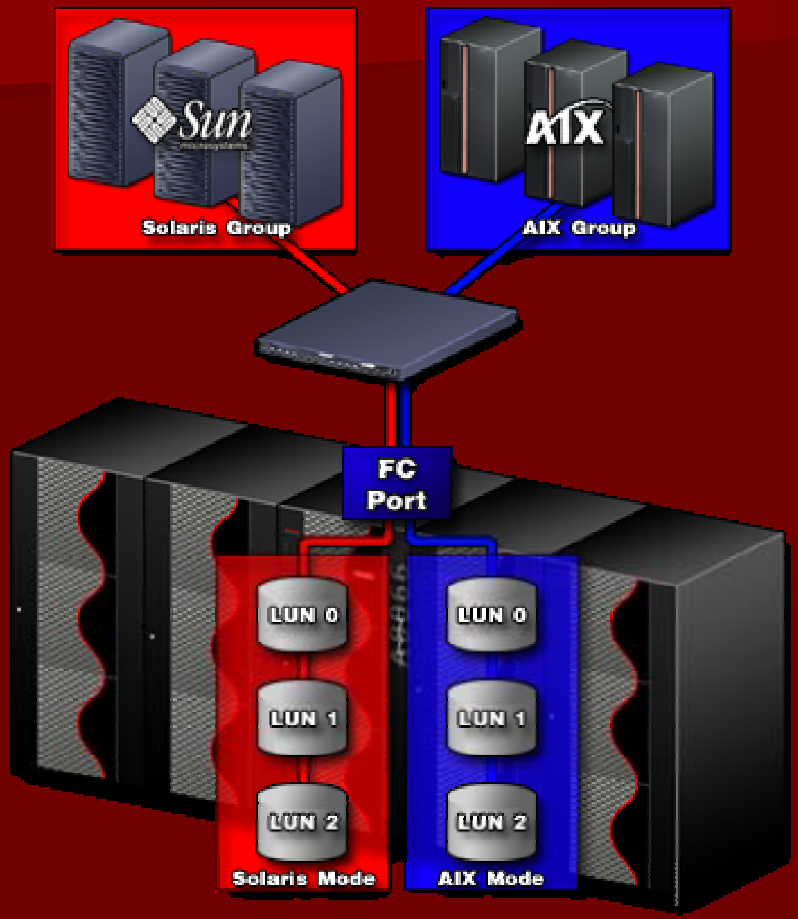


Stretching Capacity with *Capacity-Free* Vsnaps



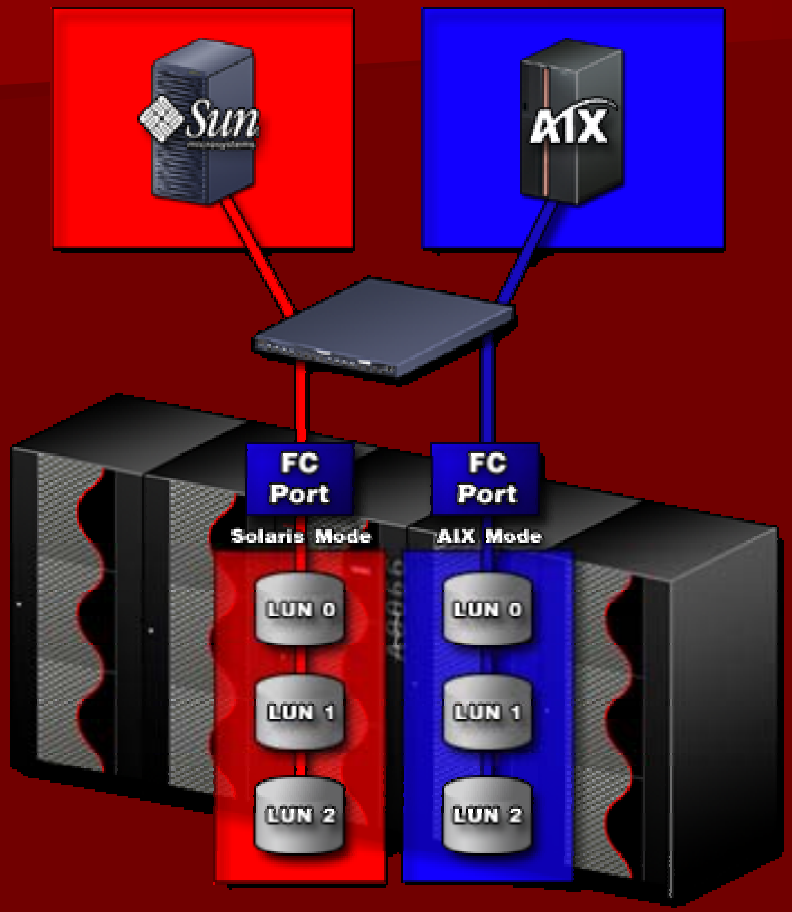
Virtually Boundless Storage

Heterogeneous Connections – The Problem



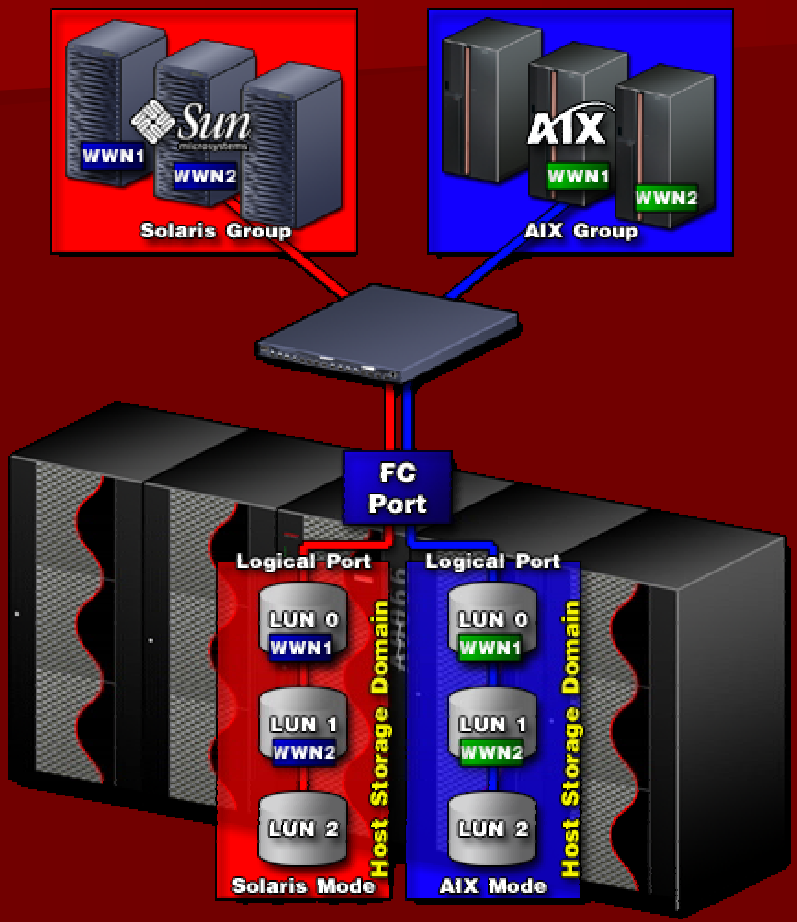
- The Problem
  - "Mode Set" determines port behavior.
  - Different mode sets for Solaris, AIX, NT, etc.
  - Port can be set in only one mode
  - Sharing same port between different host platform types creates problems.

Heterogeneous Solutions Today



- The Solution
  - Must configure separate physical ports for each host platform type
- Consequences
  - Expense of extra ports
  - Max number of channel features, especially with FC, ESCON, FICON mix limits heterogeneous consolidation.

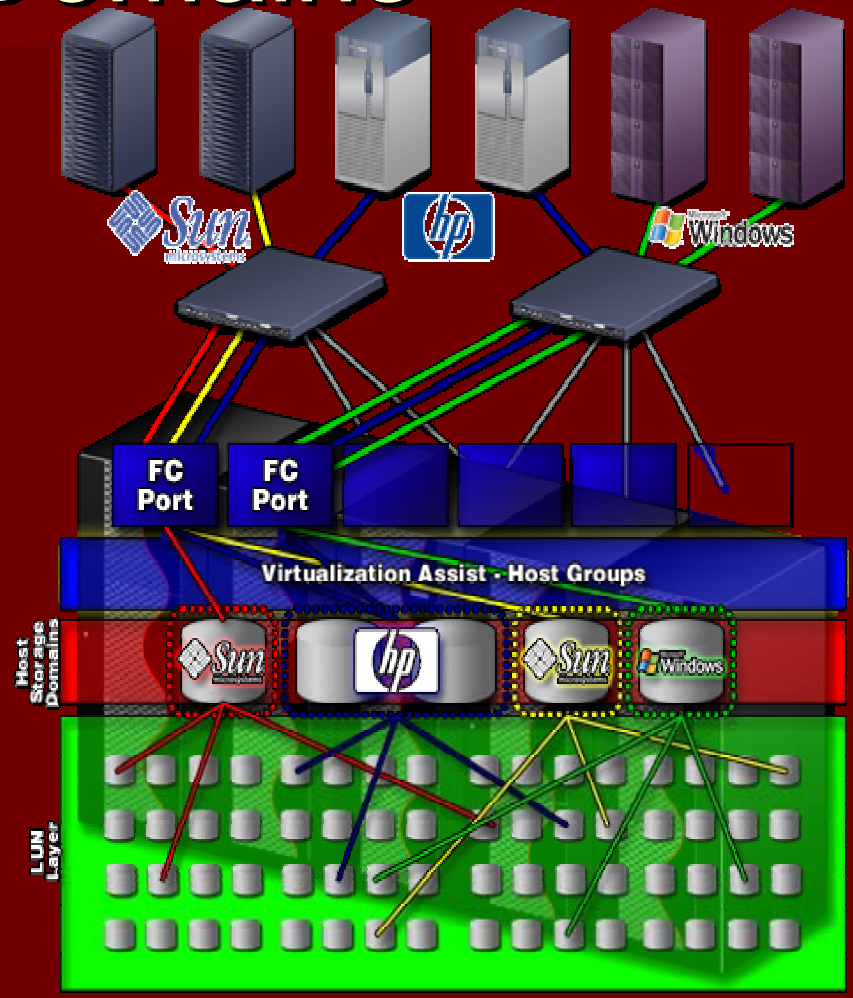
The HP XP Series Heterogeneous Solution



- Host Storage Domains (HSD)
  - Each HSD has its own logical (virtual) FC port.
  - FC port "mode set" applies to logical FC port in each HSD.
  - Each HSD has its own set of LUNs.
  - Hosts matched to their assigned HSD based upon WWN
  - Hosts can only "see" their own HSD

Virtualization Assist: Host Storage Domains

- Multiple Host Storage Domains can share same physical port.
- Each Host Storage Domain has its own logical FC port and its own independent set of LUNs.
  - Multiple LUN 0's
- Host connections routed to HSD based upon WWN.
- Fewer physical ports needed
  - Reduces complexity & cost
- More overall connections
- Enables consolidation



HP XP Virtualization Assist Summary

- Many of virtualization benefits immediately
  - Single span of control large amount of data
  - Optimum capacity utilization
    - 35% =➔ 70%+
  - Storage pooling
    - By application, by host, by host groups
    - QS to manage workload priorities
  - Significant reduction in physical complexity
    - Fewer storage units
    - Fewer physical connections (storage units and switches)
- Compatible with evolving virtualization architectures

Virtualization Assist

- Host Storage Domains storage pools
  - Better granularity than port level
  - Enables fine granularity pooling
  - Security at WWN name level
  - Performance management/control WWN level
- Works with/without higher levels virtualization
  - Complements other virtualization methods
    - Delegate to lowest common denominator
  - Protects investment as higher levels evolve
    - Many of benefits now, flexibility in future

Thank You For Attending. Questions ?

For Additional Information
Please Call Jim Gursha at HPSS
Email:JimGursha@sandisks.com
Telephone:212-831-0291 (office)
917-359-2087 (cell)