Technical Session 3718:

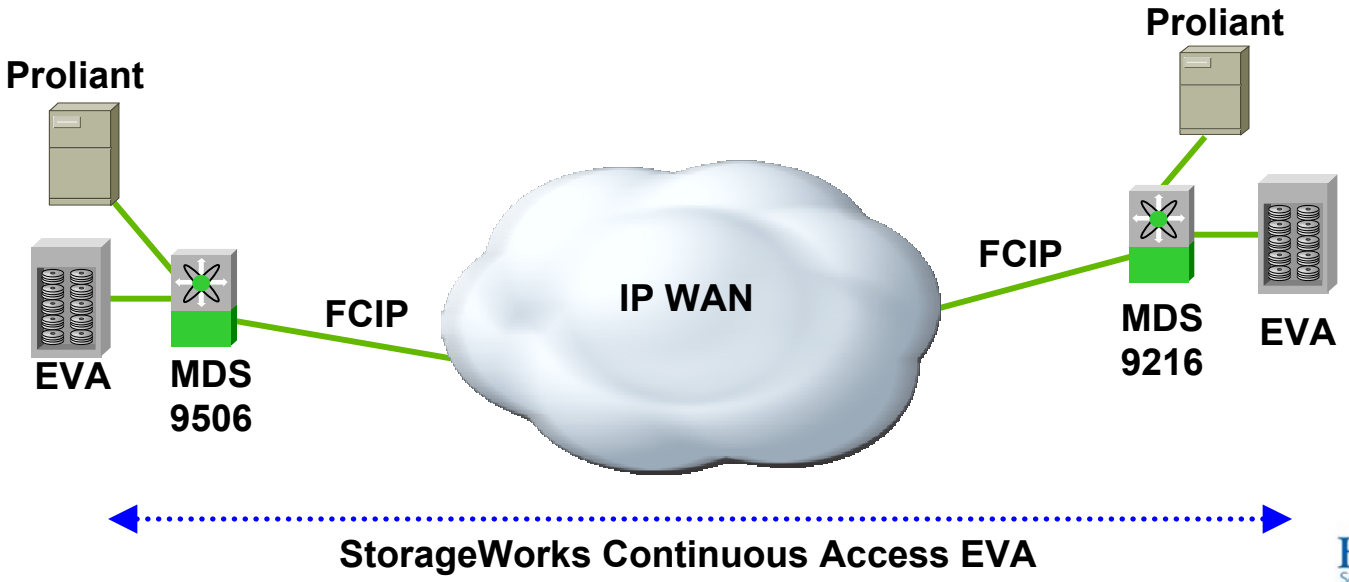
Remote Site Data Replication using the FCIP Protocol

FIT Lab Team

HP Storage Services
Hewlett-Packard

Agenda

- Introduction
- Business Case
- Fundamental Technologies
- Technical walk-through of solution



Team Members:

Jeff Lifschin : Fabric Specialist (HP)

Joe Collura : Fabric Specialist (HP)

Ravindra Neelakant : Tech Marketing (Cisco)

Christopher Greene : Service Strategy Analyst (HP)

Contact Information:

Jeff Lifschin : jeff.lifschin@hp.com

Joe Collura : joe.collura@hp.com

Christopher Greene : christopher.greene@hp.com

Data Replication

Business Case

Why the importance of Data Replication?

- Site outages
  - Power Grid August 2003
  - 1998 NE Canada Ice Storm
  - Physical site disruption
- Malicious Intent
  - Virus
  - Disruptive Hacking
- Application Development
- Data Access Efficiency
- Government Regulation
  - Sarbanes Oxley
  - Pharmaceutical – testing and research retention
  - Health Care – patient record retention
  - Financial Industry – Check 21 data retention

Revenue Loss per hour by Industry

Brokerage Operations	$7,800,000	Utilities	$643,250
Energy	$2,817,846	Healthcare	$636,030
Credit Card Sales Authorizations	$2,600,000	Metals/Natural Resources	$580,588
Telecommunications	$2,066,245	Professional Services	$532,510
Manufacturing	$1,610,654	Electronics	$477,366
Financial Institutions	$1,495,134	Construction and Engineering	$389,601
Information Technology	$1,344,461	Media	$340,432
Insurance	$1,202,444	Hospitality and Travel	$330,654
Retail	$1,107,274	Pay-Per-View TV	$150,000
Pharmaceuticals	$1,082,252	Home Shopping TV	$113,000
Banking	$996,802	Catalog Sales	$90,000
Food/Beverage Processing	$804,192	Airline Reservations	$90,000
Consumer Products	$785,719	Tele-Ticket Sales	$69,000
Chemicals	$704,101	Package Shipping	$28,000
Transportation	$668,586	ATM Fees	$14,500

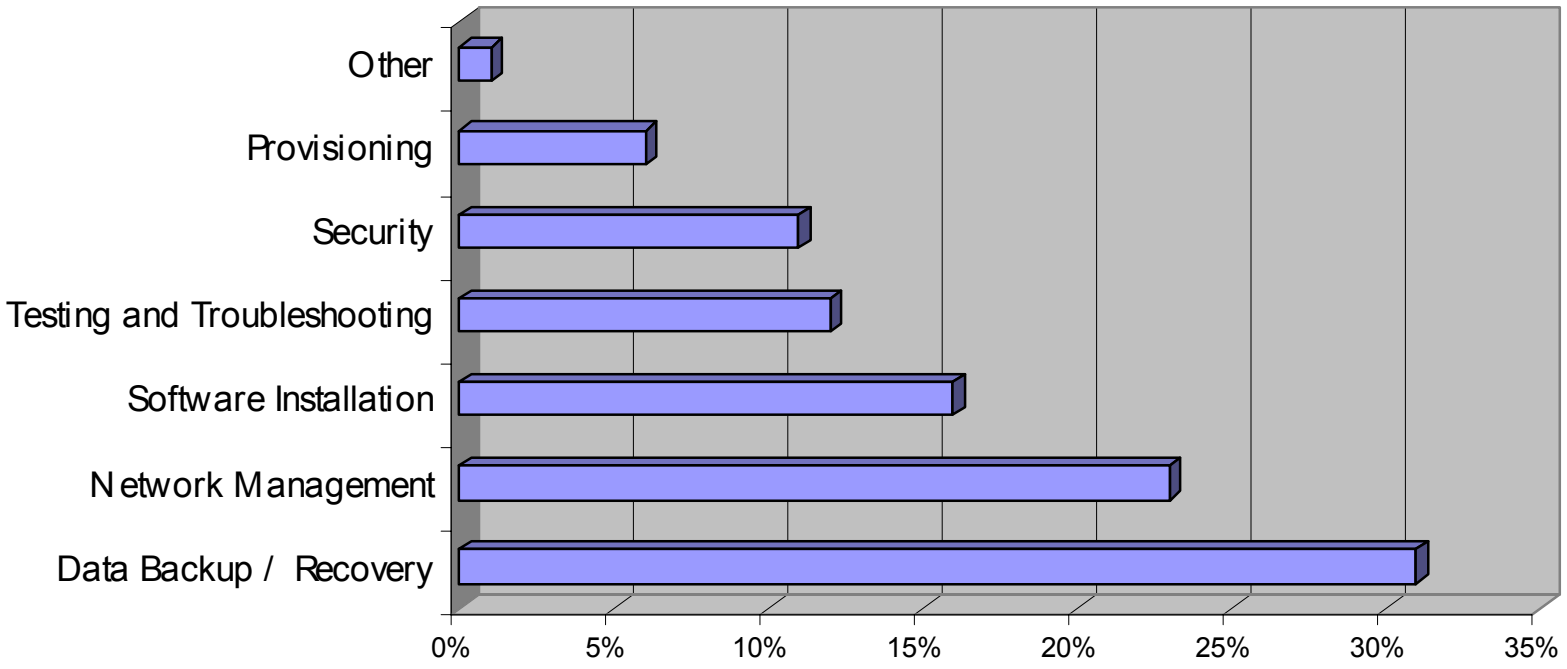
Sources: IT Performance Engineering and Measurement Strategies: Quantifying Performance and Loss, Meta Group, Oct. 2000; Fibre Channel Industry Association.

According to IDC

Storage protection is the highest priority of all planned storage projects (36%)

Source: "Top 2004 North American Storage Investment Projects"
IDC, 2003

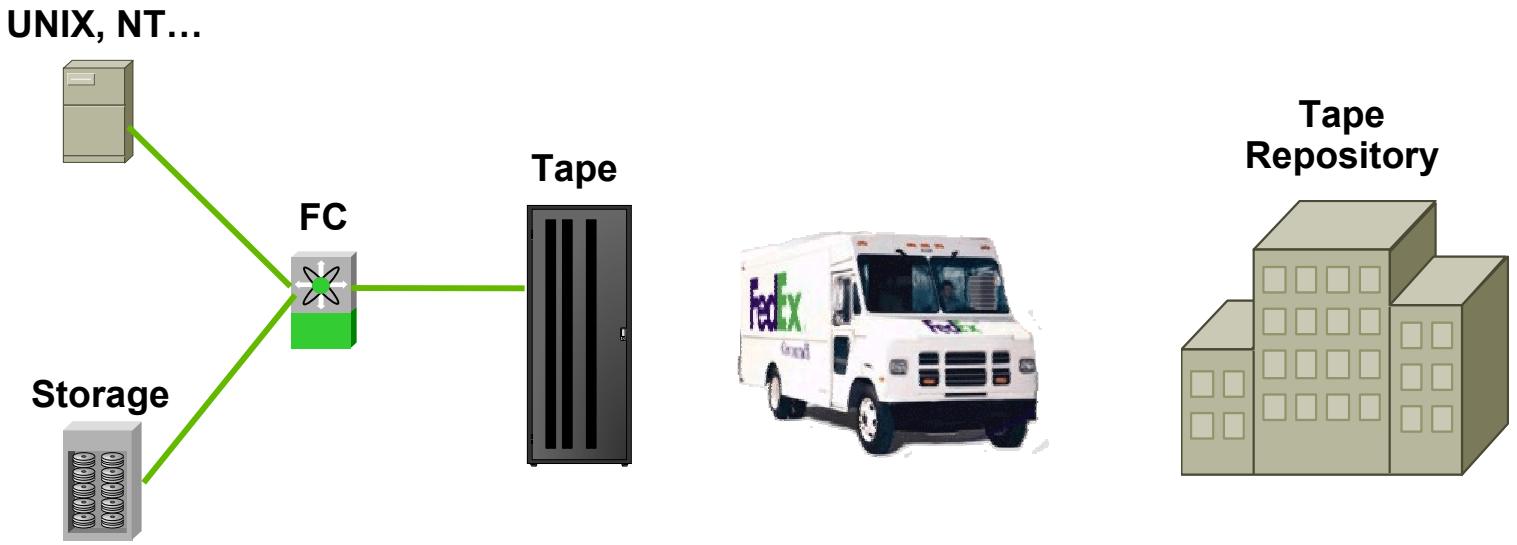
Storage Task Time Consumption



Source: Byte & Switch User Survey, 10/25/02

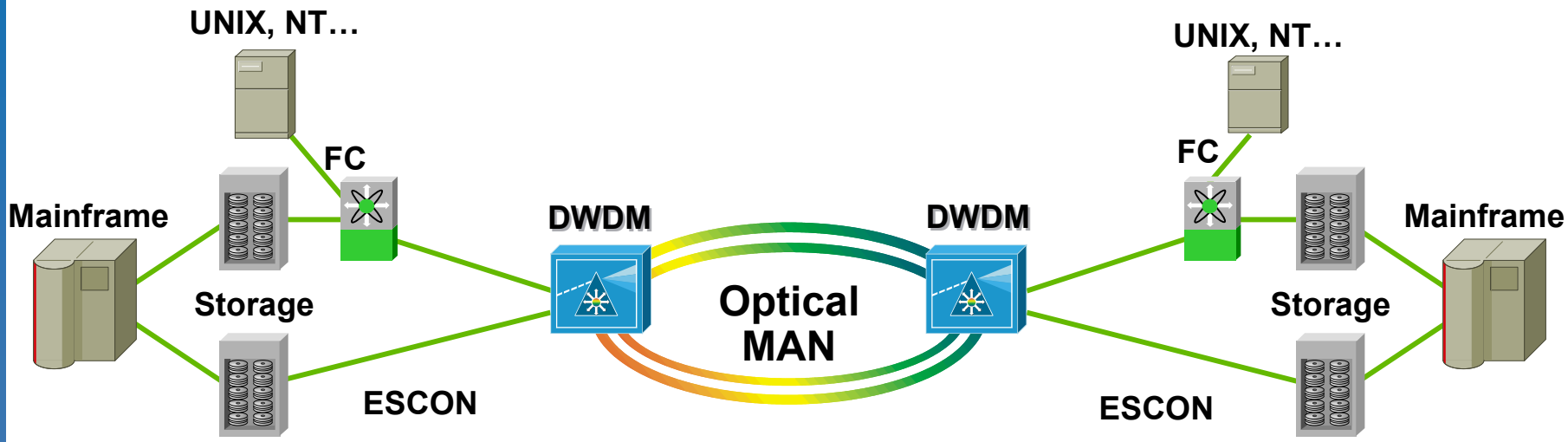
Potential Solution

Tape Backup / FedEx



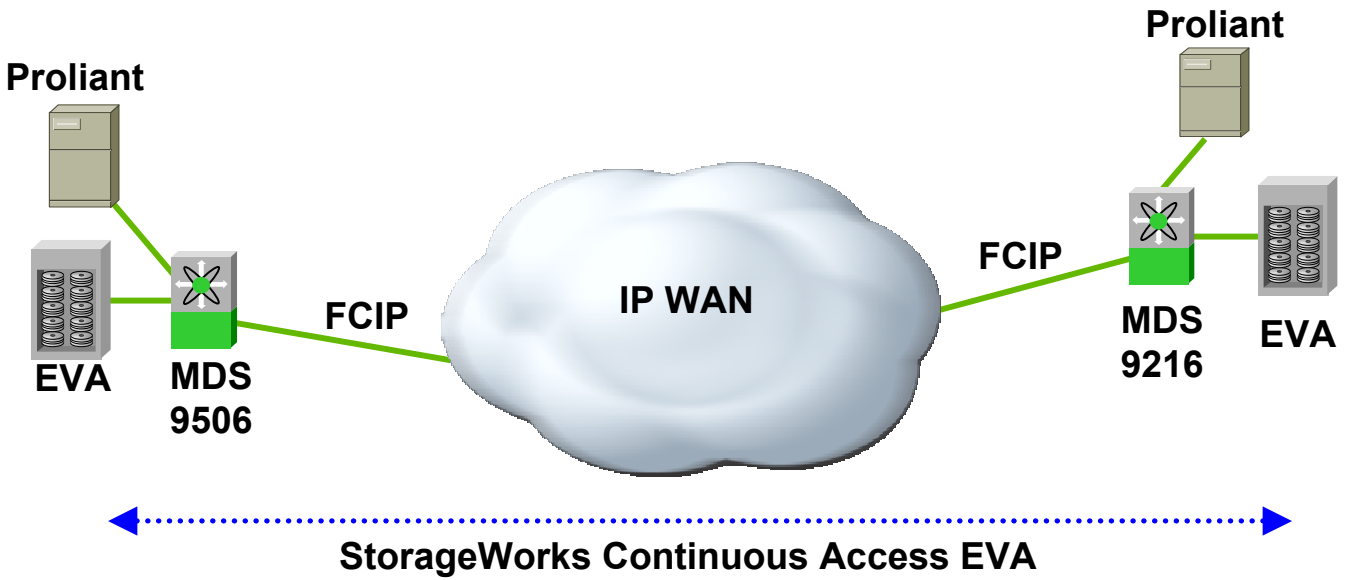
Potential Solution

Fiber Optic Replication



Potential Solution

FCIP Data Replication



Fundamental Technologies

Technologies

- Fibre Channel
- 8b/10b encoding
- TCP/IP
- FCIP
- VSAN
- IVR

Fibre Channel Overview

- Fibre Channel is a computer communications protocol designed to meet the many requirements related to the ever increasing demand for high performance information transfer. The goals of Fibre Channel include:
- Allowing many well-known existing channel and networking protocols to run over the same physical interface and media
- High bandwidth (100MB/s and beyond)
- Flexible topologies
- Connectivity over several kilometers
- Support for multiple data rates, media types, and connectors
- In general, Fibre Channel attempts to combine the benefits of both channel and network technologies.

Fibre Channel

Fibre Channel is the solution for IT professionals who need reliable, cost-effective information storage and delivery at blazing speeds. With development started in 1988 and ANSI standard approval in 1994, Fibre Channel is the mature, safe solution for one and two gigabit communications.

Fibre Channel Key Features

- Price Performance Leadership
- Reliable
- Gigabit Bandwidth Now – Both 1 and 2 gigabit solutions
- Multiple Topologies - Dedicated point-to-point, shared loops, and scaled switched topologies
- Multiple Protocols - Fibre Channel delivers data. SCSI, IP, VI, ESCON and other storage and networking protocols
- Scalable
- Congestion Free
- High Efficiency

Fibre Channel Key Features

- Despite the name, Fibre Channel can run over both copper and fiber media.
- The main tradeoffs are that although longer distances can be achieved with fiber, it is more expensive. Speeds up to 100MB/s can run on both copper and fiber; 200MB/s and 400MB/s require fiber media.
- For fiber, the choices are: 62.5mm multi-mode, 50mm multi-mode, and single-mode.
- Also for fiber, both long wave and short wave lasers can be used. Short wave seem to be most popular now.

8B/10B Encoding Scheme

- At the lowest level, Fibre Channel uses the IBM 8B/10B encoding scheme. Basically, every byte of data that is to be transmitted is first converted into a 10-bit value called a Transmission Character. Using this encoding scheme has the following benefits:
- improvement of the transmission characteristics of information to be transferred
- provides enough transitions to make clock recovery possible at the receiver
- improves the detection of single and multiple bit errors
- Some Transmission Characters contain a unique bit pattern (comma) to aid in achieving word alignment.

TCP/IP

- TCP and IP were developed by a Department of Defense (DOD) research project to connect a number different networks designed by different vendors into a network of networks (the "Internet"). It was initially successful because it delivered a few basic services that everyone needs (file transfer, electronic mail, remote logon) across a very large number of client and server systems.

TCP/IP

- TCP/IP communication is primarily point-to-point, meaning each communication is from one point (or host computer) in the network to another point or host computer. TCP/IP and the higher-level applications that use it are collectively said to be "stateless" because each client request is considered a new request unrelated to any previous one (unlike ordinary phone conversations that require a dedicated connection for the call duration).

TCP and IP

- **IP** - is responsible for moving packet of data from node to node. IP forwards each packet based on a four byte destination address (the IP number). The Internet authorities assign ranges of numbers to different organizations. The organizations assign groups of their numbers to departments. IP operates on gateway machines that move data from department to organization to region and then around the world.
- **TCP** - is responsible for verifying the correct delivery of data from client to server. Data can be lost in the intermediate network. TCP adds support to detect errors or lost data and to trigger retransmission until the data is correctly and completely received.

IP

IP is the primary layer 3 protocol in the Internet suite. In addition to internetwork routing, IP provides error reporting and fragmentation and reassembly of information units called datagrams for transmission over networks with different maximum data unit sizes. IP represents the heart of the Internet protocol suite. IP addresses are globally unique, 32-bit numbers assigned by the Network Information Center. Globally unique addresses permit IP networks anywhere in the world to communicate with each other. An IP address is divided into three parts. The first part designates the network address, the second part designates the subnet address, and the third part designates the host address.

TCP

TCP is a connection-oriented transport protocol that sends data as an unstructured stream of bytes. By using sequence numbers and acknowledgment messages, TCP can provide a sending node with delivery information about packets transmitted to a destination node. Where data has been lost in transit from source to destination, TCP can retransmit the data until either a timeout condition is reached or until successful delivery has been achieved. TCP can also recognize duplicate messages and will discard them appropriately. If the sending computer is transmitting too fast for the receiving computer, TCP can employ flow control mechanisms to slow data transfer. TCP can also communicate delivery information to the upper-layer protocols and applications it supports.

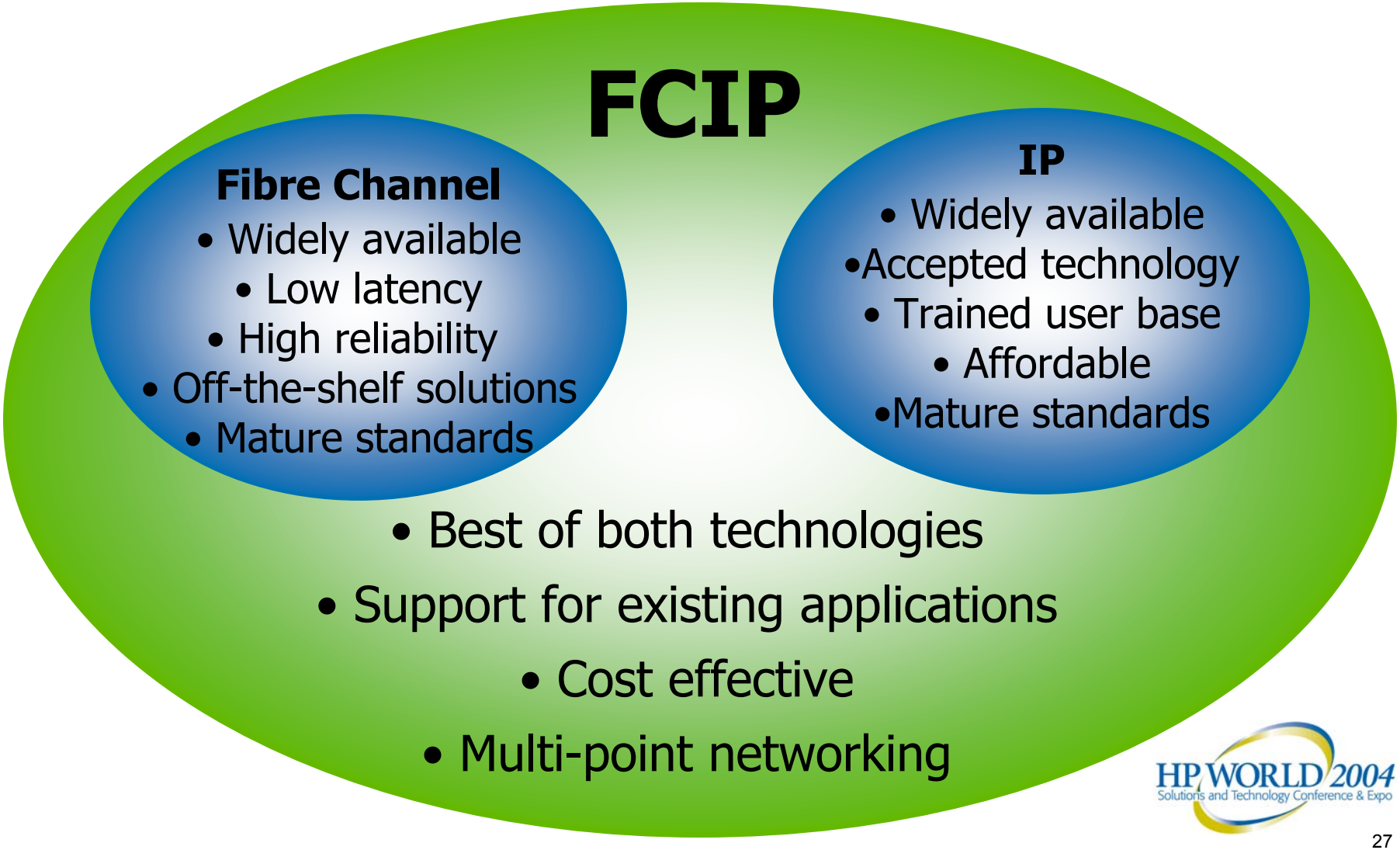
FCIP

- Technology
  - Availability of GbE and optical networking products
  - Availability of managed Ethernet / IP-based Metro and WAN connectivity services
  - Increased availability of dark fiber and xWDM solutions in the metro

FCIP

- FCIP can be widely deployed in a cost-effective manner
  - GbE links now available for less than $10K/month
- Most organizations already have IP connections
  - Extends existing management processes
  - IP provides greatest flexibility at lowest cost for latency-tolerant applications
  - Suitable for backup across campus network, MAN, or WAN
  - Additional applications with increased wire speed
  - Can be deployed within a single enterprise or in an SSP multi-tenant environment

FCIP Summary – The Best Of Both Worlds



VSAN

- Traffic isolation: Traffic is contained within VSAN boundaries and devices reside only in one VSAN, thus ensuring absolute separation between user groups, if desired.
- You can define multiple zones in a VSAN. Because two VSANs are equivalent to two nonconnected SAN fabrics
- SANs, zone A on VSAN 1 is different and separate from zone A in VSAN 2

VSAN

- Scalability: VSANs are overlaid on top of a single physical SAN. The ability to create several logical VSAN layers increases the scalability of the SAN.
- Redundancy: Several VSANs created on the same physical SAN ensure redundancy. If one VSAN fails, redundant protection is provided by a configured backup path between the host and the switch.
- Ease of configuration: Users can be added, moved, or changed between VSANs without changing the physical structure of a SAN. Moving a device from one VSAN to another only requires configuration at the port level, not at a physical level.

VSAN

- Required protocols: Every instance of a VSAN runs all required protocols such as FSPF, domain manager, and zoning.
- Independence: Fabric-related configurations in one VSAN do not affect the associated traffic in another VSAN.
- Containment: Events causing traffic disruptions in one VSAN are contained within that VSAN, and are not propagated to other VSANs.
- Isolation: No communication is possible between VSANs.

Inter VSAN Routing (IVR)

- Inter-VSAN routing allows servers in different virtual SANs to share common storage resources, such as disk or tape, located either locally or remotely, while maintaining the security, scalability, and availability of VSANs.
- Inter-VSAN Routing allows selective transfer of data traffic between devices on different VSANs without merging VSANs into a single logical fabric.
- Inter-VSAN routing can also be deployed across geographically dispersed sites for disaster recovery and business continuity applications.

Technical walk-through of Solution

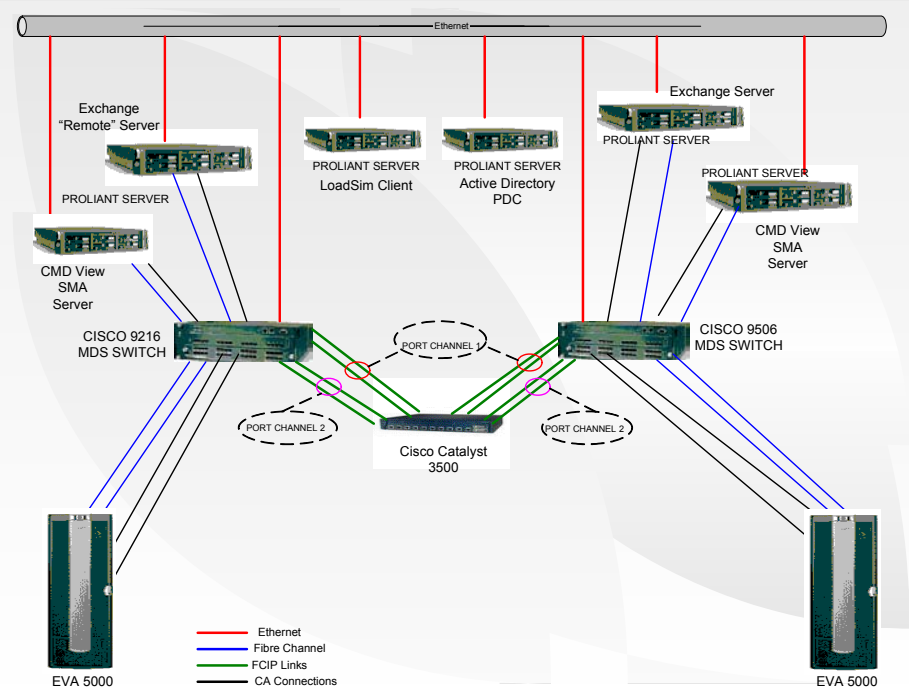
Remote Storage Replication Solution Test Plan

- For our testing, we first tested basic functionality with and without I/O load
- We then injected failures, noting the resulting behavior of the solution.
- Although our testing is not a representation of true real life implementations, we have included real life best practices from our experience in that arena. Since this is a fairly new and evolving technology, new implementation best practices are being introduced on a regular basis.
- The transport of FC traffic over IP offers many challenges to be overcome

Remote Storage Replication Solution Test Plan - continued

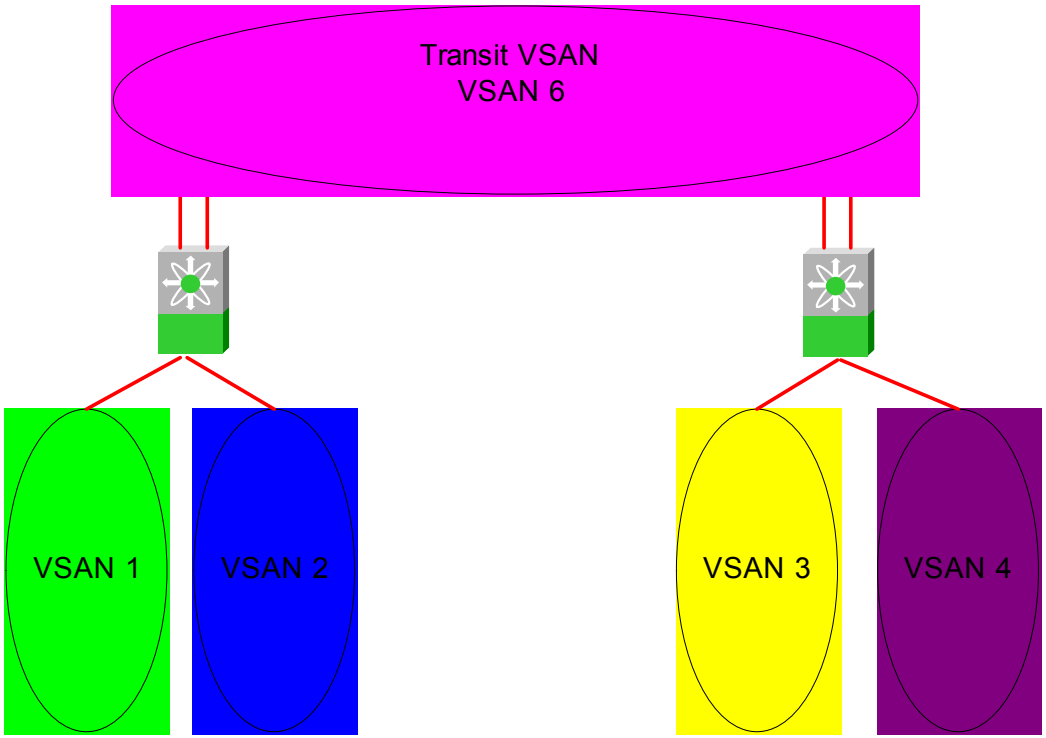
- EVA Continuous Access
  - 2 EVA 5000 with VCS 3.010
  - EVA 5000 CA License
  - CA User Interface V1.1A
  - Storage Management Appliances (SMA) Version 2.1 with Command View version 3.1
- MDS 9000 Switches - FW 1.3(4a)
  - Local Switch – MDS9506
  - Remote Switch – MDS 9216
  - FCIP using IPS-8 Blades
  - SAN_EXTN_OVER_IP License (FCIP)
  - ENTERPRISE_PKG License (IVR)
- Host Applications
  - Microsoft Exchange 2000
  - Active Directory Server – Also PDC and DNS server
  - LoadSim 2000 (Simulate exchange clients)
  - SecurePath 4.0c

Test Plan Topology



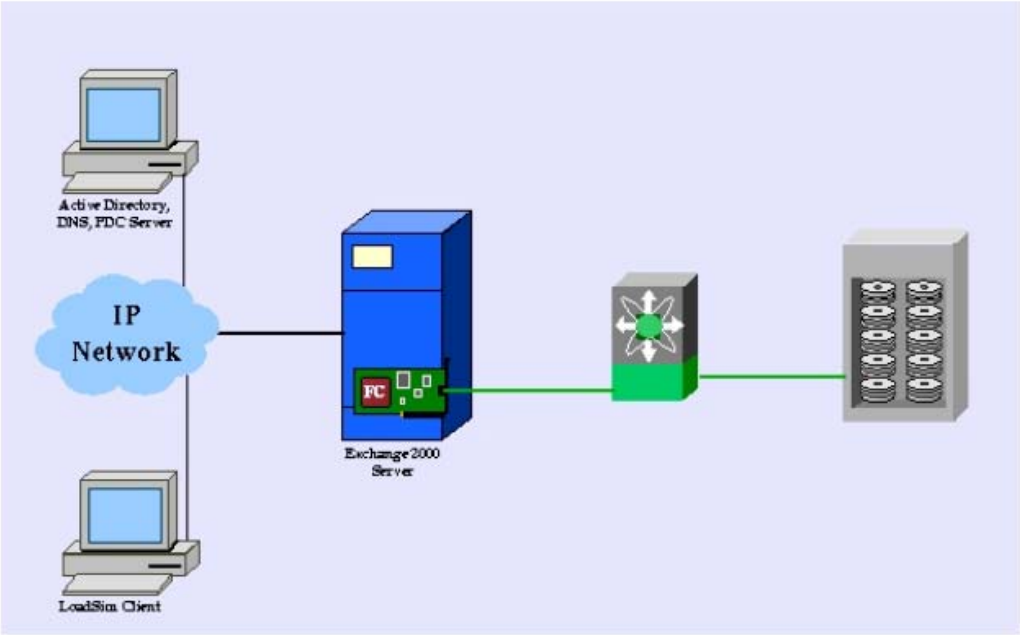
- Each MDS Switch's FC ports are divided into 2 VSANs to represent Fabrics A and B.
- Each MDS switch's Gig-E ports are are assigned to a transit VSAN.
  - This will isolate the edge VSANs from the E ports (FCIP links).
- The 4 FCIP links are divided into 2 port channels.

Test Plan Topology - continued



Exchange Topology

Setup Servers and Networks



EVA Setup Requirements

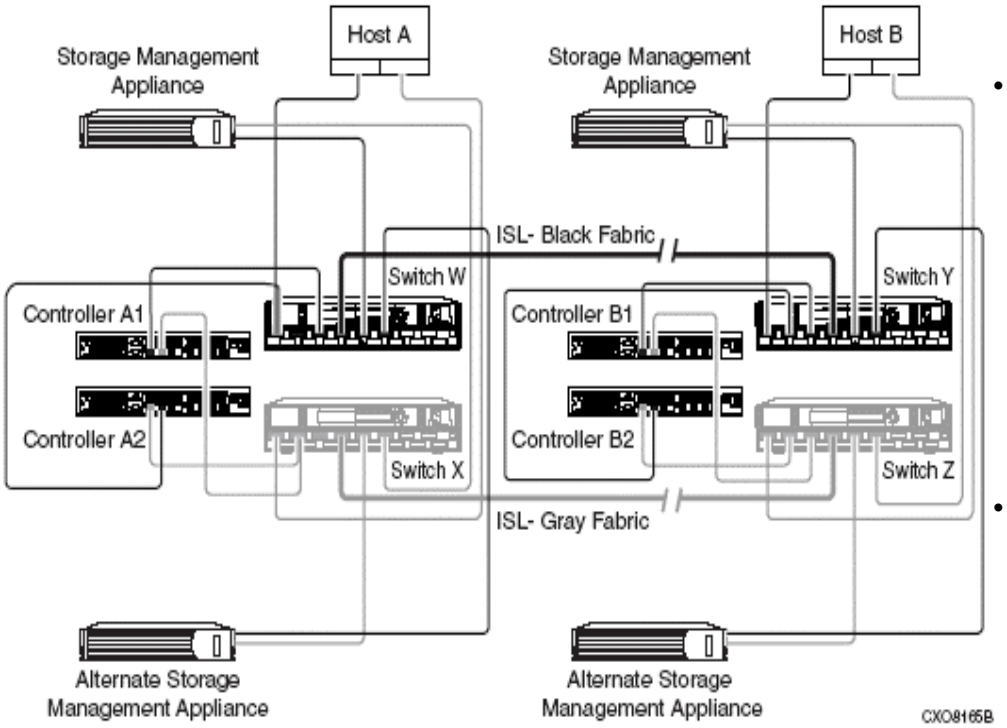


Figure 1: Continuous Access EVA basic configuration

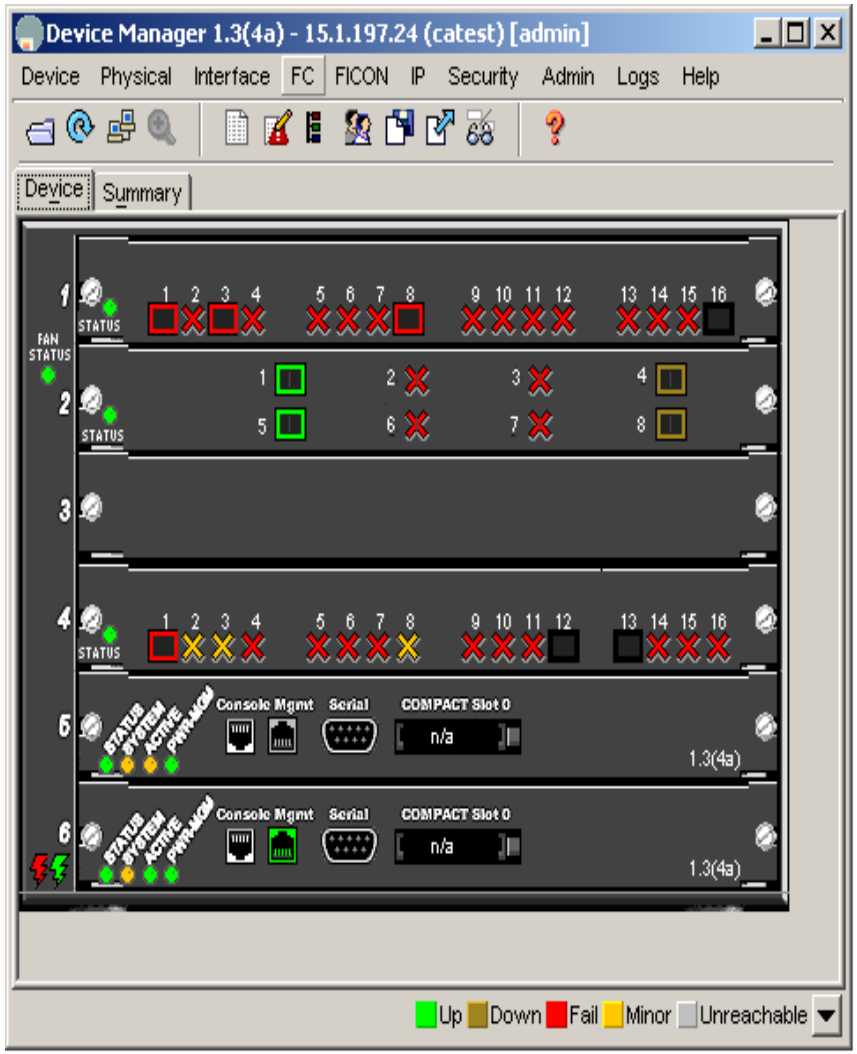
- EVA cabling must follow the below guidelines:
  - **The left hand port of the top controller is cabled to Fabric A**
  - **The right hand port of the top controller is cabled to Fabric B**
  - **The left hand port of the bottom controller is cabled to Fabric B**
  - **The right hand port of the bottom controller is cabled to Fabric A**
- One Storage Management Appliance (SMA) should be used for all CA configurations

MDS Setup Requirements

- FCIP tunnel IP addresses and management IP address must be on 2 different IP subnets
- The VSAN load balance scheme needs to be set to src-id/dst-id in order for CA EVA to work correctly The default setting on MDS switches is src-id/dst-id/oxid
- The following FCIP parameter settings must be configured
  - max-bandwidth-mbps
  - min-available-bandwidth-mbps
  - RTT (round trip time)
  - These values should be set for each FCIP profile on both local and remote switches
- Static routes should be setup between peer Gig-E interfaces. This will ensure that the IO will travel down the appropriate link between peers
- VSAN requirements
  - Four different VSANs should be setup on local and remote switches each with unique domain IDs
  - One common transit VSAN should be setup on both switches with membership assigned to the FCIP interfaces of both switches
  - Enable persistent FCIDs globally per HP best practices.

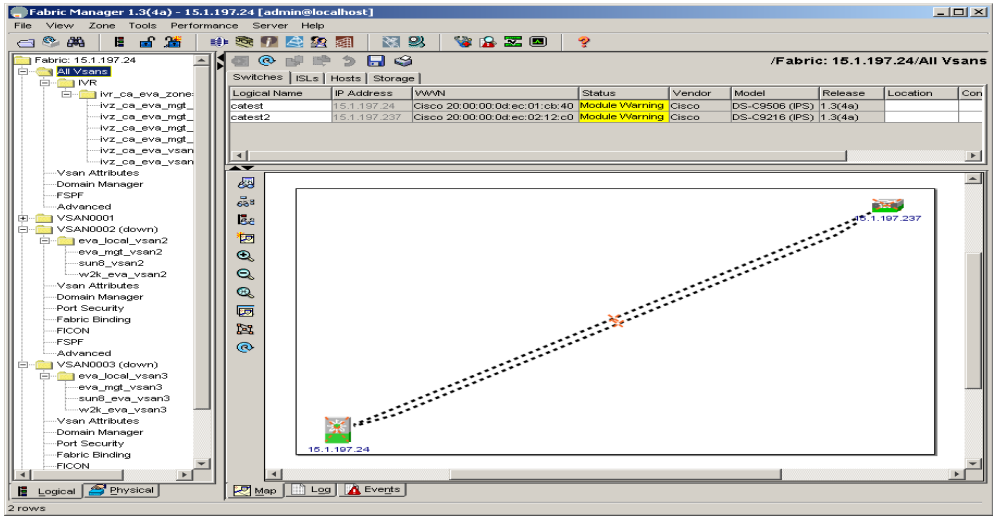
Device Manager

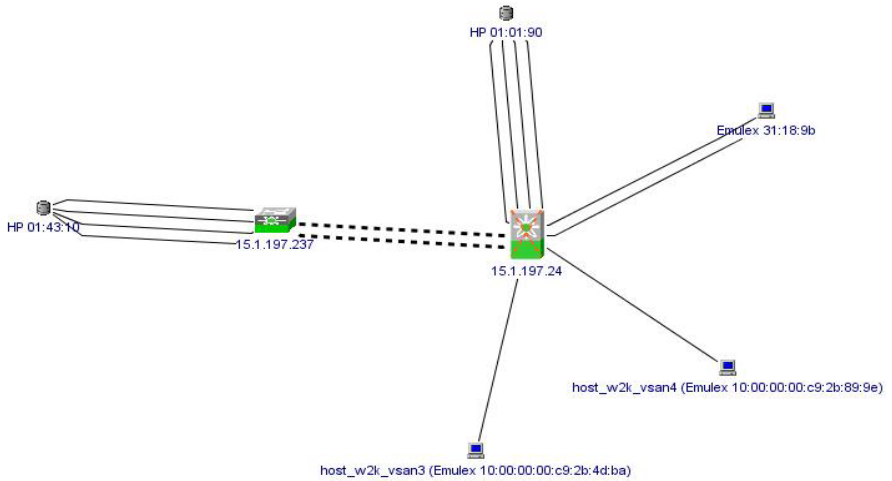
- Device Manager home page will provide you with some basic information, for example:
- active supervisor
- color coded port information



Fabric Manager

- Fabric Manager home page will also provide useful information, as well as a topology of your fabric. For example:
- All zoning, VSAN and IVR information
- Switch IP and WWN





MDS Setup requirements for Inter VSAN Routing (IVR)

- All VSANs must have unique domain IDs
- The Enterprise license is needed for IVR
- The SAN extension over IP license is needed for FCIP
- The transit VSAN will have only the FCIP links as members
- The edge VSANs will have all the n-port devices as members

Additional switch considerations

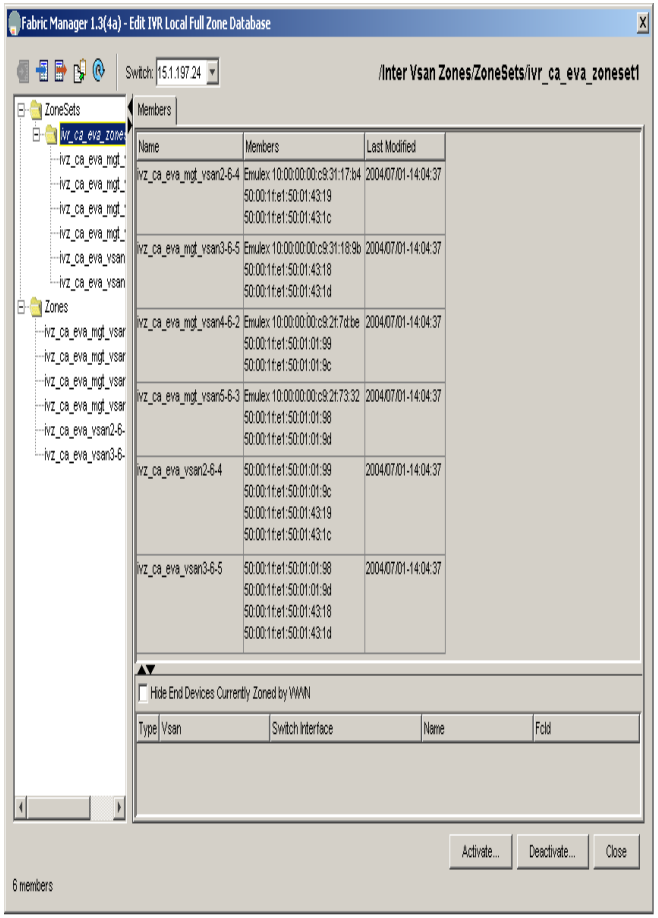
- For slow IP links in the order of 20Mbps or less, additional parameters need to be set in order to prevent congestion on the fabrics. This condition could potentially affect the storage response to host requests.
  - Configure TCP send buffer on each FCIP interface to 1024kB for one concurrent DR group copy or 2048kB for two concurrent DR group copy. Higher values are possible but not yet fully tested
  - Configure compression on FCIP interfaces
  - Do not do more than two concurrent DR group copies at a time

Test Specifications

- Local Site Connectivity test
- Stress testing using Hazard Version 6.08
- Local failure testing with LoadSim running
  - Simulated a fabric failure by removing 1 of the FC Switching Modules
  - Removed Active Supervisor Module
- Continuous Access testing using Synchronous and Asynchronous modes
- Intersite FCIP link failure. Failed 1, 2, 3 and finally all 4 links
- CA Failure Testing on local (source) switch
  - Failed active Switching Module
  - Failed active Supervisor Module
  - Failed local switch by power cycling.
    - **Confirmed that CA process resumed after switch powered up.**
- Site failover test
  - Powered down local EVA to simulate site failure
  - Failed over DR Group to remote EVA
  - Powered up remote exchange server , initialized LoadSim and verified Exchange server database was accessible

Test Results

- Local Connectivity test – test passed
  - With use of VSANs & zoning, Basic functionality was confirmed
- Stress testing – test passed
  - Hazard testing ran error free for 27, 72 and 96 hours

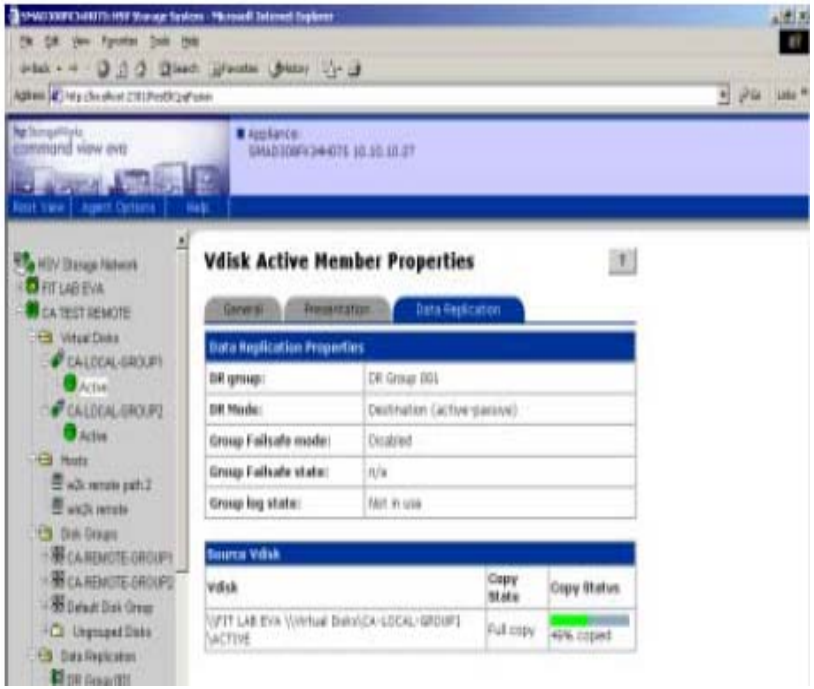


Test Results - continued

- Local failure testing – test passed
  - Removed an FC Switching Module. Observed host to storage I/O failover to redundant module with no disruption to I/O. Each blade represents a different VSAN (Fabric).

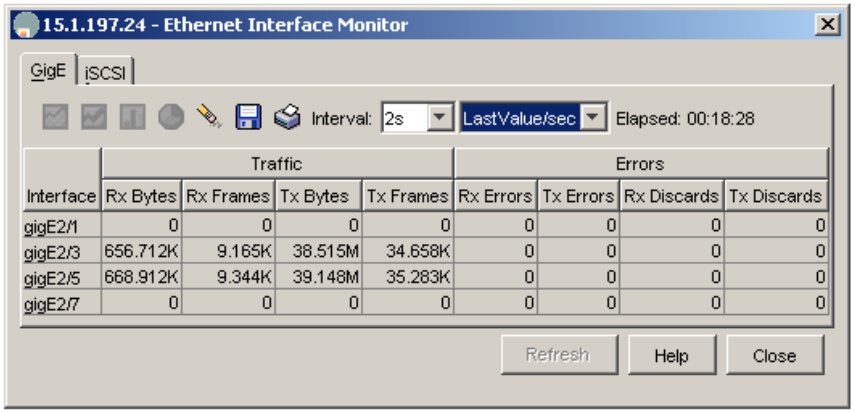
- Removed active supervisor module – test passed
  - There was no disruption to I/O. Standby supervisor seamlessly took over active role

Test Results - continued



- Continuous access testing – test passed
  - Initiated the CA process with LoadSim running
  - Monitored CA merge till completion
  - We then stopped LoadSim and shut down local exchange server
  - We failed over DR groups to remote EVA, Powered on remote exchange server and verified database was accessible by initiating LoadSim

Test Results - continued
Intersite FCIP link failure - test passed



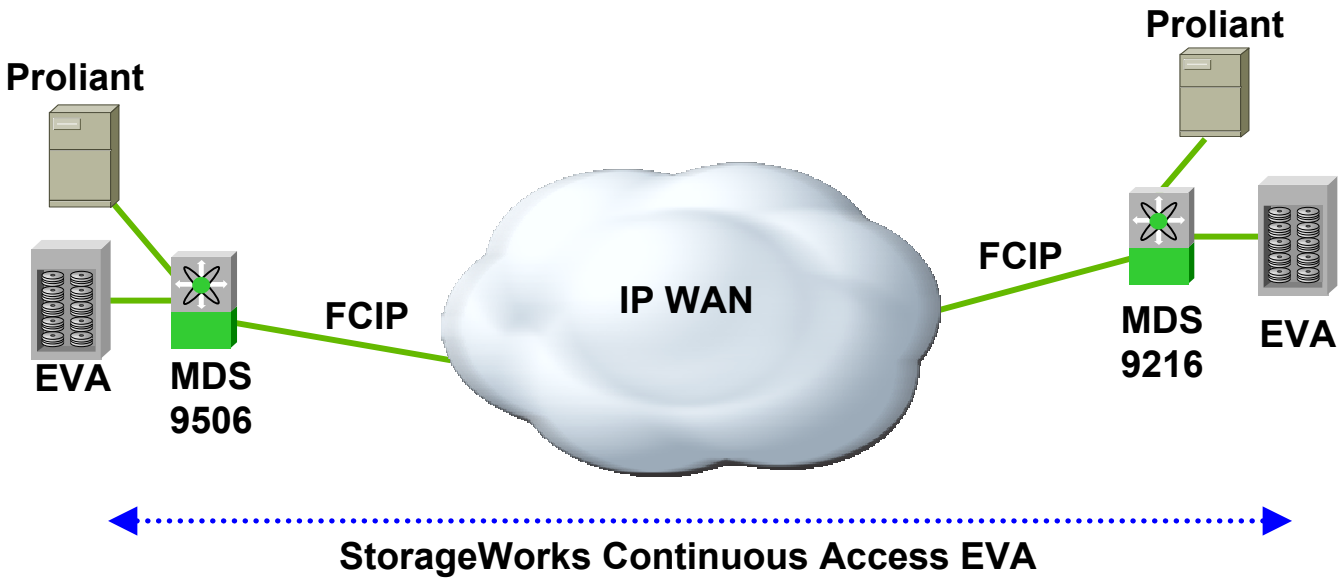
- Failed 1,then 2, then 3 FCIP links
  - Noted the redirection of I/O after each link failure
- Failed the fourth (final) FCIP link
  - CA process halted as expected
  - After link was re-established, CA process resumed where it left off

CA failure testing on local (source) switch – test passed

- Failed active Switching Module
  - Verified that I/O failed over to other host path and CA process continued without disruption
- Failed active Supervisor Module
  - Verified that standby supervisor took active role and there was no disruption to I/O
- Power cycled local switch
  - LoadSim timed out as expected due to loss of local SAN connectivity and CA process halted
  - After switch powered up, CA process resumed where it left off and LoadSim needed to be reinitialized

Site failover test – test passed

- Powered off local EVA to simulate a site failure
- Failed over DR groups to remote EVA
- Powered on remote exchange server and verified database was accessible by initiating LoadSim


Proliant
Proliant
EVA
MDS
9506
FCIP
IP WAN
FCIP
MDS
9216
EVA
StorageWorks Continuous Access EVA

HP WORLD 2004

Solutions and Technology Conference & Expo

Co-produced by:

interex

shared knowledge • shared power

encompass

AN HP USER GROUP

RECOMMENDED TRAINING VENUE FOR THE

HP Certified Professional