HP SureStore Disk Array XP Family Technical Pre-Sales Training

ESG503SG10303

invent

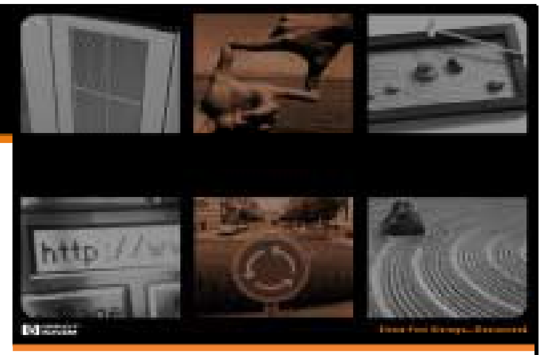
student guide

HP SureStore Disk Array XP Family Technical Pre-Sales Training

ESG503SG10303

training

HP SureStore Disk Array XP Family Technical Pre-Sales Training
Student Guide
March 2003



Welcome to

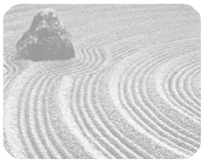
HP SureStore E XP Family Technical Pre-sales HP Channel Partner Training – SR26013

August 2001/Ver. 5

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Module Agenda



- Get to know the group exercise
- XP512 Product Feature Detail
- Course Agenda & Objectives
- XP512 Demo
- XP512/48 Product Positioning
- Wrap Up

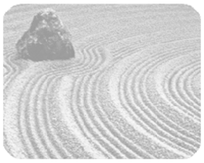
NSSO
Stress-Free Storage...Guaranteed

Get-to-Know-the-Group Exercise

- Introduce yourself to the group
  - Your name
  - The company you work for
  - Your position within the company
  - Give one or two expectations you have of this course
- What are you currently selling?
- What is one quality or skill that you possess that makes you a good at what you do?

NSSO
Stress-Free Storage...Guaranteed

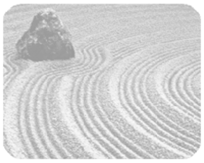
Course Agenda



- Welcome and Introduction
- XP512 Introduction (Demo)
- XP512/48 Product Introduction
- LUN Configuration Manager XP
- AutoPath XP Auto LUN XP
- Performance Advisor XP
- Secure Manager XP

NSSO
Stress-Free Storage...Guaranteed

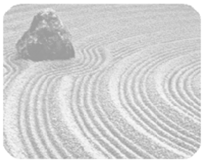
Course Agenda (cont.)



- Business Copy XP
- Command View XP
- Continuous Access XP
- RAID Manager XP
- Configuring for Performance
- Volume Concepts
- OpenView Operations Center Smart Plug-In (SPI) for XP Arrays
- Microsoft Windows Configurations
- XP Family Array Services Value Chain

N S S O
Stress-Free Storage...Guaranteed

Course Agenda (cont.)



- Cache LUN XP
- SCSI Host Array Connect

N S S O
Stress-Free Storage...Guaranteed

hp surestore disk array xp512 summary – product tour

http://www.storage.hp.com/disk_arrays/highend/xp512/index.html

Available at URL:
http://enterprisestorage.hp.com/products/disk_array/xp512.html

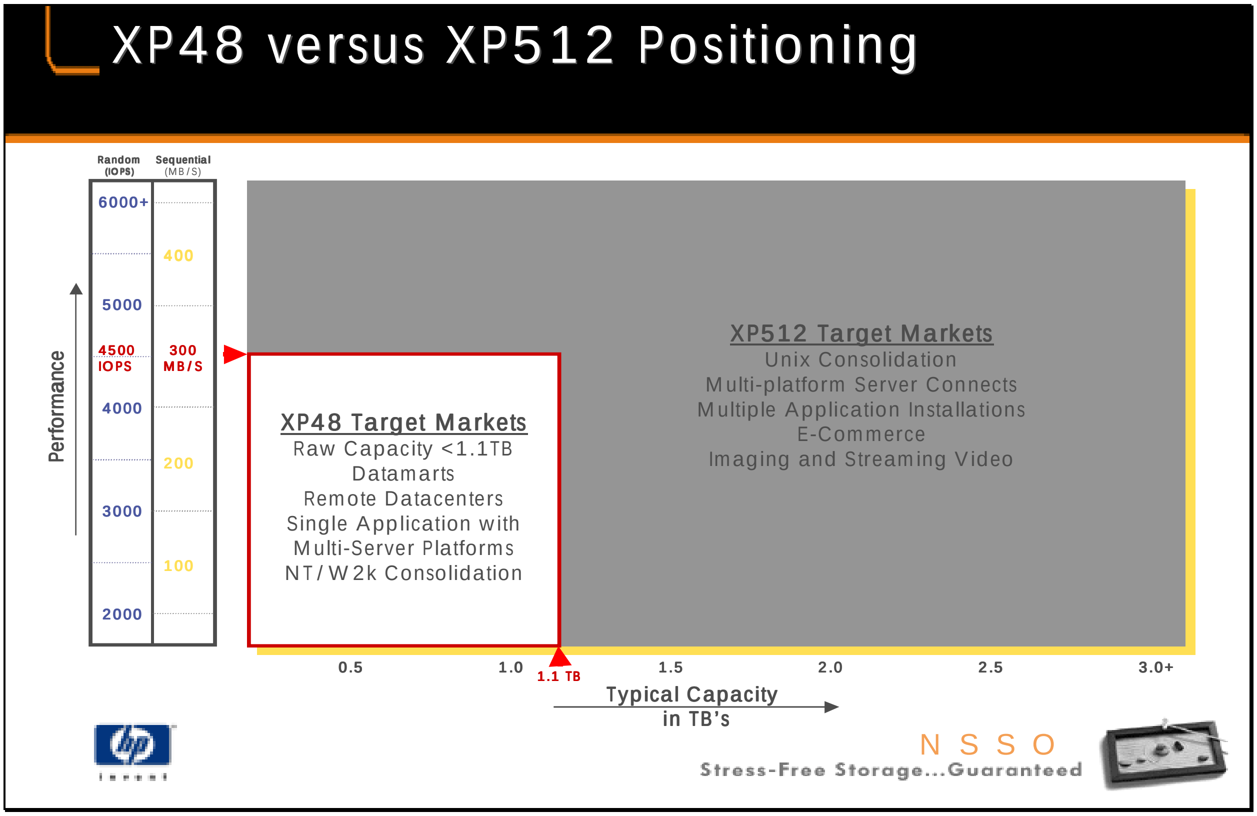
XP48 Positioning, Pricing and Product Mix Objectives

XP48 versus XP512 Positioning

- Basic Positioning Statement:
  - Sell the XP48 when:
    - raw capacity is <1.1TB with small growth expected
    - random performance is <4500 IOPS maximum
    - sequential performance is <300 MB/s
  - Sell XP512 in all other cases
  - XP256 is replaced by the combination of the XP512 and XP48

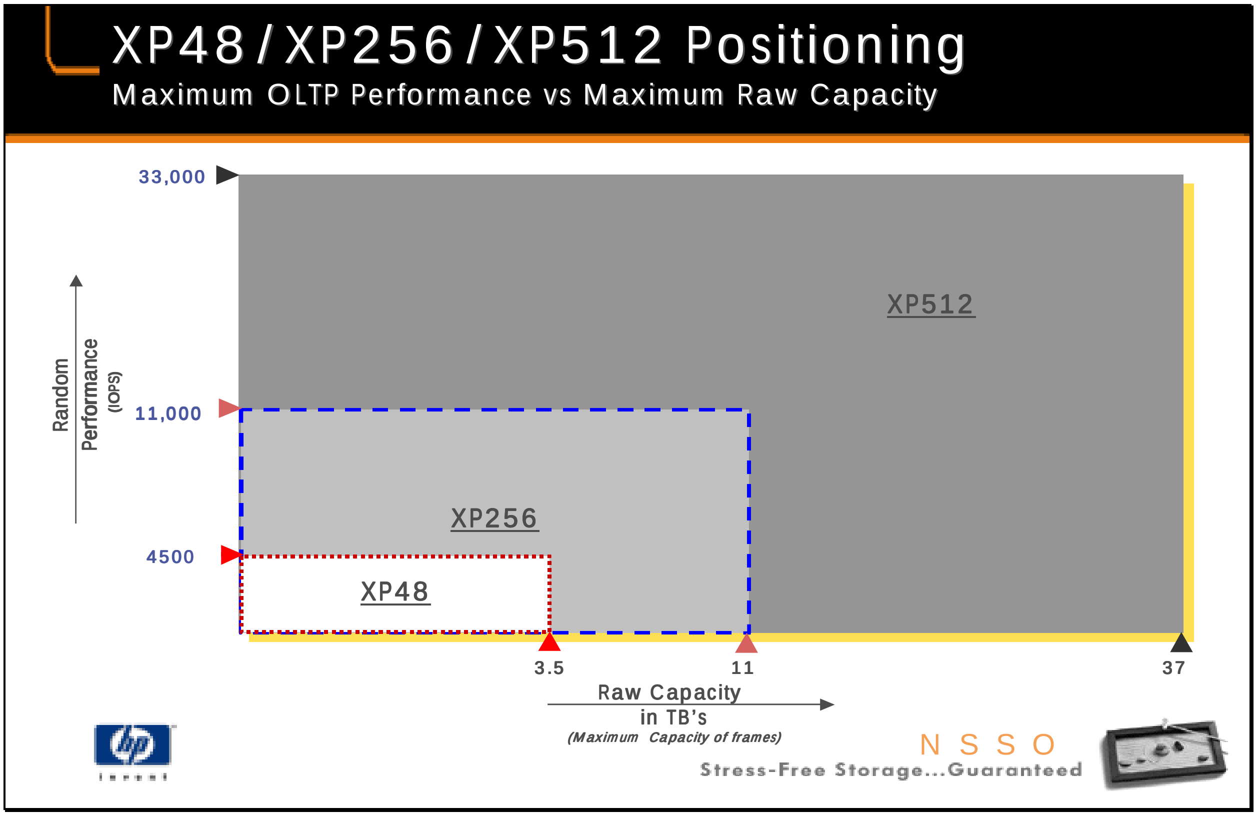
Expect that XP512 is sold in most cases!

NSSO
Stress-Free Storage...Guaranteed



Notes:

1. Small drive count limits *maximum* random performance of XP48 to a max of ~4500 IOPS with 11 array groups (44 drives total - spares cannot be used for “data”).
2. Single ACP pair limits *maximum* sequential performance of XP48 to a max of 300MB/s with 11 array groups (44 drives total - spares cannot be used for “data”).
3. *Typical* XP48 configurations expected to be ~1.1 TB (4 Array groups) which further limits performance.
4. XP512 can be configured to outperform XP48 at most price/capacity points using larger drive count and additional ACP pairs.


XP48 / XP256 / XP512 Positioning
Maximum OLTP Performance vs Maximum Raw Capacity
33,000
11,000
4500
Random
Performance
(IOPS)
XP512
XP256
XP48
3.5
11
37
Raw Capacity
in TB's
(Maximum Capacity of frames)
N S S O
Stress-Free Storage...Guaranteed

XP48 Pricing Strategy

- XP48 is priced to match the XP256 / 47GB drive price points for <1.1 TB raw configurations.
- Relative to the XP512, XP48 provides affordable entry prices for <1.1 TB configurations requiring small to moderate performance with limited growth/scalability requirements
- XP48 is targeted to provide an ~18% worldwide reference list price difference from XP512 at a 1.1 TB capacity point.

NSSO
Stress-Free Storage...Guaranteed

Product Mix Objective

- The XP256 has been replaced by a combination of the higher performing, more scalable XP512 and the smaller XP48.
- Sales objective is to sell XP512 in most situations
  - Targeted revenue mix is 80% XP512, 20% XP48 or a targeted DKC unit split of about 60% XP512, 40% XP48.
  - This mix is required to hit XP family gross margin objectives.
- XP48 is priced to provide an affordable system in the <1.1 TB range with light to moderate performance requirements

NSSO
Stress-Free Storage...Guaranteed

XP512 Product Feature Detail

- 6.4 GB /sec. Crossbar Switch
- 72 GB to 37.3 TB capacity
- 8 to 32 Fibre Channel or ESCON host ports
- 4 to 512 native FC disks
- 2 to 32 GB mirrored cache
- 512MB to 1.28 GB Shared Memory
- 18GB-15k and 73GB-10k RPM FibreChannel Disk Drives

NSSO
Stress-Free Storage...Guaranteed

XP512 Unsurpassed Data Integrity

- RAID 0/1 , RAID 5
- Mirrored write cache
- Multiple crossbar paths to all components
- CHIP and ACP provided only in pairs
- Dual, concurrently active paths
- Split power domains on data paths
- Hot pluggable boards
- Hot pluggable fans, power supplies, controllers
- “Full speed”, online firmware upgradeability

NSSO
Stress-Free Storage...Guaranteed

XP512 supported server platforms

For FibreChannel and ESCON CHIP Connectivity

- HP-UX (See Note 3)
- MPE (See Note 3)
- Windows 2000 / NT4.0 (See Note 3)
- Solaris
- AIX
- IBM S390 compatible mainframes
- Linux, Netware
- Compaq/Tru64 UNIX
- IBM NUMA-Q (Sequent Dynix)
- SGI

NOTES:

1. More details about supported server configurations may be found in the XP512 Product Software and Connectivity Summary available on ESP.
2. All FibreChannel support includes basic and HA connectivity
3. **FWD SCSI** connectivity using the A5814A router front end includes only HPUX, MPE and NT, non-HA only.

NSSO
Stress-Free Storage...Guaranteed

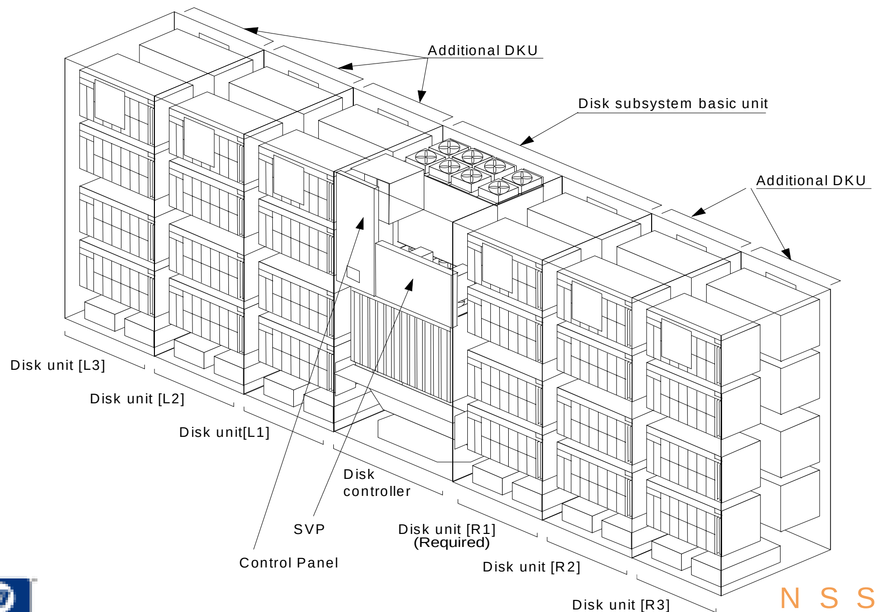
XP512 Sample Configurations

- ***Capacity 3 TB***
  - 1 DKC - A5951A
    - 2 GB cache
    - 512 MB shared memory
    - 1 ACP
  - 1 ACP - A5964A
  - 1 DKU - A5965A #001
  - 11 x 73GB groups A5968A
  - 1 spare disk A5968S
  - 3 x 2 GB cache A5962A
  - 256 MB shared memory A5963A
  - 1 FC/CA CHIP A5957A

- ***Performance 3 TB***
  - 1 DKC - A5951A
    - 2 GB cache
    - 512 MB shared memory
    - 1 ACP
  - 3 ACP - A5964A
  - 2 DKU - A5965A #001
  - 2 x 18GB groups A5970A
  - 10 x 73GB groups A5968A
  - 1 spare disk A5966S
  - 1 spare disk A5968S
  - 5 x 2 GB cache A5962A
  - 3 x 256 MB shared memory A5963A
  - 1 FC/CA CHIP A5957A

NSSO
Stress-Free Storage...Guaranteed

XP512 System Layout



NSSO
Stress-Free Storage...Guaranteed

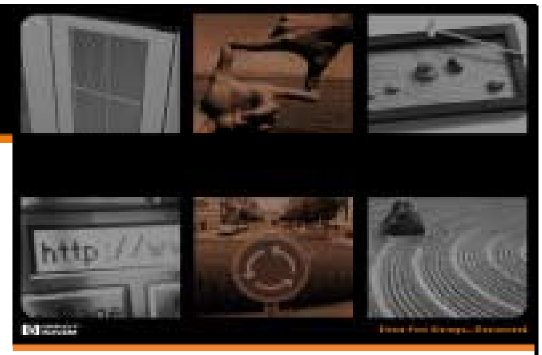
XP512 Resources Available
Sales & Technical Tools

- XP512 Product Brief: **keyword 'XP512PB'**
- XP512 Technical Whitepaper
- ESP website: **keyword "gohome"**
- CDROM: **self-paced training tool and XP512 Walkthrough**
- Configuration/Ordering Guide: **Keyword 'XP512CONFIG'**
- XP512 Product Software and Connectivity Summary : **keyword "XP512CONNECT"**
- Competitive and Misc. Product Detail: **ESP Search**
- Reference customers: **WINREF Database on ESP**

NSSO
Stress-Free Storage...Guaranteed

Module Wrap-up

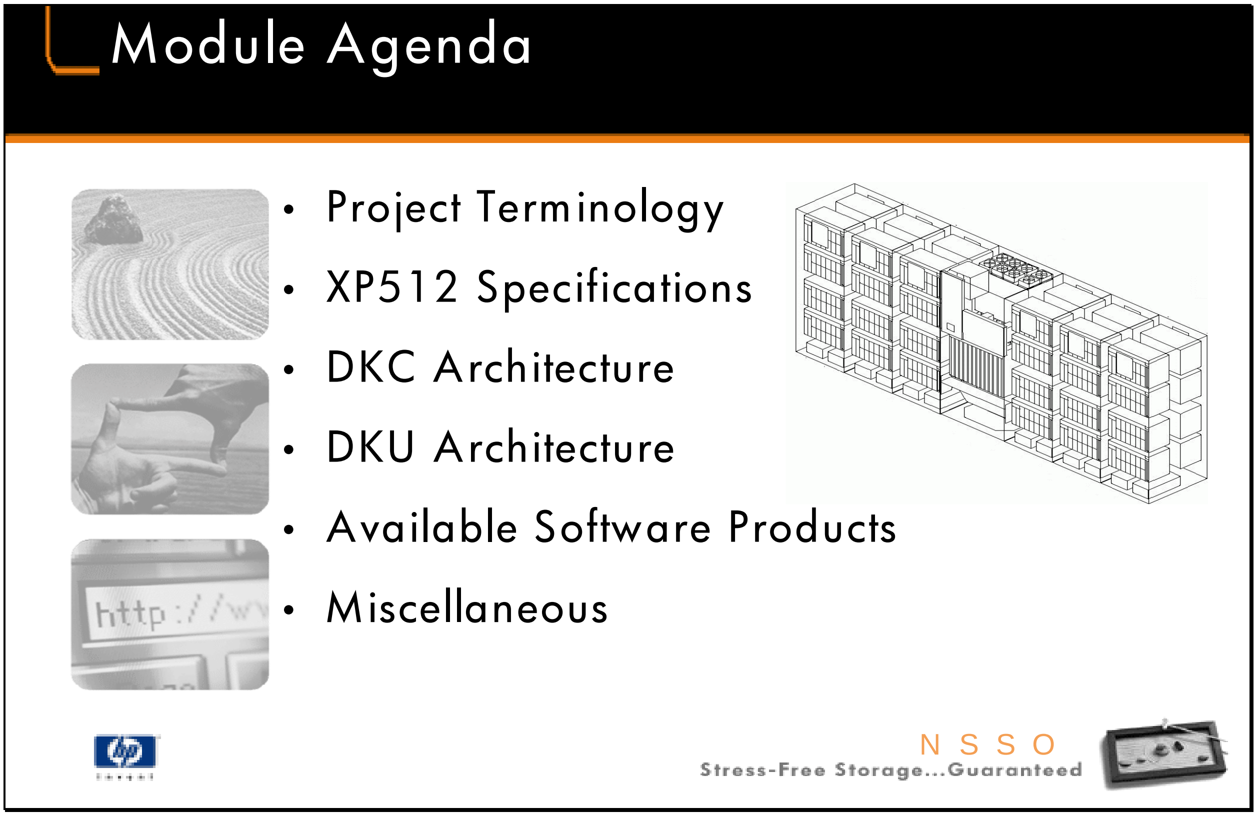
N S S O
Stress-Free Storage...Guaranteed



Module 1

XP512 & XP48 Hardware Overview Presentation

N S S O
Stress-Free Storage...Guaranteed



The focus of this training slide set is to describe only the main features that are new and/or different between the XP256 and the XP512/48. This training slide set assumes the student has a pretty clear understanding of the XP256, thus assumptions are made that repeating information that applies to both the XP256 and the XP512 are not necessary (for example, understanding of hardware redundancies throughout the array, understanding of the data and control buses, understanding of HDD slot location).

XP512 Product Feature Set

- 6.4 GB /sec. Crossbar Switch
- 72 GB to 37 TB capacity (w/72 GB disks)
- 8 to 32 Fibre Channel or ESCON host ports
- 4 to 512 native FC disks
- 2 to 32 GB mirrored cache
- 512MB to 1.5 GB Shared Memory
- 18GB, 47GB and 73GB 10k RPM FibreChannel Disk Drives

N S S O
Stress-Free Storage...Guaranteed

XP512 Project Terminology

- *Amsterdam*- HP development code name for the XP256 at launch
- *Paris*- original HP development code name for the XP512 until January 2000
- *Orca*- HP development code name for the XP512 at launch
- *Cuda*- HP development code name for the XP48 at launch
- RAID300- Hitachi development name for the XP256

  "7700E" HDS product name for the XP256
- RAID400- Hitachi development name for the XP512

  "9900E" HDS product name for XP512
- RAID401- Hitachi development name for the XP48

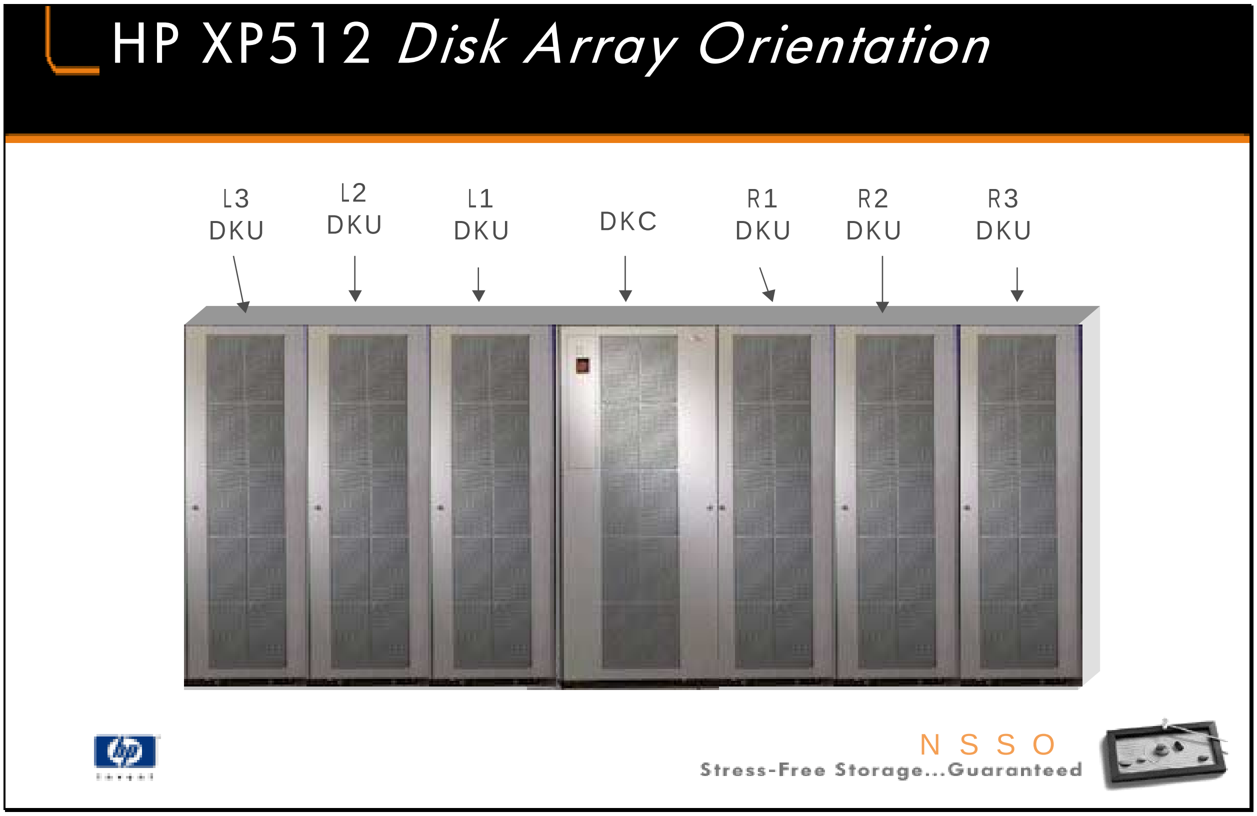
NSSO
Stress-Free Storage...Guaranteed

Hitachi Data Systems (HDS) also sells the XP array?
YES- HDS is an OEM reseller of Hitachi Ltd's disk array, just like HP (SGI also sells the XP array!)
What competitive advantages does HP have over HDS when selling the same OEM'd array from Hitachi Ltd?
For a complete listing with description, go to http://esp.mayfield.hp.com and enter keywords

"HPhds3"
"HDSkey"

Highlights from the documents listed above:
HP-UX connecting to an XP array will

-Support for up to 256 LUNs per port. The HDS array will only support 8
-LUN security via HBA WWN
-Higher queue depth (1024 verses 256)

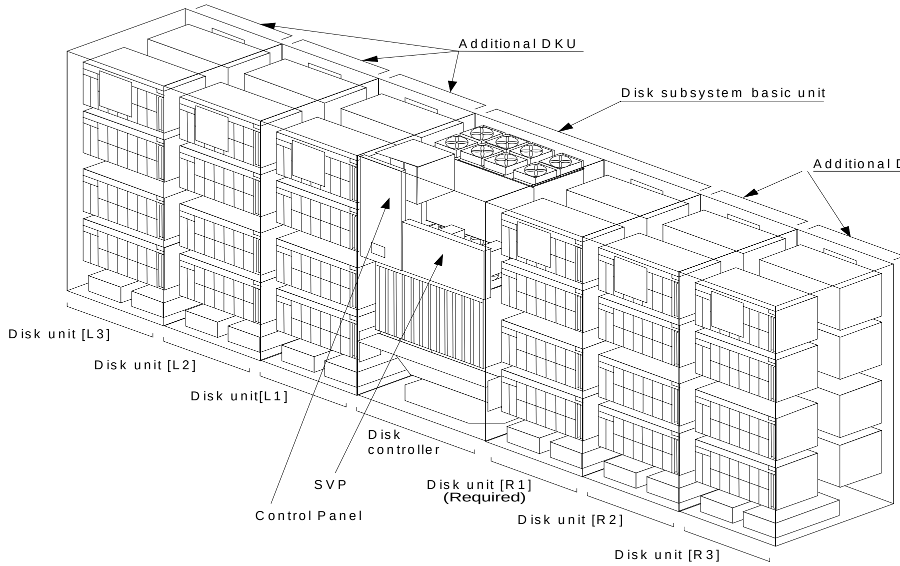


The XP512 consists of two or more cabinets with a total maximum of seven cabinets. The cabinets are referred to as DKC or DKU:

1**. DKC**- Disk Control Frame
-within each array, there will only be one DKC
-contains the primary hardware components:
  Channel host interface processor (CHIP)
  Service Processor (SVP)
  Cache
  Shared Memory
  Backup Batteries
  Power supply for the DKC and the first attached DKU
  Operator's Panel
2. **DKU**- Disk Unit Frame
-within each array, there will always be one or more DKUs totaling no more than six (for the XP512). DKU's are referenced by their respective position from the front in relation to their position to the DKC (i.e.- R1= first DKU to the right; L1= first DKU to the left).

The above illustration shows a full cabinet setup (1 DKC + 6 DKUs).

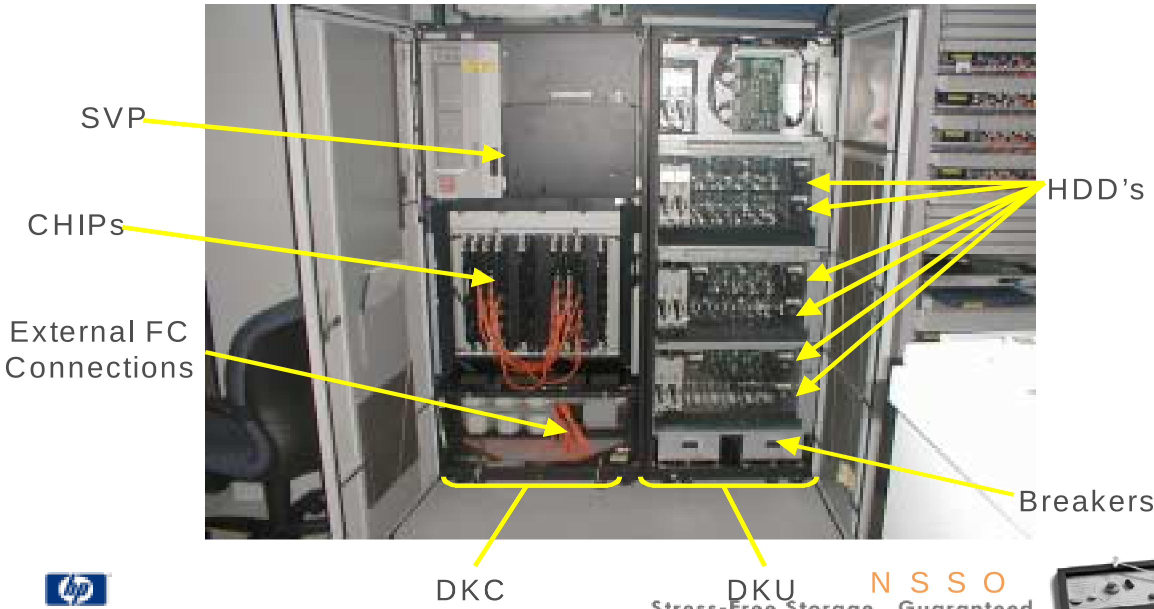
XP512 System Layout



N S S O
Stress-Free Storage...Guaranteed

The new the HP SureStore E Disk Array XP512 is a generational refinement of the successful HP SureStore E Disk Array XP256 from Hewlett-Packard. The XP512 Disk Array retains all features of the XP256 while increasing the performance and scalability of the XP256. The HP SureStore E Disk Array XP512 offers greater capacity, high-speed mass storage, continuous data availability, ease of service, scalability, and connectivity than the XP256. The array can be attached to open-system environments such as HP-UX, Windows 2000/NT, Solaris, Linux, AIX, Netware, HP 3000, and many other operating systems including S390 compatible mainframe environments. The disk array is designed to handle very large databases, data warehousing applications, datacenter consolidation and data mining applications for an e-services world.

HP XP512 *Physical View (front)*



NSSO
Stress-Free Storage...Guaranteed

XP512 Specifications

Item			Specifications
Subsystem	Max number of DisK Units (DKU) per subsystem (DKC)		6
	Maximum number of disk drives supported	Per DKU*	96
		Per DKC	512
	RAID level		RAID5 & RAID0/1
	RAID Group		RAID5:3D+1P RAID0/1:2D+2D
	Max number of array groups		126 (Spare disk: 8) 124 (Spare disk: 16)
	Max number of parity groups	All RAID0/1	126
		All RAID5	126
	Max number of spare disk drives		16
	Max number of volumes (CU:LDEV)		4096
	Max number of CUs (0 to F)		16
	Max number of SSIDs (0004 to FFFD)		64
	Support emulation type	Mainframe	3390-1/2/3/9,3380-J/K/E
		Open system	OPEN K/3/8/9/E
	Raw Data Capacity (assumes all 72GB HDDs)		37TB
	LUNs supported per port		256
Controller	Cache memory capacity (256MB DIMMS for 512-16gb; 512mb for 1gb to 32gb)		512MB to 32GB
	Shared memory capacity (256MB increments)		512MB to 1536MB
	Internal path architecture		Hiearchical Star Net (HSN)
	Internal path transfer rate		
Device I/F	DKC to DKU interface		Fibre (FC-AL)/Dual port
	Data transfer rate (MB/s)		
	Max number of HDD/FC-AL		32
	Max number of ACP pairs		4
	Max number of array control processors		Max 32
Channel I/F	Support channel option *HP FC is only short wave	Mainframe	Serial channel: 4S/8S
		Open Systems	Fibre short wavelength: 4GS/8GS Fibre long wavelength: 4GL/8GL
	Data transfer rate (MB/s)	Serial Channel	10/17
		Fibre Channel	100
	Maximum number of CHIP pairs		8
	Maximum number of FC ports		32
	WWN per FC port		128
Non Stop Maintenance	DKC	Control PCB	yes
		CM/SM memory module	yes
		Power supply, fan	yes
		Battery	yes
		Microcode	yes
	DKU	Disk drive	yes
		Power supply, fan	yes
		Control PCB	yes

*In a fully populated system, DKUs R5 and R6 only support 64 HDDs (96-32= 64).

N S S O
Stress-Free Storage...Guaranteed

Continuous Data Availability - the XP family of disk arrays is the first RAID disk array to provide truly continuous data availability. It is designed for nonstop operation and continuous access to all user data. The disk array has no single point of component failure. It is not expected to fail in any way that would interrupt user access to data. The disk array has component and function redundancy, providing full fault tolerance for disk array microprocessors, control storage, control and data buses, power supplies, and cooling fans. The disk array can sustain multiple component failures and still continue to provide full access to stored data. Nondisruptive Service and Upgrades - monitoring software detects failed disk drives or internal hardware and notifies the HP support
center automatically so a service representative can replace the faulty hardware. All hardware subassemblies can be removed, serviced, repaired, or replaced non-disruptively during disk array operation by HP service personnel. All micro-code upgrades can be performed during normal disk array operations, using the disk array's built in service processor (SVP) or the facilities of the host. Alternate pathing is achieved by host fail-over software and/or alternate Fibre Channel paths. The disk array provides up to 32 Fibre Channel ports to accommodate alternate pathing for host attachments.

XP512 DKC & DKU
(Dimensions, Heat & Power)

Item		DKC: 8GB Cache, 2 FC 8-port adapters & additional disk adapater	DKC: Full options	DKU: All 18GB	DKU: All 47GB or 72GB
Dimensions	Width	750mm	750mm	600mm	600mm
	Depth	800mm	800mm	800mm	800mm
	Height	1790mm	1790mm	1790mm	1790mm
Weight		410kg	480kg	440kg	480kg
Heat Output (kW)		1.46	2.32	2.85	3.43
Power Consumption (kVA)		1.57	2.41	3.14	3.72
Air Flow (m3/min.)		18	18	12	12

N S S O
Stress-Free Storage...Guaranteed

NOTE! FC 72GB drives are on CPL as of September 1st, 2000.
Data Integrity- the highest levels of data integrity are provided in the XP512. This is accomplished by using RAID technologies, along with redundant hardware throughout the array. Below is a list of the hardware features that allow the XP512 to achieve it's high availability and data integrity goals.

- RAID 0/1, RAID 5
- Mirrored write cache
- Dual CHIP's and ACP's
- Dual and concurrently active data and control paths through the XP512
- Split power domains on internal data paths
- All boards are hot-pluggable
- All fans, power supplies, controllers are hot-pluggable
- Online firmware upgradeability

Conversion Reference:
304.8mm=12 inches
914.4mm= 3 feet

C to F; (degrees C x 9/5)+32
F to C; (degrees f -32) x 5/9

XP512 SVP

SVP Specifications

Processor		266 or 550MHz
RAM		32 or 64MB
CDROM		TEAC 24E
Floppy		1.44mb
Hard disk (3.2GB)	C Drive	2.1GB
	D Drive	1.1GB
OS		Windows 98
Network Connections	Internal LAN	Yes
	External LAN	Yes
Universal Serial Bus (USB)		Yes
LAN Adapter #1 (Onboard laptop; left side)		Hitachi PCI 10/100 Ethernet
LAN Adapter #2 (SNMP; left side)		USB Port
64MB Flash Memory Card		PCMCIA- bottom slot
Hitachi PCI 56k Modem (Asynchronous)		PCMCIA- top slot
Non-HP Software	PCAnywhere	v9.0
	SNMP Agent	Yes
	WAR-FTPD	v1.66x4

Stress-Free Storage...Guaranteed

Beginning in **January 2001**, all SVPs were shipping with 64MB RAM (this increase in Ram was a requirement for Firmware v01.12.xx). Because the firmware had a 64MB RAM requirement, all XP512's/XP48's that only had 32MB of RAM must be upgraded to 64MB of RAM before FW v01.12.xx or higher can be installed.
Beginning in **March 2001**, all XP512's and XP48's began shipping with the new 550Hz SVP.
Beginning in **August 2001**, customers with XP512/XP48 running the old 266MHz SVP can opt for an SVP upgrade to the 550MHz SVP. The cost of this upgrade is not free (it will cost around $3-4k; please check the Corporate Price List for a definitive quote) for the customer to upgrade to the new SVP. HP is offering this upgrade solely for the performance gain that can be recognized by Command View XP software connecting to the XP array with the 550MHz SVP; it is not a requirement for the XP512/48 to function properly, regardless of whether Command View XP is in use or not.

Please see the XP512 Software training slides to learn more about the SVP's use of SNMP, the USB port and connectivity to the external network.

Please see the XP512 Software training slides in the next module to learn more about the SVP's use of SNMP, the USB port and connectivity to the external network.

XP512 DKC Architecture

Contents

- DKC Internal Architecture Overview
- CHIP
- ACP
- Cache & Shared Memory
- DKC Front & Rear Logic Gate

NSSO
Stress-Free Storage...Guaranteed

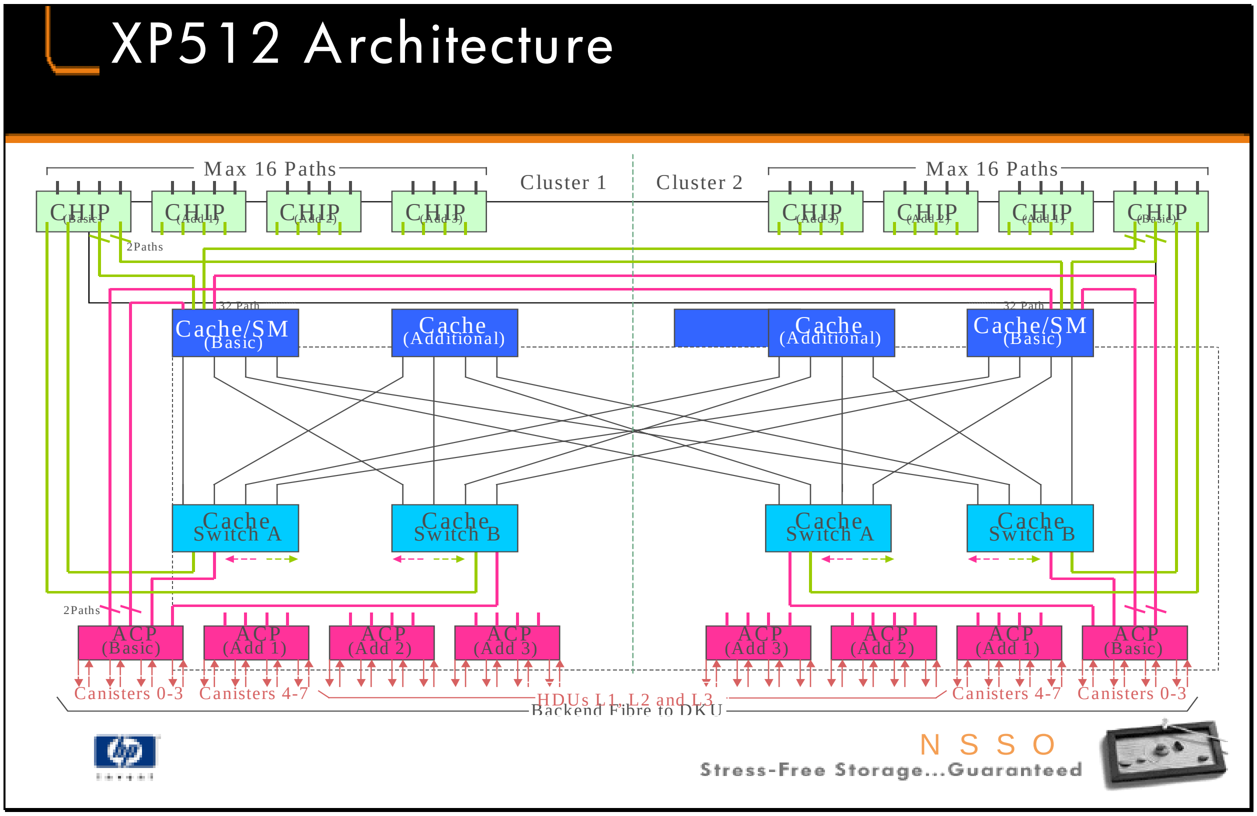
Definition of Terms for the XP512

- **DKC** - cabinet holding controller, SVP, ACP, CHIP's (one per XP512)
- DKU - cabinet holding disk canisters and disks
  (up to 96 Disks per DKU, up to 6 DKU's per XP512)
- SVP - service processor for HP support access
- CHIP pairs - ports connecting servers to the array.
  (up to 32 Fibre channel and ESCON ports per XP512)
- ACP pairs - provide connections to the disk drives .
  (Controls reads/writes, media protection, dynamic spares, RAID 0/1,5)
- Crossbar Switch - high speed (6.4 GB/sec.)point-to-point connections between internal components.
- Mirrored Cache - up to 32 GB sold in 2 GB increments
- Shared Memory - for system configuration tables, mapping
  (up to 1.5GB sold in 256 MB increments)
- Dynamic Spares - spare drives for fault resilience
  (up to 16 dynamic spares)

N S S O
Stress-Free Storage...Guaranteed

INTRODUCTORY DEFINITONS
-CHIP (Client Host Interface Processor): support the connections from the host servers to the array.
-ACP (Array Controller Processor): support the physical disk access leading from the DKC.
-SVP (Service Processor): laptop mounted inside the DKC responsible for configuring the array, monitoring and reporting status conditions.
-Shared Memory: Index of memory offsets that point to data held in cache.
-CACHE: Holds data not written to or just read from hard disk

The principle characteristics of the HP SureStore E product:
-Fault tolerant architecture: All the active components are duplexed, thus eliminating a single point of failure (CHIPs, ACPs, Cache, Memory, Disk and data path). Not shown in this diagram, but all power is also duplexed.
-Multi-controller: This device may be connected to many different computers (mainframes,NT, Unix). Connect options supported: FC, ESCON.
-Reliability, scalability: There is no disruptive operation in this machine: CPUs, memory, cache, disk may be changed, upgraded without stopping the machine.



Cache-the cache is one of the most critical portions of any disk array subsystem. All read or write operations move data through the cache. The cache must be robust to ensure that there are no data integrity problems with the flow of data through the cache, and the cache needs to be fast enough to not become a bottleneck in system performance. The XP512 has optimized both these requirements by providing a duplex crossbar interface to the cache. The crossbar interface to cache allows multiple simultaneous data paths in and out of cache. The duplex write feature between different power domains of the cache provides the robustness required to ensure no data loss will occur between the host and physical disk storage mechanisms. Both the XP256 and XP512 architectures have separate and independent shared memory from the cache memory. The shared memory is used for system configuration tables. The configuration tables are used for system components, physical to logical disk mapping of LUNs, and for identification of RAID levels for any given LUNs. The shared memory also keeps track of the cache hit and miss rates and is used to control cache pointers that allow for virtual contiguous
access of the cache to either the CHIP (or ACP) interface. The connection between shared memory and the ACP's and CHIP's is a direct point-to-point connection. The data crossbar provides the data path between the host connection and physical disk drives through the cache. Having separate shared and cache memory allows the cache to entirely be used for data transactions to and from the physical disk drives.

The XP512 supports up to 32GB of cache. A dynamic duplexed write cache is used for data transfers. All cache writes are duplexed between two portions of the cache that are on different power boundaries. This maximizes data integrity between the time that the data leaves the host and data is written onto a physical drive.
Shared Memory
Two times more cache supported than in the XP256
Four direct connections from each CHIP to shared memory.
Each port to shared memory via direct connect buses provide_____ bandwidth.
Up to 64 ports into shared memory (from ACP to SM and CHIP to SM), thus achieving _____ bandwidth (_____ x 64= _____).
Shared memory is split across power domains
Each CHIP's direct connection to shared memory is split into two connections each shared memory power domain
Inbound control path: CHIP-->Shared Memory-->ACP
Cache
Two times more cache supported than in the XP256
Point to point connection from the CHIP to the cache switch (aka crossbar).
Two ports from each CHIP and ACP into the crossbar switch.
Each crossbar switch is on a separate power domain
Each port to the crossbar switch provides multiple ______ datapaths in and out of cache.
Supports up to 16 ports to cache that achieves ______ bandwidth.
Inbound data path: CHIP--> cache switch--> ACP--> HDDs

INTRODUCTORY DEFINITONS
-**CHIP** (Client Host Interface Processor): support the connections from the host servers to the array.
-**ACP** (Array Controller Processor): support the physical disk access leading from the DKC.
-**SVP** (Service Processor): laptop mounted inside the DKC responsible for configuring the array, monitoring and reporting status conditions.
-**Shared Memory:** Index of memory offsets that point to data held in cache.
-**CACHE:** Holds data not written to or just read from hard disk

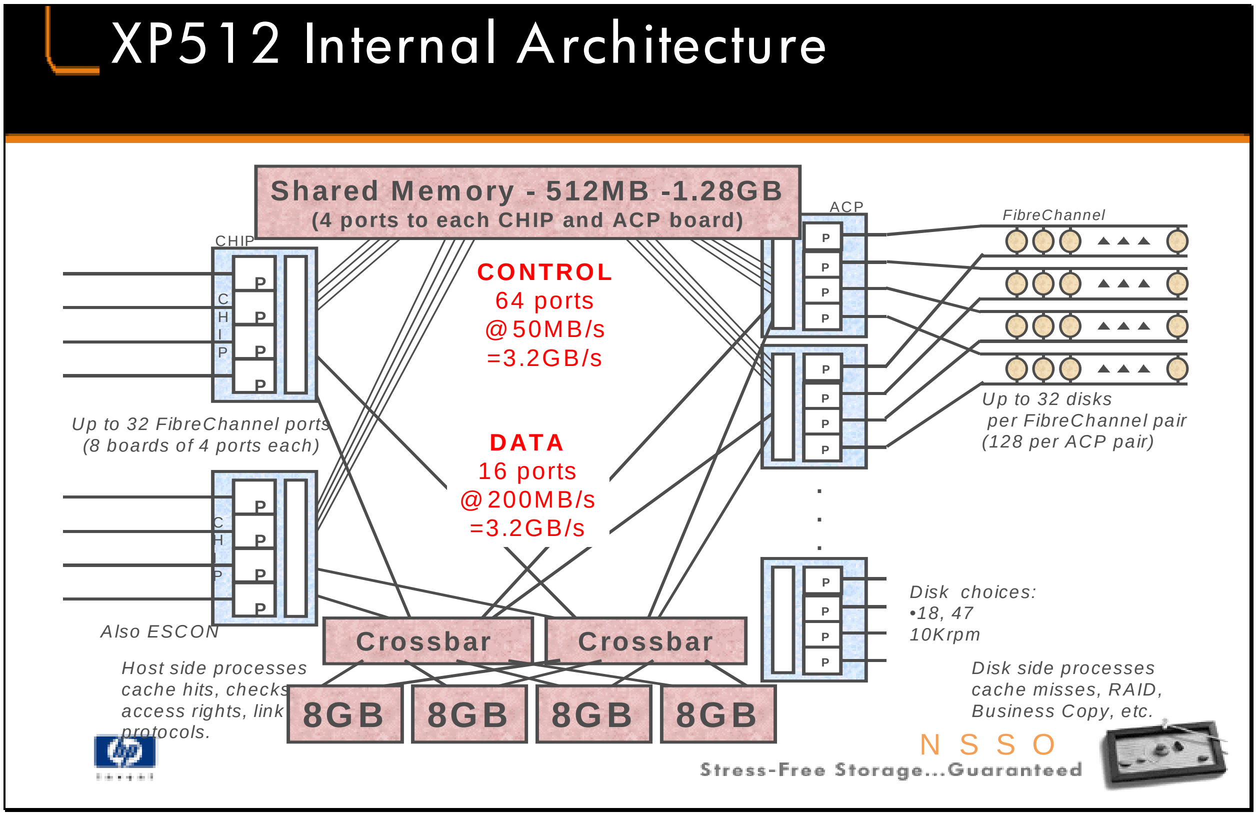
The principle characteristics of the HP SureStore E product:
-**Fault tolerant architecture**: All the active components are duplexed, thus eliminating a single point of failure (CHIPs, ACPs, Cache, Memory, Disk and data path). Not shown in this diagram, but all power is also duplexed.
-**Multi-controller**: This device may be connected to many different computers (mainframes,NT, Unix). Connect options supported: FC, ESCON.
-**Reliability, scalability**: There is no disruptive operation in this machine: CPUs, memory, cache, disk may be changed, upgraded without stopping the machine.

XP512 Crossbar Switch

- XP512 is the first array utilizing crossbar technology (provides 3x the performance of competitive EMC products)
- Commonly used in high performance computer systems requiring very high bandwidths (Consider the K vs. the V-Class Unix servers)
- Optimizes throughput of system data paths between all XP512 components
- 6.4 GB/sec total bandwidth
- Provides multiple dedicated paths to each XP512 component (cache, shared memory, ACP's, and CHIP's)

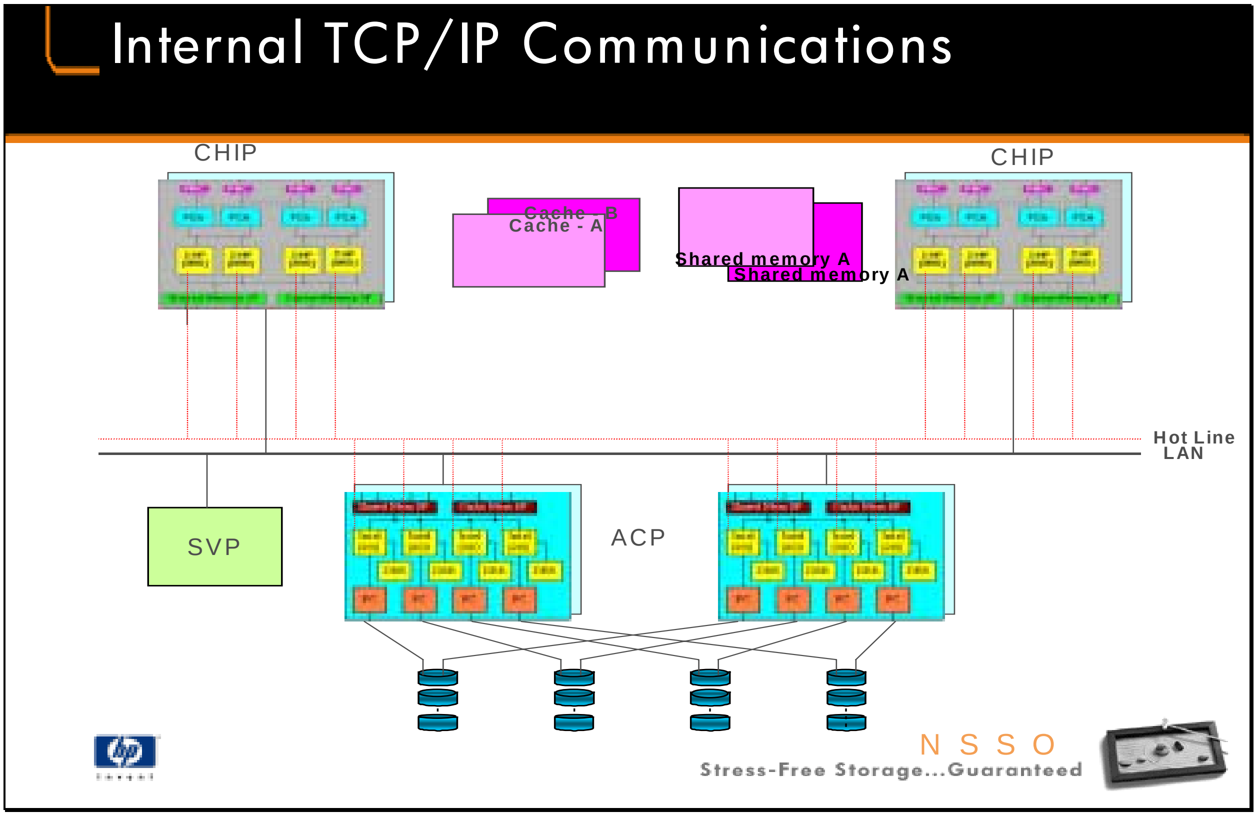
N S S O
Stress-Free Storage...Guaranteed

The XP512 uses the hardware and software architecture foundation of the successful XP256. From that foundation the internal architecture was redesigned around state of the art crossbar switches and point-to-point technologies to provide the XP512 with leadership performance and design. The internal crossbar provide the interconnect and high bandwidth links between the CHIPs, ACPs, and cache. The internal architecture of the XP512 is very different from the XP256's shared bus based backplane. The XP256 uses a back-plane architecture that has multiple CHIP's and ACP's sharing resources on the four internal buses (2 data and 2 control). The XP512 uses dedicated point-to-point and crossbar interconnect technologies for it's internal connections. Point-to-point and crossbar back-planes are today's industry leading technologies used in complex high performance computer system designs. A few years ago point-to-point and crossbar technologies were only found on high-end, high-performance data center computing platforms like the HP 9000 high-end Enterprise Servers that require very high internal bus bandwidths to meet its performance requirements. Recent VLSI technology advances have allowed crossbar and point-to-point technologies to be put into a broader range of computing applications. Today most high end server platforms use some sort of crossbar or point-to-point technologies to maximize on the performance throughput of data paths through their product while many mid-range server platforms continue to use shared bus based architectures. The XP512 is the first storage array to take advantage of point-to-point and crossbar technologies, creating a new class of high-end enterprise-class storage.



Shared Memory
Two times more shared memory supported than in the XP256
Four direct connections from each CHIP to shared memory.
Each port to shared memory via direct connect buses provide 50MBs bandwidth.
Up to 64 ports into shared memory (from ACP to SM and CHIP to SM), thus achieving 3.2GBs bandwidth (50MBs x 64= 3.2GBs).
Shared memory is split across power domains
Each CHIP's direct connection to shared memory is split into two connections each shared memory power domain
Inbound control path: CHIP-->Shared Memory-->ACP
Cache
Two times more cache supported than in the XP256
Point to point connection from the CHIP to the cache switch (aka crossbar).
Two ports from each CHIP and ACP into the crossbar switch.
Each crossbar switch is on a separate power domain
Each port to the crossbar switch provides multiple 200MBs datapaths in and out of cache.
Supports up to 16 ports to cache that achieves 3.2GBs bandwidth (200MBs x 16=3.2GBs).
Inbound data path: CHIP--> Cache Switch--> Cache---> Cache Switch--->ACP--> HDDs

6.4GB/s throughput!!
DATA= 16 ports @ 200MB/s= 3.2GB/s
CONTROL= 64 ports @ 50MB/s= 3.2GB/s

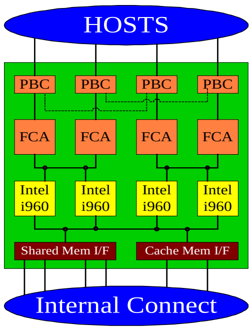


The LAN line is for TCP/IP communications from each board to the SVP. This line connects from the SVP to all major components. It is used to report back errors, firmware versions and status information. It also allows for firmware upgrades to occur. Because TCP/IP is the protocol, each board has its own IP address; hence the reason simple board swapping is not possible without first deconfiguring. A simple illustration of this setup can be seen by viewing the HOSTS file of the SVP and pinging the internal components defined within.
The SVP is not duplexed (the laptop is not duplexed, power to the single laptop is not duplexed and LAN connectivity into the laptop is not duplexed), however the array can still function without the SVP and the LAN connection to the SVP.

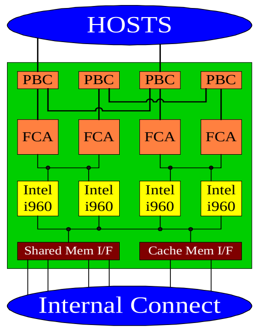
The Hot Line bus connects to all boards; it is used to detect error.

This illustration also clearly displays the data path redundancy with the XP architecture (starting from the CHIP to the physical HDD).

HP XP512 *CHIP Functions*



(Standard Mode)



(High Performance Mode)

- CHIP pairs provide host connection
  - Up to 4 CHIP Pairs per 512 Array
- CHIP Pair Hardware Components
  - 4 Agilent Tachyon TS FC Adapter (FCA) chips
  - 4 Intel i960 controllers (80 MHz)
  - 4 Point-to-point connections for Shared Memory
  - 2 Crossbar channels for Cache
- CHIP Functions
  - XP48 & 512 CHIP firmware provides 2 modes: "Standard" and "High Perf"
- XP512 CHIP firmware has simplified tuning through automatic odd/even LUN mapping
- CHIP's are parts compatible with XP512 but have different part numbers
- XP48 CHIP pairs available:

4-port FibreChannel	A5925A
8-port FibreChannel	A5926A
4-port ESCON	A5923A
8-port ESCON	A5924A

NOTE:
For SCSI connectivity, a SCSI-to-FC bridge product (A5814A) is available

N S S O
Stress-Free Storage...Guaranteed

CHIP (Client Host Interface Processor)
1. Conversion between host commands and internal commands.
2. Communicate with the ACP to use cache for read/write I/O
3. Communicates with Shared Memory to define location of data request held in cache
4. Reads and writes data to cache
5. Address mapping of logical devices to SCSI target/Logical Unit Number addresses.
6. Monitors access patterns

Note: Client Host Interface Processor or CHIPs is HP's name for the hardware of the board that connects the data host to the array via FC, or Escon (the XP512 does not support SCSI CHIPs like the XP256). HDS and much of the software's GUI will use the term Channel Host Adapter (CHA) for CHIP. The i960 processors found on the CHIP are the Client Host Processor or CHP

CHIP ***pairs*** can provide the following connections to hosts:
-4 or 8 ESCON connections (allow mainframe host connection or XP512 to XP512 connectivity with Continuous Access)
-8 FC

On each CHIP board a pair of interfaces is dynamically switchable between two microprocessors. Where ESCON CHIPs are installed, the ESCON interface can be configured either as a link to a S/390 host or for connection to another XP512 for use with the Continuos Access XP - remote mirroring product.
All host connected to a CHIP are referred to as **"Inband"**

Automatic Odd/Even LUN mapping in Firmware
4, 80 MHz i960 processors vs. 4, 66 MHz i960's with the XP256

Standard Mode

2 independent "halves" 10k IOP blocks per CHIP card

High Performance Mode

All 4 i960's work to service IO's for 20k IOPS

Ports C & D disabled

Maximum IO's achieved through LUN mapping and sharing "virtual LUN's" with A & B

(balance between all 4 ports - but ports A&B only active)

Caution one port, A for instance can take all bandwidth in HP mode

LDEV Count (total): 4096

Each type of CHIP board contains a different architecture.

PBC: Port Bypass Circuit.
FCA: Fiber Channel Adpater
Intel I960: On board micro processor.
FM= Flash Memory. Each CHIP board contains four Intel 960 processors that are upgradable via flash. The Load balancing of I/O between the i960's is done by dividing load based on odd or even numbered LUNs.

Failover does not occur within a single CHIP board; failover between CHIP pairs (all are installed in pairs). However, if one of the multiport boards fails on only one port, the remaining ports will still operate (this assumes the failure occurs in the FCA, FM or I960 processor that services that one port).

HP XP512 CHIP Configuration

- Three types of channel interfaces available (ports per pair):

8-port FibreChannel	**A5957A**
4-port ESCON	**A5953A**
8-port ESCON	**A5954A**

- Maximum of four CHIP Pairs per DKC
- Fully intermixable (Fibre Channel, ESCON)
- Up to two ESCON CHIP Pairs (16 ports) supported using Continuous Access XP

NSSO
Stress-Free Storage...Guaranteed

Client Host Interface Processor or CHIPs is HP's name for the hardware of the board that connects the data host to the array via FC, SCSI or Escon (the XP512 does not support SCSI CHIPs like the XP256). HDS and much of the software's GUI will use the term Channel Host Adapter (CHA) for CHIP. The i960 processors found on the CHIP are the Client Host Processor or CHP.
CHIP (Client Host Interface Processor)- the CHIP provides a connect point for the host to allow access to some disk space on the array. Any host connection to the XP512 array will need to connect to at least one port on a CHIP board. Typically, a host will have a dual port connection to two different CHIP boards. These two different CHIP boards that are connected to the host reside in the XP512 on independent power domains. In the event of an internal power supply failure or a CHIP board failure, the other CHIP board will be provide the path from host to the physically stored data. The CHIP's primary function is to process host commands and signal the ACP to read/write memory. Additional features of the CHIP are to access and update the cache track directory, monitor data access patterns, and to emulate host device types. Each CHIP provides four host connections except the 2 port ECSON CHIP board. The host connection goes through the FCA (Fibre Channel Adapter) interface and converts the host commands to internal commands. The FCA interface consists of a port bypass circuit (PBC) and a Tachyon TS (TacLite) Fibre Channel interface IC. The port bypass circuit is used to convert from

'Standard Mode' to 'High Performance Mode'. The host connection through the FCA is controlled by it's own i960 microprocessor that are connected in pairs. Each i960 backs the other i960 up in case of failure and they are also used together for load sharing between two input ports in Standard Mode'. The pair of i960's automatically will split the incoming transaction into even and odd LDEV's to increase the throughput through the CHIP and to cache memory. Each i960 processor also manages the host interface along with the point-to-point connections to shared and cache memory. There are three different types of CHIP boards that are supported in the XP512 array. A total of 4
CHIP pairs are available (8 total boards) for host interconnect. These boards are always added in pairs and are ordered as pairs:

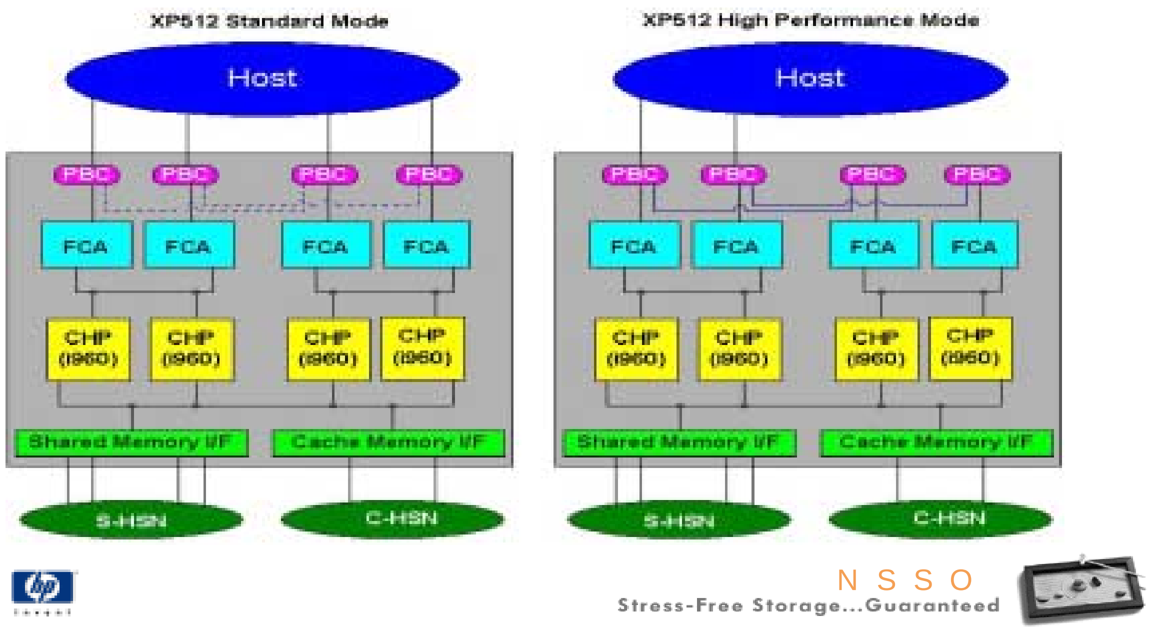
· 4 native Fibre Channel ports per CHIP board (8 Fibre Channel ports per CHIP pair)
· 2 ESCON ports per CHIP board (4 ESCON ports per CHIP pair)
· 4 ESCON ports per CHIP board (8 ESCON ports per CHIP pair)

The four-port Fibre Channel CHIP board can be configured into two different operating modes, either in 'Standard Mode' or 'High Performance Mode'. In 'Standard Mode' the CHIP is logically split into two halves with two i960's used to control two Fibre Channel ports. In "High Performance Mode' the CHIP is operating as a single unit. In "High Performance mode' the port bypass circuit is enabled and two of the four external Fibre Channel host connect ports are disabled. The circuitry on the CHIP then uses all four i960's to control just two Fibre Channel host ports. This mode provides a very high throughput through the two host ports. This cuts the number of available host connect ports in half when using high performance mode.
CHIPs must be purchased and installed in pairs, but the pairs themselves can be intermixed (i.e.- 1 FC pair and 1 ESCON pair in the same array). As we will see later in the slides, CHIPs are always installed in the front of the DKC (known as the front logic gate).

Data transfers to and from the CHIPs and CACHE is done by DMA bursts.

In general, most HP documentation and general speak talks about CHIPs in pairs and not individual CHIP boards. For some this might be confusing at first- a single FC CHIP contains four ports (this is the maximum number of ports supported on a single FC CHIP in the XP512), but the product number list this CHIP as a 8 port FC CHIP (because it is sold and must be installed in pairs). However, a CE is performing a CHIP board replacement, it would be purchased as a 2 port FC CHIP replacement board.

XP512 FC Channel Host Interface Processor

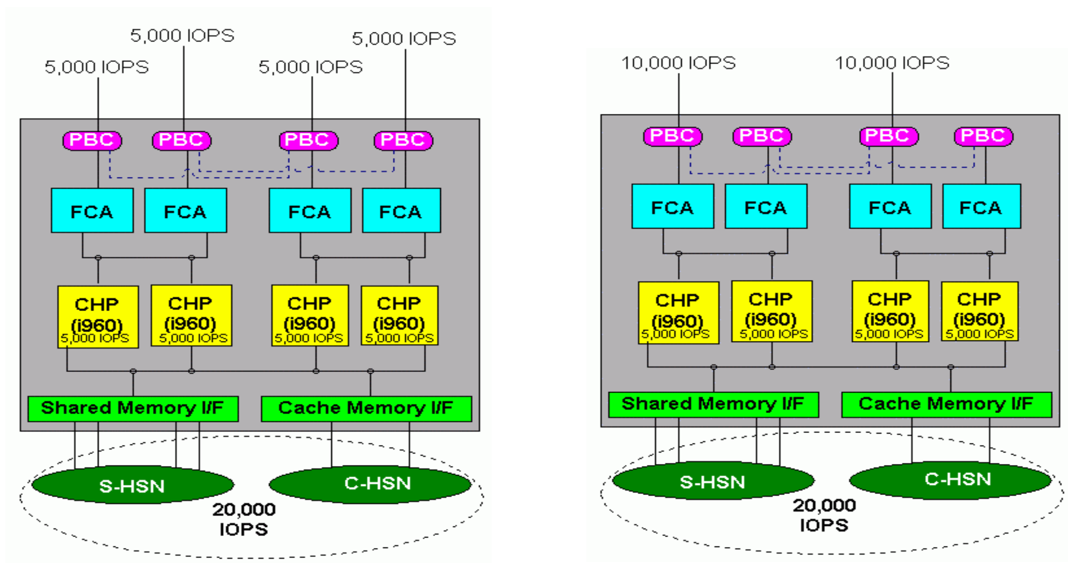


The four-port FC CHIP can be configured into two different operating modes.
1. Standard Mode: The CHIP is logically split into two halves; the individual i960s are grouped into one pair to load share and act as backup in the event of one i960's failure. Just like the XP256 CHIP, incoming throughput is split between even and odd LUNs.
2. High Performance Mode: As opposed to splitting the CHIP into two halves, this mode enables a port bypass circuit disabling two of the available four FC ports. This forces all four i960 processors to dedicate to only one FC port or two FC ports if connected.

Q: How does High Performance Mode help the CHIP gain performance?
Each i960 is capable of processing 5000 I/Os per second (IOPS). If all four ports on one CHIP are utilized (CL1A through CL1D) in standard mode the each port is capable of 5000 IOPS per port. If only every other port of a single chip were utilized (ie- CL1A and CL1C or CL1B and CL1D) the CHIP could produce 10,000 IOPS per port (this is possible because each port now has two i960s processing I/O to the port). By placing the CHIP into High Performance mode the CHIP forces only two out of four external ports to be active, thus yielding a 10,000 IOPS maximum for the two active ports. If the CHIP is in high performance mode and we only use one of the active two ports, that one port could achieve 20,000 IOPs (4 i960's x 5000iops =

20k total IOPs). High Performance CHIP mode operation is not the same thing as High Performance Cache mode of operation.

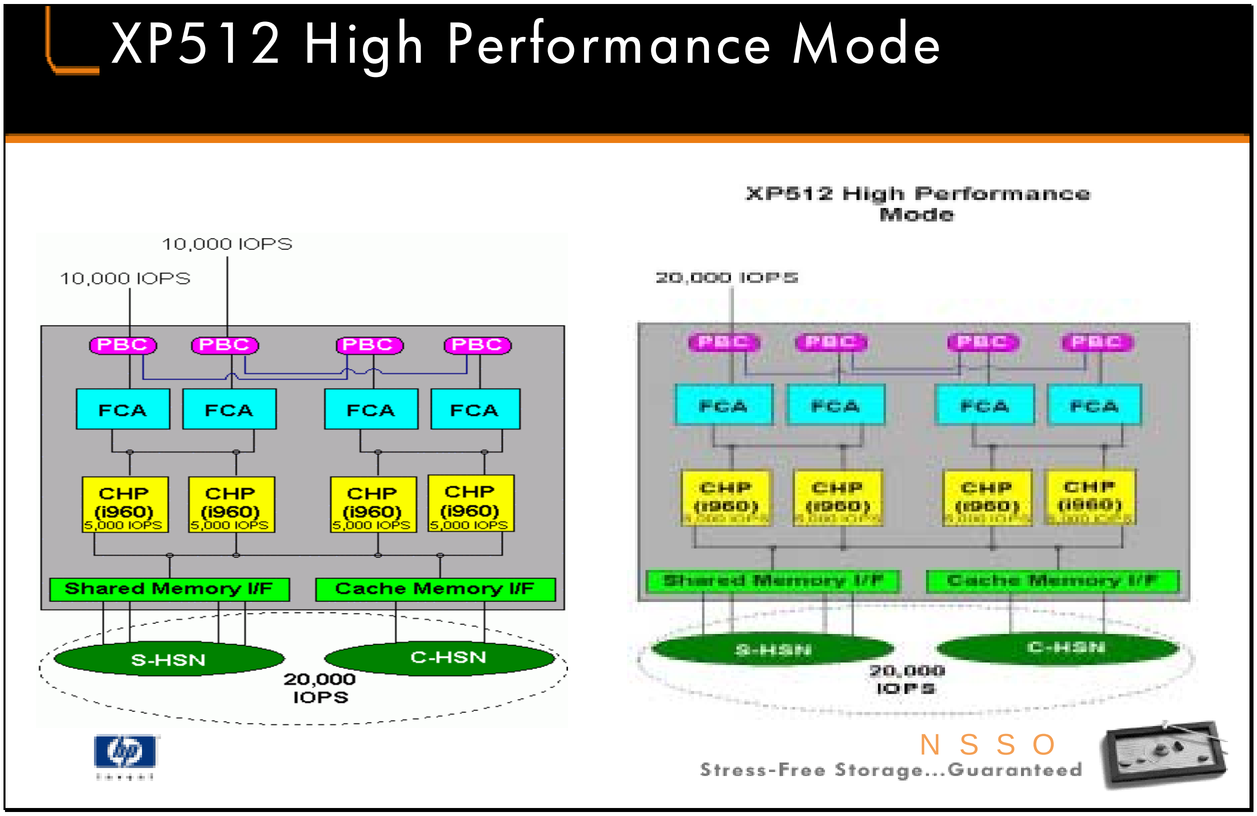
XP512 CHIP: Standard Mode



N S S O
Stress-Free Storage...Guaranteed

Shown above are maximum IOPs achievable by each CHIP in respective to active FC port connections. The increase in total IOPs is made possible by using the new TachLite chip and the increase in processor speed from 60MHz to 80MHz.

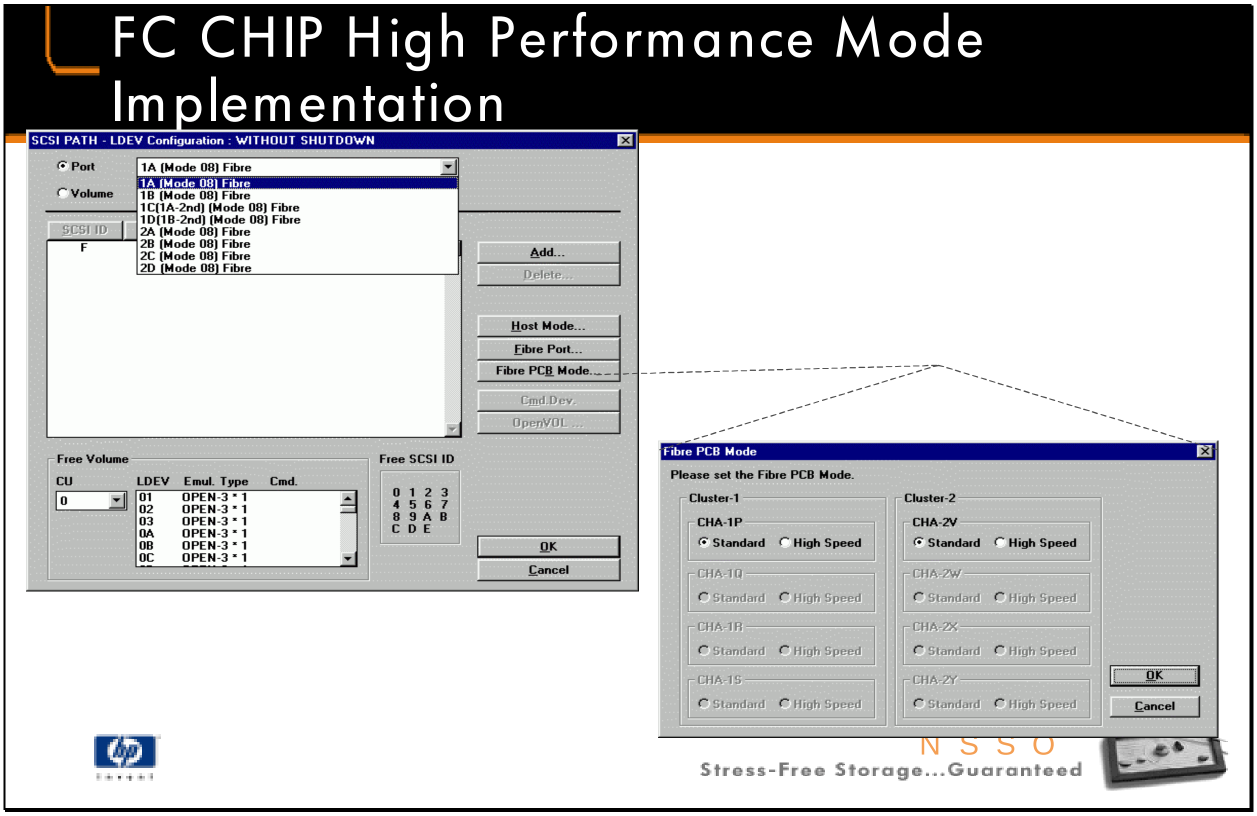
TachLite IC utilizes HP's TACHYON IC architectural approach of a hardware-based, state machine design. Unlike other implementations in the industry, the TACHYON IC architecture avoids embedded-microprocessor performance issues, such as limited parallelism and limited processor cycles per second. TACHYON designs today sustain 100MB/s throughput and deliver 10,000 I/Os per second. Additionally, TachLite IC builds on TACHYON IC performance levels by integrating 64-bit PCI interface and providing the highest levels of I/O automation in hardware. TachLite IC customers are today achieving 16,500 I/Os per second in 32-bit PCI systems.



As you can see, the XP512's high performance mode is best used if you must achieve 20,000 IOPs for one port. When using the two active ports, notice that high performance mode setting yields the same result as having the CHIP in standard mode using only the first and third ports.

Recommendations?
1. Use high performance mode if you know you need up to 20,000 IOPs for one port and you can afford to lose 3 available ports.
2. Use high performance mode if you need 10,000 to 20,000 IOPs on one port. Performance is not free. For example, in high performance mode one port can go to 18k IOPs thus leaving the other port to only go to 2k IOPs.
3. If you up to 10k IOPs on two ports, then do not set to high performance mode. Instead only physically connect FC cable to the first and third ports on the CHIP in standard mode. This allows for greater flexibility in CHIP configuration should your needs change (ie- in case you need all four ports later).
Rules for LDEV to port mapping: I/O is split between even and odd LUN numbers and not even and odd LDEV numbers.
High Performance Mode:
-Balance even/odd LUNs between the first/third and second/fourth ports.
Standard Mode:

-Balance even/odd LDEVs between the first/second and third/fourth ports.



Implementation: Set a CHIP to high performance mode by clicking on "Fibre PCB Mode" from within LUN Manager. With high performance mode CHIP implementation, the array disables or shuts down external FC ports at the third and fourth FCA on each CHIP, however the processors tied to those ports are still enabled. For example, assume we are talking about the CHIP installed in slot P. High Performance mode would leave external ports A and B active and shut down ports C and D. However, it is still possible and necessary to use LUN Manager and map LUNS to all four ports CL1-A, CL1-B, CL1-C and CL1-D. The processors that sit behind these ports are still active, thus mapping to the logical port names is necessary. The XP512 sets the physical mapping to the processors behind the port and **not** the external port input.

The two visible signs that a CHIP is in high performance mode:

1. The front panel LEDs will not hold a steady green for active ports (the first two, for example CL1-A and CL1-B) and no light on the disabled ports (the second two port, for example CL1-C and CL1-D).
2. From within LUN Manager, we will see the "-2nd" on disable FCA ports being utilized for high performance mode.

FC CHIP High Performance Mode Implementation

LUN#: 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f
PORTS: CL1-A CL1-B CL1-C CL1-D
FC CHIP Setting: High Performance Mode

Scenario: FC is physically hooked up to ports CL1-A and CL1-B and not to CL1-C and CL1-D.
Question: What ports do I map each LUN to achieve a theoretical balanced I/O of 20k for CL1-A and CL1-B?

Answer:
CL1-A: 00 01 02 03
CL1-B: 04 05 06 07
CL1-C: 08 09 0a 0b
CL1-D: 0c 0d 0e 0f

NSSO
Stress-Free Storage...Guaranteed

Why does this work?
Step1: Divide the total number of LUNs by four thus creating four separate groups (ie- Group 1 through 4); each group created should contain equal number of even/odd LUNs.
Step2: Assign Group 1 to the first port CL1-A
Step3: Assign Group 2 to the second port CL1-B
Step4: Assign Group 3 (CL1a-2nd) to port CL1-C
Step5: Assign Group 4 (CL1b-2nd) to port CL1-D

Assumptions: It must be understood that it is assumed one of the following is true…..
1. Each LUN is receiving the same number of I/Os
2. Each group of LUNs (Group 1-4) collectively is receiving the same number of I/Os.

FC CHIP High Performance Mode Implementation

LUN#: 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f
PORTS: CL1-A CL1-B CL1-C CL1-D
FC CHIP Setting: High Performance Mode

Scenario: FC is physically hooked up only to port CL1-A and not to CL1-B, CL1-C and CL1-D.
Question: What ports do I map each LUN to achieve a theoretical I/O of 20k for CL1-A?

Answer:
CL1-A: 00 01 02 03 04 05 06 07
CL1-B:
CL1-C: 08 09 0a 0b 0c 0d 0e 0f
CL1-D:

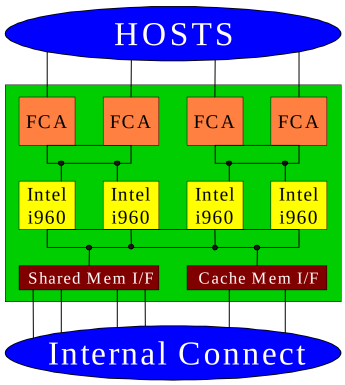
N S S O
Stress-Free Storage...Guaranteed

Why does this work?
Step1: Divide the total number of LUNs by two thus creating two separate groups (ie- Group 1 and Group 2); each group created should contain equal number of even/odd LUNs.
Step2: Assign Group 1 to CL1-A.
Step3: Assign Group 2 to CL1-C.

As I/O begins, even and odd LUN division occurs automatically between CL1-A / CL1-B and automatically between CL1-C and CL1-D via the port bypass circuit that High Performance Mode enables on the CHIP.

Assumptions: It must be understood that it is assumed one of the following is true…..
1. Each LUN is receiving the same number of I/Os
2. Each group of LUNs collectively is receiving the same number of I/Os.

XP512 CHIP Pairs (summary)



XP512 CHIP

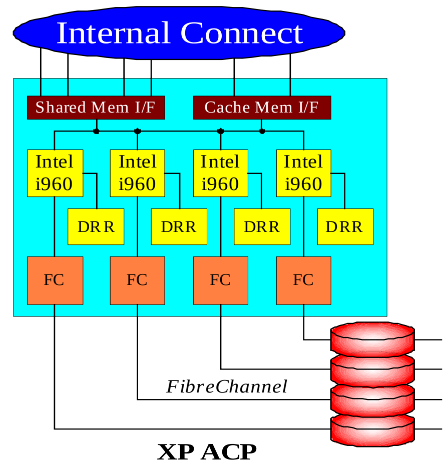
- CHIP pairs connect host servers to the DKC
  - Up to 4 CHIP Pairs per XP512
- CHIP Pair Hardware Components
  - 4 Agilent Tachyon TS FC Adaptor (FCA) chips
  - 4 Intel i960 controllers
  - 4 Point-to-point connections for Shared Memory
  - 2 Crossbar channels for Cache
- CHIP Functions
  - XP512 CHIP firmware provides 2 modes: “Standard” and “High Perf”
- XP512 CHIP firmware has simplified tuning through automatic odd/even LUN mapping
- XP512 CHIP pairs available:
  - 8-port FibreChannel A5956A
  - 4-port ESCON A5953A
  - 8-port ESCON A5954A

NOTE:
For SCSI connectivity, a SCSI-to-FC bridge product (A5814A) is available.

NSSO
Stress-Free Storage...Guaranteed

HP XP512 ACP Functions

- Connect the disk drives to the DKC
  - Up to 4 ACP Pairs per XP512 Array integral to the XP48
  - Up to 128 100 MB/s FibreChannel disks
  - 8 FC-AL loops per ACP pair (4 per card)
  - Uses Agilent Tachyon-TL FibreChannel Interface chips
- ACP Hardware Components:
  - 4 Intel i960 controllers
  - 4 Data Recovery and Restore (DRR) Parity Generator chips (RAID5 parity)
  - 4 Point-to-point connections for Shared Memory
  - 2 Crossbar channels for Cache
- ACP Functions:
  - Read/Write Control to Disks
  - Read Miss/Write Staging from Cache
  - Media Protection (disk scrubbing, Dynamic Spares, Dynamic Data rebuild)
  - Mirroring Control (RAID 0/1)
  - Hardware RAID5 parity generation
- HP Product Number: A5964A - Included in Frame with XP48



N S S O
Stress-Free Storage...Guaranteed

ACP (Array Control Processor)-the ACP functions are to deal with read/write functions to the disk, read miss staging and write de-staging from the cache. The ACP's also handle the task of media protection. Media protection is done by dynamic spares, mirrored storage (RAID 0/1), dynamic data rebuild, and hardware RAID 5 parity generation.

To accomplish the task of hardware parity generation and controlling the data flow to the physical drives, the ACP has four i960 controllers each linked to a DRR (Data Recovery and Restore) parity generator chip. The i960 controllers communicate with the shared and cache memory interface, and the Tachyon TL (TacLite) Fibre Channel I/F chips. The DRR's are the hardware RAID 5. parity generators that do the work of generating and checking the parity for RAID 5 groups, and duplicating data copies for mirrored storage (RAID 0/1). Each Fibre Channel interface has a dedicated i960 and parity circuit.

A minimum XP512 system requires a single ACP pair. The two ACP boards reside on separate power domains like the CHIP boards, and cache memory. In the event of an ACP failure the other ACP will take over all transactions to the physical disks. During normal operation both ACP's are used to load balance the data transfer between the cache and the physical disks. Each ACP pair can control up to 128 physical disks including both data and dynamic spare disks on four Fibre Channel arbitrated loops. To get 128 physical disks on a single ACP pair, the disks need to be populated in three DKU's on a single side of the DKC. In a full XP512 system configuration, up to four ACP pairs can be installed into the DKC.

ACP= Access Control Processor

Since all data I/O (reads and writes) pass through cache, the ACP pairs handle data transfers between the cache and disk. It accomplishes this by performing logical-to-physical address translation (parity controlled by the ACP).

Each DKC is sold with 1 ACP pair and expandable to 4 ACP pairs (to expand to 3 or more ACP pairs requires an additional internal FC cable set A5964A). As we will see later in the slides, ACP pairs are only installed in the back of the DKC.

Disk Scrubbing: Disk are accessed regularly to be sure they are still alive.

NOTE: The actual physical path to a full disk write or read occurs as follows...

CHIP-->CACHE-->ACP-->ALB -->disk media

***The Actuator Line Buffer (ALB) is a 4mb RAM module on the HDD. Data passes through the ALB but is not "cached" or held on the ALB- data writes straight to disk.*
Each pair of ACP's provides eight 100MB/S FC interfaces.

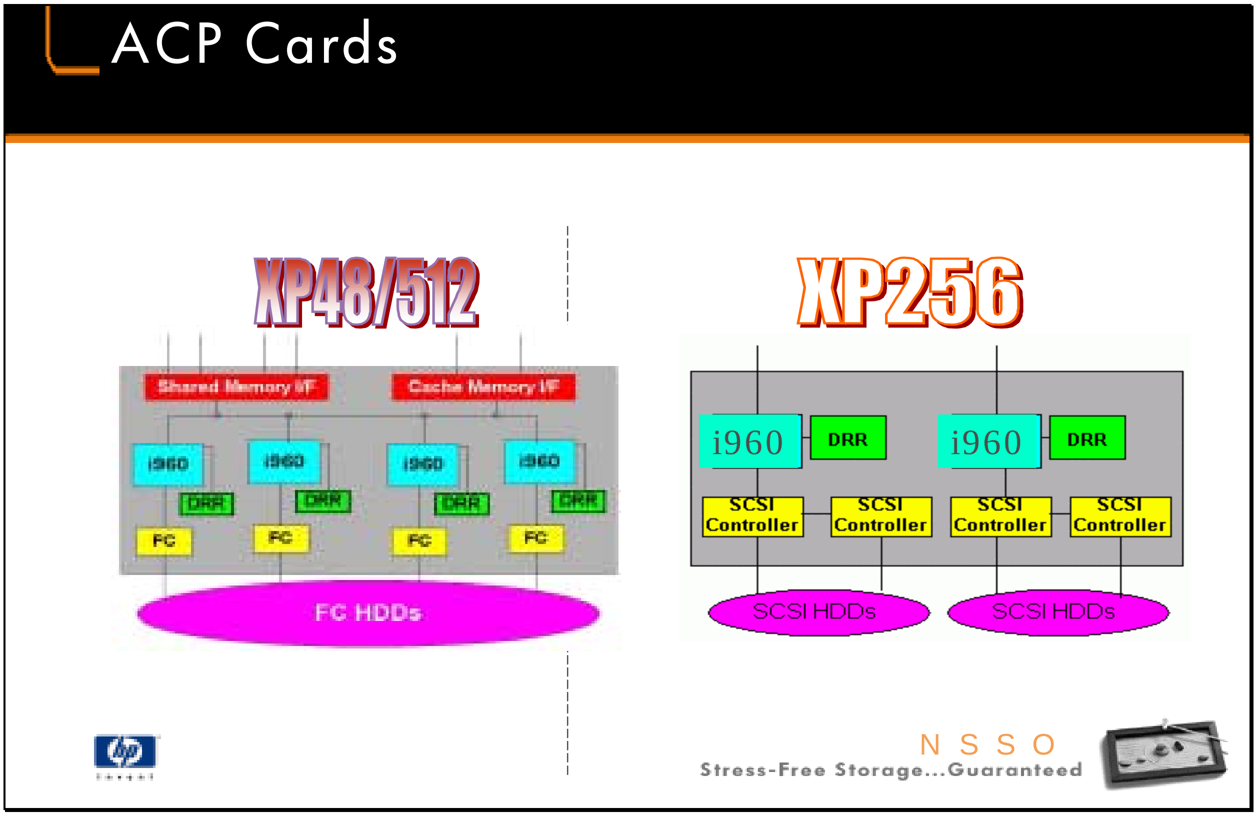
FM= Flash Memory. Each ACP chip contains four Intel 960 processors that are upgradable via flash. Using program interrupts, the i960 processors run two programs:

Disk Master Program: logical-to-physical address translation,RAID control functions and cache control.

Disk Slave Program: Parity control, drive data transfer control and SCSI control functions.

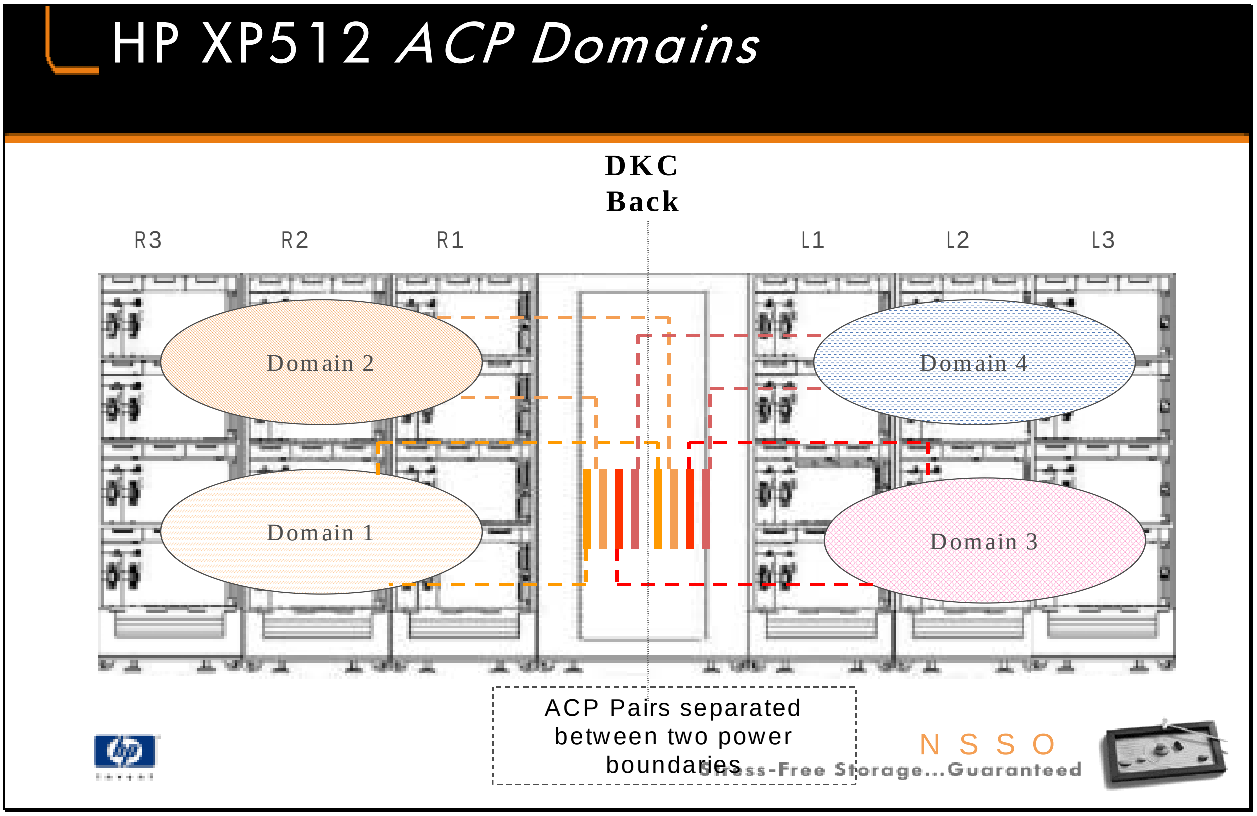
DRR= Data Recovery Reconstructor: hardware parity generators that generate and check parity for RAID5 groups and provide duplicate data copies for RAID0/1 groups.

FC- Tachyon (TacLite) FC Interface to FC HDDs.



Just like the XP256, the XP512's DKC holds up to 4 pair of ACP boards in the rear logic gate. The obvious difference in the ACP architecture between the XP256 and the XP512 is the dedicated i960 processor tied to each backend connection of the XP512's ACP board (one to one relationship). The second obvious difference is the number of paths into the DKC (paths to shared memory and cache) has more than doubled with the XP512 (the new paths are also dedicated connects to either shared memory or cache). In the XP256, only two paths with redundancy go to the data and control buses. DRR- hardware parity generators that generate and check parity for RAID5 groups and provide duplicate data copies for RAID0/1 groups.
FC- Tachyon (TachLite) FC Interface to FC HDDs.

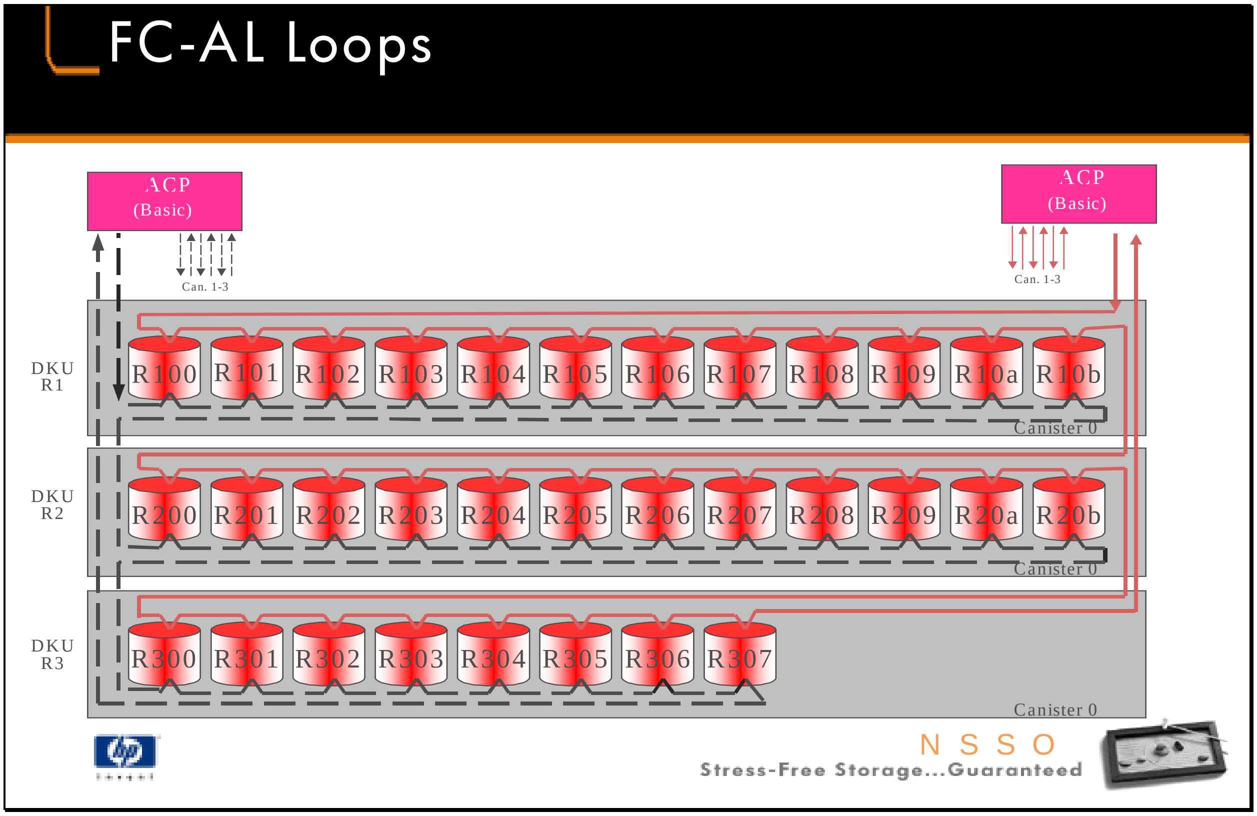
Just like the XP256, the XP512's DKC holds up to 4 pair of ACP boards in the rear logic gate. The obvious difference in the ACP architecture between the XP256 and the XP512 is the dedicated i960 processor tied to each backend connection of the XP512's ACP board (one to one relationship). The second obvious difference is the number of paths into the DKC (paths to shared memory and cache) has more than doubled with the XP512 (the new paths are also dedicated connects to either shared memory or cache). In the XP256, only two paths with redundancy go to the data and control buses.



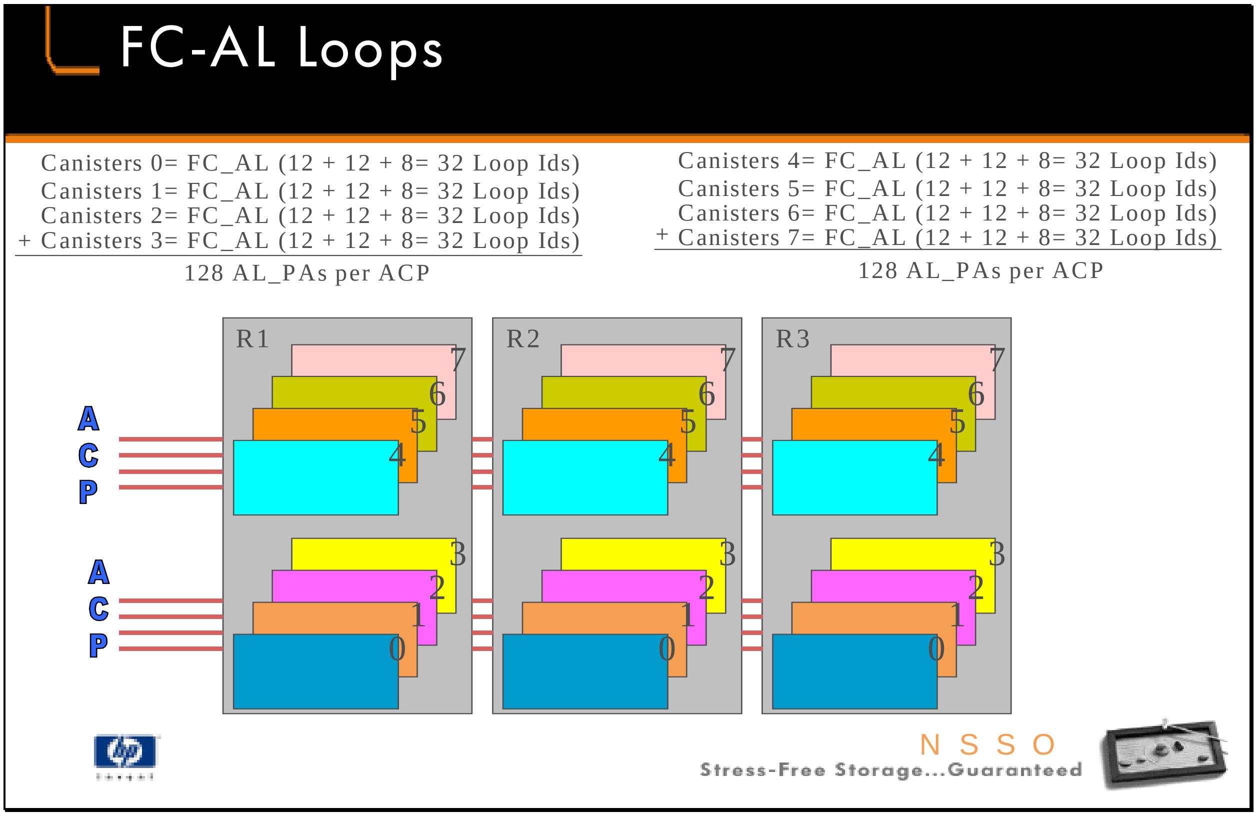
Domains: each ACP pair's connection to HDDs within its boundary is known as a domain. Domains can contain a maximum of twelve canisters (4 canisters per DKU, a maximum of 3 DKUs per domain. Each group of four canisters is known as a Block of Four or a "B4" … explained in more detail later).

CONCEPT
Each ACP pair serves the HDDs in the upper or lower portion of the DKU in only the right or left direction. Domains can be referenced as on the right or left of the DKC, in the upper or lower portion of the DKU.

RAID Definitions: Each domain normally supports one RAID configuration (either RAID 0/1 or RAID5). For XP 512/48, it is possible to have both RAID 0/1 and RAID 5 in the same domain (not possible for XP 256).



Above we see the physical connection that creates on FC-AL spanning across canisters 0 in DKUs R1, R2 and R3. The following slide will provide more detail.

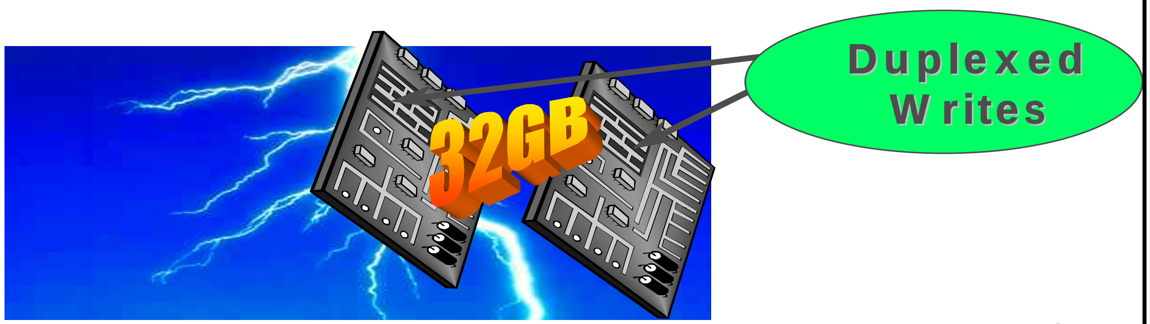


The XP512 FC-AL canister linkage is just like the XP256's SCSI bus canister linkage. The XP512 supports copper FC. Each ACP pair can control up to 128 physical disk on four FC-ALs (all in only one domain). The XP512 only allows for 32 AL_PAs or loop IDs addresses per loop. Each loop spans three canisters and up to two DKUs. These addresses are not configurable (ie- you could not make one FC loop of 32 addresses reside in only one DKU cabinet). Each ACP board allows for 128 AL_PA addresses. A single ACP pair with three DKU cabinets will have three canisters units connected on a single FC-AL pair (this is both the primary and alternate paths).
(one ACP board * 4 loops per board)*(32 AL_Pas per loop)=128 addresses
FC-AL supports up to 126 IDs per loop. However, the XP512 has a 32 IDs per loop limitation (reason is Hitachi proprietary information).

HP XP512 Write Duplexed Cache

Dynamic duplexed write cache used for data transfer

- All writes duplexed
- Separate power boundaries
- Nonvolatile Cache protected up to 48 hours with batteries



NSSO
Stress-Free Storage...Guaranteed

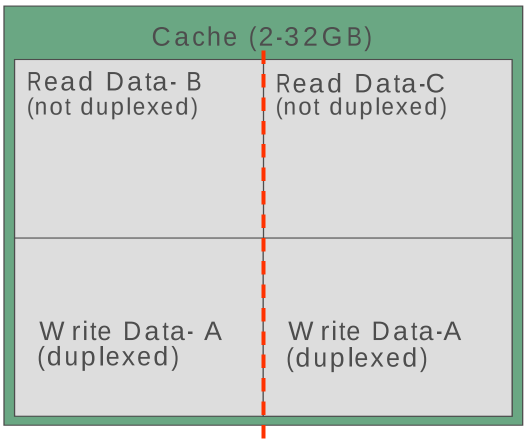
The HP XP512 subsystem can be configured with up to 32 GB of cache (minimum = 2GB). The cache modules are attached to the C-HSN (Cache Memory Hierarchical Star Net).

Without the effective use of cache, your read and write operations are limited to the speed of the disk mechanisms. With inadequate cache, the array would use time trying to de-stage writes and processing cache read misses. Primary message: when facing performance degradation, identify the amount of cache and determine if the cache hit rate adequate to meet the array's I/O demands. Most implementation recommend a minimum of 2GB cache.

What is the competitive advantage of the XP cache architecture over EMC? EMC has a single cache. The EMC array does have hooks in place to reconstruct the data lost in the cache, but the reality is they have a critical single point of failure in the array.

If a unit is expanded from 2 cache boards to 4 cache boards (the max) additional batteries must be installed (comes as a package when order cache boards).

HP XP512 *Cache*

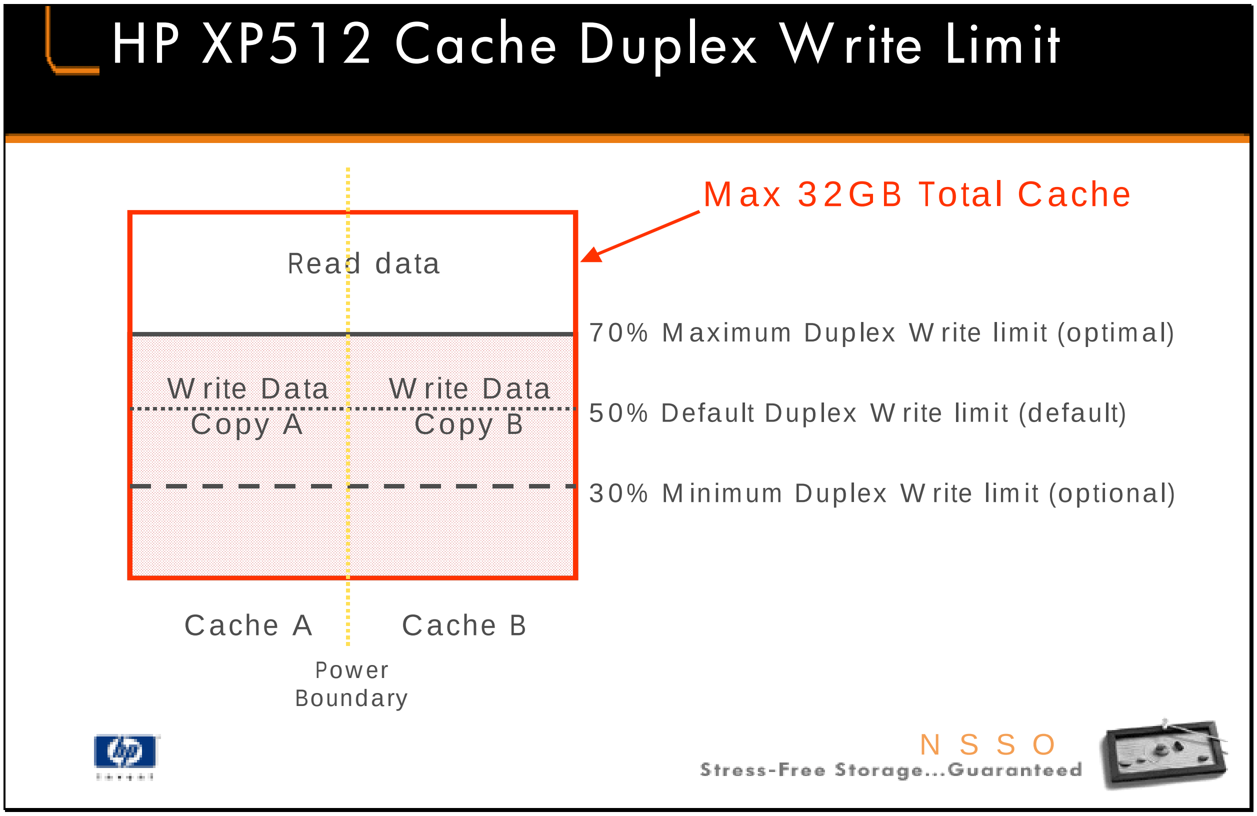


- Total Cache is divided equally on two separate power boundaries
- All write data is written to both cache segments. If one copy in cache is defective or lost the other copy is immediately de-staged to disk and cache is disabled.

N S S O
Stress-Free Storage...Guaranteed

All read and write data passes through cache. All write data is written to two WRITE cache segments with one CHIP write operation (thus the data is duplexed across power boundaries after it reaches the CACHE). Because READ data is already on disk there is no need for the XP512 to duplex this information in CACHE. The CHIP never reads or writes data directly to the disk, instead it only reads/writes data from cache.

If one or both cache segments fail, cache is no longer used. If power fails, batteries can maintain cache for at least 48 hours. Batteries should be replaced at regular scheduled intervals to maintain this 48 hour uptime.

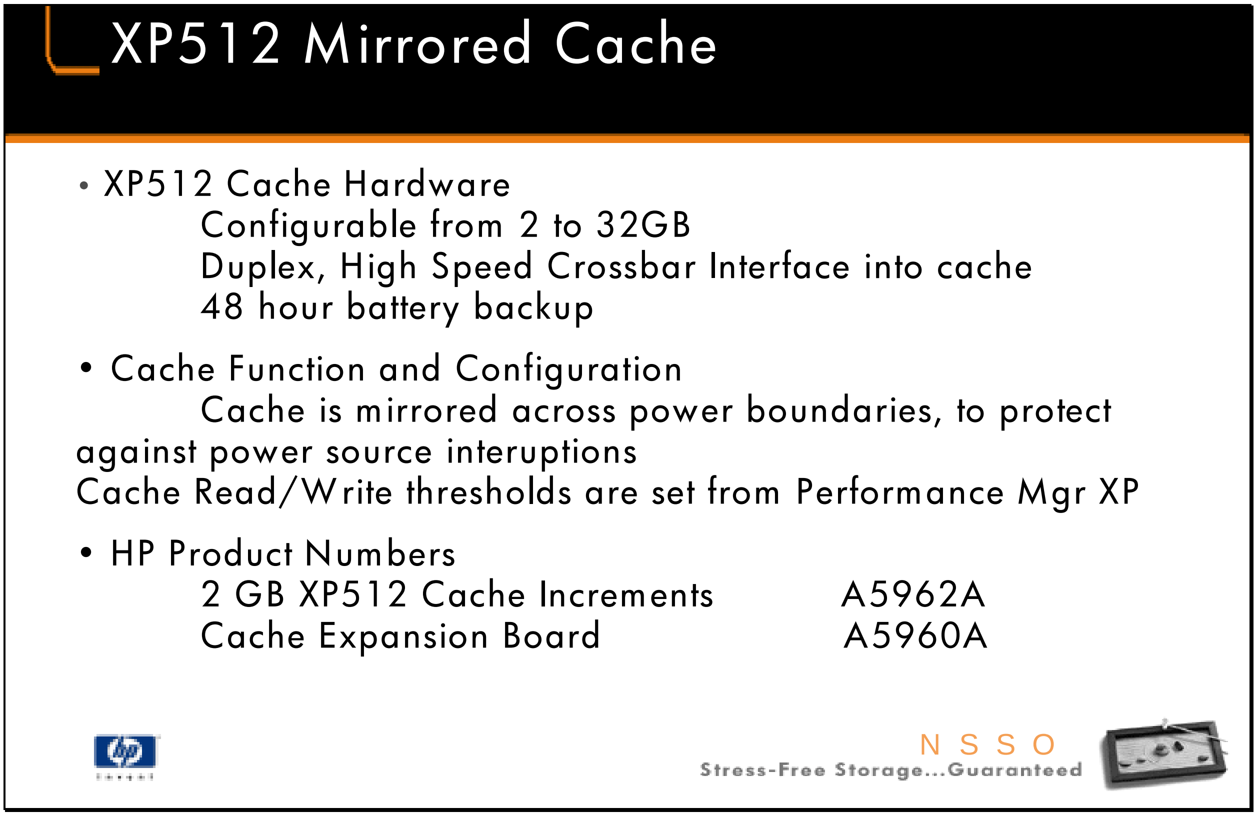


This slide illustrates the dynamic duplex cache and the duplex write line (DWL). The dynamic duplex cache is the area of cache that is dynamically allocated for write operations. The DWL is the amount of dynamic duplex cache expressed as a percentage of total cache. The amount of fast-write data stored in cache is dynamically managed by the cache control algorithms to provide the optimum amount of read and write cache based on workload I/O characteristics.

The default DWL setting allows up to 50% of cache to be allocated to fast-write data. Performance Manager XP can be used to adjust the DWL from 30% to 70% in real-time or according to a user-defined weekly schedule.

Note: If the DWL limit is ever reached, the HP XP512 sends fast-write delay or retry indications to the host until the appropriate amount of data can be de-staged from cache to the disks to make more cache slots available.

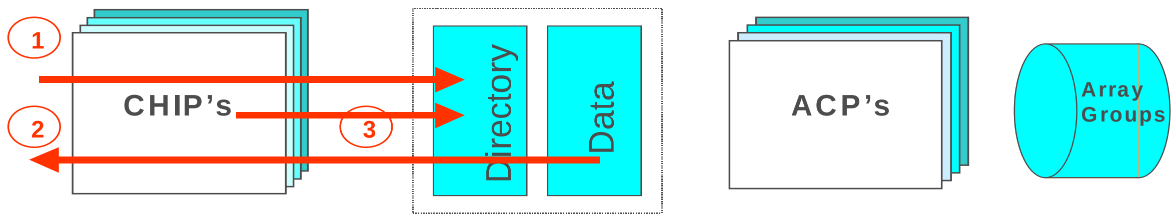
For a 32GB cache used with the default DWL, reads can use all available cache (not being used by writes) while writes would be able to use up to 8GB of the same cache (16/2= 8GB write only redundancy).



Four basic cache operations exist on the XP512:
· Read Hit
Read Miss
Fast Write
Deferred Write
A Read-hit occurs when the data requested from the host exists in cache. The CHIP initiates a search on the cache directory in shared memory. A read hit is acknowledged and the requested data is immediately transferred to the host. The directory cache is updated to reflect the most recently used data. A Read-miss occurs when the data requested from the host does not exist in cache. Like the read hit scenario, the first step is to initiate a search on the cache directory in shared memory. A read miss is signaled and the requested data is transferred from the disk to the read cache. The requested data is then transferred to the host. Also like the read hit, the cache directory in shared memory is updated to reflect the most recently used data. Unlike the write data, there is only one copy of the read data kept in cache. An algorithm that uses parameters of what physical disk the data was being read from determines what side of the duplex cache the read data is stored. A fast write occurs when the cache is not full and does not need to be destaged to the disk before the write can occur. The CHIP initiates a search on the cache directory in shared memory to identify if an old copy of the data to be written is still in cache and if the cache

space is available. The data is transferred from the host to the cache and duplexed to both Cache A and Cache B. The cache directory in shared memory is modified to reflect the most recently used data. The host is notified of an I/O completion. The data in cache is destaged to the disk in the background. The reason for writing the data to both cache areas is: the data could potentially be lost if a cache error occurs before the data has been written to the physical disk, and there was only a single copy of the data. After the data has been successfully de-staged to the disk, the cache data is switched into the read area and only one copy is kept in the cache. A deferred write occurs if the duplex write cache is at the write limits and cannot accept the new data before de-staging a cache block to the disk. In this situation the CHIP initiates a search on the cache directory in shared memory and identifies that the cache is full. The least recently used data is identified and de-staged to disk. After the least recently used data is de-staged, the data is transferred from the host to the cache and duplexed to both Cache A and Cache B. The cache directory is updated to reflect the most recently used data, and the host is notified of the I/O completion. The data in the cache is de-staged to the disk in the background. The XP512 adds data integrity codes to the host data at various points in the path from the host I/F to the physical disk. These data integrity codes are appended by hardware to allow maximum data transfer rates through the subsystem interfaces. The parity bits and integrity codes are analyzed, if the analysis identifies an error, then a cache error has occurred. The failing data in cache is identified as bad and the cache page becomes non-operative. After the failing cache page becomes non-operative the XP512's Continuous Track XP "Phone Home" capability is used to alert an HP field service representative of the failure, and a notification is sent to the Command View XP remote console.

HP XP512 Cache Management Read Hit

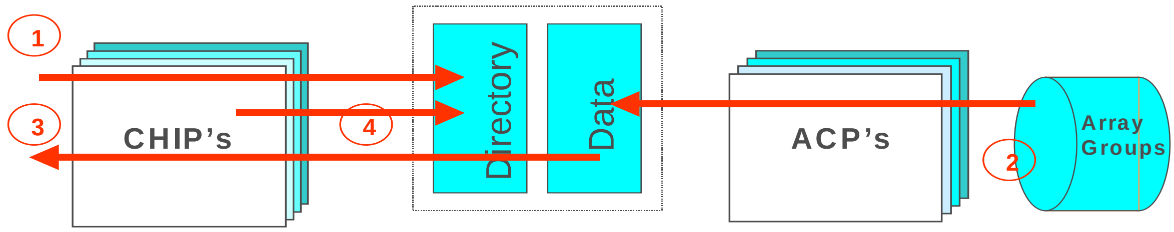


1: CHIP initiates a search on the cache directory
2: Requested data is immediately transferred to the host
3: Directory in cache is modified to reflect the most recently used data

N S S O
Stress-Free Storage...Guaranteed

Read hit. For a read I/O, when the requested data is already in cache, the operation is classified as a read hit. The CHIP searches the cache directory, determines that the data is in cache, and immediately transfers the data to the host at the channel transfer rate.

HP XP512 Cache Management Read Miss



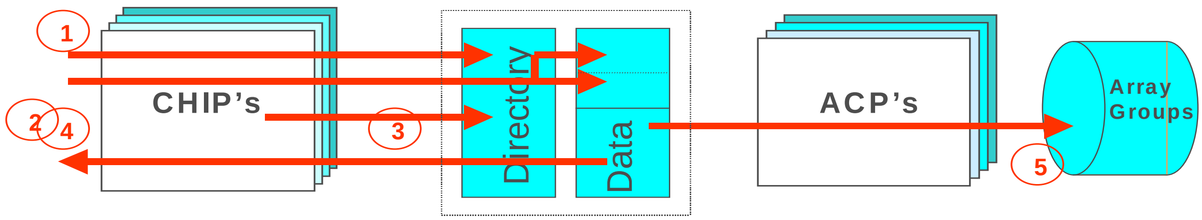
1: CHIP initiates a search on the cache directory
2: Data not in cache is transferred from the disk
3: Requested data is transferred to the hosts
4: Directory in cache is modified to reflect the most recently used data

NSSO
Stress-Free Storage...Guaranteed

Read miss. For a read I/O, when the requested data is not currently in cache, the operation is classified as a read miss. The CHIP searches the cache directory, determines that the data is not in cache, disconnects from the host, creates space in cache, updates the cache directory, and requests the data from the appropriate ACP pair. The ACP pair stages the appropriate amount of data into cache, depending on the type of read I/O (e.g., sequential).

Remember: too little cache might cause an excessive amount of cache read misses due to the requested data not residing in cache.

HP XP512 Cache Management Fast Write

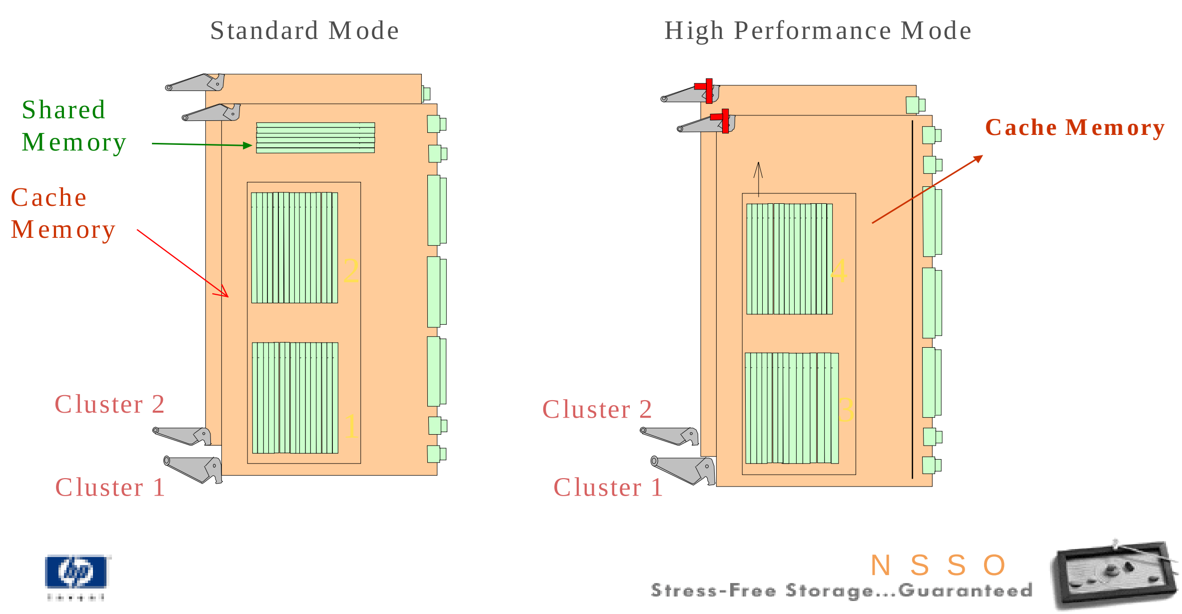


1: CHIP initiates a search on the cache directory
2: Data is transferred from the host to the cache and duplexed
3: Directory in cache is modified to reflect the most recently used data
4: Host is notified of I/O completion
5: Data is destaged to the disk in background

NSSO
Stress-Free Storage...Guaranteed

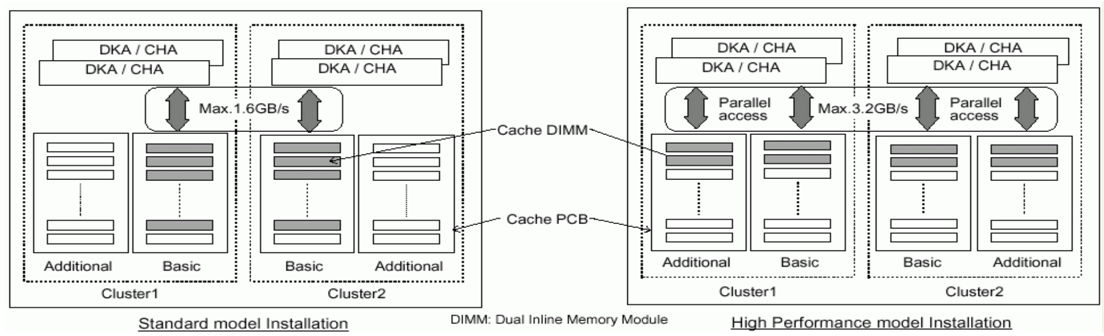
Fast write. All write I/Os to the XP512 subsystem are fast writes, because all write data is written to cache before being destaged to disk. The data is stored in two cache locations on separate power boundaries in the dynamic duplex cache. As soon as the data has been written to cache, the XP512 notifies the host that the I/O operation is complete, and then destages the data to disk in background.

Cache Board Diagrams



Shared Memory is only on the basic Cache boards; the additional High Performance cache boards (explained next slide) do not have shared memory expansion bays.

XP 512 Cache



Component Comparison		XP512	XP256
Cache memory	Capacity	512MB to 32GB	max.16GB
	Additional Capacity	512MB (x2)	512MB/1GB
Shared memory	Capacity	512MB to 1.5GB	max.512MB
	Additional Capacity	256MB (x2)	128MB

N S S O
Stress-Free Storage...Guaranteed

In the XP256, the first pair of cache boards in the rear logic gate had to be fully populated to 8GB before the second pair of cache boards in the front logic gate could be populated with more memory DIMMs. In the XP512 it is possible to install both pair of cache boards and populate each of the four boards in parallel (known as cache high performance mode). Cache high performance allows for faster parallel cache access across all four boards.

Differences in cache operations in regards to read hits, read misses, fast writes and deferred writes do not exist between the XP256 and XP512.

Standard Installation Mode: The basic board houses both standard installation cache and all shared memory.   Basic cache board only with support up to 16GB cache and 1.28GB shared memory. Total of two boards placed in slot G in the rear logic gate and slot T in the front logic gate.

High Performance Mode: Basic cache board (basic) and additional cache board (add 1 also know as high performance boards) are populated with cache memory in parallel. There is no shared memory on the High Performance boards. Basic boards remain in slots G and T and slots F and U for high performance boards. High performance mode allows for more paths to the CHIPs (CHA) and ACPs (DKA) for two times the number of parallel access paths. High Performance Mode allows for 3.2GB bandwidth.

Cache Requirements based on Usable Data

Cache Requirements in GB		
Data Storage Capacity (up to) (GB)	Mandatory Cache GB	Recommended Cache GB
up to 500	2	4
501 to 800	4	8
801 to 1300	6	12
1301 to 7400	8	16
above 7400	10	20

N S S O
Stress-Free Storage...Guaranteed

The transition points for cache requirements relative to data storage capacity for the XP512 has changed. To simplify customer presentations, data storage transition points have been rounded off to simpler increments.

OLD Table:

Data Storage Capacity	*Mandatory Cache*	*(GB)*	*(GB)*
up to 529	2		
797	4		
1122	6		
7229	8		
above 7229	10		

Shared memory Contains

Table of what is stored in the cache, by 16KB block	***Grows with size of cache***
Table of information on each LDEV	***Grows with # of LDEVs***

Information on each CHIP and ACP in the array, and other service information: relatively small, and is included in the values shown in the table.

N S S O
Stress-Free Storage...Guaranteed

XP512 Shared Memory

- XP512 Shared Memory Function
  - Stores system configuration tables
  - 16 Control Unit (CU) data tables
- Used for system configuration mapping of system components, LUN mapping, cache pointers, cache hit rates, RAID levels
- XP512 Shared Memory Hardware
  - 4 "lower bandwidth point-to-point connections for shared memory interconnect
  - Expandable from 512 MB to 1.28 GB in 256MB increments
- HP Product Number
  - 256 MB Shared Memory Module A5963A

NSSO
Stress-Free Storage...Guaranteed

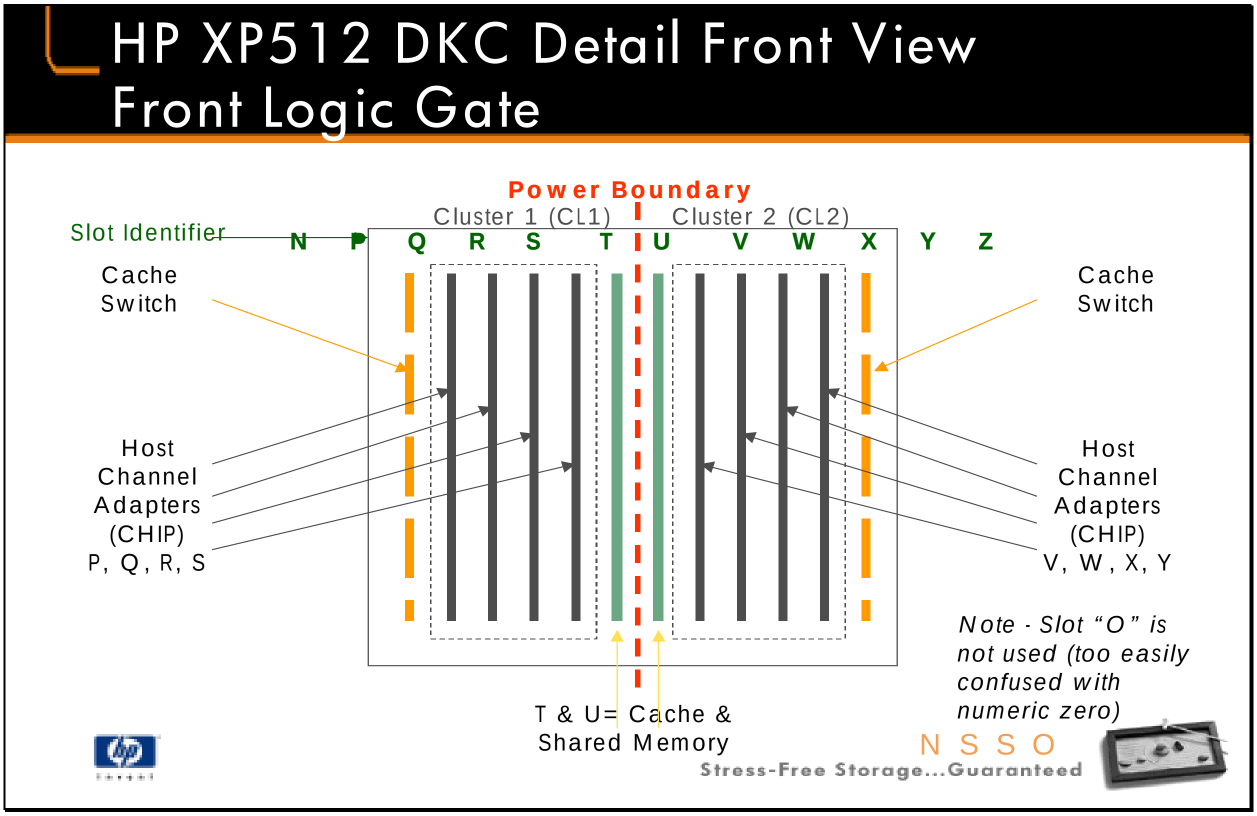
XP512 Shared Memory….. how much do I need?

1. **Determine cache size (how many Gigabytes)**
2. **Determine number of LDEVs**
3. **Look up the required shared memory size in the following table**

Shared Memory Requirements

Cache Memory (GB)	LDEV		
	0 to 1024	1025 to 2048	2049 to 4096
	SM (MB)	SM (MB)	SM (MB)
2	512	768	1024
4	512	768	1024
6	512	1024	1024
8	768	1024	1024
10	768	1024	1024
12	768	1024	1024
14	768	1024	1024
16	768	1024	1024
18	768	1024	1280
20	768	1024	1280
22	768	1280	1280
24	1024	1280	1280
26	1024	1280	1280
28	1024	1280	1280
30	1024	1280	1280
32	1024	1280	1280

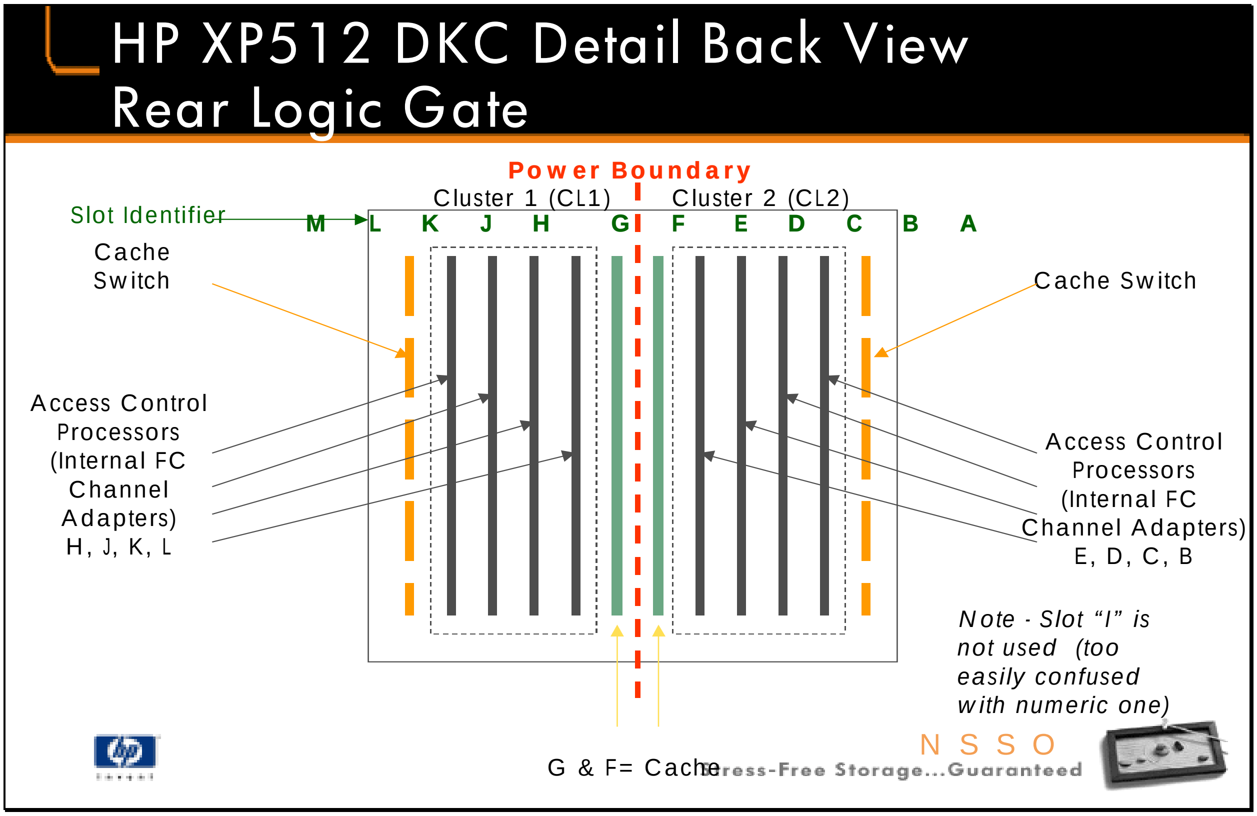
N S S O
Stress-Free Storage...Guaranteed



Here is a diagram of the DKC's front logic gate. The illustration serves a logical view of the available physical expansion slots in the front of the DKC (you must remove the large protective metal covering to view this within the DKC). The DKC has two separate locations for expansion board inserts- one in the front and one in the back (covered next). The DKC's front expansion slots P, Q, R, S, V, W, X & Y are for CHIP pairs only (FC and ESCON).

Concepts:

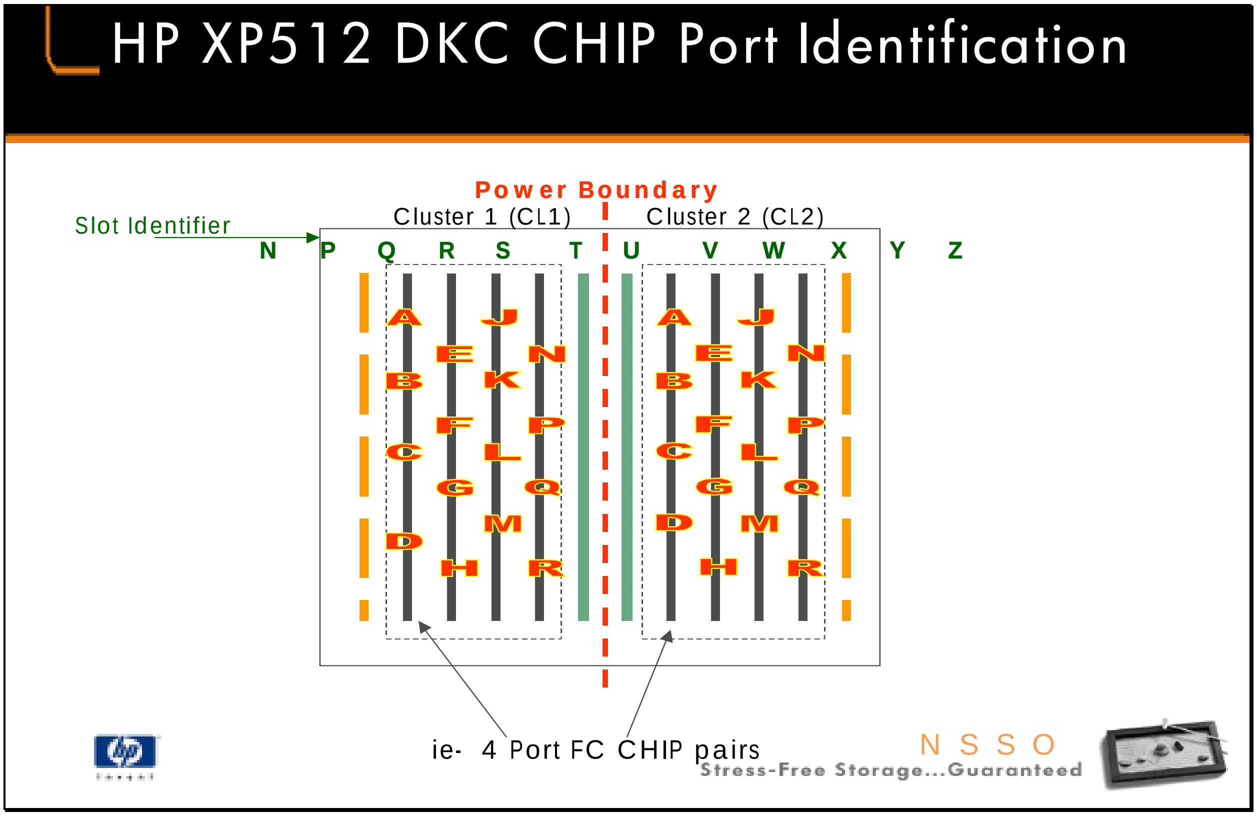
1. Cluster1 and cluster2 represent logical separation of power boundaries. If a host channel adapter is installed in slot P, a redundant host channel adapter must be installed in the slot V. Redundant pairing always occurs by establishing the following pairs: N/Z, P/V, Q/W, R/X, S/Y and T/U (cache & shared memory).
2. All pairs are physically separated across two different power sources.
3. The maximum number of host channel adapter pair expansions possible is four.
4. Cluster 1 and 2 identifiers represented in this diagram are also labeled on the DKC's front panel (toggle switches with the green LEDs). However, these clusters define the CHIP ports (A-R excluding I and O) and not slot identifiers (A-Z excluding I and O). More on this later....
5. Maximum cache in the front logic gate= 8GB.



Here is a diagram of the DKC's rear logic gate. The illustration serves a logical view of the available physical expansion slots in the back of the DKC (you must remove the large protective metal covering to view this within the DKC). Expansion slots B, C, D, E, H, J, K & L are for ACP pairs only.

Concepts:

1. Like the front, redundant pairing always occurs by establishing pairs between the rear logic gates cluster's 1 and 2: M/A, L/E, K/D, J/C, H/B, G/F (cache only).

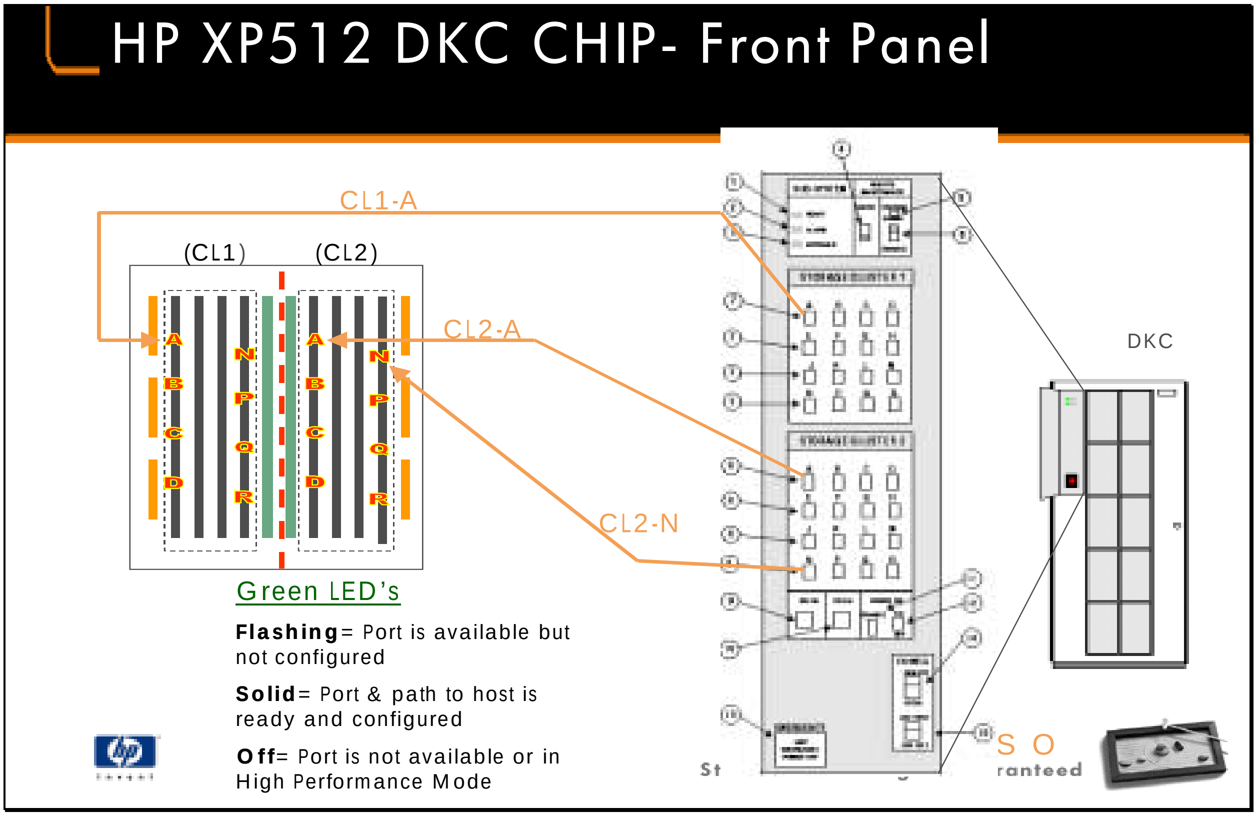


Ports on CHIP boards are lettered from the top of the board to the bottom; ports can them be identified by their respective cluster.

Example: A four port FC board installed in slot L would have, starting from the top, ports CL1-A, CL1-B, CL1-C and CL1-D. Because CHIPs are always installed in pairs, slot identifier E in cluster 2 would also have a four port FC card as well with ports CL2-A, CL2-B, CL2-C and CL2-D. You will need to understand this in order to properly assign LUNs.

Assume you had a two port ESCON card in slots Q and W. How would the XP512 identify those two ports?
Answer: The Xp512 would see the ports as CL1-E/CL1-F and CL2-E/CL2-F. In this case all ports available for use in slots K & D would never be defined (CL1-G, CL1-H, CL2-G and CL2-H would only be defined if you put in a four port ESCON or FC card in slots K & D as opposed to a two port ESCON card). Continuing this example further, slots J would start with defined ports CL1-J, CL1-K etc even though ports CL1-G and CL1-H were never "used" for port to volume mapping.

NOTE: Just like the Slot Identifiers, ports do not use alpha letters I and O in order to prevent confusion between numeric one and zero. Understanding slot lettering is necessary when installing/replacing CHIPs or stringing cable from host(s) to the array (i.e. CE responsibilities).



A previous slide mentioned the DKC's front panel LED's in relation to CHIP ports. The front panel gives a physical representation of the CHIP port identifiers (A-R excluding I and O). Use this panel- it helps orient yourself to understand what is installed without having to open the unit.

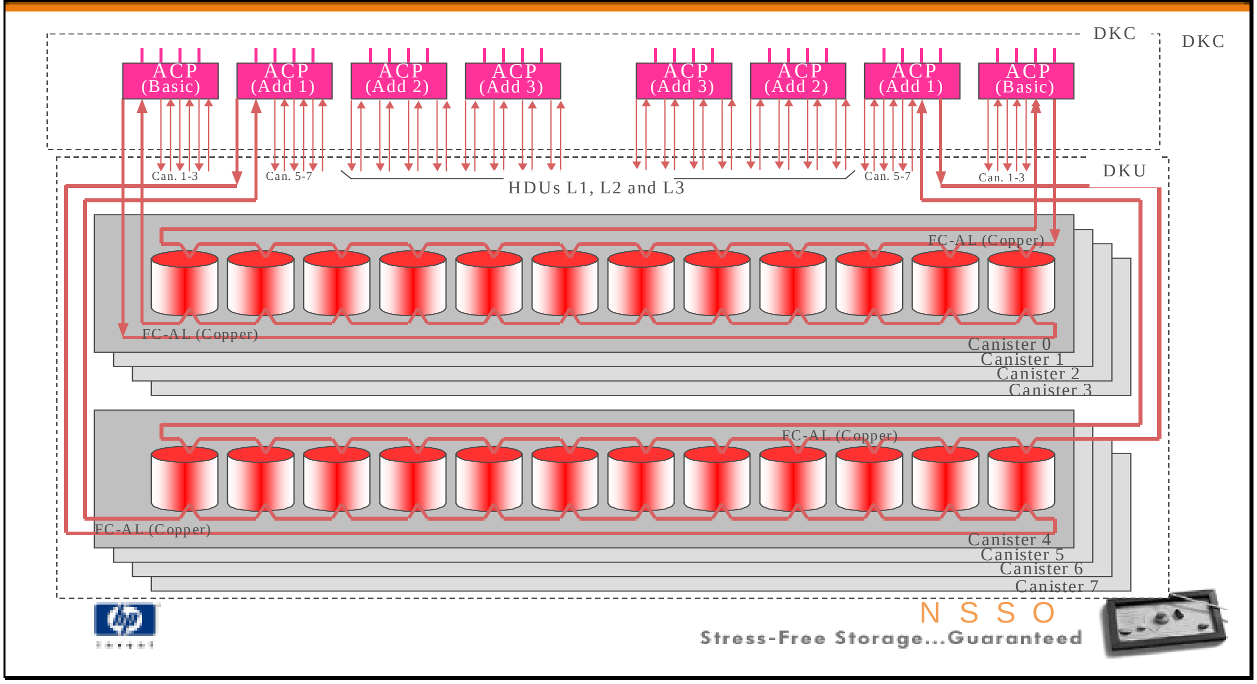
XP512 DKU Architecture

Contents

- DKU Internal Architecture Overview
- Raid Definitions
- Disk Emulations
- Parity Group & Volume Concept
- Canister & Disk Identification

NSSO
Stress-Free Storage...Guaranteed

DKU Architecture



Each DKC supports up to six DKUs. Each DKU supports up to eight canisters that have 12 HDD slots. All disk are connected to the DKC via copper Fibre Channel. Each ACP board supports up to four different FC loops supporting up to 32 AL_Pas per Fibre Channel arbitrated loop (this is a limitation of the array). Each HDD in the XP512 supports pure fibre connections. For this reason, even though the XP256 supports 47GB HDDs, same 47GB HDD that works in the XP256 would not work in the XP512.

XP512 Facts:
Number of DKUs Supported per DKC= 6
Number of canisters per DKU= 8
Number of HDD slots per canister= 12
Number of HDD slots in a six DKU array= 576
Number of functional HDD slots in a six DKU array=512
Number of HDD slots that support spares= 16
Number of HDD slots dedicated to spare HDD only= 8
Number of HDD slots that can serve as either spare or data= 8

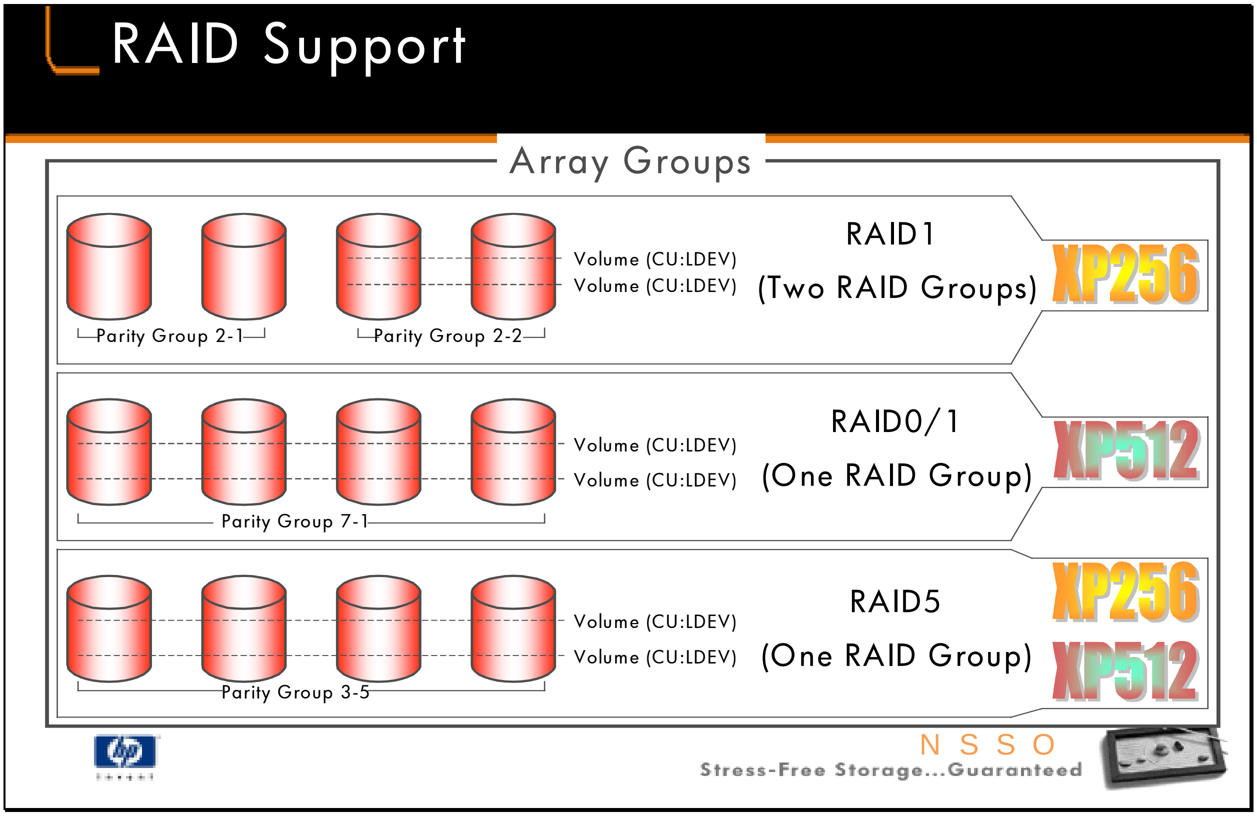
*reference slides to follow for more information of spare HDD

XP512 Supported HDD

XP512 Fibre Channel Disk Drives				
Formatted Capacity		**18.46GB**	**47.19GB**	**72.91GB**
Diameter of Disk		3 inch	3 inch	3 inch
Number of physical cyl/HDA		12027	12027	12027
Number of physical track/physical cyl		9	23	23
Number of physical disk		5	12	12
Number of physical track/HDA		60135	144324	192108
Revolution speed		10025	10025	10025
Seek time (ms)	Min	0.5/0.7	0.5/0.7	0.5/0.7
(READ/WRITE)	Ave	5.2/6.0	5.7/6.5	5.7/6.5
	Max	12.0/13.0	12.0/13.0	12.0/13.0
Average latency time (ms)		3.0	3.0	3.0
Internal data transfer rate (MB/s)		30.2 to 45.6	30.2 to 45.6	30.2 to 45.6
Interface data transfer rate (MB/s)		Max100	Max100	Max100
Number of HDD/Disk Unit		Max96	Max96	Max96
RAID 0/1 LDEVs per Parity Group	Open-K 1.881GB	19	50	77
(2D+2D)	Open-3 2.461GB	14	38	59
	Open-8 7.347GB	5	12	19
	Open-9 7.384GB	4	12	19
	Open-E 14.582GB	N/A	N/A	10
	Open-L 36.450GB	N/A	N/A	4
	Open-M 47.185GB	N/A	2	N/A
RAID 5 LDEVs per Parity Group	Open-K 1.881GB	29	75	116
(3D+1P)	Open-3 2.461GB	22	57	88
	Open-8 7.347GB	7	19	29
	Open-9 7.384GB	7	19	29
	Open-E 14.582GB	N/A	N/A	15
	Open-L 36.450GB	N/A	N/A	6
	Open-M 47.185GB	N/A	3	N/A

NOTE! FC 72GB drives and OPEN-E are supported as of September 1st, 2000. HP also ships a SCSI 72GB drive for the XP256 as well since November 2000. Although the sizes of these disk drives may be similar to disk drive sizes supported by the XP256, the XP512 disk are Fibre connected, not SCSI connected. The XP512 basic disk components are an array group. An array group depends on the RAID level that the user plans on implementing. The objectives of the RAID technology are low cost, high reliability, and high I/O performance of disk storage devices. To achieve these objectives the XP512 disk array supports two different modes of RAID control. These are RAID 0/1, and RAID 5. All RAID groups in the XP512 require four physical disks. RAID 0/1 is also called 'Dual Read RAID 1'. The user can choose between two RAID levels with the XP512. Which RAID level the user would want depends on the type of data storage solution they require. In most situations, RAID 0/1 will have better performance over RAID 5, but RAID 5 will provide more useable capacity. If the host application is write intensive there should not be much if any performance differences between either of the two RAID levels. If the host application is more read intensive, this is where the user will see some performance differences between the two RAID levels. The downside to RAID 0/1 over RAID 5 is that there is a 50% overhead on storage redundancy associated with RAID 0/1. RAID 0/1 is often referred to as mirrored storage because the host data is physically duplicated in the array. The user will get more overall useable storage out of the array with a RAID 5 implementation.

RAID 5 achieves the storage redundancy by generating parity data based on the data stored. RAID 5 only requires a 25% storage overhead. In May 2001 HP introduced a high-speed 15K 47GB disk for the XP512 array.



Array Group= always a collection of four disk (not to be confused with a Raid Group or Parity Group)
Parity Group = **Raid Group** (the terms are synonymous according to HP development)
XP512 the number of parity groups does not double if using RAID0/1 as opposed to RAID5 (in the XP256 the number of parity groups doubled if using all RAID1 as opposed to all RAID5)

XP512 only supports RAID0/1 and RAID5 (*Note: much of the XP512 specific documentation will state RAID1 as opposed to RAID0/1*)
XP256 only supports RAID1 and RAID5

Array Group= always a collection of four disk (not to be confused with a Raid Group or Parity Group). Parity Group = Raid Group (the terms are synonymous according to HP development). XP512 the number of parity groups does not double if using RAID0/1 as opposed to RAID5 (in the XP256 the number of parity groups doubled if using all RAID1 as opposed to all RAID5)

XP512 only supports RAID0/1 and RAID5 (Note: much of the XP512 specific documentation will state RAID1 as opposed to RAID0/1)
XP256 only supports RAID1 and RAID5
RAID 0/1 (2D+2D)

RAID 0/1 is new and only supported on the XP512. The XP256 supports RAID 1 and RAID 5.

RAID 0/1 uses both RAID 0 and RAID 1 technologies implemented together to achieve high reliability and high I/O performance.

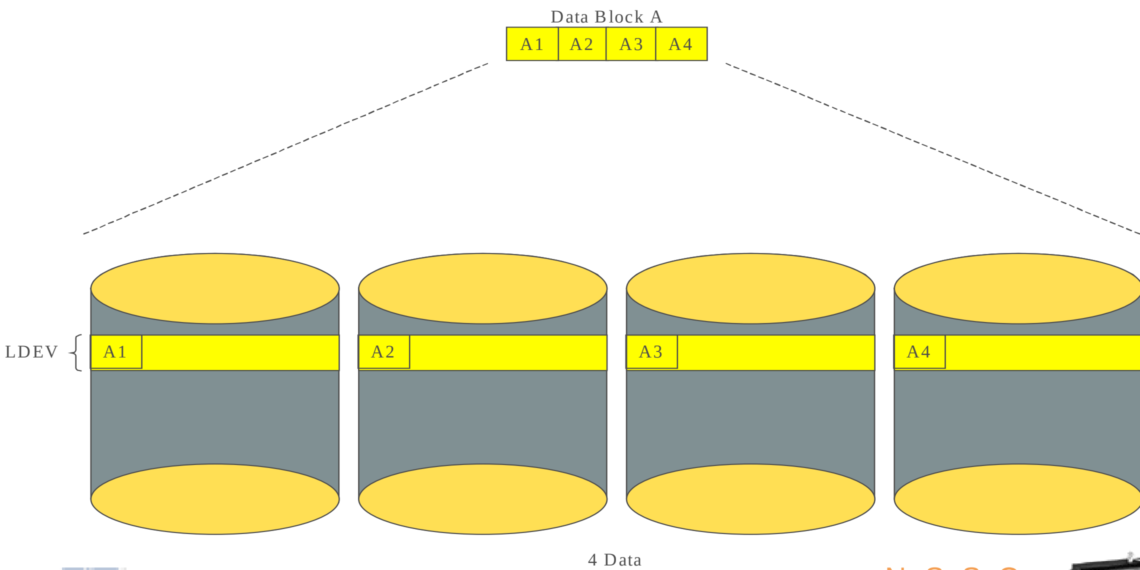
RAID 0 produces a striped drive volume. Striped data means that the stream of data from the host is split and distributed onto two or more disk devices on a byte or bit bases. RAID 0 produces a

very high performance I/O disk subsystem where fault tolerance is not required.

RAID 1 uses a disk-mirroring algorithm that requires at least two disk drives. RAID 1 produces a mirrored drive volume. RAID 1 is the simplest of RAID technologies because of the nature of mirrored storage. The RAID controller writes the data to both duplex areas of cache. After the first write to the primary disk is complete the secondary copy of data in cache is written out to the secondary disk. After the writes to both disks have completed the write cache is released. The RAID 1 is fault tolerant, but it has lower performance than RAID 0.

As mentioned earlier RAID 0/1 adds high I/O performance features of RAID 0 striping to the high reliability features of mirrored storage with RAID1. RAID 0/1 uses all four disks in an array group for the primary storage path. The I/O performance improvements are achieved from the fact that the I/O data streams are split onto all four disks in the array group. Striping the data across the four disks more than doubles the performance over the XP256 with RAID 1.

RAID-0



NSSO
Stress-Free Storage...Guaranteed

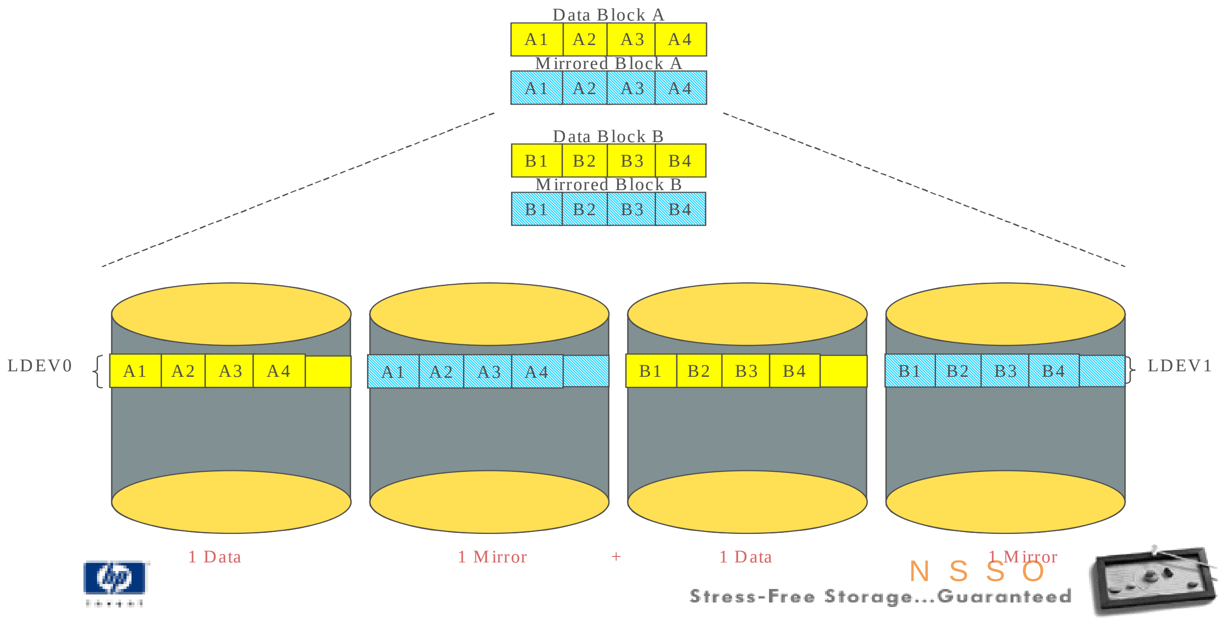
RAID0 is not supported by the XP512, but will be explained here to help in the understanding of RAID0/1.

RAID0= data from the logical volume is spread evenly across all available disk by dividing the data into blocks. The first block goes onto the first disk, the second block goes onto the second disk etc. Performance benefit is attained by having four disk work in parallel reducing queue avoidance.

BENEFITS:
1. High performance by reducing individual queue workload
2. Good random access I/O performance by reducing disk utilization and disk queues.
3. Total disk space available equals total usable disk space.

NEGATIVES:
1. No redundancy of data. Data integrity is dependent on all four disk.

RAID-1

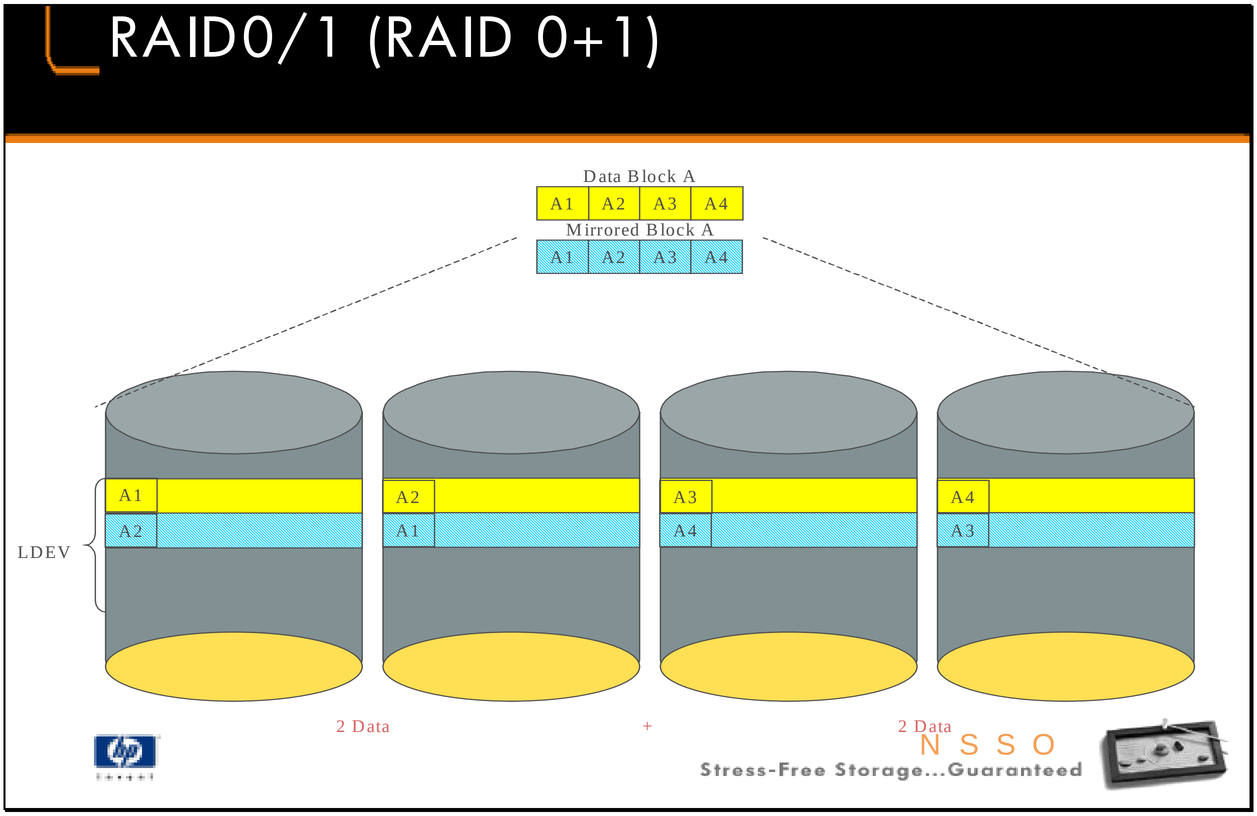


BENEFITS:

1. Mutually exclusive redundant data on mirror disk
2. Allows for two times the number of parity groups in comparison to the RAID5 or RAID0/1.

NEGATIVES:

1. 100% overhead on storage redundancy



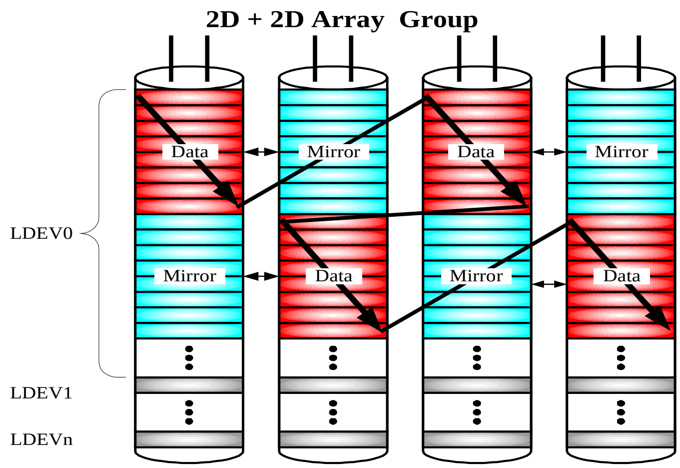
RAID0/1 creates a disk subsystem that yields increased performance by utilizing four disk to operate in parallel for writes and reads; redundant or mirrored data is dispersed evenly among the striped data across all four disk. When a host request data, four disk and multiple paths are activated during the data request. The request are dispatched serially across all four disk but are serviced concurrently. XP512's use of RAID0/1 doubles the performance from the XP256 with RAID1.

BENEFITS:
1. High performance by reducing individual queue workload. In most situations RAID0/1 will have better performance than RAID5 (benefit of RAID0 striping).
2. Good random access I/O performance by reducing disk utilization and disk queues.
3. Data integrity via mirroring.

NEGATIVES:
1. Less overall storage than RAID5.
2. In respects to the XP512, it reduces the overall number of available parity groups and LDEVs by one-half in comparison to RAID1 (assuming the same emulation type).

RAID 0/1

- RAID 0/1
  - Striped and mirrored (aka. "dual read" RAID 1
  - 50% storage overhead
  - Best for performance sensitive applications
  - Better read & write performance



RAID 0/1 is a striped and mirrored copy within a ACP pair

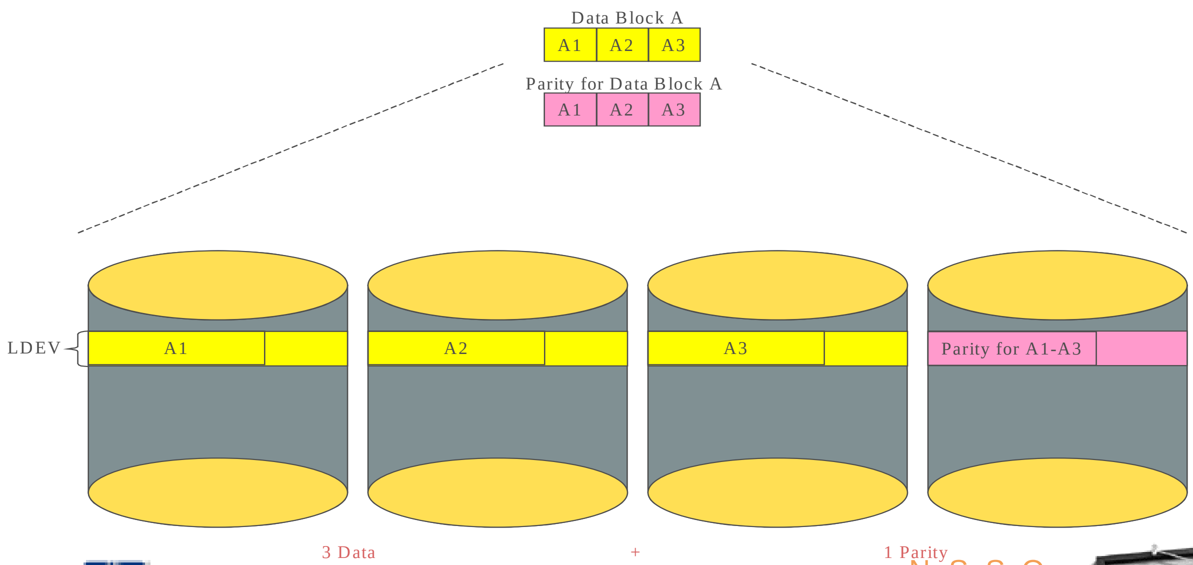
N S S O
Stress-Free Storage...Guaranteed

RAID0/1 creates a disk subsystem that yields increased performance by utilizing four disk to operate in parallel for writes and reads; redundant or mirrored data is dispersed evenly among the striped data across all four disk. When a host request data, four disk and multiple paths are activated during the data request. The request are dispatched serially across all four disk but are serviced concurrently. XP512's use of RAID0/1 doubles the performance from the XP256 with RAID1.

BENEFITS:
1. High performance by reducing individual queue workload. In most situations RAID0/1 will have better performance than RAID5 (benefit of RAID0 striping).
2. Good random access I/O performance by reducing disk utilization and disk queues.
3. Data integrity via mirroring.

NEGATIVES:
1. Less overall storage than RAID5.
2. In respects to the XP512, it reduces the overall number of available parity groups and LDEVs by one-half in comparison to RAID1 (assuming the same emulation type).

RAID-5



NSSO
Stress-Free Storage...Guaranteed

BENEFITS:

1. All four disk contain some data and some parity allowing all disk equal utilization of I/O request.
2. Sequential reads do not require parity calculations, so random and sequential reads yield high performance.
3. The most overall useable storage compared to RAID1 and RAID0/1.

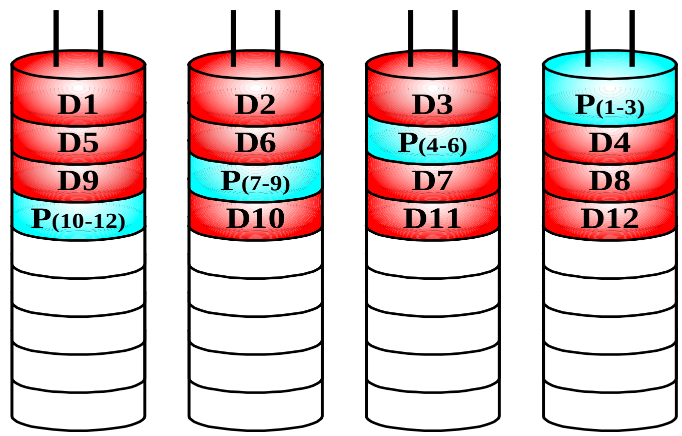
NEGATIVES:

1. Writes require read from parity for new parity calculation, calculation of new parity, then data write to disk.

RAID 5

- RAID 5
  - —Stripes parity across disks
  - —25% storage overhead
  - —More cost effective $/usable MB
  - —Excellent performance

3D + 1P Array Group *



*** Shows logical not physical view**

NSSO
Stress-Free Storage...Guaranteed

BENEFITS:
1. All four disk contain some data and some parity allowing all disk equal utilization of I/O request.
2. Sequential reads do not require parity calculations, so random and sequential reads yield high performance.
3. The most overall useable storage compared to RAID1 and RAID0/1.

NEGATIVES:
1. Writes require read from parity for new parity calculation, calculation of new parity, then data write to disk.

XP512 Mainframe Disk Emulations

- Usable Space (space available for use by a host)
  - OpenK= 1,881,538,560 (1.8GB)
  - Open3= 2,461,040,640 (2.4GB)
  - Open8= 7,347,732,480 (7.3GB)
  - Open9= 7,384,596,480 (7.4GB)
  - OpenE=14.506 GB (14.5GB)
  - OpenL= 33.94 GB
  - OpenM= 43.94 GB

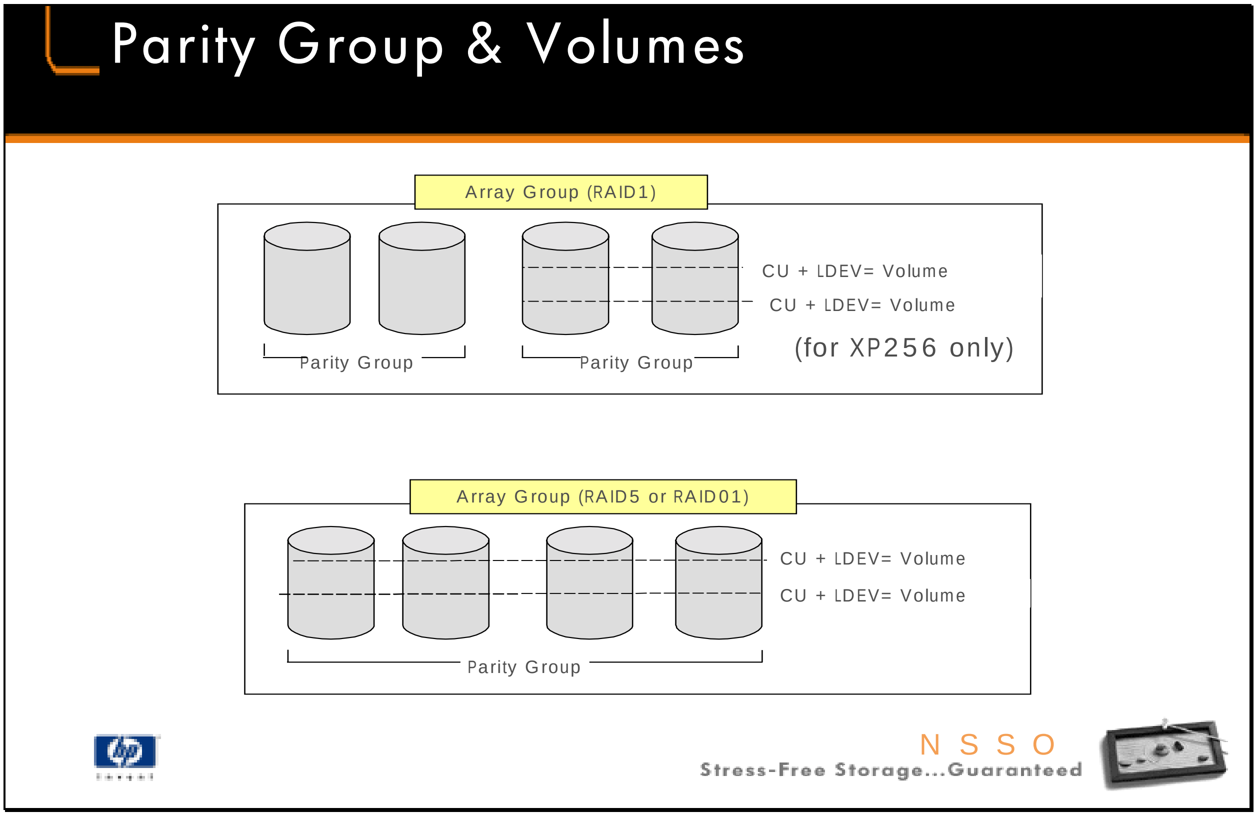
NSSO
Stress-Free Storage...Guaranteed

Disk emulation is required to map this mainframe designed storage array to be available/usable by open host. You can set all available disk emulation you wish for any type of HDDs (assuming the right microcode is installed in the DKC).

Once a set of four disk have been installed into the XP512, a definition of RAID0/1 or RAID5 must be set for the group of four disk (either 3D+1P for RAID5 or 2D+2M for RAID0/1).

Emulation Type Planning

1. Setting the emulation type to Open 3, 8, 9, E, L & M allows the array to carve up the disk into several logical devices (LDEV). The emulation type selected directly impacts the amount of storage provided to the connecting host. For example, setting up a 20 array groups with 36.9GB drives RAID5 using OPEN9 emulation allows for 14 LDEVs and 2067.7GB usable space. However, using OPEN8 emulation allows for 15 LDEVs and 2204.3GB usable space (microcode to support OPEN8 was released to optimize the 36.9GB drives)
2. The XP256 only supports a maximum of 1024 LDEVs (using 64 array groups). Using OPEN3 for high density disk will increase the total number of LDEVs.
3. Changing emulation types requires formatting



Array Group (aka Marketing Array Group): Set of four XP array hard disk drives.
Parity Group or RAID Group: A group of either two HDDs (RAID1) or four HDDs (RAID5 or RAID0/1) logically labeled in respects to their physical location in the array (based on installation order). Parity groups are a logical representation of a collection of LDEVs found within a the array. For example, parity group 6-2 might contain LDEVs 1:50-1:59.
LDEV: Logical device (LDEV) is a logical representation (hex) of a segment or slice of HDD spanning two or more disk. An example would be 0A or 20. LDEVs are created within parity groups; the number of LDEVs created within one parity group will be dictated by the defined emulation (ie- OPENK, 3, 8, or 9) and the HDD size. Thus, the number of LDEVs created in one array group will be dictated by the HDD size (determines how many LDEV of any emulation type can fit on the disks), the RAID definition (defines if our array group will be one parity group (ie RAID5 or RAID0/1) or two parity groups (ie RAID1)), and the emulation type (defines the physical size of each LDEV, the larger the LDEV the fewer number of LDEVs that will fit in the parity group). Documentation and general “speak” would be incorrect to use LDEV and Volume or LDEV and LUN as synonymous terms (LUN defined below).
Control Unit: The control unit is assigned to only one LDEV. The combination of these two values creates a volume. The control unit only acts as a logical identifier to the LDEV; the use of control units allows us to assign up to 256 LDEVs per CU (0-255). The XP256 has a total of 4 control units assigned numerically from 0 to 3 giving us a

total of 1024 volumes; the XP512 has a total of 16 control units ranging from 0 to f (hex) giving us a total of 4096 volumes.

Volume: A collective logical representation of the control unit plus the logical device number (CU:LDEV or 0:01 for example). To use the term Volume means you know or have provided the associated control unit number for the LDEV. To use the term LDEV means you do not know the control unit number for which the LDEV belongs. Remember, we can have the same LDEV number in more than one domain, so to make the terms synonymous would not be correct. This becomes confusing in the HP world of XP documentation, program interface and general speak when customers see “volume” and “LDEV” interchanged freely.

More On Control Unit

Control Units

16 Control Units (CU) per DKC

- A single CU can be combined with 256 LDEVs
- Assigning 16 different control units (0-f) to 256 LDEV ranges (00-ff) per CU allows for 4096 volumes

  16 x 256= 4096

N S S O
Stress-Free Storage...Guaranteed

The XP256 only supported four CU's per array. The XP512 supports up to 16 CUs. CUs allow for the creation of up to 256 volumes per CU. Volumes are created after assigning a CU to an LDEV (CU:LDEV); thus the range of valid CU to LDEV assignments are CU's 0x00h through 0x0fh (0-16) and LDEVs 0x 00h through0xffh (00-256). A few volume examples are illustrated below:

0:50 4:04
a:01 f:ff

With the XP256 many assumed that if they knew the control unit number of a particular device they would also know that volume's physical location. This was assumed because HP ships CU0 to span across the bottom of DKUs R1 and R2 (domain 1), CU1 to span across the top of DKU's R1 and R2 (domain 2) etc. It should be understood that CU's are only a logical identifier tied to a particular LDEV; the use of CU's does not tell us anything more than one-half of a volume name.

Can CU's span between domains? Yes. One example might be a volume migration by Auto LUN. A migration could easily result in a parity group with a non-contiguous range of volumes that contains > 1 CU identifier.

HP XP512 Disk and RAID Intermix Rules

- Cannot mix disk sizes within a four disk array group

Disks can only be installed in sets of four identical disk

- Can mix RAID types on an ACP pair

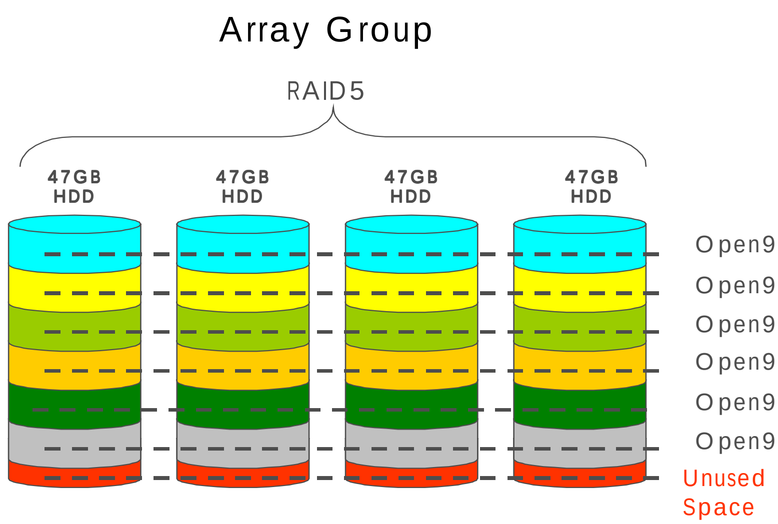
ex) RAID 0/1 and RAID5 groups can exist in the same domain within the XP512

NSSO
Stress-Free Storage...Guaranteed

1. Cannot mix disk sizes within a 4 disk array group. For example, we cannot have two 18GB and two 47GB disk defined within the same array group. All disk would have to be either 18GB or 47GB.

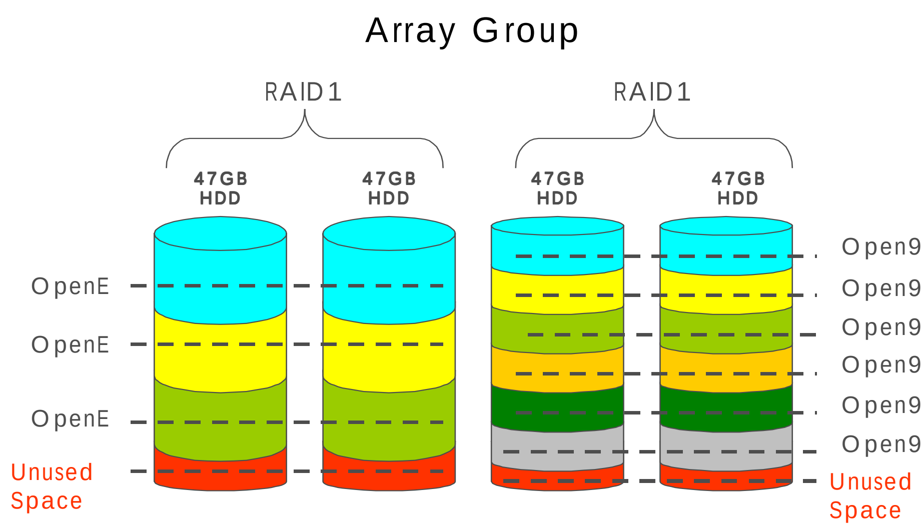
2. Can mix RAID types on an ACP pair- stated another way we can mix RAID types within a domain. This means that if we have RAID5 defined for some array groups housed in canister's 0-3 in DKU R1 and or DKU R1+R2+R3 we can define other array groups in the same domain to be RAID0/1. But you cannot do this in XP256.

Carving up an Array Group into OPEN Emulations: Illustration 1



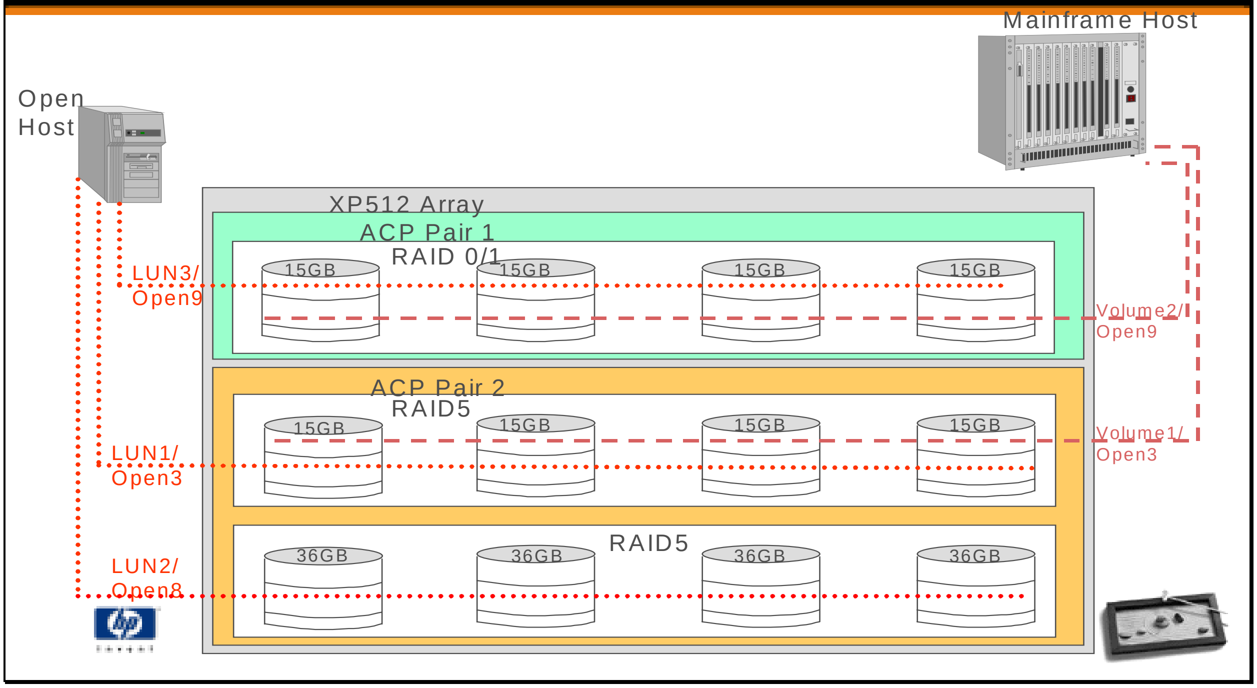
NSSO
Stress-Free Storage...Guaranteed

Carving up an Array Group into OPEN Emulations: Illustration 2



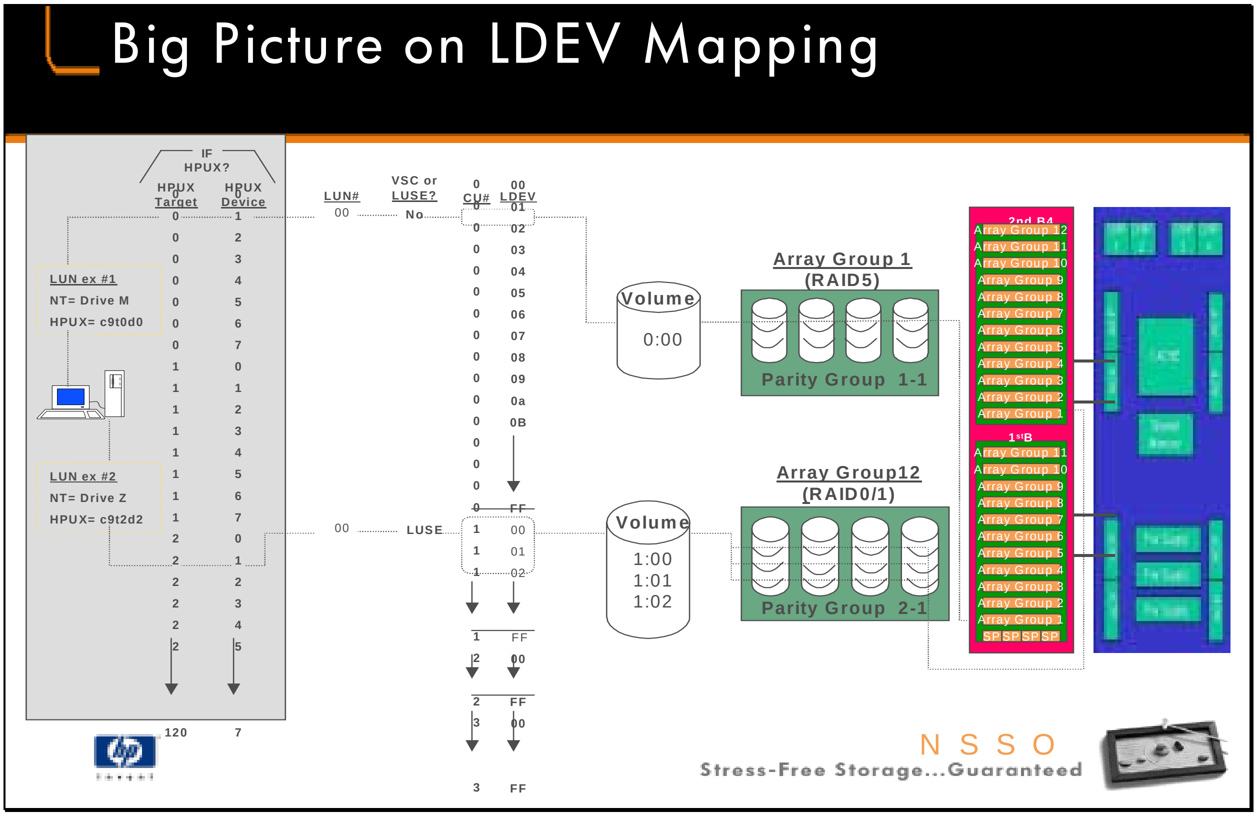
NSSO
Stress-Free Storage...Guaranteed

Host Connectivity



Concepts:

1. More than one host can connect to the same array group, regardless of their RAID definition or ACP pair connection.
2. Once a RAID type had been defined within a domain (all disk connected via the ACP pair) every LDEV within that domain must share the same RAID type.
3. Array groups connected via the same ACP pair can have different defined emulations.
4. HDD's connected via the same ACP pair can be of different physical drive types (ie 18 & 45GB).
5. Disk emulations across an array group must be the same.

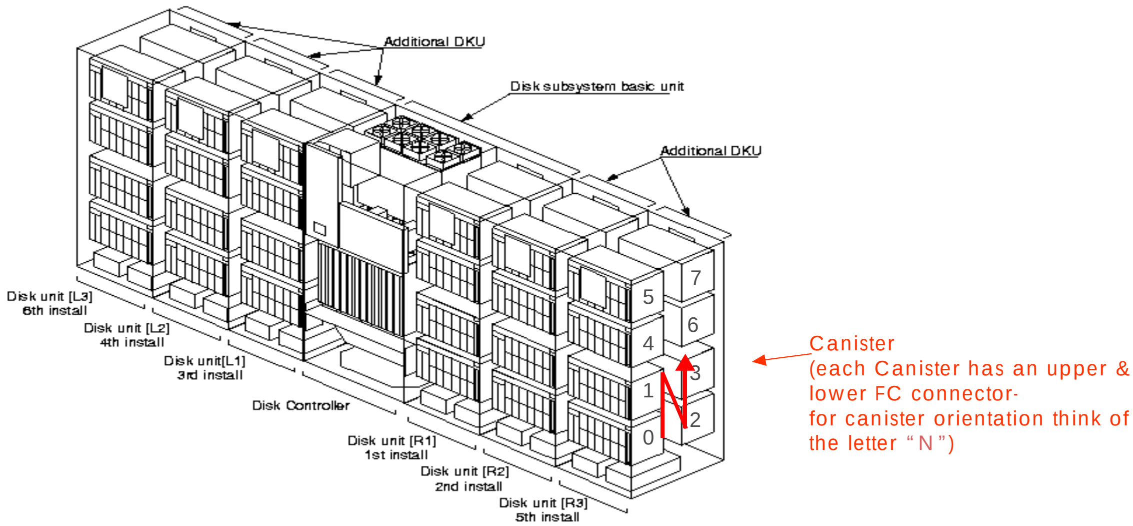


This slide set has not covered Volume Size Configuration (aka CVS) or Logical Unit Size Expansion (LUSE); below are their high level defintions:

CVS: Allows for volume creation of sizes smaller than the defined open emulation. For example, LUN 1a might be only 35MB because volume 0:10 was carved into several 35MB slices. Most common implementations of CVS are to create command devices.

LUSE: Allows for the combination of two or more volumes to create one virtual volume of a capacity larger than the defined emulation. For example, LUN 0a might be 21GB because 0a is really a collective representation of OPEN9 volumes 0:01, 0:02 and 0:03 all assigned to one LUN number.

HP XP512 Canister Identification

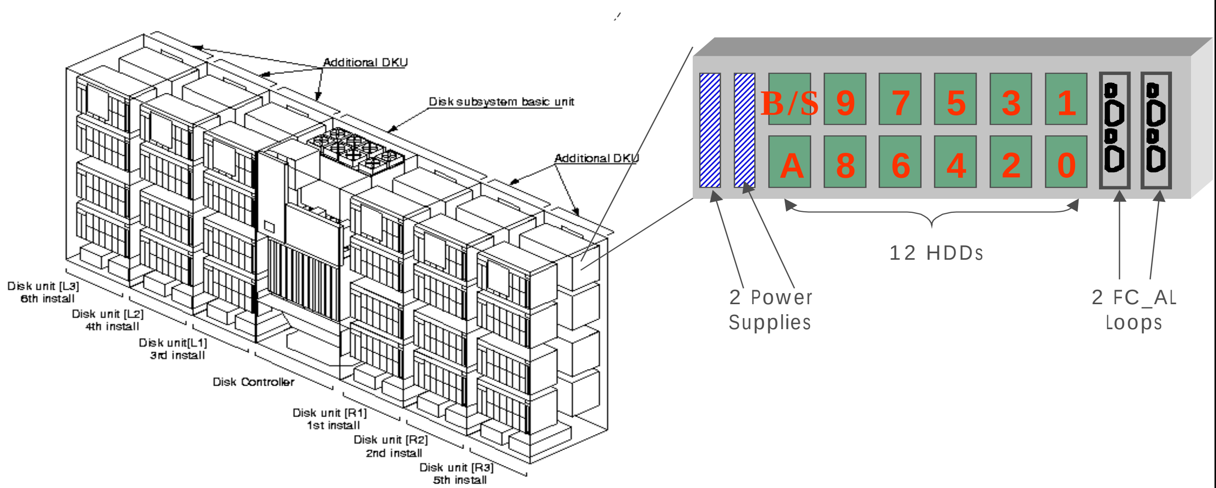


N S S O
Stress-Free Storage...Guaranteed

Hard Disk Drive (**HDD**) are stored inside of canisters. Each DKU contains a total of 8 canisters to hold data or spare disks. The canisters are numbered for logical reference from 0 to 7 (note that canister count is zero based); each canister can hold up to 12 HDDs.

To help with canister orientation, think of the letter "**N**" when trying to understand the canisters location within each of the DKUs. A group of four canisters is referred to as a “Block of Four” or B4. In the example above, we see in DKU R2 two B4’s. The first B4 is canisters 0, 1, 2 and 3 and the second is canisters 4, 5, 6 and 7.

HP XP512 DKU Detail- Canister Components

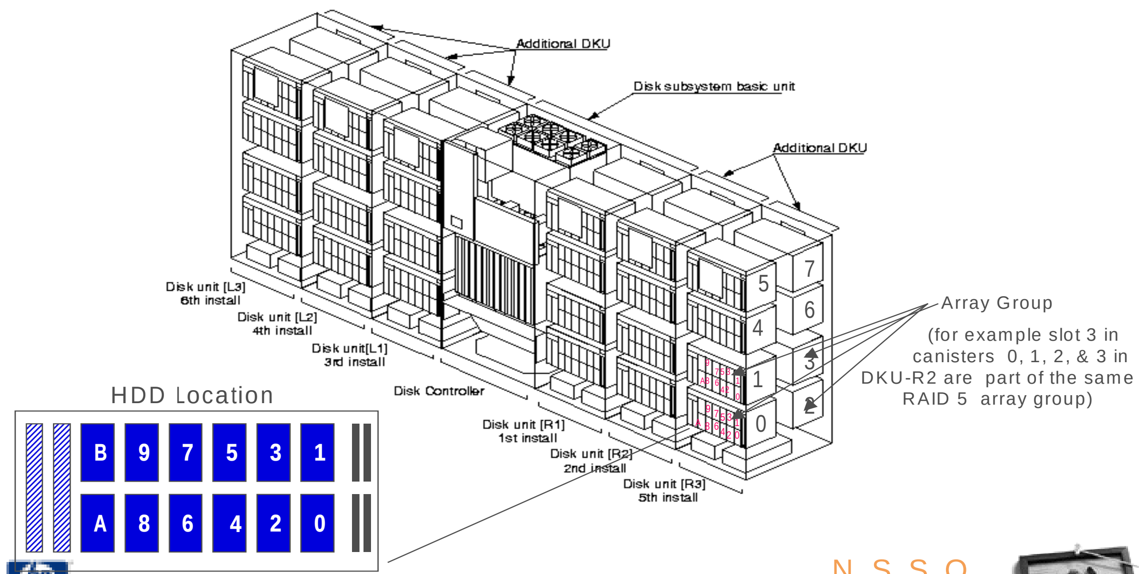


N S S O
Stress-Free Storage...Guaranteed

It is easy to see the canister's hardware redundancy in place that provides access to each hard disk unit in the array. Each disk is dual ported, so it is connected on two different FC_AL loops leading from the ACP to the HDD. Each canister is on two separate power boundaries; if one power supply fails the canister can remain operational.

B/S= Backup or Spare (the slot can be used for either).

HP XP512 Hardware Schematic



N S S O
Stress-Free Storage...Guaranteed

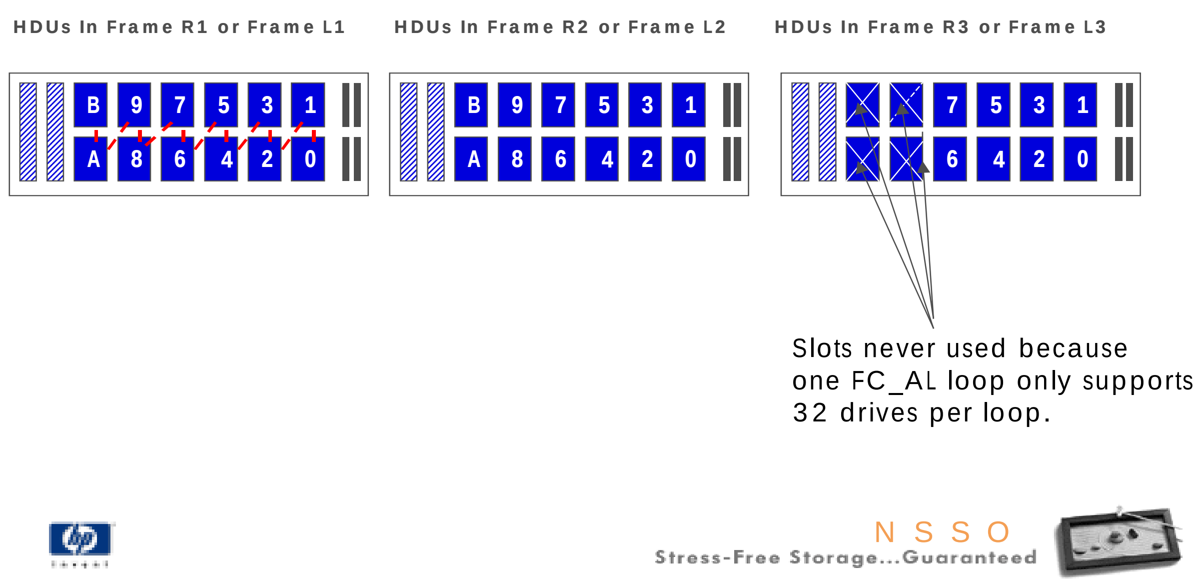
Array Group (aka Marketing Array Group)- A group of four HDDs installed into the XP512 canisters; all four are of the same size.

The slide shows slot 3 of canisters 0, 1, 2 and 3 as an example of an **array group**. All four disk are of the same size. An array group will always be four disk; two visible from the front and two visible from the back.

Disk Size
The HDDs placed in the canisters can be of any size (18, 47GB or 72GB) but each array group must have the same disk sizes. This means slot 3 in canisters 0 through 4 could have 18GB drives and slot 4 in canisters 0 through 4 could have 47GB drives. More about this later….

NOTE: Just like canisters, logical array group numbering and logical hard disk drive location numbering count from zero.

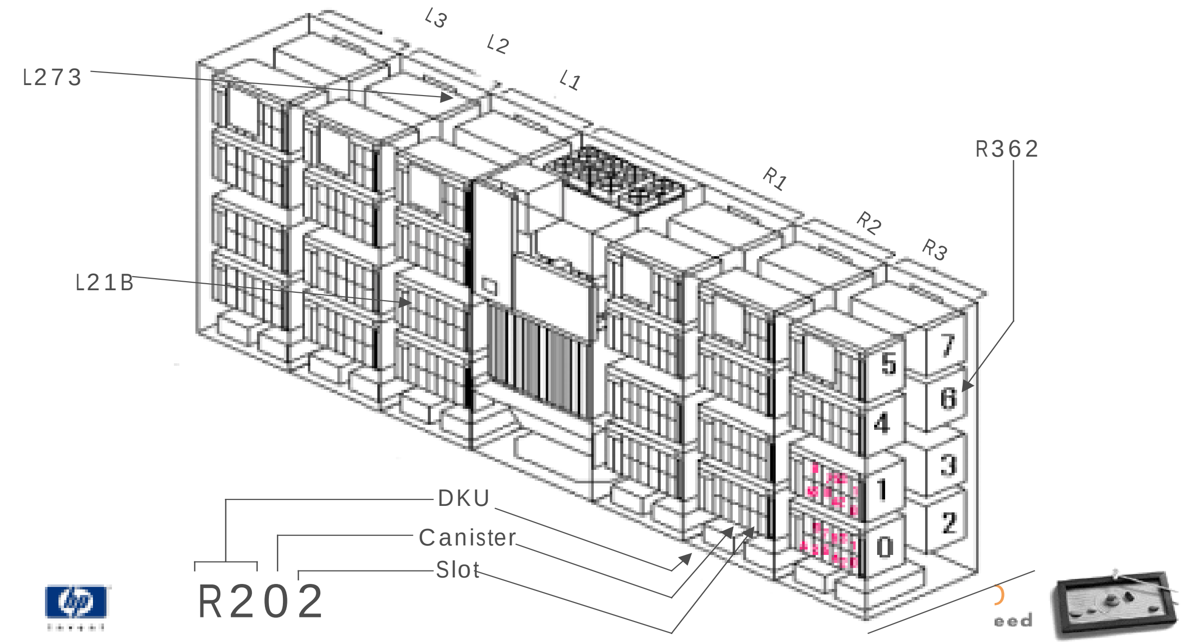
HP XP512- Canister Identification



In frames R3/L3 you would not put a disk in the 9th through 12th slots (unless you had no other place to store a disk; placing a disk in R3/L3 slot 9 through slot b is equivalent to storing that same disk in your car- if a disk is in that slot it will *not* be used).

The numbers representing the disk slots are only logical identifiers. Do not mistake these logical slot identifiers for AL_PA addresses or disk installation order.

HP XP512 HDD Identification



Hard Disk Units (HDD) housed in the DKUs are identified by a combination of their physical position relative to the DKC and their position within each canister. The best way to understand this is to review the example on the slide. Because the XP512 is drawn in three-dimension, arrows pointing to HDDs in the back of the DKUs may not appear as clear or as obvious as the arrows pointing to the HDD in the front of the DKUs.

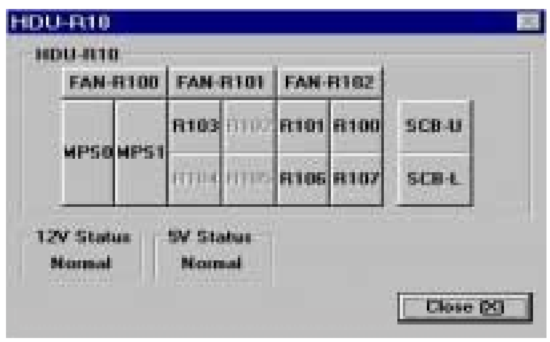
Quiz yourself- remember to think of the canisters by alpha "**N**" :

1. Locate R266?
2. Locate L147?
3. Locate L251?

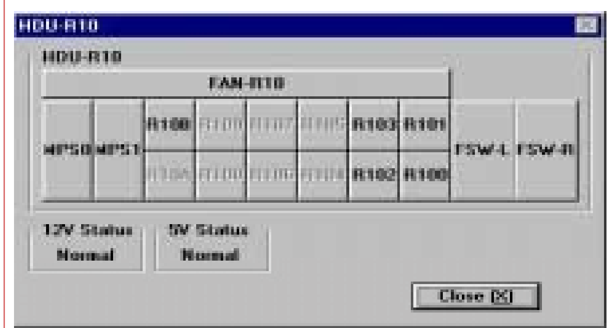
XP512 DKU Architecture

XP512's HDD slot number sequence has changed incomparison to the XP256 (it now matches the HDD installation order sequence within the DKU canisters)

XP256

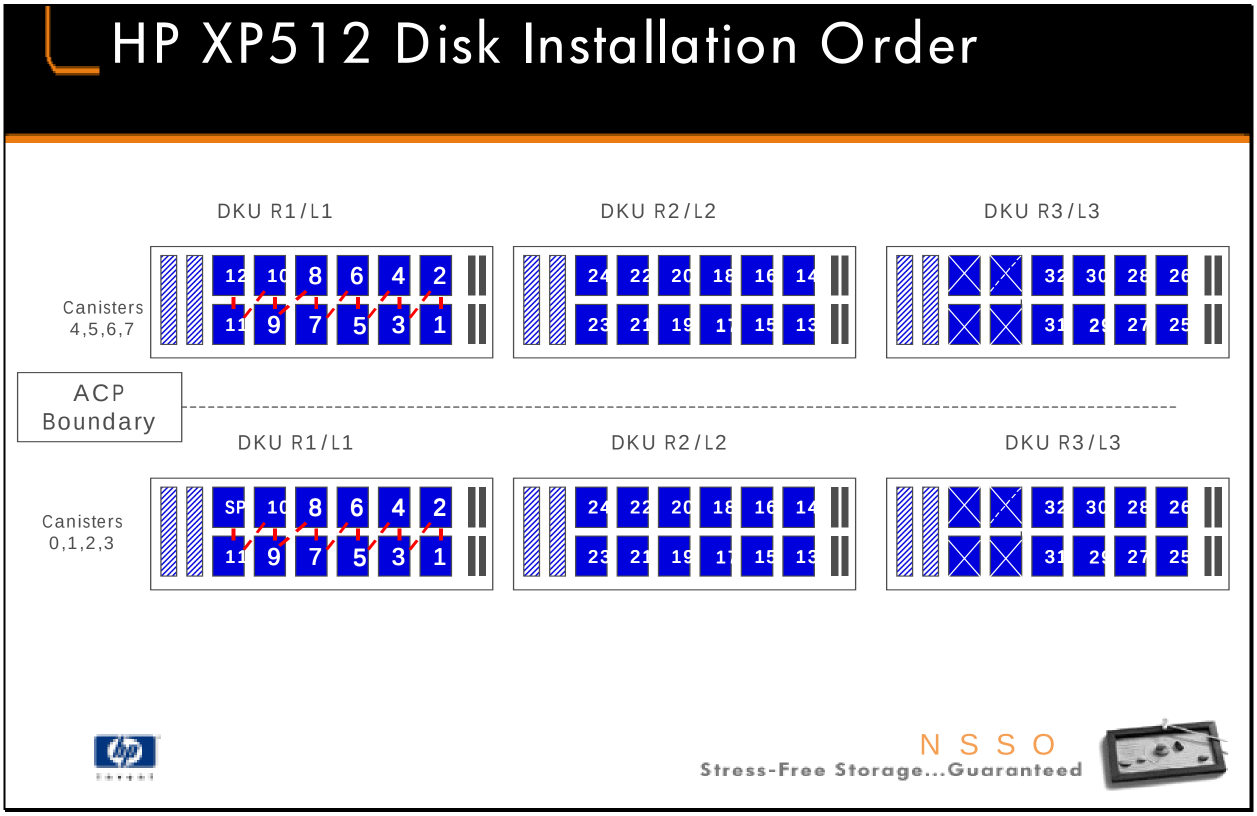


XP512



The above slides the SVP's view of installed HDD per a single canister's view within DKU R1 for each respective array (from the SVP, click MAINTENANCE-INSTALL-DKUx).

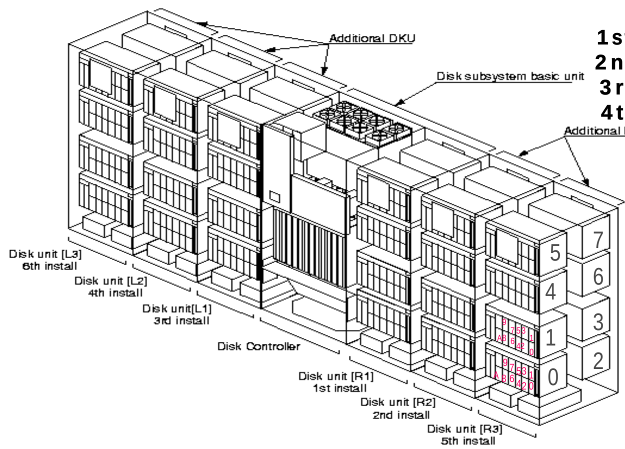
In the XP256, the HDD installation order (1 through 8) was different than the canister slot number (0 to 7). This confusion has been eliminated with the XP512; now the disk installation order sequence now matches the canister slot number (0 through B) sequence.



Disks are sold from HP in groups of 4 ("Marketing array group"). The disks of an array group are placed within a single DKU into four different canisters (two in the front & two in the back) during the installation of a marketing array group.

Notice the disk installation order numbering does not start with numeric zero. Notice that frames R3/L3 only house 8 disk per canister-due to the limitation of FC_AL loop.

HP XP512 ACP Identification



ACP Pair Connections

1st ACP pair: Canister 0, 1, 2 & 3 of R1 & R2 & R3
2nd ACP pair: Canister 4, 5, 6 & 7 of R1 & R2 & R3
3rd ACP pair: Canister 0, 1, 2 & 3 of L1 & L2 & L3
4th ACP pair: Canister 4, 5, 6 & 7 of L1 & L2 & L3

Ex) FC_AL Chains (all duplexed)

Canister 0 within R1 & R2 & R3
Canister 1 within R1 & R2 & R3
Canister 2 within R1 & R2 & R3
Canister 3 within R1 & R2 & R3
Canister 4 within R1 & R2 & R3
Canister 5 within R1 & R2 & R3
Canister 6 within R1 & R2 & R3
Canister 7 within R1 & R2 & R3

N S S O
Stress-Free Storage...Guaranteed

This slide communicates some important configuration information about ACP pairs. When trying to understand ACP pair's physical connection, it is best to think right or left of the DKC and bottom or top of the DKUs:

1st ACP Pair: Right side, DKUs R1 or R1 + R2 + R3, canisters 0-3.
2nd ACP Pairs: Right side, DKUs R1 or R1 + R2 + R3, canisters 4-6.
3rd ACP Pairs: Left side, DKUs L1 or L1 + L2 + L3, canisters 0-3.
4th ACP Pair: Left side, DKUs L1 or L1 + L2 + L3, canisters 4-7.

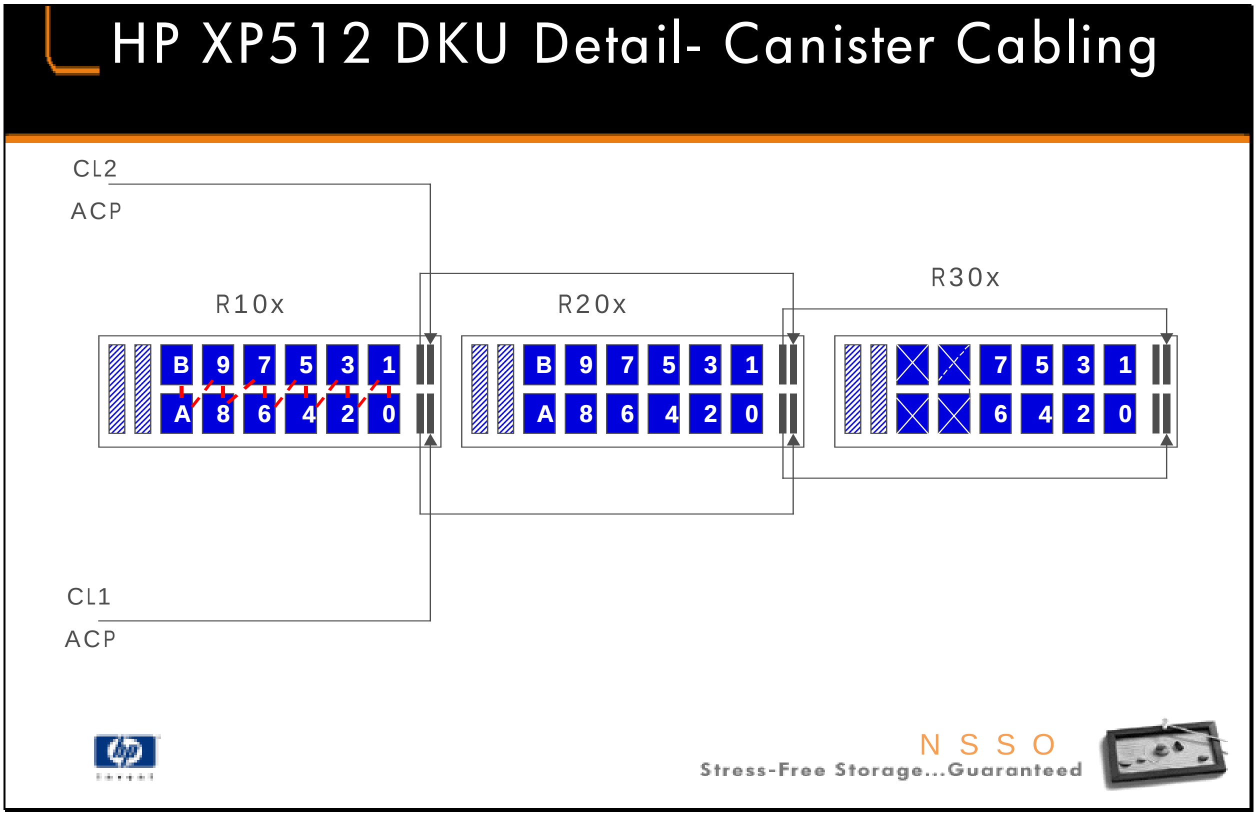
Question: If we ordered one DKC and one DKU (R1), how many ACP pairs would we need to fully utilized the DKU?
Answer: two pairs. One pair for canisters 0-3 and second pair for canisters 4-7.

Question: What if we ordered one DKC, DKU R1 and R2, how many ACP pairs would we need to fully utilized both DKUs?
Answer: two pairs. One pair for R1/R2 canisters 0-3 and second pair for R1/R2 canisters 4-7.

Question: What if we ordered one DKC, DKUs R1 and L1; we will fill R1 full of disk and only fill half of L1 with disk. How many ACP pairs do we need?

Answer: three pairs. One pair for R1 canisters 0-3, one pair for R1 canisters 4-7, and a final pair for L1 canisters 0-3.

Current ISSUE: Even if we had 1 DKC and only R1 DKU with disk only in canisters 0-3, you still would need two ACP pairs. This is under investigation.



A more granular diagram of the first ACP pair's physical connection (rear logic gate to canister) servicing all HDDs in DKU's R1 & R2 & R3, the bottom domain, canister 0 (which could be referred to as R10x ,R20x and R30x).

HDD Spares

- Number and location of dedicated spares:

  R10B, R11B, R12B, R13B, L10B, L11B, L12B, L13B
- Number and location of optional spares

  R14B, R15B, R16B, R17B, L14B, L15B, L16B, L17B

Spare Only	Spare/Data (Option)
x= canister 0, 1, 2 or 3	x= canister 4, 5, 6 or 7

L1xB
R1xB
DKU, Canister, Slot

N S S O
Stress-Free Storage...Guaranteed

Spare Disk Drive Processes:
Dynamic Sparing- if a disk read/write errors exceeds a defined threshold, the system assumes HDD failure and copies all data to a spare disk of the same size.
Correction Copy- Systems encounters an unrecoverable error and must regenerate the data (RAID5) or copy the mirror (RAID0/1) to the spare.

The XP512 can house up to 16 spare disk drives. The rules for dynamic sparing and correction copy of failed HDD has not changed in relation the XP256…
disk drive fail-over can only occur between disk of the same size, regardless of what domain the HDD failure occurred and what domain the HDD spare exist
if a HDD fails, the array can recover the target data using parity data and data stored on the other non-failing drives of the parity group
data saved on spare disk is copied back to the original HDD slot location once the failed drive had been replaced with a functioning HDD.
Spares do not count as parity groups

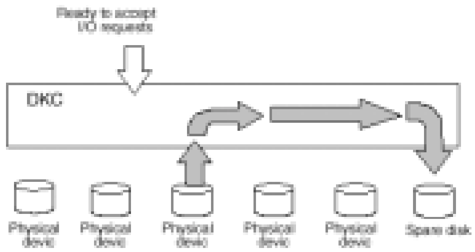
XP512 ACP Spares

Dynamic Spare Disk Detail

- Minimum of one recommended for each capacity disk drive used in the array
- Maximum of 16 spares are configurable (8 “dedicated”, 8 that require tradeoff w/data)
- Any spare of “like capacity” can backup any other disk of the same capacity anywhere in the array
- Dynamic Sparing/Rebuild causes no disruption in host I/O processing
- Dynamic Sparing is controlled by ACP firmware

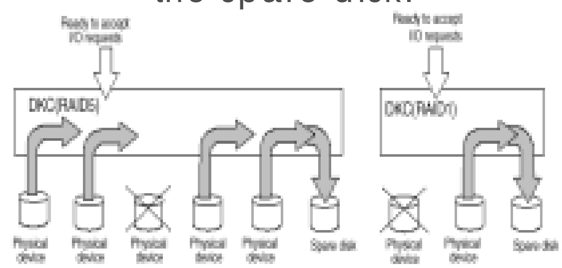
N S S O
Stress-Free Storage...Guaranteed

HP XP512 Dynamic Spare Disk



Dynamic sparing: when the number of errors exceeds a pre-fixed threshold data is copied to the spare disk

Correction copy: in case of unrecoverable error. Data is automatically reconstructed & placed on the spare disk.

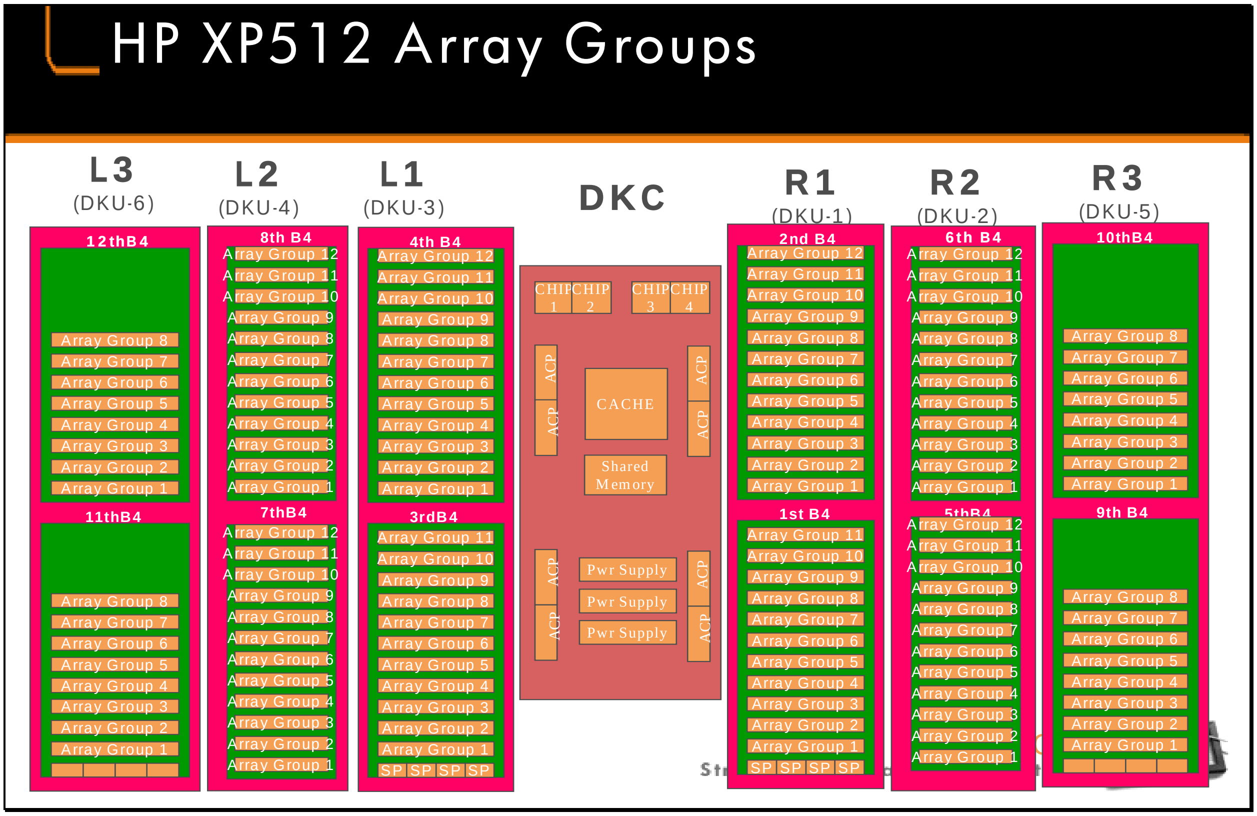


NSSO
Stress-Free Storage...Guaranteed

1. Dynamic sparing
This system keeps track of the number of errors that occurred, for each drive, when it executes normal read or write processing. If the number of errors occurring on a certain drive exceeds a predetermined value, this system considers that the drive is likely to cause unrecoverable errors and automatically copies data from that drive to a spare disk. This function is called dynamic sparing. Dynamic sparing operates the same for RAID 0/1 and RAID5 array groups.

2. Correction copy
When this system cannot read or write data from or to a drive due to an error occurring on that drive, it regenerates the original data for that drive using data from the other drives and the parity data, and copies it onto a spare disk.

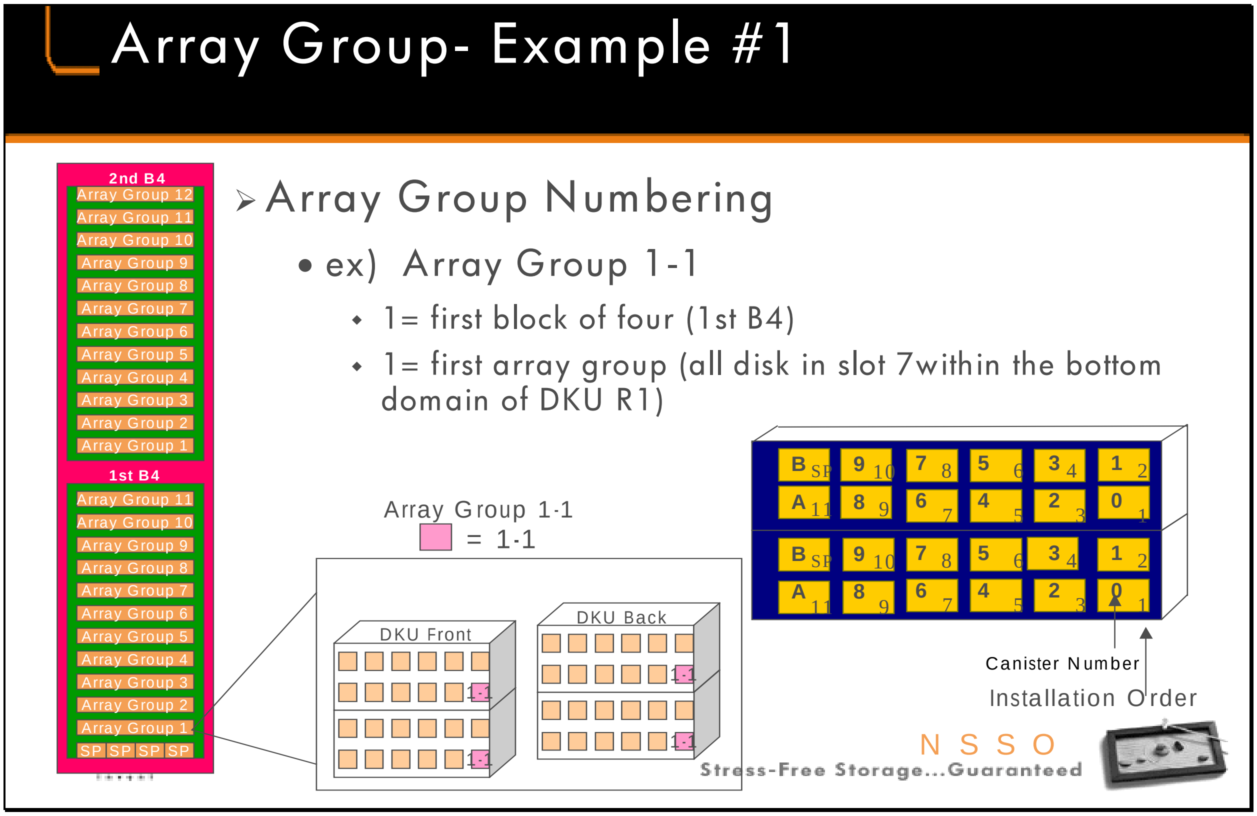
Rules for Sparing:
Drives must be of same capacity (18GB to 18GB)
Sparing can occur across ACP boundaries



A set of four canisters creates a Block of Four (B4); thus each DKU contains two B4's each.

At the bus level, array groups are spanned across DKUs; this is how we reach our maximum of 32 disk per ACP pair. This slide does a good job illustrating the maximum disk per ACP pair and why each domain only supports one raid type.

As mentioned previously there is room for up to 4 spare drives (SP= spare) in frames R1 and L1.



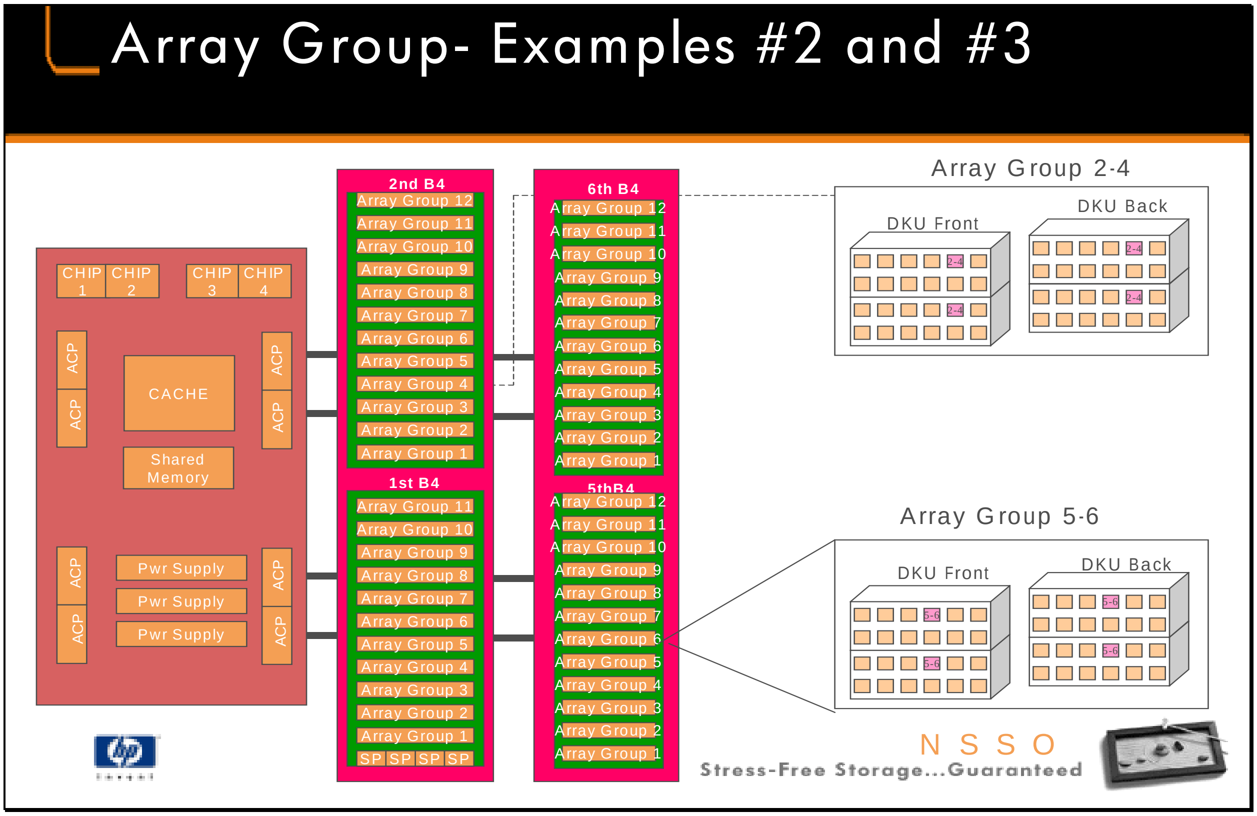
Unlike disk identification (i.e.- R112) which identifies a single disk, parity group number identifies a group of four disk (RAID0/1 or RAID5). Parity group numbers are based on two values. The first value is the B4 identifier and the second number is based on disk installation order (not the canister number).

The second digit is the disk number as defined by its **installation order**. Since disk are always sold in groups of four, the disks within that set will be installed in the same canister slots within one DKU (i.e.- canister 0, 1, 2 and 3, or canisters 4, 5, 6, and 7).

Example #1: Parity Group 1-1 would be a set of four disk, found in the first B4 of canisters 0, 1, 2 and 3 inside DKU R1 (identified as HDDs R100/R110/R120/R130). Now that we know Parity Group 1-x, how to do we know what "x" is? Because this marketing array group was install first, then we know "x" is equal to 1, thus giving us Parity Group 1-1.
Array group numbering tells us...
-disk's physical location in the array
-helps us understand which ACP pair controls the disk, therefore we know the raid type that must be in place (or defined)
Unlike disk identification (i.e.- R112) which identifies a single disk, parity group number identifies a group of four disk (RAID0/1 or RAID5). Parity group numbers are based on two values. The first value is the B4 identifier and the second number is based on disk installation order (not the canister number).

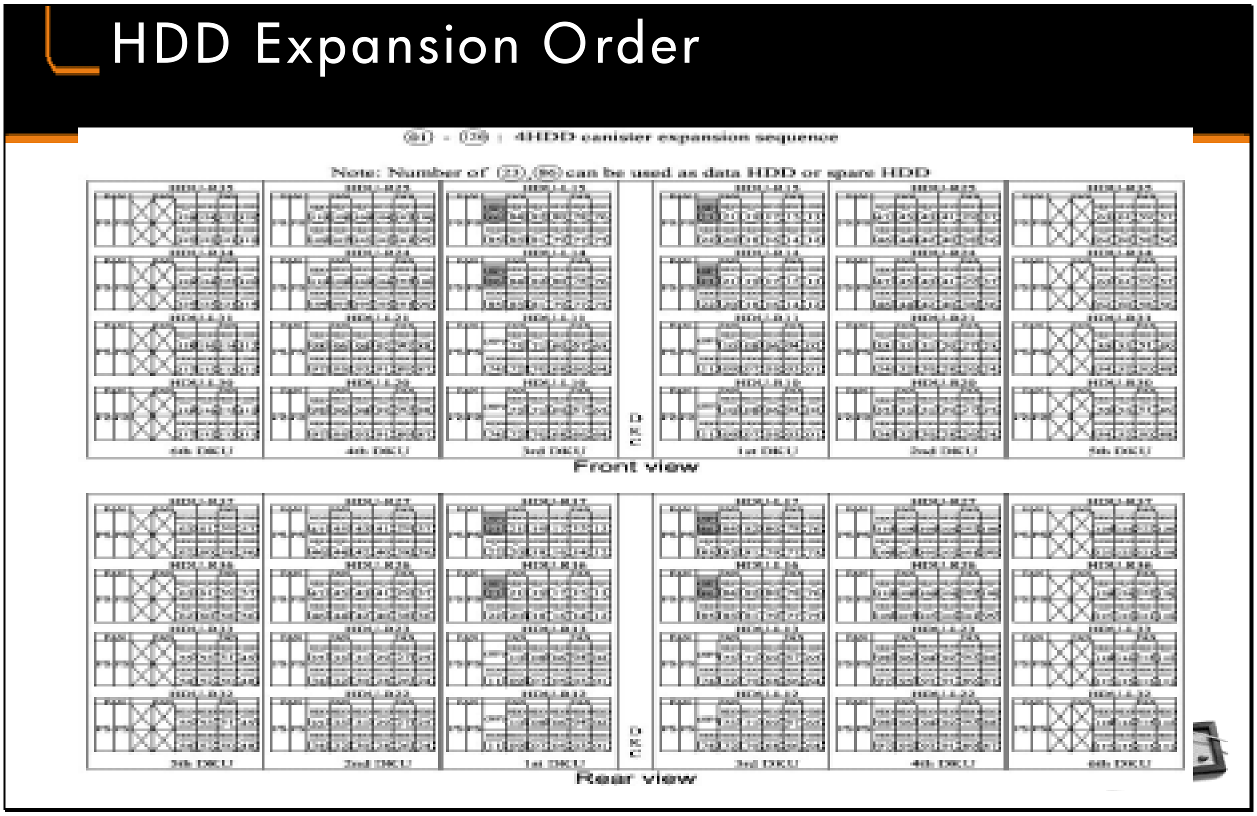
The second digit is the disk number as defined by its **installation order**. Since disk are always sold in groups of four, the disks within that set will be installed in the same canister slots within one DKU (i.e.- canister 0, 1, 2 and 3, or canisters 4, 5, 6, and 7).

Example #1: Parity Group 1-1 would be a set of four disk, found in the first B4 of canisters 0, 1, 2 and 3 inside DKU R1 (identified as HDDs R100/R110/R120/R130). Now that we know Parity Group 1-x, how to do we know what "x" is? Because this marketing array group was install first, then we know "x" is equal to 1, thus giving us Parity Group 1-1.
Array group numbering tells us...
-disk's physical location in the array
-helps us understand which ACP pair controls the disk, therefore we know the raid type that must be in place (or defined)



Example #2: Parity Group 2-4
The first number of the parity groups is 2, thus we know we are referencing the second B4 (which is always located in upper DKU R1). The second number is 4 so it should be the fourth array group installed, which includes slots R143/R153/R163/R173.

Example #3: Parity group 5-6
The first number is 5, referring to the fifth B4 (located in lower DKU R2). The second number is 6, so it should be the sixth array group installed in this B4, which includes slots R205/R215/R225/R235.



Canister numbers are 0-f

Like the XP256, the total number of HDD that can be operated by the XP512 does not match the total number of available HDD canister slots available. The XP256 could only use 240 of the available 256 slots, thus it was called the XP256. The XP512 uses only 512 of the available 576 slots.

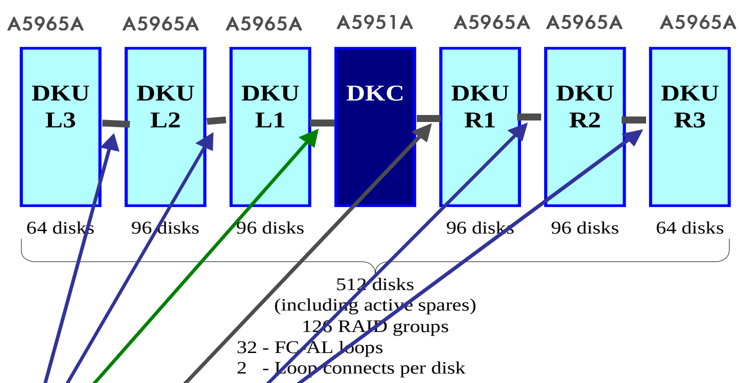
Each canister contains 12 HDDs slots; each HDU can house up to 8 canisters thus we get 96 HDD slots per canisters. Each XP512 DKC can receive up to 6 DKUs, thus the total number of HDD canister slots is 576. DKU R3 and L3 can only utilized 6 slots per canister because the 128 AL_PAs per ACP board limitation (FC limitation), therefore we arrive at 576 minus 64 equals 512.

Assuming the fully packed array is not using slots 23 and 86 for spare HDD, this allows the array to have 126 array groups. So how many parity groups do we get from this configuration? If running all RAID5, all RAID0/1 or a combination of RAID5 and RAID0/1 the total number of parity groups would be 126.
Parity Group numbering is static. This means that if we buy one DKC and only have DKU's R1 and R2 installed, we would have parity groups 1-x, 2-x, 5-x and 6-x. The

only way to truly know a volume's (CU:LDEV) physical location is to know what parity group (aka RAID group) that volume resides.

Spares do not count as parity groups, thus parity group numbering for 1-x and 3-x only range from 1 to b and parity number for 2-x and 4-x may or may not (depends if spares are utilized) range up to 2-c and 4-c.

XP512 Cabinet Capabilities



Required Cables:

1) DKC to L1: A5974A

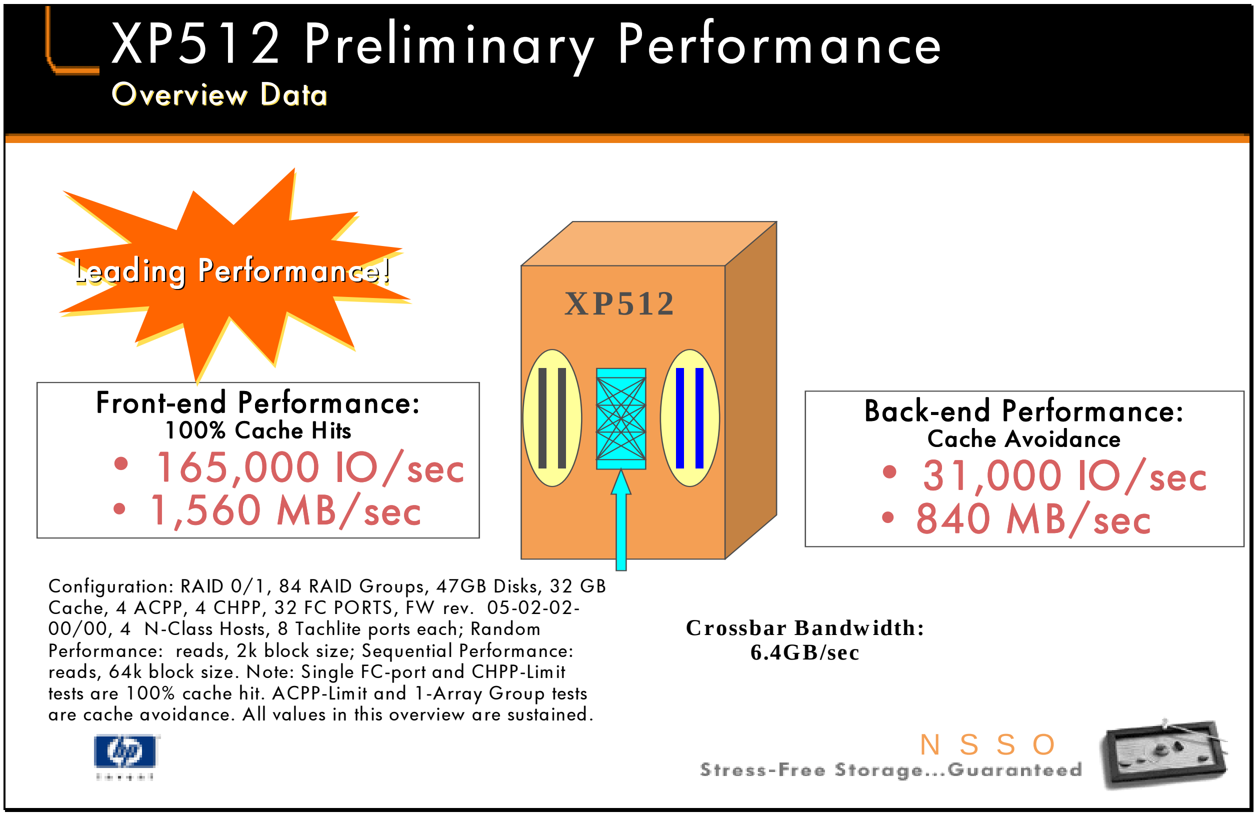
2) DKU to DKU: A5975A

3) DKC to R1: Included with Base XP512 product

NSSO Stress-Free Storage...Guaranteed

XP512 Highlights

- More Capacity
- From 72GB to 37TB capacity
- 18, 46 and 72GB 10,000 RPM Fibre Channel Disks
- Twice the number of connectivity ports
- 32 host connect ports (either FC-AL, or ESCON connect)
- Dual and concurrently active data paths
- SCSI connect through a FC-AL bridge
- More Disk Drives supported, faster drive buses
- 512 drives supported
- Denser disk drive packaging 96 disk drives/DKU
- Dense data center packaging
- Disk drives native FC-AL
- 32 FC-AL loops
- Each FC-AL has it’s own i960 ACP processor
- Large Cache
- 32GB Mirrored Cache Memory
- Same SW/FW functionality as XP256
- Continuous Access (CA) compatibility between XP512 and XP256



The XP512 uses point-to-point interconnects between both the CHIP's and ACP's to shared memory. There are four direct connect port connections between shared memory and each CHIP (and ACP) board. The XP512 can be configured with up to 64 ports to shared memory. The connection to the shared memory is a point-to-point connection. Like the independent power domains with the CHIP boards, both the shared and cache memory are split on dual power sources.
The four direct connect shared memory connections are split with two connections to each shared memory power domain. The connection from the CHIP's (and ACP's) to cache memory is a point-to-point connection that goes through a crossbar switch. Each CHIP (and ACP) has 2 ports that connect to the crossbar switch. Each crossbar switch is also on a separate power domain. The XP512 can be configured with 16 ports to cache memory. The other side of the crossbar switch connects directly to the cache memory that is grouped into four blocks in two power domains. To achieve the full cache
bandwidth between the crossbar switches and cache memory, all four cache boards need to be installed. To maximize the bandwidth of cache, the cache memory should be distributed evenly across all cache boards installed. The point-to-point direct connect architecture that the XP512 uses also adds to the robust nature of the disk array in the event of a bus path failure. The XP512 with the point-to-point direct connects

will only have minor performance degradation with a path failure. The path failure in cache will only reduce the system performance by 1/16 (~6.3%), and a path failure in shared memory will only reduce the system performance by 1/64 (~1.6%).

XP512/48 Performance Comparisons

XP48 Maximum Performance Overview

Front-end Performance:
100% Cache Hits
- 80,000 IO/sec
- 840 MB/sec

Back-end Performance:
Cache Avoidance
- 4,500 IO/sec
- 318 MB/sec

- 11 Array Groups Maximum / 4 Typical
- Up to 16 GB cache / no cache platform board
- Up to 3 FC or ESCON CHIP Pairs / Available 4-Port FC CHIP

XP512 Maximum Performance Overview

Front-end Performance:
100% Cache Hits
- 165,000 IO/sec
- 1,560 MB/sec

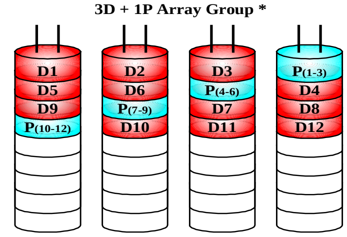
Back-end Performance:
Cache Avoidance
- 31,000 IO/sec
- 840 MB/sec

- 126 Array Groups Maximum / 11 Typical
- Up to 4 ACP Pairs
- Up to 32 GB cache / cache platform board for maximum throughput
- Up to 4 FC or ESCON CHIP Pairs

XP512 RAID Levels

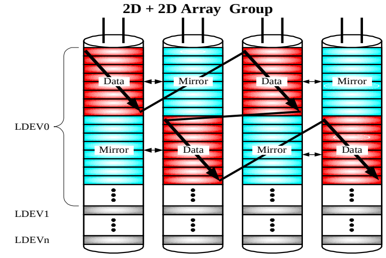
- RAID 5
  - Stripes parity across disks
  - 25% storage overhead
  - More cost effective $/usable MB
  - Excellent performance
- RAID 0/1
  - Striped and mirrored (aka. "dual read" RAID 1)
  - 50% storage overhead
  - Best for performance sensitive applications
  - Better read & write performance

3D + 1P Array Group *



* Shows logical not physical view

2D + 2D Array Group



RAID 0/1 is a striped and mirrored copy within a ACP pair

NSSO
Stress-Free Storage...Guaranteed

As mentioned earlier RAID 0/1 adds high I/O performance features of RAID0 striping to the high reliability features of mirrored storage with RAID1. RAID 0/1 uses all four disks in an array group for the primary storage path. The I/O performance improvements are achieved from the fact that the I/O data streams are split onto all four disks in the array group. Striping the data across the four disks more than doubles the performance over the XP256 with RAID 1. From the diagram above one can see how the primary data is physically written across all four drives. The diagram also shows how the secondary portion of the storage array is used for the mirrored copy. Writes to the array in RAID 0/1 mode occur the same as explained above for RAID 1. The primary side is de-staged first. After the primary data has been successfully de-staged, the secondary data is
de-staged to the mirror drive mechanisms.

LDEV Creation

- Determine usable disk and performance requirements for the XP512 array then:
- Select physical disk size: 18GB, 47 GB or 73GB
- Select RAID level: RAID 0/1 or 5
- Assign emulation types: Open 3, 8, 9, E, M, or L
- Assign LDEV's to open volumes
- Assign LUN's to LDEV's

NSSO
Stress-Free Storage...Guaranteed

XP512 Power Options

- Three phase: 50 or 60 Hz:
  Large data centers with heavy power loads
- Single phase: 50 or 60 Hz:
  Smaller data centers that may not be wired for 3-phase

	3-Phase AC		1-Phase AC	
	60Hz	50Hz	60Hz	50Hz
DKC	A5951A		A5951A Option 001	A5951A Option 002
DKU	A5965A Option 001	A5965A Option 002	A5965A Option 002	A5965A Option 003

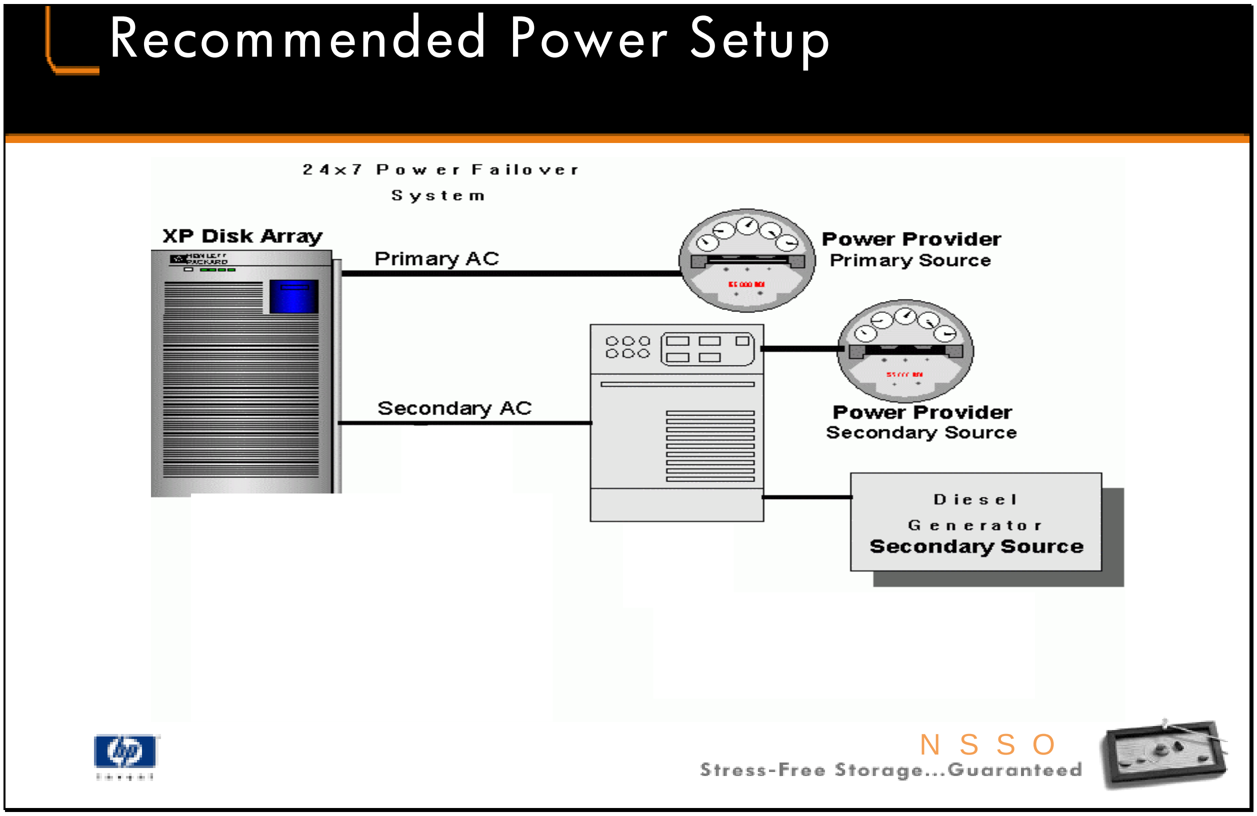
NSSO
Stress-Free Storage...Guaranteed

Battery & Power

Power supply
- Single phase= 14 AC cords (DKC has 2)
- Three phase= 12 AC cords (DKC powers R1)

XP512 battery backup is the same as the XP256
- 48 hours maximum for cache
- 96 hours maximum for shared memory

NSSO
Stress-Free Storage...Guaranteed



Illustrated above is the recommended uses of UPS and redundant power….
Dual power supply from two separate providers
Diesel generator backing up the secondary power source; both connected to the UPS

The internal architecture of the XP256 and XP512 has no single point of failure, but the XP unit(s) is reliant upon AC power. In the event of total power loss to the array, the XP will maintain data in cache via internal battery units for a total of 48 hours maintain shared memory for 96 hours. If power is not restored data held in cache and system configuration information will be lost. System configuration information can only be restored by reloading from floppy (assumes disk #31 has been updated with each change to the array) or from flash memory when the system is restarted. To extend this interval, an Uninteruptable Power Supply (UPS) is required. However, in the event of a total power failure, the array is now at mercy of the UPS. This vulnerability can be eliminated by using the xxpf program supplied by HP.
3 XPPF (XP Power Fail Program)
3.1 Description
This program will enable, disable, and query the Power Fail Mode of the XP disk array. Scripts have been provided to automate these tasks using command line parameters.
XPPF Files included:

xppf - executable program.
xppf_on - script to set power fail mode to on.
xppf_off - script to set power fail mode to off.
xppf_q - script to return the status of power fail mode.
This program requires the following:
RAID Manager Lib - 01-01-01/00 or later.
XP 256 Microcode version - 52-45-01/00 or later.
XP 512 Microcode version - 01-10-90 00/00 or later.

XP512 Base Hardware Configuration for Shipment

Base Configuration for Shipment
2 GB cache
512 MB shared memory
Redundant power supplies for CHIP pairs 1 and 2
1st ACP pair for domain #1 (right lower)
Fibre Channel Device Interconnect cables from DKC to frame R1
HP microcode
RAID Manager Library
Continuous Track XP
pcAnywhere
SNMP kit
Modem

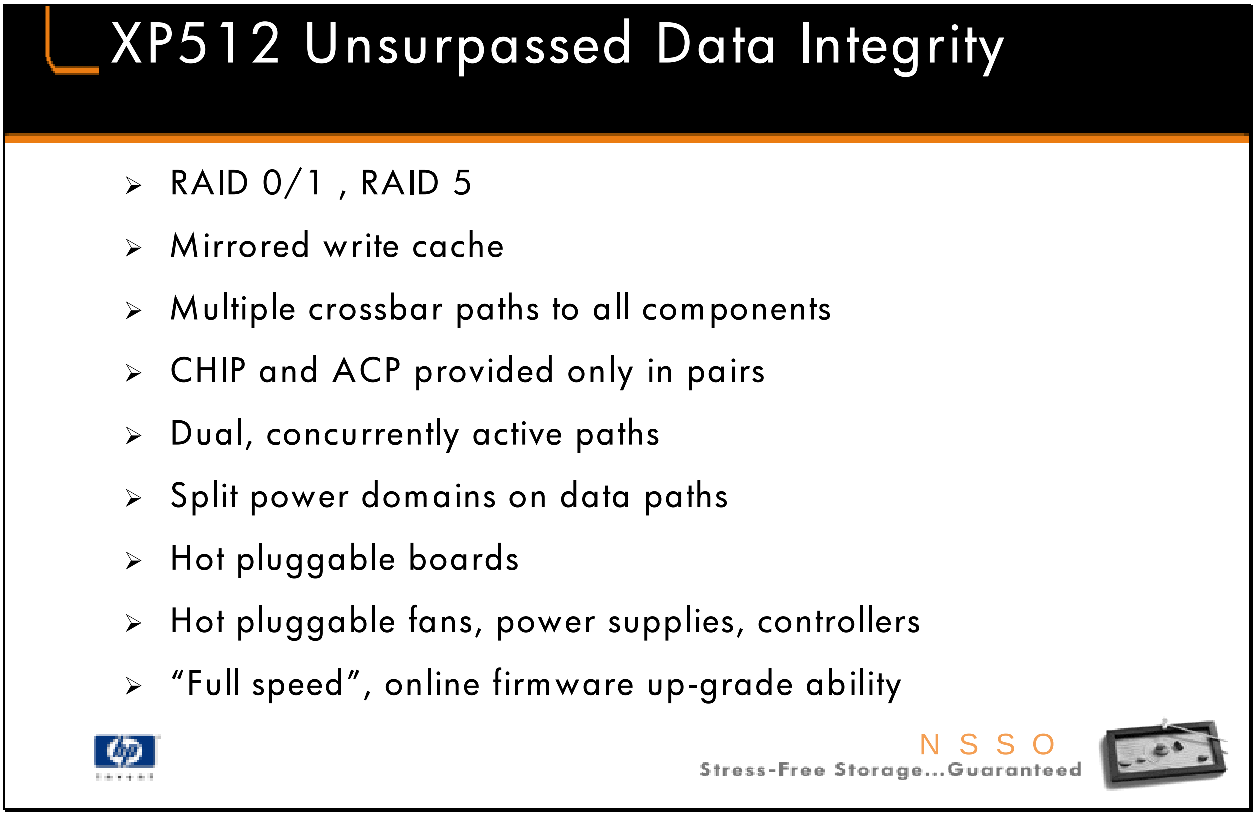
NSSO
Stress-Free Storage...Guaranteed

XP512 & XP256 Summary Comparison

#	Item	Target	RAID400 HSN		RAID300 BUS	
			Shared Memory	Cache Memory	Shared Memory	Cache Memory
			XP512		XP256	
	Performance	GB throughput	4 bit 100 MHz 64 paths/DKC 3.2 GB/s (Gross) 44 MB/s (Net)	16 bit 100 MHz 16 paths/DKC 3.2 GB/s (Gross) 2.1 GB/s (Net)	32bit 33MHz 2 paths/DKC 270 MB/s (Gross) 17 MB/s (Net)	128bit 30MHz 2 paths/DKC 480 MB/s (Gross) 320 MB/s (Net)
	Reliability	Easy to Maintenance	Reduce the number of maintenance target PCBs		Indicate many maintenance target PCBs	
			2 PCBs / max.	3 PCBs / max.	20 PCBs / max.	22 PCBs / max.
		Redundant Path	2 alternative paths for a side of SM by each PK	2 alternative paths by each MP PK	2 alternative paths as system common	2 alternative paths as system common
		Throughput under path failure	63 / 64	15 / 16	1 / 2	1 / 2

NSSO
Stress-Free Storage...Guaranteed

Note: The speed of the XP512 TachLite processor is 80MHz; the speed of the wires is 100MHz.


XP512 Unsurpassed Data Integrity
RAID 0/1 , RAID 5
Mirrored write cache
Multiple crossbar paths to all components
CHIP and ACP provided only in pairs
Dual, concurrently active paths
Split power domains on data paths
Hot pluggable boards
Hot pluggable fans, power supplies, controllers
"Full speed", online firmware up-grade ability
N S S O
Stress-Free Storage...Guaranteed

XP512 Configuration Rules

- One RAID level per ACP pair (New firmware allows mixing RAID levels on a ACP pair)
- 1-16 dynamic spares available
  - 126 array groups available (with 8 spares)
  - 124 array groups (with 16 spares)
- Cache expansion board
  - Required for >16 GB cache
  - Very large configurations only!!
- Cache: 2 GB standard in base configuration, expandable in 2 GB increments to 32 GB
- Shared Memory: 512 MB standard in base configuration, expandable in 256 MB increments to 1.28 GB

N S S O
Stress-Free Storage...Guaranteed

XP512 supported server platforms For FibreChannel and ESCON CHIP

- HP-UX (See Note 3)
- MPE (See Note 3)
- Windows 2000 / NT4.0 (See Note 3)
- Solaris
- AIX (See Note 3)
- IBM S390 compatible mainframes
- Linux, Netware

NSSO
Stress-Free Storage...Guaranteed

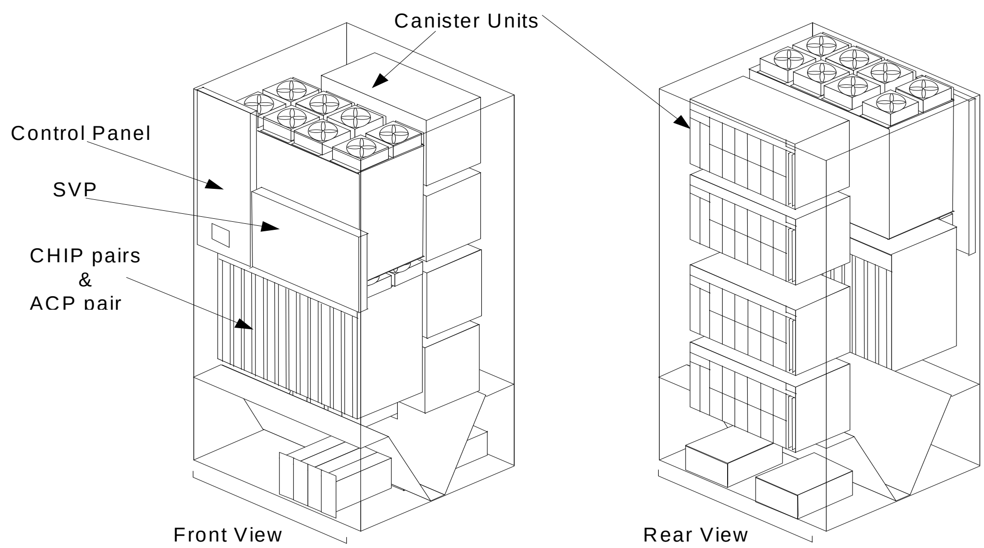
NOTES:
1. More details about supported server configurations may be found in the XP48 & 512 Product
Software and Connectivity Summary available on ESP and Partnership web.
2. All FibreChannel support includes basic and HA connectivity
3. FWD SCSI connectivity using a bridge front end includes only HPUX, MPE and NT, AIX
non-HA only.

XP48 product details

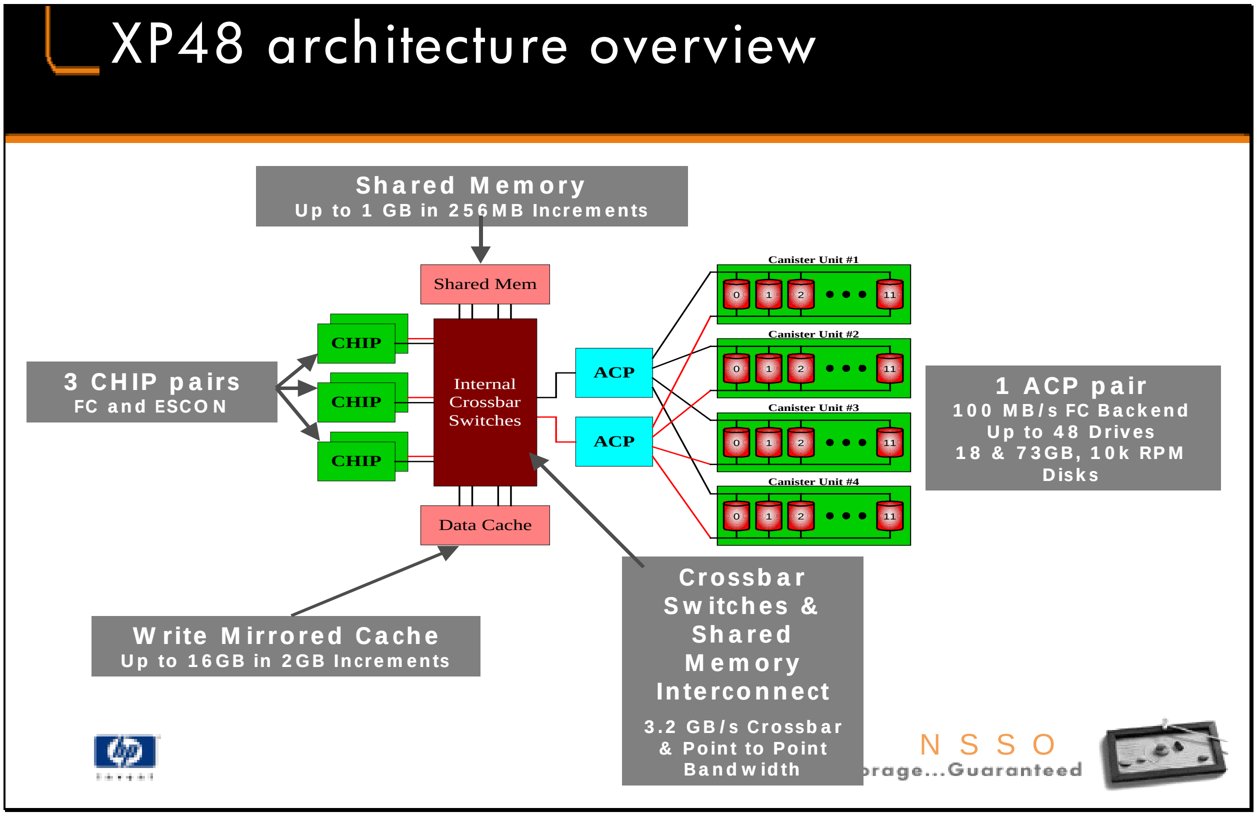
- Efficient Single Cabinet Frame
- 3.2 GB /sec. Crossbar Switch
- 72 GB to 3.5 TB capacity (w/73 GB disks)
- 8 to 24 Fibre Channel or ESCON host ports
  (Optional 4-port Fibre Channel CHIP)
- 4 to 48 native FC disks
- 2 to 16 GB mirrored cache
- 1 ACP Pair
- RAID 0/1, and 5 Capabilities
- 512MB to 1 GB Shared Memory
- 18GB and 73GB, 10k RPM FibreChannel Disk Drives
- Single Phase Power (50-60Hz) only
- Full support of all XP software titles and capabilities

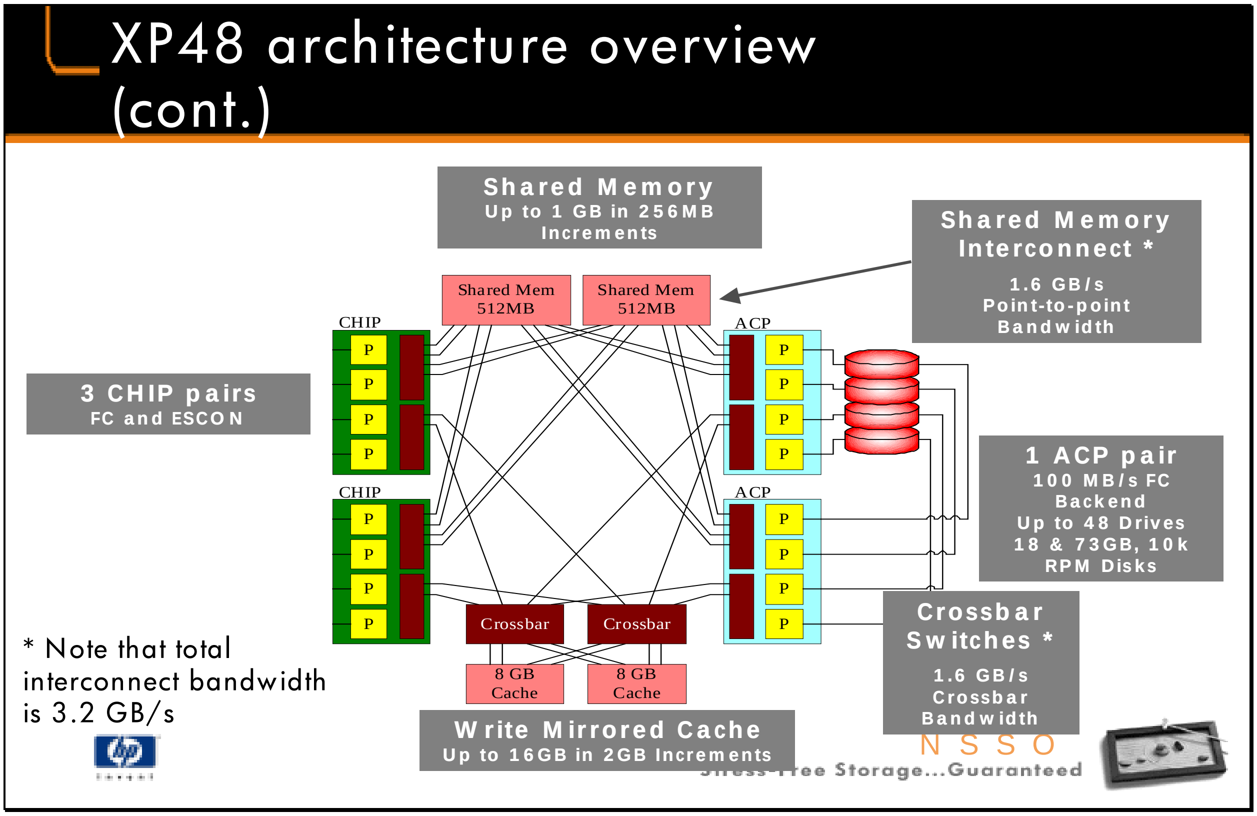
NSSO
Stress-Free Storage...Guaranteed

XP48 system layout

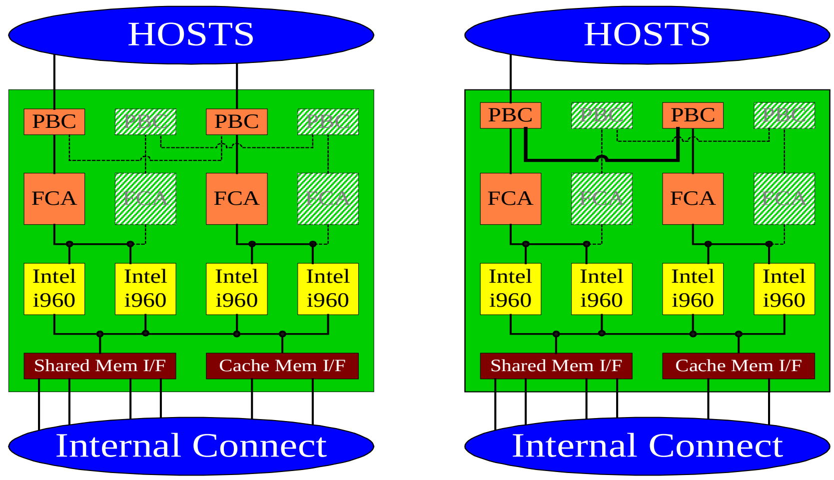


N S S O
Stress-Free Storage...Guaranteed


XP48 architecture overview
Shared Memory
Up to 1 GB in 256MB Increments
Shared Mem
3 CHIP pairs
FC and ESCON
CHIP
CHIP
CHIP
Internal Crossbar Switches
ACP
ACP
Canister Unit #1
Canister Unit #2
Canister Unit #3
Canister Unit #4
0
1
2
11
1 ACP pair
100 MB/s FC Backend
Up to 48 Drives
18 & 73GB, 10k RPM Disks
Data Cache
Write Mirrored Cache
Up to 16GB in 2GB Increments
Crossbar Switches & Shared Memory Interconnect
3.2 GB/s Crossbar & Point to Point Bandwidth
N S S O
orage...Guaranteed


XP48 architecture overview (cont.)
Shared Memory
Up to 1 GB in 256MB Increments
Shared Memory Interconnect *
1.6 GB/s Point-to-point Bandwidth
Shared Mem 512MB
Shared Mem 512MB
CHIP
CHIP
ACP
ACP
P
3 CHIP pairs
FC and ESCON
1 ACP pair
100 MB/s FC Backend
Up to 48 Drives
18 & 73GB, 10k RPM Disks
Crossbar
Crossbar
8 GB Cache
8 GB Cache
Crossbar Switches *
1.6 GB/s Crossbar Bandwidth
* Note that total interconnect bandwidth is 3.2 GB/s
Write Mirrored Cache
Up to 16GB in 2GB Increments
N S S O
Storage...Guaranteed

2-port CHIP



(Standard Mode) **(High Performance Mode)**

2 port CHIP

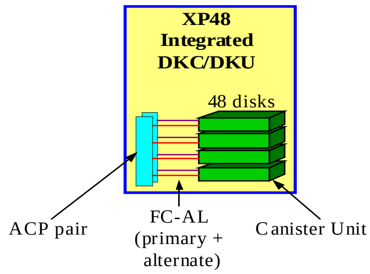
N S S O
Stress-Free Storage...Guaranteed

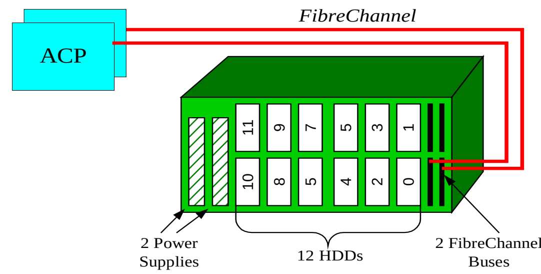
Automatic Odd/Even LUN mapping in Firmware
4, 80 MHz i960 processors vs. 4, 66 MHz i960's with the XP256
Standard Mode
- 2 independent "halves" 10k IOP blocks per CHIP card

High Performance Mode
- All 4 i960's work to service IO's for 20k IOPS
- Ports C & D disabled
- Maximum IO's achieved through LUN mapping and sharing "virtual LUN's" with A & B (balance between all 4 ports - but ports A&B only active)
- Caution one port, A for instance can take all bandwidth in HP mode

LDEV Count (total): 4096

XP48 disk mounts & canisters





- Mount inside the single integrated frame
- Canisters hold up to 12 disk drives, frame hold 4 canisters
- Connect to FC arbitrated loop to primary and secondary ACP channels
- Array groups require 4 canister units to form a disk group

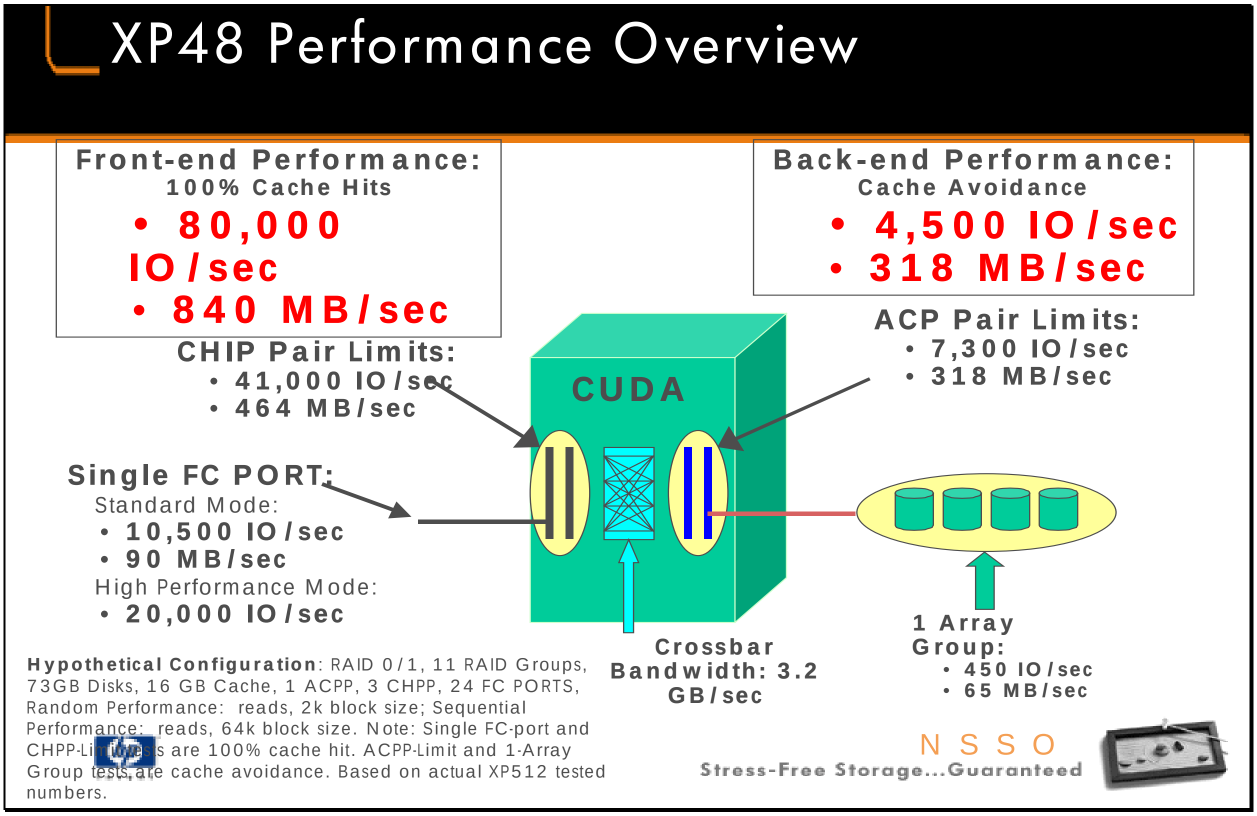
N S S O
Stress-Free Storage...Guaranteed

XP48 configuration rules

- **Multiple RAID levels per ACP pair**
  - RAID 0/1 or 5 (will be supported in XP512 also)
- **Up to 11, 4 disk array group**
  - 18 or 73 GB
- **1 to 3 Channel Host Interface Processor Pairs**
  - 4 or 8 port Fibre Channel
  - 4 or 8 Channel ESCON
- **1 Array Control Processor Pairs**
  - Integral to the XP48 frame
- **1-4 dynamic spares available**
  - 11 array groups available (with 4 spares)
- **Cache:** 2 GB standard in base XP48 frame configuration, expandable in 2 GB increments to 16 GB
- **Shared Memory:** 512 MB standard in base XP48 frame configuration, expandable in 256 MB increments to 1 GB
- **LDEV/LUN:** XP48 supports up to 4096 LDEV's and 6144 LUN's (vs. 4096 LDEVs and 8092 LUN's on the XP512)
- **Parts Compatible** with the XP512 but all parts have different product numbers

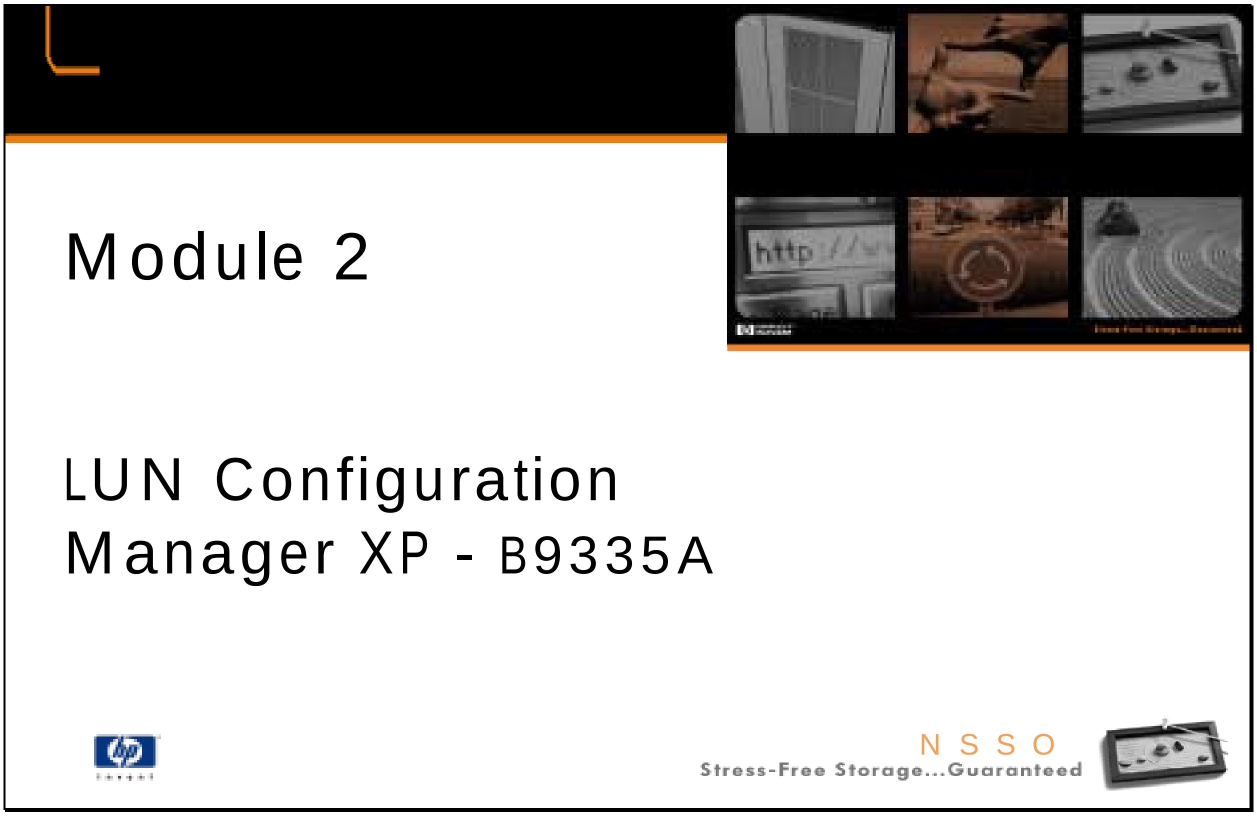
N S S O
Stress-Free Storage...Guaranteed

Multiple RAID levels on a single ACP pair also be supported in the XP512 with firmware release 01.11.22.00/00 or newer.

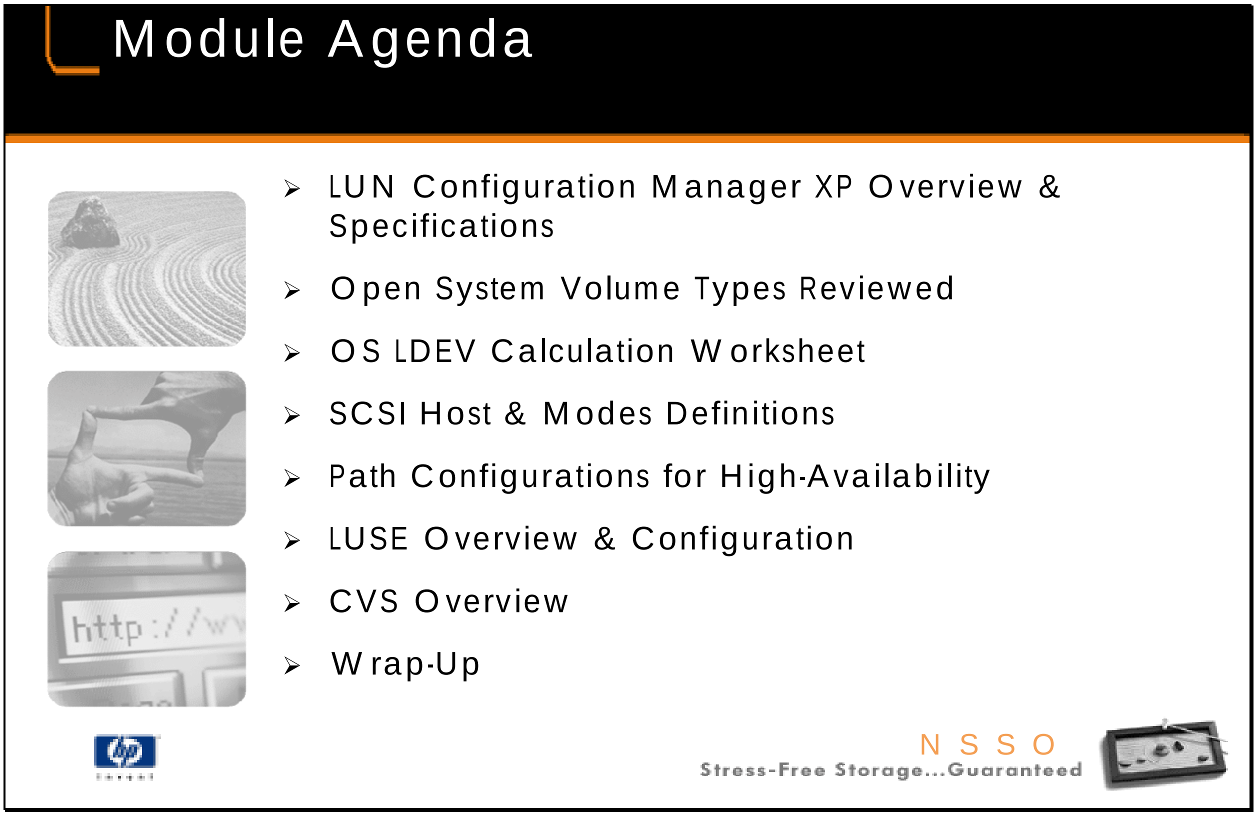

XP48 Performance Overview
Front-end Performance:
100% Cache Hits
• 80,000 IO/sec
• 840 MB/sec
Back-end Performance:
Cache Avoidance
• 4,500 IO/sec
• 318 MB/sec
CHIP Pair Limits:
• 41,000 IO/sec
• 464 MB/sec
ACP Pair Limits:
• 7,300 IO/sec
• 318 MB/sec
CUDA
Single FC PORT:
Standard Mode:
• 10,500 IO/sec
• 90 MB/sec
High Performance Mode:
• 20,000 IO/sec
Crossbar Bandwidth: 3.2 GB/sec
1 Array Group:
• 450 IO/sec
• 65 MB/sec
Hypothetical Configuration: RAID 0/1, 11 RAID Groups, 73GB Disks, 16 GB Cache, 1 ACPP, 3 CHPP, 24 FC PORTS, Random Performance: reads, 2k block size; Sequential Performance: reads, 64k block size. Note: Single FC-port and CHPP-Li s are 100% cache hit. ACPP-Limit and 1-Array Group tests are cache avoidance. Based on actual XP512 tested numbers.
NSSO
Stress-Free Storage...Guaranteed

Module Wrap-up

N S S O
Stress-Free Storage...Guaranteed



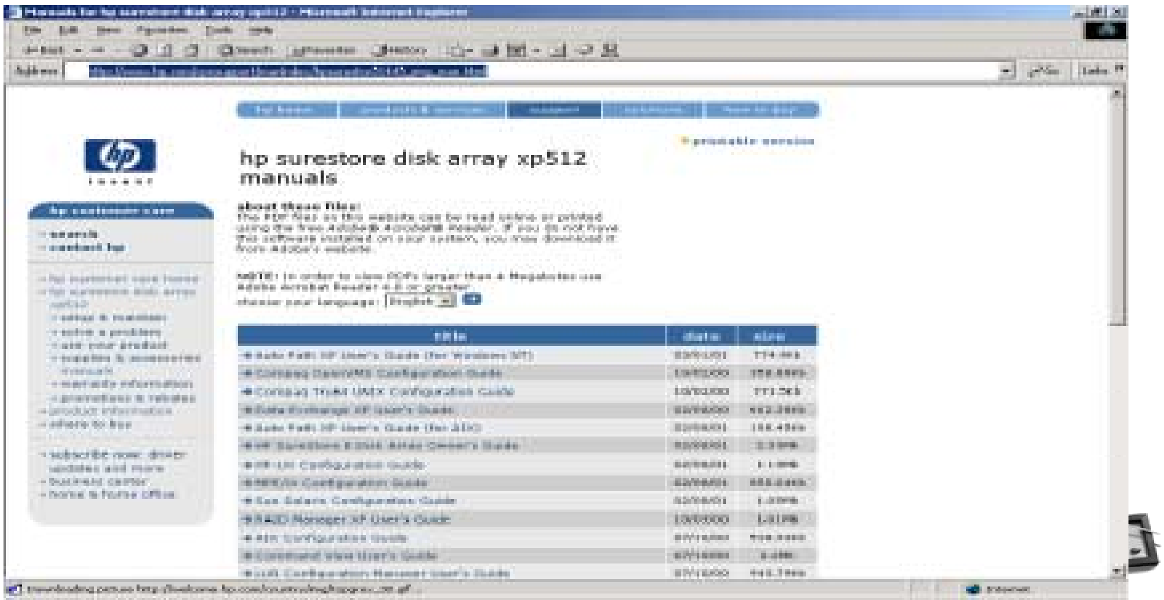
LTU in the base product (B9335A) is defined as License to Use for the software. Base product provides media, documentation, and license to use on HP SureStore E Disk Array XP 512. Additional licenses must be purchased in 1, 5 and 10 TB tiers to enable product. The base product consists of software keys for the various software functions for the SVP and Remote Console PC, along with a Software Certificate, SWUID and product manual (order number B7911-90900). On the SVP, keys are needed for LUSE and CVS only. On the Remote Console, keys are needed for all four functions.



HP SureStore E LUN Configuration Manager XP is a storage management feature of the HP SureStore E Disk Array XP Family of arrays. It consists of three software functions: LUN Size Expansion, LUN Manager and Custom Volume Size. HP SureStore E LUN Configuration Manager XP software can be run on the Remote Control PC or the SVP (Service Processor). It’s features are activated for use on the Remote Control PC with software keys that are installed on the SVP. The SVP is designed for use by authorized HP service and support personnel and is password protected from end-users.

XP Array Manuals

http://www.hp.com/cposupport/manindex/hpsurestor21443_eng_man.html



LUN Configuration Manager XP

- LUN Configuration Manager XP B9335A and the built-in Logical Unit Size Extension (LUSE) and Custom Volume Sizing (CVS) run on the Remote Console. It is the primary point for configuring the XP Array by a customer. CE does it from the SVP.
- Maps the logical disk devices (internal LDEVs) attached to the ACP pairs to logical unit numbers (external TID:LUNs) on the CHIPs for use by SAN hosts.

NSSO
Stress-Free Storage...Guaranteed

The LUN Configuration Manager XP feature enables you to define the SCSI-to-LUN paths for the disk array using the remote console PC. Each LUN can be mapped for access from multiple ports and/or multiple target IDs to provide alternate paths for continuous data availability. LUN Configuration ManagerXP allows you to reconfigure the SCSI/FC paths at any time to accommodate system configuration changes (for example, adding a host) and optimize disk array performance (for example, relieving bottlenecks of activity).

To maximize data access capabilities, LUN Configuration Manager XP allows you to create virtual logical units which are larger than standard LUNs, giving hosts access to the data on the entire disk array using fewer logical units than before. The disk array supports both SCSI and FC ports. Each SCSI port provides fifteen SCSI target IDs (TIDs) and eight lgical units (LUNs) per SCSI TID for a maximum of 120 LUNs per SCSI port to support a wide range of configurations and requirements. LUN Configuration Manager XP and the built-in Logical Unit Size Extension (LUSE) are default features of the RMCMAIN software set. No additional software key is required to activate this program. LUN Configuration Manager XP is the primary point of configuration for the XP Arrays from the Remote Console. The application maps logical disk devices (LDEVs) attached to the ACP pairs to logical unit numbers (LUNs) on the CHIPs for use by SAN hosts. LUNs are created from LDEVs. LUN to LDEV mapping scheme allows setting multiple paths to the same logical volume.

Shared volumes can enable a High Availability configurations. Other functionalities include configuration of port modes and creation of extended LDEVs.

Array Management

LUN Configuration Manager XP	Allows the following to be performed from the Remote Console PC • **LUN Manager**: assign ports to volumes (creating a LUN) and define host to port connectivity settings. • **LUSE**: combine up to 36 volumes to create one large LUN • **CVS**: create multiple volumes of smaller size than standard open emulation volumes; also allows volume creation of free space not used during LDEV creation

N S S O
Stress-Free Storage...Guaranteed

LUN Configuration Manager XP

- LDEVs are created from physical disks.
- An array group is divided into Open Volumes, or LDEVs of type OPEN-3, OPEN-8 and OPEN-9 and OPEN-E.
- These 4 emulation types are derived from mainframes, and have a certain size depending on their type.
- This size is not customizable, unless you use LUSE or CVS.

NSSO
Stress-Free Storage...Guaranteed

LUN Manager can be used to define, add and delete Disk Array XP 256 subsystems, LUNs, and SCSI paths to an open-system host. Its principle function is to assign a SCSI path to one or more array (RAID) groups.
Configure expanded-size LUNs to access data with fewer LUNs
Configure custom-sized LUNs smaller than standard LUN sizes
Define or reconfigure SCSI port to LUN mapping
Set each port to to the correct mode for the connected platform
Control Units (CU) manage the process of creating LDEVs out of array groups.
A CU is an internal table holding the backend configuration, the mapping of LDEVs to array groups.
Each CU supports 1 RAID level and 1 emulation type
From 0 to 3 CUs in one XP256, numbered as:

<CU Number>:<Logical DEVice Number>

0-16 Cus in a XP512 host. LUNs are created from LDEVs. LUN to LDEV mapping scheme allows setting multiple paths to the same logical volume. Shared volumes can enable a High Availability configuration.

LUN Configuration Manager XP Rules

- Two parts of LUN Configuration Manager XP:

  1) resident in the SVP

     - Used only by Certified CE to define LDEVs in an Array Group, assign Control Units (CU), & other low level tasks.

  2) resident in the Command View Console

     - Used by Administrator and ASEs
     - Assign SCSI or fibre Path definition for LDEV or LUN
     - Configure/remove expanded LUNs (LUSE)

NSSO
Stress-Free Storage...Guaranteed

LUN Configuration Manager XP exists in two parts: the SVP and Command View Console versions.

From the SVP, the CE can configure LDEVs - set emulation type, assign and configure control units (CUs), assign RAID levels - and other initial setup tasks. Once these are completed, all other configurations should be done from the RMC. Remember - only the CE should access the SVP directly. Once the LDEVs are created, they will be managed from the RMC's version of LUN Configuration Manager XP. From the RMC, we can utilize SCSI path definition on the logical devices. SCSI path definition is the function that defines the PORT, SCSI ID, and LUN required to connect SCSI host server or main frame host processor to our XP256. A way to remember the difference between the two functionalities:

From the SVP, LUN Configuration Manager XP configures physical devices (disks are broken down into LDEVs with specific emulation types).

From the Command View Console, LUN Configuration Manager XP configures logical devices (LDEVs)

LUN Configuration Manager XP

- Configure expanded-size LUNs to access data with fewer LUNs
- Configure custom-sized LUNs smaller than standard LUN sizes
- Define or reconfigure SCSI port to LUN mapping
- Set each port to to the correct mode for the connected platform

NSSO
Stress-Free Storage...Guaranteed

The LUN Size Expansion (LUSE) feature allows the creation of virtual LUNs that are larger than standard OPEN-3, OPEN-8, OPEN-9 and OPEN-E LUNs. Multiple individual LUNs are consolidated into a virtual LUN. With LUN Size Expansion the size of a selected LUN can be expanded to 259 GB utilizing OPEN-9 LDEVs. This capability allows open-system hosts to access data on the entire Disk Array XP arrays using fewer logical units than before expansion. The expanded LUN may consist of a mixture of different array (RAID) groups. However, the LDEVs in the array group must be contiguous in the Disk Array XP Array s CU (Control Unit) table in order to combine them into an expanded LUN. Open emulation types cannot be mixed in an expanded LUN. A minimum of two and a maximum of 36 LDEVs are required to form a LUN expansion. So one thing we'd like to point out here too is that we'd recommend using the LUSE tool to coalesce your LUNS so that you don't have quite as many, which will help with I/O scan time. If you have too many devices it will take a very long time. consider a fully loaded XP or an XP512 anyway, 128 array groups let's say. RAID 5, that's 29 LUNS per. That's about 3700 LUNS. And then you've got two device files for each of them because of the primary and the ultimate path. So that's 7400 device files that you might have to deal with. That's outrageous.

Custom Volume Size (CVS), also knows as Virtual LVI, is an optional feature of the Disk Array XP family that allows multiple virtual volumes to be defined within a RAID group. In effect, it divides one LDEV into several volumes. CVS provides a dynamic means to optimize capacity, increase productive throughput, and reduce operational overhead expenses. It consists of a function to provide variable capacity volumes and a function to provide free arrangement of logical (virtual) volumes on the physical disk drives within a RAID group. Up to eight disk arrays can be managed from a single interface. CVS allows user-defined custom-size volumes that are smaller than normal volumes to be configured. It improves data access performance by reducing device contention as well as host I/O queue times, which can occur when several frequently accessed files are located on the same physical volume. Each of these files can be stored in separate logical volumes defined with CVS. A maximum of 32 normal and CVS volumes can be defined per RAID group. CVS enables better use of the physical storage capacity of the Disk Array XP array family while reducing the amount of administrative effort required to balance direct-access storage device (DASD) I/O workloads. A data file with a high I/O frequency can be easily allocated to an independent volume thus decreasing the contention and performance bottleneck while increasing the subsystem s overall throughput and the host s performance. It can be used in conjunction with HP SureStore E Cache LUN XP to achieve better performance than when either of these options is used individually.

Control Units (CU)

- Control Units (CU) manage the process of creating LDEVs out of array groups.
- A CU is an internal table holding the backend configuration, the mapping of LDEVs to array groups.
- Each CU supports 1 RAID level and 1 emulation type
- From 0 to 3 CUs in one XP256 and 0 to15 CUs in one XP512, numbered as:

  <CU Number>:<Logical DEVice Number>

NSSO
Stress-Free Storage...Guaranteed

The expanded LUSE LUN may consist of a mixture of different array (RAID) groups. However, the LDEVs in the array group must be contiguous in the Disk Array XP array's CU (Control Unit) table in order to combine them into an expanded LUN. Open emulation types cannot be mixed in an expanded LUN. A minimum of two and a maximum of 36 LDEVs are required to form a LUN expansion.

Configuration Guidelines

- Disks are always purchased in bundles of 4 (a marketing RAID group)
- A RAID 5 array group always consists of 4 disks - RAID 5 = 3D+1P (3 data, 1 parity)
- A RAID 1 array group always consists of 2 disks - RAID 1 = 1D+1M (1 data, 1 mirror)
- A RAID 0/1 array group always consists of 4 disks – RAID 1/0 = 2D +2M (2 data, 2 mirror)*

XP512 & XP48 Only*

NSSO
Stress-Free Storage...Guaranteed

The XP512 will support RAID 0/1 and 5 at first release. A "marketing array group" is the four-disk bundle that HP sells for XP array expansion. In implementation, the definition of an array group varies depending on the RAID level. RAID 5, implemented as 3D+1P, uses four-disk array groups while RAID 1, implemented as 1D+1M, only uses a two-disk array group. RAID levels are set by the CE at disk installation and should not be changed on active array groups. The array does not use the entire disk all at once. The subsystem requires that the disk space be broken down into logical devices of four emulation types, Open-3 (2.4 GB), Open-9 (7.42 GB), Open-8 (7.347 GB) and Open-E (OPEN-E 13.58 GB). These "open volume" emulation types are based on mainframe disk types, but have been adapted for use with LDEVs on the XP256. LDEV 0xFF of CU 3 cannot be set (3:FF)
This is a special device reserved silently by Hitachi Open Remote Copy (HORC)
Test: On the remote console:
1) create a VSC volume of 36 MB to be reserved as a command device
2) try to assign this VSC to address 3:FF
message “device already reserved” will appear

OS LDEV Calculation

47.190 GByte Disc (1GByte=1000^3 Byte) **Version: 21.10.2000 21:00**

OPEN-K	1.87	GByte =	1875	MByte
OPEN-3	2.46	GByte =	2461	MByte
OPEN-8	7.35	GByte =	7348	MByte
OPEN-9	7.38	GByte =	7385	MByte
OPEN-E	14.58	GByte =	14582	MByte

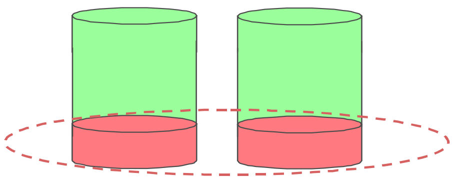
Informations in the Spreadsheet should not be used without technical guidance!

		RAID-0/1		(2D + 2M)								RAID-5		(3D + 1P)							
		OPEN-K		OPEN-3		OPEN-8		OPEN-9		OPEN-E		OPEN-K		OPEN-3		OPEN-8		OPEN-9		OPEN-E	
LDEVs / Discbundle		50		38		12		12		6		75		57		19		19		9	
Numbe Discbu	Raw Space	Usable Space	Neede LDEVs	Usable Space	Neede LDEVs	Usable Space	Neede LDEVs	Usable Space	Neede LDEVs	Usable Space	Neede LDEVs	Usable Space	Neede LDEVs	Usable Space	Neede LDEVs	Usable Space	Neede LDEVs	Usable Space	Neede LDEVs	Usable Space	Needed LDEVs
1	189	94	50	94	38	88	12	89	12	87	6	141	75	140	57	140	19	140	19	131	9
2	378	187	100	187	76	176	24	177	24	175	12	281	150	281	114	279	38	281	38	262	18
3	566	281	150	281	114	265	36	266	36	262	18	422	225	421	171	419	57	421	57	394	27
4	755	375	200	374	152	353	48	354	48	350	24	562	300	561	228	558	76	561	76	525	36
5	944	469	250	468	190	441	60	443	60	437	30	703	375	701	285	698	95	702	95	656	45
6	1133	562	300	561	228	529	72	532	72	525	36	844	450	842	342	838	114	842	114	787	54

NSSO
Stress-Free Storage...Guaranteed

Please refer to your Appendix to locate the Excel spreadsheets for calculating usable capacity.

Configuring Host Access to LDEVs



LDEV Number

SCSI ID / LUN
(SCSI-ID 0-6,8-15) (0-7)

Fibre Channel Port / LUN
(0-119)
(0-255)XP512

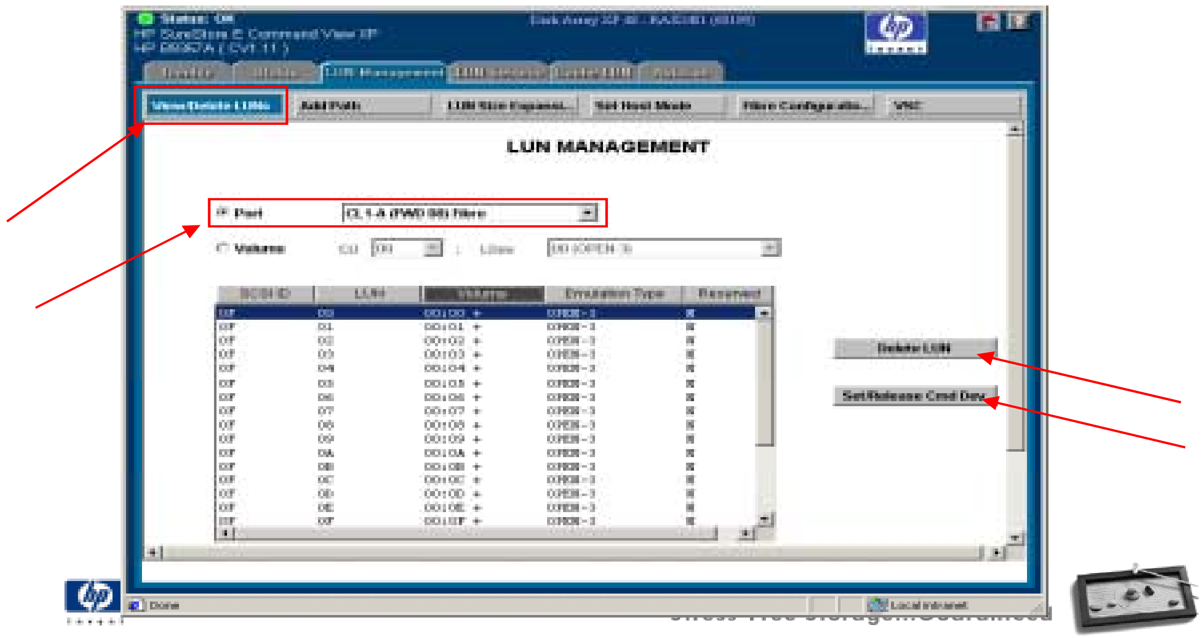
- Current firmware supports 15 SCSI addresses and 8 LUNS per SCSI port and emulation of 15 SCSI addresses and 8 LUNS per FC port (120 SCSI style addresses).
- LDEVs 00 to FF for each CU (0,1,2,3) can be used.

N S S O
Stress-Free Storage...Guaranteed

Note that emulation mode makes the output look like the SCSI addresses. Once a emulation size is chosen for an array group, that emulation must be used for the entire array group. Make sure that the appropriate drivers are loaded for the FC adapter card. SAN connections do not use the traditional FC drivers - there are special drivers configured specifically for mass storage connections. To check if the drivers are installed, run one of these command lines:
<u>10.20</u>: swlist -l product | grep -i FCMS
<u>11.00</u>: swlist -l product | grep -i FCMassStorage
Also be sure to check with HP Support Center for the most up-to-date patches for these drivers.
HP-UX is “SCSI-based”
FC LUNs have hardware addresses similar to those used by SCSI devices - under “Emulation Mode” individual FC ports are given SCSI IDs and LUN numbers
When assigning host access, LDEVs 00 to FF for each selected CU (0,1,2,3) can be used except if Continuous Access XP (CA) is used. If CA is being used, one LDEV is needed for the Command Device (discussed later in the CA module). Current firmware supports 15 SCSI addresses and 8 LUNS per SCSI port (120 addresses per port) and either 126 FC IDs or emulation of 15 SCSI address and 8 LUNS per FC port (120 SCSI style addresses). The maximum number of paths per XP256 is

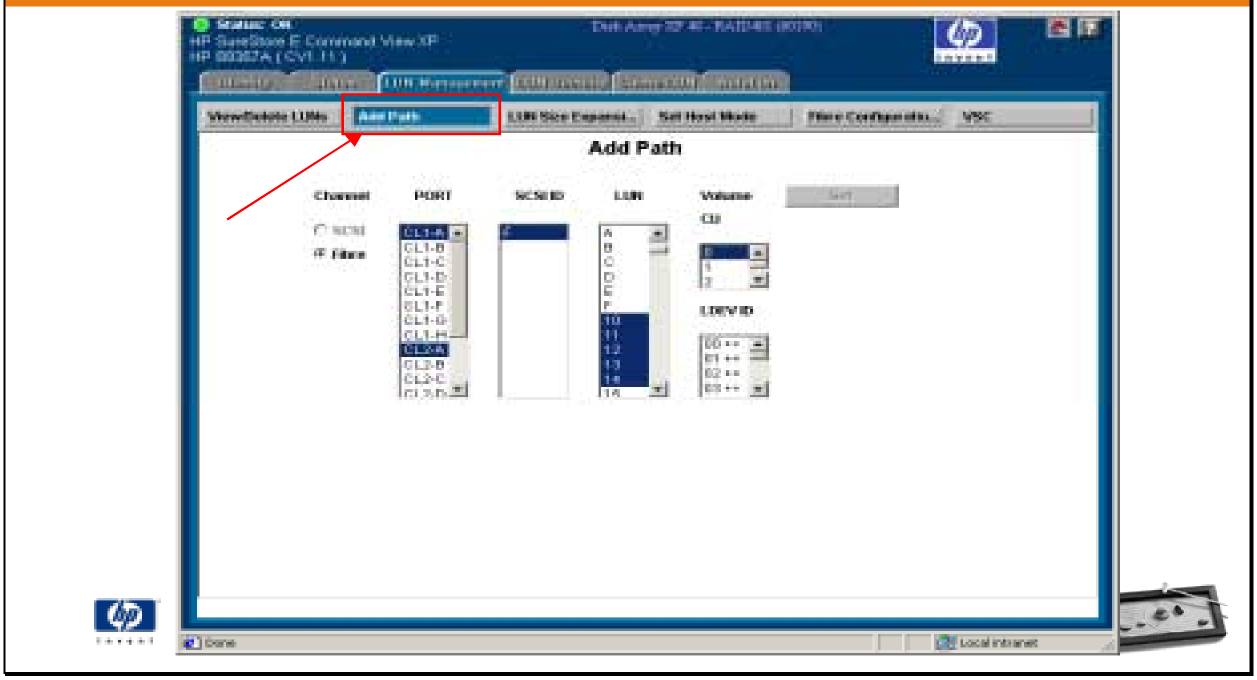
currently 2048. All values are subject to increase in later revisions of the firmware. We will discuss addressing strategies in the next few slides.

LUN Manager: Command View XP



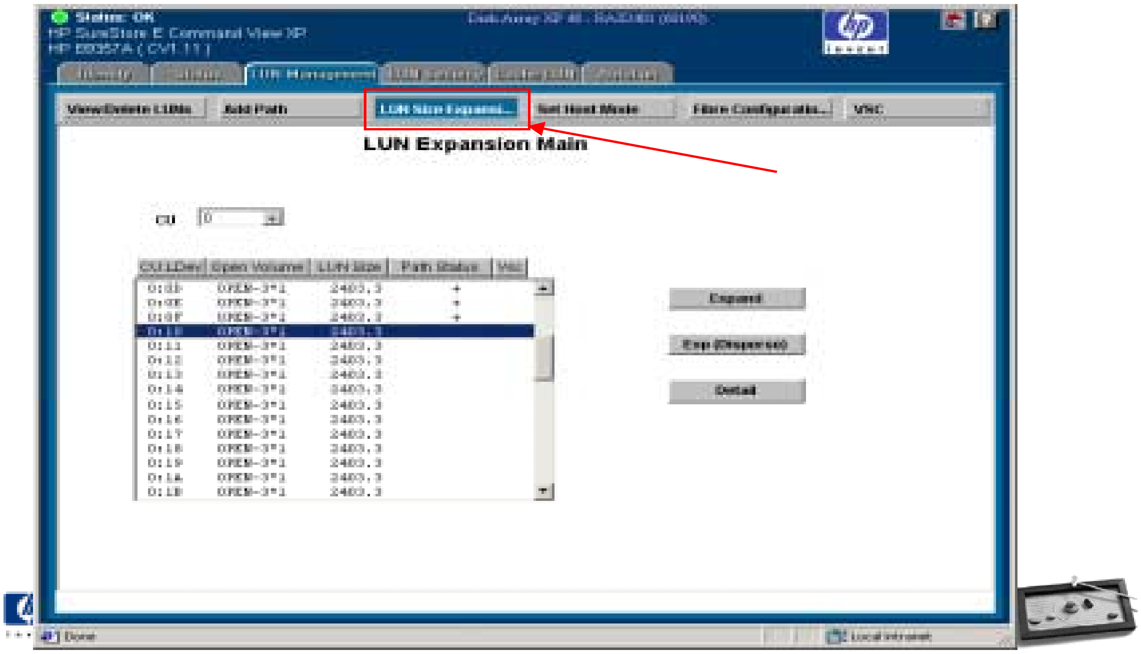
View/Delete LUN's – Click this button to view and/or delete LUN's. View LUN's by port or volume.

LUN Manager: Command View XP

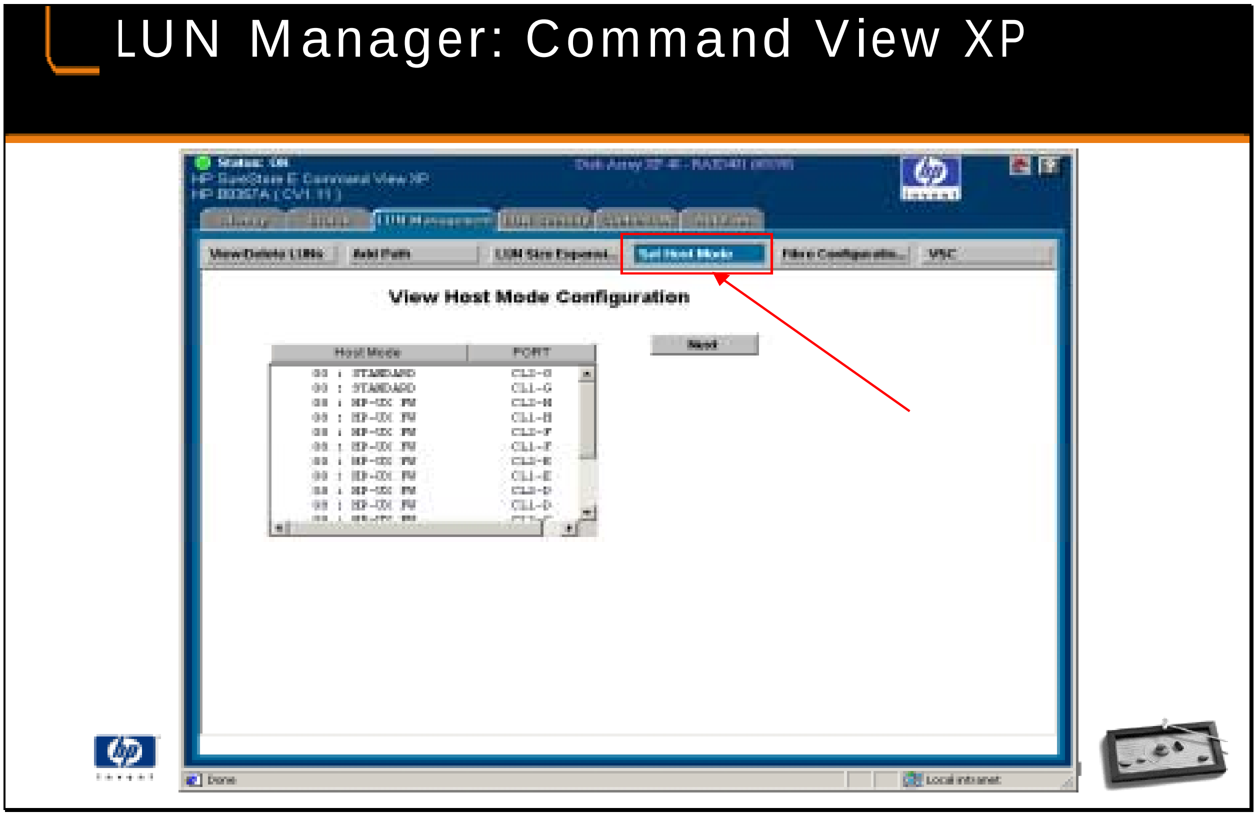


Add Path – Click this button to perform Volume Mapping

LUN Manager: Command View XP

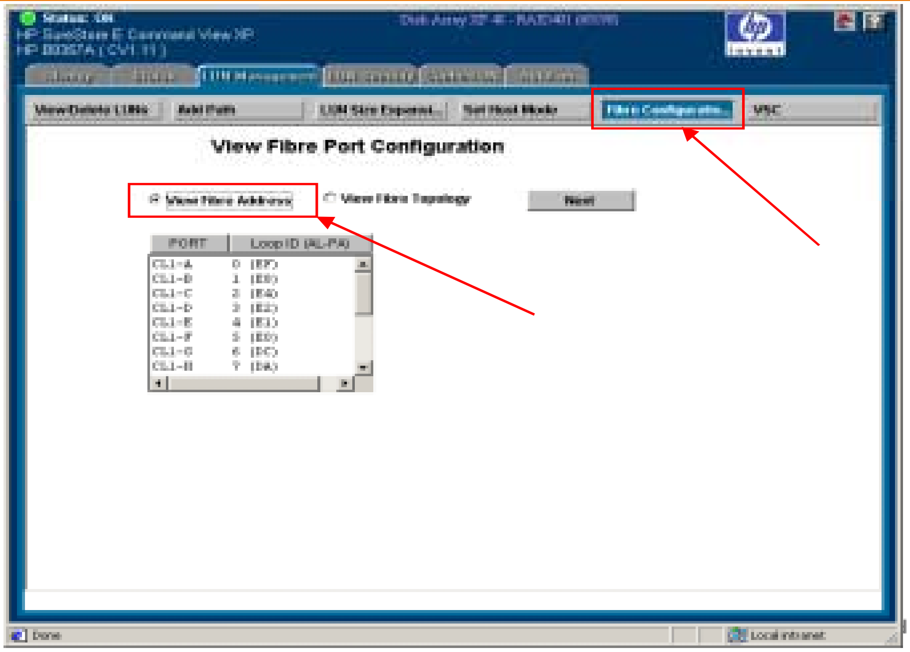


LUN Size Expansion (LUSE) – Click this button to create LUSE volumes.

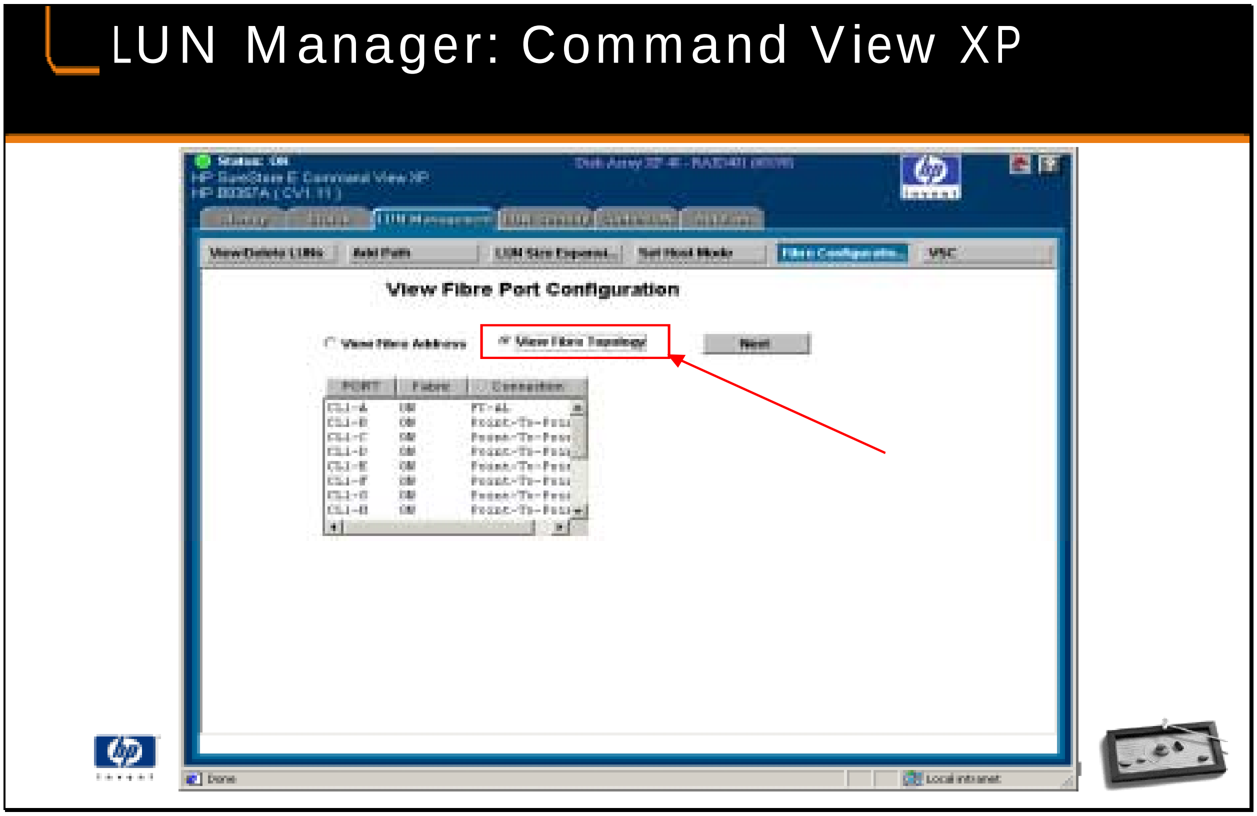


Set Host Mode – Click this button to set the host mode for a particular port on a CHIP board.

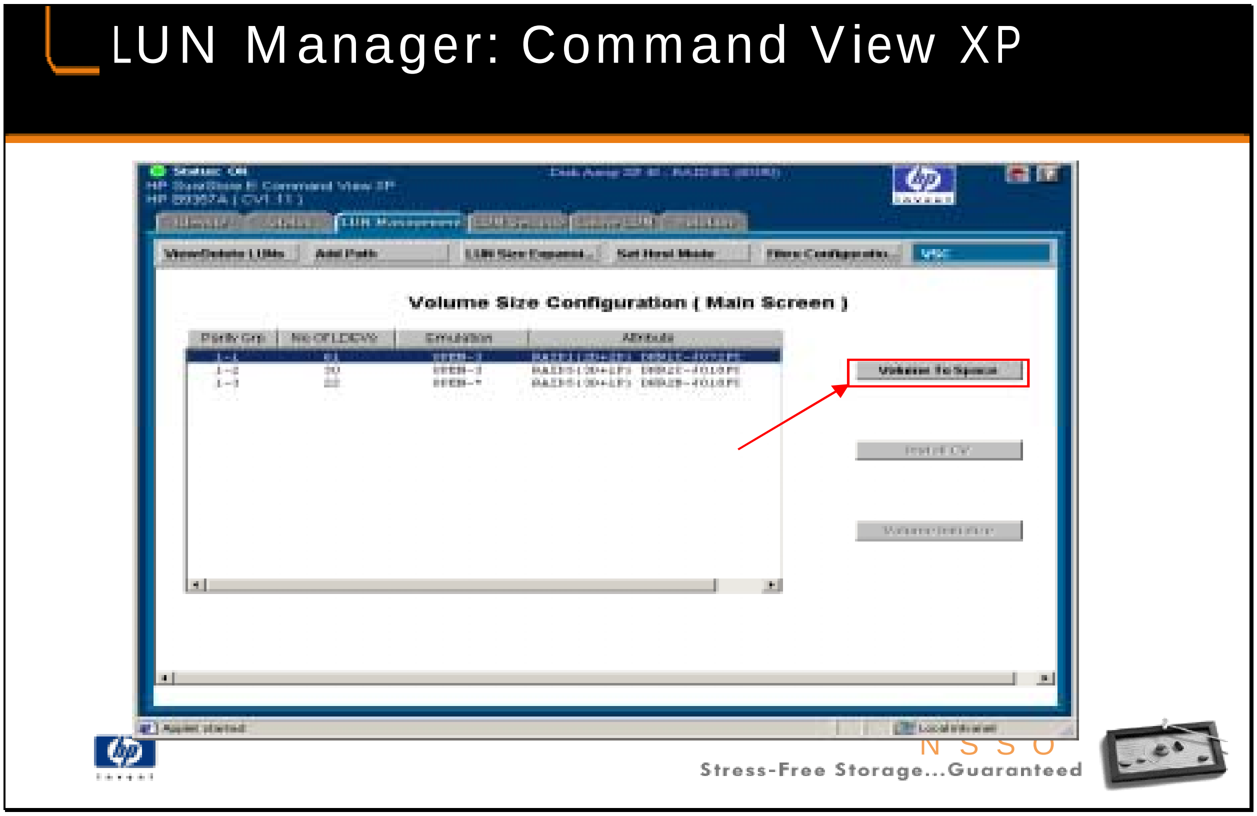
LUN Manager: Command View XP



Fibre Configuration – Click this button to set the Port Fibre Address and the Port Fibre Topology.



Fibre Configuration – Click this button to set the Port Fibre Topology.



VSC (Volume Size Configuration) – Click this button to configure CVS volumes.

Host Modes on the XP

Mode		Mode	(modes 10-1C not available on XP512/48)
00	Standard	10	Standard and Ultra SCSI Mode
01	IBM-7135 Host Mode	11	IBM-7135 Host and Ultra SCSI Mode
02	NCR Host Mode	12	NCR Host and Ultra SCSI Mode
04	Sequent Dynix Ptx Host Mode	14	Sequent Host and Ultra SCSI Mode
05	DEC OpenVMS Mode	15	DEC OpenVMS and Ultra SCSI Mode
07	Compaq Tru64 Mode (Ucode > =12-18 required)		
08	HP-UX Mode and FW Differential/FibreChannel	18	HP-UX and Ultra SCSI Mode
09	VXVM-DMP Mode (Veritas)	19	VXVM-DMP and Ultra SCSI Mode (Veritas)
0A	NetWare Mode	1A	NetWare and Ultra SCSI Mode
0C	MSCS Multipath Mode	1C	MSCS Multipath and Ultra SCSI Mode
0F	IBM AIX HACMP Mode (using FC6227 HBA)		

Notes:

- FW and Ultra SCSI is not available as an option on XP512/48. This means that host modes 10-1C (Ultra SCSI) do not apply to the XP512/48.
- Modes 00/10 are the default used for NT/Win2K (non-HP FC HBAs and all NT SCSI), Solaris, AIX, Netware, SGI, DEC, and Linux with non-HP HBAs. Mode 08 applies when the D8602 HBA is used with NT/Win2K.
- Mode 08 is used for HP-UX and for SCSI on MPE/iX. Also used for NT/Win2K FC with HP D8602 and Compaq 120186-B21 HBAs.
- Mode 08 has the special HP modifications that allow a Queue Depth of 1024 and 120 LUNS per port
- Mode 0C is used for Microsoft Wolfpack or MSCS. Use 08 if HP FC HBAs
- Standard and Ultra SCSI differ by HEX 10. Notice the parallel in the numbers. Adding “10” enables the “Ultra” mode.
- Enabling Veritas host mode also requires enabling “system mode 19” via the SVP in addition to setting the host mode for the port. This activity is accomplished with TAC assistance and is no longer required as of 43-12 (XP256) microcode and is not required on any version of the XP512/48.

N S S O
Stress-Free Storage...Guaranteed

Host Modes on the XP
Rev 2001-03-18 MWR

Changes since last version noted in blue.

The latest version of this document is always available on the NA ESSO Storage Web site at: http://sseweb.nsr.hp.com/xp256.

The following is a current list of XP Host modes. This supplements the list found in the Remote Control Manual and was assembled using information from service notes, SFC, the Hitachi TAC and other support sources..

Mode

Mode
(modes 10-1C not available on XP512/48)

00

Standard

10

Standard and Ultra SCSI Mode

01

IBM-7135 Host Mode

11

IBM-7135 Host and Ultra SCSI Mode

02

NCR Host Mode

12

NCR Host and Ultra SCSI Mode

04

Sequent Dynix Ptx Host Mode

14

Sequent Host and Ultra SCSI Mode

05

DEC OpenVMS Mode

15

DEC OpenVMS and Ultra SCSI Mode

07

Compaq Tru64 Mode (Ucode > =12-18 required)

08

HP-UX Mode and FW Differential/FibreChannel

18

HP-UX and Ultra SCSI Mode

09

VXVM-DMP Mode (Veritas)

19

VXVM-DMP and Ultra SCSI Mode (Veritas)

0A

NetWare Mode

1A

NetWare and Ultra SCSI Mode

0C

MSCS Multipath Mode

1C

MSCS Multipath and Ultra SCSI Mode

0F

IBM AIX HACMP Mode (using FC6227 HBA)

Notes:

- FW and Ultra SCSI is not available as an option on XP512/48. This means that host modes 10-1C (Ultra SCSI) do not apply to the XP512/48.
- Modes 00/10 are the default used for NT/Win2K (non-HP FC HBAs and all NT SCSI), Solaris, AIX, Netware, SGI, DEC, and Linux with non-HP HBAs. Mode 08 applies when the D8602 HBA is used with NT/Win2K.
- Mode 08 is used for HP-UX and for SCSI on MPE/iX. Also used for NT/Win2K FC with HP D8602 and Compaq 120186-B21 HBAs.
- Mode 08 has the special HP modifications that allow a Queue Depth of 1024 and 120 LUNS per port
- Mode 0C is used for Microsoft Wolfpack or MSCS. Use 08 if HP FC HBAs
- Standard and Ultra SCSI differ by HEX 10. Notice the parallel in the numbers. Adding “10” enables the “Ultra” mode.
- Enabling Veritas host mode also requires enabling “system mode 19” via the SVP in addition to setting the host mode for the port. This activity is accomplished with TAC assistance and is no longer required as of 43-12 (XP256) microcode and is not required on any version of the XP512/48.

Changes to LUN Mapping

- LUN Configuration Manager can be used to change LUN mapping.
- If there is a port for which no changes to LUN mapping are being made, I/O operation is not affected by changes to another port.
- This allows 100% availability during firmware upgrades.
  - Move all active links to a port
  - Upgrade the firmware on the inactive port
  - Move all activity to the upgraded port
  - Upgrade the remaining port
  - Move half the active links back to that port for original config.
- Have redundant links to the array from all hosts so we can down a port without stopping the host, or let customer know which ports and hosts have constraints

NSSO
Stress-Free Storage...Guaranteed

FC Addressing Restrictions

Assigning Native Fibre Channel paths to SCSI addresses needs your calculation, for example, for LDEV as (TID:LUN) 8:0 select LUN 40 (hex).

LUN	Emulated Host SCSI Target	Emulated Host SCSI LUN
0-7	0	0-7
8-F	1	0-7
10-17	2	0-7
18-1F	3	0-7
20-27	4	0-7
28-2F	5	0-7
30-37	6	0-7
38-3F	7	0-7
40-47	8	0-7
40-4F	9	0-7
50-57	10	0-7
58-5F	11	0-7
60-67	12	0-7
68-6F	13	0-7
70-77	14	0-7
78-7F	15	0-7

N S S O
Stress-Free Storage...Guaranteed

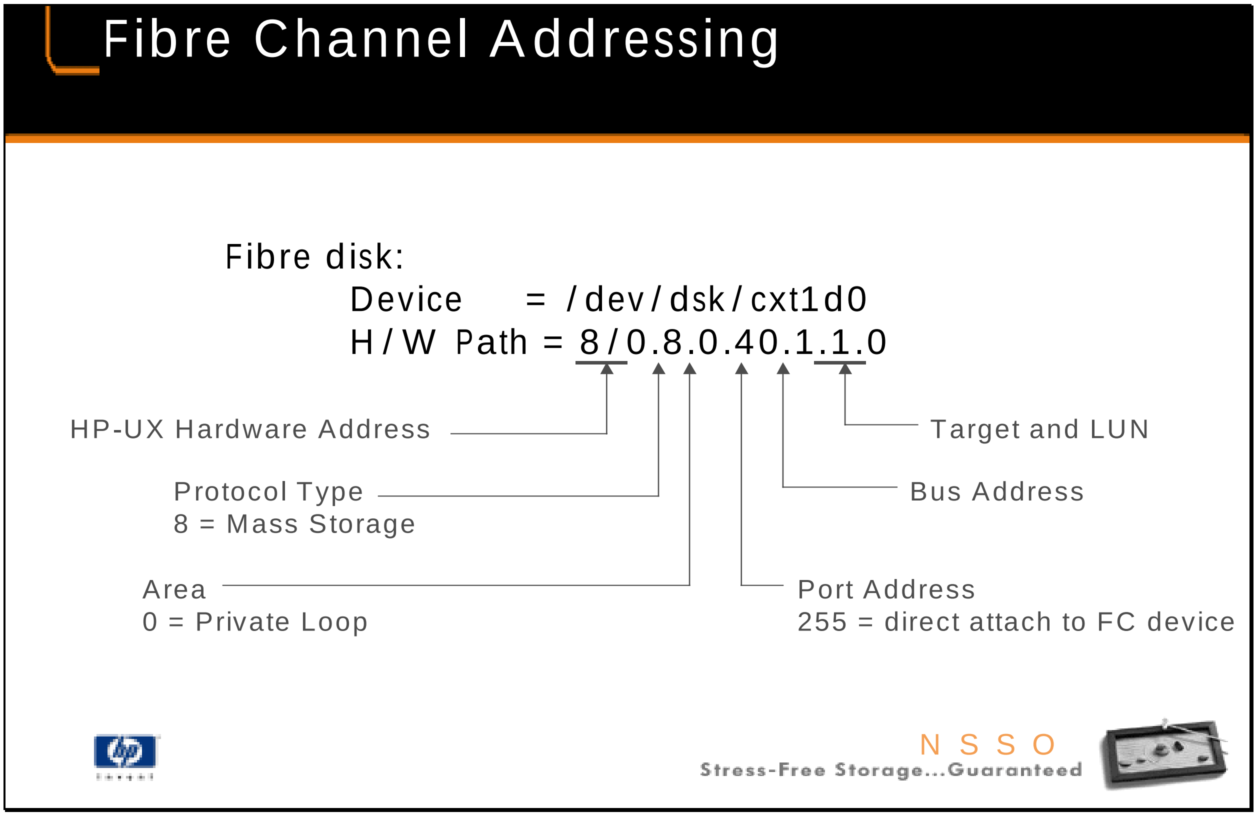
<u>FC Addressing Restrictions</u>:

Assigning Native Fibre Channel paths on the XP Family of Arrays is not always straightforward. In the XP 256 (FC-AL Mode), the FC is emulated as F/W SCSI on the HP9000 host (much like the EMC Symmetrix), however the resulting SCSI addresses must be calculated when entering them into the path screen on the SVP or Remote Control.

Set the Port Mode to “08” for all HP hosts (SCSI and FC). Note that you will only see SCSI ID “F” and 128 LUNs (00-7F Hex) from the Path "Add" screen (shown in the slide above). The LUN number will be translated by the host to SCSI target IDs and LUN numbers. You need to calculate your desired outcome before selecting your LUNs. The mapping essentially works like the table. For example, if you want your LDEV to appear as SCSI address 8, LUN 0, select LUN 40 (Hex). The basic rules that applied to the EMC also apply here:

Each SCSI target must have a LUN 0, or none of the LUNs on that target will be visible to the host (just like with SCSI addressing).

Do not break a string of LUNs or the subsequent LUNs will not be visible to the host (Ex: If LUNs 30, 31, and 36 are assigned, the host will only see LUNs 30and 31).



It is easy to deal with a XP256 in a SCSI environment. However, with the advent of Storage Area Network (SAN) environments, trying to find where disks are located becomes more complex. The host that you are connected to may be able to see several XP256s via a SAN. The only way to find where you are connected is to physically map the environment and then apply that mapping to the hardware path that is returned by the ioscan -fnC disk command. The slide shows how fibre channel addressing is built up.

With this information and your SAN map, you be able to locate where any disks are located within the SAN.

High Availability Configuration Recommendations

- Configure multiple paths to the same LDEV through different ports to maximize fault tolerance.
- It is useful to assign LDEVs in a pattern that is easy to recognize from the host in the ioscan output. Here is the beginning of a sample pattern:

LDEV	Primary Disk	Alternate Link
0:00	c0t0d0	c1t8d0
0:01	c0t0d1	c1t8d1
0:02	c0t0d2	c1t8d2
0:03	c0t0d3	c1t8d3

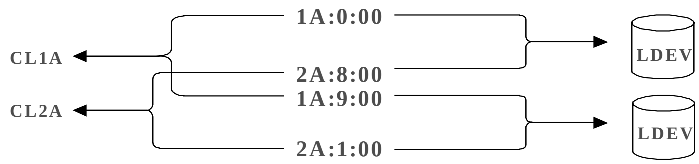
Primary Disk = c<instance_number>t<target_id>d<LUN>
Alternate Link = c<lookup_table>t<target_id+8>d<LUN>

NSSO
Stress-Free Storage...Guaranteed

A suggested configuration from "stress-free central" recommends configuring multiple paths to the same LDEV through different ports to maximize fault tolerance. Although not mandatory, it is often useful to assign your LDEVs in a pattern that is easy to recognize from the host. A common practice is to use certain target addresses for primary and others for alternate links. When the system administrator views his “ioscan” output, it will be easy to determine which volume were created with the intention of being alternate links, and which were intended as primary. This facilitates load balancing across FC (or SCSI) host adapters when creating volume groups. This process can also be used to identify volumes intended for other purposes such as Business Copy (Ex: Use target 3 for primary, and 11 for alternate links).

High Availability Configuration Recommendations

Sample configuration:



- Primary paths (PORT:TID:LUN) 1A:0:00 and 2A:1:00 to the two LDEVs are on different CHIP ports to distribute I/O loading.
- The two alternate links 2A:8:00 and 1A:9:00 are also on different CHIP ports, so that in a disaster situation, both LDEVs can be accessed via a single port.

N S S O
Stress-Free Storage...Guaranteed

The diagram above is a sample configuration. The primary paths (1A:0:00 and 2A:1:00) to the two LDEVs are on different CHIP ports to distribute I/O loading. However, the two alternate links (2A:8:00 and 1A:9:00) are also cross-connected so that in an emergency situation, both LDEVs can be accessed via a single port.

Logical Unit Size Extension (LUSE)

Logical Unit Size Extension (LUSE), aka LU Expansion, is a built-in feature of LUN Configuration Manager XP.

LUSE allows you to create virtual LUNs larger than standard OPEN-E/M/L/8/9/3 LDEVs.

- Use LUSE to combine multiple LDEVs into one.
- Use LUSE for situations where you need larger LUNs.
- Use LUSE for situations where you need fewer LUNs.

N S S O
Stress-Free Storage...Guaranteed

Logical Unit Size Extension (LUSE) can be used to create larger LUNs if the open volumes are not large enough. This is primarily for systems that do not support LVM (like NT).

LUSE cannot be used to extend the maximum number of LDEVs supported by the system. The LDEV ID numbers are not freed when a LUSE volume is created. These numbers are retained by the system to be used if the extended volume is split apart again.

WARNING: The LUSE feature is a destructive process - it will render data on an LDEV unrecoverable. The LUNs connected to make an expanded LUN are erased. You must make sure to back up the data on the LDEV before starting this process if you need the information contained on the device. Should you use LUSE? Do you want ease of management or highest performance? Using LUSE allows the XP Family of Arrays to do most of the calculations about where resources should be allocated. Once the LUSE is configured, the storage area can be used easily as nothing more than a very large physical volume. However, the sequential nature of LUSE use of space on the drives can be a major performance bottleneck, especially with really large LUSE volumes. In this case, LVM striping is a much better alternative.

LDEVs combined must be all OPEN-8, all OPEN-9 or all OPEN-3 (I.e. same emulation type).

LUNs across different array groups may be combined.

Up to 36 LDEVs can be combined into one larger LUN.

LUNs must be on the same control unit.

LDEVs you connect can be noncontiguous LUNs.

LUNs that are connected for an expanded LUN are erased!

LUNs with paths assigned cannot be expanded. Delete assigned SCSI paths before expanding.

The extended logical unit will assume the LDEV number of the lowest original LDEV number assigned to the extended unit. As long as the LDEVs are contiguous in the CU table setup by the CE when the system was configured, they can be used in a LUSE volume. We can combine up to 36 LDEVs with LUSE.

Minimum of 2 to maximum of 36 LDEVs required to form a LU size extension

LU capacity:

OPEN-8: Min 14.6 GB Max 264 GB

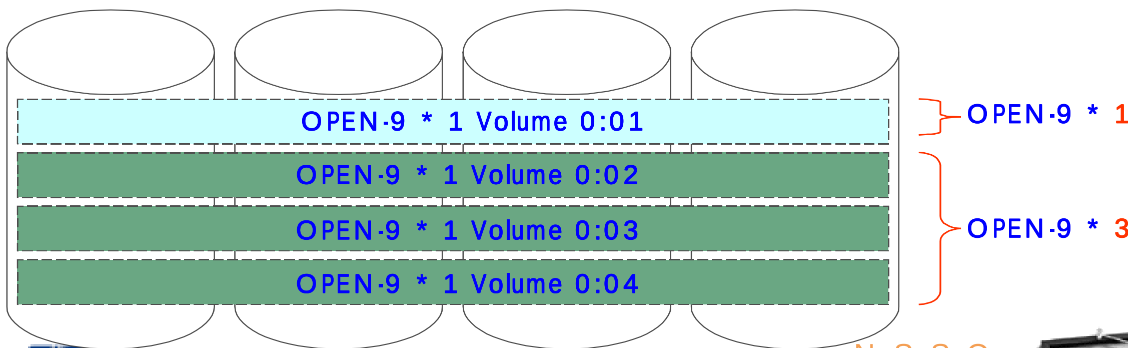
OPEN-9: Min 14.8 GB Max 267 GB

OPEN-3: Min 4.8 GB Max 86 GB

The LUN Size Expansion (LUSE) feature allows the creation of virtual LUNs that are larger than standard OPEN-3, OPEN-8 and OPEN-9 LUNs. Multiple individual LUNs are consolidated into a virtual LUN. With LUN Size Expansion the size of a selected LUN can be expanded to 259 GB. This capability allows open-system hosts to access data on the entire Disk Array XP 256/512 using fewer logical units than before expansion. The expanded LUN may consist of a mixture of different array (RAID) groups. However, the LDEVs in the array group must be contiguous in the Disk Array XP 256/512's CU (Control Unit) table in order to combine them into an expanded LUN. Open 3, Open 8 and Open 9 emulation types cannot be mixed in an expanded LUN. A minimum of two and a maximum of 36 LDEVs are required to form a LUN expansion. The LUN Size Expansion feature allows the eight LUN per port limit to be expanded to 120 LUNs per port for native FC.

LUSE Defined

- Logical Unit Size Expansion (LUSE)
- XP Array software that allows the creation of expanded Volumes that are larger than the defined emulations (that is, OPEN-3, OPEN-8, OPEN-E, etc...)
- Accessed via Remote Control XP, SVP, or Command View XP



NSSO
Stress-Free Storage...Guaranteed

LUSE Guidelines – What is Supported – What is NOT Supported

LUSE, What is supported?

- Up to 36 Volumes can be combined
- Non-contiguous Volumes can be combined with Titan microcode or later (spring 2000 release)

LUSE, What is not supported?

- Mixed RAID types within the same LUSE set
- Mixed CVS sizes
- Volumes with defined paths (LUNs)
- Combining different emulation types
- Combining volumes from different control units

NSSO
Stress-Free Storage...Guaranteed

The extended logical unit will assume the LDEV number of the lowest original LDEV number assigned to the extended unit.

LUSE- Consider the following...

Positives

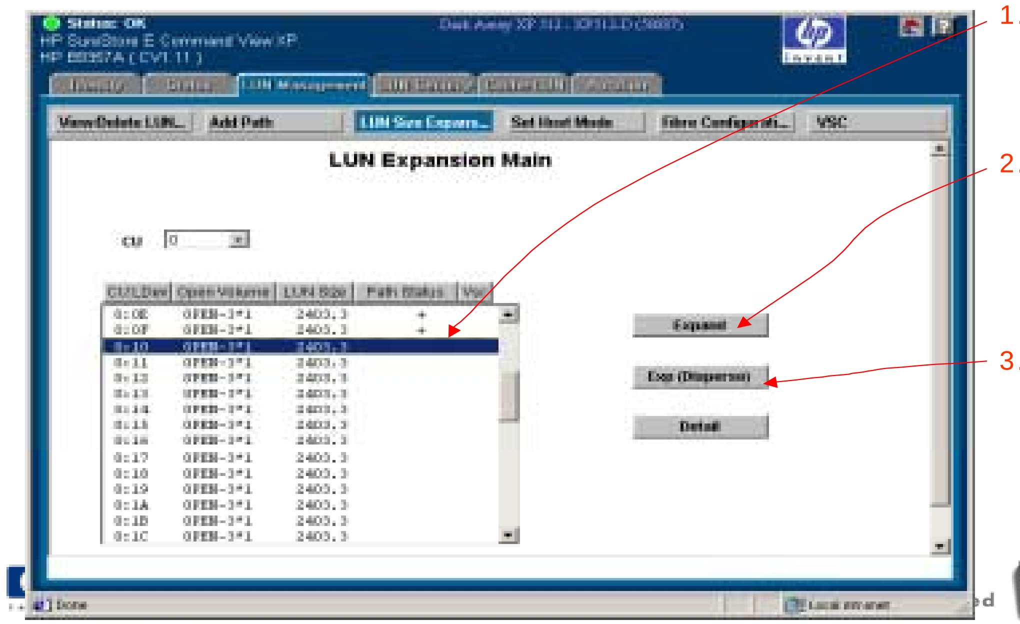
- The host can see one LUN that is larger than the maximum supported emulation
- Considering LUNS per port limitations, users can maximize the total available disk space.
- If the CU's span multiple ACP pairs, increased performance can be achieved by defining LUSE Volume members across ACP's

Negatives

- Some OS's might experience slow disk access times accessing large LUN's
- The total size of a LUSE LUN can affect the amount of time required to perform operations
- Queue depths can be adversely effected
- Maximum number of volumes supported by the array remains unchanged with or without LUSE

N S S O
Stress-Free Storage...Guaranteed

LUSE: Command View XP



Choose a volume from the LUN Size Expansion Main window
Choose Expand to combine contiguous volumes into a LUSE volume
-or-
3. Choose Exp (Disperse) to combine non-contiguous volumes into a LUSE volume
This topic is covered in greater detail in the Command View XP Module 8.

LUSE FAQ's

- Which volume label do I map to a port (LUN Mapping)?
  - The first member of the LUSE(i.e.- the lowest hex value) Volume.
- Does "non-dispersed" mean the available Volumes must be contiguous by value or contiguous by availability?
  - Both - contiguous by Volume number & availability (not mapped to a port).
- Using "dispersed" assignment of volumes, is there a way to add volumes of lower hexadecimal value than my volume number start point?
  - No

NSSO
Stress-Free Storage...Guaranteed

LUSE FAQ's (cont.)

- If data is already on a Volume that I want to include in a LUSE Volume, will the data be lost?
  - Yes. All LUSE operations are destructive.
- Will LUSE Volumes have different queue depths than normal Volumes?
  - That depends on what you mean. Queue depth is assigned on a per LUN basis, not the number of Volumes in a LUN. If a three Volume LUSE is mapped to a port, it will have the same queue depth (by default) as a regular Volume that is mapped to a port, because they each will have only one LUN number from the server's perspective. This could effect performance. For some OS's and HBA's, queue depth can be changed on the server.

NSSO
Stress-Free Storage...Guaranteed

LUSE FAQ's (cont.)

- Why would someone choose Dispersed over Non-Dispersed?
  - Dispersed could give increased performance benefit over Non-Dispersed if the Volumes are truly dispersed in the array.

N S S O
Stress-Free Storage...Guaranteed

Volume Size Configuration (CVS)

- Volume Size Configuration (CVS) is a built-in feature of LUN Configuration Manager XP.
- Formerly known as Custom Volume Size (CVS), CVS allows you to configure custom size volumes (CVs) smaller than normal volumes.
- If files on the same volume are frequently accessed, VSC can place the data on more than one volume
  - reduce logical drive contention
  - reduce host I/O queue times
  - better use physical storage capacity
  - improve balance of storage device I/O workloads
  - use with Cache LUN XP to dedicate less cache

NSSO
Stress-Free Storage...Guaranteed

Student Notes

The Volume size configuration (VSC) option on the XP48 and XP512 allows you to configure variable-size custom volumes. Creating reduced-size VSC volumes is the opposite of Logical Unit Size Expansion (LUSE).

VSC improves data access performance by reducing logical device contention and host I/O queue times, particularly when several frequently accessed files are located on the same volume. VSC enables better utilization of the physical storage capacity of the XP 512 or XP48, and reduces the amount of administrative effort required to balance I/O workloads.

Note: VSC (also referred as Custom Volume Size or CVS) used in conjunction with Cache LUN can produce greater performance improvements than when either of these options is used individually.

VSC operations can be performed by users with administrator or custom VSC access privileges. Users without administrator or VSC access privileges can only view the VSC configuration information for each attached disk array.

Note: VSC is also available for the XP256.

Volume Size Configuration (VSC), formerly known as Custom Volume Size (CVS), is used to create smaller LDEVs from standard Open-3/8/9 volumes.

VSC cannot be used on LUSE volumes.

VSC divides one logical volume into several volumes.

Every customized volume gets an LDEV#-the unused portion of the original LDEV (the “base volume”) retains the LDEV# of the normal volume. The standard (non-VSC) logical devices are called "Normal Volumes" and devices created via VSC are called “Customized (Custom) Volumes” or CVs (volumes with user-defined sizes) and "Base Volumes" (the volume being divided). Newly created CVs must be given a unique LDEV ID number. If none are available (i.e. your XP256 has been fully configured with 1024 LDEVs), a CV cannot be created. The original LDEV ID of the volume being repartitioned is maintained by the system to reference the "Base Volume". The CVs are not required to have the same emulation type as the Base Volume. VSC only creates the custom-sized volumes - the SCSI path assignment must be done from LUN Configuration Manager MC. CVs are easily recognizable from the ioscan on a host. The entry will be marked with " VSC " or " CVS " depending on the version of your software. Custom Volume Size (CVS), also knows as Virtual LVI, is an optional feature of the Disk Array XP 256/512 that allows multiple virtual volumes to be defined within a RAID group. In effect, it divides one LDEV into several volumes. CVS provides a dynamic means to optimize capacity, increase productive throughput, and reduce operational overhead expenses. It consists of a function to provide variable capacity volumes and a function to provide free arrangement of logical (virtual) volumes on the physical disk drives within a RAID group. Up to eight disk arrays can be managed from a single interface. CVS allows user-defined custom-size volumes that are smaller than normal volumes to be configured. It improves data access performance by reducing device contention as well as host I/O queue times, which can occur when several frequently accessed files are located on the same physical volume. Each of these files can be stored in separate logical volumes defined with CVS. A maximum of 32 normal and CVS volumes can be defined per RAID group. CVS enables better use of the physical storage capacity of the Disk Array XP 256 while reducing the amount of administrative effort required to balance direct-access storage device (DASD) I/O workloads. A data file with a high I/O frequency can be easily allocated to an independent volume thus decreasing the contention and performance bottleneck while increasing the subsystem s overall throughput and the host s performance. It can be used in conjunction with HP SureStore E Cache LUN XP to achieve better performance than when either of these options is used individually.

Custom Volumes

- Custom Volumes (CV's) are variable-sized custom volumes.
  - CV's can be created in sizes ranging from 36 MBytes up to the size of the OPEN emulation device size defined for the Parity Group.
  - Maximum of 64 CV's per parity group.
  - Maximum of 1024 CV's per array.

- CV's are created from Free Space.

NSSO
Stress-Free Storage...Guaranteed

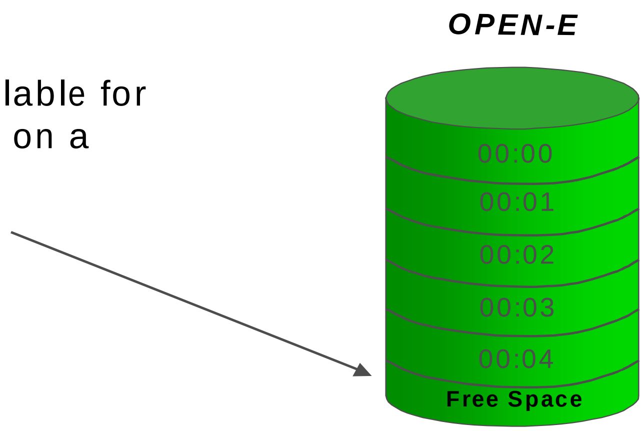
XP512 & XP48 Free Space

- For the XP512 and XP48, ALL Free Space is available for CVS use.
- If the Free Space left over after the Emulation Type definition process is insufficient for a customer's CVS needs, just free up additional unused Volumes to add to the available Free Space as necessary.

NSSO
Stress-Free Storage...Guaranteed

XP512 & XP48 Free Space

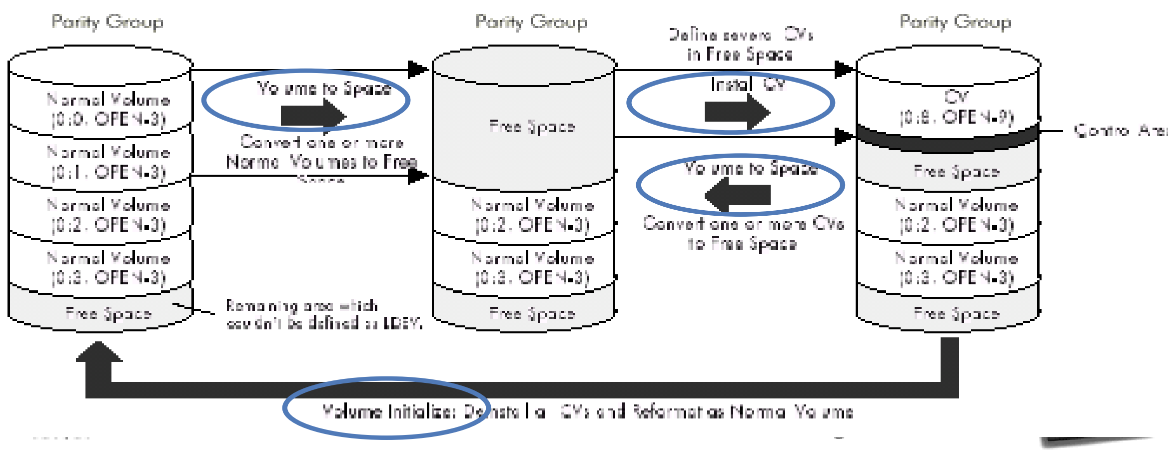
Free space available for CVS operations on a XP512/48



N S S O
Stress-Free Storage...Guaranteed

How VSC Works

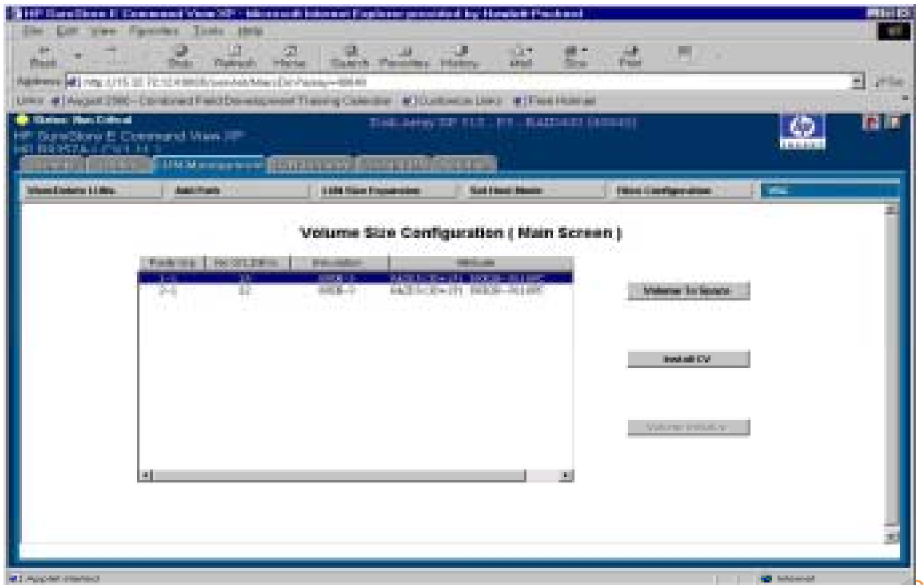
- VSC divides one logical volume into several volumes.
- Every customized volume gets an LDEV#
- The unused portion of the original LDEV (the “base volume”) retains the LDEV# of the normal volume



The standard (non-VSC) logical devices are called "Normal Volumes" and devices created via VSC are called “Customized (Custom) Volumes” or CVs (volumes with user-defined sizes) and "Base Volumes" (the volume being divided). Newly created CVs must be given a unique LDEV ID number. If none are available (i.e. your XP256 has been fully configured with 1024 LDEVs), a CV cannot be created. The original LDEV ID of the volume being repartitioned is maintained by the system to reference the "Base Volume".

The CVs are not required to have the same emulation type as the Base Volume. VSC only creates the custom-sized volumes - the SCSI path assignment must be done from LUN Configuration Manager XP. CVs are easily recognizable from the ioscan on a host. The entry will be marked with " VSC " or " CVS " depending on the version of your software. VSC should not be run on devices with live data. If you must divide a device that is already in use, back up the information and take the device offline before continuing.

Command View XP VSC Main screen



- Select the VSC button on the LUN Management screen
- Select a parity group and an operation

Stress-Free Storage...Guaranteed

Student Notes
To perform VSC operations, select the VSC button in the LUN Management screen. The VSC main screen is displayed.
All the VSC operations can be done using the following screens:

VSC Main Screen
Volume Size Configuration Screen
Volume to Space Operation Screen
Volume Initialize Screen
LDEV ID Configuration Screen

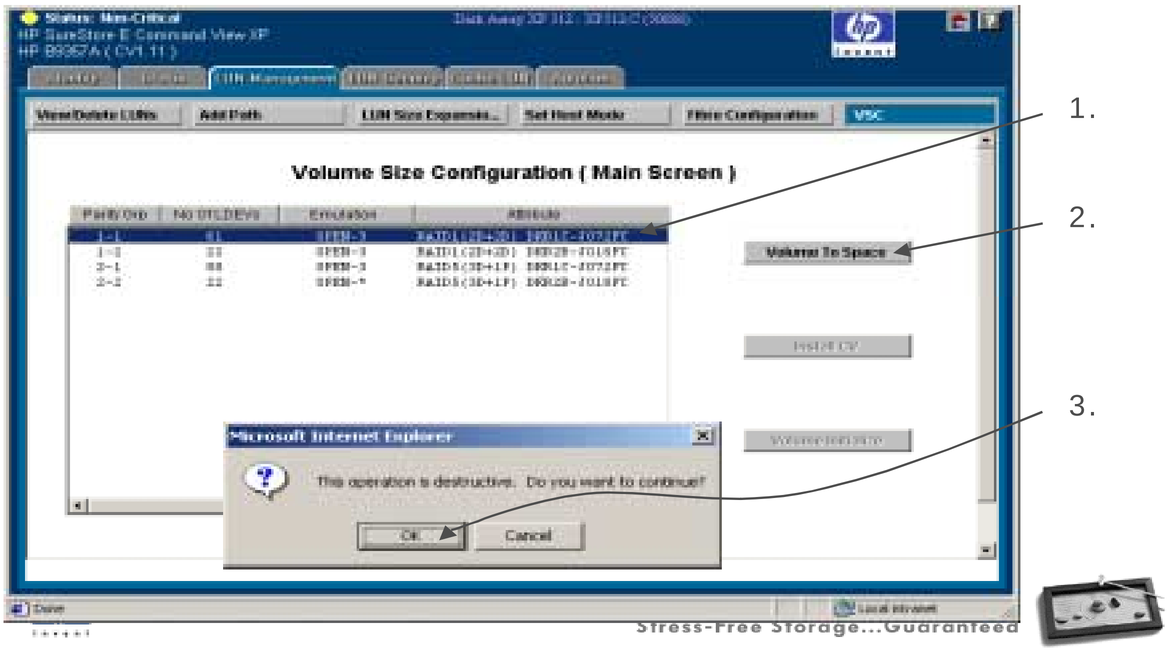
To start VSC operations:

1. Start Command View and select a disk array.
2. Select the LUN Management tab. The LUN Management page appears.
3. Select the VSC button on the LUN Management page. The VSC main screen appears.

Main Screen fields
The screen displays the Parity group, number of LDEVs in the parity group, emulation type, and attribute. RAID information is described in the attribute column as well as information about whether it is a VSC volume or not. The Emulation column will display "Open-*" if the parity group has volumes of different emulation types.
Main Screen buttons

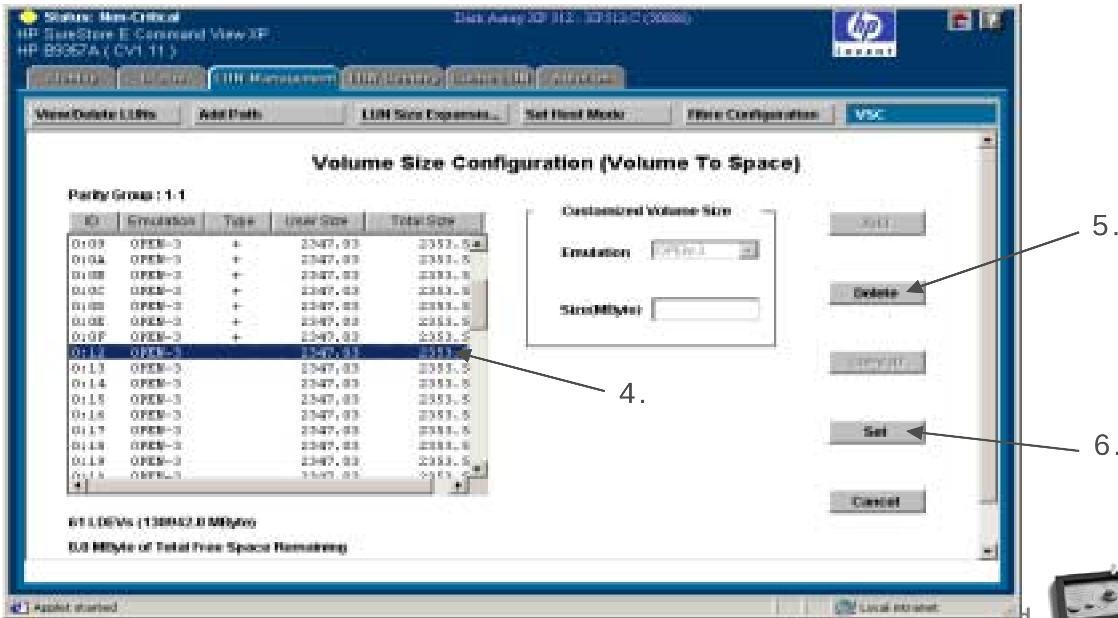
The three buttons allow you to select the three operations: Volume to Space, Install CV, and Volume Initialize.

Converting Logical Volumes into Space on a XP512/48 using Command View



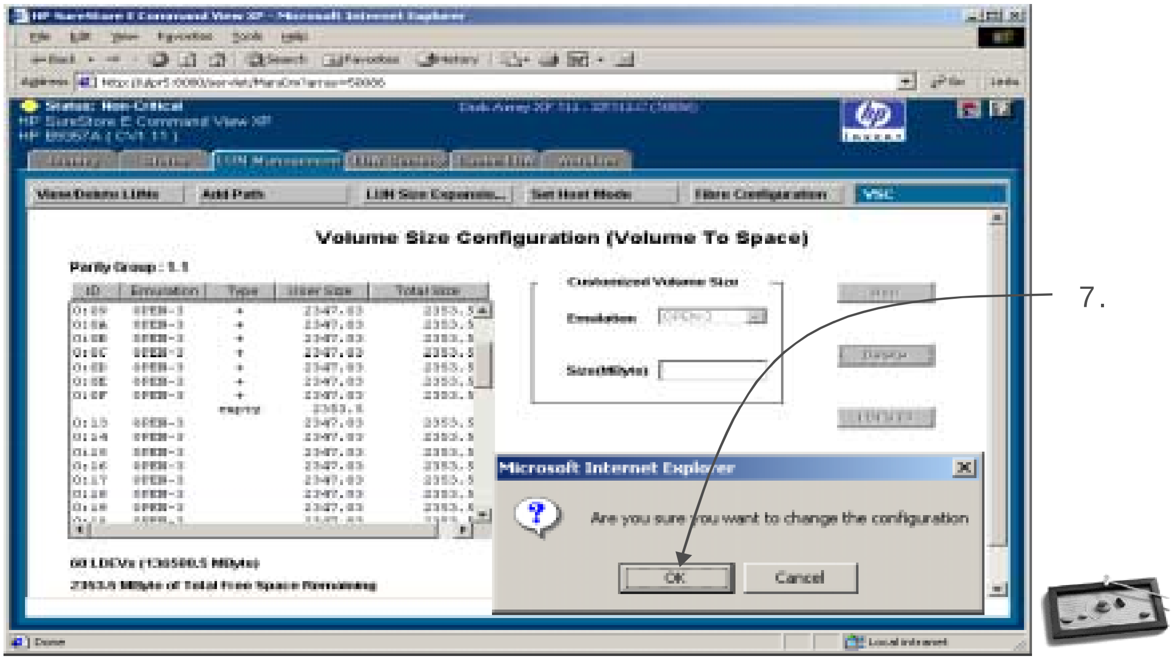
Select the parity group containing the logical volumes you want to convert to space (delete).
Select Volume To Space.
Select OK to continue the CVS Volume to Space operation.

Converting Logical Volumes into Space on a XP512/48 using Command View



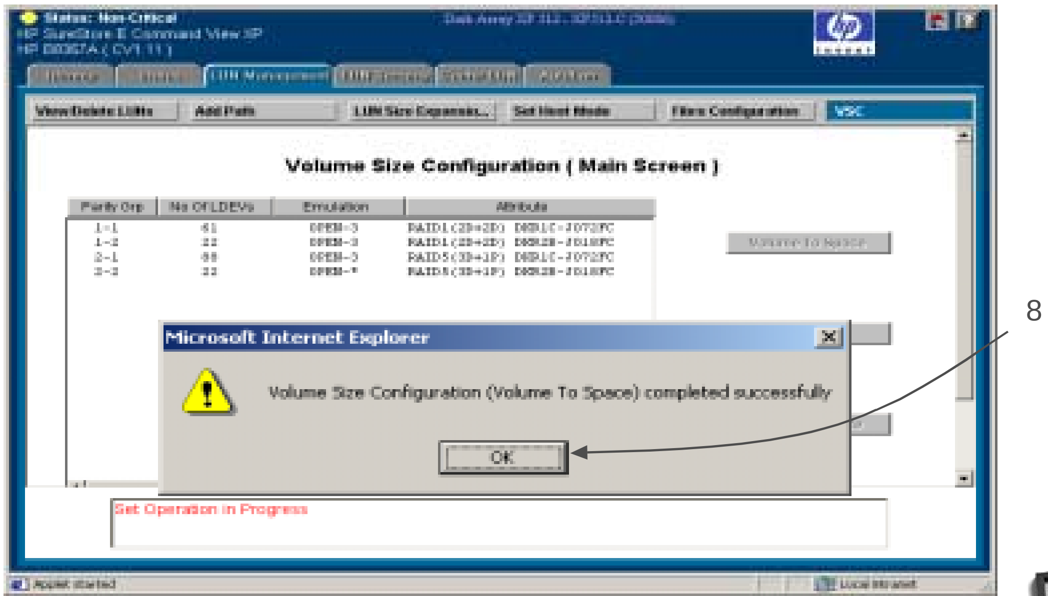
Select the volumes to be converted to space.
5. Select Delete. The Delete button is not available when either all of the LDEV's or the last remaining LDEV in the parity group are selected or the LDEV selected has a plus sign (+) next to it. If the selected volume has one or more assigned SCSI paths, or the volume is either an LDEV security volume or a LUSE volume, it is not available for VSC operations. The Customized Volume Size Define panel now displays the new CVS volume information.
6. Select Set.

Converting Logical Volumes into Space on a XP512/48 using Command View



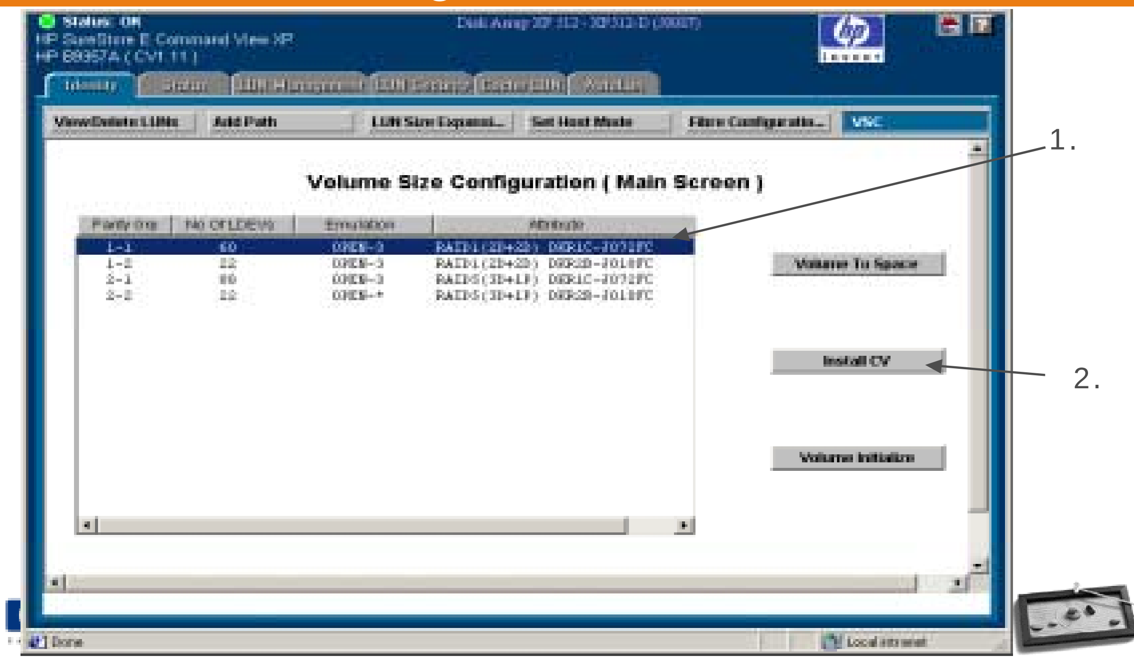
Select OK to change the volume to space.

Converting Logical Volumes into Space on a XP512/48 using Command View



Stress-Free Storage...Guaranteed

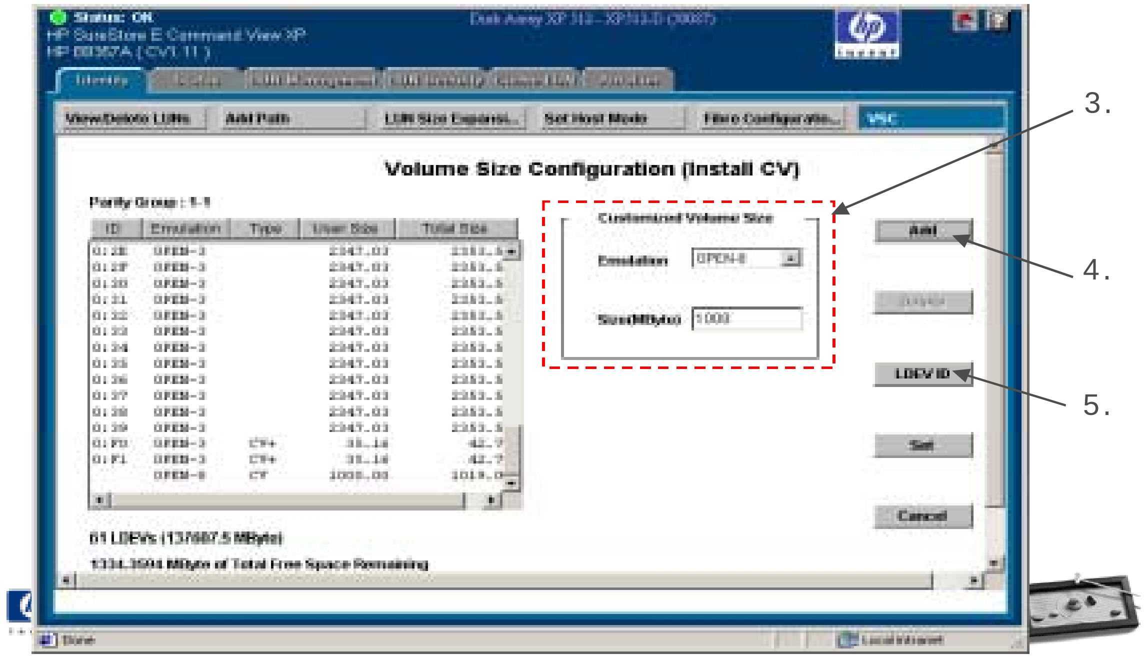
Defining and Installing CVS Volumes on a XP512/48 using Command View



Caution *This operation is destructive. If you are not certain that you want to implement the changes, select Cancel to cancel the operation.*

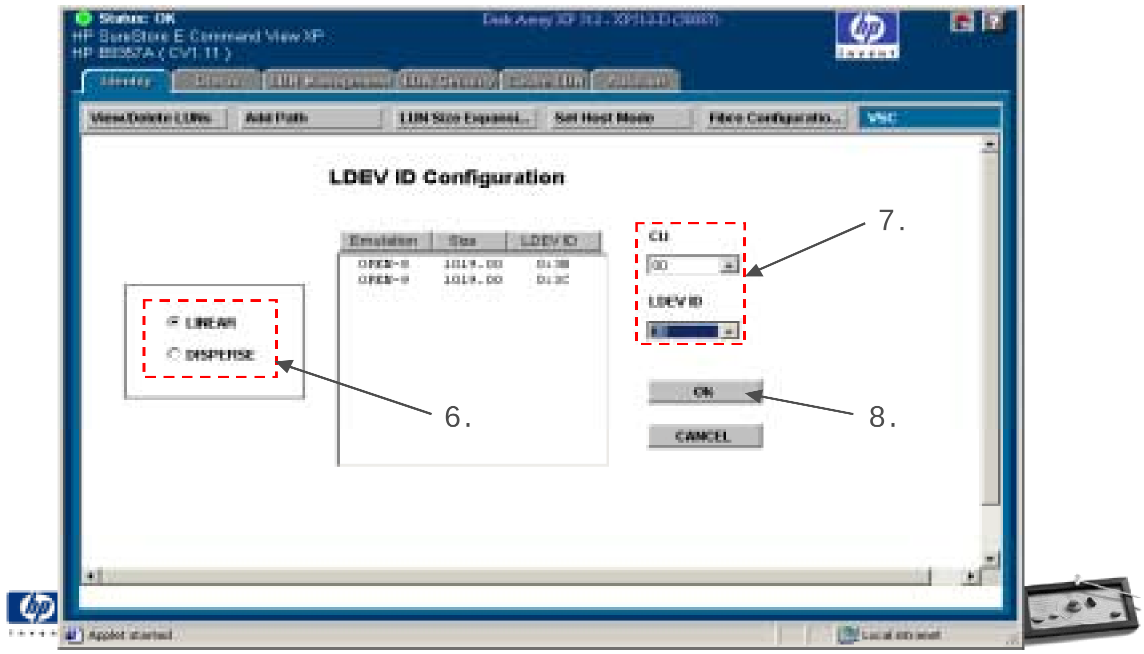
The Install CV function allows you to define and install one or more CV's under an existing volume. The installation operation does not occur until you select OK on the VSC Confirmation window at the end of the complete procedure.
Select the parity group that contains the space required to create the CV(s).
Choose Install CV.

Defining and Installing CVS Volumes on a XP512/48 using Command View



Select the emulation type for the new CV. Enter the volume size in MB.
5. Select Add to add the requested CV to the list box. Repeat the steps until you are finished adding the CV’s. If you make a mistake, select the incorrect CV and select Delete to remove it from the list box. The Cancel button restores this panel to its original configuration. When the information displayed on the Customized Volume Size Define panel is correct, select OK to continue.
6. Select Set to continue and complete the procedure for configuring CVS volumes.

Defining and Installing CVS Volumes on a XP512/48 using Command View

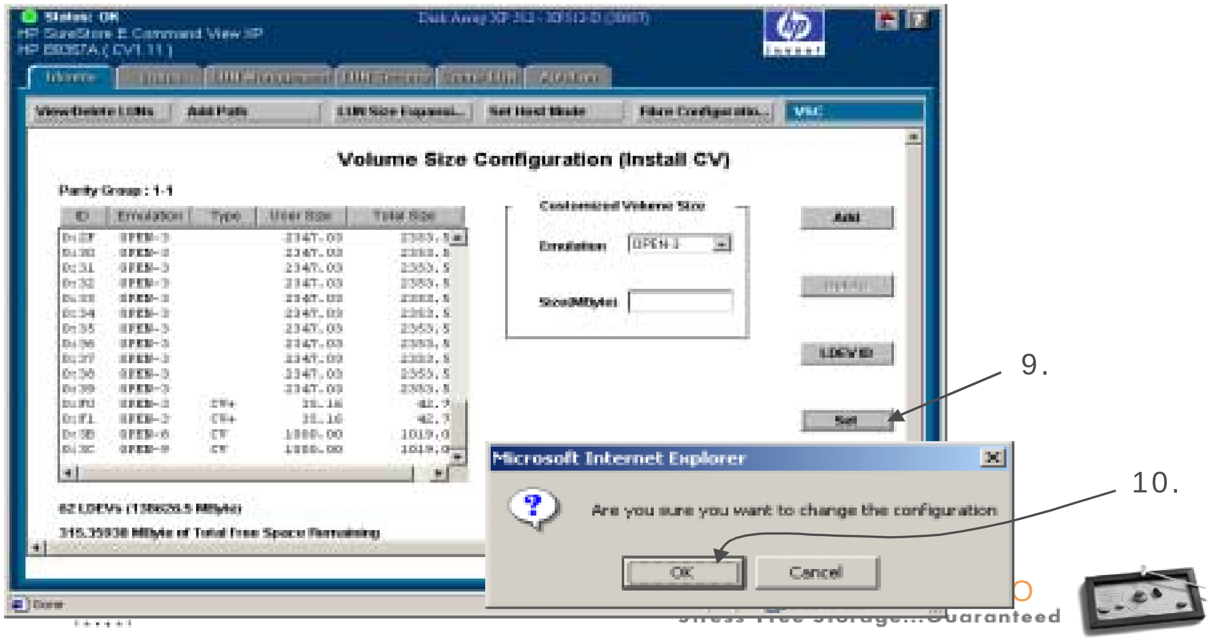


Linear: Assigns the LDEV ID’s in sequential order for all unassigned logical addresses within the parity group. The Linear option is appropriate for users who actively balance their DASD subsystem work loads.

Disperse: Randomly rotate the LDEV ID’s among multiple groups across the subsystem, eliminating hot spots and improving data access. The Disperse option is appropriate for users who do not actively balance their DASD subsystem work loads.

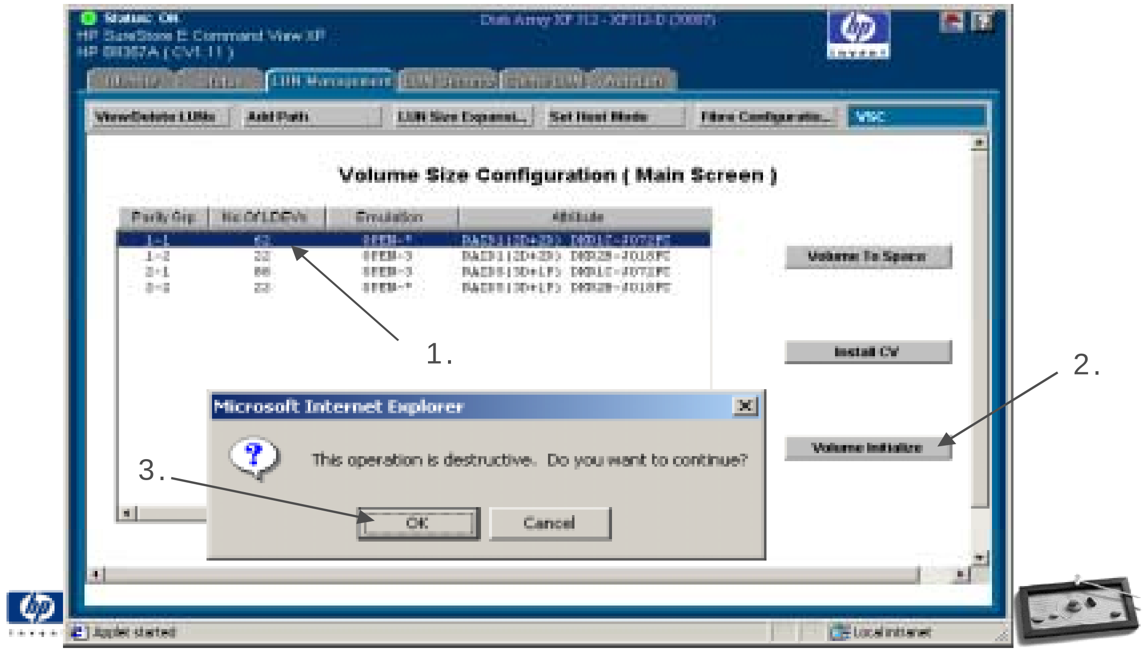
6. Choose either Linear of Disperse.
7. Select the CU, and enter the desired LDEV ID.
8. Select OK.

Defining and Installing CVS Volumes on a XP512/48 using Command View



Click Set to apply the VSC configuration settings.
10. Select OK to complete the VSC configuration operation.

Initializing a Volume on a XP512 and XP48 using Command View



Caution: The Volume Initialize function is a destructive operation. The data on the CV being converted is erased. Back up the data before performing this operation.

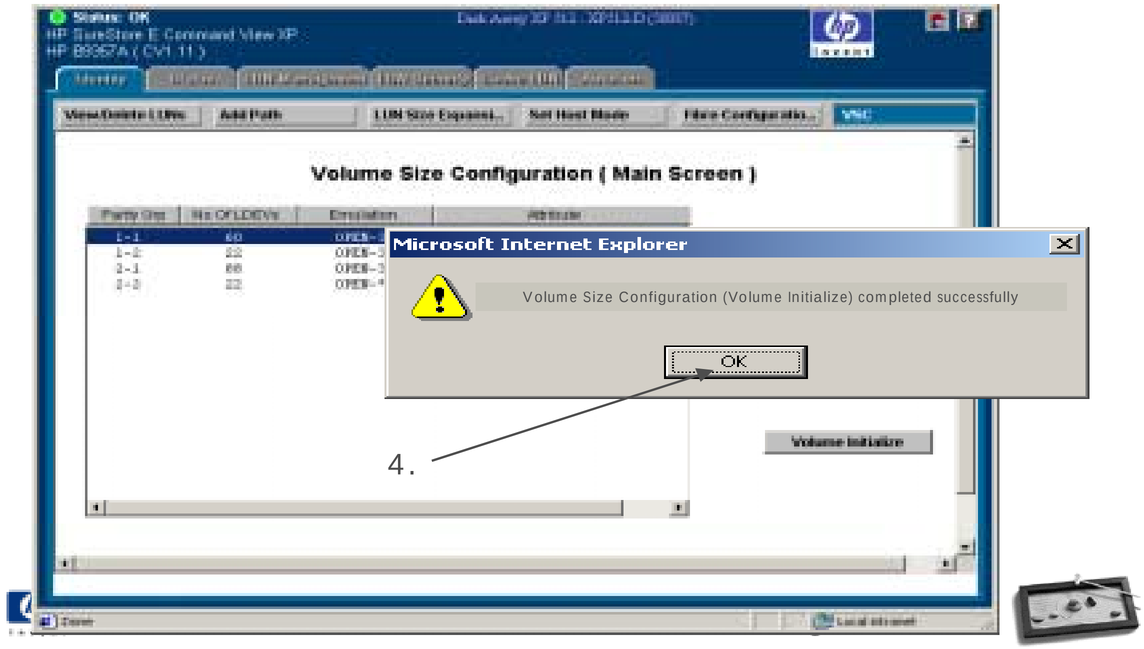
The Volume Initialize function allows you to convert a CVS volume back to a normal volume. This operation de-installs all customized volumes under a CVS volume and then reformats the CVS volume as a normal volume. The requested volume initialize operation does not occur until you select OK on the VSC Confirmation window.

Choose the parity group that contains the CVS Volumes that are to be initialized.

Select Volume Initialize to delete any CVS volumes within the selected parity group.

Click OK to complete the CVS volume initialization.

Initializing a Volume on a XP512 and XP48 using Command View



4. Click OK to close the dialog box.

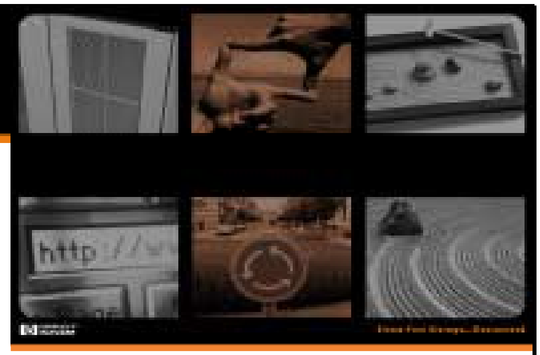
Review of Key Points

- LUN Configuration Manager XP Overview & Specifications
- Open System Volume Types Reviewed
- OS LDEV Calculation Worksheet
- SCSI Host & Modes Definitions
- Path Configurations for High-Availability
- LUSE Overview & Configuration
- CVS Overview
- Wrap-UP

N S S O
Stress-Free Storage...Guaranteed

Module Wrap-up

NSSO
Stress-Free Storage...Guaranteed



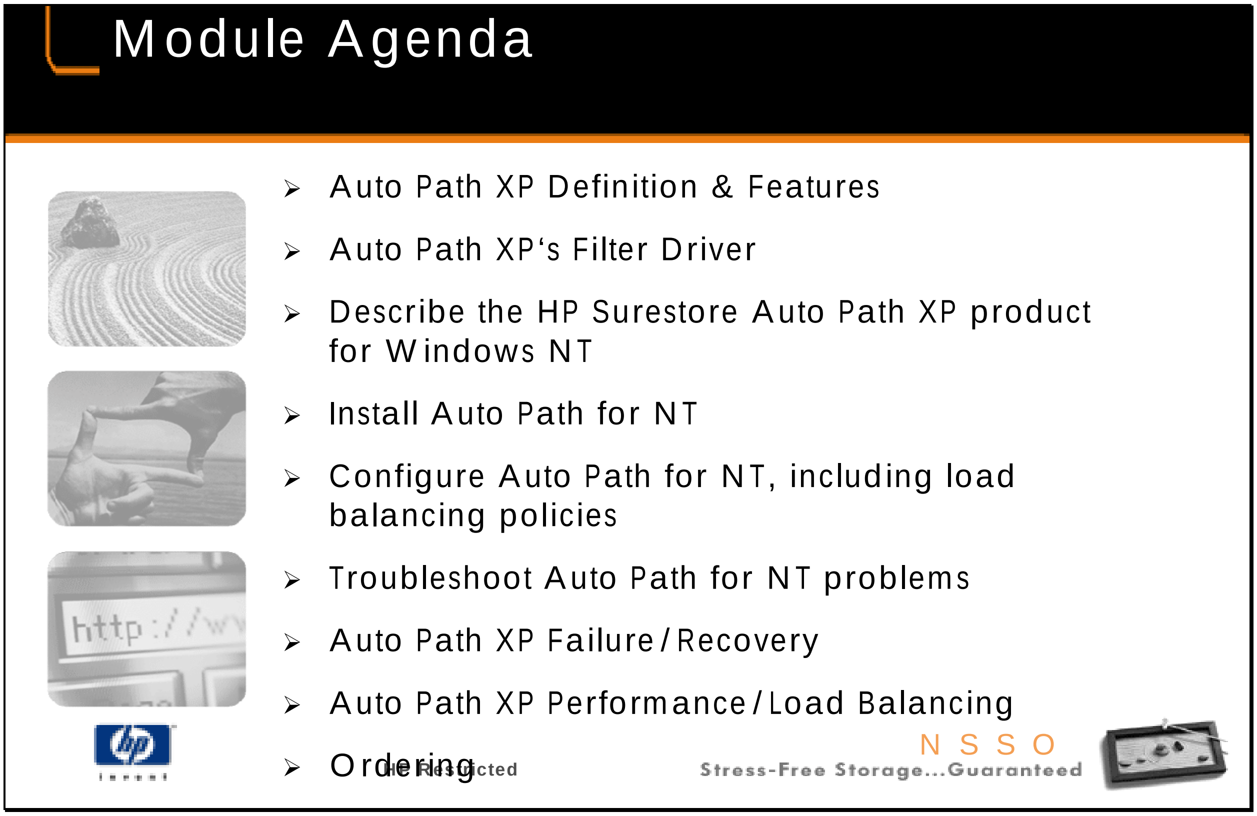
Module 3

Auto Path XP for NT Product - B7934A, AIX - B7936A & AutoPath for Windows 2000 – B9501A-04A

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

<u>One</u> Auto Path XP for NT products:
B7934A
License is per server, regardless of the number of paths between server and XP Family of arrays.
AutoPath XP is a softare application that resides on a windows NT host system attached to a SureStore E disk array family. AutoPath XP uses redundant paths between the server and the disk array to enhance reliability and performance. Because AutoPath XP balances the I/O workload over multiple paths, throughput is usually improved. Because AutoPath XP uses multiple paths, applications can continue to run without interruption if a path failure occurs.
Primary Line of Defense Against Downtime
Together with the XP family of arrays, HP Auto Path XP for NT is a primary line of defense against downtime for high-availability Windows NT installations. HP Auto Path XP has the ability to configure itself automatically and create filtering schemes or pseudo devices for partitions located on the XP Family of arrays. The application also enables automatic failover to an alternate path.



Hewlett-Packard (HP) offers high capacity, cost-effective, shared data storage with the HP SureStore E Disk Array XP Family of arrays to provide high availability for Microsoft® Windows NT®-based storage area networks (SAN). HP SureStore E Auto Path XP for Windows NT is central to HP's strategy to deliver round-the-clock data access.

Because Auto Path XP for NT features automatic recognition and dynamic load balancing of multiple paths, the software provides these additional benefits:

better performance for applications

higher service levels during times of peak operation

more satisfied users and customers

Auto Path XP for Windows 2000 Features

- Supports both stand-alone and clustered servers
- Supports both XP256 and XP512
- Supports Fibre Channel from server to array
    - HP D8602B FC host bus adapter supported at first release
      Others under investigation (Emulex LP8000, QLogic QLA2200F, Compaq 120186-B21)
- Microsoft certification in progress
- Up to 8 paths to a LUN, up to 128 LUNs / Server
- Failover
    - Operation of Auto Path XP is transparent to server applications
    - Eliminates HBA, cable, SAN and array processor as single points of failure
    - All changes to path state will be logged in the Event Viewer log
- Multi-path load balancing
    - prevents a single path from becoming overloaded causing I/O congestion
- HP designed and developed (not OEMed)

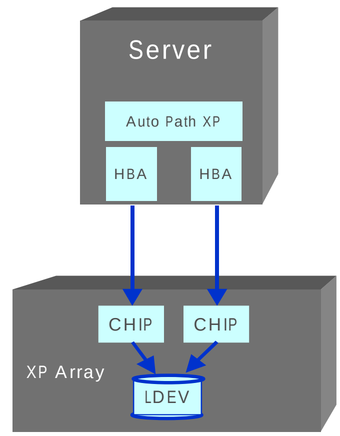
HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Auto Path XP Features

- Auto Path XP provides automatic path failover between an NT server and XP disk arrays
- Auto Path XP provides dynamic load balancing over multiple HBA paths

The result:
Enhanced data availability



Normal Operation

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

HP Surestore Auto Path XP for Windows NT has these two functions:

• automatic path failover and recovery to an alternate path
• dynamic load balancing over multiple paths

Auto Path performs dynamic load balancing while monitoring each path to ensure that the I/O is actually completing its transaction. In the event of a failure of any part of a path between the disk array and a server, Auto Path automatically switches to an alternate path, dropping the failed path out of the I/O rotation without any loss of data. The switchover is completely transparent to applications, so normal operation continues, without downtime.

Auto Path works on hosts with with multiple host bus adapters (and therefore multiple connections) to a disk array.

Benefit: enhanced data availability
Together with the XP family of arrays, HP Auto Path XP for NT is a primary line of defense against downtime for high-availability Windows NT installations. HP Auto Path XP has the ability to configure itself automatically and create filtering schemes or

pseudo devices for partitions located on the XP Family of arrays. The application also enables automatic failover to an alternate path.

HP Auto Path XP for NT recognizes multiple paths to a partition and performs dynamic load balancing on these paths. In the event of a failure of any part of a path between the XP Family of arrays and a server, HP Auto Path XP automatically switches to an alternate path, dropping the failed path out of the I/O rotation. The switchover is completely transparent to the customer’s applications, so operation continues normally, without downtime.

Auto Path XP for NT - Definition

- Host installed software that uses redundant connections between the host server and disk storage in a XP Family of arrays subsystem to provide enhanced performance and availability.
    - *Performance:* **Dynamic load balance of multiple paths**
    - *Availability:* **Automatic path recognition and failover recovery to an alternate path**

HP Restricted

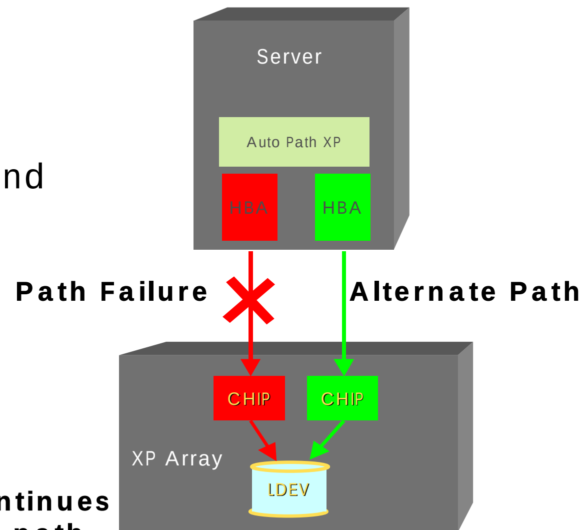
NSSO
Stress-Free Storage...Guaranteed

AutoPath XP is a host based software product that uses redundant connections between the host server and (the) storage devices contained with-in the XP Family of arrays array sub-system to provide enhanced performance and availability.
To clarify a little bit on a marketing level response regarding this definition. "Performance" in this definition means that we can have dynamic load balance (i.e. multiple paths for the XP Family of arrays). The term "availability" in this definition means that Auto Path has recognition and fail-over recovery to alternate paths to the (XP Family of arrays). Auto Path XP is a software application that resides on a Windows NT host system attached to a SureStore E disk array XP Family of arrays. AutoPath XP uses redundant paths between the server and the disk array to enhance reliability and performance. Because Auto Path XP balances the I/O workload over multiple paths, applications can continue to run without interruption if a path failure occurs. AutoPath XP provides these features:
enhanced data availability
automatic path fail-over and recover to an alternate path
dynamic load balancing over multiple paths
command line commands that display device, path and adapter information

Automatic path failover

- Failover functionality
  - Failover is transparent to applications
  - Supports Fibre Channel and Fibre Channel SANs
  - Works with MSCS clustering solutions

Path changes are logged to the Windows Event Viewer log



Operation continues even with a path failure

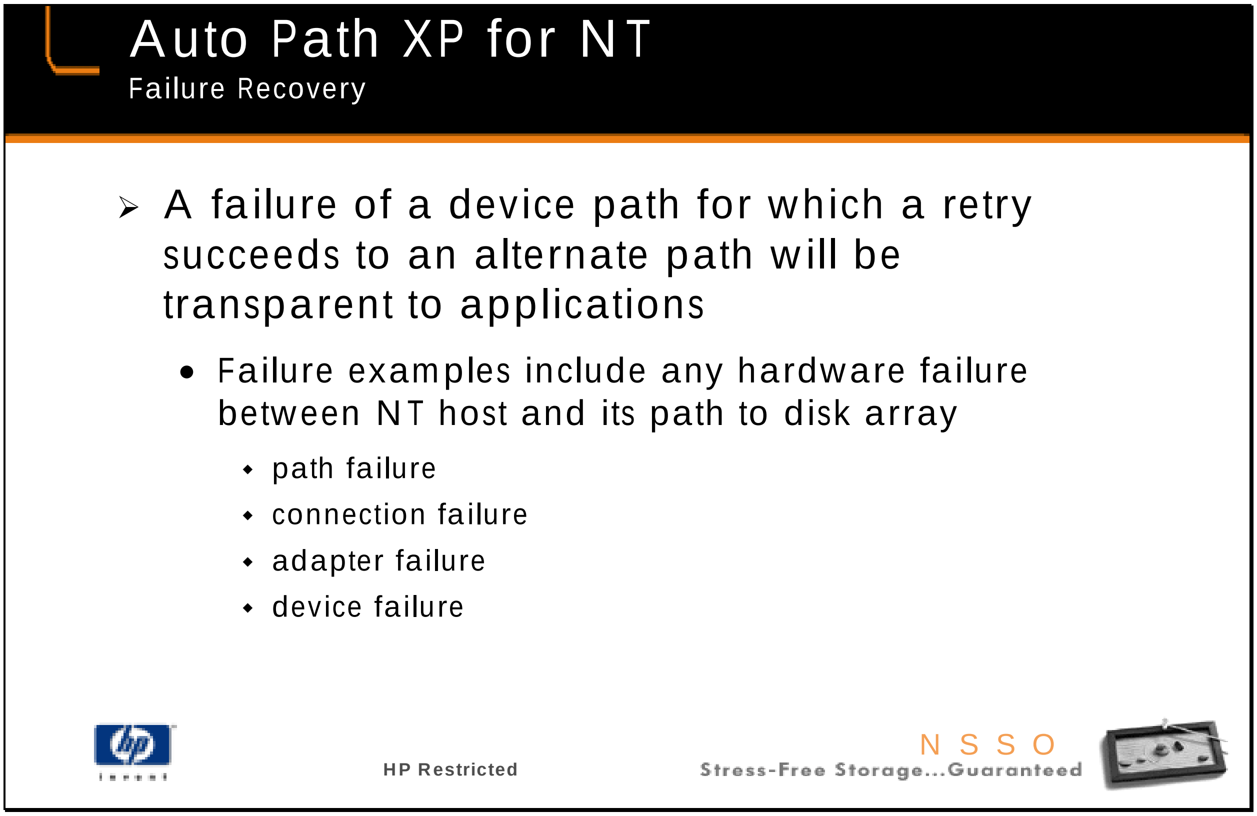
HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

If a data path fails, Auto Path automatically tries to switch to an alternate path. The failover is not visible to applications.
The amount of time required to accomplish the failover depends on the HBA used, but the average time is 60 seconds.
All path changes are logged in the Windows NT system event log, allowing service personnel to investigate and repair failed paths. The events reported by Auto Path are:

> The last path for a disk device has failed.
>
> A path to a device has failed, but other paths remain available. The device is still operational.



The failure of a device path for which it retries, succeeds to an alternate path, won't be transparent to the applications connected to the host, the applications running on that host. Failure examples include any hardware failure between the NT host and it's path to a disc array and obviously the LUN. We're going to talk a little bit more about what we see in the event of a failure and what it actually means; it is transparent, but not entirely, so we'll mention that.

Auto Path XP for NT

Event Log

- All changes to path state will be logged in the Windows NT Event Viewer log Path failures
- Example Event Viewer entries

Event Viewer - System Log on \\ATC-NT2 (Filtered)

Log View Options Help

Date	Time	Source	Category	Event	User	Computer
11/28/99	6:04:17 PM	EventLog	None	6000	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2
11/28/99	6:04:17 PM	Disk	None	15	N/A	ATC-NT2

invent

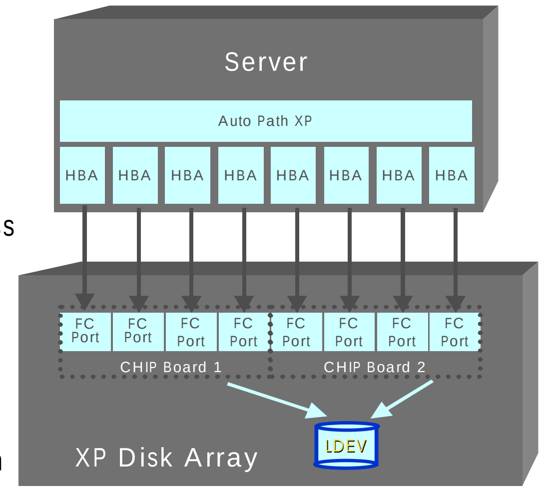
HP Restricted

NSSO
Stress-Free Storage...Guaranteed

In this test event that we have created, so all changes to path state will be logged to Windows NT event log viewer for logging path failures. That is a feature of this product that's going to remain unchanged. What's up for discussion is, how it's going to log in and how many times. In this example for this slide, HP had a NT server with I/O occurring on the sever, the cable from the back of that machine was pulled out, (I.e. no connectivity to XP256), and Auto Path XP went in and notified the system administrator over a hundred times that the path was unavailable. When you have Auto Path XP installed, it's going to allow you to actually remove all paths except for the last remaining one. If you were to remove Auto Path and disconnect all cables, NT is going to do this same thing, but it's only going to do it for one available path and not all the paths that are available.

Dynamic load balancing

- Load balancing functionality
  - Supports up to 8 HBAs (8 paths) per server, and all configured drives (i.e., drives that are recognized by the Disk Administrator. Dynamic disks are not available in Windows NT)
  - Provides dynamic load balancing across all installed HBAs in a server
  - There are 5 load balancing policies, including “no load balancing”
  - Supports the XP512, XP48 and XP256 disk arrays
  - Load balancing for MSCS supported in the next release



HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Auto Path detects multiple paths to each logical device and allocates the data load among the paths for optimum performance.
Multipath load balancing prevents any path from becoming overloaded, and it helps to prevent the congestion that occurs when many I/O operations are directed to common devices along the same path.
Load balancing is performed using a policy selected by the administrator. The policies find the least busy path, as determined by requests, bytes, or service time.

Load balancing policies

- No_Load_Balance
- Round_Robin
- Shortest_Queue_Requests
- Shortest_Queue_Bytes
- Shortest_Queue_ServiceTime

We'll see how to set a load balancing policy a little later

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

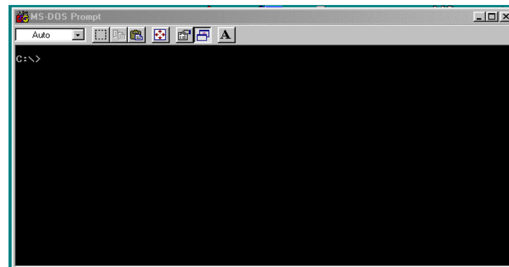
Student Notes

No_Load_Policy. No load balancing performed. All I/O is routed through a user-selected path called the Preferred Path.
Round_Robin. The requests are cycled though each path sequentially without regard to queue depth.
Shortest_Queue_Requests. Measures the current queue depth by counting the number of requests outstanding for each path.
Shortest_Queue_Bytes. Measures the current queue depth by the sum of the sizes of the I/O requests outstanding on that path.
Shortest_Queue_ServiceTime. (Default) Measures the current queue depth by the sum of the length of time each I/O request has been outstanding on that path.
"Shortest Queue" algorithms use instantaneous metrics, such as current queue depth, to select the path to be used.

The interfaces

- Graphical User Interface (all commands)

- Command Line User Interface (5 commands)



We'll look at operations a little later

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

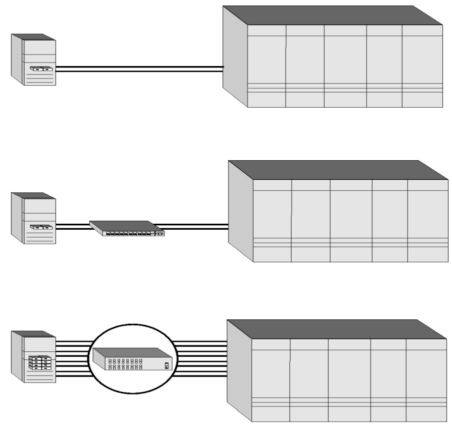
Student Notes

The Graphical User Interface (GUI) is browser-based, and provides access to all Auto Path commands.

The Command Line User Interface is executed via a DOS window. While the set of commands is limited, they are very fast and easy to use.

Configurations

- Direct Connect
- FC-AL
- Fabric Switch



Supported HBAs:	HP D8602A with 3.0.3245.1 driver (HP D8602B, QLogic QLA2200F with 7.05.05 driver coming in 04/2001 timeframe)
Configurations:	Direct Connect, FC-AL, Fabric Switch, Multi-switch
Disk arrays:	XP256, XP512, XP48
Clustering:	MSCS for failover (load balancing in future release)

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

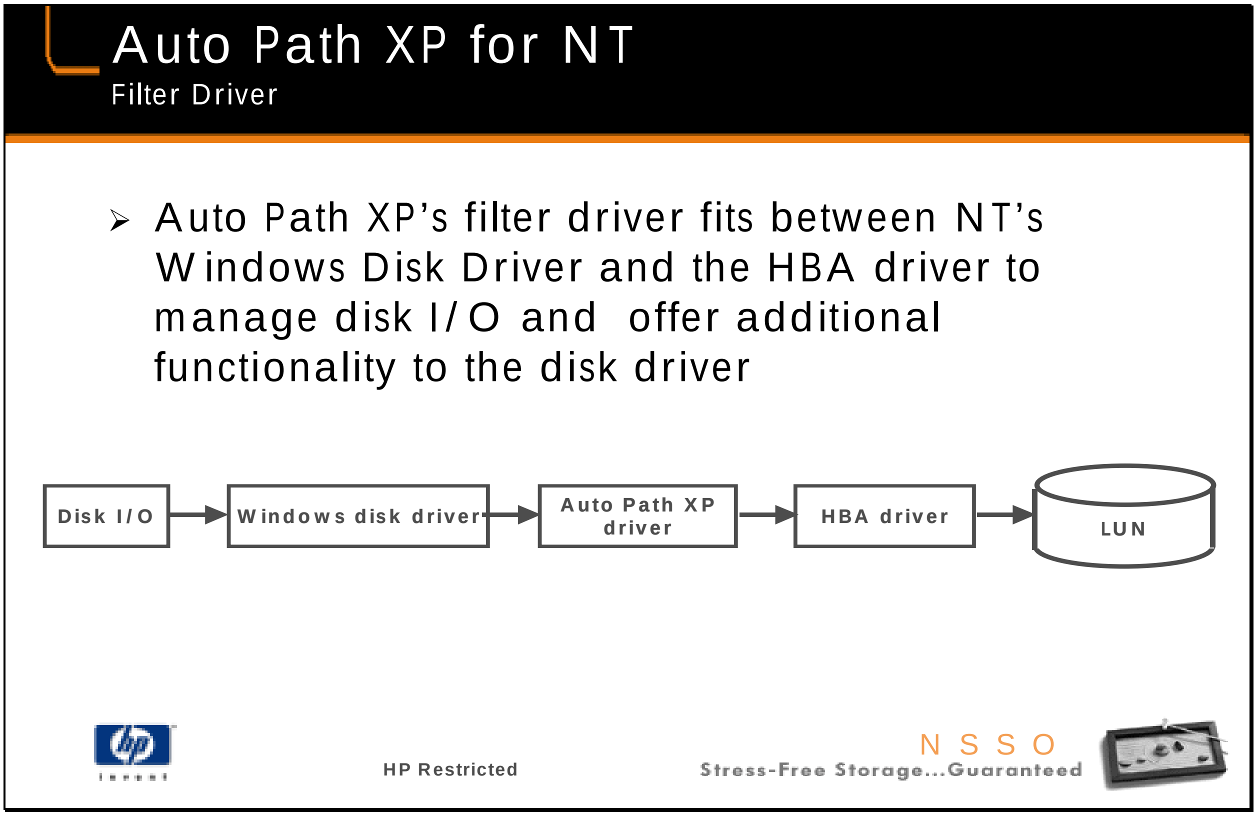
Student Notes

Generally, works with the range of Fibre Channel connections. However, hubs are not supported.

Note the limitation that the HBAs must be HP D8602A devices. The HBA driver must be 3.0.3245.1 or above for Auto Path XP version 1.00 for Windows NT 4.0.

The HP D8602B and QLogic QLA2200F with 7.05.05 driver will be supported in the 4/2001 timeframe. The latest details about supported HBAs are in the README.txt file on the installation CD.

MSCS clustering is supported in this release for failover. Support for load balancing will be in a future release.



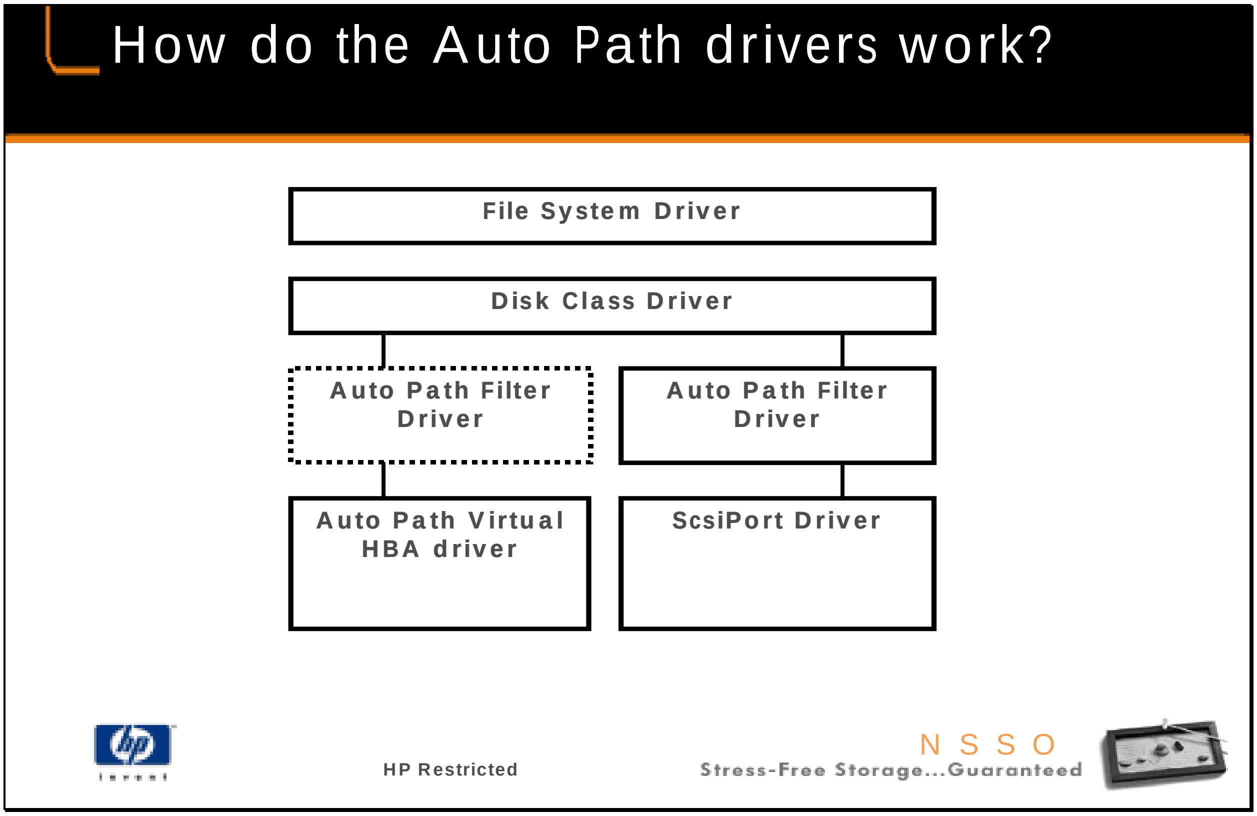
Student Notes
Normally, Windows disk I/O is sent directly to the disk driver.

However, Auto Path is a driver that fits between the Windows disk driver and the host bus adapter (HBA) driver. On the Windows NT system, Auto Path acts as a filter driver by intercepting the I/O between the driver and the disks.

Auto Path is transparent to Windows NT.

Auto Path is transparent to the disk array. The disk array requires no special software to be loaded on it.

Auto Path XP installs itself as a filter driver. Auto Path XP is going to allow us to do fail-over in the event of a path error or an adapter or device error. And as far as offering additional functionality, it will load balance between the available paths, so that's why it's actually referred to as a filter driver, it's sitting between already existing subsystems and offering us additional functionality connecting to the device. Auto Path XP is transparent to Windows NT.



Auto Path is implemented as a combination of a SCSI port filter driver and a virtual SCSI port driver. The two drivers together discover the multipath capable devices in the XP256/XP512/XP48 disk arrays and provide the load balance and failover support for the devices.

SCSI port filter driver. The SCSI port filter driver layers on top of the SCSI Port drivers (including the virtual HBA driver, shown in the dotted box) in the Windows NT 4.0 system. The filter driver discovers the multipath capable devices by processing SCSI inquiry commands to the SCSI ports on the system and passing appropriate data to the virtual SCSI port driver.

Virtual HBA driver. The Virtual HBA driver identifies the multipath capable devices from the information passed to it by the Auto Path SCSI port filter driver. The driver handles the I/O load balancing operations for the devices with multiple paths. In the event of failure of a given path, the driver routes the I/O through alternate paths available for the device.

Device discovery requirements. The Virtual HBA driver makes necessary calls to the SCSI ports to get the SCSI inquiry data about the multipath capable devices in the disk array and discovering the alternate paths to the devices and maintaining the list . The driver uses different mode pages data (Page 0x80, 0x83) for this purpose. The driver creates the device objects for the multipath capable devices. The driver can handle up to 8 paths per device. The maximum number of adapters supported is 8.

User interface requirements. The Auto Path XP driver supports device IOCTLs for providing user interface to the Auto Path XP software.

IOCTL_MPIO_QUERY_MULTIPATH_INFORMATION, for getting information regarding the multipath capable devices, paths and adapters in the system

IOCTL_MPIO_SET_LOAD_BALANCE_ALGORITHM, for setting the load balance policy for a device

IOCTL_MPIO_SELECT_PREFERRED_PATH, for changing the preferred path to a device

Performance monitoring requirements.
IOCTL_MPIO_QUERY_MULTIPATH_PERFORMANCE monitors the performance of the driver. This call returns a set of performance counters: BytesRead, BytesWritten, ReadTime, WriteTime, IdleTime, ReadCount, WriteCount, QueueDepth, QueryTime.

Auto Path XP for NT

Disk Administrator

- Device display from Disk Administrator
  - Without Auto Path XP, Disk Administrator shows each path to the LUN as a device/disk, even if each path is redundant to the same LUN (ie- you could assign > 1 drive letter the same LUN, format it, access it).
  - With Auto Path XP installed, Auto Path recognizes LUNs that contain multiple paths; in turn it marks the alternate paths as OFFLINE

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Inside of disc administrator, NT's utility for viewing your available discs or devices (ie. LUNs). NT's Disc Administrator shows disc or devices as active — active meaning it has an assigned drive letter – or as off-line. To further clarify what this means is — before we install Auto Path XP, NT is going to actually see each path to the LUN as a device or a disc. You know every path that you have defined to a LUN when you can go into the NT Disc Administrator, you will be able to see that LUN, you'll be able to assign a drive letter to that LUN for each and every path, even though it's all pointing to the same LUN.

One thing that is raising a lot of questions is that Auto Path (XP) cannot dynamically recognize new-disc partitions, deleted partitions or re-size partitions without a reboot.

Now, after we install Auto Path XP, Auto Path XP realizes that it's really just the one disc or one LUN that it's pointing to with multiple paths. So what it does, is it marks all alternate paths as off-line. Now, this probably sounds a bit confusing right now, but when we look at the next slide, we'll actually see a picture of the NT Disk Administrator illustrating that fact. Auto Path XP recognizes ***shared*** LUNs by using a unit serial number in the inquiry page. In Disk Administrator, the status of disks are displayed for the primary and secondary paths. The unused paths are displayed as ***offline***.

Installation requirements

- To install the Auto Path XP driver, Remote Access Server, and Command Line Interface
  - Windows NT 4.0 with SP6
  - 30 MB free disk space
  - An XP256/XP512/XP48 with formatted logical devices
  - At least one Fibre Channel HBA with driver installed
    - Use HP D8602A (HP D8602B, QLogic QLA2200F available soon)
    - Must have driver revision 3.0.3245.1 or above
- To install the Auto Path XP Remote Access Client
  - Windows 2000, Window NT 4.0, Windows 95, Windows 98
  - 30 MB free disk space
  - Display resolution of 1024x768 with 65536 colors

Student Notes

As stated in an earlier slide, the HP D8602B and QLogic QLA2200F with 7.05.05 driver will be supported in the 4/2001 timeframe. The latest details about supported HBAs are in the README.txt file on the installation CD.

To install Auto Path:

1. Log on as the administrator user.
2. Select **Start/Run. Enter d:\setup.exe** in the text box, where d: is the drive letter of the installation media or shared drive. Click OK to begin the installation.
3. The Auto Path XP Setup welcome screen is displayed. Click **Next**.
4. The License Agreement screen is displayed. Click **Yes** to accept the License Agreement.
5. The User Information screen is displayed. Enter the user name and company name and click **Next**.
6. The Select Destination Location screen is displayed. Accept the default destination directory or choose a directory where you want the software to be installed. Click **Next**.

7. The Select Components screen is displayed. Select the components to be installed and
click **Next**.
8. The Select Program Folder screen is displayed. Choose the folder or accept the default
and click **Next**.
The setup program will install the selected components on the computer.
9. Follow the on-screen instructions to complete the installation.
10. At the end of the installation, select the option to reboot the computer and click **Finish**.

Starting Auto Path

- Start the HP Auto Path XP server
- Start the user interface client
- The HP Auto Path Client GUI opens

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

To start the Auto Path XP server:

1. If you have Auto Path icons on your Windows desktop, double click the icon labeled
   Start HP Auto Path XP Server
2. If you do not have Auto Path icons on your Windows desktop, select:
   Start/Programs/HP Auto Path XP/Start Auto Path XP Server
   A command window appears.

To start Auto Path XP user interface client:

1. If you have Auto Path icons on your Windows desktop, double click the icon labeled
   Start HP Auto Path XP Client.
2. If you do not have the Auto Path icons on your Windows desktop, select:
   Start/Programs/HP Auto Path XP/Start Auto Path/XP Client
   The Auto Path XP GUI opens.

If this is the first time that Auto Path is opened after installation, the Servers window does not show any servers configured. To add servers, you will use the Add new server window.

Configuring Auto Path

- Configure Auto Path by doing the following:
  - Add servers
  - Specify load balancing policies or a preferred path
  - Specify the size of the error log file (optional)
  - Specify chart settings (optional)

The following slides will show how to do these tasks

HP Restricted

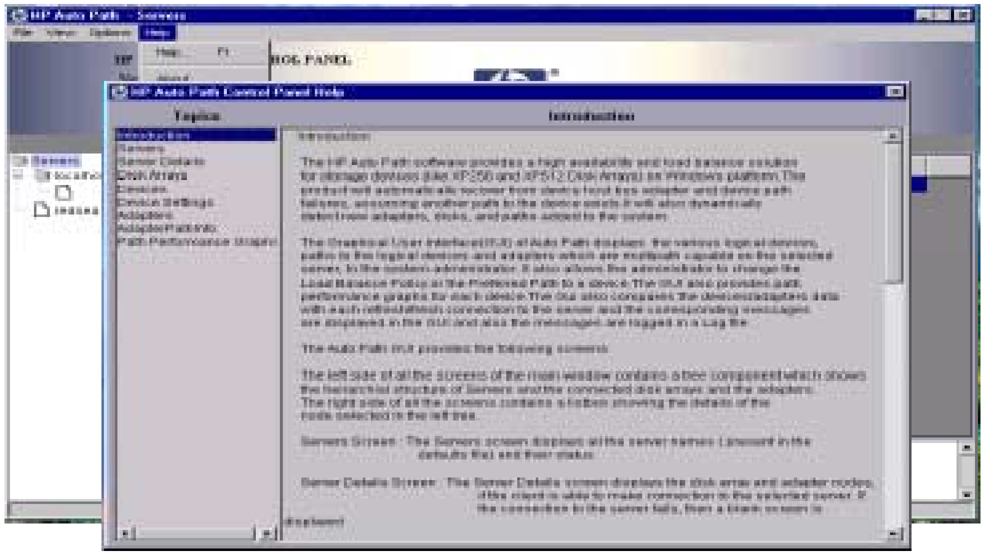
NSSO
Stress-Free Storage...Guaranteed

Student Notes

Configure Auto Path by doing the following:

1. Add the servers with configurable devices.
2. Specify load balancing policies or a preferred path for each server/disk array/device
3. Specify the size of the error log file (optional). There is a default size in place.
4. Specify chart settings (optional). There are default settings in place.

Online help

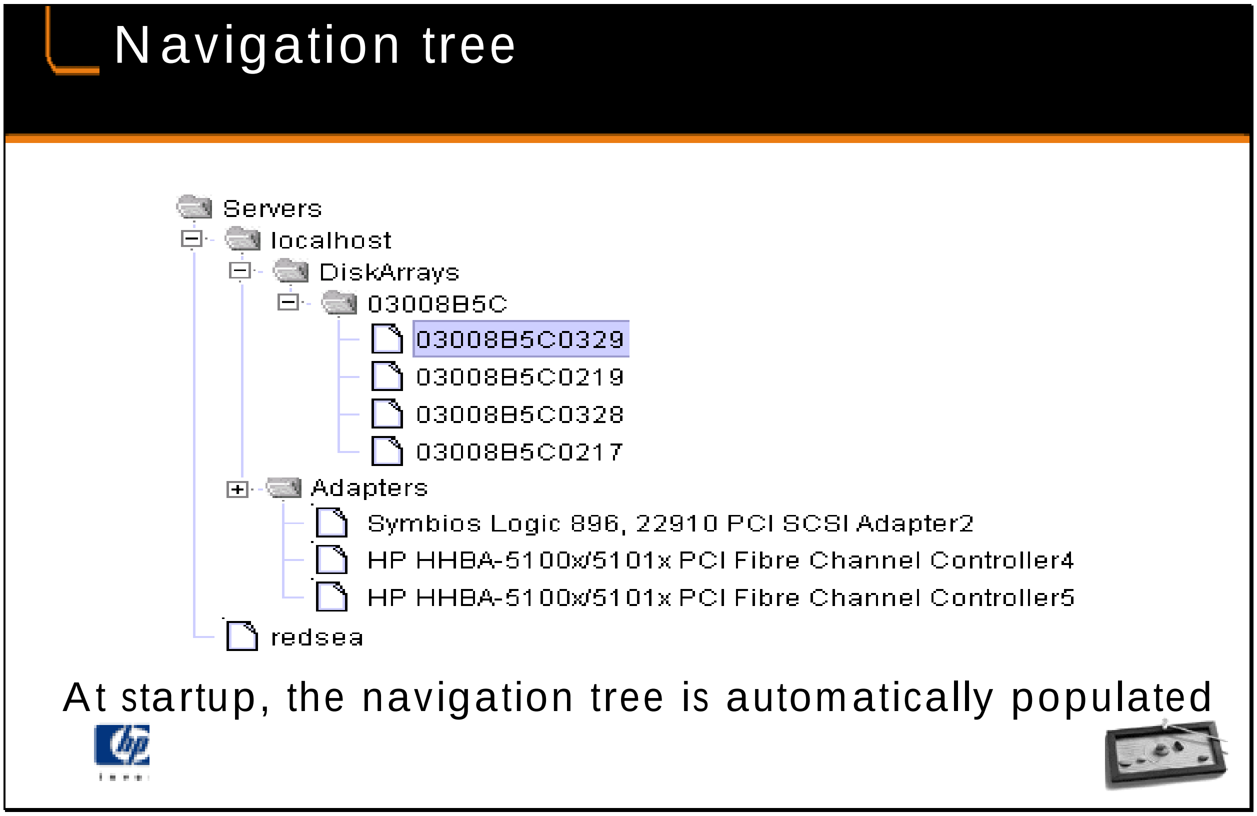


HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

Access Auto Path help using the Help dropdown menu.



Student Notes

Auto Path displays navigation tree that shows the distribution of servers, disk arrays, and adapters in the system. At startup, the tree is automatically populated. The tree structure implies a hierarchy:

Note: the example in the slide is from an AIX system. Windows NT supports different HBAs.

Tree Structure

Servers
  Disk Arrays / Adapters (same hierarchical level)
    Devices (aka LUNs)

ServersRoot of the hierarchy that contains all configured servers as nodes. Select to display the Servers window.

Server nodes Expand a server node to show disk arrays and adapters.

Disk Arrays Node that contains disk arrays on the parent server. Select to display the Disk Arrays window. Expand to show disk array nodes.

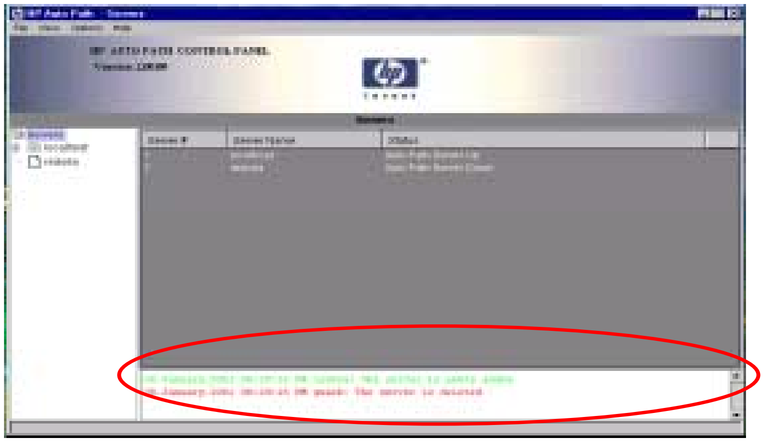
Disk Array nodes Disk array nodes are identified by disk array names. Select to display the Devices window. Expand to show devices in the disk array unit.

Device nodes Device nodes are identified by device names. Select to display the Device Settings window.

Adapters Node that contains adapters on the parent server. Select to display the Adapters window. Expand to show adapter nodes.

Adapter nodes Adapter	Adapter nodes are identified by adapter names. Select to display the Path Info window.

Diagnostic messages



- Red means a connection to a server has been lost (or a device / path has disappeared)
- Green means a connection to a server has been made (or a device / path has come up)

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

All Auto Path windows have a diagnostic message panel. The panel is located in the bottom right of the window.

The error messages are color coded:

red A connection to a server has been lost, or a device/path has disappeared.

green A connection to a server is established, or a device/path has come up.

Error log file

- The error log file is \hpss\Auto Path\data\autopath.log
- Set the log file size through the Set Log File Size command on the Options menu.

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

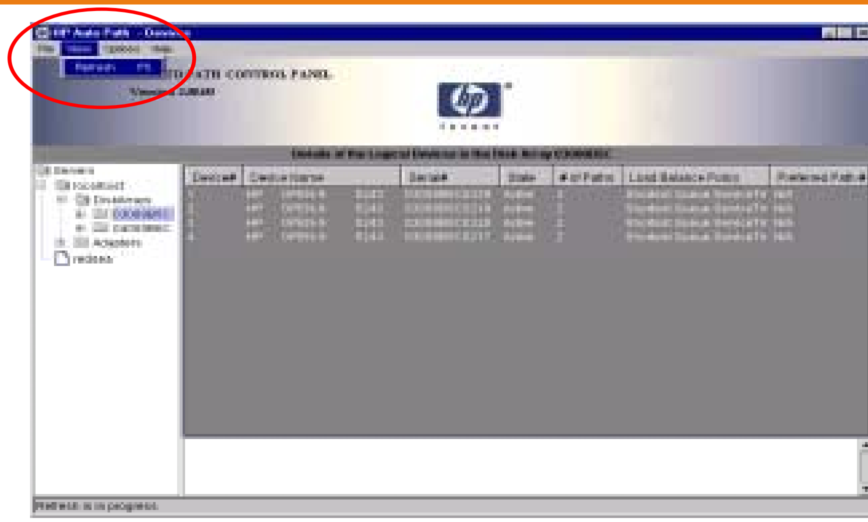
Student Notes

Error messages are recorded in an Auto Path log file. This file is stored as:

\hpss\Auto Path\data\autopath.log

Set the log file size through the Set Log File Size command on the Options menu. The example shows a log file size of 1000 bytes, but you may wish to set a larger size.

Refreshing windows



- Use the View dropdown menu to refresh window displays

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Use the View dropdown menu to refresh window displays.

For example, use the Refresh function to show the current state of a disk array.

Servers window

- Auto Path opens with this window
- To return to it, select Servers in the navigation tree

HP Restricted

Stress-Free Storage...Guaranteed

Student Notes

Auto Path opens with the Servers window. This window shows the names of the servers and the status of the Auto Path software on them.

To access this window, select Servers, the root of the tree structure.

The Servers window displays server information:

Server # An index number used by Auto Path for identifying the server.

Server Name The server name.

Status The status of the server:

Auto Path Server Up The server is operational.

Auto Path Server Down The server is not operational.

Server details window

- To see this window, select the node of a server in the navigation tree

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

The Server Details window shows the disk arrays and adapters connected to the selected server. This window is a view-only window; there are no user-configurable functions.

To access the Server Details window, select the node of a server in the tree view. The content of the window depends on the connection status of the selected server.

Successful Server Connection
If the client is able to connect to the server, the DiskArrays and Adapters folders appear in the main portion of the window.

DiskArrays display the	A folder equivalent to the DiskArrays node in the tree view. Select it to Disk Arrays window. Also, the navigation tree expands.
Adapters display the	A folder equivalent to the Adapters node in the tree view. Select it to Adapters window. Also, the navigation tree expands.

Server details window – failed server connection

- Window is empty. Status bar has a “fail” message

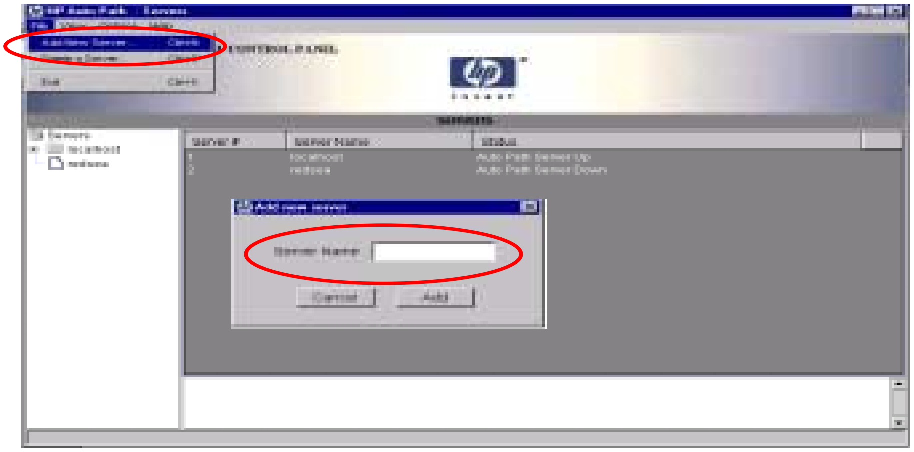
HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Failed Server Connection
If the client isn’t able to connect to the server, the main portion of the window is empty. The status bar at the bottom of the window displays the message **Failed to connect to the server.**

Adding a server



- Select File / Add New Server
- Enter the server name or IP address
- If the server name is not yet in DNS, use the IP address

HP Restricted

Stress-Free Storage...Guaranteed

Student Notes

To add a server:

1. From the Servers window, select Add New Server from the File dropdown menu.

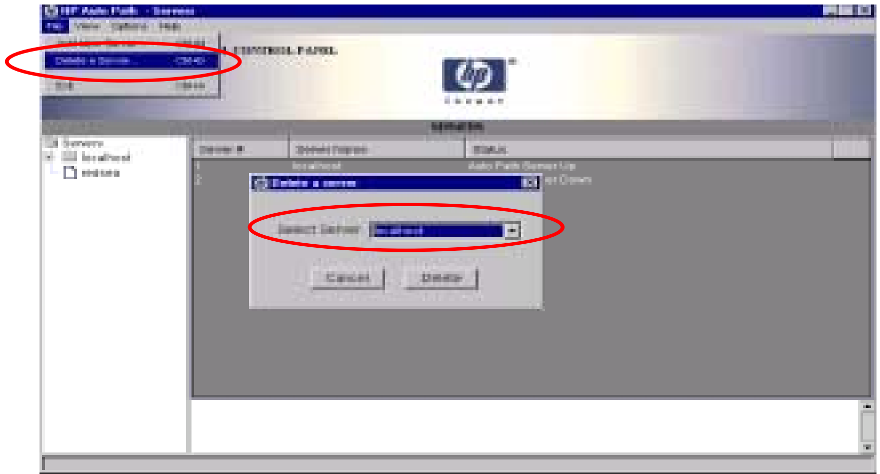
   The Add New Server dialog box opens.

2. In the Server Name box, type the server name (for example, **moh.gandhi.hp.com**) or the IP address (for example, **12.34.56.78**) for the server you want to add. If the server is not yet in DNS, user the IP address.

3. Click on **Add**.

The server name is added as a node on the Servers tree in the left panel of the Servers window. The corresponding server information displays in the right panel.

Deleting a server



- Select File / Delete a Server
- Enter the server name or IP address

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

To delete a server:

1. From the Servers window, select Delete a Server from the File dropdown menu.

   The Delete a server dialog box opens.

2. In the Select Server field, enter the name or IP address of the server you want to delete.

   You can also select a server from the dropdown combo box.

3. Click on **Delete**.

The node of the selected server is deleted from the Servers tree in the left panel of the Servers window. The corresponding server information is removed from the right panel.

Disk Arrays window

- To see this window, select the DiskArrays node of a server in the navigation tree

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes
The Disk Arrays window shows the details of the disk arrays connected to the selected server. This window is a view-only window; there are no user-configurable functions. To access the Disk Arrays window, select the DiskArrays node of a server in the navigation tree.

The Disk Arrays window displays disk array information:

Disk Array # An index number used by Auto Path for identifying the disk array.

Disk Array Name The name of the disk array.

# of Devices The number of devices on the disk array.

Devices window

- To see this window, select the disk array node in the navigation tree

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

The Devices window displays the devices for the selected disk array. This window is a view-only window; there are no user-configurable functions.

To access the Devices window, select a disk array node in the tree view.

The Devices window displays device information:

Device# An index number used by Auto Path for identifying the device.

Device Name The device name.

Serial# The logical device serial number. A device serial number has 12 hex digits (xxxxyyyyzzzz) and consists of the following elements:

xxxx: Array ID #

yyyy Array Serial #

zzzz LUN ID#

State The state of the device:

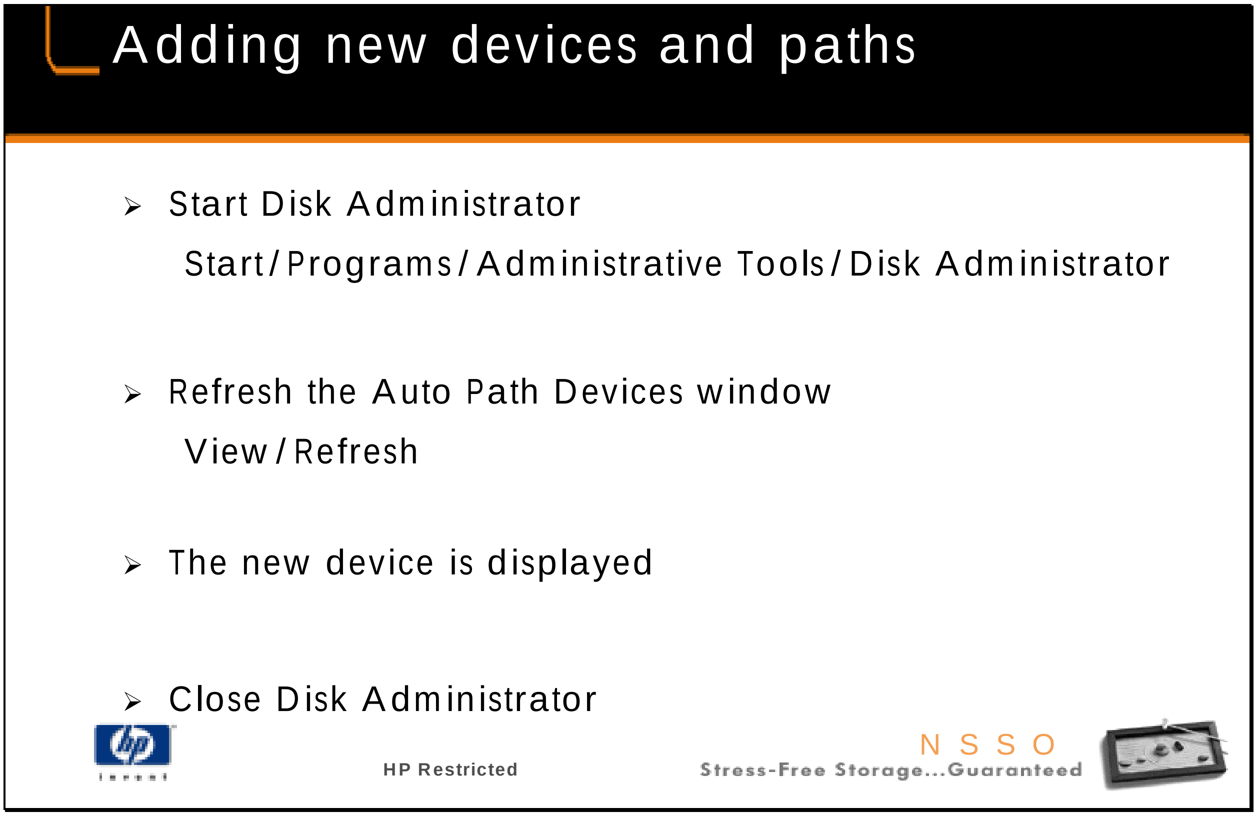
Active The device is in use.

Failed The device is no longer in use or all paths to the device have failed.

# of Paths The number of paths attached to the device.

Load Balance Policy The policy currently selected for the device.

Preferred Path # For devices with no load balance policy, this shows the selected preferred path.



Student Notes

When you add a new device or a new path to an existing device, you must run Windows NT Disk Administrator so Auto Path can recognize it.

To recognize new devices or paths in Auto Path:

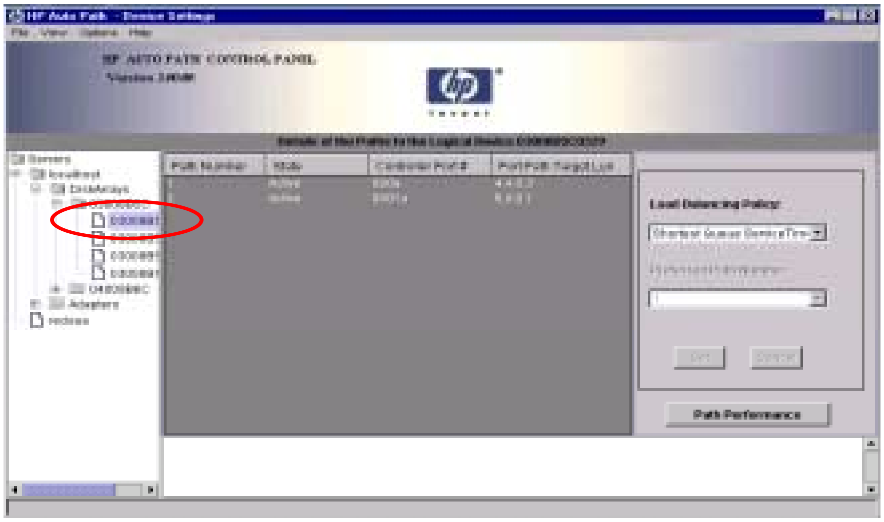
1. Start Windows NT Disk Administrator by selecting
   Start/Programs/Administrative Tools/Disk Administrator
   Disk Administrator opens.

2. After Disk Administrator is fully initialized and executed, you must refresh the Auto Path GUI.
   Return to the Devices window and select Refresh from the View menu.

   The new device or path is displayed in the Devices window.

3. Close Disk Administrator.

Device Settings window



- To see this window, select a device node in the navigation tree

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

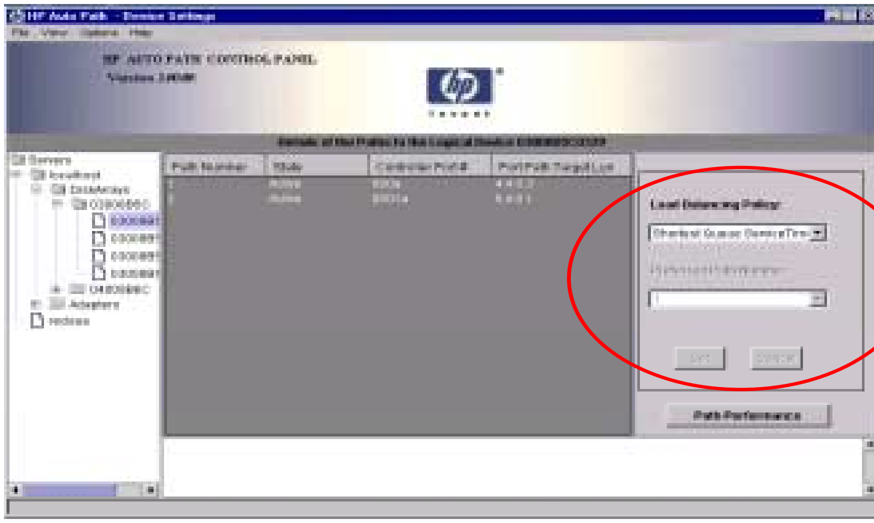
Student Notes

The Device Settings window shows the view for paths to a device and the load balance policy for the device. Access this window by selecting a device node in the tree structure.

The Device Settings window displays device information:

Path Number The index number for the path.
State The state of the path:
Active The path is in use.
Failed The path is no longer in use or all paths to the device have failed.
Controller Port # The port number of the controller on the disk array to which the device is connected.
Port.Path.Target.LUN Displays the physical path identifier.
Load Balancing Policy Displays the load balancing policy for the device. Use the dropdown list to choose from the list of policies. This is disabled if the device is not multipath-capable.
Preferred Path Number (*Displayed only when No_Load_Balancing is selected for Load Balancing Policy.*) Displays the serial number of the preferred path for the device. Select the preferred path to a device from the dropdown list.
Set Applies the changes and returns to the Devices window.
Cancel Cancels the operation and returns to the Devices window.
Path Performance Allows you to create and view performance graphs of the paths to the device.

Setting the load balancing policy



- Select a load balancing policy and click on Set

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

To set the load balancing policy:

1. From the navigation tree, select the server, disk array, and device that you want to set the
   load balancing policy for.
   The Device Settings window opens.
2. Select the load balancing policy from the Load Balancing Policy dropdown list.
   The new load balancing policy is displayed.
3. Click on **Set**.
   The load balancing policy is applied.

Load balancing policies

- No_Load_Balance
- Round_Robin
- Shortest_Queue_Requests
- Shortest_Queue_Bytes
- Shortest_Queue_ServiceTime

You were introduced to these policies in the first part of the presentation. Now you are at the point of choosing one and entering it

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

No_Load_Policy. No load balancing performed. All I/O is routed through a user-selected path called the Preferred Path. This load balancing policy does not affect path failover. If the path fails, the system uses another available, assigned path, without any preference.
Round_Robin. The requests are cycled though each path sequentially without regard to queue depth or length of time.
Shortest_Queue_Requests. Measures the current queue depth by counting the number of requests outstanding for each path.
Shortest_Queue_Bytes. Measures the current queue depth by the sum of the sizes of the I/O requests outstanding on that path.
Shortest_Queue_ServiceTime. (Default) Measures the current queue depth by the sum of the length of time each I/O request has been outstanding on that path.
"Shortest Queue" algorithms use instantaneous metrics, such as current queue depth, to select the path to be used.

Which load balance policy is best?

- The "best" policy is based on the processing environment, and that can change throughout the day
- First, check the Path Performance window
(we'll discuss that in a few minutes)
- There are three "shortest queue" policies, which use instantaneous metrics
- Try these:
  - For hit-intensive environments (for example, database with lots of I/Os) try Shortest_Queue_Requests
  - For data-size intensive environments (for example, when in contention with backup) try Shortest_Queue_Bytes
- Also, you might try scripting CLI commands to easily change policies at different times of day

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

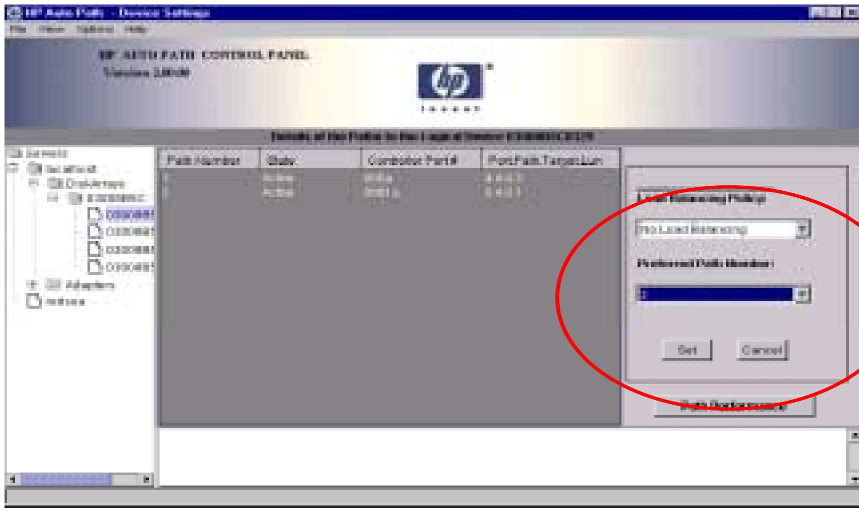
Student Notes

The load balancing policies (except Round_Robin) are based upon the current queue depths of all of the paths available to the device.

The difference between policies is how the queue depth is measured: queue requests, queue bytes, or queue service time.

All of the policies except Round_Robin shift the I/O load from heavily loaded paths or resources to more lightly loaded ones.

Setting the preferred path



- Select "No Load Balancing," enter a path and click on Set

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

To set the preferred path:

1. From the navigation tree, select the server, disk array, and device that you want to set the preferred path for.

   The Device Settings window opens.

2. Select No Load Balancing from the Load Balancing Policy dropdown list.

   The Preferred Path Number selection box is displayed.

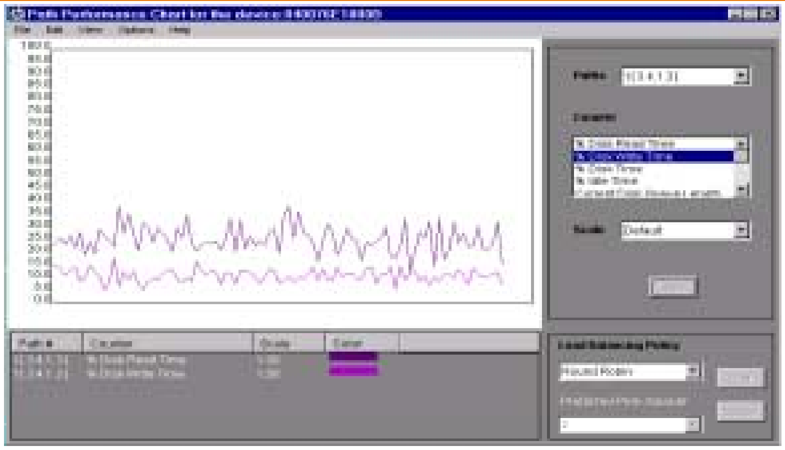
3. Select the preferred path from the Preferred Path Number dropdown menu.

   The new preferred path is displayed.

4. Click on **Set**.

   The preferred path is applied.

Path Performance window



- Allows you to create and view performance charts
- To see this window, click on Path Performance on the Device Settings window
- This is a separate window. To close it, select File/Exit

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Path Performance Chart

The y-axis shows the plot for a selected performance counter; the x-axis represents time.

Chart Parameters

Paths	Dropdown box that is used to select the index number of a path to be charted.
Counter	A list of available counters, measured on the y-axis of each chart. Choices are:
% Disk Read Time, %Disk Write Time, %Disk Time, %Idle Time
Disk Read/sec, Disk Writes/sec, Disk Transfers/sec
Current Disk Queue Length
Scale	Dropdown box used to select the scale of the chart.
Set	Starts the plotting of the chart using the selected parameters.

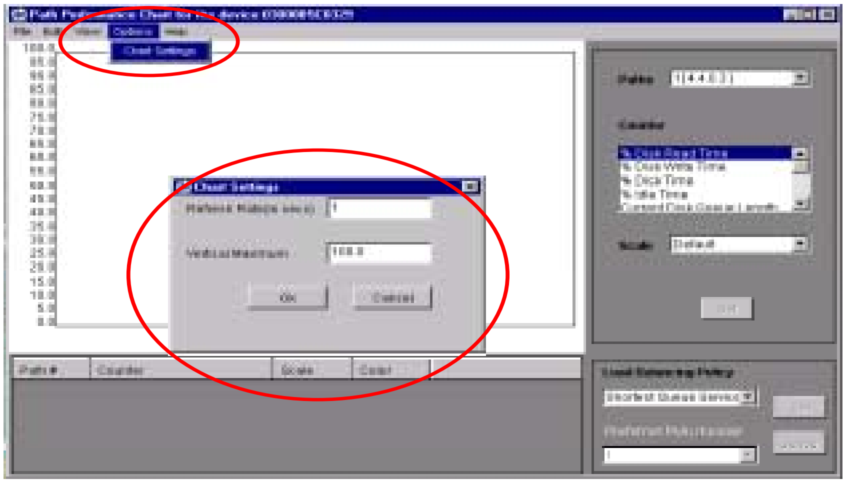
Chart Settings Information

Path #	The previously determined index number of the plotted path, followed by its SCSI address:
[Port.Path.Target.LUN].
Multiple plots, using different counters, can exist for the same path.
Counter	The selected counter of the plot.
Scale	The selected scale of the plot.

Color   The color of the plot.

When the Path Performance window opens, the Load Balancing Policy and Preferred Path Number fields are disabled.

Changing graph settings



- Select Options/Chart Settings
- Enter a new refresh rate and/or vertical maximum
- Click OK

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

To change the refresh rate or the vertical maximum of a chart:

1. Select Options/Chart Settings.

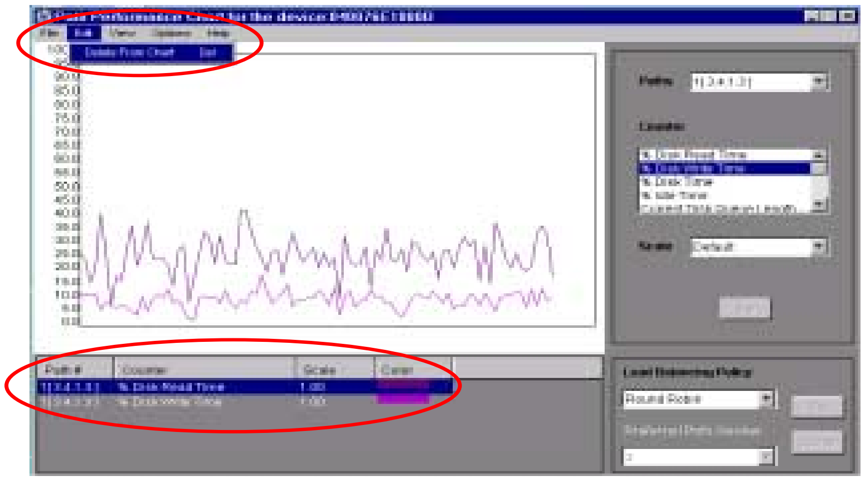
   The Chart Settings window opens.

2. Enter the new refresh rate and vertical maximum.

3. Click on **OK**.

   The new refresh rate and vertical maximum are applied.

Removing a plot From the chart



- Select the path number you want to remove
- Select Edit/Delete from Chart

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

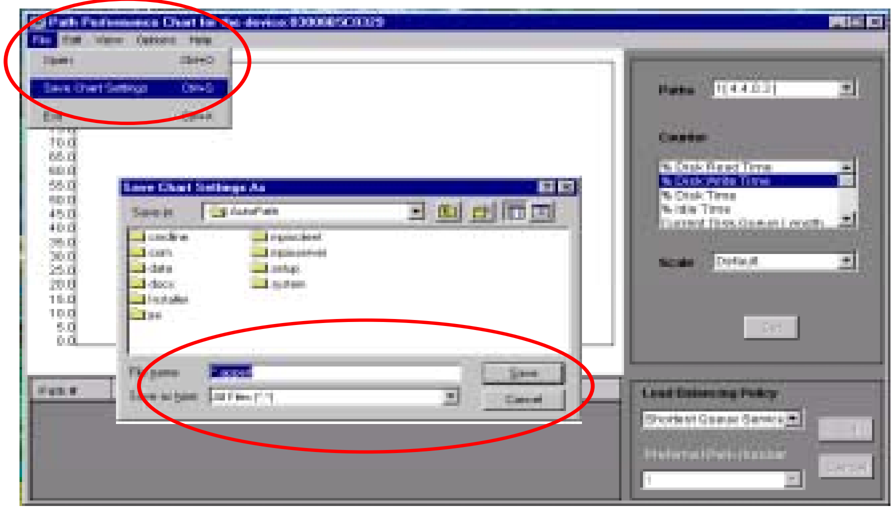
To remove a plot from the chart:

1. Select the Path # of the plot you want to delete.

2. Select Edit/Delete From Chart.

   The plot is removed from the chart.

Saving chart settings



- Select File / Save Chart Settings and name the settings file
- Click Save

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

To save chart settings:

1. Select File/Save Chart Settings.

   The Save Chart Settings As window appears.

2. Enter a filename with the extension **.apperf**.

3. Click on **Save**.

The chart settings are saved for future use.

Using Windows NT Performance Monitor

- Windows NT Performance Monitor also allows you to create and view performance charts for selected device paths
- HP Auto Path XP provides the information to Windows NT Performance Monitor via a special purpose DLL.
- The charts can be used to evaluate load balancing policies or preferred paths

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

To create a path performance chart using Performance Monitor:

1. Start Windows NT Performance Monitor by selecting Start/Programs/Administrative Tools/ Performance Monitor from the Windows Start menu.
   Performance Monitor opens.
2. Select Edit/Add to Chart.
   The Add to Chart window appears.
3. Select Object/HP Auto Path XP.
   The Instance box shows the paths to the devices on the disk arrays.
   The Counter box shows the list of counters. To view descriptions of the available counters, click Explain.
4. Select the paths and counters that you want to view the performance details for.
5. Select Add.
   The program begins plotting the chart.

Adapters window

- To see this window, select the adapters node in the navigation tree

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

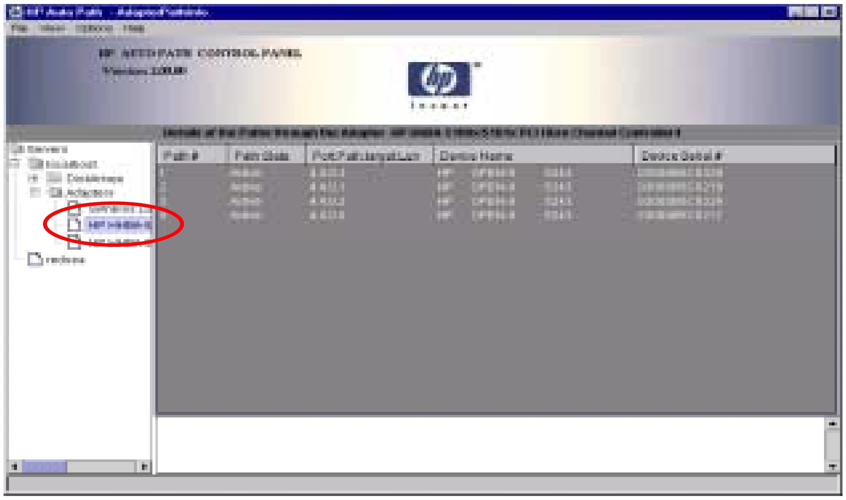
The Adapters window lists all the Auto Path-capable adapters on the system. This window is a view-only window; there are no user-configurable functions.

To access the Adapters window, select the Adapters node in the tree view.

The Adapters window displays the following adapter information:

Adapter # An index number used by Auto Path for identifying the adapter.
Adapter Name The name of the adapter.
Adapter State The state of the adapter:
Active The adapter is in use.
Failed The adapter is no longer in use.
# of Paths The number of paths through the adapter.
Adapter SCSI Port # The adapter SCSI port number.

Adapter Path Info window



- To see this window, select an adapter under the adapters node in the navigation tree

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

The Adapter Path Info window shows the details of the paths from the adapter. This window is a view-only window; there are no user-configurable functions.

To access the Adapter Path Info window, select an adapter under the Adapters node in the tree view.

The Adapter Path Info window displays path information:

Path # An index number used by Auto Path for identifying the path.
Path State The state of the path:
  Active The path is in use.
  Failed The path is no longer in use.
Port.Path.Target.LUN The physical path identifier.
Device Name The name of the device the path leads to.
Device Serial # The serial number of the device the path leads to.

Starting / Closing the Command Line User Interface

- Starting the CLUI
  - Open a DOS window
  - You can execute Auto Path XP command lines only from the directory where the installed Auto Path XP program files reside
- Closing the CLUI
  - exit the DOS window using the exit command

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

There are five CLUI commands, which we'll look at in the next slides. They are:

autopath help
autopath adapters
autopath devices
autopath path device
autopath set device

autopath help

- Opens the help for the CLI commands
- Syntax

  `autopath help`

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

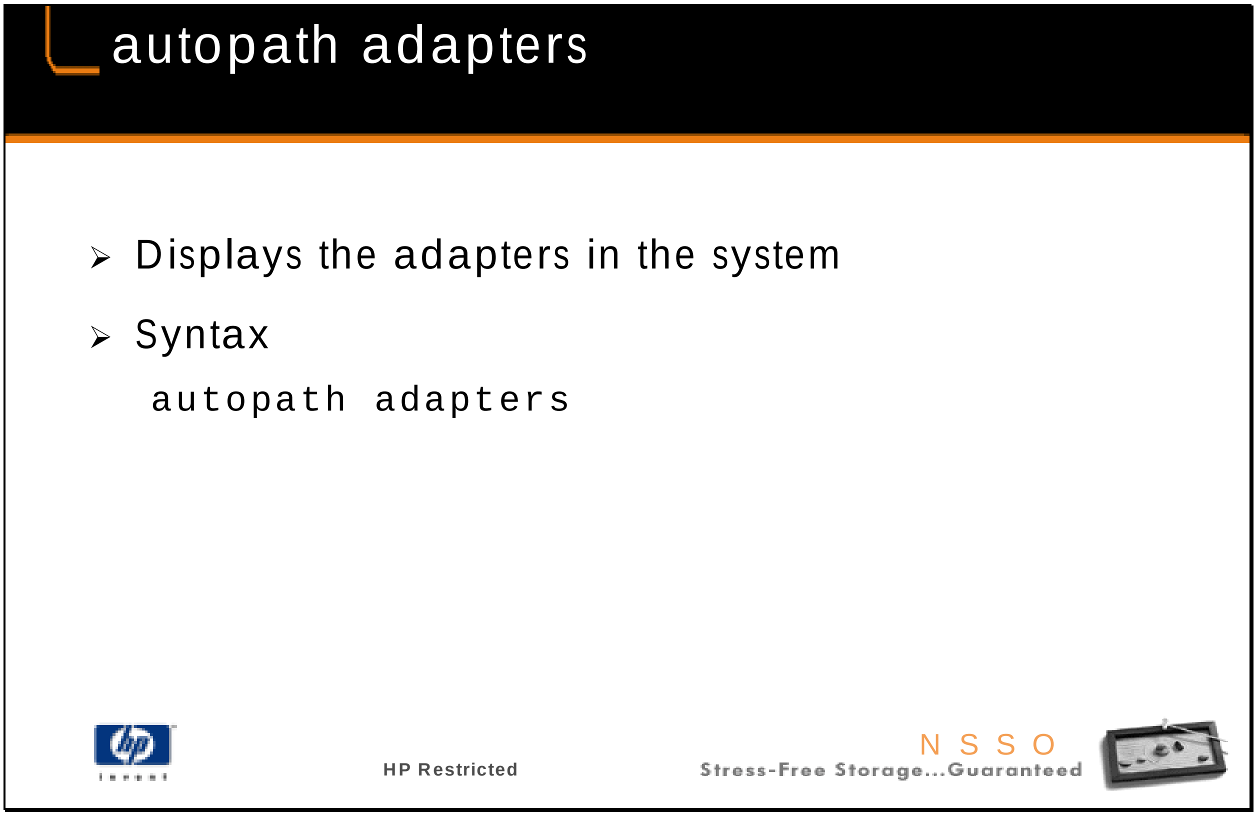
Description The **autopath help** command displays the help for the CLI commands.

Syntax **`autopath help`**

Example output

```
AUTOPATH DEVICES                                :     Displays details
of devices
AUTOPATH PATHS DEVICE=#                         :     Displays details
of paths to device#
                                 (* at the end of a path
signifies that
                                 it is the Preferred path)
AUTOPATH SET DEVICE=# PATH=#                    :     Sets preferred
path to device# as path#
AUTOPATH ADAPTERS                               :     Displays details
of adapters
AUTOPATH SET DEVICE=# POLICY=<POLICY_NAME> : Sets load balance
policy
                                             for device # to
```

```
<policy_name>
SUPPORTED LOAD BALANCE POLICIES          :     No_Load_Balance
           DefaultBalancePolicy
           Round_Robin
           Shortest_Queue_Requests
           Shortest_Queue_Bytes
           Shortest_Queue_ServiceTime
AUTOPATH HELP                            :     Displays help
```



Student Notes

Description The **autopath adapters** command is used to display information about the adapters in the system, including name, state and number of paths.

Syntax **`autopath adapters`**

Example output

```
 Adapter#  Port#              Adapter Name                          State    #
Paths
-------------------------------------------------------------------------------
------
    1        3    Agilent PCI Fibre Channel Controller (NT 4.0)     Active
4
    2        4    Agilent PCI Fibre Channel Controller (NT 4.0)     Active
4
```

autopath devices

- Displays the details of the multipath-capable devices in the system
- Syntax

  `autopath devices`

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

Description The **autopath devices** command displays the details of the multipath-capable devices in the system, including name, serial number, state, number of paths and the current load balancing policy.

Syntax **`autopath devices`**

Example output

```
 Device#     Device Name  Serial No.     State  # Paths    Policy
--------------------------------------------------------------------------------
---
   1   HP    OPEN-3      0400756000E2    Active    2
Shortest_Queue_ServiceTime
   2   HP    OPEN-3      0400756000E3    Active    2    Round_Robin
```

autopath paths device

- Display the paths to the selected device referenced by the device number
- Syntax

  `autopath paths device=`*`device_number`*

  where *device_number* is the Auto Path index number shown in the autopath devices **command line display**

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Student Notes

Description The **autopath paths device** command displays the paths to the selected device
referenced by the device number. It also shows the device name and serial number.

Syntax **`autopath paths device=`***`device_number`*

device_number The Auto Path index number shown in the **autopath devices** command line display

Example output

```
autopath paths device=1
Device Name: HP      OPEN-3
Device Serial Number: 0400756000E2

 Path#    Controller Port#      State     Scsi Port#  Bus#  TID  Lun#
-----------------------------------------------------------------------
   1             6              Active        3         4     0     0
   2            22              Active        3         4     1     0
```

autopath set device

- Set the load balance policy for the device specified by the device number
  - If "no load balancing" is set for a device, the command sets the preferred path to the device
- Syntax

```
autopath set device=device_number
  { policy=policy_name |
  path=path_number }
```

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Description The **autopath set device** command sets the load balance policy for the device specified by the device number.

If "no load balancing" is set for a device, the **autopath set device** command sets the preferred path to the device.

Syntax **autopath set device=***device_number* { **policy=***policy_name* | **path=***path_number* }

device_number The Auto Path index number shown in the **autopath devices** command line display
policy_name The load balancing policy name
path_number The preferred path number

Example output

```
    Device Name: HP      OPEN-3
Device Serial No. : 0400756000E2
Change Load Balance policy (y/n) ? y
Load Balance policy changed. Verify by issuing "autopath" command
```

autopath set device=1 path=2

```
Device Name: HP      OPEN-3
Device Serial No.   : 0400756000E2
```

```
Controller Port No. : 22
Path Scsi Address   : 3.4.1.0
Change Preferred Path (y/n) ? y
Preferred Path changed. Verify by issuing "autopath" command
```

Troubleshooting

- Troubleshooting should be more about failed paths than a failure of Auto Path
- Right after installation, perform tests to be assured that Auto Path is running correctly:
  - Run the autopath adapters and autopath devices commands
  - Set “no load balancing” with a preferred path
  - Set all paths
  - Use the Path Performance window
- Degraded performance during peak utilization is likely to be the first direct sign to a user that a path has failed
  - Use the Path Performance window
  - Make frequent inspections of the log

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

It is unlikely that Auto Path will install or run incorrectly. Troubleshooting should be devoted primarily to finding and correcting fail path conditions.
Some problem resolution can be performed by the customer. For example, customers typically replace failed GBICs on some HBAs. HBA replacement is usually a CE task.

Windows NT Troubleshooting

- Event Viewer
  - The viewer logs when a path is taken out of service, all disk and adapter errors, and I/O errors.
- Windows Diagnostics
  - The Administrative Tools utility shows all devices connected to the system.
- Windows Registry
  - Shows the devices connected to the Windows NT system. Use the Registry Editor to display the devices seen at boot time.
- Disk Administrator
  - Shows the paths that are online and offline. If all the paths to a LUN appear to be offline, it is likely that the device is not preformatted at the disk array level. Preformat the device, and the Disk Administrator will see the path of the device and the remaining paths offline.

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

To troubleshoot Auto Path, you can use these tools, as described in the slide:

Event viewer
Windows diagnostics (Administrative Tools utility)
Windows registry, using the Registry Editor
Disk Administrator

Caution *Back up the Registry before executing the Registry Editor.*

Recovering from a path failure

- Determine that a path has failed
- Diagnose and correct the failure
- Run Microsoft NT Disk Administrator or wait 60 seconds
- Display paths to be sure the path is active
- If a path was the preferred path, you must reset it manually on the Device Settings window

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

After you restore the failed path, run Windows NT Disk Administrator or wait for 60 seconds. Auto Path automatically begins using the restored path, following the designated load balancing policy.

If there is a policy of “no load balancing” for the failed path and the path was not the preferred path, no user action is necessary. However, Auto Path monitors all paths, even unused paths. You can still use the GUI to determine that there has been a path failure and take corrective action.

If failed path was the designated preferred path, the system will failover to another path. However, after you recover the failed path, you should then reset it as the preferred path. Use the Device Settings window.

Known problems and workarounds

- You need identical HBAs and drivers on a server
- You need the same type servers and HBAs on a cluster
- Path performance graphs in PerfMon may show higher than 100% I/O throughput in a few cases
- Some Auto Path messages do not get logged to the system log during system bootup
- If LUNs are removed from the SVP and you run Disk Administrator, the disk subsystem appears as a new device with an existing path number
- Over a slow network, client access to a remote server may fail

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

1. On any one server, all Host Bus Adapters must be of the same type. At this point, there is only one supported HBA.
2. In a cluster, all the servers must be of the same type and all Host Bus Adapters on each server must of the same type.
3. The path performance graphs in PerfMon may show greater than 100% I/O throughput in a few cases.
4. During system bootup, some Auto Path messages do not get logged in the system event log.
5. If LUNs are removed from the SVP and Disk Administrator is run, the disk subsystem appears as a new device with an existing path number.
6. Over a slow network, client access to a remote server may fail.

Auto Path XP for Windows 2000 Product Ordering Structure

- A structure similar to the XP512/48 software structure, i.e.
  - media product that provides array connectivity
  - licenses based on number of servers with Auto Path installed
- B9500A Auto Path for W2K/Pentium Media
- B9501A Auto Path for W2K/Pentium 1 Server LTU
- B9502A Auto Path for W2K/Pentium 5 Server LTU
- B9503A Auto Path for W2K/Pentium 10 Server LTU
- B9504A Auto Path for W2K/Pentium 25 Server LTU

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Auto Path XP for Windows 2000 Product License Details

- License allows one Windows 2000/Pentium server to connect to one or more supported arrays
- Enterprise license allows connectivity to all arrays supported by Auto Path
    - XP256 and XP512 today, other enterprise arrays in the future
    - other current and future departmental arrays
- Structure does not have server dependency for licensing classification. Therefore, structure is independent of new server introductions and subsequent classification into various tiers
- Multi-server license eliminates price as an inhibitor in large scale consolidations - equivalent of a "site" license

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Auto Path XP
Product Ordering Benefits

- Provides similar benefits as existing 512 software structure
  - Increments priced on sliding scale: larger increments (more servers) at a lower cost/server than smaller increments
  - Allows for competitive small configuration and low-end platform pricing
  - Allows customers to choose between short-term, minimum-cost strategy or long-term, license-for-value strategy
- Provides similar Selling Points/Strategy
  - Lower priced licensing for both entry configurations and high end configurations
  - More value to customer with larger increment licenses
  - Plan consolidation requirements with customer to identify best long-term licensing strategy

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Auto Path XP is a server based product, and is licensed per server.

Auto Path XP for Windows 2000 Pricing

Prod #	Description	Discount over 1 server	List Price	List Price / server *
B9500A	Auto Path XP for W2K Media	N/a	0	0
B9501A	Auto Path XP for W2K 1 server LTU	N/a	$5,995	$5,995
B9502A	Auto Path XP for W2K 5 server LTU	35%	19,500	3,900
B9503A	Auto Path XP for W2K 10 server LTU	50%	29,975	2,998
B9504A	Auto Path XP for W2K 25 server LTU	60%	59,975	2,399

Actual cost depends on user discount

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

1, 5, 10 and 25 server License products are priced on a sliding scale such that larger increment licenses are less expensive on a per server basis than small increment licenses. As a result, a single larger increment license may be significantly less expensive that multiple smaller increment licenses for a given number of servers.

Review of Key Points

- Auto Path XP Definition & Features
- Auto Path XP's Filter Driver
- Auto Path XP & Disk Administrator
- Auto Path XP Failure/Recovery
- Auto Path XP Performance/Load Balancing
- Auto Path XP Configuring & Troubleshooting
- Futures
- Ordering

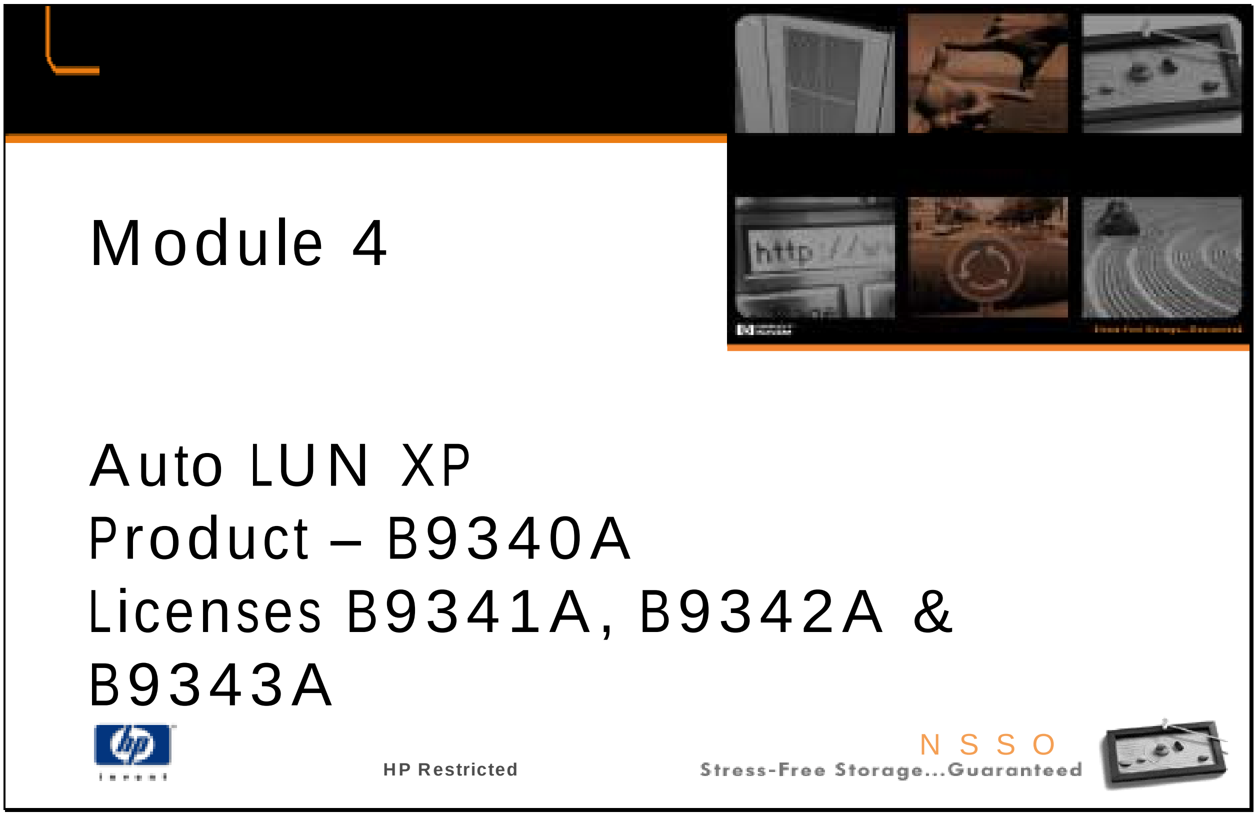
HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Module Wrap-up

HP Restricted

N S S O
Stress-Free Storage...Guaranteed



Student Notes

Auto LUN allows you to configure the disk array to monitor disk volume performance and auto-migrate the data. You can create a migration plan using a script-like interface to define the migration conditions. Then, based on this plan, the array can move the data and path automatically whenever a threshold limit is crossed. There is also a manual migration option.

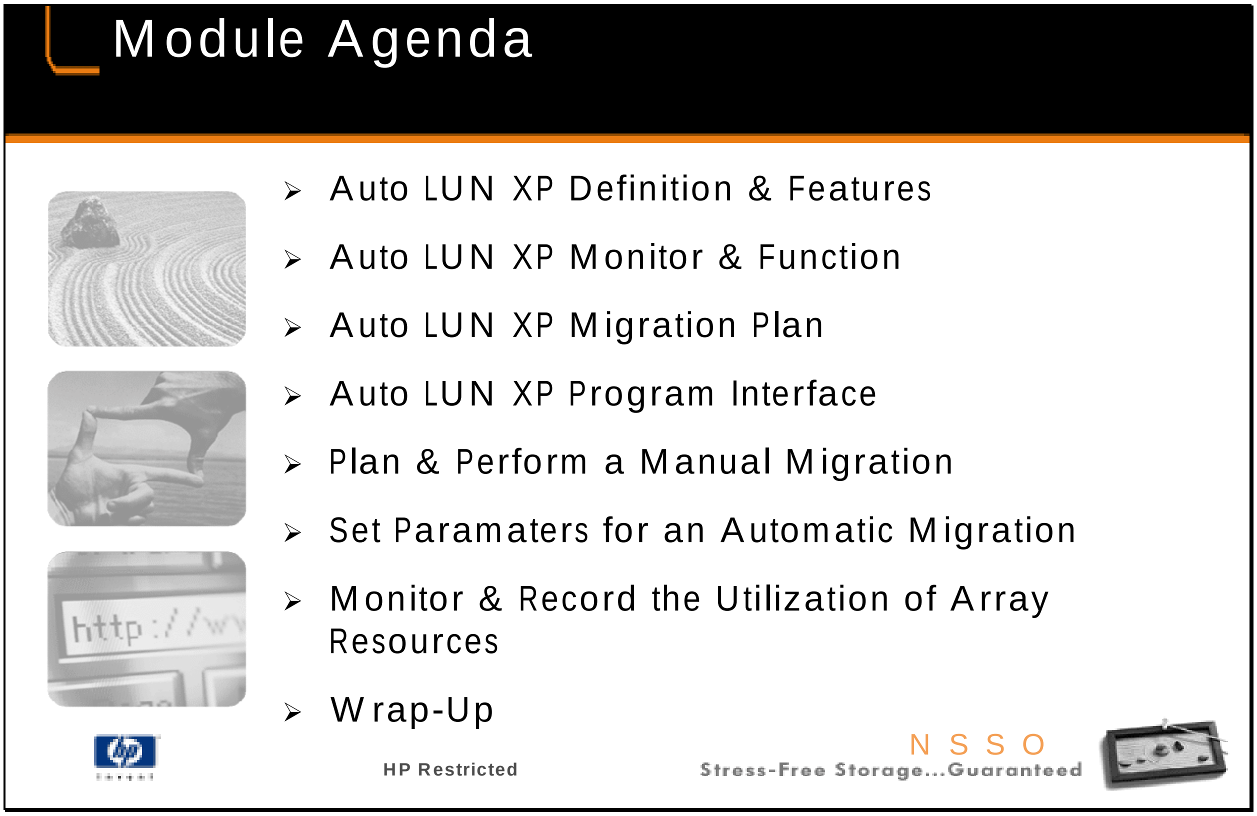
Auto LUN Operations

Disk Utilization Displays parity groups and logical volume utilization. Manual migration also takes place using these screens.

Hardware Utilization Displays CHP, DKP, MBUS and DRR utilization.

Auto Migration Displays utilization values for different logical volumes based on a migration plan. You can set threshold parameters for automatic migration of LDEVs depending on their utilization. View Configuration - Allows a user to start or stop resource monitoring, and to set the time period for viewing the information.

History Log - Shows logs of automatic migration and LDEV migration.



HP SureStore E Auto LUN XP (Auto LUN) enables you to optimize your data storage and retrieval on the disk array. Auto LUN monitors and provides detailed information on the physical disk drive usage of the disk array and enables you to optimize the logical volume allocation and RAID level configuration (RAID-1, RAID-5) of the disk array. Auto LUN operations are preformed on the Remote Console PC. The Auto LUN Remote Control software enables you to perform Auto LUN operations on the disk arrays attached to the remote console PC. The remote console PC is attached to the disk array by an internal LAN and communicates directly with the disk array service processor (SVP). The Auto LUN Remote Control also maintains and displays a history log of all Auto LUN operations. The combination of RAID level and physical drive type optimizes disk array performance for your environment.
AutoLUN XP is an HP OEM'd software product developed by Hitachi Data Systems marketed under the name Hitachi Internal Hierarchical Storage Management (HIHSM).
Auto LUN allows you to monitor disk array usage and performs automatic internal migration of logical devices (LDEVs) in efforts to optimize disk array performance. Auto LUN allows you to select the time period for the disk array monitoring, reserve LDEVs for Auto LUN target migration and specify criteria for automatic migration to occur.

Auto LUN runs on the Service Processor (SVP) found inside of the Disk Control Unit (DKC); Auto LUN's software functionality can be accessed by the user via the Remote Control PC. Like many other XP software packages, Auto LUN must be key activated from diskette on both the Remote Console and the SVP in the DKC. Before Auto LUN can run on the SVP, the DKC Main must be at microcode v43.12.00 or higher and the Remote Console must be at v10.12.00 or higher.

Auto LUN XP Definition

- *Disk array software that assists in balancing capacity with performance using the following four functions....*
  - **Monitor Function**: monitors and displays utilization of disk array resources over defined time interval
  - **Migration Function**: migration of volumes from source parity group to target parity group
  - **Estimate Function**: estimates utilization rate of the source and destination parity group's utilization assuming a volume migration were to be completed
  - **Preset Function**: automatic migration of volume(s) from one parity group to another parity group

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

AutoLUN XP is an HP OEM'd software product developed by Hitachi Data Systems marketed under the name Hitachi Internal Hierarchical Storage Management (HIHSM).

What does Auto LUN do? Auto LUN allows you to monitor disk array usage and perform either manual or automatic internal migration of volumes (CU:LDEV) in efforts to optimize disk array performance. Auto LUN collects performance data, allows the user to granulate the time period for the disk array monitoring data to be displayed, reserve volumes to serve as targets for volume migration and perform an automatic or manual migration of volumes from one parity group to another.

Auto LUN runs on the Service Processor (SVP) found inside of the Disk Control Unit (DKC); Auto LUN's software functionality can be accessed by the user via the Remote Control PC or in Command View (CV only applicable to the XP512). Like many other XP software packages, Auto LUN must be key enabled.
Auto LUN enables you to monitor disk array usage and plan and perform volume migration based on this usage data. Auto LUN provides both manual volume migration and automatic, or preset, volume migration capabilities.
• For manual operations, the user-specified volume migrations are performed immediately.

• For automatic operations, Auto LUN uses user-specified criteria to set the migration schedule and execute automatic volume migrations.
Auto LUN operations include functions that enable you to view the utilization of disk array resources, test logical volume reallocation plans, and reallocate the logical volumes on the disk array.
• The Monitor function monitors and displays the utilization of disk array resources.
• The Estimate function enables you to estimate the utilization of parity groups after a proposed Auto LUN migration of logical volumes.
• The Volume Moving (Migration) function allows you to move logical volumes to specified parity groups.
• The Preset function creates a migration plan based on the information you enter and then moves the logical volumes automatically according to your specified migration plan.

The HP SureStore E AutoLUN XP (AutoLUN) product enables you to optimize your data mangement and retrieval on the disk array. AutoLUN monitors and provides detailed information on the physical disk drive usage of the disk array and enables you to optimize the logical volume allocation and RAID level configuration (RAID-1, RAID-5) of the disk array.

Auto LUN enables you to monitor disk array usage, and plan and perform volume migration based on this usage data. Auto LUN provides both manual volume migration and automatic, or preset, volume migration capabilities.

For manual operations, the user-specified volume migrations are performed immediately. For automatic operations, Auto LUN uses user-specified criteria to set the migration schedule and execute automatic volume migrations. Auto LUN operations include the following functions which enable you to view the utilization of disk array resources, test logical volume reallocation plans, and reallocate the logical volumes on the disk array. The monitor function monitors and displays the utilization of disk array resources. The estimate function enables you to estimate the utilization of parity groups after a proposed AutoLUN migration of logical volumes. The Volume Moving (migration) function allows you to move logical volumes to specified parity groups. The Preset function creates a migration plan based on the information you enter, and then moves the logical volumes automatically according to your specified migration plan.

Auto LUN XP Monitor Function

- Utilization information is collected by the SVP from the DKC once every 24 hours (or on demand from the Remote Console) for the following components:
  - CHIP and ACP microprocessors (CHA & DKA)
  - XP512= Hierarchical Star Net (i.e.- Crossbar switch)
  - Data Recovery Reconstructor (DRR)
  - Volumes- asynchronous and synchronous usage time of each logical device in a parity group averaged by the number of HDD in the parity group

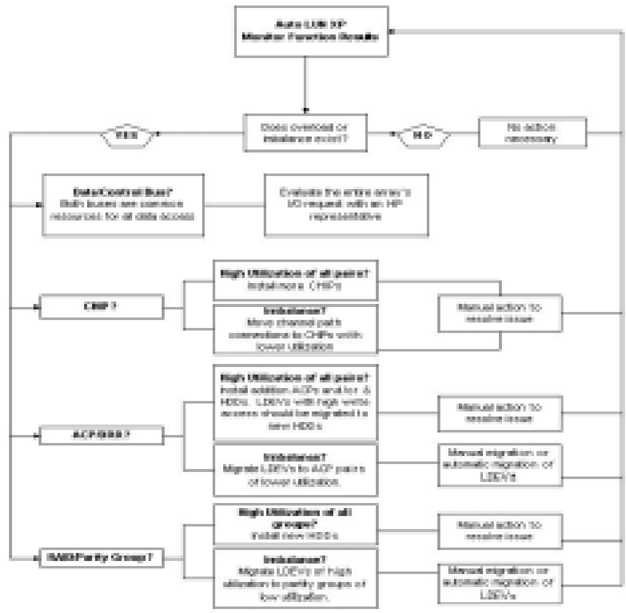
    (LDEV utilization= total use time/number HDD's)
  - Parity Groups- the average and maximum measured usage rages of all logical volumes in a parity group

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

AutoLUN monitoring enables you to evaluate disk array resource usage and determine whether resources are overloaded or out of balance. Auto LUN monitoring displays usage information for the following disk array resources:
CHIP and ACP microprocessors, F-buses and M-buses, DRRs, logical volumes
Usage (time in use) of each logical volume in a parity group, including synchronous and asynchronous access, averaged by the number of physical drives in the parity groups.
Usage (time in use) of all logical volumes in a parity group. The parity group usage is the sum of the logical volume usages for all volumes in the parity group.
Auto LUN monitor data is collected by the SVP once a day automatically and/or manually as configured and requested by the user. The SVP stores te last three months (90 days) of Auto LUN monitor information on its hard disk drive. The Auto LUN Remote Control allows you to configure, start, and stop the monitoring and collection of disk array usage data. AutoLUN allows you to specify the monitor data term (range of data) to be displayed.
The information collected daily will be stored on the SVP's hard disk for three months; if data is older than three months Auto LUN will overwrite the oldest data with the newest data. FYI- After connecting to an array via the Remote Control XP software, hitting the AUTO LUN button will ask, "Do you want to collect all statistics data now?" Answering yes to this question will download all measure data values from the SVP to the RMC for analysis within Auto LUN

Auto LUN Monitor Results Flow Chart



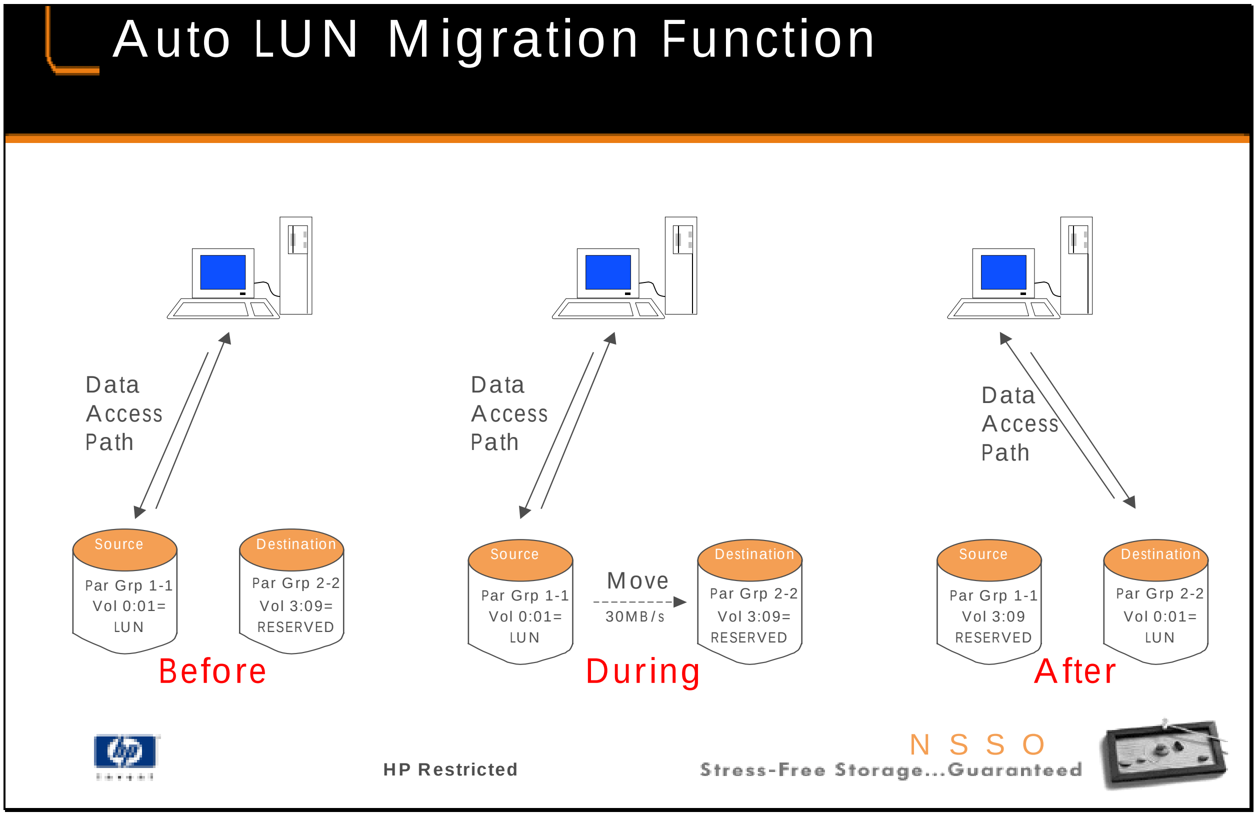
HP Restricted

SSO Stress-Free Storage...Guaranteed

The suggestions provided in this flow chart do not represent all case scenarios for reports of high utilization or imbalances. The flow chart provides high level guidelines and serves as a good starting point when interpreting the Auto LUN monitor function output. Auto LUN tuning methods should only be applied after thorough analysis and under the prospect of large improvement.

It must be understood that when you alter for performance the cost of the performance gain is not free. For example, Parity group 1-1 has a utilization rate of 20% and Parity group 3-1 has a utilization rate of 90%. After migrating a logical volume from 3-1 to 1-1 (i.e.- you are attempting to lower the utilization rate of 3-1 since it is currently 90%), RAID group 1-1 may now have a utilization rate of 55% which represents a 35% decrease in performance for parity group 3-1.

Since all I/O transactions utilize the data and control buses, volume migration will not resolve existing bottle neck issues on the buses. The proper course of action for high utilization of the buses is to evaluate what is being requested from the array and make modifications accordingly (I.e.- offload some applications or data to a different array; adjust time/variance request of data by hosts).



Migration will move data from a source Volume to a pre-defined RESERVED volume. Migration of a volumes occurs manually or via Automatic Migration using the same algorithm as Business Copy.
Destination volumes must be RESERVED in advance; setting the destination volume to RESERVED in advance prevents other hosts from accessing the volume.
(FYI- With HP's first release of AutoLUN, LUN Configuration Manager allows end users to assign paths to RESERVED LDEV's; all host I/Os to the RESERVED LDEV would receive failures). Volumes that have assigned LUN numbers cannot be defined as RESERVED by Auto LUN.

Once the data has been migrated, the source volume now becomes RESERVED. During the migration, all I/O still targeted for the SOURCE volume is maintained in CACHE until the migration has completed; if cache is unavailable, all writes to the source are written to the source volume and tracked in a bitmap. Once all original source information is migrated, the delta is migrated to the target as well (all I/O request from the host will receive a "busy state" message until the delta copy process has completed).

Concept: Notice that volume 0:01 residing in parity group 1-1 migrates to parity group 2-2 and the RESERVED attribute assigned to volume 3:09 in parity group 2-2 moved from parity group 2-2 to 1-1. Once the RESERVED attribute to a defined volume, the total number of RESERVED volumes available does not decrease by one following a volume migration. Instead, the total number of RESERVED remains the same. Both during and after the migration, the LUN number does not change (thus volume migrations during and after are transparent to connected host)

Auto LUN Migration Function

The following apply to both manual and automatic migrations

Both source and target Volumes must….

- have the same emulation type (Open 3, Open8, etc…)
- be the same size (LUSE or CVS)
- not be blocked
- be connected to the same DKC
- not be currently migrating or scheduled to migrate
- not be CA or BC pair(XP256 Only)
- not be a command device
- not be a cache LUN
- The defined RESERVED (target volume) must not be assigned a LUN number

Migrations supported…

- Volumes in a simplex state
- Same RAID level
  - ex) RAID1 to RAID1
- Different RAID levels
  - ex) RAID5 migrated to RAID1
  - ex) RAID1 migrated to RAID5
- Same drive types
  - ex) 15GB to 15GB
- Different drive types
  - ex) 15GB to 47GB
- VSC & LUSE volumes of equal size

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Execution of volume migrations can occur automatically or manually. With automatic migration, Auto LUN monitors the utilization of LDEV's and will migrate a volume if its utilization exceeds the a pre-defined threshold (explained in detail later). With manual migration, the end user selects the source volume to migrate and a RESERVED destination volume to receive the migrated volume. In either event, the migration is transparent to the host(s) (i.e.- the LUN number and CU:LDEV numbers do not change).

If the volume happens to be one component of a LUSE volume, the volume(s) that exceeds the threshold will be migrated to the RESERVED volume. Auto LUN does not allow you to RESERVE a group of volumes for migration of an entire volume collection; for this reason it becomes possible for a contiguously LUSE defined volumes to become dispersed (non-contiguous) after Auto LUN migration(s). Migration of LUSE volumes requires the same capacity and same emulation type on the source and target.

VSC volumes can also be migrated, but only to target volumes with the same capacity and emulation type as the source.

To limit the impact on performance, it is recommended to perform migration during off-peak hours.

Auto LUN Manual Migrations

- Manual migrations allow the end user to define both the source and target volumes for migration
  - Must be manually started (i.e.- there is not a method to "schedule" a manual migration to start at 2:00am)
  - If the migration fails, Auto LUN does not have an automatic notification event generated (i.e.- R-SIM or SNMP trap)
  - Only one volume can be migrated at a time

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

The first challenge with performing a manual migration is starting the migration during off-peak hours. The second challenge is if more than one migration needs to be performed, the user must be present to define and manually engage the start of each volume migration.

HINT: This slide set has not covered this topic yet (coming in slides to follow), but ensure you know in advance the parity group that contain RESERVED volumes and the emulation defined for those parity group (i.e. OPEN3 or OPEN8). When selecting a target for the volume migration, having this information in advance will reduce the time spent completing the task.

Auto LUN Manual Migration Steps

- Start Auto LUN's Monitoring Function
- Set target volume to "RESERVED"
- Check the History for past automatic and manual migrations (does a trend already exist?)
- Check to see if a migration is in progress
- Define data collection term of data to be displayed (basis for what volume to migrate)
- Define volume to migrate
- Initiate volume migration
- Monitor the migration status

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

NOTE: Manual migration is not supported during the execution of automatic migrations. Auto LUN can only perform one migration process at a time (regardless if the migration is performed manually or via preset). Always check the status of volumes to see if a migration is in progress before beginning a manual migration (as mentioned previous an * will appear in the CU field if migration is in progress).

Migrations (both manual and preset) can be cancelled during the migration process. If migration is cancelled by the user, an outage (or because the defined interval of a preset migration has been meet) all delta information stored in the cache bit-map is written to the source volume.

Auto LUN Automatic Migrations

- Automatic migrations allow the end user to define parity group thresholds to trigger a volume migration. If a volume within a parity group exceeds that defined threshold, the volume will be candidate for migration
  - Migration will only occur if Auto LUN's estimate function determines a performance gain will be yielded
  - Unlike manual migrations, automatic migration can schedule 1 to 32 volume migrations inside of one migration plan (occurs in sequential order)
  - Unlike manual migrations, the end user does not have as much granular control over which RESERVED volume will be defined to receive the volume migration

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

If automatic migration migrates 5 volumes, and prior to the migration I had 15 volumes defined as RESERVED (assuming everything is the same emulation type), which of the 15 RESERVED volumes received 5 volume migrations?
You don't know until you check the HISTORY log files. Automatic migration would have placed the 5 volume migration on the RESERVED's that were estimated to have the most performance gain for the volume(s) in question.

Auto LUN Automatic Migration Plan

- Plan creation is primarily derived from the following:

A. HDD specifications (a.k.a. Class)

DKS2A-KO18FC – 15K 18GB Drive –	Class A
DKR2B-JO18FC – 10K 18GB Drive –	Class B
DKR1B-JO47FC – 10K 47GB Drive –	Class C
DKR1B-JO72FC – 10K 72GB Drive –	Class D

> A FC 47GB drive will have slower seek time than a FC 18GB drive, even though both spin at 10,200 rpm

B. Parity and Volume utilization rates

- Historical statistic information
- Projected performance estimate calculated for RESERVED target volumes and parity groups

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Automatic - Auto LUN migration operations perform disk array tuning automatically, based on parameters entered by the user. Automatic AutoLUN migration is intended to function as the primary tuning method for the disk array.

AutoLUN can be done by following these steps repeatedly:

1.) Monitoring information

2.) Analyzing information

3.) Making volume migration plan (decision of volume migration)

4.) Moving volume (migration)

5.) Monitoring information again to confirm condition and effect of the performed tuning.
Migration plans can fail for the following reasons:
1. Auto LUN could not use monitored information from previous plan.
2. If estimated utilization rate of source parity group is over the defined maximum utilization threshold.
3. Auto LUN may fail during volume migration if the CACHE or drive replacement, installation or de-installation is executed.

If a migration plan fails, that migration will be re-scheduled to occur at the start of the next day's migration interval. Please reference slide "Auto LUN Preset Function" for more details.

Disk Type	Order (Class)
DKS2A	0 (Class A)
DKR2C	1 (Class B)
DKR2B	2 (Class C)
DKR1C	3 (Class D)
DKR1B	4 (Class E)

Auto LUN Automatic Migration Plan

- Plan creation primarily derived from the following (cont.)

C. Reserved Volume RAID definition

Is total random rate greater than total sequential rate — No → Raid5

Yes ↓

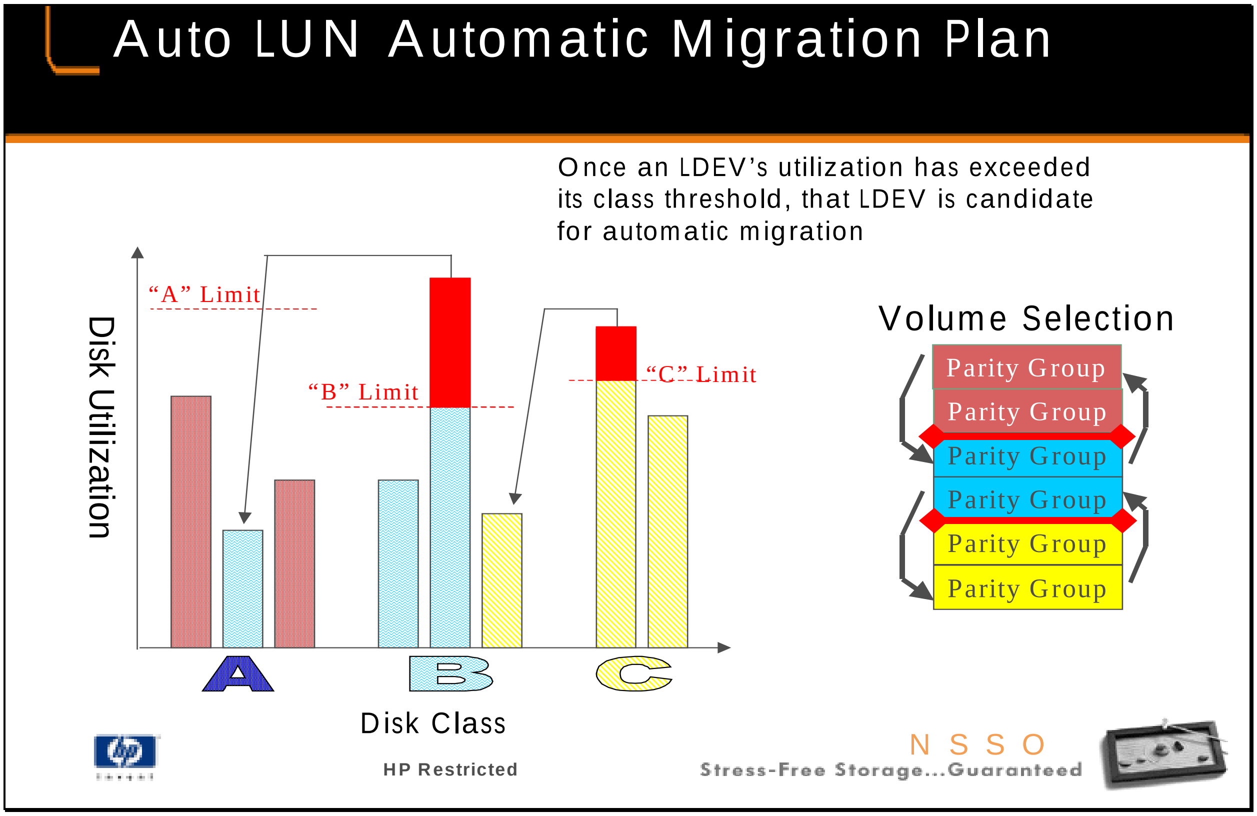
Is random read rate greater than random write sequential rate — No → Raid1

Yes ↓

Raid1 or Raid5

HP Restricted

NSSO
Stress-Free Storage...Guaranteed



Users have the option to change the default assigned maximum percentage of HDD disk utilization for each class; different utilization limits can be set for each HDD class (class B can have a different utilization threshold than C). When a volume exceeds this threshold, it acts as a trigger for Auto LUN's automatic migration function to create a migration plan at the end of AutoLUN's monitoring interval (thresholds only apply to automatic migrations). Volumes of high disk utilization that exceed the defined maximum limit become candidates for automatic migration.

Volumes within any class that exceed the threshold can be automatically migrated to any of the RESEVED volumes, but only if Auto LUN estimates that the migration will yield higher performance. When available RESERVED Volumes run out in the highest class A, parity groups of *least* utilization in the A class are selected by the estimate function to be moved one class lower in order to create more available high class RESERVE volumes. This same concept applies for medium class to low class volumes. See "Volume Selection" diagram above on the right side of the slide.

Auto LUN Estimate Function

- Using information collected via the monitor function, Auto LUN displays parity group estimates utilization rates of...
  - The parity group that contains the volume that will serve as the source. For example, if volume 0:50 were to be migrated out of 2-1, then the parity group 2-1 utilization would decrease from 67% to 44%.
  - The parity group that will serve as the target (remember, the parity group must contain a RESERVE before a migration can be accepted). For example, the parity group to accept 0:50 would increase from 12% to 53%.
- **Manual migrations?** End users can view Auto LUN's estimates for potential target parity groups by following the steps of a manual migration.
- **Automatic migrations?** Auto LUN automatically performs all calculations and decision analysis based on performance increase to be gained.

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Concept: Auto LUN's estimate function applies to both manual and automatic migrations. The difference is with manual migration the end user can see the estimate values generated by the estimate function. This can be helpful when determining if and when you might wish to setup automatic migrations (i.e.- use the manual migration function's use of the estimate function to see if the estimate function calculates a proposed migration will yield a performance gain; if it does, then you can assume an automatic migration could occur successfully).

Auto LUN Automatic Migration Preset Function

➢ How the Preset function works

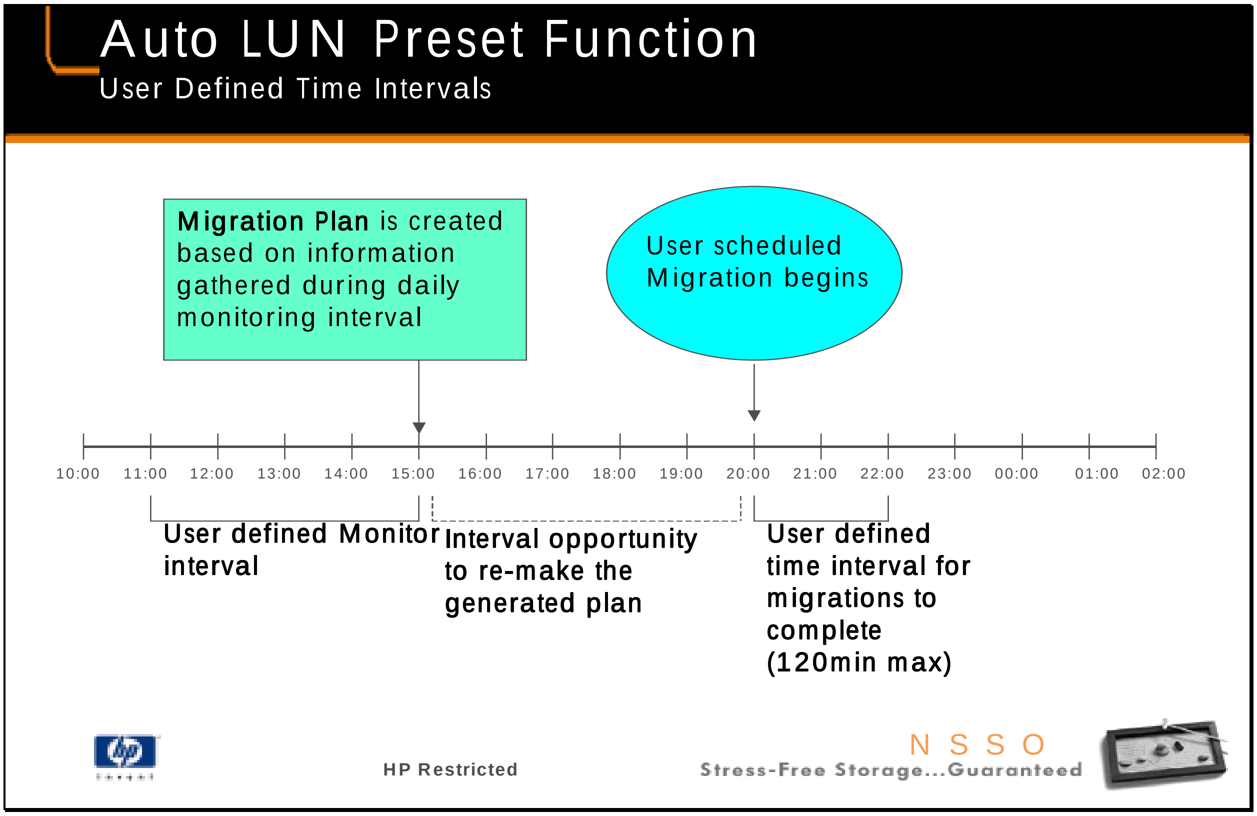
- Based on a user defined time interval for collection of monitored information, Auto LUN creates a volume migration plan. Starting at a user defined volume migration start time, Auto LUN starts the migration plan. This process can be repeated by a user defined cycle (daily, weekly, monthly etc).
  - Auto LUN can perform a predefined migration plan up to one time per day on the XP256 and two times per day for the XP512.
  - Users can re-make the plan created by Auto LUN by changing automatic migration defined criteria.
  - Auto LUN accepts up to 40 migrations in one plan; each migration will occur sequentially within that plan.

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Auto LUN automatic migration supports moving up to 40 volumes (executed sequentially; this means Auto LUN can have one plan containing up to 40 migrations). The maximum allowed time for this plan to complete is 120 minutes.

The next slide helps explain the Preset Function using a graphical timeline illustration.



Decisions that must be made by the user before implementing the Preset Functions to engage automatic migrations:
1. Set a time for Auto LUN to monitor. The information gathered during this time interval will be the basis for Auto LUN's migration plan generation.
2. Can the plan generated by Auto LUN be manually modified by the end user? No. However, the end user can force the creation of a new migration plan that replaces the current plan. For the new plan to be different, variables (i.e.- thresholds, RESERVED volumes, monitor interval, etc.) would need to be changed in order to create a plan that would be different from the existing plan.
3. Set a time for the Migration Plan created by Auto LUN to start. Since Auto LUN's primary purpose is to help distribute load for increased performance, schedule an off-peak time for the migration(s) to occur. Please read the HP white paper on AutoLUN for more information.
4. Set a time interval for Auto LUN to complete the volume migration. Again, because moving volumes can be resource intensive, Auto LUN allows you to set a time window for completion to occur (the maximum is 120 minutes). If the migration plan has not completed within this interval, Auto LUN will stop the migration. If this occurs, all data for the unfinished migration is maintained at the source volume. Migrations that did not complete during the plan interval will automatically be attempted again the following day at the same migration start time. If the migration stopped during the middle of migration, Auto LUN will restart that migration from the beginning (Auto LUN will not pick up where it left

off). For this reason, it is simple to see why the optional time interval for migration completion must be carefully set. The maximum value accepted for this interval is 120 minutes; this value is for the entire plan's completion and not 120min for each volume migration within that plan.

5. Interval cycle (if any) for Auto LUN to repeat steps 1 through 4 from above. In the Preset Function, the cycle information collection and plan creation can be for only one day, one week or several weeks. A short interval cycle (i.e.- every day) might affect Auto LUN's ability to complete migrations. For example, we define a cycle to create a plan every day. On day one the plan creates 20 migrations, but because of the 120min max time for completion, only 8 migrations completed, leaving Auto LUN to attempt the remaining 12 unfinished migrations the following day (see item #4 above). However, the interval cycle will override the execution of the 12 remaining migrations when it automatically creates a new plan for that day. Careful thought must be given when defining automatic migration plan variables.

Auto LUN Automatic Migration Example

- Set the maximum disk utilization limit (i.e.- 60%) for each disk class. By default, Auto LUN will consider all volumes that exceed this threshold as candidate for migration
- Turn on disk monitoring
- Check past migration history
- Set the fixed parity group(s) you do not want Auto LUN to migrate if the disk utilization exceeds the defined threshold
- Define volumes as RESERVE to serve as targets
- Set the monitoring schedule, migration start time and maximum migration completion time interval
- Check the migration status
- Display migration history

HP Restricted

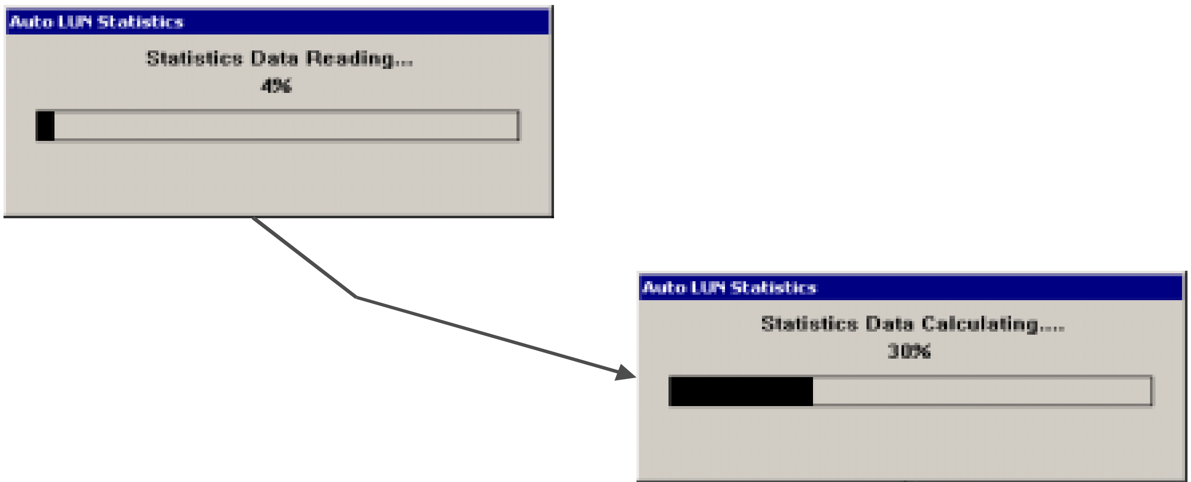
NSSO
Stress-Free Storage...Guaranteed

Setting up automatic migrations to occur should be done with caution. The migration process itself is very safe; the caution required should be in understanding why an automatic migration occurred and interpreting the information Auto LUN used to qualify a performance gain could be gained by performing a migration.

HP's recommendations for automatic migration:
After turning on automatic migration, define all parity groups as FIXED. This step prevents automatic migration from considering each and every parity group's volumes as candidate for migration.
Using past historical statistics or forecasting performance trends, define potential hot parity groups as candidate for migration and move them out of the FIXED list.
Using past historical statistics, set the maximum tolerated percentage use values for each disk class. Understand that the value defined will determine the point at which volumes with parity groups that belong in that class will be candidate for migration.
Properly disperse RESERVES within the array. RESERVEs should be defined for each emulation that will be migrated, for each RAID type, and dispersed across ACP's and CHIP connections. Understand that wherever you define your RESERVES, that parity group will be receiving HOT volumes, thus RESERVES should be tied to parity groups that can handle high I/O.

Define data term interval for decision analysis and a non-conflicting time for migration starts. For example, interval decision time should be during peak activity and may or may not be a long interval period. Also, the SVP reboots nightly at 3:30am by default, so 3:30am is not a good time for data term or migration start time.

Auto LUN Program Interface- Launching from Remote Control XP



HP Restricted

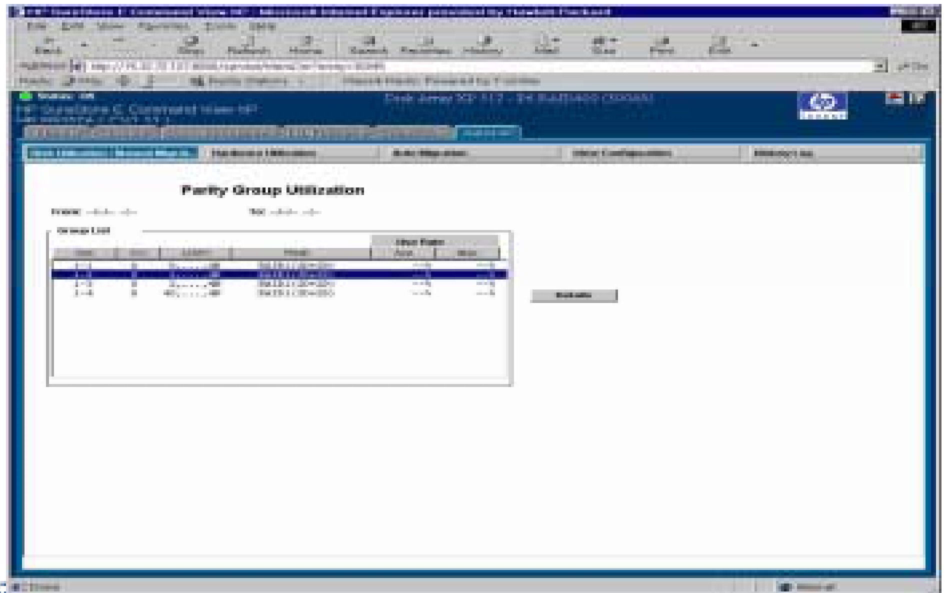
N S S O
Stress-Free Storage...Guaranteed

From the Remote Console, select Auto LUN button (Users can log on with Administrator rights if changes are planned or as a user with VIEW only permissions). Upon launch a dialogue box “do you wish to gather statistics information” is displayed. Answering “YES” allows Auto LUN to take monitor information from the SVP for data analysis to occur on the Remote Console PC.

As the SVP collects more information, the time it takes to open Auto LUN’s first screen increases. For example, after having Auto LUN collect information on an array that contained 16 parity groups for a period of 5 days, it took 8 minutes for Auto LUN to open to the first window (Remote Console was running on a 486/66, the minimum recommended PC for RMC).

As mentioned previous, both the SVP and the Remote Console must be keyed before Auto LUN functionality becomes activated on the Remote Console.

AutoLUN Main Tab in Command View XP



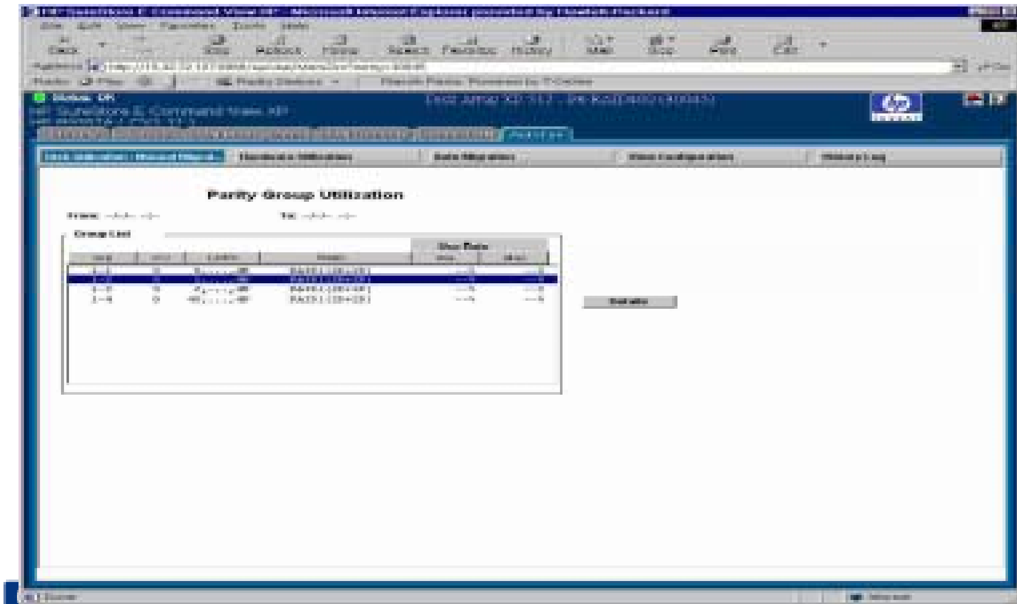
HP Restricted

Stress-Free Storage...Guaranteed

Student Notes

Auto LUN allows you to configure the disk array to monitor disk volume performance and auto-migrate the data.
You can create a migration plan using a script-like interface to define the migration conditions. Then, based on this plan, the array can move the data and path automatically whenever a threshold limit is crossed.
There is also a manual migration option.
Auto LUN Operations
Disk Utilization Displays parity groups and logical volume utilization. Manual migration also takes place using these screens.
Hardware Utilization Displays CHP, DKP, MBUS and DRR utilization.
Auto Migration Displays utilization values for different logical volumes based on a migration plan. You can set threshold parameters for automatic migration of LDEVs depending on their utilization. View Configuration - Allows a user to start or stop resource monitoring, and to set the time period for viewing the information.
History Log - Shows logs of automatic migration and LDEV migration.

Disk Utilization/Manual Migration Parity Group Utilization



HP Restricted

Stress-Free Storage...Guaranteed

Student Notes

This screen allows you to view the parity groups utilization information. It shows average and maximum use rates for various parity groups.

From, To displays the start and end date of monitoring term.

Group List:

Grp displays - Parity Group ID

CU displays - Control unit it belongs to

LDEV displays - a list of logical devices contained in this parity group

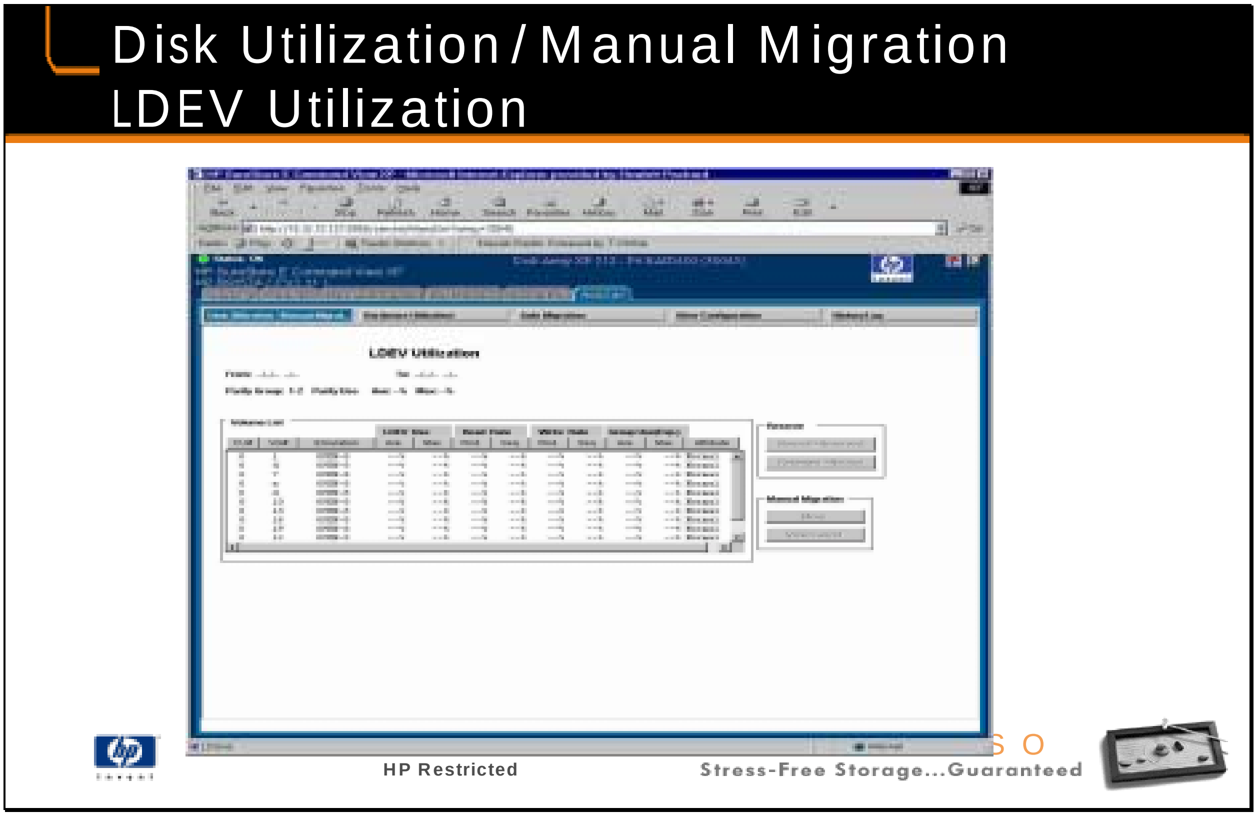
RAID displays - the RAID type of this parity group. RAID1 and RAID5 are supported

Use Rate displays - Average and Maximum Utilization of this parity group

The Group List display - allows you to sort the parity group utilization information. You can sort on a column by clicking on the column header. The name of the column on which the list is sorted will be highlighted.

Monitoring data is shown in percentages (%). If no monitoring data is available, "--%" will be shown to represent null utilization data and the From and To fields will show "--" in place of date and time.

Select the Details button to display monitoring information about the LDEVs in a parity group. Select one parity group from the Group List and select the Details button. The LDEV Utilization screen appears with detailed monitoring information about the LDEVs in the parity group. Select the Refresh button to obtain the latest monitoring information from the MIB and updates the screen.



Student Notes

This screen appears when you selects the Details button in the Parity Group Utilization screen. The utilization information about each of the LDEVs in the parity group selected is displayed along with disk serial number .

From, To displays the start and end date and time of monitoring term.

Parity Group displays the Parity group ID.

Parity Use displays the average and maximum utilization (%) of the Parity Group

Volume List.

CU displays Control Unit.

Vol# displays Logical Device.

Emulation displays Emulation of Logical device.

LDEV Use displays Average and Maximum Utilization of LDEV.

Read Rate displays the Random and Sequential Read rate utilization of LDEV.

Write Rate displays the Random and Sequential Write rate utilization of LDEV.

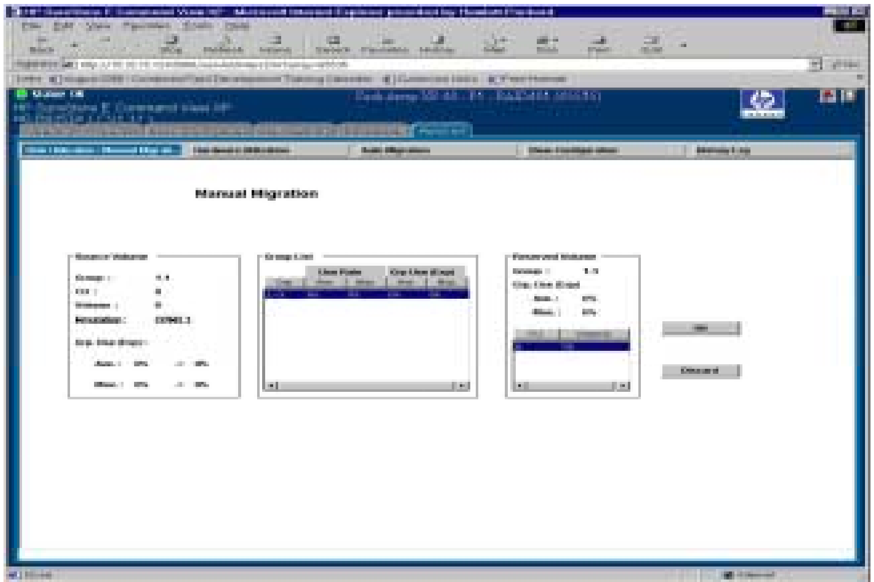
Group Use(Exp) displays Expected Average and Maximum Utilization of the parity group if this LDEV is migrated to some other parity group.

The Volume List display allows you to sort the logical volume utilization information. You can sort on a column by clicking on the column header. The name of the column on which the list is sorted will be highlighted

This screen provides the following buttons to change the LDEV attributes:

Normal -> Reserved converts a Normal to Reserved volume. This button will be disabled for View only users, and for Modify users if no rows are selected or if selected row does not have “Normal” attribute.
Reserved -> Normal converts a Reserved to a Normal Volume. This button will be disabled for View only users, and for Modify users, if no rows are selected or if selected row does not have “Reserved” attribute.
Move displays the Manual Migration - Move screen, used to create a migration plan for the volume selected in Volume List. This button will be disabled for View only users, and for Modify users, if no rows are selected or if selected row does not have “Normal” attribute.
View/Stop Migration displays the View/Stop Migration screen to view the status of the migration plans and to stop the migration, if needed. This button will be enabled for view and modify users only if any manual migration plans exist.

Disk Utilization/Manual Migration
Manual Migration



- Manual Migration screen

HP Restricted

Stress-Free Storage...Guaranteed

Student Notes

This screen shows utilization information for creating migration plans. It can only be accessed by users with modify privileges.

Source Volume displays the volume from where the migration will take place

Group The group to which the source volume belongs.

CU Shows the CU number of the source volume to be migrated.

Volume The LDEV number of the source volume to be migrated.

Emulation The emulation type of the source parity group.

Grp Use (Exp) Average and maximum utilization of the parity group before and after the migration of the volume from this parity group.

Group List displays all the parity groups that are of the same emulation type as the source parity group.

Group Possible target parity group.

Use Rate The current average and maximum utilization of this group.

Grp Use (Exp) The average and maximum utilization of this group is the source volume is moved to this group.

Reserved Volume displays the volume where the migration will take place.

Group Shows the parity group selected in the Group List.

Grp Use (Exp) The average and maximum utilization of this group if the source volume is moved to this group.

CU The CU number reserved volume in this parity group.

Volume The LDEV number of the reserved volume in this parity group.
Buttons
OK The migration plan is created for the source volume and target volume selected.
This is button is enabled only for modify users.
Discard All prediction utilization values are discarded and control returns to the LDEV
Utilization screen. A different LDEV can now be selected for migration. This button is enabled only for users with modify privileges.
The Auto LUN Manual Migration function lets you take full control of Auto LUN tuning operations. Using manual migration operations, you can fine-tune the subsystem to optimize parity group, volume, and back-end processor usage (DKPs, DRRs) for your operating environment.
Before performing Manual Migration operations, you must collect and analyze monitored data.
After collecting the data, you must decide on the source and target volumes for manual migration. This is essentially making a migration plan.
After performing manual migration operations, you can collect more monitored data to verify the tuning results of the manual volume migration.
The manual migration functions supported by the web interface allows you:
• Create a migration plan
• View the migration progress status
• Cancel of the manual migration plan

The Auto LUN Manual Migration plan involves analyzing the Auto LUN monitoring data to select the source LDEV, and then determining the target parity groups.
1. Analyze the source LDEV
The Auto LUN monitoring data is displayed in the Disk Utilization and Hardware Utilization screens.
Analyze the DKP, DRR and DKA information to select a DKA pair. The DKA pair is chosen if the DKA pair utilization is much higher than other DKA pair or if the DRR/DKP average utilization is over 60%.
Select the parity group that has the highest utilization under the selected DKA pair.
Select the LDEV with the highest utilization within the selected parity group.
2. Determine the target parity groups
The target parity groups should satisfy the following conditions:
The PG should not be the same PG that contains the source LDEV.
The PG should have reserved volumes.
The reserved volume should have the same emulation type and sizes as the source LDEV.
Determine the prediction utilization data for the selected target parity groups.
The source LDEV can be moved to any of the selected parity groups. The utilization of each of the parity groups is determined. The parity group that has the lowest prediction utilization is then selected.

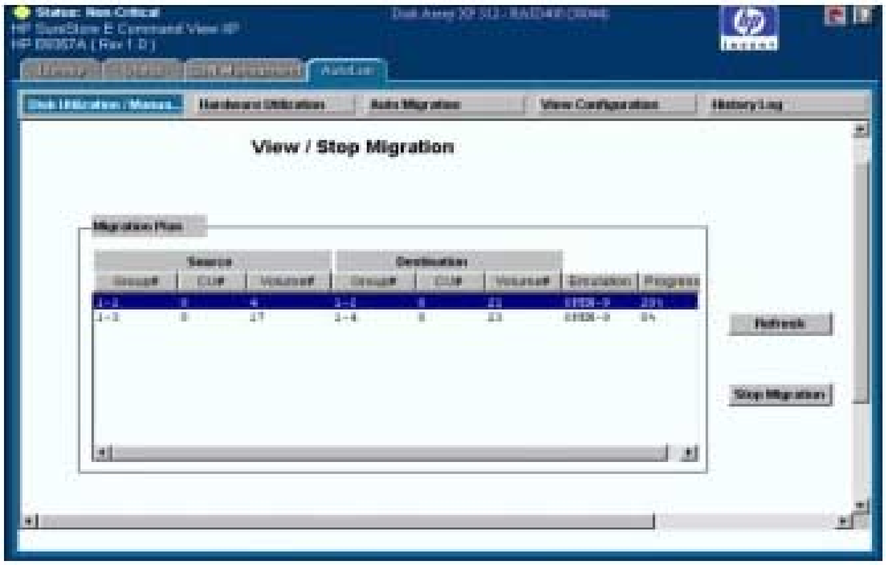
3. View Migration Progress Status. Manual migration allows the creation of up to 36 migration plans. Only one plan can be executed at any time. Manual migration displays the current migration plans and the progress status of the plan currently under execution.

Canceling Manual Migration. Manual migration allows the cancellation of a migration plan when the plan is in progress. A plan can be cancelled any time before the completion of the migration.

Precautions. Manual migration should not be done when the automatic migration is in progress.

History Log. Check the history log to view migration events.

Disk Utilization/Manual Migration
Manual Migration



- View /Stop Migration screen

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Migration Plan displays the migration plan details.

Source The source volume that has to be migrated. Shows the parity group, CU and LDEV numbers of the source volume. An '*' next to the volume number specifies that volume is currently under migration.

Destination - The target volume to which the source volume has migrated. Displays the parity group, CU, and LDEV numbers of the target volume.

Emulation The emulation type of the parity groups.

Progress Shows the percentage of the source volume migrated to the target volume for the volume currently under migration. All other plans will indicate 0%.

Drive Type - Hardware display name for drive type.

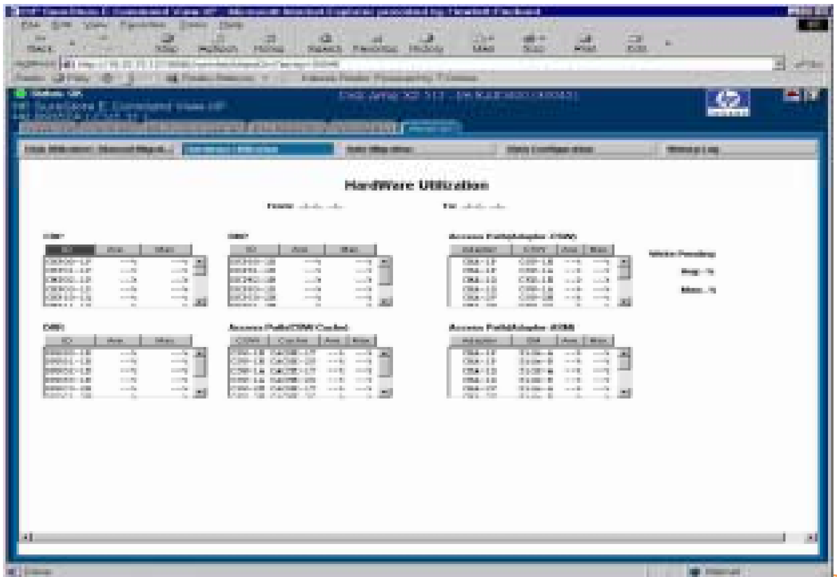
Total - Total number of LDEV in a particular parity group.

Buttons

Refresh The migration process status for the volume under migration is refreshed. This button is enabled for users with view and modify privileges.

Stop Migration Stops/Cancels the migration plan. This button is enabled only for users with modify privileges.

Hardware Utilization



HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes
For an XP 256, you can view the CHP, DKP, DRR and BUS monitoring information along with disk serial number. For an XP 512: You can view the CHP, DKP, DRR , CSW-Cache, Adapter-CSW,Write-Pending and Adapter-ASM monitoring information along with disk serial number.
Resource Utilization Values
Utilization Value. This represents the percent of time for which the resource is active (non-idle) during the monitoring term (period). The DKC monitors and records usage in 15-minute intervals. Each value it records represents a data point in the monitoring term.
Average Utilization Value is the average value across all data points for the monitoring term.
Maximum Utilization Value is the highest use percent across all data points for the monitoring term.
System Resources monitored by Auto LUN

Channel Processor (CHP) and Disk Processor (DKP) utilization

DRR (Data Recovery and Reconstruct) utilization. DRR is LSI-generating RAID parity information

Parity group utilization (physical disk drives). This is the time utilized by physical drives in parity groups. Parity groups are groupings of hard disks in a particular RAID configuration

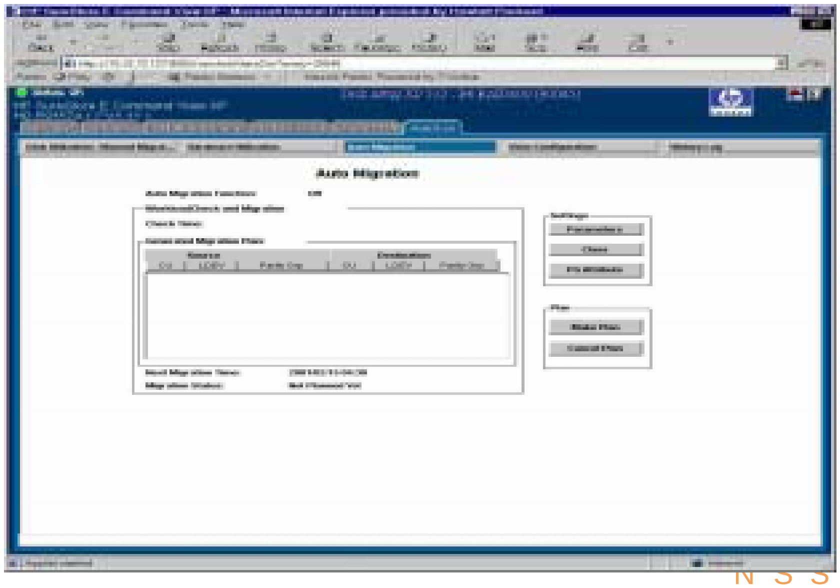
Parity group utilization of each logical volume

Access Path (CSW-Cache)
Access Path (Adapter-CSW)
Access Path (Adapter-ASM)

Write Pending. This is the time utilized by physical drives of synchronous and asynchronous access on each logical volume, averaged by the number of physical drives in the parity group.

When no monitoring data is available, Auto LUN shows null data as “--%.” The value “0%” means that the utilization is zero percent, while “--%” means that no information is available.

Auto Migration



HP Restricted

Stress-Free Storage...Guaranteed

Student Notes

The Auto LUN Auto Migration function lets you set migration plan parameters and thereby tune disk arrays. This screen allows you to view the migration plan information.

Auto Migration Function displays whether Auto Migration is ON or OFF.

Check Time displays the time at which the migration plan was created.

Generated Migration Plan shows details of the generated migration plan, if any. It displays the source and destination volume information. Next Migration Time is the time for the next migration to start. Migration Status may be one of the following:

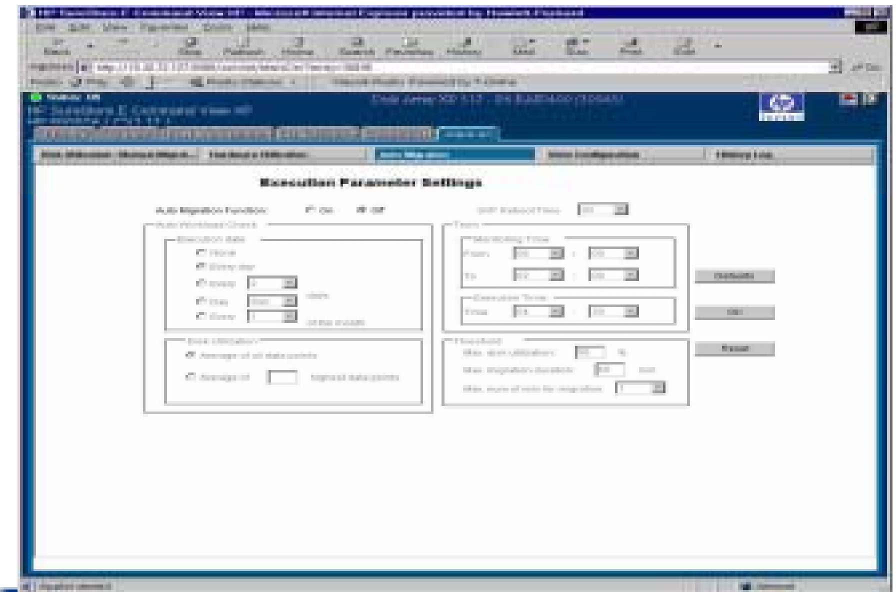
- No plan created
- Plan not performed yet
- Migration being performed
- Last migration canceled
- Failed to make a plan

To set up Auto Migration:

1. Click on the Parameters button to set Auto Migration on.
2. Click on the Class button to set the class threshold utilization.
3. Click on the PG Attributes button to set the parity group as fixed or non-fixed.
4. Click on the Make Plan button to delete the existing plan and make a new plan.

The Cancel Plan button cancels the migration plan made on the last “check time”. Canceling a plan does not mean stopping the creation of the plan. To stop creation of the plan, Auto Migration status should be set to OFF. After the next monitoring phase, the DKC will create the new plan, taking into account the current migration parameters.

Auto Migration – Execution Parameter Settings



HP Restricted

Stress-Free Storage...Guaranteed

Student Notes

Auto Migration Function - Set this radio button to ON to turn Auto Migration on.

Auto Workload Check

Execution Date. The date an Auto Migration plan begins. You can set this date to none, daily, every X days, on a particular day of the week or for every X day in a month. The limit values - for every X day is 2 to 31 days.

Disk Utilization - This parameter is used to determine the average of collected utilization information. You can derive either the average of all collected data, or the average of X highest points in collected data.

Term - Monitoring Time. This is the time when monitoring information about LDEV and parity groups is collected. Based on the utilization information collected in this period, the migration plan is made by the DKC.

Execution Time - Displays the time that the migration plan was executed.

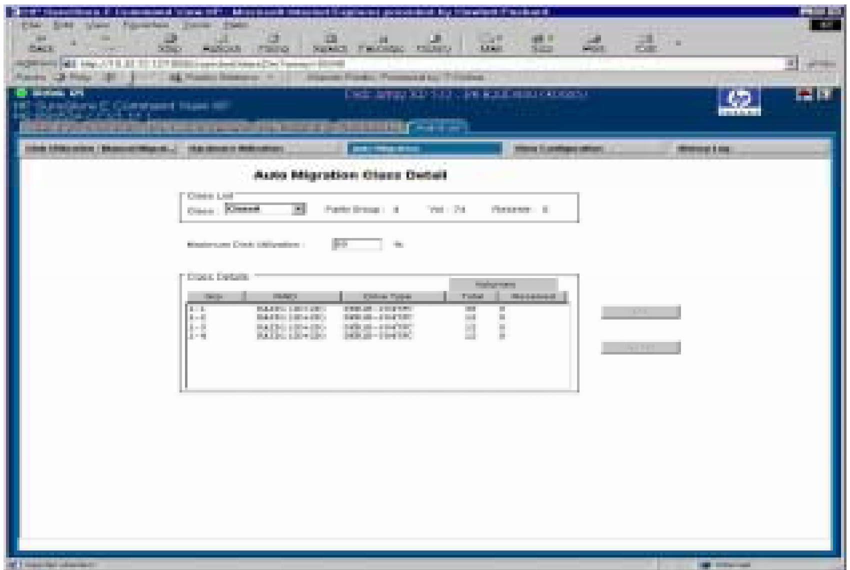
Threshold - Maximum Disk Utilization. If the destination LDEV during migration crosses this utilization threshold, migration will terminate.

Maximum Migration Duration - The Migration plan should not require more than this time. If it does, the migration stops and is attempted again at the next execution time.

Maximum Number of Volumes for Migration - This is the maximum number of volumes (1 to 40), that can be migrated at one execution time.

The Default buttons sets the migration parameters to their default values. The OK button sets the changed migration parameters. The Reset button restores the migration parameters to their previous values.

Auto Migration – Class Detail



HP Restricted

NSSO

Stress-Free Storage...Guaranteed

Student Notes

Class List

Class - The classification type for class, for example, A, B, and so forth. It is followed by a summary of class details, for example, the number of parity groups it has, total number of volumes, and how many volumes are reserved.

Maximum Disk Utilization - The threshold utilization for selected class type. When the utilization during the monitoring term exceeds this threshold, it becomes eligible for migration.

Class Details

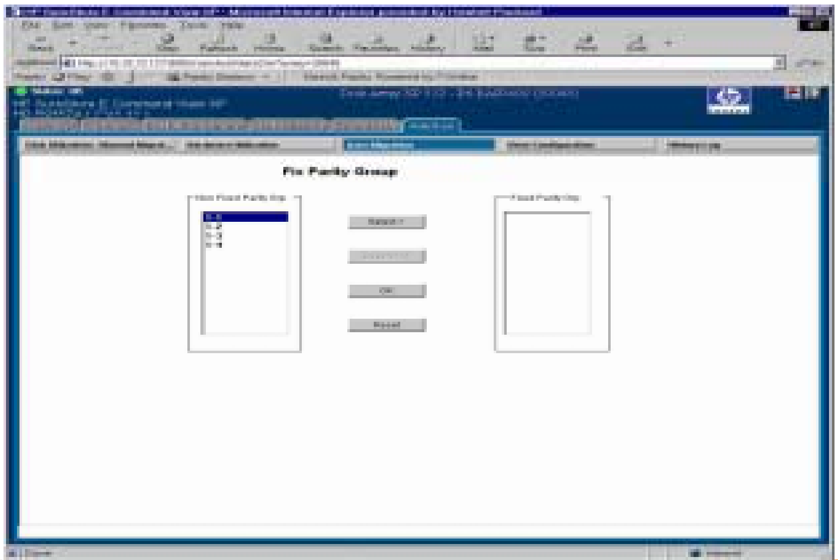
- Grp. Shows parity group name.
- RAID. RAID type of particular parity group.
- Drive. Type Hardware display name for drive type.
- Total. Total number of LDEV in a particular parity group.
- Reserved. Total number of reserved volumes in a particular parity group.

Buttons

OK. Sets the maximum utilization for the selected class. The page becomes inactive to user input until the result of changing threshold appears.

Reset. Restores the maximum threshold value for the selected class if changed before set.

Auto Migration – Fix Parity Group



HP Restricted

NSSO Stress-Free Storage...Guaranteed

Student Notes

The LDEV of a fixed parity group is not taken into consideration while making the migration plan. The LDEV cannot be migrated through Auto Migration. The LDEV of a non-fixed parity group can be migrated only through a migration plan.

Fixed Parity Group

Select a fixed parity group from the list to move it to the non-fixed list

Non-Fixed Parity Group

Select a non-fixed parity group from the list to move it to the fixed list.

Buttons

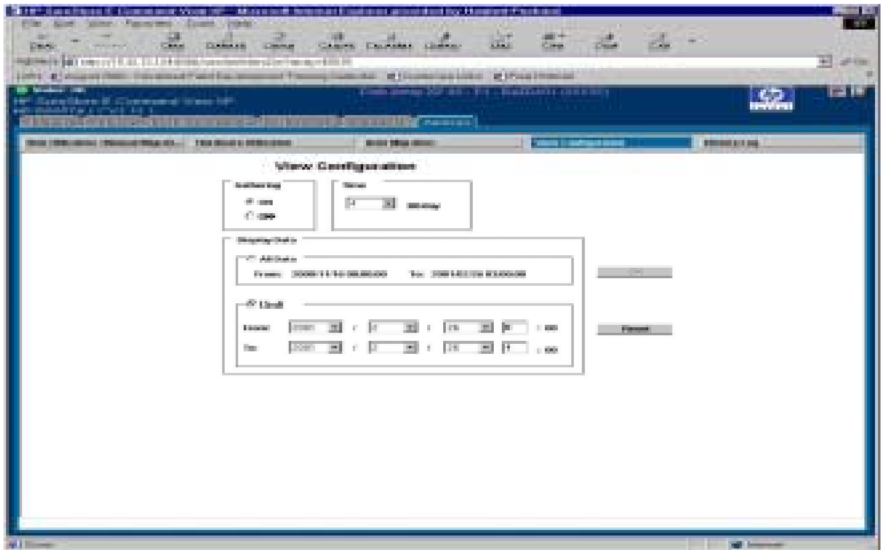
Select < Choose Select to move parity groups from fixed group to non-fixed group.

Deselect > Restores the maximum threshold value for the selected class if changed before set.

OK Select the OK button to set the changes done to parity groups. The screen becomes inactive to user input until the result of changing parity group attribute is displayed.

Reset Select Reset to restores the fixed and non-fixed group values, if changed before performing the set operation.

View Configuration



HP Restricted

N S S O

Stress-Free Storage...Guaranteed

Student Notes

The View Configuration page allows you to start/stop monitoring and set the "From/To" monitoring term.

Initially, monitoring (gathering) is set to OFF. Before the monitoring function starts for the first time and when no monitoring information is available, the Display Data component shows NULL values. The Display Time component shows a default value.

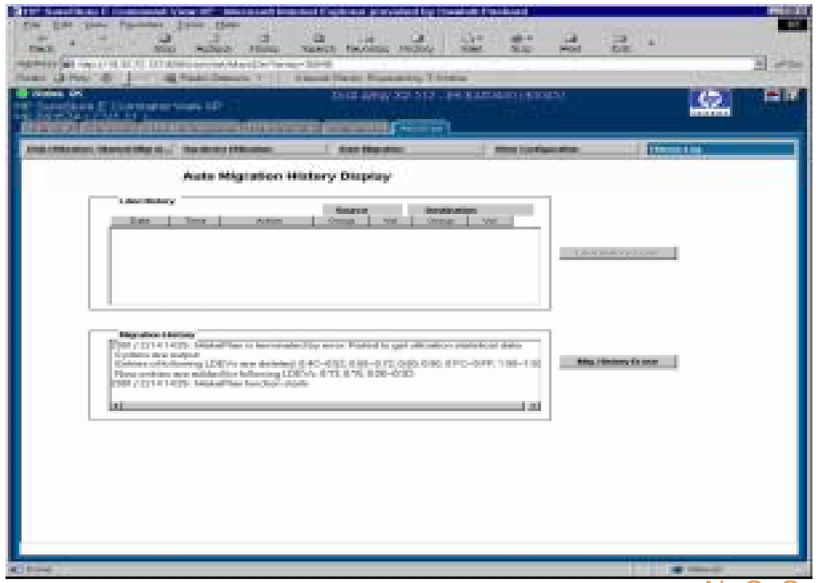
When monitoring is turned ON for the first time, only the Time field becomes enabled. Simultaneously, the Display Data component becomes disabled because no monitoring information is available. Display Data becomes enabled only when monitoring information is available.

To enable monitoring:

1. Select the ON radio button in the Gathering area.
2. Select a time in the Time area. This is the time at which the SVP will gather monitoring information from the DKC and update the MIB values. The SVP does not support "01(AM &PM), 02(AM&PM), 03(AM&PM)" hrs as time to gather information, and hence these values are not included.
3. Set your date and time choices in the Display Data areas. You can select the All Data radio button or the the Limit radio button.
4. Click on OK. The system saves your settings, and then the DKC becomes enabled and begins monitoring the system resources.

Select the Cancel button to reset the components displayed to their initial values.

History Log



HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

The AutoLUN auto migration function logs all auto migration events in the auto migration history log. An auto migration event is any event related to the creation and execution of auto migration plans.

This screen displays an Ldev History window and a Migration History window.

Ldev history

The LDEV History box displays the migration history, with source and destination group and volume details.

Migration history

The Migration History box displays messages related to the creation of a plan, changing migration parameters, and success/failure in executing migration plans.

Buttons

Ldev History Erase erases the LDEV history displayed. This button is enabled only for users with modify privileges.

Mig. History Erase erases the Migration history displayed. This button is enabled only for users with modify privileges.

To refresh the history displayed, go to the Disk Utilization screen and click the Refresh button, as there is no separate refresh operation for history.

Auto LUN Export Of Data

- Using the AutoLUN Graph Tool, statistical information collected on the SVP can be saved to floppy disk and loaded into the Graph Tool for a graphical display of the following:
  - CHIP Processors
  - ACP Processors
  - DRR
  - Bus Utilization
  - Parity Group & volume utilization

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

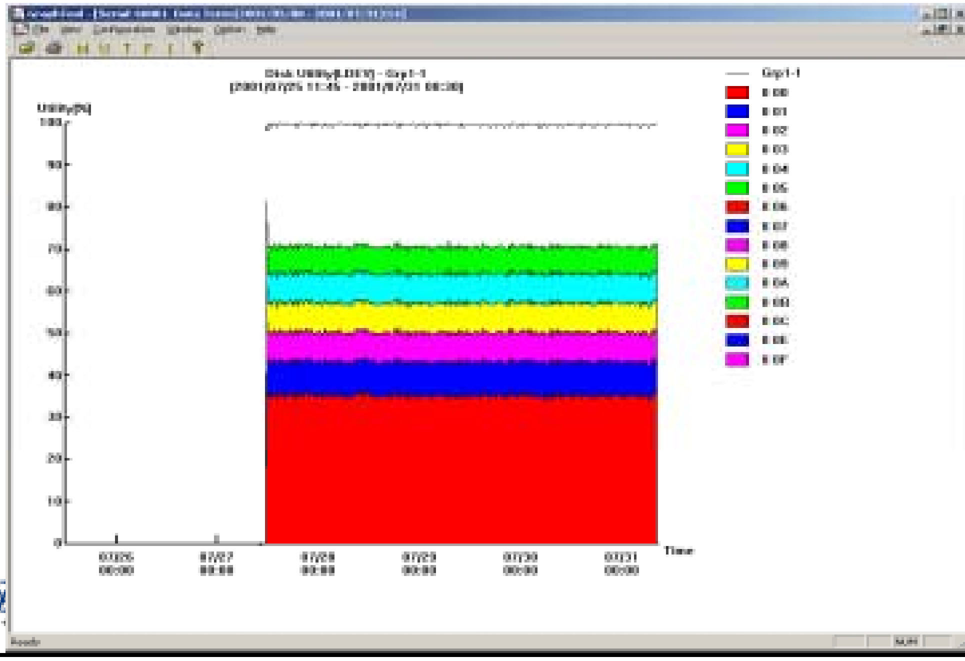
As mentioned in the Monitor Function slides at the beginning of the presentation, the information collected daily is stored on the SVP's hard disk for 90 days; if data is older than 90 days Auto LUN will overwrite the oldest data with the newest data. Auto LUN allows for export of monitored data to floppy if you have a need to keep information older than 3 months.

Auto LUN Graph Tool

The Graph Tool software requires the following operational environment:

- A Windows PC (Windows 95, 98, 2000, or WindowsNT 4.0) with a diskette drive (to access the a:\artlog.lzh file) and a CD drive (to install the Graph Tool software).
- At least eight MB of free space (three MB to install Graph Tool, about five MB when using Graph Tool). More free space may be required depending on the configuration of the disk array whose monitor data you are analyzing (for example, number of parity groups, DKPs, or volumes) and the range of monitor data being analyzed (that is, size of artlog.lzh).

Auto LUN Graph Tool Example



Illustrated above is an example of the CHP Utility graph. The Module list box lists the data modules available for graphing:
CHP Utility: Actual, average, and maximum usage rates of the CHP's.
DKP Utility: Actual, average, and maximum usage rates of the DKP's.
DRR Utility: Actual, average, and maximum usage rates of the DRR's.
Disk Utility (Parity Group): Actual, average, and maximum usage rates of all parity groups.
Disk Utility (LDEV): Actual, average, and maximum usage rates of LDEV's of a specific parity group. If you select this module, the Select Parity Group window opens automatically to allows you to specify the desired parity group.

See the Auto LUN User's Guide for more details on the Auto LUN Graph Tool.
http://www.hp.com/cposupport/manindex/hpsurestor21443_eng_man.html

FAQ

- Is there an automatic notification facility for migration failures?
  - No. Migration failures notification is not automatic but is available via the HISTORY logs.
- Can I define RAID1 to be viewed as higher performance than RAID5?
  - No
- Is there any current integration between Performance Manager XP and Auto LUN XP?
  - No
- Is there an automated solution to retrieve statistics information used by the Excel spreadsheet?
  - No, the process supplied by Auto LUN is manual retrieval

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Please reference the HP Surestore Remote Control XP User's Guide for more information on Auto LUN...

http://www.hp.com/cposupport/manindex/hpsurestor21443_eng_man.html

FAQ (cont.)

- Can I create a manual migration to start at a predefined time?
  - No. Automatic migration allows you to re-create plans using Auto LUNs estimate function which must be scheduled to start at a predefined time, but this is very different than manually defining a volume migration. All manually created plans must be started by the user and cannot be started automatically by Auto LUN.
- How long does it take to migrate an OPEN3 volume?

The following numbers assume no activity on the array, and migrating to a parity group on a different ACP pair

  - XP256: 7 minutes
  - XP512: 1 minute
- If HDD's spin at the same rate, how can I have different disk classes?
  - Automatic migration will view the lowest density disk as the fastest spinning disk. For example, a FC 15GB disk would be class A and the FC 47GB disk would be class B.

HP Restricted

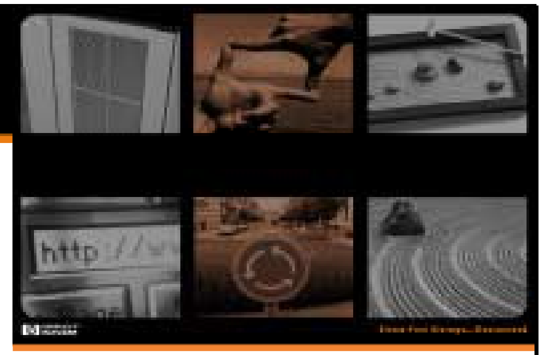
NSSO

Stress-Free Storage...Guaranteed

Module Wrap-up

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

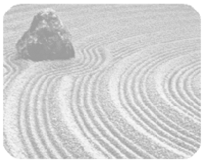


Module 5

Performance Advisor XP

N S S O
Stress-Free Storage...Guaranteed

Module Agenda



- Describe HP Surestore Performance Advisor XP
- Define “Management Station”
- Define “Host Station”
- Name four benefits of using XML
- Identify the supported operating systems
- Describe the HP Surestore Performance Advisor XP product
  - Functionality
  - Hardware and software components
- Define Performance Advisor functionality
- Install Performance Advisor
- Configure Performance Advisor
- Troubleshoot Performance Advisor problems
- Wrap-UP

N S S O
Stress-Free Storage...Guaranteed

HP SureStore E Performance Advisor XP

Real-time performance monitoring

- ✓ Internet-based application that provides real-time collection and monitoring of XP Disk Array family
- ✓ Can be used standalone or integrated into Command View XP
- ✓ Integration into VantagePoint MeasureWare and VantagePoint SPI for XP
- ✓ Command line interface for easy integration into enterprise level solutions
- ✓ Safe and secure authentication and authorization

N S S O
Stress-Free Storage...Guaranteed

Performance Advisor XP is an Internet-based application that provides real-time monitoring and data collection for the XP256/512/48. Performance Advisor consists of the following components: management station, distributed host agents and web-browser of CLI interface. The CLI provides easy integration into enterprise level infrastructure solutions.
The data collected by Performance Advisor can be presented in two different ways: host-centric or array-centric.
Performance Advisor can be used as a standalone solution or can be integrated into Command View XP, the strategic management platform for the XP family of disk arrays.
Performance Advisor provides the ability to have events/traps forwarded to the VantagePoint management platform. This can be accomplished in two ways:
integration into VantagePoint MeaureWare
Integration into the VantagePoint Operations platform via the XP SPI (available January ’01)
The command line interface also enables integration into other leading enterprise level applications
Security – Performance Advisor will provide two basic components for security: password file for user authentication and authorization via the Apache web server.
When the administrator logs on, they will be required to provide a user password which will be compared with the information contained in a password file. If they

match, the user is authenticated. The Apache web server controls authorization. Once the user has been authenticated, the web server specifies the directories, files and servlets that a particular group or user is allowed access.

Overview of Performance Advisor XP

- Collects and monitors real-time performance of the HP Surestore XP family of disk arrays
- Provides real-time and historical data on:
  - LDEV I/O
  - Front-end and back-end port utilization
  - Internal bus utilization
  - Cache usage
- Sends performance alarms (also, performance alarms can be integrated with VPO)
- Uses a simple, browser-based interface
- Easy to customize performance data collection
- Integrated with VPP (VantagePoint's Performance) and VPO / SPI

NSSO
Stress-Free Storage...Guaranteed

Features and benefits (1)

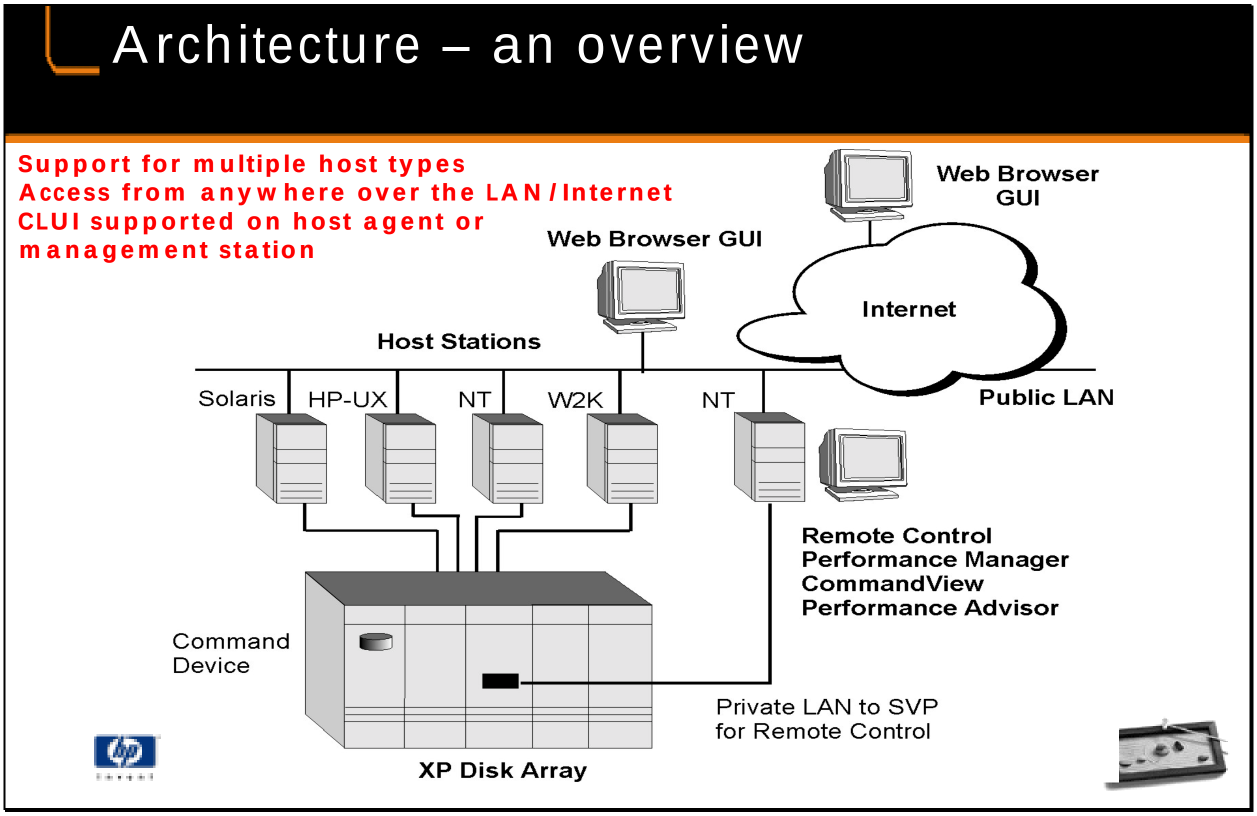
FEATURES	BENEFITS
Works standalone or integrated into Command View XP or Vantagepoint Measurware	Ensures ease of management and protects investment in tools
Proactive thresholds and e-mail notifications	Sends alerts when the storage system reaches user-defined limits before problems occur
Internet-based application	Possible to monitor the system anytime, from anywhere
Secure authentication and authorization	Keeps sensitive performance and configuration data safe
Integration with Command View XP	Enables real-time monitoring from the strategic management platform for HP Surestore XP disk arrays

NSSO
Stress-Free Storage...Guaranteed

Features and benefits (2)

FEATURES	BENEFITS
Integration with VantagePoint MeasureWare	Enables seamless performance management of globally distributed XP infrastructure
CLI or GUI interface	Two interfaces available to meet different needs. CLI interface enables easy integration into third-party solutions
Notification via e-mail or VPO management server	Notification of potential performance bottlenecks before they affect applications and users
Central database	Enables IT personnel to store historical information to identify trends or potential bottlenecks

N S S O
Stress-Free Storage...Guaranteed



Student Notes

Performance Advisor has three main components:

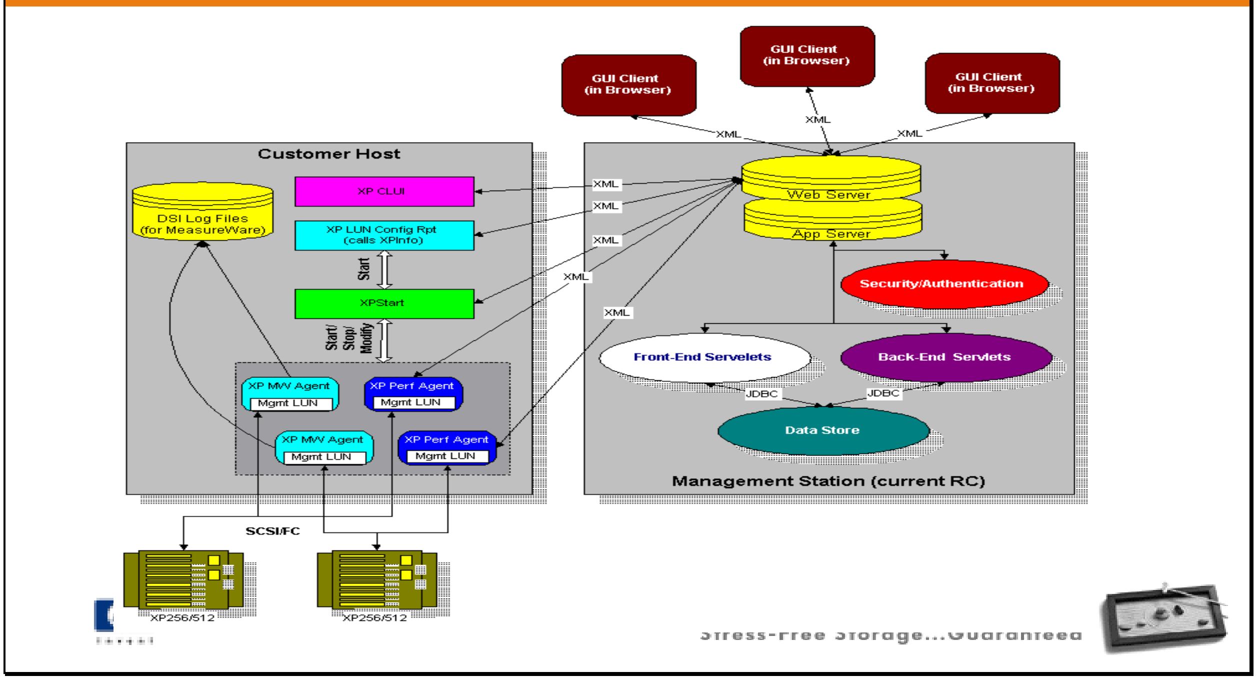
Centralized Management Station
The Performance Advisor centralized management station is compatible with HP Surestore Command View XP management station and they can co-exist on the same Remote Control PC. Performance Advisor includes the Jserv Servlet Engine and a JDBC compatible database.

Distributed Host Stations
Host agents running on the hosts are responsible for data collection, LUN mapping configuration reporting, and MeasureWare integration.

Web Browser Presentation Clients
This is the user interface to Performance Advisor.

XLM is used for data communication between all components. This enables the management station and browser-based clients to be located anywhere on the Internet.

Performance Advisor Architecture



There are three main components to Performance Advisor: centralized management station (or Command View management station), distributed Host Agents and web-browser or command line presentation clients. Data communication between all components is via XML which provides the following benefits:

Operating System interoperability. No proprietary runtime support is required

Uses standard http communications/channels which enables the web security to be leveraged with other applications

Works over firewalls enabling management from a web browser anywhere on the Internet

Provides flexibility in locating the management station via the Internet

Management Station – has four main components: Front/Back-end servlets, XML communication layer, persistent datastore for storing performance data and performance alarms.

Servlets – responsible for handling XML-based exchanges occurring between the DataStore and either a host agent or presentation client.

DataStore is a repository for all array performance data, configurations and health/status information. Performance data and configuration data is presented from the Servlets as Java objects which are translated into relational format and stored in the database.   The DataStore saves the following information:

Performance data

Configuration Information

User authentication information
Alarms/Status Health – notification of alarms/status heath is provided via notification through email and VPO (January ’01). Performance Advisor provides user definable trigger thresholds, enable/disable capabilities and multiple alarm destinations. Performance Advisor monitors the health of agents and the management station.

Management Station – co-existence with other software

- Performance Advisor can run standalone
- Performance Advisor can be integrated with Command View 1.11
- Can co-exist with Command View and Performance Manager (co-existence currently being tested) on the same remote console PC
- Components
  - DataStore (repository)
  - Servlets
- Features
  - XML
  - Performance Alarms
  - Standard Apache web server security

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Co-existence
Performance Advisor can co-exist with other management software on the same machine.
DataStore
The DataStore (repository) stores array array performance data. It maintains Performance Advisor configuration metadata. It can take up a lot of space, but is automatically purged once a week. Properties files control the station's configuration (event alarm notification, messages, etc.)
Servlets
Front-end servlets handle Client exchanges, while back-end servlets handle Host Agent exchanges.
XML
XML is used for data communication between components. It an enables the management station and browser-based clients to be located anywhere on the Internet

Operating System interoperability
Uses standard HTTP channel
Works over firewalls

Performance Alarms
There are several settable levels of performance alarms. Alarms can be issued for sequential reads, sequential writes, random reads, random writes, etc.
Security

Security is provided by the Apache Web Server.

Management Station – components

- DataStore – repository
  - Stores array performance data
  - Maintains Performance Advisor configuration metadata
  - Property files control the station's configuration (event alarm notification, messages, etc.)
- Servlets
  - Front-end servlets handle Client exchanges
  - Back-end servlets handle Host Agent exchanges

NOTE: If you need Performance Advisor integrated in VPO, then the VPO/SPI is required. VPO/SPI is located on the management station

NSSO
Stress-Free Storage...Guaranteed

Management Station – features

- XML
  - Used for data communication between components
  - Enables the management station and browser-based clients to be located anywhere on the Internet
    - Operating System interoperability
    - Uses standard HTTP channel
    - Works over firewalls

- Performance Alarms
  - LDEVs sequential/random reads and writes

- Security
  - Provided by Apache Web Server

NSSO
Stress-Free Storage...Guaranteed

Performance Advisor requirements – Management Station

- The Management Station software must be installed before the Host station software
- Hardware / Operating System
  - 500 MHZ CPU
  - 256 MB RAM
  - 2.5 GB or more free space on hard drive (FAT or NTFS)
  - Network card (two network cards required if PA is to be installed on workstation running Remote Control
  - CD-ROM drive
  - NT 4.0 SP5

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Performance Advisor does not require connection to the private LAN, and a single network card is often sufficient. A second network card is needed only if the management station is running Remote Control and is connected to this disk array over a private LAN.

Supported disk array hosts

Platform	Revision
HP-UX	11.0/10.20
Windows NT	4.0 SP5
Windows 2000	Professional Server, Advanced Server and Datacenter
Sun Solaris	2.6 or higher

NOTE: Both SCSI and FC are supported

NOTE: HP-UX 11.i is not supported at this time

NSSO
Stress-Free Storage...Guaranteed

Performance Advisor requirements – Host Station

- Microsoft NT and Windows 2000:
  - CD-ROM drive
  - 128 MB RAM or greater
  - 200 MHZ CPU or greater
  - Windows NT 4.0 SP5 or greater, or Windows 2000
  - RAID Manager Library, version 01.02.01 or greater
- HP-UX or Sun Solaris:
  - CD-ROM drive
  - HP-UX 10.20 or 11.00, or Sun Solaris 2.5 or greater
  - RAID Manager Library, version 01.02.01 or greater
  - For Sun Solaris only: Java Runtime Environment (JRE), version 1.2, or Java Development Kit (JDK), version 1.2

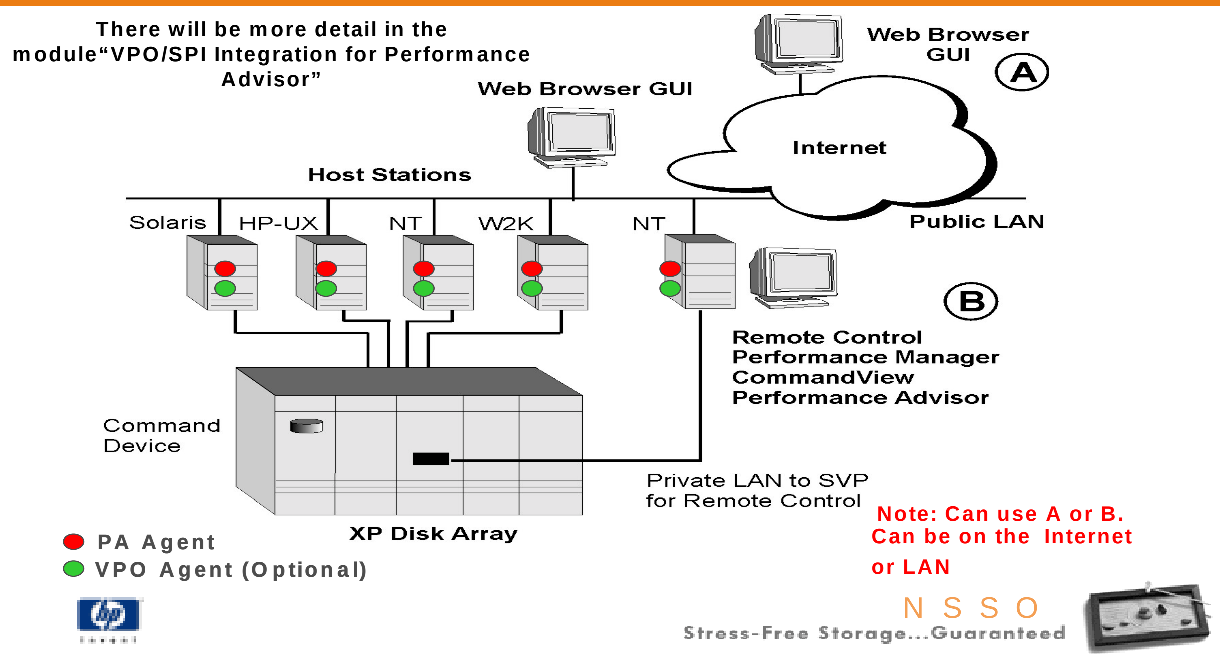
N S S O
Stress-Free Storage...Guaranteed

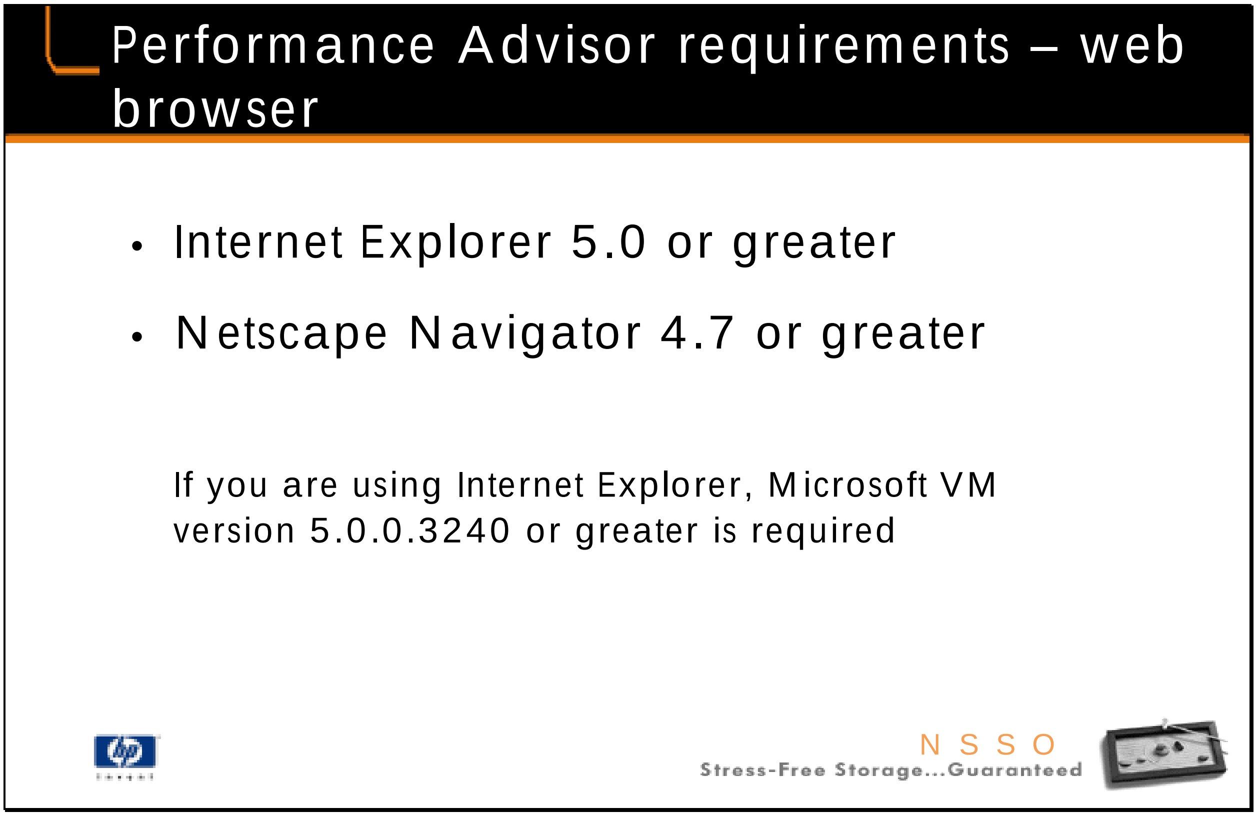
Host station

- Distributed Hosts
  - Performance Advisor agents running on the hosts are responsible for collecting array performance data, LUN mapping, configuration reporting, and MeasureWare integration
  - XML is used for data communication
- The Host Station controls the execution of:
  - ConfigCollector
    - Host configuration information via xpinfo available to the management station
  - DataCollector
    - Retrieves performance data via a command device on the XP disk array

NSSO
Stress-Free Storage...Guaranteed

Architecture – with agents





Student Notes

User Interface
Supported web browsers are Internet Explorer version 5.0 or greater, and Netscape Navigator version 4.7 or greater. Strictly speaking, it is Netscape Communicator that is at the 4.7 or greater level (4.73 is widely distributed and 6.0 is now available). For those who download the Netscape browser only, it is known as Navigator and is currently at the 4.08 level.

Note: Microsoft VM, version 5.0.0.3240 or greater is required if you are using Internet Explorer. Refer to the Microsoft web site for this download.

Performance Advisor requirements – disk arrays

- XP256 firmware ver 52.47.06.00/00 or greater
- XP512 firmware ver 01.11.22.00/00 or greater
- XP48 firmware ver 01.11.22.00/00 or greater
- A command device on the disk array

NSSO
Stress-Free Storage...Guaranteed

Student Notes
For XP256 arrays, firmware version 52.47.06.00/00 or greater is required.
For XP512 arrays, firmware version 01.11.22.00/00 or greater is required.
For XP48 arrays, firmware version 01.11.22.00/00 or greater is required.
A command device on the disk array must be created.

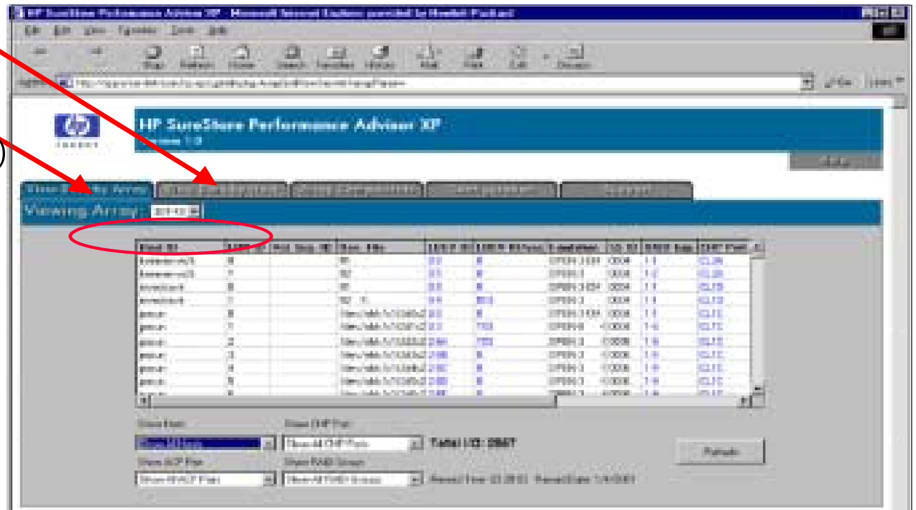
Graphical User Interface

- Graphical User Interface provides for easy entry of management components and commands for:
  - Alarm Configuration
  - Array Components
  - Data Collection Configuration
  - Historical Data
  - Host Configuration
  - View Data by Array
  - View Data by Host

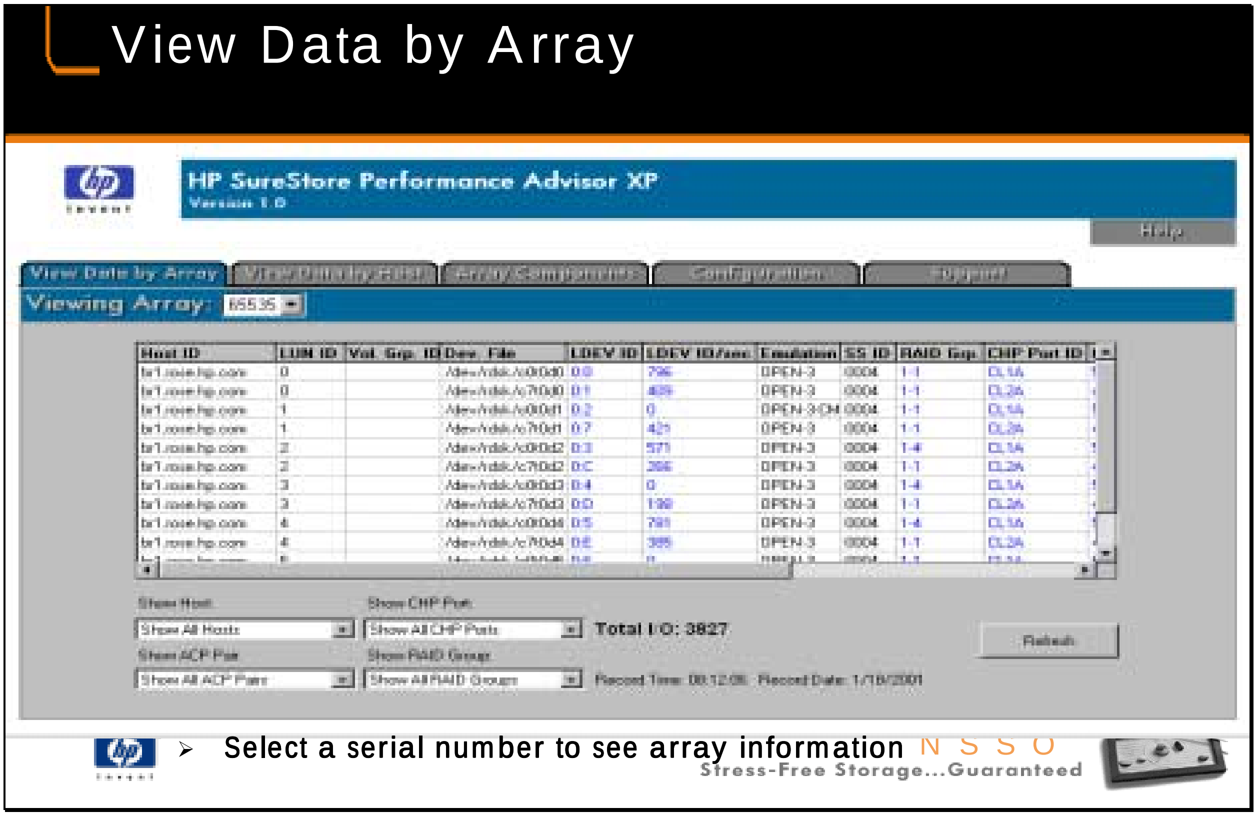
NSSO
Stress-Free Storage...Guaranteed

GUI Sample – Performance snapshot table

- By host
- By array (as shown)



NSSO
Stress-Free Storage...Guaranteed



Student Notes

Select an array serial number from the list of serial numbers. A table is generated showing a current snapshot of all connected XP LUNs with their performance numbers.

Filtering. To filter the information, select a hostname in the Show Hosts dropdown menu. You can view information associated with a particular host or with all hosts.

Refreshing. To retrieve new current information about the array from the database, click Refresh.

Date/Time values. The date and time values displayed at the bottom of the table indicate when the database was last refreshed.

Detailed information. If data is displayed in blue text, more specific information is available for that field. Double-click on the blue text to see a the detailed information. If data is displayed in black text, then no additional information is available for that field.

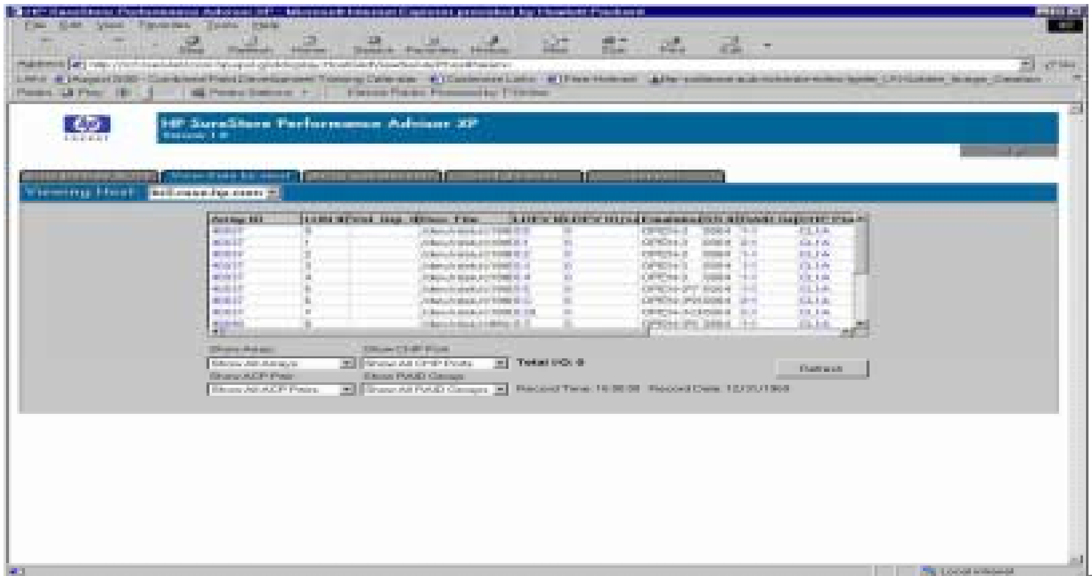
Historical data. Select a cell (for example an LDEV). Click on the column (for example, LDEV IO/Sec) to see historical data for the selected LDEV.

Viewing data. You can resize the columns to desired width.

Sorting data. To sort the data, click on the column heading. For reverse sorting, click on the column heading again.

For additional information, access the online help system

View Data by Host



N S S O
Stress-Free Storage...Guaranteed

Student Notes

Select a host stations from the list of connected hosts. A table is generated, showing a current snapshot of all connected XP LUNs, along with their performance numbers.

Filtering. To filter the information, select an array serial number in the Show Array dropdown menu. You can also select “Show All Arrays”.

Refreshing. To retrieve new current information about the array from the database, click Refresh.

Date/Time values. The date and time values displayed at the bottom of the table indicate when the database was last refreshed.

Detailed information. If data is displayed in blue text, more specific information is available for that field . Double-click on the blue text to see a the detailed information. This brings up a dialog box (the Array Components screen). If data is displayed in black text, then no additional information is available for that field.

Viewing data. You can resize the columns to desired width.

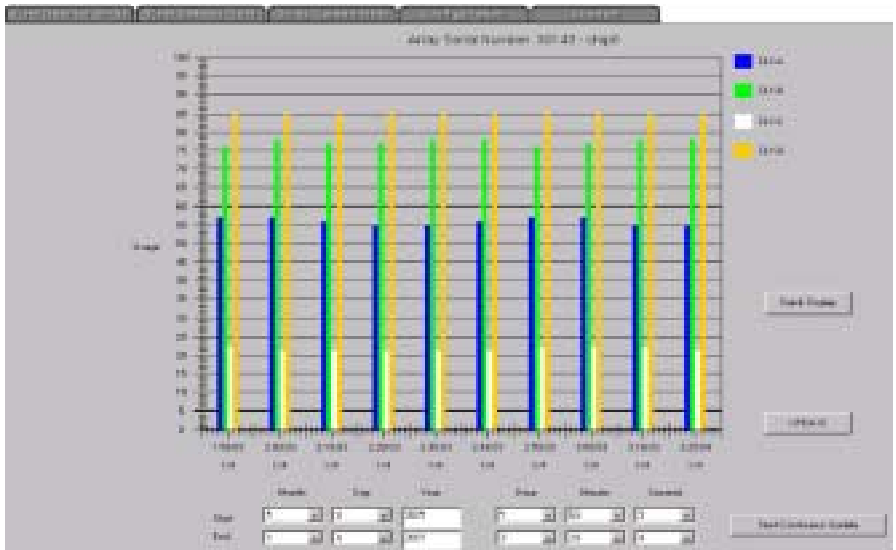
Sorting data. To sort the data, click on the column heading. For reverse sorting, click on the column heading again. (Example: If you want to see all LUNs connected to a given CHP port, sort by CHP port id (and array id). Then the table will be listed with all LUN/LDEVs belonging to the same CHP port grouped together).

For additional information, access the online help system

Host specific information – this is the menu used to configure the mapping of LUNs and LDEVs by host as well as providing details on device files.   A “Y” in the Req column causes a request for information to be issued while a “N” in the Rec column causes the information collection to be cancelled.

GUI Sample – Performance history graph

- Select item from snapshot
- Performance history available



N S S O
Stress-Free Storage...Guaranteed

Student Notes

The history screen displays historical data for many of the array components for a selected array.

The components are cache, ACP (array control processor), ACP pairs, ports, CHIPs (client-host interface processor), SM bus, CM bus, and LDEVs. Each component is broken down to its corresponding sub-components and the data is charted using a bar graph format. The sub-components vary depending upon the component that is being displayed.

Time is displayed along the bottom of the graph. The unit at the left of the bar graph depends upon the component displayed.

Start and end times. You can specify a start and end time for the sets of data to be displayed by changing the start and end times in the drop-down boxes and clicking the “Update” button.

Scroll bar. Only ten sets of data can be displayed in the graph area at any one time. Use the scroll bar to see data that does not fit on the screen.

Refreshing Information. As the screen is loaded, information is retrieved from the database to display the bar graph. The most current ten sets of data are initially displayed. After the screen is loaded, it will continue to refresh the data every 20 seconds.

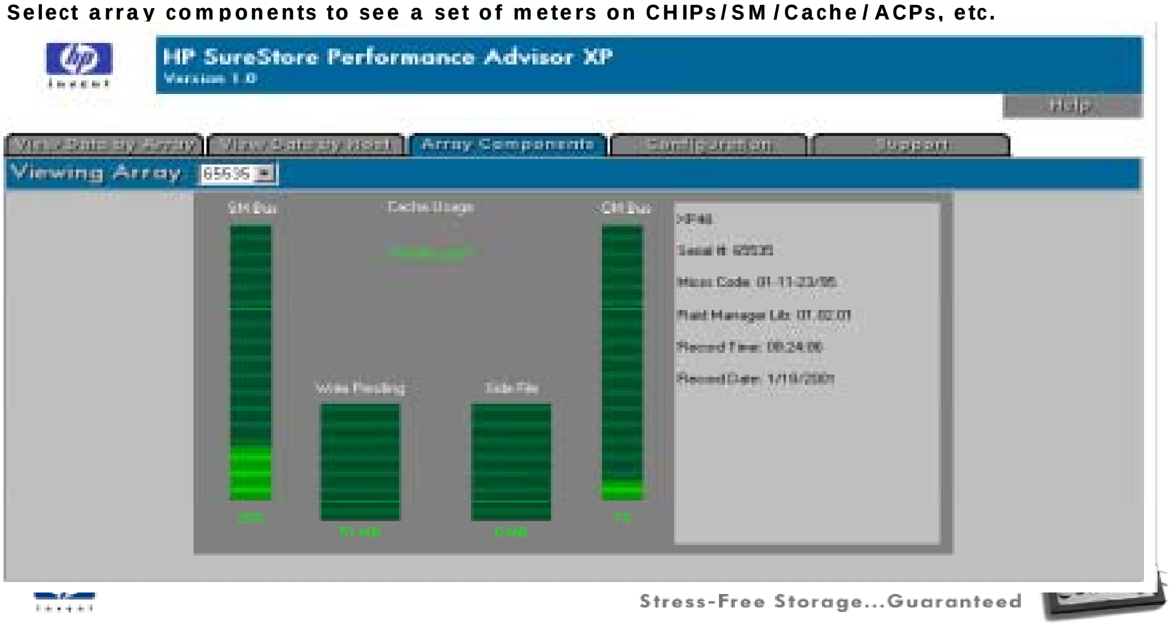
Bar colors. The color of the bar is associated with an element defined in the legend located at the side of the graph.

Stack Display/Parallel Display button. Click on this button to toggle the display between a stacked or parallel view of the data.

Update: This button allows you to gather the most recent range of data as specified in the end and start times.
Start Continuous Update: This button obtains the most recent data every 20 seconds.

GUI Sample – Array component screen

Select array components to see a set of meters on CHIPs / SM / Cache / ACPs, etc.



Student Notes

The Array Components screen displays a graphical representation of the data about the following: SM bus, CM bus, cache usage, write pending (the amount of data in cache that is waiting to be written to a disk), and side file usage (the part of the cache dedicated to HP SureStore Business Copy XP and HP Continuous Access XP).

Select an array serial number from the list of serial numbers.

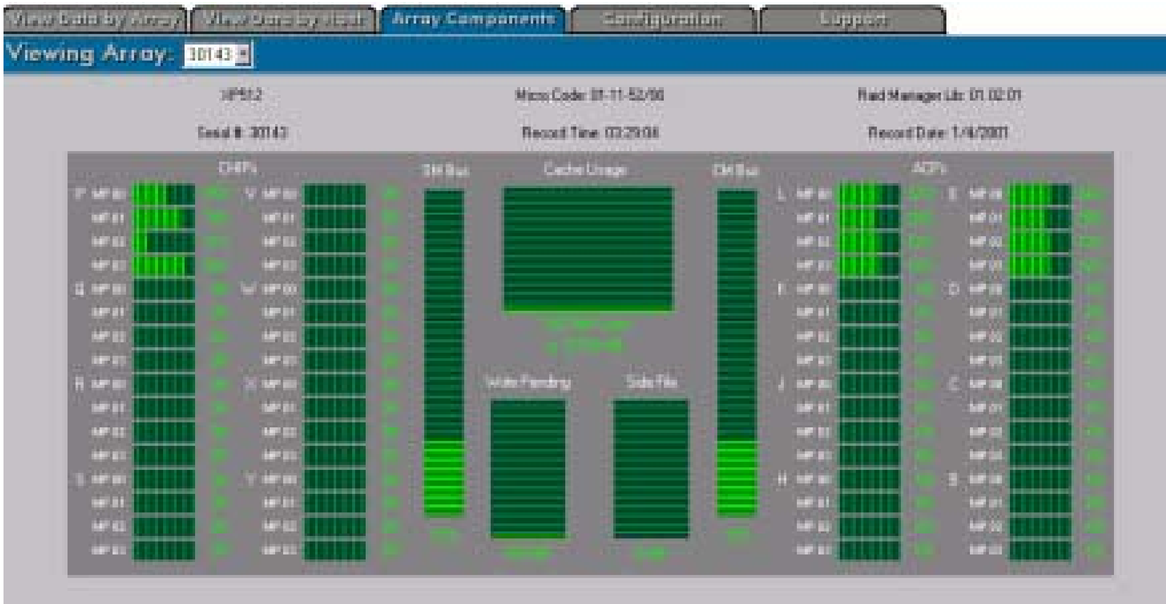
Information in this screen includes the array's serial number, the microcode version, the RAID Manager Library version number, and the date and time the displayed data was recorded.

Refreshing. As the Array Components screen is loaded, information from the database is retrieved to display in the table. The screen will refresh every 20 seconds as data is received from the database.

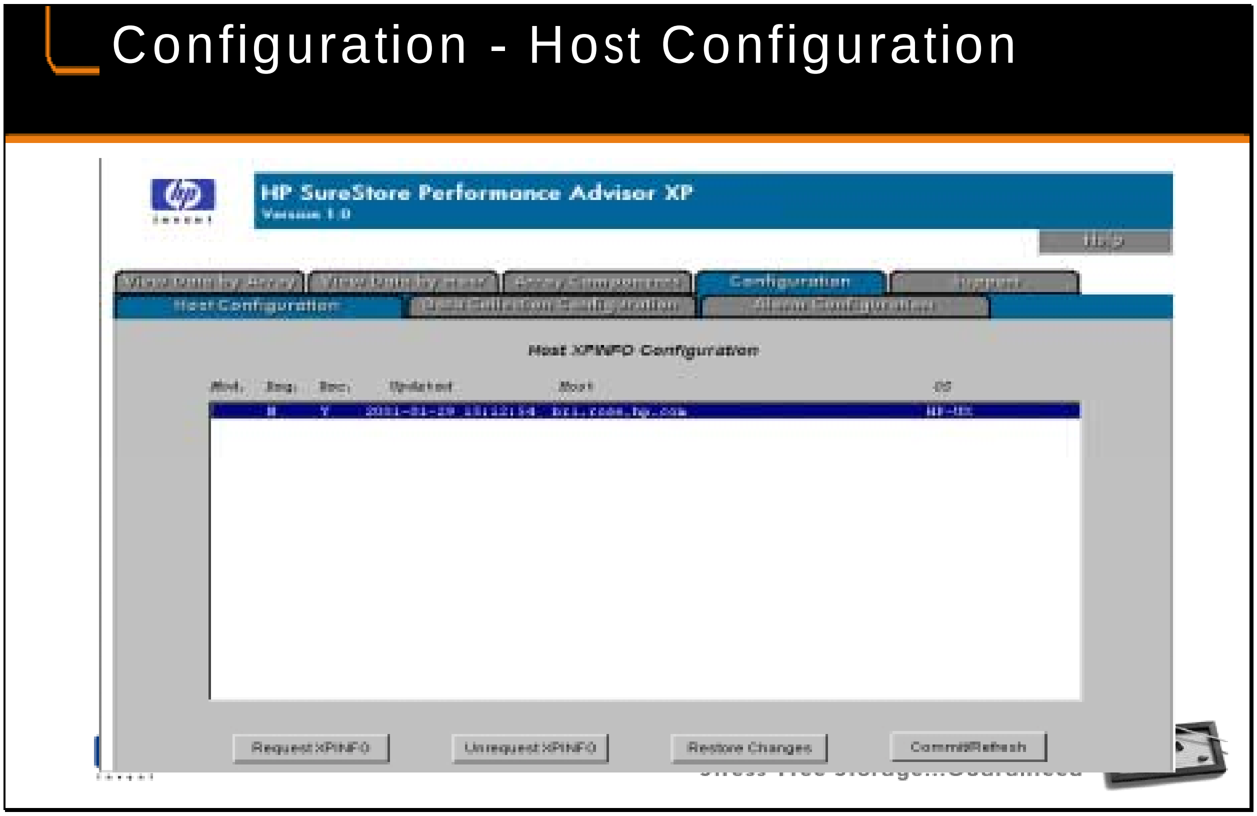
Accessing detailed data. To access more detailed information about a component, click on the component to open another browser window. For example, if you want more detailed information about a specific CHIP, double-click on the corresponding progress bar beneath the CHIPs heading. A new browser window will appear, providing a color-coded graph along with the dates and times that the information was gathered from the database.

Viewing the data. Most of the graphical data displayed is the percentage of usage of the component, is measured in I/Os per second over the maximum I/Os per second possible. The components that are measured are CHIP, ACP, Data Bus/SM Bus, and Control Bus/CM Bus. The cache is split into three categories: Cache Usage, MB Write Pending, and MB Side File Usage. Each of these categories is measured in megabytes used over the maximum cache memory the array can hold. The maximum cache memory that an XP256 can hold is 16 GB, while an XP48 can hold 16 GB and an XP512 can hold 32 GB.

Array Components (sample from future PA release 1.01.00)



Stress-Free Storage...Guaranteed



Student Notes

Use the Host Configuration screen to request or unrequest that Xpinfo be run on a specified host or to view updated information.

To request that Xpinfo be run on a particular host:

1. Click on the line for the host id you want to request data from. The line should be highlighted.

   If you want to request XpInfo on multiple hosts, click on all the additional hosts.
2. Click Request XPINFO. An asterisk will appear beside any host that was altered.
3. Click Commit/Refresh to send the request to the database.

To unrequest that Xpinfo be run on a particular host:

1. Click on the line that represents the host id that you don't want Xpinfo run on. The line should be highlighted. If you want to unrequest XpInfo on multiple hosts, click on all the additional hosts.
2. Click Unrequest XPINFO. An asterisk will appear beside any host that was altered.
3. Click Commit/Refresh to send the request to the database.

To view updated information:

1. Click Commit/Refresh. Once the request is received by the database, XpInfo is run on the selected host. When that host reports data back to the database the requested field

will be set
to false and the received field will be set to true. In this manner you can determine if XpInfo
was run. (You will need to click Commit/Refresh to see the updated fields.)
For additional information, access the online help system.

Online help

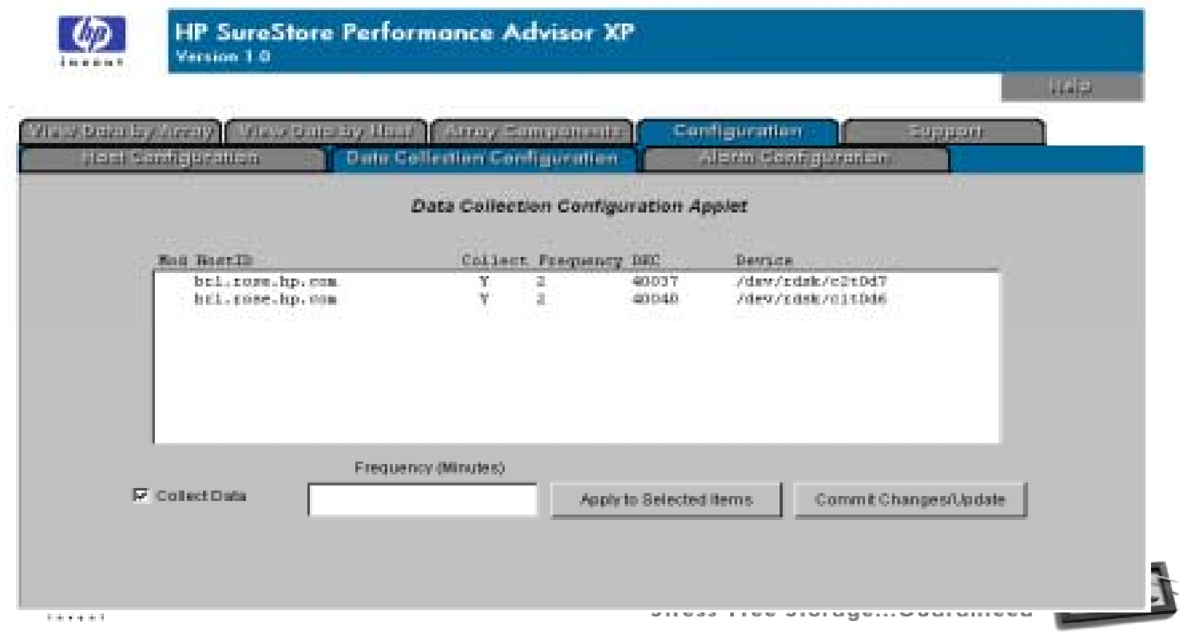
Select the Help button to launch the online help system. Help is context sensitive.
The main help page contains links to the following topics:

- Alarm Configuration
- Array Components
- Data Collection Configuration
- Historical Data
- Host Configuration
- View Data by Array
- View Data by Host

General information

Refreshing. To retrieve new current information from the database, click Refresh.
Date/Time values. The date and time values displayed at the bottom of various screens indicate when the database was last refreshed.

Configuration - Data Collection Configuration



Student Notes

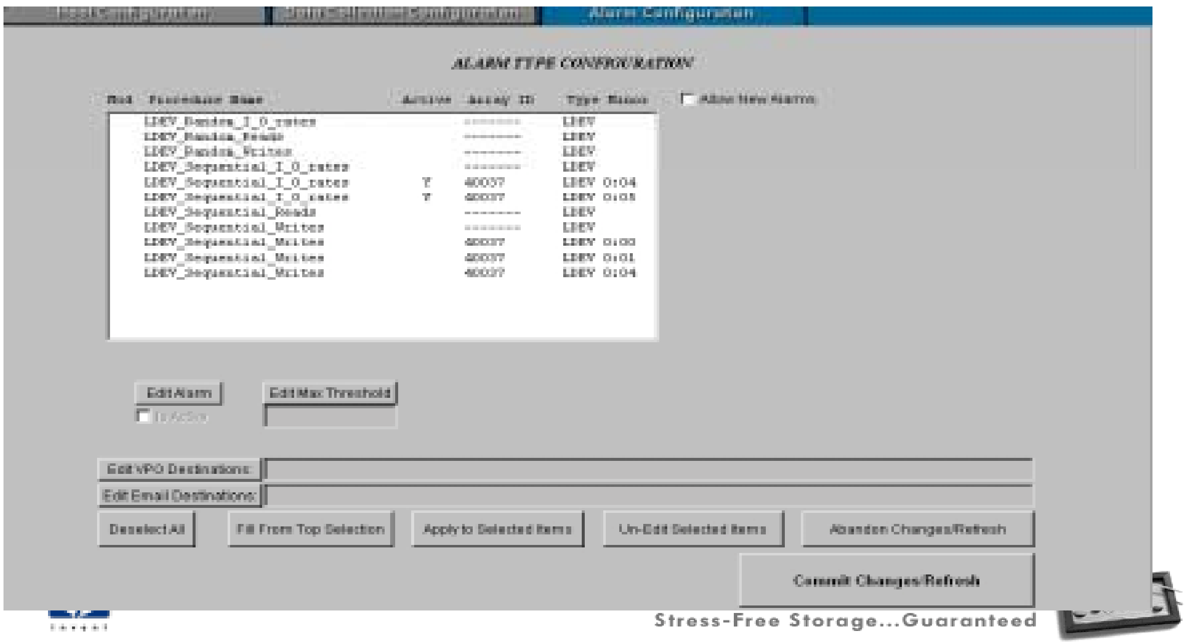
The Data Collection Configuration screen lets you specify or change the values you want to use for collecting performance and array configuration data from a particular XP disk array.

1. Click on a command device line in the window. The line will be highlighted and its values will be displayed in editing fields at the bottom of the screen.
2. To collect data, select the Collect Data checkbox. To stop collecting data, deselect the checkbox. Enter the desired number of seconds between data collections in the Frequency (seconds) field.
3. Highlight the command device in the table that you want the information applied to. An asterisk is be displayed in the Mod (modified) field.
4. After all changes are complete, click **Commit Changes/Update**. The asterisks in the Mod field will be deleted because the changes are now committed to the database.

Note: It is recommended that you set the data collection rate at one-hour intervals due to management station performance and database storage limitations. However, smaller intervals will be accepted. Follow the general rule of 1 minute per 100 LDEVs..

Note: Performance Advisor is currently limited to displaying 1000 LDEVs. Performance Advisor supports all LDEVs; only the display is limited.
For additional information, access the online help system.

Configuration - Alarm Configuration



Student Notes

Use the Alarm Configuration screen to create and configure alarm types for specified performance metrics on an XP disk arrays.

As new performance data is received by the Performance Advisor management station, the data is compared with the currently configured alarm types. A new instance of an alarm is created after comparing the received data from the host station(s) with the configured, active, alarm types.

When a value that exceeds the configured threshold) is detected, an alarm is generated and sent to the email or VPO management station that you have specified. No setup is required other than editing serverparameters.properties file to add a valid mail server.

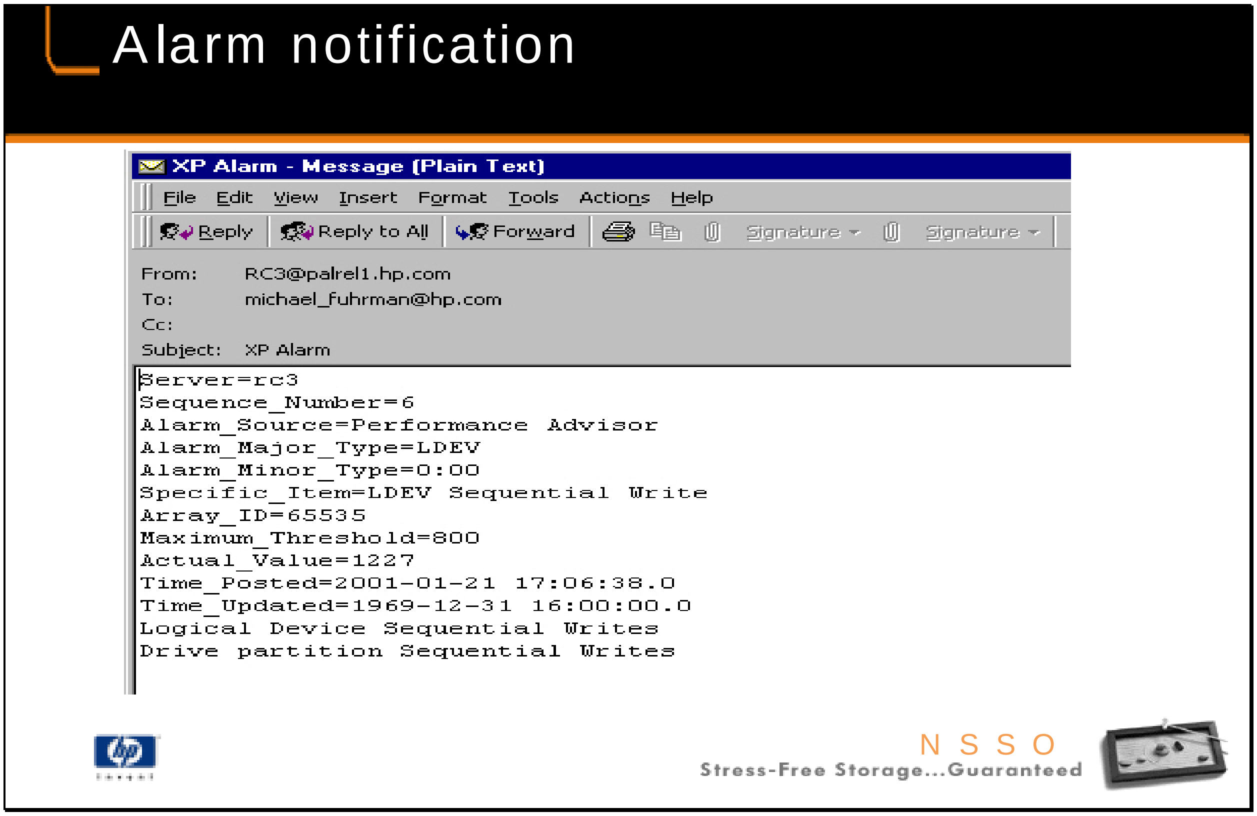
The alarm event information is sent as long as the minimum interval for successive notifications has been exceeded. For example, if the minimum reporting interval is five minutes for a given alarm, the closest amount of time between alarms being sent out is five minutes. So if a condition is out of range for 15 minutes then four alarms would be posted since the first alarm is sent out at 0 minutes.

NOTE: Due to firmware limitations, you cannot set alarms for ACP or CHIP utilization. This feature will be enabled with the next release of the HP Performance Advisor XP.

The alarm configuration screen has a variety of capabilities:

- Create alarm types from alarm definition templates and selected array components
- Configure the alarms to your own specifications
- View the state of the alarm configuration.
- Specify if an alarm is enabled or disabled

Specify the destination (Email or VPO)
Set the maximum thresholds at which alarm events will be generated



Student Notes

This is a sample of an alarm notification sent by e-mail.

It shows an LDEV sequential write alarm. The maximum threshold was set at 800. Actually performance produce a figure of 1227 writes, triggering the alarm.

Cautions

- The management station web server must be stopped before making any modifications of the configurable files located under the hpss subdirectory
- Current firmware limitations
  - ACP and CHIP utilization is not displayed
  - It is not possible to configure alarms for ACP and CHIP utilization
  - Feature will be enabled with the next release of Performance Advisor
  - When the firmware is upgraded, see if a Performance Advisor update if required

N S S O
Stress-Free Storage...Guaranteed

Student Notes

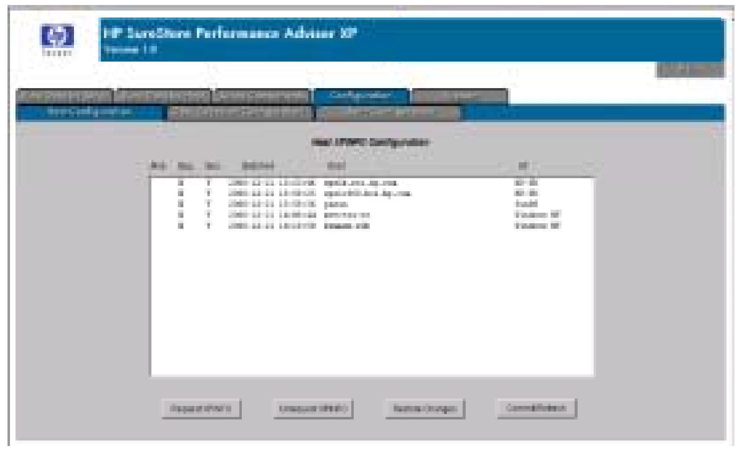
The management station web server must be stopped before any modifications are made to any of the configurable files located under the hpss subdirectory.
To stop the server. proceed as follows:

1. Select **Start/Settings/Control Panel/Services**.
2. Select **HPSS Apache** from the Services box (if it exists). Otherwise select **Apache**.
3. Click **Stop**. After modifications have been saved, click **Start**.

Current firmware limitations
ACP and CHIP utilization is not displayed, and it is not possible to configure alarms for ACP and CHIP utilization.
This feature will be enabled with the next release of the HP Performance Advisor XP.
When the firmware is modified, see if a Performance Advisor update is needed.

GUI Sample – Host configuration screen

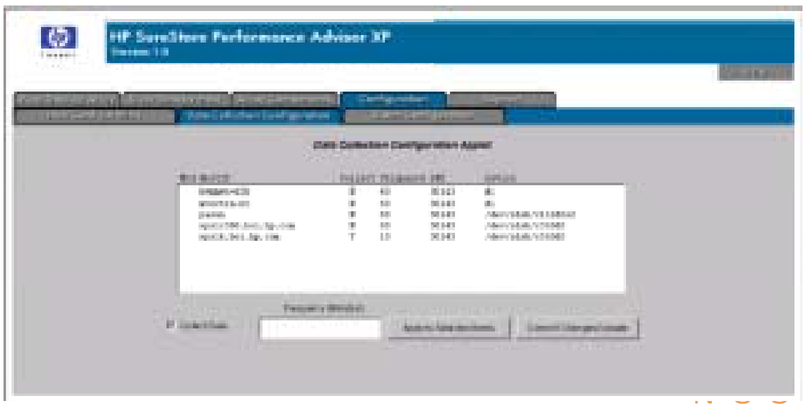
- Request xpinfo from host
- Unrequest xpinfo from host
- Restore changes/ Commit/Refresh



NSSO
Stress-Free Storage...Guaranteed

GUI Sample – Data collection configuration screen

- Controls which hosts collect performance data
- Controls the time intervals for the collection
  - Default interval is 60 minutes
  - At a minimum, allow 1 minute for each 100 LDEVs

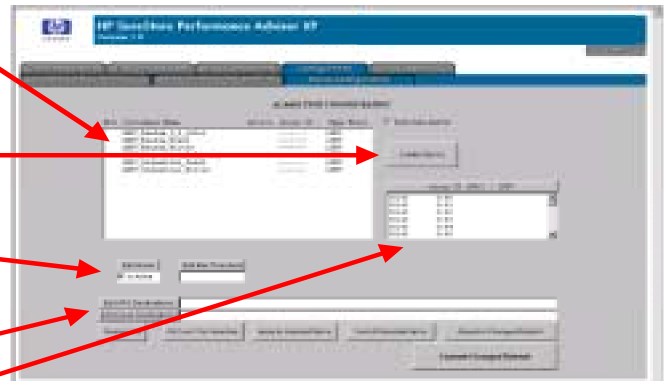


Stress-Free Storage...Guaranteed

Data Collection Configuration – this is the screen used to control which hosts collect performance data and the interval at which the information is collected.

GUI Sample – Performance alarm configuration screen

- Select alarm template
- Create alarm types for specified performance metrics
- Activate and edit alarm types
- Specify the destination (e-mail or VPO)
- LDEV table



NSSO
Stress-Free Storage...Guaranteed

Performance Advisor installation – preliminaries

- Confirm hardware and OS requirements
- Read the README.txt file on the CD
- Uninstall previous versions of Performance Advisor
- If Command View is running
  - All users should be logged off
  - The Apache Service should be stopped
- Verify that a command device has been created on the host which will act as a data collection host agent
- Make sure all hosts and management stations are networked and connected properly

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Make sure you meet the following hardware and operating system requirements.
If any copies of the HP Performance Advisor XP have been previously installed, you will need to run the uninstall program before reinstalling the software.
If Command View is running, make sure that all users are logged off and that the Apache Service is stopped before installing the Performance Advisor XP software.
Verify that a Command Device has been created.

Creating a command device using Command View

- Can be created using either Command View or Remote Control
- Using Command View:
  - Can be any OPEN-K/3/8/9/E device
  - Can use CVS to create a command device as small as 36MB
  - Use Command View's LUN Management tab
  - Note: this activity must be performed offline

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Customers who are currently running Raid Manager (Business Copy and Continuous Access) will already have command devices in place. They can be used for Performance Advisor.

Command View

1. Log in as administrator.
2. Connect to the desired array.
3. Click on the LUN Management tab.
4. Highlight the volume to be used as a command device.

CAUTION: *Be certain that there is no data on a volume you select as a command device.* Any data that might reside on the volume you select becomes inaccessible to the host. Also, make sure no file system has been mounted and no data is stored there.

5. Click the **Set/Release Cmd Dev** button.

Note: The volume designated as the command device is used only by the disk array and should note be used by the user. The command device can be any OPEN-K/3/8/9/E device that is accessible by the host.

The command device uses 16MB of space. The remaining volume space is reserved for Raid Manager and its utilities. You cannot use logical unit size expansion (LUSE) volumes as command devices; however, you can use Volume Size Configuration (VSC) devices as small as 36MB.

Installation – installing Performance Advisor on the Management Station

- Install Performance Advisor software on the Management Station before you install software on the hosts
- Windows NT is the only operating system supported by the Management Station
- Verify successful installation.

N S S O
Stress-Free Storage...Guaranteed

Student Notes

Insert the Performance Advisor XP software CD into the CD-ROM drive.
1. In the Start menu, select **Programs/Windows NT Explorer**.
2. In the Exploring window, go to the CD-ROM drive and double-click on the **CD-ROM** icon.
3. Double-click on the **jre –1.2.2.006-win.exe**. This will run the InstallShield® wizard,
which will automatically install the jre -1.2.2.006-win.exe application.
4. Click **Yes** to accept the terms of the Software License Agreement.
5. Click **Next** to install the application on the default drive, or specify an alternative drive.
6. When the setup is complete, reboot the computer.
7. Re-open the **PA** directory and double-click on the **PAServer.exe** icon. This will run the
InstallShield wizard, which will automatically install the PAServer.exe application.
8. The InstallShield wizard Welcome screen will appear.
9. Click **Next** to go to the license agreement screen.
10. Click **Yes** to accept the terms of the Software License Agreement.
11. Click **Next** to choose the destination Location.

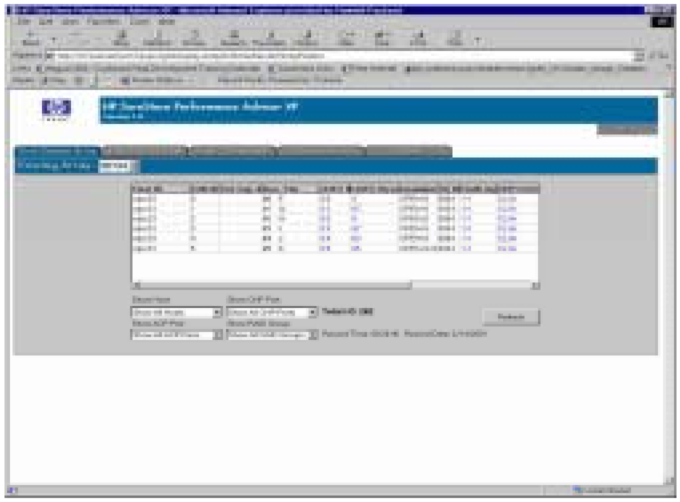
When the setup is complete, click **Finish**.

Verify that the HP Performance Advisor management station software has been successfully installed.

To do so, confirm that two major services, HPSS Apache web server and HPSS SOLIDSERVICE, are operating:

1. In the Start menu, select **Settings/Control Panel**.
2. In the Control Panel, double-click on the **Services** icon. The Services dialog box will appear.
3. If HP Surestore Performance Advisor XP has been installed, then the status of HPSS Apache
   and HPSS Solid should be in the Started mode.
4. Close the Services dialog box.

Starting HP Surestore Performance Advisor XP



- Open your browser and type: http://<systemname>/pa
- Enter ID and password (only necessary in the config screen)

NSSO
Stress-Free Storage...Guaranteed

Student Notes

1. To start Performance Advisor software, type:
    http://<systemname>/pa
2. Enter the password for the Performance Advisor. The user name is confmonxp and the password is redstar.
    Note: Refer to the Performance Advisor configuration document (PAConfig.html) for instructions about changing the password.
3. Go to the data collection configuration screen. To begin collecting data, choose a command device on the host you want to use. Perform data collection on only one host per array.

Note: It is recommended that you set the data collection rate at one-hour intervals due to management station performance and database storage limitations. Smaller intervals will be accepted. Follow the general rule of 1 minute per 100 LDEVs for the management station (PC) to keep up with collection.

Performance Advisor collects performance data on all LDEVs in the array. It is not limited to the number of LDEVs the host station is mapped to use.

Note: History records are retained for 1 month. However, you can limit the amount of disk array history recorded.

Performance Manager?

- Performance Manager XP (PMXP) is primarily a workload monitoring and display tool for the SureStore E Disk Array XP Family of arrays.
- PMXP allows the user to see trends and peaks as well as tune the array's duplex write line (DWL) for optimal performance.
- Performance Manager XP is a performance/workload management tool - not a performance benchmarking tool.
- A customer should purchase this tool if they want to monitor the XP Family of arrays's workload for load balancing and/or DWL tuning.

NSSO
Stress-Free Storage...Guaranteed

Performance Manager XP (PM) is a software application which collects and displays disk array data, and allows you to execute performance management functions. PM runs on a PC, which much like the remote console, is also attached to the internal (private LAN) of the XP Family of arrays subsystem. Using an internal LAN, the performance workstation can be attached to up to eight (8) disk arrays.

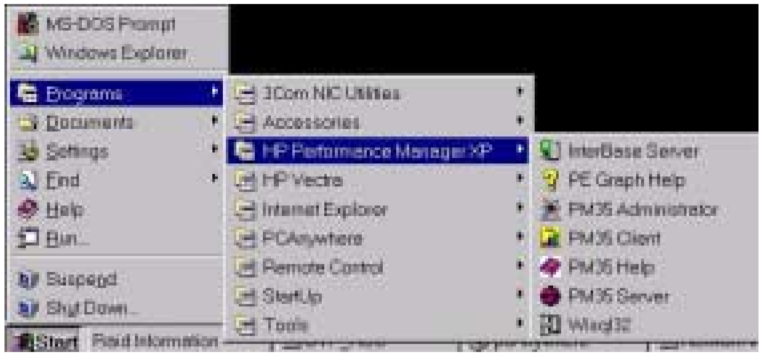
Performance Manager collects and displays detailed performance and usage data for the disk arrays. It monitors hardware performance, cache performance, and I/O statistics, and displays real-time and historical data as graphs which highlight key information such as peaks and trends. PMXP displays information about LDEVs not LUNs or Volumes.

What is an LDEV? The array is made up of disk that are grouped together into array groups of four disks each. The array groups are then set up with a RAID level of 1 or 5 for data redundancy. Next, the array group (with a RAID level) is partitioned into smaller logical disks called a LDEV. The next step is mapping the LDEV to a port of the array for use by a host, thereby creating a LUN or Volume. NOTE: an LDEV may be mapped to more than one port and referred to by more than one LUN number (from the external hosts perspective).

Simplistic LDEV Example:

If a disk is 1GByte in size, and four of these disks are in a RAID 5 array group (which means 3/4 of the array group disk space is available for data and 1/4 of the array group disk space is for redundancy), then there is 3 GBytes available for LDEVs. I could now create three 1GByte LDEVs or maybe two 1.5GByte LDEVs - all depending on how the available space is partitioned.

Starting the PMXP Administrator



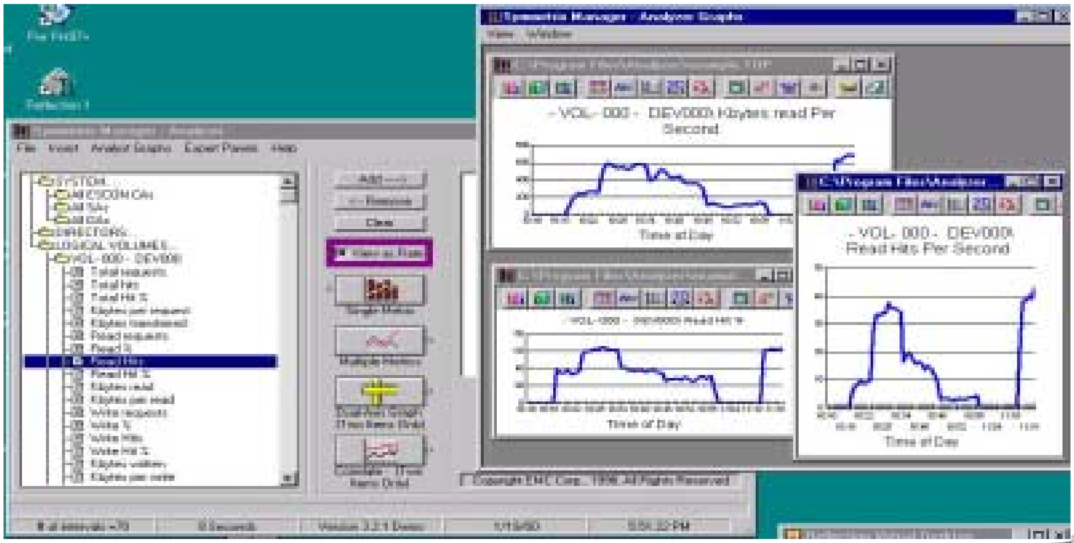
Assuming the PMXP Utility is installed in the default location, first start the PMXP Administrator:

START ➜ Programs ➜ HP Performance Manager XP
➜ PM35 Administrator

NSSO
Stress-Free Storage...Guaranteed

Run the PM administrator to create PM database files and configure environment options (for example startup mode and date display format).

Competitive Product



N S S O
Stress-Free Storage...Guaranteed

EMC's Workload Analyzer

Competitive Comparison

Features	Performance Advisor	EMC	Hitachi Data Systems
Pre-defined performance thresholds	☑	☑	☑
CLI on Host & Management Station	☑	No	No
Manage multiple arrays in multiple locations	☑	Separate version of Control Center required for each O/S	Manage 1 array at single location
Security	Internet Standard	Unknown	Private
Backend Channel Utilization	1H'01	☑	☑
Detail Cache Utilization	1H'01	☑	☑
LUN Status	1H'01	☑	☑
Integration into 3rd party applications	☑	MVS Agent	No

NSSO
Stress-Free Storage...Guaranteed

EMC: Performance metrics comes from two solutions: Symmetrix Manager (Control Center) and Workload Analyzer. Symmetrix Manager is the "core" of Control Center so the customer receives performance tuning capabilities automatically upon purchase. Information provided via Symmetrix Manager is: front-end directors, volume performance for individual disks, thresholds on cache & write ratios, disk performance against pre-defined upper/lower thresholds for I/O sec. For analysis of this data, the customer needs to purchase Workload Analyzer – which is not available as a standalone solution. Workload Analyzer provides: trend analysis, collects/analyzes/graphs historical data and monitors system I/O rates & throughput. Workload Analyzer is positioned as a complement to Symmetrix Manager and a "plug-in" to Control Center. EMC currently has an agent for MVS systems.

EMC's advantage over HP:
More robust data collection (until 1 H'01) on back-end directors and graphing
MVS agent.
HP's advantage over EMC:
Integration into VantagePoint Operations (1/01) and MeasureWare
Performance Advisor management station is platform independent
Internet standard security
No mention of security in EMC documentation
Hosts supported: HP-UX, NT, Solaris

The structure of Control Center requires the customer to have a separate version of CC for each different host O/S (Unix, Windows NT & Window NT w/browser). The only solution that provides support for Workload Analyzer is the Windows NT w/browser).

HDS:

HDS provides customers with a suite of solutions called Resource Manager. Their version of Performance Manager is called Graph Track, which is part of the Resource Manager bundle. The bundle is priced very aggressively – almost to the point of being a giveaway.

HDS advantage over HP:

Broader host support – all platforms supported by the Hitachi version of the arrays

More detailed information on LUN status, cache utilization, logical drive performance and front-end response times

Price

HP's advantage over HDS:

Usefulness of performance data. Graph Track (and Performance Manager) currently have a synchronization problem which results in the performance data not matching host activities.

Command line interface on host and management station

Ability to manage multiple device from multiple locations

Internet Standard security vs. private LAN

Integration into MeasureWare and Command View vs. Remote Control U/I

Performance Advisor

Orderable December 1st

Shipment begin1/2/02

WW Reference Price: $25K

Initial release limited to 1K LDEVs

Support for 4K LDEVs (4/01)

Performance Manager

On 11/1/00 WW Reference Price will be increased to $25K

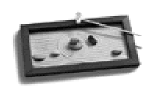
Product is "functionally complete"

Discontinuance planned for 3/1/01

Review of Key Points

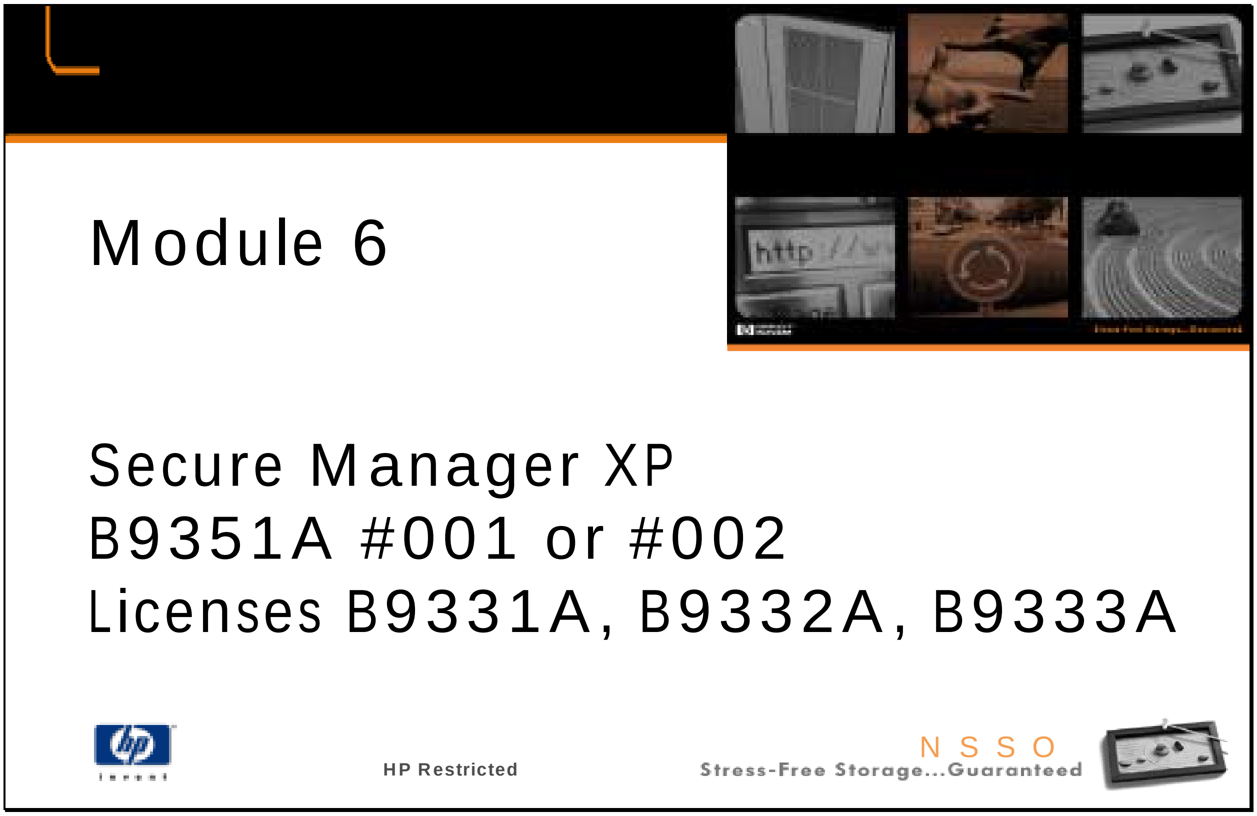
- Definition & Features
- Configuration & Software/Hardware Requirements
- GUI, Interface and Environment
- Configuring and Monitoring Capabilities
- Competitive Information

N S S O
Stress-Free Storage...Guaranteed



Module Wrap-up

N S S O
Stress-Free Storage...Guaranteed



Student Notes

LUN Security allows you to restrict LUN availability to a specified open system host (or hosts), using the host's world wide name (WWN). Other open system hosts cannot access the secured LUN and its data.

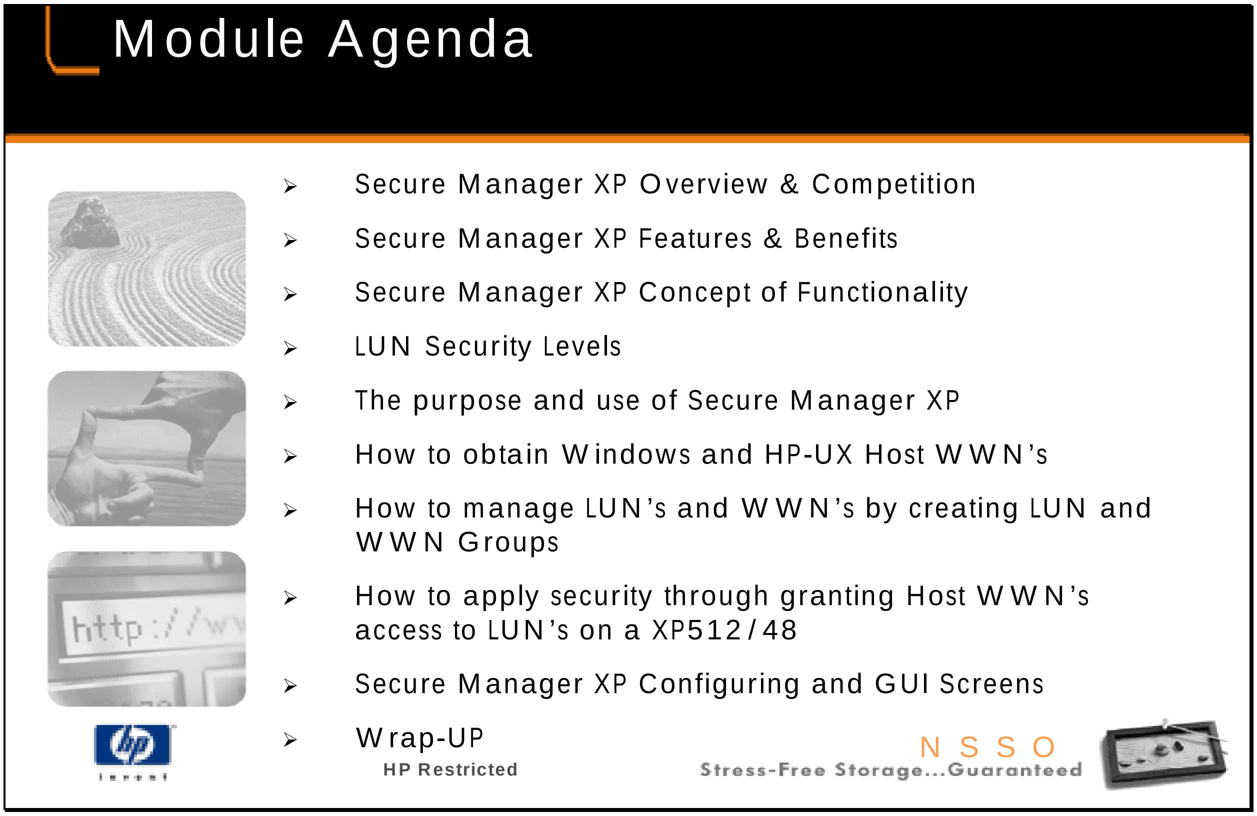
You can secure some or all of the LUNs on a port. You can create LUN groups to help accomplish this task. A LUN group allows you to assign specified LUNs to a single group name.

A WWN group gives every host in the specified WWN group access to a specified LUN or group of LUNs.

A WWN group allows you to assign up to 32 WWNs to a single group (XP256) or up to 256 WWNs to a single group (XP512 and XP48). You can assign more than one WWN to a port and one WWN to several ports.

LUN Security allows the following functions:

• enabling and disabling the security for Fibre Channel ports
• creating, modifying, and deleting a LUN group
• changing the LUN group nickname
• defining the WWN
• changing the WWN nickname and WWN
• creating, modifying, and deleting WWN groups
• changing the WWN group nickname
• setting the access permission for LUNs and LUN groups



Secure Manager allows you to restrict a LUN's availability to a specified open system host or hosts, using the host's worldwide name (WWN). Other open system hosts cannot access the secured LUN and its data. Secure Manager can be turned on or off at the port level. If you do not want to use Secure Manager on a particular port you can disable the feature so that the LUNs are not restricted to a particular host or group of hosts. If you
enable Secure Manager on a port, that specific port is restricted to a particular host or group of hosts. You can assign a WWN to as many ports as you want, and you can assign more than one WWN to each port. You can also change the WWN access for any port without disrupting the settings of that port.

Secure Manager XP Overview

- Secure Manager XP - Product Number B9351A allows you to restrict a LUNs availability to a specified host or hosts in a SAN, using the host's World Wide Name (WWN). Other hosts cannot access the secured LUN and its data.
- Secure Manager XP is part of the LUN Configuration Manager XP software and runs on the Command View Console, Remote Console and Service Processor (SVP).
- Secure Manager XP is supported only for open system Fibre Channel ports (native FC). It is not supported for mainframes or SCSI connectivity (SCSI MUX or SCSI CHIP on the XP256).

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

The HP Secure Manager XP feature of the disk array is part of the LUN Configuration Manager XP software. It is available only for open system fibre channel ports. Secure Manager allows you to restrict availability to a specified open system host or hosts, using the host's World Wide Name (WWN). Other open system hosts cannot access the secured LUN and it's associated data. You can secure some or all of the LUNs on a port. You can create LUN groups to help accomplish this task. A LUN group allows you to assign specified LUNs to a single group name. A WWN group allows you to assign up to 32 WWNs to a single group. A WWN group gives every host in the specified WWN group access to a specified LUN or group of LUNs. You can assign more than one WWN to a port and one WWN to several ports.Secure Manger XP allows you to restrict a LUNs availability to a specified host or hosts. You can prevent hosts from being able to see and access certain LUNs on the XP 256, using the host worldwide name and we'll explain what that means later. Secure Manager XP is enabled by a separate key and invoked through the LUN Configuration Manager product. List price for Secure Manger XP is $15,000.00.
You are ensured that hosts that do not have this access enabled are prohibited to the LUN(s) and the data that is contained in that LUN. Basically, the LUNs and their associated data are secured from those others hosts that are not allowed permission. Secure Manger XP is a part of the LUN Configuration manager software that runs on the remote console PC and on the service processor, the SVP, which is a part of the sub-system. Secure Manager XP is

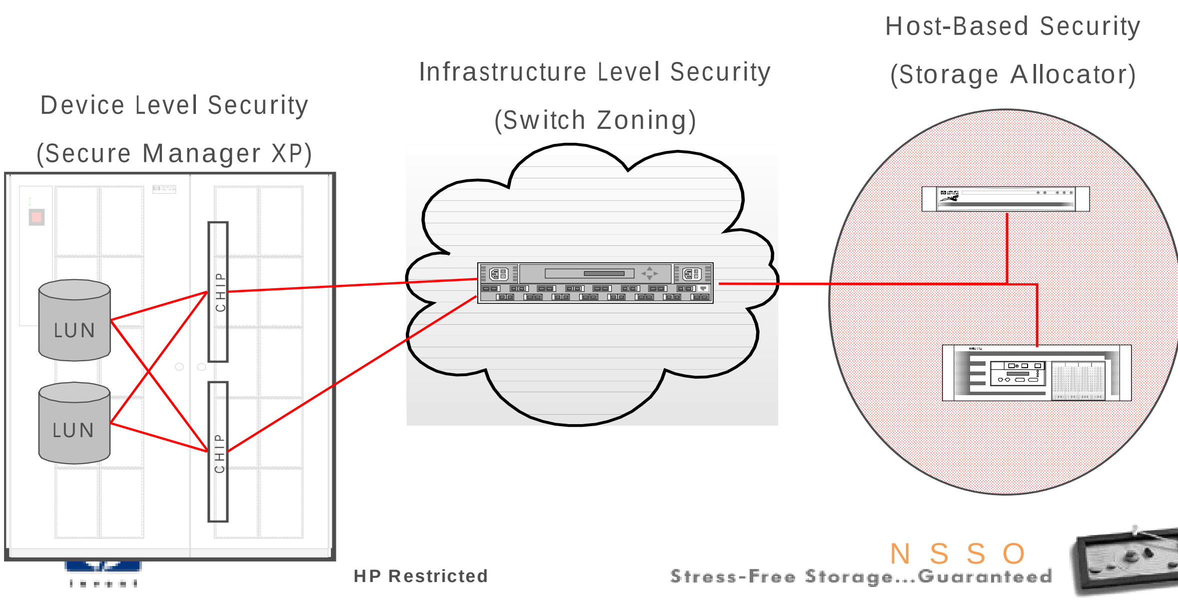
supported only for open system fiber channel ports (i.e. native fiber channel). So, it's not going to work over your SCSI. This product is just for fiber channel only using the pWWW addressing schema.

Secure Manager allows you to restrict a LUN's availability to a specified open system host or hosts, using the host's worldwide name (WWN). Other open system hosts cannot access the secured LUN and its data.

Secure Manager can be turned on or off at the port level. If you do not want to use Secure Manager on a particular port you can disable the feature so that the LUNs are not restricted to a particular host or group of hosts. If you enable Secure Manager on a port, that specific port is restricted to a particular host or group of hosts. You can assign a WWN to as many ports as you want, and you can assign more than one WWN to each port. You can also change the WWN access for any port without disrupting the settings of that port.

Because many WWNs can access each port and the same WWNs may go to additional ports in the same subsystem, Secure Manager allows you to create LUN groups and WWN groups so you can more easily manage your disk array. A LUN group allows you to assign specified LUNs to a single group name. A WWN group allows you to assign multiple WWNs to a single group. A WWN group gives every host in the specified WWN group access to the specified LUN or group of LUNs.

LUN Security Choices – Three Levels

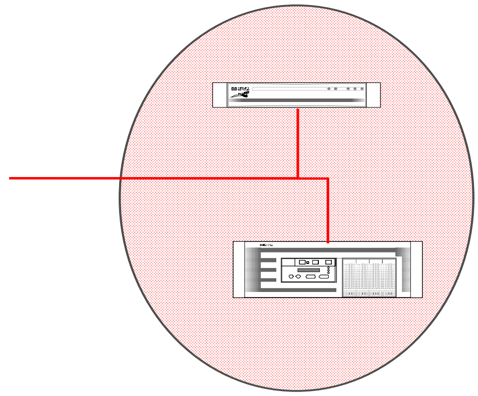


HP Restricted

NSSO
Stress-Free Storage...Guaranteed

LUN Security – Storage Allocator

Host Level Security



__Advantages:__

- Independent of target devices.
- Mixed heterogeneous devices.
- Ease the management of a storage pool.

__Disadvantages:__

- Management software is host/HBA specific

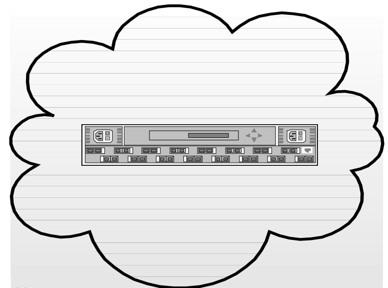
and must be present on all hosts in the SAN to be effective. It must be a Co-operated, trusted environment.

- Not a bullet proof security scheme - a host can be plugged into the SAN where it can corrupt data. This scenario is particularly likely in a multiple campus situation.

HP Restricted

NSSO Stress-Free Storage...Guaranteed

LUN Security – Switch Zoning



Infrastructure Level (Switch Zoning)

Advantages:

- Independent of Hosts and target devices.
- Safeguard SAN against unauthorized hosts.

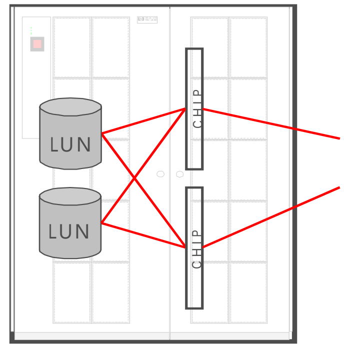
Disadvantages:

- Granularity is at port and node level (Not LUN level).
- When multiple, separated SAN is connected to the storage, zoning information may not be consistent.
- Most switch vendors use soft zoning when flexibility is required (i.e. moving physical cable connection). This level of security can be breached.

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

LUN Security – Secure Manager XP



Device Level (Secure Manager XP)

<u>Advantages:</u>

- Best granularity - LUN level.
- Best safe guarded – from anywhere.
- Control LUN access in a SAN

<u>Disadvantages:</u>

- Device dependent - low end array or JBOD may not support this function.
- Administration may become cumbersome for large node count (e.g. 200 NT servers sharing a LUN for mail database).

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Secure Manager XP Capacity

The XP256 and XP512/48 have different capacities for LUN's, LUN Groups, WWN's and WWN Groups per port.

Capacity	XP256	XP512/48
Maximum WWN's in a LUN group	31	255
Maximum LUN's in a LUN group	120	256
Maximum WWN's per port	31	255
Maximum LUN groups per port	16	128
Maximum WWN groups per port	15	127

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Using caution, Secure Manager also allows you to set the security parameters while you are online, but be very careful if you choose to do so. If any of the Secure Manager parameters are incorrect, the operation could be disruptive to the open-system host I/O activity. Always verify the Secure Manager parameters before entering them.

EMC Symmetrix Volume Logix

- Competition: EMC Volume Logix controls access to specific volumes for each server on the SAN
- Resides on a workstation or server
- Command line interface to assign Symmetrix volumes to each host on the hub or switch
- Proprietary, supports only Symmetrix subsystems

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

HP has three software components that put HP on parity with EMC:

SAN Manager LM - FC60 supported
Secure Manager XP
zoning on a switch

Managing Data Access in Enterprise Storage Networks

First software to manage data access.

Managing the complexity of connecting and accessing hundreds of Windows NT® or UNIX® servers in an enterprise requires highly specialized control. Without providing a way to control information access, there is the potential for servers to gain unauthorized entry to volumes of sensitive information. And as the enterprise evolves and grows, storage, servers, and access privileges must often be reallocated. To date, no storage or Fibre Channel supplier has been able to provide a solution that can span a heterogeneous environment.
EMC® Volume Logix™ is the first software solution to provide flexible, intuitive, easy-to-implement data protection and volume access control in a Fibre Channel-based storage network topology. It allows storage administrators of EMC Enterprise Storage-based information to centrally manage and control heterogeneous volume access in Enterprise Storage Networks (ESN). ESN extends the value of enterprise storage throughout more of the organization and enables your information infrastructure to evolve with your enterprise. ESN

simplifies access to the exponential growth of information while Volume Logix optimizes the information protection of those assets.
EMC Volume Logix provides storage-centric access control of storage volumes, in both mixed and homogeneous platform environments. When used with host computers that are connected via a Fibre Channel-Switch Fabric (FC-SW) or a Fibre Channel-Arbitrated Loop (FC-AL), Volume Logix enhances information protection through more of the organization. Protection is provided on the Symmetrix® and managed by a Graphical User Interface (GUI) on a management server that is part of the Enterprise Storage Network.

<u>EMC Volume Logix provides secure shared or selective access to designated logical Symmetrix data volumes for authorized users.</u>

Best protection through Symmetrix storage-based software.
By protecting data at the storage-level, Volume Logix provides centralized control of data, taking the burden of information protection off departments while accommodating administrators' ability to freely change host and operating systems environments. Volume Logix even shields data from inadvertent access from host servers trying to access protected volumes. Storage-based Volume Logix software eliminates high management and overhead costs associated with volume access and control. With Volume Logix, administrators manage data access once, at the storage-level instead of at every server in the environment.

Volume access control for mixed environments.
Volume Logix allows multiple UNIX, NT, and Novell server hosts to share the same Symmetrix Fibre Adapter port while providing multiple volume access controls and data protection. See the complete list of supported platforms at the end of this document.

No agent software on servers.
With Volume Logix, the software intelligence resides on Symmetrix and includes a database that tracks and controls host volume access. This design feature eliminates the necessity for device polling and host-based I/O filtering, imposing no impact on server CPU cycles. Host servers can run Volume Logix host utilities to simplify and expedite volume access configurations with the Symmetrix.
EMC Volume Logix provides secure shared or selective access to designated logical Symmetrix data volumes for authorized users.

<u>Improves performance through easy-to-read, easy-to-use display windows.</u>

Volume Logix protects data in Enterprise Storage Networks.
EMC Volume Logix manages the way servers use centralized consolidated storage on an Enterprise Storage Network. With Volume Logix, storage administrators can efficiently allocate Symmetrix storage among host servers located on a Fibre Channel switch or hub, ensuring the integrity of the data.

Together, Volume Logix and Symmetrix provide the robust data protection needed to effectively implement an Enterprise Storage Network. Today, organizations can take advantage of the additional connectivity, bandwidth, and extended distance capability offered by ESNs to connect isolated servers to storage throughout the enterprise. This creates an organizational environment where information is centralized and efficiently managed, protected, and available for sharing. Servers are deployed (and redeployed) at the department level, where they are most needed and accessible. As a result, these organizations can experience The EMC Effect™: the competitive advantage gained through EMC Enterprise Storage™ systems, software solutions, and services when an organization puts all of its critical information to work on meeting its most aggressive goals.

Ease of use.
Volume Logix features an intuitive graphical user interface (GUI) for Microsoft Windows NT, HP-UX, IBM AIX, and Sun Solaris platforms. This ease of use feature streamlines system administrator's volume access configuration tasks while providing a complete, easy to understand graphical topology of the Symmetrix volume assignments to the host computers. The GUI also contains a task list manager to flexibly stage configuration changes. Additionally, a Command Line Interface (CLI) is also supported and suitable for scripts that can easily automate procedures. Both GUI and CLI interfaces provide the best choice selection for the types of tasks to be performed when configuring volume access.

Volume Logix components and supported platforms
Volume Logix Administrator (VLA)*

VLA is a systems administrator tool used to configure Symmetrix device access control for either mixed or homogeneous Fibre Channel-based storage networks.

Supported Platforms: (GUI)
Windows NT 4.0 Intel-based SP3, SP4, SP5
Sun Solaris 2.51, 2.6
HP & NEC HP-UX 10.20, 11.0
IBM & Bull AIX 4.3.2, 4.3.3
Supported Platforms: (CLI)
Windows NT 4.0 Intel-based SP3, SP4, SP5
Sun Solaris 2.51, 2.6
HP & NEC HP-UX 10.20, 11.0
IBM & Bull AIX 4.3.2, 4.3.3

Volume Logix Host Utilities

Provides automatic World Wide Name (WWN) mapping to an ASCII World Wide Name (AWWN)

Supported Platforms:
Windows NT 4.0-Intel Based SP3, SP4, SP5
Sun Solaris 2.51, 2.6

HP & NEC HP-UX 10.20, 11.0
IBM & Bull AIX 4.3.2, 4.3.3

Supported Platforms in a Fibre Channel Storage Network

Manages volumes each computer is allowed to access on each Symmetrix FA by host computer HBA

Supported Platforms:

Windows NT 4.0 Intel-based SP3, SP4, SP5
Sun Solaris 2.51, 2.6
HP & NEC HP-UX 10.20, 11.0
IBM & Bull AIX 4.3.2, 4.3.3
Sequent DYNIX /ptx 4.4.4
Siemens Reliant UNIX 5.44C1014
Novell Netware 5.0.0
Linux Redhat 2.2.9, 2.2.11
Compaq Tru64 UNIX 4.0F
*Minimum one Volume Logix Administrator for one or more Symmetrix systems.

Highlights

GUI & Command Line Interface (CLI)
User interfaces provide best choice selection for types of tasks to be performed.

Java-based Graphical User Interface (GUI)
GUI simplifies system administrator's tasks and allows one or more operations on one or more Symmetrix systems.

Command Line Interface (CLI)
CLI is suitable for scripts to automate procedures.

Task List Manager
Streamlines multiple configuration changes before they are made.

Access Control in heterogeneous & homogeneous Fibre Channel Storage Network environments
Data protection across multiple UNIX & Windows NT platforms with a single product.

Online management
Dynamic access control allows configuration changes "on the fly" from a single management station.

Data Sheet
L814.1 - 9/99

What Benefits Does Secure Manager XP Have?

- Limit the LUNs and data a host can see
  - Control LUN access in a SAN
  - Manage a pool of LUNs between servers
- NT grabs every visible LUN. Secure Manager allows you to specify which NT-server should have access to which LUN
- LUNs do not need to be contiguous (EMC Volume Logix does need contiguous LUNs, a competitive advantage)

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

EMC targets LUNs entirely in the EMC array, but does not secure other data, for example, FC60 disk arrays connected to a SAN. With HP, FC60 data can be masked with zoning on a switch. EMC need contiguous LUNs, but does not allow the default sharing of LUN 0 data. What benefits does Secure Manager XP have? It allows you to limit the LUNs and data a host can see. You can control the LUN access in a SAN, a storage area network by using the SAN Manager LM product. And you manage a pool of LUNs between servers. For NT grabs every visible LUN it happens to see, so with Secure Manger you are allowed to specify which NT server should have access to which LUN. And one other benefit of Secure Manager is the LUNs do not to be contiguous. So, you can have LUNs in various areas and you can secure which server can see those particular LUNs.

Secure Manager XP

- Secure Manager XP uses a permission matrix to associate a host pWWN (port World Wide Name on a HBA) to LUNs defined behind a XP Family of Array's FC port.
- The WWN grants access from selected Fibre Channel host connections to selected XP Family of Array's LUNs on a per port basis.
- If the host WWN is not on the list, that host has no access when Secure manager is enabled.
- Access is either allowed or denied. You cannot restrict the type of access, for example, setting read-only permissions for a LUN.

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Secure Manager XP uses a permission matrix to associate a host, port worldwide name (pWWN), which is on the host bust adapter, HBA, and associates that to LUNs defined behind an XP 256 fiber channel port. The worldwide name grants access from selected fiber channel host connections to selected XP Family of Array's LUNs on a per port basis. And we will get into examples of that later. If the host worldwide name is not on the list, then that host has no access when Secure Manager is enabled. You also have the option of turning Secure Manager off if you want to enable access. Access is either allowed or denied. You cannot restrict the type of access. For example, by setting "read-only" in the permissions for a LUN. So, it's either access is allowed or not allowed, nothing in between.

Changes with Feb 2000 Release

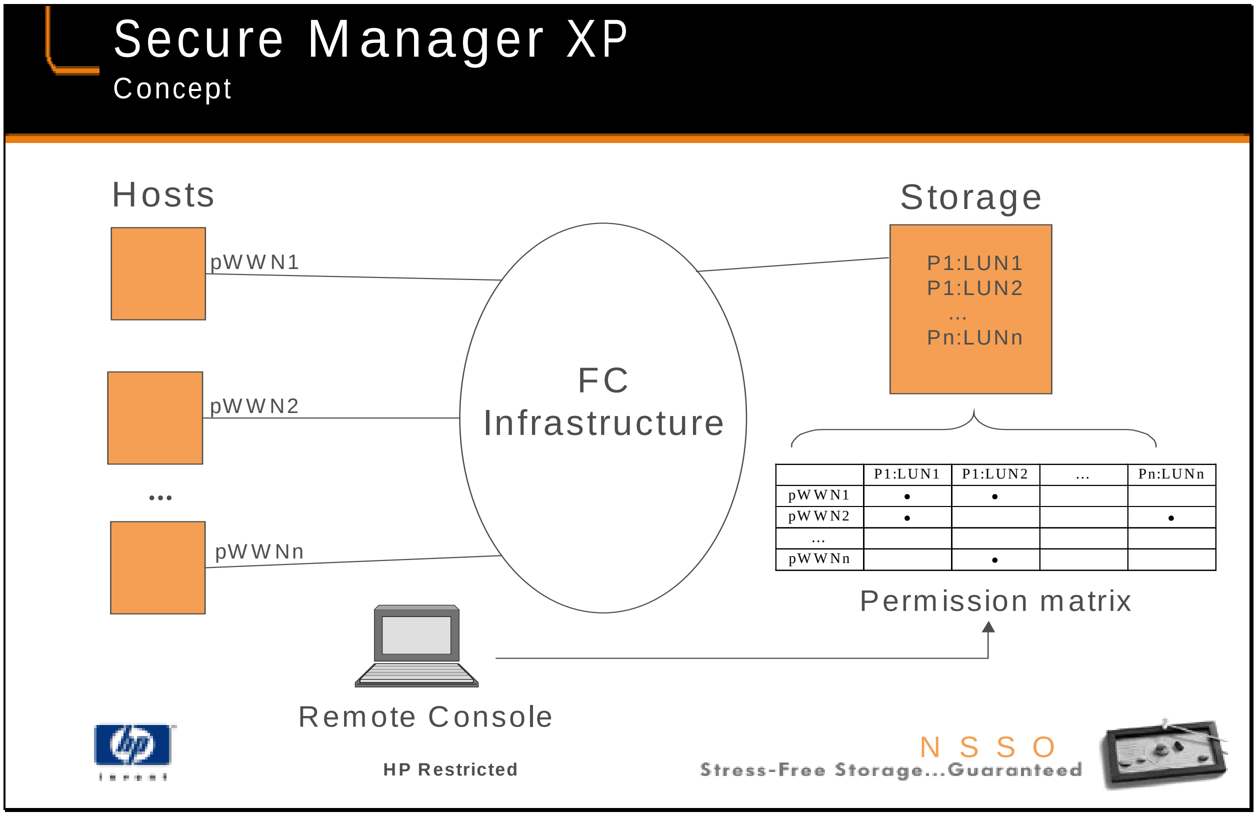
- Optional Feature: check all I/Os for security violations
  - this is not a default setting
  - without this option hides the LUN from the server during I/O scan or the boot process

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

There are two flavors of Secure Manager.

1. Standard version offering LUN security without checking by I/O
2. Optional version that checks security by I/O



This is a picture of – in general what Secure Manager can do. On the left side are your hosts and connected to the hosts are host bus adapters and these can be inserted into either a NT (Emulex) or H-P UNIX, or SUN Solaris (Jaycor) host system. And they go through some fiber channel or SAN infrastructure, and this is where the problems with security arise in that everything is now shared. Basically, every host system could see every LUN on the XP Family of Arrays because of this SAN infrastructure that ties the hosts and storage together (.e. Virtual Pool of Storage). The storage component, or array in this diagram is the XP 256. And in that storage box you see port numbers, for example, port 1, : LUN 1, port 1, LUN 2, port n, LUN n. And below that is a permission matrix, that the Secure Manager builds that grants access to selected LUNs for a selected set or singular worldwide name(s). And the arrow below on the bottom shows that Secure Manager XP is set up and managed from the Remote Console. Again, like what was said earlier, Secure Manager XP can be operated from the SVP as well (CE access only).

World Wide Name (WWN) and LUNs

- The WWN is a 64-bit physical address that uses the IEEE 48-bit format with a 12-bit extension and a 4-bit prefix. The WWN is hard-coded to each port. It is a unique name assigned only to that port. WWN is a unique name assigned only to that port.
- Obtain the WWN from each host port that will be granted access to LUNs (shown on following slide).

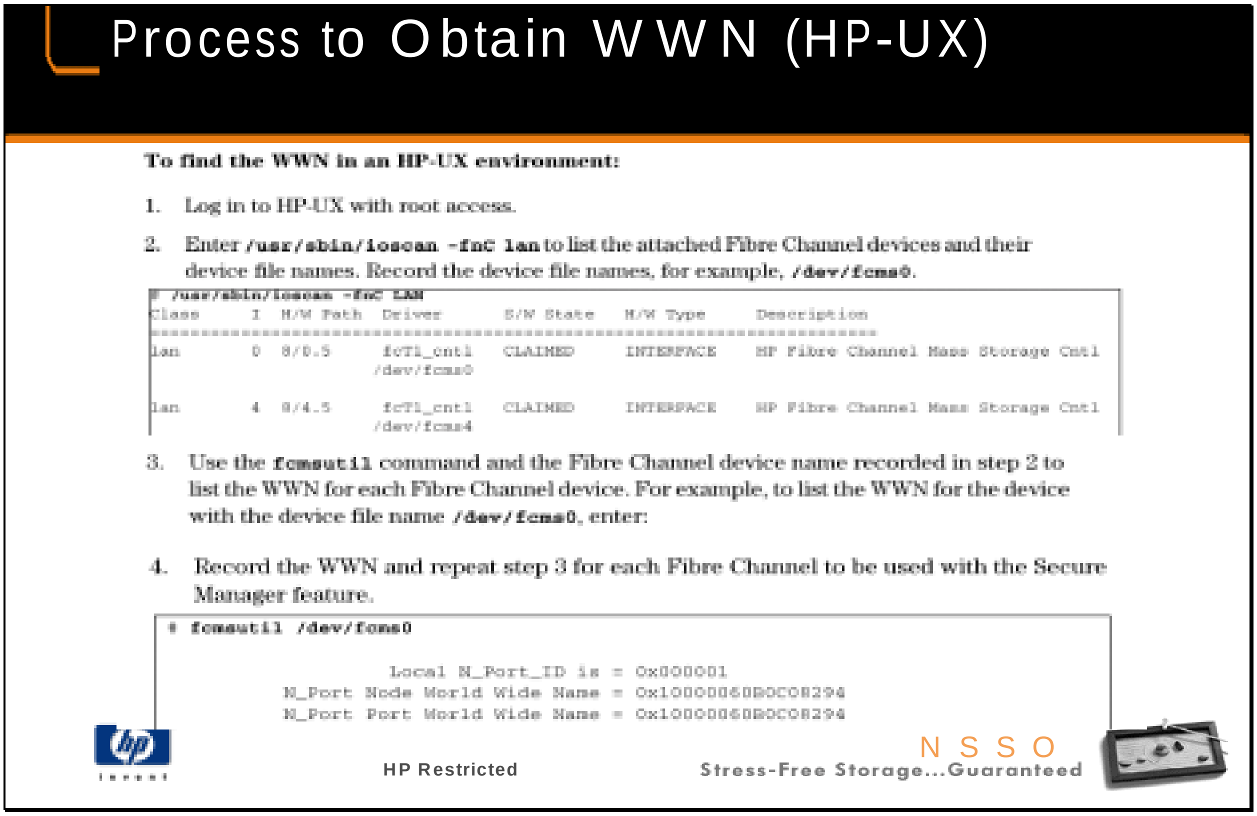
HP Restricted

NSSO
Stress-Free Storage...Guaranteed

HP-UX - copy the WWN printed on a label on the the HBA printed circuit board, or if access to the board is difficult, see below. The worldwide name is a 64-bit physical address, hard coded on HPUX fiber channel mass storage adapter cards, or host bus adapters for both HP-UX and NT adapter cards. Worldwide name is a unique name assigned and owned by that port. When you want to enable Secure Manager, you need to obtain the worldwide name for each host port that will be granted access to the different LUNs.

•WWN group
A WWN group allows you to assign multiple WWNs to a single group. Access to LUNs or LUN groups can be managed for all WWNs in the WWN group using a WWN group's name.

• worldwide name (WWN)
The WWN is a 64-bit physical address that uses the IEEE 48-bit format with a 12-bit extension and a 4-bit prefix. The WWN is hard-coded to each port. It is a unique name assigned only to that port.



To enable security to a given host, you must first determine the worldwide names (WWNs) of the host adapter ports. Finding the WWN is different depending on whether you have an HP-UX, Windows NT, or Sun Solaris system. Now for (UNIX), you can do this fairly easily by using a command called I/O scan as you can see outlined on the next slide. You can run I/O scan and identify the device file name. Then you run a Unix command called fcmsutil. And that command outputs the port node worldwide name, which is again, that 64-bit physical address. However, it shows up as 16 characters, so this is four bits per character, which is (ascii). So, you need to just copy down that 16 bit – number in hex decimal, which we'll use later to enter into Secure Manager. For a Windows NT environment the disk array currently supports the Emulex Fibre Channel adapter card in a Windows NT environment. Note whether you have installed the Emulex port driver or the Emulex mini-port driver. Follow one of the following two procedures, LightPulse for mini-port or the utility lp6dutil for the port driver.

Finding the WWN in an HP-UX environment requires that you have root access. To find the WWN in an HP-UX environment:
1. Log in to HP-UX with root access.
2. Enter /usr/sbin/ioscan -fnC lan to list the attached Fibre Channel devices and their device file names.

Example

/usr/sbin/ioscan -fnC lan

3. Record the device file names, for example, /dev/fcms0.
4. Use the fcmsutil command and the Fibre Channel device name recorded in step 3 to list the WWN for each Fibre Channel device.

Example To list the WWN for the device with the device file name /dev/fcms0, enter:

/opt/fcms/bin/fcmsutil /dev/fcms0

5. Record the WWN and repeat step 4 for each Fibre Channel to be used with the Secure Manager feature.

Operation Finding the WWN in a Solaris Environment

The disk array supports the Jaycor Fibre Channel adapter in a Sun Solaris environment.

To find the WWN in a Sun Solaris environment:

1. Log on to Solaris with root access.
2. Enter the command dmesg | grep Fibre to list the installed Fibre Channel devices and their WWNs.

Example # dmesg | grep Fibre

:

fcaw1: JNI Fibre Channel Adapter model FCW

fcaw1: Fibre Channel WWN: 200000e0694011a4

fcaw2: JNI Fibre Channel Adapter model FCW

fcaw2: Fibre Channel WWN: 200000e06940121e

#

3. Record the listed WWNs.

Finding the Worldwide Name in an HP-UX environment

- You can find the Worldwide Name in these environments:
  - HP-UX
  - Windows NT (using Emulex HBA with mini-port driver)
  - Windows NT (using Emulex HBA with port driver)
  - Sun Solaris
- The user's guide provides complete procedures for finding the Worldwide Name

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Student Notes

Note: A script is available to find the WWNs on HPUX, AIX and Solaris. Go to:

http://sseweb.nsr.hp.com/xp256/scripts/wwn_find.ksh

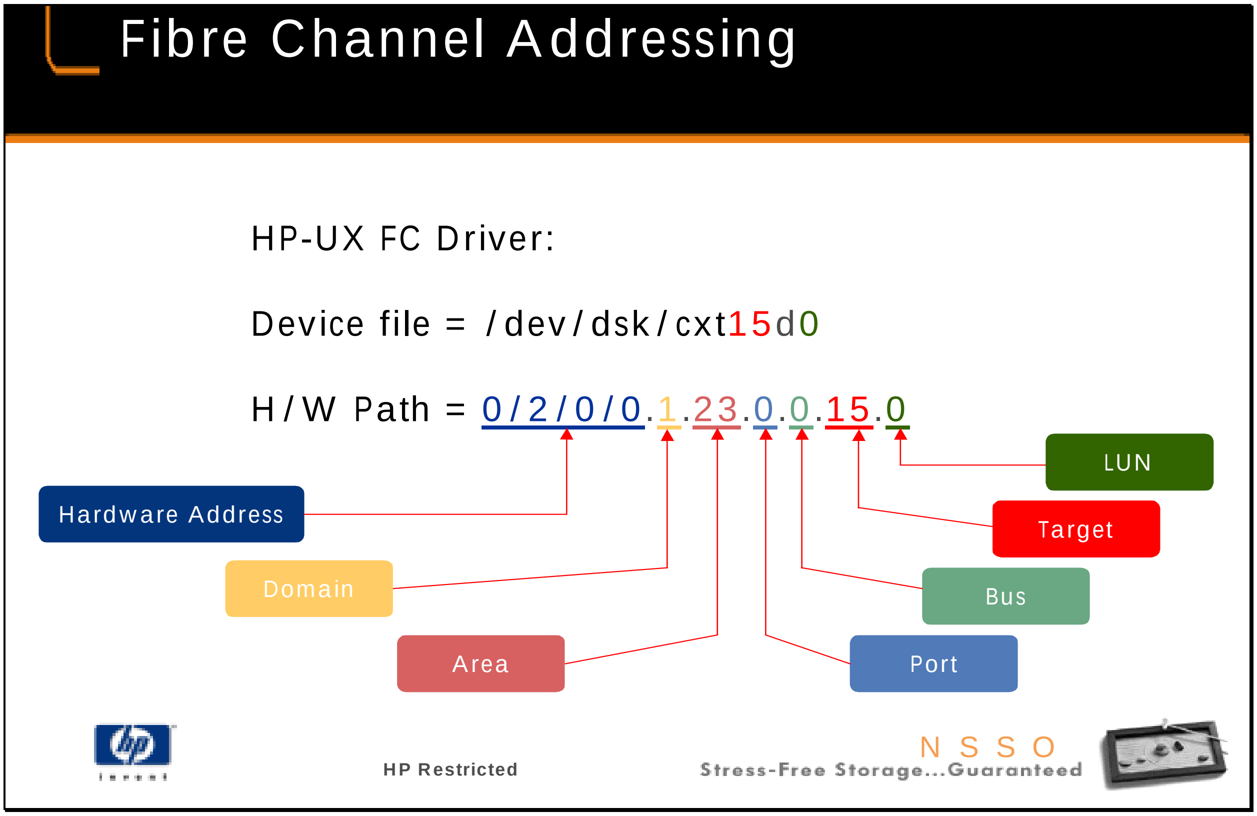
Security with Secure Manager

- Secure Manager XP is platform independent. LUNs are normally discovered by SCSI commands sent over Fibre Channel between hosts and the XP Family of Arrays.
- Secure Manager XP can prohibit access from a WWN to a Port:LUN by returning a “not present” status for the SCSI Inquiry command.
- SCSI status can prevent data access by preventing the host from building the proper device files.
- The SCSI Inquiry command is the underlying SCSI protocol command for HP-UX “ioscan” as well as for several of the special file commands (mksf, insf, etc.)

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

The WWN is printed on the edge connector/board stiffener. Sending SCSI commands over the Fibre Channel is independent of host operating systems. Secure Manager XP is platform independent. The LUNs are normally discovered by SCSI commands and the SCSI commands are sent over fiber channel between hosts and the XP 256. And these SCSI commands are an industry standard that, again, do not depend on a particular type of host you're using. Secure Manager XP can prohibit access from a worldwide name to a port : LUN by returning in what, in essence, is a “not present” status for the SCSI inquiry command. Because that's just what SCSI is doing. It’s sensing the presence of a device in that port and Secure Manager returns a “not present” to the host. SCSI status can prevent data access by preventing the host from building the proper device files. And that's the key to the way Secure Manager works because the device files on the host allow the host to talk with that particular device. A SCSI inquiry command is the underlying SCSI protocol command used by HPUX I/O scan utility, as well as for several of the special file commands like (mksf), and insf.



It is easy to deal with a XP Family of Arrays in a SCSI environment. However, with the advent of Storage Area Network (SAN) environments, trying to find where disks are located becomes more complex. The host that you are connected to may be able to see several XP Family of Arrays via a SAN. The only way to find where you are connected is to physically map the environment and then apply that mapping to the hardware path that is returned by the ioscan -fnC disk command. The slide shows how fibre channel addressing is built up. With this information and your SAN map, you would be able to locate where any disk is within the SAN. This slide shows how fiber channel addressing is accomplished. How in HP-UX fiber channel, if you're doing an I/O scan for example, you can see the device file name and the hardware path. And it shows different parts of the hardware address. What's important in this discussion are the rightmost components of the hardware path. And that indicates the SCSI target : LUN for that particular device. So, for example, the slide has marked in red the hardware path the last two numbers are 1.0. That would equate to be SCSI target 1 LUN 0, or TID 1:0.

HP-UX LUN's w/o Secure Manager XP

- 16 LUN's for this XP512 port were defined to demonstrate how LUN's appear at the UNIX host.
- Output of an ioscan from a UNIX host shows the SCSI Target ID's and LUN's for each target.

disk	231	0/4/0/0.1.17.0.0.14.0	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	296	0/4/0/0.1.17.0.0.14.1	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	297	0/4/0/0.1.17.0.0.14.2	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	232	0/4/0/0.1.17.0.0.15.0	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	233	0/4/0/0.1.17.0.0.15.1	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	234	0/4/0/0.1.17.0.0.15.2	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	235	0/4/0/0.1.17.0.0.15.3	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	236	0/4/0/0.1.17.0.0.15.4	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	237	0/4/0/0.1.17.0.0.15.5	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	238	0/4/0/0.1.17.0.0.15.6	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	239	0/4/0/0.1.17.0.0.15.7	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	240	0/4/0/0.1.17.0.1.0.0	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	241	0/4/0/0.1.17.0.1.0.1	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	242	0/4/0/0.1.17.0.1.0.2	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	243	0/4/0/0.1.17.0.1.0.3	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	244	0/4/0/0.1.17.0.1.0.4	sdisk	CLAIMED	DEVICE	HP	OPEN-3
disk	245	0/4/0/0.1.17.0.1.0.5	sdisk	CLAIMED	DEVICE	HP	OPEN-3

This slide is another example of HP-UX LUNs and as was mentioned before in a previous slide, there are 120 LUNs for a XP 256 port (SCSI or FC). We define all 120 LUNs to demonstrate what it looks like at the host by running an I/O scan. This on the left is an output of the I/O scan utility and it shows the SCSI target IDs and the LUNs for each target. And you'll see at the top, you'll see at the rightmost part of the number would be SCSI target zero and then you've got 0.0, 0.1, 0.2, 0.3, etc..., up to 0.7, so you'd have a maximum of eight LUNs for each target ID. Then the next line is SCSI target ID one, one zero, one one, etc…., for the next eight all the way up to SCSI ID 14 down at the bottom of the slide.

HP-UX LUN's with Secure Manager XP

- Secure Manager grants access to LUN's that show "CLAIMED" and denies access to LUN's that show "NO_HW" for this XP512 port.
- The no access LUN's show NO_HW at the host.

disk	231	0/4/0/0.1.17.0.0.14.0	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	296	0/4/0/0.1.17.0.0.14.1	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	297	0/4/0/0.1.17.0.0.14.2	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	232	0/4/0/0.1.17.0.0.15.0	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	233	0/4/0/0.1.17.0.0.15.1	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	234	0/4/0/0.1.17.0.0.15.2	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	235	0/4/0/0.1.17.0.0.15.3	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	236	0/4/0/0.1.17.0.0.15.4	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	237	0/4/0/0.1.17.0.0.15.5	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	238	0/4/0/0.1.17.0.0.15.6	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	239	0/4/0/0.1.17.0.0.15.7	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	240	0/4/0/0.1.17.0.1.0.0	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	241	0/4/0/0.1.17.0.1.0.1	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	242	0/4/0/0.1.17.0.1.0.2	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	243	0/4/0/0.1.17.0.1.0.3	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	244	0/4/0/0.1.17.0.1.0.4	sdisk	NO_HW	DEVICE	HP	OPEN-3
disk	245	0/4/0/0.1.17.0.1.0.5	sdisk	NO_HW	DEVICE	HP	OPEN-3

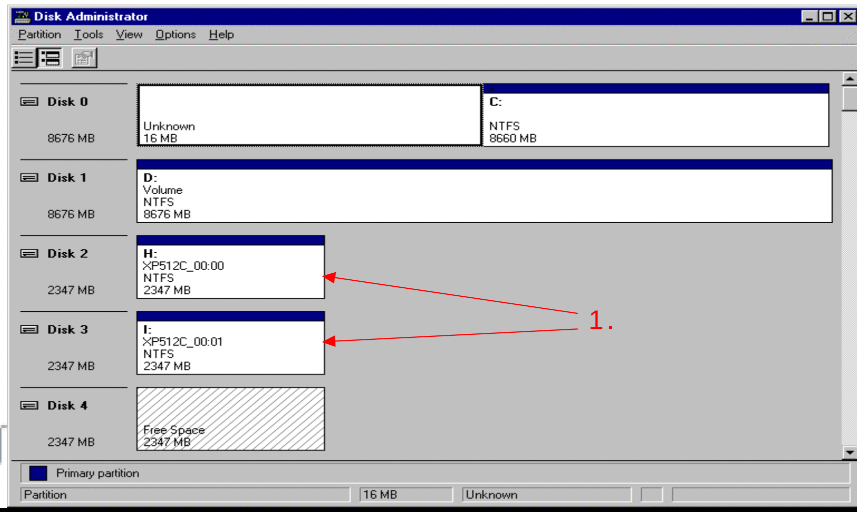
invent

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

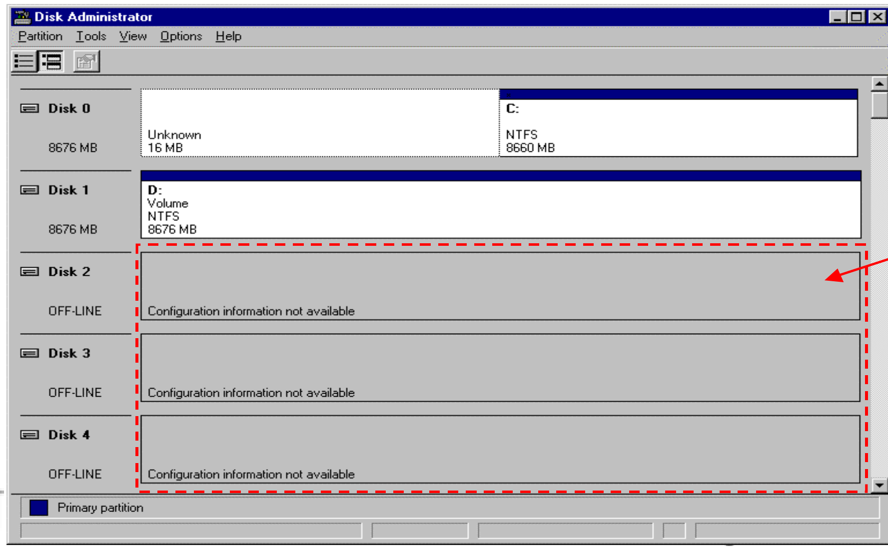
NT LUN's w/o Secure Manager XP

1. LUN's on a XP512 that have been formatted and appear as disks in Windows NT Disk Administrator.
2. Secure Manager XP has NOT been enabled, therefore this host has access to all LUN's configured to the XP connected port.



NT LUN's with Secure Manager XP

After the Secure Manager settings have been applied, the previously configured LUN's are no longer accessible.



Secure Manager XP

- XP256/XP512/XP48 Port can be online when changing LUN security.
- Optional feature XP256, default feature on XP512/48 - Mode 111- Security is checked for every I/O.
- The permission matrix will grant LUN's access only by WWN even if special device files exist on the host for a LUN.
- The LUN security of the enhanced Secure Manager XP cannot be bypassed by a HP-UX sysadmin or NT developer.
- Remote Control XP contains the software options that can be enabled by a key. The firmware version installed on the XP256/XP512/XP48 determines which Remote Control SW is installed on the Remote Console. Please check with your local HP CE or ASE to get the latest copy of the 'Microcode Release to Software Version Dependency Matrix'.

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

System Requirements

Required Equipment:

- HP LUN Configuration Manager XP installed on the system.
- HP Secure Manager XP option software key installed.
- HP XP256/512/48 Disk Array.

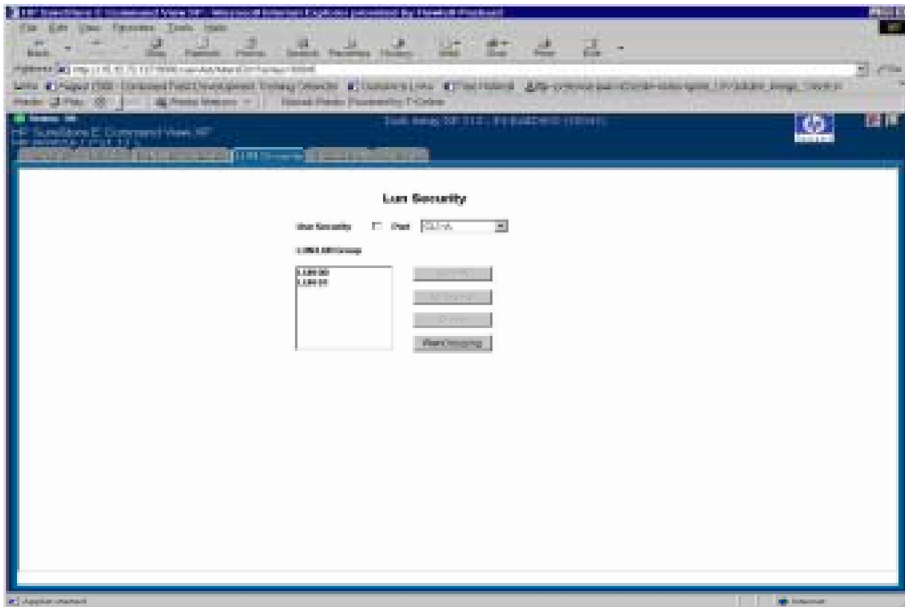
Optional Equipment:

- Remote console PC with Windows 95, Windows 98, or Windows NT.
- HP Remote Control XP (RC) installed on the system with LUN Configuration Manager acitvated.

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

LUN Security main page



- Click on the LUN Security tab

HP Restricted

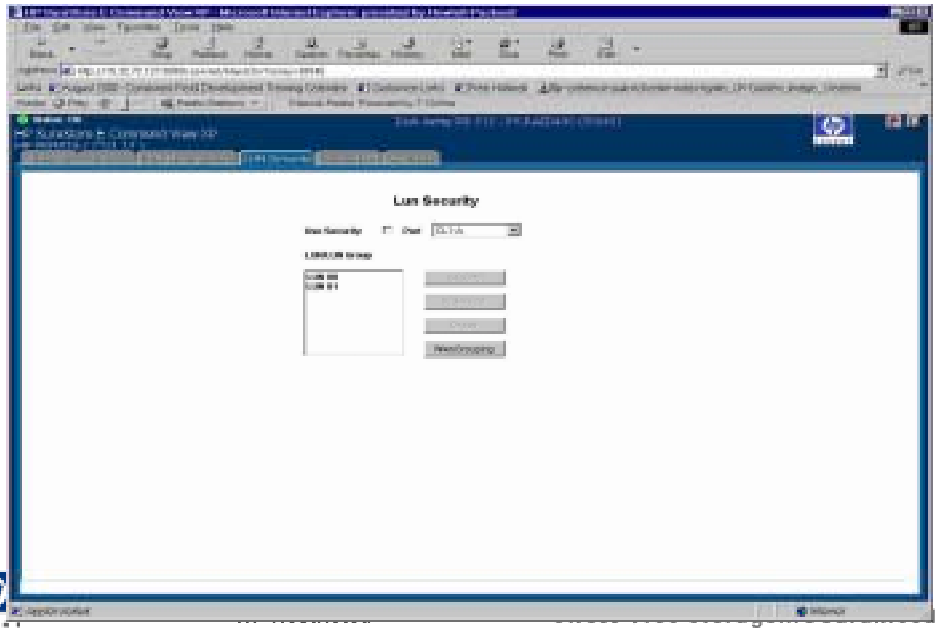
S S O
Stress-Free Storage...Guaranteed

Student Notes

To start LUN security operations:
1. Start Command View and select a disk array.
2. Select the LUN Security tab. The LUN Security page opens. All LUN security functions are performed from the LUN Security page.

The LUN Security page is divided into two parts: the top section appears when you select the LUN Security tab.
The bottom section appears when you select the WWN Grouping button.

Enabling and disabling LUN security



Student Notes

Disabling a port does not erase the LUN Security setting for that LUN. The assigned setting for the LUN is held until that LUN or its port is de-installed or until the settings are manually erased from within the LUN Security software.

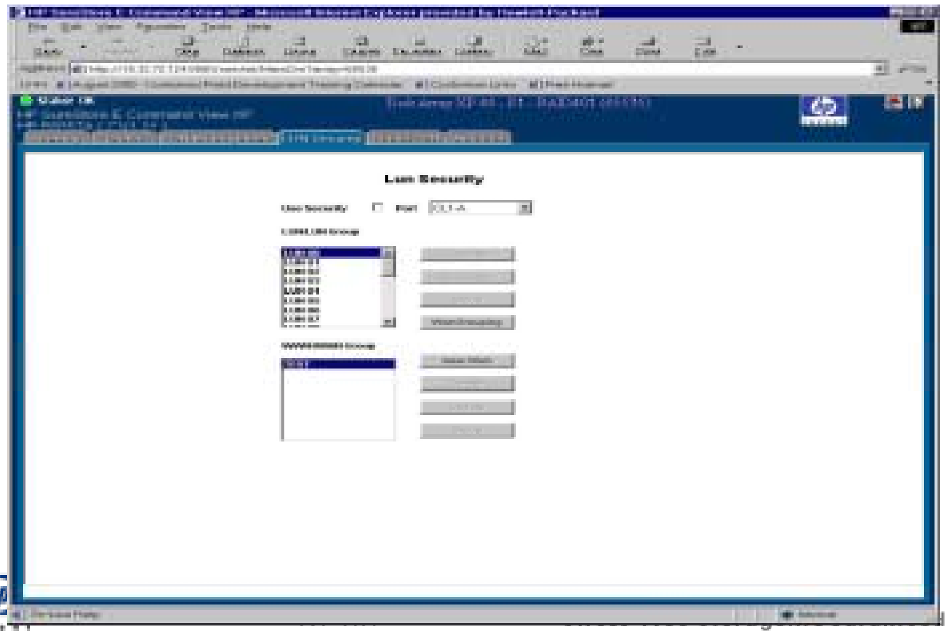
To enable LUN security:

1. Choose the LUN Security tab. The LUN Security Main page opens.
2. Select a port from the Port pulldown list box on the LUN Security Main page.
3. Select the Use Security check box to enable LUN security for the selected port. If no WWN is defined for the selected port, a dialog box asks you to define a WWN for the port before setting the security switch.

To disable LUN security:

1. Choose the LUN Security tab. The LUN Security Main page opens.
2. Select a port from the Port pulldown list box on the LUN Security Main page.
3. Deselect the Use Security check box to disable LUN security for the selected port. All security permissions for the port will be disabled and any open system can access the port.

Defining security parameters



Student Notes

When defining LUN security parameters, you can allow access to a single LUN or a group of LUNs. You must create the LUN group. Access to a LUN or a LUN group can be made available to a single host (WWN), or you can group the WWNs to allow access by multiple hosts.

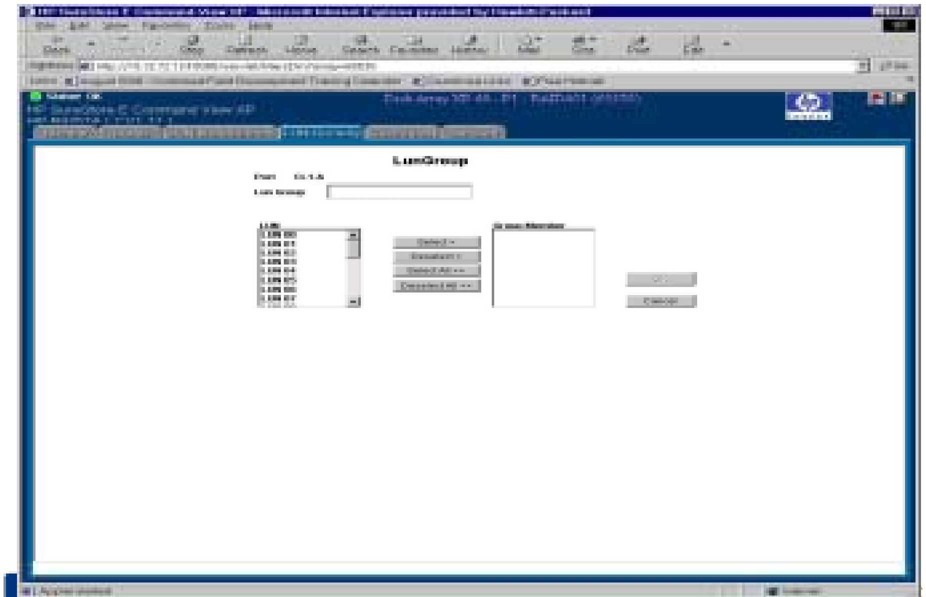
To define LUN security parameters:

1. Choose the LUN Security tab. The LUN Security Main page opens.
2. Select the port you want to secure from the Port pulldown list.
3. Select the Use Security box. The LUNs for the specified port appear in the LUN/LUN Group list box. If no LUNs are displayed, a LUN must be defined.
4. (Optional) If you want to create a LUN group, choose the Group button, and the LUN Group page opens. If you do not want to create a LUN group, continue with step 5.
5. Choose the WWN Grouping button to view the WWN/WWN group list. The WWN/WWN Group window opens. If no WWNs and WWN groups are defined for the selected port, only the New WWN button is activated. Define a new WWN and/or WWN group.
6. After you have defined the LUNs, LUN groups, WWNs, and WWN groups, choose the Security button. The LUN Security page opens.
7. Select the LUN or LUN group you want to secure from the LUN/LUN Group pulldown box.

8. Select the WWN or WWN group you want to grant access to for the selected LUN or LUN group in the WWN/WWN Group list box.
9. Choose the Select button. To select all of the WWNs in the list box, choose the Select All button. The selected WWNs are displayed in the Permitted Member list box.
10. (Optional) If you do not want to grant access for a selected WWN, select the unwanted WWN in the Permitted Member list box and choose the Deselect button. To move all of the selected WWNs out of the Permitted Member list, choose the Deselect All button. The selected WWNs are removed from the Permitted Member list box and displayed in the WWN/WWN Group list box.
11. Choose OK to confirm your entries or select Cancel to cancel your selections. In either case, you are returned to the LUN Security page.

Repeat these steps for each port you want to secure.

Creating a LUN group



HP Restricted

Stress-Free Storage...Guaranteed

Student Notes

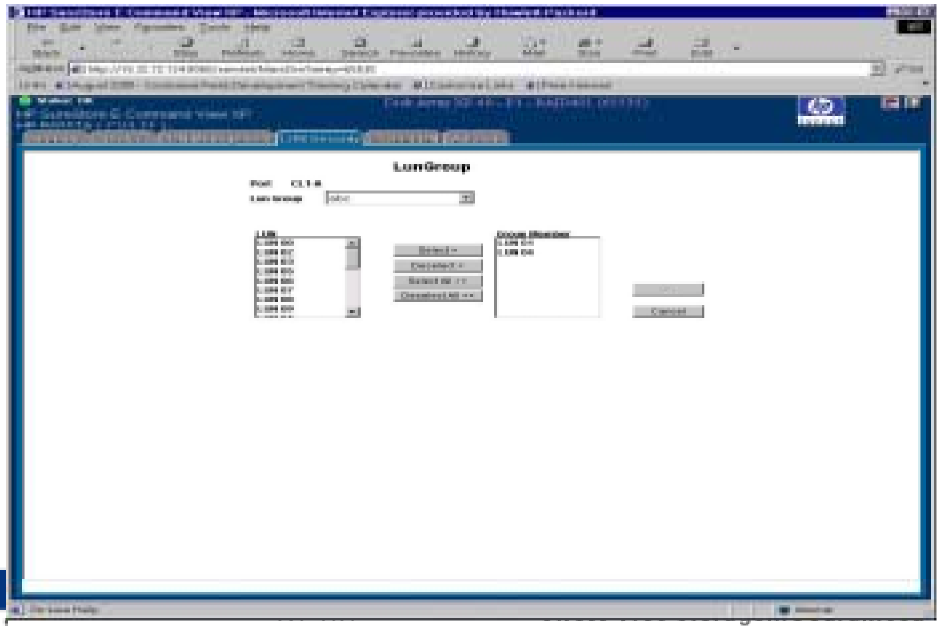
To create a LUN group:

1. Choose the LUN Security tab. The LUN Security Main page opens.
2. Select the desired port from the Port pull-down list box on the LUN Security Main page.
3. Select the Group button. The LUN Group page opens.
4. Define or select a group name in the LUN Group pull-down list box.

In the LUN list box, select the LUNs you want to assign to the specified group, and choose the Select button. To select all of the LUNs in the list box, choose the Select All button. The selected LUNs are displayed in the Group Member list box. If you do not want to assign a selected LUN, select the unwanted LUN in the Group Member list box and choose the Deselect button. To move all of the LUNs out of the defined group, choose the Deselect All button. The selected LUNs are removed from the Group Member list box and displayed in the LUN list box.

5. After assigning the desired LUNs to the group, choose OK to confirm your selections. The LUN Security page opens again. Choose Cancel to cancel your selections and return to the Security Main page.

Changing LUN or LUN group parameters



Student Notes

To change LUN or LUN group parameters:

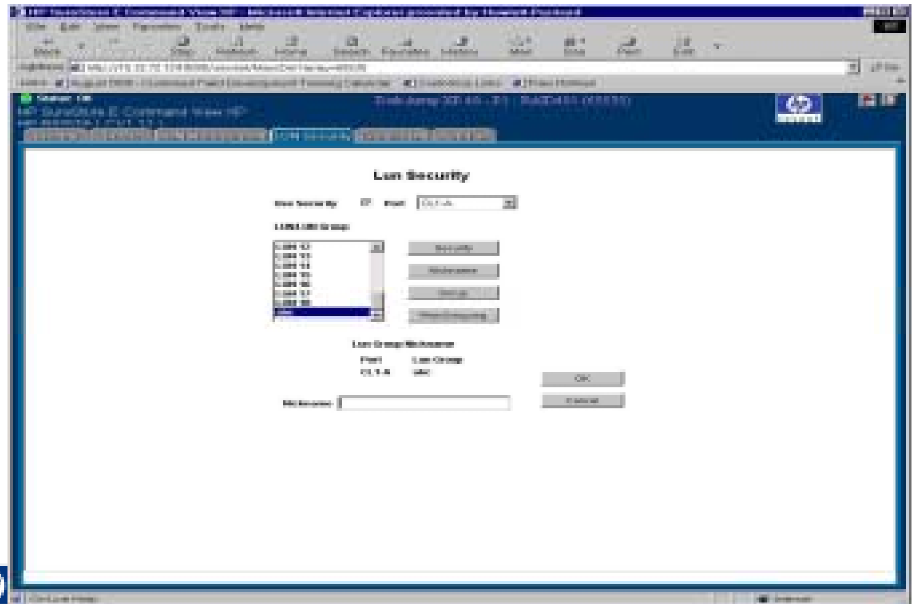
1. Select the LUN Security tab. The LUN Security Main page opens.
2. Select the port from the Port pulldown list on the LUN Security page.
3. Select the LUN group from the LUN/LUN group list box on the LUN Security Main page.
4. Choose the Group button. The LUN Group page opens.
5. Select the group name from the LUN Group pulldown list box.
6. You can add LUNs to or remove LUNs from a group.

To add LUNs, select the LUNs you want to assign and choose the Select button or to select all of the displayed LUNs choose the Select All button.

To remove LUNs, select the LUNs you want to remove and choose the Select button or to select all of the displayed LUNs choose the Deselect All button.

7. Choose the OK button to confirm your selections, or choose the Cancel button to cancel your selections.

Changing a LUN group nickname



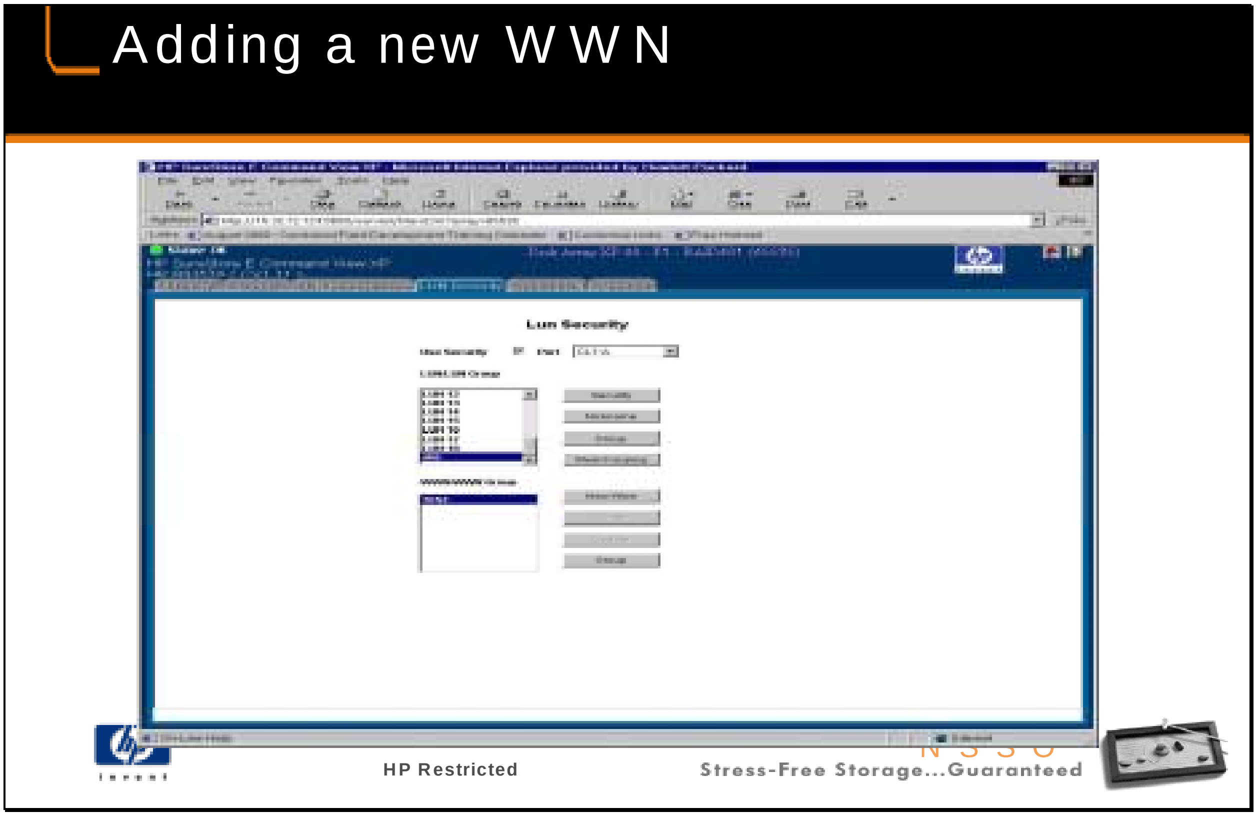
HP Restricted   Stress-Free Storage...Guaranteed

Student Notes

You may assign nicknames LUN groups. A nicknames is a familiar or easy-to-remember name.

To change a LUN group nickname:

1. Select the LUN Security tab. The LUN Security Main page opens.
2. Select the port from the Port pulldown list on the LUN Security page.
3. Select the LUN group you want to change in the LUN/LUN Group list box.
4. Choose the Nickname button. The LUN Group Nickname sub-screen opens.
5. Enter the new nickname for the LUN group, and choose the OK button to confirm your entry, or choose the Cancel button to cancel the nickname change.

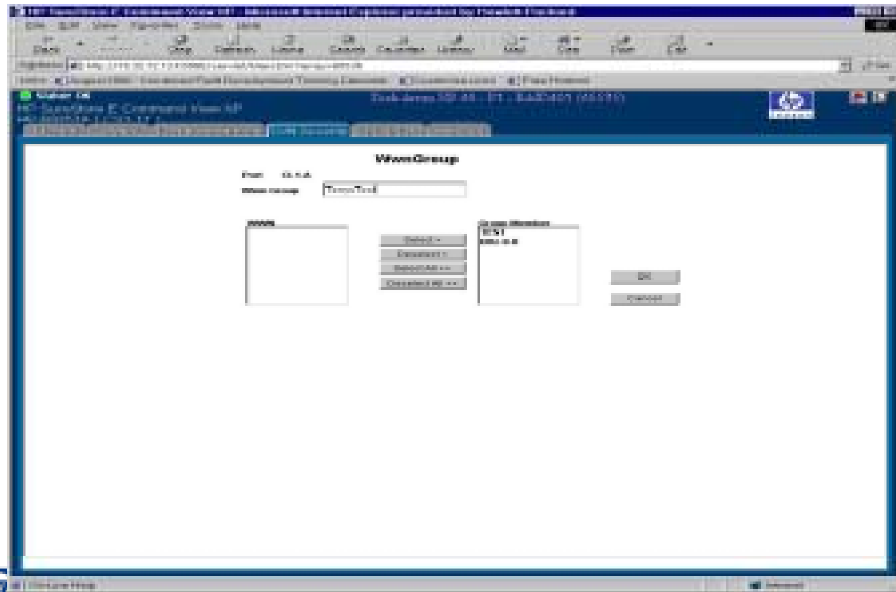


Student Notes

To add a new WWN:

1. Select the LUN Security tab. The LUN Security Main page opens.
2. Select the port from the Port pulldown list on the LUN Security page.
3. Choose the WWN Grouping button on the LUN Security Main page. The WWN/WWN Group subscreen opens.
4. Choose the New WWN button. The New WWN page opens.
5. Enter the New WWN in the WWN text box. The WWN is a hexadecimal string containing 16 characters.
6. Enter the desired Nickname in the Nickname text box. (Maximum 8 characters)
7. Choose the Add button to add the WWN to the WWN list. The WWN displays in the list box.
8. Select the WWN that you want to register from the WWN list. Multiple selections are possible.
9. Choose the Set or OK button to register the WWN, or choose Cancel to cancel your entries. If the Set button is selected, the display remains in the same screen after registering the WWN for the selected port. If the OK button is selected, the display returns the LUN Security Main page after registering the WWN.

Creating a W W N group



HP Restricted

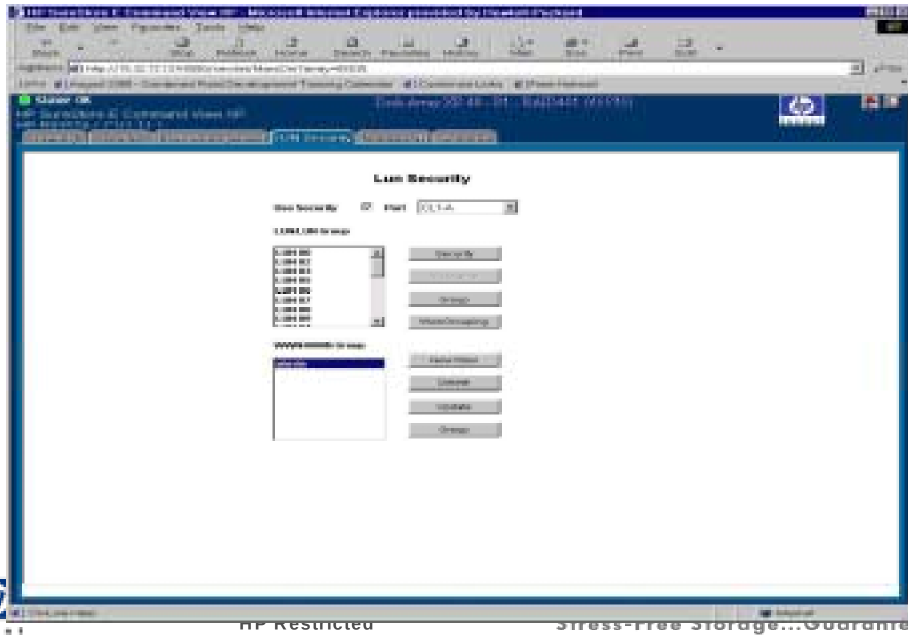
Stress-Free Storage...Guaranteed

Student Notes

To create a WWN group:

1. Select the LUN Security tab. The LUN Security main page opens.
2. Select the port from the Port pulldown list on the LUN Security page.
3. Choose the WWN Grouping button. The WWN/WWN Group subscreen opens.
4. Choose the Group button. The WWN Group page opens.
5. Enter the desired name for the WWN Group in the WWN Group text box.
6. To add a WWN to the selected WWN group, select the WWN from the WWN list box and choose the Select button. The selected WWN will display in the Group Member list box and removed from the WWN list box.
7. To undo the selection, choose the Deselect button. The WWN is displayed in the WWN list box and removed from the Group Member list box.
8. To select all the WWNs in the list box, choose the Select All button.
9. To remove all the selected WWNs from the defined group, choose the Deselect All button.

Deleting a WWN or WWN group

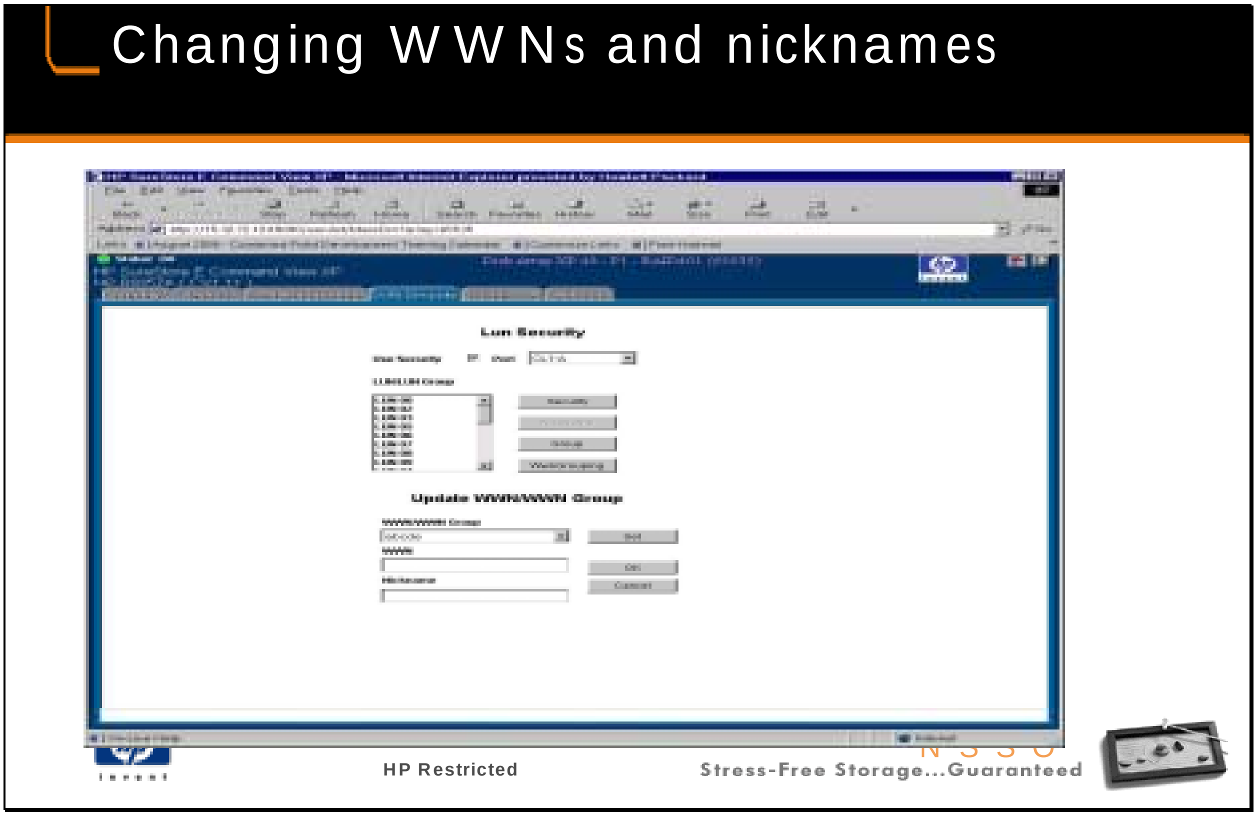


HP Restricted Stress-Free Storage...Guaranteed

Student Notes

To delete a WWN or WWN group:

1. Select the LUN Security tab. The LUN Security Main page opens.
2. Select the port from the Port pulldown list on the LUN Security page.
3. If the WWN/WWN Group list box is not displayed, choose the WWN Grouping button.
4. Choose the WWN Group button. The WWN/WWN Group sub-screen opens.
5. In the WWN/WWN Group list box, double click on the unwanted WWN or WWN group and choose the Delete button. A delete confirmation message is displayed.
6. Choose OK to delete the name, or Cancel to exit without deleting the name.

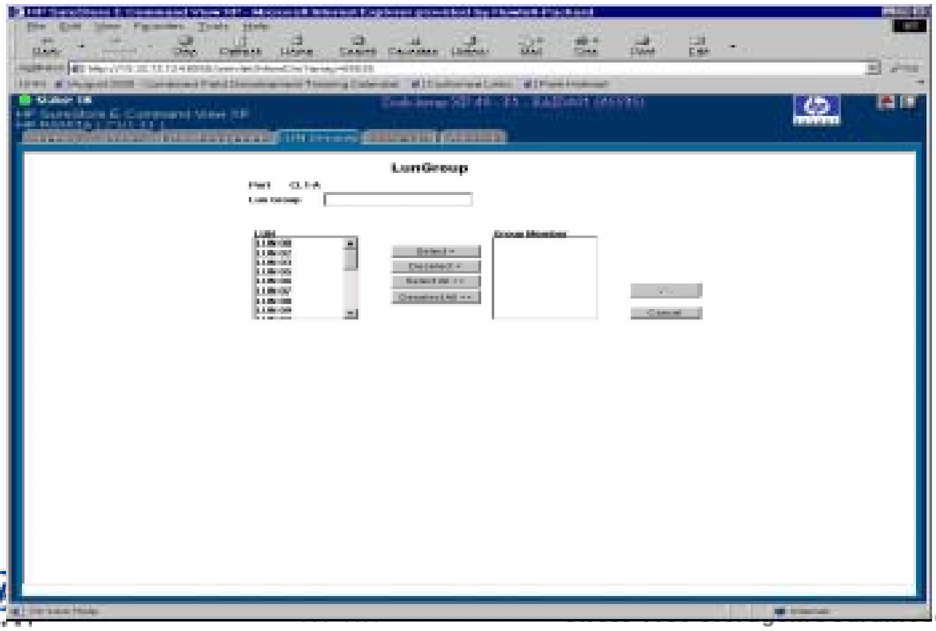


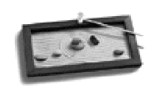
Student Notes

To change WWNs and nicknames:

1. Select the LUN Security tab. The LUN Security Main page opens.
2. Select the port from the Port pull-down list on the LUN Security page.
3. If the WWN/WWN Group list box is not displayed, choose the WWN/Grouping button.
4. Choose the Update button next to the WWN/WWN Group list box. The Update WWN Group page opens. To change the WWN, select the nickname for the WWN you want to change from the WWN/WWN Group pull-down list.
5. Enter the new WWN in the WWN box, and choose the Set button.
6. A warning box asks whether you want to change the WWN or the nickname. Choose Yes.
7. To change the WWN or WWN Group nickname, select the nickname for the WWN or WWN Group from the WWN/WWN Group pull-down list.
8. When you have finished making the changes to the WWN parameters, choose OK.

Changing the W W N group definition





Student Notes

To change the WWN group definition:

1. Select the LUN Security tab. The LUN Security Main page opens.
2. Select the port from the Port pulldown list on the LUN Security page.
3. Select the WWN Group from the WWN/WWN Group list box.
4. Choose the Group button. The WWN Group page opens.
5. Select the WWN group that you want to change in the WWN Group pulldown box.
6. To add a WWN to the selected group, select the WWNs you want to assign to the specified group in the WWN list box, and choose the Select button. Choose the Select All button to select all of the WWNs in the list box.
7. To delete a WWN from the selected group, select the WWNs you want to delete from the specified group, and choose the Deselect button. Choose the Deselect All button to delete all of the WWNs in the list box.
8. When you are satisfied with your selections, choose the OK button. To cancel your selections, choose the Cancel button.

Defining LUN Security Parameters

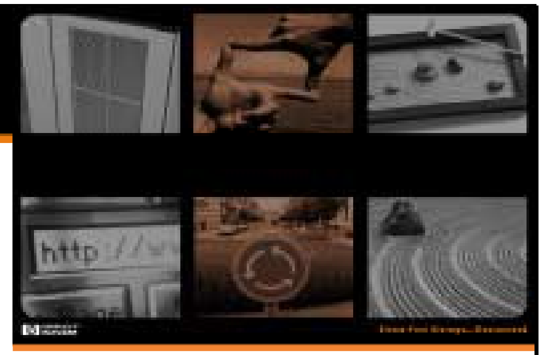
- Secure Manager XP Capacity.
- Secure Manager LUN security parameters (screen shots follow.)
- Assign LUN group names to specified LUN's.
- Assign WWN nicknames to WWN's.
- Assign WWN group names to WWN nicknames.
- Selectively grant access between hosts and LUN's.

With Secure Manager, you can assign specified LUNs to a LUN group name. You can also assign worldwide names to worldwide name nicknames. That way you don't need to use that 16 character name, you can give it a nickname which makes it easier. You can assign worldwide names to worldwide name groups. You can assign more than one worldwide name to a port and one worldwide name to several ports. Every host in the specified worldwide name group has access to a specified LUN or a specified group of LUNs and again, it does this in the permission table, and I'll get to an example of that in a few minutes.

Module Wrap-up

HP Restricted

NSSO
Stress-Free Storage...Guaranteed



Module 7

Business Copy XP - B9330A Licenses: B9331A, B9332A, B9333A & B9334A

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

**Summary:

Beginning August 1st, 2001, Continuous Access XP, Continuous Access XP Extension, and Business Copy XP Software for the XP Disk Array will be licensed based upon ‘used’ capacity.

**Background:

For 7 XP Software titles (Continuous Access XP, Continuous Access XP Extension, Business Copy XP, LUN Configuration Manager XP, Secure Manager XP, Cache LUN XP, and Auto LUN XP), current license terms dictate that in order for a customer to use these products on an XP Disk Array they must purchase the appropriate number of license-to-use (LTU) products to provide total license capacity to cover the total ‘raw’ capacity of the disk array in question (Raw capacity is defined as the total nominal capacity of all Disk Array Groups installed on the XP Array). Customers have expressed that licensing based upon the raw capacity of the Array for Business Copy XP, Continuous Access XP, and Continuous Access XP Extension titles is not equitable and does not align with their perceived value of the products owing to the fact that they are most often deployed on a subset of the Disk Array.

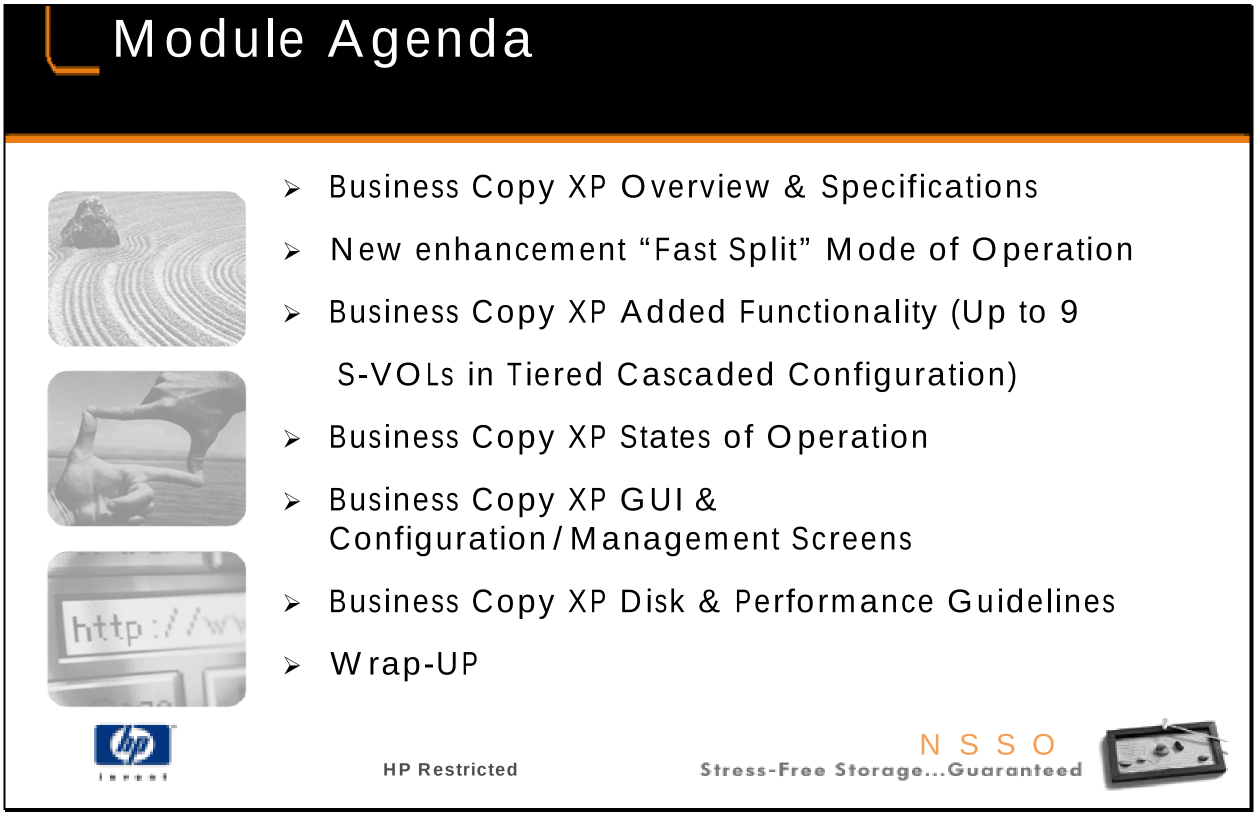
Effective August 1st, 2001, license terms for Continuous Access XP, Continuous Access XP Extension, and Business Copy XP will be changed to be based on the ‘used’ capacity of the Disk Array. This means that customers will buy license capacity to cover only the portion of the array capacity that is being ‘utilized’ by the Software.

For these Software titles, used capacity is defined as the total capacity of all volumes that are being managed as a Primary/Source Volume (P-Vol) or a Secondary/Copy Volume (S-Vol) on the Disk Array. This amount represents the total true usable capacity of these volumes corresponding with their configured LDEV sizes (i.e. the amount excludes RAID overhead). For Business Copy XP, this represents all related P-Vol and S-Vol on the Disk Array (every source volume and any associated copy volumes). For Continuous Access XP and/or Continuous Access XP Extension configuration, this represents any P-Vol or S-Vol pairs on a single Disk Array that is being managed in a P-Vol/S-Vol pair with another Disk Array.

Note the following:
- There will be no change to product structure (same products and product numbers as prior to August 1st)
- There will be no change to product pricing for the 1/5/10/25TB LTU products (no price changes in moving to used capacity license terms)
- Licensing for LUN Configuration Manager XP, Secure Manager XP, Auto LUN XP, and Cache LUN XP will continue to be based upon the raw capacity of the Disk Array.

This change results in SW License cost savings for most customers who require these products, especially for those who utilize a relatively small portion of the array for the solution.

**
A CSU Newsflash will soon be posted on ESP which details this change (ESP, keyword CABCUSED,)



The HP SureStore E Business Copy XP software (BC) maintains internal copies of open-system logical units (LUs). HP has introduced a 9-tiered Business Copy (BC) mirroring capability, either via the Remote Control (RC) GUI or a host-based RAIDManager (RM) command line interface (CLI). SureStore E Business Copy XP (B9333A), this software allows you to create high-performance, local business copies of any active application volume while benefiting from full RAID protection. These copies can be used by another application or system for a variety of purposes, including Year application testing, batch and system backup. This software allows you to maintain up to 9 copies of critical data plus the master. Business Copy XP enhances user productivity and reduces data duplication requirements through a non-disruptive resynchronization capability. When used in conjunction with Continuous Access XP, Business Copy XP enables you to maintain multiple copies of critical data at local and remote sites. Business Copy XP also allows you to perform updates from the source copy to the business copy and from the business copy back to the source copy.

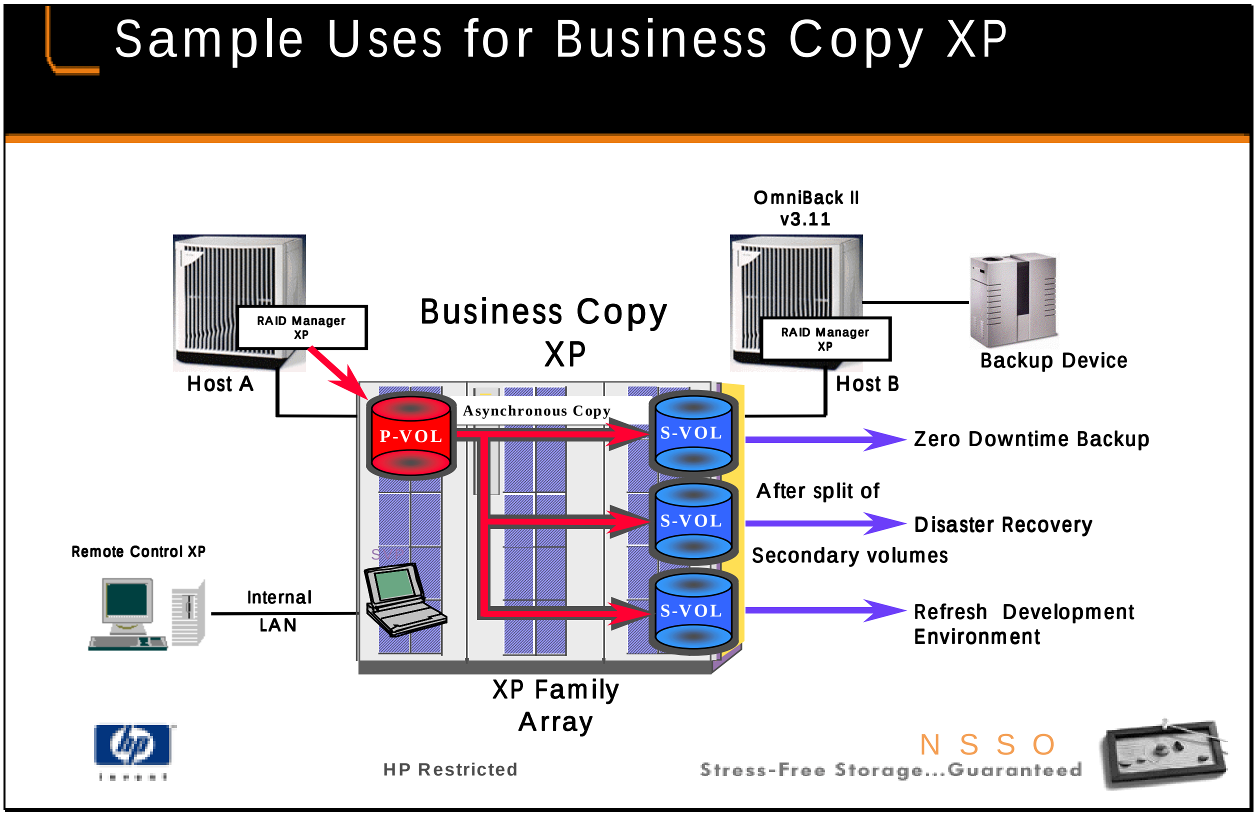
Business Copy XP

- Business Copy, B9333A, is a XP Family Array feature that allows local copies of logical units within the same disk array
- You can make up to 9 Secondary Volumes (S-VOLs) for each Primary Volume (P-VOL)
  - P-VOLs contain original data, S-VOLs contain duplicate data
- BC has Asynchronous Copy to Secondary Volumes
  - Resulting in a decrease in acknowledgement time and increase in throughput from the host
- BC operations are performed on the RMC, SVP or with command line control of RAID Manager XP
- Online split-mirror backup with OmniBack II integration
- EMC Timefinder makes only one BCV, Business Continuance Volume

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

We are going to go into detail about the actual functionality of BC, its capabilities, and detail of how you would use it, and more specifically, how would you solve customer issues and business needs with this product. First of all, let's mention that Business Copy is an in-frame copy, that means that these copies are inside the XP Family of Arrays so it uses the local buses and CPUs to do the actual copy operation and the host is, in most part, is completely unaware of any of this background copy operation. Our solution on the XP Family of Arrays supports up to nine secondary volumes. This is a selling point. We have a number of customers out there who need more than just a couple of copies of their data. This product is also integrated with OmniBack II for doing zero downtime backups. People use this for batch processing, application testing and so on. There is a number of different things that people will use the split copies for, and we have customers that have up to seven copies of their data for various kinds of report generation and so on. I also need to point out that this operation is asynchronous, which means that your primary copy and your secondary copy are not exactly synchronized until the split command is complete. So it's important to note that this is an asynchronous operation, which means every time the host does a write, it doesn't guarantee you that it gets into the secondary copy or copies immediately. This operation is controlled by either RAID Manager XP, which runs on your host, or by the remote control, which runs on a PC connected to the private network to the XP Family Array. It allows you to create, split and monitor the process. Some of the things that I've noted here on the slide that you would use a secondary copy for would be, for instance, data backup, which is reason number one. And you would do that after you've split the secondary volume. Batch processing is another good use of a split volume because you can momentarily suspend your OLTP or database operation, complete the split, resume your database or OLTP processing and then start up batch processing on a completely different set of data; therefore not contending with your primary data processing. Also if you want to do application testing, which could be destructive to the data, a copy very well suited for that because you can do destructive testing – where actually modifying the data and you can refresh it off the primary copy at any point.



General BC Concepts/Terminology
BC is an internal volume mirroring product that:
· Supports a Primary Volume plus up to 9 Secondary Volumes
· Is integrated with Omniback for fully automated split-mirror backup
· Provides (suspended) mirrors for Backup, Batch Processing and Application Testing·
. Provides high performance via asynchronous (opportunistic) copying to Secondary Volumes.
· Can be controlled via RAID Manager XP CLI, the RM library API or the Remote Control PC
· Local volumes are replicated internally using one or more DKA pairs.
· Up to 2047 BC pairs can be created (due to the need for a Control Device).
· BC Volumes can be CVS, DCR (or both), regular or LUSE (aggregated) LUNS.
· Two LUNs in a pair must be the same emulation type (e.g. OPEN-3) but RAID types can be mixed
· No additional Interface card and cable needed (as with loop-back CA)
· If LUSE or CVS, the two LDEVs must be the same capacity and same emulation type.
Vgchgid
The vgchgid command changes the LVM volume group identifier of a split BC mirror so it can be used by LVM without a conflict with the LVM volume group

residing on the pvol. At first release, this command refused to work with XP256 svols unless the "-f" force option was used. Work is now completed in the LVM Cupertino Lab to add vgchgid support for the XP256 (via patch).

Business Copy Specifications

- Each S-VOL is paired with its P-VOL independently.
  - Each S-VOL can be maintained as a copy that can be spit, suspended, resynchronized, and deleted separately from other S-VOLs from same P-VOL
  - Up to 2047 BC pairs can be created (due to the need for a Control Device)
  - BC Pairs composed of two LDEVS each with the Same Emulation Type.
  - In case of LUSE & VSC, two LDEVs must have same or larger capacity and same emulation type.
  - Supports Cache LUN XP (BIND/PRIORITY mode) for BC Paired Volumes.
  - Volumes of a BC Pair DO NOT have to match RAID Levels, mixing allowed
  - Command Device Volumes cannot be used as paired volumes

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

The source or 'primary' volumes are standard RAID 1 or RAID 5 volumes. Intermix is supported. That is, the Business Copy images can be different from the primary. For example, a RAID 5 Business Copy can be made from a primary that is RAID 1. Business copy provides non-disruptive subsystem internal replication of an open-3/9/8 or custom sized (VSC and LUSE) LDEVs (primary volume). A primary volume may have up to nine copies (secondary volumes). If the source volume is a customized, then the target must also be a customized volume of equal or greater capacity. Source and target volumes can be any combination of RAID levels (1 and 5) and must have their SCSI paths defined. The primary and each of its secondary copies are independent pairs which can be synchronized, split, suspended, resynchronized, and deleted separately.

Business Copy Specifications
(cont.)

- Shared Memory sizing needs to be considered when going past 64 BC volume pairs.
- Additional shared memory needed for more than 64 BC volume pairs (option for more BC pairs is selected on the SVP after more memory is installed)
- Track Change Tables are maintained in shared memory for Primaries and all copies
- Up to 9 secondary volumes for a single P-VOL

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Business Copy supports a maximum of 512 BC pairs (512 P-VOLs and 512 S-VOLs, for a total of 1,024 volumes). When BC volume pairs include size-extended LUs, the maximum number of pairs decreases. When BC volumes include more than one S-VOL, the maximum number of P-VOLs decreases. If you plan to mare than 64 BC volume pairs, you must have additional shared memory installed and available for use by BC. P-VOLs remain available to all hosts at all times for read and write I/O operations. All write operations to the S-VOL are prohibited. S-VOLs become available for host access only after you split the BC volume pair. While a BC volume pair is split, the disk array keeps track of all changes to the P-VOL and S-VOL. When you resynchronize the pair, the differential data in the P-VOL (due to the P-VOL and S-VOL updates) is copied to the S_VOL so that the S-VOL is again identical to the P-VOL.

Business Copy Specifications
(cont.)

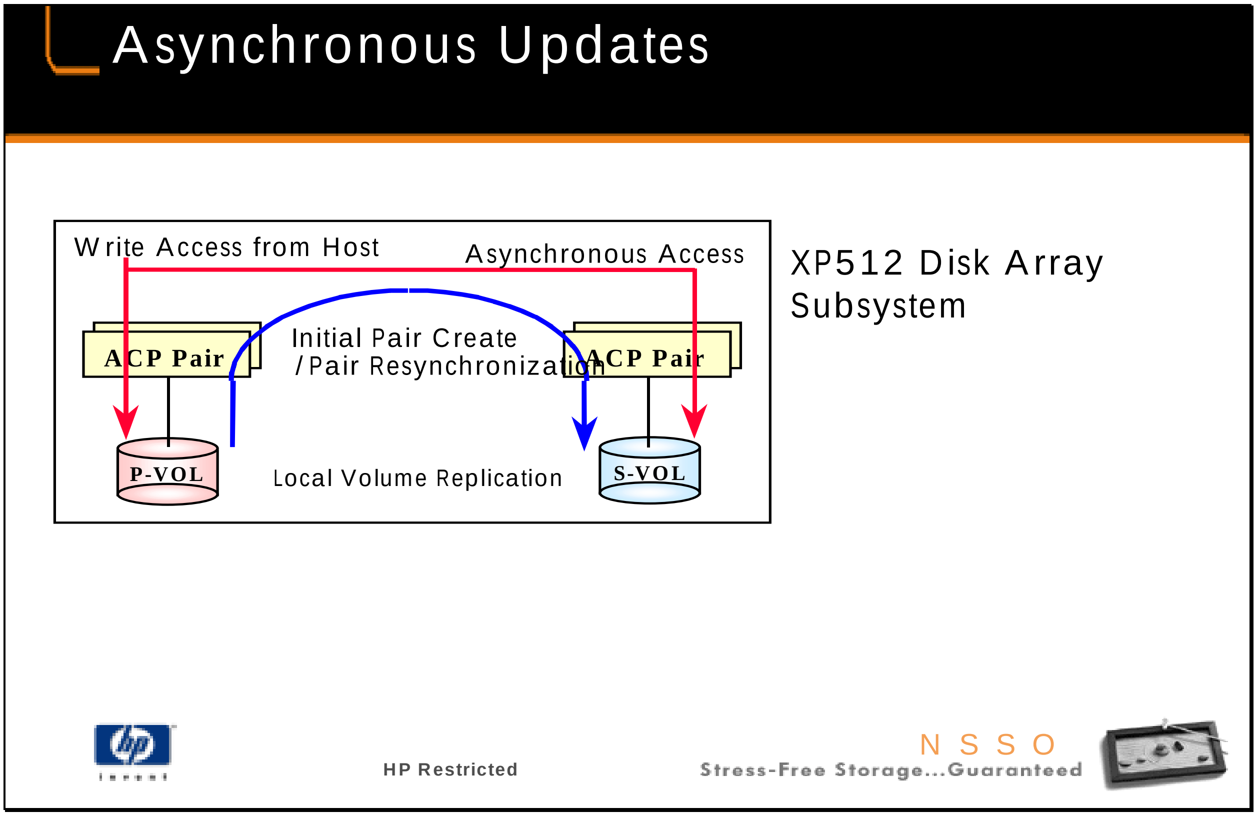
- LDEV maintenance cannot be performed on LDEVs assigned to BC pairs. The pair must first be deleted.
- When a failure of any kind prevents a copy operation from completing, BC will suspend the pair. For a LDEV failure, BC suspends the pair. If a Physical Device failure occurs, BC pair status is not affected because of the RAID architecture.
- Continuous Access XP can be used with BC to enable creation of remote copies of the data stored on the local disk array, for data backup and disaster recovery.
- High performance is achieved via asynchronous writes
  - "Write complete" is sent to the host when data is written to cache
  - Cache data is asynchronously written to disk at a ***convenient*** time

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

High performance is achieved via asynchronous (opportunistic) data movement to the mirror(s).Because the Shadow image mechanism is independent of the Continuous Access XP software, you can have an other copy available on the remote site. And of course this copy may be put in RW protection to do some specific actions. EMC Time finder/ BCV Equivalent Local Mirroring (within XP Family Array subsystem) of LUN's, High performance mirroring of local LUN's, Full RAID protection for LUN copies at all times!, LUN mirror available for local or remote Host via optional Continuous Access XP product),

Resynchronization with production LUN support, Supports primary volume plus 9 simultaneous mirrors, (for a total of 10 copies of data), Co-existence with Continuous Access XP - Management interface provided via SVP, Remote Console or host command set.

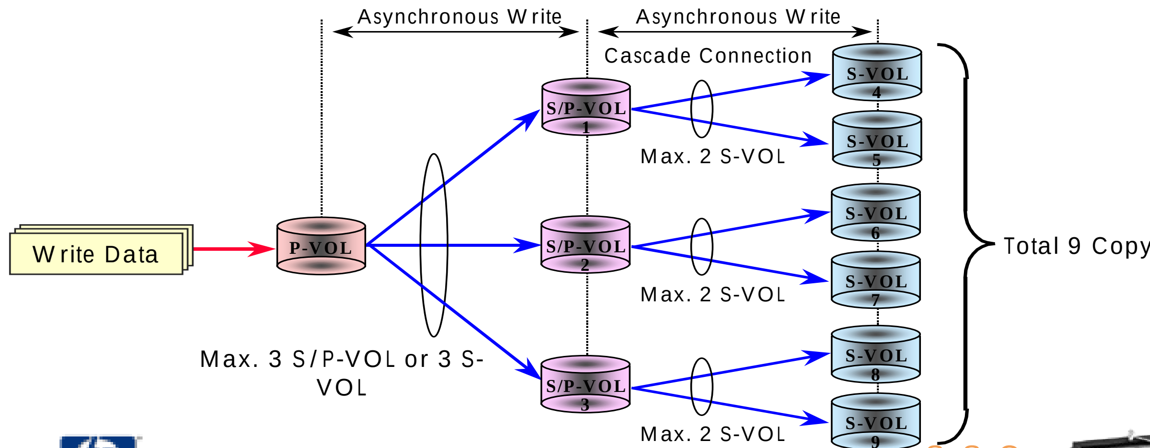


High performance is achieved via asynchronous (opportunistic) data movement to the mirror(s).

- XP replies “write complete” to the host as soon as data is written to cache memory.
- Data in cache memory is asynchronously written to the P-VOL & S-VOL at convenient times

BC Cascade Connection

- Max. 9 Secondary volumes per Primary volume.
- 9 Secondary Volumes consist of 3 S/P-VOL & 6 S-VOL (Cascade Connection).



HP Restricted

NSSO Stress-Free Storage...Guaranteed

This slide gives you a detail of how the tiering works on HP's nine copy implementation, and this is important because of the way you organize your data and organize your customer's applications. You will need to know how this tiering is implemented and what the impacts are of the internal architecture. So what you have is off your primary volume, which is where data is primarily written or de-staged from the cache, you have immediately three copies that are tiered off the primary copy. To get to the remaining copies, those are actually tiered off of the middle tier or the P/S volumes. So this is what we call a cascaded connection. Since you're running asynchronously from the primary, the data that comes out onto the very final copies of the cascade are actually derived from the middle tier. This is important because how you treat the middle tier of copies determines to what happens to the two that are cascaded off of it. So for instance, if you were to break the relationship between the primary and say that first top drive there in the middle tier, that will also break the two downstream. They will wind up being split also, from the primary. So the picture gives you a good feel for the internal architecture of this. I would recommend that if you only need three copies or less, stick with one level of cascading and that will just make your entire solution simpler to administer and simpler to handle any changes that you might have to make because once you go to a tiered architecture you'll have to careful of process and timing, when you decide to break those, and when you decide to create them. But there would be good business reasons for using cascades, for instance, you wanted to have different times that you've backed things up. For instance, you have a Monday, Tuesday, Wednesday backup, you could use the tiered architecture to cover those and then you'd have subsets of your Monday, Tuesday, Wednesday backup, for instance. One of the selling points here is the fact that we actually can accommodate a total of nine copies. This is a pretty powerful feature because again, that we do have customers that have requested up to seven copies of their data and when you size this thing, when you size this customer's solution, you have to take into account what kind of RAID implementation they want to use, and of course how much storage space that would require to implement such a solution. I would recommend in virtually every case that we use 36 gigabyte drives and use RAID 5, and what we're attempting to achieve with this solution is high density. We want to pack as much data as possible into the available drives we have. Reason being, is that you all ready say have a one terabyte working set for your customer, there is, every copy that you produce would theoretically use the equivalent one terabyte of usable space. Therefore, if your primary copy is one RAID 1, for instance, 15 gigabytes drives for maximum performance, it's OK to have your business copies on say 36 gigabytes drives running at the lower RPM on RAID 5. Because it's an asynchronous event, you do not hold up the primary disk waiting for the copies to complete. It's completely asynchronous, and there will not be any performance degradation on the primary. For backup, 36 GB drives on RAID 5 are more than sufficient to keep up with whatever tape devices you will be connecting for doing backups. Because the 36 GB drives running in RAID 5 mode read very well. And because these are sequential reads when you do a backup, you will have no issue at all with attaining a 100 percent cache read-hit rate.

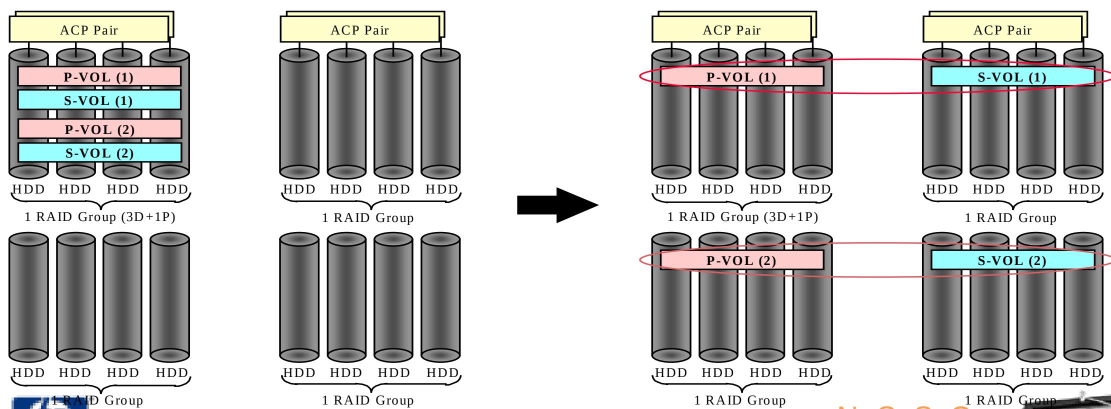
So, we will show how you should distribute these BC drives across ACP pairs.

But I want to point out here that all of these copies are protected, which means that you do not have a RAID 0 choice – it's forced to be RAID 1 or RAID 5. And the translation is done automatically for you, so that your primary can be a RAID 1 and your secondary can be a RAID 5, or reversed. The only thing that you have to match are the Open emulation sizes. So that means Open three to three, Open nine to Open nine, and with our new Open eights, it'll be to an Open eight. That's the only restriction we really have on the conversions. Keep in mind again and stress the point that BC mirrors are updated asynchronously and opportunistically (vs. synchronously for CA) which means that complete (physical) data equality is only assured just after the completion of a split or a suspend. BC can be combined with CA, but since CA is synchronous and BC is asynchronous, it is not recommended that CA receive it's data from an asynchronous BC mirror. The default amount of I/O parallelism is 4 I/Os within the entire BC hierarchy, but can be set as high as 16. After the parallelism limit is exceeded, I/Os are FIFO queued. A higher number will make BC operations faster at (potentially) the expense of host I/O performance. BC mirror copy performance can also be traded off against host I/O performance by setting the number of concurrent tracks (to be copied per I/O) within the range of 1-15. (default is 3) at pair creation time. A delta track table in shared memory will keep track of changes while the BC pair is suspended. At resync. time, (best case copy speed is ~33MB/s per pair with a default of 4 pairs updating simultaneously (as mentioned above). The P-VOL will update the S-VOL tracks that have been altered on either the P-VOL or S-VOL. BC manages the delta table bit map by track, CA manages the delta table bit map by cylinder. Pairs must be assembled and split top down only. You can't pair (or split) a third-tier S-VOL without first pairing (or splitting) the P>P/S pair above it.

BC RAID Group Configuration

• BC Paired Volumes should be dispersed across RAID Groups.

- Low BC Performance due to Bottle Neck at RAID Group.
- Well-Balanced Layout of RAID Groups and DKA Pairs.



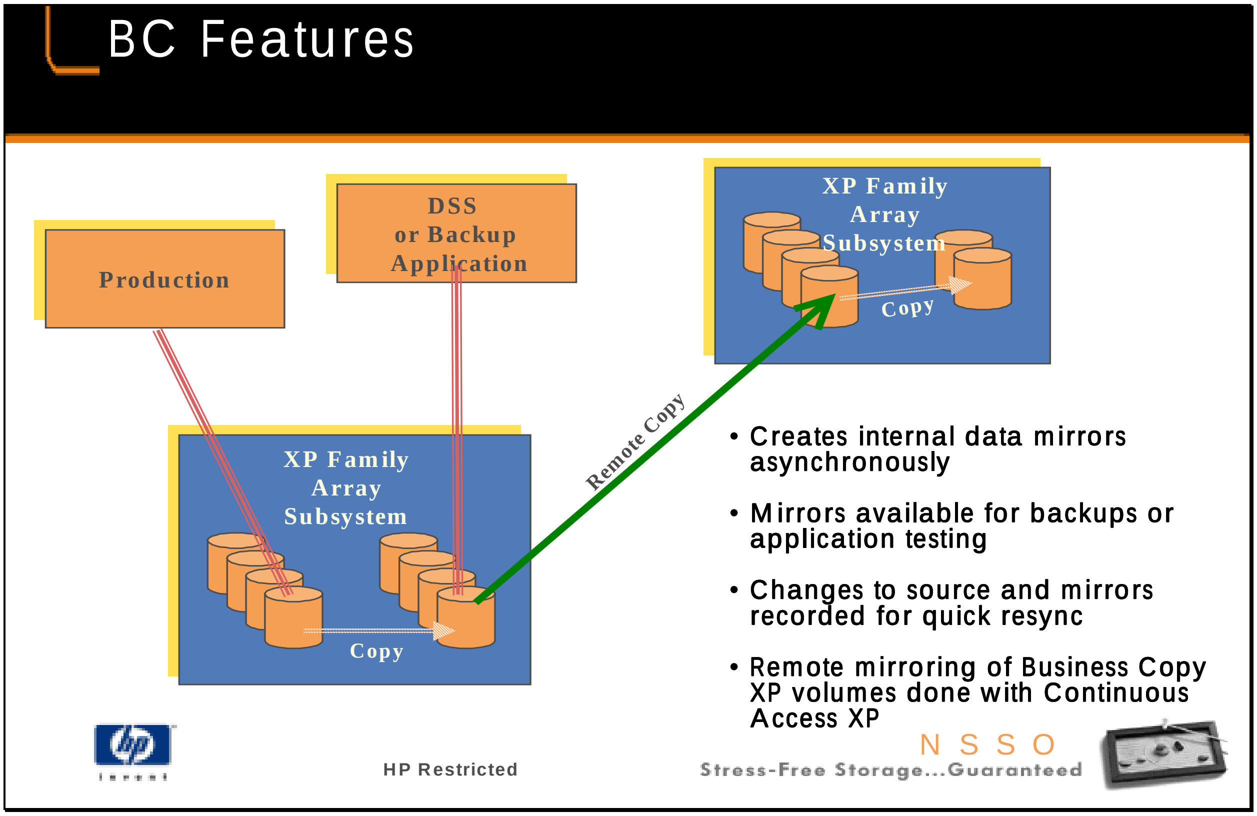
In a cascaded configuration, data updates flow outward from the Primary volume (i.e. P-VOL) to the S-VOLs. BC pairs should be spread across more than one RAID group. BC pairs should be spread across more than one ACP pair. This gives you some performance layout information that's critical to your customer's solution. The first example on the left shows you how Business Copy performance is affected due to a bottleneck in a RAID group. Most of the time, this is probably not desirable. In other words, you would want to balance and spread your copies across other ACP pairs. If you put all of your copies so the primary volume and secondary volume are all on the same ACP pair, or DKA pair as referred to in the manuals which is a HDS name, this would mean that this will run fairly slow. Because every write you do into the primary volume will also generate traffic on those buses for all the secondary volumes on the same SCSI chains that your primary was on. So this would slow down your primary volume, due to the fact that you're sharing bus bandwidth and hardware resources with your secondary copies.

The example on the right-hand side would be a more balanced approach, which means that you put your primary volumes on a DKA pair, or ACP pair, and you would push your (S-volumes) to a different pair.

One of the basic reasons for doing this is per my previous example of RAID 1 being backed up by RAID 5. We know that we cannot mix RAID 1 and RAID 5 on the same ACP pair, so you would have to use a different ACP pair group, which is exactly what we intend. So here your primary volume and your secondary volume will not compete for the same SCSI resources – SCSI bus resources and CPU cycles on those ACP pairs.

An example of how you would actually architect this is let's say that you're running a database of some sort, and you're doing quite a bit of read-write access to your database – for instance an OLTP application will probably run approximately 50 percent read-write.

So that means that our primary volume is responsible for both the read traffic and the write traffic, and we'd probably want that to be on a RAID 1, 36GB drive configuration, or for really high performance, a RAID 1, 15GB RAID group for maximum performance. But the secondary copies, the (S-VOLs) – could very easily be on another ACP group running RAID 5 on 36-gb drives, and they will be behind. Because by definition, that's an asynchronous event. It's not a problem, because you will never – should never – access the secondary volumes until after a successful split. So the fact that they're out of sync is not an issue at this point. Once you actually wish to use them, and you split the volumes, the array firmware will bring them up to date and make them identical at the time the split command was issued. At the moment you get back the status that the split has completed, then you can open those volumes for read-write accessibility, and they are exactly synchronized to the point in time where you issued the split command. So, just for review on this slide, the thing that's important is to distribute and properly balance your primary and your secondary volumes so that they do not compete for resources. The one exception where you might want them to be on the same ACP pair is if you've by chance put a development system onto the same array as your production system. I don't recommend this, but in the cases where customers are purchasing only a single array, you have no choice. In order to help isolate the load of the evelopment platform from the rest of the production platform, I would then recommend putting all of the data from the developed platform onto a single ACP pair, so that they don't compete for resources with the production volumes. It's probably one of the few cases I can think of where it would be desirable to limit them. That also means then that everybody uses the same RAID group – RAID format, which is RAID 1 or RAID 5. They're forced to, due to the fact that the ACP pairs will not mix RAID types.



Continuous Access (CA) can be used in conjunction with BC to provide multiple copies of Lus at both local and remote sites. CA enables you to create and maintain remote copies of the data stored on the local disk array for data backup and disaster recovery purposes. CA operations can be performed across distances of up to 43 km. This slide gives an example of how you can combine BC and CA and have them work together. You could have a Disaster Recovery & Backup at the local site. You could elect to have Disaster Recovery & Backup at the remote site, or you could elect to have Disaster Recovery & Backup at both the local and remote site. Because the Business Copy (BC) image mechanism is independent of the Continuous Access XP software, you can have another copy available on the remote site. And of course this copy may be put in R/W protection to do some specific actions.

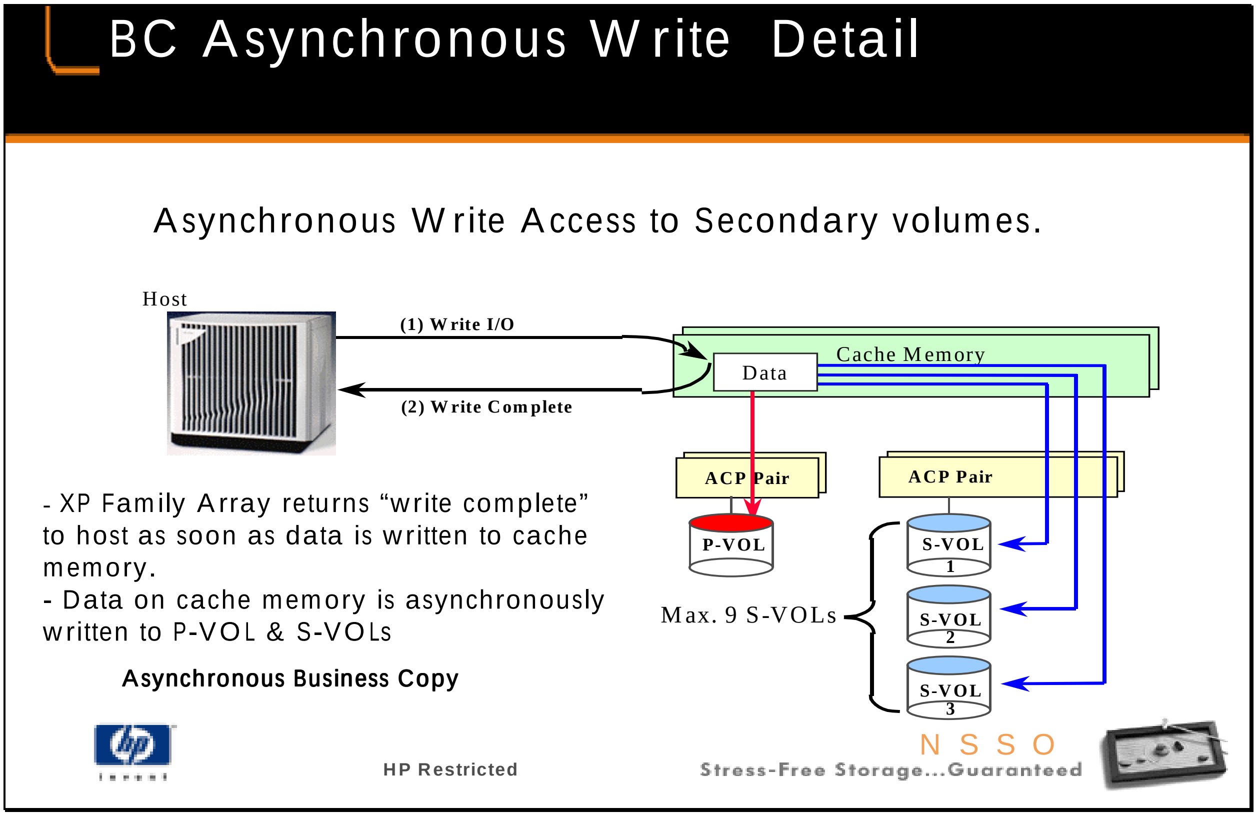
Volume Pairs (P-VOLs and S-VOLs)

- When creating volume pairs, data on the P-VOL is copied to the S-VOL to synchronize the volumes.
- P-VOLs remain available to all hosts at all times for read/write I/O.
- S-VOLs become available for host access only after you split the pair.
  - All write operations to the S-VOL are prohibited before the split.
- While the BC volume pair is split, all changes to the P-VOL and S-VOL are logged.
- When you resync the pair, the differential data in the P-VOL (due to P-VOL and S-VOL updates) is copied to the S-VOL.

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Set Reserve Attribute Operation (Attribute). The set reserve attribute operation reserves a volume so that it can be used as a Business Copy S-VOL. The disk array rejects all write I/O operations to reserved volumes. You can reserve up to 512 volumes in one disk array. Reset Reserve Attribute Operation (Attribute). The reset reserve attribute operation un-reserves a volume so that it can be mounted and accessed by all hosts. After you reset the reserve attribute, the disk array will accept all subsequent I/O operations to the volume. Add Pair Operation (Paircreate). The add pair operation establishes new volume pairs. The volume which will be the P-VOL must be in the simplex state (no pair affinity), and the volume which will be the S-VOL must be in the simplex state, reserved or not. Initial copy operation takes place when you add a new volume pair. The initial copy operation copies all data on the P-VOL to the associated S-VOL(s). The P-VOL remains available to all hosts for read and write I/Os. Write operations performed on the P-VOL during initial copy operation will be duplicated at the S_VOL by update copy operations after the initial copy is complete. When adding pairs, you can select the pace for the initial copy operation: slower (one track ata time), medium (three tracks at a tme), and faster (fifteen tracks at a time. The slower pace minimizes the impact of BC operations on disk array I/O performance, while the faster pace completes the intial copy operation as quickly as possible. Update copy operation updates the S_VOL of a volume pair after the initial copy operation is complete. Update copy operations only take place when the pair status is in duplex state. Split pair operation performs all pending S-VOL updates to make the S_VOL identical to the state of the P-VOL when the split command was issued. You will then have full read/write access to the split S-VOL. You can split volume pairs as needed, and you can also use the split pair operation to split and create pairs in one step. The split pair capability provides point-in-time backup of your data. Re-sync pair operation resynchronizes split and suspended volume pairs. For resynchronization of a split pair, the disk array merges the P-VOL track map into the S-VOL track map and copies all flagged tracks from the P-VOL to the S_VOL. This not only ensures an accurate copy of the current P_VOL is copied to the S_VOL, but also greatly reduces the time needed to resynchronize the pair. Delete pair operation terminates BC copy operations to the S-VOL of the pair and changes the pair status of each LU to simplex. After you delete a pair, the S_VOL is still not available for write operations until it is unreserved.



This is some detail on how the write actually occurs when running Business Copy. Again we'll point out that it is asynchronous. This means that when your host performs a write, it writes into cache memory, and this is true for every write whether it's Business Copy (BC) or not. Once the data is in the cache and it's properly duplexed (meaning 2 copies in cache), and the array then responds back to the host that the write is complete. The host will go and do additional work and generate future writes, but what we're interested in is what happening with-in the XP Family Array. What happens immediately, or the best possible time possible, is the data is de-staged to the primary P-VOL. The data is also de-staged to all of the S-VOL(s), on a lower priority so it will copy out there to the S-VOLs at the best possible speed but not before the primary write is made. So the important things to note here are, that the reason this is done is by looking at a worst case scenario where you've created nine total copies of your data. That means for every 8K blocks that you write into the array, you're going to generate another nine writes of that 8K block of data somewhere in the array. The effects on all of that I/O traffic taking place is that you will burn back-end response time and also burn some of the access time from all the ACP pairs on the back end that are actually controlling the drives. Therefore, if you are writing at a very high rate, for instance lets say 100 megabytes a second, for instance, there is no way that you would be able to maintain that kind of traffic across the remaining nine (S-VOLs).

Business Copy Volume States

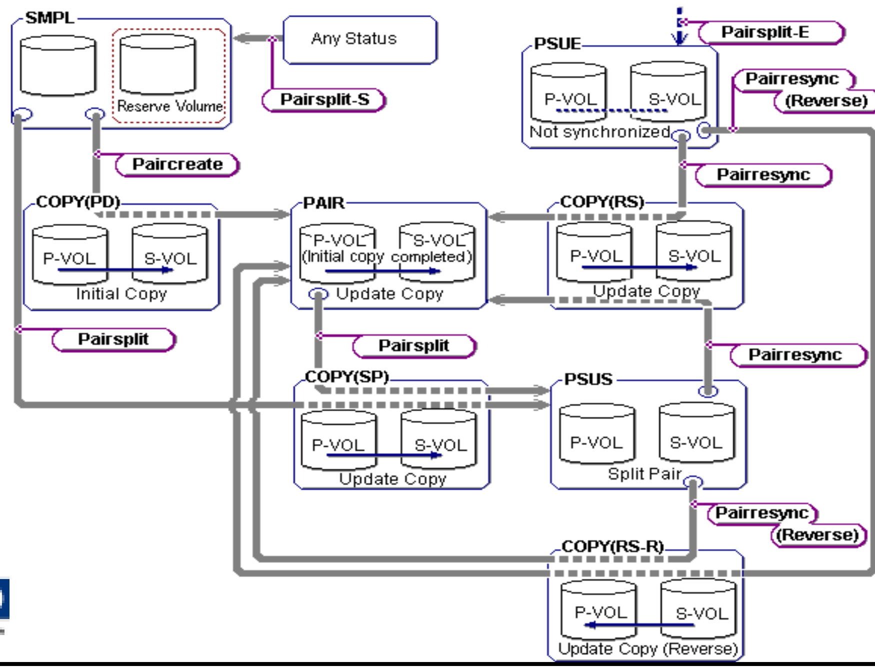
simplex . . . . . not currently paired
pending . . . . .initial copy in progress
duplex . . . . . . initial copy complete, asynchronous updates are in progress
splitpend . . . .updates issued prior to a split command are in progress
split . . . . accepting write I/Os for P-VOL and S-VOL, maintaining difference data on track maps
resync . . . . . . P-VOL and S-VOL are being resynchronized, updates are in progress
suspend . . . . asynchronous updates to S-VOL have stopped, track maps are not being maintained

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

BC Terminology (split/paired/cascading)
Within this document, “paired” will be used to mean that two volumes are linked (i.e. have an affinity or logical [if not immediately physical] data equality connection with one another) in a supplier/consumer data movement model where the consumer (S-Vol) is not writable. “Suspended” will be used to connote that the S-Vol is separately writable and the P-Vol (supplier) still has a relationship with the consumer (S-Vol) but is not currently keeping the S-Vol up to date. The P-Vol and S-Vol will keep a delta table that will be used to re-synchronize with the S-Vol at pair-resync time. “Resync” will be used to indicate that either a delta or full copy will take place to make the P-Vol logically equivalent to the S-Vol. “Delete or Pairsplit” will be used to indicate that any relationship (affinity) between a P-Vol and S-Vol is terminated.

Possible BC volume states and how do they transition?



Factors that Affect BC Operation Times

BC Operation times depend on configuration:

- Copy Performance of the Delta Data between P-VOL and S-VOL.
- Capacity of BC paired Volumes
- P-VOL & S-VOL Locations
- HDD Performance (6300rpm / 12030rpm / 10025rpm)
- RAID Level (RAID1 / RAID5)
- Time increases with increasing I/O load
- Copy Pace Parameter (Slower / Medium / Faster)

➢ Pair Re-Sync Times with 15GB drives and a medium copy pace, no host I/O and disks spread across ACP pairs

Delta Data	Time
0.5TB	4h 21m
1.0TB	8h 50m

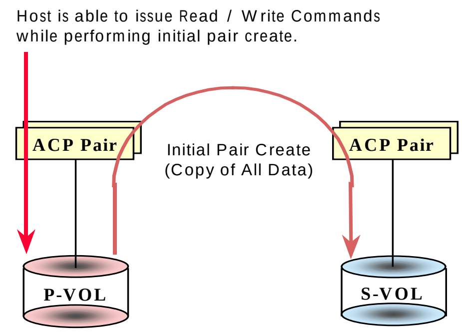
HP Restricted

N S S O
Stress-Free Storage...Guaranteed

As we reviewed on the previous slide, different aspects in designing your BC operations and the effect that they have on actual performance of the array. There are a number of factors, as indicated on this slide, that affect BC performance. Also note, on the bottom of this slide is a test scenario of several performance numbers in doing a BC operation and the time that was required to complete the copy process.

Initial BC Volume Pair Creation

- Initial Pair Creation copies all Data from P-VOL to S-VOL



HDD Type	RAID Level	Copy Pace	Initial Copy Performance	Initial Copy Time	
				500GB	1TB
47GB HDD (10000rpm)	RAID5	Medium	80MB/S	1hours 47min	3hours 39min

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

BC Pair Create Times
The time it takes to do initial pairing can be significant and depends on many factors. Initial Pair Creation time may depend on the BC configuration:

- Capacity of paired Volumes.
- P-VOL & S-VOL Location.
- HDD Type.
- RAID Level(RAID1/RAID5)
- Copy Pace Parameter(Slower/Medium/Faster), etc.

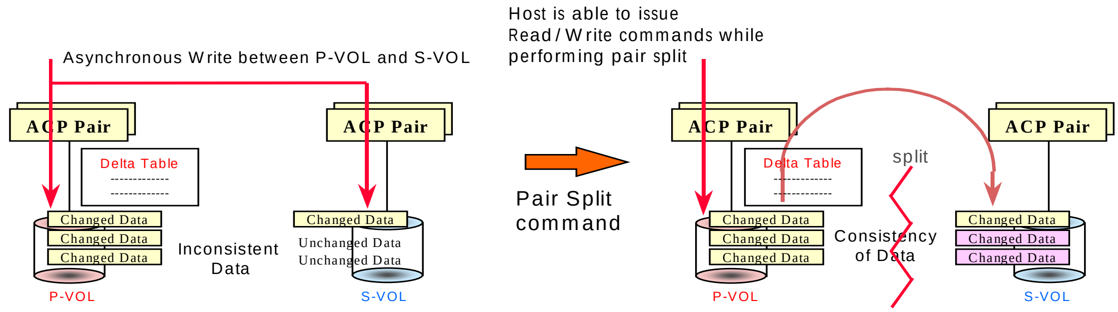
Couple of things come up when we have to do an initial pair creation. There's several things that affect how long it takes to do a pair creation. One of those factors is the full capacity of the volume that you are creating. So for example, if you have a terabyte of data, and you initiate the command to do a pair create, it will take longer than if you only had a 10-gigabyte volume to create. It sounds quite obvious and it is. We would also have to be concerned about where the S-VOLs and T-VOLs are located within the array. So for instance, the primary volumes and secondary volumes are in different areas of the machine – they will copy slightly faster than if they're competing for the same resources on the same ACP pair.
The other thing that has an effect on performance here is the speed of the drives. Twelve thousand RPM drives will copy faster than the 6,300-RPM drives. The RAID level makes a difference also. The internal mechanisms of the XP Family Array have

a concept of threads that it's copying, not actual disk mechanisms. Therefore, a RAID 5 will copy faster than a RAID 1, due to the fact that four drives are running in the case of RAID 5 versus one drive operating on the read side and two operating on the write side of the RAID 1 configuration. There's a copy-pace parameter that you can set, and my last slide gives you more description of what slow, medium and faster – what they actually do inside the array. But whatever level of the copy-pace parameter you've selected will also affect the performance. One of the things that works quite well, if you need to create a volume and at absolutely as fast as you possibly can,so you want to reduce the creation time. Two things you can do: one approach is that if you know that you're building something from scratch, for instance, you're getting ready to build a database table. As you get ready and are going to create a table, or you're going to create a brand new file system, or you're creating something you know that's going to be written in its entirety, you can do a pair create and tell it to not do a copy. That's an option that you have available, when you use the RAID managers, you can do pair create with no copy. What that does is generates the relationship between the primary and secondary volumes, but without actually copying any data. This is OK due to the fact that if you're going to create a table, 100 percent of the blocks that you're going to read will have to be re-written, which means that there's no point in copying data that you'll never use. That would be the situation if you did the pair create without the no-copy option. So therefore you would spend a lot of time waiting for that to complete, when none of that data actually ever being used. So for raw file systems for a database, just do pair create and the no copy option. The other option you have here is to use the host to copy the primary (P-VOLs) to the secondary volumes (S-VOLs) data. So therefore, in the simple case would be a simple cp of the raw space which would copy all of the simple blocks and make an exact image of the data to the S-VOLs.
Once the host is finished with the copy operation and you know the two LUNs are exactly identical, again you can do the pair, create, no copy, which tells the array that they're already synchronized. All you need to do is create the relationship and then any future blocks written to the primary volume will be essentially copied to the secondary volumes. This will be important because there isn't anything out there that will be faster than the host creating the copies. The add pair operation establishes pairing of one P-VOL and one S-VOL. In order to be added to a pair, both the P-VOL and the S-VOL must be in the simplex state. The add pair operation triggers the initial copy operation. Until the initial copy is complete, the pair is in the pending state. Once the initial copy is complete, the pair is in the duplex state, which is the "normal" state for a BC pair that is in operation and functioning properly. From the duplex state, the update copy operation is transparently performed periodically based on data changes recorded in the P-VOL "track map". Minor Notes: An LDEV must be marked by the system as "reserved" in order to be included in a pair as a S-VOL. The initial copy operation copies P-VOL to S-VOL at slow, medium, or fast.

This is configurable and should be set as desirable (fast copy - greater impact to general I/O performance versus slow copy - minimal impact on IO performance).

BC Volume Pair Split

Pair Split after Copy of only the Delta Data



HDD Type	RAID Level	Copy Pace	Copy Performance of Delta Data	Pair Split Time (1TB DB) Rate of Delta Data 1% (10GB)	Pair Split Time (1TB DB) Rate of Delta Data 5% (50GB)
47GB HDD (10000rpm)	RAID5	Medium	48MB/S	4min	18min

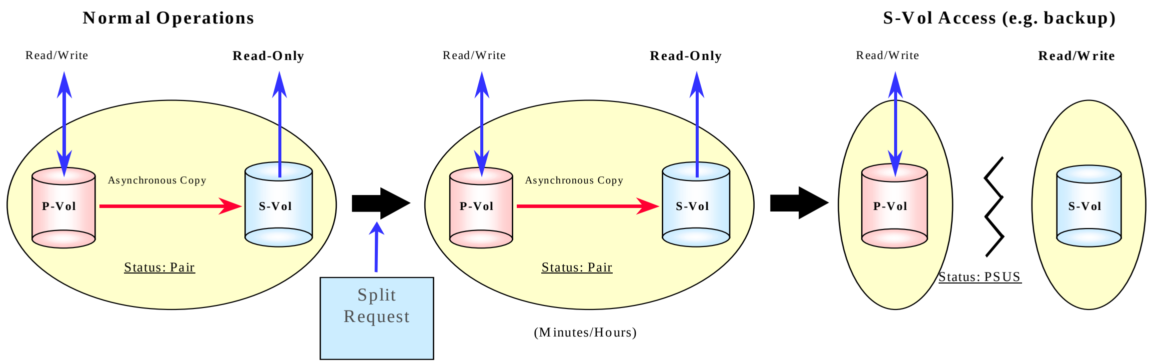
HP Restricted

N S S O
Stress-Free Storage...Guaranteed

BC Pair Split Times
It is very important that databases be in backup mode for only the shortest possible time. The fast split (2H00) functionality allows splits to complete in literally just a few seconds. Since this is a fast "virtual" split the actual background split time takes longer, although this is hidden from the user. The actual background time to complete a 1GB or 10 BG copy is shown for one combination of factors:

The split pair operation synchronizes the primary and secondary volumes by completing all pending copy operations and then splits the P-VOL and S-VOL. This operation is typically done to pairs in the duplex state; however, it can be issued to create and split pairs in one step as we had discussed on the previous slide. Until the P-VOL and S-VOL have been made identical and without any pending copy operations, the pair is in the split pending state. Once the pending copies have completed, the pair is in the split state. While in the split state, track maps for both the P-VOL and the S-VOL are maintained to record changes. One of the ways that the split pair operation might be used is to provide a "point-in-time" backup operation. Split operations cannot be performed on suspended pairs. When splitting pairs, you can select the pace for the update copy operations (slow, medium, or fast). If used to create and split a pair in one step, the initial state will be simplex. The split state is the only state which allows full read/write access to a reserved volume. These are things that will influence and affect splitting the pairs within the array. Let's review for a moment what this means. Our host has been busily writing the primary data to the P-VOL, which means we have changed data on the primary P-VOL or LUN. Our copy on the S-VOL(s) are running slightly behind the P-VOL and that depends on how fast we've been writing to the P-VOL along with how much data we have to write to the S_VOL(s). Performance numbers of five to ten percent of data that needs to be copied from the T-VOL to the S-VOL is fairly normal, which means on a one terabyte size working set, you would have five percent of that to copy or to continue copying once you initiate the split (i.e. 50GB). Most of the time you're going to be issuing the split command through RAID Manager XP on the host because you're probably going to script these solutions for things like zero downtime backup (ZDB) and so on. That split time goes into the I/O queue on the array. At that instant in time, any new write requests into the array, and we only care about writes because reads come off the primary anyway and it has no affect on the secondary. Any new write requests will be written to the primary and recorded into a delta table that is maintained in shared memory. The writes that occur after this point in time will not be queued for the S-VOLs at all. So they'll be held in the delta table in shared memory. Any I/O's, or let's say any pending I/O's of changed data, not yet copied to the S-VOL will continue and the delta table in shared memory updated. And while that's happening you should not be using the secondary volume. You should not have that mounted yet. You need to wait until that split operation completes, because only when it completes, do you know that it has actually caught up to where the P-VOL was at the that point in time where you issued the split. And then those two volumes are identical at that instant time that you had requested the split from RAID Manager XP. Two things you should note at this point is that you can continue reading and writing to the original P-VOL. This is safe to do because on the rare event, or maybe it's not a rare event, that you write on top of a block of data that was pending to be moved over to the S-VOL, there would be data corruption, if it allowed you to write on top of those blocks to be moved and then subsequently moved them over because the data would not be accurate to that point in time. What the array does is when it detects that the host is attempting to write on top of a block that is targeted to be moved, it moves it first, assuming it has not all ready been moved, it will move that block over to the S-VOL(s). Once that is completed, it will then update the block that the host requested to be written. Now, the host is unaware of this event because once the data is written in the cache, the host gets return status that says that it happened. All of this activity comes during the de-staging event of the data that the host wrote into the cache. The bottom line to take from all this is that reading and writing to the primary copy (P-VOL) during the split is completely safe and has no impact on the data integrity of the S-VOLs.

XP Family Feature Enhancement: Fast Split for Business Copy XP

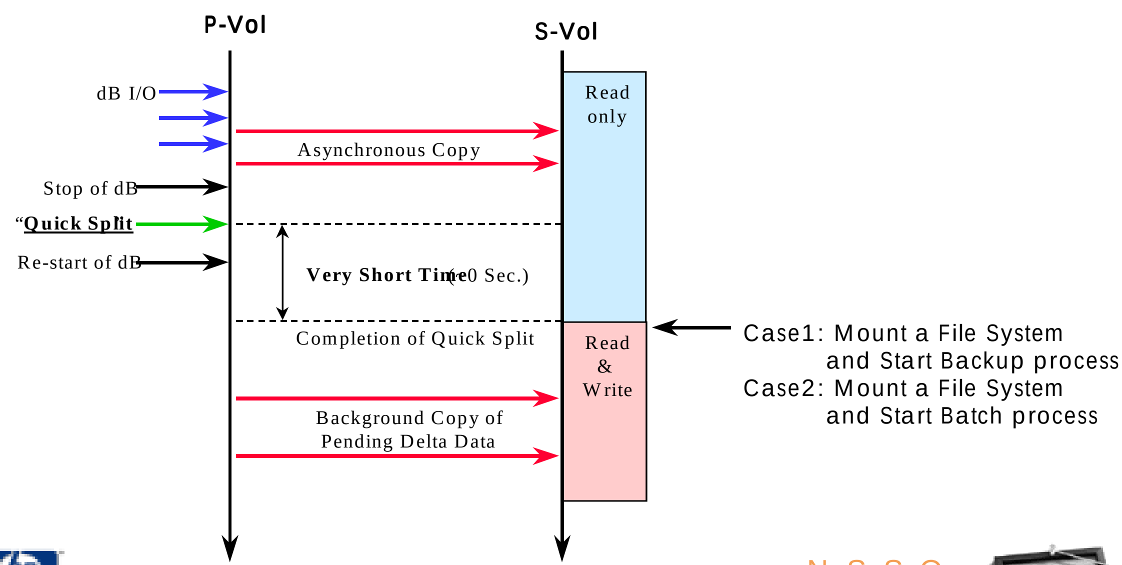


HP Restricted

N S S O
Stress-Free Storage...Guaranteed

Current Situation:
Prior to Fast Split functionality, the user must wait for the secondary copy volume (S-Vol) to be completely up-to-date or current (caught up with primary volume, P-Vol) before they could have read/write access to it. Business Copy mirrors data asynchronously and opportunistically. Furthermore, customers can choose to prioritize S-Vol currency or P-Vol I/O response time. If a customer wants to emphasize I/O response time (i.e. they do not want to impact P-Vol I/O response), or if the P-Vol is seeing a large number of write requests, the data on the S-Vol can 'fall behind'. If a customer wanted Read/Write access to the S-Vol before Fast Split functionality was present, the system would first make sure that S-Vol has 'caught up' with the P-Vol. That is, it would make sure that the S-Vol contained all data present on the P-Vol at the time of the split request. Only until the S-Vol was completely brought up to date with the P-Vol would the system release the S-Vol and grant Read/Write access to the S-Vol. This could potentially take many minutes our hours before Read/Write access to the S-Vol could begin.

Quick Split

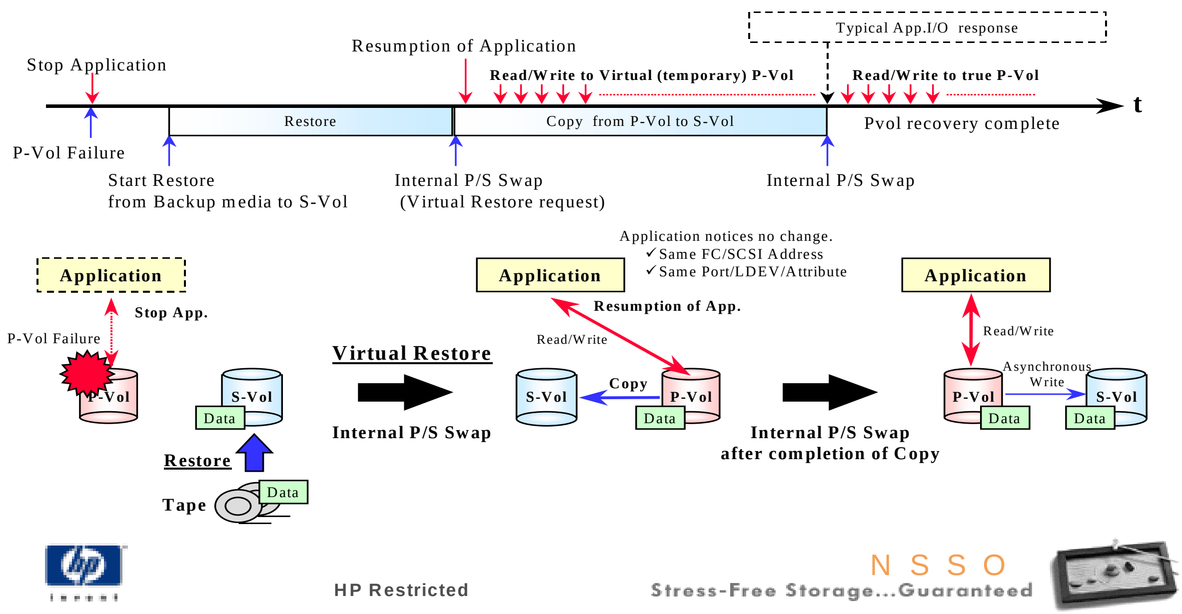


HP Restricted

NSSO Stress-Free Storage...Guaranteed

Backup/Batch processes require No waiting to accomplish "Quick Split".

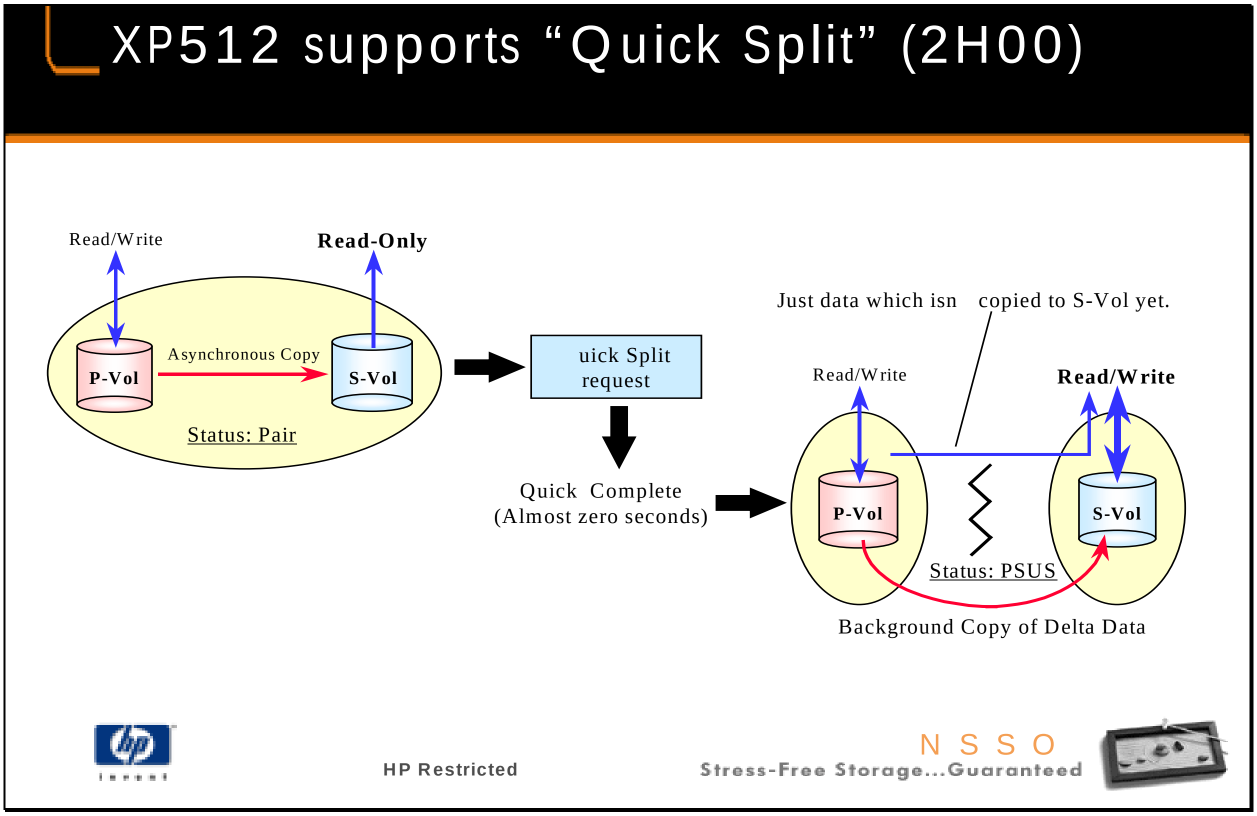
Virtual Restore 2H00



The XP512 supports "Virtual Restore" (2H00) for fast S-Vol->P-Vol restore.

• "Pairresync -restore" allows application to use the virtual P-Vol as soon as the data is restored from tape to the S-Vol. Note: as of FW rev 01-11-14/20/22 it does not do the second internal swap at it's a full copy. If the P-Vol was a faster disk than the S-Vol, this would slow down the application.

• Virtual Restore uses internal P-Vol/S-Vol swap technology.



New Situation:
With Fast Split, the user is granted immediate access the S-Vol. The system keeps track of which S-Vol blocks are up–to-date and which blocks have not yet been copied from the P-Vol. If the user wishes to Read from or Write to blocks that are up to date on the S-Vol (i.e. they have been copied from the P-Vol) then the system accesses the S-Vol directly. If the user wishes to Read from or Write to blocks which are not current on the S-Vol (i.e. the most current data is still on the P-Vol and has not yet made it to the S-Vol), then the system ‘reads through’ the S-Vol to the P-Vol for these blocks. As far as the user is concerned, they are accessing all data from the S-Vol.

Note: For Business Copy mirrors to be useful in most solutions, they must contain data that, taken as a whole, represent a usable image of the database or dataset being mirrored at a particular point in time. For instance, for a split mirror backup to be effective, Business Copy Mirrors must represent a ‘recoverable’ image of the database being backed up for them to be useful in a recovery scenario. This usually requires that the database be ‘quiesced’ to allow a stable, point-in-time image of the data of the P-Vol to be established so that a valid S-Vol image can be generated in turn. While some applications/databases can be quiesced by ‘pausing’ them for a short period of time without dramatically affecting availability, others require that they be taken offline or shut down in order to accomplish this.

Please work with your customer to understand the split-mirror compatibility of their application/database environment.

Value Proposition:

With Fast Split, the user has immediate access to Business Copy volumes. In general, Fast Split allows for more time-efficient use of data by making data access and synchronization parallel operations. In a zero-downtime backup solution, a user could begin streaming data to a tape library at a specific point in time and would not be subject to synchronization delays potentially inherent a serialized/non Fast Split mode of operation. Most applications that would use Business Copy Volumes (e.g. Data Warehousing) benefit from Fast Split capability. With parallel access and update to the Mirror volumes, the window of time required for the Business Copy volume to be 'suspended' or split-off for offline use can be reduced. This would mean that the volume pairs could in turn be returned to Pair State more quickly, thus requiring fewer outstanding updates to be made against the mirror volumes.

An issue to keep in mind with Fast Split is that immediate access to S-Vol data can result in some I/O response time impact with the P-Vol. This is the 'no free lunch' scenario (in this case, Fast Split may not come without a price). This can happen if Read/Writes made against the P-Vol are contending with Read/Writes made against the S-Vol. In other words, if a primary host using the P-Vol is also accessing the blocks being read through from the S-Vol, then the system will prioritize the read-through operation of the S-Vol before allowing access by the primary host. This is to ensure that the user accessing the S-Vol gets data 'present' on the P-Vol only up until the time of the Split and not updates that might have been made to the P-Vol subsequently. In most cases however, the net effect is positive. In other words, the time efficiencies gained through the parallel processing of Business Copy Volume access and updating outweighs the potential I/O impact experienced by users of the P-Vol.

Technical Specifics:

Fast Split functionality is now part of Business Copy XP with the release of following Firmware Revisions:

XP Platform

FW Revision

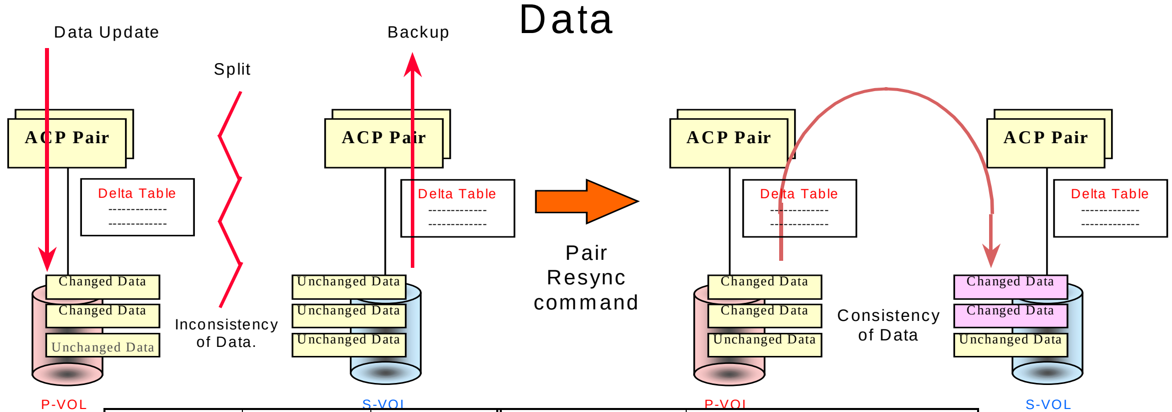
XP256

52-48-06

XP512
01-11-22
XP48
01-11-22

BC Volume Pair Resynchronization

Copy of only the Delta Data



HDD Type	RAID Level	Copy Pace	Copy Performance of Delta Data	Pair Resync Time (1TB) Rate of Delta Data	
				10% (100GB)	50% (500GB)
47GB HDD (10000rpm)	RAID5	Medium	48MB/S	36min	2hours 58min

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

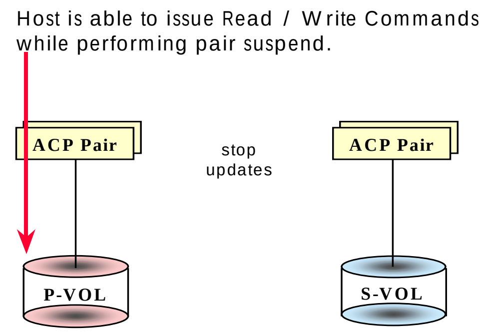
The time required for resync will depend on several factors:

- The amount of Delta Data.
- P-VOL & S-VOL Location (e.g. under same DKC?, in same RAID group?).
- HDD Type.
- RAID Level(RAID1/RAID5)
- Copy Pace Parameter(Slower/Medium/Faster), etc.

The resync pair operation resynchronizes split or suspended pairs. On a split pair, this operations merges the P-VOL and S-VOL track maps and then copies only data which needs updating. On a suspended pair, all P-VOL data is copied to the S-VOL and thus takes as long as an initial copy. While the copy is in progress, the pair is in the resync state. Once the resynchronization and copy of data is complete, the pair is in the duplex state. Normal update copy operations then resume on the pair. Minor Speaker Notes: When re-synchronizing pairs, you can select the pace for the update copy operations (slow, medium, or fast). The integrity of data read from the S-VOL while the pair is in the resync state cannot be guaranteed. This is the event where you have split the drives. And they've been split for some time and you have a delta label that's been kept in shared memory keeping track of all changed tracks to the P-VOL since the split. We then ask the array to begin to synchronize these pairs and bring them back into a consistent state, as soon as you issue the command, the array firmware goes out and looks for the delta table for the P-VOL. One note that is very important here is that it depends on which direction you want to resync. The array has the capability of resynchronizing from either direction. But the simple case, where we are resynchronizing from the primary or P-VOL, and that's the P-VOL that's been in use. When that resync command has been issued the array goes and looks at the delta table, and in the case of Business Copy (BC), it looks at tracks to see which ones need to get copied and it begins the process of copying all of that data over to the S-VOLs. In addition, the P-VOL also handles all new write requests coming in from the host. All those writes are all copied over to the secondary volumes as well. And the resynchronization will complete with the same rules we used on – on pair, create and pair, split. The delta data – so let's say for some reason you became 100 percent of the data was changed. In other words, you recreated a database table and wrote everything fresh so that every single track was marked as dirty, then you would have to copy the entire contents of the P-VOL to the S_VOL. The locations are important. Again, all the other things about the speed, the RAID, the configuration of the ACPs and the copy-pace all affect resynchronization time.

BC Volume Pair - Suspend

- BC Volume Pair - Suspend stops any attempts to update copies or to keep accurate track map difference information for the pair.
- It allows the pair to get out of sync.
- Suspend operation can be done on volume pairs in any state.



HP Restricted

N S S O
Stress-Free Storage...Guaranteed

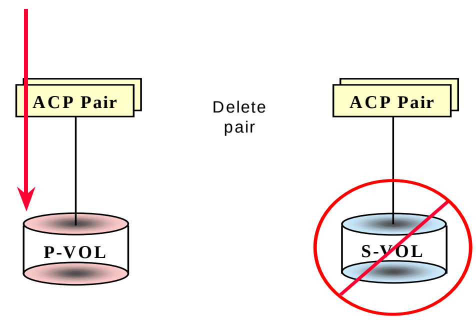
This operation suspends any attempts to update copies or to keep accurate track map difference information for the pair. In other words, it allows the pair to get out of sync. A pair may be suspended proactively from RAID Manager. The subsystem itself may perform a suspend pair operation if it detects an error during a copy operation or if the track map data becomes lost or corrupted. This is typically the only time this operation would be invoked. The suspend pair operation can be performed on pairs in any state.

Once the operation is applied, the pair is immediately in the suspended state. When a pair is suspended, the entire P-VOL track map is marked as different. This is what triggers the re-sync operation on a suspended pair to do the complete copy of P-VOL to S-VOL. The P-VOL remains fully accessible even if the pair has been suspended by the subsystem itself on detection of an error.

BC Volume Pair Delete

Host is able to issue Read / Write Commands while performing pair delete.

- BC Volume Pair - delete pair operation removes the association between the P-VOL and the S-VOL
- Delete operation can be done on volume pairs in any state.

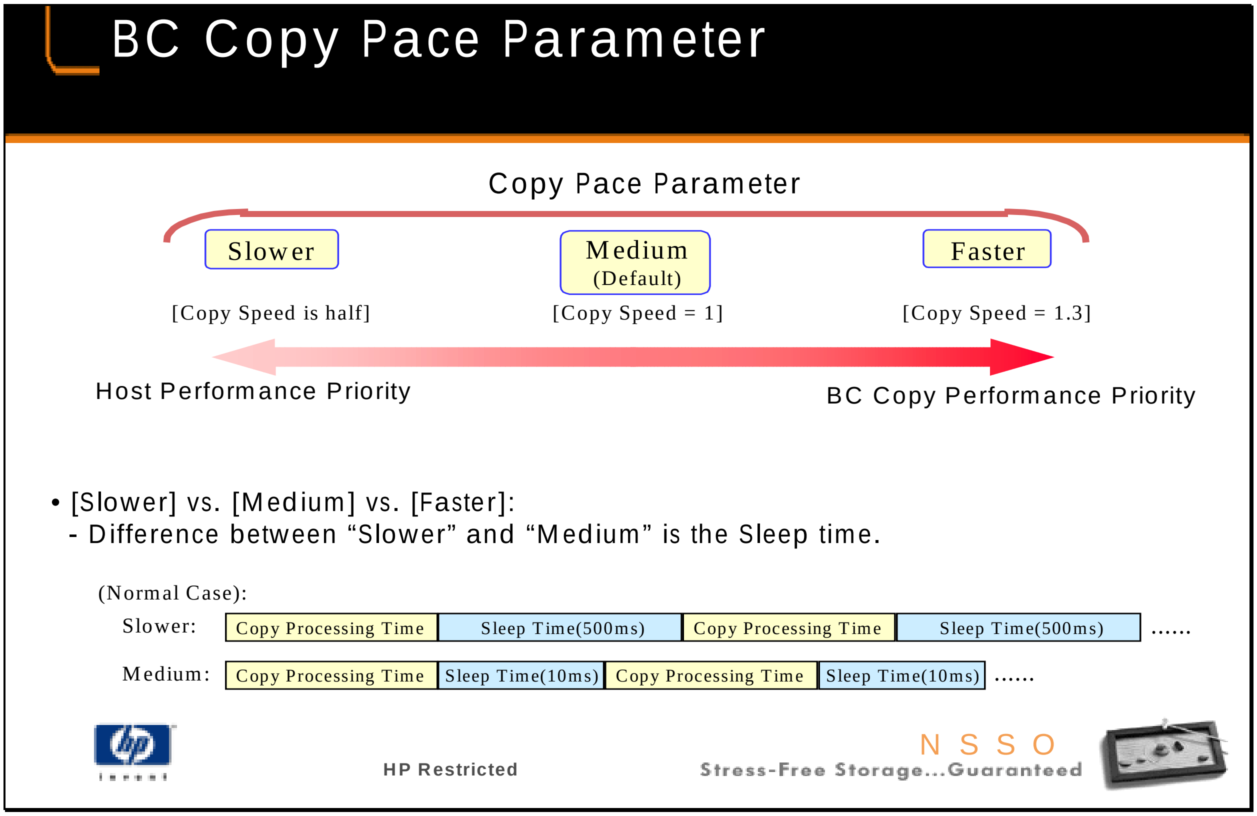


No longer S-VOL

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

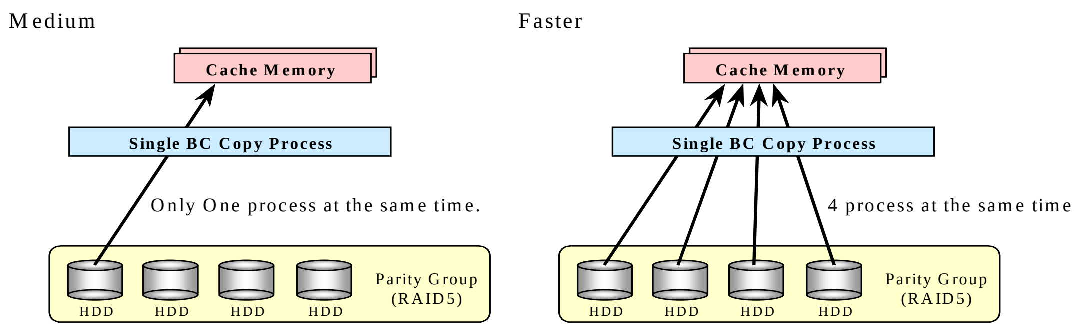
The delete pair operation removes the association between the P-VOL and the S-VOL. A delete pair operation may be performed on pairs in any state except simplex. Once performed, the pair is immediately returned to the simplex state. Note that once deleted, the volume which was the S-VOL will need to be “unreserved” to allow general I/O operations (re-enable write access) to the volume. The P-VOL and S-VOL of a duplex pair may not be identical at any given time. If it is desirable to synchronize the volumes prior to deleting the pair, a split pair operation should be performed first. The delete pair operation does not provide any synchronization or final copy updating prior to pair deletion.



The copy/paste parameter gives you some options on how quickly you want the array firmware to move data from the primary (P-VOL) to secondary (S-VOL) volumes. If you select slower, there's a process that wakes up at regular intervals to go out and do some amount of copying and then goes to sleep again. So the slower pace affects this sleep interval. And it's approximately 500 milliseconds. I don't know that you need to worry a lot about this because in every case that HP has worked on and every customer solution that has been implemented, we used nothing but fast. It has been HP's position to always go as quickly as possible on the copy operation. The medium speed here, all it does is change the amount of sleep time that it waits between each time the copy process wakes up and copies some data. Now for fast, or "faster" as what it is actually called on the service processor (SVP). It's the same as medium with one difference. And that is, it allows more copies to happen in parallel, as opposed to being just absolute linear and sequential. So it still has the 10 millisecond sleep time between copies, but it runs the threads in parallel. Also, it affects the priority. For instance, if you have the host, who is still writing to the array, it always has priority over the Business Copy operations. So, in fact what this would mean is that if you are writing a very high volume or reading a high volume of data to the array, the copy performance will slow significantly. If there aren't any host requests at all, then medium and fast will perform the same. They run, pretty much, uncontested because there isn't anything to prioritize against.

So the recommendation with this is to always use the fastest possible setting. And there is an option when you do pair-create, where you can tell it that for every time you do a copy, go ahead and do 15-tracks at a time, which is the maximum number.

BC Copy Pace Parameter



“Medium” vs ”faster” is a matter of the number of copy processes assigned.

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Business Copy General Considerations

- To determine disk capacity required for Business Copy
  - Use the P-VOL disk space and add the space used by the number of copies (1 - 9).
- Must have either Remote Control XP installed on the Remote Console PC and/or Raid Manager XP installed on the host.
- Stop the specific instance of Raid Manager before changing it's configuration file.

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

During a BC create operation a maximum of 33 MB/s will be consumed
from the array backend by 16 parallel LDEV to LDEV copy operations.
Typical LDEV copy performance for 16 threads will be 25 MB/s. Business Copy runs on the Remote Control PC, as an optional component of the Remote Control XP application or from the command line interface of RAID Manager (RM) on the host. Each copy of RAID Manager is know as a Raid Manager instance. Instances are local or remote and can run on the same host. Each BC pair which consists of a P-VOL and a S_VOL is linked to an instance or RM. Each instance of RM can manage multiple volumes and manage either P-VOLs or S-VOLs. For example, instance 0 of RM is linked to the P-VOL for pairs a,b, and c. Instance 1 is linked to the S-VOL for pairs a,b, and c. Instances can be on the same or different host systems. The host that is running the instance must have access to the volumes to which it is linked and have access to a disk array command device for the array. We will talk about the components tat comprise RAID Manager XP later in this class.

Business Copy Rules of Performance

- Each volume of a BC pair should be on a different ACP pair.
- Each volume of a BC pair should be on different Array groups.
- The Mirror Volume should be on a RAID 5 group w/36 GB drives. (For better price)
- Assume 25MB/s usage during copy operation. (Create and Resync)
- For increased performance, track copy size should be set to 15, in RAID Manager
  - paircreate ... -c 15
- When in sync mode multiply additional IOPs per copy to ACP.

HP Restricted

N S S O
Stress-Free Storage...Guaranteed

During a BC create operation a maximum of 33 MB/s will be consumed from the array backend by 16 parallel LDEV to LDEV copy operations. typical LDEV copy performance for 16 threads will be 25 MB/s.

Review of Key Points

- Overview & Specifications
- New Added Functionality (Up to 9 S-VOLs)
- New enhancement "Fast Split" Mode of Operation
- States of Operation Defined
- GUI & Configuration/Management Screens
- Disk & Performance Guidelines
- Wrap-UP

HP Restricted

NSSO
Stress-Free Storage...Guaranteed

Module Wrap-up

HP Restricted

NSSO
Stress-Free Storage...Guaranteed