XP and EVA fundamentals and case studies

Bob Ingram

Solutions Specialist
HP Storage

HP WORLD 2003
Solutions and Technology Conference & Expo

HP StorageWorks
The industry's most complete array portfolio

HP WORLD 2003
Solutions and Technology Conference & Expo

MSA1000

Low Cost SAN
≤6 T Bytes

- WEB, Exchange, SQL
- Simple DAS-to-SAN (Proliant)
- Windows, Linux, Netware + more

eva3000

Midrange Modular
≤8 T Bytes

- Mid—sized DB/Remote office applications
- Simplification through virtualization
- Windows, HP-UX, Linux, + more

eva5000

Enterprise Modular
≤35 T Bytes

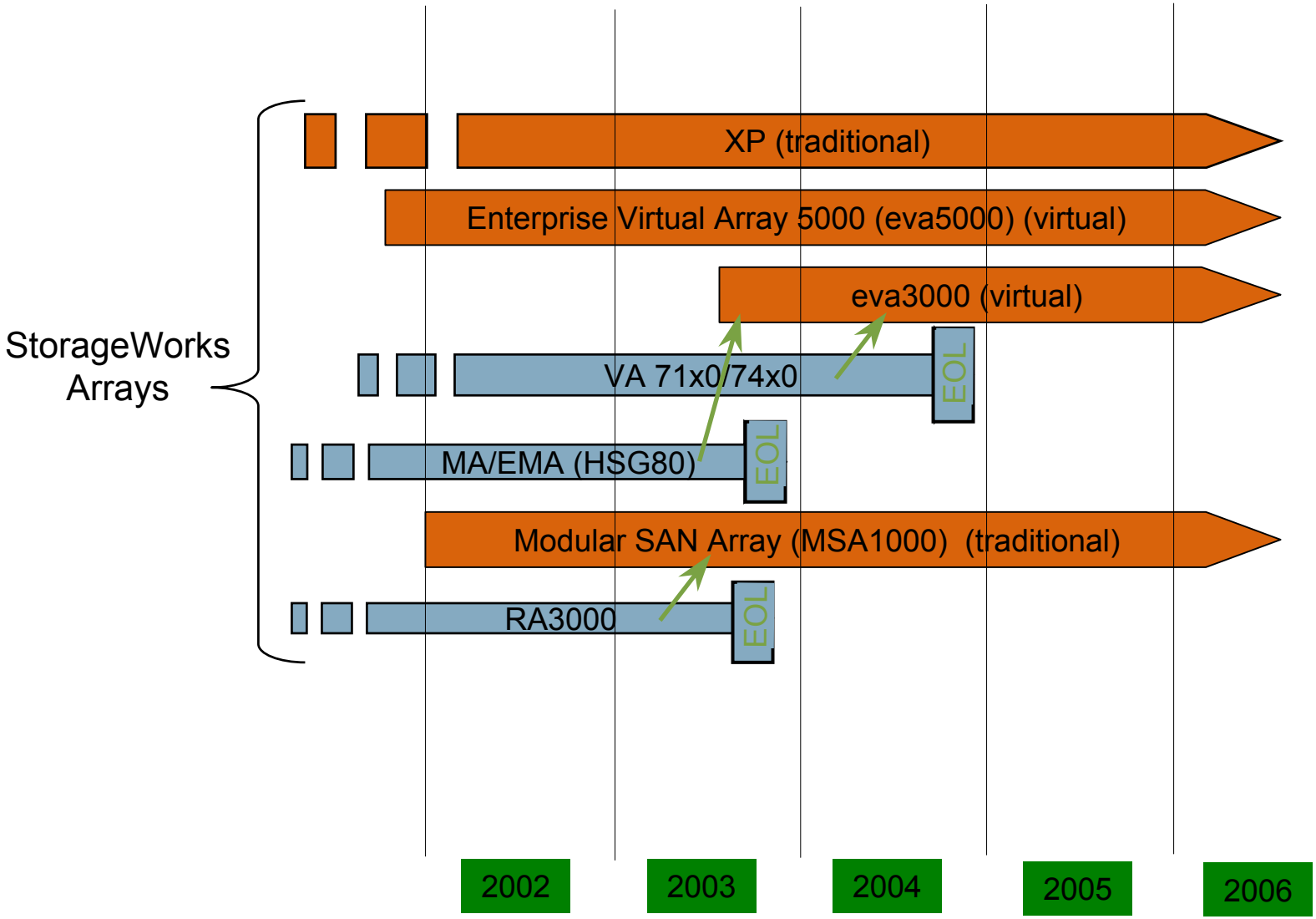
- Storage consolidation + disaster recovery
- Simplification through virtualization
- Windows, HP-UX, Linux, + more

XP

Enterprise Frame
≤149 T Bytes

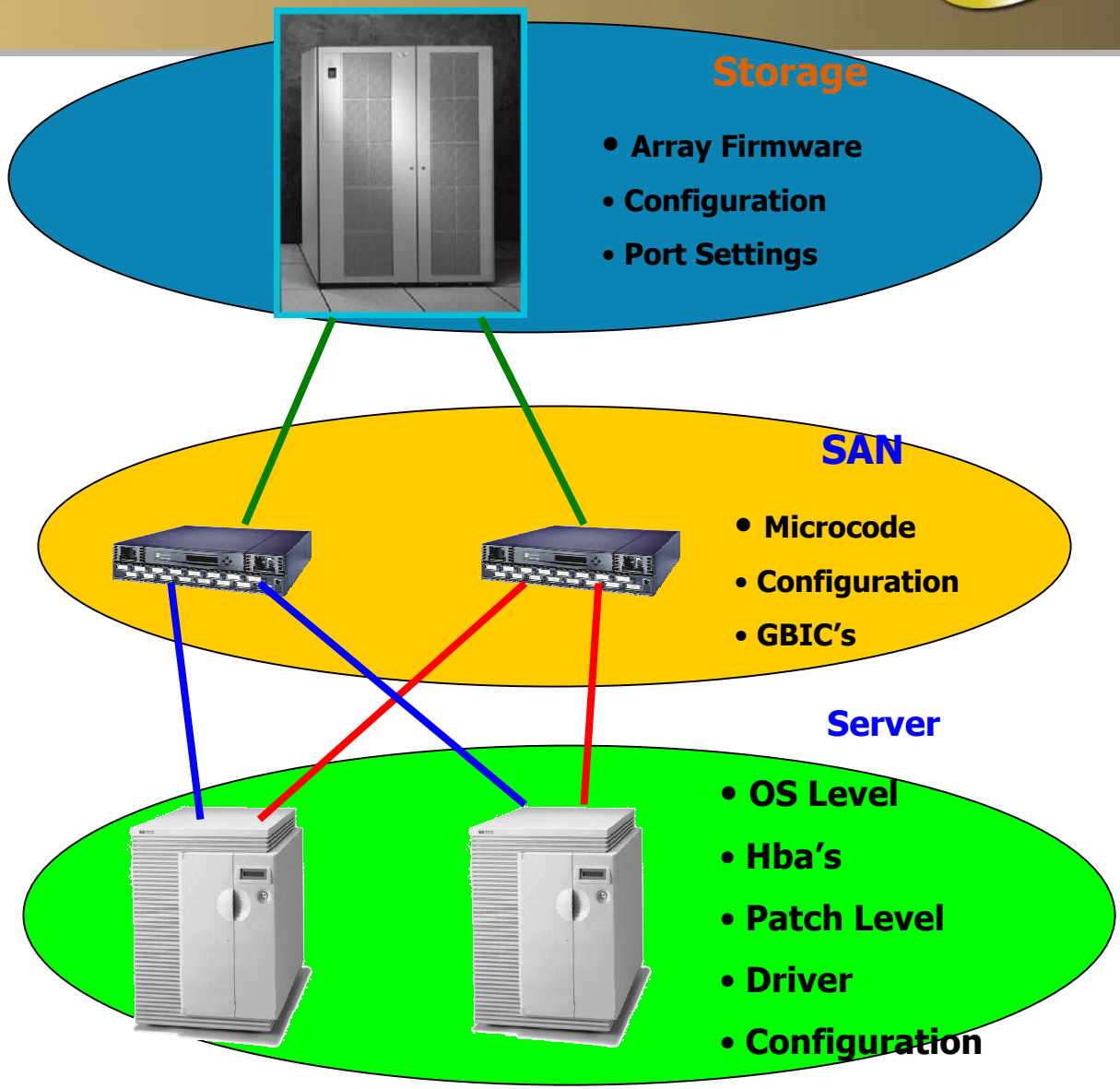
- Data center consolidation + disaster recovery
- Large scale Oracle/SAP applications
- HP-UX, Windows, + 20 more including mainframe

HP storage arrays long-term roadmap



HP SAN Testing Strategy

HP Certifies each level of the Storage "food Chain". This eliminates the guesswork other vendors subject their customers to. We do this for both HP and Non-HP products



the EVA and XP family - supported operating systems

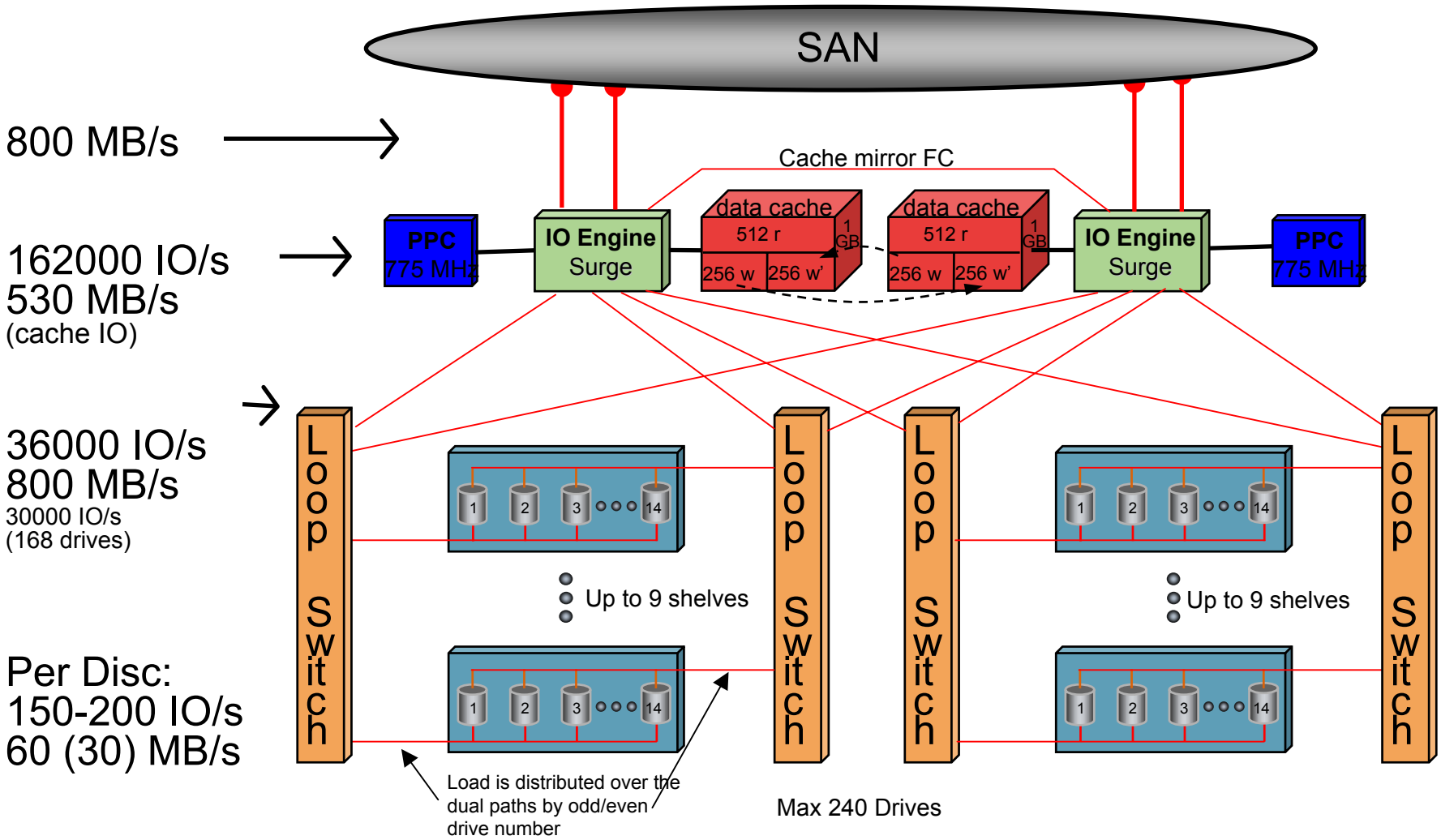
	eva3000	eva5000	xp128	xp1024
HP-UX	✓	✓	✓	✓
Tru64	✓	✓	✓	✓
SGI-IREX			✓	✓
Windows NT / 2000	✓	✓	✓	✓
Solaris	✓	✓	✓	✓
AIX	✓	✓	✓	✓
Open VMS	✓	✓	✓	✓
Linux (Intel)	✓*	✓*	✓	✓
Netware	✓*	✓*	✓ *	✓ *
Windows 2003 (.NET)	✓	✓	✓	✓
MPE/iX			✓	✓
Mainframe			✓	✓

* single path initially

EVA3000/5000

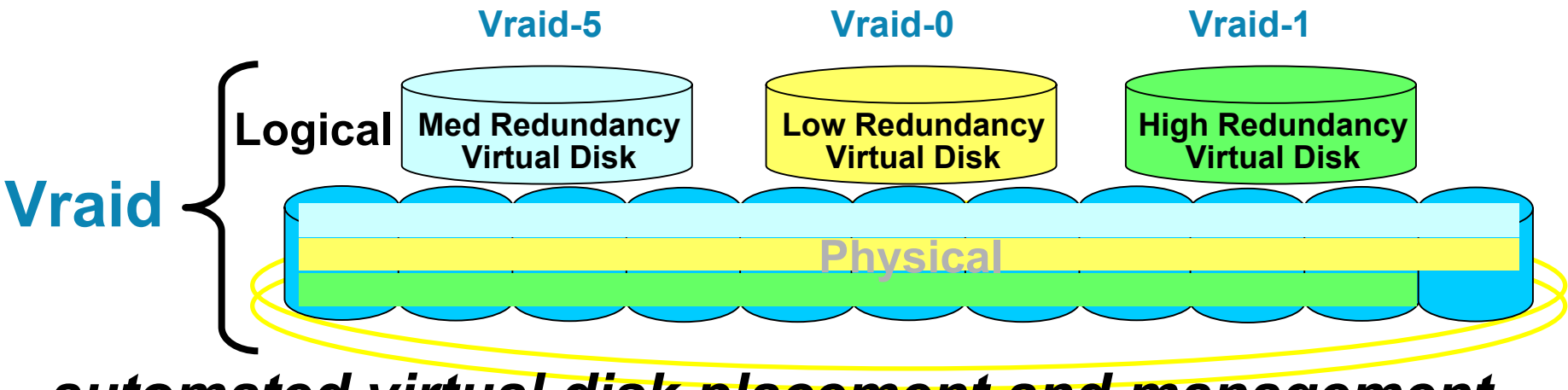
- EVA3000/5000 hardware comparison

EVA Architecture



EVA's Vraid
Bringing RAID to the next level

- JIT dynamic/seamless disk volume expansion
- avoidance of any "hot" spindles
- elimination of stranded capacity
- faster parity reconstruction
- automatic load redistribution upon pool expansion
- virtual hot sparing
- pool free space used for Vsnap update Writes



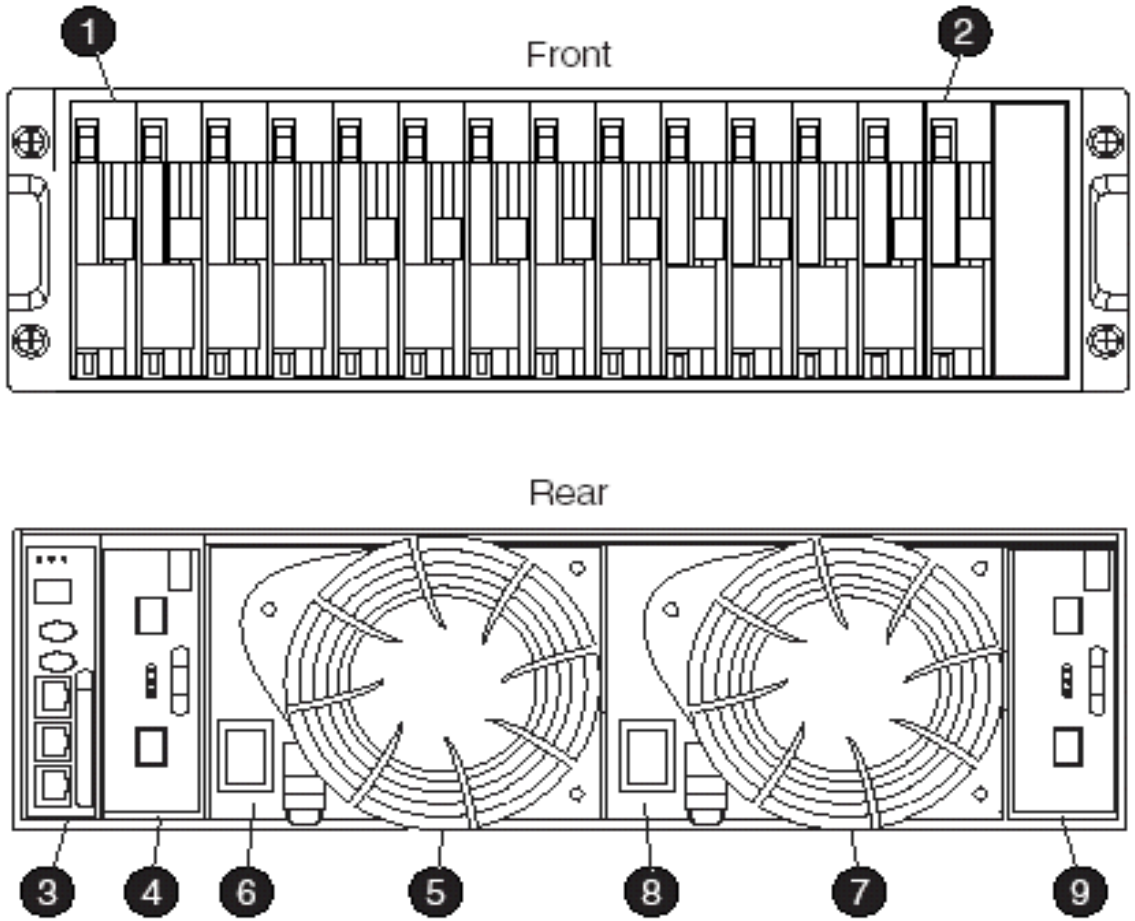
***automated virtual disk placement and management by attributes** (redundancy level, size - up to 2 TB)*

EVA 3000/5000 hardware comparison

- EVA 3000 Hardware
  - HSV100 controllers
    - Two 2Gbit Copper Back-end loops
  - M3200 Controller Enclosure
  - M5114 Drive Enclosures
  - 2 – 4 drive shelves
  - Drives
    - 36GB 10K and 15K rpm FC disk
    - 72GB 10K and 15K rpm FC disk
    - 146GB 10K rpm FC disk
  - Up to 56 drives
  - No loop switch required

- EVA 5000 Hardware
  - HSV110Controllers
    - Four 2Gbit Optical Back-end loops
  - M3220 Controller Enclosure
  - M5214 Drive Enclosure
  - 2 – 18 drive shelves
  - Drives
    - 36GB 10K and 15K rpm FC disk
    - 72GB 10K and 15K rpm FC disk
    - 146GB 10K rpm FC disk
  - Up to 240 drives
  - Supports loop switches

disk enclosure

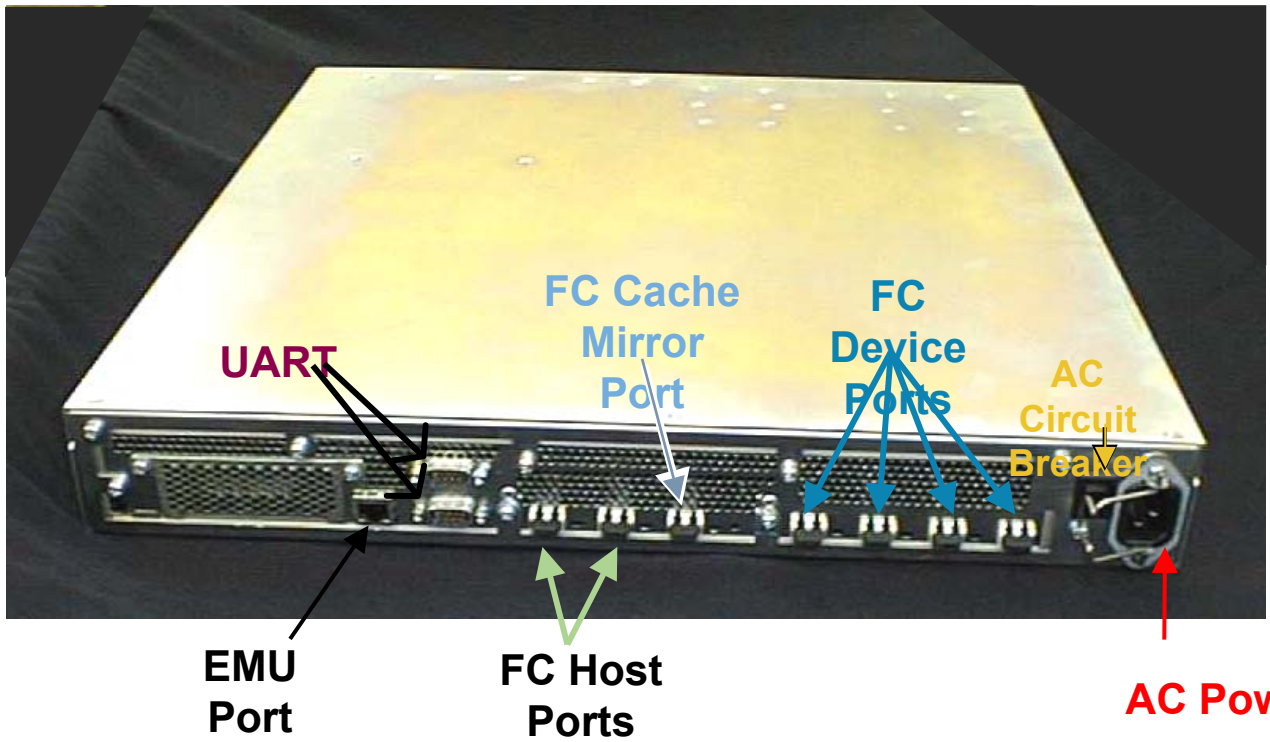


1. Drive bay 1
2. Drive bay 14
3. EMU
4. I/O module B
5. Blower 1
6. Power supply 1
7. Blower 2
8. Power supply 2
9. I/O module A

- Dual redundant copper FC loops, power supplies and fans. No SPOF
- Minimum 2 shelves and maximum 4 shelves
- Capacities of 8 to 56 FC drives for the EVA 3000
- Capacities of 8 to 240 drives for the EVA 5000

HSV110 controller enclosure – rear view

- 1-GB cache per controller-mirrored write cache
  - battery backup (up to 96 hours)
- Dedicated 2-Gbps FC cache mirroring port



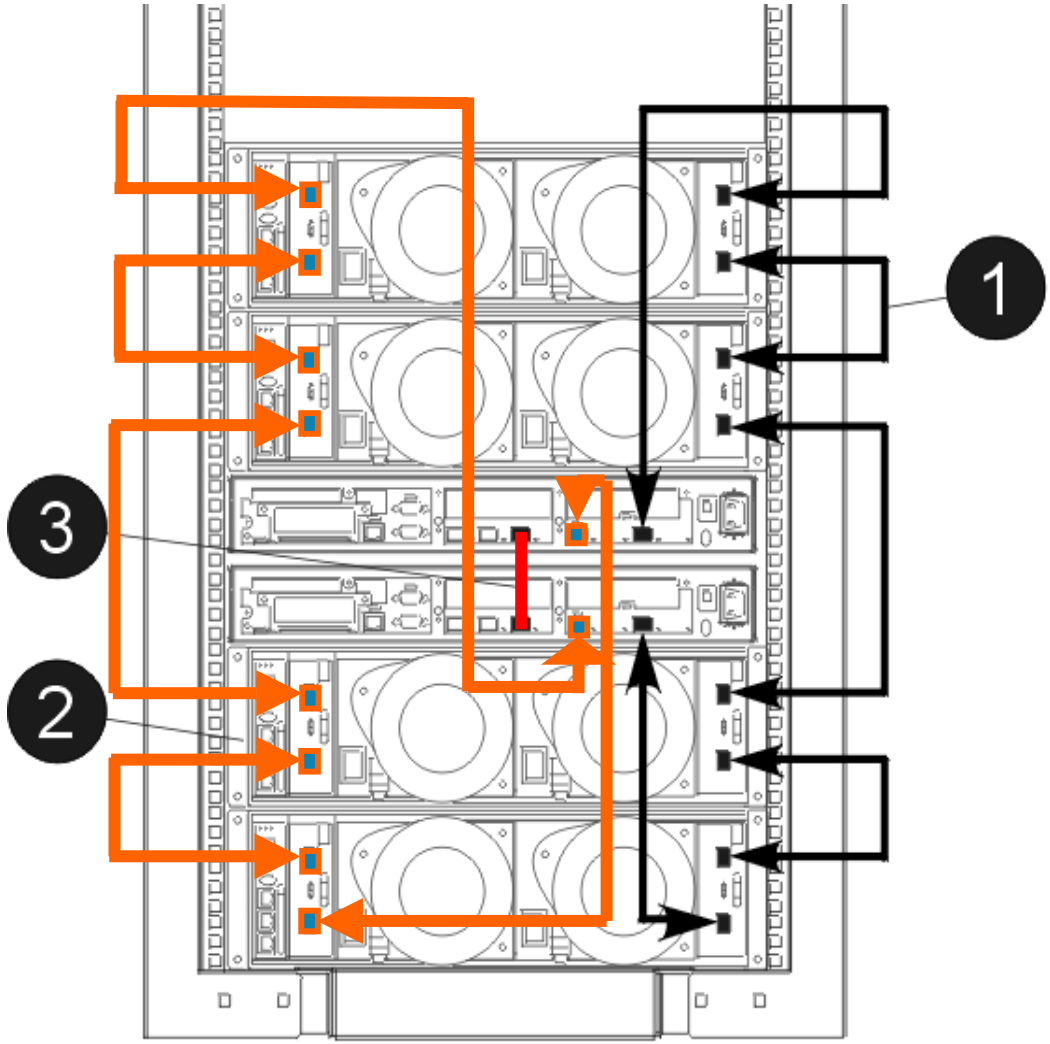
EVA 3000/5000

- EVA 3000/5000 backend comparison

EVA3000 loop configuration – 2C4D

1. Cable
2. FC drive enclosure
3. Controller-to-controller

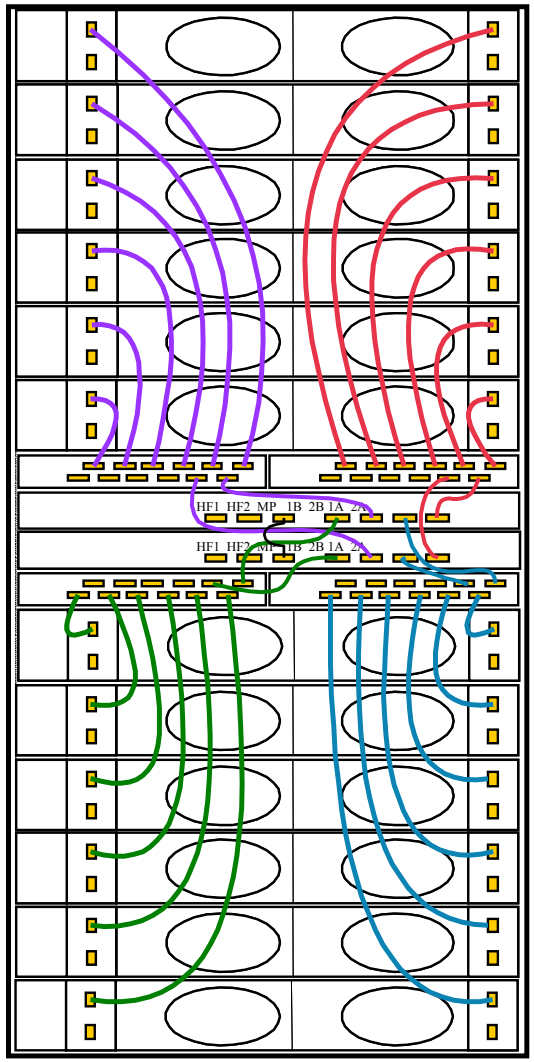
2C4D Config
2 Controllers
4 Drive shelves
Up to 56 Drives



EVA5000 loop configuration - 2C12D with FC loop switches

HP WORLD 2003
Solutions and Technology Conference & Expo

2C12D Config
2 Controllers
4 FC Loop Switches
12 Drive shelves
Up to 168 Drives



EVA 3000/5000

- EVA 3000/5000 software and OS support

EVA 3000 vs EVA 5000: Software

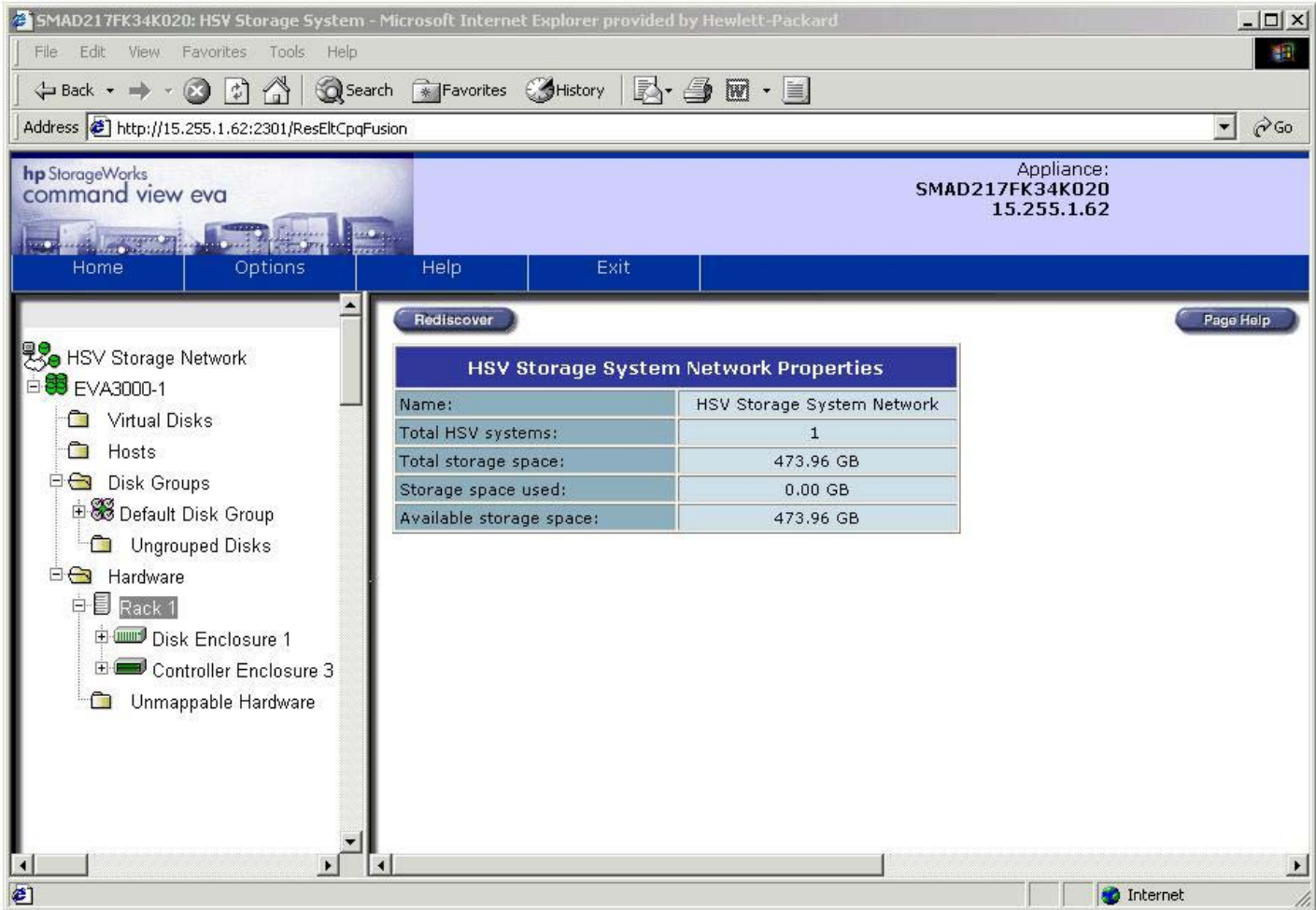
- Command View EVA 3000
  - user interface for managing the EVA 3000
- Virtual Controller Software (VCS)
  - HSV 100 VCS V2.004
- Optional Software
  - Secure Path
  - Business Copy EVA
    - Virtually Capacity-Free Snapshot (Vsnap)
    - Standard Snapshot (Snapshot)
    - Virtually Instantaneous Snapclone
  - Continuous Access EVA (DRM) 3Q2003

- Command View EVA 5000
  - user interface for managing the EVA 5000
- Virtual Controller Software (VCS)
  - HSV 110 VCS V2.002, V2.003, and V3.0 (CA-EVA)
- Optional Software
  - Secure Path
  - Business Copy EVA
    - Virtually Capacity-Free Snapshot (Vsnap)
    - Standard Snapshot (Snapshot)
    - Virtually Instantaneous Snapclone
  - Continuous Access EVA (DRM)
    - Requires VCS V3.0

EVA 3000/5000

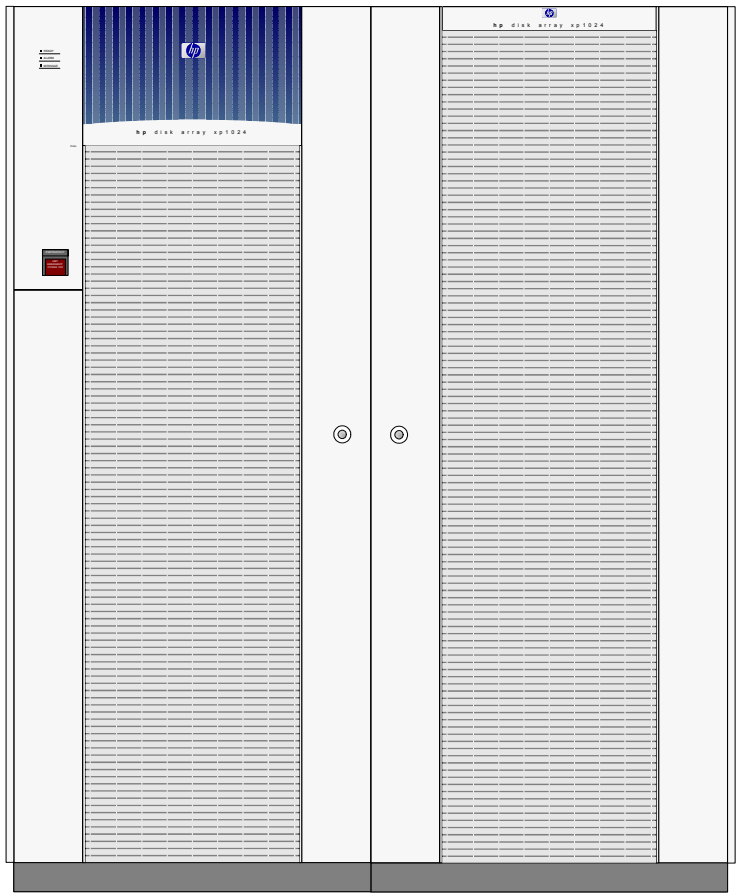
- Command View EVA V2.1
  - HSV110 and HSV100 controller management
  - EVA-Continuous Access (EVA 5000 only)
  - Previous called element Manager
- GUI to manage the EVA3000 and EVA5000
- Create disk groups
- Create virtual disks
  - Vraid level selection
  - cache write policy selection
  - host presentation
  - controller failover/failback
- Manage storage system hardware
- Create Snapclones and Snapshots of virtual disks

command view eva

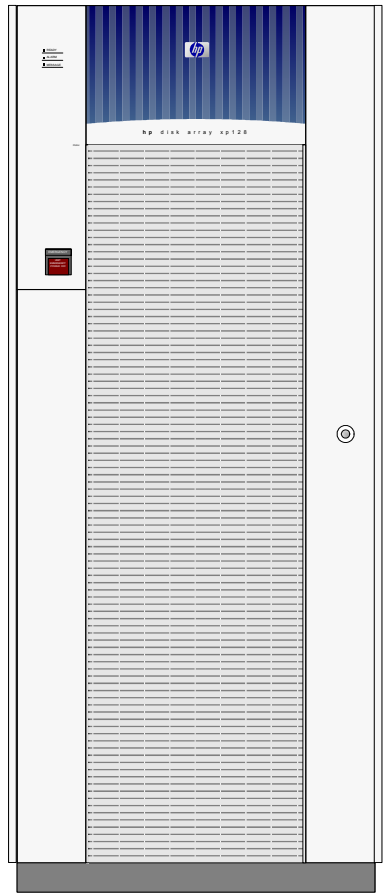


hp StorageWorks xp1024/128

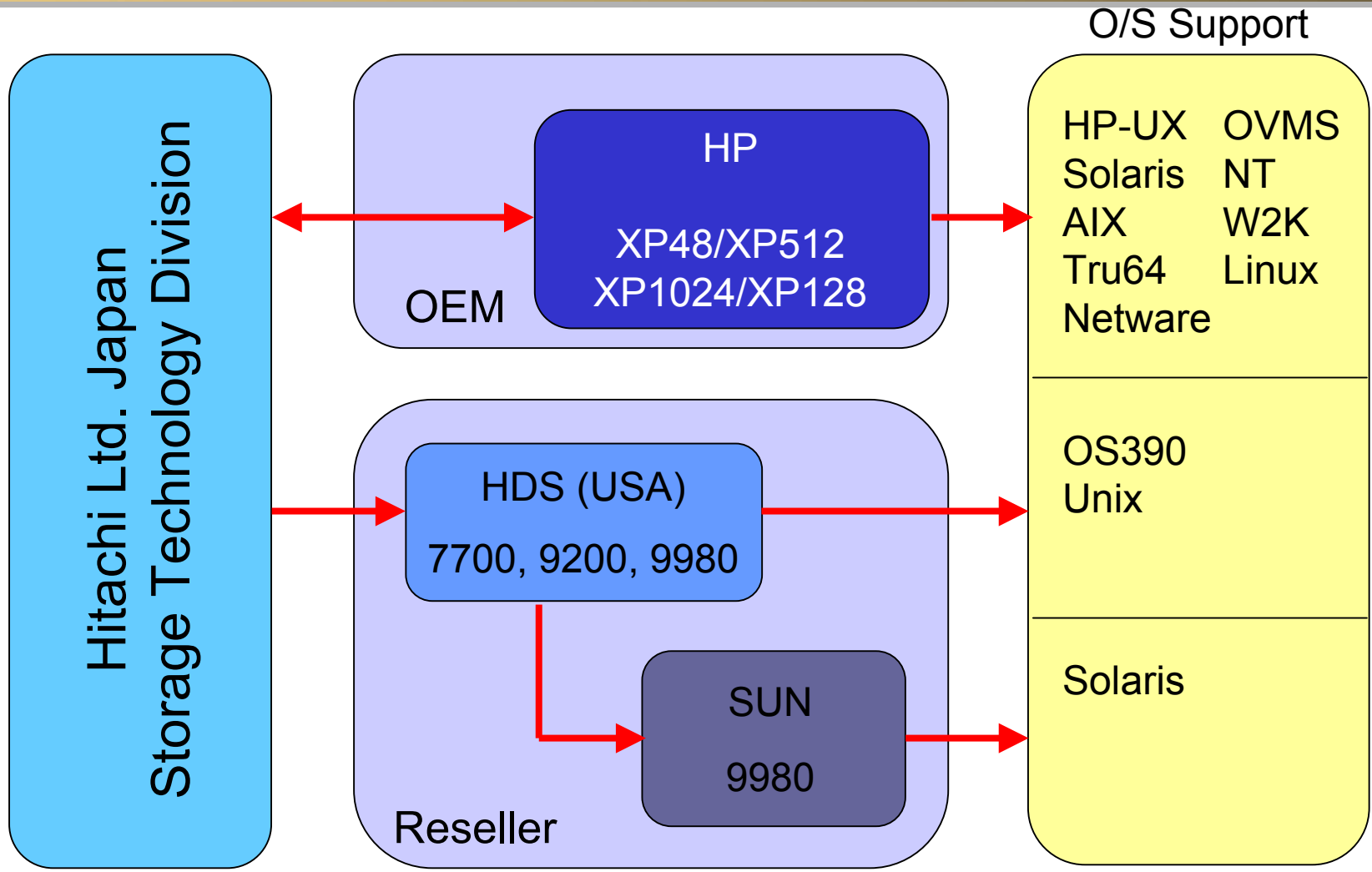
hp StorageWorks xp1024



hp StorageWorks xp128

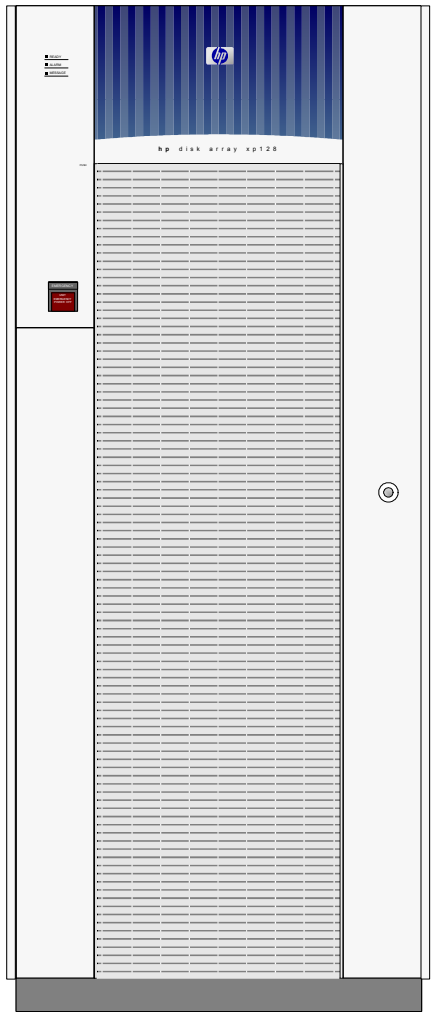


xp1024/128 – *HP & Hitachi Partnership*



hp XP128

- Single cabinet
- Controllers in front
- Disks in back
- 4-128 disks



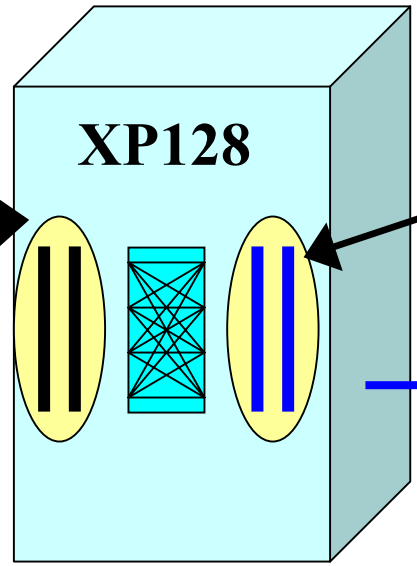
- RAID 1/0 2D+2D
- RAID 1/0 4D+4D
- RAID 5 3D+1P
- RAID 5 7D+1P

XP128 Performance Overview

HP WORLD 2003
Solutions and Technology Conference & Expo

Front-end Performance:
100% Cache Hits
- **375,000 (best case)**
- **187,000 IO/sec (typical)**
- **2,400 MB/sec**

Back-end Performance:
Cache Avoidance
- **27,000 IO/sec**
- **1,100 MB/sec**



ACP Pair Limits:
16,000 IO/sec
570 MB/sec

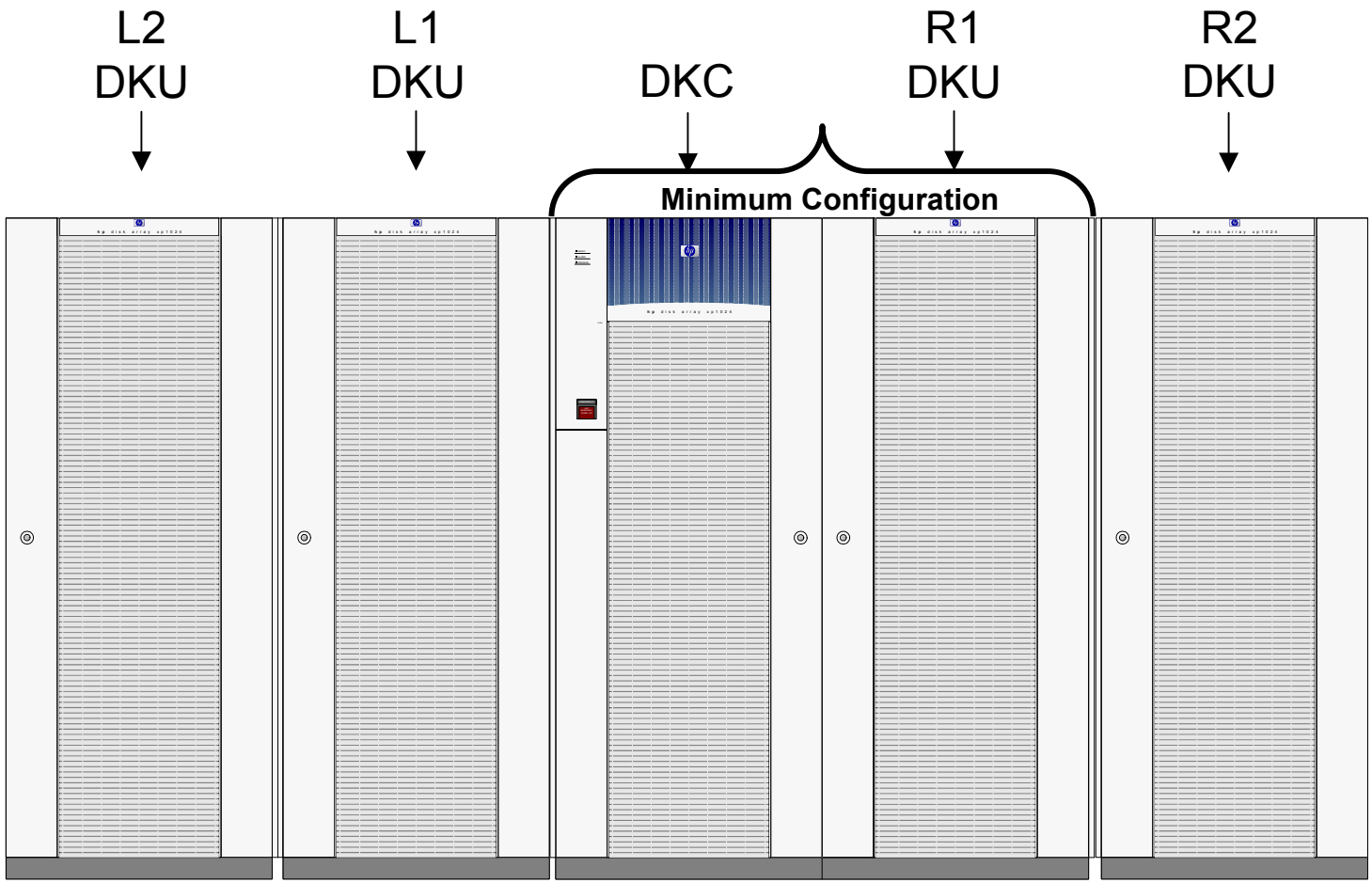
CHIP Pair Limits:

	100% Cache Hits CACHE	Cache Avoidance BACKEND
IO/sec	70,000	30,000
MB/sec	960	810

1 Array Group Limits:
850 IO/sec
90 MB/sec

Configuration: RAID 1, 32 Array Groups, 73GB Disks, 32 GB Cache, 1 and 2 ACPP, 1 and 2 CHIPP, 8 FC ports per CHIPP, FW rev. 13-02-09-00/00, N-Class Hosts, 8 Tachlite ports each; Random Performance: reads, 2k block size; Sequential Performance: reads, 64k block size. Note: ACPP-Limit and 1-Array Group tests are cache avoidance. All values in this overview are sustained.

hp XP1024

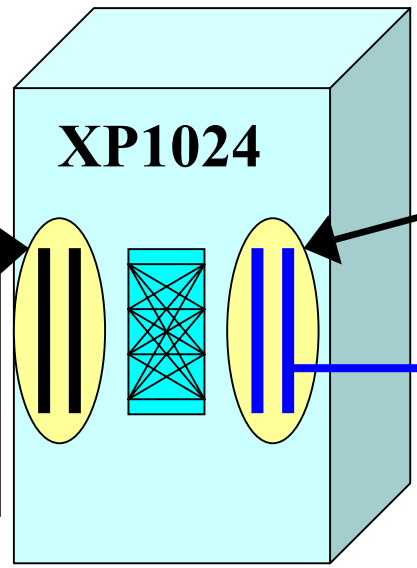


Multi Cabinet solution
4 to 1024 disks

XP1024 Performance Overview

Front-end Performance:
100% Cache Hits
- 500,000 (best case)
- 250,000 IO/sec (typical)
- 3,200 MB/sec

Back-end Performance:
Cache Avoidance
- 57,000 IO/sec
- 2 GB/sec



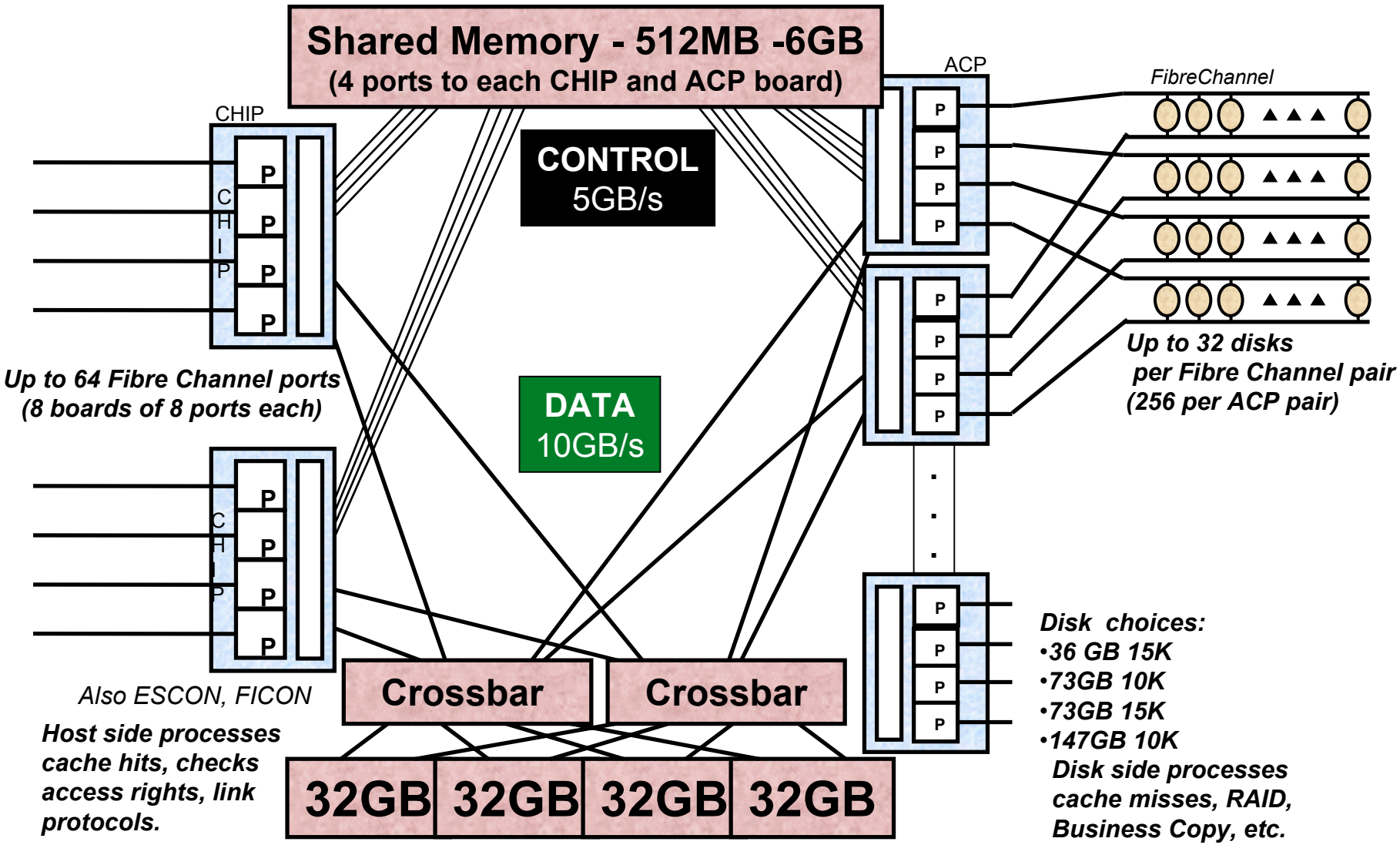
ACP Pair Limits:
16,000 IO/sec
570 MB/sec

CHIP Pair Limits:

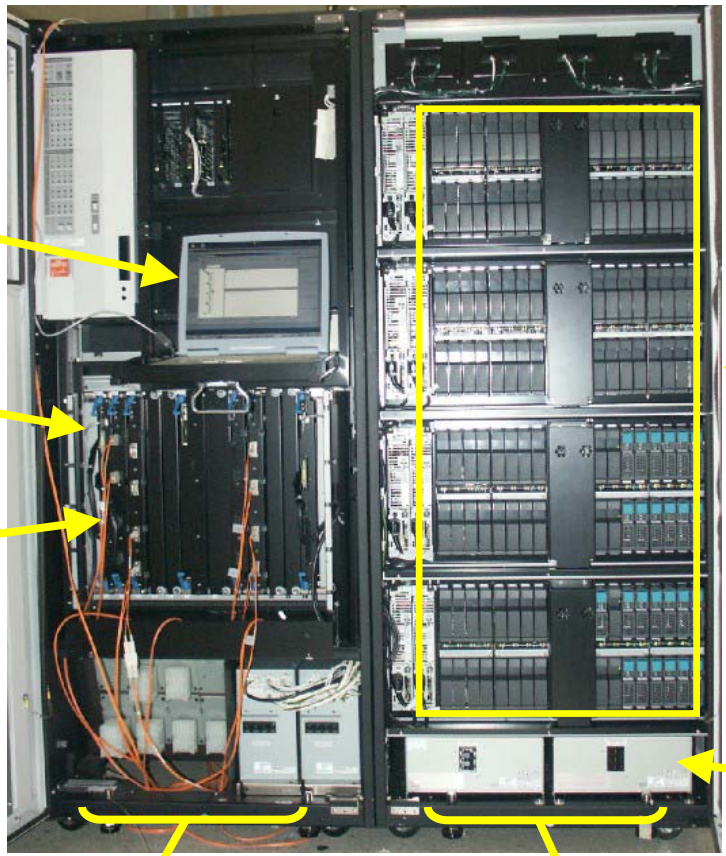
	100% Cache Hits CACHE	Cache Avoidance BACKEND
IO/sec	70,000	30,000
MB/sec	960	810

1 Array Group Limits:
850 IO/sec
90 MB/sec

hp XP1024 – *Internal Architecture*

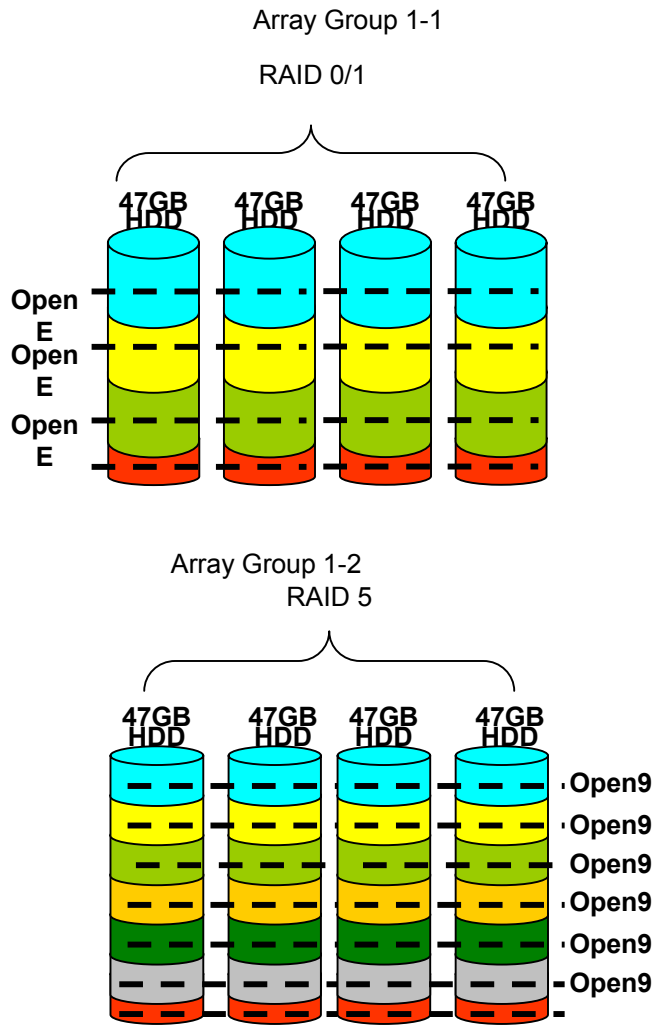


hp XP1024 – *Physical View (front)*

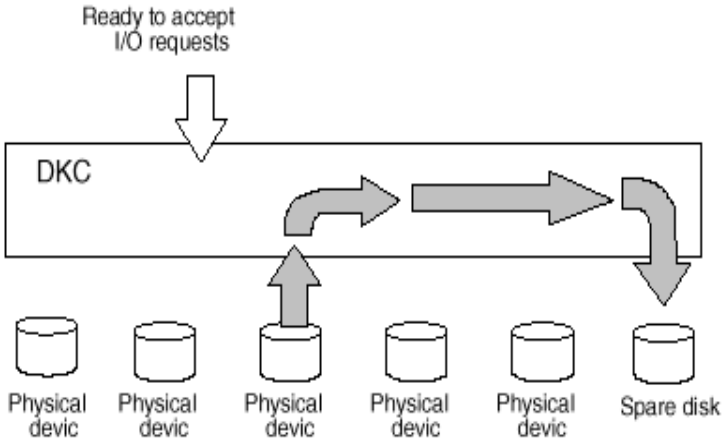


xp1024/128 – *Array Groups*

- XP Normal Volumes come in a predetermined “Emulation” sizes.
  - Normal Volume Emulation Types: OPEN-3/8/9/E/L/V
    - OPEN-3 = 2.3GB
    - OPEN-8 = 7.3GB
    - OPEN-9 = 7.4GB
    - OPEN-E = 14.5GB
    - OPEN-L = 36.5GB
    - OPEN-V = Variable sized volumes. Volumes can be created from 45M to 2TB

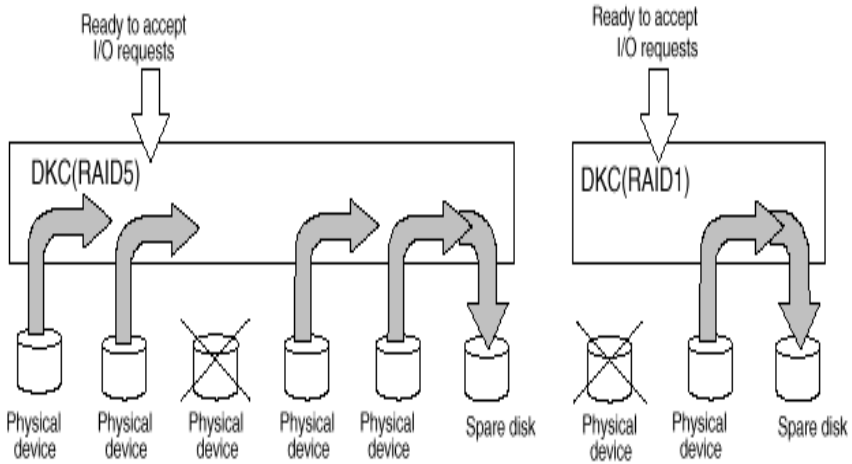


xp1024/128 – *Dynamic Spare Disk*



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



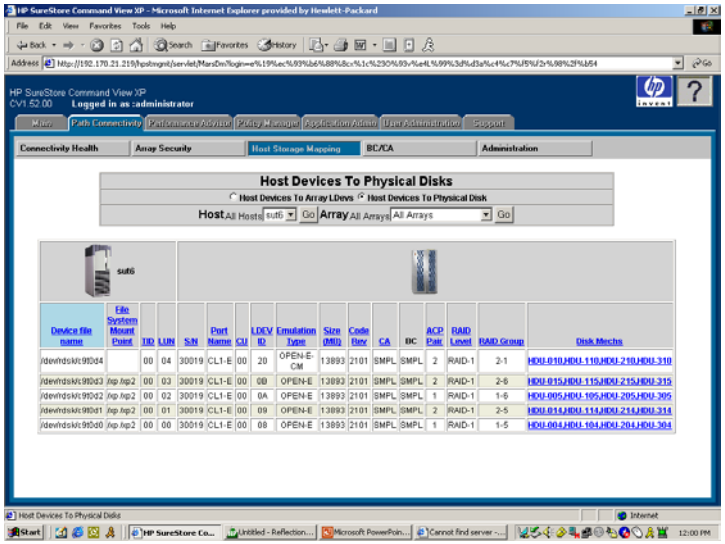
XP-1024/128

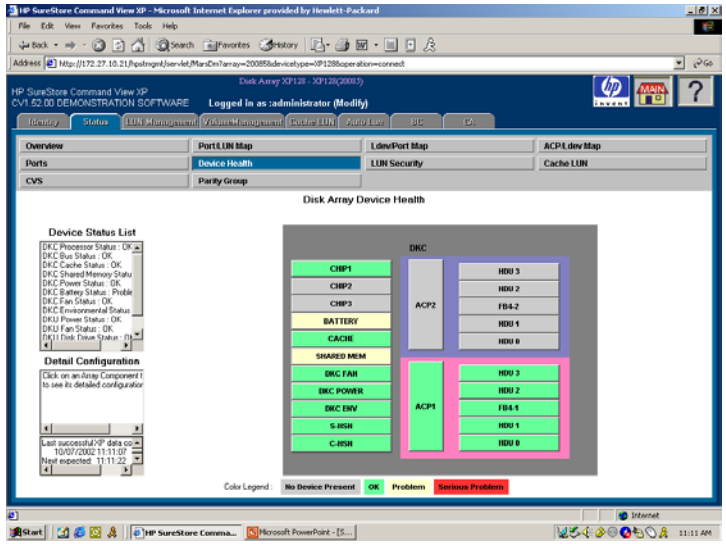
- XP-1024/128 software and OS support

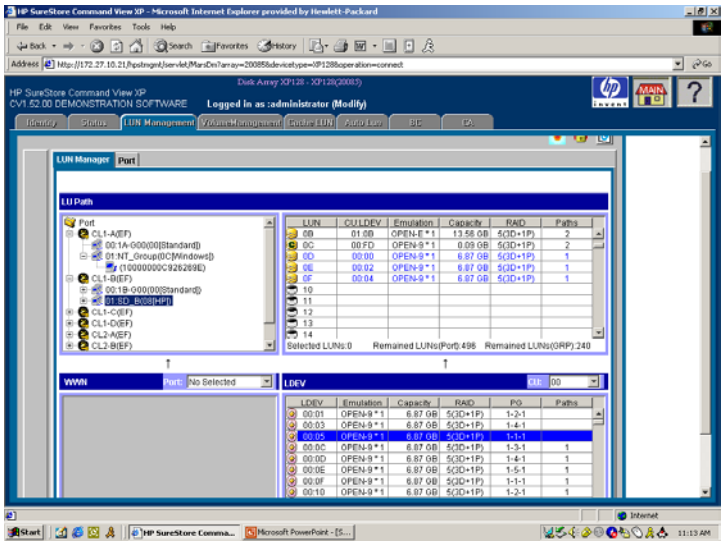
XP-1024/128 Software

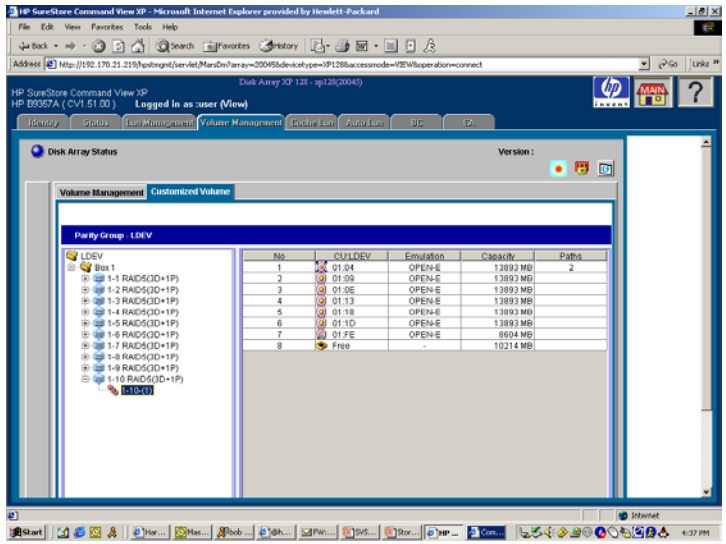
- Command View XP
- Performance Advisor XP
- AutoPath/Secure Path
- Business Copy XP
- Continuous Access XP
- Application Policy Manager
- Cache LUN XP

Command View

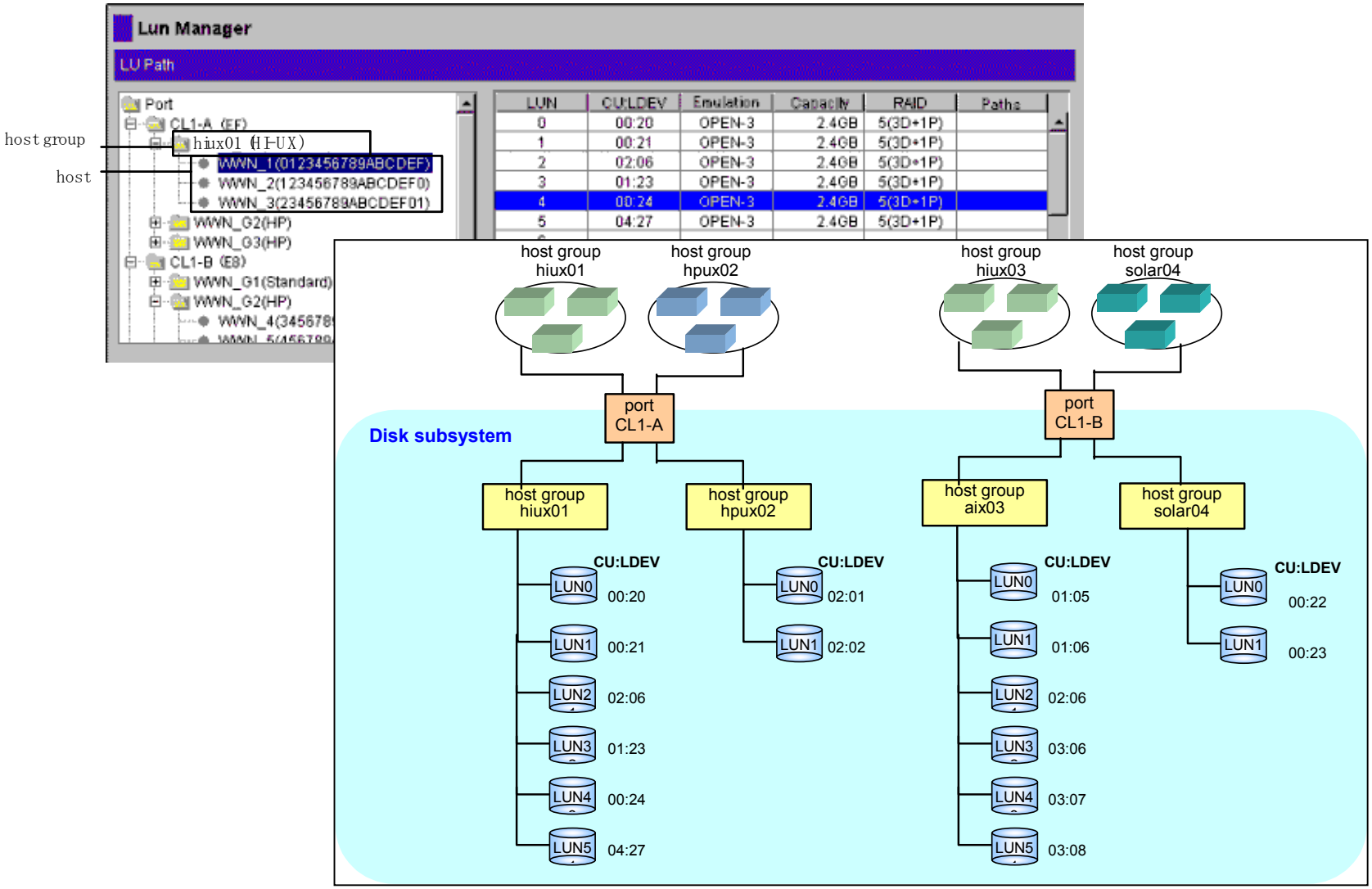








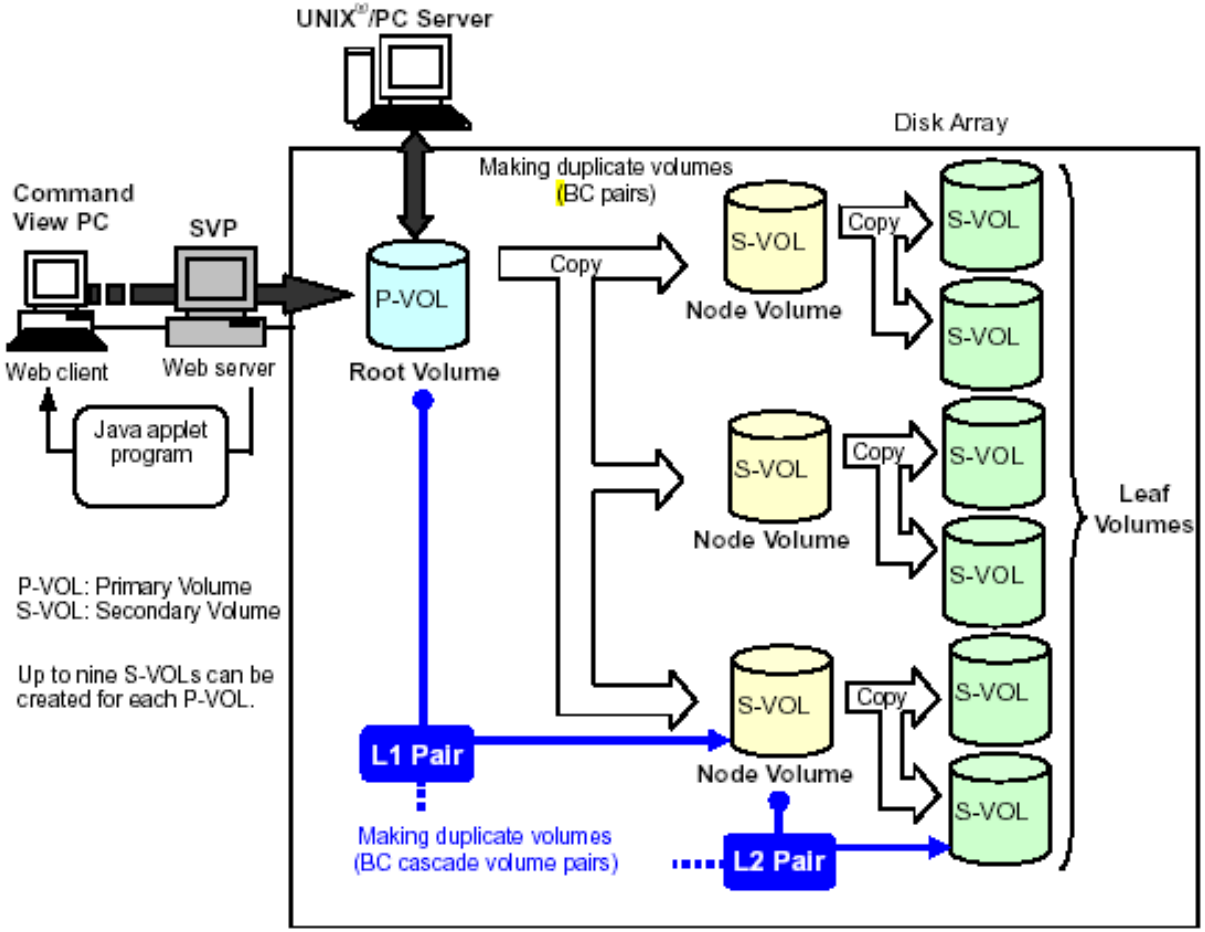
Command View/Host groups



xp1024/128 – *Business Copy XP*

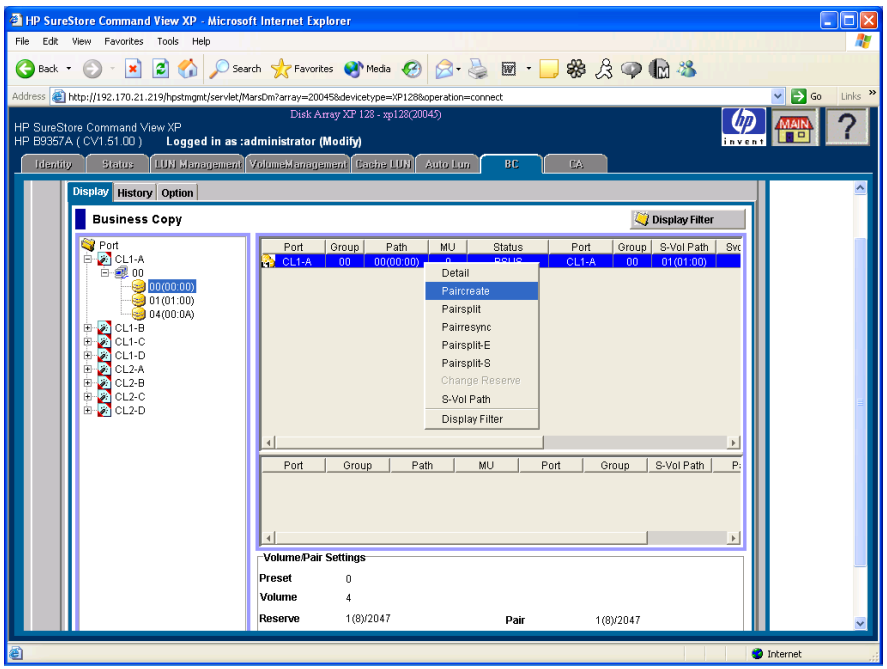
HP WORLD 2003
Solutions and Technology Conference & Expo

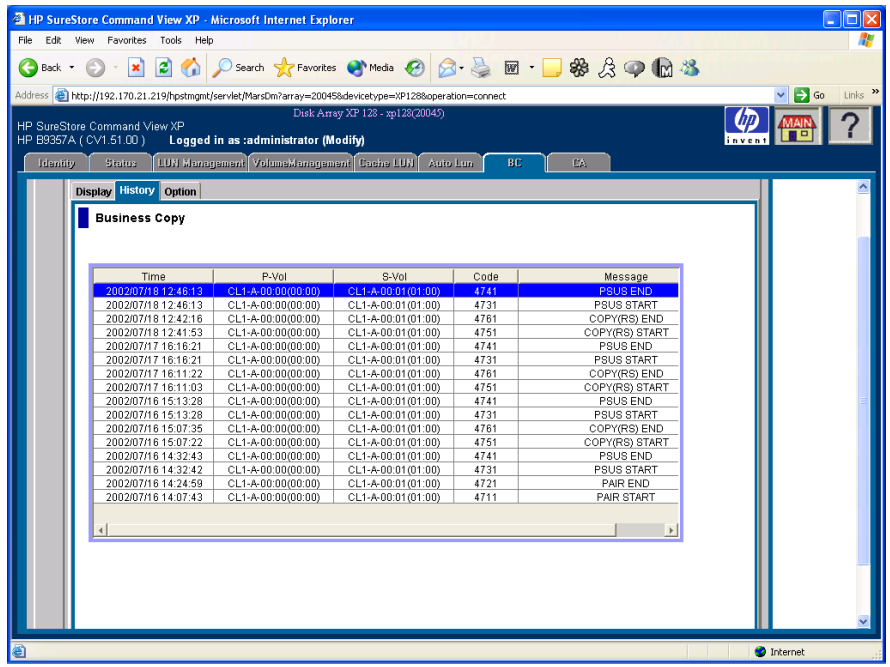
- Business Copy XP
  - local copies of logical units within the same disk array



Business Copy GUI

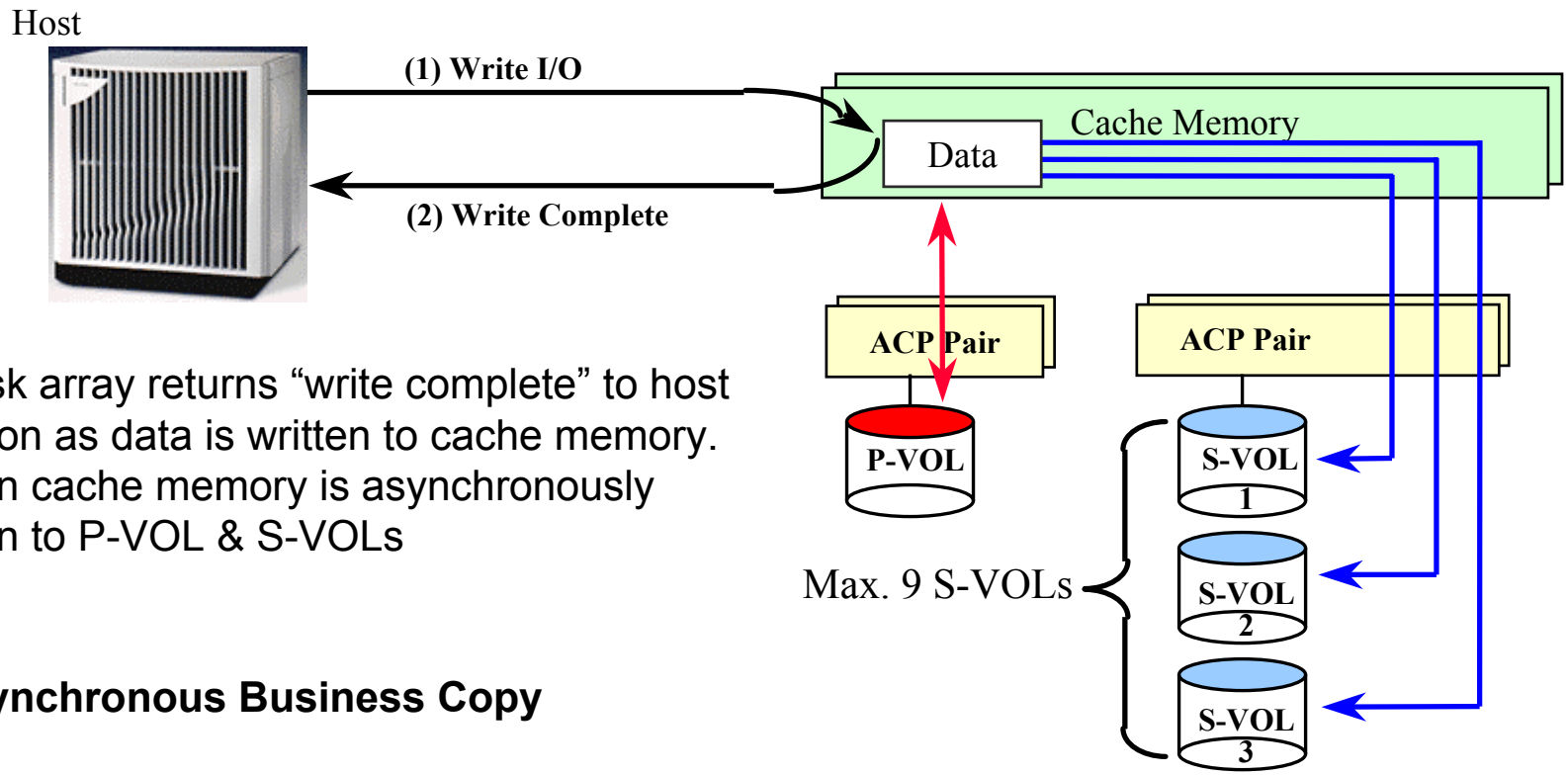
- Business Copy XP through CV
- Historical data is maintained by CV





BC Asynchronous Write Detail

Asynchronous Write Access to Secondary volumes maintains performance to Primary volumes.

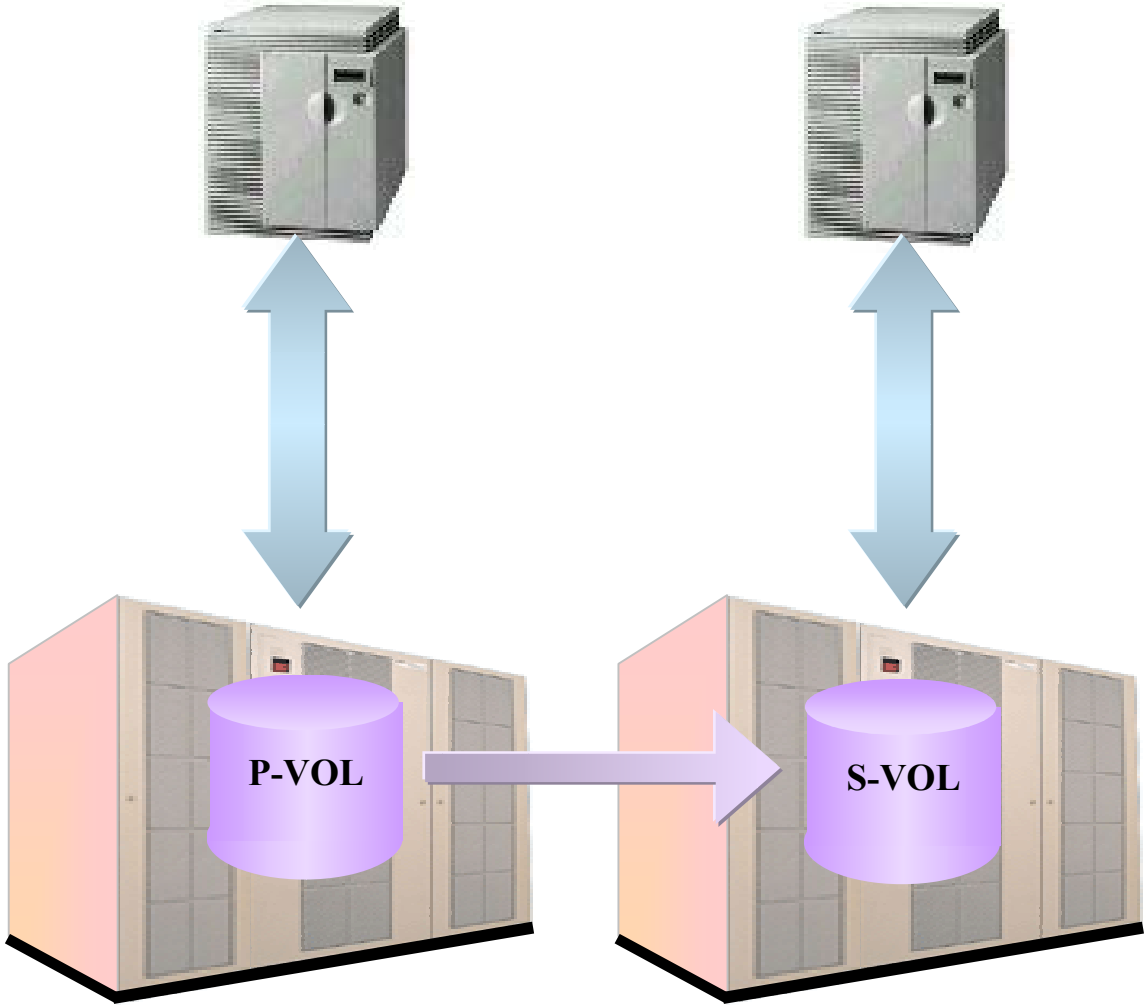


- XP disk array returns “write complete” to host as soon as data is written to cache memory.
- Data in cache memory is asynchronously written to P-VOL & S-VOLs

Asynchronous Business Copy

xp1024/128 – *Continuous Access XP*

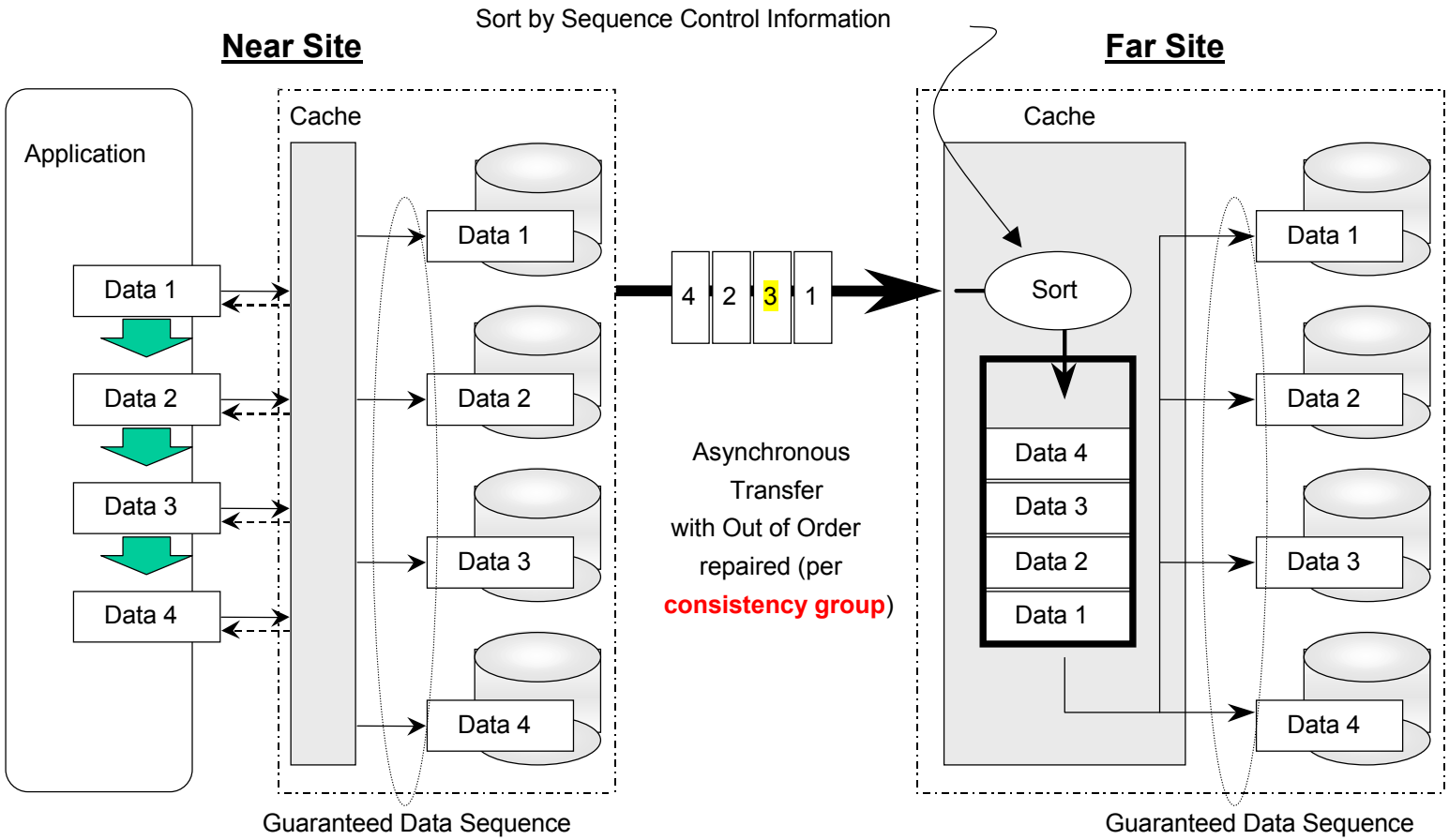
- "Synchronous" Continuous Access



xp1024/128 – *Async-CA Guaranteed Data Sequence*

HP WORLD 2003
Solutions and Technology Conference & Expo

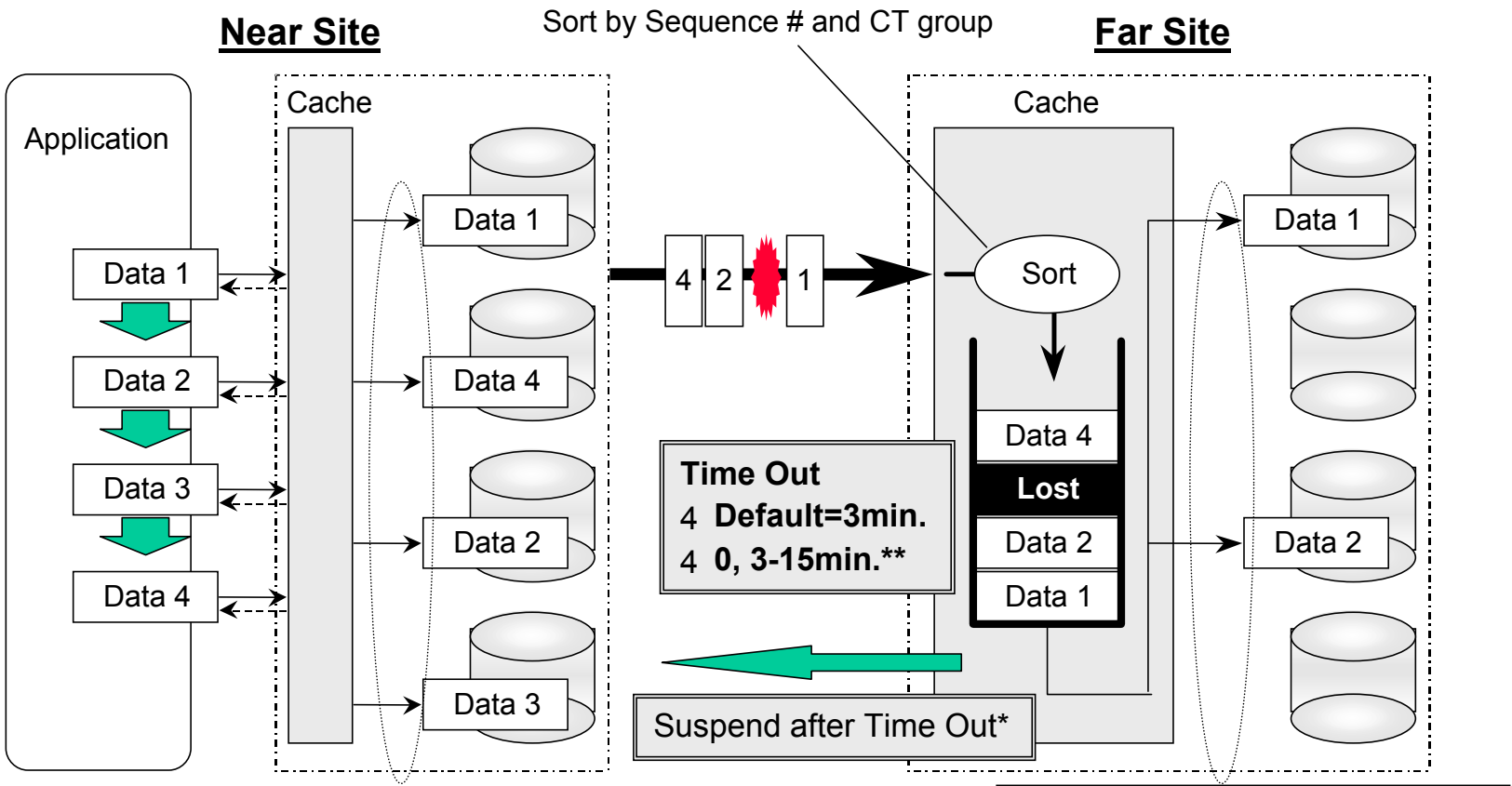
Asynchronous-CA Control



At far site, I/O #3 will be held until I/O #2 shows up

xp1024/128 – *Async-CA Lost Data in Flight*

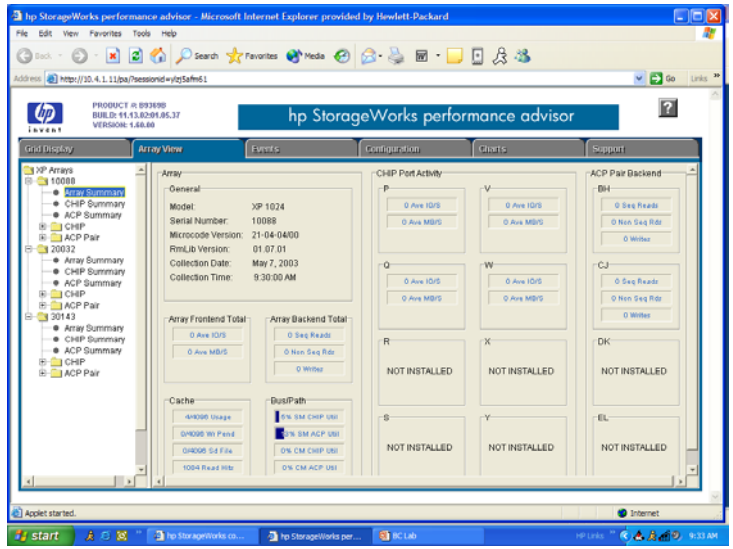
What of Lost Data in Flight?

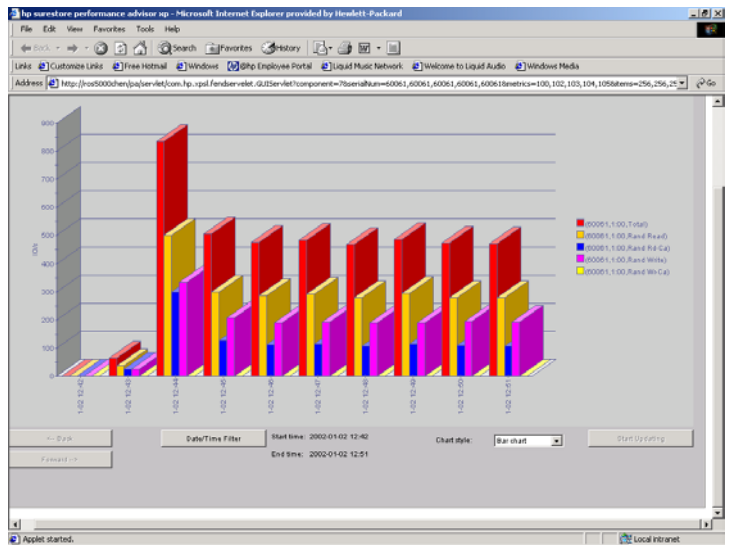


* All pending correct Record Set is completed. ** Under Consideration

Performance Advisor

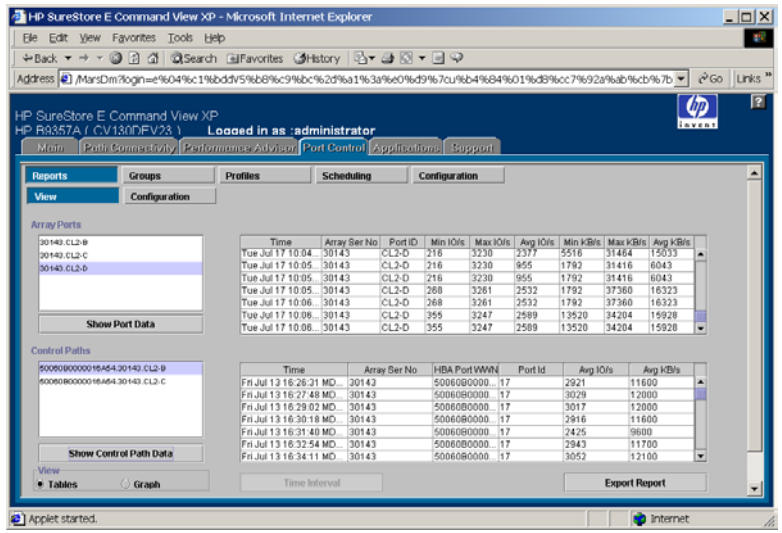
- Real time and historical performance reporting tool
- Can create charts to monitor resources based on
  - Hosts
  - Groups
  - Devices
  - And many more
- GUI and CLI
- Historical data stored in a database.

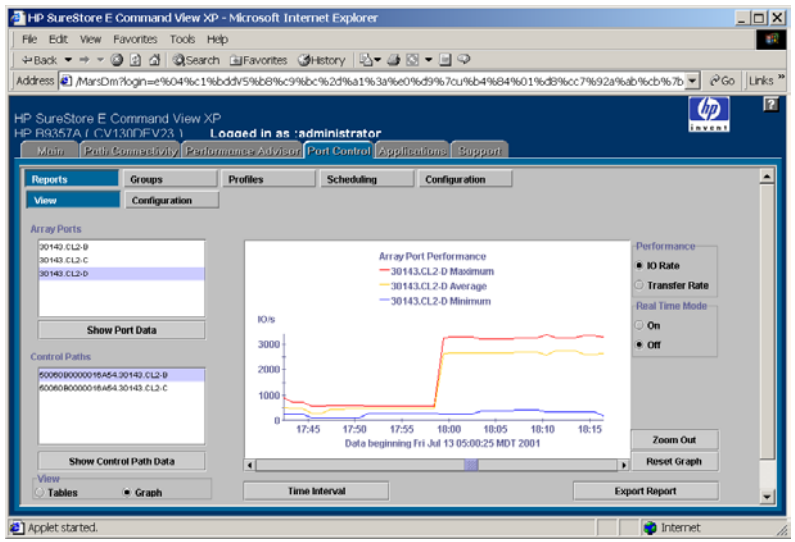




Application Policy Manager

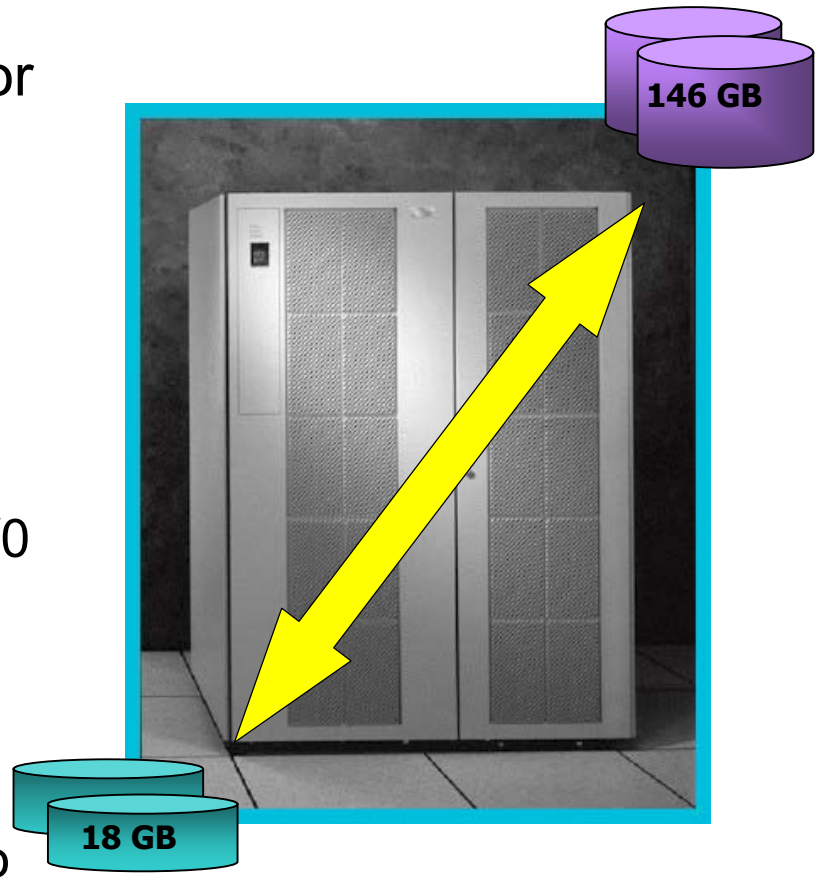
- Enables users to better manage their storage environment by prioritizing and allocating available array performance
- Control the performance available to each HBA
- Prevent lower priority hosts from impacting high priority host
  - limited the maximum performance for lower priority hosts
- User defined profile to group hosts for easy management



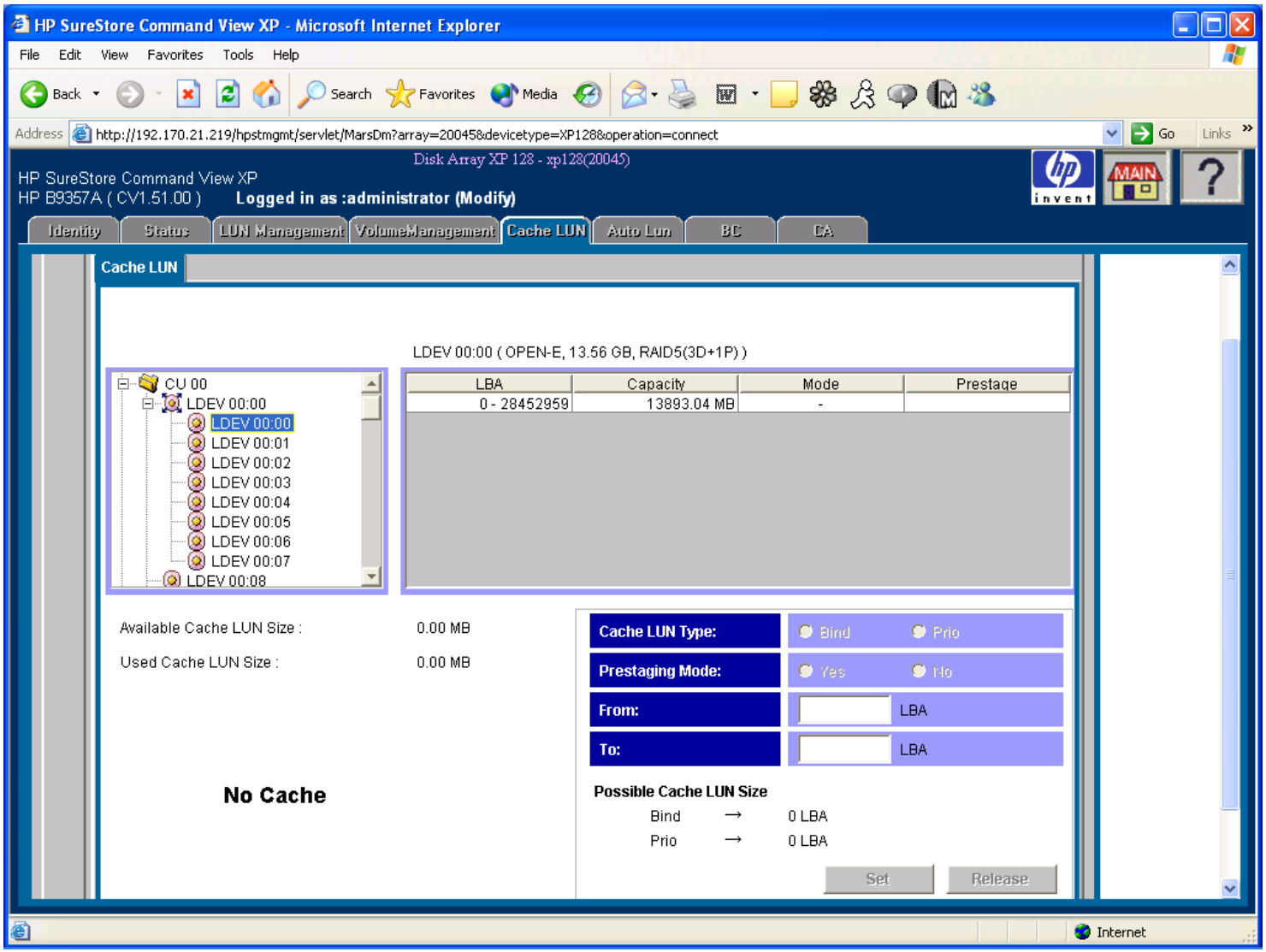


Auto LUN XP

- Allows the Array to suggest or control the appropriate type of storage space for your data.
- Has several modes of operation
  - recommend
  - automatic
- Can balance over several variables
  - Allows you to migrate between RAID 1/0 RAID5 devices.
  - Also allows you to migrate between different spindle types.
  - All changes are made without changing device files. This is done transparently to the connected servers



Cache LUN XP


HP SureStore Command View XP - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://192.170.21.219/hpstmgmt/servlet/MarsDm?array=20045&devicetype=XP128&operation=connect
Go Links
Disk Array XP 128 - xp128(20045)
HP SureStore Command View XP
HP B9357A (CV1.51.00) Logged in as :administrator (Modify)
invent
MAIN
Identity Status LUN Management VolumeManagement Cache LUN Auto Lun BC CA
Cache LUN
LDEV 00:00 (OPEN-E, 13.56 GB, RAID5(3D+1P))
CU 00
LDEV 00:00
LDEV 00:00
LDEV 00:01
LDEV 00:02
LDEV 00:03
LDEV 00:04
LDEV 00:05
LDEV 00:06
LDEV 00:07
LDEV 00:08
LBA Capacity Mode Prestage
0 - 28452959 13893.04 MB -
Available Cache LUN Size : 0.00 MB
Used Cache LUN Size : 0.00 MB
No Cache
Cache LUN Type: Bind Prio
Prestaging Mode: Yes No
From: LBA
To: LBA
Possible Cache LUN Size
Bind → 0 LBA
Prio → 0 LBA
Set Release
Internet

EVA and XP Configuration limitations

- EVA 3000 Maximum operating parameters
  - Up to 56 disks
  - Up to seven disk groups
  - Up to 512 Virtual disks
  - Up to 8192 host presentations
  - 256 LUNs/host HBA
  - 1024 FC adapter (WWN) connections per array
  - Up to seven snapshots per virtual disk
  - V-RAID 0 1 or 5

- XP128 Operating parameters
  - 4 to 128 disks
  - Up to 8192 ldevs
  - Up to 8192 LUNs
  - 4 to 24 CHIP ports
  - 1 or 2 ACP pairs
  - RAID1/0 2D+2D or 4D+4D
  - RAID5 3D+1P or 7D+1P (with 2 ACP pair)
  - Up to 64 GB Cache

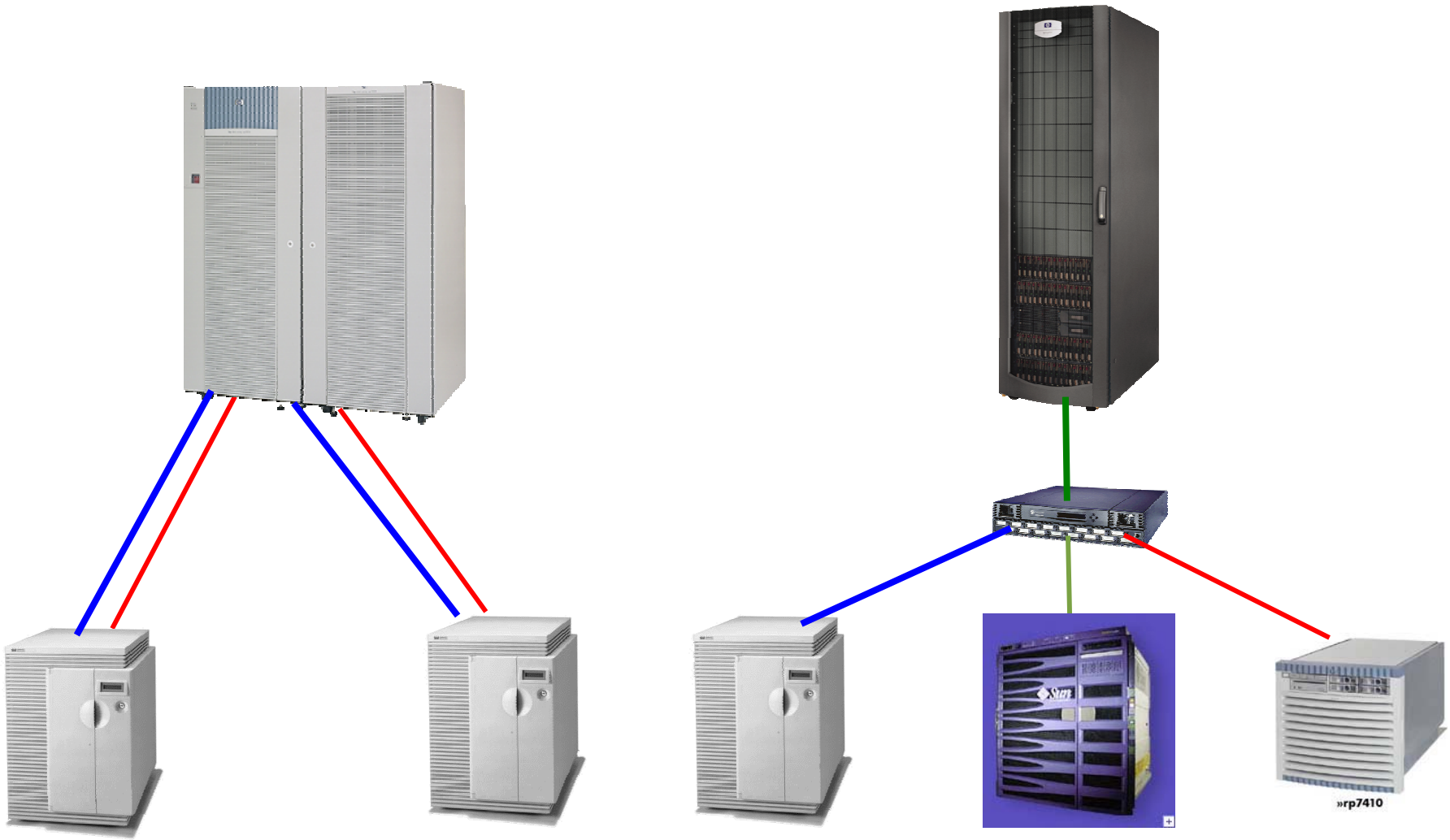
- EVA 5000 Maximum operating parameters
  - Up to 240 disks
  - Up to 16 disk groups
  - Up to 512 Virtual disks
  - Up to 8192 host presentations
  - 256 LUNs/host HBA
  - 1024 FC adapter (WWN) connections per array
  - Up to seven snapshots per virtual disk
  - V-RAID 0 1 or 5

- XP1024 Operating parameters
  - 4 to 1024 disks
  - Up to 8192 ldevs
  - Up to 8192 LUNs
  - 4 to 64 CHIP ports
  - 1 to 4 ACP pairs
  - Up to 4 Disk cabinets on one controller cabinet
  - RAID1/0 2D+2D or 4D+4D
  - RAID5 3D+1P or 7D+1P (with 2 ACP pair min.)
  - Up to 128 GB Cache

EVA and XP Comparison

- EVA and XP performance
  - EVA, High throughput direct to disk
    - Applications that utilize server buffer cache.
  - XP High cache hit rates (50 to 90% read hits typical)
    - Applications that utilize 20% of the data 80% of the time
- EVA and XP management
  - EVA, Easy quick LUN and RAID level changes, automatic load balancing.
    - Test and development environments
    - Highly dynamic production environments
  - XP slightly more management overhead.
    - More tools for tuning. I.E. Auto LUN. APM.
- EVA and XP scalability
  - EVA, scales by adding controllers and disk cabinets
  - XP scales by adding disks and cabinets to existing frame

EVA and XP Case Studies



EVA and XP Case Study 1

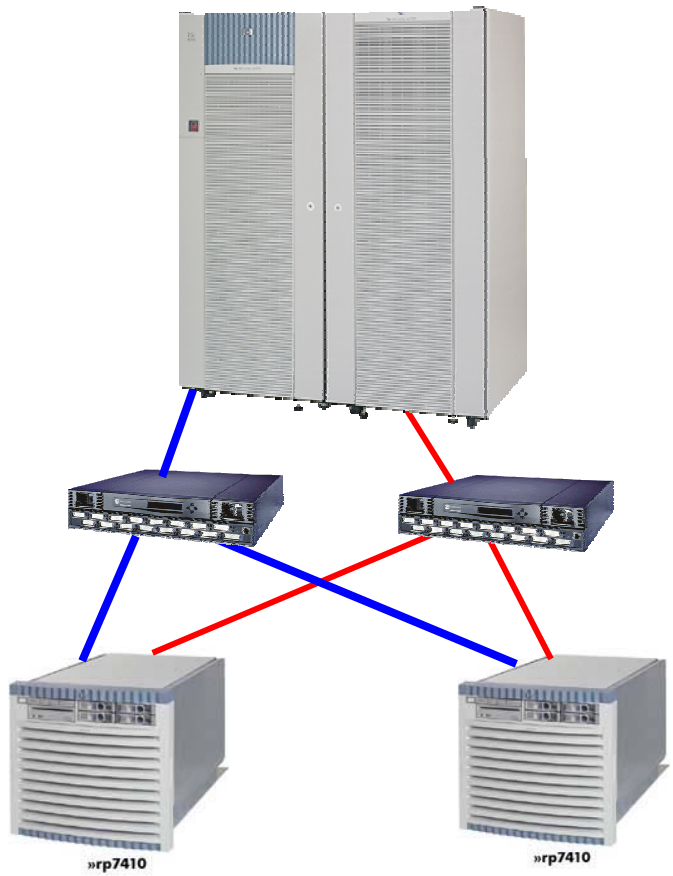
GE Real Estate EVA from XP migration

- IT problem
  - Customer has a growing development and test environment
  - IT budget is stressed
  - XP-256 currently used in Dev/Test is expensive to expand, and costly to maintain.
  - Point in time copies require disk space for every copy.
  - Constant changes to the Dev/Test environment are difficult and time consuming.
- HP solution
  - Trade in the XP-256
  - Implement an EVA 3000 on a McData SAN
  - Command View EVA ease of use allows administrators to change RAID levels, LUN sizes, number of disks etc, easily.
  - EVA Vsnap allows space efficient point in time copies.

Case Study 2

XP Benchmarking

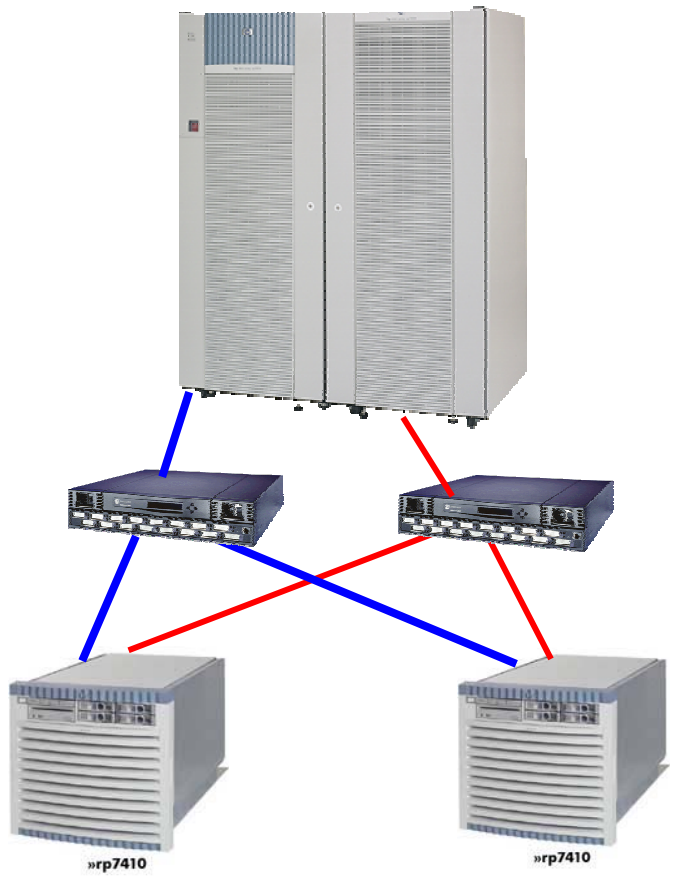
- Connected 2 servers to an XP-1024
  - Configured 4 RAID 5 RAID Groups for each server
  - Performed random r/w I/O tests
  - Incorrect conclusion Max. IOPS=2,500 per server.
- Reconfigured so that Applications shared the 4 RAID Groups
  - Application performance suffered
  - Incorrect conclusion, XP cannot share RAID Groups between applications.



Case Study 2

XP Benchmarking

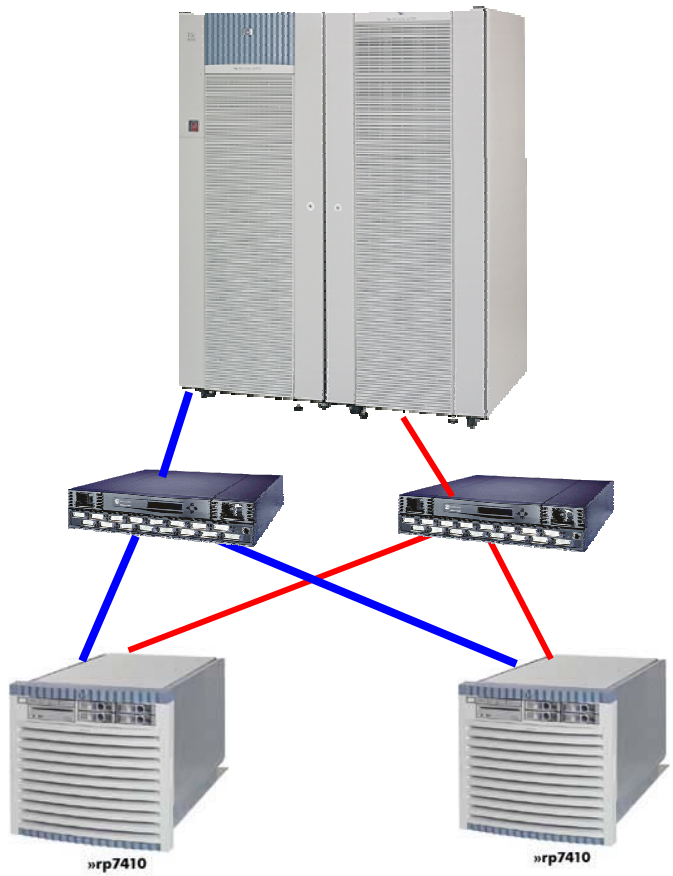
- Customer connected 2 servers to an XP-1024
  - Configured 16 RAID 5 RAID Groups for both servers
  - Configured stripe sets or 4 devices per group at 64k stripe size.
  - Performed random r/w I/O tests
  - Correct conclusion Max. IOPS=15,000+ per server.
    - (bottleneck is now CPU not disk)
  - Application performance does not suffer from sharing disks



Case Study 2

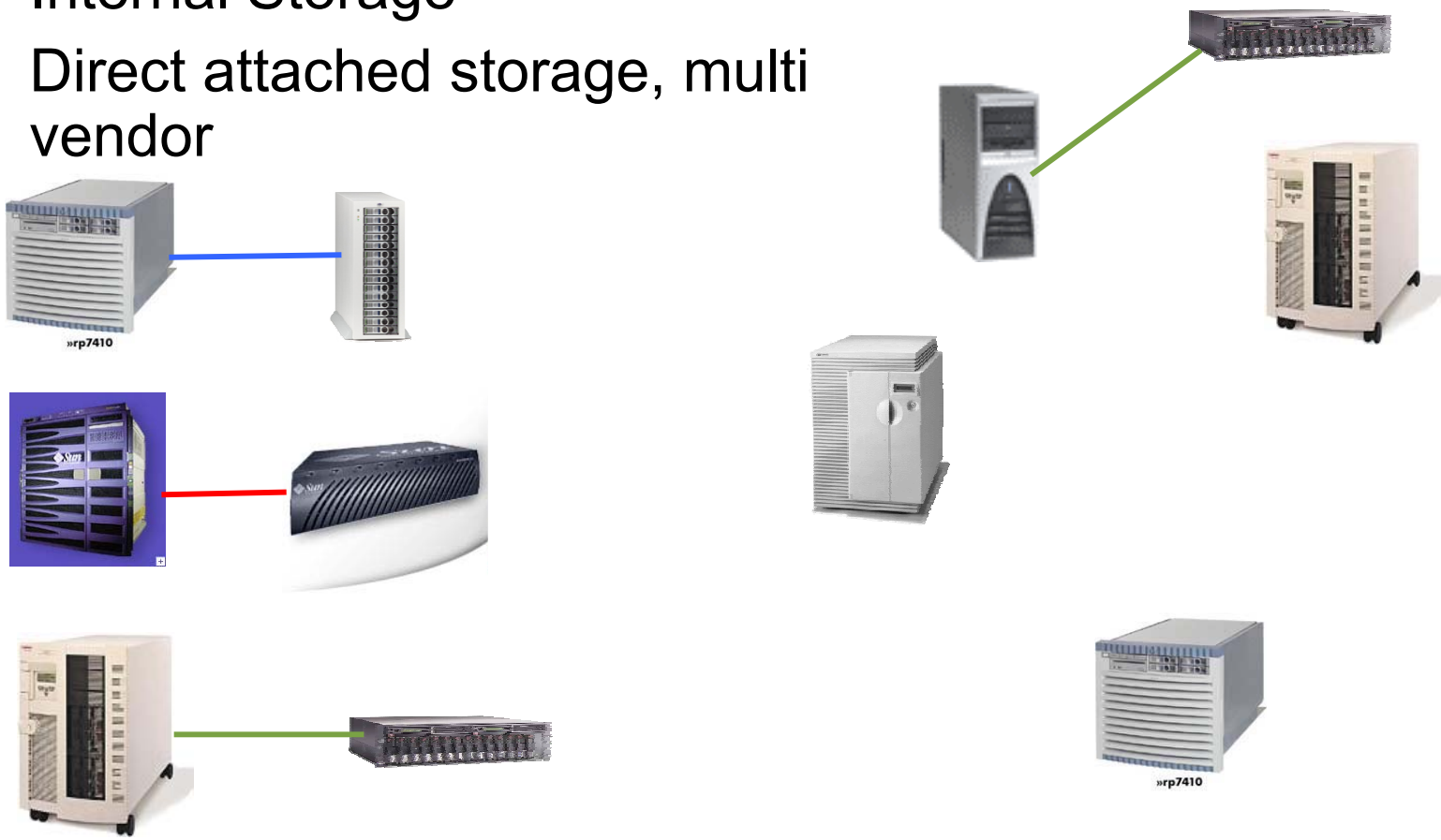
XP Benchmarking

- Conclusion. Configure the XP array “WIDE” meaning lots of RAID groups.
- The XP completes all I/O in cache. The goal is to flush cache efficiently
- Striping allows more efficient UNIX performance through more SCSI queue distribution.
- Applications do not suffer from sharing resources on the XP array.



Case Study 3

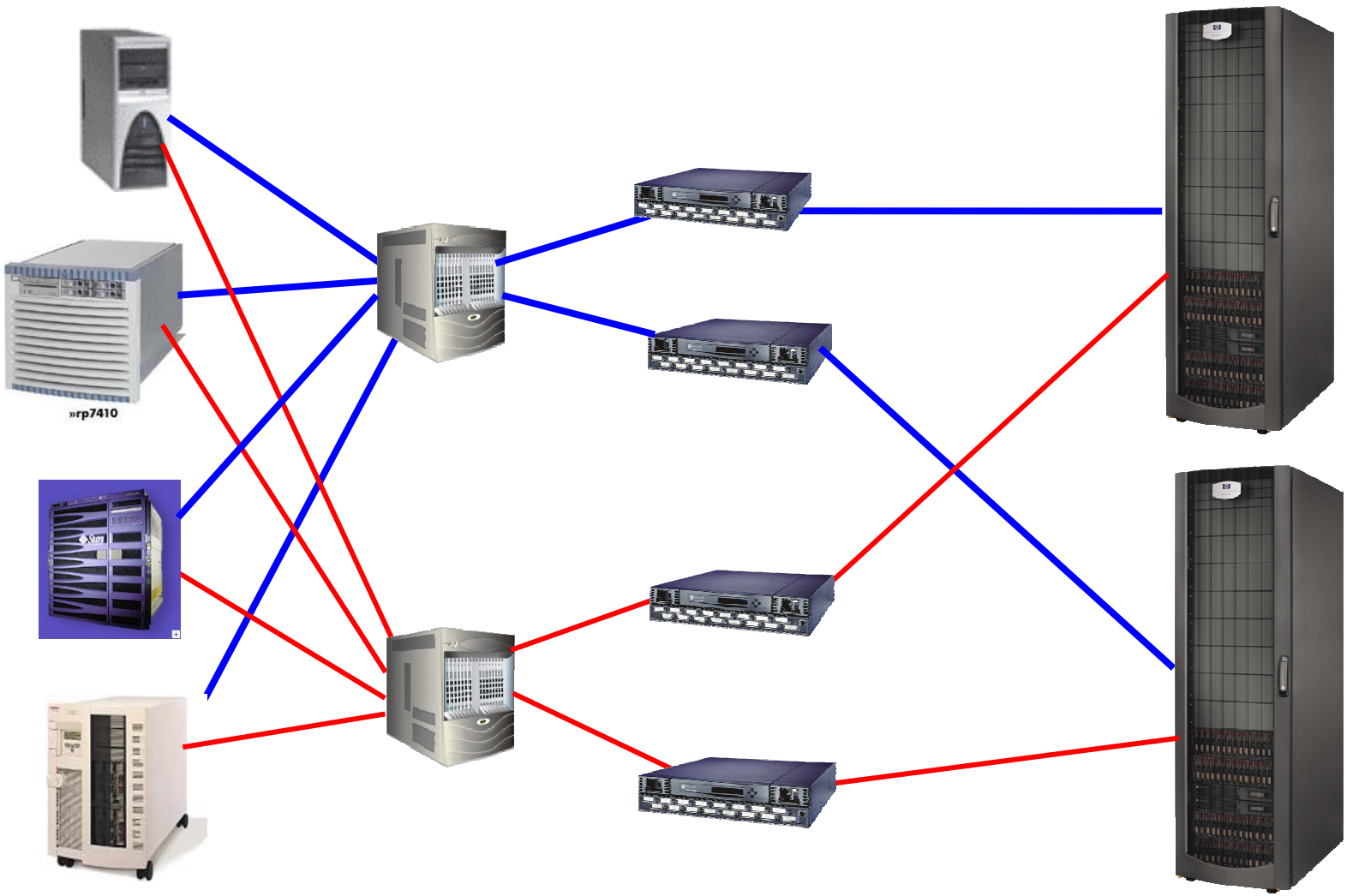
- Hackensack Medical Center
  - Internal Storage
  - Direct attached storage, multi vendor



Case study 3

- IT problem
  - Hackensack Medical Center had 300+ multi vendor servers with direct attached storage
  - Doctors could only pull up part of a patients medical records during a data retrieval.
    - I.E. Written records but not films (they are “filmless”)
  - S.A’s were overwhelmed in maintaining data storage.
  - Films need to be stored on-line for one year, and archived for another 21 years.
  - DR may have been an issue
- HP Solution
  - Create a SAN infrastructure using Brocade FC switches
  - Create a shared storage pool using EVA arrays.
  - S.A’s are able to simply request storage from the pool and the appliance/EVA find the storage.
  - The EVA virtualization and load leveling allows simplicity of administration without performance impact.
  - Film retrieval performance was decreased from 1-2.5 minutes, to 2-10 seconds.
  - Futures will include using DRM for remote replication.

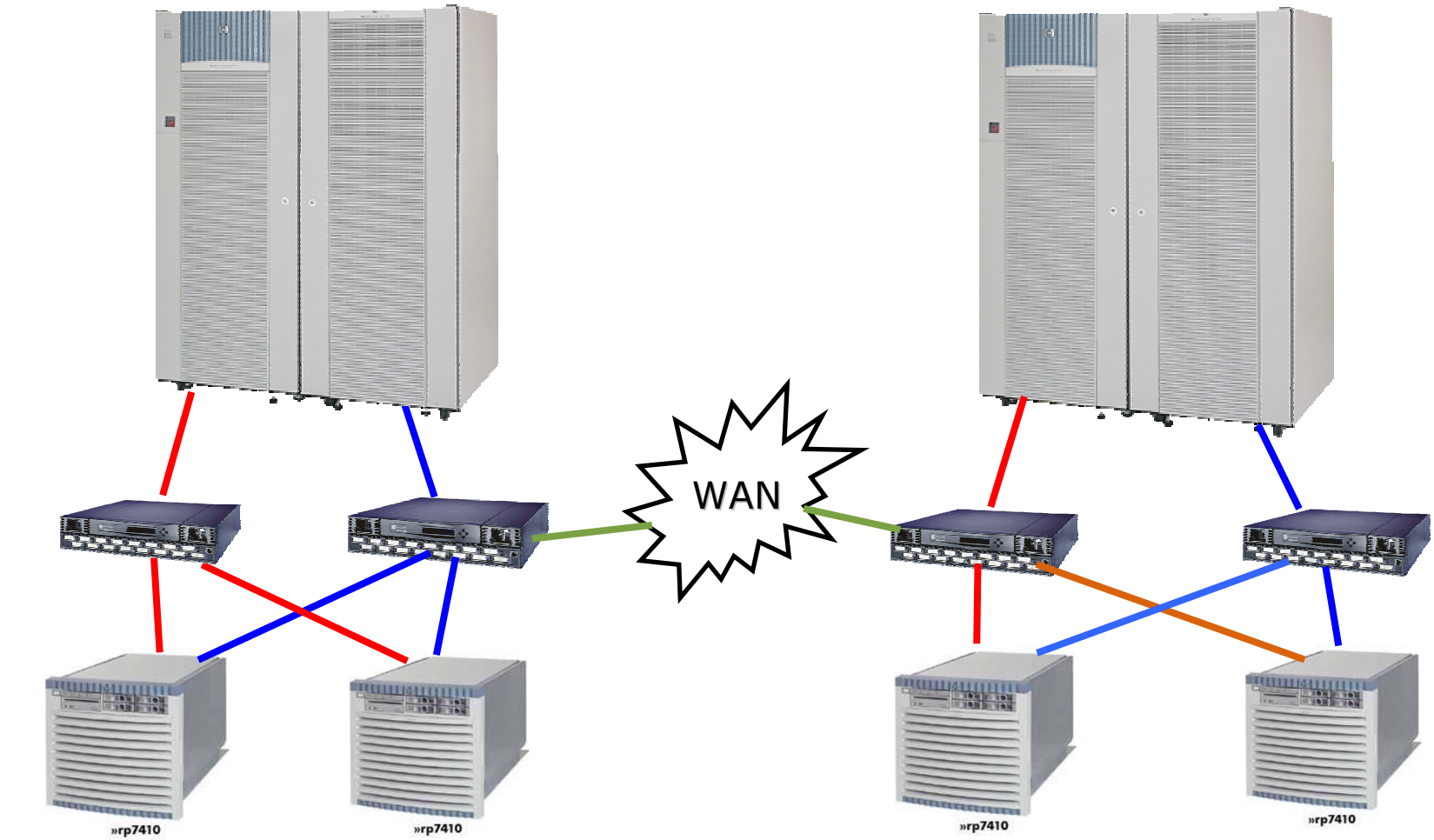
Case Study 3 EVA solution



Case Study 4

- FedEx Oracle 9iRAC testing using XP-1024
  - Customer requirement for 2000TPS
  - Expected IOPS 12,000
  - DR using CA async over distance simulator
- Volumes using HP-UX LVM striping were created for
  - Rac1-8
  - Archive logs
- Max throughput achieved
  - 1100 TPS
  - 8600 IOPS

Case study 4
FedEx Benchmark layout



Oracle 9i rac cluster

Oracle 9i rac DR cluster

Case study 4 Performance report

- Performance advisor database was extracted
  - All I/O was shown to be randomized
  - Customer's expectation was that "rac" would be random however "archive logs" would be sequential
- Recreated "archive logs" volumes as concatenated
- Total throughput of the application test went up by 112%. Service times on devices went down from 40ms average to less than 5ms average.
- Conclusion
  - Striping randomized the I/O to the sequential logs, creating high service times, lowering total throughput.
  - Performance advisor logs as well as DBA's knowledge of the application were instrumental in completing this benchmark successfully.

Case Study 5 EVA solution

IT Problem

- The need to combines more than 200 databases into a single, proprietary information system to manipulate the data for investment analysis.
- The sheer volume of this data – along with 30 percent growth annually and the addition of 40 or so new databases over the past year escalated the storage demands.
- The requirement is for relentless storage capacity, high-bandwidth performance, and the ability to scale its storage environment easily.

Solution

- Tested the Hitachi based array and the EVA5000 on the application.
- Built two storage-area networks (SANs), which include HP AlphaServer systems plus a Sun Solaris 6800-series server and two EVA 5000's.
- The Web-enabled applications require substantial server bandwidth, and the combination of high-powered AlphaServer systems and EVA arrays boosts performance considerably.

Case Study 5 EVA solution

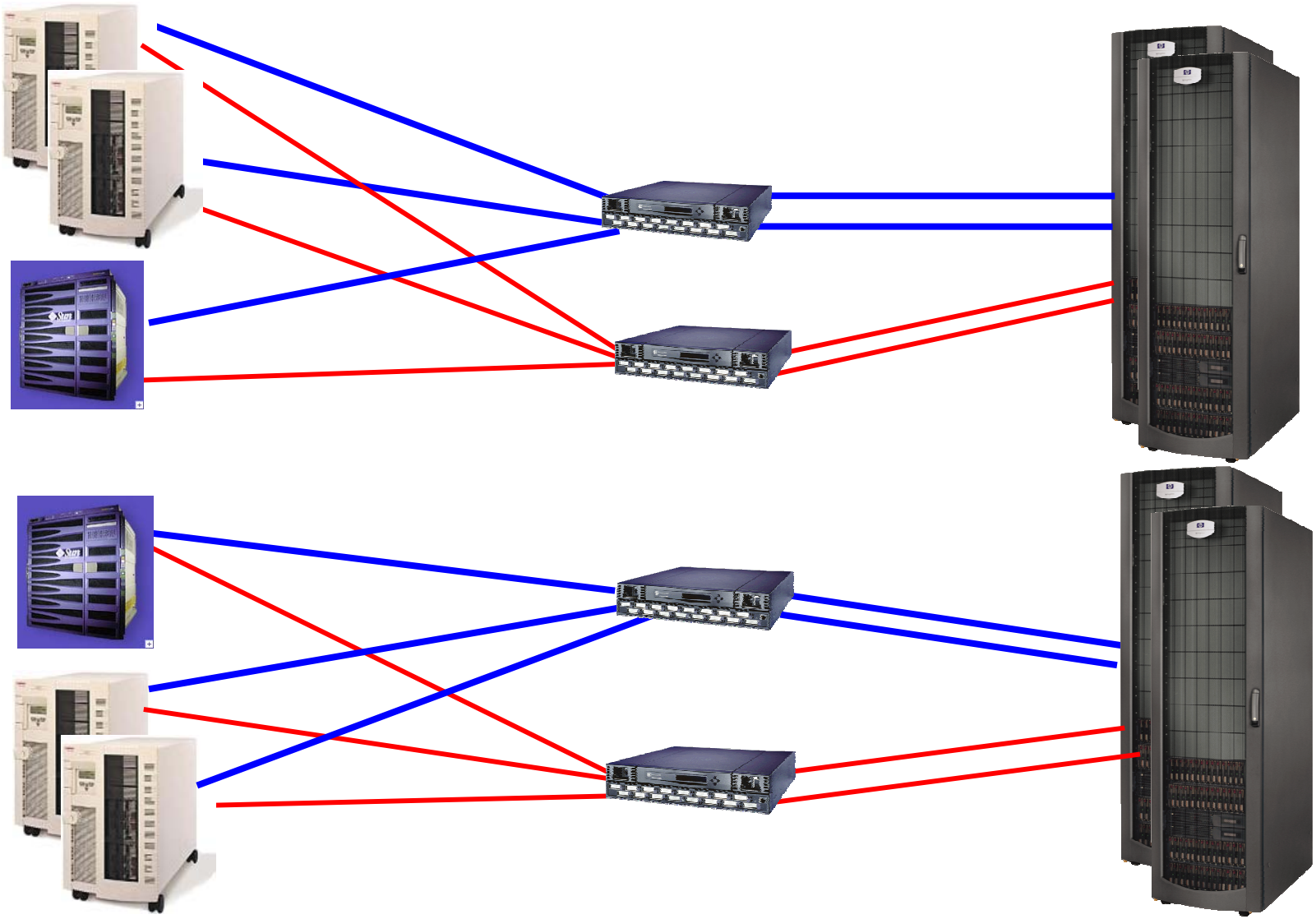
IT Problem.

- The current environment is HP storage solution consisting of computer interconnect-based (CI) storage with HSJ80 controllers and SCSI disks.
- As data grew, it became more difficult to reorganize and scale the storage environment.
- “We used to spend every weekend manually rearranging the disk arrays to maximize storage space,” says Systems Engineering Director.

Solution

- “The high throughput and data-transfer rates of the HP StorageWorks Enterprise Virtual Array 5000 system allow us to harness more power from our AlphaServer systems,” explains the SE Director.
- “The HP StorageWorks solution reduced latency of our read/write operations as much as twenty-fold.” (OVMS team director)

Case Study 5 EVA solution



Interex, Encompass and HP bring you a powerful new HP World.