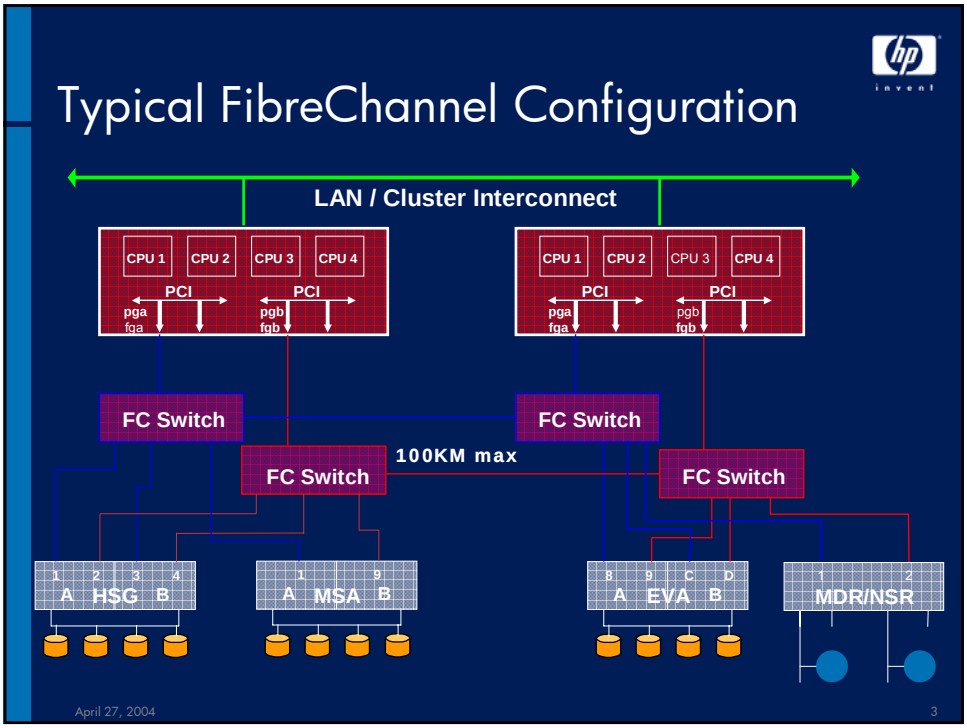
3D09 OpenVMS SAN/Storage Technical Directions

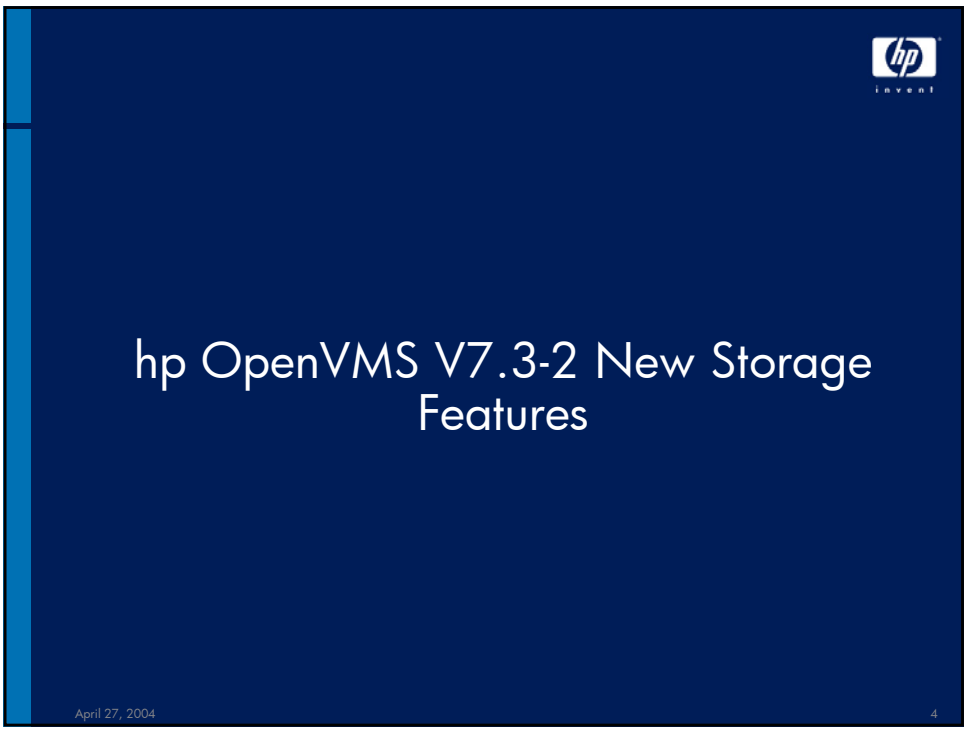
Manfred Kaser
HP Services

hp OpenVMS Storage Technology

- V7.3-2 New Storage Features
- Storage Arrays
- Tape Products
- SAN Infrastructure
- Post V7.3-2 Projects
- Itanium® Based Systems IO Plans
- Longer Term Storage Interconnects

April 27, 2004 2


Typical FibreChannel Configuration
LAN / Cluster Interconnect
CPU 1
CPU 2
CPU 3
CPU 4
PCI
pga
fga
pgb
fgb
FC Switch
100KM max
A HSG B
A MSA B
A EVA B
MDR/NSR
April 27, 2004
3


hp OpenVMS V7.3-2 New Storage Features
April 27, 2004
4

hp OpenVMS V7.3-2 Storage Features

- *Dynamic Volume Expansion*
- *Shadowing of dis-similar volumes*
- *Mount Verification Filtering*
- *Smart Array Fast Path*
- *Fabric Management Enhancements*

April 27, 2004 5

V7.3-2

Dynamic Volume Expansion (DVE)

- Storage arrays today can dynamically expand volumes
- OpenVMS could not utilize the expanded volume without re-initializing
- OpenVMS dynamic volume expansion allows allocation of extra bitmap space at init time and then later enables volume size expansion while the device is mounted
  - $ init /limit 1dga100: name *! Allocates 1TB bitmap*
  - $ mount/cluster 1dga100: name *! Mount and use volume*
  - *(expand physical volume size using management tools)*
  - $ set volume 1dga100: /size=xxxxxx *! Set new size*
- For volumes already initialized
  - $ dismount/cluster 1dga100: *! Dismount volume*
  - $ mount 1dga100: name *! Privately mount volume*
  - $ set volume 1dga100: /limit *! Allocate new bitmap*
  - $ mount/cluster 1dga100: name *! Cluster mount volume*
  - *(expand physical volume size using management tools)*
  - $ set volume 1dga100: /size=xxxxxx *! Set new size*

April 27, 2004 6

V7.3-2

Dissimilar Device Shadowing (DDS)

- Prior to V7.3-2, shadowed volumes had to be identical in size
  - This made it impractical to shadow between different types of storage arrays
- Different sized volumes can now be shadowed
  - The size of the smallest volume determines the size of the shadow set
  - The shadow set can only be mounted on V7.3-2 nodes
  - The size of the shadow set can be grown using DVE

April 27, 2004 7

V7.3-2

Quieter Mount Verification

- In a SAN there are many reasons for "normal" mount verifications
  - Path switch by another cluster node
  - Dropped FC packets (not a norm, but it does happen)
  - Transient command timeout to a busy storage device
- These result in mount verification messages that alarm system managers
- "Quiet" mount verification will allow infrequent, immediately recovered mount verifications to be suppressed from operator logs
- Sysgen parameters (setting either of these to 0 results in previous behaviour)
  - Mvsupmsg_intvl *(default 1 hour)*
  - Mvsupmsg_num *(default 5)*
- Mount verification messages are suppressed unless that are more than "Mvsupmsg_num" mount verifications within "Mvsupmsg_intvl" seconds to a single device

April 27, 2004 8

V7.3-2

Smart Array 5300 Fastpath

- Fastpath is now available for the Smart Array 5300
  - Sysgen parameter fast_path_ports bits *(set to zero to enable fastpath)*
  - Fastpath enabled for all devices by default
    - 0 Controls Fast Path for PKQDRIVER (for parallel SCSI).
    - 1 Controls Fast Path for FGEDRIVER (for Fibre Channel).
    - 2 Controls Fast Path for PKADRIVER (for Ultra3 SCSI).
    - 3 Controls Fast Path for PEDRIVER (for LAN).
    - 4 Controls Fast Path for PKRDRIVER (for SMART Array 5300).
    - $ set dev pkx /pre=x *! Changes preferred CPU*

April 27, 2004 9

V7.3-2

Fabric Management Enhancements

- Removed FC port reset on RSCN
  - Prior versions of VMS have reset the FC port (KGPSA adapter), on receipt of fabric RSCN (most common cause would be a re-zone of a switch)
  - This caused all I/O to be aborted and resulted in many device path switches (and mount verification messages)
- PLOGI retry
  - For any device on the SAN that fails PLOGI, VMS will now re-attempt PLOGI every 5 minutes
  - Allows configuration of some class of devices without a re-boot (HSG with full host table, or an XP array which had no VMS LUNs at VMS boot time)

April 27, 2004 10

V7.3-2

Fabric Management Enhancements

- Fabric name table scan
  - VMS now scans the fabric name table every 5 seconds
  - This allows device configuration of a device in the event that we loose the RSCN message associated with adding a device to the fabric
- New fibre_scan utility
  - $ mc sys$etc:fibre_scan
  - Scans the entire fabric for configured and non-configured devices
  - Used as a diagnostic tool for SAN configuration problems

April 27, 2004 11

Storage Arrays

April 27, 2004 12

Storage Products
EVA 5000

- 4 2gb host ports
- 4 2gb FC disk ports (2 redundant loops)
- 240 disks / 35TB max
- 1GB mirrored cache
- Multi-level snapshots
- Dynamic LUN expansion
- CA/DRM for OpenVMS
- SSSU host based control utility

April 27, 2004 13

Storage Products
EVA 3000

- 4 2gb host ports
- 2 2gb FC disk ports(1 redundant loop)
- 56 disks / 8TB max
- 1GB mirrored cache
- Multi-level snapshots
- Dynamic LUN expansion
- CA/DRM for OpenVMS
- SSSU host based control utility

April 27, 2004 14

Storage Products
EVA VCS 3.010

- VCS V3.010 is a mandatory upgrade
- Bug/serviceability fixes plus
  - Async CA mode
  - 128 copy sets
  - 128 DR groups
  - 100ms link limit

April 27, 2004 15

HP StorageWorks xp128

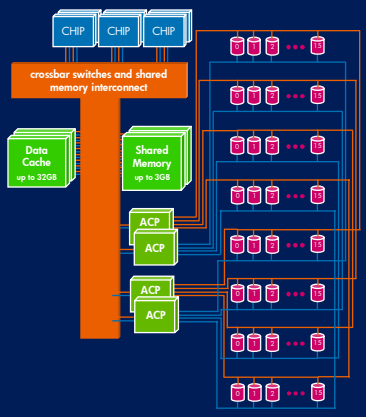
xp128	
Max Disk Drives	128
Max Capacity	18 TB raw 16 TB usable
Max Aggregated Crossbar Throughput	7.5 GB/s total 5 GB/s data 2.5 GB/s control
Max Sequential Data Transfer Rate	1.1 GB/s sustained 2.4 GB/s peak from cache
Max Random IOPS from Cache	375,000
Max LDEVs	8192

Enterprise Performance, Capacity, and Reliability in a Small Footprint

April 27, 2004 16

HP StorageWorks xp128

Start at the size you need, scale as your needs grow, staying in the same small footprint



April 27, 2004 17

HP StorageWorks xp1024

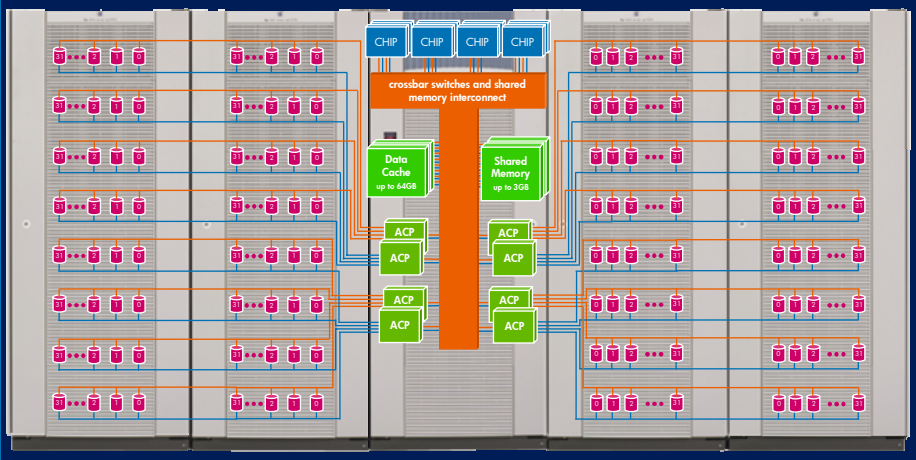
xp1024	
Max Disk Drives	1024
Max Capacity	149 TB raw 129 TB usable
Max Aggregated Crossbar Throughput	15 GB/s total 10 GB/s data 5 GB/s control
Max Sequential Data Transfer Rate	2 GB/s sustained 3.2 GB/s peak from cache
Max Random IOPS from Cache	500,000
Maximum LDEVs	8192

Unprecedented Performance, Scalability, and Reliability for the Enterprise

April 27, 2004 18

HP StorageWorks xp1024

Start at the size you need, scale to the highest usable capacity of any array on the planet



April 27, 2004 19

Storage Products

Modular Storage Array 1000

- 2gb FibreChannel front-end
- 4 U160 SCSI backend ports
- 4u rackmount with 14 drives
- 28 additional drives with 2 external storage shelves
- Works in existing SANs
- Low cost 2 node clusters
  with embedded 3 port FC-AL hub (V7.3-1+ only)
- Supported with
  - V7.2-2
  - V7.3
  - V7.3-1
  - V7.3-2

April 27, 2004 20

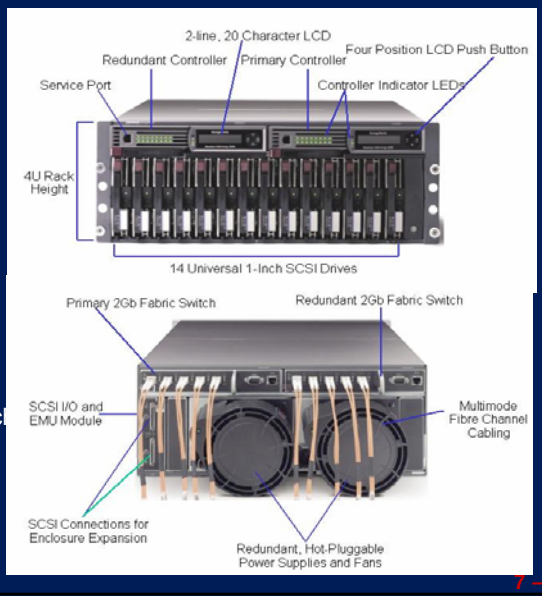
Storage Products

Modular Storage Array 1000

- Initial VMS / Tru64 support with V3.36 firmware
- Firmware must be field upgraded
  - OpenVMS V7.2-2 VMS722_FIBRE_SCSI-V0400
  - OpenVMS V7.3 VMS73_FIBRE_SCSI-V0500
  - OpenVMS V7.3-1 VMS731_MSA1000-V0100
- V4 do out any day now with windows/linux/vms/tru64 support
  - VMS will now work with an "out of the box" MSA
- FC-AL still in test

April 27, 2004 21

MSA1000 SAN Solution



FC Connection Options
Single 2Gb port for external switch
3 port 2Gb Hub
8 port 2Gb switch (Brocade)

April 27, 2004 7 – 47

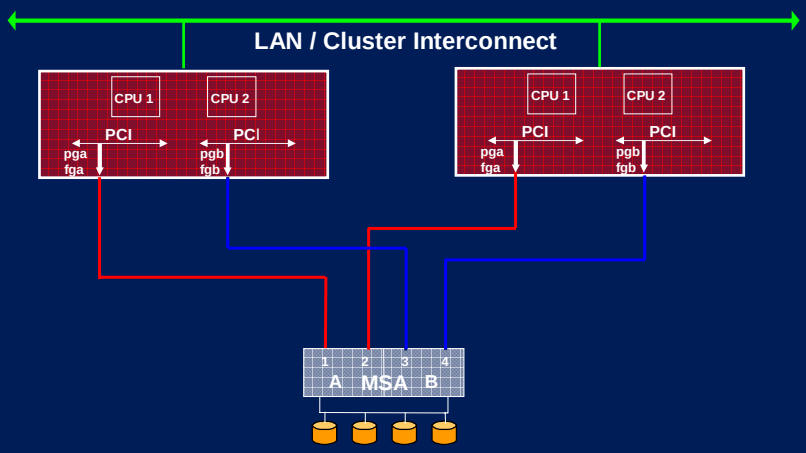
Storage Products

Modular Storage Array 1000

- High Performance Architecture
  - 200MB/sec throughput
  - 25,000 I/O per second
- Redundant controller support
  - Active/Standby in initial product
  - Active/Active in future
- RAID 0, 1, 0+1, 5 and ADG
- LUN Masking (SSP)
- 2Gb/1Gb auto-sense host ports
- Dual Cache Modules
  - Upgradeable to 512MB (per controll
- Serial line config and management

April 27, 2004 23

Low End FC-AL Based Cluster



April 27, 2004 24

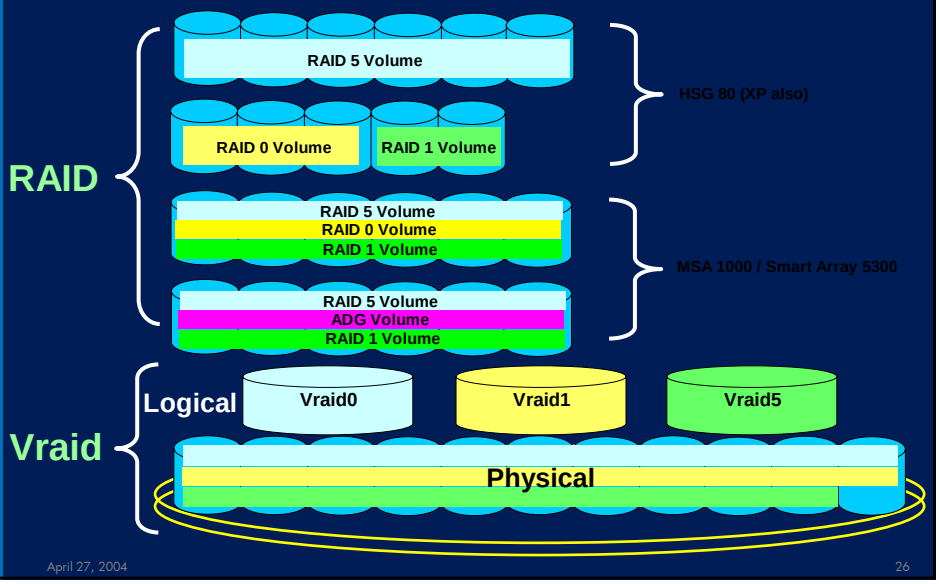
Storage Products

Smart Array 5300

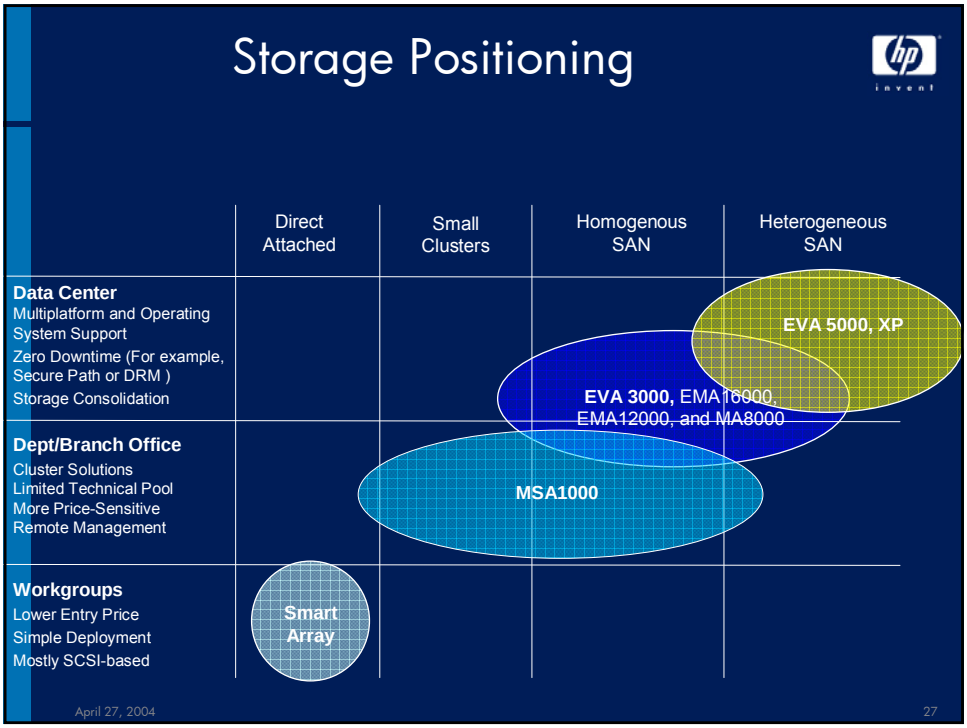
- Smart Array 5300
  - 2/4 U160 SCSI Channels
  - Up to 56 drives (4TB)
  - Raid0/1/5/ADG
  - Up to 256MB cache
- Supported on V7.3-1
  - 300MB/sec
  - 20K io/sec
- Fastpath with V7.3-2
- **Shadowing:**
  Doesn't support forced error commands so full shadowing support is an issue. Shadowing works fine but a member will be ejected if an unrecoverable disk error occurs on one member and an error cannot be forced on the shadow copy

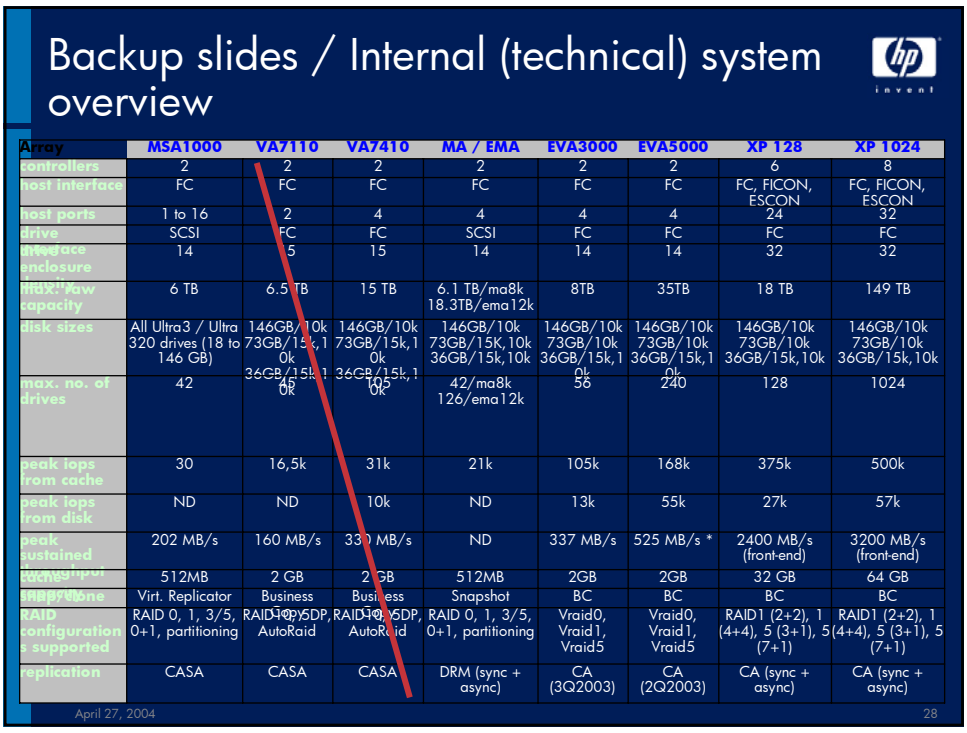
April 27, 2004 25

HSG / MSA / EVA Volume Construction

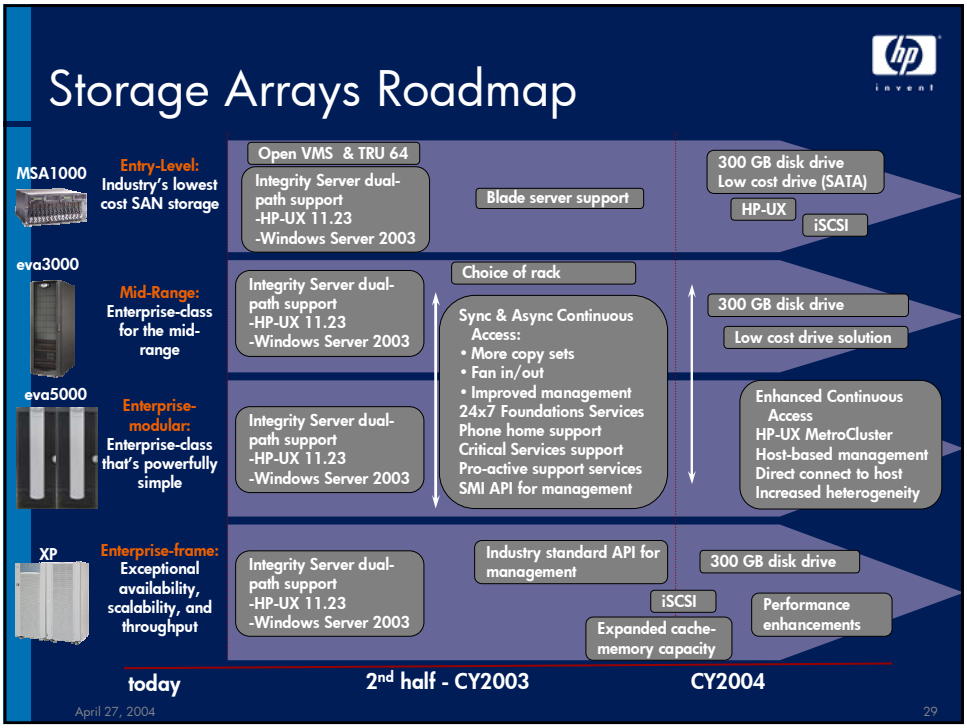


April 27, 2004 26





Array	MSA1000	VA7110	VA7410	MA / EMA	EVA3000	EVA5000	XP 128	XP 1024
controllers	2	2	2	2	2	2	6	8
host interface	FC	FC	FC	FC	FC	FC	FC, FICON, ESCON	FC, FICON, ESCON
host ports	1 to 16	2	4	4	4	4	24	32
drive	SCSI	FC	FC	SCSI	FC	FC	FC	FC
drive/ enclosure	14	15	15	14	14	14	32	32
max. raw capacity	6 TB	6.5 TB	15 TB	6.1 TB/ma8k 18.3TB/ema12k	8TB	35TB	18 TB	149 TB
disk sizes	All Ultra3 / Ultra 320 drives (18 to 146 GB)	146GB/10k 73GB/15k,10k 36GB/15k,10k	146GB/10k 73GB/15k,10k 36GB/15k,10k	146GB/10k 73GB/15K,10k 36GB/15k,10k	146GB/10k 73GB/10k 36GB/15k,10k	146GB/10k 73GB/10k 36GB/15k,10k	146GB/10k 73GB/10k 36GB/15k,10k	146GB/10k 73GB/10k 36GB/15k,10k
max. no. of drives	42	42	105	42/ma8k 126/ema12k	56	240	128	1024
peak iops from cache	30	16,5k	31k	21k	105k	168k	375k	500k
peak iops from disk	ND	ND	10k	ND	13k	55k	27k	57k
peak sustained throughput	202 MB/s	160 MB/s	330 MB/s	ND	337 MB/s	525 MB/s *	2400 MB/s (front-end)	3200 MB/s (front-end)
cache	512MB	2 GB	2 GB	512MB	2GB	2GB	32 GB	64 GB
snapvalue	Virt. Replicator	Business Copy SDP, AutoRaid	Business Copy SDP, AutoRaid	Snapshot	BC	BC	BC	BC
RAID configuration s supported	RAID 0, 1, 3/5, 0+1, partitioning			RAID 0, 1, 3/5, 0+1, partitioning	Vraid0, Vraid1, Vraid5	Vraid0, Vraid1, Vraid5	RAID1 (2+2), 1 (4+4), 5 (3+1), 5 (7+1)	RAID1 (2+2), 1 (4+4), 5 (3+1), 5 (7+1)
replication	CASA	CASA	CASA	DRM (sync + async)	CA (3Q2003)	CA (2Q2003)	CA (sync + async)	CA (sync + async)


Storage Arrays Roadmap
MSA1000
Entry-Level:
Industry's lowest cost SAN storage
Open VMS & TRU 64
Integrity Server dual-path support
-HP-UX 11.23
-Windows Server 2003
Blade server support
300 GB disk drive
Low cost drive (SATA)
HP-UX
iSCSI
eva3000
Mid-Range:
Enterprise-class for the mid-range
Integrity Server dual-path support
-HP-UX 11.23
-Windows Server 2003
Choice of rack
Sync & Async Continuous Access:
• More copy sets
• Fan in/out
• Improved management
24x7 Foundations Services
Phone home support
Critical Services support
Pro-active support services
SMI API for management
300 GB disk drive
Low cost drive solution
eva5000
Enterprise-modular:
Enterprise-class that's powerfully simple
Integrity Server dual-path support
-HP-UX 11.23
-Windows Server 2003
Enhanced Continuous Access
HP-UX MetroCluster
Host-based management
Direct connect to host
Increased heterogeneity
XP
Enterprise-frame:
Exceptional availability, scalability, and throughput
Integrity Server dual-path support
-HP-UX 11.23
-Windows Server 2003
Industry standard API for management
300 GB disk drive
iSCSI
Performance enhancements
Expanded cache-memory capacity
today
2nd half - CY2003
CY2004
April 27, 2004
29


Tape Products
April 27, 2004
30

New Storage Products

Tape Libraries

- ESL 9595
  - Up to 16 SDLT / LTO drives
  - Up to 4 2gb FC ports
  - Up to 595 cartridges
    - –95TB of uncompressed data
    - –900GB/hour uncompressed backup rate SDLT 160/320
    - –1.7TB/hour uncompressed backup rate LTO-2
- MSL 5052
  - 4 SDLT drives / 52 cartridges
- MSL 6060
  - 4 LTO drives / 60 cartridges
- MSL 5026
  - 2 SDLT drives / 26 cartridges
- MSL 6030
  - 2 LTO drives / 30 cartridges

April 27, 2004 31

New Storage Products

Network Storage Router M2402

- 1U Fibre Channel-to-SCSI router
  - 2Gb FC Support
  - 4 module slots
    - –2x 2gb Fibre Channel
    - –4x LVD/SE SCSI
    - –4x HVD SCSI
  - Web Based Management
  - Embedded product for tape libraries (E2400-160)
  - Supported in V7.3-1
    - –Tima kit for V7.3 (fibre_scsi v300)
    - –Tima kit for V7.2-2 (fibre_scsi v300)

April 27, 2004 32

New Storage Products

Network Storage Router N1200

- 1U Fibre Channel-to-SCSI router
  - High Performance - 2Gb FC Support
    - 200 MB/s of information throughput
  - 1 2Gb Fibre Channel port
  - 2 U160 LVD SCSI ports
  - Web Based Management
  - Embedded product for tape libraries (E1200-160)
  - Supported in V7.3-1
    - Tima kit for V7.3 (fibre_scsi v300)
    - Tima kit for V7.2-2 (fibre_scsi v300)

April 27, 2004 33

New Storage Products

LTO Tape Drives

- VMS will never support Ultrium 1 drives
  - No support of odd byte transfers
- VMS supports the Ultrium 2 (LTO460) drives
  - Testing completed on all current AlphaServer systems
  - Supported in both direct-attach SCSI and behind the FC Bridges (NSR, MDR)
  - Now supported in **ESL/MSL libraries**

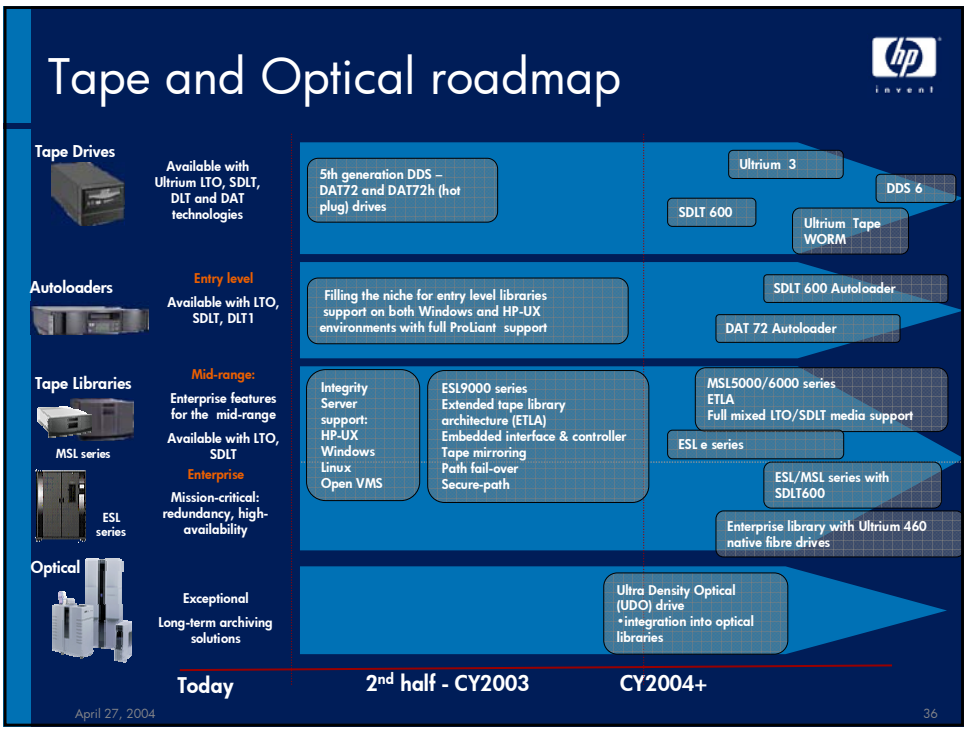
April 27, 2004 34

Recent Backup Performance Measurements

Drive	No compression	Highly compressed
SDLT 110/220	12MB/sec	49 MB/sec
STK 9940A	10MB/sec	33 MB/sec
SDLT 160/320	18MB/sec	49 MB/sec
STK 9940B	18 MB/sec	55 MB/sec
Ultrium 2	32MB/sec	60MB/sec

April 27, 2004 35

Tape and Optical roadmap

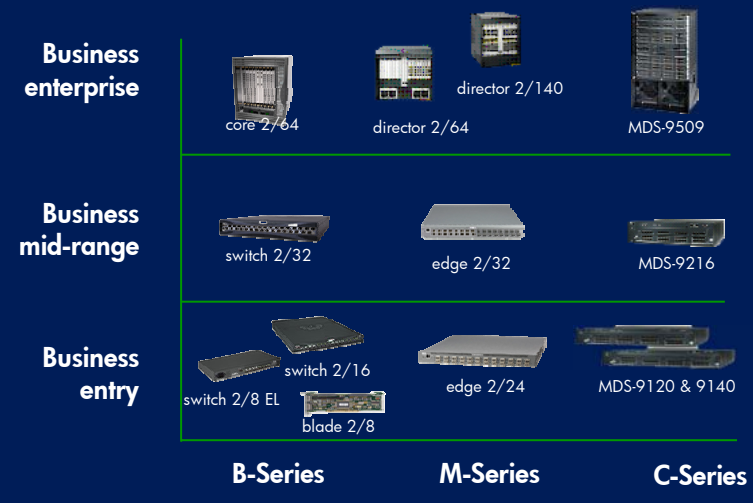


April 27, 2004 36

SAN Infrastructure

SAN Infrastructure product portfolio

Infrastructure alternatives that span the enterprise



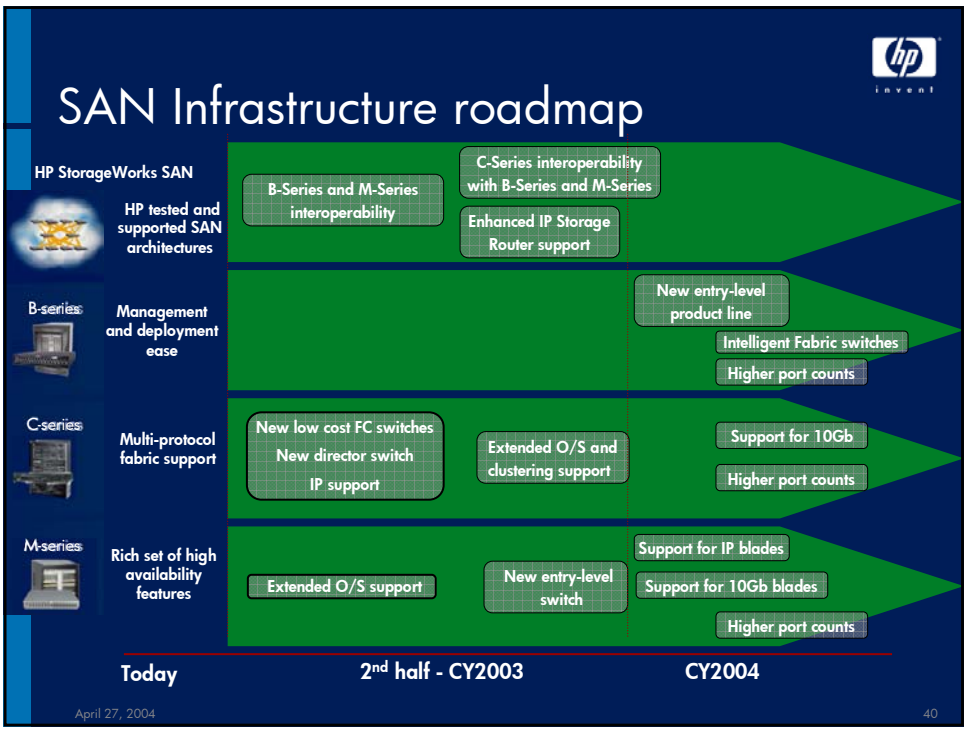
New SAN Features

Bigger SAN Switches

- SAN Switch 2/32
  - 32 port monolithic switch
    - 4 port 8Gb/sec ISL trunking available
- Core Switch 2/64
  - 2 x 64 port blade type switch
    - 16 ports per blade
    - Total of 128 2gb ports
    - 4 port 8Gb/sec ISL trunking available
    - Redundant everything
    - Highest performance, highest $$$ SAN option

April 27, 2004 39

SAN Infrastructure roadmap



April 27, 2004 40

Post V7.3-2 Projects

Post V7.3-2 Projects

Host-Based MiniMerge

- Recent cancellation of HSG80 Write History Logging project redirected the MiniMerge plans
- Available on V7.3-2 and V8.2 (Q42004)
- Will support **ALL** FC types (HSG, HSV, MSA, XP)
- Does NOT require any storage firmware assists

Post V7.3-2 Projects

Fibre Channel Cluster Interconnect

- Development project underway to send SCS traffic over FibreChannel
  - Provides LAN over FC
  - Use PEDRIVER to provide SCS communication
  - Goal is to provide stretch clusters without requiring additional cluster interconnect
    - –May also have "cleaner" failure characteristics because SCS and storage will fail as a single unit
    - –CI/MC class latency is not a goal (and not possible)
  - It is a non goal to provide general TCP/IP or DECNET over FC links (but we'll probably get it for free)
  - Prototype working since Sept-2002
  - Release commitment is OpenVMS V8.2

April 27, 2004 43

Itanium® Based Systems IO Plans

April 27, 2004 44

OpenVMS Itanium® Based Systems IO Plans

- SCSI
  - A6829A U160 controller (lvd only)
  - U320 controller (lvd only)
  - No plans for multi-host SCSI
- Host Based Raid
  - Smart array family
- FibreChannel
  - 2gb adapter
    - Storage & Cluster Interconnect
    - KGPSA-EA in V8.1 (emulex)
    - A6826A in V8.2 (qlogic)
  - Storage Arrays
    - HSG / EVA / MSA / XP

April 27, 2004 45

Longer Term Storage Interconnects

April 27, 2004 46

Long Term Storage Interconnects

- 4Gb FibreChannel
  - 2004???
  - Gaining momentum for SAN usage
- 10Gb FibreChannel
  - 2005???
  - Very expensive infrastructure costs at first
  - Probably most interesting as inter-switch links
- iSCSI
  - Industry has stagnated some in 2003
  - Has some promise as a low-cost way to connect PCs to a SAN
  - Host performance overhead is the main issue today

April 27, 2004 47

Questions???

April 27, 2004 48