Session # 1594
Tape Backup Performance Issues Exposed!

by

Charles Curtis & Andy Buckley

HP Nearline Storage
Technical Advocates.

HP WORLD 2003
Solutions and Technology Conference & Expo

Objectives of this Session

- To ensure backup performance expectations are considered at the time the system architecture is designed, and not as an afterthought.
- To show a structured way to analyse, predict and test where potential bottlenecks may occur.
- To review some practical scenarios to illustrate some of the key points in practice.
- To update you with industry trends in the backup performance arena.

Today's Agenda

- The performance jigsaw – 5 easy pieces
- Tools available - to aid diagnosis
- 5 Performance scenarios analysed
- New developments –
  - New Legislation
  - Microsoft Volume Shadow Copy
  - Disk Trends,
  - Itanium Servers,
  - FC vs SCSI tape.
  - LTO3, SDLT 600
  - Extended Library Architecture performance analysis.
- Q & A

Over to you Charles

The Big Picture and the Problem with Backup

1. There is no simple answer to performance issues (job security?).

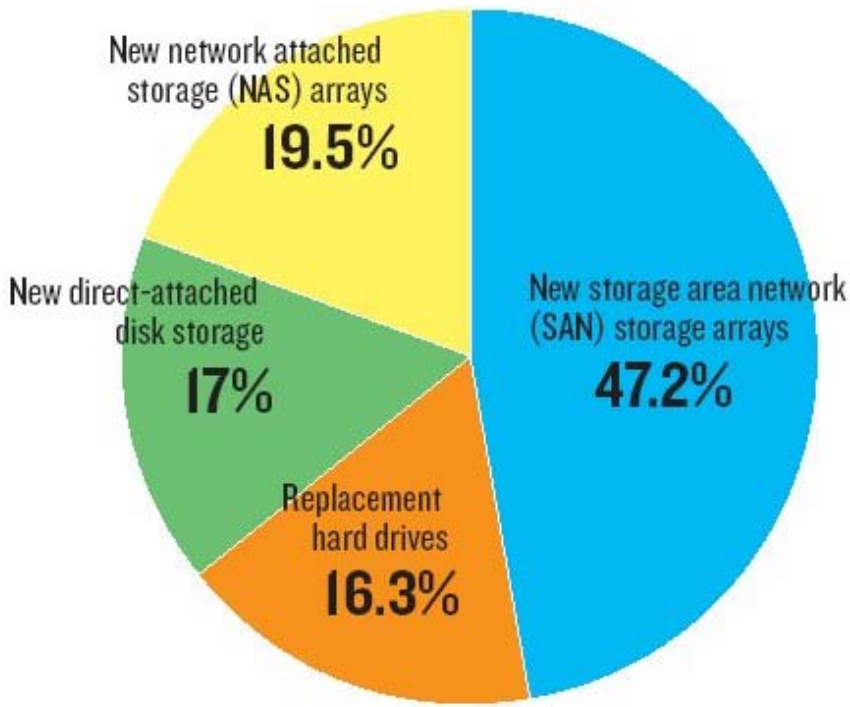
2. Most Backup Performance Issues Are Not Due to Tape Drive Performance.

3. Industry quoted Speeds are **not typical**.

4. Consider Backup & Restore as a SYSTEM

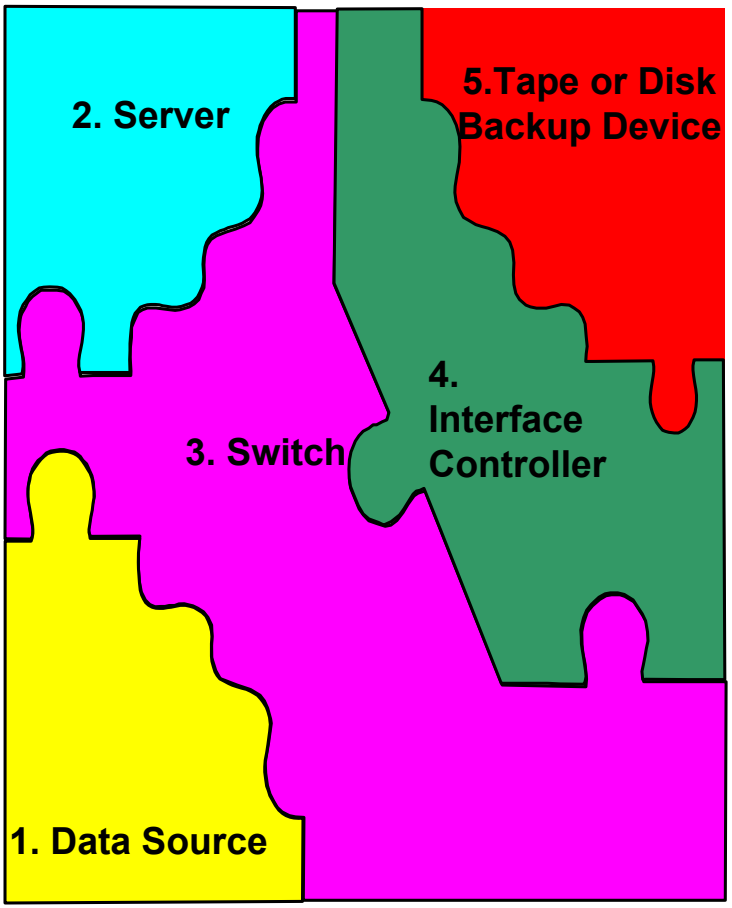
5. Suspect EVERYTHING, Trust Nobody!

Disk storage Purchasing Intentions



Source: Storage 2003 Purchasing Intentions

5 Easy Pieces of SAN Backup Tuning



1. Primary Storage – Although this could reside on a local adapter within the server, it is usually a large disk subsystem on the SAN network.

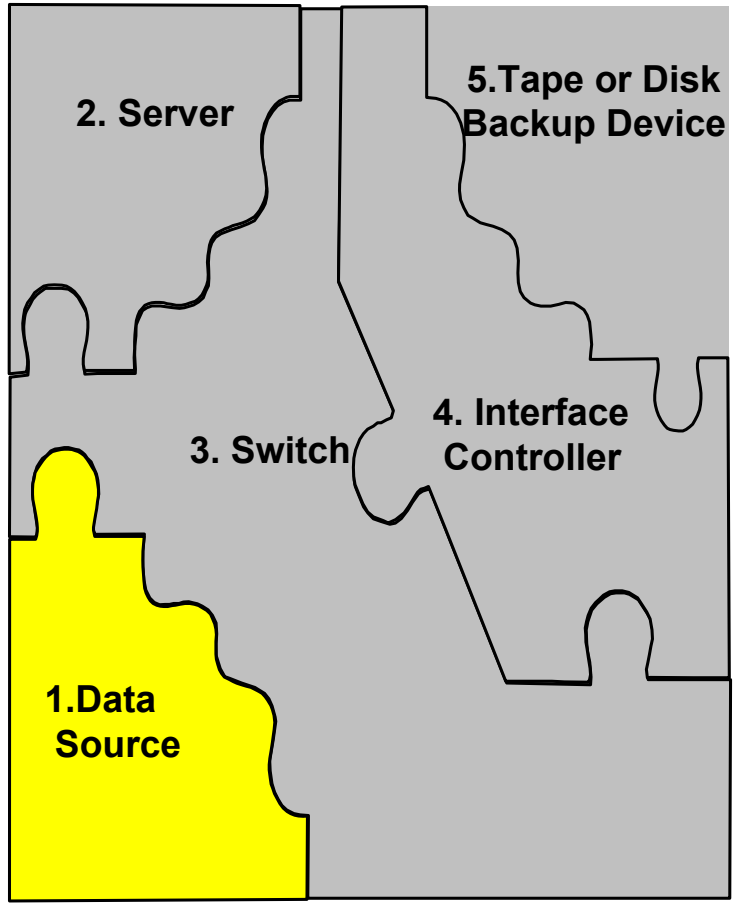
2. The Backup SERVER — connects via the SWITCH to its Primary storage and the backup target. It runs the backup application. It uses an HBA to link to the SAN

3. The Switch - The glue that holds everything together, this box permits systems and storage to be shared at high speeds.

4. The Interface Controller — This device extracts the SCSI commands out of the SAN Fibre Channel protocol and permits simple tape and disk drives to connect to the SAN.

5. The Tape or Disk Backup Target (Secondary Storage) System— SCSI devices that appear as locally connected SCSI to the Server, via the magic of Fibre Channel!.

Step 1: Data Source



Data Source considerations

- As a rule of thumb you need 3 X tape speed to maintain streaming.
- Conversely your backup speed is around 1/3 of your data source speed.
- Restore speed is generally around 40-60% of Backup speed.

Data Source considerations – RAID structure

Structure Considerations

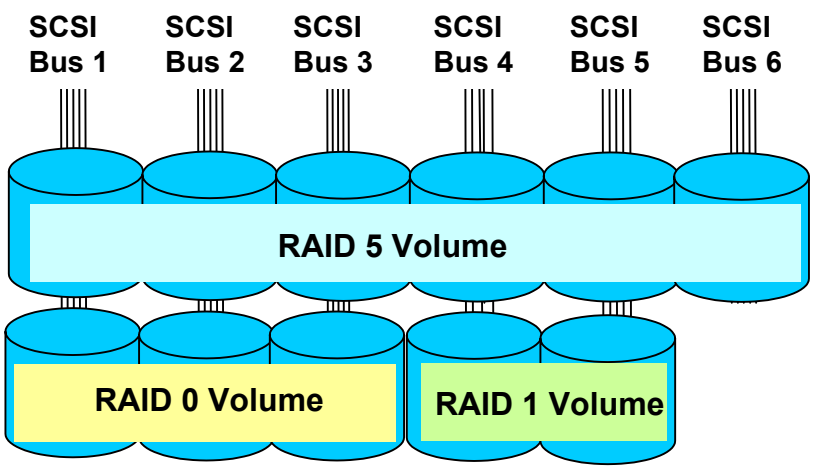
- Striping across multiple arrays controllers improves performance

- Multiple Volumes permit multiple jobs (multiple entry points for backup applications) - Parallelism

Data Source Considerations - RAID

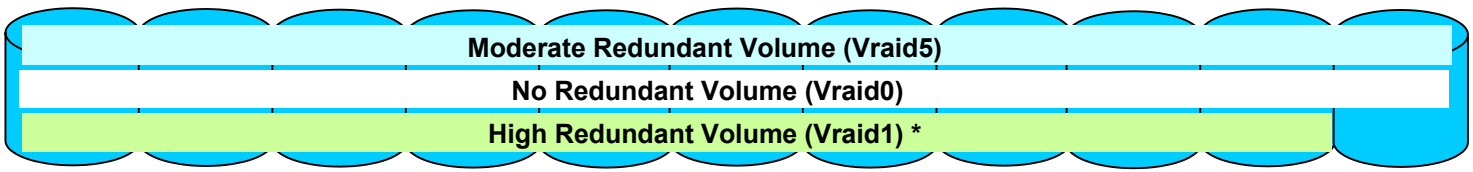
- eliminate throughput bottlenecks
- eliminate load balancing procedures for applications and databases

Yesterdays RAID Sets



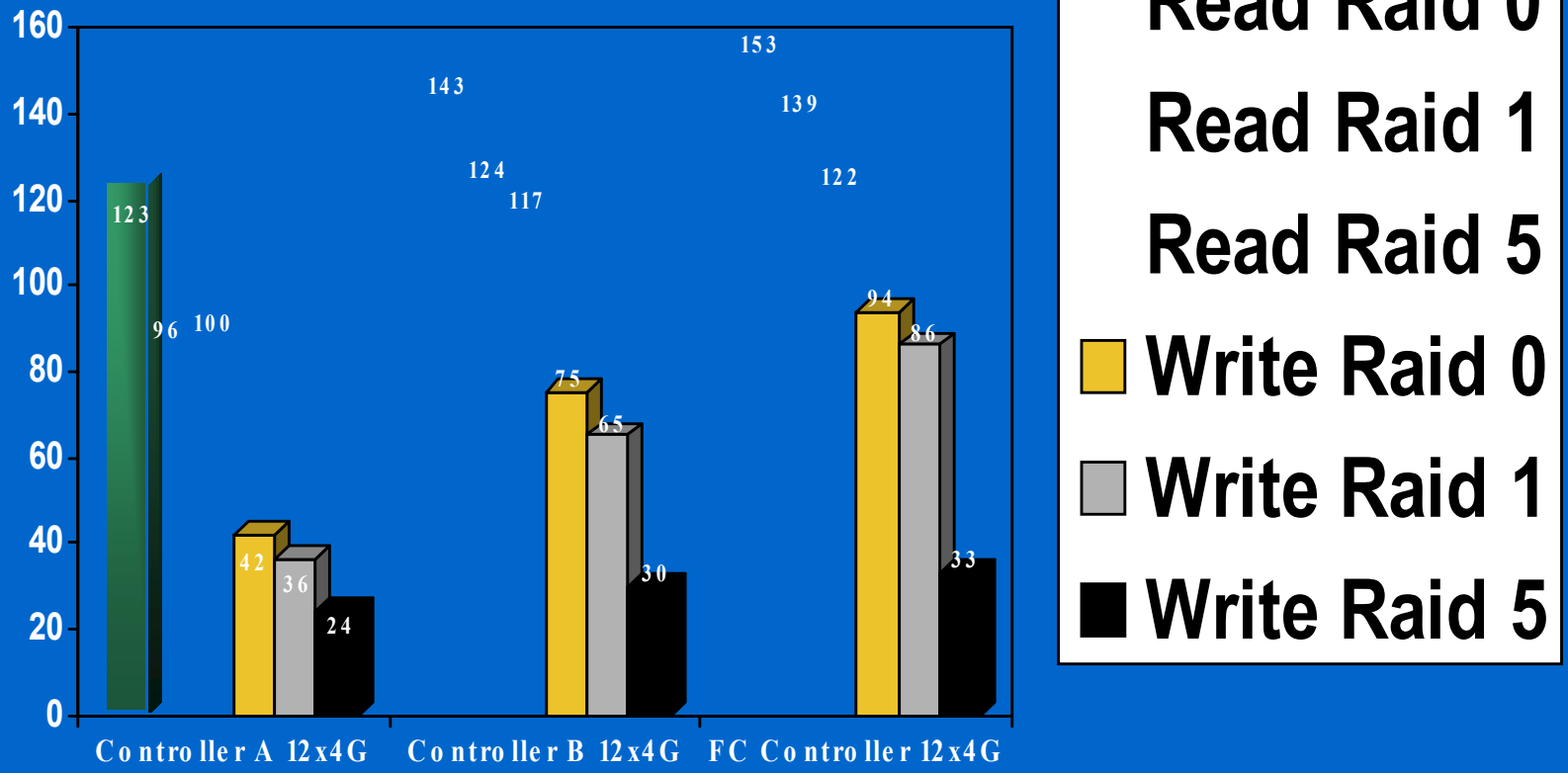
HP's Virtual Disk Technology

Workload is evenly distributed across the spindles in group



*RAID V1 uses even numbers of disks

Data Source Considerations- RAID



Disk Array Performance

	XP1024	EVA5000	EVA3000	MSA1000
# Disks in base unit	1024	240	56	14
Max Throughput MB/sec	2000	628	335	200
# FC ports on base unit	64	20	4	2

Source Data : Best Case vs. Worse Case Data

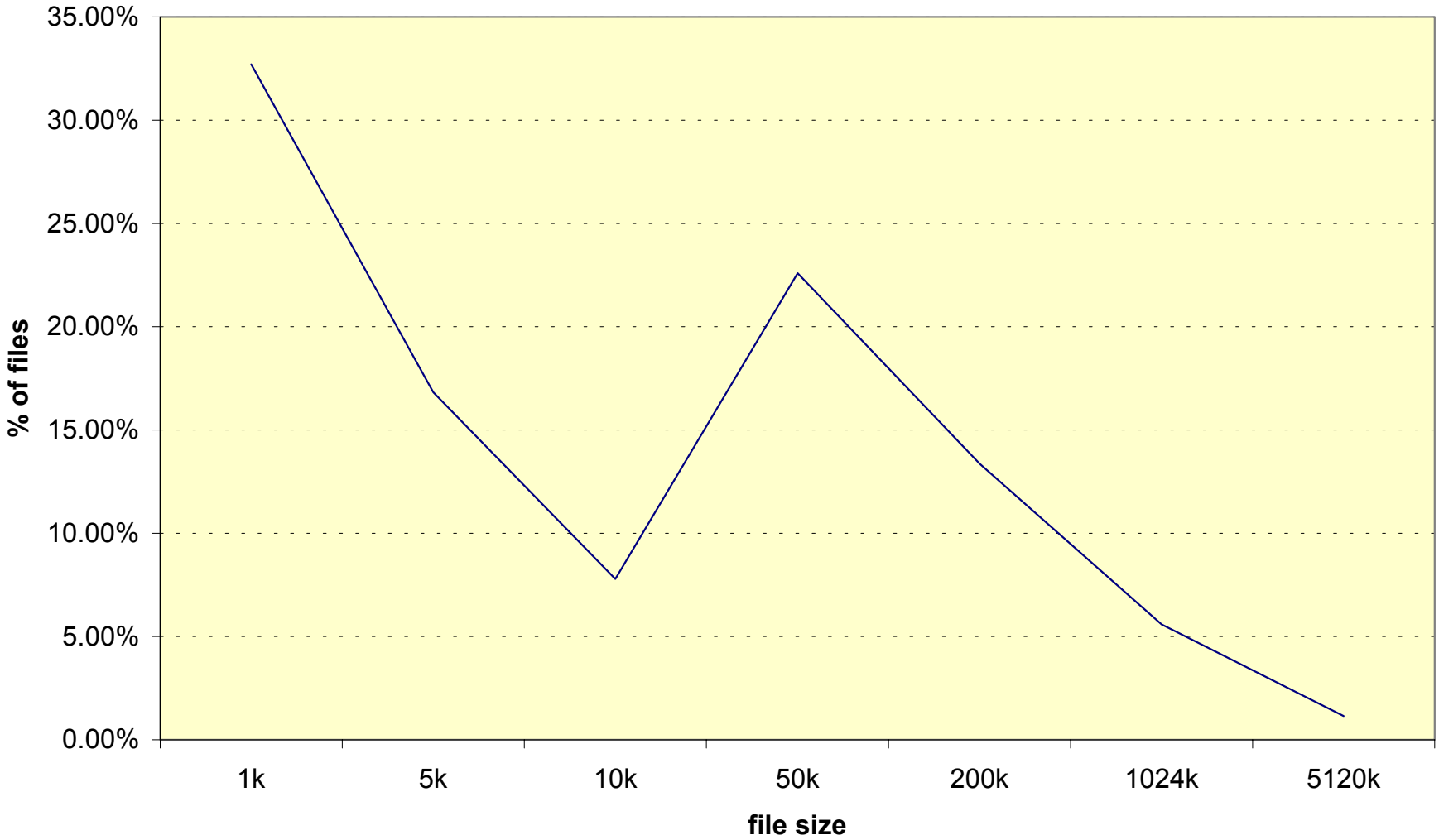
Best Case –

- Low File Count
- Large Files
- Simple File Tree Structure
- Short file names
- Compressible Data (2:1)

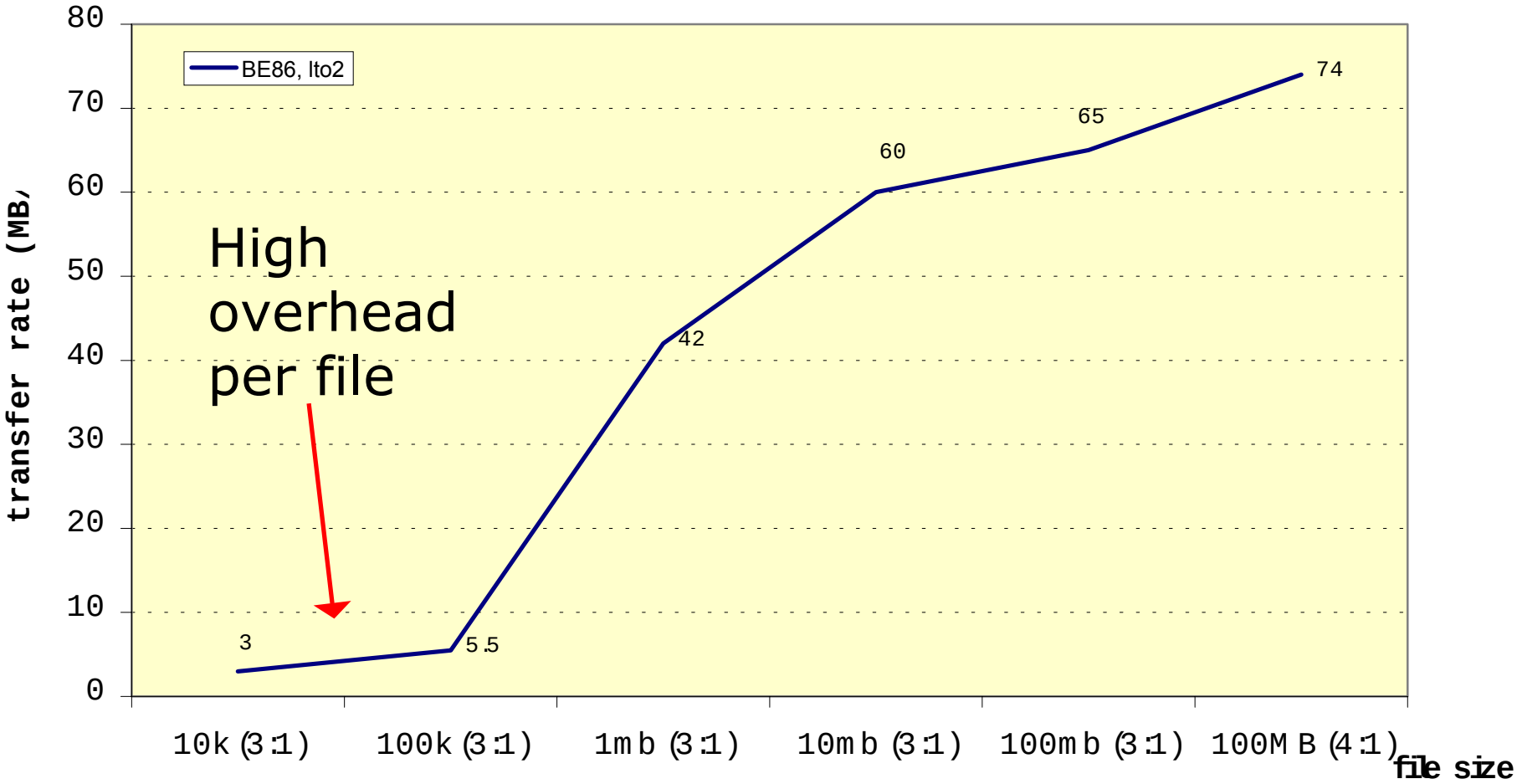
Worse Case –

- High File Count
- Tiny Files (1k byte)
- Complex directory
- Long File Names
- Non-Compressible (1:1)

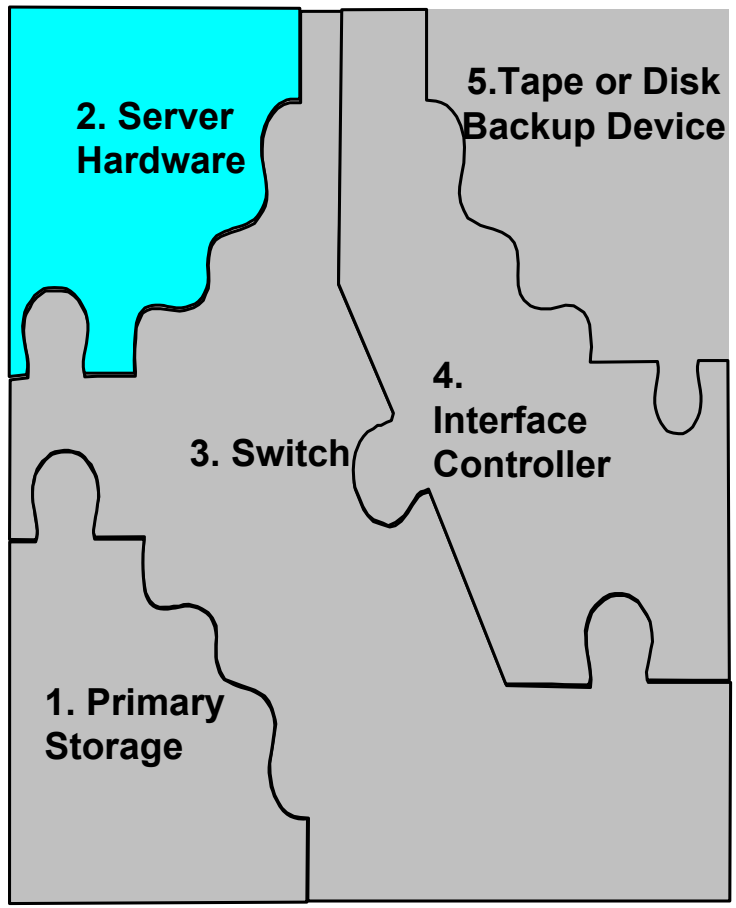
Example file size distribution



File size performance example (NT)



Step 2: Backup Server Part I
Procs/Memory/PCI



Sizing the Backup Server

- Hardware Considerations (Microsoft W2K or NT)
  - Number of Processors
  - Memory – 1GB minimum (app drivers?)
  - Boot Drives (Mirrored RAID-1)
  - Faster C: Drive can help OS and TAPE Software
  - PCI Bus(s) – 32bit, 64bit, 33-66-133Mhz, PCI-X, PCI-X 2.0
  - LAN / SAN connections

- Software Considerations (Microsoft NT/W2K)
  - Don't put the Backup app. database on a slow drive
  - Use the largest block sizes for tape drivers (64K or greater))

PCI shouldn't be the bottleneck

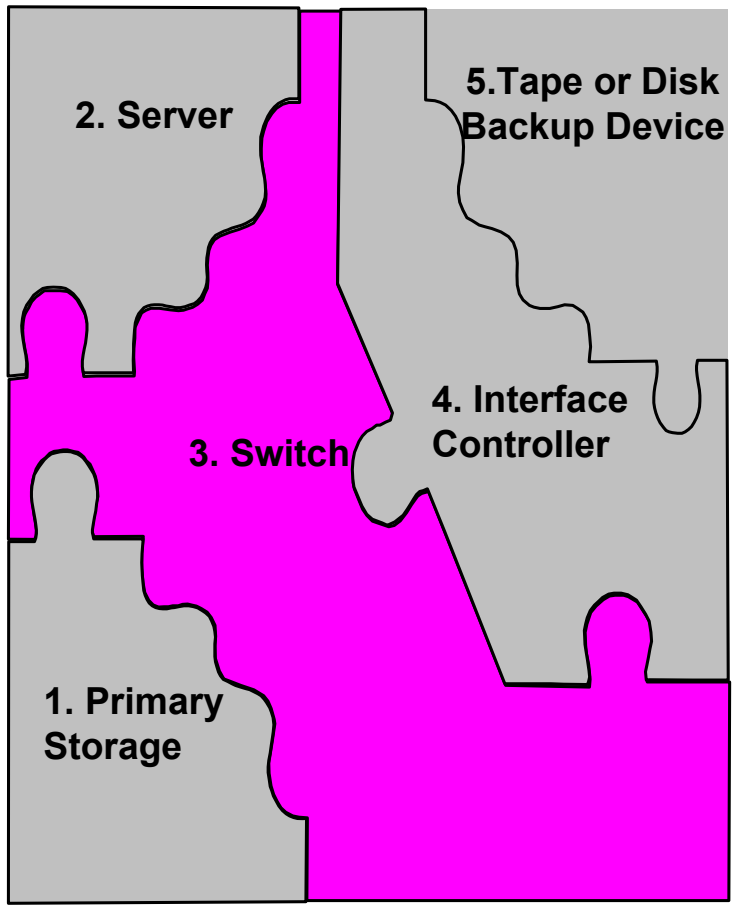
Clock Frequency	Bus Width	Burst Perf	Sustainable Perf
33MHz	32 bit	133 MB/s	115MB/s
33MHz	64 bit	266 MB/s	230MB/s
66MHz	32 bit	266MB/s	230MB/s
66MHz	64 bit	533MB/s	490Mb/s
133MHz (PCI-X)	64 bit	1066MB/s	980MB/s
266 MHzPCI-X 2.0*	64 bit	2132MB/s	1960MB/s
533 MHzPCI-X 2.0*	64 bit	4264 MB/s	3940MB/s

*-in compliance testing Aug 03

Sizing a backup server - Summary

Parameter	Rule of Thumb	Comment
Processors	Today a typical 2 proc machine can handle up to 5 single data streams @ 30Mb/sec at 75% processor load. From a datapoint of a single stream = X% processor load, then a concurrent stream = 2X% (irrespective of concurrency value)	Critical at the planning stage to understand likely growth and select a "scaleable" server. Highly scaleable servers also have better memory access speeds and more PCI bus "peer" capabilities.
Memory	1GB should be adequate unless large Nos of parallel streams.	Some Backup apps scale better in performance than others if more memory available.
PCI Architecture	Use servers with "peer" PCI buses not bridged PCI buses. Keep gigabit /FC cards in the 64/66 slots and SCSI HBA's in the "other slots"	Unlikely that PCI will be the bottleneck in most cases.

Step 3: Switch Connection



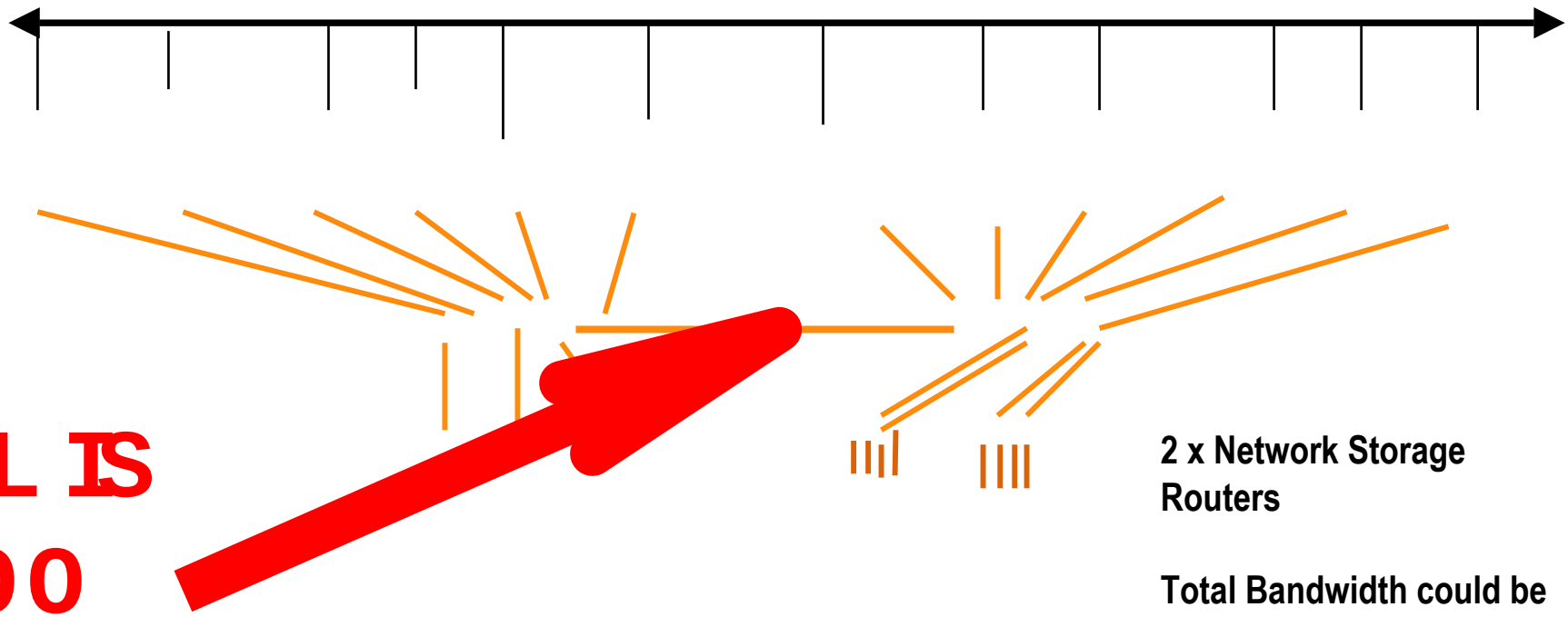
The SAN Switch

- **Functions of the SAN Switch**
  - **Interconnection of all SAN components.**
  - **Performance Monitoring device used to analyze backup performance problems.**
  - **Permits "speed matching", the interconnection of newer 2 Gig devices to older 1 Gig devices without forcing the 2 Gig device to run at a 1 Gig rate all the time.**
  - **Zoning controller/manager – zoning is much like LUN masking, in that it limits visibility of SAN devices to each other.**

The SAN Switch – TIPS

- **Tuning Considerations: None for individual connections. They usually either WORK WELL or NOT AT ALL!**
- **FABRIC TUNING – The SAN Fabric can consist of many switches that are interconnected via ISL's (Inter Switch Links). Backup tuning should always analyse the pathways from the source to the target (disk to tape) to determine if the paths through the switches are bottlenecked by insufficient ISL's.**

What Is WRONG with this Picture?



ISL IS TOO SMALL!

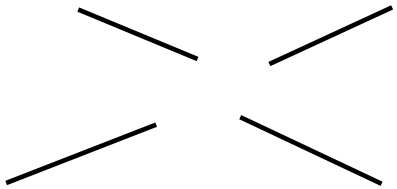
MA8000, EMA12000

2 x Network Storage Routers

Total Bandwidth could be 600 MB/sec !

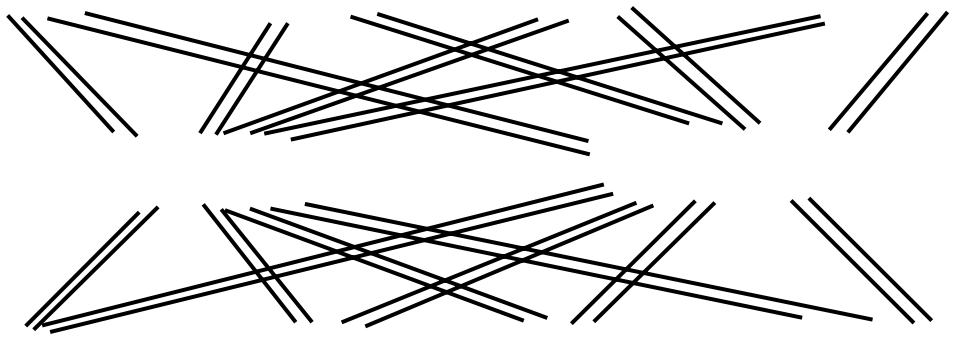
Backbone SAN - Skinny Tree Fabric ISL's

Some methods of Switch Interconnection Techniques



5 Switch Skinny Tree
(56 - 60 F-Ports)

5x2 Switch Skinny Tree (HA - 2 Fabrics)
(112 - 120 F-Ports)



10 Switch Skinny Tree
(96 - 112 F-Ports)

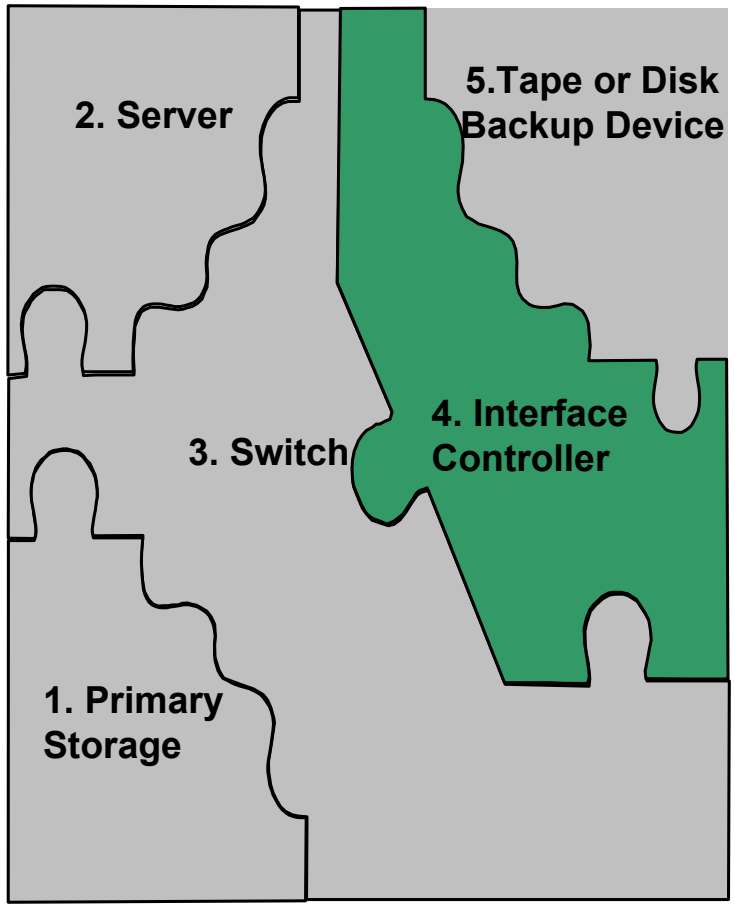
Next generation FC.

Speed	Throughput Mbps (duplex)	Line Rate Gbaud	Release date
1 GFC	200	1.0625	1998
2 GFC	400	2.125	2001
4 GFC	800	4.25	Late 2003
10 GFC	2400	10.5	2004

Recent changes in 4Gb FC

- 4Gb/sec is an extension of 2Gb
- 10Gb is a new development with no backwards compatibility
- Originally 4Gb destined for “intrabox” use only. E.g. inside disk arrays
- Recent FCIA review has decided to move forward with a 4Gb switching network – expect products early 2004
- 10Gb likely to be used for ISL’s only, no “devices” will connect to 10Gb
- Impact for Backup: to avoid wasted 4GB bandwidth a controller based approach to libraries is best placed to utilize these enhancements. (see later)

Step 4: Interface Controllers



What is an Interface Controller?

- An Interface Controller is
  - NOT just a bridge
  - NOT just a router

- An Interface Controller
  - HAS the intelligence to actually look at the data, and makes some judgements as to where it will go.
  - HAS the mission to communicate with a management system to provide performance tuning, diagnostic monitoring and reporting abilities.
  - HAS the ability to support advanced functions such as….
    - Selective Storage Presentation
    - Rapid (Serverless) Backup
    - Secure Path
    - Partitioning
    - Virtualization

Interface Controller Performance

- Ultra 2 Routers have a total bandwidth of 140MB/sec, 2 SCSI ports

E1200/E2400 – up to 2 x SDLT220/320 or 2 x Ultrium 230 per SCSI port

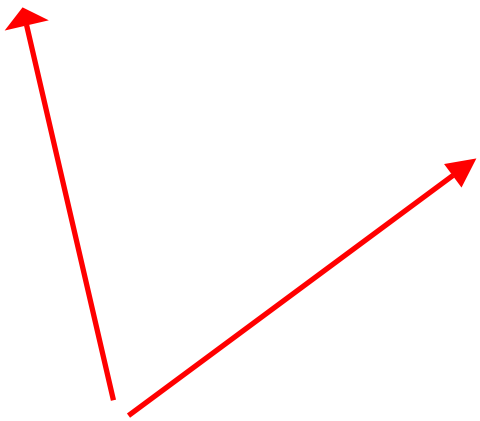
- Ultra 3 Routers have a total bandwidth of 280MB/sec, 4 SCSI ports

E2400-160 & M2402 – 1 Ultrium 460 per SCSI port

- up to 2 x SDLT220/320 or Ultrium 230 per SCSI port

DON'T EXCEED THESE VALUES OR THE ROUTER WILL BECOME THE BOTTLENECK TO BACKUP PERFORMANCE.

Selective Storage Presentation – optimises performance by maximising FC bandwidth



Bind Host HBAs to specific ports and drives on the Router

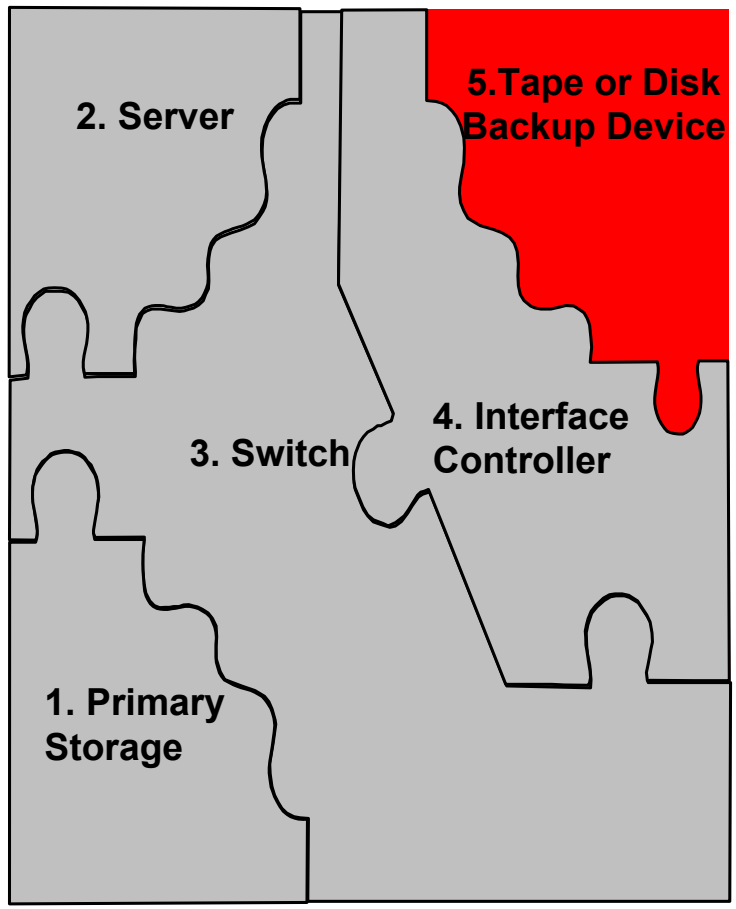
Interface Controller Bottleneck Considerations

- Don’t forget to calculate your data throughput rate based on what level of compressibility your data has on your specific tape drive technology. If you don’t know, assume at least 2:1, so that you DOUBLE your expected data rates compared to using 1:1 data.

- When you are considering how much data can go through the router at one time, consider analyzing your data sets compressibility and how you might stagger your backup jobs so that data flow based on tested transfer rates (see tools coming up!) are run in a way that doesn’t exceed the router’s performance limits.

Step 5: Tape Drives

Over to you
Andy

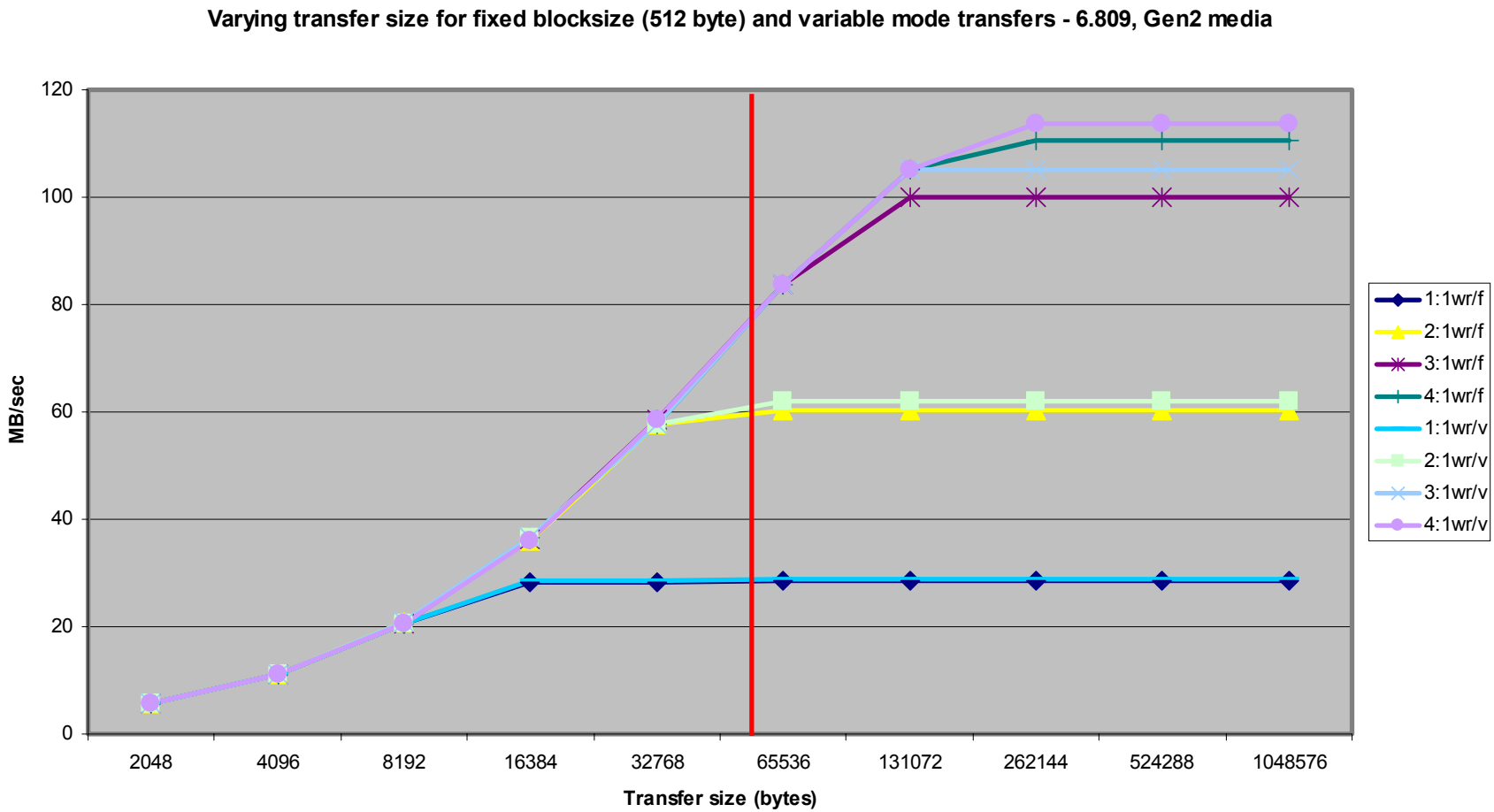


High Performance Tape

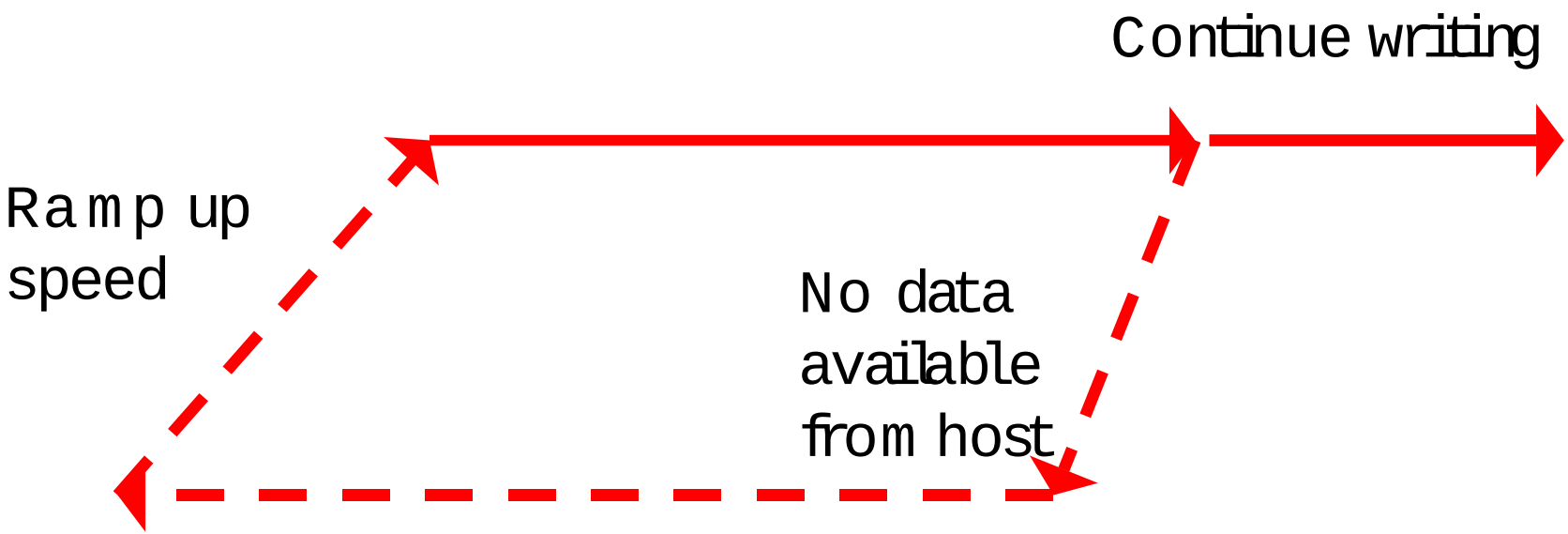
Drive	Capacity (native) GB	Transfer Rate native GB/Hr	Interface(s)
HP LTO 1	100	54	Ultra 2 SCSI
SDLT220	110	57	Ultra 2 SCSI
SDLT320	160	57	Ultra 2 SCSI
HP LTO 2	200	108 108	Ultra 3 SCSI FC- 2Gb/s
STK9940B	200	108	FC -2Gb/s
AIT 100	100	43	Ultra 2 SCSI
S-AIT1	500	108	Ultra 3 SCSI

Tape Transfer/Block Size – ensure 64K or above

HP Ultrium 460 on Solaris 9 with Ultra 160 HBA

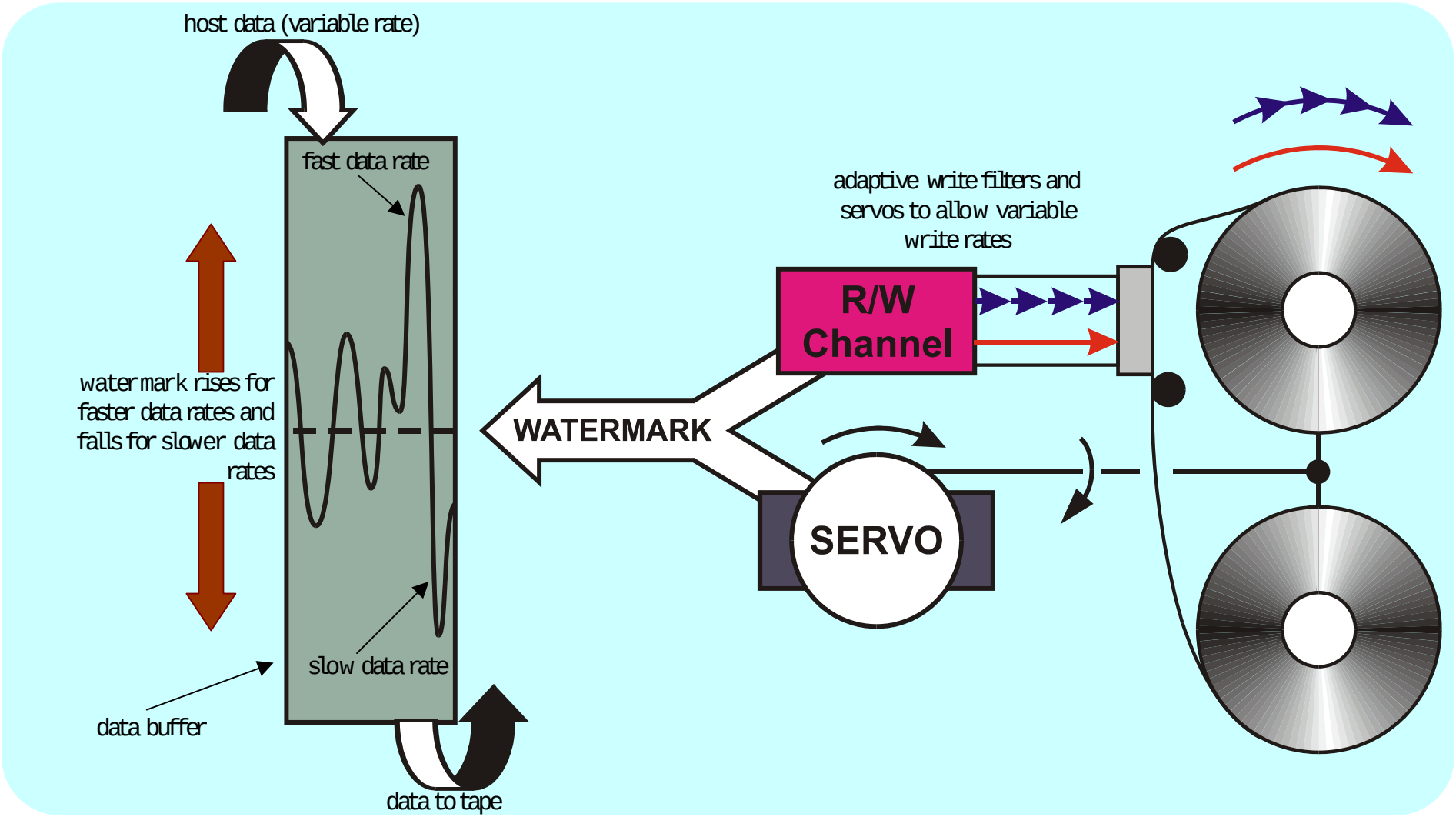


Streaming vs Repositioning



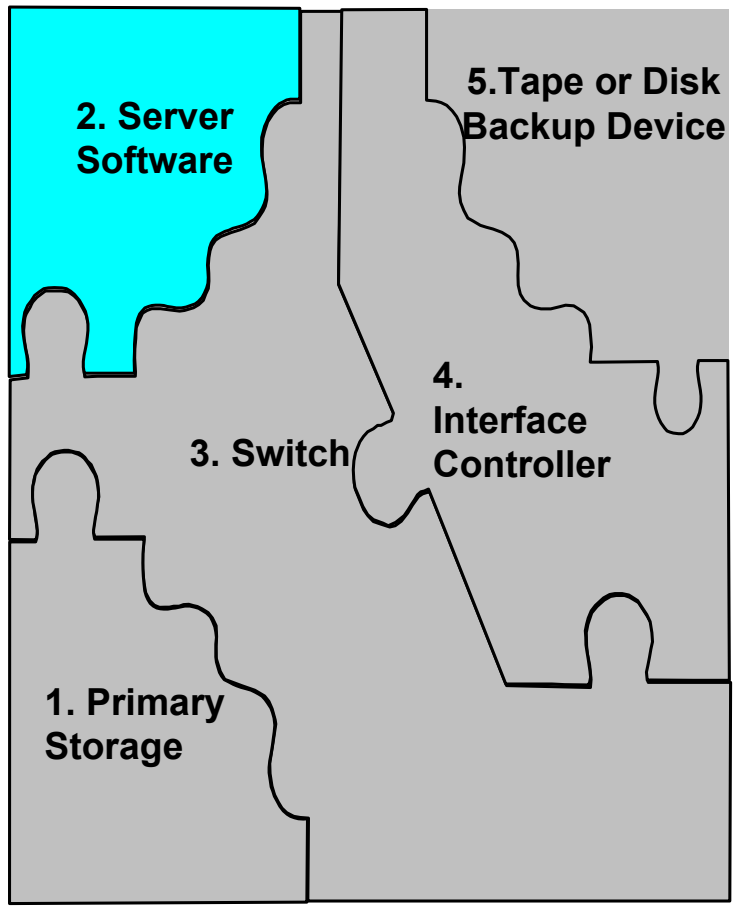
e.g. Ultrium 460 running at 7m/sec requires a full 3 seconds to reposition

Only Ultrium technology has Adaptive Tape Speed



Step 2 (backup Server) Part II

Operating System considerations and ISV performance tuning



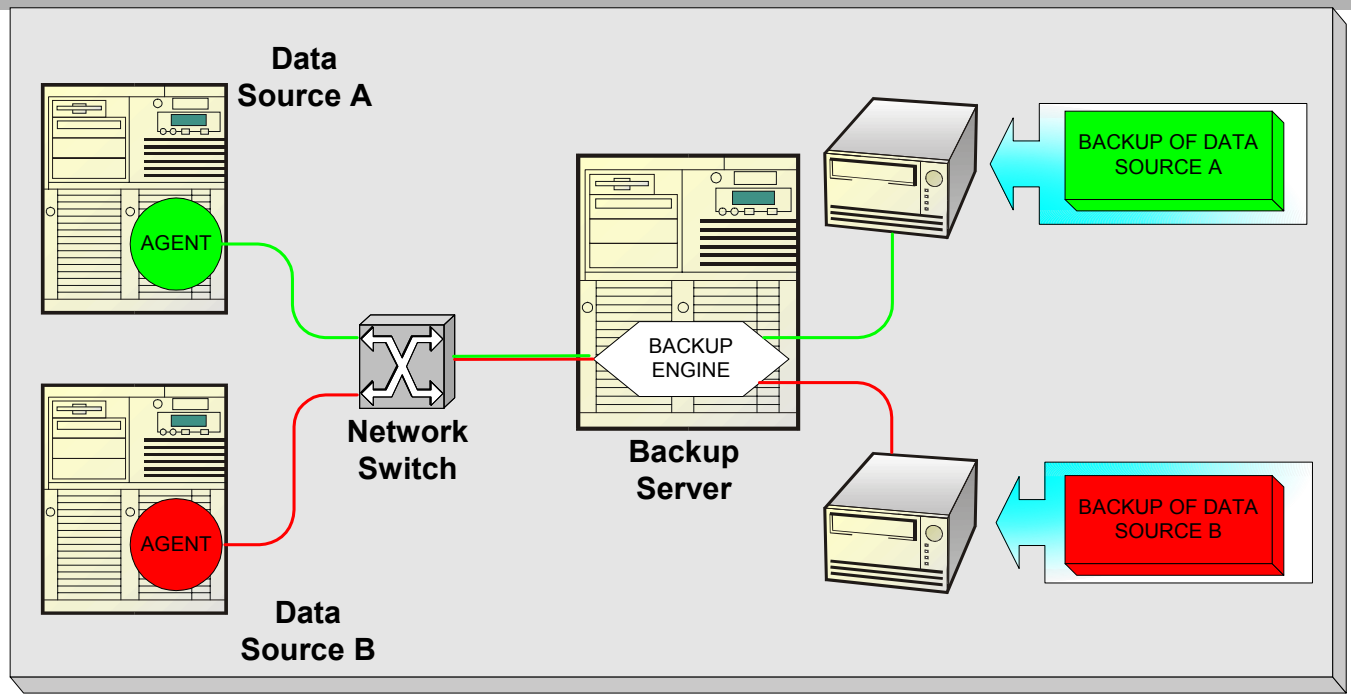
OS Transfer Size

- OS Transfer size is the maximum amount of data the operating system will transfer in a single “operation”
  - Windows is preset to 64K unless MaximumSGList is changed in the registry
  - hp-ux has a 1MB limititation (although 256K at atomic level)
    - kmtune –s scsi_maxphys = <value>
  - More recent Solaris versions have 64K default extendable to 16M via st.conf

ISV performance tuning

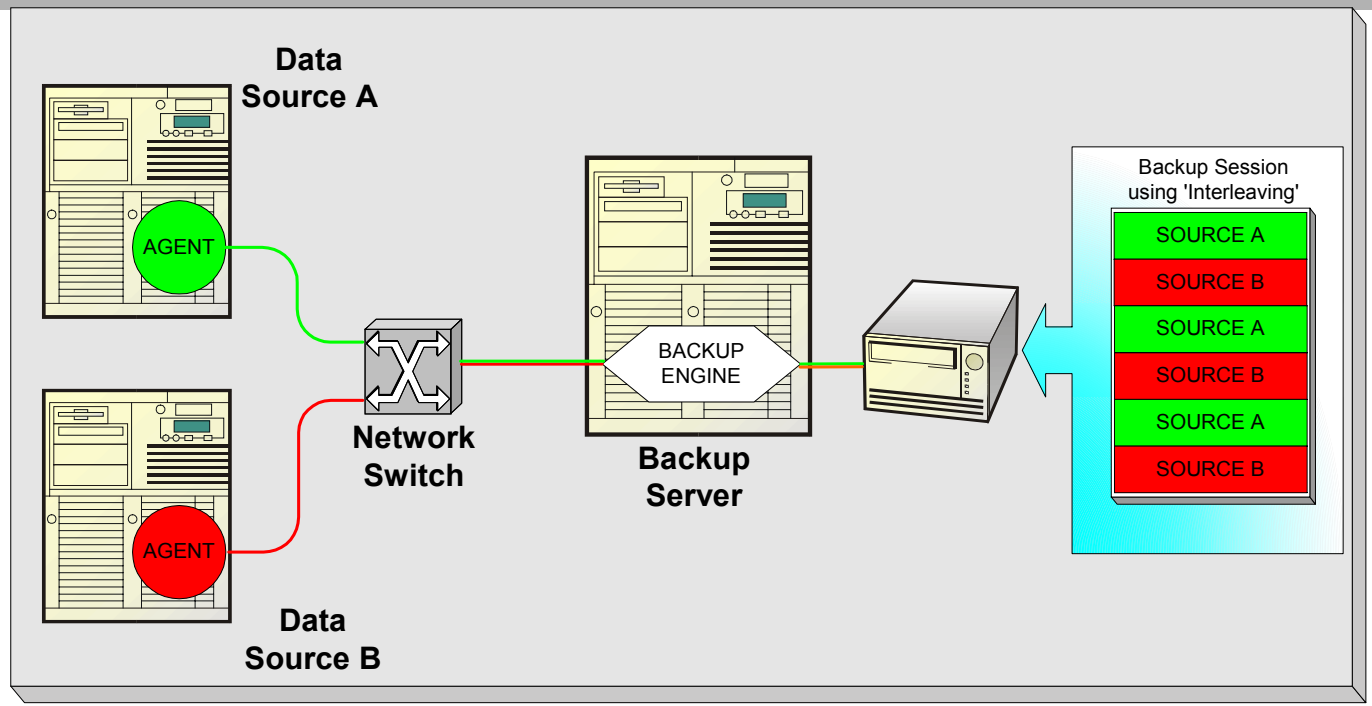
- Concepts
  - Multiple Streams
  - Multiplexing (optimising tape performance)
  - Image Backup
- Transfer Size
- Database API's (bypass filesystem)
  - RMAN (recovery manager) - Oracle
  - VDI (Virtual device interface) – SQL
  - Limitations of Exchange API
- Samples featuring HP Data Protector.
  - Tape Blocksize/Disk Buffers
  - Concurrency

Network Backup with 'Parallel' Data Transfers



- Device Parallelism
  - 'Device Parallelism' or 'Multiple device streams' uses the Principle that data from a specific source system can be routed to a dedicated device in a one-to-one relationship

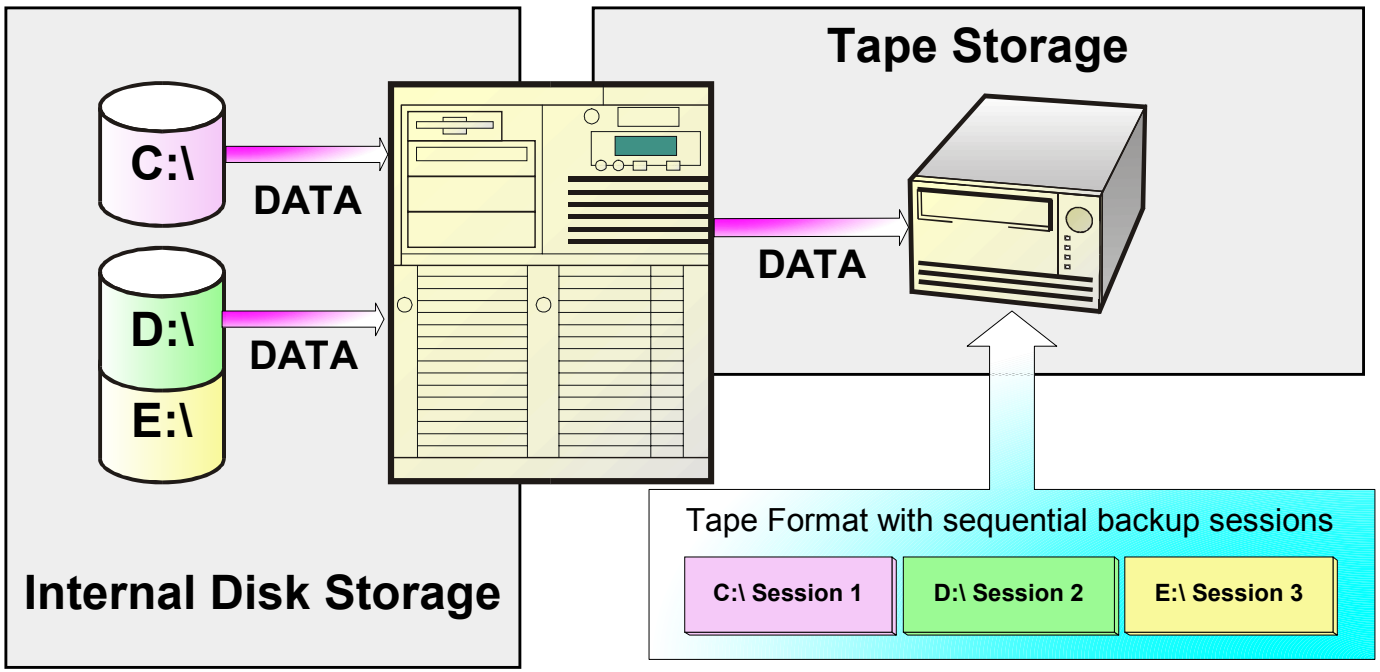
Backup with Concurrency (Interleaving/multiplexing)



- Combines data from multiple Data streams from multiple sources to a 'specific backup session' onto a single tape or backup tape set, so we can maintain tape streaming.
- Increases backup performance but can reduce restore performance.

ISV Software 'Features' – Image Option

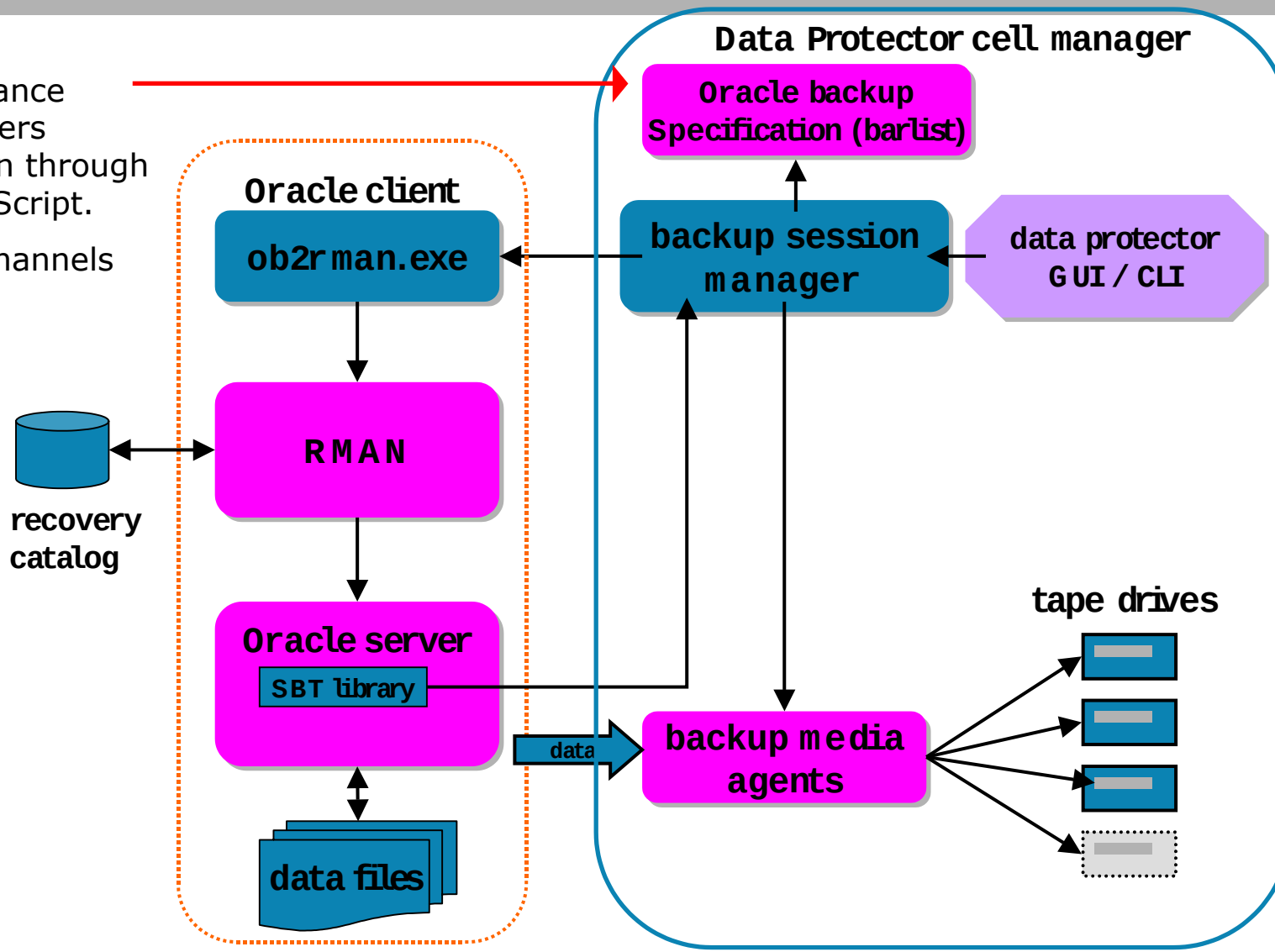
- Image backup operates at the device level rather than the file system level
- The source Drive Must be quiesced
- Lower File-System overhead resulting in higher performance
- Best used when lots of small files would give slow file access.
- Single file restore is possible but slow.



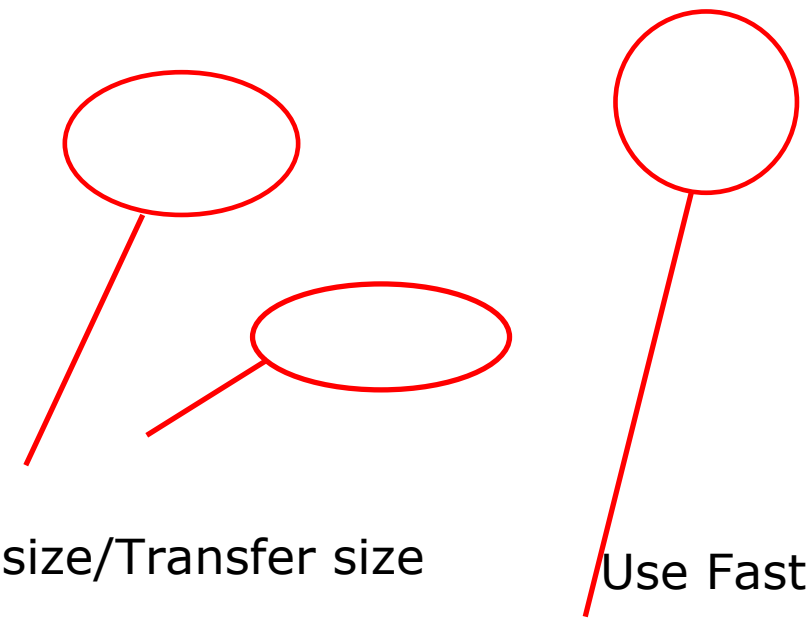
ISV performance tuning – database API's - Oracle

RMAN performance parameters passed in through Backup Script.

E.g. # channels



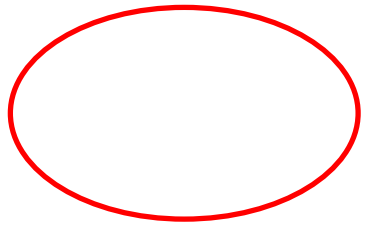
ISV performance tuning – database API's - SQL



Optimise Block size/Transfer size

Use Fast Direct Mode

ISV performance tuning
Tape Block Size/Buffers and Segments

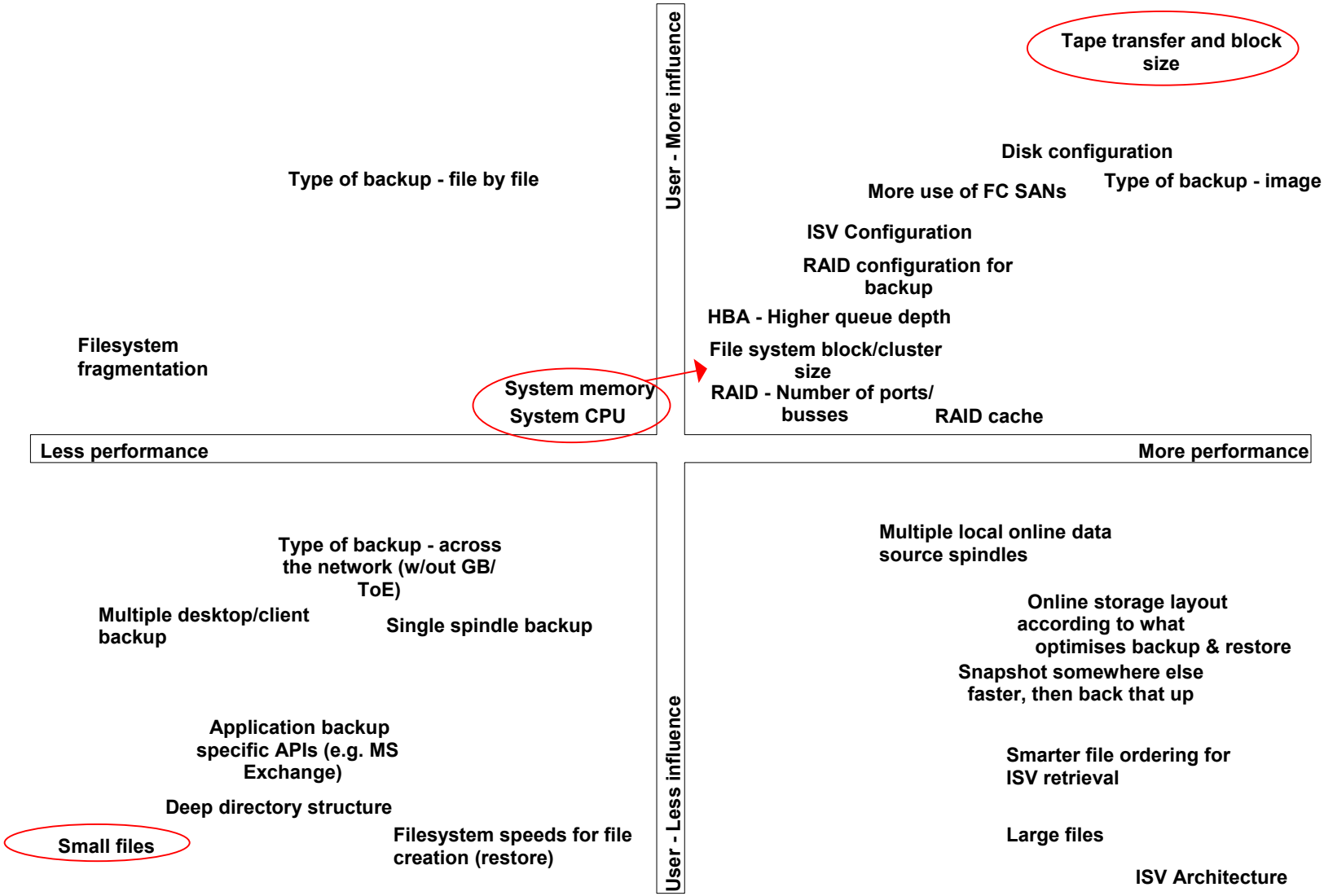


ISV performance tuning – Concurrency/Multiplexing

Set at a device level.

Calculate these concurrency levels as part of an overall backup strategy to meet specific window requirements

Tape backup performance influence map - Summary



Your turn Charles

What About Restore Performance ?

multi-level data protection

Level	Protection	Protects Against	Recovery Time
1	RAID mirroring	device failure equipment failure	instant
2	snapshots replication and clones	data corruption user error equipment failure data corruption user error site destruction	seconds to minutes
3	tape backup & restore	equipment failure data corruption user error site destruction virus & hacker attacks	minutes to hours
4	data vaulting	equipment failure data corruption user error site destruction virus & hacker attacks [illegible]	hours to days

Tape is still the foundation

Recovery considerations

- With higher performance environments recovery from catastrophic failure can take place at speeds of around 2 Terabytes/Hr Maximum. (16 drives in Library)

- Main restore bottleneck is in re-creating RAID consistency and creating files through the filesystem.

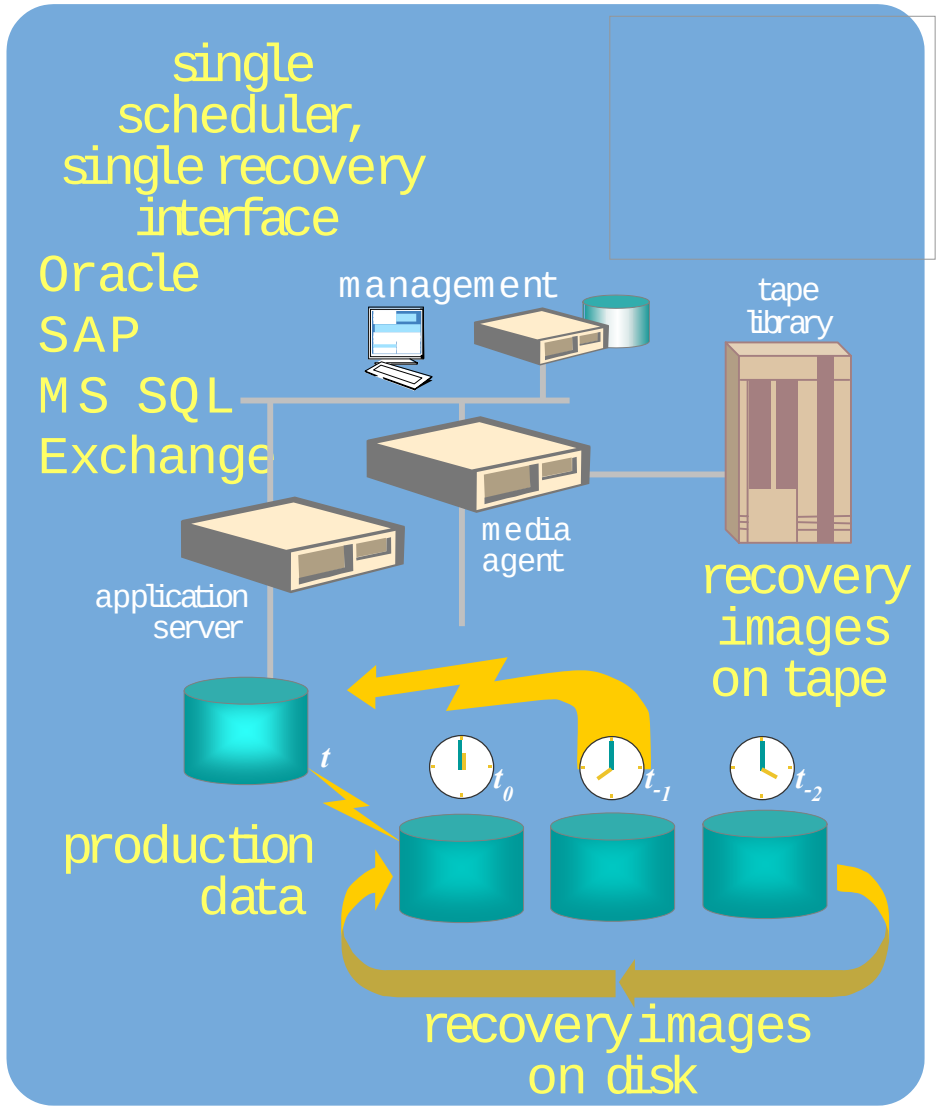
- Consider Parallelism to improve recovery times.

- Single file recovery performance can also be dependent on Drive search performance and ISV search algorithms.

- Single file restore from an Image backup can be slow.

Integrated fast recovery

- builds on zero-downtime backup techniques to retain multiple images on disk for selective recovery to any point-in-time image
- A fully automated protection process, including creation and rotation of mirrors or snapshots and regular backup to tape.
- for recovery, administrator selects a specific recovery image from the graphical user interface



Pros and cons of fast recovery technologies

- Zero-Downtime Backup
  + no impact on application performance
  - requires specific arrays and software

- Instant Recovery
  + recovery of TBs in minutes
  - requires Zero-Downtime Backup as a basis

- Volume ShadowCopy Service (VSS) – see later
  + simple snapshot on any disc
  + Integrated into applications SQL/Exchange
  - supported within Windows 2003 only

To find out more.....

Backup-to-Disk, and Multi-Level Data Protection

Session 1619
Harald Burose
Senior Architect
HP Nearline Storage Division

Session 1597
Glenn Wuenstel
Solutions Systems Engineer
HP Nearline Storage Division

Tools to help you do the job

hp tape performance assessment tools for Windows, Linux, HP-UX, Solaris, AIX , Windows(64)

A

HPTapeperf

PAT

HPCreatedata

HPScandisk

90% of backup performance issues are NOT the tape drive itself

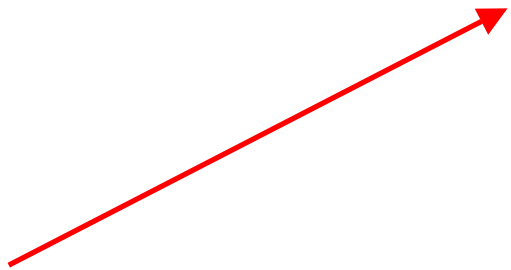
Free download from …http://www.hp.com/support/pat

HPTapePerf

Proves the tape drive is not the bottleneck by writing data from memory direct to tape.

HPScandisc

CreateData – for use in Benchmarking



Can also generate equal distribution of MB

XP Disk Performance Estimator.

For EVA & MSA – try IOmeter

MSA1000
Access pattern: 64 Kilobyte block size, 100% read, 0% Write, 100% sequential, 0% random

RAID Level	I/O per SEC	MB/sec
0	2445	152.8
1+0	2403	150.2
5	2403	150.2
ADG	2407	150.4

There's always performance monitor

Sometime Backup apps have a lot of pre & post processing that disguise the actual tape backup rate.

On windows platforms

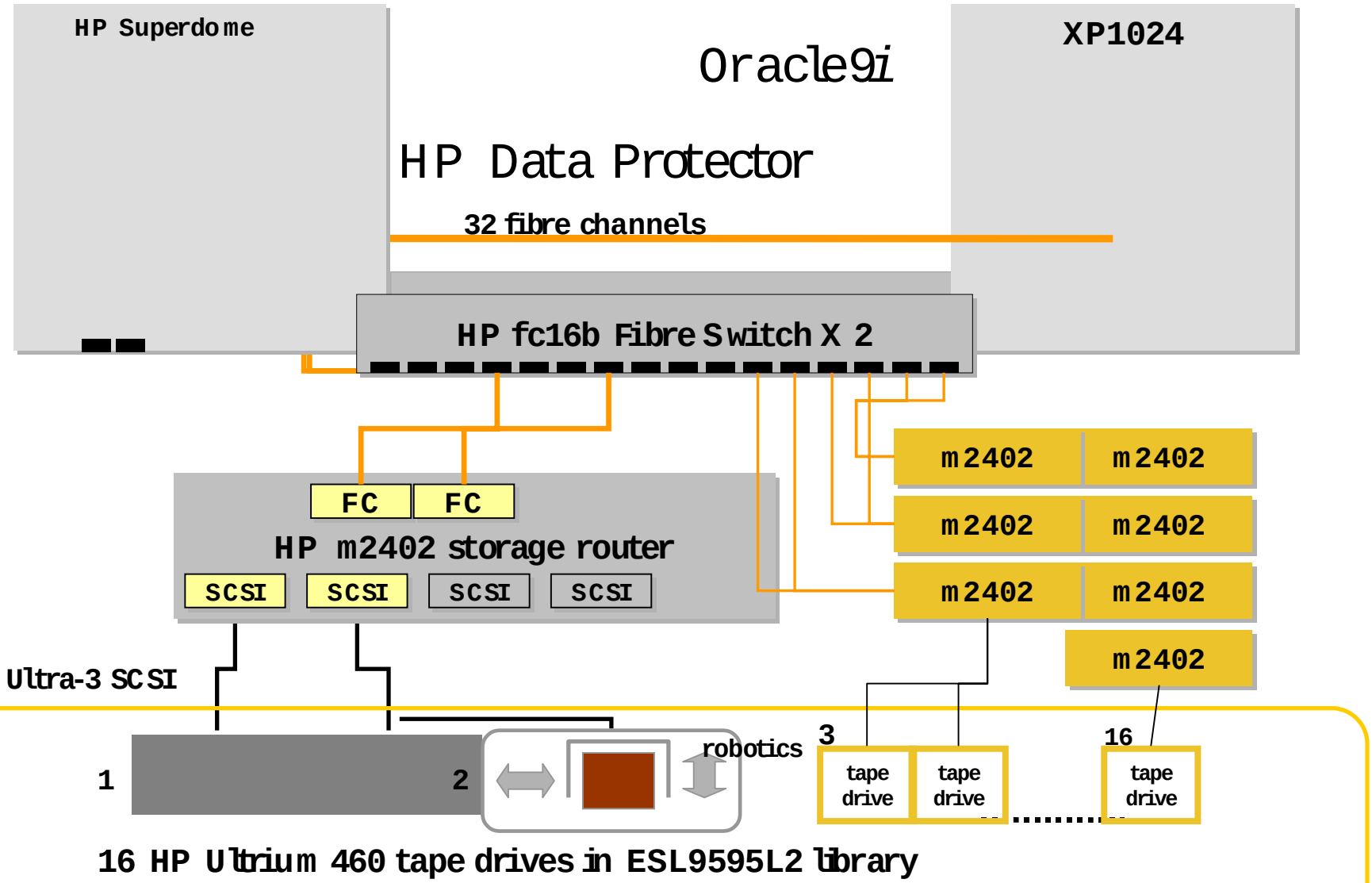
Performance Monitor set to Monitor "system" "File reads" gives a good idea of instantaneous backup rate.

Glance on hp-ux

Now Andy will tell you how it works in the real world

Sample Scenarios

Scenario 1 – XP array – hp-ux - Superdome



Scenario 1 – XP array – hp-ux - Superdome

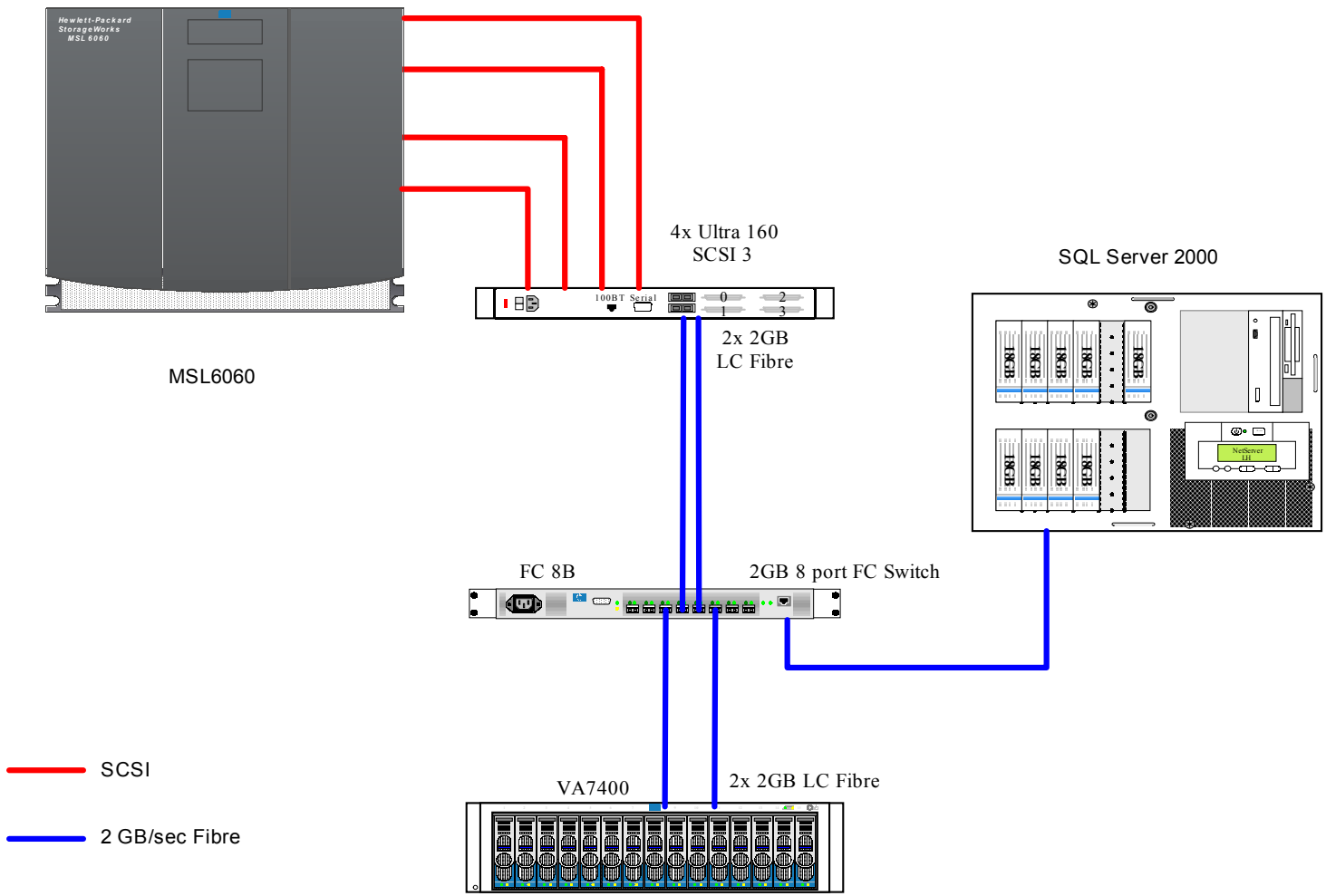
- Results: Backup 3600 GB/Hr (3.6 TB/Hr)

  Restore 1200 GB/Hr (1.2 TB/Hr)

Restore performance subsequently improved to 2.4 TB/Hr

- RAID 5 (3D + 1P) across 384 disks
- 32 Processor Superdome.
- Used Data Protector interface to Oracle RMAN
- Don't use Multiplexing use RMAN tuning
  - Use Backup_disk_ I/O_ slaves = 16
  - 1 RMAN channel per tape (16)
  - MaxopenFiles = 1
- Used 1 router for 2 tape drives since data proved to be 3:1 compressible

Scenario 2 – SQL Server 2000– Data Protector



Scenario 2 – SQL – Data Protector

- Results : 245 GB/Hr Backup (single tape drive)
  90GB/Hr Restore (single tape drive)
- RAID 0+1 (VA7400) disk config across 15 drives
- SQL API is multi-threaded and very efficient at extracting data from the database.
- Use VDI interface so data is buffered only once.
- Configure Data Protector allow large tape transfer size (644K) but only after setting MaximumSGList
- Restore: Pre-zeroing impacts restore performance

Scenario 3 - IMAGE or File-By-File Backup?

DL380 Server w/FC 200 MB/sec

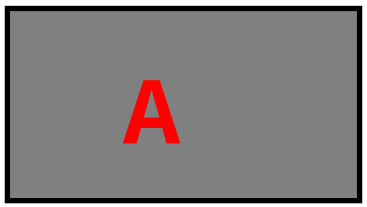


Image Backup of **A** = 40 MB/sec

Image Backup of **B** = 40 MB/sec

File-By-File Backup of **A** = 36 MB/sec

File-By-File Backup of **B** =

Less Than 1 MB/sec

2 x MSA1000 Storage Systems: 6-drive RAID 5 set of 10K RPM drives.

A. BIG FILES = 10 Gig Dataset = 10 x 1 Gig Files

B. LITTLE FILES = 10 Gig Dataset = 10 Million x 1 K Files

Sceanrio 4 – MS Exchange – Veritas Netbackup

- VA7400 – 15 x 18G(15krpm)
- RAID 0+1
- 2GB/sec SAN
- Ultrium 460 in Library
- Using Veritas NetBackup 4.5

- Single Storage Group
- Exchange Data is already pre-compressed
- Exchange Transaction Logs are 5MB and generally < 1.5CR

Scenario 4 – MS Exchange

Results: only 58GB/Hr Backup
(< 25% throughput capability of LTO2)

54GB/Hr Restore

Based on Full information Store Backup

Scenario 4 – MS Exchange – a challenge

Why so Slow?

•MS Exchange Backup API is single threaded

How to Improve

• Backup multiple exchange servers to the same tape

• Consider mailbox backup using "single instance store technology" – Veritas.

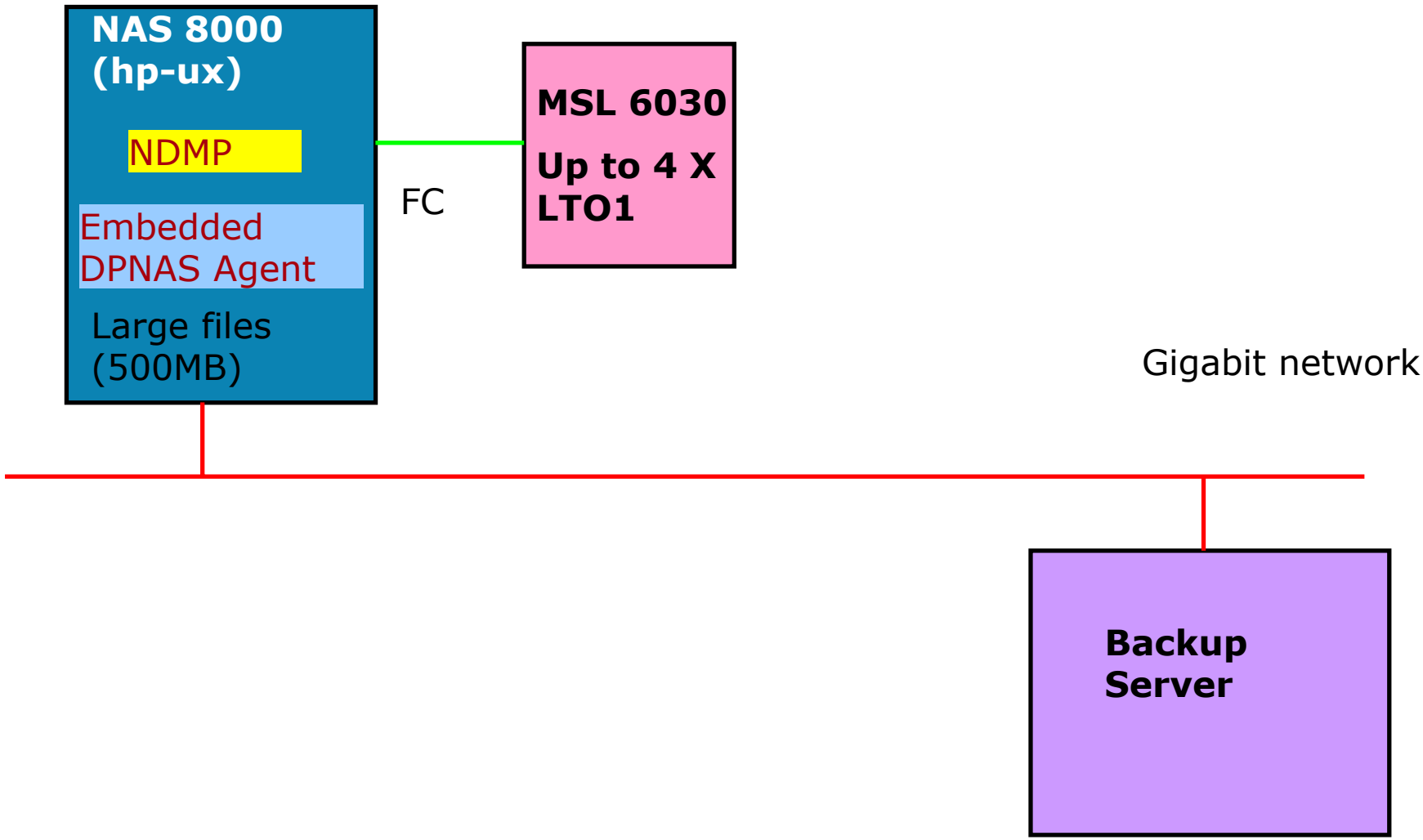
• Consolidation/distributed Email systems.

• Snapshot & Mirror capability in Late 2003

Longer Term

• Exchange will have an SQL based database with multi-pathing (late 2004 or ?)

Scenario 5 - NAS Backup Performance



Results – throughput GB/Hr

	1 drive single stream	4 drives single stream	4 drives 2 streams each	4 drives 4 streams each
NDMP (e.g. Legato or NetBackup)	**42**	**132**	**N/A**	**N/A**
NAS AGENT (Data Protector)	**42**	**132**	**144**	**144***

NDMP can generate multiple data streams from multiple "mount "points but only 1 stream per tape drive.

HP DP NAS Agent allows multiple streams to each tape drive therefore achieving greater throughput.

NDMP (alone) cannot scale

* LTO1 Drive saturated – try Ultrium 460.

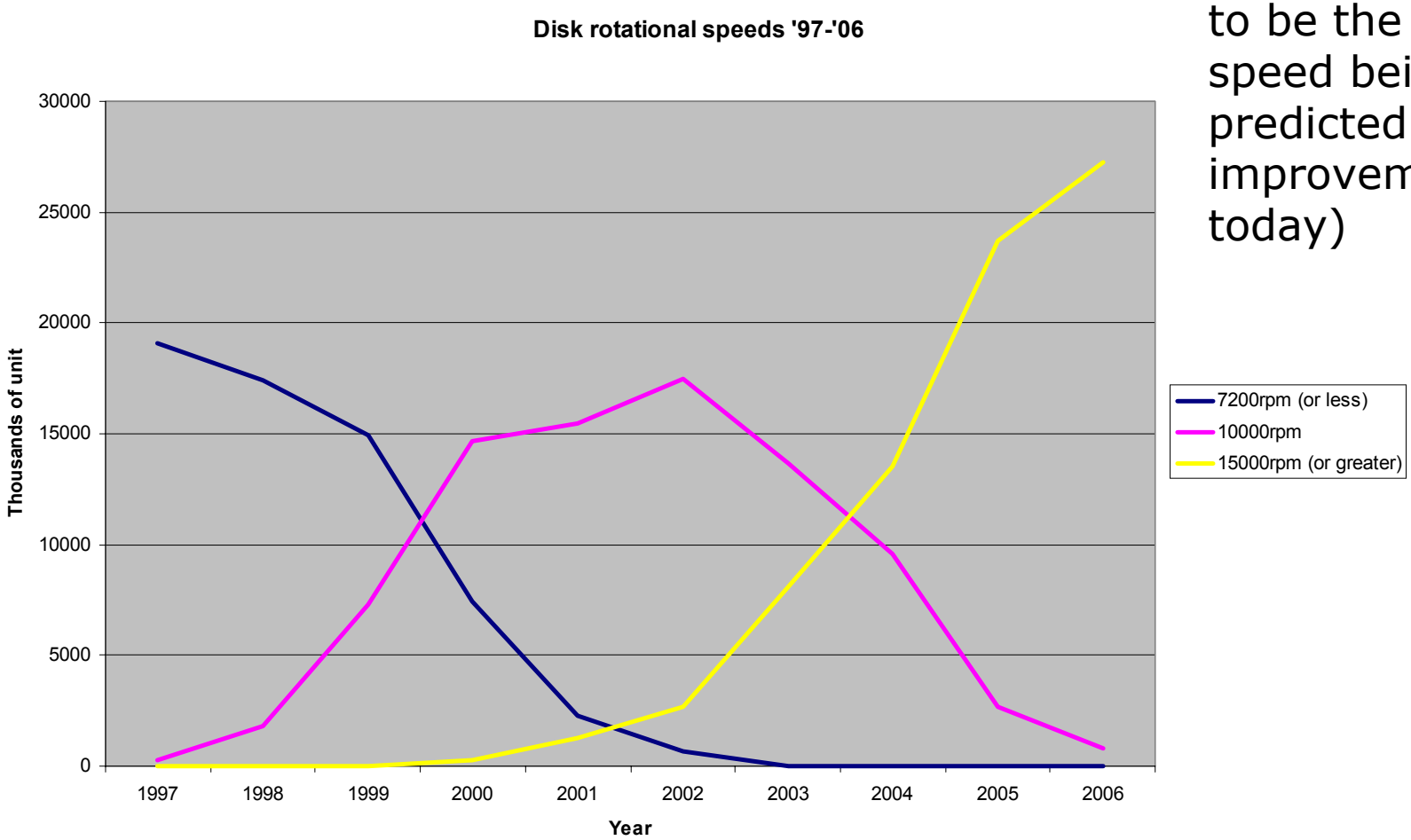
NEW DEVELOPMENTS

Legislation
Disk Performance
Integrity servers
FC vs SCSI tape drives
HP Extended Library architecture
SDLT600 & Ultrium3
Microsoft VSS
File System developments

New Data Protection Legislation Sarbanes/Oxley and HIPAA

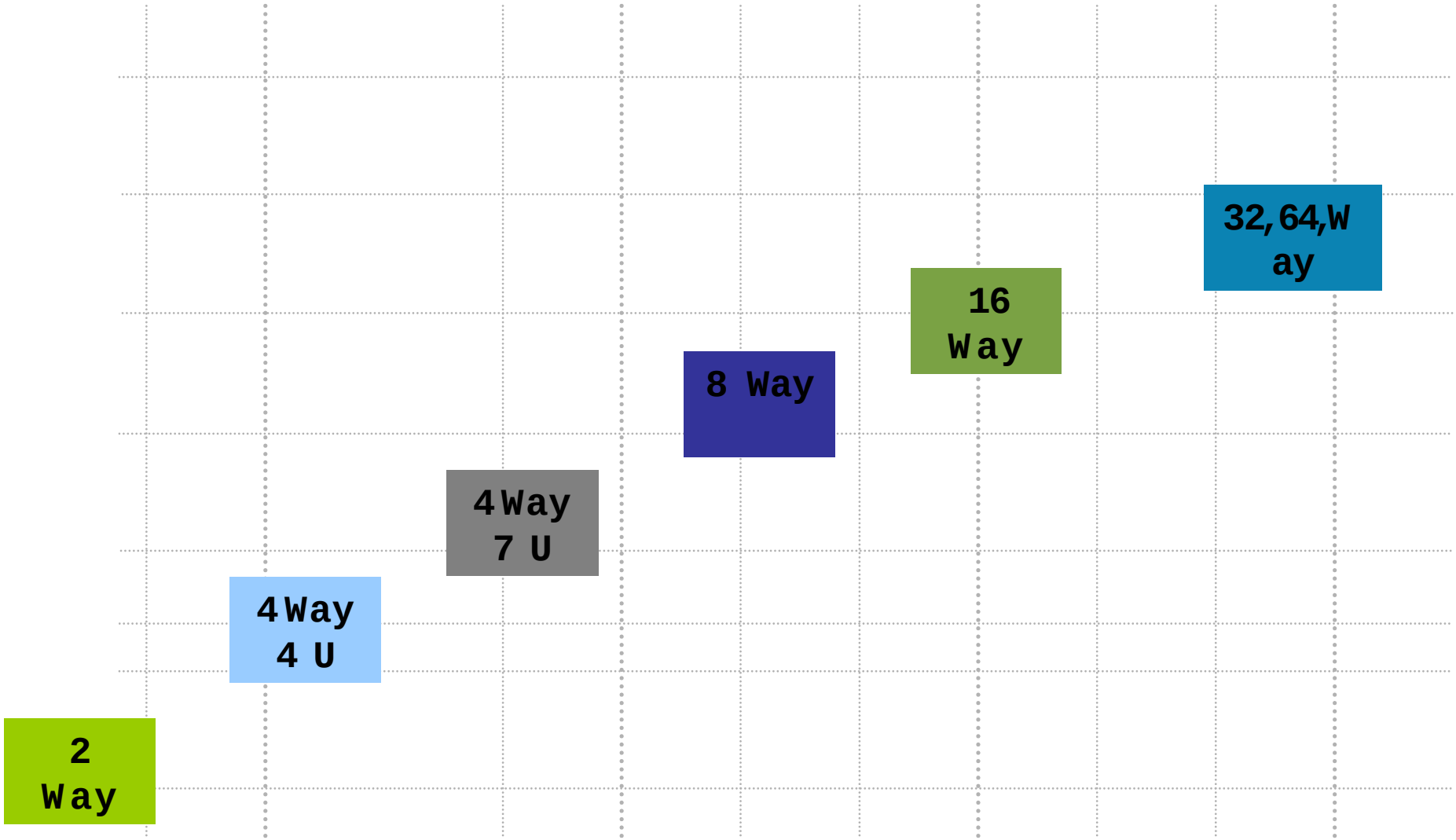
- The Sarbanes/Oxley act was enacted in July 2002, with compliance required on October 31, 2003. The act, which applies to all public companies of any size in the United States, mandates data retention policies and criminalizes tampering with or destroying corporate financial records, even prior to a subpoena.
- HIPAA was enacted in 1996, with compliance scheduled for 2002-2005. This act requires full disclosure to each patient of his or her data. All of the data. The challenge, especially for large organizations: how to find ALL of the patient records.

New Developments – Disk Trends



22K rpm appears to be the highest speed being predicted (50% improvement on today)

HP Integrity Server Roadmap



HP Nearline – Product support on Itanium

- Hardware compatibility is available NOW
- Most Backup Software ISVs are adding compatibility for Itanium based Servers as Media Servers by end of 03
- Currently Data Protector 5.1 (Windows & hp-ux) and Legato Networker 7.0 (Linux) are certified on Itanium based servers

What difference will Itanium 64 bit servers make to Backup?

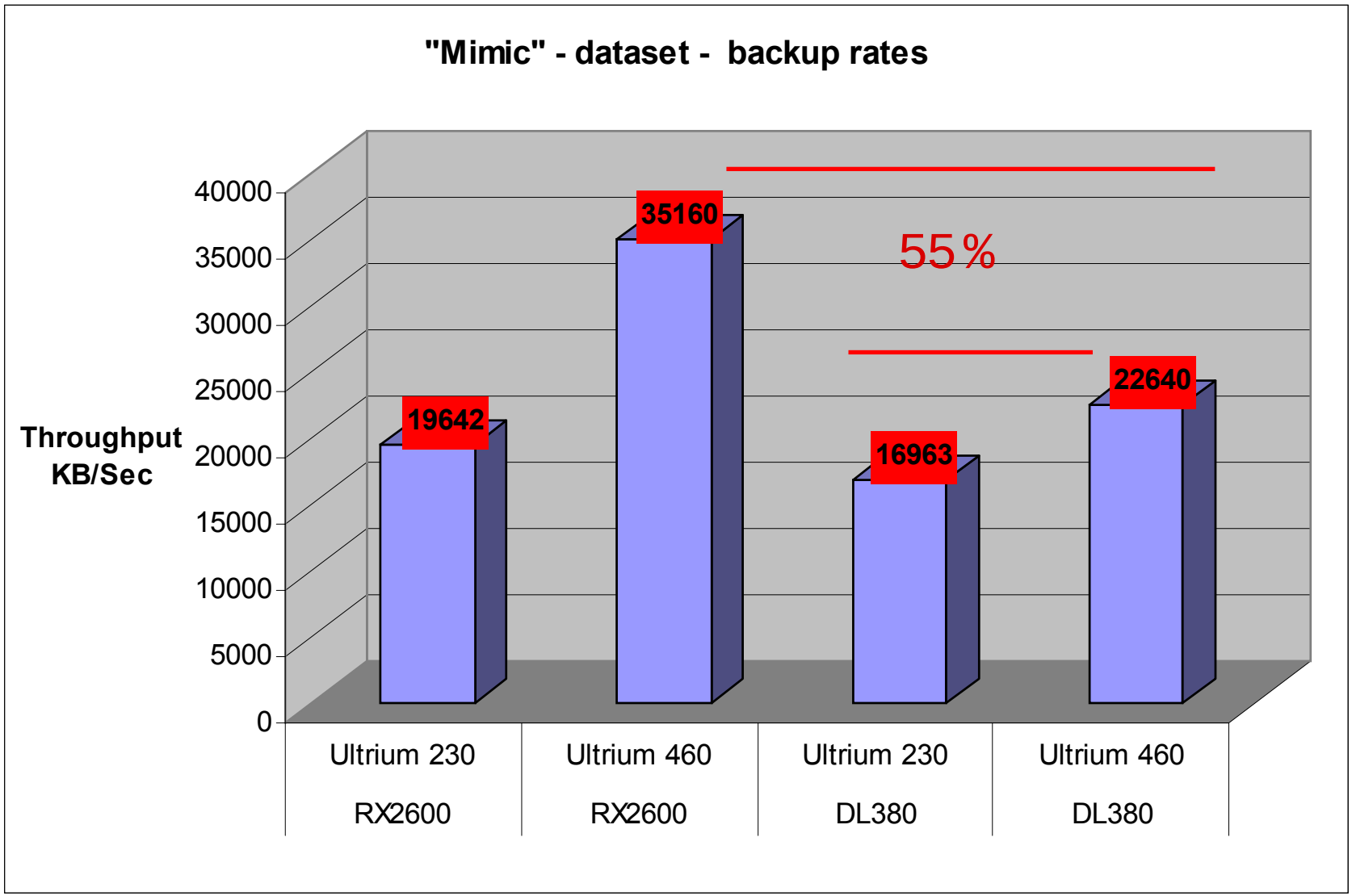
- Higher processing power
- Able to handle increased PCI-X & PCI-X 2.0 I/O throughput
- Increased Memory addressing range
- Improved memory management (Caching)
- Reduced buffer copy times
- Improved driver efficiencies

Itanium vs Pentium Shoot - out

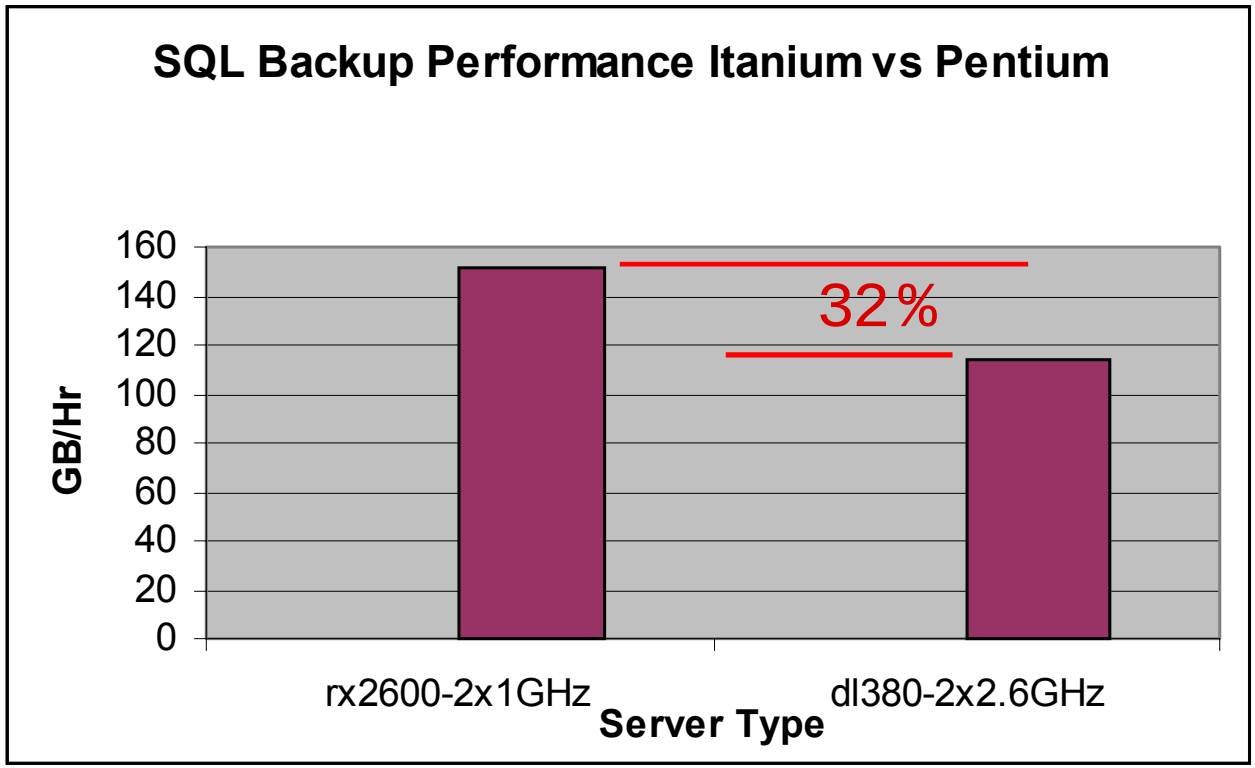
- Identical data on identical Logical Volumes of MSA100 – "typical" sample size distribution. (1K to 10MB)
- W2003(64) on rx2600
- W2000(32) on DL380
- HP Data Protector 5.1

Results – Itanium up to 55% faster

A



SQL 2000(32) vs SQL2000(64)



- Identical HPNorthwind databases on MSA1000
- DL380 – W2K+SQL(32)
- rx2600 – W2003(64) with SQL(64)
- 68K Block size
- Fast Direct mode
- HP Data Protector 5.1

Native FC vs SCSI Tape drives in Libraries

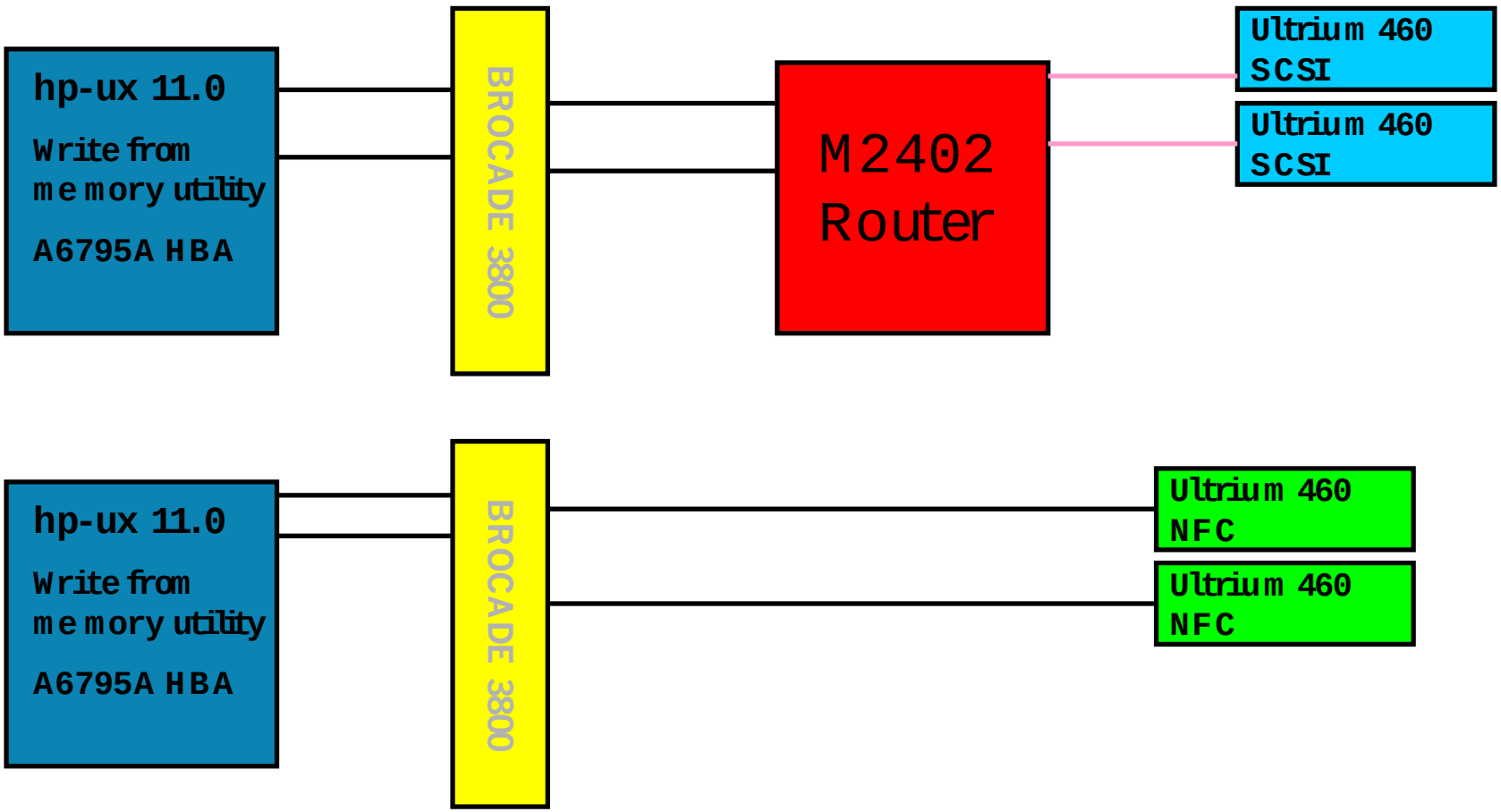
Are Native FC drives faster than SCSI tape drives?

NO!

- in linear tape drives speed is ultimately determine by speed of tape passing by the R/W head and the # of R/W channels

- Native FC drives in Libraries offer increased functionality not increased performance.

TEST SET-UP


hp-ux 11.0
Write from memory utility
A6795A HBA
BROCADE 3800
M2402 Router
Ultrium 460 SCSI
Ultrium 460 SCSI
hp-ux 11.0
Write from memory utility
A6795A HBA
BROCADE 3800
Ultrium 460 NFC
Ultrium 460 NFC

Results

- 64K transfers of 2:1 data.

Drive Type	1 Drive 1 FC HBA	2 Drives 1 FC HBA	2 Drives 2 FC HBA
Native FC	58.8 MB/s	106.7 MB/s	117.6 MB/s
	1 Drive 1 FC HBA 1 Router Port	2 Drives 1 FC HBA 1 Router Port	2 Drives 2 FC HBA 2 Router Ports
SCSI	57.8 MB/s	105 MB/s	115.7 MB/s

..and now
Charles will tell
you about….

HP Extended Library Architecture and it's implications for performance.

FC Library implementation performance option 1

Native FC drives in the library

Drives are directly connected to a FC switch

- Inoperability issues
- Lower performance per FC port
- Additional port cost for drive
- Complex management and maintenance

FC Library implementation performance option 2

FC embedded libraries

Tape Library

SCSI drives, with FC to SCSI tape controller in the library box

- Intelligent controller for better interoperability
- Better bandwidth performance through aggregation in the controller
- Enhanced manageability
- Direct FC connection to fabric with less switch ports for lower cost

FC Switch Ports
2 to 8

The hp StorageWorks extended tape library architecture

Interface Controllers

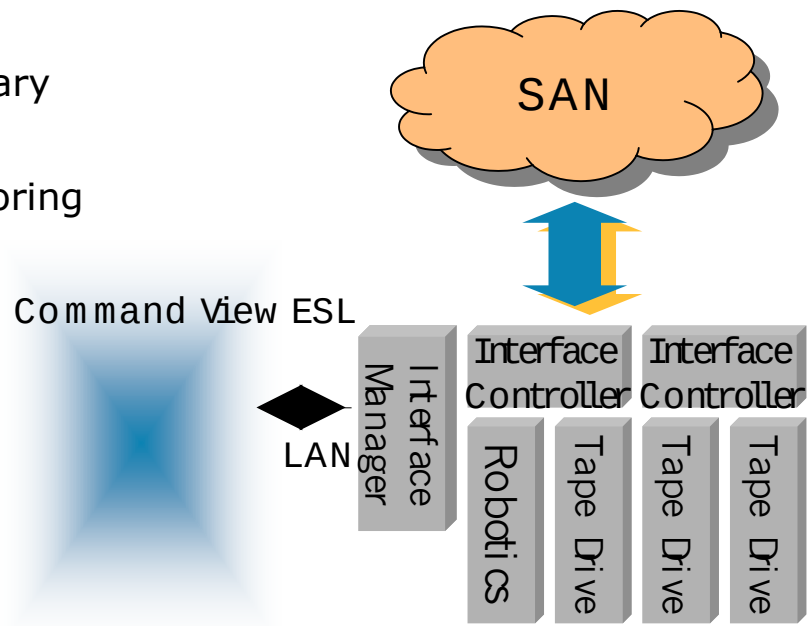
- Layer of intelligence between tape drives and the SAN
- Manages shared access to the tape library, intelligently handling conflicts and storage network events
- Similar architecture to disk-arrays with controllers in front of disk drives

Interface Manager

- Extends the intelligent management
- A central point of knowledge for the entire tape library subsystem
- Enables remote management & performance monitoring

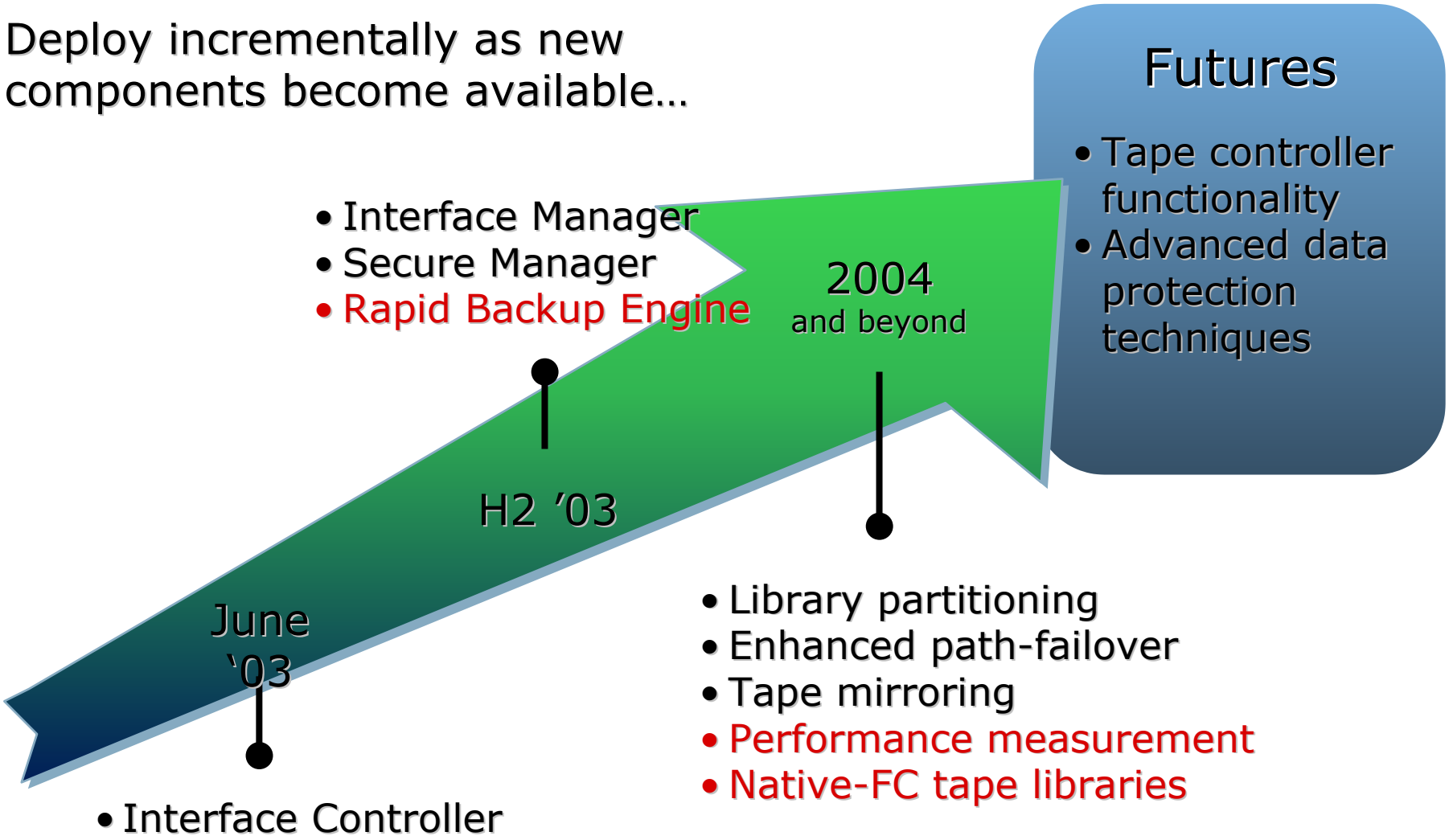
Command View ESL

- A single pane of glass view of the entire library
- Delivers easy-to-use remote management
- Provides wizards for setup and configuration
- Simplifies and automates the most complex tasks
- Stays *out* of the SAN to allow critical traffic to flow



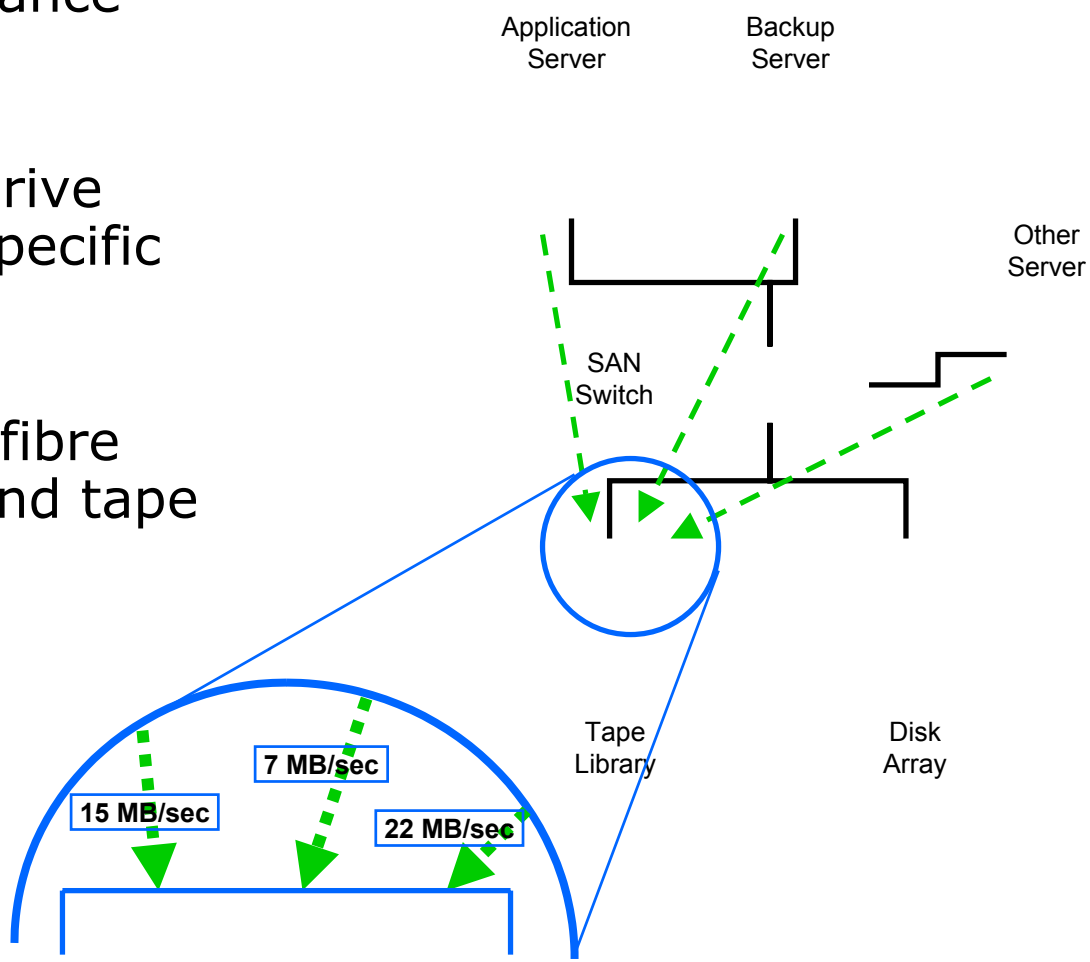
Advanced features and extended architecture timeline

Deploy incrementally as new components become available…



Performance Measurement

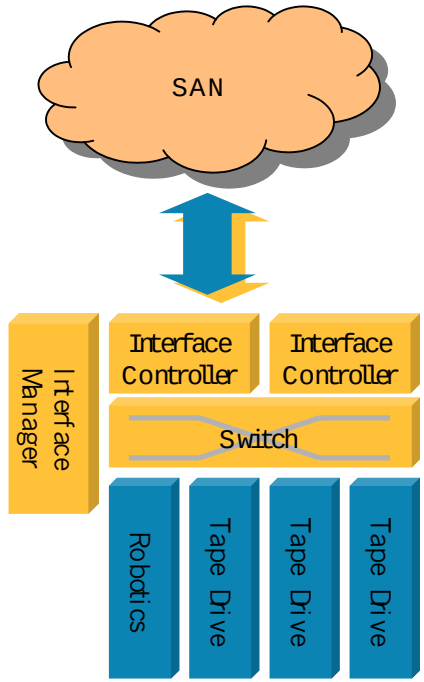
- Analyzes and identifies backup and restore performance bottlenecks
- Tracked to a host+drive combination, for a specific backup job
- Data collected from fibre channel interfaces and tape drives



Extended library architecture - Beyond 2003

Native-FC tape libraries

- Includes native-FC tape drives in future tape libraries
- Offers better flexibility and FC bandwidth utilization
- Enables advanced backup techniques and more functionality



Tape controller functionality

- Moves tape command processing into the Interface Controller
- Greatly simplified shared access model similar to disk arrays
- Further isolates hosts providing more robust SAN behavior
- Enables advanced backup techniques and functionality
  - Virtual tape emulate to disk
  - Tape mirroring

Best of both worlds – Future Option

Next generation native FC library

FC Drives are directly connected to an intelligent controller

- Intelligent FC controller for better interoperability and future enhancements
- Better performance through aggregation in the controller
- Enhanced manageability
- Direct FC connection to fabric with less switch ports
- Dual ported FC Drives with path failover

FC Switch ports
2-8

New drive release - SDLT roadmap

SDLT 600

- Due Early 2004
- 300 GB Native
- 32 + MB/sec Native
- BRC (Read only) for SDLT 220 & 320
- Media – SDLT 2 (new)
- Interfaces
  - Ultra 160
  - 2Gb/sec Fibre for Automation Drive

New drive releases - LTO roadmap

WHAT	availability	today	2003	2005	2007
	capacity	100GB	200GB	400GB	800GB
	transfer rate	10-20MB	20-40MB/s	40-80MB/s	80-160MB/s
HOW	media type	MP	MP	MP	MP
	encoding scheme	RLL 1,7	PRML	PRML	PRML
	tape speed	2.7- 5.4m/s	3.8-7.5m/s	3.8-7.5m/s	3.8-7.5m/s
	tape length	580m	580m	800m	800m
	data tracks	384	512	768	1024
	data channels	8	8	16	16
		GEN 1	GEN 2	GEN 3	GEN 4

HP Ultrium 3

- Due Early 2005
- 400GB native
- 60MB/sec
- New media
- Reads Gen 1,2,3 Writes Gen 3 and Gen 2
- Interfaces
  - Ultra 320 SCSI
  - 4GB Fibre for automation drives
  - SAS 3Gb/sec

..almost there – this next bit is difficult so – over to Andy.

Microsoft Volume Shadow Copy Service – what it means to tape.

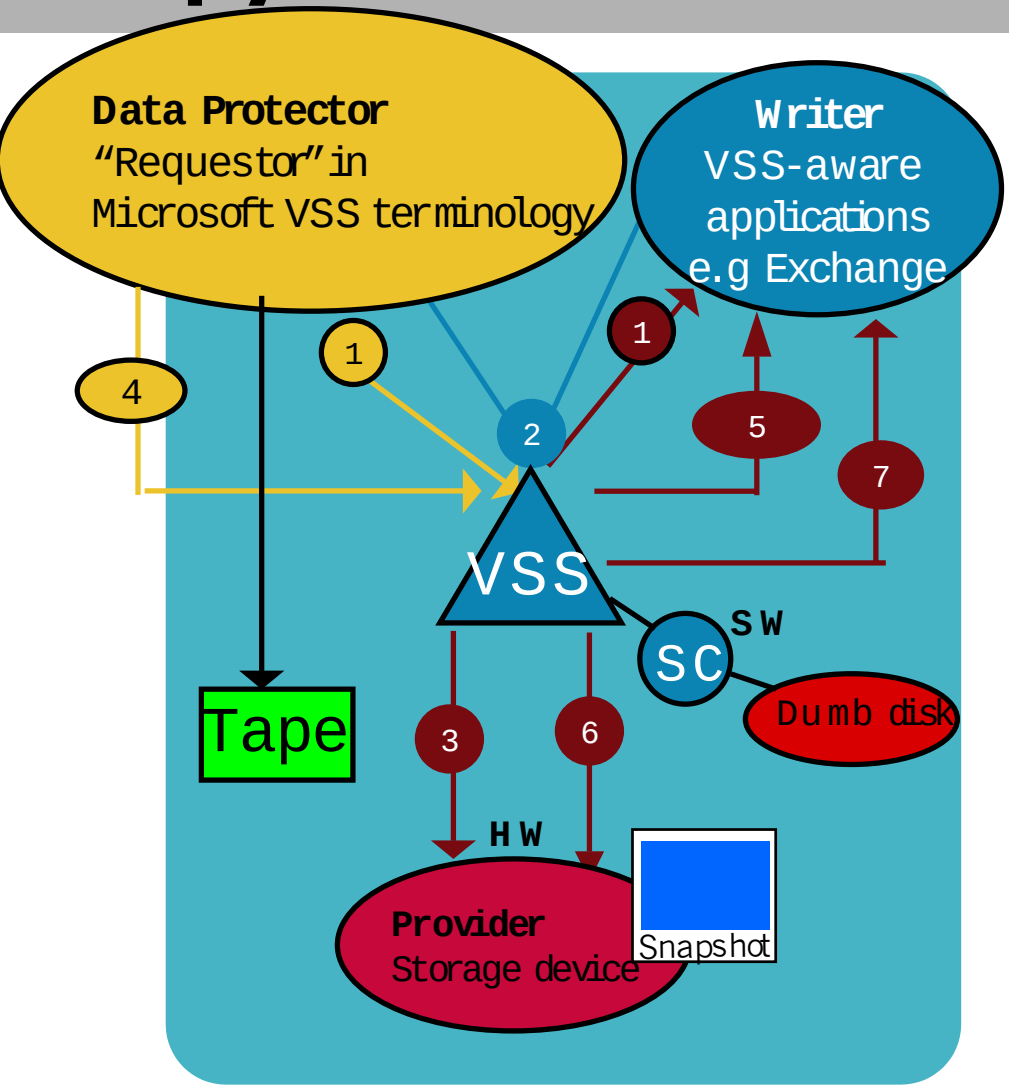
Volume Shadow Copy Service - intro

- Confusing naming
  - Shadow copy service - VSS
  - Shadow copy – hardware API support - Arrays
  - Shadow copy – software support – dumb disks
  - Shadow copies – enabled at drive level for file protection
- Primarily designed as a preparation tool for tape backup – will eventually replace 47 current backup API's with a single API.
- No concept of VSS restore – restore procedure is embedded in meta data at time snapshot is prepared.

Volume Shadow Copy Service – intro (cont)

- Provides open file backup capabilities.
- In general will reduce the impact of backup on application performance
- Currently some limitations with SQL database restores
- Will offer snapshot/mirror capability for the first time to Exchange 2003 (Oct?) – supported by HP Data Protector patch.

Microsoft's Volume Shadow Copy Service:



- VSS allows
  - to create shadow copy backups of volumes
  - coordination of providers, writers and requestors
- HP Data Protector fully integrated with VSS - providing:
  - backup of open files, databases & applications
  - consistency and integrity of backed up volumes

If the file system is the bottleneck – why don't we improve the file system performance?

A: It is not that easy.

Journaling file systems in Unix have helped

Rumour has it Microsoft's SQL based filesystem has been significantly delayed.

Grid computing and distributed file systems may be the next step.

Distributed File Systems – good or bad for tape backup & restore?

Today

- NFS(unix)
- CIFS (windows)
- NDMP (File serving devices)

Near Future

Distributed File Systems (clients connect via protocol)

- HP backing Lustre (www.lustre.org)
- Microsoft File system unlikely to change in the mid term.

Backup & Restore Issues

In theory backup speed could improve since many separate (faster) sources of the filesystem now exist, however restore will be far more complicated.

Resources

- Whitepapers

http://h18006.www1.hp.com/storage/tapewhitepapers.html

- Performance Tool downloads

http://www.hp.com/go/pat

- IOmeter

http://www.edelbyte.org/misc/iometer/

- Stand alone tape drive connectivity

http://www.hp.com/go/connect

- Enterprise Backup Solutions Guide

http://www.hp.com/go/ebs

Q&A

- Any Questions?

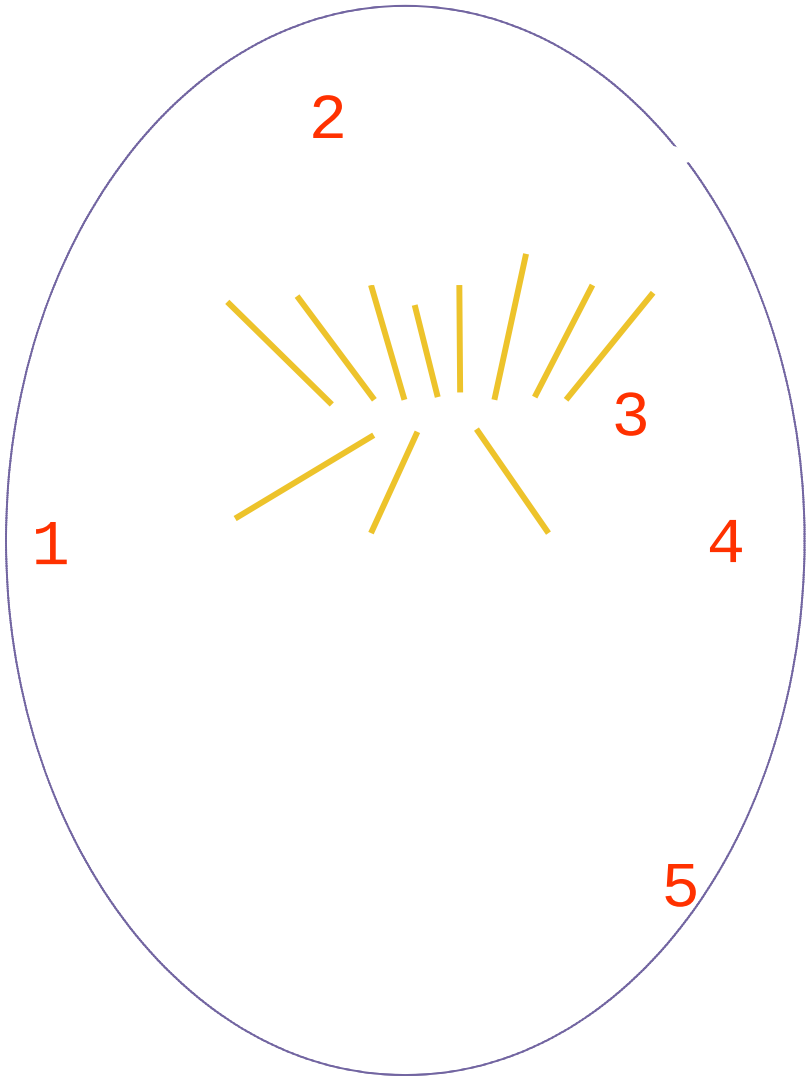
Interex, Encompass and HP bring you a powerful new HP World.

Supplementary Slides

Glossary of Terms

- Multiplex
- Multi-Stream
- Adaptive Tape Speed
- Backup Server
- ISV tuning parameters
- Useable Bandwidth
- Block Size
- Transfer Size/Block Size
- API
- Compressibility
- Rapid Backup/X Copy/Direct/Serverless
- Agent
- Volume Shadow Copy Service
- Mirrors/Split Mirrors.
- Information LifeCycle Management
- Zero DownTime Backup (ZDB)

5 building blocks to performance.



5 Easy Pieces to SAN Backup Tuning

1. A Source of Data
2. A Backup Server System
3. A Connection
4. A Connection
5. The Tape Drive

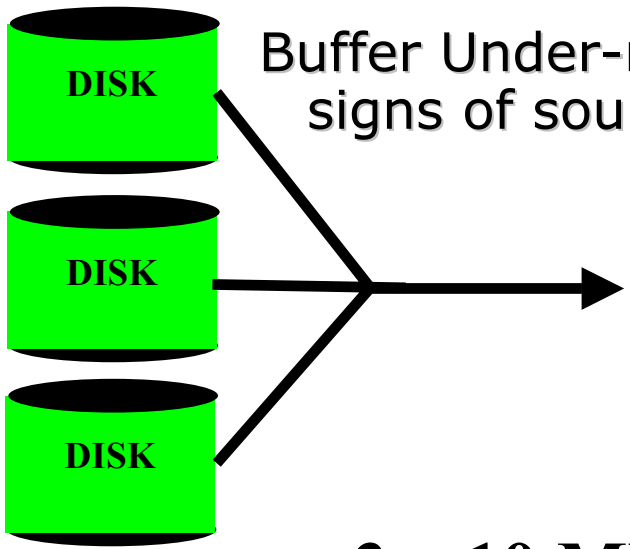
The point?

All are equal possible causes of performance issues!

Primary Storage Performance – the 3:1 rule

- SOURCE TO TARGET SPEED - You usually need 2 to 3 times the source speed as it compares to your desired backup speed.

Tape is a typically a streaming device.



Buffer Under-runs , Shoe Shining, Back Hitching and other signs of source performance issues.

Tape

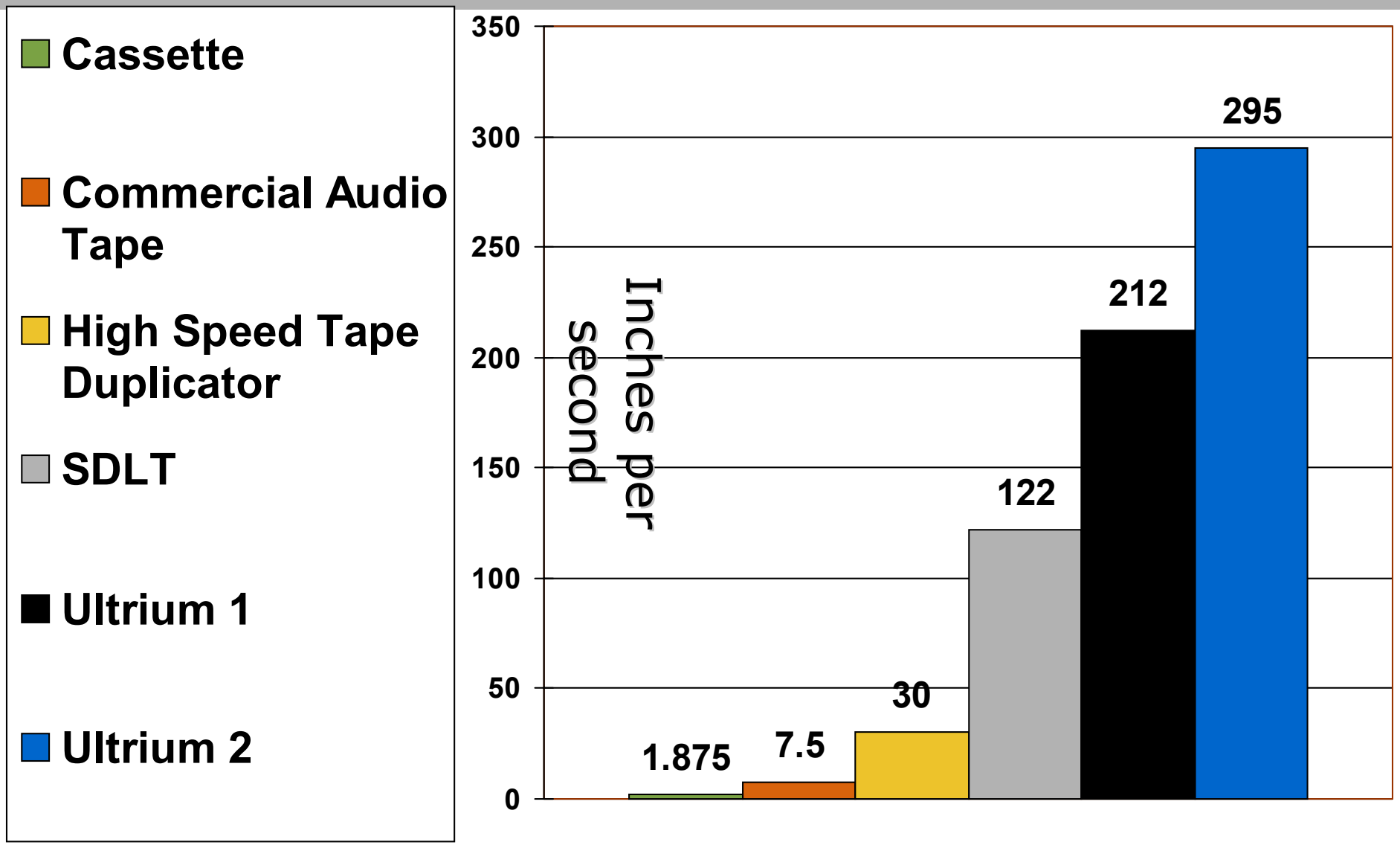
3 x 10 MBsec Hard-Drives to feed a single 10 Mbsec Tape Drive

The SAN Switch (continued)

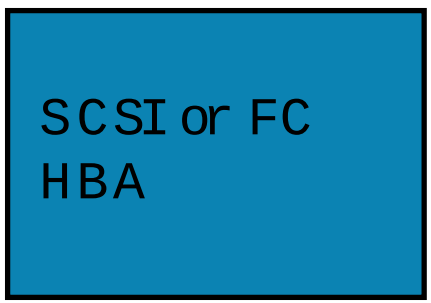
- **SPEED of a Fibre Channel port:**
  - **1.063 Gbit/Sec full duplex (100+ MBytes/Sec)**
  - **2.125 Gbit/Sec full duplex (200+ Mbytes/Sec)**

  Full duplex – an FC adapter can do 2X the above speeds if the adapter card is both transmitting AND receiving at the same time.

How FAST is the TAPE MOVING –



Transfer Size – e.g. MaximumSGList



MaximumSGList

Attribute
of the
HBA

ISV performance tuning – database API's - Exchange

- No inherent tuning parameters available for Exchange, other than basic buffer size and # of disk buffers.
- Exchange API is renowned for being a tape performance bottlenec

Example: Veritas NetBackup

C:\Veritas\Netbackup\db\config\

Two files must be created to override the default NetBackup settings:

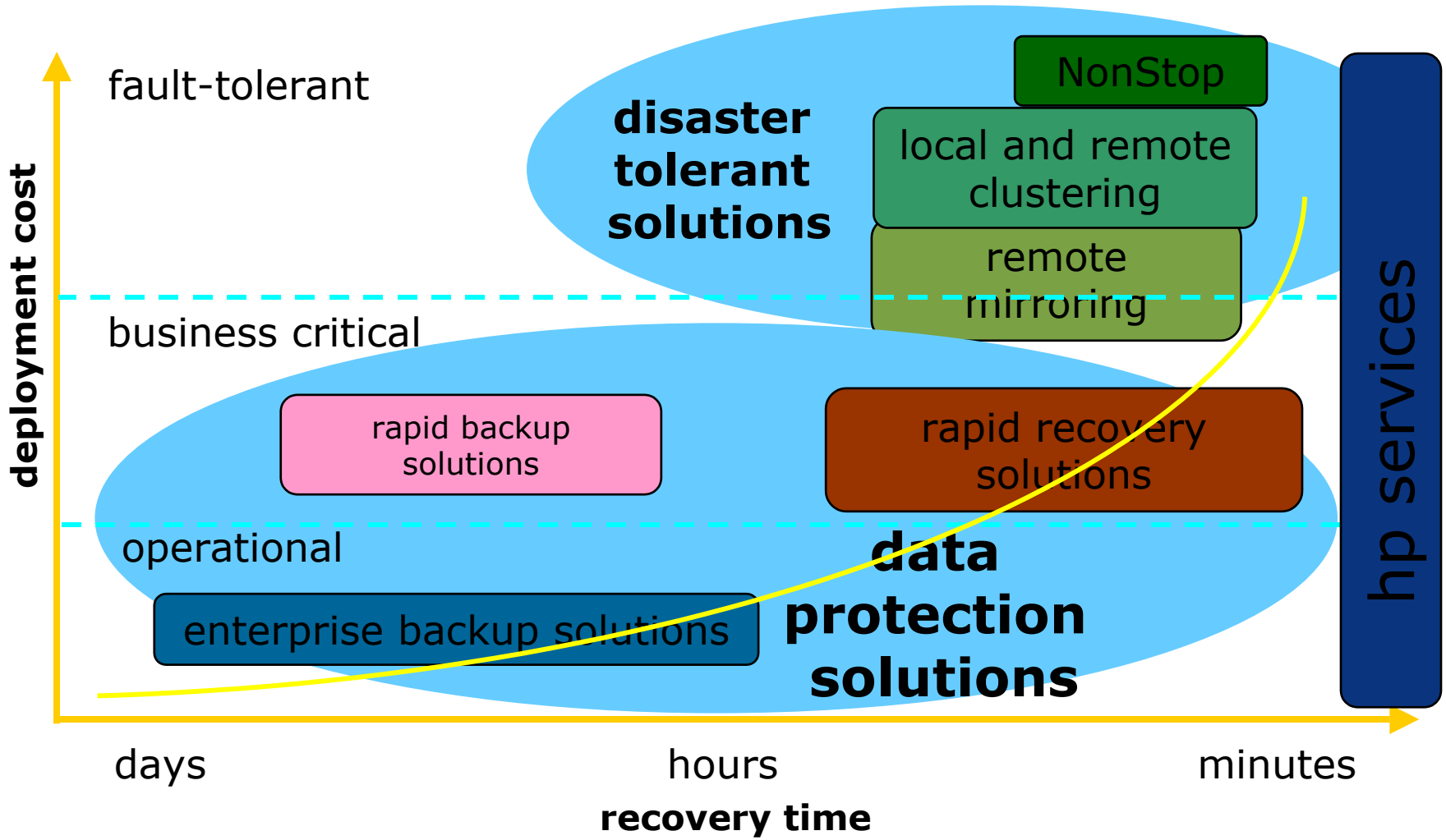
NUMBER_OF_BUFFERS (64)

BUFFER_SIZE (64)

ISV features relating to Backup performance

Feature	HP Data Protector	CA BrightStore Enterprise	Veritas NetBackup	Legato Networker
Backup to Disk (fast restore)	✔	✔	✔	✔
3PC/Serverless XCopy/Rapid Backup	✔ (oracle 9i with hp-ux)	✔ (windows & solaris)	✔ (oracle 9i with hp-ux)	✔ (oracle 9i with Network appliance & EMC hw)
Image backup	**X (raw disk via Oracle RMAN)**	✔ **(Windows only)**	✔ (Solaris & hp-ux **)**	✔ **(Windows & unix)**
Parallel Streaming	✔	✔	✔	✔
Concurrency (multiplexing)	✔ (32)	✔ (32)	✔ (32)	✔ (32)

HP: A Range Of Business Continuity Solutions



Comparing Data Protection using Zero DownTime Backup (ZDB) vs. VSS

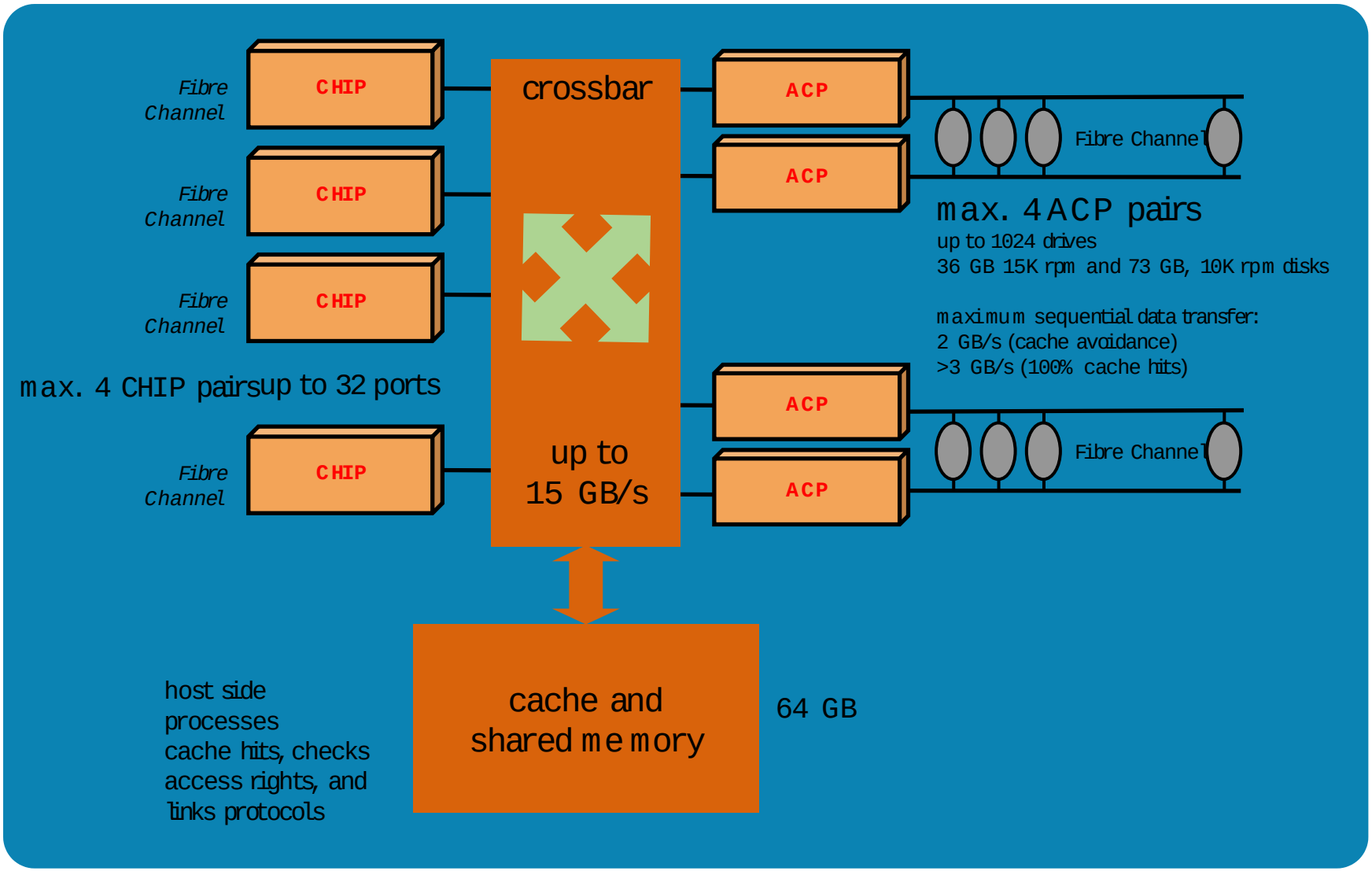
	ZDB	VSS
what it is	array-based data mirroring	host-based data mirroring
disc storage requirements	VA, XP, EVA, EMC Symmetrix, HDS	any disc
OS requirements	HP-UX, Windows 2003, Windows NT, SUN Solaris	Windows 2003 only
software requirements	Business Copy Software	VSS (included in Windows 2003)
applications supported	Oracle, Exchange, SQL, SAP	any VSS compliant application
snapshot mgmt	can manage up to 3 snapshots automatically	any storage supported under VSS (having a VSS snapshot provider)
backup mgmt	backup to tape via a seperate server (application-free backup)	backup to tape can impact performance of application
recovery mgmt	recovery from tape or disc (Instant Recovery)	recovery from tape

RAID 0+1

Drive

Requires minimum of 4 drives

xp1024 architecture

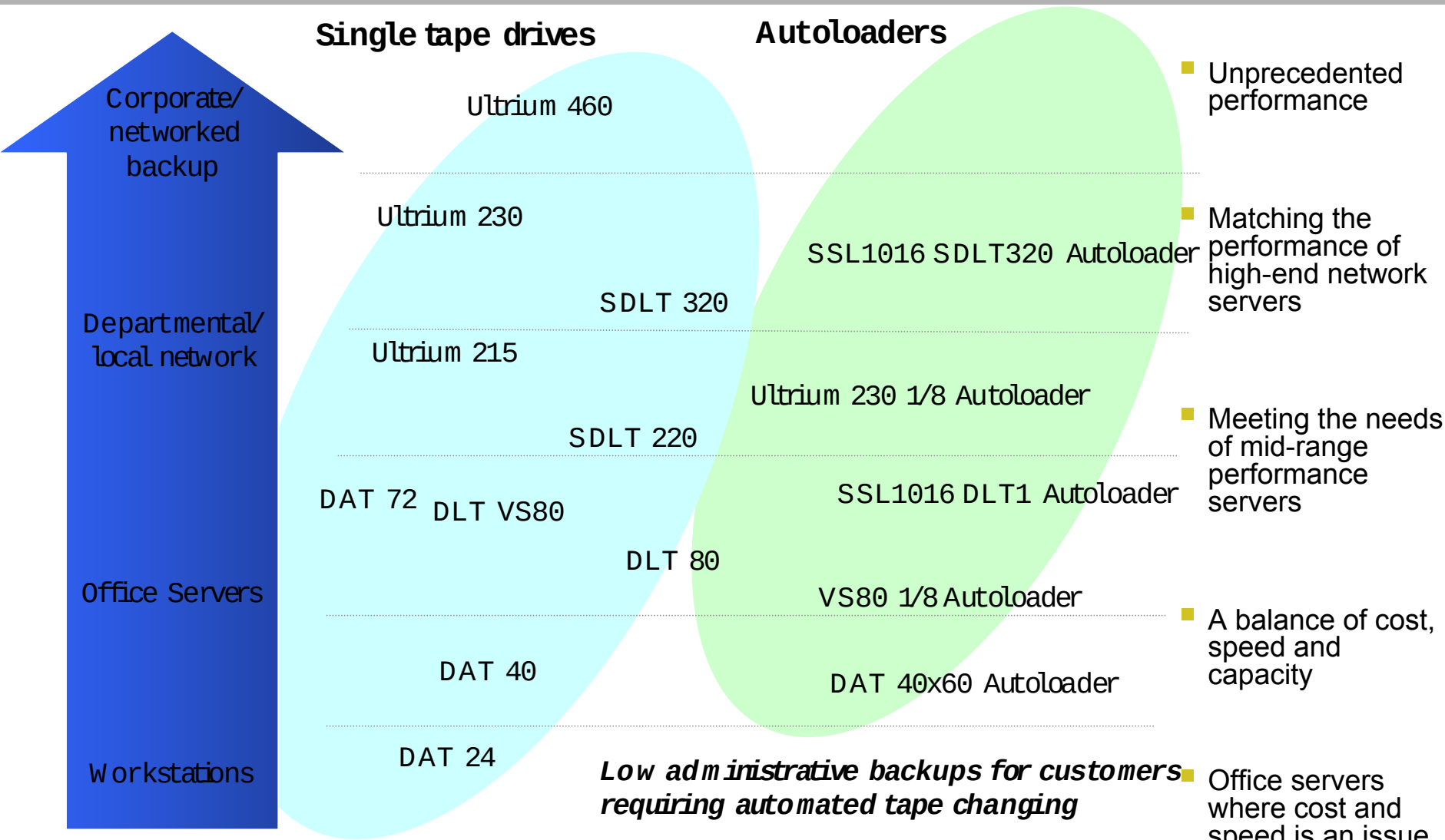


Backup to disk solutions

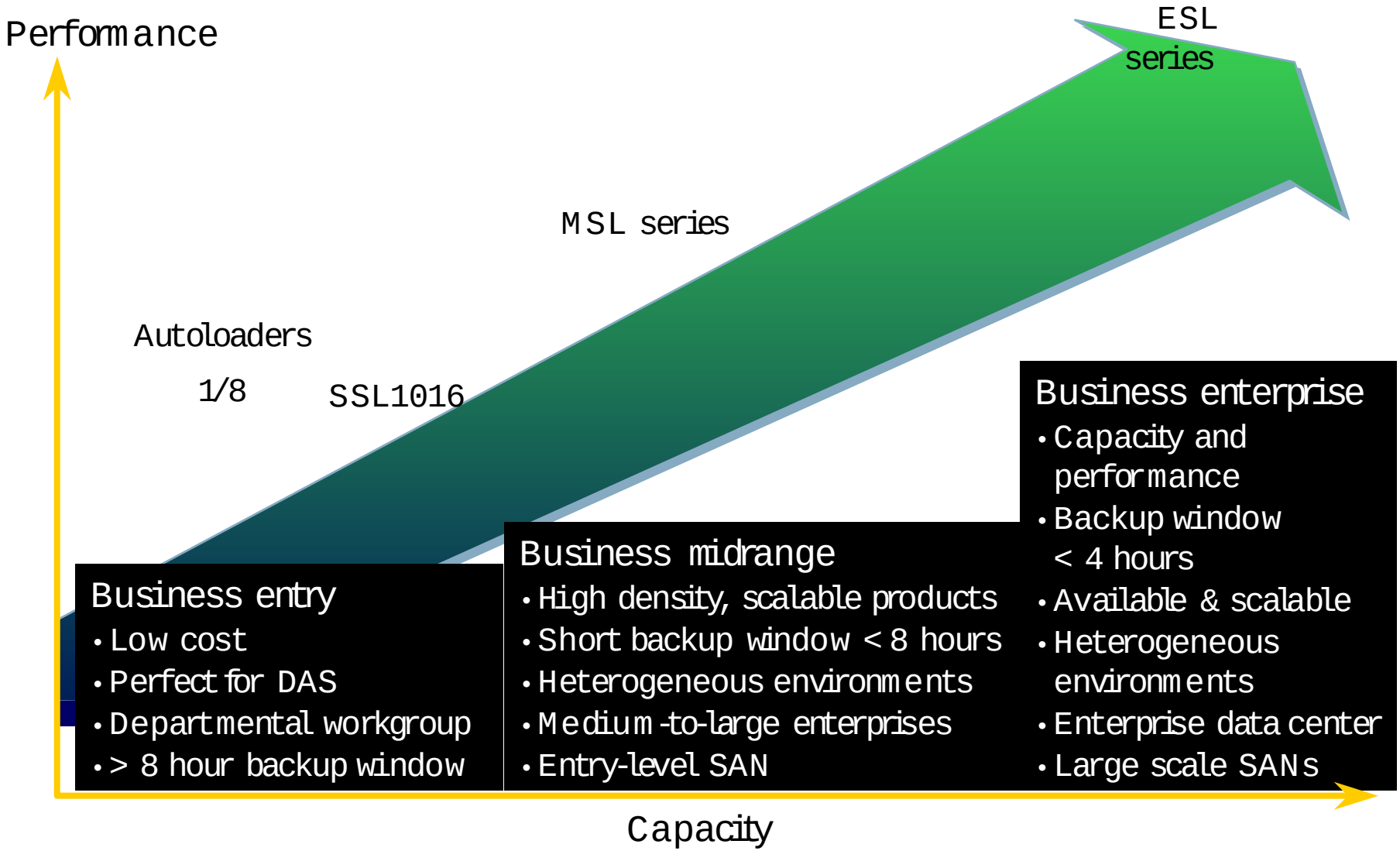
- Tape format on disk
  - Virtual tape
  - Simple backup to disk support in tape backup product
  - Advanced backup to disk support in tape backup products
- Point in Time copies on disk
  - Snapshot & mirroring on different levels
    - FS, VM, SAN, Array
  - Dedicated backup journals
- Compare by Hash technology
  - File by file based (PC backup technology)
  - Network wide (Server backup technology)

Shadow copies enablement for files

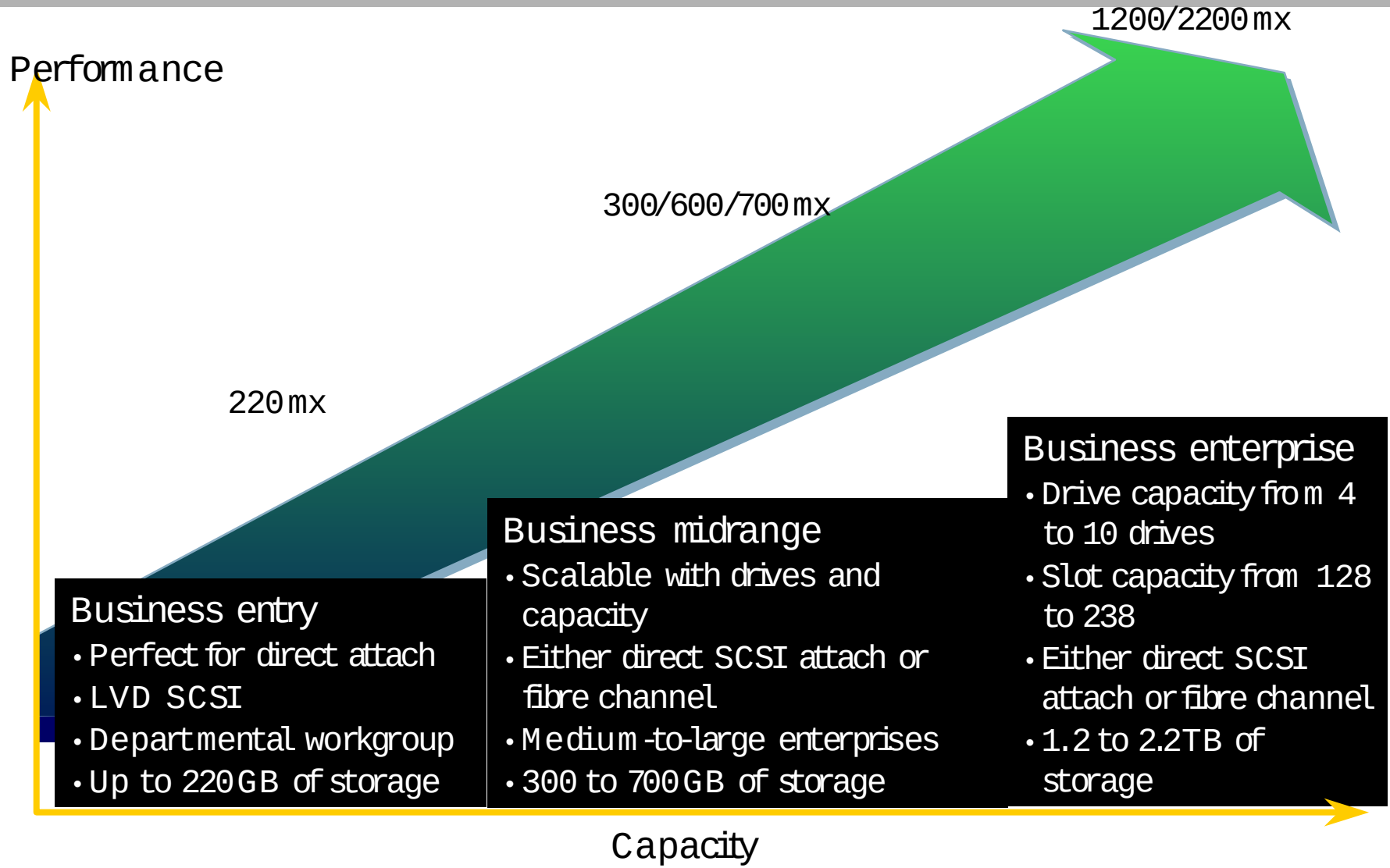
HP Tape Drives and Autoloaders Positioning



HP StorageWorks tape automation positioning



HP StorageWorks optical jukebox positioning



ISV Restore Positioning Command Usage

- Veritas BackupExec
  - Utilises locate block, then read through of data; however for nested directories, will execute a locate block command for each parent directory (e.g. 11 levels deep = 10 locate commands)
- Veritas NetBackup
  - Utilises combination of locate block and space to filemark depending on whether multiplexed backup image. [Filemarks separate multiplexed clients sessions on tape media]
  - Tape format is in 'tar' (NBU essentially GNU tar underneath)
- HP Data Protector
  - Space to filemark and read-through of unwanted data.
  - Tape format uses filemarks to separate catalog and data segments. Defaults to catalog every 2GB, however, if catalog would be bigger than 12MB it's written (e.g. lots of small files may cause more frequent writing of catalog segment).

Application Restore Performance

- Aim:
  - Determine with three major ISV platforms whether there is a fundamental difference in restore performance for a (common) randomly selected set of files out of a reasonably sized backup image
- Methodology:
  - Tests run under Windows 2000 Server, same HBA, cable, disk subsystem
  - [Same media used for LTO drives]
  - Create a dataset on two Cassini (VA7100) FC disk array spindles –
    - Dataset was a total of 70GB (2*35) in size, of varying filesizes from 64KB to 4MB of 2:1 compressible data, with an equal number of files at each size increment
  - Back this up to tape media then measure restore performance of three individual sets out of whole tape media image for each application and drive.
  - Three restores run per application, per drive, to a clean disk array LUN.

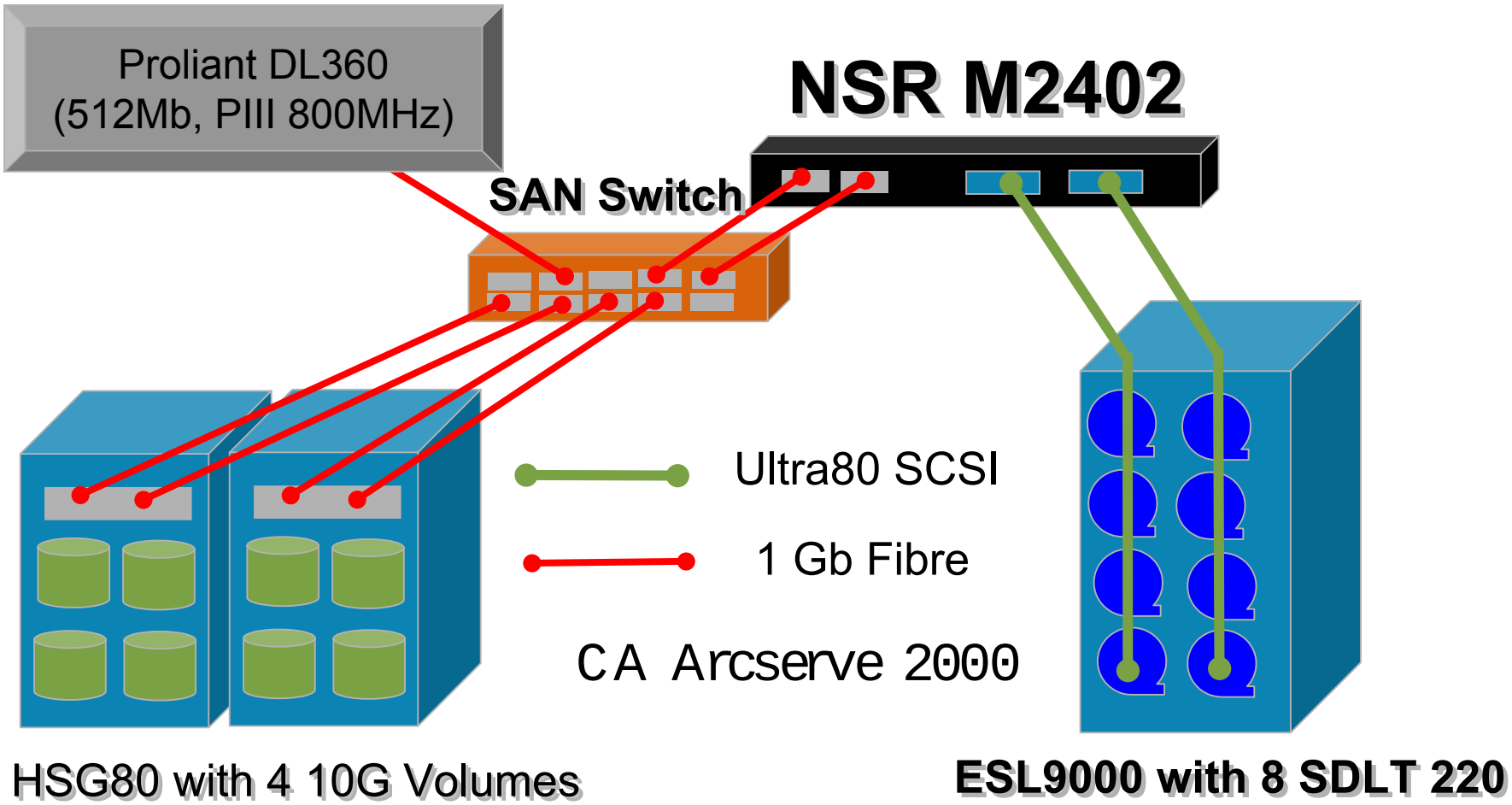
Application restore result format

- Application noted, three lines per application used:
  - 1st line per application:
    All of a sub-tree from first LUN to be restored to a clean LUN (7.2GB)
    Physically located towards beginning of tape media
  - 2nd line per application:
    All of a second sub-tree from second LUN to be restored to a clean LUN (0.23GB)
    Physically located towards middle of media
  - 3rd line per application:
    All of a selection out of one directory from second LUN (c.76MB, mixed file sizes) to a clean LUN (72.5MB)
    Physically located towards end of media
- Green background for fastest drive, red for the slowest of the three.

Application Restore Performance

	Veritas BackupExec	HP DataProtector	CA ArcServe2000
HP SCSI LTO2	7 min 36 sec	4 min 33 sec	5 min 39 sec
	2 min 2 sec	10 sec	2 min 9 sec
	3 min 28 sec	9 sec	3 min 18 sec
IBM SCSI LTO2	8 min 2 sec	6 min 27 sec	6 min 23 sec
	1 min 41 sec	12 sec	50 sec
	2 min 49 sec	8 sec	2 min 23 sec
SDLT320	10min 35 sec	10 min 56 sec	21 min 12 sec
	1 min 28 sec	20 sec	1 min 9 sec
	3 min 13 sec	22 sec	15 min 4 sec

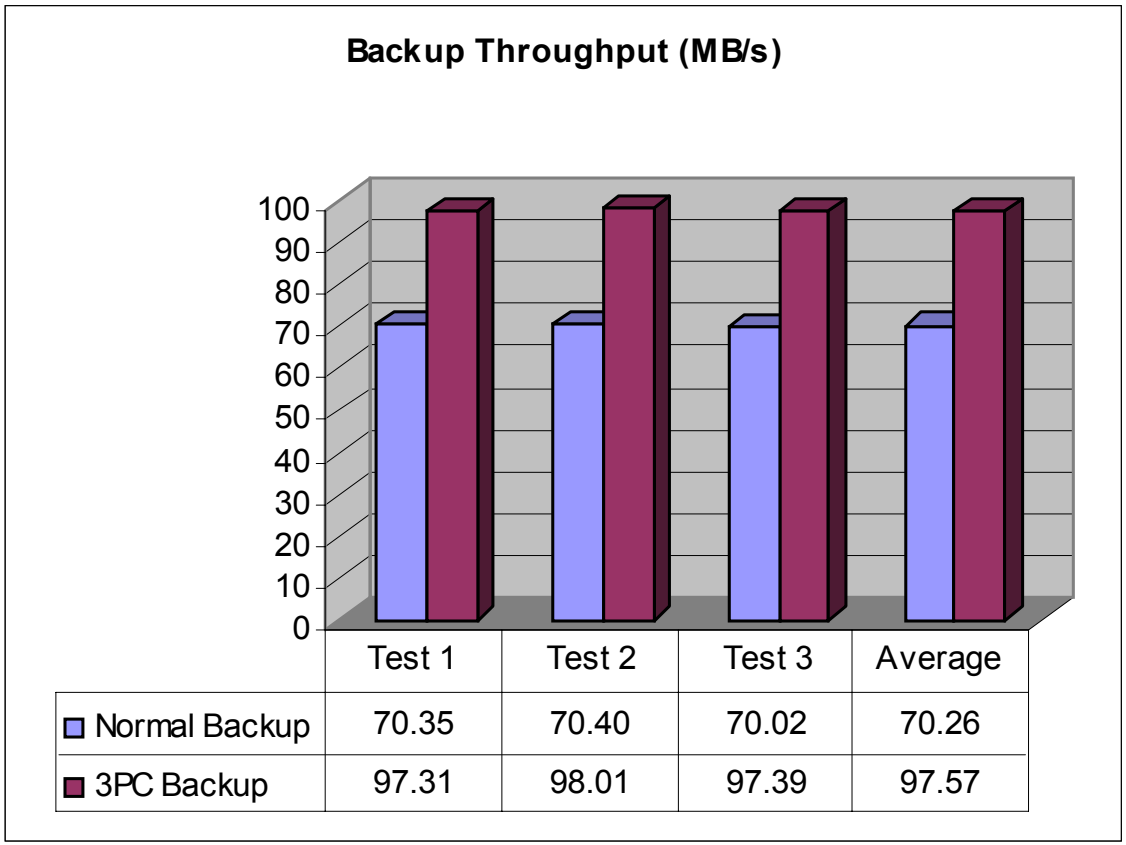
Scenario 3 – Rapid Backup (Serverless)



8 streams simultaneously.

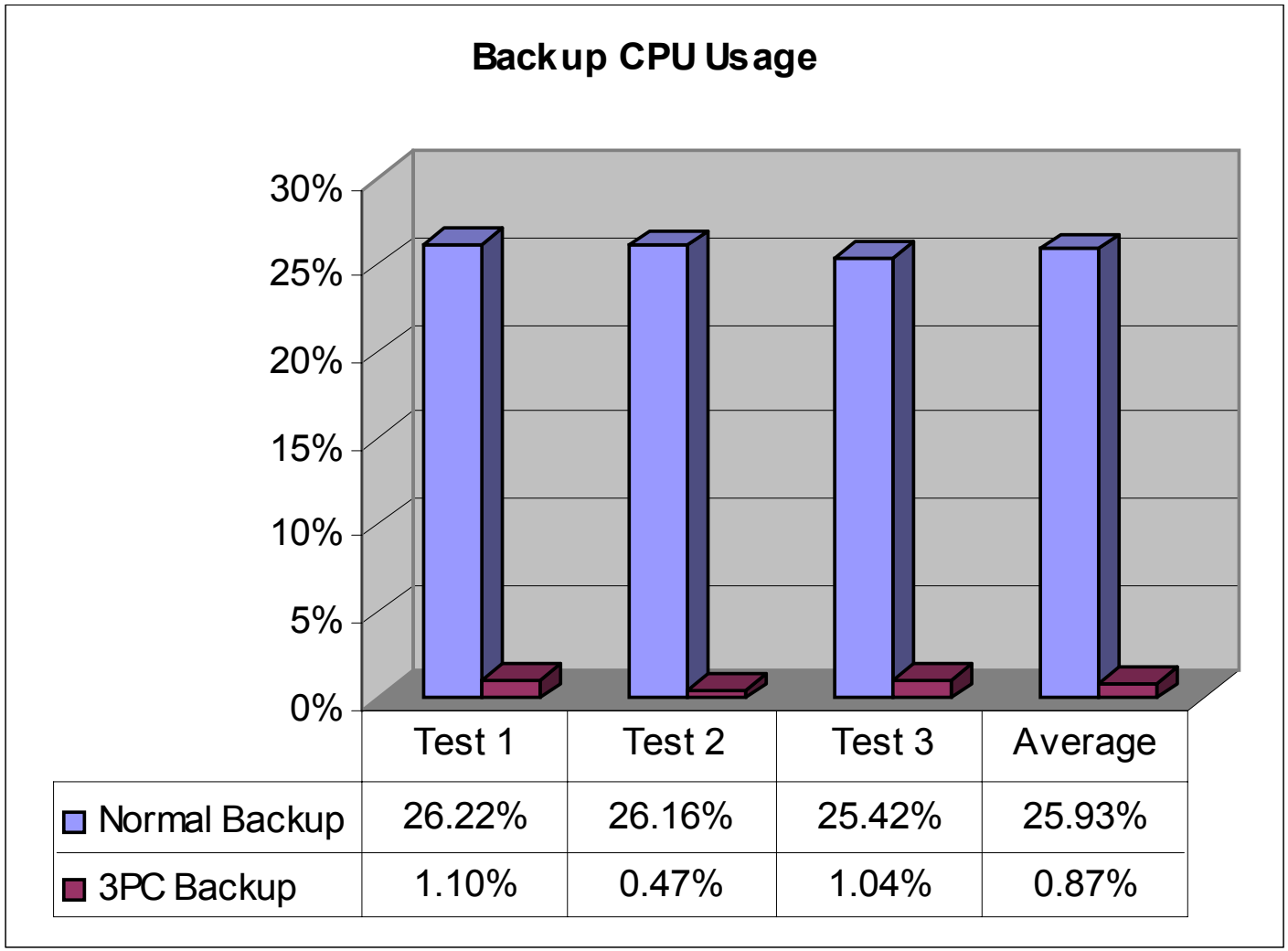
Rapid Backup – Performance

- 352 GB/Hr
- 40% faster than file system Backup



	Test 1	Test 2	Test 3	Average
Normal Backup	70.35	70.40	70.02	70.26
3PC Backup	97.31	98.01	97.39	97.57

Rapid Backup – Server Loading



	Test 1	Test 2	Test 3	Average
Normal Backup	26.22%	26.16%	25.42%	25.93%
3PC Backup	1.10%	0.47%	1.04%	0.87%

Q: Whatever happened to Serverless Backup?

A: It has just been re-named Rapid backup

XCopy/Serverless/Direct/Rapid Backup

- At last an HP official name that describes what this technique really achieves….RAPID BACKUP
  - Bypasses Filesystems to enable rapid backup direct from block level on disk to tape.
  - Also enables server offloading for the application server – increased high performance availability of application whilst backup is taking place.
  - Makes sense that this technology is best offered as a proven solution with close application integration
  - Still requires a backup server for backup and restore control. – unlikely there will ever be a rapid restore using this technology.

Certified EBS Rapid Backup Solution

- Direct Backup Engine for ESL:
  - Enables backup directly from disk to tape
  - Bypasses application server, requiring no downtime
  - Provides a high-performance backup solution
  - Restores are done through backup server at normal speed

- HP's Certified EBS Direct Backup Solution goes beyond protocol-level support to deliver a tested solution that works in the real world with:
  - ESL tape libraries
  - EVA or XP disk arrays
  - HP Data Protector 5.1 or Veritas NetBackup software 4.5
  - Oracle 9i