

How HP delivered a 3TB/hour Oracle™ backup & 1TB/hour restore

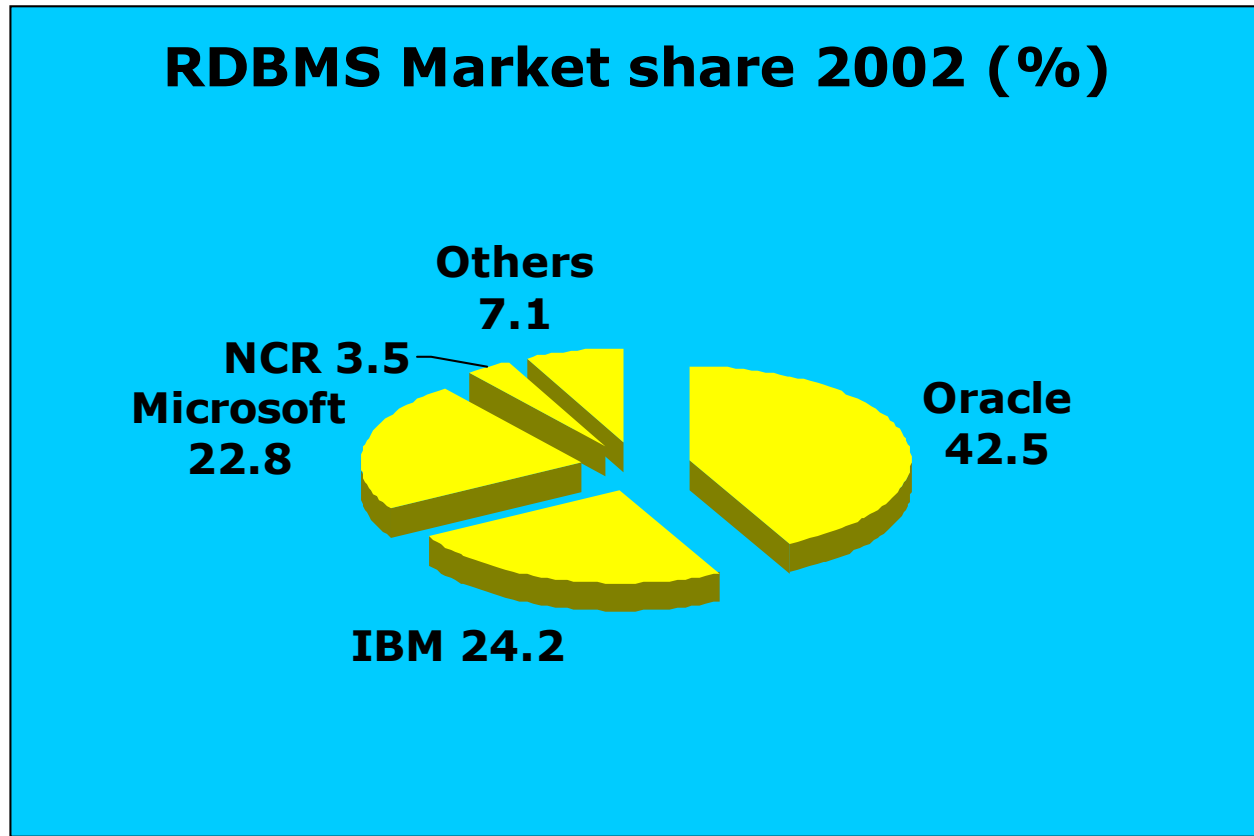
Andy Buckley

Technical Advocate

HP Network Storage Solutions



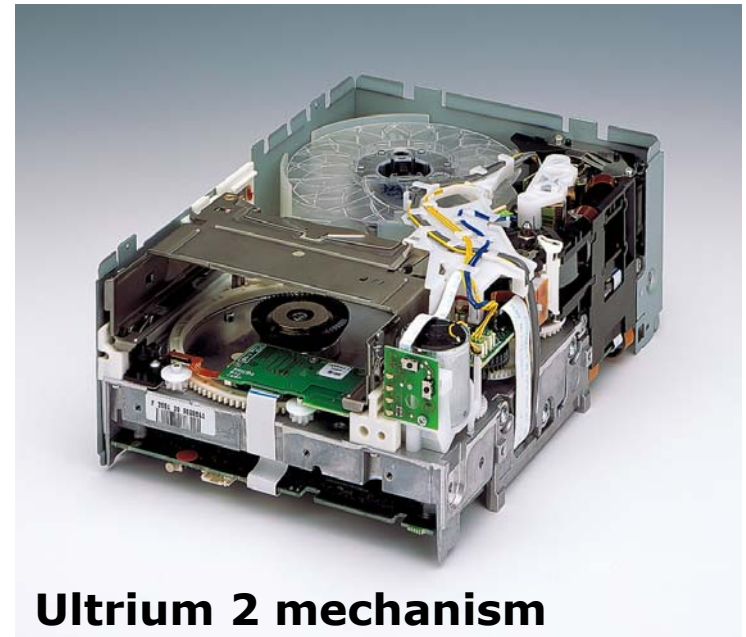
Why Oracle?



source: Gartner dataquest May 2003

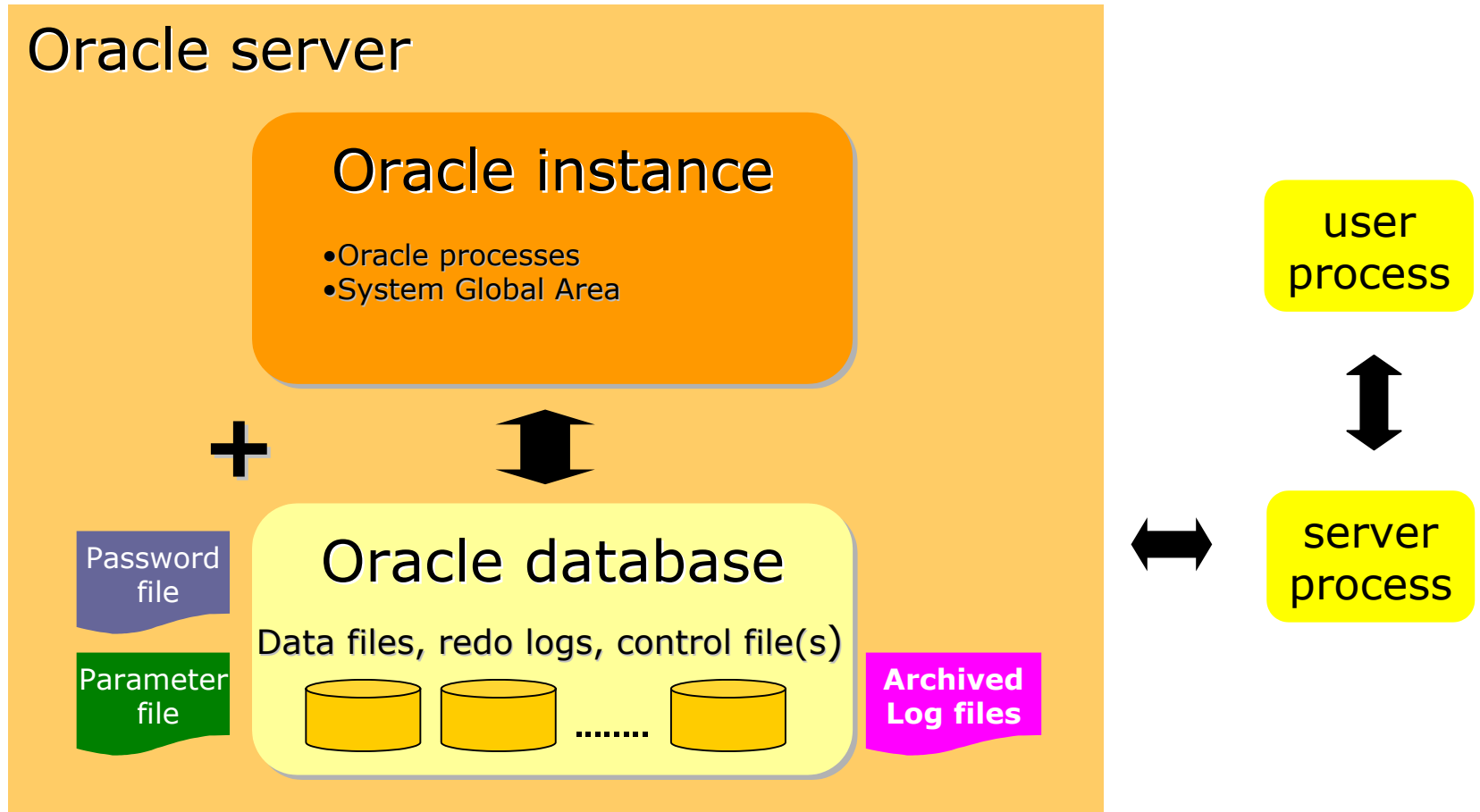
Why Ultrium tape?

- Capacity – 200GB Native
- Performance – 30MB/sec Native
 - the only technology with adaptive tape speed
- Reliability
- Industry standard
- Solid 4 generation roadmap



Ultrium 2 mechanism

Oracle Basics



Database backup options

- Disk based
 - Snapshots & Mirroring
 - Oracle Recovery Manager backup to disk

- Tape based backup
 - offline (cold) backups
 - online (hot) backups ← **Area of focus**

- Most Backup & Recovery strategies use a combination of tape and disk

To find out more.....

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

Session 1619

Harald Burose

Senior Architect

HP Nearline Storage Division

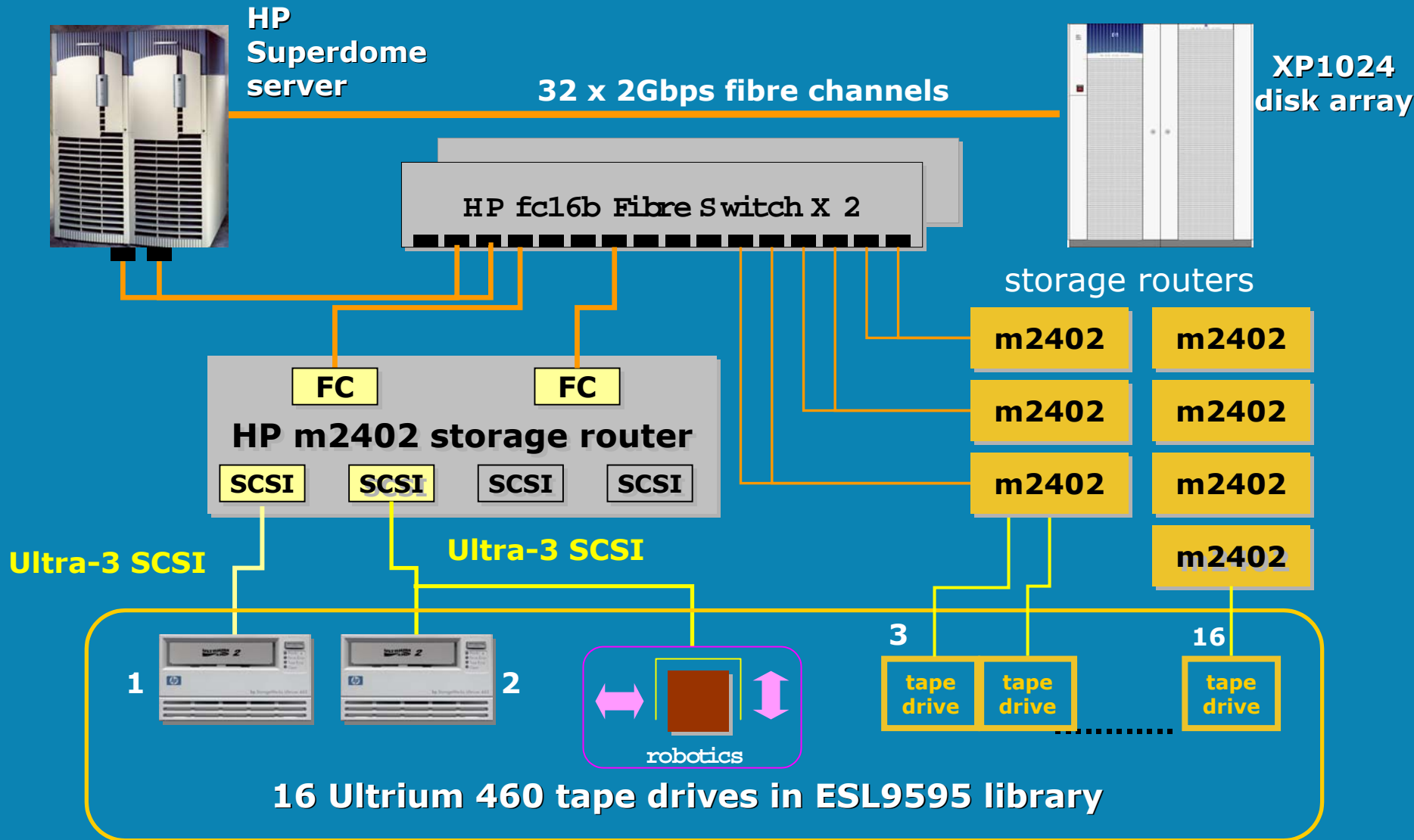
Session 1597

Glenn Wuenstel

Solutions Systems Engineer

HP Nearline Storage Division

Test lab - system diagram



Selective Storage Presentation – optimises performance by maximising FC bandwidth

Address Go



Network Storage Router M2402

To view settings, you may click on the modules. To change settings, you may click on ports and buses.



MAIN MENU

- Home
- System
- Modules
- Discovery
- Mapping
- Statistics
- Utilities
- Report
- Reboot

MAPPING MENU

FC MODULE 0

- Port 0
- Port 1

NO MODULE 1

NO MODULE 2

SCSI MODULE 3

- Bus 0
- Bus 1
- Bus 2
- Bus 3

FC MODULE 0 PORT 0 MAP SETTINGS

Host	Map
89512793 (FC Port Name (Low))	Indexed
10EDDA (FC Port Name (Low))	port0
2022D26 (FC Port Name (Low))	Indexed
10EB74 (FC Port Name (Low))	Indexed
10FDD0 (FC Port Name (Low))	Indexed

FC Map - Microsoft Internet Explorer

FC MODULE 0 PORT 0 port0

Lun	Protocol	Module	Bus	Type	Status	Device Specific Address
0	PSCSI	3	0	TAPE	UP	Target= 3 Lun= 0
1	PSCSI	3	1	TAPE	UP	Target= 3 Lun= 0

Fill Map

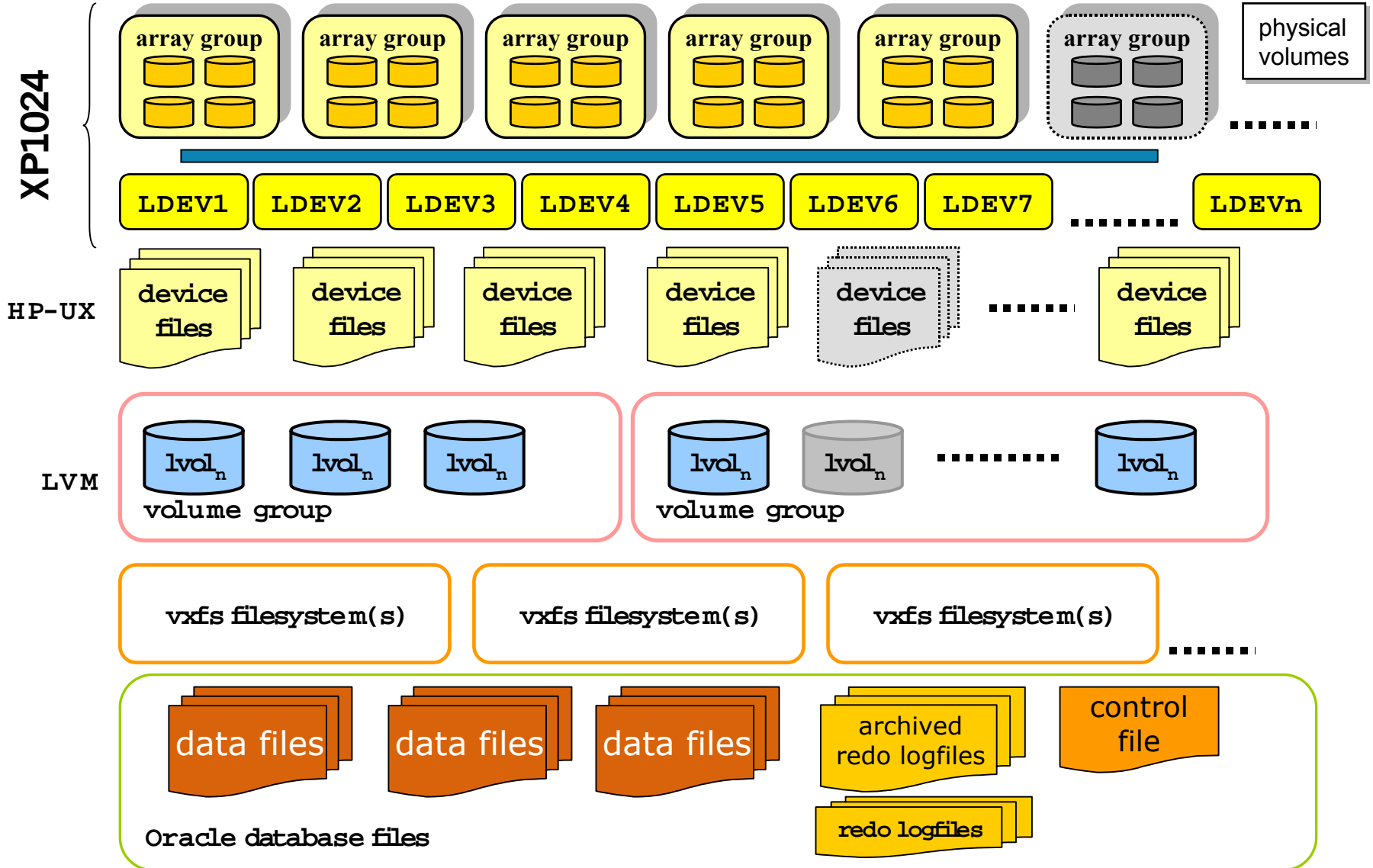
Priority

Delete Map Item(s)

Lun (from) (optional) to

Bind Host HBAs to specific ports and drives on the Router

HP-UX/Oracle9i



Generating the test database

- Test databases are available from the Transaction Processing Council TPC

<http://www.tpc.com>

- Used TPC-H dbgen tool
- Limit of 535GB per mount point as a result of the 8MB extent size. This is a limitation set by the HP-UX logical volume manager.

Test database structure

- 9TB of space allocated with 4.3TB of real data loaded
- 8 tables within the TPC schema
 - Line item 15billion records partitioned over 45 tablespaces
 - 3 datafiles per tablespace
 - Order : 4.5 billion records over 15 tablespaces
 - Partsupp : 2.4 billion records over 20 tablespaces
 - Part: 600 million records Over 20 tablespaces
 - Customer : 400 million records over 20 tablespaces
 - Supplier: 30 million records over 20 tablespaces
 - Nation: 25 records in a single tablespace
 - Region: 5 records in single tablespace
- Total: 259 datafiles and 146 tablespaces

Test Database structure - partitioned data

line item table

orderkey	partkey	supp_key	linenum ber	shipdate			comment
July orders							
Aug orders							

July orders

Aug orders

partition 1

partition 14

tablespaces:



datafiles



mount points



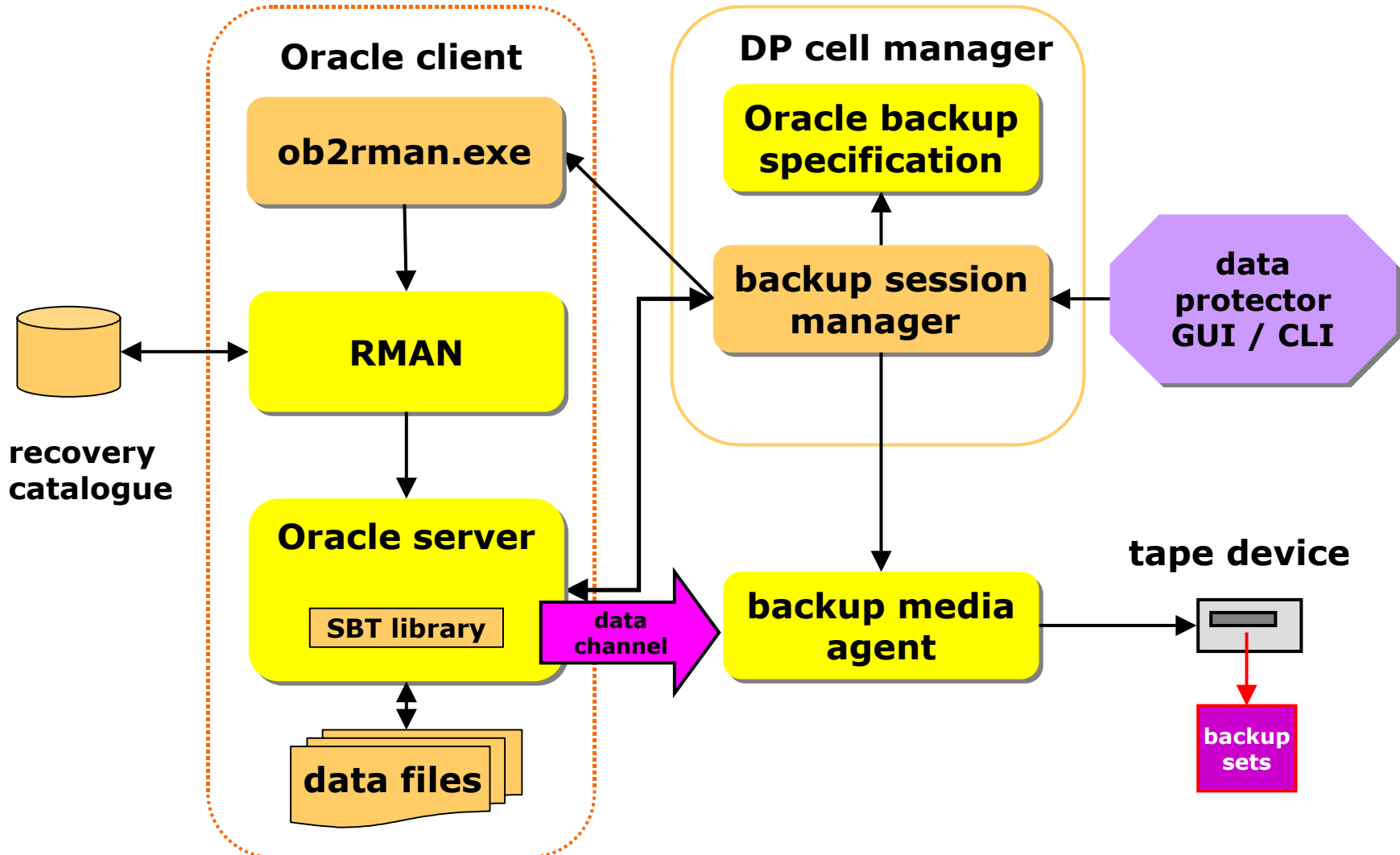
Oracle9i Recovery Manager

- commonly called RMAN
- available from Oracle 8 onwards
- can perform online & offline backups
- ability to back up 'raw disk' as well as filesystem
- provides block level data integrity checks
- eliminates backup of 'whitespace'
- provides incremental backup functionality
- provides 'point in time' backup
- platform independent
- database driven process
- scriptable

What does a media manager do?

- The Oracle recovery manager requires a media manager to backup to a tape device.
- Controls the tape library.
- Provides RMAN with access to tape devices.
- Tracks the location of media in the library
- HP Data Protector was used as the media manager for the 3 TB/hr project.

HP Data Protector & Oracle Integration



RMAN - Data Protector integration

The screenshot displays the HP OpenView Storage Data Protector Manager interface. The main window shows a tree view of backup specifications for an Oracle8 Server. An 'Application Specific Options' dialog is open, showing the 'Oracle8 integration' section. Within this dialog, the 'RMAN Script' field is highlighted with a red box. A secondary 'RMAN Script' dialog is also open, showing the script content. Red arrows and text annotations highlight specific parts of the configuration:

- Edit window**: Points to the 'RMAN Script' dialog box.
- Include control file**: Points to the line `include current controlfile` in the script.
- Full backup filesperset = 1**: Points to the line `backup incremental level <incr_level> filesperset 1` in the script.
- Channel allocation**: Points to the 'Edit...' button in the 'RMAN Script' field of the 'Application Specific Options' dialog.

The 'RMAN Script' dialog contains the following script content:

```
parms 'ENV=(OB2BARTYPE=Oracle8,OB2APPNAME=tpch,OB2BARLIST=Ora9_Database_to_
allocate channel 'dev_23' type 'sbt_tape'
parms 'ENV=(OB2BARTYPE=Oracle8,OB2APPNAME=tpch,OB2BARLIST=Ora9_Database_to_
allocate channel 'dev_24' type 'sbt_tape'
parms 'ENV=(OB2BARTYPE=Oracle8,OB2APPNAME=tpch,OB2BARLIST=Ora9_Database_to_
allocate channel 'dev_25' type 'sbt_tape'
parms 'ENV=(OB2BARTYPE=Oracle8,OB2APPNAME=tpch,OB2BARLIST=Ora9_Database_to_
allocate channel 'dev_26' type 'sbt_tape'
parms 'ENV=(OB2BARTYPE=Oracle8,OB2APPNAME=tpch,OB2BARLIST=Ora9_Database_to_
allocate channel 'dev_27' type 'sbt_tape'
parms 'ENV=(OB2BARTYPE=Oracle8,OB2APPNAME=tpch,OB2BARLIST=Ora9_Database_to_
allocate channel 'dev_28' type 'sbt_tape'
parms 'ENV=(OB2BARTYPE=Oracle8,OB2APPNAME=tpch,OB2BARLIST=Ora9_Database_to_
allocate channel 'dev_29' type 'sbt_tape'
parms 'ENV=(OB2BARTYPE=Oracle8,OB2APPNAME=tpch,OB2BARLIST=Ora9_Database_to_
backup incremental level <incr_level> filesperset 1
format Ora9_Database_to_30Dev_Conc1<ch_%s:st%p>.dbf
database
include current controlfile
}
```

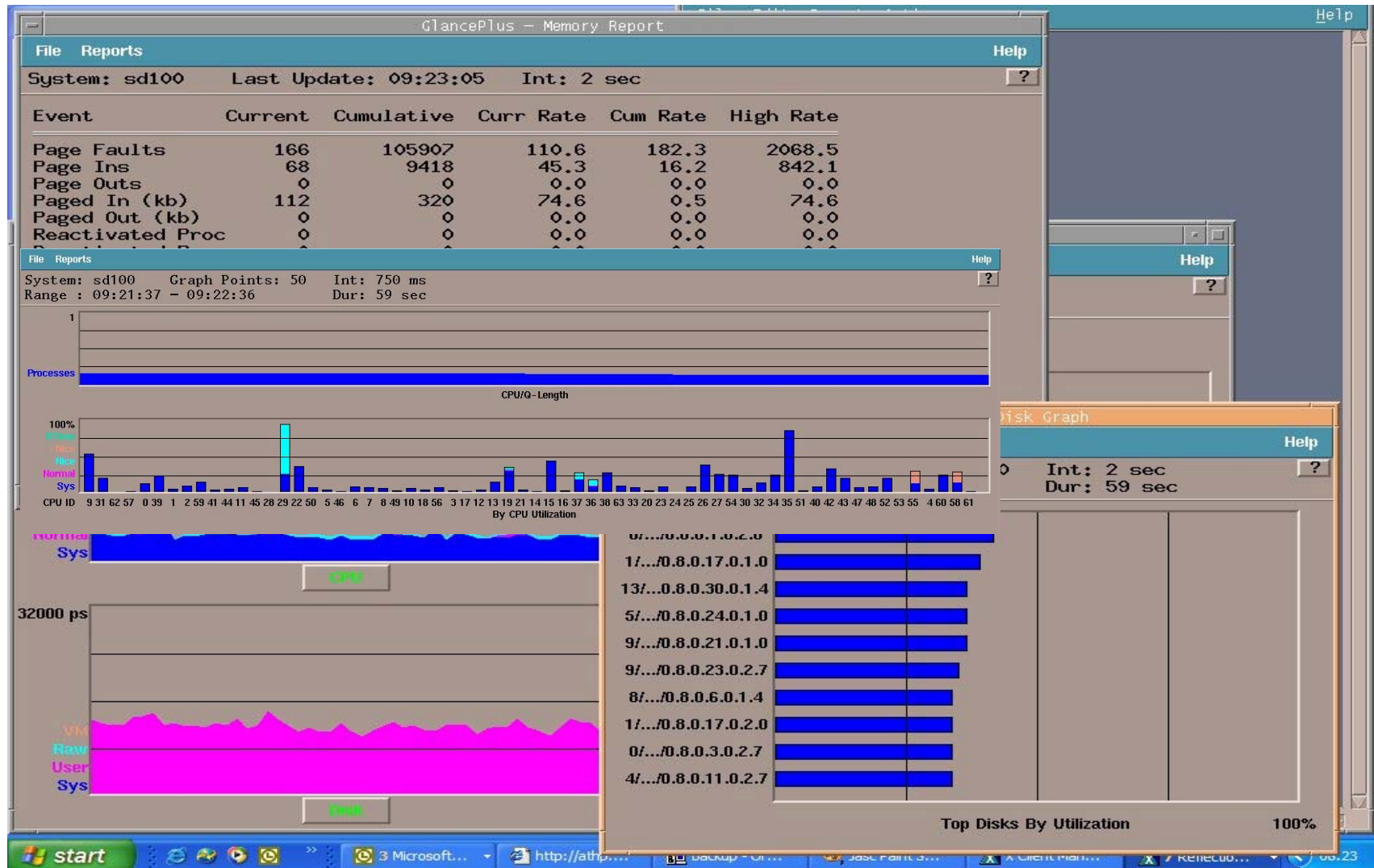

RMAN & Data Protector tuning

- concurrency
- filesperset
- blocksize
- maxopenfiles
- tape I/O slaves
- disk I/O slaves
- contofile autobackup

Measuring the performance

- Glance/XL utility monitors CPU and disk utilization
- RMAN statistics (SQL script queries the recovery catalog)
- HP Data Protector statistics
- FC Switch throughput

HP Glance/XL results



Results: 3TB/hr backup

# tape drives	# processors	CPU idle	Data protector parameters	RMAN settings	backup performance
16 in HP ESL9595 tape library	32	34%	concurrency=1 (1 RMAN channel/tape drive)	disk I/O slaves=16 filesperset=no set maxopenfiles=1 tapeblocksize=1MB backup_tape_io_slaves=disabled controlfile autobackup=ON	3.16 TB/hr (including tape load/unload time)
					3.62 TB/hr (excluding tape load/unload time)

Results: backup (16 way CPU)

# tape drives	# processors	CPU idle	Data protector parameters	RMAN settings	backup performance
16 in HP ESL9595 tape library	16	15%	concurrency=1 (1 RMAN channel/tape drive)	disk I/O slaves=16 filesperset=no set maxopenfiles=1 tapeblocksize=256 KB backup_tape_io_slaves=disabled controlfile autobackup=ON	2.87 TB/hr (includes tape load/unload time)

Results: restore

# tape drives	# processors	CPU idle	Data protector parameters	RMAN settings	backup performance
16 in HP ESL9595 tape library	32	45%	concurrency=1 (1 RMAN channel/tape drive)	disk I/O slaves=16 filesperset=no set maxopenfiles=1 tapeblocksize=1MB backup_tape_io_slaves=enabled controlfile autobackup=ON	1.23 TB/hr (including tape load/unload time)
					1.29 TB/hr (excluding tape load/unload time)

- Restore times subsequently improved to 2.46TB/hr !

Database Recovery

- Use RMAN script to recover database
- Recovery Catalogue must be available (normally resides on another server)
- Media Manager must be available
- GUI based recovery is available from Data Protector 5.1 onwards

Future directions

- Oracle Real Application Clusters
- Serverless Backup
- Advanced Filesystems
- Ultrium 3 (late 2004) +60Mb/s

Lessons Learnt

- Backup performance determined in the main by disk subsystem performance. The 3:1 rule held true.
- Use SAME for best Oracle file system performance
- Storage Router configuration could be simplified.
- Memory usage was low 19GB Max (of 256GB)
- Could have been achieved with 16 Processors (15% idle)

Lessons Learnt

- In a high performance environments set 1 RMAN channel per tape drive, don't use external "multiplexing"
- Restore performance subsequently increased to 2.4 TB/HR – restore 40-60% of backup rate.
- Data Protector 5.1 integration with Oracle is now even better – restore GUI.
- Only Ultrium technology has adaptive tape speed to optimize performance and increase media life.

Conclusions

- 3TB/hour online database backup of a real life system is possible and cost effective with standard tape library configurations
- Ultrium tape has the growth path in future generations for even greater capacity and performance
- The Oracle9i Recovery Manager provides performance, flexible and reliable backup and recovery
- Tape based backup is still essential and viable

Reference material

- Oracle9i RMAN Backup & Recovery
by RG Freeman & M Hart (ISBN 0-07-222662-5)
- HP Dataprotector Oracle 9i Integration Guide
- HP Dataprotector Unix installation Guide
(www.hp.com)
- Oracle9i Recovery Manager User Guide and Reference
(<http://otn.oracle.com>)
- Copies of HP 3TB/Hr Whitepaper available after this presentation

HP WORLD 2003

Solutions and Technology Conference & Expo

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



Supplementary slides



RMAN – Data Protector integration

The screenshot displays the HP OpenView Storage Data Protector Manager interface. The left pane shows a tree view of backup specifications under the 'Oracle8 Server' folder. The right pane shows the configuration for a selected backup job, including source and destination options.

Backup Specifications (Left Pane):

- Backup
- Backup Specifications
- Filesystem
- Oracle8 Server
 - Ora9_10DBF_to_10_Dev_Conc10
 - Ora9_150DBF_30Dev_Conc5
 - Ora9_1DBF_to_1Dev_Conc1
 - Ora9_1TS_Null
 - Ora9_20DBF_20Dev_Conc1
 - Ora9_30DBF_to_10Dev_Conc3
 - Ora9_30DBF_to_30Dev_Conc1
 - Ora9_5DBF_to_5Driv_Conc1
 - Ora9_Database_to13Dev_Conc5_Filesperset4
 - Ora9_Database_to16Dev_Conc15_8Router
 - Ora9_Database_to_13Dev_Conc1_filesperset20
 - Ora9_Database_to_16Dev_Con10_4Router
 - Ora9_Database_to_16Dev_Conc10
 - Ora9_Database_to_16Dev_Conc1_8Router
 - Ora9_Database_to_16Dev_Conc5
 - Ora9_Database_to_30Dev_Con10
 - Ora9_Database_to_30Dev_Con5
 - Ora9_Database_to_30Dev_Conc1
 - Ora9_Database_to_30Dev_Conc3
 - Ora9_Database_to_30Dev_Conc8
 - Ora9_Database_to_Lib_maxopenfiles1**
 - TEST_Ora9_Database_to_16FILEDEV_CONC10
- Templates

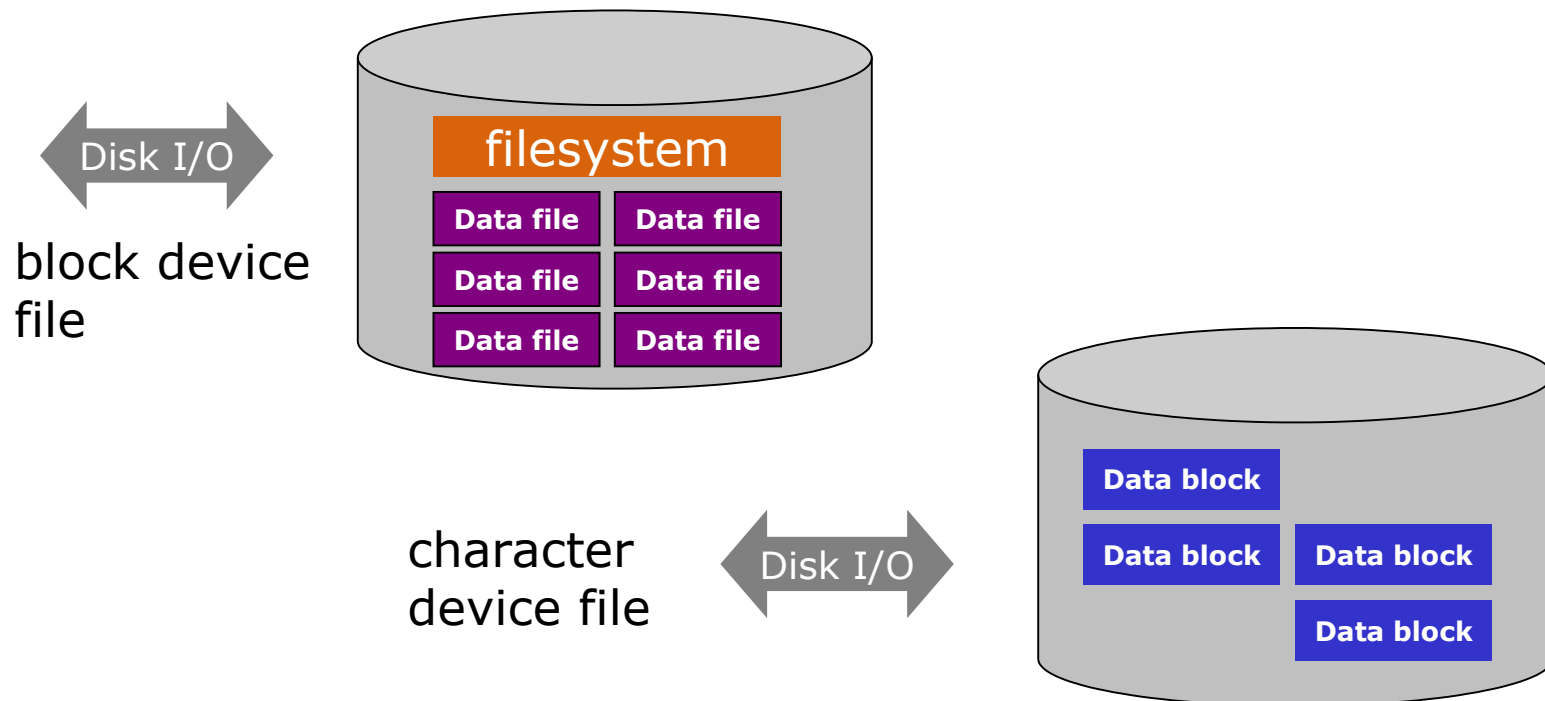
Backup Job Configuration (Right Pane):

- Source: Select application specific data that you want to back up.
- Show: Selected
- sd100.bbn.hp.com tpch
 - DATABASE

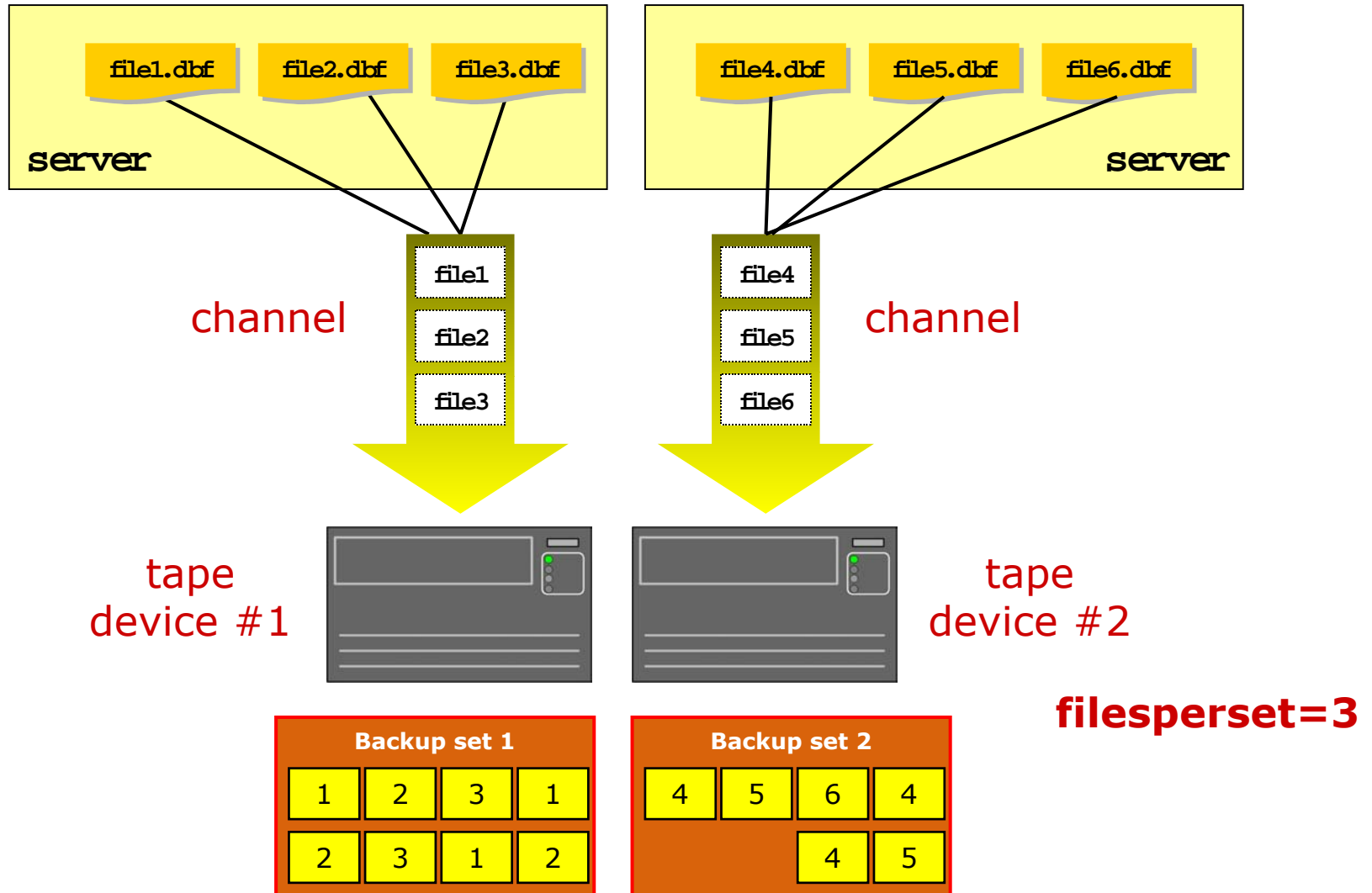
Annotation: A red arrow points from the text "These backup jobs contain prepared scripts" to the 'Ora9_Database_to_Lib_maxopenfiles1' job in the left pane.

Filesystem vs Raw disk

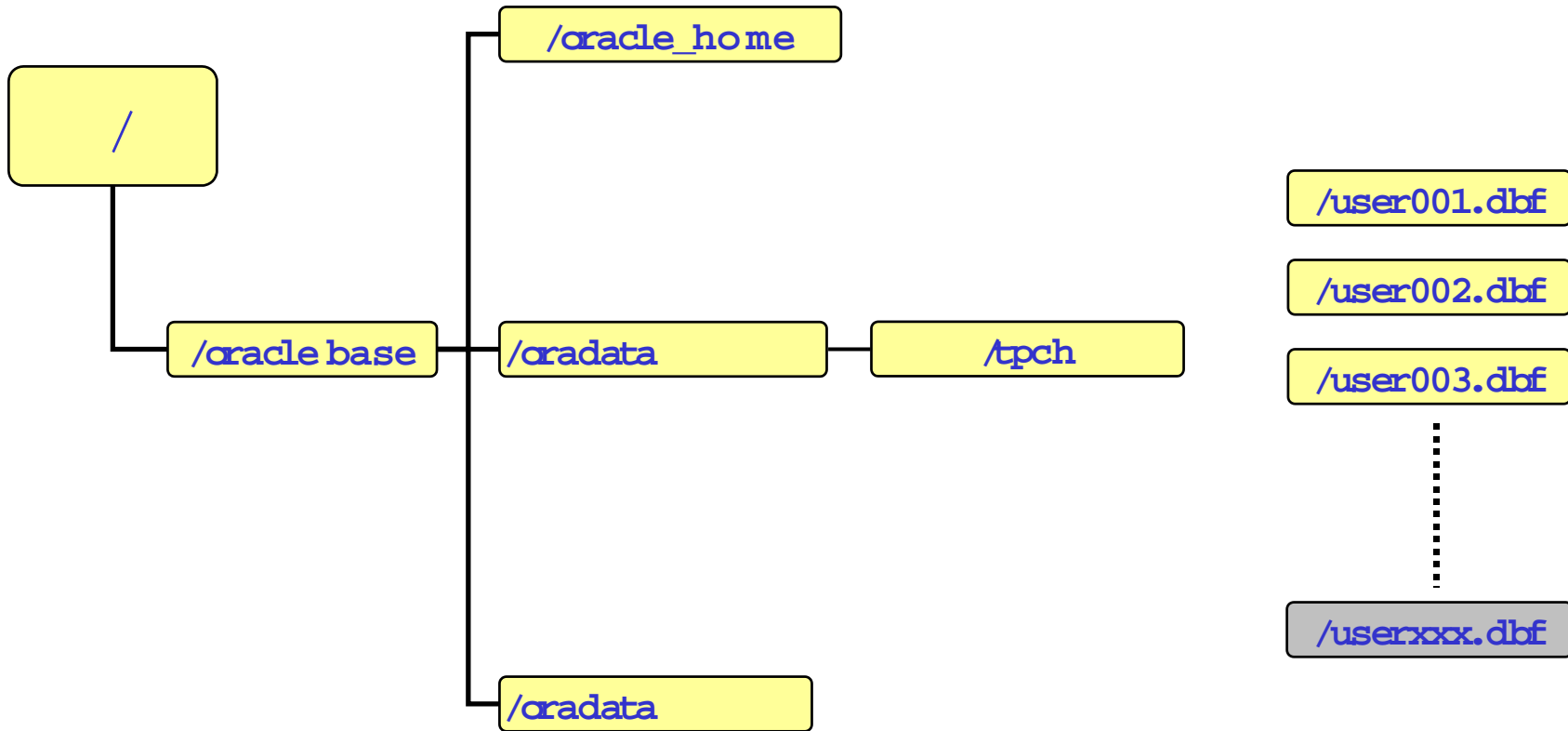
- Oracle9i can use the filesystem supplied by the OS or what is known as raw disk. Raw disk is generally considered to be faster but is less manageable.
- The 3TB/hr project used filesystem



Multiplexing backup sets



3TB/hr directory structures



as per Oracle9i OFA model (Optimal Flexible Architecture)