Enterprise
Data Backup and Recovery

W. Curtis Preston, PSA
(Pencom Systems Administration)

Do you need a commercial backup utility?
Introduction

- Thirty flavors of UNIX, Novell, W95/NT, Mac, Oracle, Informix, Sybase, Ingres, DB2, L. Notes, SQL Server
- Size and complexity of data growing rapidly
- Misinformation or no information
- Many shops do not know the danger that lies before them
Agenda

- Defining the Enterprise
- Native Utilities - what’s wrong with them?
- Databases - why are they so hard?
- Available Hardware and Software
- Putting it all together - don’t forget anything
Essential Premise

No Data Should Ever Be Lost!!!
Overview

- Native Utilities -- old pitfalls still there
- Databases make things even harder
- Get a good backup system and learn it
- *You probably need a commercial utility now*
- How to choose a commercial backup utility
TEST YOUR BACKUPS

- Please Document them
- Test the documentation
- Test your disaster recovery plan
- Prove your backups do what you think
- Test them again in 6 months
Defining the New Enterprise

- New file systems (jfs, vxfs, dfs)
- Huge Databases -- RLTP, Warehouse
- 24x7 -- Web, Batch Processing
- Linux, BSDI, NT, Novell
- MVS?
Native UNIX Utilities

- “Free”
- Almost always require some scripting
- Not always compatible, even when you think they are
- No cataloging, automatic re-routing,
- All responsibility lies on your shoulders
- tar, cpio, dump
**tar vs. cpio**

- Usually compatible*
- Include list limited to filename expansion
- Changes atime*
- No remote devices*
- No wild-cards on restore
- No sparse files*

- Often incompatible*
- Unlimited include list, read from stdin
- Change atime or ctime
- No remote devices*
- Supports wild-cards on restore
- No sparse files

*Fixed in GNU versions
“dump”

- Index on front of tape allows interactive restore and FFS
- Does not change access times
- Works directly with device (faster)
- Available on *most* UNIX systems
- *Highly dependent on the type of file system*
- File systems developing faster than dump
- Different names, options, features, format
Database backups

- A difficult, sometimes impossible, task
- DBA’s don’t know backup methodologies
- SA’s don’t know database structures and technologies
- Assumptions mean *things get left out*
Oracle

- Instances, Tablespaces, Data files, Control Files, On-line and Archived Redologs, oratab
- "begin backup," "end backup" and "backup control file"
- oraback.sh
- Enterprise Backup Utility (EBU)
- SQL Backtrack
Informix

- Servers, Dbspaces, Data files, Continuous backups, no central file
- ontape (tbtape)
- ON-archive
- onbar
- SQL Backtrack
Sybase and SQL Server

- Master database, Devices, Transaction log
- Bad setup can invalidate any backup/recovery plan
- dump database, dump transaction log
- Backup Server
- SQL Backtrack
SQL Backtrack

- Incremental Backups
- Multiple backups per tape
- remote devices
- Single interface to 3 products
Others

- MS Exchange -- Cannot restore a single mailbox
- Lotus Notes -- Difficult to do incremental
- HP Openview -- Need tie-in to not lose display
- Remedy -- usually Informix
- ??? -- know what it is
3rd Party Utilities

- Public Domain Utilities (Amanda)
- Commercial Utilities
  - Low End, concentrating on desktop
  - High End, concentrating on servers
  - Middle of the Road
Which One is the Best?

- *One that works!*
- One that works *in your environment*
- One that works *with you*
- One that *lets you work around it*
Important Things to Look At

- Database
- Supported Platforms
- Multitasking
Database Format

- size (44 Bytes per file -> 600 Bytes per file)
- Backups for your backups (How hard?)
- Distributed Databases?
- ASCII, btree, DB2
**Supported Platforms**

- *Properly* Supported Platforms
  - UNIX Special files (Device, Named Pipes)
  - NT/95 registry
  - Netware NDS

- Interfaces to Databases

- Ability to run home-grown scripts
Multitasking

- Many systems to one device
- One system to many devices
- Test effect on backups & restores
- Automated or manual
Devices

- Standard device drivers
- Jukeboxes
- Remote devices
- Peer-to-Peer backups
Tape Storage Technologies

- DLT
- Mammoth
- Magstar
- AIT
Tape Drive Technology Issues

- **Helical vs. Linear**
  - .5 ips vs. 100 ips
  - Overwrites = resync

- **Size vs. access time**
  - Large Tape ~= Long Seek Time
  - *Total access time* must be considered in high-use environments (HSM, heavy restores)
Tape Technology Advances

- Memory in Cassette
  - 75/150ips seek speed
  - Does not require rewind
  - Stores Reads/Write Passes & Cumulative ECC Rate

- DLC vs. MP
  - Extended Life
  - No cleaning
  - Higher Magnetic Flux Density
Quantum DLT

- 5/10 MB/s if streamed
- MB/s Very Logarithmic
- Linear, slow load & FFS
- Longest total access time
- Fragile cartridge interior
- Costly (purch. oper'n)
Exabyte Mammoth

- 3/6 MB/s
- Helical Scan
- Long total access time
- DLC Coating = no cleaning required
- Limited due to backward compatibility
- No upgrade announcements
IBM Magstar

- New 3570 Cartridge
- 3/8 MB/s
- Linear
- Rugged Cartridge
- Tape stays inside
- Mounts mid-point
- Fastest total access time
Sony AIT

- 3/8 MB/s
- Helical Scan
- DLC = no cleaning
- Memory in Cassette
- Redesigned Media made to support all three drive generations (6/16 MB’s in ‘98 & 12/32 MB/s in ‘00)
- As Fast as Mag, Faster than DLT, Cheaper than both
Sony AIT

The future direction of AIT

- AIT-1: 1996, 25GB, 3MB/s
- AIT-2: 1996, 50GB, 6MB/s
- AIT-3: 2009, 100GB, 12MB/s

Strength of the aluminum plate rigid cartridge

Graph showing the transformation amount (mm) versus load (Kg) for different cartridges.
## Total Access Time Comparison

<table>
<thead>
<tr>
<th></th>
<th>DLT</th>
<th>Mamm.</th>
<th>Magstar</th>
<th>AIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rew ind/Eject Time</td>
<td>79</td>
<td>58.1</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>50MB File</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Head Position (avg.)</td>
<td>66</td>
<td>72</td>
<td>8.8</td>
<td>20</td>
</tr>
<tr>
<td>Load Vol.</td>
<td>48</td>
<td>20</td>
<td>5.6</td>
<td>7</td>
</tr>
<tr>
<td>Robotics Exchange</td>
<td>24</td>
<td>10</td>
<td>6.7</td>
<td>7.5</td>
</tr>
</tbody>
</table>
Network Hardware

- IBM 8270
  - 8-16 Token Ring Ports
  - 1 UFC Slots
- IBM 8272
  - 8 UFC Slots
- Universal Feature Cards
  - 4 port Token TP
  - 2 port Token Fiber
  - 1 port 155Mbps ATM
Network Hardware

- Cisco Catalyst 5000
  - HSC has chassises
  - Auto 10/100 Modules
  - Will support 1 1Gb card
Backup Product Evaluation

- 250 question RFI
- A lot of “good” products out there
- Good products with some bad answers
- Some products with false answers
- More products coming
Putting it all Together

- Scripts can copy database data to backup disks
- Scripts can shutdown databases
- “Second opinion” monitor -- is every system/disk included?
- Success Monitor -- is everything we’re backing up being backed up?
New Things To Keep You up all Night

- DFS, EFS
- Terabyte File Systems
- New File system types
- Networks not capable
- More local data
New Things That Keep me up all Night

- Ampex
- DTF, Magstar, AIT
- DLT “Stackable” Jukeboxes
- API’s
- NDMP
Summary

- Take your backups seriously
- Understand the data you are responsible for
- Find the proper tools and make them co-exist
- Monitor the tool
- Keep up with new technology (It doesn’t)
For more help:

- I can be reached at: curtis@pencom.com
  http://www.pencom.com/psa/answer.html
- Independent Backup & Recovery info at:
  http://www.backupcentral.com
- Look for “Enterprise Backup and Recovery” from O’Reilly and Associates
  (Estimated Release Date 1Q, 1998)