Backups

tsaimh (2022, CC BY)
lwhsu (2019-2020, CC BY)
? (?-2018)

國立陽明交通大學資工系資訊中心
Computer Center, Department of Computer Science, NYCU

Outline

- Backup devices and media
- Backup philosophy
- Unix backup and archiving commands

Key Concepts

- The backup that has never been verified is the same as no backup
- Types
  - Scope
    - Full
    - Incremental
    - Differential
  - Storage
    - Hot/Cold
    - Online/Offline
    - Remote

Backup Media – By Physical Characteristics

- Hard Disk
  - Pro: common
  - Con: fragile
- Optical Disk
  - Pro: durable
  - Con: size
- Magnetic Tape
  - Pro: cost-effective
  - Con: slow

Backup Media – By Product

- RAID
  - Redundant Array of Independent Disks
- NAS
  - Network Attached Storage
- SAN
  - Storage Area Network

Backup Media – By Availability

- Off-line Storage
  - CD、DVD、MO
    - Adv:
      - Low cost, high reliability
    - Disadv:
      - Not-convenient, low speed
- Near-line Storage
  - JukeBox、Tape Library
    - Adv:
      - High capacity, high reliability
    - Disadv:
      - High malfunction rate, Not-convenient
- On-line Storage
  - Disk Array (RAID)
    - Adv:
      - Fast and high availability
    - Disadv:
      - High cost

Tape Library (3U/720TB)
Source: https://www.backupworks.com/Qualstar-Q40-LTO-8-Tape-Library.aspx

JukeBox
Source: https://kintronics.com/products/archiving/hit-storage-optical-jukebox-library-system/

Backup Media – Cloud

- Azure Backup
- AWS S3 Glacier / Deep Archive
- GCP Archival Cloud Storage

There is no cloud
it's just someone else's computer

https://www.safer-computing.com/?p=5010

Backup Philosophy

- Perform all dumps from one machine
- Label your tapes
- Pick a reasonable backup interval
- Choose filesystems carefully
- Make daily dumps fit on one tape
- Make filesystems smaller than your dump device
- Keep Tapes off-site
- Protect your backups
- Limit activity during dumps
- Check your tapes
- Develop a tape life cycle
- Design your data for backups
- Prepare for the worst

Dumping filesystems – dump command (1)

- Used to backup filesystem into a large file to archive to an external device
- Advantages:
  - Backups can span multiple output media
  - Files of any type can be backed up and restored
  - Permissions, ownerships, and modification times are preserved
  - Files with holes are handled correctly
  - Backups can be performed incrementally
- Limitations:
  - Each filesystems must be dumped individually
  - Only filesystems on the local machine can be dumped
    - NFS filesystem is not allowed

Dumping filesystems – dump command (2)

- Backup level
  - 0 ~ 9
    - **Level 0 -> full backup**
    - **Level N -> incremental backup of Level≦ N-1 for N = 1 ~ 9**
- dump command format
  - % dump [arguments] file-system
- dump command arguments
  - **u: update the /etc/dumpdates file after dump**
  - **f: the output backup file**
    - **Special device file, like /dev/nrsa0**
    - **Ordinary file**
    - **'-' to standard out**
    - **"user@host:file"**
  - **d: tape density in bytes per inch**
  - **s: tape length in feet**
  - **a: auto-size, bypass all tape length considerations (default d = 1600, s = 2300)**

Dumping filesystems – dump command (3)

- Example: Full backup

```
zfs[/mnt]  ls -lh
drwxr-xr-x   3 root   wheel      512B       Nov 22 15:34 ./
drwxr-xr-x  20 root   wheel      25B        Nov 18 20:02 ../
-rw-r--r--   1 root   wheel      512M       Nov 21 22:20 haha
zfs[/mnt]  cat /etc/dumpdates
zfs[/mnt]  df -h
Filesystem          Size      Used       Avail      Capacity   Mounted on
zfs                 15G       4.1G       11G        27%        /
devfs               1.0K      1.0K       0B         100%       /dev
/dev/da0s1a         8.7G      512M       7.G        6%         /mnt
zfs[/mnt]  sudo dump 0uLf - /dev/da0s1a > ~/dump.0
  DUMP: Date of this level 0 dump: Sun Nov 22 15:37:44 2009
  DUMP: Date of last level 0 dump: the epoch
  DUMP: Dumping snapshot of /dev/da0s1a to standard output
  DUMP: mapping (Pass I) [regular files]
  DUMP: mapping (Pass II) [directories]
  DUMP: estimated 525772 tape blocks.
  DUMP: dumping (Pass III) [directories]
  DUMP: dumping (Pass IV) [regular files]
  DUMP: DUMP: 525625 tape blocks
  DUMP: finished in 36 seconds, throughput 14600 KBytes/sec
  DUMP: level 0 dump on Sun Nov 22 15:37:44 2009
  DUMP: DUMP IS DONE
zfs[/mnt]  cat /etc/dumpdates
/dev/da0s1a                    0 Sun Nov 22 15:37:44 2009
```

11

Dumping filesystems – dump command (4)

- Example: Incremental backup

```
zfs[/mnt]  sudo cp -Rp /etc /mnt/
zfs[/mnt] -- ls -lh
drwxr-xr-x   4 root  wheel     512B Nov 22 15:48 ./
drwxr-xr-x  20 root  wheel      25B Nov 18 20:02 ../
drwxr-xr-x  20 root  wheel     2.0K Nov 22 15:35 etc/
-rw-r--r--   1 root  wheel     512M Nov 21 22:20 haha
zfs[/mnt]  sudo dump 2uLf - /dev/da0s1a > ~/dump.2
  DUMP: Date of this level 2 dump: Sun Nov 22 15:49:04 2009
  DUMP: Date of last level 0 dump: Sun Nov 22 15:37:44 2009
  DUMP: Dumping snapshot of /dev/da0s1a to standard output
  DUMP: mapping (Pass I) [regular files]
  DUMP: mapping (Pass II) [directories]
  DUMP: estimated 2267 tape blocks.
  DUMP: dumping (Pass III) [directories]
  DUMP: dumping (Pass IV) [regular files]
  DUMP: DUMP: 2124 tape blocks
  DUMP: finished in less than a second
  DUMP: level 2 dump on Sun Nov 22 15:49:04 2009
  DUMP: DUMP IS DONE
zfs[/mnt]  cat /etc/dumpdates
/dev/da0s1a                    0 Sun Nov 22 15:37:44 2009
/dev/da0s1a                    2 Sun Nov 22 15:49:04 2009
zfs[/mnt]  ls -lh ~/dump*
-rw-rw-r--  1 tsaimh  user   513M Nov 22 15:38 /home/tsaimh/dump.0
-rw-rw-r--  1 tsaimh  user   2.1M Nov 22 15:49 /home/tsaimh/dump.2
```

12

Restoring from dumps – restore command (1)

- Restore can do
  - Restoring individual files
  - Restoring entire filesystem
- Options of restore command
  - i: interactive restore
  - r: restore an entire filesystem
  - f: the backup file that restore is going to use

Restoring from dumps –
restore command (2)

● Restore individual file interactively

```
zfs[/tmp]  cat ~/dump.2 | restore if -
restore > ?
Available commands are:
        ls [arg] - list directory
        cd arg - change directory
        pwd - print current directory
        add [arg] - add `arg' to list of files to be extracted
        delete [arg] - delete `arg' from list of files to be extracted
        extract - extract requested files
        setmodes - set modes of requested directories
        quit - immediately exit program
        what - list dump header information
        verbose - toggle verbose flag (useful with ``ls'')
        help or `?' - print this list
If no `arg' is supplied, the current directory is used
```

Restoring from dumps – restore command (3)

- Restore individual file interactively (cont.)

```
zfs[/tmp]  cat ~/dump.2 | restore if -
restore > ls
.:
.snap/ etc/
restore > cd etc
restore > add make.conf
restore > extract
set owner/mode for '.'? [yn] n
restore > quit
zfs[/tmp]  ls -ld etc
drwxr-xr-x  2 tsaimh  wheel  3 Nov 22 15:35 etc/
zfs[/tmp]  ls -l etc
drwxr-xr-x   2 tsaimh  wheel    3 Nov 22 15:35 ./
drwxrwxrwt  10 root     wheel   42 Nov 22 15:58 ../
-rw-r--r--   1 tsaimh  wheel  590 Nov 19 23:04 make.conf
```

Restoring from dumps – restore command (4)

- Restore entire filesystem
  - $ restore -rf /home/temp/root.0
  - Steps
    - Restore level 0 first
    - Restore incremental dumps
      - 0 0 0 0 0
      - 0 5 5 5 5
      - 0 3 2 5 4 5
      - 0 9 9 5 9 9 3 9 9 5 9 9
      - 0 3 5 9 3 5 9

Other archiving programs

- tar command
  - Read multiple files and packages them into one file
  - Example

> # tar czvf etc.tar.gz /etc/
>
> # tar xzvf etc.tar.gz
>
> # tar cf - fromdir | tar xfp - -C todir

- dd command
  - Copy filesystems between partitions of exactly the same size
  - Example

> # dd if=/dev/rst0 of=/dev/rst1
>
> # dd if=/tmp/kern.flp of=/dev/fd0
>
> # dd if=/dev/da1 of=/dev/da2 bs=1048576

CS home backup

- Using rsync
  - rsync -a --delete
    - -a: archive mode
      - Recursive and preserve everything
    - --delete:
      - Delete any file that are not in the sending sid

```
0 4 * * 1 (cd /raid && /usr/local/bin/rsync -aH --delete cs      /backup/user/)
0 4 * * 2 (cd /raid && /usr/local/bin/rsync -aH --delete gcs     /backup/user/)
0 4 * * 3 (cd /raid && /usr/local/bin/rsync -aH --delete dcs     /backup/user/)
0 4 * * 4 (cd /raid && /usr/local/bin/rsync -aH --delete alumni  /backup/user/)
```

CS home backup

- Snapshot
  - CS home snapshot

```
tsaimh@csduty.cs.nctu.edu.tw[/u/gcs][20:14]$ ls -a
.          01        103        109        91        95        99
..         100        104        193        92        96
.snap      101        105        199        93        97
.snapshot 102        106        90        94        98
tsaimh@csduty.cs.nctu.edu.tw[/u/gcs/.snapshot][20:14]$ cd .snapshot/
tsaimh@csduty.cs.nctu.edu.tw[/u/gcs/.snapshot][20:14]$ ls
4hour.2018-01-02_0000  4hour.2018-01-02_2000  daily.2018-01-01_0010
4hour.2018-01-02_0400  daily.2017-12-28_0010  daily.2018-01-02_0010
4hour.2018-01-02_0800  daily.2017-12-29_0010  weekly.2017-12-17_0015
4hour.2018-01-02_1200  daily.2017-12-30_0010  weekly.2017-12-24_0015
4hour.2018-01-02_1600  daily.2017-12-31_0010  weekly.2017-12-31_0015
```

UFS Snapshot

```
derek[/]  df -h
Filesystem        Size      Used     Avail Capacity   Mounted on
/dev/ad4s1a        70G       16G       48G     25%     /
devfs             1.0K      1.0K       0B    100%     /dev
derek[/]  sudo mount -u -o snapshot /.snap/snapshot /
derek[/]  df -h
Filesystem        Size      Used     Avail Capacity   Mounted on
/dev/ad4s1a        70G       16G       48G     25%     /
devfs             1.0K      1.0K       0B    100%     /dev
derek[~]  sudo mdconfig -a -t vnode -f /.snap/snapshot -u 1
WARNING: opening backing store: /.snap/snapshot readonly
derek[~]  sudo mount -r /dev/md1 /mnt
derek[~]  ls /mnt/
./         COPYRIGHT   compat@      ftp/        mnt/        sys@
../         bin/        dev/        home/       proc/       tmp/
.cshrc      boot/       dist/       lib/        rescue/     usr/
.profile    cdrom/      entropy     libexec/    root/       var/
.snap/      cdrom1/     etc/        media/      sbin/
derek[~]  sudo umount /mnt
derek[~]  sudo mdconfig -d -u 1
```

20

Appendix

Backup Media – By Storage (1)

- **By Storage category**
  - **Hard disk**
    - **SATA / SAS / SSD**
      - **120 ~ 450 MB /s**
    - **1 TB SATA3: NT 1,500**
    - **2 TB SATA3: NT 2,000**
    - **4 TB SAS: NT 9,000**
    - **256 G SSD: NT 2,500**
    - **Different "types"**
      - **RAID**
      - **Cold Archive**
  - CD/DVD R RW
    - CD
      - 6 ~ 8 MB/s
    - DVD
      - 8 ~ 15 MB/s
    - CD-R 0.7G: NT 6
    - DVD-R 4.7G: NT 9
    - DVD DL 8.5GB: NT 35
  - BD
    - 4x 18 MB/s, 12x 64 MB/x
    - 6x double-layer BD-R 50GB : NT 60

Backup Media – By Storage (2)

- Types
  - **DAT (Digital Audio Tape) 4mm tapes**
    - **DDS (Digital Data Storage), Minimal Error Rate, Higher Efficiency**
    - **DDS-4 (often used)**
      - **20/40GB(compressed), about NT 400.**
      - **1.0~3.0MB/s**
 - **Travan tapes**
    - **High Transfer Rate**
    - **Travan 40 (often used)**
      - **20/40GB(compressed), about NT 2000.**
      - **Up to 8.0MB/s**
  - **DLT (Digital Linear Tape)**
    - **High Capacity, Solid Reliability**
    - **Media**
      - **Max 800 GB, about NT 4000.**
      - **Speed: Up to 60 MB/s**
  - **LTO Ultrium**
    - **Fast Transfer Rate, High Performance, and High Storage Capacity**
    - **LTO Ultrium 3 (often used)**
      - **Max 1600 GB, about NT 5000.**
      - **Speed: up to 80 MB/s**
      - **Tape Drive is much more expensive……**
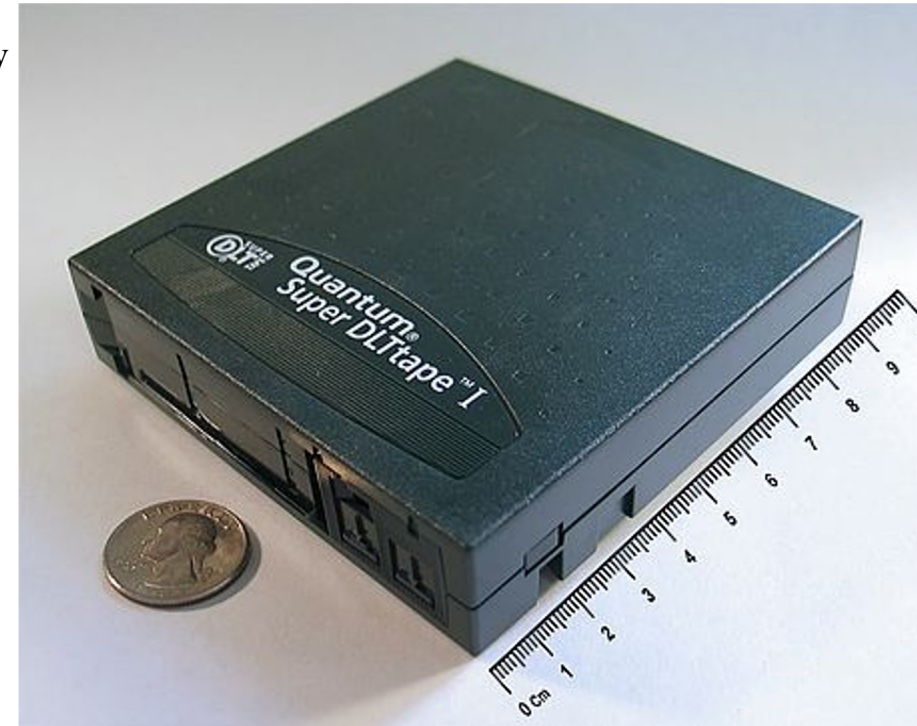
https://en.wikipedia.org/wiki/Digital_Linear_Tape

23

Backup Media – By Storage (3.1)

- Backup media compare

Medium	Capacity	Speed	Drive	Media	Cost/GB	Reuse?	Random?
CD-R	700MB	7MB/s	$15	15¢	21¢	No	Yes
CD-RW	700MB	4MB/s	$20	30¢	42¢	Yes	Yes
DVD ± R	4.7GB	30MB/s	$30	30¢	6¢	No	Yes
DVD + R DL	8.5GB	30MB/s	$30	$1	12¢	No	Yes
DVD ± RW	4.7GB	10MB/s	$30	40¢	9¢	Yes	Yes
Blu-ray	25GB	30MB/s	$100	$3	12¢	No	Yes
DDS-4 (4mm)	20GB	30MB/s	$100	$5	25¢	Yes	No
DLT/S-DLT	160GB	16MB/s	$500	$10	6¢	Yes	No
DLT-S4	800GB	60MB/s	$2,500	$100	13¢	Yes	No
AIT-4 (8mm)	200GB	24MB/s	$1,200	$40	20¢	Yes	No
AIT-5	400GB	24MB/s	$2,500	$50	13¢	Yes	No
VXA-320	160GB	12MB/s	$800	$60	38¢	Yes	No
LTO-3	400GB	80MB/s	$200	$25	6¢	Yes	No
LTO-4	800GB	120MB/s	$1,600	$40	5¢	Yes	No

Backup Media – By Storage (3.2)

- **MO (Magneto-Optical)**
  - MO 540M, 640M, 1.3G, 2.3G
- **Removable Media**
  - Floppy, ZIP, LS-120
- **Jukebox**
  - Automatically change removable media
    - DAT, DLT, CD, …
- **Tape Library**
  - Hardware backup solution for large data set

Backup Media – By Storage (4)

- Jukebox
  - Automatically change removable media
  - Available for several types of media
    - DAT, DLT, CD

| Specifications	
Number of Magazines (50-disc Magazine)	Max. 6 units (front: max. 4, rear: max. 3)
Number of Magazines (20-disc)	1
Number of Drives	Max. 8 drives
Disc Change Time	Max. 8 seconds

Backup Media – By Storage (5)

● Tape Library

機櫃定義	**L25** - TS1155、TS1150 及 TS1140（3592）磁碟機及磁帶匣的基礎機櫃 - 包含 32 個輸入/輸出（I/O）插槽 **D25** - 適用於 TS1155、TS1150 及 TS1140（3592）磁碟機及磁帶匣的可使用磁碟機及儲存擴充機櫃 **S25** - 適用於 3592 的擴充機櫃（僅限儲存設備） **L55** - 適用於 LTO 磁碟機和磁帶匣的基礎機櫃 - 包含 36 個 I/O 插槽 **D55** - 適用於 LTO 磁碟機和磁帶匣的可使用磁碟機及儲存擴充機櫃 **S55** - 適用於 LTO 磁帶匣的擴充機櫃（僅限儲存設備）
磁帶機類型	TS1155、TS1150 及 TS1140（3592）、LTO Ultrium 8、7、6 及 5 磁帶機
每個磁帶櫃的機櫃數量	一個基本機櫃,最多 17 個擴充機櫃,配備最多 7 個 Dx5 機櫃
磁碟機數目	每個機櫃多達 16 個（機櫃 1 中多達 12 個） 每個磁帶櫃字串中多達 128 個
磁帶匣數量	**L25** - 多達 660 個（機櫃 1 中多達 550 個） **D25** - 多達 740 個（機櫃 1 中多達 590 個） **S25** - 多達 1,000 個（機櫃 1 中多達 798 個） 每個磁帶櫃支援總數目:多達 17,550 個 **L55** - 多達 882 個（機櫃 1 中多達 730 個） **D55** - 多達 970 個（機櫃 1 中多達 774 個） **S55** - 多達 1,320 個（機櫃 1 中多達 1,054 個） 每個磁帶櫃支援總數目:最多 23,170 個
容量*	3592 進階磁帶匣:每個磁帶櫃高達 263.25 PB（採用 3:1 壓縮時為 789.75 PB） LTO Ultrium 8 磁帶匣:每個磁帶櫃高達 278 PB（採用 2.5:1 壓縮時,最高可達 695 PB）

Backup Media – By Enterprise Product (1)

● RAID architecture

SCSI cable

Fiber Channel

Server

Disk array
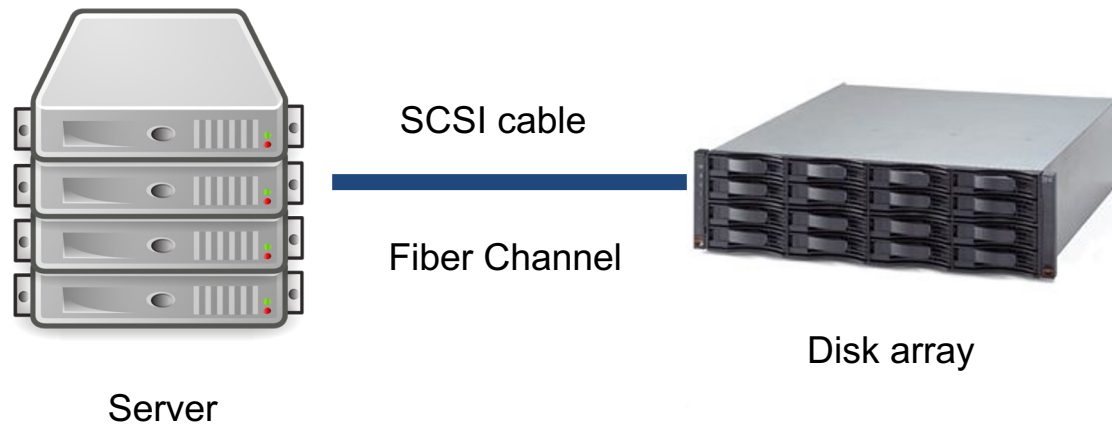
IBM TotalStorage DS6000 的目標：
- 以合理價格的儲存系統解決方案，為大中型企業提供高可用性
- 具有企業級功能、模組化、可擴充特性，能支援開放性平台與大型主機
- 提供進階複製服務，與 IBM TotalStorage DS8000 系列及 IBM TotalStorage Enterprise Storage Server® (ESS) 800 和 750 系統互通
- 提供 GUI 介面與「快捷組態 (Express Configuration)」精靈，透過隨附的 IBM TotalStorage DS Storage Manager 來簡化系統配置與管理
- 採用模組化、3U、16 個磁碟機、機架式，隨儲存需求而增，最高可達 67.2 TB 的實體容量

Backup Media – By Enterprise Product (2)

- NAS (Network Attached Storage)
  - Storage + Server + Cross-platform access OS + network access protocol

IBM NAS 300G
Supported Protocol:
NFS, HTTP, FTP, CIFS
Netware