IT Scripting and Automation

Backups

Lecturer: Art Ó Coileáin

Backups

Usually the information stored on computers is worth far more than the computers themselves. Protecting this information is one of the tasks of a system administrator.

Data Loss:

There are several ways to lose data:

- Software bugs routinely corrupt documents
- Users accidentally delete data files
- Hackers and disgruntled users erase disks
- Hardware problems
- Natural disasters

- Backups must be done carefully and on strict schedule.
- The backup system and backup media must also be tested regularly to verify that they are working correctly.

General Recommendations

- Perform all backups from a central location
  - Centralisation facilitates administration and lets you restore data to alternative servers.
  - Often, you can connect more than one media device on a server without impacting the performance.
  - If backing up across a network, the bandwidth becomes an issue. Consider creating a LAN dedicated to backup traffic.
- Label your media
  - Label each piece of backup media clearly and completely. Label to uniquely identify its contents.
  - List filesystems, backup dates, syntax of the commands used to create them.
  - Any other information that will help to restore the backups.

Backups (2)

- Pick a reasonable backup interval
  - The more often backups are done, the less data is lost in a crash.
  - An adequate data integrity must be provided at a *reasonable* cost of time and materials.
  - On busy systems is usually appropriate to back up home directories every workday.
  - On systems used less heavily performing backups several times a week is sufficient.

- Ask yourself: ***How much data are your users willing to lose?***

Backups (3)

- Choose filesystems carefully
  - Files that are rarely modified do not need to be backed up as frequently as users' home directories.
- Make daily dumps fit on one piece of media
  - If you can fit your daily backups on one tape: Buy a higher-capacity backup device, buy a stacker or library and feed multiple pieces of media to one device, change your dump sequence, use multiple backup devices.
- Keep media off-site
  - Always have an off-line copy of your data.
  - Snapshots and RAID arrays are not substitutes for real backups.
  - Some sites perform two dumps to different backup devices, one stays on-site and one is moved immediately.
- Protect your backups
  - Encryption of backup media is usually a good option.

Backups (4)

- Limit activity during backups
  - Filesystems activity should be limited during backups because changes can cause your backup utility to make mistakes (do it in the middle of night or on weekends).
  - Automatic process can be implemented (**cron** can help)
  - These days the only way to do a backup with no disk activity is to first create a snapshot.
- Develop a media life cycle
  - All media have a finite life. Follow manufacturer's recommendations regarding the life of the media.
  - Before you toss old tapes in the trash, remember to erase them.
- Design your data for backups
  - Start by taking an inventory of your storage needs: kind of data, expected volatility, backup frequency, network and political boundaries.
  - Use this information to design your site's storage architecture,

Backups (5)

Prepare for the worst:

- Explore the worst scenario: your site completely destroyed.
- Determine how much data would be lost and how long it would take to get your system back to life.
- Organisation often designate a Recovery Time Objective (**RTO**) and a Recovery Point Objective (**RPO**)
- **RTO** represents the maximum amount of time that the business can tolerate waiting for a recovery to complete
- **RPO** indicates how recent a backup is required for the restore.

Backups Devices and Media (inc. historical)

- Optical media: CD-R/RW, DVD+-R/RW, DVD-RAM, and Blu-ray
- Portable and removable hard disks
- Magnetic tapes in general
- Small tapes drives: 8mm and DDS/DAT
- DLT/S-DT: Digital Linear Tape/Super Digital Linear Tape is a mainstream backup medium.
- AIT and SAIT: Advanced Intelligent Tape is Sony's own 8mm product. SAIT is a larger media and has greater capacity.
- VXA/VXA-X: developed by Exabyte and acquired by Tandberg Data. Large capacity media with fast transfer rate.
- LTO: Linear Tape-Open developed by IBM, HP and Quantum.
- Jukeboxes, stackers, and tape libraries.
- Hard disks.
- Internet and cloud backup services.

Incremental Backups

- Almost all backup tools support at least two different kinds of backups: **Full** backup and **incremental** backup.
- **A full backup** includes all of a filesystem's contents.
- **An incremental backup** includes only files that have changed since the previous backup.
- The **dump** and **restore** commands are the most common way to create and restore from backups.
- You may have to explicitly install **dump** and **restore** on your Linux systems.
- **dump** builds a list of files that have been modified since a previous dump, then packs those files into a single large file to archive to an external device.

Example: **$ dump -0u -s 60000 -d 6250 -f /dev/nts0 /work**

Other Archiving Programs

tar:

- **tar**: Packages files.
- **tar** reads multiples files or directories and packages them into one file, often a tape device.
- **tar** is also useful for moving directory trees from place to place, especially if you are copying files as root. tar can preserve ownership and time information, but only if it is indicated.

Example:

- **$ tar -zcvf myarchive.tar.gz *folder_of_files_to_archive***
- ***Create (-c) a compressed (-z) archive file (-f) named myarchive.tar.gz and verbosely (-v) list the files as they are processed, from a specified directory or list of files. More information available from the manual page (man tar).***

Other Archiving Programs (2)

- **dd** is a file copying and conversion program. Unless you tell it to do some sort of conversion, **dd** just copies from its input file to its output file.
- **dd** is also a popular tool among forensic specialist because it creates a bit-for-bit unaltered copy of a volume, **dd** can be used to duplicate electronic evidence for use in court.
- With two tape drives, say, **/dev/st0** and **/dev/st1**, you'd use the command:

```
$ dd if=/dev/st0 of=/dev/st1 cbs=16
```

BACULA

- It is an enterprise-level client/server **backup solution** that manages backup, recovery, and verification of files over a network.
- The Bacula components run on Linux, Solaris, and FreeBSD. The Bacula client backs up data from many platforms, including Windows.
- https://www.bacula.org/
- http://blog.bacula.org/