Incremental Backup for Linux

Khalid Baheyeldin

May 2018

http://2bits.com

Agenda

- Why Backup?

- What is Backup (and what is not)?

- Incremental Backup

- Discussion

Why Backup?

- Hardware Lifetime

    - Specially moving parts: hard disk

- Human Errors

    - Deleting or overwriting files by mistake

    - Running SQL UPDATE without WHERE

- Disasters

    - Physical Theft

    - Fire, Lightening

- Increasingly Hostile Internet

    - Ransomware

- One (or more) safe copy of your *data* and *configuration*, for the purpose of recovering from certain events

Criteria

- Complete

  – Metadata, timestamps, ownerships, ...

- Portable

  – cpio on Motorola vs. Intel

- Immutable

  – Cannot be changed (e.g. ransomware)

- Off site

  – Disaster, theft

- Speed of recovery

- In Same Machine

  - Theft, disaster, human error

- Not Immutable

  - RAID, LVM, ZFS

  - File sync (e.g. Rsync, to the same remote copy. Human error)

- Not Versioned

Devices and Media

- Tapes
  - Used to be less cost per MB
  - Robotic Tape Library (rotation, eject for off site)
  - Nowadays: several standards
    - Though LTO is winning
  - Drives and media are expensive

Devices and Media

- USB Drives
    - Readily available
    - High CPU utilization
    - Speed not great with USB 2.0

Disk Dock

- With pair of high capacity disk drives

- Each used in rotation

- Keep one off site

- USB 3.0 most common now

- eSATA is what I use

- Dual disk models (second slow?)

- Two laptops

- Uses cron for daily rsync

- To the home server (/home/foo/laptop)

- The home server is incrementally backed up daily

Phone Photos

- FolderSync Android App

- Syncs to home server using SFTP

    – Only needs ssh enabled on server

- Once daily

- Only when on A/C power, and specific WiFi SSID is connected

Incremental Backup

- The **dump** and **restore** programs
- First appeared in AT&T UNIX V6 (late 1970s)
- Works at the file system level (not the file/directory level)
- Does not change read timestamp
- Supported in ext2, ext3, ext4 filesystems
- Supports 'Towers of Hanoi' scheme

Dump

- Both run from cron, on the home server

- Full Backup, Sunday:

  - dump -0 -z -f /mnt/disk1/yymmdd.dump /

- Incremental Backup (other days, files changed since lower level):

  - dump -1 -z -f /mnt/disk1/Monday.dump /

Dump

DUMP: Date of this level 0 dump: Sun Apr 29 04:01:02 2018

DUMP: Dumping /dev/sda1 (/) to /backup/0/backup-2018-04-29.dump

DUMP: Label: none

DUMP: Writing 64 Kilobyte records

DUMP: Compressing output at compression level 2 (zlib)

DUMP: mapping (Pass I) [regular files]

DUMP: mapping (Pass II) [directories]

DUMP: estimated 144277135 blocks.

DUMP: Volume 1 started with block 1 at: Sun Apr 29 04:04:14 2018

Dump

DUMP: dumping (Pass IV) [regular files]

DUMP: 2.51% done at 12049 kB/s, finished in 3:14

DUMP: 6.14% done at 14762 kB/s, finished in 2:32

...

DUMP: 98.67% done at 20630 kB/s, finished in 0:01

DUMP: 99.66% done at 19969 kB/s, finished in 0:00

DUMP: Closing /backup/0/backup-2018-04-29.dump

DUMP: Volume 1 completed at: Sun Apr 29 06:04:43 2018

DUMP: Volume 1 took 2:00:29

DUMP: Volume 1 transfer rate: 14691 kB/s

Dump

DUMP: Volume 1 144013248kB uncompressed, 106201815kB
  compressed, 1.357:1

DUMP: 144013248 blocks (140637.94MB) on 1 volume(s)

DUMP: finished in 7229 seconds, throughput 19921 kBytes/sec

DUMP: Date of this level 0 dump: Sun Apr 29 04:01:02 2018

DUMP: Date this dump completed: Sun Apr 29 06:04:43 2018

DUMP: Average transfer rate: 14691 kB/s

DUMP: Wrote 144013248kB uncompressed, 106201815kB
  compressed, 1.357:1

DUMP: DUMP IS DONE

Restore

- Interactive mode

  - Good for select files (recovery from limited human error)

- Batch mode (full restoration)

  - Used it when hard disk died on home server

Exclusions

- Movies recorded OTA to PVR

  – Raw size is too big

  – Transcoding necessary (6:1 reduction)

  – Rare changes, usually added to only

  – Separate pair of USB 3.0 disks

  – Manual process, with scripts

Exclusions

- Family Photos
  - Less frequent additions
  - Full backup using tar, weekly
  - Same disk as 'dump'

Exclusions

- Family Photos

2018-04-29 06:04:43: Renaming last exclude backup to prev

2018-04-29 06:04:45: Renaming last exclude complete

2018-04-29 06:04:45: Creating new exclude backup to: /backup/0/excludes.tar

tar: Removing leading `/' from member names

2018-04-29 08:10:22: Exclude backup complete

- Physical disk off premise (friend or relative)

  - Encryption

- Safe deposit box in a bank

Cloud Backup?

- Bandwidth intensive

- Can be expensive

- Recovery can be slow

- Not your computer ...
    - Can go out of business
    - Can leak information

Discussion

- Questions?

- What do *you* use for backup?