

Linux Backup Solutions

Christian Külker

2022-06-12

Contents

1	The 3-2-1 Rule	2
2	Methods	3
3	Storage Media	3
4	Data Selection	3
5	Custom Solutions	3
5.1	NFS + Tar + Cron	4
6	File System Snapshots	4
7	Backup Software	5
8	Details On Some Software	5
8.1	Amanda	6
8.2	Back In Time	6
8.3	BackupPC	6
8.4	Dirvish	6
8.5	Duplicity	6
8.6	Rsnapshot	7
8.7	Rdiff-backup	7
8.8	Slbackup	7
9	Links	7
10	History	8
11	Disclaimer of Warranty	8

12 Limitation of Liability 8

A data backup is a copy of computer data taken and stored somewhere else so that it may be used to restore the original after a data loss event.

Backups can be used to recover data after its loss on the original computer due to deletion or corruption, or backups can be used to recover data from an earlier time. The last case should be avoided and replaced with revision control systems.

There are different kind of backups depending the data to be backed up. A backup Of system data can be used to recover a system of well planned, called disaster recovery. While a backup of user data can recover errors (deletion for example) made by users.

No all backup systems can be uses as a disaster recovery. And sometimes the ability depends on the complexity of the system to backed up: Computer clusters and database server for example.

A backup system should only contain data considered worth saving and at least one copy of it. Some systems are able to provide more than one copy or can reconstruct data depending on a time stamp. They are suited for user data recovery.

Data storage requirements of backup systems can be very large. Due this some solutions overcome the price tag by offering expensive primary storage and less expensive secondary storage which often differs in retrieval speed. Examples are: SSD for primary storage and tapes for secondary storage.

Data to be stored needs to be selected. The selection can be up on time or location or other criteria. The backup systems has to support certain methods for different data types as: databases, open files, compression, encryption and de-duplication.

1 The 3-2-1 Rule

Every backup should be planned with a strategy in mind. The 3-2-1 rule states, there should be 3 copies of the data, stored on 2 different media types (of 2 different computers) and 1 copy should not be on site.

The reason why one should use at least 2 different computers is, that sometimes even hardware (not the media) might introduce errors to the backup. A defective raid controller might not only crash the file systems of spinning hard disks attach to it but also that of SSDs, for example.

2 Methods

- Unstructured
- Full only/ System images
- Incremental
- Near-CDP
- Reverse incremental
- Differential

3 Storage Media

- Magnetic tapes
- Hard disks (Spinning, SSD: via SCSI, USB, FireWire, eSATA, ethernet, iSCSI or fibre channel)
- optical storage
- Flash memory (thumb drives, compactFlash, SmartMedia, memory sticks via USB)
- Network storage (via NFS, ssh, rsync)

4 Data Selection

Usually the smallest entity to be backed up is a **file**. Files are organized via a **file system**. The data types are: **static** data (files, ...), **live** data (database, log file, ...) or **meta** data (boot sector, partition data, file modification data, ...).

For example, to back up the live data of a database it is often required to shut down the database to make the database file static, before making a backup.

5 Custom Solutions

Every Linux is shipped with backup tools.

- cp [files]
- tar [files]
- dump [file systems] (archaic)
- dd [block devices]
- cpio [files]

They can be used together with `crond` and/or `ssh` for example to create a small and reliable backup system.

5.1 NFS + Tar + Cron

For example, if you have mounted a directory of an external server via NFS to `/remote/server/dir` you can put this script under `/etc/cron.daily` to back up the `/etc` directory.

```
#!/usr/bin/bash

DIR=/remote/server/dir
FILE=$DIR/backup-etc-`date +%F`.tar.gz
if [ -d $DIR ]; then
    cd /
    if [ -f $FILE ]; then
        echo "ERROR: File $FILE exists"
        exit 2
    else
        /bin/tar czf $FILE etc
    fi
else
    echo "ERROR: No directory $DIR"
    exit 3
fi
```

This can of course **not** be used to make a disaster recovery of the system, but parts of the system configuration could be restored.

6 File System Snapshots

If the target is to backup a partition (which might be questionable), so called snapshots can be used for certain file systems. The following list should be considered with caution as it do not claim that the mentioned file systems are stable. Sometimes the can be done via file system specific `fsck` tools, external commands or via an interim layer, like `LVM`.

- btrfs
- zfs
- lvm (ext3, xfs)

Software that can handle LVM:

- [Snapper and LVM thin-provisioned Snapshots](#)

Software that sounds like a snapshot, but isn't:

- [rsnapshot](#) - uses `rsync` and hard links. Debian-Edu (Skolelinux) use [rsnapshot](#)

for backups in school environments.

7 Backup Software

While it is advisable to create one's own backup systems for young system administrators to understand the underlying problems of backups, there are free open source software already provided to Linux that can be used, if one understands the core principles of backups.

Package	Debian	Interface	Debian Buster	Bullseye
Amanda		CLI	1:3.3.9-5	1:3.5.1-7
Areca Backup	n.a.			
BackupPC		GUI	3.3.1-4	4.4.0-3
Back In Time		GUI	1.1.12-2	1.2.1-3
Bacula			7.4.4+dfsg-6+deb9u2	9.6.7-3
Bareos	Buster	GUI	16.2.4-3+deb9u2	1)
Cpio		CLI	2.11+dfsg-6	2.13+dfsg-4
dirvish		CLI	1.2.1-1.3	1.2.1-2.1
duplicity	Bullseye			0.8.17-1+b1
rdiff-backup			1.2.8-7	2.0.5-2
rsbackup			3.1-3+b1	6.0-2+b2
rsnapshot	Buster	CLI	1.4.2-1	2)
slbackup		CLI	0.0.12-8	0.0.12-13
snapper		CLI	0.4.1-3	0.8.15-1
Warewulf	n.a.	CLI		
Tar		CLI	1.29b-1.1	1.34+dfsg-1

- 1) Removed at 2021-12-01 due to maintainer [inactivity](#) and [9 RC bugs](#). See also [tracker](#)
- 2) Was removed at [2021-04-11](#) due to bug [#986709](#) that tells that the upstream maintainer do not intend continue to maintain it.

8 Details On Some Software

8.1 Amanda

- [amanda](#)
- Organization: full + incremental
- Storage: tape: tar, dump,

8.2 Back In Time

- [back in time](#)
- Organization: snapshots
- Storage: disk
- Dependencies: rsync

8.3 BackupPC

- [backuppc](#)
- Organization: full + incremental (hardlink shared between servers)
- Restore: tar, http, zip
- Remote: rsync, tar (over ssh/rsh/nfs) or ftp, samba, rsyncd
- Dependencies: Perl, Apache2

Even though the acronym PC is used, it seems that this tool can be used not only for desktop or personal computers, but also for servers.

```
1 + easy install (Debian Squeeze, App + Webserver)
2 + Webinterface easy to see
3
4 - Webinterface has bugs (not choosing of host)
5 - Webinterface no choice of target dir
6 - Webinterface no creation of host cfg
7 - Webinterface no comparison of files?
```

8.4 Dirvish

- [dirvish](#)
- Organization: rotating images (bank->vault->image)
- Restore: rsync, scp, cpio, tar, ...
- Storage: disk
- Dependencies: rsync, Perl, ...

8.5 Duplicity

- [duplicity](#)

- Organization: full backup + incremental
- Restore: GnuPG, rdiff, and tar.
- Remote: scp/ssh, ftp, rsync, HSI, WebDAV, Tahoe-LAFS, and Amazon S3
- Storage: encrypted tar format
- Dependencies: Python, librsync, GnuPG,...

8.6 Rsnapshot

- [rsnapshot](#)
- Organization: full + incremental (daily, weekly)
- Restore: cp
- Dependencies: Perl, rsync or librsync?

8.7 Rdiff-backup

- [rdiff-backup](#)
- Organization: server/client full backup + incremental
- Remote: ssh
- Dependencies: librsync, python

8.8 Slbackup

- [slbackup](#)
- Uses [rdiff-backup](#)
- Used by Skolelinux/ Debian-Edu

9 Links

- [amanda](#)
- [areca backup](#)
- [areca backup wikipedia](#)
- [backupper](#)
- [backupper wikipedia](#)
- [back in time](#)
- [bacula](#)
- [bareos](#)
- [bareos home](#)
- [bareos open source](#)
- [bareos github](#)
- [bareos documentation](#)

- [cpio](#)
- [dirvish](#)
- [duplicity](#)
- [duplicity gitlab](#)
- [rdiff-backup](#)
- [rsnapshot](#)
- [rsbackup](#)
- [slbackup](#)
- [snapper](#)
- [warewulf](#)
- [warewulf wikipedia](#)

10 History

Version	Date	Notes
0.1.3	2022-06-12	Add reason of package removal from Bullseye
0.1.2	2022-06-11	Shell -> bash, Bullseye
0.1.1	2021-06-08	Typos, Slbackup, Details, Dirvish, Duplicity
0.1.0	2021-06-07	Initial release

11 Disclaimer of Warranty

THERE IS NO WARRANTY FOR THIS INFORMATION, DOCUMENTS AND PROGRAMS, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE INFORMATION, DOCUMENT OR THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE INFORMATION, DOCUMENTS AND PROGRAMS IS WITH YOU. SHOULD THE INFORMATION, DOCUMENTS OR PROGRAMS PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

12 Limitation of Liability

IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MODIFIES AND/OR CONVEYS THE INFORMATION, DOCUMENTS OR PROGRAMS AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE INFORMATION, DOCUMENTS OR PROGRAMS (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE INFORMATION, DOCUMENTS OR PROGRAMS TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.