



# Faking a Snapshotting Filesystem

Providing user access to older versions of their files

Peter Edwards | Sr Systems Administrator  
DMF Users Group | February 2015

CSIRO IM&T SCIENTIFIC COMPUTING  
[www.csiro.au](http://www.csiro.au)



# Snapshots on NetApp filers

Snapshots are scheduled every four hours and users can see older versions of their files in a `~/snapshot` directory.

This allows:

- Self-service file restoration after accidents (with “cp”)
- Visibility of when files were altered (“ls -l”)
- Visibility of what changed (“diff”)

# NetApp structure

```
machine$ ls -l .bashrc
```

```
-rw----- 1 joe joe 3533 May 28 13:01 .bashrc <--- "live" file
```

```
machine$ ls -l .snapshot/*/bashrc
```

```
-rw----- 1 joe joe 3190 May 28 11:53 .snapshot/hourly.0/.bashrc  
-rw----- 1 joe joe 3285 May 27 23:37 .snapshot/hourly.1/.bashrc  
-rw----- 1 joe joe 3314 May 27 19:36 .snapshot/hourly.2/.bashrc  
-rw----- 1 joe joe 3348 May 27 15:07 .snapshot/hourly.3/.bashrc  
-rw----- 1 joe joe 3398 May 27 11:02 .snapshot/hourly.4/.bashrc  
-rw----- 1 joe joe 3388 May 26 20:11 .snapshot/hourly.5/.bashrc  
-rw----- 1 joe joe 3285 May 27 23:37 .snapshot/nightly.0/.bashrc  
-rw----- 1 joe joe 3388 May 26 20:11 .snapshot/nightly.1/.bashrc
```

# CSIRO rsync-based backups

CSIRO rsync-based “external backups” use a DMF-managed disk back-end, using a forest of directory trees containing hard-linked files. For example shiraz:/home/edw192/.bashrc is backed up to:

/backup/	private directory (no public access)
3/	mountpoint of DMF-managed FS
shiraz-dumps/	client name
home.20141129.seq.125/	timestamped and versioned FS name
edw192/	username
.bashrc	user files & directories

(repeat all but first as required)

Files (.bashrc here) in different trees are hard-linked together if identical.

# Exporting a read-only view

- On the NFS server (“datastore”), we set up a read-only bind mount  
/backup\_ro/3
- UNIX permissions of a filesystem root apply to all mounts, so to set different permissions on different mounts, need to do so on the parents of the mountpoints:
  - /backup should be administrator-only to block user access to the writable version.
  - /backup\_ro should be world-access to allow user access to the read-only one.
- On the NFS client (“shiraz”) use the automounter to mount as  
/nfs/datastore/backup\_ro/3

# Per-user access

Then for each user, create a ~/.Snapshots directory containing symlinks with names of the form YYYYMMDD pointing to the username entry in the above hierarchy.

```
shiraz$ ls -l ~edw192/.Snapshots/
```

```
...
```

```
lrwxrwxrwx 1 root root 77 Nov 30 18:38 20141129 -> /nfs/datastore/backup_ro/3/shiraz-dumps/home.20141129.seq.125/edw192
```

```
lrwxrwxrwx 1 root root 77 Nov 30 18:38 20141130 -> /nfs/datastore/backup_ro/3/shiraz-dumps/home.20141130.seq.126/edw192
```

```
-rw-r--r-- 1257 root asc 2025 Nov 30 18:37 README
```

```
shiraz$ dmls -l ~/.Snapshots/20141129/.bashrc
```

```
-rw-r--r-- 8 edw192 asc 7917 2014-10-25 10:35 (OFL) /home/edw192/.Snapshots/20141129/.bashrc
```

# Notes from README file

- This access is read-only - the snapshot files cannot be deleted or altered.
- These snapshots are unavailable when the server (“Cherax”) is down for any reason.
- The files are managed by DMF, just like the Data Store. “dmls -l” will show that all but the smallest files will be offline; “dmget” should be used to recall them to disk.
- Your .Snapshots directory will be regenerated each day to track changes inside the backup system.
- You will see that there is a link for each of the previous five days, but further back in time the coverage is sparser. This is a characteristic of the way we do backups.

# And for the sysadmins too!

- As well as making .Snapshots subdirectories in the users' home directories, for the convenience of the administrators they can be created in system directories as well, such as
  - /etc
  - /usr/local



# Assessment

- Costs:
  - Time to recreate all ~/.Snapshots directories after each backup (negligible)
- Pros:
  - Secure self-serve file restorations
  - Ability to compare different versions of a file for debugging
- Cons:
  - None

# CSIRO non-rsync-based backups

On the server (“Cherax”), users’ home directories are in a DMF-managed hybrid filesystem /datastore which is backed up with xfsdump, so there is no forest of directory trees to access.

So create one!

# Create a forest of trees

Create new subdirectory `/datastore/Snapshots.d` with world-execute permission, but not read, and create a similar hierarchy as above. The subdirectories are real, and require an inode and storage, but the files are hard-links to the user's files, so no extra space or inodes required there, and no DMF recalls.

This linking is done with (simplified):

```
rsync --archive --exclude=/Snapshots.d --link-dest=/datastore \  
/datastore/ /datastore/Snapshots.d/20141202
```

Then create per-user `~/.Snapshots` directories as above.

# Limitations

- These snapshots work as expected when you create a new file or delete an existing one. But if you alter (chmod, overwrite or extend) a file in your home directory, all the “copies” under ~/.Snapshots will also change, and vice versa. If you don't like that, then you should delete, or copy & move, the original first.
- As the created directories reside in the SSD partition of a hybrid filesystem, a limit on how many daily snapshots are retained is necessary (eg: one week).
- Deleting old files to avoid a quota limit doesn't work until that week has elapsed.

# Notes from README file

- These snapshots work as expected when you create a new file or delete an existing one. But if you alter (chmod, overwrite or extend) a file in your home directory, all copies under ~/.Snapshots are also altered. If you don't like that, then you should delete or copy/move the original first. For example:

To overwrite myfile without changing the snapshots:

```
rm -f myfile  
a.out > myfile
```

To extend myfile without changing the snapshots:

```
cp -p myfile myfile.tmp && mv myfile.tmp myfile  
a.out >> myfile
```

Similarly, all copies share any new timestamps, permissions and ownership. The copy/move procedure above will work around that if necessary.

- As far as quotas are concerned, your inode usage will be increased by 1 for every directory you have and by 1 for every altered file (see above).
- The files are managed by DMF, just like the Data Store. “dmls -l” will show that all but the smallest files will be offline; “dmget” should be used to recall them to disk.
- Your .Snapshots directory will be regenerated each day to track changes inside the filesystem.
- Only a set number of daily snapshots are kept, to minimise the space they take up.

# Assessment

- Costs:
  - Time to create a new directory tree each night, and retire the oldest one (about a minute for 18 users; speed-up techniques not used yet)
  - Time to recreate all `~/Snapshots` directories afterwards (negligible)
  - Space in SSD to hold inodes and directories for these trees (hard to measure)
- Pros:
  - Self-serve file restorations
  - Ability to compare different versions of a file for debugging
  - DMF compatible
- Cons:
  - File alterations not handled intuitively, so offer to users only on an opt-in basis
  - Only have a week or so of equally spaced snapshots
  - Time to hard-delete a migrated file extended by that week or so
  - Impact on user inode usage wrt quotas (small)

# References

- CSIRO rsync-based backups:

[http://hpsc.csiro.au/users/dmfug/Meeting\\_Oct2009/Presentations/Using\\_DM\\_Filesystem\\_as\\_target\\_for\\_backups.pdf](http://hpsc.csiro.au/users/dmfug/Meeting_Oct2009/Presentations/Using_DM_Filesystem_as_target_for_backups.pdf)

- Rsync:

“man rsync”

- Hybrid filesystems:

[http://hpsc.csiro.au/users/dmfug/Meeting\\_Feb2015/Presentations/CSIRO\\_SCSI\\_Hybrid\\_filesystems\\_revisited.pdf](http://hpsc.csiro.au/users/dmfug/Meeting_Feb2015/Presentations/CSIRO_SCSI_Hybrid_filesystems_revisited.pdf)

# Thank you

## **CSIRO IM&T Scientific Computing**

Peter Edwards

Sr Systems Administrator

**t** +61 3 8601 3812

**e** peter.edwards@csiro.au

**w** <https://wiki.csiro.au/display/ASC/Scientific+Computing+Homepage>

**CSIRO IM&T SCIENTIFIC COMPUTING**

[www.csiro.au](http://www.csiro.au)

