



Reva on CephFS

Theofilos Mouratidis, Storage Engineer, CERN Storage Group

Background

→ Reva:

- ◆ Interoperability platform for various API's
- ◆ Supports Sync & Share API for user's files
- ◆ Provides file access via two methods:
 - Dropbox style client
 - Direct filesystem mount for linux users
- ◆ Currently supports local FS (XFS/EXT3), EOS, ownCloud and S3 storage backends

→ Ceph:

- ◆ Open-source object store for scalable applications
- ◆ Provides object storage, block storage and POSIX-like file system (CephFS) in one solution
- ◆ Popular storage choice for organizations and universities

Project Goal

- Determine whether CephFS is capable to support Reva as a backend
- Add Reva's Sync&Share functionality in one of our CephFS Clusters
- Implementing an additional backend on Reva will also broaden its audience
- Create a proof-of-concept Reva module for CephFS

Proof-of-concept

- We run the local FS module on a CephFS mount
 - ◆ Use CephFS features that Reva requires through CephFS extended attributes
 - ctime propagation → faster file difference discovery
 - snapshots → fast way to keep old versions of files
- Current implementation ([github PR](#))
 - ◆ Uses localhome module on a CephFS mount with a simple configuration
 - ◆ Etag is calculated as:
 - ctime+inode number for files
 - rctime+inode number for directories
 - ◆ File versions based on fixed time interval snapshots
 - CephFS can only snapshot directories, not individual files

Manual sync test case

- One VM for Reva backend
- One VM using Ubuntu and Cernbox sync-client v2.5
- Untar linux kernel source code on the sync folder
- Run cernboxcmd to sync with backend
- Change two files in one random location and the one with the biggest path depth
 - ◆ linux/drivers/gpu/drm/amd/include/asic_reg/nbio/nbio_7_0_smn.h
 - ◆ linux/tools/testing/selftests/rcutorture/formal/srcu-cbmc/empty_includes/uapi/linux/types.h
- Run sync again
- CephFS backend found the differences immediately without checking the whole tree

Backend profind to find the different files

```
[root@cephdev8 ~]# grep -P '2021-01-19 09:4[67]' /tmp/reva.log | grep PROPFIND | grep -Po "url=.*"
url=/remote.php/webdav/
url=/remote.php/webdav/linux
url=/remote.php/webdav/linux/drivers
url=/remote.php/webdav/linux/drivers/gpu
url=/remote.php/webdav/linux/drivers/gpu/drm
url=/remote.php/webdav/linux/drivers/gpu/drm/amd
url=/remote.php/webdav/linux/drivers/gpu/drm/amd/include
url=/remote.php/webdav/linux/drivers/gpu/drm/amd/include/asic_reg
url=/remote.php/webdav/linux/drivers/gpu/drm/amd/include/asic_reg/nbio
url=/remote.php/webdav/linux/tools
url=/remote.php/webdav/linux/tools/testing
url=/remote.php/webdav/linux/tools/testing/selftests
url=/remote.php/webdav/linux/tools/testing/selftests/rcutorture
url=/remote.php/webdav/linux/tools/testing/selftests/rcutorture/formal
url=/remote.php/webdav/linux/tools/testing/selftests/rcutorture/formal/srcu-cbmc
url=/remote.php/webdav/linux/tools/testing/selftests/rcutorture/formal/srcu-cbmc/empty_includes
url=/remote.php/webdav/linux/tools/testing/selftests/rcutorture/formal/srcu-cbmc/empty_includes/uapi
url=/remote.php/webdav/linux/tools/testing/selftests/rcutorture/formal/srcu-cbmc/empty_includes/uapi/linux
url=/remote.php/webdav/MyShares
[root@cephdev8 ~]# |
```

What's next?

- Home directories and Shares
- ACLs
 - ◆ CephFS supports POSIX ACLs from linux kernel v3.10 and on
- Recycle bin for files
 - ◆ The snapshots can be used to implement both a recycle bin and file versions
- Pass all Smashbox tests
 - ◆ Smashbox is a test suite for ownCloud/Cernbox
 - ◆ We currently pass some basic sync tests
- Performance improvements
 - ◆ Direct golang api access to CephFS
 - ◆ Remove unnecessary functionality from local FS (sqlite DB for example)
- Review file level snapshots with the Ceph devs for implementation

Questions?

Thanks!

