<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migration of CVMFS stratum 0 from Netapp to Pure CEPH</td>
<td>1</td>
</tr>
<tr>
<td>Impact to the Wider User community</td>
<td>2</td>
</tr>
<tr>
<td>Early Thoughts, Ideas and Wishlist</td>
<td>2</td>
</tr>
<tr>
<td>Doubling Up</td>
<td>2</td>
</tr>
<tr>
<td>Stop Using cvmfs_config server</td>
<td>3</td>
</tr>
<tr>
<td>Need to actually look into the back up stuff</td>
<td>3</td>
</tr>
<tr>
<td>Puppet</td>
<td>3</td>
</tr>
<tr>
<td>First Candidate</td>
<td>3</td>
</tr>
<tr>
<td>Magic Failover</td>
<td>3</td>
</tr>
<tr>
<td>Steps</td>
<td>3</td>
</tr>
<tr>
<td>Extras</td>
<td>3</td>
</tr>
</tbody>
</table>
Migration of CVMFS stratum 0 from Netapp to Pure CEPH.

Status on July 9th 2015

todo

lxcvmfs37.cern.ch: * The CVMFS Stratum 0 for repo /cvmfs/atlas.cern.ch.
lxcvmfs38.cern.ch: * The CVMFS Stratum 0 for repo /cvmfs/alice.cern.ch.
lxcvmfs41.cern.ch: * The CVMFS Stratum 0 for repo /cvmfs/lhcb.cern.ch.
lxcvmfs43.cern.ch: * The CVMFS Stratum 0 for repo /cvmfs/alice-ocdb.cern.ch.

in progress

lxcvmfs40.cern.ch: * The CVMFS Stratum 0 for repo /cvmfs/cms.cern.ch.

done

lxcvmfs52.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/lhcbdev.cern.ch.
lxcvmfs54.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/cvmfs-config.cern.ch.
lxcvmfs55.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/sft.cern.ch.
lxcvmfs56.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/ams.cern.ch.
lxcvmfs58.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/test.cern.ch
lxcvmfs60.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/aleph.cern.ch.
lxcvmfs62.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/grid.cern.ch.
lxcvmfs65.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/fcc.cern.ch
lxcvmfs66.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/ganga.cern.ch
lxcvmfs67.cern.ch: The CVMFS Stratum 0 for repo /cvmfs/opal.cern.ch

Motivation: The Netapp Filer warranty runs out on of July 31st 2015 - time to stop using it.

We currently have for each CvmFS repository an openstack VM which has main storage on a netapp volume. In addition there is typically two ceph volumes for storing each transaction change for each CvmFS publication. e.g cvmfs-cms.cern.ch has volumes:

- **CVMFS-nfs01.cern.ch:/vol/CVMFS/cms** mounted on /srv/cvmfs, this is the main release files that are published to everyone.
- **ceph vdb** mounted on /var/spool/cvmfs, this is the runtime directory for preparing releases.
• **ceph vdc** mounted on `/var/spool/cvmfs/cms.cern.ch/cache`, this is a normal cvmfs cache for the existing released files.

The intention is to migrate to a Virtual Machine with one CEPH volume for each repository files dropping netapp from the system, i.e

• **ceph vdc** for `/var/spool/cvmfs/cms.cern.ch`

This directory contains the normal spool data as well as

* `/var/spool/cvmfs/cms.cern.ch/home` the home directory of the shared user. *
* `/var/spool/cvmfs/cms.cern.ch/cms.cern.ch` is the old `/srv/cvmfs/cms.cern.ch` directory.

The need for the second ceph volume in the previous case is no longer needed due to a kernel fix.

For the webserving part of the stratum 0, cvmfs-stratum-zero.cern.ch is currently two apache sitting atop the NFS mounted netapp `/srv/cvmfs` volumes. Instead each of these two apaches would be made to be reverse proxy back to each of individual stratum 0s (The -1 stratum server).

Once all repositories are migrated we can consider instructing global stratum ones to all reconfigure to use the new stratum -1s and we burn the existing stratum 0 zero web servers so the -1++.

### Impact to the Wider User community.

- For the WLCG CvmFS readers everything should be transparent, they never talk to the stratum zero ever anyway
- For software installers there will inevitability be a down time while at least a final rsync of data is done from netapp to ceph. Hopefully the data will not need to be reprocessed in any way but this is untested currently.
- It is perfectly possible to do each repository one at time starting with the smaller/less important repos and working to the LHC ones.

### Early Thoughts, Ideas and Wishlist

Having ran a stratum Zero 2.1 for 6 months there is also opportunity to improve the service with migration.

### Doubling Up

I'd like to support more than one stratum 0 per node. We currently have 10 nodes out of 14 or so that really do absolutely nothing, maybe 2 things just one thing a year.

Clearly one node for our main LHC customers but e.g belle, boss, .. should be able to coexist on a node with no impact on them.

The aim would be that we could break apart or move together stratum zeros as required.

We would almost certainly need to avoid uid/gid clashes between files of different repositories. A quick check and very new rsyncs have a `--usermap` flag which looks to be the magic for this job.

`cvmfs_server` this supports this, some work is needed in the puppet module.

### Impact to the Wider User community.
Stop Using cvmfs_config server

I’ll probably stop using the cvmfs_config server script and move to doing the configuration by hand (via puppet).

Need to actually look into the back up stuff

Need to automate a backup, if can be synced with the release process so much the better.

A precise backup policy must be written.

Puppet

I’ll probably need to rewrite much the puppet stuff since multi VO, stop using cvmfs_config server, and I probably just know more now as well about how to run a cvmfs_server 2.1.0.

First Candidate

I’ll do a test first of course but fcc would be my first candidate given a free choice.

Magic Failover

Can I keep a standby stratum0 to become one of the others automatically. Seems easy in principal. Doing it automatically is always slightly more scary.

Steps

List of steps to be done.

- Install a small puppet managed VM with two small fake repositories with ceph backed.
- Backups
- Install a big puppet managed VM with big fake repositories.
- Test it
- Test migration times.
- Write a schedule
- Do It.

Extras

IOPS Plots -
https://filer-carbon.cern.ch/graphlot/?from=-1month&until=-0hour&target=netapp.nfs01-1.qtree.CVMF*.*.ops

-- SteveTraylen - 2015-01-16