CEPHFS BEST PRACTICES

This guide provides recommendations for best results when deploying CephFS.

For the actual configuration guide for CephFS, please see the instructions at Ceph Filesystem.

WHICH CEPH VERSION?

Use at least the Jewel (v10.2.0) release of Ceph. This is the first release to include stable CephFS code and fsck/repair tools. Make sure you are using the latest point release to get bug fixes.

Note that Ceph releases do not include a kernel, this is versioned and released separately. See below for guidance of choosing an appropriate kernel version if you are using the kernel client for CephFS.

MOST STABLE CONFIGURATION

Some features in CephFS are still experimental. See Experimental Features for guidance on these.

For the best chance of a happy healthy filesystem, use a **single active MDS** and **do not use snapshots**. Both of these are the default.

Note that creating multiple MDS daemons is fine, as these will simply be used as standbys. However, for best stability you should avoid adjusting `max_mds` upwards, as this would cause multiple MDS daemons to be active at once.

WHICH CLIENT?

The FUSE client is the most accessible and the easiest to upgrade to the version of Ceph used by the storage cluster, while the kernel client will often give better performance.

The clients do not always provide equivalent functionality, for example the fuse client supports client-enforced quotas while the kernel client does not.

When encountering bugs or performance issues, it is often instructive to try using the other client, in order to find out whether the bug was client-specific or not (and then to let the developers know).

WHICH KERNEL VERSION?

Because the kernel client is distributed as part of the linux kernel (not as part of packaged ceph releases), you will need to consider which kernel version to use on your client nodes. Older kernels are known to include buggy ceph clients, and may not support features that more recent Ceph clusters support.

Remember that the “latest” kernel in a stable linux distribution is likely to be years behind the latest upstream linux kernel where Ceph development takes place (including bug fixes).

As a rough guide, as of Ceph 10.x (Jewel), you should be using a least a 4.x kernel. If you absolutely have to use an older kernel, you should use the fuse client instead of the kernel client.

This advice does not apply if you are using a linux distribution that includes CephFS support, as in this case the distributor will be responsible for backporting fixes to their stable kernel: check with your vendor.

REPORTING ISSUES

If you have identified a specific issue, please report it with as much information as possible. Especially important information:

- Ceph versions installed on client and server
- Whether you are using the kernel or fuse client
- If you are using the kernel client, what kernel version?
- How many clients are in play, doing what kind of workload?
- If a system is ‘stuck’, is that affecting all clients or just one?
- Any ceph health messages

- Any backtraces in the ceph logs from crashes

If you are satisfied that you have found a bug, please file it on the tracker. For more general queries please write to the ceph-users mailing list.