EXPERIMENTAL FEATURES

CephFS includes a number of experimental features which are not fully stabilized or qualified for users to turn on in real deployments. We generally do our best to clearly demarcate these and fence them off so they cannot be used by mistake.

Some of these features are closer to being done than others, though. We describe each of them with an approximation of how risky they are and briefly describe what is required to enable them. Note that doing so will *irrevocably* flag maps in the monitor as having once enabled this flag to improve debugging and support processes.

INLINE DATA

By default, all CephFS file data is stored in RADOS objects. The inline data feature enables small files (generally <2KB) to be stored in the inode and served out of the MDS. This may improve small-file performance but increases load on the MDS. It is not sufficiently tested to support at this time, although failures within it are unlikely to make non-inlined data inaccessible

Inline data has always been off by default and requires setting the “inline_data” flag.

MANTLE: PROGRAMMABLE METADATA LOAD BALANCER

Mantle is a programmable metadata balancer built into the MDS. The idea is to protect the mechanisms for balancing load (migration, replication, fragmentation) but stub out the balancing policies using Lua. For details, see Mantle.

SNAPSHOTS

Like multiple active MDSes, CephFS is designed from the ground up to support snapshotting of arbitrary directories. There are no known bugs at the time of writing, but there is insufficient testing to provide stability guarantees and every expansion of testing has generally revealed new issues. If you do enable snapshots and experience failure, manual intervention will be needed.

Snapshots are known not to work properly with multiple filesystems (below) in some cases. Specifically, if you share a pool for multiple FSes and delete a snapshot in one FS, expect to lose snapshotted file data in any other FS using snapshots. See the CephFS Snapshots page for more information.

Snapshots are known not to work with multi-MDS filesystems.

For somewhat obscure implementation reasons, the kernel client only supports up to 400 snapshots (http://tracker.ceph.com/issues/21420).

Snapshotting was blocked off with the “allow_new_snaps” flag prior to Firefly.

MULTIPLE FILESYSTEMS WITHIN A CEPH CLUSTER

Code was merged prior to the Jewel release which enables administrators to create multiple independent CephFS filesystems within a single Ceph cluster. These independent filesystems have their own set of active MDSes, cluster maps, and data. But the feature required extensive changes to data structures which are not yet fully qualified, and has security implications which are not all apparent nor resolved.

There are no known bugs, but any failures which do result from having multiple active filesystems in your cluster will require manual intervention and, so far, will not have been experienced by anybody else – knowledgeable help will be extremely limited. You also probably do not have the security or isolation guarantees you want or think you have upon doing so.

Note that snapshots and multiple filesystems are *not* tested in combination and may not work together; see above.

Multiple filesystems were available starting in the Jewel release candidates but were protected behind the “enable_multiple” flag before the final release.

PREVIOUSLY EXPERIMENTAL FEATURES

DIRECTORY FRAGMENTATION

Directory fragmentation was considered experimental prior to the *Luminous* (12.2.x). It is now enabled by default on new filesystems. To enable directory fragmentation on filesystems created with older versions of Ceph, set the `allow_dirfrags` flag on the filesystem:

```
ceph fs set <filesystem name> allow_dirfrags 1
```

MULTIPLE ACTIVE METADATA SERVERS

Prior to the *Luminous* (12.2.x) release, running multiple active metadata servers within a single filesystem was considered experimental. Creating multiple active metadata servers is now permitted by default on new filesystems.

Filesystems created with older versions of Ceph still require explicitly enabling multiple active metadata servers as follows:

```
ceph fs set <filesystem name> allow_multimds 1
```

Note that the default size of the active mds cluster (`max_mds`) is still set to 1 initially.