CephFS clients first contact a MON to authenticate and retrieve the cluster map. Then, the client queries an active MDS for file metadata. The client uses the metadata to access the objects that comprise the requested file or directory by communicating directly with the OSDs.

MDS features and configuration options are described in the following list:

MDS Ranks
MDS ranks define how the metadata workload is distributed over the MDS daemons. The number of ranks, which is defined by the `max_mds` configuration setting, is the maximum number of MDS daemons that can be active at a time. MDS daemons start without a rank and the MON daemon is responsible for assigning them a rank.

Subvolumes and Subvolume Groups
CephFS subvolumes are an abstraction for independent CephFS file system directory trees. When creating subvolumes, you can specify more fine-grained rights management, such as the UID, GID, file mode, size, and the subvolume group for your subvolume. Subvolume groups are abstractions at a directory level across a set of subvolumes.

Note
You can create snapshots of subvolumes, but Red Hat Ceph Storage 5 does not support creating snapshots of subvolume groups. You can list and remove existing snapshots of subvolume groups.

File System Affinity
Configure your CephFS file system to prefer one MDS over another MDS. For example, you can configure to prefer an MDS that runs on a faster server over another MDS that runs on an older server. This file system affinity is configured through the `mds_join_fs` option.

MDS Cache Size Limits
Limit the size of the MDS cache by limiting the maximum memory to use with the `mds_cache_memory_limit` option, or by defining the maximum number of inodes with the `mds_cache_size` option.

Quotas
Configure your CephFS file system to restrict the number of bytes or files that are stored by using quotas. Both the FUSE and kernel clients support checking quotas when mounting a CephFS file system. These clients are also responsible for stopping writing data to the CephFS file system when the user reaches the quota limit. Use the `setfattr` command's `ceph.quota.max_bytes` and `ceph.quota.max_files` options to set the limits.

New CephFS Capabilities

Red Hat Ceph Storage 5 removes limitations from earlier versions.

- Red Hat Ceph Storage 5 supports more than one active MDS in a cluster, which can increase metadata performance. To remain highly available, you can configure additional standby MDSes to take over from any active MDS that fails.
- Red Hat Ceph Storage 5 supports more than one CephFS file system in a cluster. Deploying more than one CephFS file system requires running more MDS daemons.

Deploying CephFS

To implement a CephFS file system, create the required pools, create the CephFS file system, deploy the MDS daemons, and then mount the file system. You can manually create the pools,