

MDS CONFIG REFERENCE

mon force standby active

Description: If true monitors force standby-replay to be active. Set under [mon] or [global].
Type: Boolean
Default: true

max mds

Description: The number of active MDS daemons during cluster creation. Set under [mon] or [global].
Type: 32-bit Integer
Default: 1

mds max file size

Description: The maximum allowed file size to set when creating a new file system.
Type: 64-bit Integer Unsigned
Default: 1ULL << 40

mds cache size

Description: The number of inodes to cache.
Type: 32-bit Integer
Default: 100000

mds cache mid

Description: The insertion point for new items in the cache LRU (from the top).
Type: Float
Default: 0.7

mds dir commit ratio

Description: The fraction of directory that is dirty before Ceph commits using a full update (instead of partial update).
Type: Float
Default: 0.5

mds dir max commit size

Description: The maximum size of a directory update before Ceph breaks it into smaller transactions) (MB).
Type: 32-bit Integer
Default: 90

mds decay halflife

Description: The half-life of MDS cache temperature.
Type: Float
Default: 5

mds beacon interval

Description: The frequency (in seconds) of beacon messages sent to the monitor.
Type: Float
Default: 4

mds beacon grace

Description: The interval without beacons before Ceph declares an MDS laggy (and possibly replace it).
Type: Float
Default: 15

mds blacklist interval

Description: The blacklist duration for failed MDSs in the OSD map.
Type: Float
Default: 24.0*60.0

mds session timeout

Description: The interval (in seconds) of client inactivity before Ceph times out capabilities and leases.
Type: Float
Default: 60

mds session autoclose

Description: The interval (in seconds) before Ceph closes a laggy client's session.
Type: Float
Default: 300

mds reconnect timeout

Description: The interval (in seconds) to wait for clients to reconnect during MDS restart.
Type: Float
Default: 45

mds tick interval

Description: How frequently the MDS performs internal periodic tasks.
Type: Float
Default: 5

mds dirstat min interval

Description: The minimum interval (in seconds) to try to avoid propagating recursive stats up the tree.
Type: Float
Default: 1

mds scatter nudge interval

Description: How quickly dirstat changes propagate up.
Type: Float
Default: 5

mds client prealloc inos

Description: The number of inode numbers to preallocate per client session.
Type: 32-bit Integer
Default: 1000

mds early reply

Description: Determines whether the MDS should allow clients to see request results before they commit to the journal.
Type: Boolean
Default: true

mds use tmap

Description: Use trivialmap for directory updates.
Type: Boolean
Default: true

mds default dir hash

Description: The function to use for hashing files across directory fragments.
Type: 32-bit Integer

Default: 2 (i.e., rjenkins)

mds log

Description: Set to true if the MDS should journal metadata updates (disabled for benchmarking only).

Type: Boolean

Default: true

mds log skip corrupt events

Description: Determines whether the MDS should try to skip corrupt journal events during journal replay.

Type: Boolean

Default: false

mds log max events

Description: The maximum events in the journal before we initiate trimming. Set to -1 to disable limits.

Type: 32-bit Integer

Default: -1

mds log max segments

Description: The maximum number of segments (objects) in the journal before we initiate trimming. Set to -1 to disable limits.

Type: 32-bit Integer

Default: 30

mds log max expiring

Description: The maximum number of segments to expire in parallels

Type: 32-bit Integer

Default: 20

mds log eopen size

Description: The maximum number of inodes in an EOpen event.

Type: 32-bit Integer

Default: 100

mds bal sample interval

Description: Determines how frequently to sample directory temperature (for fragmentation decisions).

Type: Float

Default: 3

mds bal replicate threshold

Description: The maximum temperature before Ceph attempts to replicate metadata to other nodes.

Type: Float

Default: 8000

mds bal unreplicate threshold

Description: The minimum temperature before Ceph stops replicating metadata to other nodes.

Type: Float

Default: 0

mds bal frag

Description: Determines whether the MDS will fragment directories.

Type: Boolean

Default: false

mds bal split size

Description: The maximum directory size before the MDS will split a directory fragment into smaller bits.
Type: 32-bit Integer
Default: 10000

mds bal split rd

Description: The maximum directory read temperature before Ceph splits a directory fragment.
Type: Float
Default: 25000

mds bal split wr

Description: The maximum directory write temperature before Ceph splits a directory fragment.
Type: Float
Default: 10000

mds bal split bits

Description: The number of bits by which to split a directory fragment.
Type: 32-bit Integer
Default: 3

mds bal merge size

Description: The minimum directory size before Ceph tries to merge adjacent directory fragments.
Type: 32-bit Integer
Default: 50

mds bal merge rd

Description: The minimum read temperature before Ceph merges adjacent directory fragments.
Type: Float
Default: 1000

mds bal merge wr

Description: The minimum write temperature before Ceph merges adjacent directory fragments.
Type: Float
Default: 1000

mds bal interval

Description: The frequency (in seconds) of workload exchanges between MDSs.
Type: 32-bit Integer
Default: 10

mds bal fragment interval

Description: The frequency (in seconds) of adjusting directory fragmentation.
Type: 32-bit Integer
Default: 5

mds bal idle threshold

Description: The minimum temperature before Ceph migrates a subtree back to its parent.
Type: Float
Default: 0

mds bal max

Description: The number of iterations to run balancer before Ceph stops. (used for testing purposes only)
Type: 32-bit Integer
Default: -1

mds bal max until

Description: The number of seconds to run balancer before Ceph stops. (used for testing purposes only)
Type: 32-bit Integer
Default: -1

mds bal mode

Description: The method for calculating MDS load.

- 1 = Hybrid.
- 2 = Request rate and latency.
- 3 = CPU load.

Type: 32-bit Integer
Default: 0

mds bal min rebalance

Description: The minimum subtree temperature before Ceph migrates.
Type: Float
Default: 0.1

mds bal min start

Description: The minimum subtree temperature before Ceph searches a subtree.
Type: Float
Default: 0.2

mds bal need min

Description: The minimum fraction of target subtree size to accept.
Type: Float
Default: 0.8

mds bal need max

Description: The maximum fraction of target subtree size to accept.
Type: Float
Default: 1.2

mds bal midchunk

Description: Ceph will migrate any subtree that is larger than this fraction of the target subtree size.
Type: Float
Default: 0.3

mds bal minchunk

Description: Ceph will ignore any subtree that is smaller than this fraction of the target subtree size.
Type: Float
Default: 0.001

mds bal target removal min

Description: The minimum number of balancer iterations before Ceph removes an old MDS target from the MDS map.
Type: 32-bit Integer
Default: 5

mds bal target removal max

Description: The maximum number of balancer iteration before Ceph removes an old MDS target from the MDS map.
Type: 32-bit Integer
Default: 10

mds replay interval

Description: The journal poll interval when in standby-replay mode. (“hot standby”)
Type: Float
Default: 1

mds shutdown check

Description: The interval for polling the cache during MDS shutdown.
Type: 32-bit Integer
Default: 0

mds thrash exports

Description: Ceph will randomly export subtrees between nodes (testing only).
Type: 32-bit Integer
Default: 0

mds thrash fragments

Description: Ceph will randomly fragment or merge directories.
Type: 32-bit Integer
Default: 0

mds dump cache on map

Description: Ceph will dump the MDS cache contents to a file on each MDSMap.
Type: Boolean
Default: false

mds dump cache after rejoin

Description: Ceph will dump MDS cache contents to a file after rejoining the cache (during recovery).
Type: Boolean
Default: false

mds verify scatter

Description: Ceph will assert that various scatter/gather invariants are true (developers only).
Type: Boolean
Default: false

mds debug scatterstat

Description: Ceph will assert that various recursive stat invariants are true (for developers only).
Type: Boolean
Default: false

mds debug frag

Description: Ceph will verify directory fragmentation invariants when convenient (developers only).
Type: Boolean
Default: false

mds debug auth pins

Description: The debug auth pin invariants (for developers only).
Type: Boolean
Default: false

mds debug subtrees

Description: The debug subtree invariants (for developers only).
Type: Boolean
Default: false

mds kill mdstable at

Description: Ceph will inject MDS failure in MDSTable code (for developers only).
Type: 32-bit Integer
Default: 0

mds kill export at

Description: Ceph will inject MDS failure in the subtree export code (for developers only).
Type: 32-bit Integer
Default: 0

mds kill import at

Description: Ceph will inject MDS failure in the subtree import code (for developers only).
Type: 32-bit Integer
Default: 0

mds kill link at

Description: Ceph will inject MDS failure in hard link code (for developers only).
Type: 32-bit Integer
Default: 0

mds kill rename at

Description: Ceph will inject MDS failure in the rename code (for developers only).
Type: 32-bit Integer
Default: 0

mds wipe sessions

Description: Ceph will delete all client sessions on startup (for testing only).
Type: Boolean
Default: 0

mds wipe ino prealloc

Description: Ceph will delete ino preallocation metadata on startup (for testing only).
Type: Boolean
Default: 0

mds skip ino

Description: The number of inode numbers to skip on startup (for testing only).
Type: 32-bit Integer
Default: 0

mds standby for name

Description: An MDS daemon will standby for another MDS daemon of the name specified in this setting.
Type: String
Default: N/A

mds standby for rank

Description: An MDS daemon will standby for an MDS daemon of this rank.
Type: 32-bit Integer
Default: -1

mds standby replay

Description: Determines whether a ceph-mds daemon should poll and replay the log of an active MDS (hot standby).
Type: Boolean
Default: false

