

JOURNAL CONFIG REFERENCE

Ceph OSDs use a journal for two reasons: speed and consistency.

- **Speed:** The journal enables the Ceph OSD Daemon to commit small writes quickly. Ceph writes small, random i/o to the journal sequentially, which tends to speed up bursty workloads by allowing the backing filesystem more time to coalesce writes. The Ceph OSD Daemon's journal, however, can lead to spiky performance with short spurts of high-speed writes followed by periods without any write progress as the filesystem catches up to the journal.
- **Consistency:** Ceph OSD Daemons require a filesystem interface that guarantees atomic compound operations. Ceph OSD Daemons write a description of the operation to the journal and apply the operation to the filesystem. This enables atomic updates to an object (for example, placement group metadata). Every few seconds-between `filestore max sync interval` and `filestore min sync interval`-the Ceph OSD Daemon stops writes and synchronizes the journal with the filesystem, allowing Ceph OSD Daemons to trim operations from the journal and reuse the space. On failure, Ceph OSD Daemons replay the journal starting after the last synchronization operation.

Ceph OSD Daemons support the following journal settings:

`journal dio`

Description: Enables direct i/o to the journal. Requires `journal block align` set to `true`.
Type: Boolean
Required: Yes when using `aio`.
Default: `true`

`journal aio`

Changed in version 0.61: Cuttlefish

Description: Enables using `libaio` for asynchronous writes to the journal. Requires `journal dio` set to `true`.
Type: Boolean
Required: No.
Default: Version 0.61 and later, `true`. Version 0.60 and earlier, `false`.

`journal block align`

Description: Block aligns write operations. Required for `dio` and `aio`.
Type: Boolean
Required: Yes when using `dio` and `aio`.
Default: `true`

`journal max write bytes`

Description: The maximum number of bytes the journal will write at any one time.
Type: Integer
Required: No
Default: `10 << 20`

`journal max write entries`

Description: The maximum number of entries the journal will write at any one time.
Type: Integer
Required: No
Default: `100`

`journal queue max ops`

Description: The maximum number of operations allowed in the queue at any one time.
Type: Integer
Required: No
Default: `500`

`journal queue max bytes`

Description: The maximum number of bytes allowed in the queue at any one time.

Type: Integer
Required: No
Default: 10 << 20

journal align min size

Description: Align data payloads greater than the specified minimum.
Type: Integer
Required: No
Default: 64 << 10

journal zero on create

Description: Causes the file store to overwrite the entire journal with 0's during mkfs.
Type: Boolean
Required: No
Default: false