POOL, PG AND CRUSH CONFIG REFERENCE

When you create pools and set the number of placement groups for the pool, Ceph uses default values when you don't specifically override the defaults. **We recommend** overridding some of the defaults. Specifically, we recommend setting a pool's replica size and overriding the default number of placement groups. You can specifically set these values when running pool commands. You can also override the defaults by adding new ones in the `[global]` section of your Ceph configuration file.

```
[global]

        # By default, Ceph makes 3 replicas of objects. If you want to make four
        # copies of an object the default value--a primary copy and three replica
        # copies--reset the default values as shown in 'osd pool default size'.
        # If you want to allow Ceph to write a lesser number of copies in a degraded
        # state, set 'osd pool default min size' to a number less than the
        # 'osd pool default size' value.

        osd pool default size = 3  # Write an object 3 times.
        osd pool default min size = 2 # Allow writing two copies in a degraded state.

        # Ensure you have a realistic number of placement groups. We recommend
        # approximately 100 per OSD. E.g., total number of OSDs multiplied by 100
        # divided by the number of replicas (i.e., osd pool default size). So for
        # 10 OSDs and osd pool default size = 4, we'd recommend approximately
        # (100 * 10) / 4 = 250.

        osd pool default pg num = 250
        osd pool default pgp num = 250
```

`mon max pool pg num`

- **Description:** The maximum number of placement groups per pool.
- **Type:** Integer
- **Default:** 65536

`mon pg create interval`

- **Description:** Number of seconds between PG creation in the same Ceph OSD Daemon.
- **Type:** Float
- **Default:** 30.0

`mon pg stuck threshold`

- **Description:** Number of seconds after which PGs can be considered as being stuck.
- **Type:** 32-bit Integer
- **Default:** 300

`mon pg min inactive`

- **Description:** Issue a `HEALTH_ERR` in cluster log if the number of PGs stay inactive longer than `mon_pg_stuck_threshold` exceeds this setting. A non-positive number means disabled, never go into ERR.
- **Type:** Integer
- **Default:** 1

`mon pg warn min per osd`

- **Description:** Issue a `HEALTH_WARN` in cluster log if the average number of PGs per (in) OSD is under this number. (a non-positive number disables this)
- **Type:** Integer
- **Default:** 30

`mon pg warn max per osd`

- **Description:** Issue a `HEALTH_WARN` in cluster log if the average number of PGs per (in) OSD is above this number. (a non-positive number disables this)

Type: Integer
Default: 300

`mon pg warn min objects`

Description: Do not warn if the total number of objects in cluster is below this number
Type: Integer
Default: 1000

`mon pg warn min pool objects`

Description: Do not warn on pools whose object number is below this number
Type: Integer
Default: 1000

`mon pg check down all threshold`

Description: Threshold of down OSDs percentage after which we check all PGs for stale ones.
Type: Float
Default: 0.5

`mon pg warn max object skew`

Description: Issue a `HEALTH_WARN` in cluster log if the average object number of a certain pool is greater than `mon pg warn max object skew` times the average object number of the whole pool. (a non-positive number disables this)
Type: Float
Default: 10

`mon delta reset interval`

Description: Seconds of inactivity before we reset the pg delta to 0. We keep track of the delta of the used space of each pool, so, for example, it would be easier for us to understand the progress of recovery or the performance of cache tier. But if there's no activity reported for a certain pool, we just reset the history of deltas of that pool.
Type: Integer
Default: 10

`mon osd max op age`

Description: Maximum op age before we get concerned (make it a power of 2). A `HEALTH_WARN` will be issued if a request has been blocked longer than this limit.
Type: Float
Default: 32.0

`osd pg bits`

Description: Placement group bits per Ceph OSD Daemon.
Type: 32-bit Integer
Default: 6

`osd pgp bits`

Description: The number of bits per Ceph OSD Daemon for PGPs.
Type: 32-bit Integer
Default: 6

`osd crush chooseleaf type`

Description: The bucket type to use for `chooseleaf` in a CRUSH rule. Uses ordinal rank rather than name.
Type: 32-bit Integer
Default: 1. Typically a host containing one or more Ceph OSD Daemons.

`osd crush initial weight`

Description: The initial crush weight for newly added osds into crushmap.
Type: Double
Default: `the size of newly added osd in` TB. By default, the initial crush weight for the newly added osd is set to its volume size in TB. See Weighting Bucket Items for details.

`osd pool default crush rule`

Description: The default CRUSH rule to use when creating a replicated pool.
Type: 8-bit Integer
Default: `-1`, which means "pick the rule with the lowest numerical ID and use that". This is to make pool creation work in the absence of rule 0.

`osd pool erasure code stripe unit`

Description: Sets the default size, in bytes, of a chunk of an object stripe for erasure coded pools. Every object of size S will be stored as N stripes, with each data chunk receiving `stripe unit` bytes. Each stripe of `N * stripe unit` bytes will be encoded/decoded individually. This option can is overridden by the `stripe_unit` setting in an erasure code profile.
Type: Unsigned 32-bit Integer
Default: `4096`

`osd pool default size`

Description: Sets the number of replicas for objects in the pool. The default value is the same as `ceph osd pool set {pool-name} size {size}`.
Type: 32-bit Integer
Default: `3`

`osd pool default min size`

Description: Sets the minimum number of written replicas for objects in the pool in order to acknowledge a write operation to the client. If minimum is not met, Ceph will not acknowledge the write to the client, **which may result in data loss**. This setting ensures a minimum number of replicas when operating in `degraded` mode.
Type: 32-bit Integer
Default: `0`, which means no particular minimum. If `0`, minimum is `size - (size / 2)`.

`osd pool default pg num`

Description: The default number of placement groups for a pool. The default value is the same as `pg_num` with `mkpool`.
Type: 32-bit Integer
Default: `8`

`osd pool default pgp num`

Description: The default number of placement groups for placement for a pool. The default value is the same as `pgp_num` with `mkpool`. PG and PGP should be equal (for now).
Type: 32-bit Integer
Default: `8`

`osd pool default flags`

Description: The default flags for new pools.
Type: 32-bit Integer
Default: `0`

`osd max pgls`

Description: The maximum number of placement groups to list. A client requesting a large number can tie up the Ceph OSD Daemon.
Type: Unsigned 64-bit Integer
Default: `1024`
Note: Default should be fine.

`osd min pg log entries`

Description: The minimum number of placement group logs to maintain when trimming log files.
Type: 32-bit Int Unsigned
Default: `1000`

`osd default data pool replay window`

Description: The time (in seconds) for an OSD to wait for a client to replay a request.
Type: 32-bit Integer
Default: 45

`osd max pg per osd hard ratio`

Description: The ratio of number of PGs per OSD allowed by the cluster before OSD refuses to create new PGs. OSD stops creating new PGs if the number of PGs it serves exceeds `osd max pg per osd hard ratio` * `mon max pg per osd`.
Type: Float
Default: 2