

PG state	Description
<code>inactive</code>	The PG has been inactive for too long.
<code>unclean</code>	The PG has been unclean for too long.
<code>remapped</code>	The acting set has changed, and the PG is temporarily remapped to a different set of OSDs while the primary OSD recovers or backfills.
<code>down</code>	The PG is offline.
<code>splitting</code>	The PG is being split; the number of PGs is being increased.
<code>scrubbing, deep-scrubbing</code>	A PG scrub or deep-scrub operation is in progress.

When an OSD is added to a placement group, the PG enters the `peering` state to ensure that all nodes agree about the state of the PG. If the PG can handle read and write requests after peering completes, then it reports an `active` state. If the PG also has the correct number of replicas for all of its objects, then it reports a `clean` state. The normal PG operating state after writes are complete is `active+clean`.

When an object is written to the PG's primary OSD, the PG reports a `degraded` state until all replica OSDs acknowledge that they have also written the object.

The `backfill` state means that data is being copied or migrated to rebalance PGs across OSDs. If a new OSD is added to the PG, it is gradually backfilled with objects to avoid excessive network traffic. Backfilling occurs in the background to minimize the performance impact on the cluster. The `backfill_wait` state indicates that a backfill operation is pending. The `backfill` state indicates that a backfill operation is in progress. The `backfill_too_full` state indicates that a backfill operation was requested, but could not be completed due to insufficient storage capacity.

A PG marked as `inconsistent` might have replicas that are different from the others, detected as a different data checksum or metadata size on one or more replicas. A clock skew in the Ceph cluster and corrupted object content can also cause an `inconsistent` PG state.

Identifying Stuck Placement Groups

The placement groups transition into `degraded` or `peering` states after a failure. If a placement group remains in one of these states for a long period, then the MON marks the placement group as stuck. A stuck PG might be in one or more of the following states:

- An **inactive** PG might be having a peering problem.
- An **unclean** PG might be having problems recovering after a failure.
- A **stale** PG has no OSDs reporting, which might indicate that all OSDs are down and out.
- An **undersized** PG does not have enough OSDs to store the configured number of replicas.



Note

The MONs use the `mon_pg_stuck_threshold` parameter to decide if a PG has been in an error state for too long. The default value for the threshold is 300 seconds.

Ceph marks a PG as `stale` when all OSDs that have copies of a specific PG are in `down` and `out` states. To return from a `stale` state, an OSD must be revived to have a PG copy available and for