PG

CONCEPTS

Peering Interval

See PG::start_peering_interval. See PG::up_acting_affected. See PG::RecoveryState::Reset

A peering interval is a maximal set of contiguous map epochs in which the up and acting sets did not change. PG::RecoveryMachine represents a transition from one interval to another as passing through RecoveryState::Reset. On PG;:RecoveryState::AdvMap PG::up_acting_affected can cause the pg to transition to Reset.

PEERING DETAILS AND GOTCHAS

For an overview of peering, see Peering.

- PG::flushed defaults to false and is set to false in PG::start_peering_interval. Upon transitioning to PG::RecoveryState::Started we send a transaction through the pg op sequencer which, upon complete, sends a FlushedEvt which sets flushed to true. The primary cannot go active until this happens (See PG::RecoveryState::WaitFlushedPeering). Replicas can go active but cannot serve ops (writes or reads). This is necessary because we cannot read our ondisk state until unstable transactions from the previous interval have cleared.