CERTIFICATION-BASED REPLICATION

Certification-based replication uses group communication and transaction ordering techniques to achieve synchronous replication.

Transactions execute optimistically in a single node, or replica, and then at commit time, they run a coordinated certification process to enforce global consistency. It achieves global coordination with the help of a broadcast service that establishes a global total order among concurrent transactions.
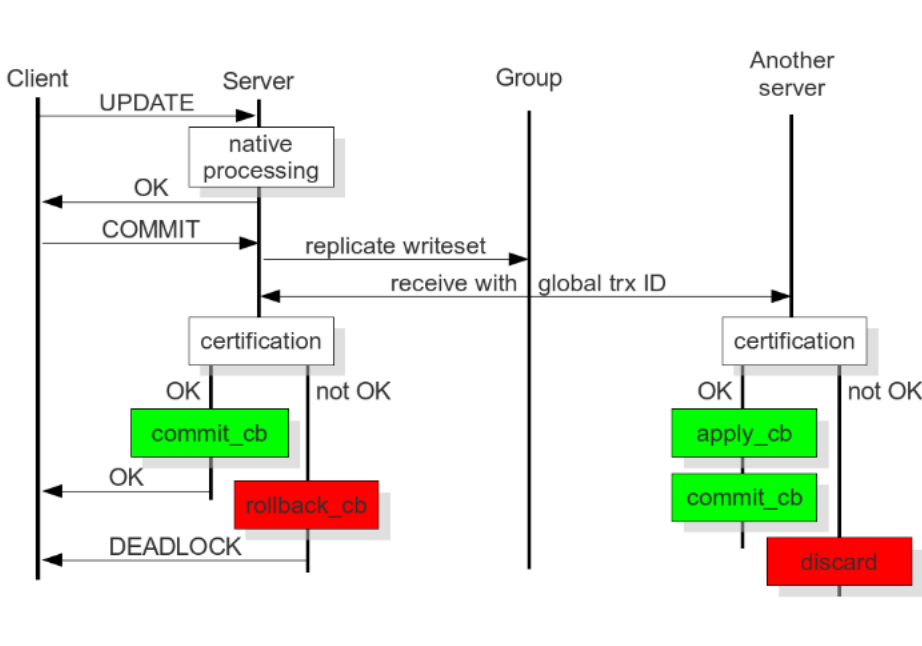
WHAT CERTIFICATION-BASED REPLICATION REQUIRES

It is not possible to implement certification-based replication for all database systems. It requires certain features of the database in order to work.

- **Transactional Database** It requires that the database is transactional. Specifically, that the database can rollback uncommitted changes.
- **Atomic Changes** It requires that replication events change the database atomically. Specifically, that the series of database operations must either all occur, else nothing occurs.
- **Global Ordering** It requires that replication events are ordered globally. Specifically, that they are applied on all instances in the same order.

HOW CERTIFICATION-BASED REPLICATION WORKS

The main idea in certification-based replication is that a transaction executes conventionally until it reaches the commit point, assuming there is no conflict. This is called optimistic execution.

Certification Based Replication

When the client issues a `COMMIT` command, but before the actual commit occurs, all changes made to the database by the transaction and primary keys of the changed rows are collected into a write-set. The database then sends this write-set to all the other nodes.

The write-set then undergoes a deterministic certification test, using the primary keys. This is done on each node in the cluster, including the node that originates the write-set. It determines whether or not the node can apply the write-set.

If the certification test fails, the node drops the write-set and the cluster rolls back the original transaction. If the test succeeds, the transaction commits and the write-set is applied to the rest of the cluster.

CERTIFICATION-BASED REPLICATION IN GALERA CLUSTER

The implementation of certification-based replication in Galera Cluster depends on the global ordering of transactions.

Galera Cluster assigns each transaction a global ordinal sequence number, or seqno, during replication. When a transaction reaches the commit point, the node checks the sequence number against that of the last successful transaction. The interval between the two is the area of concern, given that transactions that occur within this interval have not seen the effects of each other. All transactions in this interval are checked for primary key conflicts with the transaction in question. The certification test fails if it detects a conflict.

The procedure is deterministic and all replica receive transactions in the same order. Thus, all nodes reach the same decision about the outcome of the transaction. The node that started the transaction can then notify the client application whether or not it has committed the transaction.

POWERED BY GALERA:

MariaDB

PRODUCTS

Galera Cluster

Use cases

Success stories

Case studies

Technology

White papers

Benchmarks

Videos & presentations

DOWNLOADS

Software

Documentation

Partner downloads

SUPPORT

Support Partners

Support Subscription

Consulting

COMMUNITY

Mailing List

Facebook feed

COMPANY

Overview

What we believe

News

Careers

Galera Cluster Story

BLOG

CONTACT US

Images: NASA

YOUR DATA IS INVALUABLE. WE'RE HERE TO SECURE IT. ALL OF IT.