MariaDB Xpand is a distributed SQL database compatible with MariaDB MaxScale and MariaDB/MySQL connectors. It is capable of executing millions of transactions per second with distributed transactions, high availability, strong consistency and standard SQL.

**LOGICAL DATABASE**

Xpand databases consist of multiple, physical database instances (or nodes) working together as a single logical database (or cluster). This is different from traditional relational databases, whether primary/replica or multi-master, where every database instance functions as an identical, complete database storing all data and executing queries by itself. In an Xpand database, each database instance is responsible for storing and querying a subset of data, and multiple database instances may be needed to execute a query.

**DISTRIBUTED DATA AND INDEXES**

Xpand divides tables into partitions and evenly distributes them across all database nodes. By default, the number of partitions is equal to the number of nodes. However, with large multi-core processors, the number of partitions can be increased with multiple partitions per node. If nodes are added or removed, Xpand automatically splits/merges partitions to rebalance the data.

Most Distributed SQL databases have to execute the same query on every single database node within a cluster because each one has its own local index not visible to the others. However, Xpand distributed indexes too. It is able to find and look up the relevant part of an index first, then send a query to only those database nodes with matching rows (rather than to all database nodes).
HIGH AVAILABILITY AND AUTOMATIC REBALANCING

Xpand can be configured to store multiple copies of partitions, or replicas. By default, it stores two replicas of a partition, with each one on a different node. However, it can store anywhere from one copy of a partition per cluster to one copy per node. If a node fails, Xpand automatically begins reading and writing to a replica on a different one. When deployed to the cloud, it can ensure replicas are stored in different zones. If an entire zone fails, Xpand automatically begins reading and writing to replicas in the other zones. The same is true if deployed on premises and across multiple racks. In either scenario, Xpand automatically replaces any lost replicas in order to restore full fault tolerance.

Note:
Xpand can store multiple replicas of a partition, but only one replica is cached in order to maximize the total amount of data that can be read from memory – effectively creating a distributed cache with no duplicate entries.

ACID TRANSACTIONS AND STRONG CONSISTENCY

Xpand uses three-phase commit (3PC) and Paxos to perform distributed transactions, as well snapshot isolation via multi-version concurrency control (MVCC) for lock-free reads and two-phase locking (2PL) for concurrent writes. Applications can begin a transaction on any node. It will become the global transaction manager (GTM), and coordinate the transaction with all participating nodes. In addition, transactions write to all replicas (in parallel) in order to ensure strong consistency.

UPDATE table1
SET favorite=true
WHERE color=red or color=purple

mariadb.com
Americas: sales-AMER@mariadb.com
Europe, Middle East, Africa: sales-EMEA@mariadb.com
Asia Pacific: sales-APAC@mariadb.com

© Copyright 2021 MariaDB Corporation Ab, Tekniikantie 12, 02150 Espoo, Finland. MariaDB is a trademark or registered trademark of MariaDB Corporation.