19. Read Repair

Let's learn about read repair and its usage.

We'll cover the following

- Background
- Definition
- Solution
- Examples

Background#

In Distributed Systems, where data is replicated across multiple nodes, some nodes can end up having stale data. Imagine a scenario where a node failed to receive a write or update request because it was down or there was a network partition. How do we ensure that the node gets the latest version of the data when it is healthy again?

Definition#

Repair stale data during the read operation, since at that point, we can read data from multiple nodes to perform a comparison and find nodes that have stale data. This mechanism is called Read Repair. Once the node with old data is known, the read repair operation pushes the newer version of data to nodes with the older version.

Solution#

Solution#

Based on the quorum, the system reads data from multiple nodes. For example, for Quorum=2, the system reads data from one node and digest of the data from the second node. The digest is a checksum of the data and is used to save network bandwidth. If the digest does not match, it means some replicas do not have the latest version of the data. In this case, the system reads the data from all the replicas to find the latest data. The system returns the latest data to the client and initiates a **Read Repair** request. The read repair operation pushes the latest version of data to nodes with the older version.

When the read consistency level is less than 'All,' some systems perform a read repair probabilistically, for example, 10% of the requests. In this case, the system immediately sends a response to the client when the consistency level is met and performs the read repair asynchronously in the background.

Examples#

Cassandra and **Dynamo** use 'Read Repair' to push the latest version of the data to nodes with the older versions.

← Back

18. Hinted Handoff

Next →

20. Merkle Trees

Mark as Completed

Report an Issue