

Easily determine which partition contains a given key

Not evenly spaced

Range scans are easy

Hot spots

Good at distributing keys fairly

Lose the ability to do efficient range queries

In the extreme case where all reads and writes are for the same key

Add random number, allow keys to be distributed to different partitions

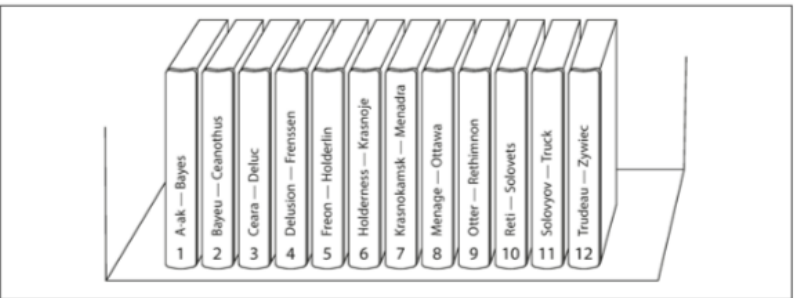


Figure 6-2. A print encyclopedia is partitioned by key range.

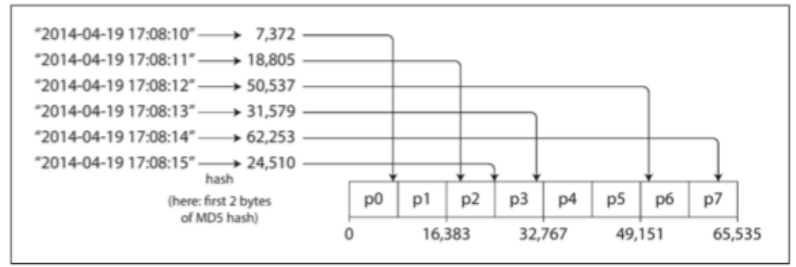


Figure 6-3. Partitioning by hash of key.

Partitioning by hash of key

Skewed workloads and relieving hot spots

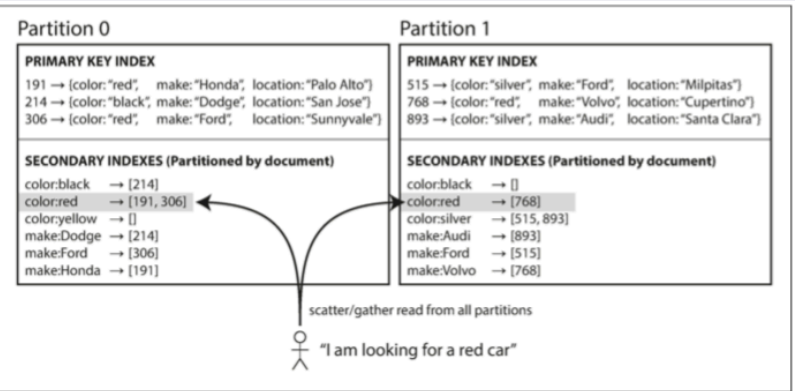


Figure 6-4. Partitioning secondary indexes by document.

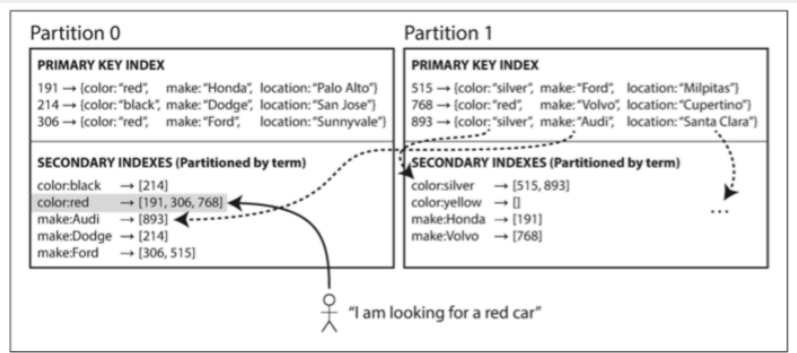


Figure 6-5. Partitioning secondary indexes by term.

By document

By term

Partitioning of key-value data

Partitioning and secondary indexes

## Partitioning

Why: scalability

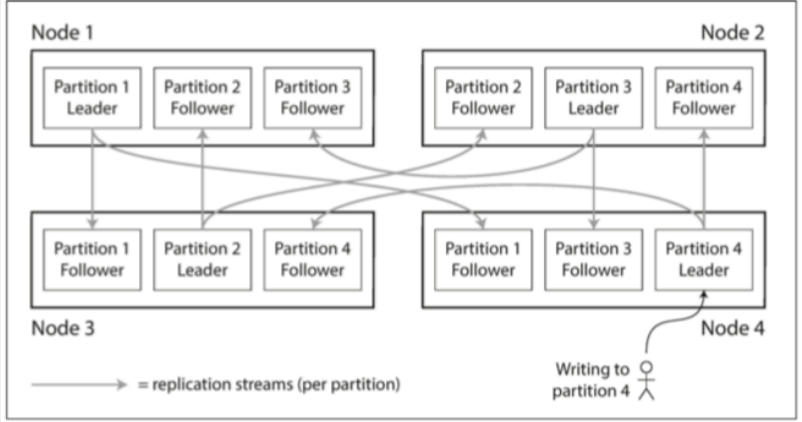


Figure 6-1. Combining replication and partitioning: each node acts as leader for some partitions and follower for other partitions.

Partitioning and Replication

Requirements

After rebalancing, the load should be shared fairly

While rebalancing is happening, the database should be continue accepting reads and writes

No more data than necessary should be moved between nodes

Hash mod N

Problem is the if the number of nodes N changes, most of the keys need to be moved.

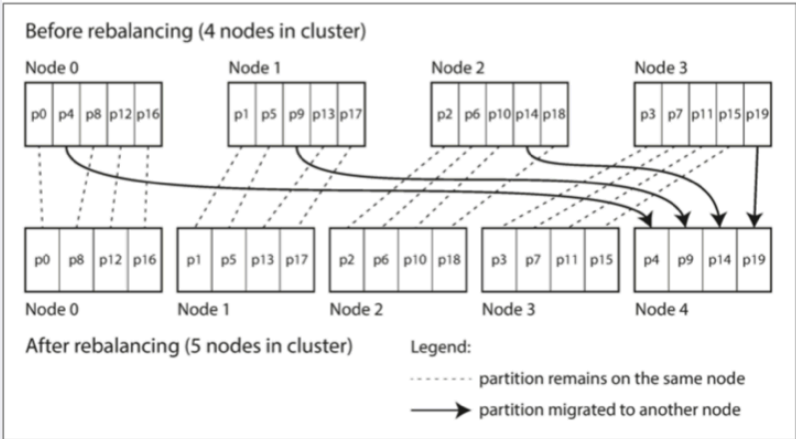


Figure 6-6. Adding a new node to a database cluster with multiple partitions per node.

Fixed number of partitions

If the boundaries is wrong, it could end up with all of the data in one partition.

Strategies for rebalancing

Dynamic partitioning

When a partition grows to exceed a configured size, it is split into two partitions. Conversely it can merge.

Partitioning proportionally to nodes

Examples : Cassandra Ketama

Have a fixed number of partitions per node.

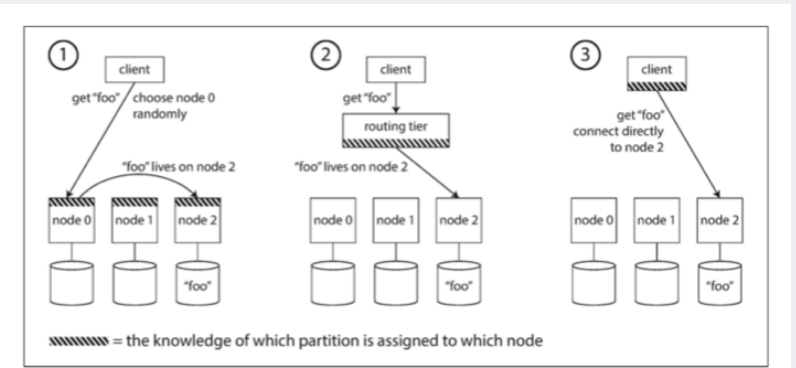


Figure 6-7. Three different ways of routing a request to the right node.

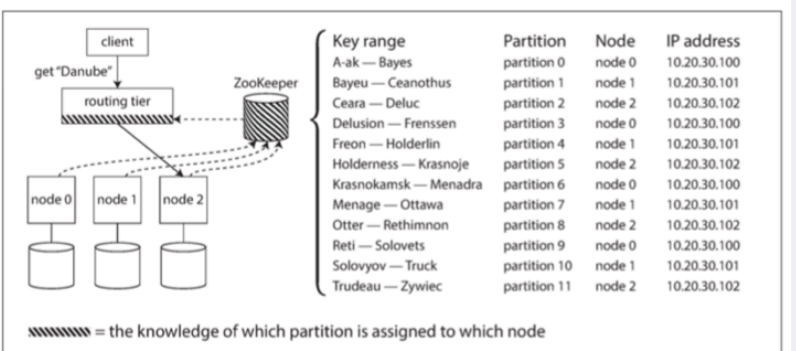


Figure 6-8. Using ZooKeeper to keep track of assignment of partitions to nodes.

Request routing

Use a separate coordination service such as ZooKeeper