

Assignment 3

- **What is a system key space and what all schema does that hold? Is the system key space configurable?**

System key space is similar to data dictionary – in the System tablespace -- in Oracle. It keeps meta information about the database. The System key space holds meta information about the cluster, which helps in running the cluster.

It holds two column family: Schema column family and Migration column family.

It contains information such as cluster name, node's token, whether or not node is bootstrapped among others.

The System key space cannot be configured.

- **We know Cassandra node are identical what others node configurations do you know off? What are their advantages and disadvantages?**

The Cassandra nodes are identical as they communicate using peer-to-peer and gossip protocol. With peer-to-peer, any node can take read/or write request.

The other node configuration is master-slave, which is used in HBase.

Advantage of P2P Protocol:

- Being highly decentralized any node can be used for read or write, which that helps in delivering better performance and scalability.
- It's easy to scale horizontally as a node can be added by using any other node as seed node.
- Since all nodes acts as peer, all of them can share contents and resources
- As it is highly decentralized, there is no single point of failure.

Disadvantage of P2P Protocol

- Since there is no master node, administration such as backup and security and other routine administration task are not easy to do.

Advantage of Master-Slave

- Since there is a master node, administration such as backup and security and other routine administration task are not easy to do.

Disadvantage of Master-Slave

- Performance of write is usually not as good as read. As all writes go through the master node.
- There is a single point of failure; if the master node goes down it can affect the operation till the master node becomes operational.

- **Describe various Cassandra Architectural layers? Please explain all the components of the middle layer?**

Broadly, Cassandra architecture can be logically – in bottom-up approach—be divided in three layers: bottom, middle, and top.

The bottom layer deals with replication, partitioning, gossiping, and messaging.

The middle layer deals with Commit Log, MemTable, SSTable, Indexes and Compaction.

The top-layer deals with bootstrapping, read repair, hinted handoff, tombstone, and monitoring tool.

Explanation of middle layer:

During a write, first the record is written to a commit log. The commit log helps in recovering the written should node fails before the record is actually written to the SSTable.

MemTable receives the record from the Commit Log, and the data from the MemTable is flush to SSTable as it crosses the spillover level. Compaction process takes the small SSTable and merges them into a single SSTable when compaction threshold is reached – default is 4.

- **What other function does Gossip protocol achieve apart from internode communication? Please explain the process?**

Gossip protocol contacts to the other node to find information about its location and state. If the node fails then state

information which the other node has received via gossip protocol helps to write missing the information.

- **What happens when you write to a failed node in Cassandra?
Please explain the process in detail?**

When information is written to a failed node, then the node uses hinted hand off to let the other node write the information as failed nodes comes up. This hinted handoff is helped by gossip process. In other words, using gossip protocol, the other node knows what all the information is missing in the failed node at it relates to replication factor.

- **How do you increase the performance of Cassandra reads?
Please explain the process in detail?**

The Random Partitioning, which is default, and Weak Consistency are the key to improving read performance.

With Random Processing, rows are assigned by generating MD5 hash of the row key (128 bit) than it divided by the number of nodes. The remainder decides which node the key would go. This process doesn't create hot spots and load of the read process is evenly divided.

Weak Consistency, which is Any or One as opposed to Quorum and All, also helps in getting better read performance. In case of Any, the row is sent to the client if it present on any node -- the timestamp is not compared to determine if the node has latest value. In case of One, the read is sent if at least one node has the latest value and read repair is performed with the background

thread to update the latest value on the other nodes based on the replication factor.

- **What are snitches? Please explain different kinds of snitches?**

Snitches help in determining relative host proximity which helps in replication – Simple and Network Topology.

There are two types: Simple and Property File

In Simple Snitch, if the second octet of the two nodes are the same it means they are in the same data center. Furthermore, if third octet is same as well then it means two nodes are on the same rack.

In the Property File Snitch, a property file is looked up to determine the node location.