Sessional 2

RJ

CASSANDRA

2 MARKS

1. What are the key features of Apache Cassandra?

ANS: Elastic Scalability, Tunable Consistency, Peer to Peer, Open Source, Decentralized, Column Oriented.

1. What is Column Store DB? Explain with an example.

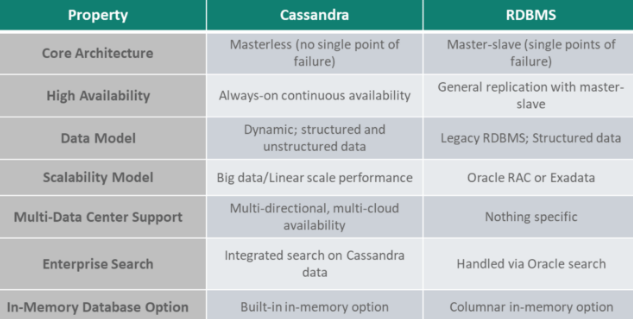
ANS: Data is stored in cells are grouped in columns of data rather than as rows of data. Columns are logically grouped into column families.

One row may have one or multiple data records, which is indexed by a partition key.

Apache Cassandra, Apache HBase

1. What are the Key Differences between Cassandra and Traditional RDBMS?

ANS:



1. What do you mean by replication factor and consistency level?

ANS: Cassandra stores copies (called replicas) of each row based on the row key. The replication factor refers to the number of nodes that will act as copies (replicas) of each row of data.

The Cassandra consistency level is defined as the minimum number of Cassandra nodes that must acknowledge a read or write operation before the operation can be considered successful.

1. What is a Keyspace ? How a Keyspace is created in Cassandra?

ANS: A keyspace is a data container in Cassandra, similar to a database in relational database management systems (RDMBS). A cluster has one keyspace per application, as many as needed, depending on requirements and system usage

CREATE KEYSPACE ABC WITH replication = { ‘class ’: ‘SimpleStrategy’, ‘replication\_factor’: ‘3’}

AND durable\_writes = ‘TRUE’;

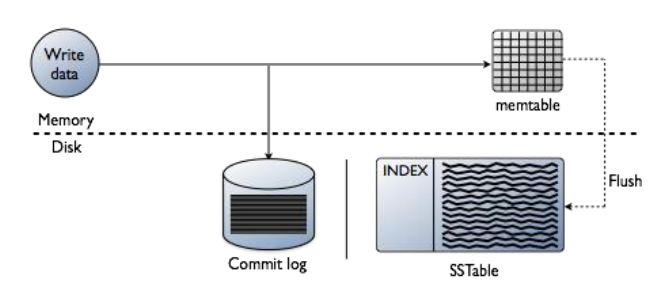
3 MARKS

1. What is Hinted Handoffs in Cassandra? Explain with Diagram.

ANS: Pg no. 175

1. How does Cassandra perform write operations? Explain with diagram.

ANS:

A write operation once completed is first written to commit log.

• Next it is pushed to a memory resident data structure called Memtable, which has a predefined threshold value.

• Only when write is completed in memtable and commitlog, a node responds successful message to coordinator.

• When no. of objects reaches a threshold, then the contents of Memtable are flushed to the disk in a file called SSTable(Sorted String Table)

• Flushing is a non-blocking operation

1. What do you mean by Replication in Cassandra? write down and explain different replication strategies.

ANS: Replication factor determines to how many

nodes a keyspace (like databases in relation

databases) will be replicated.

• RF should be >1 and <no. of nodes in cluster

• Two Strategies:

1) SimpleStrategy:

– Just move around the ring to meet the replication factor

– Stop once replication factor is met.

–2)NetworkTopologyStrategy

--Complicated

– Allows to specify different replication factor to different data centers

– Inside a datacenter, it stores replicas on different racks.

4 MARKS

1. Write a short note on Cassandra Read Path and there components.

ANS:

(Key topics supposed to be discussed)

Bloom Filter

Key Cache

Partition Summary

Partition Index

SSTable

Row Cache

1. Discuss Primary key in tables also discuss how you will use one partition key, two partition keys, sort on one column and sort on two columns.
2. How you can set and check consistency levels in Cassandra? Discuss various Read and Write Consistency levels.

6 MARKS

1. What is Cassandra CRUD Operation? Explain with suitable example.