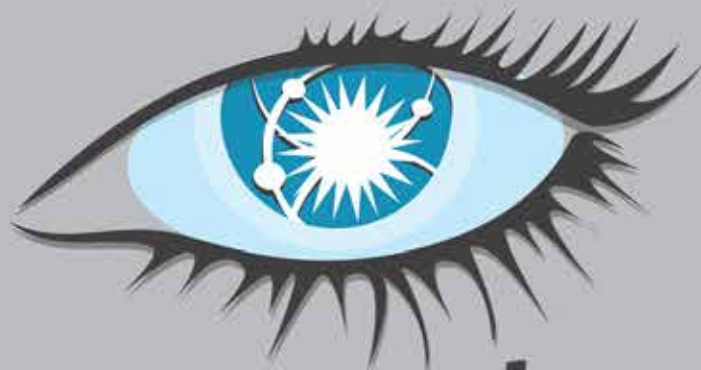


An Overview of Apache



cassandra



IMPORTANT

Copyright Infringement and Illegal Content Sharing Notice

All course content designs, video, audio, text, graphics, logos, images are Copyright© and are protected by India and international copyright laws. All rights reserved.

Permission to download the contents (wherever applicable) for the sole purpose of individual reading and preparing yourself to crack the interview only. Any other use of study materials – including reproduction, modification, distribution, republishing, transmission, display – without the prior written permission of Author is strictly prohibited.

Trendytech Insights legal team, along with thousands of our students, actively searches the Internet for copyright infringements. Violators subject to prosecution.

Cassandra Overview

1. What is Cassandra

Cassandra is a distributed column oriented database and is highly performant.

It is highly scalable also.

When we require transactional activities or quick retrieval. Low latency retrieval.

2. Cap Theorem

C - Consistency

A - Availability

P - Partition tolerance

Hbase - CP (Consistency & Partition tolerance)

Cassandra - AP (Availability & Partition tolerance)

Consider you have 100 likes on your post.

100 likes (Satish liked it)

101 likes

3. How Does a Cassandra Cluster Look Like

In case of your Hbase we run on hadoop cluster.

Also Hbase provides a master-slave architecture where master is the hmaster and slave is the region server however in case of Cassandra there is no master.

All the nodes are peers.

Cassandra has a decentralized architecture.

For communication among themselves they use gossip protocol.

In master slave architectures there can be down times if master fails.

Because of a decentralized (no master architecture) Cassandra is highly available

4. Tunable Read/Write Consistency

Cassandra is a **AP** system.

It by default compromises on the consistency to be highly available.

Step 1: Client will send the request to get value of A

Step 2: The request will go to one of the machine.
for instance node 5.

Step 3: Node 5 will go and talk to node 1 to get the results.

Step 4: Node 5 will return the result to client.

Cassandra provides you a tunable consistency.

I want the result only when 2 nodes agrees on the same result.

I want the result only when all nodes agree on the same result.

One node (default)

Availability is high and consistency is low.

All nodes (All the machines should agree on the same result)

Availability is low and consistency is high

Quorum (2 nodes)

5. Differences Between Hbase and Cassandra

Similarities:

- Both of these are nosql distributed databases and hold the data in a columnar fashion.
- Both of these are highly scalable.
- When you want to perform transaction (updates,inserts) and quick reads.
- Low latency operations.

Differences:

- Hbase has a master-slave architecture. However Cassandra has a decentralized architecture. (highly available as it has no dependency on single master)
- Hbase provides (CP) consistency and partition tolerance and Cassandra (AP) availability and partition tolerance. However cassandra also offers tunable consistency.

Where we can set the consistency level:

- One node
- All nodes
- Quorum

- Hbase runs on top of hadoop cluster. that means data is kept in hdfs.
- However cassandra has a separate cluster than your hadoop cluster.
- if you are working on hadoop then hbase is a perfect choice mostly.

4. Hbase Shell Commands.

Cassandra has its own query language called as **CQL** (Cassandra Query Language).

This CQL is very similar to normal SQL.

TRENDYTECH 9108179578

Hbase syntax looks quite hard.

Apache Phoenix:

Gives you a sql interface on top of nosql

Apache Phoenix works on top of Hbase where you can query your Hbase using SQL like syntax.



5 Star Google Rated
Big Data Course

LEARN FROM THE EXPERT



9108179578

Call for more details



Follow US

Trainer Mr. Sumit Mittal

Phone 9108179578

Email trendytech.sumit@gmail.com

Website <https://trendytech.in/courses/big-data-online-training/>

LinkedIn <https://www.linkedin.com/in/bigdatabysumit/>

Twitter @BigdataBySumit

Instagram bigdatabysumit

Facebook <https://www.facebook.com/trendytech.in/>

Youtube TrendyTech