

- Data model.
- ④ SEMANDRA
 - ⑤ mongoDB
 - ⑥ SPARQL (PySPARQL)
 - ⑦ airflow
 - ⑧ Sqoop

Cloud

AWS
GCP
Azure

Database → Data model. → RDBMS. → table (Row × Columns) → Schema. (Rigid schema)

Card name deg. card.		
→	→	→
→	→	→
→	→	→
→	→	→

HBASE → Column oriented.

emp details

→ Row id → Person detail → Official detail

→ Column family → Name → Age → empid → Designation

→ (column oriented DB) → Sunny → 25 → 1 → DS

→ → Vishal → 23 → 2 → ML

→ → Iman → 26 → 3 → BDE

Row id	Person detail		Official detail	
	Name	Age	empid	Designation
	Sunny	25	1	DS
	Vishal	23	2	ML
	Iman	26	3	BDE

HBASE V/S HDFS

- RDBMS v/s HBASE

Q. Why we are using HBASE still we are having HDFS.

- ① Schema.
- ② Performance
- ③ Speed.
- ④ Scalability. (master slave)

HBASE Architecture.

- ① Zookeeper
- ② HMaster
- ③ Region Server
- ④ Region
- ⑤ Column family.
- ⑥ HDFS

emp details		Rowid	= Person detail		Official detail	
Column family	Column		Name	Age	Empid	Designation
	-		Sunny	25	1	DS
	-		Vishal	23	2	ML
	-		Iraan	26	3	BDE

(column oriented DB)

- ① mem storage
- ② Block cache
- ③ Hfile

- ① mem storage
- ② Block cache
- ③ Hfile

- ① Read operation (Reading our data)
- ② Write operation (store q data)

Configure:

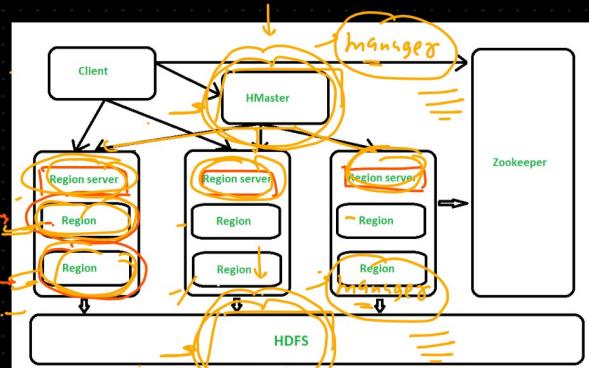
128 MB

Datanode

default

⇒ 256 MB

Node → system.



= 512 MB

HBASE

2-region

256 MB

If (region) can we store this entire data in single region?

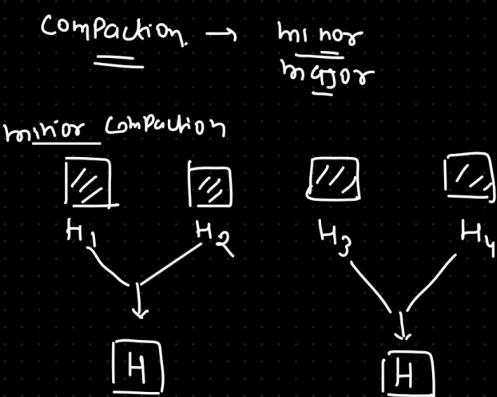
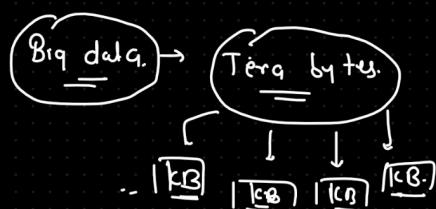
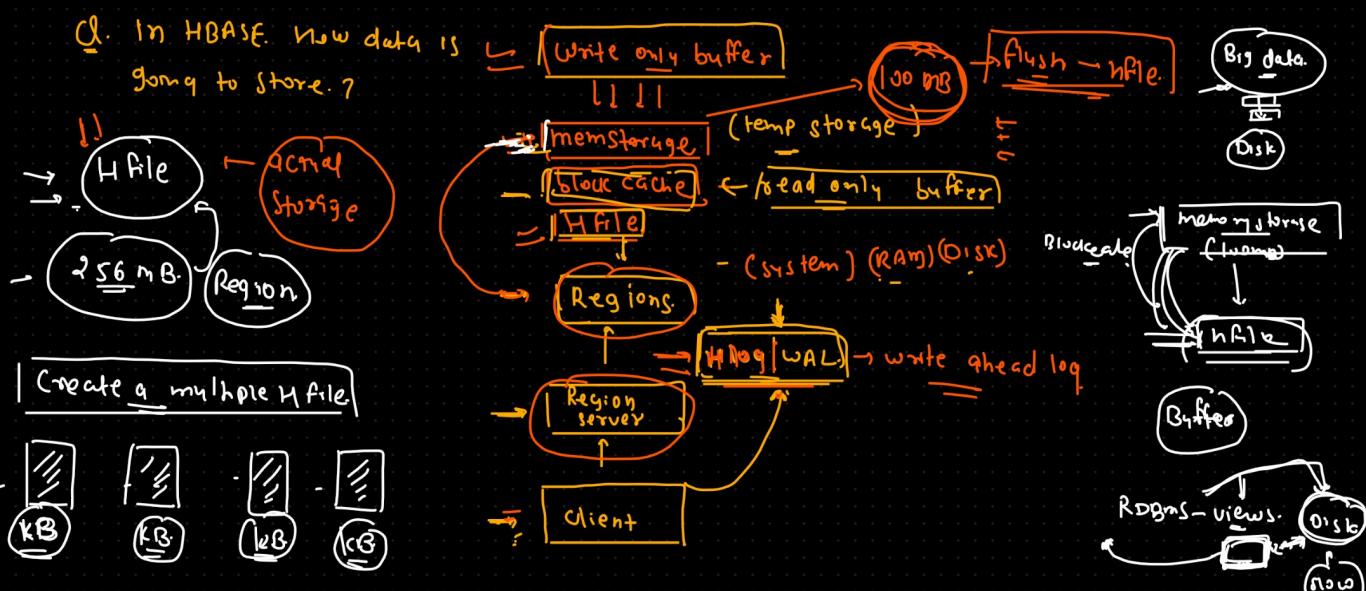
① Create / update, ② Delete, ③ Maintain region server, ④ Load balancing, ⑤ Region server

failure

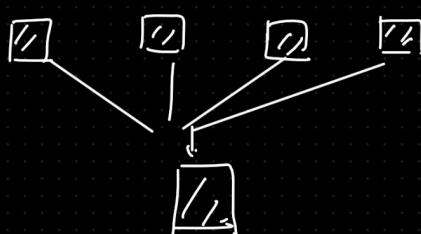


① Read. ② Write.

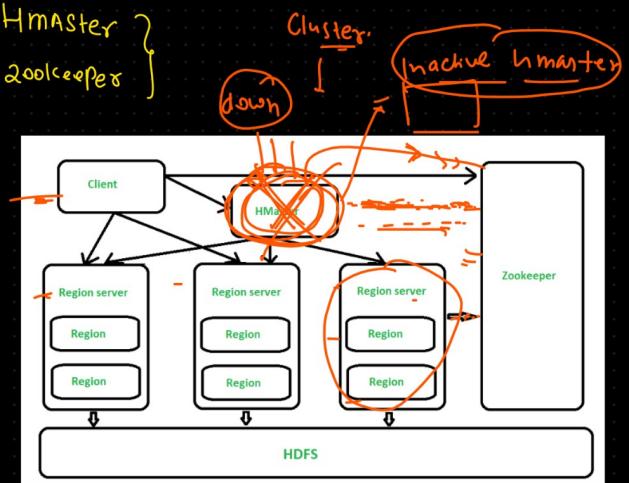
Q: In HBASE. How data is going to store?



major compaction (minor compaction)



1 HMaster
2 Zookeeper



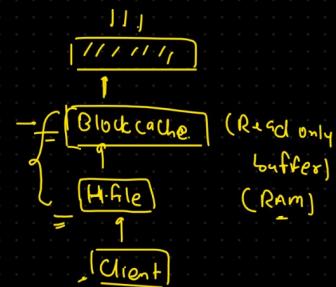
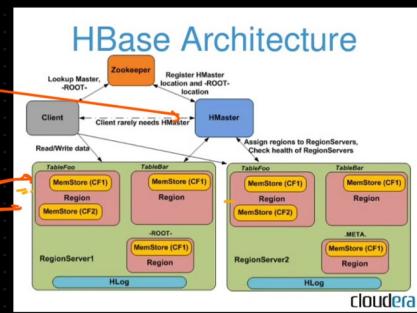
Write → **Hfile**

Read →

Single

Root

=
Meta table



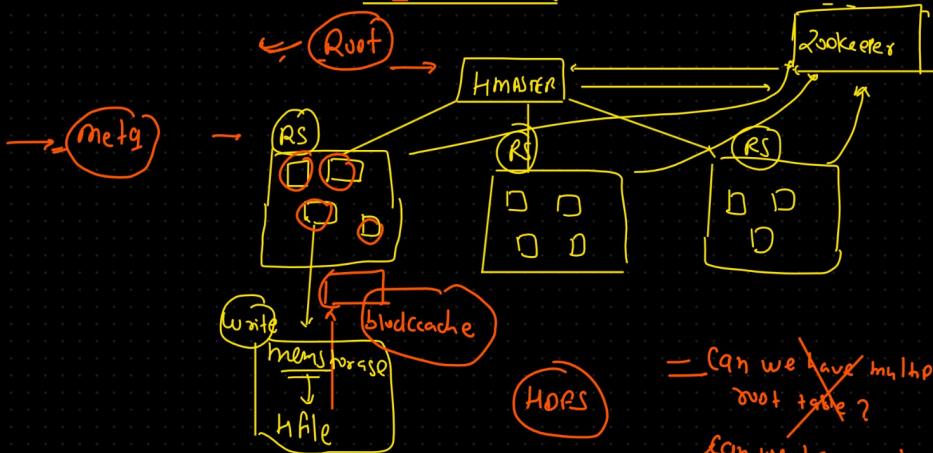
① ZooKeeper (Managing) (Heartbeat)

④ HMASTER (Manage regionserver)

③ region server

region (data node)

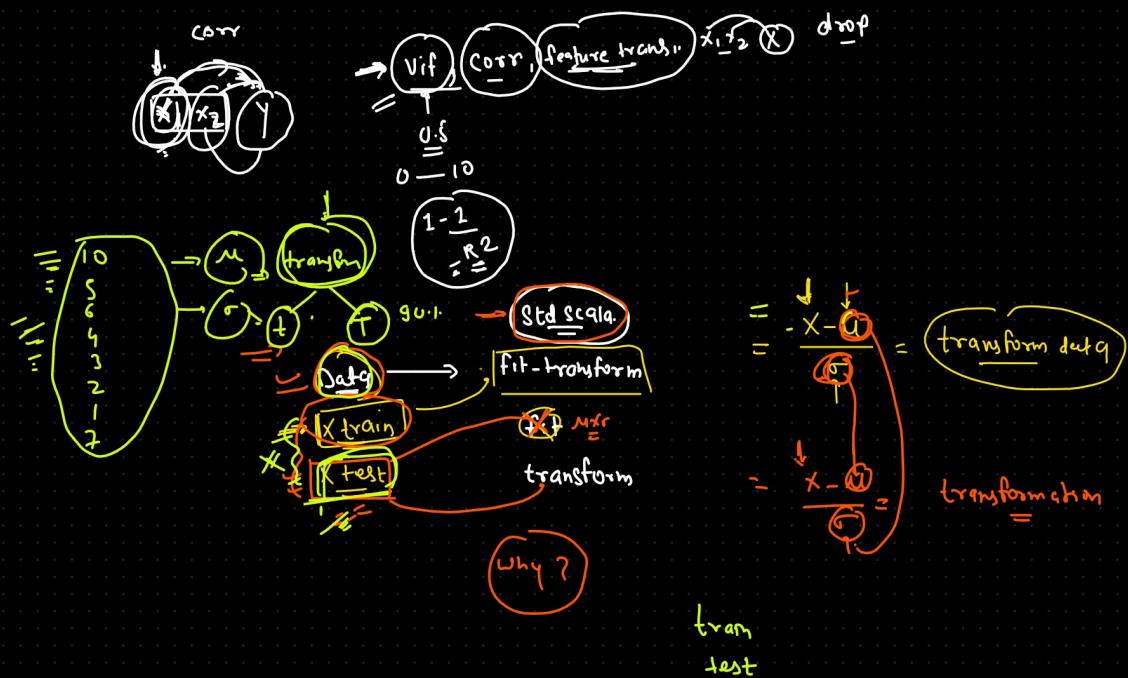
multiple region servers



Can we have multiple root table?

Can we have multiple meta table?





X_{train}

X_{test}

$$\begin{aligned}
 & \left[\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \right] \xrightarrow{\text{fitting}} \left(\begin{array}{c} \mu = 5 \\ \sigma = 2 \end{array} \right) \\
 & \xrightarrow{\text{transform}} \frac{x - \mu}{\sigma} = \frac{1 - 5}{2} = -2 \\
 & \quad = \frac{2 - 5}{2} = -1.5 \\
 & \quad = \frac{3 - 5}{2} = -1 \\
 & \quad = \frac{4 - 5}{2} = -0.5 \\
 & \quad = \frac{5 - 5}{2} = 0 \\
 & \quad = \frac{6 - 5}{2} = 0.5 \\
 & \quad = \frac{7 - 5}{2} = 1 \\
 & \quad = \frac{8 - 5}{2} = 1.5 \\
 & \quad = \frac{9 - 5}{2} = 2 \\
 & \quad = \frac{10 - 5}{2} = 2.5
 \end{aligned}$$