# HBase

HBase Architecture

Region server:

WAL: Write ahead log on disk – used for recovery

Read Request: Go to BlockCache 🡪 Get the column name information 🡪 get MenStore or Hfile location of the requested information.

Hbase vs RDBMS:

Column oriented storage

* HBase rows are sorted based on row id key
* Values are stored at cell level corresponding to each row key
* No fixed schema, can add or delete columns dynamically

Denormalized – Struct and Array can be used

Only CRUD (Create Read Update Delete) operations

* Operations can be performed using API’s or JAVA program API’s
* No joins possibles

ACID at row level

* In RDBMS when updating 10 rows for 2 columns in a single transactions, if transaction fails at 6th transaction, all 5 updateds are rolled back. In Hbase 5 rows are updated, others are not updated.

Hbase vs HIVE:

|  |  |
| --- | --- |
| Hbase | Hive |
| Data Management | Only Analytical processing |
| For transactions and Analytical processing |  |
| low latency | Not low latency |
| No SQL interface | SQL interface |
| HDFS only | HDFS MapReduce for processing engine |
| Batch and Online | Batch |