HBase

WE'LL COVER THE FOLLOWING

A Hadoop database is the Apache HBase, in which operations keep running in real time on its database instead of Map Reduce jobs, just like the Pig or Hive. Apache HBase offers random read or write access to big data. Also, it can accommodate expansive tables, for instance, billions of rows 'by' millions of columns. It is similar to other tools in the ecosystem of Hadoop, due to the fact that it is fault-tolerant and also designed to operate on commodity hardware.

One significant contrast between HBase and different other RDBMSs is that HBase relates to a NoSQL Database. Rather than SQL as the essential access language, HBase offers an API in order to read and write data.

Features

- Linear and modular scalability.
- Strictly consistent reads and writes.
- Automatic and configurable sharding of tables
- Automatic failover support between RegionServers.
- Convenient base classes for backing Hadoop MapReduce jobs with Apache HBase tables.
- Easy to use Java API for client access.
- Block cache and Bloom Filters for real-time queries.
- Query predicate push down via server side Filters

- Thrift gateway and a REST-ful Web service that supports XML, Protobuf, and binary data encoding options
- Extensible jruby-based (JIRB) shell
- Support for exporting metrics via the Hadoop metrics subsystem to files or Ganglia; or via JMX.

https://hbase.apache.org/