HBase is a data model that is similar to Google’s big table designed to provide quick random access to huge amounts of structured data. The procedures to set up HBase on Hadoop File Systems, and ways to interact with HBase shell. It also describes how to connect to HBase using java, and how to perform basic operations on HBase using java.

Limitations of Hadoop

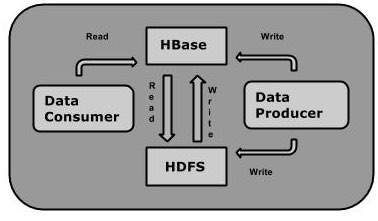
Hadoop can perform only batch processing, and data will be accessed only in a sequential manner. That means one has to search the entire dataset even for the simplest of jobs.

A huge dataset when processed results in another huge data set, which should also be processed sequentially. At this point, a new solution is needed to access any point of data in a single unit of time (random access). HBase, Cassandra, couchDB, Dynamo, and MongoDB providing

## What is HBase?

HBase is a distributed column-oriented database built on top of the Hadoop file system. It is an open-source project and is horizontally scalable.

It leverages the fault tolerance provided by the Hadoop File System (HDFS)

``````````` `

## Storage Mechanism in HBase (column-oriented database)

* Table is a collection of rows.
* Row is a collection of column families.
* Column family is a collection of columns.
* Column is a collection of key value pairs.
* Each cell value of the table has a timestamp.