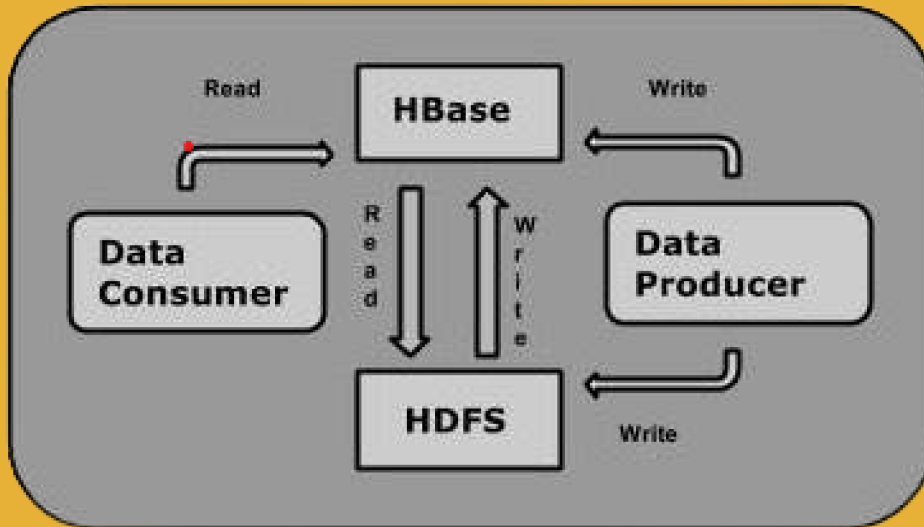


Apache HBase



Distributed Column-Oriented Database



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

Storage Mechanism in HBase

HBase is a column-oriented database, and its tables are sorted by row. The table schema defines only column families, which are the key-value pairs. A table has multiple column families; each can have a number of columns. Subsequent column values are stored contiguously on the disk. Each cell value of the table has a timestamp.

It is a part of the Hadoop ecosystem that provides random real-time read/write access to data in the Hadoop File System.

One can store the data in HDFS either directly or through HBase. Data consumer reads/accesses the data in HDFS randomly using HBase. HBase sits on top of the Hadoop File System and provides read and write access.

