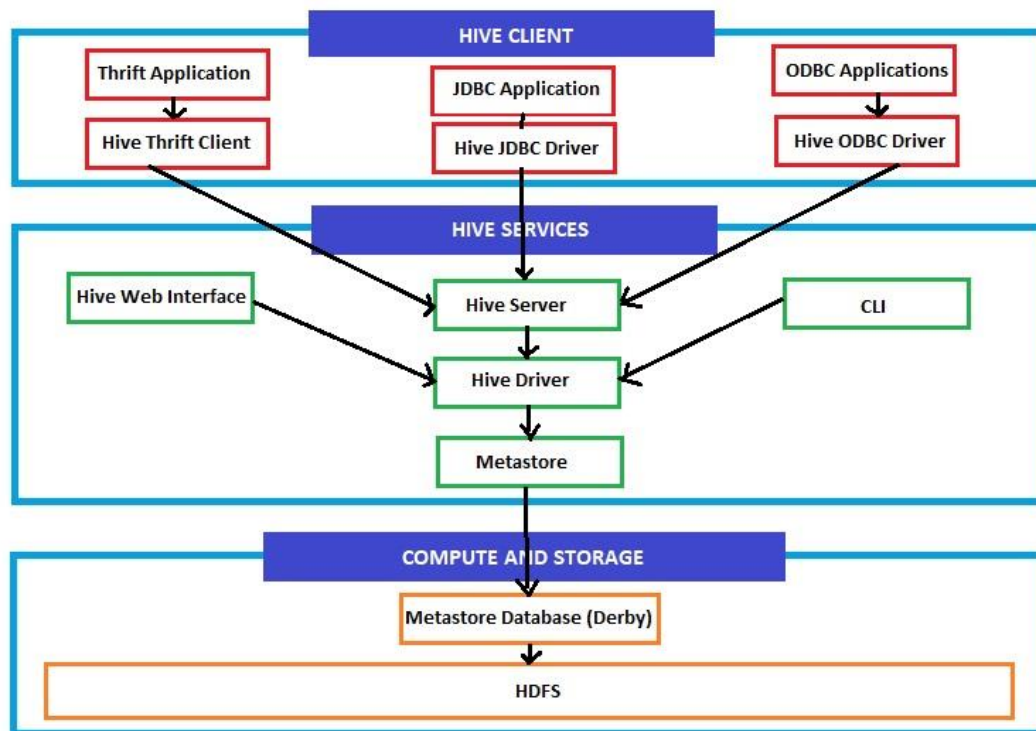


Assignment 2.9

Hive Architecture



The Hive Architecture comprises 3 main components:

1. **Hive Client**
This is where the applications get an interface to interact with hive.
2. **Hive Services**
Hive services enable the hive interactions by passing them through the hive driver which in turn uses MapReduce.
3. **Compute and Storage**
This is the workhorse of the Hive ecosystem which includes the Metastore DB and HDFS storage.

Hive Components

Hive Client

Hive client is the interface for different applications and clients which makes it possible for different applications to communicate with Hive.

Hive client comprises Thrift Client, JDBC Client and ODBC Client.

Thrift Client enables Thrift based applications to communicate with Hive. JDBC and ODBC enable several applications and languages to make connections and process with hive.

These clients in turn connect to the Hive Server to make this possible.

Hive Services

Client interactions are made possible by Hive services.

CLI service enables us to interact with Hive via command line.

Web based service makes it possible to have a web based interaction with Hive.

The Hive server service is one of the most important and most used service which most of the Client services connect to.

Hive Storage and Compute

Hive storage used MapReduce in the background which makes it possible for hive to process Big Data.

Hive stores meta information in the Hive Metastore which is the backbone of Hive.

By default, Hive uses Derby DB for storing metadata, but any other RDBMS can be used for this purpose.

However, the data resides on either Local storage or HDFS.