

Review: BDAT - TUTORIAL - HIVE (2)

Respondent

15

Dr. V. Sivakumar

06:55

Time to complete

0/24

Points

📝 **Needs review**

Score / 3 pts

Needs review

1. What is Hive? *

Apache HIVE is a data warehouse infrastructure software/system which is built on top of Hadoop MapReduce. It is used for Querying and analyzing the big data. (summarization)
It is targeted towards the users who are comfortable with SQL.

📝 **Needs review**

Score / 3 pts

Needs review

2. What is the difference between HBase and Hive?

*

Hbase is a NoSQL database whereas HIVE is a datawarehouse framework which is analytical tool.
Hbase does not allow execution of SQL queries, whereas HIVE allows the execution of SQL queries.

📝 **Needs review**

Score / 3 pts

Needs review

3. What is a Hive Metastore? *

Hive Metastore is a central repository which stores the metadata (schema of the table, ...).

📝 **Needs review**

Score / 3 pts

Needs review

4. What are the two types of tables in Hive? *

1. Managed Table (Internal Table)
2. External Table

📝 **Needs review**

Score / 3 pts

Needs review

5. List the Limitations of Hive. *

Apache Hive does not provide row level updates.
Hive is not good for OLTP.
Latency for Hive queries is very high.

📝 **Needs review**

Score / 3 pts

Needs review

6. List the advantages of Hive *

It is an efficient ETL (Extract, Transform and Load) tool.
Useful for people those who are not from programming background as it eliminates the need to write complex MapReduce program.
It is scalable to cope up with the growing volume and variety of data.

📝 **Needs review**

Score / 3 pts

Needs review

7. What are the types of data models in Hive? *

Data in Hive can be categorized into data models as:

Table

Partition

Bucket

📝 **Needs review**

Score / 3 pts

Needs review

8. State the significant differences amongst internal and external tables in HIVE. *

Refer to our Notes (Difference with regard to Structure, With regard to dropping of table, With regard to global usage)