Session14 Assignment 2- answers

**1. Is it possible to use same meta store by multiple users in case of embedded Hive, if no then why?**

By default, metastore database in hive is Debry. which has a limitation to only one Derby database to have access to files on disk anytime, meaning only one session is possible.

**2. What is SerDe in Hive?**

SerDe is short for Serializer/Deserializer. Hive uses the SerDe interface for IO. The interface handles both serialization and deserialization and also interpreting the results of serialization as individual fields for processing.

A SerDe allows Hive to read in data from a table, and write it back out to HDFS in any custom format. Anyone can write their own SerDe for their own data formats.

**3. What is the functionality of query processor in Apache Hive?**

This component implements the processing framework for converting SQL to a graph of map/reduce jobs and the execution time framework to run those jobs in the order of dependencies.

**4. How can Hive avoid MapReduce?**

Hive can avoid MapReduce, by making use of other execution engine. For example Apache Tez. In order to set execution engine, use “hive.execution.engine” property.

**5. What are the types of table in Hive?**

There are 2 types of tables : 1. Internal table and 2. External table

**6. Does Hive support record level insert, delete or update?**

Nope

**7. What are the binary storage formats supported in Hive?**   
• SEQUENCEFILE  
• RCFILE  
• ORCFILE

**8. What is the difference between external table and internal table in Hive?**

Drop table – when drop an external table, the underlying data files stay intact. Whereas internal table drop causes both meta data(underlying directories) and data get dropped.

Data availability – External Tables are available outside hive, whereas internal tables are not available outside hive.