Give brief answers to the questions below:

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

No, it is not possible to use metastore in sharing mode. It is recommended to use standalone "real" database like MySQL or PostGresSQL.

1. What is SerDe in Hive?

A SerDe is a short name for a Serializer Deserializer. Hive uses SerDe (and FileFormat) to read and write data from tables. An important concept behind Hive is that it DOES NOT own the Hadoop File System (HDFS) format that data is stored in. Users are able to write files to HDFS with whatever tools/mechanism takes their fancy("CREATE EXTERNAL TABLE" or "LOAD DATA INPATH," ) and use Hive to correctly "parse" that file format in a way that can be used by Hive. A SerDe is a powerful (and customizable) mechanism that Hive uses to "parse" data stored in HDFS to be used by Hive.

1. 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.

1. How can Hive avoid MapReduce?

Hive can use the [Apache Tez](http://hortonworks.com/hadoop/tez) execution engine instead of the venerable Map-reduce engine.

set hive.execution.engine=tez;

1. What are the types of table in Hive?

There are two types of tables in Hive ,one is Managed table and second is external table.  
the difference is , when you drop a table, if it is managed table hive deletes both data and meta data,if it is external table Hive only deletes metadata.

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

Apache Hive provides a convenient SQL query engine and table abstraction for data stored in Hadoop. Hive uses Hadoop to provide highly scaleable bandwidth to the data, but until recently did not support updates, deletes, or transaction isolation. This has prevented many desirable use cases, such as updating of dimension tables or doing data clean up. We have implemented the standard SQL commands insert, update, and delete allowing users to insert new records as they become available, update changing dimension tables, repair incorrect data, and remove individual records. This also allows very low latency ingestion of streaming data from tools like Storm and Flume.

1. What are the binary storage formats supported in Hive?

In sequence files the data is stored in a binary storage format consisting of binary key value pairs. A complete row is stored as single binary value. Sequence files are more compact than text and fit well the map-reduce output format. Sequence files do support block compression and can be compressed on value, or block level, to improve its IO profile further.

SEQUENCEFILE is a standard format that is supported by Hadoop itself and is good choice for Hive table storage especially when you want to integrate Hive with other techonolgies in the Hadoop ecosystem.

CREATE TABLE SEQUENCEFILE\_TABLE (

COLUMN1 STRING,

COLUMN2 STRING,

COLUMN3 INT,

COLUMN4 INT

) STORED AS SEQUENCEFILE

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

When we drop an internal table, it drops the data, and it also drops the metadata.

When we drop an external table, it only drops the meta data. That means hive is ignorant of that data now. It does not touch the data itself.