1. The \_\_\_\_\_\_\_\_\_\_is the heart of an HDFS file system. It keeps the metadata such as

directory tree of all files in the file system and tracks the cluster where the file data is

present. The actual data is stored on \_\_\_\_\_\_\_as HDFS blocks.

Ans :c. namenode,datanode

a) datanode, datanode

b) resourcemanager, namenode

c) namenode, datanode

d) tasktraker, jobtraker

2. When the end of the block is reached, \_\_\_\_\_\_\_closes the connection to the datanode,

then finds the best datanode for the next block.

Ans :c .DFSInputStream

a) DFSOutputStream

b) client

c) DFSInputStream

d) Resourcemanager

3. When the client finishes reading, it calls \_\_\_\_method on the close stream.

Ans :d. close()

a) read()

b) write()

c) shuffle

d) close()

**A C A D G I L D**

**Page 4**

4. Which configuration file contains Environmental variable settings used by Hadoop?

Ans :d.Hadoop-env.sh

a) Core-site.xml

b) mapred-site.xml

c) yarn-site.xml

d) Hadoop-env.sh

5. Which MapReduce daemon instantiates user code, and executes map and reduce tasks

on a cluster running MapReduce vl (MRvl)?

Ans :d. TaskTracker

a) NameNode

b) DataNode

c) JobTracker

d) TaskTracker

6. Identify the function performed by the Secondary NameNode daemon on a cluster

configured to run with a single NameNode.

Ans :a. ) In this configuration, the Secondary NameNode performs a checkpoint operation on the files by the NameNode.

a) In this configuration, the Secondary NameNode performs a checkpoint operation on the files by

the NameNode.

b) In this configuration, the Secondary NameNode is standby NameNode, ready to failover and

provide high availability.

c) In this configuration, the Secondary NameNode performs deal-time backups of the NameNode.

d) In this configuration, the Secondary NameNode servers as alternate data channel for clients to

reach HDFS, should the NameNode become too busy.

7. Hadoop administrators write a script called Topology script to determine the rack

location of nodes. It triggers to know the distance of the nodes to replicate the data and

Configures this script in \_\_\_\_\_\_\_\_.

Ans :c. coe-site.xml

a) yarn-site.xml

b) Hadoop-env.sh

c) core-site.xml

d) mapred-site.env

**A C A D G I L D**

**Page 5**

8. \_\_\_\_\_\_\_\_\_\_is the master that arbitrates all the available cluster resources and thus

helps manage the distributed applications running on the YARN system

Ans :c .ResourceManager(RM)

a) Node manager

b) data manager

c) ResourceManager (RM)

d) Taskmanager

9. \_\_\_\_\_\_\_\_\_take instructions from the ResourceManager and manage resources available

on a single node.

Ans :a. NodeManagers

a) NodeManagers

b) data manager

c) ResourceManager (RM)

d) Taskmanager

10. How does HDFS Federation help HDFS Scale horizontally?

Ans :d. HDFS Federation reduces the load on any single NameNode by using the multiple,

independent NameNode to manage individual pars of the filesystem namespace.

a) HDFS Federation improves the resiliency of HDFS in the face of network issues by

removing the NameNode as a single-point-of-failure.

b) HDFS Federation allows the Standby NameNode to automatically resume the services of

an active NameNode

c) HDFS Federation provides cross-data center (non-local) support for HDFS, allowing a

cluster administrator to split the Block Storage outside the local cluster.

d) HDFS Federation reduces the load on any single NameNode by using the multiple,

independent NameNode to manage individual pars of the filesystem namespace.