

Hadoop Test : 2017

Total points 11/20 ?

Email address *

vivek.as@media.net

AD Id, Eg. Daniella.f

vivek.as

Room

- ☒ Gandalf
- ☐ Sauron
- ☐ Mordor
- ☐ Bangalore

Test ka Aarambh

✓ Choose the correct statement

1/1

- ☐ A) Datanode manage file system namespace

- ☐ B) Namenode stores metadata

☒ B) Namenode stores metadata ✓

☐ C) NameNode stores actual data

☐ D) None of the above

✓ Point out the correct statement :

1/1

☐ a) HDFS provides low latency access to single rows from billions of records (Random access)

☒ b) HBase sits on top of the Hadoop File System and provides read and write access ✓

☐ c) HBase is a distributed file system suitable for storing large files

☐ d) None of the mentioned

✓ What data, does a Reducer reduce method process?

1/1

☐ A) All data produced by a single mapper.

☐ B) All the data in a single input file.

☐ c) All data for a given value, regardless of which mapper(s) produced it

☒ D) All data for a given key, regardless of which mapper(s) produced it. ✓

✗ Which of the following statements are true? Choose multiple answers.

0/1

☒ Distributed Systems can consistently give any two of the CAP properties. ✓

☒ When a DataNode is down, Availability is still achieved ✓

- ☐ When a DataNode is down. Eventual Consistency is achieved.
- ☐ When a DataNode is down, Partition Tolerance is not achieved.

Correct answer

- ☒ Distributed Systems can consistently give any two of the CAP properties.
- ☒ When a DataNode is down, Availability is still achieved
- ☒ When a DataNode is down. Eventual Consistency is achieved.

✗ Parts of Yarn ? (select more than one)

0/1

- ☐ A) NodeManager
- ☒ B) Scheduler and Resource manager ✓
- ☐ C) Namenode
- ☐ D) Datanode

Correct answer

- ☒ A) NodeManager
- ☒ B) Scheduler and Resource manager

✗ What happens when a DataNode goes down? Mark which false. 0/1

- ☒ All the blocks that are under-replicated get replicated to some other DataNode ✗
- ☐ NameNode stops receiving heartbeat from this DataNode

- ☐ All the jobs scheduled on this DataNode are rescheduled to somewhere else
- ☐ NameNode deletes all the references it contains for this DataNode

Correct answer

- ☒ All the jobs scheduled on this DataNode are rescheduled to somewhere else

✗ A file in HDFS that is smaller than a single block size

0/1

- ☐ A) Cannot be stored in HDFS
- ☒ B) Occupies the full block's size. ✗
- ☐ C) Occupies only the size it needs and not the full block.
- ☐ D) Can span over multiple blocks.

Correct answer

- ☒ C) Occupies only the size it needs and not the full block.

✓ The need for data replication can arise in various scenarios like : 1/1

- ☐ a) Replication Factor is changed
- ☐ b) DataNode goes down
- ☐ c) Data Blocks get corrupted

☒ d) All of the mentioned



✓ Which of the following is a valid flow in Hadoop ?

1/1

☐ a. Input -> Reducer -> Mapper -> Combiner -> -> Output

☐ b. Input -> Mapper -> Reducer -> Combiner -> Output

☒ c. Input -> Mapper -> Combiner -> Reducer -> Output



☐ d. Input -> Reducer -> Combiner -> Mapper -> Output

✗ Which component is responsible for handling containers for map and reduce tasks?

0/1

☒ Application Manager



☐ Application Master

☐ Resource Manager

☐ Scheduler

Correct answer

☒ Application Master

✓ Which one of these tools give better throughput and latency for seek operations?

1/1

☐ HDFS

☒ Redis



- ☐ MySql
- ☐ Postgres

✗ During start up, the _____ loads the file system state from the fsimage and the edits log file. 0/1

- ☐ a) DataNode
- ☐ b) NameNode
- ☐ c) ActionNode
- ☒ d) None of the above ✗

Correct answer

- ☒ b) NameNode

✓ UnderReplication in HDFS means- 1/1

- ☐ A - No replication is happening in the data nodes.
- ☐ B - Replication process is very slow in the data nodes.
- ☐ C - The frequency of replication in data nodes is very low.
- ☒ D - The number of replicated copies is less than as specified by the replication factor. ✓

✗ In HBase, table cells are indexed by 0/1

- ☐ row key + column key + timestamp
- ☐ row key + timestamp
- ☐ row key

☒ None of the above



Correct answer

☒ row key

✗ Which of the following is true about MapReduce?

0/1

☐ A) Data processing layer of hadoop

☐ B) It provides the resource management

☒ C) It is an open source data warehouse system for querying and analyzing large datasets stored in hadoop files ✗

☐ D) All of the above

Correct answer

☒ A) Data processing layer of hadoop

✓ Which among the following is the duties of the DataNodes

1/1

☐ A) Manage file system namespace

☐ B) Stores meta-data

☐ C) Regulates client's access to files

☒ D) Perform read-write operation as per request for the clients ✓

✓ Point out the wrong statement

1/1

☐ Replication Factor can be configured at a cluster level (Default is set to 3) and also at a file level

☐ Block Report from each DataNode contains a list of all the blocks that are stored on that DataNode

☐ User data is stored on the local file system of DataNodes

☒ DataNode is aware of the files to which the blocks stored on it belong to ✓

✓ For a HDFS directory the replication factor(RF) is 1/1

☐ A - same as the RF of the files in that directory

☐ B - Zero

☐ C - 3

☒ D - Does not apply. ✓

✓ For a Spark Program, driver takes 1Gb and 1 core and two instances of executors with 4Gb and 2 cores each,determine the resources and number of tasks running at a point in time. 1/1

☐ a. 8Gb 4cores and 4 tasks

☒ b. 9Gb 5cores and 4 tasks ✓

☐ c. 9Gb 5cores and 2 tasks

☐ d. 8Gb 5cores and 2 tasks



✗ The _____ is a framework-specific entity that negotiates resources from the ResourceManager 0/1

☐ a) NodeManager

☒ b) ApplicationManager ✗

- ☐ c) Resource Negotiator
- ☐ d) ApplicationMaster

Correct answer

- ☒ d) ApplicationMaster

This form was created inside of Media.Net. - Terms of Service - Additional Terms

Google Forms