

QUESTIONS & ANSWERS

Kill your exam at first Attempt



Cloudera

CCA-332

Cloudera Certified Administrator for Apache Hadoop

Reference:

http://docs.hortonworks.com/CURRENT/index.htm#Monitoring_HDP/Using_Nagios_With_Hadoop/Nagios_Alerts_For_Hadoop_Services.htm(Corrupt/missing block alert, Possible remedies, third bulleted point)

QUESTION: 26

For each job, the Hadoop framework generates task log files. Where are these log files stored?

- A. On the local disk of the slave node running the task.
- B. In HDFS, in the directory of the user who generates the job.
- C. On the local disk of the slave node running the task, then copied into HDFS.
- D. On the local disk of the slave node running the task, but logs are discarded immediately upon task completion.

Answer: B

QUESTION: 27

What additional capability does Ganglia provide to that Hadoop alone DOES NOT provide?

- A. Ability to monitor number of files in HDFS.
- B. Ability to monitor free task slots.
- C. Ability to monitor processor utilization.
- D. Ability to monitor NameNode memory usage.
- E. Ability to monitor the amount of free space on HDFS.

Answer: C

QUESTION: 28

Someone in your data center unplugs a slave node by accident. Users of the cluster notice via the Hadoop Web UI that the cluster size has shrunk and express concerns about data loss and HDFS performance. The replication factor of all the files in the cluster is unchanged from the default of 3. What can you tell the users?

- A. The HDFS filesystem is corrupt until the administrator re-adds the DataNode to the cluster. The warnings associated with the event should be reported.
- B. After identifying the outage, the NameNode will naturally re-replicate the data and there will be no data loss. The administrator can re-add the DataNode at any time. The client can disregard warnings concerned with this event. Data will be under-replicated but will become properly replicated over time.

C. The NameNode will re-replicate the data after the administrator issues a special command. The data is not lost but is under-replicated until the administrator issues this command.

D. The NameNode will identify the outage and re-replicate the data when the clients receive connection failures to the DataNode, so the end users can disregard such warnings.

Answer: A

QUESTION: 29

In a Hadoop cluster for running production jobs that require short turnaround times, the most important consideration for slave nodes is:

- A. The ratio between the amount of memory and the number of disk drives.
- B. The ratio between the amount of memory and the total storage capacity.
- C. The ratio between the number of processor cores and the amount of memory.
- D. The ratio between the number of processor cores and total storage capacity.
- E. The ratio between the number of processor cores and number of disk drives.

Answer: E

QUESTION: 30

What is the Secondary NameNode?

- A. An alternate data channel for clients to reach HDFS, should the NameNode become too busy.
- B. A process that performs a checkpoint operation on the files produced by the NameNode.
- C. A data channel between the primary name node and the tertiary NameNode.
- D. A process purely intended to perform backups of the NameNode.
- E. A standby NameNode, for high availability.

Answer: B

Reference:

[http://wiki.apache.org/hadoop/FAQ#What_is_the_purpose_of_the_secondary_name-node.3F\(3.2\)](http://wiki.apache.org/hadoop/FAQ#What_is_the_purpose_of_the_secondary_name-node.3F(3.2))

For More exams visit <https://killexams.com> -



[KILLEXAMS.COM](https://killexams.com)

Kill your exam at First Attempt....Guaranteed!