Pass2Lead
https://Pass2Lead.com

HADOOP-PR000007^{Q&As}

Hortonworks Certified Apache Hadoop 2.0 Developer (Pig and Hive Developer)

Pass Hortonworks HADOOP-PR000007 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass2lead.com/hadoop-pr000007.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Hortonworks Official Exam Center

🛠 **Instant Download** After Purchase

🛠 **100% Money Back** Guarantee

🛠 **365 Days** Free Update

🛠 **800,000+** Satisfied Customers

QUESTION 1

Which one of the following statements regarding the components of YARN is FALSE?

A. A Container executes a specific task as assigned by the ApplicationMaster

B. The ResourceManager is responsible for scheduling and allocating resources

C. A client application submits a YARW job to the ResourceManager

D. The ResourceManager monitors and restarts any failed Containers

Correct Answer: D

QUESTION 2

On a cluster running MapReduce v1 (MRv1), a TaskTracker heartbeats into the JobTracker on your cluster, and alerts the JobTracker it has an open map task slot.

What determines how the JobTracker assigns each map task to a TaskTracker?

A. The amount of RAM installed on the TaskTracker node.

B. The amount of free disk space on the TaskTracker node.

C. The number and speed of CPU cores on the TaskTracker node.

D. The average system load on the TaskTracker node over the past fifteen (15) minutes.

E. The location of the InsputSplit to be processed in relation to the location of the node.

Correct Answer: E

Explanation: The TaskTrackers send out heartbeat messages to the JobTracker, usually every few minutes, to reassure the JobTracker that it is still alive. These message also inform the JobTracker of the number of available slots, so the JobTracker can stay up to date with where in the cluster work can be delegated. When the JobTracker tries to find somewhere to schedule a task within the MapReduce operations, it first looks for an empty slot on the same server that hosts the DataNode containing the data, and if not, it looks for an empty slot on a machine in the same rack.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, How JobTracker schedules a task?

QUESTION 3

In a MapReduce job, you want each of your input files processed by a single map task. How do you configure a MapReduce job so that a single map task processes each input file regardless of how many blocks the input file occupies?

A. Increase the parameter that controls minimum split size in the job configuration.

B. Write a custom MapRunner that iterates over all key-value pairs in the entire file.

![Pass2Lead](https://Pass2Lead.com)
C. Set the number of mappers equal to the number of input files you want to process.

D. Write a custom FileInputFormat and override the method isSplitable to always return false.

Correct Answer: D

Explanation: FileInputFormat is the base class for all file-based InputFormats. This provides a generic implementation of getSplits(JobContext). Subclasses of FileInputFormat can also override the isSplitable (JobContext, Path) method to ensure input-files are not split-up and are processed as a whole by Mappers.

Reference: org.apache.hadoop.mapreduce.lib.input, Class FileInputFormat

---

QUESTION 4

Identify the MapReduce v2 (MRv2 / YARN) daemon responsible for launching application containers and monitoring application resource usage?

A. ResourceManager

B. NodeManager

C. ApplicationMaster

D. ApplicationMasterService

E. TaskTracker

F. JobTracker

Correct Answer: B

Reference: Apache Hadoop YARN ?Conceptsand; Applications

QUESTION 5

You have just executed a MapReduce job. Where is intermediate data written to after being emitted from the Mapper\\'s map method?

A. Intermediate data in streamed across the network from Mapper to the Reduce and is never written to disk.

B. Into in-memory buffers on the TaskTracker node running the Mapper that spill over and are written into HDFS.

C. Into in-memory buffers that spill over to the local file system of the TaskTracker node running the Mapper.

D. Into in-memory buffers that spill over to the local file system (outside HDFS) of the TaskTracker node running the Reducer

E. Into in-memory buffers on the TaskTracker node running the Reducer that spill over and are written into HDFS.

Correct Answer: C

Explanation: The mapper output (intermediate data) is stored on the Local file system (NOT HDFS) of each individual mapper nodes. This is typically a temporary directory location which can be setup in config by the hadoop administrator.

The intermediate data is cleaned up after the Hadoop Job completes.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, Where is the Mapper Output (intermediate kay-value data) stored ?

HADOOP-PR000007 Practice Test   HADOOP-PR000007 Study Guide   HADOOP-PR000007 Exam Questions