

APACHE-HADOOP-DEVELOPER^{Q&As}

Hadoop 2.0 Certification exam for Pig and Hive Developer

Pass Hortonworks APACHE-HADOOP-DEVELOPER Exam with 100% Guarantee

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

<https://www.leads4pass.com/apache-hadoop-developer.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

You have user profile records in your OLPT database, that you want to join with web logs you have already ingested into the Hadoop file system. How will you obtain these user records?

- A. HDFS command
- B. Pig LOAD command
- C. Sqoop import
- D. Hive LOAD DATA command
- E. Ingest with Flume agents
- F. Ingest with Hadoop Streaming

Correct Answer: C

Reference: Hadoop and Pig for Large-Scale Web Log Analysis

QUESTION 2

Examine the following Pig commands:

```
customers = LOAD 'data' USING PigStorage(',')
AS (name:chararray,age:int,zip:int,receipts:double);
s = SAMPLE customers 0.2;
DUMP s;
```

Which one of the following statements is true?

- A. The SAMPLE command generates an "unexpected symbol" error
- B. Each MapReduce task will terminate after executing for 0.2 minutes
- C. The reducers will only output the first 20% of the data passed from the mappers
- D. A random sample of approximately 20% of the data will be output

Correct Answer: D

QUESTION 3

Workflows expressed in Oozie can contain:

- A. Sequences of MapReduce and Pig. These sequences can be combined with other actions including forks, decision points, and path joins.
- B. Sequences of MapReduce job only; on Pig on Hive tasks or jobs. These MapReduce sequences can be combined

with forks and path joins.

C. Sequences of MapReduce and Pig jobs. These are limited to linear sequences of actions with exception handlers but no forks.

D. Iterntive repetition of MapReduce jobs until a desired answer or state is reached.

Correct Answer: A

Explanation: Oozie workflow is a collection of actions (i.e. Hadoop Map/Reduce jobs, Pig jobs) arranged in a control dependency DAG (Direct Acyclic Graph), specifying a sequence of actions execution. This graph is specified in hPDL (a XML Process Definition Language).

hPDL is a fairly compact language, using a limited amount of flow control and action nodes. Control nodes define the flow of execution and include beginning and end of a workflow (start, end and fail nodes) and mechanisms to control the workflow execution path (decision, fork and join nodes).

Workflow definitions Currently running workflow instances, including instance states and variables

Reference: Introduction to Oozie

Note: Oozie is a Java Web-Application that runs in a Java servlet-container - Tomcat and uses a database to store:

QUESTION 4

Given a directory of files with the following structure: line number, tab character, string: Example:

1abialkjjkaoasdfjksdlkjhweroij 2kadfjhuwqounahagtnbvaswslmnbfgy 3kjfteiomndscxeqalkzhtopedkfsikj You want to send each line as one record to your Mapper. Which InputFormat should you use to complete

the line: `conf.setInputFormat (____.class) ; ?`

A. SequenceFileAsTextInputFormat

B. SequenceFileInputFormat

C. KeyValueFileInputFormat

D. BDBInputFormat

Correct Answer: C

<http://stackoverflow.com/questions/9721754/how-to-parse-customwritable-from-text-in-hadoop>

QUESTION 5

A combiner reduces:

A. The number of values across different keys in the iterator supplied to a single reduce method call.

B. The amount of intermediate data that must be transferred between the mapper and reducer.

C. The number of input files a mapper must process.

D. The number of output files a reducer must produce.

Correct Answer: B

Explanation: Combiners are used to increase the efficiency of a MapReduce program. They are used to aggregate intermediate map output locally on individual mapper outputs. Combiners can help you reduce the amount of data that needs to be transferred across to the reducers. You can use your reducer code as a combiner if the operation performed is commutative and associative. The execution of combiner is not guaranteed, Hadoop may or may not execute a combiner. Also, if required it may execute it more than 1 times. Therefore your MapReduce jobs should not depend on the combiners execution.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, What are combiners? When should I use a combiner in my MapReduce Job?

[APACHE-HADOOP-DEVELOPER PDF Dumps](#)

[APACHE-HADOOP-DEVELOPER VCE Dumps](#)

[APACHE-HADOOP-DEVELOPER Practice Test](#)