



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.pass4itsure.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**

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. Iterative 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 2

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 YARN job to the ResourceManager
- D. The ResourceManager monitors and restarts any failed Containers

Correct Answer: D

QUESTION 3

Assuming default settings, which best describes the order of data provided to a reducer's reduce method:

- A. The keys given to a reducer aren't in a predictable order, but the values associated with those keys always are.



- B. Both the keys and values passed to a reducer always appear in sorted order.
- C. Neither keys nor values are in any predictable order.
- D. The keys given to a reducer are in sorted order but the values associated with each key are in no predictable order

Correct Answer: D

Explanation: Reducer has 3 primary phases:

1.

Shuffle

The Reducer copies the sorted output from each Mapper using HTTP across the network.

2.

Sort

The framework merge sorts Reducer inputs by keys (since different Mappers may have output the same key).

The shuffle and sort phases occur simultaneously i.e. while outputs are being fetched they are merged.

SecondarySort

To achieve a secondary sort on the values returned by the value iterator, the application should extend the key with the secondary key and define a grouping comparator. The keys will be sorted using the entire key, but will be grouped using the grouping comparator to decide which keys and values are sent in the same call to reduce.

3. Reduce

In this phase the `reduce(Object, Iterable, Context)` method is called for each in the sorted inputs. The output of the reduce task is typically written to a `RecordWriter` via `TaskInputOutputContext.write`

`(Object, Object)`.

The output of the Reducer is not re-sorted.

Reference: `org.apache.hadoop.mapreduce, Class`

Reducer

QUESTION 4

For each input key-value pair, mappers can emit:

- A. As many intermediate key-value pairs as designed. There are no restrictions on the types of those key-value pairs (i.e., they can be heterogeneous).
- B. As many intermediate key-value pairs as designed, but they cannot be of the same type as the input key-value pair.
- C. One intermediate key-value pair, of a different type.
- D. One intermediate key-value pair, but of the same type.



E. As many intermediate key-value pairs as designed, as long as all the keys have the same types and all the values have the same type.

Correct Answer: E

Explanation: Mapper maps input key/value pairs to a set of intermediate key/value pairs.

Maps are the individual tasks that transform input records into intermediate records. The transformed intermediate records do not need to be of the same type as the input records. A given input pair may map to zero or many output pairs.

Reference: Hadoop Map-Reduce Tutorial

QUESTION 5

To use a Java user-defined function (UDF) with Pig what must you do?

- A. Define an alias to shorten the function name
- B. Pass arguments to the constructor of UDFs implementation class
- C. Register the JAR file containing the UDF
- D. Put the JAR file into the user's home folder in HDFS

Correct Answer: C

[Latest HADOOP-PR000007 Dumps](#)

[HADOOP-PR000007 VCE Dumps](#)

[HADOOP-PR000007 Study Guide](#)