



Free Questions for [ACA-BigData1](#) by [certsdeals](#)

Shared by [Duffy](#) on [09-08-2024](#)

For More Free Questions and Preparation Resources

[Check the Links on Last Page](#)

Question 1

Question Type: MultipleChoice

DataWorks can be used to create all types of tasks and configure scheduling cycles as needed. The supported granularity levels of scheduling cycles include days, weeks, months, hours, minutes and seconds.

Options:

A- True

B- False

Answer:

B

Question 2

Question Type: MultipleChoice

One Alibaba Cloud account is entitled to join only one organization that uses DataWorks.

Options:

A- True

B- False

Answer:

A

Question 3

Question Type: MultipleChoice

Alibaba Cloud E-MapReduce can be easily plugged with other Alibaba Cloud services such as Log Service, ONS, MNS that act as data ingestion channels from real-time data streams. Which of the following descriptions about real-time processing are correct? (Number of correct answers: 3)

Options:

- A- This data is streamed and processed using Apache Flume or Kafka in integration with Apache Storm using complex algorithms
- B- Kafka is usually preferred with Apache Storm to provide data pipeline
- C- The final processed data can be stored in HDFS, HBase or any other big data store service in real time.
- D- Apache Sqoop is used to do the real-time data transmission of structured data

Answer:

A, B, C

Question 4

Question Type: MultipleChoice

In DataWorks, a task should be instantiated first before a scheduled task is running every time, that is, generating a corresponding instance which is executed for running the scheduled task. The status is different in each phase of the scheduling process, including _____. (Number of correct answers: 3)

Options:

A- Not running

B- Running

C- Running Successfully

Answer:

A, B, C

Question 5

Question Type: MultipleChoice

In order to improve the processing efficiency when using MaxCompute, you can specify the partition when creating a table. That is, several fields in the table are specified as partition columns. Which of the following descriptions about MaxCompute partition table are correct? (Number of correct answers: 4)

Options:

A- In most cases, user can consider the partition to be the directory under the file system

B- User can specify multiple partitions, that is, multiple fields of the table are considered as the partitions of the table, and the relationship among partitions is similar to that of multiple directories

C- If the partition columns to be accessed are specified when using data, then only corresponding partitions are read and full table scan

is avoided, which can improve the processing efficiency and save costs

D- MaxCompute partition only supports string type and the conversion of any other types is not allowed

E- The partition value cannot have a double byte characters (such as Chines.

Answer:

A, B, C, E

Question 6

Question Type: MultipleChoice

Which of the following task types does DataWorks support?

(Number of correct answers: 4)

Options:

A- Data Synchronization

B- SHELL

C- MaxCompute SQL

D- MaxCompute MR

E- Scala

Answer:

A, B, C, D

Question 7

Question Type: MultipleChoice

In DataWorks, we can configure alert policies to monitor periodically scheduled tasks, so that an alert will be issued timely. Currently DataWorks supports _____ alerts.

(Number of correct answers: 2)

Options:

A- Email

B- Text message

C- Telephone

C- Aliwangwang

Answer:

A, B

Question 8

Question Type: MultipleChoice

You are working on a project where you need to chain together MapReduce, Hive jobs. You also need the ability to use forks, decision points, and path joins. Which ecosystem project should you use to perform these actions?

Options:

A- Apache HUE

B- Apache Zookeeper

C- Apache Oozie

D- Apache Spark

Answer:

C

Question 9

Question Type: MultipleChoice

Your company stores user profile records in an OLTP databases. You want to join the serecords with web server logs you have already ingested into the Hadoop file system.

What is the best way to obtain and ingest these user records?

Options:

A- Ingest with Hadoop streaming

B- Ingest using Hive

C- Ingest with sqoop import

D- Ingest with Pig's LOAD command

Answer:

C

Question 10

Question Type: MultipleChoice

You want to understand more about how users browse your public website. For example, you want to know which pages they visit prior to placing an order. You have a server farm of 100 web servers hosting your website. Which is the most efficient process to gather these web servers across logs into traditional Hadoop ecosystem.

Options:

- A- Just copy them into HDFS using curl
- B- Ingest the server web logs into HDFS using Apache Flume
- C- Channel these clickstreams into Hadoop using Hadoop Streaming
- D- Import all user clicks from your OLTP databases into Hadoop using Sqoop

Answer:

B

Question 11

Question Type: MultipleChoice

By integrating live dashboards, DataV can present and monitor business data simultaneously. This data-driven approach enables for well-organized data mining and analysis allowing the user to seize new opportunities that otherwise might remain hidden. It can support wide range of databases and data formats. Which of the following options DataV does not support?

Options:

- A- Alibaba Cloud' s AnalyticDB, ApsaraDB
- B- Static data in CSV and JSON formats
- C- Oracle Database
- D- MaxCompute Project

Answer:

D

Question 12

Question Type: MultipleChoice

DataV is a powerful yet accessible data visualization tool, which features geographic information systems allowing for rapid interpretation of data to understand relationships, patterns, and trends. When a DataV screen is ready, it can embed works to the existing portal of the enterprise through _____.

Options:

- A- URL after the release
- B- URL in the preview
- C- MD5 code obtained after the release
- D- Jar package imported after the release

Answer:

A

To Get Premium Files for ACA-BigData1 Visit

<https://www.p2pexams.com/products/aca-bigdata1>

For More Free Questions Visit

<https://www.p2pexams.com/alibaba/pdf/aca-bigdata1>

