AlibabaCloud.ACA-BigData1.v2021-05-21.q26

Exam Code:	ACA-BigData1
Exam Name:	ACA Big Data Certification Exam
Certification Provider:	Alibaba Cloud
Free Question Number:	26
Version:	v2021-05-21
# of views:	303
# of Questions views:	1977
https://www.freecram.com/torrent/AlibabaCloud.ACA-BigData1.v2021-05-21.q26.html	

NEW QUESTION: 1

If a MySQL database contains 100 tables, and jack wants to migrate all those tables to MaxCompute using DataWorks Data Integration, the conventional method would require him to configure 100 data synchronization tasks. With _____ feature in DataWorks, he can upload all tables at the same time.

Score 2

A. Configure a MySQL Writer plug-in

B. Add data sources in Bulk Mode

C. Full-Database Migration feature

D. Configure a MySQL Reader plug-in

Answer: B (LEAVE A REPLY)

NEW QUESTION: 2

JindoFS in E-MapReduce provided by SmartData uses OSS as the storage back end.

Score 1

A. False

B. True

Answer: B (LEAVE A REPLY)

NEW QUESTION: 3

Alibaba Cloud Quick BI reporting tools support a variety of data sources, facilitating users to analyze and present their data from different data sources. _____ is not supported as a data source yet.

Score 2

A. MySQL RDS

B. Results returned from the API

C. Local Excel files

D. MaxCompute

Answer: B (LEAVE A REPLY)

NEW QUESTION: 4

DataWorks can be used to develop and configure data sync tasks. Which of the following statements are correct? (Number of correct answers: 3) Score 2

A. The data source configuration in the project management is required to add data source

B. For the extraction of source data, "where" filtering clause can be referenced as the criteria of incremental synchronization

C. Clean-up rules can be set to clear or preserve existing data before data write

D. Some of the columns in source tables can be extracted to create a mapping relationship between fields, and constants or variables can't be added

Answer: A,C,D (LEAVE A REPLY)

NEW QUESTION: 5

When odpscmd is used to connect to a project in MaxCompute, the command _____ can be executed to view the size of the space occupied by table table_a.

Score 2

A. size table_a;

B. select size from table_a;

C. show table table_a;

D. desc table_a;

Answer: (SHOW ANSWER)

NEW QUESTION: 6

MaxCompute supports two kinds of charging methods: Pay-As-You-Go and Subscription (CU cost).

Pay-As-You-Go means each task is measured according to the input size by job cost. In this charging method the billing items do not include charges due to _____.

Score 2

A. Data download

B. Storage

C. Computing

D. Data upload

Answer: A (LEAVE A REPLY)

NEW QUESTION: 7

DataWorks provides two billing methods: Pay-As-You-Go (post-payment) and subscription (prepayment).

When DataWorks is activated in pay-as-you-go mode, Which of the following billing items will not apply?

Score 2

A. Task nodes created by developer

B. Baseline instances generated by Intelligent Monitor and Data Quality checks

C. Calls and execution time of APIs compiled in DataService Studio

D. Shared resource groups for scheduling and Data Integration instances

Answer: (SHOW ANSWER)

NEW QUESTION: 8

MaxCompute is a fast and fully-managed TB/PB-level data warehousing solution provided by Alibaba Cloud. Which of the following product features are correct? _____ (Number of correct answers: 3) Score 2

A. High security and reliability

B. Efficient transaction processing

C. Fast real-time response

D. Multi-level management and authorization

E. Distributed architecture

Answer: A,C,E (LEAVE A REPLY)

NEW QUESTION: 9

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?

A. Ingest using Hive

B. Ingest with sqoop import

C. Ingest with Hadoop streaming

D. Ingest with Pig's LOAD command

Answer: B (LEAVE A REPLY)

NEW QUESTION: 10

A start-up company wants to use Alibaba Cloud MaxCompute to provide product recommendation services for its users. However, the company does not have much users at the initial stage, while the charge for MaxCompute is higher than that of ApsaraDB RDS, so the company should be recommended to use MaxCompute service until the number of its users increases to a certain size.

Score 1

A. True

B. False

Answer: A (LEAVE A REPLY)

NEW QUESTION: 11

There are various methods for accessing to MaxCompute, for example, through management console, client command line, and Java API. Command line tool odpscmd can be used to create, operate, or delete a table in a project.
Score 1
A. False
B. True
Answer: B (LEAVE A REPLY)

NEW QUESTION: 12

In order to ensure smooth processing of tasks in the Dataworks data development kit, you must create an AccessKey. An AccessKey is primarily used for access permission verification between various Alibaba Cloud products. The AccessKey has two parts, they are _____. (Number of correct answers: 2) Score 2
A. Access Username
B. Access Password
C. Access Key Secret
D. Access Key ID
Answer: (SHOW ANSWER)

NEW QUESTION: 13

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?
A. Apache HUE
B. Apache Zookeeper
C. Apache Spark
D. Apache Oozie
Answer: D (LEAVE A REPLY)

NEW QUESTION: 14

Where is the meta data(e.g.,table schemas) in Hive?
Score 2
A. Stored in ZooKeeper
B. Stored along with the data in HDFS
C. Stored in the RDBMS like MySQL
D. Stored as metadata on the NameNode
Answer: (SHOW ANSWER)

NEW QUESTION: 15

MaxCompute Tunnel provides high concurrency data upload and download services. User can use the Tunnel service to upload or download the data to MaxCompute. Which of the following descriptions about Tunnel is NOT correct:

Score 2

A. MaxCompute Tunnel provides the Java programming interface for usersB. MaxCompute provides two data import and export methods: using Tunnel Operation on the console directly or using TUNNEL written with java

B. Tunnel commands are mainly used to upload or download data.They provide the following functions:upload, download, resume, show, purge etc.

C. If data fails to be uploaded, use the restore command to restore the upload from where it was interrupted

Answer: B (LEAVE A REPLY)

NEW QUESTION: 16

MaxCompute is a general purpose, fully managed, multi-tenancy data processing platform for largescale data warehousing, and it is mainly used for storage and computing of batch structured data. Which of the following is not a use case for MaxCompute?

Score 2

A. User profile

B. Order management

C. Date Warehouse

D. Social networking analysis

Answer: B (LEAVE A REPLY)

NEW QUESTION: 17

Scenario: Jack is the administrator of project prj1. A new team member, Alice (already has an Alibaba Cloud account alice@aliyun.com), applies for joining this project with the following permissions: view table lists, submit jobs, and create tables. Which of the following SQL statements is useless:

Score 2

A. grant List, CreateTable, CreateInstance on project prj1 to user aliyun$alice@aliyun;

B. flush privileges;

C. use prj1;

D. add user aliyun$alice@aliyun.com;

Answer: B (LEAVE A REPLY)

NEW QUESTION: 18

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)

A. Apache Sqoop is used to do the real-time data

transmission of structured data

B. This data is streamed and processed using Apache

Flume or Kafka in integration with Apache Storm using

complex algorithmsB. 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.

Answer: A,B,C (LEAVE A REPLY)

NEW QUESTION: 19

MaxCompute SQL is suitable for processing less real-time massive data, and employs a
syntax similar to that of SQL. The efficiency of data query can be improved through
creating proper indexes in the table.

Score 1

A. True

B. False

Answer: B (LEAVE A REPLY)

NEW QUESTION: 20

A dataset includes the following items (time, region, sales amount). If you want to present
the information above in a chart, _____ is applicable.

Score 2

A. Tree Chart

B. Pie Chart

C. Radar Chart

D. Bubble Chart

Answer: D (LEAVE A REPLY)

NEW QUESTION: 21

In MaxCompute, you can use Tunnel command line for data upload and download. Which of the following description of Tunnel command is NOT correct:

Score 2

A. Upload: Supports file or directory (level-one) uploading. Data can only be uploaded to a single table or table partition each time.

B. Purge: Clears the table directory. By default, use this command to clear information of the last three days.

C. Resume: If an error occurs due to the network or the Tunnel service, you can resume transmission of the file or directory after interruption.

D. Download: You can only download data to a single file. Only data in one table or partition can be downloaded to one file each time. For partitioned tables, the source partition must be specified.

Answer: D (LEAVE A REPLY)

NEW QUESTION: 22

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.

A. Channel these clickstreams into Hadoop using Hadoop Streaming

B. Ingest the server web logs into HDFS using Apache Flume

C. Import all user clicks from your OLTP databases into Hadoop using Sqoop

D. Just copy them into HDFS using curl

Answer: B (LEAVE A REPLY)

NEW QUESTION: 23

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

A. True

B. False

Answer: (SHOW ANSWER)

NEW QUESTION: 24

Alibaba Cloud Elastic MapReduce (E-MapReduce) is a big data processing solution to quickly process huge amounts of data. Based on open source Apache Hadoop and Apache Spark, E-MapReduce flexibly manages your big data use cases such as trend analysis, data warehousing, and analysis of continuously streaming data.

Score 1

A. False

B. True

Answer: B (LEAVE A REPLY)

NEW QUESTION: 25

Data Migration Unit (DMU) is used to measure the amount of resources consumed by data integration, including CPU, memory, and network. One DMU represents the minimum amount of resources used for a data synchronization task.

Score 1

A. False

B. True

Answer: ([SHOW ANSWER](#))

NEW QUESTION: 26

In each release of E-MapReduce, the software and software version are flexible. You can select multiple software versions.

Score 1

A. False

B. True

Answer: A ([LEAVE A REPLY](#))