



"Please note that these files may not be up to date. However, the questions will help you understand the exam format and typical question patterns."

www.atmicnetworks.com

Warning: Keep connected with our support team for latest updates

Question:1

Scenario: Jack is the administrator of project prj1. The project involves a large volume of sensitive data such as bank account, medical record, etc. Jack wants to properly protect the data. Which of the follow statements is necessary?

- A. set ProjectACL=true;
- B. add accountprovider ram;
- C. set ProjectProtection=true;
- D. use prj1;

Answer: C

Question:2

Where is the meta data (e.g.,table schemas) in Hive?

- A. Stored as metadata on the NameNode
- B. Stored along with the data in HDFS
- C. Stored in the RDBMS likeMySQL
- D. Stored in ZooKeeper

Answer: C

Question:3

MaxCompute tasks contain computational tasks and non-computational tasks. The computational tasks require actual operations on data stored in the table. MaxCompute parses the task to obtain its execution plan, and submits the task for execution. The noncomputational tasks require substantial reading of and modification to metadata information. Therefore, the task is not parsed, and no execution plan is provided. The task is directly submitted for execution. The latter one has a faster response speed than the former one. Which of the following operations on the table t_test is a computational task?

- A. desc t_test
- B. alter table t_test add columns (comments string);
- C. select count(*) from t_test;
- D. truncate table t_test;

Answer: C

Question:4

When we use the MaxCompute tunnel command to upload the log.txt file to the t_log table, the t_log is a partition table and the partitioning column is (p1 string, p2 string). Which of the following commands is correct?

- A. tunnel upload log.txt t_log/p1="b1", p2="b2"
- B. tunnel upload log.txt t_log/(p1="b1", p2="b2")
- C. tunnel upload log.txt t_log/p1="b1"/p2="b2"

Answer: A

Question:5

A Log table named log in MaxCompute is a partition table, and the partition key is dt. A new partition is created daily to store the new data of that day. Now we have one month's data, starting from dt='20180101' to dt='20180131', and we may use to delete the data on 20180101.

- A. delete from log where dt='20180101'
- B. truncate table where dt='20180101'
- C. drop partition log (dt='20180101')
- D. alter table log drop partition(dt='20180101')

Answer: D

Question:6

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 .

- 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

Question:7

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?

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

Answer: D

Question:8

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.

Ingest the server web logs into HDFS using Apache

- A. Just copy them into HDFS using curl
- B. 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: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 with Hadoop streaming
- B. Ingest using Hive
- C. Ingest with sqoop import
- D. Ingest with Pig's LOAD command

Answer: C

Question: 10

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 Oozie
- D. Apache Spark

Answer: C

Question: 11

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)

- A. Email
- B. Text message
- C. Telephone
- C. Aliwangwang

Answer: A, B

Question: 12

Which of the following task types does DataWorks support? (Number of correct answers: 4)

- A. Data Synchronization
- B. SHELL
- C. MaxCompute SQL
- D. MaxCompute MR
- E. Scala

Answer: A, B, C,D

Question: 13

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)

- A. In most cases, user can consider the partition to be the directory under the filesystem
- 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 Chinese).

Answer: A, B, C, E

Question: 14

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)

- A. Not running
- B. Running
- C. Running Successfully

Answer: A, B,C

Question: 15

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. 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: 16

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

- A. True
- B. False

Answer: A

Question: 17

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.

- A. True
- B. False

Answer: B

Question: 18

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.

- A. True
- B. False

Answer: B

Question: 19

Function Studio is a web project coding and development tool independently developed by the Alibaba Group for function development scenarios. It is an important component of DataWorks. Function Studio supports several programming languages and platform-based function development scenarios except for .

Score 2

- A. Real-time computing
- B. Python
- C. Java
- D. Scala

Answer: D

Question: 20

A business flow in DataWorks integrates different node task types by business type, such a structure improves business code development facilitation. Which of the following descriptions about the node type is INCORRECT?

Score 2

- A. A zero-load node is a control node that does not generate any data. The virtual node is generally used as the root node for planning the overall node workflow.
- B. An ODPS SQL task allows you to edit and maintain the SQL code on the Web, and easily implement code runs, debug, and collaboration.
- C. The PyODPS node in DataWorks can be integrated with MaxCompute Python SDK. You can edit the Python code to operate MaxCompute on a PyODPS node in DataWorks.
- D. The SHELL node supports standard SHELL syntax and the interactive syntax. The SHELL task can run on the default resource group

Answer: C

Question: 21

Apache Spark included in Alibaba E-MapReduce(EMR) is a fast and general-purpose cluster computing system. It provides high-level APIs in Java, Scala, Python and R, and an optimized engine that supports general execution graphs. It also supports a rich set of higher-level tools. Which of the following tools does not be included in Spark?

Score 2

- A. Spark SQL for SQL and structured data processing
- B. MLlib for machine learning
- C. GraphX for graph processing
- D. TensorFlow for AI

Answer: D

Question: 22

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. Shared resource groups for scheduling and Data Integration instances
- B. Baseline instances generated by Intelligent Monitor and Data Quality checks
- C. Task nodes created by developer
- D. Calls and execution time of APIs compiled in DataService Studio

Answer: C

Question: 23

Users can use major BI tools, such as Tableau and FineReport, to easily connect to MaxCompute projects, and perform BI analysis or ad hoc queries. The quick query feature in MaxCompute is called allows you to provide services by encapsulating project table data in APIs, supporting diverse application scenarios without data migration.

Score 2

- A. Lightning
- B. MaxCompute Manager
- C. Tunnel
- D. Labelsecurity

Answer: A

Question: 24

If a MySQL database contains 100 tables, and Jack wants to migrate all those tables to MaxCompute using DataWorks DataIntegration, 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. Full-Database Migration feature
- B. Configure a MySQL Reader plug-in
- C. Configure a MySQL Writer plug-in
- D. Add data sources in BulkMode

Answer: D

Question: 25

Machine Learning Platform for Artificial Intelligence (PAI) node is one of the node types in DataWorks business flow. It is used to call tasks created on PAI and schedule production activities based on the node configuration. PAI nodes can be added to DataWorks only

Score 2

- A. after PAI experiments are created on PAI
- B. after PAI service is activated
- C. after MaxCompute service is activated
- D. Spark on MaxCompute Machine Learning project is created

Answer: A

Question: 26

In a scenario where a large enterprise plans to use MaxCompute to process and analyze its data, tens of thousands of tables and thousands of tasks are expected for this project, and a project team of 40 members is responsible for the project construction and O&M. From the perspective of engineering, which of the following can considerably reduce the cost of project construction and management?

Score 2

- A. Develop directly on MaxCompute and use script-timed scheduling tasks
- B. Use DataWorks
- C. Use Eclipse
- D. Use a private platform specially developed for this project

Answer: B

Question: 27

AliOrg Company plans to migrate their data with virtually no downtime. They want all the data changes to the source database that occur during the migration are continuously replicated to the target, allowing the source database to be fully operational during the migration process. After the database migration is completed, the target database will remain synchronized with the source for as long as you choose, allowing you to switch over the database at a convenient time. Which of the

following Alibaba products is the right choice for you to do it:

Score 2

- A. Log Service
- B. DTS(Data Transmission Service)
- C. Message Service
- D. CloudMonitor

Answer: B

Question: 28

There are three types of node instances in an E-MapReduce cluster: master, core, and

Score 2

- A. task
- B. zero-load
- C. gateway

D. agent

Answer: A

Question: 29

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. Bubble Chart
- B. Tree Chart
- C. Pie Chart
- D. Radar Chart

Answer: A

Question: 30

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. Results returned from theAPI
- B. MaxCompute
- C. Local Excel files
- D. MySQL RDS

Answer: A

Question: 31

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

Score 2

- A. URL after the release
- B. URL in the preview
- C. MD5 code obtained afterthe release

D. Jar package imported after therelease

Answer: A

Question: 32

Where is the meta data(e.g.,table schemas) in Hive?

Score 2

- A. Stored as metadata on the NameNode
- B. Stored along with the data in HDFS
- C. Stored in the RDBMS likeMySQL
- D. Stored in ZooKeeper

Answer: C

Question: 33

instances in E-MapReduce are responsible for computing and can quickly add computing power to a cluster. They can also scale up and down at any time without impacting the operations of the cluster.

Score 2

- A. Task
- B. Gateway
- C. Master
- D. Core

Answer: C

Question: 34

Your company stores user profile records in an OLTP databases. You want to join these records with web server logs you have already ingested into the Hadoop file system. What is the best way to obtain and ingestthese user records?

Score 2

- A. Ingest with Hadoop streaming
- B. Ingest using Hive
- C. Ingest with sqoop import
- D. Ingest with Pig's LOAD command

Answer: C

Question: 35

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?

Score 2

- A. Spark
- B. HUE
- C. Zookeeper
- D. Oozie

Answer: D

Question: 36

Which node type in DataWorks can edit the Python code to operate data in MaxCompute?

Score 2

- A. PyODPS
- B. ODPS MR Node
- C. ODPS Script Node
- D. SHELL node

Answer: A

Question: 37

DataService Studio in DataWorks aims to build a dataservice bus to help enterprises centrally manage private and public APIs. DataService Studio allows you to quickly create APIs based on data tables and register existing APIs with the DataService Studio platform for centralized management and release. Which of the following descriptions about DataService Studio in DataWorks is INCORRECT? Score 2

- A. DataService Studio is connected to API Gateway. Users can deploy APIs to API Gateway with oneclick.
- B. DataService Studio adopts the serverless architecture. All you need to care is the query logic of APIs, instead of the infrastructure such as the running environment.
- C. To meet the personalized query requirements of advanced users, DataService Studio provides the custom Python script mode to allow you compile the API query by yourself. It also supports multitable association, complex query conditions, and aggregate functions.
- D. Users can deploy any APIs created and registered in DataService Studio to API Gateway for management, such as API

authorization and authentication, traffic control, and metering

Answer: C

Question: 38

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 users
- B. MaxCompute provides two data import and export methods: using Tunnel Operation on the console directly or using TUNNEL written with java
- C. If data fails to be uploaded, use the restore command to restore the upload from where it was interrupted
- D. Tunnel commands are mainly used to upload or download data.They provide the following functions:upload, download, resume, show, purge etc.

Answer: C

Question: 39

Which of the following is not proper for granting the permission on a L4MaxCompute table to a user. (L4 is a level in MaxCompute Label-based security (LabelSecurity), it is a required MaxCompute Access Control (MAC) policy at the project space level. It allows project administrators to control the user access to column-level sensitive data with improved flexibility.)

Score 2

- A. If no permissions have been granted to the user and the user does not belong to the project, add the user to the project. The user does not have any permissions before they are added to the project.
- B. Grant a specific operation permission to the user.
- C. If the user manages resources that have labels, such as datasheets and packages with datasheets, grant label permissions to the user.
- D. The user need to create a project in simple mode

Answer: D

Question: 40

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 upload
- B. Data download
- C. Computing
- D. Storage

Answer: B

Question: 41

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. Order management
- B. Date Warehouse
- C. Social networking analysis
- D. User profile

Answer: A

Question: 42

Tomistheadministratorofaprojectprj1 inMaxCompute. Theproject involvesalargevolume of sensitive data such as user IDs and shopping records, and many data mining algorithms with proprietary intellectual property rights. Tom wants to properly protect these sensitive data and algorithms. To be specific, project users can only access the data within the project, all data flows only within the project.

What operation should he perform?

Score 2

- A. Use ACL authorization to set the status to read-only for all users
- B. Use Policy authorization to set the status to read-only for all users
- C. Allow the object creator to access the object
- D. Enable the data protection mechanism in the project, using set ProjectProtection=true;

Answer: D

Question: 43

There are multiple connection clients for MaxCompute, which of the following is the easiest way to configure workflow and scheduling for MaxCompute tasks?

Score 2

- A. Use DataWorks

- B. Use Intelij IDEA
- C. Use MaxCompute Console
- D. No supported tool yetc

Answer: A

Question: 44

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

Answer: B

Question: 45

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. use prj1;
- B. add user aliyun\$alice@aliyun.com;
- C. grant List, CreateTable, CreateInstance on project prj1 to user aliyun\$alice@aliyun;
- D. flush privileges;

Answer: D

Question: 46

commark is an open-source framework that functions on the service level to support data processing and analysis operations. Equipped with unified computing resources and data set permissions, Spark ON

MaxCompute allows you to submit and run jobs while using your preferred development methods.

Which of the following descriptions about Spark on MaxCompute is NOT correct:

Score 2

- A. Spark on MaxCompute provides you with native Spark Web UIs.
- B. Different versions of Spark can run in MaxCompute at the same time.
- C. Similar to MaxCompute SQL and MaxCompute MapReduce, Spark on MaxCompute runs in the unified computing resources activated for MaxCompute projects.
- D. Spark on MaxCompute has a separate permission system which will not allow users to query data without any additional permission modifications required.

Answer: B

Question: 47

In MaxCompute command line, if you want to view all tables in a project, you can execute command:

Score 2

- A. show tables;
- B. use tables;
- C. desc tables;
- D. select tables;

Answer: A

Question: 48

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. select size from table_a;
- B. size table_a;
- C. desc table_a;
- D. show table table_a;

Answer: C

Question: 49

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. True
- B. False

Answer: A

Question: 50

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.

Score 1

- A. True
- B. False

Answer: B

Question: 51

If a task node of DataWorks is deleted from the recycle bin, it can still be restored.

- A. True
- B. False

Answer: A

Question: 52

If the DataWorks(MaxCompute) tables in your request belong to two owners. In this case, Data Guard(DataWorks component) automatically splits your request into two by table owner.

Score 1

- A. True
- B. False

Answer: A

Question: 53

The FTP data source in DataWorks allows you to read/write data to FTP, and supports configuring synchronization tasks in wizard and script mode.

Score 1

- A. True
- B. False

Answer: A

Question: 54

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

Score 1

- A. True
- B. False

Answer: B

Question: 55

Alibaba Cloud Elastic MapReduce (E-MapReduce) is a big data processing solution to quickly process huge amounts of data.

a. 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. True
- B. False

Answer: A

Question: 56

An enterprise uses Alibaba Cloud MaxCompute for storage of service orders, system logs and management data. Because the security levels for the data are different, it is needed to register multiple

Alibaba Cloud accounts for data management.

Score 1

- A. True
- B. False

Answer: B

Question: 57

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

Score 1

- A. True
- B. False

Answer: A

Question: 58

In DataWorks table permission system, you can revoke permissions only on the fields whose security level is higher than the security level of your account.

Score 1

- A. True
- B. False

Answer: A

Question: 59

Project is an important concept in MaxCompute. A user can create multiple projects, and each object belongs to a certain project.

Score 1

- A. True
- B. False

Answer: A

Question: 60

Assume that Task 1 is configured to run at 02:00 each day. In this case, the scheduling system automatically generates a snapshot at the time predefined by the periodic node task at 23:30 each day.

That is, the instance of Task 1 will run at 02:00 the next day. If the system detects the upstream task is complete, the system automatically runs the Task 1 instance at 02:00 the next day.

Score 1

- A. True
- B. False

Answer: A

Question: 61

In MaxCompute, if error occurs in Tunnel transmission due to network or Tunnel service, the user can resume the last update operation through the command tunnel resume;

Score 1

- A. True
- B. False

Answer: A

Question: 62

A company originally handled the local data services through the Java programs. The local data have been migrated to MaxCompute on the cloud, now the data can be accessed through modifying the Java code and using the Java APIs provided by MaxCompute.

Score 1

- A. True
- B. False

Answer: A

Question: 63

MaxCompute takes Project as a charged unit. The bill is charged according to three aspects: the usage of storage, computing resource, and data download respectively. You pay for compute and storage resources by the day with no long-term commitments.

Score 1

- A. True
- B. False

Answer: B

Question: 64

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. True
- B. False

Answer: A

Question: 65

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

Question: 66

Synchronous development in DataWorks provides both wizard and script modes.

Score 1

- A. True
- B. False

Answer: A

Question: 67

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

Question: 68

Table is a data storage unit in MaxCompute. It is a two-dimensional logical structure composed of rows and columns. All data is stored in the tables. Operating objects of computing tasks are all tables.

A user can perform create table, drop table, and tunnel upload as well as update the qualified data in the table.
Score 1

- A. True
- B. False

Answer: A

Question: 69

Which of the following Hadoop ecosystem componets can you choose to setup a streaming log analysis system?(Number of correct answers: 3)

Score 2

- A. Apache Flume
- B. Apache Kafka
- C. Apache Spark
- D. Apache Lucene

Answer:A,C,D

Question: 70

A distributed file system like GFS and Hadoop are design to have much larger block(or chunk) size like 64MB or 128MB, which of the following descriptions are correct? (Number of correct answers: 4) SCORE 2

- A. It reduces clients' need to interact with the master because reads and writes on the same block(or chunk) require only one initial request to the master for block location information
- B. Since on a large block(or chunk), a client is more likely to perform many operations on a given block, it can reduce network overhead by keeping a persistent TCP connection to the metadata server over an extended

period of time

C. It reduces the size of the metadata stored on the master

D. The servers storing those blocks may become hot spots if many clients are accessing the same small files

E. If necessary to support even larger file systems, the cost of adding extra memory to the meta data server is a big price

Answer: A, B, C, D

Question: 71

MaxCompute can coordinate multiple users to operate one project through ACL authorization. The objects that can be authorized by ACL include _____ . (Number of correct answers: 3)

Score 2

A. Project

B. Table

C. Resource

D. Procedure

Answer: A, B,C

Question: 72

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. 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

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

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

Answer: A, B,D

Question: 73

The data development mode in DataWorks has been upgraded to the three-level structure comprising of _____ , and _____ .

(Number of correct answers: 3)

Score 2

- A. Project
- B. Solution
- C. Business flow
- D. Directory

Answer: A, B,D

Question: 74

The data development mode in DataWorks has been upgraded to the three-level structure comprising of _____, and _____. (Number of correct answers: 3)

Score 2

- A. Project
- B. Solution
- C. Business flow
- D. Directory

Answer: A, B

Question: 75

DataWorks provides powerful scheduling capabilities including time-based or dependency-based task trigger mechanisms to perform tens of millions of tasks accurately and punctually each day based on DAG relationships. It supports multiple scheduling frequency configurations like: (Number of correct answers: 4)

Score 2

- A. By Minute
- B. By Hour
- C. By Day
- D. By Week
- E. By Second

Answer: A, B, C, D

Question: 76

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. Distributed architecture

- B. High security and reliability
- C. Multi-level management and authorization
- D. Efficient transaction processing
- E. Fast real-time response

Answer: A, B,E