Azure Question Answers

06 October 2024   10:24

- In which service model does the cloud provider manage the underlying infrastructure, allowing users to focus on application development?

  IaaS
  PaaS
  SaaS
  MaaS

  Ans : PaaS

- Which service model provides virtualized computing resources over the internet?

  IaaS
  PaaS
  SaaS
  CaaS

  Ans : IaaS

- You can implement cloud virtualization to save time and energy as well as for decreasing the cost. Which of the following are the results of performing the same?

  A. Providing ability to manage resources effectively
  B. Provides for data loss prevention
  C. Provides secure remote access.

  Options :
  Only A and B
  Only B and C
  Only A and C
  All of these

  Ans : Only A and C

- Which among the following services does the IaaS cloud service model NOT provide?

  Compute
  Load Balancers
  URL Fetch
  Storage

  Ans : URL Fetch

- In the context of service models, what does PaaS primarly provide to users?

  Ready-made platforms for app development
  Virtualized computing resources
  Fully developed applications
  Infrastructure components

  Ans : Ready-made platforms for app development

- In which of the following service models can a cloud consumer only be able to manage authorization and entitlements ?

  IaaS
  PaaS
  SaaS
  Both 1 and 2
  All of the above

  Ans : SaaS

- What is each organization or user served by SaaS provider called?

  Tenant
  Subscriber
  Consumers
  Hosts

  Ans : Tenant

- A firm wants to implement a cloud computing service model having predefined business functionalities so that it can avoid building everything from scratch and directly start working on their projects

  Which among the following will you suggest

  IaaS
  SaaS
  PaaS
  DaaS

  Ans : SaaS

- A very small scale organization wants to switch to a cloud computing model for which only an up-front cost on a subscription basis need to be paid

  As a cloud engineer, which model would you suggest?

  Paas
  SaaS
  IaaS
  DaaS

  Ans : SaaS

- A new enterprise XYZ doesn't want to worry about managing or maintaining its application development platform. Additionally, XYZ wants the service provider to deliver and manage the backend infrastructure

  which of these cloud computing models would be most suitable for XYZ's use case described above?

  IaaS
  PaaS
  SaaS
  None of these

  Ans : PaaS

- Select the correct option regarding the statements given below.

  A) Using virtualization to create cloud systems results in a better security for public cloud users.
  B) The relationship between OS and the underlying hardware is not altered by virtualization

  Options :
  Statement A is true, Statement B is true , Statement A is the correct explanation of Statement A.
  Statement A is true, Statement B is true , Statement A is not the correct explanation of Statement A.
  Statement A is false, Statement B is true
  Statement A is false, Statement B is false

  Statement A is false, Statement B is false

- You have implemented cloud virtualization to save time and energy and to decreasing the cost. Which of the following adv can be achieved as a result of performing the same?

  A. Ability to manage resources effectively.
  B. Easier backup and disaster recovery.
  C. Increased efficiency of IT operations.

  Options :
  Only A and B
  Only B and C
  Only A and C
  All of the above

  Ans : All of the above

- Your organization wants to scale their cloud infrastructure to achieve the following capabilities:

  1. Take advantage of virtual servers, which are provisioned using predefined images.
  2. Use manual forecast or automated warning of system monitoring tooling which will then trigger operations to expand or reduce the cluster or farm of resources

  which scaling method can the organisation use in this context?

  Options :
  Manual scaling
  Semi-automated scaling
  Elastic scaling
  None of the above

  Ans : Elastic scaling

- An e-commerce company wants to switch from on-premises to cloud based model. It is observed that the servers have had unpredictable workload volumes and the following problems have been observed.

  1. Slow response times with on-premises applications
  2. Unpredictable spikes in demand for infrastructure services

  The company wants to switch to a cost-effective way to support their oerations. As a cloud expert, which of these cloud computing models would you recommend for given scenario?

  IaaS

PaaS
SaaS
None of these

Ans : Iaas

1. Read the given below statements about ADF and choose most appropriate choice.
   1. Data Factory contains a series of interconnected systems that provide a complete end-to-end platform for data engineers.
   2. Data Factory is a cloud-based ETL and data integration service that allows you to create data-driven workflows

   Options :
   Only 1 is NOT true
   Only 2 is NOT true
   Both 1 and 2 are NOT true
   Both 1 and 2 are true

   Ans : Both 1 and 2 are true

2. Which of the following tells Azure Synapse how many Spark resources to use when creating a serverless Apache Spark pool?

   The Spark pool
   The SQL pool
   The logical SQL server
   The dedicated SQL pool

   Ans : The Spark pool

3. Assume that you are copying data to and from data stores that are located on-premises when working with Azure Data Factory.
   Which of these integration runtimes should you setup in the given scenario?

   Azure integration runtime
   Self-hosted integration runtime
   Global integration runtime
   Any of the given choices

   Ans: Self-hosted integration runtime

4. Where does the provisioning take place when creating a new dedicated SQL pool in your Synapse workspace?

   On Apace Spark pool
   On the logical SQL server
   On the built-in serverless SQL pool
   None of these

   Ans: On the logical SQL server

5. Which one of these capabilities does Azure Integration runtime provide for a private network in Data Factory

1. Data Flow
2. Data Movement
3. Activity dispatch

Only 1 and 2
Only 2 and 3
Only 1 and 3
All 1,2 and 3

Ans : All 1,2 and 3

6. An Apache Spark job failed with the following error.
    Driver is temporarily unavailable
    Which of the following is the easiest way to resolve this issue in the absence of specific details?

    Increase the driver memory
    Avoid memory intensive operations
    Avoid running batch jobs on a shared interactive cluster
    Distribute the workloads into different clusters.

    Ans : Avoid running batch jobs on a shared interactive cluster

7. Which of these sharding pattern can you use to distribute the data when you define the table in Synapse SQL?
    1. Hash
    2. Round robin
    3. Replicate

    Only 1 and 2
    Only 2 and 3
    Only 1 and 3
    All 1,2 and 3

    Ans : All 1,2 and 3

8. You are attempting to rerun a Apache Spark write operation by cancelling the currently running job. When doing so the following error occurs.
    Error : org.apache.spark.sql.AnalysisException: Cannot create the manged table('testdb', 'testtable').
    The associated location ('dbfs:/user/hive/warehouse/testdb.db/metastore_cache_testtable) already exists; what could be the reason for this?

    The job is interrupted
    A temporary network issue occured
    The cluster is terminated during the execution of the write operation
    Either choice 1 or 2
    Either choice 1, 2 or 3

    Ans : Either choice 1,2 or 3

    1. Which of the following can be used as source dataset for a Power Query mash-up in Azure Data Factory?

An existing dataset
Create a new dataset
Select wrangling dataset
Either 1 or 2
Either 1 or 3

Ans : Either 1 or 2

2. Which one of the following can be used to move data from cloud source to a centralised data store in the cloud using Data Factory

   Copy Activity
   Azure Blob
   Data pipeline
   Azure Data lake

   Ans : Copy Activity

3. Which of the following represents a processing step in a Azure Data Factory Pipeline?

   Activities
   Datasets
   Linked services
   Dataflows

   Ans : Activities

4. You are performing data transformation activities using mapping data flows in Azure Data Factory How is the data flow canvas designed to make building transformation logic easy?

   It is separated into 3 parts: top bar, the graph, and the configuration panel
   It is separated into two parts : the graph, and the configuration panel.
   It is separated into 3 parts : the canvas panel, the graph , and the configuration panel.
   It is separated into two parts : the graph , and the author panel.

   Ans : It is separated into 3 parts: top bar, the graph, and the configuration panel

5. Which of the below given keys is Azure Data Factory NOT composed of
   1. Data Flows
   2. Integration Runtimes
   3. API
   4. Pipelines

   Only 1
   Only 2
   Only 3
   Only 4

   Ans : Only 3

6. Which of the following statements is true about datasets in Azure Data Factory?
   1. A data factory can have one or more pipelines.
   2. Datasets are used in source and sink in copy activity.
   Only 1
   Only 2
   Both 1 and 2
   Neither 1 nor 2

Ans : Both 1 and 2

7. Which of these Azure Data Factory control activities allows a pipeline to invoke another pipeline?
Append Pipeline
Until Activity
Set Pipeline
Execute Pipeline

Ans : Execute Pipeline

8. Which of the following roles can you assign to yourself when you prepare an existing storage account for use with Azure Synapse Analytics?
Owner role only
Storage Blob Data Owner role only
Both Owner role and Storage Blob Data Owner role
None of these

Ans : Both Owner role and Storage Blob Data Owner role

9. You are copying data between two data stores that are publicly accessible through the interest from any IP. Which of the following integration runtime can you use to perform this copy activity in Azure Data Factory?

Azure integration runtime
Self-hosted integration runtime
Sink integration runtime
Either 1 or 2
Either 2 or 3

Ans : Azure integration runtime

10. Which of the following is the fastest partitioning option for writing evenly distributed files in Azure Data Factory?

Set Partitioning
Use current partitioning
Single partition
Multiple partition

Ans : Use current partitioning

11. Which transformation activity in Azure Data Factory is used to split a dataset into multiple partitions?
Aggregate
Filter
Join
Partition

Ans : Partition

12. You have built-up a reusable library of data transformation routines and want to execute those processes in a scaled-out manner.
what can you do to achieve this?
Use ADF clusters
Use ADF Pipelines
Use ADF routines

Use ADF Copy activity

Ans : Use ADF Pipelines

13. While working in Azure Data Factory, you want to distribute data equally across several partitions in a transformation.
which of the following partitioning options can be used to achieve this?

Round Robin
Hash
Dynamic range
Fixed range
key

Ans : Round Robin

1. You have decides to use pie charts to visualize some data in Azure Databricks notebooks. Which of the given inferences is/are valid in the given scenario?
1. Pie charts can be used to show proportionality between metrics.
2. Pie charts can be used for conveying time series data.

Options :
Only 1
Only 2
Both 1 and 2
Neither 1 and 2

Ans : Only 1

2. While working on data preparation resources and information you are using the medallion lake house architecture.
Which of these choices is/are true regarding this?

The medallion lake house architecture guides you in data preparation by specifying a set of data layers of increasing quality.
The medallion lake house architecture is a framework for building reliable, maintainable, and testable data processing pipelines.
Both 1 and 2
Neither 1 not 2

Ans : Both 1 and 2

3. You are using the medallion lake house architecture to prepare your data in Azure Databricks.
How does this architecture provide efficient analysis?

It specifies a set of data layers of increasing quality
It guarantees to manage task orchestration, cluster management, monitoring, data quality, and error handling.
It maintains ACID guarantees as data passes through multiple layers of validations before being stored in a layout optimized.
All of these

Ans : All of these

4. You are designing a streaming pipeline to monitor and alert on anomalies in a continuous data stream. The pipeline must trigger alerts when anomalies are detected in real-time.
What component in Databricks can be effectively used to implement this functionality?

   Databricks Machine Learning for real-time anomaly detection
   Databricks Streaming Aggregations for alerting
   Databricks Delta for continuous data storage
   Databricks Structured Streaming for event windowing.

   Ans : Databricks Structured Streaming for event windowing.

5. You are trying to visualize some data using Azure Databricks notebooks, which of these visualization types can you use if the requirement is as follows?
   1. It should support backend aggregations
   2. It should provide support for queries returning more than 64k rows of data without truncation of the result set.

   Bar Charts
   Line Charts
   Pie Charts
   Both Choice 1 and 2
   All choices

   Ans : Both Choice 1 and 2

6. You are given two types of datasets as shown below.
   A : very large dataset containing thousands of records
   B : A dataset containing thousands of distinct fields.
   In which of the following dataset would you not be advised to use a Query filter while working with Azure Databricks

   Dataset A
   Dataset B
   Both A and B
   None of these

   Ans : Dataset B

7. You want to manipulate an Azure Databricks table visualization independently of the original cell results table.
   Which of these actions can you take when doing so?
   a) Reorder columns by dragging them up or down using Column Handle
   b) Hide columns by toggling the visibility icon
   c) Format columns using format settings

   Only a and b
   Only b and c
   All a, b and c
   None of these

   Ans : All a, b and c

8. You have decided to use a box chart visualization to show the distribution summary of some numerical data in Azure Databricks notebook.
   Which of the following can be achieved in the given scenario?
   1. You can quickly compare the value range across categories
   2. You can visualize the locality, spread and skewness groups of the values through their quartiles.

   Only 1

Only 2
Both 1 and 2
Neither 1 nor 2

Ans : Both 1 and 2

9. You want to create an interactive visualization for many groups of data points in Azure Databricks notebook.
   However, it is given that some groups of data points appear just once.
   Which of these charts be beset to use in the given scenario?

   Scatter Charts
   Line Charts
   Bar Charts
   Bubble Chart

   Ans : Scatter Charts

10. You have created a Azure Databricks workspace in portal.
    What happens when the workspace deployment fails?

    The workspace is still created in a failed state
    The workspace is deleted
    An error is thrown
    The workspace is created in a resource group

    Ans : The workspace is still created in a failed state

11. You want to add aggregations directly in your Azure Databricks visulalization rather than modifying the query to add an aggregation column. For which of these visualizations is aggregation available?
    1. Line chart
    2. Bar chart
    3. Area chart
    4. Pie chart

    Only 2, 3 and 4
    Only 2 and 4
    Only 4
    All 1, 2, 3 and 4

    Ans : All 1, 2, 3 and 4

12. Which of the following visualization types is enabled by default in Azure Databricks Notebook?

    Boxplot
    Chart
    Map
    Table
    Word cloud

    Ans : Table

13. You want to build reliable , maintainable and testable data processing pipelines when performing

data preparation tasks in Azure Databricks.
Which of these can be used to achieve this?

Delta live tables
Databricks partner connect
Databricks runtime
None of these

Ans : Delta live tables

14. Which of the following type of user can get access to special backup folder created in the workspace when a user is removed from Azure Databricks?

Databricks account owner
Databricks account admins
Admin users
Non-admin users

Ans : Databricks account admins

15. You are creating a dashboard in Azure Databricks notebook by adding a markdown cell to it. How are markdown cells useful in the given context?

Markdown cells are useful to publish visualizations
Markdown cells are useful to publish graphs
Markdown cells are useful as labels on your dashboard.
Markdown cells are useful when you want to share a graph as a presentation.

Ans : Markdown cells are useful as labels on your dashboard.

16. You are loading data into an Azure Databricks Lakehouse.
Which of these operations is best to use if you're going to ingest files in the order of thousands when doing so?

COPY.into
COPY INTO
Auto Loader
None of these

Ans : Auto Loader

17. Which of these interfaces does Azure Databricks support for accessing the assets of Azure Databricks SQL Analytics?

UI
API
SQL endpoints
Both 1 and 2
Both 2 and 3

Ans : Both 1 and 2

18. Which of the following should be created for running computations on a Azure Databricks Runtime cluster?

Cluster
Folder
Notebook
Table

Ans : Cluster