

Microsoft DP-200 Implementing an Azure Data Solution



QUESTION & ANSWERS

Microsoft

DP-200 Exam

Microsoft Implementing an Azure Data Solution Exam

**Questions & Answers
Demo**

Version: 14.0

Case Study: 1

Proseware Inc

Background

Proseware, Inc, develops and manages a product named Poll Taker. The product is used for delivering public opinion polling and analysis.

Polling data comes from a variety of sources, including online surveys, house-to-house interviews, and booths at public events.

Polling data

Polling data is stored in one of the two locations:

- An on-premises Microsoft SQL Server 2019 database named PollingData
- Azure Data Lake Gen 2

Data in Data Lake is queried by using PolyBase

Poll metadata

Each poll has associated metadata with information about the poll including the date and number of respondents. The data is stored as JSON.

Phone-based polling

Security

- Phone-based poll data must only be uploaded by authorized users from authorized devices
- Contractors must not have access to any polling data other than their own
- Access to polling data must set on a per-active directory user basis

Data migration and loading

- All data migration processes must use Azure Data Factory
- All data migrations must run automatically during non-business hours
- Data migrations must be reliable and retry when needed

Performance

After six months, raw polling data should be moved to a lower-cost storage solution.

Deployments

- All deployments must be performed by using Azure DevOps. Deployments must use templates used in multiple environments
- No credentials or secrets should be used during deployments

Reliability

All services and processes must be resilient to a regional Azure outage.

Monitoring

All Azure services must be monitored by using Azure Monitor. On-premises SQL Server performance must be monitored.

Question: 1

HOTSPOT

You need to ensure that Azure Data Factory pipelines can be deployed. How should you configure authentication and authorization for deployments? To answer, select the appropriate options in the answer choices.

NOTE: Each correct selection is worth one point.

Security requirement

Technology

Authorization

RBAC	v
DAC	
MAC	
Claims	

Authentication

Service Principal	^
Kerberos	
Certificate-based	
Bearer Token	

Answer:

Security requirement

Technology

Authorization

RBAC	v
DAC	
MAC	
Claims	

Authentication

Service Principal	^
Kerberos	
Certificate-based	
Bearer Token	

Explanation:

The way you control access to resources using RBAC is to create role assignments. This is a key concept to understand – it’s how permissions are enforced. A role assignment consists of three elements: security principal, role definition, and scope.

Scenario:

No credentials or secrets should be used during deployments

Phone-based poll data must only be uploaded by authorized users from authorized devices

Contractors must not have access to any polling data other than their own

Access to polling data must set on a per-active directory user basis

References:

<https://docs.microsoft.com/en-us/azure/role-based-access-control/overview>

Question: 2

DRAG DROP

You need to provision the polling data storage account.

How should you configure the storage account? To answer, drag the appropriate Configuration Value to the correct Setting. Each Configuration Value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Configuration values

Answer Area

LRS

Setting

Configuration value

GRS

Account type

RA-GRS

Replication type

Storage

StorageV2

Answer:

Configuration values

Answer Area

LRS

Setting

Configuration value

GRS

Account type

Replication type

Storage

Explanation:

Account type: StorageV2

You must create new storage accounts as type StorageV2 (general-purpose V2) to take advantage of Data Lake Storage Gen2 features.

Scenario: Polling data is stored in one of the two locations:

- An on-premises Microsoft SQL Server 2019 database named PollingData
- Azure Data Lake Gen 2

Data in Data Lake is queried by using PolyBase

Replication type: RA-GRS

Scenario: All services and processes must be resilient to a regional Azure outage.

Geo-redundant storage (GRS) is designed to provide at least 99.99999999999999% (16 9's) durability of objects over a given year by replicating your data to a secondary region that is hundreds of miles away from the primary region. If your storage account has GRS enabled, then your data is durable even in the case of a complete regional outage or a disaster in which the primary region isn't recoverable.

If you opt for GRS, you have two related options to choose from:

- GRS replicates your data to another data center in a secondary region, but that data is available to be read only if Microsoft initiates a failover from the primary to secondary region.
- Read-access geo-redundant storage (RA-GRS) is based on GRS. RA-GRS replicates your data to another data center in a secondary region, and also provides you with the option to read from the secondary region. With RA-GRS, you can read from the secondary region regardless of whether Microsoft initiates a failover from the primary to secondary region.

References:

<https://docs.microsoft.com/bs-cyrl-ba/azure/storage/blobs/data-lake-storage-quickstart-create-account>

<https://docs.microsoft.com/en-us/azure/storage/common/storage-redundancy-grs>

Question: 3

HOTSPOT

You need to ensure polling data security requirements are met.

Which security technologies should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Context	Security technology				
SQL Server	<table border="1"> <tr> <td data-bbox="470 1097 965 1153">Azure Active Directory user</td> <td data-bbox="965 1097 1021 1254" rowspan="3" style="text-align: center; vertical-align: middle;">▼</td> </tr> <tr> <td data-bbox="470 1153 965 1198">Domain Active Directory user</td> </tr> <tr> <td data-bbox="470 1198 965 1254">Managed Identity</td> </tr> </table>	Azure Active Directory user	▼	Domain Active Directory user	Managed Identity
Azure Active Directory user	▼				
Domain Active Directory user					
Managed Identity					
PolyBase	<table border="1"> <tr> <td data-bbox="470 1288 965 1332">Database scoped credential</td> <td data-bbox="965 1288 1021 1444" rowspan="3" style="text-align: center; vertical-align: middle;">▼</td> </tr> <tr> <td data-bbox="470 1332 965 1377">Database encryption key</td> </tr> <tr> <td data-bbox="470 1377 965 1444">Application role</td> </tr> </table>	Database scoped credential	▼	Database encryption key	Application role
Database scoped credential	▼				
Database encryption key					
Application role					

Answer:

Context	Security technology				
SQL Server	<table border="1"> <tr> <td>Azure Active Directory user</td> <td rowspan="3">v</td> </tr> <tr> <td>Domain Active Directory user</td> </tr> <tr> <td>Managed Identity</td> </tr> </table>	Azure Active Directory user	v	Domain Active Directory user	Managed Identity
Azure Active Directory user	v				
Domain Active Directory user					
Managed Identity					
PolyBase	<table border="1"> <tr> <td>Database scoped credential</td> <td rowspan="3">v</td> </tr> <tr> <td>Database encryption key</td> </tr> <tr> <td>Application role</td> </tr> </table>	Database scoped credential	v	Database encryption key	Application role
Database scoped credential	v				
Database encryption key					
Application role					

Explanation:

Box 1: Azure Active Directory user

Scenario:

Access to polling data must set on a per-active directory user basis

Box 2: DataBase Scoped Credential

SQL Server uses a database scoped credential to access non-public Azure blob storage or Kerberos-secured Hadoop clusters with PolyBase.

PolyBase cannot authenticate by using Azure AD authentication.

References:

<https://docs.microsoft.com/en-us/sql/t-sql/statements/create-database-scoped-credential-transact-sql>

Question: 4

You need to ensure that phone-based poling data can be analyzed in the PollingData database. How should you configure Azure Data Factory?

- A. Use a tumbling schedule trigger
- B. Use an event-based trigger
- C. Use a schedule trigger
- D. Use manual execution

Answer: C

Explanation:

When creating a schedule trigger, you specify a schedule (start date, recurrence, end date etc.) for the trigger, and associate with a Data Factory pipeline.

Scenario:

All data migration processes must use Azure Data Factory

All data migrations must run automatically during non-business hours

References:

<https://docs.microsoft.com/en-us/azure/data-factory/how-to-create-schedule-trigger>

Question: 5

You need to process and query ingested Tier 9 data.
Which two options should you use? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Azure Notification Hub
- B. Transact-SQL statements
- C. Azure Cache for Redis
- D. Apache Kafka statements
- E. Azure Event Grid
- F. Azure Stream Analytics

Answer: EF

Explanation:

Event Hubs provides a Kafka endpoint that can be used by your existing Kafka based applications as an alternative to running your own Kafka cluster.

You can stream data into Kafka-enabled Event Hubs and process it with Azure Stream Analytics, in the following steps:

Create a Kafka enabled Event Hubs namespace.

Create a Kafka client that sends messages to the event hub.

Create a Stream Analytics job that copies data from the event hub into an Azure blob storage.

Scenario:

Internal Distribution and Sales	9	Yes, once ingested at branches	Data ingested from Contoso branches
---------------------------------	---	--------------------------------	-------------------------------------

Tier 9 reporting must be moved to Event Hubs, queried, and persisted in the same Azure region as the company's main office

References:

<https://docs.microsoft.com/en-us/azure/event-hubs/event-hubs-kafka-stream-analytics>

Question: 6

You need to set up Azure Data Factory pipelines to meet data movement requirements.
Which integration runtime should you use?

- A. self-hosted integration runtime
- B. Azure-SSIS Integration Runtime
- C. .NET Common Language Runtime (CLR)
- D. Azure integration runtime

Answer: A

Explanation:

The following table describes the capabilities and network support for each of the integration runtime types:

IR type	Public network	Private network
Azure	Data movement Activity dispatch	
Self-hosted	Data movement Activity dispatch	Data movement Activity dispatch
Azure-SSIS	SSIS package execution	SSIS package execution

Scenario: The solution must support migrating databases that support external and internal application to Azure SQL Database. The migrated databases will be supported by Azure Data Factory pipelines for the continued movement, migration and updating of data both in the cloud and from local core business systems and repositories.

References:

<https://docs.microsoft.com/en-us/azure/data-factory/concepts-integration-runtime>

Question: 7

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some questions sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You need to implement diagnostic logging for Data Warehouse monitoring.

Which log should you use?

- A. RequestSteps
- B. DmsWorkers
- C. SqlRequests
- D. ExecRequests

Answer: C

Explanation:

Scenario:

The Azure SQL Data Warehouse cache must be monitored when the database is being used.

Metric	Description
A	Low cache hit %, high cache usage %
B	Low cache hit %, low cache usage %
C	High cache hit %, high cache usage %

References:

<https://docs.microsoft.com/en-us/sql/relational-databases/system-dynamic-management-views/sys-dm-pdw-sql-requests-transact-sq>

Question: 8

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some questions sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You need setup monitoring for tiers 6 through 8.

What should you configure?

- A. extended events for average storage percentage that emails data engineers
- B. an alert rule to monitor CPU percentage in databases that emails data engineers
- C. an alert rule to monitor CPU percentage in elastic pools that emails data engineers
- D. an alert rule to monitor storage percentage in databases that emails data engineers
- E. an alert rule to monitor storage percentage in elastic pools that emails data engineers

Answer: E

Explanation:

Scenario:

Tiers 6 through 8 must have unexpected resource storage usage immediately reported to data engineers. Tier 3 and Tier 6 through Tier 8 applications must use database density on the same server and Elastic pools in a cost-effective manner.

Question: 9

On which data store you configure TDE to meet the technical requirements?

- A. Cosmos DB
- B. SQL Data Warehouse
- C. SQL Database

Answer: B

Explanation:

Scenario: Transparent data encryption (TDE) must be enabled on all data stores, whenever possible.

The datacentre for Mechanical Workflow must be moved to Azure SQL data Warehouse.

Question: 10

HOTSPOT

You are building the data store solution for Mechanical Workflow.

How should you configure Table1? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Table Type:

	▼
Hash distributed	
Replicated	
Round-robin	

Index type:

	▼
Clustered	
Clustered columnstore	
Heap	
Nonclustered	

Answer:

Table Type:

	▼
Hash distributed	
Replicated	
Round-robin	

Index type:

	▼
Clustered	
Clustered columnstore	
Heap	
Nonclustered	

Explanation:

Table Type: Hash distributed.

Hash-distributed tables improve query performance on large fact tables.

Index type: Clusted columnstore

Scenario:

Mechanical Workflow has a named Table1 that is 1 TB. Large aggregations are performed on a single column of Table 1.

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-distribute>

