



Microsoft

Exam Questions DP-200

Implementing an Azure Data Solution

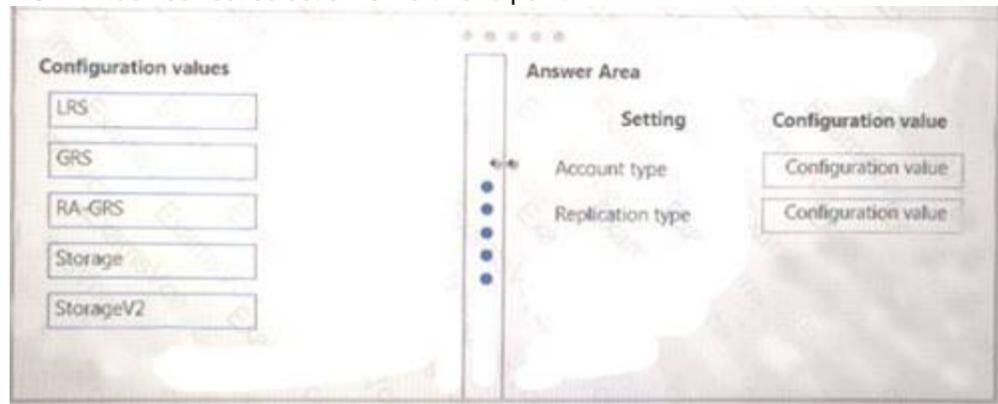
NEW QUESTION 1

- (Exam Topic 1)

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.

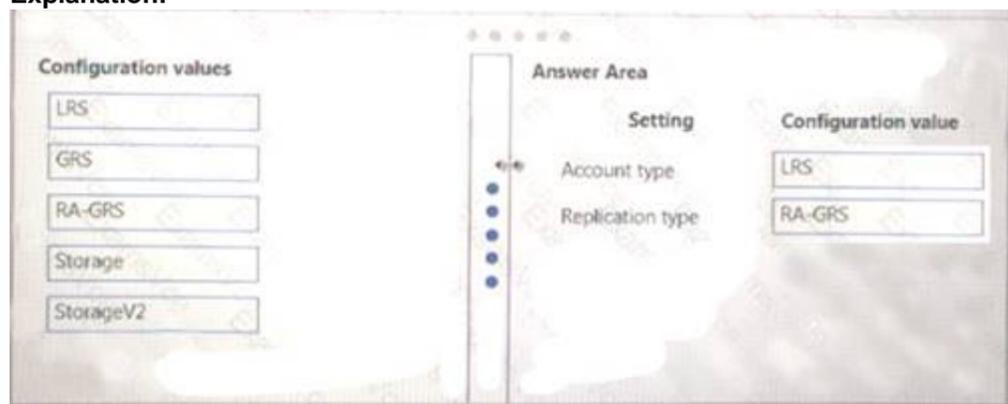
NOTE: Each correct selection is worth one point.



- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

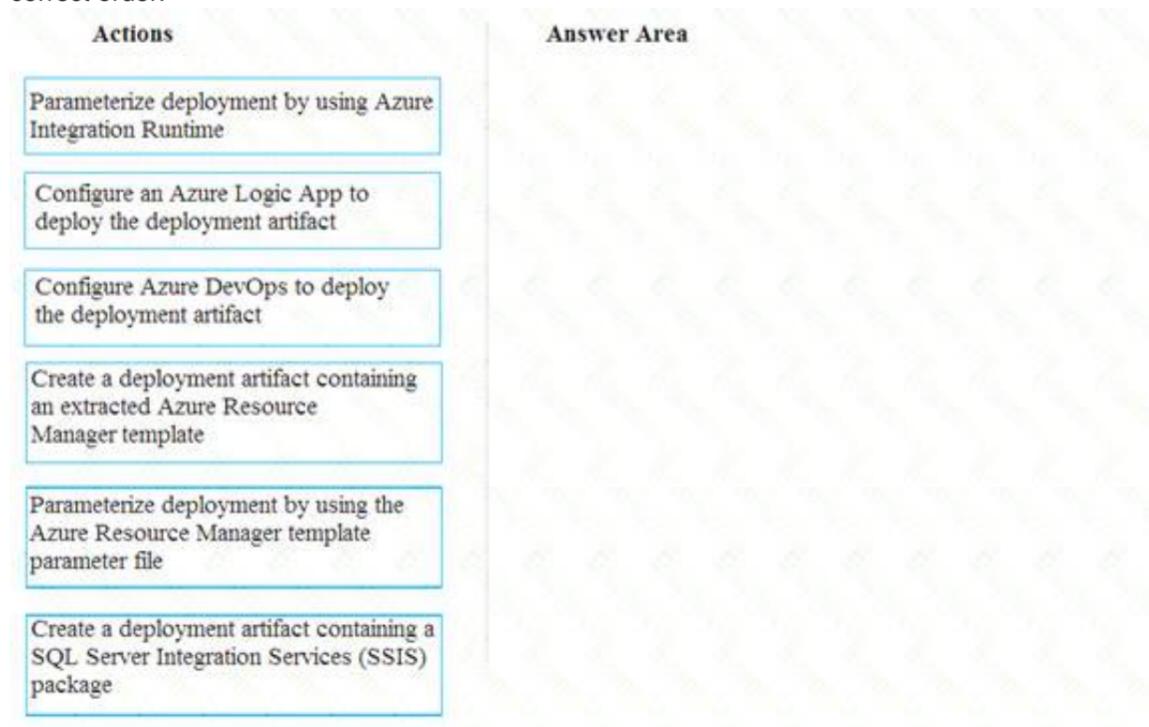


NEW QUESTION 2

- (Exam Topic 1)

You need to ensure that phone-based polling data can be analyzed in the PollingData database.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer are and arrange them in the correct order.



- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Create a deployment artifact containing an extracted Azure Resource Manager template

Parameterize deployment by using the Azure Resource Manager template parameter file

Configure Azure DevOps to deploy the deployment artifact

Scenario:

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

NEW QUESTION 3

- (Exam Topic 1)

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	<table border="1"> <tr><td>RBAC</td><td><input checked="" type="checkbox"/></td></tr> <tr><td>DAC</td><td><input type="checkbox"/></td></tr> <tr><td>MAC</td><td><input type="checkbox"/></td></tr> <tr><td>Claims</td><td><input type="checkbox"/></td></tr> </table>	RBAC	<input checked="" type="checkbox"/>	DAC	<input type="checkbox"/>	MAC	<input type="checkbox"/>	Claims	<input type="checkbox"/>
RBAC	<input checked="" type="checkbox"/>								
DAC	<input type="checkbox"/>								
MAC	<input type="checkbox"/>								
Claims	<input type="checkbox"/>								
Authentication	<table border="1"> <tr><td>Service Principal</td><td><input checked="" type="checkbox"/></td></tr> <tr><td>Kerberos</td><td><input type="checkbox"/></td></tr> <tr><td>Certificate-based</td><td><input type="checkbox"/></td></tr> <tr><td>Bearer Token</td><td><input type="checkbox"/></td></tr> </table>	Service Principal	<input checked="" type="checkbox"/>	Kerberos	<input type="checkbox"/>	Certificate-based	<input type="checkbox"/>	Bearer Token	<input type="checkbox"/>
Service Principal	<input checked="" type="checkbox"/>								
Kerberos	<input type="checkbox"/>								
Certificate-based	<input type="checkbox"/>								
Bearer Token	<input type="checkbox"/>								

- A. Mastered
- B. Not Mastered

Answer: A

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>

NEW QUESTION 4

- (Exam Topic 1)

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>Azure Active Directory user</td><td><input checked="" type="checkbox"/></td></tr> <tr><td>Domain Active Directory user</td><td><input type="checkbox"/></td></tr> <tr><td>Managed Identity</td><td><input type="checkbox"/></td></tr> </table>	Azure Active Directory user	<input checked="" type="checkbox"/>	Domain Active Directory user	<input type="checkbox"/>	Managed Identity	<input type="checkbox"/>
Azure Active Directory user	<input checked="" type="checkbox"/>						
Domain Active Directory user	<input type="checkbox"/>						
Managed Identity	<input type="checkbox"/>						
PolyBase	<table border="1"> <tr><td>Database scoped credential</td><td><input checked="" type="checkbox"/></td></tr> <tr><td>Database encryption key</td><td><input type="checkbox"/></td></tr> <tr><td>Application role</td><td><input type="checkbox"/></td></tr> </table>	Database scoped credential	<input checked="" type="checkbox"/>	Database encryption key	<input type="checkbox"/>	Application role	<input type="checkbox"/>
Database scoped credential	<input checked="" type="checkbox"/>						
Database encryption key	<input type="checkbox"/>						
Application role	<input type="checkbox"/>						

- A. Mastered
- B. Not Mastered

Answer: A

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>

NEW QUESTION 5

- (Exam Topic 2)

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-r>

NEW QUESTION 6

- (Exam Topic 2)

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>

NEW QUESTION 7

- (Exam Topic 2)

You need set up the Azure Data Factory JSON definition for Tier 10 data.

What should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Data factory component	Value								
Connector	<table border="1"> <tbody> <tr><td>connection string</td><td><input type="checkbox"/></td></tr> <tr><td>linked service name string</td><td><input type="checkbox"/></td></tr> <tr><td>gateway connection string</td><td><input type="checkbox"/></td></tr> <tr><td>data store name string</td><td><input type="checkbox"/></td></tr> </tbody> </table>	connection string	<input type="checkbox"/>	linked service name string	<input type="checkbox"/>	gateway connection string	<input type="checkbox"/>	data store name string	<input type="checkbox"/>
connection string	<input type="checkbox"/>								
linked service name string	<input type="checkbox"/>								
gateway connection string	<input type="checkbox"/>								
data store name string	<input type="checkbox"/>								
Data movement activity	<table border="1"> <tbody> <tr><td>Azure SQL Data Warehouse</td><td><input type="checkbox"/></td></tr> <tr><td>Azure Files</td><td><input type="checkbox"/></td></tr> <tr><td>Azure Blob</td><td><input type="checkbox"/></td></tr> <tr><td>Azure SQL Database</td><td><input type="checkbox"/></td></tr> </tbody> </table>	Azure SQL Data Warehouse	<input type="checkbox"/>	Azure Files	<input type="checkbox"/>	Azure Blob	<input type="checkbox"/>	Azure SQL Database	<input type="checkbox"/>
Azure SQL Data Warehouse	<input type="checkbox"/>								
Azure Files	<input type="checkbox"/>								
Azure Blob	<input type="checkbox"/>								
Azure SQL Database	<input type="checkbox"/>								

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Connection String

To use storage account key authentication, you use the ConnectionString property, which specifies the information needed to connect to Blob Storage.

Mark this field as a SecureString to store it securely in Data Factory. You can also put account key in Azure Key Vault and pull the accountKey configuration out of the connection string.

Box 2: Azure Blob

Tier 10 reporting data must be stored in Azure Blobs

External Distribution and Sales	10	Yes, once ingested at Contoso main office	Data is ingested from multiple sources
---------------------------------	----	---	--

References:

<https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-blob-storage>

NEW QUESTION 8

- (Exam Topic 2)

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 configure data encryption for external applications.

Solution:

1. Access the Always Encrypted Wizard in SQL Server Management Studio
2. Select the column to be encrypted
3. Set the encryption type to Deterministic
4. Configure the master key to use the Windows Certificate Store
5. Validate configuration results and deploy the solution Does the solution meet the goal?

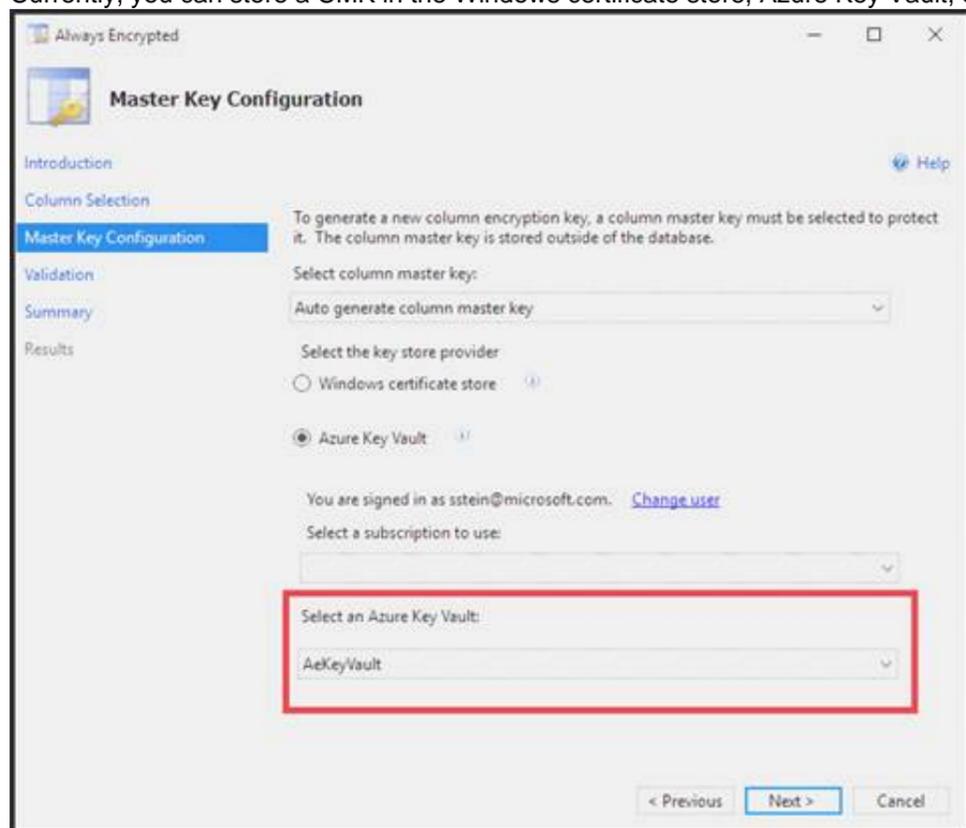
- A. Yes
- B. No

Answer: B

Explanation:

Use the Azure Key Vault, not the Windows Certificate Store, to store the master key.

Note: The Master Key Configuration page is where you set up your CMK (Column Master Key) and select the key store provider where the CMK will be stored. Currently, you can store a CMK in the Windows certificate store, Azure Key Vault, or a hardware security module (HSM).



References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-always-encrypted-azure-key-vault>

NEW QUESTION 9

- (Exam Topic 2)

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>

NEW QUESTION 10

- (Exam Topic 2)

You need to mask tier 1 data. Which functions should you use? To answer, select the appropriate option in the answer area.

NOTE: Each correct selection is worth one point.

Data type	Masking function								
A	<table border="1"> <tr><td>custom text</td><td><input type="checkbox"/></td></tr> <tr><td>default</td><td><input type="checkbox"/></td></tr> <tr><td>email</td><td><input type="checkbox"/></td></tr> <tr><td>random number</td><td><input type="checkbox"/></td></tr> </table>	custom text	<input type="checkbox"/>	default	<input type="checkbox"/>	email	<input type="checkbox"/>	random number	<input type="checkbox"/>
custom text	<input type="checkbox"/>								
default	<input type="checkbox"/>								
email	<input type="checkbox"/>								
random number	<input type="checkbox"/>								
B	<table border="1"> <tr><td>custom text</td><td><input type="checkbox"/></td></tr> <tr><td>default</td><td><input type="checkbox"/></td></tr> <tr><td>email</td><td><input type="checkbox"/></td></tr> <tr><td>random number</td><td><input type="checkbox"/></td></tr> </table>	custom text	<input type="checkbox"/>	default	<input type="checkbox"/>	email	<input type="checkbox"/>	random number	<input type="checkbox"/>
custom text	<input type="checkbox"/>								
default	<input type="checkbox"/>								
email	<input type="checkbox"/>								
random number	<input type="checkbox"/>								
C	<table border="1"> <tr><td>custom text</td><td><input type="checkbox"/></td></tr> <tr><td>default</td><td><input type="checkbox"/></td></tr> <tr><td>email</td><td><input type="checkbox"/></td></tr> <tr><td>random number</td><td><input type="checkbox"/></td></tr> </table>	custom text	<input type="checkbox"/>	default	<input type="checkbox"/>	email	<input type="checkbox"/>	random number	<input type="checkbox"/>
custom text	<input type="checkbox"/>								
default	<input type="checkbox"/>								
email	<input type="checkbox"/>								
random number	<input type="checkbox"/>								

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

A: Default

Full masking according to the data types of the designated fields.

For string data types, use XXXX or fewer Xs if the size of the field is less than 4 characters (char, nchar, varchar, nvarchar, text, ntext).

B: email

C: Custom text

Custom StringMasking method which exposes the first and last letters and adds a custom padding string in the middle. prefix,[padding],suffix

Tier 1 Database must implement data masking using the following masking logic:

Data type	Masking requirement
A	Mask 4 or less string data type characters
B	Mask first letter and domain
C	Mask everything except characters at the beginning and end

References:

<https://docs.microsoft.com/en-us/sql/relational-databases/security/dynamic-data-masking>

NEW QUESTION 10

- (Exam Topic 3)

Your company manages on-premises Microsoft SQL Server pipelines by using a custom solution.

The data engineering team must implement a process to pull data from SQL Server and migrate it to Azure Blob storage. The process must orchestrate and manage the data lifecycle.

You need to configure Azure Data Factory to connect to the on-premises SQL Server database.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions	Answer Area
Create an Azure Data Factory resource.	
Configure a self-hosted integration runtime.	
Create a virtual private network (VPN) connection from on-premises to Microsoft Azure.	
Create a database master key on SQL Server.	
Backup the database and send it Azure Blob storage.	
Configure the on-premises SQL Server instance with an integration runtime.	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: Create a virtual private network (VPN) connection from on-premises to Microsoft Azure.

You can also use IPSec VPN or Azure ExpressRoute to further secure the communication channel between your on-premises network and Azure.

Azure Virtual Network is a logical representation of your network in the cloud. You can connect an on-premises network to your virtual network by setting up IPSec VPN (site-to-site) or ExpressRoute (private peering).

Step 2: Create an Azure Data Factory resource. Step 3: Configure a self-hosted integration runtime.

You create a self-hosted integration runtime and associate it with an on-premises machine with the SQL Server database. The self-hosted integration runtime is the component that copies data from the SQL Server database on your machine to Azure Blob storage.

Note: A self-hosted integration runtime can run copy activities between a cloud data store and a data store in a private network, and it can dispatch transform activities against compute resources in an on-premises network or an Azure virtual network. The installation of a self-hosted integration runtime needs on an on-premises machine or a virtual machine (VM) inside a private network.

References:

<https://docs.microsoft.com/en-us/azure/data-factory/tutorial-hybrid-copy-powershell>

NEW QUESTION 14

- (Exam Topic 3)

You develop data engineering solutions for a company. The company has on-premises Microsoft SQL Server databases at multiple locations.

The company must integrate data with Microsoft Power BI and Microsoft Azure Logic Apps. The solution must avoid single points of failure during connection and transfer to the cloud. The solution must also minimize latency.

You need to secure the transfer of data between on-premises databases and Microsoft Azure.

What should you do?

- A. Install a standalone on-premises Azure data gateway at each location
- B. Install an on-premises data gateway in personal mode at each location
- C. Install an Azure on-premises data gateway at the primary location
- D. Install an Azure on-premises data gateway as a cluster at each location

Answer: D

Explanation:

You can create high availability clusters of On-premises data gateway installations, to ensure your organization can access on-premises data resources used in Power BI reports and dashboards. Such clusters allow gateway administrators to group gateways to avoid single points of failure in accessing on-premises data resources. The Power BI service always uses the primary gateway in the cluster, unless it's not available. In that case, the service switches to the next gateway in the cluster, and so on.

References:

<https://docs.microsoft.com/en-us/power-bi/service-gateway-high-availability-clusters>

NEW QUESTION 17

- (Exam Topic 3)

A company uses Microsoft Azure SQL Database to store sensitive company data. You encrypt the data and only allow access to specified users from specified locations.

You must monitor data usage, and data copied from the system to prevent data leakage.

You need to configure Azure SQL Database to email a specific user when data leakage occurs.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions	Answer Area
In Auditing, enable Auditing .	
Configure the service to create alerts for threat detections of type Data Exfiltration .	
In Firewalls and virtual networks, enable Allow access to Azure services .	
Enable advanced threat protection.	
Configure the service to send email alerts to security@contoso.com	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Actions	Answer Area
In Auditing, enable Auditing .	Enable advanced threat protection.
Configure the service to create alerts for threat detections of type Data Exfiltration .	Configure the service to send email alerts to security@contoso.com
In Firewalls and virtual networks, enable Allow access to Azure services .	Configure the service to create alerts for threat detections of type Data Exfiltration .
Enable advanced threat protection.	
Configure the service to send email alerts to security@contoso.com	

NEW QUESTION 19

- (Exam Topic 3)

You manage security for a database that supports a line of business application. Private and personal data stored in the database must be protected and encrypted. You need to configure the database to use Transparent Data Encryption (TDE).

Which five actions should you perform in sequence? To answer, select the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions	Answer Area
Create a database encryption key using a certificate generated with the master key.	
Create a certificate and then create the master key using a password.	
Set the context to the master database.	
Create a master key using a password.	
Set the context to the company database.	
Enable encryption.	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: Create a master key

Step 2: Create or obtain a certificate protected by the master key Step 3: Set the context to the company database

Step 4: Create a database encryption key and protect it by the certificate Step 5: Set the database to use encryption

Example code: USE master; GO

CREATE MASTER KEY ENCRYPTION BY PASSWORD = '<UseStrongPasswordHere>';

go

```
CREATE CERTIFICATE MyServerCert WITH SUBJECT = 'My DEK Certificate'; go
USE AdventureWorks2012; GO
CREATE DATABASE ENCRYPTION KEY WITH ALGORITHM = AES_128
ENCRYPTION BY SERVER CERTIFICATE MyServerCert; GO
ALTER DATABASE AdventureWorks2012 SET ENCRYPTION ON;
GO
```

References:

<https://docs.microsoft.com/en-us/sql/relational-databases/security/encryption/transparent-data-encryption>

NEW QUESTION 20

- (Exam Topic 3)

You are developing a solution to visualize multiple terabytes of geospatial data. The solution has the following requirements:

- Data must be encrypted.
- Data must be accessible by multiple resources on Microsoft Azure. You need to provision storage for the solution.

Which four actions should you perform in sequence? To answer, move the appropriate action from the list of actions to the answer area and arrange them in the correct order.

The screenshot shows an exam question interface with two main sections: 'Actions' and 'Answer Area'. The 'Actions' list contains six items:

- Enable encryption on the Azure Data Lake using the Azure portal.
- Add an access policy for the new Azure Data Lake account to the key storage container.
- Create a new Azure Data Lake Storage account with Azure Data Lake managed encryption keys.
- Select and configure an encryption key storage container.
- Create a new Azure Data Lake Storage account with Azure Key Vault managed encryption keys.
- Create a new Azure Data Lake Storage account with encryption disabled.

The 'Answer Area' is currently empty. Navigation arrows are visible between the two sections.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

The screenshot shows the same exam question interface as above, but with the correct sequence of four actions highlighted in the 'Answer Area' and moved from the 'Actions' list:

- Select and configure an encryption key storage container.
- Add an access policy for the new Azure Data Lake account to the key storage container.
- Create a new Azure Data Lake Storage account with Azure Data Lake managed encryption keys.
- Create a new Azure Data Lake Storage account with encryption disabled.

The remaining two actions in the 'Actions' list are 'Enable encryption on the Azure Data Lake using the Azure portal.' and 'Create a new Azure Data Lake Storage account with Azure Key Vault managed encryption keys.'

NEW QUESTION 23

- (Exam Topic 3)

Your company has on-premises Microsoft SQL Server instance.

The data engineering team plans to implement a process that copies data from the SQL Server instance to Azure Blob storage. The process must orchestrate and manage the data lifecycle.

You need to configure Azure Data Factory to connect to the SQL Server instance.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

NEW QUESTION 24

- (Exam Topic 3)

You implement an event processing solution using Microsoft Azure Stream Analytics. The solution must meet the following requirements:

- Ingest data from Blob storage
- Analyze data in real time
- Store processed data in Azure Cosmos DB

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

NEW QUESTION 29

- (Exam Topic 3)

You develop data engineering solutions for a company.

You need to deploy a Microsoft Azure Stream Analytics job for an IoT solution. The solution must:

- Minimize latency.
- Minimize bandwidth usage between the job and IoT device.

Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

NEW QUESTION 32

- (Exam Topic 3)

A company plans to develop solutions to perform batch processing of multiple sets of geospatial data. You need to implement the solutions. Which Azure services should you use? To answer, select the appropriate configuration in the answer area. NOTE: Each correct selection is worth one point.

Scenario	Tool
Use a native client application to run interactive queries and batch processes.	<ul style="list-style-type: none"> HDInsight Tools for Visual Studio Hive View HDInsight REST API Azure Data Factory
Use a web browser to run interactive queries and batch processes.	<ul style="list-style-type: none"> HDInsight Tools for Visual Studio Hive View HDInsight REST API Azure PowerShell
Develop batch processing applications that use Azure HDInsight.	<ul style="list-style-type: none"> HDInsight Tools for Visual Studio Hive View HDInsight REST API NoSQL database

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Scenario	Tool
Use a native client application to run interactive queries and batch processes.	<ul style="list-style-type: none"> HDInsight Tools for Visual Studio Hive View HDInsight REST API Azure Data Factory
Use a web browser to run interactive queries and batch processes.	<ul style="list-style-type: none"> HDInsight Tools for Visual Studio Hive View HDInsight REST API Azure PowerShell
Develop batch processing applications that use Azure HDInsight.	<ul style="list-style-type: none"> HDInsight Tools for Visual Studio Hive View HDInsight REST API NoSQL database

NEW QUESTION 33

- (Exam Topic 3)

A company has a real-time data analysis solution that is hosted on Microsoft Azure the solution uses Azure Event Hub to ingest data and an Azure Stream Analytics cloud job to analyze the data. The cloud job is configured to use 120 Streaming Units (SU).

You need to optimize performance for the Azure Stream Analytics job.

Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Implement event ordering
- B. Scale the SU count for the job up
- C. Implement Azure Stream Analytics user-defined functions (UDF)
- D. Scale the SU count for the job down
- E. Implement query parallelization by partitioning the data output
- F. Implement query parallelization by partitioning the data input

Answer: BF

Explanation:

Scale out the query by allowing the system to process each input partition separately.

F: A Stream Analytics job definition includes inputs, a query, and output. Inputs are where the job reads the data stream from.

References:

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-parallelization>

NEW QUESTION 38

- (Exam Topic 3)

A company uses Azure SQL Database to store sales transaction data. Field sales employees need an offline copy of the database that includes last year's sales on their laptops when there is no internet connection available.

You need to create the offline export copy.

Which three options can you use? Each correct answer presents a complete solution.
 NOTE: Each correct selection is worth one point.

- A. Export to a BACPAC file by using Azure Cloud Shell, and save the file to an Azure storage account
- B. Export to a BACPAC file by using SQL Server Management Studi
- C. Save the file to an Azure storage account
- D. Export to a BACPAC file by using the Azure portal
- E. Export to a BACPAC file by using Azure PowerShell and save the file locally
- F. Export to a BACPAC file by using the SqlPackage utility

Answer: BCE

NEW QUESTION 39

- (Exam Topic 3)

A company is deploying a service-based data environment. You are developing a solution to process this data. The solution must meet the following requirements:

- Use an Azure HDInsight cluster for data ingestion from a relational database in a different cloud service
- Use an Azure Data Lake Storage account to store processed data
- Allow users to download processed data

You need to recommend technologies for the solution.

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

Data process	Technology								
Ingest	<table border="1"> <tr><td>RevoScaleR</td><td><input type="checkbox"/></td></tr> <tr><td>Apache Sqoop</td><td><input type="checkbox"/></td></tr> <tr><td>Apache DistCp</td><td><input type="checkbox"/></td></tr> <tr><td>Azure CLI</td><td><input type="checkbox"/></td></tr> </table>	RevoScaleR	<input type="checkbox"/>	Apache Sqoop	<input type="checkbox"/>	Apache DistCp	<input type="checkbox"/>	Azure CLI	<input type="checkbox"/>
RevoScaleR	<input type="checkbox"/>								
Apache Sqoop	<input type="checkbox"/>								
Apache DistCp	<input type="checkbox"/>								
Azure CLI	<input type="checkbox"/>								
Process	<table border="1"> <tr><td>Apache DistCp</td><td><input type="checkbox"/></td></tr> <tr><td>Apache Kafka</td><td><input type="checkbox"/></td></tr> <tr><td>C#</td><td><input type="checkbox"/></td></tr> <tr><td>Apache Hive</td><td><input type="checkbox"/></td></tr> </table>	Apache DistCp	<input type="checkbox"/>	Apache Kafka	<input type="checkbox"/>	C#	<input type="checkbox"/>	Apache Hive	<input type="checkbox"/>
Apache DistCp	<input type="checkbox"/>								
Apache Kafka	<input type="checkbox"/>								
C#	<input type="checkbox"/>								
Apache Hive	<input type="checkbox"/>								
Download	<table border="1"> <tr><td>Apache Sqoop</td><td><input type="checkbox"/></td></tr> <tr><td>MapReduce</td><td><input type="checkbox"/></td></tr> <tr><td>RevoScaleR</td><td><input type="checkbox"/></td></tr> <tr><td>Ambari Hive View</td><td><input type="checkbox"/></td></tr> </table>	Apache Sqoop	<input type="checkbox"/>	MapReduce	<input type="checkbox"/>	RevoScaleR	<input type="checkbox"/>	Ambari Hive View	<input type="checkbox"/>
Apache Sqoop	<input type="checkbox"/>								
MapReduce	<input type="checkbox"/>								
RevoScaleR	<input type="checkbox"/>								
Ambari Hive View	<input type="checkbox"/>								

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Apache Sqoop is a tool designed for efficiently transferring bulk data between Apache Hadoop and structured datastores such as relational databases. Azure HDInsight is a cloud distribution of the Hadoop components from the Hortonworks Data Platform (HDP).

NEW QUESTION 41

- (Exam Topic 3)

You develop data engineering solutions for a company.

A project requires analysis of real-time Twitter feeds. Posts that contain specific keywords must be stored and processed on Microsoft Azure and then displayed by using Microsoft Power BI. You need to implement the solution.

Which five actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions	Answer Area
Create an HDInsight cluster with the Hadoop cluster type.	
Create a Jupyter Notebook.	
Run a job that uses the Spark Streaming API to ingest data from Twitter.	
Create a Runbook.	
Create an HDInsight cluster with the Spark cluster type.	
Create an table.	
Load the hvac table into Power BI Desktop.	

- A. Mastered

B. Not Mastered

Answer: A

Explanation:

Step 1: Create an HDInsight cluster with the Spark cluster type Step 2: Create a Jupyter Notebook

Step 3: Create a table

The Jupyter Notebook that you created in the previous step includes code to create an hvac table. Step 4: Run a job that uses the Spark Streaming API to ingest data from Twitter

Step 5: Load the hvac table into Power BI Desktop

You use Power BI to create visualizations, reports, and dashboards from the Spark cluster data. References:

<https://acadgild.com/blog/streaming-twitter-data-using-spark>

<https://docs.microsoft.com/en-us/azure/hdinsight/spark/apache-spark-use-with-data-lake-store>

NEW QUESTION 44

- (Exam Topic 3)

An application will use Microsoft Azure Cosmos DB as its data solution. The application will use the Cassandra API to support a column-based database type that uses containers to store items.

You need to provision Azure Cosmos DB. Which container name and item name should you use? Each correct answer presents part of the solutions.

NOTE: Each correct answer selection is worth one point.

- A. table
- B. collection
- C. graph
- D. entities
- E. rows

Answer: AE

Explanation:

Depending on the choice of the API, an Azure Cosmos item can represent either a document in a collection, a row in a table or a node/edge in a graph. The following table shows the mapping between API-specific entities to an Azure Cosmos item:

Cosmos entity	SQL API	Cassandra API	Azure Cosmos DB's API for MongoDB	Gremlin API	Table API
Azure Cosmos item	Document	Row	Document	Node or Edge	Item

An Azure Cosmos container is specialized into API-specific entities as follows:

Azure Cosmos entity	SQL API	Cassandra API	Azure Cosmos DB's API for MongoDB	Gremlin API	Table API
Azure Cosmos container	Collection	Table	Collection	Graph	Table

References:

<https://docs.microsoft.com/en-us/azure/cosmos-db/databases-containers-items>

NEW QUESTION 45

- (Exam Topic 3)

A company plans to analyze a continuous flow of data from a social media platform by using Microsoft Azure Stream Analytics. The incoming data is formatted as one record per row.

You need to create the input stream.

How should you complete the REST API segment? To answer, select the appropriate configuration in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

```
{
  "properties":{
    "type":"stream",
    "serialization":{
      [dropdown]
      "properties":{
        "fieldDelimiter":",",
        "encoding":"UTF8"
      }
    },
    "datasource":{
      [dropdown]
      "properties":{
        "serviceBusNamespace":"sampleServiceBus",
        "sharedAccessPolicyName":"SampleReceiver",
        "sharedAccessPolicyKey":"<PolicyKey>"
        "eventHubName":"sampleEventHub"
      }
    },
    "compression":{
      "type":"GZip"
    }
  }
}
```

Answer Area

```
{
  "properties":{
    "type":"stream",
    "serialization":{
      [dropdown]
      "type":"CSV",
      "type":"Avro",
      "type":"JSON",
      "properties":{
        "fieldDelimiter":",",
        "encoding":"UTF8"
      }
    },
    "datasource":{
      [dropdown]
      "type":"Microsoft.Storage/Blob",
      "type":"Microsoft.ServiceBus/EventHub",
      "type":"Microsoft.Devices/IotHubs",
      "properties":{
        "serviceBusNamespace":"sampleServiceBus",
        "sharedAccessPolicyName":"SampleReceiver",
        "sharedAccessPolicyKey":"<PolicyKey>"
        "eventHubName":"sampleEventHub"
      }
    }
  }
}
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

```
{
  "properties":{
    "type":"stream",
    "serialization":{
      [dropdown]
      "type":"CSV",
      "type":"Avro",
      "type":"JSON",
      "properties":{
        "fieldDelimiter":",",
        "encoding":"UTF8"
      }
    },
    "datasource":{
      [dropdown]
      "type":"Microsoft.Storage/Blob",
      "type":"Microsoft.ServiceBus/EventHub",
      "type":"Microsoft.Devices/IotHubs",
      "properties":{
        "serviceBusNamespace":"sampleServiceBus",
        "sharedAccessPolicyName":"SampleReceiver",
        "sharedAccessPolicyKey":"<PolicyKey>"
        "eventHubName":"sampleEventHub"
      }
    }
  }
}
```

NEW QUESTION 50

- (Exam Topic 3)

You are developing a data engineering solution for a company. The solution will store a large set of key-value pair data by using Microsoft Azure Cosmos DB. The solution has the following requirements:

- Data must be partitioned into multiple containers.
- Data containers must be configured separately.

- Data must be accessible from applications hosted around the world.
- The solution must minimize latency. You need to provision Azure Cosmos DB

- A. Configure account-level throughput.
- B. Provision an Azure Cosmos DB account with the Azure Table API Enable geo-redundancy.
- C. Configure table-level throughput
- D. Replicate the data globally by manually adding regions to the Azure Cosmos DB account.
- E. Provision an Azure Cosmos DB account with the Azure Table AP
- F. Enable multi-region writes.

Answer: A

NEW QUESTION 52

- (Exam Topic 3)

Note: This question is part of series of questions that present the same scenario. Each question in the series contain a unique solution. Determine whether the solution meets the stated goals.

You develop data engineering solutions for a company.

A project requires the deployment of resources to Microsoft Azure for batch data processing on Azure HDInsight. Batch processing will run daily and must: Scale to minimize costs

Be monitored for cluster performance

You need to recommend a tool that will monitor clusters and provide information to suggest how to scale. Solution: Monitor cluster load using the Ambari Web UI.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Ambari Web UI does not provide information to suggest how to scale.

Instead monitor clusters by using Azure Log Analytics and HDInsight cluster management solutions. References:

<https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-hadoop-oms-log-analytics-tutorial> <https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-hadoop-manage-ambari>

NEW QUESTION 57

- (Exam Topic 3)

A company runs Microsoft Dynamics CRM with Microsoft SQL Server on-premises. SQL Server Integration Services (SSIS) packages extract data from Dynamics CRM APIs, and load the data into a SQL Server data warehouse.

The datacenter is running out of capacity. Because of the network configuration, you must extract on premises data to the cloud over https. You cannot open any additional ports. The solution must implement the least amount of effort.

You need to create the pipeline system.

Which component should you use? To answer, select the appropriate technology in the dialog box in the answer area.

NOTE: Each correct selection is worth one point.

Action	Technology					
Extract SQL data on-premises	<table border="1"> <tr><td>Self-hosted integration runtime</td><td rowspan="4" style="text-align: center;">▼</td></tr> <tr><td>Azure-SSIS integration runtime</td></tr> <tr><td>Azure integration runtime</td></tr> <tr><td>Source</td></tr> </table>	Self-hosted integration runtime	▼	Azure-SSIS integration runtime	Azure integration runtime	Source
Self-hosted integration runtime	▼					
Azure-SSIS integration runtime						
Azure integration runtime						
Source						
Load SQL data warehouse	<table border="1"> <tr><td>Self-hosted integration runtime</td><td rowspan="4" style="text-align: center;">▼</td></tr> <tr><td>Azure-SSIS integration runtime</td></tr> <tr><td>Azure integration runtime</td></tr> <tr><td>Sink</td></tr> </table>	Self-hosted integration runtime	▼	Azure-SSIS integration runtime	Azure integration runtime	Sink
Self-hosted integration runtime	▼					
Azure-SSIS integration runtime						
Azure integration runtime						
Sink						

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Source

For Copy activity, it requires source and sink linked services to define the direction of data flow. Copying between a cloud data source and a data source in private network: if either source or sink linked

service points to a self-hosted IR, the copy activity is executed on that self-hosted Integration Runtime.

Box 2: Self-hosted integration runtime

A self-hosted integration runtime can run copy activities between a cloud data store and a data store in a private network, and it can dispatch transform activities against compute resources in an on-premises network or an Azure virtual network. The installation of a self-hosted integration runtime needs on an on-premises machine or a virtual machine (VM) inside a private network.

References:

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

NEW QUESTION 61

- (Exam Topic 3)

You configure monitoring for a Microsoft Azure SQL Data Warehouse implementation. The implementation uses PolyBase to load data from comma-separated value (CSV) files stored in Azure Data Lake Gen 2 using an external table.

Files with an invalid schema cause errors to occur. You need to monitor for an invalid schema error. For which error should you monitor?

- A. EXTERNAL TABLE access failed due to internal error: 'Java exception raised on call to HdfsBridge_Connect: Error[com.microsoft.polybase.client.KerberosSecureLogin] occurred while accessing external files.'
- B. EXTERNAL TABLE access failed due to internal error: 'Java exception raised on call to HdfsBridge_Connect: Error [No FileSystem for scheme: wasbs] occurred while accessing external file.'
- C. Cannot execute the query "Remote Query" against OLE DB provider "SQLNCLI11": for linked server "(null)", Query aborted- the maximum reject threshold (0 rows) was reached while regarding from an external source: 1 rows rejected out of total 1 rows processed.
- D. EXTERNAL TABLE access failed due to internal error: 'Java exception raised on call to HdfsBridge_Connect: Error [Unable to instantiate LoginClass] occurredwhile accessing external files.'

Answer: C

Explanation:

Customer Scenario:

SQL Server 2016 or SQL DW connected to Azure blob storage. The CREATE EXTERNAL TABLE DDL points to a directory (and not a specific file) and the directory contains files with different schemas.

SSMS Error:

Select query on the external table gives the following error: Msg 7320, Level 16, State 110, Line 14

Cannot execute the query "Remote Query" against OLE DB provider "SQLNCLI11" for linked server "(null)". Query aborted-- the maximum reject threshold (0 rows) was reached while reading from an external source: 1 rows rejected out of total 1 rows processed.

Possible Reason:

The reason this error happens is because each file has different schema. The PolyBase external table DDL when pointed to a directory recursively reads all the files in that directory. When a column or data type mismatch happens, this error could be seen in SSMS.

Possible Solution:

If the data for each table consists of one file, then use the filename in the LOCATION section prepended by the directory of the external files. If there are multiple files per table, put each set of files into different directories in Azure Blob Storage and then you can point LOCATION to the directory instead of a particular file. The latter suggestion is the best practices recommended by SQLCAT even if you have one file per table.

NEW QUESTION 64

- (Exam Topic 3)

You develop data engineering solutions for a company.

A project requires the deployment of data to Azure Data Lake Storage.

You need to implement role-based access control (RBAC) so that project members can manage the Azure Data Lake Storage resources.

Which three actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Assign Azure AD security groups to Azure Data Lake Storage.
- B. Configure end-user authentication for the Azure Data Lake Storage account.
- C. Configure service-to-service authentication for the Azure Data Lake Storage account.
- D. Create security groups in Azure Active Directory (Azure AD) and add project members.
- E. Configure access control lists (ACL) for the Azure Data Lake Storage account.

Answer: ADE

NEW QUESTION 67

- (Exam Topic 3)

You are a data architect. The data engineering team needs to configure a synchronization of data between an on-premises Microsoft SQL Server database to Azure SQL Database.

Ad-hoc and reporting queries are being overutilized the on-premises production instance. The synchronization process must:

Perform an initial data synchronization to Azure SQL Database with minimal downtime Perform bi-directional data synchronization after initial synchronization

You need to implement this synchronization solution. Which synchronization method should you use?

- A. transactional replication
- B. Data Migration Assistant (DMA)
- C. backup and restore
- D. SQL Server Agent job
- E. Azure SQL Data Sync

Answer: E

Explanation:

SQL Data Sync is a service built on Azure SQL Database that lets you synchronize the data you select bi-directionally across multiple SQL databases and SQL Server instances.

With Data Sync, you can keep data synchronized between your on-premises databases and Azure SQL databases to enable hybrid applications.

Compare Data Sync with Transactional Replication

	Data Sync	Transactional Replication
Advantages	<ul style="list-style-type: none"> - Active-active support - Bi-directional between on-premises and Azure SQL Database 	<ul style="list-style-type: none"> - Lower latency - Transactional consistency - Reuse existing topology after migration
Disadvantages	<ul style="list-style-type: none"> - 5 min or more latency - No transactional consistency - Higher performance impact 	<ul style="list-style-type: none"> - Can't publish from Azure SQL Database single database or pooled database - High maintenance cost

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-sync-data>

NEW QUESTION 69

- (Exam Topic 3)

You manage the Microsoft Azure Databricks environment for a company. You must be able to access a private Azure Blob Storage account. Data must be available to all Azure Databricks workspaces. You need to provide the data access.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions	Answer Area
Upload a certificate	
Add secrets to the scope	
Use Blob Storage access key	
Create a secret scope	
Configure a JDBC connector	
Mount the Azure Blob Storage container	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: Create a secret scope Step 2: Add secrets to the scope

Note: `dbutils.secrets.get(scope = "<scope-name>", key = "<key-name>")` gets the key that has been stored as a secret in a secret scope.

Step 3: Mount the Azure Blob Storage container

You can mount a Blob Storage container or a folder inside a container through Databricks File System - DBFS. The mount is a pointer to a Blob Storage container, so the data is never synced locally.

Note: To mount a Blob Storage container or a folder inside a container, use the following command:

```
Python dbutils.fs.mount(
source = "wasbs://<your-container-name>@<your-storage-account-name>.blob.core.windows.net", mount_point = "/mnt/<mount-name>",
extra_configs = {"<conf-key>":dbutils.secrets.get(scope = "<scope-name>", key = "<key-name>")}) where:
dbutils.secrets.get(scope = "<scope-name>", key = "<key-name>") gets the key that has been stored as a secret in a secret scope.
```

References:

<https://docs.databricks.com/spark/latest/data-sources/azure/azure-storage.html>

NEW QUESTION 71

- (Exam Topic 3)

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 question 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.

A company uses Azure Data Lake Gen 1 Storage to store big data related to consumer behavior. You need to implement logging.

Solution: Use information stored in Azure Active Directory reports.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

NEW QUESTION 76

- (Exam Topic 3)

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution. Determine whether the solution meets the stated goals.

You develop a data ingestion process that will import data to a Microsoft Azure SQL Data Warehouse. The data to be ingested resides in parquet files stored in an Azure Data lake Gen 2 storage account.

You need to load the data from the Azure Data Lake Gen 2 storage account into the Azure SQL Data Warehouse.

Solution:

1. Create an external data source pointing to the Azure storage account
2. Create a workload group using the Azure storage account name as the pool name
3. Load the data using the INSERT...SELECT statement

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

You need to create an external file format and external table using the external data source. You then load the data using the CREATE TABLE AS SELECT statement.

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-load-from-azure-data-lake-store>

NEW QUESTION 80

- (Exam Topic 3)

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution. Determine whether the solution meets the stated goals.

You develop a data ingestion process that will import data to a Microsoft Azure SQL Data Warehouse. The data to be ingested resides in parquet files stored in an Azure Data Lake Gen 2 storage account. You need to load the data from the Azure Data Lake Gen 2 storage account into the Azure SQL Data Warehouse.

Solution:

1. Create an external data source pointing to the Azure storage account
2. Create an external file format and external table using the external data source
3. Load the data using the INSERT...SELECT statement

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

You load the data using the CREATE TABLE AS SELECT statement. References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-load-from-azure-data-lake-store>

NEW QUESTION 83

- (Exam Topic 3)

A company has a SaaS solution that uses Azure SQL Database with elastic pools. The solution contains a dedicated database for each customer organization. Customer organizations have peak usage at different periods during the year.

You need to implement the Azure SQL Database elastic pool to minimize cost. Which option or options should you configure?

- A. Number of transactions only
- B. eDTUs per database only
- C. Number of databases only
- D. CPU usage only
- E. eDTUs and max data size

Answer: E

Explanation:

The best size for a pool depends on the aggregate resources needed for all databases in the pool. This involves determining the following:

- Maximum resources utilized by all databases in the pool (either maximum DTUs or maximum vCores depending on your choice of resourcing model).
- Maximum storage bytes utilized by all databases in the pool.

Note: Elastic pools enable the developer to purchase resources for a pool shared by multiple databases to accommodate unpredictable periods of usage by individual databases. You can configure resources for the pool based either on the DTU-based purchasing model or the vCore-based purchasing model.

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-pool>

NEW QUESTION 86

- (Exam Topic 3)

A company has a SaaS solutions that will uses Azure SQL Database with elastic pools. The solution will have a dedicated database for each customer organization Customer organizations have peak usage at different periods during the year.

Which two factors affect your costs when sizing the Azure SQL Database elastic pools? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

- A. maximum data size
- B. number of databases
- C. eDTUs consumption
- D. number of read operations
- E. number of transactions

Answer: AC

NEW QUESTION 90

- (Exam Topic 3)

You are a data engineer. You are designing a Hadoop Distributed File System (HDFS) architecture. You plan to use Microsoft Azure Data Lake as a data storage repository.

You must provision the repository with a resilient data schema. You need to ensure the resiliency of the Azure Data Lake Storage. What should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Requirement	Node			
Provide data access to clients.	<table border="1"> <tr> <td>DataNode</td> <td rowspan="2">V</td> </tr> <tr> <td>NameNode</td> </tr> </table>	DataNode	V	NameNode
DataNode	V			
NameNode				
Run operations on files and directories of the file system.	<table border="1"> <tr> <td>DataNode</td> <td rowspan="2">V</td> </tr> <tr> <td>NameNode</td> </tr> </table>	DataNode	V	NameNode
DataNode	V			
NameNode				
Perform block creation, deletion, and replication.	<table border="1"> <tr> <td>DataNode</td> <td rowspan="2">V</td> </tr> <tr> <td>NameNode</td> </tr> </table>	DataNode	V	NameNode
DataNode	V			
NameNode				

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: NameNode

An HDFS cluster consists of a single NameNode, a master server that manages the file system namespace and regulates access to files by clients.

Box 2: DataNode

The DataNodes are responsible for serving read and write requests from the file system's clients. Box 3: DataNode

The DataNodes perform block creation, deletion, and replication upon instruction from the NameNode.

Note: HDFS has a master/slave architecture. An HDFS cluster consists of a single NameNode, a master server that manages the file system namespace and regulates access to files by clients. In addition, there are a number of DataNodes, usually one per node in the cluster, which manage storage attached to the nodes that they run on. HDFS exposes a file system namespace and allows user data to be stored in files. Internally, a file is split into one or more blocks and these blocks are stored in a set of DataNodes. The NameNode executes file system namespace operations like opening, closing, and renaming files and directories. It also determines the mapping of blocks to DataNodes. The DataNodes are responsible for serving read and write requests from the file system's clients. The DataNodes also perform block creation, deletion, and replication upon instruction from the NameNode.

References: https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html#NameNode+and+DataNodes

NEW QUESTION 94

- (Exam Topic 3)

A company has a Microsoft Azure HDInsight solution that uses different cluster types to process and analyze data. Operations are continuous.

Reports indicate slowdowns during a specific lime window.

You need to determine a monitoring solution to track down the issue in the least amount of time. What should you use?

- A. Azure Log Analytics log search query
- B. Ambari REST API
- C. Azure Monitor Metrics
- D. HDInsight .NET SDK
- E. Azure Log Analytics alert rule query

Answer: B

Explanation:

Ambari is the recommended tool for monitoring the health for any given HDInsight cluster.

Note: Azure HDInsight is a high-availability service that has redundant gateway nodes, head nodes, and ZooKeeper nodes to keep your HDInsight clusters running smoothly. While this ensures that a single failure will not affect the functionality of a cluster, you may still want to monitor cluster health so you are alerted when an issue does arise. Monitoring cluster health refers to monitoring whether all nodes in your cluster and the components that run on them are available and functioning correctly.

Ambari is the recommended tool for monitoring utilization across the whole cluster. The Ambari dashboard shows easily glanceable widgets that display metrics such as CPU, network, YARN memory, and HDFS disk usage. The specific metrics shown depend on cluster type. The "Hosts" tab shows metrics for individual nodes so you can ensure the load on your cluster is evenly distributed.

References:

<https://azure.microsoft.com/en-us/blog/monitoring-on-hdinsight-part-1-an-overview/>

NEW QUESTION 97

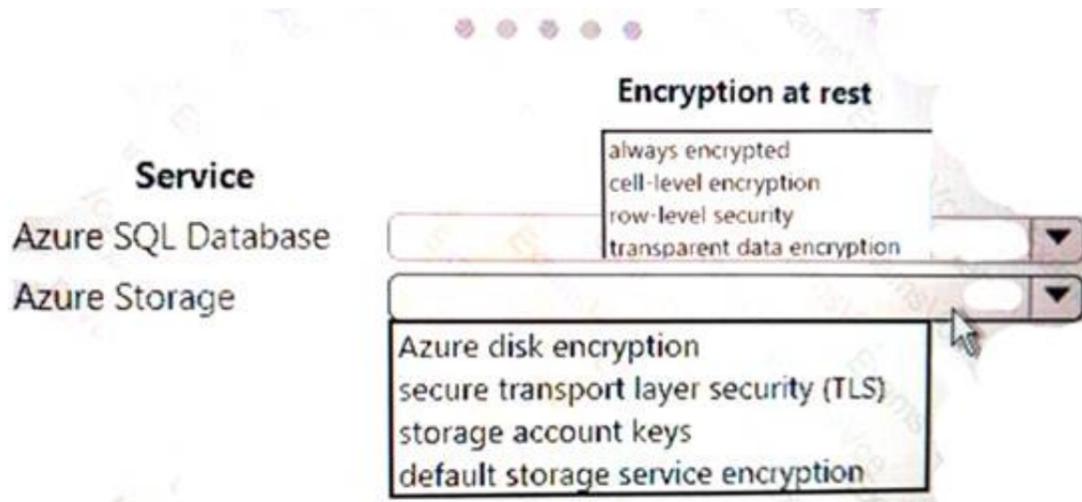
- (Exam Topic 3)

Your company uses Azure SQL Database and Azure Blob storage.

All data at rest must be encrypted by using the company's own key. The solution must minimize administrative effort and the impact to applications which use the database.

You need to configure security.

What should you implement? To answer, select the appropriate option in the answer area. NOTE: Each correct selection is worth one point.



- A. Mastered
- B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 99

- (Exam Topic 3)

You develop data engineering solutions for a company.

You need to ingest and visualize real-time Twitter data by using Microsoft Azure.

Which three technologies should you use? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Event Grid topic
- B. Azure Stream Analytics Job that queries Twitter data from an Event Hub
- C. Azure Stream Analytics Job that queries Twitter data from an Event Grid
- D. Logic App that sends Twitter posts which have target keywords to Azure
- E. Event Grid subscription
- F. Event Hub instance

Answer: BDF

Explanation:

You can use Azure Logic apps to send tweets to an event hub and then use a Stream Analytics job to read from event hub and send them to PowerBI.

References:

<https://community.powerbi.com/t5/Integrations-with-Files-and/Twitter-streaming-analytics-step-by-step/td-p/95>

NEW QUESTION 104

.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

DP-200 Practice Exam Features:

- * DP-200 Questions and Answers Updated Frequently
- * DP-200 Practice Questions Verified by Expert Senior Certified Staff
- * DP-200 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * DP-200 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

100% Actual & Verified — Instant Download, Please Click
[Order The DP-200 Practice Test Here](#)