

Microsoft

Exam Questions DP-200

Implementing an Azure Data Solution



NEW QUESTION 1

- (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 2

- (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 Randomized
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?

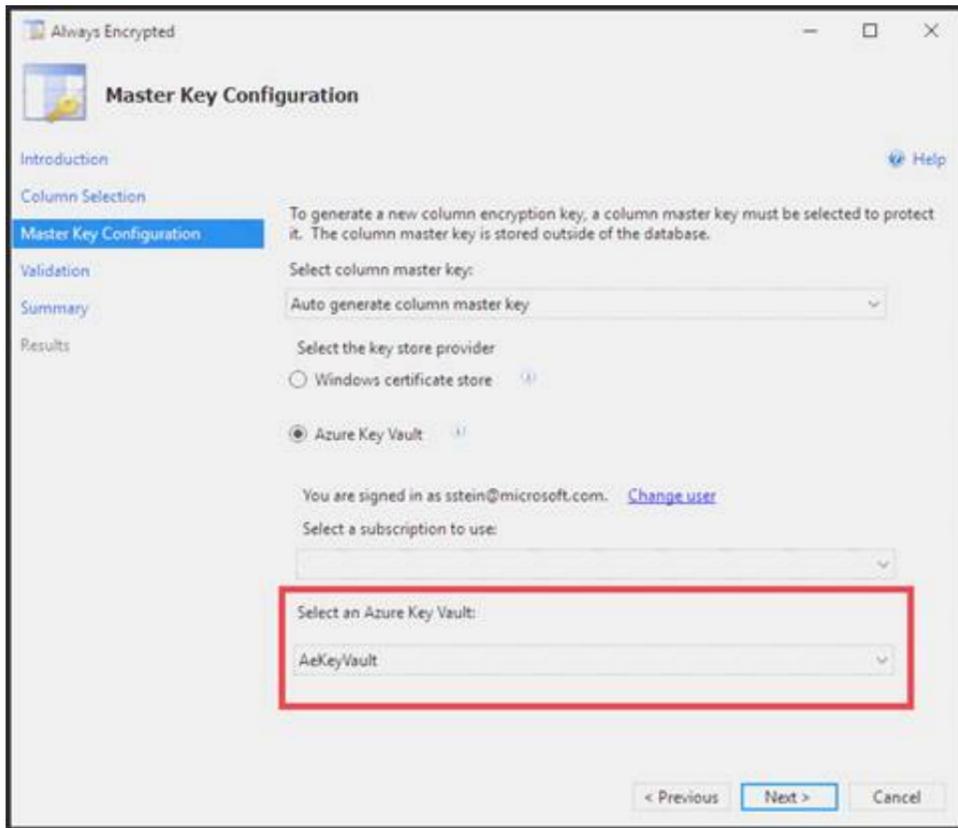
- A. Yes
- B. No

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 3

- (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 Azure Key Vault
5. Validate configuration results and deploy the solution Does the solution meet the goal?

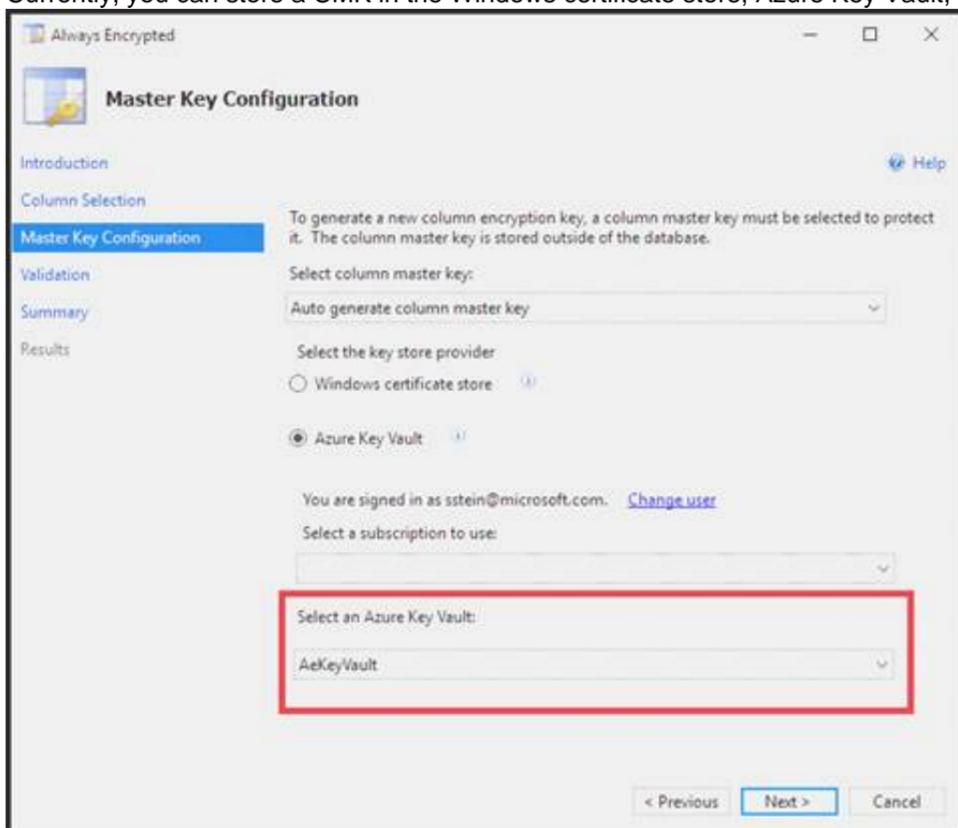
- A. Yes
- B. No

Answer: A

Explanation:

We 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 4

- (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 5

- (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 6

- (Exam Topic 3)

A company is planning to use Microsoft Azure Cosmos DB as the data store for an application. You have the following Azure CLI command:

```
az cosmosdb create --name "cosmosdbdev1" --resource-group "rgdev"
```

You need to minimize latency and expose the SQL API. How should you complete the command? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Parameter	Value					
<code>--default-consistency-level</code>	<table border="1"> <tr><td>Strong</td><td rowspan="4">v</td></tr> <tr><td>Session</td></tr> <tr><td>Eventual</td></tr> <tr><td>Bounded staleness</td></tr> </table>	Strong	v	Session	Eventual	Bounded staleness
Strong	v					
Session						
Eventual						
Bounded staleness						
<code>--kind</code>	<table border="1"> <tr><td>Parse</td><td rowspan="3">v</td></tr> <tr><td>MongoDB</td></tr> <tr><td>GlobalDocumentDB</td></tr> </table>	Parse	v	MongoDB	GlobalDocumentDB	
Parse	v					
MongoDB						
GlobalDocumentDB						

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Eventual

With Azure Cosmos DB, developers can choose from five well-defined consistency models on the consistency spectrum. From strongest to more relaxed, the models include strong, bounded staleness, session, consistent prefix, and eventual consistency. The following image shows the different consistency levels as a spectrum.



Box 2: GlobalDocumentDB

Select Core(SQL) to create a document database and query by using SQL syntax.

Note: The API determines the type of account to create. Azure Cosmos DB provides five APIs: Core(SQL) and MongoDB for document databases, Gremlin for graph databases, Azure Table, and Cassandra.

References:

<https://docs.microsoft.com/en-us/azure/cosmos-db/consistency-levels> <https://docs.microsoft.com/en-us/azure/cosmos-db/create-sql-api-dotnet>

NEW QUESTION 7

- (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 8

- (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: Configure Azure Data Lake Storage diagnostics to store logs and metrics in a storage account. Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

NEW QUESTION 9

- (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 10

- (Exam Topic 3)

You are a data engineer implementing a lambda architecture on Microsoft Azure. You use an open-source big data solution to collect, process, and maintain data. The analytical data store performs poorly.

You must implement a solution that meets the following requirements:

- Provide data warehousing
- Reduce ongoing management activities
- Deliver SQL query responses in less than one second

You need to create an HDInsight cluster to meet the requirements. Which type of cluster should you create?

- A. Interactive Query
- B. Apache Hadoop
- C. Apache HBase
- D. Apache Spark

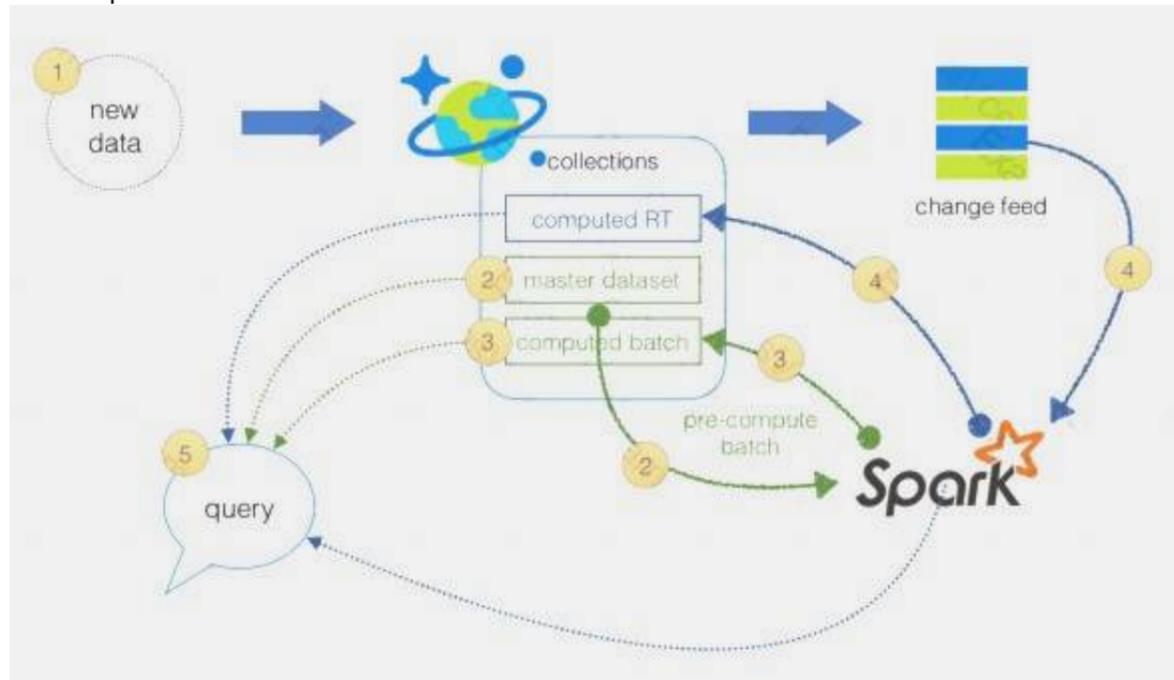
Answer: D

Explanation:

Lambda Architecture with Azure:

Azure offers you a combination of following technologies to accelerate real-time big data analytics:

- Azure Cosmos DB, a globally distributed and multi-model database service.
- Apache Spark for Azure HDInsight, a processing framework that runs large-scale data analytics applications.
- The Spark to Azure Cosmos DB Connector



Note: Lambda architecture is a data-processing architecture designed to handle massive quantities of data by taking advantage of both batch processing and stream processing methods, and minimizing the latency involved in querying big data.

References:

<https://sqlwithmanoj.com/2018/02/16/what-is-lambda-architecture-and-what-azure-offers-with-its-new-cosmos->

NEW QUESTION 10

- (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.

Actions
Answer Area

- Configure routes.
- Create an Azure Data Lake Storage container.
- Create an IoT Hub and add the Azure Stream Analytics module to the IoT Hub namespace.
- Create an Azure Stream Analytics edge job and configure job definition save location.
- Create an Azure Stream Analytics cloud job and configure job definition save location.
- Create an Azure Blob storage container.
- Configure Streaming Units

➤

➤

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

NEW QUESTION 13

- (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 clusters by using Azure Log Analytics and HDInsight cluster management solutions. Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

HDInsight provides cluster-specific management solutions that you can add for Azure Monitor logs. Management solutions add functionality to Azure Monitor logs, providing additional data and analysis tools. These solutions collect important performance metrics from your HDInsight clusters and provide the tools to search the metrics. These solutions also provide visualizations and dashboards for most cluster types supported in HDInsight. By using the metrics that you collect with the solution, you can create custom monitoring rules and alerts.

NEW QUESTION 17

- (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 22

- (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 27

- (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":{
      
      "properties":{
        "fieldDelimiter":",",
        "encoding":"UTF8"
      }
    },
    "datasource":{
      
      "properties":{
        "serviceBusNamespace":"sampleServiceBus",
        "sharedAccessPolicyName":"SampleReceiver",
        "sharedAccessPolicyKey":"<PolicyKey>"
        "eventHubName":"sampleEventHub"
      }
    },
    "compression":{
      "type":"GZip"
    }
  }
}

```

Answer Area

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

NEW QUESTION 28

- (Exam Topic 3)

You develop data engineering solutions for a company. You must migrate data from Microsoft Azure Blob storage to an Azure SQL Data Warehouse for further transformation. You need to implement the solution.

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.

Actions	Answer Area
Provision an Azure SQL Data Warehouse instance.	
Connect to the Blob storage container by using SQL Server Management Studio.	
Provision an Azure Blob storage container.	
Run Transact-SQL statements to load data.	
Connect to the Azure SQL Data Warehouse by using SQL Server Management Studio.	
Build external tables by using Azure portal.	
Build external tables by using SQL Server Management Studio.	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: Provision an Azure SQL Data Warehouse instance. Create a data warehouse in the Azure portal.

Step 2: Connect to the Azure SQL Data warehouse by using SQL Server Management Studio Connect to the data warehouse with SSMS (SQL Server Management Studio)

Step 3: Build external tables by using the SQL Server Management Studio

Create external tables for data in Azure blob storage.

You are ready to begin the process of loading data into your new data warehouse. You use external tables to load data from the Azure storage blob.

Step 4: Run Transact-SQL statements to load data.

You can use the CREATE TABLE AS SELECT (CTAS) T-SQL statement to load the data from Azure Storage Blob into new tables in your data warehouse.

References:

<https://github.com/MicrosoftDocs/azure-docs/blob/master/articles/sql-data-warehouse/load-data-from-azure-blo>

NEW QUESTION 33

- (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 37

- (Exam Topic 3)

You need to develop a pipeline for processing data. The pipeline must meet the following requirements.

- Scale up and down resources for cost reduction.
- Use an in-memory data processing engine to speed up ETL and machine learning operations.
- Use streaming capabilities.
- Provide the ability to code in SQL, Python, Scala, and R.
- Integrate workspace collaboration with Git. What should you use?

- A. HDInsight Spark Cluster
- B. Azure Stream Analytics
- C. HDInsight Hadoop Cluster
- D. Azure SQL Data Warehouse

Answer: B

NEW QUESTION 39

- (Exam Topic 3)

The data engineering team manages Azure HDInsight clusters. The team spends a large amount of time creating and destroying clusters daily because most of the data pipeline process runs in minutes.

You need to implement a solution that deploys multiple HDInsight clusters with minimal effort. What should you implement?

- A. Azure Databricks
- B. Azure Traffic Manager
- C. Azure Resource Manager templates
- D. Ambari web user interface

Answer: C

Explanation:

A Resource Manager template makes it easy to create the following resources for your application in a single, coordinated operation:

- HDInsight clusters and their dependent resources (such as the default storage account).
- Other resources (such as Azure SQL Database to use Apache Sqoop).

In the template, you define the resources that are needed for the application. You also specify deployment parameters to input values for different environments.

The template consists of JSON and expressions that you use to construct values for your deployment.

References:

<https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-hadoop-create-linux-clusters-arm-templates>

NEW QUESTION 44

- (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: Create an Azure Automation runbook to copy events. Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

NEW QUESTION 48

- (Exam Topic 3)

Your company uses Microsoft Azure SQL Database configured with Elastic pool. You use Elastic Database jobs to run queries across all databases in the pod. You need to analyze, troubleshoot, and report on components responsible for running Elastic Database jobs. You need to determine the component responsible for running job service tasks.

Which components should you use for each Elastic pool job services task? To answer, drag the appropriate component to the correct task. Each component 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 53

- (Exam Topic 3)

A company manages several on-premises Microsoft SQL Server databases.

You need to migrate the databases to Microsoft Azure by using the backup process of Microsoft SQL Server. Which data technology should you use?

- A. Azure SQL Database Managed Instance
- B. Azure SQL Data Warehouse
- C. Azure Cosmos DB
- D. Azure SQL Database single database

Answer: D

NEW QUESTION 55

- (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] occurred while 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 57

- (Exam Topic 3)

You develop data engineering solutions for a company.

You must integrate the company's on-premises Microsoft SQL Server data with Microsoft Azure SQL Database. Data must be transformed incrementally.

You need to implement the data integration solution.

Which tool should you use to configure a pipeline to copy data?

- A. Use the Copy Data tool with Blob storage linked service as the source
- B. Use Azure PowerShell with SQL Server linked service as a source
- C. Use Azure Data Factory UI with Blob storage linked service as a source
- D. Use the .NET Data Factory API with Blob storage linked service as the source

Answer: C

Explanation:

The Integration Runtime is a customer managed data integration infrastructure used by Azure Data Factory to provide data integration capabilities across different network environments.

A linked service defines the information needed for Azure Data Factory to connect to a data resource. We have three resources in this scenario for which linked services are needed:

- On-premises SQL Server
- Azure Blob Storage
- Azure SQL database

Note: Azure Data Factory is a fully managed cloud-based data integration service that orchestrates and automates the movement and transformation of data. The key concept in the ADF model is pipeline. A pipeline is a logical grouping of Activities, each of which defines the actions to perform on the data contained in Datasets. Linked services are used to define the information needed for Data Factory to connect to the data resources.

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/move-sql-azure-adf>

NEW QUESTION 59

- (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 62

- (Exam Topic 3)

A company plans to use Azure Storage for file storage purposes. Compliance rules require: A single storage account to store all operations including reads, writes and deletes

Retention of an on-premises copy of historical operations You need to configure the storage account.

Which two actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Configure the storage account to log read, write and delete operations for service type Blob
- B. Use the AzCopy tool to download log data from \$logs/blob
- C. Configure the storage account to log read, write and delete operations for service-type table
- D. Use the storage client to download log data from \$logs/table
- E. Configure the storage account to log read, write and delete operations for service type queue

Answer: AB

Explanation:

Storage Logging logs request data in a set of blobs in a blob container named \$logs in your storage account. This container does not show up if you list all the blob containers in your account but you can see its contents if you access it directly.

To view and analyze your log data, you should download the blobs that contain the log data you are interested in to a local machine. Many storage-browsing tools enable you to download blobs from your storage account; you can also use the Azure Storage team provided command-line Azure Copy Tool (AzCopy) to download your log data.

References:

<https://docs.microsoft.com/en-us/rest/api/storageservices/enabling-storage-logging-and-accessing-log-data>

NEW QUESTION 66

- (Exam Topic 3)

Your company plans to create an event processing engine to handle streaming data from Twitter. The data engineering team uses Azure Event Hubs to ingest the streaming data.

You need to implement a solution that uses Azure Databricks to receive the streaming data from the Azure Event Hubs.

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

The screenshot shows an exam question interface with a progress indicator at the top (5 dots, 4th is filled). On the left, under the heading "Actions", there is a list of six actions in rectangular boxes:

- Create and configure a Notebook that consumes the streaming data.
- Import data from Blob storage.
- Use Environment variables to define the Apache Spark connection.
- Configure an ODBC or JDBC Connector.
- Deploy the Azure Databricks service.
- Deploy a Spark cluster and then attach the required libraries to the cluster.

To the right of the actions list are two circular navigation buttons: a right-pointing arrow (>) and a left-pointing arrow (<). On the right side, under the heading "Answer Area", there is an empty space for the user to place and order the selected actions.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

This screenshot shows the same exam question interface as above, but with the "Answer Area" populated. A dashed red box highlights the three actions that have been moved from the "Actions" list to the "Answer Area" in the following sequence:

1. Import data from Blob storage.
2. Configure an ODBC or JDBC Connector.
3. Use Environment variables to define the Apache Spark connection.

A right-pointing arrow (>) is shown between the second and third actions in the answer area, indicating the correct sequence. The other three actions from the original list remain in the "Actions" list on the left.

NEW QUESTION 71

- (Exam Topic 3)

You are creating a managed data warehouse solution on Microsoft Azure.

You must use PolyBase to retrieve data from Azure Blob storage that resides in parquet format and load the data into a large table called FactSalesOrderDetails. You need to configure Azure SQL Data Warehouse to receive the data.

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.

The screenshot shows an exam question interface. On the left, under the heading "Actions", there is a list of seven actions in a scrollable container:

- Create an external file format to map the parquet files.
- Load the data to a staging table.
- Create the external table FactSalesOrderDetails.
- Enable Transparent Data Encryption.
- Create an external data source for Azure Blob storage.
- Create a master key on database.
- Configure PolyBase to use Azure Blob storage.

On the right, there is an "Answer Area" with a dashed red border. Below the actions list, there are two circular arrows: a right-pointing arrow above a left-pointing arrow.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

This screenshot shows the same exam question interface as above, but with the correct sequence of four actions highlighted in the "Answer Area" with a dashed red border:

- Enable Transparent Data Encryption.
- Configure PolyBase to use Azure Blob storage.
- Load the data to a staging table.
- Create an external file format to map the parquet files.

Below the answer area, there is a left-pointing circular arrow.

NEW QUESTION 75

- (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 Data Lake Gen 2 storage account.
2. Create an external file format and external table using the external data source.
3. Load the data using the CREATE TABLE AS SELECT statement.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

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 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 79

- (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><input type="checkbox"/></td> </tr> <tr> <td>NameNode</td> <td><input type="checkbox"/></td> </tr> </table>	DataNode	<input type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input type="checkbox"/>				
NameNode	<input type="checkbox"/>				
Run operations on files and directories of the file system.	<table border="1"> <tr> <td>DataNode</td> <td><input type="checkbox"/></td> </tr> <tr> <td>NameNode</td> <td><input type="checkbox"/></td> </tr> </table>	DataNode	<input type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input type="checkbox"/>				
NameNode	<input type="checkbox"/>				
Perform block creation, deletion, and replication.	<table border="1"> <tr> <td>DataNode</td> <td><input type="checkbox"/></td> </tr> <tr> <td>NameNode</td> <td><input type="checkbox"/></td> </tr> </table>	DataNode	<input type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input type="checkbox"/>				
NameNode	<input type="checkbox"/>				

- 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 82

- (Exam Topic 3)

You manage a financial computation data analysis process. Microsoft Azure virtual machines (VMs) run the process in daily jobs, and store the results in virtual hard drives (VHDs.)

The VMs product results using data from the previous day and store the results in a snapshot of the VHD. When a new month begins, a process creates a new VHD.

You must implement the following data retention requirements:

- Daily results must be kept for 90 days
- Data for the current year must be available for weekly reports
- Data from the previous 10 years must be stored for auditing purposes
- Data required for an audit must be produced within 10 days of a request. You need to enforce the data retention requirements while minimizing cost.

How should you configure the lifecycle policy? To answer, drag the appropriate JSON segments to the correct locations. Each JSON segment may be used once, more than once, or not at all. You may need to drag the split bat between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Code segments	Answer Area
<input type="text" value="delete"/>	{
<input type="text" value="blockBob"/>	"version": "0.5",
<input type="text" value="baseBlob"/>	"rules": [
<input type="text" value="snapshot"/>	{
<input type="text" value="tierToCool"/>	"name": "dataRetention",
<input type="text" value="tierToArchive"/>	"type": "Lifecycle",
	"definition": {
	"actions": {
	" <input type="text" value=""/>
	" <input type="text" value=""/> ": {daysAfterModificationGreaterThan": 365),
	" <input type="text" value=""/> ": {daysAfterModificationGreaterThan": 3650)
	},
	" <input type="text" value=""/> ": {
	" <input type="text" value=""/> ": {"daysAfterCreationGreaterThan": 90)
	}
	}
	}
	}

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

The Set-AzStorageAccountManagementPolicy cmdlet creates or modifies the management policy of an Azure Storage account.

Example: Create or update the management policy of a Storage account with ManagementPolicy rule objects.

Action -BaseBlobAction Delete -daysAfterModificationGreaterThan 100

PS C:\>\$action1 = Add-AzStorageAccountManagementPolicyAction -InputObject \$action1 -BaseBlobAction TierToArchive -daysAfterModificationGreaterThan 50

PS C:\>\$action1 = Add-AzStorageAccountManagementPolicyAction -InputObject \$action1 -BaseBlobAction TierToCool -daysAfterModificationGreaterThan 30

PS C:\>\$action1 = Add-AzStorageAccountManagementPolicyAction -InputObject \$action1 -SnapshotAction Delete -daysAfterCreationGreaterThan 100

PS C:\>\$filter1 = New-AzStorageAccountManagementPolicyFilter -PrefixMatch ab,cd

PS C:\>\$rule1 = New-AzStorageAccountManagementPolicyRule -Name Test -Action \$action1 -Filter \$filter1

PS C:\>\$action2 = Add-AzStorageAccountManagementPolicyAction -BaseBlobAction Delete

-daysAfterModificationGreaterThan 100

PS C:\>\$filter2 = New-AzStorageAccountManagementPolicyFilter References:

<https://docs.microsoft.com/en-us/powershell/module/az.storage/set-azstorageaccountmanagementpolicy>

NEW QUESTION 84

.....

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](#)