



Microsoft

Exam Questions DP-200

Implementing an Azure Data Solution

NEW QUESTION 1

- (Exam Topic 1)

You need to ensure phone-based polling data upload reliability requirements are met. How should you configure monitoring? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Setting	Value						
Metric	<table><tr><td>FileCount</td><td></td></tr><tr><td>BlobCapacity</td><td></td></tr><tr><td>FileCapacity</td><td></td></tr></table>	FileCount		BlobCapacity		FileCapacity	
FileCount							
BlobCapacity							
FileCapacity							
Aggregation	<table><tr><td>Avg</td><td></td></tr><tr><td>Sum</td><td></td></tr></table>	Avg		Sum			
Avg							
Sum							

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: FileCapacity

FileCapacity is the amount of storage used by the storage account's File service in bytes. Box 2: Avg

The aggregation type of the FileCapacity metric is Avg.

Scenario:

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

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

References:

<https://docs.microsoft.com/en-us/azure/azure-monitor/platform/metrics-supported>

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 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 3

- (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 4

- (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	<div><div>connection string</div><div>linked service name string</div><div>gateway connection string</div><div>data store name string</div></div> <div>▼</div>
Data movement activity	<div><div>Azure SQL Data Warehouse</div><div>Azure Files</div><div>Azure Blob</div><div>Azure SQL Database</div></div> <div>▼</div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Connection String

To use storage account key authentication, you use the `ConnectionString` property, which xpecify the information needed to connect to Blobl 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 5

- (Exam Topic 2)

You need to set up access to Azure SQL Database for Tier 7 and Tier 8 partners.

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

Actions	Answer Area
Connect to the Database and use Azure PowerShell to create a database firewall rule	
Set the Allow Azure Services to Access Server to Disabled	
In thier Azure portal, create a database firewall rule	
In the Azure portal, create a server firewall rule	
Connect to the database and use Transact-SQL to create a database firewall rule	
Set the Allow Azure Services to Access Server setting to Enabled	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Tier 7 and 8 data access is constrained to single endpoints managed by partners for access Step 1: Set the Allow Azure Services to Access Server setting to

Disabled

Set Allow access to Azure services to OFF for the most secure configuration.

By default, access through the SQL Database firewall is enabled for all Azure services, under Allow access to Azure services. Choose OFF to disable access for all Azure services.

Note: The firewall pane has an ON/OFF button that is labeled Allow access to Azure services. The ON setting allows communications from all Azure IP addresses and all Azure subnets. These Azure IPs or subnets might not be owned by you. This ON setting is probably more open than you want your SQL Database to be. The virtual network rule feature offers much finer granular control.

Step 2: In the Azure portal, create a server firewall rule Set up SQL Database server firewall rules

Server-level IP firewall rules apply to all databases within the same SQL Database server. To set up a server-level firewall rule:

➤ In Azure portal, select SQL databases from the left-hand menu, and select your database on the SQL databases page.

➤ On the Overview page, select Set server firewall. The Firewall settings page for the database server opens.

Step 3: Connect to the database and use Transact-SQL to create a database firewall rule

Database-level firewall rules can only be configured using Transact-SQL (T-SQL) statements, and only after you've configured a server-level firewall rule.

To setup a database-level firewall rule:

➤ In Object Explorer, right-click the database and select New Query.

➤ EXECUTE sp_set_database_firewall_rule N'Example DB Rule','0.0.0.4','0.0.0.4';

➤ On the toolbar, select Execute to create the firewall rule. References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-security-tutorial>

NEW QUESTION 6

- (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

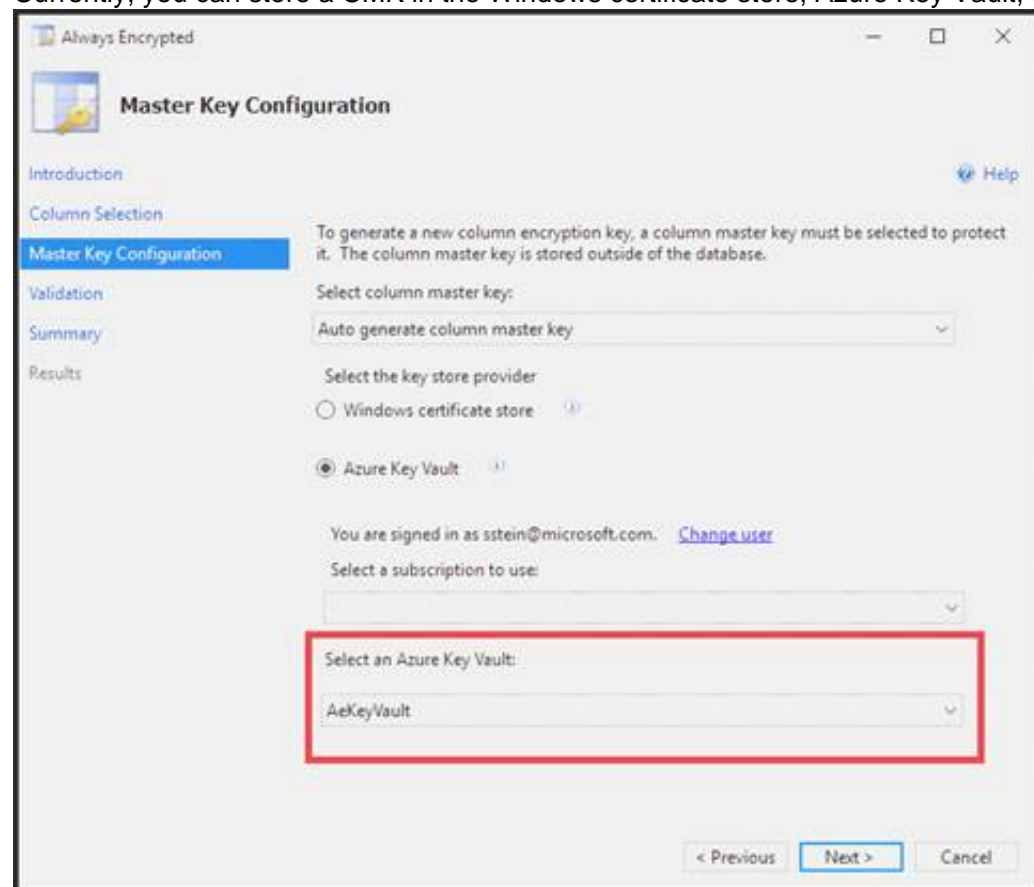
Answer: A

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 7

- (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?

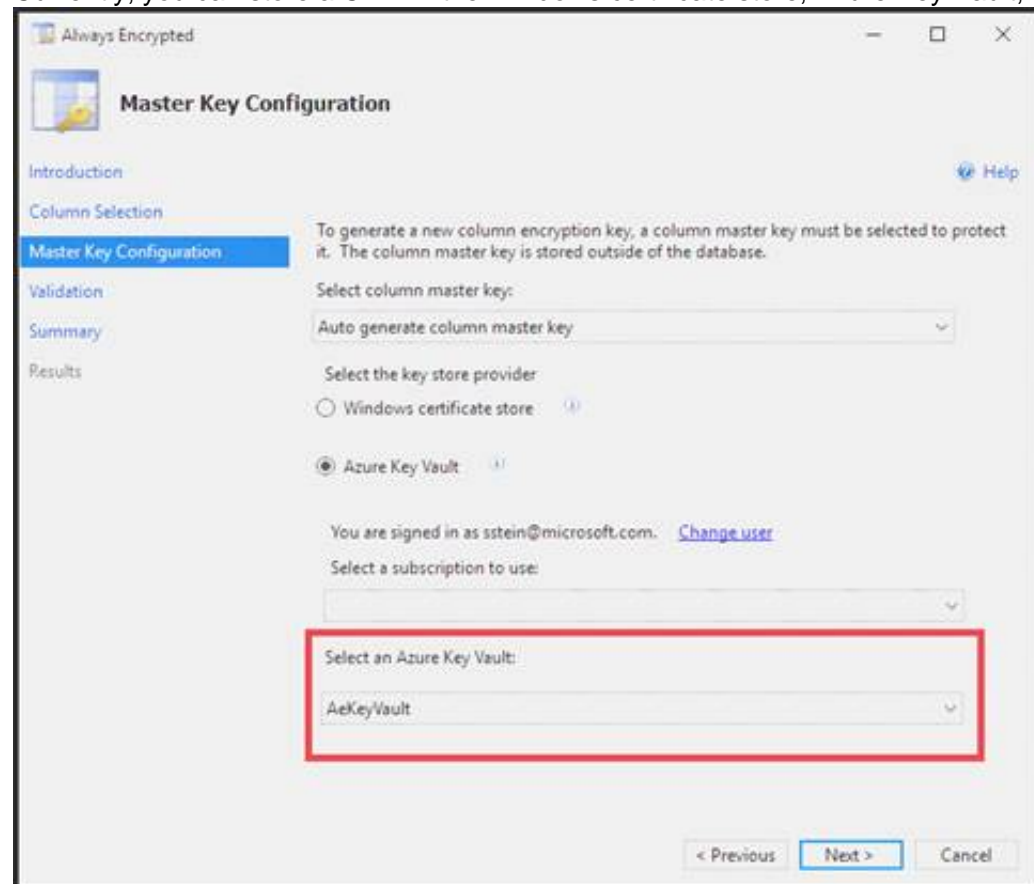
- A. Yes
- B. No

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 8

- (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 9

- (Exam Topic 3)

Each day, company plans to store hundreds of files in Azure Blob Storage and Azure Data Lake Storage. The company uses the parquet format.

You must develop a pipeline that meets the following requirements:

- ▶ Process data every six hours
- ▶ Offer interactive data analysis capabilities
- ▶ Offer the ability to process data using solid-state drive (SSD) caching
- ▶ Use Directed Acyclic Graph(DAG) processing mechanisms
- ▶ Provide support for REST API calls to monitor processes
- ▶ Provide native support for Python
- ▶ Integrate with Microsoft Power BI

You need to select the appropriate data technology to implement the pipeline. Which data technology should you implement?

- A. Azure SQL Data Warehouse
- B. HDInsight Apache Storm cluster
- C. Azure Stream Analytics
- D. HDInsight Apache Hadoop cluster using MapReduce

E. HDInsight Spark cluster

Answer: B

Explanation:

Storm runs topologies instead of the Apache Hadoop MapReduce jobs that you might be familiar with. Storm topologies are composed of multiple components that are arranged in a directed acyclic graph (DAG). Data flows between the components in the graph. Each component consumes one or more data streams, and can optionally emit one or more streams.

Python can be used to develop Storm components. References:

<https://docs.microsoft.com/en-us/azure/hdinsight/storm/apache-storm-overview>

NEW QUESTION 10

- (Exam Topic 3)

You manage a solution that uses Azure HDInsight clusters.

You need to implement a solution to monitor cluster performance and status. Which technology should you use?

- A. Azure HDInsight .NET SDK
- B. Azure HDInsight REST API
- C. Ambari REST API
- D. Azure Log Analytics
- E. Ambari Web UI

Answer: E

Explanation:

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.

The Apache Ambari project is aimed at making Hadoop management simpler by developing software for provisioning, managing, and monitoring Apache Hadoop clusters. Ambari provides an intuitive, easy-to-use Hadoop management web UI backed by its RESTful APIs.

References:

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

NEW QUESTION 10

- (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

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

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Actions

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

Answer Area

Enable advanced threat protection.

Configure the service to send email alerts to security@contoso.com

Configure the service to create alerts for threat detections of type **Data Exfiltration**.

NEW QUESTION 12

- (Exam Topic 3)

You are developing the data platform for a global retail company. The company operates during normal working hours in each region. The analytical database is used once a week for building sales projections.

Each region maintains its own private virtual network.

Building the sales projections is very resource intensive and generates upwards of 20 terabytes (TB) of data. Microsoft Azure SQL Databases must be provisioned.

- ▶ Database provisioning must maximize performance and minimize cost
- ▶ The daily sales for each region must be stored in an Azure SQL Database instance
- ▶ Once a day, the data for all regions must be loaded in an analytical Azure SQL Database instance

You need to provision Azure SQL database instances. How should you provision the database instances? To answer, drag the appropriate Azure SQL products to the correct databases. Each Azure SQL product 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.

Azure SQL products	Database	Azure SQL product
Azure SQL Database elastic pools	Daily Sales	Azure SQL product
Azure SQL Database Premium	Weekly Analysis	Azure SQL product
Azure SQL Database Managed Instance		
Azure SQL Database Hyperscale		

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: Azure SQL Database elastic pools

SQL Database elastic pools are a simple, cost-effective solution for managing and scaling multiple databases that have varying and unpredictable usage demands. The databases in an elastic pool are on a single Azure

SQL Database server and share a set number of resources at a set price. Elastic pools in Azure SQL Database enable SaaS developers to optimize the price performance for a group of databases within a prescribed budget while delivering performance elasticity for each database.

Box 2: Azure SQL Database Hyperscale

A Hyperscale database is an Azure SQL database in the Hyperscale service tier that is backed by the Hyperscale scale-out storage technology. A Hyperscale database supports up to 100 TB of data and provides high throughput and performance, as well as rapid scaling to adapt to the workload requirements. Scaling is transparent to the application – connectivity, query processing, and so on, work like any other SQL database.

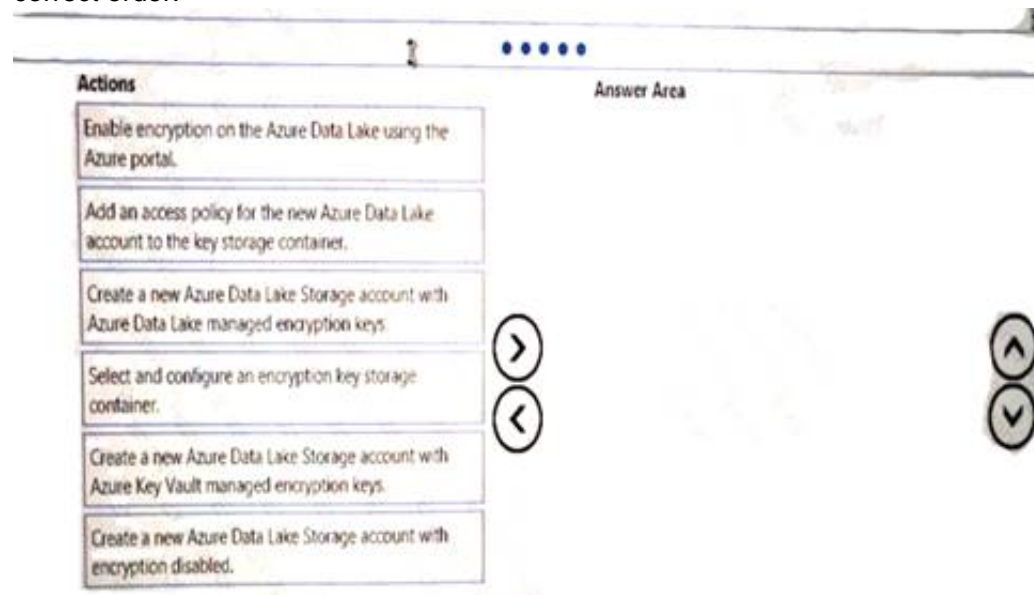
NEW QUESTION 17

- (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.

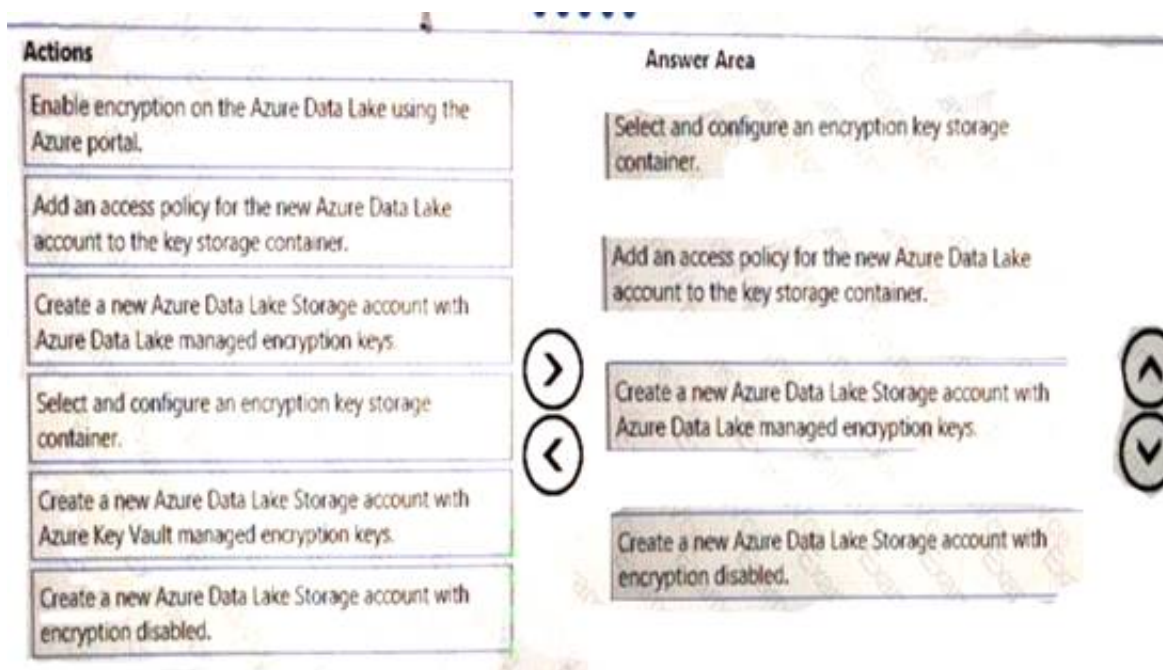
Which four actions should you perform in sequence? To answer, move the appropriate action from the list of actions to the answer area and arrange them in the correct order.



- A. Mastered
B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 22

- (Exam Topic 3)

Your company uses several Azure HDInsight clusters.

The data engineering team reports several errors with some application using these clusters. You need to recommend a solution to review the health of the clusters.

What should you include in you recommendation?

- A. Azure Automation
- B. Log Analytics
- C. Application Insights

Answer: C

NEW QUESTION 25

- (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 30

- (Exam Topic 3)

A company plans to use Azure SQL Database to support a mission-critical application.

The application must be highly available without performance degradation during maintenance windows. You need to implement the solution.

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

NOTE: Each correct selection is worth one point.

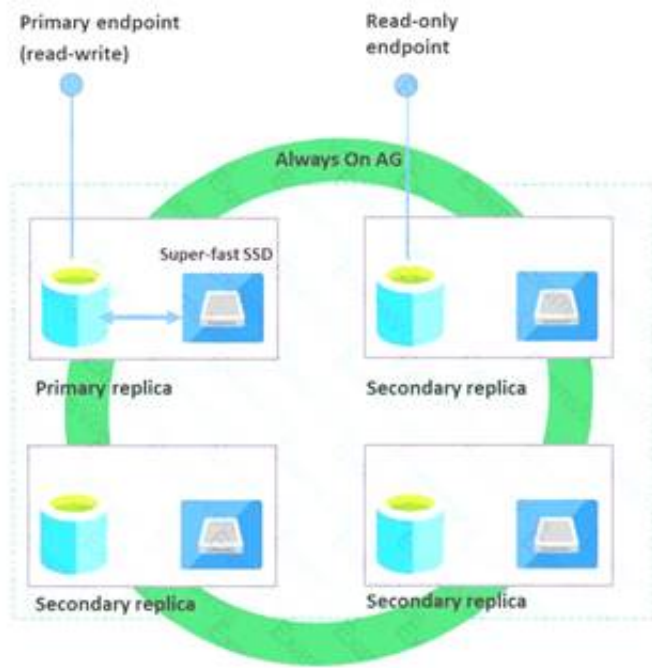
- A. Premium service tier
- B. Virtual machine Scale Sets
- C. Basic service tier
- D. SQL Data Sync
- E. Always On availability groups
- F. Zone-redundant configuration

Answer: AEF

Explanation:

Premium/business critical service tier model that is based on a cluster of database engine processes. This architectural model relies on a fact that there is always a quorum of available database engine nodes and has minimal performance impact on your workload even during maintenance activities.

In the premium model, Azure SQL database integrates compute and storage on the single node. High availability in this architectural model is achieved by replication of compute (SQL Server Database Engine process) and storage (locally attached SSD) deployed in 4-node cluster, using technology similar to SQL Server Always On Availability Groups.



Business Critical service tier: colocated compute and storage

Zone redundant configuration

By default, the quorum-set replicas for the local storage configurations are created in the same datacenter. With the introduction of Azure Availability Zones, you have the ability to place the different replicas in the quorum-sets to different availability zones in the same region. To eliminate a single point of failure, the control ring is also duplicated across multiple zones as three gateway rings (GW).

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-high-availability>

NEW QUESTION 35

- (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.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> Azure Data Factory
Use a web browser to run interactive queries and batch processes.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> Azure PowerShell
Develop batch processing applications that use Azure HDInsight.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> NoSQL database

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Scenario	Tool
Use a native client application to run interactive queries and batch processes.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input checked="" type="checkbox"/> HDInsight REST API <input type="checkbox"/> Azure Data Factory
Use a web browser to run interactive queries and batch processes.	<input checked="" type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> Azure PowerShell
Develop batch processing applications that use Azure HDInsight.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input checked="" type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> NoSQL database

NEW QUESTION 39

- (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 43

- (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 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":{
      
    },
    "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 48

- (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 50

- (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 54

- (Exam Topic 3)

Your company uses Microsoft Azure SQL Database configure 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.

Components

Control Database

Azure Service Bus

Azure Storage

Job Service

Answer Area

Task

Execution results and diagnostics

Job launcher and tracker

Job metadata and state

Component

Component

Component

Component

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Components

Control Database

Azure Service Bus

Azure Storage

Job Service

Answer Area

Task

Execution results and diagnostics

Job launcher and tracker

Job metadata and state

Component

Azure Service Bus

Job Service

Control Database

NEW QUESTION 56

- (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

Extract SQL data on-premises

Technology

Self-hosted integration runtime

Azure-SSIS integration runtime

Azure integration runtime

Source

Action

Load SQL data warehouse

Technology

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 59

- (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 64

- (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 67

- (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: Download Azure HDInsight cluster logs by using Azure PowerShell.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Reference:

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>

NEW QUESTION 69

- (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.

Actions

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.

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Actions

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.

Answer Area

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

NEW QUESTION 70

- (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 tile 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 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. 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 73

- (Exam Topic 3)

You manage a Microsoft Azure SQL Data Warehouse Gen 2.

Users report slow performance when they run commonly used queries. Users do not report performance changes for infrequently used queries. You need to monitor resource utilization to determine the source of the performance issues. Which metric should you monitor?

- A. Cache used percentage
- B. Local tempdb percentage
- C. WU percentage
- D. CPU percentage

Answer: B

NEW QUESTION 75

- (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><tr><td>DataNode</td><td><input checked="" type="checkbox"/></td></tr><tr><td>NameNode</td><td><input type="checkbox"/></td></tr></table>	DataNode	<input checked="" type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input checked="" type="checkbox"/>				
NameNode	<input type="checkbox"/>				
Run operations on files and directories of the file system.	<table><tr><td>DataNode</td><td><input checked="" type="checkbox"/></td></tr><tr><td>NameNode</td><td><input type="checkbox"/></td></tr></table>	DataNode	<input checked="" type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input checked="" type="checkbox"/>				
NameNode	<input type="checkbox"/>				
Perform block creation, deletion, and replication.	<table><tr><td>DataNode</td><td><input checked="" type="checkbox"/></td></tr><tr><td>NameNode</td><td><input type="checkbox"/></td></tr></table>	DataNode	<input checked="" type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input checked="" 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 80
.....

About Exambible

[Your Partner of IT Exam](#)

Found in 1998

Exambible is a company specialized on providing high quality IT exam practice study materials, especially Cisco CCNA, CCDA, CCNP, CCIE, Checkpoint CCSE, CompTIA A+, Network+ certification practice exams and so on. We guarantee that the candidates will not only pass any IT exam at the first attempt but also get profound understanding about the certificates they have got. There are so many alike companies in this industry, however, Exambible has its unique advantages that other companies could not achieve.

Our Advances

* 99.9% Uptime

All examinations will be up to date.

* 24/7 Quality Support

We will provide service round the clock.

* 100% Pass Rate

Our guarantee that you will pass the exam.

* Unique Gurantee

If you do not pass the exam at the first time, we will not only arrange FULL REFUND for you, but also provide you another exam of your claim, ABSOLUTELY FREE!

NEW QUESTION 1

- (Exam Topic 1)

You need to ensure phone-based polling data upload reliability requirements are met. How should you configure monitoring? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Setting	Value						
Metric	<table><tr><td>FileCount</td><td></td></tr><tr><td>BlobCapacity</td><td></td></tr><tr><td>FileCapacity</td><td></td></tr></table>	FileCount		BlobCapacity		FileCapacity	
FileCount							
BlobCapacity							
FileCapacity							
Aggregation	<table><tr><td>Avg</td><td></td></tr><tr><td>Sum</td><td></td></tr></table>	Avg		Sum			
Avg							
Sum							

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: FileCapacity

FileCapacity is the amount of storage used by the storage account's File service in bytes. Box 2: Avg

The aggregation type of the FileCapacity metric is Avg.

Scenario:

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

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

References:

<https://docs.microsoft.com/en-us/azure/azure-monitor/platform/metrics-supported>

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 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 3

- (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 4

- (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	<div><div>connection string</div><div>linked service name string</div><div>gateway connection string</div><div>data store name string</div></div>
Data movement activity	<div><div>Azure SQL Data Warehouse</div><div>Azure Files</div><div>Azure Blob</div><div>Azure SQL Database</div></div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Connection String

To use storage account key authentication, you use the `ConnectionString` property, which xpecify the information needed to connect to Blobl 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 5

- (Exam Topic 2)

You need to set up access to Azure SQL Database for Tier 7 and Tier 8 partners.

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

Actions	Answer Area
Connect to the Database and use Azure PowerShell to create a database firewall rule	
Set the Allow Azure Services to Access Server to Disabled	
In thier Azure portal, create a database firewall rule	
In the Azure portal, create a server firewall rule	
Connect to the database and use Transact-SQL to create a database firewall rule	
Set the Allow Azure Services to Access Server setting to Enabled	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Tier 7 and 8 data access is constrained to single endpoints managed by partners for access Step 1: Set the Allow Azure Services to Access Server setting to

Disabled

Set Allow access to Azure services to OFF for the most secure configuration.

By default, access through the SQL Database firewall is enabled for all Azure services, under Allow access to Azure services. Choose OFF to disable access for all Azure services.

Note: The firewall pane has an ON/OFF button that is labeled Allow access to Azure services. The ON setting allows communications from all Azure IP addresses and all Azure subnets. These Azure IPs or subnets might not be owned by you. This ON setting is probably more open than you want your SQL Database to be. The virtual network rule feature offers much finer granular control.

Step 2: In the Azure portal, create a server firewall rule Set up SQL Database server firewall rules

Server-level IP firewall rules apply to all databases within the same SQL Database server. To set up a server-level firewall rule:

➤ In Azure portal, select SQL databases from the left-hand menu, and select your database on the SQL databases page.

➤ On the Overview page, select Set server firewall. The Firewall settings page for the database server opens.

Step 3: Connect to the database and use Transact-SQL to create a database firewall rule

Database-level firewall rules can only be configured using Transact-SQL (T-SQL) statements, and only after you've configured a server-level firewall rule.

To setup a database-level firewall rule:

➤ In Object Explorer, right-click the database and select New Query.

➤ EXECUTE sp_set_database_firewall_rule N'Example DB Rule','0.0.0.4','0.0.0.4';

➤ On the toolbar, select Execute to create the firewall rule. References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-security-tutorial>

NEW QUESTION 6

- (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

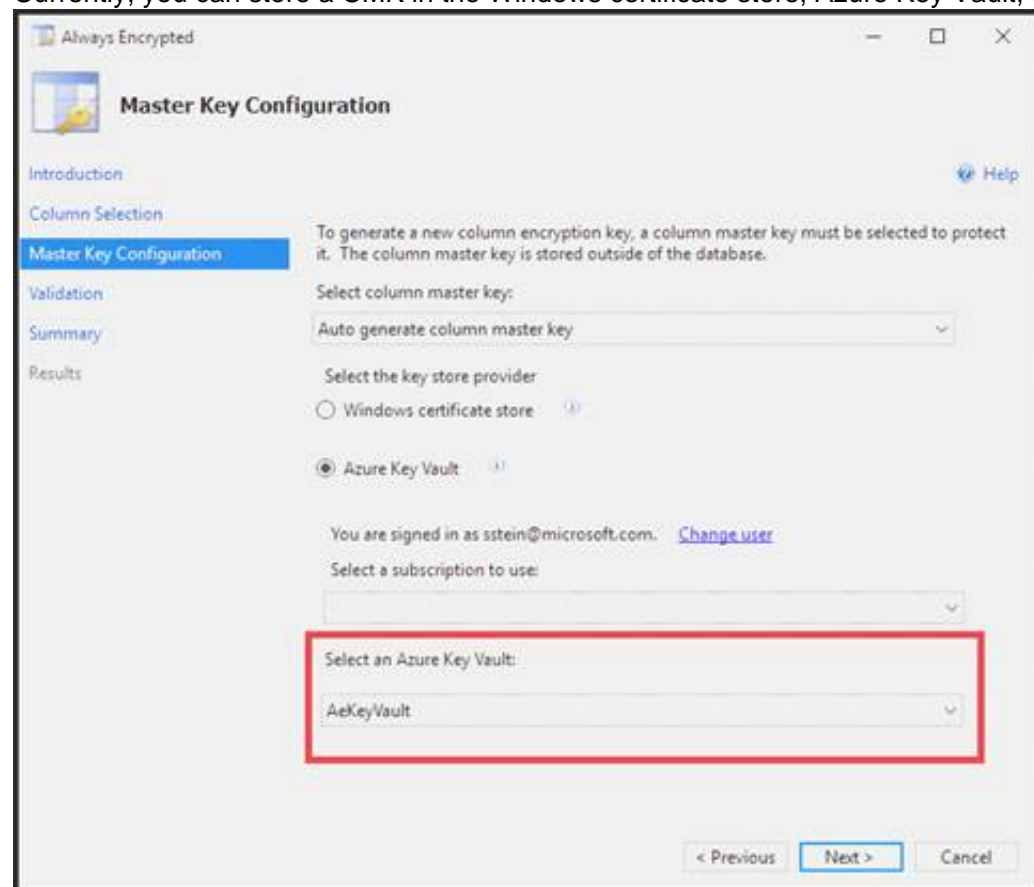
Answer: A

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 7

- (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?

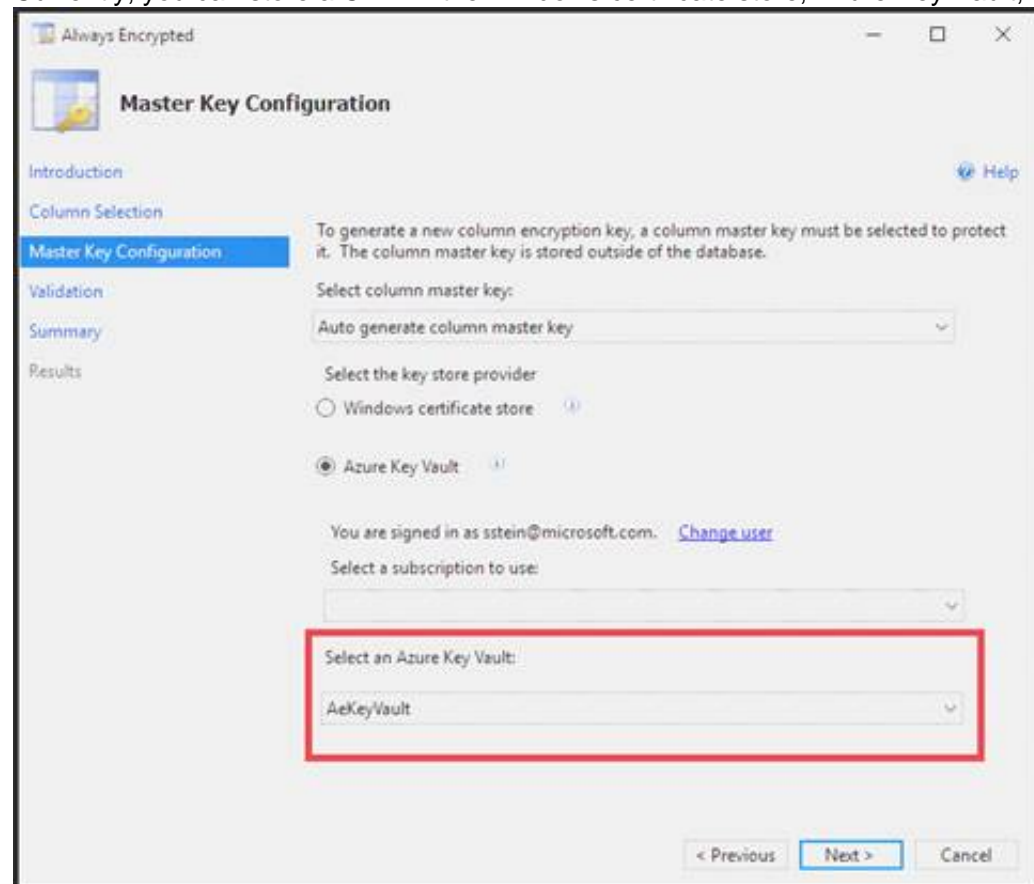
- A. Yes
- B. No

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 8

- (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 9

- (Exam Topic 3)

Each day, company plans to store hundreds of files in Azure Blob Storage and Azure Data Lake Storage. The company uses the parquet format.

You must develop a pipeline that meets the following requirements:

- ▶ Process data every six hours
- ▶ Offer interactive data analysis capabilities
- ▶ Offer the ability to process data using solid-state drive (SSD) caching
- ▶ Use Directed Acyclic Graph(DAG) processing mechanisms
- ▶ Provide support for REST API calls to monitor processes
- ▶ Provide native support for Python
- ▶ Integrate with Microsoft Power BI

You need to select the appropriate data technology to implement the pipeline. Which data technology should you implement?

- A. Azure SQL Data Warehouse
- B. HDInsight Apache Storm cluster
- C. Azure Stream Analytics
- D. HDInsight Apache Hadoop cluster using MapReduce

E. HDInsight Spark cluster

Answer: B

Explanation:

Storm runs topologies instead of the Apache Hadoop MapReduce jobs that you might be familiar with. Storm topologies are composed of multiple components that are arranged in a directed acyclic graph (DAG). Data flows between the components in the graph. Each component consumes one or more data streams, and can optionally emit one or more streams.

Python can be used to develop Storm components. References:

<https://docs.microsoft.com/en-us/azure/hdinsight/storm/apache-storm-overview>

NEW QUESTION 10

- (Exam Topic 3)

You manage a solution that uses Azure HDInsight clusters.

You need to implement a solution to monitor cluster performance and status. Which technology should you use?

- A. Azure HDInsight .NET SDK
- B. Azure HDInsight REST API
- C. Ambari REST API
- D. Azure Log Analytics
- E. Ambari Web UI

Answer: E

Explanation:

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.

The Apache Ambari project is aimed at making Hadoop management simpler by developing software for provisioning, managing, and monitoring Apache Hadoop clusters. Ambari provides an intuitive, easy-to-use Hadoop management web UI backed by its RESTful APIs.

References:

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

NEW QUESTION 10

- (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

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

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Actions

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

Answer Area

Enable advanced threat protection.

Configure the service to send email alerts to security@contoso.com

Configure the service to create alerts for threat detections of type **Data Exfiltration**.

NEW QUESTION 12

- (Exam Topic 3)

You are developing the data platform for a global retail company. The company operates during normal working hours in each region. The analytical database is used once a week for building sales projections.

Each region maintains its own private virtual network.

Building the sales projections is very resource intensive and generates upwards of 20 terabytes (TB) of data. Microsoft Azure SQL Databases must be provisioned.

- Database provisioning must maximize performance and minimize cost
 - The daily sales for each region must be stored in an Azure SQL Database instance
 - Once a day, the data for all regions must be loaded in an analytical Azure SQL Database instance
- You need to provision Azure SQL database instances.

How should you provision the database instances? To answer, drag the appropriate Azure SQL products to the correct databases. Each Azure SQL product 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.

Azure SQL products	Database	Azure SQL product
Azure SQL Database elastic pools	Daily Sales	Azure SQL product
Azure SQL Database Premium	Weekly Analysis	Azure SQL product
Azure SQL Database Managed Instance		
Azure SQL Database Hyperscale		

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: Azure SQL Database elastic pools

SQL Database elastic pools are a simple, cost-effective solution for managing and scaling multiple databases that have varying and unpredictable usage demands. The databases in an elastic pool are on a single Azure

SQL Database server and share a set number of resources at a set price. Elastic pools in Azure SQL Database enable SaaS developers to optimize the price performance for a group of databases within a prescribed budget while delivering performance elasticity for each database.

Box 2: Azure SQL Database Hyperscale

A Hyperscale database is an Azure SQL database in the Hyperscale service tier that is backed by the Hyperscale scale-out storage technology. A Hyperscale database supports up to 100 TB of data and provides high throughput and performance, as well as rapid scaling to adapt to the workload requirements. Scaling is transparent to the application – connectivity, query processing, and so on, work like any other SQL database.

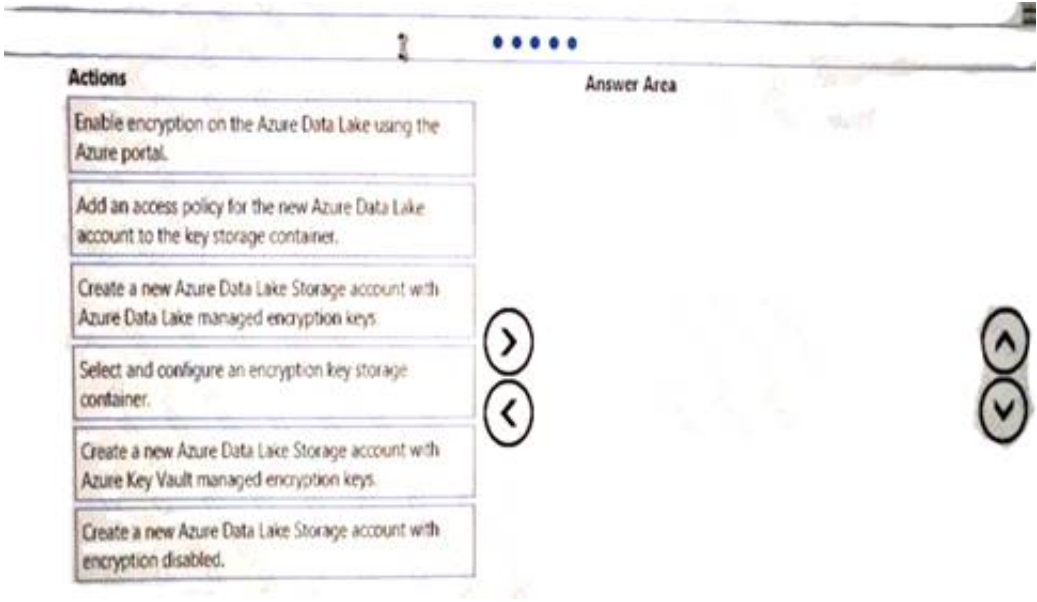
NEW QUESTION 17

- (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 interface with a list of actions on the left and an answer area on the right. The actions are:

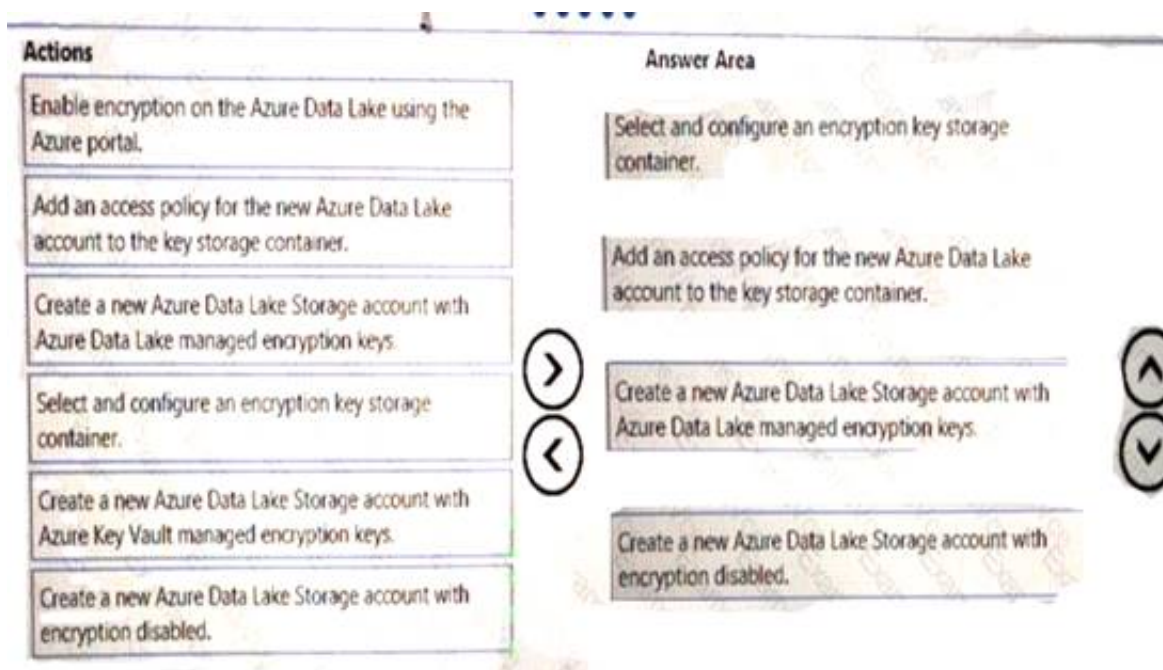
- Enable encryption on the Azure Data Lake using the Azure portal.
- Add an access policy for the new Azure Data Lake account to the key storage container.
- Create a new Azure Data Lake Storage account with Azure Data Lake managed encryption keys.
- Select and configure an encryption key storage container.
- 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 empty, and there are arrows indicating that actions can be moved from the list to the answer area.

- A. Mastered
B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 22

- (Exam Topic 3)

Your company uses several Azure HDInsight clusters.

The data engineering team reports several errors with some application using these clusters. You need to recommend a solution to review the health of the clusters.

What should you include in you recommendation?

- A. Azure Automation
- B. Log Analytics
- C. Application Insights

Answer: C

NEW QUESTION 25

- (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 30

- (Exam Topic 3)

A company plans to use Azure SQL Database to support a mission-critical application.

The application must be highly available without performance degradation during maintenance windows. You need to implement the solution.

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

NOTE: Each correct selection is worth one point.

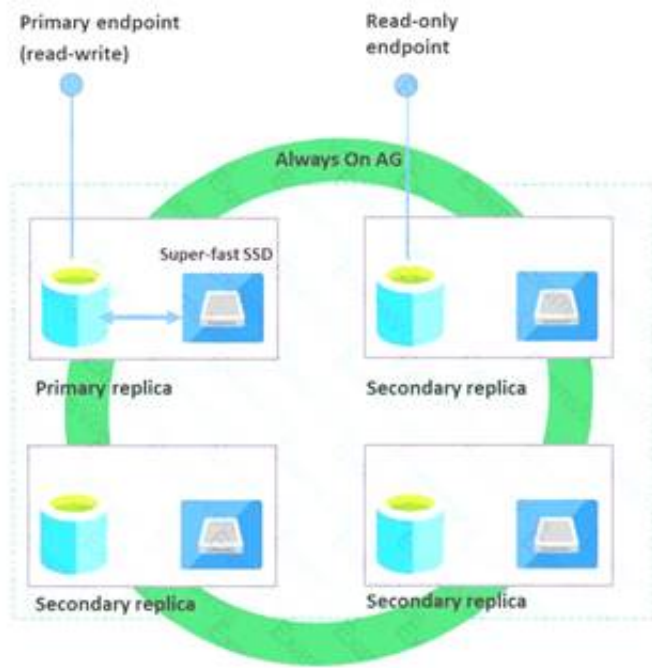
- A. Premium service tier
- B. Virtual machine Scale Sets
- C. Basic service tier
- D. SQL Data Sync
- E. Always On availability groups
- F. Zone-redundant configuration

Answer: AEF

Explanation:

Premium/business critical service tier model that is based on a cluster of database engine processes. This architectural model relies on a fact that there is always a quorum of available database engine nodes and has minimal performance impact on your workload even during maintenance activities.

In the premium model, Azure SQL database integrates compute and storage on the single node. High availability in this architectural model is achieved by replication of compute (SQL Server Database Engine process) and storage (locally attached SSD) deployed in 4-node cluster, using technology similar to SQL Server Always On Availability Groups.



Business Critical service tier: collocated compute and storage

Zone redundant configuration

By default, the quorum-set replicas for the local storage configurations are created in the same datacenter. With the introduction of Azure Availability Zones, you have the ability to place the different replicas in the quorum-sets to different availability zones in the same region. To eliminate a single point of failure, the control ring is also duplicated across multiple zones as three gateway rings (GW).

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-high-availability>

NEW QUESTION 35

- (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.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> Azure Data Factory
Use a web browser to run interactive queries and batch processes.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> Azure PowerShell
Develop batch processing applications that use Azure HDInsight.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> NoSQL database

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Scenario	Tool
Use a native client application to run interactive queries and batch processes.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input checked="" type="checkbox"/> HDInsight REST API <input type="checkbox"/> Azure Data Factory
Use a web browser to run interactive queries and batch processes.	<input checked="" type="checkbox"/> HDInsight Tools for Visual Studio <input type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> Azure PowerShell
Develop batch processing applications that use Azure HDInsight.	<input type="checkbox"/> HDInsight Tools for Visual Studio <input checked="" type="checkbox"/> Hive View <input type="checkbox"/> HDInsight REST API <input type="checkbox"/> NoSQL database

NEW QUESTION 39

- (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 43

- (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 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":{
      
    },
    "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 48

- (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 50

- (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 54

- (Exam Topic 3)

Your company uses Microsoft Azure SQL Database configure 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.

Components

Control Database

Azure Service Bus

Azure Storage

Job Service

Answer Area

Task

Execution results and diagnostics

Job launcher and tracker

Job metadata and state

Component

Component

Component

Component

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Components

Control Database

Azure Service Bus

Azure Storage

Job Service

Answer Area

Task

Execution results and diagnostics

Job launcher and tracker

Job metadata and state

Component

Azure Service Bus

Job Service

Control Database

NEW QUESTION 56

- (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

Extract SQL data on-premises

Technology

Self-hosted integration runtime

Azure-SSIS integration runtime

Azure integration runtime

Source

Action

Load SQL data warehouse

Technology

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 59

- (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 64

- (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 67

- (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: Download Azure HDInsight cluster logs by using Azure PowerShell.
Does the solution meet the goal?

A. Yes
B. No

Answer: B

Explanation:

Reference:
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>

NEW QUESTION 69

- (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.

...

Actions

Answer Area

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.

>

<

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

...

Actions

Answer Area

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.

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

<

NEW QUESTION 70

- (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 tile 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 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. 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 73

- (Exam Topic 3)

You manage a Microsoft Azure SQL Data Warehouse Gen 2.

Users report slow performance when they run commonly used queries. Users do not report performance changes for infrequently used queries. You need to monitor resource utilization to determine the source of the performance issues. Which metric should you monitor?

- A. Cache used percentage
- B. Local tempdb percentage
- C. WU percentage
- D. CPU percentage

Answer: B

NEW QUESTION 75

- (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><tr><td>DataNode</td><td><input checked="" type="checkbox"/></td></tr><tr><td>NameNode</td><td><input type="checkbox"/></td></tr></table>	DataNode	<input checked="" type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input checked="" type="checkbox"/>				
NameNode	<input type="checkbox"/>				
Run operations on files and directories of the file system.	<table><tr><td>DataNode</td><td><input checked="" type="checkbox"/></td></tr><tr><td>NameNode</td><td><input type="checkbox"/></td></tr></table>	DataNode	<input checked="" type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input checked="" type="checkbox"/>				
NameNode	<input type="checkbox"/>				
Perform block creation, deletion, and replication.	<table><tr><td>DataNode</td><td><input checked="" type="checkbox"/></td></tr><tr><td>NameNode</td><td><input type="checkbox"/></td></tr></table>	DataNode	<input checked="" type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input checked="" 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 80
.....

Relate Links

100% Pass Your DP-200 Exam with ExamBible Prep Materials

<https://www.exambible.com/DP-200-exam/>

Contact us

We are proud of our high-quality customer service, which serves you around the clock 24/7.

Viste - <https://www.exambible.com/>