

Exam Questions DP-200

Implementing an Azure Data Solution

<https://www.2passeasy.com/dumps/DP-200/>



NEW QUESTION 1

- (Exam Topic 1)

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

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

Answer: C

Explanation:

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

Scenario:

All data migration processes must use Azure Data Factory

All data migrations must run automatically during non-business hours

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

NEW QUESTION 2

- (Exam Topic 1)

You need to ensure polling data security requirements are met.

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

NOTE: Each correct selection is worth one point.

Context	Security technology
SQL Server	Azure Active Directory user
	Domain Active Directory user
	Managed Identity
PolyBase	Database scoped credential
	Database encryption key
	Application role

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Azure Active Directory user

Scenario:

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

Box 2: DataBase Scoped Credential

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

PolyBase cannot authenticate by using Azure AD authentication. References:

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

NEW QUESTION 3

- (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	FileCount
	BlobCapacity
	FileCapacity
Aggregation	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 4

- (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 5

- (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:

- 1. Create a Kafka enabled Event Hubs namespace.
- 2. Create a Kafka client that sends messages to the event hub.
- 3. Create a Stream Analytics job that copies data from the event hub into an Azure blob storage.

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 6

- (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 specifies the information needed to connect to Blob Storage.

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

Box 2: Azure Blob

Tier 10 reporting data must be stored in Azure Blobs

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

References:

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

NEW QUESTION 7

- (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

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 the 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

Answer Area

- 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 8

- (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> <tr><td>custom text</td><td rowspan="4">V</td></tr> <tr><td>default</td></tr> <tr><td>email</td></tr> <tr><td>random number</td></tr> </table>	custom text	V	default	email	random number
custom text	V					
default						
email						
random number						
B	<table> <tr><td>custom text</td><td rowspan="4">V</td></tr> <tr><td>default</td></tr> <tr><td>email</td></tr> <tr><td>random number</td></tr> </table>	custom text	V	default	email	random number
custom text	V					
default						
email						
random number						
C	<table> <tr><td>custom text</td><td rowspan="4">V</td></tr> <tr><td>default</td></tr> <tr><td>email</td></tr> <tr><td>random number</td></tr> </table>	custom text	V	default	email	random number
custom text	V					
default						
email						
random number						

- 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 9

- (Exam Topic 3)

Contoso, Ltd. plans to configure existing applications to use Azure SQL Database. When security-related operations occur, the security team must be informed.

You need to configure Azure Monitor while minimizing administrative efforts

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

- A. Create a new action group to email alerts@contoso.com.
 B. Use alerts@contoso.com as an alert email address.
 C. Use all security operations as a condition.
 D. Use all Azure SQL Database servers as a resource.
 E. Query audit log entries as a condition.

Answer: ACE

NEW QUESTION 10

- (Exam Topic 3)

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

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

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

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

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

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

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

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

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

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

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

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

References:

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

NEW QUESTION 10

- (Exam Topic 3)

You develop data engineering solutions for a company.

A project requires an in-memory batch data processing solution.

You need to provision an HDInsight cluster for batch processing of data on Microsoft Azure.

How should you complete the PowerShell segment? To answer, select the appropriate option in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

New-AzureStorageContainer
New-AzureRmHDInsightClusterConfig
New-AzureRmHDInsightCluster

```
-Name $clusterName -Context $defaultStorageContext
$objectConfig = New-Object "System.Collections.Generic.D
$objectConfig.Add (
(
"spark"
"hadoop"
-ResourceGroupName $resourceGroupName `
-ClusterName $clusterName `
-Location $location `
-ClusterSizeInNodes $clusterSizeInNodes `
-ClusterType "spark"
-OSType $clusterOS `
-Version $clusterVersion `
-ComponentVersion $objectConfig
```

New-AzureRmHDInsightCluster
New-AzureRmHDInsightClusterConfig
New-AzureStorageContainer

"spark"
"hadoop"
"HBase"
"Storm"

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

New-AzureStorageContainer
New-AzureRmHDInsightClusterConfig
New-AzureRmHDInsightCluster

```
-Name $clusterName -Context $defaultStorageContext
$objectConfig = New-Object "System.Collections.Generic.D
$objectConfig.Add (
(
"spark"
"hadoop"
-ResourceGroupName $resourceGroupName `
-ClusterName $clusterName `
-Location $location `
-ClusterSizeInNodes $clusterSizeInNodes `
-ClusterType "spark"
-OSType $clusterOS `
-Version $clusterVersion `
-ComponentVersion $objectConfig
```

New-AzureRmHDInsightCluster
New-AzureRmHDInsightClusterConfig
New-AzureStorageContainer

"spark"
"hadoop"
"HBase"
"Storm"

NEW QUESTION 14

- (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 16

- (Exam Topic 3)

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

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

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: Create a master key

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

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

Example code: USE master; GO

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

go

CREATE CERTIFICATE MyServerCert WITH SUBJECT = 'My DEK Certificate'; go

USE AdventureWorks2012; GO

CREATE DATABASE ENCRYPTION KEY WITH ALGORITHM = AES_128

ENCRYPTION BY SERVER CERTIFICATE MyServerCert; GO

ALTER DATABASE AdventureWorks2012 SET ENCRYPTION ON;

GO

References:

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

NEW QUESTION 18

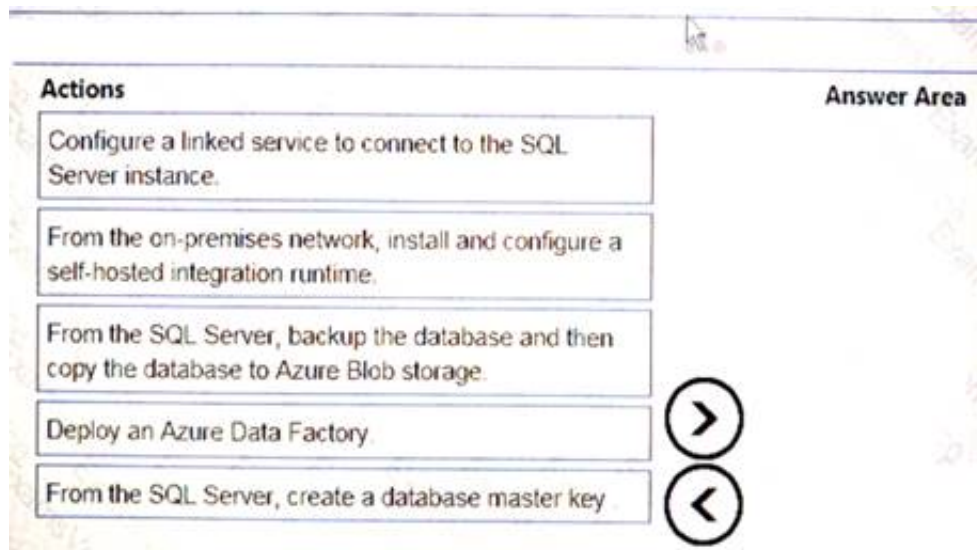
- (Exam Topic 3)

Your company has on-premises Microsoft SQL Server instance.

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

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

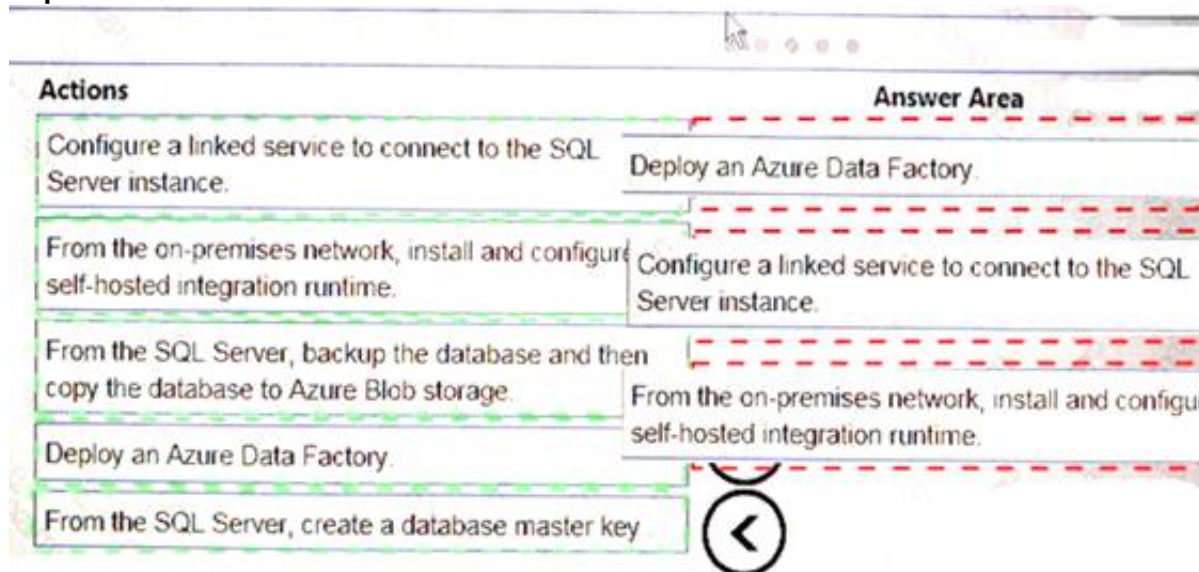
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



- A. Mastered
- B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 22

- (Exam Topic 3)

A company runs Microsoft SQL Server in an on-premises virtual machine (VM). You must migrate the database to Azure SQL Database. You synchronize users from Active Directory to Azure Active Directory (Azure AD). You need to configure Azure SQL Database to use an Azure AD user as administrator. What should you configure?

- A. For each Azure SQL Database, set the Access Control to administrator.
- B. For the Azure SQL Database server, set the Active Directory to administrator.
- C. For each Azure SQL Database, set the Active Directory administrator role.
- D. For the Azure SQL Database server, set the Access Control to administrator.

Answer: A

NEW QUESTION 24

- (Exam Topic 3)

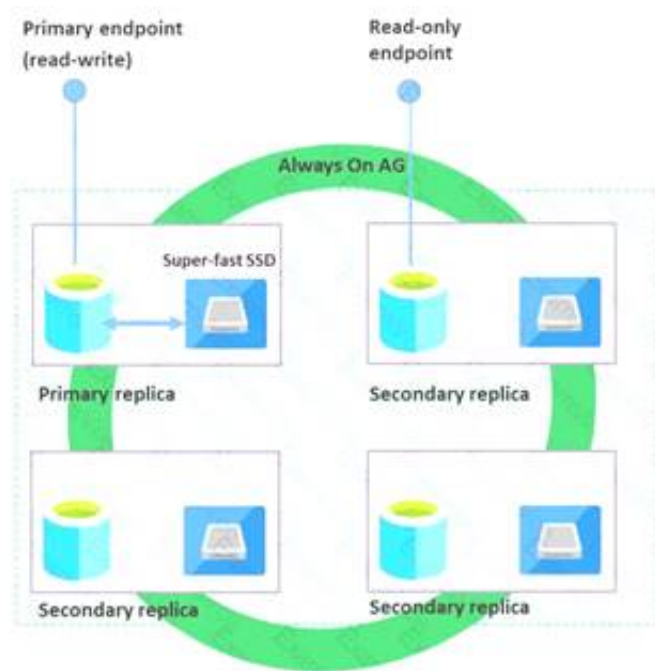
A company plans to use Azure SQL Database to support a mission-critical application. The application must be highly available without performance degradation during maintenance windows. You need to implement the solution. Which three technologies should you implement? Each correct answer presents part 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 29

- (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 port.

- 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 30

- (Exam Topic 3)

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

You need to create the offline export copy.

Which three options can you use? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

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

Answer: BCE

NEW QUESTION 34

- (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 36

- (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 38

- (Exam Topic 3)

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

You need to create the input stream.

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

NOTE: Each correct selection is worth one point.

Answer Area

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

Answer Area

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

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

NEW QUESTION 43

- (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 44

- (Exam Topic 3)

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

You develop data engineering solutions for a company.

A project requires the deployment of resources to Microsoft Azure for batch data processing on Azure

HDInsight. Batch processing will run daily and must: Scale to minimize costs

Be monitored for cluster performance

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

Does the solution meet the goal?

- A. Yes
 B. No

Answer: B

Explanation:

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

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

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

NEW QUESTION 45

- (Exam Topic 3)

A company is designing a hybrid solution to synchronize data and on-premises Microsoft SQL Server database to Azure SQL Database.

You must perform an assessment of databases to determine whether data will move without compatibility issues.

You need to perform the assessment. Which tool should you use?

- A. Azure SQL Data Sync
 B. SQL Vulnerability Assessment (VA)
 C. SQL Server Migration Assistant (SSMA)
 D. Microsoft Assessment and Planning Toolkit
 E. Data Migration Assistant (DMA)

Answer: E

Explanation:

The Data Migration Assistant (DMA) helps you upgrade to a modern data platform by detecting compatibility issues that can impact database functionality in your new version of SQL Server or Azure SQL Database. DMA recommends performance and reliability improvements for your target environment and allows you to move your schema, data, and uncontained objects from your source server to your target server.

References:

<https://docs.microsoft.com/en-us/sql/dma/dma-overview>

NEW QUESTION 49

- (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 51

- (Exam Topic 3)

You plan to create a new single database instance of Microsoft Azure SQL Database.

The database must only allow communication from the data engineer's workstation. You must connect directly to the instance by using Microsoft SQL Server Management Studio.

You need to create and configure the Database. Which three Azure PowerShell cmdlets should you use to develop the solution? To answer, move the appropriate cmdlets from the list of cmdlets to the answer area and arrange them in the correct order.

Azure PowerShell cmdlets	Answer Area
New-AzureRmSqlElasticPool	
New-AzureRmSqlServerFirewallRule	
New-AzureRmSqlServer	
New-AzureRmSqlServerVirtualNetworkRule	
New-AzureRmSqlDatabase	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: New-AzureSqlServer Create a server.

Step 2: New-AzureRmSqlServerFirewallRule

New-AzureRmSqlServerFirewallRule creates a firewall rule for a SQL Database server. Can be used to create a server firewall rule that allows access from the specified IP range. Step 3: New-AzureRmSqlDatabase

Example: Create a database on a specified server

PS C:\>New-AzureRmSqlDatabase -ResourceGroupName "ResourceGroup01" -ServerName "Server01"

-DatabaseName "Database01"

References:

<https://docs.microsoft.com/en-us/azure/sql-database/scripts/sql-database-create-and-configure-database-powershell>

NEW QUESTION 54

- (Exam Topic 3)

A company builds an application to allow developers to share and compare code. The conversations, code snippets, and links shared by people in the application are stored in a Microsoft Azure SQL Database instance. The application allows for searches of historical conversations and code snippets.

When users share code snippets, the code snippet is compared against previously share code snippets by using a combination of Transact-SQL functions including SUBSTRING, FIRST_VALUE, and SQRT. If a match is found, a link to the match is added to the conversation.

Customers report the following issues:

- Delays occur during live conversations
- A delay occurs before matching links appear after code snippets are added to conversations

You need to resolve the performance issues.

Which technologies should you use? To answer, drag the appropriate technologies to the correct issues. Each technology 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.

Technologies	Answer Area	
	Issue	Technology
columnstore index	Delays in conversations	
non-durable table	Delays in match links	
materialized view		
memory-optimized table		

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: memory-optimized table

In-Memory OLTP can provide great performance benefits for transaction processing, data ingestion, and transient data scenarios.

Box 2: materialized view

To support efficient querying, a common solution is to generate, in advance, a view that materializes the data in a format suited to the required results set. The Materialized View pattern describes generating prepopulated views of data in environments where the source data isn't in a suitable format for querying, where generating a suitable query is difficult, or where query performance is poor due to the nature of the data or the data store.

These materialized views, which only contain data required by a query, allow applications to quickly obtain the information they need. In addition to joining tables or combining data entities, materialized views can include the current values of calculated columns or data items, the results of combining values or executing transformations on the data items, and values specified as part of the query. A materialized view can even be optimized for just a single query.

References:

<https://docs.microsoft.com/en-us/azure/architecture/patterns/materialized-view>

NEW QUESTION 57

- (Exam Topic 3)

You develop data engineering solutions for a company.

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

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

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

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

Answer: ADE

NEW QUESTION 60

- (Exam Topic 3)

You manage a process that performs analysis of daily web traffic logs on an HDInsight cluster. Each of 250 web servers generates approximately gigabytes (GB) of log data each day. All log data is stored in a single folder in Microsoft Azure Data Lake Storage Gen 2.

You need to improve the performance of the process.

Which two changes should you make? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point.

- A. Combine the daily log files for all servers into one file
- B. Increase the value of the mapreduce.map.memory parameter
- C. Move the log files into folders so that each day's logs are in their own folder
- D. Increase the number of worker nodes
- E. Increase the value of the hive.tez.container.size parameter

Answer: AC

Explanation:

A: Typically, analytics engines such as HDInsight and Azure Data Lake Analytics have a per-file overhead. If you store your data as many small files, this can negatively affect performance. In general, organize your data into larger sized files for better performance (256MB to 100GB in size). Some engines and applications might have trouble efficiently processing files that are greater than 100GB in size.

C: For Hive workloads, partition pruning of time-series data can help some queries read only a subset of the data which improves performance.

Those pipelines that ingest time-series data, often place their files with a very structured naming for files and folders. Below is a very common example we see for data that is structured by date:

\DataSet\YYYY\MM\DD\datafile_YYYY_MM_DD.tsv

Notice that the datetime information appears both as folders and in the filename. References:

<https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-performance-tuning-guidance>

NEW QUESTION 62

- (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 63

- (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 64

- (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 65

- (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

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.

Create a master key on database

Create the external table FactSalesOrderDetails.

- A. Mastered
 B. Not Mastered

Answer: A

Explanation:

NEW QUESTION 66

- (Exam Topic 3)

You are the data engineer for your company. An application uses a NoSQL database to store data. The database uses the key-value and wide-column NoSQL database type.

Developers need to access data in the database using an API.

You need to determine which API to use for the database model and type.

Which two APIs should you use? Each correct answer presents a complete solution. NOTE: Each correct selection s worth one point.

- A. Table API
 B. MongoDB API
 C. Gremlin API
 D. SQL API
 E. Cassandra API

Answer: BE

Explanation:

B: Azure Cosmos DB is the globally distributed, multimodel database service from Microsoft for mission-critical applications. It is a multimodel database and supports document, key-value, graph, and columnar data models.

E: Wide-column stores store data together as columns instead of rows and are optimized for queries over large datasets. The most popular are Cassandra and HBase.

References:

<https://docs.microsoft.com/en-us/azure/cosmos-db/graph-introduction> <https://www.mongodb.com/scale/types-of-nosql-databases>

NEW QUESTION 69

- (Exam Topic 3)

You implement 3 Azure SQL Data Warehouse instance.

You plan to migrate the largest fact table to Azure SQL Data Warehouse The table resides on Microsoft SQL Server on-premises and e 10 terabytes (TB) in size.

Incoming queues use the primary key Sale Key column to retrieve data as displayed in the following table:

SaleKey	CityKey	CustomerKey	StockItemKey	InvoiceDateKey	Quantity	UnitPrice	TotalExcludingTax
49309	90858	70	89	10/22/13	8	16	128
49313	55710	126	89	10/22/13	2	16	32
49343	44710	234	88	10/22/13	10	16	160
49352	66109	183	70	10/22/13	4	16	64
49448	65312	230	70	10/22/13	8	16	128
49646	85877	271	70	10/24/13	1	16	16
49798	41238	288	89	10/24/13	1	16	16

You need to distribute the fact table across multiple nodes to optimize performance of the table. Which technology should you use?

- A. hash distributed table with clustered ColumnStore index
- B. hash distributed table with clustered index
- C. heap table with distribution replicate
- D. round robin distributed table with clustered index
- E. round robin distributed table with clustered ColumnStore index

Answer: A

NEW QUESTION 74

- (Exam Topic 3)

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

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

NOTE: Each correct selection is worth one point.

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

Answer: AC

NEW QUESTION 78

- (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 83

- (Exam Topic 3)

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

Reports indicate slowdowns during a specific time window.

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

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

Answer: B

Explanation:

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

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

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

References:

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

NEW QUESTION 86

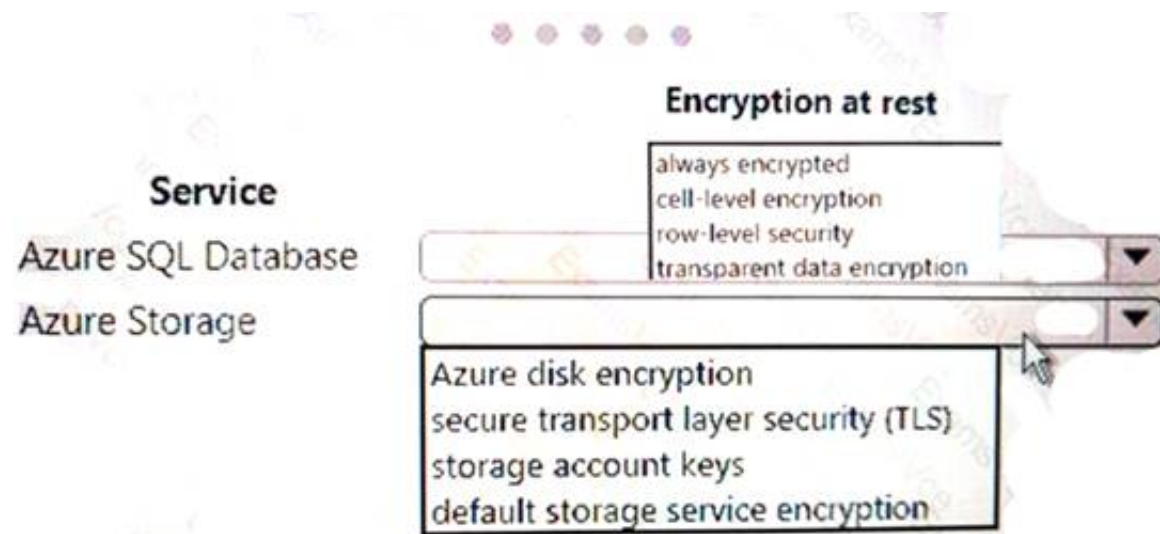
- (Exam Topic 3)

Your company uses Azure SQL Database and Azure Blob storage.

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

You need to configure security.

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



- A. Mastered
- B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 87

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual DP-200 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the DP-200 Product From:

<https://www.2passeasy.com/dumps/DP-200/>

Money Back Guarantee

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