



## **Microsoft**

### **Exam Questions DP-200**

Implementing an Azure Data Solution

## 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 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 2

- (Exam Topic 2)

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

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

You need to configure data encryption for external applications. Solution:

1. Access the Always Encrypted Wizard in SQL Server Management Studio
2. Select the column to be encrypted
3. Set the encryption type to 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?

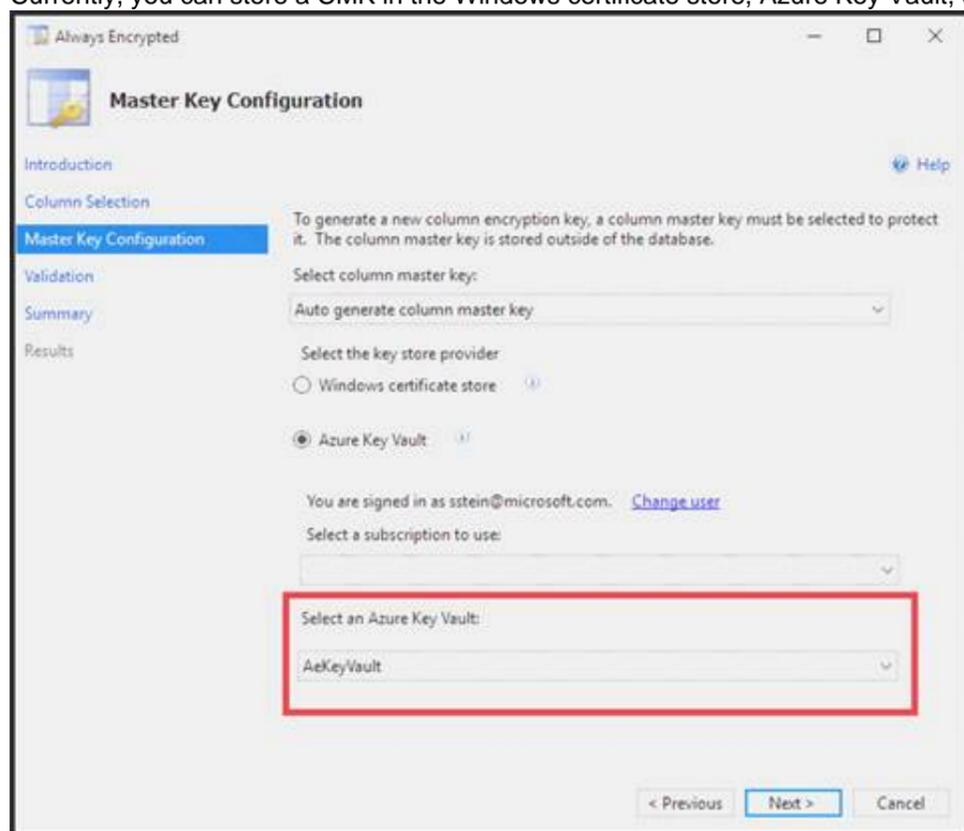
- A. Yes
- B. No

**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 3

- (Exam Topic 2)

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

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

You need to configure data encryption for external applications.

Solution:

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

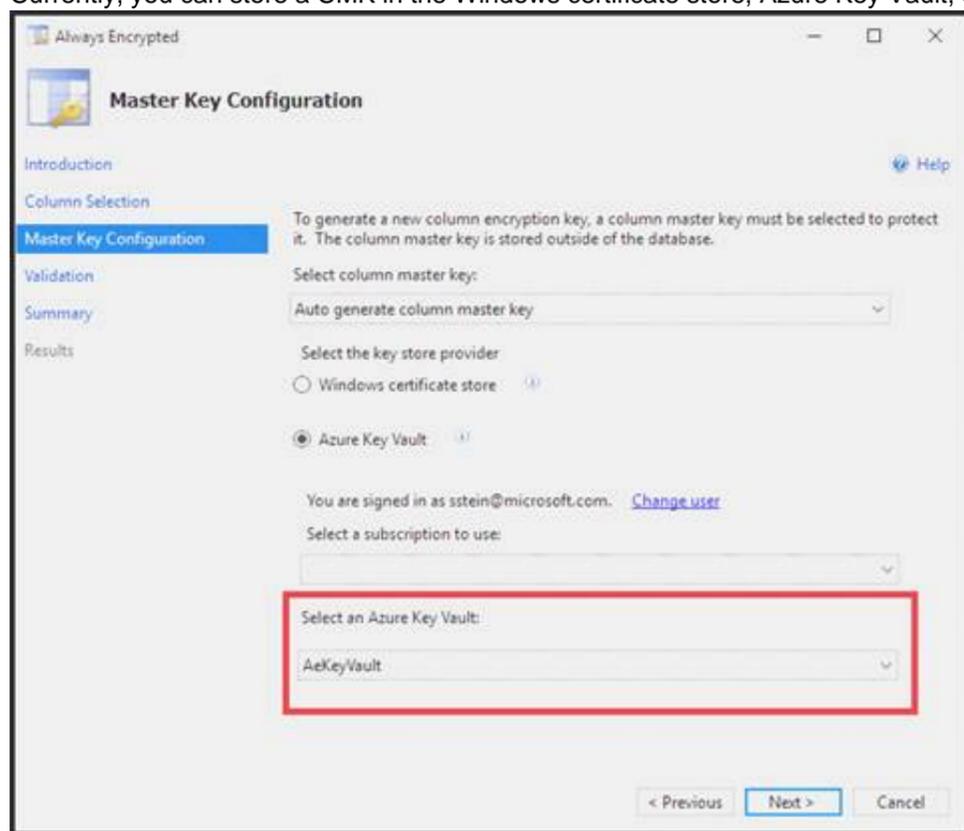
- A. Yes
- B. No

**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 4**

- (Exam Topic 2)

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

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

**Answer: A**

**Explanation:**

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

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

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

References:

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

**NEW QUESTION 5**

- (Exam Topic 2)

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

NOTE: Each correct selection is worth one point.

Data type	Masking function					
A	<table border="1"> <tr><td>custom text</td><td 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 border="1"> <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 border="1"> <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 6**

- (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 7**

- (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 (
(
-ResourceGroupName $resourceGroupName
-ClusterName $clusterName
-Location $location
-ClusterSizeInNodes $clusterSizeInNodes
-ClusterType
-OSType $clusterOS
-Version $clusterVersion
-ComponentVersion $objectConfig

```

Dropdown menu options for the first Add method call: "spark", "hadoop"

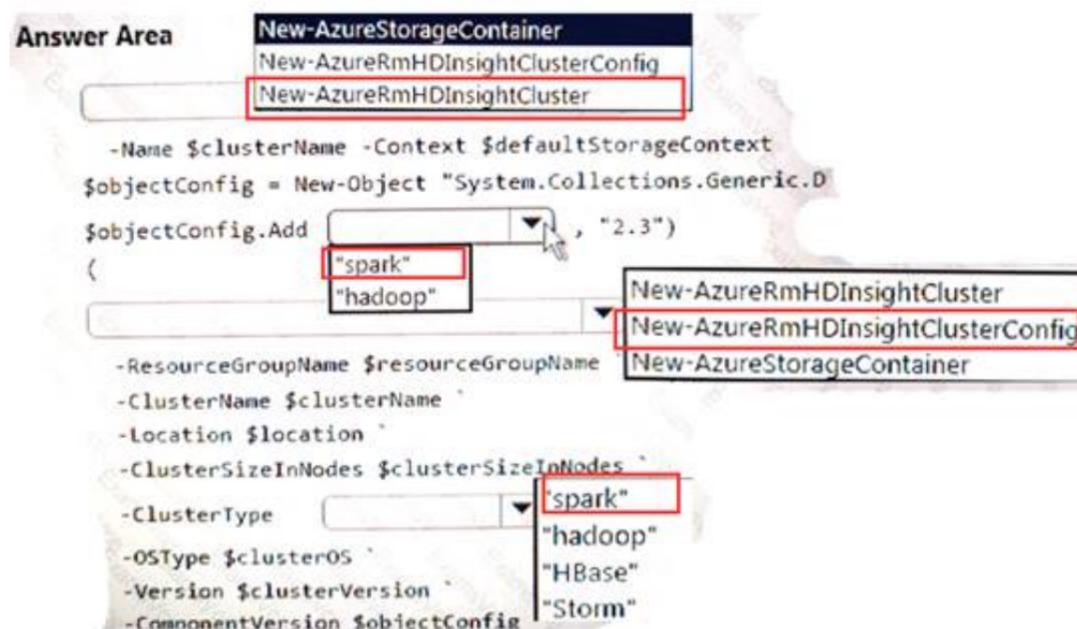
Dropdown menu options for the ClusterType parameter: "spark", "hadoop", "HBase", "Storm"

Dropdown menu options for the ComponentVersion parameter: New-AzureRmHDInsightCluster, New-AzureRmHDInsightClusterConfig, New-AzureStorageContainer

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:



### NEW QUESTION 8

- (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 9

- (Exam Topic 3)

You plan to use Microsoft Azure SQL Database instances with strict user access control. A user object must:

- Move with the database if it is run elsewhere
- Be able to create additional users

You need to create the user object with correct permissions.

Which two Transact-SQL commands should you run? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. ALTER LOGIN Mary WITH PASSWORD = 'strong\_password';
- B. CREATE LOGIN Mary WITH PASSWORD = 'strong\_password';
- C. ALTER ROLE db\_owner ADD MEMBER Mary;
- D. CREATE USER Mary WITH PASSWORD = 'strong\_password';
- E. GRANT ALTER ANY USER TO Mary;

Answer: CD

Explanation:

C: ALTER ROLE adds or removes members to or from a database role, or changes the name of a user-defined database role.

Members of the db\_owner fixed database role can perform all configuration and maintenance activities on the database, and can also drop the database in SQL Server.

D: CREATE USER adds a user to the current database.

Note: Logins are created at the server level, while users are created at the database level. In other words, a login allows you to connect to the SQL Server service (also called an instance), and permissions inside the database are granted to the database users, not the logins. The logins will be assigned to server roles (for example, serveradmin) and the database users will be assigned to roles within that database (eg. db\_datareader, db\_backupoperator).

References:

<https://docs.microsoft.com/en-us/sql/t-sql/statements/alter-role-transact-sql> <https://docs.microsoft.com/en-us/sql/t-sql/statements/create-user-transact-sql>

### NEW QUESTION 10

- (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 10**

- (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 14

- (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 your recommendation?

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

**Answer: C**

#### NEW QUESTION 19

- (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 24

- (Exam Topic 3)

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

The solution has the following requirements:

- Data must be partitioned into multiple containers.
- Data containers must be configured separately.
- Data must be accessible from applications hosted around the world.
- The solution must minimize latency. You need to provision Azure Cosmos DB

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

**Answer: A**

#### NEW QUESTION 29

- (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 32**

- (Exam Topic 3)

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

Which components should you use for each Elastic pool job services task? To answer, drag the appropriate component to the correct task. Each component may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content. NOTE: Each correct selection is worth one point.

- A. Mastered
- B. Not Mastered

**Answer: A**

**Explanation:**

**NEW QUESTION 34**

- (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.

- 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-powersh
```

### NEW QUESTION 35

- (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 36

- (Exam Topic 3)

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

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

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

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

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

**Answer:** E

#### Explanation:

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

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

Compare Data Sync with Transactional Replication

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

References:

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

### NEW QUESTION 38

- (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 40**

- (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 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. Use Azure Data Factory to convert the parquet files to CSV files
2. Create an external data source pointing to the Azure storage account
3. Create an external file format and external table using the external data source
4. Load the data using the INSERT...SELECT statement Does the solution meet the goal?

- A. Yes
- B. No

**Answer:** B

**Explanation:**

There is no need to convert the parquet files to CSV files.

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>

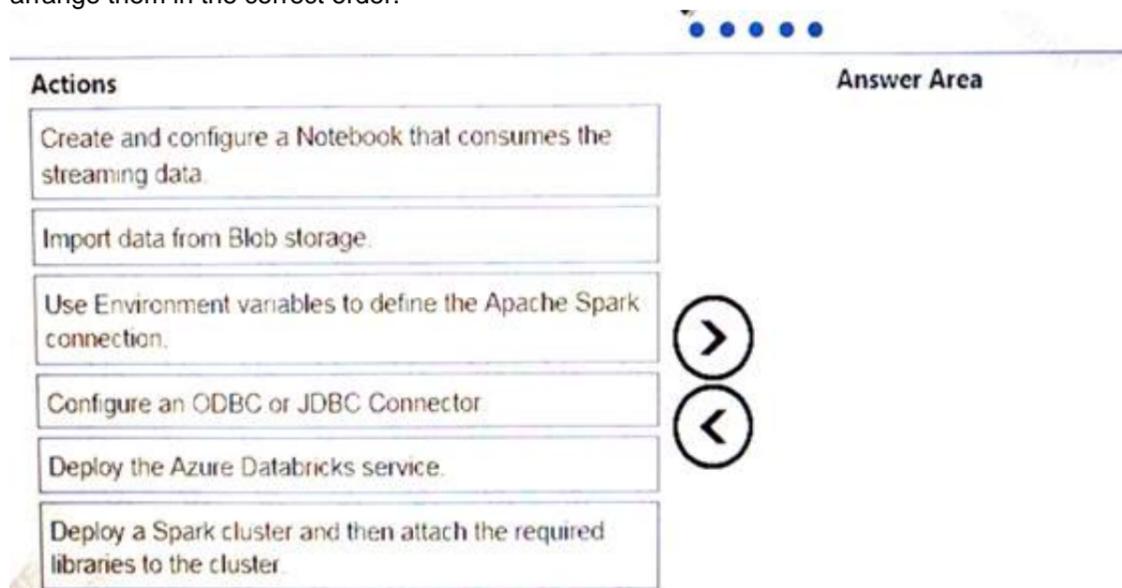
**NEW QUESTION 43**

- (Exam Topic 3)

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

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

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



The screenshot shows an exam question interface. At the top, there are five blue dots, with the first one filled, indicating the current question. Below this, there are two main sections: 'Actions' on the left and 'Answer Area' on the right. The 'Actions' section contains a list of six actions in rectangular boxes:

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

To the right of the 'Actions' list, there are two circular arrows: a right-pointing arrow (>) above a left-pointing arrow (<). The 'Answer Area' is currently empty.

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

**NEW QUESTION 48**

- (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	69	10/22/13	2	16	32
49343	44710	234	88	10/22/13	10	16	160
49352	66109	163	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 53**

- (Exam Topic 3)

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

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

NOTE: Each correct selection is worth one point.

Requirement	Node				
Provide data access to clients.	<table border="1"> <tr> <td>DataNode</td> <td><input type="checkbox"/></td> </tr> <tr> <td>NameNode</td> <td><input checked="" type="checkbox"/></td> </tr> </table>	DataNode	<input type="checkbox"/>	NameNode	<input checked="" type="checkbox"/>
DataNode	<input type="checkbox"/>				
NameNode	<input checked="" type="checkbox"/>				
Run operations on files and directories of the file system.	<table border="1"> <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 border="1"> <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](https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html#NameNode+and+DataNodes)

#### NEW QUESTION 54

- (Exam Topic 3)

You develop data engineering solutions for a company.

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

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

NOTE: Each correct selection is worth one point.

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

**Answer:** BDF

#### **Explanation:**

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

References:

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

#### NEW QUESTION 58

.....

## 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/>