



Microsoft

Exam Questions DP-200

Implementing an Azure Data Solution

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 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 3

- (Exam Topic 3)

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

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

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

NOTE: Each correct selection is worth one point.

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Eventual

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

The following image shows the different consistency levels as a spectrum.



Box 2: GlobalDocumentDB

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

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

References:

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

NEW QUESTION 4

- (Exam Topic 3)

You are designing a new Lambda architecture on Microsoft Azure. The real-time processing layer must meet the following requirements: Ingestion:

- Receive millions of events per second
- Act as a fully managed Platform-as-a-Service (PaaS) solution
- Integrate with Azure Functions

Stream processing:

- Process on a per-job basis
- Provide seamless connectivity with Azure services
- Use a SQL-based query language

Analytical data store:

- Act as a managed service
- Use a document store
- Provide data encryption at rest

You need to identify the correct technologies to build the Lambda architecture using minimal effort. Which technologies should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Architecture requirement	Answer Area					
Ingestion	<table border="1"> <tr> <td>HDInsight Kafka</td><td rowspan="4">▼</td></tr> <tr> <td>Azure Event Hubs</td></tr> <tr> <td>HDInsight Storm</td></tr> <tr> <td>HDInsight Spark</td></tr> </table>	HDInsight Kafka	▼	Azure Event Hubs	HDInsight Storm	HDInsight Spark
HDInsight Kafka	▼					
Azure Event Hubs						
HDInsight Storm						
HDInsight Spark						
Stream Processing	<table border="1"> <tr> <td>Azure Stream Analytics</td><td rowspan="4">▼</td></tr> <tr> <td>HDInsight with Spark Streaming</td></tr> <tr> <td>Azure Cosmos DB Change Feed</td></tr> <tr> <td>Azure Analysis Services</td></tr> </table>	Azure Stream Analytics	▼	HDInsight with Spark Streaming	Azure Cosmos DB Change Feed	Azure Analysis Services
Azure Stream Analytics	▼					
HDInsight with Spark Streaming						
Azure Cosmos DB Change Feed						
Azure Analysis Services						
Analytical Data Store	<table border="1"> <tr> <td>Hive LLAP on HDInsight</td><td rowspan="4">▼</td></tr> <tr> <td>Azure Analysis Services</td></tr> <tr> <td>Azure Cosmos DB</td></tr> <tr> <td>SQL Data Warehouse</td></tr> </table>	Hive LLAP on HDInsight	▼	Azure Analysis Services	Azure Cosmos DB	SQL Data Warehouse
Hive LLAP on HDInsight	▼					
Azure Analysis Services						
Azure Cosmos DB						
SQL Data Warehouse						

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Azure Event Hubs

This portion of a streaming architecture is often referred to as stream buffering. Options include Azure Event Hubs, Azure IoT Hub, and Kafka.

NEW QUESTION 5

- (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 6

- (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 7

- (Exam Topic 3)

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

You must implement a solution that meets the following requirements:

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

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

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

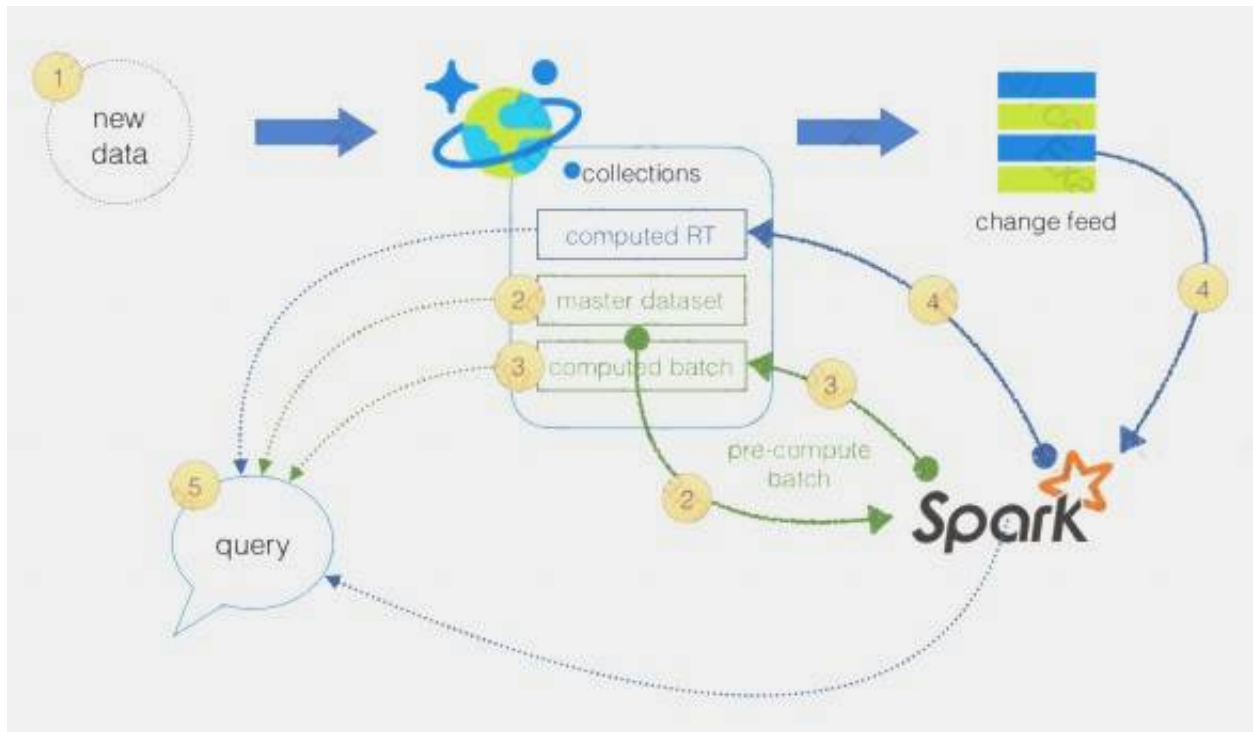
Answer: D

Explanation:

Lambda Architecture with Azure:

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

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



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

References:

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

NEW QUESTION 8

- (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":{
      "type": "CSV",
      "properties":{
        "fieldDelimiter": ",",
        "encoding": "UTF8"
      }
    },
    "datasource":{
      "type": "Microsoft.Storage/Blob",
      "properties":{
        "serviceBusNamespace": "sampleServiceBus",
        "sharedAccessPolicyName": "SampleReceiver",
        "sharedAccessPolicyKey": "<PolicyKey>",
        "eventHubName": "sampleEventHub"
      }
    },
    "compression":{
      "type": "GZip"
    }
  }
}
```

Answer Area

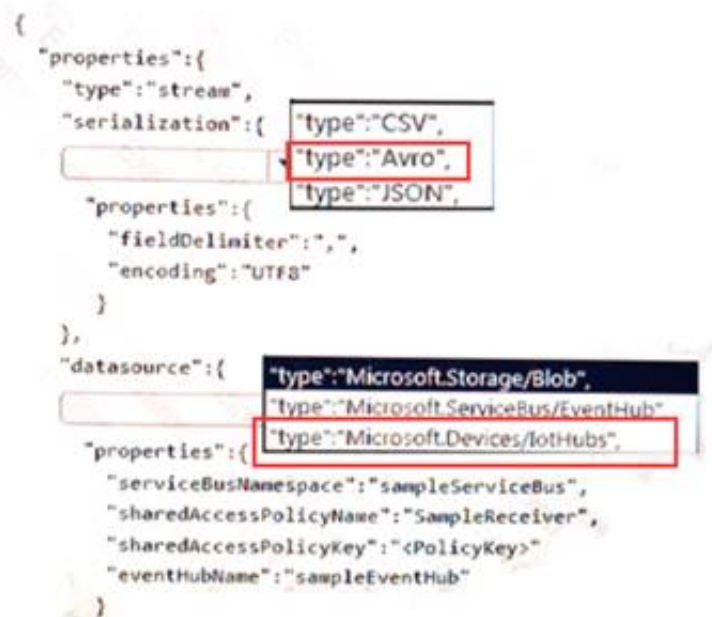
```
{
  "properties":{
    "type":"stream",
    "serialization":{
      "type": "CSV",
      "type": "Avro",
      "type": "JSON",
      "properties":{
        "fieldDelimiter": ",",
        "encoding": "UTF8"
      }
    },
    "datasource":{
      "type": "Microsoft.Storage/Blob",
      "type": "Microsoft.ServiceBus/EventHub",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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 9

- (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 10

- (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 10

- (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-powersh>

NEW QUESTION 13

- (Exam Topic 3)

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

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

You need to create the pipeline system.

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

NOTE: Each correct selection is worth one point.

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

- A. Mastered
 B. Not Mastered

Answer: A

Explanation:

Box 1: Source

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

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

Box 2: Self-hosted integration runtime

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

References:

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

NEW QUESTION 15

- (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 18

- (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 21

- (Exam Topic 3)

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

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

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

Solution: Use information stored in Azure Active Directory reports.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

NEW QUESTION 25

- (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 is 10 terabytes (TB) in size.

Incoming queries use the primary key SaleKey column to retrieve data as displayed in the following table:

SaleKey	CityKey	CustomerKey	StockItemKey	InvoiceDateKey	Quantity	UnitPrice	TotalExcludingTax
49309	90858	70	69	10/22/13	8	16	128
49313	55710	126	69	10/22/13	2	16	32
49343	44710	234	68	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	69	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 30

- (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.

Service

Azure SQL Database

Azure Storage

Encryption at rest

always encrypted

cell-level encryption

row-level security

transparent data encryption

Azure disk encryption

secure transport layer security (TLS)

storage account keys

default storage service encryption

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Service

Azure SQL Database

Azure Storage

Encryption at rest

always encrypted

cell-level encryption

row-level security

transparent data encryption

Azure disk encryption

secure transport layer security (TLS)

storage account keys

default storage service encryption

NEW QUESTION 35
.....

Thank You for Trying Our Product

We offer two products:

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

2nd - Questons and Answers in PDF Format

DP-200 Practice Exam Features:

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

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