

# 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 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 <input type="checkbox"/>
	Domain Active Directory user <input type="checkbox"/>
	Managed Identity <input type="checkbox"/>
PolyBase	Database scoped credential <input type="checkbox"/>
	Database encryption key <input type="checkbox"/>
	Application role <input type="checkbox"/>

- 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 2)

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

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

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

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

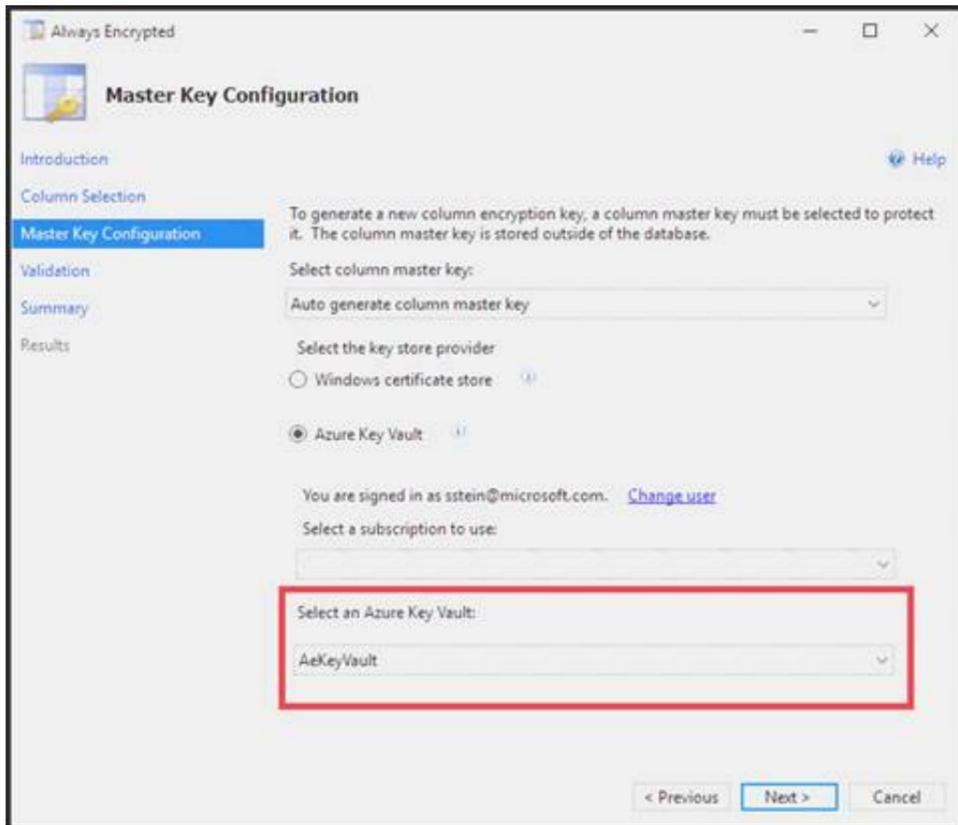
- A. Yes
- B. No

**Answer: B**

**Explanation:**

Use the Azure Key Vault, not the Windows Certificate Store, to store the master key.

Note: The Master Key Configuration page is where you set up your CMK (Column Master Key) and select the key store provider where the CMK will be stored. Currently, you can store a CMK in the Windows certificate store, Azure Key Vault, or a hardware security module (HSM).



References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-always-encrypted-azure-key-vault>

#### NEW QUESTION 4

- (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. Scenario:

Internal Distribution and Sales	9	Yes, once ingested at branches	Data ingested from Contoso branches
---------------------------------	---	--------------------------------	-------------------------------------

Tier 9 reporting must be moved to Event Hubs, queried, and persisted in the same Azure region as the company's main office

References:

<https://docs.microsoft.com/en-us/azure/event-hubs/event-hubs-kafka-stream-analytics>

#### NEW QUESTION 5

- (Exam Topic 2)

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

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

Actions	Answer Area
Connect to the Database and use Azure PowerShell to create a database firewall rule	
Set the Allow Azure Services to Access Server to Disabled	
In 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	

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

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

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

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

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

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

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

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

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

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

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

To setup a database-level firewall rule:

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

- EXECUTE sp\_set\_database\_firewall\_rule N'Example DB Rule','0.0.0.4','0.0.0.4';

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

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

**NEW QUESTION 6**

- (Exam Topic 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					
<b>Ingestion</b>	<table border="1"> <tr><td>HDInsight Kafka</td><td rowspan="4">v</td></tr> <tr><td>Azure Event Hubs</td></tr> <tr><td>HDInsight Storm</td></tr> <tr><td>HDInsight Spark</td></tr> </table>	HDInsight Kafka	v	Azure Event Hubs	HDInsight Storm	HDInsight Spark
HDInsight Kafka	v					
Azure Event Hubs						
HDInsight Storm						
HDInsight Spark						
<b>Stream Processing</b>	<table border="1"> <tr><td>Azure Stream Analytics</td><td rowspan="4">v</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	v	HDInsight with Spark Streaming	Azure Cosmos DB Change Feed	Azure Analysis Services
Azure Stream Analytics	v					
HDInsight with Spark Streaming						
Azure Cosmos DB Change Feed						
Azure Analysis Services						
<b>Analytical Data Store</b>	<table border="1"> <tr><td>Hive LLAP on HDInsight</td><td rowspan="4">v</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	v	Azure Analysis Services	Azure Cosmos DB	SQL Data Warehouse
Hive LLAP on HDInsight	v					
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 7**

- (Exam Topic 3)

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

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

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

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

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

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

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

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

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

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

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

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

References:

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

**NEW QUESTION 8**

- (Exam Topic 3)

You develop data engineering solutions for a company. The company has on-premises Microsoft SQL Server databases at multiple locations.

The company must integrate data with Microsoft Power BI and Microsoft Azure Logic Apps. The solution must avoid single points of failure during connection and transfer to the cloud. The solution must also minimize latency.

You need to secure the transfer of data between on-premises databases and Microsoft Azure.  
 What should you do?

- A. Install a standalone on-premises Azure data gateway at each location
- B. Install an on-premises data gateway in personal mode at each location
- C. Install an Azure on-premises data gateway at the primary location
- D. Install an Azure on-premises data gateway as a cluster at each location

**Answer: D**

**Explanation:**

You can create high availability clusters of On-premises data gateway installations, to ensure your organization can access on-premises data resources used in Power BI reports and dashboards. Such clusters allow gateway administrators to group gateways to avoid single points of failure in accessing on-premises data resources. The Power BI service always uses the primary gateway in the cluster, unless it's not available. In that case, the service switches to the next gateway in the cluster, and so on.

References:

<https://docs.microsoft.com/en-us/power-bi/service-gateway-high-availability-clusters>

**NEW QUESTION 9**

- (Exam Topic 3)

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
        
```

spark  
hadoop

spark  
hadoop  
HBase  
Storm

New-AzureRmHDInsightCluster  
 New-AzureRmHDInsightClusterConfig  
 New-AzureStorageContainer

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

spark  
hadoop

spark  
hadoop  
HBase  
Storm

New-AzureRmHDInsightCluster  
 New-AzureRmHDInsightClusterConfig  
 New-AzureStorageContainer

**NEW QUESTION 10**

- (Exam Topic 3)

You are developing a solution using a Lambda architecture on Microsoft Azure. The data at test layer must meet the following requirements:

Data storage:

- Serve as a repository (or high volumes of large files in various formats).
- Implement optimized storage for big data analytics workloads.
- Ensure that data can be organized using a hierarchical structure. Batch processing:
- Use a managed solution for in-memory computation processing.

- Natively support Scala, Python, and R programming languages.
- Provide the ability to resize and terminate the cluster automatically. Analytical data store:
- Support parallel processing.
- Use columnar storage.
- Support SQL-based languages.

You need to identify the correct technologies to build the Lambda architecture.

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	Technology
Data storage	<ul style="list-style-type: none"> <li>Azure SQL Database</li> <li>Azure Blob Storage</li> <li>Azure Cosmos DB</li> <li>Azure Data Lake Store</li> </ul>
Batch processing	<ul style="list-style-type: none"> <li>HDInsight Spark</li> <li>HDInsight Hadoop</li> <li>Azure Databricks</li> <li>HDInsight Interactive Query</li> </ul>
Analytical data store	<ul style="list-style-type: none"> <li>HDInsight HBase</li> <li>Azure SQL Data Warehouse</li> <li>Azure Analysis Services</li> <li>Azure Cosmos DB</li> </ul>

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

Architecture requirement	Technology
Data storage	<ul style="list-style-type: none"> <li>Azure SQL Database</li> <li>Azure Blob Storage</li> <li>Azure Cosmos DB</li> <li>Azure Data Lake Store</li> </ul>
Batch processing	<ul style="list-style-type: none"> <li>HDInsight Spark</li> <li>HDInsight Hadoop</li> <li>Azure Databricks</li> <li>HDInsight Interactive Query</li> </ul>
Analytical data store	<ul style="list-style-type: none"> <li>HDInsight HBase</li> <li>Azure SQL Data Warehouse</li> <li>Azure Analysis Services</li> <li>Azure Cosmos DB</li> </ul>

**NEW QUESTION 10**

- (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\_bckupoperator).

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 14**

- (Exam Topic 3)

A company uses Microsoft Azure SQL Database to store sensitive company data. You encrypt the data and only allow access to specified users from specified

locations.

You must monitor data usage, and data copied from the system to prevent data leakage.

You need to configure Azure SQL Database to email a specific user when data leakage occurs.

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

Actions	Answer Area
In Auditing, enable <b>Auditing</b> .	
Configure the service to create alerts for threat detections of type <b>Data Exfiltration</b> .	
In Firewalls and virtual networks, enable <b>Allow access to Azure services</b> .	
Enable advanced threat protection.	
Configure the service to send email alerts to security@contoso.com	

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

Actions	Answer Area
In Auditing, enable <b>Auditing</b> .	Enable advanced threat protection.
Configure the service to create alerts for threat detections of type <b>Data Exfiltration</b> .	Configure the service to send email alerts to security@contoso.com
In Firewalls and virtual networks, enable <b>Allow access to Azure services</b> .	Configure the service to create alerts for threat detections of type <b>Data Exfiltration</b> .
Enable advanced threat protection.	
Configure the service to send email alerts to security@contoso.com	

**NEW QUESTION 15**

- (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 18**

- (Exam Topic 3)

You implement an event processing solution using Microsoft Azure Stream Analytics. The solution must meet the following requirements:

- Ingest data from Blob storage
- Analyze data in real time
- Store processed data in Azure Cosmos DB

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

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

**NEW QUESTION 23**

- (Exam Topic 3)

A company plans to develop solutions to perform batch processing of multiple sets of geospatial data. You need to implement the solutions.

Which Azure services should you use? To answer, select the appropriate configuration in the answer area. NOTE: Each correct selection is worth one point.

Scenario	Tool
Use a native client application to run interactive queries and batch processes.	<ul style="list-style-type: none"> <li>HDInsight Tools for Visual Studio</li> <li>Hive View</li> <li>HDInsight REST API</li> <li>Azure Data Factory</li> </ul>
Use a web browser to run interactive queries and batch processes.	<ul style="list-style-type: none"> <li>HDInsight Tools for Visual Studio</li> <li>Hive View</li> <li>HDInsight REST API</li> <li>Azure PowerShell</li> </ul>
Develop batch processing applications that use Azure HDInsight.	<ul style="list-style-type: none"> <li>HDInsight Tools for Visual Studio</li> <li>Hive View</li> <li>HDInsight REST API</li> <li>NoSQL database</li> </ul>

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

Scenario	Tool
Use a native client application to run interactive queries and batch processes.	<ul style="list-style-type: none"> <li>HDInsight Tools for Visual Studio</li> <li>Hive View</li> <li><b>HDInsight REST API</b></li> <li>Azure Data Factory</li> </ul>
Use a web browser to run interactive queries and batch processes.	<ul style="list-style-type: none"> <li><b>HDInsight Tools for Visual Studio</b></li> <li>Hive View</li> <li>HDInsight REST API</li> <li>Azure PowerShell</li> </ul>
Develop batch processing applications that use Azure HDInsight.	<ul style="list-style-type: none"> <li>HDInsight Tools for Visual Studio</li> <li><b>Hive View</b></li> <li>HDInsight REST API</li> <li>NoSQL database</li> </ul>

**NEW QUESTION 28**

- (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 32**

- (Exam Topic 3)

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

You need to create the input stream.

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

NOTE: Each correct selection is worth one point.

**Answer Area**

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

**Answer Area**

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

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

**Answer Area**

```

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

```

**NEW QUESTION 37**

- (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 42**

- (Exam Topic 3)

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

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

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

F. Enable multi-region writes.

**Answer:** A

**NEW QUESTION 43**

- (Exam Topic 3)

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

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

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

**Answer:** C

**Explanation:**

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

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

In the template, you define the resources that are needed for the application. You also specify deployment parameters to input values for different environments. The template consists of JSON and expressions that you use to construct values for your deployment.

References:

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

**NEW QUESTION 47**

- (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)

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 56**

- (Exam Topic 3)

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

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

You need to create the pipeline system.

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

NOTE: Each correct selection is worth one point.

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

- 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 60**

- (Exam Topic 3)

You configure monitoring for a Microsoft Azure SQL Data Warehouse implementation. The implementation uses PolyBase to load data from comma-separated value (CSV) files stored in Azure Data Lake Gen 2 using an external table.

Files with an invalid schema cause errors to occur. You need to monitor for an invalid schema error. For which error should you monitor?

- A. EXTERNAL TABLE access failed due to internal error: 'Java exception raised on call to HdfsBridge\_Connect: Error[com.microsoft.polybase.client.KerberosSecureLogin] occurred while accessing external files.'
- B. EXTERNAL TABLE access failed due to internal error: 'Java exception raised on call to HdfsBridge\_Connect: Error [No FileSystem for scheme: wasbs] occurred while accessing external file.'
- C. Cannot execute the query "Remote Query" against OLE DB provider "SQLNCLI11": for linked server "(null)", Query aborted- the maximum reject threshold (0 rows) was reached while regarding from an external source: 1 rows rejected out of total 1 rows processed.
- D. EXTERNAL TABLE access failed due to internal error: 'Java exception raised on call to HdfsBridge\_Connect: Error [Unable to instantiate LoginClass] occurredwhile accessing external files.'

**Answer: C**

**Explanation:**

Customer Scenario:

SQL Server 2016 or SQL DW connected to Azure blob storage. The CREATE EXTERNAL TABLE DDL points to a directory (and not a specific file) and the directory contains files with different schemas.

SSMS Error:

Select query on the external table gives the following error: Msg 7320, Level 16, State 110, Line 14

Cannot execute the query "Remote Query" against OLE DB provider "SQLNCLI11" for linked server "(null)". Query aborted-- the maximum reject threshold (0 rows) was reached while reading from an external source: 1 rows rejected out of total 1 rows processed.

Possible Reason:

The reason this error happens is because each file has different schema. The PolyBase external table DDL when pointed to a directory recursively reads all the files in that directory. When a column or data type mismatch happens, this error could be seen in SSMS.

Possible Solution:

If the data for each table consists of one file, then use the filename in the LOCATION section prepended by the directory of the external files. If there are multiple files per table, put each set of files into different directories in Azure Blob Storage and then you can point LOCATION to the directory instead of a particular file. The latter suggestion is the best practices recommended by SQLCAT even if you have one file per table.

**NEW QUESTION 65**

- (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 68**

- (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.

- A. Mastered
- B. Not Mastered

**Answer: A**

**Explanation:**

**NEW QUESTION 71**

- (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 75**

- (Exam Topic 3)

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

You develop a data ingestion process that will import data to a Microsoft Azure SQL Data Warehouse. The data to be ingested resides in parquet files stored in an Azure Data lake Gen 2 storage account.

You need to load the data from the Azure Data Lake Gen 2 storage account into the Azure SQL Data Warehouse.

Solution:

1. Create an external data source pointing to the Azure storage account
2. Create a workload group using the Azure storage account name as the pool name
3. Load the data using the INSERT...SELECT statement

Does the solution meet the goal?

- A. Yes
- B. No

**Answer: B**

**Explanation:**

You need to create an external file format and external table using the external data source. You then 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 80**

- (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 checked="" type="checkbox"/></td> </tr> <tr> <td>NameNode</td> <td><input type="checkbox"/></td> </tr> </table>	DataNode	<input checked="" type="checkbox"/>	NameNode	<input type="checkbox"/>
DataNode	<input checked="" type="checkbox"/>				
NameNode	<input type="checkbox"/>				
Run operations on files and directories of the file system.	<table 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 83

- (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 86

.....

## **Thank You for Trying Our Product**

### **We offer two products:**

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

2nd - Questions and Answers in PDF Format

### **DP-200 Practice Exam Features:**

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

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