

Exam Questions DP-200

Implementing an Azure Data Solution

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



NEW QUESTION 1

- (Exam Topic 1)

You need to ensure 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 2

- (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 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 Azure Key Vault
5. Validate configuration results and deploy the solution Does the solution meet the goal?

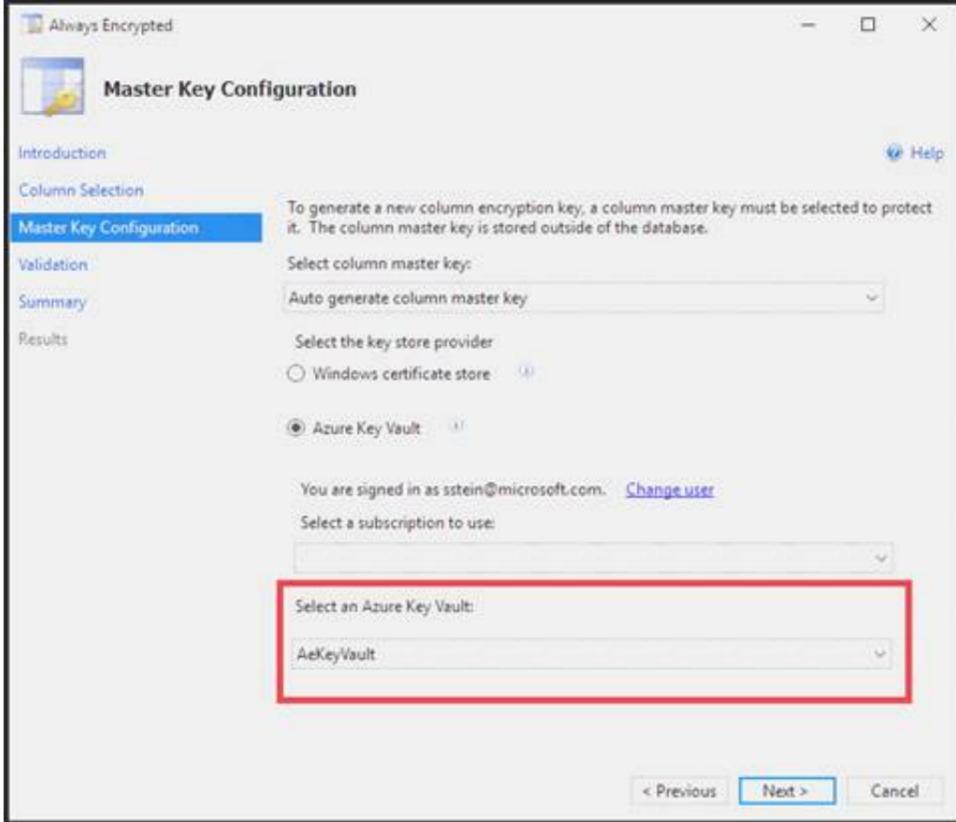
- A. Yes
- B. No

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 4

- (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 5

- (Exam Topic 3)

Each day, company plans to store hundreds of files in Azure Blob Storage and Azure Data Lake Storage. The company uses the parquet format. You must develop a pipeline that meets the following requirements:

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

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

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

Answer: B

Explanation:

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

Python can be used to develop Storm components. References:

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

NEW QUESTION 6

- (Exam Topic 3)

You are developing a solution to visualize multiple terabytes of geospatial data. The solution has the following requirements:

- Data must be encrypted.
- Data must be accessible by multiple resources on Microsoft Azure. You need to provision storage for the solution.

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

The screenshot shows an exam interface with two main sections: 'Actions' and 'Answer Area'. The 'Actions' list contains six items:

- Enable encryption on the Azure Data Lake using the Azure portal.
- Add an access policy for the new Azure Data Lake account to the key storage container.
- Create a new Azure Data Lake Storage account with Azure Data Lake managed encryption keys.
- Select and configure an encryption key storage container.
- Create a new Azure Data Lake Storage account with Azure Key Vault managed encryption keys.
- Create a new Azure Data Lake Storage account with encryption disabled.

The 'Answer Area' is currently empty. There are navigation arrows (left and right) between the two sections.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

The screenshot shows the same exam interface as above, but with four actions moved from the 'Actions' list to the 'Answer Area' in the following sequence:

- Select and configure an encryption key storage container.
- Add an access policy for the new Azure Data Lake account to the key storage container.
- Create a new Azure Data Lake Storage account with Azure Data Lake managed encryption keys.
- Create a new Azure Data Lake Storage account with encryption disabled.

The remaining two actions in the 'Actions' list are: 'Enable encryption on the Azure Data Lake using the Azure portal.' and 'Create a new Azure Data Lake Storage account with Azure Key Vault managed encryption keys.'

NEW QUESTION 7

- (Exam Topic 3)

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

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

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

Solution: Configure Azure Data Lake Storage diagnostics to store logs and metrics in a storage account. Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

NEW QUESTION 8

- (Exam Topic 3)

A company runs Microsoft SQL Server in an on-premises virtual machine (VM).

You must migrate the database to Azure SQL Database. You synchronize users from Active Directory to Azure Active Directory (Azure AD).

You need to configure Azure SQL Database to use an Azure AD user as administrator. What should you configure?

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

Answer: A

NEW QUESTION 9

- (Exam Topic 3)

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

You need to create the offline export copy.

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

NOTE: Each correct selection is worth one point.

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

Answer: BCE

NEW QUESTION 10

- (Exam Topic 3)

You develop data engineering solutions for a company. An application creates a database on Microsoft Azure. You have the following code:

Which database and authorization types are used? To answer, select the appropriate option in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Component	Technology								
Azure database type	<table border="1"> <tr><td>Azure Cosmos DB</td><td><input checked="" type="checkbox"/></td></tr> <tr><td>Azure SQL Database</td><td><input type="checkbox"/></td></tr> <tr><td>files</td><td><input type="checkbox"/></td></tr> <tr><td>Blob</td><td><input type="checkbox"/></td></tr> </table>	Azure Cosmos DB	<input checked="" type="checkbox"/>	Azure SQL Database	<input type="checkbox"/>	files	<input type="checkbox"/>	Blob	<input type="checkbox"/>
	Azure Cosmos DB	<input checked="" type="checkbox"/>							
	Azure SQL Database	<input type="checkbox"/>							
	files	<input type="checkbox"/>							
Blob	<input type="checkbox"/>								
Key type	<table border="1"> <tr><td>resource token</td><td><input checked="" type="checkbox"/></td></tr> <tr><td>Master key</td><td><input type="checkbox"/></td></tr> <tr><td>certificate</td><td><input type="checkbox"/></td></tr> </table>	resource token	<input checked="" type="checkbox"/>	Master key	<input type="checkbox"/>	certificate	<input type="checkbox"/>		
	resource token	<input checked="" type="checkbox"/>							
	Master key	<input type="checkbox"/>							
certificate	<input type="checkbox"/>								

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Azure Cosmos DB

The DocumentClient.CreateDatabaseAsync(Database, RequestOptions) method creates a database resource as an asynchronous operation in the Azure Cosmos DB service.

Box 2: Master Key

Azure Cosmos DB uses two types of keys to authenticate users and provide access to its data and resources: Master Key, Resource Tokens

Master keys provide access to the all the administrative resources for the database account. Master keys: Provide access to accounts, databases, users, and permissions.

- Cannot be used to provide granular access to containers and documents.
- Are created during the creation of an account.
- Can be regenerated at any time.

NEW QUESTION 10

- (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 14

- (Exam Topic 3)

A company manages several on-premises Microsoft SQL Server databases.

You need to migrate the databases to Microsoft Azure by using the backup process of Microsoft SQL Server. Which data technology should you use?

- A. Azure SQL Database Managed Instance
- B. Azure SQL Data Warehouse
- C. Azure Cosmos DB
- D. Azure SQL Database single database

Answer: D

NEW QUESTION 18

- (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 19

- (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 (orows) 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 23

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

References:

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

NEW QUESTION 24

- (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 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. 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 30

- (Exam Topic 3)

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

Reports indicate slowdowns during a specific lime window.

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

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

Answer: B

Explanation:

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

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

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

References:

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

NEW QUESTION 35

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

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

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

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

Money Back Guarantee

DP-200 Practice Exam Features:

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