# Practice Test 2

### **Question 1**

Domain :Monitor and optimize data solutions

Your team has created a new Azure Data Factory environment. You have to analyze the pipeline executions. Trends need to be identified in execution duration over the past 30 days. You need to create a solution that would ensure that data can be queried from Azure Log Analytics.

Which of the following would you choose as the Log type when setting up the diagnostic setting for Azure Data Factory?

]A.

**ActivityRuns**

]B.

**AllMetrics**

]C.

**PipelineRuns**

]D.

**TriggerRuns**

**Explanation:**

Answer – C

Since you need to measure the pipeline execution, consider storing the data on pipeline runs.

The Microsoft documentation gives the schema of the log attributes for pipeline runs. Here there are properties for the start and end time for all activities that run within the pipeline.



Option A is incorrect since this will store the log for each activity execution within the pipeline itself.

Option B is incorrect since this will store all the metrics for the Azure Data Factory resource.

Option D is incorrect since this will store each trigger run for the Azure Data Factory resource.

For more information on monitoring Azure Data Factory, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/monitor-using-azure-monitor>

### **Question 2**

Domain :Monitor and optimize data solutions

Your team has created a new Azure Data Factory environment. You have to analyze the pipeline executions. Trends need to be identified in execution duration over the past 30 days. You need to create a solution that would ensure that data can be queried from Azure Log Analytics.

Which of the following would you use as the storage location when setting up the diagnostic setting for Azure Data Factory?

]A.

**Azure Event Hub**

]B.

**Azure Storage Account**

]C.

**Azure Cosmos DB**

]D.

**Azure Log Analytics**

**Explanation:**

Answer – D

Since we have to query the logs via Log Analytics, we need to choose the storage option as Azure Log Analytics.

Since this is clearly mentioned as a requirement in the question, all other options are incorrect.

For more information on monitoring Azure Data Factory, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/monitor-using-azure-monitor>

### **Question 3**

Domain :Manage and develop data processing

You have to develop a solution that will make use of Azure Stream Analytics. The solution will perform data streaming and will also need a reference data store. Which of the following could be used as the input type for the reference data store?

]A.

**Azure Cosmos DB**

]B.

**Azure Event Hubs**

]C.

**Azure Blob storage**

]D.

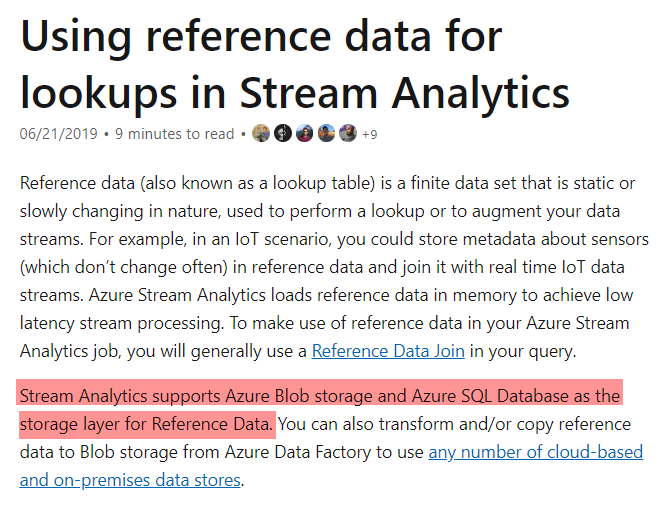
**Azure IoT Hub**

**Explanation:**

Answer – C

You can use Azure Blob storage as an input type for the reference data.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on using reference data, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-use-reference-data>

### **Question 4**

Domain :Manage and develop data processing

You have to develop a solution using Azure Stream Analytics. The stream will be sued to receive Twitter data from Azure Event Hubs. The output would be sent to an Azure Blob storage account. The key requirement is to output the number of tweets during the last 3 minutes every 3 minutes. Each tweet must be counted only once. Which of the following would you use as the windowing function?

]A.

**A three-minute Session window**

]B.

**A three-minute Sliding ion window**

]C.

**A three-minute Tumbling window**

]D.

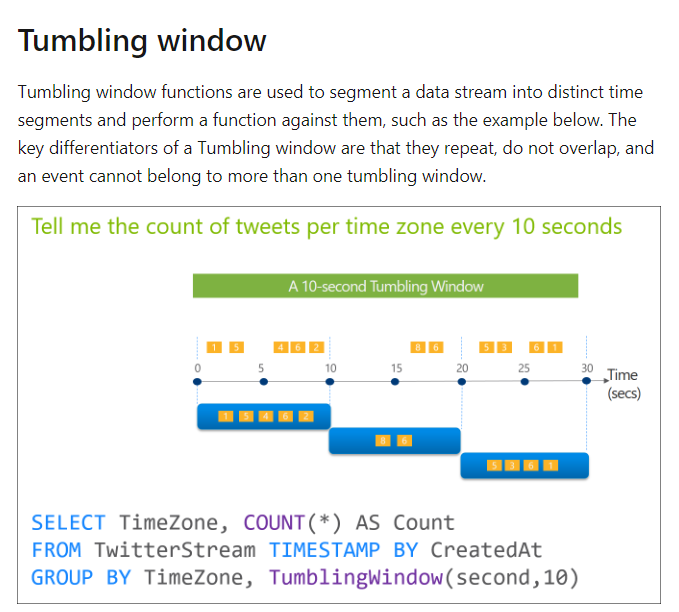
**A three-minute Hopping window**

**Explanation:**

Answer – C

The Tumbling window guarantees that data gets segmented into distinct time segments. And they do not repeat or overlap.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on stream analytics window functions, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-window-functions>

### **Question 5**

Domain :Implement data storage solutions

A company currently has an Azure SQL database. The company wants to create an offline exported copy of the database. This is so that users can work with the data offline when they don’t have any Internet connection on their laptops. Which of the following are ways that can be used to create the exported copy? Choose 3 answers from the options given below.

A.

**Export to a BACPAC file by using Azure Cloud Shell and save the file to a storage account.**

B.

**Export to a BACPAC file by using SQL Server Management Studio. Save the file to a storage account.**

C.

**Export to a BACPAC file by using the Azure portal.**

D.

**Export to a BACPAC file by using Azure PowerShell and save the file locally.**

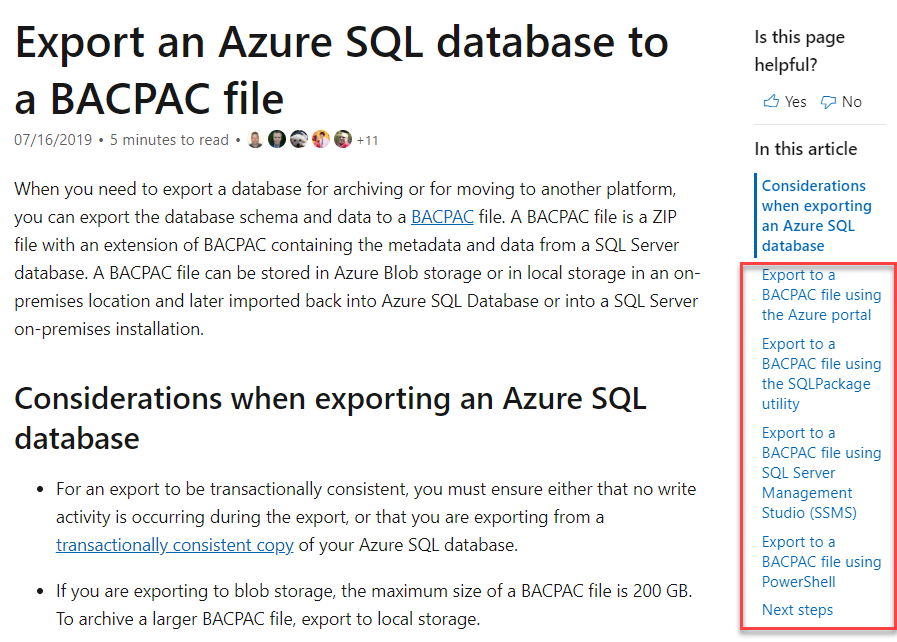
E.

**Export to a BACPAC file by using the SqlPackage utility.**

**Explanation:**

Answer – B, D and E

The Microsoft documentation mentions the different ways in which you can export a BACPAC file of a SQL database.



Option A is incorrect because there is no mention in the Microsoft documentation of being able to create a backup from Azure Cloud Shell.

Option C is incorrect because even though you can create a backup using the Azure Portal, the backup won’t be available locally.

For more information on SQL database export, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-export>

### **Question 6**

Domain :Manage and develop data processing

A company has an Azure Databricks workspace. The workspace will contain three types of workloads.

* One workload for data engineers that would make use of Python and SQL.
* One workload for jobs that would run notebooks that would make use of Python, Spark, Scala and SQL.
* One workload that data scientists would use to perform ad hoc analysis in Scala and R.

The following standards need to be adhered to the different Databricks environments.

* The data engineers need to share a cluster.
* The cluster that runs jobs would be triggered via a request. The data scientists and data engineers would provide package notebooks that would need to be deployed to the cluster.
* There are three data scientists currently. Every data scientist has to be assigned their own cluster. The cluster needs to terminate automatically after 120 minutes of inactivity.

You have to create new Databrick clusters for the workloads.

You decide to create a standard cluster for each data scientist, a standard cluster for the data engineers, and a High Concurrency cluster for the jobs.

Would this implementation fulfill the requirement?

]A.**Yes**

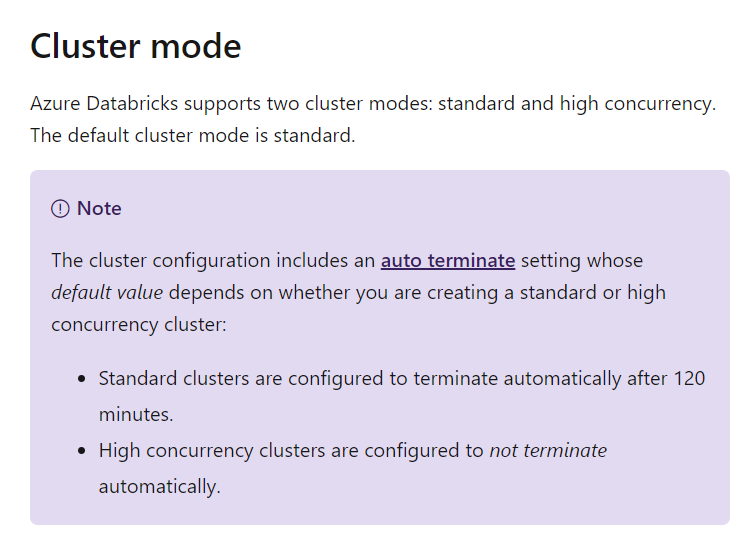
]B.**No**

**Explanation:**

Answer - B

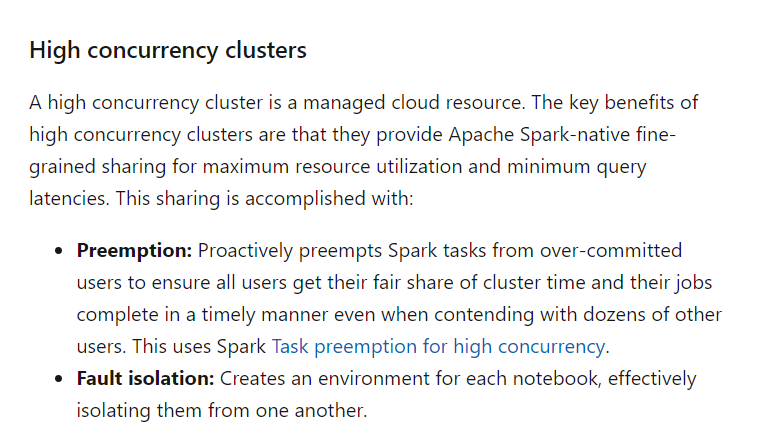
Here each Data scientist must be assigned a standard cluster. This is configured to terminate automatically after 120 minutes.

The Microsoft documentation mentions the following.



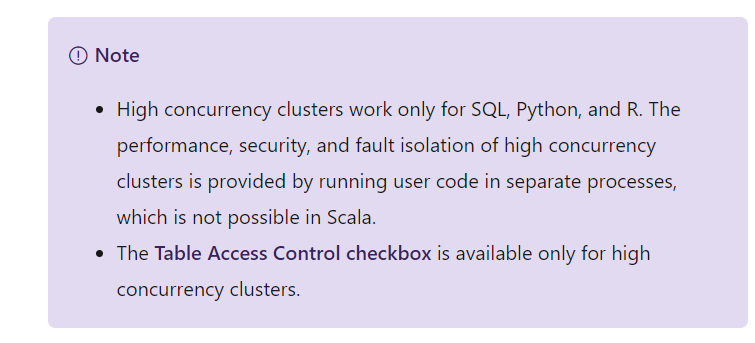
For the data engineers, we can assign a High concurrency cluster. This is beneficial for multiple users who need to use the same cluster.

The Microsoft documentation mentions the following.



For the jobs, we will have to assign a standard cluster because the high concurrency cluster does not support all programming languages for the job types.

The Microsoft documentation mentions the following.



For more information on configuring clusters, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/databricks/clusters/configure>

### **Question 7**

Domain :Manage and develop data processing

A company has an Azure Databricks workspace. The workspace will contain three types of workloads.

* One workload for data engineers that would make use of Python and SQL.
* One workload for jobs that would run notebooks would make use of Python, Spark, Scala and SQL.
* One workload that data scientists would use to perform ad hoc analysis in Scala and R.

The following standards need to be adhered to the different Databricks environments.

* The data engineers need to share a cluster.
* The cluster that runs jobs would be triggered via a request. The data scientists and data engineers would provide package notebooks that would need to be deployed to the cluster.
* There are three data scientists currently. Every data scientist has to be assigned their own cluster. The cluster needs to terminate automatically after 120 minutes of inactivity.

You have to create new Databrick clusters for the workloads.

You decide to create a High Concurrency cluster for each data scientist, a High Concurrency cluster for the data engineers and a standard cluster for the jobs.

Would this implementation fulfill the requirement?

]A.**Yes**

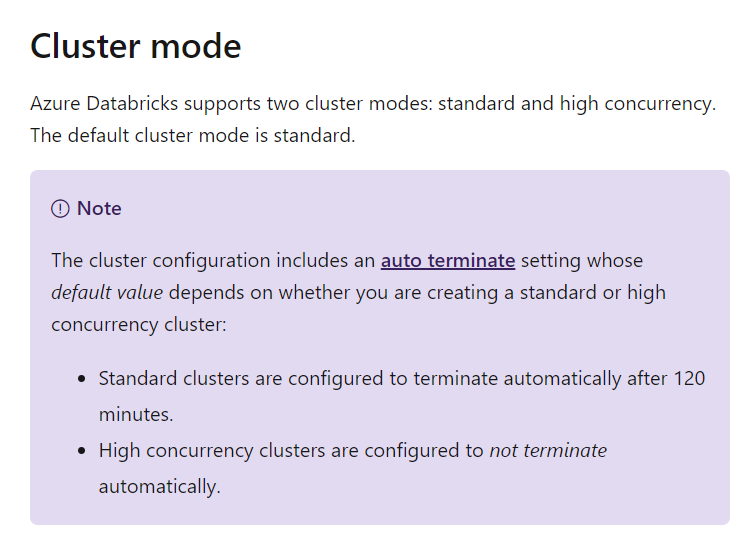
]B.**No**

**Explanation:**

Answer - B

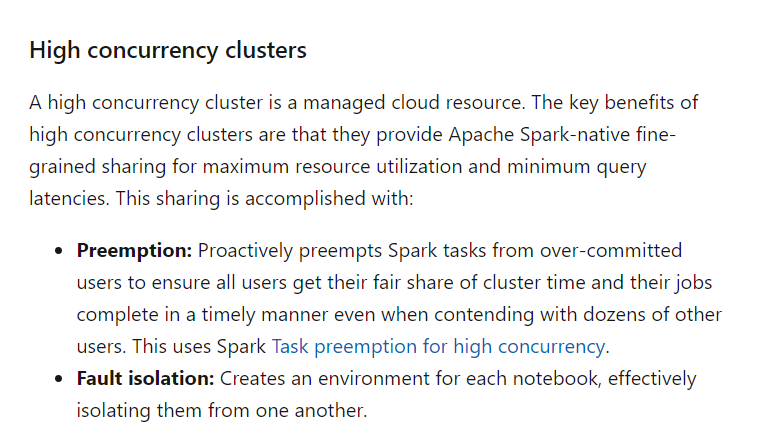
Here each Data scientist must be assigned a standard cluster. This is configured to terminate automatically after 120 minutes.

The Microsoft documentation mentions the following.



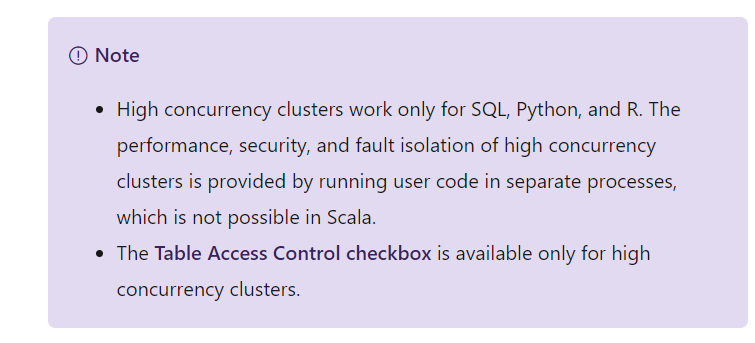
For the data engineers, we can assign a High concurrency cluster. This is beneficial for multiple users who need to use the same cluster.

The Microsoft documentation mentions the following.



For the jobs, we will have to assign a standard cluster because the high concurrency cluster does not support all programming languages for the job types

The Microsoft documentation mentions the following



For more information on configuring clusters, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/databricks/clusters/configure>

### **Question 8**

Domain :Manage and develop data processing

A company has an Azure Databricks workspace. The workspace will contain three types of workloads.

* One workload for data engineers that would make use of Python and SQL.
* One workload for jobs that would run notebooks would make use of Python, Spark, Scala and SQL.
* One workload that data scientists would use to perform ad hoc analysis in Scala and R.

The following standards need to be adhered to the different Databricks environments.

* The data engineers need to share a cluster.
* The cluster that runs jobs would be triggered via a request. The data scientists and data engineers would provide package notebooks that would need to be deployed to the cluster.
* There are three data scientists currently. Every data scientist has to be assigned their own cluster. The cluster needs to terminate automatically after 120 minutes of inactivity.

You have to create new Databrick clusters for the workloads.

You decide to create a Standard cluster for each data scientist, a High Concurrency cluster for the data engineers and a Standard cluster for the jobs.

Would this implementation fulfill the requirement?

]A.**Yes**

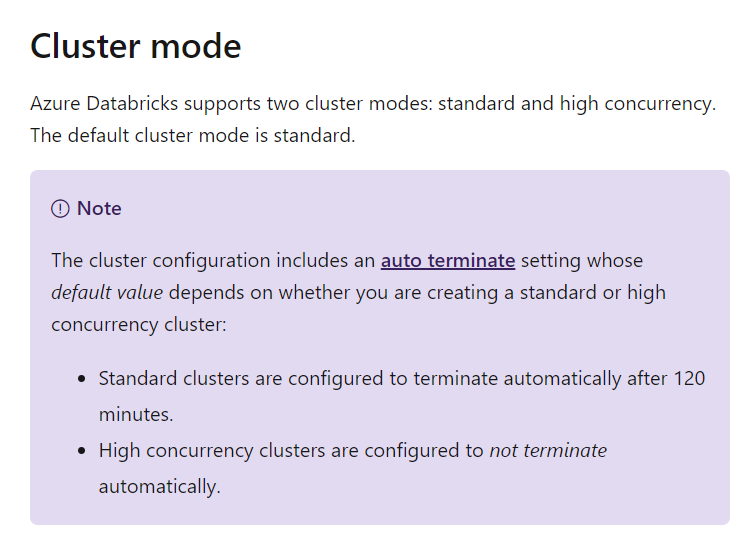
]B.**No**

**Explanation:**

Answer - A

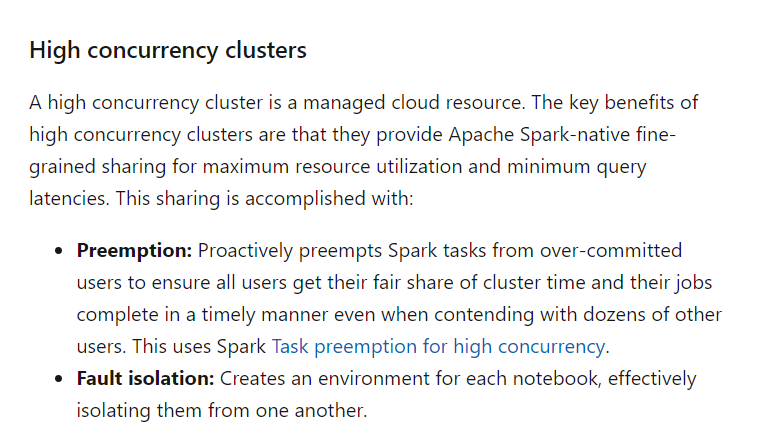
Here each Data scientist must be assigned a standard cluster. This is configured to terminate automatically after 120 minutes.

The Microsoft documentation mentions the following.



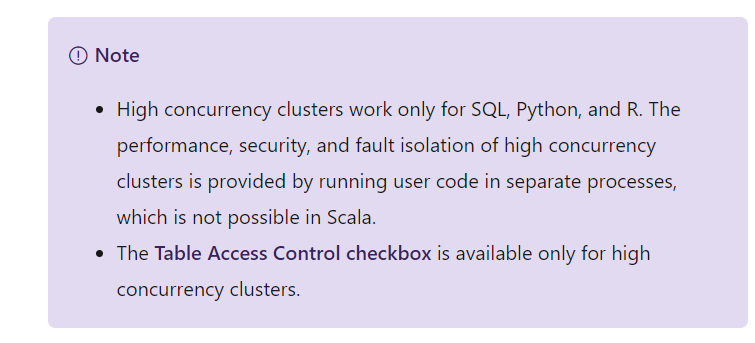
For the data engineers, we can assign a High concurrency cluster. This is beneficial for multiple users who need to use the same cluster.

The Microsoft documentation mentions the following.



For the jobs, we will have to assign a standard cluster because the high concurrency cluster does not support all programming languages for the job types.

The Microsoft documentation mentions the following.



For more information on configuring clusters, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/databricks/clusters/configure>

### **Question 9**

Domain :Manage and develop data processing

A company has an Azure Databricks workspace. The workspace will contain three types of workloads.

* One workload for data engineers that would make use of Python and SQL.
* One workload for jobs that would run notebooks that would make use of Python, Spark, Scala and SQL.
* One workload that data scientists would use to perform ad hoc analysis in Scala and R.

The following standards need to be adhered to the different Databricks environments.

* The data engineers need to share a cluster.
* The cluster that runs jobs would be triggered via a request. The data scientists and data engineers would provide package notebooks that would need to be deployed to the cluster.
* There are three data scientists currently. Every data scientist has to be assigned their own cluster. The cluster needs to terminate automatically after 120 minutes of inactivity.

You have to create new Databrick clusters for the workloads.

You decide to create a High Concurrency cluster for each data scientist, a High Concurrency cluster for the data engineers, and a High Concurrency cluster for the jobs.

Would this implementation fulfill the requirement?

]A.**Yes**

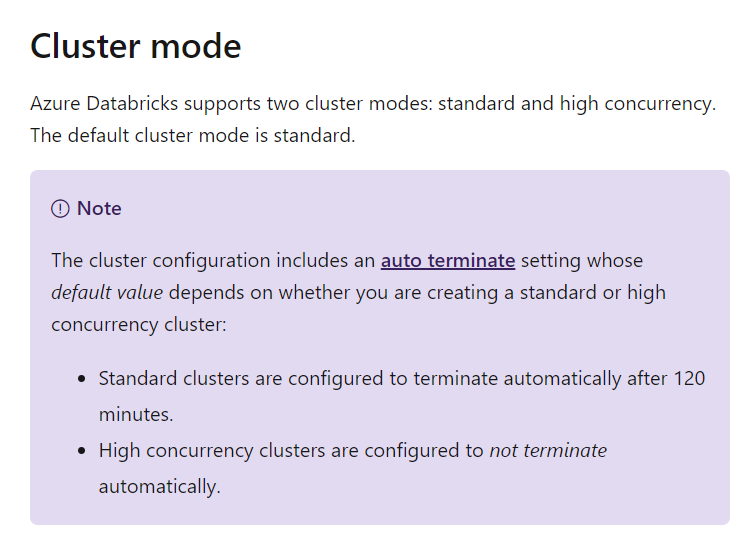
]B.**No**

**Explanation:**

Answer - B

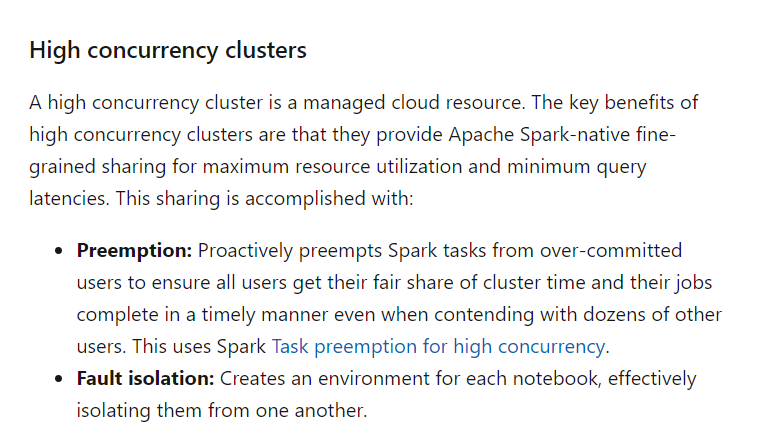
Here each Data scientist must be assigned a standard cluster. This is configured to terminate automatically after 120 minutes.

The Microsoft documentation mentions the following.



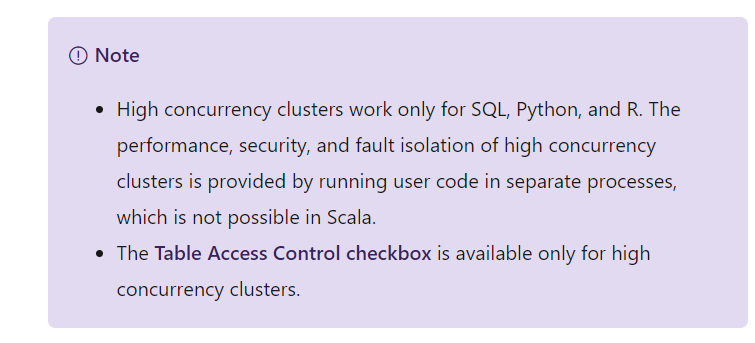
For the data engineers, we can assign a High concurrency cluster. This is beneficial for multiple users who need to use the same cluster.

The Microsoft documentation mentions the following.



For the jobs, we will have to assign a standard cluster because the high concurrency cluster does not support all programming languages for the job types.

The Microsoft documentation mentions the following.



For more information on configuring clusters, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/databricks/clusters/configure>

### **Question 10**

Domain :Manage and develop data processing

You have to develop a solution which would perform the following activities.

* Ingest twitter-based data into Azure.
* Give the ability to visualize real-time Twitter data.

Which of the following would you use to implement this solution? Choose 3 answers from the options given below.

A.

**Make use of an Event Grid Topic.**

B.

**Make use of Azure Stream Analytics to query twitter data from an Event Hub.**

C.

**Make use of Azure Stream Analytics to query twitter data from an Event Grid.**

D.

**Have a Logic App in place that would send twitter data to Azure.**

E.

**Create an Event Grid subscription.**

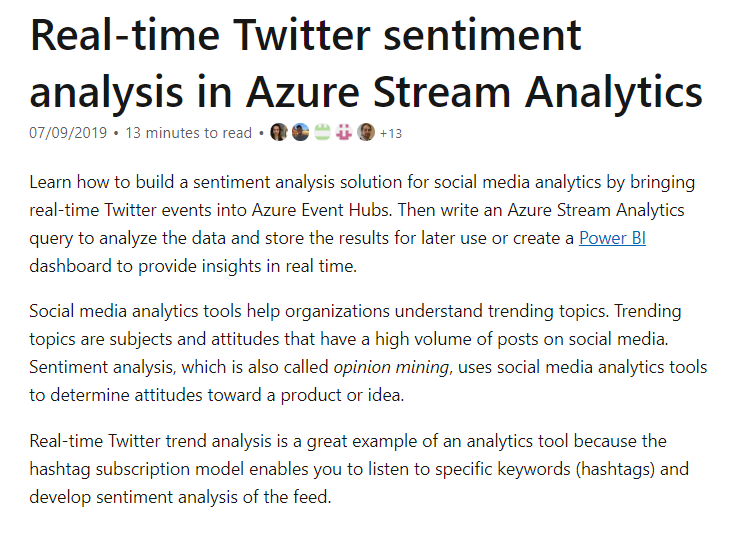
F.

**Create an Event Hub Instance.**

**Explanation:**

Answer – B, D and F

There is an example in the Microsoft documentation, which showcases how to use Azure Stream Analytics to process Twitter data.



Option A is incorrect because this is more of a messaging-based system.

Options C and E are incorrect because the Event Grid service is used for event-based processing.

For more information on the implementation, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-twitter-sentiment-analysis-trends>

### **Question 11**

Domain :Manage and develop data processing

A company wants to pull data from an on-premise SQL Server and migrate the data to Azure Blob storage. The company is planning to use Azure Data Factory. Which of the following are steps that would be required to implement this solution?

A.

**Create a new Azure Data Factory resource.**

B.

**Create a Virtual Private Network Connection from the on-premise network to Azure.**

C.

**Create a self-hosted integration runtime.**

D.

**Create a database master key.**

E.

**Backup the database.**

F.

**Configure the on-premise server to use an integration runtime.**

**Explanation:**

Answer – A , B and C

First, you have to create a Virtual Private Network Connection from the on-premise network to Azure. This is to ensure that you have connectivity between your on-premises data center and Azure.

Next, create a new Azure Data Factory resource and then have a self-hosted integration runtime in Azure Data Factory.

Option D is incorrect because we don’t need a database master key for this process.

Option E is incorrect because we are using Azure Data Factory.

Option F is incorrect because we need to configure the integration runtime in Azure Data Factory.

For more information on how to copy data using Azure Data Factory, one can visit the below URL-

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

### **Question 12**

Domain :Manage and develop data processing

A company wants to integrate their on-premise Microsoft SQL Server data with Azure SQL database. Here the data must be transformed incrementally. Which of the following can be used to configure a pipeline to copy the data?

]A.

**Make use of the AzCopy tool with Blob storage as the linked service in the source.**

]B.

**Make use of Azure PowerShell with SQL Server as the linked service in the source.**

]C.

**Make use of Azure Data Factory UI with Blob storage as the linked service in the source.**

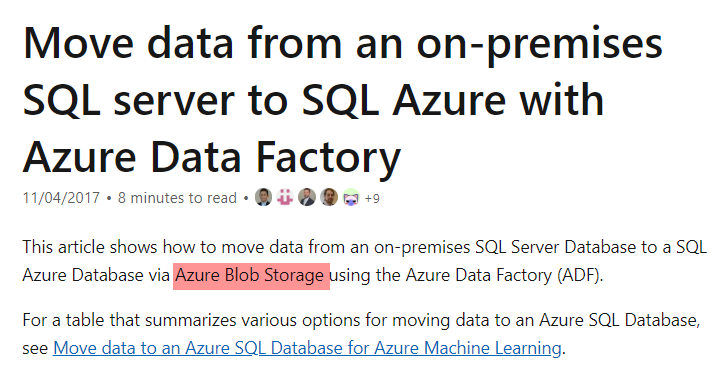
]D.

**Make use of .Net Data Factory API with Blob storage as the linked service in the source.**

**Explanation:**

Answer – C

You can use Azure Data Factory which utilizes Azure Blob storage. An example of this is also given in the Microsoft documentation.



All other options are incorrect since you need to use the Azure Data Factory UI tool to develop a pipeline.

For more information on how to copy data using Azure Data Factory for an on-premise SQL Server, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/move-sql-azure-adf>

### **Question 13**

Domain :Monitor and optimize data solutions

Your company is currently using Azure Stream Analytics to monitor devices.

The company is now planning to deploy more devices, and all of these devices need to be monitored via the same Azure Stream Analytics instance. You have to ensure that there are enough processing resources to handle the load of the additional devices.

Which of the following metric for the Stream Analytics job should you track for this requirement?

]A.

**Input Deserialization Errors**

]B.

**Early Input Events**

]C.

**Late Input Events**

]D.

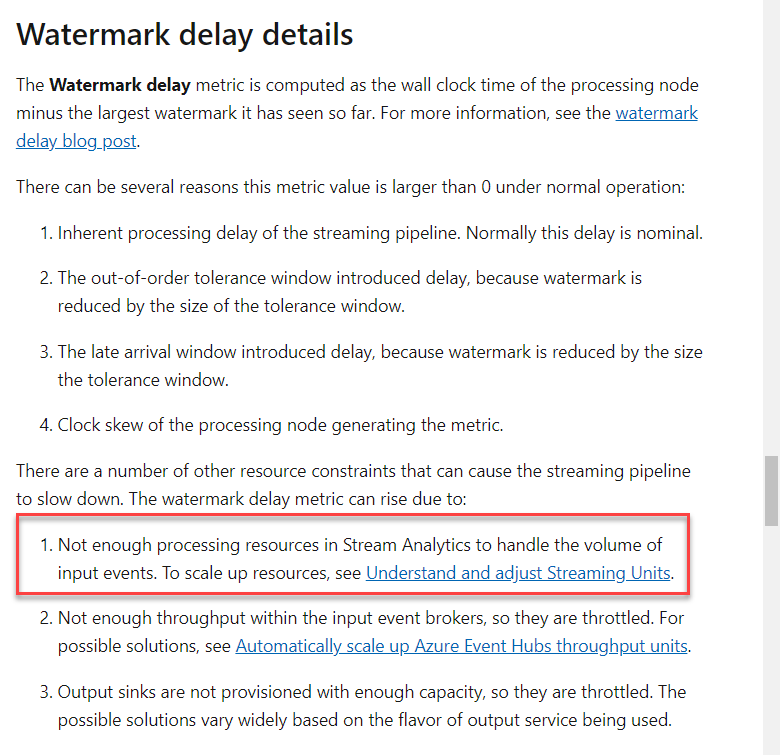
**Watermark delay**

**Explanation:**

Answer – D

You should monitor the Watermark delay. This would indicate if there are not enough processing resources for the input events.

The Microsoft documentation mentions the following.



Option A is incorrect since this is related to deserialization of the input events.

Options B and C are incorrect since this is related to the arrival time of input events.

For more information on monitoring stream analytics, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-time-handling>

### **Question 14**

Domain :Implement data storage solutions

A company wants to migrate a set of on-premise Microsoft SQL Server databases to Azure. They want to migrate the databases as a simple lift and shift process by using backup and restore processes.

Which of the following would they use in Azure to host the SQL databases?

]A.

**Azure SQL Database single database**

]B.

**Azure SQL data warehouse**

]C.

**Azure Cosmos DB**

]D.

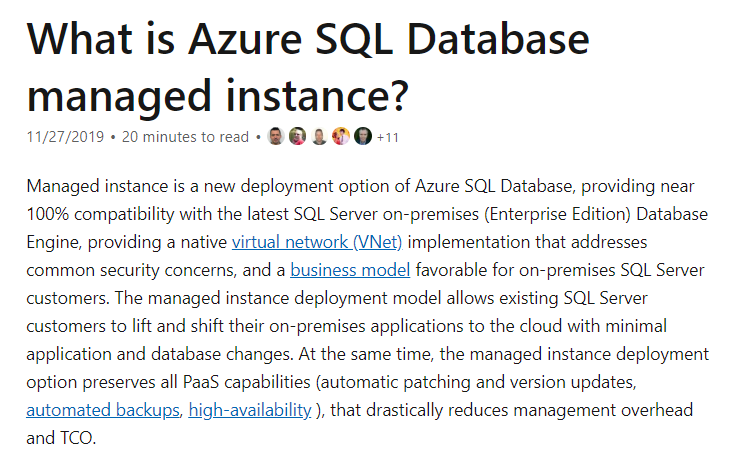
**Azure SQL Database managed instance**

**Explanation:**

Answer – D

For easy migration of on-premise databases, consider migrating to Azure SQL Database managed instance.

The Microsoft documentation mentions the following.



Option A is incorrect since this is a better option if you just want to host a single database on the Azure platform.

Option B is incorrect since this is a data warehousing solution available on the Azure platform.

Option C is incorrect since this is a NoSQL based database solution.

For more information on Azure SQL Database managed instance, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-managed-instance>

### **Question 15**

Domain :Implement data storage solutions

You have to design a Hadoop Distributed File System architecture. You are going to be using Microsoft Azure Data Lake as the data storage repository. You have to ensure that the data repository has a resilient data schema.

Which of the following would you use to provide data access to clients?

]A.

**DataNode**

]B.

**NameNode**

]C.

**PrimaryNode**

]D.

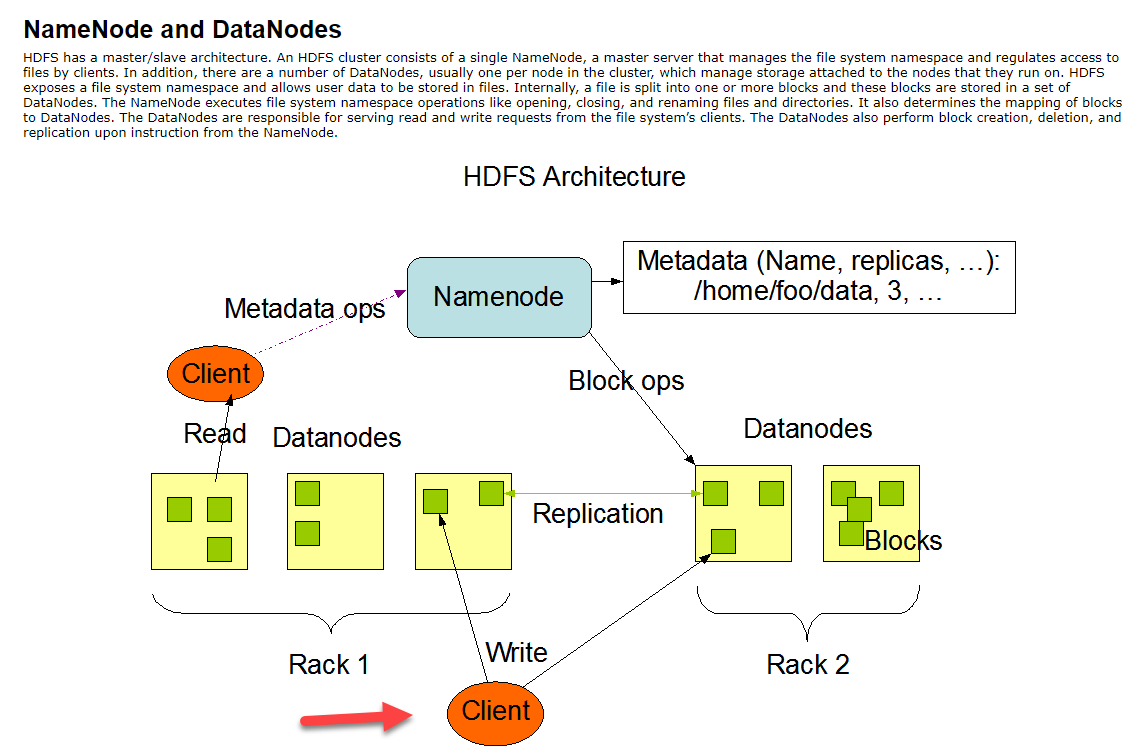
**SecondaryNode**

**Explanation:**

Answer – A

If you look at the architecture of the Hadoop Distributed File System, you will see that clients connect to the Data Nodes.

The Hadoop documentation mentions the following.



Since this is clear from the documentation, all other options are incorrect.

For more information on HDFS design, one can visit the below URL-

* <https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html#NameNode+and+DataNodes>

### **Question 16**

Domain :Implement data storage solutions

You have to design a Hadoop Distributed File System architecture. You are going to be using Microsoft Azure Data Lake as the data storage repository. You have to ensure that the data repository has a resilient data schema.

Which of the following would be used to run operations on files and directories on the file system?

]A.

**DataNode**

]B.

**NameNode**

]C.

**PrimaryNode**

]D.

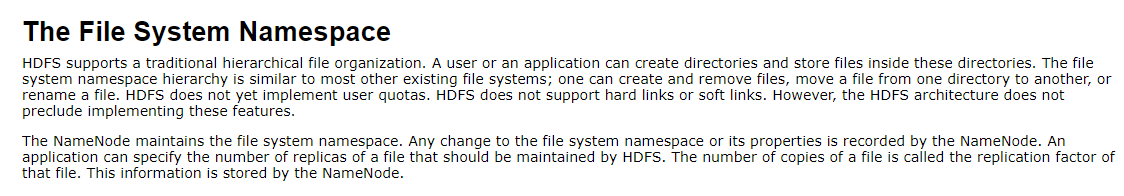
**SecondaryNode**

**Explanation:**

Answer – B

The file system namespace resides on the NameNode.

The Hadoop documentation mentions the following.



Since this is clear from the documentation, all other options are incorrect.

For more information on HDFS design, one can visit the below URL-

* <https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html#NameNode+and+DataNodes>

### **Question 17**

Domain :Implement data storage solutions

You have to design a Hadoop Distributed File System architecture. You are going to be using Microsoft Azure Data Lake as the data storage repository. You have to ensure that the data repository has a resilient data schema.

Which of the following is used to perform block creation, deletion and replication?

]A.

**DataNode**

]B.

**NameNode**

]C.

**PrimaryNode**

]D.

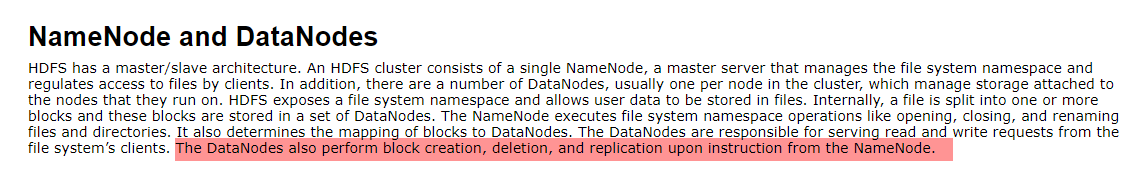
**SecondaryNode**

**Explanation:**

Answer – A

Here this is carried out by the Data Nodes.

The Hadoop documentation mentions the following.



Since this is clear from the documentation, all other options are incorrect.

For more information on HDFS design, one can visit the below URL-

* <https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html#NameNode+and+DataNodes>

### **Question 18**

Domain :Monitor and optimize data solutions

A company wants to make use of Azure SQL Database with Elastic Pools. They have different customers who will have their own database in the pool. Each customer database has its own peak usage during different periods of the year. You need to consider the best way to implement Azure SQL Database elastic pools to minimize costs. Which of the following is an option you would need to consider when configuring elastic pools?

]A.

**Number of transactions only**

]B.

**eDTUs per database only**

]C.

**Number of databases only**

]D.

**CPU usage only**

]E.

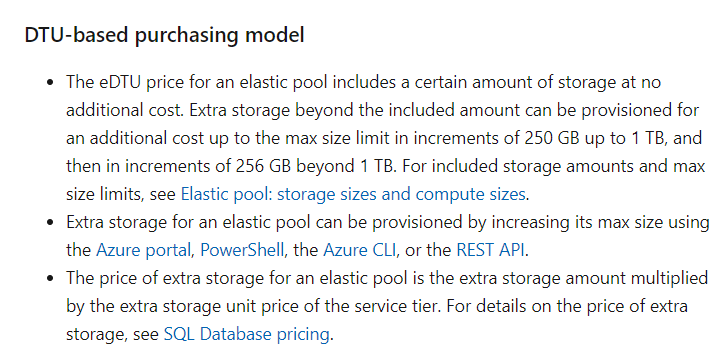
**eDTUs and maximum data size**

**Explanation:**

Answer – E

When you implement Elastic Pools using the DTU-based purchasing model, you have to consider both the eDTU’s and the storage size for the databases.

The Microsoft documentation mentions the following.



Since this is clear from the documentation, all other options are incorrect.

For more information on SQL database elastic pools, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-pool-scale>

### **Question 19**

Domain :Implement data storage solutions

A company needs to configure data synchronization between their on-premise Microsoft SQL Server database and Azure SQL database. The synchronization process must include the following.

* Be able to perform an initial data synchronization to the Azure SQL Database with minimal downtime.
* Be able to perform bi-directional synchronization after the initial synchronization is complete.

Which of the following would you consider as the synchronization solution?

]A.

**Data Migration Assistant**

]B.

**Backup and restore**

]C.

**SQL Server Agent Job**

]D.

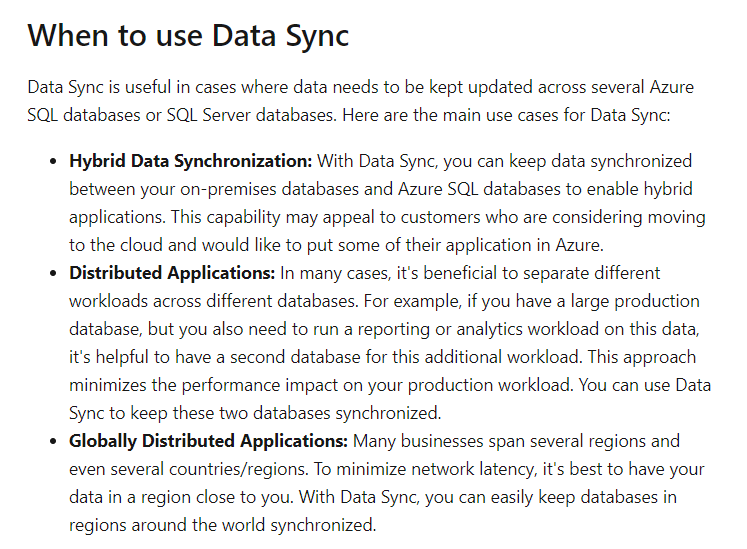
**Azure SQL Data Sync**

**Explanation:**

Answer – D

Azure SQL Data Sync can be used to synchronize data between the on-premise SQL Server and the Azure SQL database.

The Microsoft documentation mentions the following.



Option A is incorrect since this is just used to assess databases for the migration process.

Option B is incorrect since this would be the initial setup activity.

Option C is incorrect since this is used to run administrative tasks on on-premise SQL databases.

For more information on SQL database Sync, one can visit the below URL-

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

### **Question 20**

Domain :Monitor and optimize data solutions

A company has on-premise Microsoft SQL Server databases at several locations. The company wants to integrate the data in the databases with Microsoft Power BI and Microsoft Azure Logic Apps. You need to implement a solution that would avoid any single point of failure during the connection and transfer of data to the cloud. Latency must also be minimized. The transfer of data between the on-premise databases and Microsoft Azure must be secure. Which of the following would you implement for this requirement?

]A.

**Install a standalone on-premise Azure data gateway at each company location.**

]B.

**Install an on-premise data gateway in personal mode at each company location.**

]C.

**Install an Azure on-premise data gateway at the primary company location.**

]D.

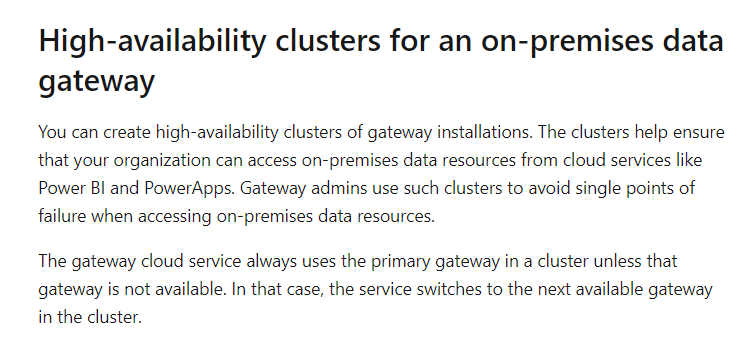
**Install an Azure on-premise data gateway as a cluster at each location.**

**Explanation:**

Answer – D

If you need a high availability solution, then you can install the on-premise data gateway as a cluster.

The Microsoft documentation mentions the following.



Since this is clear from the documentation, all other options are incorrect.

For more information on high available clusters for the gateway, one can visit the below URL-

* <https://docs.microsoft.com/en-us/data-integration/gateway/service-gateway-high-availability-clusters>

### **Question 21**

Domain :Manage and develop data processing

You need to migrate data from an Azure Blob storage account to an Azure SQL Data warehouse. Which of the following actions do you need to implement for this requirement? Choose 4 answers from the options given below.

A.

**Provision an Azure SQL Data Warehouse instance**

B.

**Connect to the Blob storage container via SQL Server Management Studio**

C.

**Create an Azure Blob storage container**

D.

**Run the T-SQL statements to load the data**

E.

**Connect to the Azure SQL Data warehouse via SQL Server Management Studio**

F.

**Build external tables by using Azure portal**

G.

**Build external tables by using SQL Server Management Studio**

**Explanation:**

Answer – A, D, E and G

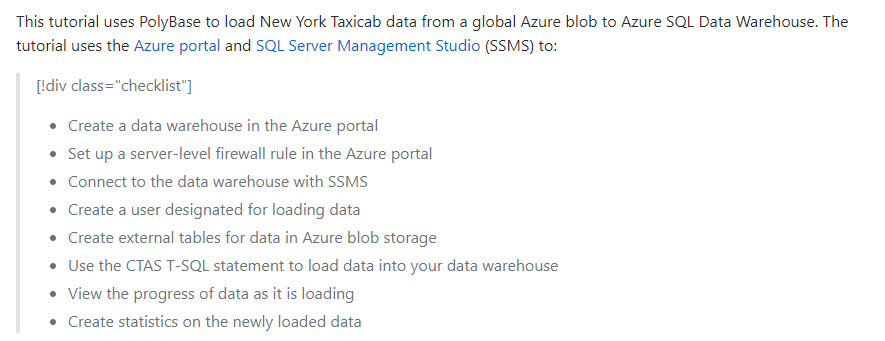
You first need to create an Azure SQL Data Warehouse instance.

Then you need to connect to the data warehouse via SQL Server Management Studio.

Then create external tables to the Azure Blob storage account.

And then finally use T-SQL statements to load the data.

This is also given as an example in GitHub as part of the Microsoft documentation on loading data from Azure Blob to an Azure SQL data warehouse.



Option B is incorrect because you can’t connect to Blob storage from SQL Server Management Studio.

Option C is incorrect because you already have the blob data in place.

Option F is incorrect because you need to build the external tables in SQL Server Management Studio.

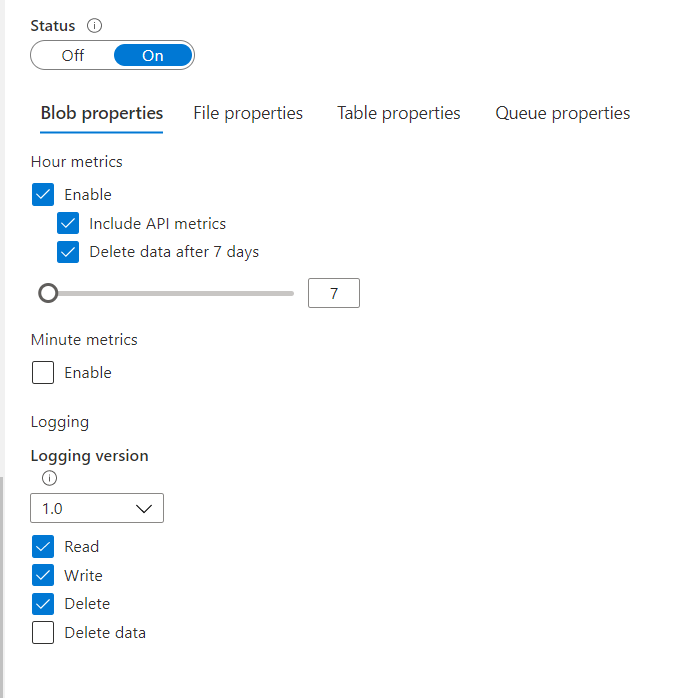
For more information on the example, one can visit the below URL-

* <https://github.com/MicrosoftDocs/azure-docs/blob/master/articles/sql-data-warehouse/load-data-from-azure-blob-storage-using-polybase.md>

### **Question 22**

Domain :Monitor and optimize data solutions

You have an Azure storage account named compstore4000. Below are the Diagnostic settings configured for the storage account.



How long will the logging data be retained for?

]A.

**7 days**

]B.

**365 days**

]C.

**Indefinitely**

]D.

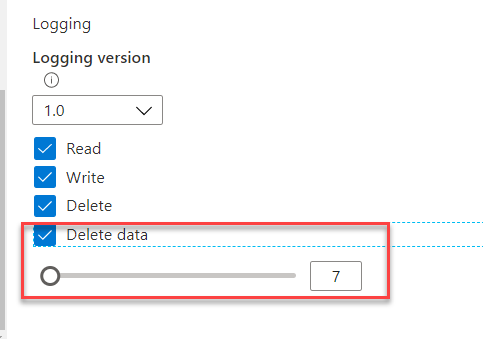
**90 days**

**Explanation:**

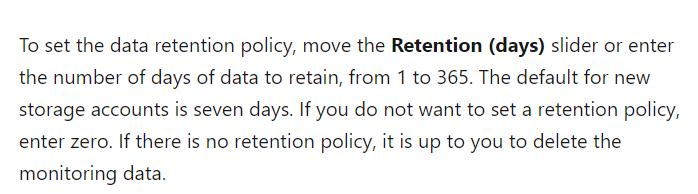
Answer – C

Here since we have not specified an option for Delete data, the data will be stored Indefinitely.

If you choose the Delete data option, you can then mention a retention period.



The Microsoft documentation mentions the following.



Since this is clear from the implementation, all other options are incorrect.

For more information on monitoring storage accounts, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/storage/common/storage-monitor-storage-account>

### **Question 23**

Domain :Implement data storage solutions

Your company has an Azure Data Lake storage account. They want to implement role-based access control (RBAC) so that project members can manage the Azure Data Lake Storage resources. Which of the following actions should you perform for this requirement? Choose 3 answers from the options given below.

A.

**Ensure to assign Azure AD security groups to Azure Data Lake Storage.**

B.

**Make sure to configure end-user authentication to the Azure Data Lake Storage account.**

C.

**Make sure to configure service-to-service authentication to the Azure Data Lake Storage account.**

D.

**Create security groups in Azure AD and then add the project members.**

E.

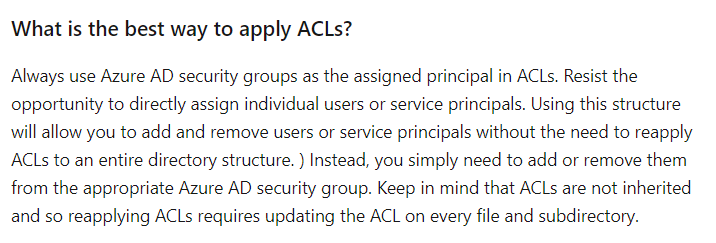
**Configure Access control lists for the Azure Data Lake Storage account.**

**Explanation:**

Answer – A, D and E

You can assign users and service principals, but the Microsoft documentation recommends giving Azure AD group permissions for Azure Data Lake Storage account. For the storage account itself, you can manage the permissions via Access Control Lists.

The Microsoft documentation mentions the following.



Since this is clear from the documentation, all other options are incorrect.

For more information on Azure Data Lake storage access control, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-access-control>

### **Question 24**

Domain :Implement data storage solutions

A company has an Azure SQL Database and an Azure Blob storage account. They want data to be encrypted at rest on both systems. The company should be able to use their own key.

Which of the following would they use to configure security for the Azure SQL Database?

]A.

**Always Encrypted**

]B.

**Cell-level encryption**

]C.

**Row-level encryption**

]D.

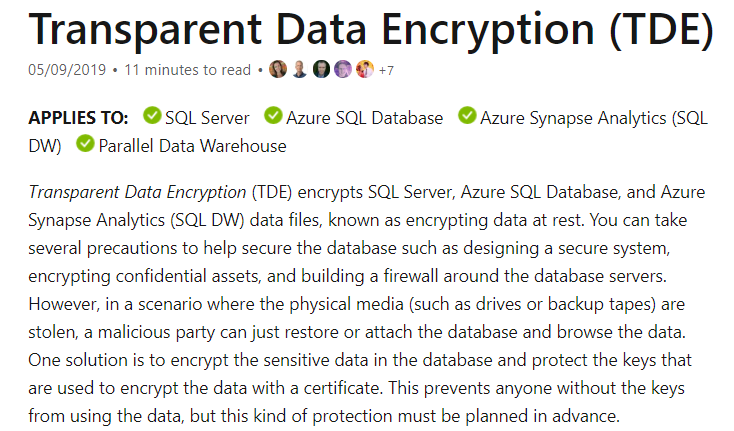
**Transparent data encryption**

**Explanation:**

Answer – D

Transparent Data Encryption is used to encrypt data at rest for Azure SQL Server databases.

The Microsoft documentation mentions the following.



All other options are incorrect as they would not give the facility to encrypt data at rest for the entire database.

For more information on Transparent Data Encryption, one can visit the below URL-

* <https://docs.microsoft.com/en-us/sql/relational-databases/security/encryption/transparent-data-encryption?view=sql-server-ver15>

### **Question 25**

Domain :Implement data storage solutions

A company has an Azure SQL Database and an Azure Blob storage account. They want data to be encrypted at rest on both systems. The company should be able to use their own key.

Which of the following would they use to configure security for the Azure Blob storage account?

]A.

**Azure Disk Encryption**

]B.

**Secure Transport Layer Security**

]C.

**Storage Account Keys**

]D.

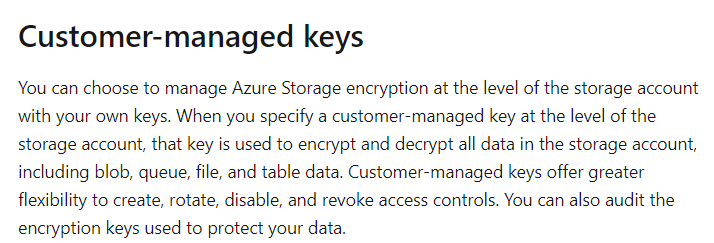
**Default Storage Service Encryption**

**Explanation:**

Answer – D

You can manage the encryption of data at rest for Azure storage accounts using the default storage service encryption.

The Microsoft documentation mentions the following.



Option A is incorrect since this is used for encrypting data at rest for Azure Virtual machines.

Option B is incorrect since this is used to encrypt data in transit.

Option C is incorrect since this is used for authorization to storage accounts.

For more information on Storage Service Encryption, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/storage/common/storage-service-encryption>

### **Question 26**

Domain :Monitor and optimize data solutions

A company has a set of Azure SQL Databases. They want to ensure that their IT Security team is informed when any security-related operation occurs on the database. You need to configure Azure Monitor while ensuring administrative efforts are reduced. Which of the following actions would you perform for this requirement? Choose 3 answers from the options given below.

A.

**Create a new action group which send email alerts to the IT Security team.**

B.

**Make sure to use all security operations as the condition.**

C.

**Ensure to query audit log entries as the condition.**

D.

**Use all the Azure SQL Database servers as the resource.**

**Explanation:**

Answer – A, B and D

You can set up alerts based on all the security conditions in Azure Monitor. When any security operation is performed, an alert can be sent to the IT Security team.

Option C is incorrect since we need to monitor all security related events.

For more information on alerts for Azure SQL Databases, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-insights-alerts-portal>

### **Question 27**

Domain :Manage and develop data processing

You need to deploy a Microsoft Azure Stream Analytics job for an IoT based solution. The solution must minimize latency. The solution must also minimize the bandwidth usage between the job and the IoT device. Which of the following actions must you perform for this requirement? Choose 4 answers from the options given below.

A.

**Ensure to configure routes.**

B.

**Create an Azure Blob storage container.**

C.

**Configure Streaming Units.**

D.

**Create an IoT Hub and add the Azure Stream Analytics modules to the IoT Hub namespace.**

E.

**Create an Azure Stream Analytics edge job and configure job definition save location.**

F.

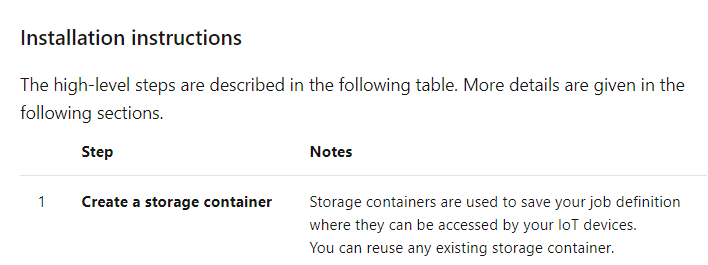
**Create an Azure Stream Analytics cloud job and configure job definition save location.**

**Explanation:**

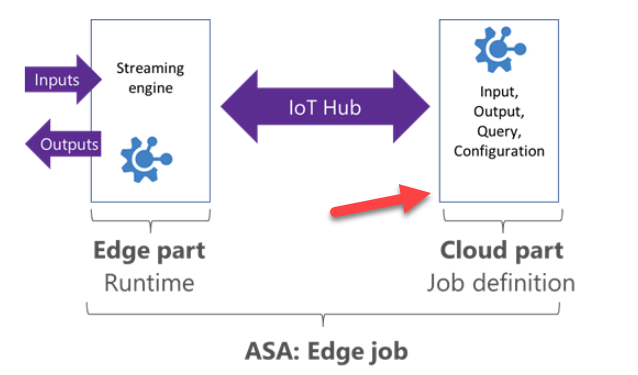
Answer – A, B, D and E

There is an article in the Microsoft documentation on configuring Azure Stream Analytics on IoT Edge devices.

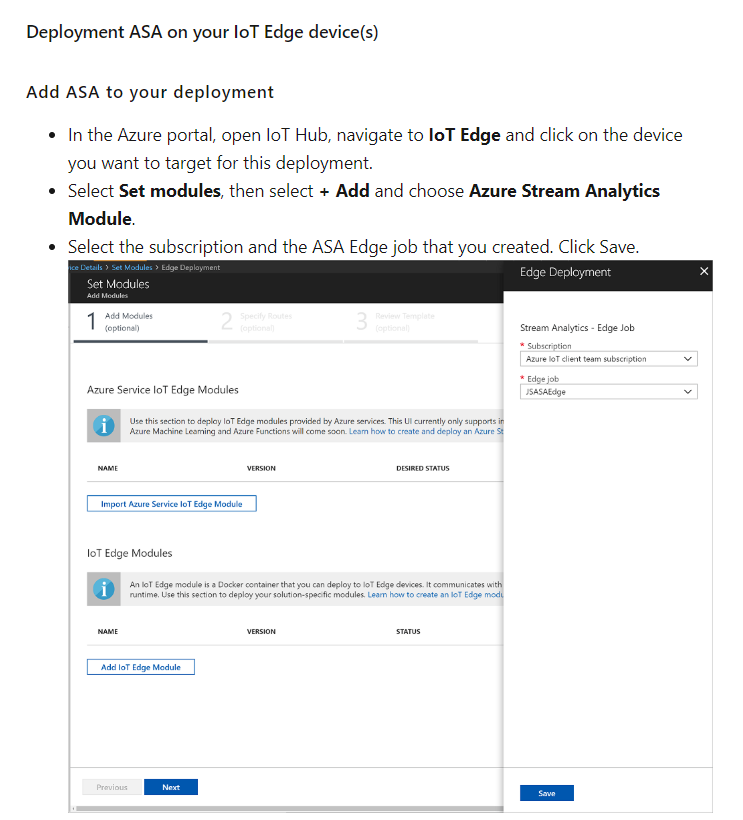
You need to have a storage container for the job definition.



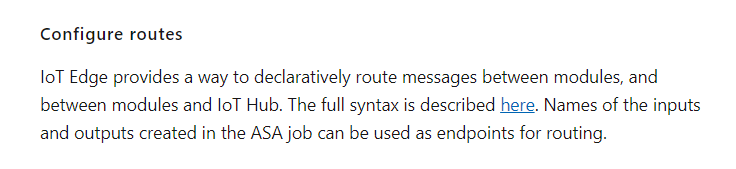
You also need to create a cloud part job definition.



You also need to set the modules for your IoT edge device.



You also need to configure the Routes.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on Stream Analytics on edge devices, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-edge>

### **Question 28**

Domain :Implement data storage solutions

Your company has 2 Azure SQL Databases named compdb1 and compdb2. Access needs to be configured for these databases from the following nodes

* A workstation which has an IP address of 5.78.99.4
* A set of IP addresses in the range of 5.78.99.6 - 5.78.99.10

The access needs to be set based on the following permissions

* Connections to both of the databases must be allowed from the workstation
* The specified IP address range must be allowed to connect to the database compdb1 and not compdb2
* The Web services in Azure must be able to connect to the database compdb1 and not compdb2

Which of the following must be set for this requirement? Choose 3 answers from the options given below

A.

**Create a firewall rule on the database compdb1 that has a start IP address of 5.78.99.6 and end IP address of 5.78.99.10**

B.

**Create a firewall rule on the database compdb1 that has a start and end IP address of 0.0.0.0**

C.

**Create a firewall rule on the server hosting both of the databases that has a start IP address of 5.78.99.6 and end IP address of 5.78.99.10**

D.

**Create a firewall rule on the database compdb1 that has a start and end IP address of 5.78.99.4**

E.

**Create a firewall rule on the server hosting both of the databases that has a start and end IP address of 5.78.99.4**

**Explanation:**

Answer – A, B, E

We can configure firewall rules at the database level.

The action of “Create a firewall rule on the database compdb1 that has a start IP address of 5.78.99.6 and end IP address of 5.78.99.10” will fulfil the requirement

**“The specified IP address range must be allowed to connect to the database compdb1 and not compdb2”**

The action of “Create a firewall rule on the database compdb1 that has a start and end IP address of 0.0.0.0” will fulfil the requirement

**“The Web services in Azure must be able to connect to the database compdb1 and not compdb2”**

The action of “Create a firewall rule on the server hosting both of the databases that has a start and end IP address of 5.78.99.4” will fulfil the requirement

**“Connections to both of the databases must be allowed from the workstation”**

Option C is incorrect since connections from this IP address should not be allowed on compdb2 as per the requirement

Option D is incorrect since we have to configure a server firewall rule to allow traffic from the workstation on both databases

For more information on working with the database firewall, please refer to the following link

* <https://docs.microsoft.com/en-us/azure/azure-sql/database/firewall-configure>

### **Question 29**

Domain :Monitor and optimize data solutions

A company is using an Azure SQL Data Warehouse Gen2. Users are complaining that performance is slow when they run commonly used queries. They do not report such issues for infrequently used queries. Which of the following should they monitor to find out the source of the performance issues?

]A.

**Cache used percentage**

]B.

**Memory percentage**

]C.

**CPU percentage**

]D.

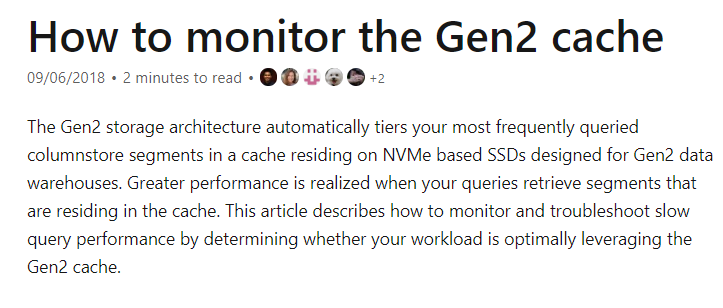
**Failed connections**

**Explanation:**

Answer - A

To check for issues on frequently used queries, you can look at the cache percentage used.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on monitoring Gen2 cache, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-how-to-monitor-cache>

### **Question 30**

Domain :Monitor and optimize data solutions

A company has implemented a real-time data analysis solution. This solution is making use of Azure Event Hub to ingest the data. The data is then sent to the Azure Stream Analytics cloud job. The cloud job has been configured to use 100 Streaming Units. Which of the following two actions can be performed to optimize the Azure Stream Analytics job's performance?

A.

**Scale up the Streaming Units of the job.**

B.

**Make use of event ordering.**

C.

**Make use of Azure Stream Analytics user-defined functions.**

D.

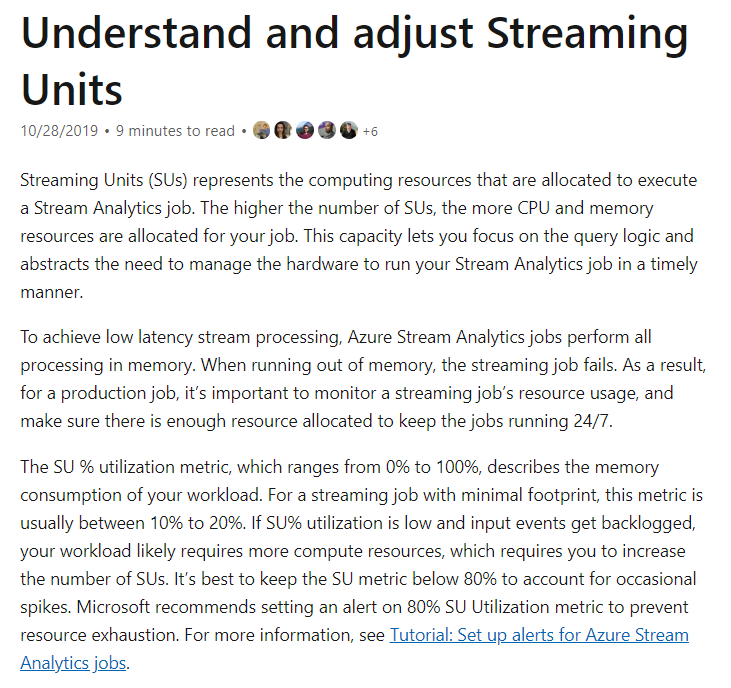
**Implement query parallelization by partitioning the data input.**

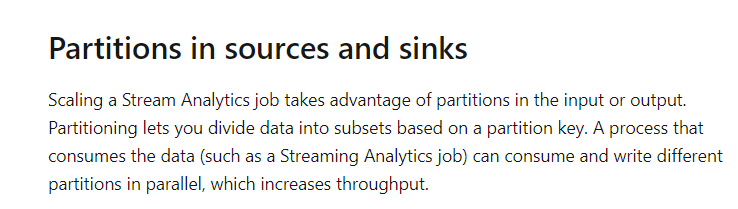
**Explanation:**

Answer – A and D

You can scale up the streaming units and also implement parallelization.

The Microsoft documentation mentions the following.





Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on stream analytics parallelization and scaling of stream analytic jobs, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-parallelization>
* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-scale-jobs>

### **Question 31**

Domain :Manage and develop data processing

[**View Case Study**](javascript:;)

A company wants to use a set of services on Azure. They want to make use of Platform-as-a-service products to create a new data pipeline process. They have the following requirements

**Data Ingestion**

* This layer must provide access to multiple sources
* This layer must provide the ability to orchestrate a workflow
* It must also provide the capability to run SQL Server Integration Service packages

**Storage**

* The storage layer must be optimized for Big Data workloads
* It must provide encryption of data at rest
* There must be no size constraints

**Prepare and Train**

* This layer must provide a fully managed interactive workspace for exploration and visualization
* Here you should be able to program in R, SQL or Scala
* It must provide seamless user authentication with Azure Active Directory

**Model and Service**

* This layer must provide support for SQL language
* It must implement native columnar storage

Which of the following should be used as a technology for the “Data Ingestion” layer?

]A.

**Azure Logic Apps**

]B.

**Azure Data Factory**

]C.

**Azure Automation**

]D.

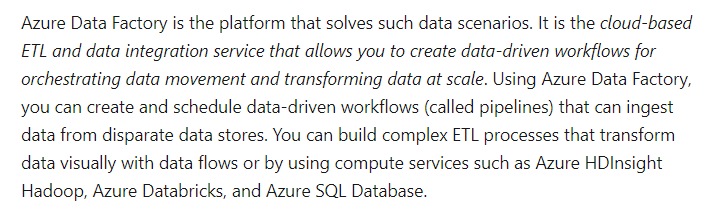
**Azure Functions**

**Explanation:**

Answer – B

Since you are looking at a data pipeline process, you must consider using Azure Data Factory. This can connect to multiple sources. You can define a workflow or pipeline and it can also run SQL Server Integration Service packages.

The Microsoft documentation mentions the following.



Since this is the perfect fit for the requirement, all other options are incorrect.

For more information on Azure Data Factory, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/introduction>

### **Question 32**

Domain :Manage and develop data processing

[**View Case Study**](javascript:;)

A company wants to use a set of services on Azure. They want to make use of Platform-as-a-service products to create a new data pipeline process. They have the following requirements

**Data Ingestion**

* This layer must provide access to multiple sources
* This layer must provide the ability to orchestrate a workflow
* It must also provide the capability to run SQL Server Integration Service packages

**Storage**

* The storage layer must be optimized for Big Data workloads
* It must provide encryption of data at rest
* There must be no size constraints

**Prepare and Train**

* This layer must provide a fully managed interactive workspace for exploration and visualization
* Here you should be able to program in R, SQL or Scala
* It must provide seamless user authentication with Azure Active Directory

**Model and Service**

* This layer must provide support for SQL language
* It must implement native columnar storage

Which of the following should be used as a technology for the “Storage” layer?

]A.

**Azure Data Lake Storage**

]B.

**Azure Blob Storage**

]C.

**Azure Files**

]D.

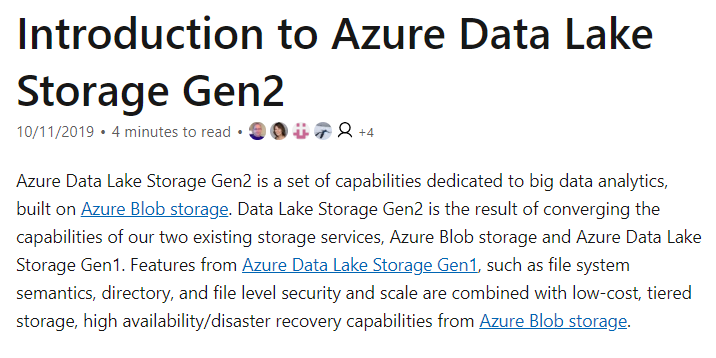
**Azure SQL Data warehouse**

**Explanation:**

Answer – A

Azure Data Lake Storage fulfills all of the right aspects as being built for Big Data Analytics. It can also scale in terms of storage.

The Microsoft documentation mentions the following.



Since this is the perfect fit for the requirement, all other options are incorrect.

For more information on Azure Data Lake Storage, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-introduction>

### **Question 33**

Domain :Manage and develop data processing

[**View Case Study**](javascript:;)

A company wants to use a set of services on Azure. They want to make use of Platform-as-a-service products to create a new data pipeline process. They have the following requirements

**Data Ingestion**

* This layer must provide access to multiple sources
* This layer must provide the ability to orchestrate a workflow
* It must also provide the capability to run SQL Server Integration Service packages

**Storage**

* The storage layer must be optimized for Big Data workloads
* It must provide encryption of data at rest
* There must be no size constraints

**Prepare and Train**

* This layer must provide a fully managed interactive workspace for exploration and visualization
* Here you should be able to program in R, SQL or Scala
* It must provide seamless user authentication with Azure Active Directory

**Model and Service**

* This layer must provide support for SQL language
* It must implement native columnar storage

Which of the following should be used as a technology for the “Prepare and Train” layer?

]A.

**HDInsight Apache Spark Cluster**

]B.

**Azure Databricks**

]C.

**HDInsight Apache Storm Cluster**

]D.

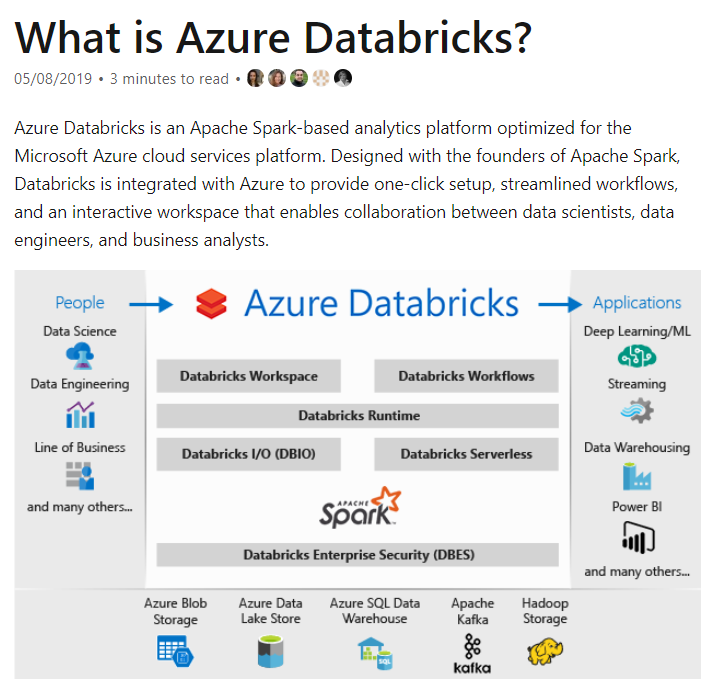
**Azure SQL Data warehouse**

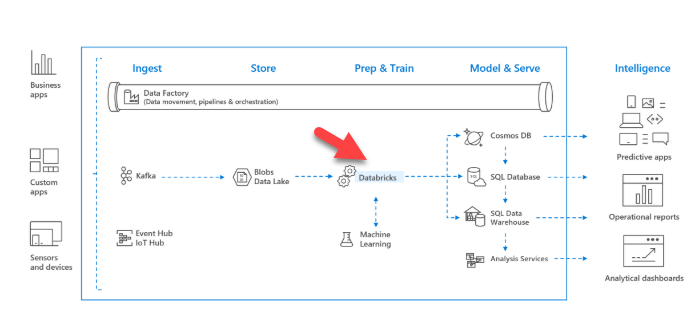
**Explanation:**

Answer – B

Azure Databricks is perfect for the Prepare and Train layer. Here you can perform interactive analysis using different programming languages.

The Microsoft documentation mentions the following.





Since this is the perfect fit for the requirement, all other options are incorrect.

For more information on Azure Databricks, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/azure-databricks/what-is-azure-databricks>

### **Question 34**

Domain :Manage and develop data processing

[**View Case Study**](javascript:;)

A company wants to use a set of services on Azure. They want to make use of Platform-as-a-service products to create a new data pipeline process. They have the following requirements

**Data Ingestion**

* This layer must provide access to multiple sources
* This layer must provide the ability to orchestrate a workflow
* It must also provide the capability to run SQL Server Integration Service packages

**Storage**

* The storage layer must be optimized for Big Data workloads
* It must provide encryption of data at rest
* There must be no size constraints

**Prepare and Train**

* This layer must provide a fully managed interactive workspace for exploration and visualization
* Here you should be able to program in R, SQL or Scala
* It must provide seamless user authentication with Azure Active Directory

**Model and Service**

* This layer must provide support for SQL language
* It must implement native columnar storage

Which of the following should be used as a technology for the “Model and Service” layer?

]A.

**HDInsight Apache Kafta cluster**

]B.

**Azure SQL Data warehouse**

]C.

**Azure Data Lake Storage**

]D.

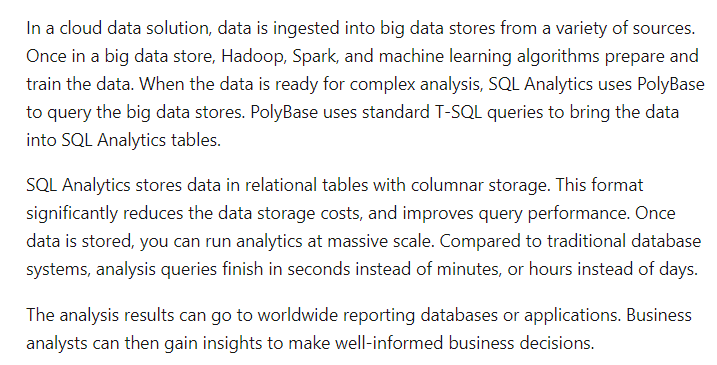
**Azure Blob Storage**

**Explanation:**

Answer – B

For columnar storage, you can make use of Azure SQL data warehouse.

The Microsoft documentation mentions the following.



Since this is the perfect fit for the requirement, all other options are incorrect.

For more information on Azure SQL data warehouse, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-overview-what-is>

### **Question 35**

Domain :Monitor and optimize data solutions

Your company has an Azure Cosmos DB Account that makes use of the SQL API. You have to ensure that all stale data is deleted from the database automatically.

Which of the following feature would you use for this requirement?

]A.

**Soft delete**

]B.

**Schema Read**

]C.

**Time to Live**

]D.

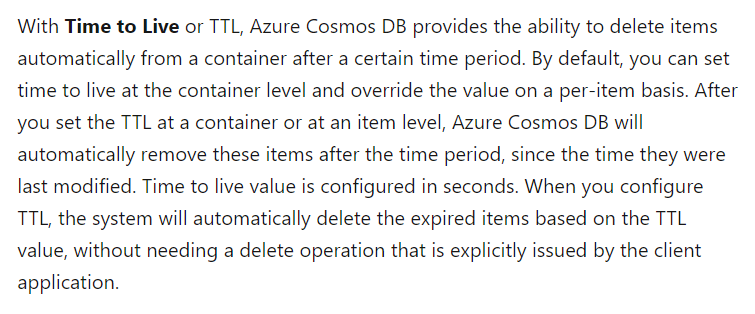
**CORS**

**Explanation:**

Answer – C

You can set a time to live for the items in a Cosmos DB database.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on the time to live feature, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/time-to-live>

### **Question 36**

Domain :Monitor and optimize data solutions

A company wants to make use of Azure Data Lake Gen 2 storage account. This would be used to store Big Data related to an application. The company wants to implement logging.

They decide to create an Azure Automation runbook which would be used to copy events.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

You need to make use of Azure Data Lake storage diagnostics for this purpose.

For more information on Azure Data Lake Gen 1 storage diagnostics, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-diagnostic-logs>

### **Question 37**

Domain :Monitor and optimize data solutions

A company wants to make use of Azure Data Lake Gen 2 storage account. This would be used to store Big Data related to an application. The company wants to implement logging.

They decide to use the information that is stored in Azure Active Directory reports.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

You need to make use of Azure Data Lake storage diagnostics for this purpose.

For more information on Azure Data Lake Gen 1 storage diagnostics, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-diagnostic-logs>

### **Question 38**

Domain :Monitor and optimize data solutions

A company wants to make use of Azure Data Lake Gen 2 storage account. This would be used to store Big Data related to an application. The company wants to implement logging.

They decide to configure Azure Data Lake Storage diagnostics to store the logs and metric data in a storage account.

Would this fulfill the requirement?

]A.**Yes**

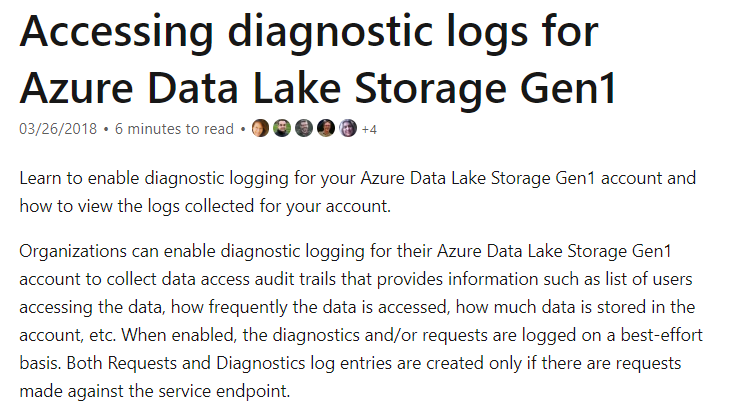
]B.**No**

**Explanation:**

Answer – A

Yes, this is the right approach.

The Microsoft documentation mentions the following.



For more information on Azure Data Lake Gen 1 storage diagnostics, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-diagnostic-logs>

### **Question 39**

Domain :Manage and develop data processing

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

Which of the following can be used to process and query the ingested data for the Tier 9 data?

]A.

**Azure Notification Hubs**

]B.

**Apache Cache for Redis**

]C.

**Azure Functions**

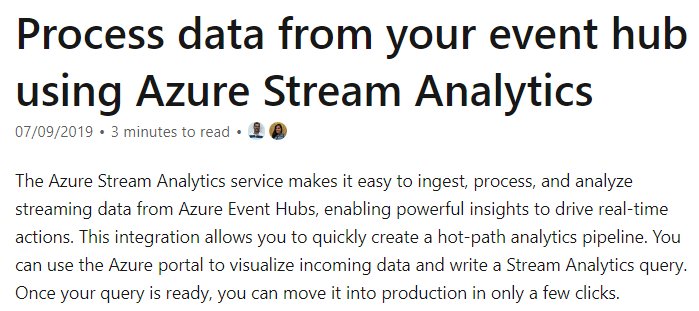
]D.

**Azure Stream Analytics**

**Explanation:**

Answer – D

One way is to use Azure Stream Analytics. The Microsoft documentation mentions the following.



Option A is incorrect since this is a Notification service.

Option B is incorrect since this is a cache service.

Option C is incorrect since this is a serverless compute service.

For more information on Azure Stream Analytics, one can visit the below URL-

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

### **Question 40**

Domain :Manage and develop data processing

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

The Azure Data Factory instance must meet the requirements to move the data from the On-premise SQL Servers to Azure. Which of the following would you use as the integration runtime?

]A.

**Self-hosted integration runtime**

]B.

**Azure-SSIS Integration runtime**

]C.

**.Net Common Language Runtime**

]D.

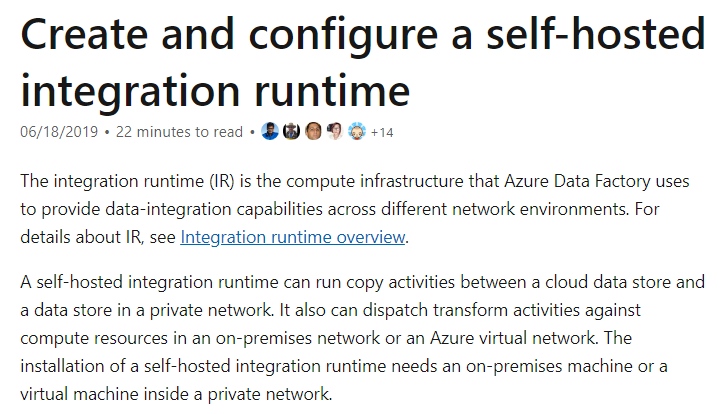
**Azure Integration runtime**

**Explanation:**

Answer – A

The self-hosted integration runtime can be used to move data between on-premise data stores to Azure cloud data stores.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the Microsoft documentation, all other options are incorrect.

For more information on self-hosted runtime environments, one can visit the below URL-

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

### **Question 41**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

The data for the external applications needs to be encrypted at rest. You decide to implement the following steps.

* Use the Always Encrypted Wizard in SQL Server Management Studio.
* Select the column that needs to be encrypted.
* Set the encryption type to Randomized.
* Configure the master key to be used from the Windows Certificate Store.
* Confirm the configuration and deploy the solution.

Would these steps fulfill the requirement?

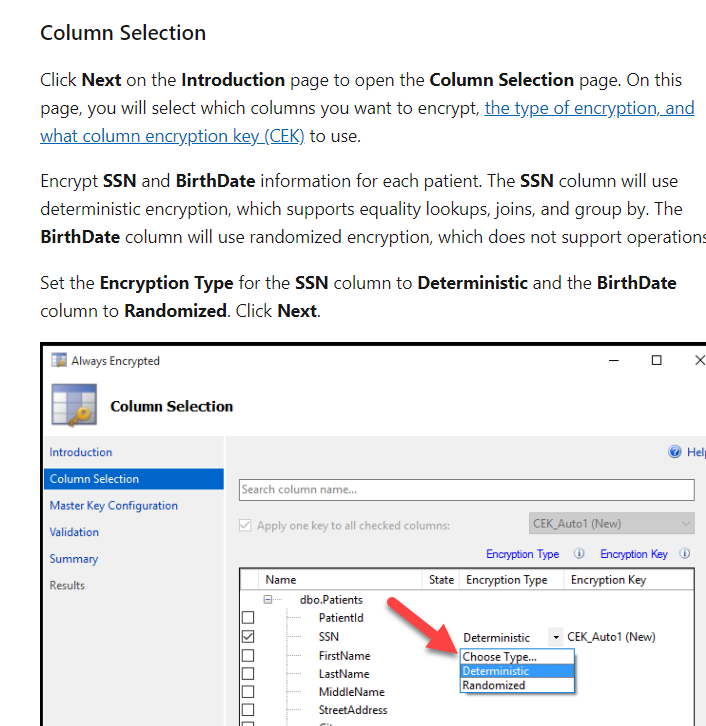
]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

As per the documentation, the encryption type needs to set as Deterministic when enabling Always Encrypted.



For more information on implementing Always Encrypted, one can visit the below URL-

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

### **Question 42**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

The data for the external applications needs to be encrypted at rest. You decide to implement the following steps.

* Use the Always Encrypted Wizard in SQL Server Management Studio.
* Select the column that needs to be encrypted.
* Set the encryption type to Deterministic.
* Configure the master key to be used from the Windows Certificate Store.
* Confirm the configuration and deploy the solution.

Would these steps fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – A

Yes, this is the right series of steps to implement Always Encrypted.

For more information on implementing Always Encrypted, one can visit the below URL-

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

### **Question 43**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

The data for the external applications needs to be encrypted at rest. You decide to implement the following steps.

* Use the Always Encrypted Wizard in SQL Server Management Studio.
* Select the column that needs to be encrypted.
* Set the encryption type to Deterministic.
* Configure the master key to be used from the Azure Key Vault.
* Confirm the configuration and deploy the solution.

Would these steps fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

As per the case study, all keys and certificates need to be managed in on-premise data stores.

For more information on implementing Always Encrypted, one can visit the below URL-

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

### **Question 44**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

Which of the following should you use as the masking function for Data type compA?

]A.

**Custom Text**

]B.

**Default**

]C.

**Email**

]D.

**Random number**

**Explanation:**

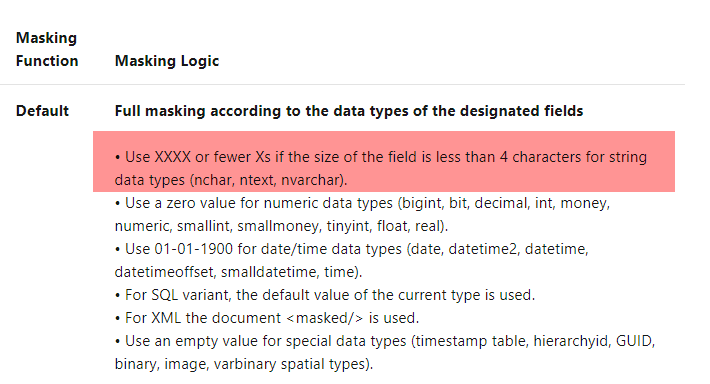
Answer – B

As per the case study, below is the requirement for the Data type.

* For Data type compA – Mask 4 or less string data type characters.

You can use the “Default” masking function for this requirement.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on dynamic data masking, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-dynamic-data-masking-get-started>

### **Question 45**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

Which of the following should you use as the masking function for Data type compB?

]A.

**Custom Text**

]B.

**Default**

]C.

**Email**

]D.

**Random number**

**Explanation:**

Answer C

As per the case study, below is the requirement for the Data type.

* For Data type compB – Expose the first letter and mask the domain.

You can use the “Email” masking function for this requirement.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on dynamic data masking, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-dynamic-data-masking-get-started>

### **Question 46**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

Which of the following should you use as the masking function for Data type compC?

]A.

**Custom Text**

]B.

**Default**

]C.

**Email**

]D.

**Random number**

**Explanation:**

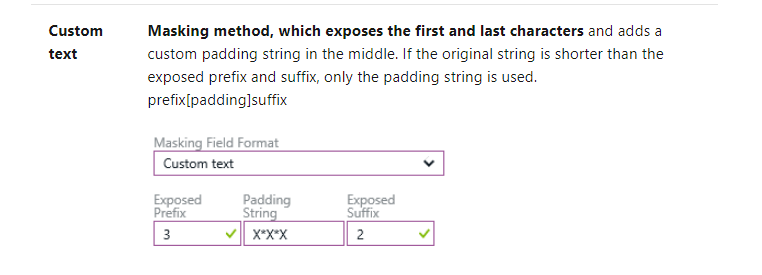
Answer - A

As per the case study, below is the requirement for the Data type.

* For Data type compC – Mask everything except characters at the beginning and the end.

You can use the “Custom Text” masking function for this requirement.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on dynamic data masking, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-dynamic-data-masking-get-started>

### **Question 47**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

You need to implement the following requirement as per the case study.

* The Application access for Tier 7 and 8 must be restricted to the database only.

Which of the following steps would you implement for this requirement? Choose 3 answers from the options given below.

A.

**Use Azure PowerShell to create a database firewall rule.**

B.

**Configure the setting of “Allow Azure Services to Access Server” to Disabled.**

C.

**Configure the setting of “Allow Azure Services to Access Server” to Enabled.**

D.

**Create a database firewall rule from the Azure portal.**

E.

**Create a server firewall rule from the Azure portal.**

F.

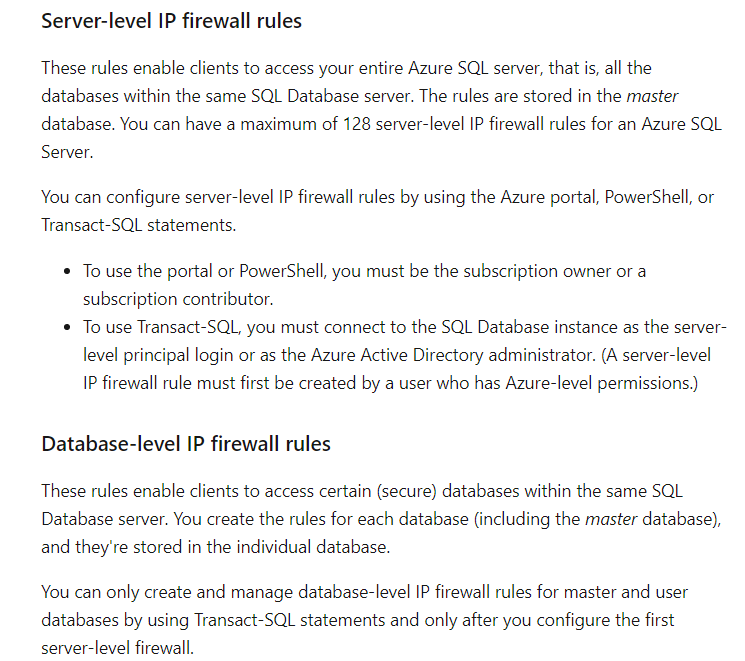
**Use Transact-SQL to create a database firewall rule.**

**Explanation:**

Answer – B, E and F

You can set database and firewall rules to restrict access to the server and the database.

The Microsoft documentation mentions the following.



Also, ensure to set the “Allow Azure Services to Access Server.” It is set to Disabled so that no other service can access the database.

Options A and D are incorrect since you can only create a database firewall rule via Transact-SQL.

Option C is incorrect since the setting “Allow Azure Services to Access Server” should be Disabled.

For more information on server and database rules for Azure SQL databases, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-firewall-configure>

### **Question 48**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

You have to implement logging for monitoring the data warehousing solution. Which of the following would you log?

]A.

**RequestSteps**

]B.

**DmsWorkers**

]C.

**SQLRequests**

]D.

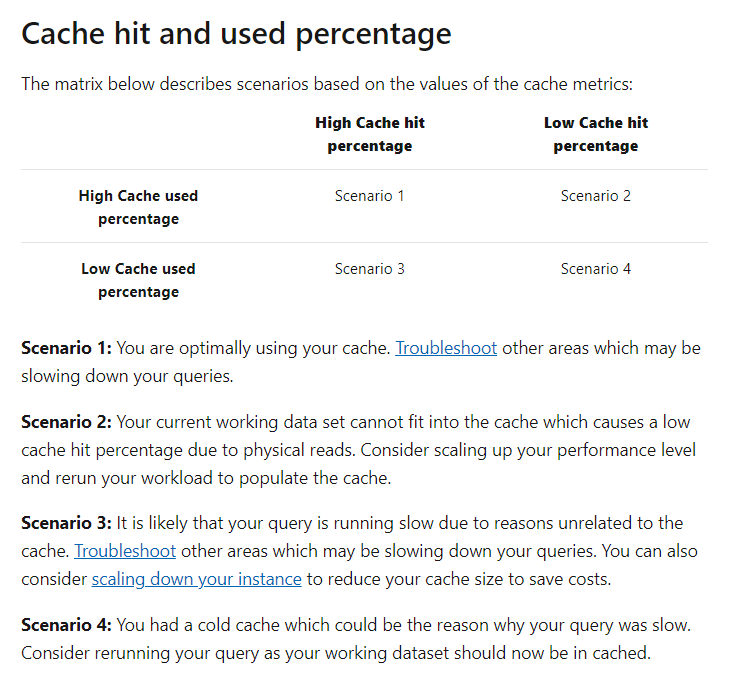
**ExecRequests**

**Explanation:**

Answer – C

Since the SQL requests would affect the cache, these requests need to be monitored.

The Microsoft documentation mentions the following on caching.



Since this is the ideal metric to monitor, all other options are incorrect.

For more information on monitoring the cache, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-how-to-monitor-cache>

### **Question 49**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider.

**Current Environment**

The company currently has Microsoft SQL databases that are split into different categories or tiers. Some of the databases are used by Internal users, some by external partners and external distributions.

Below is the list of applications, tiers and their individual requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Applications** | **Tier** | **Replication** | **Comments** |
| **Internal Comp** | 1 | Yes |  |
| **Internal Comp** | 2 | Using SQL Data Sync |  |
| **Internal Partner** | 3 | Yes | Data is replicated to the Partner |
| **External Comp** | 4,5,6 | Yes |  |
| **External Partner** | 7,8 | No | This is a Partner managed database |
| **Internal Distribution and Sales** | 9 | Yes, but only when the data is ingested at one of the branch offices | Data is ingested from Comp branch offices |
| **External Distribution and Sales** | 10 | Yes, but only once the data is ingested at the Comp main office | Data is ingested from multiple sources |

Below are the current requirements of the company

* The databases in Tier 3, Tier 6 to 8 must use a database density on the same server and Elastic pools in cost effective manner
* The Applications must have access to data from internal and external sources whilst ensuring data is encrypted at rest and in transit
* The databases in Tier 3, Tier 6 to 8 must have a recovery strategy for in case whenever the server goes offline
* The Tier 1 applications must have their databases stored on Premium P2 tier
* The Tier 1 applications must have their databases stored on Standard S4 tier
* Data will be migrated from the on-premise databases to Azure SQL Databases using Azure Data Factory. The pipeline must support continued data movement and migration.
* The Application access for Tier 7 and 8 must be restricted to the database only
* For Tier 4 and Tier 5 databases, the backup strategy must include the following
  + Transactional log backup every hour
  + Differential backup every day
  + Full backup every week
* Backup strategies must be in place for all standalone Azure SQL databases using methods available with Azure SQL databases.
* Tier 1 database must implement the following data masking logic
  + For Data type compA – Mask 4 or less string data type characters
  + For Data type compB – Expose the first letter and mask the domain
  + For Data type compC – Mask everything except characters at the beginning and the end
* All certificates and keys are internally managed in on-premise data stores
* For Tier 2 databases, if there are any conflicts between the data transfer from on-premise, preference should be given to on-premise data.
* Monitoring must be setup on every database
* Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.
* Azure SQL Data warehouse would be used to gather data from multiple internal and external databases.
* The Azure SQL Data warehouse must be optimized to use data from its cache
* The below metrics must be available when it comes to the cache
  + Metric compA – Low cache hit %, high cache usage %
  + Metric compB – Low cache hit %, low cache usage %
  + Metric compC – high cache hit %, high cache usage %
* The reporting data for external partners must be stored in Azure storage. The data should be made available during regular business hours in connecting regions.
* The reporting for Tier 9 needs to be moved to Event Hubs.
* The reporting for Tier 10 needs to be moved to Azure Blobs.

The following issues have been identified in the setup

* The External partners have control over the data formats, types and schemas
* For External based clients, the queries can’t be changed or optimized
* The database development staff are familiar with T-SQL language
* Because of the size and amount of data, some applications and reporting features are not performing at SLA levels.

You need to fulfill the below requirement of the case study.

**“Applications with Tiers 6 through 8 must ensure that unexpected resource storage usage is immediately reported to IT data engineers.”**

Which of the following would you implement for this requirement?

]A.

**An alert rule that would be used to monitor CPU percentage for the database and then alert the IT Engineers**

]B.

**An alert rule that would be used to monitor CPU percentage for the elastic pool and then alert the IT Engineers**

]C.

**An alert rule that would be used to monitor storage percentage for the database and then alert the IT Engineers**

]D.

**An alert rule that would be used to monitor storage percentage for the elastic pool and then alert the IT Engineers**

**Explanation:**

Answer – D

Since the requirement asks for monitoring the storage, we should set this. Also, since the databases are going to be part of an elastic pool, we need to set it to monitor the percentage for the entire elastic pool.

For more information on working with alerts, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/azure-monitor/platform/alerts-metric>

### **Question 50**

Domain :Implement data storage solutions

You have to access Azure Blob Storage from Azure Databricks using secrets stored in a key vault. You already have the storage account, the blob container and Azure key vault in place.

You decide to implement the following steps.

* Add the secret to the storage container.
* Create a Databricks workspace and add the access keys.
* Access the blob container from Azure Databricks.

Would these steps fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

You need to add the secret to Azure Key Vault and add the secret scope to the Databricks workspace.

For more information on accessing Azure Blob storage from Azure Databricks using Azure Key Vault, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/azure-databricks/store-secrets-azure-key-vault>

### **Question 51**

Domain :Implement data storage solutions

You have to access Azure Blob Storage from Azure Databricks using secrets stored in a key vault. You already have the storage account, the blob container and Azure key vault in place.

You decide to implement the following steps.

* Add the secret to the key vault.
* Create a Databricks workspace and add the secret scope.
* Access the blob container from Azure Databricks.

Would these steps fulfill the requirement?

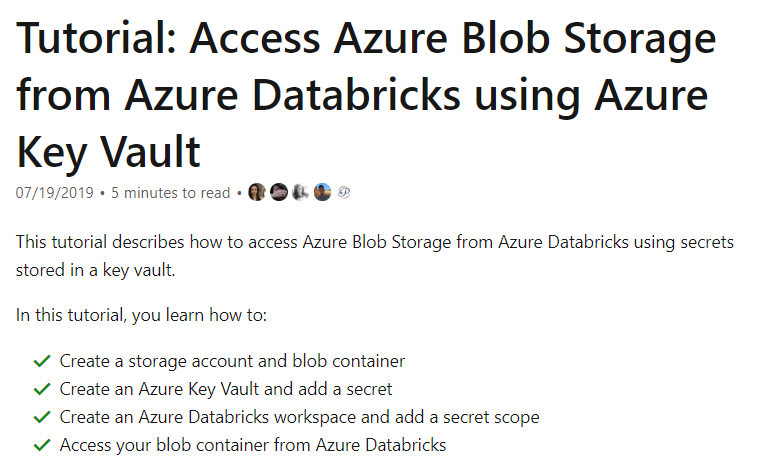
]A.**Yes**

]B.**No**

**Explanation:**

Answer – A

Yes, this would fulfill the requirement. The Microsoft documentation mentions the following.



For more information on accessing Azure Blob storage from Azure Databricks using Azure Key Vault, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/azure-databricks/store-secrets-azure-key-vault>

### **Question 52**

Domain :Implement data storage solutions

You have to access Azure Blob Storage from Azure Databricks using secrets stored in a key vault. You already have the storage account, the blob container and Azure key vault in place.

You decide to implement the following steps.

* Add the secret to the key vault.
* Create a Databricks workspace and add the access keys.
* Access the blob container from Azure Databricks.

Would these steps fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

You are supposed to add a secret scope to the Databricks workspace and not the access keys.

For more information on accessing Azure Blob storage from Azure Databricks using Azure Key Vault, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/azure-databricks/store-secrets-azure-key-vault>

### **Question 53**

Domain :Manage and develop data processing

A company has created an Azure Data Lake Gen 2 storage account. They want to ingest data into the storage account from various data sources.

Which of the following can they use to ingest data from a relational data store?

]A.

**Azure Data Factory**

]B.

**AzCopy Tool**

]C.

**Azure Event Hubs**

]D.

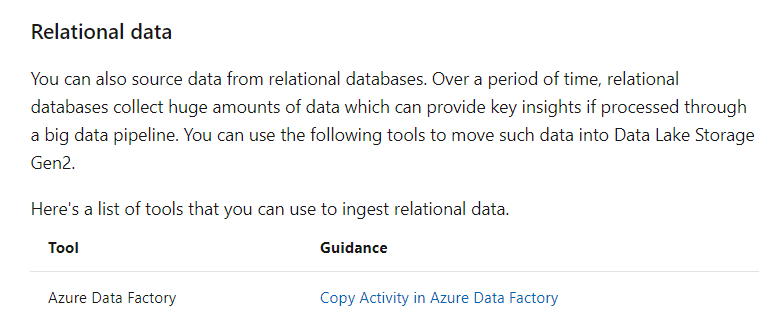
**Azure Event Grid**

**Explanation:**

Answer – A

You can use Azure Data Factory for this requirement.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the Microsoft documentation, all other options are incorrect.

For more information on data lake storage scenarios, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-data-scenarios>

### **Question 54**

Domain :Manage and develop data processing

A company has created an Azure Data Lake Gen 2 storage account. They want to ingest data into the storage account from various data sources.

Which of the following can they use to ingest data from a local workstation?

]A.

**Azure Data Factory**

]B.

**AzCopy Tool**

]C.

**Azure Event Hubs**

]D.

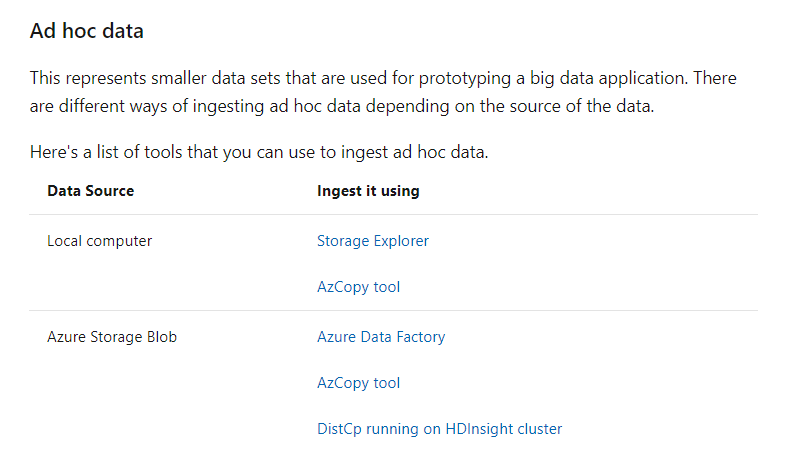
**Azure Event Grid**

**Explanation:**

Answer – B

You can use the AzCopy tool for this requirement.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the Microsoft documentation, all other options are incorrect.

For more information on data lake storage scenarios, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-data-scenarios>

### **Question 55**

Domain :Manage and develop data processing

A company has created an Azure Data Lake Gen 2 storage account. They want to ingest data into the storage account from various data sources.

Which of the following can they use to ingest data from log data stored on web servers?

]A.

**Azure Data Factory**

]B.

**AzCopy Tool**

]C.

**Azure Event Hubs**

]D.

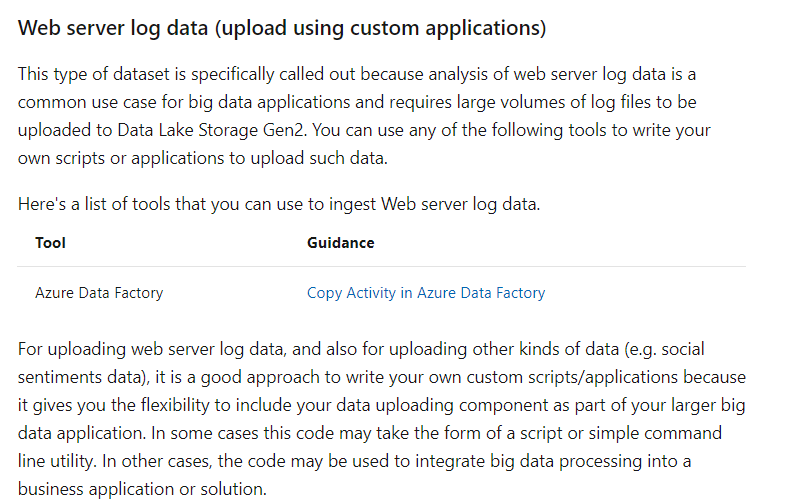
**Azure Event Grid**

**Explanation:**

Answer – A

You can use Azure Data Factory for this requirement.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the Microsoft documentation, all other options are incorrect.

For more information on data lake storage scenarios, one can visit the below URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-data-scenarios>