

Exam Questions DP-201

Designing an Azure Data Solution

<https://www.2passeasy.com/dumps/DP-201/>



NEW QUESTION 1

- (Exam Topic 1)

You need to design the vehicle images storage solution. What should you recommend?

- A. Azure Media Services
- B. Azure Premium Storage account
- C. Azure Redis Cache
- D. Azure Cosmos DB

Answer: B

Explanation:

Premium Storage stores data on the latest technology Solid State Drives (SSDs) whereas Standard Storage stores data on Hard Disk Drives (HDDs). Premium Storage is designed for Azure Virtual Machine workloads which require consistent high IO performance and low latency in order to host IO intensive workloads like OLTP, Big Data, and Data Warehousing on platforms like SQL Server, MongoDB, Cassandra, and others. With Premium Storage, more customers will be able to lift-and-shift demanding enterprise applications to the cloud.

Scenario: Traffic sensors will occasionally capture an image of a vehicle for debugging purposes. You must optimize performance of saving/storing vehicle images. The impact of vehicle images on sensor data throughout must be minimized. References:

<https://azure.microsoft.com/es-es/blog/introducing-premium-storage-high-performance-storage-for-azure-virtual>

NEW QUESTION 2

- (Exam Topic 1)

You need to design the runtime environment for the Real Time Response system. What should you recommend?

- A. General Purpose nodes without the Enterprise Security package
- B. Memory Optimized Nodes without the Enterprise Security package
- C. Memory Optimized nodes with the Enterprise Security package
- D. General Purpose nodes with the Enterprise Security package

Answer: B

NEW QUESTION 3

- (Exam Topic 1)

STION NO: 5 HOTSPOT

You need to design the authentication and authorization methods for sensors.

What should you recommend? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Requirement	Method
Authentication	<div style="border: 1px solid gray; padding: 5px;"> <div style="display: flex; justify-content: space-between; align-items: center;"> ▼ </div> <ul style="list-style-type: none"> HMAC header Resource Token Azure Managed Identity Storage account connection string </div>
Authorization	<div style="border: 1px solid gray; padding: 5px;"> <div style="display: flex; justify-content: space-between; align-items: center;"> ▼ </div> <ul style="list-style-type: none"> Custom RBAC role Cosmos DB user Azure Active Directory user IoT device identity </div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Sensor data must be stored in a Cosmos DB named treydata in a collection named SensorData Sensors must have permission only to add items to the SensorData collection

Box 1: Resource Token

Resource tokens provide access to the application resources within a Cosmos DB database.

Enable clients to read, write, and delete resources in the Cosmos DB account according to the permissions they've been granted.

Box 2: Cosmos DB user

You can use a resource token (by creating Cosmos DB users and permissions) when you want to provide access to resources in your Cosmos DB account to a client that cannot be trusted with the master key.

References:

<https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data>

NEW QUESTION 4

- (Exam Topic 2)

You need to design the image processing solution to meet the optimization requirements for image tag data. What should you configure? To answer, drag the appropriate setting to the correct drop targets.

Each source may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Tagging data must be uploaded to the cloud from the New York office location.
 Tagging data must be replicated to regions that are geographically close to company office locations.

NEW QUESTION 5

- (Exam Topic 3)

You need to design a solution to meet the SQL Server storage requirements for CONT_SQL3. Which type of disk should you recommend?

- A. Standard SSD Managed Disk
- B. Premium SSD Managed Disk
- C. Ultra SSD Managed Disk

Answer: C

Explanation:

CONT_SQL3 requires an initial scale of 35000 IOPS.

Disk size (GiB)	4	8	16	32	64	128	256	512	1,024-65,536 (in increments of 1 TiB)
IOPS range	100-1,200	100-2,400	100-4,800	100-9,600	100-19,200	100-38,400	100-76,800	100-153,600	100-160,000
Throughput Cap (MBps)	300	600	1,200	2,000	2,000	2,000	2,000	2,000	2,000

The following table provides a comparison of ultra solid-state-drives (SSD) (preview), premium SSD, standard SSD, and standard hard disk drives (HDD) for managed disks to help you decide what to use.

	Ultra SSD (preview)	Premium SSD	Standard SSD	Standard HDD
Disk type	SSD	SSD	SSD	HDD
Scenario	IO-intensive workloads such as SAP HANA, top tier databases (for example, SQL Oracle), and other transaction-heavy workloads.	Production and performance sensitive workloads	Web servers, lightly used enterprise applications and dev/test	Backup, non-critical, infrequent access
Disk size	65,536 gibibyte (GiB) (Preview)	32,767 GiB	32,767 GiB	32,767 GiB
Max throughput	2,000 MiB/s (Preview)	900 MiB/s	750 MiB/s	500 MiB/s
Max IOPS	160,000 (Preview)	20,000	6,000	2,000

References:

<https://docs.microsoft.com/en-us/azure/virtual-machines/windows/disks-types>

NEW QUESTION 6

- (Exam Topic 3)

You need to recommend an Azure SQL Database service tier. What should you recommend?

- A. Business Critical
- B. General Purpose
- C. Premium
- D. Standard
- E. Basic

Answer: C

Explanation:

The data engineers must set the SQL Data Warehouse compute resources to consume 300 DWUs. Note: There are three architectural models that are used in Azure SQL Database:

- General Purpose/Standard
- Business Critical/Premium
- Hyperscale

NEW QUESTION 7

- (Exam Topic 3)

You are designing an Azure SQL Data Warehouse for a financial services company. Azure Active Directory will be used to authenticate the users.

You need to ensure that the following security requirements are met:

- Department managers must be able to create new database.
- The IT department must assign users to databases.
- Permissions granted must be minimized.

Which role memberships should you recommend? To answer, drag the appropriate roles to the correct groups. Each role may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

Roles	Group	Role
dbmanager	Department managers	<input type="text"/>
loginmanager	IT	<input type="text"/>
dc_admin		
db_securityadmin		
db_owner		
db_accessadmin		

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: dbmanager

Members of the dbmanager role can create new databases. Box 2: db_accessadmin

Members of the db_accessadmin fixed database role can add or remove access to the database for Windows logins, Windows groups, and SQL Server logins.

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-manage-logins>

NEW QUESTION 8

- (Exam Topic 4)

You design data engineering solutions for a company.

You must integrate on-premises SQL Server data into an Azure solution that performs Extract-Transform-Load (ETL) operations have the following requirements:

- Develop a pipeline that can integrate data and run notebooks.
- Develop notebooks to transform the data.
- Load the data into a massively parallel processing database for later analysis. You need to recommend a solution.

What should you recommend? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Requirement	Service
Integrate the on-premises data into the cloud.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch
Develop notebooks to transform the data.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch
Run notebooks.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch
Load the data.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch
Store the transformed data.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Requirement	Service
Integrate the on-premises data into the cloud.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch
Develop notebooks to transform the data.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch
Run notebooks.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch
Load the data.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch
Store the transformed data.	<ul style="list-style-type: none"> Azure Databricks Azure Data Factory Azure SQL Data Warehouse Azure Batch

NEW QUESTION 9

- (Exam Topic 4)

You are designing a data processing solution that will implement the lambda architecture pattern. The solution will use Spark running on HDInsight for data

processing.

You need to recommend a data storage technology for the solution.

Which two technologies should you recommend? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

- A. Azure Cosmos DB
- B. Azure Service Bus
- C. Azure Storage Queue
- D. Apache Cassandra
- E. Kafka HDInsight

Answer: AE

Explanation:

To implement a lambda architecture on Azure, you can combine the following technologies to accelerate realtime big data analytics:

Azure Cosmos DB, the industry's first globally distributed, multi-model database service.

Apache Spark for Azure HDInsight, a processing framework that runs large-scale data analytics applications

Azure Cosmos DB change feed, which streams new data to the batch layer for HDInsight to process The Spark to Azure Cosmos DB Connector

E: You can use Apache Spark to stream data into or out of Apache Kafka on HDInsight using DStreams. References:

<https://docs.microsoft.com/en-us/azure/cosmos-db/lambda-architecture>

NEW QUESTION 10

- (Exam Topic 4)

A company installs IoT devices to monitor its fleet of delivery vehicles. Data from devices is collected from Azure Event Hub.

The data must be transmitted to Power BI for real-time data visualizations. You need to recommend a solution.

What should you recommend?

- A. Azure HDInsight with Spark Streaming
- B. Apache Spark in Azure Databricks
- C. Azure Stream Analytics
- D. Azure HDInsight with Storm

Answer: C

Explanation:

Step 1: Get your IoT hub ready for data access by adding a consumer group.

Step 2: Create, configure, and run a Stream Analytics job for data transfer from your IoT hub to your Power BI account.

Step 3: Create and publish a Power BI report to visualize the data. References:

<https://docs.microsoft.com/en-us/azure/iot-hub/iot-hub-live-data-visualization-in-power-bi>

NEW QUESTION 10

- (Exam Topic 4)

You are designing a data processing solution that will run as a Spark job on an HDInsight cluster. The solution will be used to provide near real-time information about online ordering for a retailer.

The solution must include a page on the company intranet that displays summary information. The summary information page must meet the following requirements:

- ▶ Display a summary of sales to date grouped by product categories, price range, and review scope.
- ▶ Display sales summary information including total sales, sales as compared to one day ago and sales as compared to one year ago.
- ▶ Reflect information for new orders as quickly as possible. You need to recommend a design for the solution.

What should you recommend? To answer, select the appropriate configuration in the answer area.

Use case	Technology
Data abstraction	<div style="border: 1px solid gray; padding: 2px;"> ▼ Resilient Distributed Dataset (RDD) Dataset DataFrame </div>
Data format	<div style="border: 1px solid gray; padding: 2px;"> ▼ Avro parquet </div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: DataFrame

DataFrames

Best choice in most situations.

Provides query optimization through Catalyst. Whole-stage code generation.

Direct memory access.

Low garbage collection (GC) overhead.

Not as developer-friendly as DataSets, as there are no compile-time checks or domain object programming. Box 2: parquet

The best format for performance is parquet with snappy compression, which is the default in Spark 2.x. Parquet stores data in columnar format, and is highly optimized in Spark.

NEW QUESTION 14

- (Exam Topic 4)

You are designing an Azure Databricks cluster that runs user-defined local processes. You need to recommend a cluster configuration that meets the following requirements:

- Minimize query latency.
- Reduce overall costs.
- Maximize the number of users that can run queries on the cluster at the same time. Which cluster type should you recommend?

- A. Standard with Autoscaling
- B. High Concurrency with Auto Termination
- C. High Concurrency with Autoscaling
- D. Standard with Auto Termination

Answer: C

Explanation:

High Concurrency clusters allow multiple users to run queries on the cluster at the same time, while minimizing query latency. Autoscaling clusters can reduce overall costs compared to a statically-sized cluster.

References:

<https://docs.azuredatabricks.net/user-guide/clusters/create.html> <https://docs.azuredatabricks.net/user-guide/clusters/high-concurrency.html#high-concurrency>

<https://docs.azuredatabricks.net/user-guide/clusters/terminate.html> <https://docs.azuredatabricks.net/user-guide/clusters/sizing.html#enable-and-configure-autoscaling>

NEW QUESTION 18

- (Exam Topic 4)

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

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

You are designing an HDInsight/Hadoop cluster solution that uses Azure Data Lake Gen1 Storage. The solution requires POSIX permissions and enables diagnostics logging for auditing.

You need to recommend solutions that optimize storage.

Proposed Solution: Implement compaction jobs to combine small files into larger files. Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

Depending on what services and workloads are using the data, a good size to consider for files is 256 MB or greater. If the file sizes cannot be batched when landing in Data Lake Storage Gen1, you can have a separate compaction job that combines these files into larger ones.

Note: POSIX permissions and auditing in Data Lake Storage Gen1 comes with an overhead that becomes apparent when working with numerous small files. As a best practice, you must batch your data into larger files versus writing thousands or millions of small files to Data Lake Storage Gen1. Avoiding small file sizes can have multiple benefits, such as:

Lowering the authentication checks across multiple files
Reduced open file connections

Faster copying/replication

Fewer files to process when updating Data Lake Storage Gen1 POSIX permissions

References:
<https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-best-practices>

NEW QUESTION 21

- (Exam Topic 4)

You design data engineering solutions for a company.

A project requires analytics and visualization of large set of data. The project has the following requirements:

- Notebook scheduling
- Cluster automation
- Power BI Visualization

You need to recommend the appropriate Azure service. Which Azure service should you recommend?

- A. Azure Batch
- B. Azure Stream Analytics
- C. Azure ML Studio
- D. Azure Databricks
- E. Azure HDInsight

Answer: D

Explanation:

A databrick job is a way of running a notebook or JAR either immediately or on a scheduled basis.

Azure Databricks has two types of clusters: interactive and job. Interactive clusters are used to analyze data collaboratively with interactive notebooks. Job clusters are used to run fast and robust automated workloads using the UI or API.

You can visualize Data with Azure Databricks and Power BI Desktop.

References:

<https://docs.azuredatabricks.net/user-guide/clusters/index.html> <https://docs.azuredatabricks.net/user-guide/jobs.html>

NEW QUESTION 24

- (Exam Topic 4)

You need to design the storage for the telemetry capture system. What storage solution should you use in the design?

- A. Azure Databricks

- B. Azure SQL Data Warehouse
- C. Azure Cosmos DB

Answer: C

NEW QUESTION 26

- (Exam Topic 4)

You are designing an application. You plan to use Azure SQL Database to support the application. The application will extract data from the Azure SQL Database and create text documents. The text documents will be placed into a cloud-based storage solution. The text storage solution must be accessible from an SMB network share. You need to recommend a data storage solution for the text documents. Which Azure data storage type should you recommend?

- A. Queue
- B. Files
- C. Blob
- D. Table

Answer: B

Explanation:

Azure Files enables you to set up highly available network file shares that can be accessed by using the standard Server Message Block (SMB) protocol.

References:

<https://docs.microsoft.com/en-us/azure/storage/common/storage-introduction> <https://docs.microsoft.com/en-us/azure/storage/tables/table-storage-overview>

NEW QUESTION 29

- (Exam Topic 4)

A company stores large datasets in Azure, including sales transactions and customer account information. You must design a solution to analyze the data. You plan to create the following HDInsight clusters:

You need to ensure that the clusters support the query requirements.

Which cluster types should you recommend? To answer, select the appropriate configuration in the answer area.

NOTE: Each correct selection is worth one point.

Cluster	Cluster type
Sales	<div style="border: 1px solid black; padding: 5px;"> <div style="display: flex; justify-content: space-between; align-items: center;"> Storm ▼ </div> <div style="border: 1px solid black; padding: 2px;"> <p>Storm</p> <p>Hadoop</p> <p>Interactive Query</p> <p>Kafka</p> </div> </div>
Accounts	<div style="border: 1px solid black; padding: 5px;"> <div style="display: flex; justify-content: space-between; align-items: center;"> Spark ▼ </div> <div style="border: 1px solid black; padding: 2px;"> <p>Spark</p> <p>Hadoop</p> <p>Interactive Query</p> <p>Kafka</p> </div> </div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Interactive Query

Choose Interactive Query cluster type to optimize for ad hoc, interactive queries. Box 2: Hadoop

Choose Apache Hadoop cluster type to optimize for Hive queries used as a batch process.

Note: In Azure HDInsight, there are several cluster types and technologies that can run Apache Hive queries. When you create your HDInsight cluster, choose the appropriate cluster type to help optimize performance for your workload needs.

For example, choose Interactive Query cluster type to optimize for ad hoc, interactive queries. Choose Apache Hadoop cluster type to optimize for Hive queries used as a batch process. Spark and HBase cluster types can also run Hive queries.

References:

<https://docs.microsoft.com/bs-latn-ba/azure/hdinsight/hdinsight-hadoop-optimize-hive-query?toc=%2Fko-kr%2>

NEW QUESTION 32

- (Exam Topic 4)

You are designing a Spark job that performs batch processing of daily web log traffic.

When you deploy the job in the production environment, it must meet the following requirements:

- Run once a day.
- Display status information on the company intranet as the job runs. You need to recommend technologies for triggering and monitoring jobs.

Which technologies should you recommend? To answer, drag the appropriate technologies to the correct locations. Each technology may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Technologies

Livy

Beeline

Azure Logic App

Azure API App

Requirement

Technology

Triggering of jobs

Monitoring of jobs

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Livy

You can use Livy to run interactive Spark shells or submit batch jobs to be run on Spark. Box 2: Beeline

Apache Beeline can be used to run Apache Hive queries on HDInsight. You can use Beeline with Apache Spark.

Note: Beeline is a Hive client that is included on the head nodes of your HDInsight cluster. Beeline uses JDBC to connect to HiveServer2, a service hosted on your HDInsight cluster. You can also use Beeline to access Hive on HDInsight remotely over the internet.

References:

<https://docs.microsoft.com/en-us/azure/hdinsight/spark/apache-spark-livy-rest-interface> <https://docs.microsoft.com/en-us/azure/hdinsight/hadoop/apache-hadoop-use-hive-beeline>

NEW QUESTION 34

- (Exam Topic 4)

You need to design the system for notifying law enforcement officers about speeding vehicles.

How should you design the pipeline? To answer, drag the appropriate services to the correct locations. Each service may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.



- A. Mastered
- B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 39

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual DP-201 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the DP-201 Product From:

<https://www.2passeasy.com/dumps/DP-201/>

Money Back Guarantee

DP-201 Practice Exam Features:

- * DP-201 Questions and Answers Updated Frequently
- * DP-201 Practice Questions Verified by Expert Senior Certified Staff
- * DP-201 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * DP-201 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year