BDS-C00^{Q&As}

AWS Certified Big Data - Speciality (BDS-C00)

Pass Amazon BDS-C00 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.leads4pass.com/aws-certified-big-data-specialty.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Amazon Official Exam Center

Instant Download After Purchase

100% Money Back Guarantee

365 Days Free Update

800,000+ Satisfied Customers

QUESTION 1

Because of the extensibility limitations of striped storage attached to Windows Server, Amazon RDS does not currently support increasing storage on a _____ DB Instance.

A. SQL Server

B. MySQL

C. Oracle

Correct Answer: A

QUESTION 2

Company A operates in Country X. Company A maintains a large dataset of historical purchase orders that contains personal data of their customers in the form of full names and telephone numbers. The dataset consists of 5 text files, 1TB each. Currently the dataset resides on-premises due to legal requirements of storing personal data in-country. The research and development department needs to run a clustering algorithm on the dataset and wants to use Elastic Map Reduce service in the closest AWS region. Due to geographic distance, the minimum latency between the on-premises system and the closet AWS region is 200 ms.

Which option allows Company A to do clustering in the AWS Cloud and meet the legal requirement of maintaining personal data in-country?

A. Anonymize the personal data portions of the dataset and transfer the data files into Amazon S3 in the AWS region. Have the EMR cluster read the dataset using EMRFS.

B. Establish a Direct Connect link between the on-premises system and the AWS region to reduce latency. Have the EMR cluster read the data directly from the on-premises storage system over Direct Connect.

C. Encrypt the data files according to encryption standards of Country X and store them on AWS region in Amazon S3. Have the EMR cluster read the dataset using EMRFS.

D. Use AWS Import/Export Snowball device to securely transfer the data to the AWS region and copy the files onto an EBS volume. Have the EMR cluster read the dataset using EMRFS.

Correct Answer: B

QUESTION 3

A Redshift data warehouse has different user teams that need to query the same table with very different query types. These user teams are experiencing poor performance.

Which action improves performance for the user teams in this situation?

A. Create custom table views.

B. Add interleaved sort keys per team.

C. Maintain team-specific copies of the table.

D. Add support for workload management queue hopping.

Correct Answer: D

Reference: https://docs.aws.amazon.com/redshift/latest/dg/cm-c-implementing-workloadmanagement.html

QUESTION 4

A medical record filing system for a government medical fund is using an Amazon S3 bucket to archive documents related to patients. Every patient visit to a physician creates a new file, which can add up millions of files each month. Collection of these files from each physician is handled via a batch process that runs ever night using AWS Data Pipeline. This is sensitive data, so the data and any associated metadata must be encrypted at rest.

Auditors review some files on a quarterly basis to see whether the records are maintained according to regulations. Auditors must be able to locate any physical file in the S3 bucket for a given date, patient, or physician. Auditors spend a significant amount of time location such files.

What is the most cost- and time-efficient collection methodology in this situation?

A. Use Amazon Kinesis to get the data feeds directly from physicians, batch them using a Spark application on Amazon Elastic MapReduce (EMR), and then store them in Amazon S3 with folders separated per physician.

B. Use Amazon API Gateway to get the data feeds directly from physicians, batch them using a Spark application on Amazon Elastic MapReduce (EMR), and then store them in Amazon S3 with folders separated per physician.

C. Use Amazon S3 event notification to populate an Amazon DynamoDB table with metadata about every file loaded to Amazon S3, and partition them based on the month and year of the file.

D. Use Amazon S3 event notification to populate an Amazon Redshift table with metadata about every file loaded to Amazon S3, and partition them based on the month and year of the file.

Correct Answer: A

QUESTION 5

An organization is soliciting public feedback through a web portal that has been deployed to track the number of requests and other important data. As part of reporting and visualization, AmazonQuickSight connects to an Amazon RDS database to virtualize data. Management wants to understand some important metrics about feedback and how the feedback has changed over the last four weeks in a visual representation.

What would be the MOST effective way to represent multiple iterations of an analysis in Amazon QuickSight that would show how the data has changed over the last four weeks?

A. Use the analysis option for data captured in each week and view the data by a date range.

B. Use a pivot table as a visual option to display measured values and weekly aggregate data as a row dimension.

C. Use a dashboard option to create an analysis of the data for each week and apply filters to visualize the data change.

D. Use a story option to preserve multiple iterations of an analysis and play the iterations sequentially.

Correct Answer: D

QUESTION 6

Which Amazon storage do you think is the best for my database-style applications that frequently encounter many random reads and writes across the dataset?

A. None of these.

B. Amazon Instance Storage

C. Any of these

D. Amazon EBS

Correct Answer: D

QUESTION 7

In the Amazon RDS Oracle DB engine, the Database Diagnostic Pack and the Database Tuning Pack are only available with _____

A. Oracle Standard Edition

B. Oracle Express Edition

C. Oracle Enterprise Edition

D. None of these

Correct Answer: C

QUESTION 8

An online gaming company uses DynamoDB to store user activity logs and is experiencing throttled writes on the company\\'s DynamoDB table. The company is NOT consuming close to the provisioned capacity. The table contains a large number of items and is partitioned on user and sorted by date. The table is 200GB and is currently provisioned at 10K WCU and 20K RCU.

Which two additional pieces of information are required to determine the cause of the throttling? (Choose two.)

A. The structure of any GSIs that have been defined on the table

B. CloudWatch data showing consumed and provisioned write capacity when writes are being throttled

C. Application-level metrics showing the average item size and peak update rates for each attribute

D. The structure of any LSIs that have been defined on the table

E. The maximum historical WCU and RCU for the table

Correct Answer: AD

QUESTION 9

When will you incur costs with an Elastic IP address (EIP)?

A. When an EIP is allocated.

B. When it is allocated and associated with a running instance.

C. When it is allocated and associated with a stopped instance.

D. Costs are incurred regardless of whether the EIP is associated with a running instance.

Correct Answer: C

QUESTION 10

Is there a limit to the number of groups you can have?

A. Yes for all users except root

B. No

C. Yes unless special permission granted

D. Yes for all users

Correct Answer: D

QUESTION 11

You are currently hosting multiple applications in a VPC and have logged numerous port scans coming in from a specific IP address block. Your security team has requested that all access from the offending IP address block be denied for the next 24 hours.

Which of the following is the best method to quickly and temporarily deny access from the specified IP address block?

A. Create an AD policy to modify Windows Firewall settings on all hosts in the VPC to deny access from the IP address block

B. Modify the Network ACLs associated with all public subnets in the VPC to deny access from the IP address block

C. Add a rule to all of the VPC 5 Security Groups to deny access from the IP address block

D. Modify the Windows Firewall settings on all Amazon Machine Images (AMIs) that your organization uses in that VPC to deny access from the IP address block

Correct Answer: B

QUESTION 12

A data engineer is running a DWH on a 25-node Redshift cluster of a SaaS service. The data engineer needs to build a dashboard that will be used by customers. Five big customers represent 80% of usage, and there is a long tail of dozens of smaller customers. The data engineer has selected the dashboarding tool.

How should the data engineer make sure that the larger customer workloads do NOT interfere with the smaller customer workloads?

A. Apply query filters based on customer-id that can NOT be changed by the user and apply distribution keys on customer-id.

B. Place the largest customers into a single user group with a dedicated query queue and place the rest of the customers into a different query queue.

C. Push aggregations into an RDS for Aurora instance. Connect the dashboard application to Aurora rather than Redshift for faster queries.

D. Route the largest customers to a dedicated Redshift cluster. Raise the concurrency of the multi-tenant Redshift cluster to accommodate the remaining customers.

Correct Answer: D

QUESTION 13

A user plans to use RDS as a managed DB platform. Which of the below mentioned features is not supported by RDS?

A. Automated backup

B. Automated scaling to manage a higher load

C. Automated failure detection and recovery

D. Automated software patching

Correct Answer: B

QUESTION 14

An administrator needs to design a distribution strategy for a star schema in a Redshift cluster. The administrator needs to determine the optimal distribution style for the tables in the Redshift schema.

In which three circumstances would choosing Key-based distribution be most appropriate? (Select three.)

A. When the administrator needs to optimize a large, slowly changing dimension table.

B. When the administrator needs to reduce cross-node traffic.

C. When the administrator needs to optimize the fact table for parity with the number of slices.

D. When the administrator needs to balance data distribution and collocation data.

E. When the administrator needs to take advantage of data locality on a local node for joins and aggregates.

Correct Answer: ACD

QUESTION 15

When attached to an Amazon VPC which two components provide connectivity with external networks? Choose 2 answers

A. Elastic IPS (EIP)

B. NAT Gateway (NAT)

C. Internet Gateway {IGW) D. Virtual Private Gateway (VGW)

Correct Answer: CD

Latest BDS-C00 Dumps   BDS-C00 Study Guide   BDS-C00 Exam Questions