

Amazon-Web-Services

Exam Questions BDS-C00

AWS Certified Big Data -Speciality



NEW QUESTION 1

You are working with customer who has 10 TB of archival data that they want to migrate to Amazon Glacier. The customer has a 1Mbps connection to the Internet. Which service or feature provide the fastest method of getting the data into Amazon Glacier?

- A. Amazon Glacier multipart upload
- B. AWS Storage Gateway
- C. VM Import/Export
- D. AWS Import/Export

Answer: D

NEW QUESTION 2

A us-based company is expanding their web presence into Europe. The company wants to extend their AWS infrastructure from Northern Virginia (us-east-1) into the Dublin (eu-west-1) region. Which of the following options would enable an equivalent experience for users on both continents?

- A. Use a public-facing load balancer per region to load-balancer web traffic, and enable HTTP health checks
- B. Use a public-facing load balancer per region to load balancer web traffic, and enable sticky sessions
- C. Use Amazon Route S3, and apply a geolocation routing policy to distribution traffic across both regions
- D. Use Amazon Route S3, and apply a weighted routing policy to distribute traffic across both regions

Answer: C

NEW QUESTION 3

You have started a new job and are reviewing your company's infrastructure on AWS You notice one web application where they have an Elastic Load Balancer (&B) in front of web instances in an Auto Scaling Group When you check the metrics for the ELB in CloudWatch you see four healthy instances in Availability Zone (AZ) A and zero in AZ B There are zero unhealthy instances. What do you need to fix to balance the instances across AZs?

- A. Set the ELB to only be attached to another AZ
- B. Make sure Auto Scaling is configured to launch in both AZs
- C. Make sure your AMI is available in both AZs
- D. Make sure the maximum size of the Auto Scaling Group is greater than 4

Answer: B

NEW QUESTION 4

You have launched an Amazon Elastic Compute Cloud (EC2) instance into a public subnet with a primary private IP address assigned, an internet gateway is attached to the VPC, and the public route table is configured to send all internet-based internet. Why is the internet unreachable from this instance?

- A. The Internet gateway security group must allow all outbound traffic
- B. The instance does not have a public IP address
- C. The instance "Source/Destination check" property must be enabled
- D. The instance security group must allow all inbound traffic

Answer: B

NEW QUESTION 5

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)

Answer: CD

NEW QUESTION 6

A customers needs to capture all client connection information from their load balancer every five minutes. The company wants to use data for analyzing traffic patterns and troubleshooting their applications. Which of the following options meets the customer requirements?

- A. Enable access logs on the load balancer
- B. Enable AWS CloudTrail for the load balancer
- C. Enable Amazon CloudWatch metrics on the load balancer
- D. Install the Amazon CloudWatch Logs agent on the load balancer

Answer: B

NEW QUESTION 7

You have been tasked with implementing an automated data backup solution for your application servers that run on Amazon EC2 with Amazon EBS volumes. You want to use a distributed data store for your backups to avoid single points of failure and to increase the durability of the data. Daily backups should be retained for 30 days so that you can restore data within an hour. How can you implement this through a script that a scheduling daemon runs daily on the application servers?

- A. Write the script to call the ec2-create-volume API, tag the Amazon EBS volume with the current data time group, and copy backup data to a second Amazon EBS volum
- B. Use the ec2-describe- volumes API to enumerate existing backup volume
- C. Call the ec2-delete-volume API to prune backup volumes that are tagged with a date-time group older than 30 days
- D. Write the script to call the Amazon Glacier upload archive API, and tag the backup archive with the current date-time grou
- E. Use the list vaults API to enumerate existing backup archive
- F. Call the delete vault API to prune backup archives that are tagged with a date-time group older than30 days
- G. Write the script to call the ec2-create-snapshot API, and tag the Amazon EBS snapshot with the current date-time group
- H. Use the ec2-describe-snapshot API to enumerate existing Amazon EBS snapshot
- I. Call the ec2-delete-snapshot API to prune Amazon EBs snapshots that are tagged with a date-time group older than 30 days
- J. Write the script to call the ec2-create-volume API, tag the Amazon EBS volume with the current date-time group, and use the ec2-copy-snapshot API to backup data to the new Amazon EBS volum
- K. Use the ec2-describe-snapshot API to enumerate existing backup volume
- L. Call the ec2- delete- snapshot API to prune backup Amazon EBS volumes that are tagged with a date-time group older than 30 days

Answer: C

NEW QUESTION 8

An existing application stores sensitive information on a non-boot Amazon EBS data volume attached to an Amazon Elastic Compute Cloud instance. Which of the following approaches would protect the sensitive data on an Amazon EBS volume?

- A. Snapshot the current Amazon EBS volum
- B. Restore the snapshot to a new, encrypted Amazon EBS volume Mount the Amazon EBS volume
- C. Create and mount a new, encrypted Amazon EBS volum
- D. Move the data to the new volum
- E. Delete the old Amazon EBS volume
- F. Unmount the EBS volum
- G. Toggle the encryption attribute to Tru
- H. Re-mount the Amazon EBs volume
- I. Upload your customer keys to AWS CloudHS
- J. Associate the Amazon EBS volume with AWS CloudHS
- K. Re-mount the Amazon EBS volume

Answer: A

NEW QUESTION 9

When using the following AWS services, which should be implemented in multiple Availability Zones for high availability solutions? Choose 2 answers

- A. Amazon Simple Storage Service
- B. Amazon Elastic Load Balancing
- C. Amazon Elastic Compute Cloud
- D. Amazon Simple Notification Service
- E. Amazon DynamoDB

Answer: BC

NEW QUESTION 10

An organization has configured a VPC with an Internet Gateway (IGW). Pairs of public and private subnets (each with one subnet per Availability Zone), and an Elastic Load Balancer (ELB) configured to use the public subnets. The application's web tier leverages the ELB. Auto Scaling and a multi-AZ RDS database instance the organization would like to eliminate any potential single points of failure in this design. What step should you take to achieve this organization's objective?

- A. Nothing, there are no single points of failure in this architecture.
- B. Create and attach a second IGW to provide redundant internet connectivity.
- C. Create and configure a second Elastic Load Balancer to provide a redundant load balancer.
- D. Create a second multi-AZ RDS instance in another Availability Zone and configure replication to provide a redundant database.

Answer: A

NEW QUESTION 10

A company has reproducible data that they want to store on Amazon Web Services. The company may want to retrieve the data on a frequent basis. Which Amazon web services storage option allows the customer to optimize storage costs and still achieve high availability for their data?

- A. Amazon S3 Reduced Redundancy Storage
- B. Amazon EBS Magnetic Volume
- C. Amazon Glacier
- D. Amazon S3 Standard Storage

Answer: A

NEW QUESTION 15

You currently run your infrastructure on Amazon EC2 instances behind on Auto Scaling group. All logs for your application are currently written to ephemeral storage. Recently your company experienced a major bug in code that made it through testing and was ultimately deployed to your fleet. This bug triggered your Auto Scaling group to scale up and back down before you could successfully retrieve the logs off your server to better assist you in troubleshooting the bug. Which technique should you use to make sure you are able to review your logs after your instances have shut down?

- A. Configure the ephemeral policies on your Auto Scaling group to back up on terminate
- B. Configure your Auto Scaling policies to create a snapshot of all ephemeral storage on terminate
- C. Install the CloudWatch logs Agent on your AMI, and configure CloudWatch Logs Agent to stream your logs
- D. Install the CloudWatch monitoring agent on your AMI, and set up a new SNS alert for CloudWatch metrics that triggers the CloudWatch monitoring agent to backup all logs on the ephemeral drive
- E. Install the CloudWatch Logs Agent on your AM
- F. Update your Scaling policy to enable automated CloudWatch Log copy

Answer: C

NEW QUESTION 20

Your company wants to start working with AWS, but has not yet opened an account. With which of the following services should you begin local development?

- A. Amazon DynamoDB
- B. Amazon Simple Queue Service
- C. Amazon Simple Email Service
- D. Amazon CloudSearch

Answer: A

NEW QUESTION 23

Your social media marketing application has a component written in Ruby running on AWS Elastic BeanStalk. This application component posts messages to social media sites in support of various marketing campaigns. Your management now requires you to record replies to these social media messages to analyze the effectiveness of the marketing campaign in comparison to past and future efforts. You have already developed a new application component to interface with the social media site APIs in order to read the replies. Which process should you use to record the social media replies in a durable data store that can be accessed at any time for analysis of historical data?

- A. Deploy the new application component in an Auto Scaling group of Amazon Elastic Compute Cloud (EC2) instances, read the data from the social media sites, store it with Amazon Elastic Block Store, and use AWS Data Pipeline to publish it to Amazon Kinesis for analytics
- B. Deploy the new application component as a Elastic BeanStalk application, read the data from the social media sites, store it in Amazon DynamoDB, and use Apache Hive with Amazon Elastic MapReduce for analytic
- C. Deploy the new application component in an Auto Scaling group of Amazon EC2 instances, read the data from the social media sites, store it in Amazon Glacier, and use AWS Data Pipeline to publish it to Amazon Redshift for analytics
- D. Deploy the new application component as an Amazon Elastic Beanstalk application, read the data from the social media site, store it with Amazon Elastic Block Store, and use Amazon Kinesis to stream the data to Amazon CloudWatch for analytics

Answer: B

NEW QUESTION 27

A customer wants to track access to their Amazon Simple Storage Service (S3) buckets and also use this information for their internal security and access audits. Which of the following will meet the Customer requirement?

- A. Enable AWS CloudTrail to audit all Amazon S3 bucket access.
- B. Enable server access logging for all required Amazon S3 bucket
- C. Enable the Requester Pays option to track access via AWS Billing
- D. Enable Amazon S3 event notifications for Put and Post.

Answer: B

NEW QUESTION 29

A photo sharing service stores pictures in Amazon Simple Storage Service (S3) and allows application signin using an Open ID Connect compatible identity provider. Which AWS Security Token approach to temporary access should you use for the Amazon S3 operations?

- A. SAML-based identity Federation
- B. Cross-Account Access
- C. AWS identity and Access Management roles
- D. Web identity Federation

Answer: A

NEW QUESTION 30

You are deploying an application to collect votes for a very popular television show. Millions of users will submit votes using mobile devices. The votes must be collected into a durable, scalable, and highly available data store for real-time public tabulation. Which service should you use?

- A. Amazon DynamoDB
- B. Amazon Redshift
- C. Amazon Kinesis
- D. Amazon Simple Queue Service

Answer: C

NEW QUESTION 34

A user has setup an RDS DB with Oracle. The user wants to get notifications when someone modifies the security group of that DB. How can the user configure that?

- A. It is not possible to get the notifications on a change in the security group
- B. Configure SNS to monitor security group changes
- C. Configure event notification on the DB security group
- D. Configure the CloudWatch alarm on the DB for a change in the security group

Answer: C

NEW QUESTION 36

You have a web application that is currently running on a collection of micro instance types in a single AZ behind a single load balancer. You have an Auto Scaling group configured to scale from 2 to 64 instances. When reviewing your CloudWatch metrics, you see that sometimes your Scaling group is running 64 micro instances. The web application is reading and writing to a DynamoDB-configured backend and configured with 800 Write Capacity units and 800 Read Capacity units. Your customers are complaining that they are experiencing load times when viewing your website. You have investigated the DynamoDB CloudWatch metrics; you are under the provisioned read and Write Capacity units and there is no throttling. How do you scale your service to improve the load times and ensure the principles of high availability?

- A. Change your Auto Scaling group configuration to include multiple AZs
- B. Change your Auto Scaling group configuration to include multiple AZs, and increase the number of Read Capacity units in your DynamoDB table by a factor of three, because you will need to be calling DynamoDB from three AZs
- C. Add a second load balancer to your Auto Scaling group so that you can support more inbound connections per second
- D. Change your Auto Scaling group configuration to use larger instances and include multiple AZs instead of one

Answer: D

NEW QUESTION 37

A user is trying to setup a recurring Auto Scaling process. The user has setup one process to scale up every day at 8 am and scale down at 7 PM. The user is trying to setup another recurring process which scales up on the 1st of every month at 8 AM and scales down the same day at 7 PM. What will Auto Scaling do in this scenario?

- A. Auto Scaling will execute both processes but will add just one instance on the 1st
- B. Auto Scaling will add two instances on the 1st of the month
- C. Auto Scaling will schedule both the processes but execute only one process randomly
- D. Auto Scaling will throw an error since there is a conflict in the schedule of two separate Auto Scaling Processes

Answer: D

NEW QUESTION 40

You run a small online consignment marketplace. Interested sellers complete an online application in order to allow them to sell their products on your website. Once approved, they can their product using a custom interface. From that point, you manage the shopping cart process so that when a buyer decides to buy a product, you handle the billing and coordination the shipping. Part of this process requires sending emails to the buyer and the seller at different stages. Your system has been running on AWS for a few months. Occasionally, products are shipped before payment has cleared and emails are sent out of order. Furthermore, sometimes credit cards are being charged twice. How can you resolve these problems?

- A. Use the Amazon Simple Queue Service (SQS), and use a different set of workers for each task
- B. Use the Amazon Simple Workflow Service (SWF), and use a different set of workers for each task.
- C. Use the Simple Email Service (SES) to control the correct order of email delivery
- D. Use the AWS Data Pipeline service to control the process flow of the various tasks
- E. Use the Amazon Simple Queue Service (SQS), and use a single set of workers for each task

Answer: E

NEW QUESTION 41

You are designing a web application that stores static assets in an Amazon Simple Storage Service (S3) bucket. You expect this bucket to immediately receive over 150 PUT requests per second. What should you do to ensure optimal performance?

- A. Use multi-part upload.
- B. Add a random prefix to the key names.
- C. Amazon S3 will automatically manage performance at this scale.
- D. Use a predictable naming scheme, such as sequential numbers or date time sequences, in the key names

Answer: B

NEW QUESTION 45

A user is planning to setup infrastructure on AWS for the Christmas sales. The user is planning to use Auto Scaling based on the schedule for proactive scaling. What advise would you give to the user?

- A. It is good to schedule now because if the user forgets later on it will not scale up
- B. The scaling should be setup only one week before Christmas
- C. Wait till end of November before scheduling the activity
- D. It is not advisable to use scheduled based scaling

Answer: C

NEW QUESTION 50

You have an ASP.NET web application running in Amazon Elastic BeanStalk. Your next version of the application requires a third-party Windows installer package to be installed on the instance on first boot and before the application launches. Which options are possible? Choose 2 answer

- A. In the application's Global.asax file, run msixexec.exe to install the package using Process.Start() in the Application_Start event handler
- B. In the source bundle's .ebextensions folder, create a file with a .config extension
- C. In the file, under the "packages" section and "msi" package manager, include the package's URL
- D. Launch a new Amazon EC2 instance from the AMI used by the environment
- E. Log into the instance, install the package and run sysprep
- F. Create a new AMI
- G. Configure the environment to use the new AMI
- H. In the environment's configuration, edit the instances configuration and add the package's URL to the "Packages" section
- I. In the source bundle's .ebextensions folder, create a "Packages" folder
- J. Place the package in the folder

Answer: BC

NEW QUESTION 52

Which of the following are true regarding AWS Cloud Trail? Choose 3 answers

- A. Cloudtrail is enabled globally
- B. Cloudtrail is enabled by default
- C. Cloudtrail is enabled on a per-region basis
- D. Cloudtrail is enabled on a per-service basis
- E. Logs can be delivered to a single Amazon S3 bucket for aggregation
- F. Logs can only be processed and delivered to the region in which they are generated

Answer: ACE

NEW QUESTION 53

A user has created an ELB with Auto Scaling. Which of the below mentioned offerings from ELB helps the user to stop sending new requests traffic from the load balancer to the EC2 instance when the instance is being deregistered while continuing in-flight requests?

- A. ELB sticky session
- B. ELB deregistration check
- C. ELB connection draining
- D. ELB auto registration Off

Answer: C

NEW QUESTION 54

Customers have recently been complaining that your web application has randomly stopped responding. During a deep dive of your logs, the team has discovered a major bug in your Java web application. This bug is causing a memory leak that eventually causes the application to crash. Your web application runs on Amazon EC2 and was built with AWS CloudFormation.

Which techniques should you see to help detect these problems faster, as well as help eliminate the server's unresponsiveness? Choose 2 answers

- A. Update your AWS CloudFormation configuration and enable a CustomResource that uses cfn-signal to detect memory leaks
- B. Update your CloudWatch metric granularity config for all Amazon EC2 memory metrics to support five-second granularity
- C. Create a CloudWatch alarm that triggers an Amazon SNS notification to page your team when the application memory becomes too large
- D. Update your AWS CloudFormation configuration to take advantage of Auto Scaling group
- E. Configure an Auto Scaling group policy to trigger off your custom CloudWatch metrics
- F. Create a custom CloudWatch metric that you push your JVM memory usage to create a CloudWatch alarm that triggers an Amazon SNS notification to page your team when the application memory usage becomes too large
- G. Update your AWS CloudFormation configuration to take advantage of CloudWatch metrics Agent
- H. Configure the CloudWatch Metrics Agent to monitor memory usage and trigger an Amazon SNS alarm

Answer: CD

NEW QUESTION 58

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.

Answer: C

NEW QUESTION 63

A company needs to deploy services to an AWS region which they not previously used. The company currently has an AWS identity and Access Management (IAM) role for their Amazon EC2 instances, which permits the instance to have access to Amazon DynamoDB. The company wants their EC2 instances in the new region to have the same privileges. How should the company achieve this?

- A. Create a new IAM role and associated policies within the new region
- B. Assign the existing IAM role to the Amazon EC2 instances in the new region
- C. Copy the IAM role and associated policies to the new region and attach it to the instances
- D. Create the Amazon Machine Image of the instance and copy it to the desired region using the AMI Copy feature

Answer: B

NEW QUESTION 65

The Trusted Advisor service provides insight regarding which four categories of an AWS account?

- A. Security, fault tolerance, high availability, and connectivity
- B. Security, access control, high availability, and performance
- C. Performance, cost optimization, security, and fault tolerance
- D. Performance, cost optimization, access control, and connectivity

Answer: C

NEW QUESTION 68

A user has created a launch configuration for Auto Scaling where CloudWatch detailed monitoring is disabled. The user wants to now enable detailed monitoring. How can the user achieve this?

- A. Update the Launch config with CLI to set InstanceMonitoringDisabled = false
- B. The user should change the Auto Scaling group from the AWS console to enable detailed monitoring
- C. Update the Launch config with CLI to set InstanceMonitoring.Enabled = true
- D. Create a new Launch Config with detail monitoring enabled and update the Auto Scaling group

Answer: D

NEW QUESTION 69

You have a video Trans coding application running on Amazon EC2. Each instance pools a queue to find out which video should be Trans coded, and then runs a Trans coding process.

If this process is interrupted, the video will be Trans coded by another instance based on the queuing system. You have a large backlog of videos which need to be Trans coded and would like to reduce this backlog by adding more instances. You will need these instances only until the backlog is reduced. Which type of Amazon EC2 instance should you use to reduce the backlog in the most cost- effective way?

- A. Dedicated instances
- B. Spot instances
- C. On-demand instances
- D. Reserved instances

Answer: B

NEW QUESTION 71

A systems engineer for a company proposes digitalization and backup of large archives for customers.

The systems engineer needs to provide users with a secure storage that makes sure that data will never be tampered with once it has been uploaded. How should this be accomplished?

- A. Create an Amazon Glacier Vault
- B. Specify a “Deny” Vault lock policy on this vault to block “glacier:DeleteArchive”.
- C. Create an Amazon S3 bucket
- D. Specify a “Deny” bucket policy on this bucket to block “s3:DeleteObject”.
- E. Create an Amazon Glacier Vault
- F. Specify a “Deny” vault access policy on this Vault to block “glacier:DeleteArchive”.
- G. Create a secondary AWS containing an Amazon S3 bucket
- H. Grant “s3:PutObject” to the primary account.

Answer: A

NEW QUESTION 76

A customer is collecting clickstream data using Amazon Kinesis and is grouping the events by IP address into 5-minute chunks stored in Amazon S3.

Many analysts in the company use Hive on Amazon EMR to analyze this data. Their queries always reference a single IP address. Data must be optimized for querying based on IP address using Hive running on Amazon EMR. What is the most efficient method to query the data with Hive?

- A. Store an index of the files by IP address in the Amazon DynamoDB metadata store for EMRFS
- B. Store the Amazon S3 objects with the following naming scheme: bucketname/source=ip_address/year=yy/month=mm/day=dd/hour=hh/filename
- C. Store the data in an HBase table with the IP address as the row key
- D. Store the events for an IP address as a single file in Amazon S3 and add metadata with key:Hive_Partitioned_IPAddress

Answer: B

NEW QUESTION 81

A solutions architect for a logistics organization ships packages from thousands of suppliers to end customers. The architect is building a platform where suppliers can view the status of one or more of their shipments. Each supplier can have multiple roles that will only allow access to specific fields in the resulting information.

Which strategy allows the appropriate level of access control and requires the LEAST amount of management work?

- A. Send the tracking data to Amazon Kinesis Stream
- B. Use AWS Lambda to store the data in an Amazon DynamoDB Table
- C. Generate temporary AWS credentials for the supplier's users with AWS STS, specifying fine-grained security policies to limit access only to their application data.
- D. Send the tracking data to Amazon Kinesis Firehose
- E. Use Amazon S3 notifications and AWS Lambda to prepare files in Amazon S3 with appropriate data for each supplier's role
- F. Generate temporary AWS credentials for the suppliers' users with AWS STS
- G. Limit access to the appropriate files through security policies.

- H. Send the tracking data to Amazon Kinesis Stream
- I. Use Amazon EMR with Spark Streaming to store the data in HBas
- J. Create one table per supplie
- K. Use HBase Kerberos integration with the suppliers' user
- L. Use HBase ACL-based security to limit access to the roles to their specific table and columns.
- M. Send the tracking data to Amazon Kinesis Firehos
- N. Store the data in an Amazon Redshift cluste
- O. Create views for the supplier's users and role
- P. Allow suppliers access to the Amazon Redshift cluster using a user limited to the application view.

Answer: B

NEW QUESTION 82

An administrator tries to use the Amazon Machine Learning service to classify social media posts that mention the administrator's company into posts that requires a response and posts that do not. The training dataset of 10,000 posts contains the details of each post including the timestamp, author, and full text of the post. The administrator is missing the target labels that are required for training. Which Amazon Machine Learning model is the most appropriate for the task?

- A. Unary classification model, where the target class is the require-response post
- B. Binary classification model, where the two classes are require-response and does-not-require- response
- C. Multi-class prediction model, with two classes require-response and does-not-require response
- D. Regression model where the predicted value is the probability that the post requires a response

Answer: B

NEW QUESTION 86

A company's social media manager requests more staff on the weekends to handle an increase in customer contacts from a particular region. The company needs a report to visualize the trends on weekends over the past 6 months using QuickSight. How should the data be represented?

- A. A line graph plotting customer contacts v
- B. time, with a line for each region
- C. A pie chart per region plotting customer contacts per day of week
- D. A map of the regions with a heatmap overlay to show the volume of customer contacts
- E. A bar graph plotting region vs volume of social media contacts

Answer: A

NEW QUESTION 88

A company that manufactures and sells smart air conditioning units also offers add-on services so that customers can see real-time dashboards in a mobile application or a web browser. Each unit sends its sensor information in JSON format every two seconds for processing and analysis. The company also needs to consume this data to predict possible equipment problems before they occur. A few thousand pre-purchased units will be delivered in the next couple of months. The company expects high market growth in the next year and needs to handle a massive amount of data and scale interruption.

Which ingestion solution should the company use?

- A. Write sensor data records to Amazon Kinesis Stream
- B. Process the data using KCL applications for the end-consumer dashboard and anomaly detection workflows.
- C. Batch sensor data Amazon Simple Storage Service (S3) every 15 minute
- D. Flow the data downstream to the end-consumer dashboard and to the anomaly detection application.
- E. Write sensor data records to Amazon Kinesis Firehose with Amazon Simply Storage Service (S3) as the destinatio
- F. Consume the data with a KCL application for the end-consumer dashboard and anomaly detection.
- G. Write sensor data records to Amazon Relational Database Service (RDS). Build both the end- consumer dashboard application on top of Amazon RDS.

Answer: A

NEW QUESTION 91

An administrator receives about 100 files per hour into Amazon S3 and will be loading the files into Amazon Redshift. Customers who analyze the data within Redshift gain significant value when they receive data as quickly as possible. The customers have agreed to a maximum loading interval of 5 minutes. Which loading approach should the administrator use to meet this objective?

- A. Load each file as it arrives because getting data into the cluster as quickly as possible is the priority.
- B. Load the cluster as soon as the administrator has the same number of files as nodes in the cluster.
- C. Load the cluster when the administrator has an even multiple of files relative to Cluster Slice Count, or 5 minutes whichever comes first.
- D. Load the cluster when the number files is less than the Cluster Slice Count.

Answer: C

NEW QUESTION 92

A game company needs to properly scale its game application, which is backed by DynamoDB.

Amazon Redshift has the past two years of historical data. Game traffic varies throughout the year based on various factors such as season, movie release, and holiday season. An administrator needs to calculate how much read and write throughput should be provisioned for DynamoDB table for each week in advance. How should the administrator accomplish this task?

- A. Feed the data into Amazon Machine Learning and build a regression model
- B. Feed the data into Spark Mlib and build a random forest model
- C. Feed the data into Apache Mahout and build a multi-classification model
- D. Feed the data into Amazon Machine Learning and build a binary classification model

Answer: B

NEW QUESTION 95

A data engineer chooses Amazon DynamoDB as a data store for a regulated application. This application must be submitted to regulators for review. The data engineer needs to provide a control framework that lists the security controls from the process to follow to add new users down to the physical controls of the data center, including items like security guards and cameras.

How should this control mapping be achieved using AWS?

- A. Request AWS third-party audit reports and/or the AWS quality addendum and map the AWS responsibilities to the controls that must be provided
- B. Request data center Temporary Auditor access to an AWS data center to verify the control mapping
- C. Request relevant SLAs and security guidelines for Amazon DynamoDB and define these guidelines within the application's architecture to map to the control framework
- D. Request Amazon DynamoDB system architecture designs to determine how to map the AWS responsibilities to the controls that must be provided

Answer: A

NEW QUESTION 98

A customer has an Amazon S3 bucket. Objects are uploaded simultaneously by a cluster of servers from multiple streams of data. The customer maintains a catalog of objects uploaded in Amazon S3 using an Amazon DynamoDB table. This catalog has the following fields StreamName, TimeStamp, and ServerName, TimeStamp, and ServerName, from which ObjectName can be obtained.

The customer needs to define the catalog to support querying for a given stream or server within a defined time range.

Which DynamoDB table scheme is most efficient to support these queries?

- A. Define a Primary Key with ServerName as Partition Key and TimeStamp as Sort Key
- B. Don NOT define a Secondary Index or Global Secondary Index.
- C. Define a Primary Key with StreamName as Partition Key and TimeStamp followed by ServerName as Sort Key
- D. Define a Global Secondary Index with ServerName as Partition Key and TimeStamp followed by StreamName.
- E. Define a Primary Key with ServerName as Partition Key
- F. Define a Local Secondary Index with StreamName as Partition Key
- G. Define a Global Secondary Index with TimeStamp as Partition Key.
- H. Define a Primary Key with ServerName as Partition Key
- I. Define a Local Secondary Index with TimeStamp as Partition Key
- J. Define a Global Secondary Index with StreamName as Partition key and TimeStamp as Sort Key.

Answer: A

NEW QUESTION 99

A company is using Amazon Machine Learning as part of a medical software application. The application will predict the most likely blood type for a patient based on a variety of other clinical tests that are available when blood type knowledge is unavailable.

What is the appropriate model choice and target attribute combination for the problem?

- A. Multi-class classification model with a categorical target attribute
- B. Regression model with a numeric target attribute
- C. Binary Classification with a categorical target attribute
- D. K-Nearest Neighbors model with a multi-class target attribute

Answer: C

NEW QUESTION 102

A data engineer in a manufacturing company is designing a data processing platform that receives a large volume of unstructured data. The data engineer must populate a well-structured star schema in Amazon Redshift.

What is the most efficient architecture strategy for this purpose?

- A. Transform the unstructured data using Amazon EMR and generate CSV data
- B. COPY data into the analysis schema within Redshift.
- C. Load the unstructured data into Redshift, and use string parsing functions to extract structured data for inserting into the analysis schema.
- D. When the data is saved to Amazon S3. Use S3 Event Notifications and AWS Lambda to transform the file content
- E. Insert the data into the analysis schema on Redshift.
- F. Normalize the data using an AWS Marketplace ETL tool persist the result to Amazon S3 and use AWS Lambda to INSERT the data into Redshift.

Answer: B

NEW QUESTION 103

A customer needs to determine the optimal distribution strategy for the ORDERS fact table in its Redshift schema. The ORDERS table has foreign key relationships with multiple dimension tables in this schema.

How should the company determine the most appropriate distribution key for the ORDERS table?

- A. Identify the largest and most frequently joined dimension table and ensure that it and the ORDERS table both have EVEN distribution
- B. Identify the target dimension table and designate the key of this dimension table as the distribution key of the ORDERS table
- C. Identify the smallest dimension table and designate the key of this dimension table as the distribution key of ORDERS table
- D. Identify the largest and most frequently joined dimension table and designate the key of this dimension table as the distribution key for the orders table

Answer: D

NEW QUESTION 104

The department of transportation for a major metropolitan area has placed sensors on roads at key locations around the city. The goal is to analyze the flow of traffic and notifications from emergency services to identify potential issues and to help planners correct trouble spots.

A data engineer needs a scalable and fault-tolerant solution that allows planners to respond to issues within 30 seconds of their occurrence. Which solution should the data engineer choose?

- A. Collect the sensor data with Amazon Kinesis Firehose and store it in Amazon Redshift for analysis
- B. Collect emergency services events with Amazon SQS and store in Amazon DynamoDB for analysis
- C. Collect the sensor data with Amazon SQS and store in Amazon DynamoDB for analysis. Collect emergency services events with Amazon Kinesis Firehose and store in Amazon Redshift for analysis
- D. Collect both sensor data and emergency services events with Amazon Kinesis Streams and use Amazon DynamoDB for analysis
- E. Collect both sensor data and emergency services events with Amazon Kinesis Firehose and use Amazon Redshift for Analysis

Answer: A

NEW QUESTION 107

An online photo album app has a key design feature to support multiple screens (e.g. desktop, mobile phone, and tablet) with high quality displays. Multiple versions of the image must be saved in different resolutions and layouts. The image processing Java program takes an average of five seconds per upload, depending on the image size and format. Each image upload captures the following image metadata: user, album, photo label, upload timestamp
The app should support the following requirements:

- Hundreds of user image uploads per second
- Maximum image metadata size of 10 MB
- Maximum image metadata size of 1 KB
- Image displayed in optimized resolution in all supported screens no later than one minute after image upload

Which strategy should be used to meet these requirements?

- A. Write images and metadata to Amazon Kinesis, Use a Kinesis Client Library (KCL) application to run the image processing and save the image output to Amazon S3 and metadata to the app repository DB
- B. Write image and metadata RDS with BLOB data type
- C. Use AWS Data Pipeline to run the image processing and save the image output to Amazon S3 and metadata to the app repository DB
- D. Upload image with metadata to Amazon S3 use Lambda function to run the image processing and save the image output to Amazon S3 and metadata to the app repository DB
- E. Write image and metadata to Amazon Kinesis
- F. Use Amazon Elastic MapReduce (EMR) with Spark Streaming to run image processing and save image output to Amazon

Answer: D

NEW QUESTION 112

A solutions architect works for a company that has a data lake based on a central Amazon S3 bucket. The data contains sensitive information. The architect must be able to specify exactly which files each user can access. Users access the platform through SAML federation Single Sign On platform. The architect needs to build a solution that allows fine grained access control, traceability of access to the objects, and usage of the standard tools (AWS Console, AWS CLI) to access the data. Which solution should the architect build?

- A. Use Amazon S3 Server-Side Encryption with AWS KMS-Managed Keys for strong data. Use AWS KMS to allow access to specific elements of the platform
- B. Use AWS CloudTrail for auditing
- C. Use Amazon S3 Server-Side Encryption with Amazon S3 Managed Key
- D. Set Amazon S3 ACL to allow access to specific elements of the platform
- E. Use Amazon S3 to access logs for auditing
- F. Use Amazon S3 Client-Side Encryption with Client-Side Master Key
- G. Set Amazon S3 ACL to allow access to specific elements of the platform
- H. Use Amazon S3 access logs for auditing
- I. Use Amazon S3 Client-Side Encryption with AWS KMS-Managed keys for storing data. Use AWS KMS to allow access to specific elements of the platform
- J. Use AWS CloudTrail for auditing

Answer: B

NEW QUESTION 113

A media advertising company handles a large number of real-time messages sourced from over 200 websites. The company's data engineer needs to collect and process records in real time for analysis using Spark Streaming on Amazon Elastic MapReduce (EMR). The data engineer needs to fulfill a corporate mandate to keep ALL raw messages as they are received as a top priority. Which Amazon Kinesis configuration meets these requirements?

- A. Publish messages to Amazon Kinesis Firehose backed by Amazon Simple Storage Service (S3). Pull messages off Firehose with Spark Streaming in parallel to persistence to Amazon S3
- B. Publish messages to Amazon Kinesis Stream
- C. Pull messages off Stream with Spark Streaming in parallel to AWS messages from Streams to Firehose backed by Amazon Simple Storage Service (S3)
- D. Publish messages to Amazon Kinesis Firehose backed by Amazon Simple Storage (S3). Use AWS Lambda messages from Firehose to Streams for processing with Spark Streaming
- E. Publish messages to Amazon Kinesis Streams, pull messages off with Spark Streaming and write data new data to Amazon Simple Storage Service (S3) before and after processing

Answer: D

NEW QUESTION 115

A clinical trial will rely on medical sensors to remotely assess patient health. Each physician who participates in the trial requires visual reports each morning. The reports are built from aggregations of all the sensor data taken each minute. What is the most cost-effective solution for creating this visualization each day?

- A. Use Kinesis Aggregators Library to generate reports for reviewing the patient sensor data and generate a QuickSight visualization on the new data each

morning for the physician to review

- B. Use a Transient EMR cluster that shuts down after use to aggregate the patient sensor data each night and generate a QuickSight visualization on the new data each morning for the physician to review
- C. Use Spark streaming on EMR to aggregate the sensor data coming in every 15 minutes and generate a QuickSight visualization on the new data each morning for the physician to review
- D. Use an EMR cluster to aggregate the patient sensor data each right and provide Zeppelin notebooks that look at the new data residing on the cluster each morning

Answer: AD

NEW QUESTION 116

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

Answer: ADE

NEW QUESTION 118

A company is building a new application is AWS. The architect needs to design a system to collect application log events. The design should be a repeatable pattern that minimizes data loss if an application instance fails, and keeps a durable copy of all log data for at least 30 days. What is the simplest architecture that will allow the architect to analyze the logs?

- A. Write them directly to a Kinesis Firehos
- B. Configure Kinesis Firehose to load the events into an Amazon Redshift cluster for analysis.
- C. Write them to a file on Amazon Simple Storage Service (S3). Write an AWS lambda function that runs in response to the S3 events to load the events into Amazon Elasticsearch service for analysis.
- D. Write them to the local disk and configure the Amazon cloud watch Logs agent to lead the data into CloudWatch Logs and subsequently into Amazon Elasticsearch Service.
- E. Write them to CloudWatch Logs and use an AWS Lambda function to load them into HDFS on an Amazon Elastic MapReduce (EMR) cluster for analysis.

Answer: A

NEW QUESTION 123

An administrator is processing events in near real-time using Kinesis streams and Lambda. Lambda intermittently fails to process batches from one of the shards due to a 5 –minute time limit. What is a possible solution for this problem?

- A. Add more Lambda functions to improve concurrent batch processing
- B. Reduce the batch size that lambda is reading from the stream
- C. Ignore and skip events that are older than 5 minutes and put them to Dead Letter Queue (DLQ)
- D. Configure Lambda to read from fewer shards in parallel

Answer: BD

NEW QUESTION 124

Which data store should the organization choose?

- A. Amazon Relational Database Service (RDS)
- B. Amazon Redshift
- C. Amazon DynamoDB
- D. Amazon Elasticsearch

Answer: C

NEW QUESTION 125

An organization needs to design and deploy a large-scale data storage solution that will be highly durable and highly flexible with respect to the type and structure of data being stored. The data to be stored will be sent or generated from a variety of sources and must be persistently available for access and processing by multiple applications.

What is the most cost-effective technique to meet these requirements?

- A. Use Amazon Simple Storage Service (S3) as the actual data storage system, coupled with appropriate tools for ingestion/acquisition of data and for subsequent processing and querying.
- B. Deploy a long-running Amazon Elastic MapReduce (EMR) cluster with Amazon Elastic Block Store (EBS) volumes for persistent HDFS storage and appropriate Hadoop ecosystem tools for processing and querying.
- C. Use Amazon Redshift with data replication to Amazon Simple Storage Service (S3) for comprehensive durable data storage, processing and querying.
- D. Launch an Amazon Relational Database Service (RDS), and use the enterprise grade and capacity of the Amazon Aurora Engine for storage processing and querying.

Answer: A

NEW QUESTION 128

.....

Thank You for Trying Our Product

We offer two products:

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

2nd - Questions and Answers in PDF Format

BDS-C00 Practice Exam Features:

- * BDS-C00 Questions and Answers Updated Frequently
- * BDS-C00 Practice Questions Verified by Expert Senior Certified Staff
- * BDS-C00 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * BDS-C00 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

100% Actual & Verified — Instant Download, Please Click
[Order The BDS-C00 Practice Test Here](#)