



DAS-C01^{Q&As}

AWS Certified Data Analytics - Specialty (DAS-C01)

Pass Amazon DAS-C01 Exam with 100% Guarantee

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

<https://www.lead4pass.com/das-c01.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

An airline has been collecting metrics on flight activities for analytics. A recently completed proof of concept demonstrates how the company provides insights to data analysts to improve on-time departures. The proof of concept used objects in Amazon S3, which contained the metrics in .csv format, and used Amazon Athena for querying the data. As the amount of data increases, the data analyst wants to optimize the storage solution to improve query performance.

Which options should the data analyst use to improve performance as the data lake grows? (Choose three.)

- A. Add a randomized string to the beginning of the keys in S3 to get more throughput across partitions.
- B. Use an S3 bucket in the same account as Athena.
- C. Compress the objects to reduce the data transfer I/O.
- D. Use an S3 bucket in the same Region as Athena.
- E. Preprocess the .csv data to JSON to reduce I/O by fetching only the document keys needed by the query.
- F. Preprocess the .csv data to Apache Parquet to reduce I/O by fetching only the data blocks needed for predicates.

Correct Answer: ACE

QUESTION 2

A retail company leverages Amazon Athena for ad-hoc queries against an AWS Glue Data Catalog. The data analytics team manages the data catalog and data access for the company. The data analytics team wants to separate queries and manage the cost of running those queries by different workloads and teams. Ideally, the data analysts want to group the queries run by different users within a team, store the query results in individual Amazon S3 buckets specific to each team, and enforce cost constraints on the queries run against the Data Catalog.

Which solution meets these requirements?

- A. Create IAM groups and resource tags for each team within the company. Set up IAM policies that control user access and actions on the Data Catalog resources.
- B. Create Athena resource groups for each team within the company and assign users to these groups. Add S3 bucket names and other query configurations to the properties list for the resource groups.
- C. Create Athena workgroups for each team within the company. Set up IAM workgroup policies that control user access and actions on the workgroup resources.
- D. Create Athena query groups for each team within the company and assign users to the groups.

Correct Answer: A

QUESTION 3



A company developed a new elections reporting website that uses Amazon Kinesis Data Firehose to deliver full logs from AWS WAF to an Amazon S3 bucket. The company is now seeking a low-cost option to perform this infrequent data analysis with visualizations of logs in a way that requires minimal development effort.

Which solution meets these requirements?

A. Use an AWS Glue crawler to create and update a table in the Glue data catalog from the logs. Use Athena to perform ad-hoc analyses and use Amazon QuickSight to develop data visualizations.

B. Create a second Kinesis Data Firehose delivery stream to deliver the log files to Amazon Elasticsearch Service (Amazon ES). Use Amazon ES to perform text-based searches of the logs for ad-hoc analyses and use Kibana for data visualizations.

C. Create an AWS Lambda function to convert the logs into .csv format. Then add the function to the Kinesis Data Firehose transformation configuration. Use Amazon Redshift to perform ad-hoc analyses of the logs using SQL queries and use Amazon QuickSight to develop data visualizations.

D. Create an Amazon EMR cluster and use Amazon S3 as the data source. Create an Apache Spark job to perform ad-hoc analyses and use Amazon QuickSight to develop data visualizations.

Correct Answer: D

QUESTION 4

A company has collected more than 100 TB of log files in the last 24 months. The files are stored as raw text in a dedicated Amazon S3 bucket. Each object has a key of the form year-monthday_log_HHmmss.txt where HHmmss represents the time the log file was initially created. A table was created in Amazon Athena that points to the S3 bucket. One-time queries are run against a subset of columns in the table several times an hour.

A data analyst must make changes to reduce the cost of running these queries. Management wants a solution with minimal maintenance overhead.

Which combination of steps should the data analyst take to meet these requirements? (Choose three.)

A. Convert the log files to Apache Avro format.

B. Add a key prefix of the form date=year-month-day/ to the S3 objects to partition the data.

C. Convert the log files to Apache Parquet format.

D. Add a key prefix of the form year-month-day/ to the S3 objects to partition the data.

E. Drop and recreate the table with the PARTITIONED BY clause. Run the ALTER TABLE ADD PARTITION statement.

F. Drop and recreate the table with the PARTITIONED BY clause. Run the MSCK REPAIR TABLE statement.

Correct Answer: BCF

Reference: <https://docs.aws.amazon.com/athena/latest/ug/msck-repair-table.html>

QUESTION 5



A media company is using Amazon QuickSight dashboards to visualize its national sales data. The dashboard is using a dataset with these fields: ID, date, time_zone, city, state, country, longitude, latitude, sales_volume, and number_of_items.

To modify ongoing campaigns, the company wants an interactive and intuitive visualization of which states across the country recorded a significantly lower sales volume compared to the national average.

Which addition to the company's QuickSight dashboard will meet this requirement?

- A. A geospatial color-coded chart of sales volume data across the country.
- B. A pivot table of sales volume data summed up at the state level.
- C. A drill-down layer for state-level sales volume data.
- D. A drill through to other dashboards containing state-level sales volume data.

Correct Answer: B

QUESTION 6

A company uses the Amazon Kinesis SDK to write data to Kinesis Data Streams. Compliance requirements state that the data must be encrypted at rest using a key that can be rotated. The company wants to meet this encryption requirement with minimal coding effort.

How can these requirements be met?

- A. Create a customer master key (CMK) in AWS KMS. Assign the CMK an alias. Use the AWS Encryption SDK, providing it with the key alias to encrypt and decrypt the data.
- B. Create a customer master key (CMK) in AWS KMS. Assign the CMK an alias. Enable server-side encryption on the Kinesis data stream using the CMK alias as the KMS master key.
- C. Create a customer master key (CMK) in AWS KMS. Create an AWS Lambda function to encrypt and decrypt the data. Set the KMS key ID in the function's environment variables.
- D. Enable server-side encryption on the Kinesis data stream using the default KMS key for Kinesis Data Streams.

Correct Answer: B

Reference: <https://aws.amazon.com/kinesis/data-streams/faqs/>

QUESTION 7

A team of data scientists plans to analyze market trend data for their company's new investment strategy. The trend data comes from five different data sources in large volumes. The team wants to utilize Amazon Kinesis to support their use case. The team uses SQL-like queries to analyze trends and wants to send notifications based on certain significant patterns in the trends. Additionally, the data scientists want to save the data to Amazon S3 for archival and historical re-processing, and use AWS managed services wherever possible. The team wants to implement the lowest-cost solution.

Which solution meets these requirements?



- A. Publish data to one Kinesis data stream. Deploy a custom application using the Kinesis Client Library (KCL) for analyzing trends, and send notifications using Amazon SNS. Configure Kinesis Data Firehose on the Kinesis data stream to persist data to an S3 bucket.
- B. Publish data to one Kinesis data stream. Deploy Kinesis Data Analytic to the stream for analyzing trends, and configure an AWS Lambda function as an output to send notifications using Amazon SNS. Configure Kinesis Data Firehose on the Kinesis data stream to persist data to an S3 bucket.
- C. Publish data to two Kinesis data streams. Deploy Kinesis Data Analytics to the first stream for analyzing trends, and configure an AWS Lambda function as an output to send notifications using Amazon SNS. Configure Kinesis Data Firehose on the second Kinesis data stream to persist data to an S3 bucket.
- D. Publish data to two Kinesis data streams. Deploy a custom application using the Kinesis Client Library (KCL) to the first stream for analyzing trends, and send notifications using Amazon SNS. Configure Kinesis Data Firehose on the second Kinesis data stream to persist data to an S3 bucket.

Correct Answer: A

QUESTION 8

A financial company hosts a data lake in Amazon S3 and a data warehouse on an Amazon Redshift cluster. The company uses Amazon QuickSight to build dashboards and wants to secure access from its on-premises Active Directory to Amazon QuickSight.

How should the data be secured?

- A. Use an Active Directory connector and single sign-on (SSO) in a corporate network environment.
- B. Use a VPC endpoint to connect to Amazon S3 from Amazon QuickSight and an IAM role to authenticate Amazon Redshift.
- C. Establish a secure connection by creating an S3 endpoint to connect Amazon QuickSight and a VPC endpoint to connect to Amazon Redshift.
- D. Place Amazon QuickSight and Amazon Redshift in the security group and use an Amazon S3 endpoint to connect Amazon QuickSight to Amazon S3.

Correct Answer: B

QUESTION 9

A company has developed several AWS Glue jobs to validate and transform its data from Amazon S3 and load it into Amazon RDS for MySQL in batches once every day. The ETL jobs read the S3 data using a DynamicFrame. Currently, the ETL developers are experiencing challenges in processing only the incremental data on every run, as the AWS Glue job processes all the S3 input data on each run.

Which approach would allow the developers to solve the issue with minimal coding effort?

- A. Have the ETL jobs read the data from Amazon S3 using a DataFrame.



- B. Enable job bookmarks on the AWS Glue jobs.
- C. Create custom logic on the ETL jobs to track the processed S3 objects.
- D. Have the ETL jobs delete the processed objects or data from Amazon S3 after each run.

Correct Answer: D

QUESTION 10

A mobile gaming company wants to capture data from its gaming app and make the data available for analysis immediately. The data record size will be approximately 20 KB. The company is concerned about achieving optimal throughput from each device. Additionally, the company wants to develop a data stream processing application with dedicated throughput for each consumer.

Which solution would achieve this goal?

- A. Have the app call the PutRecords API to send data to Amazon Kinesis Data Streams. Use the enhanced fan-out feature while consuming the data.
- B. Have the app call the PutRecordBatch API to send data to Amazon Kinesis Data Firehose. Submit a support case to enable dedicated throughput on the account.
- C. Have the app use Amazon Kinesis Producer Library (KPL) to send data to Kinesis Data Firehose. Use the enhanced fan-out feature while consuming the data.
- D. Have the app call the PutRecords API to send data to Amazon Kinesis Data Streams. Host the stream-processing application on Amazon EC2 with Auto Scaling.

Correct Answer: D

QUESTION 11

A company that monitors weather conditions from remote construction sites is setting up a solution to collect temperature data from the following two weather stations.

1.

Station A, which has 10 sensors

2.

Station B, which has five sensors

These weather stations were placed by onsite subject-matter experts.

Each sensor has a unique ID. The data collected from each sensor will be collected using Amazon Kinesis Data Streams.



Based on the total incoming and outgoing data throughput, a single Amazon Kinesis data stream with two shards is created. Two partition keys are created based on the station names. During testing, there is a bottleneck on data coming from Station A, but not from Station B. Upon review, it is confirmed that the total stream throughput is still less than the allocated Kinesis Data Streams throughput.

How can this bottleneck be resolved without increasing the overall cost and complexity of the solution, while retaining the data collection quality requirements?

- A. Increase the number of shards in Kinesis Data Streams to increase the level of parallelism.
- B. Create a separate Kinesis data stream for Station A with two shards, and stream Station A sensor data to the new stream.
- C. Modify the partition key to use the sensor ID instead of the station name.
- D. Reduce the number of sensors in Station A from 10 to 5 sensors.

Correct Answer: A

QUESTION 12

A bank operates in a regulated environment. The compliance requirements for the country in which the bank operates say that customer data for each state should only be accessible by the bank's employees located in the same state. Bank employees in one state should NOT be able to access data for customers who have provided a home address in a different state.

The bank's marketing team has hired a data analyst to gather insights from customer data for a new campaign being launched in certain states. Currently, data linking each customer account to its home state is stored in a tabular .csv file within a single Amazon S3 folder in a private S3 bucket. The total size of the S3 folder is 2 GB uncompressed. Due to the country's compliance requirements, the marketing team is not able to access this folder.

The data analyst is responsible for ensuring that the marketing team gets one-time access to customer data for their campaign analytics project, while being subject to all the compliance requirements and controls.

Which solution should the data analyst implement to meet the desired requirements with the LEAST amount of setup effort?

- A. Re-arrange data in Amazon S3 to store customer data about each state in a different S3 folder within the same bucket. Set up S3 bucket policies to provide marketing employees with appropriate data access under compliance controls. Delete the bucket policies after the project.
- B. Load tabular data from Amazon S3 to an Amazon EMR cluster using s3DistCp. Implement a custom Hadoop-based row-level security solution on the Hadoop Distributed File System (HDFS) to provide marketing employees with appropriate data access under compliance controls. Terminate the EMR cluster after the project.
- C. Load tabular data from Amazon S3 to Amazon Redshift with the COPY command. Use the built-in row-level security feature in Amazon Redshift to provide marketing employees with appropriate data access under compliance controls. Delete the Amazon Redshift tables after the project.
- D. Load tabular data from Amazon S3 to Amazon QuickSight Enterprise edition by directly importing it as a data source. Use the built-in row-level security feature in Amazon QuickSight to provide marketing employees with appropriate data access under compliance controls. Delete Amazon QuickSight data sources after the project is complete.



Correct Answer: C

QUESTION 13

A company is migrating its existing on-premises ETL jobs to Amazon EMR. The code consists of a series of jobs written in Java. The company needs to reduce overhead for the system administrators without changing the underlying code. Due to the sensitivity of the data, compliance requires that the company use root device volume encryption on all nodes in the cluster. Corporate standards require that environments be provisioned through AWS CloudFormation when possible.

Which solution satisfies these requirements?

- A. Install open-source Hadoop on Amazon EC2 instances with encrypted root device volumes. Configure the cluster in the CloudFormation template.
- B. Use a CloudFormation template to launch an EMR cluster. In the configuration section of the cluster, define a bootstrap action to enable TLS.
- C. Create a custom AMI with encrypted root device volumes. Configure Amazon EMR to use the custom AMI using the CustomAmiId property in the CloudFormation template.
- D. Use a CloudFormation template to launch an EMR cluster. In the configuration section of the cluster, define a bootstrap action to encrypt the root device volume of every node.

Correct Answer: C

Reference: <https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-custom-ami.html>

[DAS-C01 PDF Dumps](#)

[DAS-C01 Exam Questions](#)

[DAS-C01 Braindumps](#)



To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

Try our product !

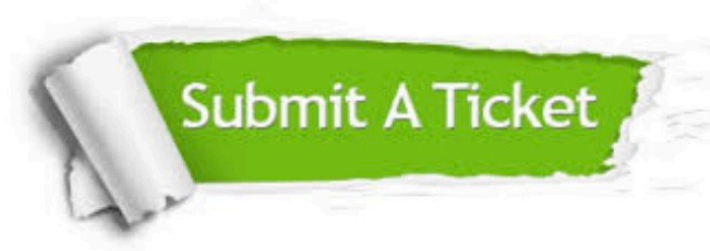
100% Guaranteed Success
100% Money Back Guarantee
365 Days Free Update
Instant Download After Purchase
24x7 Customer Support
Average 99.9% Success Rate
More than 800,000 Satisfied Customers Worldwide
Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [Kindle](#)

We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications. You can view Vendor list of All Certification Exams offered:

<https://www.lead4pass.com/allproducts>

Need Help

Please provide as much detail as possible so we can best assist you.
To update a previously submitted ticket:



 <p>One Year Free Update Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p>Money Back Guarantee To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p>Security & Privacy We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information & peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.
All trademarks are the property of their respective owners.
Copyright © lead4pass, All Rights Reserved.