


**Question 1****Domain :Continuous Improvement for Existing Solutions**

Your company has a logging microservice which is used to generate logs when users have entered certain commands in another application. This logging service is implemented via an SQS standard queue that an EC2 instance is listening to. However, you have found that on some occasions, the order of the logs are not maintained. As a result, it becomes harder to use this service to trace users' activities. How should you fix this issue in a simple way?

- ☐ A. Convert the existing standard queue into a FIFO queue. Add a deduplication ID for the messages that are sent to the queue.
- ☒ B. Delete the existing standard queue and recreate it as a FIFO queue. As a result, the order for the messages to be received is ensured. 
- ☐ C. Migrate the whole microservice application to SWF so that the operation sequence is guaranteed.
- ☐ D. The wrong order of timestamps is a limitation of SQS, which does not have a fix.

Explanation:**Correct Answer – B**

The FIFO queue improves upon and complements the standard queue. The most important features of this queue type are FIFO (First-In-First-Out) delivery and exactly-once processing. The FIFO queue is mainly used when the order to process the messages in the queue needs to be guaranteed without any items being out of order or duplicated.




- Option A is incorrect: Because you can't convert an existing standard queue into a FIFO queue. This is clarified in
 - <https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/FIFO-queues.html>.
- Option B is CORRECT: Because the FIFO queue is able to guarantee the sequence for users' operations so that the issue of the logging system is fixed.

- Option C is incorrect: Because this is not a straightforward method by changing the whole microservice to SWF. Option B is much simpler than this option.
- Option D is incorrect: Refer to the explanations in Option B.

Question 2

Domain :Migration Planning

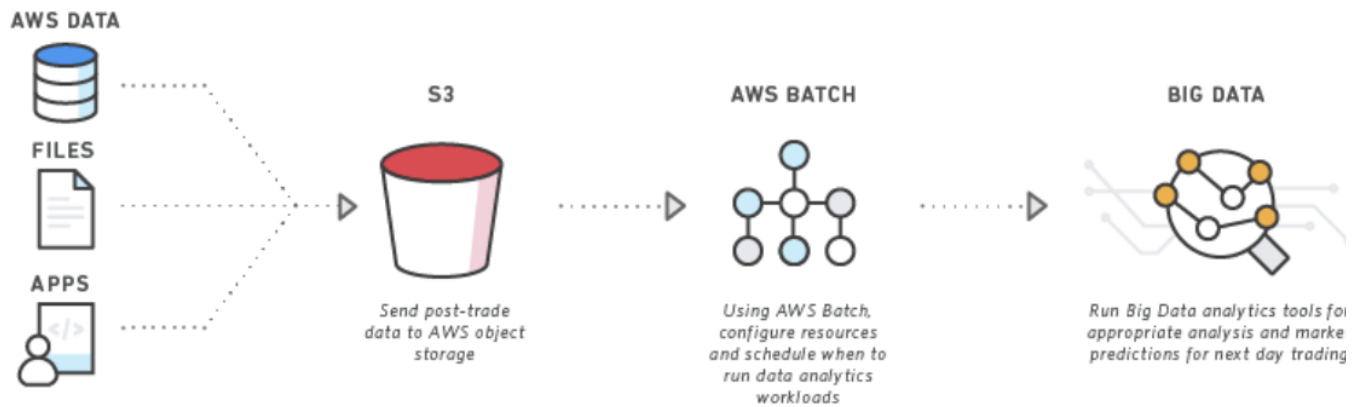
A large trading company is using an on-premise system to analyze the trade data. After the trading day closes, the data including the day's transaction costs, execution reporting, and market performance is sent to a Redhat server which runs big data analytics tools for predictions for next day trading. A bash script is used to configure resource and schedule when to run the data analytics workloads. How should the on-premise system be migrated to AWS with appropriate tools? (Select THREE)

- ☐ A. Create an S3 bucket to store the trade data that is used for post-processing. 
- ☐ B. Send the trade data from various sources to a dedicated SQS queue.
- ☐ C. Use AWS Batch to execute the bash script using a proper job definition. 
- ☐ D. Create EC2 instances with auto-scaling to handle the big data analytics workloads. 
- ☐ E. Use CloudWatch Events to schedule data analytics jobs.

Explanation:

Correct Answer – A, C and D

There are several parts of the on-premise system. The first is the place to store the data from several sources. The second is the bash script that is used to schedule the data to analyze the task. And the third part is the big data analysis. All of these three parts need to be considered when being migrated. Refer to the below chart as a reference:



- Option A is CORRECT: Because S3 is an ideal place to store trade data as it is highly available, durable, and cost-efficient.
- Option B is incorrect: Because the SQS queue is inappropriate to store source data. The trade data is very large every day which needs a durable store such as S3.
- Option C is CORRECT: Because AWS Batch is suitable to run a bash script using a job. The AWS Batch scheduler evaluates when, where, and how to run jobs.
- Option D is CORRECT: Because EC2 instances with auto-scaling can install big data analysis tools and process the data.
- Option E is incorrect: This case needs a batch job to allocate resources and execute the script to prepare for the downstream processing. CloudWatch Events cannot accomplish this mission.

Question 3



Domain :Migration Planning

A large IT company has an on-premise website which provides real-estate information such as renting, house prices and latest news to users. The website has a Java backend and a NoSQL MongoDB database that is used to store subscribers data. You are a cloud analyst and need to migrate the whole application to AWS platform. Your manager requires that a similar structure should be deployed in AWS for high availability.

Moreover, a tracing framework is essential which can record data from both the client request and the downstream call to the database in AWS. Which AWS services should you choose to implement the migration?

Select 3 Options.

- ☐ A. Deploy an autoscaling group of Java backend servers to provide high availability

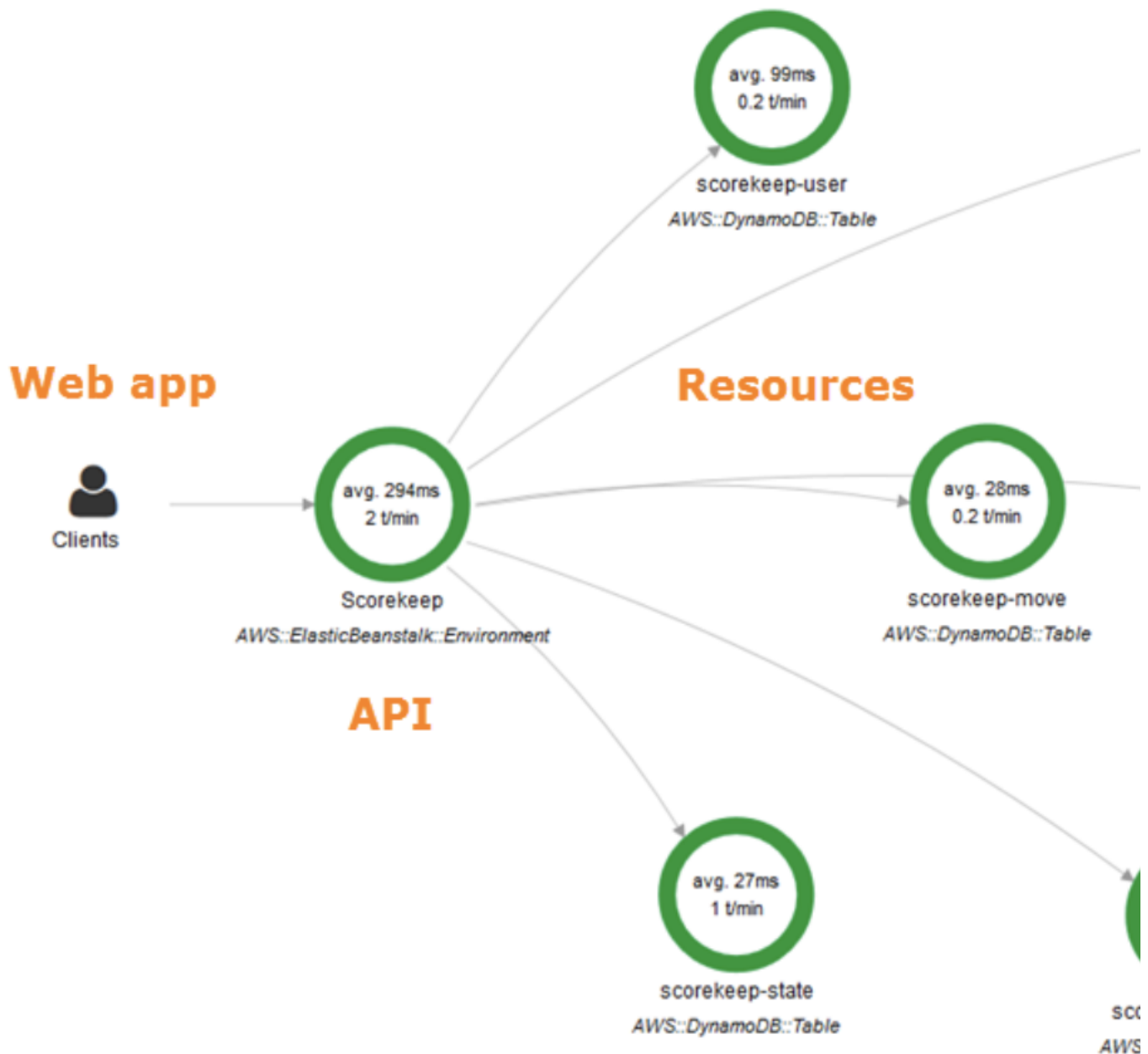
- ☐ B. Use RDS Aurora as the database for the subscriber data because it is highly available and can scale up to 15 Read Replicas.
- ☐ C. Create a DocumentDB database to hold subscriber data. Set up an autoscaling policy for the read/write throughput. 
- ☐ D. Use AWS X-Ray SDK to record data about incoming and outgoing requests. View the statistics graph in X-Ray console. 
- ☐ E. Trace the requests using AWS JAVA SDK and send logs to AWS CloudWatch Events. Create a CloudWatch dashboard to view the statistics.

Explanation:

Correct Answer – A, C, D

As this case needs to migrate the on-premise system to AWS using a similar structure, DocumentDB is more appropriate than RDS Aurora as DocumentDB is also a NoSQL database compatible with MongoDB. Besides, AWS X-Ray is a service that collects data about requests that your application serves, and provides tools you can use to view, filter, and gain insights into that data to identify issues and opportunities for optimization. Reference for AWS X-Ray is in <https://docs.aws.amazon.com/xray/latest/devguide/aws-xray.html>.

- Option A is CORRECT: Because autoscaling would provide high availability
- Option B is incorrect: Because RDS Aurora is a SQL database which is inappropriate in this scenario and brings extra unnecessary efforts.
- Option C is CORRECT: Because Amazon DocumentDB is a fully managed NoSQL database service that provides fast and predictable performance with seamless scalability and compatible with MongoDB
- Option D is CORRECT: Because AWS X-Ray is suitable to work as a tracing framework. It can monitor the requests from frontend and requests to DocumentDB database. Below is a graph that AWS X-Ray can provide:
- **Note:** Please read "DynamoDB" as "DocumentDB" instead in the diagram given below. Please refer the below link for "DocumentDB"
- <https://docs.aws.amazon.com/documentdb/latest/developerguide/developerguide.pdf>



- Option E is incorrect: Theoretically CloudWatch can be used to trace the incoming and outgoing request although it may bring extra efforts. However, the service should be CloudWatch Logs rather than CloudWatch Events. Amazon CloudWatch Events describe changes in Amazon Web Services (AWS) resources.


Question 4

Domain :Cost Control

You work in a video game company and your team is working on a feature that tells how many times that certain web pages have been viewed or clicked. You also created an AWS Lambda function to show some

key statistics of the data. You tested the Lambda function and it worked perfectly.

However, your team lead requires you to show the statistics every day at 8:00AM GMT on a big TV screen so that when employees come in to the office every morning, they have a rough idea of how the feature runs. What is the most cost efficient and straightforward way for you to make this happen?

- ☒ A. Create an AWS CloudWatch Events rule that is scheduled using a cron expression. Configure the target as the Lambda function. 
- ☐ B. Create an Amazon linux EC2 T2 instance and set up a Cron job using Crontab. Use AWS CLI to call your AWS Lambda every 8:00AM.
- ☐ C. Use Amazon Batch to set up a job with a job definition that runs every 8:00AM for the Lambda function.
- ☐ D. In AWS CloudWatch Events console, click "Create Event" using the cron expression "* * ? * * 08 00". Configure the target as the Lambda function.

Explanation:

Correct Answer– A

Potentially, more than one option may work. However this question asks the most cost efficient and straightforward method which is something that needs to be considered.

- Option A is CORRECT: Because AWS CloudWatch Events rule is free and quite easy to begin with. To schedule a daily event at 8:00AM GMT, you just need to set up a cron rule as given in the below screenshot

Event Source

Build or customize an Event Pattern or set a Schedule to invoke Targets.

☐ Event Pattern ⓘ ☒ Schedule ⓘ

☐ Fixed rate of

☒ Cron expression

Next 10 Trigger Date(s)

1. Fri, 08 Mar 2019 08:00:00 GMT
2. Sat, 09 Mar 2019 08:00:00 GMT
3. Sun, 10 Mar 2019 08:00:00 GMT
4. Mon, 11 Mar 2019 08:00:00 GMT
5. Tue, 12 Mar 2019 08:00:00 GMT
6. Wed, 13 Mar 2019 08:00:00 GMT
7. Thu, 14 Mar 2019 08:00:00 GMT
8. Fri, 15 Mar 2019 08:00:00 GMT
9. Sat, 16 Mar 2019 08:00:00 GMT
10. Sun, 17 Mar 2019 08:00:00 GMT

[Learn more](#) about CloudWatch Events schedules.

Targets

Select Target to invoke when an event is triggered.

[+ Add target*](#)

- Option B is incorrect: Because launching a new EC2 instance for this task is not cost efficient.
- Option C is incorrect: Because this is not something AWS Batch works. For AWS Batch, it runs as a containerized application on an Amazon EC2 instance in your compute environment.
- Option D is incorrect: Because firstly it should be "Create rule" rather than "Create Event". Secondly, the Cron expression of " * ? * * 08 00 " is incorrect.

For More information, Please check below AWS Docs:

- <https://docs.aws.amazon.com/AmazonCloudWatch/latest/events/ScheduledEvents.html>

Question 5

Domain :Continuous Improvement for Existing Solutions

A supermarket chain had a big data analysis system deployed in AWS. The system has the raw data such as clickstream or process logs in S3. A m3.large EC2 instance transformed the data to other formats and saved it to another S3 bucket. It was then moved to Amazon Redshift

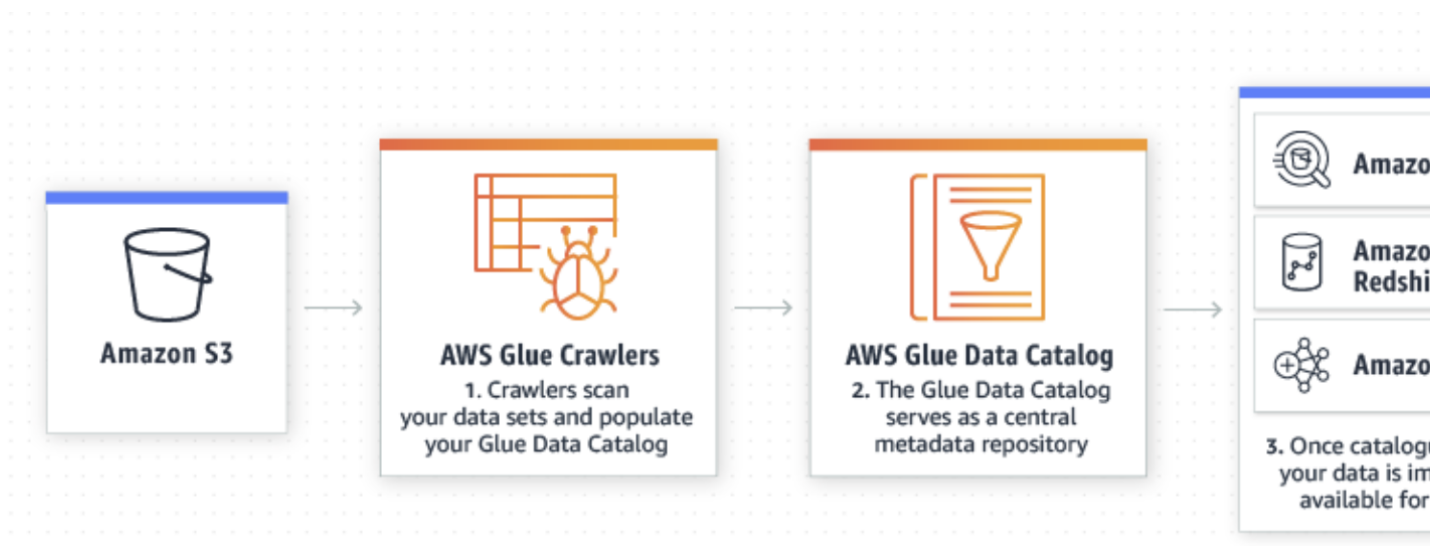
Your team is in charge of improving the system using AWS Glue which is a fully managed ETL (extract, transform, and load) service. Which tasks can AWS Glue simplify during re-establishing the big data system? (Select TWO)

- ☐ A. AWS Glue contains a crawler that connects to the S3 bucket and scans the dataset. Then the service creates metadata tables in the data catalog. ✓
- ☐ B. AWS Glue automatically generates code in Java to extract data from the source and transform the data to match the target schema.
- ☐ C. By default, AWS Glue creates a scheduler to trigger the activated tasks every minute.
- ☐ D. AWS Glue has a central metadata repository (data catalog). The data in the catalog is available for analysis immediately. ✓

Explanation:

Correct Answer – A, D

AWS Glue is a service to discover data, transform it, and make it available for search and querying. AWS Glue can make all your data in S3 immediately available for analytics without moving the data:





- Option A is CORRECT: Because Crawler is a key component in AWS Glue that can scan data in all kinds of repositories, classify it, extract schema information from it, and store the metadata automatically in the AWS Glue Data Catalog.
- Option B is incorrect: Because AWS Glue will generate ETL code in Scala or Python rather than Java.
- Option C is incorrect: Because AWS Glue does not generate trigger by default. Moreover, Cron expressions that lead to rates faster than 5 minutes are not supported.
- Option D is CORRECT: Because the data catalog of AWS Glue stores the metadata in the AWS Cloud which is readily available for analysis

Question 6

Domain :Continuous Improvement for Existing Solutions

An AWS Solution architect who is using EBS General Purpose SSD (gp2) volume type for his EBS volumes, now wants to modify some of these volumes without affecting the performance. What options would you suggest? (Select TWO)

- ☐ A. A 50GB gp2 root volume can be modified to an EBS Provisioned IOPS SSD (io1) without stopping the instance. 
- ☐ B. A gp2 volume that is attached to an instance as a root volume needs can be modified to a Throughput Optimized HDD (st1) volume.
- ☐ C. A 1GB gp2 volume that is attached to an instance as a non-root volume can be modified to a Cold HDD (sc1) volume.
- ☐ D. A 1TB gp2 volume that is attached to an instance as a non-root volume can be modified to a Throughput Optimized HDD (st1) volume without stopping the instance or detaching the volume. 

Explanation:

Correct Answer A, D

The EBS volume types can be modified in flight without the volume being detached or the instance being restarted. However, there are some limitations that need to be noticed. The details are in

- <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/modify-volume-requirements.html>

- Option A is CORRECT: Because the root volume can be changed to io1 such as:

Volume Type

 Filter by attributes



Size

General Purpose SSD (gp2)

Provisioned IOPS SSD (io1)

Magnetic (standard)

GiB, Ma

- Option B is incorrect: Because a gp2 volume that is attached to an instance as a root volume cannot be modified to an st1 or sc1 volume.
- Option C is incorrect: Because there is an limitation of minimum volume size for Cold HDD (sc1):

Volume Type

Cold HDD (sc1) ▼



Size

1

(Min: 500 GiB, Max: 16384



A Cold HDD (sc1) volume must be at least 500 GiB

- Option D is CORRECT: Because the volume size 1TB is suitable for Throughput Optimized HDD (st1):

Volume Type

Throughput Optimized HDD (st1) ▼





Size

1000

(Min: 500 GiB, Max: 16384

Which of the following are associated with using the "HLS" method of viewing the Kinesis video stream?
(Select TWO)

- ☐ A. A web application that is able to display the video stream using the third-party player Video.js. 
- ☐ B. In order to process Kinesis video streams, a SAAS provider needs to build a new video player which is integrated into their major online product.
- ☐ C. Able to view only live video, not archived video.
- ☐ D. Playback video by typing in the HLS streaming session URL in the location bar of the "Apple Safari Technology" browser for debug purpose. 

Explanation:

Correct Answer – A, D

Explanation:

For differences between GetMedia API and HLS, please refer to




<https://docs.aws.amazon.com/kinesisvideostreams/latest/dg/how-hls.html#how-hls-ex1-display>.

- Option A is CORRECT: Because third party player that supports HLS can be used to integrate with Kinesis Video Streams.
- Option B is incorrect: Because if a new player is needed which means that you have to build your own player, GetMedia API will be suitable.
- Option C is incorrect: You can use HLS to view an Amazon Kinesis video stream, either for live playback or to view archived video
 - Please refer to the following link on HLS
 - <https://docs.aws.amazon.com/kinesisvideostreams/latest/dg/how-hls.html#how-hls-ex1-display>
 - <https://docs.aws.amazon.com/kinesisvideostreams/latest/dg/how-playback.html>
- Option D is CORRECT: Because the Apple Safari Technology browser can play back video if the HLS streaming session URL is typed in the location bar.

Question 8

Domain :Design for New Solutions

A team has just received a task to build an application that needs to recognize faces in streaming videos. They will get the source videos from a third party which use a container format (MKV). The APP should be able to quickly address faces through the video in real time and save the output in a suitable manner for downstream to process. As recommended by the AWS Solutions Architect colleague, they would like to develop the service using AWS Rekognition. Which below options are needed to accomplish the task? Select 3.

- ☐ A. S3 buckets to store the source MKV videos for AWS Rekognition to process. S3 should be used in this case as it has provided an unlimited, highly available and durable storing space. Make sure that the third party has the write access to S3 buckets.
- ☐ B. A Kinesis video stream for sending streaming video to Amazon Rekognition Video. This can be done by using Kinesis "PutMedia" API in Java SDK. The PutMedia operation writes video data fragments into a Kinesis video stream that Amazon Rekognition Video consumes. 
- ☐ C. An Amazon Rekognition Video stream processor to manage the analysis of the streaming video. It can be used to start, stop, and manage stream processors according to needs. 
- ☐ D. Use EC2 or Lambda to call Rekognition API "DetectFaces" with the source videos saved in S3 bucket. For each face detected, the operation returns face details. These details include a bounding box of the face, a confidence value, and a fixed set of attributes such as facial landmarks, etc.
- ☐ E. After the APP has utilized Rekognition API to fetch the recognized faces from live videos, use S3 or RDS database to store the output from Rekognition. Another lambda can be used to post-process the result and present to UI.
- ☐ F. A Kinesis data stream consumer to read the analysis results that Amazon Rekognition Video sends to the Kinesis data stream. It can be an Amazon EC2 instance by adding to one of Amazon Machine Images (AMIs). The consumer can be autoscaled by running it on multiple Amazon EC2 instances under an Auto Scaling group. 

Explanation:

Correct Answer – B, C, F

Explanation:

For facial recognition in live videos, it is different from that in photos. Kinesis is required to meet the needs of real time process. Amazon Rekognition Video uses Amazon Kinesis Video Streams to receive and process a video stream. The analysis results are output from Amazon Rekognition Video to a Kinesis data stream and then read by your client application.

Amazon Rekognition Video provides a stream processor ([CreateStreamProcessor](#)) that you can use to start and manage the analysis of streaming video.

As a summary, the below 3 items are needed for Amazon Rekognition Video with streaming video:

- A Kinesis video stream for sending streaming video to Amazon Rekognition Video. For more information, see [Kinesis video stream](#).
- An Amazon Rekognition Video stream processor to manage the analysis of the streaming video. For more information, see [Starting Streaming Video Analysis](#).
- A Kinesis data stream consumer to read the analysis results that Amazon Rekognition Video sends to the Kinesis data stream. For more information, see [Consumers for Amazon Kinesis Streams](#).
- Option A is incorrect: Because the source videos should be put into Kinesis video stream instead of S3. Afterwards, Rekognition processor will pick up records in Kinesis stream to process.
- Option B is CORRECT: Because it is the step to convert source data into Kinesis video stream.
- Option C is CORRECT: A stream processor can be created by calling [CreateStreamProcessor](#). The request parameters include the Amazon Resource Names (ARNs) for the Kinesis video stream, the Kinesis data stream, and the identifier for the collection that's used to recognize faces in the streaming video. It also includes the name that you specify for the stream processor.

Below is an example:

```
{
  "Name": "streamProcessorForCam",
  "Input": {
    "KinesisVideoStream": {
      "Arn": "arn:aws:kinesisvideo:us-east-1:nnnnnnnnnnnn:stream/inputVideo"
    }
  },
  "Output": {
    "KinesisDataStream": {
      "Arn": "arn:aws:kinesis:us-east-1:nnnnnnnnnnnn:stream/outputData"
    }
  },
  "RoleArn": "arn:aws:iam::nnnnnnnnnnnn:role/roleWithKinesisPermission",
  "Settings": {
    "FaceSearch": {
      "CollectionId": "collection-with-100-faces",
      "FaceMatchThreshold": 85.5
    }
  }
}
```

- Option D is incorrect: Because for video processing, Rekognition API "DetectFaces" should not be used. "DetectFaces" is used to detect faces within an image that is provided as input. Instead, stream

processor relevant APIs should be used.

- Option E is incorrect: Because the output from Rekognition should be stored in Kinesis data stream. When Rekognition stream processor is created, the Rekognition output (Kinesis Data Stream) is defined:


```
"Output": {  
  "KinesisDataStream": {  
    "Arn": "arn:aws:kinesis:us-east-1:nnnnnnnnnnnnn:stream/outputData"  
  }  
}
```

- Option F is CORRECT: Because it describes correctly on how to consume the Kinesis data stream. You can use the Amazon Kinesis Data Streams Client Library to consume analysis results that are sent to the Amazon Kinesis Data Streams output stream. Details can be found in <https://docs.aws.amazon.com/rekognition/latest/dg/streaming-video-kinesis-output.html>.

Question 9

Domain :Design for Organizational Complexity

A large company starts to use AWS organizations with consolidated billing feature to manage its separate departments. The AWS operation team has just created 3 OUs (organization units) with 2 AWS accounts each. To be compliant with company-wide security policy, CloudTrail is required for all AWS accounts which is already been set up. However after some time, there are cases that users in certain OU have turned off the CloudTrail of their accounts. What is the best way for the AWS operation team to prevent this from happening again?

- ☒ A. Update the AWS Organizations feature sets to "All features" and then create a Service Control Policies (SCP) to Prevent Users from Disabling AWS CloudTrail. This can be achieved by a deny policy with cloudtrail:StopLogging denied. 
- ☐ B. This can be achieved by Service Control Policies (SCP) in "All features" set. The team needs to delete and recreate the AWS Organizations with "All features" enabled and then use a proper control policy to limit the operation of cloudtrail:StopLogging.
- ☐ C. In each AWS account in this organization, create an IAM policy to deny cloudtrail:StopLogging for all users including administrators.
- ☐ D. Use a Service Control Policies (SCP) to prevent users from disabling AWS CloudTrail. This can be done by a allow policy which denies cloudtrail:StopLogging

Explanation:

Correct Answer – A

Explanation:

AWS Organizations has provided two feature sets:

- Consolidated billing – This feature set provides shared billing functionality, but does not include the more advanced features of AWS Organizations.
- All features – The complete feature set that is available to AWS Organizations. It includes all the functionality of consolidated billing, plus advanced features that give you more control over accounts in your organization. For example, when all features are enabled the master account of the organization has full control over what member accounts can do. The master account can apply SCPs to restrict the services and actions that users (including the root user) and roles in an account can access, and it can prevent member accounts from leaving the organization.


In this case, we should use "All features". One thing to note is that the feature sets can be upgraded in flight. It does not need to delete/recreate the AWS Organizations.

- Option A is CORRECT: Because SCP is suitable for limiting actions that AWS accounts in an Organization can do. Below is an example for a deny policy:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Deny",
      "Action": "cloudtrail:StopLogging",
      "Resource": ""
    }
  ]
}
```

- Option B is incorrect: Because it does not need to delete/recreate the AWS Organizations in order to upgrade feature sets.
- Option C is incorrect: Because although it can potentially work, it has lots of repeatable work and is not straightforward if compared with Option A.
- Option D is incorrect: Because it does not mention the upgrade of feature sets. Secondly, the allow policy is incorrect as this case only requires limiting CloudTrail deletion. allow policy implicitly prevents everything except for several allow items.

A mobile App developer just made an App in both IOS and Android that has a feature to count step numbers. He has used AWS Cognito to authorize users with a user pool and identity pool to provide access to AWS DynamoDB table. The App uses the DynamoDB table to store user subscriber data and number of steps. Now the developer also needs Cognito to integrate with Google to provide federated authentication for the mobile application users so that user does not need to remember extra login access. What should the developer do to make this happen for the IOS and Android App?

- ☒ A. **Amazon Cognito Identity pools (federated identities) support user authorization through federated identity providers—including Amazon, Facebook, Google, and SAML identity providers. The developer just needs to set up the federated identities for Google access** 
- ☐ B. Only Android works for federated identities if Google access is required for AWS Cognito. This can be done by configuring Cognito identity pools with a Google Client ID.
- ☐ C. Amazon Cognito User pools support user authentication through federated identity providers—including Amazon, Facebook, Google, and SAML identity providers. The developer just needs to set up the federated identities for Google access in Cognito User pool.
- ☐ D. Only IOS (Objective-C and Swift) works for federated identities if Google access is required for AWS Cognito. This can be done by configuration Cognito identity pools with a Google Client ID. Google federated access does not work for android app.

Explanation:

Correct Answer – A

Explanation:

One common use case for Amazon Cognito is to access AWS Services with an Identity Pool. For the Identity pool itself, it can include:


- Users in an Amazon Cognito identity pool.
- Users who authenticate with external identity providers such as Facebook, Google, or a SAML-based identity provider.
- Users authenticated via your own existing authentication process.
- Option A is CORRECT because the User Pool is where the federated identity would be set-up and the Identity Pool is where permissions would be granted.
- Option B is incorrect: Because Google federated identities work for both Android and IOS. Refer to

- <https://docs.aws.amazon.com/cognito/latest/developerguide/google.html> on the details.
- Option C is incorrect: In order to set up federated identities, the developer needs to configure Identity Pool instead of User Pool.
- <https://docs.aws.amazon.com/cognito/latest/developerguide/getting-started-with-identity-pools.html> describes on how to use Amazon Cognito Identity Pools (Federated Identities).
- Option D is incorrect: Same reason as Option B.

Question 11

Domain :Migration Planning

A big company has a service to process gigantic clickstream data sets which are often the result of holiday shopping traffic on a retail website, or sudden dramatic growth on the data network of a media or social networking site. It is becoming more and more complex to analyze these clickstream datasets for its on-premise infrastructure. As the sample data set keeps growing, fewer applications are available to provide a timely response. The service is using a Hadoop cluster with Cascading. How can they migrate the applications to AWS in the best way?

- ☐ A. Put the source data to S3 and migrate the processing service to an AWS EMR hadoop cluster with Cascading. Enable EMR to directly read and query data from S3 buckets. Write the output to RDS database
- ☐ B. Put the source data to a Kinesis stream and migrate the processing service to AWS lambda to utilize its scaling feature. Enable lambda to directly read and query data from Kinesis stream. Write the output to RDS database
- ☐ C. Put the source data to a S3 bucket and migrate the processing service to AWS EC2 with auto scaling. Ensure that the auto scaling configuration has proper maximum and minimum number of instances. Monitor the performance in Cloudwatch dashboard. Write the output to DynamoDB table for downstream to process.
- ☒ D. Put the source data to a Kinesis stream and migrate the processing service to an AWS EMR cluster with Cascading. Enable EMR to directly read and query data from Kinesis streams. Write the output to Redshift. 

Explanation:

Correct Answer – D

Explanation:

The application needs to process data timely therefore Kinesis stream should be considered first. After a click event happens, a message should be put into Kinesis stream in real time. Moreover, Cascading is a proven, highly extensible application development framework for building massively parallelized data applications on EMR. By using EMR, the application does not need to change a lot for the migration. Refer to <https://aws.amazon.com/blogs/big-data/integrating-amazon-kinesis-amazon-s3-and-amazon-redshift-with-cascading-on-amazon-emr/>.

- Option A is incorrect: Because the output to RDS database is improper as RDS does not scale well when the traffic is high. Redshift is much more appropriate.
- Option B is incorrect: AWS Lambda can potentially work for Hadoop however EMR provides native support for Hadoop. Also RDS is incorrect.
- Option C is incorrect: Because EC2 is not suitable for Hadoop processing if compared with EMR. This question asks for the best option so EMR should be chosen.
- Option D is CORRECT: Because EMR cluster with Cascading can process the data from Kinesis stream in real time and Redshift is also a proper place to store the output data.

Question 12

Domain :Continuous Improvement for Existing Solutions

An Artificial Intelligence startup company has used lots of EC2 instances. Some instances use the SQL Server database while others use Oracle. As the data needs to be kept secure, regular snapshots are required. They want SQL Server EBS volume to take a snapshot every 12 hours. However, for Oracle, it only needs a snapshot every day. Which option below is the best one that the company should choose?

- ☐ A. Use a free third-party tools such as Clive to Manage EC2 instance lifecycle. It can design various backup policies for EC2 EBS volumes. Add a 12 hours backup policy to SQL Server EBS volumes and a 24 hours backup policy to Oracle EBS volumes.
- ☐ B. Add a prefix to the name of both SQL Server and Oracle EBS volumes. In the AWS Data Lifecycle Management console, create two management policies based on the name prefix. For example, add a 12 hours backup schedule to EBS volumes with a name starting with "SQL" and add a 24 hours backup schedule to EBS volumes with a name starting with "oracle".
- ☐ C. Create a dedicate Lambda function to differentiate EC2 EBS volumes and take snapshots. Set up Cloudwatch Events Rules to call the lambda so that the function runs every 12 hours for SQL Server and 24 hours for Oracle.

- D. Add different tags for SQL Server and Oracle EBS volumes. In the AWS Data Lifecycle Management console, create two management policies based on the tags. Add a 12 hours schedule to SQL Server lifecycle policy and a 24 hours schedule to Oracle lifecycle policy



Explanation:

Correct Answer – D

Explanation:

Amazon Data Lifecycle Manager (Amazon DLM) should be considered if automating snapshot management is required (<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/snapshot-lifecycle.html>). One thing to note is that DLM policy has used Tags to choose EBS volumes:

[Policies](#) > Create Snapshot Lifecycle Policy

Create Snapshot Lifecycle Policy

Data Lifecycle Manager for EBS Snapshots will help you automate the creation and deletion of EBS snapshots based on a schedule. Volumes are targeted by tags

Description*



Target volumes with tags This policy will be applied to volumes with **any** of the following tags.



You cannot use tags that are in use by another enabled or disabled lifecycle policy.

*



A value is required

- Option A is incorrect: Because although this may work, it is not a straightforward solution as DLM.
- Option B is incorrect: Because prefix in the name is incorrect. Tags (name and value) are used to choose EBS volumes.
- Option C is incorrect: Because it brings extra cost and is not as easy as DLM. The question asks for the solution without extra cost.
- Option D is CORRECT: Because two management policies in DLM can meet the needs.

For example:

For SQL Server EBS volume:

Schedule name*

Schedule-SQL



Create snapshots every

12



Hours



Snapshot creation start time

09 : 00

UTC

For Oracle EBS volume:

Schedule name*

Schedule-Oracle



Create snapshots every

24



Hours



Snapshot creation start time


09 : 00

UTC

Question 13

Domain :Design for New Solutions

API gateway and Lambda non-proxy integrations have been chosen to implement an application by a software engineer. The application is a data analysis tool that returns some statistic results when the HTTP endpoint is called. The lambda needs to communicate with some back-end data services such as Keen.io however there are chances that error happens such as wrong data requested, bad communications, etc. The lambda is written using Java and two exceptions may be returned which are `BadRequestException` and `InternalServerErrorException`. What should the software engineer do to map these two exceptions in API gateway with proper HTTP return codes? For example, `BadRequestException` and `InternalServerErrorException` are mapped to HTTP return codes 400 and 500 respectively. Select 2.

- ☐ A. Add the corresponding error codes (400 and 500) on the Integration Response in API gateway
- ☐ B. Add the corresponding error codes (400 and 500) on the Method Response in API gateway. 
- ☐ C. Put the mapping logic into Lambda itself so that when exception happens, error codes are returned at the same time in a JSON body.

D. Add Integration Responses where regular expression patterns are set such as BadRequest or InternalError. Associate them with HTTP status codes



E. Add Method Responses where regular expression patterns are set such as BadRequest or InternalError. Associate them with HTTP status codes 400 and 500.

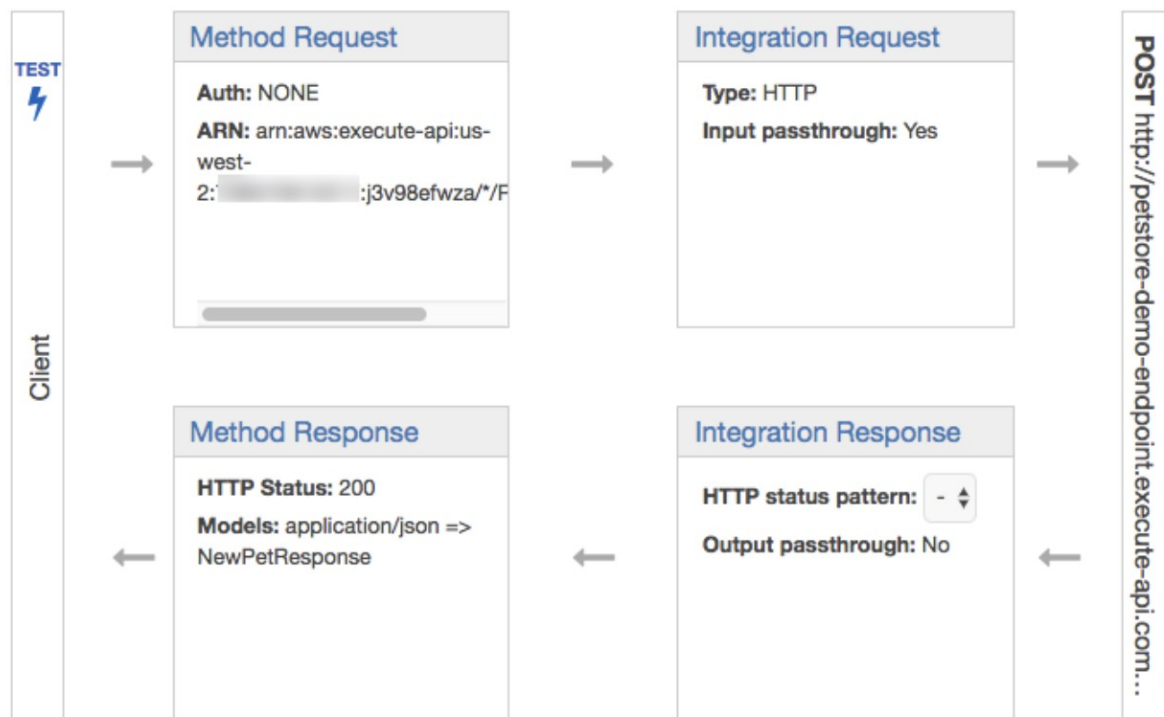
Explanation:

Correct Answer – B, D

Explanation:

When an API gateway is established, there are four parts:

/pets - POST - Method Execution



Method Request/Method Response are part mainly deal with API gateways and they are the API's interface with the API's frontend (a client), whereas Integration Request and Integration Response are the API's interface with the backend. In this case, the backend is a lambda.

For the mapping of exceptions that come from Lambda, Integration Response is the correct place to configure. However, the corresponding error code (400) on the method response should be created first. Otherwise, API Gateway throws an invalid configuration error response at runtime. The below is an example to map `BadRequestException` to HTTP return code 400:

First, declare response types using [Method Response](#). Then, map the possible responses from the backend to this method's response types.

	Lambda Error Regex	Method response status	Output model	Default mapping
-				

Map the output from your Lambda function to the headers and output model of the method response.

Lambda Error Regex ⓘ

Method response status

Content handling ⓘ



[Cancel](#) [Save](#)

- Option A is incorrect: Because HTTP error codes are defined firstly in Method Response instead of Integration Response.
- Option B is CORRECT: Because HTTP error codes are defined firstly in Method Response instead of Integration Response. (Same reason as A).
- Option C is incorrect: Because Integration Response in API gateway should be used. Refer to <https://docs.aws.amazon.com/apigateway/latest/developerguide/handle-errors-in-lambda-integration.html> on "how to Handle Lambda Errors in API Gateway".
- Option D is CORRECT: Because BadRequest or InternalError should be mapped to 400 and 500 in Integration Response settings.
- Option E is incorrect: Because Method Response is the interface with frontend. It does not deal with how to map the response from Lambda/backend.

Question 14

Domain :Continuous Improvement for Existing Solutions

An IT company owns a web product in AWS that provides discount restaurant information to customers. It has used one S3 Bucket (my-bucket) to store restaurant data such as pictures, menus, etc. The product is deployed in VPC subnets. The company's Cloud Architect decides to configure a VPC endpoint for this S3 bucket so that the performance will be enhanced. To be compliance to security rules, it is required that the new VPC endpoint is only used to communicate with this specific S3 Bucket and on the other hand, the S3 bucket only allows the read/write operations coming from this VPC endpoint. Which two options should the Cloud Architect choose to meet the security needs?

- ☐ A. Use a VPC Endpoint policy for Amazon S3 to restrict access to the S3 Bucket "my-bucket" so that the VPC Endpoint is only allowed to perform S3 actions on "my-bucket". 
- ☐ B. Modify the security group of the EC2 instance to limit the outbound actions to the VPC Endpoint if the outgoing traffic destination is the S3 bucket "my-bucket".
- ☐ C. In the S3 bucket "my-bucket", add a S3 bucket policy in which all actions are denied if the source IP address is not equal to the EC2 public IP (use "NotIpAddress" condition).
- ☐ D. For the S3 bucket "my-bucket", use a S3 bucket policy that denies all actions if the source VPC Endpoint is not equal to the endpoint ID that is created. 
- ☐ E. Create a S3 bucket policy in the S3 bucket "my-bucket" which denies all actions unless the source IP address is equal to the EC2 public IP (use "IpAddress" condition).

Explanation:

Correct Answer – A, D

In this case, two restrictions are required:

1, For the VPC endpoint, restricting access to the specific S3 Bucket "my-bucket". A VPC Endpoint policy is needed:

```
{
  "Statement": [
    {
      "Sid": "Access-to-my-bucket-only",
      "Principal": "*",
      "Action": [
        "s3:GetObject",
        "s3:PutObject"
      ],
      "Effect": "Allow",
      "Resource": ["arn:aws:s3:::my-bucket",
        "arn:aws:s3:::my-bucket/*"]
    }
  ]
}
```

2, For the S3 bucket "my-bucket", restricting access to the new VPC Endpoint. S3 Bucket policy is required:

```
{
  "Version": "2012-10-17",
```

```

"Id": "Policy1415115909152",
"Statement": [
  {
    "Sid": "Access-to-specific-VPCE-only",
    "Principal": "*",
    "Action": "s3:*",
    "Effect": "Deny",
    "Resource": ["arn:aws:s3:::my-bucket",
      "arn:aws:s3:::my-bucket/*"],
    "Condition": {
      "StringNotEquals": {
        "aws:sourceVpce": "vpce-1a2b3c4d"
      }
    }
  }
]
}

```

In terms of S3 bucket policy for VPC Endpoint, the `aws:SourceIp` condition can not be used as for either `NotIpAddress` or `IpAddress`, the condition fails to match any specified IP address or IP address range. Instead, the specific endpoint ID should be used for the S3 bucket policy.


Explanation:

- Option A is CORRECT: Because VPC Endpoint policy helps on restricting which entity is able to use the VPC Endpoint. It is an IAM resource policy that you attach to an endpoint when you create or modify the endpoint.
- Option B is incorrect: Because security group cannot limit the actions for VPC endpoints.
- Option C is incorrect: Because for S3 bucket policy, `NotIpAddress` condition is always met for VPC endpoint so that it cannot help on restricting the traffic from VPC endpoint.
- Option D is CORRECT: Because in S3 bucket policy, a rule can be set up to deny all actions if the incoming traffic is not from the VPC Endpoint ID.
- Option E is incorrect: Same reason as option C.

Question 15

Domain :Design for New Solutions

You work for an e-commerce retailer as an AWS Solutions Architect. Your company is looking to improve customer loyalty programs by partnering with other third-parties to offer a more comprehensive selection of customer rewards. You plan to use Amazon Managed Blockchain to implement a blockchain network that allows your company and third-parties to share and validate rewards information quickly and transparently. How do you add members for this blockchain?

- ☒ A. When Amazon Managed Blockchain is set up, there is an initial member in the AWS account. Then new members can be added in this AWS account by sending an invitation or a network invitation can be created for a member in a different AWS account 
- ☐ B. While Amazon Managed Blockchain is configured, there is an initial member in the AWS account. Then new members can be added in this AWS account without having to send an invitation. You cannot add new members for other AWS accounts
- ☐ C. When Amazon Managed Blockchain is created, there would be no member in the AWS account. Then new members can be added in this AWS account or other accounts by sending out an an invitation.
- ☐ D. When Amazon Managed Blockchain is first created, there would be no member in the AWS account. Then new members can be added in this AWS account. For other accounts, they can join this net blockchain network by using the network ID.

Explanation:

Correct Answer – A

Explanation:

By using Amazon Managed Blockchain, it is very convenient to manager members. New members can be added in your own account without having to send an invitation to yourself, or you can create a network invitation for a member in a different AWS account. Refer to <https://docs.aws.amazon.com/managed-blockchain/latest/managementguide/get-started-joint-channel.html> on how to "Invite Another AWS Account to be a Member and Create a Joint Channel".

- Option A is CORRECT: Because for members in other AWS accounts, you need to send out an invitation.
- Option B is incorrect: Because you can invite members from other AWS accounts.
- Option C is incorrect: Because there is already an initial account while the blockchain is set up.
- Option D is incorrect: Because there is already an initial account while the blockchain is set up. Also an invitation is required for members in other AWS accounts.