Application Integration
(SQS, SNS, SWF, Step Functions)

AWS Certified Solutions Architect– Professional (SAP-C01) - Study notes - Sep'2019

Simple Workflow Service (SWF)

The fundamental concept in Amazon SWF is the workflow.

- A workflow is a set of activities that carry out some objective, together with logic that coordinates the activities.
- Each workflow runs in an AWS resource called a domain, which controls the workflow's scope.
- An AWS account can have multiple domains, each of which can contain multiple workflows, but workflows in different domains can't interact.

Worker(s)

- is a program that receives activity tasks, performs them, and provides results back.
- Note that the task itself might actually be performed by a person, in which case the person would use the activity worker software for the receipt and disposition of the task.
  - An example might be a statistical analyst, who receives sets of data, analyzes them, and then sends back the analysis.

When using Amazon SWF, workers are implemented to perform tasks.

- These workers (workers are code specific to carry out a function or more) can run either on cloud infrastructure, such as Amazon EC2 instances, Lambda functions, or Mobile or tablet devices
- Or they can be also **on premise** behind corporate firewalls

Activity tasks

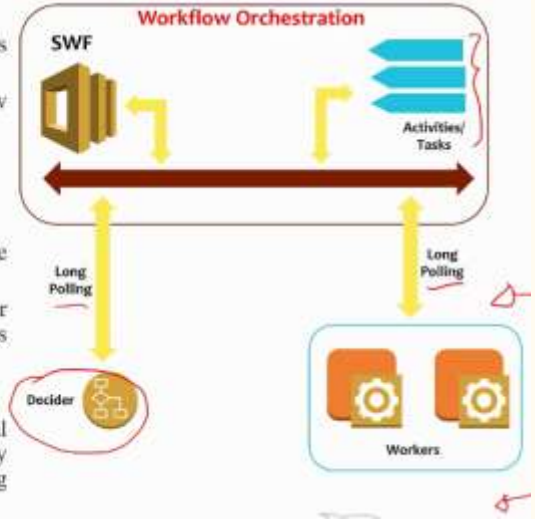
- Are performed by workers and can run synchronously or asynchronously.
- They can be distributed across multiple computers, **potentially in different geographic regions**, or they can all run on the same computer.

- Tasks can be created to be long-running, or that may fail, time out, or require restarts
- Amazon SWF stores tasks and assigns them to workers when they are ready, tracks their progress, and maintains their state, including details on their completion.
- To coordinate tasks, programs needs to coded that gets the latest state of each task from Amazon SWF and uses it to initiate subsequent tasks.
- Amazon SWF maintains an application's execution state durably so that the application is resilient to failures in individual components.
  - With Amazon SWF, you can implement, deploy, scale, and modify these application components independently.

Simple Workflow Service (SWF)

Decider:

- Is the implementation of the coordination logic in a workflow and is contained in a software program.
  - Deciders control the flow of activity tasks in a workflow execution.
- The decider schedules activity tasks,
  - It provides input data to the activity workers,
  - It processes events that arrive while the workflow is in progress,
  - and ultimately ends (or closes) the workflow when the objective has been completed.
- The mechanism by which both the activity workers and the decider receive their tasks (activity tasks and decision tasks respectively) is by polling the Amazon SWF service.



The role of the Amazon SWF service is to function as a reliable central hub through which data is exchanged between the decider, the activity workers, and other relevant entities such as the person administering the workflow.

The decider directs the workflow by receiving decision tasks from Amazon SWF and responding back to Amazon SWF with decisions.

- A decision represents an action or set of actions which are the next steps in the workflow.
- A typical decision would be to:
  - Schedule an activity task.
  - Set timers to delay the execution of an activity task,
  - Request cancellation of activity tasks already in progress, and
  - Complete or close the workflow.

Execution History

- Amazon SWF also maintains the state of each workflow execution,
  - This saves the burden of storing the state by the application itself

Amazon SWF informs the decider of the state of the workflow by including, with each decision task, a copy of the current workflow execution history.

There are three types of tasks in Amazon SWF:

- Activity task
  - An Activity task tells an activity worker to perform its function, such as to check inventory or charge a credit card.
  - The activity task contains all the information that the activity worker needs to perform its function.
- Lambda task
  - A Lambda task is similar to an Activity task, but executes a Lambda function instead of a traditional Amazon SWF activity.
- Decision task
  - A Decision task tells a decider that the state of the workflow execution has changed so that the decider can determine the next activity that needs to be performed.
  - The decision task contains the current workflow history.

Each workflow runs in an AWS resource called a domain, which controls the workflow's scope.

To reduce latency and to store data in a location that meets your requirements, Amazon SWF provides endpoints in different regions.

Each endpoint in Amazon SWF is completely independent; any domains, workflows and activities registered in one region don't share any data or attributes with those in another.

When Amazon SWF domain workflow or activity is registered, it exists only within the region you registered it in.

Simple Workflow Service (SWF) vs SQS

Both Amazon SQS and Amazon SWF are services that facilitate the integration of applications or microservices:

The following are the main differences between Amazon SQS and Amazon SWF:

- Amazon SWF API actions are task-oriented.
  - **Amazon SQS API actions are message-oriented.**
- Amazon SWF keeps track of all tasks and events in an application.
  - **Amazon SQS requires customers to implement their own application-level tracking, especially if the application uses multiple queues.**
- The Amazon SWF Console and visibility APIs provide an application-centric view that allows to search for executions, drill down into an execution's details, and administer executions.
  - **Amazon SQS requires implementing such additional functionality.**
- Amazon SWF offers several features that facilitate application development, such as passing data between tasks, signaling, and flexibility in distributing tasks.
  - **Amazon SQS requires customers to implement some application-level functionality.**

AWS Step Functions (enhanced version of SWF)

AWS Step Functions is a fully managed service that makes it easy to coordinate the components of distributed applications and microservices using **visual workflows**.

- Step Functions is a reliable way to coordinate components and step through the functions of the application.

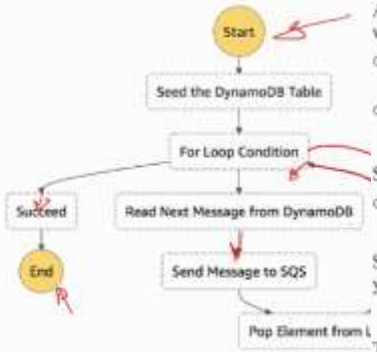
Step Functions provides a graphical console to arrange and visualize the components of the application as a series of steps.

- Makes it simple to build and run multi-step applications.

Step Functions service automatically triggers and tracks each step, and retries when there are errors, so the application executes in order and as expected.

- Step Functions logs the state of each step for easier diagnosis and debugging

It is possible to change and add steps without even writing code, so application can easily evolve and innovate faster.



AWS Step Functions allows to coordinate individual tasks by expressing the workflow as a finite state machine, written in the Amazon States Language.

- Using AWS Step Functions, define state machines that describe the workflow as a series of steps, their relationships, and their inputs and outputs.
- A finite state machine can express an algorithm as a number of states, their relationships, and their input and output.

State machines contain a number **of states**.

- Each state represents an individual step in a workflow diagram.

States can perform work, make choices, pass parameters, initiate parallel execution, manage timeouts, or terminate your workflow with a success or failure.

The visual console automatically graphs each state in the order of execution, making it easy to design multi-step applications.

The console highlights the real-time status of each step and provides a detailed history of every execution.

AWS Step Functions – Activity Tasks and Service Tasks

State machines work by creating/using activity tasks and service tasks.

- A task performs work by using an activity or an AWS Lambda function, or by passing parameters to the API actions of other services.

Activity Tasks

- Is how a specific step (task) in the workflow (state machine) can be assigned to an activity worker (code running some where else)
  - Activity workers can be any application, anywhere, that can communicate via HTTP
  - Activity workers can be run on AWS EC2 instances, a mobile device, in an on-premise DC
  - Activity workers poll work from and takes input from Step Functions, does the work, and returns results to Step functions.

Service Tasks

- Is how a step in the state machine (workflow) can connect to a supported AWS service.
- Step Functions pushes requests to other services so they can perform actions for the workflow, waits for the service task to complete, and then continues to the next step.

Workflows that are created with AWS Step Functions can connect and coordinate other supported AWS services using service tasks, this includes:

- Invoke an AWS Lambda function
- Run an Amazon ECS or AWS Fargate task
- Get or Put an item from an Amazon DynamoDB table
- Submit an AWS Batch job and wait for it to complete
- Publish a message to an Amazon SNS topic
- Send a message to an Amazon SQS queue
- Start an AWS Glue job run
- Create an Amazon SageMaker job to train a machine learning model or batch transform a data set
- Have Step Functions wait for a task to return a specific task token.

AWS Step Functions (enhanced version of SWF)

AWS Step Functions – Use Cases

AWS Step Functions are suitable for any computational problem or business process that can be subdivided into a series of steps (or a State Machine as the workflow is called under Step Functions).

- It's also useful for creating end-to-end workflows to manage jobs with interdependencies.

Here are some common use cases:

- **Data processing:**
  - Consolidate data from multiple databases into unified reports, refine and reduce large data sets into useful formats, or coordinate multi-step analytics and machine learning workflows
- **DevOps and IT automation:**
  - Build tools for CI/CD, or create event-driven applications that automatically respond to changes in infrastructure
- **E-commerce:**
  - Automate mission-critical business processes, such as order fulfillment and inventory tracking
- **Web applications:**
  - Implement robust user registration processes and sign-on authentication

Amazon SWF vs. AWS Step Functions

AWS Step Functions is a fully managed service that makes it easy to coordinate the components of distributed applications and microservices using visual workflows.

Instead of writing a Decider program as in SWF, Customers define state machines in JSON.

Consider using Step Functions for new applications. If Step Functions does not fit the needs, then consider Amazon SWF

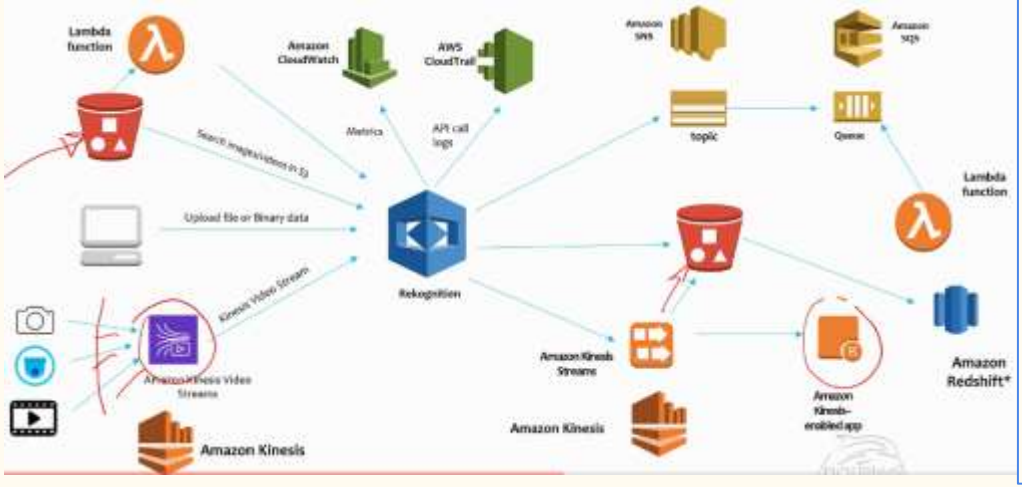
Use AWS SWF if:

- An external signal in the workflow is required
- Or, when it is required to launch child processes that return a result to a parent

Amazon SWF provides complete control over the orchestration logic, but increases the complexity of developing applications.

- AWS will continue to provide the Amazon SWF service, Flow framework, and support all Amazon SWF customers.

AWS Rekognition



Common use cases

Searchable image and video libraries
- Amazon Rekognition makes images and stored videos searchable

Face-based user verification
- Can be used in building access or similar applications
- Compares a live image to a reference image

Sentiment and demographic analysis
- Amazon Rekognition detects emotions such as happy, sad, or surprise, and demographic information such as gender from facial images.
- Rekognition can analyze images, and send the emotion and demographic attributes to Amazon Redshift for periodic reporting on trends such as in store locations and similar scenarios.

Facial recognition
- Images, Stored Videos, and Streaming videos can be searches for faces that match those in a face collection.
- A face collection is an index of faces that you own and manage.

Unsafe Content Detection
- Amazon Rekognition can detect explicit and suggestive adult content in images and in videos.
- For example, social and dating sites, photo sharing platforms, blogs and forums, apps for children, e-commerce sites, entertainment and online advertising services.

Celebrity recognition
- Amazon Rekognition can recognize thousands of celebrities (politicians, sports, business, entertainment and media) within supplied images and in videos.

Amazon Rekognition is a service that makes it easy to add powerful, image and video based, visual analysis to your applications.
- Rekognition Image lets you easily build powerful applications to search, verify, and organize millions of images.
- Rekognition Video lets you extract motion-based context from stored or live stream videos and helps you analyze them.

You just provide an image or video to the Rekognition API, and the service can identify objects, people, text, scenes, and activities. It can detect any inappropriate content as well.

Amazon Rekognition also provides highly accurate facial analysis and facial recognition. You can detect, analyze, and compare faces for a wide variety of use cases, including user verification, cataloging, people counting, and public safety.

Amazon Rekoginition is a HIPAA eligible service

You need to ensure that the Amazon S3 bucket you want to use is in the same region as your Amazon Rekognition API endpoint.

Amazon Rekognition provides two API sets, they are:
- They are Amazon Rekognition Image, for analyzing images, and
- Amazon Rekognition Video, for analyzing videos.

Both API sets perform detection and recognition analysis of images and videos.

Amazon Rekognition Video can be used to track the path of people in a stored video.

Amazon Rekognition Video to search a streaming video for persons whose facial descriptions match facial descriptions already stored by Amazon Rekognition.

RecognizeCelebrities API returns information for up to 100 celebrities detected in an image.
- This includes information about where celebrity faces are detected on the image and where to get further information about the celebrity.

AWS Rekognition

```
Image (S3 or bytes) → DetectFaces() → Response in JSON
Stored video (S3) → StartFaceDetection() → Result to SNS topic
Streaming video (Kinesis Video stream) → CreateStreamProcessor() → Kinesis Data stream
```

Amazon Rekognition – Types of Detection and Recognition

The types of detection and recognition that the **Amazon Rekognition Image API** and **Amazon Rekognition Video API** can perform are:

- **Labels**
  - A label refers to any of the following: objects (flower, tree, or table), events (a wedding, graduation, or birthday party), concepts (a landscape, evening, and nature) or activities (getting out of a car).
  - Amazon Rekognition can detect labels in images and videos.
  - Activities are not detected in images.
  - To detect labels in images, use **DetectLabels.**
  - To detect labels in stored videos, use **StartLabelDetection.**

- **Faces**
  - Amazon Rekognition can detect faces in images and stored videos.
  - Information can be obtained about where faces are detected in an image or video, facial landmarks such as the position of eyes, and detected emotions such as happy or sad.
  - It is possible to compare a face in an image with faces detected in another image.
  - Information about faces can also be stored for later retrieval.
  - To detect faces in images, use **DetectFaces**
  - To detect faces in stored videos, use **StartFaceDetection.**

- **Face Search**
  - Amazon Rekognition can search for faces.
  - Facial information is indexed into a container known as a collection.
    - Face information in the collection can then be matched with faces detected in images, stored videos, and streaming video.
  - To search for known faces in images, use **DetectFaces.**
  - To search for known faces in stored videos, use **StartFaceDetection.**
  - To search for known faces in streaming videos, use **CreateStreamProcessor.**

- **People Paths**
  - Amazon Rekognition can track the paths of people detected **in a stored video**.
  - Amazon Rekognition Video provides path tracking, face details, and in-frame location information for people detected in a video.
  - To detect people in stored videos, use **StartPersonTracking.**

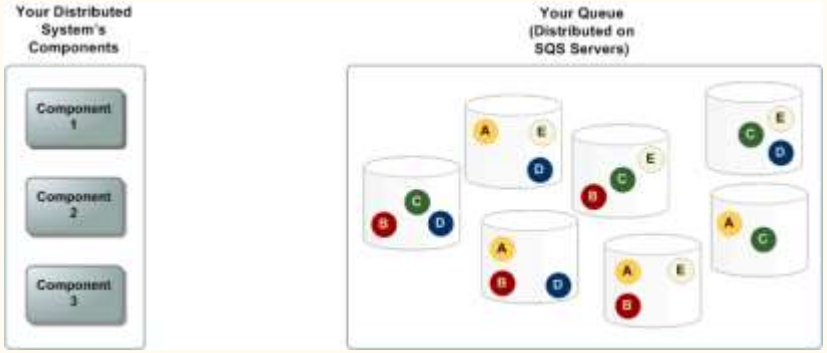
- **Celebrities**
  - Amazon Rekognition can recognize thousands of celebrities **in images and stored videos.**
  - Information can be obtained about where a celebrity's face is located on an image, facial landmarks and the pose of a celebrity's face.
  - Tracking information can also be obtained for celebrities as they appear throughout a stored video.
  - To recognize celebrities in images, **use RecognizeCelebrities.**
  - To recognize celebrities in stored videos, **use StartCelebrityRecognition**

- **Text Detection**
  - Amazon Rekognition Text in Image can detect text in images and convert it into machine-readable text.
  - To detect text in images, use **DetectText**

- **Unsafe Content**
  - Amazon Rekognition can analyze **images and stored videos** for explicit or suggestive adult content.
  - To detect unsafe images, **use DetectModerationLabels**
  - To detect unsafe stored videos, **use StartContentModeration**

Images	**Detect**Labels DetectFaces RecognizeCelebrities DetectText DetectModerationLabels
Stored Video	**Start**LabelDetection StartFaceDetection StartPersonTracking StartCelebrityRecognition StartContentModeration
Streaming Video	**Create**StreamProcessor (face detection)

AWS SQS



Standard queues -

- Order not maintained,
- At least one delivery,
- High throughput

FIFO queues

- Message Order maintained (**within message group and within 5-minute dedup interval**),
- Exact once delivery (**No duplicates, provide msg dedup id**)
- Throughput limited (300 TPS per Api)

```
Messages are ordered based on message group ID. If
multiple hosts (or different threads on the same host)
send messages with the same message group ID to a FIFO
queue, SQS stores the messages in the order in which
they arrive for processing. To ensure that SQS preserves
the order in which messages are sent and received,
ensure that each producer uses a unique message group ID
to send all its messages.
```