Chapter 11
AWS Application Deployment

Episode 11.01
Application and Deployment Services

No Slide

- There are no slides for this episode.

SAGEFOX

Episode 11.02
Lambda

Lambda

- AWS compute service that runs code without servers
- Runs code only when needed
- Scales automatically
  - Up to thousands of requests per second
- Billed by compute time

Automated Management

- Server maintenance
- Operating system maintenance
- Capacity scaling
- Code monitoring
- Logging

Languages Supported

- Node.js
- Java
- C#
- Go
- Python

Lambda Use Process

- Customer builds the code
- Customer launches the code as Lambda function
- AWS selects server
- Customer calls Lambda function as needed from applications

Episode 11.03
API Gateway

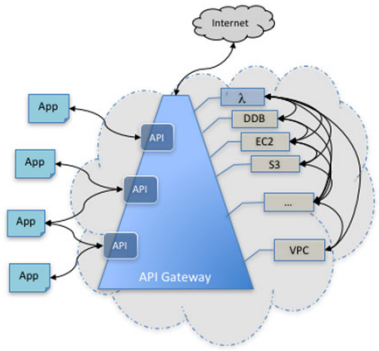
API Gateway

- API management in the cloud
  - Create
  - Publish
  - Maintain
  - Monitor
  - Secure

API Gateway

- APIs can interact with many targets
  - AWS services
  - Other web services
  - Data stored in AWS

API Gateway Architecture

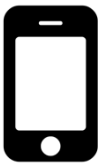


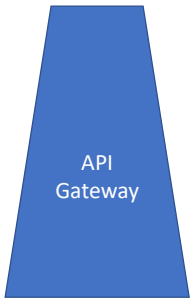
Serverless Architecture

- Moves data in and out of the cloud without instances
- Process functions without instances
- Two primary services
  - Lambda
  - API Gateway

Serverless Architecture

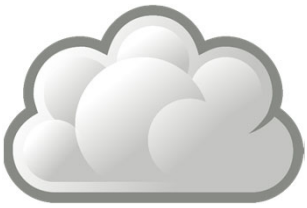
Client talks to API Gateway





Lambda function exposed as API

Lambda function processes input



Lambda function forwards results to another Internet service

Cross Origin Resource Sharing (CORS)

- Can be enabled for the API gateway
- Allows receipt of requests from other domains
  - Default is internal domain requests only
- Without it, errors will occur

Episode 11.04
Kinesis

Kinesis

- Processes streaming data
- Real-time analytics
- Multi-tier enabler
- Very DevOps focused
- Conceptual importance for an architect

Operating Modes

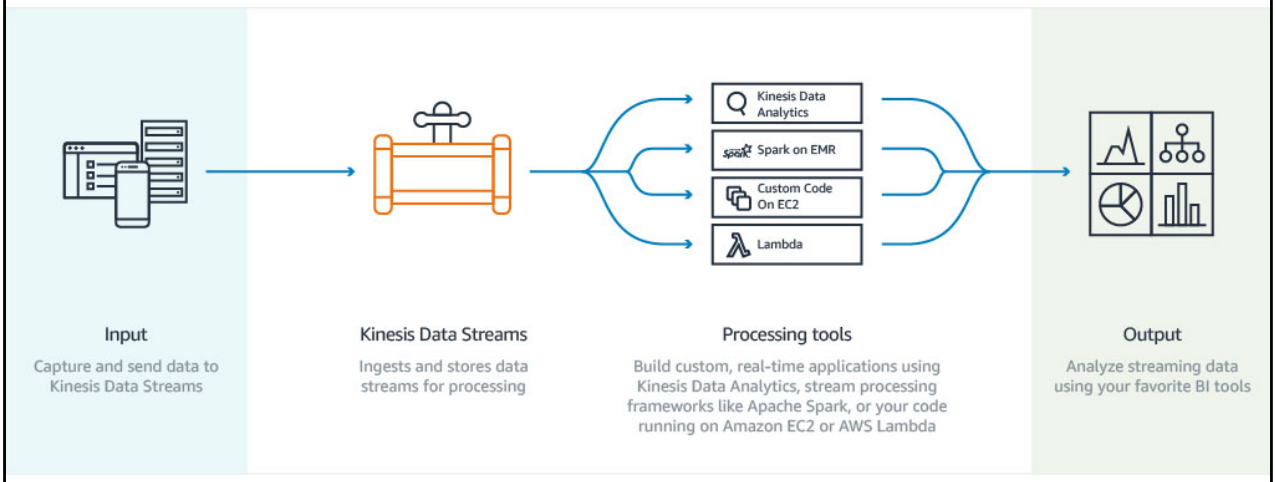
- Kinesis Data Streams
- Kinesis Data Firehose
- Kinesis Data Analytics
- Kinesis Video Streams
  - Media Services

Kinesis Benefits

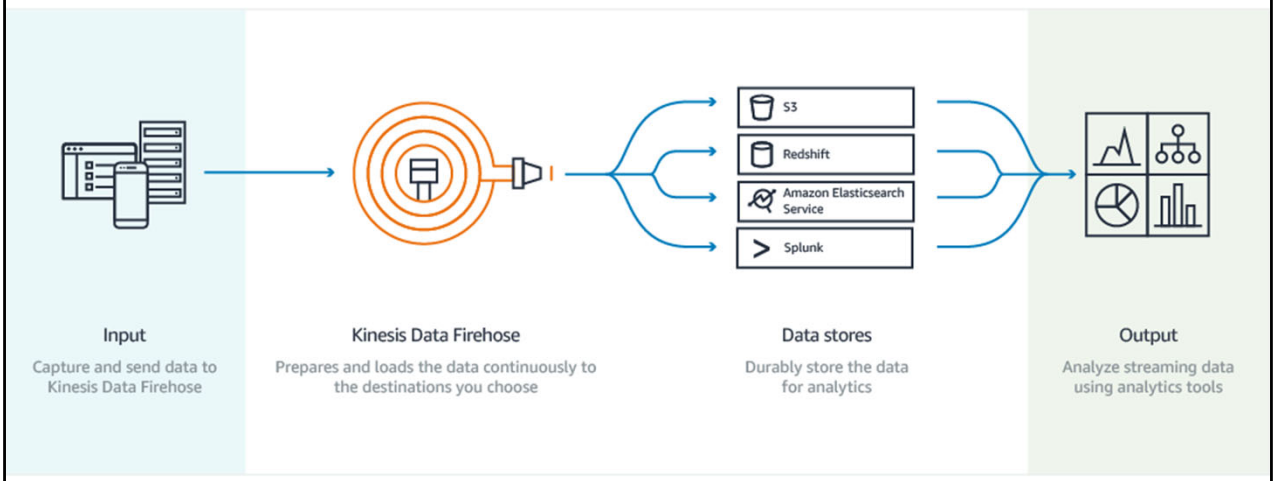
- Architecture fully managed
- No custom coding required
  - Configure producers
  - Configure consumers
  - Focus is on the analytics

Episode 11.05
Kinesis Data Streams and Firehose

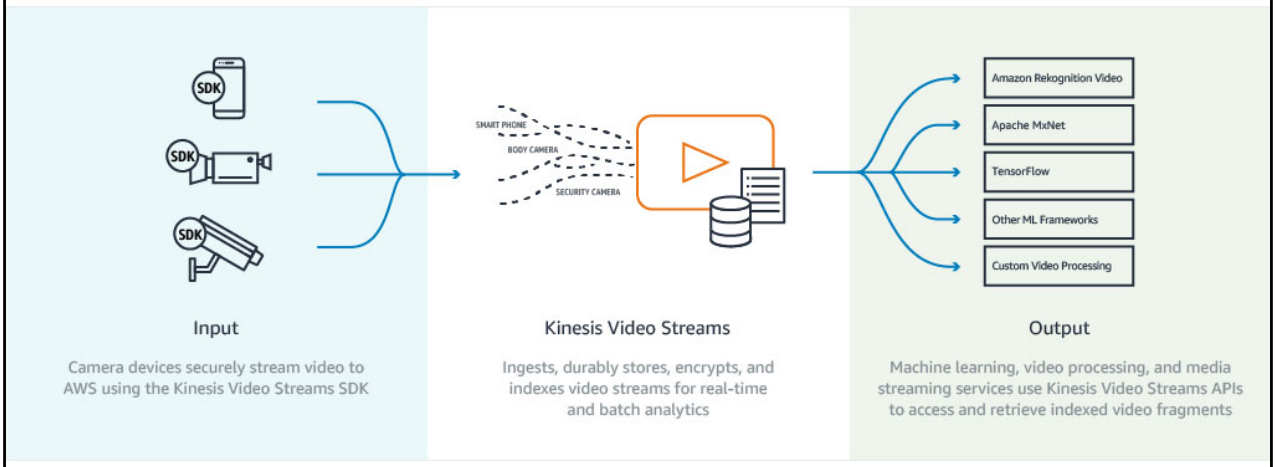
Kinesis Data Streams



Kinesis Data Firehose



Kinesis Video Streams

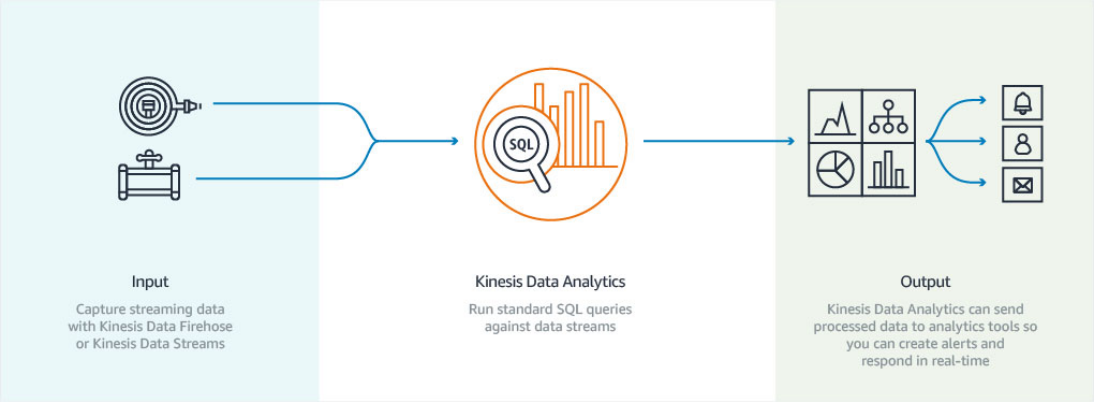


Episode 11.06
Kinesis Data Analytics

Kinesis Data Analytics

- Analyzes real-time data streams
- Based on standard SQL queries
- Supports concurrent consumers
  - Redshift
  - S3
  - Elastisearch
  - Lambda
  - Kinesis Data Streams

Kinesis Data Analytics

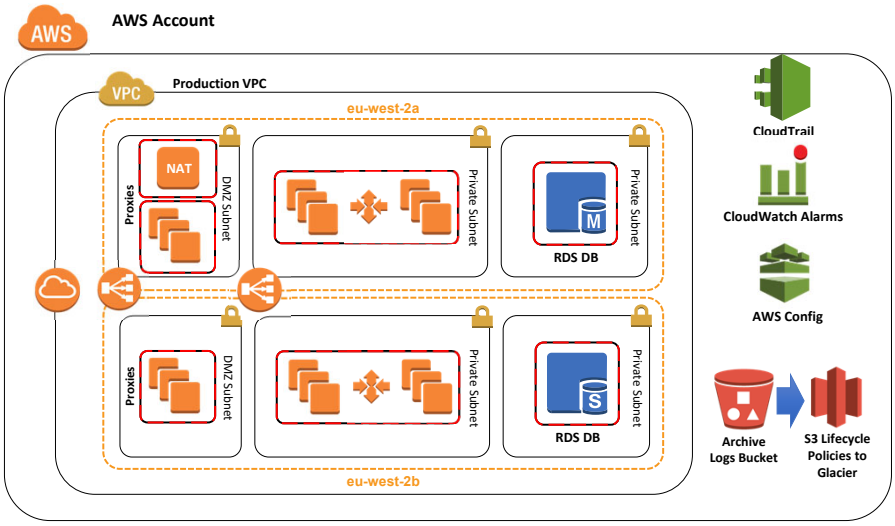


Episode 11.07
Reference Architectures

Reference Architectures

- Well-architected frameworks
- AWS created architecture plans for specific scenarios
  - HIPAA
  - PCI-DSS
  - UK-OFFICIAL

Reference Architecture



https://docs.aws.amazon.com/quickstart/latest/compliance-hipaa/overview.html

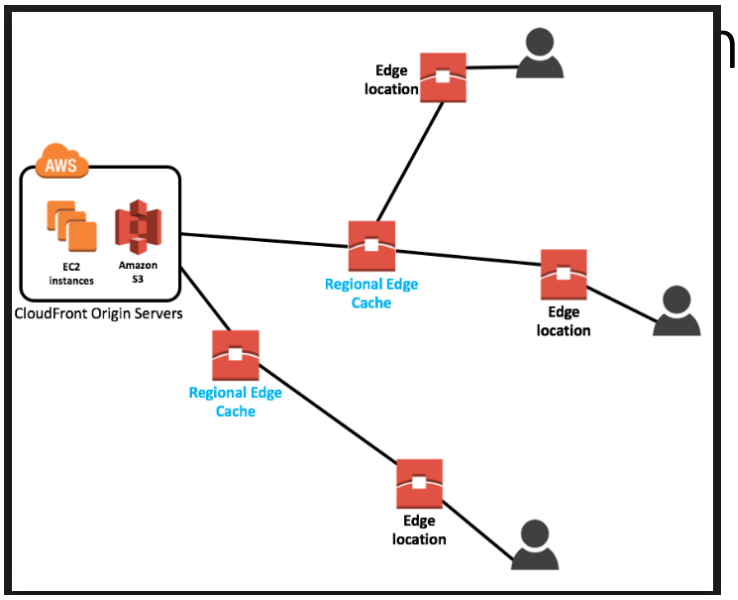
DEMO

- PCI-DSS Architecture
  - https://docs.aws.amazon.com/quickstart/latest/compliance-pci/welcome.html
- All the QuickStart Reference Architectures
  - https://aws.amazon.com/quickstart/

Episode 11.08
CloudFront

CloudFront

- Content delivery network (CDN)
  - Distributes content to localized regions
  - Reduces latency
  - Provides high data transfer speeds



https://cloudacademy.com/blog/aws-global-infrastructure/

Demo

- Use Cases
- https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/IntroductionUseCases.html

Implementation Considerations

- Content source
  - S3
  - MediaPackage channel
    - AWS Elemental MediaPackage is a just-in-time video packaging and origination service that runs in the AWS Cloud. With MediaPackage, you can deliver highly secure, scalable, and reliable video streams to a wide variety of playback devices and content delivery networks (CDNs).
  - HTTP server
- Content access
  - Public
  - Restricted
- Content constraints
  - HTTPS required
  - Geo-restrictions

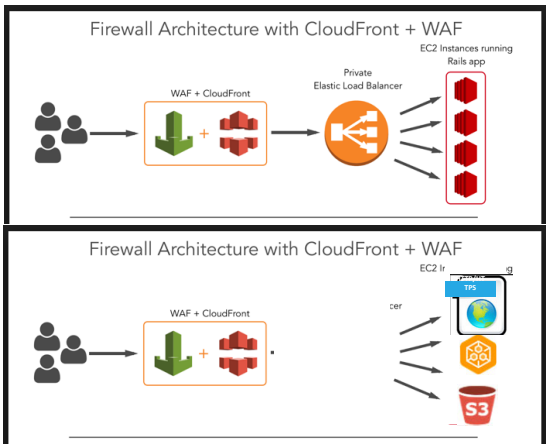
Episode 11.09
Web Application Firewall (WAF)

Web Application Firewall (WAF)

- Controls access to HTTP and HTTPS servers
  - Based on requests
  - Based on source IPs

Web Application Firewall (WAF)

- Works with CloudFront and/or Load Balancers



WAF Behaviors

- Allow all requests
  - Except the ones that you specify
- Block all requests
  - Except the ones that you specify
- Monitoring
  - Requests that match specified parameters

WAF Operations

- Error handling
  - HTTP 403 error (forbidden)
- Configurable default behavior
  - What happens when the request doesn't match any rules?
    - Allow
    - Deny

Episode 11.10
Simple Queue Service (SQS)

Simple Queue Service (SQS)

- Used to decouple applications
  - Break application into separate processing tasks
  - Allows many small processes to form a complete solution

SQS Messages

- Outputs from other processes
- Inputs to other processes
- Queued and processed asynchronously
  - Non-linear
- Up to 256 KB data
  - Record pointers, directives, parameters

SQS Participants

- Message producers
- Message consumers
- Messaging service
  - SQS

SQS Features

- Redundant across multiple AZs
  - Queued until processed
  - Retention up to 14 days
- Automatically scales

SQS Queue Types

- Standard
  - Default queue type
  - Doesn't guarantee sequential delivery of messages
- First-In-First-Out (FIFO)
  - Guarantees sequential delivery of messages
  - Supports fewer transactions per second

Episode 11.11
Simple Notification Service (SNS)

Simple Notification Service (SNS)

- Paging in the cloud
- Uses the publish-subscribe mechanism based on "topics"
  - Called pub-sub messaging

Publishers

- Publishers push messages to topics
  - Topic examples:
    - Admin alerts, performance alerts, etc.
  - Publisher examples:
    - CloudWatch, Cost Explorer, etc

Subscribers

- Clients receiving notifications
- Receive all messages broadcasted to the topic
- Publishers and subscribers aren't "aware" of each other

SNS Features

- Stored across multiple AZs
- Several delivery options supported
  - HTTP/HTTPS
  - E-Mail
  - SMS (Short Message Service)
  - Lambda
  - SQS

SNS Message Limits

- Up to 256 KB of data
- Special SMS constraints:
  - Max size of single SMS is 140 bytes
    - Larger messages sent as multiple transmissions
    - Aggregate SMS size is 1600 bytes

Episode 11.12
Simple Workflow (SWF)

Simple Workflow (SWF)

- Defines the sequence of events required to achieve a workflow
- Used in decoupled applications

Workflow

- Activities that result in a desired objective
- Logic that controls the activities
  - Decider function determines best workflow
- Operates in a domain
  - Created logical boundary in SWF to constrain the scope of the activities

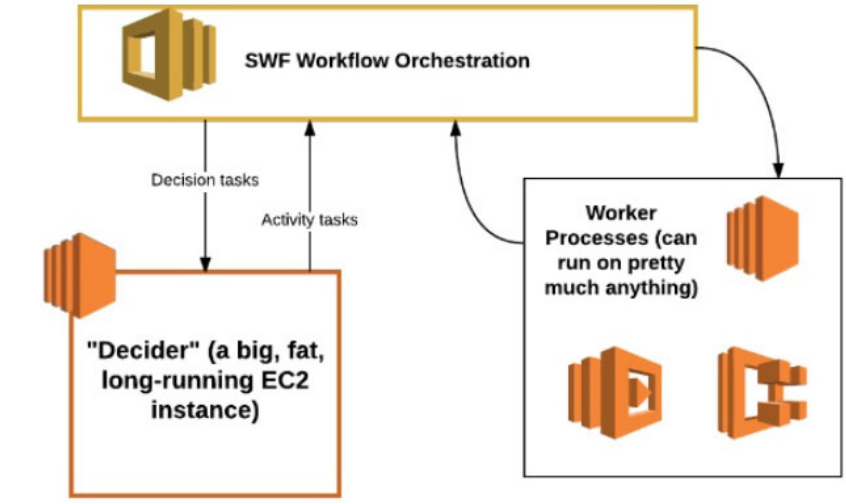
SWF Activity Tasks

- One invocation of an activity
  - For example, processing an order
- May be invoked multiple times
  - For example, processing a multi-item order

SWF Activity Workers

- The applications that receives and processes activity tasks

Traditional SWF Architecture



https://www.slideshare.net/ForrestBrazeal/serverless-workflows-on-aws-a-journey-from-swf-to-step-functions

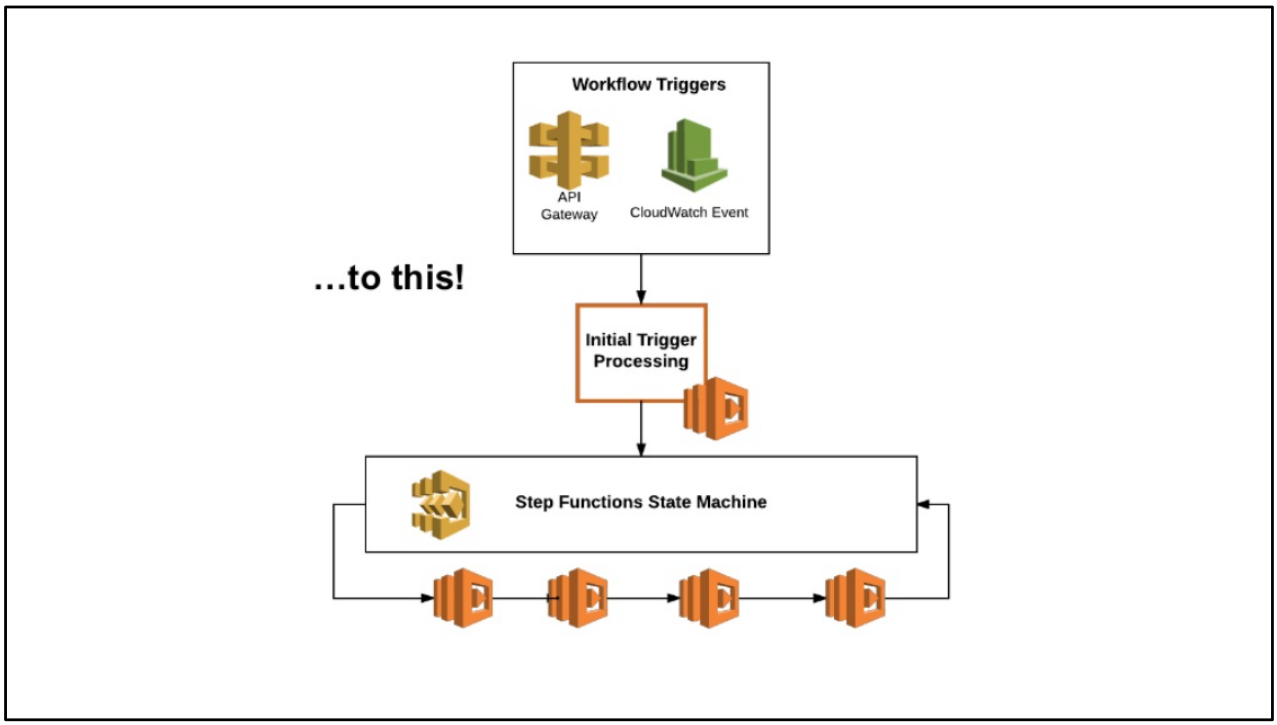
Episode 11.13
Step Functions

Step Functions

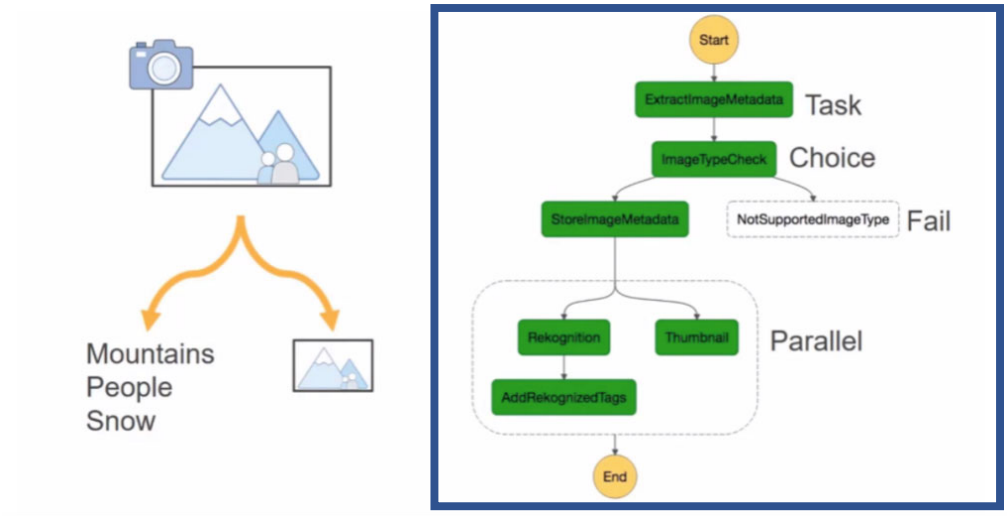
- AWS recommended practice
  - Eventually replacing SWF
- Similar to SWF in functionality
  - Uses state machines
    - Decider
    - Activity tasks
    - Worker tasks

Step Function Concepts

- Task
  - Single unit of work
- Choice
  - Provides branching logic
- Parallel
  - Process input through multiple tasks and combine the output


Workflow Triggers
API Gateway
CloudWatch Event
...to this!
Initial Trigger Processing
Step Functions State Machine

Step Functions State Machine



Episode 11.14
OpsWorks

OpsWorks

- Configuration management service
  - Configure (code-based)
    - Instance deployment
    - Service deployment
    - Application deployment
  - Operate
    - Application updates
    - Infrastructure updates
- Automated deployment

OpsWorks Stacks

- Initial OpsWorks mode
- Collection of layers
  - Any AWS service
  - Any runtime environment

OpsWorks Chef Automate

- Cookbooks contain recipes
- Recipes equivalent to layers
  - Defined configuration settings
    - Admin defined
    - AWS defined
    - Third-party defined

OpsWorks Puppet

- Master servers
  - Pre-configured modules
  - Modules equivalent to layers

Prebuilt Layers

- Ruby
- PHP
- Node.js
- Java
- Amazon RDS
- HA Proxy
- MySQL

Use Cases

- In the cloud:
  - Chef
  - Puppet
- On-premises (local):
  - Stacks

Episode 11.15
Cognito

Cognito

- User identity and data synchronization service
  - SSO
- Public identity providers
  - Google
  - Facebook
  - Amazon
- Private identity providers
  - Active Directory with SAML

Identity Management

- Based on open standards
  - OAuth 2.0
  - SAML 2.0
  - OpenID connect
- Profile management
- Scales to millions of users

AWS Integration

- Cognito controls access to AWS resources
  - Define roles
  - Map users to roles

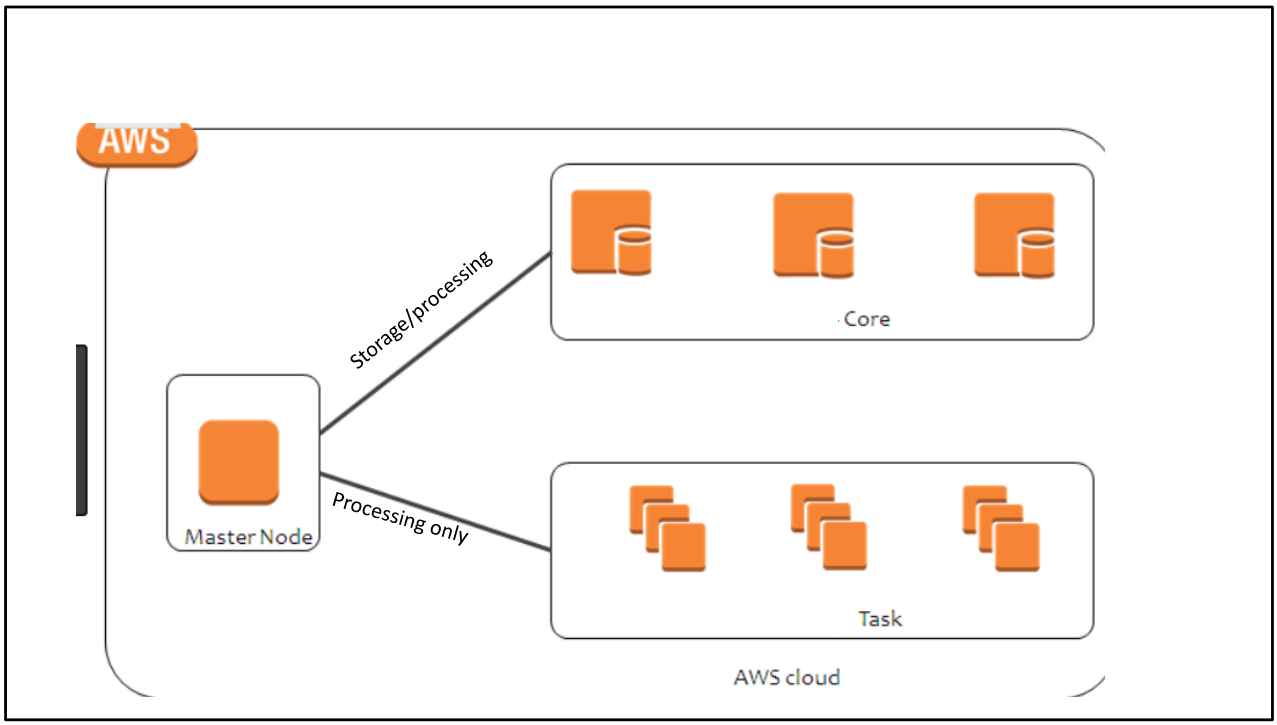
Episode 11.16
Elastic MapReduce (EMR)

Elastic MapReduce (EMR)

- Distributes processing across clusters
  - Implements a managed Hadoop framework
- Pulls data from S3 buckets
- Uses EC2 instances
- User defines the number of needed clusters

EMR Cluster Nodes

- Master
  - Coordinates job distribution across core and task nodes
- Core
  - Runs tasks assigned by the master node
  - Stores data in the cluster
- Task
  - Runs only tasks that do not store data


AWS
Storage/processing
Core
Master Node
Processing only
Task
AWS cloud

Episode 11.17
CloudFormation

CloudFormation

- Graphical interface
- Uses templates to build entire solutions in AWS
  - EC2 instances
  - Security groups
  - Subnets
  - IAM users and roles
  - Etc.

CloudFormation

- Components:
  - Templates
    - Build the stacks
  - Stacks
    - Implementation of templates
  - Change sets
    - Modify stacks

CloudFormation

- Why use it?
  - Rapid deployment
  - Mirror existing internal architectures
  - Take advantage of templates created by others

DEMO

- CloudFormation


Episode 11.18
CloudFormation Properties
SAGEFOX

No Slide

- There are no slides for this episode.

SAGEFOX

Episode 11.19
CloudWatch

CloudWatch

- Monitors the cloud and on-premises systems
- Dashboards
- Logs
- Events
- Alarms

CloudWatch

- Why use it?
  - Monitor critical systems
  - Receive notifications related to performance and security
  - Push on-premises logs into the cloud
  - Take automatic actions based on alarms

Episode 11.20
Trusted Advisor

Trusted Advisor

- Recommendations for security and other issues
- Scans the AWS cloud for recommendations

DEMO

- The Trusted Advisor interface

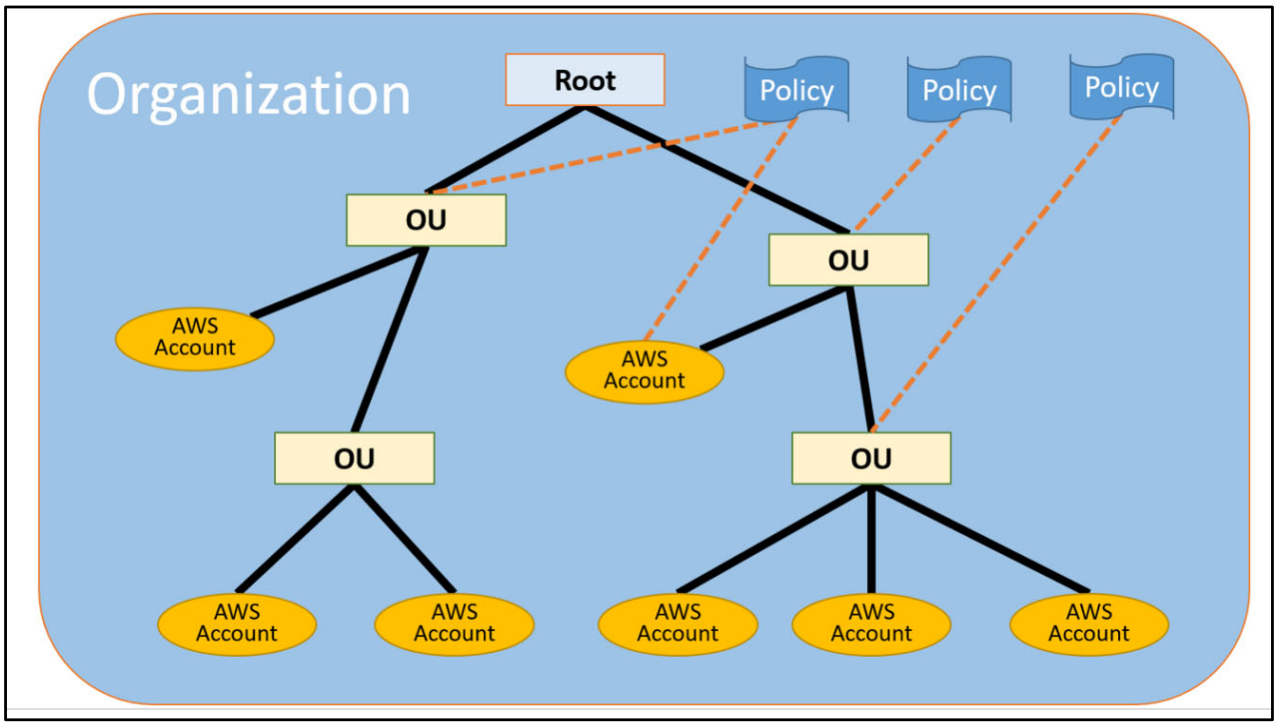
Episode 11.21
Organizations

Organizations

- Collection of AWS accounts
- Centralized
  - One management interface
  - Billing
  - Account management
- No additional charge for use

Organizational Units (OUs)

- Hierarchical account management
- Nest OUs up to five levels deep
- Policies attached for permissions


Organization
Root
Policy
Policy
Policy
OU
OU
AWS Account
AWS Account
OU
OU
AWS Account
AWS Account
AWS Account
AWS Account
AWS Account