AWS Cloud practinioner notes

1. Cloud computing

Different types of cloud computing

2. IAM - Identity and Access Management

Users: mapped to a physical person, has a username and password for the AWS console Groups: collection of users under one set of permissions Policies: JSON documents that define permissions Roles: create roles and assign them to AWS resources Security: MFA + password policy AWS CLI: manage your AWS services using the command line AWS SDK: manage your AWS services using a programming language Access keys: access your AWS account using the CLI or SDK Audit: IAM Credential Report & IAM Access Advisor

3. EC2 - Elastic Compute Cloud

EC2 Instances

- EC2 - Virtual servers in the cloud
- EC2 instances are virtual servers that can be resized and scaled up or down
- EC2 composed of AMI (OS), Instance size (CPU + RAM), Storage (EBS or Instance Store), Security Groups, EC2 User Data (bootstrap scripts that is started when the instance is launched)
- Security Groups: act as a virtual firewall for your instance to control inbound and outbound traffic
- EC2 User Data: bootstrap scripts that are run when the instance is launched
- SSH: start a terminal session on a Linux instance (port 22)
- EC2 Instance Role: link to IAM roles
- Purchasing Options:
    - On-demand: pay as you go
    - Spot: bid for unused capacity
    - Reserved Standard: 1-3 years, up to 75% discount
    - Reserved Convertible: 1-3 years, up to 54% discount, can change the EC2 instance type
    - Dedicated Host: physical server dedicated to your use
    - Dedicated Instance: instance running on a physical server dedicated to your use

4. Storage for EC2 Instance

EBS

- EBS - Elastic Block Store, block storage volumes that you can attach to your EC2 instances
- It allows your instance to persist data even after the instance is terminated
- They can only be mounted at one instance at a time
- Bound to a specific Availabilty Zone
- Types of EBS:
    - General Purpose SSD (gp2): balances price and performance
    - Provisioned IOPS SSD (io1): high-performance SSD volume for mission-critical low-latency or high-throughput workloads

    - Throughput Optimized HDD (st1): low-cost HDD volume designed for frequently accessed, throughput-intensive workloads
    - Cold HDD (sc1): low-cost HDD volume designed for less frequently accessed workloads
    - Magnetic (standard): previous generation HDD volume, low cost, low performance
- Analogy: think of them as a network USB stick
- exam keywords: block storage, persistent, EBS volumes

EFS

- EFS - Elastic File System, scalable file storage for use with Amazon EC2
- It's a network drive (i.e. not a physical drive):
    - it uses the network to communicate with the instance -> latency
    - it can be detached from an instance and attached to another one quickly
- It's locked to an AZ, but can be mounted on multiple instances in the same AZ
- To move to another AZ, you need to create a snapshot of the EFS and copy it to another AZ
- EFS have a provisioned capacity (size in GBs and IOPS)
- You get billed for the provisioned capacity
- Delete on termination: you can choose to keep the EBS volume or delete it when the EC2 instance is terminated (by default: the root EBS volume is deleted, but additional volumes are kept)
- exam keywords: network drive, scalable, multiple instances

Amazon FSx for Windows File Server

- Amazon FSx for Windows File Server - Fully managed Windows file system
- Windows File Server fully managed by AWS
- Supports the SMB protocol & Windows NTFS
- Integrates with Microsoft Active Directory
- can be accessed from AWS or on-premises

Amazon FSx for Lustre

- Amazon FSx for Lustre - Fully managed file system optimized for compute-intensive workloads
- Used for ML, video processing, financial modeling, etc.
- High-performance file system: 100GBs/s of throughput, sub-millisecond latencies
- exam keywords: compute-intensive workloads, high-performance file system

EC2 Instance Store

- EC2 Instance Store - Temporary block storage for your EC2 instance
- Instance store volumes are temporary and will be deleted if the instance is stopped or terminated
- Better I/O performance than EBS volumes
- Use cases: buffer, cache, scratch data, temporary content
- exam keywords: temporary, high I/O performance

EC2 Image Builder

- EC2 Image Builder - Fully managed service that makes it easier to automate the creation, management, and deployment of customized, secure, and up-to-date server images that are pre-installed and pre-configured with software and settings

- exam keywords: automate, creation, management, deployment, server images

5. ELB & ASG - Elastic Load Balancing & Auto Scaling Groups

High Availability, Scalability, elasticity, and agility in the cloud

- High Availability: ensure your application remains operational during component failure
- Scalability: ability to handle increased load by adding resources
- Elasticity: automatically add or remove resources based on load
- Agility in the cloud: quickly provision resources as needed

Elastic Load Balancer

- ELB - Distribute incoming application or network traffic across multiple targets, such as EC2 instances, containers, and IP addresses
- supports health checks
- 4 types:
    - application load balancer: HTTP/HTTPS, Layer 7, static DNS
    - network load balancer: ultra-high performance TCP, TLS, UDP, Layer 4, static IP
    - gateway load balancer: route traffic to firewalls that you manage on EC2 instances, Layer 3
    - classic load balancer: retired in 2023, layer 4 and 7 (replaced by ALB and NLB)
- exam keywords: distribute, incoming traffic, health checks

Auto Scaling Groups

- ASG - Scale EC2 instances based on demand
- can scale out (add instances) or scale in (remove instances) based on the demand on your system or replace unhealthy instances
- integrated with the ELB
- scaling strategies:
    - manual: you manually add or remove instances
    - dynamic: based on demand:
        - target tracking scaling: target a specific average utilization: eg. CPU at 70%
        - simple / step scaling: scale based on CloudWatch alarms
        - sheduled actions: scale based on a schedule
        - predictive scaling: use machine learning to predict future usage
- exam keywords: scale, demand

6. Amazon S3

S3 - Simple Storage Service

- S3 - Object storage service that offers industry-leading scalability, data availability, security, and performance
- S3 is object-based storage, not block-based storage
- S3 is a universal namespace, i.e. each bucket name must be unique globally
- S3 is tied to a region, and the data remains in the region unless you transfer it

- S3 security:
    - User-based: IAM policies
    - Resource-based: bucket policies, ACLs
    - Encryption: in transit (SSL/TLS), at rest (SSE-S3, SSE-KMS, SSE-C)
    - S3 websites: host a static website using S3
    - S3 versioning: multiple versions for files, prevent accidental deletes
    - S3 Replication: CRR (Cross-Region Replication), SRR (Same-Region Replication), must enable versioning
    - S3 Lifecycle Rules: automate moving objects between storage classes or deleting them
    - S3 Storage Classes:
        - Standard: general-purpose storage of frequently accessed data eg. for content delivery, big data analytics, mobile and gaming applications
        - IA (Infrequent Access): for data that is accessed less frequently, but requires rapid access when needed. Lower fee than S3, but you are charged a retrieval fee. Suitable for long-term storage, backups, and disaster recovery
        - IZ-IA (One Zone-Infrequent Access): for data that is accessed less frequently, but requires rapid access when needed. Lower fee than IA, but data is stored in a single AZ
        - Glacier: low-cost storage class for data archiving. Retrieval times configurable from minutes to hours
            - Instant retrieval: miliseconds retrieval, minimum storage duration: 90 days
            - Flexible retrieval:
                - expedited retrieval: 1-5 minutes
                - standard retrieval: 3-5 hours
                - Bulk retrieval: 5-12 hours
                - minimum storage duration: 90 days
            - Deep archive:
                - standard retrieval: 12 hours
                - bulk retrieval: 48 hours
                - minimum storage duration: 180 days
        - Intelligent Tiering: designed to optimize costs by automatically moving data to the most cost-effective access tier, without performance impact and small monthly monitoring and automation fee:
            - frequent access: default
            - infrequent access: data not accessed for 30 days
            - archive instant access: data not accessed for 90 days
            - archive access: configurable from 90 days to 700+ days
            - deep archive access: configurable from 180 days to 700+ days

Snow Family

- Snow Family - Physical devices to transfer data to and from AWS
- Snowcone: small, portable, rugged device for edge computing
- Snowball: large, rugged device for large data transfers
- Snowmobile: exabyte-scale data transfer service
- exam keywords: physical devices, edge computing, large data transfers

Edge computing

- Edge computing - Process data closer to the source of data generation
- reduce latency, save bandwidth, process data closer to the source
- Snowball Edge: compute optimized or storage optimized devices with EC2 instances
- Snowcone edge: computing optimized device for edge computing
- exam keywords: reduce latency, save bandwidth, process data closer to the source

Storage Gateway

- Storage Gateway - Hybrid cloud storage service that gives you on-premises access to virtually unlimited cloud storage

7. Databases & Analytics

RDS (Relational Database Service)

- RDS - Managed relational database service that supports multiple database engines: MySQL, PostgreSQL, MariaDB, Oracle, SQL Server, and Aurora
- suited for OLTP workloads: Online Transaction Processing

RDS Deployments options

1. Read replicas

- scale the read workload of your DB (in the same AZ)
- can create oup to 15 read replicas
- data is only written to the main DB

2. Multi-AZ

- failover DB in cas of AZ outage (high availability)
- data is only read & written to the main database
- can only have 1 other AZ as failover, failover is used only when main is down

3. Multi-Region (Read replicas)

- Disater recovery in cas of region issue
- local performance for global reads (application reads in db replica of nearest AZ)
- write is done in main DB (1 AZ)
- replication cost

Amazon Elasticache

ElastiCache is to get managed Redis or Memcached in-memory databases with high performance and low latency. --> Helps to reduce the load of databases (RDS DB: Postgres, or other) for read intensive workloads by storing cache in in-memory database

insert schema architecture

DynamoDB

keywords: serverless, low latency

- Fully managed highly available with replication across 3 AZ
- NoSQL database -> not a relational database -> stores primary key and product pairs (cfr key value)
- scales to massive worklaods because it is a distributes serverless database (no instances of the db are required)
- millions of requests per sections, trillions of rows, 100s of TB of storage
- single digit millisecond latency - low latency retrieval

DynamoDB is a flagship product of AWS

DynamoDB Global Tables

DynamoDB Global Tables makes DynamoDB table accessible with low latency in multiple-regions by using Active-Active replciation 2-way replication --> users can read and write in any table specific to the AZ

insert schema architecture

DynamoDB Accelerator - DAX

- same as Elasticache but for `DynamoDB` only: in-memory cache for DynamoDB only
- 10x performance improvement

DAX

- Amazon DynamoDB Accelerator (DAX) is a fully managed, highly available, in-memory cache for Amazon DynamoDB that delivers up to a 10 times performance improvement—from milliseconds to microseconds—even at millions of requests per second.

Redshift

- Redshift - Fully managed, petabyte-scale data warehouse service and analytics tool.
- data is stored in colnmar format (columns instead of rows)
- exam keywords: analytics, data warehouse, colunmar, OLAP
- OLAP (Online Analytical Processing) is used for complex queries and data analysis

Redshift serverless

- Redshift serverless - Serverless data warehouse that automatically scales the datawarehouse based on the workload.
- pay only for compute and storage used during analysis -> very cost efficient to run Redshift
- use cases: reporting, dashboards, real-time analytics

Amazon EMR (Elastic MapReduce)

- Amazon EMR - helps creating Hadoop clusters to analyze and process large amounts of data (Big Data)

- cluster can be made up of EC2 instances and EMR takes care of the provisioning and configuration of the cluster
- automatically scales the cluster based on the workload
- supports Apache Spark, HBasem Presto, Flink, and other big data frameworks
- exam keywords: Hadoop clusters, Big Data, data processing

Amazon Athena

- Amazon Athena - serverless query service that makes it easy to analyze data in Amazon S3 using standard SQL
- S3 objects supported are CSV, JSON, ORC, Avro, and Parquet (built on Presto)
- no need to set up or manage any infrastructure
- use cases: Business intelligence, analytics, reportingm ELB Logs, CloudTrail Logs, VPC Flow Logs
- Amazon QuickSight can be used to visualize the data
- pricing: $5 per TB of data scanned
- cost saving: use compressed or columnar data formats(Parquet, ORC) to reduce the amount of data scanned

exam keywords: analyze data in S3, serverless, SQL

Amazon QuickSight

- Amazon QuickSight - Serverless machine learning-powered Business Intelligence tool that allows you to create interactive dashboards and reports
- integrates with AWS services like RDS, Redshift, Athena, Aurora and S3
- pay per session or per user
- exam keywords: Business Intelligence, dashboards, reports, machine learning

Neptune

- Neptune - Fully managed graph database service that supports both RDF and property graph models
- use cases: social networking, fraud detection, recommendation engines, network security, knowledge graphs (Wikipedia)
- exam keywords: graph database, RDF, property graph
- build and rund applications working with highly connected datasets - optimized for these complex and hard querries
- highly available with replication across 3 AZs

TimeStream

- TimeStream - Fully managed, serverless, time-series database service for IoT and operational applications
- optimized for IoT and operational applications
- store and analyze trillions of events per day at 1/10th the cost of relational databases
- exam keywords: time-series data, IoT, operational applications

QLDB

- QLDB - Fully managed ledger database that provides a transparent, immutable, and cryptographically verifiable transaction log
- Quantun Ledger Database
- A ledger is a record of (financial) transactions
- immutable system: no entry can be removed or modified
- cryptographically verifiable: you can verify the integrity of the data -> because of the cryptographic hash of the data after each transaction
- difference with Amazon Managed Blockchain: QLDB is a ledger database, not a blockchain and it is not decentralized (all data is stored at AWS) -> in accordance with financial regulations
- manipulate data using PartiQL (SQL-compatible query language)
- 2-3x faster than common ledger blockchain frameworks
- exam keywords: ledger database, financial transactions, (immutable, cryptographically verifiable)

Managed Blockchain

- Managed Blockchain - Create and manage scalable blockchain networks using popular open-source frameworks: Hyperledger Fabric and Ethereum
- join public blockhain networks or create your own private blockchain networks
- exam keywords: blockchain, Hyperledger Fabric, Ethereum

Glue

- Glue - Fully managed extract, transform, and load (ETL) service that makes it easy to prepare and load your data for analytics
- fully serverless
- exam keywords: ETL, extract, transform, load
- use case example: EXTRACT data from S3 and RDS -> TRANSFORM data with Glue script -> LOAD data into Redshift

Glue Data Catalog

- Glue Data Catalog - Fully managed metadata repository that makes it easy to discover, search, and query metadata across your data lake and data warehouse
- catalog of datasets in AWS structure
- can be used with Athena, Redshift, EMR, and other services to build schemas for data
- probably not on exam

DMS (Database Migration Service)

- DMS - Migrate your databases to AWS with minimal downtime
- supports homogeneous migrations (Oracle to Oracle) and heterogeneous migrations (Microsoft SQL to Aurora)
- Quickly and securly migrate databases to AWS
- Source database remains available during the migration
- exam keywords: database migration

Overview

- Relational Databases: RDS and Aurora

    - Difference between multi-AZ, read replicas, multi-region
- In-memory Database: ElastiCache
- Key/Value Database: DynamoDB (serverless) & DAX (cache for DynamoDB) if cache is needed
- Data Warehouse or OLAP: Redshift(SQL)
- Hadoop Cluster: EMR
- Athena: query data on Amazon S3 (serverless & SQL)
- Quicksight: dashboards and visualization of data (serverless)
- DocumentDB: "Aurora of MongoDB" (JSON type of datasets - NoSQL database)
- Amazon QLDB: Financial transactions ledger (immutable journal, cryptographically verifiable): centralized!
- Amazon Managed Blockchain: managed Hyperledger Fabric and Ethereum blockchains
- Glue: Managed ETL (Extract, Transform, Load) and Data Catalog service
- DMS: Database Migration Service
- Neptune: Graph database
- TimeStream: Time-series database

8. Other compute services: ECS, lambda, Batch and Lightsail

ECS (Elastic Container Service)

- ECS - Highly scalable, high-performance container orchestration service that supports Docker containers
- ECS is a container management service that makes it easy to run, stop, and manage Docker containers on a cluster
- You must provision the infrastructure (EC2 instances) that run the containers

Fargate

- Fargate - Serverless compute engine for containers that works with both ECS and EKS
- With Fargate, you no longer have to provision, configure, or scale clusters of virtual machines to run containers --> fully serverless

ECR (Elastic Container Registry)

- ECR - Fully managed Docker container registry that makes it easy to store, manage, and deploy Docker container images
- Private Docker container registry on AWS

Lambda

- Lambda - Serverless compute service that lets you run code without provisioning or managing servers
- Lambda runs your code only when needed and scales automatically
- Functions are triggered by events -> event driven : S3 upload, DynamoDB update, API Gateway request
- Pioneer of serverless computing on AWS
- Supports multiple programming languages: Node.js, Python, Ruby, Java, Go, .NET
- Pricing:

    - pay per request (number of invocations)
    - compute time (GB-seconds) = memory in GB provisioned * total runtime in seconds
- Easy monitoring with CloudWatch Logs
- use cases:
    - create thumbnails from images uploaded to S3
    - run serverless CRON jobs
- exam keywords: serverless, event-driven, functions

API Gateway

- API Gateway - Fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs at any scale
- Create serverless RESTful APIs and WebSocket APIs
- Integrates with Lambda, DynamoDB, and other AWS services
- Proxies requests from endpoint to other services. Eg: Allow users to upload files to S3 or access data in DynamoDB or trigger Lambda functions from HTTP requests
- exam keywords: serverless RESTful APIs, WebSocket APIs

Batch

- Batch - Fully managed batch processing at any scale
- Is not serverless, relies on EC2 instances
- Batch computing is the processing of a large amount of data in a programmatic way

Batch vs. Lambda

- Lambda is serverless, Batch is not
- Lambda is event-driven, Batch is not
- Lambda has limited runtime (15 minutes), Batch can run for hours or days
- Lambda has limited storage (512MB), Batch can use EBS volumes
- Lambda is for small, short-lived functions, Batch is for long-running batch jobs

Lightsail

- Lightsail - Virtual private server (VPS) service that offers everything needed to build an application or website
- For people with no cloud experience
- low and predictable pricing
- use cases:
    - simple websites: Wordpress, Joomla, Drupal
    - Dev/Test environments
    - simple webapps: has templates for MEAN, LAMP, Nginx, Node.js

Overview

- ECS: Run Docker containers on EC2 instances (not serverless)
- Fargate: Serverless compute engine for containers
- ECR: Docker container registry for storing, managing, and deploying Docker container images
- Lambda: Serverless compute service

- API Gateway: Serverless RESTful APIs
- Batch: Run batch jobs on AWS using EC2 instances (not serverless)
- Lightsail: Virtual private server (VPS) service

9. Deployments

CloudFormation

- CloudFormation - Infrastructure as Code (IaC) service that helps you model and set up your AWS resources so you can spend less time managing those resources and more time focusing on your applications
- CloudFormation templates are written in YAML or JSON
- repeat architecture in different environments, regions or AWS accounts
- exam keywords: Infrastructure as Code, templates, YAML, JSON,

CDK (Cloud Development Kit)

- CDK - Software development framework for defining cloud infrastructure in code and provisioning it through AWS CloudFormation
- Define cloud infrastructre through code (Python, TypeScript, Java, C#)
- Code is then synthesized into CloudFormation templates
- Deploy infrastructure and application runtime together
- Great of Lambda functions or Docker containers in ECS / EKS
- exam keywords: Infrastructure as Code, CloudFormation, Python, TypeScript, Java, C#

Elastic Beanstalk

- Elastic Beanstalk - Platform as a Service (PaaS) that allows you to deploy and manage web applications and services
- Supports multiple programming languages: Java, .NET, PHP, Node.js, Python, Ruby, Go, Docker
- Platform as a Service (PaaS)
- You upload your code and Elastic Beanstalk automatically handles the deployment using the architecture model you choose
- Health monitoring is also built-in (health agent pushes metrics to CloudWatch)
- 3 architecture models:
    - Single instance : good for dev environments
    - ELB + ALB : great for web applications
    - ALB only : good for workers, microservices
- exam keywords: PaaS, multiple programming languages

CodeDeploy

- CodeDeploy - Automates code deployments to any instance, including Amazon EC2 instances and instances running on-premises
- exam keywords: code deployments, hybrid(On-premises and AWS)

Systems Manager (SSM)

- Systems Manager - Gives you visibility and control of your infrastructure on AWS

- Patch, configure and run commands at scale (at multiple instances at the same time)
- exam keywords: patching, configuration, automation

SSM Session Manager

- SSM Session Manager - Provides a secure and auditable instance management
- No need for SSH keys, bastion hosts, or open inbound ports
- exam keywords: secure, auditable, instance management

SSM Parameter Store

- SSM Parameter Store - Securely store configuration data and secrets
- exam keywords: secure, configuration, secrets

Developer services

CodeCommit

- CodeCommit - Fully managed source control service that makes it easy for teams to host secure and highly scalable private Git repositories
- exam keywords: source control, Git
- No longer supported in the future

CodeBuild

- CodeBuild - Fully managed build service that compiles source code, runs tests, and produces software packages that are ready to deploy
- exam keywords: build service, compile, tests

CodePipeline

- CodePipeline - Continuous integration and continuous delivery (CI/CD) service for fast and reliable application and infrastructure updates
- from code commit to build to deployment
- exam keywords: CI/CD, continuous integration, continuous delivery

10 Global appplications

Route 53

- Route 53 - Scalable Domain Name System (DNS) web service
- routing policies:
    - simple : one record with multiple IP addresses (no health checks)
    - failover : primary and secondary record (health checks)
    - latency-based : direct users to the region with the lowest latency (health checks)
    - weighted : split traffic based on pre-defined weights (health checks)
- exam keywords: DNS, latency-based routing, failover routing

CloudFront

- CloudFront - Content Delivery Network (CDN) service that securely delivers data, videos, applications, and APIs to customers globally with low latency and high transfer speeds
- Replicate part of your application or S3 bucket content to edge locations - decrease latency
- Cache content of frequent requests at edge locations - improved user experience by decreasing latency
- exam keywords: CDN, edge locations, low latency

Cloudfront vs S3 Cross Region Replication

- CloudFront:
  - caches content at edge locations to decrease latency
  - great for static content that needs to be available everywhere
- S3 Cross Region Replication:
  - replicates all data to another region for disaster recovery
  - great for dynamic content that needs to be available in another region

##3 S3 Transfer Acceleration

- S3 Transfer Acceleration - Speeds up transferring files to and from Amazon S3 using the AWS CloudFront globally distributed edge locations
- exam keywords: S3, CloudFront, edge locations

##3 AWS Global Accelerator

- AWS Global Accelerator - Improve global application availability and performance AWS global network
- traffic is routed to the nearest AWS edge location and then to the application (can be running in a different region)
- exam keywords: availability, performance, local, global

CloudFront vs Global Accelerator

Both use edge locations to decrease latency and protection against DDoS attacks

- CloudFront:
  - caches content at edge locations to decrease latency
  - great for static content that needs to be available everywhere
- Global Accelerator:
  - no caching: routes traffic to the nearest AWS edge location and then to the application
  - improves global application availability and performance

AWS Outposts

- AWS Outposts - Fully managed service that extends AWS infrastructure, AWS services, APIs, and tools to virtually any datacenter, co-location space, or on-premises facility for a truly consistent hybrid experience
- AWS provides servers, storage, and networking equipment to run AWS services on-premises
- benfits: low latency, local data processing, data residency, and seamless integration with AWS services

- exam keywords: hybrid, on-premises, consistent

AWS Wavelength

- AWS Wavelength - Deliver ultra-low latency applications for 5G devices using AWS compute and storage at the edge of the 5G network
- AWS Wavelength embeds AWS compute and storage services at the edge of telecommunications providers' 5G networks
- Goal: reduce latency for 5G applications
- exam keywords: ultra-low latency, 5G, edge

AWS Local Zones

- AWS Local Zones - AWS infrastructure deployment that places compute, storage, database, and other select services closer to large population, industry, and IT centers
- Extend AWS region to a Local Zone in a specific geographic location so you can run applications that require single-digit millisecond latency directly at the Local Zone
- exam keywords: latency-sensitive, large population, industry centers

Global application architecture

- Single Region + single AZ :
    - low availability
    - low latency
    - low difficulty
- Single Region + multiple AZs :
    - high availability
    - low latency
    - medium difficulty
- Multi-Region - Active/Passive :
    - high availability
    - medium latency
    - higher difficulty
- Multi-Region - Active/Active :
    - high availability
    - low latency
    - highest difficulty

11 Cloud Integrations

SQS (Simple Queue Service)

- SQS - Fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications

- exam keywords: message queuing, decouple, microservices

SNS (Simple Notification Service)

- SNS - Fully managed messaging service for both application-to-application and application-to-person communication
- exam keywords: messaging service, application-to-application, application-to-person, topics

Kinesis

- Kinesis - Fully managed service for real-time processing of streaming data at massive scale
- exam keywords: real-time processing, streaming data

Amazon MQ

- Amazon MQ - Managed message broker service for Apache ActiveMQ and RabbitMQ that makes it easy to set up and operate message brokers in the cloud
- exam keywords: message broker, Apache ActiveMQ, RabbitMQ

12 Cloud Monitoring

CloudWatch Metrics & Alarms

- CloudWatch - Monitor AWS resources and applications in real-time
- CloudWatch Metrics: collect and track metrics, monitor log files, set alarms
- CloudWatch Alarms: send notifications or take actions based on defined rules
- exam keywords: monitoring, metrics, alarms

CloudWatch Logs

- CloudWatch Logs - Monitor, store, and access log files from Amazon EC2 instances, AWS CloudTrail, Route 53, and other sources
- exam keywords: log files, monitoring

EventBridge

- EventBridge - Serverless event bus that makes it easy to connect applications together using data from your own applications, integrated SaaS applications, and AWS services
- default event bus: receive events from AWS services
- partner event bus: receive events from SaaS applications (AWS parntners: Zendesk, Datadog, PagerDuty)
- custom event bus: receive events from your own applications
- ex: Cron jobs, S3 uploads, CodePipeline, CloudWatch Alarms trigger events that are sent to EventBridge and then to Lambda functions or other services
- exam keywords: serverless, event bus

CloudTrail

- CloudTrail - Record API calls for your account and delivers log files to you
- CloudTrail is enabled by default

- log files can be stored in S3 or CloudWatch Logs
- exam keywords: API calls, log files, tracking activity

X-Ray

- X-Ray - Analyze and debug distributed applications in production or under development
- allows you to trace requests from beginning to end with a visual map and latency data
- troubleshooting performance issues (latency, errors)
- exam keywords: analyze, debug, distributed applications

Amazon CodeGuru

- Amazon CodeGuru - Automated code reviews and application performance recommendations
- like: SonarQube, Checkmarx, Lint
- ML-powered powered service that helps you write better code and troubleshoot issues
- 2 components:
  - CodeGuru Reviewer: automated code reviews
  - CodeGuru Profiler: application performance recommendations: eg. reduce CPU usage, remove unused code
- exam keywords: code reviews, performance recommendations

AWS Health Dashboard

- AWS Health Dashboard - Provides alerts and remediation guidance for AWS services
- provides alerts and remediation guidance when AWS is experiencing events that may impact you
- exam keywords: alerts, remediation guidance

13 VPC - Virtual Private Cloud

VPC

- VPC - Virtual network that you create in an AWS region
- exam keywords: virtual network, region

Subnets

- Subnets - Range of IP addresses in your VPC
- subnets can be public or private
- subnets are defined per AZ
- subnets can be associated with route tables
- internet gateway is required for public subnets
- NAT gateway (AWS managed) or NAT instance (self managed) is required for private subnets to access the internet while remaining private(outbound only)
- exam keywords: IP addresses, public, private, AZ

NACLs

- NACLs - Act as a firewall for associated subnets, controlling inbound and outbound traffic
- NACLs are stateless: return traffic is not automatically allowed
- can have allow and deny rules

- are attached at the Subnet level
- exam keywords: firewall, inbound, outbound, stateless, subnet

Security Groups

- Security Groups - Act as a firewall for associated EC2 instances, controlling inbound and outbound traffic
- Security Groups are stateful: return traffic is automatically allowed, regardless of any rules
- can have allow rules only
- are attached at the EC2 instance level
- exam keywords: firewall, inbound, outbound, stateful, EC2 instance

VPC Flow Logs

- VPC Flow Logs - Capture information about the IP traffic going to and from network interfaces in your VPC
- VPC flow log, subnet flow log or Elastic Network Interface (ENI) flow log
- Help, monitor and troubleshoot connectivity issues in your VPC
- VPC Flow Logs can be published to CloudWatch Logs, S3, or Kinesis Data Firehose
- captures network information from AWS managed interfaces (eg. RDS, ElastiCache), Elastic Load Balancers, ENI's, etc.
- exam keywords: IP traffic, network interfaces, monitor, troubleshoot

VPC Peering

- VPC Peering - Connect two VPCs to route traffic between them using private IP addresses, so they behave as if they are on the same network
- Must not have overlapping CIDR blocks (IP ranges)

VPC Endpoints

- VPC Endpoints - Enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by PrivateLink without requiring an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection
- VPC Enpoint Gateway: S3 and DynamoDB only
- VPC Endpoint Interface: rest of the services
- exam keywords: private connection, VPC, AWS services

PrivateLink

- PrivateLink - Enables you to privately connect your VPC to services hosted by other AWS accounts (SaaS providers) and services hosted by other VPCs in your own account
- exam keywords: private connection, VPC, AWS services, SaaS providers

Direct Connect

- Direct Connect - Establish a dedicated network connection from your premises to AWS
- exam keywords: dedicated, network connection, on-premises

Site-to-Site VPN

- Site-to-Site VPN - Connect your on-premises network to your VPC over an IPsec VPN tunnel
- on premise you need a Custumer Gateway (CGW) and on AWS a Virtual Private Gateway (VGW) to establish the VPN connection
- exam keywords: on-premises, network, IPsec, VPN, fast implementation, public internet

Direct Connect(DAX) vs. Site-to-Site VPN

- Direct Connect:
  - dedicated network connection from your premises to AWS
  - private connection
  - higher bandwidth
  - more expensive
  - time-consuming to set up
- Site-to-Site VPN:
  - connect your on-premises network to your VPC over an IPsec VPN tunnel
  - uses the public internet
  - lower bandwidth
  - less expensive
  - faster to set up

Client VPN

- Client VPN - Managed client-based VPN service that enables you to securely access your AWS resources and resources in your on-premises network
- Connect to your VPC from your computer using OpenVPN
- exam keywords: client-based, VPN, secure

Transit Gateway

- Transit Gateway - Connect thousands of VPCs and on-premises networks using a single gateway
- hub-and-spoke model (star): connect multiple VPCs and on-premises networks to a central hub (Transit Gateway)
- exam keywords: thousands of VPCs, on-premises networks, single gateway

VPC Summary

- VPC: Virtual network in an AWS region
- Subnets: Range of IP addresses in your VPC
- NACLs: Firewall for subnets, stateless
- Security Groups: Firewall for EC2 instances, stateful
- VPC Flow Logs: Capture IP traffic going to and from network interfaces
- VPC Peering: Connect two VPCs to route traffic between them
- VPC Endpoints: Privately connect your VPC to supported AWS services
- PrivateLink: Privately connect your VPC to services hosted by other AWS accounts
- Direct Connect: Dedicated network connection from your premises to AWS
- Site-to-Site VPN: Connect your on-premises network to your VPC over an IPsec VPN tunnel
- Client VPN: Managed client-based VPN service
- Transit Gateway: Connect thousands of VPCs and on-premises networks using a single gateway

14 Security & Compliance

Shared responsibility model

- AWS is responsible for security of the cloud -> hardware, software, networking, and facilities that run AWS services
- Customer is responsible for security in the cloud -> data, identity, access management, platform, applications, and network configuration

AWS Shield

- AWS Shield is a managed DDoS protection service that safeguards applications running on AWS
- DDoS protection
- Standard: protection against most common DDoS attacks
- Advanced: protection against more sophisticated attacks
- exam keywords: DDoS protection

AWS WAF - Web Application Firewall

- Protects web applications from common web exploits
- Filter specific requests based on rules
- Rules can be based on IP addresses, HTTP headers, HTTP body, or URI strings
- exam keywords: WAF, web application firewall, protect web applications

AWS Network Firewall

- Managed firewall service that makes it easy to deploy essential network protections for all of your Amazon Virtual Private Clouds (VPCs)
- From Layer 3 to Layer 7 protection
- exam keywords: protect entire VPC, network firewall

AWS Firewall Manager

- Manage security rules in all accounts of an AWS Organization
- exam: manage VPC security groups on multiple accounts in an organisation --> AWS Firewall Manager
- Security policy: common set of security rules:
    - VPC Security Groups for EC2, ALB, RDS, ElastiCache, etc.
    - WAF rules for API Gateway, CloudFront, App Load Balancer
    - AWS Shield Advanced for DDoS protection
    - AWS Network Firewall for VPCs

Penetration Testing

- AWS customers are allowed to perform penetration testing on their AWS infrastructure within certain limits
- DDOS testing other attacks that can harm the AWS infrastructure are not allowed

Encryption in rest vs in transit

- Encryption in rest: data is encrypted when it is stored (data is not moving)
- Encryption in transit: data is encrypted when it is moving from one place to another (eg data moving from EC2 to S3)

Encryption with KMS

- AWS Key Management Service (KMS) is a managed service that makes it easy for you to create and control the encryption keys used to encrypt your data
- AWS manages the encryption keys for you, you define the policies
- type of KMS keys:
    - Customer Managed Keys: you create and manage the keys
    - AWS Managed Keys: AWS creates and manages the keys (aws/s3, aws/ebs)
    - AWS Owned Keys: Collection of CMKs (Customer Master Keys) that an AWS service owns and manages to use in multiple accounts
    - CloudHSM Keys: keys are stored in a hardware security module (CloudHSM)
- exam keywords: KMS, encryption, key management service

CloudHSM

- CloudHSM is a cloud-based hardware security module (HSM) that enables you to easily generate and use your own encryption keys on the AWS Cloud
- you manage the encryption keys entirely
- exam keywords: CloudHSM, hardware security module, encryption keys

KMS vs CloudHSM

- KMS: AWS manages the software for encryption, pay per use
- CloudHSM: AWS provisions encryption hardware, dedicated hardware, pay per hour

AWS Certificate Manager (ACM)

- AWS Certificate Manager is a service that lets you easily provision, manage, and deploy public and private Secure Sockets Layer/Transport Layer Security (SSL/TLS) certificates for use with AWS services and your internal connected resources

AWS Secrets Manager

- AWS Secrets Manager helps you protect access to your applications, services, and IT resources without the upfront investment and on-going maintenance costs of operating your own infrastructure
- Integration with Amazon RDS
- Possibility to rotate secrets automatically
- exam keywords: Secrets Manager, rotate secrets, RDS secrets

AWS Artifact

- AWS Artifact is your go-to, central resource for compliance-related information that matters to you
- eg. PCI DSS, HIPAA, ISO, SOC, etc.
- exam keywords: compliance, AWS Artifact

AWS GuardDuty

- AWS GuardDuty is a threat detection service that continuously monitors for malicious activity and unauthorized behavior to protect your AWS accounts and workloads
- uses Machine Learning to detect anomalies
- input data: VPC Flow Logs, CloudTrail Logs, DNS Logs
- exam tip: can protect against CryptoCurrency attacks
- exam keywords: GuardDuty, threat detection, malicious activity

Amazon Inspector

- Amazon Inspector is an automated security assessment service that helps improve the security and compliance of applications deployed on AWS
- For EC2 instances:
    - leveraging the AWS System Manager (SSM) agent
    - Analyze against unintended network accessibility, vulnerabilities, and deviations from best practices
    - analyze the running OS against known vulnerabilities
- For container image push to Amazon ECR:
    - assesment of the container image as they are pushed
- for Lambda functions:
    - assesment of the Lambda function code and package dependencies as it is deployed
- Reporting & integration with AWS Security Hub
- send findings to Amazon Event bridge
- exam keywords: Amazon Inspector, security assessment, compliance

AWS Config

- AWS Config is a service that enables you to assess, audit, and evaluate the configurations of your AWS resources
- Helps with auditing and recording compliance of your AWS resources based on the configurations you have set
- Helpful for big enterprises to make sure all resources are compliant with the company's security policies
- exam keywords: AWS Config, compliance, audit

Amazon Macie

- Amazon Macie is a security service that uses machine learning to automatically discover, classify, and protect sensitive data in AWS
- Macie helps and alerts you to sensitive data such as personally identifiable information (PII)
- can be integrated with Amazon EventBridge
- exam keywords: Macie, sensitive data, PII

AWS Security Hub

- Central security tool to manage security across multiple AWS accounts and automate security checks
- Integrated dashboards showing current security and compliance status to quickly take actions

- Automatically aggregate alerts in predefined or personal findings format from various AWS serivces & AWS partner tools:
    - Config
    - GuardDuty
    - Inspector
    - Macie
    - IAM Access Analyzer
    - AWS Systems Manager
    - AWS Firewall Manager
    - AWS Health
- must enable AWS Config Service to use Security Hub
- exam keywords: Security Hub, security checks, compliance

Amazon Detective

- GuardDuty, Macie and Security Hub are great tools to detect security issues, but they do not provide a full investigation
- Amazon Detective is a service that helps you to investigate and identify the root cause of potential security issues or suspicious activities
- automatically collects and process events from VPC Flow Logs, CloudTrail, GuardDuty to create a visual graph of the resources and the interactions between them. So you can easily identify the root cause of the issue.
- exam keywords: Detective, investigate, root cause

AWS Abuse

- AWS Abuse is a service that helps you to report any abuse of the AWS services
- eg. phishing, malware, spam, DoS or DDoS etc. that is hosted on AWS
- exam keywords: AWS Abuse, report abuse

Root user privileges

- Root user has full access to all AWS services and resources in the account
- Some actions can only be performed by the root user:
    - `change account settings( account name, email, root user password, root user access keys)`
    - `close the account`
    - `change or cancel AWS support plan`
    - restore IAM user permissions
    - view certain tax invoices
    - `register as a seller in the Reserved Instance Marketplace`
    - Configure and Amazon S3 bucket to enable MFA
    - edit or delete an Amazon S3 bucket policy that includes an invalid VPC ID or VPC endopoint ID
    - sign up for GovCloud

IAM Access Aanlyzer

- IAM Access Analyzer helps identify which resources are shared externally:

    - S3 Buckets
    - KMS Keys
    - IAM Roles
    - Lambda Functions
    - SQS Queues
    - Secrets Manager Secrets
- Define Zone of Trust = AWS account or AWS Organization
- Access outiside zone of trust is flagged as a finding
- exam keywords: IAM Access Analyzer, shared resources, Zone of Trust, analyze access

15 Machine Learning

Rekognition

- Rekognition - Deep learning-based image and video analysis service
- Image recognition, face detection, face analysis, face comparison, text in image, unsafe content detection
- exam keywords: image recognition, face detection, text in image

Amazon Transcribe

- Transcribe - Automatic speech recognition service
- Convert speech to text (multiple languages)
- Remove PII (Personal Identifiable Information) from the transcript automatically
- exam keywords: speech to text

Amazon Polly

- Polly - Text-to-speech service
- Convert text into lifelike speech
- exam keywords: text to speech

Amazon Translate

- Translate - Neural machine translation service
- Translate text between languages
- exam keywords: machine translation

Amazon Lex

- Lex - Conversational interfaces for your applications (same tech as Alexa)
- Automatic Speech Recognition to convert speech to text
- Natural Language Understanding to recognize the intent of the text

Connect

- Connect - Cloud-based contact center service
- Receive calls, create contact flows, cloud-based virtual contact center
- can integrate with other CRM (Customer Relationship Systems) systems or AWS services
- exam keywords: contact center

- No upfront payments, no long-term contracts, 80% less than traditional contact centers

Amazon Comprehend

- Comprehend - Natural Language Processing (NLP) service
- Sentiment analysis, key phrase extraction, language detection, entity recognition
- fully managed and serverless
- exam keywords: NLP, sentiment analysis, key phrase extraction

Amazon SageMaker

- SageMaker - Fully managed service that enables developers and data scientists to quickly and easily build, train, and deploy machine learning models at any scale
- Typically diffucult to do all the processes in one place + provision servers

Forecast

- Forecast - Fully managed service for time-series forecasting
- exam keywords: time-series forecasting

Kendra

- Kendra - Intelligent search service powered by machine learning
- Fully managed document search service powered by machine learning
- extract answers from within a document
- exam keywords: intelligent search, document search service

Amazon Personalize

- Personalize - Real-time personalization and recommendation service
- exam keywords: real-time personalization, recommendation

Textract

- Textract - Extract text and data from documents
- exam keywords: extract text, data, documents

Summary

- Rekognition: image and video analysis
- Transcribe: speech to text
- Polly: text to speech
- Translate: machine translation
- Lex: conversational bots - chatbots
- Connect: cloud-based contact center
- Comprehend: natural language processing
- SageMaker: machine learning for every developer and data scientist
- Forecast: time-series forecasting
- Kendra: ML powered search engine
- Personalize: real-time personalization and recommendation

- Textract: extract text and data from documents

16 Account Management & Billing

1 AWS Organizations

- AWS Organizations helps you centrally manage and govern your environment as you grow and scale your AWS resources
- Organize accounts into Organizational Units (OUs):
    - OUs can contain other OUs
    - OUs can contain accounts
    - eg. dev, test, prod, finance,
- Main account is the master account
- Apply service control policies (SCPs) to OUs
- Consolidated billing across all accounts
- Benefits:
    - Centralized management of multiple AWS accounts
    - Cost savings through consolidated billing
    - Fine-grained control over access, security, and compliance: SCPs
    - Automation of account creation and management: through API
    - exam tip: Organizations is used to manage multiple AWS accounts
- exam keywords: Organizations, OUs, SCPs, consolidated billing

A. Multi account strategies

- Single Account:
    - simple, easy to manage
    - no isolation between environments
    - no cost separation
- Multi Account:
    - separate accounts for dev, test, prod
    - isolation between environments
    - cost separation
    - more complex to manage

B. SCPs

- Service Control Policies (SCPs) are a type of organization policy that you can use to manage permissions in your organization
- applied to OUs
- SCPs are applied to all accounts in the OU
- SCPs are used to restrict permissions (eg. deny access to S3) or to allow permissions (eg. allow access to EC2)
- JSON based policies like IAM policies
- exam keywords: SCPs, Service Control Policies, organization policy

C. Consolidated billing

- Consolidated billing is a feature that enables you to consolidate payment for multiple AWS accounts within your organization by designating one of the accounts as the payer account
- Benefits:
    - volume pricing: combined usage for all accounts
    - easy tracking of charges: one bill
    - easy to allocate costs: cost allocation tags
    - exam tip: consolidated billing is used to consolidate billing across multiple accounts

AWS Control Tower

- Easy way to se up and govern a secure and compliant multi-account AWS environment based on best practices
- Benefits:
    - set up a secure multi-account environment
    - automate the setup of accounts
    - centrally manage security and compliance
    - reduce the time to set up new accounts
    - exam keywords: Control Tower, multi-account, secure environment
- AWS Control Tower runs on top of AWS Organizations: it automatically sets up AWS Organizations and implements the SCPs (Service Control Policies) for you
- exam keywords: Automated multi-account setup, best practices, security, compliance

AWS Resource Access Manager (RAM)

- Share AWS resources that you own with other AWS accounts
- Share with any account or within your organization
- Avoid resource duplication
- Supported resources: Aurora, VPC Subnets, Transit Gateway Route 53, EC2 Dedicated Hosts, License Manager, etc.

AWS Service Catalog

- Users that are new to AWS have too many options and may create stacks that are not compliant with company policies
- AWS Service Catalog allows you to create and manage catalogs of IT services that are approved for use on AWS
- Users can launch only the products that are approved by the IT department
- The products are CloudFormation templates
- Use IAM permissions to control access to the products
- exam keywords: Service Catalog, IT services, approved products

Pricing models in AWS

- Pay as you go:
    - pay for what you use
    - no upfront cost
    - no long-term commitment
- Save when you reserve:

    - commit to a specific instance configuration
    - pay upfront or pay partially upfront
    - 1 year or 3 year term
    - significant discount compared to on-demand
- Pay less by using more:
    - volume discounts
    - tiered pricing
- Pay less as AWS grows:
    - AWS lowers prices over time
    - AWS passes the savings to customers

Free services

- IAM
- VPC
- Consolidated billing
- Elastic Beanstalk (*)
- CloudFormation (*)
- Auto Scaling Groups (*) (*) you pay for the resources that are created by these services
- Free tier:
    - 12 months free: 750 hours of EC2, 5GB of S3, 25GB of DynamoDB
    - Always free: 1M Lambda requests, 1M free requests on API Gateway, 25GB of EBS

EC2 pricing

- On-demand instances:
    - minimum of 60s
    - pay per second after the first minute (Linux/Windows) or per hour (other)
    - no long-term commitment
- Reserved instances:
    - 1 year or 3 year term
    - significant discount compared to on-demand
    - pay upfront or pay partially upfront
    - convertible reserved instances: change the instance type
    - scheduled reserved instances: launch within time window
- Spot instances:
    - bid for unused EC2 capacity
    - can be terminated by AWS with 2 minutes notice
    - great for batch jobs, big data analysis, etc.
- Dedicated hosts:
    - physical server with EC2 instance capacity fully dedicated to your use
    - can help you reduce costs by allowing you to use your existing server-bound software licenses
    - can be purchased On-Demand (hourly)
    - can be purchased as a Reservation for up to 70% off the On-Demand price
- Savings plans: future lecture

Lambda and ECS

- Lambda:
    - pay per call
    - pay per duration
    - free tier: 1M free requests per month, 400,000 GB-seconds of compute time per month
- ECS:
    - EC2 Launch Type Model: No additional fees, you pay for AWS resources stored and created in you application
- Fargate Launch Type Model: pay for vCPU and memory used by the containers

Storage pricing - S3

Pricing depends on:

- Storage class:
    - S3 Standard: general purpose storage of frequently accessed data
    - S3 Infrequent Access (IA): long-lived, but less frequently accessed data
    - S3 One Zone-IA: long-lived, infrequently accessed, non-critical data
    - S3 Intelligent Tiering: designed to optimize costs by automatically moving data to the most cost-effective access tier
    - S3 Glacier: long-term archive
    - S3 Glacier Deep Archive: lowest cost storage class for long-term retention and digital preservation
- Number and size of objects: Price can be tiered (base on volume)
- Number and type of requests
- Data transfer OUT of S3 region (IN is free)
- Transfer acceleration: fast, easy, and secure transfers of files over long distances between your client and an S3 bucket
- Lifecycle transitions: move objects to different storage classes

Similar pricing model for EFS pay per use, has infrequent access & lifecycle rules

Storage pricing - EBS

Pricing depends on:

- Type of volume:
    - General Purpose SSD (gp2)
    - Provisioned IOPS SSD (io1)
    - Throughput Optimized HDD (st1)
    - Cold HDD (sc1)
    - Magnetic (standard)
- Size of volume
- IOPS:
    - Generl purupose SSD: included
    - Provisioned IOPS SSD: you pay for the IOPS provisioned

    - Magnetic: number or requests
- Snapshots: incremental backups of EBS volumes:
    - Added data cost per GB per month
- Data transfer:
    - Outbound data transfer are tiered for volume discounts
    - Inbound data transfer is free

Pricing for RDS

- per hour billing
- Database charateristics:
    - Engine
    - Size
    - Memory class
- Purhcase type:
    - On-demand
    - Reserved (1 year or 3 year term)
- Backup storage: There is no additional charge for backup storage up to 100% of your total database storage for a region
- Additional storage: you pay for the storage you provision
- Number of input and output requests per month
- Deployment type: Single-AZ or Multi-AZ
- Data transfer: Outbound data transfer are tiered for volume discounts, Inbound data transfer is free

Content Delivery Network (CDN) - CloudFront

- Pricing is different accross different regions
- Aggregated for each edge location, then applied to your bill
- Data Transfer out (volume discount)
- Number of HTTP/HTTPS requests

Networking costs in AWS per GB - Simplified

- Free for traffic into Availability Zones
- Free for traffic between EC2 instances in the same Availability Zone
- Traffic between 2 EC2 instances in different Availability Zones:
    - $0.01 per GB using private IP
    - $0.02 per GB using public IP / Elastic IP
- Traffic between 2 EC2 instances in different regions: $0.02 per GB inter region data transfer

AWS Savings plan

- AWS Savings Plans is a flexible pricing model that offers low prices on EC2 and Fargate usage, in exchange for a commitment to a consistent amount of usage (measured in $/hour) for a 1 or 3 year term
- Easiest way to setup long-term commitments on AWS
- 2 types of Savings Plans:
    - Compute Savings Plans:

        - up to 66% discount compared to on-demand
        - regardless of family, region, size, OS, tenacy or compute options
        - compute options: EC2, Fargate, Lambda
    - EC2 Instance Savings Plans:
        - up to 72% discount compared to on-demand
        - commit to usage of individual instance families in a region (eg. C5 or M5 instances in US West)
        - regardless of AZ, size (eg. M5.2xlarge), OS or tenacy
        - All upfront, partial upfront, no upfront
- Machine Learling Savings Plans: Sagemaker
- Setup from the AWS Cost Explorer console

AWS Compute Optimizer

- AWS Compute Optimizer is a service that recommends optimal AWS resources for your workloads to reduce costs and improve performance by using machine learning
- Analyzes resource configurations and CloudWatch metrics and uses machine learning to recommend optimal resources
- supported resources: EC2 instances, Auto Scaling Groups, ECS tasks, Lambda functions, RDS databases

Billing and Costing Tools

- Estitmate costs in the cloud:
    - AWS Pricing Calculator
- Tracking costs in the cloud:
    - Billing Dashboard
    - Cost Explorer
    - Cost Allocation Tags
    - Cost and Usage Reports
- Monitoring agains cost plans:
    - Billing Alarms
    - Budgets

Billing & Costing tools

Pricing Calculator

- Pricing Calculator - Estimate your monthly bill using AWS products
- exam keywords: estimate, monthly bill

AWS Billing Dashboards

- Billing Dashboards - Monitor your AWS billing and usage
- High level overview of your AWS billing and usage
- Gives estimates of your monthly bill

Free Tier Dashboard

- Free Tier Dashboard - Monitor your free tier usage

Cost Allocation tags

- Cost Allocation tags - Tags that you can assign to your AWS resources
- Use cost allocation tags to track your AWS cost on a detailed level
- Use AWS or User-Defined tags

Tagging and Resource Groups

- Tagging and Resource Groups - Create and manage groups of resources
- Use tags to group resources together
- exam keywords: tags, resource groups
- Resources created by CloudFormation are all tagged the same way

Cost and Usage Reports

- Cost and Usage Reports - Access detailed billing reports
- The AWS Cost and Usage Report contains the most comprehensive set of AWS cost and usage data available
- The lists AWS usage for each service category used by an account and its IAM users in hourly or daily line items, as well as any tags that you have activated for cost allocation purposes
- Can be integrated with Athena, Redshift or QuickSight
- Export to S3 bucket in CSV or Parquet format

Cost Explorer

- Cost Explorer - Visualize, understand, and manage your AWS costs and usage over time
- Create custom reports that analyze cost and usage data
- High level overview of your AWS costs and usage
- Forecast usage up to 12 months (exam question)
- Choos an optimal Savings Plan (to lower prices on your bill)

Monitoring: Billing Alarms in CloudWatch

- Billing Alarms in CloudWatch - Set up billing alarms to monitor your estimated AWS charges
- helpful for email notifications when your bill exceeds a certain threshold

AWS Budgets

- Budgets - Set custom cost and usage budgets that alert you when you exceed your thresholds
- Up to 5 SNS notifications per budget
- First 2 budgets are free
- exam keywords: budgets, cost and usage budgets

AWS Cost Anomaly Detection

- Cost Anomaly Detection - Detects unusual spending patterns

- Continuously monitors your AWS usage and cost patterns and uses machine learning to identify your normal and anomalous spend patterns
- You don't need to define any rules or set up any thresholds
- Sends you the anomaly detection findings via email with root cause analysis
- exam keywords: anomaly detection, unusual spending patterns

AWS Service Quotas

- Service Quotas - View and manage your quotas for AWS services
- Notify you when you're approaching your quota (eg. number of EC2 instances, Lambda executions, etc.)
- Request a quota increase if needed
- exam keywords: quotas, request increase

AWS Trusted Advisor

- Trusted Advisor - Provides real-time guidance to help you provision your resources following AWS best practices
- No need to install anything - high level AWS account check
- For full set of checks, you need Business or Enterprise support

AWS Support Plans Pricing

- Support Plans Pricing - AWS Support Plans Pricing
- Basics: free - good for exploring AWS
    - Customer Service 24/7, documentation, whitepapers, best practices
    - AWS Trusted Advisor: core checks
    - AWS Personal Health Dashboard: alerts and remediation guidance
- Developer: Good for development and test
    - All Basic features
    - Business hours email support
    - Unlimited cases and contacts
    - General guidance: <24h business hours response time
    - System impaired: <12h business hours response time
- Business: Good for production workloads
    - All Developer features
    - 24/7 phone, email, chat support
    - Full set of Trusted Advisor checks + API access
    - Infrastructure event management for additional fee
    - General guidance: <24h business hours response time
    - System impaired: <12h business hours response time
    - Production system impaired: <4h business hours response time
    - Production system down: <1h business hours response time
- Enterprise On-Ramp Support Plan: production or business critical workloads
    - All Business features
    - Access to a pool of Technical Account Managers (TAMs)
    - Concierge support team (for billing and account best practices)

    - Infrastructure event management included, Well-Architected & Operations Reviews
    - General guidance: <24h business hours response time
    - System impaired: <12h business hours response time
    - Production system impaired: <4h business hours response time
    - Production system down: <1h business hours response time
    - Business critical system down: <30min business hours response time
- Enterprise Support plan: mission critical workloads
    - All Enterprise On-Ramp features
    - Access to a designated Technical Account Manager (TAM)
    - General guidance: <24h business hours response time
    - System impaired: <12h business hours response time
    - Production system impaired: <4h business hours response time
    - Production system down: <1h business hours response time
    - Business critical system down: <15min business hours response time

Account Best practices Summary

- Use Organizations to manage multiple accounts
- Use SCP (Service Control Policies) to restrict account power
- Easily setup multiple accounts with best practices using AWS Control Tower
- Use Tags & Cost Allocation Tags for easy management and billing
- IAM guidelines: MFA, least privilege, password policies, password rotation
- Config to
- CloudFormation to deploy stacks across multiple accounts and regions
- Trusted Advisor to get insights, Support Plan adapted to your needs
- Sends Service Logs and Access Logs to S3 or CloudWatch Logs
- CloudTrail to record API calls made within your account
- If your Account is compromised: change the root password, delete and rotate all passwords/keys, contact AWS Support
- Use AWS Service catalog to create pre-defined stacks defined by your organization

17 Advanced Identity

Security Token Service (STS)

- STS - Create temporary, limited-privileges security credentials to access AWS services
- Short-term credentials: you configure the expiration time
- Use cases:
    - Identity federation: manage user identities in external systems and proved them with STS tokens to access AWS resources
    - Cross-account access
    - IAM Roles for EC2 instances: provide temporary credentials for EC2 instances to access other AWS services
    - exam keywords: temporary credentials, limited-privileges

Cognito

- Cognito - Identity management for Web and Mobile applications users (potentially millions)

- Sign-up, sign-in, and access control for web and mobile apps quickly and easily
- login with social identity providers (Google, Facebook, Amazon, Apple, etc.)
- exam keywords: identity management, sign-up, sign-in

Directory Service

- Directory Service - Managed Microsoft Active Directory in the AWS Cloud
- 3 types of directory services: (not required for exam)
  - AD Connector: proxy to connect to on-premises AD
  - Simple AD: AD-compatible managed directory
  - Managed AD: fully managed Microsoft AD
- exam keywords: managed Microsoft Active Directory

AWS IAM Identity Center (successor of)

- IAM Identity Center - Centralized place to manage all your AWS identities
- One login (single sign-on) to access all your accounts from the IAM Identity Center portal, type of accounts:
  - AWS accounts in AWS Organizations (exam question)
  - Business cloud applications (Salesforce, Office 365, etc.)
  - SAML 2.0 compliant applications
  - EC2 Windows Instances
- Identity providers:
  - Build-in identity store in IAM Identity Center
  - 3rd party: Active Directory, OneLogin, Okta, Ping Identity
- exam keywords: one login, centralized, AWS Organizations

18 Other Services

Workspaces

- Workspaces - Desktop-as-a-Service (DaaS) solution
- Secure, managed, cloud-based virtual desktops
- exam keywords: Desktop-as-a-Service, virtual desktops

AppStream

- AppStream - Fully managed application streaming service
- Stream desktop applications to users and they can access it from a web browser (ex. stream Photoshop)

IoT Core

- IoT Core - Connect IoT devices to the AWS Cloud
- serverless, scalable, and secure
- integrates with other AWS services (eg. Lambda, S3, SageMaker, etc.)
- exam keywords: IoT, devices, AWS Cloud

Amazon Elastic Transcoder

- Elastic Transcoder - Convert media files stored in S3 into media files in the formats required by consumer playback devices

AppSync

- AppSync - Fully managed GraphQL service
- store and sync data across devices in real-time --> build backend for web and mobile applications
- uses GraphQL to make it easy for applications to get exactly the data they need
- exam keywords: GraphQL, real-time, web, mobile

Amplify

- Amplify - Full-stack framework for developing web and mobile applications
- Amplify provides a library, CLI toolchain, and UI components
- exam keywords: full-stack, web, mobile

AWS Application Composer

- Application Composer - Build applications without writing any code
- Generates Infrastructure as Code (IaC) for you using CloudFormation
- Ability to import existing CloudFormation templates to build applications
- exam keywords: no code, build applications

AWS Device Farm

- Device Farm - Test your web and mobile applications on real devices in the AWS Cloud
- exam keywords: test, web, mobile, real devices

AWS Backup

- Backup - Centralized backup service for backing up data across AWS services
- on-demand and scheduled backups
- exam keywords: backup, centralized

Disaster Recovery Strategies

- Backup and Restore: backup data to S3, EBS snapshots, RDS snapshots, etc. -> cheapest solution
- Pilot Light: small version of your application is always running in another region -> more expensive
- Warm Standby: full version of your application is always running in another region but at minimum size -> more expensive
- Multi-Site: active-active setup in multiple regions -> most expensive

Disaster Recovery for Cloud Deployments

- failover to another region (us-east-1 -> us-west-2) using route 53

AWS Elastic Disaster Recovery (DRS) (before CloudEndure)

- Elastic Disaster Recovery - Disaster recovery service that helps you recover quickly your phisical, virtual, and cloud-based servers into AWS

- protect against ransomware, data corruption, and natural disasters
- exam keywords: disaster recovery, recover quickly

AWS DataSync

- DataSync - Automate data transfer between on-premises storage and Amazon S3 or Amazon EFS
- move large amount of data from on-premises to AWS
- the replication tasks are incremental after the first full load
- exam keywords: data transfer, on-premises, S3, EFS, incremental

Cloud Migration Strategies: 7Rs

- Retire: turn of things you don't need (eg. old servers) or as result of Re-architecting
- Retain: keep as is: do nothing for now
- Relocate: move apps from on-premises to the cloud version
- Rehost (lift and shift): move apps from on-premises to the cloud without changes, databases and data. Use AWS Application Migration Service for example.
- Replatform (lift and reshape): dont change core architecture, but optimize for cloud (eg. use managed services)
- Repurchase (drop and shop): moving to a diffrent product (eg. SaaS) while moving to the cloud. Expensive in the short term but quick to deploy.
- Refactor (re-architect): re-architect the application to be cloud-native. Use AWS services to optimize the application. Most expensive but most benefits.

AWS Application Discovery Service

- Application Discovery Service - plan migration projects by gathering information about on-premises data centers
- exam keywords: migration projects, on-premises data centers

AWS Application Migration Service (MGN)

- Application Migration Service - Migrate applications from on-premises to AWS
- lift and shift (rehost) or replatform (rehost and optimize)
- exam keywords: migrate applications, on-premises, AWS

AWS Migration Evaluator

- Migration Evaluator - Analyze on-premises workloads to estimate costs and identify potential savings
- exam keywords: analyze, on-premises workloads, estimate costs

AWS Migration Hub

- Migration Hub - Track the progress of application migrations to the cloud
- Central location to collect servers and applications inventory data for the assessment planning and tracking of migrations to AWS
- integrated with AWS Application Migration Service (MGN) and Database Migration Service (DMS)
- exam keywords: centralized, track progress, application migrations

AWS Fault Injection Simulator

- Fault Injection Simulator - Test the resilience of your applications by injecting faults
- chaos engineering: test how your system behaves under stress (eg high CPU load, network failure)
- exam keywords: test resilience, inject faults

AWS Step functions

- Step functions - Serverless orchestration service that lets you coordinate multiple AWS services into serverless workflows
- exam keywords: serverless, orchestration, workflows

AWS Ground Station

- Ground Station - Fully managed service that lets you control satellite communications, process data, and scale your operations without having to worry about building or managing your own ground station infrastructure
- Allows you to download data from satellites into AWS (S3, EC2, etc.)
- exam keywords: satellite communications, download satelite data

Amazon Pinpoint

- Pinpoint - Targeted push notifications, emails, SMS, and voice messages
- scalable 2 way (outbound and inbound) communication
- exam keywords: push notifications, emails, SMS, voice messages

19 AWS Well-Architected Framework

AWS Well-Architected Framework

- Well-Architected Framework - Best practices for designing and operating reliable, secure, efficient, and cost-effective systems in the cloud
- 6 pillars:
    - Operational Excellence
    - Security
    - Reliability
    - Performance Efficiency
    - Cost Optimization
    - Sustainability
- exam keywords: best practices, reliable, secure, efficient, cost-effective

AWS Cloud Adoption Framework (CAF)

- Cloud Adoption Framework - Guidance to help organizations develop an efficient and effective plan for their cloud adoption journey
- exam keywords: guidance, cloud adoption journey

CAF Perspectives

- (will probably have exam question: identify the perspective)
- CAF groups capabilities into 6 perspectives:

    - `Business`: helps ensure that cloud investments accelerate digital transformation ambitions and business outcomes
    - `People`: serves as a bridge between technology and business, helps organizations to evolve to a culture of continous growth, learning and where change becomes business as usual with focus on culture, organizational structure, leadership and workfore
    - `Governance`: helps orchestrate cloud initiatives
    - `Platform`: helps you build an enterprise-grade, scalable, hybrid cloud platform
    - `Security`: helps achieve confidentiality, integrity, and availability of data and systems in the cloud
    - `Operations`: helps ensure that cloud services are delivered at a level that meets the need of your business
- Business: Business, People, Governance
- Technical: Platform, Security, Operations

CAF Tranformation Domains

- `Technology`: using the cloud to migrate and modernize legacy infrastructure, applications, data and analytics platforms
- `Process`: digigitizing and automating business processes:
- leverage new data and analytics platforms
- use machine learning and AI to drive innovation
- `Organization`: reimagining your operation model
    - organizing your team around products and value streams
    - leveraging agile methods to rapidly iterate and evolve
- `Product`: reimagining your business model by creating new value propositions (products and services) and revenue models

CAF Transformation Phases

- `Envision`: demonstrate how the cloud will accelerate your business strategy
- `Align`: identify cabality gaps acrosse the 6 AWS CAF perspectives which result in an action plan
- `Launch`: build and deliver pilot initiatives in production and demonstrate incremental business value
- `Scale`: expand pilot initiatives to the desired scale while realizing the desired business benefits

AWS Right sizing

- Right sizing - important to choose the right instance type for your workload. The most powerfull instance is not always the best choice, because the cloud is elastic and you can scale up and down as needed
- scaling up is easy so start small and scale up as needed
- Imporant to right size in 2 moments:
    - before cloud migration
    - continously after the cloud onboarding process because requirements change over time

AWS re:Post

- re:Post - AWS re:Post is a community-driven collection of best practices, articles, and tools that have been shared by AWS Solutions Architects

- exam keywords: community-driven, best practices

AWS Knowledge Center

- Knowledge Center - AWS Knowledge Center provides answers to frequently asked questions, best practices, and troubleshooting tips from AWS Support
- part of re:Post because it is a community-driven collection of best practices
- exam keywords: frequently asked questions, best practices

AWS Managed Services (AMS)

- Managed Services - Managed Services for AWS
- AMS Team operates AWS on your behalf, providing a secure and compliant AWS Landing Zone, a proven enterprise operating model, on-going cost optimization, and day-to-day infrastructure management
- exam keywords: managed services, secure, compliant