03/12/25 - AWS Introduction

Amazon Web Services

Leading cloud platform with over 200 different services available

Globally available via its massive networks of regions and availability zones with their massive data centers

Based on a pay-as-you-use cost model

- *Theoretically* cheaper than renting rackspace/servers in a data center

History

Originally launched in 2006 with only 2 services: S3 and EC2

By 2010: SimpleDB, Elastic Block Store, RDS, DynamoDB, and others

Amazon had competitions with big prizes to spur the adoption of AWS in its early days

AWS Service Categories

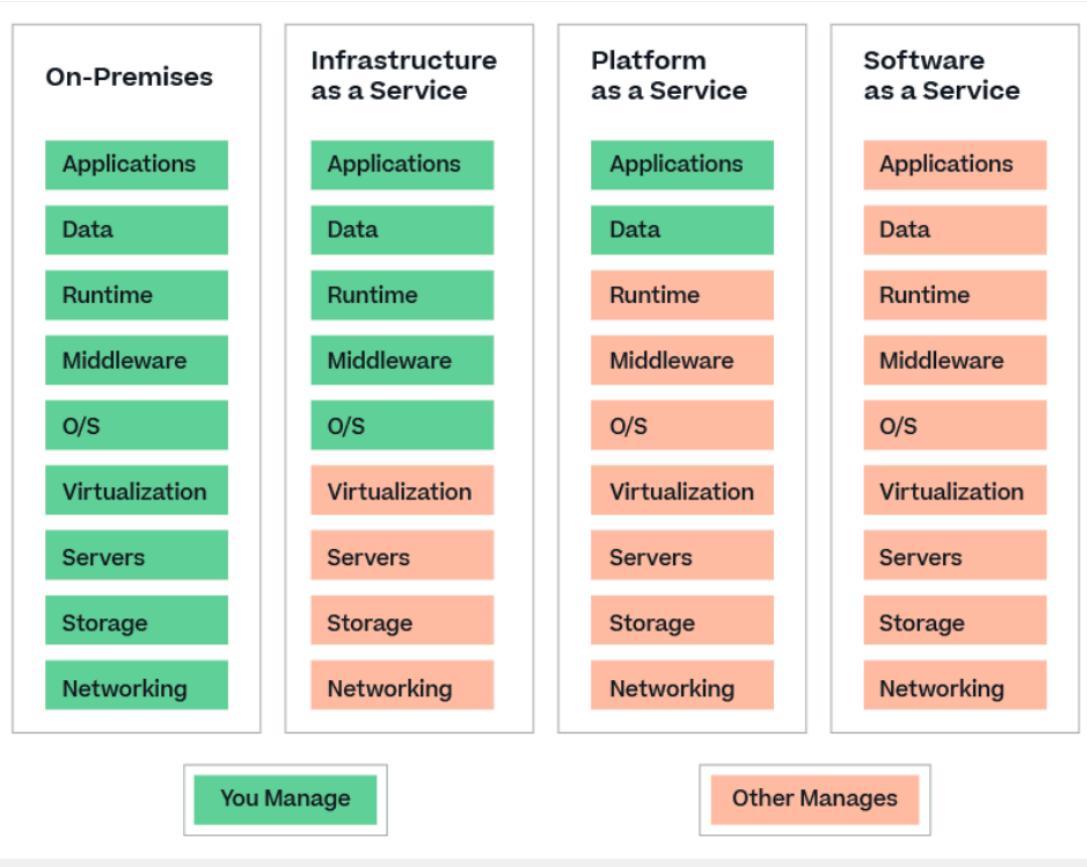
Cloud Models

IaaS (Infrastructure as a Service)

- Contains the basic services that are needed to build an IT infrastructure

PaaS (Platform as a Service)

SaaS (Software as a Service)



Shared Responsibility Model

AWS Responsibilities (Security of the cloud)

- Security of physical infrastructure (infra) and network
- Hypervisor and Host OSs
    - Hypervisor: essentially software that takes one machine and allows multiple clients to feel that they have their own individual, isolated spot
- Maintaining managed services

Client Responsibilities (Security in the cloud)

- Control of Data/Content
- Access Management and IAM
  - Principle of Least Privilege
- Manage self-hosted Apps and associated OSs
- Ensure network security to its VPC
- Handle compliance and governance policies and procedures

AWS Global Infrastructure

Regions: distinct geographical areas

- us-east-1, us-west-1, etc.

Availability Zones (AZs)

- Each region has multiple AZs
- Roughly equivalent to isolated data centers

Edge Locations

- Locations for CDN (content delivery networks) and other types of caching services
- Allows content to be closer to end user

Compute Services

VM-based

- EC2 and EC2 Spot - Elastic Cloud Compute

Container-based

- ECS - Elastic Container Service
  - Like Docker
- ECR - Elastic Container Registry
- EKS - Elastic Kubernees Service
- Fargate - Serverless container service

Serverless

- AWS Lambda

Storage Services

Amazon S3 - Simple Storage Service

- Object storage in buckets; highly scalable; different storage classes

Amazon EFS - Elastic File System

- Simple, serverless, elastic, “set-and-forget” file system

Amazon EBS - Elastic Block Storage

- High-performance block storage service

Amazon File Cache

AWS Backup

Database Services

Relational

- Amazon RDS
- Amazon Aurora

Key-Value

- Amazon DynamoDB

In-Memory

- Amazon MemoryDB
- Amazon ElastiCache

Document

- Amazon DocumentDB (compatible with MongoDB)

Graph

- Amazon Neptune

Analytics Services

Amazon Athena

- Analyze petabyte scale data where it lives (S3, for example)

Amazon EMR

- Elastic MapReduce: transform values and then amalgamating the results in some way

AWS Glue

Amazon Redshift

Amazon Kinesis

Amazon QuickSight

ML and AI Services

Amazon SageMaker

- Fully-managed ML platform, including Jupyter NBs
- Built, train, deploy ML models

AWS AI Services with pre-trained models

- Amazon Comprehend
- Amazon Rekognition
- Amazon Textract
- Amazon Translate

AWS Free Tier

Allows you to gain hands-on experience with a subset of the services for 12 months (service limitations apply as well)

- Amazon EC2: 750 hours/month (specific OSs and Instance Sizes)
- Amazon S3: 5GB (20K GETs, 2K Puts)
- Amazon RDS: 750 hours/month of DB use (within certain limits)