DS 4300

AWS Introduction

Mark Fontenot, PhD
Northeastern University

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… *Theoretically*.

History of AWS

- Originally launched in 2006 with only 2 services: S3 & EC2.
- By 2010, services had expanded to include SimpleDB, Elastic Block Store, Relational Database Service, DynamoDB, CloudWatch, Simple Workflow, CloudFront, Availability Zones, and others.
- Amazon had competitions with big prizes to spur the adoption of AWS in its early days
- They’ve continuously innovated, always introducing new services for ops, dev, analytics, etc… (200+ services now)

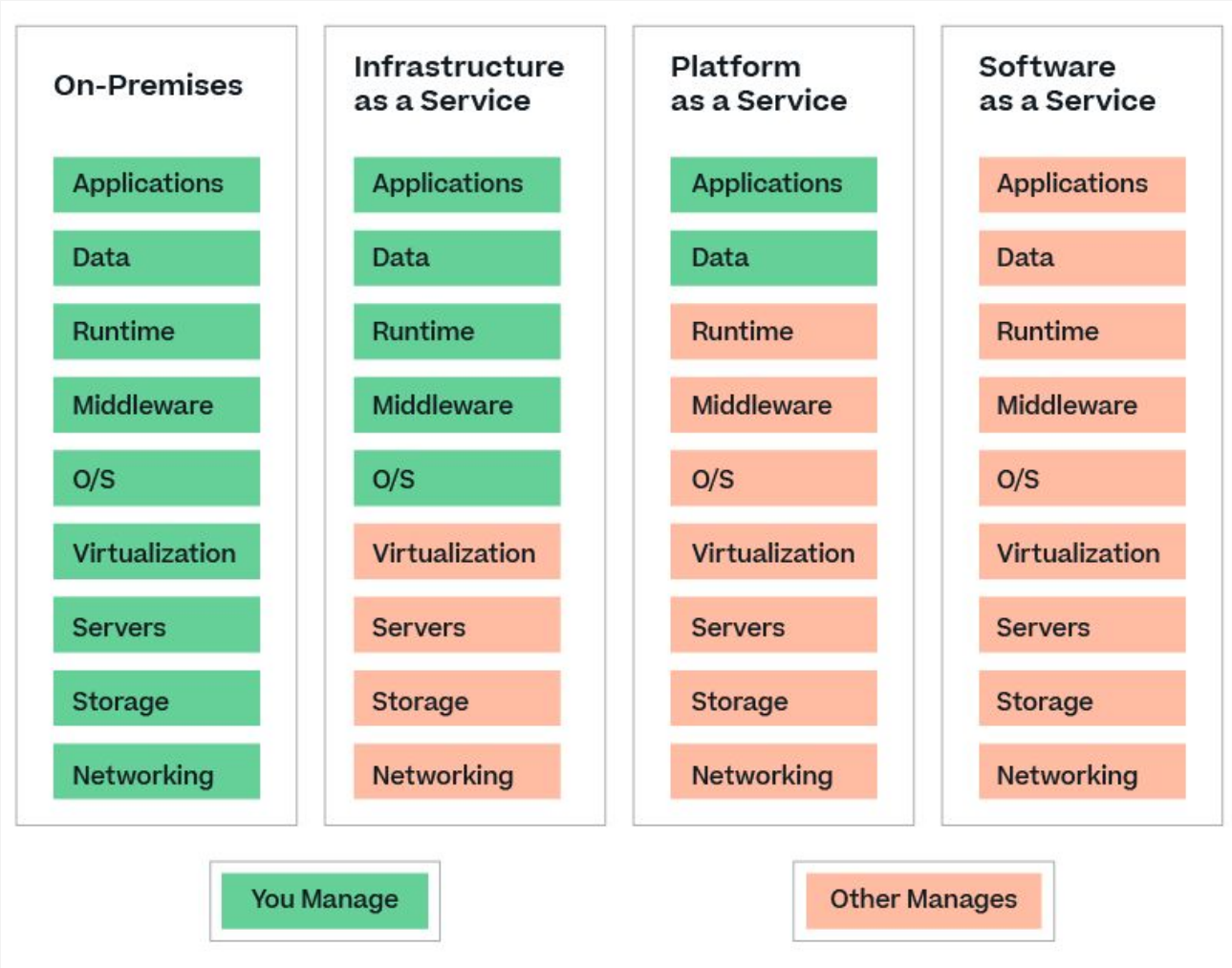
AWS Service Categories

Analytics	Application integration	Blockchain	Business applications	Cloud Financial Management	Compute
Customer enablement	Containers	Databases	Developer tools	End user computing	Front-end web and mobile
Game tech	Internet of Things (IoT)	Machine Learning (ML) and Artificial Intelligence (AI)	Management and governance	Media	Migration and transfer
Networking and content delivery	Quantum technologies	Robotics	Satellite	Security, identity, and compliance	Storage

Cloud Models

- IaaS (more) - Infrastructure as a Service
  - Contains the basic services that are needed to build an IT infrastructure
- PaaS (more) - Platform as a Service
  - Remove the need for having to manage infrastructure
  - You can get right to deploying your app
- SaaS (more) - Software as a Service
  - Provide full software apps that are run and managed by another party/vendor

Cloud Models



The Shared Responsibility Model - AWS

- AWS Responsibilities (Security OF the cloud):
  - Security of physical infrastructure (*infra*) and network
    - keep the data centers secure, control access to them
    - maintain power availability, HVAC, etc.
    - monitor and maintain physical networking equipment and global infra/connectivity
  - Hypervisor & Host OSs
    - manage the virtualization layer used in AWS compute services
    - maintaining underlying host OSs for other services
  - Maintaining managed services
    - keep infra up to date and functional
    - maintain server software (patching, etc)

The Shared Responsibility Model - Client

- Client Responsibilities (Security IN the cloud):
  - Control of Data/Content
    - client controls how its data is classified, encrypted, and shared
    - implement and enforce appropriate data-handling policies
  - Access Management & IAM
    - properly configure IAM users, roles, and policies.
    - enforce the *Principle of Least Privilege*
  - Manage self-hosted Apps and associated OSs
  - Ensure network security to its VPC
  - Handle compliance and governance policies and procedures

The AWS Global Infrastructure

- Regions - distinct geographical areas
  - us-east-1, us-west 1, etc
- Availability Zones (AZs)
  - each region has multiple AZs
  - roughly equiv to isolated data centers
- Edge Locations
  - locations for CDN and other types of caching services
  - allows content to be closer to end user.

Regions
Coming soon
36 launched Regions
each with multiple Availability Zones
114 Availability Zones
700+ CloudFront POPs
and 13 Regional edge caches

Compute Services

- VM-based:
  - EC2 & EC2 Spot - Elastic Cloud Compute
- Container-based:
  - ECS - Elastic Container Service
  - ECR - Elastic Container Registry
  - EKS - Elastic Kubernetes Service
  - *Fargate* - Serverless container service
- Serverless: AWS Lambda

https://aws.amazon.com/products/compute/

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
  - High-speed cache for datasets stored anywhere
- AWS Backup
  - Fully managed, policy-based service to automate data protection and compliance of apps on AWS

Database Services

- Relational - Amazon RDS, Amazon Aurora
- Key-Value - Amazon DynamoDB
- In-Memory - Amazon MemoryDB, Amazon ElastiCache
- Document - Amazon DocumentDB (Compat with MongoDB)
- Graph - Amazon Neptune

Analytics Services

- Amazon Athena - Analyze petabyte scale data where it lives (S3, for example)
- Amazon EMR - Elastic MapReduce - Access Apache Spark, Hive, Presto, etc.
- AWS Glue - Discover, prepare, and integrate all your data
- Amazon Redshift - Data warehousing service
- Amazon Kinesis - real-time data streaming
- Amazon QuickSight - cloud-native BI/reporting tool

ML and AI Services

- Amazon SageMaker
  - fully-managed ML platform, including Jupyter NBs
  - build, train, deploy ML models

- AWS AI Services w/ Pre-trained Models
  - Amazon Comprehend - NLP
  - Amazon Rekognition - Image/Video analysis
  - Amazon Textract - Text extraction
  - Amazon Translate - Machine translation

Important Services for Data Analytics/Engineering

- EC2 and Lambda
- Amazon S3
- Amazon RDS and DynamoDB
- AWS Glue
- Amazon Athena
- Amazon EMR
- Amazon Redshift

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)
  - ….. So many free services

??