CLF-C01

Overview

AWS Cloud Computing

Cloud Computing is the on-demand delivery of compute power, database storage, applications, and other IT resources through a cloud services platform via the internet with pay-as-you-go pricing

AWS Cloud Services Platform provides rapid access to flexible and low cost IT resources

Pay-as-you-go pricing? Pay only for what you use, when you use it !

3 major serivce model: IaaS, PaaS, SaaS, the differences between them are Functionalityand Tasks' Ownership & Flexibility

Deployment Model

On-Premises, also known as Private Cloud

Resources are deployed in your on-premises DC, using virtualization and resource management tools – VMWare, Hyper-V, OpenStack

Offfers the ability to provide dedicated resources, not split between users or end customers (only your Apps sit on the actual hardware)

You have full control over your infrastructure and are responsible for management and OS patching

Hybrid

Can be an intermediate step, whiel you are on the way to fully migrating to the AWS cloud

a way to connect infra and apps between cloud-based resources and existing resources that are not located in the cloud

The most common method of hybrid deployment is between the cloud and your existing on-premises infra in order to extend or grow your organizations' infra

Cloud

application is fully deployed in the cloud and all components of the application run in the cloud

Applications in the cloud have either been created in the cloud or have been migrated from an existing infra to take advantage of the cloud benefits

Migrating an App from on-prem to cloud is typically called "lift-and-shift"; this refers to taking the App as is, without modifying it, and running it on cloud-native resources

6 Advantages of AWS Cloud Computing

1) Trade capital expense for variable expense

You can now pay only when you consume computing resources, and pay only for how much you consume

No upfront commitment "pay-as-you-use"

2) Benefit from massive economies of scale

You can achieve a lower variable cost than you can get on your own

Because usage from hundreds of thousands of customers is aggregated in the cloud, providers such as AWS can achieve higher economies of scale, which translates into lower pay-as-you-go prices

3) Stop guessing about capacity

Eliminate guessing on your infra capacity needs

While guessing, you often end up either sitting on expensive idel resources or dealing with limited capacity

You can access as much as little capacity as you need and scale up and down as required

4) Increase speed and agility

Reduce the time to make IT resources available to your developers from weeks to just minutes

This results in a dramatic in a dramatic increase in agility for the org, since the cost and time it takes to experiment and develop is significantly lower

New server to PROD time?

5) Stop spending money running and maintaining DCs

You focus on project that differentiate your business, not the infra; let AW take care of the infra! STOP Wasting Money!

AWS will take care of the actual room (DC), power, cooling, racks, servers, cabling, storage, network, security equipment, guards

Focus on your business!

6) Go global in minutes

Easily deploy your app in multiple regions aournd the world with just a few clicks

This means you can provide lower latency and a better experience for your customers at minimal cost

Global Reach: R (Regions) | AZ (Availability Zones)
Low Latency: E (Edge Locations_

AWS Global Infra, 3 building blocks

Regions

An AWS Region is a physical location in the world that consists of multiple (2 or more) Availability Zones

All AWS Regions are completely isolated one from each other: Highest Standards fault tolerance and stability

Regions are isolated one from each other, AZs are isolated one from each other, BUT ... the AZs in the same Region are connected through low-latency links (2 or more)!

AZs (Availability Zones)

Represents one or more discrete data centers, each DC with redundant power, networking, and connectivity, housed in separate facilities

Running your Apps or services in multiple AZs, you can easily achieve high availability, fault tolerance and scalability

This is not possible if running Apps in a single on-prem data center

One Availability Zone = One Data Center

What's inside the "box"?

- Servers
- Networking
- Storage
- Balancers equipment

Edge Locations

Amazon CloudFront is a fast content delivery network (CDN) service that securely delivers data, videos, apps to customers globally with low latency and high transfer speeds

Amazon CloudFront uses a global network of 166 Points of Presense (155 Edge Locations and 11 Regional Edge Caches) in 65 cities across 29 countries)

Edge Locations vs. Regional Edge Caches

CloudFront helps you deliver your web content faster to your end users, thus providing a better user experience

CloudFront Edge Locations bring the web content closer to your viewers and make sure that popular content can be served quickly

CloudFront Regional Edge Caches really help when the content is not popular enough to stay at a CloudFront Edge Location and improve delivery performance for that content

AWS Mgmt Interfaces: AWS provides 3 distinct options in order to interact with the AWS Cloud Platform

AWS Management Console

A graphical user interface (GUI) for accessing a wide range of AWS Cloud services and managing compute, storage, and other cloud resources

A web applications that comprises and refers to a broad collection of software consoles for managing Amazon Web Services

AWS Command Line Interface (CLI)

A unified tool to manage AWS services

With just one tool to download and configure, you can control multiple AWS services from the command line and automate them through scripts

After AWS CLI tool installation, you can begin making calls to your AWS services from the command line

AWS Software Development Kits (SDKs)

A set of tools that allow developers to create software or apps for a specific platform, OS, computer system or device

Using SDKs, you can access and manage AWS services with your preferred development language or platform

AWS Core Services

Billing Alarm for AWS

Identity Access Management (IAM)

AWS IAM is a web service that helps you securely control acces to AWS resources

You use IAM to control who is authenticated (signed in) and authorized (have permissions) to use what resources

The key is that IAM is representated by these two concetps

- Authenticaton
- Authorization

4 key contents

User: a permanent named operator; can be a human or it can be a machine, or another AWS service

Authentication credentials are permanent

Group: a collection of users and usually contains multiple users; a user can belong to multiple groups

Authentication credentials are permanent

Role: an operator too, another authentication method just like a user; a role can be as well a human or another AWS service

Authentication credentials are temporary

Policy Document

Once a user/role is authenticated by AWs, it will be given permissions (authorized) based on policy document(s) that are attached to it

Policy Documents (JSON Format) can be attached to a user, group or role; if policy is attached to group, once a user joins the group, it will inherit the attached policies

JSON – JavaScript Object Notation

Full Picture

A principal (or operator), human or AWS service, makes a request for an action on an AWS resource (API call)

First, the user is authenticated, based on username/password pair or access key ID / secret access key (programmatic access – CLI, API, SDK)

The user's action will be permitted (authorized) based on attached policies

Every API call will be recorded in AWS by CloudTrail

Virtual Private Cloud (VPC)

Elastic Compute Cloud (EC2)

Security Groups (SGs)

Elastic Block Store (EBS)

Simple Storage Service (S3)