



Cloud Computing

Date: 2024/7/29 (Sharwan 14, 2081)

Asmita's Note for AWS Cloud Cohort 1.0 Fellowship

Module 1

Cloud computing is the on-demand delivery of compute power, database, storage, applications, and other IT resources via the internet with pay-as-you-go pricing.

These resources run on server computers that are located in large data centers in different location around the world.

When you use a cloud service provider like AWS, you are using the computers owned by that service provider.

These resources can be used together like building blocks to manufacture solutions that will help you meet your business goals and satisfy technology requirements.

Traditional computing model infrastructure as hardware. Cloud Computing model infrastructure as software.

Cloud service models:

- IaaS (Infrastructure as a service)
- PaaS (platform as a service)
- SaaS (software as a service)

Examples:

- **IaaS:** Facebook (partially), Alibaba Cloud
- **PaaS:** Discord, Alibaba Cloud (PaaS services)
- **SaaS:** Facebook, Instagram, Payment Systems

Cloud computing deployment models

- Cloud
- Hybrid
- On-premises (private cloud)

Advantages of Cloud Computing

- Trade capital expense for variable expense.
- Massive economies of scale
- Stop guessing capacity
- Increase speed and agility
- Stop spending money on running and maintaining data centers
- Go global in minutes

Introduction to Amazon Web Services (AWS)

A web service is any piece of software that makes itself available over the internet and uses a standardized format such as Extensible Markup language (XML) or JavaScript Object Notation (JSON) for request and response of an application programming interface (API) interaction.

'I don't care space'

AWS is a secure cloud platform that offers a broad set of global cloud-based products.

AWS provides you with on-demand access to compute, storage, network, database, and other IT resources and management tools.

AWS offers flexibility.

You pay only for the individual services you need, for as long as you use them.

AWS services work together like building blocks.

Categories of AWS services: analytics, cost management, application integration, customer engagement, database, internet of things, networking and content delivery, robotics, developer tools, game tech, block chain, storage, mobile and many more.

The service you select depends on your business goals and technology requirements.

There are three ways to interact with AWS services

- AWS Management Console
- Command Line Interface (AWS CLI)
- Software Development kits (SDKs): Access services directly from your code (such as Java, Python, and others)

AWS Cloud Adaption framework (AWS CAF)

For any organization to successfully migrate its IT portfolio to the cloud, three elements, that is, people, process and technology, must be in alignment. Business and technology leaders in an organization must understand the organization's current state, target state, and the transition that is needed to achieve the target state so that they can set goals and create processes for the staff.

At the highest level, the AWS Cloud Adoption Framework organizes guidance into six areas of focus called Perspectives. Perspectives consist of sets of capabilities and different responsibilities that are owned or managed by cross-functional teams.

Focus on Business Capabilities	Focus on technical capabilities
Business Perspective	Platform Perspective
People Perspective	Security Perspective
Governance Perspective	Operations Perspective

AWS CAF provides guidance and best practices to help organizations build a comprehensive approach to cloud computing across the organization and throughout the IT lifecycle to accelerate successful cloud adoption.

Stakeholders from Business prospective	Members within the People Perspective	Stakeholders from Governance Perspective	Stakeholders from Platform Perspective	Security Perspective members	Operations team members
Business Managers	Human Resources	Chief Information officer	Chief Technology Officer	Chief Information security officer	IT operations manager
Finance Managers	Staffing	Program Managers	IT managers	IT security managers	IT support managers
Budget owners	People Managers	Enterprise architects	Solutions architect	IT security analyst	
Strategy stakeholders		Business Analysts			
		Portfolio managers			

Module 2 : Cloud Economics and Billing

AWS pricing model : Three fundamental drivers of cost with AWS

Compute	Storage	Data transfer
Charged per hour/second	Charged per GB	Outbound is aggregated and charged
Various by instance type		Inbound has no charge (with some exceptions)
		Charged typically per GB

How do you pay for AWS ?

Pay for what you use Pay less when you reserve Pay less when you use more and as AWS grows

At the end of each month, you pay for what you use. You can start or stop using a service at any time. There are no long term contract required.

Pay for what you use

Pay only for the services that you consume, with no large upfront expenses.

Note:

Upfront expenses are initial costs paid before using a service or product, like purchase prices or setup fees.

Pay less by using more

Volume based discounts: savings as usage increase, tiered pricing for services like Amazon Simple Storage Service (Amazon S3), Amazon Elastic Block Store (Amazon EBS),or Amazon Elastic File System (Amazon EFS) ; the more you use, the less you pay per GB

Pay even less as AWS grows

Since 2006, AWS has lowered pricing 75 times (as of September 2019). Future higher-performing resources replace current resources for no extra charge.

Custom pricing

Meet varying needs through custom pricing. Available for high-volume projects with unique requirements.

AWS Free tier: Free for 1 year for new customers using these services: Amazon VPC, Elastic Beanstalk, AWS identity and access management (IAM)

Total Cost of Ownership

On-premises/Traditional Infrastructure	Cloud /AWS cloud
An on-premises infrastructure is installed on a company's own computers and servers.	A cloud infrastructure is purchased from a service provider who builds and maintains the facilities, hardware, and maintenance staff.

There are several fixed costs, also known as Capital expenses, that are associated with this traditional infrastructure.	A customer pays for what is used.
Capital expenses include facilities, hardware, licenses, and maintenance staff.	Scaling up or down is simple.
Scaling up can be expensive and time consuming.	Costs are easy to estimate because they depend on service use.
Scaling down does not reduce fixed costs.	

How can we identify the best option?

We can identify the best option by comparing the on-premises solution to a cloud solution. Total Cost of Ownership (TCO) is a financial estimate intended to help buyers and owners determine the direct and indirect cost of a product or system.

TCO includes cost of a service and all the costs that are associated with owning the service. In the cloud environment, TCO is used to compare the cost of running an entire infrastructure environment for a specific workload in an on-premises or co-location facility to the same workload running on a cloud-based infrastructure. Companies use this comparison for budgeting purposes or to make decisions about the optimal deployment solution.

Some of the costs associated with data center management includes;

1. Servers costs for both hardware and software and facilities costs to house the equipment,
2. Storage cost for hardware, administration and facilities
3. Network cost for hardware, administration and facilities
4. IT labor costs that are required to administer the entire solution

Note: "facilities cost" typically refers to the costs associated with physical infrastructure such as space, power, and cooling, building security, taxes, maintenance.

AWS Pricing Calculator to estimate monthly costs. You can use this tool to explore AWS services and create an estimate for the cost of your use cases on AWS. AWS pricing calculator helps to estimate monthly cost of AWS services, identify opportunities for cost reduction and model your solutions before building them.

Additional benefit considerations



Hard benefits

- Reduced spending on compute, storage, networking, security
- Reductions in hardware and software purchases (capex)
- Reductions in operational costs, backup, and disaster recovery
- Reduction in operations personnel



Soft Benefits

- Reuse of service and applications that enable you to define (and redefine solutions) by using the same cloud service
- Increased developer productivity
- Improved customer satisfaction
- Agile business processes that can quickly respond to new and emerging opportunities
- Increase in global reach

They also include reductions in hardware



AWS Organizations

AWS organizations is a free AWS service that enables you to consolidate multiple AWS accounts into an organizational tree with each branch representing a department or team.

AWS organizations include consolidated billing and organization security management capabilities.

Key features and benefits

- Policy-based account management
- Group-based account management
- Application programming interfaces (APIs) that automate account management
- Consolidated billing

Security with AWS Organizations

- Control access with AWS identity and Access Management (IAM)
- IAM policies enable you to allow or deny access to AWS services for users, groups, and roles.
- Service control policies (SCPs) enable you to allow or deny access to AWS services for individuals or group accounts in an organizational unit (OU).

Set up AWS Organizations

1. Create Organization with your current AWS account as the primary account.

You can also invite one AWS account to join your organization and create another account as a member account.

2. Create two organizational units(OU) in your new organization and place the member accounts in those OUs.
3. Create service control policies (SCPs), which enable you to apply restrictions to what actions can be delegated to users and roles in the member accounts. A service control policy is a type of organizational control policy.
4. Test your organization's policies. Sign in as a user for each of the roles such as OU1 or OU2 and see how organizational control policies impact account access.

Alternatively, you can use the IAM policy simulator to test and troubleshoot IAM and resource-based policies that are attached to IAM users, groups, or roles in your AWS account.

Different interfaces to access AWS Organizations

1. AWS Management Console
2. AWS command line interface (AWS CLI) tools
3. Software development kits (SDKs)
4. HTTPS Query application programming interfaces (API)

AWS Billing

AWS Billing and Cost Management is a service that you use to pay your AWS bill, monitor your usage, and budget your expenses.

We can also forecast and obtain a better idea of what our costs and usage might be in the future, so that we can plan ahead.

With filtering and grouping functionality, we can further analyze our data using a variety of available dimensions. The AWS cost and usage report tool enables us to identify opportunities for optimization, by understanding your cost and usage data trends and how you are using your AWS implementation.

AWS Technical Support Plans

AWS Support can provide you with an unique combination of tools and expertise based on your current and future use cases. Support was developed to provide the right resources to help you succeed.

- Proactive guidance:

- Technical Account Manager (TAM) : who are designated as your primary point of contact. TAM can provide guidance , architectural review, and continuous ongoing communication to keep you informed and prepared as you plan, deploy, and optimize your solutions. Please note that a Technical Account Manager is only available via the Enterprise Support Plan.
- Best practices:
 - AWS Trusted Advisor : If you want to ensure that you follow best practices in your AWS environment, you can use AWS Trusted Advisor. ATA is an automated service that acts like a customized cloud expert. It is an online resource that checks for opportunities to reduce monthly expenditures and increase productivity.
- Account Assistance
 - AWS Support Concierge : It is a billing and account expert who provide quick and efficient analysis on billing and account issues. The concierge addresses all non-technical billing and account level inquiries.

Support Plans

AWS support offers four support plans:

- Basic Support : free of charge, resource center access, service health dashboard, product FAQs, discussion forums, and support for health checks
- Developer Support : Support for early development on AWS
- Business Support : customers that run production workload
- Enterprise Support : Customers that run business and mission-critical workloads

Case severity and response times

- Critical
- Urgent
- High
- Normal
- Low

Module 3 : AWS Global Infrastructure Overview

Section 1 : AWS Global Infrastructure

- 22 current AWS regions, a few regions that will become available soon, including Milan, Cape Town and Indonesia as of October 2019.

- An AWS region is a physical geographical location with one or more availability zones.
- An availability zone in turn consists of one or more data centers.
- To achieve fault tolerance and stability AWS regions are isolated from one another.
- Resources in one region are not automatically replicated to another region. Therefore when you store data in a specific region, it is not replicated outside of that region. Data replication across Regions is controlled by you if your business or project needs require it.

Selection a Region

Determine the right Region for your services, applications, and data based on these factors:

- Data governance, legal requirements
- Proximity to customers (latency) : cloumping
- Services available within the Region
- Costs (vary by Region)

Availability Zones (AZ given by me just for shortcut)

An AWS region has multiple isolated locations that are known as Availability Zones. Each AZ provides the ability to operate applications and databases that are more highly available fault tolerant and scalable, than they would be in a single data center.

AZ can include multiple datacenters, typically three. And at full scale, they can include hundreds of thousands of servers.

They are fully isolated partitions of the database AWS Global Infrastructure.

Availability zones are separated from other AZ by several kilometers. They are interconnected with other AZ by using high-speed private networking.

You are responsible for selecting AZ where your system will reside. Systems can span multiple AZs and AWS recommends replicating across several AZs for resiliency. You should design your systems to survive the temporary prolonged failure of single AZs in the event of a disaster.

AWS Data centers

The foundation for AWS infrastructure is the data centers. Data centers are where the data resides and data processing occurs.

Amazon operates state-of-the art highly available data centers.

AWS data centers are designed for security. Data center locations are not disclosed and all access to them is restricted.

A data center typically has 50,000 to 80,000 physical servers.

Points of Presence

Amazon CloudFront is a content delivery network used to distribute content to end-users in order to reduce latency.

Amazon Route 53 is a Domain Name System , DNS service where request going to either one of these services will be routed to the nearest Edge location automatically in order to lower latency.

AWS points of presence are located in most of the major cities across 30 countries around the world.

To deliver an improved near real-time user experience, points of presence continuously measure Internet connectivity, performance, and computing to find the best possible way to route the traffic requests.

AWS Infrastructure features

- Elasticity and scalability
- Fault-tolerance
- High availability

Section 2 : AWS foundational Services

The AWS global infrastructure can be broken down into three elements

1. Regions
2. Availability Zones
3. Points of presence, which include edge locations

This platform provides broad set of services such as networking, storage, compute services, and databases.

There are 23 different product or services categories, and each category consists of one or more services.

Widely used services categories :

- Computer
- Cost Management
- Database
- Management and Governance
- Networking and Content delivery
- Security , Identity and compliance
- Storage

Storage Services:

Storage Service Category also known as Amazon Simple Storage Service, also known as Amazon S3, which is an object storage service that offers scalability data availability, security and performance.

Amazon Elastic Block Store, also known as Amazon EBS, which is a high performance block storage designed for use with Amazon EC2 for both throughput and transaction-intensive workloads.

Amazon Elastic file system, also known as Amazon EFS. This storage service provides a scalable fully managed elastic network file system, also known as NFS, for use within AWS cloud services and on-premises resources.

Amazon Simple Storage Services Glacier, which is a secure, durable and extremely low cost AWS S3 cloud storage class for data archiving and long-term backup.

AWS Computer Services

- Amazon Elastic Computer Cloud (Amazon EC2)
- Amazon EC2 Auto scaling
- Amazon Elastic Container Service (Amazon ECS)
- Amazon Elastic Container Registry Service (Amazon ECR)
- AWS Elastic Beanstalk
- AWS Lambda
- Amazon Elastic Kubernetes Service (Amazon EKS)
- AWS Fargate

Database service category

- Amazon Relational Database Service (Amazon RDS)
- Amazon Aurora
- Amazon Redshift
- Amazon DynamoDB

Networking and content delivery service category

- Amazon Virtual Private Cloud (Amazon VPC)
- Elastic Load Balancing
- Amazon CloudFront
- AWS Transit Gateway

- Amazon Route 53
- AWS Direct Connect
- AWS VPN

Security , identity, and compliance service category

- AWS Identity and Access Management (IAM)
- AWS Organizations
- Amazon Cognito
- AWS Artifact
- AWS Key Management Service
- AWS Shield

AWS cost management service category

- AWS Cost and Usage Report
- AWS Budgets
- AWS Cost Explorer

Management and governance service category

- AWS Management Console
- AWS Config
- Amazon CloudWatch
- AWS Auto Scaling
- AWS Command Line Interface
- AWS Trusted Advisor
- AWS Well-Architected Tool
- AWS CloudTrail

That's all for today: 2024/08/02 (2081 Sharwan 18)