Chapter 1:
Introduction to AWS

What is Cloud Computing

- AWS is the on-demand delivery of IT resources and Services

- Pay as you go

- Quick access, scalable, highly reliable, low cost

- No upfront cost and no HW maitinance cost

- You can have:

  - Servers

  - Storage

  - Databases

  - Broad set of application services

Advantages ... Continues

- Ability to to quickly and efficiently reconfigure the computing environment to adapt to business requirements.

- Capacity can easily scaled up and down as business needs

- Services can be taken offline or shutdown as business demands.

Advantages ... Continues

1. **Variable vs Capital:** Only invest when the need is apparent. No need to guess the demand when purchasing hardware anymore.

2. **Economies of Scale:** Achieve lower prices as a result of thousands of customers sharing the same resources.

3. **Stop Guessing Capacity:** When purchasing hardware prior to deploying an application, you often end up with limited capacity or extra unnecessary hardware. With CC you can scale up and down almost instantly.

4. **Increased Speed and Agility:** New IT resources are just a few clicks away.

Advantages

5. Focus on Business Differentiators: CC saves tons of time that is spent on racking, stacking, powering, and managing servers. Saves resources and time spent on managing data centers. More time to focus on more important aspects of your business.

6. Global in Minutes: Go global in minutes. Deploy applications around the globe to provide redundancy and deliver lower latency.

Cloud Computing Deployment Models

Two Primary Deployment Models

- All in cloud based deployments

- Hybrid Deployments

All in Cloud Deployment: An all-in cloud based application is fully deployed in the cloud, it is either created in the cloud or migrated to the cloud from an existing infrastructure.

Hybrid Deployment: Connects infrastructure and applications between cloud based resources and existing resources. (Extending a organization's existing infrastructure while connecting cloud resources to internal systems)

Security and Compliance

Cloud Security is the number one priority.

Compliance is necessary between AWS and its customers. Customers are responsible for setting up their environment in a secure and controlled manner. Customers also need to govern the overall network.

Customers retain complete control and ownership over all the data on AWS.

Global Infrastructure

- The AWS Cloud infrastructure is built around Regions and Availability Zones.

- A Region is a physical location in the world where there are multiple Availability Zones.

- Availability Zones consist of one or more discrete data centers, each with redundant power, networking and connectivity, housed in separate facilities, and separately powered/connected to a power grid.

- You can choose to replicate data between availability zones, or even regions for lower latency access across the globe.

AWS offers Various Services

Enterprise Applications	Virtual Desktops			Sharing and Collaboration	
Platform Services	**Databases** Relational NoSQL Caching	**Analytics** Hadoop Real-Time Data Warehouses Data Workflows	**App Services** Queuing Orchestration App Streaming Transcoding Email Search	**Deployment and Management** Containers DevOps Tools Resources Templates Usage Tracking Monitoring and Logs	**Mobile Services** Identity Syns Mobile Analytics Notifications
Foundation Services	Compute (VMs. Auto Scaling and load Balancing)		Storage (Object, Block and Archive)	Security and Access Control	Networking
Infrastructure	Regions	Availability Zones		Content Delivery Networks and Points of Presence	

Accessing The Platform

To access AWS cloud services you can use:

- AWS Management Console: Web Application(User Interface)

- AWS Command Line (CLI): Unified tool that is downloaded and configured (You can control multiple services from the command line and automate them through scripts)

- AWS Software development kits(SDKs): Provide Application Programming Interface that interacts with web services to make up the AWS platform. Support multiple programming languages/platforms.

Compute and Networking Services

Amazon EC2

AWS provides a variety of compute and networking services that can be used with the storage, database, and application services to provide a complete solution.

Amazon Elastic Compute Cloud: Amazon EC2

A web service that provides the ability to compute and control usage.

Resizing capacity, make use of resources to build and host softwares,

ability to select from a variety of OS, memory, CPU, storage etc.

AWS Lambda

- AWS Lambda is a zero administration compute service that lets you run code.

- AWS Lambda executes your code only when needed and scales automatically. Pay for what you use.

- Zero administration.

- Languages that AWS Lambda supports are currently Node.js, Java, C# and Python.

- You can use AWS Lambda to run your code in response to events, or changes elsewhere

Auto Scaling

- Auto scaling allows to scale EC2 capacity according to the condition and workload. It will help in optimizing costs & maintain application availability and ensure the required amount of EC2 instances are running.

- Auto scaling is good for applications that have stable demand patterns and also for applications that may have hourly demand changes.
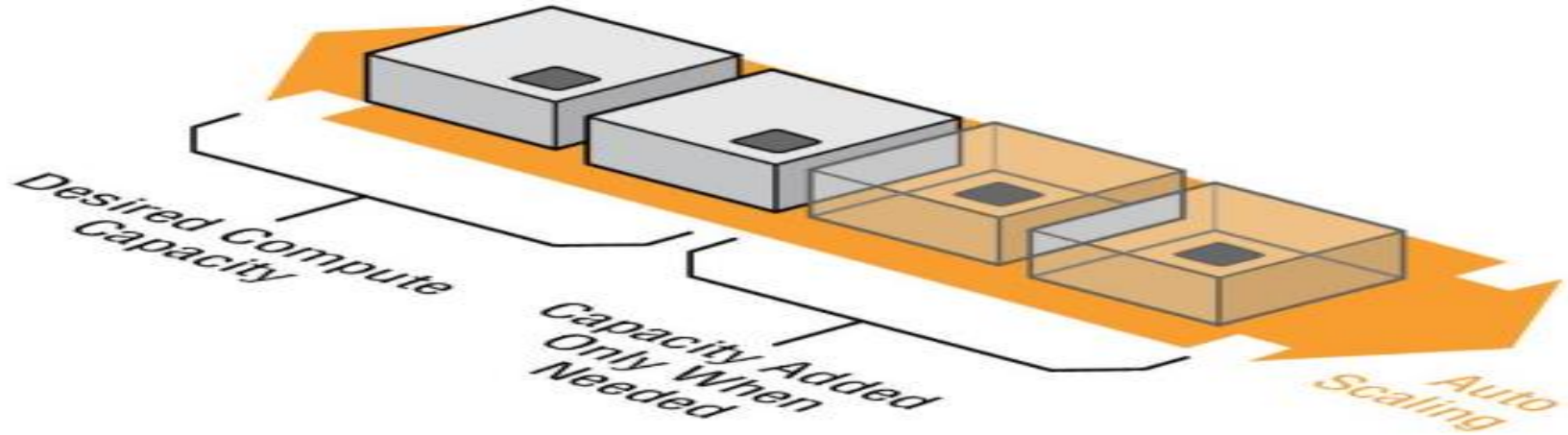
Elastic Load Balancing

- Elastic Load Balancing automatically distributes incoming traffic across multiple EC2 instances.

- Greater level of fault tolerance

- A lot easier than setting up load balancers yourself

Amazon Elastic Beanstalk

- Fastest and simplest way to get a web application up and running on AWS.

- Just upload your application code.

- AWS elastic beanstalk takes care of resource provisioning, load balancing, Auto scaling, and monitoring.

- Organizations retain full control over underlying AWS resources powering the application

- Beanstalk provides support for various platforms, example Java, Python.

VPC Virtual Private Cloud

Amazon VPC allows customers to run and manage a logically isolated section of the cloud where they can launch AWS resources of their need.

Customer has complete control on:

- Private IP address range to be used

- Creation of Subnets

- Configuration of route tables and network gateways.

AWS Direct Connect

- Allows organizations to establish a dedicated network connection from their network to AWS.

- Provides private connectivity between an organization's network and AWS.

- This can reduce network costs, increase bandwidth throughput and provide a more consistent network experience than internet based VPN connections.

Amazon Route 53

- Amazon Route 53 is a highly available and scalable cloud Domain Name System (DNS) web service.

- Gives developers and businesses an extremely reliable and cost effective way to route end users to Internet applications by translating names like www.example.com into the numeric IP addresses like 192.0.2.1 that computers use to connect to each other.

- Amazon Route 53 is fully compliant with IPv6 as well.

Amazon Simple Storage (Amazon S3)

- Provides developers and IT teams with highly durable and scalable object storage.

- Handles virtually unlimited amounts of data and large number of concurrent users

- Cost effective object storage for use case such as backup and recovery, nearline archive, big data analytics, disaster recovery, cloud applications, and content distribution.

- Any number and type of object can be stored including HTML pages, source code files, image files, and encrypted data and can be accessed using HTTP-based protocols.

Amazon Glacier

- Secure, durable, and extremely low cost data archive/long-term backup.

- In order to keep cost low, Glacier is optimized infrequent access to data, where a retrieval time of several hours is suitable.

- A range of options exist when it comes to retrieval time, from minutes to hours.

EBS Elastic Block Storage

- Provides persistent block level storage volumes for use with Amazon EC2 instances.

- Each Amazon EBS volume is automatically replicated within its availability zone, offering high availability and durability and protect from failure.

- Delivering consistent and low latency performance, EBS provides the disk storage needed to run a variety of workloads.

AWS Storage Gateway

- AWS Storage Gateway connects an on-premises software appliance with cloud-based storage to provide seamless integration with data security features between your on-premises IT environment and the AWS storage infrastructure.

- It provides very low latency performance, by maintaining a cache of frequently accessed data on premises, while storing all your data encrypted on S3 or Glacier.

Amazon CloudFront

- Is a content delivery web service.

- Integrates with other AWS cloud services allowing developers and organizations to distribute content to users across the globe with low latency, high data transfer speeds and no minimum usage commitments.

- Can be used to deliver your entire website to end users using a global network of edge locations.

Data Services

RDS Relational Database Service

Amazon RDS is a web service that makes it easier to set up, operate, and scale a relational database in the cloud.

It offers cost-efficient, resizeable capacity for an industry-standard relational database and manages common database administration tasks, such as:

- Backups

- Patching

- Monitoring

- Replication

Amazon DynamoDB

- Is a fast and flexible NoSQL database service for all applications that need consistent, single-digit millisecond latency at any scale.

- Fully managed database

- Supports both document and key/value data models.

- Great fit for mobile, web gaming, ad tech, internet of things, and other applications.

Amazon RedShift

- Amazon Redshift is a fully managed, petabyte-scale data warehouse service that makes it simple and cost-effective to efficiently analyze all your data

It's a Relational Database which:

- Leverages columnar storage technology (normally used by nosql databases optimized for better Read and Write Performance) to improve IOPS and parallelizing queries across multiple nodes (fast query performance).

- Automates most of the common administrative tasks associated with provisioning, configuring, and monitoring a data warehouse.

ElastiCache

- Amazon ElastiCache is a web service that makes it easy to set up, manage, and scale distributed in-memory cache environments he cloud.

- Allows organizations to retrieve information from fast, managed, in-memory cache, instead of lower, disk based databases.

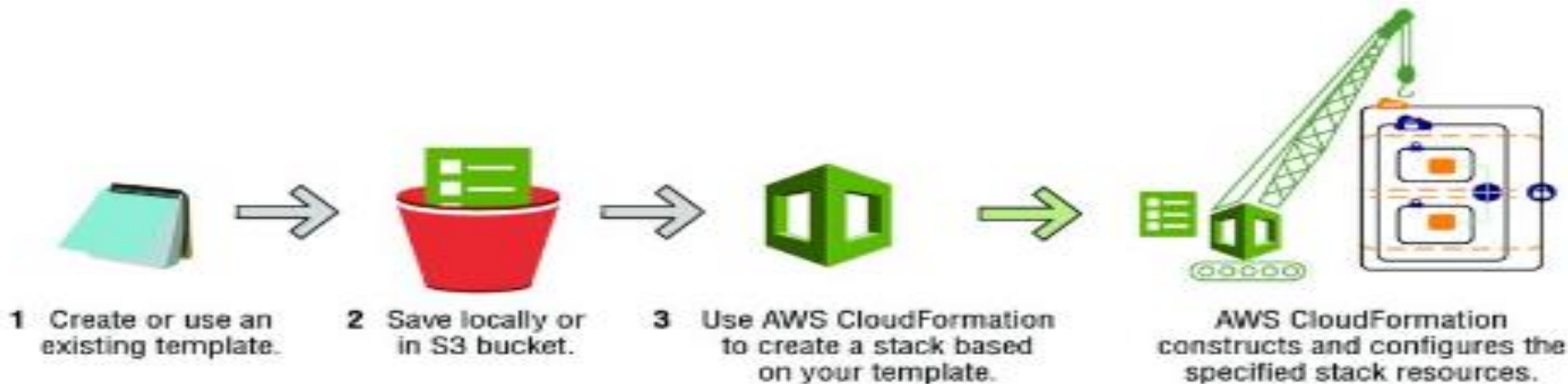- Amazon ElastiCache supports both Memcached and Redis cache engines For more info on Redis https://redis.io/

Management Tools

CloudWatch

- Amazon CloudWatch is a monitoring service for AWS cloud resources and the applications you run on AWS.

- You can use Amazon CloudWatch to collect and track metrics, collect and monitor log files, set alarms, and automatically react to changes in your AWS resources.

- Monitor custom metrics generated by your applications and services, and any log files your applications generate.

- You can use Amazon CloudWatch to gain system-wide visibility into resource utilization, application performance, and operational health. You can use these insights to react and keep your application running smoothly.

CloudFormation

- Allows developers and systems administrators to effectively create and manage a number of AWS resources, provisioning and updating them in orderly and predictable fashion.

- Templates can be submitted to CloudFormation and the service will take care of provisioning and configuring these resources.

1 Create or use an existing template.
2 Save locally or in S3 bucket.
3 Use AWS CloudFormation to create a stack based on your template.

AWS CloudFormation constructs and configures the specified stack resources.

CloudTrail

- AWS CloudTrail is a web service that records AWS API calls for your account and delivers log files to you.

The recorded information includes

- the identity of the API caller

- the time of the API call

- the source IP address of the API caller

- the request parameters

- and the response elements returned by the AWS service.

AWS Config

- Provides organizations with an AWS resource inventory, configuration history, and configuration change notifications to enable security and governance.

Capabilities with AWS Config

- Discover existing AWS resources

- Export an inventory of their AWS resources with configuration details

- Determine how a resource was configured at any time

These capabilities allow security analysis, resource change tracking, and troubleshooting.

Security and Identity

IAM Identity and Access Management

- Allows organizations to securely control access to AWS cloud services and resources for their users.

Capabilities:

- Create and manage AWS users and groups

- Use permission allow or deny access to AWS resources

KMS Key Management Service

- Is a management service that allows organizations to create and control encryption keys that are used to encrypt their data.

- Uses hardware security models (HSMs) to protect the security of your keys.

- It integrates with other AWS cloud services to protect data stored with these services.

AWS Directory Service

- Allows organizations to set up and run Microsoft Active Directory on AWS or connect your AWS resources to an existing on premises AWS Microsoft Active Directory.

Capabilities:

- Manage user and groups

- Provide single sign on to application to services

- Amazon EC2 instances join domains

- Simplify the deployment and management of cloud based Linux and Microsoft Windows workloads.

AWS Certificate Manager

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

- SSL/TLS certificates are used to secure network communications and establish the identity of websites over the Internet.

- Auto Renew and quick setup saves time.

- You can provision an SSL/TLS certificate and deploy it with your ELB, Amazon CloudFront distribution or an API in API Gatewy

- You pay only for the AWS resources you create to run your application. Certificates are free.

WAF Web Application Firewall

- WAF helps protect web applications from common attacks that could affect application availability, compromise security, or consume excessive resources.

- WAF gives organizations control over which traffic to allow or block to their web applications by defining customizable web security rules.

Application Services

AWS API Gateway

- An **API** is a software intermediary that allows two applications to talk to each other

- Amazon API Gateway is a fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs at any scale.

- Quickly setup in AWS Management Console, APIs act as the "front door" for applications to access data, business logic, or functionality from your back-end services, such as workloads running on Amazon EC2, code running on AWS Lambda, or any Web application.

- Amazon API Gateway handles all the tasks involved in accepting and processing up to thousands of concurrent API calls, traffic management, authorization and access control, monitoring, and API version management.

Elastic Transcoder

- Allows media transcoding in the cloud - Convert media files from one type to another

- Highly scalable and cost effective way for developers or organizations to convert or transcode media files from their source formats into versions that will playback on devices.

  Ex: Phones, Tablets, Etc...

SNS Simple Notification Service

- Is a web service that coordinates and manages the delivery of messages in the form of publish/subscribe notifications to recipients.

Two types of clients publishers and subscribers (AKA producers and consumers)

- Publishers communicate asynchronously with subscribers by sending a message to a topic, which is Logical Access Point and Communication Channel

- Only subscribers that are subscribed to the topic, receive the message over one of the supported protocols.

SES Simple Email Service

- Amazon Simple Email Service (Amazon SES) is an email service that organizations can use to send emails, marketing messages, or any other type of content to their customers.

SES can also be used:

- to receive messages which can be stored in an S3 bucket

- or publish notifications to Amazon SNS

SWF Simple Workflow Service

- Helps developers build, run, and scale background jobs that have parallel or sequential steps.

- Fully manages state tracker or task coordinator on the cloud.

- It is important to track the state of processing and to provide the ability to recover or retry if a task fails.

- It is vitally important to track the state of processing, if an application's steps take more than 500 milliseconds to complete, as well as to retry and recover if a task fails.

SQS Simple Queue Service

- Amazon Simple Queue Service (SQS) is a fast, reliable, scalable, fully managed message queuing service.

- You can use Amazon SQS to transmit any volume of data, without losing messages or requiring other services to be always available.