

An Overview of AWS

This section identifies the core AWS areas.

When you think about Amazon Web Services (AWS) think IaaS & PaaS. AWS offers both.

IaaS allows organizations to utilize AWS instead of owning and operating your own datacenter. You can simply rent VMs or physical servers from AWS.

The Platform as a Service (PaaS) on the other hand removes the need for your organization to manage the underlying platforms like a database, streaming services, etc.

This allows you to focus on the deployment and management of your core applications and not worry about the IaaS and PaaS layers.

It, in turn, gets organizations to be more efficient and focused as you don't need to worry about resource procurement, capacity planning, software maintenance, patching, or any of the other undifferentiated heavy lifting involved in running your application.

The Amazon Web Services Infrastructure consists of 4 primary areas which are a combination of IaaS & PaaS :-

1. Compute (EC2, LightSail, ECS, Lambda, Batch)
2. Storage (EBS, EFS, S3, Glacier, Storage Gateway, Storage Migration Services)
3. Database (RDS, Redshift)
4. Network (CloudFront, VPC, Direct Connect, Load Balancing, Route 53)

In this course we will go into the details of the various services that would fall into one of these 4 main buckets.

Compute



Compute(EC2): This is where you create/deploy your own virtual machine. At AWS you have a wide variety of compute instances you can choose from. This ranges from the type of operating system you would choose to the RAM or CPU you would want your compute instance to have.



Elastic Container Services(ECS): Are used to run and manage your Docker containers. Think about this as something like a managed Kubernetes service.



Light Sail (VPS Service – Virtual Private service): AWS Lightsail launches virtual private servers, which are VMs with individual operating systems but restricted access to physical server resources.



Lambda: Is where you upload a function on to AWS and you pay every time the function is executed or called.

You do not need to think about managing the OS or the VM. Lambda does it all for you.

e.g. Think about a Tax calculator. In the traditional model, you would write the code. Procure a virtual machine to deploy your code to the VM. Then you need to maintain, manage your deployment. However, with Lambda you don't need to pay or manage the VM. Amazon does this for you.



Batch: Batch computing is used for Batch processing. AWS Batch dynamically provisions the optimal quantity and type of compute resources (e.g., CPU vs memory optimized instances) based on the volume and specific resource requirements of the batch jobs submitted. It also manages the batch process, i.e re-starting jobs that fail, scheduling jobs, etc,

Storage



Simple Storage Service (S3). It is an Object / Bucket type of storage.



Network Attached Storage (EFS): In the Elastic file System you upload your files to an EFS and then mount that on to multiple virtual machines



Glacier: Is a storage service used for Data archival. Primarily used to store data that you do not need to use right away. The cost of storage on Glacier is significantly lower.

It takes 3 – 5 hours to restore from Glacier. Costs - \$0.01 per gigabyte, per month.



Snow Ball: Is used to transport a large amount of data on to AWS or to take it out of AWS and move it to your data center.

AWS ships you a hardware device that you plug into your data center and then upload your encrypted data on to the snowball (Hardware device). You will then ship it to AWS and they would upload your data on to AWS. This way you do not need to move your data through the internet as it might take months to move petabytes of data on to AWS.



Storage Gateway: VM you install in your data center and this replicates data Back into S3. This is used when you have an on-premises data-center and would like to replicate the data on to AWS. Once you have a Storage gateway setup you can replicate to S3 on AWS.

Database

1. RDS – Relational Database Service
2. Aurora
3. DynamoDB
4. Neptune
5. ElasticCache



Relational Database Service (RDS): Provides cost-efficient and resizable capacity while managing time-consuming database administration tasks,

freeing you to focus on your applications and business.

Amazon RDS gives you access to several familiar database engines, including Amazon Aurora, MySQL, PostgreSQL, MariaDB, Oracle, and SQL Server. This means that the code, applications, and tools you already use with your existing databases can be used with RDS.

RDS automatically patches the database software and backs up your database, storing the backups for a user-defined retention period and enabling point-in-time recovery.

Aurora is an AWS proprietary relational Database that is compatible with MySQL and PostgreSQL it combines the performance and availability of high-end commercial databases with the simplicity and cost-effectiveness of open source databases.



DynamoDB is a nonrelational database i.e a NoSQL DB. It is a Managed Service, i.e you do not need to tune or manage it in any way, AWS does this for you.

Neptune is a fully managed graph database. Relationships are first-class citizens in graph databases, and most of the value of graph databases is derived from these relationships. Graph databases use nodes to store data entities, and edges to store relationships between entities

A graph in a graph database can be traversed along specific edge types or across the entire graph. In graph databases, traversing the joins or relationships is very fast because the relationships between nodes are not calculated at query times but are persisted in the database. Graph databases have advantages for use cases such as social networking, recommendation engines, and fraud detection, when you need to create relationships between data and quickly query these relationships.





ElasticCache offers fully managed Redis and Memcached. Seamlessly deploy, operate, and scale popular open source compatible in-memory data stores.