The Flash Guide to AWS

A succinct, comprehensive and practical companion text designed for students and young professionals getting started with AWS.

By: Sujal Thapa

Dedicated to my parents,

Sharmila Homagai and Raj Kumar Thapa

Preface

This document is intended for educational purposes. It is designed as a study companion guide to aid individuals getting started with AWS, especially those preparing for AWS Practitioner and Associate level examinations. While this guide provides a succinct yet comprehensive overview of key AWS concepts and services, it is not intended to replace primary sources of information such as the official AWS documentation.

The goal of this document is to simplify the learning process and make the daunting task of diving into the AWS ecosystem more manageable, especially for those who may not have a background in IT. It serves as a quick reference and a supplemental study aid, providing a structured summary of crucial topics.

Some of the images included in this document have been sourced from external web platforms. There was no intention to infringe on copyright; and have been utilized while adhering to Creative Commons licensing guidelines. These images were chosen because they were the best infographics available. The use of these images is solely to enhance the educational value of this document.

This guide covers various AWS services, categorized into sections and chapters for the ease and convenience of the reader. From compute and storage services to security and machine learning, it aims to provide a broad overview of the major AWS services and develop a certain level of theoretical sophistication in the reader.

Though exact sources have not been mentioned throughout the book, all information presented within it can be verified by searching through the official AWS Documentation for the service and/or service suite in question.

INTRO
TO
AWS

Cloud Computing

In this section we will discuss what cloud computing is, how it came to be, and the varying methods in which we take advantage of it.

Cloud Computing: A Brief History

Computers have become an integral part of businesses environments around the world, and different businesses have resorted to different methods of acquiring computers and computer resources. However, computers are a relatively recent invention and their current ubiquity was obviously not always the case.

In the beginning, computing was seen mostly as a tool for large governments and universities, with private companies adopting computing technology mostly as a novelty, something that could aid their worker's productivity and give them a potential competitive advantage against less “modern” competitors.

Over time though, as the adoption of computers became more widespread, technology became not only a competitive advantage against other companies but a competitive necessity, an essential resource without which companies could simply not be as productive as the rest of the economy.

Thus, as organizations became more reliant on computer-based workflows, they built larger and larger computers often with very specialized networking, security, computing, software and hardware resources: the very first **data centers**.

These data centers were often located either near or exactly on the premises of corporations offices, occupying expensive land and were managed by an in-house team of IT technicians and electricians, both of which did not come cheap. And that is not to mention the cost of electricity accrued by the operation of the computers and the cooling required to maintain them.

Referred to as an **on-premises cloud environment**, data centers like the one mentioned above were brilliantly secure and performant, but were a huge drain on the business's financials and a pain to manage. While many businesses were willing to front the costs associated with these private clouds, there were many that sought other options: enter **Cloud Computing**.

Cloud Computing: What it is, Tradeoffs and Characteristics

In 1961, MIT Professor John McCarthy said in a now-famous speech that someday "Computing can be sold as a utility, like water and electricity", an idea that was realized with the advent of cloud computing platforms, platforms that allow us to lease IT infrastructure and software solutions from third-party providers, usually over the internet.

These platforms often utilize pay-as-you-go pricing models, enabling flexible and cost-effective access to computing resources and organizations that had so far been maintaining extensive arsenals of hardware and software resources gave those arsenals up for the convenience and cheapness of renting their infrastructure from companies like AWS, Azure, Salesforce and Google.

Though this period of mass transition from on-premises environments to cloud computing is seen mostly as a net-positive, there were some trade-offs to the cloud. Firstly, the ability of organizations to take advantage of the cloud was dependent on the internet, and therefore necessitated organizations to adopt and invest in high speed internet connections.

And secondly, since organizations went from owning to simply renting resources under cloud computing, they had to relinquish a great deal of the control over the underlying infrastructure and software in the process, especially when compared to the degree of control offered by on-prem cloud environments.

That second point actually was and continues to be a point of particular frustration for cloud engineers, who often have to obey illogical rules and protocols within AWS and other cloud platforms simply because that is the way the parent company has decided things should work.

Variations of this Simpsons meme is extremely popular among Cloud engineers online

Engineers and businesses like the comparatively greater control that on-prem environments offer, but the price tag associated with them is often enough to make them unattractive. Well, this has led to many organizations adopting a sort of third way, a middle-ground between the two: **Hybrid Cloud Environments**.

Hybrid cloud environments combine a standard data center with an outsourced cloud computing platform. For many organizations, the hybrid cloud is the perfect migration to the cloud. In a hybrid architecture, the organization can run its applications and systems in its local data center and offload part of the computing to the cloud. This provides an opportunity for the organization to leverage its investment in more traditional data centers while simultaneously being able to take advantage of modern cloud solutions.

Now, depending on what exactly is being rented/leased from the platform, cloud computing services are typically categorized into three main types:

1. **Infrastructure as a Service (IaaS):** Provides virtualized hardware resources for the operation of digital applications/products, such as virtual machines, storage, and networks.

2. **Software as a Service (SaaS):** Delivers software applications, often related to extremely specific use-cases, and often on a subscription basis.

3. **Platform as a Service (PaaS):** Provides both hardware and software tools, typically make the hosting of application development and/or testing more convenient and simpler.

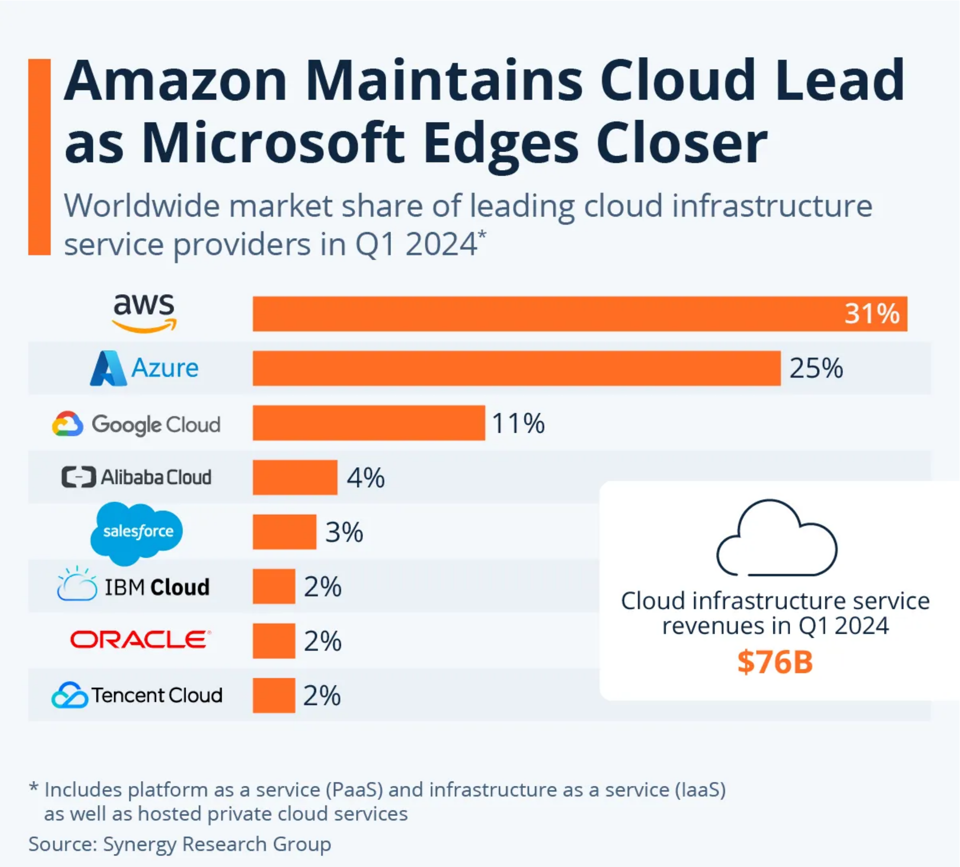
All three categories of cloud computing services, and their examples will be discussed in future sections of the book. The short descriptions above were written for the express purpose of making it clear to the reader that cloud computing and all of its constituent services amount to nothing more than renting agreements for hardware and/or software resources.

Cloud computing is therefore nothing complicated to understand, it is not quantum physics, it is not advanced mathematics and it is not rocket science, it is just **rent**. A fundamental practice that humanity has been performing for thousands upon thousands of years.

The Big Three

Though there are actually many cloud providers, the cloud computing market is mostly dominated by three large players, sometimes termed as the big three: **Amazon Web Services (AWS)**, **Microsoft Azure** and **Google Cloud Platform (GCP)**.

The big three services cumulatively control about 65% of the entire worldwide market for cloud computing with AWS alone accounting for a little more than 30% of the market all by itself at the time of writing. An infographic on the global cloud computing market and the largest players within it is provided below:



Amazon Maintains Cloud Lead as Microsoft Edges Closer

Worldwide market share of leading cloud infrastructure service providers in Q1 2024*

* Includes platform as a service (PaaS) and infrastructure as a service (IaaS) as well as hosted private cloud services

Source: Synergy Research Group

The bulk of this book deals with AWS, the cloud provider with the single-largest industry market share and the services indigenous to it but this does not mean that I endorse blindly relying on the service for business operations. In fact, I am highly against it.

A case for why companies should almost never rely on a single cloud provider and easy plug-and-play alternatives for all of the major AWS services, including offerings from both other Big 3 companies as well as other third-party providers will be considered and delved into the final portion of the book.

Readers are therefore to treat the concluding chapter of the book as a mandatory read lest they limit their range of knowledge of cloud computing to AWS, which though quite dominant in the industry, is only one of a myriad of cloud providers.

As such, the reader may progress through the book without worry of building a very Amazon-centric base of knowledge, as the concepts and type of services discussed throughout the book can be generalized to encompass the other cloud providers and it should be little hassle for someone comfortable with AWS to make the switch to say, Microsoft Azure or GCP.

AWS

Amazon Web Services is the largest and most widely adopted cloud provider in the world, used by everyone from the smallest of startups and hobbyists to well, the United States government and 90% of Fortune 500 companies.

Originally released to the public in July 2002, AWS (then called Amazon.com web services) started out as a service that allowed to incorporate information from Amazon.com into their own websites. The unexpectedly large positive response to the service from developers took Amazon by surprise and by fall 2003, ideas for expansion of Amazon.com web services into a platform for developers to rent database, storage and compute infrastructure had already been put in place.

These ideas would finally come into fruition starting from September 2006, with the release of **S3** and subsequently, **EC2** and **RDS**. The first generation of AWS Cloud Computing offerings. (All mentioned services will be discussed in detail in future sections)

It has since then grown into a behemoth of a platform with over 200 services (most of which are fully-featured and have built quite comprehensive ecosystems in and of themselves) and a massive global geographical presence, with infrastructure in all the seven continents serving 245 countries and territories. The dominance that AWS has over the cloud computing industry will be further explored in the next section.

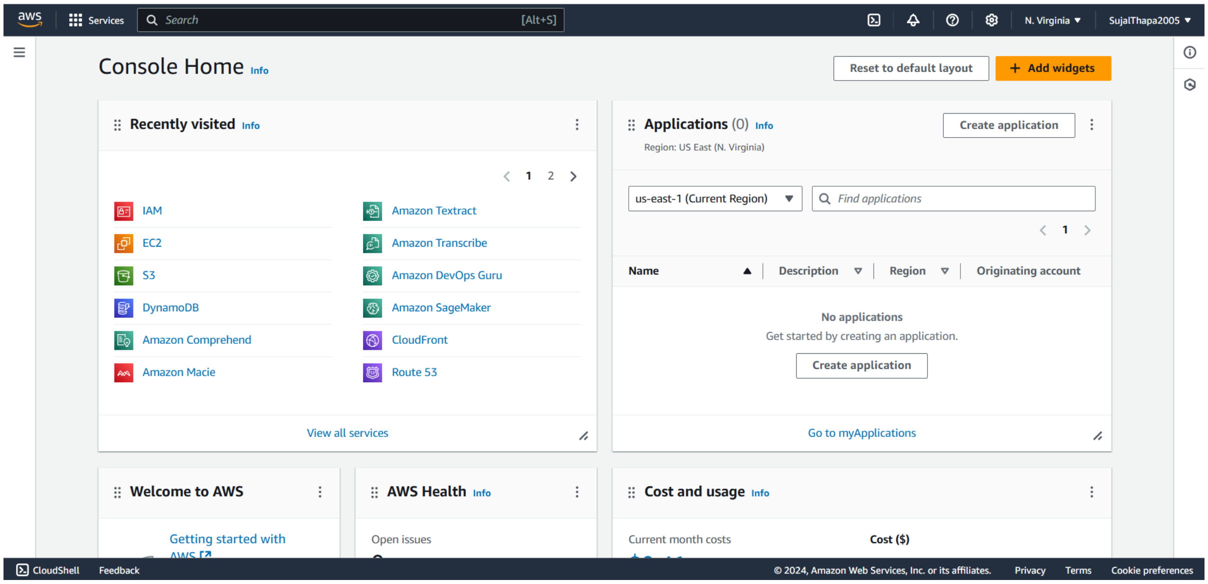
AWS is also obviously the primary subject of this book and honestly, diving into the ecosystem can be a psychologically daunting task, especially for those not from an IT or computing background. But it is precisely in order to make that task seem less daunting that this book has been written.

Connecting to AWS

Now there are three main interfaces by which the reader may utilize and administer resources in AWS: **The AWS Management Console**, **The Command-line interface** and **SDKs**. A more in-depth look at all three is given below:

The AWS Management Console

The main suggested method of interacting with AWS for people not of a technical persuasion, the AWS Management Console is a simple-to-use web interface provided to us by AWS in order to both create and manage AWS infrastructure and its related components. An image of how the Console Home page should look like for a relatively new AWS account is given below:



No applications
Get started by creating an application.

Having an account with AWS is a pre-requisite for logging in, a process that will yes, require entering the reader's credit card details. This may sound scary, especially for readers from the third world but do not be afraid, Amazon charges you a minimal amount and the information can be removed from your account once it is created. (I personally know many brilliant people who have refused to make an AWS account simply because they will not enter their credit card information digitally).

As the reader progresses through the book, it would be prudent for them to log in to console, search for, and play around with any and all services they find interesting. That is after all, the ultimate aim of the book, to make the reader comfortable understanding the cloud and to get them started building real solutions using AWS.

AWS Command Line Interface

The preferred method of handling AWS resources for people familiar with the Linux operating system, the AWS Command Line Interface (or CLI) is a text-based user interface (like Command prompt, Windows Powershell etc) enables us to manage infrastructure via Linux commands and JavaScript Object Notation (JSON) scripting.

It is more efficient than using the console but also requires administrators with a wider set of knowledge and training, and is ideal organizations that already aware of what exactly is needed and how to configure properly.

With the CLI, the administrator is working in an environment with much less guardrails and guidance, so ensure that a certain base level of expertise has been acquired before engaging with it.

AWS Software Development Kit

A method for modifying and provisioning AWS resources programmatically, the AWS Software Development Kit (or SDK) allows us to build applications that leverage AWS resources for its functioning.

With support for over a dozen major programming languages like Python and Java, the SDK is going to be the major method of interacting with AWS for the readers who plan on being software engineers or developers. Utilizing SDKs properly however, will obviously require familiarity with an established programming language alongside knowledge of AWS, its resources and its 200+ services.

While all three of the mentioned interface types operate using radically different method from one another, it is important to remember that they are simply different ways of achieving the same thing: **The creation and management of AWS resources**. Therefore, having at least a basic understanding of what those resources are, and the services they relate to must be treated as a pre-requisite to operating the mentioned interfaces.

Once that has been achieved, the reader may find that there is actually very little difference between say, using AWS through the CLI or using AWS through the console, as the soft-skills and knowledge required to do both are quite similar, it is merely a matter of preference which interface the reader prefers.

Types of AWS Services

With over 200 different AWS Services, it might seem intimidating to start learning AWS. After all, where to start?

To make this task seem less insurmountable, the services have been divided into the following broad categories:

Compute Services

Compute services provide the processing power required to run various applications and manage servers. These services are crucial for tasks that need high computational power, scalability, and flexibility. They enable businesses to run their applications without needing to maintain physical hardware, allowing for more efficient resource management and cost-effectiveness. Compute services also support a variety of architectures and environments, ensuring compatibility with different types of workloads. Example services in this category include:

- Amazon EC2 (Elastic Compute Cloud)
- AWS Lambda
- Amazon ECS (Elastic Container Service)

Storage Services

Storage services offer scalable and durable storage solutions to meet the needs of data-intensive applications and workloads. These services provide secure and reliable storage for data of all types and sizes, including structured and unstructured data. They ensure data availability and integrity, supporting various use cases such as backup and recovery, data archiving, and content distribution. Storage services also facilitate efficient data management and retrieval, enabling users to access their data from anywhere at any time. Example services in this category include:

- Amazon S3 (Simple Storage Service)
- Amazon EBS (Elastic Block Store)
- Amazon EFS (Elastic File System)
- AWS Glacier

Database Services

Database services provide managed database solutions for storing and managing data. These services include support for relational, NoSQL, and in-memory databases, catering to a wide range of applications and use cases. Database services offer high availability, scalability, and security, ensuring optimal performance for database operations. They also simplify database management tasks, such as backups, patching, and monitoring, allowing users to focus on their application logic rather than database administration. Example services in this category include:

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

Networking and Content Delivery Services

Networking and content delivery services enhance the performance, security, and availability of applications. These services enable efficient communication between different components of an application and optimize the delivery of content to users worldwide. They also provide tools for managing network traffic, securing data in transit, and ensuring high availability through global distribution and load balancing. Example services in this category include:

- Amazon VPC (Virtual Private Cloud)

- Amazon CloudFront
- AWS Direct Connect
- AWS Elastic Load Balancing (ELB)

Security, Identity, and Compliance Services

Security, identity, and compliance services help protect AWS resources and data. These services provide tools for managing access to resources, encrypting data, and ensuring compliance with regulatory requirements. They enable users to implement security best practices, monitor security events, and respond to security incidents, ensuring the integrity and confidentiality of their data. Example services in this category include:

- AWS IAM (Identity and Access Management)
- AWS KMS (Key Management Service)
- AWS Shield
- AWS WAF (Web Application Firewall)

Management and Governance Services

Management and governance services provide tools for monitoring, managing, and optimizing AWS environments. These services help users gain visibility into their resources, automate management tasks, and ensure compliance with organizational policies. They support efficient resource allocation, cost management, and performance optimization, enabling users to maintain control over their AWS environments. Example services in this category include:

- AWS CloudFormation
- AWS CloudTrail
- AWS Config
- AWS CloudWatch

Machine Learning Services

Machine learning services provide tools and frameworks for building, training, and deploying machine learning models. These services enable users to leverage advanced machine learning techniques for various applications, such as predictive analytics, natural language processing, and image recognition. Machine learning services simplify the process of developing and deploying models, allowing users to focus on creating innovative solutions without needing extensive machine learning expertise. Example services in this category include:

- Amazon SageMaker
- Amazon Comprehend
- Amazon Rekognition

Analytics Services

Analytics services provide tools for analyzing and processing large volumes of data. These services support various data analytics tasks, such as data warehousing, real-time analytics, and big data processing. They enable users to gain insights from their data, make data-driven decisions, and improve business outcomes. Analytics services also offer scalable and cost-effective solutions for managing and analyzing data, ensuring optimal performance and reliability. Example services in this category include:

- Amazon EMR (Elastic MapReduce)
- Amazon Kinesis
- AWS Glue
- Amazon QuickSight

Application Integration Services

Application integration services facilitate communication and data exchange between different applications and systems. These services enable users to build complex workflows, automate processes, and integrate various applications seamlessly. They support various integration patterns, such as messaging, event-driven architecture, and service orchestration, ensuring reliable and efficient application integration. Example services in this category include:

- Amazon SQS (Simple Queue Service)
- Amazon SNS (Simple Notification Service)
- AWS EventBridge

We will explore each of the mentioned type of AWS service, either dedicating an entire chapter to them or pairing them with a complementary category of service. Each chapter is comprised of sections, with each section being devoted to a service or set of services that we have deemed to be important.

Some chapters also includes a background section, designed to provide readers with the necessary technical concepts, history and background knowledge required to understand the contents of the chapter and the services described within them.

The Background sections are a mandatory read for any and all readers from a non-technical field and/or for people not familiar with Computer Science fundamentals. More learned readers may however choose to either read or skip the section as they see fit.