HANDBOOK OF AWS (Amazon Web Services) DCC

Guided by

Dr. Madhura Phatak

Group Members:

Roll Number	PRN	Name
2	1032220158	Kush Fadia
8	1032220251	Yuvraj Khamkar
11	1032220263	Aaditya Mehta
12	1032220271	Rujuta Kulkarni
14	1032220274	Ayush Patil

Contents

1. Introduction to AWS

Amazon Web Services (AWS) is the world's leading cloud computing platform, offering a broad range of services that help businesses scale efficiently and securely. Launched by Amazon in 2006, AWS provides a variety of computing resources, including storage, networking, databases, artificial intelligence, and security services. The platform enables organizations to deploy applications faster and reduce infrastructure costs compared to traditional on-premises solutions.

AWS operates in a global cloud infrastructure, with data centers spread across multiple regions and availability zones. This ensures high availability, fault tolerance, and disaster recovery. AWS serves millions of customers, including startups, enterprises, and government agencies, allowing them to build and run applications with improved agility.

With AWS, businesses benefit from a pay-as-you-go pricing model, which eliminates upfront investments in hardware and reduces operational costs. The cloud platform also provides tools for automation, security compliance, and analytics, making it a preferred choice for modern digital transformation initiatives.

1.1 Major Services of AWS

AWS offers a vast portfolio of cloud services tailored to different business needs. The five major services covered in this handbook include:

1. **Amazon EC2 (Elastic Compute Cloud)** Amazon EC2 provides resizable compute capacity in the cloud. It allows users to launch virtual servers, known as instances, with different configurations of CPU, memory, and storage. EC2 supports multiple operating systems and enables auto-scaling to handle varying workloads. It integrates with other AWS services for security, monitoring, and automation, making it ideal for hosting applications, databases, and machine learning workloads.

2. **Amazon S3 (Simple Storage Service)** Amazon S3 is an object storage service that offers industry-leading scalability, durability, and security. Users can store and retrieve any amount of data, from simple text files to large multimedia content. S3 provides different storage classes, such as Standard, Intelligent-Tiering, and Glacier, catering to various use cases like real-time access, cost optimization, and long-term archiving. It also offers encryption, access control, and data lifecycle management features.

3. **Amazon RDS (Relational Database Service)** Amazon RDS simplifies the setup, operation, and scaling of relational databases in the cloud. It supports multiple database engines, including MySQL, PostgreSQL, MariaDB, Oracle, and SQL Server. RDS automates routine database tasks such as backups, patching, and replication, ensuring high availability and security. Businesses use RDS to run mission-critical applications, data warehousing, and online transaction processing (OLTP) systems.

4. **AWS Lambda** AWS Lambda is a serverless computing service that allows developers to run code without provisioning or managing servers. It executes code in response to events such as API requests, file uploads, or database changes. Lambda automatically scales based on demand and charges only for the actual compute time used. This makes it an efficient choice for building event-driven applications, backend services, and real-time data processing solutions.

5. **Amazon CloudFront** Amazon CloudFront is a content delivery network (CDN) that accelerates the delivery of web content, videos, APIs, and other digital assets. It distributes content through a global network of edge locations, reducing latency and improving performance for users worldwide. CloudFront integrates with AWS security services like AWS Shield and AWS Web Application Firewall (WAF) to protect against cyber threats. Businesses leverage CloudFront for website acceleration, streaming media, and global application distribution.

2. Amazon EC2 - Service and Usage

Amazon Elastic Compute Cloud (EC2) is a web service provided by Amazon Web Services (AWS) that offers scalable and resizable computing capacity in the cloud. It eliminates the need to invest in physical hardware by allowing users to rent virtual servers, called instances, on demand.

2.1 Benefits of Using CloudFront

Resizable Compute Capacity

- Users can increase or decrease the number of virtual machines (instances) based on their needs.

- Helps businesses handle traffic spikes efficiently.

Wide Range of Instance Types

- EC2 offers different instance families optimized for specific workloads (e.g., compute-intensive, memory-optimized, GPU-based).

- Example:

  - t2.micro – Good for small applications like websites.

  - m5.large – Balanced compute and memory for general workloads.

  - g4dn.xlarge – GPU instances for AI/ML workloads.

Supports Multiple Operating Systems

- EC2 allows users to choose from various OS options, including:

  - Linux distributions (Ubuntu, Amazon Linux, CentOS, etc.)

  - Windows Server (for applications requiring a Windows environment)

Security and Networking

- EC2 provides security through AWS Identity and Access Management (IAM).

- Virtual Private Cloud (VPC) allows users to define networking rules.

- Security Groups and firewalls restrict access to instances.

Auto Scaling and Load Balancing

- Auto Scaling: Automatically increases or decreases the number of instances based on traffic patterns.

- Elastic Load Balancing (ELB): Distributes incoming traffic across multiple instances for better performance.

2.2 Amazon EC2 in Commercial Usage

Example Use Case

Example 1: Hosting a Web Application

A company developing an e-commerce website can use EC2 instances to:

- Host the website on a **Linux/Windows** EC2 instance.

- Store product images and user data in **Amazon S3** and **Amazon RDS** (for databases).

- Scale instances up during Black Friday sales to handle high traffic.

- Reduce the number of instances when demand decreases to save costs.

Example 2: Big Data Processing

A data analytics company running large-scale processing jobs can:

- Use **EC2 Spot Instances** for cost-efficient batch processing.

- Deploy **high-memory instances** for running machine learning (ML) models.

- Scale computing resources automatically using **Auto Scaling Groups**.

Amazon EC-2 Workflow

Features of EC2

Features of Amazon EC2

- Bare Metal Instances
- Optimize with Amazon EC2 Fleet
- Pause and Resume Your Instances
- GPU Compute Instances
- GPU Graphics Instances
- High I/O Instances
- Dense HDD Storage Instances
- Optimized CPU Configurations
- Flexible Storage Options
- Paying for What You Use
- Multiple Locations
- Elastic IP Addresses
- Auto Scaling
- HPC Clusters
- Enhanced Networking
- Elastic Fabric Adapter
- Available on AWS PrivateLink
- Amazon Time Sync Service

WORKFALL

3. Amazon S3 - Service and Usage

Amazon Simple Storage Service (Amazon S3) is a **highly scalable, durable, and secure object storage service** offered by AWS. It allows users to store, retrieve, and manage large amounts of data over the internet.

Unlike traditional file systems or block storage, S3 uses an **object storage** model, meaning data is stored as objects in **buckets** (containers for data).

3.1 Service Overview

Key Features of Amazon S3

1. **Highly Scalable Storage**

   o Supports **unlimited data storage** for objects ranging from a few bytes to terabytes in size.

   o Can handle millions of requests per second.

2. **Object Storage**

   o Data is stored as objects within **buckets**.

   o Each object consists of:

       ▪ **Data** (the actual file or content).

       ▪ **Metadata** (information like last modified date, permissions).

       ▪ **A unique key (identifier)**.

3. **Data Durability & Availability**

   o S3 provides **99.999999999% (11 nines) durability** by replicating data across multiple AWS Availability Zones.

   o Offers **high availability** by automatically distributing stored data across regions.

4. **Security and Access Control**

   o **Encryption**: Data can be encrypted using AWS-managed keys or customer-provided keys.

   o **Access Management**: Supports IAM policies, bucket policies, and Access Control Lists (ACLs).

   o **Public & Private Access**: Users can set bucket policies to control who can read or write data.

5. **Flexible   Storage   Classes**
   Amazon S3 offers multiple **storage classes** to optimize cost based on data access patterns:

   o **S3 Standard**: For frequently accessed data.

   o **S3 Intelligent-Tiering**: Automatically moves data to lower-cost storage based on usage.

3.2 Usage Examples

Example 1: Media Storage & Distribution

A media company uses S3 for:

- Storing **high-resolution images, videos, and documents** in S3 buckets.

- Using **CloudFront (CDN)** to distribute images globally with low latency.

- Storing older content in **S3 Glacier** to reduce costs.
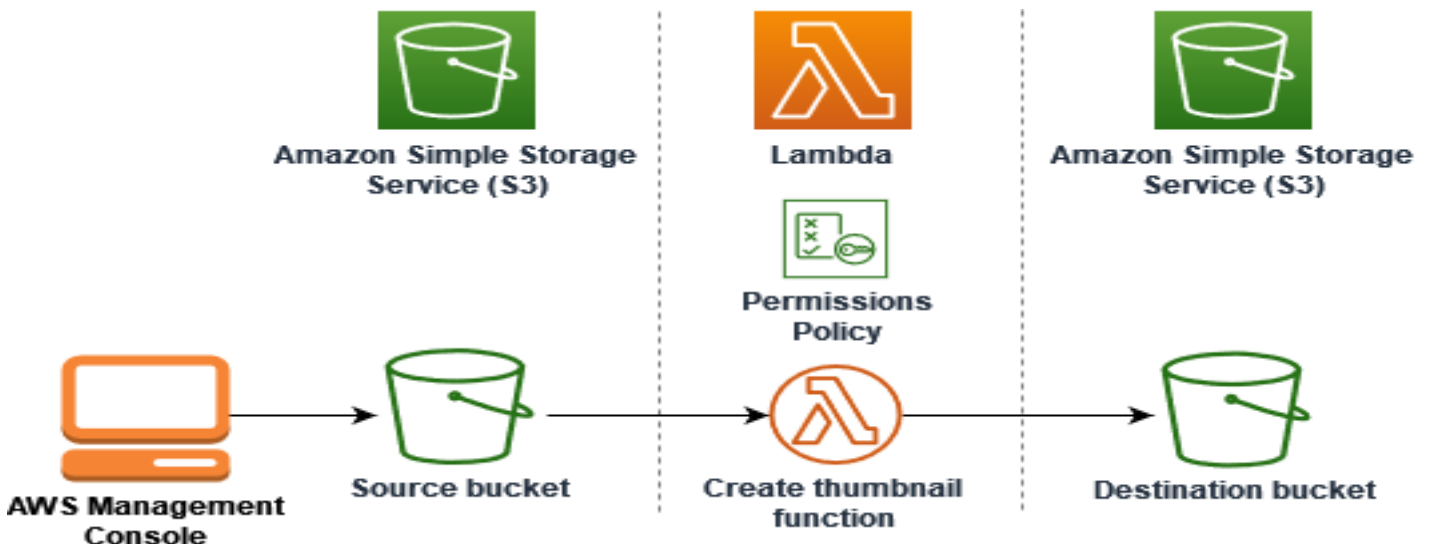
Example 2: Website Hosting

S3 can host **static websites** (HTML, CSS, JavaScript) by:

- Storing website files in an S3 bucket.

- Enabling **static website hosting** with a custom domain.

- Using **CloudFront** for faster global access.

Amazon S3 Workflow: -
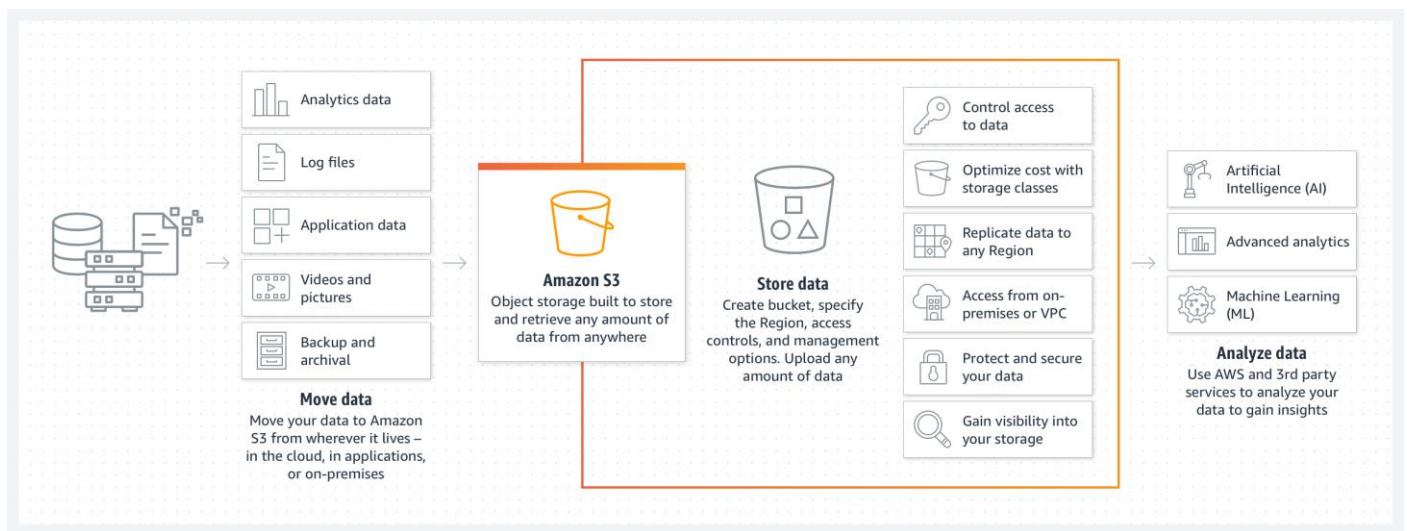
Amazon S3 Integration with AWS: -

4. Amazon RDS- Service Overview

Amazon Relational Database Service (Amazon RDS) is a fully managed database service that makes it easy to set up, operate, and scale a relational database in the cloud. It automates time-consuming administrative tasks such as hardware provisioning, database setup, patching, and backups.

4.1 Key Features:

- **Managed Infrastructure**: Automatically handles maintenance, backups, and software patching.

- **High Availability**: Provides Multi-AZ (Availability Zone) deployments for failover support.

- **Scalability**: Storage and compute capacity can be easily adjusted to meet workload demands.

- **Security**: Offers encryption at rest and in transit, VPC support, and IAM integration.

- **Monitoring**: Integrated with Amazon CloudWatch for metrics, alarms, and dashboards.

- **Automated Patching**: Ensures the database runs on the latest stable and secure version with minimal downtime.

4.2 Supported Database Engines:

- Amazon Aurora (MySQL and PostgreSQL-compatible)

- MySQL

- PostgreSQL

- MariaDB

- Oracle Database

- Microsoft SQL Server

- IBM Db2 (through partner solutions)

4.3 Usage Example

Use Case: Financial Institution Handling Transactions

A large financial institution uses Amazon RDS to manage sensitive customer transaction records. These records include deposits, withdrawals, and transfers, and require high availability and secure storage.

Implementation Details:

- **Database Engine**: PostgreSQL on Amazon RDS.

- **Deployment Model**: Multi-AZ for automatic failover and durability.

- **Security Measures**:

  o Encrypted storage using AWS KMS.

  o Data access managed through IAM roles and VPC security groups.

  o SSL-enabled connections for secure data transmission.

- **Backup Strategy**: Automated daily snapshots and point-in-time recovery enabled.

- **Scalability**: Configured to autoscale storage as transaction volume grows.

- **Performance Optimization**: Integrated performance insights used for identifying and resolving query bottlenecks.

4.4 Benefits Observed:

- **Reduced Operational Overhead**: No manual intervention required for backups or patching.

- **Improved Performance**: Optimized read replicas are used for reporting and analytics.

- **Enhanced Reliability**: Seamless failover in case of infrastructure failure ensures continuous uptime.

- **Cost Efficiency**: Pay-as-you-go pricing with reserved instances option leads to optimized cost over time.
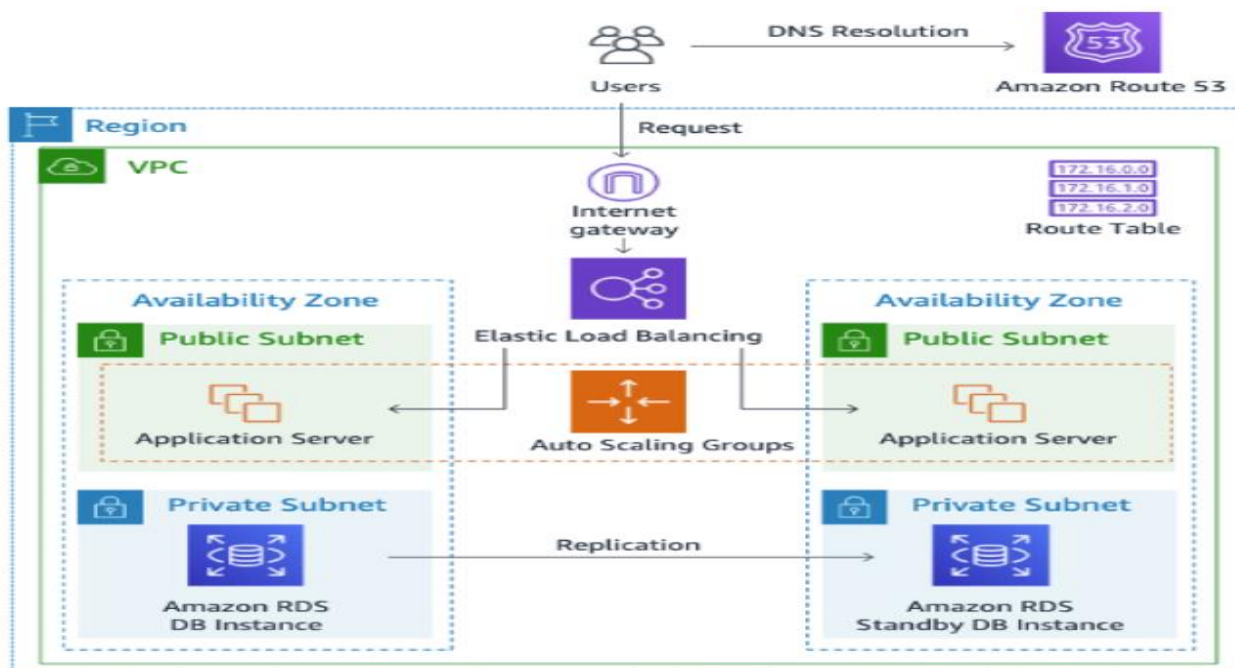
Use Case of a dynamic website that uses Amazon RDS DB

5. AWS Lambda – Service Overview

AWS Lambda is a serverless computing service that runs backend code in response to events. It allows developers to focus solely on writing code without worrying about server provisioning, scaling, or maintenance.

AWS Lambda automatically manages the compute fleet, scaling precisely with the size of the workload, and charges only for the compute time consumed, making it highly cost-efficient.

5.1 Key Features:

- **Event-Driven Execution**: Runs code in response to triggers such as S3 uploads, API Gateway calls, or DynamoDB streams.

- **No Server Management**: Developers don't need to manage any infrastructure—just write and deploy the code.

- **Automatic Scaling**: Scales up automatically to match the volume of incoming requests.

- **Pay-per-Use**: Charged based on the number of requests and duration the code runs.

- **Integrated Security**: Supports IAM for access control and secure execution.

- **Multiple Runtime Support**: Supports popular languages including Python, Node.js, Java, Go, and .NET.

5.2 Supported Event Sources:

- Amazon S3 (e.g., on object creation or deletion)

- Amazon API Gateway

- Amazon DynamoDB Streams

- Amazon SNS & SQS

- Amazon CloudWatch Events & Logs

- Amazon Kinesis Data Streams

- Amazon EventBridge

5.3 Usage Example

Use Case: Automated Image Resizing

A developer uses AWS Lambda to automate image resizing whenever a new image is uploaded to an S3 bucket. This is commonly used in content-heavy applications like social media platforms, e-commerce sites, and photo-sharing apps.

Implementation Details:

- **Trigger**: S3 bucket upload event triggers the Lambda function.

- **Functionality**: The Lambda function resizes the uploaded image into predefined dimensions (e.g., thumbnail, medium, large).

- **Language Used**: Python with Pillow library for image processing.

- **Storage**: Resized images are saved in a separate S3 bucket or prefix.

- **Permissions**: IAM role grants Lambda access to read/write from S3.

- **Monitoring**: CloudWatch is used to log function invocations and failures.

5.4 Benefits Observed:

- **Reduced Infrastructure Complexity**: No need for dedicated image processing servers.

- **Real-Time Processing**: Images are resized and available within seconds of upload.

- **Lower Costs**: Pay only for the time the Lambda function runs, no idle charges.

- **Easy Integration**: Seamlessly works with other AWS services like S3 and CloudWatch.

- **High Availability**: Lambda functions run across multiple Availability Zones by default.
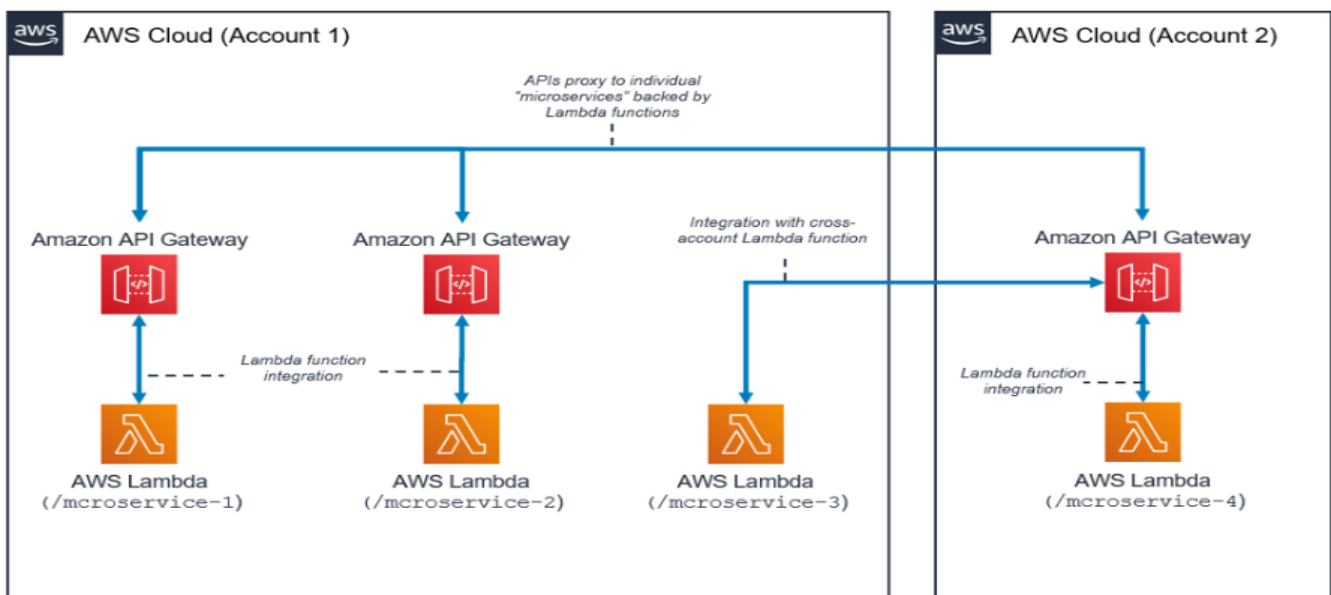
Architectural pattern for microservices with Lambda

6. Amazon CloudFront - Service & Usage

Amazon CloudFront is a Content Delivery Network (CDN) service provided by AWS, designed to distribute content globally with high performance and low latency. It caches and delivers static and dynamic web content, including images, videos, applications, and APIs, from AWS edge locations worldwide. CloudFront integrates with other AWS services like S3, EC2, and Lambda to optimize performance and security.

6.1 Benefits of Using CloudFront

- Faster Content Delivery: Content is distributed closer to users via global edge servers, minimizing load times.

- Enhanced Security: Supports SSL/TLS encryption and AWS Shield for protection against DDoS attacks.

- Efficient Handling of Dynamic Content: Optimized for APIs, real-time applications, and dynamic websites.

- Cost Optimization: Reduces bandwidth costs through efficient caching strategies.

- Seamless AWS Integration: Works effectively with Route 53, AWS WAF, and other AWS security and networking services.

6.2 Real-World Use Case: Optimized Video Streaming

- A video streaming service aiming to provide high-quality content to users across various regions can leverage Amazon CloudFront to achieve the following advantages:

- Reduced Buffering: Videos are stored at edge locations closest to viewers, ensuring smooth playback.

- Minimized Server Load: Frequently accessed content is cached, reducing the burden on origin servers.

- Adaptive Streaming Support: Delivers video in different formats such as HLS and MPEG-DASH, adjusting quality based on user bandwidth.

- Robust Security Measures: Implements signed URLs and AWS Shield to safeguard content and users.

- Cost Efficiency: Reduces data transfer expenses by optimizing content delivery across global locations.

6.3 Companies Using Amazon CloudFront

- Spotify: Uses CloudFront to instantly update new features and provide access to a vast library of songs.

- Jio Saavn: Delivers 15 petabytes of audio and video content to subscribers globally with CloudFront.

- Sky News: Utilizes CloudFront to unify content and enable faster distribution to its audience.

7. Amazon EC2 in Commercial Usage

Amazon Elastic Compute Cloud (EC2) provides businesses with flexible, scalable computing power in the cloud. It allows organizations to deploy applications without worrying about managing physical servers, offering a reliable and cost-effective solution for IT infrastructure.

7.1 Common EC2 Applications

- Web Hosting: Powers websites and applications with high availability and scalability.

- Big Data Analytics: Supports large-scale data processing and machine learning workloads.

- Artificial Intelligence: Enables AI model training using powerful GPU-based instances.

- Disaster Recovery: Ensures business continuity with backup and failover capabilities across regions.

7.2 Example Use Case

Netflix: Powering Global Streaming with EC2

Netflix relies heavily on Amazon EC2 to handle its vast streaming infrastructure. EC2 instances allow Netflix to dynamically scale its compute power based on demand, ensuring a smooth viewing experience even during peak times.

- Auto-Scaling: Netflix automatically adjusts EC2 instances based on user demand, preventing outages and optimizing costs.

- Content Delivery & Transcoding: EC2 processes and encodes vast amounts of video content to different resolutions and formats before distribution.

- Machine Learning & Personalization: EC2 powers AI-driven recommendations, analyzing user behavior to suggest relevant content in real-time.
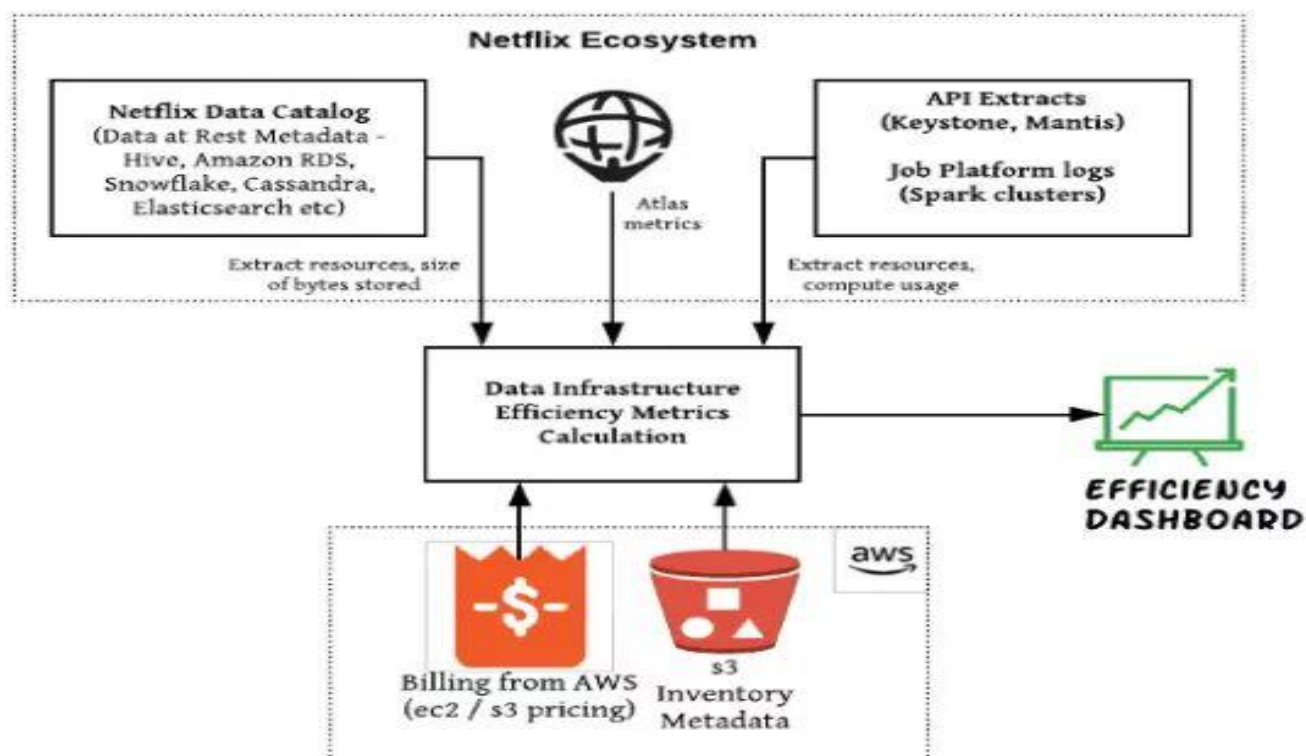
8. AWS Use Cases

AWS is widely adopted across various industries, offering scalable and cost-effective cloud solutions. Below are some of the key use cases and industries where AWS plays a crucial role:

1. Cloud Hosting & Web Applications:

AWS provides scalable infrastructure for web hosting, e-commerce platforms, and content delivery. Netflix, Airbnb, and LinkedIn use AWS to handle millions of user requests seamlessly.

2. Artificial Intelligence & Machine Learning

AWS offers powerful AI/ML services, including Amazon SageMaker, AWS Rekognition, and AWS Lex, which help businesses build and deploy AI models for image recognition, natural language processing, and predictive analytics.

3. Internet of Things (IoT) & Smart Devices

AWS IoT services, such as AWS IoT Core, AWS Greengrass, and AWS IoT Analytics, enable businesses to connect and manage IoT devices at scale. Smart homes, industrial automation, and fleet management use AWS to process real-time sensor data for analytics and decision-making.

4. Big Data & Analytics

Organizations use AWS to store, process, and analyze large datasets using services like Amazon Redshift, AWS Glue, and Amazon EMR. AWS allows industries like finance, healthcare, and research to gain insights from massive data volumes and enhance decision-making.

5. Media & Entertainment

AWS is used for video streaming, game development, and content delivery. Companies like Disney+ and Twitch leverage AWS services such as Amazon CloudFront (CDN), AWS Elemental Media Services, and Amazon S3 to stream high-quality video content to millions of users globally.

6. Financial Services & Banking

Banks and financial institutions use AWS for fraud detection, real-time transaction monitoring, and risk assessment. AWS services like Amazon Kinesis, AWS Lambda, and AWS Shield help ensure secure and compliant financial transactions.

7. Healthcare & Life Sciences

AWS enables secure storage, processing, and analysis of healthcare data. Services like AWS HealthLake, Amazon Comprehend Medical, and AWS Lambda help hospitals and researchers manage electronic health records, perform genomic analysis, and develop AI-driven diagnostics.

8.1 Future of AWS

As cloud technology evolves, AWS continues to innovate. Key trends include:
   1. Quantum Computing with AWS Braket – Allows researchers to develop quantum algorithms, benefiting cybersecurity and financial modeling.

2. Growth of Serverless Computing – AWS Lambda adoption will increase, reducing infrastructure management for businesses.

3. Strengthening Cybersecurity – AWS will enhance services like AWS Shield and IAM to improve data protection.

4. Sustainability Initiatives – AWS is investing in energy-efficient data centers and renewable energy to reduce its carbon footprint.

5. Edge Computing & 5G Integration – Services like AWS Wavelength will enhance low-latency applications for IoT and smart cities.

6. Hybrid & Multi-Cloud Strategies – AWS is expanding AWS Outposts and AWS EKS Anywhere for seamless hybrid cloud adoption.
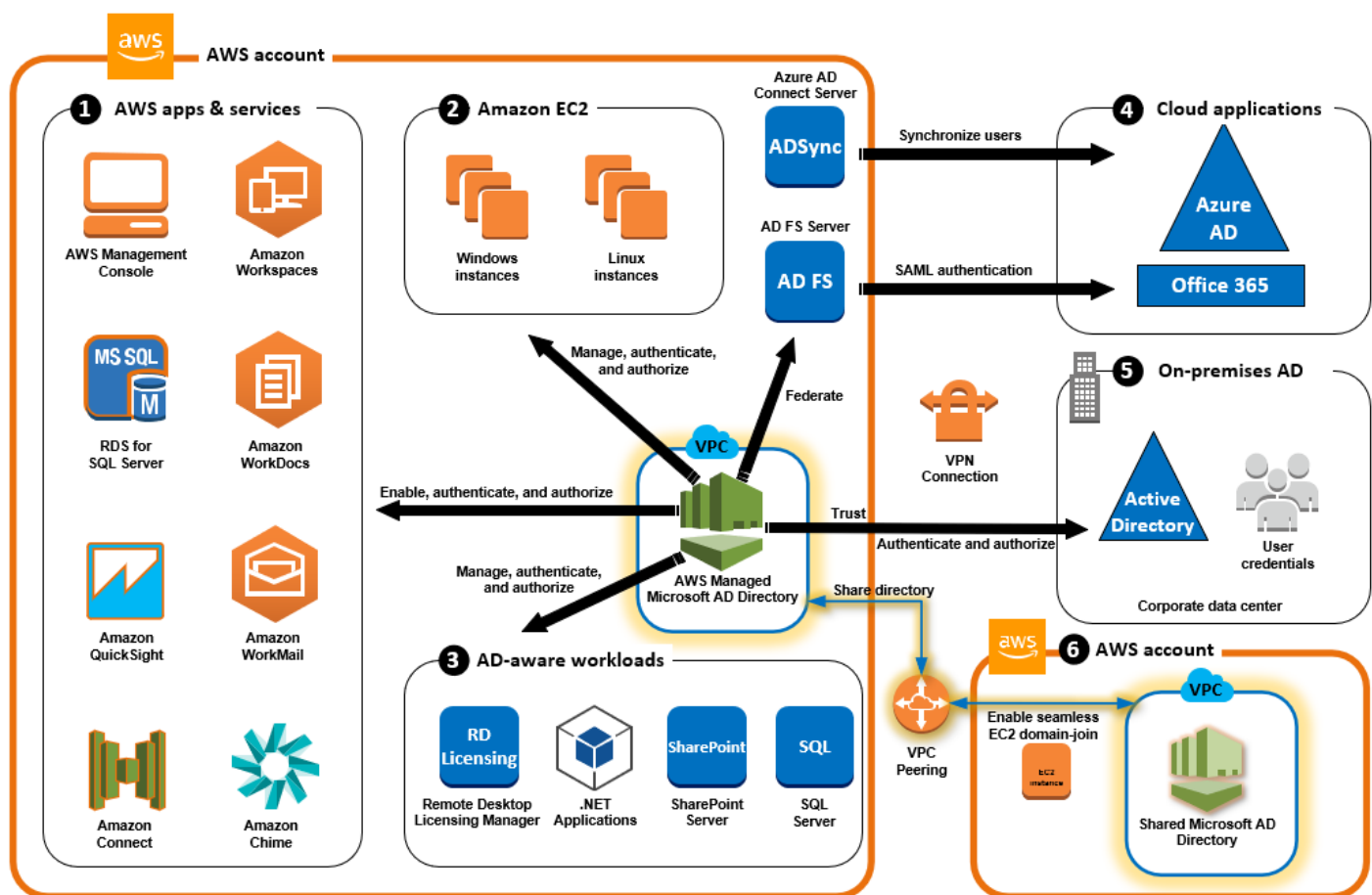
9. Conclusion

AWS has transformed the cloud computing landscape by providing scalable, secure, and cost-efficient cloud solutions across various industries. From startups to Fortune 500 companies, businesses rely on AWS for compute power, storage, networking, AI/ML, and security. Its global infrastructure, continuous innovation, and flexible pricing models make it the most popular cloud platform today. AWS is not just a cloud provider but a technology enabler, allowing organizations to innovate faster and operate efficiently.