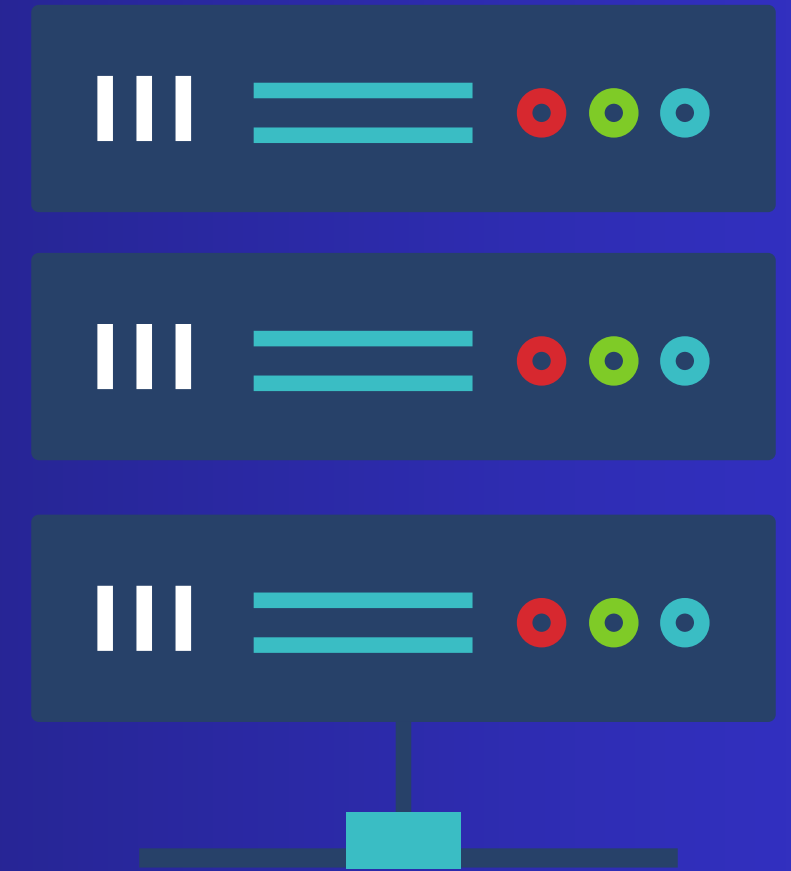


MPRASHANT

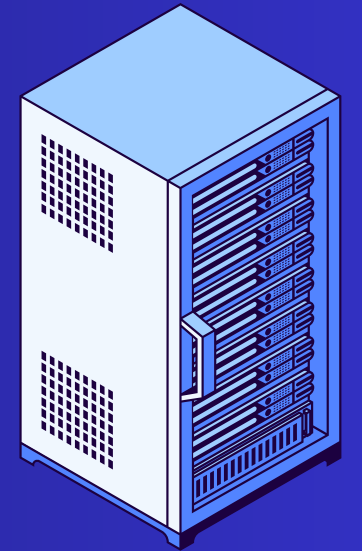


AWS EC2



Elastic Cloud Compute

AWS EC2 (Amazon Elastic Compute Cloud) is a cloud service that provides resizable virtual servers, called instances, which you can use to run applications.



Imagine you're running a business and need a server to host your website or application.

*Instead of buying and managing physical servers, **AWS EC2** lets you rent virtual servers in the cloud. These virtual servers are called instances.*



MPRASHANT



You Can Configure several options...





OS

RAM

CPU

Disk Space

Network/Firewall

How to access the machine?





- **Instance Type:** Select the hardware capacity (e.g., CPU, memory).
- **AMI (Amazon Machine Image):** Choose the **operating system** and software (linux, mac, windows).
- **Storage:** Configure the type and size of storage (e.g., EBS volume).
- **Security Groups:** Set up firewall rules to control inbound/outbound traffic.
- **Key Pair:** Create or use an existing key pair for SSH access.



- **Network Settings:** *Configure VPC, subnet, and assign public or private IP addresses.*
- **IAM Role:** *Attach an IAM role for permissions to access other AWS resources.*
- **User Data:** *Add scripts to be executed when the instance starts.*
- **Elastic IP:** *Optionally associate a static IP address for consistent public access.*

Creating our EC2 Instance

Userdata script to install and run Apache Webserver

```
#!/bin/bash
```

```
sudo yum update -y
```

```
# Install Apache web server (httpd)
```

```
sudo yum install -y httpd
```

```
sudo systemctl start httpd
```

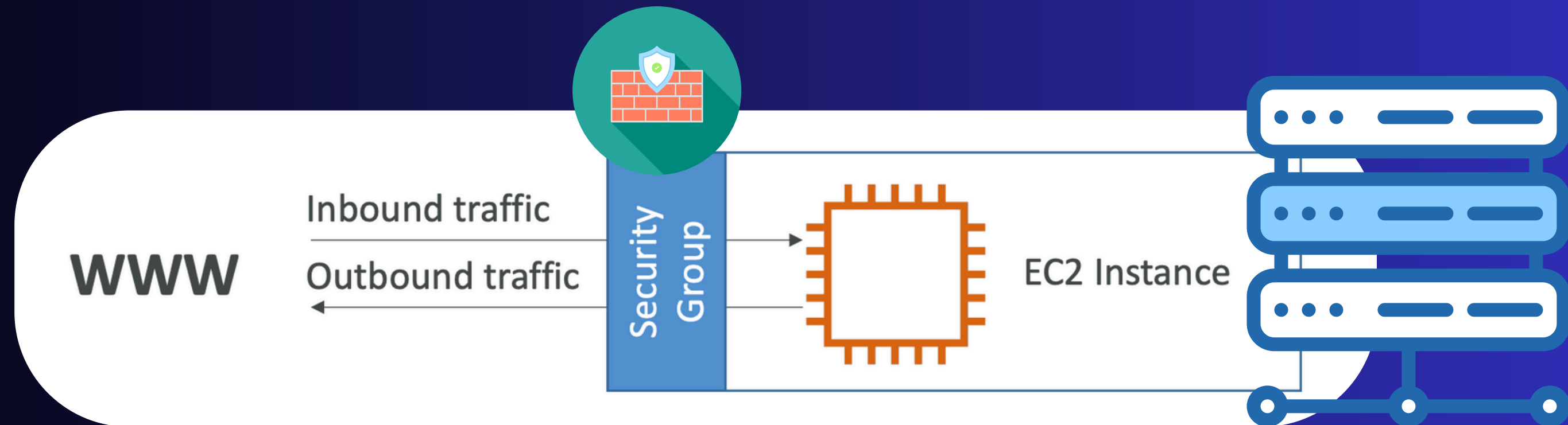
```
sudo systemctl enable httpd
```

```
# Create a simple HTML file to verify the web server is running
```

```
echo "<html><h1>Welcome to Apache Web Server on Amazon Linux!</h1></html>" >
```

```
/var/www/html/index.html
```

Security Groups: Network firewall rules that control inbound and outbound traffic for instances.



Important points about security groups

- Region specific
- Only 'Allow' rule (but no deny rule)
- All inbound traffic blocked and outbound allowed by default.
- You define rules for specific:
 - Protocols (like HTTP, HTTPS, SSH, etc.).
 - Port numbers (e.g., port 80 for HTTP, port 22 for SSH).
 - IP addresses or ranges (e.g., allow traffic only from a specific IP or range of IPs).
- If you allow incoming traffic on a specific port (e.g., port 80 for HTTP), the outgoing response traffic is automatically allowed without an explicit outbound rule.

Some ports you should be aware of:

- HTTP (Port 80) – Unencrypted web traffic.
- HTTPS (Port 443) – Encrypted web traffic (SSL/TLS).
- SSH (Port 22) – Secure remote access to servers (Linux/Unix).
- FTP (Port 21) – File Transfer Protocol (unsecured).
- SFTP (Port 22) – Secure File Transfer Protocol.
- SMTP (Port 25) – Simple Mail Transfer Protocol (email sending).
- RDP (Port 3389) – Remote Desktop Protocol (Windows remote access).
- MySQL (Port 3306) – MySQL database connections.
- PostgreSQL (Port 5432) – PostgreSQL database connections.
- DNS (Port 53) – Domain Name System (converts domain names to IP addresses).

How to SSH into EC2 instance?

SSH allows you to control/access a remote machine.





Cloud Compute Instances – Amazon EC2 Instance Types – AWS

Amazon EC2 instance types comprise varying combinations of CPU, memory, storage, and networking capacity. This gives you the flexibility to choose an instance that best meets your needs.

 Amazon Web Services

Instance Types

- **Case 1: Small Website or Blog**
 - Suitable Type: t3.micro or t3.small (General Purpose)
- **Case 2: E-Commerce Application**
 - Suitable Type: m5.large or m5.xlarge (General Purpose)
- **Case 3: Real-Time Video Rendering and Streaming (Accelerated Computing)**
 - Instance Type: g5.12xlarge or g5.24xlarge
- **Case 4: In-Memory Database for Real-Time Analytics (Memory Optimized)**
 - r6g.16xlarge or x2idn.32xlarge (Memory Optimized)

Purchasing Options

Purchasing Option	Best for	Cost	Commitment	Flexibility
On-Demand	Short-term, unpredictable workloads	High	None	High
Reserved Instances	Long-term, steady workloads (1-3 years)	Medium to Low (Up to 75% off)	1 or 3 years	Low
Spot Instances	Fault-tolerant, flexible, batch jobs, big data, CI/CD, distributed computing	Very Low (Up to 90% off)	None (subject to termination)	Medium (can be interrupted)

Savings Plans	Consistent workloads but needing more flexibility than Reserved Instances	Medium to Low (Up to 72% off)	1 or 3 years	High (more flexible than RI)
Dedicated Hosts	Workloads requiring complete hardware control (compliance, software licensing)	High	1 or 3 years	Medium
Dedicated Instances	Workloads needing physical isolation without full control over the hardware	High	None	Medium

- **For short-term, unpredictable workloads, On-Demand is the best.**
- **For long-term, predictable workloads, Reserved Instances or Savings Plans can save you significant costs.**
- **If you have flexible, fault-tolerant workloads, Spot Instances offer the most cost-effective option, while Dedicated Hosts or Dedicated Instances provide hardware isolation and compliance benefits.**

