

ONE PAGE SUMMARY

IT & Software > AWS > EC2

Introduction to EC2

Unlock the Power of Cloud Computing with EC2!

DEFINITION

Amazon EC2 (Elastic Compute Cloud) is a web service that provides resizable compute capacity in the cloud. It allows users to run virtual servers on-demand, making it easy to scale applications.

KEY CONCEPTS

Instances

Instances are virtual servers that run applications on EC2, and they come in various types to suit different needs.

AMIs

Amazon Machine Images (AMIs) are pre-configured templates used to create EC2 instances, containing the OS and software.

Elastic Load Balancing

This feature automatically distributes incoming application traffic across multiple EC2 instances for better availability.

Auto Scaling

Auto Scaling automatically adjusts the number of EC2 instances based on demand, ensuring optimal performance and cost-efficiency.

Security Groups

Security Groups act as virtual firewalls for your EC2 instances, controlling inbound and outbound traffic.

EXAMPLES

- Launching a web server using a pre-configured AMI.
- Scaling up the number of instances during high traffic periods.
- Using Elastic Load Balancing to distribute user requests.

MEMORY TIPS

- ★ Remember 'EC2' as 'Elastic Cloud Computing' for flexibility.
- ★ Think of AMIs as 'Application Meal Ingredients' for building instances.
- ★ Use 'Auto' in Auto Scaling to remember it adjusts automatically.

COMMON MISTAKES

- ✗ Not configuring security groups properly, leading to access issues.
- ✗ Overlooking instance types, resulting in performance bottlenecks.
- ✗ Failing to set up Auto Scaling, causing downtime during traffic spikes.

QUICK RECAP

Amazon EC2 provides scalable virtual servers in the cloud, enabling flexible application deployment. Key features include instances, AMIs, and Auto Scaling for efficient resource management.