Course Title : Auto Scaling

Trainer Name: Sohaib Ur Rehman

Amazon EC2 Auto Scaling

What is Auto Scaling?

"Auto-scaling also called Automatic scaling, is a method used in cloud computing, whereby the amount of computational resources in a server farm, typically measured in terms of the number of active servers, scales automatically based on the load on the farm. It is closely related to, and builds upon, the idea of load balancing"

Web Services

Amazon EC2 Auto Scaling

Benefits

SETUP SCALING QUICKLY

Target utilization levels for multiple resources in a single interface. Quickly see the average utilization of all of your scalable resources without having to navigate to other consoles.

AUTOMATICALLY MAINTAIN PERFORMANCE

Maintain optimal application performance and availability, even when workloads are periodic, unpredictable, or continuously changing. Continually monitors your applications to make sure that they are operating at your desired performance levels.

MAKE SMART SCALING DECISIONS

Automatically creates all the scaling policies and sets targets for you based on your preference. Monitors your application and automatically adds or removes capacity from your resource groups in real-time as demands change.

Web Services

Amazon EC2 Auto Scaling

Benefits

Improve Fault Tolerance

AWS EC2 Auto Scaling can detect when an instance is unhealthy, terminate it, and replace it with a new one.

Increase Application Availability

AWS EC2 Auto Scaling ensures that application always has the right amount of compute capacity.

Lower Costs

Amazon EC2 Auto Scaling adds new instances only when necessary, and terminates them when no longer needed.

Web Services

Amazon EC2 Auto Scaling

Features

Choose When And How To Scale

Scale dynamically based on AWS CloudWatch metrics, or predictably according to a schedule that defined. Receive notifications via Amazon SNS to be alerted.

Monitor The Health Of Running Instances

Amazon EC2 Auto Scaling ensures that application is able to receive traffic and that EC2 instances are working properly. Amazon EC2 Auto Scaling periodically performs health checks to identify any instances that are unhealthy.

Amazon EC2 Auto Scaling

Features

Fleet Management

To automate fleet management for EC2 instances, EC2 Auto Scaling monitors the health of running instances, automatically replaces impaired instances, and balances capacity across Availability Zones.

Support For Multiple Instance Types

User can run On-Demand or Spot Instances inside an Amazon EC2 Auto Scaling group, including inside VPC.

Amazon EC2 Auto Scaling

Functionality

Launch Configurations

Whenever an Auto Scaling group launches a new instance, it uses the currently associated launch configuration as a template for the launch.

Lifecycle Hooks

Lifecycle hooks take action before an instance goes into service or before it gets terminated. This can be especially useful if software environment not baking software into an AMI.

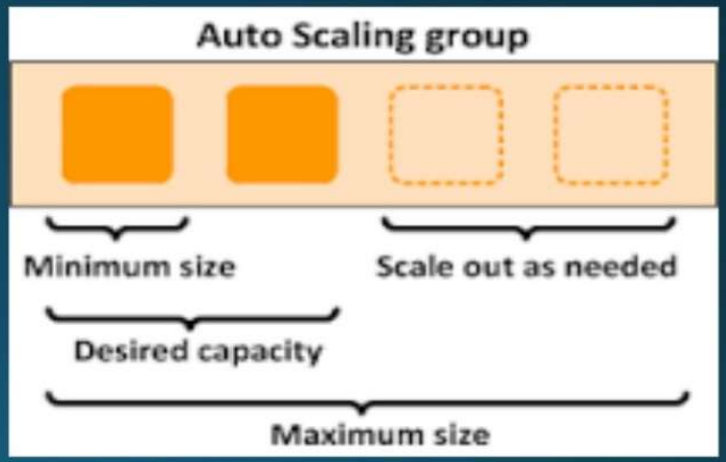
Terminate Lifecycle Hooks

Terminate hooks can be useful for collecting important data from an instance before it goes away.

Amazon EC2 Auto Scaling

Auto Scaling Group

- An Auto Scaling group starts by launching enough EC2 instances to meet its desired capacity.
- The Auto Scaling group maintains this number of instances by performing periodic health checks on the instances in the group.
- If an instance becomes unhealthy, the group terminates the unhealthy instance and launches another instance to replace it.



Amazon EC2 Auto Scaling

Scaling Policy Types

Amazon EC2 Auto Scaling supports the following types of scaling policies:

Target tracking scaling: Increase or decrease the current capacity of the group based on a target value for a specific metric.

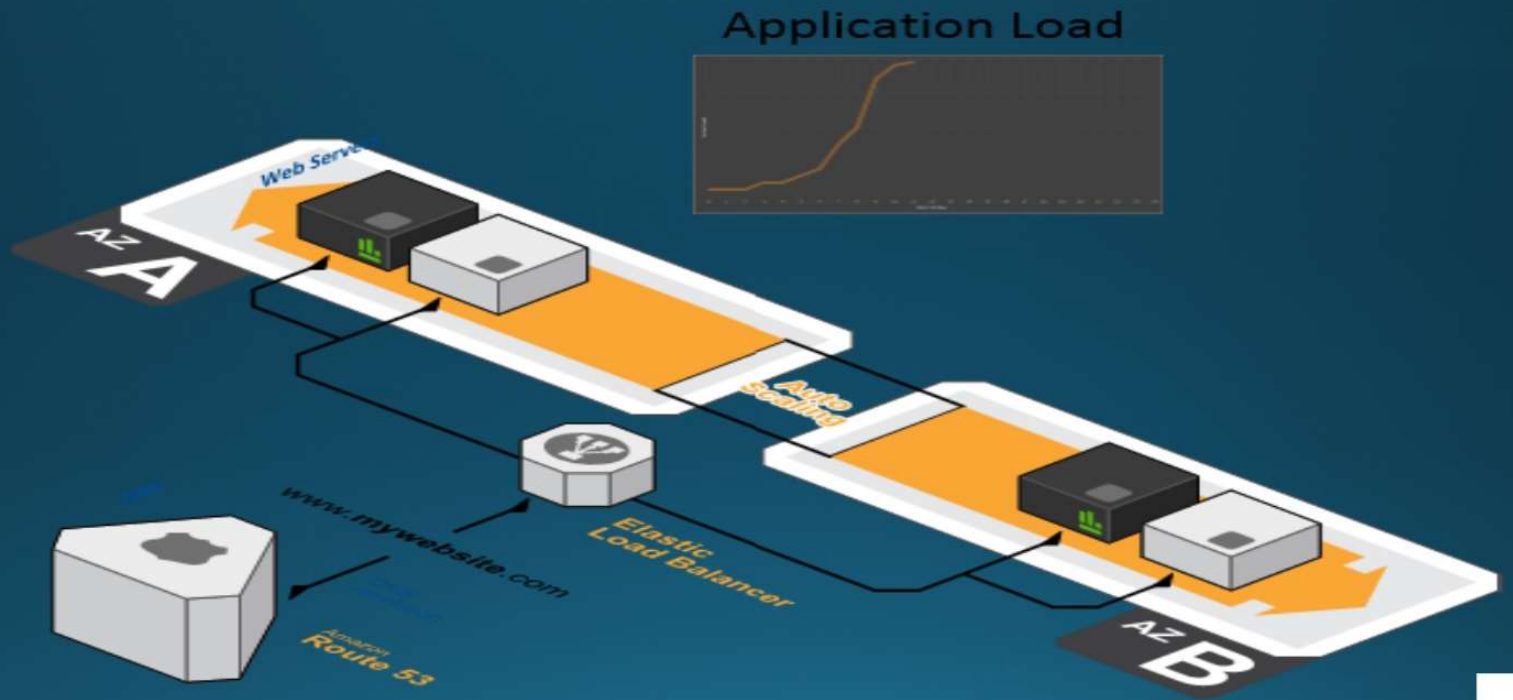
Step scaling: Increase or decrease the current capacity of the group based on a set of scaling adjustments, known as step adjustments.

Simple scaling: Increase or decrease the current capacity of the group based on a single scaling adjustment.

Scheduled scaling: The scheduled action tells Amazon EC2 Auto Scaling to perform a scaling action at specified times.

Web Services

Amazon EC2 Auto Scaling



Web Services

Contact Us

0213-498-OMNI(6664), 0213-534-460
0337-7222191, 0333-3808 376

0312-216-9325

info@omni-academy.com