

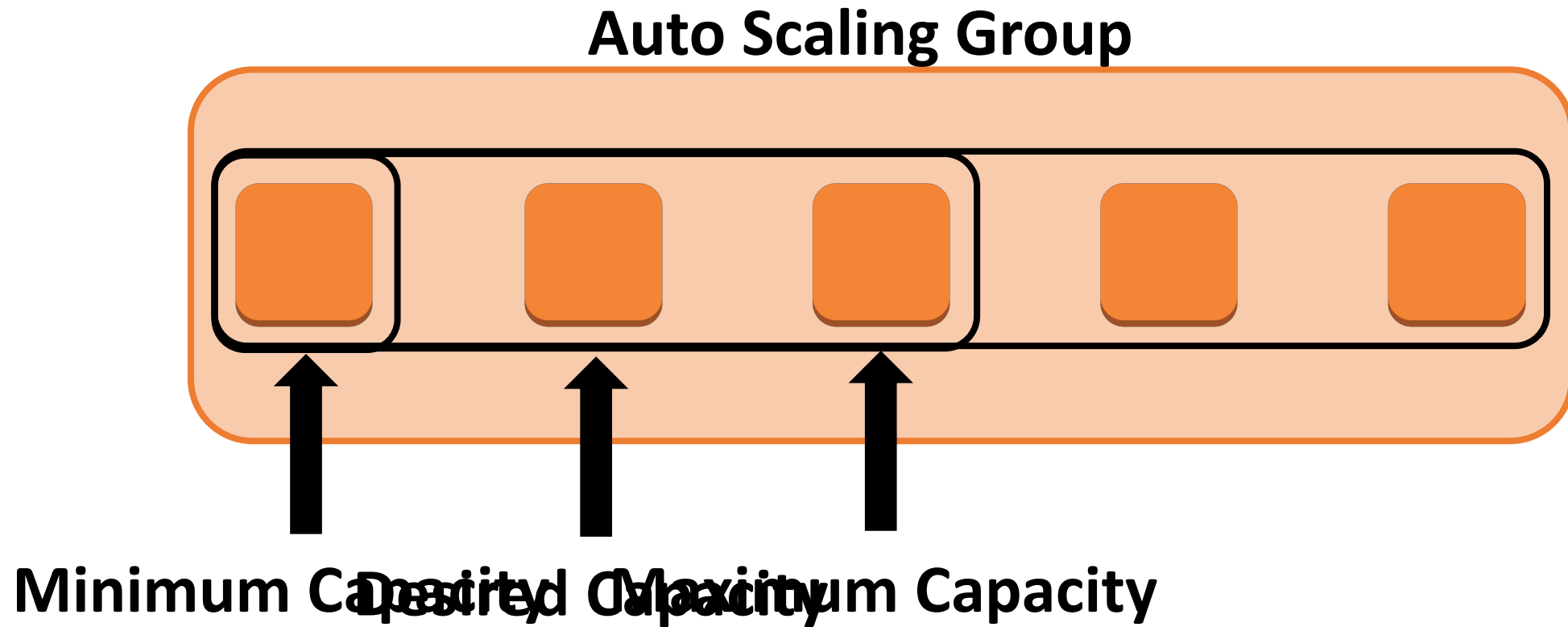
AWS Certified Cloud Practitioner  
Training Bootcamp

# AWS EC2 Auto Scaling Basics 101

# What is Amazon EC2 Auto Scaling ?

- Amazon EC2 Auto Scaling helps you ensure that you have the correct number of Amazon EC2 instances available to handle the load for your application (up or down)
- EC2 instances are grouped in *Auto Scaling Groups*:
  - Minimum number of EC2 instances
  - Desired number of EC2 instances
  - Maximum number of EC2 instances
- *Scaling policies* will automatically launch or terminate instances as your application demands

# Minimum, Maximum and Desired EC2 Capacity



# Launch Configurations

- A launch configuration is an instance configuration template that an ASG uses to launch EC2 instances
- What's included in a launch configuration:
  - AMI ID
  - Instance type
  - Key pair
  - Security group
  - Block (Volume)
- ... just as if you would launch an EC2 instance

# Auto Scaling Groups

- An *Auto Scaling Group* contains a collection of Amazon EC2 instances that are treated as a logical grouping for the purposes of automatic scaling and management
- Maintaining the number of instances in an Auto Scaling group and automatic scaling are the core functionalities of the Amazon EC2 Auto Scaling service
- The size of an Auto Scaling group depends on the number of instances you set as the desired capacity

# Auto Scaling Groups

- The ASG starts by launching the desired number of EC2 instances (desired capacity; i.e. 3 EC2 instances)
- You can use scaling policies to increase or decrease the number of instances in your group dynamically to meet changing conditions
- When the scaling policy is in effect, the Auto Scaling group adjusts the desired capacity of the group, between the min and max capacity values and launches or terminates the instances as needed (i.e. between 1 to 5)

## In the next section ...

- Use WebServer 01 and create our first AMI – Amazon Machine Image
- Create Launch Configuration – Desired 3 EC2 instances
- Create Auto Scaling Group with min. 1 EC2 and max. 5 EC2 and use the Launch Configuration
- Define dynamic Auto Scaling capabilities through Scaling Policies

AWS Certified Cloud Practitioner  
Training Bootcamp

Thank you