DATE : 31.01.2025

DT/NT : DT

LESSON: AWS

SUBJECT: EC2 AUTOSCALING

BATCH : B 303

AWS-DEVOPS











+1 (585) 304 29 59



# EC2 Auto Scaling



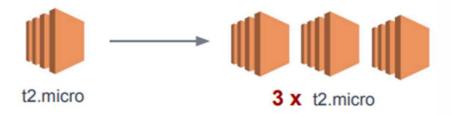
What is Scaling?

Vertical Scaling



Scale Up / Down

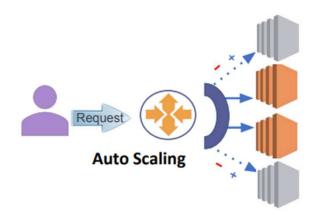
**Horizontal Scaling** 



Scale Out / In



What is Auto Scaling?



- Amazon EC2 Auto Scaling is a component that helps you ensure that you have the correct number of Amazon EC2 instances available to handle the load for your application.
- Auto Scaling adds or removes instances to keep your system steady state.
- You can automate the increasing or decreasing of virtual machines depending on your policy.

Features of Auto Scaling



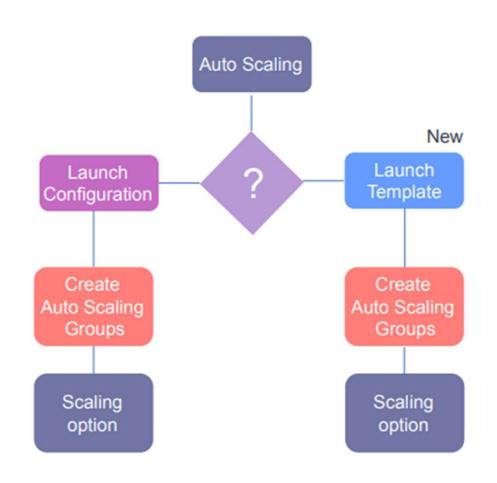


- Auto Scaling Policy
- Launch Configurations or Launch Templates.
- Fault tolerance.
- Compatible with Elastic Load Balancer
- Better Cost Management.



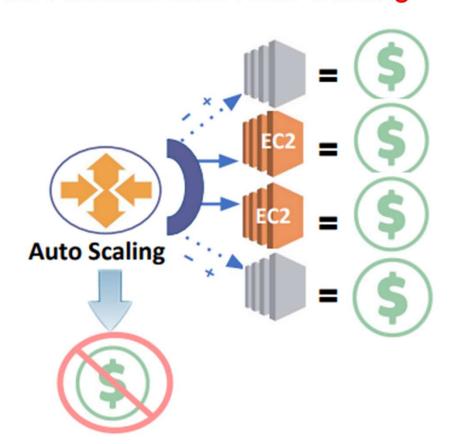
#### **Auto Scaling Creating Process**

- First, you need to select either the Launch Template or the Launch Configuration option and create it.
- Then, create an Auto Scaling Group.
- Finally, Finish Creating Auto Scaling





Pricing for Amazon EC2 Auto Scaling

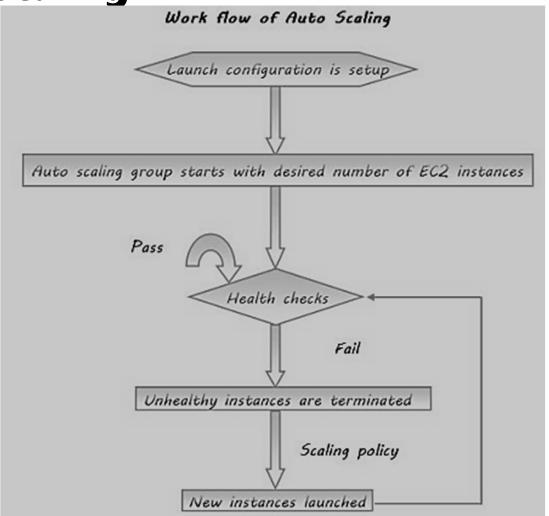






**Load Balancer** 



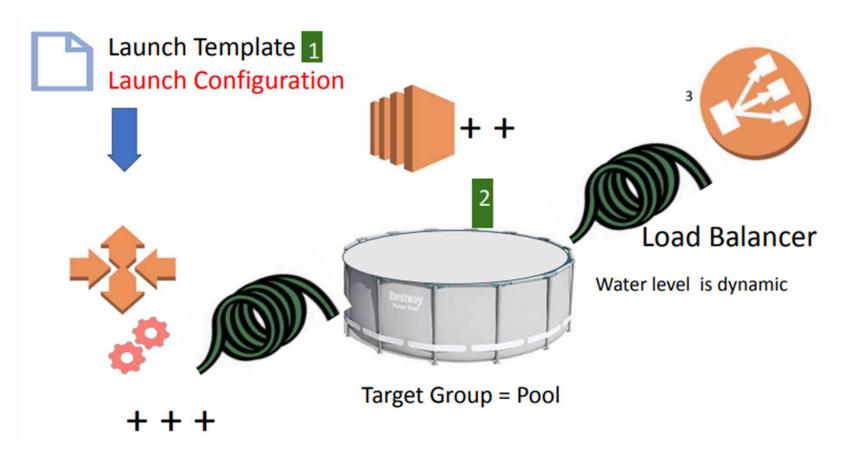




Instances	Desired	Min	Max
Current 2	Will be/ Goal 2	Min limit for Polcy 1	Max limit for Policy 4

# Auto Scaling group Minimum size Scale out as needed Desired capacity Maximum size



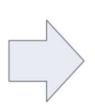




- triggered by exceeding the limits
- we watch out the limits



















#### **Health Check Grace Period & Cooldown Period**

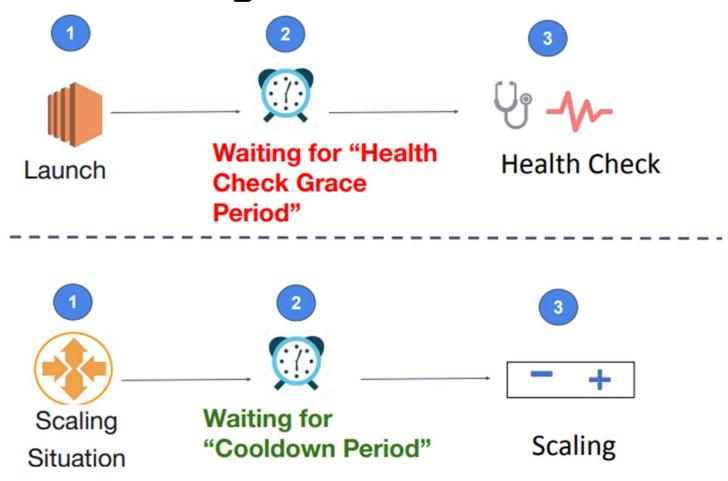
Muddy Sandy Sea Bottom in first seconds of diving



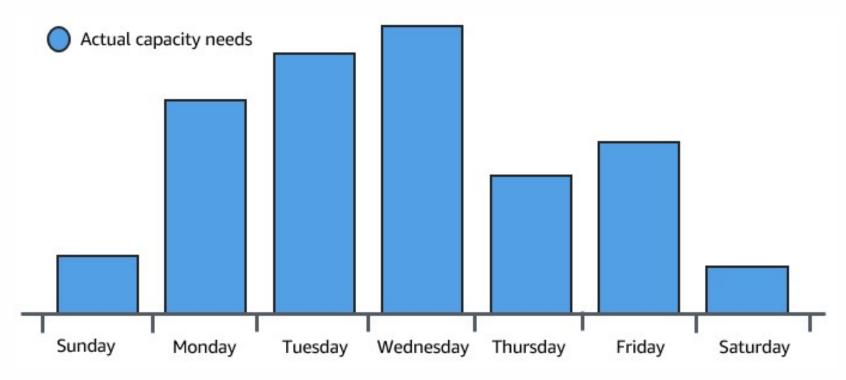
Clear Sea Bottom after waiting for a while



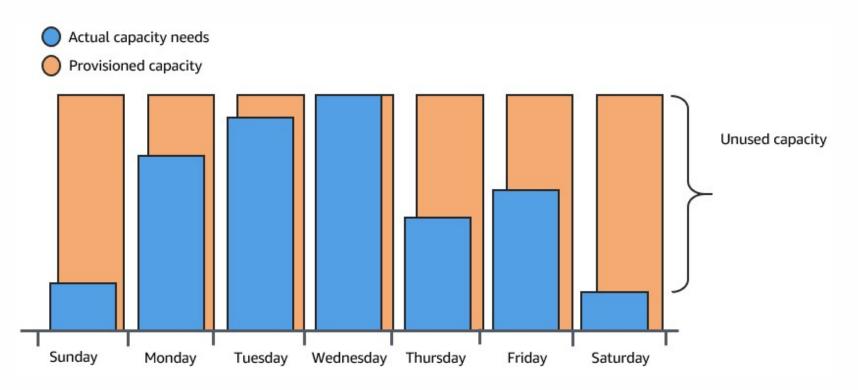




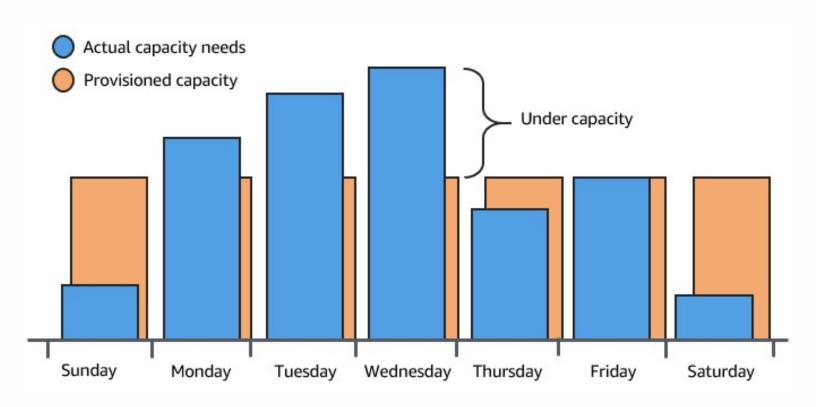




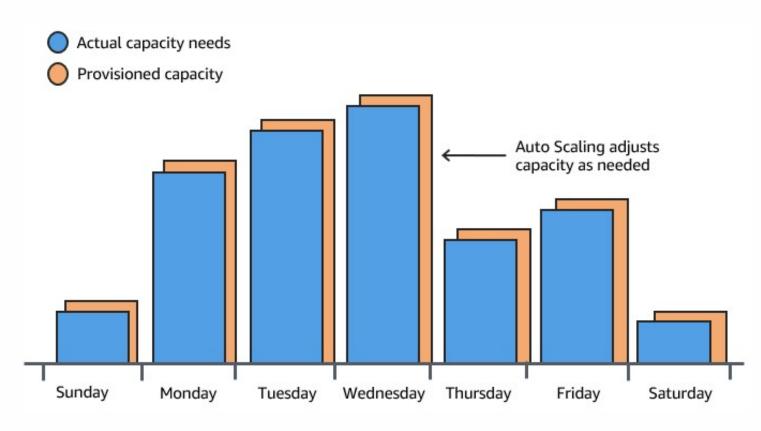
















# Do you have any questions?

Send it to us! We hope you learned something new.

