

# AWS

## Scaling

### what is it?

---

Auto Scaling launches and terminates Amazon EC2 instances automatically according to userdefined policies.

You can use Auto Scaling to maintain a fleet of AWS EC2 instances that can adjust to any presented load. You can also use Auto Scaling to bring up multiple instances in a group at one time.

### Scaling Plans

- **Maintain current level** - this plan checks all running EC2 instances and if it sees that any are unhealthy it will terminate the instance and launch a replacement
- **Manual** - this option allows you to specify the minimum and maximum capacity
- **Scheduled scaling** - if you know when demand is highest you can specify when to scale up
- **Demand based scaling** - build a policy that specifies which parameters need to be met to scale up and down based on demand