

Scaling in AWS

There are 2 type

1. Vertical Scaling

Ability to be able to rightsize a single EC2 instance
1. Rightsizing means

- Upsizing/Upgrading
 - **Example:**
 - When you change an instance type from a smaller one to larger one (moving from a t2.micro to a t3.medium)
- Downsizing/Downgrading
 - **Example:**
 - When you change an instance type from a larger one to smaller one (moving from a t3.medium to a t2.micro)

2. Horizontal Scaling

Ability to rightsize a cluster of ec2 instance cluster

- Scaling Out
 - By increasing the number of instances required to perform a certain task
 - **Example:**
 - Applying Auto Scaling group to a cluster of instances in order to improve speed / decrease latency or a website
- Scaling In
 - By decreasing the number of instances required to perform a certain
 - **Example:**
 - Configuring Auto Scaling Group to reduce number of instances because of less traffic