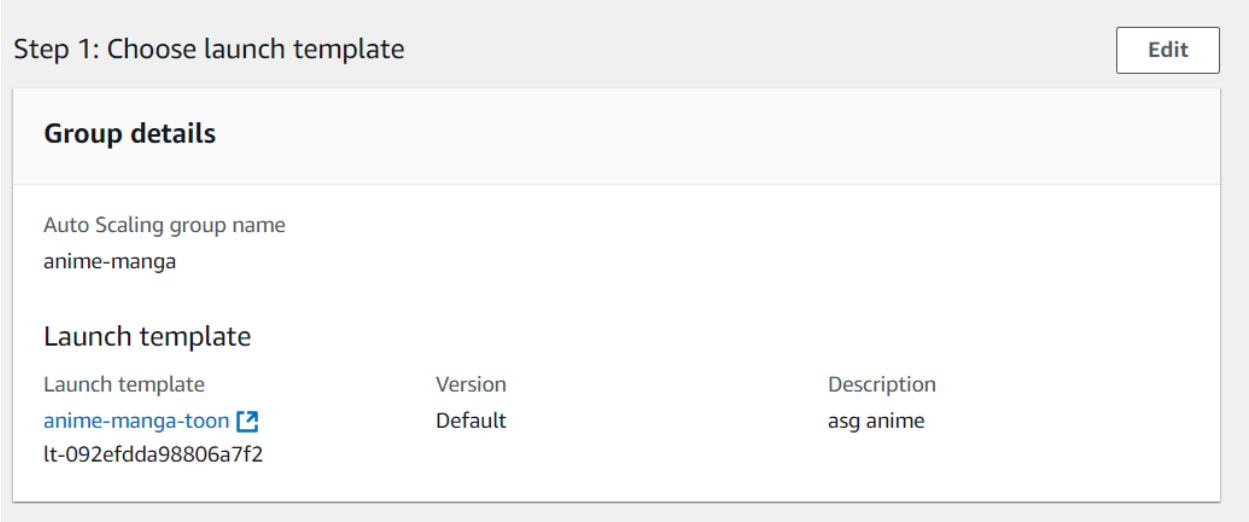
Anime Manga Toon assignment

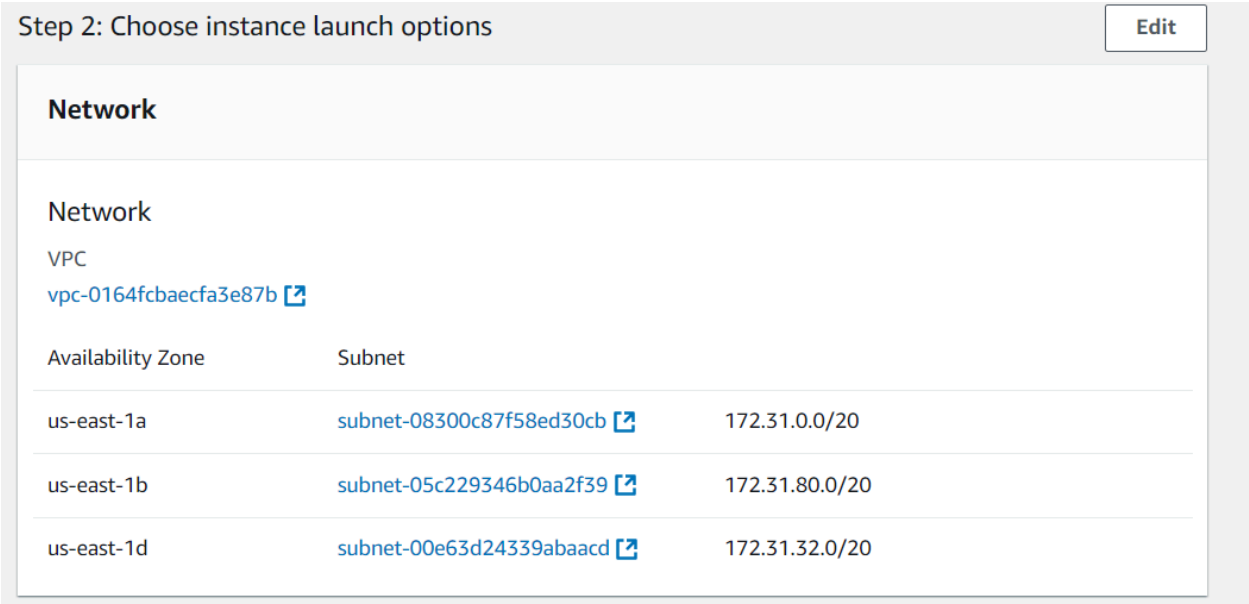
NAME ; DEV SHELKE
EMAIL : shelkedev777@gmail.com

Configuration of Application Load Balancer and Auto Scaling group

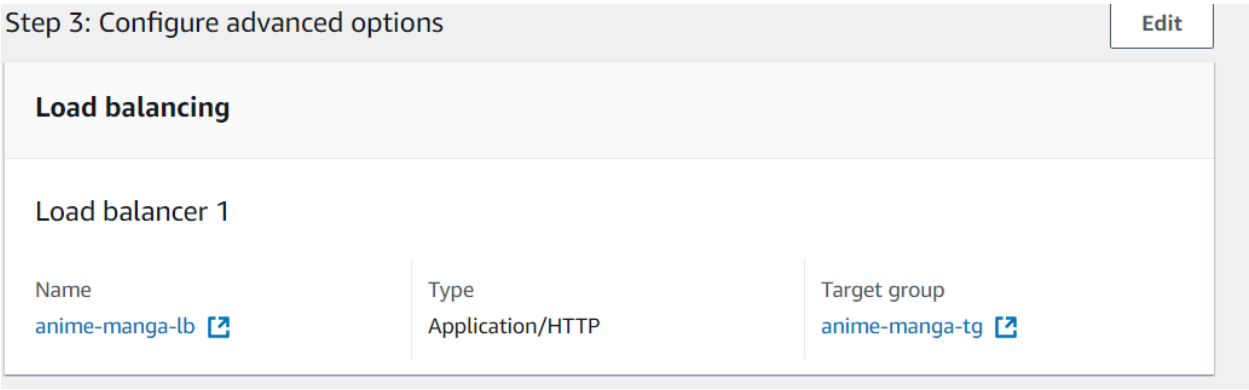
STEP 1: We make an image and template of the existing EC2 instance which has our application running on the nginx server



STEP 2: We select the Availability zones, this makes our application highly available and scalable



STEP 3: We add the Application Load balancer which was created using the target group. The target group has out existing EC2 instance listening our application on port 80.



STEP 4: Here we define policies for our AutoScalingGroup, I have set minimum 1 instance and maximum 3 instances for the webapp.

This policy executes when the CPU reaches utilization of 60%.

Step 4: Configure group size and scaling policies

Edit

Group size

Desired capacity	Desired capacity type
1	Units (number of instances)

Scaling

Minimum desired capacity	Maximum desired capacity	
1	3	
Target tracking policy		
Policy type	Scaling policy name	Execute policy when
Target tracking scaling	Target Tracking Policy	As required to maintain Average CPU utilization at 60
Take the action	Instances need	Scale in
Add or remove capacity units as required	300 seconds to warm up before including in metric	Enabled

Set limits on how much your desired capacity can be increased or decreased.

Min desired capacity

1

Equal or less than desired capacity

Max desired capacity

3

Equal or greater than desired capacity

Automatic scaling - *optional*

Choose whether to use a target tracking policy | Info

You can set up other metric-based scaling policies and scheduled scaling after creating your Auto Scaling group.

- ○ No scaling policies
  Your Auto Scaling group will remain at its initial size and will not dynamically resize to meet demand.
- ◉ Target tracking scaling policy
  Choose a CloudWatch metric and target value and let the scaling policy adjust the desired capacity in proportion to the metric's value.

Scaling policy name

Target Tracking Policy

Metric type | Info

Monitored metric that determines if resource utilization is too low or high. If using EC2 metrics, consider enabling detailed monitoring for better scaling performance.

Average CPU utilization

Target value

60

Instance warmup | Info

300 seconds

☐ Disable scale in to create only a scale-out policy