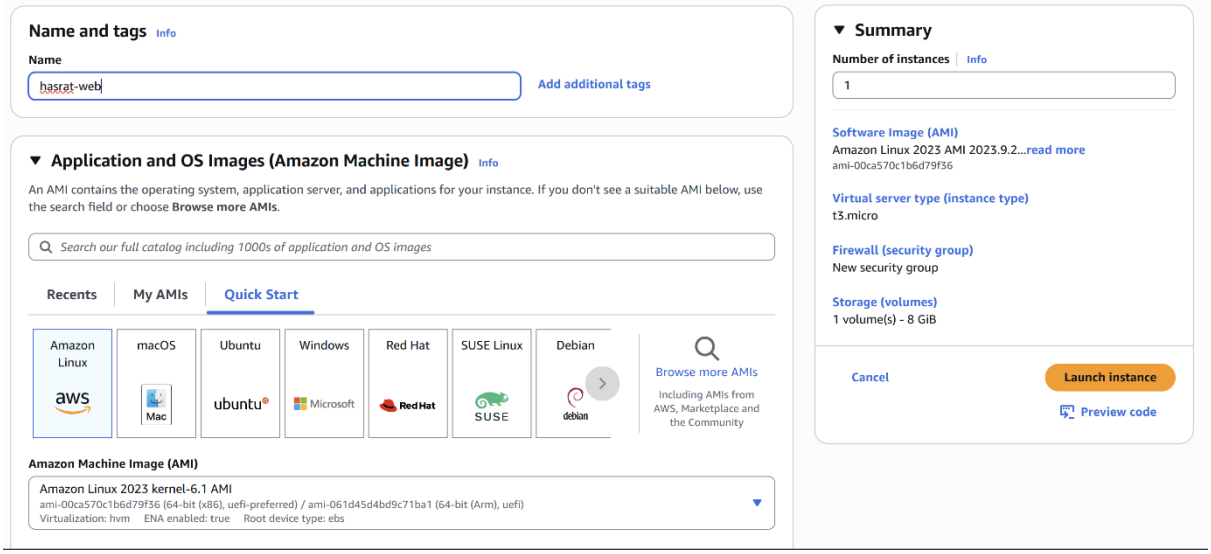
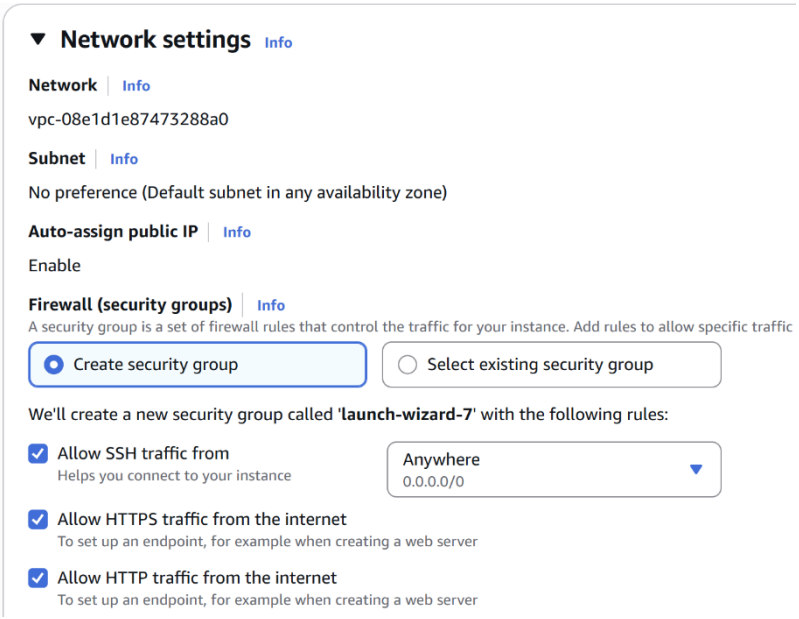
LOAD BALANCER

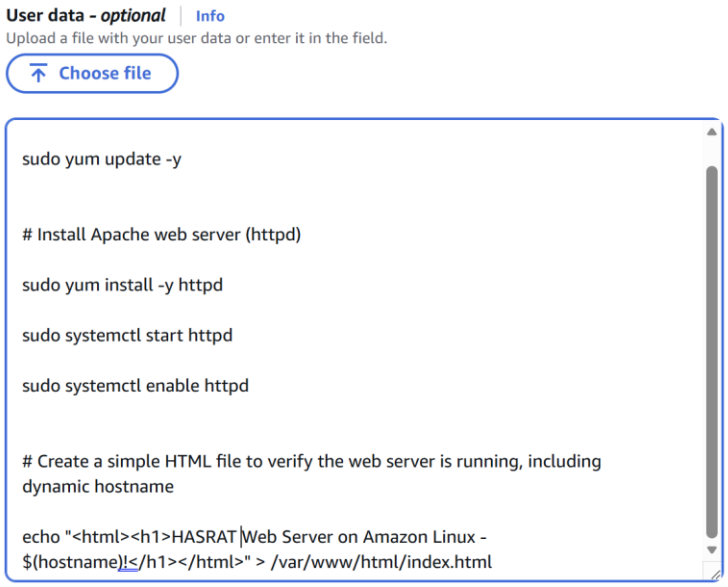
1. Launch 2 EC2 intances



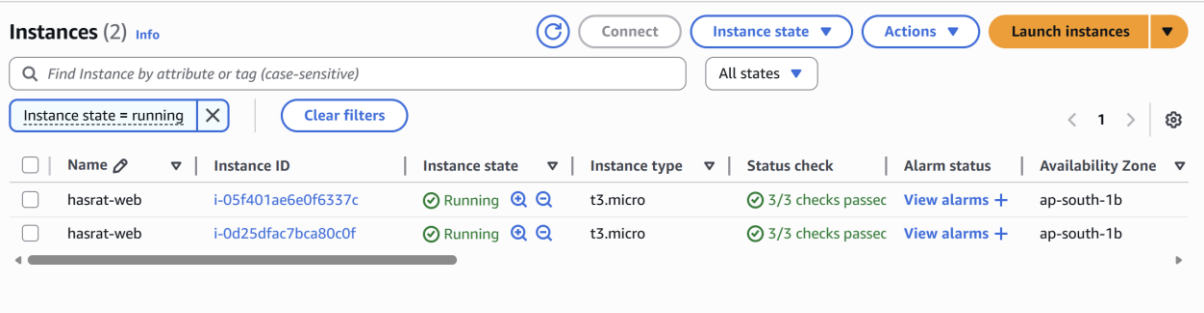
2. Allow http and https from the internet.



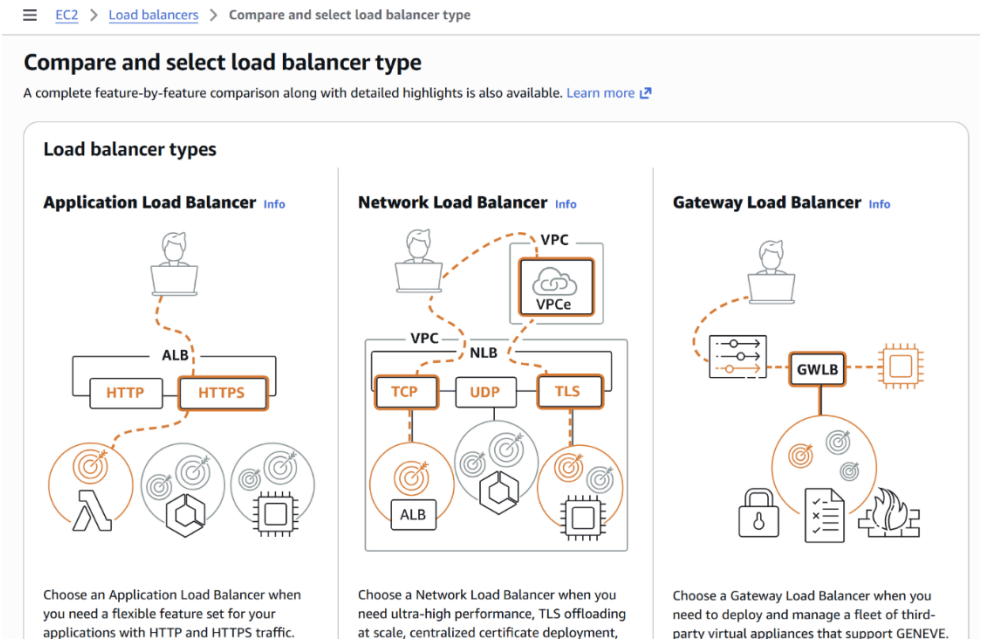
3. In advanced details paste the website script in user data . This will install httpd and create a simple website for both the Ec2 instance



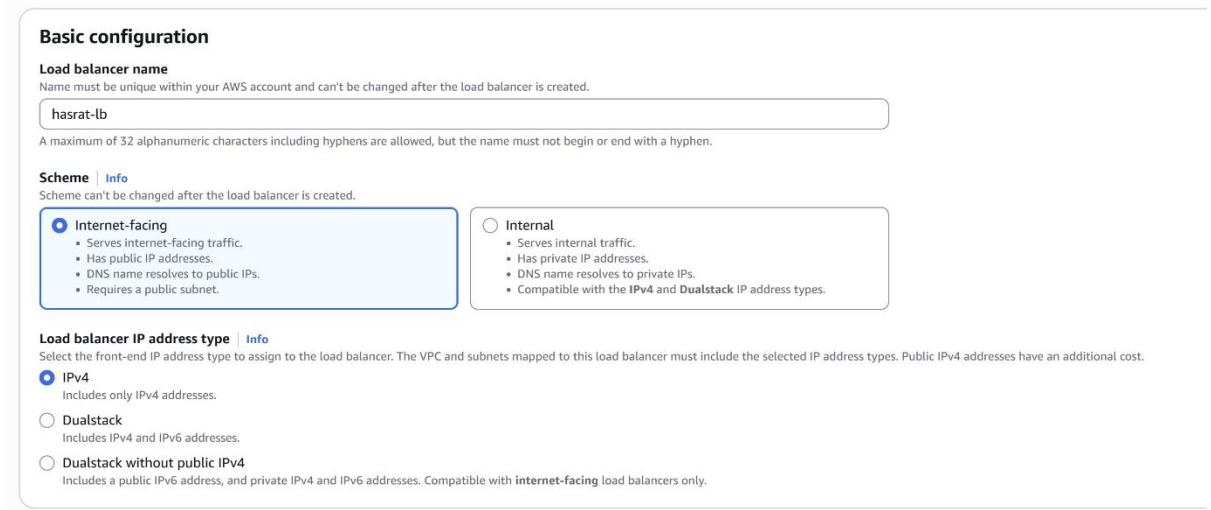
4. Verify



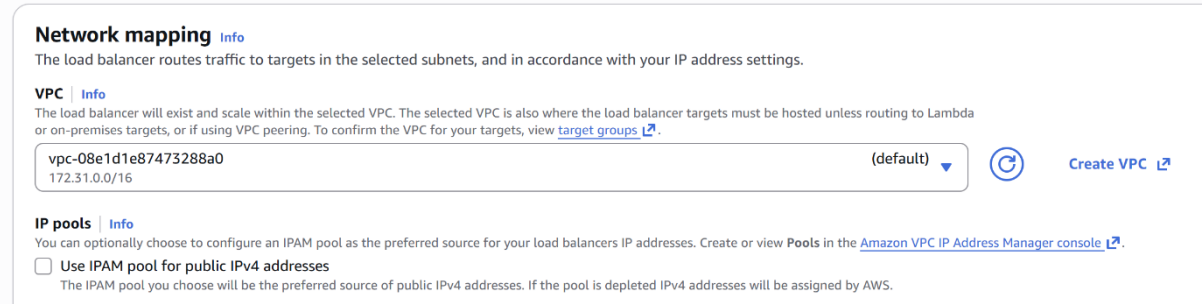
5. Go to Load balancer and select Application Load balancer



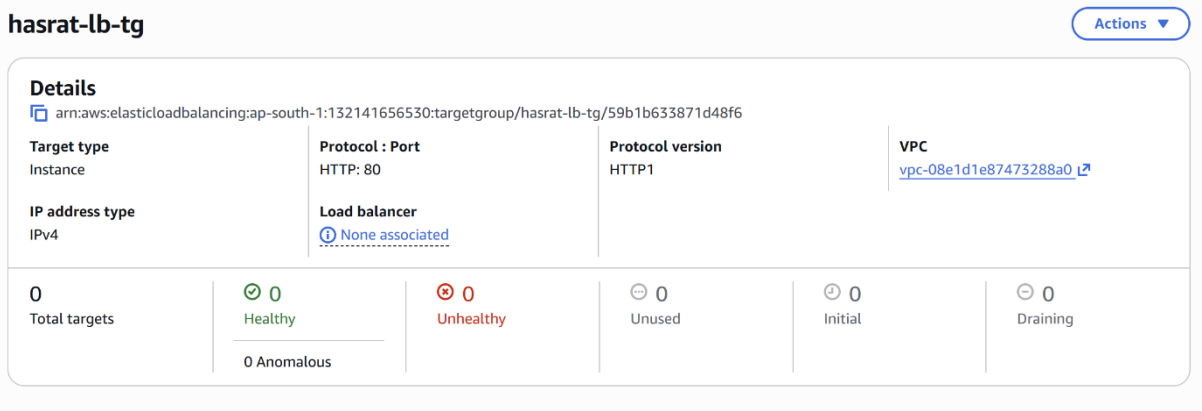
6. Give name :



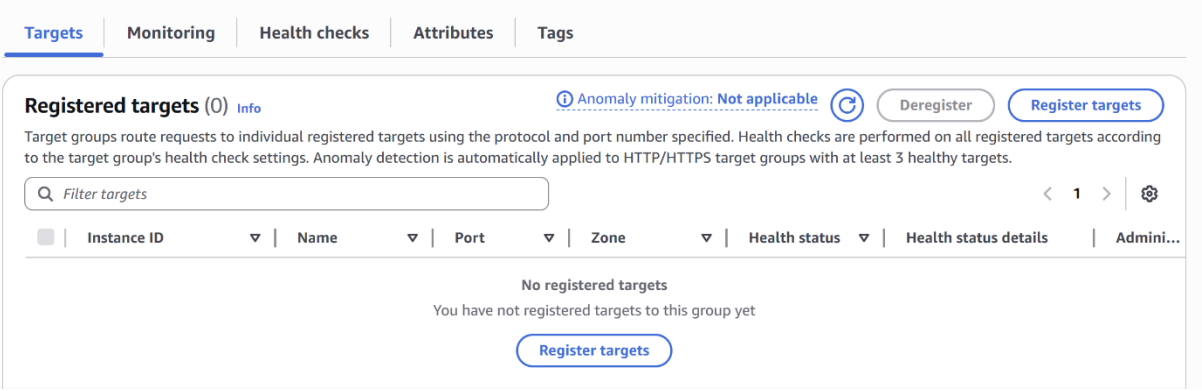
7. In Network mapping click on target group



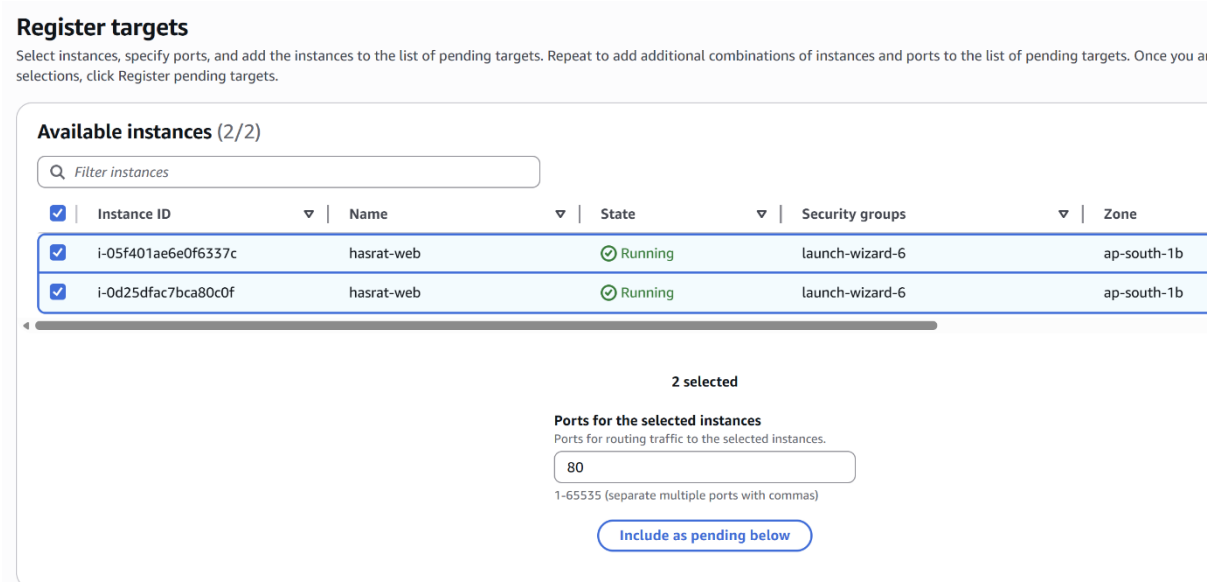
8. Create the Target group with port 80 enabled



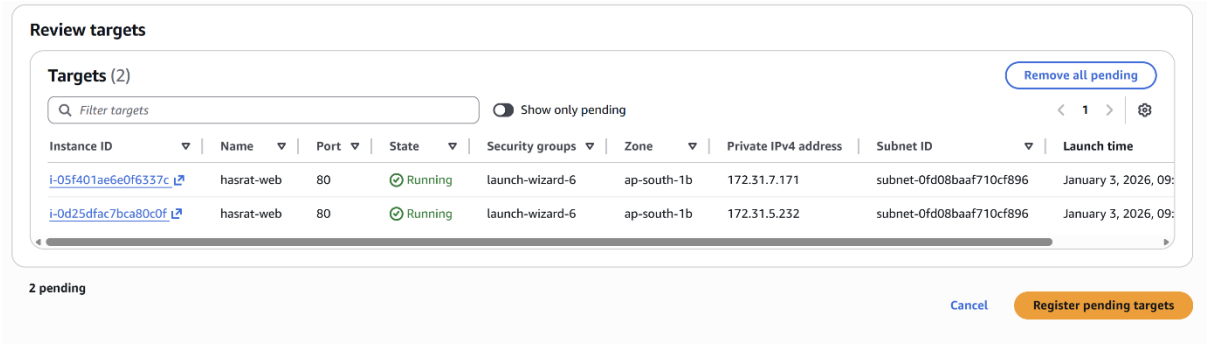
9. Inside Target group clilck on Register Targets



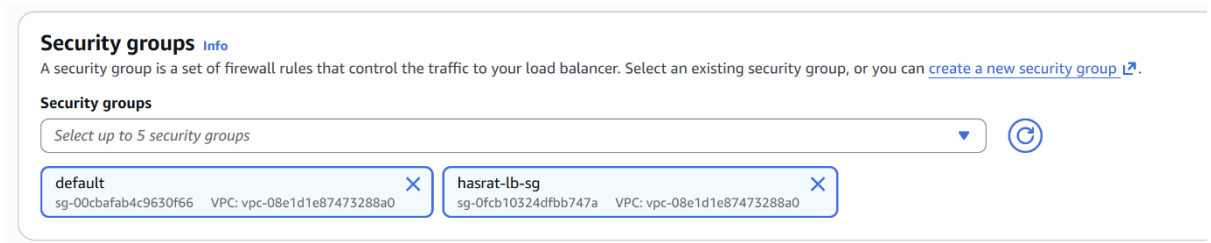
10. Select the Target (EC2)



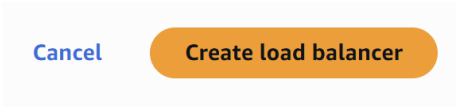
11. Click Register pending targets



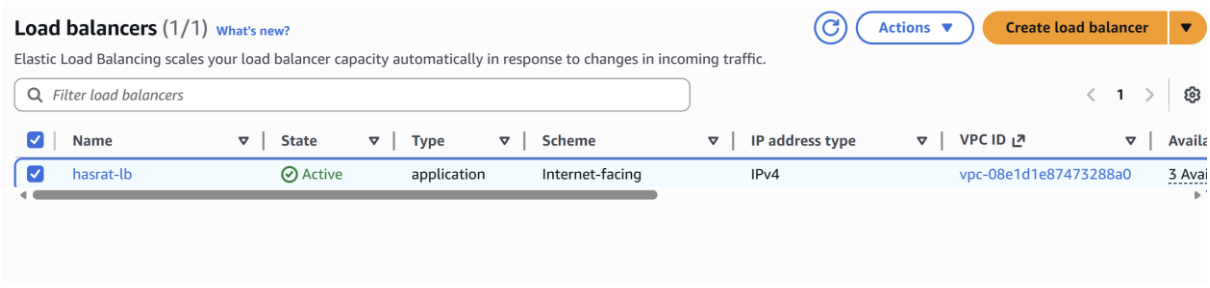
12. Go Back to Load balancer and Create security group with Inbound port 80



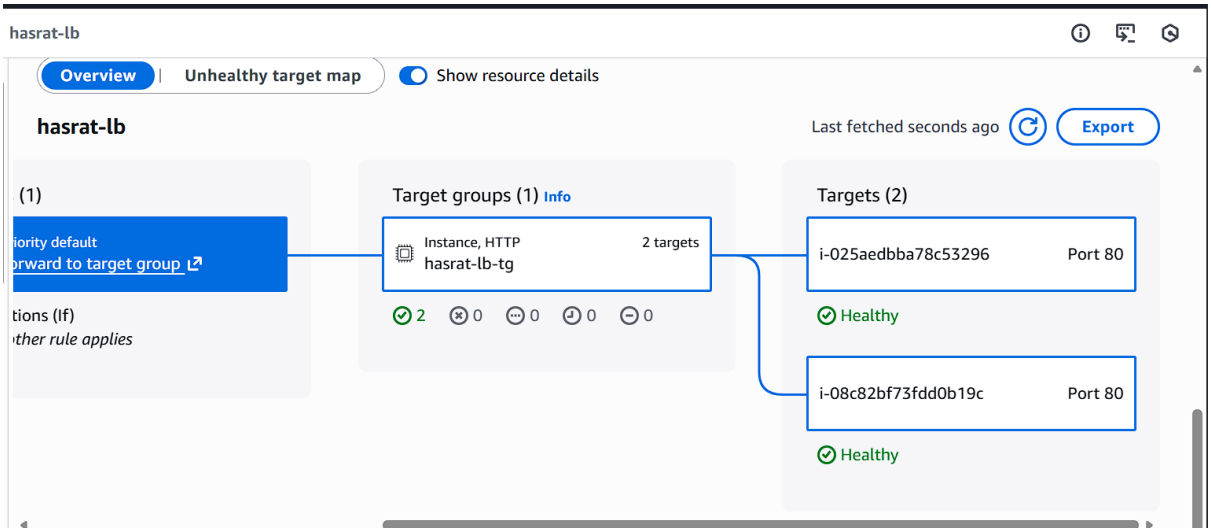
13. Click Create load balancer



14. Click on load balancer we created



15. Go to Resource Mapping to check if the Targeted Ec2 are in Healthy state

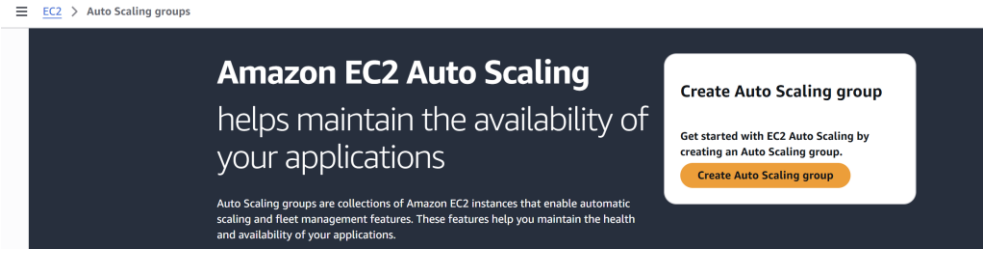


16. Copy the Public IP from EC2 and check if the Website is running and refresh the website to check if the Load balancer is working

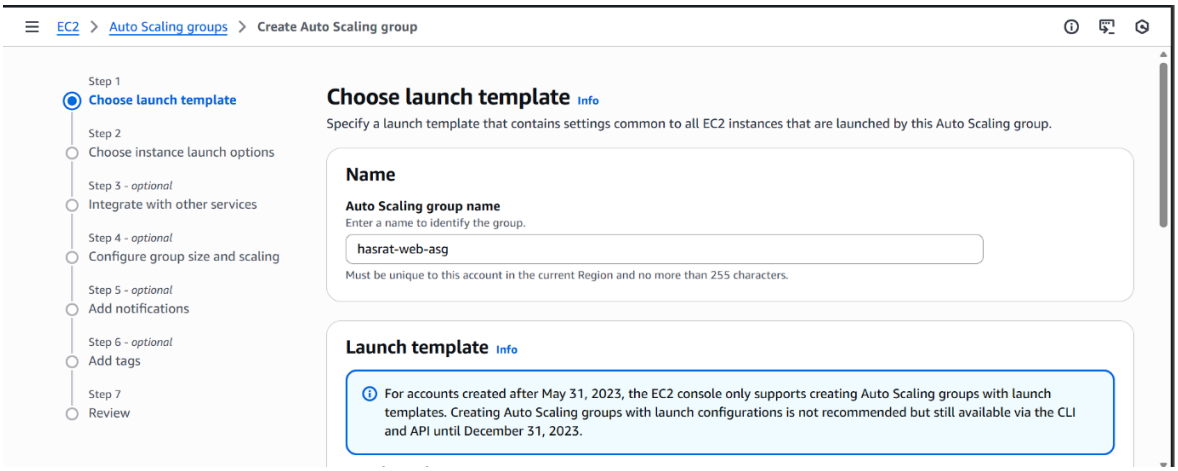
    Here we can see that the Hostname is Changing that means the load balancer is balancing the traffic between two targeted Ec2 intances

AUTO SCALING

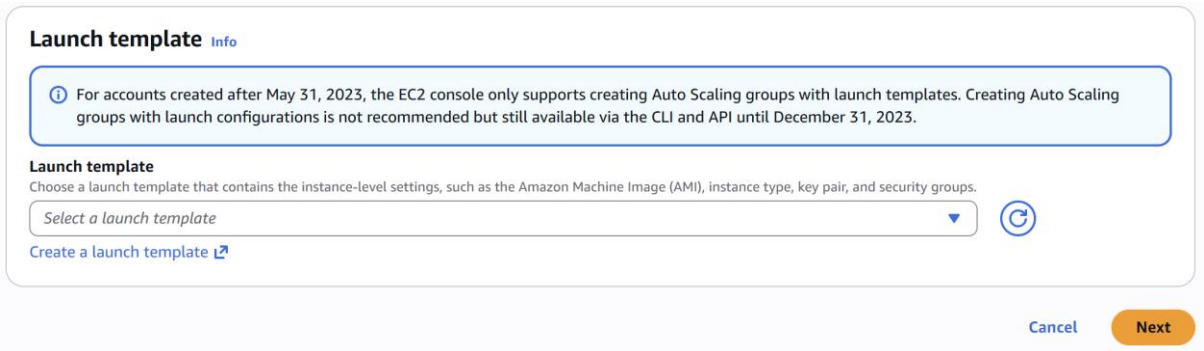
1. Inside Ec2 , Create Auto scaling Group



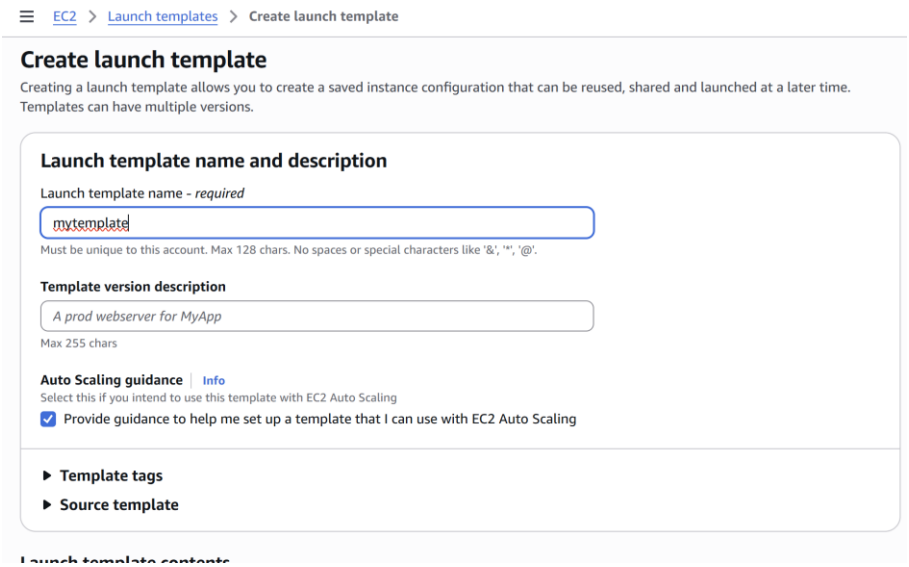
2. Give Name :



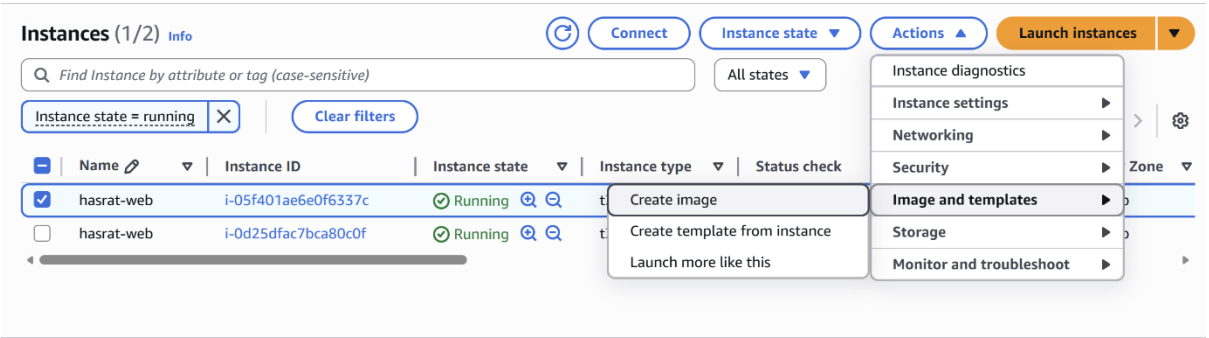
3. Click Create a Launch Template

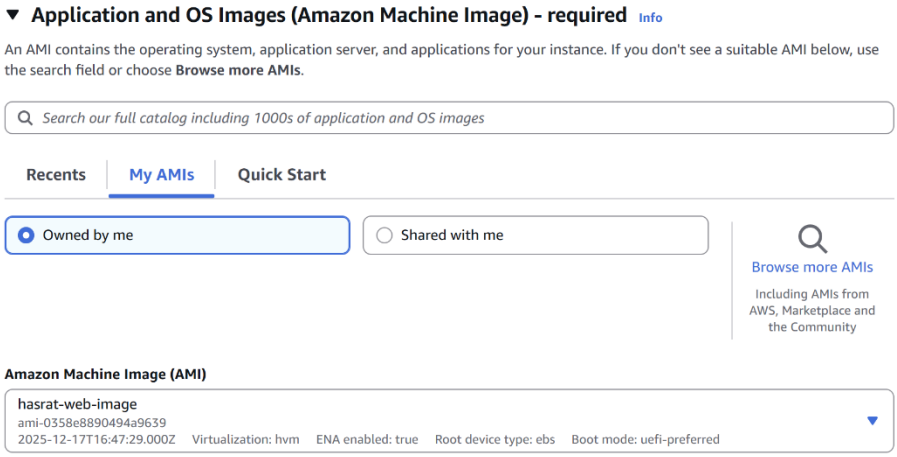


4. Give name to the template

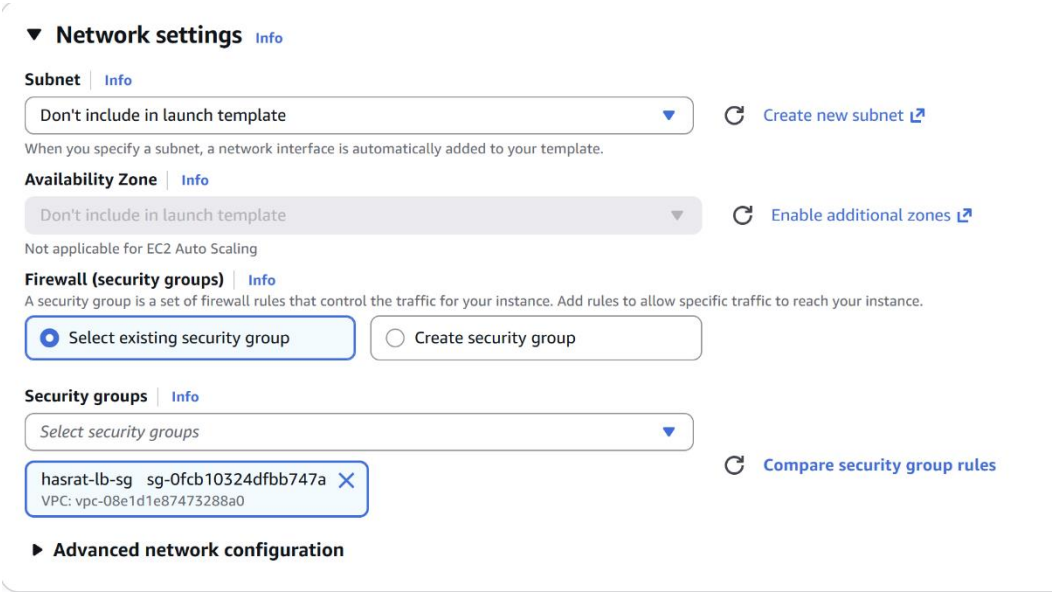


5. Create the AMI of any one of the Ec2 intance and select the AMI

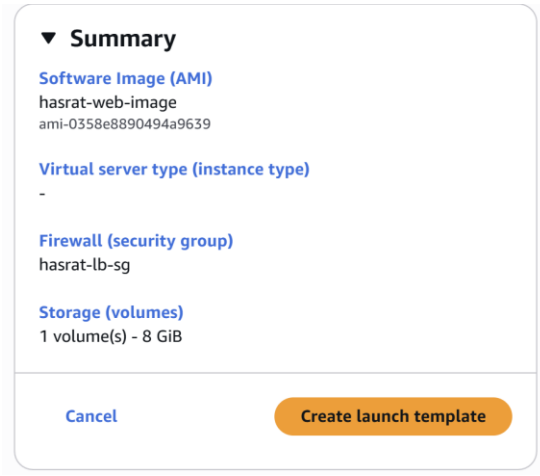




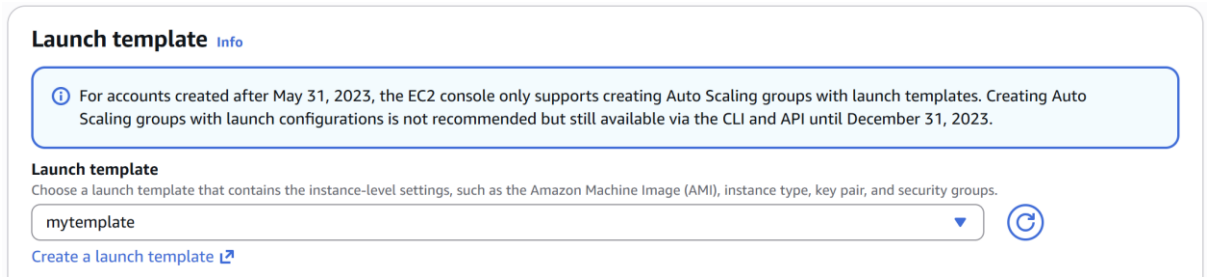
6. In Network Setting , Select the existing load balancer security group



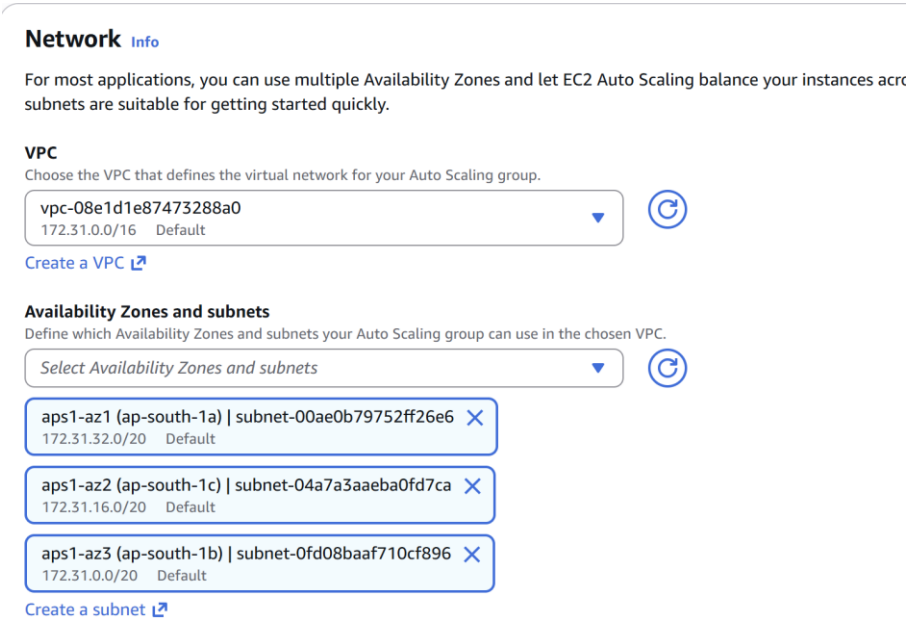
7. Review and Create the Template



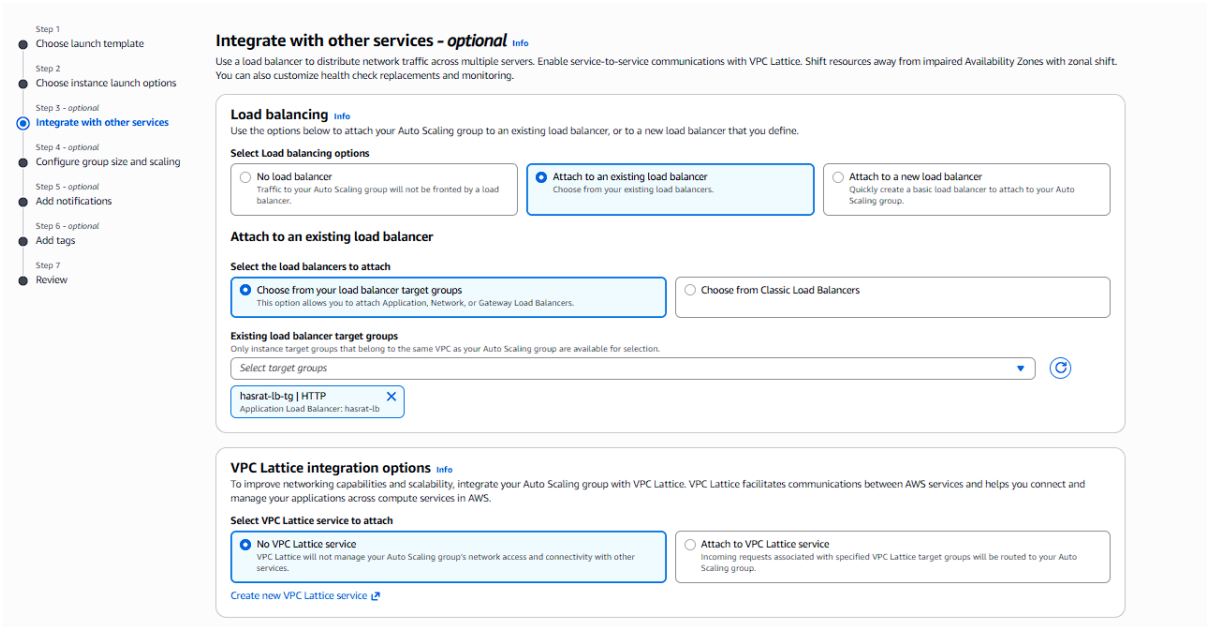
8. Go back to Auto Scaling and Select the Template we created and Click Next



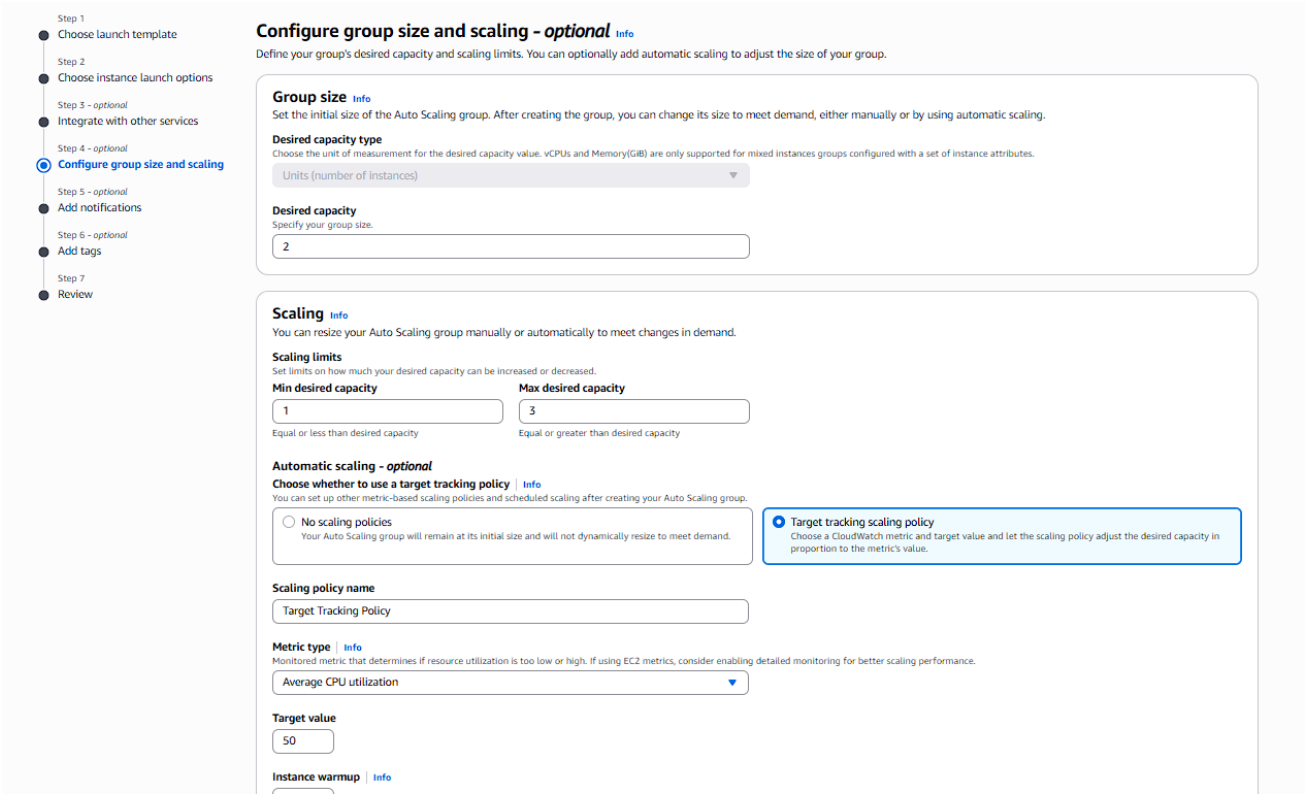
9. In Network Tab select the VPC same as Load balancer and choose the availability zones and click Next



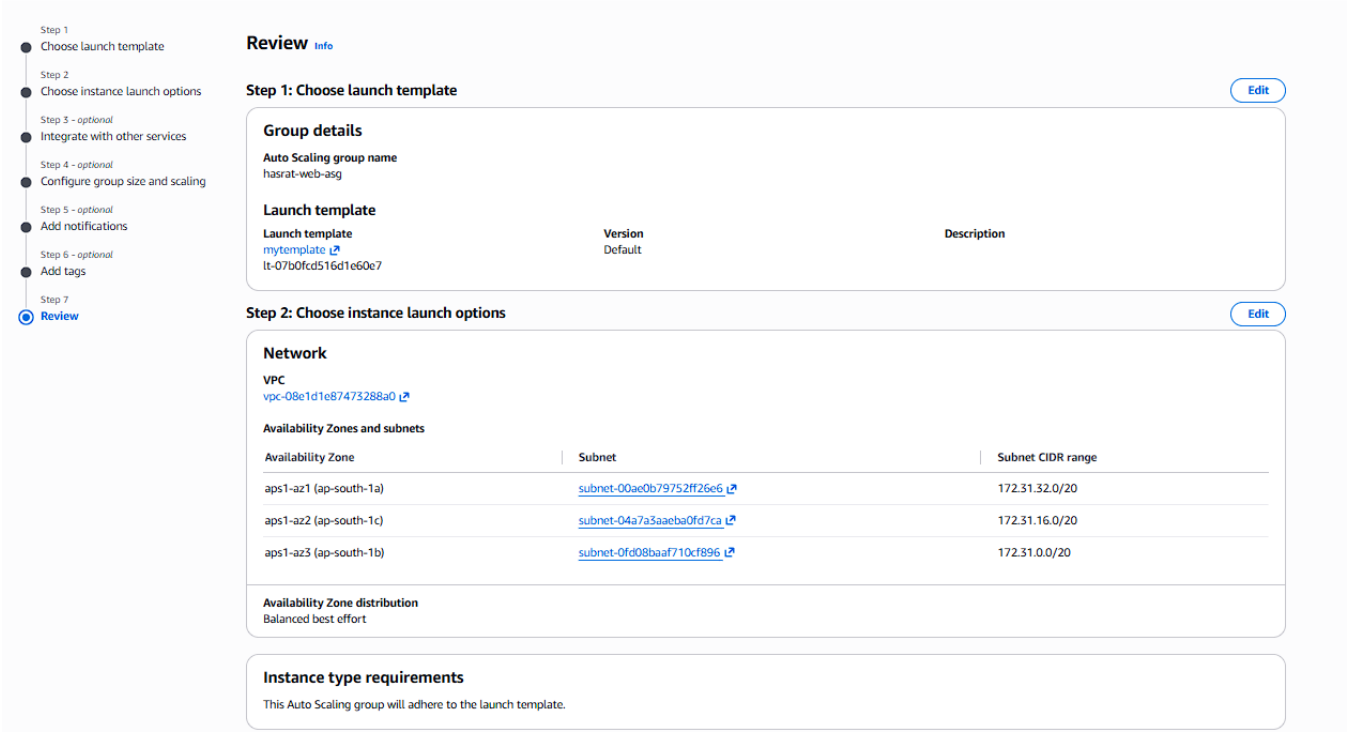
10. Select the Following Options



11. Select 2 Desired Capacity for auto scaling



12. Review and Create



Step 3: Integrate with other services

Edit

Load balancing

Load balancer 1

Name	Type	Target group
hasrat-lb	Application/HTTP	hasrat-lb-tg

VPC Lattice integration options

VPC Lattice target groups
-

Application Recovery Controller (ARC) zonal shift

ARC zonal shift
Disabled

Health checks

Health check type	Health check grace period
EC2	300 seconds

Step 4: Configure group size and scaling policies

Edit

Group size

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

Step 4: Configure group size and scaling policies

Edit

Group size

Desired capacity	Desired capacity type
2	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 50
Take the action	**Instances need**	**Scale in**
Add or remove capacity units as required	300 seconds to warm up before including in metric	Enabled

Instance maintenance policy

Replacement behavior	Min healthy percentage	Max healthy percentage
No policy	-	-

Additional settings

Instance scale-in protection	Monitoring	Default instance warmup
Disabled	Disabled	Disabled

Capacity Reservation preference

Preference	Capacity Reservation IDs	Resource Groups
Default	-	-

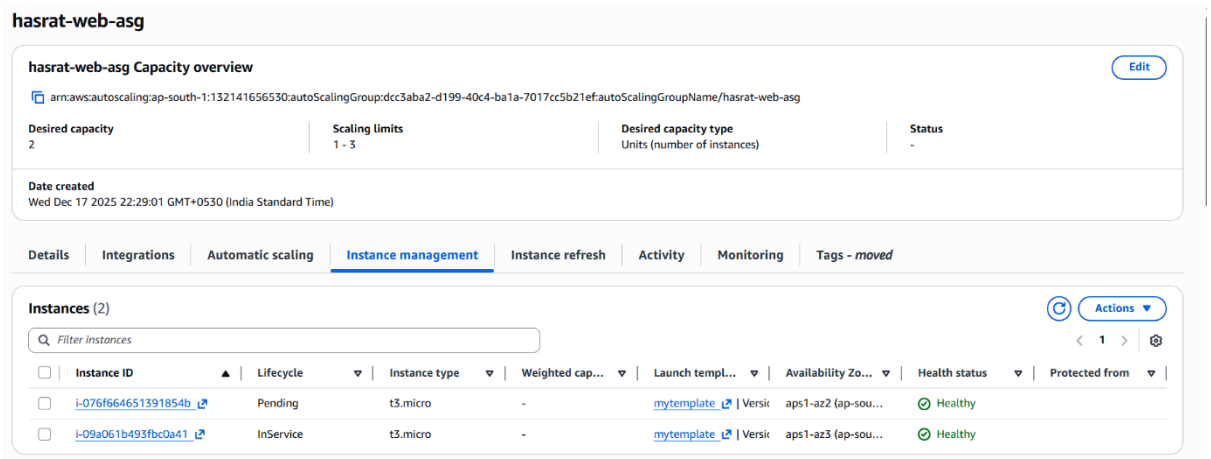
Step 5: Add notifications

Edit

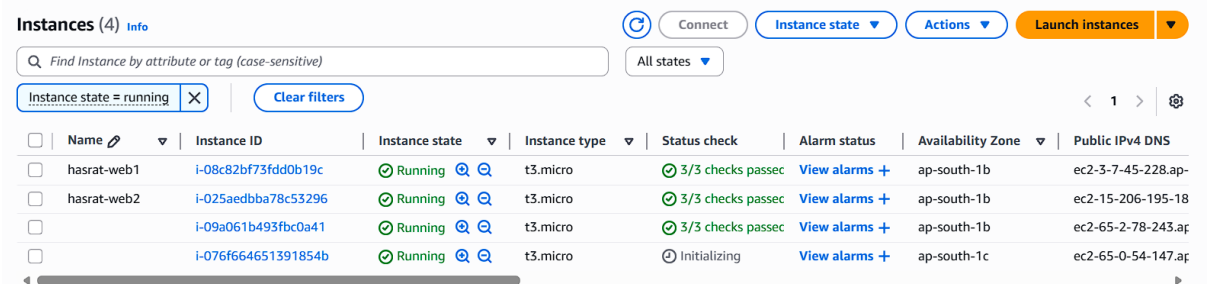
Notifications

No notifications

13. Here we can see the Auto scaling group has created two intances because we have ordered 2 desired capacity



14. We can also see the extra intance created by auto scaling group without a name inside the intances tab



15. Here is the Logs of auto scaling group , maintaining the number of instance. when one instance is terminated ASG has launched another instance in response