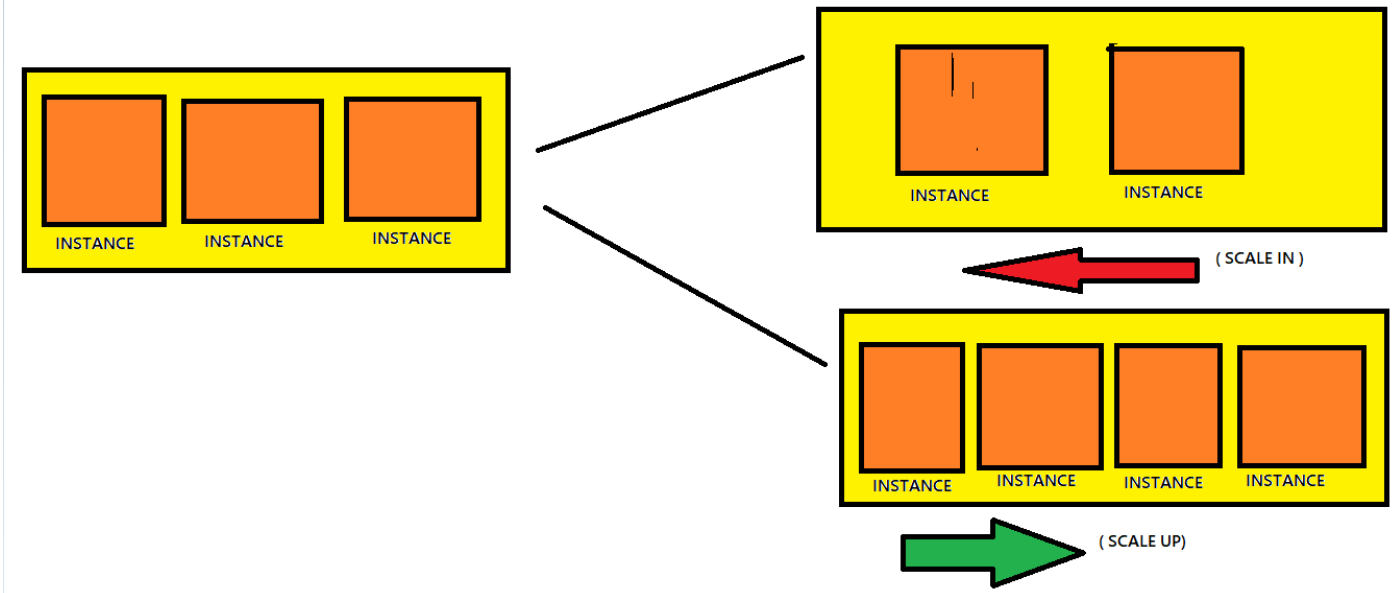
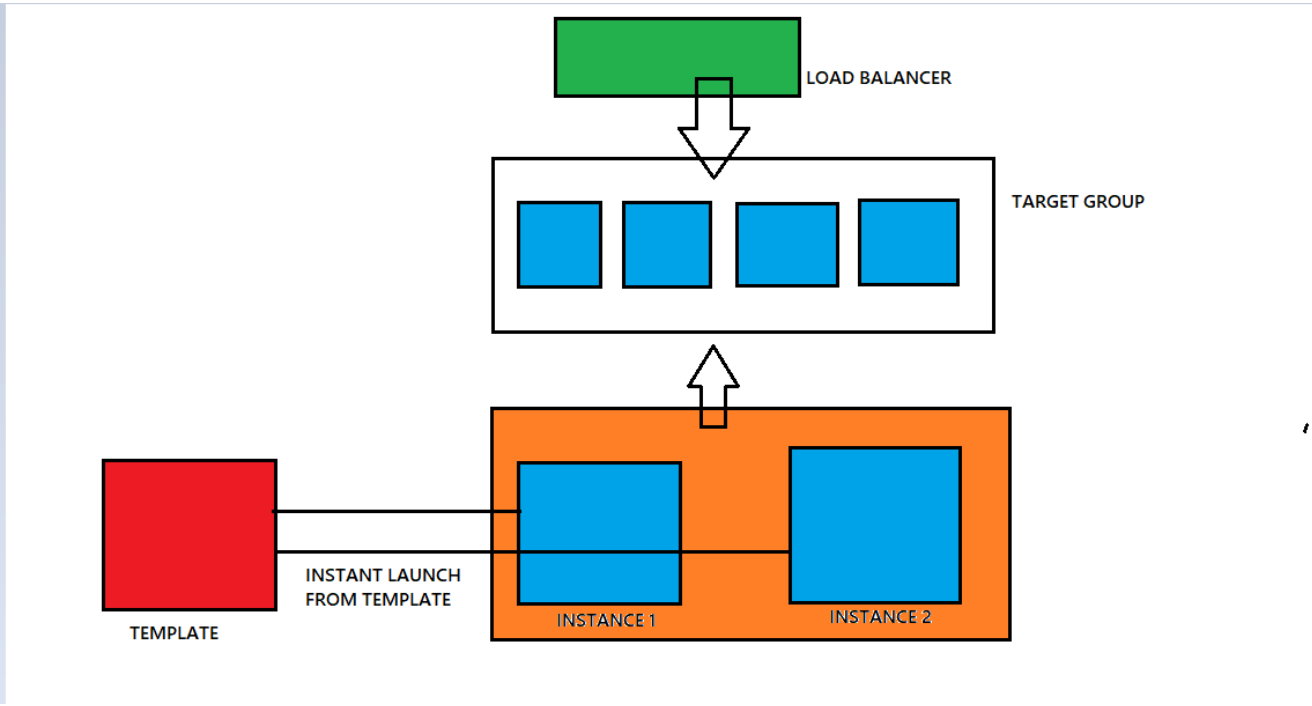
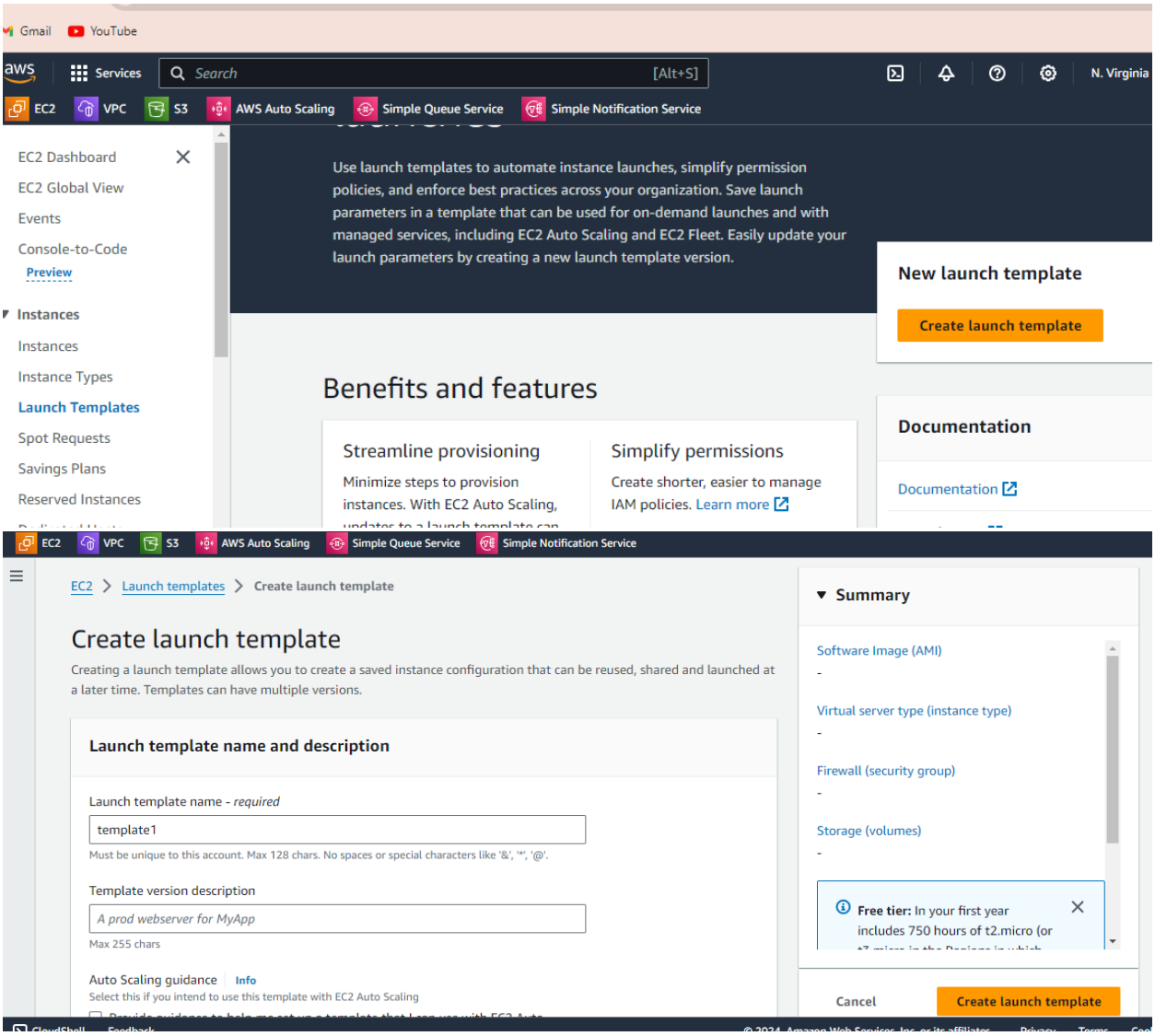
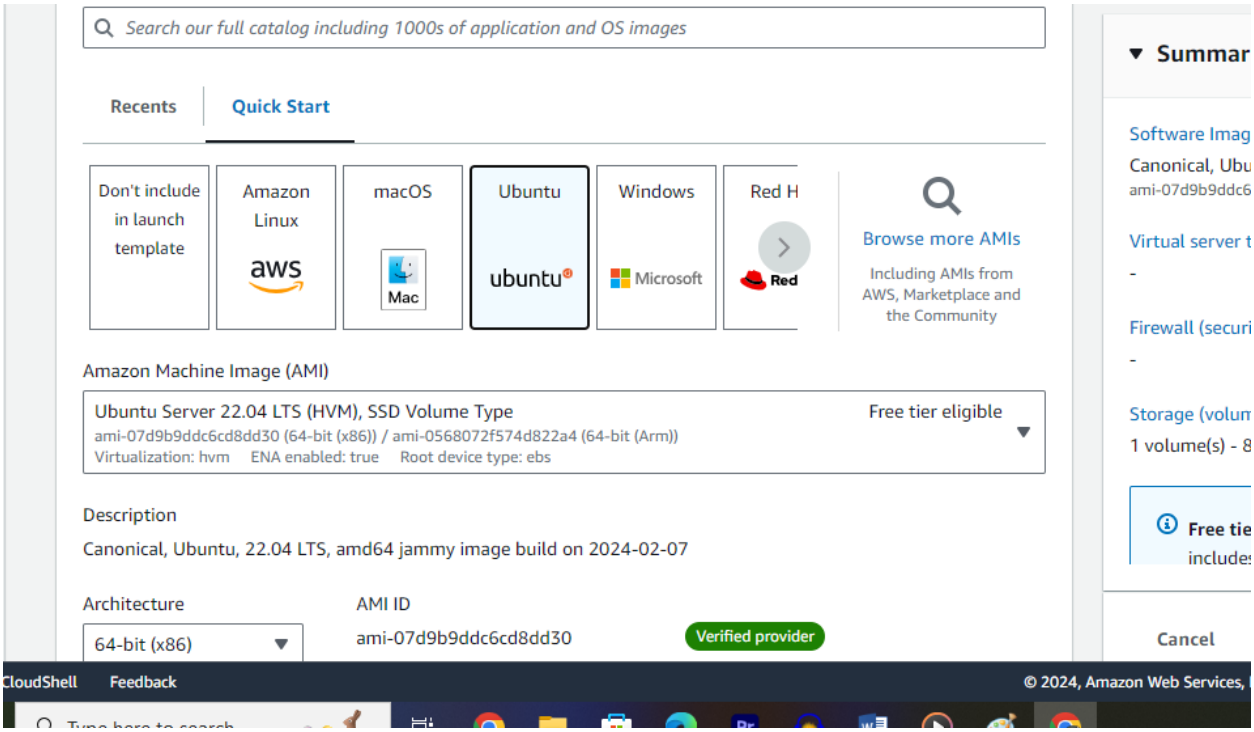
AIM- CONCEPT OF AUTO-SCALING





STEP 1- LAUNCH A TEMPLATE


Gmail
YouTube
aws
Services
Search
[Alt+S]
N. Virginia
EC2
VPC
S3
AWS Auto Scaling
Simple Queue Service
Simple Notification Service
EC2 Dashboard
EC2 Global View
Events
Console-to-Code
Preview
Instances
Instances
Instance Types
Launch Templates
Spot Requests
Savings Plans
Reserved Instances
Use launch templates to automate instance launches, simplify permission policies, and enforce best practices across your organization. Save launch parameters in a template that can be used for on-demand launches and with managed services, including EC2 Auto Scaling and EC2 Fleet. Easily update your launch parameters by creating a new launch template version.
New launch template
Create launch template
Benefits and features
Streamline provisioning
Minimize steps to provision instances. With EC2 Auto Scaling,
Simplify permissions
Create shorter, easier to manage IAM policies. Learn more
Documentation
Documentation
EC2
VPC
S3
AWS Auto Scaling
Simple Queue Service
Simple Notification Service
EC2 > Launch templates > Create launch template
Create launch template
Creating a launch template allows you to create a saved instance configuration that can be reused, shared and launched at a later time. Templates can have multiple versions.
Launch template name and description
Launch template name - required
template1
Must be unique to this account. Max 128 chars. No spaces or special characters like '&', '*', '@'.
Template version description
A prod webserver for MyApp
Max 255 chars
Auto Scaling guidance
Info
Select this if you intend to use this template with EC2 Auto Scaling
Summary
Software Image (AMI)
-
Virtual server type (instance type)
-
Firewall (security group)
-
Storage (volumes)
-
Free tier: In your first year includes 750 hours of t2.micro (or
Cancel
Create launch template



Canonical, Ubuntu, 22.04 LTS, amd64 jammy image build on 2024-02-07

Architecture

64-bit (x86)

AMI ID

ami-07d9b9ddc6cd8dd30

Verified provider

Instance type Info | Get advice

Advanced

Instance type

t2.micro

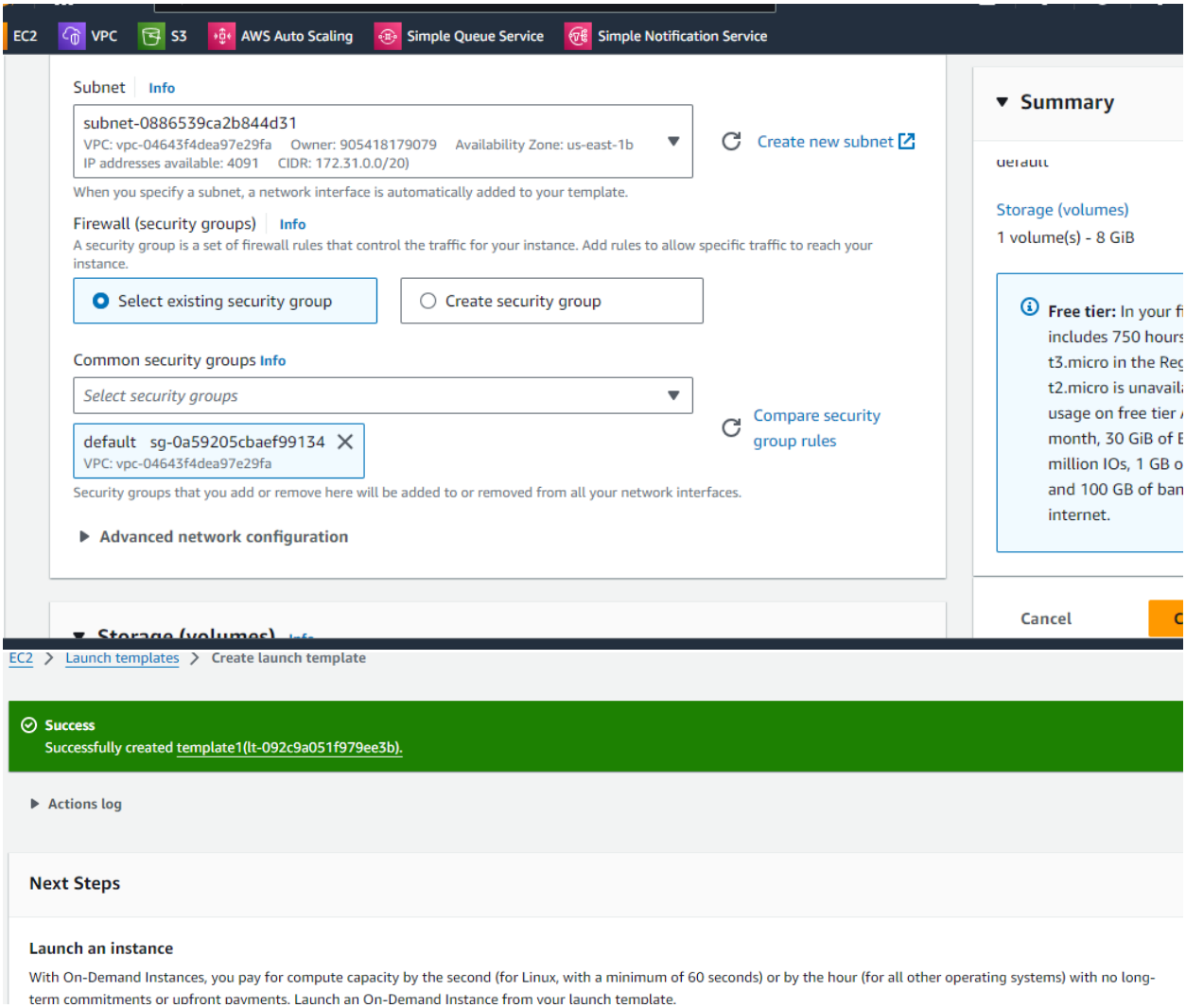
Free tier eligible

Family: t2 1 vCPU 1 GiB Memory Current generation: true
On-Demand Windows base pricing: 0.0162 USD per Hour
On-Demand SUSE base pricing: 0.0116 USD per Hour
On-Demand RHEL base pricing: 0.0716 USD per Hour
On-Demand Linux base pricing: 0.0116 USD per Hour

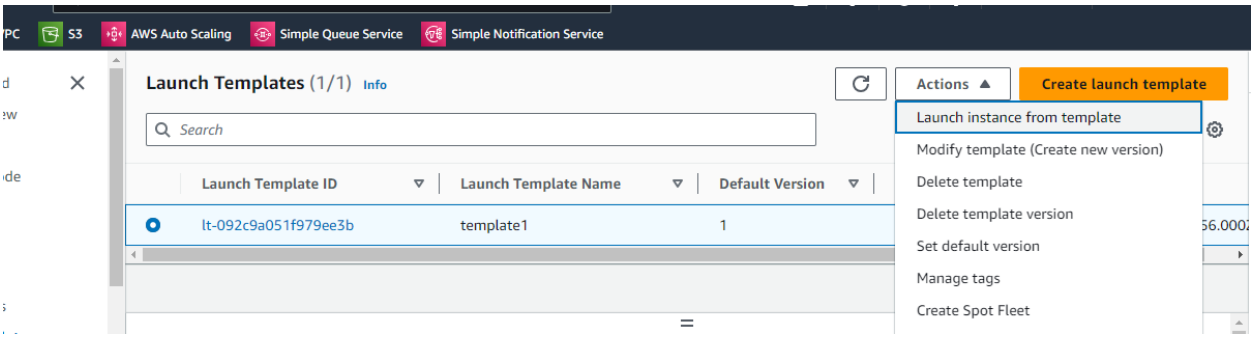
All generations

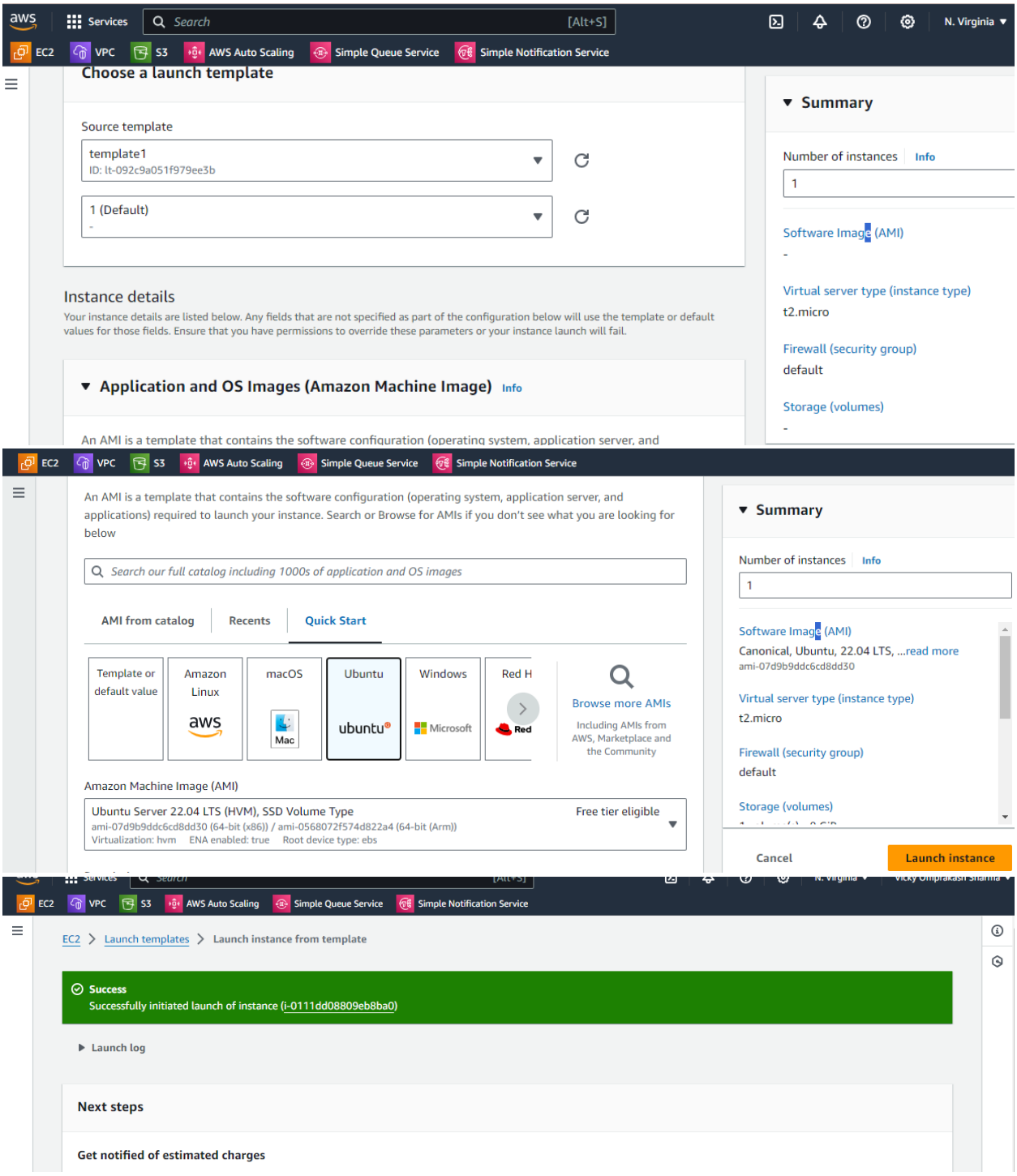
Compare instance types

Additional costs apply for AMIs with pre-installed software

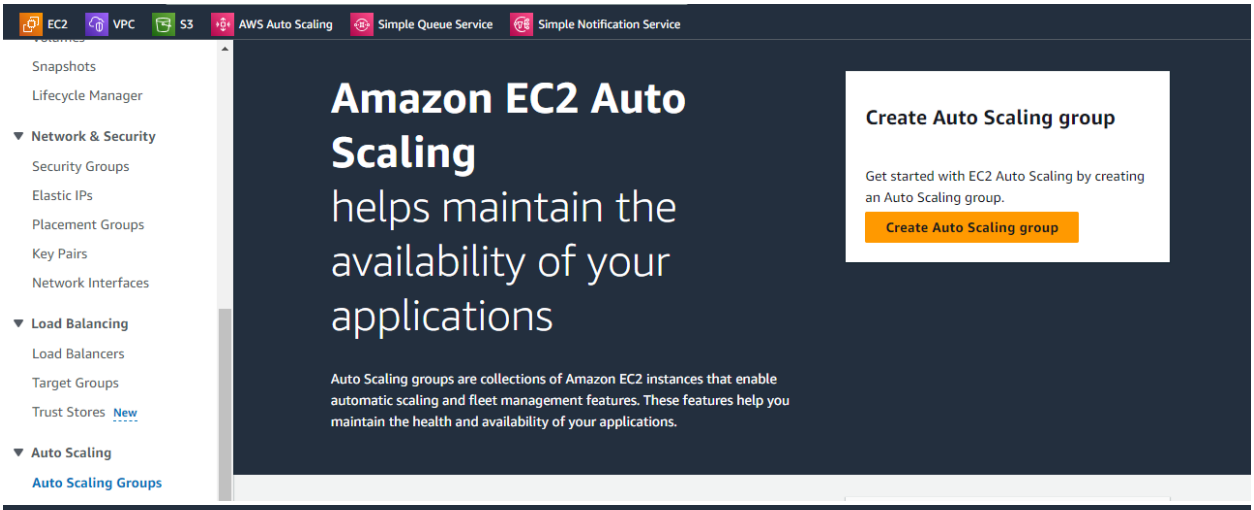


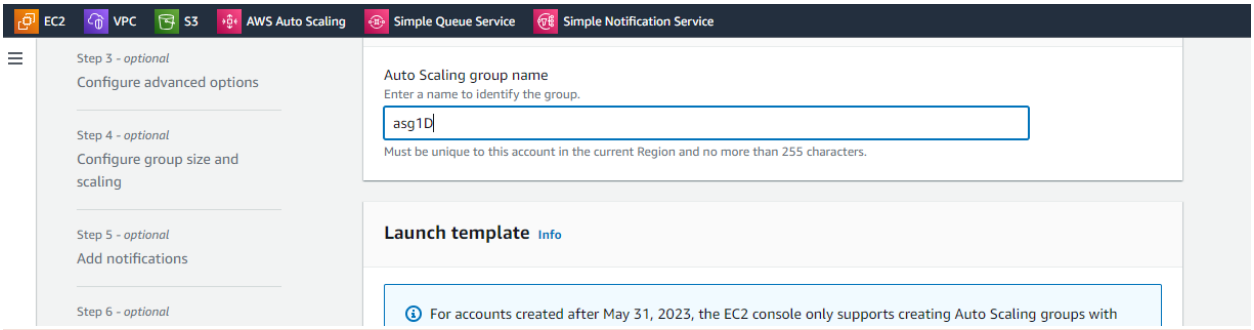
Step 2- launch instance from template

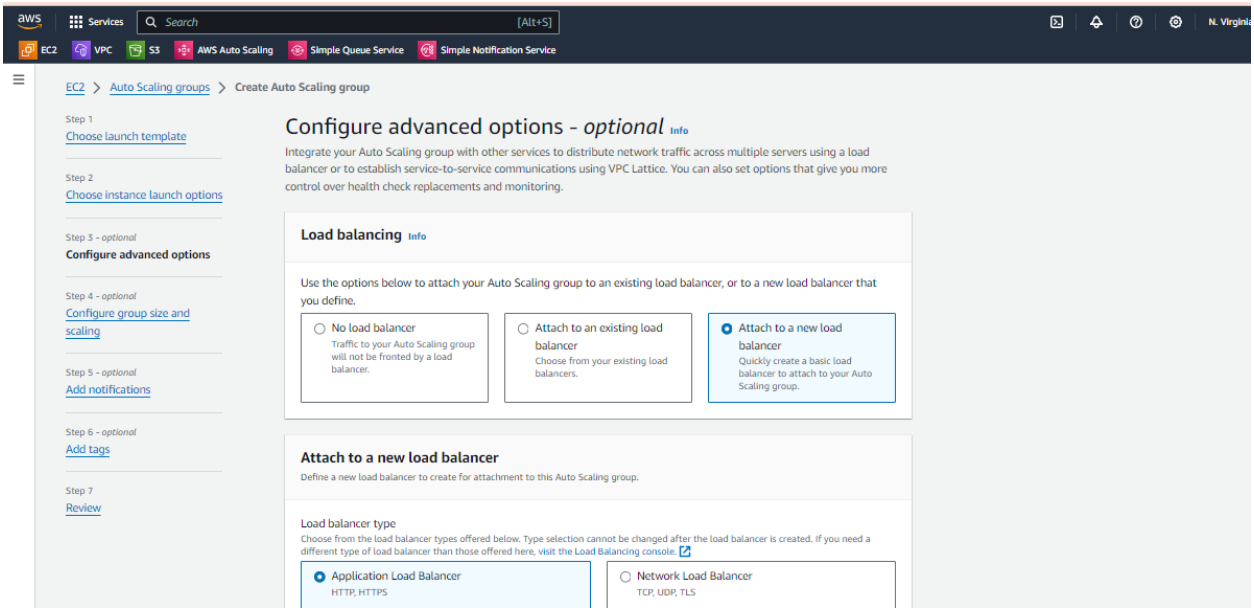




Step 3- now go to auto-scaling


EC2
VPC
S3
AWS Auto Scaling
Simple Queue Service
Simple Notification Service
Snapshots
Lifecycle Manager
Network & Security
Security Groups
Elastic IPs
Placement Groups
Key Pairs
Network Interfaces
Load Balancing
Load Balancers
Target Groups
Trust Stores New
Auto Scaling
Auto Scaling Groups
Amazon EC2 Auto Scaling
helps maintain the availability of your applications
Auto Scaling groups are collections of Amazon EC2 instances that enable automatic scaling and fleet management features. These features help you maintain the health and availability of your applications.
Create Auto Scaling group
Get started with EC2 Auto Scaling by creating an Auto Scaling group.
Create Auto Scaling group


EC2
VPC
S3
AWS Auto Scaling
Simple Queue Service
Simple Notification Service
Step 3 - optional
Configure advanced options
Step 4 - optional
Configure group size and scaling
Step 5 - optional
Add notifications
Step 6 - optional
Auto Scaling group name
Enter a name to identify the group.
asg1D
Must be unique to this account in the current Region and no more than 255 characters.
Launch template Info
For accounts created after May 31, 2023, the EC2 console only supports creating Auto Scaling groups with


aws
Services
Search
[Alt+S]
N. Virginia
EC2
VPC
S3
AWS Auto Scaling
Simple Queue Service
Simple Notification Service
EC2 > Auto Scaling groups > Create Auto Scaling group
Step 1
Choose launch template
Step 2
Choose instance launch options
Step 3 - optional
Configure advanced options
Step 4 - optional
Configure group size and scaling
Step 5 - optional
Add notifications
Step 6 - optional
Add tags
Step 7
Review
Configure advanced options - optional Info
Integrate your Auto Scaling group with other services to distribute network traffic across multiple servers using a load balancer or to establish service-to-service communications using VPC Lattice. You can also set options that give you more control over health check replacements and monitoring.
Load balancing Info
Use the options below to attach your Auto Scaling group to an existing load balancer, or to a new load balancer that you define.
No load balancer
Traffic to your Auto Scaling group will not be fronted by a load balancer.
Attach to an existing load balancer
Choose from your existing load balancers.
Attach to a new load balancer
Quickly create a basic load balancer to attach to your Auto Scaling group.
Attach to a new load balancer
Define a new load balancer to create for attachment to this Auto Scaling group.
Load balancer type
Choose from the load balancer types offered below. Type selection cannot be changed after the load balancer is created. If you need a different type of load balancer than those offered here, visit the Load Balancing console.
Application Load Balancer
HTTP, HTTPS
Network Load Balancer
TCP, UDP, TLS

Attach to a new load balancer

Define a new load balancer to create for attachment to this Auto Scaling group.

Load balancer type

Choose from the load balancer types offered below. Type selection cannot be changed after the load balancer is created. If you need a different type of load balancer than those offered here, visit the Load Balancing console.

- (●) Application Load Balancer — HTTP, HTTPS
- ( ) Network Load Balancer — TCP, UDP, TLS

Load balancer name

Name cannot be changed after the load balancer is created.

asg1D-1

Load balancer scheme

Scheme cannot be changed after the load balancer is created.

- ( ) Internal
- (●) Internet-facing

Network mapping

Your new load balancer will be created using the same VPC and Availability Zone selections as your Auto Scaling group. You can select different subnets and add subnets from additional Availability Zones.

VPC

vpc-04643f4dea97e29fa

Availability Zones and subnets

You must select a single subnet for each Availability Zone enabled. Only public subnets are available for selection to support DNS

Network mapping

Your new load balancer will be created using the same VPC and Availability Zone selections as your Auto Scaling group. You can select different subnets and add subnets from additional Availability Zones.

VPC

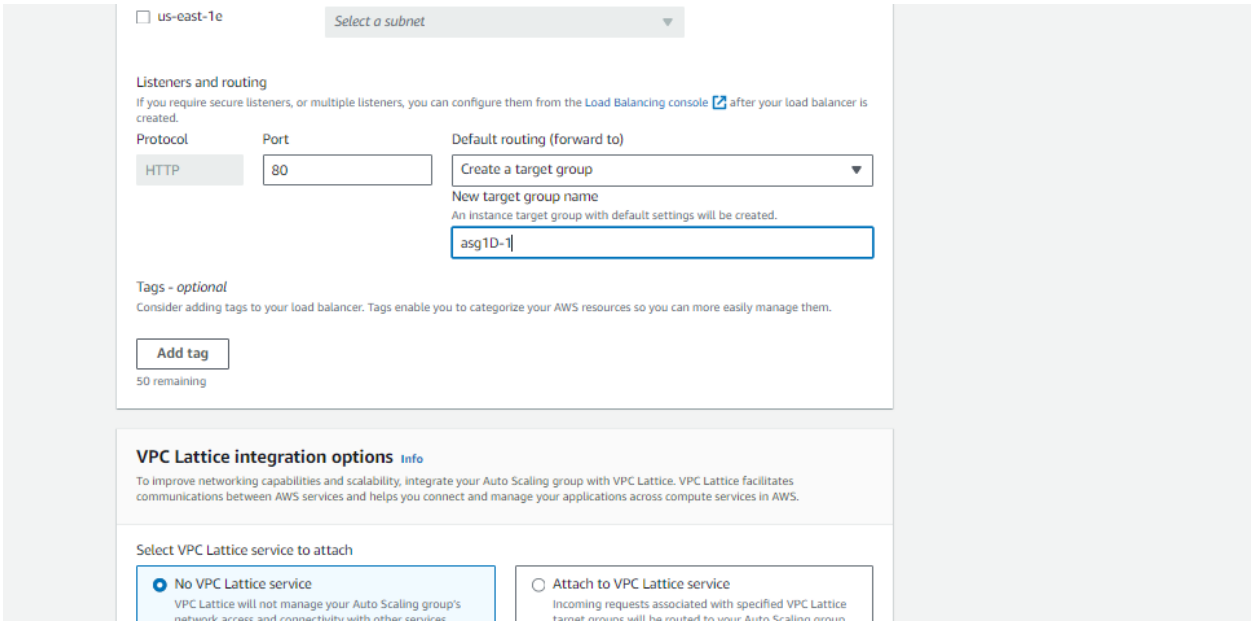
vpc-04643f4dea97e29fa

Availability Zones and subnets

You must select a single subnet for each Availability Zone enabled. Only public subnets are available for selection to support DNS resolution.

[x] us-east-1b	subnet-0886539ca2b844d31
[x] us-east-1a	subnet-0e3cc9d94044f8c92
[ ] us-east-1f	Select a subnet
[ ] us-east-1d	Select a subnet
[ ] us-east-1c	Select a subnet
[ ] us-east-1e	Select a subnet

Listeners and routing



communications between AWS services and helps you connect and manage your applications across compute services in AWS.

Select VPC Lattice service to attach

No VPC Lattice service
VPC Lattice will not manage your Auto Scaling group's network access and connectivity with other services.

Attach to VPC Lattice service
Incoming requests associated with specified VPC Lattice target groups will be routed to your Auto Scaling group.

Create new VPC Lattice service

Health checks

Health checks increase availability by replacing unhealthy instances. When you use multiple health checks, all are evaluated, and if at least one fails, instance replacement occurs.

EC2 health checks

Always enabled

Additional health check types - *optional* | Info

☐ Turn on Elastic Load Balancing health checks **Recommended**

Elastic Load Balancing monitors whether instances are available to handle requests. When it reports an unhealthy instance, EC2 Auto Scaling can replace it on its next periodic check.

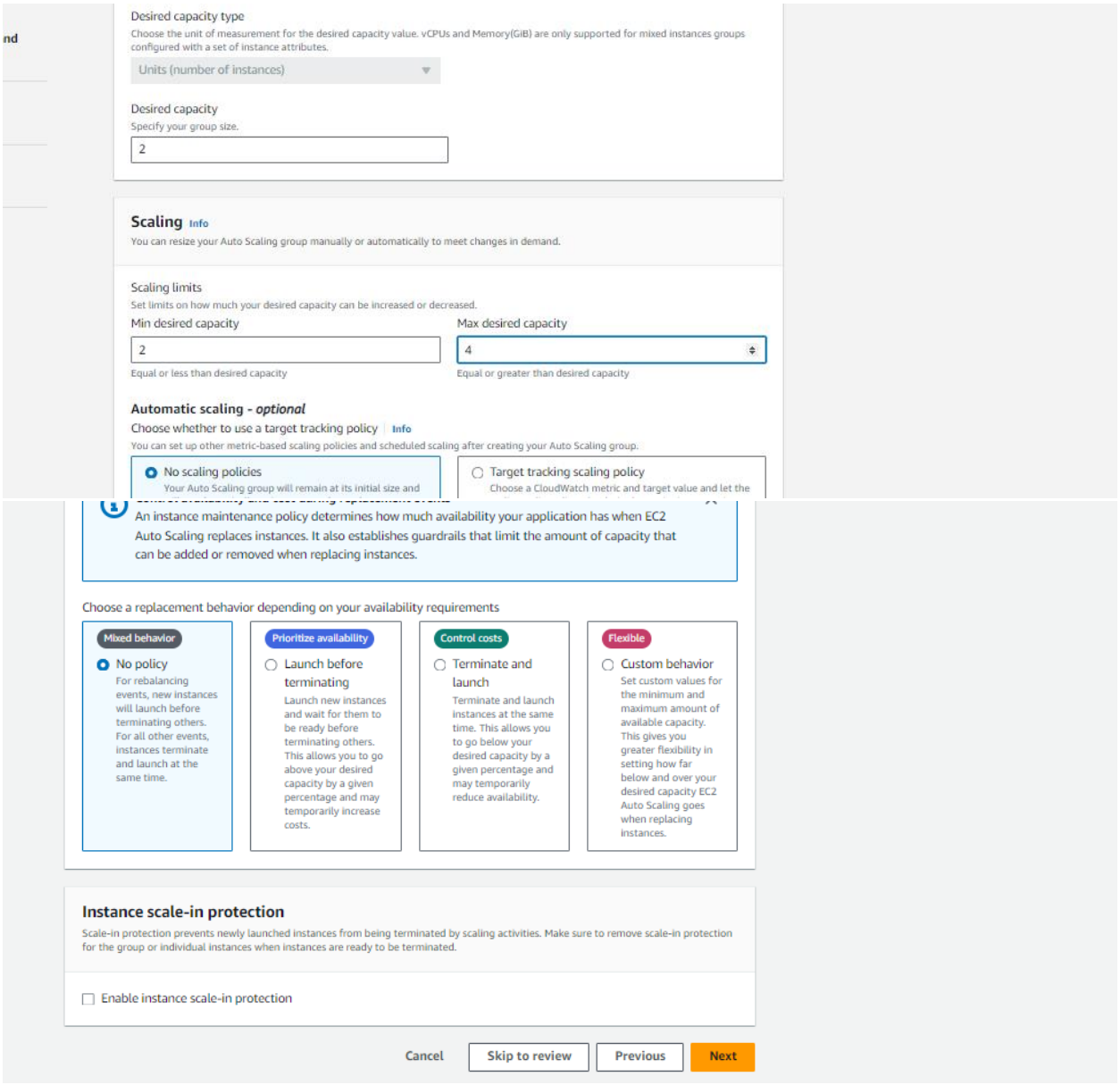
☐ Turn on VPC Lattice health checks

VPC Lattice can monitor whether instances are available to handle requests. If it considers a target as failed a health check, EC2 Auto Scaling replaces it after its next periodic check.

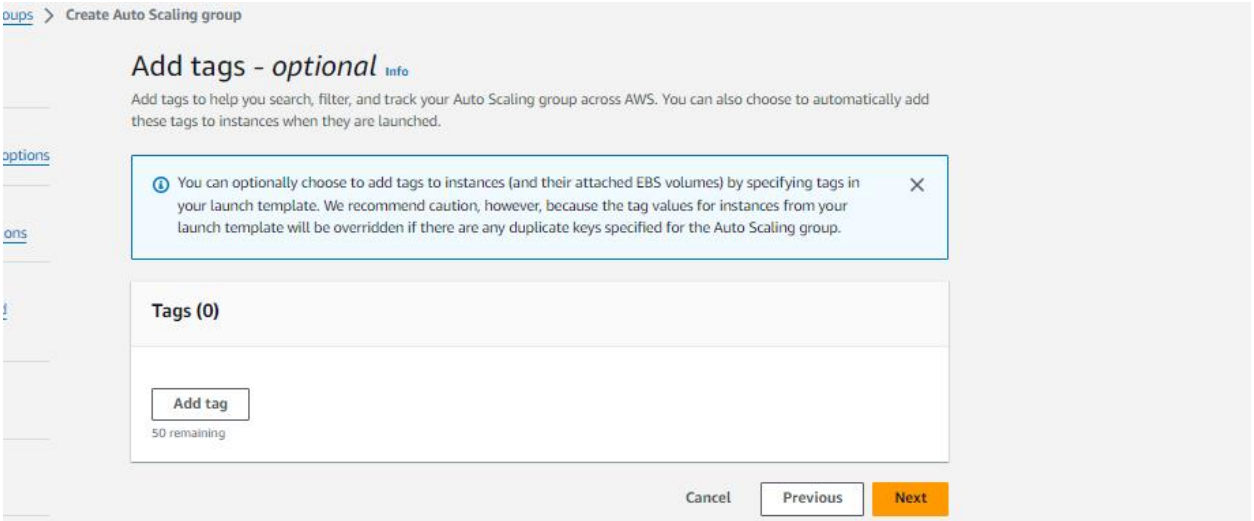
Health check grace period | Info

This time period delays the first health check until your instances finish initializing. It doesn't prevent an instance from terminating when placed into a non-running state.

300 seconds



Cancel | Skip to review | Previous | Next



No policy	-	-

Instance scale-in protection

Instance scale-in protection

Enable instance protection from scale in

Step 5: Add notifications — Edit

Notifications

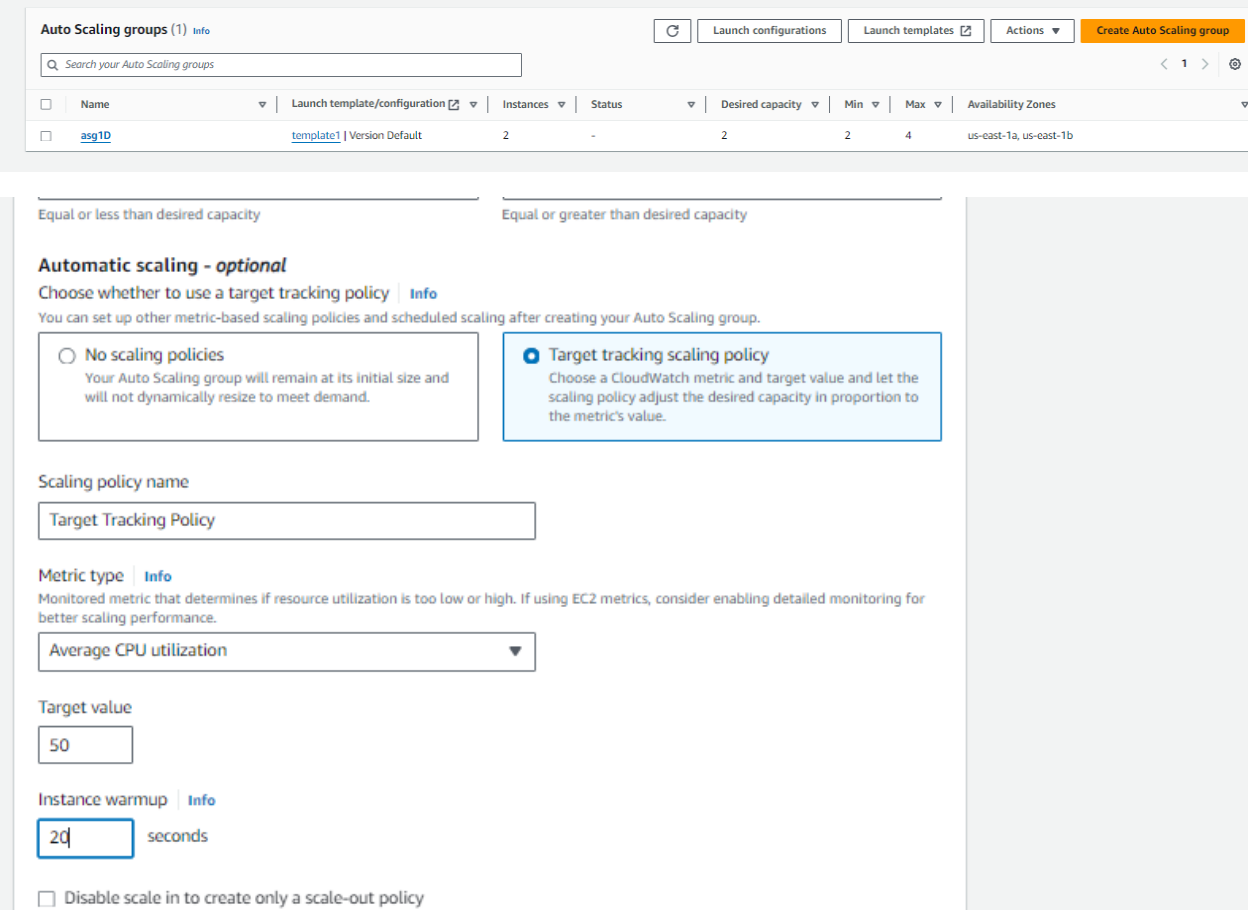
No notifications

Step 6: Add tags — Edit

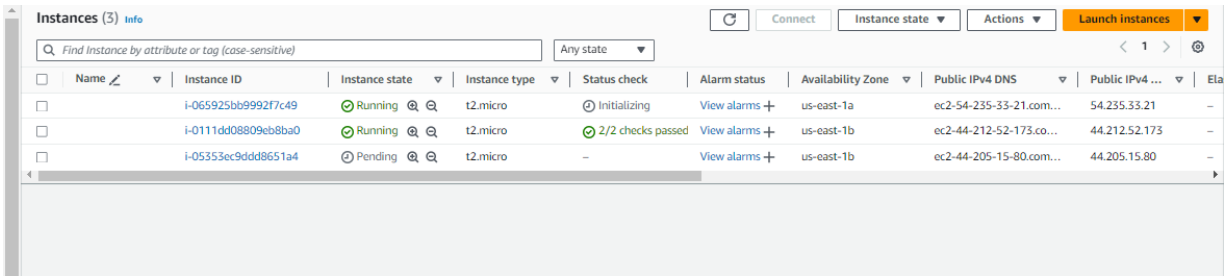
Tags (0)

Key	Value	Tag new instances
No tags		

Cancel | Previous | Create Auto Scaling group

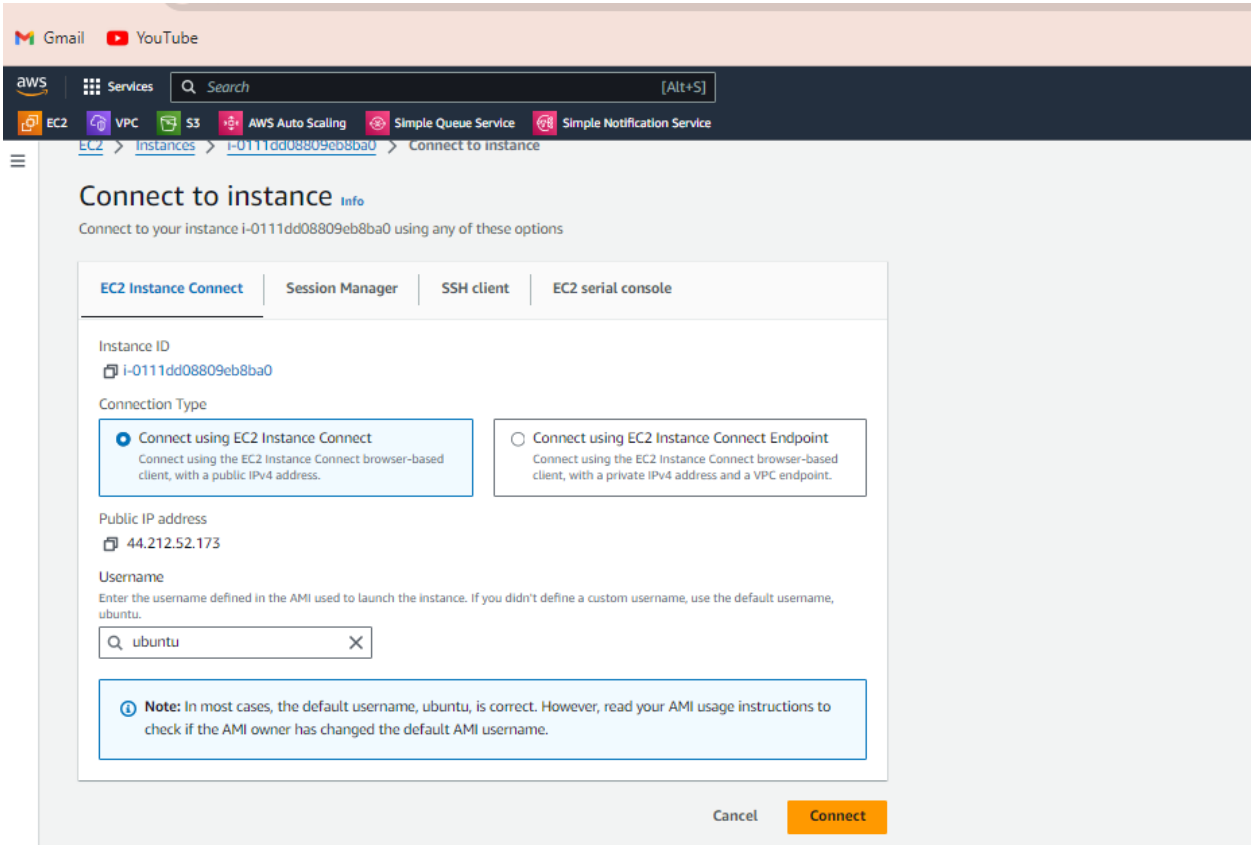


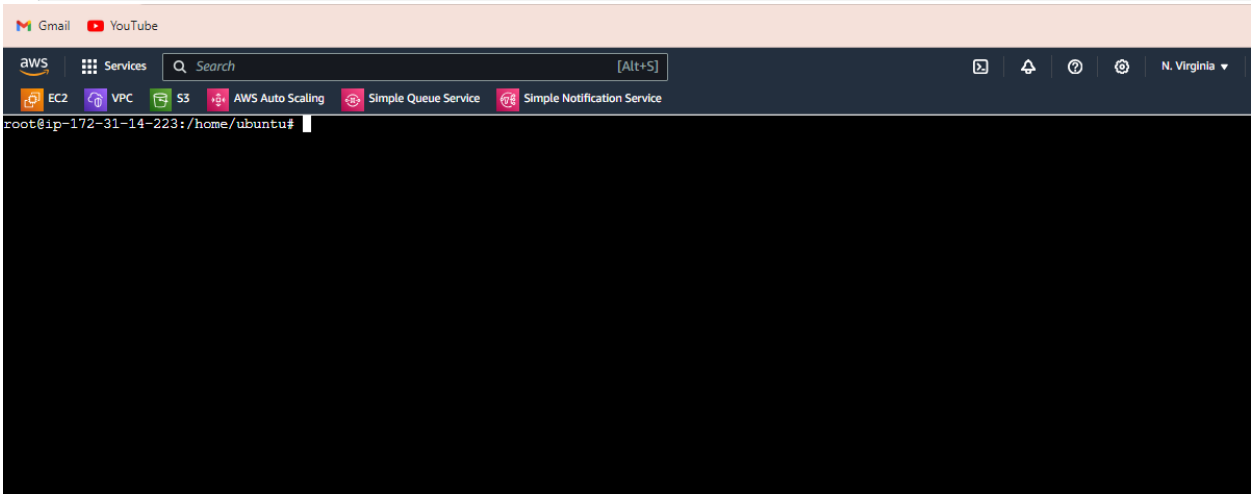
Step 4- now go to instance and check the no of min instance



Since we set min 2 instance so we got here now auto scaling is working now its time to set manually high traffic so our auto scale will work and instance will increase

Step 5- now connect to any instance





Run some command to increase traffic

Command – top

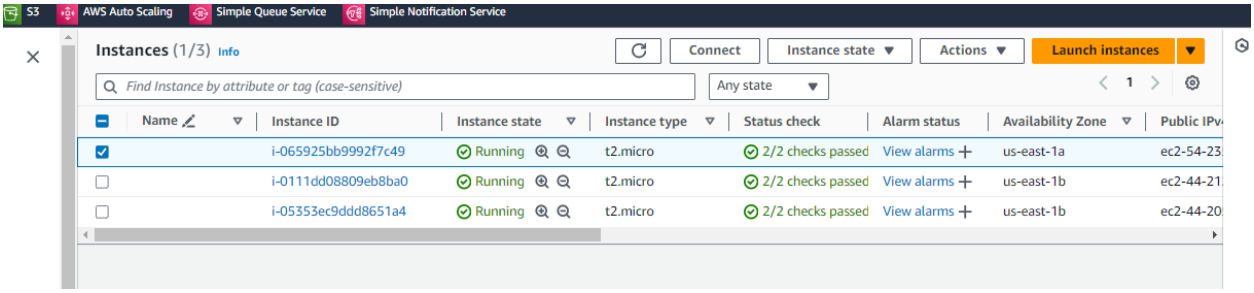
```
>/dev/null &
```

EC2 VPC S3 AWS Auto Scaling Simple Queue Service Simple Notification Service

```
%Cpu(s):  5.0 us,  5.0 sy,  0.0 ni, 90.0 id,  0.0 wa,  0.0 hi,  0.0 si,  0.0 st
MiB Mem :    949.7 total,    180.5 free,    183.1 used,    586.1 buff/cache
MiB Swap:      0.0 total,      0.0 free,      0.0 used.    580.4 avail Mem
top - 19:59:24 up 29 min,  2 users,  load average: 0.74, 0.24, 0.09
Tasks: 106 total,   2 running, 103 sleeping,   1 stopped,   0 zombie
%Cpu(s): 64.5 us, 35.5 sy,  0.0 ni,  0.0 id,  0.0 wa,  0.0 hi,  0.0 si,  0.0 st
MiB Mem :    949.7 total,    180.5 free,    183.0 used,    586.1 buff/cache
MiB Swap:      0.0 total,      0.0 free,      0.0 used.    580.4 avail Mem

    PID USER      PR  NI    VIRT    RES    SHR S  %CPU  %MEM     TIME+ COMMAND
  13872 root      20   0    6188   1792   1792 R  99.7   0.2   1:18.52 yes
  13861 root      20   0   11896   5504   4736 S   0.3   0.6   0:00.04 sudo
      1 root      20   0  167524  12000   7392 S   0.0   1.2   0:08.05 systemd
      2 root      20   0       0      0      0 S   0.0   0.0   0:00.00 kthreadd
      3 root       0 -20       0      0      0 I   0.0   0.0   0:00.00 rcu_gp
      4 root       0 -20       0      0      0 I   0.0   0.0   0:00.00 rcu_par_gp
      5 root       0 -20       0      0      0 I   0.0   0.0   0:00.00 slub_flushwq
      6 root       0 -20       0      0      0 I   0.0   0.0   0:00.00 netns
      8 root       0 -20       0      0      0 I   0.0   0.0   0:00.00 kworker/0:0H-events_highpri
     10 root       0 -20       0      0      0 I   0.0   0.0   0:00.00 mm_percpu_wq
     11 root      20   0       0      0      0 I   0.0   0.0   0:00.00 rcu_tasks_rude_kthread
     12 root      20   0       0      0      0 I   0.0   0.0   0:00.00 rcu_tasks_trace_kthread
     13 root      20   0       0      0      0 S   0.0   0.0   0:00.33 ksoftirqd/0
```

>/dev/null &



NOW WAIT FOR FEW MINT TO NOTICE YOUR INSTANCE IF IT IS INCREASE OR NOT