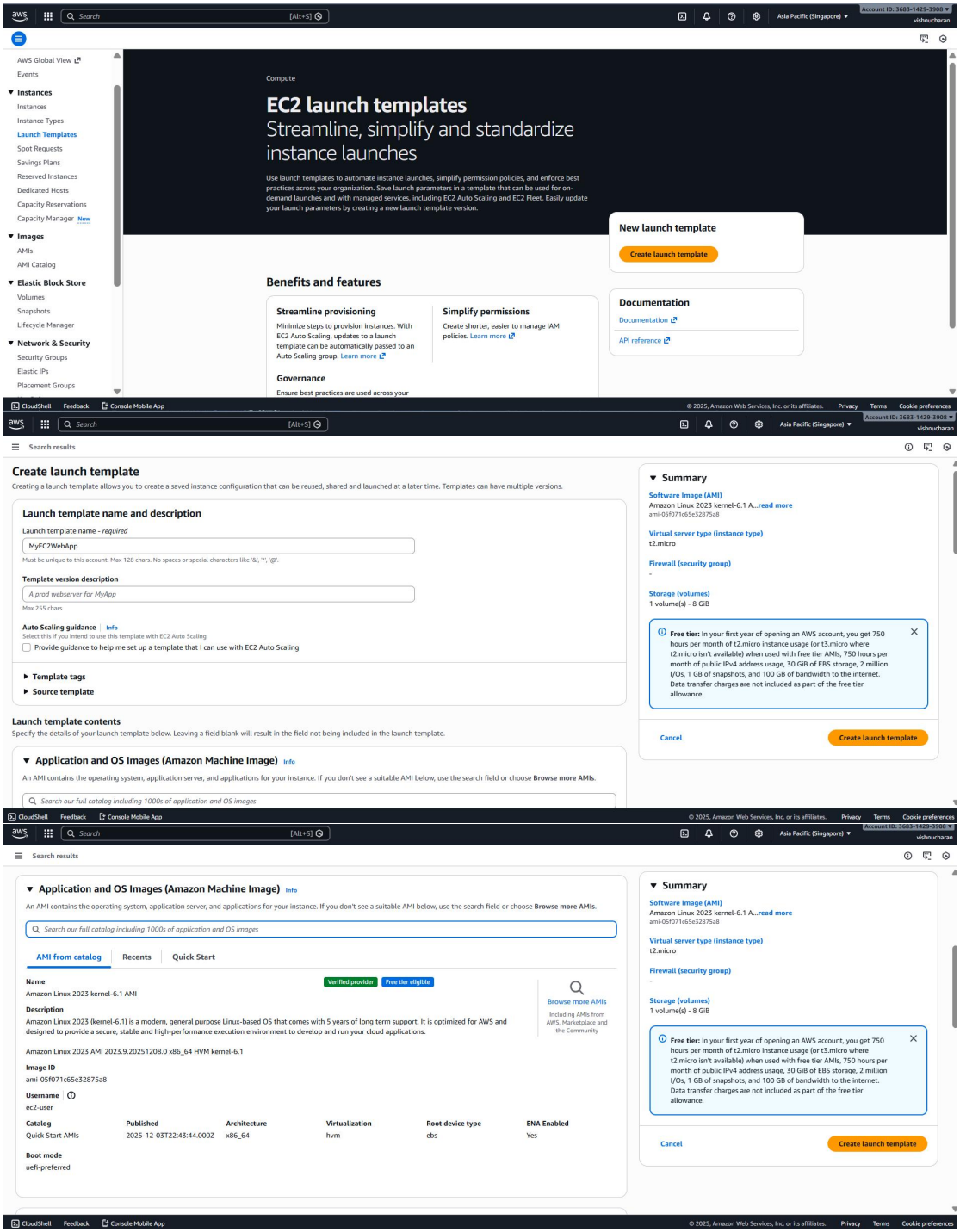
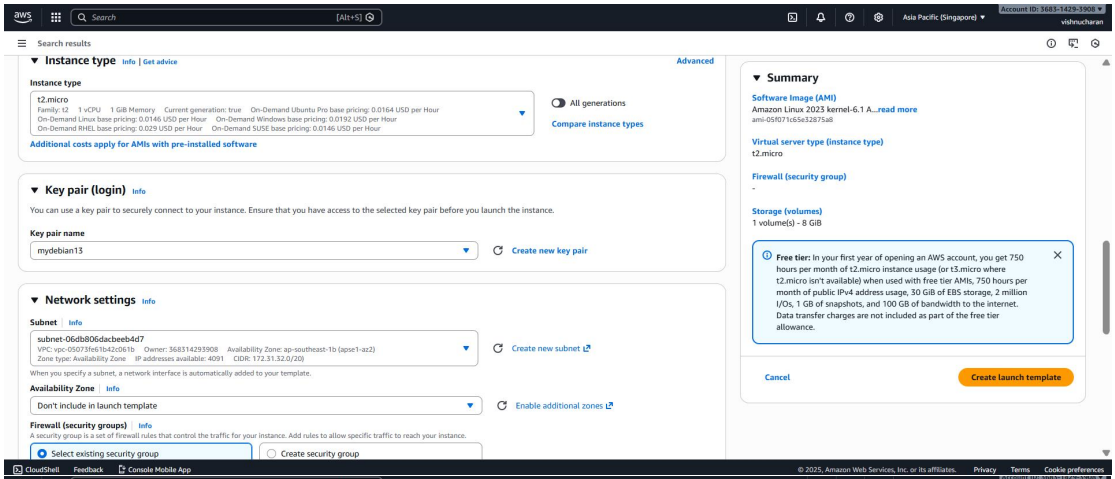
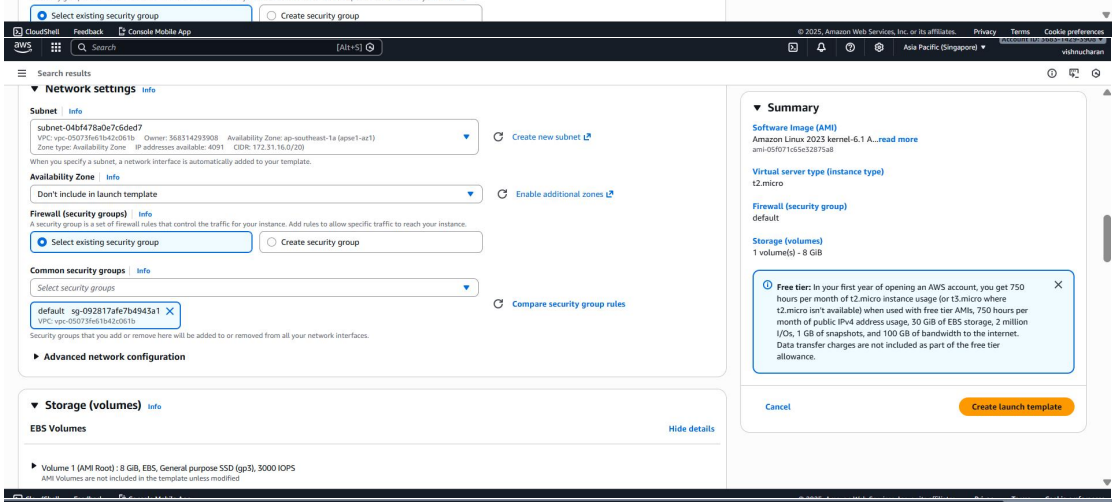
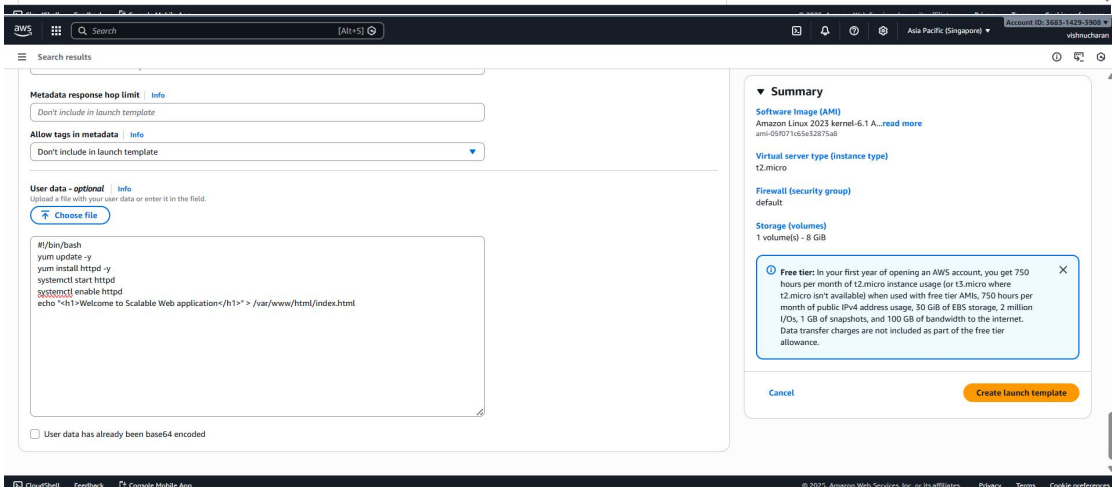
Scalable Web App with ALB & Auto Scaling – Use EC2, ALB, and Auto Scaling for high availability.

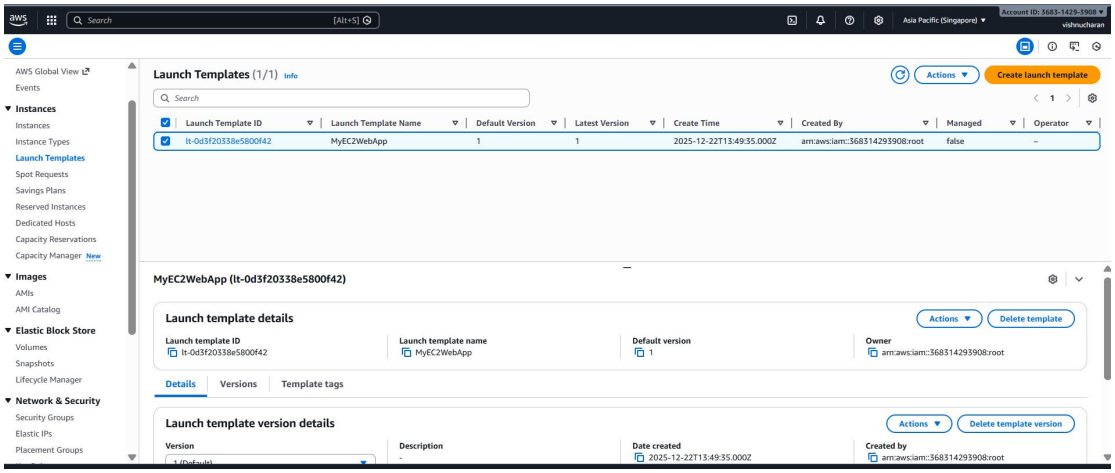
1. Create a Launch Template (Blueprint for EC2):-
2. AWS Console → EC2 → Launch Templates → Create launch template



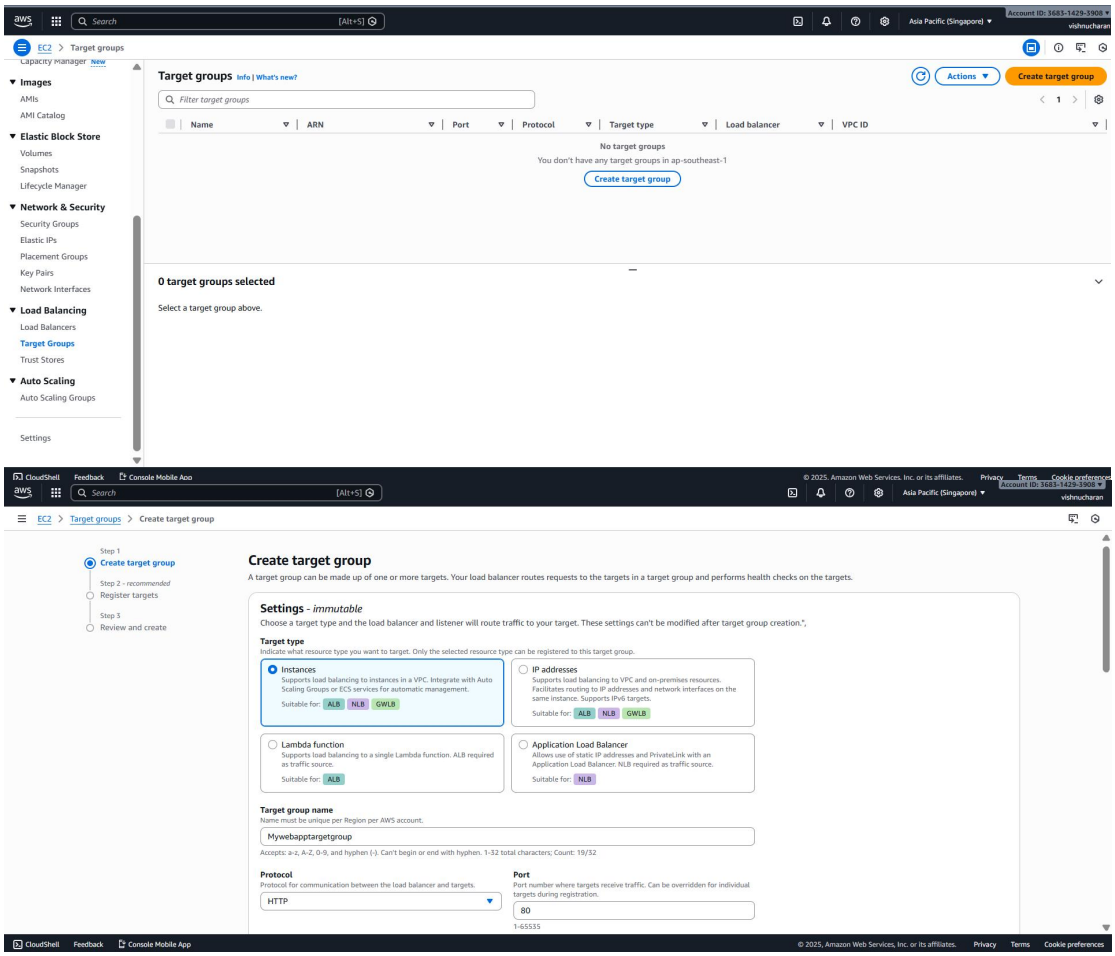

Search results
Instance type Info | Get advice
Advanced
Instance type
t2.micro
Family: t2 1 vCPU 1 GiB Memory Current generation: true On-Demand Ubuntu Pro base pricing: 0.0164 USD per Hour
On-Demand Linux base pricing: 0.0146 USD per Hour On-Demand Windows base pricing: 0.0192 USD per Hour
On-Demand RHEL base pricing: 0.029 USD per Hour On-Demand SUSE base pricing: 0.0146 USD per Hour
All generations
Compare instance types
Additional costs apply for AMIs with pre-installed software
Key pair (login) Info
You can use a key pair to securely connect to your instance. Ensure that you have access to the selected key pair before you launch the instance.
Key pair name
mydebian13
Create new key pair
Network settings Info
Subnet Info
subnet-06db806dacbeeb4d7
VPC: vpc-05073fe61b42c061b Owner: 368314293908 Availability Zone: ap-southeast-1b (apse1-az2)
Zone type: Availability Zone IP addresses available: 4091 CIDR: 172.31.32.0/20)
Create new subnet
When you specify a subnet, a network interface is automatically added to your template.
Availability Zone Info
Don't include in launch template
Enable additional zones
Firewall (security groups) Info
A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.
Select existing security group
Create security group
Summary
Software Image (AMI)
Amazon Linux 2023 kernel-6.1 A...read more
ami-05f071c65e52875a6
Virtual server type (instance type)
t2.micro
Firewall (security group)
-
Storage (volumes)
1 volume(s) - 8 GiB
Free tier: In your first year of opening an AWS account, you get 750 hours per month of t2.micro instance usage (or t3.micro where t2.micro isn't available) when used with free tier AMIs, 750 hours per month of public IPv4 address usage, 30 GiB of EBS storage, 2 million I/Os, 1 GB of snapshots, and 100 GB of bandwidth to the internet. Data transfer charges are not included as part of the free tier allowance.
Cancel
Create launch template
CloudShell Feedback Console Mobile App
© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

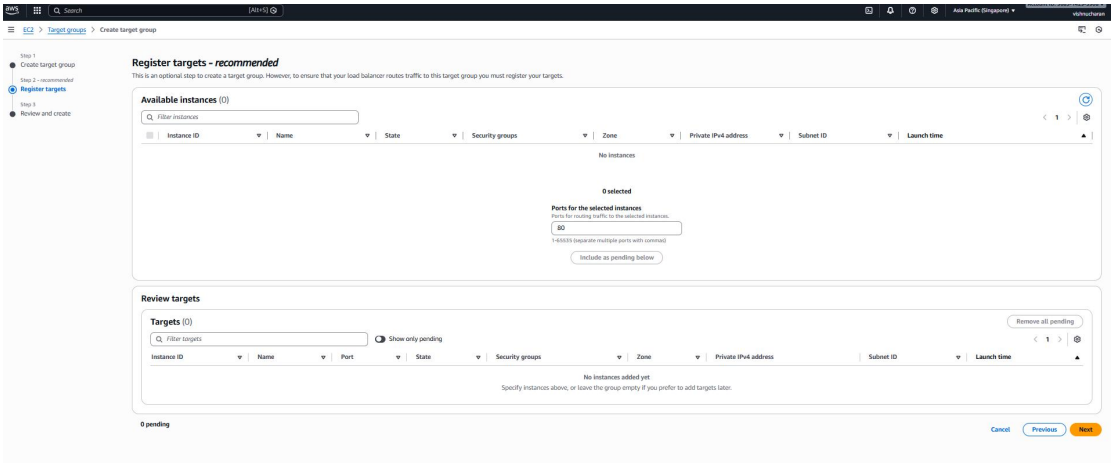

Search results
Network settings Info
Subnet Info
subnet-04bf478a0e7c6ded7
VPC: vpc-05073fe61b42c061b Owner: 368314293908 Availability Zone: ap-southeast-1a (apse1-az1)
Zone type: Availability Zone IP addresses available: 4091 CIDR: 172.31.16.0/20)
Create new subnet
When you specify a subnet, a network interface is automatically added to your template.
Availability Zone Info
Don't include in launch template
Enable additional zones
Firewall (security groups) Info
A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.
Select existing security group
Create security group
Common security groups Info
Select security groups
default sg-092817afe7b4943a1
VPC: vpc-05073fe61b42c061b
Compare security group rules
Security groups that you add or remove here will be added to or removed from all your network interfaces.
Advanced network configuration
Storage (volumes) Info
EBS Volumes
Hide details
Volume 1 (AMI Root) : 8 GiB, EBS, General purpose SSD (gp3), 3000 IOPS
AMI Volumes are not included in the template unless modified
Summary
Software Image (AMI)
Amazon Linux 2023 kernel-6.1 A...read more
ami-05f071c65e52875a8
Virtual server type (instance type)
t2.micro
Firewall (security group)
default
Storage (volumes)
1 volume(s) - 8 GiB
Free tier: In your first year of opening an AWS account, you get 750 hours per month of t2.micro instance usage (or t3.micro where t2.micro isn't available) when used with free tier AMIs, 750 hours per month of public IPv4 address usage, 30 GiB of EBS storage, 2 million I/Os, 1 GB of snapshots, and 100 GB of bandwidth to the internet. Data transfer charges are not included as part of the free tier allowance.
Cancel
Create launch template

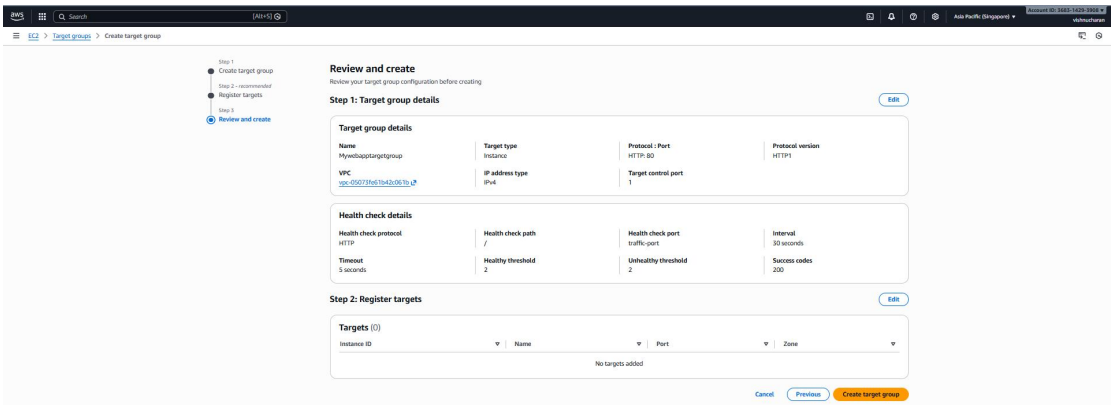

Search results
Metadata response hop limit Info
Don't include in launch template
Allow tags in metadata Info
Don't include in launch template
User data - optional Info
Upload a file with your user data or enter it in the field.
Choose file
#!/bin/bash
yum update -y
yum install httpd -y
systemctl start httpd
systemctl enable httpd
echo "<h1>Welcome to Scalable Web application</h1>" > /var/www/html/index.html
User data has already been base64 encoded
Summary
Software Image (AMI)
Amazon Linux 2023 kernel-6.1 A...read more
ami-05f071c65e52875a8
Virtual server type (instance type)
t2.micro
Firewall (security group)
default
Storage (volumes)
1 volume(s) - 8 GiB
Free tier: In your first year of opening an AWS account, you get 750 hours per month of t2.micro instance usage (or t3.micro where t2.micro isn't available) when used with free tier AMIs, 750 hours per month of public IPv4 address usage, 30 GiB of EBS storage, 2 million I/Os, 1 GB of snapshots, and 100 GB of bandwidth to the internet. Data transfer charges are not included as part of the free tier allowance.
Cancel
Create launch template

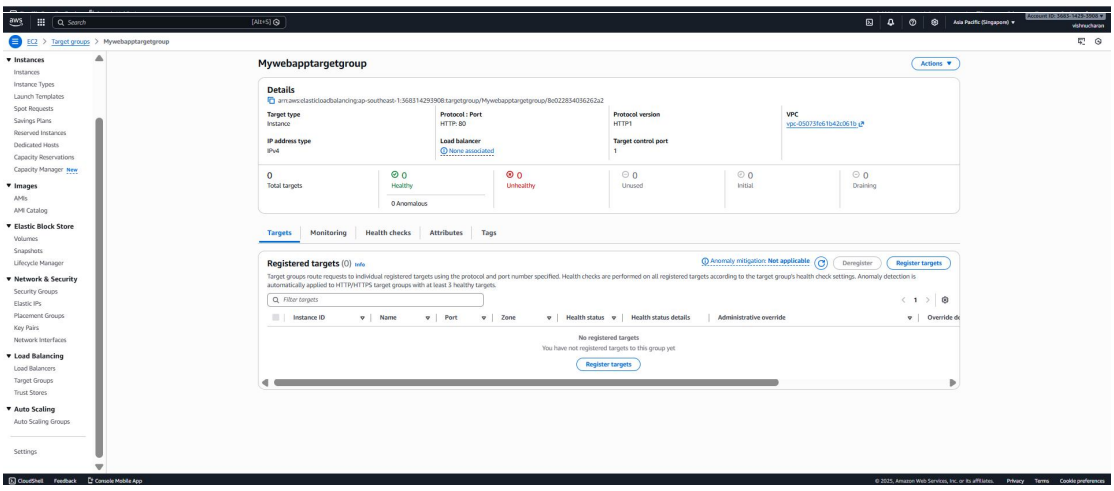


3. Create Target Group (Health & Routing)

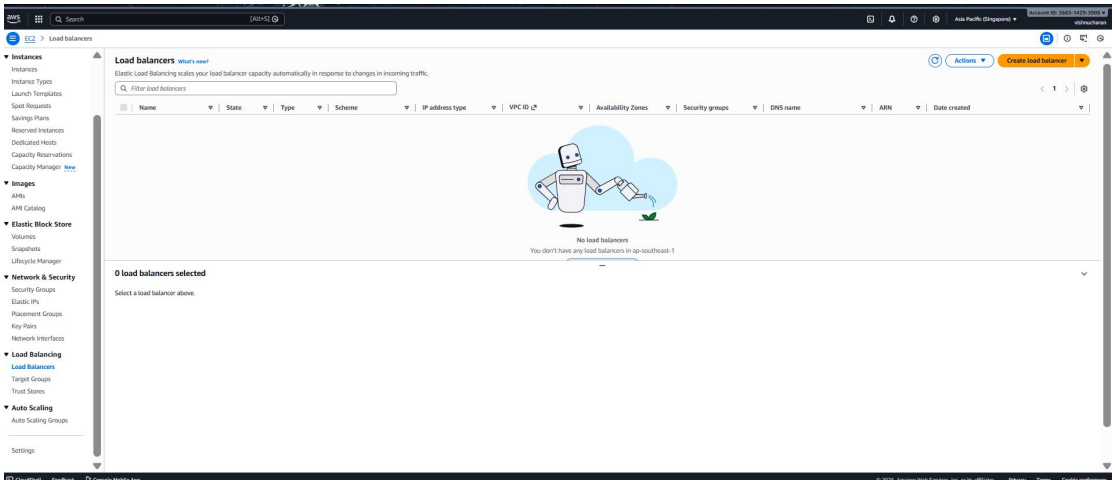


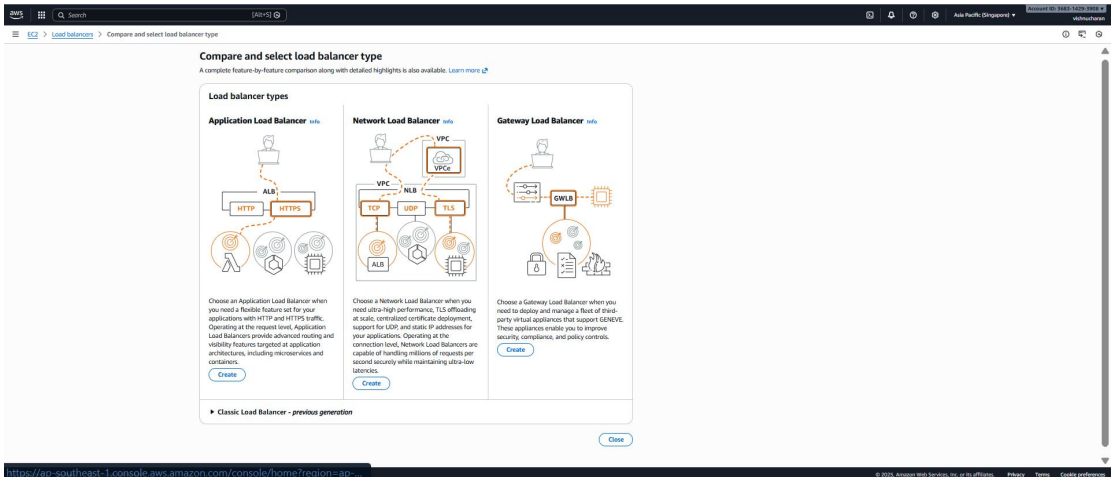

Register targets - recommended
This is an optional step to create a target group. However, to ensure that your load balancer routes traffic to this target group you must register your targets.
Available instances (0)
No instances
0 selected
Ports for the selected instances
80
Include as pending below
Review targets
Targets (0)
No instances added yet
Specify instances above, or leave the group empty if you prefer to add targets later.
0 pending
Cancel
Previous
Next

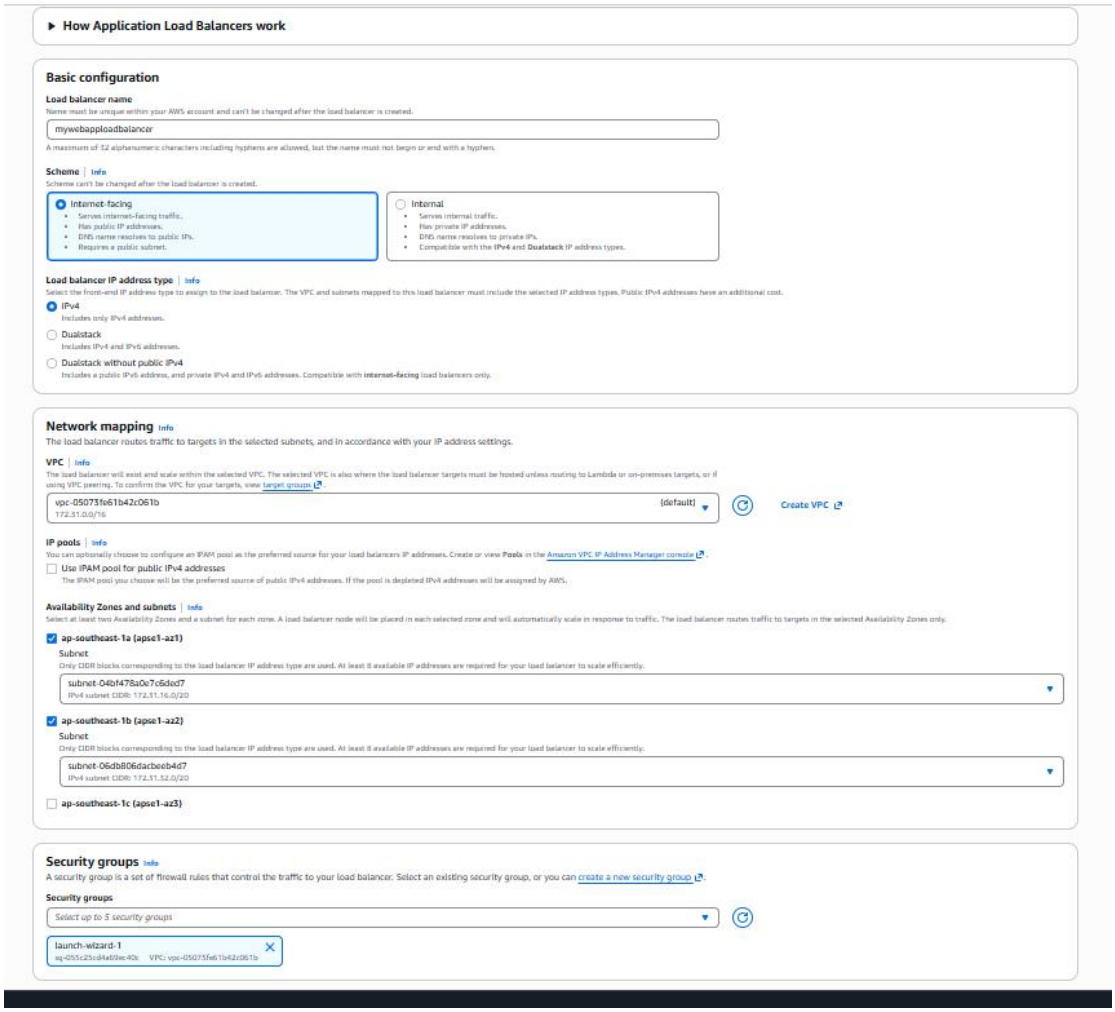

Review and create
Review your target group configuration before creating
Step 1: Target group details
Edit
Target group details
Name
Mywebapptargetgroup
Target type
Instance
Protocol : Port
HTTP: 80
Protocol version
HTTP1
VPC
vpc-05073fe61b42c061b
IP address type
IPv4
Target control port
1
Health check details
Health check protocol
HTTP
Health check path
/
Health check port
traffic-port
Interval
30 seconds
Timeout
5 seconds
Healthy threshold
2
Unhealthy threshold
2
Success codes
200
Step 2: Register targets
Edit
Targets (0)
No targets added
Cancel
Previous
Create target group

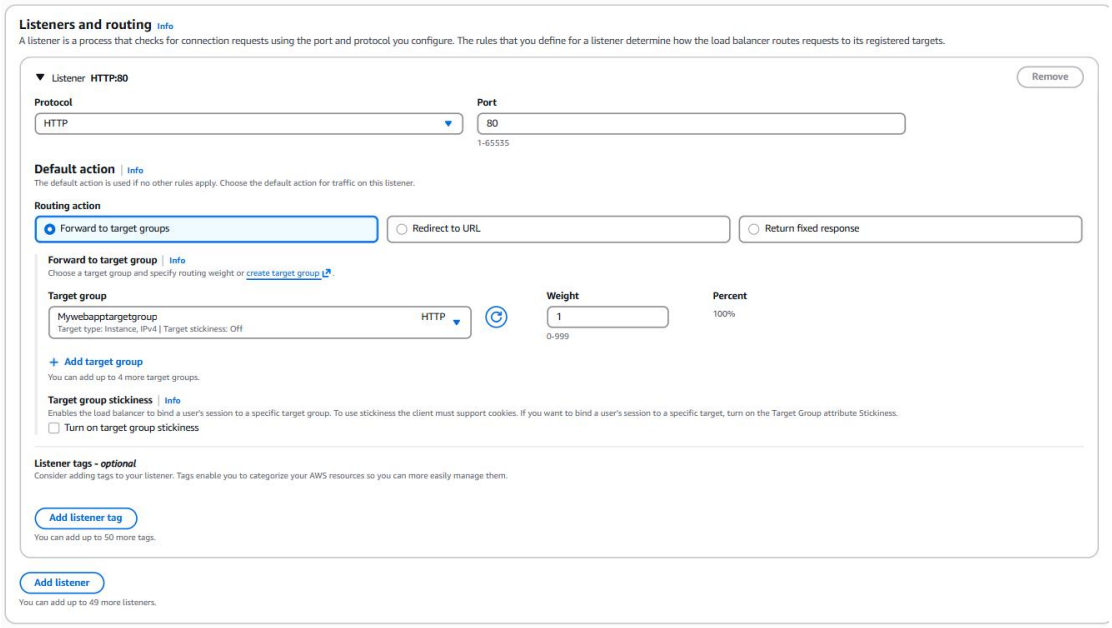

Mywebapptargetgroup
Actions
Details
arn:aws:elasticloadbalancing:ap-southeast-1:368314293908:targetgroup/Mywebapptargetgroup/8e022834036262a2
Target type
Instance
Protocol : Port
HTTP: 80
Protocol version
HTTP1
VPC
vpc-05073fe61b42c061b
IP address type
IPv4
Load balancer
None associated
Target control port
1
0
Total targets
0
Healthy
0 Anomalous
0
Unhealthy
0
Unused
0
Initial
0
Draining
Targets
Monitoring
Health checks
Attributes
Tags
Registered targets (0)
Anomaly mitigation: Not applicable
Deregister
Register targets
No registered targets
You have not registered targets to this group yet
Register targets

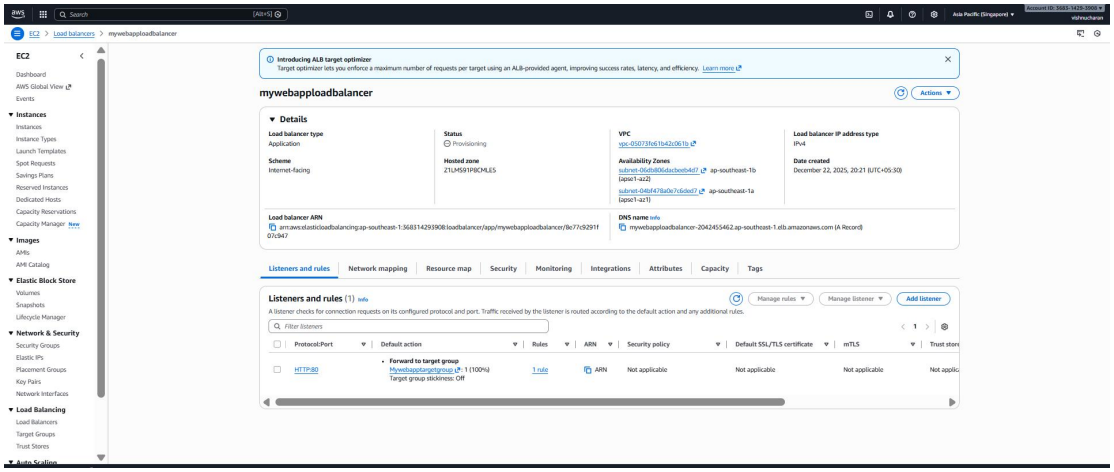
4. Create Application Load Balancer (ALB)



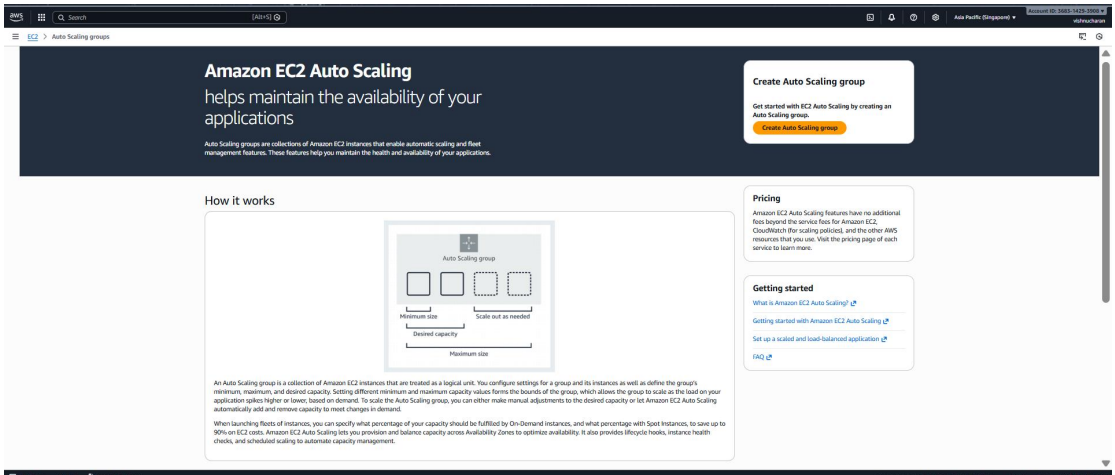


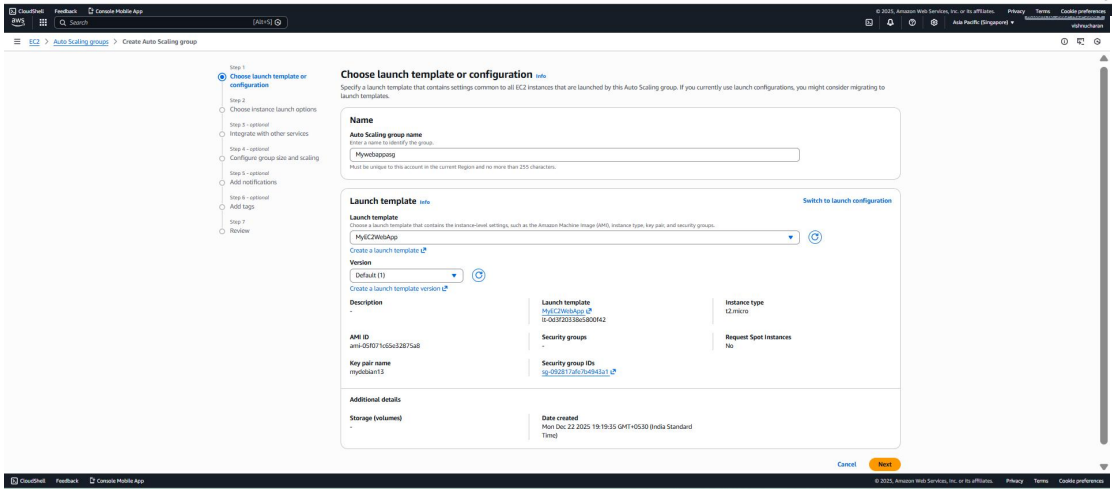

How Application Load Balancers work
Basic configuration
Load balancer name
mywebapploadbalancer
Scheme | Info
Internet-facing
Internal
Load balancer IP address type | Info
IPv4
Dualstack
Dualstack without public IPv4
Network mapping Info
The load balancer routes traffic to targets in the selected subnets, and in accordance with your IP address settings.
VPC | Info
vpc-05073fe61b42c061b
172.31.0.0/16
[default]
Create VPC
IP pools | Info
Use IPAM pool for public IPv4 addresses
Availability Zones and subnets | Info
ap-southeast-1a (apse1-az1)
Subnet
subnet-04bf478a0e7c6ded7
IPv4 subnet CIDR: 172.31.16.0/20
ap-southeast-1b (apse1-az2)
Subnet
subnet-06db806dacbeeb4d7
IPv4 subnet CIDR: 172.31.32.0/20
ap-southeast-1c (apse1-az3)
Security groups Info
A security group is a set of firewall rules that control the traffic to your load balancer. Select an existing security group, or you can create a new security group.
Security groups
Select up to 5 security groups
launch-wizard-1
sg-055c23cd4a69ec40c VPC: vpc-05073fe61b42c061b

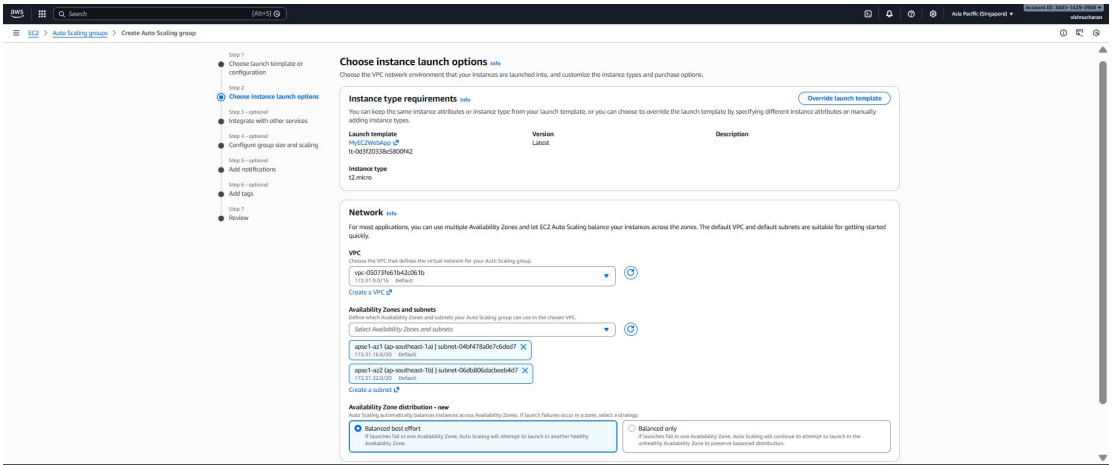

Listeners and routing Info
A listener is a process that checks for connection requests using the port and protocol you configure. The rules that you define for a listener determine how the load balancer routes requests to its registered targets.
Listener HTTP:80
Remove
Protocol
HTTP
Port
80
1-65535
Default action | Info
The default action is used if no other rules apply. Choose the default action for traffic on this listener.
Routing action
Forward to target groups
Redirect to URL
Return fixed response
Forward to target group | Info
Choose a target group and specify routing weight or create target group.
Target group
Mywebapptargetgroup
HTTP
Target type: Instance, IPv4 | Target stickiness: Off
Weight
1
0-999
Percent
100%
Add target group
You can add up to 4 more target groups.
Target group stickiness | Info
Enables the load balancer to bind a user's session to a specific target group. To use stickiness the client must support cookies. If you want to bind a user's session to a specific target, turn on the Target Group attribute Stickiness.
Turn on target group stickiness
Listener tags - optional
Consider adding tags to your listener. Tags enable you to categorize your AWS resources so you can more easily manage them.
Add listener tag
You can add up to 50 more tags.
Add listener
You can add up to 49 more listeners.

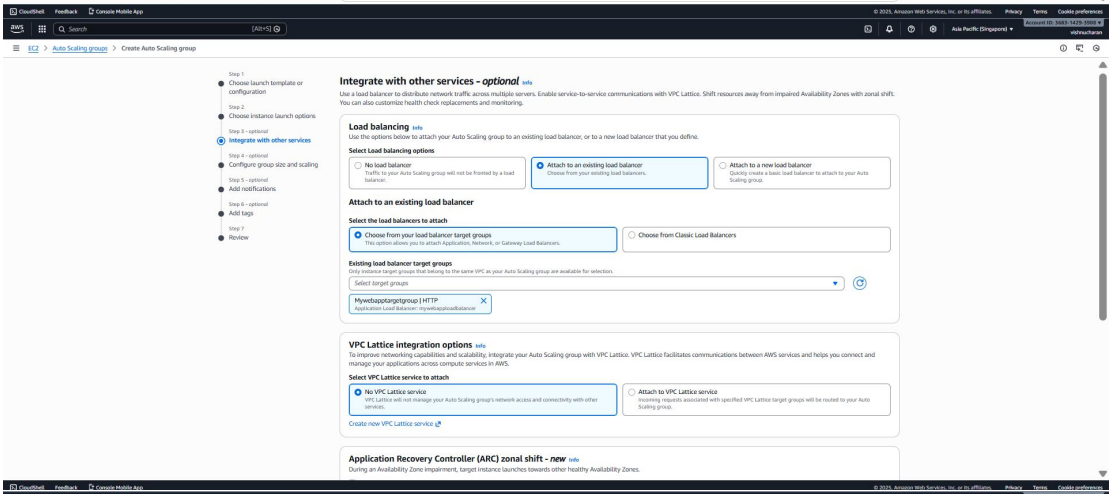


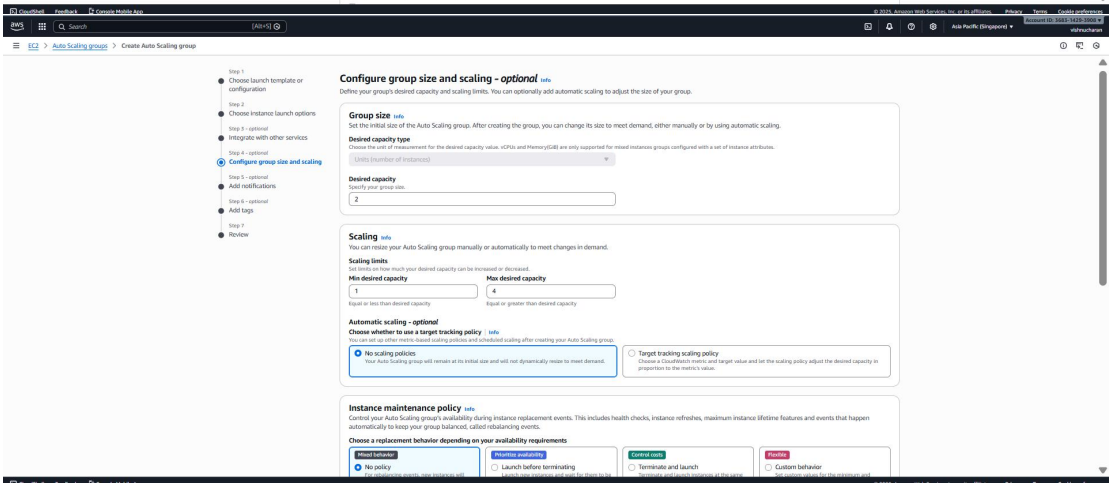
5. Create Auto Scaling Group

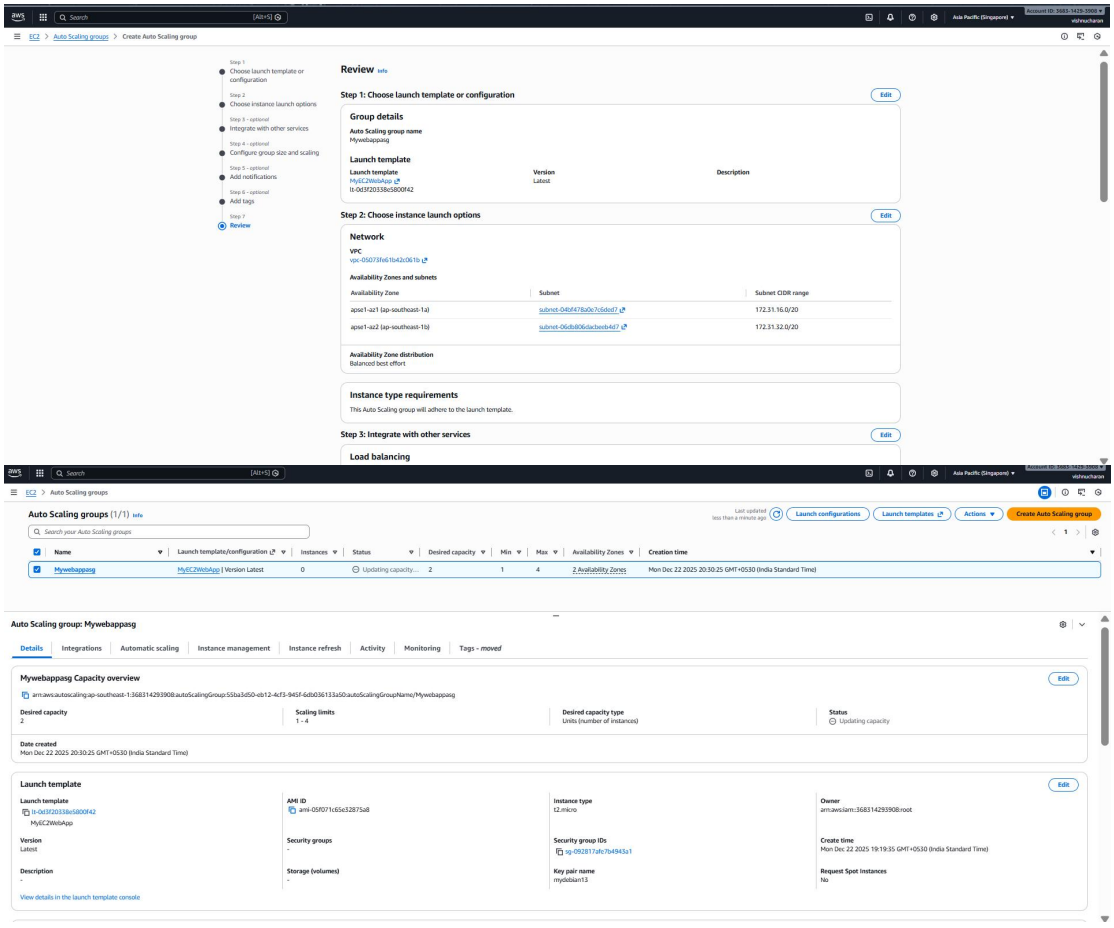




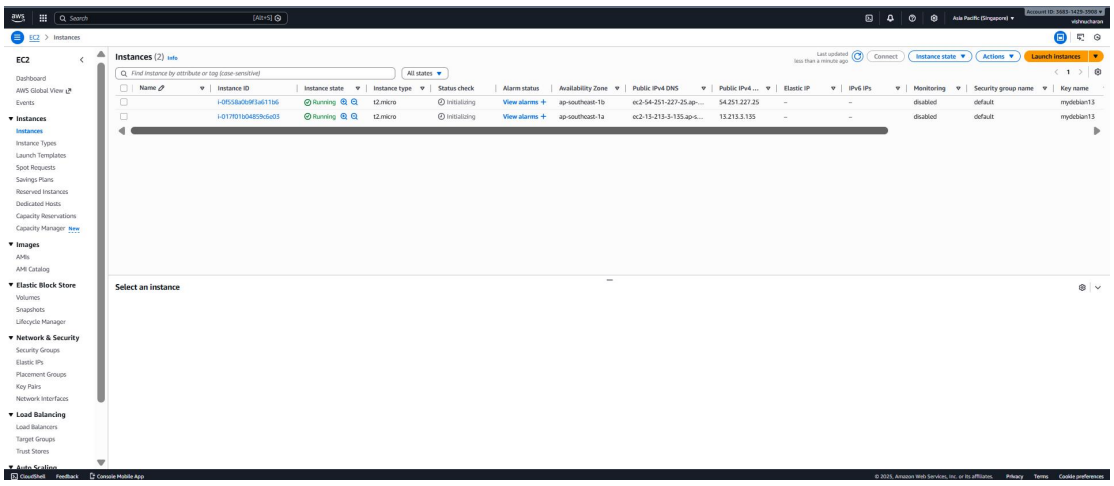

Search
[Alt+S]
Asia Pacific (Singapore)
vishnucharan
EC2 > Auto Scaling groups > Create Auto Scaling group
Step 1 Choose launch template or configuration
Step 2 Choose instance launch options
Step 3 - optional Integrate with other services
Step 4 - optional Configure group size and scaling
Step 5 - optional Add notifications
Step 6 - optional Add tags
Step 7 Review
Choose instance launch options Info
Choose the VPC network environment that your instances are launched into, and customize the instance types and purchase options.
Instance type requirements Info
Override launch template
You can keep the same instance attributes or instance type from your launch template, or you can choose to override the launch template by specifying different instance attributes or manually adding instance types.
Launch template
MyEC2NitiApp
lt-0d3f20338e5800f42
Version
Latest
Description
Instance type
t2.micro
Network Info
For most applications, you can use multiple Availability Zones and let EC2 Auto Scaling balance your instances across the zones. The default VPC and default subnets are suitable for getting started quickly.
VPC
Choose the VPC that defines the virtual network for your Auto Scaling group.
vpc-050736d1b42c061b
172.31.0.0/16 Default
Create a VPC
Availability Zones and subnets
Define which Availability Zones and subnets your Auto Scaling group can use in the chosen VPC.
Select Availability Zones and subnets
apse1-az1 (ap-southeast-1a) | subnet-04bf478a0e7c6ded7
172.31.16.0/20 Default
apse1-az2 (ap-southeast-1b) | subnet-06dbd06dacbeeb4d7
172.31.32.0/20 Default
Create a subnet
Availability Zone distribution - new
Auto Scaling automatically balances instances across Availability Zones. If launch failures occur in a zone, select a strategy.
Balanced best effort
If launches fail in one Availability Zone, Auto Scaling will attempt to launch in another healthy Availability Zone.
Balanced only
If launches fail in one Availability Zone, Auto Scaling will continue to attempt to launch in the unhealthy Availability Zone to preserve balanced distribution.


CloudShell
Feedback
Console Mobile App
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
[Alt+S]
Asia Pacific (Singapore)
vishnucharan
EC2 > Auto Scaling groups > Create Auto Scaling group
Step 1 Choose launch template or configuration
Step 2 Choose instance launch options
Step 3 - optional Integrate with other services
Step 4 - optional Configure group size and scaling
Step 5 - optional Add notifications
Step 6 - optional Add tags
Step 7 Review
Integrate with other services - optional Info
Use a load balancer to distribute network traffic across multiple servers. Enable service-to-service communications with VPC Lattice. Shift resources away from impaired Availability Zones with zonal shift. You can also customize 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.
Select Load balancing options
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 an existing load balancer
Select the load balancers to attach
Choose from your load balancer target groups
This option allows you to attach Application, Network, or Gateway Load Balancers.
Choose from Classic Load Balancers
Existing load balancer target groups
Only instance target groups that belong to the same VPC as your Auto Scaling group are available for selection.
Select target groups
Mywebapptargetgroup | HTTP
Application Load Balancer: mywebapploadbalancer
VPC Lattice integration options Info
To improve networking capabilities and scalability, integrate your Auto Scaling group with VPC Lattice. VPC Lattice facilitates 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
Application Recovery Controller (ARC) zonal shift - new Info
During an Availability Zone impairment, target instance launches towards other healthy Availability Zones.


CloudShell
Feedback
Console Mobile App
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
[Alt+S]
Asia Pacific (Singapore)
vishnucharan
EC2 > Auto Scaling groups > Create Auto Scaling group
Step 1 Choose launch template or configuration
Step 2 Choose instance launch options
Step 3 - optional Integrate with other services
Step 4 - optional Configure group size and scaling
Step 5 - optional Add notifications
Step 6 - optional Add tags
Step 7 Review
Configure group size and scaling - optional Info
Define your group's desired capacity and scaling limits. You can optionally add automatic scaling to adjust the size of your group.
Group size Info
Set the initial size of the Auto Scaling group. After creating the group, you can change its size to meet demand, either manually or by using automatic scaling.
Desired capacity type
Choose the unit of measurement for the desired capacity value. vCPUs and Memory(GiB) are only supported for mixed instances groups configured with a set of instance attributes.
Units (number of instances)
Desired capacity
Specify your group size.
2
Scaling Info
You can resize your Auto Scaling group manually or automatically to meet changes in demand.
Scaling limits
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
4
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.
Instance maintenance policy Info
Control your Auto Scaling group's availability during instance replacement events. This includes health checks, instance refreshes, maximum instance lifetime features and events that happen automatically to keep your group balanced, called rebalancing events.
Choose a replacement behavior depending on your availability requirements
Mixed behavior
No policy
Prioritize availability
Launch before terminating
Control costs
Terminate and launch
Flexible
Custom behavior



Now my EC2 instances are created automatically because of Auto Scaling Group.



Welcome to Scalable Web App