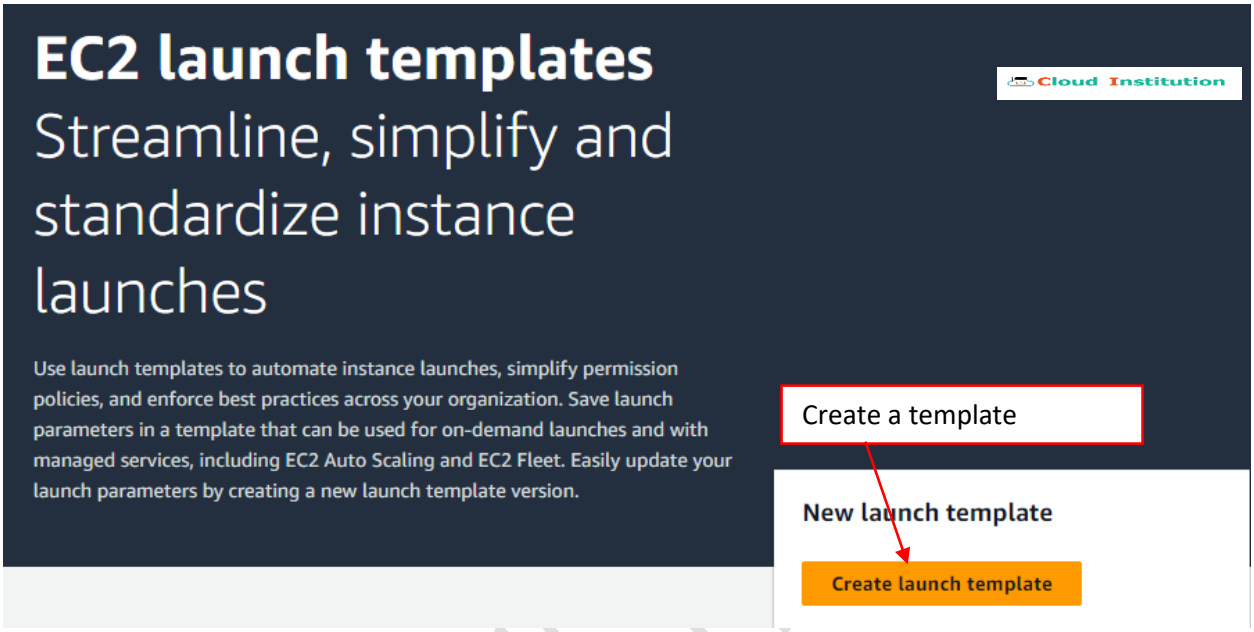
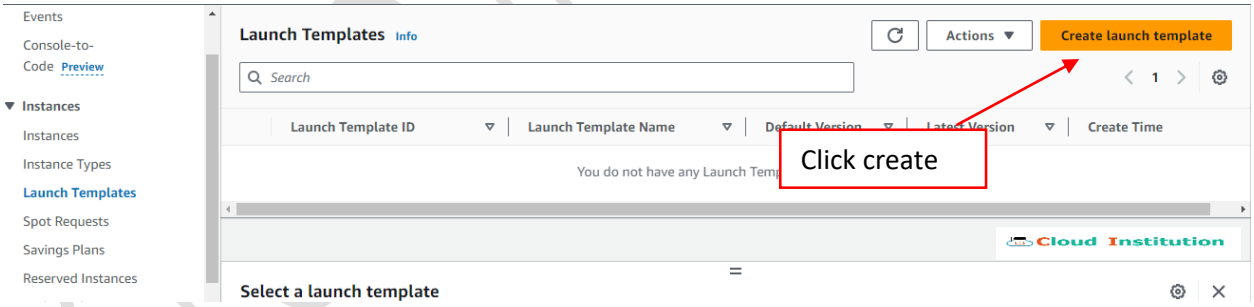
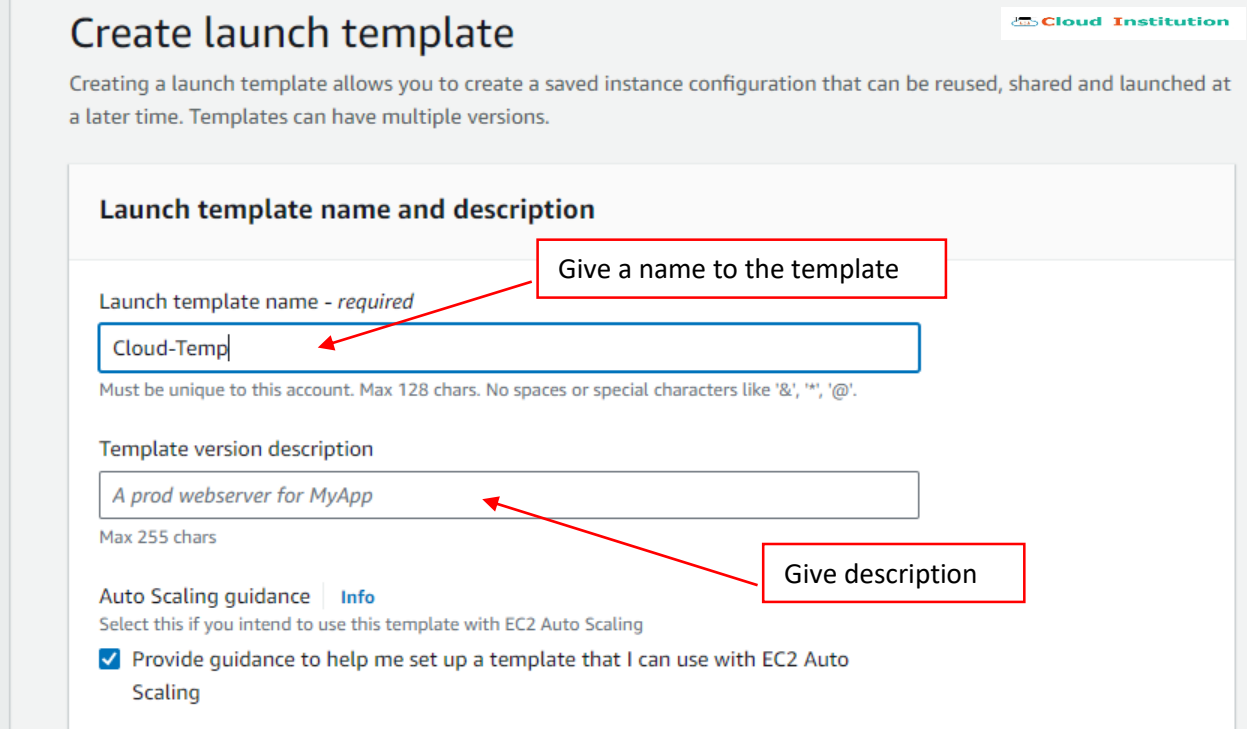
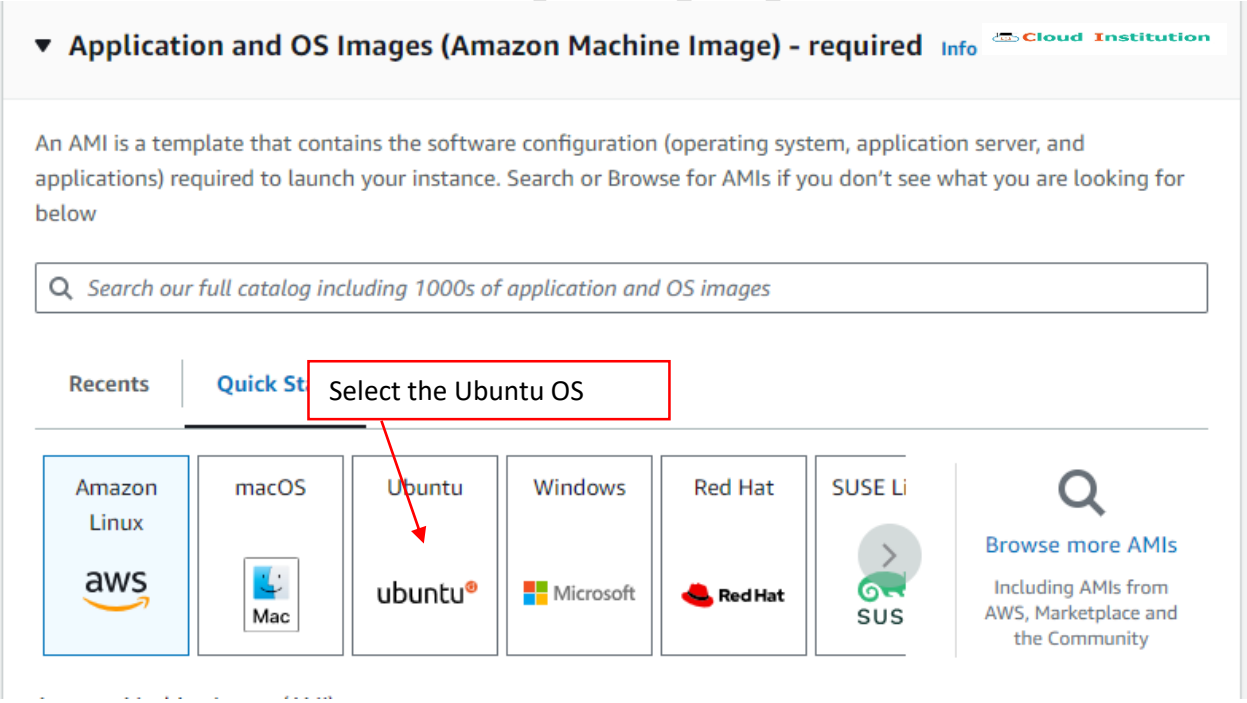
AUTO SCALING

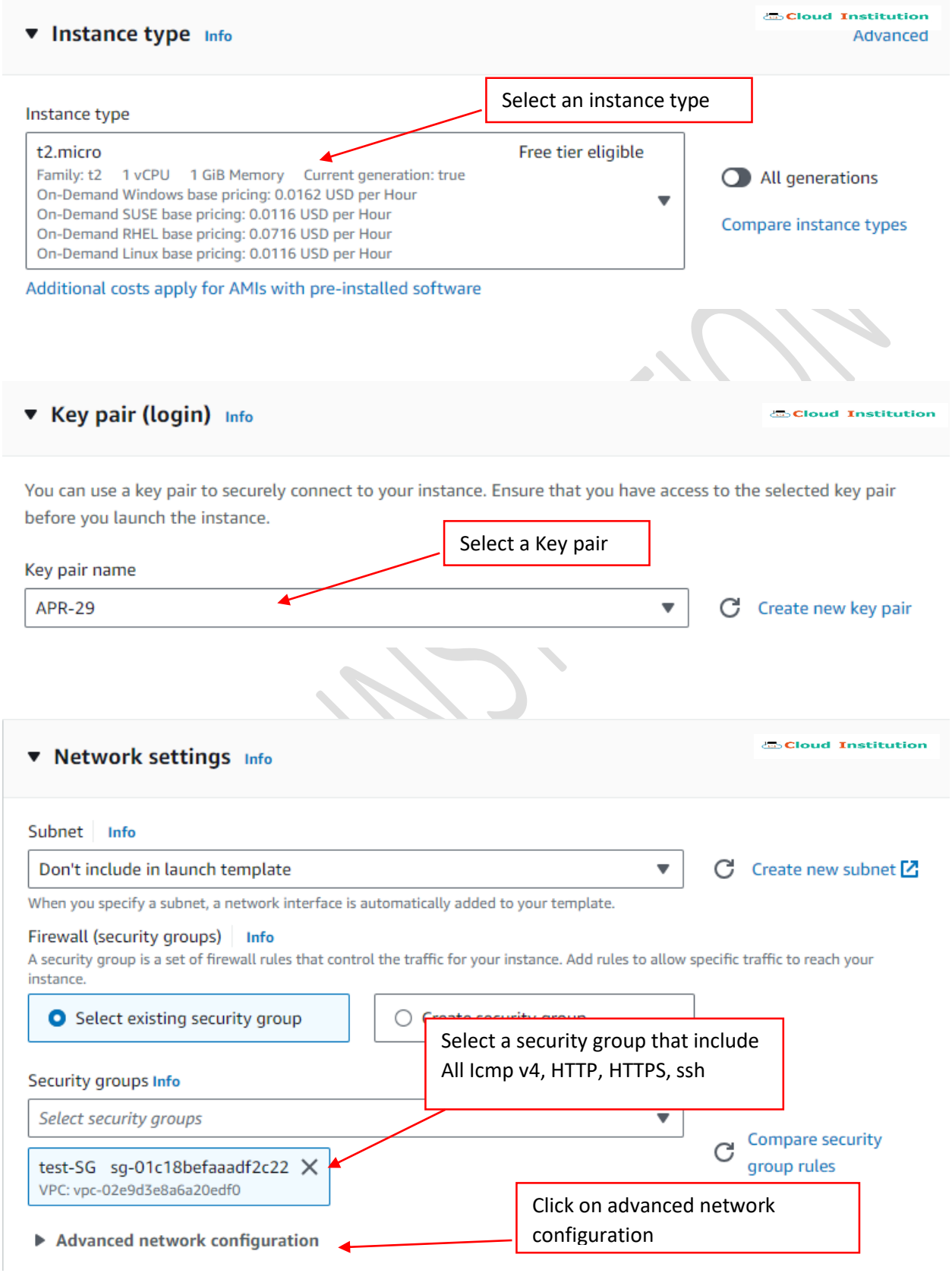
Step 1: Create a launch template

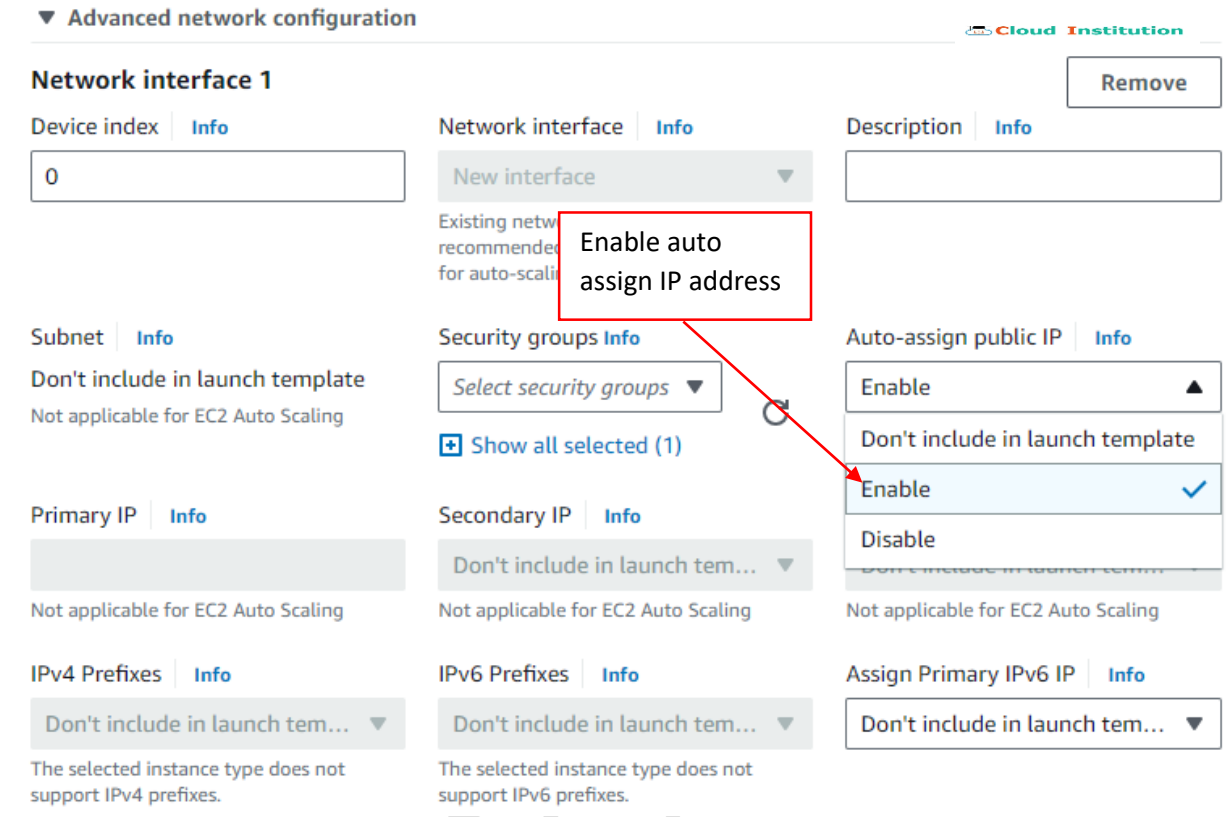


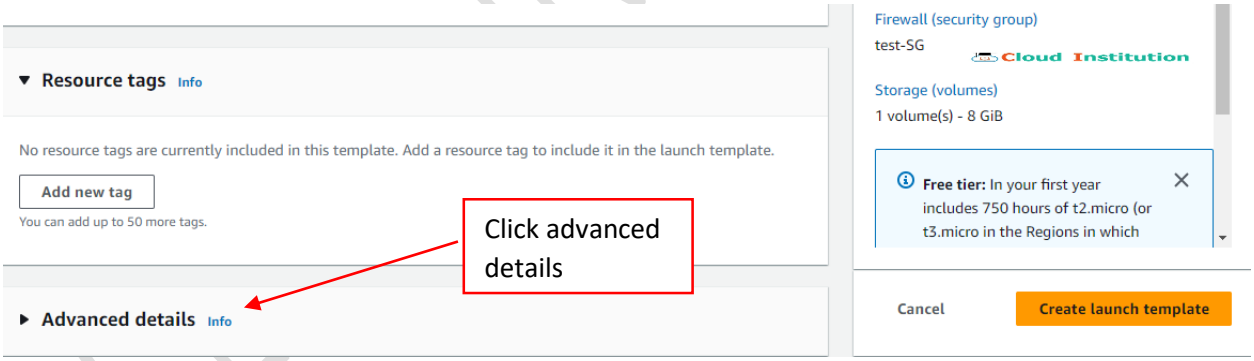



Create launch template
Cloud Institution
Creating a launch template allows you to create a saved instance configuration that can be reused, shared and launched at a later time. Templates can have multiple versions.
Launch template name and description
Give a name to the template
Launch template name - required
Cloud-Temp
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
Give description
Auto Scaling guidance
Info
Select this if you intend to use this template with EC2 Auto Scaling
Provide guidance to help me set up a template that I can use with EC2 Auto Scaling


Application and OS Images (Amazon Machine Image) - required
Info
Cloud Institution
An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below
Search our full catalog including 1000s of application and OS images
Recents
Quick St
Select the Ubuntu OS
Amazon Linux
aws
macOS
Mac
Ubuntu
ubuntu
Windows
Microsoft
Red Hat
Red Hat
SUSE Li
SUS
Browse more AMIs
Including AMIs from AWS, Marketplace and the Community


Instance type Info
Cloud Institution
Advanced
Select an instance type
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
Key pair (login) Info
Cloud Institution
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.
Select a Key pair
Key pair name
APR-29
Create new key pair
Network settings Info
Cloud Institution
Subnet Info
Don't include in launch template
Create new subnet
When you specify a subnet, a network interface is automatically added to your template.
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
Select a security group that include
All Icmp v4, HTTP, HTTPS, ssh
Security groups Info
Select security groups
test-SG sg-01c18befaaadf2c22
VPC: vpc-02e9d3e8a6a20edf0
Compare security group rules
Click on advanced network configuration
Advanced network configuration





Scroll down and find USER DATA

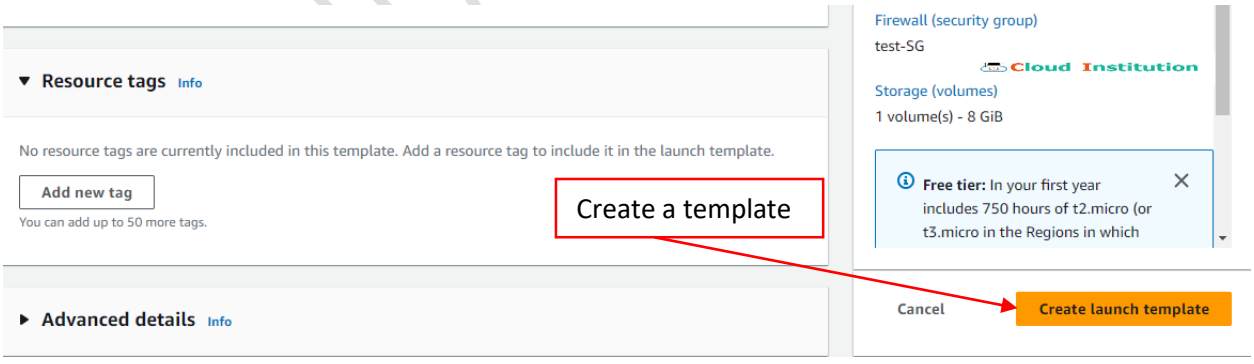
User data - *optional* Info

Upload a file with your user data or enter it in the field.

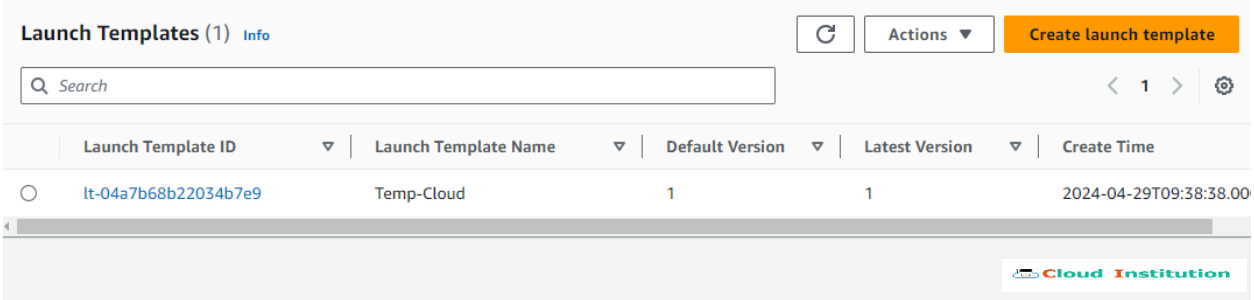
Cloud Institution

Choose file

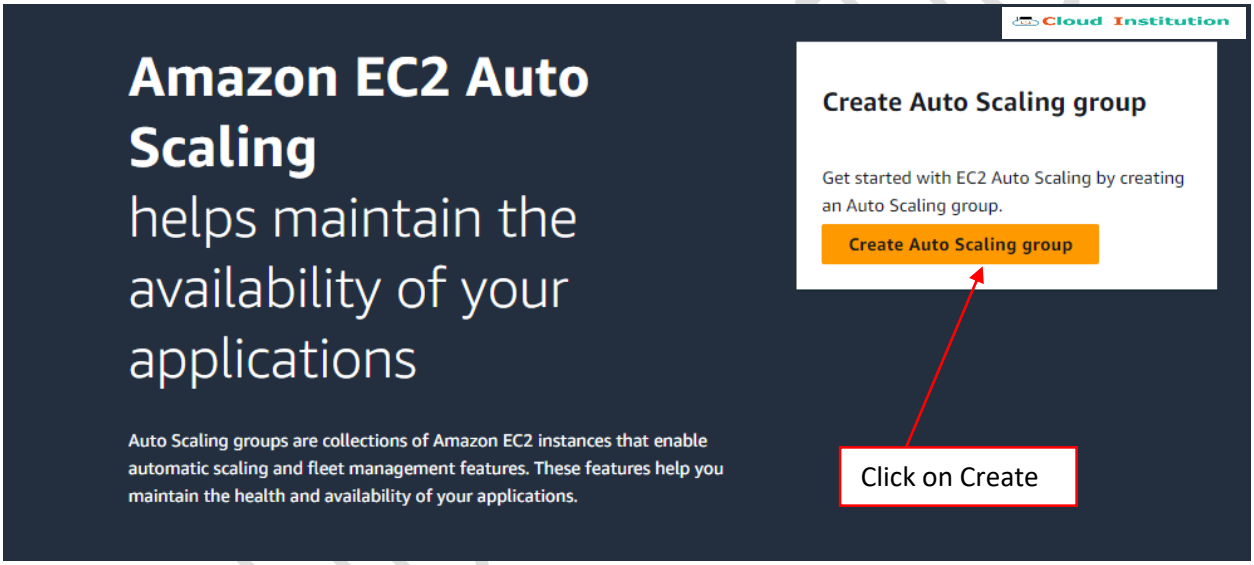
```
#!/bin/bash
yes | sudo apt update
yes | sudo apt install apache2
echo "<h1>Server Details</h1><p><strong>Hostname:</strong> $(hostname)
</p><p><strong>IP Address:</strong> $(hostname -I | cut -d" " -f1)</p>" >
/var/www/html/index.html
sudo systemctl restart apache2
```

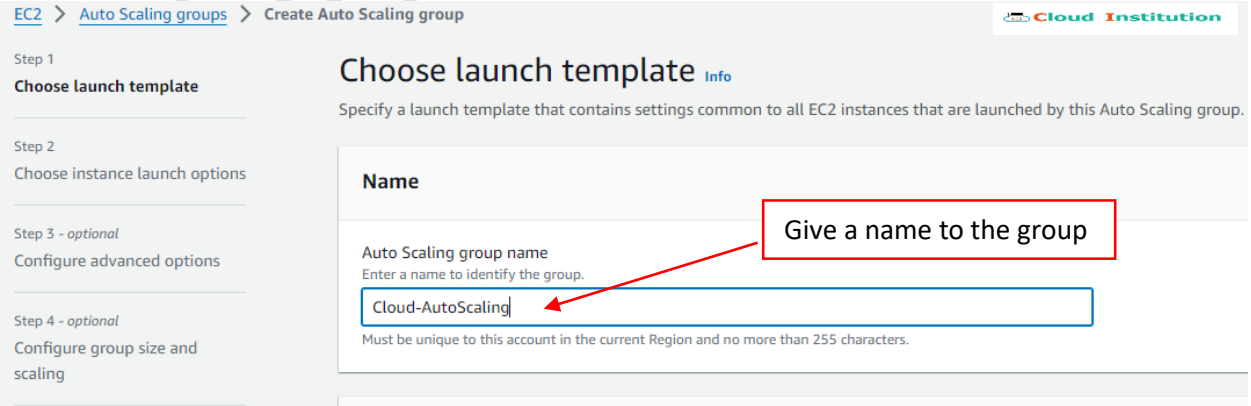


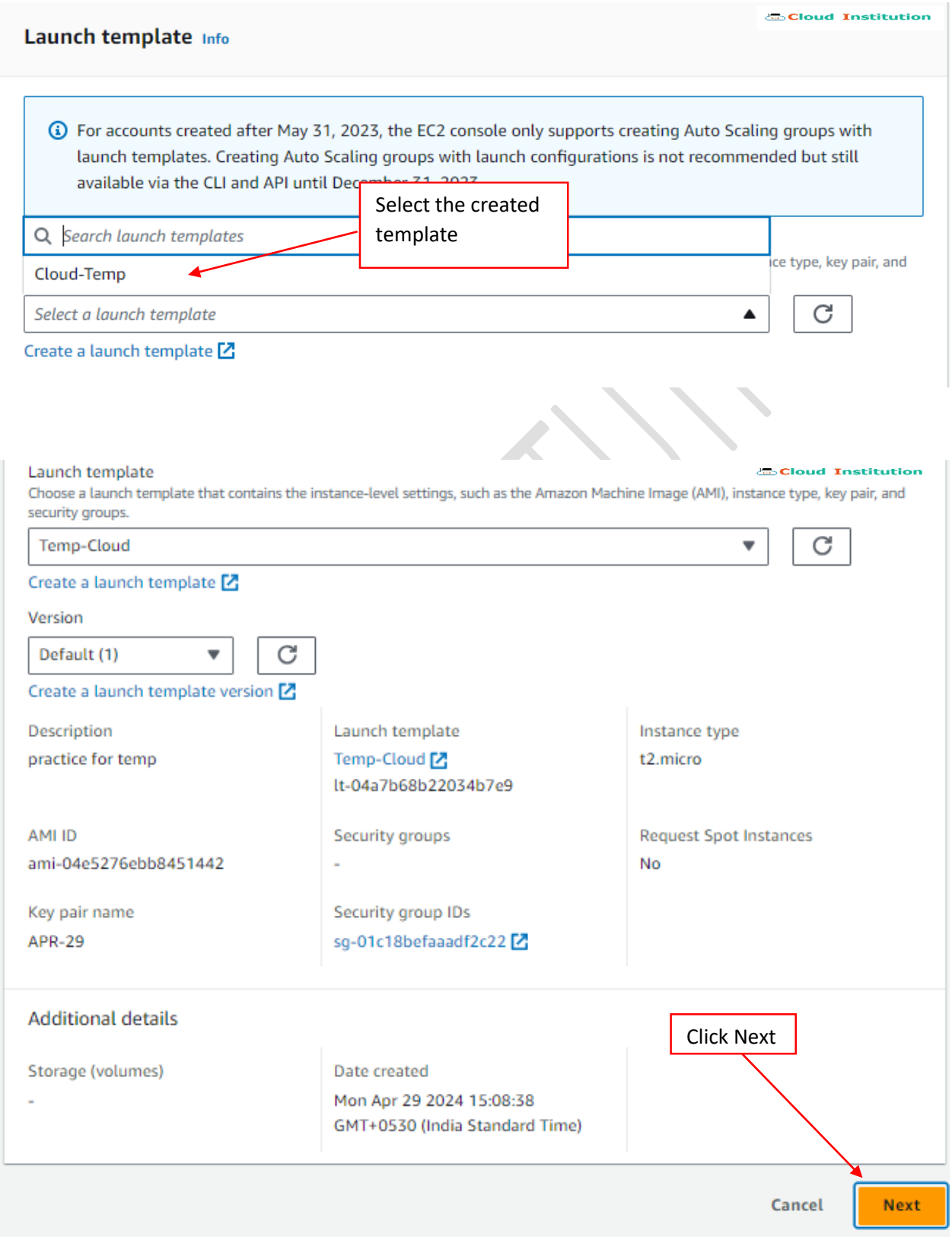
Template Created

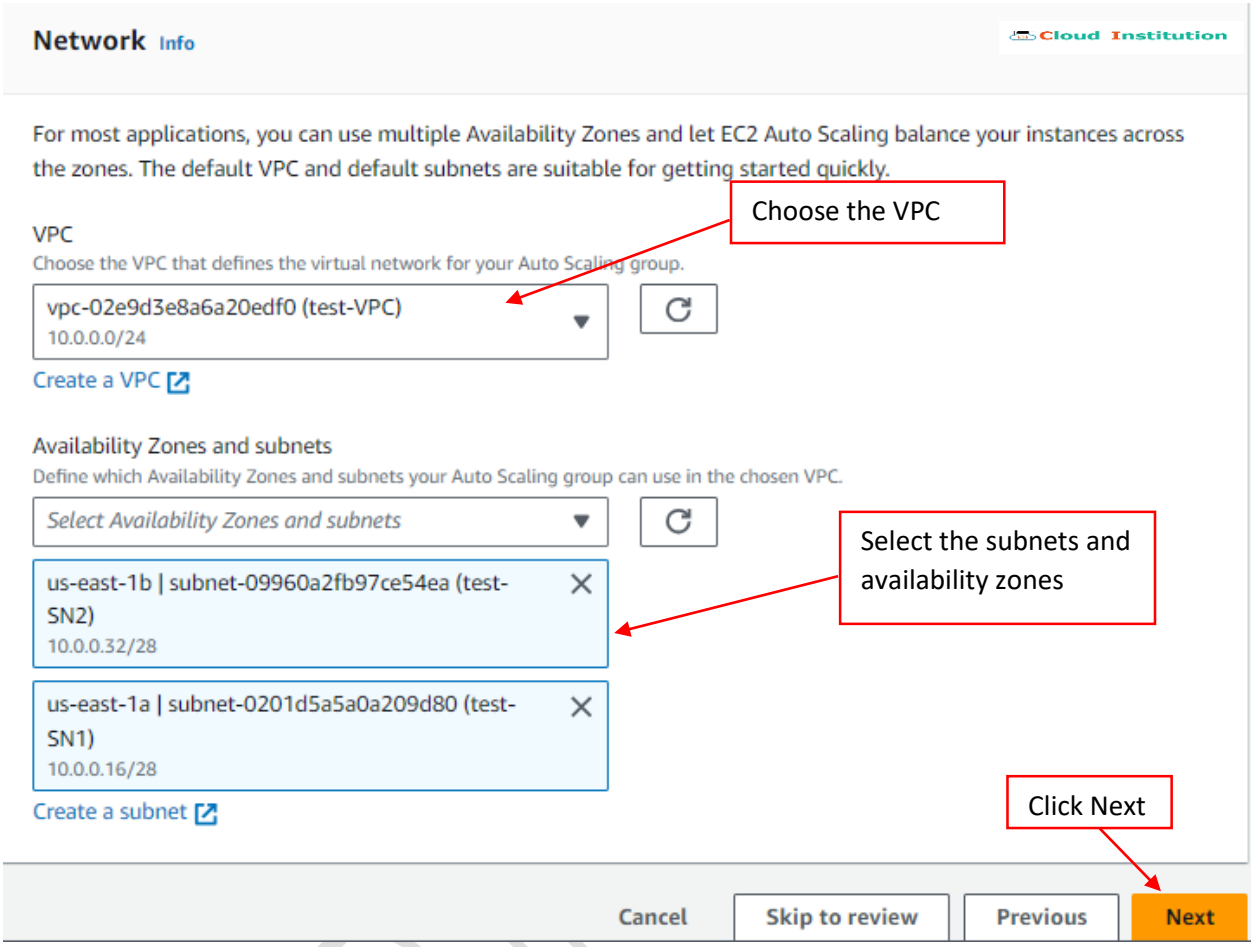


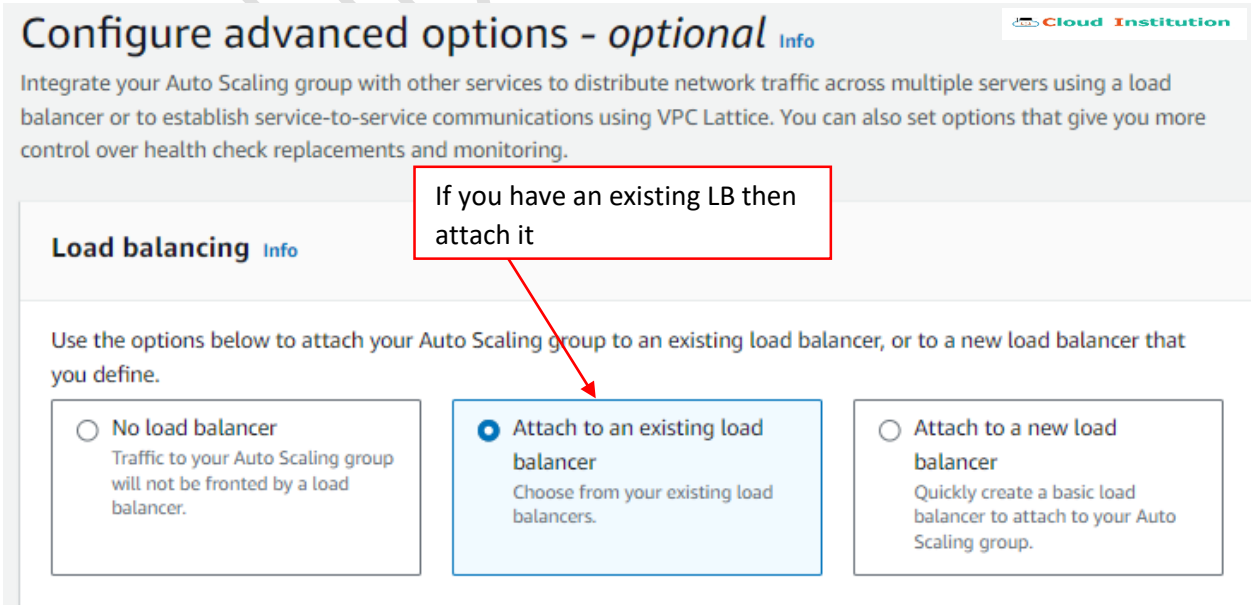
Step 2 : Create auto scaling group

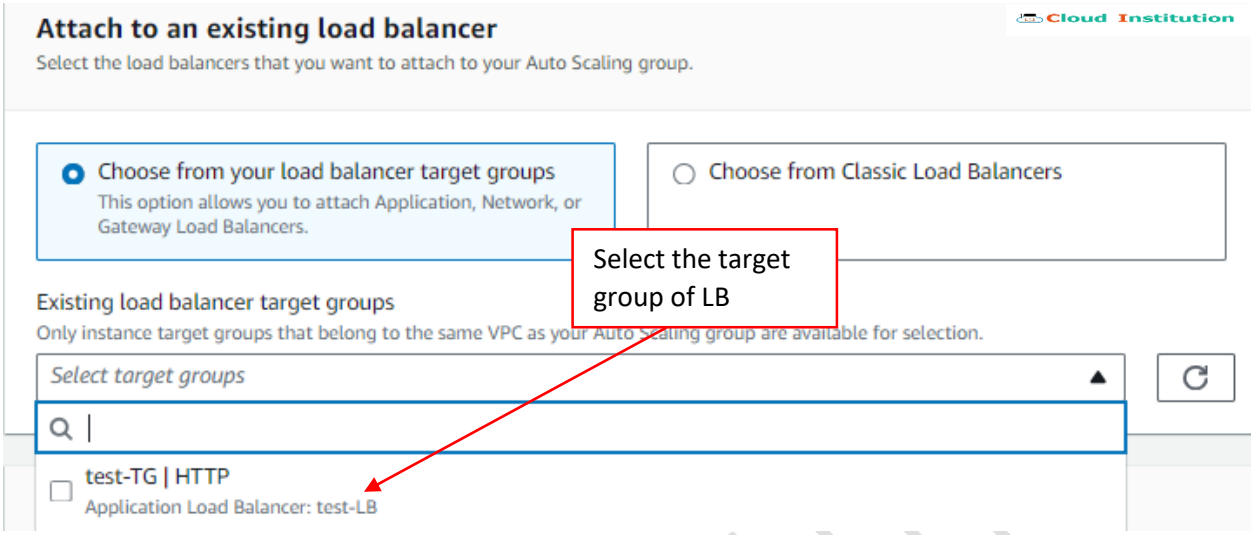


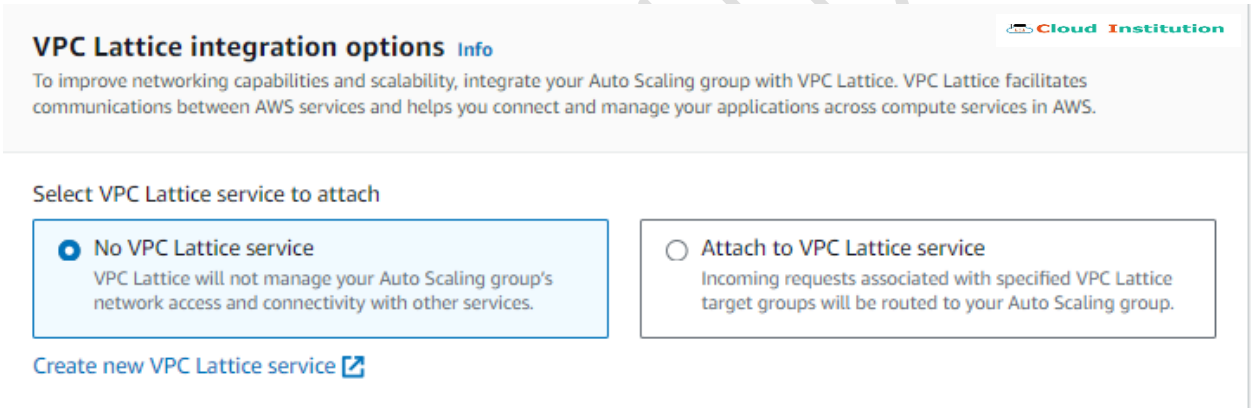


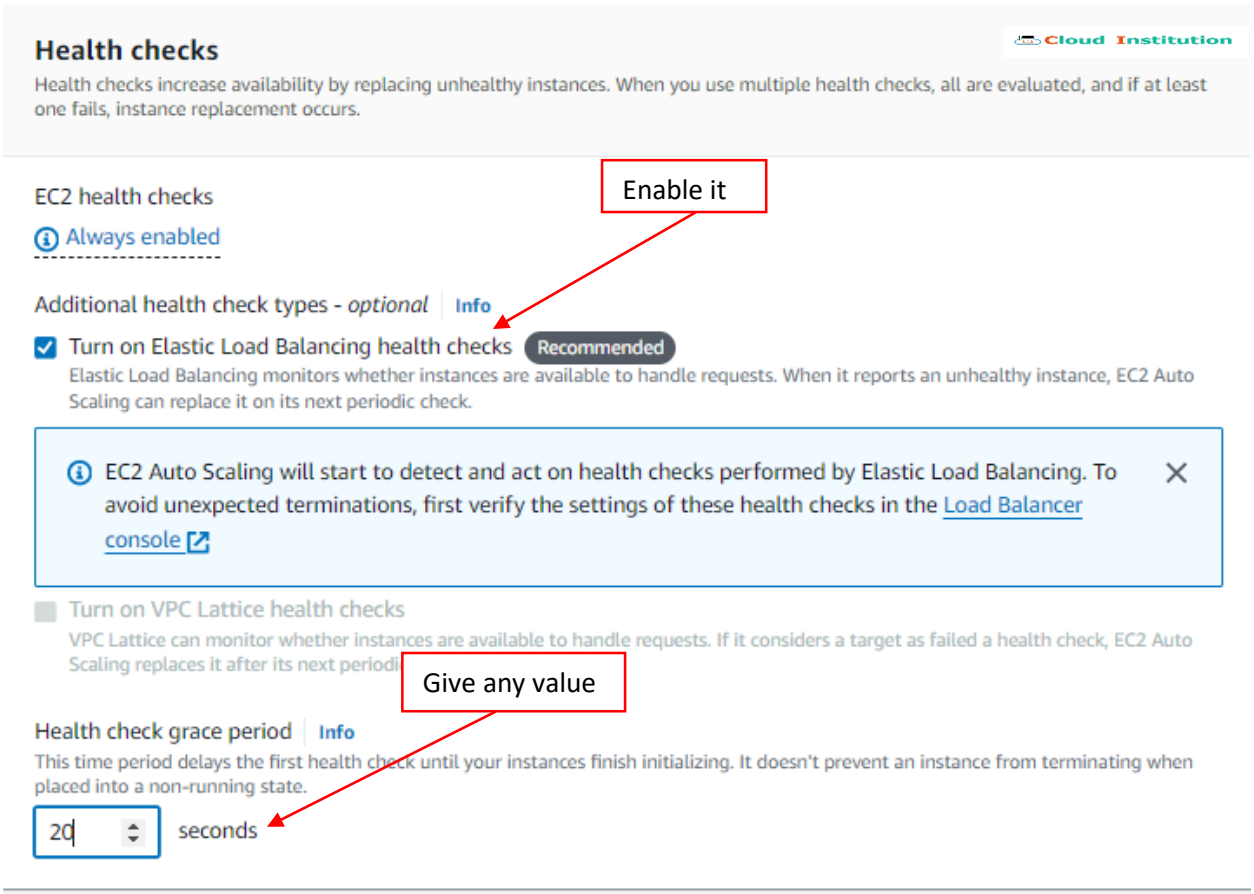

Launch template Info
Cloud Institution
For accounts created after May 31, 2023, the EC2 console only supports creating Auto Scaling groups with launch templates. Creating Auto Scaling groups with launch configurations is not recommended but still available via the CLI and API until
Select the created template
Search launch templates
Cloud-Temp
Select a launch template
Create a launch template
Launch template
Choose a launch template that contains the instance-level settings, such as the Amazon Machine Image (AMI), instance type, key pair, and security groups.
Temp-Cloud
Create a launch template
Version
Default (1)
Create a launch template version
Description
practice for temp
Launch template
Temp-Cloud
lt-04a7b68b22034b7e9
Instance type
t2.micro
AMI ID
ami-04e5276ebb8451442
Security groups
-
Request Spot Instances
No
Key pair name
APR-29
Security group IDs
sg-01c18befaaadf2c22
Additional details
Storage (volumes)
-
Date created
Mon Apr 29 2024 15:08:38
GMT+0530 (India Standard Time)
Click Next
Cancel
Next

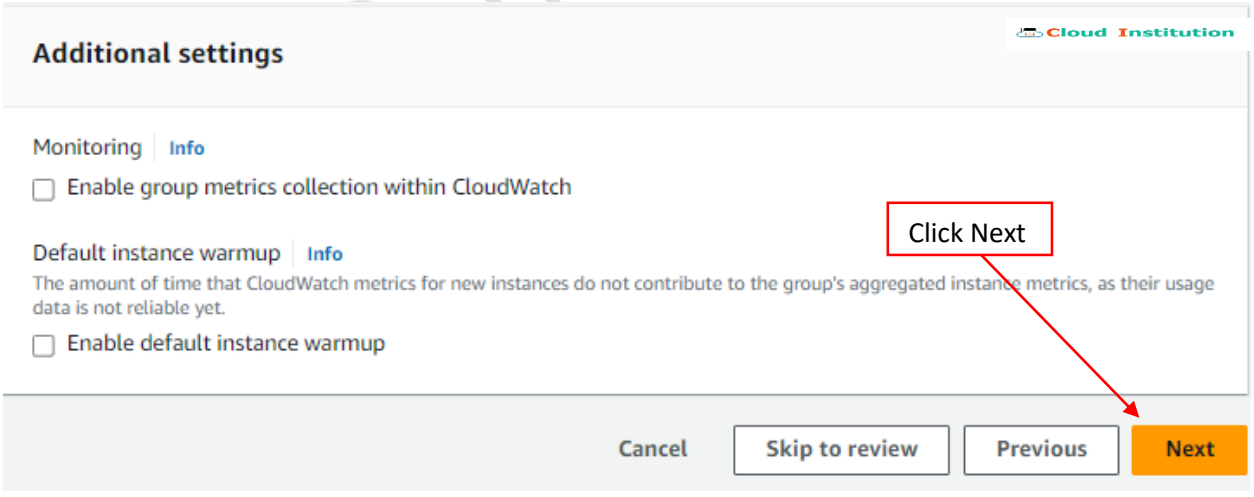

Network Info
Cloud Institution
For most applications, you can use multiple Availability Zones and let EC2 Auto Scaling balance your instances across the zones. The default VPC and default subnets are suitable for getting started quickly.
Choose the VPC
VPC
Choose the VPC that defines the virtual network for your Auto Scaling group.
vpc-02e9d3e8a6a20edf0 (test-VPC)
10.0.0.0/24
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
Select the subnets and availability zones
us-east-1b | subnet-09960a2fb97ce54ea (test-SN2)
10.0.0.32/28
us-east-1a | subnet-0201d5a5a0a209d80 (test-SN1)
10.0.0.16/28
Create a subnet
Click Next
Cancel
Skip to review
Previous
Next

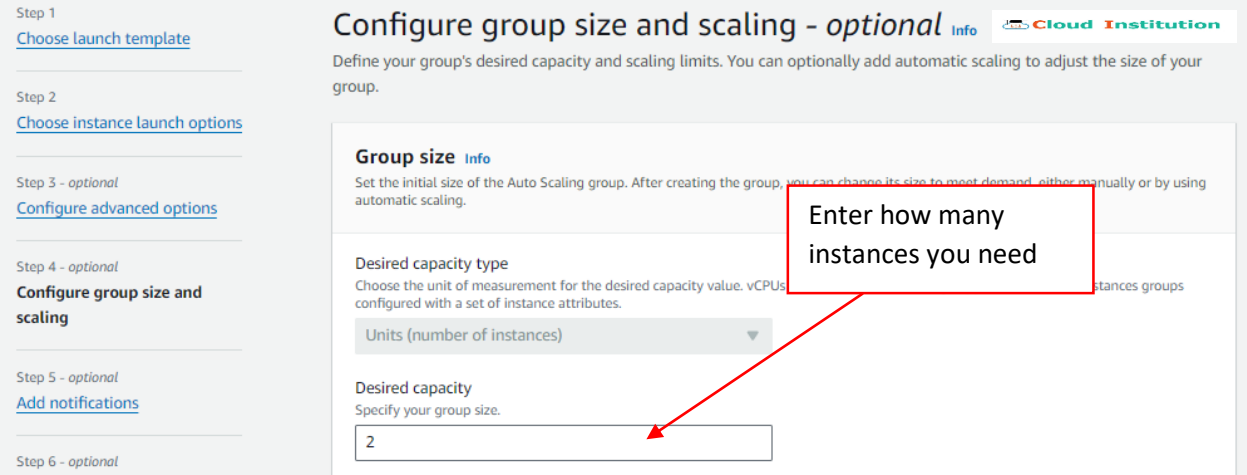

Configure advanced options - optional Info
Cloud Institution
Integrate your Auto Scaling group with other services to distribute network traffic across multiple servers using a load balancer or to establish service-to-service communications using VPC Lattice. You can also set options that give you more control over health check replacements and monitoring.
If you have an existing LB then attach it
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.

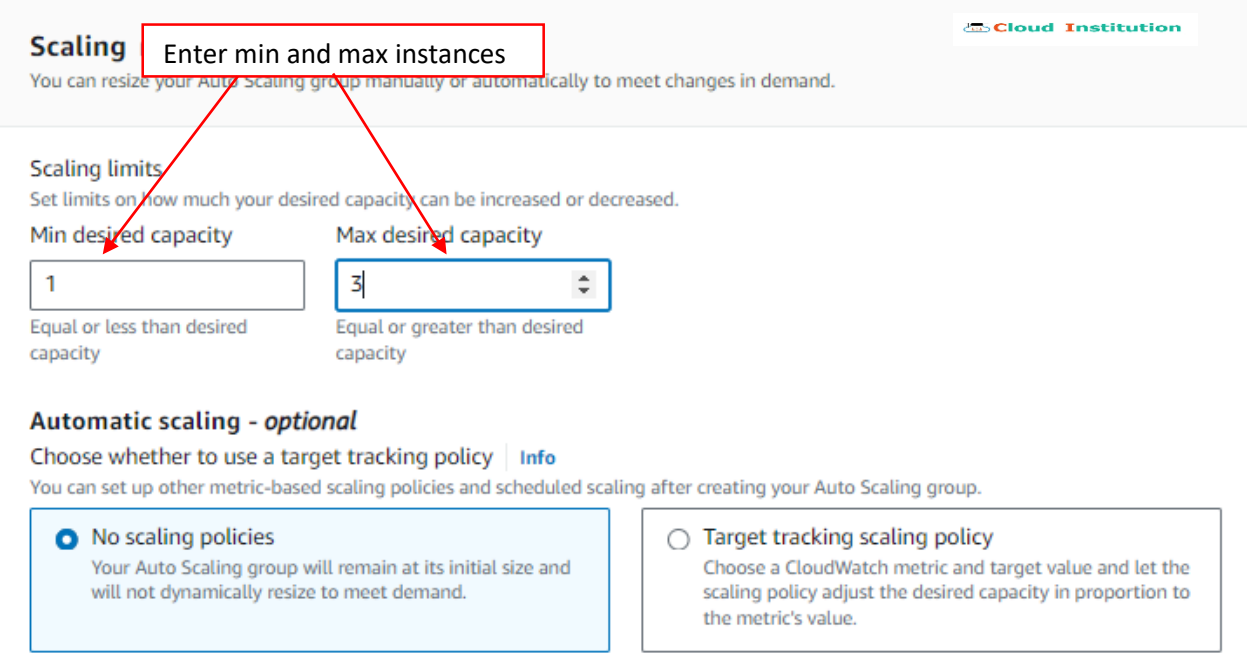

Attach to an existing load balancer
Cloud Institution
Select the load balancers that you want to attach to your Auto Scaling group.
Choose from your load balancer target groups
This option allows you to attach Application, Network, or Gateway Load Balancers.
Choose from Classic Load Balancers
Select the target group of LB
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
test-TG | HTTP
Application Load Balancer: test-LB

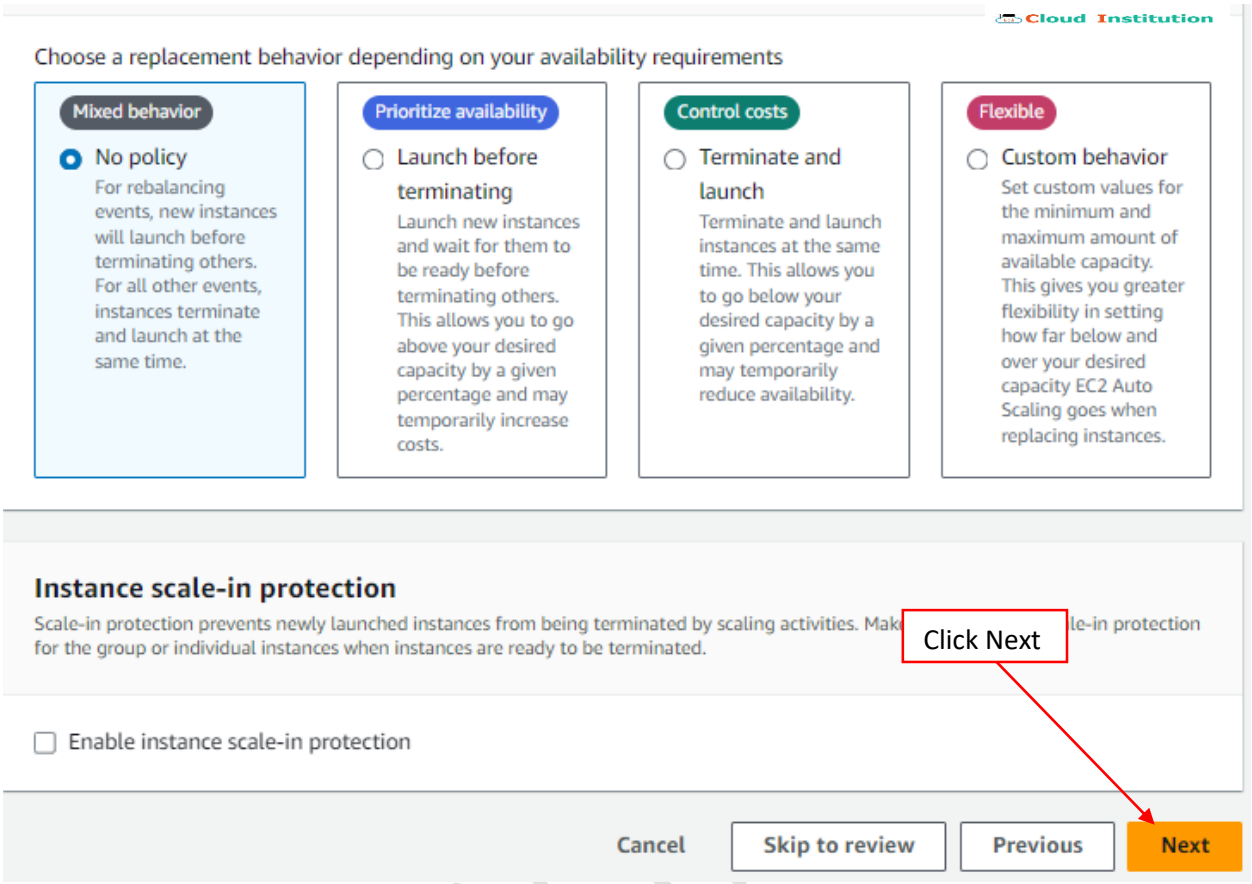

VPC Lattice integration options Info
Cloud Institution
To improve networking capabilities and scalability, integrate your Auto Scaling group with VPC Lattice. VPC Lattice facilitates communications between AWS services and helps you connect and manage your applications across compute services in AWS.
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

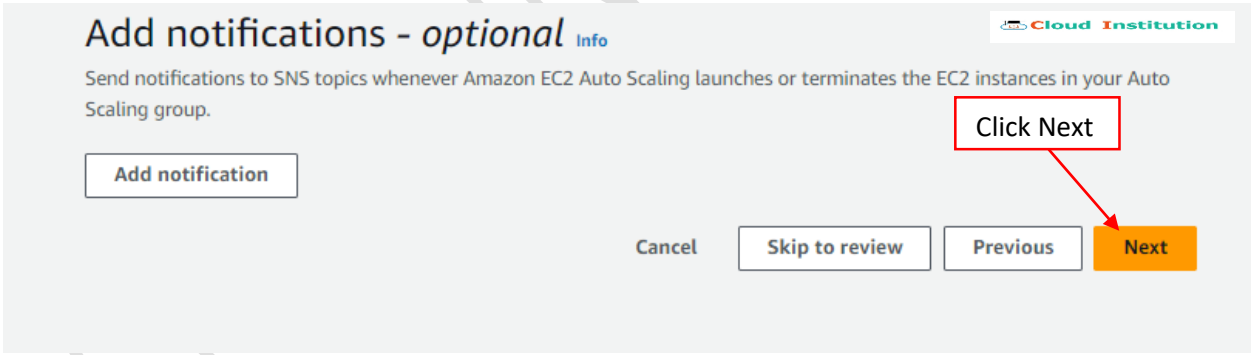

Health checks
Cloud Institution
Health checks increase availability by replacing unhealthy instances. When you use multiple health checks, all are evaluated, and if at least one fails, instance replacement occurs.
EC2 health checks
Always enabled
Enable it
Additional health check types - optional
Info
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.
EC2 Auto Scaling will start to detect and act on health checks performed by Elastic Load Balancing. To avoid unexpected terminations, first verify the settings of these health checks in the Load Balancer console
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 periodi
Give any value
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.
20
seconds

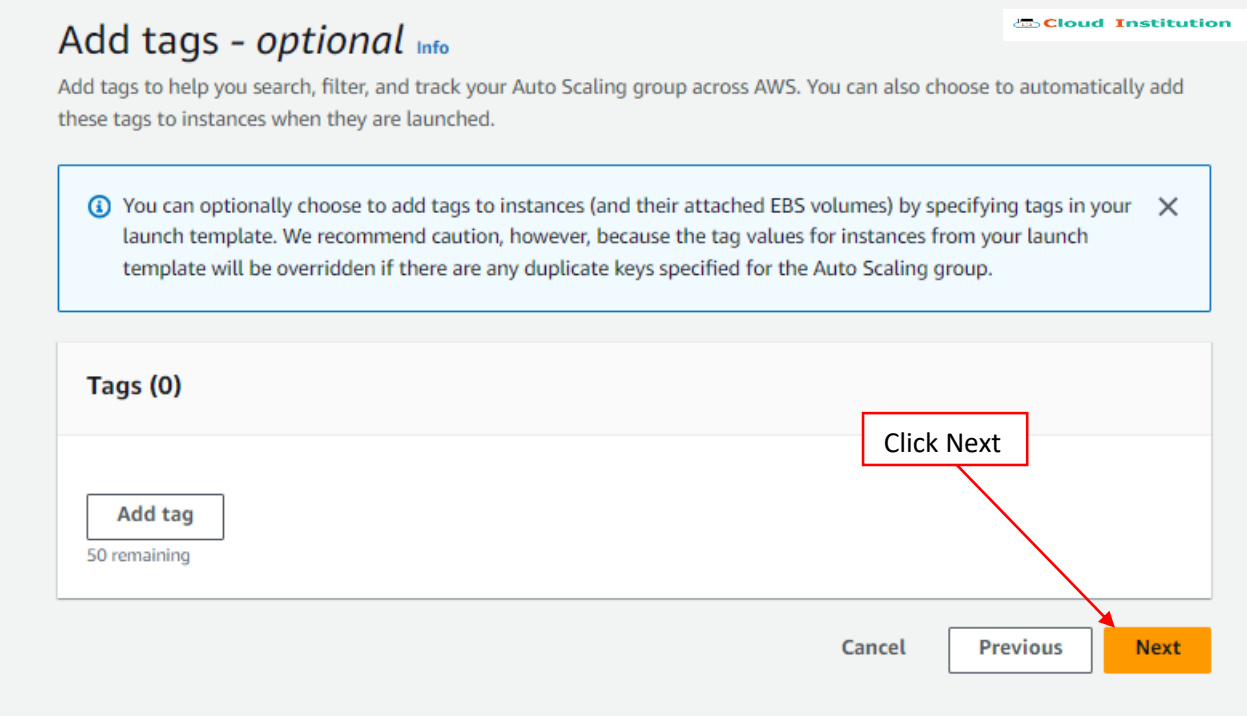

Additional settings
Cloud Institution
Monitoring
Info
Enable group metrics collection within CloudWatch
Default instance warmup
Info
The amount of time that CloudWatch metrics for new instances do not contribute to the group's aggregated instance metrics, as their usage data is not reliable yet.
Enable default instance warmup
Click Next
Cancel
Skip to review
Previous
Next

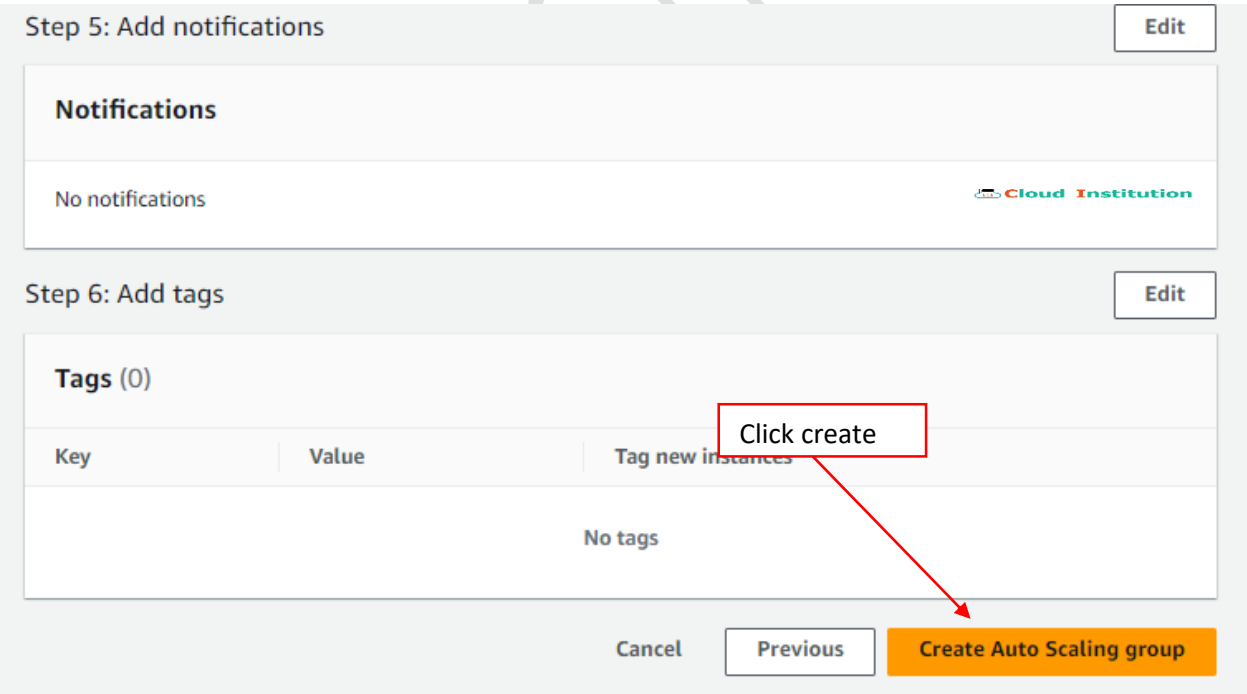

Step 1
Choose launch template
Step 2
Choose instance launch options
Step 3 - optional
Configure advanced options
Step 4 - optional
Configure group size and scaling
Step 5 - optional
Add notifications
Step 6 - optional
Configure group size and scaling - optional Info
Cloud Institution
Define your group's 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.
Enter how many instances you need
Desired capacity type
Choose the unit of measurement for the desired capacity value. vCPUs
configured with a set of instance attributes.
Units (number of instances)
Desired capacity
Specify your group size.
2


Cloud Institution
Scaling
Enter min and max instances
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
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.

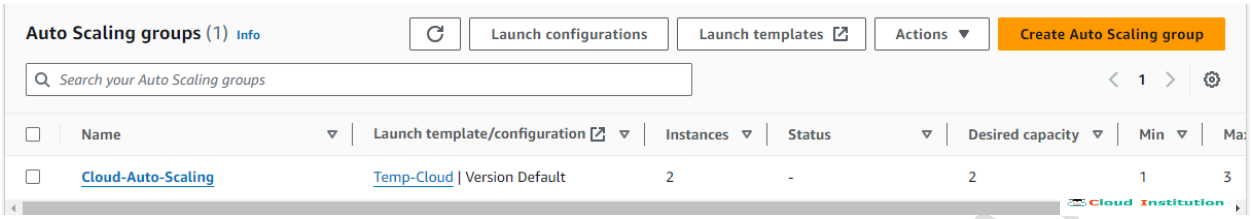

Cloud Institution
Choose a replacement behavior depending on your availability requirements
Mixed behavior
No policy
For rebalancing events, new instances will launch before terminating others. For all other events, instances terminate and launch at the same time.
Prioritize availability
Launch before terminating
Launch new instances and wait for them to be ready before terminating others. This allows you to go above your desired capacity by a given percentage and may temporarily increase costs.
Control costs
Terminate and launch
Terminate and launch instances at the same time. This allows you to go below your desired capacity by a given percentage and may temporarily reduce availability.
Flexible
Custom behavior
Set custom values for the minimum and maximum amount of available capacity. This gives you greater flexibility in setting how far below and over your desired capacity EC2 Auto Scaling goes when replacing instances.
Instance scale-in protection
Scale-in protection prevents newly launched instances from being terminated by scaling activities. Mak
le-in protection for the group or individual instances when instances are ready to be terminated.
Click Next
Enable instance scale-in protection
Cancel
Skip to review
Previous
Next


Add notifications - optional Info
Cloud Institution
Send notifications to SNS topics whenever Amazon EC2 Auto Scaling launches or terminates the EC2 instances in your Auto Scaling group.
Click Next
Add notification
Cancel
Skip to review
Previous
Next

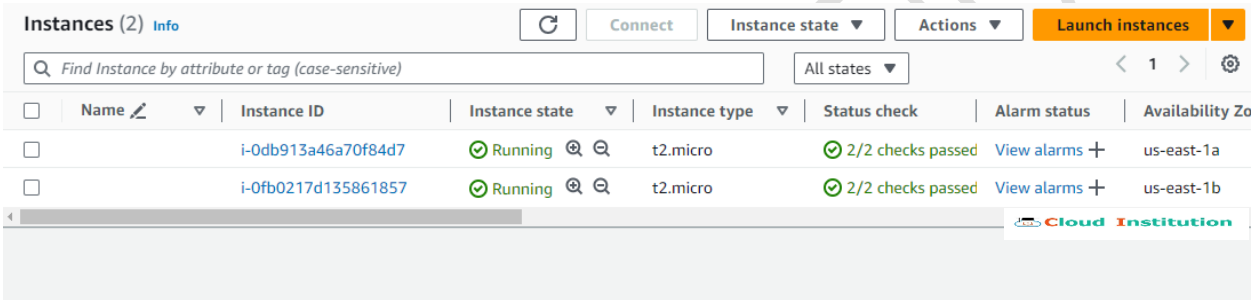

Cloud Institution
Add tags - optional Info
Add tags to help you search, filter, and track your Auto Scaling group across AWS. You can also choose to automatically add these tags to instances when they are launched.
You can optionally choose to add tags to instances (and their attached EBS volumes) by specifying tags in your launch template. We recommend caution, however, because the tag values for instances from your launch template will be overridden if there are any duplicate keys specified for the Auto Scaling group.
Tags (0)
Click Next
Add tag
50 remaining
Cancel
Previous
Next


Step 5: Add notifications
Edit
Notifications
No notifications
Cloud Institution
Step 6: Add tags
Edit
Tags (0)
Click create
Key
Value
Tag new instances
No tags
Cancel
Previous
Create Auto Scaling group

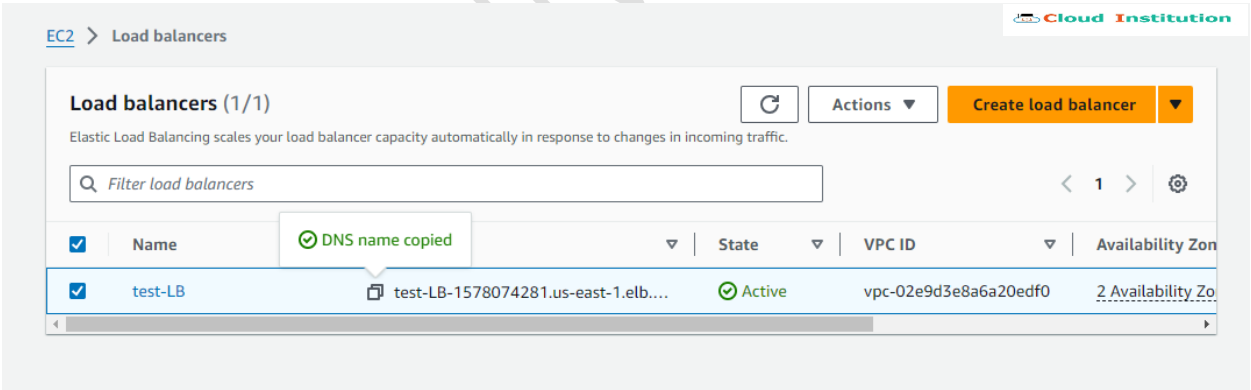
Auto Scaling group created



Now you can see the desired number of instances running



Go to the Load Balancers -> Copy the DNS name and paste it in the browser



Result :

Server Details

Hostname: ip-12-0-3-127

IP Address: 12.0.3.127