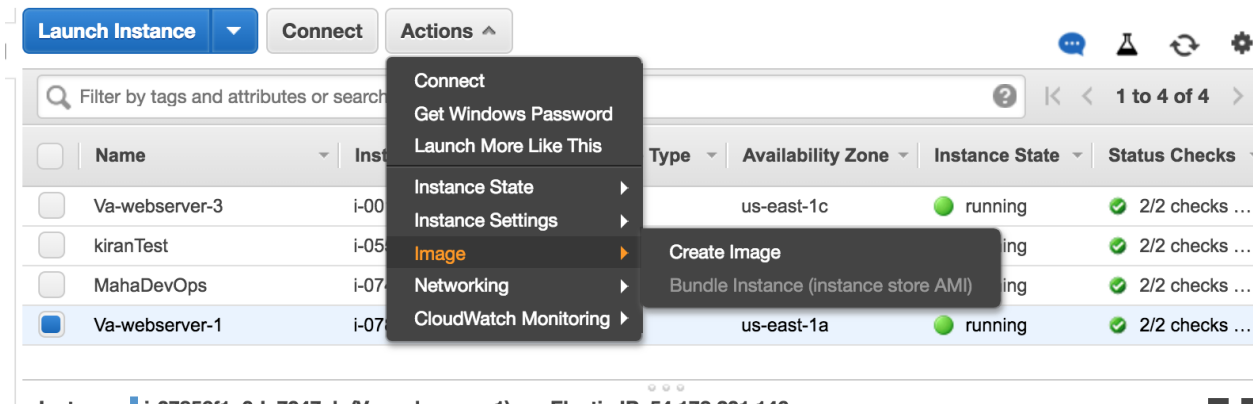


Step 0: Start by creating a fresh Security Group in our VPC.

*****Important: Use this SG for both the LB and Autoscaling group*****

Step 1: Create an AMI from one of your web servers. Make sure the web server is working (showing the index page when you browse the IP on a web browser)



Step2: Give the AMI a name and description and click create Image

The screenshot shows the 'Create Image' dialog in the AWS Management Console. The dialog has a title bar with a close button. Below the title bar, there are fields for 'Instance ID' (i-07856f1a0da7947cb), 'Image name' (my-va-web-server1), 'Image description' (f our webserver to use with the autoscaling Group), and 'No reboot' (checkbox). Below these fields is a section for 'Instance Volumes' which contains a table with columns: Volume Type, Device, Snapshot, Size (GiB), Volume Type, IOPS, Throughput (MB/s), Delete on Termination, and Encrypted. The table has one row for the 'Root' volume. At the bottom of the dialog, there is a 'Total size of EBS Volumes: 30 GiB' message and a note about EBS snapshots. The 'Create Image' button is highlighted in blue.

Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encrypted
Root	/dev/sda1	snap-071b20d015998ae85	30	General Purpose 5	100 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypted

Step 3: Wait until the AMI is created

Step 4: Click on the Autoscaling section on the left navigation Pane. Then select create Launch Configuration

Bundle Tasks

ELASTIC BLOCK STORE

Volumes

Snapshots

NETWORK & SECURITY

Security Groups

Elastic IPs

Placement Groups

Key Pairs

Network Interfaces

LOAD BALANCING

Load Balancers

Target Groups

AUTO SCALING

Launch Configurations

Auto Scaling Groups

Welcome to Auto Scaling

You can use Auto Scaling to manage Amazon EC2 capacity automatically, maintain the right number of instances for your application, operate a healthy group of instances, and scale it according to your needs.
[Learn more](#)

You have the following Auto Scaling resources in the US East (N. Virginia) region

Auto Scaling Group: 1

Create Auto Scaling group


Launch Configurations: 2

Create launch configuration

Note: To create your Auto Scaling groups in a different region, select your region from the navigation bar.


Benefits of Auto Scaling

Reusable Instance Templates




Provision instances based on

Automated Provisioning



Keep your Auto Scaling group healthy and balanced, whether

Adjustable Capacity



Maintain a fixed group size or adjust dynamically based on

Additional I

[Getting Started](#)
[Documentation](#)
[All EC2 Resourc](#)
[Forums](#)
[Pricing](#)
[Contact Us](#)

Step 5: Choose My AMI in the left Navigation Pane. Then select your AMI

Step 6: On Page 3 give the launch config a name. Then Click advanced details.

1. Choose AMI 2. Choose Instance Type 3. Configure details 4. Add Storage 5. Configure Security Group 6. Review

Create Launch Configuration

Name ⓘ

My-launch-config

Purchasing option ⓘ

☐ Request Spot Instances

IAM role ⓘ

None

Monitoring ⓘ

☐ Enable CloudWatch detailed monitoring

[Learn more](#)

► Advanced Details

Step 7: Select assign a Public IP to every instance

▼ Advanced Details

Kernel ID	<input type="button" value="i"/>	Use default
RAM Disk ID	<input type="button" value="i"/>	Use default
User data	<input type="button" value="i"/>	<input checked="" type="radio"/> As text <input type="radio"/> As file <input type="checkbox"/> Input is already base64 encoded <div>(Optional)</div>
IP Address Type	<input type="button" value="i"/>	<input type="radio"/> Only assign a public IP address to instances launched in the default VPC and subnet. (default) <input checked="" type="radio"/> Assign a public IP address to every instance. <input type="radio"/> Do not assign a public IP address to any instances. Note: this option only affects instances launched into an Amazon VPC

Step 8: Click Next

Step 9: Create a new Security Group which comes with RDP and add rule HTTP source 0.0.0.0/0

Create Launch Configuration

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group: ☒ Create a new security group
☐ Select an existing security group

Security group name: AutoScaling-Security-Group-1

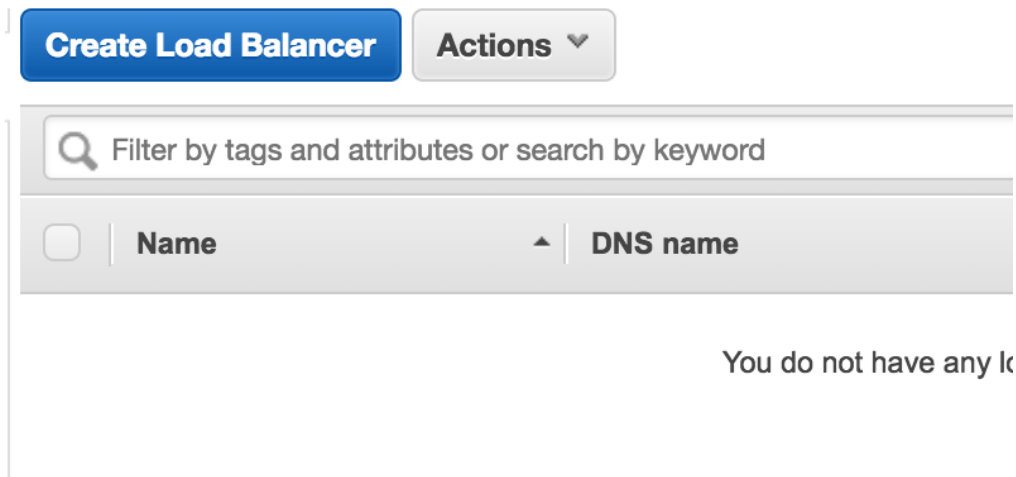
Description: AutoScaling-Security-Group-1 (2018-06-01 13:38:23.154-04:00)

Type <input type="button" value="i"/>	Protocol <input type="button" value="i"/>	Port Range <input type="button" value="i"/>	Source <input type="button" value="i"/>
RDP	TCP	3389	Anywhere 0.0.0.0/0
HTTP	TCP	80	Anywhere 0.0.0.0/0
Add Rule			

Step 10: Review and create Launch Configuration

Step 12: Go to the Load Balancer Page

Step 13: Create a load balancer



Step 14:
Select the classic Load Balancer

Step 15: Give the ELB a name. Pick the VPC that you created earlier and add the public subnets by clicking the + sign

Step 1: Define Load Balancer

Basic Configuration

This wizard will walk you through setting up a new load balancer. Begin by giving your new load balancer a unique name so that you can identify it from other load balancers you might create. You will also need to configure ports and protocols for your load balancer. Traffic from your clients can be routed from any load balancer port to any port on your EC2 instances. By default, we've configured your load balancer with a standard web server on port 80.

Load Balancer name:

Create LB Inside:

Create an internal load balancer: ☐ [\(what's this?\)](#)

Enable advanced VPC configurations: ☐

Listener Configuration:

Load Balancer Protocol	Load Balancer Port	Instance Protocol	Instance Port
HTTP	80	HTTP	80

[Add](#)

Select Subnets

You will need to select a Subnet for each Availability Zone where you wish traffic to be routed by your load balancer. If you have instances in only one Availability Zone, please select at least two Subnets in different Availability Zones to provide higher availability for your load balancer.

VPC vpc-37d7ec4c (10.0.0.0/16) | VA-VPC-10.0.0.0/16

Available subnets

Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
	us-east-1b	subnet-ac468582	10.0.2.0/24	VA-Subnet-2-Private-10.0.2.0/24

Selected subnets

Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
	us-east-1a	subnet-661bd201	10.0.1.0/24	VA-Subnet-1-Public-10.0.1.0/24
	us-east-1c	subnet-4bf66201	10.0.3.0/24	VA-Subnet-3-Public-10.0.3.0/24

Step 16: Select a SG on the same VPC and click next

Step 17:

Step 4: Configure Health Check

Your load balancer will automatically perform health checks on your EC2 instances and only route traffic to instances that pass the health check. If an instance fails the health check, it is automatically removed from the load balancer. Customize the health check to meet your specific needs.

Ping Protocol	HTTP
Ping Port	80
Ping Path	/index.html

Advanced Details

Response Timeout	5	seconds
Interval	10	seconds
Unhealthy threshold	2	
Healthy threshold	3	

Step 18: ON page 5, don't add any instances. Click next add tags. Tags are optional. Click review and create and Finally click create.

Step 19: Select the Launch Config and click on create autoscaling group

Step 20: Give the autoscaling group a name. Select 2 for no of instances. Select the VPC we created earlier. Then select the public subnets.

Create Auto Scaling Group

[Canc](#)

Launch Configuration	My-launch-config
Group name	myAutoscaleELB
Group size	Start with 2 instances
Network	vpc-37d7ec4c (10.0.0.0/16) VA-VPC-10.0.0.0/16 Create new VPC
Subnet	<div>subnet-661bd201(10.0.1.0/24) VA-Subnet-1-Public-10.0.1.0/24 us-east-1a x</div> <div>subnet-4bf66201(10.0.3.0/24) VA-Subnet-3-Public-10.0.3.0/24 us-east-1c x</div> <div>Create new subnet</div>

Each instance in this Auto Scaling group will be assigned a public IP address. [i](#)

Step 21: Click on Advanced details

Step 22: Make sure you select the receive traffic from one or more load balancers and select the Classic load balancer we created earlier.

▼ Advanced Details

Load Balancing ⓘ ☒ Receive traffic from one or more load balancers [Learn about Elastic Load Balancing](#)

Classic Load Balancers ⓘ

Target Groups ⓘ

Health Check Type ⓘ ☐ ELB ☒ EC2

Health Check Grace Period ⓘ seconds

Step 23: Click Configure Scaling policies

Create Auto Scaling Group

You can optionally add scaling policies if you want to adjust the size (number of instances) of your group automatically. A scaling policy is a set of instructions for making such adjustments in response to an Amazon CloudWatch alarm that you assign to it. In an existing policy, you can choose to add or remove a specific number of instances or a percentage of the existing group size, or you can set the group to an exact size. When the alarm triggers, it will execute the policy and adjust the size of your group accordingly. [Learn more about scaling policies.](#)

- ☐ Keep this group at its initial size
- ☒ Use scaling policies to adjust the capacity of this group

Scale between and instances. These will be the minimum and maximum size of your group.

Scale Group Size

Name:

Metric type:

Target value:

Instances need: seconds to warm up after scaling

Disable scale-in: ☐

[Click here](#)

[Scale the Auto Scaling group using step or simple scaling policies](#) ⓘ

Step 24: Create an alarm for increasing and decreasing instance size. First let's setup the increase group size settings.

Increase Group Size

Name:

Execute policy when: [Add new alarm](#)

Take the action: instances [Add step](#) ⓘ

Instances need: seconds to warm up after each step

[Create a simple scaling policy](#) ⓘ

Decrease Group Size

Name:

Execute policy when: [Add new alarm](#)

Take the action: instances [Add step](#) ⓘ

[Create a simple scaling policy](#) ⓘ

[Scale the Auto Scaling group using a target tracking scaling policy](#) ⓘ

[Cancel](#) [Previous](#)

Then a popup window for setting up the alarm will show up. Select a setting like this and click save.

Edit Alarm

You can use CloudWatch alarms to be notified automatically whenever metric data reaches a level you define.
To edit an alarm, first choose whom to notify and then define when the notification should be sent.

☐ Send a notification to: billingsms (+15404971048)

Whenever: Average of CPU Utilization

Is: >= 10 Percent

For at least: 1 consecutive period(s) of 5 Minutes

Name of alarm: awsec2-ssds-CPU-Utilization

Cancel Save

CPU Utilization Percent

asssds

Step 26: Repeat similar step for decreasing the group size and take an action

Increase Group Size

Name: Increase Group Size

Execute policy when: awsec2-ssds-CPU-Utilization [Edit](#) [Remove](#)
breaches the alarm threshold: CPUUtilization >= 10 for 300 seconds
for the metric dimensions AutoScalingGroupName = asssds

Take the action: Add 1 instances when 10 <= CPUUtilization < +infinity

Instances need: 300 seconds to warm up after each step

[Add step](#) [Create a simple scaling policy](#)

Decrease Group Size

Name: Decrease Group Size

Execute policy when: No alarm selected [Add new alarm](#)

Take the action: Remove 0 instances

[Add step](#) [Create a simple scaling policy](#)

[Scale the Auto Scaling group using a target tracking scaling policy](#)

1) Select the action

2) create similar alarm to decrease group size.

3) select the action(no of instances to remove)

Create Alarm

X

You can use CloudWatch alarms to be notified automatically whenever metric data reaches a level you define.

To edit an alarm, first choose whom to notify and then define when the notification should be sent.

☐ Send a notification to:

billingsms (+15404971048)

Whenever:

Average

of

CPU Utilization

Is:

<=

20

Percent

For at least:

1

consecutive period(s) of

5 Minutes

Name of alarm:

awsec2-ssds-High-CPU-Utilization

CPU Utilization Percent

Time	CPU Utilization Percent
7/11 20:00	20
7/11 22:00	20
7/12 00:00	20

assds

Cancel

Create Alarm

Step 27: Click Next Configure Instances→ Click Next again ---> Provide Tags(makes things easier to distinguish)

Click next until you create the auto scaling group.

Step 28: Go to the Load Balancer Page. Copy the DNS name and paste it in a browser.

You should see “Hello from VA web server 1”

Filter by tags and attributes or search by keyword

Name	DNS name	Status
myAutoscaleELB	myAutoscaleELB-1789673783.us-east-1.elb.amazonaws.com	

Load balancer: myAutoscaleELB

Description

Instances

Health Check

Listeners

Monitoring

Tags

Migration

Basic Configuration

Name:	myAutoscaleELB	Creation time:	June 1, 2018 at 1:56:49 PM UTC-4
* DNS name:	myAutoscaleELB-1789673783.us-east-1.elb.amazonaws.com (A Record)	Hosted zone:	Z35SXDOTRQ7X7K
		Status:	2 of 2 instances in service

Step 27: Terminate both EC2 instance and see what happens.

***they should come back up.

Step 28: Delete Autoscaling Group(Not launch config) to delete the Autoscaling group. This way the EC2 instances will be deleted forever and new ones won't be provisioned again.

Create Auto Scaling group

Actions

Filter Auto Scaling groups...

1 to 1 of 1 Auto Scaling Groups

Name	Launch Configuration /	Instances	Desired	Min	Max	Availability Zones
(Deleting) auto...	myAutoscale-elb	1	0	0	0	us-east-1a, us-east-1c