

Part 1: EC2 with ELB and ASG

Objective: Learn how to create a scalable and highly available web application environment using Amazon EC2 instances, ELB, and ASG.

Approach:

1. **Launch EC2 Instances:** Start by launching two or more EC2 instances. These instances will run a simple web application (e.g., a "Hello World" page or any basic web service).
2. **Configure Load Balancer:** Set up an Elastic Load Balancer (ELB) to distribute incoming web traffic across your EC2 instances. This step ensures high availability and fault tolerance.
3. **Set Up Auto Scaling Group (ASG):** Create an ASG that uses the launched EC2 instances. Configure ASG policies to automatically scale the number of instances up or down based on criteria like CPU usage or network traffic.
4. **Test Your Setup:** Simulate traffic to test the scaling policies and the load balancer. Observe how ASG adds or removes instances and how ELB distributes traffic.
5. **Verify Website Functionality:** Ensure that the website hosted on EC2 instances remains accessible and functional during scaling operations.

Goal: By the end of this lab, students will have a hands-on understanding of setting up a load-balanced and auto-scaled web application using AWS services.

Auto scaling and ELB

Creating AMI for the given EC2 instance:

The screenshot displays the AWS Management Console's 'Instances' page. At the top, there's a search bar and a filter for 'Instance state = running'. Below this, a table lists two instances: 'Web Server 1' (ID: i-0e9267e9c14525fef) and 'Bastion Host' (ID: i-09767724e3006555b). Both are 'Running' and using 't2.micro' instances. The 'Web Server 1' instance is selected, and the 'Actions' menu is open, showing options like 'Connect', 'View details', 'Manage instance state', 'Instance settings', 'Networking', 'Security', 'Image and templates', and 'Monitor and troubleshoot'. The 'Create image' option is highlighted.

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 ...
Web Server 1	i-0e9267e9c14525fef	Running	t2.micro	2/2 checks passed	View alarms +	us-east-1a	-	3.82.226.173
Bastion Host	i-09767724e3006555b	Running	t2.micro	2/2 checks passed	View alarms +	us-east-1a	-	3.90.11.209

Instance ID
i-0e9267e9c14525fef (Web Server 1)

Image name
WebServerAMI
Maximum 127 characters. Can't be modified after creation.

Image description - optional
Lab AMI for Web Server
Maximum 255 characters

No reboot
☐ Enable

Creating Load balancer:
First create a target group.

Remove all pending

Show only pending < 1 > ⚙

Port ▼	State ▼	Security groups ▼	Zone ▼	Private IPv4 address	Subnet ID ▼	Launch time ▲
No instances added yet y instances above, or leave the group empty if you prefer to add targets later.						

Cancel Previous Create target group

Now create load Balancer:

⌂ Actions ▼ Create load balancer ▼

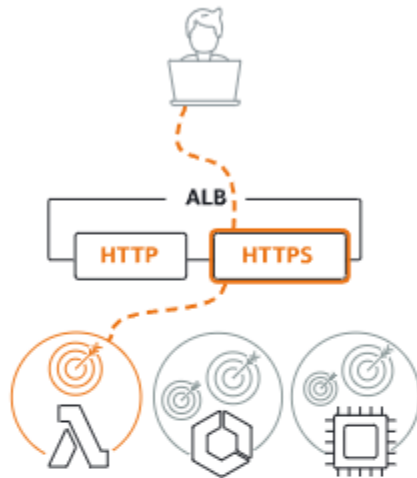
< 1 > ⚙

▼	Type ▼	Date created ▼
---	--------	----------------

st-1

Create Application Load balancer:

Application Load Balancer [Info](#)



Choose an Application Load Balancer when you need a flexible feature set for your applications with HTTP and HTTPS traffic. Operating at the request level, Application Load Balancers provide advanced routing and visibility features targeted at application architectures, including microservices and containers.

Create

Mappings [Info](#)

Select at least two Availability Zones and one subnet per zone. The load balancer routes traffic to targets in these Availability Zones only. Availability Zones that are not supported by the balancer or the VPC are not available for selection.

☒ us-east-1a (use1-az2)

Subnet

subnet-0595a05bc29148d7d

Public Subnet 1 ▼

IPv4 address

Assigned by AWS

☒ us-east-1b (use1-az4)

Subnet

subnet-08358b7cb592559e8

Public Subnet 2 ▼

IPv4 address

Assigned by AWS

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



Web Security Group

sg-0d157249119e0d417 VPC: vpc-07888501c4117a3da



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](#)

Forward to

LabGroup

Target type: Instance, IPv4

HTTP ▼



[Create target group](#)

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

Load Balancer created:

Load balancers (1)

Elastic Load Balancing scales your load balancer capacity automatically in response to changes in incoming traffic.

Filter load balancers

1 match

LabELB

Clear filters

<

1

>

<input type="checkbox"/>	Name	DNS name	State	VPC ID	Availability Zones	Type	Date created
<input type="checkbox"/>	LabELB	LabELB-261549413.us-eas...	Provisioning..	vpc-07888501c4117a...	2 Availability Zones	application	January 20, 2024, 20:1...

0 load balancers selected

=

X

Create a Launch Template and an Auto Scaling Group

For new and existing customers, we
ers better performance at a lower cost.

to the internet.

Cancel

Create launch template

Launch Templates (1/1) Info

Search

< 1 >

Launch Template ID	Launch Template Name	Default Version	Latest Version	Create Time	Created By
lt-084279ea5a1703bc8	LabConfig	1	1	2024-01-20T14:31:08.000Z	arn:aws:sts::367735330430:assumed-...

Create Auto Scaling Group

Launch Templates (1/1) Info

Q Search

Launch Template ID	Launch Template Name	Default Version	Latest Version	Create Time
lt-084279ea5a1703bc8	LabConfig	1	1	2024-01-20T14:31:08.0

Actions

Create launch temp

Launch instance from template

Modify template (Create new version)

Delete template

Delete template version

Set default version

Manage tags

Create Spot Fleet

Create Auto Scaling group

View details

Create auto scaling group:

Auto Scaling groups (1) Info

Q Search your Auto Scaling groups

< 1 >

Launch configurations

Launch templates

Actions

Create Auto Scaling group

Name	Launch template/configuration	Instances	Status	Desired capacity	Min	Max	Availability Zones
Lab Auto Scaling Group	LabConfig Version Default	0	Updating capacity...	2	2	6	us-east-1a, us-east-1b

Auto scaling group create 2 desire instance:

Q Find Instance by attribute or tag (case-sensitive)

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone
Lab Instance	i-05e4a2c31ff7bd050	Running	t2.micro	Initializing	View alarms +	us-east-1a
Web Server 1	i-0e9267e9c14525fef	Running	t2.micro	2/2 checks passed	View alarms +	us-east-1a
Bastion Host	i-09767724e3006555b	Running	t2.micro	2/2 checks passed	View alarms +	us-east-1a
Lab Instance	i-00ed7a5805fafac81	Running	t2.micro	Initializing	View alarms +	us-east-1b

Testing load balancer:

Not secure | labelb-261549413.us-east-1.elb.amazonaws.com

7eehrrv2arxgrhnc.p... Savittri_2021_J_Phy... word search in doc...

aws Load Test RDS

Meta-Data	Value
Instanceid	i-00ed7a5805fafac81
Availability Zone	us-east-1b

Current CPU Load: 0%

Not secure | labelb-261549413.us-east-1.elb.amazonaws.com

7eehrrv2arxgrhnc.p... Savittri_2021_J_Phy... word search in doc...

aws Load Test RDS

Meta-Data	Value
Instanceid	i-05e4a2c31ff7bd050
Availability Zone	us-east-1a

Current CPU Load: 0%

note : 2 different instance is returned