ASSESSMENT ON:
AUTO SCALING
AND
LOAD BALANCER

TO THE NEW

1.Differences between ELB, ALB, and NLB. Where will you use which one?

ANS:

	Application Load Balancer	**Network Load Balancer**	**Classic Load Balancer**
Quick Info	Application Load Balancer is one of the newer types of load balancers. AWS recommends that you use this instead of the classic load balancer. This ELB operates at OSI Layer 7 (HTTP)	Network Load Balancer is one of the newer types of load balancers. AWS recommends that you use this instead of the classic load balancer. This ELB operates at OSI Layer 4 (TCP)	The classic load balancer is the older type of load balancer in AWS. AWS recommends that you use ALB or NLB instead for your load balancing requirements. Operates at either Layer 4 or Layer 7
Use Case	If you need flexible application management, use an Application Load Balancer.	If extreme performance and static IP is needed for your application, use a Network Load Balancer.	If you have an existing application that was built within the EC2-Classic network, then use a Classic Load Balancer.
Supported Protocols	HTTP and HTTPS	TCP, UDP and TLS	TCP, SSL/TLS, HTTP, HTTPS
Supported Platform	VPC	VPC	VPC and EC2-Classic (deprecated)
Can load balance to multiple ports on the same instance	Yes (Useful for containerized applications)	Yes	No

Unique features	• You can set Lambda functions as load balancing targets
	• Only ALB supports the following content-based routing methods:
	  • Path based routing
	  • Host-based routing
	  • HTTP header-based routing
	  • HTTP method-based routing
	  • Query string parameter-based routing
	  • Source IP address CIDR-based routing
	• Natively supports HTTP/2, IPv6
	• Support for multiple SSL certificates on the ALB using
	• Server Name Indication (SNI)
	• Allows tag-based IAM permission policies
	• Can be configured for slow start (linearly increases the number of requests sent to targets)
	• Supports round-robin load balancing
	• You can offload the authentication functionality from your apps into ALB
	• Can redirect an incoming request from one URL to another URL, including HTTP to HTTPS
	• You can set HTTP or custom responses for incoming requests to the ALB, offloading this task from your application
		• High throughput/low latency ELB
		• Can be assigned a static IP address
		• Can be assigned an elastic IP address
		• Preserves source IP address of non-HTTP applications on EC2 instances
		• Offer multi-protocol listeners, allowing you to run applications such as DNS that rely on both TCP and UDP protocols on the same port behind a Network Load Balancer.
		• TLS Termination
			• You can create custom security policies detailing which ciphers and protocols are supported by the ELB
			• Supports both IPv4 and IPv6 for EC2-Classic network

****

2.Differences between step scaling and target scaling.

ANS:

With step scaling and simple scaling, you choose scaling metrics and threshold values for the CloudWatch alarms that trigger the scaling process. You also define how your Auto Scaling group should be scaled when a threshold is in breach for a specified number of evaluation periods.

AWS recommends that you use a target tracking scaling policy to scale on a metric like average CPU utilization or the RequestCountPerTarget metric from the Application Load Balancer. Metrics that decrease when capacity increases and increase when capacity decreases can be used to proportionally scale out or in the number of instances using target tracking. This helps ensure that Amazon EC2 Auto Scaling follows the demand curve for your applications closely. With target tracking scaling policies, you select a scaling metric and set a target value. Amazon EC2 Auto Scaling creates and manages the CloudWatch alarms that trigger the scaling policy and calculates the scaling adjustment based on the metric and the target value. The scaling policy adds or removes capacity as required to keep the metric at, or close to, the specified target value. In addition to keeping the metric close to the target value, a target tracking scaling policy also adjusts to the changes in the metric due to a changing load pattern.

3.Differences between Launch configuration and launch template.

ANS:

A *launch configuration* is an instance configuration template that an Auto Scaling group uses to launch EC2 instances. When you create a launch configuration, you specify information for the instances. Include the ID of the Amazon Machine Image (AMI), the instance type, a key pair, one or more security groups, and a block device mapping.

A launch template is similar to a launch configuration, in that it specifies instance configuration information. Included are the ID of the Amazon Machine Image (AMI), the instance type, a key pair, security groups, and the other parameters that you use to launch EC2 instances.

*****However, defining a launch template instead of a launch configuration allows you to have multiple versions of a template. With versioning, you can create a subset of the full set of parameters and then reuse it to create other templates or template versions. For example, you can create a default template that defines common configuration parameters such as tags or network configurations, and allow the other parameters to be specified as part of another version of the same template.

****

4.Differences between EC2 health check and load balancer health check

ANS:

EC2 health check watches for instance availability from hypervisor and networking point of view. For example, in case of a hardware problem, the check will fail. Also, if an instance was misconfigured and doesn't respond to network requests, it will be marked as faulty.

ELB health check verifies that a specified TCP port on an instance is accepting connections OR a specified web page returns 2xx code. Thus ELB health checks are a little bit smarter and verify that actual app works instead of verifying that just an instance works.

EC2 instance health check

- Amazon EC2 performs automated checks on every running EC2 instance to identify hardware and software issues.
- Status checks are performed every minute and each returns a pass or a fail status.
  - If all checks pass, the overall status of the instance is OK.
  - If one or more checks fail, the overall status is impaired.
- Status checks are built into EC2, so they cannot be disabled or deleted.
- You can create or delete alarms that are triggered based on the result of the status checks.
- There are two types of status checks

  System Status Checks
  - These checks detect underlying problems with your instance that require AWS involvement to repair. When a system status check fails, you can choose to wait for AWS to fix the issue, or you can resolve it yourself.

  Instance Status Checks
  - Monitor the software and network configuration of your individual instance. Amazon EC2 checks the health of an instance by sending an address resolution protocol (ARP) request to the ENI. These checks detect problems that require your involvement to repair.

Tutorials Dojo

Elastic Load Balancer (ELB) health check

- To discover the availability of your registered EC2 instances, a load balancer periodically sends pings, attempts connections, or sends requests to test the EC2 instances.
- The status of the instances that are healthy at the time of the health check is InService. The status of any instances that are unhealthy at the time of the health check is OutOfService.
- When configuring a health check, you would need to provide the following:
  - a specific port
  - protocol to use
    - HTTP/HTTPS health check succeeds if the instance returns a 200 response code within the health check interval.
    - A TCP health check succeeds if the TCP connection succeeds.
    - An SSL health check succeeds if the SSL handshake succeeds.
  - ping path
- ELB health checks do not support WebSockets.
- The load balancer routes requests only to the healthy instances. When an instance becomes impaired, the load balancer resumes routing requests to the instance only when it has been restored to a healthy state.
- The load balancer checks the health of the registered instances using either
  - the default health check configuration provided by Elastic Load Balancing or
  - a health check configuration that you configure (auto scaling or custom health checks for example).
- Network Load Balancers use active and passive health checks to determine whether a target is available to handle requests.
  - With active health checks, the load balancer periodically sends a request to each registered target to check its status. After each health check is completed, the load balancer node closes the connection that was established.
  - With passive health checks, the load balancer observes how targets respond to connections, which enables it to detect an unhealthy target before it is reported as unhealthy by active health checks. You cannot disable, configure, or monitor passive health checks.

Auto Scaling and Custom health checks

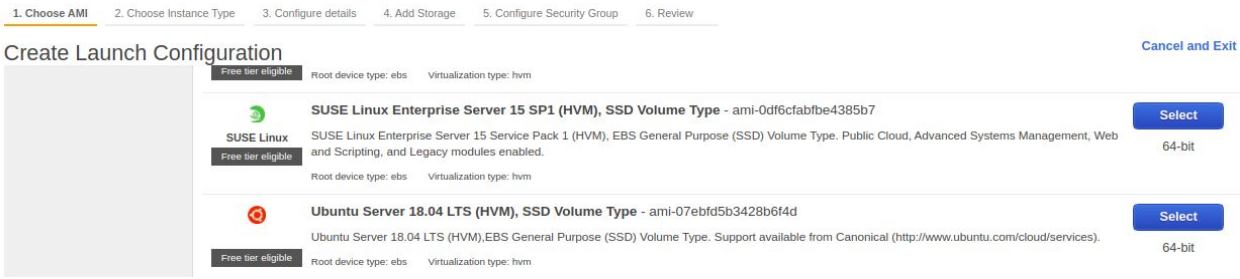
- All instances in your Auto Scaling group start in the healthy state. Instances are assumed to be healthy unless EC2 Auto Scaling receives notification that they are unhealthy. This notification can come from one or more of the following sources:
  - Amazon EC2 (default)
  - Elastic Load Balancing
  - A custom health check.
- After Amazon EC2 Auto Scaling marks an instance as unhealthy, it is scheduled for replacement. If you do not want instances to be replaced, you can suspend the health check process for any individual Auto Scaling group.
- If an instance is in any state other than running or if the system status is impaired, Amazon EC2 Auto Scaling considers the instance to be unhealthy and launches a replacement instance.
- If you attached a load balancer or target group to your Auto Scaling group, Amazon EC2 Auto Scaling determines the health status of the instances by checking both the EC2 status checks and the Elastic Load Balancing health checks.
- Amazon EC2 Auto Scaling waits until the health check grace period ends before checking the health status of the instance. Ensure that the health check grace period covers the expected startup time for your application.
- Health check grace period does not start until lifecycle hook actions are completed and the instance enters the InService state.
- With custom health checks, you can send an instance's health information directly from your system to Amazon EC2 Auto Scaling.

****

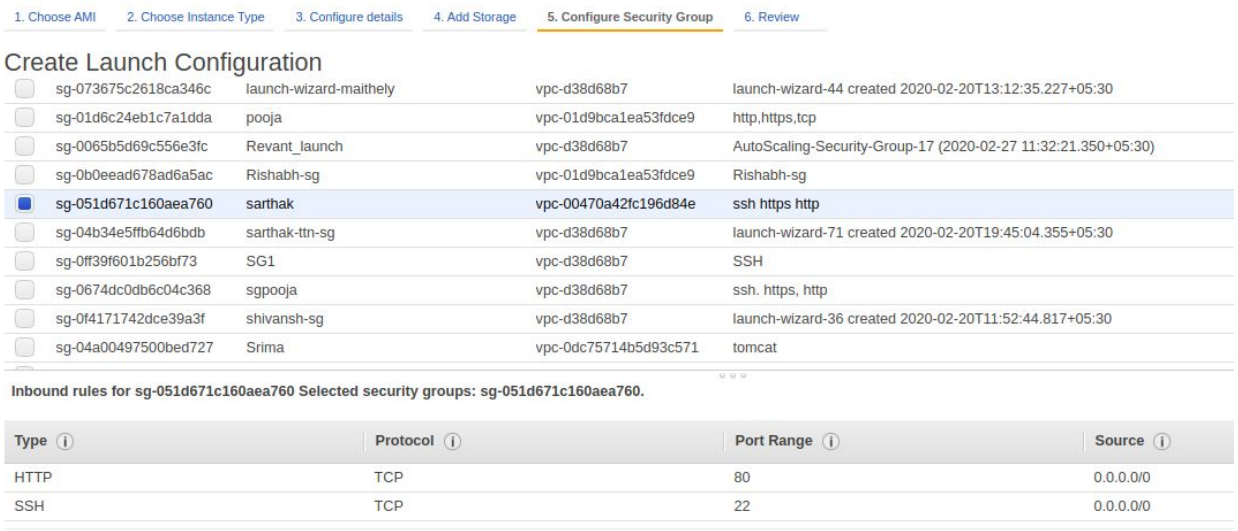
5.Create 2 auto-scaling groups with

- **launch configuration and**

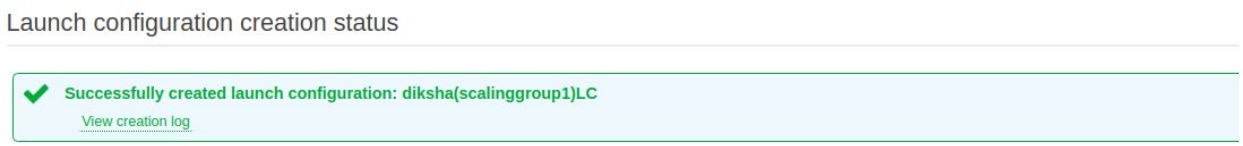
STEP 1: EC2 > Launch Configuration



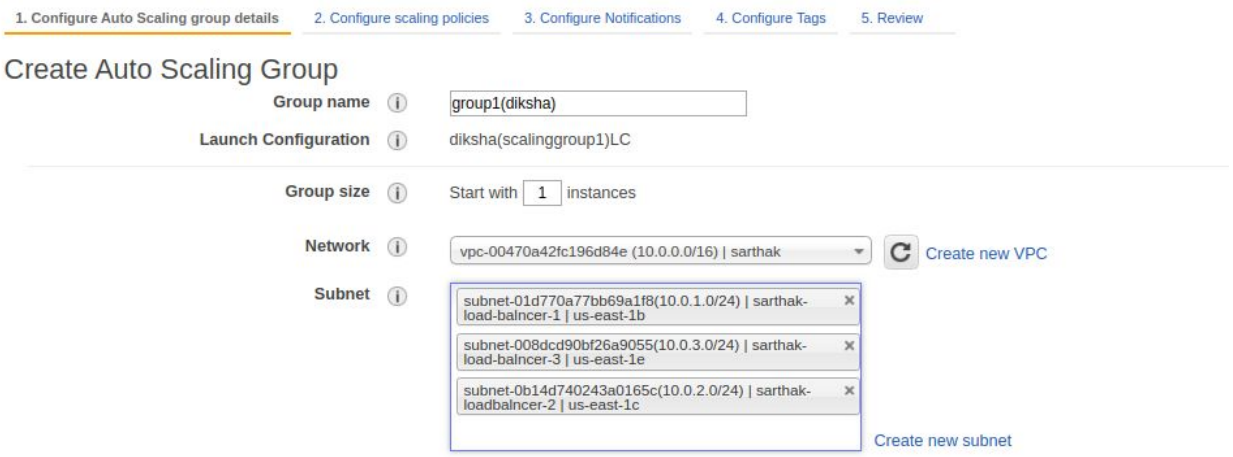
STEP 2: Provide security group



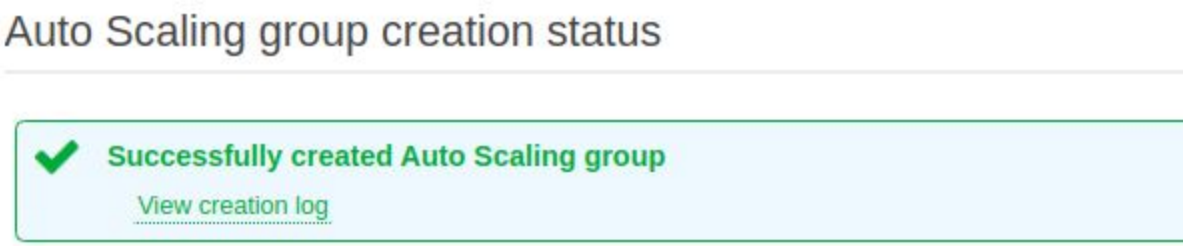
STEP 3: Launch configuration is being created



STEP 4: Select your VPC and all the subnets



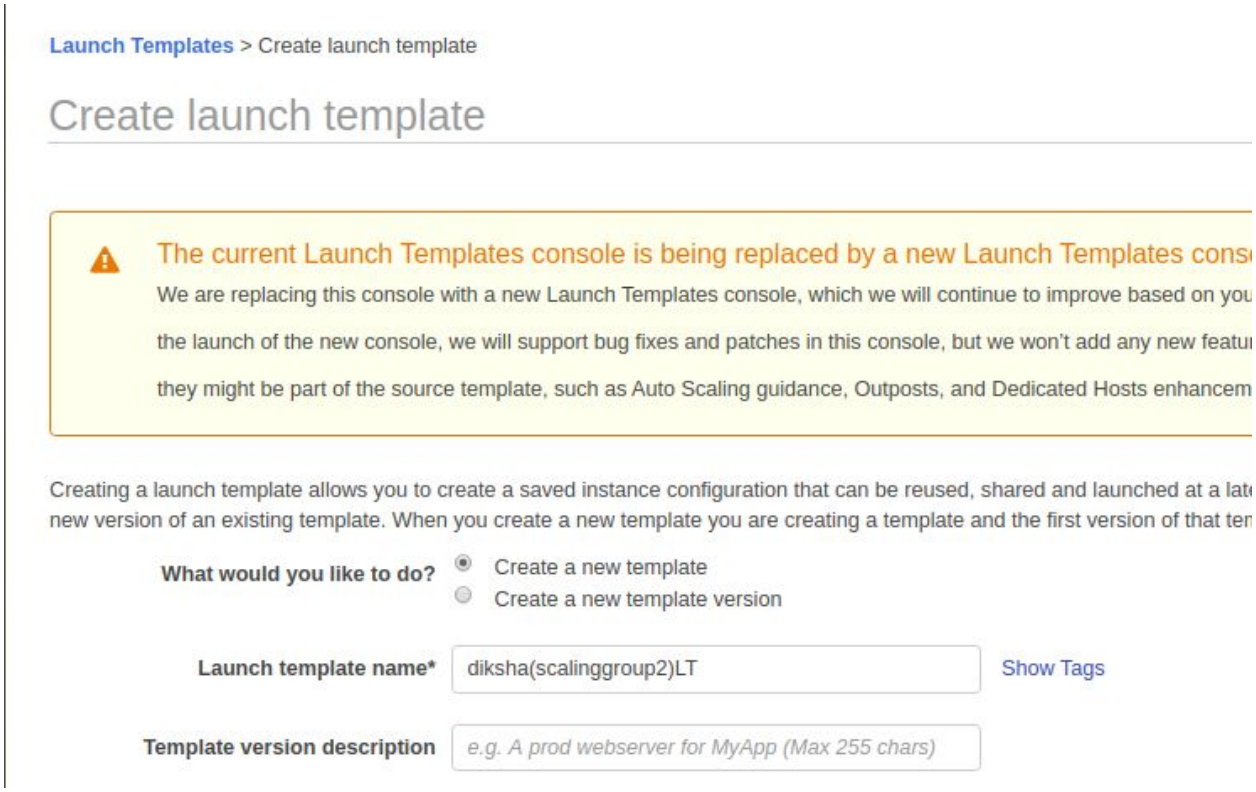
STEP 5: Your Auto scaling group has been created



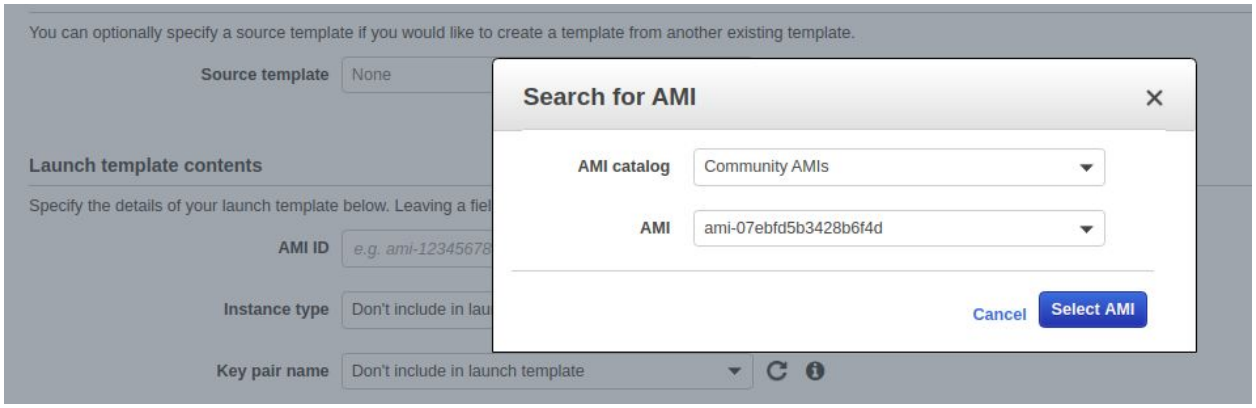
- **launch template**

STEP 1: EC2 > Launch Template

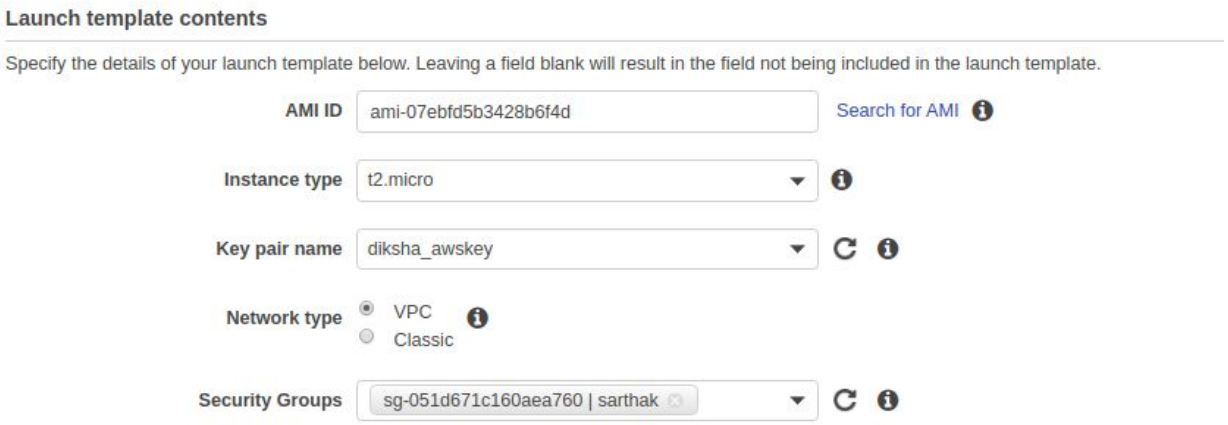
STEP 2:



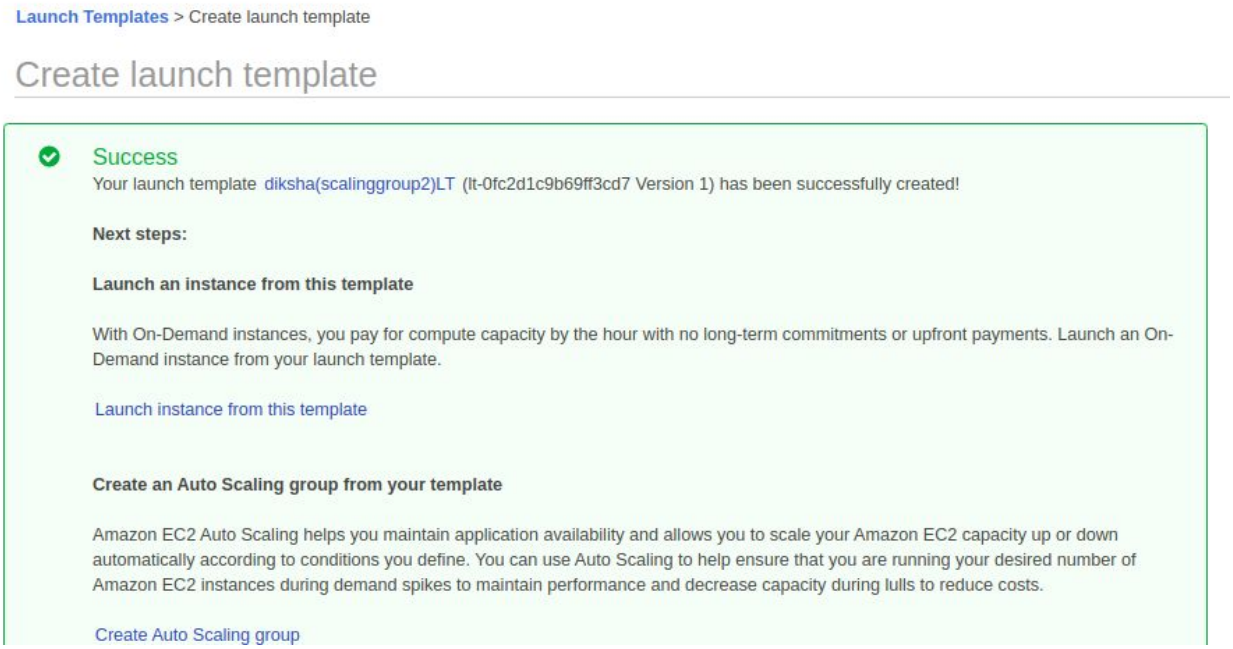
STEP 3: Provide AMI



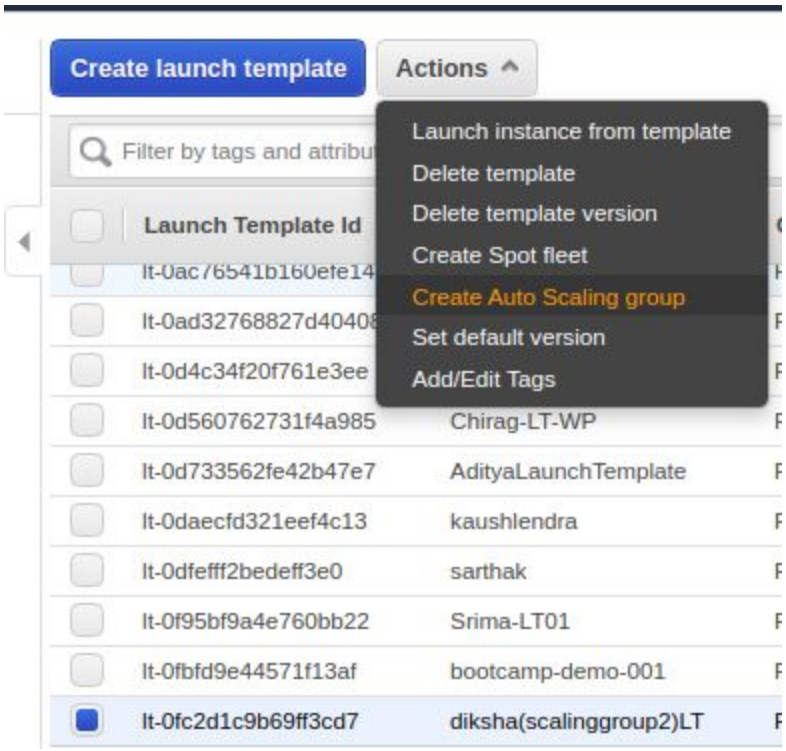
STEP 4: Select your security group

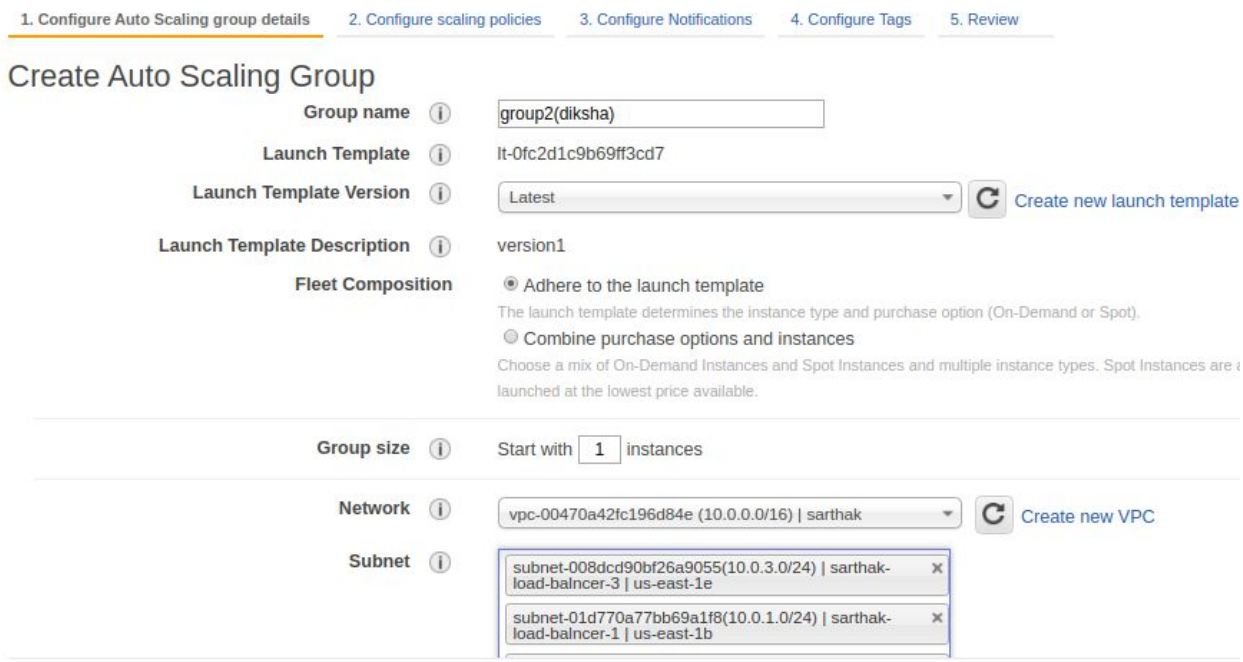


STEP 5: Your template has been created

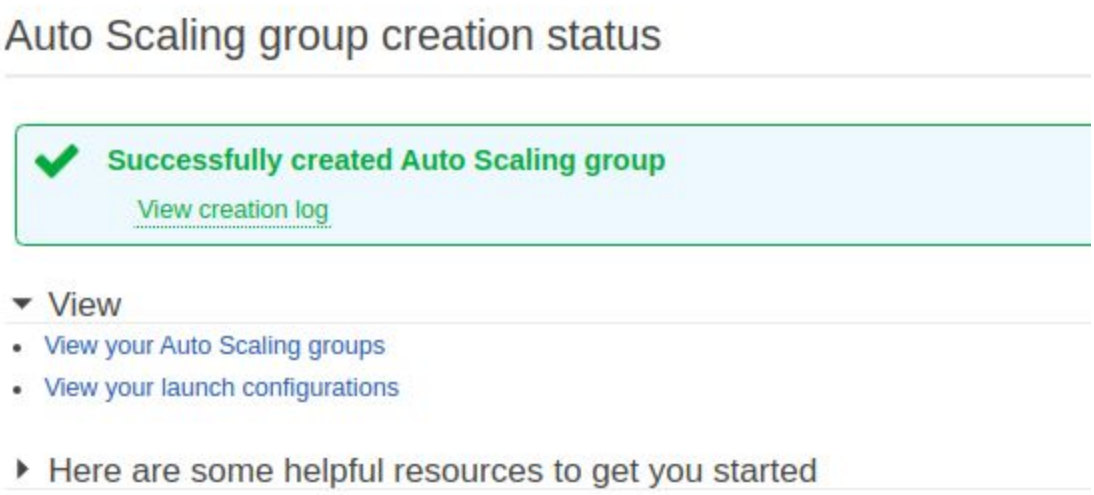


STEP 6: Now create auto scaling group from the template created





STEP 7: Your auto scaling group has been created



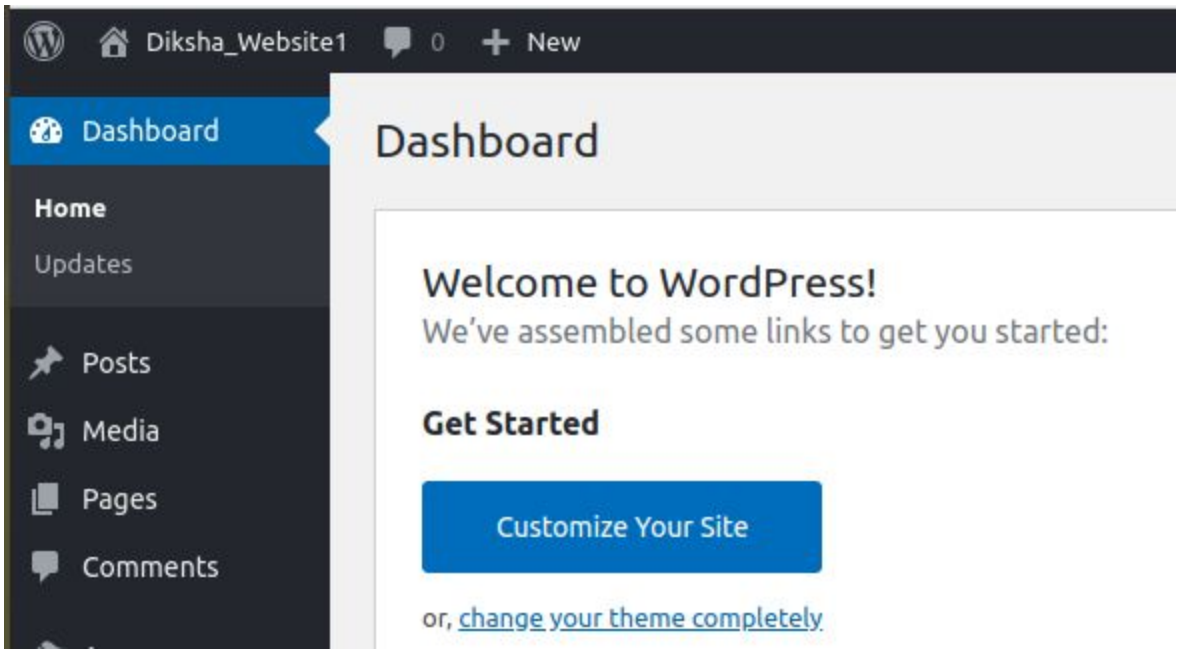
****

6.Setup autoscaling Wordpress application with the Application load balancer. Auto-scaling should be triggered based on CPU usage of EC2 instances.

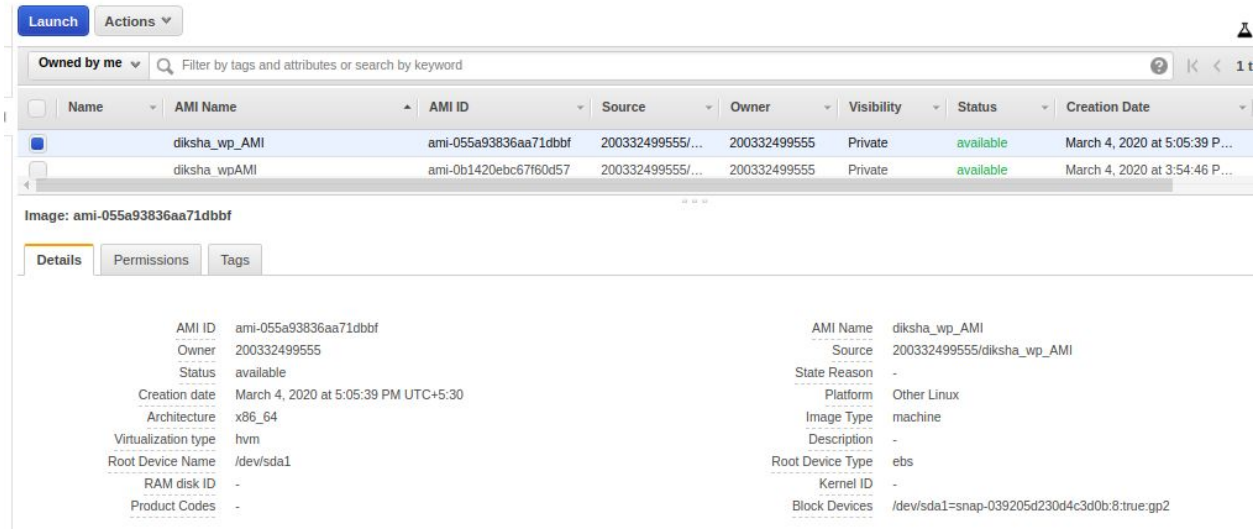
ANS:

Prerequisite (VPC >Subnet >IGW >Associate your subnet > Make route Table for your VPC and make entry for your IGW)

STEP 1: Launch an instance with wordpress installed “diksha_wp”



STEP 2: Create launch Template with the AMI of above instance.



STEP 3: Create a target group

Create target group | Actions

Filter by tags and attributes or search by keyword

Name	Port	Protocol	Target type	Load Balan	VPC ID	Monitoring
dikshaWpTG	80	HTTP	instance		vpc-09921cb2d9e87f2eb	

Target group: dikshaWpTG

Description | Targets | Health checks | Monitoring | Tags

Basic Configuration

Name	dikshaWpTG
ARN	arn:aws:elasticloadbalancing:us-east-1:200332499555:targetgroup/dikshaWpTG/2f405135ab8b5d99
Protocol	HTTP
Port	80
Target type	instance
VPC	vpc-09921cb2d9e87f2eb
Load balancer	

Register and deregister targets

Registered targets

To deregister instances, select one or more registered instances and then click Remove.

Remove

Instance	Name	Port	State	Security groups	Zone
No instances available.					

Instances

To register additional instances, select one or more running instances, specify a port, and then click Add. The default port is the port specified for the target group. If the instance is already registered on the specified port, you must specify a different port.

Add to registered on port 80

Search Instances

Instance	Name	State	Security groups	Zone	Subnet ID	Subnet CIDR
i-0c2be1129e02ac...	diksha_wp	running	launch-wizard-3	us-east-1a	subnet-066f3806db4573748	10.0.1.0/24
i-015828677e4f4ff85		running	launch-wizard-3	us-east-1a	subnet-066f3806db4573748	10.0.1.0/24

STEP 4: Create Auto Scaling Group

1. Configure Auto Scaling group details | 2. Configure scaling policies | 3. Configure Notifications | 4. Configure Tags | 5. Review

Create Auto Scaling Group

Launch Template Description -

Fleet Composition
- (•) Adhere to the launch template
  The launch template determines the instance type and purchase option (On-Demand or Spot).
- ( ) Combine purchase options and instances
  Choose a mix of On-Demand Instances and Spot Instances and multiple instance types. Spot Instances are automatically launched at the lowest price available.

Group size Start with 1 instances

Network vpc-00470a42fc196d84e (10.0.0.0/16) | sarthak — Create new VPC

Subnet
- subnet-0c3f9366d2e1133a8(10.0.11.0/24) | diksha_lb_wp | us-east-1f
- subnet-008dcd90bf26a9055(10.0.3.0/24) | sarthak-load-balncer-3 | us-east-1e

Create new subnet

▸ Advanced Details

Launch Template wordpressLT

Launch Template Version 1 (Default)

Launch Template Description -

Fleet Composition
- (•) Adhere to the launch template
  The launch template determines the instance type and purchase option (On-Demand or Spot).
- ( ) Combine purchase options and instances
  Choose a mix of On-Demand Instances and Spot Instances and multiple instance types. Spot Instances are automatically launched at the lowest price available.

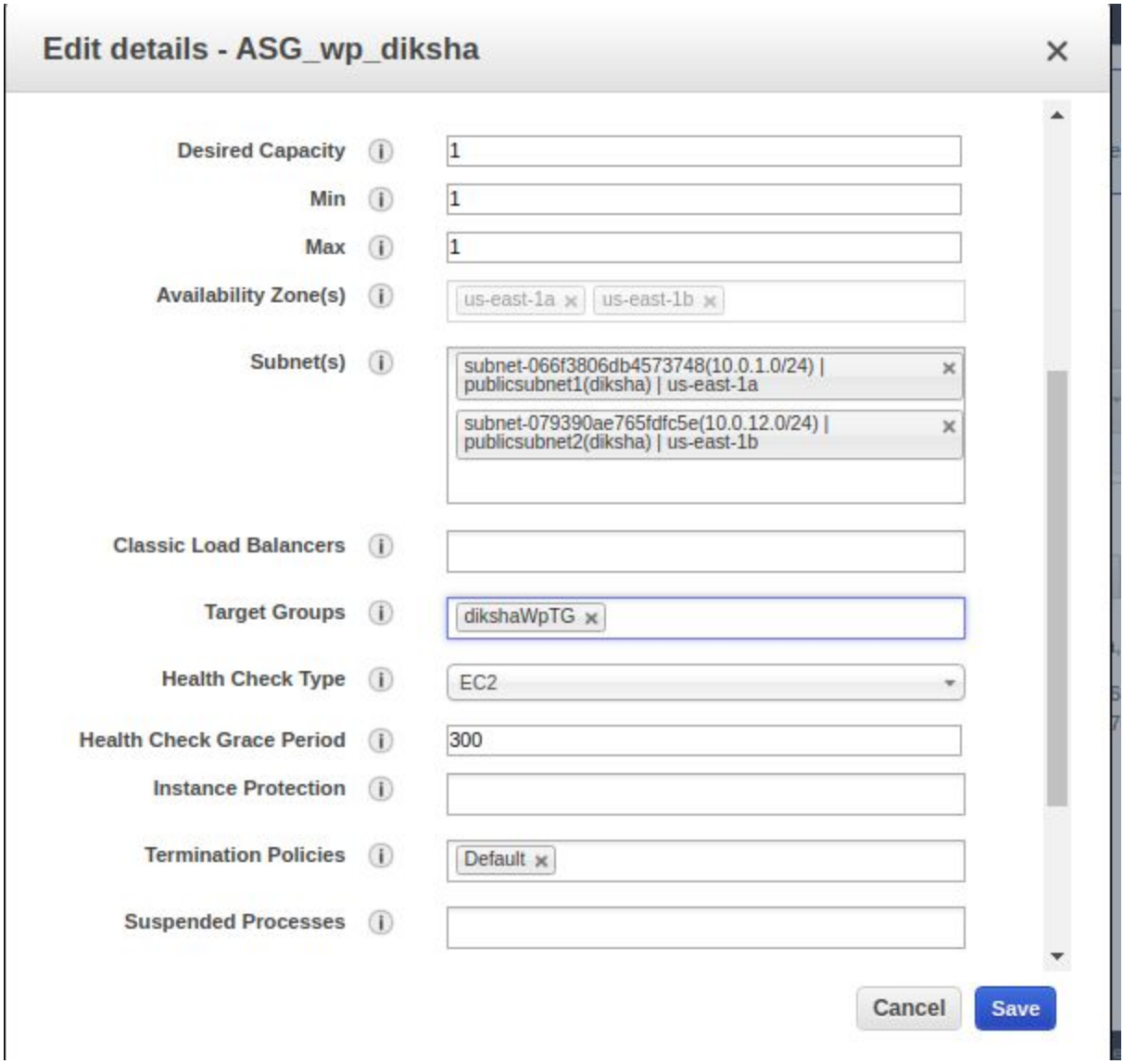
Desired Capacity 1

Min 1

Max 1

Availability Zone(s) us-east-1a, us-east-1b

Subnet(s)
- subnet-066f3806db4573748(10.0.1.0/24) | publicsubnet1(diksha) | us-east-1a
- subnet-079390ae765fdfc5e(10.0.12.0/24) | publicsubnet2(diksha) | us-east-1b



STEP 5: Create an Application Load balancer

Select load balancer type

Elastic Load Balancing supports three types of load balancers: Application Load Balance
for you

Application Load Balancer

HTTP
HTTPS

Create

Choose an Application Load Balancer when you need a flexible feature set for your web 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.

Learn more >

1. Configure Load Balancer | 2. Configure Security Settings | 3. Configure Security Groups | 4. Configure Routing | 5. Register Targets | 6. Review

Step 1: Configure Load Balancer

Basic Configuration

To configure your load balancer, provide a name, select a scheme, specify one or more listeners, and select a network. The default configuration is an Internet-facing load ba
80.

Name dikshaLBwordpress

Scheme internet-facing
internal

IP address type ipv4

Listeners

A listener is a process that checks for connection requests, using the protocol and port that you configured.

Load Balancer Protocol	Load Balancer Port
HTTP	80

Add listener

Availability Zones

Specify the Availability Zones to enable for your load balancer. The load balancer routes traffic to the targets in these Availability Zones o
Zones to increase the availability of your load balancer.

VPC vpc-09921cb2d9e87f2eb (10.0.0.0/16) | dikshaVPC

Availability Zones

- [x] **us-east-1a** subnet-066f3806db4573748 (publicsubnet1(diksha))
  - **IPv4 address** Assigned by AWS
- [x] **us-east-1b** subnet-079390ae765fdfc5e (publicsubnet2(diksha))
  - **IPv4 address** Assigned by AWS

1. Configure Load Balancer | 2. Configure Security Settings | **3. Configure Security Groups** | 4. Configure Routing | 5. Register Targets | 6. Review

Step 3: Configure Security Groups

A security group is a set of firewall rules that control the traffic to your load balancer. On this page, you can add rules to allow specific traffic to reach your load balancer. First,

Assign a security group:
- ○ Create a **new** security group
- ◉ Select an **existing** security group

	Security Group ID	Name	Description
☐	sg-02728bf96d082a4fa	default	default VPC security group
☑	sg-0ecfdca3b5cd4885b	launch-wizard-3	launch-wizard-3 created 2020-03-04T15:48:13.105+05:30

1. Configure Load Balancer | 2. Configure Security Settings | 3. Configure Security Groups | 4. Configure Routing | 5. Register Targets | 6. Review

Step 4: Configure Routing

Your load balancer routes requests to the targets in this target group using the protocol and port that you specify, and performs health checks on the targets using balancer.

Target group

Target group	Existing target group
Name	dikshaWpTG
Target type	Instance IP Lambda function
Protocol	HTTP
Port	80

Health checks

Protocol	HTTP
Path	/

Advanced health check settings

Step 5: Register Targets

Register targets with your target group. If you register a target in an enabled Availability Zone, the load balancer starts routing requests to the targe

Registered targets

The following targets are registered with the target group that you selected. You can only modify this list after you create the load balancer.

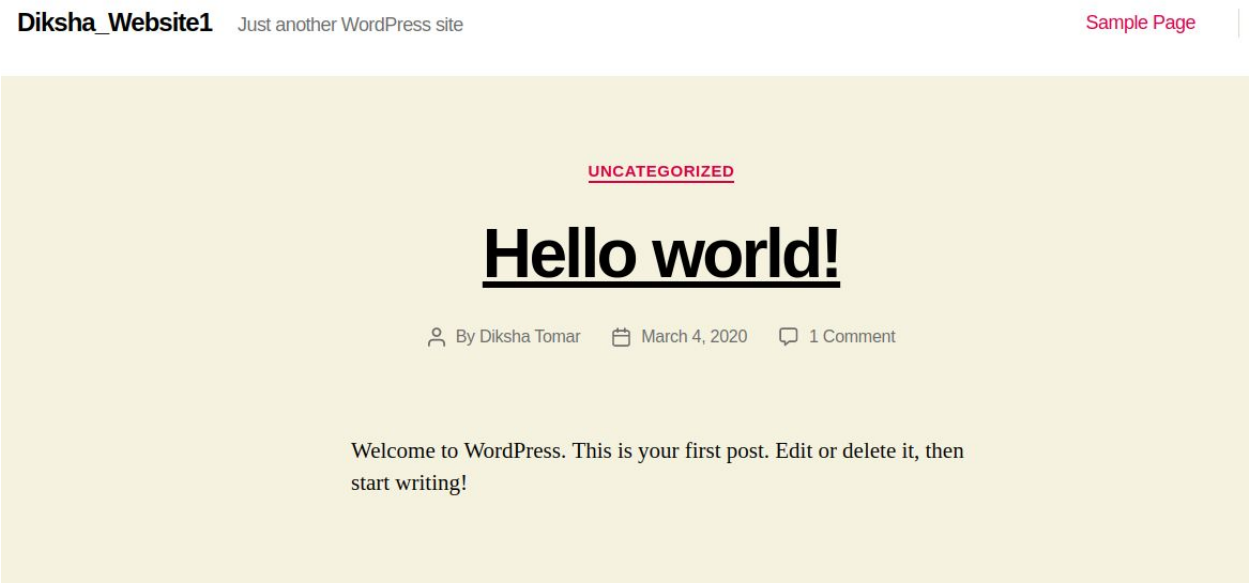
Instance	Port
i-015828677e4f4ff85	80

Create Load Balancer | Actions

Filter by tags and attributes or search by keyword

Name	DNS name	State	VPC ID	Availability Zones	Type
dikshaLBwordpress	dikshaLBwordpress-175355...	provisioning	vpc-09921cb2d9e87f2eb	us-east-1a, us-east-1b	application

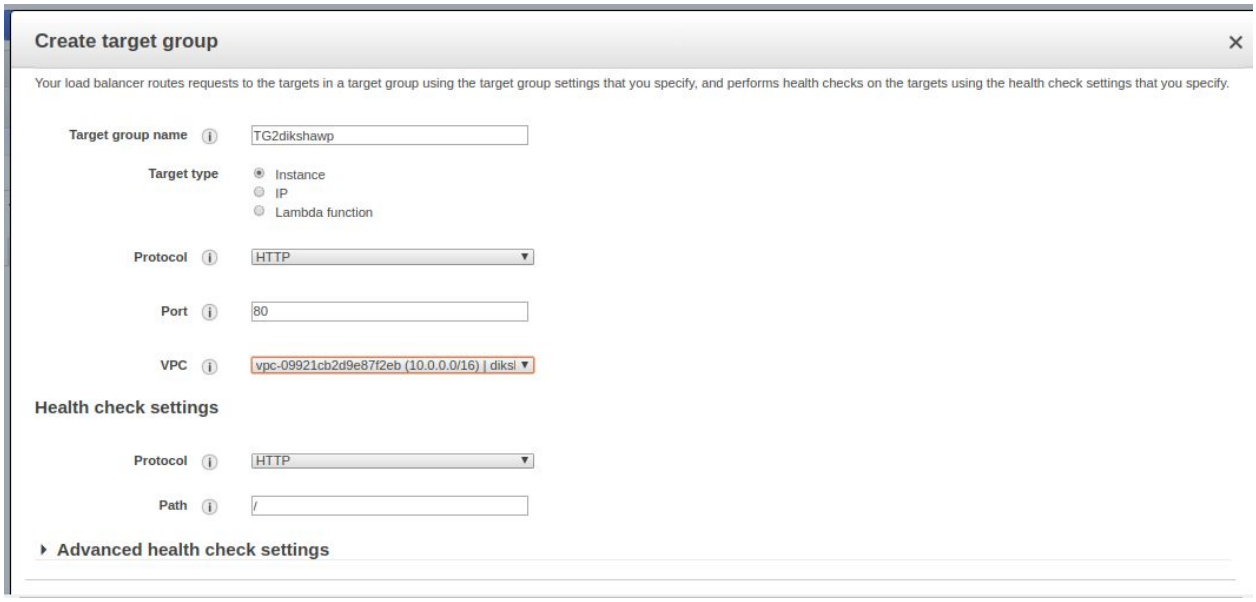
STEP 6: Copy public ip of ALB and check that wordpress page shows.

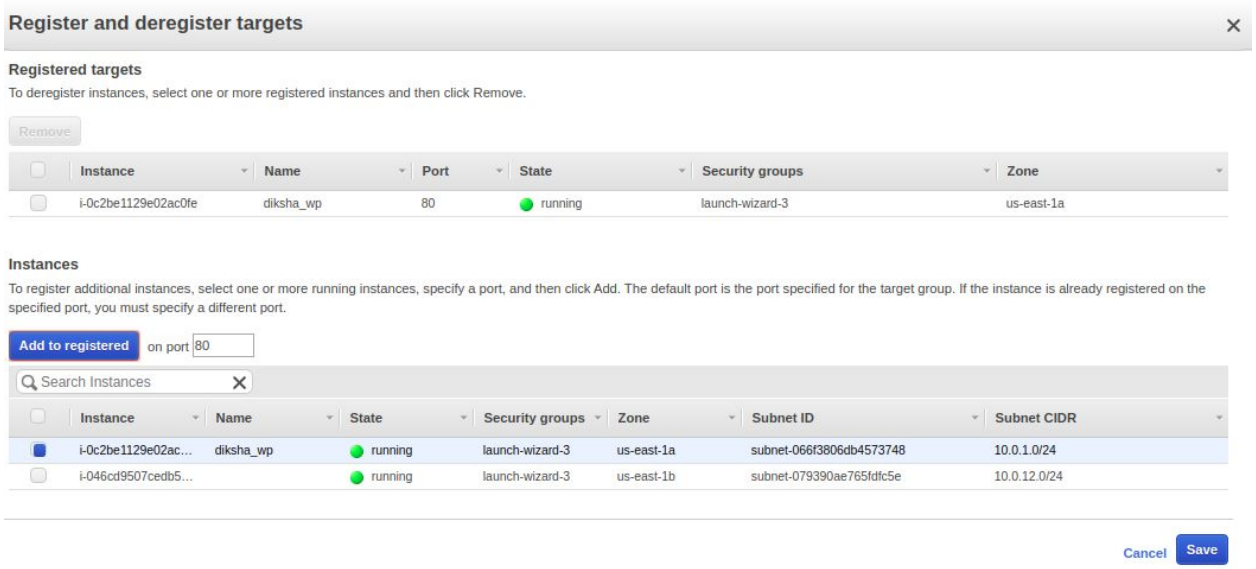


****

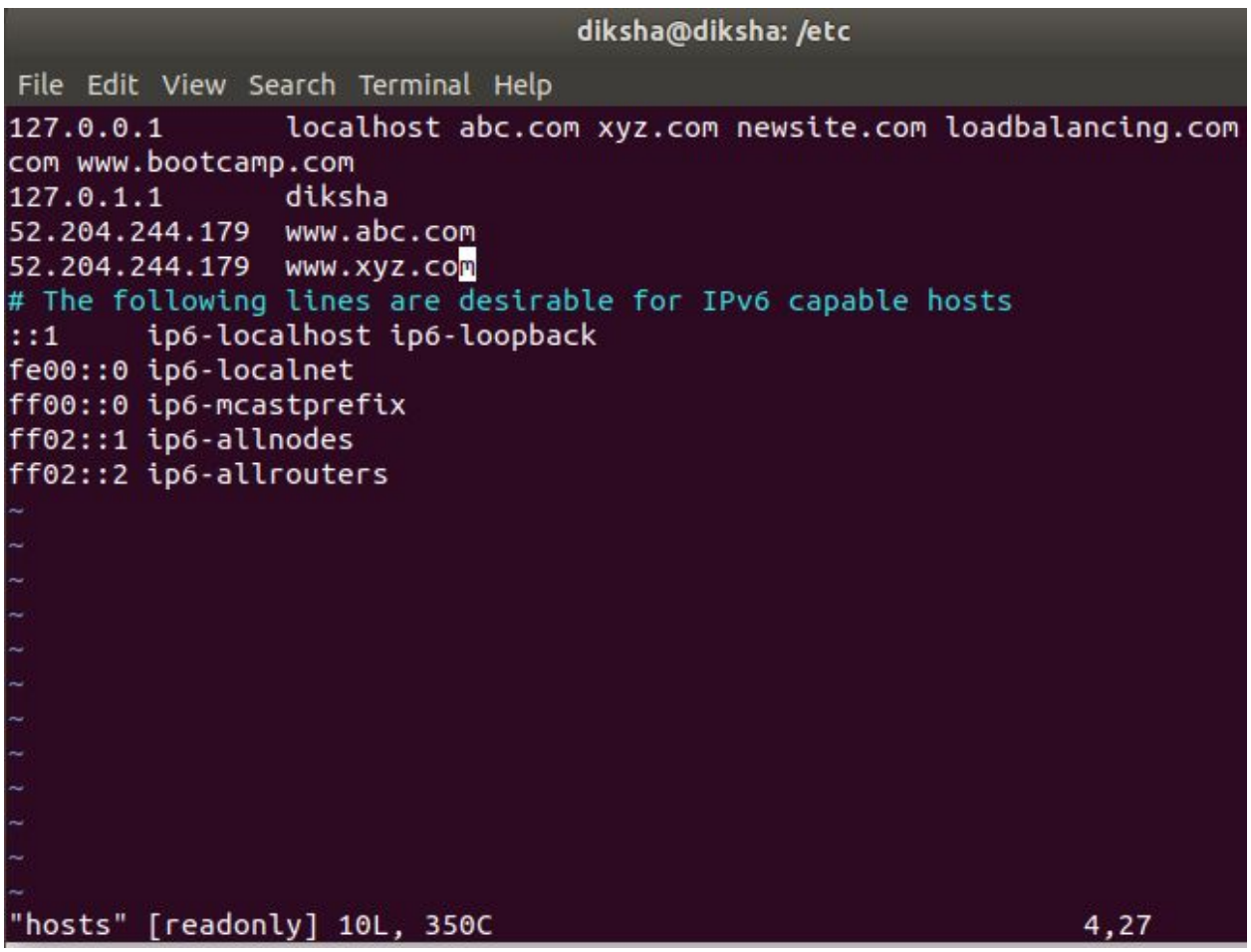
7.Create another Wordpress website and use the ALB created above to send traffic to this website based on the hostname

STEP 1: Register the instance with wordpress to a target group.

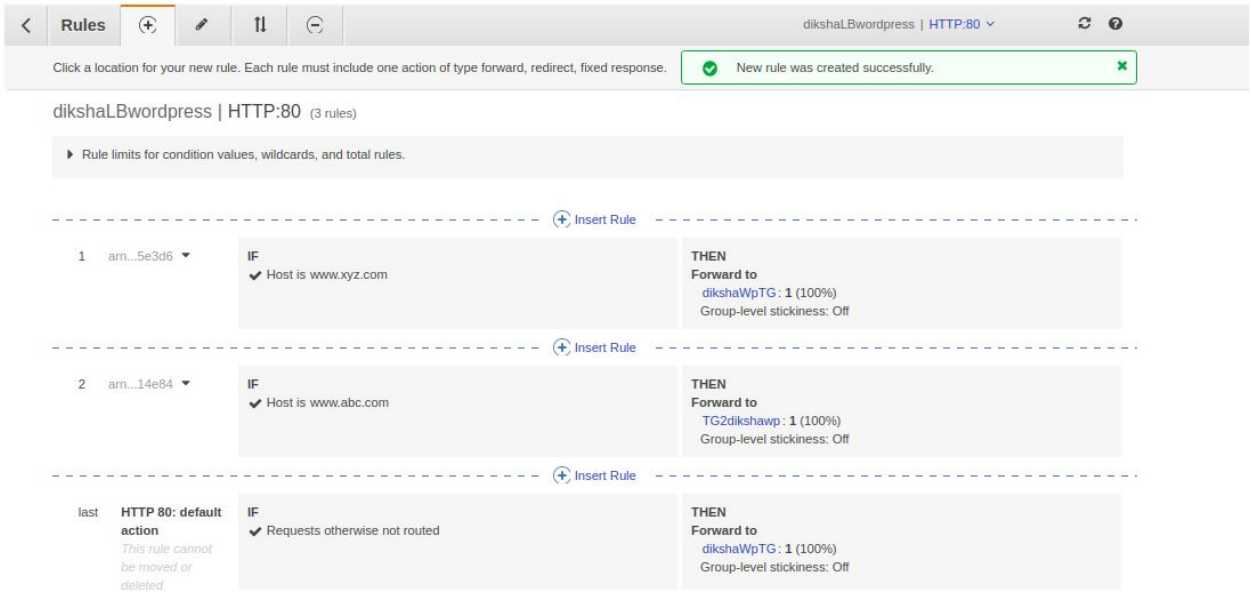




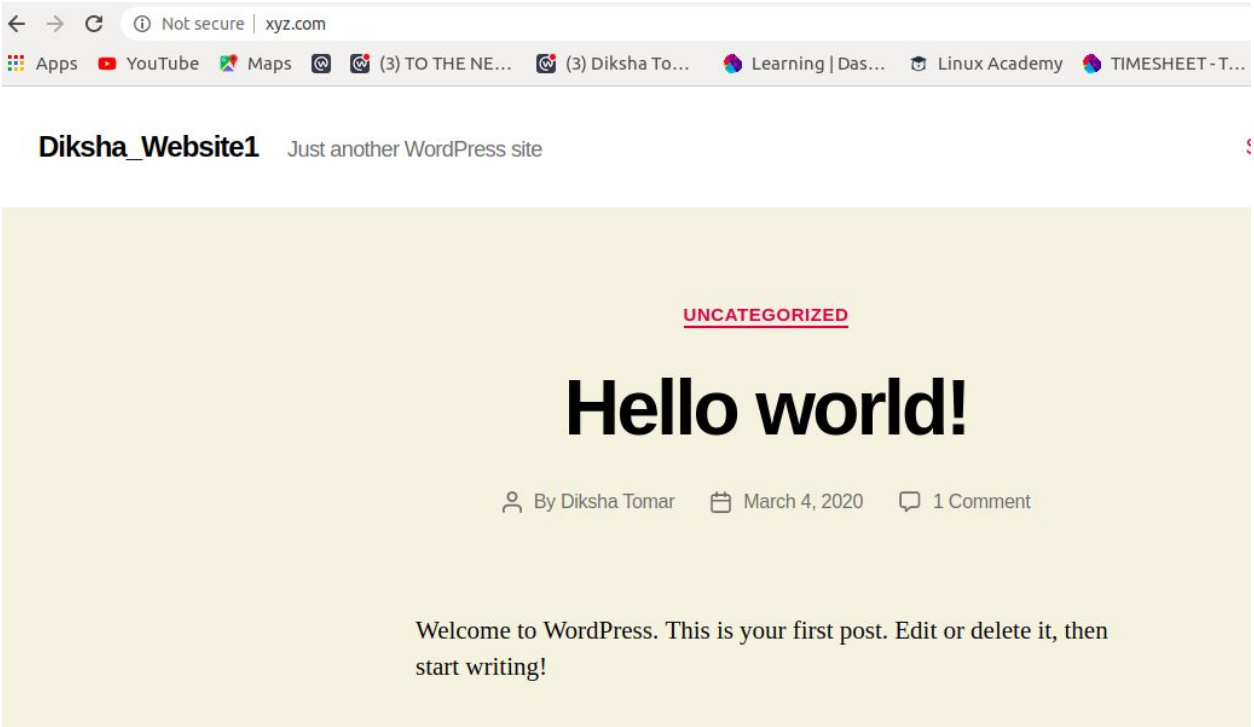
STEP 2: Make entry in the hosts file of your local against your load balancers ip(dig DNS of your LB)



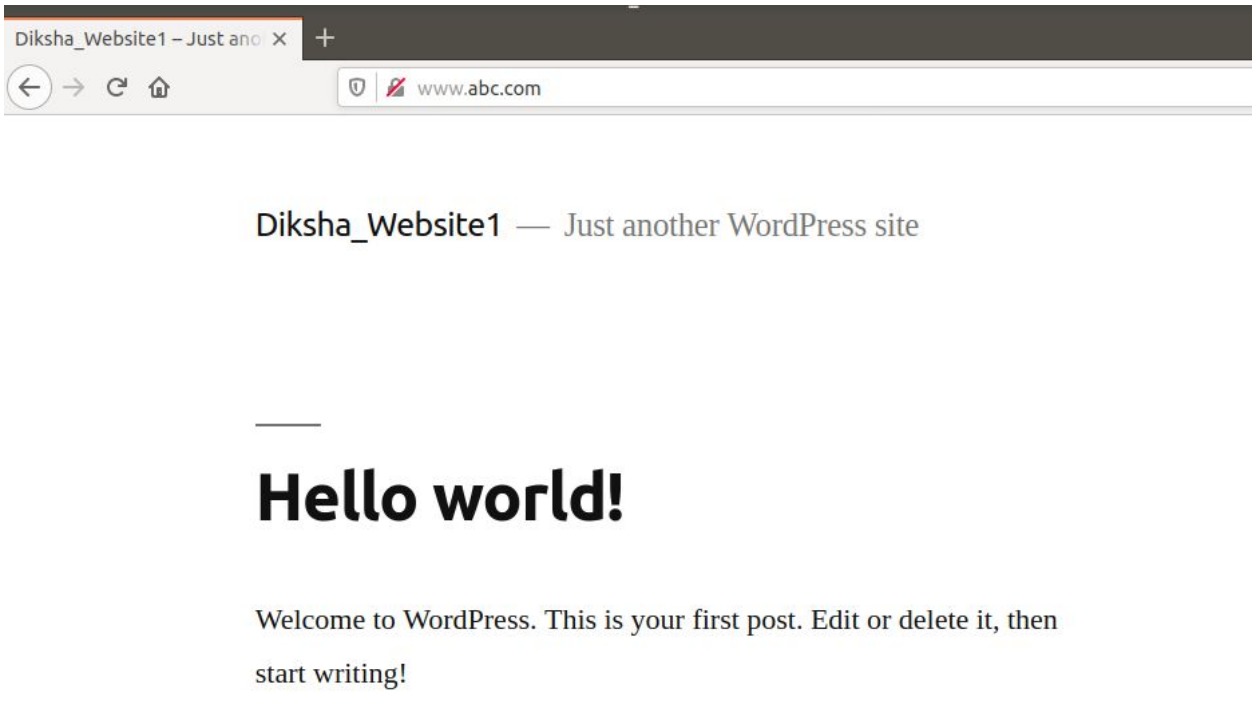
STEP 3: Edit rules and make entry in Your Load balancers rule.



STEP 4: Hit www.xyz.com in your browser



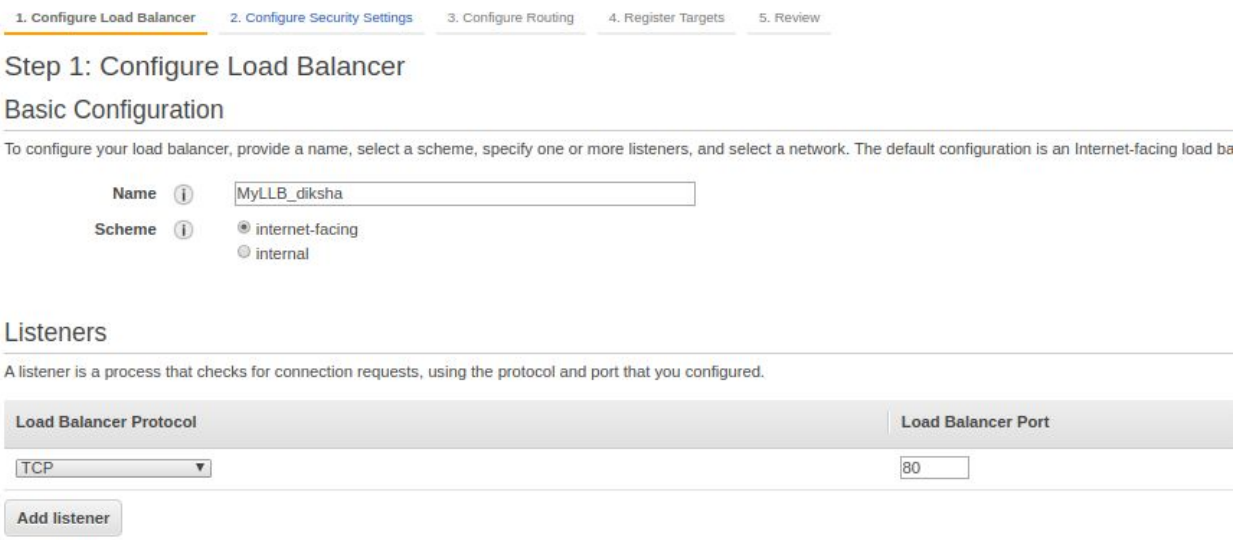
STEP 5: Hit www.abc.com in your browser



****

8.Use NLB that replaces the ALB in the above setup.

STEP 1: Launch NLB



1. Configure Load Balancer | 2. Configure Security Settings | 3. Configure Routing | 4. Register Targets | 5. Review

Step 4: Register Targets

Register targets with your target group. If you register a target in an enabled Availability Zone, the load balancer starts routing requests to the targets as soon as the registration process completes and the target passes the initial health checks.

Registered targets

To deregister instances, select one or more registered instances and then click Remove.

Remove

Instance	Name	Port	State	Security groups	Zone
i-0c2be1129e02ac0fe	diksha_wp	80	running	launch-wizard-3	us-east-1a

Instances

To register additional instances, select one or more running instances, specify a port, and then click Add. The default port is the port specified for the target group. If the instance is already registered on the specified port, you must specify a differ

Add to registered on port 80

Search Instances

Instance	Name	State	Security groups	Zone	Subnet ID	Subnet CIDR
i-0c2be1129e02ac0fe	diksha_wp	running	launch-wizard-3	us-east-1a	subnet-066f3806db4573748	10.0.1.0/24

Load Balancer Creation Status

Successfully created load balancer

Load balancer MyLLBdiksha was successfully created.

Note: It might take a few minutes for your load balancer to be fully set up and ready to route traffic, and for the targets to complete the registration process and pass the initial health checks.

Suggested next steps

- Discover other services that you can integrate with your load balancer. Visit the **Integrated services** tab within MyLLBdiksha
- Consider using AWS Global Accelerator to further improve the availability and performance of your applications. AWS Global Accelerator console

Not secure | myllbdiksha-6855976e36c55e0c.elb.us-east-1.amazonaws.com

Apps | YouTube | Maps | (3) TO THE NE... | (3) Diksha To... | Learning | Das... | Linux Academy | TIMESHE

Diksha_Website1 — Just another WordPress site

Hello world!

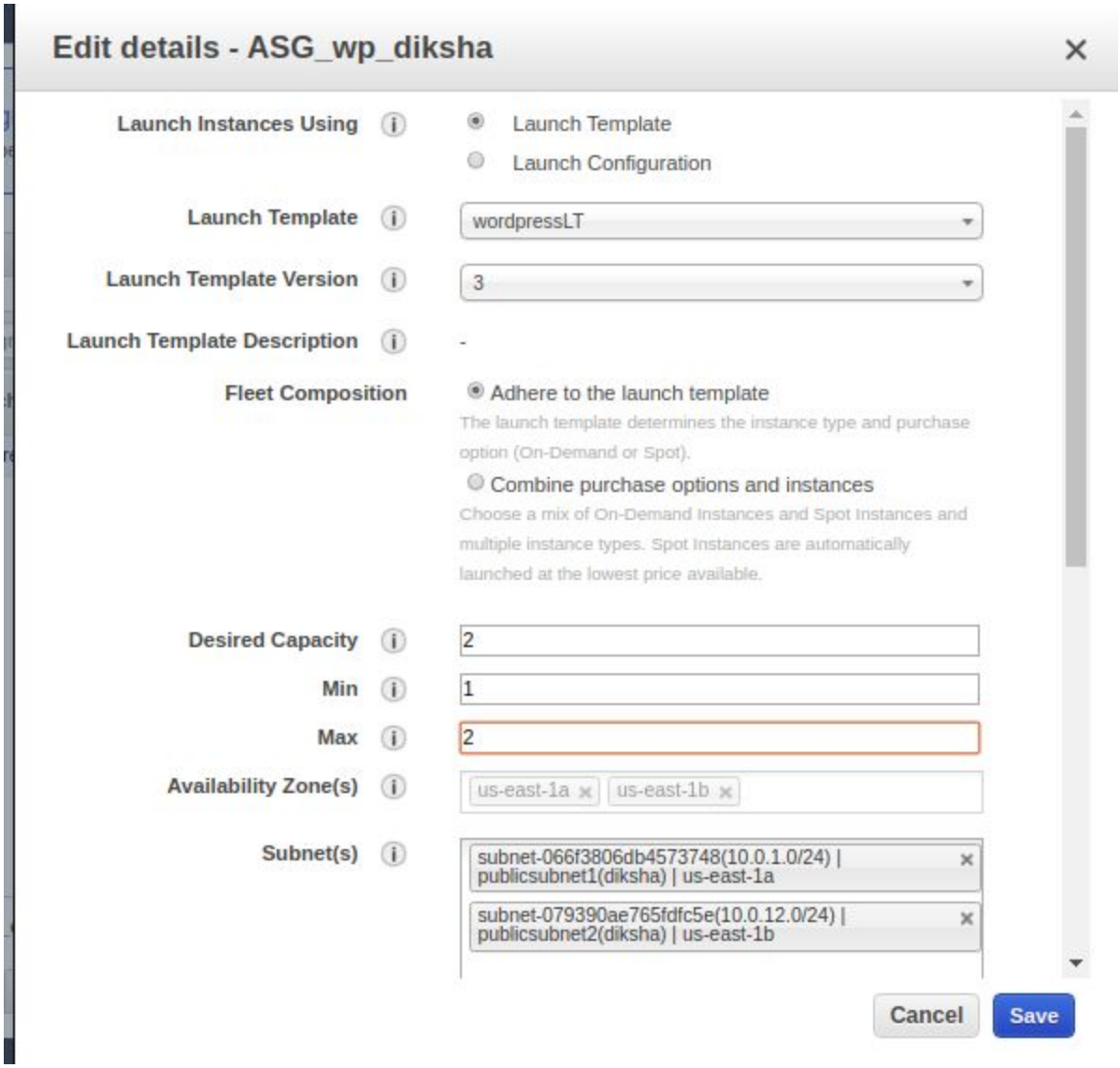
Welcome to WordPress. This is your first post. Edit or delete it, then start writing!

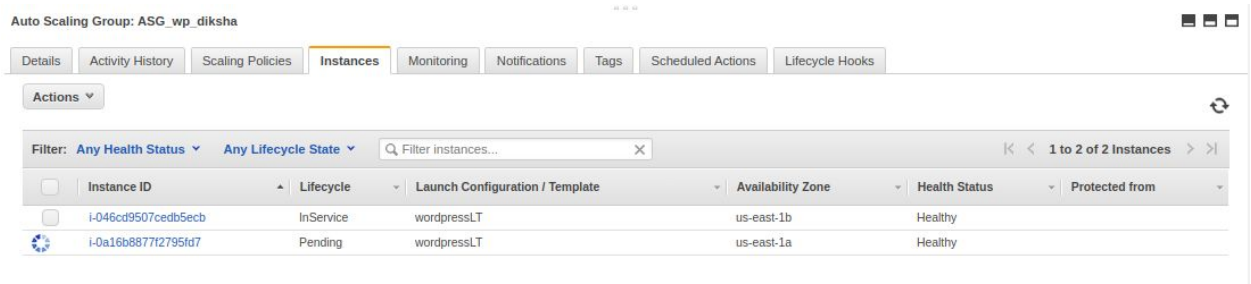
Diksha Tomar | March 4, 2020 | Uncategorized | 1 Comment

****

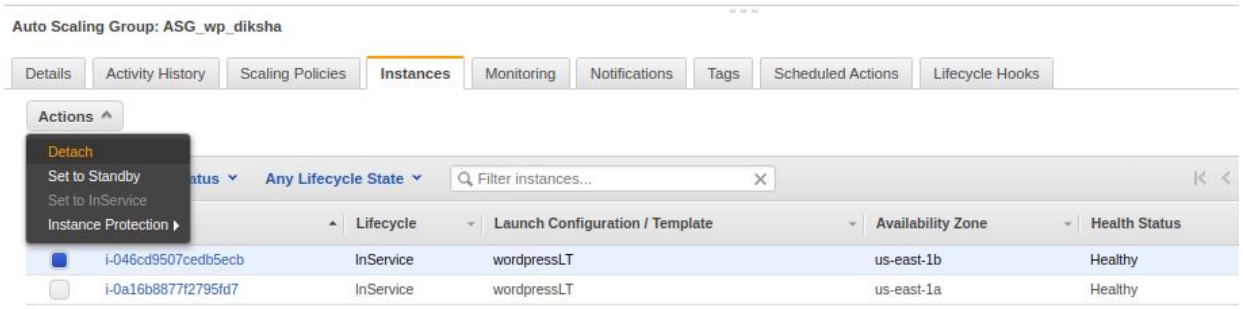
9.Take an instance out of the ASG.

<u>STEP 1:</u> Edit your ASG and edit desired and max as 2

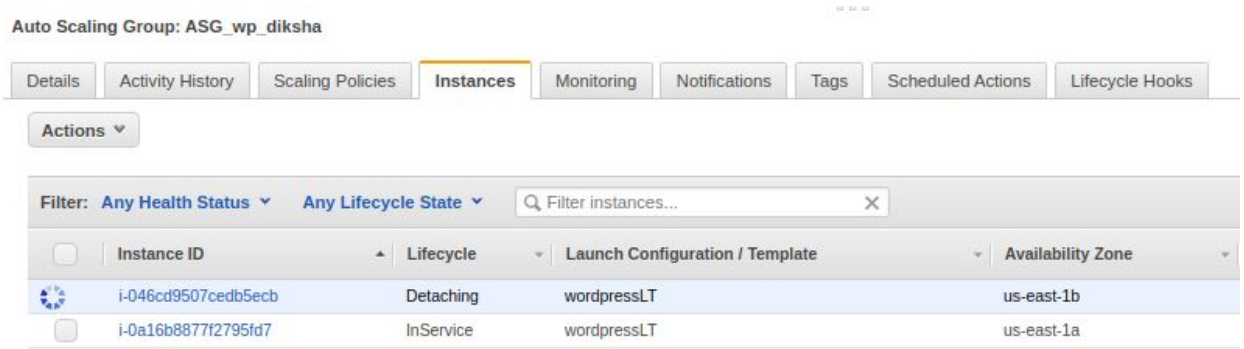




<u>STEP 2:</u> Select the instance you want to detach



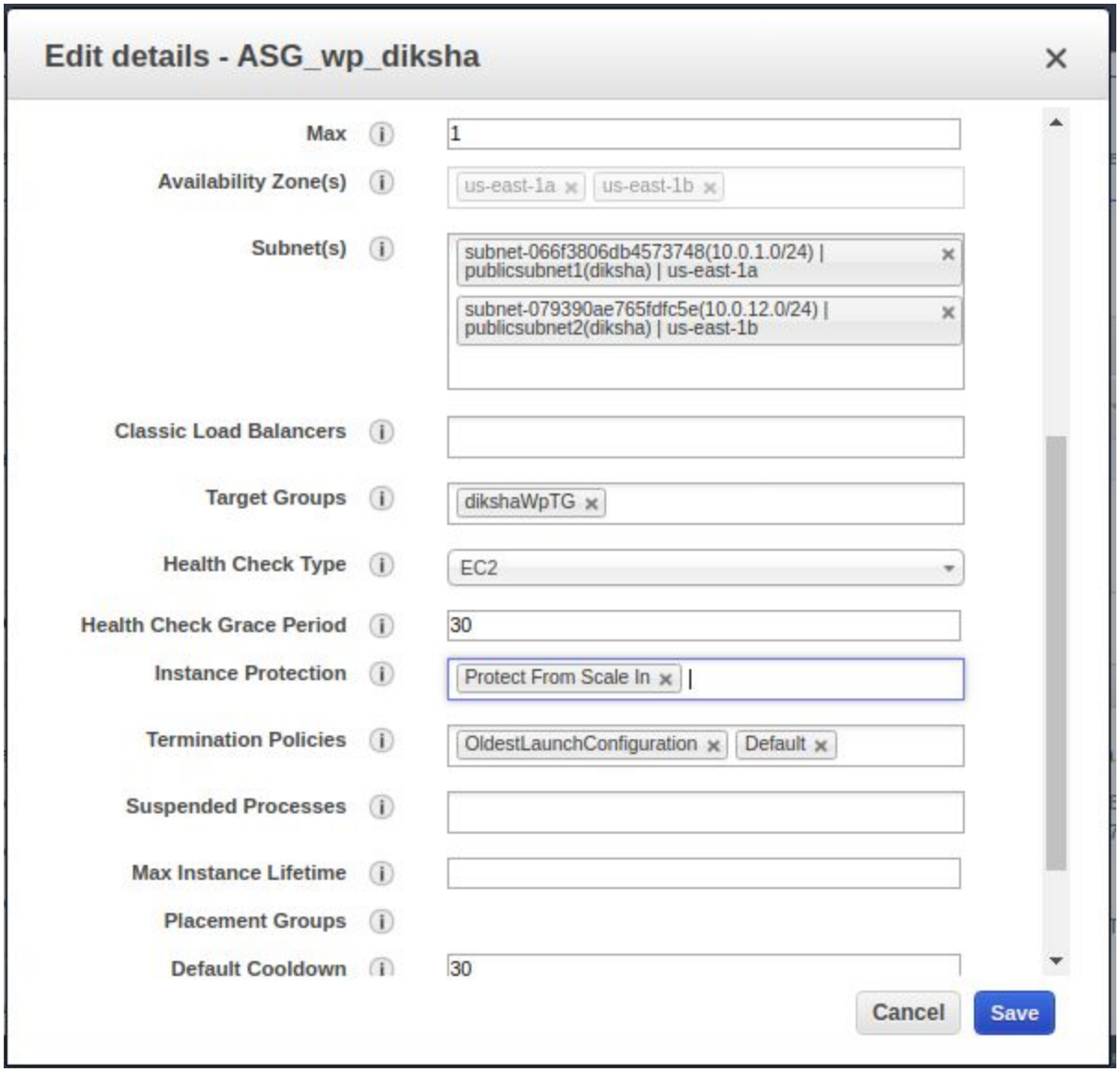
STEP 3: Your instance has been detached



****

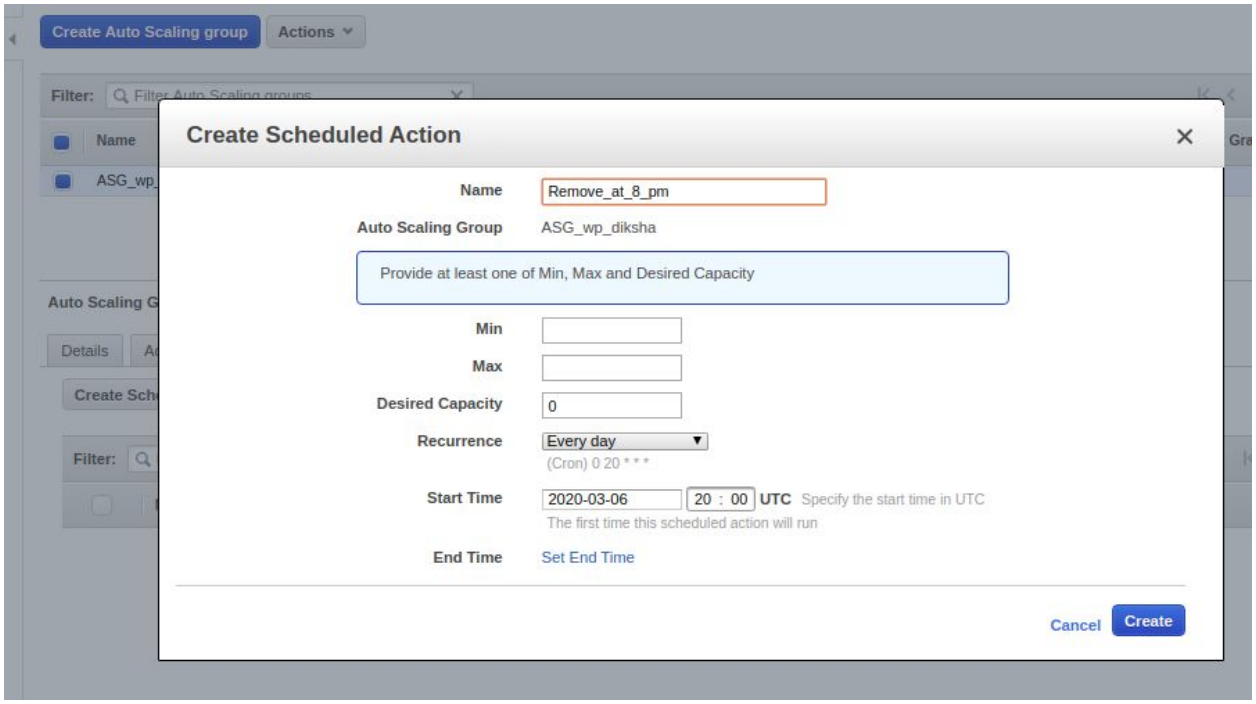
10.Put scale-in protection on an instance in the ASG.

STEP 1: Click on you ASG and then edit it. In instance protection select “Protect from scale in ”



11.Put Schedules in ASG to:

- **Remove all instances of the ASG at 8 PM**



- **Launch a minimum of 2 instances at 10 AM**