Swinburne University of Technology

*COS20019 Cloud Computing Architecture*

*­­­*­­­­

Lab 6 - Scale & Load Balance your Architecture

*Saturday 21th October, 2023*

**Task 1: Create an AMI for Auto Scaling**

In the **AWS Management Console**, in the search box next to **Services** , search for and select **EC2**.

A screenshot of a computer

Description automatically generated

In the left navigation pane, choose **Instances**.

First, you will confirm that the instance is running.

Wait until the **Status Checks** for **Web Server 1** display ***2/2 checks passed***.

A screenshot of a computer

Description automatically generated

Select **Web Server 1**.

In the **Actions** menu, choose **Image and templates** > **Create image**, then configure:

* **Image name:** WebServerAMI
* **Image description:** Lab AMI for Web Server

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Choose **Create image**

A confirmation banner displays the **AMI ID** for your new AMI.

A screenshot of a computer

Description automatically generated

**Task 2: Create a Load Balancer**

In the left navigation pane, choose **Target Groups**.

* Choose **Create target group**
* Choose a target type: **Instances**
* **Target group name**, enter: LabGroup
* Select **Lab VPC** from the **VPC** drop-down menu.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Choose **Next**. The **Register targets** screen appears.

Review the settings and choose **Create target group**

In the left navigation pane, choose **Load Balancers**.

At the top of the screen, choose **Create load balancer**.

A screenshot of a computer

Description automatically generated

Under **Application Load Balancer**, choose **Create**

A screenshot of a computer

Description automatically generated

Under **Load balancer name**, enter: LabELB

A screenshot of a computer

Description automatically generated

Scroll down to the **Network mapping** section, then:

* For **VPC**, choose **Lab VPC**
* Choose the **first** displayed Availability Zone, then select **Public Subnet 1** from the Subnet drop down menu that displays beneath it.
* Choose the **second** displayed Availability Zone, then select **Public Subnet 2** from the Subnet drop down menu that displays beneath it.

You should now have two subnets selected: **Public Subnet 1** and **Public Subnet 2**.

A screenshot of a computer

Description automatically generated

In the **Security groups** section:

* Choose the Security groups drop down menu and select **Web Security Group**
* Below the drop down menu, choose the **X** next to the default security group to remove it.

The **Web Security Group** security group should now be the only one that appears.

A screenshot of a computer

Description automatically generated

For the Listener HTTP:80 row, set the Default action to forward to **LabGroup**.

A screenshot of a computer

Description automatically generated

Scroll to the bottom and choose **Create load balancer**

The load balancer is successfully created.

* Choose **View load balancer**

A screenshot of a computer

Description automatically generated

**Task 3: Create a Launch Template and an Auto Scaling Group**

In the left navigation pane, choose **Launch Templates**.

Choose **Create launch template**

A screenshot of a computer

Description automatically generated

Configure the launch template settings and create it:

* **Launch template name:** LabConfig
* Under **Auto Scaling guidance**, select *Provide guidance to help me set up a template that I can use with EC2 Auto Scaling*

A screenshot of a computer

Description automatically generated

* In the Application and OS Images (Amazon Machine Image) area, choose *My AMIs*.
* **Amazon Machine Image (AMI)**: choose *Web Server AMI*

A screenshot of a computer

Description automatically generated

* **Instance type:** choose *t2.micro*
* **Key pair name**: choose *vockey*

A screenshot of a computer

Description automatically generated

* **Firewall (security groups)**: choose *Select existing security group*
* **Security groups**: choose *Web Security Group*

A screenshot of a computer

Description automatically generated

* Scroll down to the **Advanced details** area and expand it.
* Scroll down to the **Detailed CloudWatch monitoring** setting. Select *Enable*

A screenshot of a computer

Description automatically generated

Choose **Create launch template**

A white background with black lines

Description automatically generated

In the Success dialog, choose the **LabConfig** launch template.

A screenshot of a computer

Description automatically generated

From the **Actions** menu, choose *Create Auto Scaling group*

A screenshot of a computer

Description automatically generated

Configure the details in Step 1 (Choose launch template or configuration):

* **Auto Scaling group name:** Lab Auto Scaling Group
* **Launch template**: confirm that the *LabConfig* template you just created is selected.
* Choose **Next**

A screenshot of a computer

Description automatically generated

Configure the details in Step 2 (Choose instance launch options):

* **VPC**: choose *Lab VPC*
* **Availability Zones and subnets**: Choose *Private Subnet 1* and then choose *Private Subnet 2*.
* Choose **Next**

A screenshot of a computer

Description automatically generated

Configure the details in Step 3 (Configure advanced options):

* Choose **Attach to an existing load balancer**
  + **Existing load balancer target groups**: select *LabGroup*.
* In the **Additional settings** pane:
  + Select **Enable group metrics collection within CloudWatch**
* Choose **Next**

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated

Configure the details in Step 4 (Configure group size and scaling policies - optional):

* Under **Group size**, configure:
  + **Desired capacity:** 2
  + **Minimum capacity:** 2
  + **Maximum capacity:** 6

A screenshot of a computer

Description automatically generated

* + Under **Scaling policies**, choose *Target tracking scaling policy* and configure:
* **Scaling policy name:** LabScalingPolicy
* **Metric type:** *Average CPU Utilization*
* **Target value:** 60

A screenshot of a computer screen

Description automatically generated

Configure the details in Step 5 (Add notifications - optional):

A screenshot of a computer

Description automatically generated

Configure the details in Step 6 (Add tags - optional):

Tags applied to the Auto Scaling group will be automatically propagated to the instances that are launched.

* + Choose **Add tag** and Configure the following:
    - **Key:** Name
    - **Value:** Lab Instance
* Choose **Next**

A screenshot of a computer

Description automatically generated

Configure the details in Step 6 (Review):

* Review the details of your Auto Scaling group
* Choose **Create Auto Scaling group**

**Task 4: Verify that Load Balancing is Working**

In the left navigation pane, choose **Instances**.

You should see two new instances named **Lab Instance**. These were launched by Auto Scaling.

A screenshot of a computer

Description automatically generated

In the left navigation pane, choose **Target Groups**.

Select *LabGroup*

Choose the **Targets** tab.

A screenshot of a computer

Description automatically generated

In the left navigation pane, choose **Load Balancers**.

Select the *LabELB* load balancer.

In the Details pane, copy the **DNS name** of the load balancer, making sure to omit "(A Record)".

A screenshot of a computer

Description automatically generated

Open a new web browser tab, paste the DNS Name you just copied, and press Enter.

A screenshot of a computer

Description automatically generated

**Task 5: Test Auto Scaling**

In the search box next to **Services** , search for and select **CloudWatch**.

A screenshot of a computer

Description automatically generated

In the left navigation pane, choose **All alarms**.

A screenshot of a computer

Description automatically generated

Choose the **OK** alarm, which has *AlarmHigh* in its name.

Return to the browser tab with the web application.

Choose **Load Test** beside the AWS logo.

A screenshot of a computer

Description automatically generated

Return to browser tab with the **CloudWatch** console.

Wait until the **AlarmHigh** alarm enters the *In alarm* state.

A screenshot of a computer

Description automatically generated

In the search box next to **Services** , search for and select **EC2**.

In the left navigation pane, choose **Instances**.

A screenshot of a computer

Description automatically generated

**Task 6: Terminate Web Server 1**

Select **Web Server 1** (and ensure it is the only instance selected).

In the **Instance state** menu, choose **Instance State** > **Terminate Instance**.

Choose **Terminate**

A screenshot of a computer

Description automatically generated

A screenshot of a computer error

Description automatically generated

**END LAB.**