Setup Compute Resources

# What needs to be done

Compute resources config:

1. 2x EC2 Instances - t2.medium
2. 1x AutoScaling Group - Min 1 instance running at all times - Simple Scaling Policy
3. 3. 1x Elastic Load Balancer - simple routing policy between the two EC2 instances

# My Steps

Guide that I followed → [Tutorial: Set up a scaled and load-balanced application](https://docs.aws.amazon.com/autoscaling/ec2/userguide/tutorial-ec2-auto-scaling-load-balancer.html)

1. **Create a Target Group**
   1. Use Default VPC
   2. Skip register targets for now
2. **Create a Load Balancer**
   1. Default scheme and IP address type
   2. Use Default VPC
   3. For mappings, use us-east-1a and us-east-1b
   4. Use default security group
   5. In Listeners and Routing, set default action to target group created
3. **Create a Security Group for the EC2 instances**
   1. Why? Allows the compute-resources load balancer to communicate with EC2 instances on the listener port (and the health check port, but we will not set that up now).
   2. Add an inbound rule with type HTTP, port 80, and source as the load balancer’s security group
4. **Set up a launch template**
   1. Use Amazon Linux 2 AMI (HVM) - Kernel 5.10, SSD Volume Type
   2. Use t2.medium instances
   3. Use fhir\_server Key pair
   4. Use security group created above
   5. Be sure to install a web server, such as Apache or Internet Information Services (IIS), on each EC2 instance.
   6. Expand Advanced network configuration
      1. Add network interface
      2. Select Yes for Delete on termination
5. **Create an Auto Scaling group from launch template**
   1. In Choose instance launch options
      1. Use Default VPC
      2. Select us-east-1a and us-east-1b for Availability Zones and subnets
   2. Select Attach to an existing load balancer and choose the created load balancer
   3. In Group Size
      1. Desired Capacity set to 2
      2. Min Capacity set to 1
      3. Max Capacity set to 2
   4. In Scaling Policies, select None
6. **Verify that load balancer is attached**
7. **How to test**
   1. Go into one instance and shut it down/stop it
   2. Auto scaling should automatically create another instance
   3. Try different scaling to test different scenarios
8. **Clean up**
   1. Delete Auto Scaling group
   2. Delete launch template or launch configuration
   3. Delete load balancer
   4. Delete target group
   5. Delete EC2 instances?

# Helpful Documentation

## EC2

[What is Amazon EC2?](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/concepts.html)

[Tutorial: Get started with Amazon EC2 Linux instances](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EC2_GetStarted.html)

[Video walkthrough](https://drive.google.com/file/d/1rfxh8QBhZHi36RCqrMzsSGp1dbiSIK9d/view?usp=sharing)

## Load Balancing

[What is an Application Load Balancer?](https://docs.aws.amazon.com/elasticloadbalancing/latest/application/introduction.html)

[Getting started with Application Load Balancers](https://docs.aws.amazon.com/elasticloadbalancing/latest/application/application-load-balancer-getting-started.html)

[Target security groups](https://docs.aws.amazon.com/elasticloadbalancing/latest/application/target-group-register-targets.html#target-security-groups)

## EC2 Auto Scaling

[What is Amazon EC2 Auto Scaling?](https://docs.aws.amazon.com/autoscaling/ec2/userguide/what-is-amazon-ec2-auto-scaling.html)

[Get started with Amazon EC2 Auto Scaling](https://docs.aws.amazon.com/autoscaling/ec2/userguide/get-started-with-ec2-auto-scaling.html)