**Lab 6 Activity Report (FINAL EXAM): Scale and Load Balance your Architecture (100 points)**

**ITELEC 4103**

**Overview:**

Scaling and load balancing are essential components of cloud infrastructure that ensure high availability, performance, and reliability of applications. This activity focuses on leveraging AWS tools and services to dynamically manage workloads and optimize resources. It demonstrates how to use scaling and load balancing to handle increasing or fluctuating application demands effectively.

In this lab, you will use Elastic Load Balancing and Amazon EC2 Auto Scaling to load balance and scale your infrastructure. An initial infrastructure setup will be provided.

**REPORT TEMPLATE:**

The report **should not exceed 8 pages**. For each additional page beyond the limit, a 5-point deduction will be applied.

Please **use screenshots** to enhance your explanations or clarify specific points. Screenshots should be used to support the discussion and illustrate key concepts.

**I. Brief Rationale for the Activity Conducted:**

Provide a short introduction explaining the purpose of this activity. Discuss the importance of scaling and load balancing in cloud infrastructure, to ensure high availability, performance, and reliability of cloud applications.

**II. Discussion (based on the tasks performed):**

* **Creating an AMI:**
  + Discuss how you created an AMI from a running AMI instance to serve as a template to replicate instances with the same configurations and applications, enabling quick scaling
* **Creating an Application Load Balancer:**
  + Discuss how you setup to distribute incoming traffic across multiple EC2 instances.
* **Creating a Launch Configuration and an Auto-scaling Group:**
  + Discuss how you created the launch configuration to define the specifications of instances that the auto-scaling group will manage.
  + Discuss how you configured the Auto-scaling group to automatically launch and terminate instance bases on defined scaling policies, ensuring the architecture can adjust to workload changes.
* **Automatically Scaling New Instances within a Private Subnet:**
  + Discuss how you configured auto-scaling to launch instances in a private subnet to maintain security while ensuring that resources are dynamically provisioned as needed.
* **Creating Amazon CloudWatch Alarms and Monitoring Performance:**
  + Discuss how alarms are setup to monitor infrastructure performance metrics to trigger scaling actions based on pre-defined thresholds.

**III. Analysis Scenarios (based on the tasks performed):**

**Scenario 1**: What happens to application performance when there is a sudden increase in user traffic? Analyze how the Auto Scaling group responds to the increased load and ensures availability by launching additional instances.

**Scenario 2**: Simulate a scenario where one or more instances become unhealthy. Observe how the application load balancer routes traffic only to healthy instances, ensuring uninterrupted service.

**Scenario 3**: Evaluate the impact of removing unused instances during periods of low traffic. Discuss how this approach optimizes cost and resource usage.

**Scenario 4**: Examine CloudWatch metrics and alarms to understand performance trends and determine whether the scaling policies and thresholds are effectively meeting application needs.

**RUBRICS FOR THE ACTIVITY**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **5 pts** | **10 pts** | **15 pts** | **25** **pts** |
| **Focus** | The discussion lacked clarity and direction, with little to no connection to the main topic or activity. | Some topics were addressed, but the discussion was not comprehensive or lacked depth in several areas. | All relevant topics were addressed with a clear and organized focus, though some areas may have lacked detail. | Each topic was addressed with thorough, detailed, and well-organized discussion, providing a deep focus on the key areas. |
| **Clarity** | The answer or discussion was unclear and difficult to follow, lacking a direct connection to the topic. | The answer or discussion addressed the topic, but in a roundabout or vague manner, making it harder to understand. | The answer or discussion directly addressed the topic, providing clarity and coherence. | The answer or discussion was exceptionally clear, concise, and to the point, making the topic easy to understand. |
| **Organization** | The answer or discussion had no clear structure, leaving gaps for readers to fill in. | The answer or discussion had some structure and painted a complete picture, but the organization was not orderly, making it less effective. | Details were presented in a logical order, but the structure could have been more engaging. Some loose ends remained, and the flow of the discussion was somewhat uneven. | Details were logically ordered and presented in engaging ways. The discussion flowed smoothly, covering all main points clearly, and concluded effectively. |
| **Application** | The answer or discussion showed a lack of understanding of basic concepts and principles, with no clear connection to the material. | The answer or discussion was able to state a few concepts and principles but did not fully apply them. | The answer showed a clear understanding and application of basic concepts and principles, demonstrating comprehension of the material. | The answer integrated and applied concepts and principles from the module materials effectively, showing a deep understanding and practical application of the ideas. |