

Intelligent Cloud Infrastructure Review with GenAI Assistance

(AWS Cloud Security & Optimization Project)

Name: SUMITHA A

Reg. No.: 727823TUCY048

1. Executive Summary

Modern cloud environments generate a massive volume of infrastructure telemetry, configuration data, and operational logs. Manually analyzing this data to identify security risks, inefficiencies, and misconfigurations is time-consuming and error-prone. This project demonstrates an Intelligent Cloud Infrastructure Review using Amazon Web Services (AWS) combined with Generative AI (GenAI) in an advisory, human-in-the-loop role.

The objective of this project is to design, deploy, monitor, and review a containerized cloud environment while leveraging GenAI to:

- Summarize infrastructure behavior
- Identify security risks and misconfigurations
- Suggest optimization and hardening strategies

The environment consists of an Amazon EC2 instance running Docker, hosting an Nginx web application, deployed inside a custom Amazon VPC. Amazon CloudWatch Logs is used for centralized logging of system and container logs, which are later reviewed using GenAI for insights.

Importantly, GenAI is used only for analysis and decision support. It has no direct access to the AWS environment, ensuring security and compliance. This approach aligns with enterprise best practices and the AWS Well-Architected Framework.

2. Problem Statement

2.1 Background

Traditional cloud monitoring and observability tools primarily focus on collecting metrics, events, and raw log data from cloud resources. While these tools are effective at providing visibility, they often lack the contextual intelligence required to interpret the data and derive meaningful insights. As cloud environments grow in complexity, manually correlating logs and configurations across multiple services becomes increasingly difficult.

IT leadership and cloud administrators require smarter, high-level insights into key areas such as:

- Infrastructure usage and resource efficiency
- Security exposure and access control weaknesses
- Operational inefficiencies affecting performance and cost
- Misconfigurations across compute, networking, storage, and containerized workloads

Although Artificial Intelligence can assist in analyzing this data, fully automated AI systems with direct write access to production infrastructure introduce significant security, compliance, and governance risks. Therefore, a controlled, advisory approach is necessary, where AI supports decision-making while critical actions remain under human oversight.

2.2 Problem Definition

The core challenge addressed in this project is the design of a GenAI-assisted cloud infrastructure review system that enhances visibility and decision-making without compromising security or control. The system must be capable of analyzing infrastructure logs and configuration data collected from cloud resources to extract meaningful operational and security insights.

Specifically, the solution should be able to:

- Analyze system and application logs along with configuration metadata
- Identify risky patterns, security gaps, and misconfigurations
- Suggest actionable improvements and optimization measures
- Maintain full human control by ensuring GenAI operates strictly in an advisory role

This approach balances the analytical strengths of Generative AI with the need for human oversight, making it suitable for real-world enterprise cloud environments.

2.3 Scope of the Project

The scope of this project is limited to the analysis and review of a controlled AWS-based cloud environment using GenAI as a decision-support tool. The project focuses on evaluating infrastructure behavior, security posture, and operational efficiency based on logs and configuration data.

The key areas covered within the scope include:

- Reviewing AWS infrastructure logs and configuration settings related to compute, networking, and containers
- Identifying potential security risks, misconfigurations, and operational inefficiencies
- Utilizing GenAI strictly for advisory analysis, insight generation, and recommendation purposes
- Emphasizing documentation, analysis, and reporting rather than automated remediation
- Ensuring that no real production workloads, sensitive information, or confidential data are used

By clearly defining these boundaries, the project maintains security, compliance, and ethical use of AI while effectively demonstrating GenAI-assisted cloud infrastructure review capabilities.

3. Objectives of the Project

The primary objective of this project is to demonstrate how Generative AI can be effectively integrated into cloud infrastructure reviews while maintaining security and human oversight. To achieve this, the project is guided by the following specific objectives:

1. To design and implement a secure AWS cloud architecture that follows industry best practices
2. To deploy a containerized web application using Docker on an Amazon EC2 instance
3. To enable centralized logging and monitoring using Amazon CloudWatch Logs
4. To review infrastructure behavior and log data using GenAI for analytical insights
5. To identify security misconfigurations and operational risks within the environment
6. To propose security hardening and optimization recommendations based on the review
7. To document the complete implementation, analysis, and findings with supporting evidence

These objectives ensure a structured approach to building, analyzing, and documenting a GenAI-assisted cloud infrastructure review.

4. High-Level Architecture

4.1 Architecture Overview

The cloud environment is designed using a modular and isolated architecture, ensuring clear separation between compute, networking, storage, and monitoring components. This approach improves security, scalability, and manageability by allowing each layer of the infrastructure to be independently monitored and controlled. The architecture follows AWS best practices and supports effective logging and analysis for GenAI-assisted reviews.

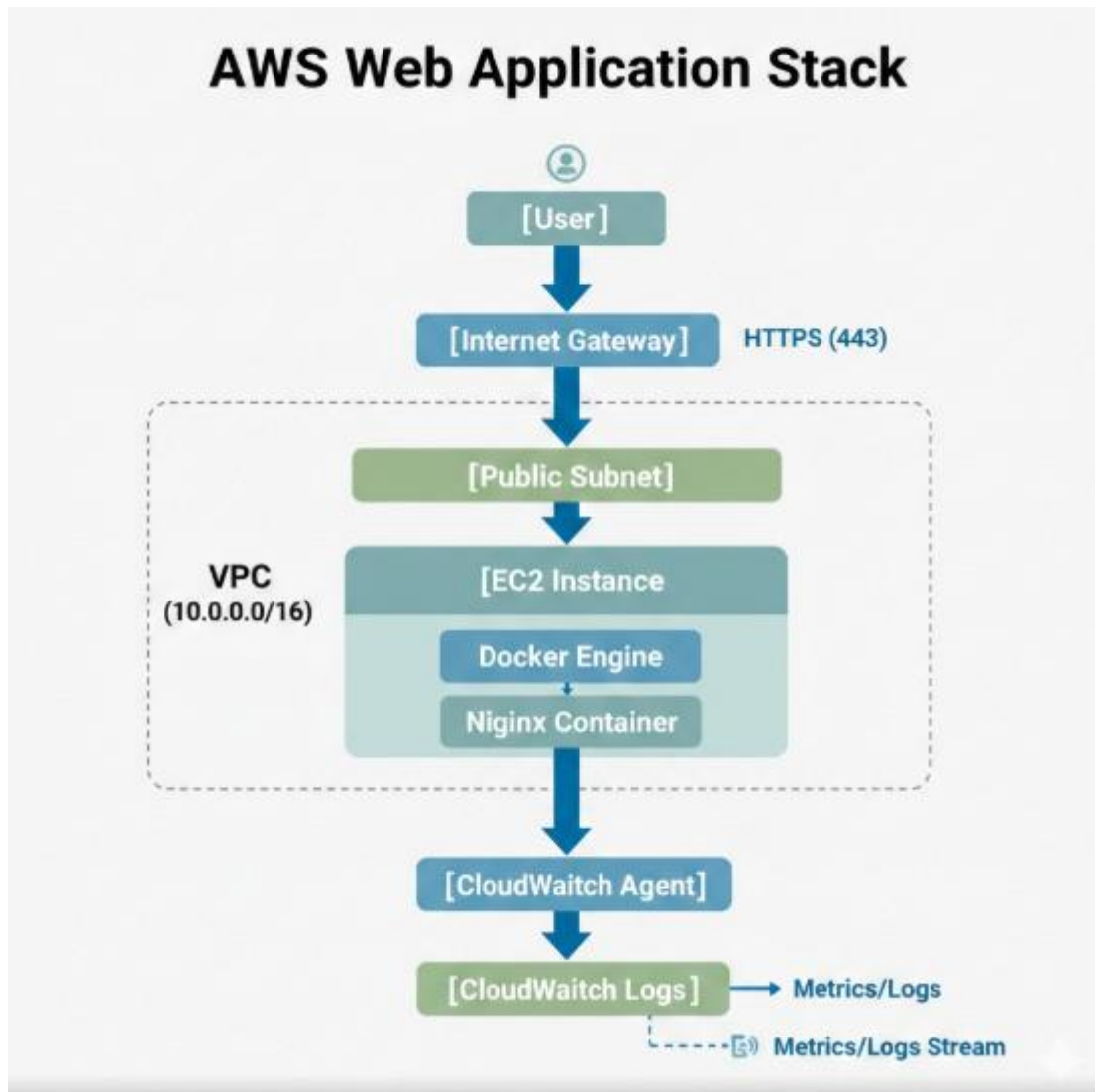
4.2 Architecture Components

The major components of the architecture are described below:

- **Amazon Virtual Private Cloud (VPC)**
 - CIDR Block: 10.0.0.0/16
 - Provides an isolated virtual network for hosting cloud resources
- **Public Subnet**
 - Hosts the EC2 instance running the containerized application
 - Enables controlled access to the application through defined security rules
- **Amazon EC2 (t2.micro)**
 - Acts as the primary compute resource
 - Runs the Docker Engine and hosts the Nginx container
- **Docker (Container Runtime)**
 - Provides a lightweight and portable environment for running the Nginx web server
 - Simplifies application deployment and management
- **Amazon CloudWatch Logs**
 - Collects and centralizes system and container logs

- Enables monitoring, troubleshooting, and analysis
- **Amazon S3 (Optional)**
 - Used for long-term archival of logs and audit data
 - Supports cost-effective storage and compliance requirements

4.3 Architecture Diagram



5. Infrastructure Implementation

5.1 VPC and Network Configuration

A custom Amazon VPC was created to provide network isolation for the application environment. Public subnets, route tables, and an Internet Gateway were configured to enable controlled external access. Security Groups acted as virtual firewalls, allowing only required inbound and outbound traffic.

VPC Resource Map:

The image displays two screenshots of the AWS VPC console, illustrating the resource map for a VPC named 'GenAI-VPC' (VPC ID: vpc-0de4b28da1af61542).

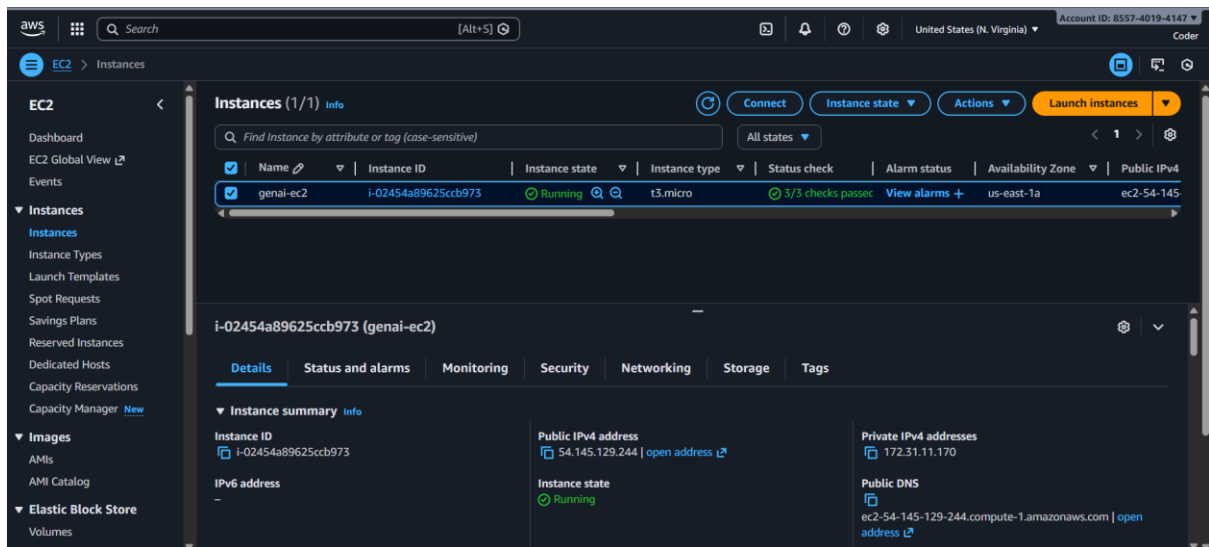
Top Screenshot: The 'Resource map' section shows the VPC connected to a single subnet named 'us-east-1d' (subnet1) and a single route table named 'rtb-008cb826eec69d960'. The VPC is in the 'Available' state.

Bottom Screenshot: The 'Resource map' section shows the VPC connected to a single subnet named 'us-east-1d' (subnet1) and a single route table named 'rtb-008cb826eec69d960'. Additionally, a 'Network Connections (1)' section is visible, showing a connection to an 'igw' (Internet Gateway).

5.2 EC2 Instance Setup

An Amazon EC2 instance running Amazon Linux 2023 was launched within the public subnet. Docker was installed using user-data scripts to ensure consistency and repeatability across deployments.

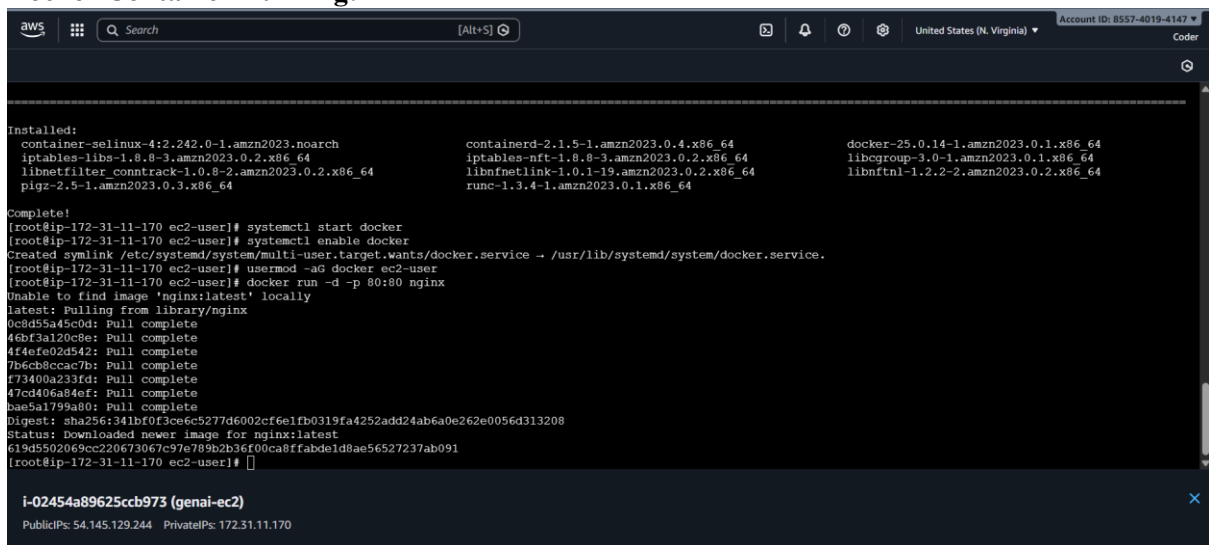
EC2 Instance Summary:



5.3 Docker and Application Deployment

Docker was configured on the EC2 instance to run an Nginx container. The container exposes port 80, enabling HTTP access to the web application and validating successful container deployment.

Docker Container Running:



6. Logging and Monitoring Configuration

6.1 Need for Centralized Logging

In cloud-native environments, logs are generated across multiple layers such as operating systems, applications, containers, and network components. Centralized logging consolidates these logs into a single platform, making analysis and monitoring more efficient.

Centralized logging enables:

- Faster troubleshooting by correlating events across services
- Improved security auditing and forensic analysis
- Easier compliance reporting and log retention
- Effective AI-assisted analysis by providing structured and complete data

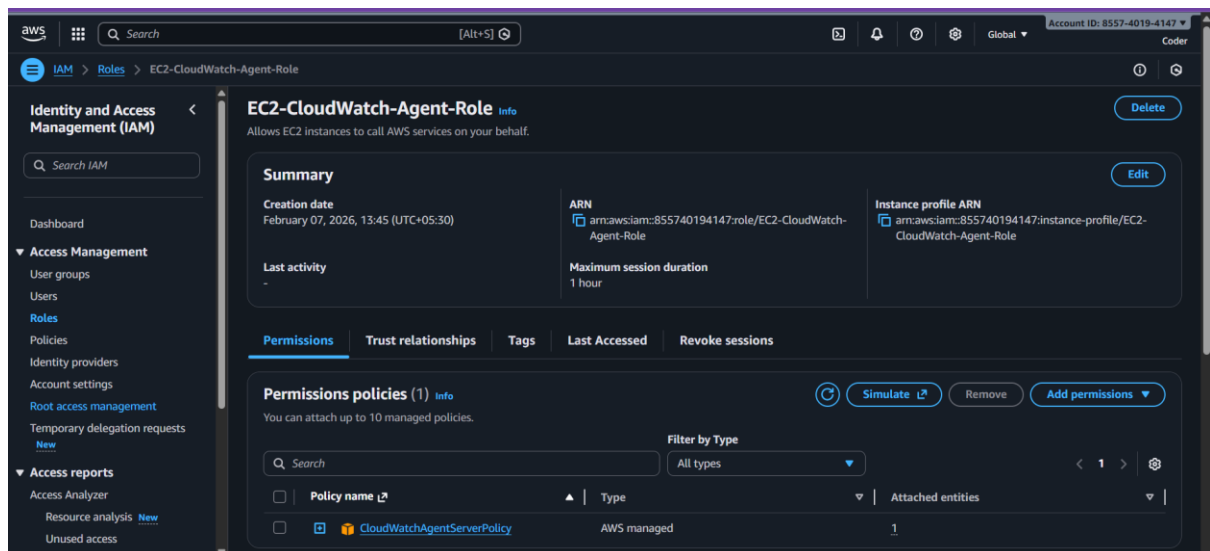
Amazon CloudWatch Logs was selected as it integrates seamlessly with AWS services and supports scalable log ingestion and analysis.

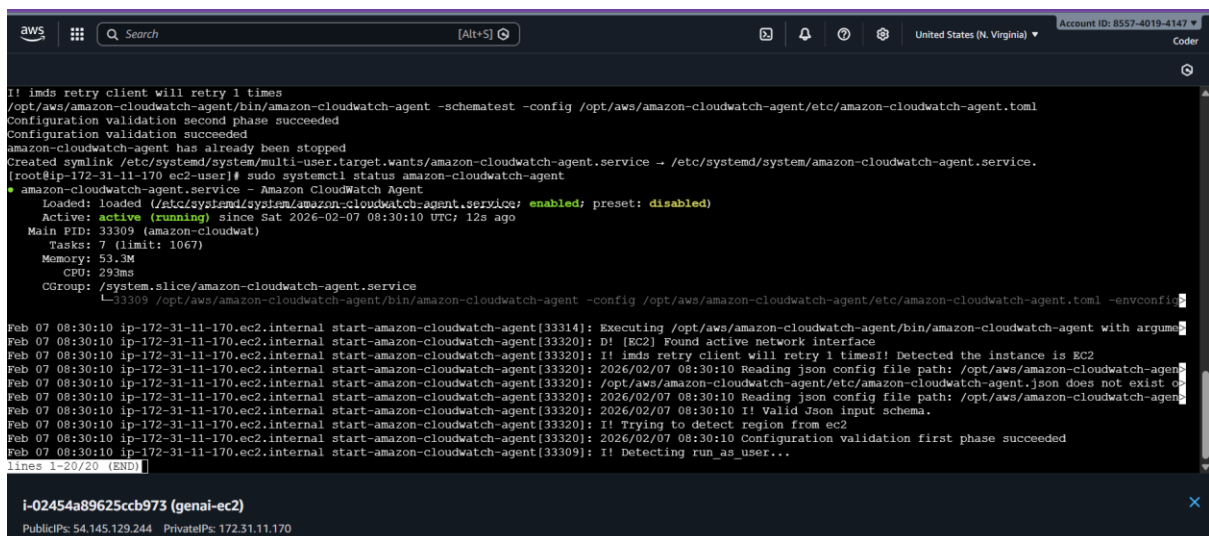
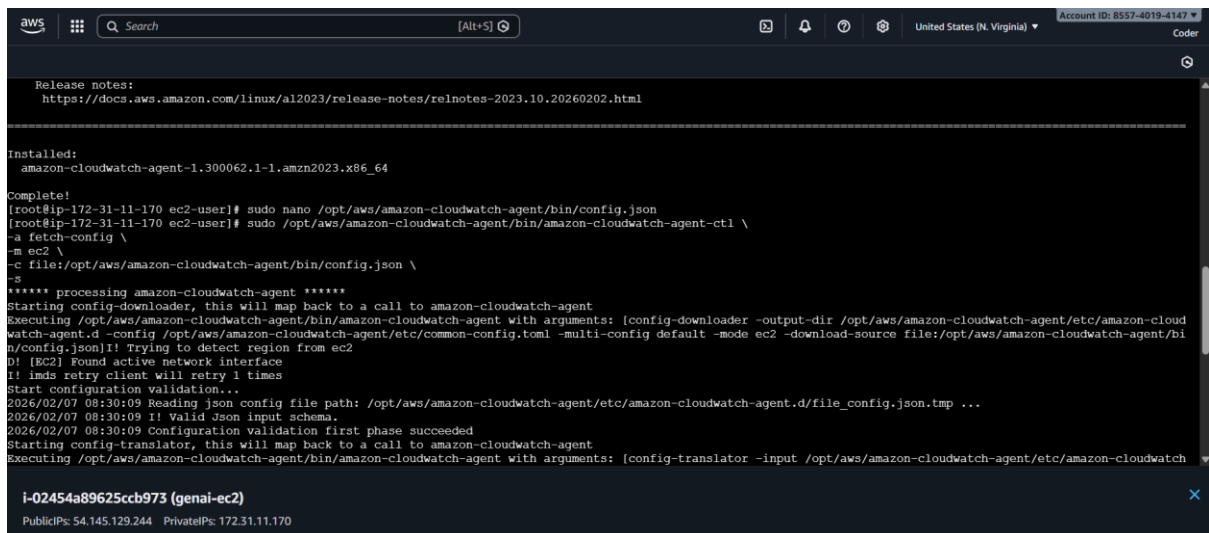
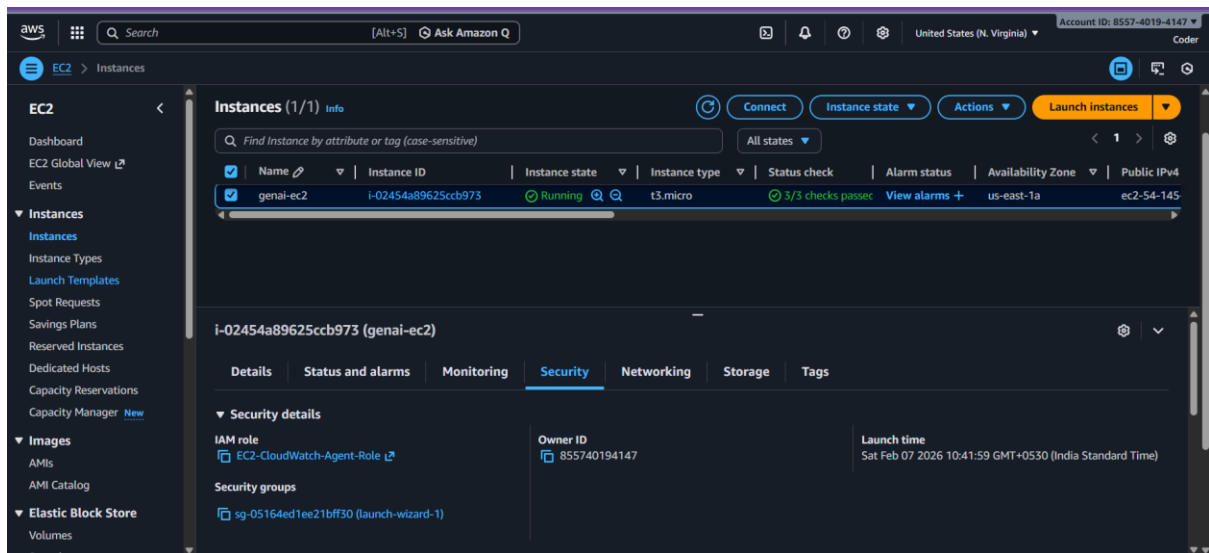
6.2 CloudWatch Agent Configuration

The Amazon CloudWatch Agent was installed and configured on the EC2 instance to collect system-level and application logs. Since Amazon Linux 2023 uses journald by default, rsyslog was enabled to forward logs into /var/log/messages, ensuring compatibility with the CloudWatch Agent configuration.

The agent configuration defined log file paths, log group names, and stream naming conventions based on the EC2 instance ID. This ensured logs were reliably collected and easily traceable to the source instance.

CloudWatch Agent Running:





6.4 Amazon S3 Log Archival Configuration

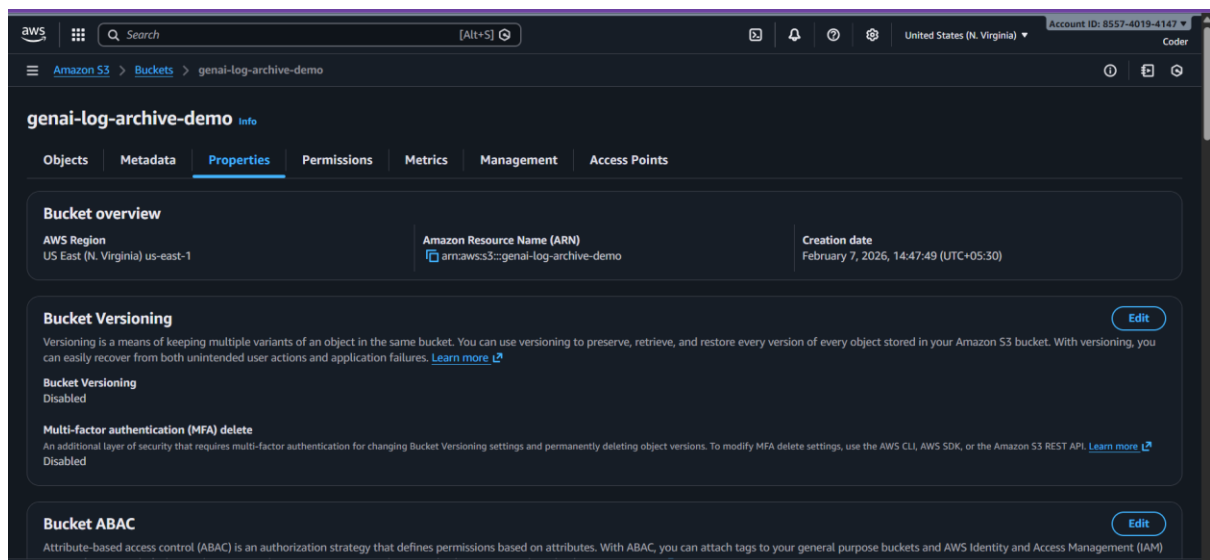
To support long-term log retention and cost-effective storage, an Amazon S3 bucket was created for archiving CloudWatch Logs. While CloudWatch is suitable for real-time monitoring and short-term analysis, S3 provides durable, scalable, and low-cost storage for historical data.

The S3 bucket was configured with:

- Private access enabled by default
- Server-side encryption (SSE-S3) for data at rest
- Block Public Access enabled to prevent accidental exposure

CloudWatch Logs can be exported periodically to the S3 bucket for compliance audits, forensic investigations, and offline analysis using analytics tools such as Amazon Athena.

Amazon S3 Bucket Configuration:



6.5 Benefits of Using Amazon S3 for Log Storage

Using Amazon S3 for log archival offers several advantages:

- Long-term retention of logs beyond CloudWatch limits
- Reduced operational cost for older log data
- High durability and availability
- Easy integration with analytics and AI-based review tools

This approach ensures that critical infrastructure and security logs remain accessible for future audits and GenAI-assisted analysis without impacting real-time system performance.

6.3 Log Groups Created

Two primary CloudWatch Log Groups were created for structured monitoring:

- **/ec2/system/logs** → Captures operating system and service logs
- **/ec2/docker/logs** → Captures Docker container logs

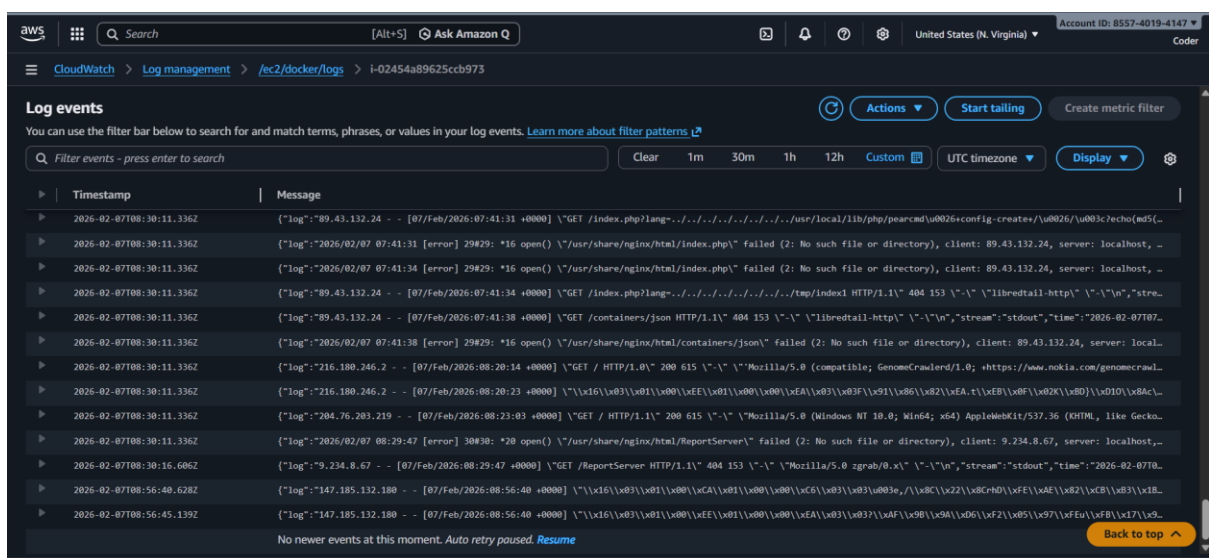
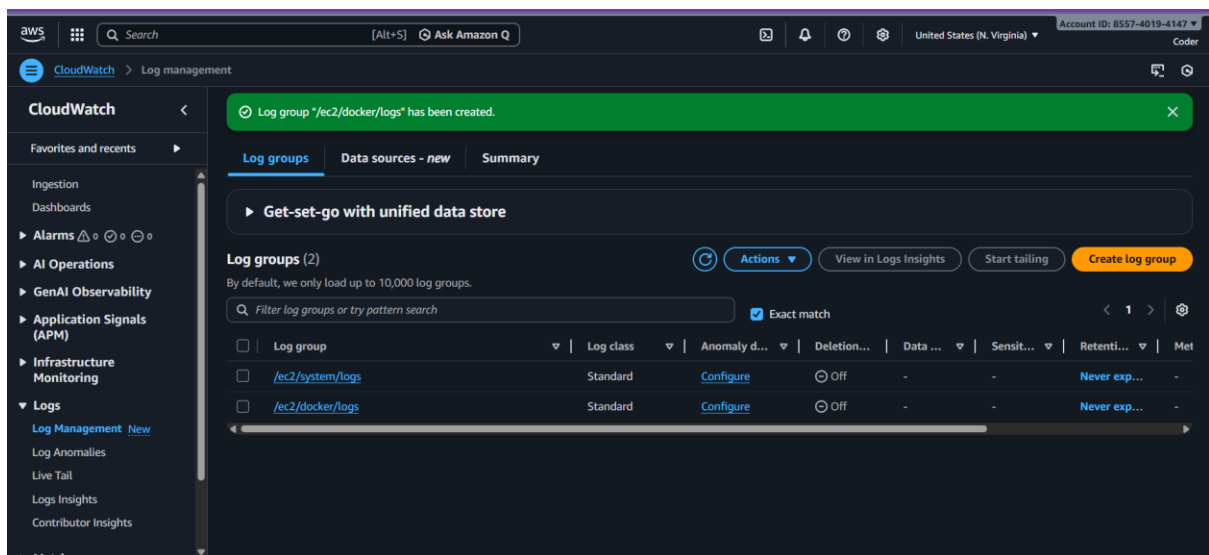
Log streams were automatically generated using the EC2 instance ID, enabling clear identification of log sources and simplifying multi-instance scalability.

7. Log Verification and Evidence

7.1 System Logs

System logs from `/var/log/messages` were successfully ingested into Amazon CloudWatch Logs. These logs include system boot events, service status updates, and security-related messages. The availability of these logs confirms correct agent configuration and network connectivity.

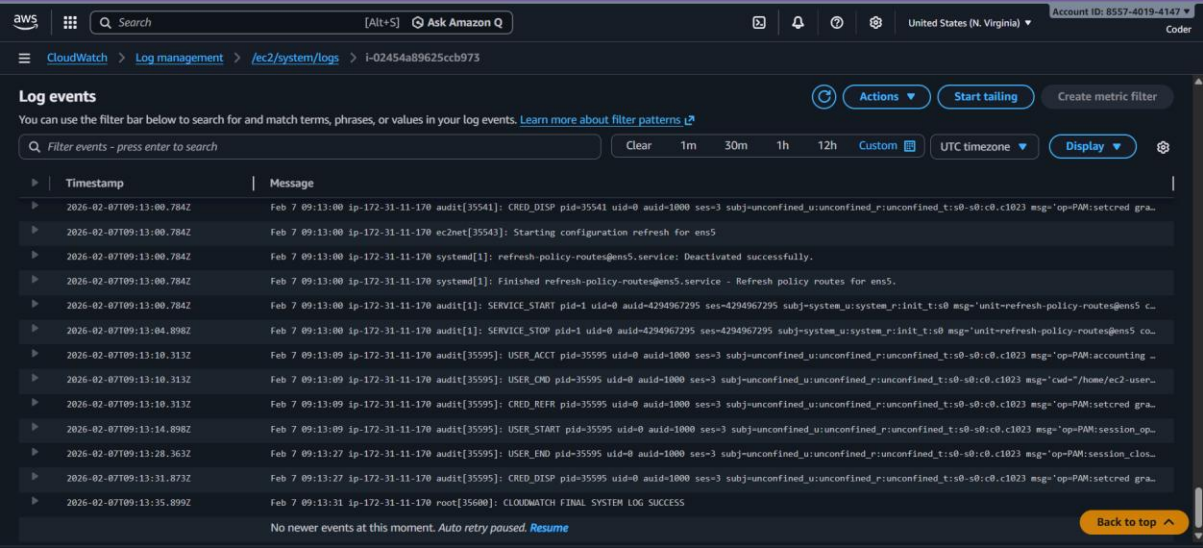
System Log Stream:



7.2 Docker Logs

Docker container logs were collected from /var/lib/docker/containers/*/*.log. These logs capture application-level events generated by the Nginx container, including startup messages and HTTP request handling. This ensures visibility into container behavior and application health.

Docker Log Stream:



8. GenAI-Assisted Security Review

8.1 Methodology

Infrastructure metadata, architectural details, and summarized log outputs were manually provided to GenAI (ChatGPT) for analysis. GenAI was used strictly in an advisory, human-in-the-loop role to:

- Identify risky configurations
- Highlight deviations from AWS best practices
- Suggest security and optimization improvements

GenAI had no direct access to the AWS environment, ensuring full human control and eliminating the risk of unintended automated changes.

8.2 Findings and Risk Analysis

Based on the review, the following key findings were identified:

Category	Finding	Risk Level	GenAI Recommendation
Network	SSH (port 22) open to 0.0.0.0/0	High	Restrict IP range or use SSM
IAM	Broad permissions assigned	Medium	Apply least-privilege policies
Logging	Logs manually reviewed	Low	Enable alerts and automation

Security Group Inbound Rules:

Security group name - *required*
launch-wizard-1

This security group will be added to all network interfaces. The name can't be edited after the security group is created. Max length is 255 characters. Valid characters: a-z, A-Z, 0-9, spaces, and _-./!@,[@!+=&(){}]*

Description - *required* [Info](#)
launch-wizard-1 created 2026-02-07T04:59:02.474Z

Inbound Security Group Rules

▼ Security group rule 1 (TCP, 22, 0.0.0.0/0) [Remove](#)

Type Info	Protocol Info	Port range Info
ssh	TCP	22
Source type Info	Source Info	Description - <i>optional</i> Info
Anywhere	<input type="text" value="0.0.0.0/0"/> Add CIDR, prefix list or security group	e.g. SSH for admin desktop

Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only. [X](#)

[Add security group rule](#)

Summary

Number of instances [Info](#)
1

Software Image (AMI)
Amazon Linux 2023 AMI 2023.10...[read more](#)
ami-0532be01f26a3de55

Virtual server type (instance type)
t3.micro

Firewall (security group)
New security group

Storage (volumes)
1 volume(s) - 8 GiB

[Cancel](#) [Launch instance](#) [Preview code](#)

9. Security Hardening Recommendations

9.1 Identity and Access Management

To strengthen access control, the following IAM best practices are recommended:

- Enable Multi-Factor Authentication (MFA) for all IAM users
- Use IAM roles instead of long-term access keys
- Apply least-privilege permissions to reduce attack surface

9.2 Network Security

Network-level protections significantly reduce exposure to attacks:

- Restrict SSH access to specific IP addresses
- Replace SSH access with AWS Systems Manager Session Manager
- Enable VPC Flow Logs to monitor network traffic patterns

9.3 Container Security

Securing containers is essential in modern cloud deployments:

- Store images in Amazon ECR with vulnerability scanning enabled
- Avoid running containers with root privileges
- Regularly update base images and dependencies

9.4 Logging and Monitoring Enhancements

Further improvements to observability include:

- Creating CloudWatch Alarms for error patterns and anomalies
- Archiving logs to Amazon S3 with lifecycle policies
- Enabling Amazon GuardDuty for continuous threat detection

10. Benefits of GenAI Integration

Integrating GenAI into cloud reviews provides multiple benefits:

- Faster analysis of large and complex log datasets
- Reduced human error during manual reviews
- Improved decision-making through contextual insights
- Scalable advisory support for growing environments

GenAI acts as a security and operations sidekick, enhancing human expertise rather than replacing it.

11. Limitations of the Approach

Despite its benefits, the approach has certain limitations:

- GenAI analysis depends heavily on the quality of provided inputs
- No real-time or automated remediation is performed
- Requires skilled engineers to interpret and act on recommendations

12. Conclusion

This project successfully demonstrates an Intelligent Cloud Infrastructure Review using AWS services combined with GenAI assistance. By implementing centralized logging, secure architecture, and AI-driven analysis, the system delivers actionable security and operational insights while maintaining full human control.

The project aligns closely with enterprise-grade practices and the AWS Well-Architected Framework, proving that GenAI can significantly enhance cloud governance when used responsibly.

13. Future Enhancements

Future improvements to this project may include:

- Integrating AWS Config for continuous compliance monitoring
- Using Amazon Athena for advanced log analytics
- Adding automated alerting and incident workflows
- Extending the solution to multi-account AWS environments