iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XI **Month of publication:** November 2023

DOI: https://doi.org/10.22214/ijraset.2023.56690

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

Automated Dashboard for AWS Services Monitoring

Mr. Roopesh Kumar B N[1], Anirudha K K[2], Dhanush Y[3], Dhanvin C Bhargav[4], H A Sankeerthan[5]
[1]*Associate Professor, Dept Of Computer Science, K S Institute of Technology, Bengaluru, Karnataka*
[2, 3, 4, 5]*Dept Of Computer Science, K S Institute of Technology, Bengaluru, Karnataka*

Abstract: ***Cloud computing has become an integral part of modern business operations, offering flexibility, scalability, and cost-efficiency. However, the migration to the cloud brings forth a new set of challenges, with security being a top concern. Organizations need to continuously monitor and audit their cloud environments to identify vulnerabilities, detect threats, and ensure compliance with security standards. In today's digital landscape, cloud computing has become the backbone of many organizations, offering scalability, flexibility, and cost-effectiveness.***
However, with the increasing reliance on cloud services comes the crucial need for robust security measures to protect sensitive data and ensure compliance with industry regulation. Traditional security audits and monitoring processes are often manual, time-consuming, and prone to human error. They lack the agility and real-time insights required to effectively protect cloud assets. To address these limitations, the Automated Reporting and Dashboards for Cloud Security Audits project leverages Artificial Intelligence (AI) and Machine Learning (ML) technologies to revolutionize cloud security management.
Keywords: Cloud Computing, Auditing, Cloud Security, User-friendly Dashboard, AWS(Amazon Web Services), S3(Simple Storage Service), EC2(Elastic Compute Cloud), Amazon CloudWatch, Isolation Forest, AI(Artificial Intelligence, ML(Machine Learning), Python, API(Application Programming Interface), Microsoft Power BI

I. INTRODUCTION

This project, "Automated Reporting and Dashboards for Cloud Security Audits," seeks to revolutionize cloud security auditing by harnessing the capabilities of Artificial Intelligence (AI) and Machine Learning (ML). The primary objective is to automate the auditing process within cloud environments, alleviating the time-consuming and error-prone nature of manual audits.
This automation involves collecting data from cloud service providers through APIs, subjecting it to AI and ML analysis to detect anomalies and potential threats, and presenting the findings through user-friendly dashboards and reports. Real-time alerts and notifications are integrated to enable proactive responses to security incidents, while compliance monitoring ensures adherence to industry regulations and internal security policies. The project's ultimate aim is to enhance cloud security, reduce operational overhead, and provide organizations with actionable insights into the health of their cloud infrastructure.

II. RELATED WORKS

1) [Intelligent Role-Based Access Control Model and Framework Using Semantic Business Roles in Multi-Domain Environments]
This paper [1] introduces an access control framework, utilizing semantic business roles and intelligent agents in an Intelligent RBAC (I-RBAC) model. Occupational entitlements from real-world roles are integrated, while intelligent agents automate ontology creation. The model's efficiency is validated through implementation results in dynamic multi-domain environments.

2) [A Lightweight Identity - Based Remote Data Auditing Scheme for Cloud Storage]
The paper [2] introduces an identity-based data auditing (IBDA) scheme for secure cloud storage. The scheme utilizes data owner generated tags and data blocks, while the CSP ensures data integrity by concealing data during the challenge-proof phase, preventing TPA data theft. The proposed scheme's security is proven in the random oracle model, and efficiency analysis demonstrates its superiority over other schemes.

3) [An Efficient Data Auditing Protocol With a Novel Sampling Verification Algorithm]
The paper [3] elucidates that existing data auditing schemes, following Ateniese etal.'s framework, face challenges like repeated sampling leading to detection delays and data loss risk. This paper presents an efficient sampling verification algorithm that optimizes the scheme, enhancing data integrity in the cloud. The proposed scheme is secure, swift in detecting corrupted blocks, and offers dynamic auditing capabilities.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

4) [Privacy-Preserving Cloud Auditing for Multiple Users Scheme with Authorization and Traceability]
The paper [4] introduces a privacy-preserving cloud auditing scheme for multiple users using certificate less signature technology. It ensures user identity anonymity, collaborative traceability by managers, and prevents denial-of-service attacks. The scheme supports user revocation, maintains security without certificate management complexities, and is proven secure and efficient in analyses.

5) [A Survey on Securing Federated Learning Analysis of Applications, Attacks, Challenges, and Trends]
This paper [5] discusses Federated Learning (FL) as a privacy-preserving approach for training machine learning models. It outlines vulnerabilities impacting user privacy and model performance, presents mitigation strategies, analyzes FL applications, and highlights the role of security strategies in protecting user privacy and model performance in FL applications.

6) [Machine Learning for Cloud Security: A Systematic Review]
The paper [6] conducts a Systematic Literature Review (SLR) on the use of Machine Learning (ML) for Cloud security. The SLR covers 63 studies, highlighting Cloud security threats, ML techniques (SVM being prominent), and outcomes. Key findings include 11Cloud security categories, focus on DDoS and data privacy, model efficiency comparisons, and varied evaluation metrics. KDD and KDD CUP'99 datasets are notably popular.

III. OBJECTIVES

1) *User-Friendly Dashboard:* Create a user friendly dashboard to display security audit results, vulnerability reports, remediation progress.
2) *Continuous Monitoring:* Implement continuous monitoring to assess the security posture of AWS resources and applications regularly.
3) *AI-Driven Anomaly Detection:* Utilize machine learning or AI algorithms to detect security anomalies and suspicious activities in the AWS environment.
4) *Documentation and Reporting:* Generate comprehensive reports and documentation for security audits and compliance purposes.

IV. METHODOLOGY

A. Cloud Architecture Setup

This is the initial step where we establish the foundation for the cloud - based infrastructure. It involves launching the services that we will
monitor and setting up the necessary cloud resources, configuring security, and defining the overall structure of our architecture to support the subsequent steps. Some of the main AWS Services we are focusing on will be S3 and EC2 as they are the most popularly used services.

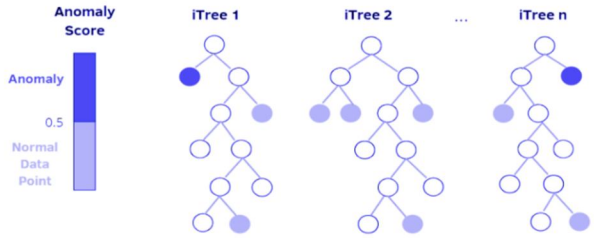
B. Setup AWS CloudWatch to Monitor Metrics

In this step, we will configure AWS CloudWatch, a monitoring service provided by Amazon Web Services, to monitor various performance and operational metrics from our AWS resources in focus. These metrics could include data related to the health and performance of our applications, servers, and other cloud resources which are running in
the services.


Amazon CloudWatch

C. Fetch the data out of AWS using CloudWatch APIs

Once AWS CloudWatch is set up, we will use its APIs (Application Programming Interfaces) to access and fetch the collected metrics and data out of AWS. This involves programmatically querying CloudWatch to extract specific information or time-series data relevant to the monitoring and analytics requirements.



Isolation Forest Anomaly Detection

D. Processing and Storing the data into a database using MongoDB

In this step, you take the data fetched from AWS CloudWatch and process it. This processing will involve data transformation, aggregation, or conversion into a more usable format, such as CSV (Comma-Separated Values). After processing, the processed data will then be stored in a MongoDB database. MongoDB is a NoSQL database management system known for its flexibility and scalability. It will definitely be an advantage to use this software.

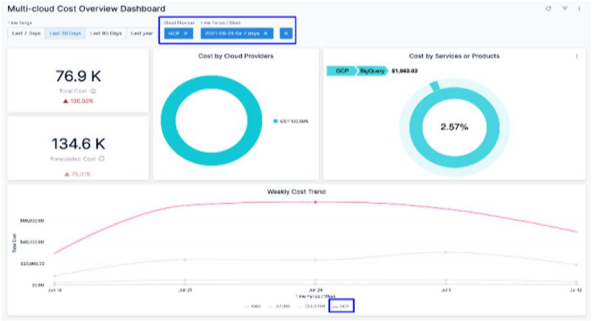
MongoDB

E. Use Isolation Forest for Anomaly Detection

Isolation Forest is a machine learning algorithm used mainly for detecting outliers or anomalies in data. It is based on the idea that outliers are more likely to be isolated from the rest of the data by random splits on the features. The algorithm builds an ensemble of binary trees, called isolation trees, that recursively partition the data until each point is isolated or reaches a maximum depth. In this step, we apply this algorithm to the data stored in MongoDB to identify unusual or anomalous data points. The metrics fetched in the previous step will act as the data-points here. The Isolation Forest works by isolating anomalies that are distant from the majority of the data, making it useful for identifying outliers and potential issues.

F. Present the Visualized Data on the Dashboard

To make the insights and results of the anomaly detection accessible and understandable, we create a data visualization dashboard. This dashboard can be implemented using tools like Power BI, or custom web-based dashboards. It displays the processed and analyzed data in a visually informative way, making it easier for users to interpret and take actions based on the detected anomalies or trends in the data.

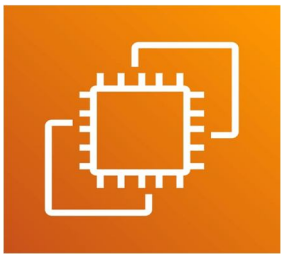


Cloud Dashboard

G. Application Requirements

1) *Amazon S3 (Simple Storage Service) and Amazon EC2 (Elastic Compute Cloud):* These are core AWS services that play distinct roles in cloud computing. Amazon S3 is a highly scalable object storage service, designed for the secure and durable storage of data and files, making it an ideal solution for storing assets like images, videos, and backups. It offers excellent durability, data replication across Availability Zones, and strong security features for access control. On the other hand, Amazon EC2 provides resizable virtual machines known as EC2 instances. It allows you to run various operating systems and software, making it a versatile choice for a wide range of computing tasks, from hosting web applications to running databases and machine learning models. EC2 instances can be easily scaled up or down to adapt to changing workloads and are under your complete control, enabling you to customize and manage your computing environment.


Amazon S3


Amazon EC2

2) *AWS CloudWatch:* AWS CloudWatch is a comprehensive monitoring and observability service designed to help users gain insights into their AWS resources and applications. It collects and stores various performance metrics, enabling users to track the health, performance, and operational state of their AWS infrastructure. CloudWatch offers features like dashboards to create custom views of metrics, alarms for automated notifications and actions, and the ability to capture and analyze log data from applications and resources. It also supports event-driven responses, allowing you to react to changes in your resources or application states. With CloudWatch, you can identify and address issues promptly, ensuring the reliability and performance of your AWS environment. It's a crucial component for maintaining the operational excellence of your AWS infrastructure and applications.
3) *Matplotlib:* Matplotlib is a popular Python library used for creating 2D and basic 3D visualizations and plots. It provides a wide range of tools for generating various types of graphs, charts, and plots, making it a valuable tool for data visualization, scientific research, and data analysis. Matplotlib allows users to customize the appearance of plots and can be used for creating static, animated, or interactive visualizations. It is widely employed by data scientists, researchers, and engineers for conveying data and insights in a graphical form.

Matplotlib

4) *Pandas:* Pandas is a powerful open-source Python library used for data manipulation and analysis. It provides easy-to-use data structures, primarily DataFrames and Series, which allow you to work with structured data efficiently. Pandas is widely used for tasks such as data cleaning, transformation, exploration, and analysis. It is an essential tool in the toolkit of data scientists, analysts, and researchers, enabling them to work with tabular data, handle missing values, filter, aggregate, and perform a wide range of operations on datasets.

Pandas

5) *NumPy:* NumPy, which stands for "Numerical Python," is a fundamental Python library for numerical and mathematical operations. It provides support for working with large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays. NumPy is a crucial component of the Python data science and scientific computing ecosystem.

NumPy

6) *Jupyter Notebook:* Jupyter Notebook is an open-source, web-based interactive computing environment that allows users to create and share documents that contain live code, equations, visualizations, and narrative text. It is particularly popular in data science and scientific research. Jupyter Notebooks support various programming languages, with Python being one of the most commonly used. Users can write and execute code in a cell-by-cell manner, making it easy to experiment, visualize data, and document their work. Jupyter Notebook is widely used for tasks such as data analysis, machine learning, data visualization, and collaborative research, as it provides an interactive and reproducible platform for working with code and data.
7) *Seaborn:* Seaborn is a Python data visualization library based on Matplotlib that provides a high-level interface for creating informative and attractive statistical graphics. It is designed to work seamlessly with Pandas DataFrames and simplifies the process of creating complex, aesthetically pleasing visualizations for data analysis and exploration.

Seaborn

V. ACKNOWLEDGEMENT

We would like to express our deep gratitude to Mr. ROOPESH KUMAR B N, Associate Professor, Department of Computer Science and Engineering, KSIT for his valuable and constructive suggestions during the planning and development of this project. His willingness to give his time so generously has been very much appreciated. We would also like to thank all the professors of KSIT for their continuous support and encouragement.

REFERENCES

[1] RUBINA GHAZAL, AHMAD KAMRAN MALIK, NAUMAN QADEER, BASIT RAZA , AHMAD RAZA SHAHID, HANI ALQUHAYZ “Intelligent Role-Based Access Control Model And Framework Using Semantic Business Roles In Multi-Domain environments”

- COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan
- University Institute of Information Technology, Pir Maher Ali Shah (PMAS) Arid Agriculture University, Rawalpindi 46300, Pakistan
- Department of Computer Science, Federal Urdu University of Arts, Science, and Technology at Islamabad, Islamabad 44080, Pakistan

- Department of Computer Science and Information, College of Science Al-Zulfi, Majmaah University, Al Majmaah 11952, Saudi Arabia - January 9, 2020

https://ieeexplore.ieee.org/Xplore/home.jsp

[2] LUNZHI DENG, BENJUAN YANG, AND XIANGBIN WANGA "A Lightweight Identity-Based Remote Data Auditing Scheme for Cloud Storage"

- School of Mathematical Sciences, Guizhou Normal University, Guiyang 550001, China
- College of Computer Science and Technology, Guizhou University, Guiyang 550025, China
- School of Big Data and Computer Science,
  Guizhou Normal University, Guiyang 550001, China - November 7, 2020

https://ieeexplore.ieee.org/Xplore/home.jsp

[3] XUELIAN LI,LISHA CHEN ,AND JUNTAO GAO, "An Efficient Data Auditing Protocol With a Novel Sampling Verification Algorithm"

- School of Mathematics and Statistics, Xidian University, Xi'an, Shaanxi 710071, China
- Guangxi Key Laboratory of Cryptography
  and Information Security, School of Telecommunications Engineering, Xidian University, Xi'an, Shaanxi 710071, China - July 2, 2021

https://ieeexplore.ieee.org/Xplore/home.jsp

[4] XIAODONG YANG, (Member, IEEE), MEIDING WANG, TING LI , RUI LIU1, AND CAIFEN WANG, "Privacy-Preserving Cloud Auditing for Multiple Users Scheme With Authorization and Traceability"

- College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, China
- College of Big Data and Internet, Shenzhen Technology University, Shenzhen 518118, China - July 15, 2020

https://ieeexplore.ieee.org/Xplore/home.jsp

[5] HELIO N. CUNHA NETO , JERNEJ HRIBAR2, IVANA DUSPARIC ,
DIOGO MENEZES FERRAZANI MATTOS, AND NATALIA C. FERNANDES, "A Survey on Securing Federated Learning: Analysis of Applications, Attacks, Challenges, and Trends "

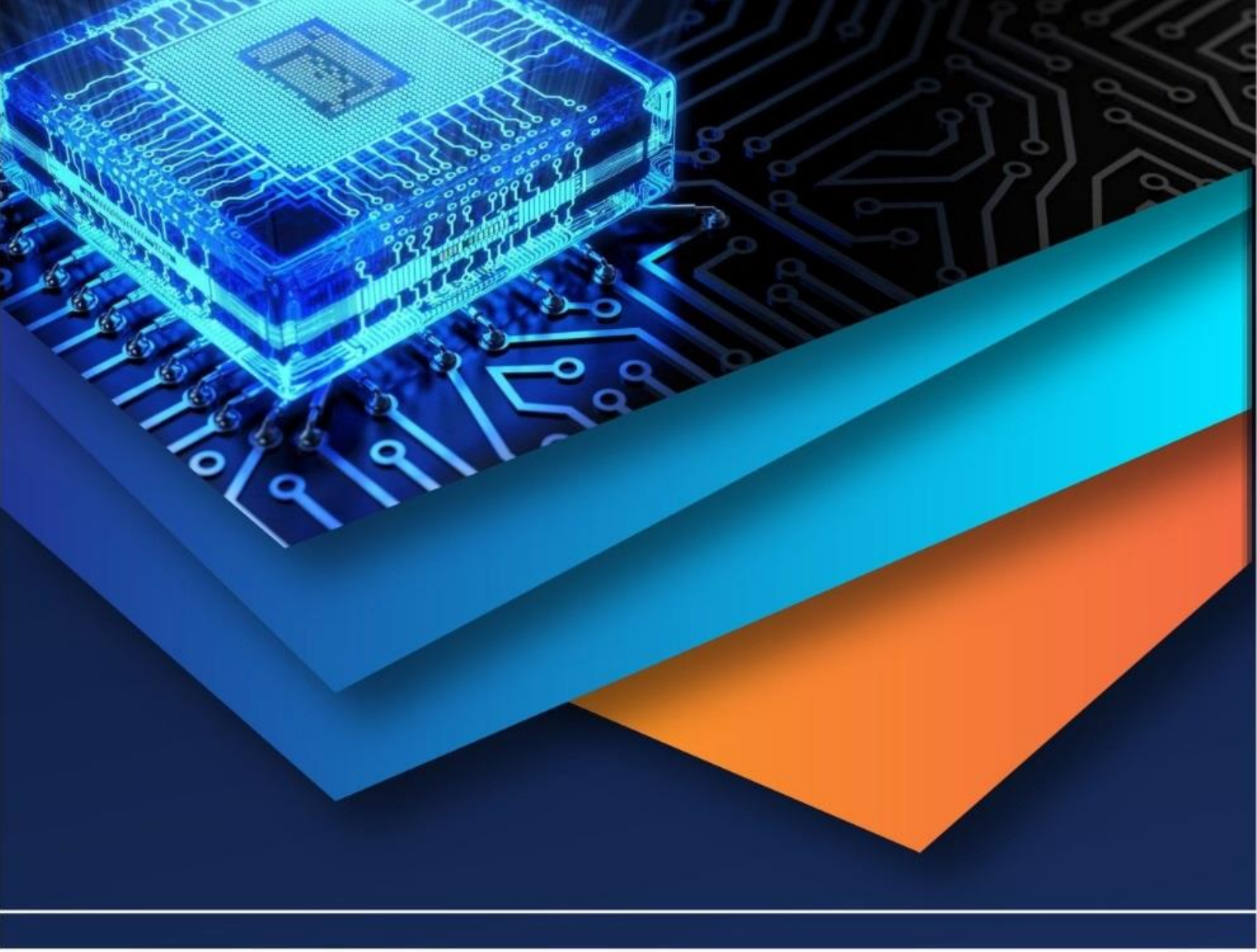
- MídiaCom, PPGEET, Universidade Federal Fluminense (UFF), Niterói 24210-240, Brazil,
- Department for Communication Systems, Jožef Stefan Institute, 1000 Ljubljana, Slovenia
- School of Computer Science, Trinity College Dublin, Dublin 2, D02 PN40 Ireland - 24 April 2023

https://ieeexplore.ieee.org/Xplore/home.jsp

[6] ALI BOU NASSIF , MANAR ABU TALIB, QASSIM NASIR, HALAH ALBADANI, AND FATIMA MOHAMAD DAKALBAB "Machine Learning for Cloud Security: A Systematic Review "

- Department of Computer Engineering, University of Sharjah, Sharjah, United Arab Emirates
- Department of Computer Science, University of Sharjah, Sharjah, United Arab Emirates
- Department of Electrical Engineering, University of Sharjah, Sharjah, United Arab Emirates - January 25, 2021

https://ieeexplore.ieee.org/Xplore/home.jsp



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY