IJARSCT

ISSN (Online) 2581-9429

IJARSCT

Impact Factor: 7.53

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 4, March 2024

Real-Time AWS Resource Surveillance with Reporting and Dashboard Solution

Nuthana M[1] and Dr M N Nachappa[2]

PG Student, Department of MSc CS-IT[1]

Professor, School of CS & IT[2]

Jain (Deemed-to-be University), Bangalore, India

shettynuthana9@gmail.com[1] and mn.nachappa@jainuniversity.ac.in[2]

Abstract*:* *Cloud computing, which offers cost-effectiveness, scalability, and flexibility, has become a crucial component of contemporary company operations. However, moving to the cloud presents a new set of difficulties, one of which is security. Continuous monitoring and auditing of cloud environments is critical for organizations to detect attacks, identify vulnerabilities, and verify adherence to security requirements. Because it is scalable, flexible, and affordable, cloud computing has emerged as the mainstay of many businesses in today's digital environment.*

But as the use of cloud services grows, so does the necessity for strong security measures to safeguard confidential information and guarantee adherence to industry rules. Conventional methods for security audits and monitoring are frequently labor-intensive, manual, and prone to human mistake. They are not agile enough or have the real-time information necessary to efficiently.

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

Cloud computing has become an integral part of modern business operations, offering flexibility, scalability, and cost-efficiency. However, moving to the cloud presents a new set of difficulties, one of which is security. Companies must constantly monitor and audit their cloud systems to find security flaws, identify risks, and make sure security requirements are being followed. Because it is scalable, flexible, and affordable, cloud computing has emerged as the mainstay of many businesses in today's digital environment. But as the use of cloud services grows, so does the necessity for strong security measures to safeguard confidential information and guarantee adherence to industry rules. Conventional methods for security audits and monitoring are frequently labor-intensive, manual, and prone to human mistake.

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.

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 errorprone 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.

DOI: 10.48175/IJARSCT-15936

II. LITERATURE SURVEY

[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 multidomain environments.

[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.

[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.

[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 denialof-service attacks. The scheme supports user revocation, maintains security without certificate management complexities, and is proven secure and efficient in analyses.

[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.

[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

User-Friendly Dashboard: Create a user friendly dashboard to display security audit results, vulnerability reports, remediation progress.

Continuous Monitoring: Implement continuous monitoring to assess the security posture of AWS resources and applications regularly.

AI-Driven Anomaly Detection: Utilize machine learning or AI algorithms to detect security anomalies and suspicious activities in the AWS environment.

Documentation and Reporting: Generate comprehensive reports and documentation for security audits and compliance purposes.

IV. METHODOLOGY

Cloud Architecture Setup:

This is the initial stage of creating the foundation for the cloud-based system. It involves launching the services that we will monitor, configuring security, putting up the necessary cloud resources, and creating the overall architecture structure in order to support the next steps. S3 and EC2 are two of the most popular AWS services, thus those will be the services we focus on initially.

Setup AWS CloudWatch to monitor metrics:

This phase involves configuring Amazon Web Services' CloudWatch monitoring service to track a variety of operational and performance indicators from the AWS resources we are specifically interested in. These metrics may contain information about the functionality and state of our servers, apps, and other cloud resources that are used by the services.

Amazon CloudWatch

Fetch the data out of AWS using CloudWatch APIs:

We will access and retrieve the gathered metrics and data from AWS using its APIs (Application Programming Interfaces) after AWS CloudWatch is configured. In order to retrieve certain data or time-series data pertinent to the monitoring and analytics requirements, this entails programmatically querying CloudWatch.

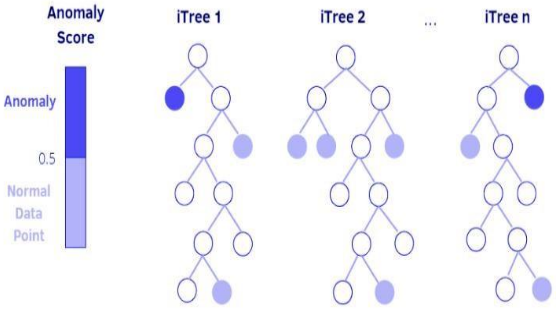
Processing and Storing the data into a database using MongoDB:

Processing the data that was obtained from AWS CloudWatch is the focus of this step. This process will include data processing, aggregation, and translation into a more understandable format, like CSV (Comma-Separated Values). After processing, the data will be stored in a MongoDB database. MongoDB is a well-known NoSQL database management system that is highly versatile and scalable. Utilizing this program will prove advantageous.

MongoDB

.Use Isolation Forest for anomaly detection:

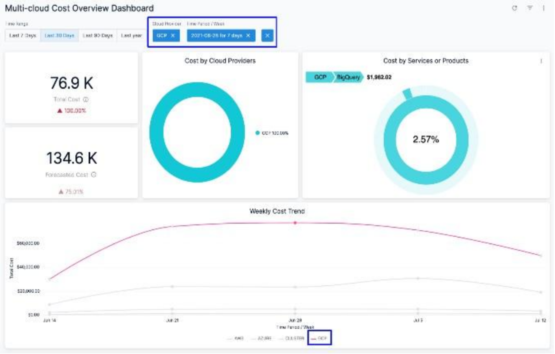
The main purpose of the machine learning technique known as Isolation Forest is to identify anomalies or outliers in data. It is based on the idea that feature-based random splits are more likely to isolate outliers from the rest of the data. The procedure divides the data recursively until each point is either isolated or reaches a predetermined depth by generating a collection of binary trees known as isolation trees. We apply this technique on the MongoDB data to identify unusual or anomalous data points. In this instance, the metrics that were obtained in the earlier phase will serve as the data points. Since the Isolation Forest separates anomalies that are distant from the norm, it might be useful for identifying outliers and potential anomalies.



Isolation Forest Anomaly Detection

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.

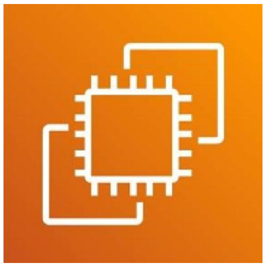


Cloud Dashboard

V. APPLICATION REQUIREMENTS

Amazon S3 (Simple Storage Service) and Amazon EC2 (Elastic Compute Cloud) : These are core AWS services that serve particular cloud computing purposes. Amazon S3 is an object storage service that is ideal for storing media such as images, videos, and backups since it is extremely scalable. It is designed to preserve files and data securely over an extended period of time. It offers great durability, data replication across Availability Zones, and strong security capabilities for access control. On the other hand, resizable virtual machines known as EC2 instances are provided by Amazon EC2. Because it can run many operating systems and applications, it is a versatile alternative for a range of computing workloads, from hosting web apps to running databases and machine learning models. In order to adjust to shifting workloads, EC2 instances are easily scaled up or down and are under

Amazon S3



Amazon EC2

DOI: 10.48175/IJARSCT-15936

AWS CloudWatch : Amazon CloudWatch is a full-featured observability and monitoring service that helps users understand their AWS apps and resources. It allows users to monitor the functionality, performance, and overall state of their AWS infrastructure by gathering and storing a variety of performance indicators. A few of the capabilities that CloudWatch provides are the ability to capture and analyze log data from applications and resources, alarms for automated notifications and actions, and dashboards for creating personalized views of metrics. Additionally, it offers event-driven replies, which let you respond when your resources or application states change. By using CloudWatch, you can quickly detect and resolve problems, maintaining the dependability and efficiency of your AWS environment. It's an essential part of keeping your AWS apps and infrastructure operating at peak performance.

Matplotlib A well-liked Python package called Matplotlib is used to create simple 2D and rudimentary 3D plots and visualizations. It is an invaluable tool for data visualization, scientific research, and data analysis since it offers a broad range of capabilities for creating different kinds of graphs, charts, and plots. Matplotlib can be used to create static, animated, or interactive visualizations, and it also lets users change the look of plots. It is widely employed by data scientists, researchers, and engineers for conveying data and insights in a graphical form.

Matplotlib

Pandas : An effective open-source Python library for data analysis and manipulation is called Pandas. It offers user-friendly data structures that make working with structured data easier, mainly DataFrames and Series. Pandas is a popular tool for activities including exploration, analysis, transformation, and cleansing of data. It is a crucial tool for researchers, analysts, and data scientists to have in their toolbox since it allows them to work with tabular data, manage missing values, filter, aggregate, and execute a variety of operations on datasets.

Pandas

NumPy :"Numerical Python," or NumPy, is a core Python library for mathematical and numerical operations. Large, multidimensional arrays and matrices can be worked with, and a number of mathematical functions can be applied to these arrays. An essential part of the Python scientific computing and data science ecosystem is NumPy.

NumPy

Jupyter Notebook : With Jupyter Notebook, an open-source, web-based interactive computing environment, users may create and share documents that include narrative prose, equations, pictures, and real-time code. In data science and

ISSN
2581-9429
IJARSCT

scientific research, it is highly favored. One of the most popular programming languages that Jupyter Notebooks supports is Python. Write code, execute it cell by cell, and experiment, display data, and publish discoveries with ease. Jupyter Notebook is widely used for tasks like data analysis, machine learning, data visualization, and collaborative research because it provides an interactive and reproducible environment for working with code and data.

Seaborn: Based on Matplotlib, Seaborn is a Python data visualization package that offers an elegant and user-friendly high-level interface for producing statistical visuals. Its seamless integration with Pandas DataFrames design makes it easier to create intricate, aesthetically beautiful visuals for data exploration and analysis.

Seaborn

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

[2] 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