COMP3010HK Security Operations and Incident Management 2025/2026

Coursework 2: BOTSv3 Incident Analysis and Presentation

Student ID: 10953435

Github repo: https://github.com/aka331/COMP3010HK/tree/main/CW2

YouTube: https://www.youtube.com/watch?v=Knpgs-4_tVI

Word count: 3243

Table of Contents

1.0 Introduction

This report will explore what AWS and endpoint-related activity looks like in the BOTSv3 dataset using Splunk, in addition to answering 200-level guiding questions with the information provided by answering these questions, which will then give us additional detection logic and event handling insights. The current scope is limited to potential IAM misuse with the Frothly environment, potential S3 bucket issues, and potential issues with critical endpoints. Other BOTSv3 scenarios and non-AWS sources are out of scope.

The analysis takes the data provided and assumes it is complete and accurate, and the timestamps are of the correct format and timezone, as provided. No other validation takes place. In performing the analysis, only the data provided by BOTSv3 and the Search Processing Language (SPL) provided by the course are used. No modifications are made to the environment provided in the Frothly simulation for this investigation. Further subsections discuss the role of the SOC and the reflections on handling incidents, the implementation of Splunk, questions provided, and the overall findings and recommendations for improvement.

A Security Operations Center (SOC) is responsible for continuous monitoring, detection, and response to security threats across an organization's systems. In practice, SOC teams focus on three core functions:
(1) tracking activity across networks, servers, applications, and endpoints
(2) investigating and responding to suspicious events
(3) supporting compliance while improving the organization's security posture.

Depending on scale and operational needs, SOC functions may be handled by a single internal team, distributed global teams (GSOCs), or outsourced providers.

2.0 SOC Roles & Incident Handling Reflection

2.1 SOC Roles

In the SOC structure, it has 3 tiers of security analysts and operators. They have different responsibilities and technical skills in the monitoring, analysis, and response to security incidents. In the BOTSv3 scenario, it can be regarded as simulating a complete incident in which an endpoint and cloud resources are compromised, allowing analysts at different levels to carry out their roles across the four phases of prevention, detection, response, and recovery.

SOC Role	Responsibilities	Relevance to the BOTSv3 exercise
Tier 1 (T1) — Monitoring & Triage	Continuous alert monitoring, review the initial alert, and alert escalation if needed.	Monitor the alert from the Splunk dashboard. Use an SPL query to search CloudTrail logs.
Tier 2 (T2) — Investigation & Scoping	In-depth investigation, event correlation, attack scoping, timeline reconstruction, and incident classification.	Investigate the S3 bucket misconfiguration, unauthorized access, and the change of Access Control List.
Tier 3 (T3) — Threat Hunting & Detection Engineering	Threat hunting, detection rule tuning, playbook development, advanced analysis, and post-incident improvement.	Active reverse engineering to search attacker that perform network forensics investigation.
SOC Manager / SOC Lead	Incident coordination, risk assessment, escalation decisions, communication, and strategic oversight.	Oversees investigation severity, ensures correct escalation, and uses BOTSv3 outcomes to inform SOC strategy and policy improvements.

Table 1. SOC role in BOTSv3

2.1.1 Tier 1 (T1) — Monitoring & Triage

For Tier-1, usually referred to as a “front-tier” or “junior” security analyst who responds to 24x7 monitoring of alerts and performs initial triage using Splunk. For example, when the T1 operator validates the alert from Splunk Triggered Alerts, they could determine and review BOTSv3’s basic enrichment**,** such as AWS CloudTrail, the hostname, source IP, logs, timestamps, affected asset, etc. Once reviewed, they could take action (runbooks or playbooks), such as blocking the intruder's IP address, escalating the alert/issue to Tier-2

security analysis for further investigation and handling malware alerts or account hacking attempts. After that, they were required to generate the report for the user or T2 analysts for notification and record.

2.1.2 Tier 2 (T2) — Investigation & Scoping

For Tier-2, usually referred to as "second-line" or "senior" security analysts who have deeper technical knowledge and experience. In a real SOC, they typically understand the organization's environment and what "normal" looks like, which helps them separate true threats from false positives. They respond to more complex and in-depth security incident investigations. In the BOTSv3 exercise, Tier-2 would build on the Tier-1 handover by running deeper SPL searches, tuning or developing new queries, and correlating multiple data sources (e.g., cloud, authentication, and endpoint logs). Their main tasks are to reconstruct the attack path, validate whether suspicious activity is actually malicious, and determine the full scope of impact—who was affected, what resources were accessed, and how far the activity spread. They also document findings in a structured investigation report, including evidence, timeline, and recommended containment steps.

2.1.3 Tier 3 (T3) — Threat Hunting & Detection Engineering

Tier-3 analysts are the "third-line" experts in a SOC. They handle the most complex, high-impact incidents and provide deep technical capability in areas such as threat hunting, advanced log analysis, incident response support, malware analysis, and digital forensics. In many organizations, Tier-3 also plays a specialist role when evidence must be preserved and explained rigorously—especially in cases that may involve legal review—where a qualified forensic expert may be required to justify methodology, interpret technical findings, and clearly present evidence.

In the BOTSv3 scenario, Tier-3's focus is less about answering individual alerts and more about long-term improvement. They translate investigation results into durable, reusable capabilities: building and tuning correlation searches, defining high-fidelity detection logic, and developing baselines for "normal" CloudTrail and endpoint behavior. For example, they can extract repeatable IAM misuse patterns, risky S3 configuration indicators, and endpoint

anomaly baselines from the dataset, then package these as detection rules and playbook guidance for Tier-1 and Tier-2 to use in daily operations. This turns one investigation into sustained SOC maturity: better alert quality, faster triage, and fewer repeat incidents.

2.1.4 SOC Manager / SOC Lead

The position of the SOC department head coordinates the overall response, manages communication, and makes risk-based decisions during an incident. In the BOTSv3 context, they confirm severity and likely business impact, assign ownership to the right teams (e.g., cloud or endpoint), and ensure containment actions—such as disabling IAM keys or isolating endpoints—follow organizational policy and minimize operational disruption.

2.2 Incident Handling Reflection

A computer security incident refers to any violation or impending violation of computer security policies, acceptable use policies, or standard security practices. For example, an employee might be tricked into opening a "quarterly report" sent via email that is actually malware; running the program infects the employee's computer, establishes a connection to an external host, and spreads to other computers.

In this stage, Incident Handling Reflection is an important post-incident process where a security operation team reviews an event to evaluate their performance and identify improvements. One of the best practices is to establish an incident planning and response framework based on the industry standard NIST SP 800-61 Incident Handling framework, which emphasises continuous improvement across the incident lifecycle.

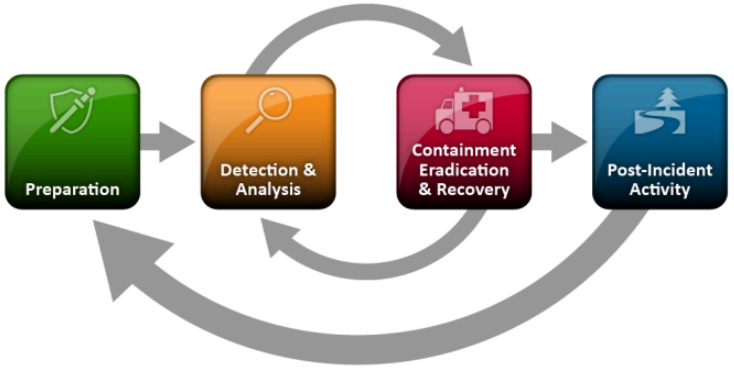

Diagram 1. Previous incident response life cycle model

Phase (NIST Phase)	Apply to the BOTSv3 Investigation
Prevention (Preparation)	During the preparation phase, SOC ensures that appropriate logging, monitoring, and access control mechanisms are in place before an incident. In the context of the BOTSv3 scenario, appropriate logging is ensured via Cloudtrail, Splunk dashboards, IAM role policy, MFA, and asset inventories. Baseline configuration of bucket permissions for S3 and hardening of the endpoint could have avoided the public ACL misconfiguration.
Detection (Detection and Analysis)	Suspicious events are identified and validated during detection and analysis. In BOTSv3, this ranged from detecting abnormal PutBucketAcl events to identifying API calls performed without MFA and also observing unusual endpoint behavior, such as BSTOLL-L running Windows 10 Enterprise. For determining if the activity was malicious and the extent of it, IAM activity, CloudTrail logs, and endpoint telemetry were correlated using Splunk SPL queries.
Response (Containment)	Containment actions are designed to limit the effects of an attack well and prevent further exploitation. In the case of BOTSv3, this would involve immediately reverting the S3 bucket ACL to private, disabling or rotating compromised IAM credentials, enforcing MFA policies, and isolating the affected endpoint BSTOLL-L. Rapid containment reduces Time-to-Contain (TTC), minimizing data exposure risk.
Recovery (Eradication and Recovery)	During the recovery phase, the organization restores systems back to a secure end-state. For BOTSv3, that means verifying that the S3 bucket is no longer publicly accessible, along with IAM policy integrity, scanning endpoints for persistence mechanisms, and confirming no further unauthorized uploads have taken place. Lessons learned are then implemented as improved detection rules and hardened configurations to prevent reoccurrences.

Table 2. Incident Handling by Phase

3.0 Installation & Data Preparation

3.1 Environment Setup

The SOC lab environment was deployed in a virtualized environment using VMware Workstation / Oracle VirtualBox, hosting a Linux virtual machine (VM), such as Ubuntu / CentOS / Red Hat Enterprise Linux (RHEL). Virtualization provides isolation and reproducibility, which is aligned with SOC lab practices for safely testing datasets and repeated investigations.

Splunk Enterprise (Linux x86_64) was installed under /opt/splunk following common Linux deployment conventions. The installation was performed by downloading the Splunk tarball, extracting it into /opt, and starting Splunk with --accept-license to automate the license acceptance step.

```
wget -O splunk-10.0.1.tgz
"https://download.splunk.com/products/splunk/releases/10.0.1/linux
/splunk-10.0.1-c486717c322b-linux-amd64.tgz"
sudo tar -xzf splunk-10.0.1.tgz -C /opt
sudo /opt/splunk/bin/splunk start --accept-license
```

Bash command 1: Download and install (extract) Splunk

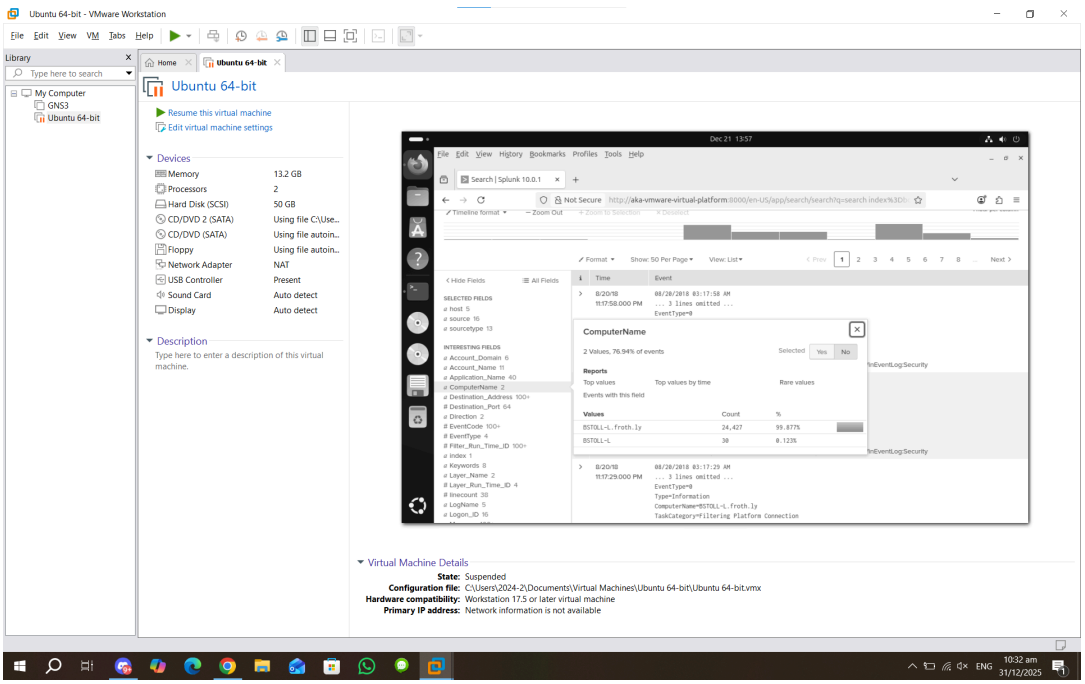

Fig 1. Install the Ubuntu environment in VMware

3.2 Dataset Ingestion

After Splunk was successfully installed, the BOTSv3 dataset was downloaded and ingested by extracting the dataset package into Splunk's apps directory. BOTSv3 is packaged as a Splunk app containing the required data configurations, dashboards, and supporting objects for SOC-style investigations.

Dataset source: BOTSv3 dataset archive (downloaded separately):
https://botsdataset.s3.amazonaws.com/botsv3/botsv3_data_set.tgz

```
cd ~/Downloads
wget -O botsv3_data_set.tgz
"https://botsdataset.s3.amazonaws.com/botsv3/botsv3_data_set.tgz"
```

Bash command 2: Download the dataset

```
sudo tar zxvf botsv3_data_set.tgz -C /opt/splunk/etc/apps/
```

Bash command 3: Extract BOTSv3 dataset into Splunk apps path

After extraction, Splunk was restarted to ensure the app and knowledge objects were properly loaded.

```
sudo /opt/splunk/bin/splunk restart
```

Bash command 4: Restart Splunk to load the BOTSv3 app

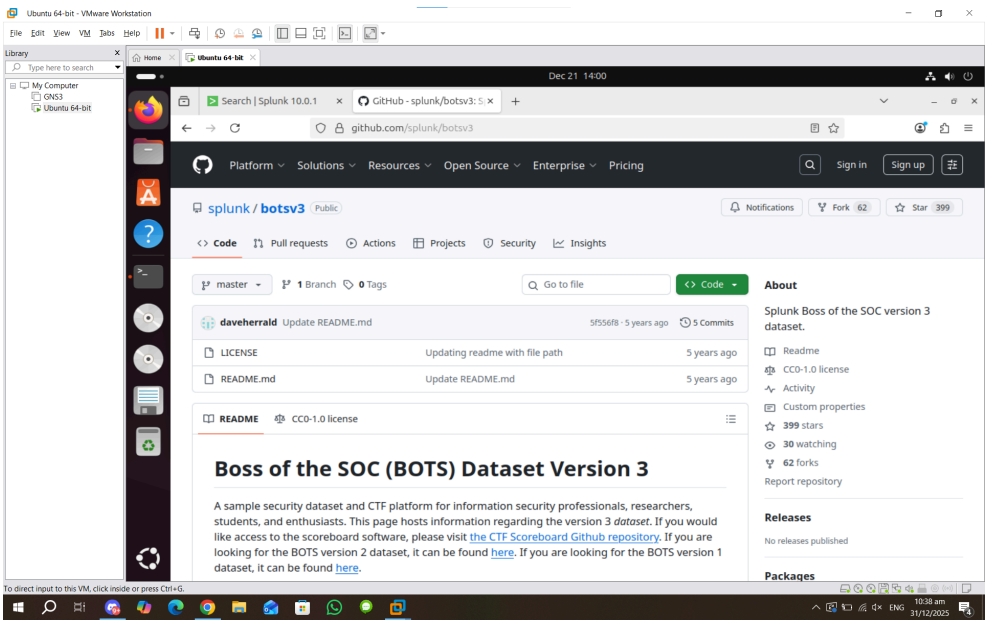

Fig 2. BOTSv3 Github repo

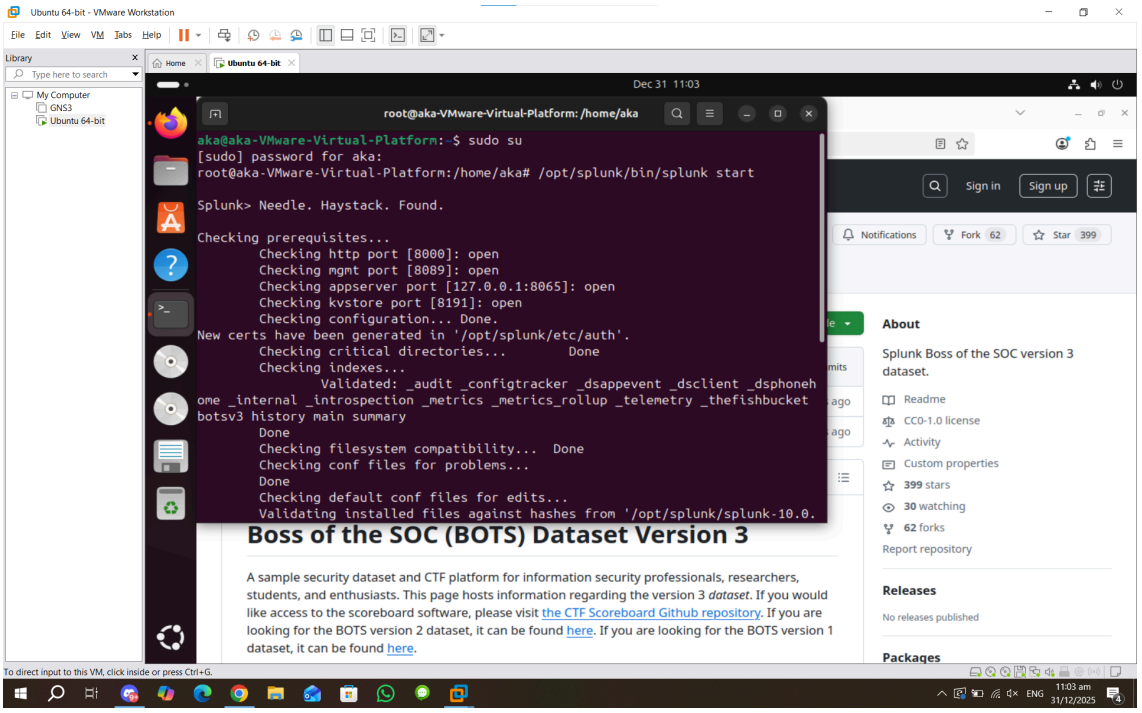

Fig 3. Start the Splunk environment in the terminal (1)

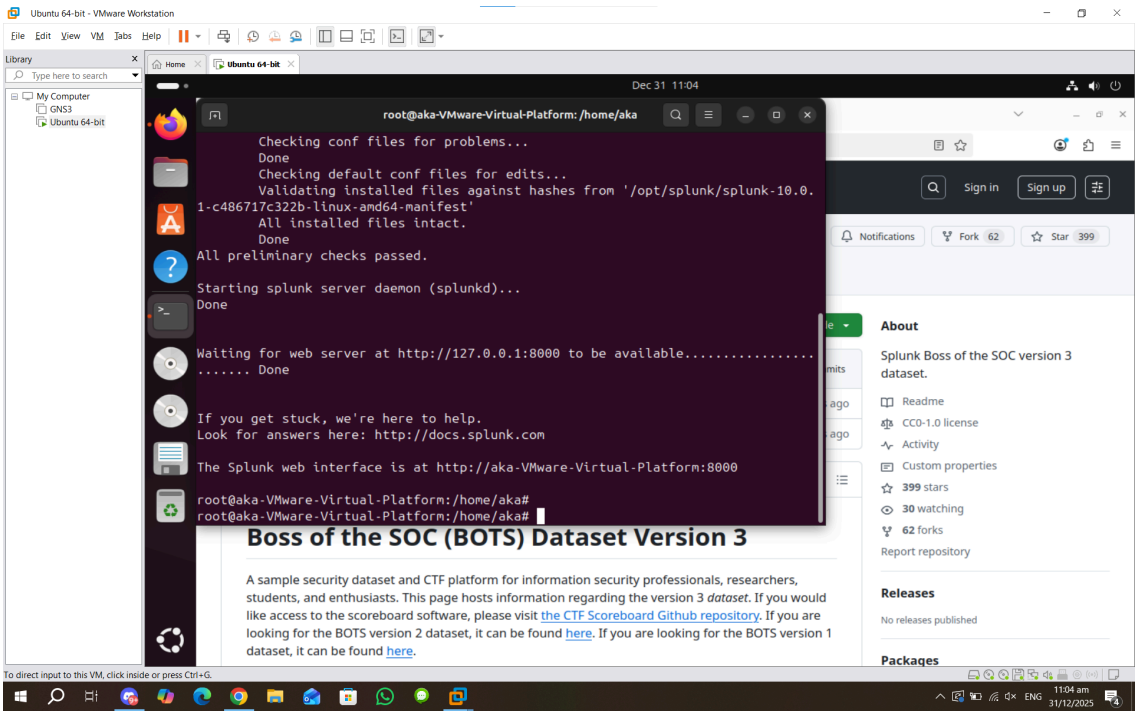

Fig 4. Start the Splunk environment in the terminal (2)

Finally, Splunk Web can be accessed via:

Inside the VM: http://127.0.0.1:8000

From host machine (if needed): http://<VM_IP>:8000 (depends on NAT/bridged setup)

To validate that the BOTSv3 dataset was successfully ingested and searchable, the following SPL query was executed in the Splunk Search interface:

```
index=botsv3 earliest=0
```

SPL 1: Validate BOTSv3 data availability (full time range)

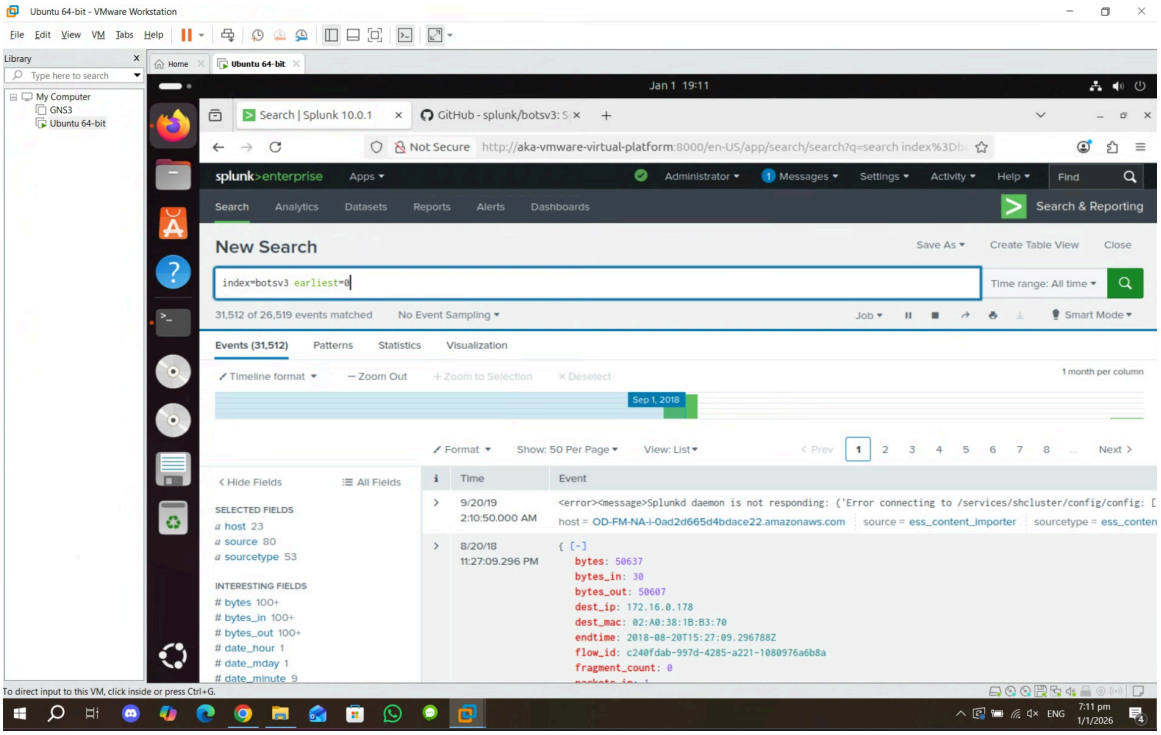

Fig 5. Screenshot of validating the BOTSv3 data availability

4.0 Investigation

This section details the investigation of unauthorized IAM activity and S3 bucket compromise within the Frothly AWS environment. The analysis uses Splunk to identify the attacker, the method of compromise, and the extent of data exposure.

4.1 Key Findings

Question	Finding / Answer	SOC Significance
Q1	Active IAM Users: bstoll, btun, splunk_access, web_admin	Identity Baselining: Distinguishes human users from service accounts to tune anomaly detection.
Q2	MFA Alert Field: userIdentity.sessionContext.attributes.mfaAuthenticated	API Security: reliably detects MFA bypass on programmatic API calls, unlike console logs.
Q3	Web Server CPU: E5-2676	Asset Inventory: Identifies hardware-level vulnerabilities.
Q4	Public Access Event ID: ab45689d-69cd-41e7-8705-5350402cf7ac	Incident Timeline: Pinpoints the exact moment of exposure for accurate Time-to-Detect (TTD) metrics.
Q5	Bud's username: bstoll	Attribution: Links the security breach to a specific identity for insider threat investigation.
Q6	S3 bucket: frothlywebcode	Containment: Identifies the specific compromised asset for surgical remediation (ACL rollback).
Q7	text file that was uploaded: OPEN_BUCKET_PLEASE_FIX.txt	Impact Assessment: Confirms unauthorized write access (integrity loss), not just data leakage.
Q8	FQDN: BSTOLL-L.froth.ly	Outlier Detection: Highlights non-standard endpoints that may lack server-grade security controls.

4.2 Splunk Analysis & Evidence

4.2.1 IAM User Enumeration

Objective: Identify all IAM users interacting with AWS services.

Query:

```
index=botsv3 earliest=0 sourcetype=aws:cloudtrail
```

SPL 2. Filter the information from the AWS CloudTrail log

```
index=botsv3 earliest=0 sourcetype=aws:cloudtrail | stats count by
userIdentity.userName
```

SPL 3. View statistics about userIdentity.userName in AWS CloudTrail log

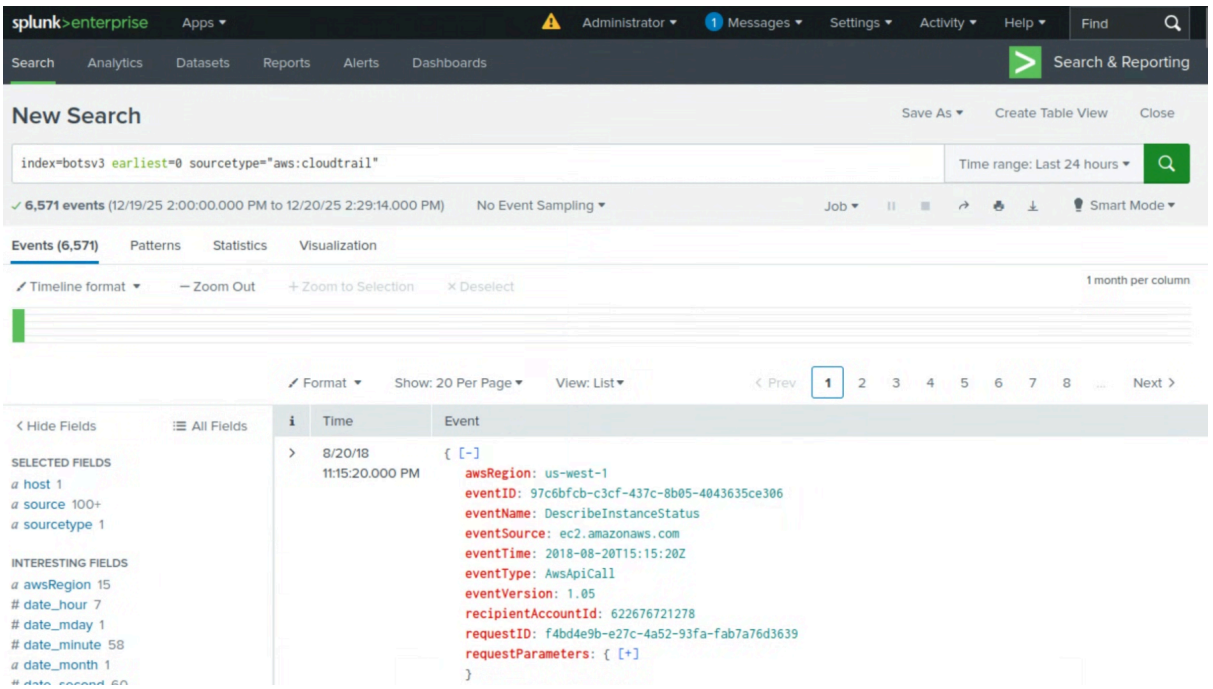

Fig 1. Capture to filter the information from the AWS CloudTrail log

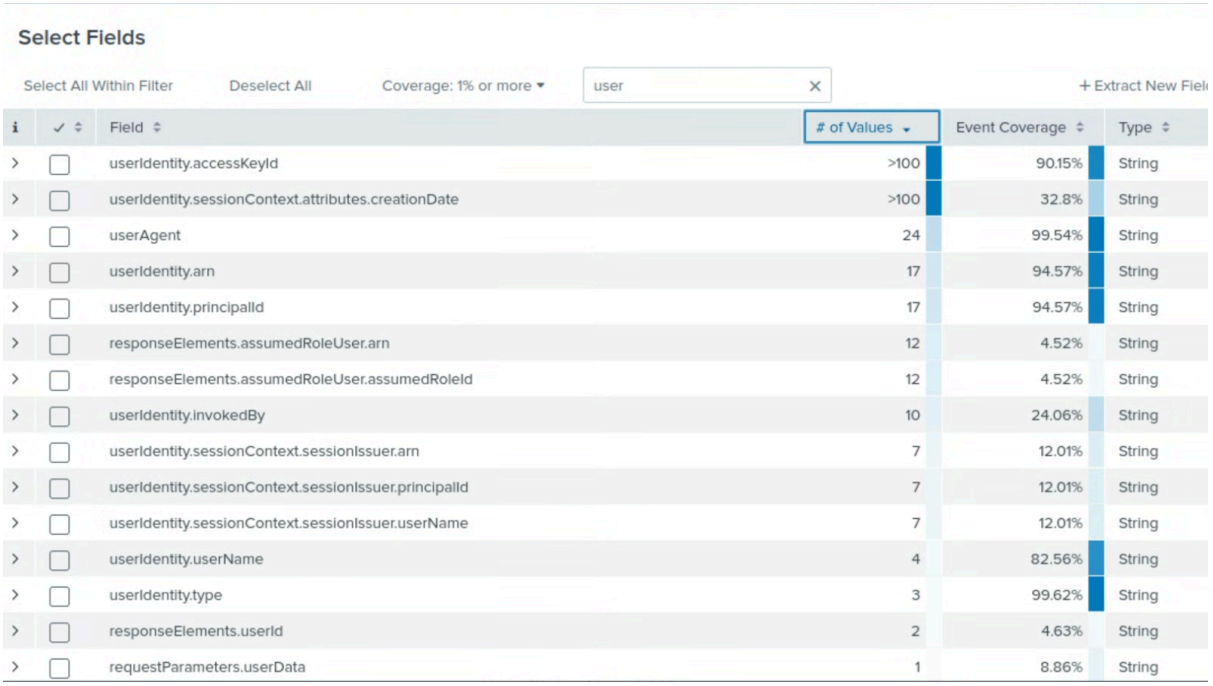

Fig 2. View the available field in the AWS CloudTrail log

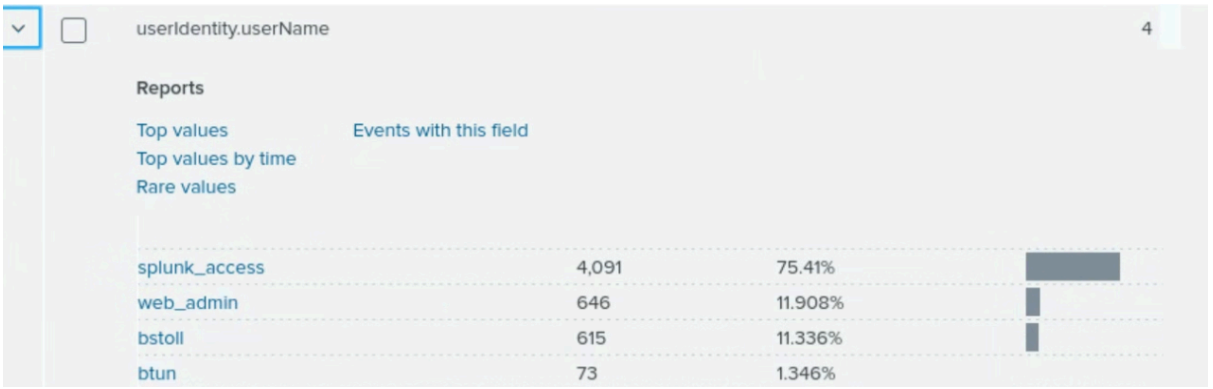

Fig 3. Show the IAM username

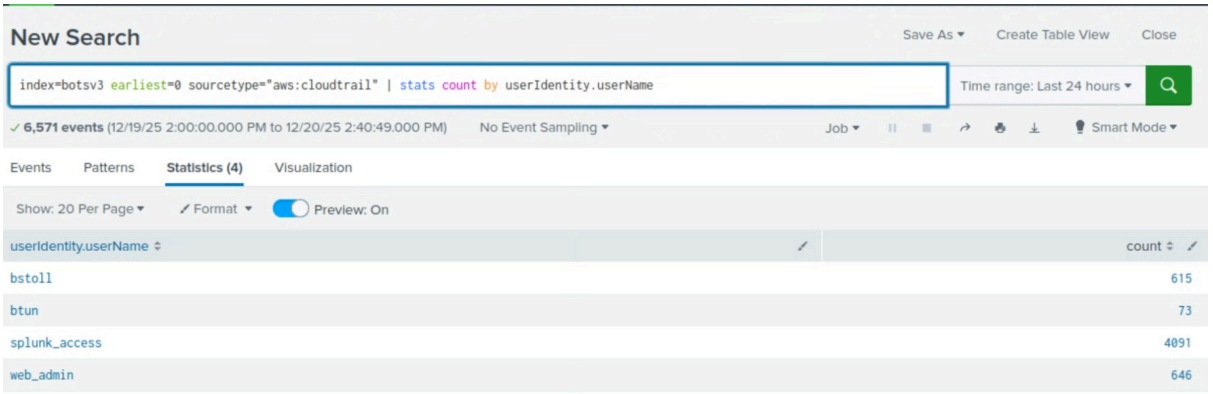

Fig 4. List the IAM username

In the investigation, it reviews the source type based on the hint (sourcetype:"aws:cloudtrail"). Using the "user" in the search field to filter information related to the username of IAM. Then, it found the list of IAM users that accessed an AWS service, which is bstoll, btun, splunk_access, web_admin.

Result:

bstoll, btun, splunk_access, web_admin

4.2.2 MFA Alerting Logic

Objective: Identify the correct field to detect API activities performed without MFA.

Query:

```
index=botsv3 earliest=0 sourcetype=aws:cloudtrail | regex "MFA"
```

SPL 4. View statistics that contain the word "MFA" in AWS CloudTrail log

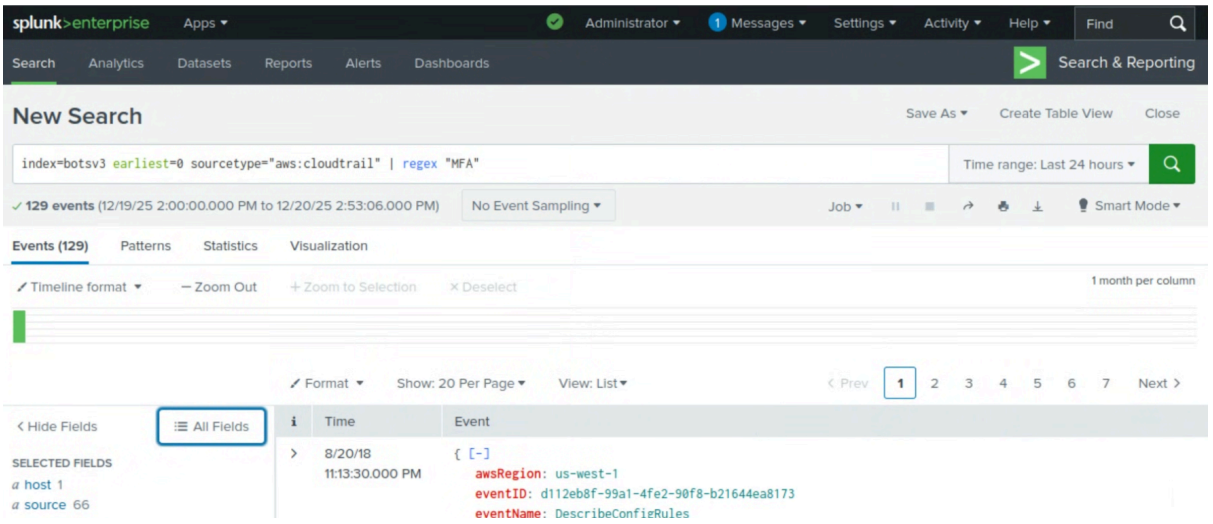

Fig 5. Capture of statistics contains the word "MFA" in the AWS CloudTrail log

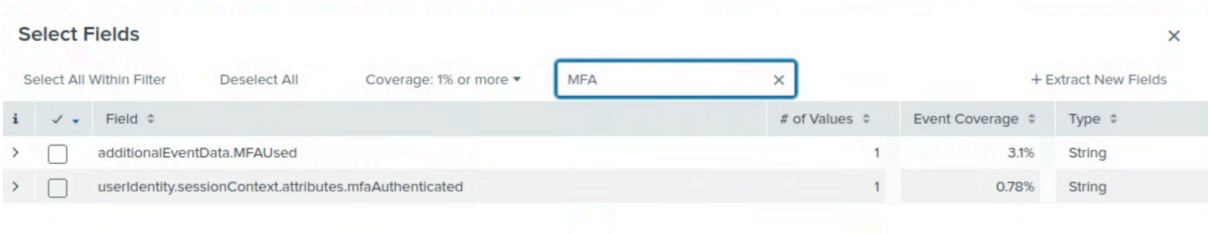

Fig 6. Search for the MFA in the select field option

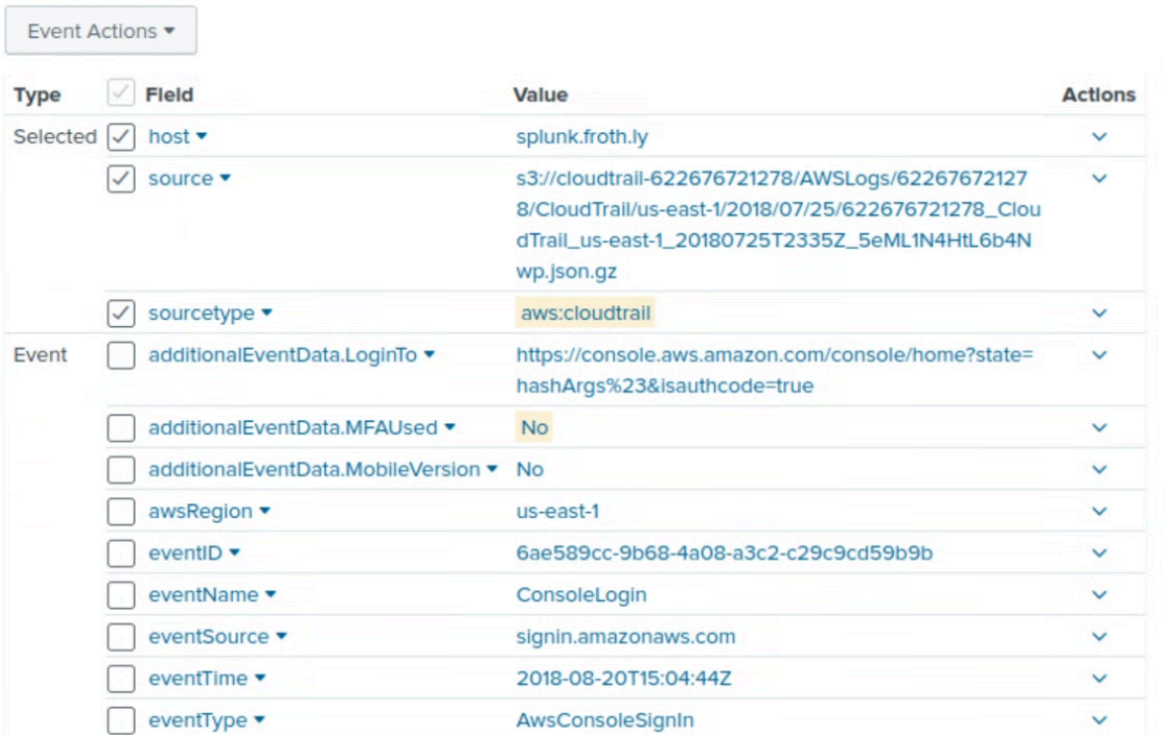

Fig 7. The log contain MFA field/attribute

During the research, it displays 2 fields: “additionalEventData.MFAUsed” and “userIdentity.sessionContext.attributes.mfaAuthenticated”. Then, I try to do the research for those 2 values.

The “additionalEventData.MFAUsed” is to determine if the user uses MFA during this specific login attempt. It used to record whether the user successfully provided an MFA token at the moment they signed into the AWS Management Console. The values use “Yes” or “No”.

On the other hand, “userIdentity.sessionContext.attributes.mfaAuthenticated” is to determine if the user performing the command has a valid MFA-verified session. It indicates the state of the session that is making the request. It tells you if the temporary security credentials (session token) being used were originally issued with MFA verification. The values use "True" or “False”.

The question is to find the field that is used to alert that AWS API activity has occurred without MFA. The **“userIdentity.sessionContext.attributes.mfaAuthenticated”** meets the requirement to alert the AWS API activity that has occurred without MFA because the “**additionalEventData.MFAUsed**” is only for login behaviour, excluding API activity.

Result:

userIdentity.sessionContext.attributes.mfaAuthenticated

4.2.3 Web Server Hardware Identification

Objective: Determine the specific processor model used on web servers.

Query:

```
index=botsv3 earliest=0 sourcetype=”hardware”
```

SPL 5. Filter the information from the hardware log

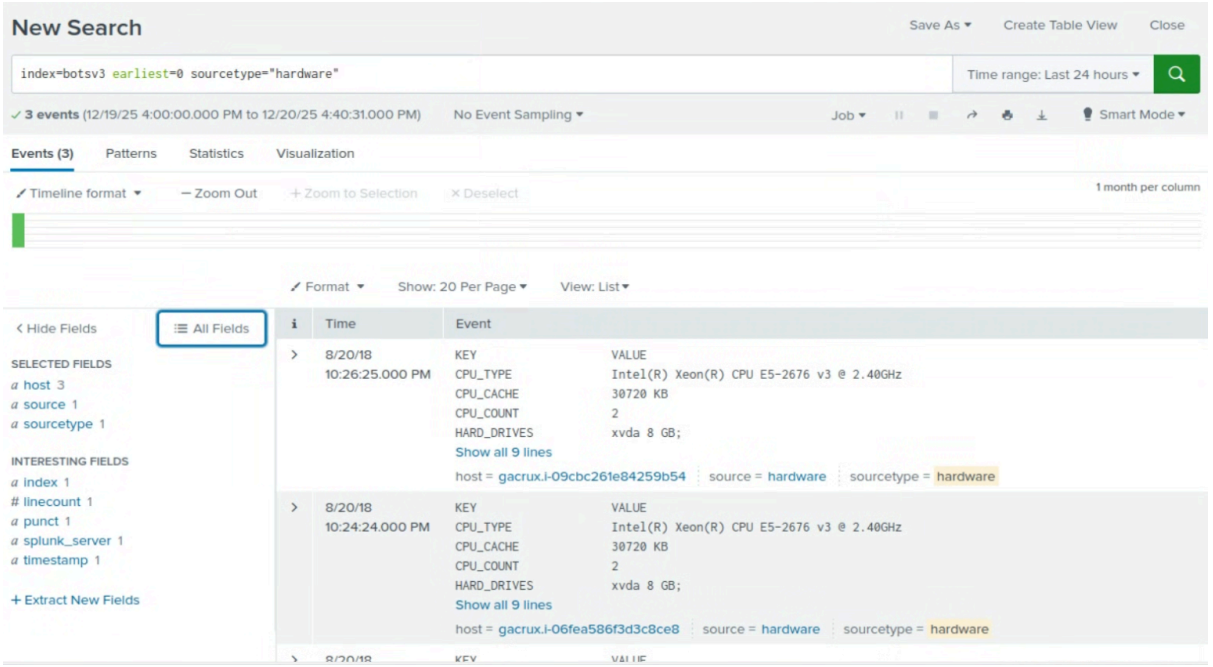

Fig 8. Capture of information from the hardware log

Hardware logs reveal the full CPU specification string.

Result:

E5-2676

4.2.4 S3 Public Access Incident

Objective: Investigate the S3 bucket exposure event (Event ID, Actor, and Bucket Name).

Q4

Query:

```
index=botsv3 earliest=0 sourcetype=aws:cloudtrail | regex
“PutBucketAcl”
```

SPL 6. Filter the information from the AWS CloudTrail log and view the log that contains “PutBucketAcl”

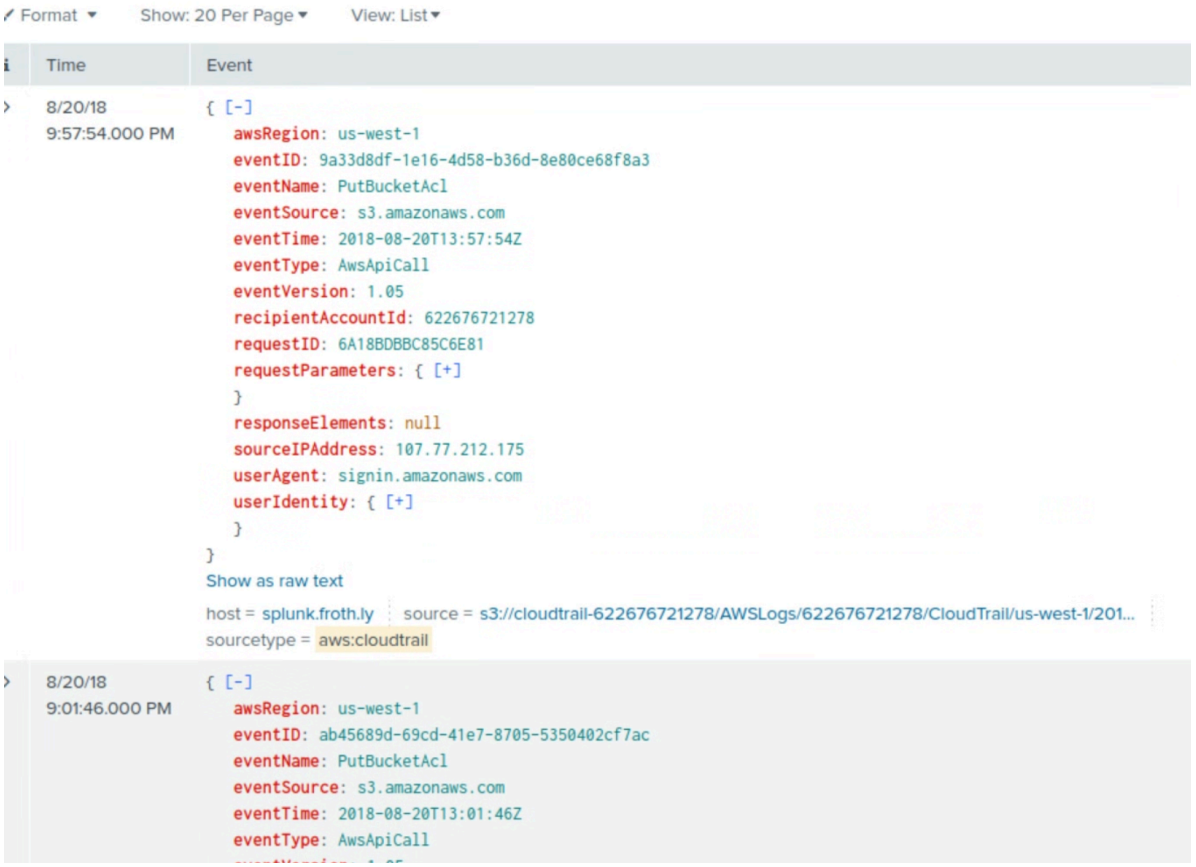

Fig 9. Capture of the log that contains “PutBucketAcl”

According to the timestamp. There are 2 events named “PutBucketAcl” that appeared at 8/20/18 9:01:46 PM and 9:57:54 PM.

Result:

ab45689d-69cd-41e7-8705-5350402cf7ac

Q5

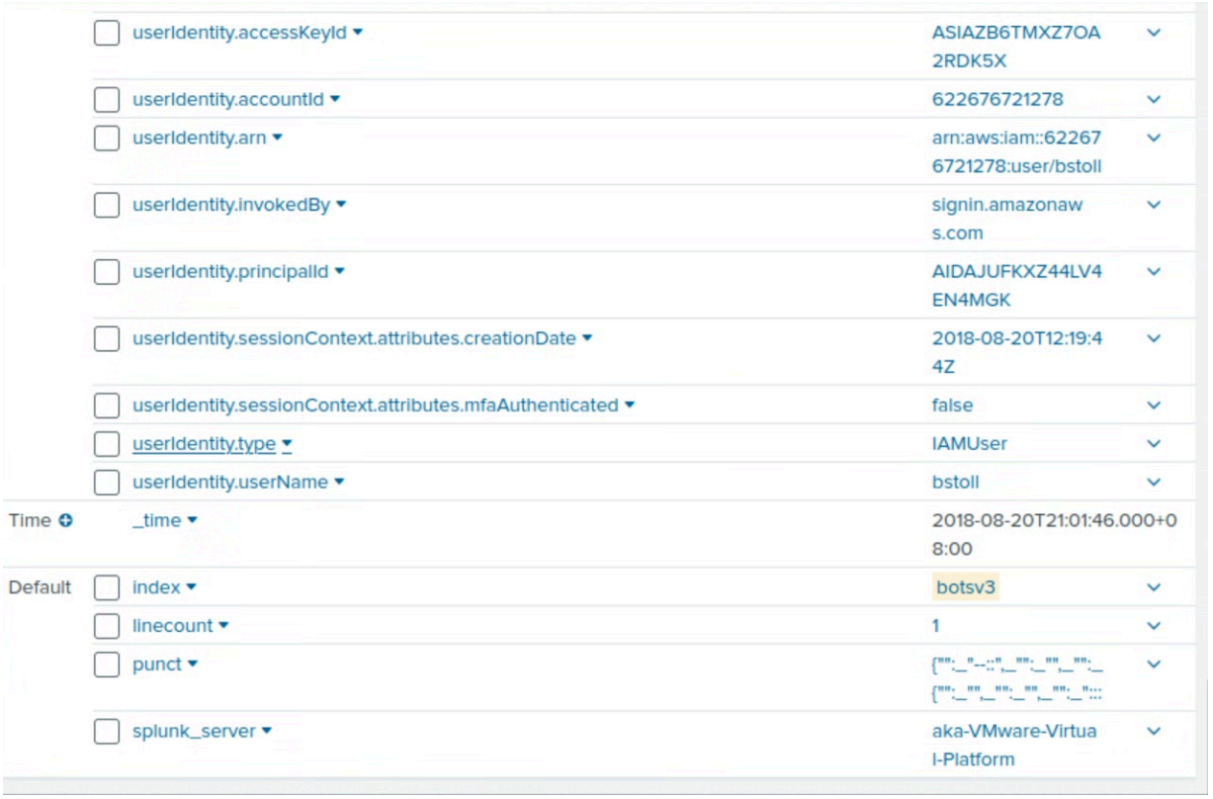

Fig 10. The user who updates the ACL in the S3 bucket

When opening the event details, the userIdentity.userName showed the action belongs to bstoll.

Result:

bstoll

Q6

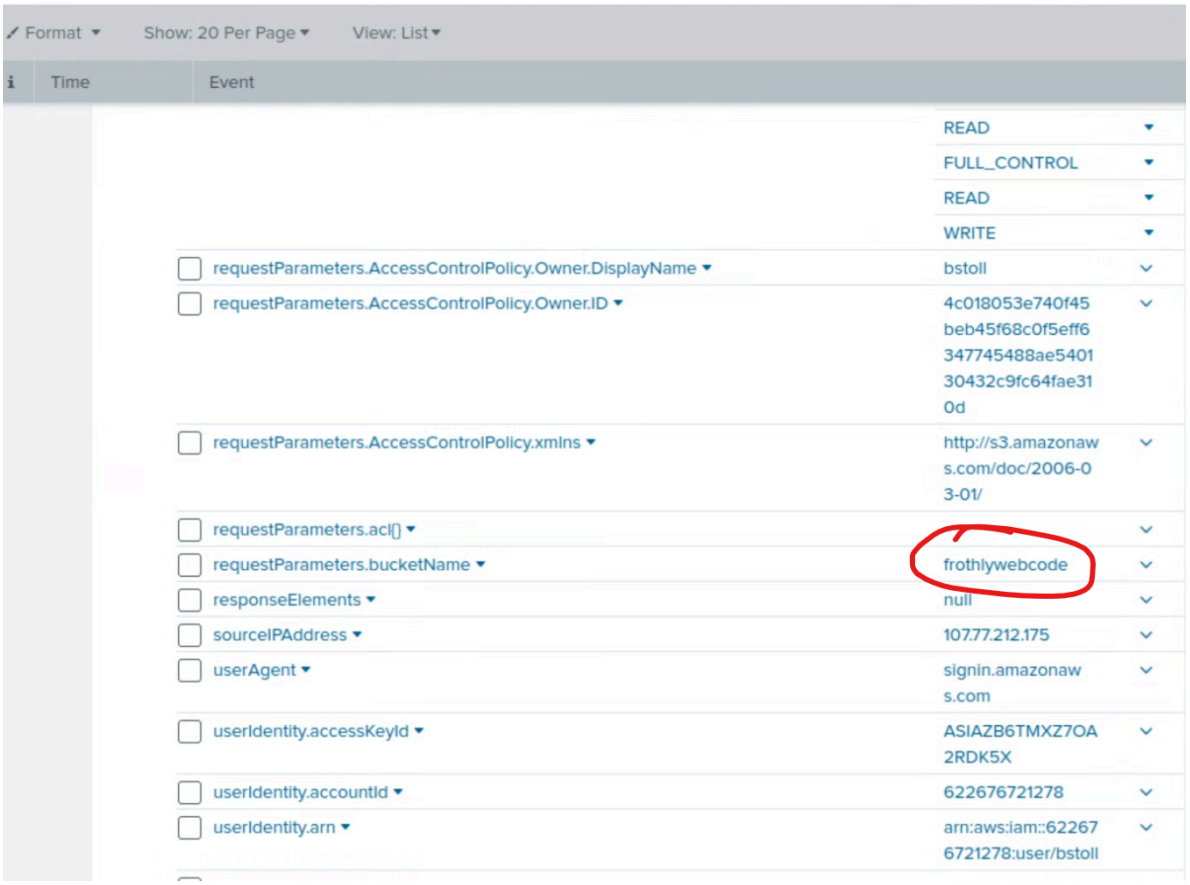

Fig 11. The S3 bucket name

Result:

frothlywebcode

4.2.5 Suspicious File Upload

Objective: Identify the text file uploaded during the public exposure window.

Query:

```
index=botsv3 earliest=0 frothlywebcode | regex “.txt”
```

SPL 7. Filter the log related to the S3 bucket and containing “.txt” in the regex

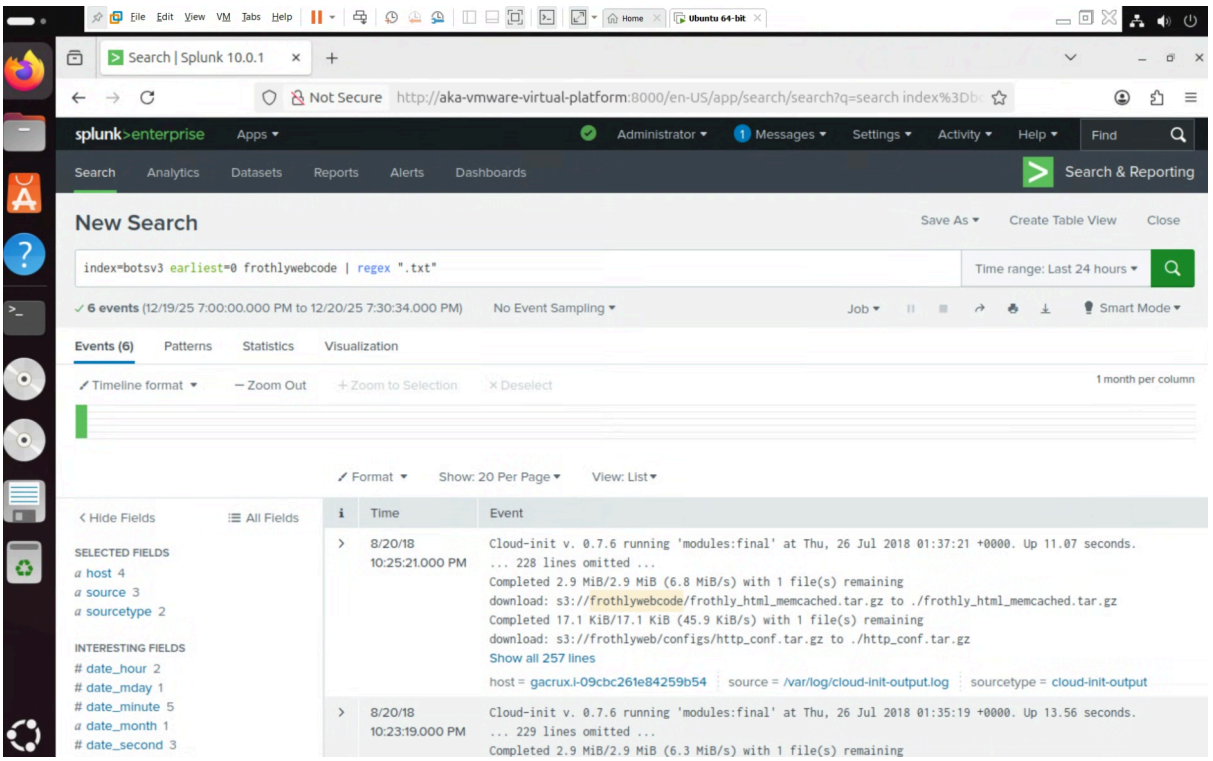

Fig 12. Capture of the log related to the S3 bucket and containing ".txt" in regex (1)

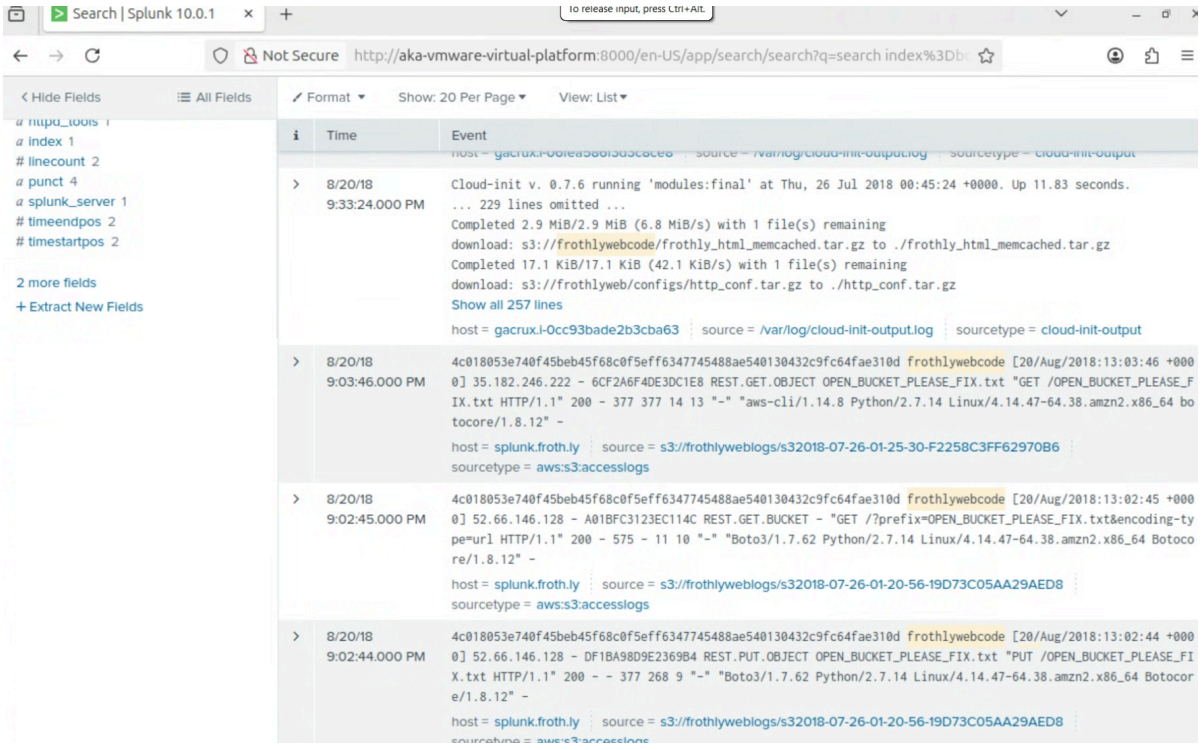

Fig 13. Capture of the log related to the S3 bucket and containing ".txt" in regex (2)

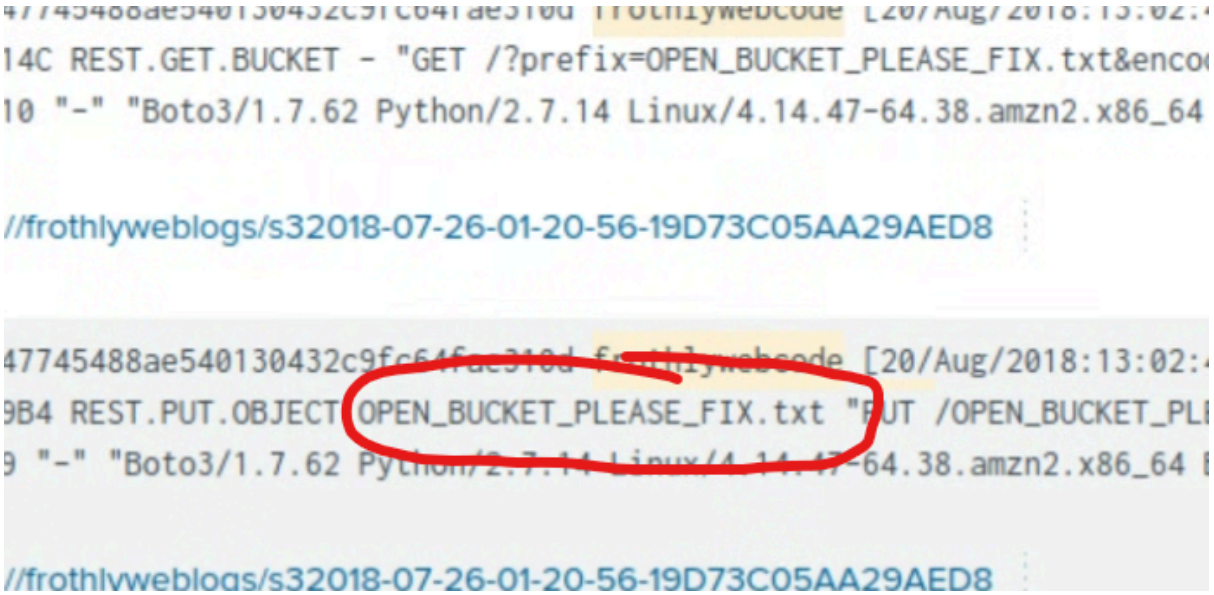

Fig 14. File uploaded into the S3 bucket

Analysis of S3 access logs shows an external PUT request for a .txt file with a 200 OK status.

Result:

OPEN_BUCKET_PLEASE_FIX.txt

4.2.6 Endpoint Outlier Detection

Objective: Identify the endpoint running a different OS edition than the standard fleet.

Query:

```
index=botsv3 earliest=0 sourcetype=winhostmon
source=”operatingsystem” | stats count by host OS
```

SPL 8. Search the operating system in WinHostMon

```
index=botsv3 earliest=0 BSTOLL-L
```

SPL 9. Search the data related to BSTOLL-L

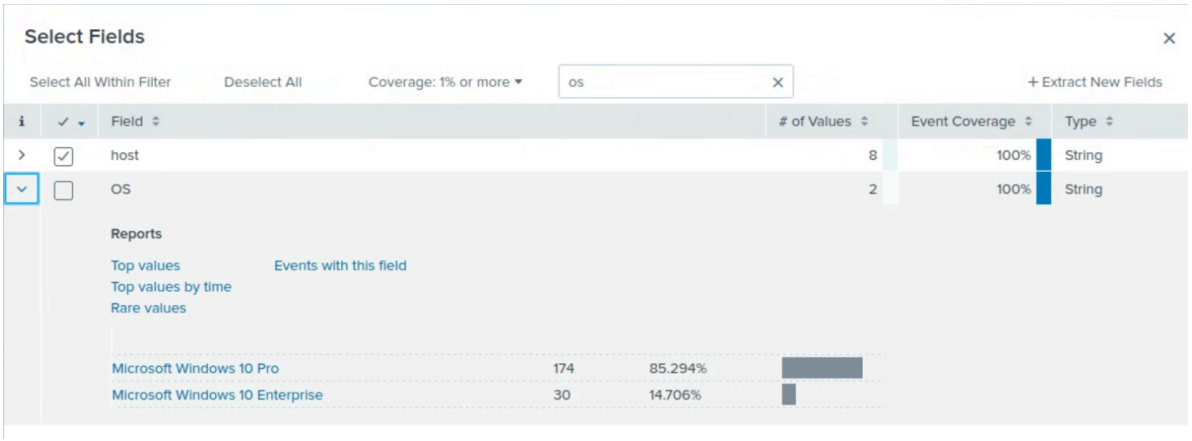

Fig 15. View the OS information in “Select Fields”

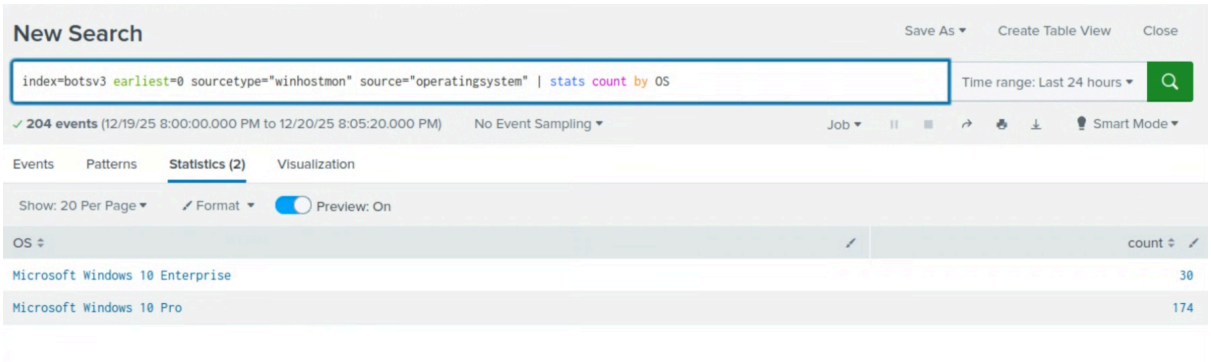

Fig 16. View statistics for operationsystem in winhostmon source

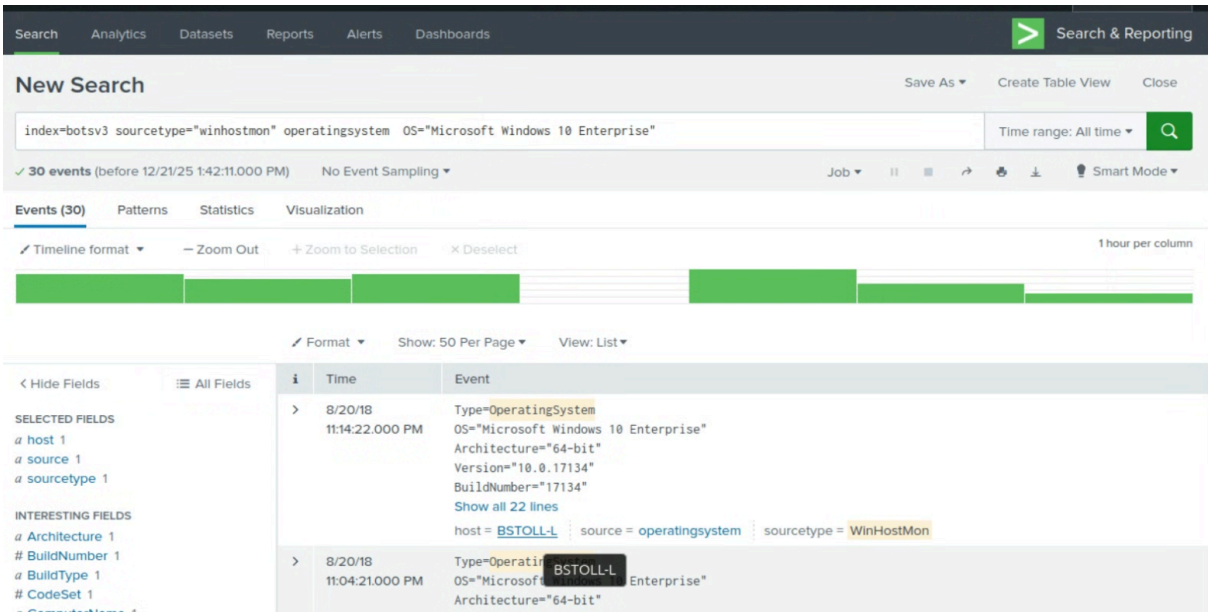

Fig 17. Windows 10 Enterprise hosted by BSTOLL-L

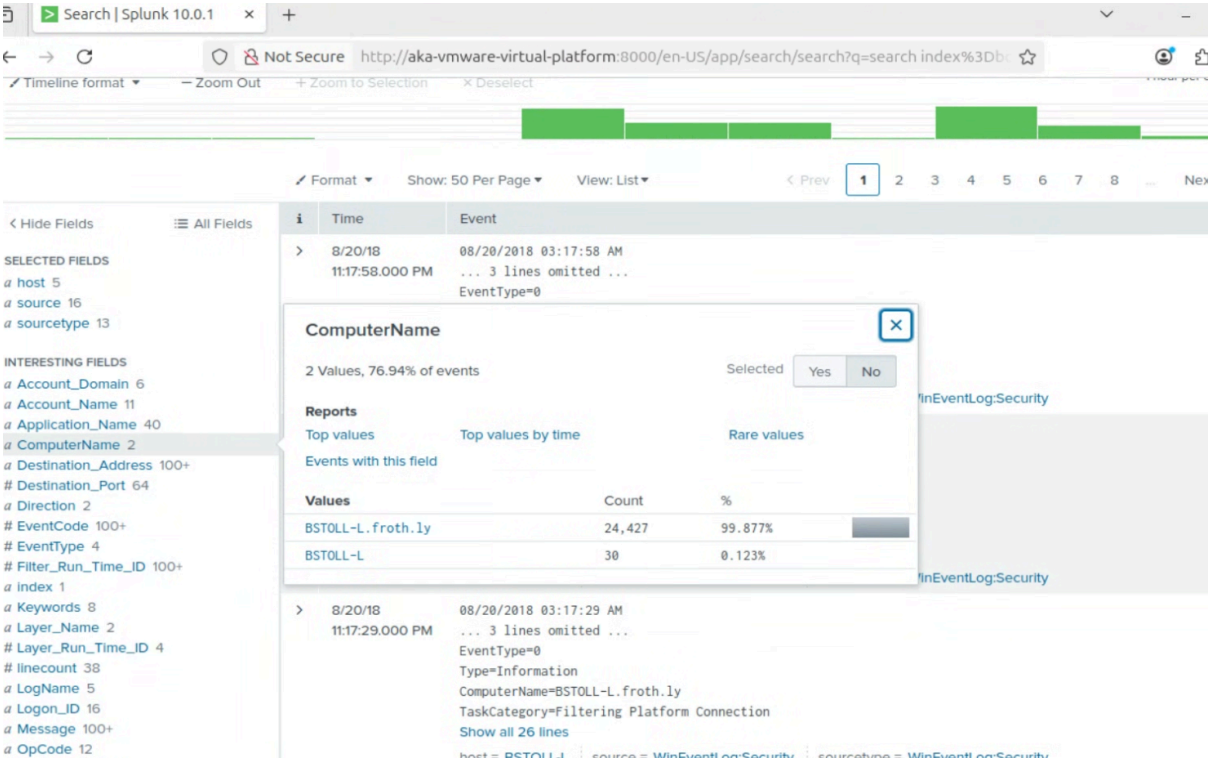

Fig 18. BSTOLL-L.froth.ly (ComputerName) in any data related to BSTOLL-L

While most hosts run Windows Server, one specific host is running Windows 10 Enterprise.

Result:

BSTOLL-L.froth.ly

5.0 Conclusion

The investigation successfully identified a critical data exposure incident caused by a misconfigured S3 ACL on the frothlywebcode bucket. The breach, initiated by user bstoll from the endpoint BSTOLL-L.froth.ly, allowed unauthorized public access and resulted in the upload of OPEN_BUCKET_PLEASE_FIX.txt.

Key Recommendations:

1. Prevention: Enable "Block Public Access" at the AWS Account level and enforce mandatory MFA for all privileged API calls to prevent accidental exposure.
2. Detection: Implement real-time Splunk alerts for high-risk events like PutBucketAcl targeting "AllUsers" to minimize Time-to-Detect (TTD).
3. Response: Deploy SOAR playbooks to automatically revert unauthorized public bucket policies and isolate non-standard endpoints (like BSTOLL-L) that deviate from the security baseline.

6.0 Reference

1. Cloudflare. (n.d.). What is a security operations center (SOC)? https://www.cloudflare.com/zh-tw/learning/security/glossary/what-is-a-security-operations-center-soc/

2. Tutorialspoint. (n.d.). Splunk - Search Processing Language. Retrieved December 24, 2025, from https://www.tutorialspoint.com/splunk/splunk_search_language.htm

3. National Institute of Standards and Technology (NIST). (2012). Computer security incident handling guide (SP 800-61 Rev. 2). https://csrc.nist.gov/publications/detail/sp/800-61/rev-2/final

4. Splunk. (n.d.). Botsv3: Boss of the SOC dataset (Version 3). GitHub. Retrieved February 13, 2026, from https://github.com/splunk/botsv3

Student Declaration of AI Tool use in this Assessment

Please indicate your level of usage of generative AI for this assessment - please tick the appropriate category(s).

If the "Assisted Work" or "Partnered Work" category is selected, please expand on the usage and in which elements of the assignment the usage refers to.

Solo Work	**S1 - Generative AI tools have not been used for this assessment.**	☐
Assisted Work	**A1 – Idea Generation and Problem Exploration** Used to generate project ideas, explore different approaches to solving a problem, or suggest features for software or systems. Students must critically assess AI-generated suggestions and ensure their own intellectual contributions are central.	☐
	A2 - Planning & Structuring Projects AI may help outline the structure of reports, documentation and projects. The final structure and implementation must be the student's own work.	☐
	A3 – Code Architecture AI tools maybe used to help outline code architecture (e.g. suggesting class hierarchies or module breakdowns). The final code structure must be the student's own work.	☐
	A4 – Research Assistance Used to locate and summarise relevant articles, academic papers, technical documentation, or online resources (e.g. Stack Overflow, GitHub discussions. The interpretation and integration of research into the assignment remain the student's responsibility.	☐
	A5 - Language Refinement Used to check grammar, refine language, improve sentence structure in documentation not code. AI should be used only to provide suggestions for improvement. Students must ensure that the documentation accurately reflects the code and is technically correct.	☐

	A6 – Code Review AI tools can be used to check comments within the code and to suggest improvements to code readability, structure or syntax. AI should be used only to provide suggestions for improvement. Students must ensure that the code accurately reflects their knowledge and is technically correct.	☐
	A7 - Code Generation for Learning Purposes Used to generate example code snippets to understand syntax, explore alternative implementations, or learn new programming paradigms. Students must not submit AI-generated code as their own and must be able to explain how it works.	☐
	A8 - Technical Guidance & Debugging Support AI tools can be used to explain algorithms, programming concepts, or debugging strategies. Students may also help interpret error messages or suggest possible fixes. However, students must write, test, and debug their own code independently and understand all solutions submitted.	☐
	A9 - Testing and Validation Support AI may assist in generating test cases, validating outputs, or suggesting edge cases for software testing. Students are responsible for designing comprehensive test plans and interpreting test results.	☐
	A10 - Data Analysis and Visualization Guidance AI tools can help suggest ways to analyse datasets or visualize results (e.g. recommending chart types or statistical methods). Students must perform the analysis themselves and understand the implications of the results.	☐
	A11 - Other uses not listed above Please specify:	☐

Partnered Work	**P1 - Generative AI tool usage has been used integrally for this assessment** Students can adopt approaches that are compliant with instructions in the assessment brief. Please Specify: - The analysis of the botsv3 log, and guide the splunk's query - Report drafting and improvement - Verify the output - Understand the SOC environment	☒

Please provide details of AI usage and which elements of the coursework this relates to:

The analysis of botsv3 log, and guide the splunk's query, report improvement and the soc report structure

I understand that the ownership and responsibility for the academic integrity of this submitted assessment falls with me, the student.	☒
I confirm that all details provide above are an accurate description of how AI was used for this assessment.	☒