Frothly Incident Analysis

Abstract

Analysis of AMW S3 bucket incident resulting in unauthorized upload of file

(s) Nikodem Drabik

Table of Contents

Introduction

This report is an investigation into a Frothly's security incident identified within the Boss of the SOC v3 (BOTSv3) dataset from GitHub. The incident was an AWS S3 bucket misconfiguration allowing public access that resulted in an uploading of a text file by an external user. This report will be using SOC methodologies and industry incident handling frameworks. Furthermore, this investigation will evaluate the root causes for incident to be possible and steps that could have been taken to prevent it from happening.

The BOTSv3 dataset simulates a real SOC environment for "Frothly" a fictional organisation. This dataset contains logs for AWS CloudTrail, S3 access and endpoints. These data sources provide a realistic environment like what a real SOC team may encounter. For the analysis I will be using Splunk and its search function, which is common in industry SOC environments.

The Scope of the report is limited to a single 200-level question from this dataset. Other incidents are present but will be excluded. I also assume that the logs present are complete, accurate and untampered as any deviation from this would impact the accuracy of the incident analysis. Additionally, no evidence relevant to this incident is present to indicate the accounts are compromised or fictitious.

The following objectives for the report are:

1. Reconstruct the timeline of events of the incident
2. Uncover and asses the SOC detection and response
3. Evaluate any failures
4. Propose cost effective, realistic solutions improvements for the incident

SOC & Incident Handling and Response

In this incident review I will be following Security Operation Centre (SOC) three-tier incident handling procedures of prevention, detection, response and recovery. Which closely align with the Cyber Incident Response Cycle (CIRC) prevention, detection and analysis, containment eradication and recovery and post incident. These two procedures are closely related and provide guidance on the best method to analyse an incident.

The Prepare and Detection stage is performed by SOC tier 1 analysts that monitor the activity, systems logs and alerts to gather activity baseline and statistics to identify anomalies. Their tasks in this incident would have involved detection of the misconfiguration incident, identification of IAM users, continuous identification and review of IAM privileges as stated in *"PR.AA-05"*[5] and verification of MFA utilisation. These provided foundational evidence to support higher tier teams' response as reinforced by NIST SP800-61r3[5] outlining best practice for Cyber Risk Management.

The Respond stage is performed by SOC Tier 2 analysts; they conduct in-depth investigation of the reported incident. They correlate gathered data such as CloudTrail logs and S3 access logs;

to create a timeline of events and get an understanding of the root of the incident and respond by implement short-term containment strategies outlined by the company policy.

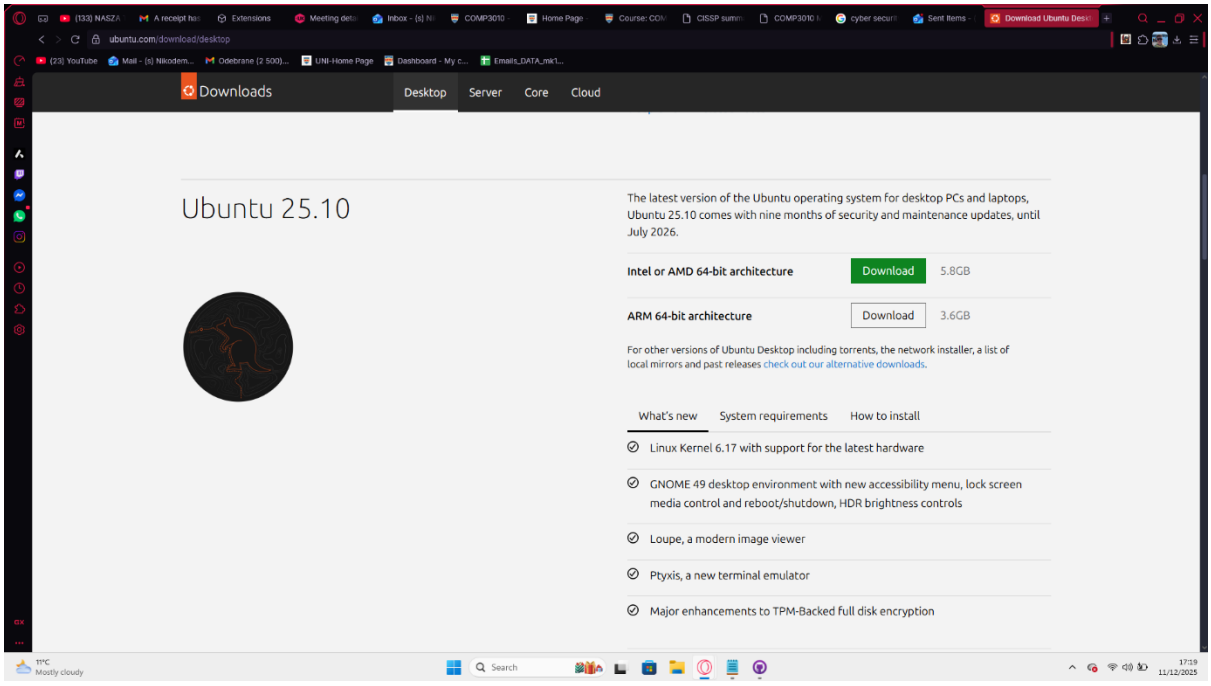
Lastly the Post-incident Follow up stage is performed by SOC tier 3, where they focus on the root cause of the vulnerability and long-term remediation for this incident and recovery of the organisations to pre-incident state.

Installation & Data Preparation

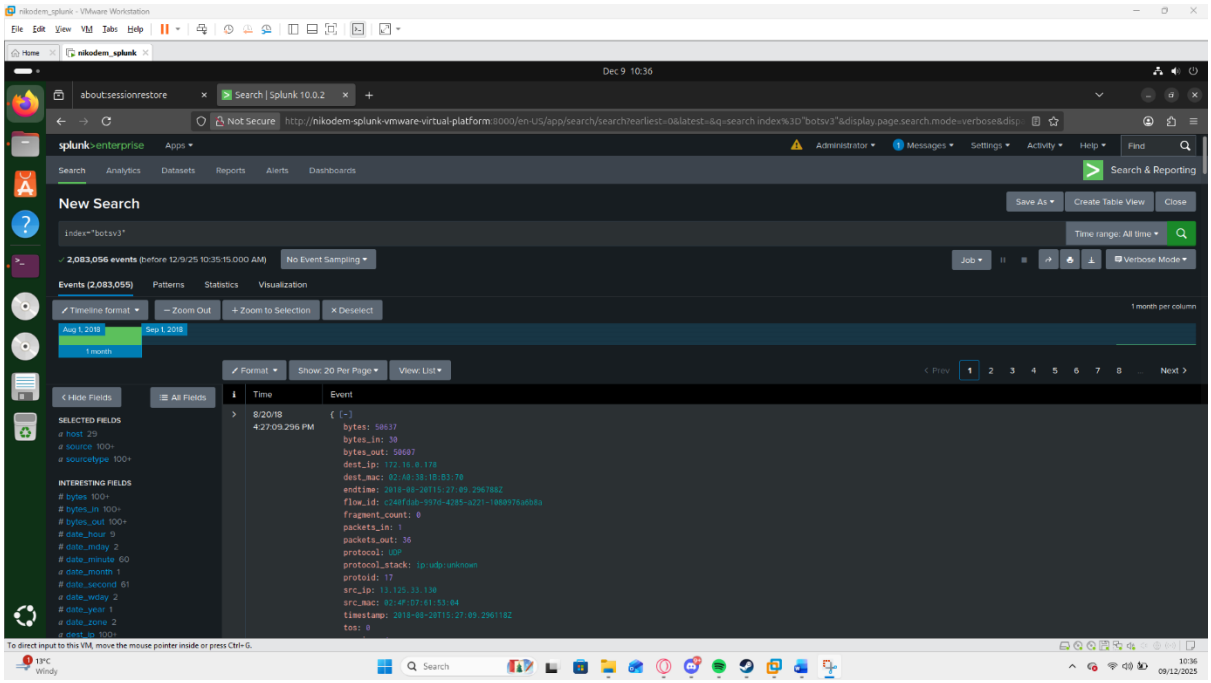
For my environment I utilised VMware software to host my Ubuntu 25.10 virtual machine with 8.6gb of RAM and 8 processors. I began by installing Splunk, followed by downloading the BOtsV3 data set from GitHub and then having it ingested. Finally, I validated the BOTSv3 dataset was validated through Splunk's UI by the event count of 2,083,056 when querying *index="botsv3"* to ensure I had a complete dataset. It is important to have all relevant facts for a quality analysis.

Splunk is a good choice for this SOC monitoring due to its proven application in industry and easy to learn user interface. The Splunk environment allows identification of anomalies using graphs and logs. It is also the monitoring software that the fictional organization is using based on the dataset.

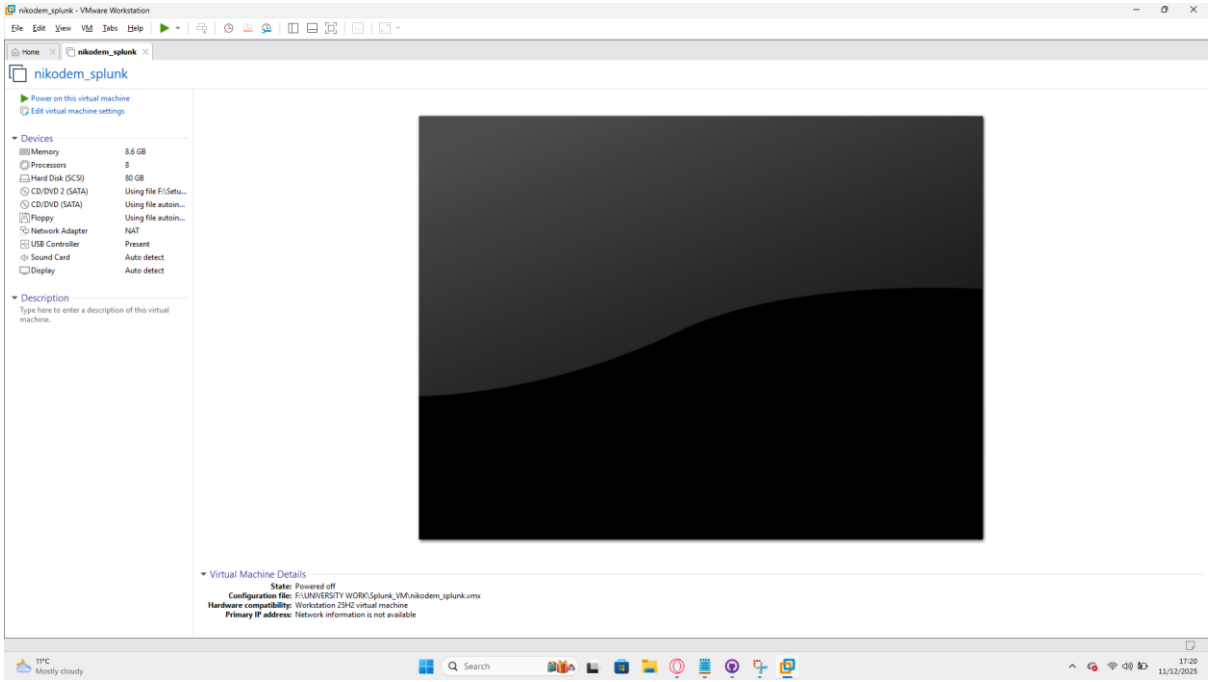
For a SOC environment it is crucial for the environment to be able to perform all the required duties and be able to ingest multiple data sources and be able to display them efficiently and completely to allow effective action to be taken. For my task Splunk has all the resources I require indicating it is a good choice.



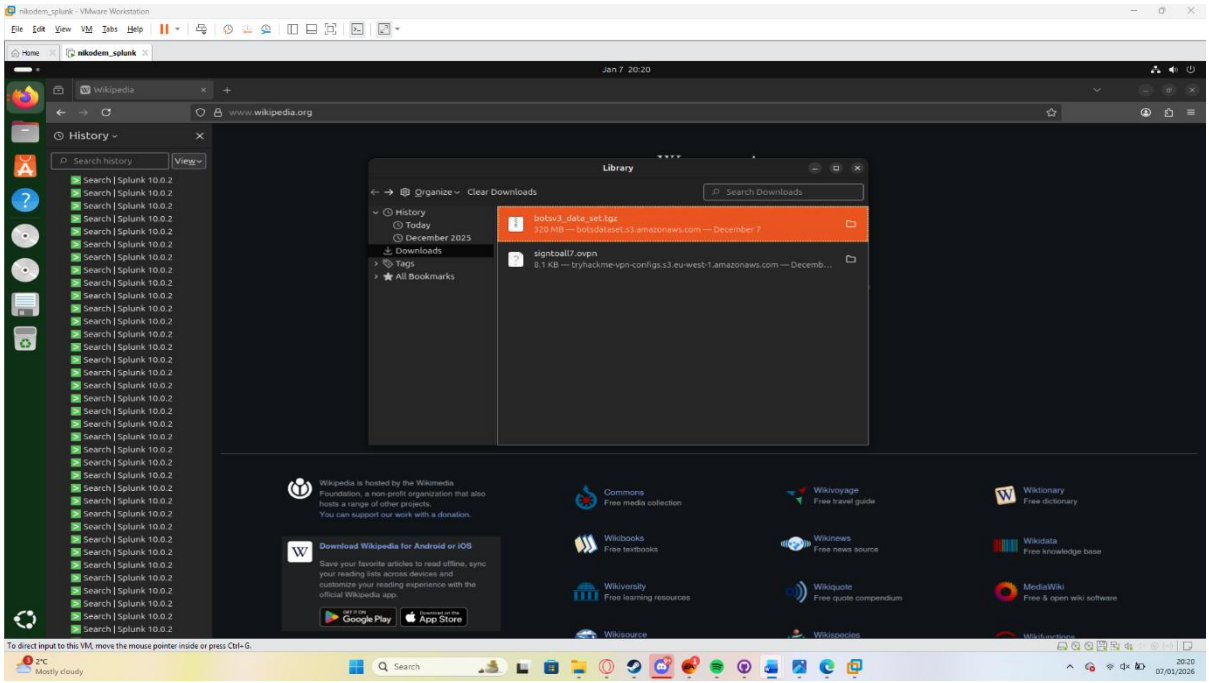
[UBUNTU install Page]



[Verified data ingestion]



[Hardware]



[Ubuntu Downloading BOTSv3]

Guided Questions (BOTSv3 Q&A) Walkthrough

Initial SOC tier 1 duties focused on identifying raw data and prepare it for upper tiers to review and identify the issue.

Starting by identifying IAM (Identity and Access Management) users. As seen in Fig 1 there are 4 users (bstoll,btun,splunk_access,web_admin) that have access to the system in some capacity. It is important to identify active users and asses if they have the least-privilege implemented.

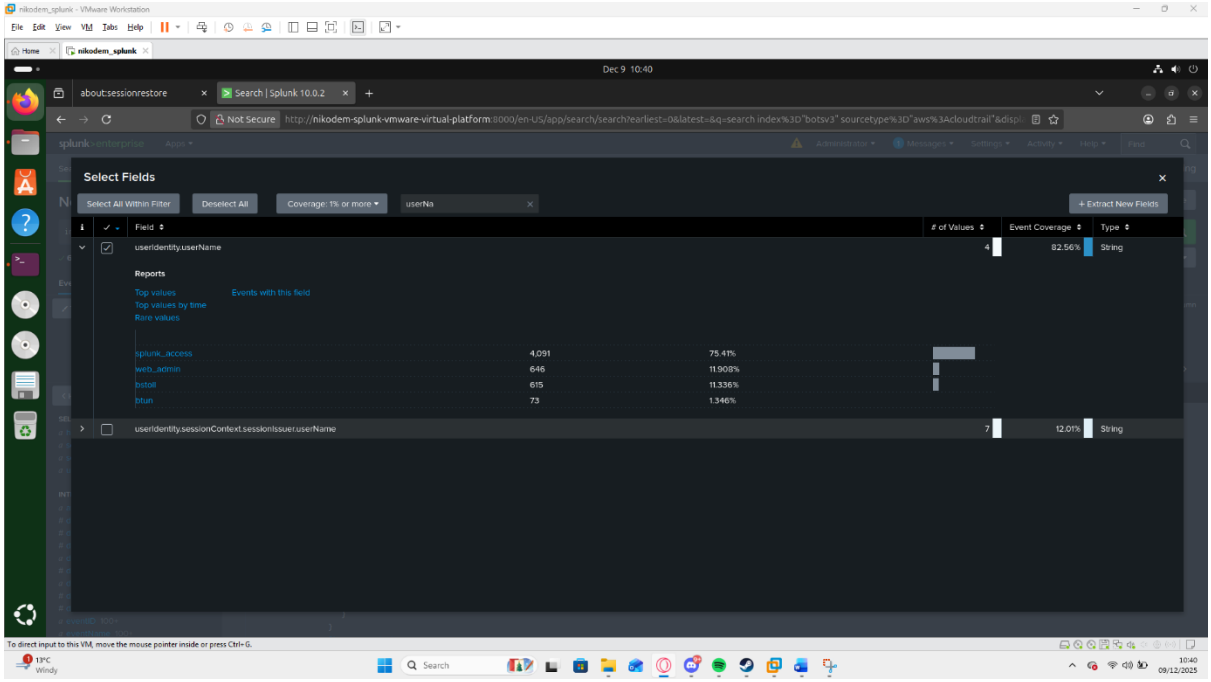


Fig 1 – 4 users

Secondly, review of MFA if it was enabled when the AWS API activity occurred. MFA ensures that the change is performed by an authorized individual with valid access reducing risk in erroneous critical changes. Having MFA also alerts SOC tier 1 team of changes before/while they happen.

I used “userIdentity.sessionContext.attributes.mfaAuthenticated”=* to identify all logs with MFA when AWS API activity occurred. Fig 2 and Fig 3 show 2,155 events with MFA being false indicating it is disabled and 0 with it enabled. Indicating its disabled for all events in this logged dataset. This is may be caused by MFA not enable by default in security settings in 2018 [2] which is the age of these logs.

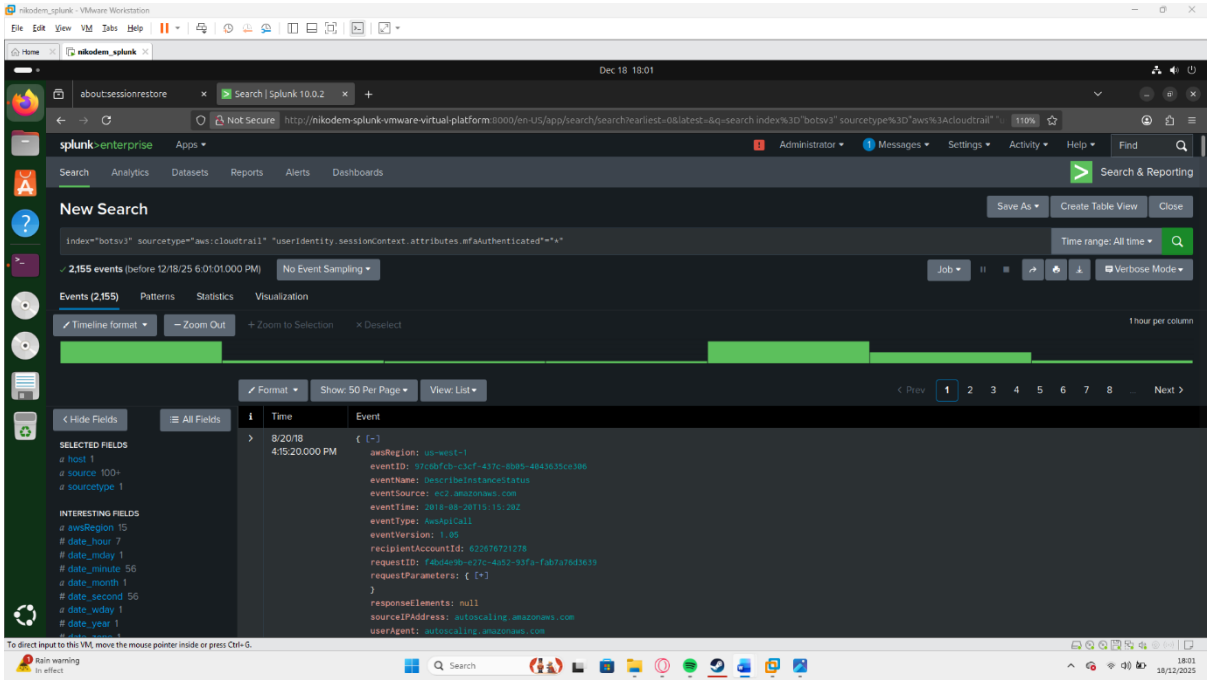


Fig 2 – all MFA attributes

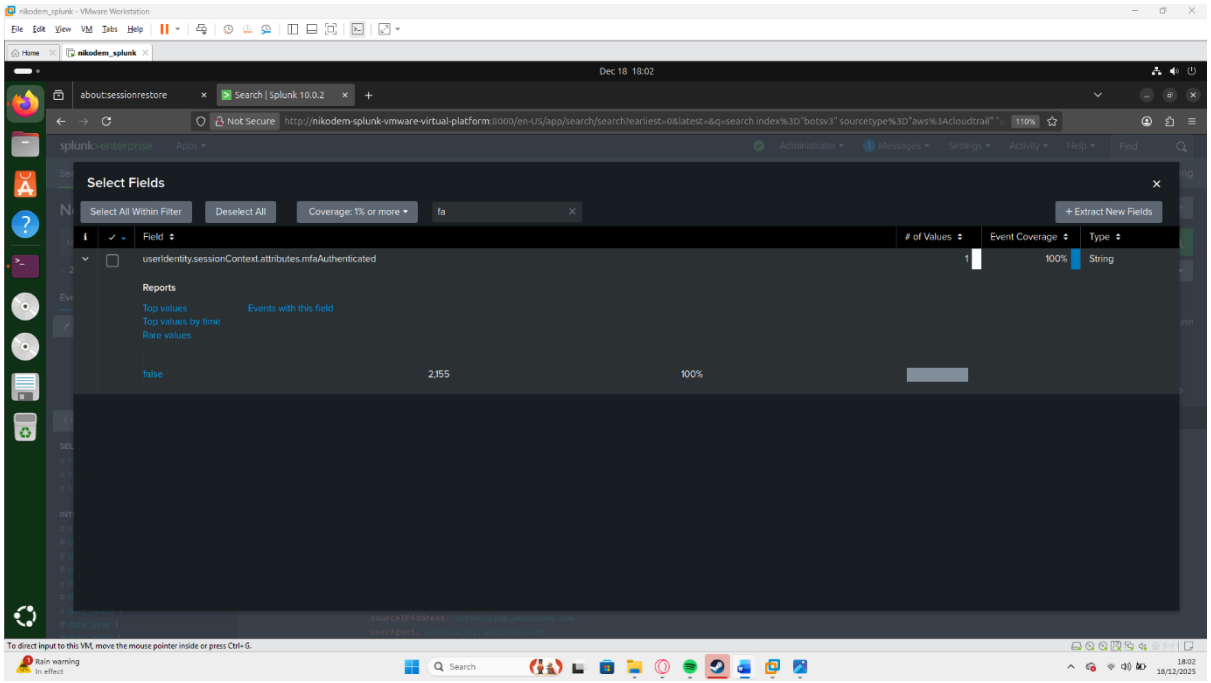


Fig 3 – MFA only false

Following this I mapped the web server's processor to be Intel Xeon E5-2676 (fig 4) a server grade CPU. Identifying the hardware is critical for analysis of vulnerability in the detection stage of CIRC, as the vulnerability may be hardware related and is crucial for the containment and recovery stage. However, in this incident further evidence suggests that this was not relevant to this incident.

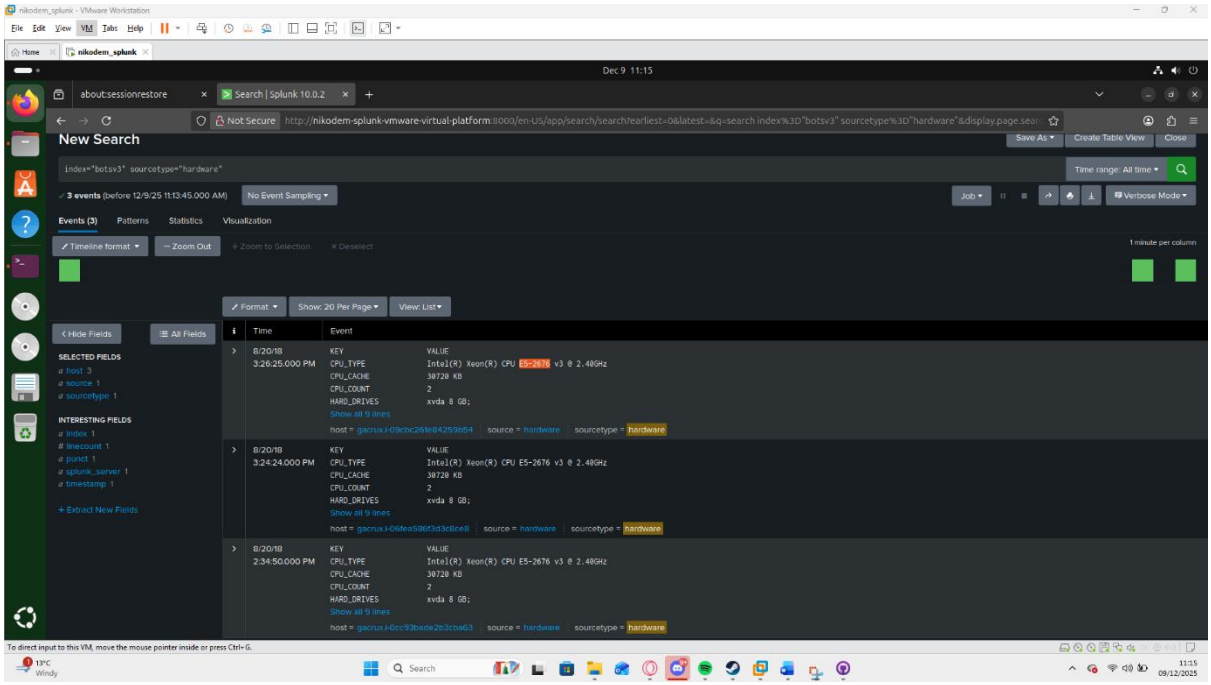


Fig 4 – Hardware

Around this section of the investigation SOC Tier 2 Incident Response team would take over due to the it being a higher risk incident. As further threat intelligence gathering and analysis

identified the CloudTrail with event id: “ab45689d-69cd-41e7-8705-5350402cf7ac” contained the API call for misconfiguration of the S3 bucket “frothlywebcode” (fig 7) publicity status. As seen in fig 6 user bstoll gave AllUsers group Read and Write permission at 13:01 on 20/08/18 (fig 5). This allowed non-authorised users to read this bucket and modify it after this period. No remediation events were observed following the permission change, suggesting that the misconfiguration was not detected or acted upon until the incident. However, in the same event bstol granted what seems to be a legitimate read and write permission to Log Delivery group. This may indicate that due to this being in the same event as the misconfiguration it may have been an accidental Masquerade as the illegitimate configuration is masked by legitimate use.

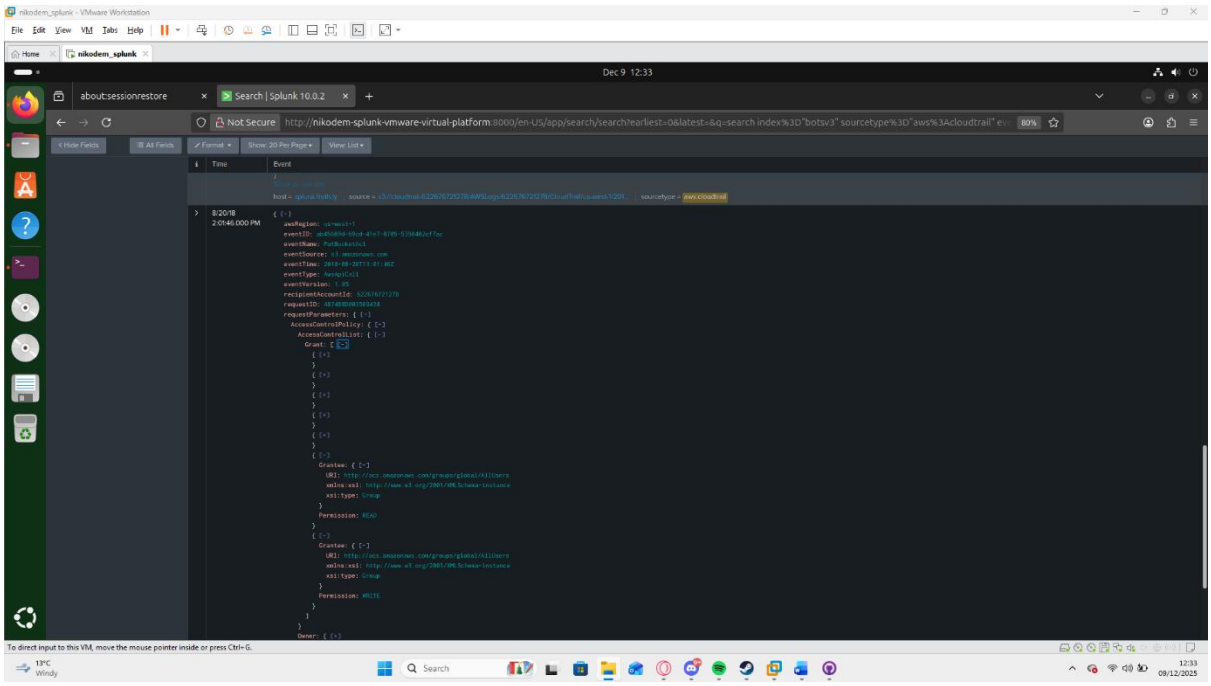


Fig 5 – Changes performed

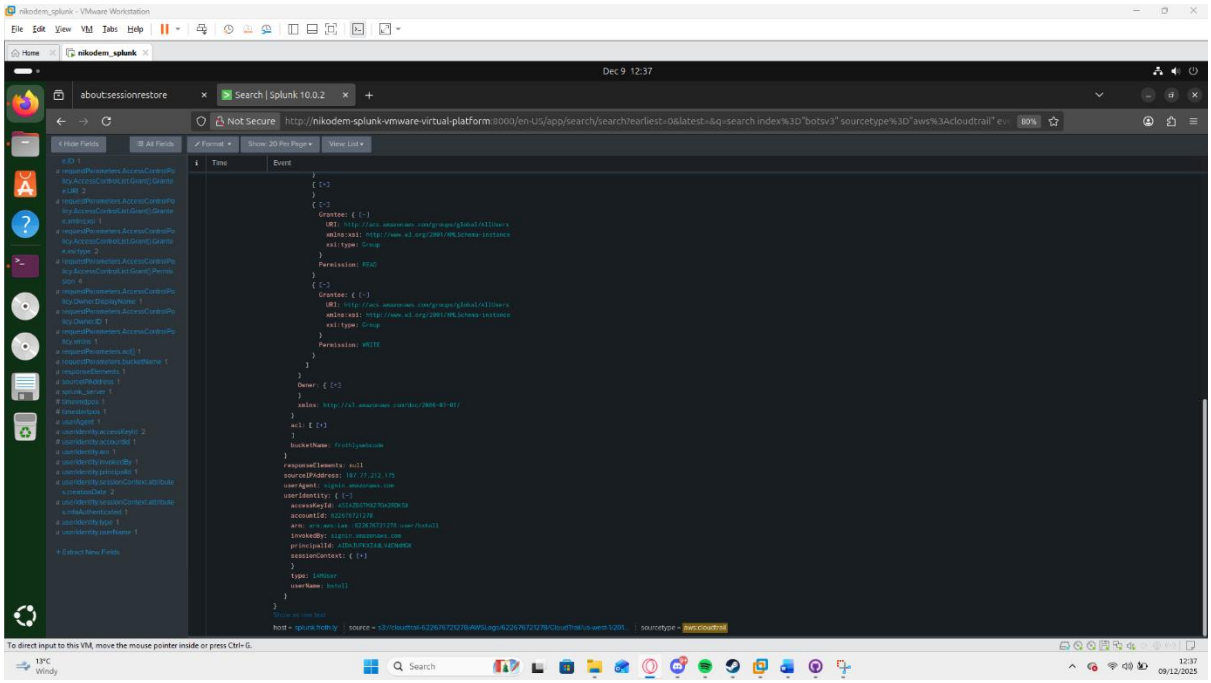
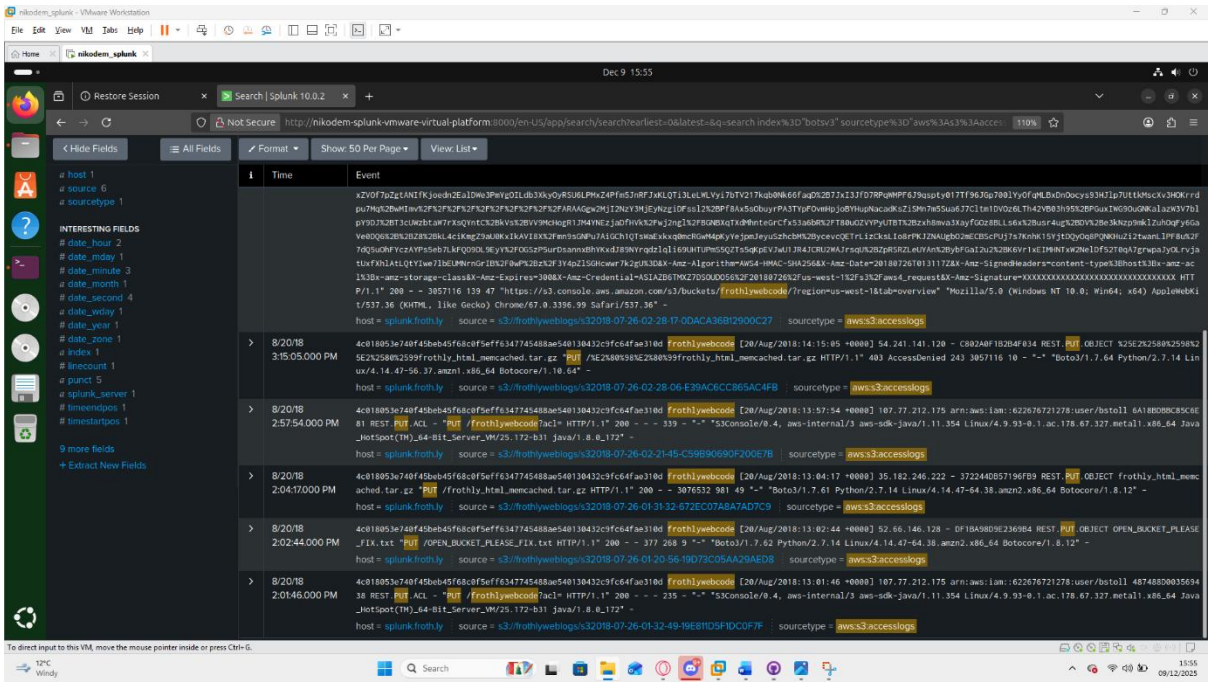


Fig 6 – User responsible

However as mentioned MFA was disabled meaning users accessing the S3 bucket were not verified to be whitelisted by the organisation. Lack of this control resulted in an unauthorised user uploading an unauthorised file OPEN_BUCKET_PLEASE_FIX.txt at 14:02 20/08/18 to the bucket (Fig 7).



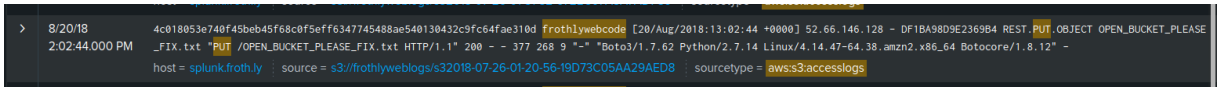


Fig 7 – Bucket Name and txt upload

Lastly, I analysed the user BSTOLL to review the user and their endpoint to understand the scope of the vulnerability and identify why this bstol was able to create it. Searching " ‘sourcetype=winhostmon’ OS” allowed me to identify two varieties of OS Windows 10 Pro and Enterprise (Fig 9). By Isolating only events from devices using Windows 10 Enterprise; BSTOLL-L was the outlier (Figs 10). Further analysis of data relevant to bstol identified his FQDN (Fully Qualified Domain Name) as BSTOL-L.froth.ly (Fig 8) following the pattern of “ ‘*user*’.froth.ly ” . With bstol having a different OS from the rest of the user may have resulted in the alert system not understanding the syntax of the requests sent on the machine.

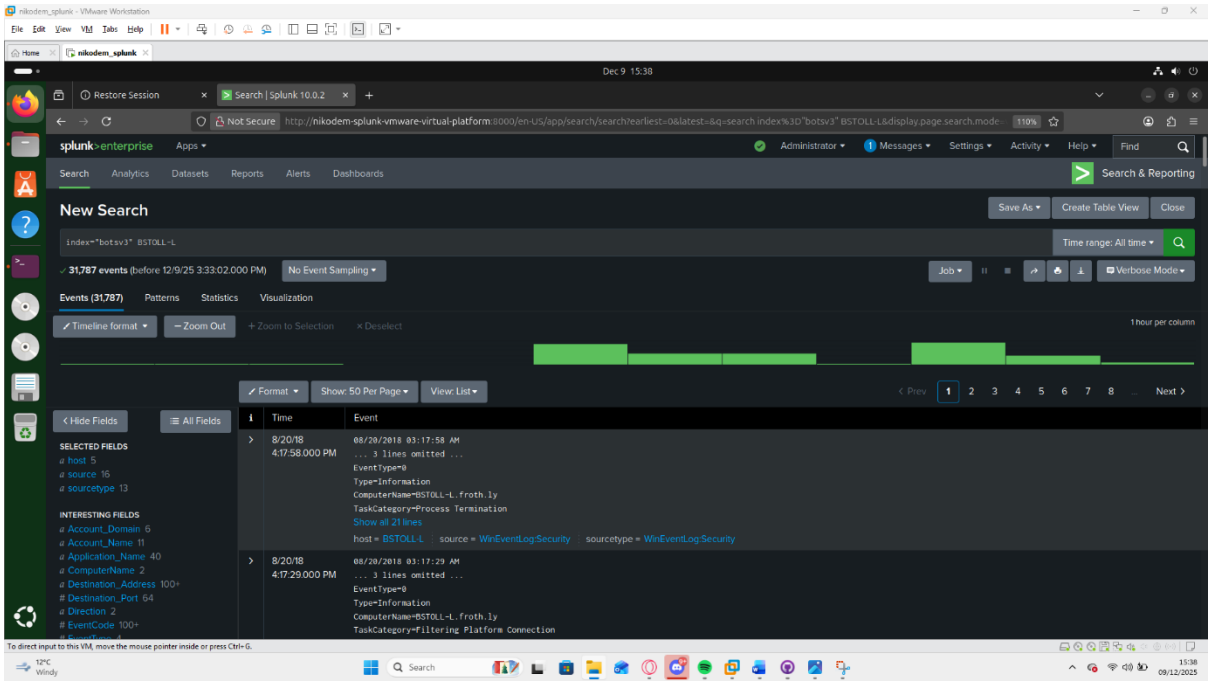


Fig 8 – BSTOL FQDN

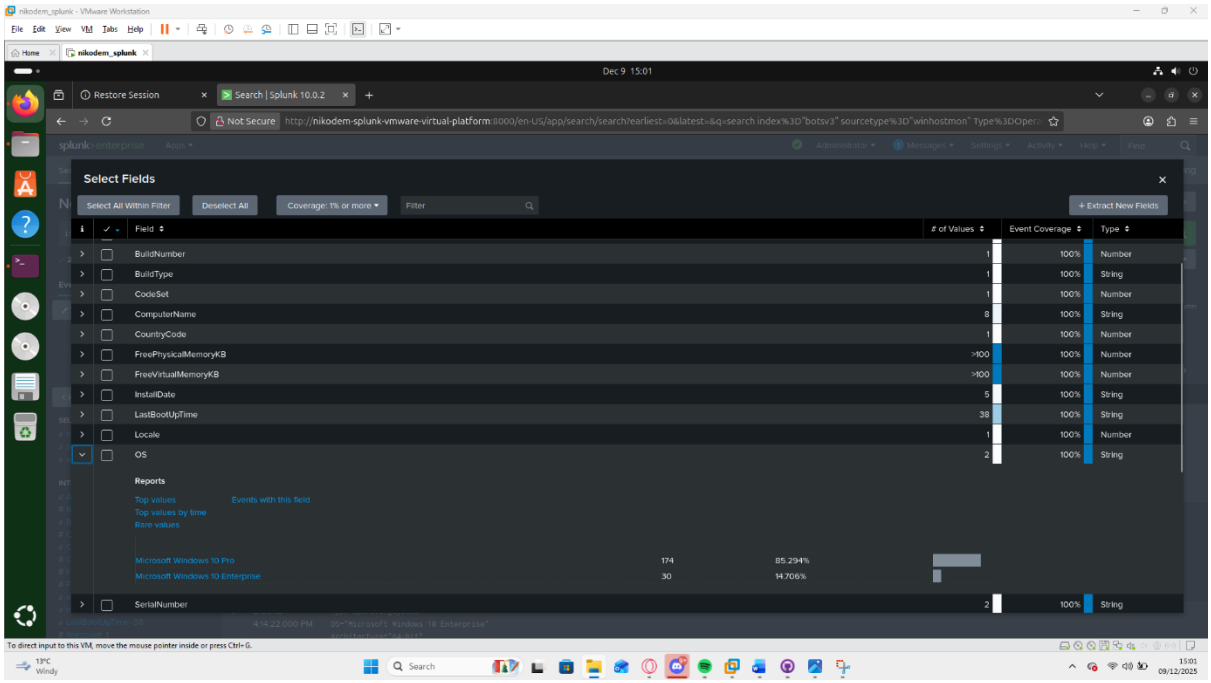


Fig 9 – OS types

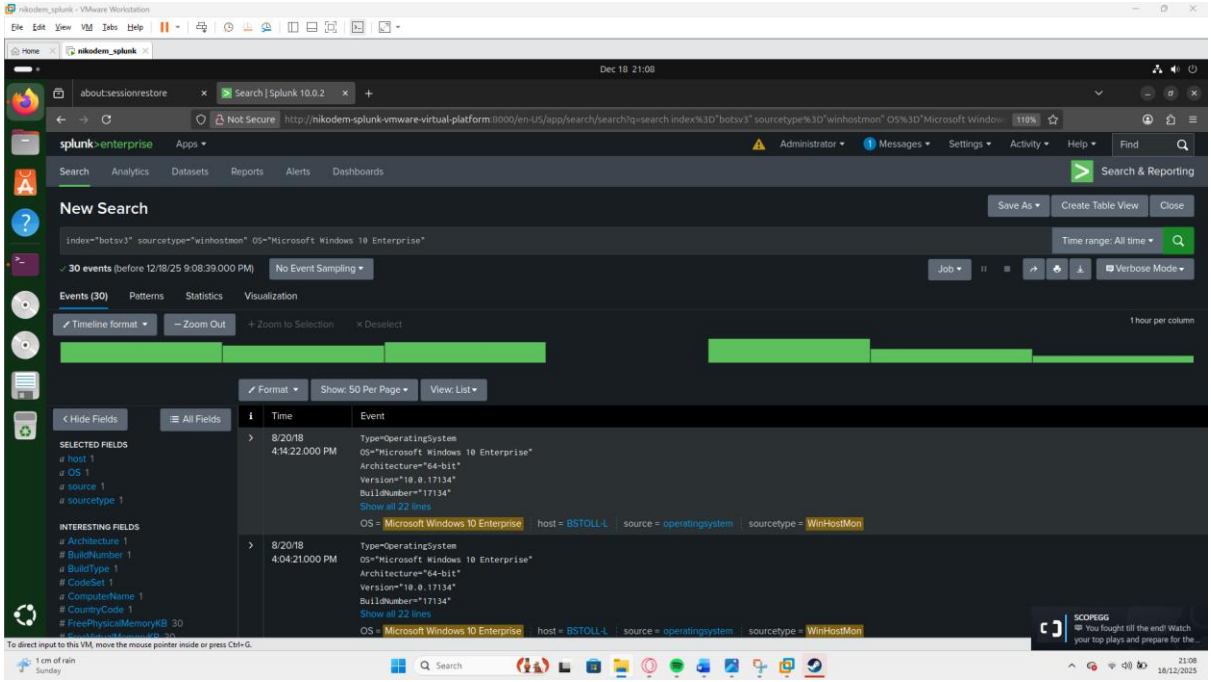


Fig 10.1 SPL isolating OS

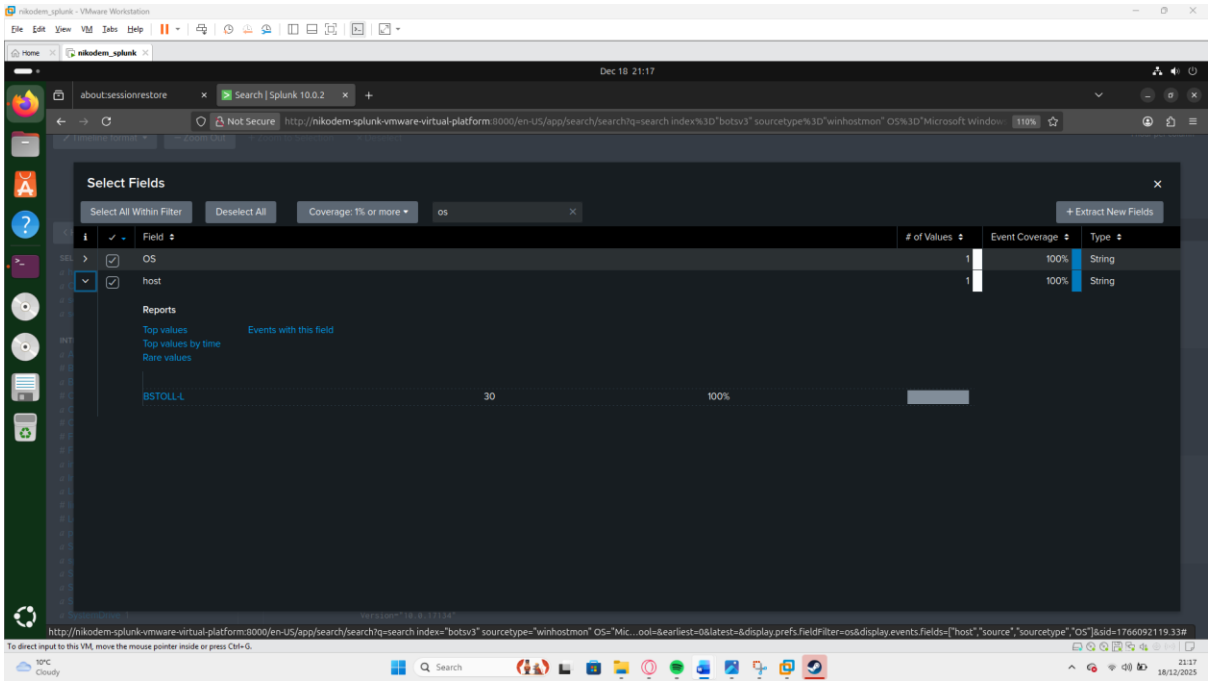


Fig 10.2 – BSTOL-L's identified in filters.

After the identifying the vulnerability and possible causes. The SOC team should now contain or fix the vulnerability. By viewing the logs, I was unable to any containment or remediation, which indicates a potential gap in incident response process.

Outcome

After the CIRC response step the tier 3 SOC team member should perform the follow up and post incident review. This includes reviewing the data that was gathered regarding the vulnerability and asses the containment and remediation techniques implemented, review other possible reasons this vulnerability evolved and updating the risk assessments based on the incident to remove the vulnerability and improve security measures.

I have not been provided the implemented prevention methods or any risk assessments that Frothly may be using, so the next step is to analyse the relevant risks identified in this incident and implement prevention techniques:

Risk level is assigned on a qualitative scale of 0–10 based off the combination off impact and likelihood of happening again without changes.

RISK	Risk Level	Consequence	Mitigation
MFA Disable	10	User can perform actions without verification of permission	Enable MFA for all users
Excessive IAM permissions	10	Users have access that is beyond their required operational scope	Review permissions and instate least-privilege

Non-Uniform OS	4	Certain network wide settings, restrictions or alerts may not register on the machines due to different architectures	Utilize Uniform OS and versioning.
No SOC alerts	10	Critical changes to S3 Bucket's security were not verified nor reported.	Implement alerting for critical events.
Lack of monitoring and incident response.	10	Critical vulnerabilities are left open for prolonged time.	Impose monitoring and incident response procedures.
Publicly available S3 Bucket	10	Critical data, codebase and access is made public.	Remove Public access and disable it from happening.

First step to take is to disable Public Access as it's the highest priority and the vulnerability that caused the incident.

Secondly MFA being disabled is Risk level 10 requiring instant mitigation. Enabling MFA would require the user to verify identity removing the possibility of a unverified user using the bucket. Secondly Enable the Block Public Access setting [1], which is responsible for disabling access to everybody without explicit permissions. These two settings would prevent this incident due to the action of modifying the bucket would be blocked and the unverified user would not have access as they would not pass MFA.

Thirdly review of excessive IAM permissions is Risk level 10 requiring instant mitigation. Based on the industry naming convention of user's web_admin and splunk_access likely indicate these are admin accounts for Splunk software and the company website with elevated permission. Standard users bstol and btun are likely intended to have minimal user permissions as that is best practice; and not be allowed to change permission of all users as that is a high-level permission not for normal users.

OS variance is a risk of 4 at this stage due to different OS often having different syntax and security settings. As bstol has a different OS, this may result in security measures implemented by SOC not correctly be implementing on his OS due to the different versions.

It is possible alerts from his inputs may contain different syntax resulting in the alert software misinterpreting them leading to no alert and no action form SOC team. Standardising OS for the whole organization would reduce the risk further however this is a low risk due to both OS being Windows Ecosystem. However, SOC tier 3 should review logs of alerts and if review of alerting systems identifies that the OS version is responsible for lack of alerts this should increase the risk level on risk assessment and increase its priority to level 10.

Conclusion

Review of this incident and analysis of the evidence, indicates this is likely caused by the inadequate implementation of security measures and monitoring by the organisation Frothly. AWS was not at fault as outlined in their Shared Responsibility Model [3] as IAM and Network

Traffic Protection are the Customers responsibility. The analysis of AWS CloudTrail and S3 access logs revealed misuse of IAM permissions, MFA enforcement not present and lack of S3 bucket security settings resulting in misconfiguration and a vulnerability. The available evidence indicated cost effective solutions were available before the breach that would likely prevent this incident. Overall, these findings highlighted the need for continuous monitoring within SOC and timely remediation plans put in place to prevent future security failures and a review of the current procedures. However, I have not found any evidence of this incident affecting Frothly's infrastructure long term.

Supporting Information

References

[1] AWS, *Blocking public access to your Amazon S3 storage*, accessed 19 December 2025, https://docs.aws.amazon.com/AmazonS3/latest/userguide/access-control-block-public-access.html.

[2] Amazon Web Services, *Secure API access with MFA*, AWS Identity and Access Management User Guide, accessed 19 December 2025, https://docs.aws.amazon.com/IAM/latest/UserGuide/id_credentials_mfa_configure-api-require.html

[3] Amazon Web Services, *Shared Responsibility Model*, AWS Compliance, accessed 22 December 2025, https://aws.amazon.com/compliance/shared-responsibility-model/.

[4] Alex Nelson, Sanjay Rekhi, Murugiah Souppaya and Karen Scarfone, *Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile*, NIST Special Publication 800-61r3, National Institute of Standards and Technology, April 2025, accessed 27 December 2025, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r3.pdf

[5] Government Security Group, *Principle B2: Identity and Access Control*, Government Cyber Security Policy Handbook, accessed 27 December 2025, https://www.security.gov.uk/policy-and-guidance/government-cyber-security-policy-handbook/principle-b2-identity-and-access-control/

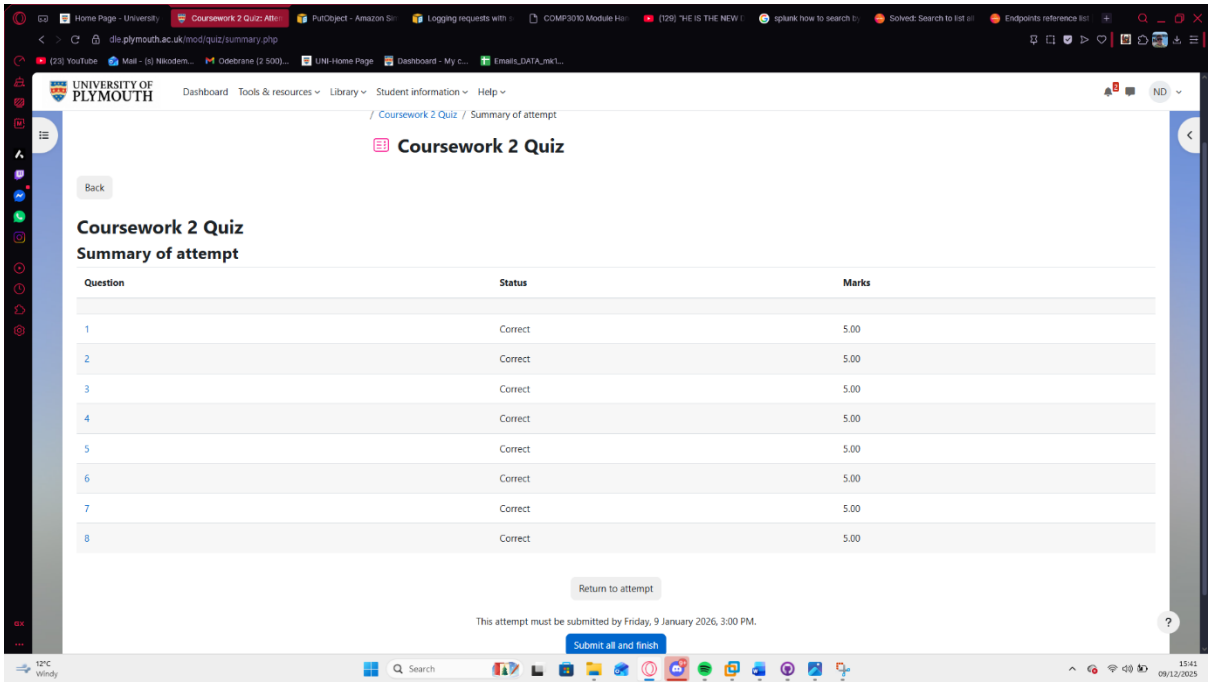
Video Link

https://youtu.be/4OlrZ5YUPSg

GitHub Link

https://github.com/Niko-PL/Comp3010-2025-Nikodem-v1

Proof Of Quiz



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:	☐

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

A5- Language Refinement:

Used for identifying syntax errors.

Used to provide suggestions for academic language.

Used to provide suggestions for better flow of sentences

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.	☒