SECURITY MONITORING TOOL SYSTEM USING THREAT INTELLIGENCE vs THREAT HUNTING

WAN IKBAL ISMAT BIN WAN KAMAL

OPEN UNIVERSITY MALAYSIA

2021

SECURITY MONITORING TOOL SYSTEM USING THREAT INTELLIGENCE vs THREAT HUNTING

WAN IKBAL ISMAT BIN WAN KAMAL

A Final Year Project submitted in fulfilment of the requirements
for the degree of
Bachelor of Information Technology in Network Computing with Honours (BITN)

Open University Malaysia

2021

SECURITY MONITORING TOOL SYSTEM USING THREAT INTELLIGENCE vs THREAT HUNTING

ABSTRACT

This project is about developing a Security Monitoring Tool System using Graylog SIEM (Security Information Event Management) with a combination of Threat Intelligence and an expected outcome for Threat Hunting results. This is built in accordance to specific ruleset been made for threat hunting purposes with an automation of logs from Windows endpoint host and Network activity. A datasets of Threat Intelligence enrichment will be integrated to the provided platform which is Graylog. Main objective is to ensure Security Analyst or Network Analyst to have a look at any suspicious behavior of attacks by hackers and act to it in a timely manner. Most organizations normally ingesting network and endpoint logs to the SIEM tools and integrating with some commercial tools to detect or trigger anomalies and directly send them notifications via email or 3rd party channel like Slack channel. Bear in mind that, the commercial tools is highly expensive and not really cost effective, however with this development definitely will help them to deploy the same approach with very limited budget or could be at zero cost for small medium enterprise but for big enterprise it will only cost $1500 at fixed price which considered as cheaper than the other tools. There are many developments out there whereby they are using well-known open-source IDS like Suricata and open source SIEM like elastic stack comprises of Elasticsearch, Kibana and Logstash. However, in this development, Graylog been used with the usage of Elasticsearch and MongoDB as a database server and to store, search and analyze huge volumes of data ingested. Generally, the Graylog is introduced as a powerful logging tool with a simple user-friendly interface visualized with Grafana as well as offering minimal effort to configure with very low maintenance. Due to that, creating a ruleset for Threat Hunting and Threat Intelligence enrichment, it will be much easier to configure and straight forward to compare with other competitors in the market.

Keywords: Graylog, Threat Intelligence, Threat Hunting, Centralized Log Solution, Security Monitoring Tool

ACKNOWLEDGEMENT

I would like to take this opportunity to express my gratitude and appreciation to my supervisor, Puan Zarinah Mohammed Mohaideen for her guidance, patience of guiding me whilst identified as covid-19 patient and invaluable advice throughout this project.

I also would like to express my appreciation to my family and friends for their endless support whenever I face problems. Without the mentioned parties, it is impossible for me to complete this project report successfully.

THANK YOU.

WAN IKBAL ISMAT BIN WAN KAMAL
30 AUGUST, 2021

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS	DESCRIPTIONS
SIEM	Security Information Event Management
SOC	Security Operation Centre
IR	Incident Response
EPS	Events Per Seconds
CTI	Cyber Threat Intelligence
IOC	Indicator of Compromise
IOA	Indicator of Attacks
C2	Command & Control
GB	Gigabyte
MB	Megabyte
IDS	Intrusion Detection System
NIDS	Network Intrusion Detection System
HIDS	Host Intrusion Detection System
CRM	Customer Relationship Management
LTS	Long Term Support
OPENSSH	OpenBSD Secure Shell
TOR	The Onion Router
UDP	User Datagram Protocol

CHAPTER 1

INTRODUCTION

1.1 Background to the study

Cybercriminals are getting sophisticated nowadays in terms of developing new techniques, tools and tactics in cyber hacking world. With several of tactics that is commonly known as Phishing attacks or also known as one of the techniques called Social Engineering to deliver malicious script embedded in the link and executing malware or virus or retrieving user's credentials like usernames and passwords, end users are prone to this attack every day. Further to this, one way to spot if users are impacted with these kinds of attacks, the introduction of Graylog Log Monitoring Tool and data enrichment of Threat Intelligence and Threat Hunting is useful so then the security analyst or network analyst can be fully aware of what is actually happening within their network perimeters.

Generally, the Threat Intelligence is known as threat information that is known historically. Threat intelligence data are known as artifacts generally retrieving an IP address, domains, registry key, mutex, hashes of executable files and several others. If found, all of this relevant information is called as an IOC's known as (Indicator of Compromise). As for Threat Hunting, it is an IOA's (Indicator of Attacks) that is not known historically and meant to know possible suspicious attacks residing in the host machine.

In this study of developing Graylog operated with Ubuntu operating system, there are two approaches that will be developed. The first one is Threat Intelligence enrichment powered by third parties that will be enabled and the logs of host endpoint network activities will be logged and to be received by the Graylog via Zeek Intrusion Detection System. A span network also known as port mirroring is connected from Cisco ASA firewall to Zeek Intrusion Detection System operated with Ubuntu operating system. Secondly, for Threat Hunting, a log shipper of NXLog will be installed and configured on Windows 7 machine and push the log to the Graylog directly. Once all of these enabled, the threat hunting and threat intelligence rules are created to ensure an analyst to be notified and the analyst can act or respond to this threat accordingly.

1.2 Problem Statement

The problem statement as follows:

- The evasive techniques like fileless malware attack to bypass Cyber Security tools or operating system security levels is increasing every day by hackers and therefore Threat Hunting is introduced to reduce the risk.

- A database signature from Anti-Virus or any other security tools is just a database of the sign of attacks which is known from the past and cannot detect a behaviour-based attacks on the present day by hackers. In order to overcome and reduce this risk, a Threat Intelligence is one of the ways to act as an added value tool.

1.3 Objectives of the study

Objectives of this project are:

- To configure and deploy Graylog logging tool and Zeek Intrusion Detection System for Threat Intelligence and Threat Hunting method.
- To create and enable Threat Intelligence enrichment and Threat Hunting rules, and to receive alerts from it.
- To enable the mechanism to push logs from Zeek Intrusion Detection System and Windows operating system to Graylog.
- To configure network threat types of Zeek Intrusion Detection System and NXlog from Windows operating system.
- To monitor traffic flow and windows logs, and detect suspicious activities.

Generally, an intrusion may happen externally and internally and this could resulting to data breaches of sensitive information, modifying of vital system or disrupting vital services that may impact financially to organization.

The most highly dangerous intrusion is actually within internal networking system. This commonly called as insider threats. The insider threats comprise of employees, partners, vendors, friends or customers. External intrusion is an intrusion coming outside of the in-house networking system whereby most of the attacks are coming from the internet. In fact, for this deployment, it gives value to companies of using Graylog logging tools alongside NXLog shipper and Zeek Intrusion Detection System as an added value tool.

1.4 Scope and Limitation

Graylog logging tool and Zeek Intrusion Detection System comprise of lot of scopes. In this study of configuring and deploying of these two for Threat Intelligence and Threat Hunting purposes, the scope is

- Network traffic of endpoint host activities.
- User's activities on Windows Operating System.

1.4.1 System Scope

At least two users involved in using the said system of Graylog Monitoring Tool:

i) Administrator – the administrator who have highest privilege over the system.

ii) User – normal user created by Administrator who has read, edit and write access to perform any configuration or fine tuning within the said system.

1.4.2 User Scope

Graylog	Zeek IDS	Filebeat	NXLog	VMWare ESXi
OS: Ubuntu	OS: Ubuntu	OS: Ubuntu	OS: Windows 7	OS: VMKernel
Administrator – The administrator who administered the configuration with full privilege access	root – The root access running on Ubuntu which can read,write,read+write with full privilege access configuring Zeek IDS configuration for Graylog integration	root – The root access running on Ubuntu which can read,write,read+write with full privilege access configuring Filebeat configuration for Zeek IDS integration		
User – The member who can read and edit several key functions with least privilege access	No other user created for this development	No other user created for this development	No other user created for this development	No other user created for this development

Table 1: User & Administrator Scope for System Development

1.4.3 Limitation

The major part of this development is definitely the Graylog logging tool and since we have created another user with least privilege access who can read, write and edit several key functions, it is sufficient for the user to look at necessary functions relevant to the requirements as a Security or Network Analyst.

However, the only issue is that since we are using free trial access for 30 days which given more log storage size, in the end of trial license ended, it will move back to default state which is a community edition. If this is the case, the log storage size will be decreased to smaller size and when the log reaches the limit, the old log data will be overwritten.

Additionally, for Threat Intelligence integration part, Graylog version 4.1 only provides two sources which is different than the previous version of Graylog 3.x which offers three of them. In order to create an additional custom Threat Intelligence data enrichment, it will be quite time consuming and therefore it will be not implemented during this development phase.

1.5 Implementation Plan

Gantt chart as shown below:

Figure 1: Implementation Pre-Planning

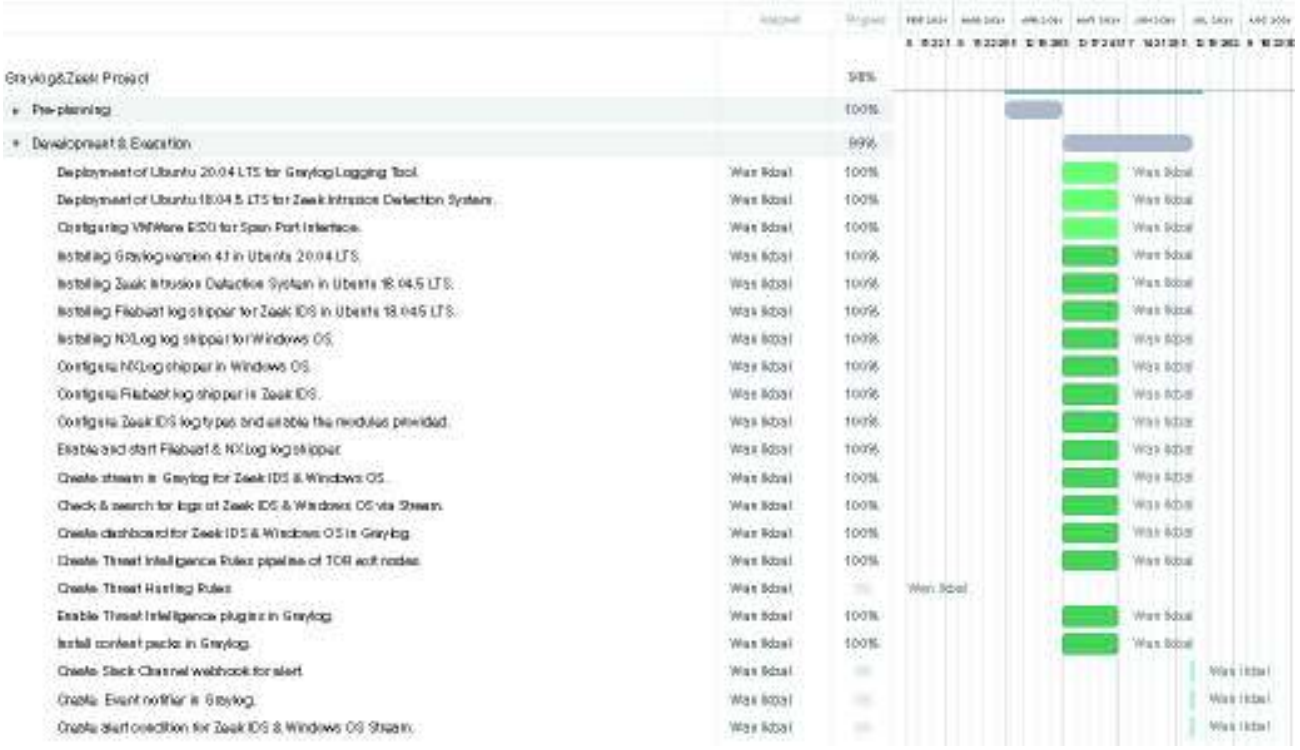

Figure 2: Development & Execution Plan

Figure 3: Pilot-Testing Plan

CHAPTER 2

LITERATURE REVIEW

2.1 Security Monitoring Tool Review

In this chapter 2.1, we will be discussing and review about the platform about Security Monitoring Tool as known as SIEM (Security Information Event Management. Take note that, most enterprise are exposed to a growing number of security threats, the visibility knowing the inside out about network traffic and user activity is a mandatory. According to (Mitkovskiy, Ponomarev and Proletarskiy, 2019), to form a holistic view of any organization's IT security is necessary to process a large volumes of log information from multitude of systems with different kind of formats that normally difficult without an automated system. In order to solve this problem, a Security Information Event Management (SIEM) been designed to pinpoint to potential security breaches.

According to (David, Amanda, David Sutton, Andy, 2020), a SIEM is to be facilitated by using an automated tools is to analyze logs, system activity and network traffic with the objective for identification and investigation. Further to this, since there's so much of log data and system activity, it is impossible for one person to monitor it all in real time and therefore the only practical solution is an automation work.

Due to that, a SIEM or also known as Security Information Event Management is introduced with the purpose to have full visibility knowing what is happening in network environment. We live in an era with Information Technology comprise of computers, servers, mobile phones, routers, switches, firewalls, clouds and many others. Each one of this assets or technology has its own threats whereby hackers can take advantage to exploit it and achieving their objectives. Further to this, let's review on several platform that implementing the same concept as Graylog Monitoring Tool (Free License). For this review, it will enlighten us to understand the differences of features introduced from other platforms.

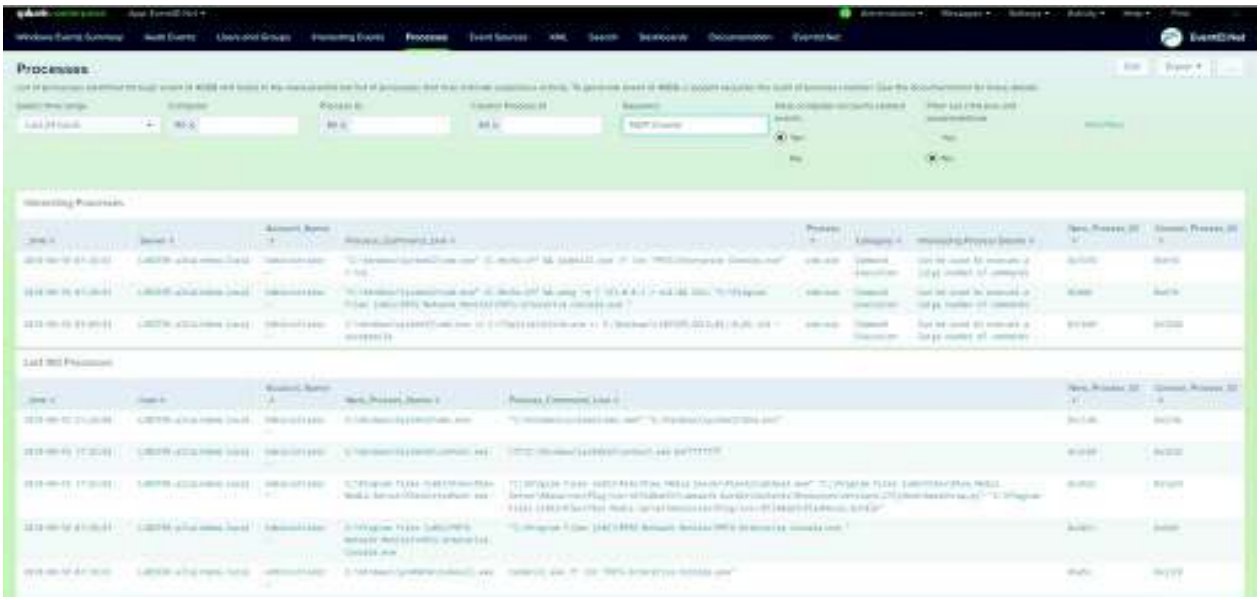

Figure 4: Splunk Log Monitoring Tool

Splunk in general is very well-known log monitoring solutions that is established since year 2003. It has nearly 90% similarities in terms of features and functions if we compare with the Graylog Monitoring Tool where it can ingest logs, allowing to search and visualizing said logs. Splunk also working well in terms of scalability and flexibility giving to its users when it comes to resizing, reducing or increasing CPU resources, RAM size or Data Storage. From here, the abilities between Graylog and Splunk is pretty similar.

Generally, the obstacles of using Splunk is that the solution given is considered as the most expensive log monitoring tool in the world. The charges normally been charged as per GB of storage logged, per node (per server), or based on unlimited plan for the highest subscription. Splunk currently only offers cloud-based deployment solution for the paid subscription and the free license of Splunk Enterprise is on premise whilst Graylog Small Business provides on-premise. Shown below a table of pricing between Splunk Enterprise (Free License) and Graylog Small Business (Free License).

Platform	Price Per Year	Deployment Type	Log Volume
Splunk Enterprise (Free License)	Free	On Premise	500 MB / Day Max
Graylog Small Business (Free License)	Free	On Premise	Up to 5GB / Day

Figure 5: Price Differences between Splunk Enterprise and Graylog Small Business

Retrieved at:

https://splunkpricing.com/, https://www.splunk.com/view/SP-CAAAAEQ & https://www.graylog.org/pricing

As we can see as shown at Figure 5 above, an identical in terms of pricing offering which given free and the same of deployment type. In fact, the offering of the log volume ingestion is slightly difference whereby Splunk Enterprise capped to only 500 MB and Graylog Small Business is giving more than that, but the solution given is similar to Graylog Small Business. Moreover, Graylog Small Business is given free of charge for the logs ingested with On-Premise deployment.

Figure 6: SIEM Elastic Log Monitoring Tool

Another competitor is based on ELK with the usage of (Elasticsearch, Logstash, Kibana) commonly known as SIEM Elastic Stack as shown at Figure 6 above, is established since year 2000.

Take note that, Graylog framework is also using Elasticsearch for its data processing as well but for MongoDB are for searching and storage purposes. Therefore, it has similarities to SIEM Elastic Stack.

Elastic is quite popular since the day it was introduced. It was first known as Elasticsearch and then changed its name to Elastic in 2015. During those years, it gives users to use Elasticsearch, Kibana and Logstash for free but the paid subscription based is quite limited. The reasons it was known as Elastic Stack because the solution is comprising of Elasticsearch, Kibana and Logstash whereby Elasticsearch is for open-source full-text search and analytics engine, Kibana is for data visualization and exploration tool and Logstash act as a server-side data processing pipeline that allows collecting data from variety of sources, transform and send to desired destination.

Nowadays, since most of users are using it for SIEM, it is known as SIEM Elastic and they are offering four types of subscriptions named as Standard, Gold, Platinum and Enterprise. As we know, the SIEM Elastic Standard (Free License) subscription is the lowest and we can compare it to Graylog Small Business (Free License) subscription accordingly as shown at Figure 7 below.

Platform	Price Per Year	Deployment Type	Log Volume
Elastic Standard (Free/Non-Free License)	Free / $679	On Premise	60 GB Max
Graylog Small Business (Free License	Free	On Premise	Up to 5GB / Day

Figure 7: Price Differences between Elastic Standard and Graylog Small Business

Retrieved at:

https://cloud.elastic.co/pricing

https://www.graylog.org/pricing

If we compare the features offering between SIEM Elastic Standard and Graylog Small Business, the SIEM Elastic Standard offer more features than the Graylog Small Business, however, the only drawbacks by SIEM Elastic Standard are that even it can be used with Free under trial plan, the subscription only offer one node of server deployment and this will lead to missing events during log shipping. If the Standard plan with standard license, the user can scale and install multiple nodes of clustering servers while Graylog Small Business can be scale with additional clustering servers for free. Moreover, since that the objective using the Log Monitoring Tool is to let user to have looks and feels of UI/UX that facilitate them using the platform. In this case, Graylog Small Business is far easier to be used compare to SIEM Elastic Standard.

Figure 8: SIEMonster Log Monitoring Tool

The last competitor of the SIEM platform is named as SIEMonster. The SIEMonster is founded in the year of 2016 and if we compare with another competitors, the SIEMonster is quite new in the market.

The SIEMonster also using Elasticsearch for data analytics engine but it is known platform for integration flexibility whereby most of the components are integrated with multiple solutions for the purpose of Security Operation Centre, Incident Response, Threat Intelligence and several others.

Additionally, the SIEMonster is yet to known worldwide because often enterprise is looking for more established SIEM solution. The pricing model for this SIEMonster is known not that cheap as well. It offers variety of solutions well integrated with another solution that makes it not that cheap nor expensive. It also gives user a free version named as Community Edition but it has limitations.

Platform	Price Per Year	Deployment Type	Log Volume
SIEMonster Community Edition (Free License)	Free	On Premise	5K EPS* = 18GB / Day
Graylog Small Business (Free License)	Free	On Premise	Up to 5GB / Day

Figure 9: Price Differences between SIEMonster and Graylog Small Business
Retrieved at: https://siemonster.com/pricing/ & https://www.graylog.org/pricing

From the price comparison as shown above at Figure 9, clearly shown that the price is given free which is identical to Graylog Small Business (Free License), the only drawbacks is that SIEMonster (Free License) only given for on-premise deployment type with limitation of five thousand EPS (Events Per Seconds). This EPS is the event of logs that is sent from data sources that is ingested to the SIEM platform.

If the data sources are the busiest device like firewalls or routers, it is not enough to receives all of the logs due to the limitation given. If it hits more than five thousand events per seconds, the rest of the log of events will be discarded and will not be received. The rough estimation is pretty simple, the higher devices monitored, the higher of Events Per Seconds will be generated.

#	Device	EPS	Count
1	Windows Servers - HIGH EPS (~50 eps)	50	20
2	Windows Servers - MED EPS (~3 eps)	3	30
3	Windows Servers - LOW EPS (~1 eps)	1	20
4	Windows Workstations	1	300
5	Windows AD Servers	10	10
6	Linux Servers	1	20
7	IBM AIX Unix Servers	2	5
8	HP-UX Unix Servers	2	3
9	Sun Solaris Unix Servers	2	3
10	IBM Mainframe / Midrange	2	1
11	Network Routers	1	7
12	Network Switches	2	1
13	Network Switches (Netflow)	30	10
14	Network Wireless LAN	5	4
15	Network Load balancers	5	6
16	WAL Accelerator	14	3

#	Device	EPS	Count
17	Other Network Devices	5	40
18	Network Firewalls (Check Point - Internal)	10	3
19	Network Firewalls (Check Point - DMZ)	50	6
20	Network Firewalls (Cisco - Internal)	10	6
21	Network Firewalls (Cisco - DMZ)	30	6
22	Network IPS/IDS	15	6
23	Network VPN	2	2
24	Network Antispam	10	5
25	Network Web Proxy	15	5
26	Other Security Devices	10	20
27	Web Servers (IIS, Tomcat, Apache)	1	20
28	Database (MSSQL, Oracle, Sybase - # of instances)	1	20
29	Email Servers (Exchange, Sandmail, etc)	2	8
30	AntiVirus Server (indicate number of AV clients)	5	300
31	Other Applications (Email, DB, AV, etc)	5	20

Figure 10: EPS Analysis and Log Volume Result

Retrieved at: http://www.aspiretss.com/tools

An example of calculation of Events Per Seconds (EPS) with numerous types of monitored devices with quantity as shown at Figure 10 above. Count means quantity of devices to be monitored and from the calculation shown above, the total of EPS is equivalent to five thousand Events Per Seconds or equivalent to 18 GB per day. In this case, if it exceeds more than that per day, the logs will be discarded. Graylog even it says up to 5GB per day, the events will still be stored and will not be discarded.

2.2 Threat Intelligence Review

In this chapter of 2.2, we will be discussing and reviewing features that related to Threat Intelligence for each of the platforms such as a Splunk Enterprise (Free License), Elastic Stack Standard (Free License), SIEMonster Community Edition (Free License) against Graylog Small Business (Free License). Also, we will be discussing what Threat Intelligence is all about.

Generally, Threat Intelligence in this context is called as Cyber Threat Intelligence where all information about the known or confirmed threats are collected from cyber world. Furthermore, this attacks information from the hackers are based on how they perform the attacks with the development phases comprise of tactics, techniques and tools in order to achieve their objective.

According to (Kure and Islam, 2019), most of the attacks nowadays more sophisticated, multi-vectored and less predictable and critical infrastructure needs a new line of security defense to control these threats and reducing risks. A Cyber Threat Intelligence provides more evidence-based information about the threats particularly to prevent threats. An example attribution of Cyber Threat Intelligence that we can use in this study for better understanding as shown below at Figure 11:

IOC Type	Terms	General Type
IP	ReceivedFromIP, Port/remoteIP	IP, Received, Destination Remote, Port
Hash	Service Item, DLL, EXE Md5sum, FileItem Md5sum	Service, dll, sha1 / md5
Domain / URL	ReceivedFromDomain/URL, Port/remoteDomain/URL	Domain/URL, Received, Destination Remote, Port
Registry	Service Item, DLL, EXE and any other malicious file extensions, File Item	WrittenFiles to Disk, Persistence Communication Destination Remote, Port
File Type	Service Item, xls, docx and any other file extensions, File Item	Drop to Disk, User Interaction or Execution Required

Figure 11: Common Types of IOC (Indicator of Compromise) in Cyber Threat Intelligence

According to (Liao, Xiaojing, Yuan, Kan, Wang, Xiaofeng, Li, Zhou, Xing, Luyi, Beyah and Raheem, 2016), with the use of these types of IOC's, it will enable organization to analyze the attack once it happens or counter during its execution. In fact, with such information of IOCs with overview of full context, it will help an organization to understand the security posture, detecting early signs of the threats and can continuously improving their security controls in place.

The table below describing the features offered in Threat Intelligence by Splunk Enterprise, Elastic Standard, SIEMonster Community Edition and Graylog Small Business;

.

Platform	Threat Intelligence Feature	Limitation
Splunk Enterprise (Free License)		Free license allows only 500 MB data per day. No Cloud deployment provided.
Elastic Standard (Free License)		The greater events generated; the more node is needed to provide High Availability to ensure zero downtime. Additional node requires paid license. Also, Elastic Threat Intelligence feature is complex to setup. No cloud deployment provided.
SIEMonster Community Edition (Free License)		Free license allows only Events Per Seconds capped to five thousand per day. No cloud deployment provided.
Graylog Small Business (Free License)		Free license allows only logs data to be stored only 90 days. No cloud deployment provided.

Figure 12: Comparison of Threat Intelligence feature between Four Competitors

Retrieved from:
https://docs.splunk.com/Documentation/Splunk/8.2.2/Admin/TypesofSplunklicenses
https://www.elastic.co/blog/establish-robust-threat-intelligence-with-elastic-security
https://siemonster.com/products-overview/
https://www.graylog.org/pricing

To summarize this, we have to ensure that the features of using Log Monitoring Tool with the use of Threat Intelligence to be met. From the Figure 12 as shown above, we know that all provided Log Monitoring providers are offering Threat Intelligence feature. However, we need an unlimited log size with unlimited data sources to ensure no events is missing after log events is shipped. In this case, it is clearly shown that Graylog Small Business has more advantages.

2.3 Threat Hunting Review

In this chapter of 2.3, we will be discussing and reviewing features that related to Threat Hunting for each of the platforms such as a Splunk Enterprise (Free License), Elastic Stack Standard (Free License), SIEMonster Community Edition (Free License) against Graylog Small Business (Free License). Also, we will be discussing what threat hunting is all about.

According to (Lemos, 2018), they found out that, with the threat hunting approach, it is giving their organization a valuable exercise in order to improve their organization's security posture. In addition, instead of waiting for an incident to occur, the approach of using threat hunting proactively will definitely help in minimizing cyber risk in their organization's network perimeter.

Overall, threat hunting is not just passively waiting for an incident to occur, but it works with many ways. One of the ways is by leveraging an IOCs (Indicator of Compromise) and query or search the relevant artifacts within the tool provided. Another way is leveraging an IOA's (Indicator of Attacks) by researching a threat actor profiles in knowing their modus operandi by looking at their tactics, techniques and procedures.

An example of formulating a hypothesis before creating a threat hunting detection logic is shown below at Figure 13.

Hypothesis	Datasources	Common Network / Event Session	Common Application Protocol
Threat actor using a C2 channel that uses common protocol on a common network port	Zeek IDS, Firewall, Netflow, NDR, Proxy, Windows Event	IP address, URL/Domain, Port, Event ID	Domain (HTTP/DNS, SSL), URL (HTTP), User Agent String (HTTP), Email Address (SMTP)

Figure 13: An example of Threat Hunting Hypothesis

From the hypothesis as shown above, we know the detection logic if we want to query a host machine running with Windows Operating System, for example it would be <ipaddress> or <domain> and <port80> or <port53> or <port 443> or port <25> or port<587>.

Herewith a feature in Threat Hunting by Splunk Enterprise, Elastic Standard, SIEMonster Community Edition and Graylog Small Business as shown below at Figure 14.

Platform	Threat Hunting Feature	Limitation
Splunk Enterprise (Free License)		With only log ingested maximum 500 MB per day. Threat Hunting will become useless due to missing events.
Elastic Standard (Free License)		The greater events generated; the more node is needed to provide High Availability to ensure zero downtime. Additional node requires paid license. Possible missing events might be occurred if the log ingested exceeds 60GB or no clustering servers in place.
SIEMonster Community Edition (Free License)		Log ingested only allows Events Per Seconds capped to five thousand per day. Threat Hunting will become useless due to missing events.
Graylog Small Business (Free License)		Free license allows only logs data to be stored only 90 days.

Figure 14: Comparison of Threat Hunting Feature between Four Competitors

As what is shown at Figure 14 shown above, under column limitation, each of the competitors has its own limitations. It seems like Graylog Small Business is giving more advantages to compare when the log can be received up to 5GB per day and log stored only 90 days. Even it only given 90 days, it is more than enough to handle any alert pertaining to threat hunting rules triggered and investigate the possible incident in a timely manner.

2.4 Intrusion Detection in General

Generally, an Intrusion Detection System (IDS) is a program or a software application embedded with a functionality to monitor a network for malicious or suspicious activities within internal network or external network. It is typically finding traffic that is deemed malicious depending on signature database that is stored in the program itself or anomalies-based to detect unknown threat with the help of machine learning. Further to this, best practice to leverage its usage is to reported or collected to the SIEM (Security Information Event Management).

Typically, there are two types of IDS:

- Host Based Intrusion Detection System (HIDS)
  This is a system where it monitors key items or types residing within operating system. Most common operating systems are Microsoft Windows, Linux, Macintosh and Unix.

- Network Intrusion Detection System (NIDS)
  This is a system where it monitors and analyzed incoming network traffic within internal networks.

Additionally, between those two types of IDS, it has its own subset of IDS types. Typically known based on anomaly based and signature based.

- Anomalies: This anomaly based is intended to identify or detect to unknown threats within internal networks which could be related to worm, viruses, malware, botnet and other attacks.

- Signature: This signature-based detecting and identifying threats by looking for specific attack's types, patterns or bytes when the IDS monitoring is in place. However, this signature based is depending on the signature database that is saved and stored in the Zeek IDS for this system development project. It is therefore easy for signature-based IDS to detect known attacks but it would not be able to detect unknown threats with the Zeek IDS.

2.5 Conclusion of Review

This is to conclude the overall review from the perspective using the Graylog Monitoring Tool instead using other competitors in the market. The main objective is to defined and prioritized:

a. Lowest cost solution
b. Ability to have a Threat Intelligence function with ease without possible issues
c. Ability to have a Threat Hunting function with ease without possible issues

From the reviewed results, it is clearly shown that Elastic Stack Standard (Free License) is the close competitor to Graylog Small Business (Free License) whereby they are offering more features other than Threat Intelligence and Threat Hunting. In fact, the Elastic Stack Standard is offering more log storage capped at 60GB maximum whilst the Graylog Small Business is offering up to 5GB per day.

As for SIEMonster Community Edition (Free License) and Splunk Enterprise (Free License) are only limited based on 5000 events per seconds generated and daily capped ingested log volume at 500 MB per day. This is too limited and low to compare with Elastic Stack Standard and Graylog Small Business. Due to this, this are the only drawbacks for these two platforms

Putting two competitors in place Elastic Stack Standard and Graylog Small Business, one of the biggest issues for Elastic Stack Standard are that, the complexity of integrating with Threat Intelligence is very complex to setup. It does not provide one click function for Threat Intelligence function. The only good feature from Elastic Stack Standard is the usage of Threat Hunting features, whilst the Graylog Small Business setup is pretty straightforward and simple under one single click for both features.

Overall, the Graylog Log Monitoring with the subscription of Small Business (Free License) has more advantages to compare with other competitors. Even the other competitors offering the same features for Threat Intelligence and Threat Hunting, the main limitations are the setup complexity, log retention, events per seconds generated and maximum log storage.

CHAPTER 3

SYSTEM ANALYSIS AND DESIGN

3.1 Feasibility Studies

There are about four feasibility study that were defined in this system development:

1. Technical

Technical wise, in this context, the Graylog Monitoring Tool platform are downloaded and installed from Graylog official website on Ubuntu Operating System. Once downloaded and installed, the platform can be accessed directly through web-based platform. Logging in as user is simple similar to any other web-based system that has username and password logging area.

As for Zeek IDS, this will be downloaded and installed on Ubuntu Operating System separately from Graylog Monitoring Tool. This Zeek IDS will act as a traffic receiver from Cisco ASA firewall. Typically, due to Zeek IDS instance will be running through virtual machine instance under VMWare, the Zeek IDS will need to enable promiscuous mode in order to receive traffic from Cisco firewall. It is indirectly act as a sniffing device.

The Cisco ASA Firewall itself, will need to enable the port mirroring which also known as switch port mirroring. In this stage, the Cisco ASA Firewall is physically connected to the VMWare ESXi and requires configuration virtually.

Lastly, the NXLog will be installed on the Windows Operating System in order to send logs to the Graylog Monitoring Tool platform.

2. Scope

This system is developed with the objective to use and leverage the method of Threat Intelligence and Threat Hunting with the suggested Log Monitoring Tool platform named as Graylog Log Monitoring Tool. Due to that, a pre-assessment is required in order to choose the best platform before go to on-production phase. However, in this

development stage, the final stage only can be done until pilot-testing phase. Moreover, this development stage is developed in a smaller scale.

From here, this can be taken as a prototype for the said solution of Log Monitoring Tool with Threat Intelligence and Threat Hunting. Overall, from this development, it will only evolve the website development and other integrations between Graylog Log Monitoring, Zeek IDS and NXLog which are provided by respective entities. Further to this, we will need to perform installation, configuration and fine tuning by our own. Even though, the steps may look quite complex, but we do have a guideline provided by those three entities for reference purposes.

3. **Economical**

   The development of Graylog Monitoring Tool for the purpose by leveraging Threat Intelligence and Threat Hunting are to gives insights and awareness to user or customer who might be looking for low cost-effective solutions focusing in SIEM. Whilst, in the end of the development stage, it can be considered as a pre-assessment stage before making final decision in selecting appropriate platform.

   Prior developing the said platform, research had been conducted and by leveraging our own expertise in building up the said platform, the cost of development can be minimal. During this development stage, we use our own company VMWare ESXi hyper-virtualization as a hosted machine for Graylog Monitoring Tool and Zeek IDS, whilst the Windows Operating System acted as an endpoint windows machine to send the logs to the said platform.

   Looking at the bigger picture, we can conclude that, to create and build this platform can be considered as zero cost. Only if, the user or customer do not have the VMWare ESXi hyper-virtualization to run a virtual machine, it still can be considered as low cost due to that the license of the said platform is given free.

4. **Time**

   This project development is divided into three phases;

   - Implementation Pre-Planning – For research
   - Development and Execution Plan – For development and execution

- Pilot-Testing Plan – For testing of the said platform

Referring to the implementation pre-planning for research purposes, it took about at least 2 months to complete. This is to ensure that we have a solid understanding before developing the said platforms leveraging Threat Intelligence and Threat Hunting.

As for development and execution plan for the said platform, it took us about 3 months. During the development stages, we had encountered numerous types of non-functional services which includes system errors, application errors, downtime due to improper configuration and several other issues.

Moreover, it also requires us to complete the configuration phases that most of the consumed was coming from Graylog Monitoring Tool and Zeek IDS.

Ultimately, the final phase is the pilot-testing plan. From here, we have to ensure that the key functions are workable as intended and as designed. This includes the configuration, validation and testing for Threat Intelligence as well as Threat Hunting. All of these key function's coverage are coming from the Graylog Log Monitoring, Zeek IDS, NXLog and Slack Alert Channel for notification and it is therefore, requires close attention to these three entities.

3.2 System Requirement

Graylog	Zeek IDS	Filebeat	NXLog	VMWare ESXi
OS: Ubuntu 20.04 LTS	OS: Ubuntu 18.04.5 LTS	OS: Ubuntu 18.04.5 LTS	OS: Windows 7 Professional	OS: VMKernel
Store traffic and raw logs from host endpoint	Received traffics from span port from VMWare ESXi	Forward and centralizing log data from Zeek IDS to Graylog	Forward and centralizing log data from Windows	

			machine to Graylog	
Detect & improve system security	Installed on Ubuntu OS running with Filebeat agent and Zeek IDS provide real time traffic analysis	Installed on Ubuntu OS running with Zeek IDS as agent and monitor log files or location specified	Deployed on Windows 7 operating system with NXLog agent and monitor log files or location specified	Port mirroring sending traffic to Zeek IDS instance
Regulatory compliance	Deep packet inspection for each network logs	Collect log events and forwards them to Graylog elasticsearch for indexing	Collect Windows log events and forwards them to Graylog elasticsearch for indexing	
Compliance with security policies	Network data logging			

Figure 15: System Requirements List to Develop the Platform

Referring to the Figure 15 shown above, in this deployment method, about two instances are required for Zeek Intrusion Detection System and Graylog Logging Tool which is separately installed. The Graylog operated with Ubuntu 20.04 LTS and Zeek Intrusion Detection System altogether with Filebeat on the same machine operated with Ubuntu 18.04.5 LTS, and these two instances running in VMWare ESXi.

These virtual instances are deployed within internal networks. The Zeek Intrusion Detection System are installed with two network interfaces, one is for management with private IP address and another one is listening private IP address in order to receive span traffic from Cisco ASA firewall. Shown below Figure 16, a network diagram for basic understanding.

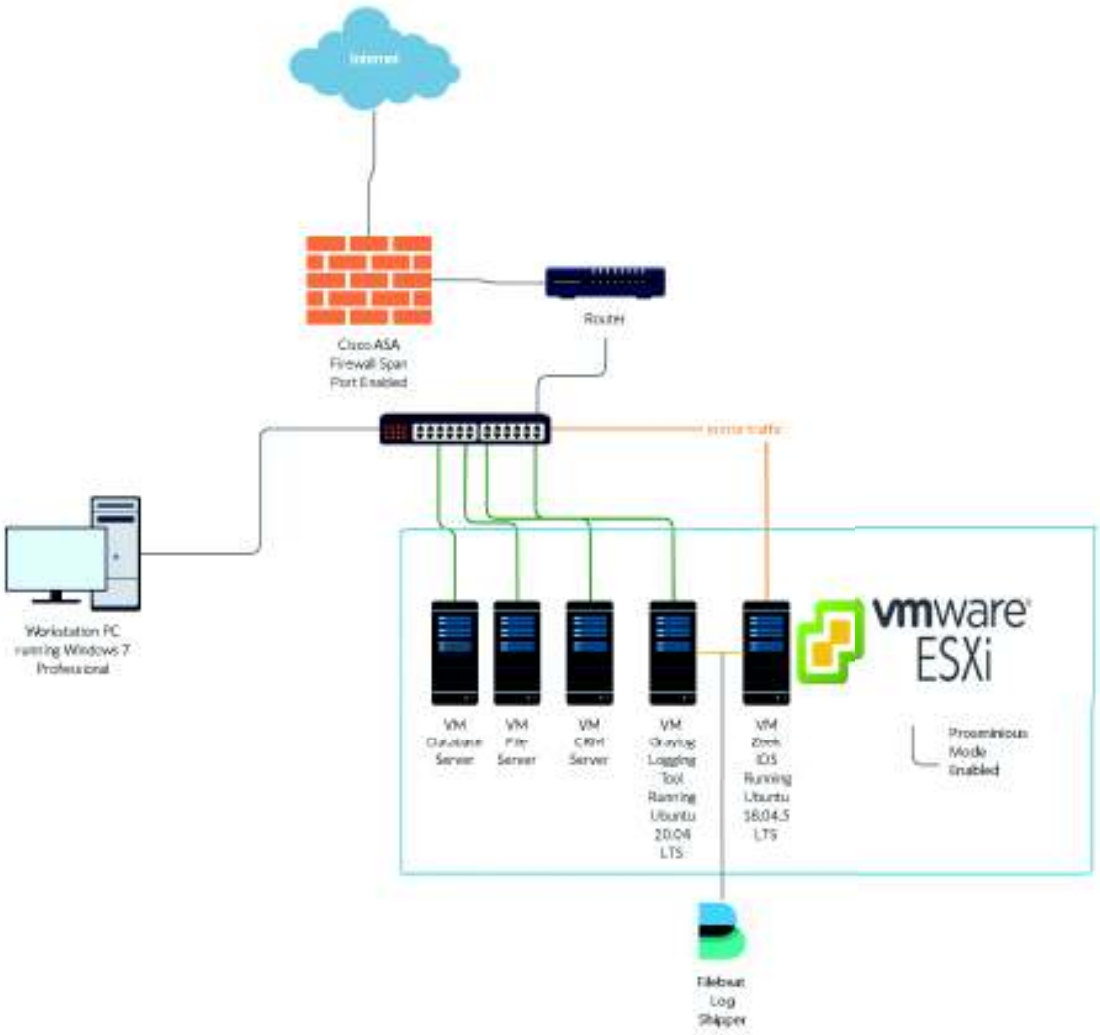

Figure 16: An overview of Network Architecture Diagram for the Platform

Based on Figure 16 above, this is basic overview to understand the objective of using the Zeek IDS and Graylog Logging Tool. The Zeek IDS are receiving data via mirror traffic from device known as Cisco ASA Firewall and it will route through router and switch.

Therefore, in this development phase, we relying on the system development of Zeek IDS and hardware dependent device known as VMWare ESXi, a hyper virtualization for virtual machines. The VMWare ESXi is capable allowing virtual machine to allow promiscuous mode to receive the mirror traffic from Cisco ASA Firewall. In computer networking, this mode is a mode for wired network interface controller that causes the controller to pass all traffic it receives to the destination network interface.

For this situation, the Zeek IDS is created under VMWare ESXi as a virtual machine and setup the promiscuous mode to receives the traffic from Cisco ASA Firewall. From the architecture shown above, the Cisco ASA Firewall is the heart of the internal network perimeter and the one to handles, receives and seeing every traffics internally and externally from all other devices within the internal networks.

To have a complete understanding how the devices play the roles based on the diagram shown above at Figure 16, have a look at the table shown below at Figure 17:

Device Name	Sent Traffic To	Received Traffic From
VM: Graylog Logging Tool	Not Sending Any	VM: Zeek IDS via Filebeat Log Shipper and Workstation PC
VMWare ESXi	Not Sending Any	Cisco ASA Firewall
VM: Zeek IDS	VM: Graylog Logging Tool	
Cisco ASA Firewall	VM: Zeek IDS via mirror traffic	Workstation PC, VM Database Server, VM File Server, VM CRM Server
Workstation PC	VM: Graylog Logging Tool	Not Sending Any

Figure 17: Basic Overview Traffic Communication between Devices

Referring to Figure 17 shown above, the device named VM:Graylog Logging Tool is not given privilege to send any traffics because it only acts as a SIEM and it is therefore only work as a standalone device to receive data logs that is sent from VM:Zeek IDS via Filebeat and Workstation PC.

The VMWare ESXi, is not sending any traffic but only receive port mirror of traffic that is coming from Cisco ASA Firewall and then VM: Zeek IDS will be sending the data logs to VM: Graylog Logging Tool.

As for Cisco ASA Firewall, its role is to send the traffic mirror to VM: Zeek IDS and receive traffic from other devices such as Workstation PC, VM Database Server, VM File Server and

VM CRM Server. This is because we know that Cisco ASA Firewall is the heart of the internal networks that can see every traffics inbound and outbound connections from other devices.

Lastly, the Workstation PC logs will be sent to VM: Graylog Logging Tool and not receiving traffic from other devices.

3.3 System Development Method

For this system development of Log Monitoring Tool with Threat Intelligence and Threat Hunting, we were using a complete Open-Source system like Ubuntu based on Debian, Elasticsearch, MongoDB, Zeek, Logstash through Filebeat.

Ubuntu is for operating system-based Debian. Elasticsearch is used to run with Graylog. MongoDB is to support Graylog for database. Zeek is for intrusion detection system and Filebeat with Logstash is to ship the Zeek IDS log events to Graylog.

All of these are installed separately and manually.

3.4 System Design Attack Based

Here is the system design architecture for this system development involving attacker. A close example whereby hacker could potentially perform an attack to Workstation-PC and the rest of the servers and compromise them with several ways.

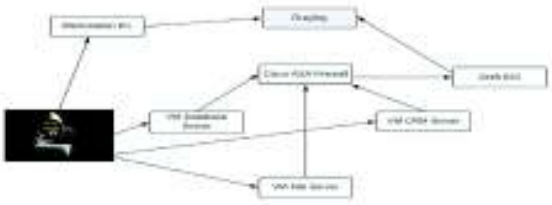

Figure 18: System Design Attack Based

CHAPTER 4

SYSTEM IMPLEMENTATION AND TESTING

4.1 System Guides / Manual

After the implementation of pre-planning is clearly defined comprising of information collected from the research phase and expertise requirements especially in hands-on skills, in order to develop the said platform, we are ready to proceed for the next stage which is under development and execution phase.

The system development steps are structured with the following order:

- Ensure the Ubuntu operating system up and running. One for Graylog and another one for Zeek IDS.
- Download, install and configure Graylog.
- Download, install and configure Zeek IDS and Filebeat.
- Download and install NXLog agent in Workstation PC running with Windows 7 Professional. Also, configure the types of system events required.
- Configure and fine-tuning data sources required from Zeek IDS and NXLog in Graylog.
- Enable and configure the feature for Threat Intelligence in Graylog.
- Create and test Threat Intelligence and Threat Hunting rules in Graylog.
- Create and test Slack Alert Channel for notifications.

4.2 Installation Manual

Firstly, we will make sure that the Ubuntu Operating System 20.04 LTS is installed for Graylog, refer Figure 20. Once installed in ESXi, we can see that the host is up and running, refer Figure 19.

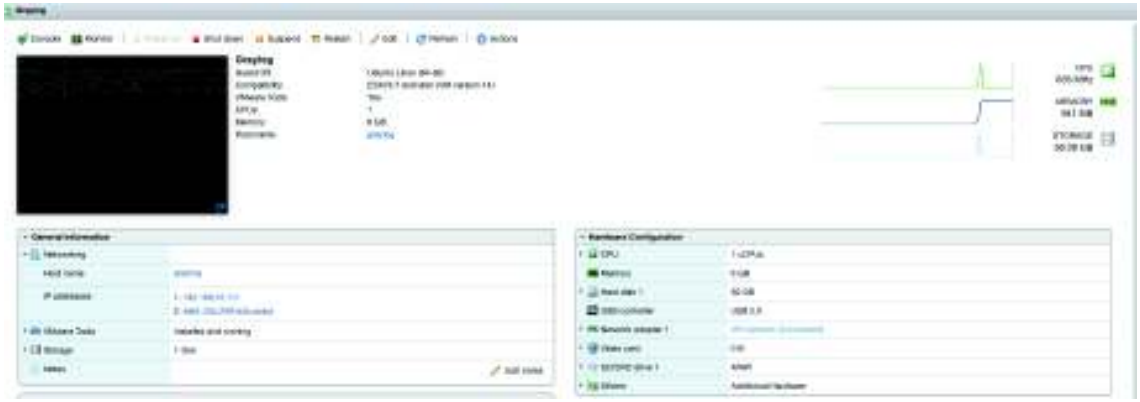

Figure 19: Graylog Virtual Machine Instance

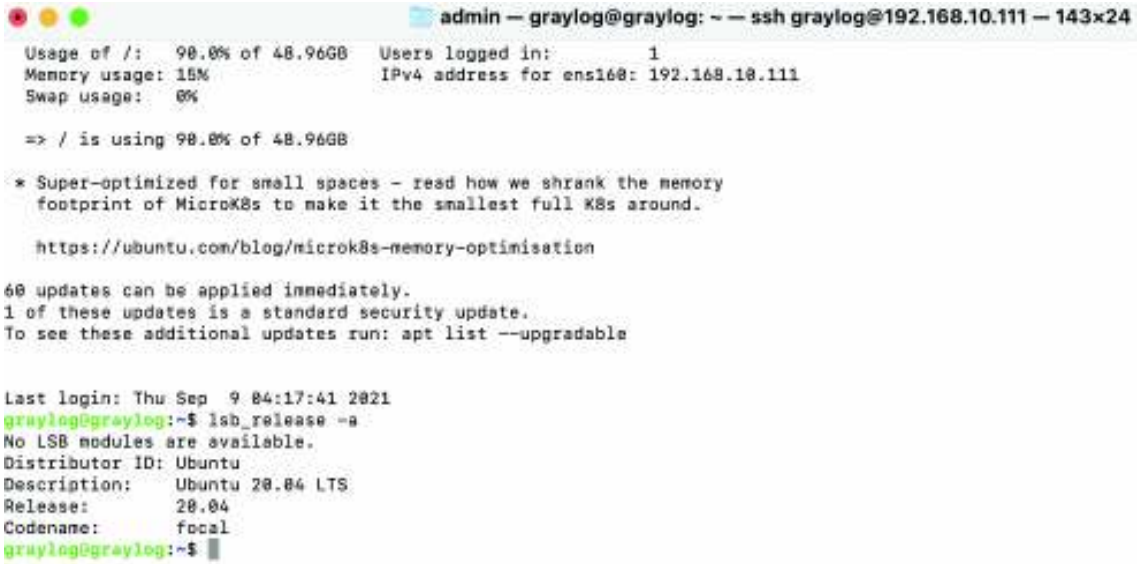

Figure 20: Graylog Instance Running with Ubuntu 20.04 LTS

Secondly, we will make sure that the Ubuntu Operating System 18.04.5 LTS is installed for Zeek IDS, refer Figure 22. Once installed in ESXi, we can see that the host is up and running, refer Figure 21.

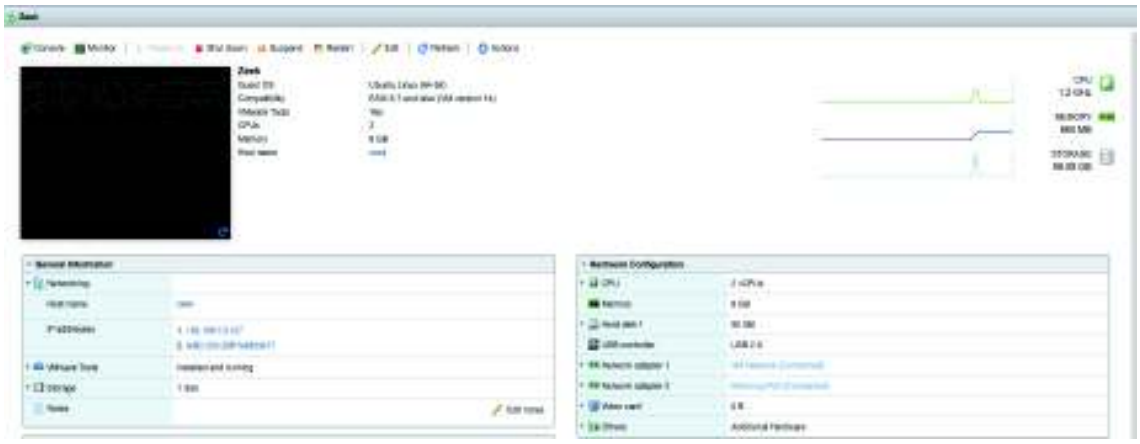

Figure 21: Zeek Virtual Machine Instance

Figure 22: Zeek Instance Running with Ubuntu 18.04.5 LTS

4.2.1 Download, Install and Configure Graylog

When the Graylog been installed with the Ubuntu OS, we need to access the instance with ease. In order to do that, we needed to perform SSH to the instance from the host machine. In this situation, we were using Macintosh Macbook. Therefore, in Graylog instance, we have installed OpenSSH so then we can perform SSH to the instance from Macbook terminal.

In the instance, we performed the following command to install OpenSSH.

```
$ sudo apt update
$ sudo apt install openssh-server

Press Y to continue with the installation
```

Then, we can perform SSH from the Macbook machine to the instance.

After successfully logged into Graylog instance via SSH. We started downloading and installed Graylog version 4.1.1 within the instance.

For the Graylog instance, running with Ubuntu 20.04 LTS, we chose DEB / APT options to install.

```
$ sudo apt-get install apt-transport-https
$ wget https://packages.graylog2.org/repo/packages/graylog-4.1-
repository_latest.deb
$ sudo dpkg -i graylog-4.1-repository_latest.deb
$ sudo apt-get update
$ sudo apt-get install graylog-server -y
```

Then, before starting the Graylog, we had to make sure the prerequisites are installed. Graylog 4.1.1 are depending and supported on Java version 8 and above, Elasticsearch version 6.x or 7x and MongoDB version 4, 4.2 or 4.4 only.

Before proceeding with the said dependencies, we installed Java for Graylog with the following command:

```
$ sudo apt-get install openjdk-11-jre-headless -y
```

We will also require to generate a secret to secure password for Graylog user.

```
$ pwgen -N 1 -s 96
```

And we got the following output:

```
oefAXKFoU93q18b2H7tLrOgATaB2wOCDJlIiP3fA5yO8dqa7ENLv7zg4e8epuF4pOqQQVVRmoArMIvtpSLMHXdvFo4fHOdVq
```

Next thing, is to generate a password for the admin in order to login into the Graylog web system. From here, we generated a secure password with the following command:

```
$ echo -n Cr@cker42 | sha256sum
```

And we got the following output:

```
973e1b07645e98444bc9491e3a568f0327649ee57ec4a71a77fc0e0fad132d1e
```

Then, we edit the existing configuration file and put those two passwords.

Password_secret =

```
oefAXKFoU93q18b2H7tLrOgATaB2wOCDJlIiP3fA5yO8dqa7ENLv7zg4e8epuF4pOqQQVVRmoArMIvtpSLMHXdvFo4fHOdVq
```

Root_password_sha2 =

```
973e1b07645e98444bc9491e3a568f0327649ee57ec4a71a77fc0e0fad132d1e
```

Then, saved and exited the changes.

Also, in order for allowing us to access the web system, we had to bind the local server with the propose port number:

```
$ http_bind_address = localhost:9000
```

Then, save and exit the changes. After that, we had to start the Graylog service and enabled it accordingly with the following command:

```
$ sudo systemctl daemon-reload
$ sudo systemctl start graylog-server.service
$ sudo systemctl enable graylog-server.service
```

After performing those commands, we had to verify and validate whether the Graylog is working as intended with the following command:

```
$ sudo systemctl status graylog-server.service
```

And the output was like this:

```
graylog@graylog:~$ sudo systemctl status graylog-server.service
● graylog-server.service - Graylog server
     Loaded: loaded (/lib/systemd/system/graylog-server.service; enabled; vendor preset: enabled)
     Active: active (running) since Thu 2021-09-09 04:05:36 UTC; 4h 18min ago
       Docs: http://docs.graylog.org/
   Main PID: 781 (graylog-server)
      Tasks: 175 (limit: 7029)
     Memory: 1.4G
     CGroup: /system.slice/graylog-server.service
             ├─ 781 /bin/sh /usr/share/graylog-server/bin/graylog-server
             └─1087 /usr/bin/java -Xms1g -Xmx1g -XX:NewRatio=1 -server -XX:+ResizeTLAB -XX:-OmitStackTraceInFastThrow -Djdk.tls.

Sep 09 04:05:36 graylog systemd[1]: Started Graylog server.
```

Now for Elasticsearch, we started installing and configuring Elasticsearch. With this case, we installed Elasticsearch version 7.x.

```
$ sudo wget -qO - https://artifacts.elastic.co/GPG-KEY-elasticsearch | apt-
key add -
$ echo "deb https://artifacts.elastic.co/packages/oss-6.x/apt stable main" |
sudo tee -a /etc/apt/sources.list.d/elastic-6.x.list
$ sudo apt-get update -y$
$ sudo apt-get install elasticsearch-oss -y
```

After installing Elasticsearch, we edited the Elasticsearch config file and inserted the cluster name.

```
$ sudo nano /etc/elasticsearch/elasticsearch.yml

Insert the cluster name and following settings to:

cluster.name: graylog
action.auto_create_index: false
```

After save and close the config file, we started the Elasticsearch service and enabled it.

```
$ systemctl daemon-reload
$ systemctl start elasticsearch.service
$ systemctl enable elasticsearch.service
```

From here, we needed to ensure the service is up and running without any error.

```
$ systemctl status elasticsearch.service
```

Then, to verify whether if we send request to Elasticsearch, we will be able to receive the response, if there is no response, the installation is consider failed.

```
$ curl -X GET http://localhost:9200
```

```
graylog@graylog:~$ curl -X GET http://localhost:9200
{
  "name" : "graylog",
  "cluster_name" : "graylog",
  "cluster_uuid" : "rkF2QnZBR8yJTEfIqIxujQ",
  "version" : {
    "number" : "7.13.3",
    "build_flavor" : "default",
    "build_type" : "deb",
    "build_hash" : "5d21bea28db1e89ecc1f66311ebdec9dc3aa7d64",
    "build_date" : "2021-07-02T12:06:10.804015202Z",
    "build_snapshot" : false,
    "lucene_version" : "8.8.2",
    "minimum_wire_compatibility_version" : "6.8.0",
    "minimum_index_compatibility_version" : "6.0.0-beta1"
  },
  "tagline" : "You Know, for Search"
}
graylog@graylog:~$
```

Next, we installed MongoDB, because Graylog requires MongoDB as a database and works well with it. The following command:

```
$ sudo apt-get install mongodb-server -y
```

Once installed, we started the MongoDB and enabled it.

```
$ sudo systemctl start mongod.service
$ systemctl enable mongod.service
```

From here, we needed to ensure the service is up and running without any error.

```
$ sudo systemctl status mongod.service
```

And the output was like this:

```
graylog@graylog:~$ sudo systemctl status mongod.service
● mongod.service - MongoDB Database Server
     Loaded: loaded (/lib/systemd/system/mongod.service; enabled; vendor preset: enabled)
     Active: active (running) since Thu 2021-09-09 04:05:36 UTC; 4h 37min ago
       Docs: https://docs.mongodb.org/manual
   Main PID: 785 (mongod)
     Memory: 190.4M
     CGroup: /system.slice/mongod.service
             └─785 /usr/bin/mongod --config /etc/mongod.conf

Sep 09 04:05:36 graylog systemd[1]: Started MongoDB Database Server.
graylog@graylog:~$
```

4.2.2 Download, Install and Configure Zeek IDS and Filebeat

Similar to items 4.2.1, we installed OpenSSH in this Zeek IDS instance with the following command.

```
$ sudo apt update
$ sudo apt install openssh-server

Press Y to continue with the installation
```

Then, we can perform SSH to the instance from Macbook terminal.

```
admin — zeek@zeek: ~ — ssh zeek@192.168.10.127 — 143×24
admins-MacBook-Pro:~ admin$ ssh zeek@192.168.10.127
zeek@192.168.10.127's password:
```

After successfully logged into Zeek instance via SSH. We started downloading, installing and configuring Zeek IDS within the instance.

For the Zeek instance running with Ubuntu 18.04.5 LTS, we performed the update and did some upgrade with the following command.

```
$ sudo apt-get update && sudo apt-get dist-upgrade
```

Then, to install Zeek, we needed to download and install few dependencies before installing, with the following commands:

```
$ sudo apt-get install cmake make gcc g++ flex bison libpcap-dev libssl-dev
python3 python3-dev swig zlib1g-dev
```

In order to download Zeek, we needed to download it from Zeek official GitHub. Therefore, we installed git with the following command.

```
$ sudo apt-get install git
```

Then we cloned the Zeek from Zeek GitHub with the following command:

```
$ sudo git clone --recursive
https://github.com/zeek/zeek
```

Then, we used cd command to the Zeek directory after downloaded from GitHub, then we built and installed the Zeek IDS with this command:

```
$ sudo ./configure && make && sudo make install
```

After built and installation is complete, we had to configure the run-time environment for the path environment, so then when we executed the Zeek the platform or shell and it were able to run under that specific path. Therefore, we had to open the following configuration file with the command:

```
$ sudo nano ~/.bashrc
```

Then, we have inserted the shell syntax, save and exit after changes.

```
Export PATH=/usr/local/zeek/bin:$PATH
```

Then, we had to manage the node for Zeek. In our case, we are only depending to one server for Zeek, it is therefore, we needed the ZeekControl to be installed as standalone. ZeekControl is an interactive shell for operating or managing Zeek on the system.

Thus, we needed to configure something inside node configuration file for ZeekControl with the following command and things to change.

```
$ sudo nano /usr/local/zeek/etc/node.cfg
```

Then, we changed the interface to monitor the traffic. This is mirror traffic that we received from Cisco ASA Firewall.

```
[zeek]
type=standalone
host=localhost
interface=ens192   # change this according to your
listening interface in ifconfig
```

In this case, our interface was ens192. Then changed it and saved accordingly.

After saving the node config file. We needed to enable the interface ens192 in order to receive mirror traffic with the following command.

```
$ sudo ip link set dev ens192 up
```

To verify whether the ens192 was up or not with the following interfaces, we needed to perform this command:

```
$ ip link
```

Then the results shown as:

```
zeek@zeek:~$ ip link
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN mode DEFAULT group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
2: ens160: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP mode DEFAULT group default qlen 1000
    link/ether 00:0c:29:88:b9:17 brd ff:ff:ff:ff:ff:ff
3: ens192: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP mode DEFAULT group default qlen 1000
    link/ether 00:0c:29:88:b9:21 brd ff:ff:ff:ff:ff:ff
```

So, from here we know that ens160 was the management interface or known as localhost interface and ens192 was the mirroring interface to receive mirror traffic from Cisco ASA Firewall.

From here, we had to start the ZeekControl and executed this command:

```
$ sudo /usr/local/zeek/bin/zeekctl
```

Then we executed this command by order:

```
install
```

1

```
deploy
```

2

After deploying, we needed to check the status of the Zeek by executing:

```
[ZeekControl] > status
Name         Type       Host          Status    Pid    Started
zeek         standalone localhost     running   16562  09 Sep 10:24:18
```

Then, the next step is to define what kind of logs that we should select once the mirror traffic ingested into Zeek. As we know, the mirror traffic that is sent by Cisco ASA Firewall might be coming from multiple devices within the network. It is therefore important for us to defined what kind of logs that is critical for monitoring.

As we know, the Zeek is already enabled and now we needed to see what kind of logs Zeek IDS received thus far.

By using cd command in ubuntu to the path:

```
$ cd /usr/local/zeek/logs/current
```

We see the logs as shown below:

```
zeek@zeek:/usr/local/zeek/logs/current$ ls
capture_loss.log  dhcp.log  files.log  known_hosts.log     notice.log  ocsp.log  software.log  ssl.log    stderr.log  syslog.log  x509.log
conn.log          dns.log   http.log   known_services.log  ntp.log     sip.log   ssh.log       stats.log  stdout.log  weird.log
zeek@zeek:/usr/local/zeek/logs/current$
```

It is decided that we chose the capture_loss log, conn.log, dce_rpc.log, dhcp.log, dns.log, dpd.log, files.log, ftp log, http.log, intel.log, kerberos.log, mysql.log, ntlm.log, oscp.log, pe.log, rdp.log, smb_files.log, socks.log, ssh.log, ssl.log, syslog.log, tunnel.log and weird.log.

Even some of these logs were not in the current logs, it still can be ingested to Graylog Logging Tool when there is an activity pertaining to that network protocol occurs. However, this is depending on the Filebeat module that we needed to configure.

After this stage, we had to install a log shipper, in this case, we installed a Filebeat log shipper. Filebeat does have the Zeek module after the installation. We executed this command to download Filebeat, installed and enabled them.

```
$ sudo apt-get update && sudo apt-get install filebeat
$ sudo systemctl enable filebeat
$ sudo filebeat modules enable zeek
```

The next stage, it requires us to configure the Zeek module in Filebeat path. We executed and modified them accordingly. Then we saved and closed it.

```
$ sudo nano /etc/filebeat/modules.d/zeek.yml
```

The modified configuration in the zeek.yml as shown below:

```
- module: zeek
  capture_loss:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/capture_loss.log"]
  connection:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/conn.log"]
  dce_rpc:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/dce_rpc.log"]
  dhcp:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/dhcp.log"]
  dns:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/dns.log"]
  dpd:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/dpd.log"]
  files:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/files.log"]
  ftp:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/ftp.log"]
  http:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/http.log"]
  intel:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/intel.log"]
  kerberos:
    enabled: true
    var.paths: ["/usr/local/zeek/logs/current/kerberos.log"]
```

```
mysql:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/mysql.log"]
ntlm:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/ntlm.log"]
ocsp:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/ocsp.log"]
pe:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/pe.log"]
rdp:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/rdp.log"]
smb_files:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/smb_files.log"]
socks:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/socks.log"]
ssh:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/ssh.log"]
ssl:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/ssl.log"]
syslog:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/syslog.log"]
tunnel:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/tunnel.log"]
weird:
  enabled: true
  var.paths: ["/usr/local/zeek/logs/current/weird.log"
```

Zeek has its own policy of translating log extension to another extension and also it does provide policy tuning in it. Thus, in order to secure all protocols to be monitored by Zeek, we had to add all types of alerts in Zeek config file and change the log types received to. json format. We executed and added this up by the following:

```
$ sudo nano /usr/local/zeek/share/zeek/site/local.zeek
```

```
@load alert-all-notices
@load policy/tuning/json-logs.zeek
```

Then saved and closed.

Lastly, we needed to configure the Filebeat.yml file to change from default path to the destination path logs for retrieval as well as the destination IP of Graylog. We executed and made changes of the following:

```
$ sudo nano /etc/filebeat/filebeat.yml
```

```
Paths: - /usr/local/zeek/logs/current/*.log
Output.logstash:
Hosts: ["192.168.10.111:5044"]
```

Then we restarted the Filebeat.

```
$ sudo systemctl restart filebeat
```

After all configurations been made, the final stage is to deploy new config through ZeekControl.

Executed.

```
$ sudo /usr/local/zeek/bin/zeekctl
```

```
[ZeekControl] > deploy
checking configurations ...
installing ...
removing old policies in /usr/local/zeek/spool/installed-scripts-do-not-touch/site ...
removing old policies in /usr/local/zeek/spool/installed-scripts-do-not-touch/auto ...
creating policy directories ...
installing site policies ...
generating standalone-layout.zeek ...
generating local-networks.zeek ...
generating zeekctl-config.zeek ...
generating zeekctl-config.sh ...
stopping ...
stopping zeek ...
starting ...
starting zeek ...
[ZeekControl] >
```

4.2.3 Download, Install and Configure NXLog Agent for Workstation PC

For Windows computer named as Workstation PC, this is the one requires an agent installation. This agent based is called as NXLog. This NXLog is a log shipper and look similar to Filebeat. It does have a configuration file that needs to be modified and some fine tuning of log types. We needed this to be modified to ensure appropriate logs is ingested into Graylog Log Monitoring Tool.

The Workstation PC is running with Windows 7 Professional and its running as a virtual machine. We configured this Windows machine with remote desktop enabled. We need to remotely control this machine from our Macbook machine and installed the NXLog.

Once we downloaded the NXLog from https://nxlog.co/system/files/products/files/348/nxlog-ce-2.11.2190.msi , we installed and modified few things inside the NXLog config file. The following changes as shown below.

```
define ROOT     C:\Program Files (x86)\nxlog
define CERTDIR  %ROOT%\cert
define CONFDIR  %ROOT%\conf
define LOGDIR   %ROOT%\data
define LOGFILE  %LOGDIR%\nxlog.log
LogFile %LOGFILE%

Moduledir %ROOT%\modules
CacheDir  %ROOT%\data
Pidfile   %ROOT%\data\nxlog.pid
SpoolDir  %ROOT%\data

<Extension _syslog>
    Module      xm_syslog
</Extension>

<Extension _charconv>
    Module      xm_charconv
    AutodetectCharsets iso8859-2, utf-8, utf-16, utf-32
</Extension>

<Extension _exec>
    Module      xm_exec
</Extension>
```

```
<Extension _fileop>
   Module      xm_fileop

   # Check the size of our log file hourly, rotate if larger than 5MB
   <Schedule>
      Every   1 hour
      Exec    if (file_exists('%LOGFILE%') and \
              (file_size('%LOGFILE%') >= 5M)) \
               file_cycle('%LOGFILE%', 8);
   </Schedule>

   # Rotate our log file every week on Sunday at midnight
   <Schedule>
      When    @weekly
      Exec    if file_exists('%LOGFILE%') file_cycle('%LOGFILE%', 8);
   </Schedule>
</Extension>

########################################################
############### Extensions ###############################

<Extension _gelf>
   Module      xm_gelf
</Extension>
########## INPUTS ###########
<Input in>
   Module      im_mseventlog
# For windows 2003 and earlier use the following:
#   Module      im_mseventlog

</Input>
########################################
################## OUTPUTS ##############
<Output graylog>
   Module      om_udp
   Host        192.168.10.111
   Port        3514
   #Exec       to_syslog_snare();
   OutputType  GELF

</Output>
########################################
#################### ROUTE  ###########
<Route 1>
   Path in => graylog
</Route>
```

From the config file as shown above, the most important things are the inputs from xm_gelf, im_mseventlog and the Graylog output. The input conditions are the one that NXLog will collect from the Windows 7 machine and send the logs to the output of Graylog. From here we can see the most important module is im_mseventlog. This is the Microsoft event logs that we normally see in Windows event viewer. It consists of many information pertaining to the application, system and security.

For the output of Graylog here, the events will be sent to the Graylog Monitoring Tool destination host IP through UDP protocol.

4.3 Testing Plan and Test Output

The following are the steps of testing and the output in order to receive logs from Zeek IDS and Windows machine which also requires fine tuning once Graylog able to receive from those two sources.

a. Open Web System at https://192.168.10.111:9000 and log in as admin
b. Open System / Inputs, then go to Inputs
c. Select Beats for Zeek IDS and GELF UDP for Windows machine to launch new input
d. Enter the following information for both sources as shown below.
    i. Tick Global
    ii. For Windows insert as Windows-input and Zeek as Zeek
    iii. Bind address leave as 0.0.0.0
    iv. Port for Windows is 3514 and Zeek is 5044
    v. And click save

The two data sources we configured for the above is shown as below:

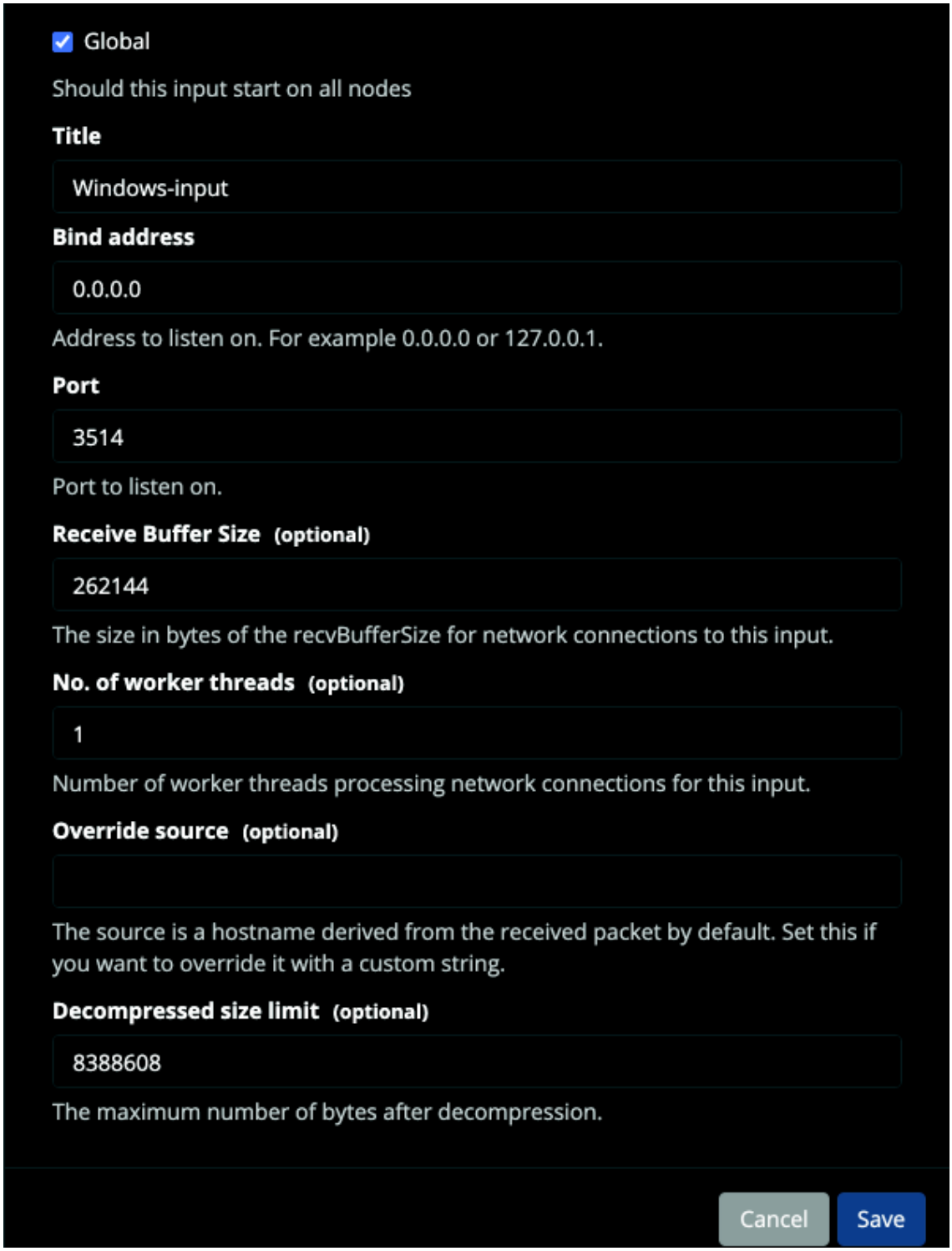

Figure 23: From Windows NXLog

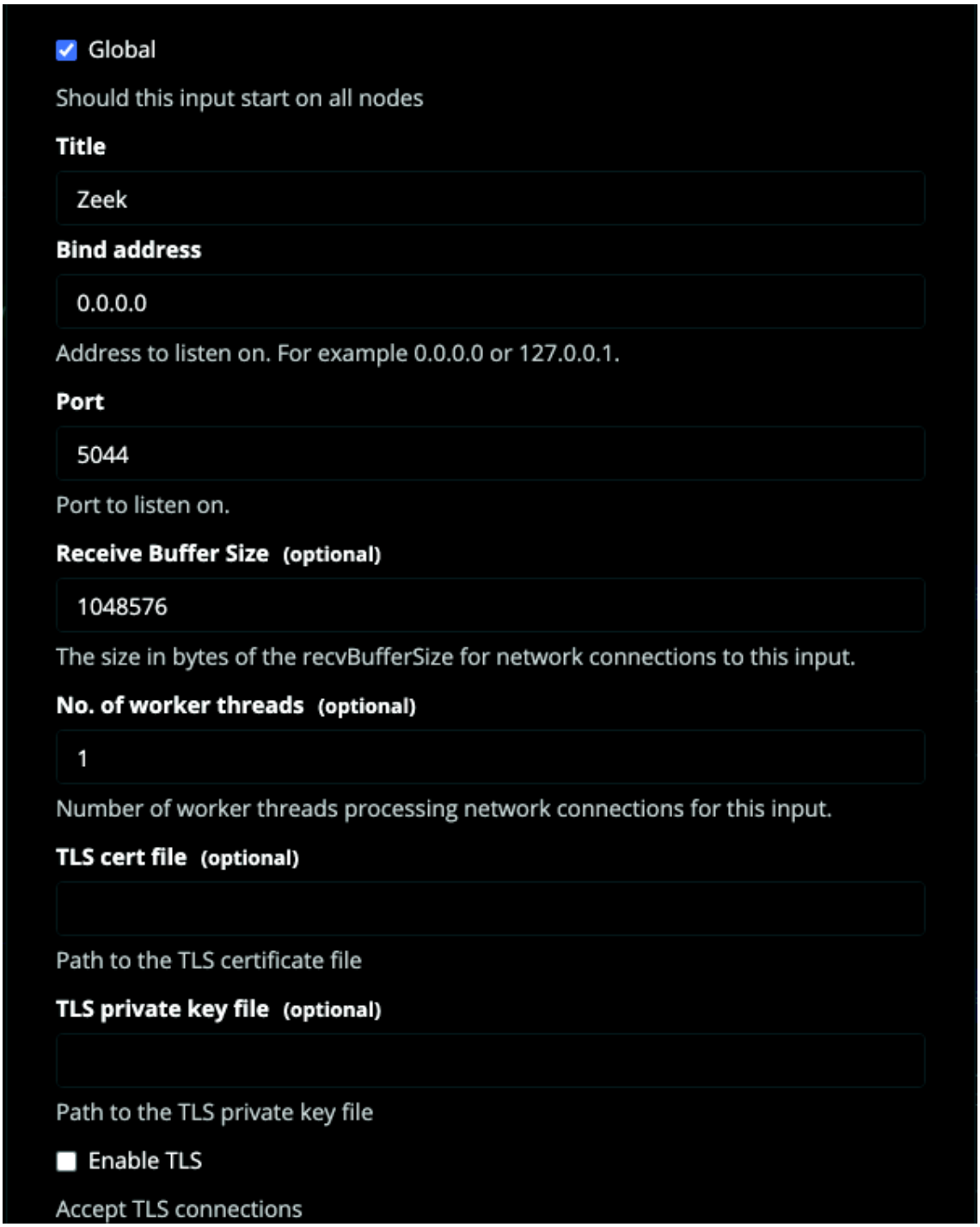

Figure 24: From Zeek IDS Filebeat

Now, after the inputs configured, we needed to validate by looking at the data streams that will be sent to Graylog. Therefore, we created the stream for those two sources of Zeek IDS and Windows machine.

a. Open https://192.168.10.111:9000 and log in as admin.
b. Then go to streams and Create Stream
c. Insert title and description and leave Index set as default for Zeek IDS and Windows separately.
d. Tick remove matches from “All messages” stream and click save

And we can see those two stream sources been created as shown below.

Figure 25: Event Streams Configuration for Windows and Zeek

Then, for each of these, we had to click Manage Rules and select an input to type Server Input and select an input for those input sources Windows and Zeek and click Load Message.

From here we can see it has been loaded as shown in the following page.

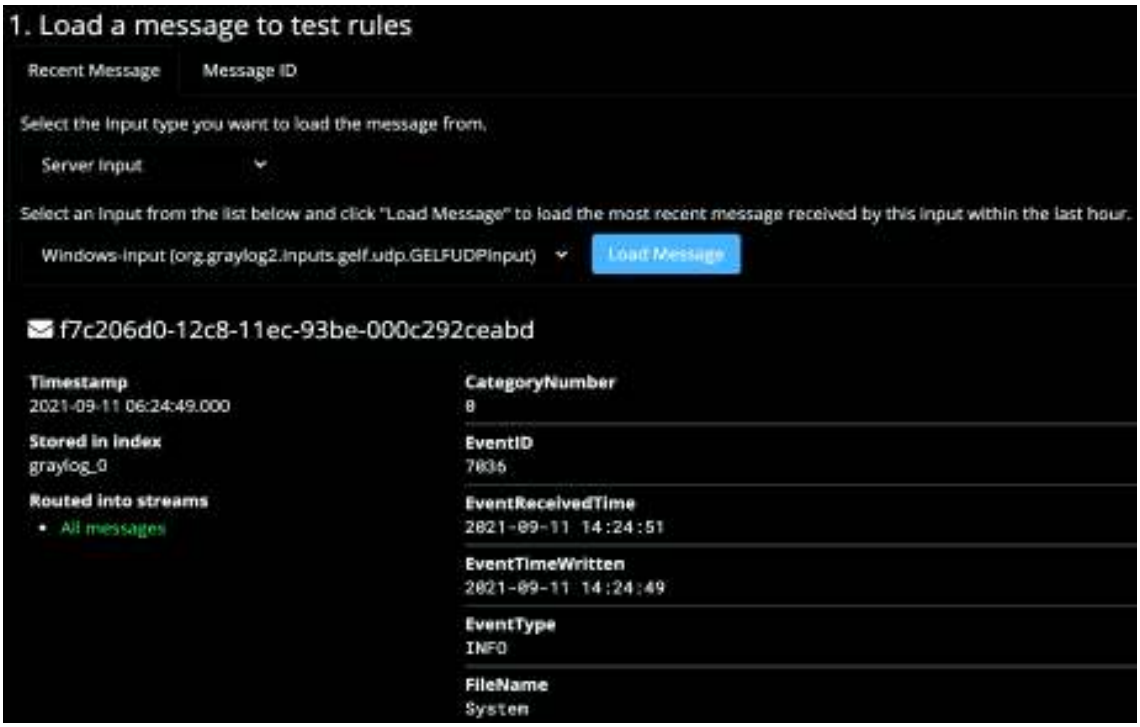

Figure 26: Event Test Rules for Windows Logs

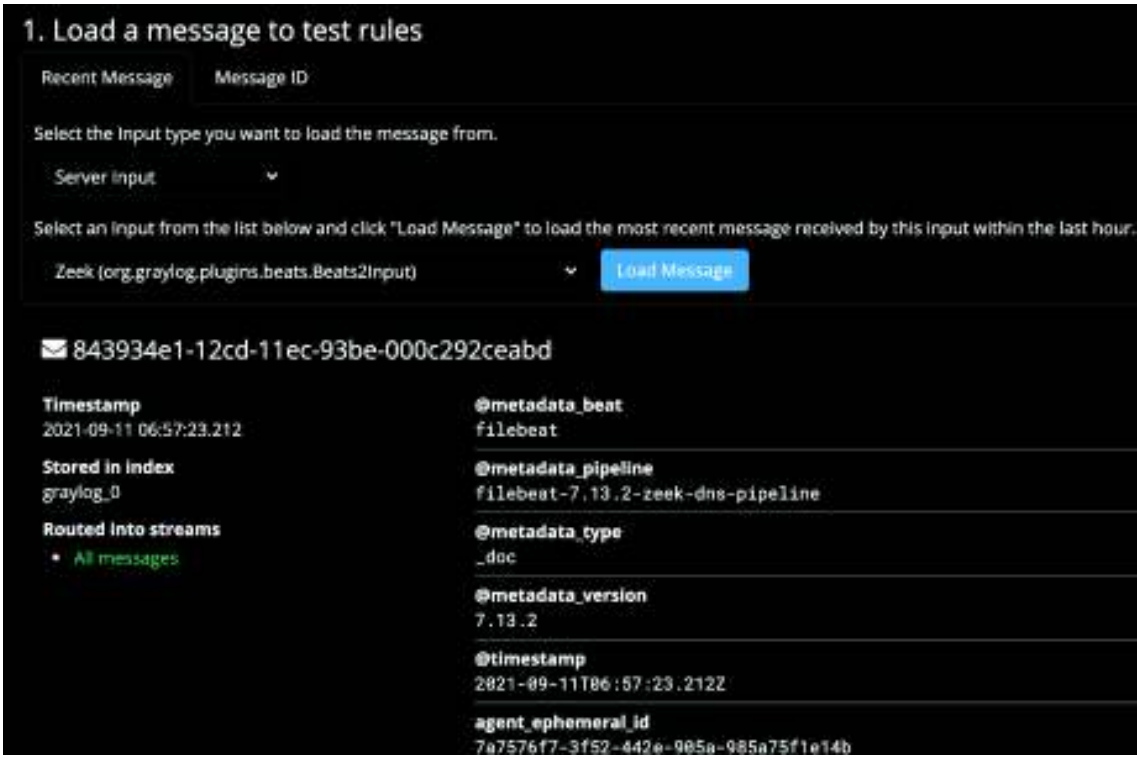

Figure 27: Event Test Rules for Zeek IDS

Then, we can see all events shown in Graylog. From here had to Search and aggregate these two sources as shown below:

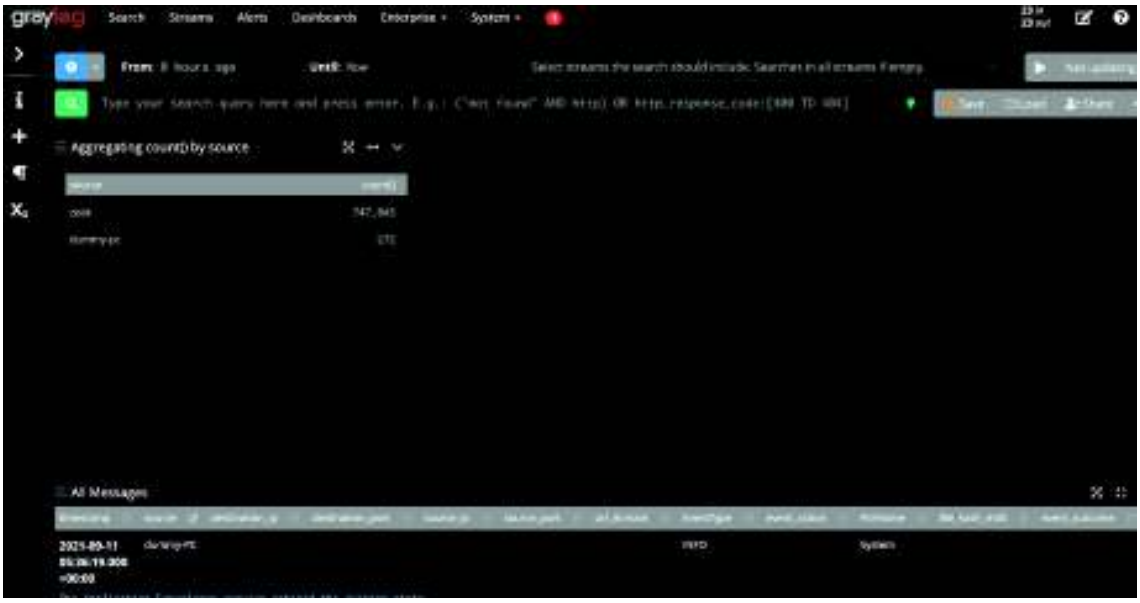

Figure 28: Aggregated Events Dummy-PC from Windows NXLog

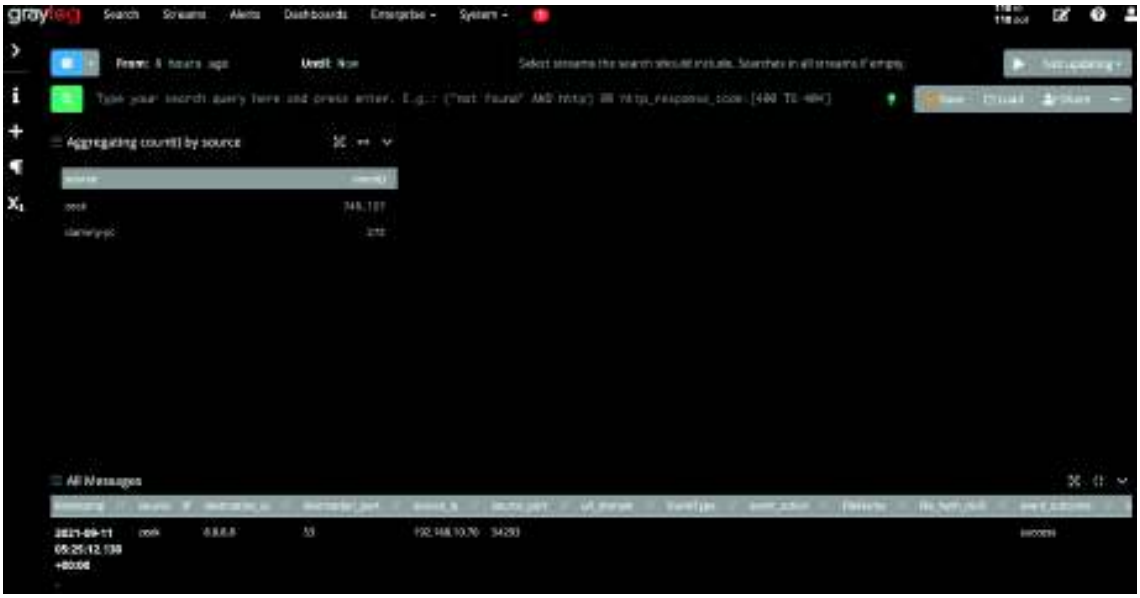

Figure 29: Aggregated Events Zeek IDS from Filebeat

4.3.1 Enable and Configure Threat Intelligence Feature in Graylog

In this part, we enabled and configured the Threat Intelligence features in Graylog. For this part, we will only be relying on Zeek IDS data source and not from Windows NXLog because Windows NXLog source of events is suitable for Threat Hunting.

We go the Web System and go to System / Configurations and enabled the following and click save:

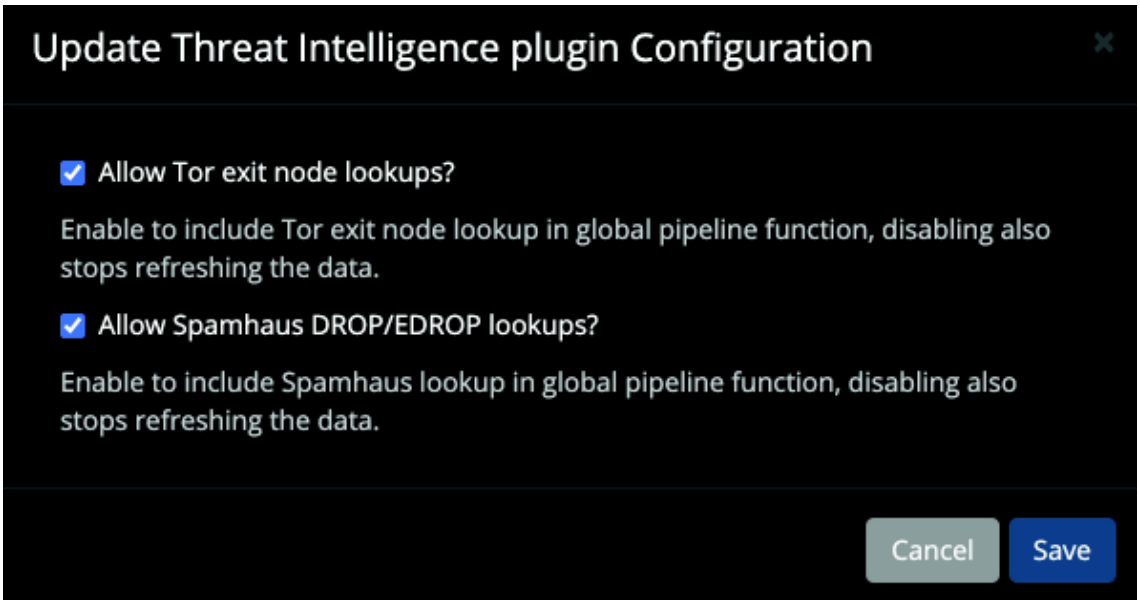

Figure 30: Threat Intelligence Features

Then, still under System / Configurations, we went to Content Packs and installed the following:

Figure 31: Threat Intelligence Content Packs

4.3.2 Create and Test Threat Intelligence and Threat Hunting Rules in Graylog

Now, the most important thing in this development is to ensure the rules for Threat Intelligence and Threat Hunting workable as intended.

First, we created the Threat Intelligence rules and test to see whether will it work as intended.

Still in the Web System, we go to System / Configurations settings and click on Update at Message Processors Configuration. From here, we moved the pipeline processor to the bottom and enabled all the processors listed.

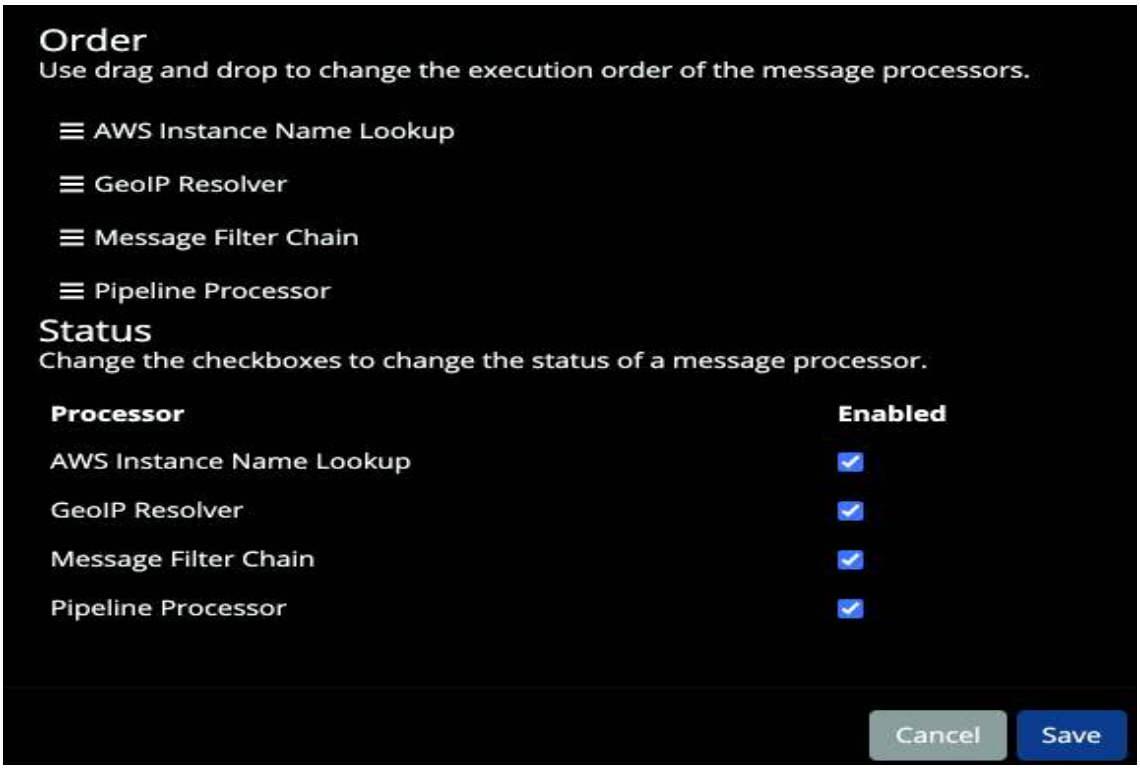

Figure 32: Message Processors Configuration for Threat Intelligence

This pipeline processor is the processor for retrieving, aggregating and processing the third-party Threat Intelligence feeds which in this case, we used the Tor Exit Node List – Threat Intel Plugin as shown at page 51.

Then, we create a Threat Intelligence pipeline and insert a title, its description and saved it. After that, we edited the pipeline connection and clicked on edit connections and select the Zeek stream for streams and saved.

Still under the same pipeline named Threat Intel: Tor Exit Nodes, we configured the rules by clicking on Manage rules at the top and Create rule. Once here, we inserted the description and the rule for Tor Lookup and saved it.

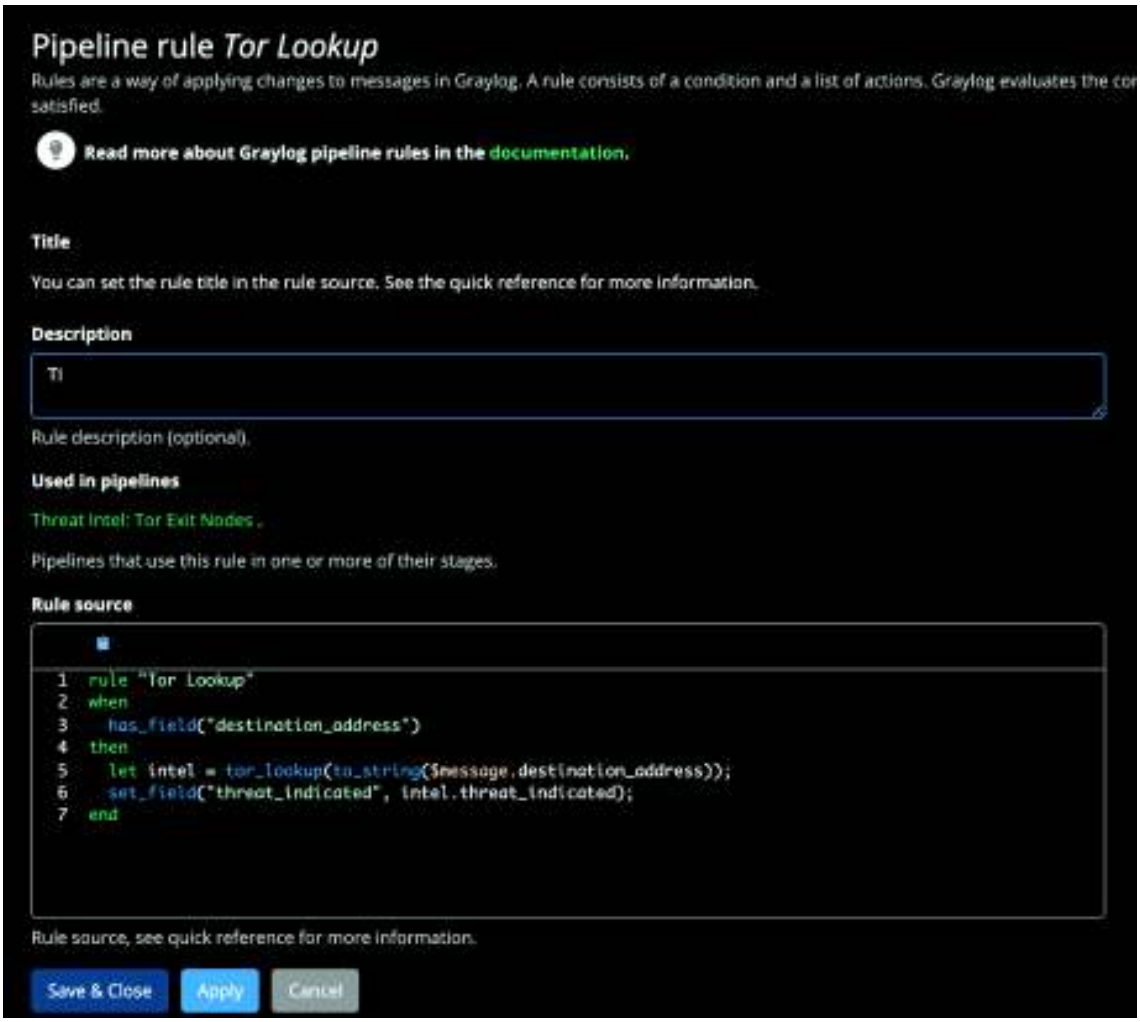

Figure 33: Threat Intelligence Rule Feature

Next, we had to test the rule and see whether it works or not. Open the Workstation PC running Windows 7 machine and open the browser and go to Tor Exit Node list. This Tor Exit node list can be retrieved at https://www.dan.me.uk/tornodes.

Then, we see the outcome of the threat_indicated it returned to true, which means it has been identified as true positive events mapped to Tor Exit Node Threat Intelligence list. If it shown as false, that means, it is not a Tor Exit Node.

Figure 34: True Positive Event

Figure 35: False Positive Event

Second, now we need to create a Threat Hunting rule for Windows machine that running with Windows 7 Professional.

For this Threat Hunting, we will need to create a Threat Hunting pipeline and insert a title and its description and saved it. After that, we need to edit the pipeline connection and click on edit connections and select the Windows stream for streams and saved.

Still under the same pipeline named Threat Hunting: Failed Login, we configured the rules by clicking on Manage rules at the top and Create rule. Once here, we need to insert the description and the rule for Threat Hunting Failed Login and saved it.

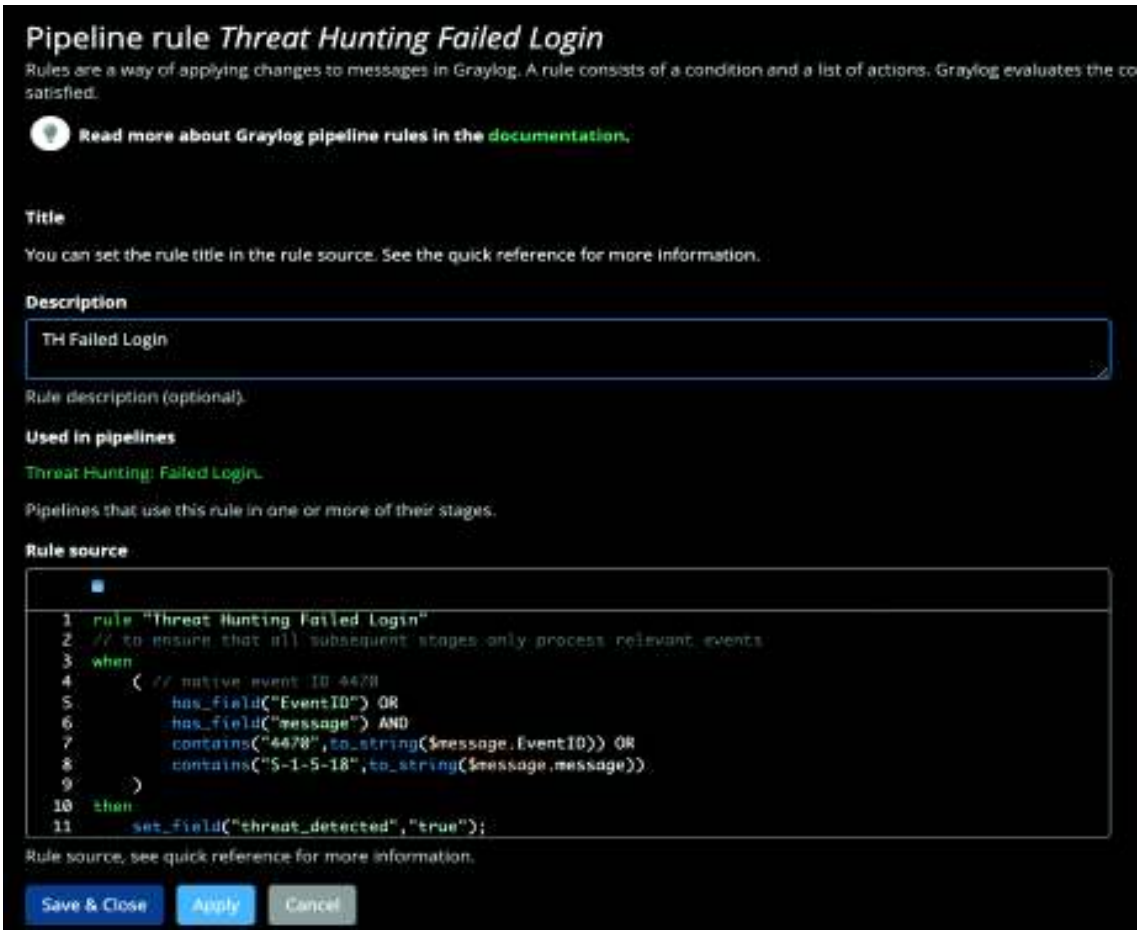

Figure 36: Threat Hunting Rule Feature

Now, in order to test the rule, we tried to attempt to the Windows 7 machine with the wrong password three times. Once exceeded than 3 times, the login will be locked. By right, the result in Graylog of threat_detected should be returned to true condition.

Figure 37: True Positive Event

4.3.3 Create and Test Slack Alert Channel for Notifications

In this part, we will be creating the alert and notifications for Threat Intelligence and Threat Hunting rules. The medium which was used is Slack channel.

In order to create this Slack, we will need to register ourselves at https://slack.com and we created it via Slack App Directory. After created, we need to copy the Webhook URL and put into Graylog Alert feature.

Figure 38: Slack App Directory Configuration

Figure 39: Graylog Alert & Events Feature

And here were the output results of Threat Intelligence from Slack channel that we received.

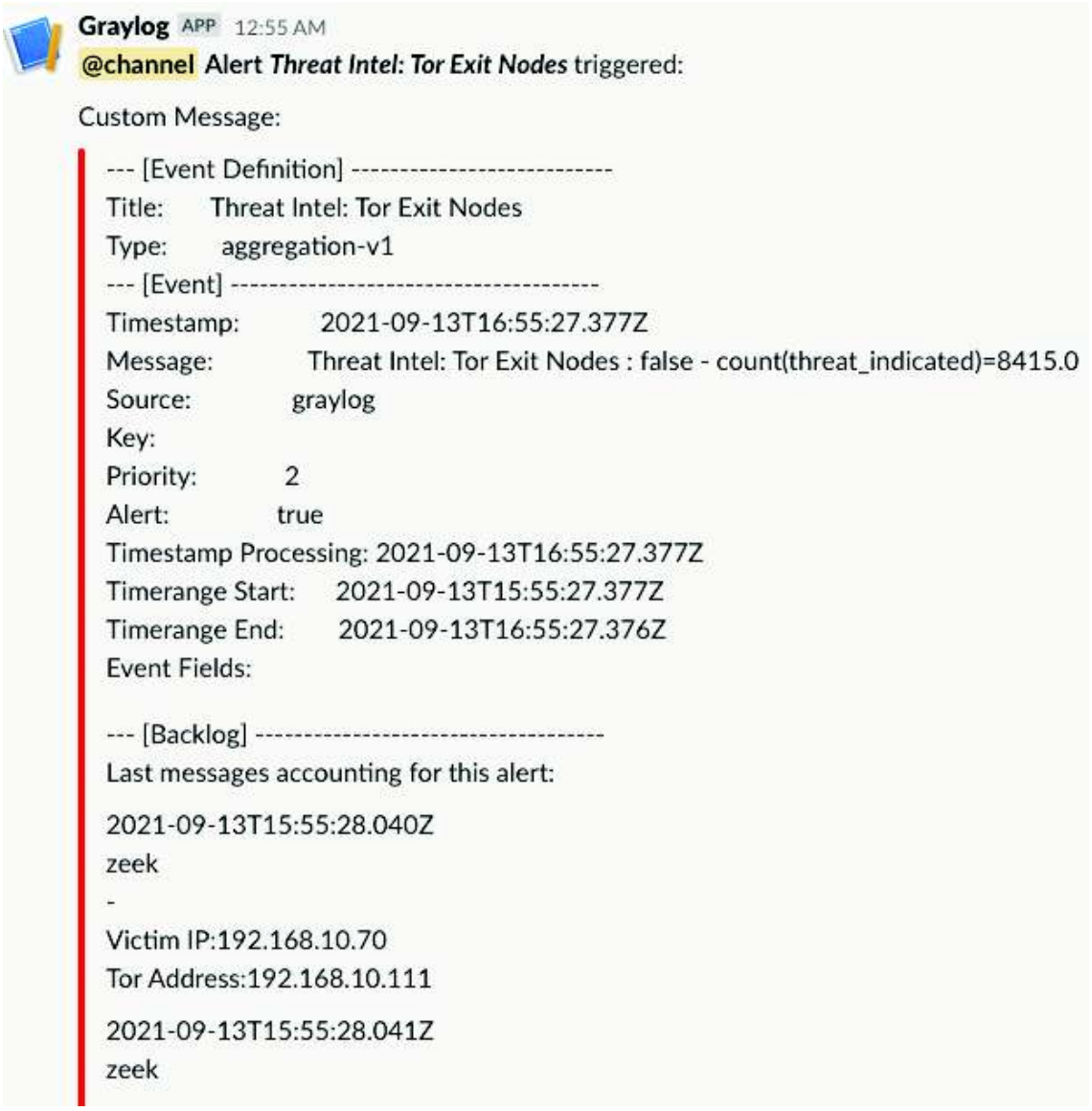

Figure 40: Threat Intelligence Alert

The following is the output result of Threat Hunting from Slack channel that we have received.

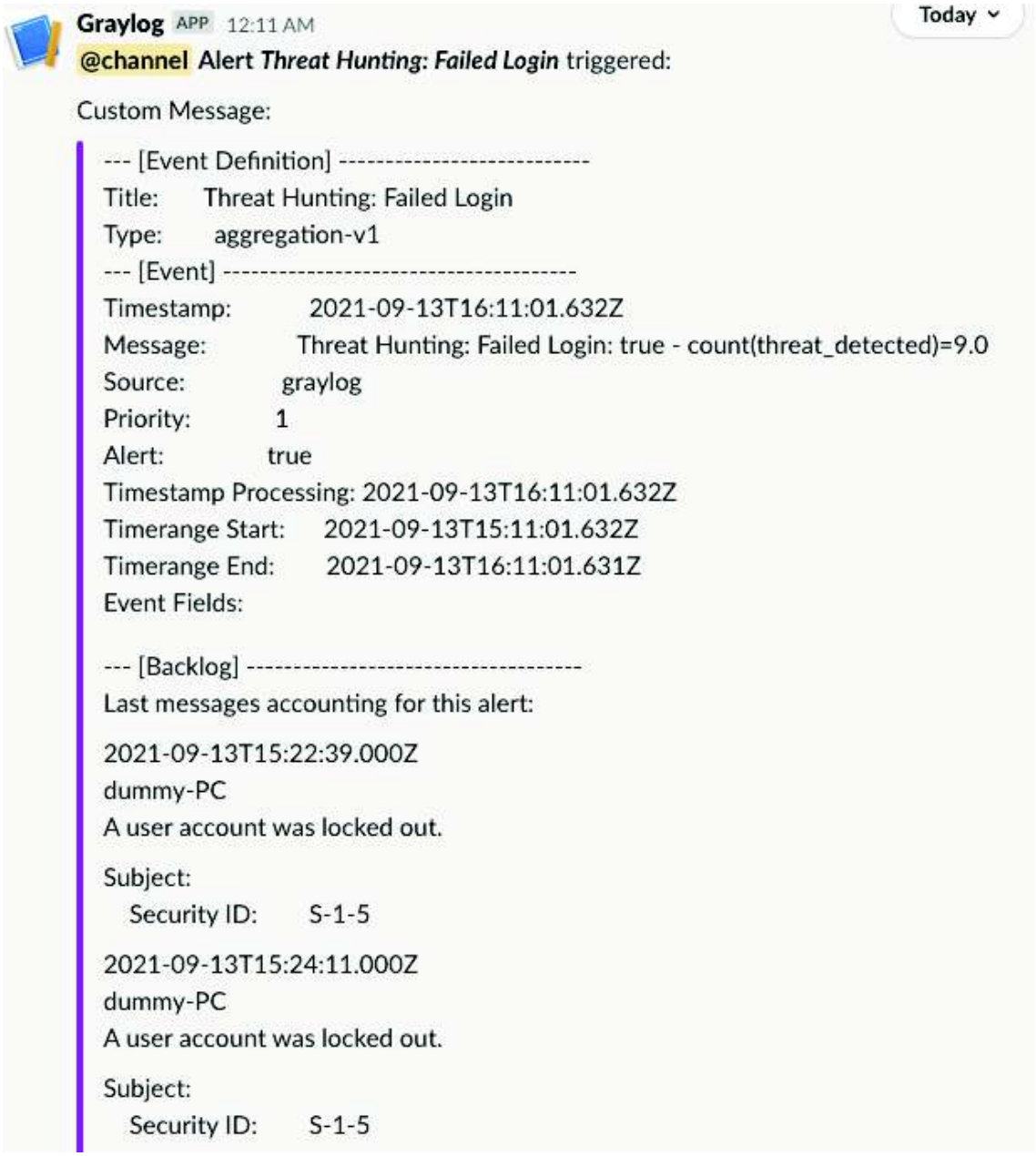

Figure 41: Threat Hunting Alert

CHAPTER 5

SUMMARY AND CONCLUSION

5.1 Summary of main findings

As we know, most of the cyber-attacks are sophisticated and it comes with many ways. It could be coming from phishing techniques, social engineering techniques, vulnerabilities and many others. As cyber-attacks evolving every single day, we as a cyber defender in an organization are required to equip ourselves with correct tools and technologies in place in order to combat cyber-attacks. Despite that, a very skilful staff is also required to ensure that any attacks found are analysed properly. Organization nowadays is having difficulties finding the right people in protecting their network perimeter from attacker. However, the tool is vital to facilitate the objective in defending cyber-attacks with ease. With this development system of Log Monitoring Tool equipped of Threat Intelligence and Threat Hunting, it could help a small medium enterprise or small organization with an idea to protect their organization from cyber-attacks with minimal cost. It is therefore, we decided to develop this platform to overcome their issues at very low cost. As we know, with the covid-19 as a new-norm to communities around the world, most of small business enterprise are impacted of this virus economically.

5.2 Discussion and Implications

Before we decided to use Graylog Log Monitoring Tool as a platform for this system development, we had reviewed the other competitors in the market. After taking into consideration of finding the best platform with a very minimal cost as well as finding platform that could give less complexity in configuring or installing it, we decided to choose Graylog Log Monitoring Tool.

We also compared the modules provided between the other competitors with Graylog. In terms of UI and UX satisfaction, to us Graylog is better. We also think, the close competitor to Graylog with similar solution is Elastic SIEM. The only part they cannot win is the feature of Threat Intelligence as a whole.

5.3 Limitations of the system

Even we know most of the platforms are offering similar features to Graylog, we seen that, the Graylog offer less complexity in developing it to compare to other competitors. However, throughout the system development took place, we faced challenges. One of the challenges are the rulesets we created for Threat Intelligence and Threat Hunting. When we created the rules, we are required to test it out and validate it. In the beginning, we found nothing is triggered which by right it should be sent to Slack channel for notification. After hundred times of reading the platform documentations, configuration and testing, we managed to get it done.

Also, as we know this platform which given free does have its limitation if we compare with the paid version. One of it, is correlation of the events and creating complex events and alerts. For the free version that we used, is only given aggregation and not correlation.

Even this system may not perfect to compared to other competitors which are quite expensive and easier to use, it suits the objective in achieving it at very minimal cost to develop.

5.4 Future Development

For future development, we think the Graylog can still consume more events from multiple devices as long as the 5 Gigabyte maximum storage stay intact per day. If this is done, we can create more use cases for Threat Intelligence rules and Threat Hunting rules. As we know, in this development, we only created two rules in total that is for Threat Intelligence and Threat Hunting.

Moreover, with the existing devices and rulesets in place, we could also create more rulesets based on continuous use-case development. Generally, building a SIEM is not one or two destinations. It is a marathon and never-ending destination. It requires a lot of efforts in maintaining, developing and configuring it. The cyber attackers never sleep in offending or compromising business assets and therefore, we as a cyber defender would need to do the same to ensure we are defending against them with modern technology, right process and right people.

REFERENCES

Lemos, R. (2018). *On the Hunt for Security: Tired of waiting for signs of an attack, companies large and small add threat hunting capabilities to their playbooks*.

Information Security, 20(6), 4–9.

Retrieved online: http://eds.a.ebscohost.com.newdc.oum.edu.my/eds.

David Alexander, Amanda Finch, David Sutton, & Andy Taylor. (2020). *Information Security Management Principles.*

BCS, The Chartered Institute for IT.

Retrieved online: http://eds.a.ebscohost.com.newdc.oum.edu.my/eds.

Liao, Xiaojing & Yuan, Kan & Wang, Xiaofeng & Li, Zhou & Xing, Luyi & Beyah, Raheem. (2016). *Acing the IOC Game: Toward Automatic Discovery and Analysis of Open-Source Cyber Threat Intelligence*.

755-766. 10.1145/2976749.2978315.

Retrieved online: http://eds.a.ebscohost.com.newdc.oum.edu.my/eds.

Kure, H., & Islam, S. (2019). *Cyber Threat Intelligence for Improving Cybersecurity and Risk Management in Critical Infrastructure*.

JUCS - Journal of Universal Computer Science.

Retrieved online: http://eds.a.ebscohost.com.newdc.oum.edu.my/eds.

Mitkovskiy, A., Ponomarev, A., & Proletarskiy, A. (2019). *SIEM-Platform for Research and Educational Tasks on Processing of Security Information Events.*

ELearning & Software for Education, 3, 48–56.

Retrieved online: http://eds.a.ebscohost.com.newdc.oum.edu.my/eds.