

National College of Ireland

BSc (Honours) in Computing

Cyber Security

2020/2021

Umar Rafique

x18142061

x18142061@student.ncirl.ie

Cloud-based Research Honeypots

Technical Report

Contents

1. Executive Summary.....	2
2. Introduction	3
Background	4
Aim:	5
Technology:.....	5
Structure	12
3. System.....	12
Requirements.....	12
Functional Requirements	12
Use Case Diagram:	13
Requirement 1 <Create Account/Login >	13
Description & Priority.....	13
Use Case	13
Requirement 2 <Deploy Sensor/Droplet>	15
Description & Priority.....	15
Use Case	15
Requirement 3 <Destroy Sensor/Droplet>	17
Description & Priority.....	17
Use Case	18
Requirement 4 <Monitor Activity>	19
Description & Priority.....	19
Use Case	19
Requirement 5 < View Details >.....	21
Description & Priority.....	21
Use Case	21
Data Requirements	23
User Requirements	23
Implementation	23
Graphical User Interface (GUI).....	38
4. MHN Analysis Using Splunk	44
Cowrie:	48
Dionaea:	54
Combined Analysis:.....	60

5. Testing.....	63
6. Conclusions	66
7. Further Development or Research	66
8. References	66
9. Appendices.....	68
Project Plan	68
Ethics Approval Application (only if required)	72
Reflective Journals	72
Survey.....	79
Splunk Overview	83
Digital Ocean Invoices	83

1. Executive Summary

"Cybercriminals are developing and boosting their attacks at an alarming pace, exploiting the fear and uncertainty caused by the unstable social and economic situation created by COVID-19" (Stock, 2020) INTERPOL Secretary-General.

In today's world, everything is connected to the Internet, which increases the exploitation of devices and systems.

As we live in the middle of the pandemic, Company employees work remotely, making them even more vulnerable. The intrusions also increase on the same scale as the Internet multiplies. In the past few months, cyber-attacks are at their peak.

The National College of Ireland and Technology University Dublin was attacked by malware. Then we have probably the most significant cyberattack on USA oil pipeline due to which oil prices surge and about two days ago the most significant cyber-attack in the history of the Republic of Ireland attacking HSE. The systems were forced to shut down for avoiding future damage.

Minister of State for Public Procurement and eGovernment Ossian Smyth said, "This is a very significant attack, possibly the most significant cyberattack on the Irish State". (Gráinne Ní Aodha, 2021)

Throughout the year, the project I have worked on is to research how these attacks took place in the first place, which are commonly attacked ports and gathering the information of the attacker, such as IP address and if it is a brute-force. The most common passwords used to get into the systems.

I decided to work on this project because the main reason was the study of attackers' behaviour. During the Christmas break each year, we have seen a new cyber-attack. This

gives me the idea that cyber attackers will be more active during the pandemic, and we can see the results.

The project's complexity is a bit high because I used third-party services. The reason to do this was to mitigate the risk on my machine. If I run server and pots on my machine, I knew I would be gathering the attacks. Simple attacks could disrupt while working on the project.

The project is built on Linux (Ubuntu) using the Digital Ocean cloud platform, running an MHN server and attaching the sensors/droplets (Honeypots) to the MHN.

2. Introduction

This report will describe the use of Honeypots, what honeypots are used for—a guide to set up your own Modern Honey Network (MHN). We can deploy the MHN server for a limited time. The Ubuntu has been upgraded & a few MHN features are not compatible anymore with the new versions of Ubuntu whilst the developers of MHN are constantly working to improve their project.

Companies of significant scale to small scale have implemented strong firewalls policies and antivirus software to prevent such intrusions. However, attackers usually break the firewall rules by using different ports or by opening the back door for devices. In this scenario, the implementation of honeypots could capture the attacker's activities and alert the security members to block those ports for future damage. The companies demand countermeasures for specific periods of times, e.g., which ports been attacked and what payloads have been used, to prevent such attacks in the future.

This work uses honeypot scenarios of Modern Honeypot Network to detect future attacks. I will deploy honeypot, gather logs in real-time in the above network scenarios and take decisions based on machine learning techniques. Currently, work carried out in this semester is divided into three parts (1) Selection of MHN (2) Deploying Honeypots (3) Debugging of MHN server.

- (1) Selection of MHN, In the early stages of the project. I had to research and decide to either run honeypots on an actual system or virtually. I decided to run honeypots virtually because it is more secure than deploying on the existing system rather than putting my system at risk.
- (2) Deploying Honeypots when it comes to deploying honeypots, there is a wide variety of honeypots to deploy for the project. I have decided to run Dionaea and Cowrie as the main honeypots for the final report.
- (3) Debugging of MHN sever, Whilst the deployment of droplets/sensors. I ran into few errors explained in detail under the Technology section. I had to debug those error to gather the logs from attackers.

The work status is that honeypot is successfully deployed in the digital ocean, and real-time logs are already collected to make real-time decisions on attacks. The subsequent immediate work is to apply machine learning techniques to classify attacks.

Using MHN enables the person who has deployed the MHN server to study the different attacks by attackers by trapping them in the honeypot and logging their attacks to the server. This will be explained in greater detail, and there are three types of interactions low, medium, and high. This report will focus on low interaction, medium interaction, and high interaction. The report will also outline my own Modern Honey Network analysis, the troubleshooting whilst developing the network, and the results of my analysis on different droplets/sensors attached to the server. The analysis will focus on the multiple results from attackers' locations, IP addresses, most used passwords to get in the system.

This report will also explain the requirements needed to deploy the project, the design, the architecture of the project, how I implemented and troubleshoot whilst implementing the server and the sensors/droplets for MHN, and the implementation of the Splunk tool for analysis. To illustrate how exactly this project operates.

Background

As my specialisation stream is cybersecurity and for my final year project. I wanted to something related to my stream when we are studying a module called Security Principles. That was the first time I heard about honeypots and how they work. It took my attention, and I started researching and working on honeypots.

Honeypots are decoy systems used in real networks to divert attacker from systems to another to study their behaviour, tools and techniques used by hackers. The level of interaction allowed by systems helps us categorise the honeypots into three categories. Usually, honeypots are designed to detect and report attacks against network and network systems such as Dos/DDOS DNS, DHCP etc.

Honeypots came with an idea and approach of catching the attackers understanding their tools to be ready for countermeasures to protect the existing organisation network. The honeypots are defined as a security resource. As we know, the number of cyber attackers is increasing rapidly and leaving the organisation vulnerable and sometimes bankrupt with the new GDPR. We must be extra cautious when handling users' data on Networks. The typical methodologies used to prevent attacks were IDS/IPS systems or penetration testing to find vulnerabilities within the network and systems. The downfall for Penetration testing could be unexpected results while a member is trying to find vulnerabilities which put the system in a very insecure mode to do countermeasures. As we know, the attackers are finding alternative ways to exploit the system, so we should also find an alternative way of learning their behaviour without risking our existing system. I will do so by deploying the MHN server on the cloud and attaching the sensors/droplets. This will attract the attackers to exploit. In the meantime, we would understand their behaviour what ports they are attacking the cloud platform is Digital Ocean for time interaction and analysis.

Mainly Honeypots are broken into two categories: production-based honeypots and research-based honeypots. For this project, I will deploy research-based honeypots whose primary purpose is to attract and lure attackers worldwide to gather the data for analysis. The sole purpose of researched-based honeypots is to gather the data from different locations worldwide whilst deploying the sensors worldwide, which helps us understand

where the attackers are most attracted to attacks and what countries the attackers are usually attracted to that part of the world. For the trial, I had deployed Honeypots Cowrie and Dionaea in the USA & India to understand where the most common IP addresses and the ratio of attacks hit the sensors.

Aim:

Many honeypots can be deployed, but the aim to deploy and analyse for my BSc (honours) project is mainly based on Cowrie, Dionaea. In the later stages of my master's degree, I would love to work on snort and Passive OS Fingerprinting (POf).

As mentioned earlier, since the servers update, the MHN project was fully compatible with earlier versions of Ubuntu. While deploying the cowrie, I ran into unexpected errors that were resolved using troubleshoot guide for MHN. The ticket is open for snort and POf. The developers of MHN are working on it to resolve the issue, including the script changes. Once the ticket is resolved, I would successfully deploy other Honeypots. In the meantime, the core focus is on Cowrie and Dionaea. As these two are mainly target networks

Dionaea Honeypot:

Dionaea honeypot is to trap malware exploiting vulnerabilities. The goal is a gathering the data copy of the malware. Since the pandemic people are working remotely for meeting VOIP, Dionaea is a low interaction honeypot that works around the server-side that supports such protocols SMB, HTTP, FTP, TFTP, MSSQL, MySQL, SIP for Voice Over IP (VOIP).

Cowrie Honeypot:

Cowrie honeypots are designed to work as a proxy to capture Secure Shell interaction (SSH) and Telnet connections to log brute force attacks. Since people are working remotely and these are the most used ports. The cowrie records the session information and often connected to the Internet to monitor the tools, scripts and host use by password guessing attacks such as Brute Force Attacks. Cowrie is a medium Interaction Honeypot.

I will be working on POf, a high-level interaction for fingerprints and Snort, an intrusion detection system, if the issue from the server-side of MHN gets resolved by tickets to the developers. Snort help us to monitor TCP/IP networks. Which usually are used by attackers for Denial of Service and Distributed Denial of Service.

Technology:

What Technology will you use to achieve what you have set out to do and how will you use it?

For the technologies to achieve the goal, I researched to start the deployment process for honeypots. There were two options either I do physical or virtual. The physical deployment

of honeypots on the laptop was that I would have required an extra laptop as the attackers could deploy the payload into my actual laptop or hack my logs. I realise the second option, which is running virtually, is more secure and cost-effective.

Following are the technology I am using for the project to achieve its primary goal:

1. Digital Ocean
2. Modern Honey Network
3. Splunk

The first step to achieving this project's goal was to find a compatible cloud network. Where I could deploy my MHN server and sensors/droplets. After few suggestions, I came across Digital Ocean Network, allowing students to use 60 days free trial of 100 US dollars.

After creating an account with Digital Ocean, the next step was to deploy the Modern Honey Network and sensors to gather the data from attackers. I ran into few unexpected errors whilst deploying the MHN server. The screenshots are following, which would give an overview of errors and how I manage to troubleshoot those errors:

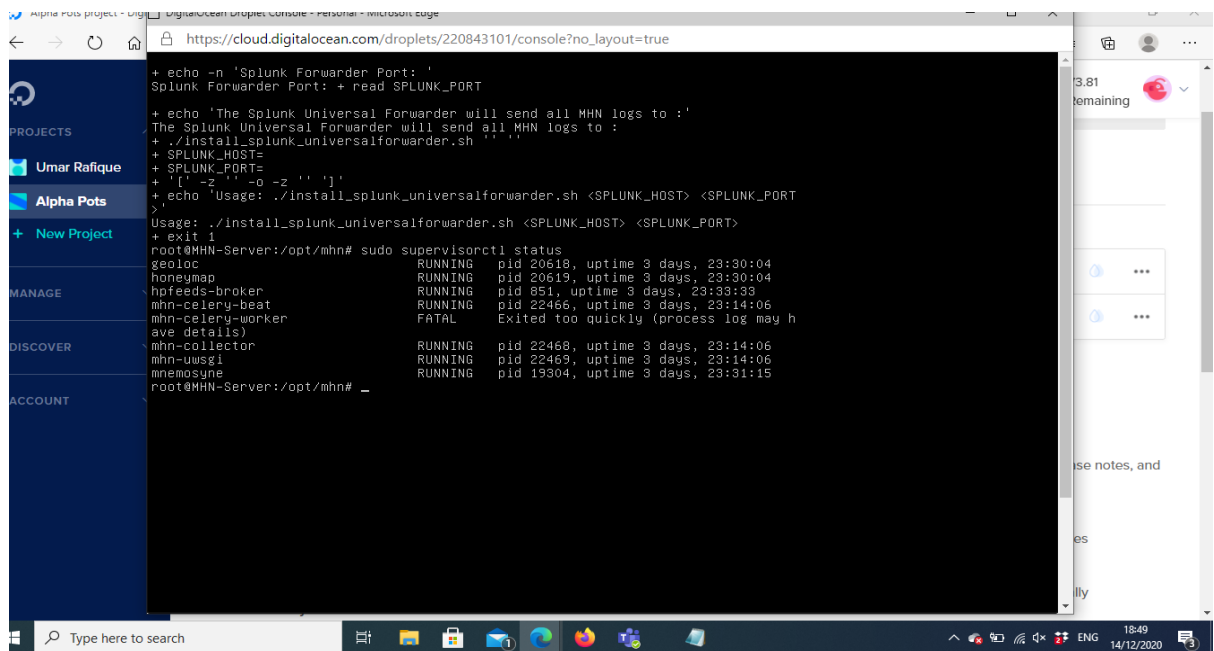


Figure: 1

Figure 1 shows us that the mhn-celery worker was FATAL. It is vital to log the attacks as I will not be able to work on the project without resolving it. The issue directly impacted the MHN server dashboard. After the first attack, the rest of the logs were not stored and logged into the server.

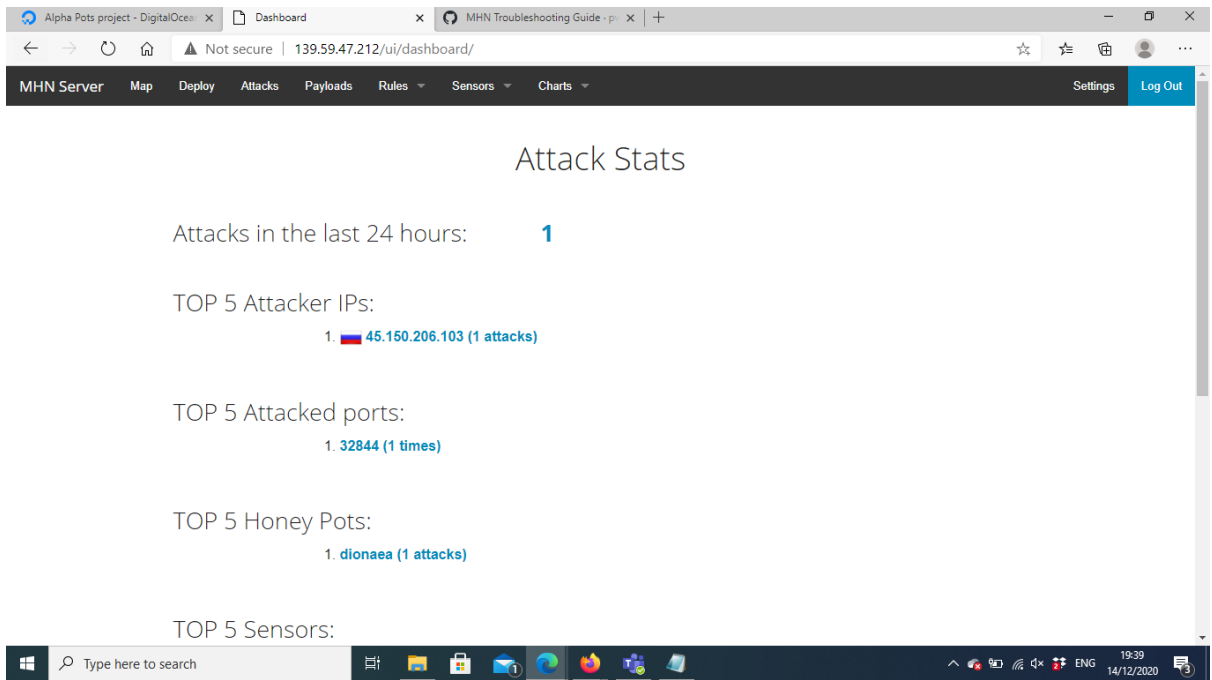


Figure: 2

Figure 2. shows us exactly what I was getting on my dashboard.

After researching the issue and how to troubleshoot the issue. The following screenshots will illustrate the Linux commands I ran to resolve the error.

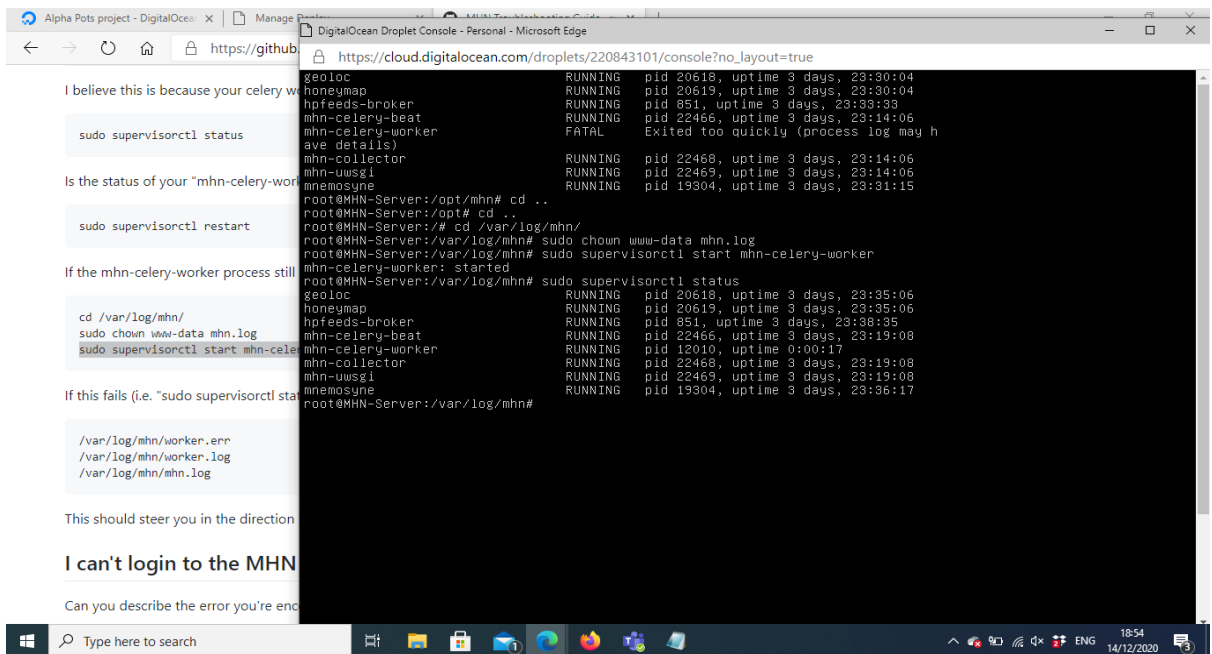


Figure 3.

After researching the effects of error, I also researched how to resolve the causing issues. The resolving of the error was based on the commands on the link for the troubleshooting guide for MHN.

I also ran into an error from the digital ocean platform where their data was lost for virtual machines due to that issue from the platform provider. I could not do much, so I had to create a new MHN server and droplets/sensors to restart gathering the data. Figure 4 screenshot will provide you with an overview of the issue.

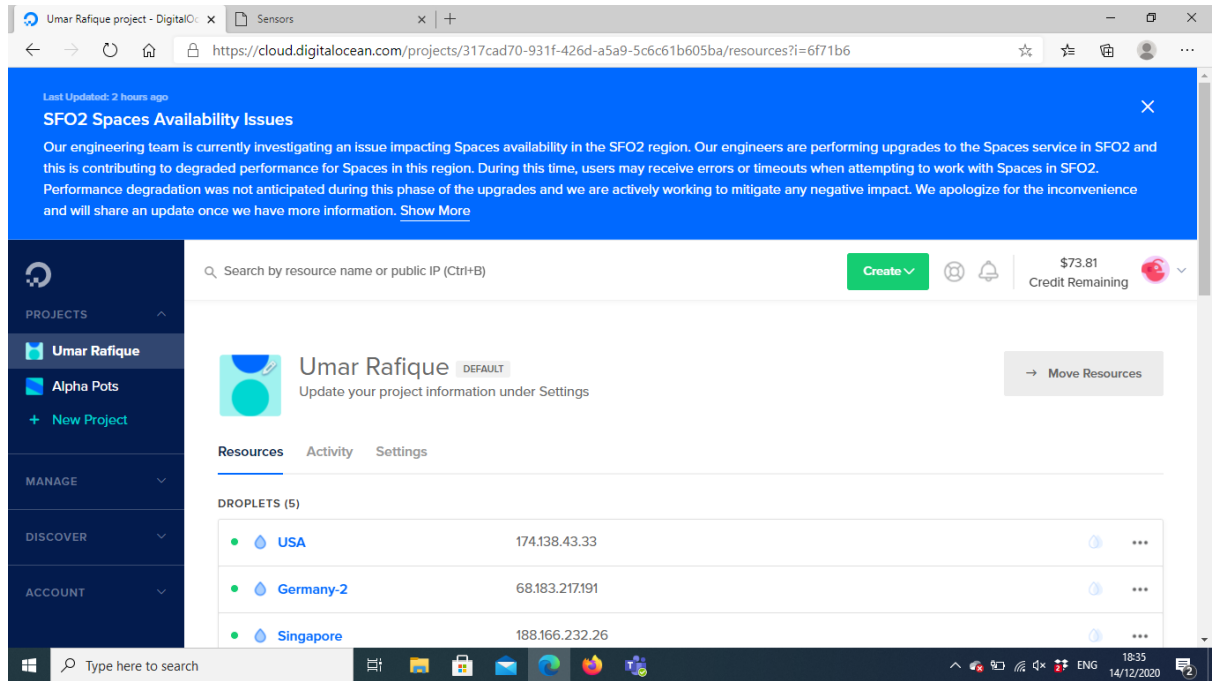


Figure: 4

As you can see, the virtual machines for the USA, Germany-2, Singapore were running, but due to the error in the SF02 region where my MHN server was running, I lost all the gathered data from all the virtual machines. After opening a ticket with Digital Ocean, I realised as the server was not on backup because I had to pay extra for backup. So, I could have lost all the data. Rather than waiting for them to resolve the issue, I created a new server and droplets/sensors.

Figure 5 and 6 screenshots will explain the nature of error on the Digital Ocean.

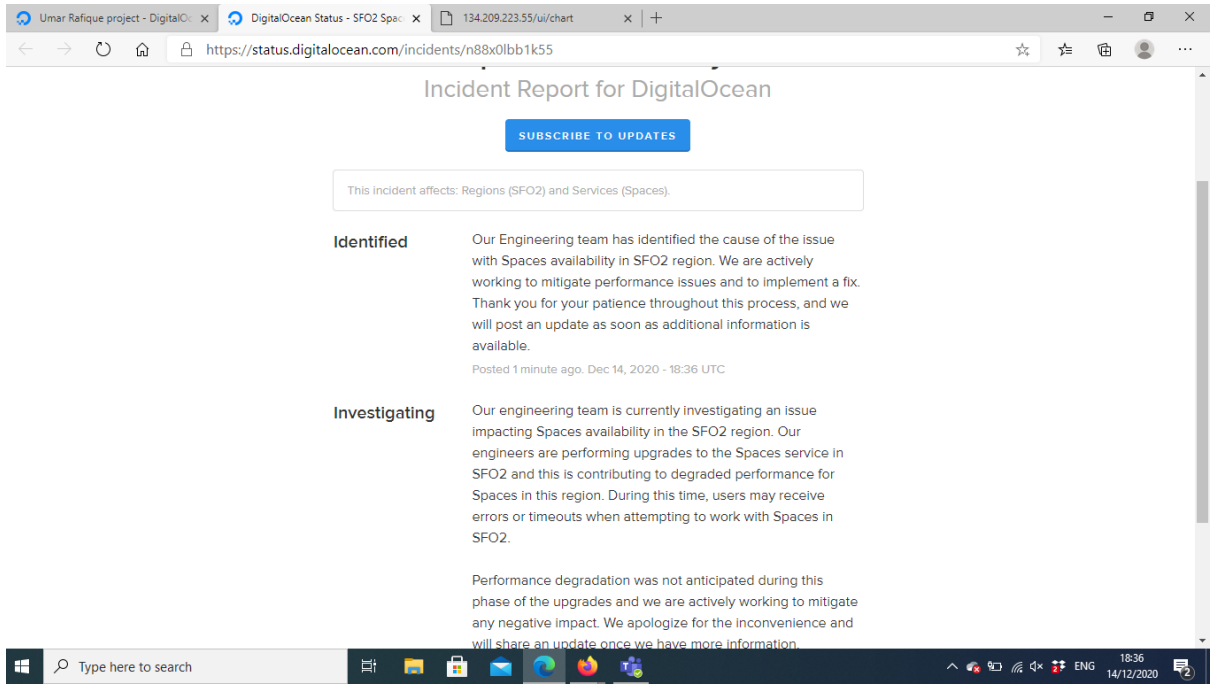


Figure 5

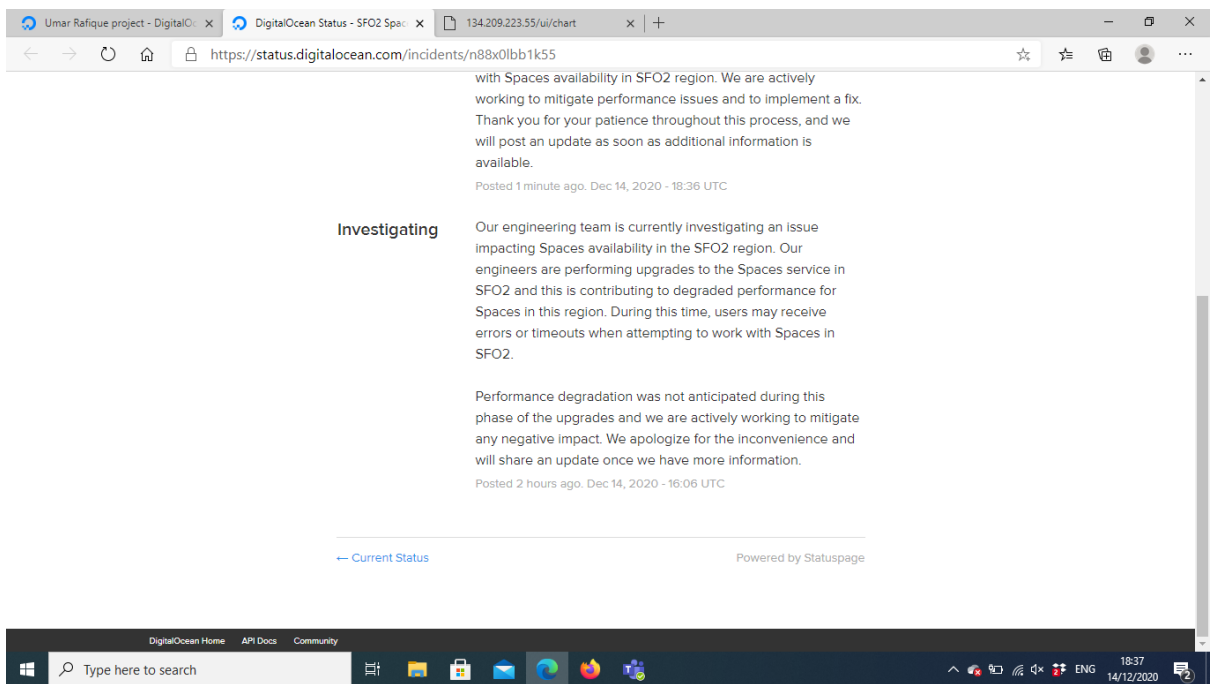


Figure 6

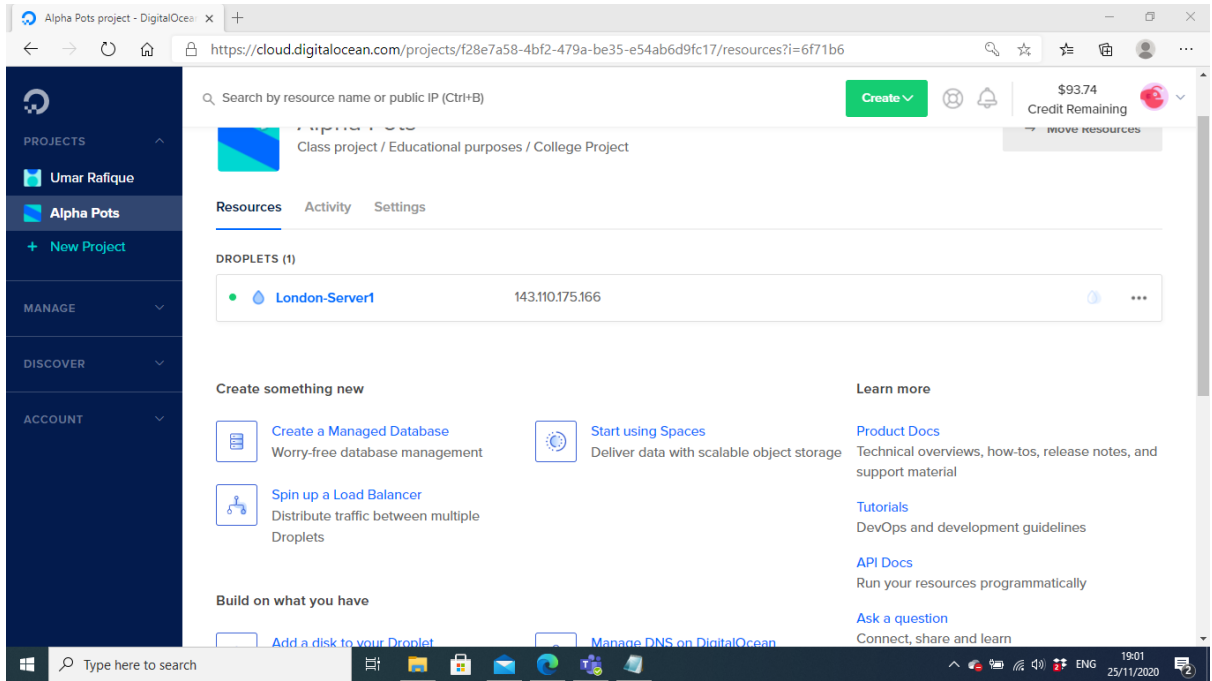


Figure 7

Figure 7 shows the new New MHN server, which is now running on 143.110.175.166 instead of 139.59.47.212.

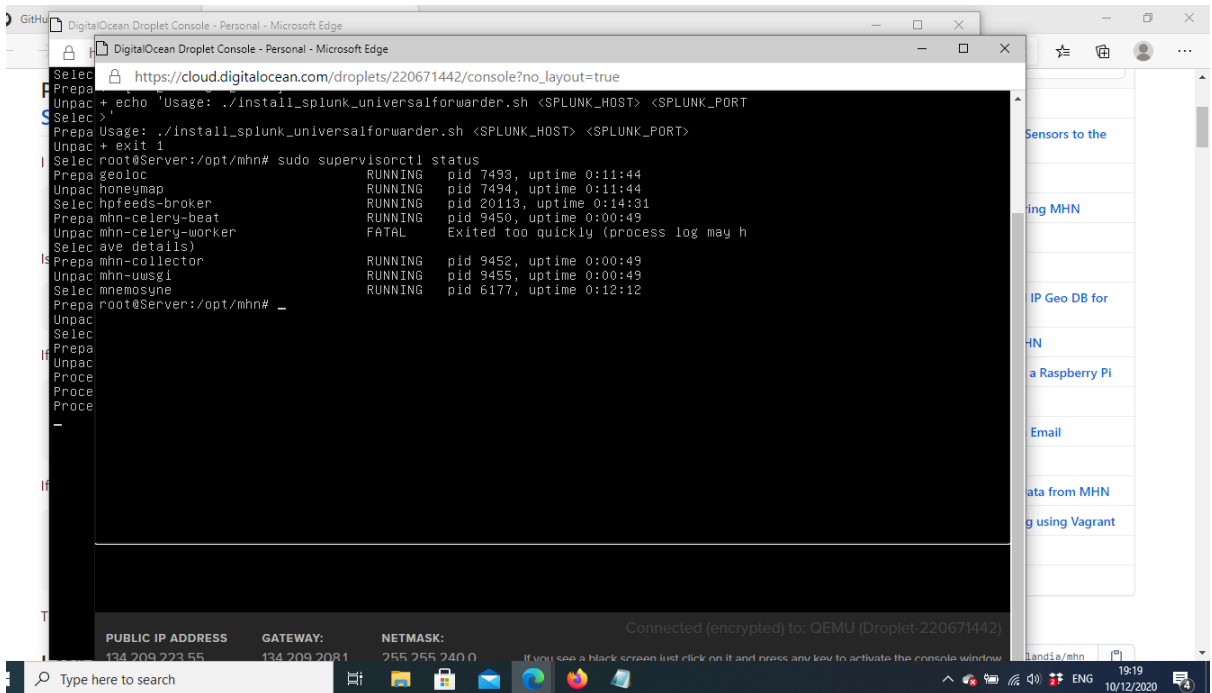


Figure 8

Figure 8 screenshots show that when I was running into errors and the day, I resolved the error from the MHN server and ran into an error from Digital Ocean.

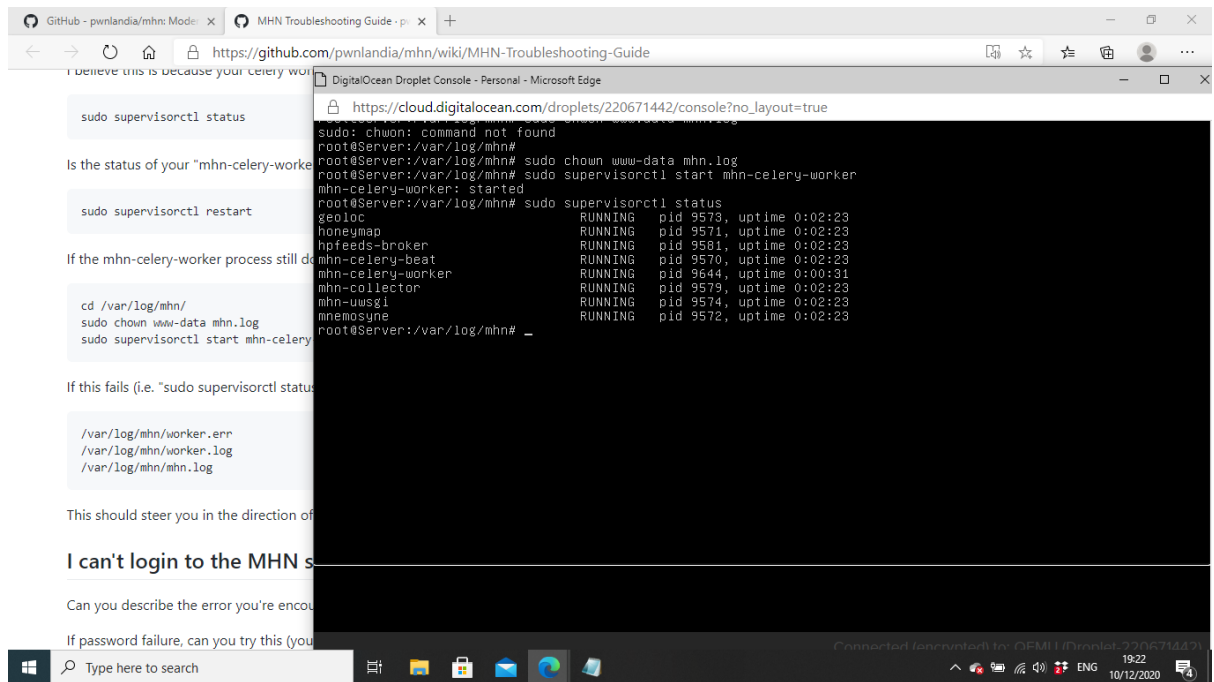


Figure 9

Figure 9 shows that I successfully resolved the issues I was encountering. The next step was to integrate Splunk into the MHN server. So, I would be able to get the logs when I integrate Splunk to analyse data. Figure 10 screenshot shows where I successfully integrated Splunk with the MHN server.

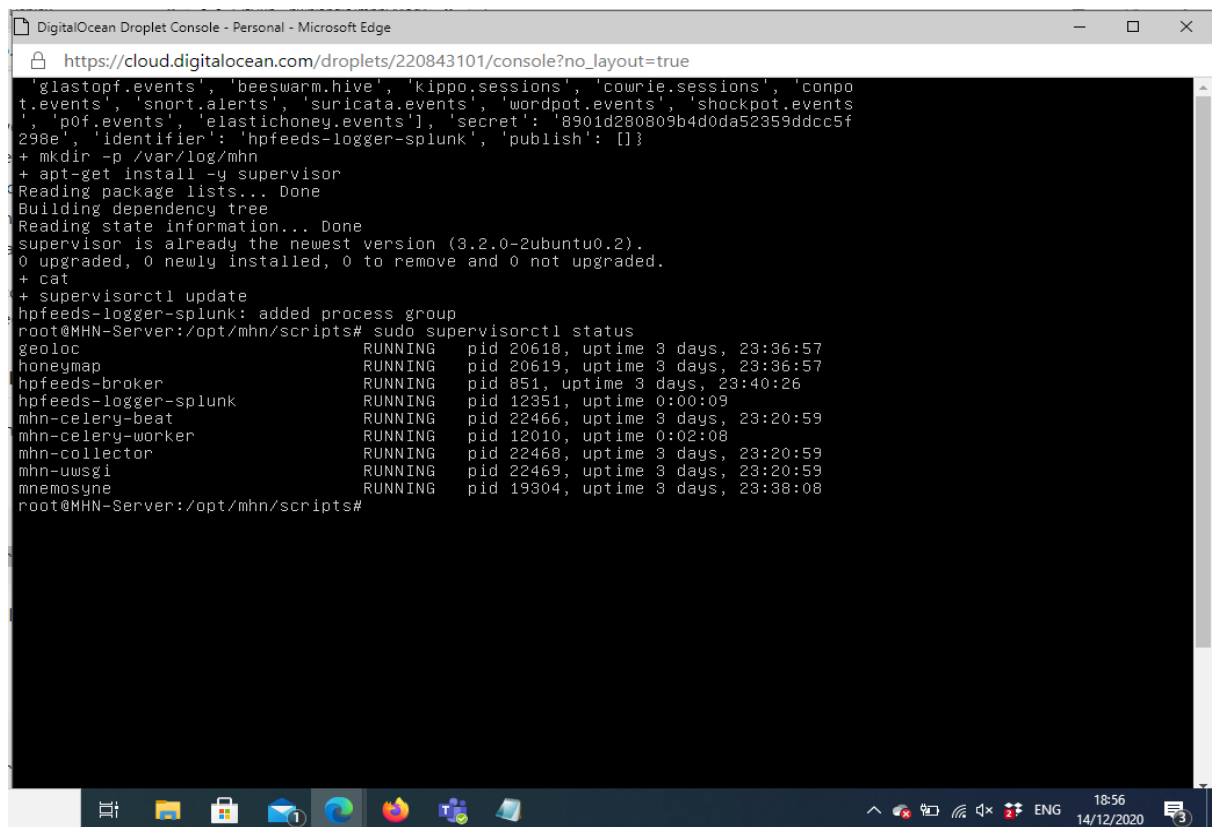


Figure 10

So far, I have successfully achieved the first main steps for the project, and now I will start adding sensors to gather and learn the behaviour of attackers. The final step will be creating an account with Splunk using their free trial for a month to analyse the gathered data.

Structure

Section 1 was an executive summary, and section 2 is based on the project introduction. I have categorised the introduction into different heading, including the brief introduction, background, aims, technology, and structure.

The honeypots are explained in detail, why they are vital, and the types of honeypots. I have also explained the project's aim, what honeypots sensors/droplets I am currently running, and will be running in the future if the developers of MHN resolve the issue. I have explained the technologies I am using to host the MHN server in the technology section. I ran into it while deploying the server with screenshots and resolving those issues with screenshots. In the technologies section, I have explained the next step of the project and how I will achieve the project results.

Section 3 of the system will grow through the system requirement, functional requirements (Use Case), Data requirement, User requirement, Implementation and Graphic User Interface (GUI) of the Project.

Section 4 is based on an analysis of attacks gathered using MHN and Splunk's machine learning tool. Section 4 demonstrate the testing while working on the project. Section 5 is the conclusion. Section 6 is briefly written about further development or research. Section 7 is about references. The last section is about Appendices. I have attached the submitted Project Plan, Reflective Journals, a survey I conducted, Splunk Overview by exporting the pdf file from Splunk and Digital Ocean invoices.

3. System

Requirements

The Requirement section contains all the functional requirements of the project. It explains the description of the project and features how a user is going to interact with the system features.

Functional Requirements

This section lists the functional requirements in **ranked order**. Functional requirements describe the possible effects of a software system, in other words, *what* the system must accomplish. Other kinds of requirements (such as interface requirements, performance requirements, or reliability requirements) describe *how* the system accomplishes its functional requirements.

A short, imperative sentence stating highest ranked functional requirement.

Use Case Diagram:



Requirement 1 <Create Account/Login >

Description & Priority

A description of the requirement and its priority. Describes how essential this requirement is to the overall system.

This is essential as only authorised users can log in and access all the features of the system/project. All the user have to create an account before they log in to the system. The user's information will be saved on a cloud-based server database which could be SQL or NoSql, depending on the platform used by the providers.

Use Case

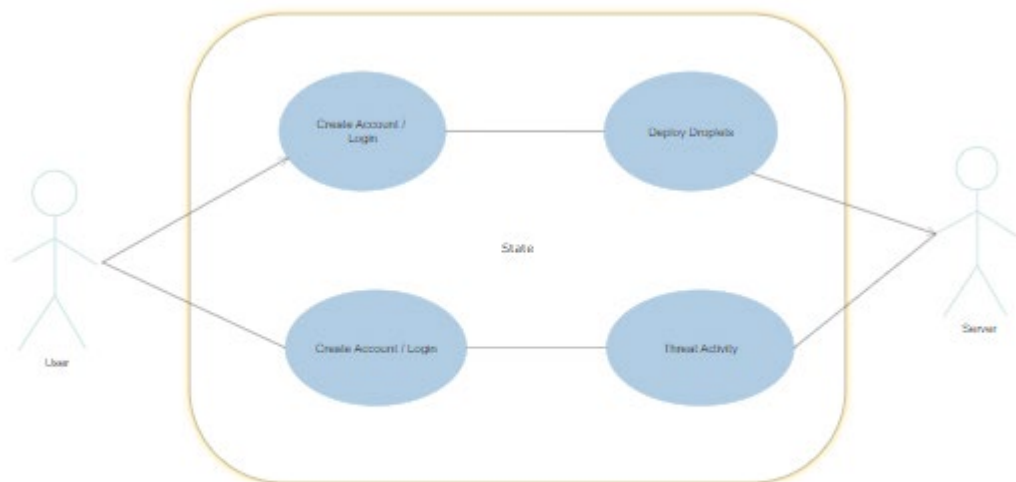
UR 1

Scope

The scope of this use case is to explain how exactly the users access the system and how the system interacts with the users.

Description

Create Account/Login



Use Case Diagram

The user is brought to the login screen whilst the system is in initialisation mode.

Activation

This use case starts when a <Actor> presses the register button /sign up and then brought up to the register page of the system.

Main flow

1. The user presses the register button. The system triggers the sign-up.
2. The user fills out the necessary details on the sign-up page.
3. The user clicks the register button.
4. The system sends an authentication email to the user account to validate the user.
5. The user clicked on the email link and brought it up to the login page.
6. The user entered the credentials, and the system brought them to the homepage of the cloud-based platform (Digital Ocean).

Alternate flow

A1 : <Incorrect Password >

1. The system compares the passwords entered.
2. The system notifies the user that the password is incorrect and did not match.
3. The use case continues at position 2 of the main flow.

A2 : <Email Check >

1. The system checks email entered is not already registered.
2. If the typed-in email is already registered, the user must choose another email or use the recovery option for their account.
3. The use case continues at position 2 of the main flow.

A3 : <Checks the data is being filled as required >

1. The system checks that all the mandatory fields are filled.
2. The system informs the user if they forgot to fill any mandatory field.
3. The use case continues at position 2 of the main flow.

A4 : <Already registered >

1. The use case continues at position 2 of the main flow.

Termination

The system presents the next to the login page. Therefore the sign-up page will be terminated.

Post condition

Successful:

- The user has created an account.
- The user can log in.
- The user credentials and personal data is stored in the server database.

Failure:

- The user cannot create an account.
- The user cannot log in to the account.
- The user credentials and personal data is not stored in the server database.

[Requirement 2 <Deploy Sensor/Droplet>](#)

[Description & Priority](#)

When the user has successfully created an account and logged in to the system, the user is then brought to the homepage. On the home page, the user can do several things, and one of the vital things to is deploying a Sensor / Droplet -Honeypot.

[Use Case](#)

UD 1

Scope

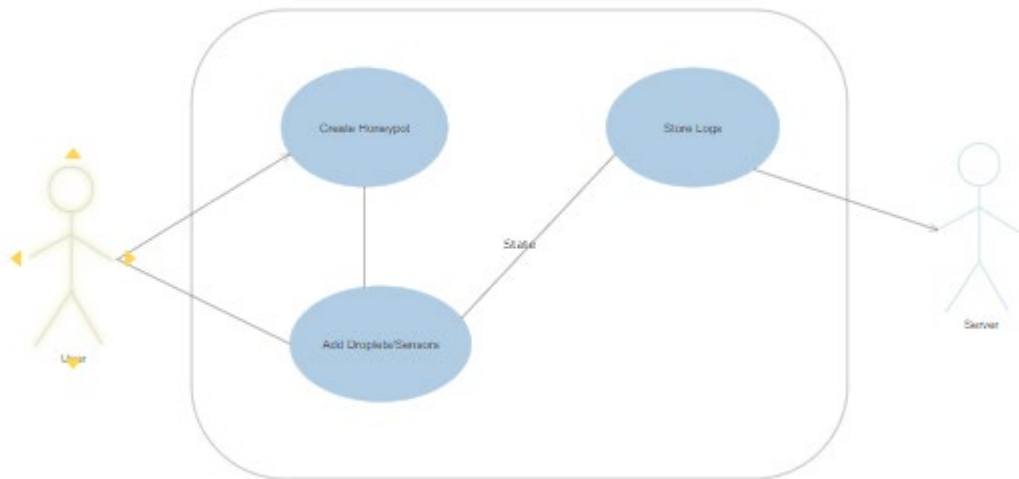
The scope of this use case is to explain how the user interacts with the system and how the droplets/sensors (honeypots) interacts with the MHN server.

Description

This use case describes the deployment of the droplets/sensors or honeypot. It also shows us how the honeypot communicates with the MHN server.

Use Case Diagram

Deploy Honeypots



Flow Description

Precondition

The user can do several things once the user logs in to the system. One of the vital things is creating droplets/ sensors (Honeypot).

Activation

This use case starts when a <Actor> presses the 'create new droplet' on Digital Ocean.

Main flow

- 1.0 The user clicks on 'create a new droplet'.
- 2.0 The system brings the user to the 'create new droplet page'.
- 3.0 The user selects the Operating System & version of the operating system wants to deploy.
- 4.0 The user then selects the size and location where in the world, wants to deploy the virtual machine.
- 5.0 The user type in the password for the virtual machine.
- 6.0 The user gives a password to the virtual machine to access it later.
- 7.0 The user clicks the deploy button.
- 8.0 The system brings the user back to the main page, where the user can access the droplet/sensors (Honeypot).
- 9.0 The user accesses the sensor by clicking on the droplet and selecting the console.

- 10.0 The user accesses the droplet, and login credentials typed earlier whilst creating a droplet.
- 11.0 The system responds with several commands.
- 12.0 The user adds sensors to the droplet (HoneyPot) to detect attacks using the Linux script.
- 13.0 The MHN server starts storing the attack logs.

Alternate flow

A1 : <Text not filled >

- 1.0 The system checks that all the mandatory fields have been filled correctly.
- 2.0 The system notifies the user of mandatory fields.
- 3.0 The use case continues at position 3.0.

A2 : <Incorrect Password>

- 1.0 The system checks the entered password with the saved password.
- 2.0 The system notifies the user invalid password or username.
- 3.0 The use case continues at position 10.0.

A3 : <Invalid Sensor Script>

- 1.0 The system run the Linux commands by the user.
- 2.0 The system notifies the user that the script copied by the MHN server no longer support due to the upgrade.
- 3.0 The use case continues at position 11.0.

Termination

The system presents added droplet on the MHN server dashboard. The sensor starts picking up attacks once it deployed and shown on the MHN server dashboard under the sensors.

Post condition

Successful:

- The droplet/sensor (HoneyPot) deployed successfully.
- The logs are stored in the MHN server.

Failure:

- The droplet/sensor (HoneyPot) fails to deploy successfully.
- The logs are not stored in the MHN server.

Requirement 3 <Destroy Sensor/Droplet>

Description & Priority

This use case describes how a user destroy the exciting sensor/droplets (HoneyPot). This is done for several reasons: the storage is filled, or the droplet is not attracting any attackers as it is supposed to do.

Use Case

UD 2

Scope

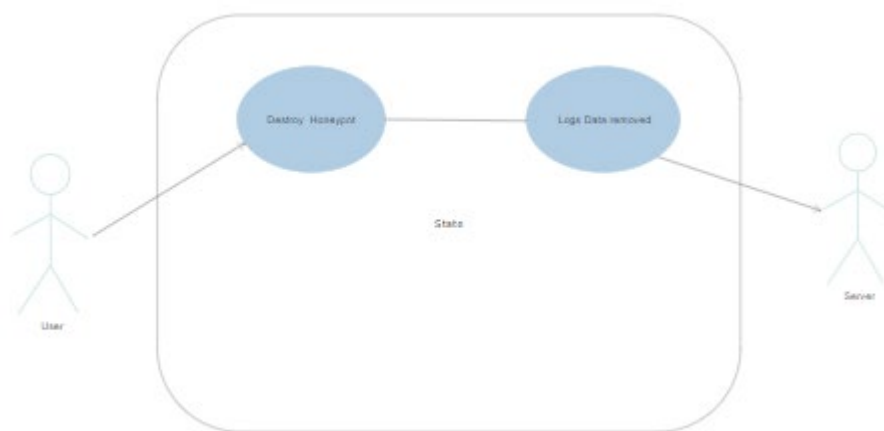
The scope of this use case is to explain how the user can destroy an exciting honeypot.

Description

This use case describes how a user destroy their droplets/sensors (Honeypot)

Use Case Diagram

Destroy Honeypot



Flow Description

Precondition

The user can do several things once the user logs in to the system. One of the vital things is to destroy the honeypot from the main screen.

Activation

This use case starts when a <Actor> presses the 'destroy droplet' on the Digital Ocean.

Main flow

1. The user clicks on 'destroy droplet'.
2. The system prompts the user to another page for confirmation.
3. The user clicks on the 'destroy' button.
4. The system destroys the droplet/sensor (Honeypot.)

Alternate flow

A1 : <The user changes their mind >

1. The system prompts the user to destroy the droplet page.
2. The user changes their mind and exits the page without clicking destroy button.
3. The use case continues at position 1.

Termination

The system destroys the droplet/sensor (Honeypot) and brings the user back to the main page.

Post condition

Successful:

- The droplet/sensor (Honeypot) destroyed successfully.
- The droplet /sensor (Honeypot) is removed from the homepage.

Failure:

- The droplet/sensor (Honeypot) fails to destroy successfully.
- The droplet /sensor (Honeypot) is still running on the homepage.

Requirement 4 <Monitor Activity>

Description & Priority

This use case describes how a user monitors the threat activity on created honeypots and gets the idea of which honeypots are attacks mostly and where the attacks are coming.

Use Case

MA 1

Scope

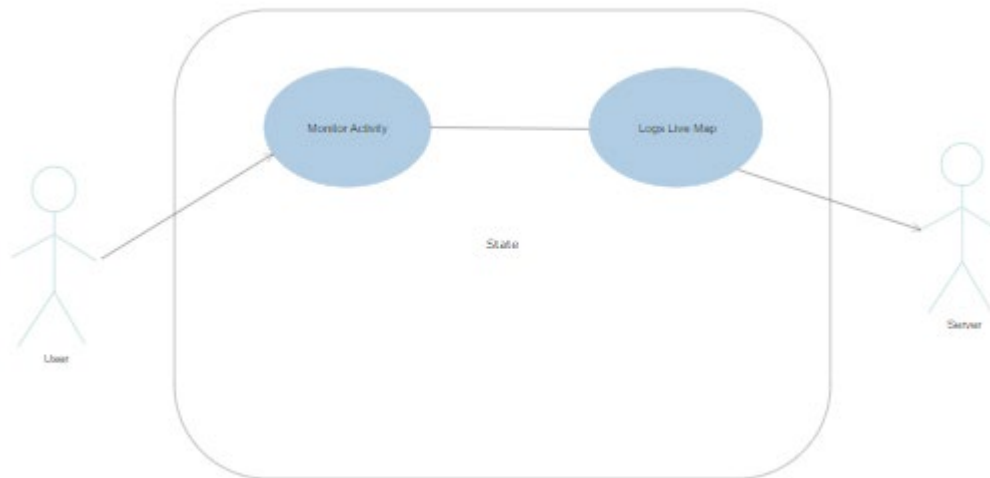
The scope of this use case is to explain how the user can monitor the threat activity in real-time.

Description

This use case describes how a user views the threat activity on the MHN server.

Use Case Diagram

Monitor Activity



Flow Description

Precondition

The user must have successfully deployed droplets/sensors (Honeypots) on the MHN server. The user is logged into the MHN server, and the homepage user clicks the Map button. Then the system shows all the attacks in real-time coming from worldwide.

Activation

This use case starts when a <Actor> presses the 'Map' button on the MHN server.

Main flow

- 1.0 The user clicks on the 'Map' button.
- 2.0 The system prompts the user to the Map page, where the user can monitor the attacks.
- 3.0 The user views the attacks coming from different locations and targeting the droplets.
- 4.0 The user monitors the attacks and finishes the task.

Alternate flow

A1 : < No Attacks >

- 1.0 The system displays the Map page.
- 2.0 The user sees there are no current attacks.
- 3.0 The use case continues at position 4 of the main flow.

Termination

The system successfully displays the “Map” page to monitor the attacks.

Post condition

Successful:

- The map page is displayed.
- Attacks are displayed.
- User monitors attacks.

Failure:

- The map page is displayed.
- No attacks are displayed.
- User cannot monitor attacks.

[Requirement 5 < View Details >](#)

[Description & Priority](#)

This use case describes how a user views the activity log. The user would view the log to make the attacks readable. Therefore, do research and analysis on the logs.

[Use Case](#)

VD 1

Scope

The scope of this use case is to explain how the user views the attack logs.

Description

This use case describes how a user views the logged logs.

Use Case Diagram Monitor Activity



Flow Description

Precondition

The user must have successfully deployed droplets/sensors (Honeypots) on the MHN server. The user has deployed droplets and has already stored the logs. The user has integrated the Splunk with the MHN server then the user clicks the “View log files” button on the Splunk application. The system shows all log files with every attack since the deployment of droplets/sensors (Honeypots).

Activation

This use case starts when a <Actor> presses the ‘View log files’ button.

Main flow

- 1.0 The system identifies that the “View log files” button is clicked.
- 2.0 The system prompts the user to a page where the user can view the stored log files.
- 3.0 The user extracts the log files using Splunk.

4.0 The user downloads the files for research & analysis to make the final report.

Termination

The system successfully extracts and downloads all the stored log.

Post condition

Successful:

- Log files are extracted.
- Logfile are downloaded.
- Research and analysis are conducted for the final report.

Failure:

- Log files are not extracted.
- A log file is not downloaded.
- Research and analysis are not conducted for the final report.

[Data Requirements](#)

The Data requirements are essential in this project, as we are storing the logs, which is raw data. I decided to go for an 80 GB SSD disk for the server as all the logs will be stored in the server, and for sensor/droplets (Honeypots), it is 25GB per sensor.

[User Requirements](#)


As we know, the server and droplets/sensors are hosted online and run virtually on a cloud-based platform. A fast, reliable internet connection is required to access, monitor and analysis the data.

[Implementation](#)

This section will implement the technologies used for this project and some issues that I had to debug to run this project successfully.

The first step to implement the MHN server is by creating an account with the cloud provider and use their facilities to deploy the MHN server. I have mentioned that for the project, I used Digital Ocean. The cloud provider came with the facilities of security and recovery. Once I created the account. I was promoted to the main dashboard to create the project. In the project, I can add the MHN server.

Create new project



Name your project

Enter name

New Project ✓

Add a description
Helpful for teams or differentiating between projects with similar names.

Enter description

College

Tell us what it's for
This will help us to provide a more relevant experience.

Class project / Educational purposes * ▼

Create Project

Figure:3.1

Figure 3.1 shows how to create the project.

PROJECTS

- x18142061
- New Project
- College Project
- + New Project

MANAGE

DISCOVER

ACCOUNT

Q Search by resource name or public IP (Ctrl+B)

Create ▼
🔒
🔔
USAGE \$12.67
🔴

College Project

Class project / Educational purposes / Server & Pots

→ Move Resources

Resources Activity Settings

DROPLETS (5)

● USA	167.99.61.5	🔗 ⋮
● Germany	167.71.42.67	🔗 ⋮
● Canada	138.197.159.36	🔗 ⋮
● India	167.71.228.187	🔗 ⋮
● Server-Lon	46.101.21.157	🔗 ⋮

Figure 3.2

Figure 3.2 shows us the project created in the main dashboard. As you can see, I am using a project name “College Project”, in which I am currently running my MHN server and few honeypots.

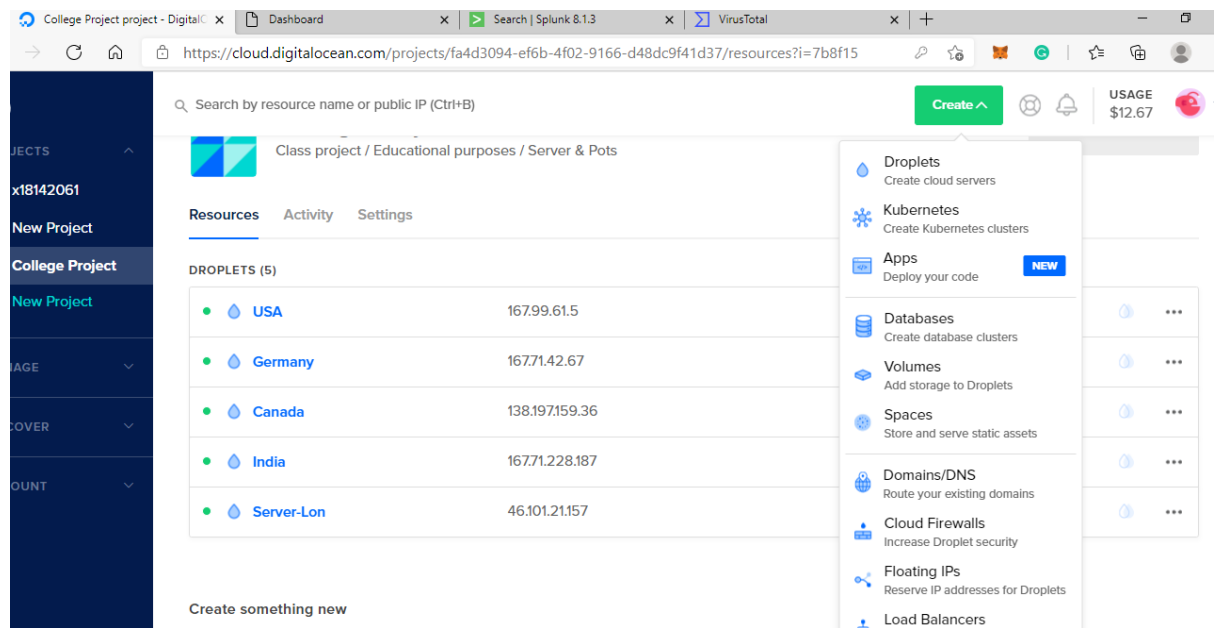







Figure: 3.3

The next step is creating a sensor, creating a virtual machine. In figure 3.3, it shows us the button of creating. Once we clicked the button, it prompts us to create the virtual machine. We need to click the first option, which is droplets. Once I clicked the create droplets. The page takes us to another page shown in figure 3.4 and 3.5.

Create Droplets

Choose an image [?](#)

[Distributions](#) [Container distributions](#) [Marketplace](#) [Custom images](#)

 Ubuntu 20.04 (LTS) x64	 FreeBSD Select version	 Fedora Select version	 Debian Select version	 CentOS Select version
--	--	---	--	---

Choose a plan [Help me choose](#)

SHARED CPU	DEDICATED CPU			
Basic	General Purpose	CPU-Optimized	Memory-Optimized	Storage-Optimized NEW

Basic virtual machines with a mix of memory and compute resources. Best for small projects that can handle variable levels of CPU performance, like blogs, web apps and dev/test environments.

CPU options: Regular Intel with SSD Premium Intel with NVMe SSD **NEW** Premium AMD with NVMe SSD **NEW**


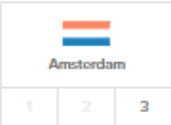
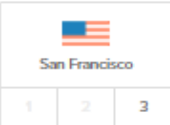





\$6/mo \$0.009/hour	\$12/mo \$0.018/hour	\$18/mo \$0.027/hour	\$24/mo \$0.036/hour	\$48/mo \$0.071/hour	\$96/mo \$0.143/hour
1 GB / 1 AMD CPU 25 GB NVMe SSDs 1000 GB transfer	2 GB / 1 AMD CPU 50 GB NVMe SSDs 2 TB transfer	2 GB / 2 AMD CPUs 60 GB NVMe SSDs 3 TB transfer	4 GB / 2 AMD CPUs 80 GB NVMe SSDs 4 TB transfer	8 GB / 4 AMD CPUs 160 GB NVMe SSDs 5 TB transfer	16 GB / 8 AMD CPUs 320 GB NVMe SSDs 6 TB transfer

i Our Basic Droplet plans, formerly called Standard Droplet plans, range from 1 GB of RAM to 16 GB of RAM. [General Purpose Droplets](#) have more overall resources and are best for production environment, and [Memory-Optimized Droplets](#) have more RAM and disk options for RAM intensive applications.

Each Droplet plan includes free outbound data transfer which is shared between all Droplets each billing cycle. Inbound bandwidth to Droplets is always free. [Learn more](#) or [try our price calculator](#).

Figure 3.4

Choose a datacenter region

 New York 1 2 3	 Amsterdam 1 2 3	 San Francisco 1 2 3	 Singapore 1	 London 1	 Frankfurt 1
 Toronto 1	 Bangalore 1				

VPC Network

default-nyc1 DEFAULT

All resources created in this datacenter will be members of the same VPC network. They can communicate securely over their Private IP addresses. [What does this mean?](#)

Heads up

Private networking is now automatically enabled. You can create new networks or just use the default.

OK [Learn more](#)

Select additional options ?

IPv6 User data Monitoring

Authentication ?

SSH keys
A more secure authentication method

Password
Create a root password to access Droplet (less secure)

Create root password *

PASSWORD REQUIREMENTS

- Must be at least 8 characters long
- Must contain 1 uppercase letter (cannot be first or last character)
- Must contain 1 number
- Cannot end in a number or special character

▲ Please store your password securely. You will not be sent an email containing the Droplet's details or password.

Figure 3.5

Figure 3.4 & 3.5 shows the option how to create a sensor. The first option is to get the list of the operating systems, and if we click on them, we can use them. The next option is regarding the cloud facilities; either we want to share them or the machine or the rack of others. Then we have an option of different CPU and SSD option. After selecting the plan, the next option is for the datacentre region, which means we can locate our VM on their data centre and by default, we will get the IP of that region. After that, the last option is for authentication. We can use SSH Keys or password.

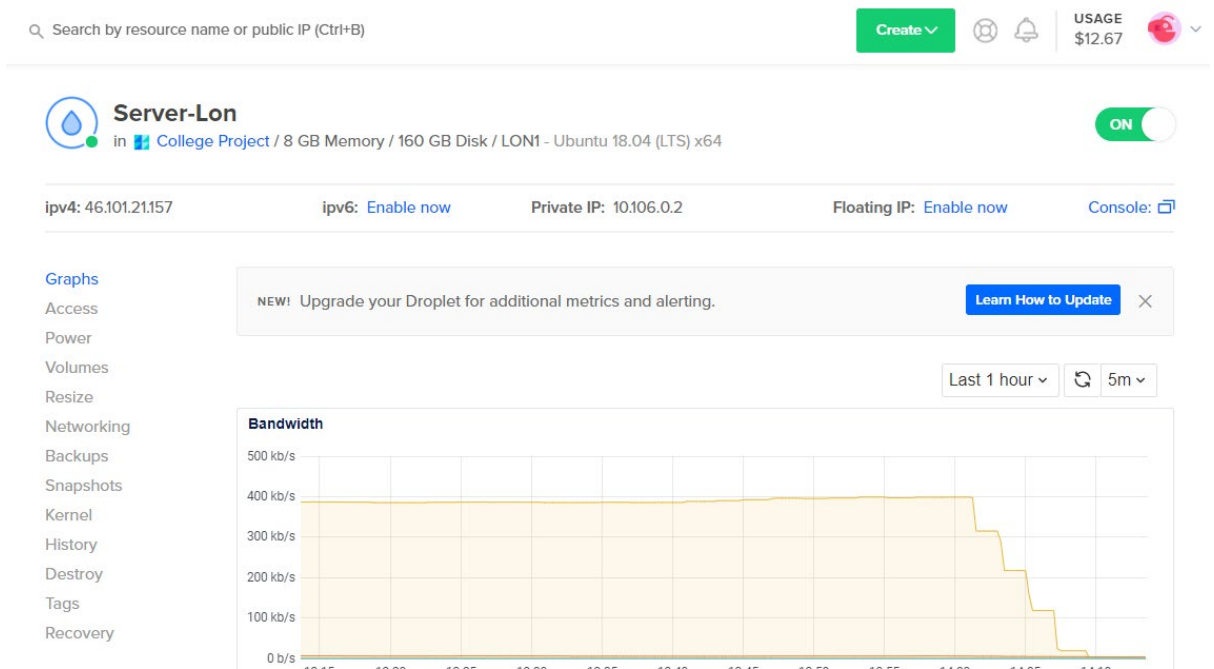


Figure:3.6

Figure 3.6 shows us the technology I am using to run my server. I am running my MHN server at the datacentre of London, which is on 8GB RAM and 160 GB SSD. The Operating system I am using is Ubuntu 18.04.

The following screenshots show the implementation of the MHN server on Digital Ocean.

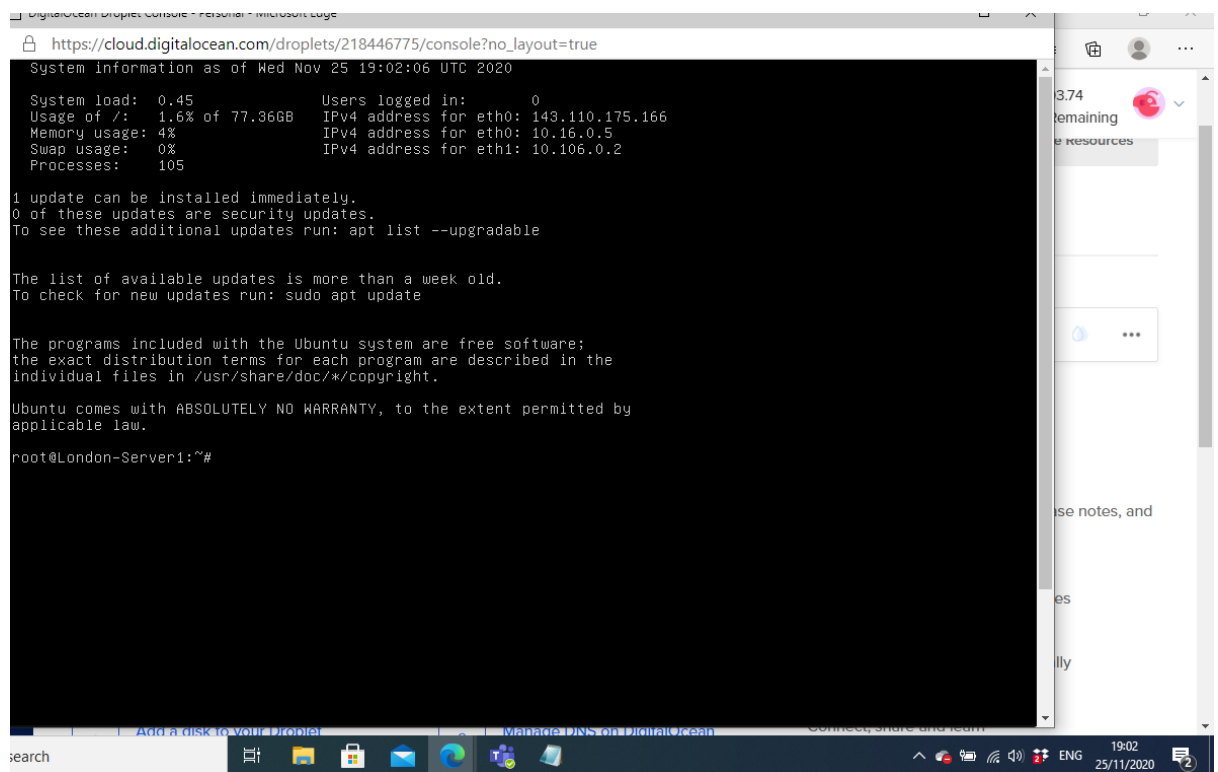
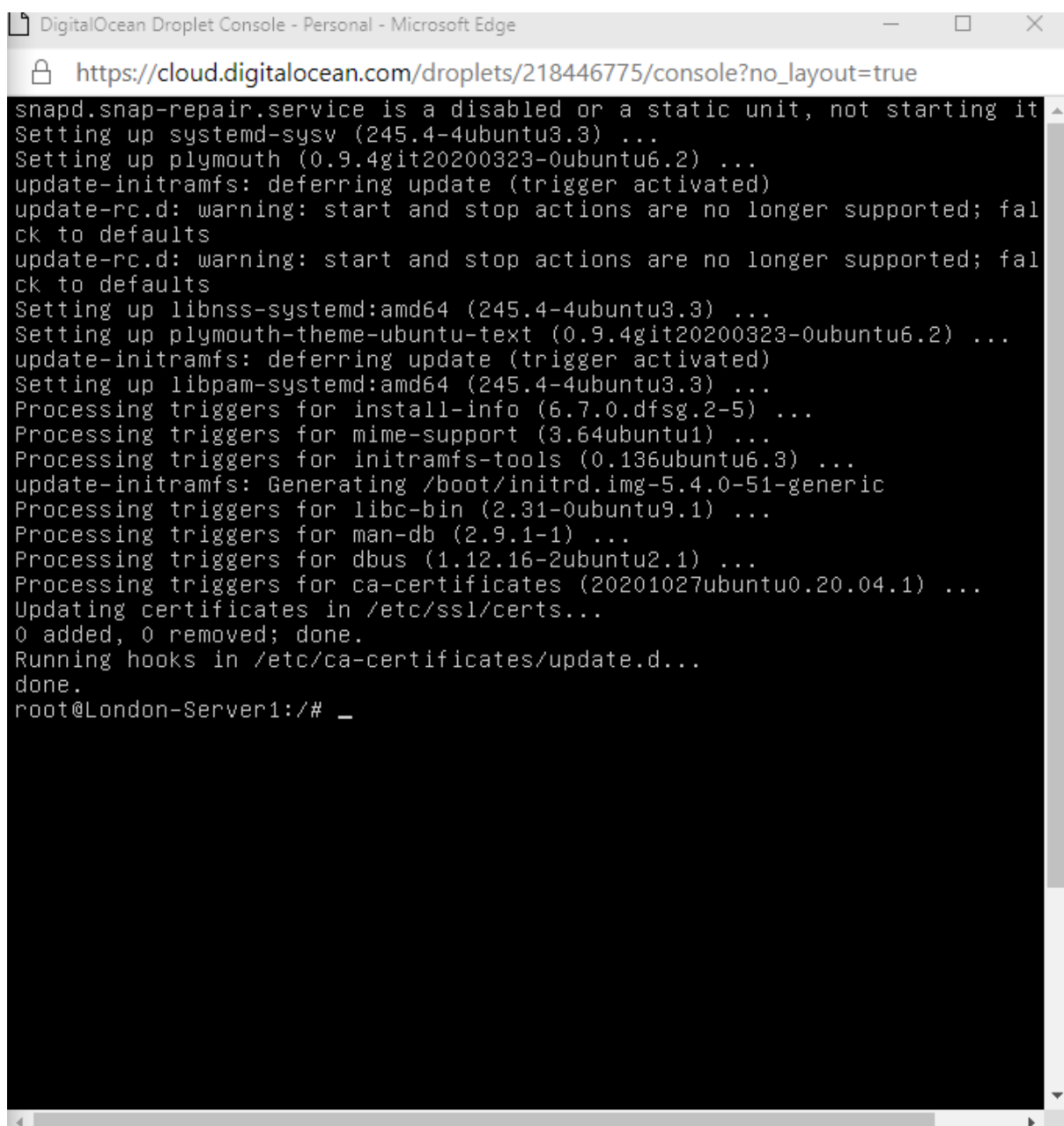


Figure:3.7

The image shows a terminal window from a DigitalOcean Droplet Console. The window title is "DigitalOcean Droplet Console - Personal - Microsoft Edge". The address bar shows the URL "https://cloud.digitalocean.com/droplets/218446775/console?no_layout=true". The terminal output displays the following text:

```
snapped.service is a disabled or a static unit, not starting it
Setting up systemd-sysv (245.4-4ubuntu3.3) ...
Setting up plymouth (0.9.4git20200323-0ubuntu6.2) ...
update-initramfs: deferring update (trigger activated)
update-rc.d: warning: start and stop actions are no longer supported; fall
back to defaults
update-rc.d: warning: start and stop actions are no longer supported; fall
back to defaults
Setting up libnss-systemd:amd64 (245.4-4ubuntu3.3) ...
Setting up plymouth-theme-ubuntu-text (0.9.4git20200323-0ubuntu6.2) ...
update-initramfs: deferring update (trigger activated)
Setting up libpam-systemd:amd64 (245.4-4ubuntu3.3) ...
Processing triggers for install-info (6.7.0.dfsg.2-5) ...
Processing triggers for mime-support (3.64ubuntu1) ...
Processing triggers for initramfs-tools (0.136ubuntu6.3) ...
update-initramfs: Generating /boot/initrd.img-5.4.0-51-generic
Processing triggers for libc-bin (2.31-0ubuntu9.1) ...
Processing triggers for man-db (2.9.1-1) ...
Processing triggers for dbus (1.12.16-2ubuntu2.1) ...
Processing triggers for ca-certificates (20201027ubuntu0.20.04.1) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
root@London-Server1:/# _
```

Figure:3.8

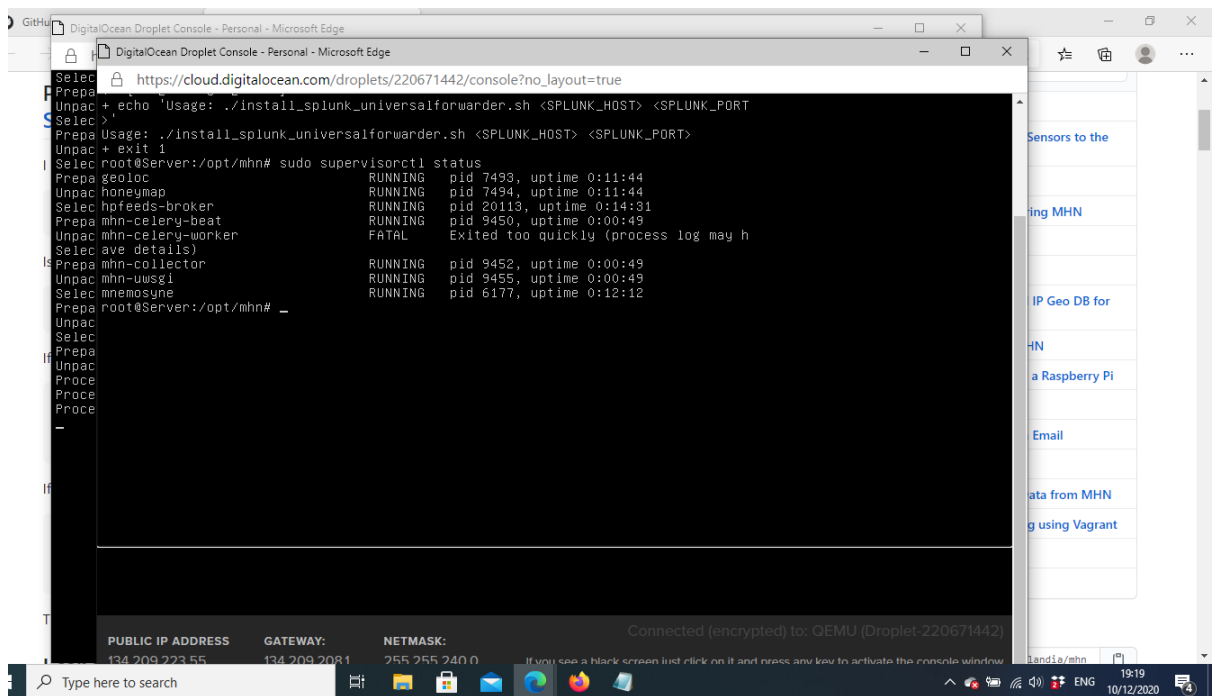
```
update-initramfs: deferring update (trigger activated)
Setting up libpam-systemd:amd64 (245.4-4ubuntu3.3) ...
Processing triggers for install-info (6.7.0.dfsg.2-5) ...
Processing triggers for mime-support (3.64ubuntu1) ...
Processing triggers for initramfs-tools (0.136ubuntu6.3) ...
update-initramfs: Generating /boot/initrd.img-5.4.0-51-generic
Processing triggers for libc-bin (2.31-0ubuntu9.1) ...
Processing triggers for man-db (2.9.1-1) ...
Processing triggers for dbus (1.12.16-2ubuntu2.1) ...
Processing triggers for ca-certificates (20201027ubuntu0.20.04.1) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
root@London-Server1:/# sudo apt-get install git
Reading package lists... Done
Building dependency tree
Reading state information... Done
git is already the newest version (1:2.25.1-1ubuntu3).
git set to manually installed.
0 upgraded, 0 newly installed, 0 to remove and 4 not upgraded.
root@London-Server1:/# cd/opt
-bash: cd/opt: No such file or directory
root@London-Server1:/# cd /opt
root@London-Server1:/opt# _
```

Figure:3.9

```
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
root@London-Server1:~# sudo apt-get install git
Reading package lists... Done
Building dependency tree
Reading state information... Done
git is already the newest version (1:2.25.1-1ubuntu3).
git set to manually installed.
0 upgraded, 0 newly installed, 0 to remove and 4 not upgraded.
root@London-Server1:~# cd /opt
-bash: cd /opt: No such file or directory
root@London-Server1:~# cd /opt
root@London-Server1:~# sudo git clone https://github.com/threatstream/mhn
Cloning into 'mhn'...
remote: Enumerating objects: 60, done.
remote: Counting objects: 100% (60/60), done.
remote: Compressing objects: 100% (57/57), done.
remote: Total 7515 (delta 6), reused 47 (delta 1), pack-reused 7455
Receiving objects: 100% (7515/7515), 3.71 MiB | 875.00 KiB/s, done.
Resolving deltas: 100% (4068/4068), done.
root@London-Server1:~#
```

Figure:3.10

Figure 3.8 – 3.10 shows us the screenshots of updating the system and cloning the git repository at cd/opt. The cloned git is [GitHub - pwnlandia/mhn: Modern Honey Network](https://github.com/pwnlandia/mhn).



```
Selec
Prepa
Unpac + echo "Usage: ./install_splunk_universalforwarder.sh <SPLUNK_HOST> <SPLUNK_PORT>
Selec >
Prepa Usage: ./install_splunk_universalforwarder.sh <SPLUNK_HOST> <SPLUNK_PORT>
Unpac + exit 1
Selec root@Server:/opt/mhn# sudo supervisorctl status
Prepa geoloc RUNNING pid 7493, uptime 0:11:44
Unpac honeymap RUNNING pid 7494, uptime 0:11:44
Selec hpfeeds-broker RUNNING pid 20113, uptime 0:14:31
Prepa mhn-celery-beat RUNNING pid 9450, uptime 0:00:49
Unpac mhn-celery-worker FATAL Exited too quickly (process log may h
Selec ave details)
Selec mhn-collector RUNNING pid 9452, uptime 0:00:49
Unpac mhn-uiscgi RUNNING pid 9455, uptime 0:00:49
Selec mmosyne RUNNING pid 6177, uptime 0:12:12
Prepa root@Server:/opt/mhn# _
Unpac
Selec
Prepa
Unpac
Proce
Proce
Proce
```

Figure:3.11

Figure 3.11 shows us the error I received after running a supervisorctl status to check if the MHN server is functioning as it supposes to be. Unfortunately, the error I got was complicated as the mhn-celery worker was not working. I could not check the attacker's location after the first attack, and the rest of the attacks were not getting stored on var/log.

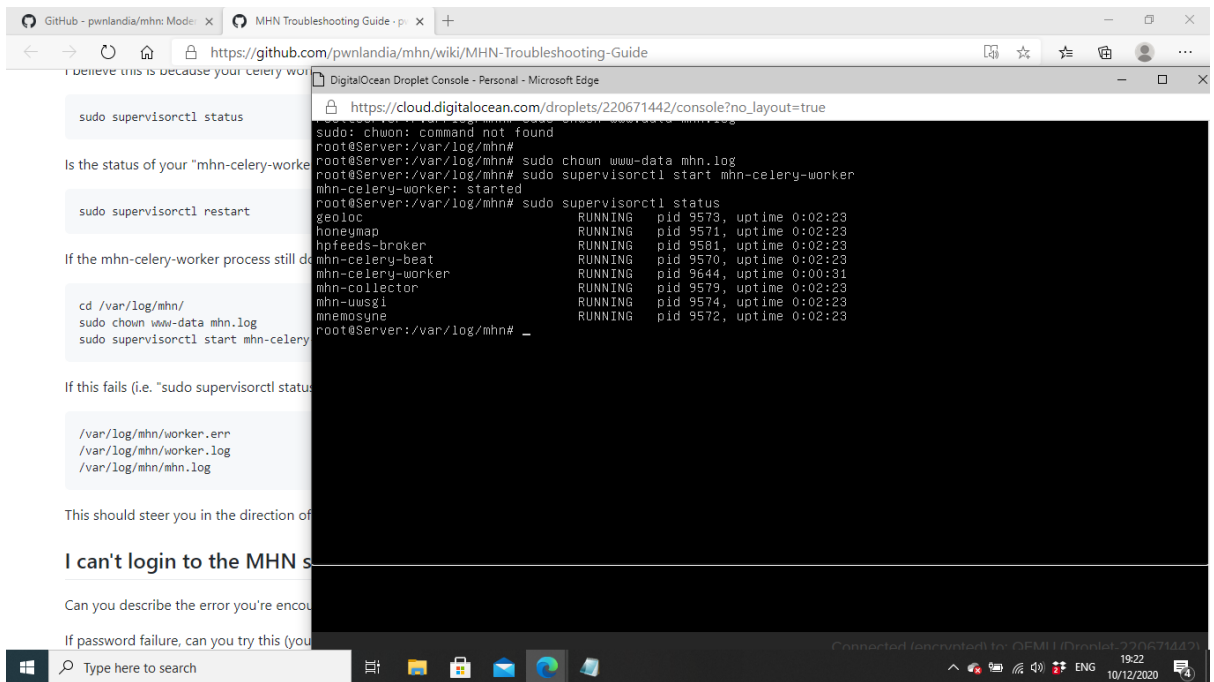


Figure:3.12

Figure 3.12 shows that I was able to debug the error, and mhn-celery-worker is running the commands I ran was `sudo chown www-data mhn.log` and then restart the server by `sudo supervisor start mhn-celery-worker`.

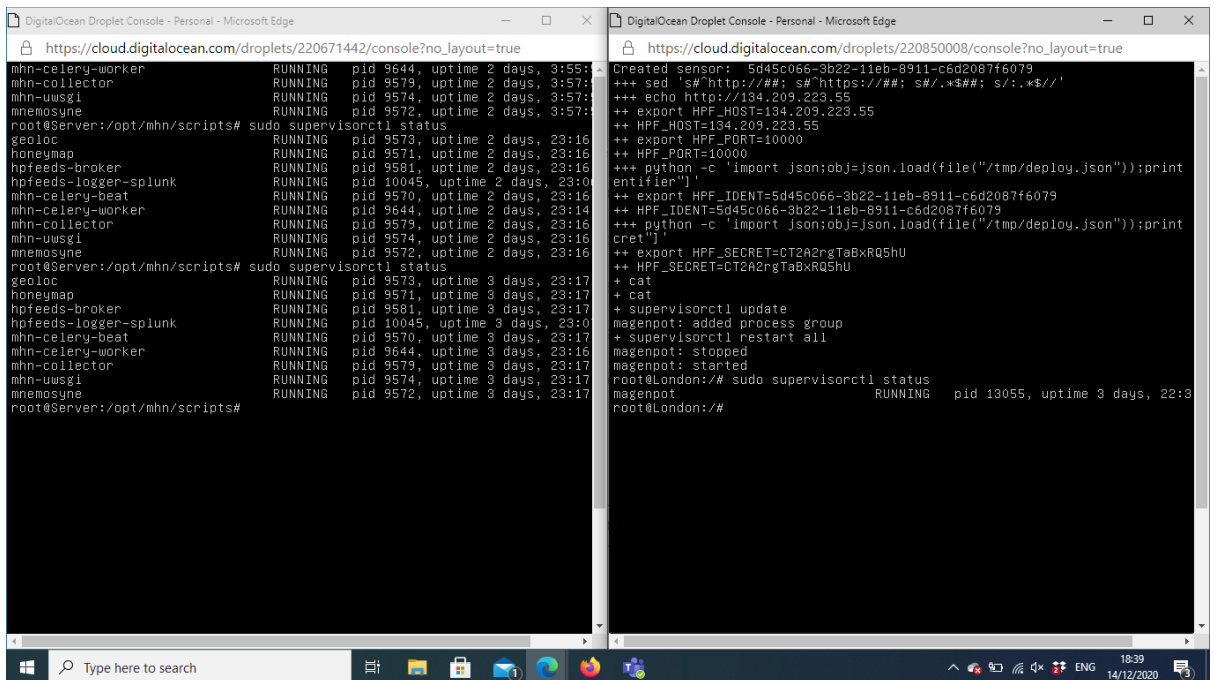


Figure:3.13

Figure: 3.13 shows us the integration of Splunk. Hpfeeds-logger-Splunk is running, enabling Splunk to collect the data from `var/log/mhn` when Splunk is integrated.

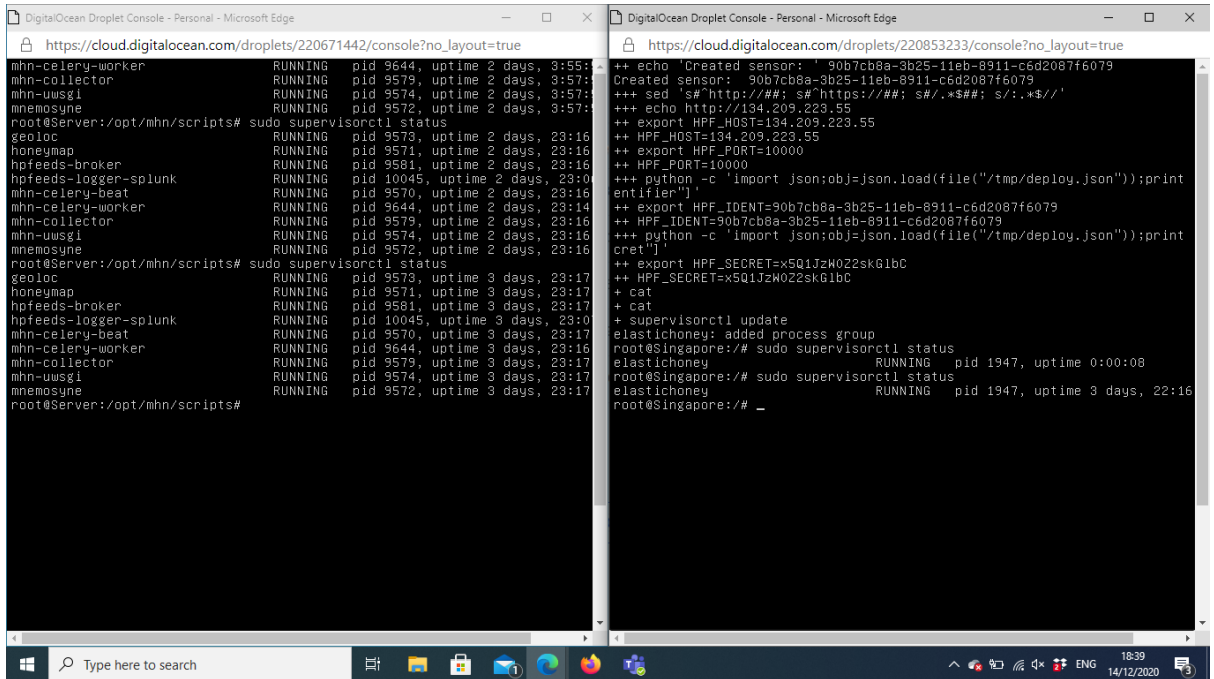


Figure:3.14

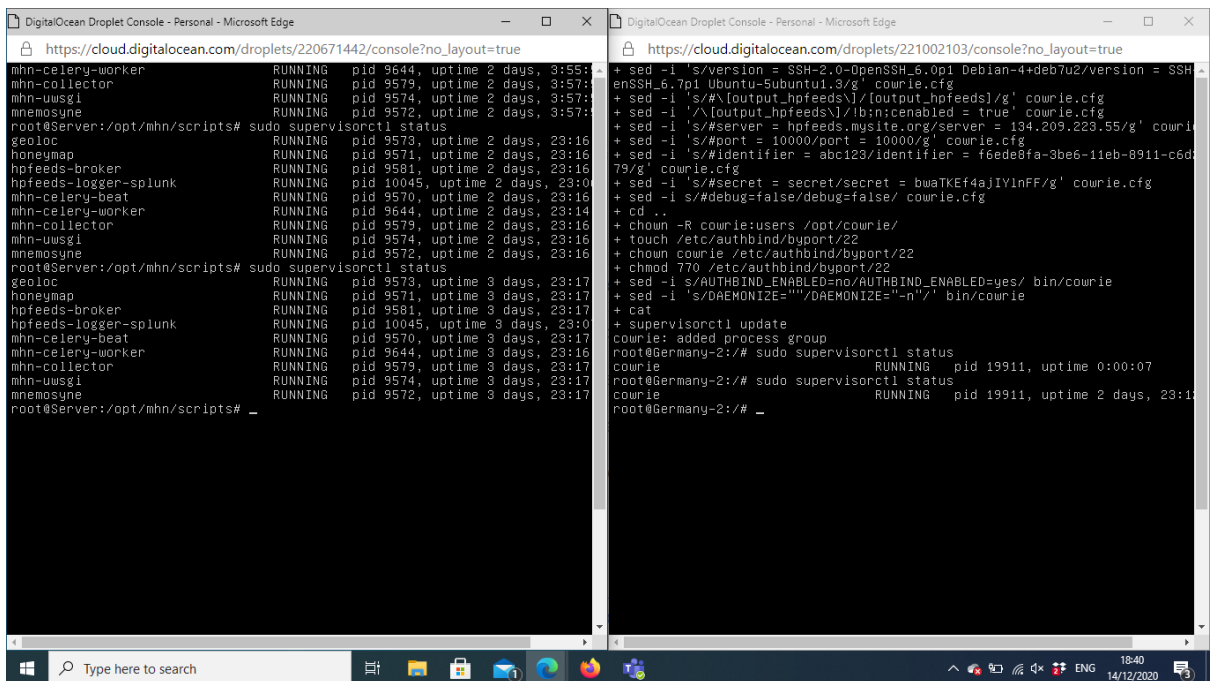


Figure:3.15

Figure 3.14 & 3.15 shows the honeypots' successful installation, but I ran into an error when I tried to install cowrie as a honeypot. The cowrie was one of the vital honeypots I planned to use for my project because it logs the data from brute force.

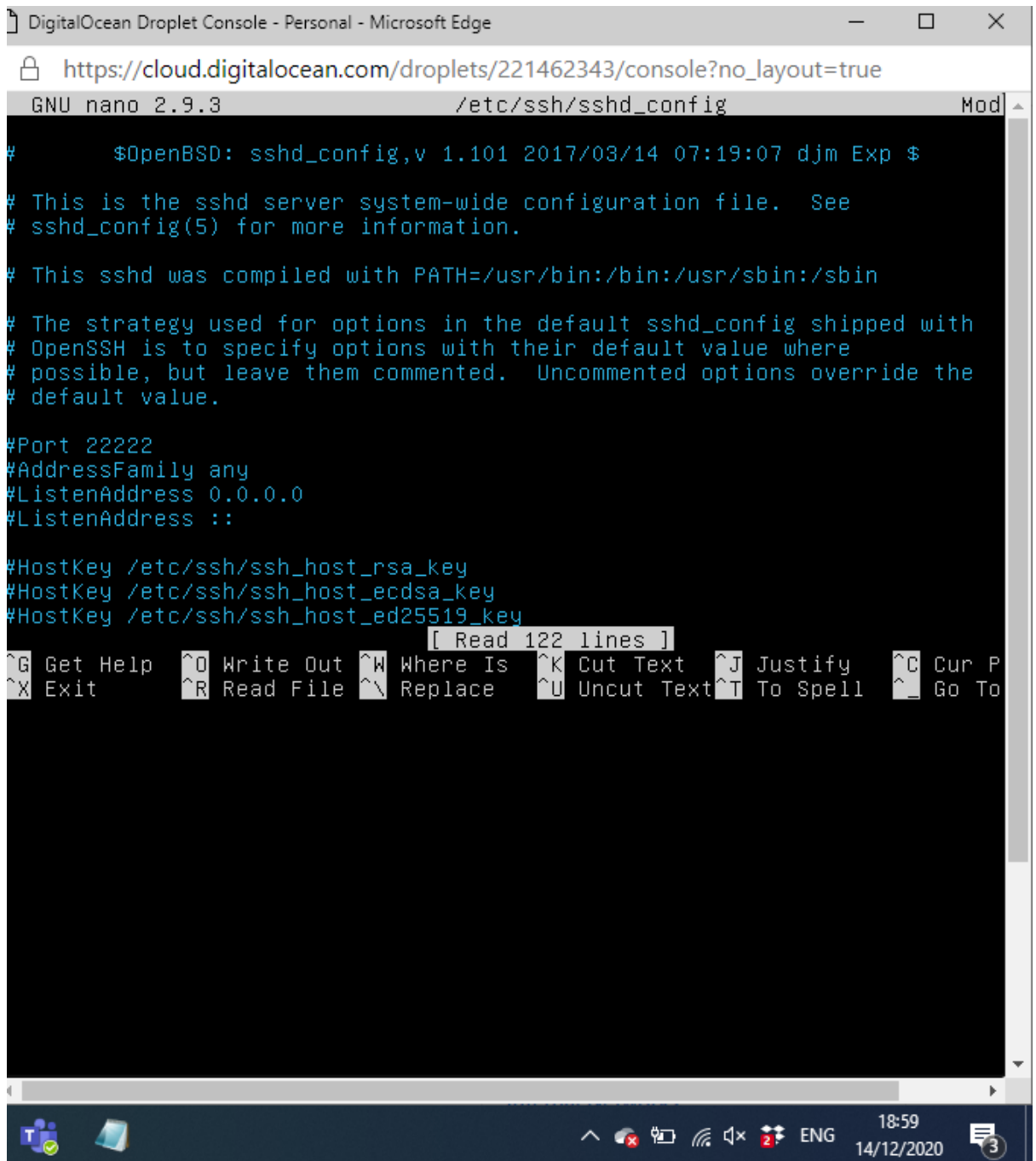


Figure:3.16

```
DigitalOcean Droplet Console - Personal - Microsoft Edge
https://cloud.digitalocean.com/droplets/221462343/console?no_layout=true

#AddressFamily any
#ListenAddress 0.0.0.0
#ListenAddress ::

#HostKey /etc/ssh/ssh_host_rsa_key
#HostKey /etc/ssh/ssh_host_ecdsa_key
#HostKey /etc/ssh/ssh_host_ed25519_key

root@Apots-USA:~# sudo systemctl restart ssh
root@Apots-USA:~# systemctl status ssh
* ssh.service - OpenBSD Secure Shell server
  Loaded: loaded (/lib/systemd/system/ssh.service; enabled; vendor preset: enab
  Active: active (running) since Mon 2020-12-14 18:59:57 UTC; 8s ago
  Process: 1529 ExecStartPre=/usr/sbin/sshd -t (code=exited, status=0/SUCCESS)
  Main PID: 1540 (sshd)
  Tasks: 1 (limit: 1152)
  CGroup: /system.slice/ssh.service
          └─1540 /usr/sbin/sshd -D

Dec 14 18:59:57 Apots-USA systemd[1]: Starting OpenBSD Secure Shell server...
Dec 14 18:59:57 Apots-USA sshd[1540]: Server listening on 0.0.0.0 port 22.
Dec 14 18:59:57 Apots-USA sshd[1540]: Server listening on :: port 22.
Dec 14 18:59:57 Apots-USA systemd[1]: Started OpenBSD Secure Shell server.
lines 1-13/13 (END)
```

Figure:3.17

```
DigitalOcean Droplet Console - Personal - Microsoft Edge
https://cloud.digitalocean.com/droplets/221462343/console?no_layout=true

+ sed -i 's/listen_endpoints = tcp:2222:interface=0.0.0.0/listen_endpoints = tcp
:22:interface=0.0.0.0/g' cowrie.cfg
+ sed -i 's/version = SSH-2.0-OpenSSH_6.0p1 Debian-4+deb7u2/version = SSH-2.0-Op
enSSH_6.7p1 Ubuntu-Subuntu1.3/g' cowrie.cfg
+ sed -i 's/#[output_hfeeds\]/[output_hfeeds]/g' cowrie.cfg
+ sed -i 's/#server = hfeeds.musite.org/server = 139.59.47.212/g' cowrie.cfg
+ sed -i 's/#port = 10000/port = 10000/g' cowrie.cfg
+ sed -i 's/#identifier = abc123/identifier = 0bc4a052-3e3f-11eb-907e-aa99a75422
d0/g' cowrie.cfg
+ sed -i 's/#secret = secret/secret = XnDg1Jp5aC28N1M9/g' cowrie.cfg
+ sed -i 's/#debug=false/debug=false/' cowrie.cfg
+ cd ..
+ chown -R cowrie:users /opt/cowrie/
+ touch /etc/authbind/byport/22
+ chown cowrie /etc/authbind/byport/22
+ chmod 770 /etc/authbind/byport/22
+ sed -i 's/AUTHIND_ENABLED=no/AUTHIND_ENABLED=yes/' bin/cowrie
+ sed -i 's/DAEMONIZE=""/DAEMONIZE="-n"/' bin/cowrie
+ cat
+ supervisorctl update
cowrie: added process group
root@Apots-USA:~# sudo supervisorctl status
cowrie          RUNNING pid 19985, uptime 0:00:51
root@Apots-USA:~#
```

Figure:3.18

Figure 3.16 - 3.18 shows the debugging of the cowrie, in which I changed the port and cowrie was added successfully.

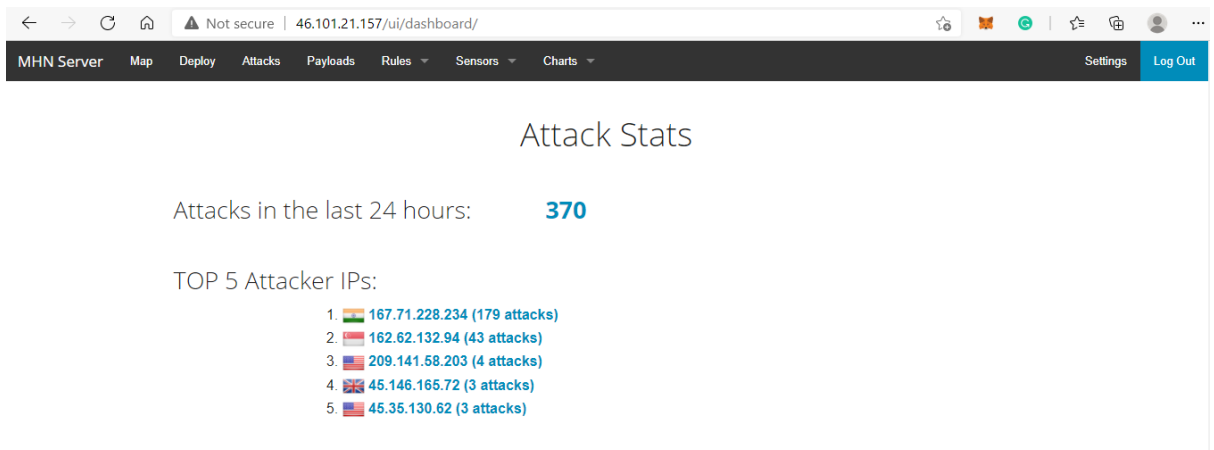


Figure:3.20

Figure 3.20 shows us the MHN main dashboard of the MHN server. I am currently running the Ip address of my server is 46.101.21.157, which the following screenshot can confirm.



Figure:3.21

As the first two-step was successfully installed/configured and debugged, the final and vital part of the project was the installation of Splunk.

As mentioned earlier, the Splunk hpfeed logger was installed in the early stages of the project by the following commands:

```
cd /opt/mhn/scripts/
sudo ./install_hpfeeds-logger-splunk.
```

The next step was to install Splunk on the MHN server to do; I required to sign up with Splunk. Once I created the Splunk account, I was given two options to run an enterprise trial or cloud trial. I picked cloud. On the page, the option where given which operating system I am using. After selecting the operating system, which was

Linux ubuntu, I was given a wget command, which I pasted in the server to get the Splunk.

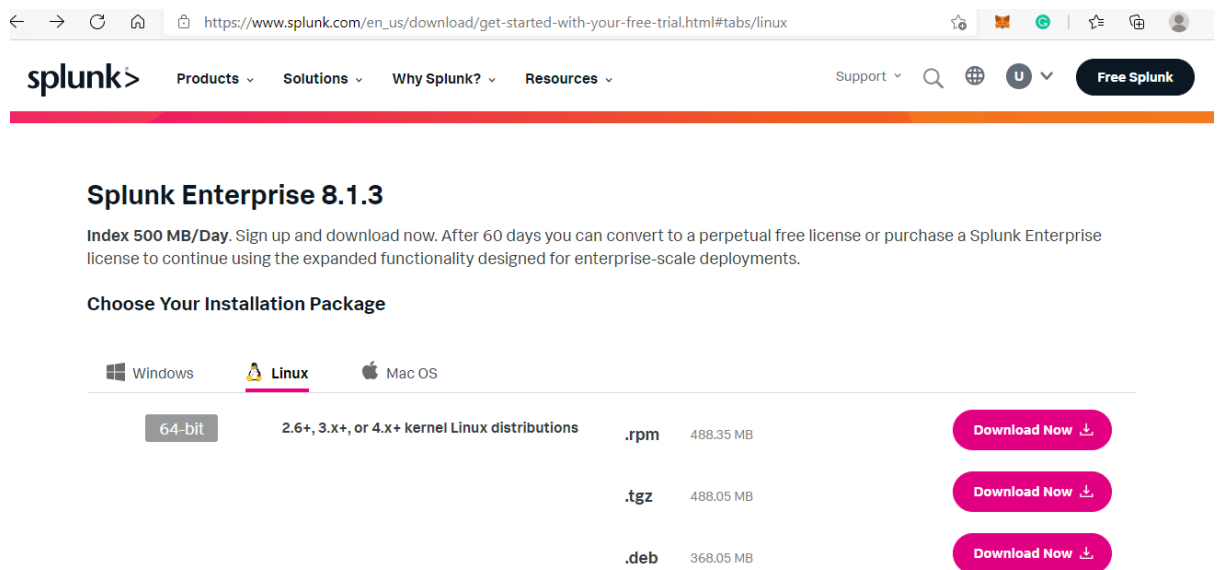


Figure: 3.22

Figure 3.22 shows us the page after the sign-up. The .tgz file was downloaded into the MHN server and then extracted.

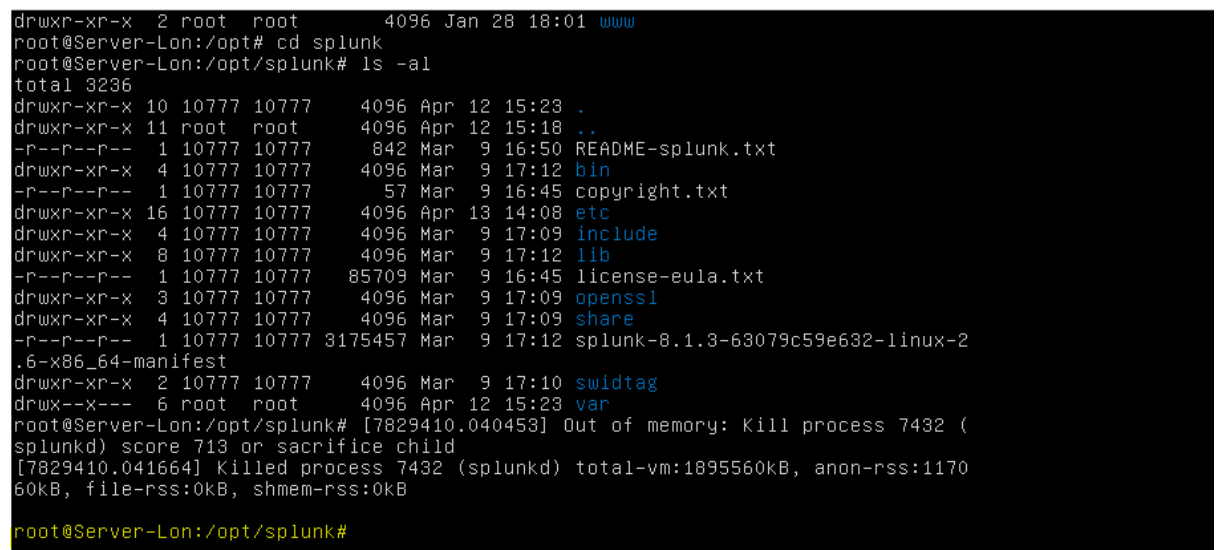


Figure: 3.23

In figure 3.23 the Splunk was successfully installed. The following screenshot will confirm it.

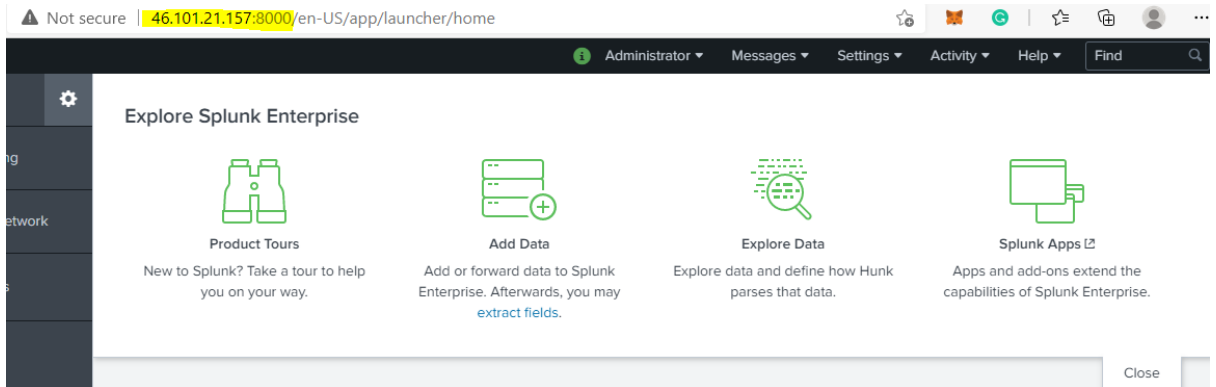


Figure: 3.24

Figure 3.21 and Figure: 3.24, the Ip address is the same. So, it shows that the MHN server I am working on the Splunk, is integrated successfully to the server.

Graphical User Interface (GUI)

I will be going through the GUI of MHN Server, Digital Ocean and Splunk GUI in the GUI section.

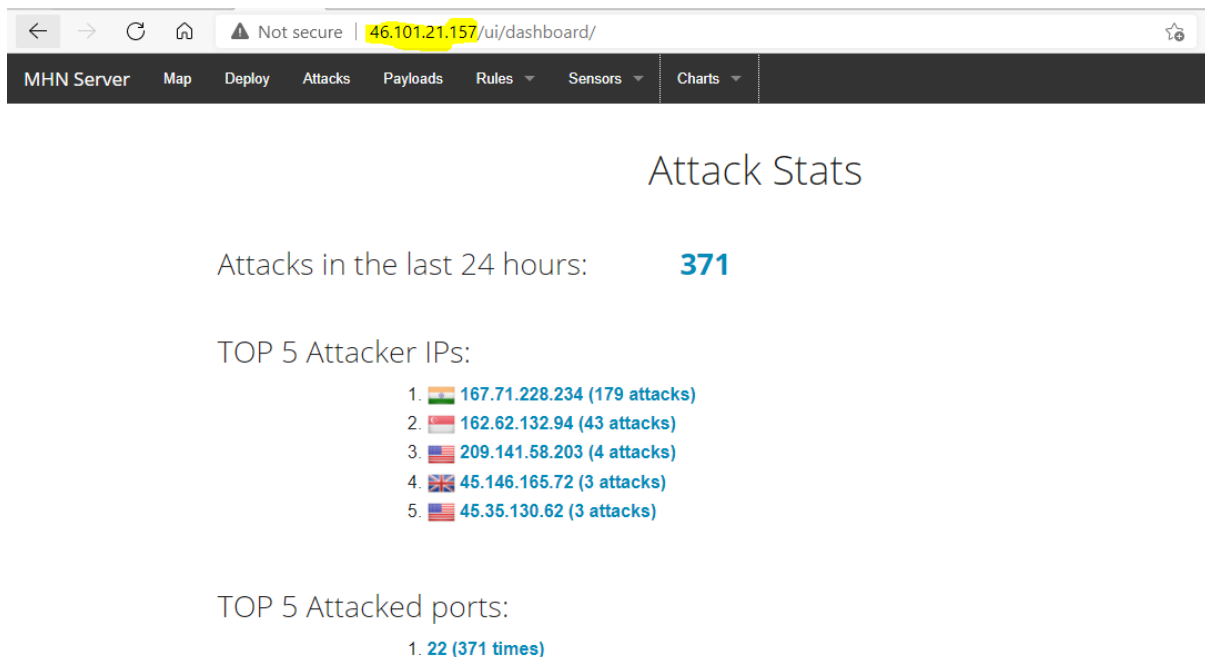


Figure 3.1.1

Figure 3.1.1 shows us the main dashboard of the MHN server, where we can see the stats of the past 24 hours, top 5 attackers, top 5 attacked ports, top 5 honey pots, top 5 sensors to attract the attackers.

Attacks Report

Search Filters

Sensor: Honeypot: Date: Port: IP Address:

	Date	Sensor	Country	Src IP	Dst port	Protocol	Honeypot
1	2021-05-08 14:04:20	USA		209.141.47.246	22	ssh	cowrie
2	2021-05-08 14:04:17	USA		209.141.47.246	22	ssh	cowrie
3	2021-05-08 14:04:14	USA		209.141.47.246	22	ssh	cowrie
4	2021-05-08 14:04:09	USA		209.141.47.246	22	ssh	cowrie
5	2021-05-08 14:04:05	USA		209.141.47.246	22	ssh	cowrie
6	2021-05-08 14:03:48	India		217.170.206.138	22	ssh	cowrie
7	2021-05-08 14:03:13	USA		209.141.47.246	22	ssh	cowrie

Figure 3.1.2

Figure 3.1.2 once the attacker successfully launches the attack, it logs it on var/log/mhn and then shows us the date, the name of the honeypot, the country from where the attack was launched, the IP address of the attacker, the attacked port and protocol.

Sensors

	Name	Hostname	IP	Honeypot	UUID	Attacks
1-	<input type="text" value="Dionaea-Frankfurt-dionaea"/>	Dionaea-Frankfurt	64.227.126.27	dionaea	88f750a6-6194-11eb-99bd-aa50c7416afa	207705
2-	<input type="text" value="Dionaea-NewYork-dionaea"/>	Dionaea-NewYork	164.90.142.233	dionaea	c726a4ac-6195-11eb-99bd-aa50c7416afa	3450161
3-	<input type="text" value="Dionaea-Bangalore-dionaea"/>	Dionaea-Bangalore	139.59.42.145	dionaea	fa8aa55e-6196-11eb-99bd-aa50c7416afa	37173
4-	<input type="text" value="Cowrie-Canada-cowrie"/>	Cowrie-Canada	159.203.14.64	cowrie	457da44e-67eb-11eb-99bd-aa50c7416afa	232056
5-	<input type="text" value="test-agave"/>	test	164.90.140.28	agave	d9b3b5ba-722a-11eb-99bd-aa50c7416afa	4
6-	<input type="text" value="16-agave"/>	16	157.245.118.36	agave	0e5ca196-722b-11eb-99bd-aa50c7416afa	10
7-	<input type="text" value="conp-conpot"/>	conp	143.198.12.63	conpot	d5253f54-722b-11eb-99bd-aa50c7416afa	0
8-	<input type="text" value="ubuntu-s-dionaea"/>	ubuntu-s	139.59.68.147	dionaea	de8d1b74-722c-11eb-99bd-aa50c7416afa	16
9-	<input type="text" value="pie-Drugpot-USA-agave"/>	pie-Drugpot-USA	64.225.29.75	agave	474d8944-722f-11eb-99bd-aa50c7416afa	8627
10-	<input type="text" value="Cowrie-Singapore-cowrie"/>	Cowrie-Singapore	206.189.94.106	cowrie	6fea3c30-722f-11eb-99bd-aa50c7416afa	131664

Figure 3.1.3

Figure 3.1.3 shows us the sensor or the honeypots attached to our system with the number of attacks to the virtual machine.

uri	daddr	saddr	dport	sport	sha512
178.128.228.248	109.64.47.164	445	59453	053e31bb5d469a2c4b1c5ab658d87051168c8a0b8d55d1709bcc4c11faf16fd1617263c5c30c4c9bfb5b319ead6d2712fafc5a40888f5a6b46d1eb0630	
178.128.228.248	200.71.185.97	445	59573	8062093734b11fdd2a8650bfc22f36aa679103e7a7ebee74db1eectdbf9d9b76d105f395308db713746dbadacc5796db85ab883a4187587f03b2d3cf71	
178.128.228.248	117.220.100.97	445	62269	3b3ff7e0e6dd3f1bebe0346068aab8560730d8674fc3f9edeb79e704d312a0a6d8737778eb37af7e2992fe822e192cf8be377870f2c47dd075325e0c0a7b9	
178.128.228.248	191.31.226.159	445	65120	05ee932752ae00a3e07f673b777249ee6bda150e864fc7bbfcee13d02e2da99838b708120a76610ac49fba3b0a686d61a39f188719b302716332286a4	
178.128.228.248	117.2.24.53	445	54590	64f223e762c17b750790a8ec483319e851e317164d662e8d1d5e8b3551e297f115e51ef52dfc350d5f9f29ae4ad146501c343778671dbb934fc0116b50	
178.128.228.248	80.24.131.111	445	65505	d083def96ec407ea201974ba26068c492f1d7db42107ac455dd2550ce74d261758304eeddb6b65fac40a813bec64c8a3ef0b5a1e97995e8a59bd2db125	
178.128.228.248	191.31.226.159	445	61445	05ee932752ae00a3e07f673b777249ee6bda150e864fc7bbfcee13d02e2da99838b708120a76610ac49fba3b0a686d61a39f188719b302716332286a4	
178.128.228.248	168.194.72.50	445	63208	8cacc0e9de57754fe52720d0e48aef081a48558120886fb92719428dc7bcb1928d6f049dc006ea4d07844dd8ae58d76b36d7739731e9bdf2232ed639c	
178.128.228.248	189.111.192.80	445	39503	53ace4673ac8ab17b3aee3360697e4f70296d7c0b951feddb77ccccf41c794fa3045b680eafb076061d7b09b554fc24fe4755041e411f7ee43d2bee83b7	
178.128.228.248	181.115.185.34	445	11402	8a1c007f51ec141ce7bacc5966eb27296ae2c443bd62ebf3471ae7a4e879f324d2949ca70d73f740a5e9d22947ba7eb229f639b5a30001eaf5cfd7cb68	

Figure 3.1.4

Figure 3.1.4 shows us the payload report of attacks, as I worked with cowrie and dionaea. The dionaea is known for attracting malware attacks. So the payload report is very beneficial for analysis to check which type of payloads are commonly used nowadays.



Figure 3.1.5

Figure 3.1.5 shows us the map. The map's purpose is to see live attacks from where the attackers are attacking and on which honeypot. When I took the screenshot, no honeypot was added to my MHN server. That is why it shows no attack but fetching the old data of attacks. I was able to show the attacks on Splunk on the following screenshot:

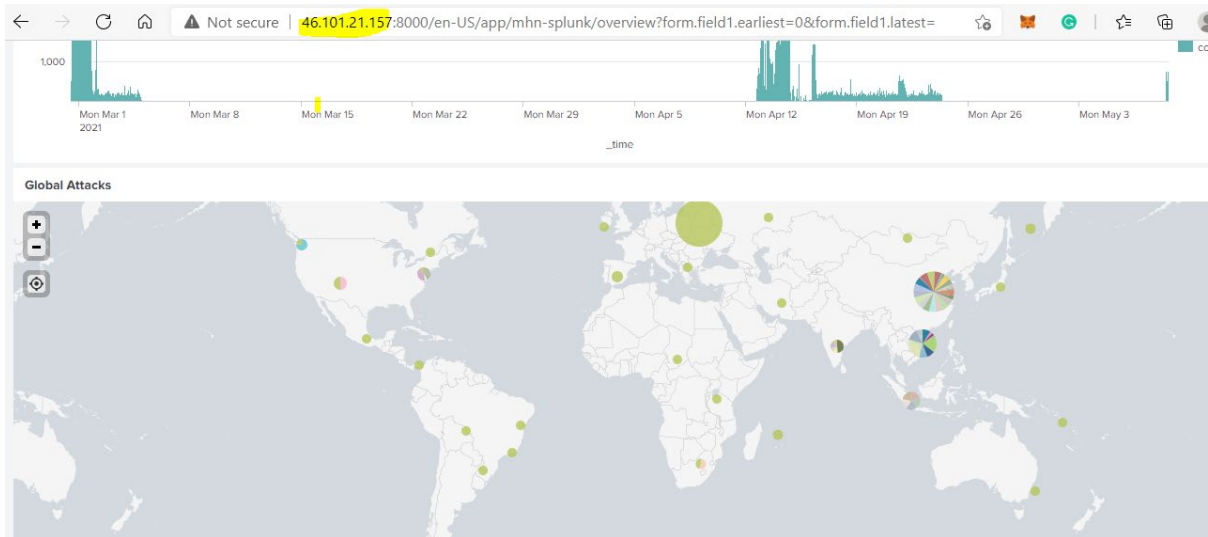


Figure 3.1.6



Figure 3.1.7

Figure 3.1.7 shows us the main dashboard of the Digital Ocean. Where we can create a different project to centralised things. As you can see, I am running my server on a College project.

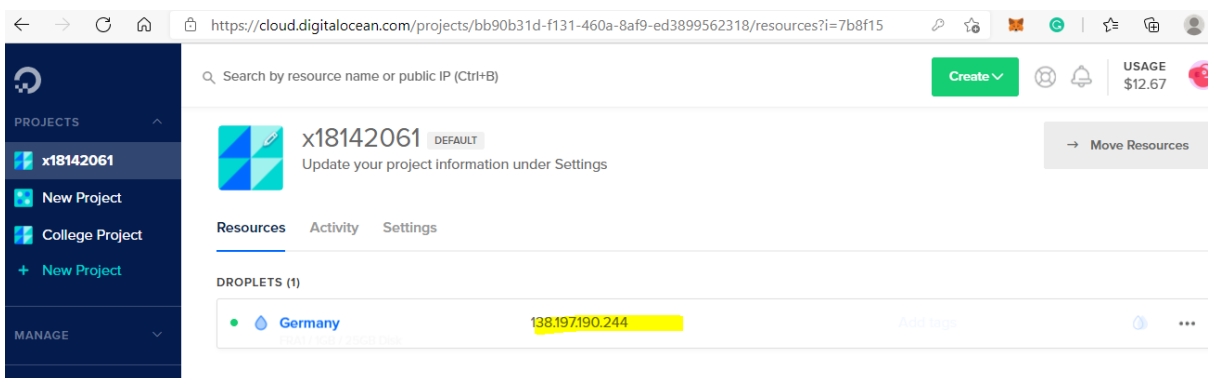


Figure 3.1.8

In figure 3.1.8, I created a cowrie honeypot under my student id project and added it to the MHN server.

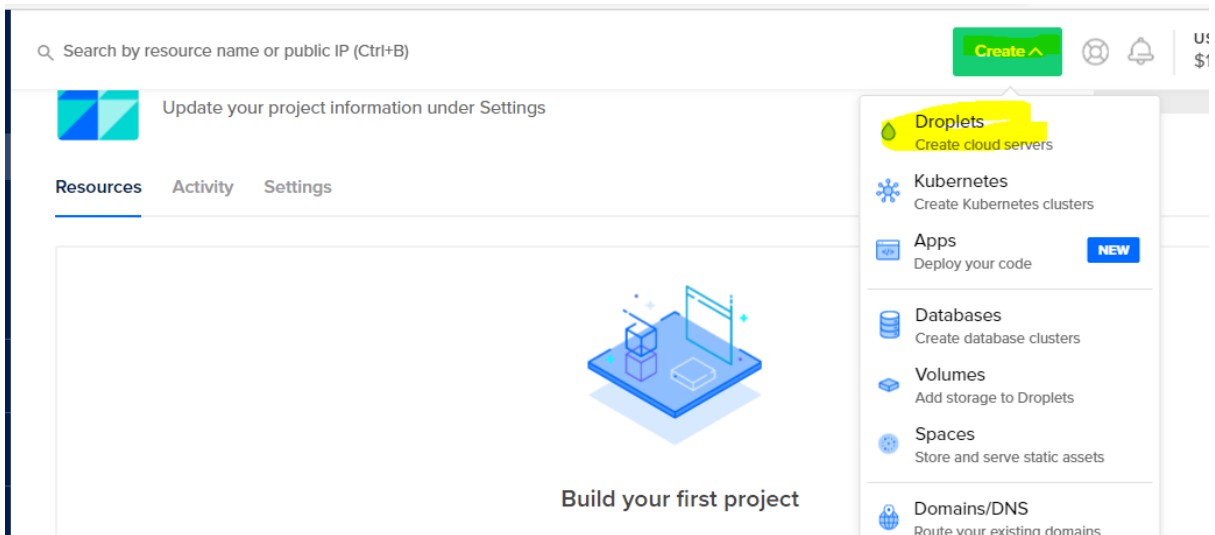


Figure 3.1.9

The create option for honeypots.

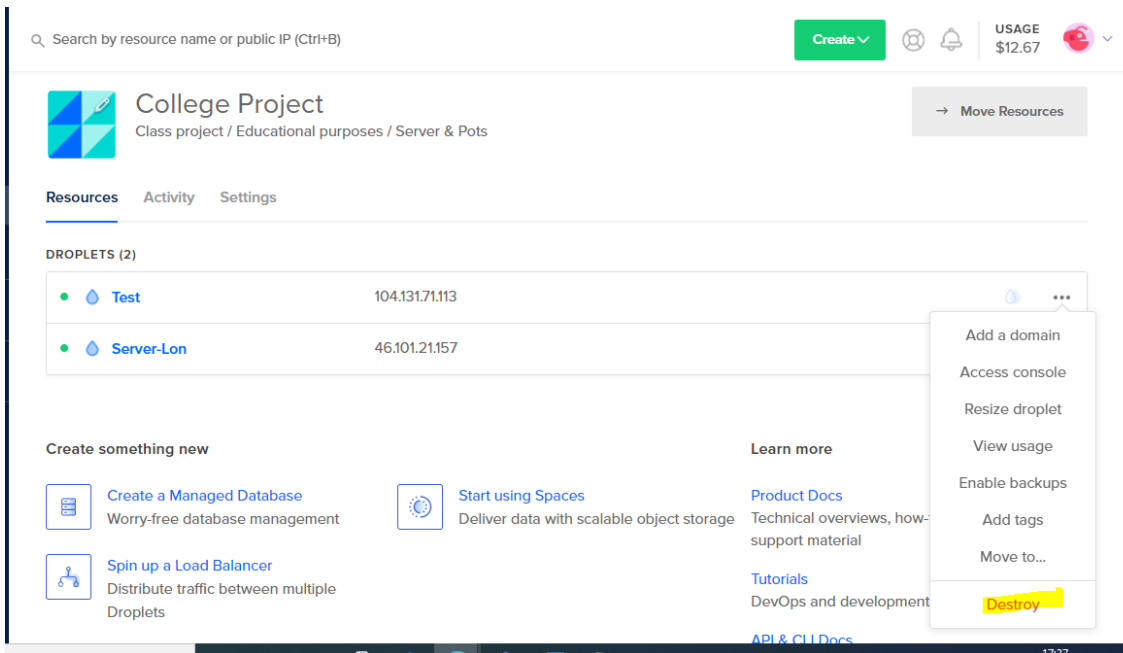


Figure 3.1.10

The delete option. Once we decided to delete the honeypot, we need to click on the three dots available on the right-hand side option. After clicking, a dropdown menu will give us the option to destroy. Once we click the destroy, we will be promoted to the following page:

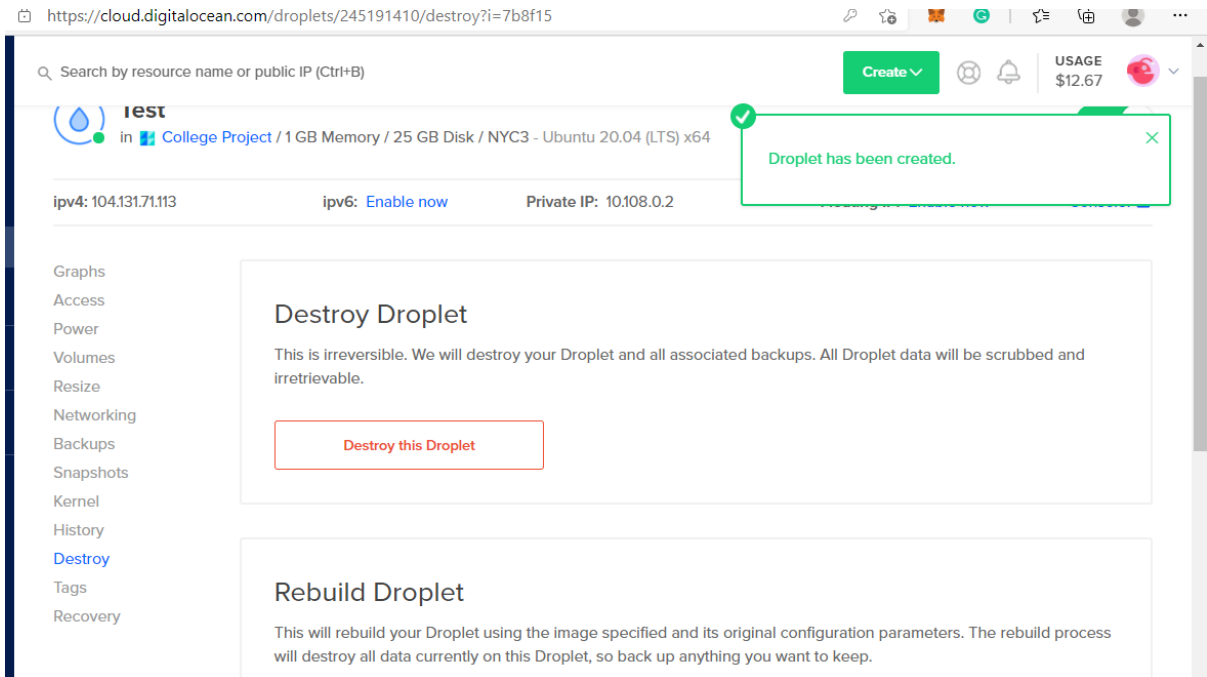


Figure 3.1.11

Then we need to click on the “Destroy this droplet”. Once we hit the button, we will be promoted by the following option:

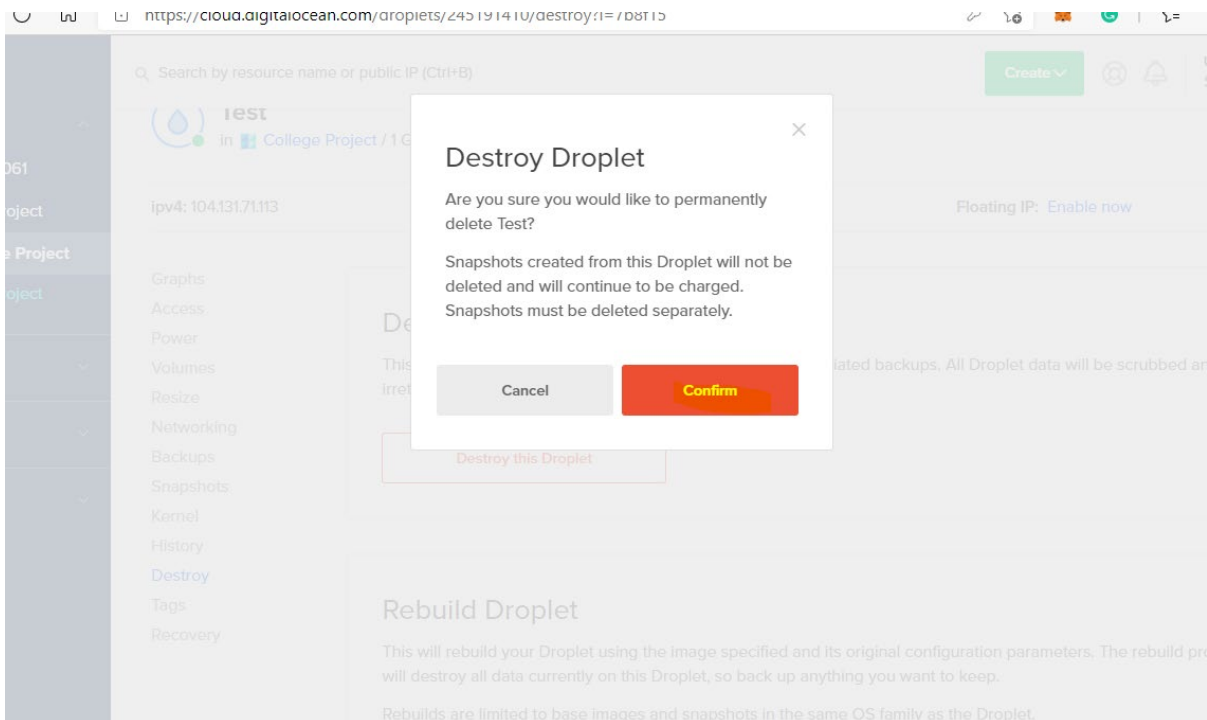


Figure 3.1.12

To delete, we need to click on confirm. Once it has been clicked, the droplet/sensor/honeypot is deleted and unable to retrieve our decision.

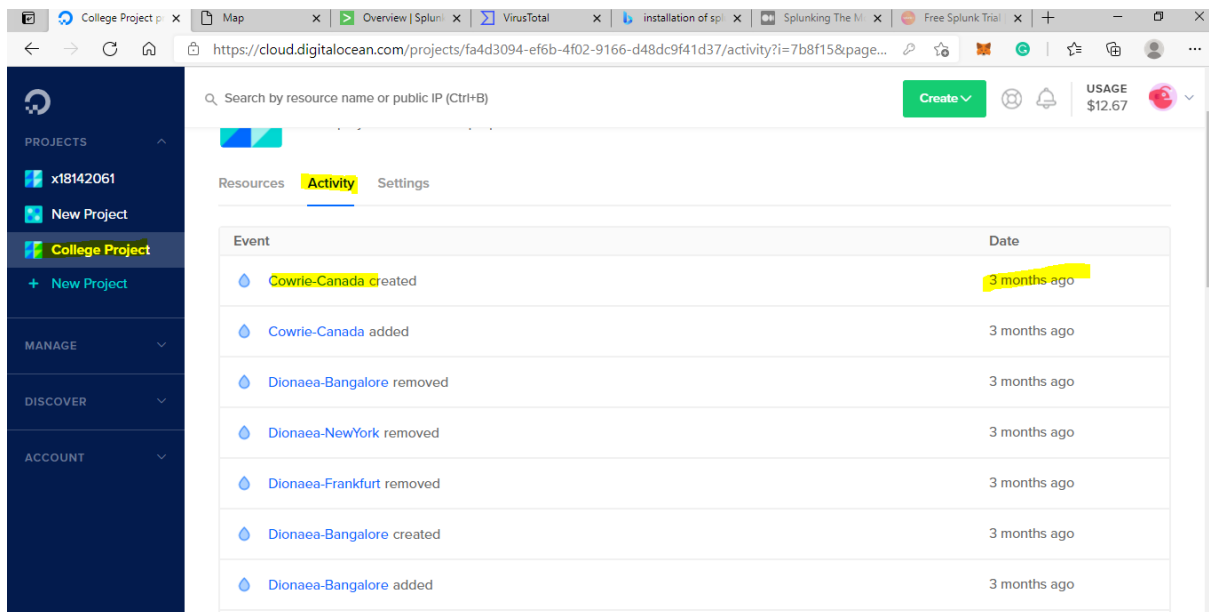


Figure 3.1.13

Shows us the activity within the project: the added and deleted sensors.

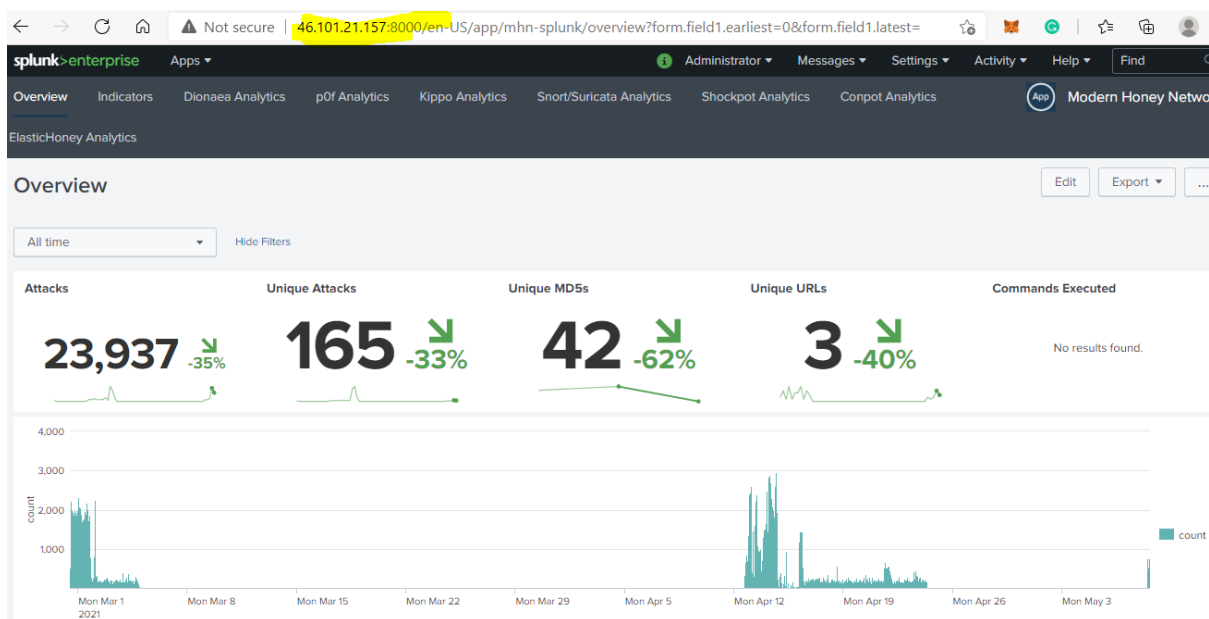


Figure 3.1.14

The overview of the Splunk page. The highlighted part shows the IP address associated with the MHN server.

4. MHN Analysis Using Splunk

The project was 75% done, but something was missing. I successfully learned how to deploy the MHN server, debug it and add honeypots. The outcome of the project I wanted was to get the geolocation of the attackers and most commonly attacked ports. I must gather the raw data and turn it into useful information to do the task. I used Splunk, which is a machine learning tool. I could perform the analysis on the MHN server, but it will be minimal as it

shows only the past 24 hours. In Splunk, I had the facility upload all the log files I stored in var/log/mhn to perform the analysis.

While researching, I came across an article, “[Splunking The Modern Honey Network: Getting the Value From Your Honeypots Data \(Part 1\)](#)” The article is a step by step guide to the integration of the MHN with Splunk. The advantage I got from this article was a premade MHN Splunk App that can be downloaded using this [link](#). The MHN Splunk App can be integrated into Splunk once it successfully downloaded.

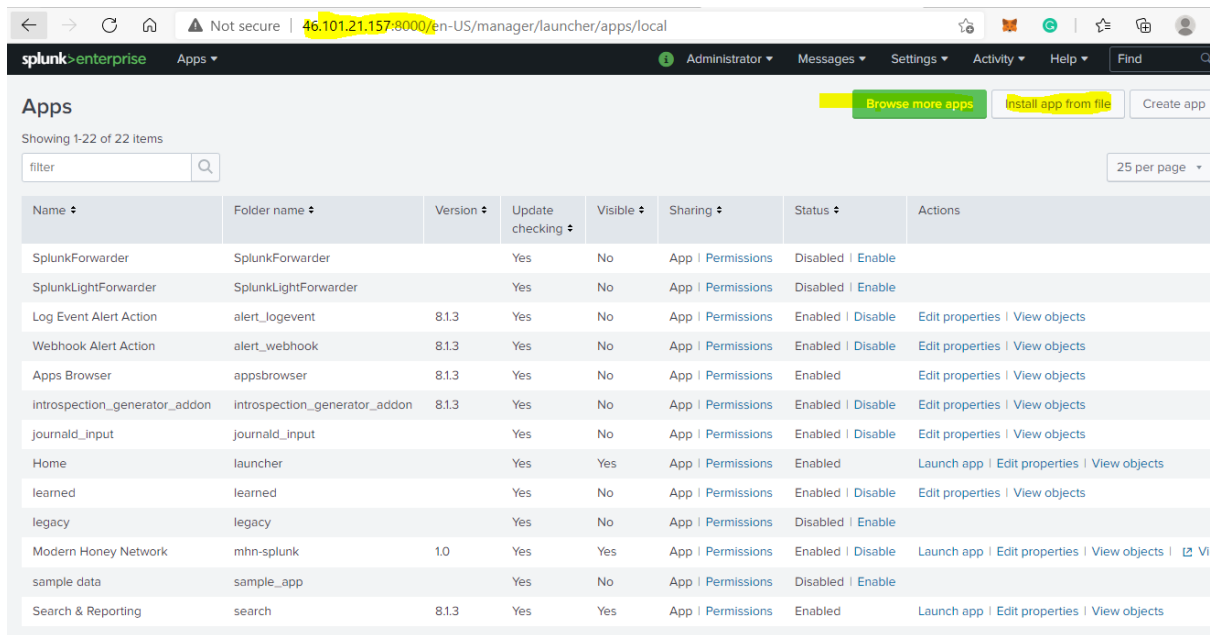


Figure 4.1

In figure 4.1 it shows that the MHN Splunk App was added successfully. To install the app, we need to go to the setting, and once we are on the setting, this page will be shown with the given option. We can browse and Install the app from the file because I downloaded the file I clicked on “Install the app from file”.

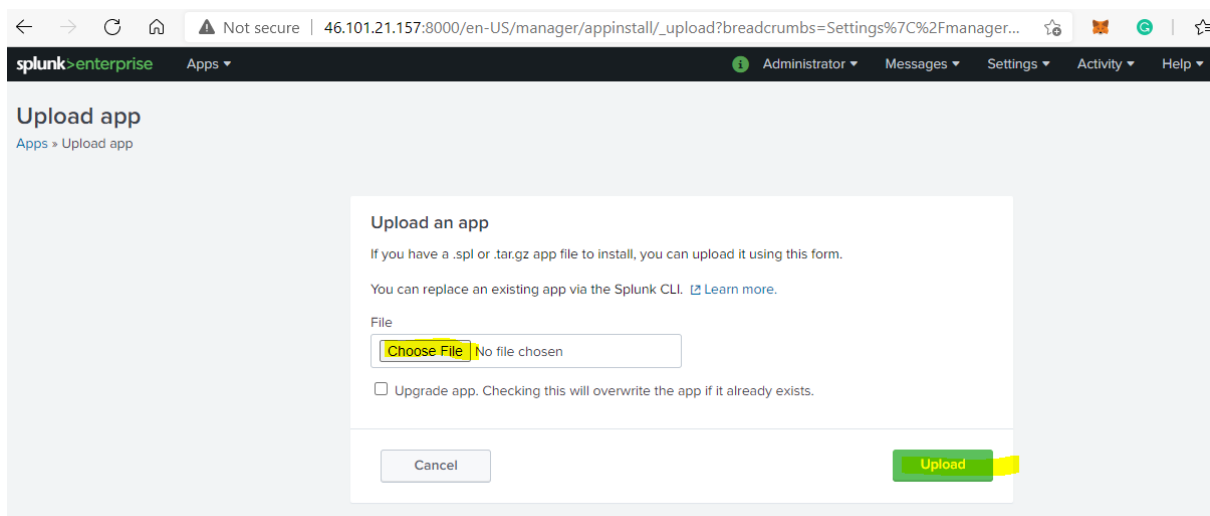


Figure 4.2

After clicking the install app from the file, we will be promoted to this page to choose a file location and upload.

The next step after installing the app is to point to the application form where we will fetch the data.

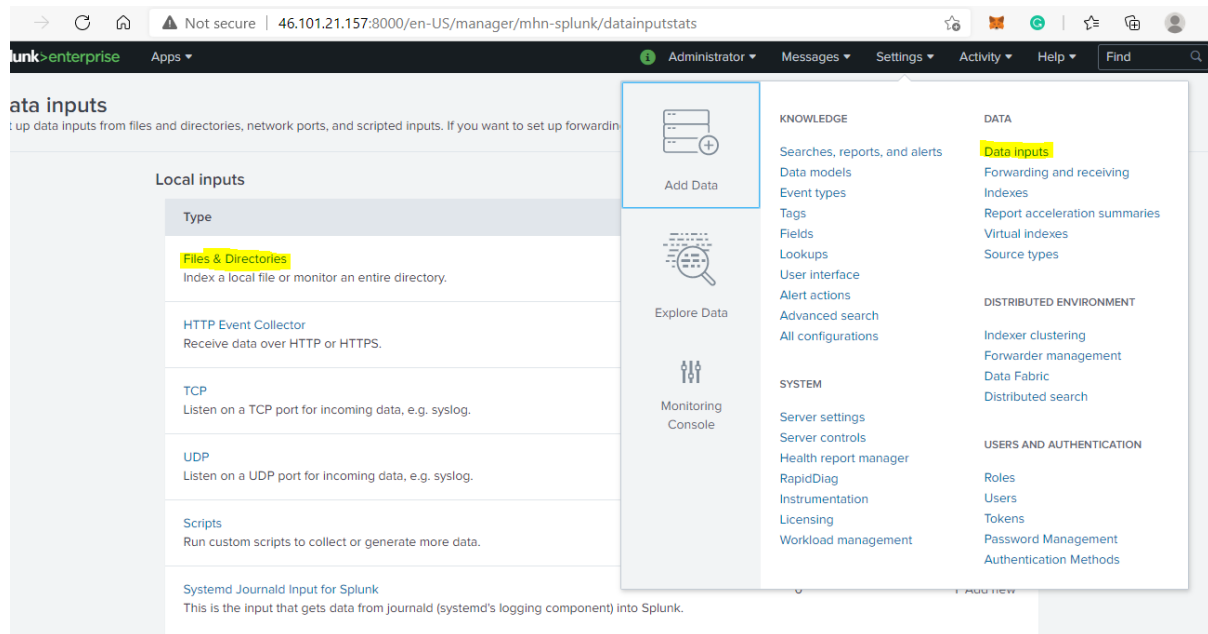


Figure 4.3

With the help of figure 4.3, I will illustrate how data can be fetched. On settings, we require to click on data inputs. I stored the data into var/log/mhn, so I had to point to the application where the data is located. To do so, I required to click on the Files & Directories.

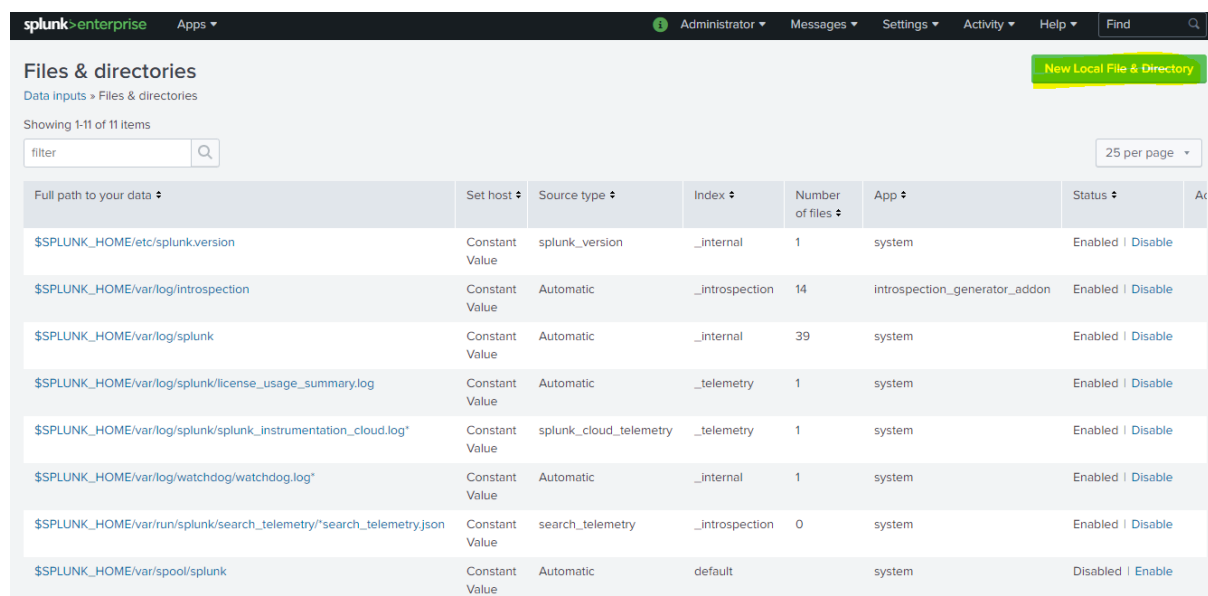


Figure 4.4

Figure 4.4 shows the existing files and directories in my Splunk App. We require to click on New Local File and Directory to add new.

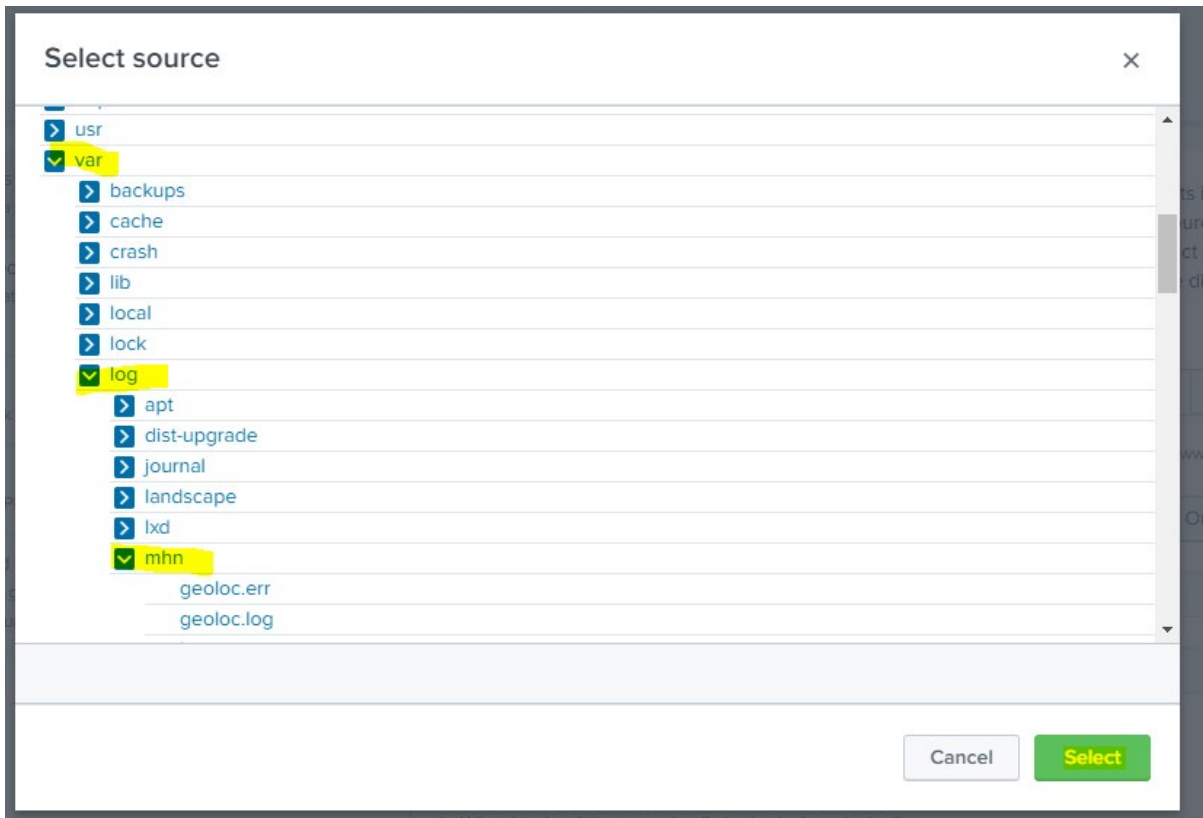


Figure 4.5

Once we are on the page, we can select the directory by clicking on it. For the project, everything is working as requires.

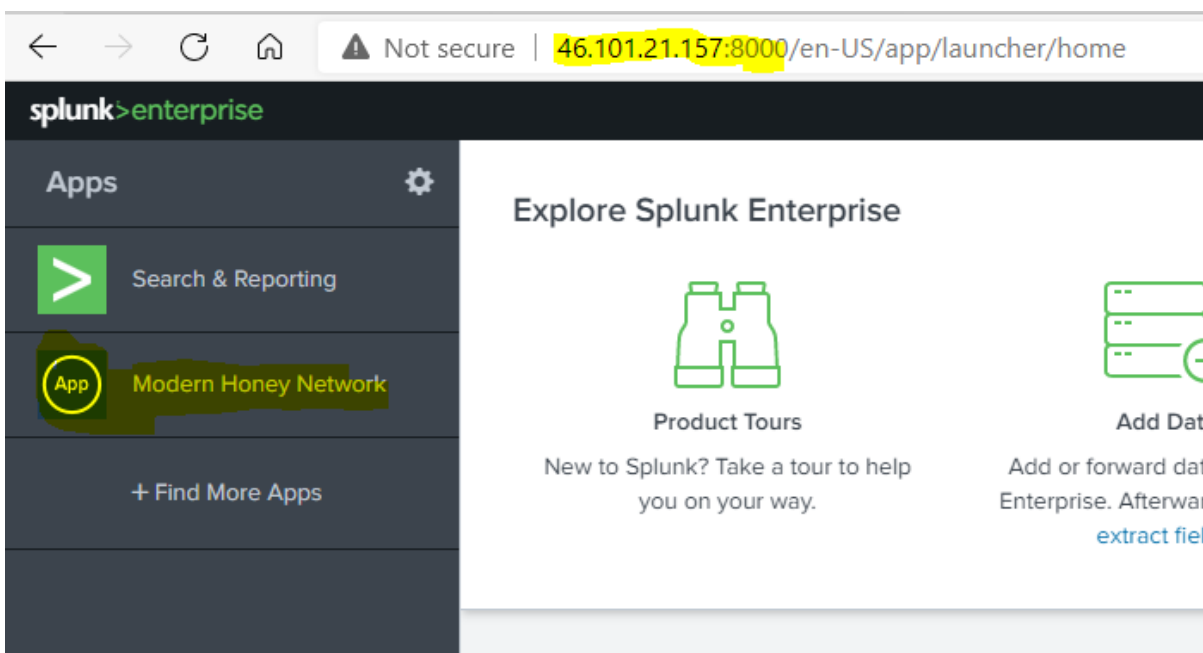


Figure 4.6

Figure 4.6 shows us the added app into the Splunk dashboard.

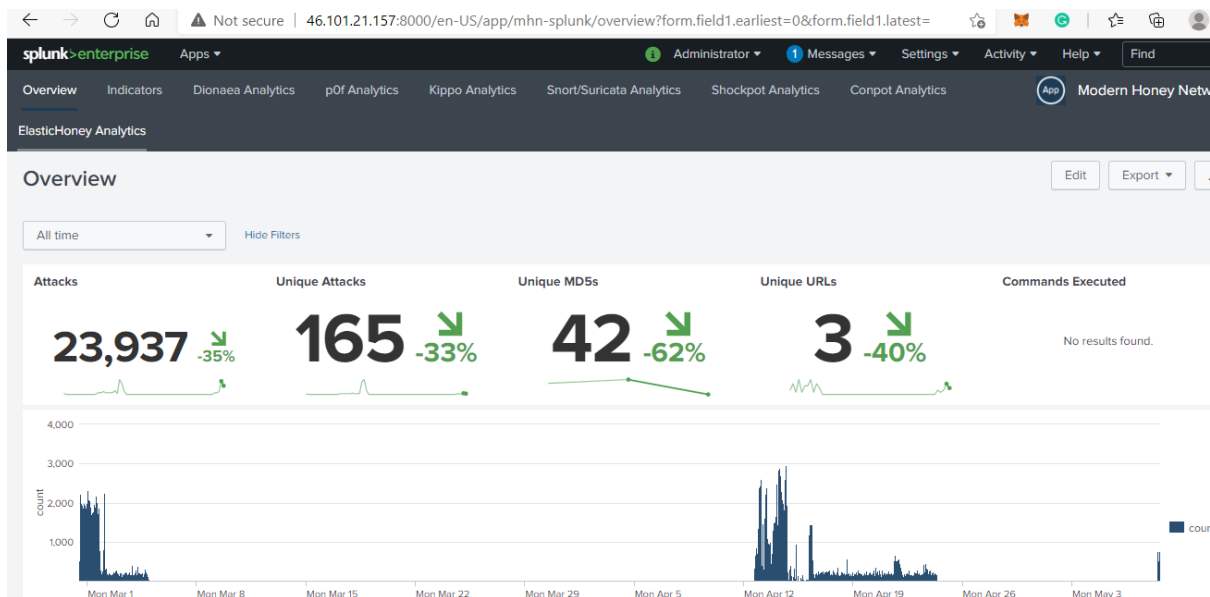


Figure 4.7

The overview of Splunk's dashboard also shows the time I was running my honeypots. As you can view, I ran few honeypots in March to gather the data and at last while writing the report. To show everything is functional.

The next step was analysing the Honeypots, as my project evolves around the Cowrie and Dionaea.

Cowrie:

Cowries are medium-level interaction honeypots. They operate on the SSH. Their primary purpose is to detect the attackers' attempts to connect to the SSH server through brute-forcing attacks. In my case, the default password was 'password'.

I deployed many cowries during the project, and even in penetration class, the deployed cowries IP address was given to classmates to perform attacks. So, I could gather the data.

Before analysing the whole gather data, I decided to run few attacks from my VM and through two cloud-based machines.



Figure 4.1.0

To perform the attacks, I deployed a Cowrie in Germany datacentre. The Ip address I got was 138.68.105.180.

Once I have the Ip address, I logged in through my Kali Linux and other virtual machines to perform some attacks.

```

(umar@kali)-[~]
└─$ ssh 138.68.105.180
The authenticity of host '138.68.105.180 (138.68.105.180)' can't be established.
RSA key fingerprint is SHA256:p9JZeNdQlPdVj7CdY3XNwXcTkWfFL4Lxbb9dzmLKmHg.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added '138.68.105.180' (RSA) to the list of known hosts.
umar@138.68.105.180's password:
Permission denied, please try again.
umar@138.68.105.180's password:
Permission denied, please try again.
umar@138.68.105.180's password:
umar@138.68.105.180: Permission denied (password,publickey).

(umar@kali)-[~]
└─$ Home

(umar@kali)-[~]
└─$ ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.226.128 netmask 255.255.255.0 broadcast 192.168.226.255
    inet6 fe80::20c:29ff:fe54:98a3 prefixlen 64 scopeid 0<link>
    ether 00:0c:29:54:98:a3 txqueuelen 1000 (Ethernet)
    RX packets 2401 bytes 209895 (204.9 KiB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 2161 bytes 164082 (160.2 KiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
    inet 127.0.0.1 netmask 255.0.0.0
    inet6 ::1 prefixlen 128 scopeid 0<host>
    loop txqueuelen 1000 (Local Loopback)
    RX packets 917 bytes 317543 (310.1 KiB)
    RX errors 0 dropped 0 overruns 0 frame 0

```

Figure 4.1.1

Figure 4.1.1 shows that the Ip address I am attacking from is 192.168.226.128. To confirm if my attacks being logged, I went back to the dashboard of MHN.

The screenshot shows the MHN Server dashboard with the following search filters: Sensor: Germany-Cowrie, Honeypot: All, Date: MM-DD-YYYY, Port: 445, IP Address: 8.8.8.8. The Attacks Report table contains the following data:

Date	Sensor	Country	Src IP	Dst port	Protocol	Honeypot
2021-05-09 09:40:01	Germany-Cowrie		109.79.171.28	22	ssh	cowrie
2021-05-09 09:39:41	Germany-Cowrie		109.79.171.28	22	ssh	cowrie
2021-05-09 09:37:11	Germany-Cowrie		109.79.171.28	22	ssh	cowrie
2021-05-09 09:32:49	Germany-Cowrie		194.165.16.89	22	ssh	cowrie

Figure 4.1.2

Figure 4.1.2 shows the attacks were logged, the Ip address and the country where the attacks were deployed.

So, I decided to go one step future, in which I will deploy a VM on the cloud platform and run and attack from there.

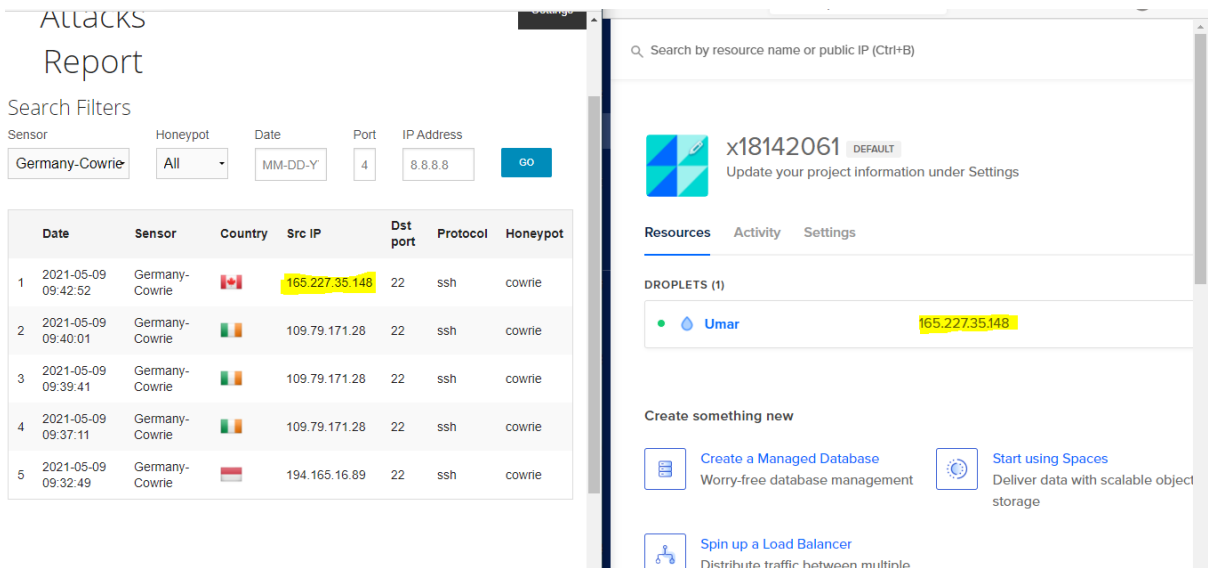


Figure 4.1.3

Figure 4.1.3 shows us the attacker IP and the country. In this case, it was the VM I deployed in Canada. After few successful attempts, I decided to deploy another VM in the USA and run few more attacks.

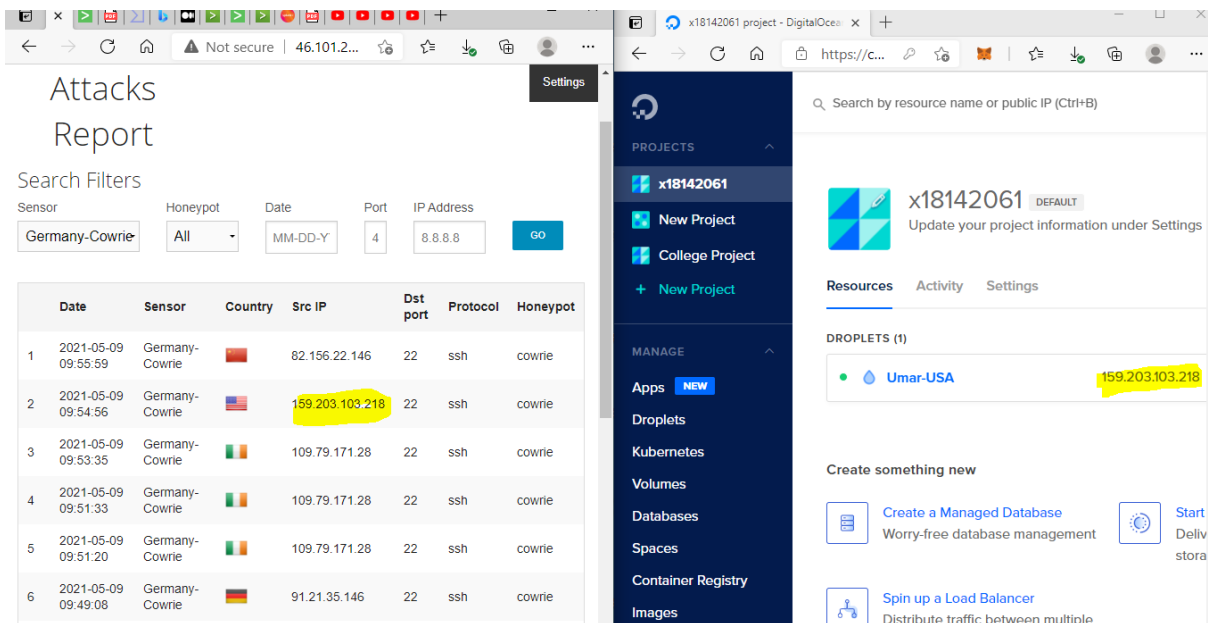


Figure 4.1.4

Figure 4.1.4 shows us the attack I ran from a virtual machine-based in the USA. I successfully logged into the cowrie and ran few unfamiliar commands, which was my name and student Id shown in Figure 4.1.5.

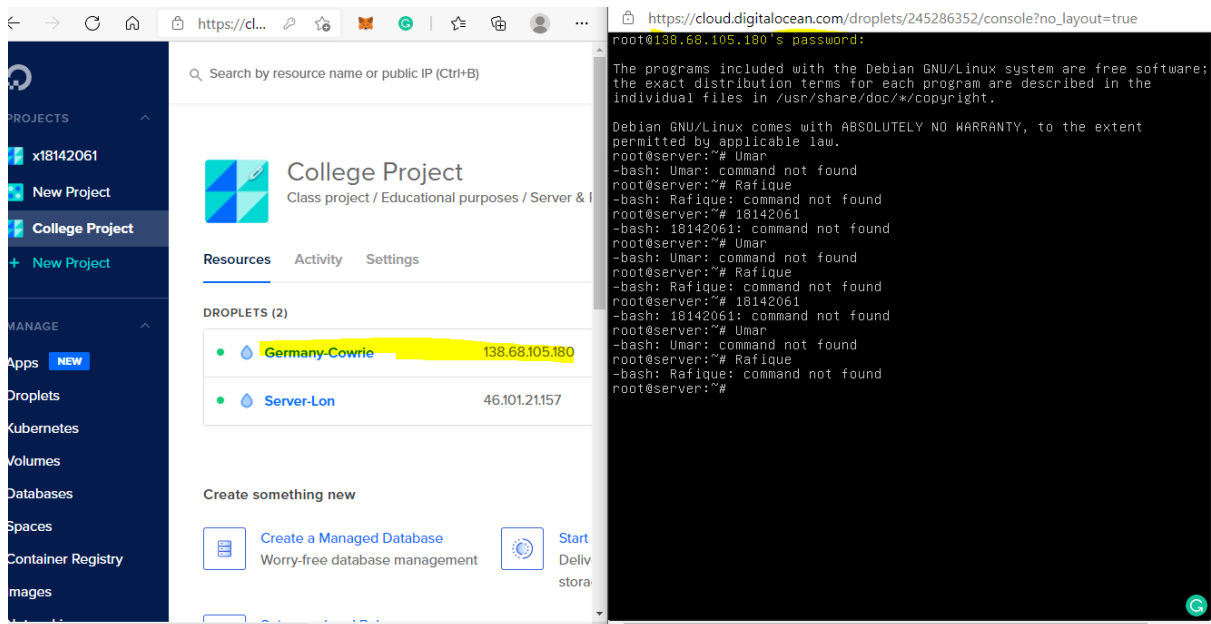


Figure 4.1.5

The next step will perform analysis on Splunk if it has been logged or not.

Top Username/Password Combinations

ssh_username	ssh_password	count
umar	password	6
umar	12345	6
umar	Password	4
umar	123456	2
umar	[APass1234	2
zhangnt	zhangnt	1
umar	root	1
umar	password1	1
umar	pass1234	1
umar	Pass1234	1

Figure 4.1.6

Figure 4.1.6 shows the username and password of the attacker used to get the access. The username is shown as umar because I created virtual machines under my first name.

It illustrates the recent attack and logs, but now, I wanted to investigate what password has been used frequently to access the honeypot.

Top Passwords	
ssh_password ↕	count ↕
123456	646
nproc	629
123	211
password	158
admin	118
1234	117
1	109
test	106
12345	91
user	90

Figure 4.1.7

Figure 4.1.7 shows the Top password. On top of the list was 123456, which was attempted 646 times, then nproc with the second-highest attempts of 629 so on and so forth.

Top Username/Password Combinations		
ssh_username ↕	ssh_password ↕	count ↕
nproc	nproc	629
user	user	78
admin	admin	76
test	test	64
pi	raspberry	61
postgres	postgres	57
pi	raspberryraspberry993311	53
oracle	oracle	53
user	1	49
ubuntu	ubuntu	43

Figure 4.1.8

Then I wanted to investigate the combination with includes the username and password. In this case, the nproc username and nproc password were used 629, the second-highest username and password used were user, user, and so forth.

At the start of the project, my investigation was limited, but I realised a lot could be done when I went through the project. As mentioned in my proposal and technical report, I wanted to get geolocation, so the following chart gives us the attackers countries.

Top Attacker Countries

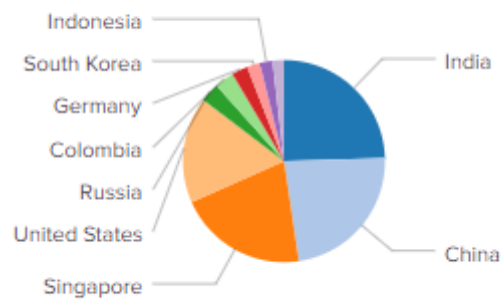


Figure 4.1.9

After receiving the Countries, I wanted to investigate to top cities of the attacker.

Top Attacker Cities

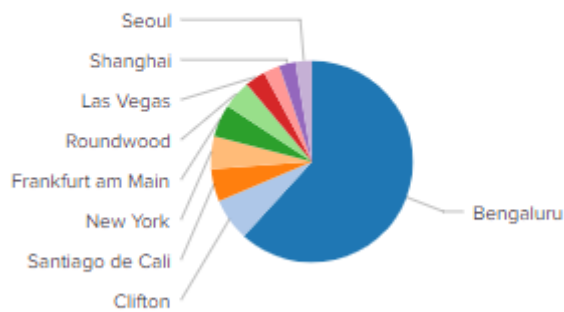


Figure 4.1.10

Figure 4.1.10 shows us the top cities of the attackers.

Top Attackers		
Source ↕	Sparkline ↕	count
167.71.228.234		472
162.62.132.94		382
109.79.171.28		35
159.203.103.218		32
61.155.2.142		24
118.89.237.20		24
1.15.100.43		20
13.65.16.18		20
14.63.213.72		20
43.128.3.101		20

Figure 4.1.11

Figure 4.1.11 show us how Splunk was able to determine the raw data and converted it into a helpful graph. We were able to retrieve the top counties and cities of the attackers based on the above data. For manual computation, it would have taken month or years to come up with figures with the probability of fifty per cent of wrong information because the attackers were attacking in real-time.

Dionaea:

Dionaea was the second and the most crucial honeypot. It comes under the low interaction category of honeypots. It matches vulnerabilities inside the network servers. Dionaea supports HTTP, SIP, SMB, FTP protocol. Dionaea runs on a restricted mode with no administrative privileges to keep the impact of attacks at a minimum. If an attacker exploits any of the port mentioned above, it creates an alert while making a copy. Dionaea gathers malware attacks, which are nowadays on the peak. When dionaea is attacked, it logs the payload into URLDownloadToFile, Command Exec, Multi-Stage Payloads, and shells, Once the payload from the attacker has been received. Dionaea analysis the payload using 'Libemu', then the extracted information is sent to and stored in the SQLite database.

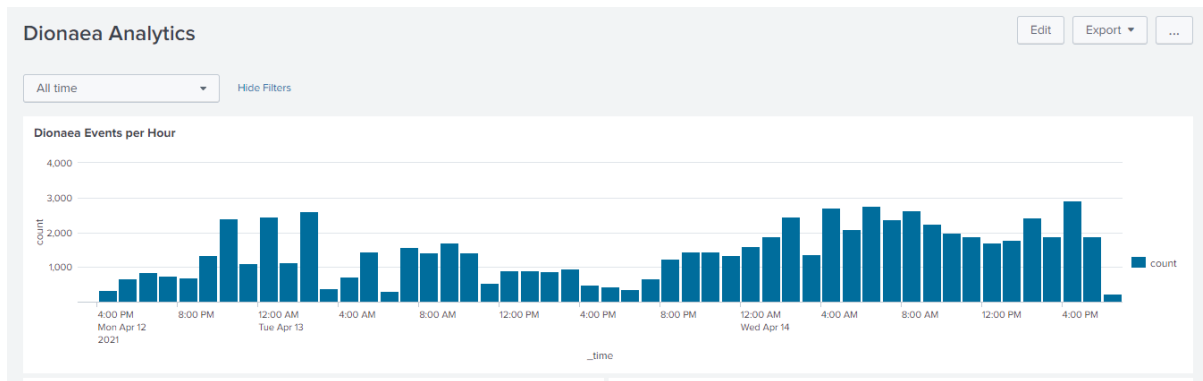


Figure 4.2.1

Figure 4.2.1 shows us the attacks I was receiving on the Dionaea honeypot. The minimum per day I have received was 204, and the maximum goes to 2772. It shows the trend of malware attacks.

Top Ports	
dest_port ↕	count ↕
445	28459
80	8056
1433	2260
23	2062
8080	224
65533	178
6379	165
25	134
443	123
3389	119

Figure 4.2.2

Figure 4.2.2 shows us the top ports being hit, SMB was on top of the list with the count of 28459, and then we have HTTP with 8056, so on and so forth.

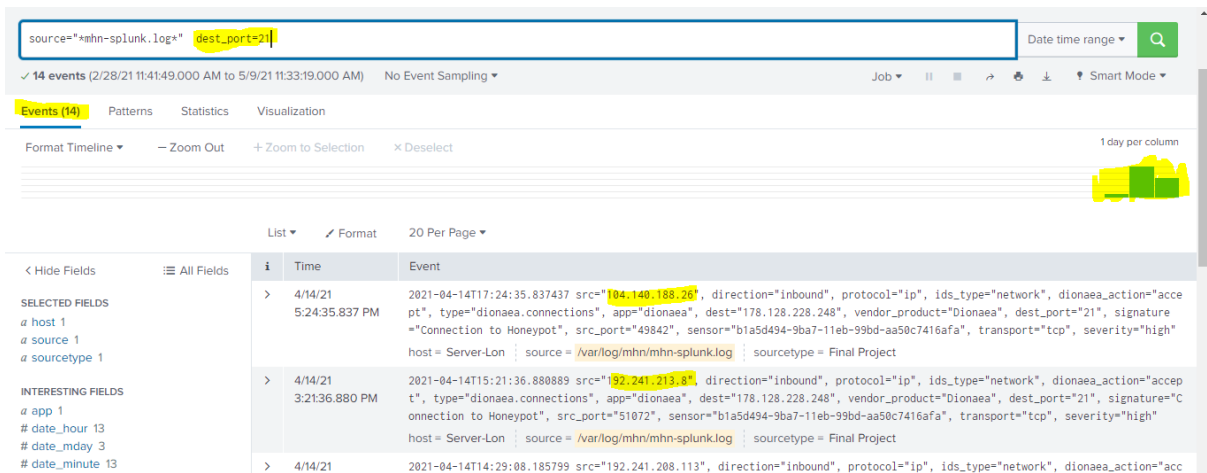


Figure 4.2.3

Figure 4.2.3 shows that the port 21 FTP was exploited 14 times.

Top Dionaea Attackers

src	Country	count	DShield
45.125.65.74	Hong Kong	24034	DShield
120.68.252.73	China	7167	DShield
117.4.43.220	Vietnam	3320	DShield
36.77.78.7	Indonesia	2847	DShield
201.249.191.114	Venezuela	2785	DShield
109.224.30.178	Iraq	1862	DShield
218.56.161.69	China	1498	DShield
81.28.175.43	Russia	1159	DShield
42.112.229.185	Vietnam	1090	DShield
14.171.138.86	Vietnam	771	DShield

Figure 4.2.4

Figure 4.2.4 shows us the top Dionaea attackers IP address with the country's name and total count of how many attacks have been launched on the honeypot.

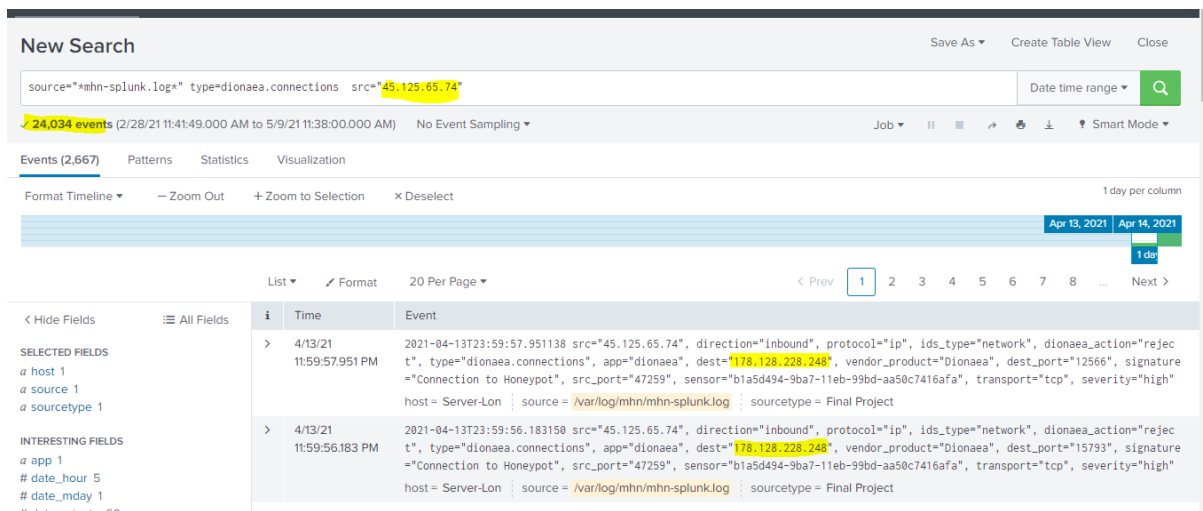


Figure 4.2.5

Figure 4.2.5 shows us the top attacker and the IP address of the honeypot he/she was trying to exploit. Only within two days, April 13 and April 14, 2021, a total number of 24034 attacks were logged from the IP address of 45.125.65.74.

Top MD5s Captured

md5	count	TotalHash
ae12bb54af31227017feffd9598a6f5e	105	TotalHash
996c2b2ca30180129c69352a3a3515e4	54	TotalHash
0ab2aeda90221832167e5127332dd702	46	TotalHash
414a3594e4a822cfb97a4326e185f620	23	TotalHash
95ae8e32eb8635e7eabe14ffbf77b	14	TotalHash
cd99e5e4f44621978faf8df0e01d2d2b	13	TotalHash
a55b9addb2447db1882a3ae995a70151	7	TotalHash
fc6b0f95853dfda72d5535a424b3a29	5	TotalHash
6e72ad805b4322612b9c9c7673a45635	5	TotalHash
59b5090fad3d62f05572470f0c79c9a4	5	TotalHash

Figure 4.2.6

Figure 4.2.6 shows us the hashes of files. To run a further analysis and check what is inside the files, I copied the hash to virustotal.

Analyze suspicious files and URLs to detect types of malware, automatically share them with the security community

FILE
URL
SEARCH

ae12bb54af31227017feffd9598a6f5e|

By submitting data above, you are agreeing to our [Terms of Service](#) and [Privacy Policy](#), and to the sharing of your Sample submission with the security community. Please do not submit any personal information; VirusTotal is not responsible for the contents of your submission. [Learn more.](#)

Want to automate submissions? [Check our API](#), free quota grants available for new file uploads

Figure 4.2.7

59 / 66

59 security vendors flagged this file as malicious

c05e2dab77349cd639aa837e7e121710b8a0718d8fc93fb4cc6458ae90e5c597
 VirusShare_ae12bb54af31227017feffd9598a6f5e

5.02 MB Size | 2021-05-09 06:35:18 UTC 5 hours ago

armadillo cve-2017-0147 exploit overlay pedll via-tor

DETECTION	DETAILS	RELATIONS	BEHAVIOR	COMMUNITY
Acronis	Suspicious		Ad-Aware	Trojan.Agent.CZTF
AegisLab	Trojan.Win32.Wanna.tpxd		AhnLab-V3	Trojan/Win32.WannaCryptor.R200894
Alibaba	Ransom:Win32/CVE-2017-0147.96af7fc8		ALYac	Trojan.Ransom.WannaCryptor
Antiy-AVL	Trojan/Generic.ASMalwS.2041E4F		SecureAge APEX	Malicious
Arcabit	Trojan.Agent.CZTF		Avast	Sf:WNCryLdr-A [Trj]
AVG	Sf:WNCryLdr-A [Trj]		Avira (no cloud)	TR/AD.WannaCry.zbqny
Baidu	Win32.Worm.Rbot.a		BitDefender	Trojan.Agent.CZTF
BitDefenderThreat	Gen:NN.Zadla.F.34688.8vE@eCOW7Zal		Blkai Pro	W32.DikasaDGE.Trojan

Figure 4.2.8

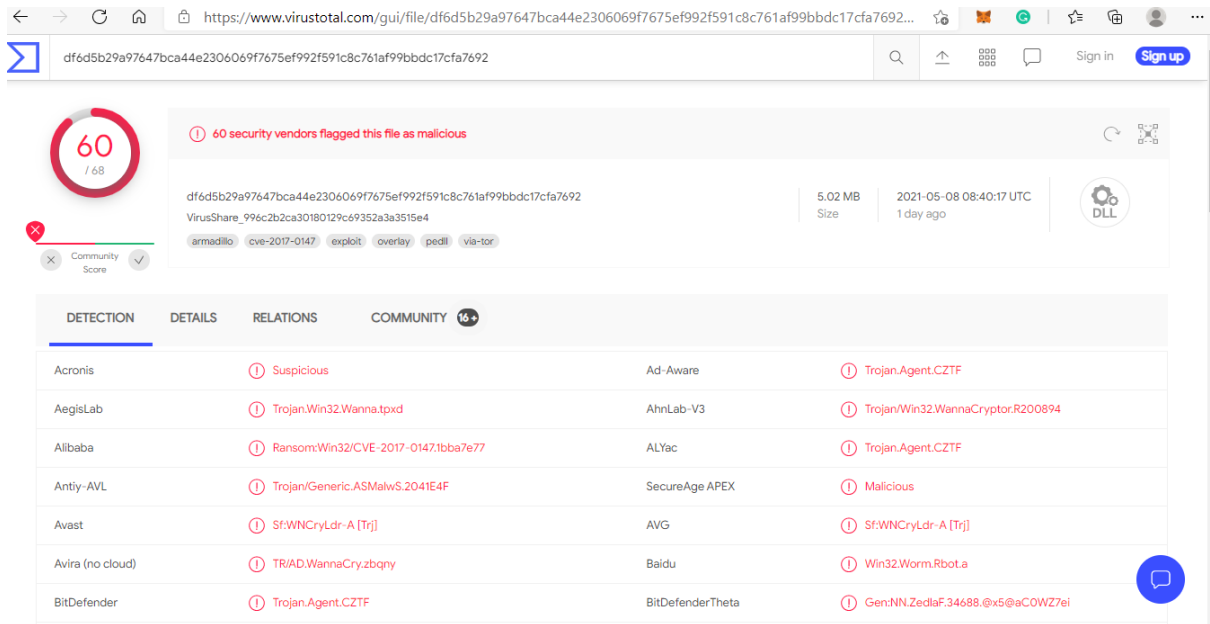


Figure 4.2.9

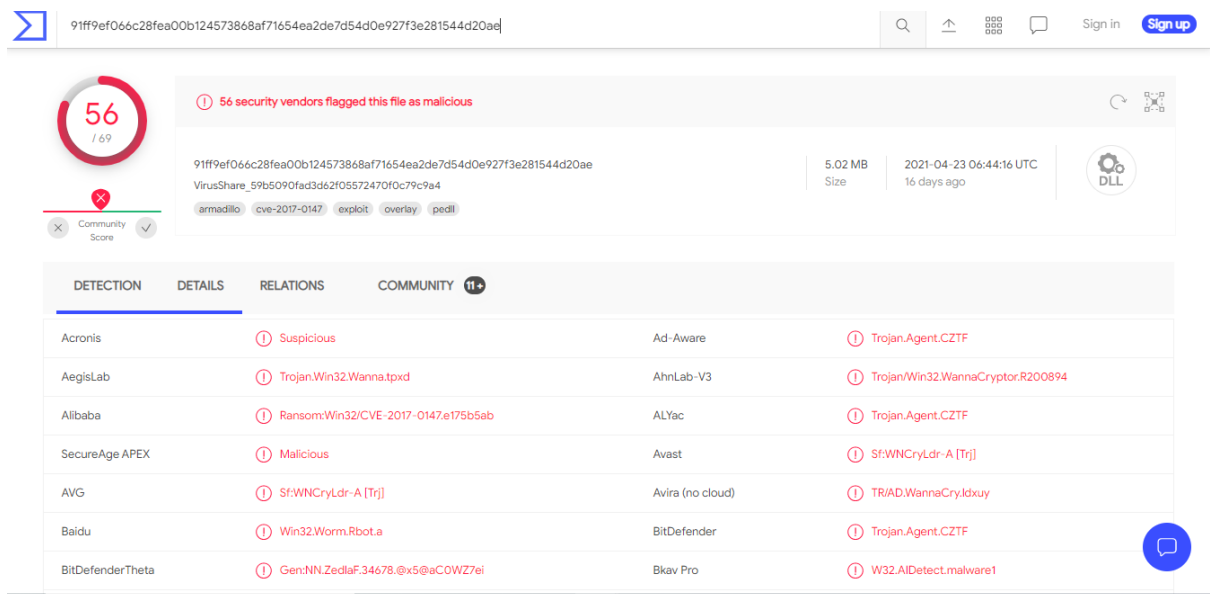


Figure 4.2.10

The above figures show I copied the top two and the last hash of the file. After investigating on VirusTotal, it gives us the conclusion that the files were associate with malware.

Combined Analysis:

The global map of attackers.

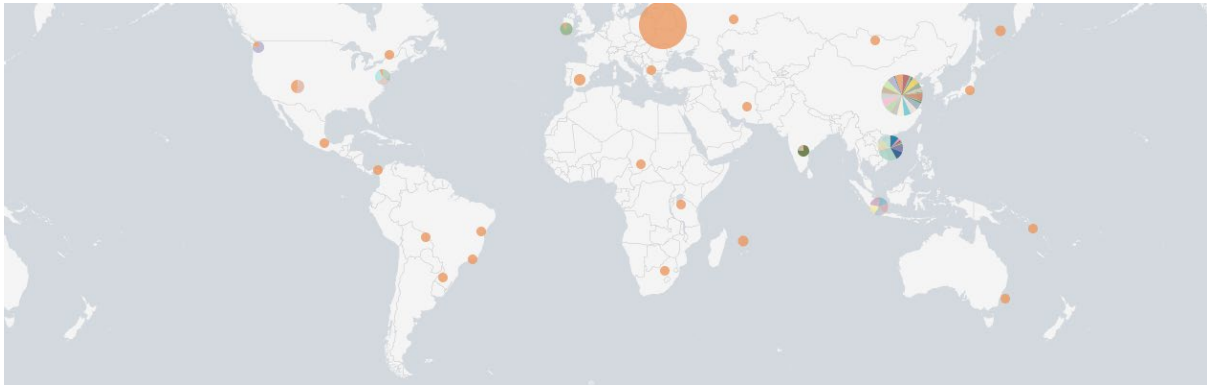


Figure 4.3.1

Top countries and cities of attackers.

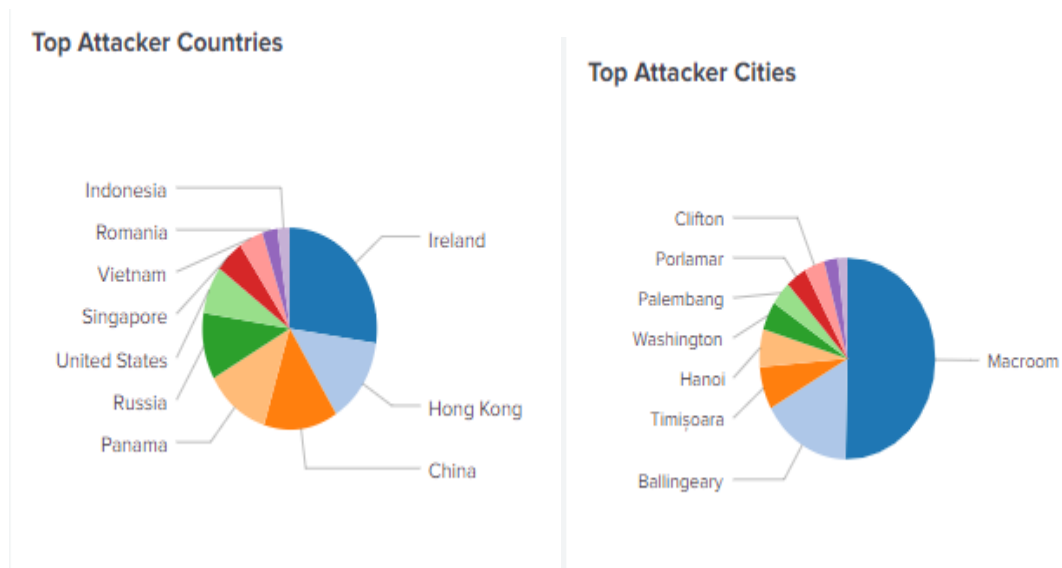


Figure 4.3.2

Top ports being attacked.

Top Ports

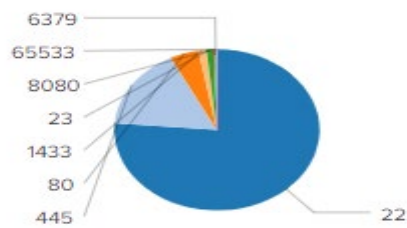


Figure 4.3.3

Top attackers and the honeypots that attracted the attackers.

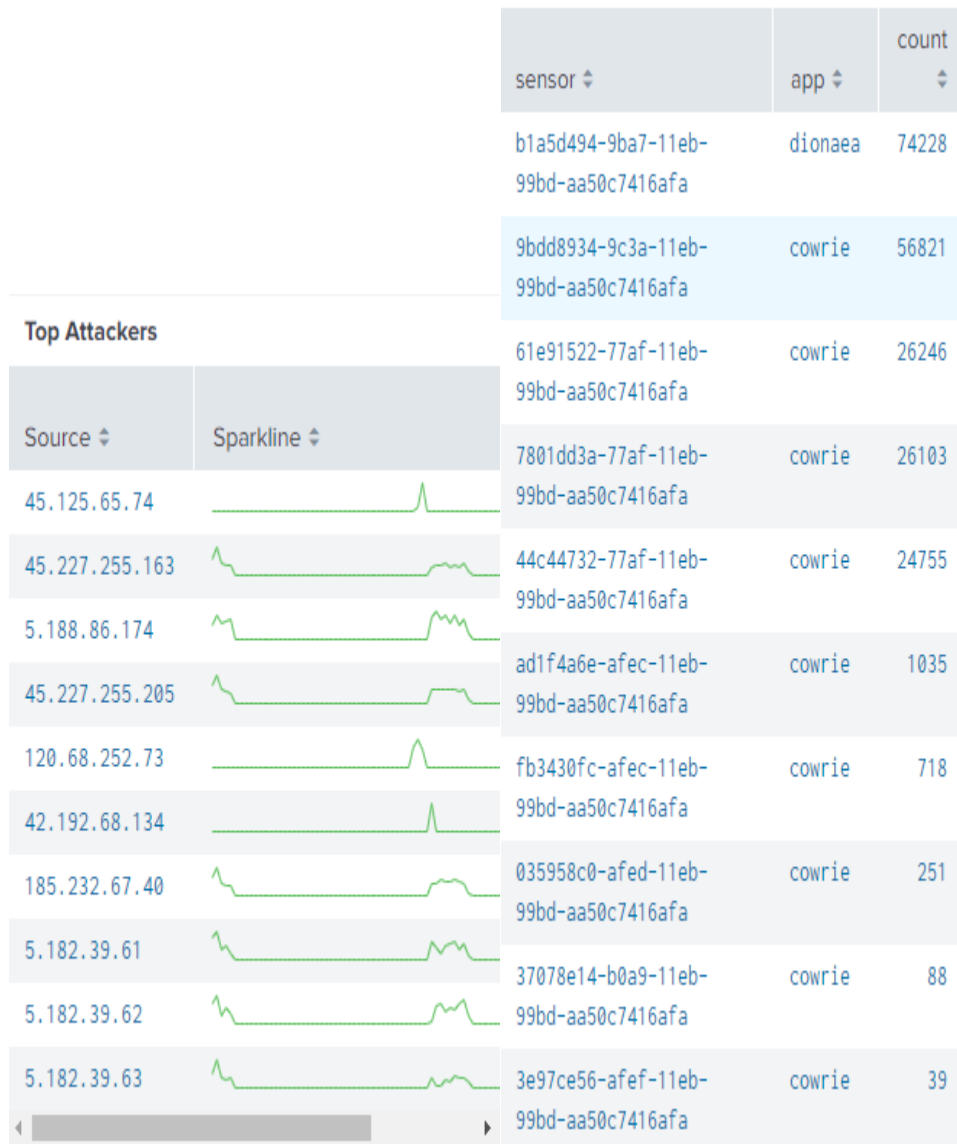


Figure 4.3.4

Figure 4.3.5 shows total Dionaea attacks running on different datacentres. The graph below shows the New York datacentre was the hit most and the Frankfurt was the second highest.

Dionaea	Attacks
Dionaea-Frankfurt	207705
Dionaea-NewYork	3450161
Dionaea-Bangalore	37173
Dionaea-Canada	73738

Figure 4.3.5

Data visualisation of numbers.

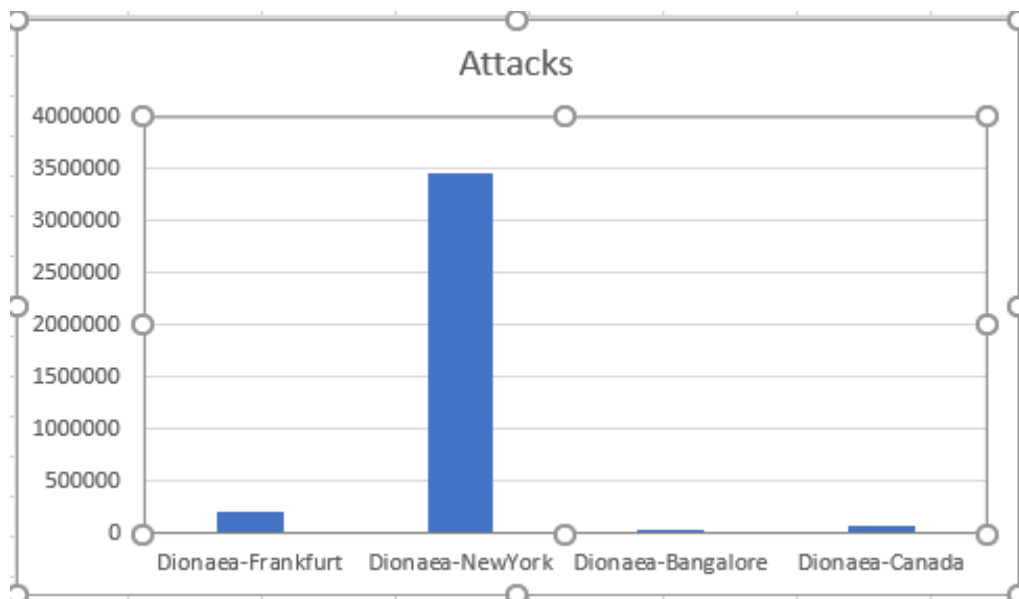


Figure 4.3.6

In the below graph for Cowrie, Canada datacentre was the one hit mostly, Singapore took the second position than Germany and was the USA.

Cowrie	Attacks
Cowrie-Germany	70421
Cowrie-USA	28517
Cowrie-Singapore	131664
Cowrie-Canada	232056

Figure 4.3.7

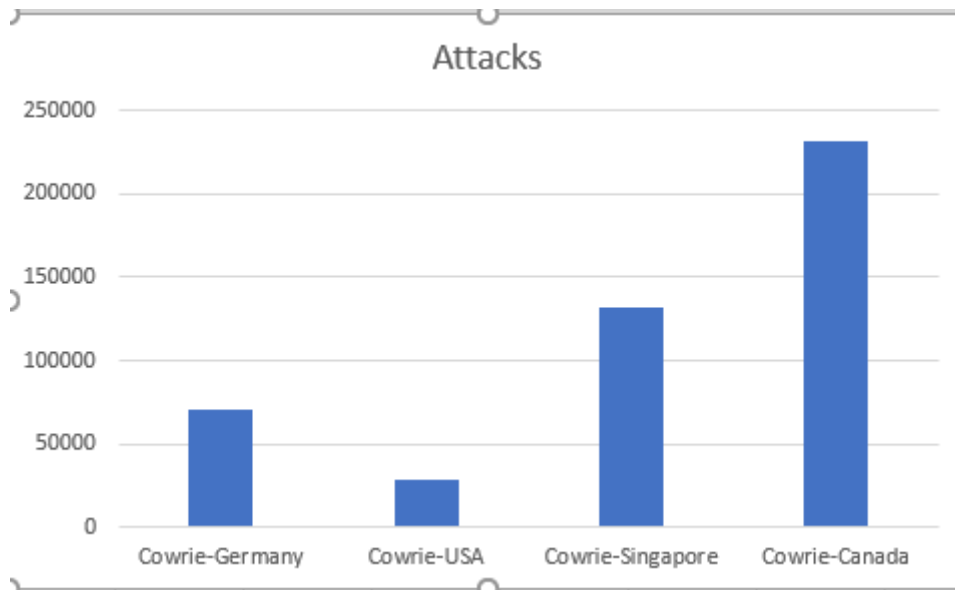


Figure 4.3.8

5. Testing

The testing plays a vital role in this project. I used waterfall methodology while installing each of the honeypots. This means I tested every single server and droplet (Honeypot) when installed. If the server and honeypots are not deployed properly, I will not have achieved the results of this project.

Following screenshots are the demonstration of testing that when the celery worker was not installed properly, I could not log the attacks after the first attack.

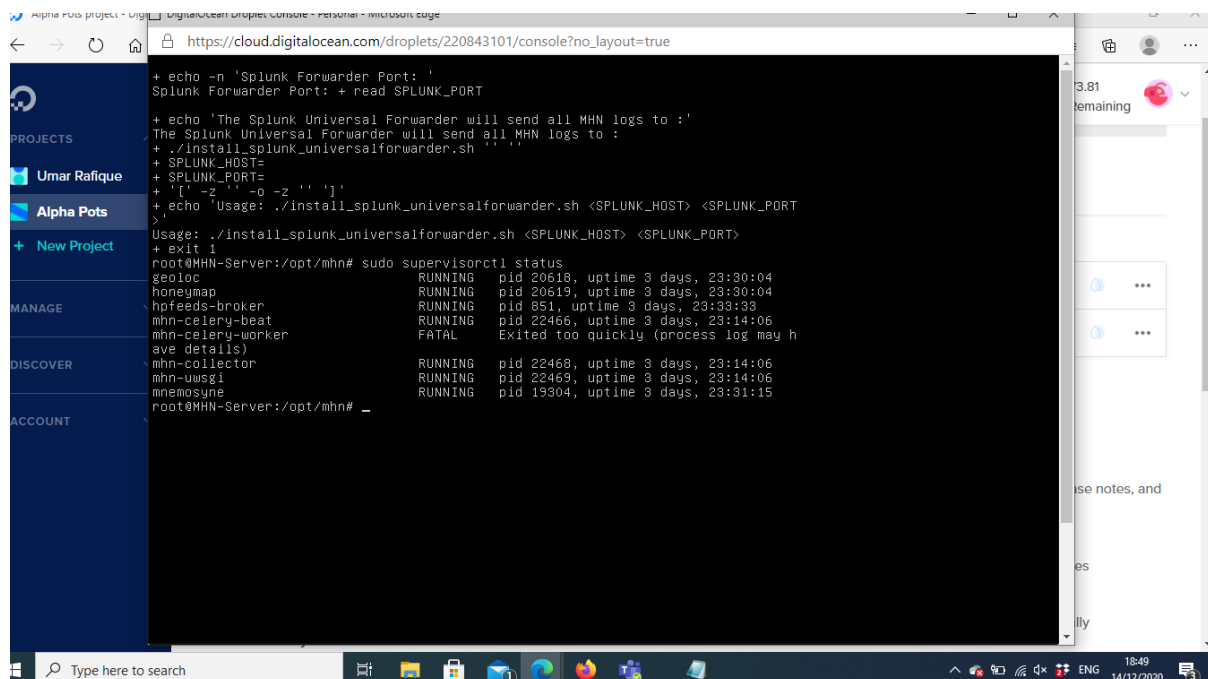


Figure 5.1

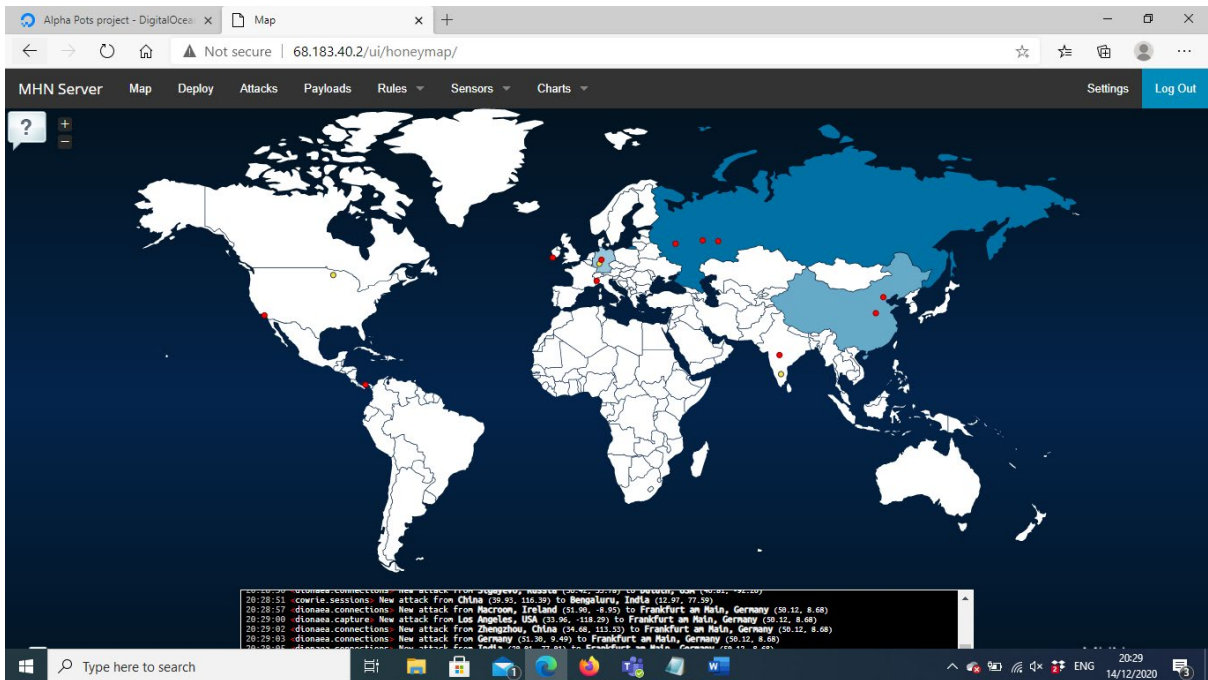


Figure 5.2

Figure 5.1 show us that mhn-celery-worker was Fatal. I received a couple of attacks worldwide, but the server logged only one of the attacks, shown in Figure 5.3.

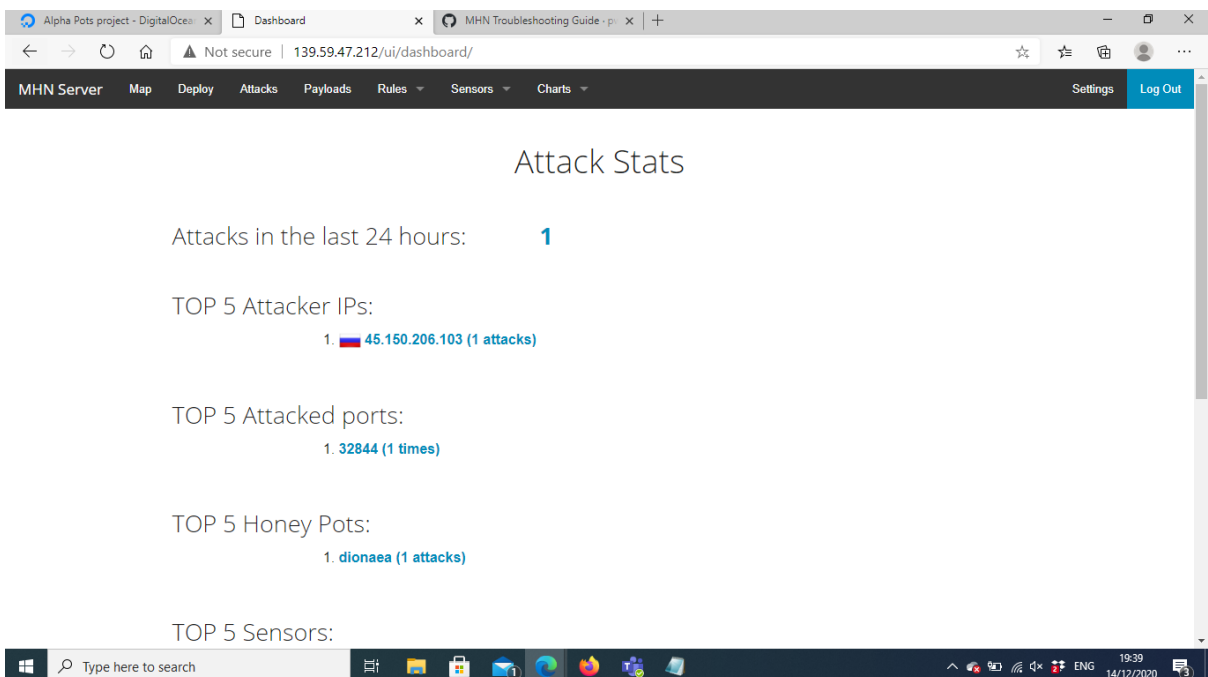


Figure 5.3

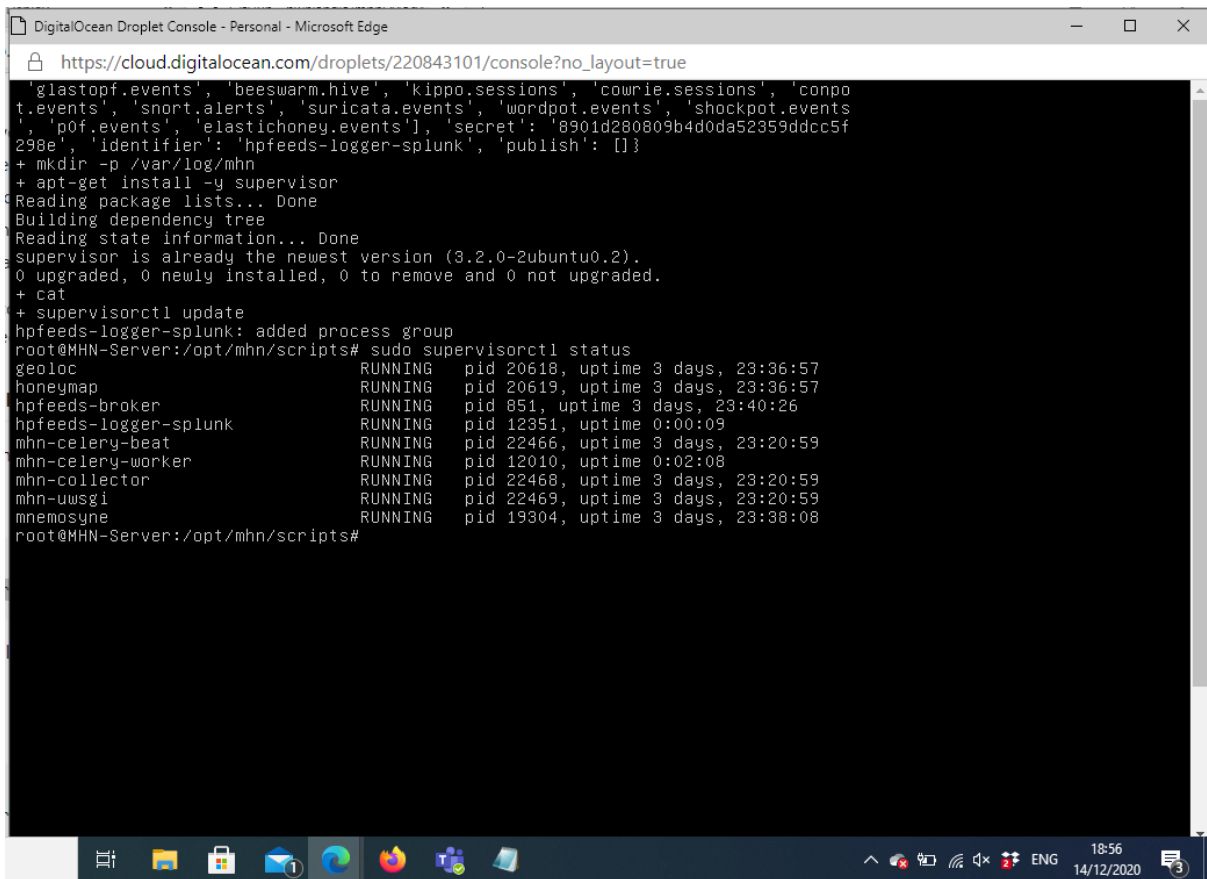


Figure 5.4

After researching, I restored the celery-worker via few Linux commands. Figure 5.4 show us that the mhn-celery-worker is generally functioning because the issue was resolved in figure 5.5. The screenshot shows the other attackers are also getting logged into the server.

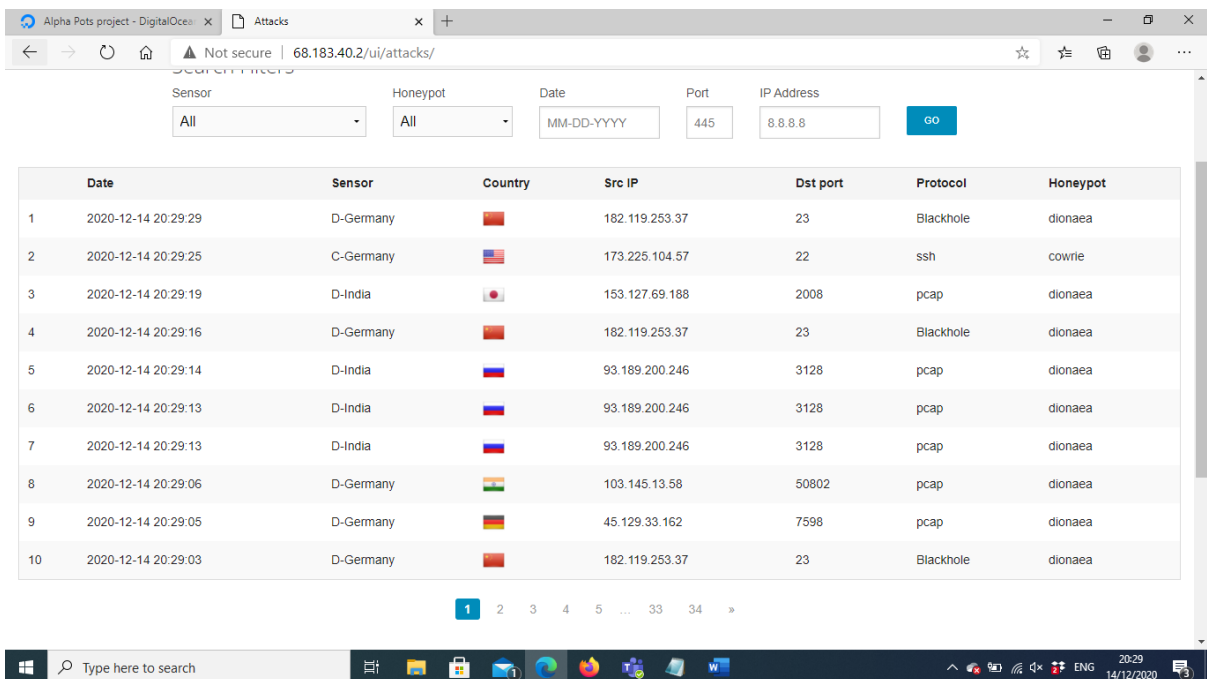


Figure 5.5

As you can see, the issue was resolved in figure 5.5. I also tested for the MHN server and Splunk if they worked properly. The results could be confusing if the MHN server does not log the attacks coming from the same IP address, location, and destination port.

The testing for the MHN server by running few attacks were done under the heading of MHN Analysis using Splunk can be seen from figure 4.1.0 to 4.1.7.

I also ran Nmap commands to see the open ports.

6. Conclusions

This project has enhanced my knowledge in the cybersecurity field, especially how the attackers launched the attacks on different systems. Where to find the system's vulnerabilities using different tools and operating system.

Overall, it was a successful project for me. With much learning, including different tools and operating systems, I successfully achieved my project aim, which was mentioned under the sub-heading of the introduction. I successfully deployed the honeypots, gathered the data and analysed the data.

I was also able to check the hashes using total virus if the files include any malware, which has broadened my skills, especially when dealing with unfamiliar email addresses.

The only limitation of the project it works virtually, which could cause us few issues from the cloud provider side. In that scenario, we could lose all the gathered information.

7. Further Development or Research

With additional time and resources, this project could lead us to a company where we provide organisation knowledge about cybersecurity and prevent it. I would require few members who are expert in Cyber Security and tools.

The other alternative is to work as a freelance Penetration tester by creating a phishing email or checking which unsecured ports are open. The attackers could deploy the payload. The most accessible demonstration would be creating a simple malware and deploying it into the organisation. The password to unlock the malware will be given to the organisation before deploying the attack, not to affect the organisation's work.

Also, if we research correctly in this project, we might find many new ways the hackers/attackers are using to exploit the system and learn how to defend the new techniques the attackers/hackers are using.

8. References

Stock, J., 2020. *INTERPOL report shows alarming rate of cyberattacks during COVID-19*. [Online] Available at: <https://www.interpol.int/en/News-and-Events/News/2020/INTERPOL-report-shows-alarming-rate-of-cyberattacks-during-COVID-19> [Accessed April 17 2021].

Arora, M., 2021. *What Are Honeypots? Various Types Of HoneyPots*. [Online]
Available at: <https://catchupdates.com/honeypots/>
[Accessed March 5 2021].

David Millward, 2021. *Biden to step up cybersecurity after hackers hit vital oil pipeline*. [Online]
Available at: <https://www.independent.ie/world-news/north-america/biden-to-step-up-cybersecurity-after-hackers-hit-vital-oil-pipeline-40406446.html>
[Accessed May 14 2021].

Fortinet, 2020. *What Are Honeypots (Computing)?*. [Online]
Available at: <https://www.fortinet.com/resources/cyberglossary/what-is-honeypot>
[Accessed September 19 2020].

Gatlan, S., 2021. *Ransomware hits TU Dublin and National College of Ireland*. [Online]
Available at: <https://www.bleepingcomputer.com/news/security/ransomware-hits-tu-dublin-and-national-college-of-ireland/>
[Accessed May 13 2021].

Gráinne Ní Aodha, 2021. *HSE confirms ransom has been sought over cyber attack but says it will not be paid*. [Online]
Available at: <https://www.thejournal.ie/hse-cyber-attack-5436981-May2021/>
[Accessed May 15 2021].

Mukherjee, L., 2020. *What Is a Honeypot in Network Security? Definition, Types & Uses*. [Online]
Available at: <https://sectigostore.com/blog/what-is-a-honeypot-in-network-security-definition-types-uses/>
[Accessed March 17 2021].

National College of Ireland, 2021. *IT Systems Outage*. [Online]
Available at: <https://www.ncirl.ie/News/ArtMID/748/ArticleID/587/IT-Systems-Outage>
[Accessed May 13 2021].

Oxf0x.com, 2019. *Setting up dionaea & cowrie with mhn*. [Online]
Available at: <https://neil-fox.github.io/Setting-up-Dionaea-&-Cowrie-with-MHN/>
[Accessed October 10 2020].

Pwnlandia, 2014. *Repositories*. [Online]
Available at: <https://github.com/pwnlandia>
[Accessed September 14 2020].

pwnlandia, 2017. *MHN Troubleshooting Guide*. [Online]
Available at: <https://github.com/pwnlandia/mhn/wiki/MHN-Troubleshooting-Guide>
[Accessed November 10 2020].

Symanovich, S., 2020. *What is a honeypot? how it can lure cyberattackers*. [Online]
Available at: <https://us.norton.com/internetsecurity-iot-what-is-a-honeypot.html>
[Accessed September 20 2020].

wikipedia, 2014. *Honeypot (computing)*. [Online]
Available at: [https://en.wikipedia.org/wiki/Honeypot_\(computing\)](https://en.wikipedia.org/wiki/Honeypot_(computing))
[Accessed September 20 2020].

9. Appendices

This section should contain information that is supplementary to the main body of the report.

Project Plan

National College of Ireland

Project Proposal

Cloud-Based Research Honeypot

08-11-2020

BSc (Hons) in Computing

Cyber Security

Academic Year 2020/2021

Umar Rafique

18142061

x18142061@student.ncirl.ie

Contents

1.0	Objectives	2
2.0	Background	2
3.0	Technical Approach	3
4.0	Special Resources Required	3
5.0	Project Plan	4

6.0	Technical Details	4
7.0	Evaluation	5

1.0 Objectives

The main objective of this project is to create a cloud-based honeypot, gather the data from attackers to create a more potent defence mechanism against online attacks. Last but not least, as a Cyber Security student, I need to learn and develop the habit of research to keep myself up to date with online attacks.

I choose cloud-based honeypot because of few online articles. It was suggested not to run honeypot on the same network & machine we use for daily activities because it is very likely that the machine (PC/Laptop) can be easily compromised. Also, the online honeypots assure us that the data of log files we will be receiving from attackers will be safe in a secure manner.

2.0 Background

As we know, the different types of cyber-attacks are increasing daily. The following studies have shown the cyber-attacks are increased due to the Covid-19 lockdown. Please find the detail by clicking on the link:

Cyber Attacks Studies on the extended enterprise.

The link also discussed; the type of attacks differs from countries. The security of systems and networks should be enhanced as the attacks are opportunist and luring people by Covid-19 related information. The typical methods used by organisations to defend their systems and networks are Firewall, Intrusion Detection System, Intrusion Prevention System, Penetration tests, or using third-party applications to prevent and detect cyber-attacks. We should always keep in mind that all of these systems have vulnerabilities and

can be compromised at some stage. To enhance the security of systems and networks from existing vulnerabilities currently present in systems, I decided to create cloud-based honeypots.

Honeypot is a computer or computer system intended to attract cybercriminals by showing them it is an easy target for cybercriminals to gain data. Whilst it operates oppositely, the honeypot is used to detect attacks or divert cybercriminals from legitimate targets. The honeypot is also used to gain cybercriminals' information and behaviour learning. There are two main types of a honeypot, classified as Production honeypots and Research honeypots.

Production honeypots: are deployed inside the production network with other production servers by an organisation to enhance the network's overall security. The production honeypots are easy to use, but it catches only limited information.

Research honeypots: comparing to production honeypots, research honeypots are complex to install and maintain. The research honeypots are run to gather information from a broad audience to learn their techniques and motives. Research honeypots do not add values directly to a company or organisation. While learning the new cyber-attacks and their behaviour, researchers, military, and government organisations usually use the research honeypots.

Research honeypots can be classified as:

Pure honeypots

High-interaction honeypots

Low-interaction honeypots

[https://en.wikipedia.org/wiki/Honeypot_\(computing\)](https://en.wikipedia.org/wiki/Honeypot_(computing))

3.0 Technical Approach

In order to develop and complete this project, First, I need to research the requirements for the project. I have learned from past experiences that if project goals are well defined for the short and long term, it improves the overall quality of work and help to finish the project on time successfully.

As this is a research-based project, I have decided to finish the project's documentation, which includes the project proposal, specifications, project plan and research requirements of the project. Then I will begin creating my prototype of the project, which will be used in the mid-point presentation.

Then the first step of this project is to deep dive into honeypots. As I have decided to go cloud-based, the research will be required to find a suitable and authenticated platform for deploying the honeypots. Once the honeypots are deployed, I will require checking and keeping them updated. So, I can gather as many as possible logs of cybercriminals after that will require to store those logs in a safe place where the logs cannot be altered.

After receiving the required data, the next step of the project will be to analyse the gathered data and get the user data, requiring an application that can deal with big data.

Once the data are refined, I will progress into learning and researching the mechanism used to defend such attacks and the available options to protect us from such attacks.

4.0 Special Resources Required

I will require much research in this project but not a piece of hardware. My research will evolve around honeypots, cyberattacks, cloud-based platform to host honeypots, machine learning applications, cybercriminals behaviour, existing and alternative ways of preventing such attacks.

I will require to learn Linux as most cloud platforms are run on the Linux Operating System and pull the attackers' logs filled on the stored server. The further challenge will be deep dive into an existing and alternative way of defending systems.

5.0 Project Plan

The Microsoft Project License key did not work also tried to use Microsoft Project Citrix. The Citrix account got temporary locked after attempting few logins try. The message show on Citrix to contact the administrator.

So, I decided to create a table on word rather than Grant chart. The table for project plan is following:

	Tasks	Duration	Start Date	Finished Date
1	Project Pitch	1 Day	18-10-2020	18-10-2020
2	Proposal Report	13 Days	26-10-2020	07-11-2020
3	Ethical Form	1 Day	07-10-2020	08-11-2020
4	Feasibility Study	8 Days		
5	Technical feasibility	4 Days	11-10-2020	15-10-2020
6	Behaviour feasibility	2 Days	15-10-2020	17-10-2020
7	Requirement Analysis	11 Days		
8	Requirement Gathering	5 Days	19-10-2020	23-10-2020
9	Group interaction	1 Day	25-10-2020	25-10-2020
10	Analysis	4 Days	26-10-2020	30-10-2020
11	Design & code	1 Month+		
12	Cloud Platform	10 Days	31-10-2020	10-11-2020
13	Debugging MHN Server	1 Month+	10-11-2020	14-12-2020
14	Mid-term Presentation/ Documentation	8 Days	14-12-2020	22-12-2020
15	Data Collection	20 Days		
16	Data Analysis	10 Days		
17	Data Pre-processing	10 Days		
18	Applying Machine Learning	14 Days		
19	Learning python & R	14 Days		
20	Testing	1 Month		
21	Unit Testing	1 Month	Working	
	Evaluation	1 Month+		
22	Integration Testing & System integration	1 Month		
23	System testing	10 Days		
24	Documentation & Analysis	4 Months		
25	Most Common Attacks	1 Month		
26	Analysis on Attacks	2 Months		
27	Defence Mechanisms for Attacks	1 Month		

6.0

6.0 Technical Details

The technical details will be uploaded later. The only technical detail I can forward right now is:

I will be using a cloud-based platform to host honeypots and machine learning tools, which will allow me to pull the logs of the attack from the server.

I have decided to use Digital Ocean to host the MHN server and droplets (Honeypots), and I will be using Splunk (ML tool) to analyse the data.

7.0 Evaluation

I plan to use my experience that I have gained so far in my academic years. One of the best scenarios while moving forward with the project is to test whenever a stage of application is developed. I will be running almost daily, if not then differently weekly test to keep the project bug free, which is relief on its own. I will be getting feedback from my Supervisor after passing each step to the final product, which would allow me to add or fix the application in a scenario where things can be out of control. By each step, I will prepare a document to keep me on track if I lose the VM or data.

When the project is finalised by myself and Supervisor, I plan to ask colleagues and friends to test the application and get the survey based on their experience, including people currently working within Information Technology Industry and people who are not very familiar with the technology.

The survey will enhance my knowledge and what people think about the project. To see if the project can be launched in Technology Industry and what additional features I will add within the project.

Ethics Approval Application (only if required)

Reflective Journals

Reflective Journal 1:

As I differed my final year Software project due to Covid-19. I decided to go back to college and enrol myself for the academic year of 2020-2021. It was refreshing to keep on track for a final year project in the first few weeks.

For the academic year of 2020-2021, I decided to change my software project. I realised I want to work with something related to my final year Stream, Cyber Security. I came up with creating a honeypot and gathering the data of attackers.

I came across honeypots last year while studying a module called Security Principles and taught by Professor Sachin Sharma. I spoke to Sachin regards the project. We discussed the project, and he authorised it as a good project for a fourth-year software project.

I started researching honeypots. While the Project pitch video was due, I finished and uploaded the project pitch. Whilst waiting on results for the project pitch, I start watching YouTube tutorials. I found few excellent tutorials on honeypots. One showed how to run a honeypot on a virtual machine within your pc or laptop. I also discussed this video with

Sachin as there was a risk factor of running a honeypot on your machine. I do not have another laptop and an internet connection. Because of that, I was suggested to run a cloud-based honeypot.

I am working on Project Proposal and Ethical form, which is due this weekend.

Umar Rafique

18142061. Reflective Journal

Student name: Umar Rafique

Programme: BSc (Honours) in Computing

Month: November

My Achievements

In November

As mentioned in my previous month Journal, I worked on Project Proposal and Ethical form due on November 8. After submitting the first draft of the project proposal and ethical form. I started working on a cloud-based form to deploy my server & honeypots. I received a suggestion from one of my friends to use the Digital Ocean platform because they gave 100 US Dollar free trial for up to 60 days to students. I created an account with Digital form and started working on Honeypots. I required a server to store all the logs. I started working on a server. The MHN-Server was successfully deployed, but unfortunately, I tried to deploy sensor/ honeypots with the server. The server was unable to detect all the attacks for almost a week. The server showed me only one attack. Whilst on the dashboard, I could see different attacks from different countries. After almost a week, I decided to delete the server and associated sensors and honeypots. I created another server with sensors/ honeypots. I am still facing the same issue. Right now, I am trying to troubleshoot this issue

My Reflection

The Cloud-based platform is searched and working on a server.

Intended Changes

Next month, I am working on troubleshooting for the server. If the same issue kept showing up, I might change the cloud platform and try to find an alternative.

Supervisor Meetings

Date of Meeting: last meeting 24/11/2020.

I discuss the so far progress of a project. The next step from Sachin was to try to sort the server issue, and once I start receiving the data as required, connect the server with any machine learning technique and start analysing the data, which would help for final year report and start mentioning everything on the report with the help of screenshots.

Student name: Umar Rafique

Programme: BSc (Honours) in Computing

Month: November

My Achievements

In December

As mentioned in my previous month Journal, I worked with the MHN server and tried to resolve the error. After a deep dive into research, I figured out what exactly was the issue. The MHN celery worker was FATAL. Then I read the troubleshooting guide for MHN Server. After few items, I was successfully able to troubleshoot.

I was also working on a mid-term report and presentation. In mid of December, I received an error message from Digital Ocean, the cloud platform I was using to host my machines. The message indicated that a few of Digital Ocean virtual machines had an issue storing the data on the server and due to it. I might have lost all the data because I did not enable backup. The backup cost almost five dollars on each machine. Due to the issue, I lost my all logs and created the new MHN server from scratch. I have added the screenshots of all the issues and how I could resolve them on my mid-term report.

My Reflection

Researched oo Celery-worker and why the logs were not showing up on the dashboard. Debug the issue with the MHN server.

Intended Changes

Next month, I am working on logs and expand my research with ML tools.

Supervisor Meetings

I discuss the so far progress of the project. The next step from Sachin was to try to sort the server issue, and once I start receiving the data as required, connect the server with any machine learning technique and start analysing the data, which would help for final year report and start mentioning everything on the report with the help of screenshots.

Student name: Umar Rafique

Programme: BSc (Honours) in Computing

Month: January

My Achievements

In January

At the start of this month, my server was crashed due to a large volume of dump files dumped into pots, and the attackers probably got access to my server. I decided to delete and create a new one.

Currently, I am working on data gathering. Three Dionaea pots are up since last week. I will destroy tonight, 05-02-2012 and I will be adding new pots. Dionaea and Crowie.

My Reflection

I have created a new server and researched Splunk's use of it.

Intended Changes

Next month, I am working it will also be data gathering.

Supervisor Meetings

I had an early meeting with my new Supervisor. I discussed the project idea—David like the idea of the project and what I am doing. During the meeting, it was suggested to narrow down the research when the project comes towards the end.

Student name: Umar Rafique

Programme: BSc (Honours) in Computing

Month: February

My Achievements

In February


As mentioned in the previous journal. I mentioned that I would keep working on gathering data for February. My final year software Project Supervisor (David Collins) luckily teaching us a module called Penetration Testing. Which goes side by side, especially the project I am working on for my final year.


David decides to run a few tests on the honeypots during the penetration labs deployed by me for my project. It helps us understand the tools' dept and enhanced the knowledge for professional testing tools.

Also, how to make pdf reports of vulnerabilities through tools. Which I will be using and adding into my final report.

Following are the screenshots of the IP addresses of honeypots and the conversation with my Supervisor regards the scan for finding the vulnerabilities within the pots:

 Umar Rafique
Fri 12/02/2021 09:15
To: David Collins
For brute force
IP address: 159.203.14.64

 Umar Rafique
Thu 18/02/2021 21:32
To: David Collins
Hi David,
As promised, I am attaching the IP addresses of new honeypots.
Raspberry Pie: 142.93.150.151
Raspberry Pie: 64.225.29.75
SSH: 206.189.94.106
Thanks,
Kind Regards,
Umar Rafique

 David Collins
Fri 19/02/2021 11:33
To: Umar Rafique
Hi Umar,
running a scan against 64.225.29.75 taking a long time to start.
thanks,
David



Umar Rafique
Fri 26/02/2021 12:57
To: David Collins



Hi David,

Apology for delay for ending the ip addresses in short notice. It's because I wasn't feeling well yesterday.

Cowrie Honeypots SSH:
67.205.128.94
164.90.170.52
167.99.183.95

Thanks,

Kind Regards,
Umar Rafique

My Reflection

Deployed new honeypots, run few tests with the help of David (Project Supervisor).

Intended Changes

Next month, I plan to work on integrating the Splunk ML tool.

Supervisor Meetings

As David is my Penetration Testing lecturer and my final year software project Supervisor. I feel fortunate to have him as my Project Supervisor because of the knowledge he has about Computer security and vulnerabilities.

After each Penetration lab, David gives us time from his precious time to discuss our module and Software Project. David is always offering a helping hand for module and project. We are having meetings each Friday.

Student name: Umar Rafique

Programme: BSc (Honours) in Computing

Month: February

My Achievements

In March

At the start of this month, I completed the data gathering for my final year project (Honeypots). I researched several techniques suggested by my Supervisor as March was way busy comparing to others. Submitted my CA1 for Penetration testing, Advance Secure Programming and Digital Forensics.

My Reflection

Gathered the data.

Intended Changes

In April, I will be working on Machine Learning Techniques & another way to make the data readable.

Supervisor Meetings

I had an early meeting with my new Supervisor, I discussed the project process and the next step I attended to take. Few suggestions were made by David and alternative ways how the data can be fetched and used in future steps—currently researching those methods, achieving the max performance, and how to mitigate the risk to the lowest level.

Student name: Umar Rafique

Programme: BSc (Honours) in Computing

Month: April

My Achievements

In April

In March, after submitting all the CA and TABA. I worked on Splunk integrated Splunk into my final year project. I tried my best to fetch all the data. After collecting the data, I took the raw data and tried to fetch the information.

The information is mentioned in my final documentation. As currently, I am working on the final documentation.

My Reflection

Transformation of data from raw to readable.

Intended Changes

I will be working on a video presentation in mid-April, and there are no intended changes.

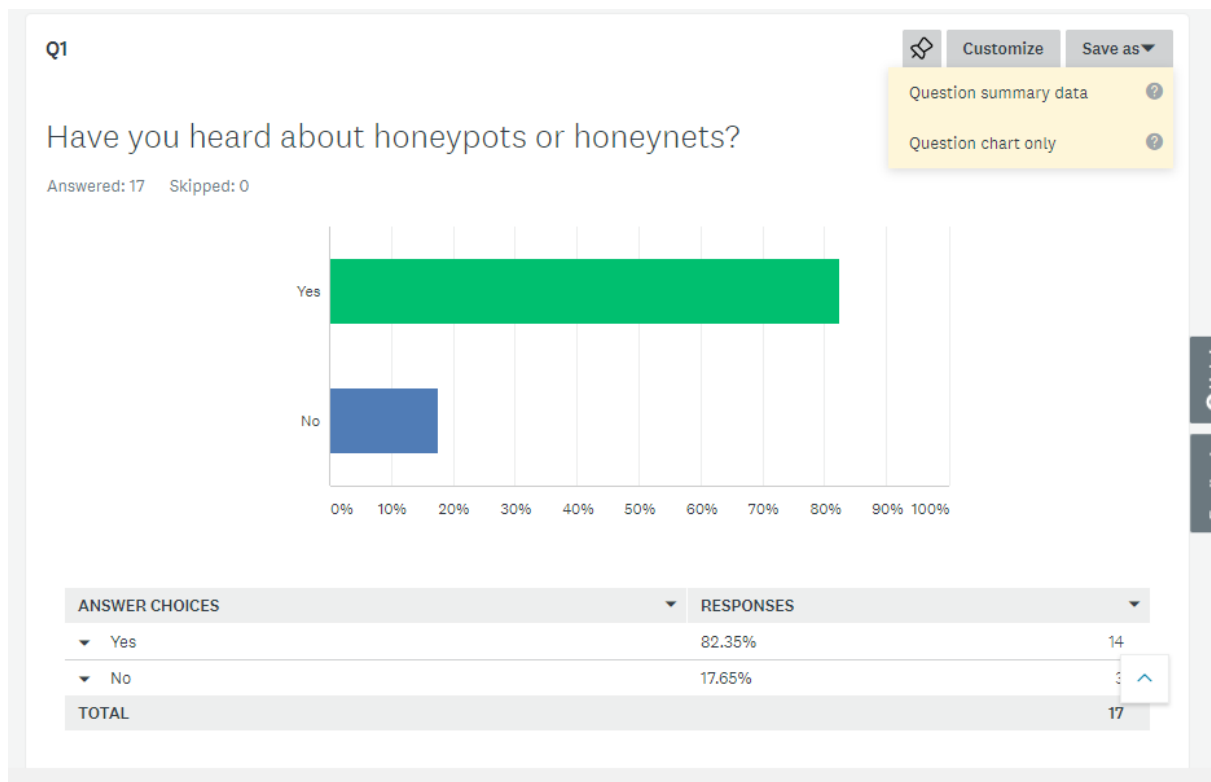
Supervisor Meetings

Yesterday 07/05/2021, I had a meeting with my Supervisor. Explained the progress and welcomed the suggestion. The next supervisor meeting is set for 09/05/2021. The meeting is

to go through the project and documentation to see if any changes are required before final submission.

Survey

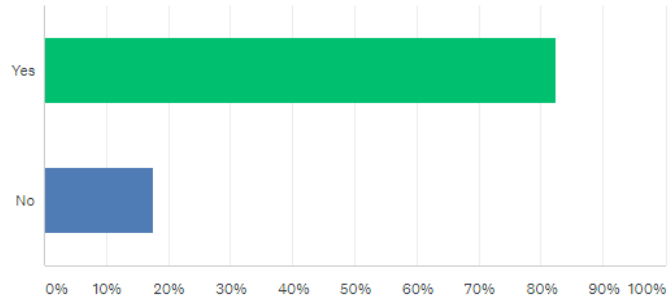
I conducted a survey on Survey Monkey asking Cyber Security students from BSc Honours & Master students about the knowledge and deployment of Honeypots without collecting any personal data. Following are the results from the survey:



As you can see, before starting the actual survey. I wanted to ask students their knowledge regards honeypots, and I was surprised that almost 18% of students did not know about honeypots.

Do you know honeypots or honeynets are used to gather the data from attackers by storing their logs?

Answered: 17 Skipped: 0

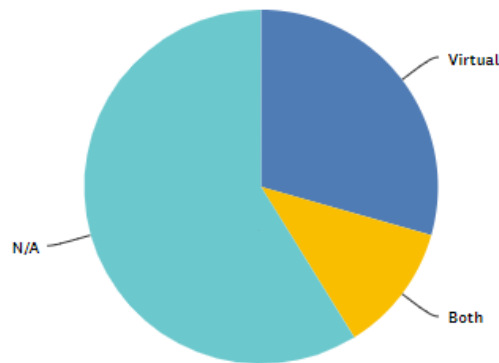


ANSWER CHOICES	RESPONSES	
▼ Yes	82.35%	14
▼ No	17.65%	3
TOTAL		17

The question was related to getting the general knowledge about honeypots, and the ratio remains the same in the above questions.

If you have ever deployed honeypots or honeynets, were they physical or virtual?

Answered: 17 Skipped: 0

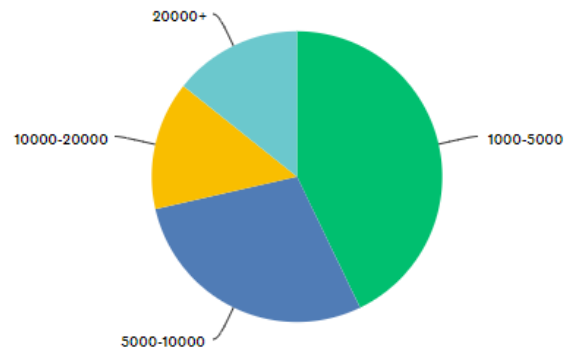


ANSWER CHOICES	RESPONSES
Physical	0.00% 0
Virtual	29.41% 5
Both	11.76% 2
N/A	58.82% 10
TOTAL	17

Then I went ahead and asked students about the deployment of honeypots. Almost 60% of students have not deployed any honeypot. The ratio of deployment for physical remain 0%, the virtual deployment was high as it was almost 30%, and both were quite surprising as two members have done virtual and physical deployment of honeypots.

If you have deployed honeypots or honeynets, Approximately how often your machines received an attack?

Answered: 14 Skipped: 3

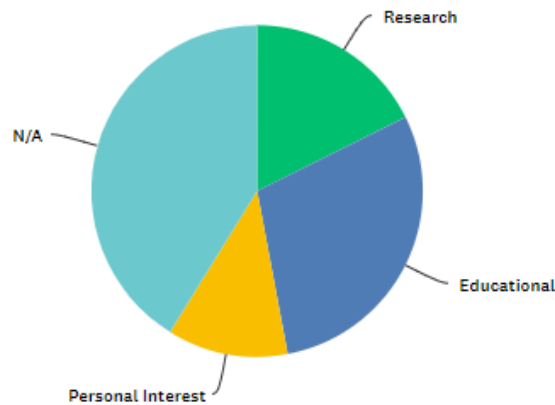


ANSWER CHOICES	RESPONSES	
▼ 1000-5000	42.86%	6
▼ 5000-10000	28.57%	4
▼ 10000-20000	14.29%	2
▼ 20000+	14.29%	2
TOTAL		14

Then I asked how many approximately their machines were getting an attack. The 1000-5000 remains relatively high. As it depends on the honeypot they have deployed, Dionea and annum honeypots get many more attacks than the others.

If you have deployed honeypots or honeynets, what was your primary reason?

Answered: 17 Skipped: 0





ANSWER CHOICES	RESPONSES
Research	17.65% 3
Educational	29.41% 5
Personal Interest	11.76% 2
N/A	41.18% 7
TOTAL	17

My final question was related to deployment, and I was amazed to see mostly it was for educational purpose than research. Only two users clicked the personal interest, and I believe these were the user who deployed both virtual and physical honeypots.

Splunk Overview

-  Past 24hours overview.pdf
-  dionaea overview.pdf
-  Cowrie overview .pdf
-  indicators overview.pdf
-  kippo_analytics-2021-05-09.pdf

Digital Ocean Invoices

-  DigitalOcean Invoice 2021 Mar (8311695-424146355).pdf
-  DigitalOcean Invoice 2021 Apr (8311695-424904618).pdf