CYB102 Milestone 1

(🔗 Instructions Page)

Team Members (Required)

Reminder: Make sure to provide **edit access** for this Milestone document to **everyone on your team!**

👤 Student Name:	Dennys Antunish
💬 Student Pronouns:	He/Him
✉️ Student Email:	dantunish2@gmail.com
🐹Favorite Animal:	Sharks
👤 Student Name:	Evgeniia Yeroshkina
💬 Student Pronouns:	she/her
✉️ Student Email:	mironova.eug2016@gmail.com
🍦Favorite Flavor:	Vanilla
👤 Student Name:	Angie Rivera
💬 Student Pronouns:	She/Her
✉️ Student Email:	angiervr9@gmail.com
🎡Favorite Park:	McCarren Park
👤 Student Name:	Aliya Jones
💬 Student Pronouns:	She/Her
✉️ Student Email:	aliya.jones@macaulay.cuny.edu
🎮Favorite Game:	Fortnite
👤 Student Name:	Navruz Asatullaev
💬 Student Pronouns:	
✉️ Student Email:	navruz.college@gmail.com
☕Favorite Drink:	Water

What are pronouns / Why are they included here?

Select one (or more) open-source Datasets to analyze (Required)

Data Set Chosen: The data set we have chosen to analyze for The Data Dig is...

Name:	Intrusion detection evaluation dataset (CIC-IDS2017)
Primary Link:	https://www.unb.ca/cic/datasets/ids-2017.html

Data Set Description: Where does the data come from? Who generated it? What kind of devices / technologies does it target? What format is the data in?

The CIC-IDS2017 dataset was created by the Canadian Institute for Cybersecurity (CIC) at the University of New Brunswick (UNB). It was developed by Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, and published in the paper *"Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization"* (ICISSP 2018). The goal was to provide a modern and realistic dataset for evaluating intrusion detection systems. The data was captured over five days and includes both normal (benign) and attack traffic based on real-world scenarios such as DoS/DDoS, infiltration, web attacks, and botnets. The dataset targets various devices and technologies, including Windows (7, 8.1, 10, Vista), Linux (Ubuntu 12, 14.4, 16.4), macOS, routers, switches, and firewalls. It focuses on both traditional and modern cyberattacks and covers widely used protocols like HTTP, HTTPS, FTP, SSH, and email. The data is available in both PCAP format and CSV files.

Hypothesis: What are 3 things you expect to find when you analyze the data?
Tip: You won't lose points if these hypotheses turn out to be wrong! Make educated guesses!

Finding #1:	We expect to find multiple victims on Friday on July 7, 2017
Finding #2:	We expect to find a lot of logs of an attempted attack at a company
Finding #3:	We expect to find continuous use of the same IP address for multiple attacks

Select an incident-response playbook to follow (Required)

Playbook Chosen: The playbook we have decided to follow for The Data Dig is...

Name:	GSPBC-1080 - Impact - Network Denial of Service.pdf
Primary Link:	https://github.com/guardsight/gsvsoc_cirt-playbook-battle-cards/blob/master/GSPBC-1080%20-%20Impact%20-%20Network%20Denial%20of%20Service.pdf

Playbook Description: Who wrote this playbook? Who is the target audience? Does it make any specific assumptions about the data set? If so, do those match your data, or will you have to adapt the playbook?

The playbook was created by GuardSight, a U.S.-based cybersecurity company specializing in managed detection and response (MDR), cybersecurity operations, and cyber incident response. It follows their CIRT (Computer Incident Response Team) methodology and is built on the PICERL model (Preparation, Identification, Containment, Eradication, Recovery, Lessons).The target audience would be Security Operations Center Analysts, Cybersecurity Engineers, or basically, anyone involved in defensive security. It does not assume any specific dataset or toolset. It includes DoS attacks and flow-based features which can be mapped to the playbook's identification and containment steps.

Tools we Plan to Use: Based on your dataset and playbook, what blue-team tools from this course will you use to analyze the incident? (MINIMUM of 2)

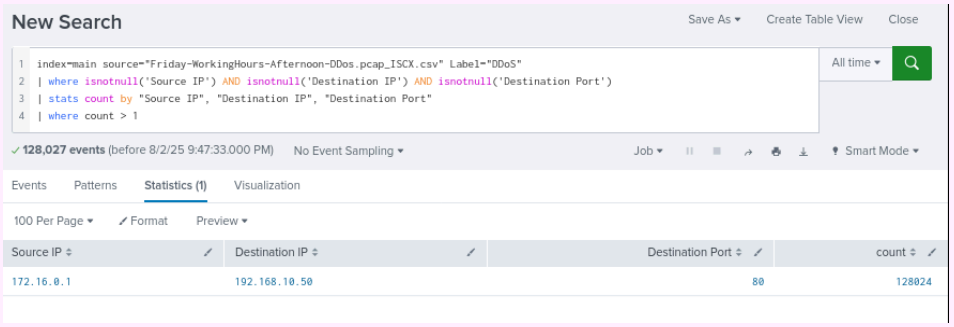
Tool #1:	Splunk
Tool #2:	VirusTotal
Tool #3:	Abuseipdb
Tool # 4:	Catayst

Answer each of the *key aspect* questions (Required)

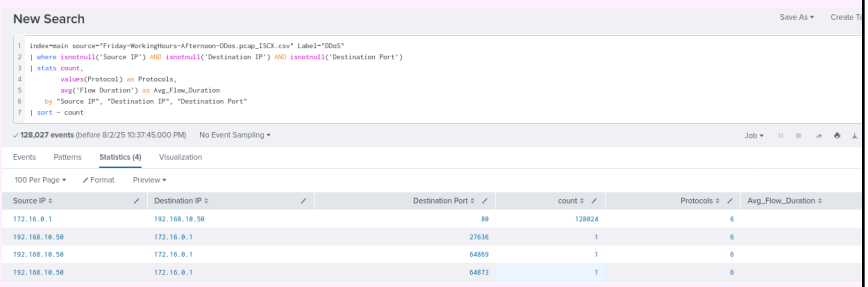
Instructions: *For each of the key aspects below, include a few sentences explaining how your project is demonstrating that aspect. Please include at least one specific example.*

For a full definition of each of the key aspects, please view the Data Dig Project page on the Course Portal.

Monitoring Sources	
How it relates to our project:	We had two zip files labeled Machine Learning and Generative Labeling which we then unzipped to download the Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv which we entered into Splunk to monitor network logs. Network logs let us identify high volume traffic from many IPs toward a single destination — a typical DDoS pattern. We filtered based on "Label=DDoS" and "Destination IP"="192.168.10.50". Network logs were the most important data source for identifying this DDoS incident. These logs include detailed packet-level metadata such as: • Source IP (e.g., the attacker) • Destination IP (target of the attack) • Destination Port (e.g., port 80, 443)

	• Protocol (TCP, UDP, ICMP) • Flow Duration and Timestamps
Example(s):	• In Splunk, we ran a search filtering by Label=DDoS, Source IP=”*”and *Destination IP=””*. This revealed consistent attack behavior originating from a single source IP (172.16.0.1) targeting HTTP services on port 80.  - **index=main source="Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv" Label=* | stats count by Label** - This will output how many events are associated with each label (BENIGN, DDoS).  - **index=main source="Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv" Label="DDoS" | where isnotnull('Source IP') AND isnotnull('Destination IP') AND isnotnull('Destination Port') | stats count by "Source IP", "Destination IP", "Destination Port" | where count > 1** - This helped us spot repeated patterns of DDoS traffic going to the same IP/port from the same source • By analyzing fields like Protocol and Flow Duration, we detected abnormally high traffic volumes, confirming the severity of the activity and validating its classification as

a DDoS attack.



- **index=main source="Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv" Label="DDoS"**
  | where isnotnull('Source IP') AND isnotnull('Destination IP') AND isnotnull('Destination Port')
  | stats count,
  values(Protocol) as Protocols,
  avg('Flow Duration') as Avg_Flow_Duration
  by "Source IP", "Destination IP", "Destination Port"
  | sort - count

- This query groups DDoS-labeled flows by source IP, destination IP, and port, counting occurrences and averaging flow duration. It also lists the protocols used, helping identify patterns of suspicious high volume traffic.

- By organizing this data chronologically (_time), we were able to *track repeated attack attempts over a 5-year span*, which simulated what a real-world persistent threat might look like in a production environment.

- **index=main source="Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv" Label="DDoS"**
  | where isnotnull('Source IP') AND isnotnull('Destination IP') AND isnotnull('Destination Port')
  | bucket _time span=1d
  | stats count, values(Protocol) as Protocols, avg('Flow Duration') as Avg_Flow_Duration
  by _time, "Source IP", "Destination IP", "Destination Port"
  | sort _time

splunk>enterprise Apps ▾ Messages ▾ Settings

Search Analytics Datasets Reports Alerts Dashboards

New Search

```
index=main source="Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv" Label="DDoS"
| where isnotnull('Source IP') AND isnotnull('Destination IP') AND isnotnull('Destination Port')
| bucket _time span=1d
| stats count, values(Protocol) as Protocols, avg('Flow Duration') as Avg_Flow_Duration
    by _time, "Source IP", "Destination IP", "Destination Port"
| sort _time
```

✓ **128,027 events** (before 8/2/25 10:48:48.000 PM) No Event Sampling ▾ Job ▾

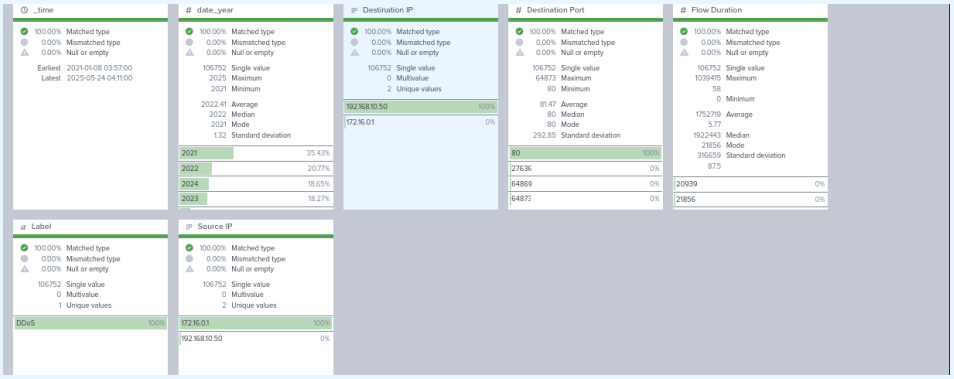
Events Patterns **Statistics (91)** Visualization

_time ⇕	Source IP ⇕	Destination IP ⇕	Destination Port ⇕	count ⇕	Protocols ⇕
2020-02-13	172.16.0.1	192.168.10.50	80	531	
2020-03-31	172.16.0.1	192.168.10.50	80	106	
2020-04-28	172.16.0.1	192.168.10.50	80	566	
2020-06-04	172.16.0.1	192.168.10.50	80	1839	
2020-08-01	172.16.0.1	192.168.10.50	80	1	
2020-09-30	172.16.0.1	192.168.10.50	80	2583	
2020-10-01	172.16.0.1	192.168.10.50	80	1855	
2020-10-12	172.16.0.1	192.168.10.50	80	220	
2020-10-14	172.16.0.1	192.168.10.50	80	1	
2020-10-15	172.16.0.1	192.168.10.50	80	553	
2020-10-31	172.16.0.1	192.168.10.50	80	91	
2020-11-11	172.16.0.1	192.168.10.50	80	3623	
2020-11-15	172.16.0.1	192.168.10.50	80	2012	
2020-11-23	172.16.0.1	192.168.10.50	80	7290	
2020-12-10	172.16.0.1	192.168.10.50	80	2	
2020-12-30	172.16.0.1	192.168.10.50	80	2	
2021-01-08	172.16.0.1	192.168.10.50	80	1238	
2021-04-03	172.16.0.1	192.168.10.50	80	4455	
2021-05-26	172.16.0.1	192.168.10.50	80	5816	
2021-07-18	172.16.0.1	192.168.10.50	80	4147	
2021-08-21	172.16.0.1	192.168.10.50	80	2300	
2021-09-17	172.16.0.1	192.168.10.50	80	251	
2021-09-27	172.16.0.1	192.168.10.50	80	3002	
2021-10-01	172.16.0.1	192.168.10.50	80	2461	
2021-10-07	172.16.0.1	192.168.10.50	80	2021	
2021-10-13	172.16.0.1	192.168.10.50	80	88	
2021-10-18	172.16.0.1	192.168.10.50	80	2045	
2021-10-22	172.16.0.1	192.168.10.50	80	3017	
2021-10-26	172.16.0.1	192.168.10.50	80	2723	
2021-11-04	172.16.0.1	192.168.10.50	80	901	
2021-11-18	172.16.0.1	192.168.10.50	80	1920	
2021-11-27	172.16.0.1	192.168.10.50	80	2	
2021-12-03	172.16.0.1	192.168.10.50	80	986	
2021-12-30	172.16.0.1	192.168.10.50	80	452	
2022-01-14	172.16.0.1	192.168.10.50	80	2020	
2022-02-08	172.16.0.1	192.168.10.50	80	2004	
2022-03-17	172.16.0.1	192.168.10.50	80	1097	
2022-09-09	172.16.0.1	192.168.10.50	80	1660	
2022-10-09	172.16.0.1	192.168.10.50	80	1341	
2022-10-19	172.16.0.1	192.168.10.50	80	1968	
2022-11-05	172.16.0.1	192.168.10.50	80	2709	
2022-11-07	172.16.0.1	192.168.10.50	80	6	
2022-11-11	172.16.0.1	192.168.10.50	80	67	
2022-11-13	192.168.10.50	172.16.0.1	27636	1	
2022-11-13	192.168.10.50	172.16.0.1	64869	1	
2022-11-13	192.168.10.50	172.16.0.1	64873	1	
2022-11-19	172.16.0.1	192.168.10.50	80	4992	
2022-11-30	172.16.0.1	192.168.10.50	80	2437	
2022-12-01	172.16.0.1	192.168.10.50	80	511	
2022-12-04	172.16.0.1	192.168.10.50	80	268	
2023-01-16	172.16.0.1	192.168.10.50	80	2050	
2023-01-26	172.16.0.1	192.168.10.50	80	460	
2023-02-06	172.16.0.1	192.168.10.50	80	734	
2023-02-11	172.16.0.1	192.168.10.50	80	2	
2023-04-05	172.16.0.1	192.168.10.50	80	627	
2023-04-18	172.16.0.1	192.168.10.50	80	4118	
2023-06-29	172.16.0.1	192.168.10.50	80	425	
2023-08-09	172.16.0.1	192.168.10.50	80	3186	
2023-10-03	172.16.0.1	192.168.10.50	80	73	
2023-10-06	172.16.0.1	192.168.10.50	80	757	
2023-10-11	172.16.0.1	192.168.10.50	80	1416	
2023-10-20	172.16.0.1	192.168.10.50	80	1778	
2023-11-07	172.16.0.1	192.168.10.50	80	597	
2023-11-15	172.16.0.1	192.168.10.50	80	491	
2023-11-26	172.16.0.1	192.168.10.50	80	546	
2023-11-30	172.16.0.1	192.168.10.50	80	1394	
2023-12-06	172.16.0.1	192.168.10.50	80	688	

100 Per Page ▾ Format Preview ▾

_time ⇕	Source IP ⇕	Destination IP ⇕	Destination Port ⇕	count ⇕	Protocols ⇕
2024-06-12	172.16.0.1	192.168.10.50	80	432	
2024-07-28	172.16.0.1	192.168.10.50	80	2053	
2024-10-15	172.16.0.1	192.168.10.50	80	518	
2024-10-22	172.16.0.1	192.168.10.50	80	5328	
2024-10-23	172.16.0.1	192.168.10.50	80	4154	
2024-10-27	172.16.0.1	192.168.10.50	80	1196	
2024-11-02	172.16.0.1	192.168.10.50	80	6	
2024-11-04	172.16.0.1	192.168.10.50	80	1592	
2024-11-15	172.16.0.1	192.168.10.50	80	2577	
2024-11-18	172.16.0.1	192.168.10.50	80	8	
2024-11-20	172.16.0.1	192.168.10.50	80	802	
2024-11-28	172.16.0.1	192.168.10.50	80	127	
2024-12-05	172.16.0.1	192.168.10.50	80	83	
2024-12-29	172.16.0.1	192.168.10.50	80	822	
2025-01-11	172.16.0.1	192.168.10.50	80	18	
2025-01-19	172.16.0.1	192.168.10.50	80	1009	
2025-01-20	172.16.0.1	192.168.10.50	80	4	
2025-02-07	172.16.0.1	192.168.10.50	80	4539	
2025-03-07	172.16.0.1	192.168.10.50	80	485	
2025-03-27	172.16.0.1	192.168.10.50	80	702	
2025-04-02	172.16.0.1	192.168.10.50	80	1	
2025-05-24	172.16.0.1	192.168.10.50	80	583	

Identified Assets	
How it relates to our project:	In our project, we looked at a real DDoS attack using Splunk. The data shows us which systems were involved both the attackers and the victim. Understanding this helps us see what was affected, where weaknesses might be, and how tools like Splunk help spot and track these kinds of cyber attacks.
Example(s):	**Target Asset:** • **192.168.10.50** This internal server is consistently targeted on port 80 (HTTP), making it the primary victim of the DDoS attack. ○ *Why it matters*: This server likely runs a web application, and it couldn't handle the huge amount of requests. It may need better protection against heavy traffic like DDoS attacks. **Attacker Assets:** • **Source IPs:** Mostly **172.16.0.1**, repeated many times in the dataset as the attacker's IP. ○ *Why it matters*: It was the machine (or one of many) sending the flood of traffic. Even one attacker can cause serious problems if not blocked. **Network Protocols & Ports:** • **Protocol 6** (TCP) with **Destination Port 80** is dominant, confirming that the attack leveraged standard web traffic to mask malicious behavior. ○ *Why it matters*: Shows potential weakness in perimeter defenses, where standard ports are often less scrutinized, making it easier for attackers

	to slip DDoS traffic past filters. **Systems & Applications Potentially Affected:** • Any **web application** or **HTTP service** running on 192.168.10.50. • **Firewall or IDS systems**, if not properly tuned, could have failed to detect or block this attack. • **Network infrastructure (switches, routers)** could experience performance degradation due to excessive traffic volume. 

Impact Analysis and Triage	
How it relates to our project:	• **This project analyzes a labeled DDoS dataset (CIC-IDS2017) to show SOC detection, impact assessment, and triage. Using Splunk to query Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv and filter Label=DDoS for Destination IP=192.168.10.50, we discovered a high-volume disruption event. The impact analysis illustrates how we judged severity and prioritized mitigation, while triage describes how we scoped the attack and looked for further vulnerabilities.** • **Impact Analysis — How we determined severity** **Data sources & metrics used** - Primary: Network flow/packet logs (CSV from CIC-IDS2017). - Key fields inspected: Source IP, Destination IP, Destination Port, Protocol, Timestamp, packet/flow

	counts, flow duration. • **Severity criteria and findings** - **Event volume:** ~128,000 DDoS events identified — indicates *high* attack intensity. - **Sustained spikes on timechart**: Repeated high counts per minute → indicates sustained service impact (not a short burst). - **Single attacker IP:** Attack originated solely from `172.16.0.1` - **Targeted critical ports (e.g., 80/443)**: If true, this increases business impact (web service disruption). - **Conclusion: Severity = High** because of service availability impact and volume of traffic. Priority: **Immediate mitigation** to restore availability.

Example(s):	**🔎 Triage — How we scoped the incident and what we found**
	
	Triage objectives
	
	• Confirm the incident type (DDoS)
	• Determine scope: isolated or widespread
	Check for signs of lateral movement or further compromise
	• Identify attacker IPs for blocking/containment
	
	Triage steps performed
	
	• Filtered to:
	
	`Label=DDoS "Destination IP"="192.168.10.50"`
	
	• Identified attacker:
	
	`| top "Source IP"`
	`→ Result: Only `172.16.0.1``
	
	• Time-based analysis:
	`| timechart span=1m count`

	→ Showed sustained spikes
	
	• Checked if other systems were targeted:
	
	`index=main source="Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv" Label=DDoS`
	`| stats count by "Destination IP"`
	`| sort - count`
	
	→ **Only** `192.168.10.50` was affected
	
	• Hypothetical log analysis (not available in dataset): check system logs on `172.16.0.1` and `192.168.10.50` for performance issues, service crashes, or signs of compromise
	
	• Threat intelligence lookup (optional in a real-world SOC): Verify if `172.16.0.1` is internal and trusted or compromised
	
	Triage results summary
	
	• **Scope:** Single attacker (`172.16.0.1`) targeting single destination (`192.168.10.50`)
	Attack type: Single-source DDoS — continuous, high-traffic flood
	• **No lateral movement** or multiple internal hosts involved
	• **Key indicators captured:** Attacker IP, timeline, protocol/port usage

Threat Intelligence	
How it relates to our project:	If you came across any relevant threat intelligence during your analysis, we will be sure to discuss this in your presentation. This might include information about the threat actors involved in the incident, the tactics, techniques, and procedures (TTPs) used, and any indicators of compromise (IOCs) that were identified. -Who owns the destination IP address? –targeting private company, the attacker used someone within the company

-VirusTotal and the other one

destination ip - companies private ip address source ip - copying the company's ip address but covering it up

Example(s):

Type	Example	Why it matters
Source IP	`172.16.0.1`	Repeated source of attack traffic
Destination IP	`192.168.10.50`	Targeted internal server (likely hosting a web service)
Destination Port	`80`	HTTP — often abused in DDoS due to open access
Protocol	`6 (TCP)`	Used to mimic legitimate HTTP traffic
Flow patterns	High counts, short duration	Indicates flood behavior (common in DDoS attacks)

Tactic	Technique	Details from Dataset
Initial Access	External Remote Services (T1133)	Traffic flooded from external IP toward open HTTP port
Impact	Network Denial of Service (T1498.001)	Repeated, large-volume TCP traffic over port 80
Evasion	Abuse of Legitimate Protocol (T1071.001)	Using HTTP/TCP to disguise the attack as regular web traffic

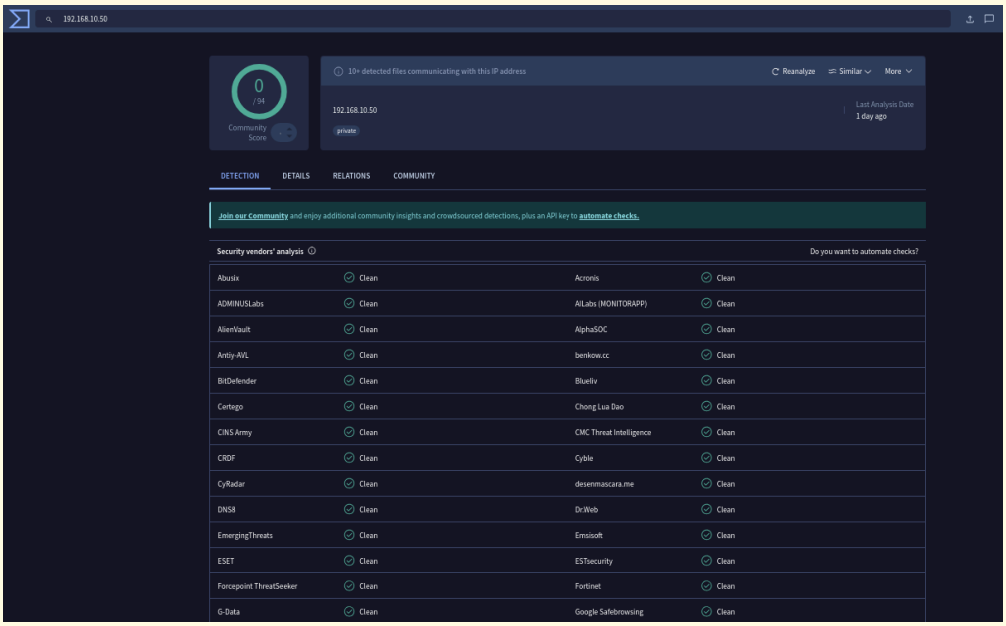
From our analysis, we extracted key indicators like the attacker's IP address 172.16.0.1, the targeted server 192.168.10.50, and the fact that the attack used

standard HTTP traffic on port 80. These are known as indicators of compromise and help security teams write firewall or IDS rules to block or flag such activity.

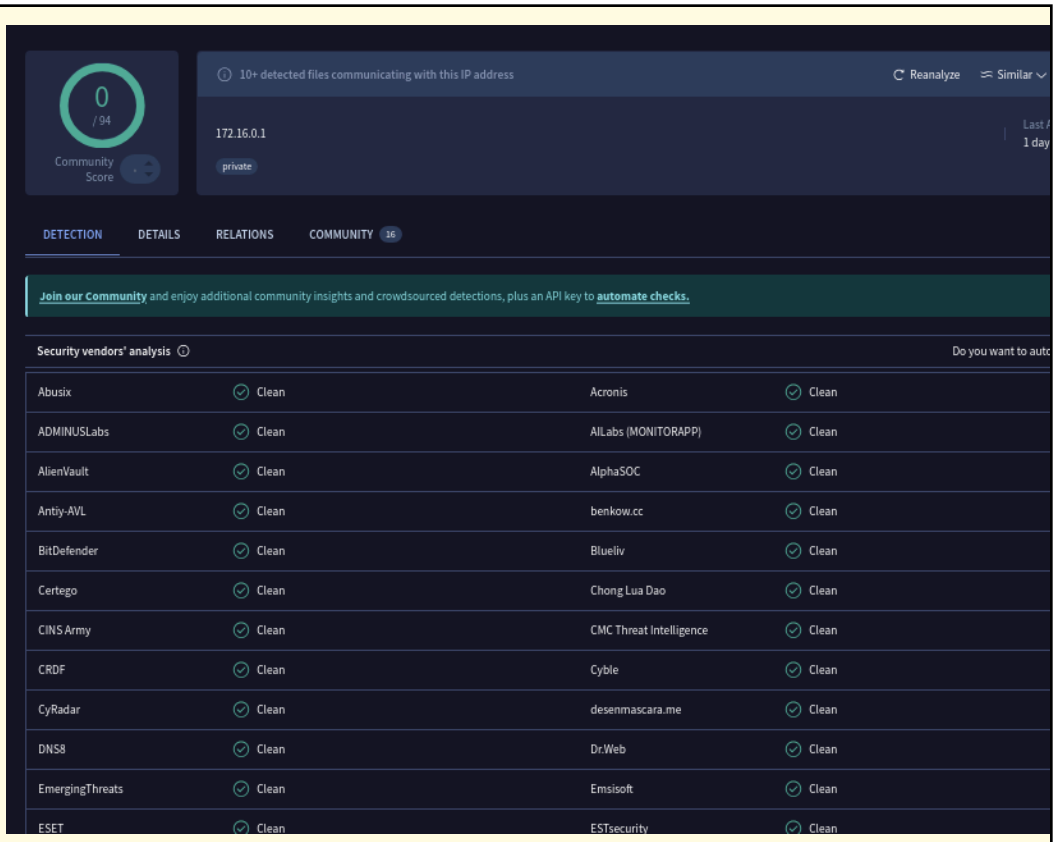
In terms of attacker behavior, or what's known as TTPs, we observed techniques commonly associated with denial-of-service campaigns.

In a real-world SOC scenario, this behavior would resemble a botnet-driven DDoS, where compromised machines (often part of a botnet like Mirai) flood a victim server. While we only saw one IP in this dataset, real DDoS attacks often involve hundreds or thousands of IPs coordinated across the globe.

- Private IP `192.168.10.50`

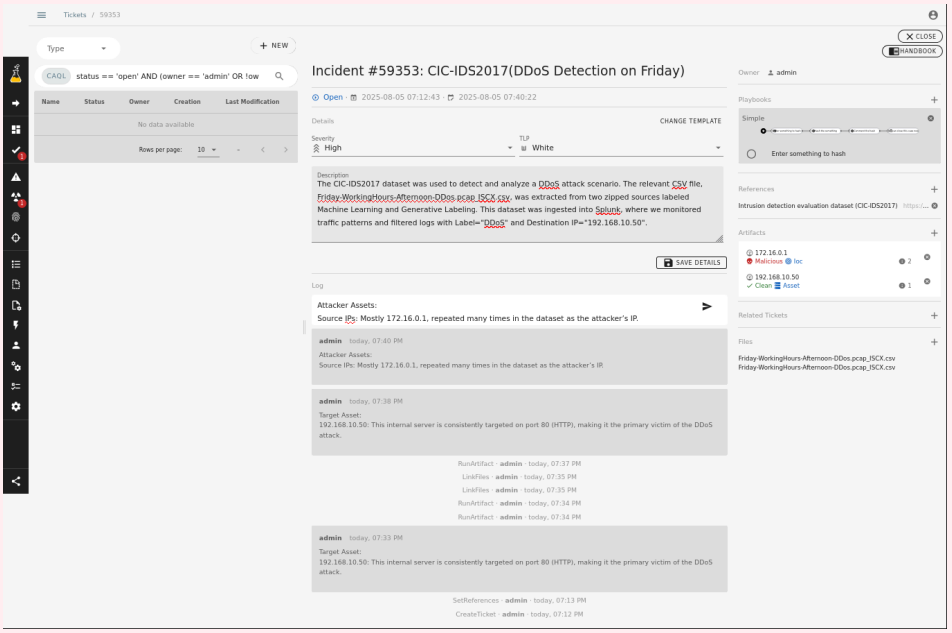


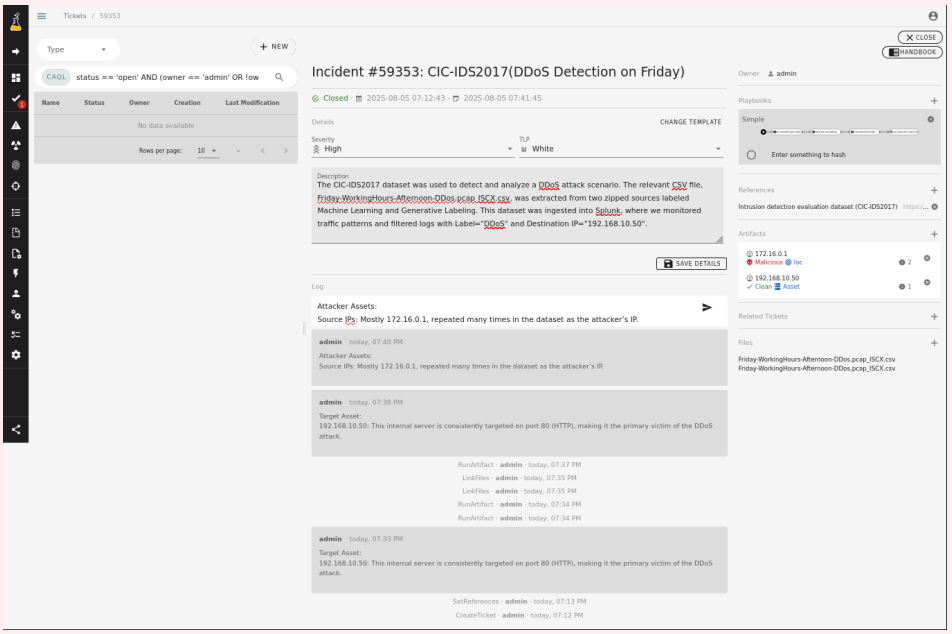
- Private IP `172.16.0.1`



Recommended Remediation	
How it relates to our project:	Based on what we found, we came up with the steps below to help protect the server and reduce the chance of this happening again in the future.
Example(s):	**Block Attacker IP (172.16.0.1)** Use a firewall rule to block the IP that generated the DDoS traffic. **Limit HTTP Requests per IP** Set rules to block IPs making more than 100 HTTP requests per minute to prevent flooding. **Deploy Cloudflare WAF** Add a Web Application Firewall in front of 192.168.10.50 to block suspicious HTTP traffic. **Set Up Splunk Alert for DDoS Pattern** Set up Splunk alerts to notify the team when a single IP sends over 100 requests per minute or when there's a sudden spike in

	HTTP traffic. **Force HTTPS and Disable HTTP on Port 80** Force HTTPS by installing an SSL certificate, redirecting port 80 traffic to 443, and disabling HTTP if it's not needed. **Update Devices** Update the OS, firewall, and web server software on 192.168.10.50, and disable any open ports or services not being used.

Case Management System (and screenshots)	
How it relates to our project:	We used the Catalyst case management system to log each step in the incident response process. This made it easier to track analysis, SOCs, containment actions, and recovery.
Example(s):	** CATALYST**  * Incident Log  * OPEN



* Closed

Presentation Prep (Required)

Presentation Plan: What is your plan for the presentation? Please include a roadmap, flowchart, diagram, or outline.

Things to consider:

- ☐ What will you talk about, and in what order?
- ☐ Who will be talking at what times?
- ☐ What visual-aids will you use?

Introduction - Dennys 15 sec
Dataset - Dennys 30 sec
Hypothesis - Aliya 15 sec
Playbook - Aliya 30 sec
Monitoring Sources - Angie 60 sec
Identified Assets - Angie 60 sec
Impact Analysis and Triage - Navruz 45 sec
Threat Intelligence - Dennys 45 sec
Recommended Remediation - Evgeniia 45 sec
Case Management System - Angie 45 sec
Conclusion - Aliya 30 sec
We will have a PowerPoint presentation

Submission Checklist

👉*Check off each of the features you have completed.* ***You will only be graded on the features you check off.***

Required Features

- ☑ ~~Select one (or more) open-source Datasets to analyze~~
  - ☑ ~~Data Set Chosen (Name & Link)~~
  - ☑ ~~Data Set Description~~
  - ☑ ~~3 Hypotheses Made~~
- ☑ ~~Select an incident-response playbook to follow~~
  - ☑ ~~Playbook Chosen (Name & Link)~~
  - ☑ ~~Playbook Description~~
  - ☑ ~~2+ Tools Identified~~
- ☑ ~~Answer each of the key aspect questions:~~
  - ☑ ~~Monitoring Sources~~
  - ☑ ~~Identified Assets~~
  - ☑ ~~Impact Analysis and Triage~~
  - ☑ ~~Threat Intelligence~~
  - ☑ ~~Recommended Remediation~~
  - ☑ ~~Case Management System~~
- ☑ ~~Your presentation plan: A roadmap, outline, or diagram~~

Submit your work!