SOC ATTACKER CENTRIC - Analysis of a prevention oriented SOC

Master's Degree Programme in Information and Communication Technology
Department of Computing, Faculty of Technology
Master of Science in Technology Thesis
Cyber Security

Author:
Mirko Ioris

Supervisors:
Tahir Mohammad
Petri Sainio

UNIVERSITY OF TURKU
Department of Computing, Faculty of Technology
Master's Degree Programme in Information and Communication Technology

Subject: Cyber Security
Title: SOC ATTACKER CENTRIC - Analysis of a prevention oriented SOC
Author: Mirko Ioris
Number of pages: 62 p.
Date: December 2022

This thesis will explain what a Security Operation Center (SOC) is and how it works, analyzing all the different phases and modules that make up the final product. Typically, a SOC centralizes all of the company's information in one place where it can constantly keep an eye on the data and monitor the system. The IT infrastructure is analyzed in real time for anomalies, malicious activities, or intrusion attempts. Not only the data sent from one machine to another, but also the physical state and resources (e.g., memory and CPU) are constantly monitored. Through the creation and use of multiple detection rules, various alerts are generated and are then reviewed by the SOC analyst team, which promptly informs customers in case of need.

The State of the Art will be explored to study current SOCs and best practices adopted. Then the innovative **SOC Attacker Centric** developed by the company Wuerth Phoenix will be analyzed. The functioning of the SOC-AC will be studied and explained, highlighting how it adds to the classic suite of services offered by a SOC an extra part, focused on the attacker's point of view. This SOC-AC is capable of covering the reconnaissance phase, usually neglected by SOCs, in which attackers gather information about a target in order to find the best strategy to break in and successfully carry out the attack.

In the last part of the thesis, the design and implementation of an automatic SOC reporting functionality will be shown. An important feature is to have an efficient communication channel with the customer and to provide them with data on the status of the SOC they are paying for. Initially, this procedure was a static, manually executed, error-prone process. The procedure was improved by creating a semi-automatic system of report generation and delivery using the Elastic SIEM and several languages such as python, bash, Lucene, Elastic, and Kibana Query Languages, leaving the reporter with fewer parts to analyze and document, saving time and resources.

Keywords: SOC, prevention, monitoring, attacker's perspective, report

Contents

List of acronyms

ANSSI	French national cybersecurity agency
API	Application Programming Interface
APT	Advanced Persistent Threat
CERT-EU	Computer Emergency Response Team of the European Union
CSIRT	Computer Security Incident Response Team
CSRF	Cross-Site Request Forgery
CSS	Cascading Style Sheets
CTI	Cyber Threat Intelligence
CVE	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Scoring System
DB	Data Base
DDoS	Distributed Denial of Service
DMARC	Domain-based Message Authentication, Reporting & Conformance
DNS	Domain Name System
ECS	Elastic Common Schema

ENISA	The European Union Agency for Cybersecurity
GUI	Graphical User Interface
IDS	Intrusion Detection System
ILM	Index Lifecycle Management
IoA	Indicator of Attack
IoC	Indicator of Compromise
IoT	Internet of Things
IP	Internet Protocol
IPS	Intrusion Prevention System
IT	Information Technology
JSON	JavaScript Object Notation
MTDR	Managed Threat Detection and Response
NDJSON	Newline Delimited JSON
OSINT	Open Source Intelligence
PDF	Portable Document Format
RIR	Regional Internet Registry
RSS	Really Simple Syndication
SATAYO	Search All Things About Your Organization
SIEM	Security Information and Event Management
SoA	State of Art

SOC-AC	Security Operation Center - Attacker Centric
SOCaaS	SOC As A Service
SOC	Security Operation Center
SPF	Sender Policy Framework
SQL	Structured Query Language
SSL	Secure Sockets Layer
SSO	Single Sign On
TLD	Top Level Domain
TLS	Transport Layer Security
TTP	Tactics Techniques Procedures
URL	Uniform Resource Locator
VM	Virtual Machine
VPN	Virtual Private Network
WAF	Web Application Firewall

1 Introduction

A Security Operation Center (SOC) is a complex structure that provides an overview and situational awareness of an enterprise and enhances security with the ability to detect anomalies, threats and possible intrusion attempts through constant monitoring. It collects all events generated by IT components in one place, where they are analyzed by a team of security analysts. Searching for “SOC” on Google returns pictures of rooms full of large screens, used by employees or hanging on the walls, that continuously display data and graphs. This is the typical structure of a SOC, but its elements are more complicated than just a couple of charts. The SOC is a relatively new concept, developed over the last fifteen years as a countermeasure to effectively detect cyber-attacks that are becoming increasingly complex. There are different types of SOCs, from the small ones run in-house by a company to huge ones where a large number of analysts work 24/7. Most SOCs are managed services because building and maintaining your own SOC is quite expensive. Every company should have a SOC to protect itself, especially nowadays, but the costs to afford it are often too high for small organizations.

According to Internet World Statistics [1], there are more than 5.4 billion users connected to the Internet in 2022. With so many people interconnected, the attack surface (i.e., the number of possible victims of a cyber-attack) increase tremendously. Most people use their smartphone connected to the Internet, every household usually has a laptop or personal computer that they use at home, and in every company, regardless industry in which the organization operates, there must be computers or servers to store information and access the Internet to keep up with the digital transformation that affects all businesses. IoT is also becoming increasingly popular, and a wide variety of devices are connected in a way that has never been seen before. Unfortunately, many technologies have been developed without security in mind, and with more interconnected devices the likelihood of finding one with a vulnerability that can be exploited is higher.

Protecting the organization is no longer enough; after the COVID-19 pandemic, hybrid work has become the new normal, and many IT companies are allowing employees to work from home. Flexible working broadens the firm's perimeter, with attackers likely to find it easier to target home networks and then slowly gain access to sensitive information. Humans are often seen as the weakest link in the chain and are therefore attractive targets.

"It takes 20 years to build a reputation and few minutes of cyber-incident to ruin it." - Stephane Nappo

In cybersecurity, in a matter of seconds all your data can be stolen or encrypted by threat actors. Data show a 50 percent increase in weekly attacks on corporate networks in 2021 compared to 2020, with a peak of 925 cyber-attacks a week per organization, globally [2]. The frequency and variety of cyberattacks continue to grow each year, with a sustained increase in ransomware with 623.3 million attacks globally in 2021 [3], resulting in tons of gigabytes of sensitive and private data being posted on the internet and accessible by everyone. Not only do ransomware ruin reputations, but it is also responsible for revenue losses caused by the ransom itself and the time it takes to recover all the data and put everything back to work as before. According to estimates, 71.1 million people are victims of cybercrime yearly [4] and the global annual cost of cybercrime is estimated at 6 trillion dollars per year

[5]. In addition, the war in Ukraine has contributed to an increase in cyber-attacks worldwide. Figure 1.1 shows the list of the involved hacker groups.

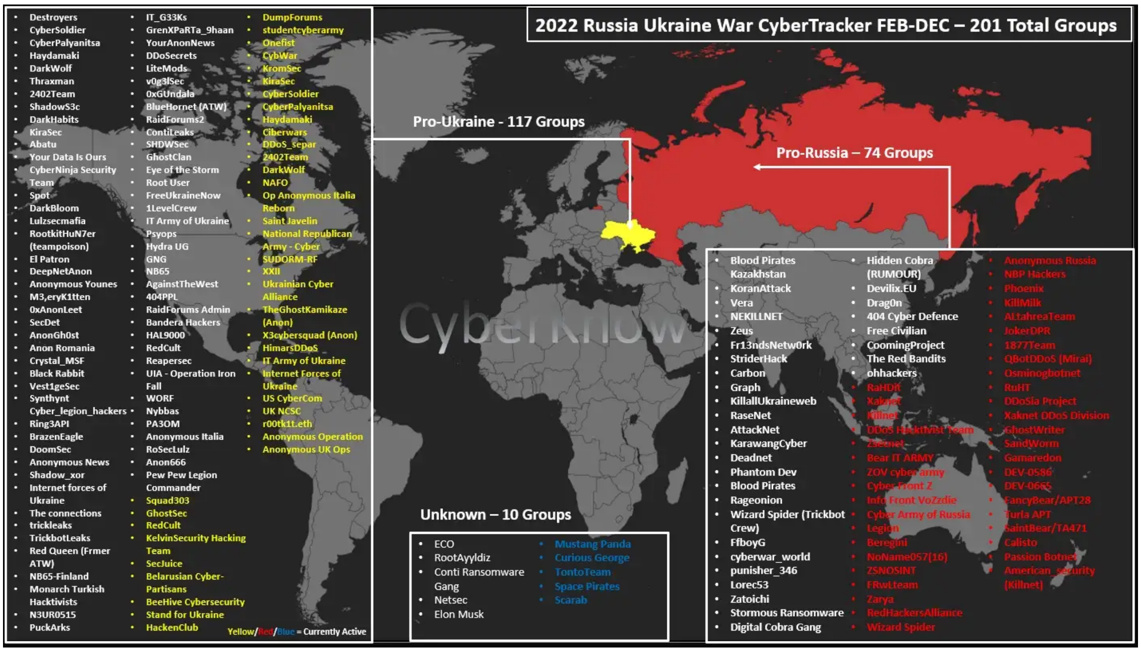


Figure 1.1: List of Pro-Ukraine and Pro-Russia hacker groups. Courtesy of Cyberknow [6]

Defenders have always been at a disadvantage: they have to protect everything and make sure there are no security holes to avoid catastrophic scenarios, while an attacker only needs to find a way in to win and penetrate the system. Threat agents have become smarter and attacks are increasingly sophisticated to avoid being detected as malicious by antivirus or detection tools. Malevolent content can be encrypted, or the virus can be polymorphic or metamorphic and transform itself once it reaches the victim's computer to pass through the firewall among other genuine packets. Detecting and preventing attacks became difficult; virus policies and signatures must be constantly updated, and the vulnerabilities patched, which requires a lot of effort and time.

After all, devices run software, and software is made by humans, who can make mistakes. Programs used today often have millions of lines of code and it would be impossible to guarantee that they are foolproof. In fact, in the first part of 2022, an average of 2000 new vulnerabilities were discovered each month [7]. These are classified with a severity score called CVSS (Common Vulnerability Scoring System) that ranges from 0 to 10. Security researchers and software engineers do their best to find these vulnerabilities and inform vendors who try to release official fixes on time. However, in most cases, companies do not immediately update their software because of compatibility issues, and the window an attacker has to launch a successful attack is wider.

This is why a Security Operation Center could be useful. A SOC can improve and strengthen security controls and protect an organization more efficiently. By having all the data accessible from a unique point and being able to analyze and correlate every event, the protection offered by a SOC is of a high level and strongly recommended. Unfortunately, not many vendors offer this solution, and it is usually expensive for small companies, which decide not to purchase it.

This thesis will be structured as follows: first we will have a look at SOCs in general, with information on their structure and how they are used, then we will examine more in detail the main components that make up the SOC used at Wuerth Phoenix. In the last part we will focus on a feature that was implemented by me to improve the product.

2 Literature review

2.1 Introduction

A study of the State of Art (SoA) was conducted to understand what the current knowledge about SOCs, and cyber threats is and to identify differences between a traditional SOC and the Attacker Centric approach adopted by Wuerth Phoenix.

As Onwubiko states in his paper [8], a good approach to protecting an organization's critical services is to continuously monitor enterprise technologies and be **ready to react** to security breaches or violations with an incident response plan. This **persistent control** can be provided by a SOC, which identifies, monitors, and protects the infrastructure. The SOC can identify threat actors and attempts to exploit vulnerabilities and offers a **real-time monitoring** against cyber-attacks. Sadly, cyber-attacks are becoming more frequent, and the target of the next attack can be anyone

"It's not a matter of "if" but "when" there will be a breach" - Popular saying on cybersecurity

Investing in a SOC can **reduce the losses** a company may suffer after an attack. With proactive detection, it is possible to tell when a system component has been breached by an unauthorized person and act quickly to mitigate the damage. Attacks can come from insiders or outsiders and often go unnoticed and are not discovered until several months later.

IBM publishes an annual report in which the average cost of data breaches is calculated and displayed with various statistics [9]. In the last year there was a 10 percent increase in the average cost compared to 2020, with a value of $4.24 million, that becomes $4.62 million in the case of a ransomware breach, not including the cost of the ransom. The average time to detect a breach was 212 days, plus another 75 to contain it. Investing in a SOC or a SIEM helps minimize this huge time frame and allows organizations to act faster and reduce the financial impact, saving several million dollars.

To emphasize the importance of the topic nowadays, **ENISA** (The European Union Agency for Cybersecurity) published a report in late 2020 [10] describing how to set up **CSIRT** (Computer Security Incident Response Team) and SOC, pointing out a guide for all those interested in implementing such a service and protecting their assets.

The structure of a SOC has been defined in several papers, often with some differences. Mallory documents in his article that a proper SOC should include a SIEM, a Threat Intelligence platform, an Incident Response process and 24/7 operational monitoring time [11]. Comprehensive research on the State of the Art was conducted by Vielberth et al. [12] a couple of years ago. Their work is impressive and was able to summarize in one document all the **building blocks** of a SOC, from general aspects such as architecture and operational models to detailed explanations of people and their roles within the SOC, processes and how they are divided and organized, technology and how to collect and analyze data and finally, compliance and how to audit based on various metrics.

2.2 SOC Structure

According to Vielberth et al. [12] a SOC can be divided into four components, all of which are interconnected to make the SOC work.

2.2.1 People

Running a SOC requires a staff of several people with various roles and responsibilities. Of course, the number and skills involved may vary depending on the type of SOC. The workers that never lack are **SOC Analysts**, whose job is to review and analyze alerts, identify whether they are false or true positives and act accordingly. Analysts are divided into **three levels** based on knowledge, skills, and responsibilities. Third level analysts are the most experienced and handle critical events forwarded to them from lower levels. The **SOC manager** oversees the activity of the team, and other figures such as malware analysts, threat intelligence researchers, and security engineers can collaborate with the analysts to increase the level of security offered by the SOC. An ongoing **training process** for all personnel is always highly recommended because threats evolve rapidly.

2.2.2 Processes

The initial process on top of which the SOC is built is the **data collection** process. This process serves to normalize data collected from multiple sources into a standard format to allow an easier analysis and correlation of logs. It also filters out some unnecessary data, such as entire events or only specific fields in a log. Logs can be also aggregated and ranked with priority levels to facilitate further analysis. The **detection** and **analysis** process follows the previous one, in which incidents are identified and appropriate damage control measures are taken.

2.2.3 Technology

A SOC contains a combination of technologies that work together. There are a security information and event management system (SIEM), intrusion detection or prevention systems (IDSs/IPSs) platforms. Modern SOCs have also begun to integrate some **artificial intelligence** algorithms, which are likely to be heavily used in the future.

2.2.4 Governance and Compliance

A SOC can ensure that **compliance** with regulations, standards, and guidelines is met. It is an important source of information for **auditors** and can help analyze and evaluate the IT infrastructure of a company. This can be done with the use of metrics. Assets that can be monitored are related to the SOC performance, such as the number of devices covered, average analysis time, number of incidents handled, number of policy violations, remediation enforcement time, and many others.

2.3 SOC Functions and Aspects

An entity, to be classified as a SOC, must satisfy certain requirements, and possess specific aspects. Primary aspects are mandatory, while secondary aspects are extra functions offered in addition to the primary services. These aspects allow us to determine the level of service and compare different SOC offerings. Jacobs et al. [13] identify the following **primary aspects**.

- Log Management
  - Log collection: the ability of retrieving all logs produced by events and having them in one place, ready to be analyzed
  - Log retention and archival: the capability of storing logs for a different number of times. It may vary depending on the content of the log and legal constraints
  - Log analysis: the ability to analyze collected and archived logs in detail, with the possibility to create dashboards or calculate metrics on them.

- Monitoring
  - Operational time: at least 8/5 service, but most SOCs offer a 24/7 service.
  - Different devices: the ability to monitor multiple devices from different vendors provides better insights and increased security.
  - Event correlation: the ability to correlate events happened on different machines or at different times to detect an anomaly.
- Threat analysis
  - Threat identification: the capability to identify threats present on one or more devices.
  - Reaction to threats: reacting in real-time to potential threats and taking measures to mitigate damage.
  - Incident management: being able to react to incidents and escalate them if necessary.
- Reporting
  - Security reports: are used to keep the customer informed. They usually include graphs and statistics along with analysis of threat actors.

Incident analysis is carried out by SOC analysts. An alert is investigated to understand its nature, which can be **false positive**, **false negative** or **true positive**. In the first scenario there was no real threat, and the alert is closed. It may be documented for future analysis, but not escalated. The second scenario is probably the worst, because there was a malicious activity that didn't trigger any alarm. It is important to keep detection rules up to date to avoid such situations. The third scenario occurs when a threat is correctly detected in the system. Corrective measures should be applied as soon as possible to mitigate the damage.

The secondary aspects in SOC comprise [13]:

- Malware analysis
- Vulnerability scanning and analysis
- Penetration testing
- Integration with other controls (e.g., physical security)
- Device management
- Identity Attestation

Secondary aspects are many and can change over time, but most of them are related to Threat Intelligence. Being aware in time of potential data breaches and threats is **valuable** and helps prevent attacks.

2.4 SOC and Cyber Threat Intelligence

Cyber Threat Intelligence (CTI) adds another layer of information to a classic SOC and helps in finding true positives. CTI focuses on collecting information to characterize technical threats that can affect a business, providing better cyber resilience with proactive defense. CTI reveals the **tactics**, **techniques**, and **procedures** (TTPs) of attackers and knowledge of their behavior enables a faster response to incidents and perhaps even anticipate their next move. In a SOC environment, CTI is used to **enrich alerts** and link them to incidents, along with the implementation of new detection rules and controls.

CTI is usually divided into three categories [14]:

- Tactical Threat Intelligence: Identifies **new threats** with IoC and IoA. Analysis of these two indicators is mainly automatic and constantly updated. They are explained in detail in the next section.

- Operational Threat Intelligence: Identifies the **context** in which threat actors conduct their operations. It is used in the SOC to derive use cases from TTPs.
- Strategic Threat Intelligence: Identifies the **risks** of cyber threats and their impact on organizations. It informs business processes on how to protect companies.

We can find some advice about the usage of CTI in a recent guide on how to build a SOC written by AT&T [15], but searching for a product on the market that supports this functionality is not so easy. There are some solutions, but it is a relatively new concept and still far from being adopted by everyone. The market is still saturated with conventional solutions that are outdated and ineffective [16]. Exploits are more common and dangerous today than in the past, and enhanced situational awareness is needed to detect both common and sophisticated exploits [17]. This can be achieved with next-generation SOCs, which include a CTI or threat hunting module and are based on artificial intelligence algorithms.

A popular platform intended for processing and sharing knowledge for cyber threat intelligence purposes is OpenCTI [18]. Initially developed by **ANSSI** (French national cybersecurity agency) along with **CERT-EU** (Computer Emergency Response Team of the European Union), it is now an open-source product, complemented by connectors for many services such as AbuseIPDB, AlienVault, CrowdStrike, Mitre Att&ck, VirusTotal, Triage, Shodan and many others [19]. With the platform it is possible to consult the cyber knowledge of the entire community and visualize and share TTP data of threat actors. Analysts can use it to fully understand a threat and perform a more accurate analysis.

Finally, social media and information channels should not be underestimated. Vendors publish new vulnerabilities through official channels, and news about the latest breaches or ransomware attacks can be found on Twitter or in different online newspapers.

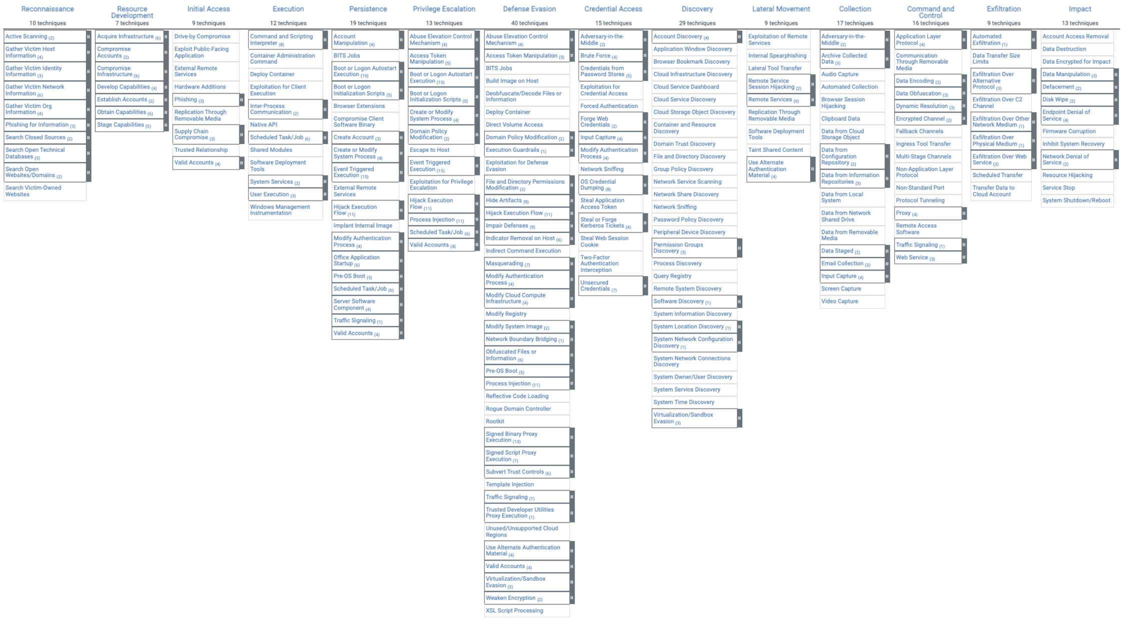
2.5 SOC Security

A SOC works with sensitive data coming from many different companies and therefore must be protected accordingly. Security ranges from a technological point of view such as data encryption, network level protection with firewalls, IPS/IDS, VPN, antivirus, WAF and anti-phishing training, to a physical point of view, such as backup management, secure data storage, incident response plan, data recovery, physical badges or keys controlled access, video surveillance, biometric controls for data centers, access cards, badge-only authorized visitors, and entry control policies.

To see the protection offered by a product, we can use the **MITRE ATT&CK Matrix** [20]. It is globally recognized and utilized as a collection of tactics, techniques and procedures that **threat agents** can use. It is divided into 14 categories or columns that collect techniques used in each phase of a cyber-attack, from the beginning (called reconnaissance) to the end (called impact). Traditional SOCs usually cover the fields from the third column (Initial Access) onwards, because they analyze a company's internal traffic to detect potential anomalies. Figure 2.1 below represents the complete Matrix. More precise pictures will be shown in the chapter 3 of the thesis.

2.6 SOC Challenges

A SOC is a service that an enterprise adds in a second time to protect itself and is usually not part of the original infrastructure. We have seen previously that a SOC comes with many components, and it can be challenging to integrate it in into the company's environment, but necessary to realize the full potential of the SOC. Mutemwa et al. [21] suggest how incident management and other processes can be integrated with the people in the organization, with the creation of a council to monitor the merge and evaluate changes made to the ITenvironment. SOC play-



Reconnaissance
10 techniques
Active Scanning (2)
Gather Victim Host Information (4)
Gather Victim Identity Information (3)
Gather Victim Network Information (6)
Gather Victim Org Information (4)
Phishing for Information (3)
Search Closed Sources (2)
Search Open Technical Databases (5)
Search Open Websites/Domains (2)
Search Victim-Owned Websites

Resource Development
7 techniques
Acquire Infrastructure (6)
Compromise Accounts (2)
Compromise Infrastructure (6)
Develop Capabilities (4)
Establish Accounts (2)
Obtain Capabilities (6)
Stage Capabilities (5)

Initial Access
9 techniques
Drive-by Compromise
Exploit Public-Facing Application
External Remote Services
Hardware Additions
Phishing (3)
Replication Through Removable Media
Supply Chain Compromise (3)
Trusted Relationship
Valid Accounts (4)

Execution
12 techniques
Command and Scripting Interpreter (8)
Container Administration Command
Deploy Container
Exploitation for Client Execution
Inter-Process Communication (2)
Native API
Scheduled Task/Job (6)
Shared Modules
Software Deployment Tools
System Services (2)
User Execution (3)
Windows Management Instrumentation

Persistence
19 techniques
Account Manipulation (4)
BITS Jobs
Boot or Logon Autostart Execution (15)
Boot or Logon Initialization Scripts (5)
Browser Extensions
Compromise Client Software Binary
Create Account (3)
Create or Modify System Process (4)
Event Triggered Execution (15)
External Remote Services
Hijack Execution Flow (11)
Implant Internal Image
Modify Authentication Process (4)
Office Application Startup (6)
Pre-OS Boot (5)
Scheduled Task/Job (6)
Server Software Component (4)
Traffic Signaling (1)
Valid Accounts (4)

Privilege Escalation
13 techniques
Abuse Elevation Control Mechanism (4)
Access Token Manipulation (5)
Boot or Logon Autostart Execution (15)
Boot or Logon Initialization Scripts (5)
Create or Modify System Process (4)
Domain Policy Modification (2)
Escape to Host
Event Triggered Execution (15)
Exploitation for Privilege Escalation
Hijack Execution Flow (11)
Process Injection (11)
Scheduled Task/Job (6)
Valid Accounts (4)

Defense Evasion
40 techniques
Abuse Elevation Control Mechanism (4)
Access Token Manipulation (5)
BITS Jobs
Build Image on Host
Deobfuscate/Decode Files or Information
Deploy Container
Direct Volume Access
Domain Policy Modification (2)
Execution Guardrails (1)
Exploitation for Defense Evasion
File and Directory Permissions Modification (2)
Hide Artifacts (9)
Hijack Execution Flow (11)
Impair Defenses (9)
Indicator Removal on Host (6)
Indirect Command Execution
Masquerading (7)
Modify Authentication Process (4)
Modify Cloud Compute Infrastructure (4)
Modify Registry
Modify System Image (2)
Network Boundary Bridging (1)
Obfuscated Files or Information (6)
Pre-OS Boot (5)
Process Injection (11)
Reflective Code Loading
Rogue Domain Controller
Rootkit
Signed Binary Proxy Execution (13)
Signed Script Proxy Execution (1)
Subvert Trust Controls (6)
Template Injection
Traffic Signaling (1)
Trusted Developer Utilities Proxy Execution (1)
Unused/Unsupported Cloud Regions
Use Alternate Authentication Material (4)
Valid Accounts (4)
Virtualization/Sandbox Evasion (3)
Weaken Encryption (2)
XSL Script Processing

Credential Access
15 techniques
Adversary-in-the-Middle (2)
Brute Force (4)
Credentials from Password Stores (5)
Exploitation for Credential Access
Forced Authentication
Forge Web Credentials (2)
Input Capture (4)
Modify Authentication Process (4)
Network Sniffing
OS Credential Dumping (8)
Steal Application Access Token
Steal or Forge Kerberos Tickets (4)
Steal Web Session Cookie
Two-Factor Authentication Interception
Unsecured Credentials (7)

Discovery
29 techniques
Account Discovery (4)
Application Window Discovery
Browser Bookmark Discovery
Cloud Infrastructure Discovery
Cloud Service Dashboard
Cloud Service Discovery
Cloud Storage Object Discovery
Container and Resource Discovery
Domain Trust Discovery
File and Directory Discovery
Group Policy Discovery
Network Service Scanning
Network Share Discovery
Network Sniffing
Password Policy Discovery
Peripheral Device Discovery
Permission Groups Discovery (3)
Process Discovery
Query Registry
Remote System Discovery
Software Discovery (1)
System Information Discovery
System Location Discovery (1)
System Network Configuration Discovery (1)
System Network Connections Discovery
System Owner/User Discovery
System Service Discovery
System Time Discovery
Virtualization/Sandbox Evasion (3)

Lateral Movement
9 techniques
Exploitation of Remote Services
Internal Spearphishing
Lateral Tool Transfer
Remote Service Session Hijacking (2)
Remote Services (6)
Replication Through Removable Media
Software Deployment Tools
Taint Shared Content
Use Alternate Authentication Material (4)

Collection
17 techniques
Adversary-in-the-Middle (2)
Archive Collected Data (3)
Audio Capture
Automated Collection
Browser Session Hijacking
Clipboard Data
Data from Cloud Storage Object
Data from Configuration Repository (2)
Data from Information Repositories (3)
Data from Local System
Data from Network Shared Drive
Data from Removable Media
Data Staged (2)
Email Collection (3)
Input Capture (4)
Screen Capture
Video Capture

Command and Control
16 techniques
Application Layer Protocol (4)
Communication Through Removable Media
Data Encoding (2)
Data Obfuscation (3)
Dynamic Resolution (3)
Encrypted Channel (2)
Fallback Channels
Ingress Tool Transfer
Multi-Stage Channels
Non-Application Layer Protocol
Non-Standard Port
Protocol Tunneling
Proxy (4)
Remote Access Software
Traffic Signaling (1)
Web Service (3)

Exfiltration
9 techniques
Automated Exfiltration (1)
Data Transfer Size Limits
Exfiltration Over Alternative Protocol (3)
Exfiltration Over C2 Channel
Exfiltration Over Other Network Medium (1)
Exfiltration Over Physical Medium (1)
Exfiltration Over Web Service (2)
Scheduled Transfer
Transfer Data to Cloud Account

Impact
13 techniques
Account Access Removal
Data Destruction
Data Encrypted for Impact
Data Manipulation (3)
Defacement (2)
Disk Wipe (2)
Endpoint Denial of Service (4)
Firmware Corruption
Inhibit System Recovery
Network Denial of Service (2)
Resource Hijacking
Service Stop
System Shutdown/Reboot

Figure 2.1: The Mitre Att&ck Matrix, fully visualized [20]

books are also mentioned as useful guides for informing workers how to respond to and resolve an incident.

One of the biggest problems for analysts working in a SOC is the **large number of alerts** they must handle every single day. SOC operations are far from automated, and alarms generated by detection tools are manually validated by analysts and their cognitive skills and knowledge. Most of the alerts are **false positives** and involve legitimate behavior. Analysts must spend time analyzing them all, but this validation is a **tedious task** that can cause alarm burnout and eventually desensitization [22].

With too many false positives there is a greater probability that an attack will go unnoticed. This is an issue that modern SOCs seek to solve by integrating data with threat intelligence sources and using artificial intelligence algorithms to better classify events and improve the overall security detection process [23]. With alert correlation, elimination of irrelevant alarms and other automated solutions, many companies are trying to reduce the volume of alerts that need to be analyzed by analysts, but for this solution to be efficient, the quality of alerts must be high. Each organization's networks and systems are unique, and detection rules and associated alarms must be customized to fit the monitored environment.

Setting up a SOC is not a one day process. When a company decides to purchase a SOC, there is an initial phase called onboarding where all technologies and business services are added to the SOC platform. All events and logs generated are forwarded to the SOC to enable real time monitoring. As stated by Onwubiko [24] current approaches to onboarding present some difficulties, and can be time consuming, expensive and onerous. This is especially true for on-premises SOC services. By leveraging cloud infrastructure, a new, faster, more efficient and cost optimized onboarding procedure can be performed. In the future SOC processes and technologies will be more automated, and the cloud solution will contribute improve the speed and delivery of this service.

Moreover, like in every modern job, humans can be the weakest link in the chain. SOC analysts are cyber security experts and probably think they cannot be the target of an attack. Some statistics on threat actors [25] point out that phishing is still the most common tactic used by criminals to carry out attacks, and is involved in 70% of data breaches as per 2021. Employee training can increase security awareness and reduce attacks, but in 2020 only 38% of U.S. state and local government employees received general training on ransomware prevention [26]. There is still much work to do to reduce the attack surface exploitable by criminals.

Lastly, a lot of analysts needed to offer a SOC service, especially if it is a 24/7 monitoring solution, so the cost of the salaries can be high.

2.7 SOC Benefits

There are indeed some challenges and points of improvement in this type of technology, but as stated by János et al. [27], a SOC has more advantages than disadvantages and the operation of a SOC is worthwhile and efficient, although a perfect system is impossible and too expensive to achieve.

Some benefits offered by a SOC are as follows [28]:

- Faster incident response times: From a single location the SOC can monitor all devices in real time and intervene promptly when needed. Alerts are displayed in the SIEM with all relevant data related to the incident.
- Reduced cost: Setting up a SOC is not something that comes for free, but the money invested in this type of protection is a few orders of magnitude less than the cost of a ransomware attack.
- Operational efficiencies: With a 24/7 monitoring service, the ratio of detected incidents can approach 100%. Cooperation among SOC team members allow many experts to contribute their own capabilities during analysis.
- Enhanced visibility: Keeping track of all devices connected to the corporate network has become difficult with several employees working from home. SOC facilitates monitoring because all the data is collected and passes through it

SIEM is a critical and vital component for a SOC. It offers a great deal of functionalities and provides a "big picture" of what is happening in the monitored IT infrastructure, conserving the data for audit operations. It can track unauthorized access between different systems, detect violation of internal policies, attempted web applications attacks, malwares, botnet activity, DDoS attacks, intrusion attempts, ransomware, data exfiltration, system bottlenecks and vulnerabilities exploitation.

It also monitors connections and detect suspicious ones, issuing alerts if anomalies occur.

2.8 Types of SOC

There are different Security Operation Centers models and they can be chosen to best suit the needs of the organization, as companies companies have different infrastructures and budgets.

- Internal or dedicated SOC: This is a dedicated area within the company that focuses completely on cybersecurity. It provides continuous, 24/7, 365-day-a-year monitoring and centralized visibility into every network activity. Threat response time is the fastest because employees know the infrastructure they are monitoring. The main disadvantage is cost, which can be prohibitive for small organizations, but totally sustainable by large enterprises and governments.

- Virtual SOC: It is hosted on a web portal in the cloud, so it is easily accessible, very scalable and affordable. It has no physical hardware, so it costs less than a traditional SOC, but it is also less reliable. It usually does not offer 24/7 service and is a mostly reactive approach.

- Hybrid SOC: A hybrid SOC model attempts to bring together the best of both worlds. Alerts are analyzed by internal and external analysts, and costs are reduced compared to a fully dedicated SOC. It is especially useful in cases where a company lacks expertise because it relies on third-party staff.

- SOC As A Service (SOCaaS): This is the most popular option today. Outsourcing the SOC and relying on a company that provides this service to multiple clients offers many advantages. It reduces the high costs of a dedicated SOC by providing a 24/7 monitoring solution and high-quality protection.

> Many SOCaaS integrate CTI modules, multiple levels of analysts, and can intervene when problems arise, with a proactive approach to threats. They also produce regular reports to keep you updated on the security status of your environment.

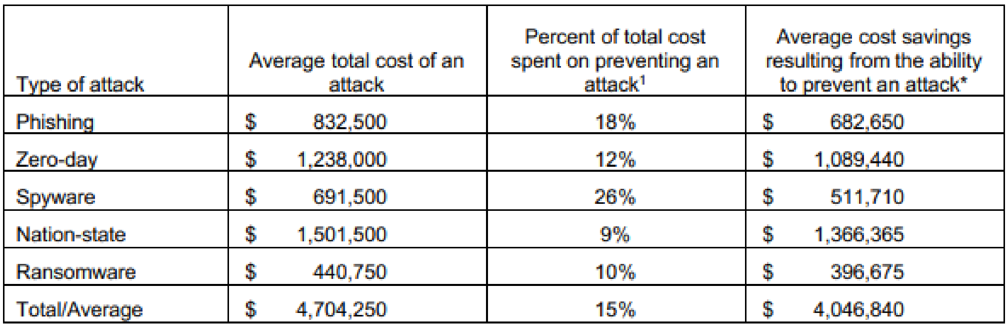
SOCs have become very popular in recent times, and there are many industry-leading SOC vendors. Some of them will now be analyzed to understand the current features offered, the best practices adopted, and their limitations. It is difficult to summarize these services in a table and each would require a chapter of its own. All offer competitive and high-quality solutions. Some are American products, others European or focused on the Italian market. This choice was made because the SOC presented in this thesis is located in Italy and that is where most of the customers come from.

SOCaaS and MTDR services			
Product	24/7/365	CTI	OSINT
ConnectWise	Yes	CRU	Yes
Critical Start	Yes	ZTAP	No
AT&T MDTR	Yes	Alien Labs	No
Arctic Wolf	Yes	Yes	Yes
Ascend Technologies	Yes	No	No
Binary Defense	Yes	Yes	No
HWG	Yes	Yes	No
Reevo	Yes	Yes	No
Wuerth Phoenix	Yes	SATAYO	Yes

Table 2.1: Comparison between some SOC providers

Today's high-end SOCs offer 24/7 service, and most rely on Threat Intelligence enrichment, sometimes with custom solutions. OSINT analysis, however, is still lacking in most products. A standard SOC usually analyzes logs and event information and, through correlation and the use of detection rules, can raise alerts if something suspicious appears. This is sufficient to detect attacks in time if the rules are constantly updated, but the offered protection is always a reaction to an attack, and although it is possible to isolate the network and contain the damage, remediation is applied only after the attack has begun. SOCs are seen only as monitoring and detection tools and do not perform prevention analysis. A great improvement would be to anticipate threat actors and spot them when they are still planning the attack.

Prevention can hugely reduce the cost of an attack, but companies are not investing in it. As reported by the Ponemon Institute in its latest report “The Economic Value of Prevention in the Cybersecurity Lifecycle” [29], prevention is considered the hardest thing to accomplish, and most of the budget is usually spent on detection, containment, recovery and remediation of the attack.



Type of attack	Average total cost of an attack	Percent of total cost spent on preventing an attack[1]	Average cost savings resulting from the ability to prevent an attack*
Phishing	$ 832,500	18%	$ 682,650
Zero-day	$ 1,238,000	12%	$ 1,089,440
Spyware	$ 691,500	26%	$ 511,710
Nation-state	$ 1,501,500	9%	$ 1,366,365
Ransomware	$ 440,750	10%	$ 396,675
Total/Average	$ 4,704,250	15%	$ 4,046,840

Figure 2.2: The table shows that prevention would save up to the 91% of the cost of the attack [29].

A successful attack conducted by a threat agent requires planning and research to find the potential victim, write the necessary exploit and penetrate inside the network, usually with social engineering skills or stolen credentials. A widely known method for differentiating the various phases of an attack is the **Cyber Kill Chain**, developed by Lockheed Martin, an American defense company. It identifies what adversaries must complete to achieve their objective.

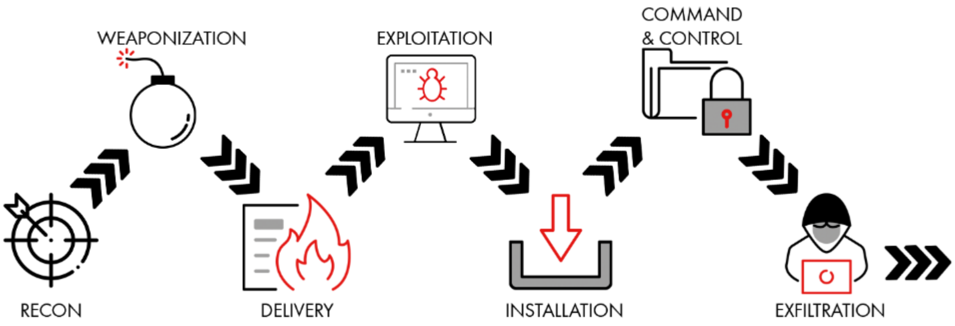


Figure 2.3: The Lockheed Martin Cyber Kill Chain

The first phase of the Kill Chain coincides with the first column of the Mitre Att&ck Framework, the Reconnaissance [30], and is vital for attackers but usually not covered by commercially available SOCs. The SOC developed by Wuerth Phoenix aims to fill this gap and uses and approach that has been defined **Attacker-Centric**, because it gives the ability to cover the Reconnaissance area and recover the attacker's point of view. This can be done by collecting and correlating Open Source Intelligence sources from around the world, just as an attacker would, to find out as much as possible about their victim.

SATAYO (Search All Things About Your Organization) [31] is an OSINT and CTI platform, developed by Wuerth Phoenix's cybersecurity team, that can verify an organization's exposure to possible cybercriminal actions. It is a major component of the SOC and has the purpose of gathering all information publicly available information about a company. This exposure assessment is performed using OSINT, the branch of intelligence that analyzes data coming from public sources. SATAYO shows a company from an attacker's perspective and flags possible attack vectors such as phishing domains, exposed credentials, stolen accounts, vulnerabilities, and many others.

The goal of the platform is to make customers understand the attack surface that

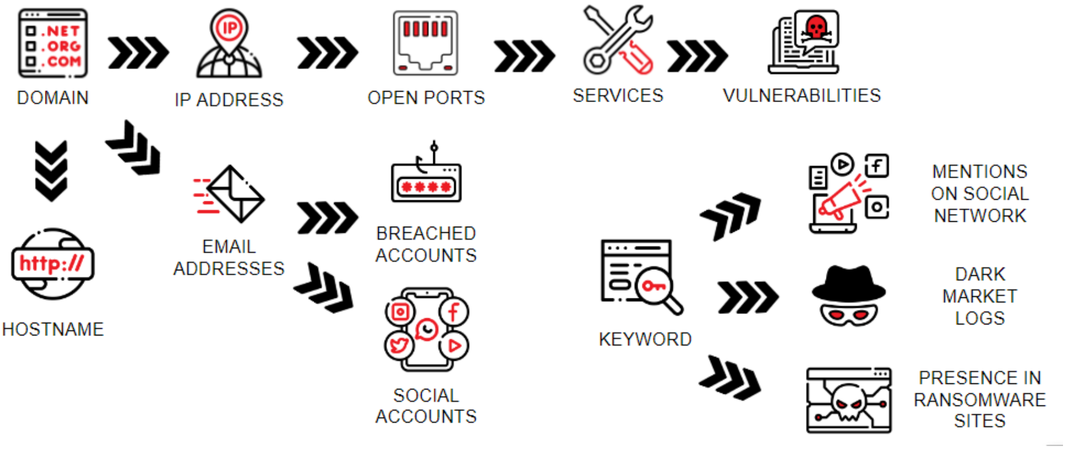
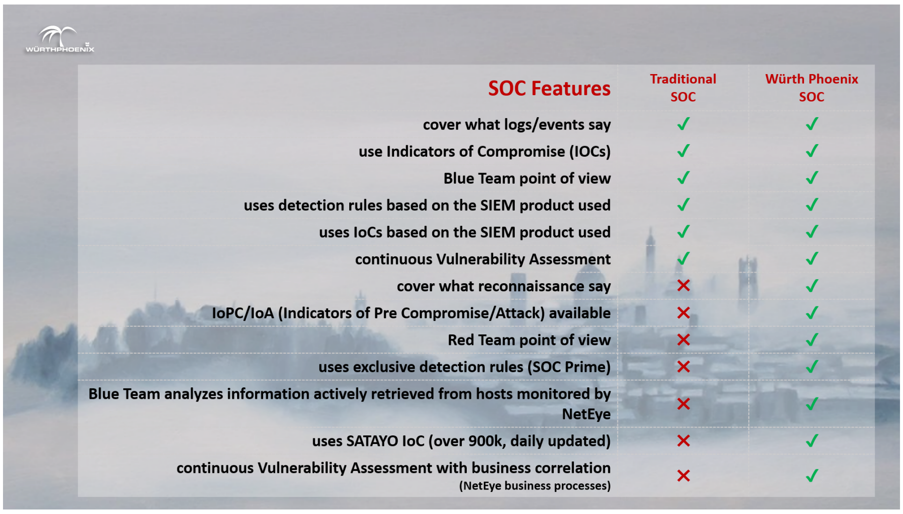


Figure 2.4: A summary of the exposure assessment analysis performed by SATAYO

a third party would have at its disposal should it decide to choose the customer's organization as the target victim [32].

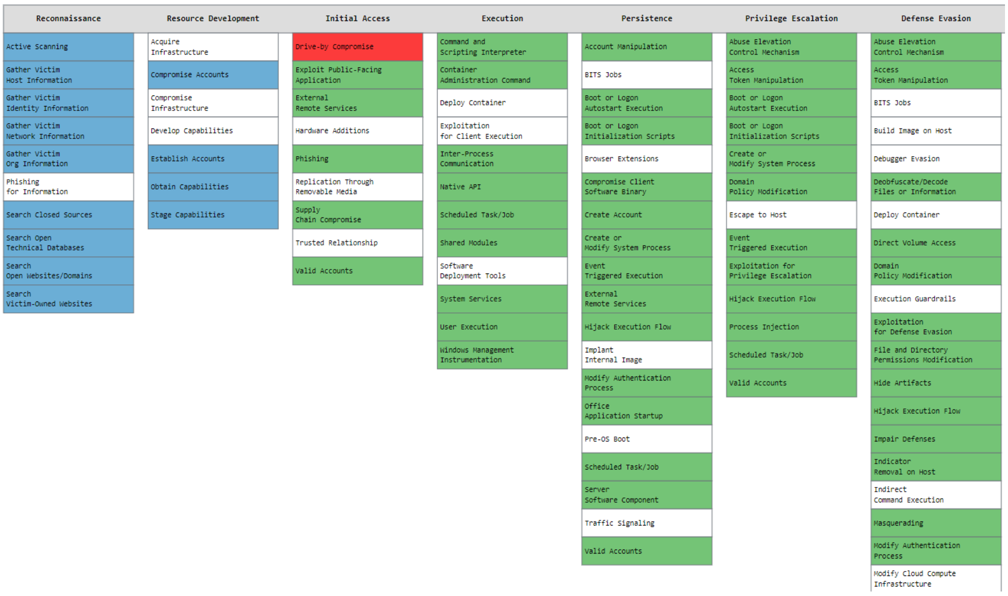
The Attacker Centric SOC offered by Wuerth Phoenix and empowered by SATAYO's analysis is able to cover the first two columns of the Mitre Att&ck matrix, offering a more comprehensive service than the competition. It is distinguished by the following elements:



SOC Features	Traditional SOC	Würth Phoenix SOC
cover what logs/events say	✓	✓
use Indicators of Compromise (IOCs)	✓	✓
Blue Team point of view	✓	✓
uses detection rules based on the SIEM product used	✓	✓
uses IoCs based on the SIEM product used	✓	✓
continuous Vulnerability Assessment	✓	✓
cover what reconnaissance say	✗	✓
IoPC/IoA (Indicators of Pre Compromise/Attack) available	✗	✓
Red Team point of view	✗	✓
uses exclusive detection rules (SOC Prime)	✗	✓
Blue Team analyzes information actively retrieved from hosts monitored by NetEye	✗	✓
uses SATAYO IoC (over 900k, daily updated)	✗	✓
continuous Vulnerability Assessment with business correlation (NetEye business processes)	✗	✓

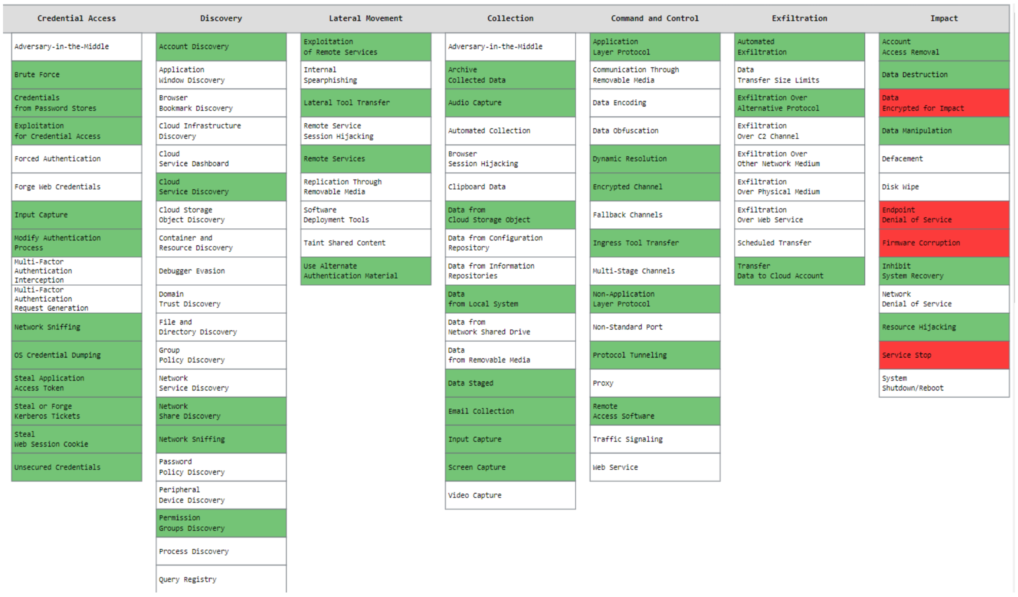
Figure 2.5: Traditional SOC vs Wuerth Phoenix SOC

The following images show the Mitre Att&ck matrix, colored **green** in areas covered by the SIEM and its standard detection rules, **red** in areas covered by custom rules added to the SIEM by our SOC team, and **blue** in areas covered by SATAYO, which covers areas that are not covered at all by other SOCs. The Matrix has been divided into three separate images to better show the content of each cell. The SOC Attacker Centric approach will be studied and discussed in the next chapter.



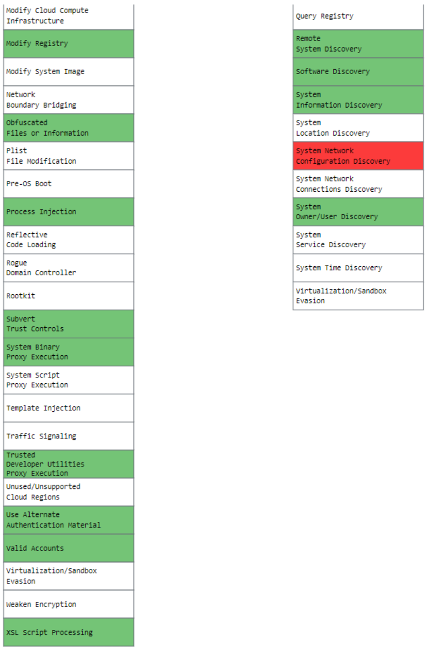
Reconnaissance	Resource Development	Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion
Active Scanning	Acquire Infrastructure	Drive-by Compromise	Command and Scripting Interpreter	Account Manipulation	Abuse Elevation Control Mechanism	Abuse Elevation Control Mechanism
Gather Victim Host Information	Compromise Accounts	Exploit Public-Facing Application	Container Administration Command	BITS Jobs	Access Token Manipulation	Access Token Manipulation
Gather Victim Identity Information	Compromise Infrastructure	External Remote Services	Deploy Container	Boot or Logon Autostart Execution	Boot or Logon Autostart Execution	BITS Jobs
Gather Victim Network Information	Develop Capabilities	Hardware Additions	Exploitation for Client Execution	Boot or Logon Initialization Scripts	Boot or Logon Initialization Scripts	Build Image on Host
Gather Victim Org Information	Establish Accounts	Phishing	Inter-Process Communication	Browser Extensions	Create or Modify System Process	Debugger Evasion
Phishing for Information	Obtain Capabilities	Replication Through Removable Media	Native API	Compromise Client Software Binary	Domain Policy Modification	Deobfuscate/Decode Files or Information
Search Closed Sources	Stage Capabilities	Supply Chain Compromise	Scheduled Task/Job	Create Account	Escape to Host	Deploy Container
Search Open Technical Databases		Trusted Relationship	Shared Modules	Create or Modify System Process	Event Triggered Execution	Direct Volume Access
Search Open Websites/Domains		Valid Accounts	Software Deployment Tools	Event Triggered Execution	Exploitation for Privilege Escalation	Domain Policy Modification
Search Victim-Owned Websites			System Services	External Remote Services	Hijack Execution Flow	Execution Guardrails
			User Execution	Hijack Execution Flow	Process Injection	Exploitation for Defense Evasion
			Windows Management Instrumentation	Implant Internal Image	Scheduled Task/Job	File and Directory Permissions Modification
				Modify Authentication Process	Valid Accounts	Hide Artifacts
				Office Application Startup		Hijack Execution Flow
				Pre-OS Boot		Impair Defenses
				Scheduled Task/Job		Indicator Removal on Host
				Server Software Component		Indirect Command Execution
				Traffic Signaling		Masquerading
				Valid Accounts		Modify Authentication Process
						Modify Cloud Compute Infrastructure

Figure 2.6: The first 7 columns of the Mitre Att&ck Matrix



Credential Access	Discovery	Lateral Movement	Collection	Command and Control	Exfiltration	Impact
Adversary-in-the-Middle	Account Discovery	Exploitation of Remote Services	Adversary-in-the-Middle	Application Layer Protocol	Automated Exfiltration	Account Access Removal
Brute Force	Application Window Discovery	Internal Spearphishing	Archive Collected Data	Communication Through Removable Media	Data Transfer Size Limits	Data Destruction
Credentials from Password Stores	Browser Bookmark Discovery	Lateral Tool Transfer	Audio Capture	Data Encoding	Exfiltration Over Alternative Protocol	Data Encrypted for Impact
Exploitation for Credential Access	Cloud Infrastructure Discovery	Remote Service Session Hijacking	Automated Collection	Data Obfuscation	Exfiltration Over C2 Channel	Data Manipulation
Forced Authentication	Cloud Service Dashboard	Remote Services	Browser Session Hijacking	Dynamic Resolution	Exfiltration Over Other Network Medium	Defacement
Forge Web Credentials	Cloud Service Discovery	Replication Through Removable Media	Clipboard Data	Encrypted Channel	Exfiltration Over Physical Medium	Disk Wipe
Input Capture	Cloud Storage Object Discovery	Software Deployment Tools	Data from Cloud Storage Object	Fallback Channels	Exfiltration Over Web Service	Endpoint Denial of Service
Modify Authentication Process	Container and Resource Discovery	Taint Shared Content	Data from Configuration Repository	Ingress Tool Transfer	Scheduled Transfer	Firmware Corruption
Multi-Factor Authentication Interception	Debugger Evasion	Use Alternate Authentication Material	Data from Information Repositories	Multi-Stage Channels	Transfer Data to Cloud Account	Inhibit System Recovery
Multi-Factor Authentication Request Generation	Domain Trust Discovery		Data from Local System	Non-Application Layer Protocol		Network Denial of Service
Network Sniffing	File and Directory Discovery		Data from Network Shared Drive	Non-Standard Port		Resource Hijacking
OS Credential Dumping	Group Policy Discovery		Data from Removable Media	Protocol Tunneling		Service Stop
Steal Application Access Token	Network Service Discovery		Data Staged	Proxy		System Shutdown/Reboot
Steal or Forge Kerberos Tickets	Network Share Discovery		Email Collection	Remote Access Software		
Steal Web Session Cookie	Network Sniffing		Input Capture	Traffic Signaling		
Unsecured Credentials	Password Policy Discovery		Screen Capture	Web Service		
	Peripheral Device Discovery		Video Capture			
	Permission Groups Discovery					
	Process Discovery					
	Query Registry					

Figure 2.7: The last 7 columns of the Mitre Att&ck Matrix



Modify Cloud Compute Infrastructure	Query Registry
Modify Registry	Remote System Discovery
Modify System Image	Software Discovery
Network Boundary Bridging	System Information Discovery
Obfuscated Files or Information	System Location Discovery
Plist File Modification	System Network Configuration Discovery
Pre-OS Boot	System Network Connections Discovery
Process Injection	System Owner/User Discovery
Reflective Code Loading	System Service Discovery
Rogue Domain Controller	System Time Discovery
Rootkit	Virtualization/Sandbox Evasion
Subvert Trust Controls	
System Binary Proxy Execution	
System Script Proxy Execution	
Template Injection	
Traffic Signaling	
Trusted Developer Utilities Proxy Execution	
Unused/Unsupported Cloud Regions	
Use Alternate Authentication Material	
Valid Accounts	
Virtualization/Sandbox Evasion	
Weaken Encryption	
XSL Script Processing	

Figure 2.8: The lower part that is not visible above (on the left the column No. 7 - Defense Evasion and on the right the column No. 9 - Discovery)

3 SOC-AC Structure and Implementation

3.1 SIEM

A SIEM is a system that collects, process, stores, and queries data. It analyzes events and logs coming from every IT component of the company and gives the ability to detect and block a potential attack. Not all logs are relevant and must be tuned accordingly to avoid unwanted noise and analyze unnecessary traffic useless for detection purposes.

> “To stay ahead of the attackers your organizations IT, security and architecture specialists should think like attackers when it comes to creating a comprehensive and layered cybersecurity solution.” [33]

The SIEM used by Wuerth-Phoenix is the Elastic Stack, also widely known as ELK Stack. The name comes from the three main components of the SIEM, **Elasticsearch**, **Logstash** and **Kibana** [34].

Elasticsearch is the core component of the Stack. It is a fast, scalable, and flexible **search engine** that allows all data to be stored, searched, and analyzed. It is a distributed DB written in Java that runs in **cluster** mode, where multiple instances called **nodes** communicate and interact with each other. Data is saved in **indexes**, which are themselves divided into **shards**, distributed among the nodes, and replicated for greater efficiency. There is always a **primary** shard and a **replica** shard that serves as a backup in case of a node is malfunctioning. The usage of the shards is done by the cluster autonomously and is completely transparent to the end user.

Logstash is a server side **data processing pipeline** that ingests data from multiple sources, manipulates it, and returns an output that will be collected by Elasticsearch. It is divided in three section called **plugins**. The input plugin, the filter plugin (where the data is edited and enriched), and the output plugin that sends the final data to the DB.

Kibana is the **visualization layer** of the stack and allows the user to create charts, diagrams, graphs, dashboards, canvases, and to model and shape the data in different and customized ways. It is the Graphical User Interface (GUI) of the ELK Stack. Some of its features will be explained in the sections 3.1.2 and 3.1.3.

The last component of the stack are the **Beats**, a family of **data shippers** that collect logs, metric and network data, windows events, audit data and more. Their purpose is to gather data from thousands of machines connected to the network and send those data back to the ELK stack. They communicate directly with Logstash or Elastic. In the latest version, the Beats are grouped together into the so called Elastic Agent to simplify their use.

The architecture of the Elastic Stack can considerably vary from implementation to implementation. A standard architecture is shown in the image 3.1.

In contrast, Wuerth Phoenix's architecture is based on **NetEye**, a monitoring system for IT environments developed in-house. The Elastic SIEM is integrated on top of NetEye and is available as a module for customers interested in the SOC service. The schema is shown in the image 3.2.

NetEye is as a monitoring system and and IT management software. It can monitor and control the performance of company's IT assets and services, increas-

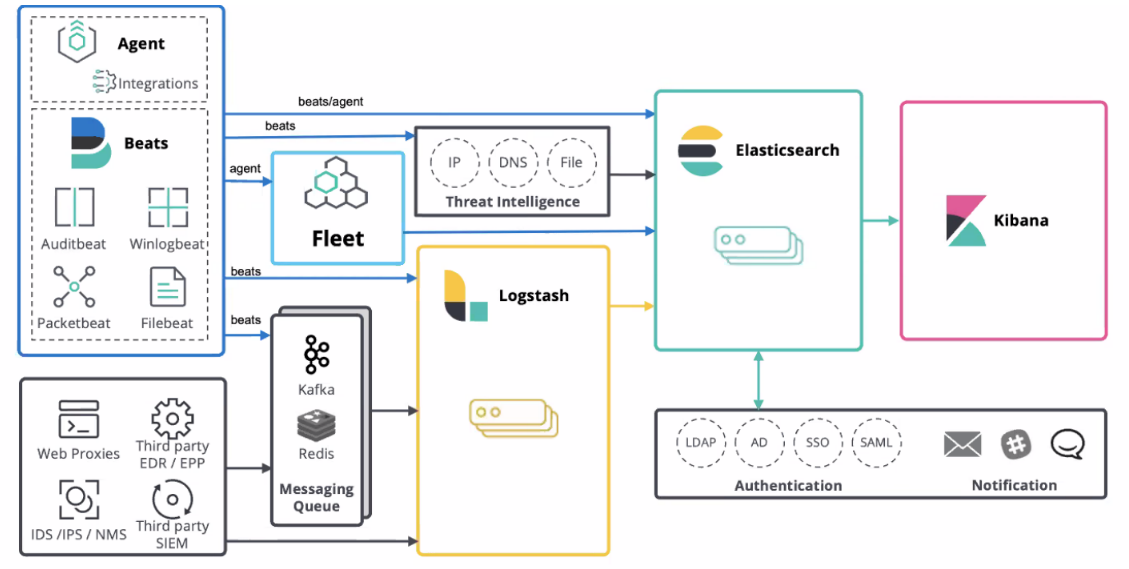


Figure 3.1: The Elastic Stack architecture, taken as a reference

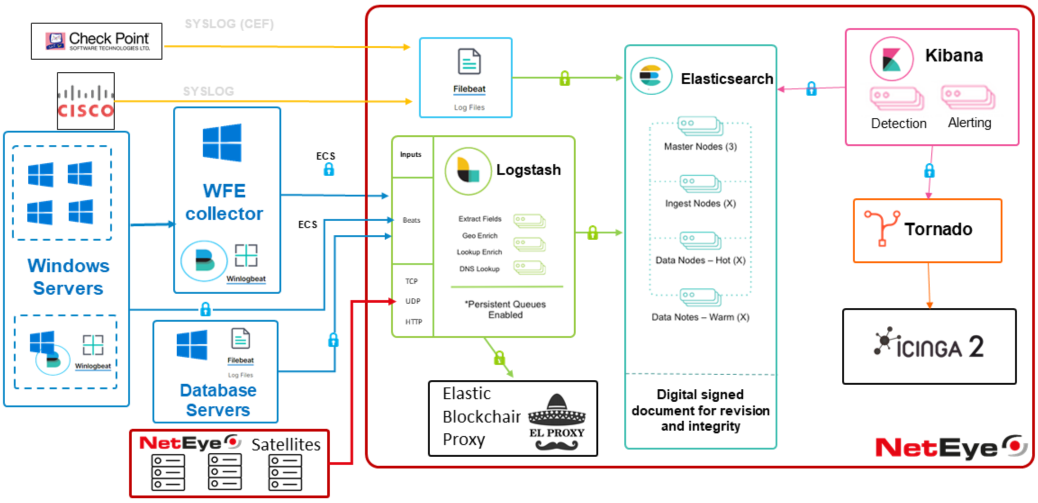
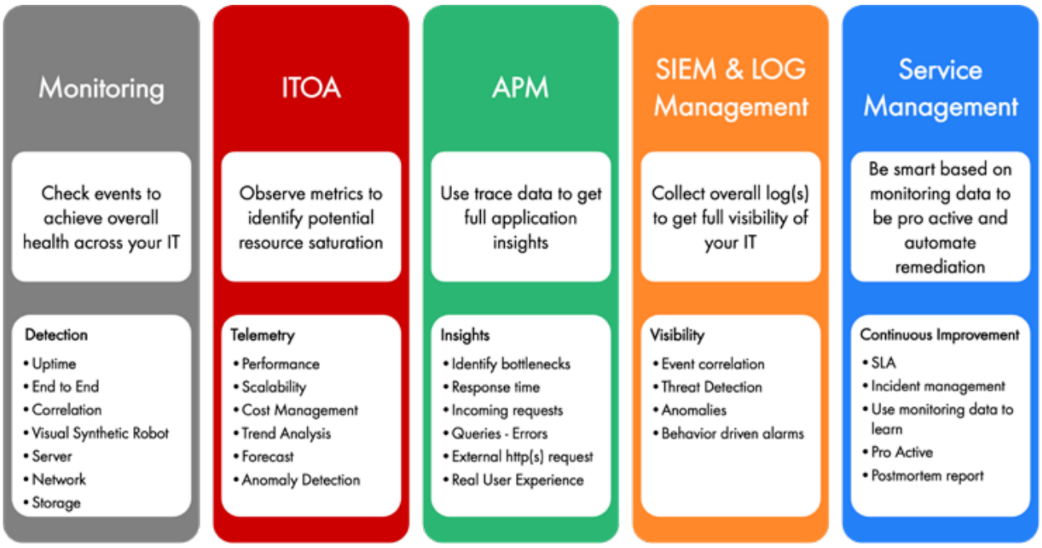


Figure 3.2: The Log Management architecture used at Wuerth Phoenix

ing productivity. It can proactively detect components that cause disservices and highlights the health status of the IT ecosystem. Icinga, a network monitoring application, is integrated in NetEye, and with it we can schedule multiple checks that inform us about the status of the monitored components. When a computer has a problem with some detection rules, for example, NetEye informs us via a Grafana dashboard. Grafana is an open source data visualization platform, used for the creation of graphs and charts thanks to its flexibility. It can monitor infrastructure and improve operational efficiency.



Monitoring

Check events to achieve overall health across your IT

Detection
- Uptime
- End to End
- Correlation
- Visual Synthetic Robot
- Server
- Network
- Storage

ITOA

Observe metrics to identify potential resource saturation

Telemetry
- Performance
- Scalability
- Cost Management
- Trend Analysis
- Forecast
- Anomaly Detection

APM

Use trace data to get full application insights

Insights
- Identify bottlenecks
- Response time
- Incoming requests
- Queries - Errors
- External http(s) request
- Real User Experience

SIEM & LOG Management

Collect overall log(s) to get full visibility of your IT

Visibility
- Event correlation
- Threat Detection
- Anomalies
- Behavior driven alarms

Service Management

Be smart based on monitoring data to be pro active and automate remediation

Continuous Improvement
- SLA
- Incident management
- Use monitoring data to learn
- Pro Active
- Postmortem report

Figure 3.3: All the functionalities offered by NetEye

3.1.1 Index Management

Within the Elastic Stack, data are organized using the following logic. Each log is called **document** and has a unique ID inside the DB. Documents are saved in logical entities called **indexes** and every index is divided into **shards**. The latter are the tiniest piece of Elastic, where the data is written to disk. Elastic takes care of the management and distribution of the shards. There are primary shards and replica shards to improve performance and increase resilience. A shard is divided into read only portion of data called **segments** that can be queried.

Users can manage Elastic indexes. They can have three states, colored like a traffic light.

- Green: means that all primary shards and replica shards are available for that index.
- Yellow: means that at least one replica shard is not available. It may exist but is not reachable by Elastic.
- Red: means that at least one primary shard is not reachable, and the data is not searchable.

There are different types of logs, and they are associated with one or more log sources. A single host can have multiple log sources, for example, a Windows server has system, security, and application events. These sources help us determine the type and the structure of the log.

Elastic has defined a standard to normalize all the fields of a document and have a common vocabulary. It is called **Elastic Common Schema** (ECS) and defines a common set of fields to be used when storing event data in Elasticsearch, to be aligned across different vendors and products. The common schema helps correlate and use data from various sources. It is very important that data coming from different log sources be consistent between each other. If the same field is interpreted as a number once and as a string another time, there will be a conflict in the DB.

Everything present in Elastic is somehow traced back to one of the fields chosen by ECS to recognize it. There are some general guidelines to follow, for example every document must have a **timestamp**. Then there are four fields that define the typology of data and are called **categorization fields**. Their purpose is to group similar events from multiple data sources and are event.category, event.type, event.outcome and event.kind. These special fields are managed by Elastic and it is not possible to assign custom values to them. For instance, there are only three values allowed for event.outcome and they can be failure, success or unknown.

Logs are stored in indexes, and an index is considered to be a database with **settings** and **mappings**. The settings are specific for each index and tell Elastic where to store the data, how long to keep it, how many shards to use, etc. Settings can be static and no longer editable if they are defined during index creation (such as the number of shards) or dynamic if they can vary during the life cycle of the index. The mappings define how a document and the fields it contains are stored and indexed. Each field is mapped to its own data type. ECS forces the mapping procedure so that the data is congruent with the best practices and general guidelines.

Data cannot be kept forever and after a certain amount of time the old indexes must be discarded. Within Elastic there is an **Index Lifecycle Management** (ILM) that defines the rotation of old indexes. It is highly customizable according to business policies and logic. There are five different phases to manage the life of an index: hot, warm, cold, frozen, and delete.

- Hot: in this phase the data is fresh, up to date, and it is easy and fast to make queries on it. This is the only mandatory phase, and it is possible to specify several parameters such as the age or size of the index, the maximum number of documents or dimension of the shard, that once satisfied will move the index to the next phase.

- Warm: the data here is no longer updated but can still be searched. There are some optimizations like read only indexes and reduced segment size to allow better queries.

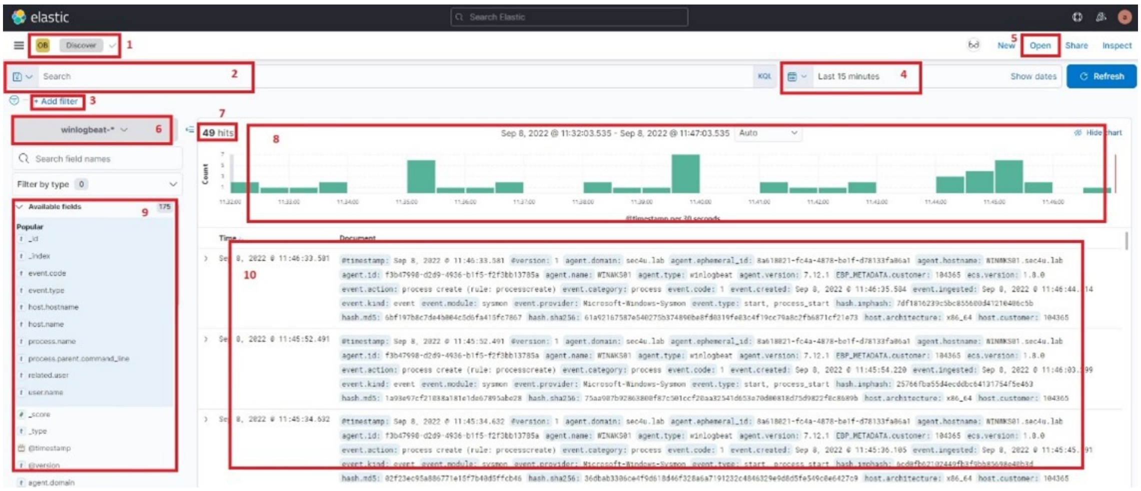
- Cold: queries in this stage are very slow because the data are not needed immediately. Used when the data needs to be kept for multiple years, there are some storage media specifically made to save the data present in this phase.

- Frozen: this phase was used to increase performance of unused indexes, but was removed in Elastic version 8 because heap memory usage was improved and frozen indexes were no longer useful.
- Delete: the data here will be deleted after a specified number of days. It is possible to create a snapshot and save the data to external storage before the cancellation.

In Elastic there is the possibility of creating multiple spaces and associating each of them with one client. Unfortunately, **multitenancy** is not fully supported but creating an index for each tenant is a simple solution for isolation. In any case, admin users have visibility to all data, no matter what space they are in. It would have been more logical to see data related to space currently viewed. To achieve this, we have created index patterns that match the specific data for each tenant, so that even when using administrative accounts, dashboards and charts are loaded correctly and show data only for the current customer.

3.1.2 Discover

The Discover is the section of the SIEM where it is possible to view all the logs present in Elastic and query the different documents. This function is used daily by SOC analysts to investigate incidents, usually after an alert has been triggered. The aspect looks as follow:



```
Sep 8, 2022 @ 11:46:33.581   @timestamp: Sep 8, 2022 @ 11:46:33.581 @version: 1 agent.domain: sec4u.lab agent.ephemeral_id: 8a618821-fc4a-4878-be1f-d78133fa86a1 agent.hostname: WINAKS01.sec4u.lab
                             agent.id: f3b47998-d2d9-4936-b1f5-f2f3bb13785a agent.name: WINAKS01 agent.type: winlogbeat agent.version: 7.12.1 EBP_METADATA.customer: 104365 ecs.version: 1.8.0
                             event.action: process create (rule: processcreate) event.category: process event.code: 1 event.created: Sep 8, 2022 @ 11:46:35.504 event.ingested: Sep 8, 2022 @ 11:46:44.114
                             event.kind: event event.module: sysmon event.provider: Microsoft-Windows-Sysmon event.type: start, process_start hash.imphash: 7df1816239c5bc855680d41210486c5b
                             hash.md5: 6bf197b8c7de4b004c5d6fa415fc7867 hash.sha256: 61a92167587e540275b374890be8fd0319fe83c4f19cc79a8c2fb6871cf21e73 host.architecture: x86_64 host.customer: 104365
Sep 8, 2022 @ 11:45:52.491   @timestamp: Sep 8, 2022 @ 11:45:52.491 @version: 1 agent.domain: sec4u.lab agent.ephemeral_id: 8a618821-fc4a-4878-be1f-d78133fa86a1 agent.hostname: WINAKS01.sec4u.lab
                             agent.id: f3b47998-d2d9-4936-b1f5-f2f3bb13785a agent.name: WINAKS01 agent.type: winlogbeat agent.version: 7.12.1 EBP_METADATA.customer: 104365 ecs.version: 1.8.0
                             event.action: process create (rule: processcreate) event.category: process event.code: 1 event.created: Sep 8, 2022 @ 11:45:54.220 event.ingested: Sep 8, 2022 @ 11:46:03.199
                             event.kind: event event.module: sysmon event.provider: Microsoft-Windows-Sysmon event.type: start, process_start hash.imphash: 25766fba55d4ecddbc64131754f5e463
                             hash.md5: 1a93e97cf21838a181e1de67895abe28 hash.sha256: 75aa987b92863808f87c501ccf28aa32541d653a70d00818d75d9822f8c8689b host.architecture: x86_64 host.customer: 104365
Sep 8, 2022 @ 11:45:34.632   @timestamp: Sep 8, 2022 @ 11:45:34.632 @version: 1 agent.domain: sec4u.lab agent.ephemeral_id: 8a618821-fc4a-4878-be1f-d78133fa86a1 agent.hostname: WINAKS01.sec4u.lab
                             agent.id: f3b47998-d2d9-4936-b1f5-f2f3bb13785a agent.name: WINAKS01 agent.type: winlogbeat agent.version: 7.12.1 EBP_METADATA.customer: 104365 ecs.version: 1.8.0
                             event.action: process create (rule: processcreate) event.category: process event.code: 1 event.created: Sep 8, 2022 @ 11:45:36.105 event.ingested: Sep 8, 2022 @ 11:45:45.191
                             event.kind: event event.module: sysmon event.provider: Microsoft-Windows-Sysmon event.type: start, process_start hash.imphash: 6cd8fb62102449fb3f9bb85698e48b3d
                             hash.md5: 02f23ec95a886771e15f7b40d5ffcb46 hash.sha256: 36dbab3306ce4f9d618d46f328a6a7191232c4846329e9d8d5fe549c0e6427c9 host.architecture: x86_64 host.customer: 104365
```

Figure 3.4: Elastic Discover, divided into sections for easier explanation

1. The name of your space and the section you are in

2. The search bar through which it will be possible to make queries and filter data

3. Additional filters that can be applied to the query or used individually to filter the data

4. The filter to select the time range in which the results will be shown

5. The ability to open previously saved search queries

6. The Elastic index where the data to be displayed are contained

7. The number of logs (also called documents) found within the chosen time range and query written

8. The chart of documents that match the query sorted by time

9. The fields of the documents that were found

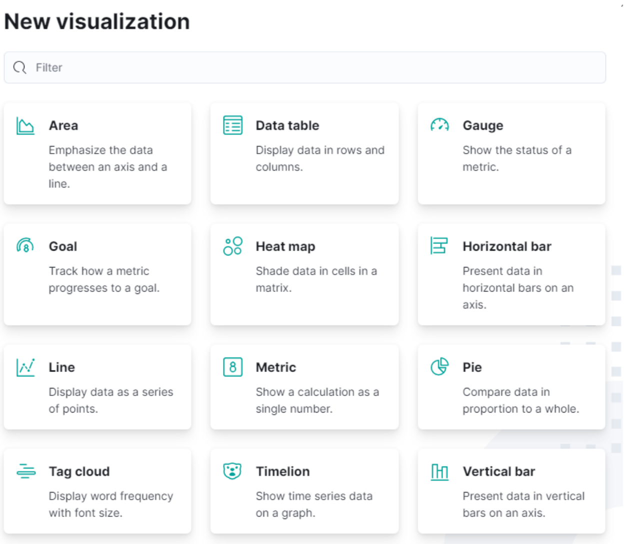
10. The complete list of documents, sorted from newest to oldest. You can expand each individual document and view all the fields present inside it

The search bar at the top is where we can type queries and search for something specific. If we leave it blank, Elastic will show us all the logs present in the given time interval. The latter can be specified with the time filter on the right. Additional filters can be applied to the query, to filter out certain values. In the center there is a histogram showing us the frequency of the documents with a counter for the number of logs displayed. Below the graph we can find the list of documents, each with its own fields. By default, only the last 500 are listed in the table. Every document can be expanded to see the fields of the log in detail. Values can be of different types, such as timestamp, text, keyword, IP, and geographic values.

3.1.3 Dashboard and Visualization

Dashboards are a collection of visualization or saved searches. They are useful because they are interactive and adapt to our needs. We can click on an element of a chart and drill down into the data. Dashboards are used daily in the SOC to receive real time updates on alerts, IOCs, events and machines status, and so on. Visualizations can be numerous and can be created from one of the options shown in the image 3.5.

Elastic provides a tool called Lens that offers the ability to create most of the visualizations shown above in a guided and easy to use manner, perfect for beginners. There is also an option to create more complicate graphs, with the Vega or Timelion



New visualization

Filter

Area
Emphasize the data between an axis and a line.

Data table
Display data in rows and columns.

Gauge
Show the status of a metric.

Goal
Track how a metric progresses to a goal.

Heat map
Shade data in cells in a matrix.

Horizontal bar
Present data in horizontal bars on an axis.

Line
Display data as a series of points.

Metric
Show a calculation as a single number.

Pie
Compare data in proportion to a whole.

Tag cloud
Display word frequency with font size.

Timelion
Show time series data on a graph.

Vertical bar
Present data in vertical bars on an axis.

Figure 3.5: The list of possible visualization

language, where you program every single aspect of the graph, axes and values, and the degree of customization is very high. Some examples of dashboards are shown in images 3.6 and 3.7 to give the reader an idea of what is possible with the SIEM. Everything that is contained in the indexes can be displayed graphically.

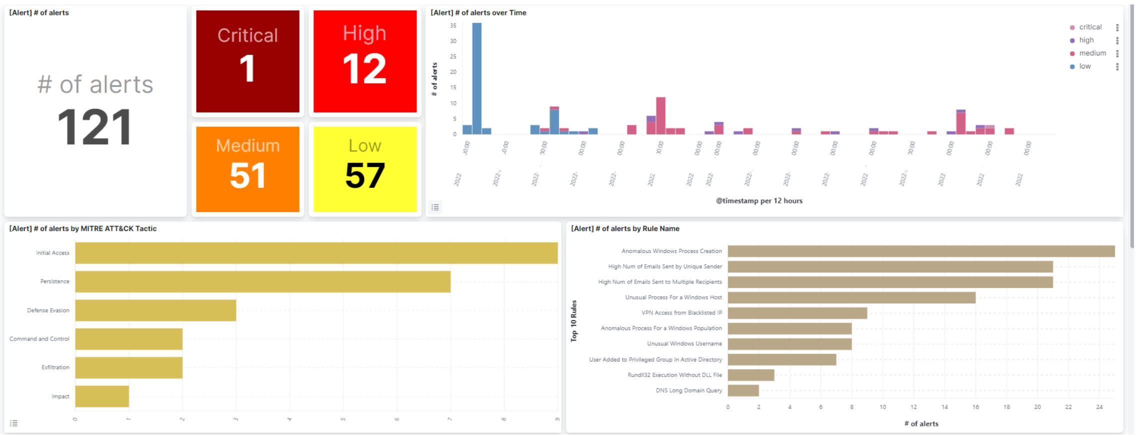

[Alert] # of alerts
# of alerts
121
Critical
1
High
12
Medium
51
Low
57
[Alert] # of alerts over Time
# of alerts
35
30
25
20
15
10
5
0
critical
high
medium
low
@timestamp per 12 hours
[Alert] # of alerts by MITRE ATT&CK Tactic
Initial Access
Persistence
Defense Evasion
Command and Control
Exfiltration
Impact
0
1
2
3
4
5
6
7
8
9
[Alert] # of alerts by Rule Name
Top 10 Rules
Anomalous Windows Process Creation
High Num of Emails Sent by Unique Sender
High Num of Emails Sent to Multiple Recipients
Unusual Process For a Windows Host
VPN Access from Blacklisted IP
Anomalous Process For a Windows Population
Unusual Windows Username
User Added to Privileged Group in Active Directory
Rundll32 Execution Without DLL File
DNS Long Domain Query
0
2
4
6
8
10
12
14
16
18
20
22
24
# of alerts

Figure 3.6: A dashboard with details on alerts triggered by detection rules

Dashboards are used by SOC analysts and are also provided to customers so they can see what is happening with their data. Each customer is assigned a read-only

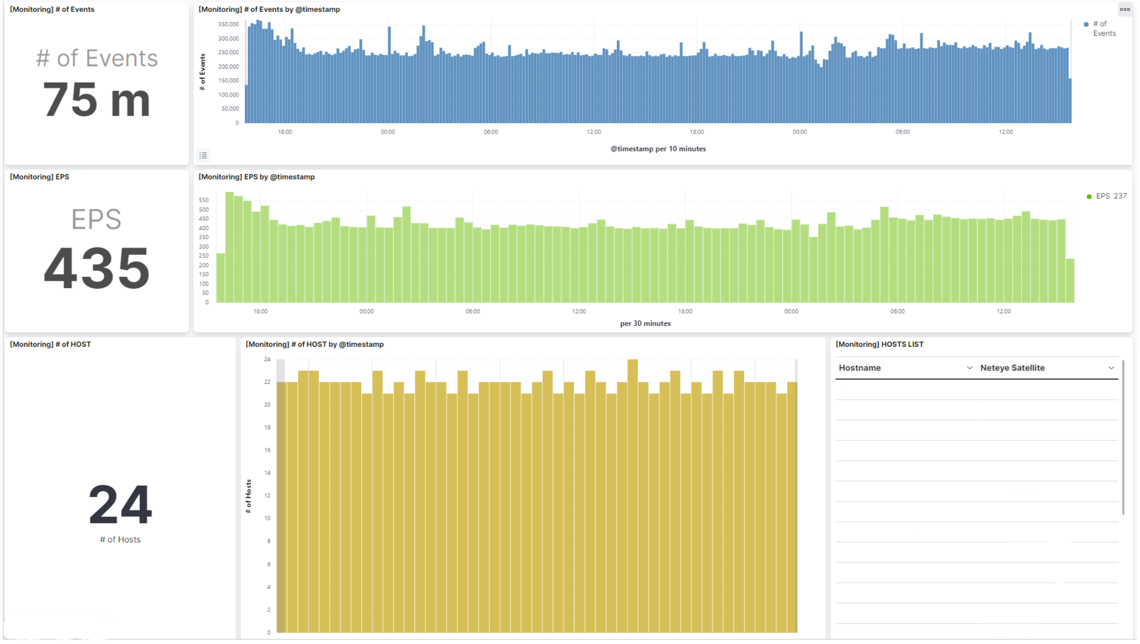


Figure 3.7: A dashboard with information on EPS (events per second) ingested by Elastic

account to access their Elastic space in the SIEM and from which they can inspect their data in detail if they wish.

3.1.4 Canvas

This section allows us to create poster, slides, presentations, and pdf documents. We use it to deliver high-level, customized monthly reports to customers that are updated in real-time with data present in the Elastic environment. This section is further explored in the last chapter of this thesis.

3.1.5 OSINT Search and Enrichment with SATAYO

OSINT stands for **Open Source Intelligence** and is the branch of intelligence that collects and analyzes data from sources that are publicly available. It is the opposite of secret and covert sources, those about which you have no visibility or which require authorization of some kind to access. In the internet environment OSINT defines all the tools that allow one to find information about a person, a company, or more generally, an entity. The goal of OSINT is to investigate a specific topic. OSINT activity is carried out daily by both attackers and defenders. Attackers attempt to gain knowledge about a target and evaluate a company's attack surface in order to efficiently execute an aimed attack. Defenders perform the same activity with the purpose of anticipating attacks by continuously monitoring the exposure assessment and receiving notifications if anything changes. OSINT software can combine data from multiple sources to produce a more informative output. It is important for companies to be aware of what kind of data about them is available on the Internet in order to take preventive protective measures. OSINT has many advantages. For example, because OSINT information are public, it is not illegal to search for it. Even the United States Department of Justice has published a document containing guidelines for collecting open source data and purchasing it from illicit sources [35]. It is also a cost-effective technique because little amount of money is spent to retrieve a lot of useful information. Many open source projects already exist and can be used for free. The OSINT framework [36] or the awesome-osint repository [37] contain a collection of tools to introduce people to OSINT.

Indicator of Compromise (IoCs) and **Indicator of Attack** (IoAs) are indicators that helps security experts to identify malicious activities [38]. IoCs are **reactive controls** such as malicious IP addresses used in a cyber-attack, malware signatures, malicious DNS, and many others. If the same IP is used in another attack, it is detected from the IoC list and blocked. They are not so efficient because they change often, and the same IP can be used the next day by a totally legitimate company. IoAs, on the other hand, are the new generation of indicators as they try to offer **proactive control**. They do not detect the malware used in the attack but the attacker's behavior. They monitor a series of action that a threat actor must follow to successfully execute the attack and are more reliable than IoCs which can be obsolete very soon and must be constantly updated.

Starting from the organization's domain, SATAYO performs daily scans of the surface, deep and dark web to capture evidence related to the customer and enrich the data analyzed by the SOC with an updated list of IOCs. This OSINT research is related to the reconnaissance activities described by the Mitre Att&ck that threat agents (i.e., cybercriminals) usually perform before carrying out an attack. Various sources are monitored, such as ransomware gang websites, twitter accounts and telegram channels. The data collected by SATAYO is used to populate an Elastic index with IOCs and keep it updated daily. Data from this index is used to enrich logs and efficiently evaluate communications coming from malicious IPs.

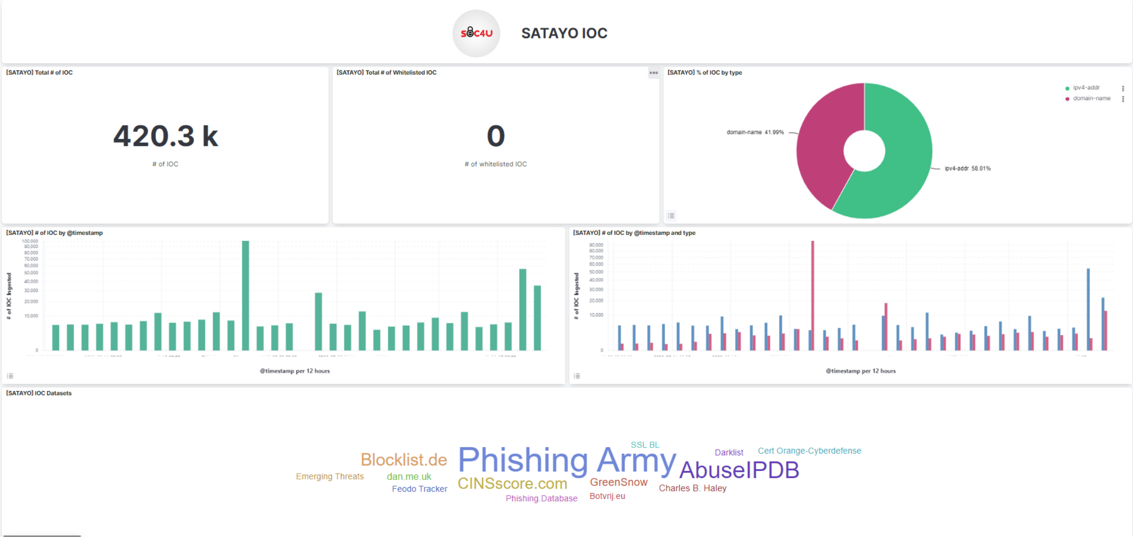
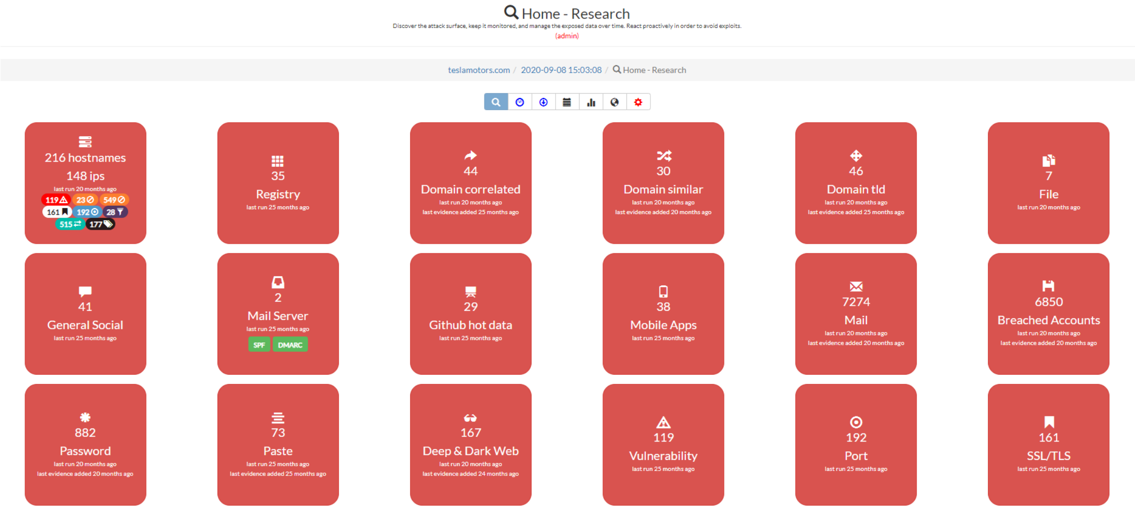


Figure 3.8: A dashboard showing the IOCs collected by SATAYO

IOCs are only a small part of the elements collected by SATAYO. The platform assesses a company's exposure and tries to retrieve as much as data available online related to the target. The information collected is classified into different items, as shown in the figure 3.9.



Home - Research

Discover the attack surface, keep it monitored, and manage the exposed data over time. React proactively in order to avoid exploits.

(admin)

teslamotors.com / 2020-09-08 15:03:08 / Home - Research

Figure 3.9: An example of the findings for the domain teslamotors.com

The items collected by SATAYO may vary from organization to organization. Some companies are larger than others and leave a wider fingerprint on the Internet. The initial vectors that attackers use to penetrate the network are many, but compromised credentials and phishing account for more than one third of the total breaches. SATAYO collects information of breached accounts and suspicious websites that may be used for phishing and can detect a potential attack before it happens. All the different evidences gathered will now be briefly explained.

Hostname: It is one of the starting points for SATAYO's analysis. This page shows the hostnames found for the selected domain. The list of hostnames present within the infrastructure is compiled and all their IP addresses are identified.

Vulnerability: This page shows the existence of vulnerabilities, identified by a CVE (Common Vulnerabilities and Exposures) number and a CVSS (Common Vulnerability Scoring System) score, on resources exposed and related to the domain.

Blacklist host: It shows the presence of hostnames within blacklists. Situation like these can compromise the provision of the service (the connection to the blacklisted IP may be refused by security policies) and ruin the reputation.

Unencrypted protocols: This page shows the exposure of services over protocols that transmit information in cleartext. This cleartext protocols may simplify the activity of network sniffing and consequent capture of confidential information.

Interesting services: it shows exposed services that might be of interest to a malicious user. Web server ports: it shows exposed encrypted and unencrypted web servers. Wayback machine: this page shows previous snapshots of websites. Old versions of sites may contain confidential information.

Technologies: This page shows the technologies used within the exposed web resources.

robots.txt: This page shows the details of the robots.txt file exposed by the website. Very often within this file, attackers can find information and paths access to restricted areas.

HTTP method: This page shows the exposed http methods of the website hosted on the analyzed IP. In case of vulnerabilities attackers can exploit them.

SSL/TLS: This page shows the robustness of the TLS / SSL protocols of various web sites of the domain. Checks performed may return evidence of expired SSL certificates or the use of obsolete and insecure cryptographic algorithms.

Registry: This page shows the subnet blocks where the found IP address reside. The records are managed by various Regional Internet Registries (RIRs). If some IP blocks are managed directly by the analyzed organization, the addresses inside are scanned to see if there are other resolvable hostnames.

Domain suspicious: it shows domains classified as suspicious because they contain the company's name inside.

Domain correlated: it shows domains related to the main one. The correlation may result from similar elements present in the DNS record.

Domain phishing: it shows known malicious domains or URLs that are currently performing phishing activities.

Domain similar: it shows all registered domains similar to the main one. These domains could be used for phishing activities. If a domain was registered recently, chances are that it can be used for malicious purposes.

Domain TLD: it shows all top-level domain registered with the same base name as the main domain.

File: This page shows all files found within the domain. Some of these files may contain confidential information so it is recommended to check the content and remove them in that case.

Github hot data: This page shows information considered interesting obtained from Github repositories related to the scanned domain. It is possible that some files contain confidential information.

Bucket: This page shows the Amazon, Google and Azure buckets and containers that belong to the organization. They may contain sensitive data.

Mail Server: This page shows the mail servers used by IPs within the scanned domain. The presence of the SPF (Sender Policy Framework) and DMARC (Domain-based Message Authentication, Reporting & Conformance) is checked. These are email validation system designed to detect email spoofing attempts.

Mobile Apps: This page shows organization-related mobile applications uploaded to the Play Store or other third-party stores. Attackers are increasingly exploiting the field of mobile devices to carry out their attacks, cloning legitimate applications.

Phone number: This page shows every phone number published on the institutional website of the analyzed domain. It is suggested to remove personal phone numbers if present, as they may facilitate social engineering activities.

General Social: This page shows all the social presence of accounts named after the domain in analysis. Attackers may create accounts to simulate the identity of an organization, with the goal of establishing trust relationships with victims.

Mail: This page shows the email accounts belonging to the domain under analysis. It is reported whether the account was used to subscribe to an online service and if it is present in one or more data breaches.

Breached accounts: This page shows the presence of corporate email addresses in different data breach scenarios. At this moment SATAYO monitors the data breach evidence of accounts within a database of **27,607,593,488** entries and new data breaches are constantly being added.

Password: This page shows the passwords detected in the various data breaches, with an indication of the type of password. For any hashes present the equivalent in plain text is shown if it was possible to crack it.

Paste: This page shows the presence of corporate email accounts within various Paste Sites. The presence in multiple paste sites may indicate that a data breach has occurred, and that sensitive information may be at risk.

Open Bug Bounty: This page shows evidence of occurrences related to domain resources found within the Open Bug Bounty portal. The portal allows an organization to manage the Vulnerability Disclosure activity in a coordinated way with the researchers who discover it.

Deep & Dark Web: This page contains evidence retrieved from Deep & Dark Web sources. The analysis is performed with several keywords related to the analyzed domain.

Market: On this page SATAYO shows evidence related to credentials, cookies and sessions offered for sale within various marketplaces and originating from attacks carried out using log stealer malware. We try as much as possible to maintain privacy and not let the market know our interests. For example, to find amazon-related data, searches are done by searching for the keyword “azon” and the results are filtered so that they contain only the items we are really interested in.

Sandboxes: This page shows the evidence found within sandboxes and related to the monitored organization. Evidence is detected using special YARA rules preconfigured by the team of analysts. A sandbox detonates files into controlled virtual environments to track their activities and communications, producing detailed reports that include files opened, created and written, registry keys set, domains contacted, and more.

3.1.6 Onboarding Process

A SOC must receive data from the client in order to function properly. To carry out all the initial configurations, a period called onboarding is scheduled. This process is composed of multiple steps that end with the ingestion by the SIEM of all the data that the customer wants to monitor. The onboarding phase for Wuerth Phoenix customers lasts two months.

During onboarding, NetEye satellites are installed in the customer's environment and VPN connections are established between them and the NetEye master node present at Wuerth Phoenix. The satellites have the duty of collecting as input logs, events and streams generated by the hosts from applications and system detected within the monitoring perimeter of the client organization, and then forwarding the data to the NetEye Master. The latter receives and elaborates the data, sending it in real-time to the SIEM. The data is parsed and placed into a dedicated index, depending on the beat that sent it. Data is logically separated among different customers, with the use of multiple indexes customized with the appropriate tenant number. The NetEye master is placed internally in the Wuerth Phoenix cloud infrastructure, while the NetEye satellites are installed in the client's network in the form of Virtual Machines (VM). To ensure that all connected sources are send logs correctly, NetEye uses Icinga monitoring checks that sends alerts if a source fails to send events in a predefined time window.

3.2 Security

3.2.1 Detection Rules

Elastic has a Security category where detection rules can be defined and managed. Elastic already has more than **600** detection rules internally, which are kept constantly updated thanks to a script that downloads new versions every day from the official Elastic GitHub repository. Usually not all rules are used at the same time, and there is a tuning process for each customer based on the technologies deployed.

In addition to Elastic's rules, custom rules are also supported and can be added within the SIEM. What we typically do is monitor the global cyber threat landscape. Analysts must be informed about new zero-days vulnerabilities or possible exploits available online. There is a large community of security researchers and pen testers who are constantly looking for vulnerabilities in known applications, and when something is found, a rule is usually published to detect the malicious behavior. In the past, each SIEM had its own rule format, but cyber threats evolve rapidly, and to ensure a rapid response **Sigma Rules** were invented. The Sigma project allows defenders to write and share detection rules in a **common format** that can be understood by every SIEM and that grants faster access to security for all [39].

We often add Sigma rules inside our SIEM, but we do not take them for granted, even when they are published by authoritative sources. Before the activation in the production environment, every rule goes through a testing process in our lab. The procedure is simple: we write the rule, download the exploit associated with a vulnerability from which the rule is supposed to protect us, run the malware in a controlled environment, and monitor the behavior of the SIEM and the correct detection of the malicious process. If everything is okay the rule can be added to production and start protecting customers as well. If some adjustments are needed modifications are applied. Wuerth Phoenix is also a partner of **SOC Prime** [40], one of the most important **Threat Detection Marketplace** available, which offers a database of detection rules developed by Threat Hunters internationally.

There is always the possibility that a rule will generate false positives. The job of the SOC analysts is to analyze all the alerts and figure out whether the process that triggered the rule was malicious. The analysis may vary from alert to alert, for example it can be done by first reading the details of the alert, then browsing with Elastic SIEM's Discover function to the moment when the alert was generated, making different queries to isolate incriminating events and inspecting them in detail, correlating them with others and searching the Internet for information. A large number of false positives distracts the analyst from his work, so it is desirable to whitelist some harmless processes when they are identified, so that they do not generate more alerts. Adding something to the exception list should be done carefully, because future evidence will go unnoticed. Therefore, before doing so, we discuss with the customer, informing them about our findings and leaving the final decision of whether or not to whitelist something to them. They know their internal network architecture better than we do, and what looks suspicious could be a legitimate process.

3.2.2 Alerts

Detection rules, once activated, run continuously in the background, depending on the chosen timing. There is an option to add additional lookback time to scan more events a do not miss anything. The SIEM examines and correlates data to find anomalies and when something suspicious is detected an alert is raised. They are divided into 4 categories: **critical**, **high**, **medium**, and **low**. It is important to differentiate and assign different severities to understand what has the priority and needs to be handled first, especially in cases where a large number of warnings occur simultaneously. SOC analysts open the alert and analyze it further using the SIEM functionalities and their personal experience. For every investigation a case is opened in Elastic. Cases are another section of the Elastic Stack and are used when alerts require documentation or there is a need to notify the customer. When a new type of alert appears, our standard procedure is to create a case and document the performed analysis. We have created some templates to facilitate the writing of cases and share a standard process adopted by each analyst. Cases can be used for internal purposes to keep track of events and see if they recur in the future, or they can used as a link between the SOC and the client. When events are suspicious and require more attention, the case is sent to the customer. This is done through the Jira platform, explained in the next paragraph 3.2.3.

3.2.3 Ticketing System

The **Help Center** of Wuerth Phoenix is integrated with **Atlassian Jira**. By using Jira as a ticket platform, we can establish a connection with the customer and inform them about important evidence gathered from SATAYO or Elastic. Each ticket has a predefined workflow, starting from the open state and then moving from *waiting for support* to *waiting for customer* until it is *resolved* or *canceled.* To simplify our work, we have created scripts that create tickets directly from the platforms we use. The SOC evidence is analyzed with **Elastic Cases**, which can be automatically pushed to Jira if the corresponding tag is added. Automation takes care of opening the ticket, associating the respective client, notifying customers via email, etc. Similar to the SOC, SATAYO evidence is parsed in the platform and tickets are opened in Jira with another integration. Customer accounts saved in our system are automatically added as participants in a ticket related to their company and receive an email notification. Tickets for suspicious evidence are usually opened by us, but customers can browse SATAYO's findings and open tickets to request further analysis on the selected evidence. Jira can also be used by customers for other purposes, for example they can open tickets to raise questions, doubts or concerns, talk about problems they encountered, etc. We provide them with the link to the support portal where a guided procedure guides the user to the correct section for the ticket they wish to open.

3.2.4 Red Team Activities

The client onboarding process takes two months, and soon after it is completed some red team activities are performed to check the security of the client's infrastructure. These activities are aimed at finding vulnerabilities in the protocols used and checking for the possible presence of a flaw or an exploitable component in the network. Among the activities performed are user enumeration, credential sniffing, password cracking, and privilege escalation. Privilege escalation, for example, is performed at different levels, first without any access, then, if nothing vulnerable is found, additional permissions are requested and exploits are verified at a higher level. If nothing is found here either, the last step is to temporarily add our account as system administrators to test the infrastructure from the higher vantage point. These activities are a very important exercise for the SOC, which should be able to detect all of them, as they simulate attack scenarios. After onboarding, periodic checks are performed using **Greenbone Security Manager**, a continuous **vulnerability assessment platform** integrated into the SIEM. It contains more than **100,000** vulnerability tests and helps the SOC protect customers.

Threat Hunting, a process of proactively and iteratively searching within networks to identify and isolate advanced threats that elude current security solutions, is also frequently applied. Cyber attackers can remain hidden within an organization's network after a successful breach, silently collecting data and seeking to gain greater privileges with lateral movement tactics. To identify these **Advanced**

Persistent Threats (APTs), threat hunting is a fundamental instrument. When a new TTP emerges, threat hunters try to find out if that type of behavior is present within the network. Potential hidden activity can also be detected with the use of IoC and IoA, as well as the application of machine learning analytics to investigate large amounts of data and find irregularities that could lead to malicious activity.

The cybersecurity environment is often checked, and whenever a major new vulnerability is discovered, rules are updated or added to detect fresh threats. The tests are performed in a laboratory environment and then the rules are progressively deployed for each tenant monitored by the SOC.

4 Challenges and New Features

4.1 Introduction

A SOC maintains contact with the customer because when an anomaly is detected it needs to be reported quickly and efficiently. Opening tickets seems to be a good approach for this purpose, with an e-mail being sent to all required participants. In the case of more dangerous discoveries, a phone call may be the best option, along with direct access to the compromised machine to block and prevent further damage. In addition to direct contacts such as those just mentioned, keeping the customer informed of the status of the SOC and the IT world is another important aspect that should not be underestimated. To this end, monthly reports full of detailed and useful information are produced and delivered to clients. Reports such as these are also important from a legal point of view and can be delivered to auditors who verify compliance with the law. In Italy, for example, the privacy guarantor stipulates that logs containing administrators' authentication activities must be collected and stored in an unalterable manner and must be handed over in case of audits by auditors. A report containing this information is therefore desirable and required by the client who must obey current regulations. When the SOC service at Wuerth Phoenix was about to be initiated, having a monthly report to be delivered was considered an important feature. Some sort of report already existed, but it was poorly and hurriedly written. This chapter will talk about the work performed to create a better report to satisfy customers and increase the efficiency.

It was obvious from the beginning that the creation and delivery of these reports had to be somewhat automatized. It would be impossible to handle many different reports belonging to multiple clients each month, but at the same time not everything can be automated because human attention is needed to perform certain analyses. So, the problem was apparently simple: find a way to generate what is required and send it to the appropriate recipient with some sort of automation. The solution that was ultimately implemented and that is used today utilizes several tools and technologies to finally achieve the intended goal. The whole process will be explained in the following paragraphs. A flowchart that summarizes the process is shown in figure 4.1.

4.2 Step 1: Writing

Within Elastic's SIEM all client data are present and it is the perfect place to start to create a report. Elastic offers a feature, called Canvas, that is typically used to create large posters or infographics. You can import visualizations into it and create charts, tables, and counters based on the logs collected by Elastic. Each visualization has its own time interval and is dynamically updated with the logs present in that fraction of time. This feature seemed the most interesting to use to create a report, because along with static text and images there was the possibility of inserting content that would be automatically populated with the data of the current month.

The idea is to divide the report into three main categories, based on the three types of Threat Intelligence outlined in section 2.4. Operational Threat Intelligence will contain information about what the SOC has analyzed to monitor and protect the customer, Tactical Threat Intelligence will focus instead on information about the Tactics Techniques and Procedures (TTPs) used by cyber criminals to execute

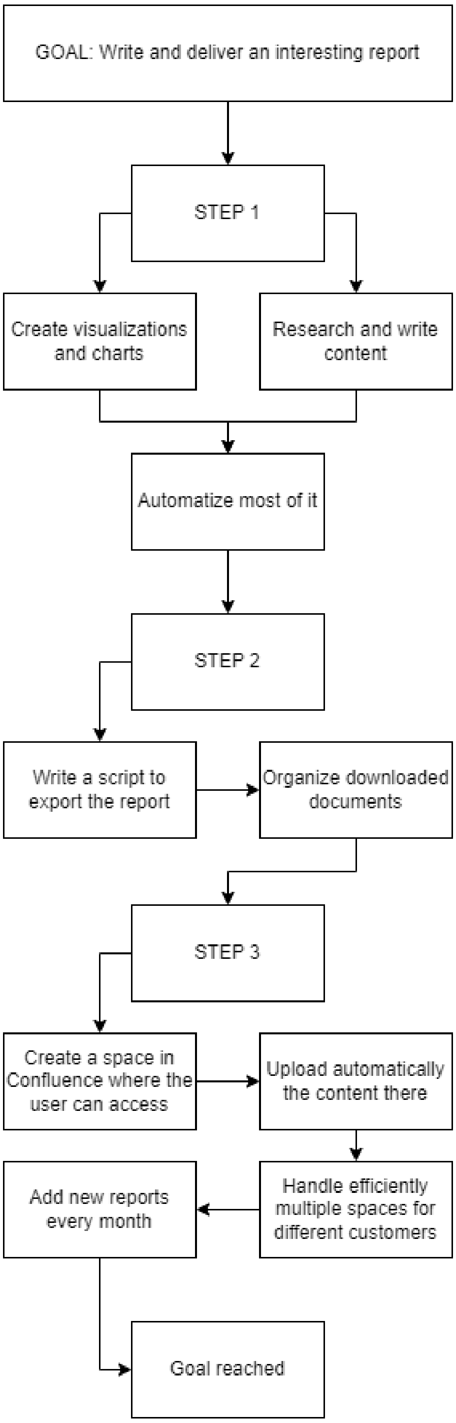

GOAL: Write and deliver an interesting report
STEP 1
Create visualizations and charts
Research and write content
Automatize most of it
STEP 2
Write a script to export the report
Organize downloaded documents
STEP 3
Create a space in Confluence where the user can access
Upload automatically the content there
Add new reports every month
Handle efficiently multiple spaces for different customers
Goal reached

Figure 4.1: The flowchart of the process

attacks, and finally, Strategic Threat Intelligence will talk about the overall cyber situation, threats, and impact on businesses.

It is not possible to fill all sections with charts and fully automate the process, especially the Tactical requires a monthly analysis of threat actors and their activities, which has to be done physically by a human being. Still, I tried to improve the process as best I could.

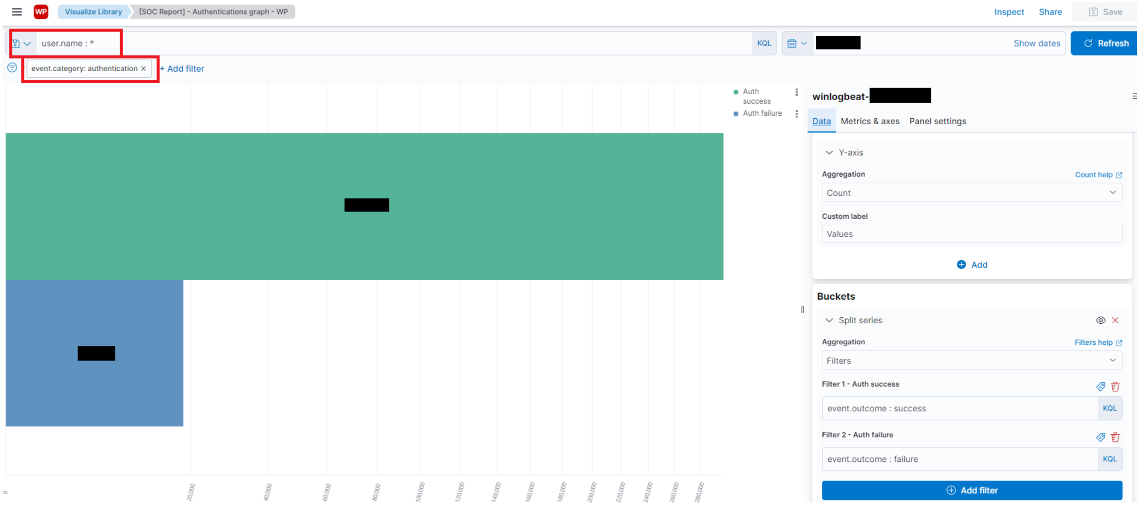
When editing a canvas there are two different ways to add graphics. One is to create them directly from the canvas, and in this case the queries must be written in Elasticsearch SQL (Structured Query Language), a component that allows SQL-type queries to be executed within Elastic. The other is to create a visualization in Kibana and then import it directly into the canvas.

There are some differences between the two approaches: Elasticsearch SQL allows you to query the data present in a DB and returns the results. It is able to match events in Elastic coming from different sources and across different categories and time intervals.

```
SELECT count(user.name) as Logout FROM "*name_ of_ index-*" where
event.code = 4634 AND (user.is_ admin = true) AND "@timestamp" > today() -
INTERVAL '45' day
```

Unfortunately, it is difficult to visualize the queried data in a good-looking graph. However, this query can be useful for showing a single number or data in a table.

Kibana's visualizations, on the other hand, offer a wide range of possibilities, from GUI wizards for creating graphs to complex syntax for customizing every aspect of the result, as shown in the chapter 3.1.3. From the GUI, you can make queries and apply filters to limit the amount of data you want to display in the visualization.



user.name : *

event.category: authentication

[redacted]

[redacted]

[redacted]

winlogbeat-[redacted]

Aggregation: Count

Custom label: Values

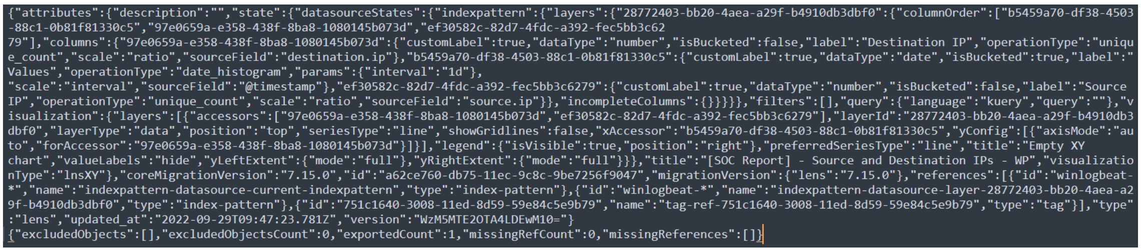
Aggregation: Filters

Filter 1 - Auth success: event.outcome : success

Filter 2 - Auth failure: event.outcome : failure

Figure 4.2: An example of Kibana visualization

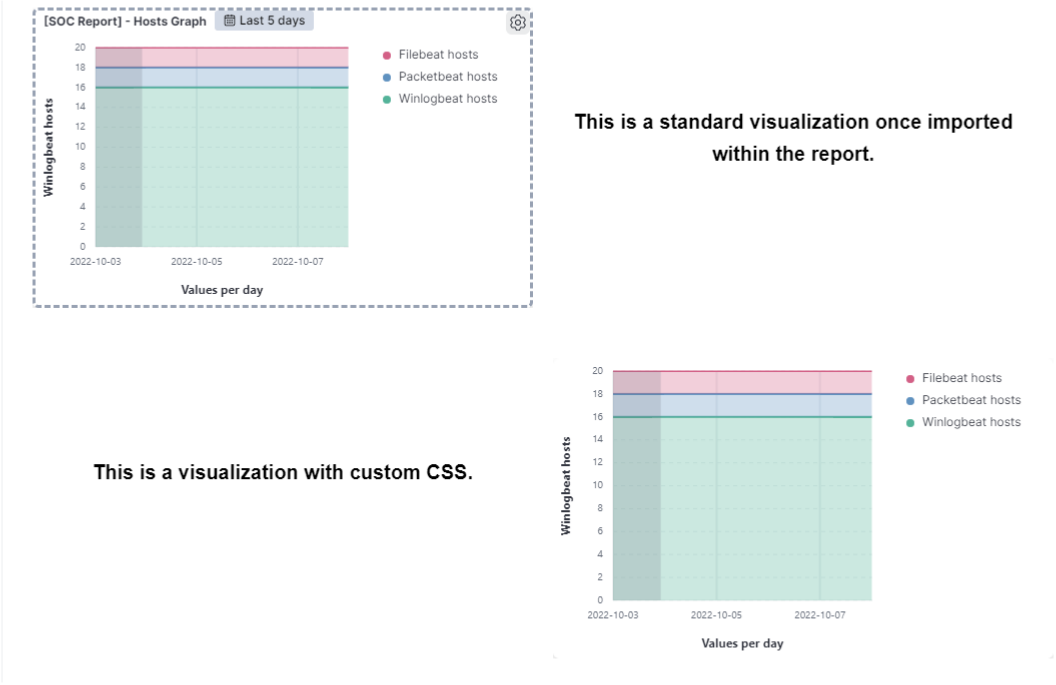
All the elements can be designed with the tools provided by the GUI and can be further customized with JSON input. They can also be exported as NDJSON, a flexible format often used for log files.



```
{"attributes":{"description":"","state":{"datasourceStates":{"indexpattern":{"layers":{"28772403-bb20-4aea-a29f-b4910db3dbf0":{"columnOrder":["b5459a70-df38-4503
-88c1-0b81f81330c5","97e0659a-e358-438f-8ba8-1080145b073d","ef30582c-82d7-4fdc-a392-fec5bb3c62
79"],"columns":{"97e0659a-e358-438f-8ba8-1080145b073d":{"customLabel":true,"dataType":"number","isBucketed":false,"label":"Destination IP","operationType":"uniqu
e_count","scale":"ratio","sourceField":"destination.ip"},"b5459a70-df38-4503-88c1-0b81f81330c5":{"customLabel":true,"dataType":"date","isBucketed":true,"label":"
Values","operationType":"date_histogram","params":{"interval":"1d"},
"scale":"interval","sourceField":"@timestamp"},"ef30582c-82d7-4fdc-a392-fec5bb3c6279":{"customLabel":true,"dataType":"number","isBucketed":false,"label":"Source
IP","operationType":"unique_count","scale":"ratio","sourceField":"source.ip"}},"incompleteColumns":{}}}}},"filters":[],"query":{"language":"kuery","query":""},"v
isualization":{"layers":[{"accessors":["97e0659a-e358-438f-8ba8-1080145b073d","ef30582c-82d7-4fdc-a392-fec5bb3c6279"],"layerId":"28772403-bb20-4aea-a29f-b4910db3
dbf0","layerType":"data","position":"top","seriesType":"line","showGridlines":false,"xAccessor":"b5459a70-df38-4503-88c1-0b81f81330c5","yConfig":[{"axisMode":"au
to","forAccessor":"97e0659a-e358-438f-8ba8-1080145b073d"}]}],"legend":{"isVisible":true,"position":"right"},"preferredSeriesType":"line","title":"Empty XY
chart","valueLabels":"hide","yLeftExtent":{"mode":"full"},"yRightExtent":{"mode":"full"}}},"title":"[SOC Report] - Source and Destination IPs - WP","visualizatio
nType":"lnsXY"},"coreMigrationVersion":"7.15.0","id":"a62ce760-db75-11ec-9c8c-9be7256f9047","migrationVersion":{"lens":"7.15.0"},"references":[{"id":"winlogbeat-
*","name":"indexpattern-datasource-current-indexpattern","type":"index-pattern"},{"id":"winlogbeat-*","name":"indexpattern-datasource-layer-28772403-bb20-4aea-a2
9f-b4910db3dbf0","type":"index-pattern"},{"id":"751c1640-3008-11ed-8d59-59e84c5e9b79","name":"tag-ref-751c1640-3008-11ed-8d59-59e84c5e9b79","type":"tag"}],"type"
:"lens","updated_at":"2022-09-29T09:47:23.781Z","version":"WzM5MTE2OTA4LDEwM10="}
{"excludedObjects":[],"excludedObjectsCount":0,"exportedCount":1,"missingRefCount":0,"missingReferences":[]}
```

Figure 4.3: The ndjson code for the graph shown in the picture 4.2

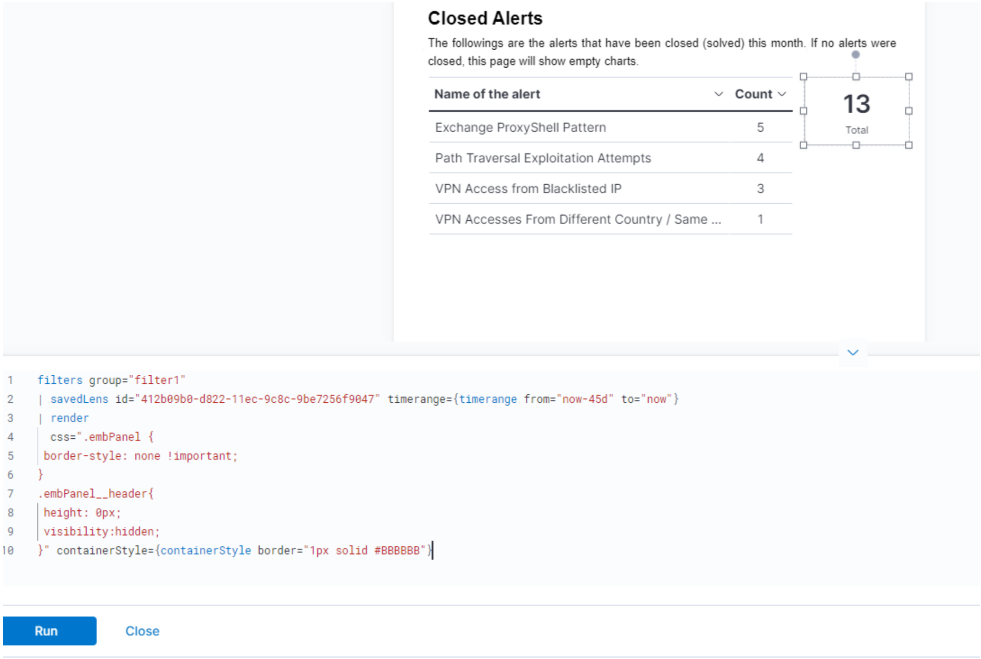
Once all the desired visualizations have been created, they can be imported into the canvas. The default appearance is rather ugly and bulky, but thankfully CSS can be customized to change the appearance of the graphic. Figure 4.4 shows a comparison between the default appearance and the customized one.



This is a standard visualization once imported within the report.

This is a visualization with custom CSS.

Figure 4.4: Some examples of editing


Closed Alerts

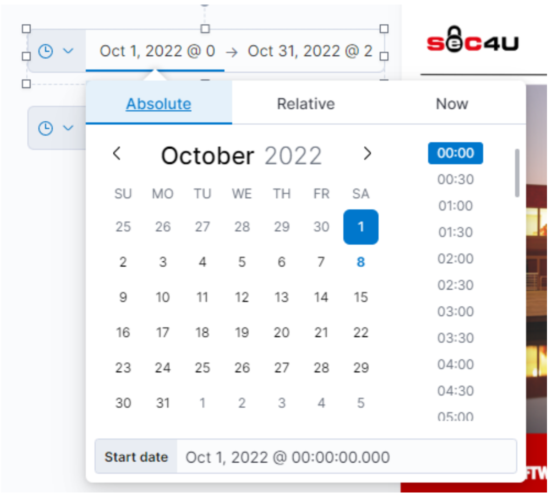
The followings are the alerts that have been closed (solved) this month. If no alerts were closed, this page will show empty charts.

Name of the alert	Count
Exchange ProxyShell Pattern	5
Path Traversal Exploitation Attempts	4
VPN Access from Blacklisted IP	3
VPN Accesses From Different Country / Same ...	1

```
filters group="filter1"
| savedLens id="412b09b0-d822-11ec-9c8c-9be7256f9047" timerange={timerange from="now-45d" to="now"}
| render
  css=".embPanel {
 border-style: none !important;
}
.embPanel__header{
 height: 0px;
 visibility:hidden;
}" containerStyle={containerStyle border="1px solid #BBBBBB"}
```

Figure 4.5: Above a report page containing two different visualizations, a metric box and a table. At the bottom, the expression editor of the metric box.

Modification can vary from view to view. Each element can be edited using a custom language that is present in Elastic. In the example shown in picture 4.5 we can see the visualization's expression editor showing us how the metric box that counts the number of closed alerts is structured. We notice a "timerange" attribute that displays data for the previous 45 days. Each visualization added has its own timerange, which means that when you want to see the next month's data, you have to edit them all manually. This is a waste of time, and fortunately there is the possibility of associating all visualizations with a single time filter. To do this, you must first create a time filter, shown in 4.6, give it a name, and associate the existing visualizations with that filter. In the image 4.5 this has been done in the first row of the expression editor. To see different data now simply change the date in the filter and all visualizations will be automatically updated. Some bugs were encountered and it was not possible to use the time filter correctly, but you can fix this by customizing the CSS of the entire report and allowing the filter to be selectable, as shown in 4.7.



Oct 1, 2022 @ 0 → Oct 31, 2022 @ 2

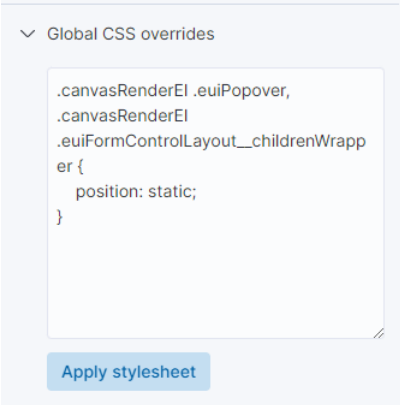
SU	MO	TU	WE	TH	FR	SA
25	26	27	28	29	30	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

00:00
00:30
01:00
01:30
02:00
02:30
03:00
03:30
04:00
04:30
05:00

Start date Oct 1, 2022 @ 00:00:00.000

Figure 4.6: The global time filter

Not only can CSS be changed, but variables can also be defined. I used them on the first page to list customer details and elsewhere in the report. Each month all you need to do is change the variable associated with the month and the time filter



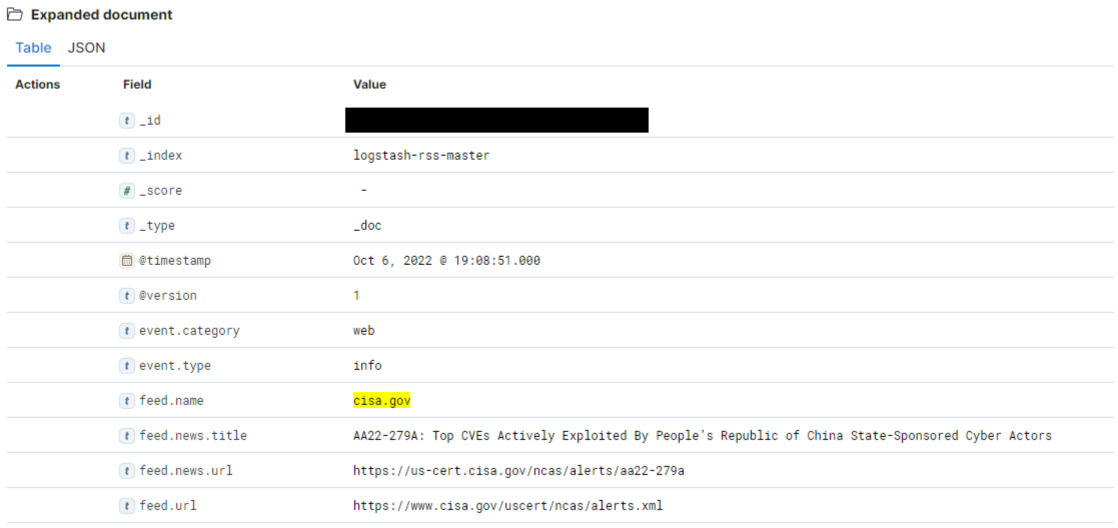
```
.canvasRenderEl .euiPopover,
.canvasRenderEl
.euiFormControlLayout__childrenWrapp
er {
   position: static;
}
```

Figure 4.7: The CSS override

to get the data you are looking for.

For the second and third parts of the report (Strategic and Tactical Threat Intelligence) it is not so easy to automate all the information we provide because it comes from different sources, such as Twitter accounts, cyber news websites, vulnerability databases, etc. One simple integration possible with Elastic is to import RSS feed data from various sources. There is already a Logstash input plugin that handles RSS; simply install it and edit the configuration file in the Logstash folder to add the desired RSS sites. Figure 4.8 shows an example of RSS log. From there you can set the update interval, which adjusts how often the data should be retrieved, and you can customize the fields. The log will then be available within Elastic along with other documents.

Other parts of the report still require manual writing for the moment, and future improvements will try to automate these sections and integrate them into Elastic. Automation helps us deliver an efficient report, but to be effective and interesting the research part on hacker activities should be conducted by a human. The creation of a new version of the report is usually done in one Elastic space. When the report is finished, the canvas and its visualization are copied to the other client spaces and automatically retrieve the correct data. The indexes are categorized with a tenant number so that each space has visibility only to its own data.



Expanded document

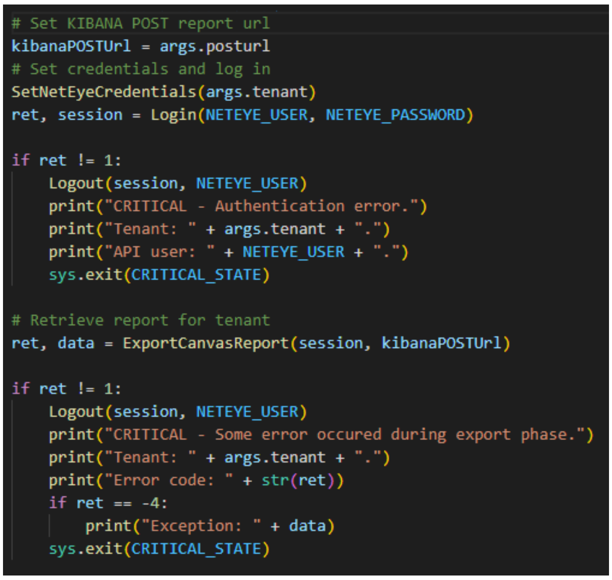
Table JSON

Actions	Field	Value
	t _id	[redacted]
	t _index	logstash-rss-master
	# _score	-
	t _type	_doc
	📅 @timestamp	Oct 6, 2022 @ 19:08:51.000
	t @version	1
	t event.category	web
	t event.type	info
	t feed.name	cisa.gov
	t feed.news.title	AA22-279A: Top CVEs Actively Exploited By People's Republic of China State-Sponsored Cyber Actors
	t feed.news.url	https://us-cert.cisa.gov/ncas/alerts/aa22-279a
	t feed.url	https://www.cisa.gov/uscert/ncas/alerts.xml

Figure 4.8: A log extracted from the RSS security bulletins published by cisa.gov

4.3 Step 2: Export

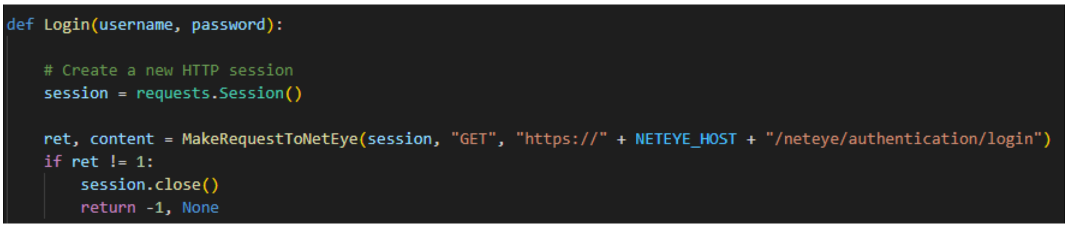
After the reports have been placed correctly in each space and loaded with client-specific information, they can be exported and sent. Exporting is easy to do, just open the canvas you want to export and click the *Share* button. Elastic needs some time to process the report, which can vary depending on the size. Then it will be available in the Reporting section of the SIEM in PDF format. So at this point the procedure was to enter each client's space, generate the report, download it, and write an e-mail. This takes time and the effort required increases with the number of clients, along with the possibility of getting something wrong. The idea was to improve the process by using a script that can generate the report via API, to reduce human intervention and prevent errors. The simplest thing to do was to initiate the export remotely and then let the user log into his Elastic space and navigate to the report section to see everything. Unfortunately, this does not work because the reports are only visible to the person who generated them. So instead of generating the report only, after the PDF is ready it is downloaded and sent to the client in some other way. Each canvas offers the option of using its POST URL to initiate an API export procedure. Since in Wuerth Phoenix environment the Elastic Stack is built on top of NetEye, it is necessary to authenticate to it before accessing the SIEM. For these authentication operations a special account with limited permissions was created, to avoid including personal information such as user credentials inside a script.



```python
# Set KIBANA POST report url
kibanaPOSTUrl = args.posturl
# Set credentials and log in
SetNetEyeCredentials(args.tenant)
ret, session = Login(NETEYE_USER, NETEYE_PASSWORD)

if ret != 1:
    Logout(session, NETEYE_USER)
    print("CRITICAL - Authentication error.")
    print("Tenant: " + args.tenant + ".")
    print("API user: " + NETEYE_USER + ".")
    sys.exit(CRITICAL_STATE)

# Retrieve report for tenant
ret, data = ExportCanvasReport(session, kibanaPOSTUrl)

if ret != 1:
    Logout(session, NETEYE_USER)
    print("CRITICAL - Some error occured during export phase.")
    print("Tenant: " + args.tenant + ".")
    print("Error code: " + str(ret))
    if ret == -4:
        print("Exception: " + data)
    sys.exit(CRITICAL_STATE)
```

Figure 4.9: We set the URL and perform the log in to NetEye. We then export the report.

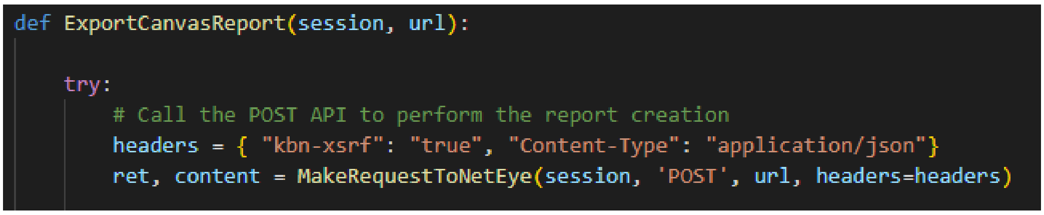
NetEye operates as a SSO provider, and after the login it is possible to use the other functionalities without having to log in again.



```python
def Login(username, password):

    # Create a new HTTP session
    session = requests.Session()

    ret, content = MakeRequestToNetEye(session, "GET", "https://" + NETEYE_HOST + "/neteye/authentication/login")
    if ret != 1:
        session.close()
        return -1, None
```

Figure 4.10: A request to the NetEye login page

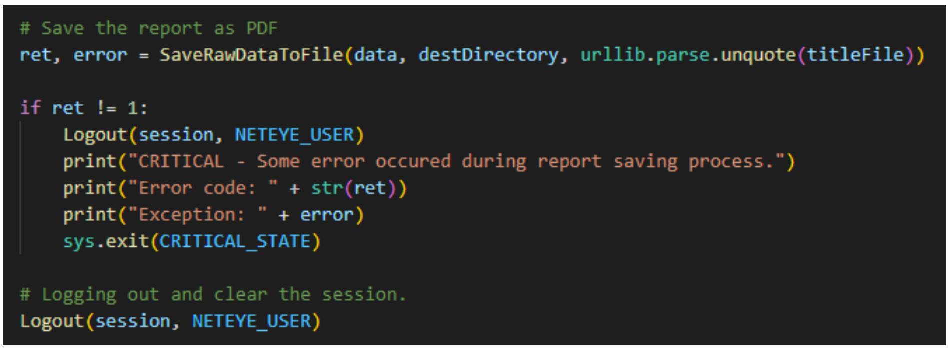
For the login process username and password are not enough; to increase security a CSRF (cross-site request forgery) token is extracted and sent through a POST request along with the credentials.



```
def ExportCanvasReport(session, url):

    try:
        # Call the POST API to perform the report creation
        headers = { "kbn-xsrf": "true", "Content-Type": "application/json"}
        ret, content = MakeRequestToNetEye(session, 'POST', url, headers=headers)
```

Figure 4.11: The function that sends the POST URL to Kibana

After logging in, it is possible to request Kibana to export the report. Some headers are required for this request and must be set correctly. Then a periodic request is sent to check if the export process is finished and the report is available for download. When it is ready, the PDF data is saved and sent to the function that will generate the document.



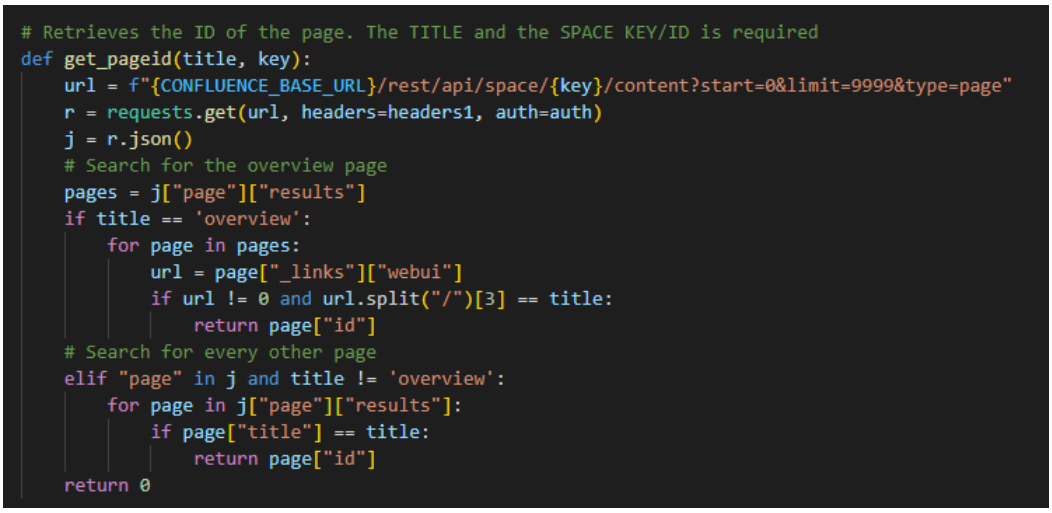
```
# Save the report as PDF
ret, error = SaveRawDataToFile(data, destDirectory, urllib.parse.unquote(titleFile))

if ret != 1:
    Logout(session, NETEYE_USER)
    print("CRITICAL - Some error occured during report saving process.")
    print("Error code: " + str(ret))
    print("Exception: " + error)
    sys.exit(CRITICAL_STATE)

# Logging out and clear the session.
Logout(session, NETEYE_USER)
```

Figure 4.12: Saving the report and logging out

This script takes as input parameters the tenant number and the POST URL of the report you want to download. The URL does not change if the report is modified, so you only need to retrieve it once.

4.4 Step 3: Delivery

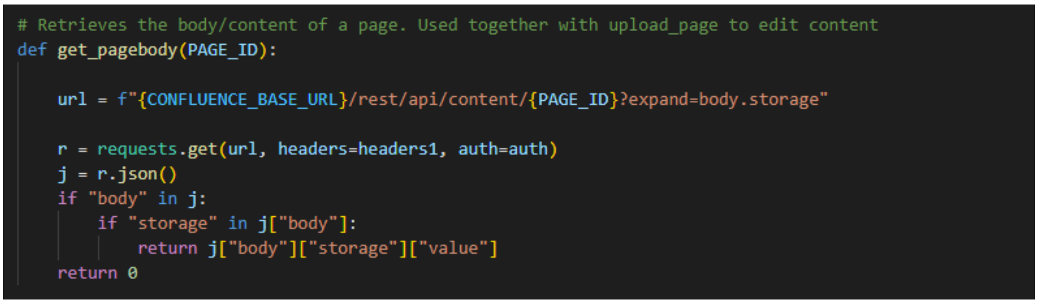
Reports are saved with meaningful names containing the tenant ID, title, and month to which the report refers. They are organized in separate folders, each for each tenant. Several ideas were analyzed and discussed to understand what the most efficient process might be to use to deliver the report to the client. The use of e-mail, adopted earlier, is not scalable and was discarded. A trial was done with Jira, where a new ticket was created monthly for each report, with the PDF attached. Unfortunately, there are some limitations in Jira that do not allow this attachment to be loaded correctly and clients were not able to see the file. It was necessary to open each ticket and attach the report manually, without much difference from the emailing process discussed earlier. Atlassian offers various products, and while Jira focuses on tickets and communication, Confluence was created to write documentation. So we have created a Confluence space for each client, where reports will be uploaded and where all the history can be easily managed in one place. Customers can access this space whenever they want and find all the information inside, always available for download. Since Confluence was created for documentation, the space was the perfect location where to add other useful information for the client, in addition to the report. We started writing a guide on how to use the SOC and some of its features, because we provide each customer with read-only accounts to access their Elastic space in the SIEM and from which they can inspect their data in detail if they wish. We are completely transparent and allow them to see and inspect the data that we are processing that belongs to their organization. Depending on the type of contract you sign, you can choose between a managed service (a complete SOC) and a non-managed one. In the case of a managed SOC, everything is handled by the SOC analysts and the customer does not need to log into Elastic, but we still provide them with accounts. The monthly report gives a static picture of the month, but it is limited by the number and width of pages. From Elastic, on the other hand, all data can be viewed dynamically and the time interval can be adjusted to fit custom needs. Access to this data allows clients to make custom queries. The unmanaged SOC option is frequently sold in Italy because there is a law that requires every company to collect and store all the operations performed by system administrators in an unalterable manner. For companies that are interested only in this but not in a fully operational SOC, we offer a reduced version that does not monitor the entire infrastructure but only authentication events. For these types of customers, it can be very useful to have guidance on how to use some basic Elastic functionality. Therefore, providing a user-friendly guide with simple directions on where to start was considered a good feature. To make this process more efficient, we wrote a script that uses the Confluence API and is able to keep all spaces updated to the latest version of the guide. Each page has a unique ID, which is needed for copy, edit or delete operations. The first step is to retrieve it.



```
# Retrieves the ID of the page. The TITLE and the SPACE KEY/ID is required
def get_pageid(title, key):
    url = f"{CONFLUENCE_BASE_URL}/rest/api/space/{key}/content?start=0&limit=9999&type=page"
    r = requests.get(url, headers=headers1, auth=auth)
    j = r.json()
    # Search for the overview page
    pages = j["page"]["results"]
    if title == 'overview':
        for page in pages:
            url = page["_links"]["webui"]
            if url != 0 and url.split("/")[3] == title:
                return page["id"]
    # Search for every other page
    elif "page" in j and title != 'overview':
        for page in j["page"]["results"]:
            if page["title"] == title:
                return page["id"]
    return 0
```

Figure 4.13: The function that gets the ID of a specific page. Of course, you also need the key to the space you want to examine

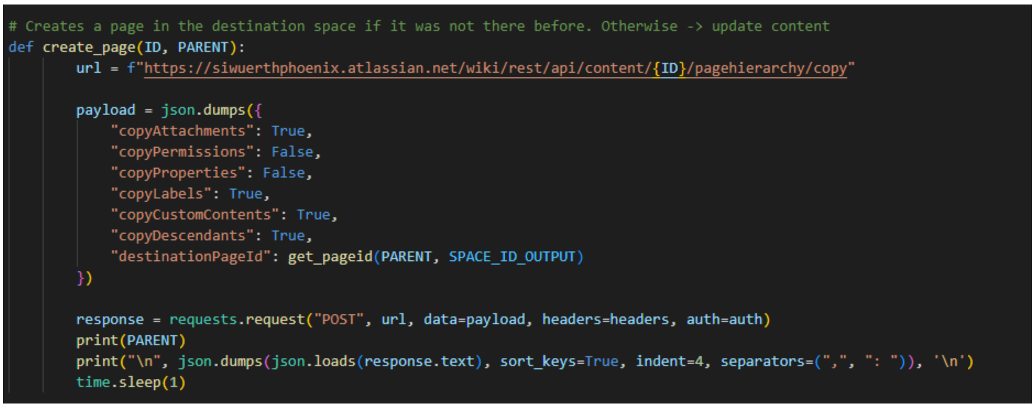
The contents of the Confluence page it is rendered as HTML, so you can retrieve it, edit it and push the modifications.



```
# Retrieves the body/content of a page. Used together with upload_page to edit content
def get_pagebody(PAGE_ID):

    url = f"{CONFLUENCE_BASE_URL}/rest/api/content/{PAGE_ID}?expand=body.storage"

    r = requests.get(url, headers=headers1, auth=auth)
    j = r.json()
    if "body" in j:
        if "storage" in j["body"]:
            return j["body"]["storage"]["value"]
    return 0
```

Figure 4.14: The function that retrieves the HTML content of a specific page

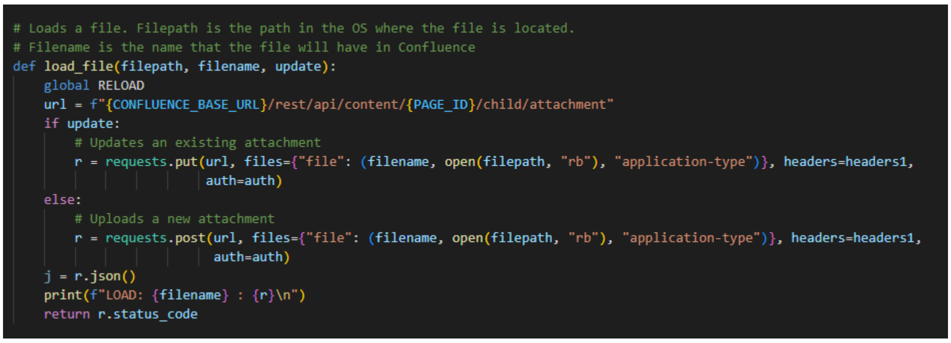
The content and hierarchy of the page are finally copied to the other spaces. The version of the page is checked first, so that the new content can be copied only if there is a previous version to be updated.



```
# Creates a page in the destination space if it was not there before. Otherwise -> update content
def create_page(ID, PARENT):
        url = f"https://siwuerthphoenix.atlassian.net/wiki/rest/api/content/{ID}/pagehierarchy/copy"

        payload = json.dumps({
            "copyAttachments": True,
            "copyPermissions": False,
            "copyProperties": False,
            "copyLabels": True,
            "copyCustomContents": True,
            "copyDescendants": True,
            "destinationPageId": get_pageid(PARENT, SPACE_ID_OUTPUT)
        })

        response = requests.request("POST", url, data=payload, headers=headers, auth=auth)
        print(PARENT)
        print("\n", json.dumps(json.loads(response.text), sort_keys=True, indent=4, separators=(",", ": ")), '\n')
        time.sleep(1)
```

Figure 4.15: The function that creates a new page. Used to keep pages updated between spaces

One page of the Confluence space is dedicated to the collection of reports. Monthly reports downloaded in the previous step are uploaded here, under the correct category. SOC customers, for example, receive two monthly reports and one weekly report. The script handles the loading of all reports, based on the inputs provided.



```
# Loads a file. Filepath is the path in the OS where the file is located.
# Filename is the name that the file will have in Confluence
def load_file(filepath, filename, update):
    global RELOAD
    url = f"{CONFLUENCE_BASE_URL}/rest/api/content/{PAGE_ID}/child/attachment"
    if update:
        # Updates an existing attachment
        r = requests.put(url, files={"file": (filename, open(filepath, "rb"), "application-type")}, headers=headers1,
                         auth=auth)
    else:
        # Uploads a new attachment
        r = requests.post(url, files={"file": (filename, open(filepath, "rb"), "application-type")}, headers=headers1,
                          auth=auth)
    j = r.json()
    print(f"LOAD: {filename} : {r}\n")
    return r.status_code
```

Figure 4.16: The function that uploads the reports in the correct section

5 Conclusion and Future Work

5.1 Future Improvements

Regarding the SIEM, the current situation at Wuerth Phoenix is the one explained in the steps in the chapter 4. With this procedure, reports can be exported and delivered more efficiently compared to doing everything by hand, but it is not yet fully automated. As shown, manual intervention is required to launch the scripts for each tenant. A possible future improvement would be to integrate the script into the NetEye cloud and let it run periodically once a month. This is possible by leveraging the capabilities of Icinga. You can create a template where you specify on which instance of Icinga you want to run the script (it can be the host, the satellites that retrieve logs, or the master) and the time interval at which the script should run. As for the sections that are not yet automated, we are integrating other platforms, such as OpenCTI, within Elastic. With more sources of Threat Intelligence, it will be easier to generate the report in the future.

5.2 Conclusions

This thesis explored what a Security Operation Center is and how it works, taking the one offered by Wuerth Phoenix as the main example. There are multiple types of SOCs and may differ a lot from the one described. There exist many SIEM and CTI platforms and each environment is different. Other SOCs may be more advanced in terms of artificial intelligence analysis or other aspects, but they all aim to increase the security of the monitored clients. The last chapter of this thesis analyzed a problem the author worked on during their assignment. A SOC is an enormous entity and the one mentioned is only a small but important part of the whole ecosystem. Little and big improvements are made on a daily basis to provide a reliable service to customers. SOCs are not perfect and still have several limitations today, but they are one of the best weapons of defense we possess against an ever-growing world of cybercrime. The author hopes they have succeeded in arousing interest in this topic especially to those who were not aware of it and to show how important security is and that it should not be underestimated. Together, we all can give a contribution to live in safer world.

References

[1] *World Internet Users Statistics and 2022 World Population Stats*. [Online]. Available: https://www.internetworldstats.com/stats.htm (visited on 09/24/2022).

[2] gmcdouga, *Check Point Research: Cyber Attacks Increased 50% Year over Year*, en-US, Jan. 2022. [Online]. Available: https://blog.checkpoint.com/2022/01/10/check-point-research-cyber-attacks-increased-50-year-over-year/ (visited on 09/26/2022).

[3] C. Edwardson, *SonicWall Threat Intelligence Confirms Alarming Surge in Ransomware, Malicious Cyberattacks as Threats Double in 2021*, en-US, Section: Press Release. [Online]. Available: https://www.sonicwall.com/news/sonicwall-threat-intelligence-confirms-alarming-surge-in-ransomware-malicious-cyberattacks-as-threats-double-in-2021/ (visited on 09/26/2022).

[4] *2022 Cyber Security Statistics Trends & Data*, en-US. [Online]. Available: https://purplesec.us/resources/cyber-security-statistics/ (visited on 09/26/2022).

[5] B. Robb, *The State of Ransomware in 2021*, en-US, Section: Ransomware, Jan. 2022. [Online]. Available: https://www.blackfog.com/the-state-of-ransomware-in-2021/ (visited on 09/26/2022).

[6] Cyberknow, *Update 21. 2022 russia-ukraine war — cyber group tracker. december 19*, en, Dec. 2022. [Online]. Available: https://cyberknow.medium.com/update-21-2022-russia-ukraine-war-cyber-group-tracker-december-19-24c61f3349e3 (visited on 01/26/2023).

[7] *NVD - Full Listing.* [Online]. Available: https://nvd.nist.gov/vuln/full-listing (visited on 07/30/2022).

[8] C. Onwubiko, "Cyber security operations centre: Security monitoring for protecting business and supporting cyber defense strategy", in *2015 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, Jun. 2015, pp. 1–10. DOI: 10.1109/CyberSA.2015.7166125.

[9] IBM, "Cost of a Data Breach Report 2021", en, *Risk quantification*, p. 73,

[10] ENISA, *How to set up CSIRT and SOC*, en, Report/Study. [Online]. Available: https://www.enisa.europa.eu/publications/how-to-set-up-csirt-and-soc (visited on 07/24/2022).

[11] P. Mallory, *Security operations center*, en-US. [Online]. Available: https://resources.infosecinstitute.com/career/security-operations-center/ (visited on 07/30/2022).

[12] M. Vielberth, F. Böhm, I. Fichtinger, and G. Pernul, "Security Operations Center: A Systematic Study and Open Challenges", *IEEE Access*, vol. 8, pp. 227 756–227 779, 2020, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3045514.

[13] P. Jacobs, A. Arnab, and B. Irwin, "Classification of Security Operation Centers", in *2013 Information Security for South Africa*, ISSN: 2330-9881, Aug. 2013, pp. 1–7. DOI: 10.1109/ISSA.2013.6641054.

[14] *What is Cyber Threat Intelligence? [Beginner's Guide]*, en. [Online]. Available: https://www.crowdstrike.com/cybersecurity-101/threat-intelligence/ (visited on 09/10/2022).

[15] A. Cybersecurity, *How to Build a SOC: Threat Intelligence | AT&T Cybersecurity.* [Online]. Available: https://cybersecurity.att.com/solutions/security-operations-center/building-a-soc/threat-intelligence (visited on 07/24/2022).

[16] T. Bikov, D. Radev, T. Iliev, and D. Stankovski, "Threat Hunting as Cyber Security Baseline in the Next-Generation Security Operations Center", in *2021 29th Telecommunications Forum (TELFOR)*, Nov. 2021, pp. 1–4. DOI: 10.1109/TELFOR52709.2021.9653361.

[17] C. Onwubiko, *Security operations centre: Situation awareness, threat intelligence and cybercrime*, en-US. [Online]. Available: https://ieeexplore.ieee.org/document/8074844 (visited on 07/31/2022).

[18] *OpenCTI – The open source solution for processing and sharing threat intelligence knowledge*, en-GB. [Online]. Available: https://www.ssi.gouv.fr/actualite/opencti-the-open-source-solution-for-processing-and-sharing-threat-intelligence-knowledge/ (visited on 09/20/2022).

[19] *OpenCTI Ecosystem.* [Online]. Available: https://www.notion.so (visited on 09/20/2022).

[20] *MITRE ATT&CK®.* [Online]. Available: https://attack.mitre.org/ (visited on 08/07/2022).

[21] M. Mutemwa, J. Mtsweni, and L. Zimba, "Integrating a Security Operations Centre with an Organization's Existing Procedures, Policies and Information Technology Systems", in *2018 International Conference on Intelligent and In-*

novative Computing Applications (ICONIC), Dec. 2018, pp. 1–6. DOI: `10.1109/ICONIC.2018.8601251`.

[22] B. A. Alahmadi, L. Axon, and I. Martinovic, "99% False Positives: A Qualitative Study of SOC Analysts' Perspectives on Security Alarms", en, p. 19,

[23] A. Perera, S. Rathnayaka, N. D. Perera, W. Madushanka, and A. N. Senarathne, "The Next Gen Security Operation Center", in *2021 6th International Conference for Convergence in Technology (I2CT)*, Apr. 2021, pp. 1–9. DOI: `10.1109/I2CT51068.2021.9418136`.

[24] C. Onwubiko, "Rethinking Security Operations Centre Onboarding", in *2021 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, Jun. 2021, pp. 1–9. DOI: `10.1109/CyberSA52016.2021.9478245`.

[25] R. Sobers, *86 Ransomware Statistics, Data, Trends, and Facts [updated 2022]*, en. [Online]. Available: `https://www.varonis.com/blog/ransomware-statistics` (visited on 09/25/2022).

[26] IBM and T. H. Poll, *Public sector security research - IBM-Harris Poll survey 2020*. [Online]. Available: `https://www.ibm.com/downloads/cas/74JKYWZQ` (visited on 09/25/2022).

[27] F. D. János and N. Huu Phuoc Dai, "Security Concerns Towards Security Operations Centers", in *2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, May 2018, pp. 000 273–000 278. DOI: `10.1109/SACI.2018.8440963`.

[28] C. Chipeta, *What is a Security Operations Center (SOC)? | UpGuard*, en. [Online]. Available: `https://www.upguard.com/blog/security-operations-center` (visited on 09/23/2022).

[29] D. Instinct and P. Institute, *The economic value of prevention in the report.pdf*. [Online]. Available: https://info.deepinstinct.com/hubfs/The_Economic_Value_of_Prevention_in_the_Report.pdf (visited on 09/25/2022).

[30] *Reconnaissance, Tactic TA0043 - Enterprise | MITRE ATT&CK®*. [Online]. Available: https://attack.mitre.org/tactics/TA0043/ (visited on 08/07/2022).

[31] *SATAYO*. [Online]. Available: https://satayo.cloud/ (visited on 08/07/2022).

[32] R. D. Manager, *I vantaggi di un SOC Attacker Centric*, it-IT, Mar. 2022. [Online]. Available: https://www.datamanager.it/2022/03/i-vantaggi-di-un-soc-attacker-centric/ (visited on 08/07/2022).

[33] O. Podzins and A. Romanovs, “Why SIEM is Irreplaceable in a Secure IT Environment?”, in *2019 Open Conference of Electrical, Electronic and Information Sciences (eStream)*, Apr. 2019, pp. 1–5. DOI: 10.1109/eStream.2019.8732173.

[34] *Elastic Stack: Elasticsearch, Kibana, Beats & Logstash*, en-us. [Online]. Available: https://www.elastic.co/elastic-stack (visited on 09/26/2022).

[35] U. D. of Justice, *Legal Considerations when Gathering Online Cyber Threat Intelligence and Purchasing Data from Illicit Sources*. [Online]. Available: https://www.justice.gov/criminal-ccips/page/file/1252341 (visited on 08/14/2022).

[36] *OSINT Framework*. [Online]. Available: https://osintframework.com/ (visited on 08/14/2022).

[37] *Jivoi/awesome-osint: A curated list of amazingly awesome OSINT*. [Online]. Available: https://github.com/jivoi/awesome-osint (visited on 08/14/2022).

[38] Crowdstrike, *IOA vs IOC: Defining & Understanding The Differences | CrowdStrike*, en. [Online]. Available: `https://www.crowdstrike.com/cybersecurity-101/indicators-of-compromise/ioa-vs-ioc/` (visited on 08/15/2022).

[39] A. Swan, *SIGMA Rules: The Beginner's Guide*, en-US, Section: Blog, May 2022. [Online]. Available: `https://socprime.com/blog/sigma-rules-the-beginners-guide/` (visited on 10/15/2022).

[40] *Sigma Rules Search Engine for Threat Detection, Threat Hunting, and CTI*, en. [Online]. Available: `https://socprime.com` (visited on 10/15/2022).