Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя
(повне найменування вищого навчального закладу)
Факультет комп'ютерно-інформаційних систем і програмної інженерії
(назва факультету)
Кафедра кібербезпеки
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

бакалавр
(освітній рівень)

на тему: *"Створення автоматизованої системи аналізу журналів для Виявлення аномалій і загроз безпеки в комп'ютерній системі"*

Виконав: студент (ка)
Спеціальності:
125 «Кібербезпека»
(шифр і назва напряму підготовки, спеціальності)
Микитюк Т.В.
підпис (прізвище та ініціали)

Керівник *Козак Р.О.*
підпис (прізвище та ініціали)

Нормоконтроль
підпис (прізвище та ініціали)

Завідувач кафедри *Загородна Н.В.*
підпис (прізвище та ініціали)

Рецензент
підпис (прізвище та ініціали)

м. Тернопіль – 2023

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра кібербезпеки
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Загородна Н.В.
(підпис) (прізвище та ініціали)
«__» _____ 2023 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Бакалавр
(назва освітнього ступеня)

за спеціальністю 125 Кібербезпека
(шифр і назва спеціальності)

Студенту Микитюку Тарасу Володимировичу
(прізвище, ім'я, по батькові)

1. Тема роботи Створення автоматизованої системи аналізу журналів для виявлення аномалій і загроз безпеки в комп'ютерній системі

Керівник роботи Козак Руслан Орестович
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «03» 04 2023 року № 4/7-349

2. Термін подання студентом завершеної роботи 17.06.2023

3. Вихідні дані до роботи Журнали безпеки операційної системи Linux і Windows

4. Зміст роботи (перелік питань, які потрібно розробити)

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності, основи хорони праці			

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	16.02 – 19.02	*Виконано*
2.	Підбір джерел про методи та засоби аналізу журналів	20.02 – 27.02	*Виконано*
3.	Підбір джерел про механізми, структуру та архітектуру вистем виявлення вторгнень	28.02 – 16.03	*Виконано*
4.	Опрацюванн джерел в галузі досліджень	17.03 – 20.03	*Виконано*
5.	Вибір мов програмування та інструментарію розробки	20.03-05.04	*Виконано*
6.	Розробка архітектури системи та її складових	06.03 – 17.04	*Виконано*
7.	Розробка модулів для зчитування і обробки даних	18.04 – 29.04	*Виконано*
8.	Аналіз результатів тестування та їх інтерпретація	30.04 – 13.05	*Виконано*
9.	Оцінка ефективності розробленої системи	14.05 – 21.05	*Виконано*
10.	Виконання розділу «Безпека життєдіяльності, основи охорони праці»	22.05 – 05.06	*Виконано*
11.	Оформлення кваліфікаційної роботи	06.06 – 12.06	*Виконано*
12.	Нормоконтроль	10.06 – 15.06	*Виконано*
13.	Перевірка на плагіат	16.06 – 18.06	*Виконано*
14.	Попередній захист кваліфікаційної роботи	19.06 - 20.06	*Виконано*
15.		20.06.2022	

Студент _____ _____
(підпис) (прізвище та ініціали)

Керівник роботи _____ *Лечаченко Т.А.*
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Створення автоматизованої системи аналізу журналів для виявлення аномалій і загроз безпеки в комп'ютерній системі // Кваліфікаційна робота ОР «Бакалавр» //Микитюк Тарас Володимирович// Тернопільський національний технічний університет імені Івана Пулюя, факультет комп’ютерно-інформаційних систем і програмної інженерії, кафедра кібербезпеки, група СБ-41 // Тернопіль, 2023 // С. – 52, рис. – 25, ліст. – 3.

Ключові слова: автоматизація, аналіз, виявлення аномальної поведінки.

Кваліфікаційна робота присвячена розробці системи виявлення аномалій поведінки користувачів для виявлення вразливостей, атак та інших загроз безпеки комп’ютерної системи. Розроблена система допоможе забезпечити безпеку комп’ютерної системи та підвищить ефективність процесу виявлення вразливостей та реагування на можливі загрози.

У першому розділі кваліфікаційної роботи було розглянуто функції систем виявлення і запобігання вторгнень а також SIEM, види журналів та методи аналізу журналів подій.

У другому розділі було розглянуто питання автоматизції аналізу журналів подій, а також визначено критерії виявлення аномалій і загроз безпеки.

У третьому розділі було розроблено і протестовано систему автоматичного аналізу журналів подій.

ANNOTATION

Creation of an automated log analysis system for detecting anomalies and security threats in a computer system// Thesis of educational level "Bachelor" // Mykytiuk Taras Volodymyrovych // Ternopil Ivan Puluj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Cybersecurity, SB-41 group // Ternopil, 2023 // P. 52, fig. – 25, list. – 3.

Keywords: automation, analysis, detection of abnormal behavior.

The qualification work is devoted to the development of a system for detecting anomalies in user behavior to detect vulnerabilities, attacks and other threats to the security of the computer system. The developed system will help to ensure the security of the computer system and increase the efficiency of the process of detecting vulnerabilities and responding to possible threats.

In first section of the qualification work, the functions of intrusion detection and prevention systems, as well as SIEM, types of logs and methods of analyzing event logs were considered.

In the second section, the issue of automating the analysis of event logs was considered, as well as the criteria for detecting anomalies and security threats were defined.

In the third section, a system for automatic analysis of event logs was developed and tested.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

IDS – Intrusion detection system (система виявлення вторгнень)

IPS – Intrusion prevention system (система запобігання вторгнень)

SIEM – Security Information and Event Management (інформація про безпеку та управління подіями)

ELK-stack – Elasticsearch, Logstash, Kibana stack (сукупність інструментів для роботи з журналами подій)

Бд – база даних

Дашборд – інформаційна панель в системі ELK-stack

Зміст

ВСТУП

В сучасному світі, де комп’ютерні системи використовуються у всіх сферах, від медицини до військової справи, кібербезпека стала надзвичайно важливою проблемою для багатьох організацій, компаній та користувачів. Хакери та кіберзлочинці не рідко діють в інтересах держави, проти інших країн, постійно шукаючи нові і більш ефективні способи для несанкціонованого доступу до інформації що можуть становити в тому числі державну таємницю. Щоб забезпечити захист від таких загроз, необхідно вести постійний моніторинг вразливостей та аномалій комп’ютерних систем для вчасного реагування на інциденти, що може бути досягнуто з допомогою аналізу журналів подій.

Одним з ефективних способів виявлення аномалій та загроз безпеки в комп’ютерних системах є аналіз журналів. Комп’ютерні системи зберігають журнали подій, які відображають всі операції, що відбуваються в системі. Аналіз журналів може допомогти виявити потенційні загрози та небезпеки в системі, а також забезпечити вчасну реакцію на можливі атаки, в цьому полягає актуальність роботи.

Саме тому, метою даної бакалаврської роботи є розробка автоматизованої системи збору і аналізу журналів для виявлення аномалій та загроз безпеки в комп’ютерній системі. За допомогою цієї системи буде забезпечено постійний моніторинг і автоматичний аналіз журналів веб сервера, що дозволить оперативно виявляти потенційні загрози та аномалії в роботі системи.

Для досягнення мети були поставленні такі завдання:

1) Вивчення методик та інструментів для збору і аналізу журналів сервера.

2) Розробка методики і алгоритму для аналізу журналів сервера з метою виявлення аномалій та загроз безпеки.

3) Розробка програмного забезпечення мовою програмування Python для автоматизації збору і аналізу журналів веб сервера.

4) Тестування розробленої системи на реальних журналах сервера та оцінка ефективності роботи системи.

1 ТЕОРЕТИЧНІ ЗАСАДИ АНАЛІЗУ ЖУРНАЛІВ ПОДІЙ БЕЗПЕКИ

1.1 Функції ips\ids і SIEM

Виявлення вторгнень — це процес моніторингу вашого мережевого трафіку та його аналізу на наявність ознак можливих вторгнень, таких як спроби використання та інциденти, які можуть бути безпосередньою загрозою для вашої мережі. Зі свого боку, запобігання вторгненню — це процес виявлення вторгнення з подальшим припиненням виявлених інцидентів, що зазвичай здійснюється шляхом відкидання пакетів або припинення сеансів. Ці заходи безпеки доступні як системи виявлення вторгнень (IDS) і системи запобігання вторгненням (IPS), які є частиною заходів безпеки мережі , вжитих для виявлення та припинення потенційних інцидентів, і включені до функцій брандмауерів наступного покоління (NGFW) .

IDS/IPS відстежує весь трафік у мережі, щоб виявити будь-яку відому шкідливу поведінку. Одним із способів, якими зловмисник намагатиметься скомпрометувати мережу, є використання вразливості в пристрої або програмному забезпеченні. IDS/IPS ідентифікує ці спроби використання та блокує їх до того, як вони успішно скомпрометують будь-яку кінцеву точку в мережі. IDS/IPS є необхідними технологіями безпеки як на межі мережі, так і в центрі обробки даних , саме тому, що вони можуть зупинити зловмисників, коли вони збирають інформацію про вашу мережу.

Для виявлення інцидентів зазвичай використовуються три методології виявлення IDS:

- виявлення на основі сигнатур порівнює сигнатури з спостережуваними подіями, щоб визначити можливі інциденти. Це найпростіший метод виявлення, оскільки він порівнює лише поточну одиницю активності (наприклад, пакет або запис у журналі зі списком підписів) за допомогою операцій порівняння рядків;
- виявлення на основі аномалій порівнює визначення того, що вважається нормальною активністю, із спостережуваними подіями, щоб виявити значні

відхилення. Цей метод виявлення може бути дуже ефективним для виявлення раніше невідомих загроз;

- аналіз протоколу з урахуванням стану порівнює попередньо визначені профілі загальноприйнятих визначень доброякісної активності протоколу для кожного стану протоколу з спостережуваними подіями, щоб виявити відхилення.

Security Information and Event Management (SIEM) — це програмне забезпечення, яке збирає та аналізує інформацію з кількох різних джерел по всій інфраструктурі(рис.1.1). Якщо дивитися глобально, можна уявити SIEM як систему, яка збирає всю мережеву активність в одному місці у вигляді зрозумілого набору даних. Таким чином, замість перегляду консолі кожного пристрою безпеки, ви отримуєте одне централізоване рішення для пошуку попереджень і реагування відповідно. Це рішення не лише централізує інформацію, але й виконує глибокий аналіз інформації, щоб знайти приховані закономірності та побачити, чи ваша компанія піддається атаці. Термін був придуманий компанією Gartner в далекому за мірками інформаційних технологій, 2005 році. Однак, за минулий час, як поняття, так і функції SIEM зазнали великої кількість змін. Спочатку абревіатура складалася з двох термінів: SIM (Security Information Management) і SEM (Security Event Management). Отже, SIEM збирає, об’єднує і аналізує інормацію з різних пристроїв в умовах реального часу тим самим допомагаючи аналітикам і фахівцям з безпеки виявляти і блокувати атаки на ранній стадії. Інструменти SIEM працюють за наперед заданими правилами, що дозволяє визначати загрози і сповіщати про них адміністраторам безпеки.

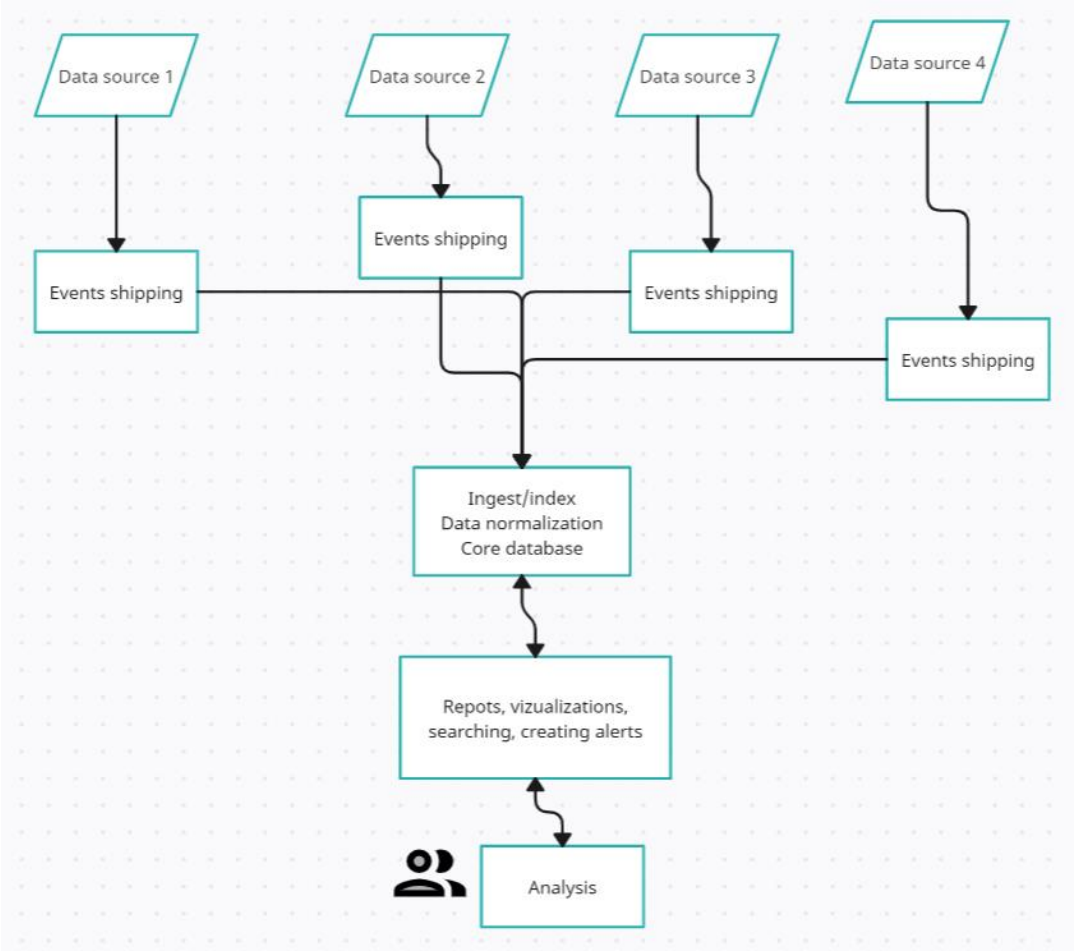

Рисунок 1.1 – Базова архітектура SIEM

Функції SIEM:

- збір логів з різних джерел;
- нормалізація логів;
- таксономія отриманої інформації ;
- кореляція подій;
- створення сповіщення, надання інструментів для розлідування інциденту;
- зберігання інформації про події;
- пошук і фільтрація збережених даних.

Джерелам даних для SIEM можуть бути будь які програмні засоби чи технічні пристрої які мають функцію записування подій. До таких відносяться:

- сервери. Сервери обробляють велику кількість інформації кожного дня тому логи які вони генерують можуть бути використанні для виявлення аномалій і потенційних загроз;
- мережеві пристрої. Мережеві пристрої такі як маршрутизатори, комутатори, вай фаї і мережеві фаєрволи можуть генерувати логи різного типу які можуть бути надіслані до SIEM і використані для аналізу;
- antivirus / antispy програмне забезпечення. Журнали подій цих програм можуть дати велику кількість інформації яка буде корисною аналітикам безпеки для виявлення атак чи загроз;
- хости. Журнали з робочих комп'ютерів організації містять інформацію про дій користувачів, автентифікацію, програмне забезпечення ті іншу важливу інформацію;
- IDS/IPS. Логи цих систем містять інформацію про атаки, аномалії та інші загрози безпеки.

Як запевняють експерти, такі системи призначені для виявлення системних аномалій та підтримки систем виявлення шкідливого програмного забезпечення. Завдяки SEIM, можна спостерігати більш чітку картину що відбувається інформацією в мережі. Така система необхідна, коли звичайні звичні засоби не справляються зі своїм завданням, SIEM корелює отриману інформацію з еталоном і виявляє невідповідності. З цієї причини, компанії різного розміру, розглядають SIEM як важливий, додатковий бар'єр для захисту мережі від направлених атак.

Завдання SIEM:

- відстеження аутентифікації і компрометації користувальницьких акаунтів;
- відстеження заражень системи. Відстеження шкідливого ПЗ на підставі логів і журналів антивірусів і брандмауера;
- перевірка вихідного трафіку, в тому числі підозрілого. Моніторинг логів проксі, брандмауера, NetFLow. Крім цього, виявляються підозрілі зовнішні з'єднання, і потенційна крадіжка даних;

- відстеження змін в системі, редагування реєстру, зміна системних файлів і перехоплення процесів. Перевірка відповідності дій дозволеним політикам;
- відстеження веб-додатків і атак на них. Аналіз журналів WAF і веб-сервера. Аналіз звітів на предмет компроментації веб-додатків.

У деяких компаній гостро стоїть питання – чи потрібна така система, або такий підхід застарів і не ефективний. Потрібно розуміти, що SIEM система безпосередньо не буде протидіяти хакерам, вона тільки аналізує велику кількість вхідної інформації і надає звіт і доводи про небезпечності певної області, повідомляючи користувача. Таким чином, SIEM потрібно додавати в комплексний підхід для забезпечення безпеки мережі. Обов'язково в компанії повинен бути фахівець, який зможе відреагувати на повідомлення системи і в найкоротші терміни вжити дії для запобігання зараження або крадіжки конфіденційної інформації.

Крім усього іншого, впроваджуючи SIEM систему, потрібно особливо ретельно ознайомитися з інфраструктурою компанії в кожному конкретному випадку, враховуючи встановлену систему безпеки, архітектуру мережі. Правильно налаштована система, з правильно налаштованими правилами, дозволяє адміністратору реагувати тільки на дійсно важливі події та інциденти. Основною ідеєю таких систем є можливість передачі на них всі рутинні процеси і можливість приймати рішення по рівню загрози події для мережі.

1.2 Визначення журналів та їх види

Журнали сервера – це прості текстові документи які містять інформацію про події які стались на сервері. Ці файли автоматично створюються і підтримуються сервером і не потребують втручання адміністратора і може надати вам детальну інформацію про дату, час, користувача і саму суть події. Зазвичай сервери створюють файли CLF або необроблені файли журналів які представляють собою

набір, часто, неструктурованої і не обробленої інформації. В залежності від налаштувань сервера, файли журналів видаляються через деякий проміжок часу.

Файли CLF – це файли, у яких кожен рядок відповідає одному запиту. При взаємодії користувача з веб сайтом який розміщений на цьому веб сервері у файлі журналу буде створено декілька рядків тексту, кожен з них може містити різну інформацію, наприклад:

- адреса веб сайту;
- IP адреса користувача;
- час і дата;
- http запит, статус і т.д.

Файли журналів чудовий інструмент для аналізу роботи сервера, тому він доступний лише адміністраторам сервера.

Існують декілька видів журналів сервера, основні з них включають:

Журнал доступу. Журнали доступу записують інформацію про те, які HTML-файли та графічні зображення запитуються з вашого сервера. Журнали доступу повідомляють вам про кількість відвідувачів, а також про їхнє походження (наприклад, чи прийшли вони з сайту .com, .edu або .gov). Крім того, ви можете отримати шаблони використання та інформацію про те, які сторінки відвідувачі запитували найчастіше.

Журнали агента. Журнали агентів можуть записувати і надавати вам інформацію про те, які веб-клієнти робили запити на вашому сервері.

Журнали адресата. Журнали адресата записують інформацію про URL-адресу, за якою відвідувач перебував безпосередньо перед тим, як перейти на вашу веб-сторінку і зробити запит на ваш сервер. Цей тип журналів надзвичайно корисний, коли ви хочете точно визначити походження запитів на вашому веб-сервері, а також те, з яких веб-сторінок надходить трафік на ваш сервер.

Журнали помилок. Журнали помилок записують і надають вам інформацію про невдалі запити до сервера. По суті, вони автоматично генерують

повідомлення про помилку, як тільки хтось намагається отримати доступ до неіснуючого файлу на вашому сервері.

В залежності від конфігурації сервера, можуть бути доступні інші види журналів, наприклад журнали аудиту. Адміністратори можуть використовувати ці журнали для аналізу роботи чи безпеки сервера.

1.3 Аналіз інструментів та методів аналізу журналів

Існує досить багато інструментів для збору та аналізу журналів сервера, одними з найпопулярніших є splunk, ELK stack (elasticsearch, logstash, kibana), fluentd.

Splunk – це платформа для збору, індексації та аналізу великих обсягів даних, включаючи журнали сервера. Splunk дозволяє аналізувати дані в реальному часі, знаходити аномалії та загрози безпеки, а також створювати звіти та графіки. До переваг можна віднести:

1) Широкий функціонал: splunk підтримує багато джерел даних і може аналізувати великі обсяги даних.

2) Легкість використання: інтерфейс splunk простий та зрозумілий, що дозволяє швидко навчитися користуватись платформою

3) Розширюваність: Splunk дозволяє створювати свої власні додатки та розширювати функціонал за допомогою сторонніх додатків.

Недоліки Splunk:

1) Високі витрати: Splunk є комерційною платформою, що може призвести до значних витрат на ліцензії та обладнання.

2) Обмежені можливості безкоштовної версії: Безкоштовна версія Splunk має обмежені можливості та обсяги даних, що може обмежити її використання.

3) Системні вимоги: Для використання Splunk необхідно мати достатньо потужний сервер та мережу з високою пропускною здатністю.

Порівняно з іншими схожими інструментами, Splunk має більш широкий функціонал та легкість використання. Проте, в порівнянні з безкоштовними інструментами, Splunk має значні витрати та обмежені можливості безкоштовної версії.

ELK Stack - це відкритий стек для аналізу і візуалізації журналів, що складається з трьох основних компонентів: Elasticsearch, Logstash та Kibana. Elasticsearch є потужним інструментом для зберігання та пошуку даних, він забезпечує швидкий пошук та аналіз великих обсягів даних. Logstash - це інструмент для обробки інформації та логів, який здатний збирати, обробляти та передавати дані в Elasticsearch. Kibana - це інтерактивний інтерфейс для відображення та аналізу даних, він забезпечує можливість візуалізації даних та створення дашбордів для моніторингу та аналізу.

Однією з переваг ELK Stack є його відкритість та безкоштовність, що зробило його популярним в ряді сфер. Також, стек є досить простим та легким у використанні, що робить його доступним для більш широкої аудиторії. Крім того, стек має гнучку архітектуру, що дозволяє легко налаштовувати та розширювати його функціонал.

Серед недоліків можна виділити те, що в ELK Stack немає вбудованих функцій безпеки, тому її доведеться додавати окремо. Також, можуть виникати проблеми зі швидкодією при великих обсягах даних, які потребують використання додаткових ресурсів.

Fluentd - це відкритий інструмент для збору, агрегації та передачі журналів. Він дозволяє збирати дані з різних джерел, включаючи сервери, додатки та сервіси хмарних обчислень, та направляти їх до різних призначених місць, таких як бази даних, сховища даних та інші системи аналізу даних.

Переваги:

1) Fluentd має широкий спектр вбудованих розширень та плагінів для різноманітних задач.

2) Fluentd може працювати з різними форматами журналів, включаючи JSON, syslog та CSV.

3) Fluentd має гнучку архітектуру, що дозволяє налаштовувати його під різні потреби.

Недоліки:

1) Fluentd може бути вимогливий до ресурсів та мати високу витрату пам'яті, особливо при обробці великої кількості даних.

2) Конфігурація може бути дещо складною для новачків.

Порівняно з ELK Stack, Fluentd пропонує ширший спектр розширень та плагінів, що дозволяє більш ефективно налаштовувати систему аналізу журналів. У порівнянні з Splunk, Fluentd є відкритим інструментом, що може бути більш привабливим для компаній, що прагнуть уникнути високих витрат на ліцензії та підтримку.

Для аналізу журналів сервера існує кілька методів, серед яких:

1) Ручний аналіз: це метод, в якому журнали переглядаються вручну, щоб знайти відповідні запити, події та помилки. Цей метод може бути досить часоємним, але може бути корисним для знаходження важливих проблем.

2) Аналіз журналів з використанням текстового пошуку: цей метод включає в себе використання інструментів для пошуку ключових слів та фраз в журналах. Це дозволяє знайти важливі події та помилки швидше, ніж при ручному пошуку.

3) Використання машинного навчання: цей метод включає в себе використання алгоритмів машинного навчання для автоматичного аналізу журналів сервера. Моделі машинного навчання можуть навчитися розпізнавати важливі події та помилки, що дозволяє ефективно знаходити та вирішувати проблеми.

1.4 Системи запобігання вторгнень: механізми, структура і архітектура

Система виявлення вторгнень (IDS) — це система, яка відстежує мережевий трафік на наявність підозрілої активності та сповіщає, коли така активність виявлена.

IDS можна порівняти з системою запобігання вторгненням (IPS), яка відстежує мережеві пакети на предмет потенційно шкідливого мережевого трафіку, як IDS, але її основною метою є *запобігання* загрозам після виявлення, на відміну від первинного виявлення та запису загроз.

Системи виявлення вторгнень використовуються для виявлення аномалій з метою зловити хакерів до того, як вони завдадуть реальної шкоди мережі. IDS можуть бути як мережевими, так і хостовими. Система виявлення вторгнень на основі хоста встановлена на клієнтському комп'ютері, тоді як мережева система виявлення вторгнень знаходиться в мережі.

Системи виявлення вторгнень працюють, шукаючи ознаки відомих атак або відхилення від нормальної діяльності. Ці відхилення або аномалії надходять у стек і перевіряються на протокольному та прикладному рівнях. Вони можуть ефективно виявляти такі події, як сканування ялинок і отруєння системи доменних імен (DNS).

IDS може бути реалізовано як програмне забезпечення, що працює на апаратному забезпеченні клієнта, або як пристрій безпеки мережі . Хмарні системи виявлення вторгнень також доступні для захисту даних і систем у хмарних розгортаннях.

IDS діляться на типи за принципом роботи і за місцем розташування, за місцем розташування системи виявлення вторгнень поділяються на:

Network Intrusion Detection System (NIDS) – система виявлення вторгнень яка виявляє зловмисний трафік у мережі. Як правило, NIDS встановлюється на апаратне забезпечення в межах мережевої інфраструктури організації(рис 1.2). NIDS це пасивний пристрій який зазвичай не заважає проходженню трафіку, але може зменшити пропускну здатність мережі за певних умов. NIDS «сніфить»

внутрішній інтерфейс фаєрвола в режимі читання та надсилає сповіщення на сервер менеджменту NIDS.

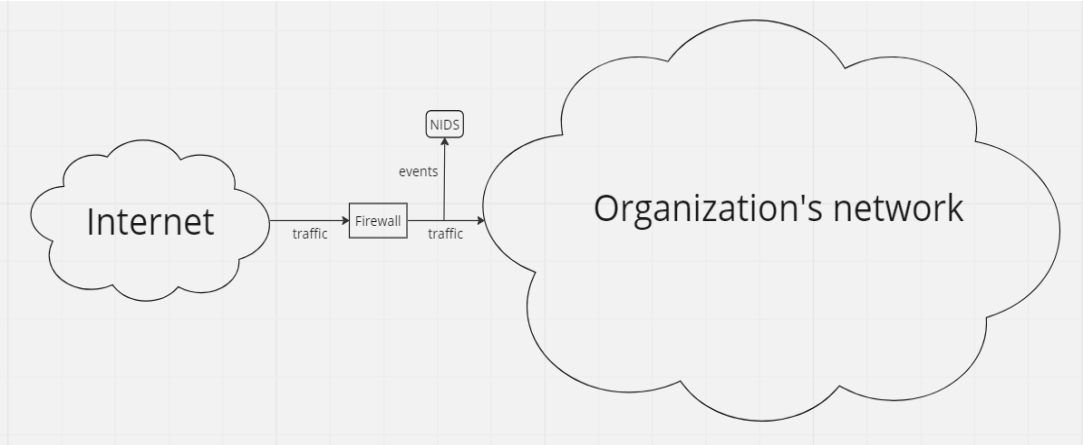

Рисунок 1.2 – Спрощена схема встановлення NIDS

Переваги NIDS полягають у:

- аналіз всього вхідного і вихідного трафіку. Система аналізує весь вхідний і вихідний трафік, що збільшує шанси на успішне виявлення вторгнення;
- робота в режимі реального часу. Такий режим роботи дозволяє аналітикам кібербезпеки виявляти атаки на ранній стадії та швидко реагувати на інциденти;
- складність виявлення зловмисниками. NIDS працює в пасивному режимі і не впливає на трафік, тому її складно виявити;
- можуть бути розміщені в критичних областях. NIDS можуть бути розміщені будь де в мережі, що дозволяє використовувати її можливості на максимум.

Недоліки:

- обслуговування. Оскільки NIDS зазвичай встановлюються на виділений апаратний засіб, потрібно буде виділити більше часу на ручну взаємодію;
- низька специфіка. Чим більша трафіку аналізує NIDS, тим більша ймовірність генерації хибних сповіщень чи ігнорування вторгнень.

Network Node Intrusion Detection System (NNIDS) – технічно, NNIDS це різновид NIDS, але вони відрізняються принципом роботи, тому її можна віднести до окремого типу. Так само як і NIDS цей тип системи виявлення вторгнень аналізує мережевий трафік, але замість того щоб покладатись на централізований девайс як моніторить мережу, NNIDS аналізує трафік з кожного вузла окремо(Рис 1.3).

Такий підхід має свої переваги, а саме:

- швидкість. Кількість трафіку який припадає на кожен вузол менший, тому система буде працювати швидше;
- менші витрати ресурсів. Логіка проста – менше трафіку, менше ресурсів потрібно на їх опрацювання.

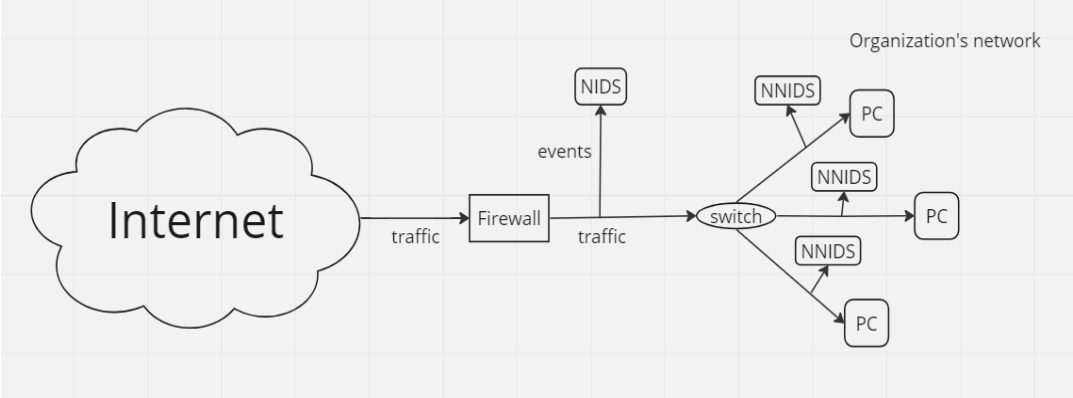

Рисунок 1.3 – Спрощена схема встановлення NIDS

Host Intrusion Detection System (HIDS) – встановлюється на кожен пристрій який має підключення до мережі(рис 1.4). Хостова система виявлення вторгнень працює за принципом порівняння поточних снапшотів цільової системи з минулими записами, таким чином система виявляє вторгнення. До переваг цього типу системи виявлення вторгнень можна віднести:

- можуть бути встановлені на сервери і комп’ютери;
- можуть точно визначити скомпрометований пристрій;
- можуть генерувати сповіщення при зміні або видалені системних файлів;
- дуже ефективні проти інсайдерських атак.

До недоліків HIDS можна віднести її низьку швидкість реагування на вторгнення.

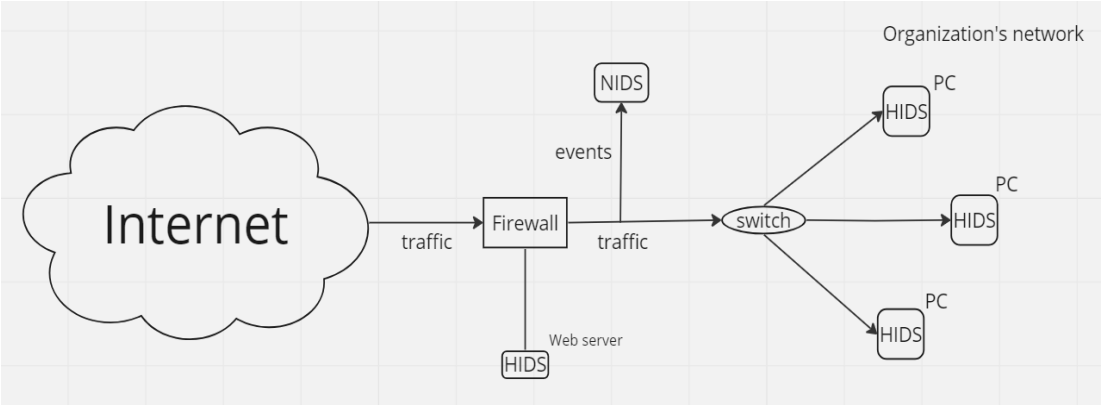

Рисунок 1.4 – Спрощена схема встановлення HIDS

Protocol-Based Intrusion Detection System (PIDS) – спеціальна система виявлення вторгнень яка контролює використовуваний протокол. На практиці цей тип IDS контролює потік HTTP і HTTPS трафіку між клієнтом і веб-сервером. На рисунку 1.5 зображено приклад схеми встановлення PIDS у мережі.

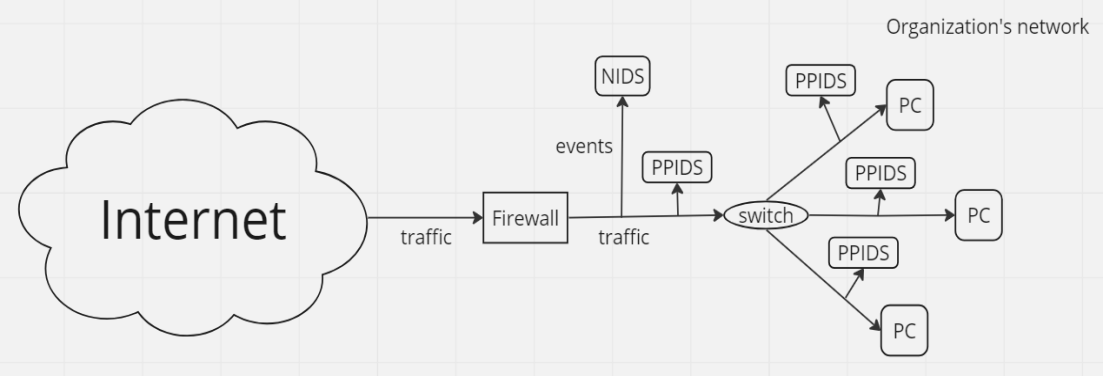

Рисунок 1.5 – Спрощена схема встановлення PIDS

Application Protocol-Based Intrusion Detection System (APIDS) – тип системи виявлення вторгнень який спеціалізується на моніторингу та аналізі специфічного протоколу, що використовуються в комп’ютерній системі. Зазвичай система складається з агентів що розміщуються між групою серверів для моніторингу протоколу програми між двома з’єднаними пристроями. Наприклад APIDS може бути встановлено між веб сервером і сервером баз даних(рис 1.6).

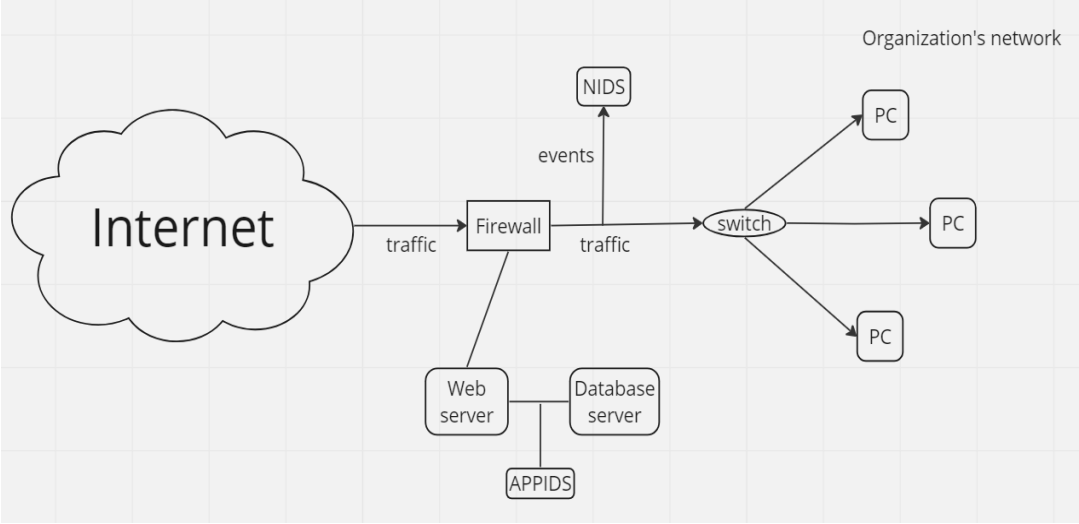

Рисунок 1.6 – Спрощена схема встановлення APIDS

За принципом роботи системи виявлення вторгнень поділяються на:

Signature-Based Intrusion Detection System (SIDS) – націлена на пошук патернів поведінки для їхнього порівняння з патернами вже відомих атак. Цей тип системи виявлення вторгнень покладається на базу даних попередніх атак. Як тільки активність в мережі співпадає з патерном з бази даних, система генерує сповіщення і надсилає його адміністратору. Оскільки основою роботоздатності SIDS є база даних сигнатур, постійне оновлення бази даних життєво необхідне. Крім того, якщо організація була атакована невідомою технікою, то цей тип системи виявлення вторгнень не захистить систему.

Anomaly-Based Intrusion Detection system (AIDS) – система виявлення вторгнень на основі аномалій спрямована на виявлення аномальної активності, що може свідчити про атаку. На відміну від SIDS, ця система здатна виявляти атаки нульового дня (zero-day intrusions). З допомогою машинного навчання і статистичних даних, система створює модель «нормальної» поведінки. В разі якщо трафік не під визначення нормального з точку зору AIDS, система помітить його як підозрілий. Проблемою цього типу системи виявлення вторгнень є велика кількість хибних спрацьовувань.

Гібридна система виявлення вторгнень – поєднання найкращого з AIDS і SIDS. Так само як пошук патернів, ця система аналізує поведінку з метою

виявлення аномалій, тому може виявляти як і відомі атаки, так і атаки нульового дня. Єдиним недоліком цієї системи є ще більша кількість хибних спрацьовувань.

2 МЕТОДИКА ТА АЛГОРИТМ АНАЛІЗУ ЖУРНАЛІВ ПОДІЙ

2.1 Автоматизація аналізу журналів подій

Автоматизований аналіз – це здатність системи без людського втручання виявляти аномалії, закономірності при обробці журналів, що створюються серверами або іншими компонентами інформаційної системи. Журнали можуть містити різноманітну інформацію про систему, наприклад запити, помилки, доступи та інше. В автоматизації аналізу інформації використовуються сучасні технології машинного навчання, штучного інтелекту, обробки природної мови, добування знань і інші. До задач автоматизованого аналізу входить генерування репорту про помилки, аномалії чи незвичну поведінку та надсилання його системному адміністратору. Алгоритми машинного навчання можна використовувати для швидкого і досить точного аналізу великих даних. В цьому полягає одна з ключових переваг автоматизованого аналізу даних, автоматизована система виявляє проблеми які ймовірно залишились би не поміченими, загубленими серед величезної кількості даних при ручному аналізі.

Автоматизована аналітика дозволяє дозволяє швидше робити висновки на підставі отриманих даних. При такому підході автоматизований аналіз дає ширші можливості для аналізу даних і виявлення атак на ранніх стадіях що дозволяє зекономити час і прийняти міри набагато швидше ніж це було б зроблено при ручному аналізі. Таким чином автоматизований аналіз стає чудовим доповненням дозволяючи зекономити людські ресурси, а для бізнесу велику кількість грошей які можна витратити для інших цілей тим самим підвищити ефективність інвестицій за рахунок автоматизованої аналітики. Це дозволяє системним адміністраторам виявляти і вирішувати проблеми до того як вони стануть серйозними настільки, що призведуть до зайвих витрат. Можливість швидко реагувати на зміни підвищує гнучкість системи і дозволяє системним адміністраторам і користувачам приділити більше часу і уваги проблемам які не можуть бути вирішені за допомогою автоматичної аналітики.

Такий підхід до аналізу даних розширює можливості від постачальників програмного забезпечення до звичайних користувачів. Застосування автоматизованого аналізу має величезний спектр застосування і варіюється від виявлення атак до вибору стратегії ведення бізнесу.

Для звичайних користувачів автоматизована аналітика надає більш персоналізовану і актуальну статистику, оскільки гнучкість такого методу збору і аналізу інформації дозволяє налаштувати її під себе отримавши статистику яка є найважливішою для кожного користувача і зробити це набагато швидше ніж при ручному підході до аналізу інформації. За допомогою алгоритмів машинного навчання можна створити персоналізовані сценарії для найважливіших даних. З часом поповнюючи базу сценаріїв можна отримати найбільш релевантну статистику збільшуючи ефективність і зменшуючи час і зусилля.

2.2 Визначення критеріїв виявлення аномалій та загроз безпеки

IDS забезпечує виявлення неавторизованого доступу в комп’ютерну систему або мережу або факт несанкціонованого управління ними через мережу інтернет. Системи виявлення вторгнень встановлюється зазвичай на важливих точках комп’ютерної мережі, часто в демілітаризованій зоні або на кінцевих точках мережі. Один з компонентів системи виявлення вторгнень перехоплює весь мережевий трафік з метою подальшого його аналізу на предмет зловмисної поведінки.

IDS переслідує дві основні мети: Аналіз інформації та прийняття рішення на основі проаналізованих даних. Рішення в цьому контексті це: надсилат чи не надсилати сповіщення адміністратору безпеки. Для виконання цих завдань система IDS здійснює наступні дії:

- моніторить мережевий трафік і відслідковує активність користувачів;
- в залежності від типу IDS, проводить регулярний аудит важливих системних файлів;

- здійснює статистичний аналіз снапшотів цільової системи порівнюючи їх показники вже з відомим атаками.

IDS намагається виявити порушення політики безпеки. До особливостей віднести те що система виявлення вторгнень не завжди блокує виявлену атаку, натомість вона створює звіт і сповіщає адміністратора безпеки.

Для виявлення вторгнень і загроз безпеці система IDS використовує різні критерії і методи:

Підписи атак. IDS має базу даних відомих сигнатур, які складаються з певних шаблонів або прикладів поведінки відомих атак, це може бути ключова фраза або команда. Система виявлення вторгнень порівнює вхідний трафік з цими сигнатурами і в разі збігу генерує сповіщення.

Аномалії. Система виявлення вторгнень аналізує загальну поведінку мережі або окремої системи для виявлення аномалій в поведінці. До цього можна віднести різку зміну обсягу трафіку, незвичні шаблони трафіку, зміни в конфігурації і т.д.

Відхилення від норми. Система виявлення вторгнень може мати певні встановлені норми, згідно яких вона буде аналізувати мережу. Відхилення від цих норм може свідчити про атаку.

Виявлення незвичних дій. Система виявлення вторгнень може аналізувати поведінку користувачів або компонентів мережі, щоб визначити незвичні або підозрілі дії. До таких дій можна віднести спроби несанкціонованого доступу, незвичний трафік, маніпуляції з запитами і т.д.

2.3 Алгоритм автоматичного аналізу журналів

З метою створення автоматизованої системи аналізу журналів буде використано комплекс засобів які будуть взаємодіяти між собою. Алгоритм включає в себе:

- збір лог-файлів. Система постійно моніторить та збирає лог файли з метою їх подальшого аналізу. Лог-файли будуть зберігатись в базі даних для створення звітності;
- нормалізація лог файлів. Сирі журнали важкі для читання і аналізу, тому необхідно виділити з них основу інформацію яка буде корисна під час аналізу;
- виявлення аномалій. На цьому етапі будуть застосовані методи аналізу з метою виявлення загроз інформаційної безпеки і аномалій в системі;
- сповіщення. У разі виявлення загрози безпеці буде автоматично створене сповіщення для інформування аналітика безпеки;
- зберігання данних. Незалежно від того який пріоритет і інформацію містить лог-файл його буде збережено в централізовану бази даних для подальшого аналізу, виявлення трендів та патернів поведінки для подальшого вдосконалення системи. На рисунку 2.1 зображено блок-схему алгоритму роботи системи.

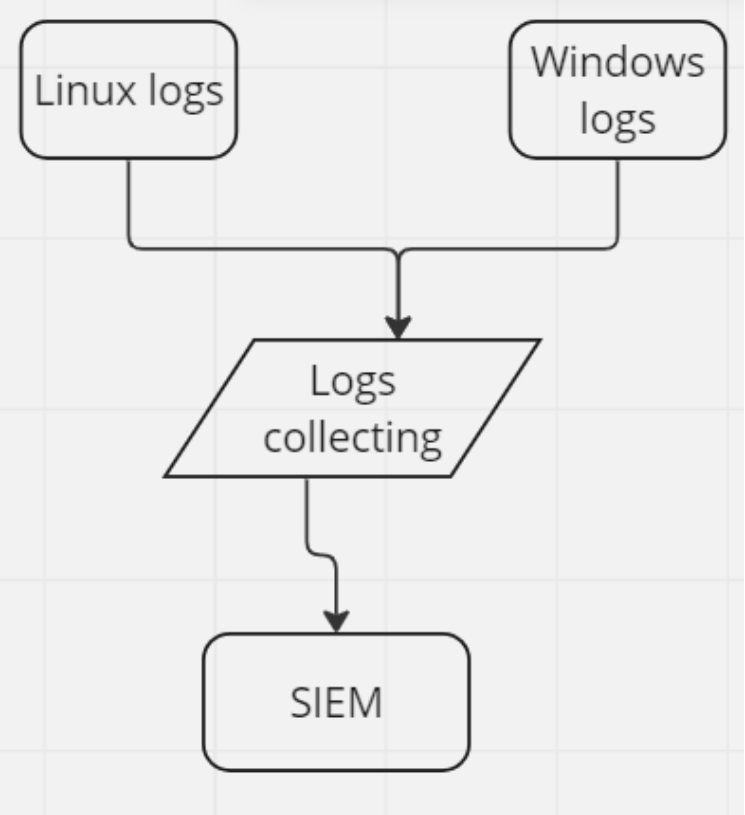

Рисунок 2.1 – Алгоритм роботи системи

Для побудови ефективної і вигідної системи, компоненти системи будуть обрані опираючись на три критерії:

Складність використання. Компоненти системи з якими буде взаємодіяти аналітик мають бути легкі у використанні, а інтерфейс інтуїтивно зрозумілий.

Використання ресурсів. Важливо побудувати таку систему яка буде ефективно використовувати ресурси. Системні ресурси, такі як пам’ять, оперативна пам’ять та процесор мають використовуватись ефективно щоб досягти максимальної продуктивності, при мінімальних затратах, а також важливо забезпечити ефективний розподіл ресурсів з метою уникнення перевантаження системи.

Ефективність роботи. За цим критерієм буде оцінено наскільки точно та швидко компоненти системи виконують свої задачі. Важливо щоб система ефективно обробляла велику кількість логів, при найменших втратах оскільки навіть найменші втрати можуть призвести до неправильного аналізу та інтерпретації отриманих даних.

Перед тим як почати розробку компонентів системи, розглянемо особливості роботи з лог-файлами на двох найпопулярніших операційних системах: Windows і Linux.

Операційна система windows має вбудовані засоби збору і відображення лог-файлів які можна налаштувати під свої потреби. Основні типи журналів операційної системи windows включають в себе:

- event logs;
- audit logs;
- service logs.

Кожен з цих типів лог файлів містять в собі різну інформацію специфічну відносно типу логу. Для того шоб переглянути лог-файли можна скористатись одним з методів:

Powershell. PowerShell — це мова сценаріїв і набір інструментів для автоматизації завдань у Windows. Використовуючи PowerShell, ви можете виконувати операції з файлами журналів. Для того щоб запустити засіб перегляду журналів достатньо запустити командний рядок і виконати наступну команду: eventvwr, після виконання даної команди відкриється стандартний засіб перегляду лог-файлів(рис. 2.1).

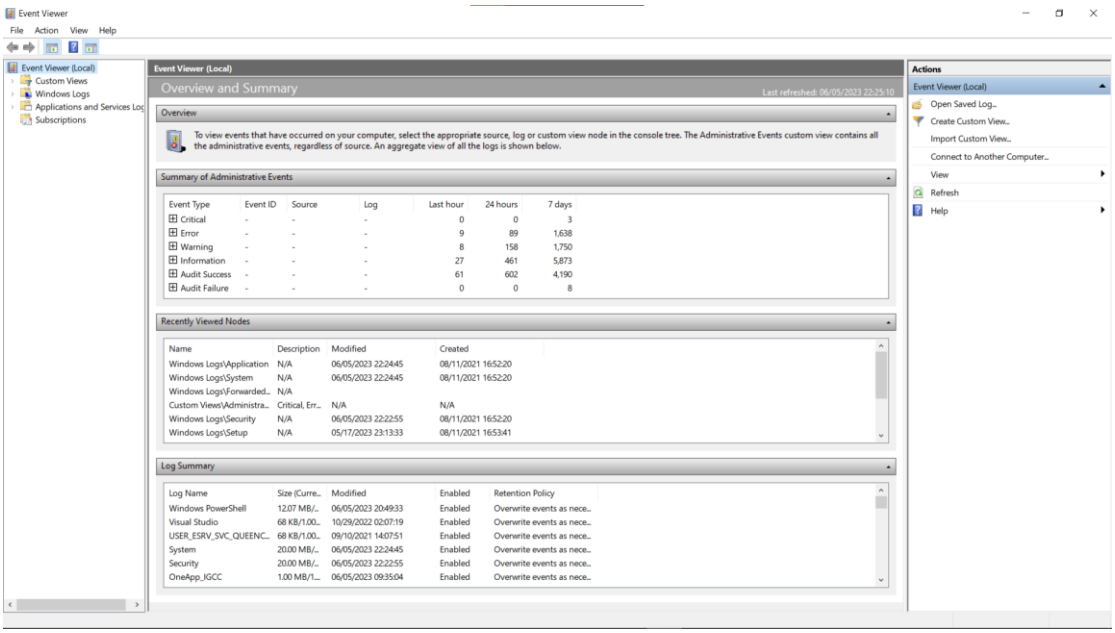

Рисунок 2.1 – Інтерфейс Event Viewer

Event Viewer дозволяє зберігати лог-файли в форматі .evtx, .xml, .csv і .txt.

Файли журналів операційної системи linux можна знайти за шляхом /var/log(рис. 2.2)

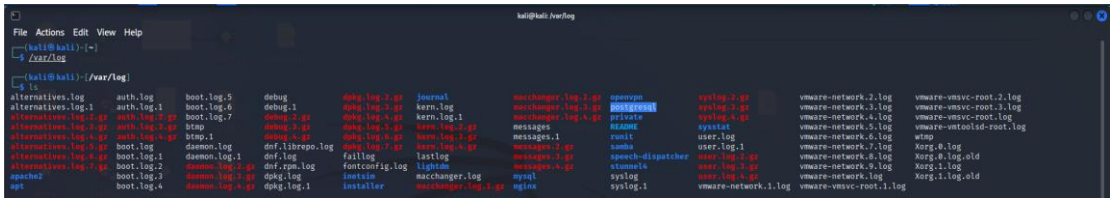

Рисунок 2.3 – Лог-файли linux

В операційній системі linux логи мають стандартний вигляд (рис 2.3)

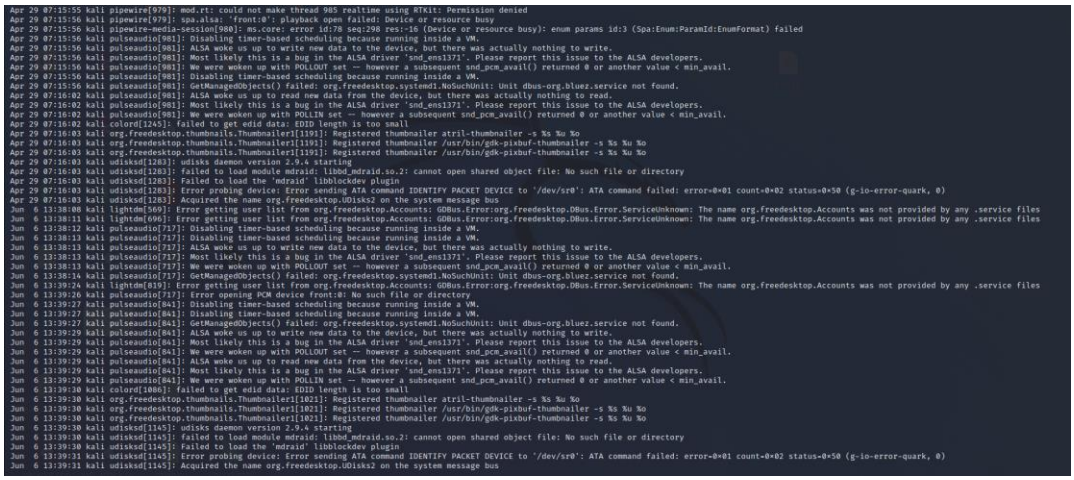

Рисунок 2.4 – Загальний вигляд журналів linux.

За допомогою скрипта мовою програмування python буде здійснюватись читання цих лог-файлів. Оскільки стандартний вигляд логів не зручний для аналізу і візуалізації даних які в ньому містяться, існують різні способи видобування даних. В цьому контексті – видобування даних з журналів це діставання окремих корисних даних які будуть цікаві аналітику і записування цих даних в поле з певною назвою, таким чином створюється словник. Одним з способів такого розбиття логів на словник є grok.

Grok – це збірник шаблонів для структурування неструктурованих даних отриманих з логів або інших текстових даних. Інструмент яким зручно можна працювати з цими патернами називається grok debugger (рис 2.5).

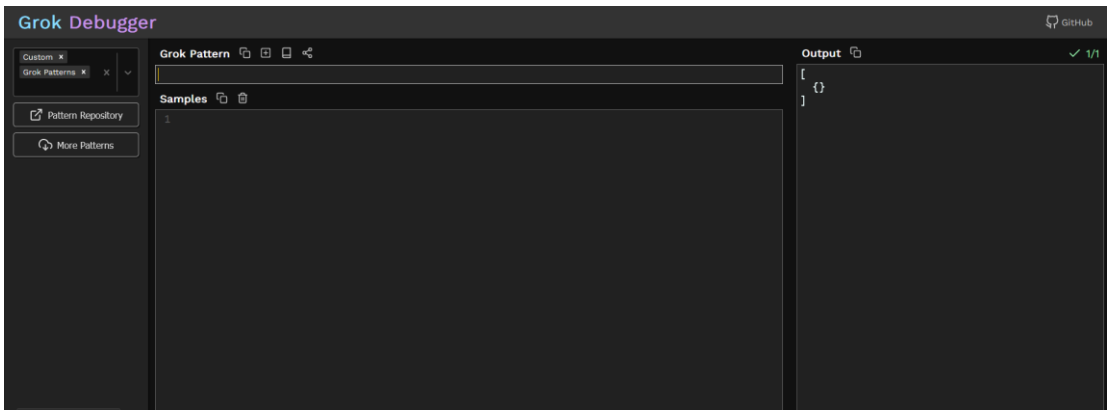

Рисунок 2.5 – Grok debugger

В залежності від структури журналу і даних в ньому використовуються різні патерни. Grok-патерни широко використовуються в автоматизованому аналізі

даних, зокрема в logstash що є частиною ELK stack (Elasticsearch, Logstash, Kibana). Сервіс надає певний список стандартних патернів який можна знайти на репозиторії сервісу (рис.2.6). Також ви можете додавати власні патерни за бажанням. Самі ж патерни являють собою регулярний вираз.

Рисунок 2.6 – Приклад стандартних патернів

Така кількість різних патернів зумовлена тим що в файлах журналів є різні типи даних, наприклад ip-адреси, числа і простий текст. Коли наперед відомо що в лозі буде IP-адреса хоста а не його ім’я краще використати “IP” патерн, якщо значення наперед передбачити неможливо, краще використати “NOTSPACE”, цей патерн опрацює стрічку до пробілу, цей патерн підійде, оскільки IP-адреси і назви хостів зазвичай закінчуються пробілом (рис.2.7).

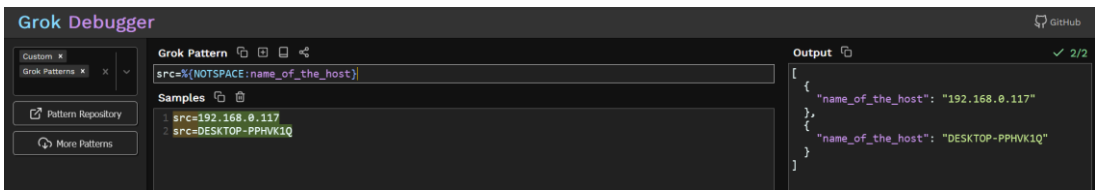

Рисунок 2.7 – Приклад використання патерну

Logstash – це конвеєр даних з відкритим кодом який дозволяє збирати, обробляти і зберігати структуровані дані з різних ресурсів в режимі реального часу. Хоч він і потребує більше обчислювальних ресурсів, але натомість пропонує велику кількість переваг.

Python-скрипт вимагає набагато менше обчислювальних ресурсів і легше встановлюється, але не має стільки можливостей як Logstash.

3 РОЗРОБКА ТА ТЕСТУВАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ АВТОМАТИЗОВАНОГО АНАЛІЗУ ЖУРНАЛІВ ПОДІЙ

3.1 Вибір мов програмування та інструментарію розробки

В якості мови програмування для розробки скриптів було обрано мову програмування Python. Це одна з найпопулярніших мов програмування яка має ряд переваг і відносно невелику кількість недоліків. До переваг обраної мови програмування можна віднести:

Простота синтаксису. Синтаксис мови програмування Python простий і зрозумілий, особливо добре це видно у порівняні з, наприклад, C++.

Велике і активне ком'юніті. Python – одна з найпопулярніших мов програмування яка використовується великою кількістю розробників. Це означає що для цієї мови набагато знайти потрібну бібліотеку, документацію чи відповіді на можливі питання на тематичних форумах.

Переносимість і інтеграція. Python працює на всіх популярних платформах, таких як windows, linux, mac os. Також завдяки активному ком'юніті мова програмування python має бібліотеки для роботи з багатьма інструментами які можуть знадобитись в розробці, наприклад бібліотеки для роботи з базами даних.

Для зберігання журналів буде використана база даних під назвою elasticsearch. Elasticsearch – документоорієнтована база даних яка використовує документи як одиницю даних. Основною причиною вибору цієї бази даних це її відкритий код та широка спільнота.

Для візуалізації і нормалізації журналів подій буде використано ELK-stack. Ця SIEM має ряд переваг які забезпечують зручну і ефективну роботу з автоматизації аналізу журналів подій.

3.2 Розробка архітектури системи та її складових

Система виявлення вторгнень буде складатись з 4 основних частин які взаємодіють між собою, кожен елемент якої виконує свою роль і по своєму важливий для повноцінного функціонування системи:

Аналіз даних. Модуль збирає дані і надсилає їх на SIEM. Як альтернатива python-скрипту може бути використана вбудоване в ELK-stack рішення під назвою filebeat, але оскільки цей конвеєр даних вимагає значно більших ресурсів, він не підходить.

Формування сповіщення. Модуль формує і надсилає сповіщення аналітику безпеки в разі виявлення аномалії чи загрози безпеці.

Зберігання даних. Після збору і аналізу, структуровані дані зберігаються в базі даних з метою їх подальшої візуалізації.

Візуалізація даних. Отримані дані які зберігаються в базі даних використовуються для створення дашборду.

Для того щоб система виявлення вторгнень працювала централізовано і могла збирати дані з декількох пристроїв одночасно, було прийнято рішення розділити систему на клієнтську і серверну частину. Завдання клієнтської частини системи це збір і надсилання даних журналів. В свою чергу, завдання серверної частини це приймання даних від клієнта, їх аналіз, зберігання в базу даних і подальша візуалізація.

Розробка клієнтської частини. Клієнтська частина буде працювати на linux і windows машинах, це потрібно врахувати оскільки під кожну операційну систему є окремий алгоритм. Збір лог-файлів буде проводитись шляхом пересилання файлу /var/log/syslog з клієнтської машини на сервер використовуючи python бібліотеку paramiko. Paramiko забезпечує безпечне надсилання файлів за допомогою протоколу SFTP. SFTP – це розширена версія протоколу SSH для безпечного доступу, надсилання і керування файлами.

Для передавання файлів журналів з операційної системи windows буде використана та ж сама бібліотека. Оскільки алгоритм збору лог файлів

відрізняється в залежності від операційної системи клієнта, необхідно зробити перевірку типу операційної системи, для цього буде використано бібліотеку platform (Лістинг 3.1).

Лістинг 3.1 – Перевірка типу операційної системи клієнта

```
def check_os_type():
    os_type = platform.system()
    return os_type
```

В залежності від операційної системи в функцію передачі файлів будуть передаватись різні аргументи (Лістинг 3.2).

Лістинг 3.2 – Передача файлу за допомогою безпечного протоколу SFTP

```
    def transfer_file_from_windows(remote_host, remote_username, 
remote_password, remote_file, local_path):
        # Ініціалізування SSH клієнту
        ssh = paramiko.SSHClient()
        ssh.set_missing_host_key_policy(paramiko.AutoAddPolicy())

        try:
            # Створення SSH з'єднання
            ssh.connect(hostname=remote_host, 
username=remote_username, password=remote_password)
            # Створення SFTP сесії
            sftp = ssh.open_sftp()

            # Передача файлу
            sftp.get(remote_file, local_path)

            print("File transfer successful!")

        except paramiko.AuthenticationException:
            print("Authentication failed. Please check your 
credentials.")
        except paramiko.SSHException as e:
            print("SSH error occurred:", str(e))
        finally:
            sftp.close()
            ssh.close()
```

Після того як файл журналів було отримано сервером, наступним етапом буде розбивання лог-файлів з допомогою Grok, це завдання серверної частини. Для кращої спостережності для кожного клієнта створюється окремий індекс. В якості бази даних для зберігання нормалізованих логів було використано

elasticsearch, що є частиною ELK-stack. Для написання патерну для логів буде використано сервіс “Grok debugger” (Рис. 3.1).

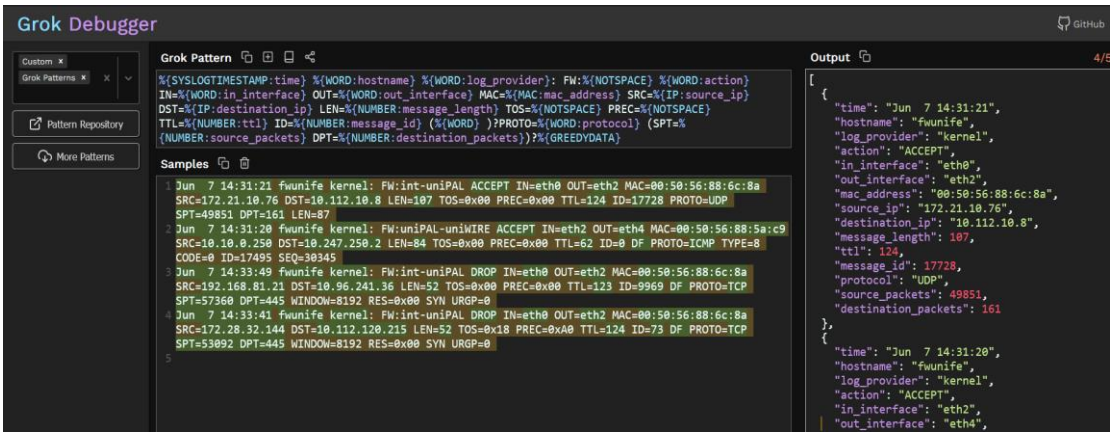

Рисунок 3.1 – Розбивання лог-файлу

Для того щоб парсити логи буде використано плагін “grok” що є функціональною частиною Logstash (Лістинг 3.2).

Лістинг 3.2 – скрипт з зразком патерну для аналізу лог-файлів операційної системи linux.

Лістинг 3.2 – Нормалізація логу з допомогою grok

```
grok {
        match => {
            "message" => [
                '%{SYSLOGTIMESTAMP:time} %{WORD:hostname}
%{WORD:log_provider}: FW:%{NOTSPACE} %{WORD:action}
IN=%{WORD:in_interface} OUT=%{WORD:out_interface}
MAC=%{MAC:mac_address} SRC=%{IP:source_ip} DST=%{IP:destination_ip}
LEN=%{NUMBER:message_length} TOS=%{NOTSPACE} PREC=%{NOTSPACE}
TTL=%{NUMBER:ttl} ID=%{NUMBER:message_id} (%{WORD}
)?PROTO=%{WORD:protocol} (SPT=%{NUMBER:source_packets}
DPT=%{NUMBER:destination_packets})?%{GREEDYDATA}',

            ]
        }
        tag_on_failure => ["_failure"]
    }
```

За допомогою цього скрипта відбувається автоматичне розбивання лог-файлу на ключ-значення, таким чином подаючи на вхід сирі лог-файли ми отримуємо структуризовані дані які можуть бути збережені в базу даних для подальшого використання (Рис 3.2).

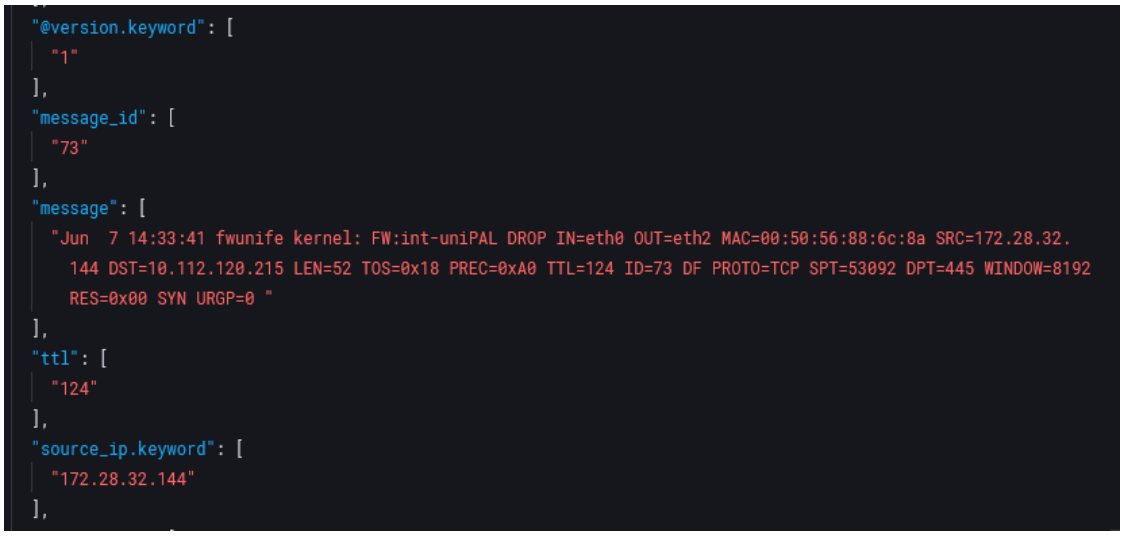

Рис 3.2 – Нормалізований лог-файл

Після нормалізації логів їх буде збережено до бази даних (рис 3.3), для логів з різних операційних систем було створено окремі колекції.

Time (@timestamp)	action	destinatio...	destina...	event.mod...	hostname	in_interface	mac_addre...	log_provider	message_id	message_l...	protocol	out_interfa...	source_ip	source_pa...
Jun 16, 2023 @ 14:25:42.972	DROP	10.112.120.2...	445	linux	fwunife	eth0	00:50:56:88:...	kernel	73	52	TCP	eth2	172.28.32.144	53092
Jun 16, 2023 @ 14:25:35.072	DROP	10.96.241.36	445	linux	fwunife	eth0	00:50:56:88:...	kernel	9969	52	TCP	eth2	192.168.81.21	57360
Jun 16, 2023 @ 14:25:24.710	ACCEPT	10.247.250.2	-	linux	fwunife	eth2	00:50:56:88:...	kernel	0	84	ICMP	eth4	10.10.0.250	-
Jun 16, 2023 @ 14:24:07.021	ACCEPT	10.112.10.8	161	linux	fwunife	eth0	00:50:56:88:...	kernel	17728	107	UDP	eth2	172.21.10.76	49851

Рисунок 3.3 – Збереження нормалізованих логів до бази даних

База даних повинна бути добре структурована щоб забезпечити спостережність даних для кожного клієнта. Для того щоб перевірити чи збереглись пропаршені логи до бази даних необхідно скористатись веб інтерфейсом і переглянути індекс в який було збережено нормалізовані логи (рис.3.4).

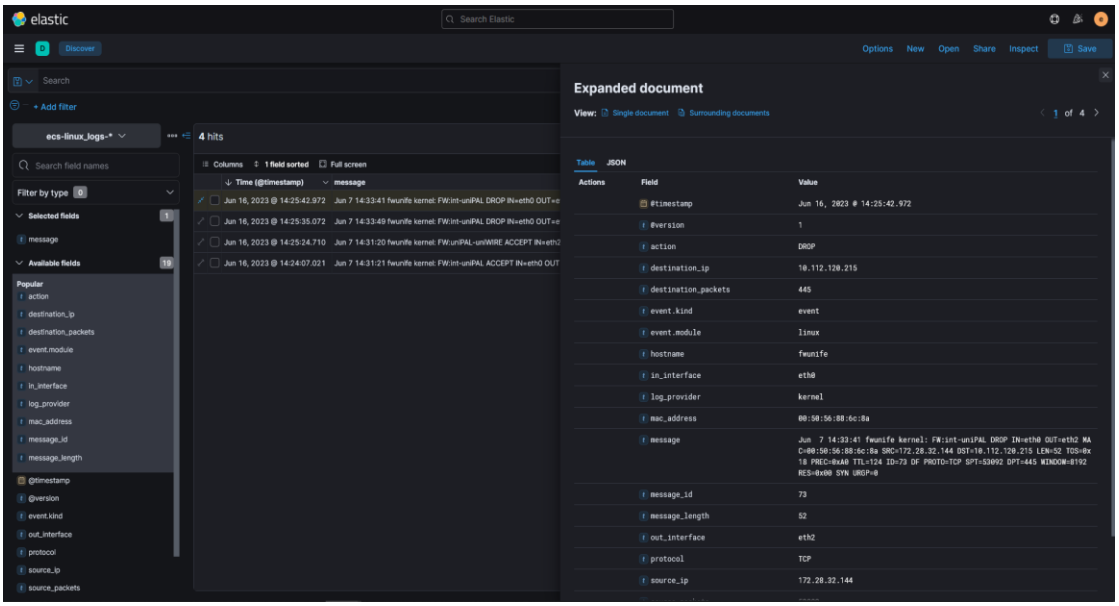

Рисунок 3.4 – Відображення вмісту бд з пропаршеними логами

Тепер коли лог-файли з Unix-подібних систем автоматично збираються, аналізуються і зберігаються в базу даних, настав час зробити те саме для лог-файлів операційної системи Windows, алгоритм відрізнятись не буде, але методи треба використовувати інші.

Почнем з того що визначимо які саме логи потрібні, в event viewer є можливість фільтрації логів за рівнем важливості. Ця особливість значно підвищить ефективність роботи системи, адже логи з рівнем “info” рідко бувають корисні в контексті виявлення аномалії чи вторгнень. Отже для ефективної роботи, будуть аналізуватись лише логи з рівнем warning, critical і error. Також логи операційної системи windows мають іншу структуру, тому патерни для неї треба створити інші (рис 3.5).

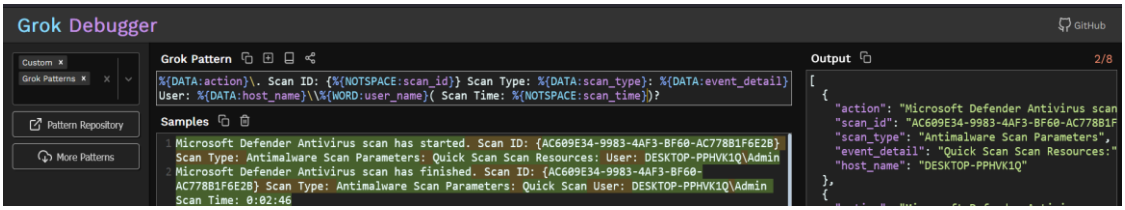

Рисунок 3.5 – Приклад патерну для логів Windows

Після цього створений патерн потрібно додати у нову конфігурацію для того щоб він був застосований до відповідних логів і згодом збережений у базу даних.

Важливо створити нову конфігурацію для клієнтів з операційною системою windows, оскільки ці операційні системи мають різну структуру логів та інформацію що в них міститься. Як видно з структури логів поле “message” має структуру json, в такому випадку можна скористатись спеціальним плагіном для полегшення роботи (лістинг 3.3)

Лістинг 3.3 – приклад конфігурації з патерном для Windows

```
json {
    source => "message"
}
```

Таким чином писати безліч патернів не потрібно, все зробить json плагін. Після проходження через цей плагін, нормалізований лог як і у випадку з журналами операційної системи Linux зберігаються у базу даних(рис 3.6)

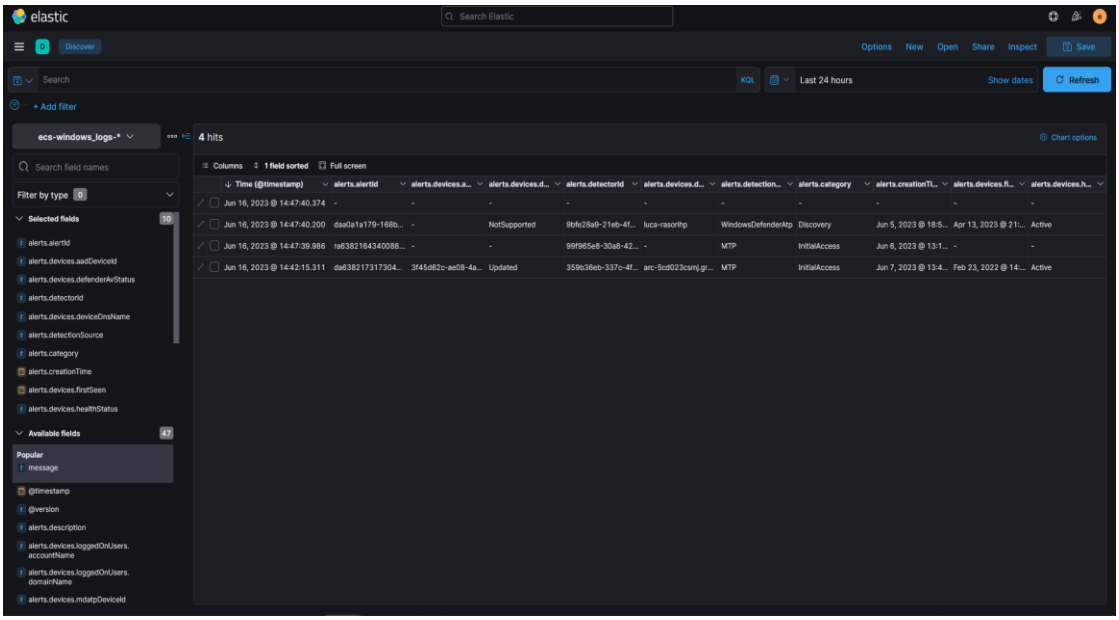

Рисунок 3.6 – Відображення вмісту бд з нормалізованими логами операційної системи Windows

3.4 Аналіз результатів тестування та їх інтерпретація

Після запуску скрипта і конвеєра даних Logstash нормалізовані логи будть автоматично збиратись і зберігатись в відповідній базі даних. Для полегшення аналітики і виявлення аномалій буде створено дашборд.

Завдання дашборду – відображення зібраної інформації для аналізу аналітиком кібербезпеки. Для початку створення дашборду необхідно перейти на відповідний розділ під назвою “Dashboard”. Після створення нового дашборду можна почати зповнювати його візуалізаціями(рис 3.7)

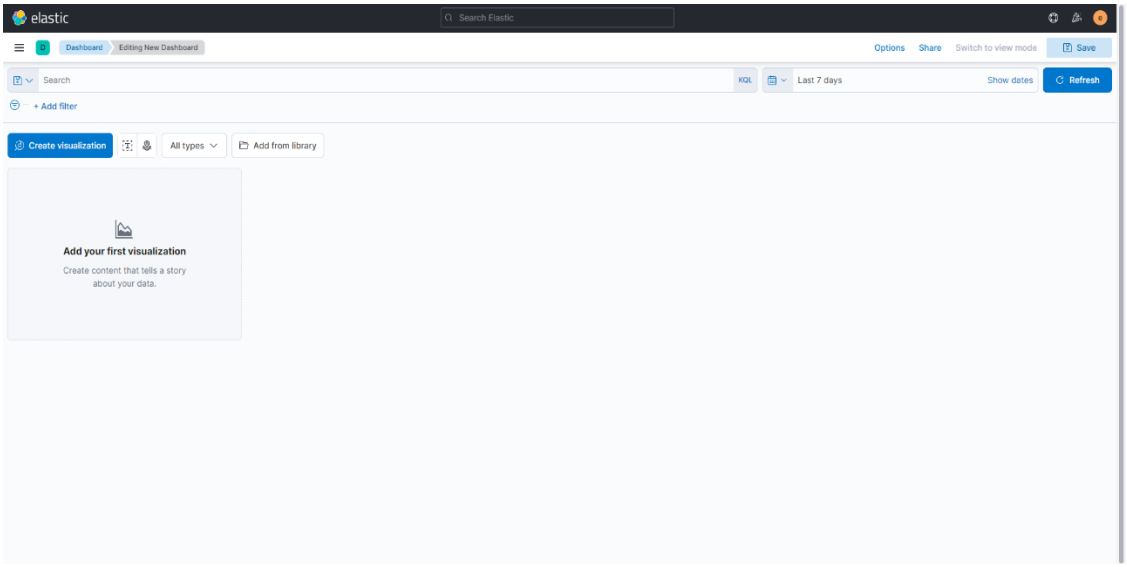

Рис 3.7 – створення дашборду

Kibana надає широкий набір інструментів для створення різних видів візуалізацій, таких як графіки, діаграми, мапи та діаграми. Ви можете налаштовувати параметри візуалізацій, такі як вибір поля даних, агрегації, фільтри та колірна палітра. Створення візуалізацій відбувається за допомогою редакторів, доступ до цих редакторів можна отримати різними способами, серед них: бібліотека візулізацій чи панель інструментів. Щоб створити візуалізацію спочатку потрібно натиснути “All types”(рис 3.8).

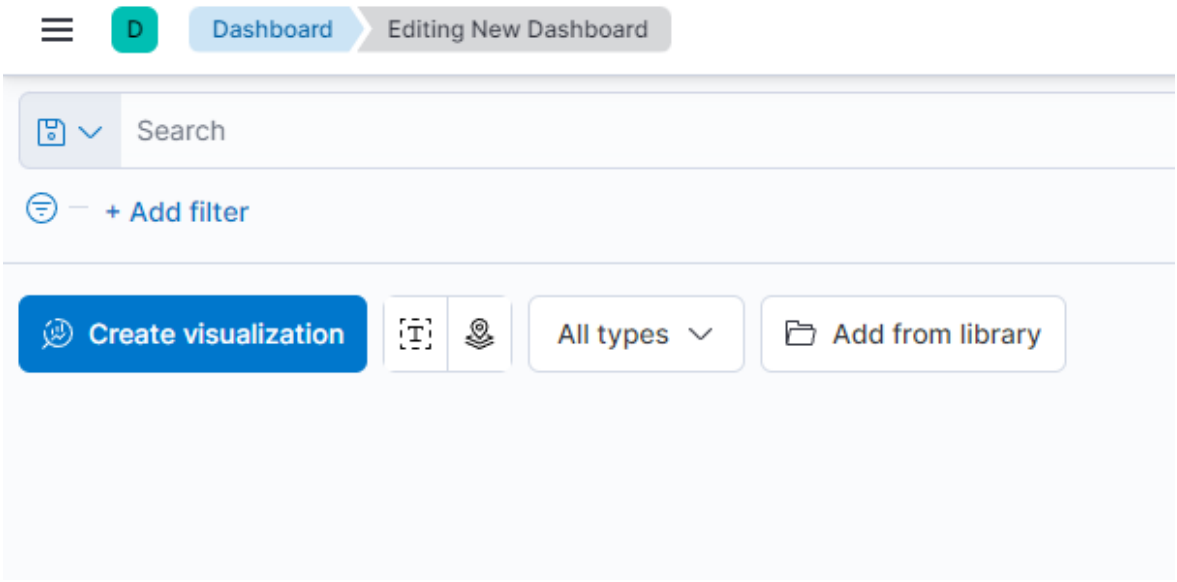

Рисунок 3.8 – Перший етап створення візуалізації

Після цього у випадаючому меню на вибір буде запропоновано один з методів створення візулізацій: Lens, Maps, Machine Learning, Log Stream, TSBV, Custom visualization, Aggregation based, Controls і text. Першою візуалізацією буде pie-chart створений за допомогою “Aggregation based”, який буде відображати поле alerts.severity(рис 3.8)

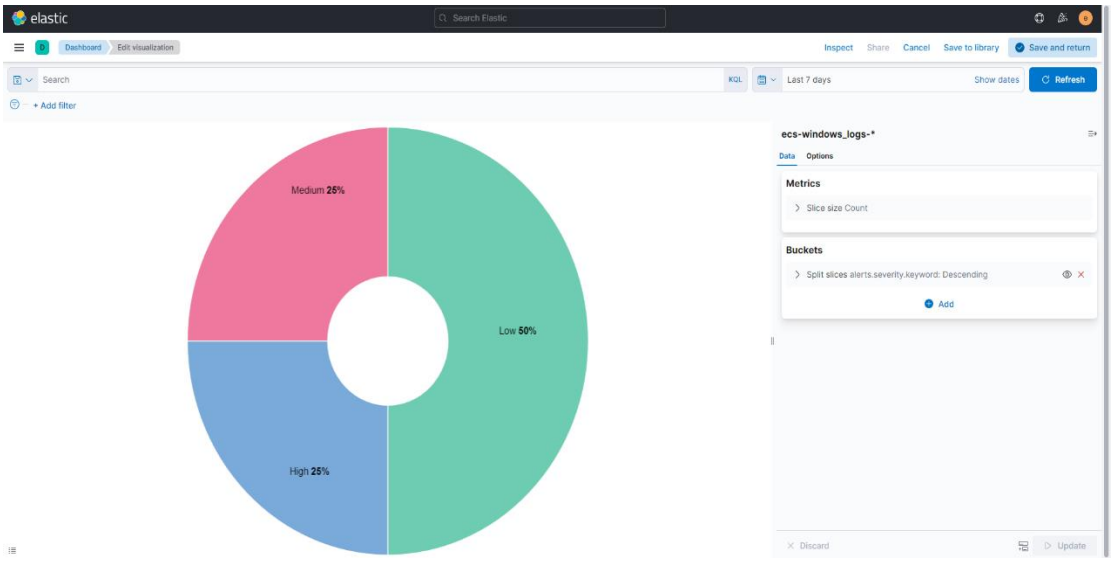

Рисунок 3.8 – Створення першої візуалізації

За допомогою цих інструментів було створено дашборд, який може бути використаний для зручного аналізу даних (рис 3.9).

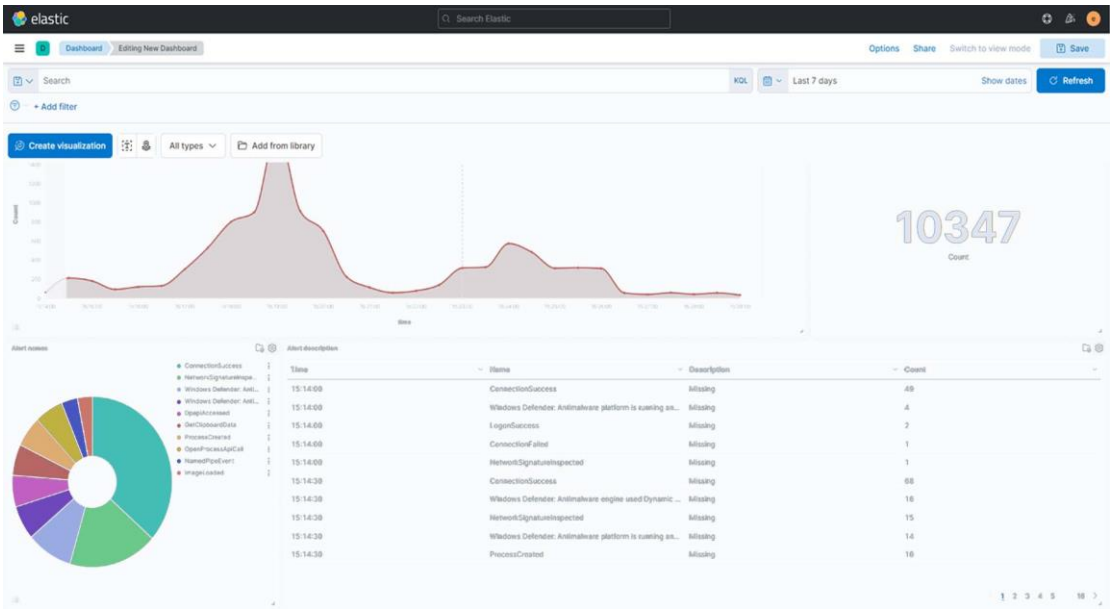

Рисунок 3.9 – Створена інформаційна панель

3.5 Оцінка ефективності розробленої системи

Система використовує методи машинного навчання для виявлення аномалій і загроз безпеки комп’ютерної системи. Для того щоб провести оцінку ефективності роботи розробленої системи необхідно створити штучні загрози і перевірити наскільки ефективно система їх виявляє. В інформаційній панелі було створено візуалізацію яка буде відображати інформацію про виявлені загрози, поки що вона пуста (рис 3.10), але після того як буде запущена тестова атака там повинна з’явитись інформація про подію.

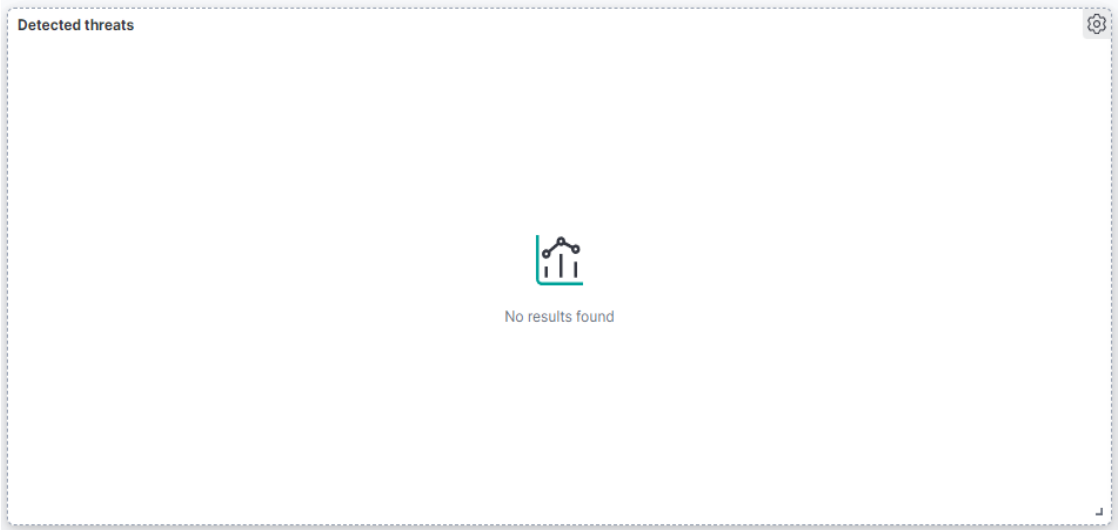

Рисунок 3.10 – Візуалізація в якій буде відображатись інформація про атаку

В рамках перевірки ефективності розробленої системи було проведено просту атаку розвідки, внаслідок цієї атаки має бути згенерований відповідний лог який система обробить. Для цього було використано команду: «sudo nmap -Pn -p62500,51653,53643,12660,16601,17237 -sC -sV -A -T4 192.168.1.4». Ця атака в агресивному режимі (агресивний режим збільшує шанс виявлення атаки) перевіряє задані порти по різних пунктах. Атака була спрямована на перевірку ефективності розробленої системи. Для того щоб перевірити чи система виявила атаку потрібно перевірити дашборд на предмет нових записів(рис. 3.11).

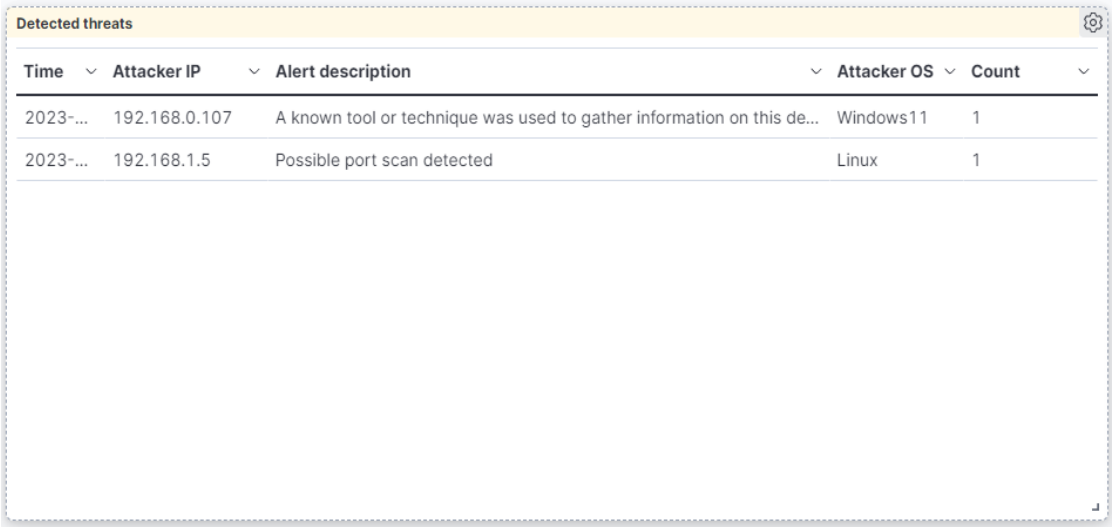

Рисунок 3.11 – Результат роботи системи

Як видно з рисунку 3.11 в таблиці з’явився новий запис з айпі адресою Linux машини з якої було запущено атаку, це свідчить про успішне спрацьовування системи автоматичного аналізу журналів подій.

4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ

4.1 Загальні вимоги безпеки з охорони праці для користувачів ПК

Користування комп'ютерами стало необхідністю в сучасному світі, особливо на робочому місці. Проте, тривале використання ПК може негативно впливати на здоров'я та самопочуття людей. З метою підтримки здоров'я та забезпечення безпеки користувачів ПК існують загальні вимоги безпеки з охорони праці.

Один з міжнародних стандартів, який визначає вимоги до робочого місця користувачів ПК, - це стандарт ISO 9241-5 "Ergonomic requirements for office work with visual display terminals (VDTs) - Part 5: Workstation layout and postural requirements". Цей стандарт надає рекомендації щодо розміщення та організації робочого місця. Згідно з цим стандартом, монітор повинен бути розташований на відстані 50-70 см від очей користувача. Висота столу має бути на рівні ліктів, а кут нахилу клавіатури повинен бути комфортним для зап'ястя. Крісло повинно мати підтримку для спини та регульовану висоту.

Крім стандарту ISO 9241-5, існують інші документи та рекомендації, що визначають вимоги безпеки для користувачів ПК. В багатьох країнах такі документи видані органами з охорони праці, а саме, Occupational Safety and Health Administration (OSHA) у США. OSHA надає настанови та рекомендації щодо безпеки та охорони праці під час роботи з комп'ютером.

Охорона зору є важливим аспектом безпеки користувачів ПК. У зв'язку з цим рекомендується дотримуватися наступних правил:

- забезпечити достатнє освітлення на робочому місці, уникати надмірного блиску та відблисків на екрані;
- регулярно робити перерви для відпочинку очей, зосереджуючись на далеких предметах протягом 10-15 секунд;

Ергономіка клавіатури та миші також має значення для безпеки та здоров'я користувачів ПК. Для цього необхідно:

- використовуйте клавіатуру з поділом на зони, що дозволяє знизити напругу на зап'ястях;
- миша повинна бути під рукою на такій висоті, щоб зап'ястя були в нейтральному положенні. використання миші з підтримкою зап'ястя допоможе уникнути травм та напруги.

Забезпечення безпеки даних також є важливим аспектом роботи з пк. Для цього потрібно:

- регулярно робити резервне копіювання важливих даних для запобігання їх втраті;
- встановлювати паролі для захисту доступу до комп'ютера та інших облікових записів;
- використовувати антивірусне програмне забезпечення для захисту від шкідливих програм та вразливостей.

Ці загальні вимоги безпеки з охорони праці, поєднані зі стандартами та рекомендаціями міжнародних страндартів, є спрямовані на зниження ризиків травм та покращення комфорту та продуктивності користувачів ПК. Дотримання цих вимог сприяє створенню безпечного та здорового робочого середовища для користувачів ПК.

4.2 Долікарська допомога при опіках.

Опіки є серйозними травмами, які можуть призвести до значних ушкоджень шкіри та негативно вплинути на здоров'я постраждалої особи. Надання належної долікарської допомоги при опіках є критично важливим для забезпечення швидкого загоєння, запобігання інфекціям та зменшення болю та дискомфорту.

Один із найважливіших стандартів, що визначає вимоги до долікарської допомоги при опіках, - це "Основні принципи надання долікарської допомоги" Червоного Хреста та Червоного Півмісяця. Цей стандарт надає важливі рекомендації та процедури для допомоги постраждалим в разі опіків. Згідно з цим стандартом, основні принципи долікарської допомоги при опіках включають:

- забезпечення безпеки: перш за все, необхідно усунути постраждалого з небезпечного середовища, вимкнути джерело опіку та забезпечити безпеку для самої постраждалої особи та допомагаючих осіб;
- оцінка тяжкості опіку: необхідно оцінити глибину і площу опіку, а також визначити, чи потрібна медична допомога. це допоможе визначити необхідні кроки для подальшого лікування;
- застосування охолодження: одразу після опіку необхідно здійснити охолодження ураженої області. нанесення прохолодної (не льоду) води на опік допоможе знизити біль та запобігти подальшому ушкодженню тканин;
- захист ураженої області: опік необхідно накрити чистою, негрубою тканиною або спеціальними пов'язками, щоб запобігти інфекції та подразненню.

Окрім стандарту Червоного Хреста та Червоного Півмісяця, існують інші важливі документи, які надають вказівки та рекомендації щодо долікарської допомоги при опіках, це організація "Всесвітня організація охорони здоров'я" (World Health Organization, WHO) надає детальні рекомендації з приводу надання першої медичної допомоги при опіках. Їхні рекомендації базуються на міжнародних клінічних протоколах та доказовій медицині. Документ ВОЗ "Надання першої медичної допомоги при опіках" надає детальний огляд процедур, пов'язаних з першою допомогою при опіках, включаючи очищення, охолодження та накриття опіку.

Організація "Американське товариство пластичних хірургів" (American Society of Plastic Surgeons) також надає рекомендації щодо долікарської допомоги та подальшого лікування при серйозних опіках. Ці рекомендації базуються на експертному досвіді пластичних хірургів та передових методиках лікування опіків.

Важливі принципи долікарської допомоги при опіках:

- безпека: безпека постраждалого і допомагаючих осіб є найважливішою. необхідно усунути постраждалого з небезпечного середовища та забезпечити безпеку перед наданням допомоги;

- охолодження: негайно після опіку необхідно нанести прохолодну воду на уражену область. це допоможе зменшити біль та запобігти подальшому ушкодженню тканин. важливо пам'ятати, що вода не повинна бути холоднішою за 15-25 ºc;
- накриття опіку: опік слід накрити чистою, негрубою тканиною або спеціальними пов'язками. це допоможе запобігти інфекції та забезпечити оптимальні умови загоєння;
- психологічна підтримка: опіки можуть бути дуже болісними та травматичними для постраждалих. надання психологічної підтримки є важливою складовою долікарської допомоги. необхідно спокійно спілкуватися з постраждалою особою, заспокоювати її та надавати впевненість у тому, що допомога надходить.

Долікарська допомога при опіках є критично важливою для швидкого загоєння, запобігання ускладненням та полегшення болю та дискомфорту постраждалих. Стандарти та документи, такі як "Основні принципи надання долікарської допомоги" Червоного Хреста та Червоного Півмісяця, рекомендації Всесвітньої організації охорони здоров'я (ВОЗ) та інших медичних організацій, надають важливі вказівки та протоколи для ефективного надання допомоги при опіках. Дотримання цих вимог та принципів є ключовим для забезпечення належної медичної допомоги та мінімізації ушкоджень та ускладнень у постраждалих осіб.

ВИСНОВКИ

В результаті виконання кваліфікаційної роботи було створено автоматизовану систему для аналізу журналів подій з метою виявлення зловмисної діяльності, аномалій та загроз безпеки комп’ютерній системі.

В ході виконання першого розділу кваліфікаційної роботи було детально розглянуто функції, що виконують системи виявлення і запобігання вторгнень, а також Security Information and Event Management (SIEM) - їх основні принципи та функції. Крім цього, було проаналізовано різні типи журналів, що використовуються для зберігання інформації про події, що відбуваються у системі, а також розглянуто різні методи аналізу цих журналів.

У другому розділі було проведено докладний аналіз питань, пов'язаних з автоматизацією процесу аналізу журналів подій. Було розглянуто різні підходи та методи, що використовуються для ефективного виявлення аномалій і потенційних загроз безпеці. Крім цього, були визначені ключові критерії, за якими можна встановити наявність аномалій і потенційних загроз у журналах подій.

Третій розділ кваліфікаційної роботи було присвячено розробці повноцінної системи автоматичного аналізу журналів подій. Було розроблено модуль для зчитування і пересилання журналів подій з клієнтської машини на сервер для подальшої обробки. Це було зроблено з метою зменшення навантаження на клієнтську машину, що забезпечить вищу ефективність роботи. Для обробки, зберігання і візуалізації даних було розгорнуто SIEM під назвою ELK-stack що включає в себе інструменти для обробки(Logstash), зберігання(Elasticsearch) та візуалізації(Kibana) даних. За допомогою Logstash було проведено нормалізацію журналів подій для їх ефективного аналізу. Elasticsearch було використано в якості бази даних для зберігання нормалізованих логів. З допомогою Kibana було реалізовано візуалізацію даних.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. SERVER LOG FILES IN NUTSHELL URL: https://www.graylog.org/post/server-log-files-in-a-nutshell/

2. Splunk vs ELK URL: https://logz.io/blog/splunk-vs-elk/

3. Splunk review URL: https://www.gartner.com/reviews/market/security-information-event-management/vendor/splunk/product/splunk-enterprise

4. Splunk Enterprise URL: https://www.splunk.com/en_us/software/splunk-enterprise.html

5. Офіційна документація Fluentd URL: https://www.fluentd.org/

6. Fluentd vs Logstash: A Comparison of Log Collectors URL: https://www.fluentd.org/

7. Fluentd for Centralized Logging URL: https://www.youtube.com/watch?v=Jb0fPN5u1IY

8. Elastic. ELK Stack: Introduction to Elastisearch, Logstash and Kibana URL: https://www.elastic.co/elk-stack

9. ELK Stack Tutorial URL: https://www.tutorialspoint.com/elk_stack/index.htm

10. What Is Server Security – and Why Should You Care? URL: https://www.avast.com/c-b-what-is-server-security

11. Ризик і Загроза URL: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf

12. Основні загрози безпеці комп'ютерної системи URL: https://sites.google.com/site/daindividualnosti/zagrozi-informacijnij-bezpeci

13. Intrusion Detection System (IDS) URL: https://www.techtarget.com/searchsecurity/definition/intrusion-detection-system

14. IDS – що це таке? Система виявлення вторгнень (IDS) як працює? URL: https://poradumo.com.ua/49510-ids-sho-ce-take-sistema-viiavlennia-vtorgnen-ids-iak-pracuye/

15. What is IDS and IPS? | Juniper Network US URL: https://www.juniper.net/us/en/research-topics/what-is-ids-ips.html

16. Що таке SIEM система? URL: https://ua.softlist.com.ua/articles/chto-takoe-siem-sistema/

17. IPS/IDS – системи виявлення і запобігання вторгнень URL: https://selectel.ru/blog/ips-and-ids/

18. Windows Logging Guide: The basics – Crowdstrike URL: https://www.crowdstrike.com/guides/windows-logging/

19. Перша допомога при опіках URL: http://ssmp.health.kiev.ua/index.php/porady-likaria/191-persha-dopomoga-pri-opikakh

20. Допомога при опіках URL: https://www.if.gov.ua/storage/app/sites/24/uploaded-files/uns-pp-dopomoga-pri-opikakh.pdf

21. Правила безпечної роботи на комп'ютері URL: https://pedcollege.kiev.ua/index.php/77-robota-koledzhu/okhorona-pratsi/589-pravyla-bezpechnoi-roboty-na-kompiuteri

22. Grok debbuger pattern repository URL: https://github.com/cjslack/grok-debugger/tree/master/public/patterns

23. Paramiko URL: https://www.paramiko.org/

24. APIDS URL: http://www.devopswiki.net/index.php/APIDS

25. Logichub SIEM Arhictecture URL: https://help.logichub.com/docs/logichub-siem-architecture

26. Intrusion Detection System (IDS) Types: The complete guide URL: https://nira.com/intrusion-detection-systems-ids-types/

27. Dashboard and visualizations URL: https://www.elastic.co/guide/en/kibana/current/dashboard.html