DE-42 (DC & SE)

YEAR 2024

NUST COLLEGE OF
ELECTRICAL AND MECHANICAL ENGINEERING

Arkon - Security Information & Event Management

PROJECT REPORT

DE-42 (DC & SE)

Submitted by

PC MUHAMMAD SALMAN

NS UZAIR SULTAN

ASC ABDULLAH TARIQ

NS HINA HAQ

BACHELORS

IN

COMPUTER ENGINEERING

YEAR

2024

Project Supervisor

DR. WASI HAIDER BUTT

DEPARTMENT OF COMPUTER & SOFTWARE ENGINEERING
COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING
NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,
ISLAMABAD, PAKISTAN

Certification

This is to certify that Muhammad Salman, Uzair Sultan, Abdullah Tariq and Hina Haq have successfully completed the final project Arkon Security Information & Event Management, at the College of Electrial & Mechanical Engineering, to fulfill the partial requirement of the degree Computer Engineering.

Dr. Wasi Haider Butt
Associate Professor

Sustainable Development Goals (SDGs)

SDG No	**Description of SDG**	**SDG No**	**Description of SDG**
SDG 1	No Poverty	**SDG 9**	**Industry, Innovation, and Infrastructure**
SDG 2	Zero Hunger	SDG 10	Reduced Inequalities
SDG 3	Good Health and Well Being	SDG 11	Sustainable Cities and Communities
SDG 4	Quality Education	SDG 12	Responsible Consumption and Production
SDG 5	Gender Equality	SDG 13	Climate Change
SDG 6	Clean Water and Sanitation	SDG 14	Life Below Water
SDG 7	Affordable and Clean Energy	SDG 15	Life on Land
SDG 8	Decent Work and Economic Growth	SDG 16	Peace, Justice and Strong Institutions
		SDG 17	Partnerships for the Goals

1 NO POVERTY
2 ZERO HUNGER
3 GOOD HEALTH AND WELL-BEING
4 QUALITY EDUCATION
5 GENDER EQUALITY
6 CLEAN WATER AND SANITATION
7 AFFORDABLE AND CLEAN ENERGY
8 DECENT WORK AND ECONOMIC GROWTH
9 INDUSTRY, INNOVATION AND INFRASTRUCTURE
10 REDUCED INEQUALITIES
11 SUSTAINABLE CITIES AND COMMUNITIES
12 RESPONSIBLE CONSUMPTION AND PRODUCTION
13 CLIMATE ACTION
14 LIFE BELOW WATER
15 LIFE ON LAND
16 PEACE, JUSTICE AND STRONG INSTITUTIONS
17 PARTNERSHIPS FOR THE GOALS
SUSTAINABLE DEVELOPMENT GOALS

Sustainable Development Goals

Complex Engineering Problem

Range of Complex Problem Solving

	Attribute	**Complex Problem**	
1	Range of conflicting requirements	Involve wide-ranging or conflicting technical, engineering and other issues.	✓
2	Depth of analysis required	Have no obvious solution and require abstract thinking, originality in analysis to formulate suitable models.	
3	Depth of knowledge required	Requires research-based knowledge much of which is at, or informed by, the forefront of the professional discipline and which allows a fundamentals-based, first principles analytical approach.	✓
4	Familiarity of issues	Involve infrequently encountered issues	
5	Extent of applicable codes	Are outside problems encompassed by standards and codes of practice for professional engineering.	✓
6	Extent of stakeholder involvement and level of conflicting requirements	Involve diverse groups of stakeholders with widely varying needs.	
7	Consequences	Have significant consequences in a range of contexts.	
8	Interdependence	Are high level problems including many component parts or sub-problems	

Range of Complex Problem Activities

	Attribute	**Complex Activities**	
1	Range of resources	Involve the use of diverse resources (and for this purpose, resources include people, money, equipment, materials, information and technologies).	✓
2	Level of interaction	Require resolution of significant problems arising from interactions between wide ranging and conflicting technical, engineering or other issues.	
3	Innovation	Involve creative use of engineering principles and research-based knowledge in novel ways.	
4	Consequences to society and the environment	Have significant consequences in a range of contexts, characterized by difficulty of prediction and mitigation.	
5	Familiarity	Can extend beyond previous experiences by applying principles-based approaches.	

Dedicated to our Parents, Project Supervisor Dr Wasi Haider Butt,& Our friends.

Acknowledgment

We would like to express our heartfelt gratitude to Dr. Wasi Haider Butt for his invaluable support and guidance throughout our journey. His assistance with every issue we faced was critical to our success, and his encouragement fueled our determination. We are extremely grateful for the positive impact he has had on our lives, and his contributions will never be forgotten.

Abstract

Security Information Event Management system abbreviated as SIEM is a cybersecurity software that helps organizations collect, analyze and correlates log activity data from various sources for the detection of potential security disruptions and provide real time alerts. It helps businesses in enhancing cybersecurity posture by managing security incidents, mitigating potential risks and ensuring compliance. However, the high costs and complexity of these advanced services often make it difficult for many Small and Medium-sized Enterprises (SMEs) to take advantage of them.

Our project intend to close this gap by providing SMEs with a SIEM solution that is both affordable and easy to use. This gives them the ability to see their IT infrastructure in real time, make threat detection easier with pre-established rules, and react to security threats more quickly with tools for centralized log management and incident response. In essence, our solution offers SMEs enterprise-grade security without the expense or complexity associated with enterprise-level solutions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

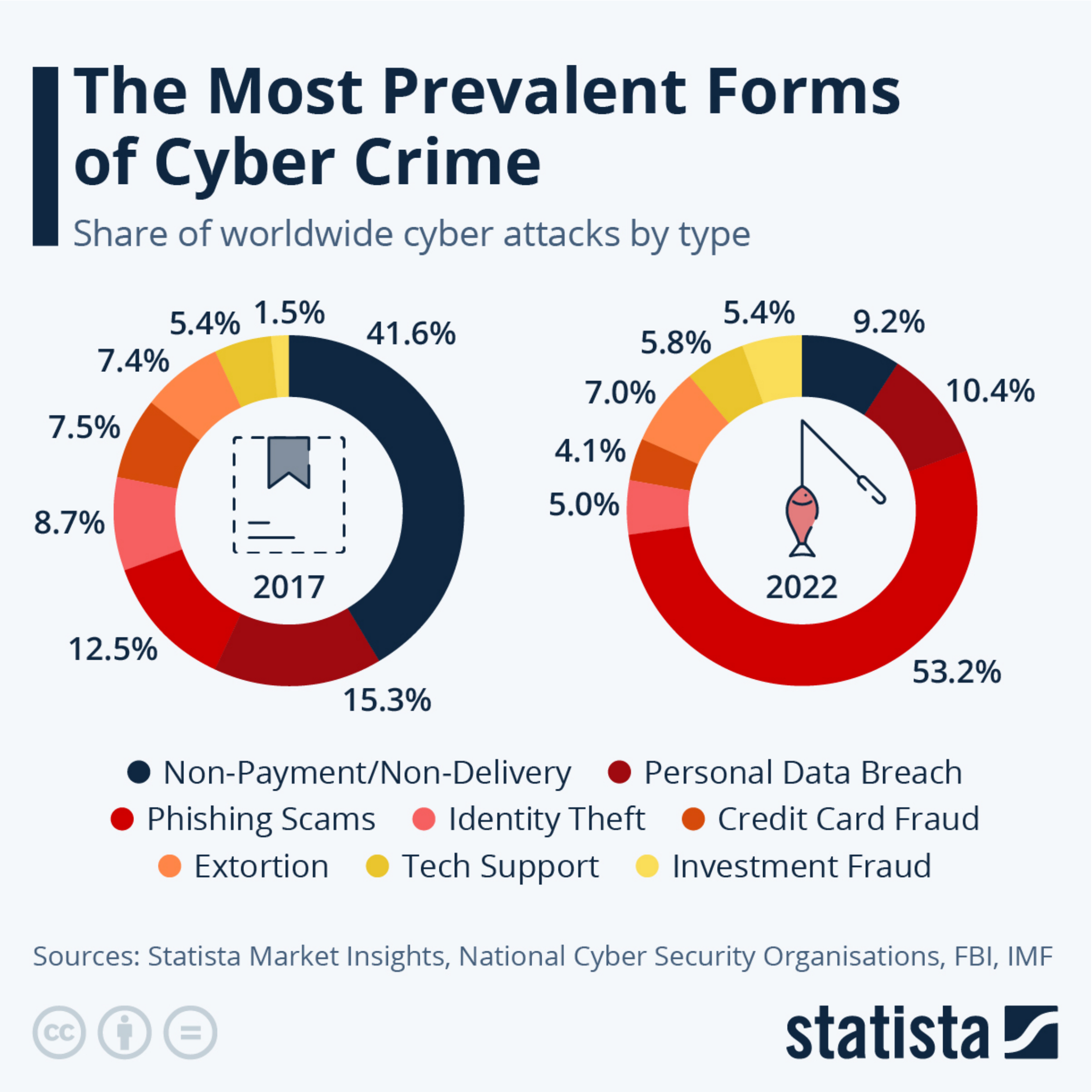
Small and Medium Enterprises(SMEs) represents 90% of total companies in the vast majority of economies worldwide and more provide more than 50% of the employment.61% of SMEs were the target of a cyber attacks in 2021.To defend themselves against more sophisticated cyber attacks in the digital economy, businesses must monitor and secure their data. Your organization probably has more information to gather and analyze than ever before. Before SIEM, security analysts would manually sift through millions of fragmented and siloed data bits for every app and security point. In summary, SIEM may accelerate the response and detection of cyber attacks, making security analysts' investigations more efficient and accurate.

1.2 Problem Statement

Small Medium Businesses face a range of cybersecurity risks including ransomware attacks,malware infections, phishing attacks, insider threats and data breaches.The successful execution of these risk can cause significant financial and reputation la damage to a business. Even though SIEM is an extensive of cybersecurity defence tool-set, most of the businesses cannot set it up on their own stating the reasons like it is time consuming to setup, it has too many complex configurations, high cost of initial setup and the intensive resources required. After installation the major issue is they don't have any idea which information is the most important to monitor. With no SIEM installed, SMEs become more vulnerable to cyber attacks due to lack of visibility into their IT infrastructure.The motivation for this project is to build a cost effective and user friendly SIEM solution for SMEs and bridge the security gap.

1.3 Scope

Cyber security threats are always on the rise and this makes the task of protecting organizations' IT systems a never-ending battle. Although SIEM solutions help large businesses to provide full-featured protection against threats, these systems are expensive and difficult to implement for SMEs. This means that SMEs are more vulnerable due to the lack of access to advanced security tools. Typical on-premise SIEM systems are expensive, require significant installation and configuration efforts and constant updates, making them unattainable for most SMEs due to the high costs involved.



Sources: Statista Market Insights, National Cyber Security Organisations, FBI, IMF

Figure 1.1: Statista report for 2022

1.4 Aims and Objectives

- Provide web activity tracking and logging in real-time
- Developing and Deploying Efficient data collection agents
- Easy agent installation
- Fast Detection of threats based on predefined rule sets
- User friendly interface

1.5 Outcomes

We have successfully built a system which collects event as soon as they are generated from window devices and stores them on database. Admin is able to access a website to view all logs and alerts generated by predefined queries.

1.6 Report Organization

The organization of the thesis is as follows:

- Chapter 2, explain available solutions and identify the gap created by them
- Chapter 3, briefly explain the technologies used in development of the product
- Chapter 4, contains the step by step development of the software
- Chapter 5 includes validation about the product
- Chapter 6 include conclusion and future work

Chapter 2

Literature Review

2.1 Security for Small & Medium Businesses

As the world shift toward IT enabled and IT controlled environment, it surely helps the organization to optimize their workloads,make better decisions and increase their productivity. That change introduce new vulnerabilities which if not mitigated properly, can become a major issue for Small & Medium Businesses.The consequences of not implementing cyber-security can lead to financial losses, significant reputation damage and even business closure.From data breaches to multiple day long shutdowns can occur if no security practices and tools are used to protect the organization.Many SMEs think that their size makes them less attractive for cyber-attacks in comparison to large organizations.Most attackers target SMEs because of their naivety to cyber-security practices, their limited access to advanced defence tools and the cyber fatigue caused by the complex layers of IT infrastructure.

Accenture's cyber crime study reveals that 43% of cyber-attacks were on small and medium businesses and only 18% of them were prepared to face them.

2.2 Popular Threat and Attack Types

With advent of new technologies like Generative AI, sophisticated attacks are done by the cyber criminals and to defend organizations now require more effort then ever.Being unaware from simple defence tools like Multi Factor Authentication(MFA) and Password managers lead to irreversible damage to SMEs.With respect to information, SMEs are also dealing with large amount of user data which they need to protect under regulations like GDPR,CCPA.

2.2.1 Supply Chain Attacks

A vast majority of devices and software used by small and large enterprises can be same, luring cyber criminals towards attacking small businesses learn about a system vulnerability and attack a larger corporation.Cyber criminal groups deliberately target a large enterprise's smaller partners in supply chain attacks, also known as "island hopping" attacks, with the goal of gaining access to the larger company's data.

2.2.2 Phishing

Phishing is one of the most common and successful cyber assaults against small and medium-sized organizations. Phishing has many types but the most common ones are spear-phishing and business email compromise. Since 2020, 81% of companies worldwide have experienced an upsurge in phishing attempts. It is believed that 82% of all data breaches can be traced back to an initial phishing assault.[1]

2.2.3 Ransomware

Ransomware is a type of malicious software designed to encrypt a victim's file, making them inaccessible to access until a ransom is paid to the attacker usually in a cryptocurrency. Ransomware attacks have expanded to include double- and triple-extortion attacks, raising the stakes significantly.[2]Double-extortion attacks include the threat of stealing and leaking the victim's data online. Furthermore, triple-extortion attacks target victim's business partners or customers using teh stolen data.

Figure 2.1: image depicting ransomware attacks

2.2.4 Malware

Malware, or malicious software, is any intrusive program developed by cybercriminals (often referred to as hackers) with the intent to steal data and harm or destroy computers and computer systems. Common types of malware include viruses, worms, Trojan horses, spyware, adware, and ransomware.

2.2.5 Patching

Patch management involves ensuring that all IT devices within your organization are updated with the latest software versions. Outdated software can be compromised by known loopholes, which malware can exploit. Consequently, inadequate patch management can significantly increase the risk of a data breach for your business. According to a report by Heimdal Security, 18% of all vulnerabilities result from unpatched software.[3]

2.2.6 Insider Threats

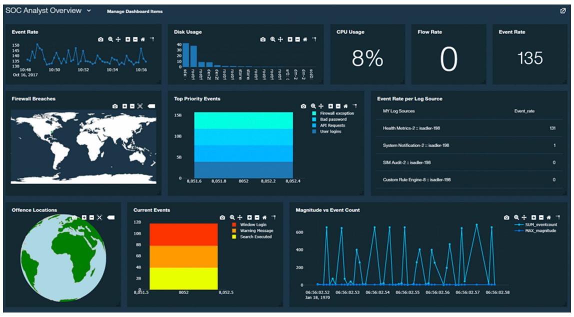
An insider threat is a risk to an organization stemming from the actions of current or former employees, business contractors, or associates. These individuals have access to sensitive company information and can potentially cause harm through greed, malice, ignorance, or carelessness.[2]

2.3 Available Products

There are lot of SIEM solutions available to store logs in a Central repository, detect threats and secure organizations.Sadly, there is not a one-size fits all product available. Here is a list of popular software commonly used by corporations

2.3.1 IBM QRadar SIEM

IBM QRadar SIEM allows you to supervise your IT systems in real time. It is the type whose design enables threat identification and distinction based on levels of risk. It logs the activity of multiple protocols and allows for multiple settings, along with having features such as analytics.[4]



Event Rate per Log Source

MY Log Sources	Event_rate
Health Metrics-2 :: isadler-198	131
System Notification-2 :: isadler-198	1
SIM Audit-2 :: isadler-198	0
Custom Rule Engine-8 :: isadler-198	0

Figure 2.2: IBM Qradar SOC Analyst Dashboard

2.3.2 Splunk

Splunk offers a popular SIEM solution. It handles both security and application/network monitoring scenarios making it popular among Security Professionals. Splunk's SIEM provides real-time information and has a user-friendly interface.

Figure 2.3: Splunk Logo

2.3.3 LogRhythm

LogRhythm is one of the oldest SIEM dealers and it has earned its name over the years. Popularly known as log data analysis and SIEM, open source LogRythm also uses numerous analytical tools, AI, and log correlation. Despite this, integration with LogRhythm is easy which makes it quite easier as compared to other SIEMs but as compared to other SIEMs, this has steeper learning curve.

Figure 2.4: LogRythm Logo

2.3.4 Microsoft Azure Sentinel

Released in 2019, Microsoft Azure Sentinel is a relatively new product in SIEM market.It is a popular choice for users of Microsoft products giving them ability to analyze access under one shade.

Figure 2.5: Azure sentinel Logo

2.4 Problems Faced by SMEs While Selecting SIEM

Even though existing solution helps us organizations safeguard their infrastructure but it still have some shortcomings.

- **High Costs:** Existing SIEM solutions such as Splunk, QRadar, and Azure Sentinel frequently come at a high cost for small and medium businesses. Licensing, maintenance, and operational costs can be significant financial burdens.

- **Complex Implementation:** Many enterprise-level SIEM systems have complex implementation processes that require specialized knowledge and significant time investment, making it difficult for SMEs with limited IT staff.

- **Resource Intensive:** SIEM systems can be resource-intensive, requiring significant computational power and storage, which may exceed the capabilities of SMEs' current infrastructure.

- **High False Positives:** Enterprise SIEM solutions frequently generate a large number of false positives, which leads to alert fatigue. SMEs with a small staff may struggle to manage and triage these alerts efficiently.

- **User Unfriendliness:** Many SIEM tools have complex user interfaces that are not intuitive, making it difficult for SMEs to navigate and use without extensive training.

Chapter 3

Material & Component

3.1 System Overview

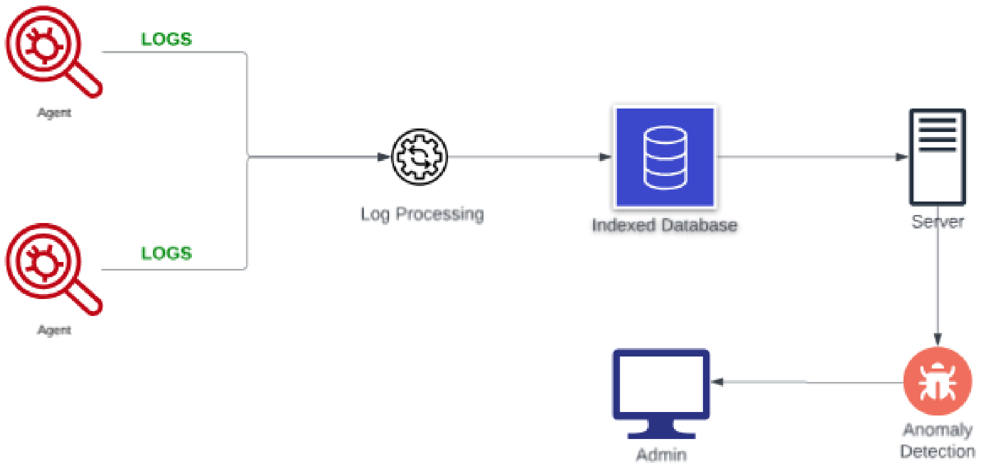


Figure 3.1: System Level Diagram

This is a simple explanation of our project. Logs are collected from windows devices and browser activity is collected via a chrome extension built using JavaScript. Agents collect logs and store them into relevant databases from where threat detection queries are used to extract elements from event details which can be dangerous.

3.1.1 Objective

To build an easy to use and effective security system for Small and Medium Enterprises helping them secure their IT infrastructure.

3.2 Tech Stack

Here is a little introduction to every language we used in developing the system. These technologies are widely used in modern web development to ensure compatibility and ease of maintenance.

3.2.1 ReactJS

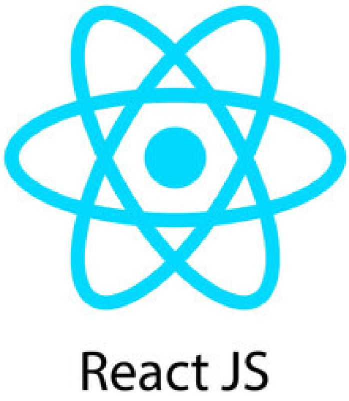


Figure 3.2: React Logo

React JS is a free and open source JavaScript library for front end maintained by meta formerly Facebook. It is used by developers to build Single Page Applications(SPAs), mobile applications and used with frameworks like NEXTJS to build server rendered applications.React works using principle of using modular and reusable components to build pages.According to Stack Overflow Trends,React is the most popular tool among developers according to no of questions asked monthly,

3.2.2 Redux

Figure 3.3: Redux LOgo

As the pages in react is build using components which can be readily used whenever we need them in another page.We also need to maintain the them state-full and stateless components.Stateless are those whose values are static and does not change with user inputs and interactions.But for maintaining state-full components inside react, Redux is used. It provides complex process optimization,encapsulated Application Programming Interfaces,and predictable values in the components with changing states.

3.2.3 NodeJS

Figure 3.4: Node JS Logo

Node.js is an open-source, cross-platform JavaScript runtime and a popular tool for a wide range of tasks. It uses Google Chrome's V8 JavaScript engine outside of the browser, which makes Node.js extremely fast. A Node.js application runs in a single process, rather than creating a new thread for each request. The Node.js standard library provides a set of asynchronous I/O primitives that prevent JavaScript code from blocking. Libraries in Node.js are typically designed following non-blocking paradigms, making blocking behavior the exception rather than the rule.

3.2.4 ExpressJS

Figure 3.5: Express Logo

Express JS, or simply Express, is a back end web application framework for developing representational state transfer(REST)ful Application Programming Interfaces with

NodeJS.It has been called the de-facto standard server framework for node JS.Developers can define routes to handle incoming requests and data and send out responses, based on handling different URLs and HTTP verbs like GET,POST.In short, express ought to help developers clean,scalable and maintainable applications faster.

3.2.5 MongoDB

Figure 3.6: MongoDB Logo

MongoDB is a popular No-SQL database that offers a flexible and salable solution for data management and storage. It employs a document-based model in which data is stored in JSON-like documents, making it simple to work with for developers. It is ideally used for agile development due to its natural dynamic schema allowing developers to easily modify data structure without any downtime.It enables efficient data retrieval by its indexing capabilities and powerful query language.

3.2.6 FastAPI

Figure 3.7: fastAPI Logo

FastAPI is a cutting-edge, highly efficient web framework used for creating Python APIs. Its asynchronous features and automatic OpenAPI and JavaScript Object Notation(JSON) Schema production make it the perfect option for building reliable back end systems. FastAPI's user-friendly syntax and compatibility with Python 3.6+ make Swift API development and maintenance easier, ensuring compatibility with the newest language features.

3.2.7 Clickhouse

Figure 3.8: ClickHouse Logo

Click-House is an open-source powerhouse built for handling massive datasets and complex analytical queries.As it is a column-oriented database,it is way faster in performing aggregation and filtering tasks. It is mainly used in scenarios where faster online analytical processing(OLAP) is required, OLAP main requirement is used to analyze billions of data points within seconds and click-house fulfills it very comfortably..

3.2.8 Heroku

Figure 3.9: Heroku Logo

Heroku is a cloud provider that specializes in delivering cloud services for consumer-focused applications with high speeds, low latency, and the ability to scale without performance degradation. It's an excellent choice for companies on a tight budget or those looking to explore cloud opportunities due to its straightforward setup. The Heroku platform is built on a managed container system (referred to as dynos in Heroku's terminology) and includes integrated data services along with a robust ecosystem for deploying and running modern applications.

Chapter 4

Methodology

4.1 Data Collection

The first part of the system is collecting logs from devices and endpoints.Collected data is then parsed and normalized to extract the important information.All version of windows maintain three main event logs:

- Application
- Security
- System

Each channel has its own importance and help in detecting related errors and threats. This picture depicts the structure of agent files executing on the back-end collecting events and sending them to server.

4.1.1 win32evtlog library

The **win32evtlog** module is a pywin32 library which was primarily created by Mark Hammond and to this day still maintained by him.It helps developers accessing and interact with windows operating system via Windows Application Programming Interfaces.The

PyWin32 library has been developed and maintained for many years, with contributions from a variety of open-source developers. We followed multiple strategies for collecting window events,mentioned here with full functional details. Functions that I used from win32evtlog module.

- **OpenEventLog** accept two arguments:(Server,Source) where server is the computer location where event log is stored,null means local computer.Source is the name of specific log channel.The function returns a handle to the specified event log, which is used by other event log functions to perform operations on the log.
- **EVENTLOG-BACKWARDS-READ** is a constant used as a flag to read events in reverse chronological order i.e most recent to oldest event.
- **EVENTLOG-SEQUENTIAL-READ** specifying that events should be read in sequential order
- **ReadEventLog** accepts three arguments:handle which is returned by OpenEventLog,event flags specifying the read orders,and offset displaying record number where to start from.
- **GetOldestEventLogRecord** obtains the oldest record's record number from the event log.It is helpful in determining the starting point of reading the logs
- **CloseEventLog** closes an opened event log handle and release resources associated with it.

4.1.2 Collection via Comma Supported Value(CSV) files

In this method, we were collecting window events using win32evtlog and dumping them into a csv file.After the dump is completed,upload using clickhouse file upload option.Due to naivety of the approach, it was disregarded.

4.1.3 Collection via channel specific scripts

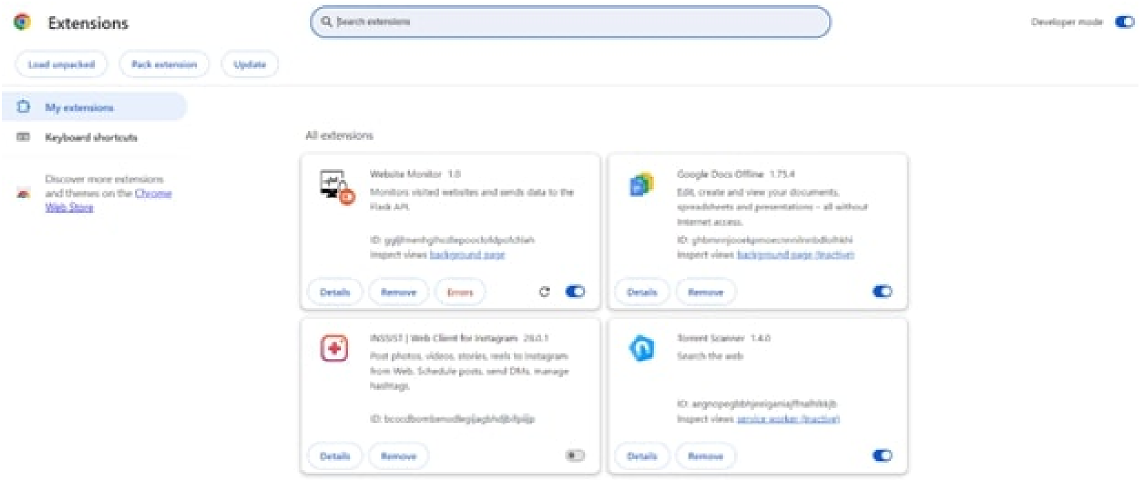
By installing ClickHouse-Connect module, python allow devices to connect with database using http protocol.After execution of script logs were successfully transferred to ClickHouse table making it ready for querying and other operations.The drawback of this approach was that separate files were running in the background and data writes to database were not concurrent.

4.1.4 Collection via executable

After trial and error, the best suitable approach was to build an executable of python scripts to combine the different log extraction process.Threading was used for collecting events simultaneously.Now a simple executable can be run on any PC for collecting logs and storing into the respective tables.

4.1.5 Chrome Extension:

The Chrome extension is used to get information about the websites user is visiting . It allows the system to directly track and record the user's interactions and the websites they visit in the browser.It does not make the user experience bad by showing notifications time by time, This ensures the tracking happens with minimal disruption to the user. The extension's ability to interact with web pages and collect relevant data gives the system valuable insights into user behavior and website usage patterns.



Website Monitor 1.0
Monitors visited websites and sends data to the Flask API.
ID: ggjlhmenhghcdlepooclofdpofchlah
Inspect views background page

Google Docs Offline 1.75.4
Edit, create and view your documents, spreadsheets and presentations – all without Internet access.
ID: ghbmnnjooekpmoecnnnilnnbdlolhkhi
Inspect views background page (Inactive)

INSSIST | Web Client for Instagram 28.0.1
Post photos, videos, stories, reels to Instagram from Web. Schedule posts, send DMs, manage hashtags.
ID: bcocdbombenodlegijagbhdjbifpiijp

Torrent Scanner 1.4.0
Search the web
ID: aegnopegbbhjeeiganiaffnalhlkkjb
Inspect views service worker (Inactive)

Figure 4.1: Extension installed

4.2 Log Storage

Within a SIEM system,a crucial component for comprehensive security analysis is log storage.Data generated by end devices is growing exponentially and so does the need of storing that data, along with the high volume of generated data another issue is the velocity of data creation.So for SIEM systems, storage server has to make sure that information must be readily accessible, provides efficient querying capabilities without flow breakdown and can scale with the inflow of data.

Keeping these tough requirements in mind,we have opted to use click-house which is a columnar database used for online analytical processing(OLAP) tasks.Its high performance capabilities enable it to handle large-scale data ingestion with low latency, ensuring that logs are stored in real time.

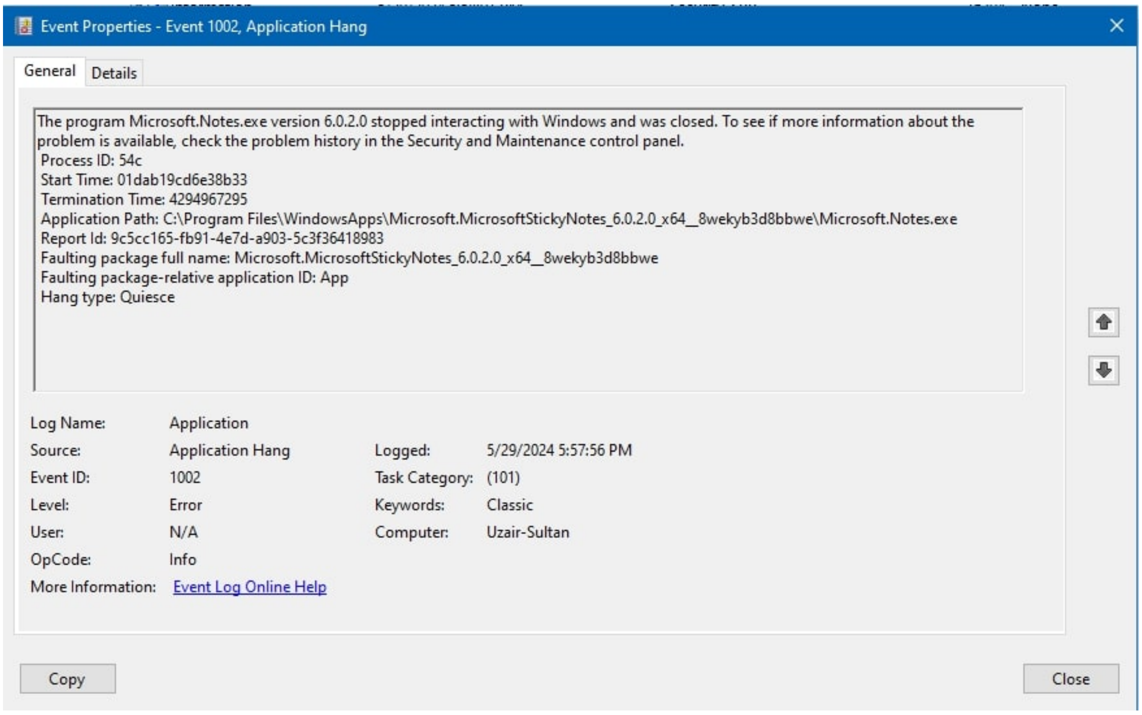
4.2.1 Tables

Instead of creating one big database table ,Separate tables were created for collected logs to store them as they come.This allow faster data retrieval of queried data as the overall size of aggregation is reduced.

- Application
- Security
- System

4.2.2 Log Format

The format for all events is same.The categorization of logs is done via special keywords related to the Event ID. Log in event viewer looks like this:

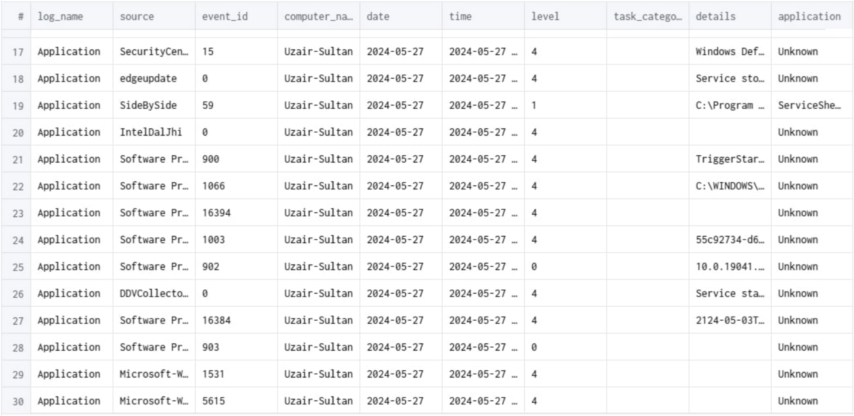


```
The program Microsoft.Notes.exe version 6.0.2.0 stopped interacting with Windows and was closed. To see if more information about the
problem is available, check the problem history in the Security and Maintenance control panel.
 Process ID: 54c
 Start Time: 01dab19cd6e38b33
 Termination Time: 4294967295
 Application Path: C:\Program Files\WindowsApps\Microsoft.MicrosoftStickyNotes_6.0.2.0_x64__8wekyb3d8bbwe\Microsoft.Notes.exe
 Report Id: 9c5cc165-fb91-4e7d-a903-5c3f36418983
 Faulting package full name: Microsoft.MicrosoftStickyNotes_6.0.2.0_x64__8wekyb3d8bbwe
 Faulting package-relative application ID: App
 Hang type: Quiesce

Log Name:          Application
Source:            Application Hang        Logged:          5/29/2024 5:57:56 PM
Event ID:          1002                    Task Category:   (101)
Level:             Error                   Keywords:        Classic
User:              N/A                     Computer:        Uzair-Sultan
OpCode:            Info
More Information:  Event Log Online Help
```

Figure 4.2: Image displaying an event in EventViewer

One event file can be easily read but to read thousands of them and apply queries the process becomes slow and resource consuming due to continuous opening and closing of event files.To cover this issue, all of the log files are extracted into database using python agents running in the background, working without input after short intervals.Here is an example of row being stored:



#	log_name	source	event_id	computer_na…	date	time	level	task_catego…	details	application
17	Application	SecurityCen…	15	Uzair-Sultan	2024-05-27	2024-05-27 …	4		Windows Def…	Unknown
18	Application	edgeupdate	0	Uzair-Sultan	2024-05-27	2024-05-27 …	4		Service sto…	Unknown
19	Application	SideBySide	59	Uzair-Sultan	2024-05-27	2024-05-27 …	1		C:\Program …	ServiceShe…
20	Application	IntelDalJhi	0	Uzair-Sultan	2024-05-27	2024-05-27 …	4			Unknown
21	Application	Software Pr…	900	Uzair-Sultan	2024-05-27	2024-05-27 …	4		TriggerStar…	Unknown
22	Application	Software Pr…	1066	Uzair-Sultan	2024-05-27	2024-05-27 …	4		C:\WINDOWS\…	Unknown
23	Application	Software Pr…	16394	Uzair-Sultan	2024-05-27	2024-05-27 …	4			Unknown
24	Application	Software Pr…	1003	Uzair-Sultan	2024-05-27	2024-05-27 …	4		55c92734-d6…	Unknown
25	Application	Software Pr…	902	Uzair-Sultan	2024-05-27	2024-05-27 …	0		10.0.19041.…	Unknown
26	Application	DDVCollecto…	0	Uzair-Sultan	2024-05-27	2024-05-27 …	4		Service sta…	Unknown
27	Application	Software Pr…	16384	Uzair-Sultan	2024-05-27	2024-05-27 …	4		2124-05-03T…	Unknown
28	Application	Software Pr…	903	Uzair-Sultan	2024-05-27	2024-05-27 …	0			Unknown
29	Application	Microsoft-W…	1531	Uzair-Sultan	2024-05-27	2024-05-27 …	4			Unknown
30	Application	Microsoft-W…	5615	Uzair-Sultan	2024-05-27	2024-05-27 …	4			Unknown

4.3 User Interface

4.3.1 Login

Existing users can log themselves into the application using this simple sign-in page.This is done to make sure only authorized users access our system.It asks user for two things:

1. email
2. password

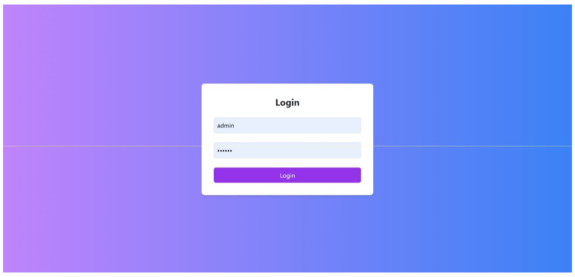
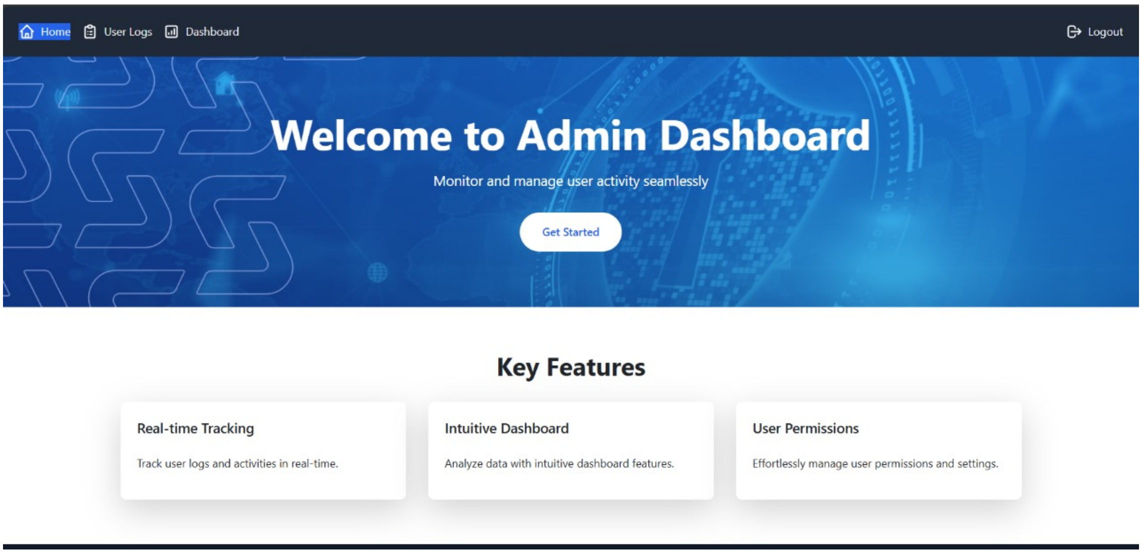


Figure 4.3: Login Page

4.3.2 Homepage



Welcome to Admin Dashboard

Monitor and manage user activity seamlessly

Key Features

Real-time Tracking

Track user logs and activities in real-time.

Intuitive Dashboard

Analyze data with intuitive dashboard features.

User Permissions

Effortlessly manage user permissions and settings.

Figure 4.4: Homepage of web monitor

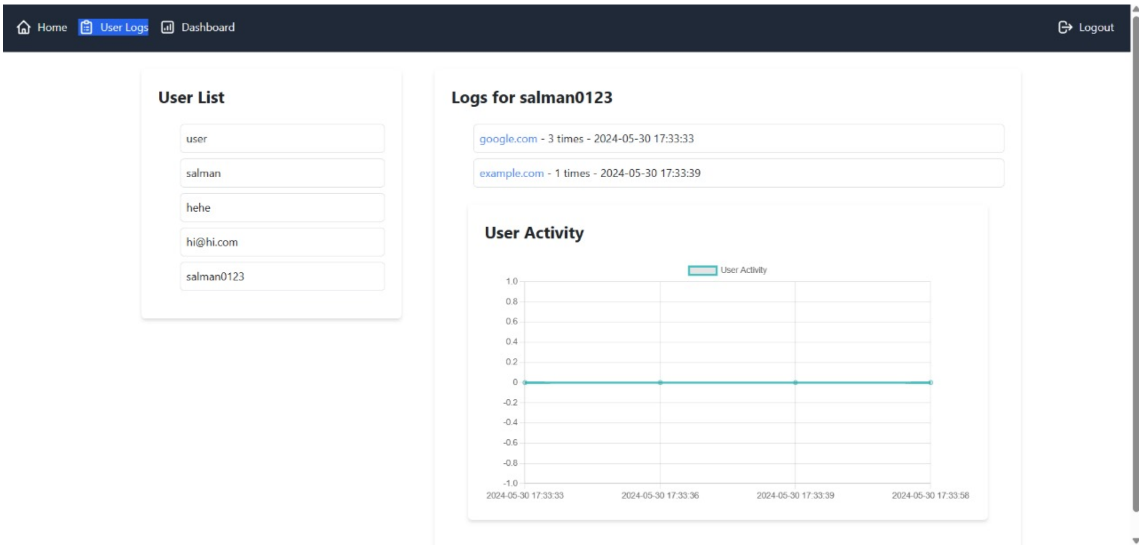
Home page consist of following buttons:

- **Home:** showing the current page
- **Log Out:** allowing the user to logout of the application
- **User Logs:** keep and display track of web activity logs

4.3.3 User-log

This page has three parts:

- **User List:** This list show the devices which are registered as users using web monitor extension
- **Logs for selected user:** When a user is selected this button shows what are the specific websites the user has visited and at what time.
- **ine chart:** The graph is raised when user access a restricted websites.



User List

- user
- salman
- hehe
- hi@hi.com
- salman0123

Logs for salman0123

- google.com - 3 times - 2024-05-30 17:33:33
- example.com - 1 times - 2024-05-30 17:33:39

User Activity

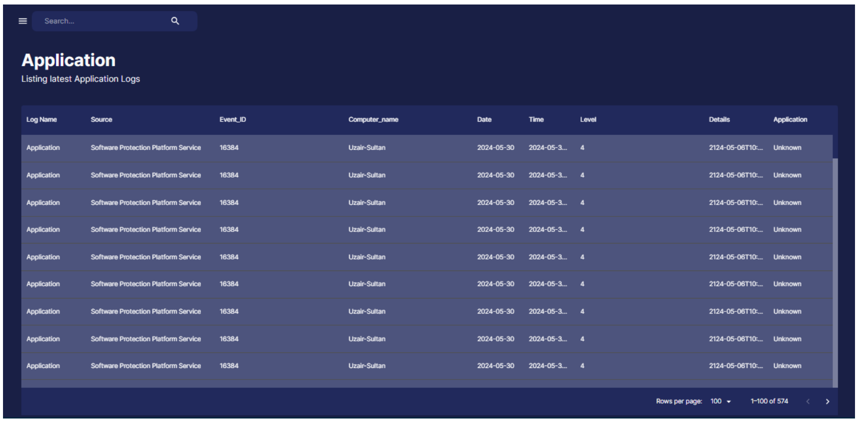
4.3.4 Logs View

A separate dashboard has been created to display latest logs from these channels:Application,Security,Sys

4.3.4.1 Application Logs

This image is pulling live logs from click-house tables and displaying them in the form of table for user to view.For the table Material UI data grids are used, as click-house index every 8192 row to be precise and this method of indexing is called as sparse indexing. MUI Data Grid needs row index for every entry, for this case row-id is created by concatenation of

- Event Id
- Computer Name
- Date
- Time



Application

Listing latest Application Logs

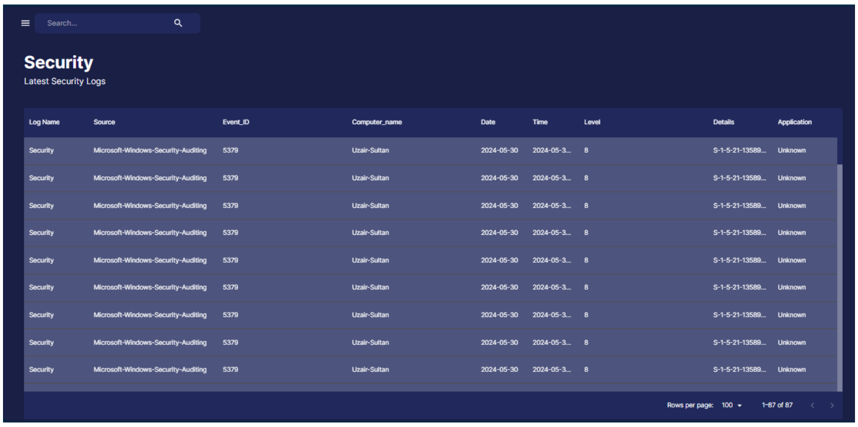
Log Name	Source	Event_ID	Computer_name	Date	Time	Level	Details	Application
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown
Application	Software Protection Platform Service	16384	Uzair-Sultan	2024-05-30	2024-05-3...	4	2124-05-06T10:...	Unknown

Rows per page: 100 1–100 of 574

Figure 4.5: Image displaying latest application logs

4.3.4.2 Security Logs

Pulling logs from security table



Security

Latest Security Logs

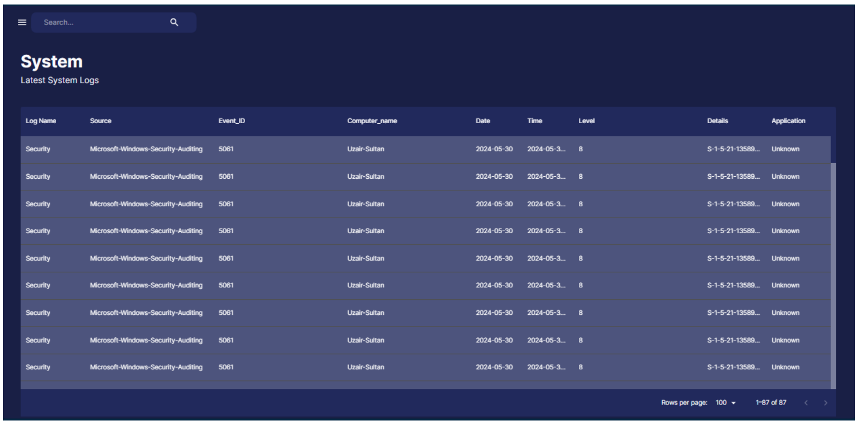
Log Name	Source	Event_ID	Computer_name	Date	Time	Level	Details	Application
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5379	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown

Rows per page: 100 1–87 of 87

Figure 4.6: Image displaying latest security logs

4.3.4.3 System Logs

Pulling logs from system table



System

Latest System Logs

Log Name	Source	Event_ID	Computer_name	Date	Time	Level	Details	Application
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown
Security	Microsoft-Windows-Security-Auditing	5061	Uzair-Sultan	2024-05-30	2024-05-3...	8	S-1-5-21-13589...	Unknown

Rows per page: 100 1–87 of 87

Figure 4.7: Image displaying latest system logs

4.4 Detected Threats

MITRE's Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) architecture provides a curated knowledge base and model for cyber adversary behavior. It represents the stages of an adversary's attack life-cycle and the platforms they target. MITRE ATTCK, created by Malware Archaeology, can identify 83% of Windows-specific attack tactics based on event IDs for organizations employing the Windows operating system.

These are some tactic areas where techniques are mapped to:

- Collection
- Command and Control
- Credential Access
- Defense Evasion
- Initial Access
- Execution
- Persistence
- Privilege Escalation

- Discovery
- Lateral Movement
- Ex-filtration

4.4.1 Collection

Collection refers to methods used in acquiring data from a target network, including files and other information that may be of importance before transmitting them. This embeds locations or areas on a system or a network where a threat actor can find information to steal.

No.	Tactic
1	Audio Capture
2	Automated Collection
3	Clipboard Data
4	Data from Information Repositories
5	Data from Local System
6	Data from Network Shared Drive
7	Data from Removable Media
8	Data Staged
9	Email Collection
10	Man in the Browser
11	Screen Capture

Table 4.1: Collection Tactics

4.4.2 Command and Control

Command and Control contains information about the commuincation done by threat actors to the systems under their control.

No.	Tactic
1	Commonly Used Port
2	Communication Through Removable Media
3	Connection Proxy
4	Custom Command and Control Protocol
5	Custom Cryptographic Protocol
6	Data Encoding
7	Data Obfuscation
8	Domain Fronting
9	Fallback Channels
10	Multiband Communication
11	Multi-hop Proxy
12	Multilayer Encryption

Table 4.2: Command and Control Tactics

4.4.3 Credential Access

Credential access is necessary for accessing applications, but its security is more important than ever.When any outside organization has gotten credential access, then the risk of every other cyber risk goes up.

No.	Tactic
1	Account Manipulation
2	Brute Force
3	Credential Dumping
4	Credentials in Files
5	Credentials in Registry
6	Exploitation for Credential Access
7	Forced Authentication
8	Kerberoasting

Table 4.3: Credential Access Methods

4.4.4 Defence Evasion

Defense Evasion are techniques that actors use to avoid detection during their compromise. Encrypting data and scripts ,disabling security software are common defense evasion techniques.

No.	Tactic
1	File Deletion
2	File System Logical Offsets
3	Indicator Blocking
4	Indicator Removal from Tools
5	Indicator Removal on Host
6	Indirect Command Execution
7	Install Root Certificate
8	Masquerading
9	Modify Registry
10	Network Share Connection Removal

Table 4.4: Defence Evasion Tactics

4.4.5 Discovery

Threat actors have to learn how the internal and external components of system and networks work, the techniques they use to identify these fall under Discovery tactic area.

No.	Tactic
1	Application Window Discovery
2	Browser Bookmark Discovery
3	File and Directory Discovery
4	Network Service Scanning
5	Network Share Discovery
6	Password Policy Discovery
7	Peripheral Device Discovery
8	Permission Groups Discovery
9	Process Discovery
10	Query Registry

Table 4.5: Discovery Tactics

4.4.6 Execution

Execution involves strategies for running code on a local or remote system. Malicious code execution techniques are often combined with other strategies to achieve larger goals, such as network exploration or data theft. For instance, an attacker might use a remote access tool to execute a PowerShell script for performing Remote System Discovery.

No.	Tactic
1	Command-Line Interface
2	Dynamic Data Exchange
3	Execution through API
4	Execution through Module Load
5	Exploitation for Client Execution
6	Graphical User Interface
7	PowerShell

Table 4.6: Execution Tactics

4.4.7 Exfiltration

Stealing data from end points and storing it somewhere else is known as Exfiltration. Collected data is frequently package to avoid detection.

No.	Tactic
1	Data Compressed
2	Data Encrypted
3	Exfiltration Over Alternative Protocol
4	Exfiltration Over Command and Control Channel
5	Exfiltration Over Other Network Medium

Table 4.7: Exfiltration Tactics

4.4.8 Initial Access

Initial Access refers to techniques that leverage various entry vectors to gain an initial foothold within a network.

No.	Tactic
1	Data Compressed
2	Data Encrypted
3	Exfiltration Over Alternative Protocol
4	Exfiltration Over Command and Control Channel
5	Exfiltration Over Other Network Medium

Table 4.8: Initial Access Tactics

4.4.9 Lateral Movement

Lateral movement is the process that the attackers employ to achieve access to as well as oversight of other systems on the network. In order to accomplish their main objective regularly, adversaries often need to navigate through the network to find the object of their interest and then, get to it.

No.	Tactic
1	Pass the Ticket
2	Remote Desktop Protocol
3	Remote Services
4	Shared Webroot
5	Taint Shared Content
6	Windows Admin Shares
7	Replication Through Removable Media
8	Logon Scripts

Table 4.9: Lateral Movement Tactics

4.4.10 Persistence

Persistence refers to techniques used by adversaries to maintain access to systems despite restarts, changed credentials, and other interruptions that could terminate their access.

No.	Tactic
1	Modify Existing Service
2	Netsh Helper DLL
3	Office Application Startup
4	Registry Run Keys/Start Folder
5	Screensaver
6	Security Support Provider
7	Shortcut Modification
8	System Firmware
9	Time Providers
10	Windows Management Instrumentation
11	Event Subscription
12	Winlogon Helper DLL

Table 4.10: Persistence Tactics

4.4.11 Privelege Escalation

Privilege escalation refers to techniques used by adversaries to gain higher-level permissions on a system or network. Adversaries can frequently enter and explore a network with unprivileged access but require elevated permissions to carry out their objectives.

No.	**Tactic**
1	File System Permissions Weakness
2	Software Misconfiguration
3	Service Misconfiguration
4	Insufficient File/Directory Permissions
5	Token Impersonation
6	Hardware Vulnerabilities
7	Poorly Privileged Services
8	Valid Accounts with Excessive Privileges
9	Exploiting Vulnerable Applications
10	Brute Force Attacks
11	Supply Chain Attacks

Table 4.11: Privelege Escalation

Chapter 5

Validation

Here is the step by step implementation and of the system.

5.0.1 Installing Agent

The collector files are bundles into one executable which can be run on any window device.

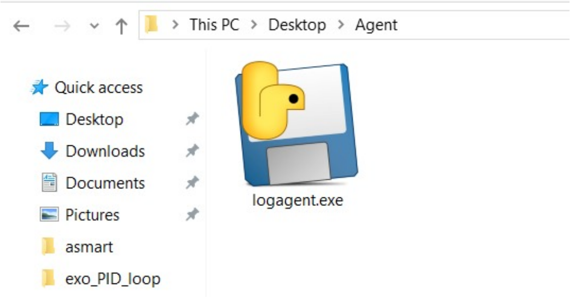
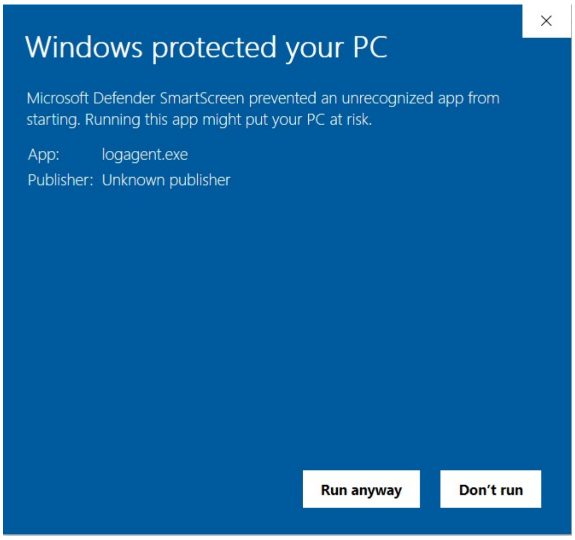


Figure 5.1: Locating Agent Executable

After double clicking windows defender gives a safety prompt.Pressing Run anyway button:



Windows protected your PC

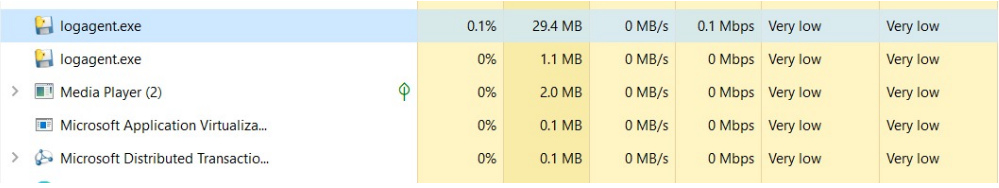
Microsoft Defender SmartScreen prevented an unrecognized app from starting. Running this app might put your PC at risk.

App: logagent.exe

Publisher: Unknown publisher

Figure 5.2: Defender Safety Message

As the button is pressed application starts running in the background collecting logs and sending them to click-house database.

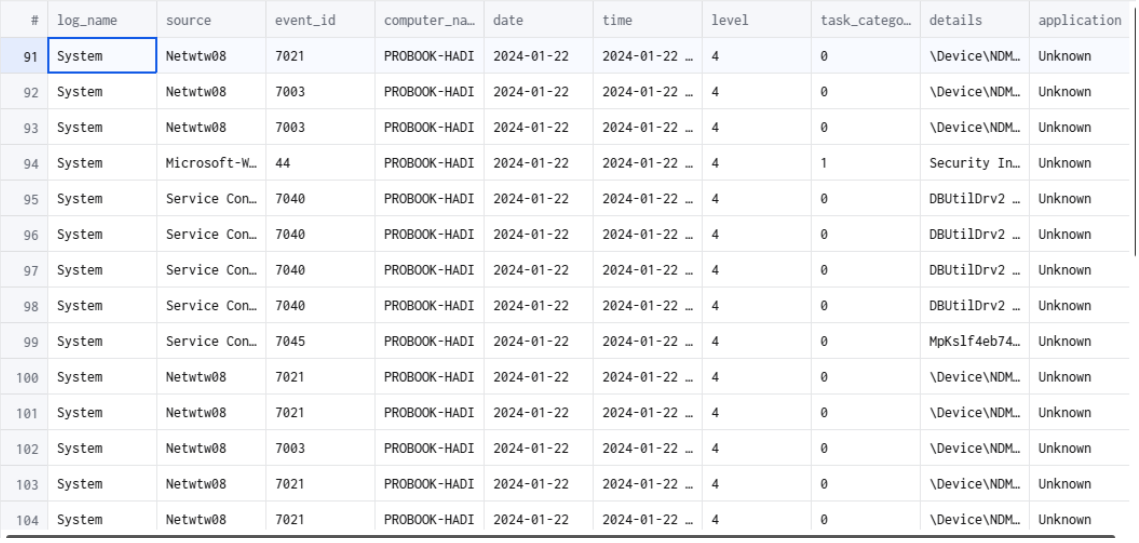


logagent.exe	0.1%	29.4 MB	0 MB/s	0.1 Mbps	Very low	Very low
logagent.exe	0%	1.1 MB	0 MB/s	0 Mbps	Very low	Very low
Media Player (2)	0%	2.0 MB	0 MB/s	0 Mbps	Very low	Very low
Microsoft Application Virtualiza...	0%	0.1 MB	0 MB/s	0 Mbps	Very low	Very low
Microsoft Distributed Transactio...	0%	0.1 MB	0 MB/s	0 Mbps	Very low	Very low

Figure 5.3: Task Manager showing log agent is running in the background

5.0.2 Log Check

After agent start running we need to check whether it is sending logs or not.Before checking in the front end it is viewed in the click-house database . Here is the collected log from device where log agent was just installed.



#	log_name	source	event_id	computer_na…	date	time	level	task_catego…	details	application
91	System	Netwtw08	7021	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	\Device\NDM…	Unknown
92	System	Netwtw08	7003	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	\Device\NDM…	Unknown
93	System	Netwtw08	7003	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	\Device\NDM…	Unknown
94	System	Microsoft-W…	44	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	1	Security In…	Unknown
95	System	Service Con…	7040	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	DBUtilDrv2 …	Unknown
96	System	Service Con…	7040	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	DBUtilDrv2 …	Unknown
97	System	Service Con…	7040	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	DBUtilDrv2 …	Unknown
98	System	Service Con…	7040	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	DBUtilDrv2 …	Unknown
99	System	Service Con…	7045	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	MpKslf4eb74…	Unknown
100	System	Netwtw08	7021	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	\Device\NDM…	Unknown
101	System	Netwtw08	7021	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	\Device\NDM…	Unknown
102	System	Netwtw08	7003	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	\Device\NDM…	Unknown
103	System	Netwtw08	7021	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	\Device\NDM…	Unknown
104	System	Netwtw08	7021	PROBOOK-HADI	2024-01-22	2024-01-22 …	4	0	\Device\NDM…	Unknown

Figure 5.4: Database image displaying newly inserted logs

Chapter 6

Conclusion & Future Work

6.1 Conclusion

We were able to collect Windows events and web activity logs successfully using custom python agents using pywin32 library.The logs were securely transferred from endpoints to database using Hyper Text Transfer Protocol. The logs are then accessed via user specific queries for investigation and analysis for customer's use case. The results are stored in tables which can be accessed later. Chrome extension helps in pulling users' web activity and notify the admin for any organization rule negation.

6.2 Future Work

We were able to collect logs and events from multiple devices and store them in a central storage hub.Along with it, we were able to detect specific threats live and show in a table.Here are some extra things which can be added to the project for making it an industry suitable.

- Re-configurable and update-able agent
- Creating specific tables for each new installed agents

- Analytics and insights graphs
- Using cloud functions check for new entries in database and classify them using rule based detection and ML models, and store in respective threat databases
- Adding new log sources
- Adding notification functionality for admins
- allowing forensic analysis

References

[1] C. Jones, “50 phishing stats you should know in 2024,” Dec. 2023.

[2] J. Witts, “The top 5 biggest cybersecurity threats that small businesses face and how to stop them,” Feb. 2024.

[3] E. Georgescu, “Software patching statistics: Common practices and vulnerabilities,” 2022.

[4] “Ibm qradar security intelligence platform 7.4,” May 2024.