IBM Project Report On IT System Log Analyzer

Developed By: -

Dhruv Patel(20162171015)

Jay Sapra(20162171009)

Tanishk Patel(20162121020)

Guided By: -

Prof. Neha Sisodiya (Internal)

Mr. Anoj Dixit (External)

Submitted to

Department of Computer Science & Engineering

Institute of Computer Technology

Year: 2024

Ganpat University
॥ विद्यया समाजोत्कर्षः ॥

Institute of Computer Technology

National Assessment and Accreditation Council
NAAC
A
Grade
Accredited University

CERTIFICATE

This is to certify that the **IBM** Project work entitled **"IT System Log Analyzer"** by Dhruv Patel(Enrolment No.20162171015), Jay Sapra(Enrolment No.20162171009) and Tanishk Patel (Enrolment No.20162121020) of Ganpat University, towards the partial fulfillment of requirements of the degree of Bachelor of Technology – Computer Science and Engineering. The results/findings contained in this Project have not been submitted in part or full to any other University / Institute for award of any other Degree/Diploma.

Name & Signature of Internal Guide

Name & Signature of Head

Place: ICT - GUNI

Date:

ACKNOWLEDGEMENT

IBM Internship project is a golden opportunity for learning and self-development. I consider myself very lucky and honored to have so many wonderful people lead me through in completion of this project. First and foremost, I would like to thank Dr. Hemal Shah, Principal, ICT, and Prof. Dharmesh Darji, Head, ICT who gave us an opportunity to undertake this project. My grateful thanks to Prof. Neha Sisodiya & Mr. Anoj Dixit (Internal & External Guides) for their guidance in project work Predicting Application Rating of Google Play Store, who despite being extraordinarily busy with academics, took time out to hear, guide and keep us on the correct path. We do not know where would have been without his/her help. CSE department monitored our progress and arranged all facilities to make life easier. We choose this moment to acknowledge their contribution gratefully.

Dhruv Patel (Enrolment No. 20162171015)
Jay Sapra (Enrolment No. 20162171009)
Tanishk Patel (Enrolment No. 20162121020)

ABSTRACT

CRPF units/offices and personnel are deployed in different location of CRPF. There is no centralized system to analyze the log of IT system by the experts to access threats and breaches. Proposed Solution: Centralized system should be developed for analyzing the systems deployed at the different location of the country Expert per problems statement.

INDEX

CHAPTER 1: INTRODUCTION

CHAPTER 1 INTRODUCTION

In the contemporary landscape of technological advancements, the Central Reserve Police Force (CRPF) plays a pivotal role in ensuring national security with its widespread deployment across diverse locations in the country. The seamless operation of information technology (IT) systems within CRPF units is imperative for efficient communication, data management, and overall mission success. However, a critical gap exists in the current infrastructure - the absence of a centralized system for comprehensive analysis of IT system logs.

As our nation faces evolving cybersecurity threats and potential breaches, it becomes essential to address this vulnerability within the CRPF. The lack of a unified approach to scrutinize IT system logs poses a significant challenge for identifying and mitigating potential threats in a timely manner. CRPF units and offices, scattered across different regions, operate independently, lacking a consolidated mechanism for experts to analyze system logs systematically.

This project, titled "IT System Log Analyzer," emerges as a strategic response to the identified issues within the existing CRPF IT infrastructure. The primary goal is to develop a centralized system that empowers experts to analyze IT system logs efficiently, irrespective of the geographical dispersion of CRPF units. This proposed solution aligns with the need for a systematic and expert-driven approach to assess and address potential threats and breaches within the CRPF IT network.

The forthcoming sections of this report will delve into the specific challenges faced by CRPF units regarding IT system log analysis, outlining the proposed solution in detail. By establishing a centralized system, this project aims to enhance the overall cybersecurity posture of the CRPF, ensuring a proactive and responsive approach to emerging threats.

In the subsequent chapters, we will explore the current state of IT system log analysis within CRPF, elucidate the intricacies of the proposed centralized system, and present a comprehensive roadmap for its implementation. Through this project, we aspire to contribute to the robustness of CRPF's IT infrastructure, fostering a secure environment for operations and data management.

CHAPTER 2: PROJECT SCOPE

CHAPTER 2 PROJECT SCOPE

The scope of the "IT System Log Analyzer" project is defined by the critical need to establish a centralized system for analyzing the IT system logs within the Central Reserve Police Force (CRPF). Currently, the dispersed deployment of CRPF units and offices across various locations has resulted in a decentralized approach to IT system log analysis. This fragmentation presents a formidable challenge, as there is no centralized mechanism for experts to systematically assess and identify potential threats and breaches. The project aims to bridge this gap by designing, developing, and implementing a comprehensive centralized system that will facilitate the efficient analysis of IT system logs, ensuring a unified and expert-driven approach to cybersecurity.

The primary focus of the project is to create a centralized platform capable of collecting, monitoring, and analyzing IT system logs from different CRPF locations across the country. This scope encompasses the integration of advanced analytical tools and methodologies to enable experts to identify patterns, anomalies, and potential security breaches promptly. The system will not only serve as a repository for log data but will also provide a user-friendly interface for experts to conduct in-depth analyses, generate reports, and take proactive measures to safeguard the integrity and security of CRPF's IT infrastructure. By addressing the current lack of centralized log analysis, this project seeks to enhance the overall cybersecurity resilience of CRPF, promoting a more secure and cohesive IT environment for its personnel and operations

CHAPTER 3: SOFTWARE AND HARDWARE REQUIREMENTS

CHAPTER 3 SOFTWARE AND HARDWARE REQUIREMENTS

Minimum Hardware Requirements

Processor	2.0 GHz
RAM	4GB
HDD	40GB

Table 3.1 Minimum Hardware Requirements

Minimum Software Requirements

Operating System	Any operating system which can support an internet browser.
Programming language	-
Other tools & tech	Splunk, Ngrok

Table 3.2 Minimum Software Requirements

CHAPTER 4: IMPLEMENTATION DETAILS

CHAPTER 4 IMPLEMENTATION DETAILS

4.1 Install Windows 10 for System Logs

We have to install windows 10 machine in VMware for the system logs.

4.2 Install Splunk in Main System

We have to install a Splunk in our main system to get log of the other machine.

Splunk is an excellent choice for IT systems log analysis. It's a powerful platform designed to make sense of machine-generated data, including logs, by providing real-time visibility and actionable insights. Here's how Splunk can be beneficial for IT systems log analysis:

- Centralized Log Management: Splunk allows you to aggregate logs from various sources, including servers, applications, network devices, and more, into a centralized repository. This makes it easier to manage and analyze logs from multiple sources.

- Real-Time Monitoring: Splunk enables real-time monitoring of log data, allowing you to detect and respond to issues as they occur. You can set up alerts and notifications based on predefined criteria to proactively address potential problems.

- Search and Analysis: Splunk's powerful search capabilities allow you to quickly search and analyze large volumes of log data. You can use complex search queries, filters, and statistical functions to gain insights into system behavior, performance, security events, and more.

- Visualization and Reporting: Splunk provides various visualization tools to help you understand log data more easily. You can create dashboards and reports to visualize trends, anomalies, and key performance indicators (KPIs), making it easier to communicate insights to stakeholders.

- Troubleshooting and Root Cause Analysis: Splunk can help streamline troubleshooting and root cause analysis by providing visibility into system events and dependencies. You can trace the flow of events across different components and identify the underlying causes of issues more efficiently.

- Compliance and Security: Splunk can assist with compliance adherence and security monitoring by providing visibility into user activities, access logs, and security events. You can use Splunk to detect security threats, analyze security incidents, and ensure compliance with regulatory requirements.

- Scalability and Flexibility: Splunk is highly scalable and flexible, capable of handling large volumes of log data from diverse sources. Whether you're managing a small IT environment or a large enterprise infrastructure, Splunk can scale to meet your needs.

4.3 Setup Splunk Forwarder in VMware Machine

We have to setup a splunk forwarder in a VMware machine to forward logs of the VMware

machine to main system. Whatever activity we will perform in VMware machine the data would be forward to my main system splunk.

Splunk Universal Forwarder is a lightweight, dedicated component of the Splunk platform designed specifically for sending data to Splunk Enterprise or Splunk Cloud. It plays a crucial role in IT systems log analysis by collecting log data from various sources and forwarding it to a centralized Splunk deployment for indexing, search, and analysis. Here's how Splunk Universal Forwarder contributes to IT systems log analysis:

- Data Collection: Universal Forwarder collects log data from a wide range of sources, including servers, applications, network devices, sensors, and custom data sources. It can monitor log files, listen on network ports, receive syslog messages, and ingest data from other inputs.

- Agent-Based Approach: Universal Forwarder operates as an agent installed on each data source or system where log data needs to be collected. This lightweight agent consumes minimal system resources and can be easily deployed across distributed IT environments.

- Secure Data Transmission: Universal Forwarder securely transmits log data to the Splunk indexing tier using encrypted communication protocols, such as HTTPS or TCP/TLS. This ensures the confidentiality and integrity of log data during transmission over the network.

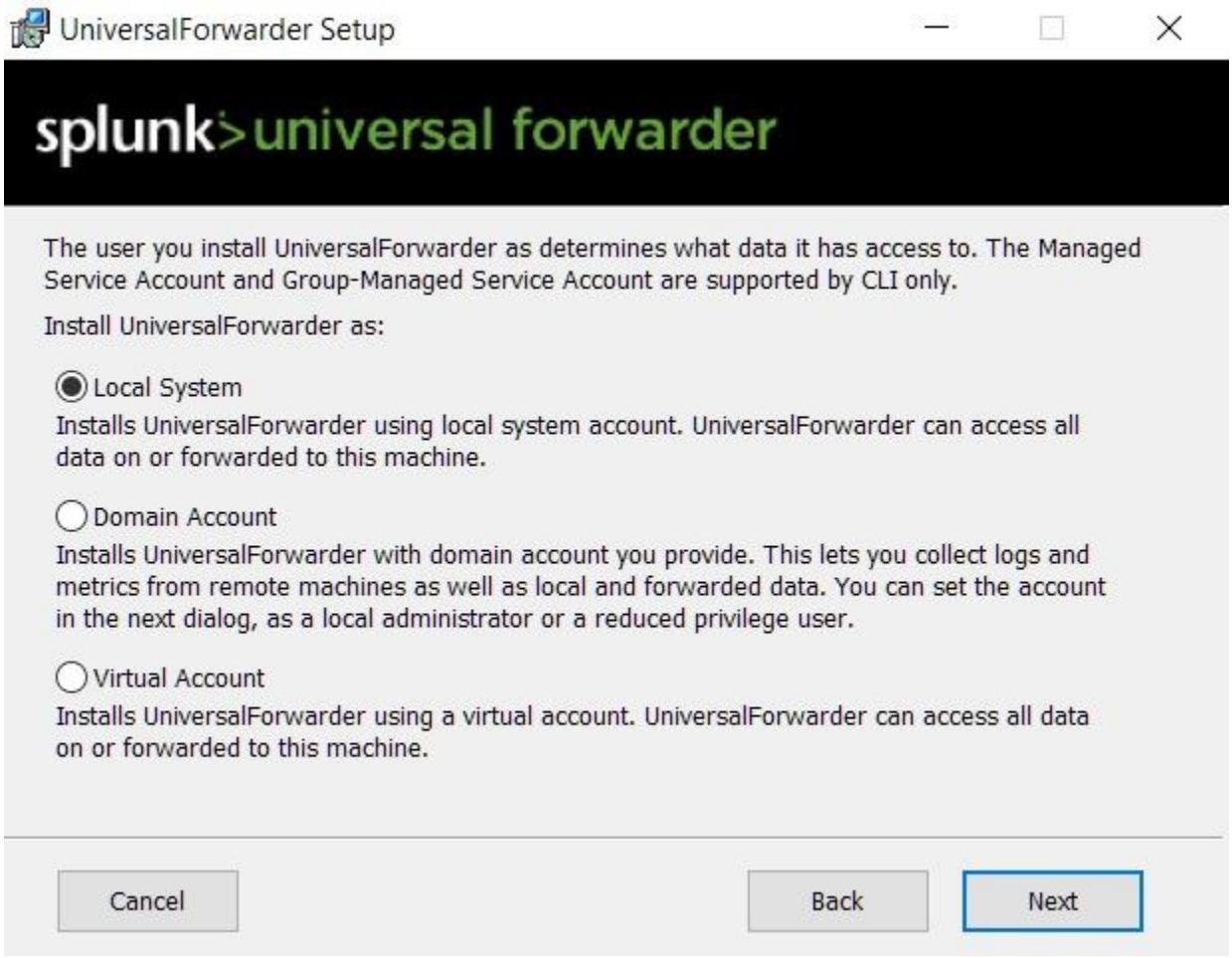
- Indexing Efficiency: By offloading the task of log collection to Universal Forwarder, Splunk indexing tier can focus on efficiently indexing and storing log data without putting undue strain on system resources. This helps maintain indexing performance and scalability, especially in large-scale deployments.

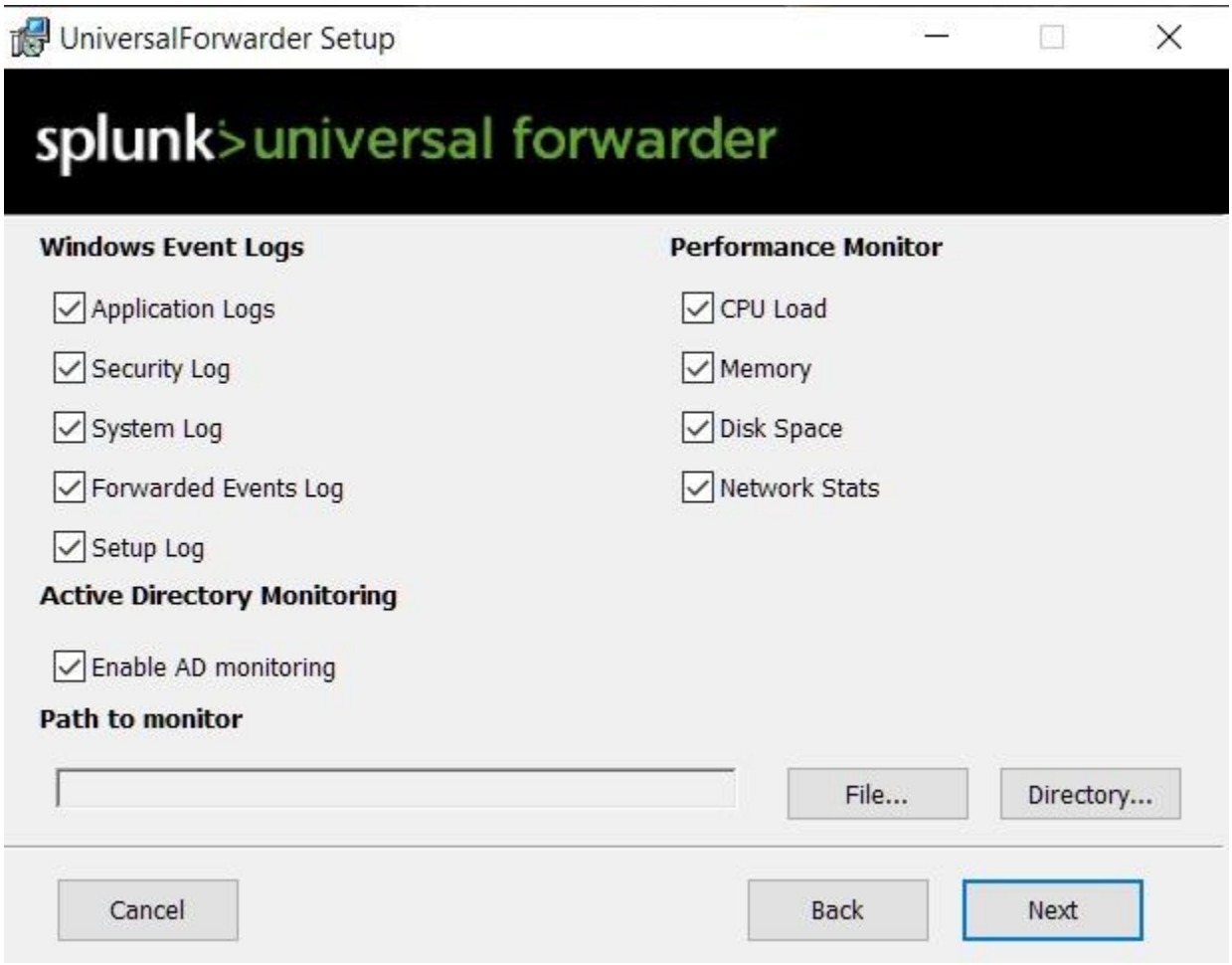
- Real-Time Data Streaming: Universal Forwarder can stream log data in real-time to Splunk, enabling near real-time visibility into system events and activities. This facilitates real-time monitoring, alerting, and response to critical events as they occur.

- Deployment Management: Splunk provides centralized management capabilities for deploying, configuring, and monitoring Universal Forwarder instances across distributed environments. Administrators can centrally manage configurations, updates, and health monitoring of Universal Forwarder deployments.

- Flexibility and Customization: Universal Forwarder supports a wide range of customization options, allowing administrators to tailor data collection configurations based on specific requirements. This includes filtering, data transformation, field extractions, and other advanced processing tasks.

- Compatibility and Integration: Universal Forwarder seamlessly integrates with other Splunk components, such as Splunk Enterprise, Splunk Cloud, and Splunk Data Stream Processor, to provide end-to-end log analysis solutions. It also supports integration with third-party solutions and ecosystem tools through APIs and plugins.

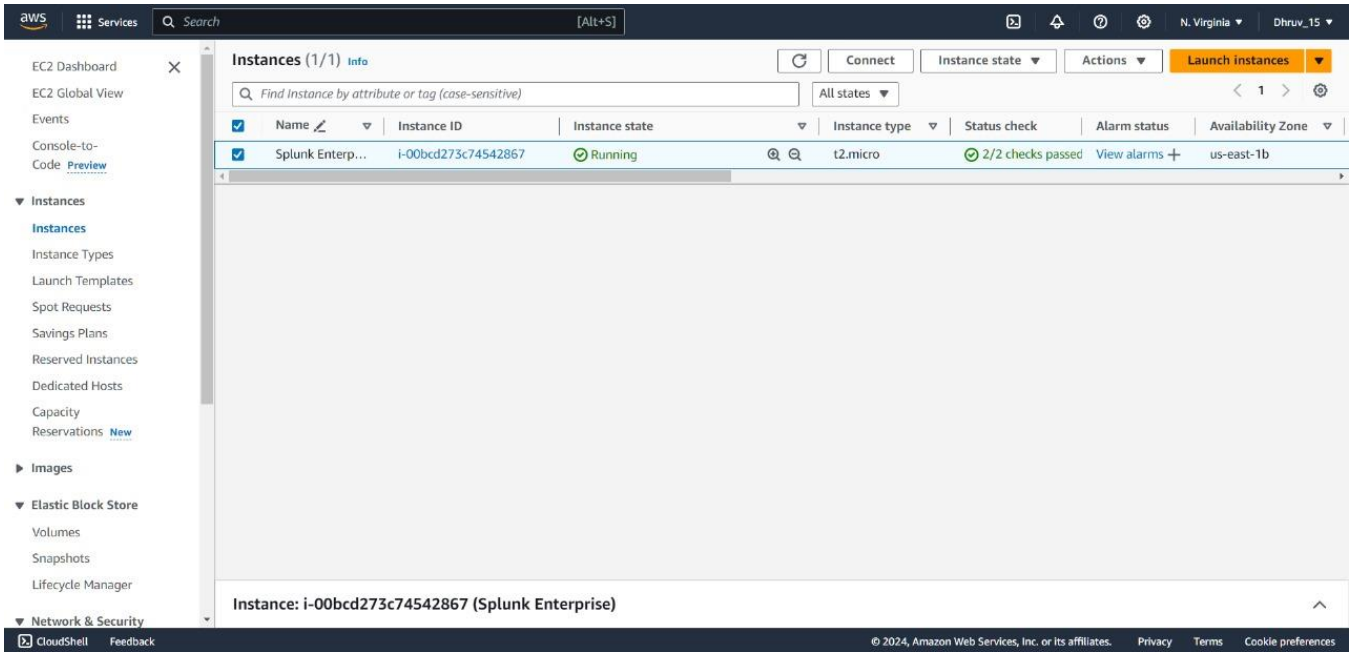

UniversalForwarder Setup
splunk>universal forwarder
The user you install UniversalForwarder as determines what data it has access to. The Managed Service Account and Group-Managed Service Account are supported by CLI only.
Install UniversalForwarder as:
Local System
Installs UniversalForwarder using local system account. UniversalForwarder can access all data on or forwarded to this machine.
Domain Account
Installs UniversalForwarder with domain account you provide. This lets you collect logs and metrics from remote machines as well as local and forwarded data. You can set the account in the next dialog, as a local administrator or a reduced privilege user.
Virtual Account
Installs UniversalForwarder using a virtual account. UniversalForwarder can access all data on or forwarded to this machine.
Cancel
Back
Next

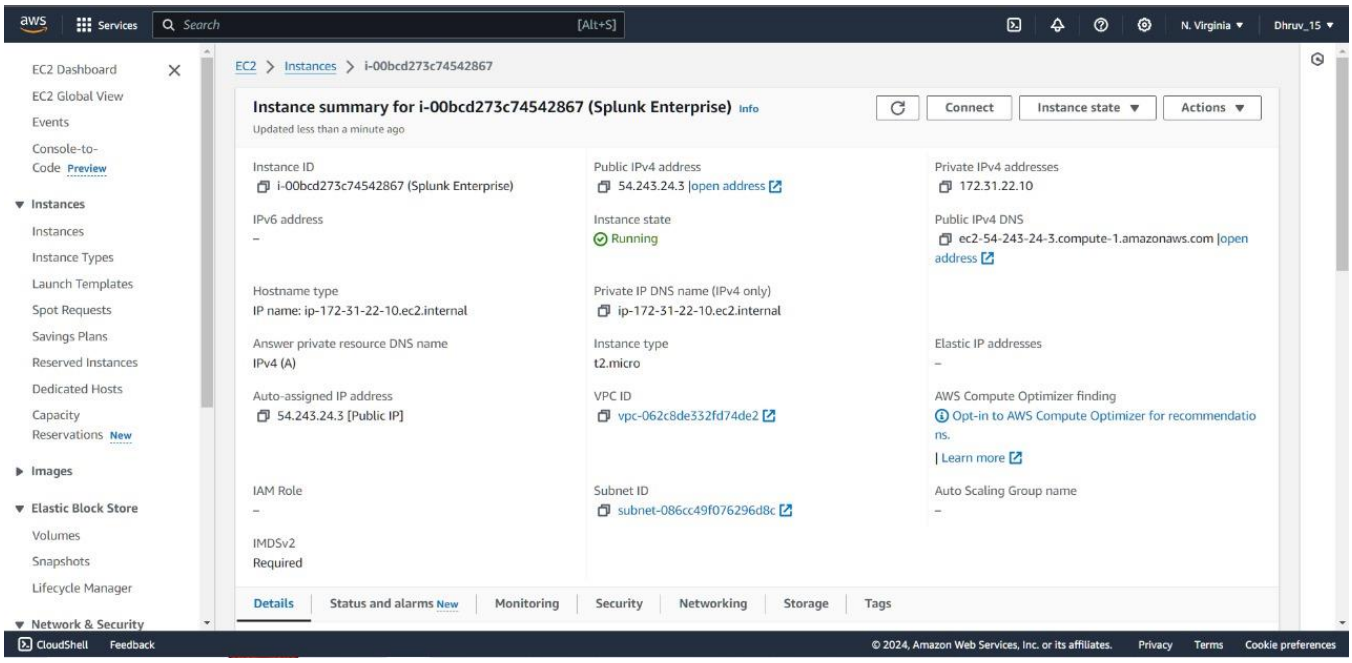

UniversalForwarder Setup
splunk>universal forwarder
Windows Event Logs
Application Logs
Security Log
System Log
Forwarded Events Log
Setup Log
Performance Monitor
CPU Load
Memory
Disk Space
Network Stats
Active Directory Monitoring
Enable AD monitoring
Path to monitor
File...
Directory...
Cancel
Back
Next

4.4 Create EC2 Instance and Integrate Splunk to EC2 Instance

4.4.1 Create EC2 Instance

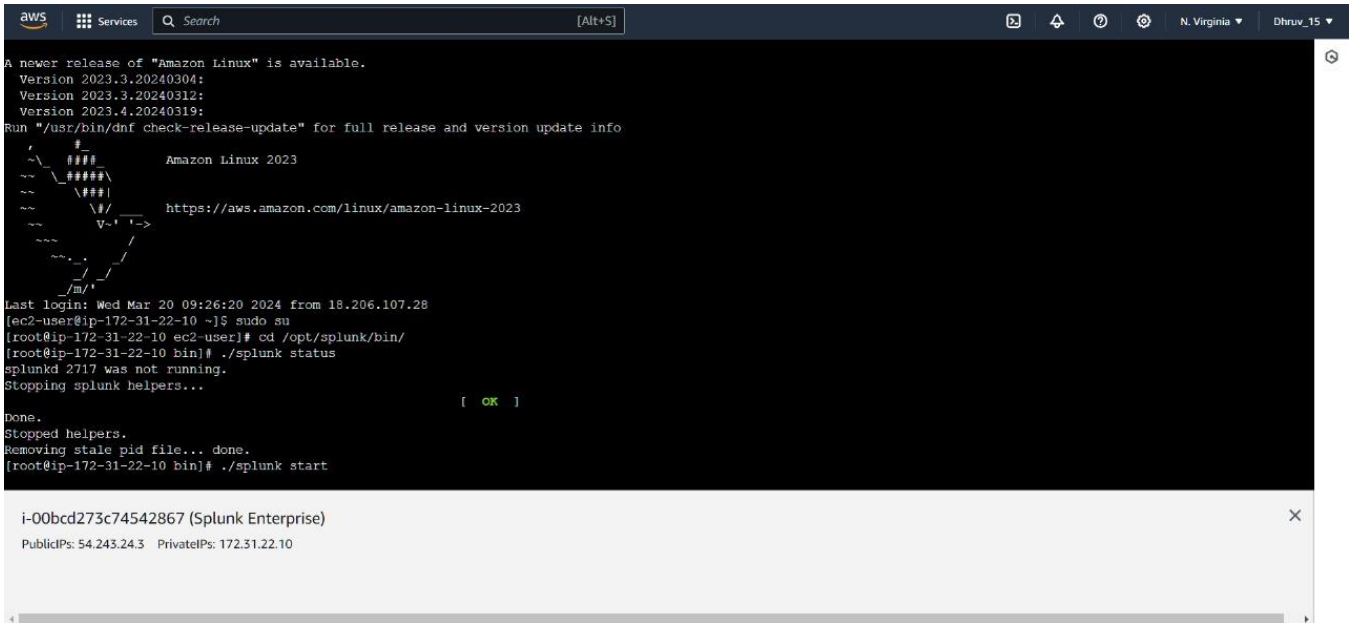
Create EC2 instance to make centralized system. In centralized system we get all the logs from all the connected system at EC2 instance.





4.4.2 Integrate Splunk to EC2 instance

We integrate splunk to EC2 instance.



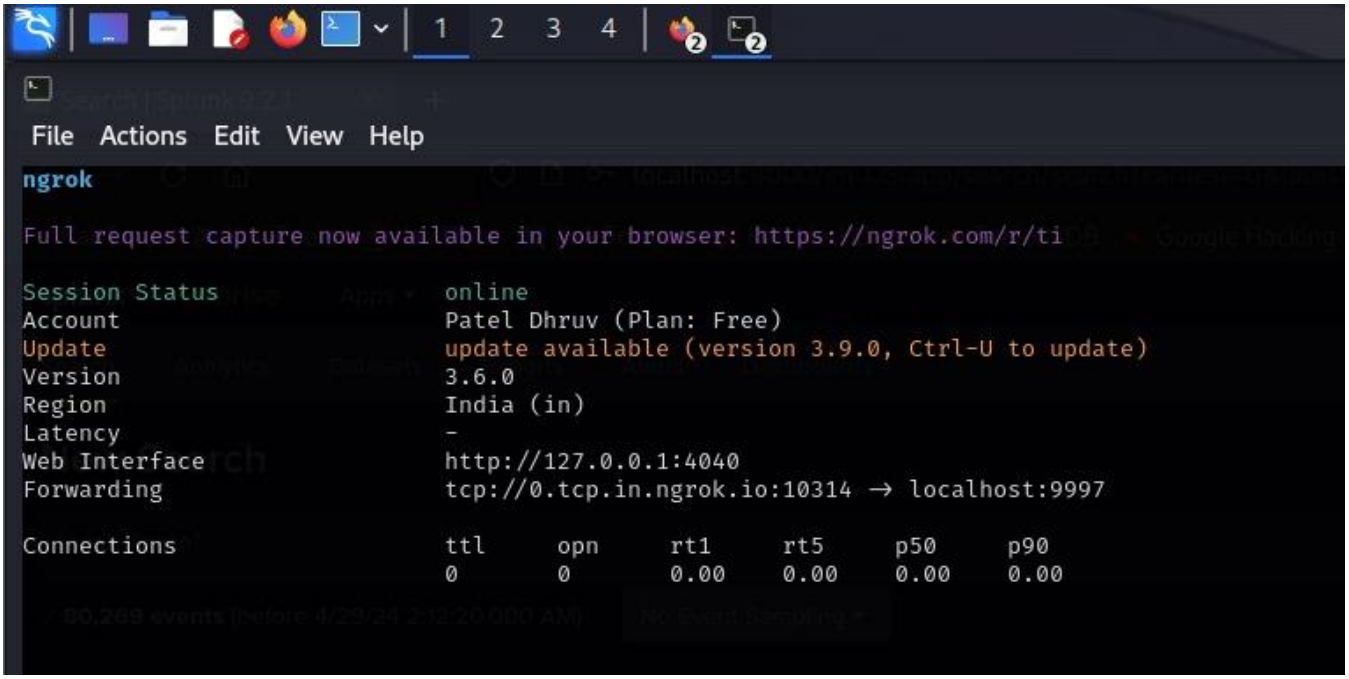
4.4 Using Ngrok instead of AWS EC2

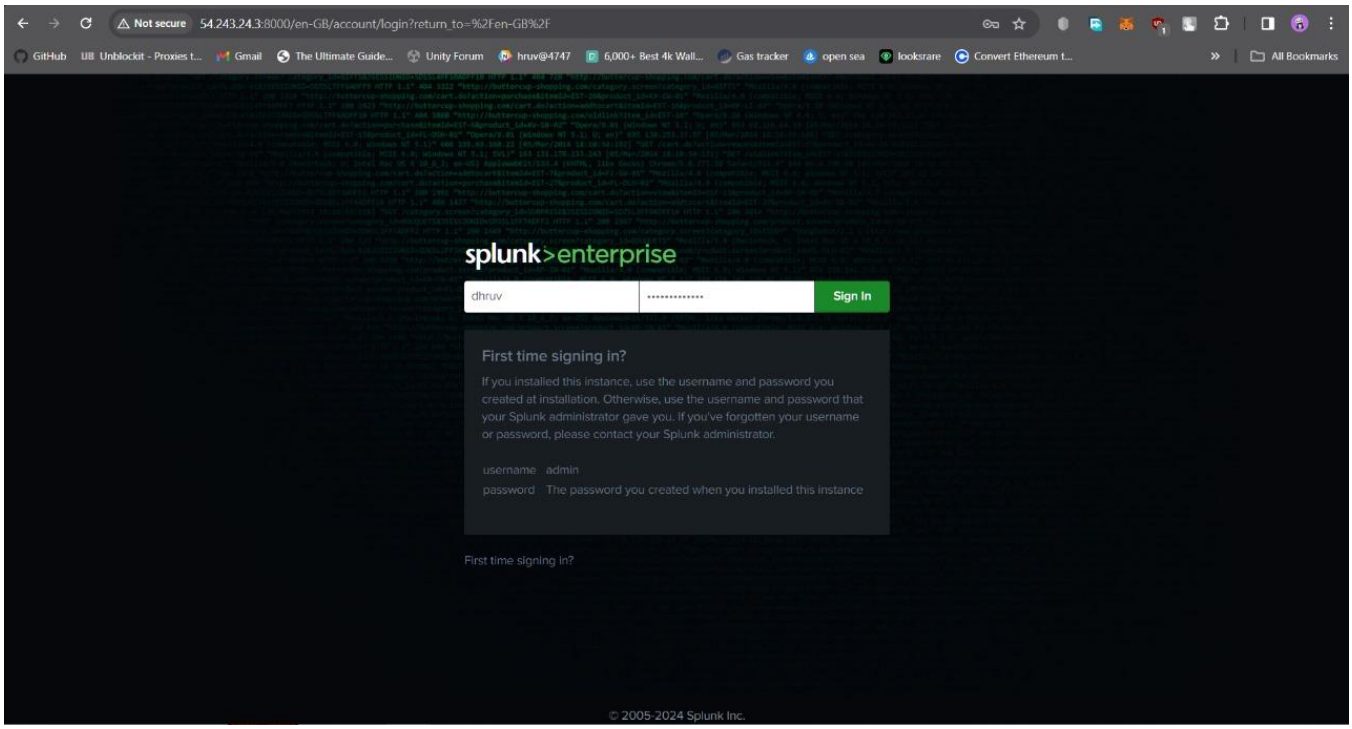
We got a storage problem in the AWS EC2 instance. That's why we use Ngrok instead of AWS EC2.

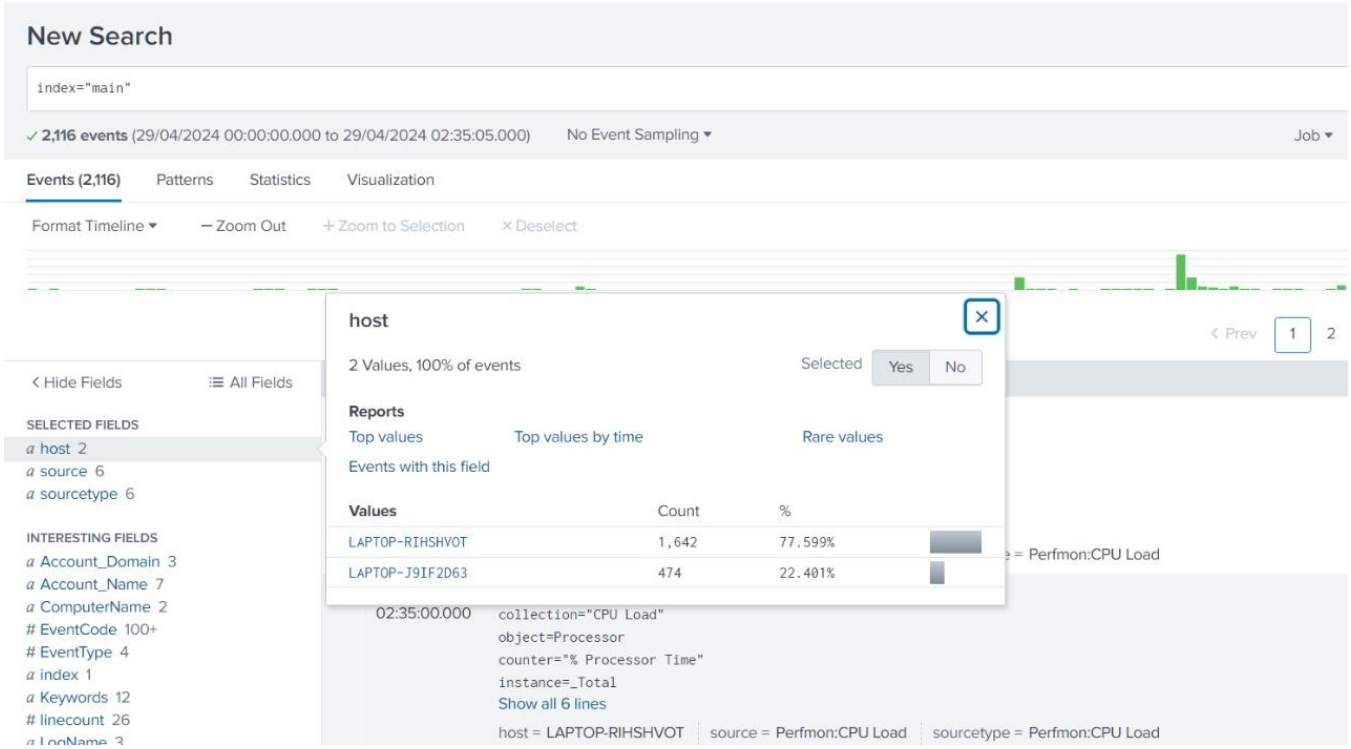
- Ngrok will be used to address the storage limitations posed by AWS EC2.
- Ngrok provides secure tunneling capabilities, allowing access to local services behind firewalls and NATs.
- By using Ngrok, we can securely expose the Splunk instance to the internet without the need for extensive storage infrastructure.

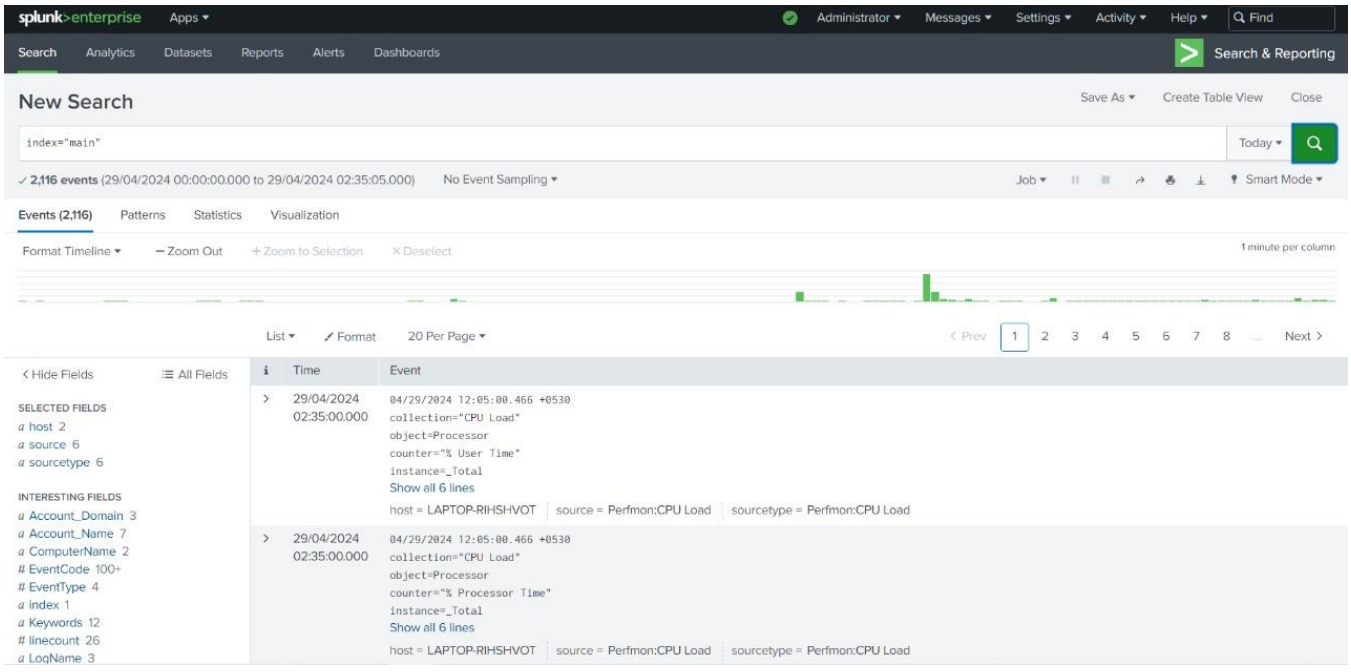
```
┌──(root㉿kali)-[/opt]
└─# ls
microsoft  splunk  Storm-Breaker

┌──(root㉿kali)-[/opt]
└─# cd /opt/splunk/bin

┌──(root㉿kali)-[/opt/splunk/bin]
└─# ls
2to3-3.7                    compsup                genWebCert.py  mongod                      pcregextest   pripngtopam  rapidDiag                       shc_upgrade_template.py  splunk-tlsd
bloom                       copyright.txt          genWebCert.sh  mongod-3.6                  pid_check.sh  priweavepng  recover-metadata                slgtool                  supervisor-simulator
bottle.py                   dbmanipulator.py       idle3          mongod-4.0                  pip3          pydoc3       rest_handler.py                 slim                     tarit.py
btool                       easy_install-3.7       idle3.7        mongodump                   pip3.7        pydoc3.7     runScript.py                    spl2-orchestrator        tocsv.py
btprobe                     exporttool             importtool     mongorestore                prichunkpng   python       S3benchmark                     spl-lang-server-sockets  tsidxprobe
bzip2                       fill_summary_index.py  installit.py   noah_self_storage_archiver.py priforgepng python3      safe_restart_cluster_master.py  splunk                   tsidxprobe_plo
classify                    genAuditKeys.py        jars           node                        prigreypng    python3.7    scripts                         splunkd                  tsidx_scan.py
ColdStorageArchiver_GCP.py  genRootCA.sh           jsmin          openssl                     pripalpng     python3.7m   scrubber.py                     splunkmon                untarit.py
ColdStorageArchiver.py      genSignedServerCert.py locktest       parse_xml_buckets.py        pripamtopng   pyvenv       searchtest                      splunk-optimize          walklex
coldToFrozenExample.py      genSignedServerCert.sh locktool       pcre2-config                pripnglsch    pyvenv-3.7   setSplunkEnv                    splunk-optimize-lex      wheel

┌──(root㉿kali)-[/opt/splunk/bin]
└─# ./splunk start

Splunk> Australian for grep.

Checking prerequisites...
        Checking http port [8000]: open
        Checking mgmt port [8089]: open
        Checking appserver port [127.0.0.1:8065]: open
        Checking kvstore port [8191]: open
        Checking configuration... Done.
        Checking critical directories...        Done
        Checking indexes...
                Validated: _audit _configtracker _dsappevent _dsclient _dsphonehome _internal _introspection _metrics _metrics_rollup _telemetry _thefishbucket history main summary
        Done
        Checking filesystem compatibility...  Done
        Checking conf files for problems...
        Done
        Checking default conf files for edits...
        Validating installed files against hashes from '/opt/splunk/splunk-9.2.1-78803f08aabb-linux-2.6-x86_64-manifest'
        All installed files intact.
```

```
File  Actions  Edit  View  Help
        Checking critical directories...        Done
        Checking indexes...
                Validated: _audit _configtracker _dsappevent _dsclient _dsphonehome _internal _introspection _metrics _metrics_rollup _telemetry _thefishbucket history main summary
        Done
        Checking filesystem compatibility...  Done
        Checking conf files for problems...
        Done
        Checking default conf files for edits...
        Validating installed files against hashes from '/opt/splunk/splunk-9.2.1-78803f08aabb-linux-2.6-x86_64-manifest'
        All installed files intact.
        Done
All preliminary checks passed.

Starting splunk server daemon (splunkd)...
PYTHONHTTPSVERIFY is set to 0 in splunk-launch.conf disabling certificate validation for the httplib and urllib libraries shipped with the embedded Python interpreter; must be set to "1" for increased security
Done


Waiting for web server at http://127.0.0.1:8000 to be available..... Done


If you get stuck, we're here to help.
Look for answers here: http://docs.splunk.com

The Splunk web interface is at http://kali:8000


┌──(root㉿kali)-[/opt/splunk/bin]
└─# ./splunk status
splunkd is running (PID: 4876).
splunk helpers are running (PIDs: 4877 5092 5157 5207 5234 5238 5240 5525).

┌──(root㉿kali)-[/opt/splunk/bin]
└─# ngrok tcp 9997

┌──(root㉿kali)-[/opt/splunk/bin]
└─# ngrok tcp 9997

┌──(root㉿kali)-[/opt/splunk/bin]
└─# ngrok tcp 9997

┌──(root㉿kali)-[/opt/splunk/bin]
└─# systemctl start apache2

┌──(root㉿kali)-[/opt/splunk/bin]
└─# ngrok http 80

┌──(root㉿kali)-[/opt/splunk/bin]
└─# ngrok tcp 9997

┌──(root㉿kali)-[/opt/splunk/bin]
└─#
```


1 2 3 4
File Actions Edit View Help
ngrok
Full request capture now available in your browser: https://ngrok.com/r/ti
Session Status online
Account Patel Dhruv (Plan: Free)
Update update available (version 3.9.0, Ctrl-U to update)
Version 3.6.0
Region India (in)
Latency -
Web Interface http://127.0.0.1:4040
Forwarding tcp://0.tcp.in.ngrok.io:10314 → localhost:9997
Connections ttl opn rt1 rt5 p50 p90
0 0 0.00 0.00 0.00 0.00


Not secure 54.243.24.3:8000/en-GB/account/login?return_to=%2Fen-GB%2F
splunk>enterprise
dhruv
Sign In
First time signing in?
If you installed this instance, use the username and password you created at installation. Otherwise, use the username and password that your Splunk administrator gave you. If you've forgotten your username or password, please contact your Splunk administrator.
username admin
password The password you created when you installed this instance
First time signing in?
© 2005-2024 Splunk Inc.


New Search
index="main"
2,116 events (29/04/2024 00:00:00.000 to 29/04/2024 02:35:05.000)
No Event Sampling
Job
Events (2,116)
Patterns
Statistics
Visualization
Format Timeline
Zoom Out
Zoom to Selection
Deselect
host
2 Values, 100% of events
Selected
Yes
No
Reports
Top values
Top values by time
Rare values
Events with this field
Values
Count
%
LAPTOP-RIHSHVOT
1,642
77.599%
LAPTOP-J9IF2D63
474
22.401%
Hide Fields
All Fields
SELECTED FIELDS
a host 2
a source 6
a sourcetype 6
INTERESTING FIELDS
a Account_Domain 3
a Account_Name 7
a ComputerName 2
EventCode 100+
EventType 4
a index 1
a Keywords 12
linecount 26
a LogName 3
Prev
1
2
= Perfmon:CPU Load
02:35:00.000
collection="CPU Load"
object=Processor
counter="% Processor Time"
instance=_Total
Show all 6 lines
host = LAPTOP-RIHSHVOT
source = Perfmon:CPU Load
sourcetype = Perfmon:CPU Load


splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
Search
Analytics
Datasets
Reports
Alerts
Dashboards
Search & Reporting
New Search
Save As
Create Table View
Close
index="main"
Today
2,116 events (29/04/2024 00:00:00.000 to 29/04/2024 02:35:05.000)
No Event Sampling
Job
Smart Mode
Events (2,116)
Patterns
Statistics
Visualization
Format Timeline
Zoom Out
Zoom to Selection
Deselect
1 minute per column
List
Format
20 Per Page
Prev
1
2
3
4
5
6
7
8
Next
Hide Fields
All Fields
SELECTED FIELDS
a host 2
a source 6
a sourcetype 6
INTERESTING FIELDS
a Account_Domain 3
a Account_Name 7
a ComputerName 2
EventCode 100+
EventType 4
a index 1
a Keywords 12
linecount 26
a LogName 3
Time
Event
29/04/2024 02:35:00.000
04/29/2024 12:05:00.466 +0530
collection="CPU Load"
object=Processor
counter="% User Time"
instance=_Total
Show all 6 lines
host = LAPTOP-RIHSHVOT
source = Perfmon:CPU Load
sourcetype = Perfmon:CPU Load
29/04/2024 02:35:00.000
04/29/2024 12:05:00.466 +0530
collection="CPU Load"
object=Processor
counter="% Processor Time"
instance=_Total
Show all 6 lines
host = LAPTOP-RIHSHVOT
source = Perfmon:CPU Load
sourcetype = Perfmon:CPU Load

CHAPTER 5: CONCLUSION AND FUTURE WORK

CHAPTER 5 CONCLUSION AND FUTURE WORK

5.1 Conclusion

The development of the IT system log analyzer using Splunk, Splunk Forwarder, and Ngrok has provided a robust solution for monitoring and analyzing log data from various sources within the IT infrastructure. By leveraging Splunk's powerful capabilities for indexing, searching, and visualizing data, alongside Splunk Forwarder for data forwarding, and Ngrok for secure tunneling, the system offers real-time insights into system performance, security incidents, and operational issues.

Key accomplishments of the project include:

- Efficient Log Monitoring: The system efficiently collects and indexes log data from diverse sources, enabling IT administrators to gain insights into system health and performance in real-time.
- Enhanced Security: By analyzing logs for security events and anomalies, the system helps in identifying potential security threats and breaches, allowing for timely mitigation actions.
- Operational Insights: The system provides valuable operational insights by correlating log data from different sources, facilitating proactive maintenance and troubleshooting.
- Scalability and Flexibility: With Splunk's scalability and flexibility, the system can easily accommodate the increasing volume and complexity of log data as the IT infrastructure grows.

5.2 Future Work

While the current implementation of the IT system log analyzer lays a strong foundation, there are several areas for future enhancement and expansion:

- Machine Learning Integration: Integrate machine learning algorithms into the system to enable predictive analytics and anomaly detection for proactive issue resolution and threat identification.
- Automation: Implement automation features for the system to automatically trigger responses or actions based on predefined rules and thresholds, reducing manual intervention and response time.
- Integration with External Systems: Extend the system's capabilities by integrating with other IT management systems, such as ticketing systems or configuration management databases (CMDBs), to streamline incident response and resolution processes.
- Custom Dashboard Development: Develop custom dashboards tailored to specific user roles and requirements, providing relevant insights and visualization to different stakeholders within the organization.

- Enhanced Security Features: Continuously improve the system's security features to adapt to evolving threats and compliance requirements, such as implementing advanced threat detection mechanisms and enhancing data encryption protocols.

- Performance Optimization: Optimize the performance of the system by fine-tuning indexing and search configurations, minimizing resource utilization, and ensuring smooth operation even under high loads.

By focusing on these areas of future work, the IT system log analyzer can evolve into a more advanced and comprehensive solution, further empowering organizations to effectively monitor, analyze, and secure their IT infrastructure.

CHAPTER 6: REFERENCE

CHAPTER 6 REFERENCE

6.1 Reference

- Splunk Documentation: https://docs.splunk.com/
- Ngrok Documentation: https://ngrok.com/docs
- https://community.splunk.com/t5/Getting-Data-In/Forwarding-windows-event-viewer-logs-to-Splunk/m-p/124085#M25589