Team - Enigma Hunters
PS ID - SIH1408
Problem Statement - IT System Log Analyzer

Proposal Document

Table Of Content

1. Dissecting the Problem Statement

The situation at hand involves CRPF units, offices, and personnel deployed across various locations within the CRPF organization. Currently, there is no centralized system in place to analyze the logs generated by the IT systems at these different locations. This absence of a centralized log analysis system gives rise to several challenges:

- **Geographical Dispersion:** CRPF units and offices are spread across different geographical locations within the country. These locations may include urban areas as well as remote and isolated regions. The diverse geography poses logistical challenges.

- **Lack of Centralization:** The core issue is the absence of a centralized system for log analysis. This means that logs generated by IT systems at different CRPF locations are not systematically collected, analyzed, or monitored in one central location.

- **Lack of Analysis Capability:** Even if a centralized system were in place, the existing SIEM solutions may not provide the depth of log analysis required to detect and respond to security threats adequately. This limitation can hinder the overall cybersecurity posture of CRPF and may result in missed or delayed threat detection.

2. Why are existing solutions not enough

Existing SIEM (Security Information and Event Management) solutions, while valuable in many contexts, may fall short in meeting the unique requirements of CRPF's log analysis needs.

- **Geographical Dispersion:** The diverse geographical locations of CRPF units, including remote and isolated areas, pose a significant challenge. Collecting logs from these dispersed units, especially those separated by vast distances, requires a solution capable of reaching these locations efficiently. None of existing log analysis solutions are able to collect log data from geographically isolated devices.

- **Connectivity Limitations:** Many CRPF units operate in regions with inconsistent or limited network connectivity. This presents a fundamental challenge in achieving real-time log collection; sometimes logs will not be transmitted to the central repository because of tracking issues and lack of local storage utilization. Ensuring that the solution functions reliably in such contexts is critical. Which is not taken care of in many existing log analysis solutions according to our study.

- **Data Transport & Security:** Transmitting logs securely over the internet is paramount to protect sensitive information. The proposed solution must incorporate robust encryption and security measures to safeguard log data during transmission. Point to be noted, no existing log solution or protocols allows log transmission over the internet. They are limited to internal networks.

- **Integration Complexity:** Integrating the new log analysis solution with existing IT systems at CRPF units can be complex. The diversity of technologies and infrastructure across units necessitates a flexible and adaptable solution.

According to our research on existing log analysis solutions none of them are made for handling the complexity.

- **Scalability:** As the CRPF continues to expand and evolve, the log analysis solution must be scalable to handle increasing data volumes and accommodate new units seamlessly. We are not so sure how scalable existing solutions are.

- **Resource Efficiency:** Efficiency in resource utilization is crucial. The solution should optimize resource allocation, prevent redundancy in log analysis efforts, and avoid straining limited resources at CRPF units.

- **Comprehensive Log Analysis:** The solution must provide CRPF experts with the tools needed to detect threats, investigate incidents, and respond effectively. All existing log analysis solutions are best for doing some particular operations. None of them are best for all types of operations needed. Which is not an acceptable risk for a sensitive organization like CRPF.

- **Regulatory Compliance:** Meeting cybersecurity regulations and reporting requirements is a legal imperative. The solution must ensure compliance, preventing any potential legal repercussions due to non-compliance. The CRPF units cannot assure that the log data is handled via some 3rd party solution(s) or service providers.

- **Adaptation to Cyber Threats:** The evolving landscape of cyber threats demands a solution that can adapt and evolve alongside these threats. The log analysis system must have the ability to detect and respond to increasingly sophisticated attacks. CRPF units need to depend on the service provider. Which will lead to serious consequences. CRPF has its own department for threat intelligence as “CoBRA” & “RAF”.

A Quick Insight – *Organizations like CRPF have a profound need for a solution that they know the underlying functionality, can fully understand and control. They cannot entrust the analysis of their critical logs or rely on third-party infrastructures for log transfers without raising concerns about data security.*

In essence, agencies like CRPF cannot afford to jeopardize sensitive data. It's important to highlight that CRPF units possess their own threat intelligence, exemplified by "CoBRA", which adds an extra layer of effectiveness to the solution in terms of threat intelligence and stands as a knowledge-base for the ML systems. This localized threat intelligence enables them to identify and address specific threats pertinent to their organization.

However, the sensitivity of this intelligence makes it impractical to share with third parties. Consequently, CRPF finds itself in a situation where complete reliance on third-party solutions is unfeasible. Their foremost priority is safeguarding their data, especially given that logs hold valuable insights that could be exploited by malicious actors. Any data breach could potentially trigger national emergencies, underscoring the critical nature of their data security concerns.

3. Proposed Solution

In response to the pressing need for a comprehensive threat detection strategy for CRPF units, We propose a 7-fold strategic framework designed to fortify the organization against evolving cyber threats. Our solution encompasses and addresses the need of log file collection, analysis, threat detection and reporting. Here's a detailed breakdown of our proposed solution:

- **Collection of Log files:** A meticulous log file collection process utilizing the Syslog Protocol, ensuring comprehensive coverage of log files across all CRPF units. Utilizing a dual-layer encrypted channel with SSH and TLS protocols for secure and confidential data transmission to centralize the collected data in a secured repository to establish a strong foundation for analysis and response.

- **Intelligent Threat Identification:** Within this secure repository, our Monitoring & Analysis Layer springs into action. Equipped with static signature-based threat detection mechanisms, it rigorously scrutinizes the log files, identifying potential threats in real-time.

- **Expert Analysis & Manual Precision:** Known threats are promptly countered through automated playbooks (can be enabled), guaranteeing immediate responses. However, for novel or intricate threats like 0-Days, our cybersecurity experts can step in. Armed with meticulous analysis, they can craft tailored responses, ensuring each threat is met with a precise and effective countermeasure.

- **Data Enrichment & Proactive Detection:** Before reaching the expert's domain, each piece of data undergoes a transformative process. Our data enrichment layer augments the raw data with invaluable insights, converting it into actionable intelligence. Empowering experts to devise custom detection rules, it enables proactive measures against specific threats.

- **Real-time Vigilance & Collaborative Response:** An advanced alerts management system operates in real-time, ensuring our experts are kept informed. Collaboration stands at the core of our strategy. Our experts provide on-site personnel with real-time guidance, fostering a collaborative, harmonized, and most importantly, effective response strategy.

- **Fortified Data Management:** Our system's objective extends beyond merely responding to threats; it's about securing CRPF's data legacy. A meticulously designed data retention engine ensures compliance with lifecycle policies. Log files are securely backed up and archived, guaranteeing data integrity and availability.

- **Continuous Learning & Fortification:** In the ever-changing landscape of cybersecurity, stagnation equates to vulnerability. Hence, our system is perpetually learning. Machine Learning algorithms meticulously analyze expert interactions and external databases, fortifying the system's intelligence against future threats. Regular updates and patches are seamlessly integrated into the network, ensuring CRPF units are perpetually armed with the latest threat detection capabilities.

4. Why Our Solution Stand Out from the Crowd

Usage of Agents

We will be installing agents in CRPF Units which are geographically isolated and then the agents will be transporting the logs using **syslog (**protocol for log collection) and **SSH** (for the transportation). We make these work in conjunction to achieve the transmission over the internet.

Advanced Encryption

Our solution uses a dual layer of encryption as we are using **SSH** (Secure Shell) and **TLS** (Transport Layer Security) using **origin-to-end certificates** for secure transportation of log files from the CRPF units to the Centralised System where we will be performing the analysis.

Comprehensive Threat detection

We will be achieving threat detection with two different methodologies or mechanisms

- **Static Signature-Based Threat Detection Mechanisms:**

  This aspect of our system uses **predefined patterns or signatures** which are already known threats to detect any activity that could be malicious. When such activity is detected it raises alerts or triggers **predefined responses** ensuring rapid action against threat.

- **Custom Detection Rules and Logics:**

  Here in this aspect we will be providing flexibility to our experts for creating custom detection rules and logics which means they themselves can define certain **criteria for**

identifying threats addressing the need of CRPF Units.This particular aspect will be putting us in advantage for **detecting zero-day threat**.

- **Machine Learning-Based Threat Detection:**

  The incorporation of machine learning algorithms that **learn from expert interactions, external and internal databases** (CoBRA, RAF) to ensure that the system evolves and adapts to new and evolving threats, enhancing its threat intelligence over time. This capability adds an additional layer of sophistication to the threat detection system, making it more robust and capable of handling evolving threats.

Efficient Automation

Automating **responses through playbooks** for known threats ensures swift and efficient responses, enabling the system to handle routine security incidents effectively. This feature will allow faster reaction times and reduce human errors.

5. Co-relating problem statement and solution

Centralized Log Analysis

- **Problem Statement:** Lack of a centralized system for log analysis.
- **Solution:** We proposed a centralized system where log files from various CRPF units are collected, stored, and analyzed centrally. This directly addresses the problem of decentralized log analysis. We are using a combination of SSH and SYSLOG to do that.

Secure Data Transmission

- **Problem Statement:** Ensuring secure log transmission.
- **Solution:** Our solution includes a dual-layer encrypted channel using SSH and TLS protocols for transmitting log data. This ensures the secure transmission of data, addressing the security concern mentioned in the problem statement.

Expert Analysis and Interaction

- **Problem Statement:** Lack of expert analysis for assessing threats and breaches.
- **Solution:** Our solution incorporates expert interaction at multiple levels, including static signature-based threat detection, data enrichment, and custom detection rules & logics. This addresses the need for expert analysis as mentioned in the problem statement.

Automated Responses and Manual Intervention

- **Problem Statement:** Handling both known threats (requiring automated responses) and novel threats (requiring expert interaction).

- **Solution:** Proposed solution have provisions for triggering automated threat remedies via playbooks for known threats and allowing expert interaction for 0-Day threats. This dual approach aligns with the problem statement's requirement for handling different types of threats.

Continuous Learning and Improvement

- **Problem Statement:** Need for continuous improvement in threat intelligence.
- **Solution:** Our solution incorporates machine learning, learning from expert interaction, and external databases. This ensures continuous learning, enhancing the system's threat intelligence over time.

Data Retention and Updates

- **Problem Statement:** Managing data retention and ensuring software updates.
- **Solution:** Our system includes a data retention engine with lifecycle policies and ensures that Agent Software can receive updates and patches from the central system. This addresses the concerns related to data retention and system updates.

Our proposed solution not only aligns with the problem statement but also addresses the specific concerns mentioned in a comprehensive manner. It provides a centralized, secure, and expert-driven approach to log analysis, threat detection, and response, thereby effectively resolving the challenges faced by CRPF units in handling IT system logs and cybersecurity threats.

6. Conclusion

This solution not only addresses current vulnerabilities but also ensures adaptability to future threats. By empowering CRPF personnel with knowledge, tools, and support, we pave the way for a more secure and confident organization. Together, we forge a path toward unwavering cybersecurity, safeguarding CRPF's mission-critical operations and upholding its legacy of excellence.