ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 9, September 2023

Enhancing Incident Response with Live Logs: The Significance and Challenges of Maintaining Sufficient Log Retention for Mitigating Cyber Attacks

[1] Ertuğrul AKBAŞ

[1] İstanbul Esenyurt University, SureLog SIEM, Istanbul, Türkiye
Corresponding Author Email: [1] ertugrul.akbas@surelogsiem.com

*Abstract— **In today's rapidly evolving cyber threat landscape, incident response plays a crucial role in safeguarding organizations against cyber attacks. Live logs, real-time records of system activities, have emerged as essential tools for incident response teams to detect and respond promptly to security incidents. However, archive log search speed is often insufficient for mitigating cyber attacks in a timely manner. This research paper explores the significance of live logs in incident response and the challenges associated with maintaining sufficient log retention to effectively mitigate cyber attacks. The paper also examines real-world case studies, recommendations, regulations, and RFP requirements related to live log retention. Additionally, it delves into the benefits of live log monitoring and emphasizes the importance of adhering to industry best practices to strengthen an organization's cybersecurity defense.***

*Index Terms— **Cyber Attacks, Hot Logs, Log Retention, Incident Response.***

I. INTRODUCTION

The escalating frequency and sophistication of cyber attacks underscore the need for robust incident response strategies. Incident response teams rely on live logs, which provide real-time visibility into an organization's IT infrastructure. This real-time monitoring allows security professionals to identify potential threats, track attacker activities, and initiate timely responses. The paper delves into the importance of integrating live logs into incident response protocols and the complexities faced in retaining logs for extended periods.

II. LIVE LOGS: FOUNDATION OF EFFECTIVE INCIDENT RESPONSE

A. Real-time Threat Detection

Live logs enable proactive monitoring, allowing incident response teams to detect security threats as they occur. Timely identification of anomalous behavior, unauthorized access attempts, or suspicious network activities empowers organizations to thwart attacks before significant damage occurs.

B. Enhanced Forensics and Investigation

The availability of live logs facilitates detailed forensic analysis of security incidents. By analyzing log data, incident responders can reconstruct the attack chain, determine the extent of compromise, and gather critical evidence for identifying the threat actors.

C. Rapid Incident Response and Mitigation

Live logs streamline incident response workflows by providing real-time insights. Immediate access to log data empowers security teams to initiate mitigation measures promptly, limiting the attacker's dwell time and minimizing the impact of the breach.

Live logs directly affect TTR (Time to Respond) values.

TTR = (Time of Incident Resolution) - (Time of Event Detection)

Time to Respond calculates the time it takes to resolve a security incident from the moment it was detected. A shorter TTR indicates a quicker response time, which is essential for minimizing the impact of security breaches. TTR values are directly proportional to the live log retention time. If you keep logs live for a longer period, your TTR values get shorter.

III. CHALLENGES IN MAINTAINING SUFFICIENT LOG RETENTION

A. Volume and Scalability

The exponential growth of log data poses challenges in managing and retaining large volumes of logs. There are different technologies in the market. For example, Apache Lucene’s indexed (hot, live) log growth formula:

disk space used (original) = 1/3 original for each indexed field + 1 * original for stored + 2 * original per field with term vectors

There are other technologies utilized by some SIEM vendors that compress both the indexes and raw logs. Organizations must contend with scalability issues and invest in robust log storage and management solutions to accommodate the influx of log data. Log volume increases can be unmanageable both in terms of price and disk size.

B. Log Security and Access Controls

Maintaining the security and integrity of live logs is critical to prevent tampering or unauthorized access. Proper access controls and encryption mechanisms are essential to safeguard log data from both external and internal threats.

C. Compliance and Regulatory Requirements

Various industries and regions have specific data retention and privacy regulations. Organizations must navigate compliance challenges and adhere to legal requirements when retaining live logs for incident response purposes.

IV. CASE STUDIES: REAL-WORLD CONSEQUENCES OF INSUFFICIENT LOG RETENTION PERIOD

A. MITRE and Government Recommendation

Various cybersecurity authorities and government agencies recommend specific log retention periods to support effective incident response. MITRE suggests a minimum of six (6) months to 2+ years of online log retention within the SOC [1]. The Memorandum for the Heads of Executive Departments and Agencies mandates 12 Months of Active Storage (hot log) and 18 Months of Cold Data Storage [2]. The Event Logging Guidance from the Treasury Board of Canada Secretariat established log retention times of 90 days to two years [3].

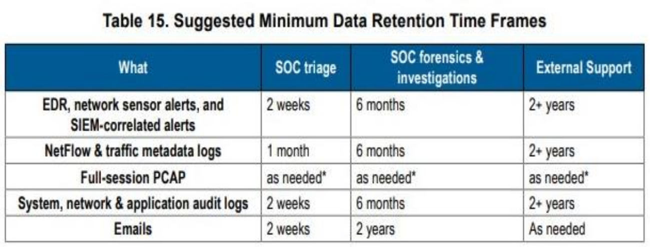

Table 15. Suggested Minimum Data Retention Time Frames

What	SOC triage	SOC forensics & investigations	External Support
EDR, network sensor alerts, and SIEM-correlated alerts	2 weeks	6 months	2+ years
NetFlow & traffic metadata logs	1 month	6 months	2+ years
Full-session PCAP	as needed*	as needed*	as needed*
System, network & application audit logs	2 weeks	6 months	2+ years
Emails	2 weeks	2 years	As needed

Figure 1. MITRE Log Retentions

B. Real-world Attack Cases

Real-world attack cases, such as the Stuxnet, The Verizon Data Breach, The Dominion National Data Breach, SolarWinds hack, demonstrate the criticality of hot-log usage in detecting and responding to cyberattacks. The timeline of the attacks highlights the importance of maintaining long-term hot logs to detect and respond to threats effectively.

The Stuxnet Worm Attack on Iran's Nuclear Program:

The Stuxnet worm serves as a poignant example of how inadequate logging facilitated a prolonged and covert cyber attack [4].

The Verizon Data Breach (2017):

A breach at Verizon Communications highlights the significance of comprehensive monitoring in preventing data exposure [5].

The Dominion National Data Breach (2019):

A long-lasting breach at Dominion National showcases the damaging effects of insufficient logging and monitoring over an extended period [6].

Solarwinds Hack:

After the investigation of the incident, it was discovered that there were not enough logs available. The "Memorandum for the Heads of Executive Departments and Agencies," published by the Executive Office of the President, Office of Management and Budget, mandates 12 months of active storage (hot logs) and 18 months of cold data storage [7].

C. RFP Requirements

Many Request for Proposals (RFPs) require that logs, even if stored in an archive, must be made available within 24 hours [4]. This requirement emphasizes the significance of having logs readily accessible and live to meet response timeframes.

D. Google's Perspective

Google highlights the importance of one-year log retention as both a compliance requirement and a key resource for detecting top-tier threats [5].

E. SANS and Log Management

SANS emphasizes the importance of live logs as a criterion for next-generation SIEM solutions, providing online access to current and archived log data [6].

V. MITIGATING CYBER ATTACKS THROUGH EFFECTIVE LIVE LOG RETENTION

A. Real-time Threat Hunting and Analysis

A well-designed live log retention strategy empowers threat hunters to identify emerging attack trends and patterns. Proactive analysis of live logs helps detect previously unknown threats and improve overall cybersecurity resilience.

B. Incident Response Training and Preparedness

Live logs are invaluable in incident response training exercises. Organizations can simulate real-world attack scenarios using historical log data, ensuring incident response teams are well-prepared to handle future security incidents effectively.

C. Continuous Improvement and Optimization

Organizations must continually optimize their live log retention practices to strike a balance between log storage costs and the need for historical data. Regular review and improvement of log retention policies help enhance incident response capabilities. The biggest challenge here is the size of disks and their associated costs. Organizations can implement solutions that can compress hot (live) logs to efficiently utilize disk space. There are different technologies available that extend disk size or enable log compression [11,12]. For example, Apache Lucene-based solutions like Elasticsearch, Logrthym, and Exabeam offer disk usage extensions, following specific formulas [13]. Many vendors came up with proprietary compression solutions like SureLog [14]

VI. LIVE LOG RETENTION CAPABILITIES IN CURRENT SIEM SOLUTIONS

A. What is SIEM?

SIEM (Security Information and Event Management) is a comprehensive cybersecurity solution that plays a critical role in log collection and incident management for organizations. Its primary purpose is to centralize and analyze logs and security events from various sources within an organization's IT infrastructure.

Log Collection: SIEM solutions gather logs and data from diverse sources, such as servers, network devices, applications, firewalls, antivirus systems, and more. These logs contain valuable information about security events, user activities, system health, and potential threats.

Log Normalization and Correlation: SIEM tools normalize the collected logs, converting them into a consistent format for easy analysis. They also correlate events from different sources to identify relationships and patterns that might indicate security incidents.

Real-time Monitoring: SIEM platforms continuously monitor incoming logs in real-time. This allows security teams to promptly detect suspicious activities or potential threats, enabling faster response and mitigation.

Event Analysis: SIEM engines use predefined rules, algorithms, and machine learning techniques to analyze log data for signs of security breaches, anomalies, or policy violations. These rules can be customized based on an organization's specific security requirements.

Incident Management: When potential security incidents are detected, SIEM systems automatically generate alerts or notifications to inform the security operations team. These alerts are prioritized based on severity, helping analysts focus on critical issues first.

Incident Response: SIEM assists in incident response by providing contextual information about the detected incident. Analysts can access detailed logs and related data to understand the scope and impact of the incident, facilitating quicker and more effective responses.

Log Retention: Retaining log data is a crucial aspect of SIEM and overall cybersecurity strategy. It refers to the practice of storing log data for a specific period, enabling organizations to maintain historical records of security events and incidents.

Compliance Requirements: Many industries and regions have specific regulations and compliance standards that dictate how long organizations must retain log data. Adhering to these requirements is essential for avoiding legal and financial repercussions.

Incident Investigation: Log retention allows organizations to investigate past security incidents thoroughly. By accessing historical logs, security analysts can reconstruct events leading up to a breach, identify the root cause, and implement measures to prevent similar incidents in the future.

Forensics and Auditing: In the event of a security breach or suspected unauthorized access, log data serves as digital evidence for forensic analysis. Detailed logs can help determine the extent of the breach, the methods used by attackers, and any data that may have been compromised.

Trend Analysis and Pattern Recognition: Retained logs offer valuable insights into long-term security trends. By analyzing historical data, organizations can identify recurring patterns, spot potential weaknesses in their defenses, and make proactive improvements.

Log Retention Policies: Organizations should establish clear log retention policies that define the specific types of logs to be retained, the retention periods for each log type, and the mechanisms for secure storage and disposal of logs after the retention period expires.

B. Advantages of SIEM

Properly deployed and optimized centralized log file processing and analysis by the SIEM solution will give companies the following benefits (but is not limited to):

- Centralized log storage - in case of primary log source unavailability, there always will be second log storage with high file availability, integrity and confidentiality;
- Analysis of large amount of logs which will allow to detect:
    - Anomaly activities from all systems
    - activities.
    - Reconnaissance activities.
    - All kinds of cloud platform anomalies.
    - DDoS attacks.
    - Botnet activities.
    - Intrusion attempts.
    - Post intrusion activities.
    - Ransomware.
    - Data theft (data exfiltration) – from both
    - internal and external data theft threats.
    - If someone is trying to modify or access
    - logs in any way.
    - Internal policy misuse.
    - System misconfiguration.
    - System vulnerabilities and exploits as they
    - are being used.
    - System bottlenecks.

- Streamlined compliance reporting for GDPR, PCIDSS, ISO27001, HIPPA and others.
- Increased incident response efficiency.
- "Big picture" of what is happening in the IT environment at any given moment.
- Information on historical events for forensic purposes. In the same time conserving audit log integrity.

To sum up, SIEM will give ability to identify almost all activities within the IT environment from a "single pane of glass". If only system will be configured to look for all of these activities and will understand "context" of these files.

C. Log Retentions of Current SIEM Solutions

There Log retention policies represent a crucial aspect of SIEM (Security Information and Event Management) solutions and have a significant impact on an organization's cybersecurity practices. Varying across different SIEM vendors, these policies are often customizable to align with specific organizational needs and compliance requirements. The determination of log retention periods is influenced by a multitude of factors, including industry regulations, security best practices, data storage capacity, and the organization's incident response requirements.

One critical consideration in log retention is the usage of live logs, which directly affects data storage capacity. Many SIEM solutions initially retain logs in a live state for a specific time period (typically up to 90 days) before transitioning them to archival storage. However, a 90-day retention period may not be sufficient for effective incident response, especially for organizations facing complex and persistent cyber threats.

To further enhance storage efficiency, SIEM solutions often employ log compression and aggregation techniques. Through log compression, log data is condensed without compromising vital information, effectively reducing the overall storage footprint.

Despite the general understanding of log retention policies and their impact on storage, the landscape of SIEM technologies is highly diverse and dynamic. Various SIEM vendors offer a wide array of solutions, each equipped with unique features, performance capabilities, and storage requirements.

For instance, one SIEM solution may require a 10 Terabytes (TB) live log disk size to handle 10,000 Events Per Second (EPS) over a specified time frame. In contrast, another SIEM solution might demand a substantially larger 500 TB live log disk size to accommodate the same EPS volume and time frame. These variations can be attributed to the underlying architecture, log processing algorithms, and data storage methodologies employed by different SIEM products.

The impact of log retention on data storage capacity is a vital aspect that organizations must carefully evaluate when selecting a SIEM solution. Factors such as EPS volume, log volume fluctuations, retention periods, and available storage resources must all be taken into account to ensure the chosen SIEM technology aligns with the organization's specific needs.

VII. CONCLUSION

In conclusion, the evidence presented in this paper underscores the critical importance of live log retention as an indispensable pillar of a robust and proactive cybersecurity strategy. The ability to capture, analyze, and retain logs in real-time empowers organizations to face the ever-evolving cyber threat landscape with increased agility and resilience.

The advantages of live log retention are manifold. Swift detection and response to cyber threats are facilitated, enabling security teams to identify and neutralize potential attacks before they can inflict significant harm. Additionally, the availability of detailed logs bolsters forensic investigations, ensuring a thorough examination of security incidents and aiding in the attribution of cybercriminal activity.

Moreover, live log retention streamlines the process of compliance audits, easing the burden of demonstrating adherence to stringent regulatory requirements. Organizations armed with comprehensive log data can furnish concrete evidence of their security practices, thus instilling trust and confidence among stakeholders.

To fully harness the potential of live log retention, organizations must carefully tailor their log retention policies, considering both the specific nature of their operations and the regulatory landscape they operate within. The implementation of sophisticated SIEM solutions, complemented by the support of a competent Security Operations Center (SOC), is vital in achieving optimal log analysis, correlation, and real-time monitoring.

However, we recognize that the feasibility of live log retention may vary depending on the financial resources and security budget of each organization. While its benefits are indisputable, the decision to invest in live log retention requires a prudent assessment of cost-effectiveness and risk mitigation.

The research undertaken in this study aimed to not only shed light on the escalating cybersecurity risks faced by organizations but also to present a compelling view on the efficacy of live logs in fulfilling crucial cybersecurity needs, particularly in the domain of identification. The conclusions drawn from the research provide a foundation for organizations to make informed decisions on their cybersecurity strategies, laying the groundwork for a proactive defense against cyber threats.

In summary, live log retention serves as a dynamic safeguard, enabling organizations to respond rapidly to emerging threats, conduct thorough investigations, and adhere to regulatory obligations. By embracing live log retention, organizations can reinforce their cyber resilience and emerge as stalwarts in safeguarding sensitive data and preserving the trust of their stakeholders. As cybersecurity challenges continue to evolve, it is imperative for organizations to embrace live log retention as a steadfast ally in their unyielding pursuit of digital security.

REFERENCES

[1] https://www.mitre.org/sites/default/files/2022-04/11-strategies-of-a-world-class-cybersecurity-operations-center.pdf

[2] https://www.whitehouse.gov/wp-content/uploads/2021/08/M-21-31-Improving-the-Federal-Governments-Investigative-and-Remediation-Capabilities-Related-to-Cybersecurity-Incidents.pdf

[3] https://www.canada.ca/en/government/system/digital-government/online-security-privacy/event-logging-guidance.html

[4] https://en.wikipedia.org/wiki/Stuxnet

[5] https://www.foxnews.com/tech/verizon-data-breach-14-million-customers-reportedly-exposed

[6] https://www.hipaajournal.com/dominion-national-proposes-2-million-settlement-to-resolve-class-action-data-breach-lawsuit/

[7] https://www.wired.com/story/the-untold-story-of-solarwinds-the-boldest-supply-chain-hack-ever

[8] http://vadodarasmartcity.in/vscdl/assets/tenders/17.09.2020/2021_499-1.pdf

[9] https://chroniclesec.medium.com/retaining-logs-for-a-year-boring-or-useful-9b04c1e55fba

[10] https://www.sans.org/media/vendor/evaluator-039-s-guide-nextgen-siem-38720.pdf

[11] https://www.peerspot.com/articles/features-of-today-s-siems-requirements-for-today-s-attacks-and-breaches

[12] https://www.peerspot.com/questions/why-hot-data-and-cold-data-differences-in-siem-solutions-are-not-discussed-sufficiently

[13] https://lucidworks.com/post/estimating-memory-and-storage-for-lucenesolr

[14] https://surelogsiem.com/features/

[15] J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility,” IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, Jan. 1959.

[16] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, “Rotation, scale, and translation resilient public watermarking for images,” IEEE Trans. Image Process., vol. 10, no. 5, pp. 767-782, May 2001.