Overview of Security Information and Event Management Systems

Cosmin MĂCĂNEAȚĂ
Omega Trust, Bucharest, Romania
cosmin.macaneata@omega-trust.ro

Organizations face continuous risks of cybersecurity breaches and malicious activities. Reviewing logs is a critical activity to identify these threats, but the large volume of systems and data often presents an insurmountable challenge. As IT infrastructures expand, logs multiply exponentially, making the traditional manual log analysis extremely difficult and prone to missing key events. Also, this high volume of information combined with the uncorrelation of logs makes traditional approaches ineffective, especially in detecting sophisticated attacks. The solution to this challenge is deploying a Security Information and Event Management (SIEM) system, who collects, correlates, and analyze disparate logs from various sources in real-time, offering a complete overview of the organization's security. By automating the log review and providing detailed information, SIEM not only resolves the problem of log overload but also significantly improves threat detection and incident response capabilities. This paper presents the overview of SIEM systems and highlights the ways in which they can overcome these issues.

Keywords: *SIEM, Security Event Management, Cybersecurity, Information Security, Security Logs analysis*
DOI: 10.24818/issn14531305/28.1.2024.02

1 Introduction

In today's digital age, the exponential rise in cyber threats has become an unavoidable reality for organizations across industries. The relentless advancement of technology brings with it a myriad of opportunities but also exposes entities to an ever-evolving spectrum of cybersecurity risks. From sophisticated phishing attacks to ransomware, data breaches, and insider threats, the arsenal of potential cyber threats is expansive and formidable.

As organizations strive to safeguard their sensitive data, intellectual property, and customer information, the need to proactively detect and mitigate these threats has never been more critical. The current evolution of Internet of Things (IoT) is greatly changing the impact of cyber-attacks, because these have nowadays a much larger attacking surface [1].

Central to identifying these threats lies the analysis of system logs, which serve as digital footprints recording various activities across networks, applications, and devices. The review of logs is an activity very important for organization and their auditors, as well [2].

However, the sheer volume and diversity of these logs present a monumental challenge. Manual review and analysis become laborious, time-consuming, and error-prone due to the massive amounts of data generated by numerous interconnected systems. For security teams, looking through these logs to pinpoint anomalies or potential threats becomes a Herculean task, often resulting in critical events going unnoticed or delays in response.

A SIEM platform acts as a central point in the organization, who collects, aggregate, correlate logs, and security events from disparate sources in real-time. The primary function of a SIEM system is to automate the analysis of logs and security events. This automation significantly reduces the burden on cybersecurity professionals, allowing them to focus on critical tasks rather than drowning in an ocean of data. By employing correlation rules and algorithms, SIEM systems can identify patterns, anomalies, and potential threats that might otherwise go unnoticed within the deluge of logs.

Moreover, SIEM solutions offer invaluable benefits beyond log management. They facilitate proactive threat detection by continuously monitoring for suspicious activities, deviations from normal behavior, or indicators of compromise. This real-time visibility enables rapid response and mitigation, thwarting potential threats before they escalate into full-fledged security incidents.

Additionally, the scalability and flexibility of SIEM platforms make them adaptable to diverse organizational infrastructures. Whether an organization operates on-premises, in the cloud, or follows a hybrid model, SIEM systems can ingest and analyze logs from disparate sources, providing a unified view of the security landscape.

Furthermore, compliance requirements and regulatory standards mandate robust log management and threat detection practices. SIEM solutions aid in fulfilling these obligations by providing comprehensive audit trails, facilitating incident investigations, and generating compliance reports, thus ensuring adherence to industry regulations.

The mounting complexity and volume of cybersecurity threats demand a paradigm shift in how organizations approach threat detection and response. The manual review of logs, once a conventional method, is no longer feasible or effective in today's intricate digital ecosystems. The implementation of a SIEM system emerges as the indispensable solution, streamlining log management, enhancing threat detection capabilities, and empowering organizations to fortify their defenses against an evolving landscape of cyber threats.

2 SIEM architecture overview

The architecture of a SIEM system involves a set of agents installed on hosts throughout an organization's infrastructure. The agent represents a software component deployed on individual hosts or endpoints within an organization's network. Their main function consists in continuous monitoring, ensuring that they capture every relevant activity on the host.

Agents monitor files and logs generated on respective hosts, in real-time, collecting data relevant to information security (for example successful or unsuccessful logins, unauthorized user behaviors, modification of system files etc.).

After the data is collected by the agent, it is transferred to the server, where it is correlated and analyzed in real-time using detection rules.

This architecture enables robust monitoring, allowing identification of potential security threats and anomalies.

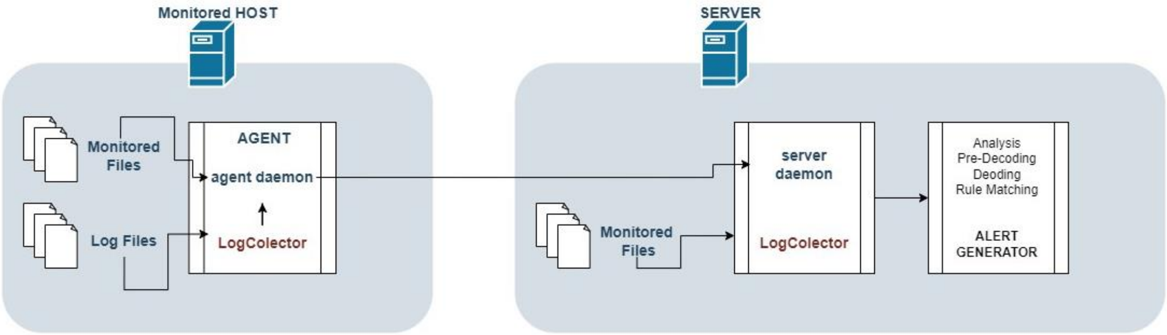

Fig. 1. SIEM architecture overview

2.1 Technical Processes of SIEM Log Collection and Processing

The main objective of a SIEM is to gather, compile, and examine log data produced by the whole IT infrastructure of an enterprise. Logs must be collected and processed through a number of technological procedures.

Log sources and collection

Logs are produced by a variety of IT environment devices and applications. These can include servers, databases, firewalls, routers, antivirus programs, and more. Different formats, including syslog, Common Event Format (CEF), and proprietary formats unique to vendors, can be used to generate logs.

To collect log data, SIEM solutions usually install agents or collectors on every server or device. Local logs are gathered by these agents and then sent to a centralized SIEM server. As an alternative, some systems route logs to the SIEM server directly without the requirement for local agents using protocols like SNMP or Syslog. This is useful for devices (for example network routers or firewalls) that do not allow installation of agents. Also, the SIEM can be configured to allow the manual ingestion of logs files, a situation useful if agents cannot be installed on particular systems or if these cannot be configured to send logs automatically.

Log normalization and tokenization

Distinct log sources frequently have distinct forms. These logs are normalized by SIEM systems into a standard format to facilitate analysis. This procedure includes gathering pertinent data and organizing it into a consistent format.

Tokens are individual pieces of information that are separated out of logs. A log entry could be tokenized into the following fields, for instance: event type, source IP, destination IP, and timestamp.

Processed logs are kept in a database for compliance and historical analysis. Depending on the SIEM solution, either relational or NoSQL databases can be used.

Log indexing, searching, correlation and analysis

Logs are indexed using a variety of parameters, including time, source IP, event type, etc., to enable quick search and retrieval. SIEM solutions give security analysts a way to browse and query the indexed logs through an interface.

In terms of correlation, SIEM systems look for patterns and connections among various log items using pre-established correlation criteria. For instance, if several things happen at once and point to a security issue, an alert might be sent out.

Statistical models are utilized by certain SIEM systems to identify anomalies that can point to a security risk.

A SIEM can come with built-in rules, which allow generation of alerts based on predefined criteria, but it is recommended that the SIEM allow the security team to add new rules or customize existing ones, to match the context of the organization. For example, a built-in rule detecting brute force attempts that can trigger an alert if it identifies more than 10 unsuccessful login attempts in 100 milliseconds, should be modified by the security team who wants to have stricter security to generate the alert if it identifies less unsuccessful login attempts in less time (e.g. 5 login attempts in 30 milliseconds).

Alerting, notification, and incident response

Alerts are generated by the SIEM system when it finds a security incident based on correlation rules or abnormalities. Security analysts receive alerts for additional research via a variety of methods, including dashboard notifications, SMS, and email.

Security teams may respond quickly in response to threats that are detected when SIEM solutions are integrated with incident response protocols.

SIEM solutions include executive summaries and compliance reports for auditing needs, along with reporting capabilities.

Configurable data retention policies are a standard feature of SIEM systems, which help to control log storage over time.

When combined, these components allow the SIEM system to efficiently gather, handle, and evaluate log data in order to locate and address security issues in the IT infrastructure of a company. Depending on the SIEM system being used, the precise details could change.

Auditing and compliance

Due to the fact that it uses agents which are installed on the monitored devices, the SIEM can collect relevant information about the security configuration of the respective devices and can compare it versus standards or regulations. For example, the SIEM can highlight if the passwords rules match the industry standards in terms of complexity, can highlight uninstalled security patches, can identify security vulnerabilities depending on the software version, can detect active insecure protocols etc.

Therefore, the SIEM is a valuable instrument for performing audits to evaluate the organizations' compliance with security benchmarks and regulatory requirements. Through this comprehensive analysis, the organizations can identify areas for improvement and take proactive measures to improve their overall security posture, in general.

Reporting

The SIEM systems offer powerful reporting functionalities, generating detailed and customizable reports based on correlated security data. These reports include critical information about security events, compliance status and incidents trends. By presenting these reports, the SIEM systems significantly improve organization's ability to proactively address security challenges and enhance the resilience against cybersecurity risks.

2.2 Log Collection Methods in SIEM Systems

Logging data from a variety of sources within an enterprise's IT infrastructure is one of the core components of SIEM systems. The various techniques used for log collection – including agent-based and agentless approaches – are examined in this section. The significance of log normalization and parsing in guaranteeing correct and significant data for analysis within the SIEM platform will also be covered.

SIEM systems gather logs from a variety of sources, including as servers, endpoints, network devices, and apps. Integrating logs from many sources presents many difficulties, including dealing with problems like data format inconsistencies, log volume, and the requirement for ongoing updates to accommodate new technologies. Please find below the methods which may be used for logs collection:

a) Agent-Based Log Collection:

One popular technique in SIEM systems is agent-based log collecting, which involves placing small software agents on certain IT infrastructure components. The agents hold the responsibility of keeping an eye on and gathering log data locally, following which it is transmitted to the central SIEM server. Among the agent-based approach's main benefits are:

- **Granular Visibility:** Agents provide a granular level of visibility into the activities and events on each endpoint, allowing for detailed analysis and correlation of security events.
- **Real-time Monitoring:** Since agents continuously monitor and transmit data, real-time analysis and response to security incidents become feasible, enhancing the system's overall efficacy.
- **Reduced Network Load:** By processing and compressing data locally, agent-based collection minimizes the impact on network bandwidth, making it a suitable option for organizations with bandwidth constraints.
- **Secure Communication:** Agents often employ secure communication protocols to transmit log data, ensuring the confidentiality and integrity of sensitive information during transmission.

b) Agentless Log Collection (Syslog):

Devices transmit their log data to a central SIEM server through agentless log collection, which is frequently accomplished through protocols like syslog and eliminates the need for specialized software agents. The agentless log collection is particularly useful for devices where agents cannot be installed, such as networking devices or Industrial IoT [3].

The following are some important aspects of this type of collection:

- **Simplified Installation:** This approach significantly eases the installation process since there is no need to install and manage software agents on each individual device.
- **Interoperability:** Many devices and applications support standard log formats, such as syslog, facilitating interoperability and ease of integration into SIEM systems.

- **Reduced Resource Overhead:** Agentless methods often result in lower resource overhead on individual devices, as they do not require additional software components running locally.
- **Challenges with Data Normalization:** Despite its simplicity, agentless log collection may face challenges related to data normalization, as logs from different sources may vary in format and structure.

c) Manual Log File Collection:

SIEM systems may occasionally use manual log file collection, in which security staff members or system administrators compile log files by hand from a variety of sources. This approach has a unique set of use cases and considerations, despite being less popular and requiring more work:

- **Isolation of Critical Systems:** Manual log file collection might be preferred for critical systems or devices that require special handling, allowing administrators to carefully control the transfer of log data.
- **Audit Trail:** Human intervention in log collection can provide an additional layer of oversight and auditability, ensuring that specific logs are intentionally collected and reviewed.
- **Challenges with Scalability:** Manual log file collection becomes impractical in large-scale environments due to its inherent scalability limitations and the increased likelihood of errors.

Table 1. Type of logs collection for each data source [4]

Log source	**Agent-based log collection**	**Agentless log collection**
Core Windows infrastructure	✓	✓
Database platforms	✗	✓
Endpoint security solutions	✗	✓
Firewalls, NGFWs, IDSs, and IPSs	✗	✓
Hypervisors	✗	✓
Linux and Unix systems	✓	✓
Routers and switches	✗	✓
Vulnerability scanners	✗	✓
Web servers	✗	✓
Servers	✓	✓

Log source	**Agent-based log collection**	**Agentless log collection**
Workstations	✓	✓
Cloud platforms	✗	✓

In general, agent-based approach best suits collecting logs from systems which allow installation such as Windows, Linux/Unix, servers, and workstations.

The agentless approach best fits firewalls, hypervisors, routers, switches, vulnerability scanners, web servers, and cloud platforms, offering simplicity. Choice depends on the possibility to install the agent, as the agent-based approach usually provides more information than the agentless one.

2.3 Example Scenario of how a log entry is transformed into security incident

Consider a web server that generates access logs. Each log entry includes information such as the timestamp, source IP address, requested URL, HTTP status code, and user agent.

Log Entry: *[2023-01-15T12:34:56] [192.168.1.100] "/admin/login" 401 "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/63.0.3239.132 Safari/537.36"*

The transformation into alert and later in security incident is done as follows:

- The SIEM agent or collector on the web server captures this log entry.
- Log parsing extracts key fields: timestamp, source IP, requested URL, HTTP status code, and user agent.
- Regular expressions are applied to normalize and standardize fields.
- For example, the user agent field might be normalized to identify the browser type and version.
- The log entry is tokenized into specific pieces of information: timestamp, source IP, URL, status code, and user agent.
- Each token is now a discrete piece of data for analysis.
- Correlation rules are defined to identify suspicious activities. In this case, a rule might be triggered if multiple failed login attempts (HTTP status code 401) happen in a short period of time from the same source IP.
- Statistical models may also detect anomalies, such as a user agent string that deviates from the usual patterns.
- The correlation engine identifies a pattern of failed login attempts within a specified time window from a specific source IP.
- An alert is generated, indicating a potential brute-force attack on the "/admin/login" URL.
- The alert includes details from the log entry, such as the timestamp, source IP, and URL.
- Severity levels may be assigned based on the number of failed attempts and other factors.
- The SIEM system sends an alert notification to security analysts via their preferred channels, such as email, SMS, or integration with a collaboration platform.
- Security analysts review the alert and initiate an incident response process.
- Automated playbooks may be triggered to block the source IP, update firewall rules, or perform other predefined actions.
- Details of the alert and the incident response actions are logged for auditing purposes.
- Reports may be generated to analyze trends, track incidents, and measure the effectiveness of the security measures taken.

In this way, the security event data captured in the log file, is transferred to the server where it is matched against the set of rules which may trigger the alert and raise the security incident.

2.4 Common challenges of SIEM Systems

Systems for managing Security Information and Events (SIEM) are essential for cybersecurity, but they also provide a unique set of difficulties. Organizations looking to deploy and maintain efficient SIEM solutions must comprehend these difficulties. These challenges, commonly relate to:

a. False Positives

False positives, or situations in which the system mistakenly perceives routine activity as a security threat, are one of the ongoing problems with SIEM systems. To overcome this obstacle, correlation criteria must be adjusted, and threat intelligence must be used to increase warning accuracy.

b. Alert Fatigue

Security analysts may become alert fatigued due to the massive number of alerts issued by SIEM systems. One of the biggest operational challenges is sifting through a large number of alerts to find real dangers. The root cause of alert fatigue is due to filtering and correlation problems [5]. This problem can be lessened by putting automated response systems in place and ranking notifications according to their seriousness. The use of machine Learning and AI as well as data visualization tools can be a solution [6].

c. Complexity of Rule Management

It can become difficult to manage correlation rules in a SIEM system, particularly in large-scale settings. To have an effective SIEM, it is critical to have a set of detection rules adapted for the needs of the organization. Therefore, effort should be placed especially at the beginning of the deployment to customize the set of rules and, continuously during the operation of the system, to refine and adapt the rules to the evolving situation. Retaining a current set of rules in line with the changing threat environment calls for specialized knowledge and resources.

d. Scalability Issues

The amount of log data generated by various sources increases as businesses expand. To manage this growing log volume, SIEM systems need to grow without compromising functionality. The deployment of cloud based SIEM or the allocation of extra hardware resources may be required to address scalability issues [7].

e. Integration of New Technologies

The SIEM systems may face challenges when required to integrate logs from emerging platforms and new technologies. To accommodate various log formats and sources, IT infrastructures, cloud environments, and IoT devices necessitate ongoing updates and customization. To overcome this, continuous research should be made to update the SIEM to the latest technological changes.

II.5. Future developments in the SIEM systems

As the cybersecurity threats are continuously evolving it is expected that SIEMs will develop accordingly and will be an important defense mechanism in the future ecosystems. There can be identified a few trends in the evolution of the SIEM systems:

a. Development of AI/ML capabilities

It is expected that the next generation of SIEMs will continue the trend to integrate AI/ML technologies in their core engines to offer improve detection, reduce alert fatigue, and provide predictions based on traffic and behavior. One step forward for cyber threat detection, mitigation, and prevention is to consider AI in SOAR solutions which would be ideally integrated in SIEM platforms [8].

b. Cloud integration

Cloud technology is an enabler for current systems because it solves the problem of scalability. SIEM systems will be no exception, and it is expected that more of these systems will use cloud technologies. Traditionally this has been done in the past especially for storage reasons, as the SIEM is highly demanding in this regard. However, it is expected that in the future, SIEMs will run natively in cloud environments to benefit from the scalability and flexibility advantages of these kinds of infrastructures.

c. Extended Detection and Response (XDR)

XDR platforms, which integrate and correlate data from multiple security sources beyond traditional SIEM data sources, may become more prevalent. This allows for more comprehensive threat detection and response capabilities across the entire IT environment.

d. Integration with Security Orchestration Automation and Response (SOAR)

SIEMs will further the trend of SOAR by integrating automation and orchestration capabilities. This advancement streamlines incident response processes and enhances threat remediation efficiency. By automating repetitive tasks and orchestrating security workflows, SIEMs empower security teams to respond to threats more rapidly and effectively. This synergy between SIEM and SOAR technologies maximizes the effectiveness of security operations, enabling proactive threat detection and mitigation. As cybersecurity landscapes evolve, the integration of SIEM and SOAR becomes increasingly vital for bolstering organizational resilience against emerging threats.

e. Mobile technologies integration

The increase of mobile devices usage presents security challenges to organizations which need to be addressed by the future SIEM systems. In today’s environment, mobile devices are used more and more for business as many tools and services run on this kind of devices: email, cloud storage, collaboration tools, mobile banking, video conferencing etc.

In addition, many organizations use BYOD (Bring Your Own Device) policies. There’s a current tendency for employees to utilize both company-issued and personal devices for work purposes. [9].

This trend brings forth numerous potential security issues, as BYOD devices fall outside the management of the organizations IT and usually lack key security measures such as encryption or monitoring. To solve these issues, it is expected that the SIEMs systems will increase their range to monitor mobile devices to provide real time alerts.

f. Privacy Protection

Amid increasing concerns surrounding data privacy, SIEM systems are expected to integrate features ensuring the protection of sensitive information and compliance with privacy regulations such as the EU Regulation – 2016/679 – GDPR. As organizations wrestle with heightened scrutiny over data handling practices, SIEM platforms may evolve to incorporate functionalities that fortify data protection measures such as personal data pseudonymization.

3 Conclusions

In conclusion, the SIEM system is the primary instrument for supervising the cybersecurity environment of organizations. The SIEM system provides a centralized platform for monitoring, aggregating, and analyzing security-related data across an organization’s network. This is done through the deployment of agents on hosts, who monitor files and logs generated by the applications residing on the hosts. These logs are transferred to a centralized server where the data is correlated and matched against a set of rules to identify patterns of cyber threats. By applying this approach, the SIEM can detect anomalies, identify events and generate alerts who can potentially represent security incidents. Using the SIEM, the security teams may have complete overview of the security within their organization and may respond to threats in real-time, thus minimizing the impact of cyber-attacks and reducing risks, in general.

The main advantage of a SIEM is its invaluable ability to correlate and analyze huge volumes of data in real-time, collected from large numbers of systems, activity which cannot be done using the traditional manual approach.

Besides threat detection, by providing detailed audit trails and compliance reports, a SIEM system also helps organizations achieve and monitor their compliance with regulations and standards.

While a SIEM system brings significant security benefits to an organization, it also presents challenges like false positives, alert fatigue, and scalability issues. False positives can lead to wasted resources, alert fatigue can overwhelm security personnel, and scalability issues can hinder system performance as the organization grows. However, these challenges can be effectively managed by fine-tuning correlation rules, implementing automated response mechanisms, and ensuring the SIEM system is designed for scalability. Through proactive strategies and adjustments, organizations can optimize their SIEM systems to better detect and respond to security threats while minimizing the impact of these challenges. A key point to have an efficient SIEM is to keep it updated to the ever-evolving cyber threats and emerging technologies.

In the future, SIEMs will evolve to keep pace with cyber threats, leveraging advanced AI and machine learning algorithms and will continue the trends of integrating XDR and SOAR capabilities. This evolution will ensure proactive threat detection and real-time response capabilities, ensuring enhanced cybersecurity posture for organizations in the face of evolving threats.

References

[1] M-G. Ioniță, V-V Patriciu, "Secure Threat Information Exchange across the Internet of Things for Cyber Defense in a Fog Computing Environment", Available: http://revistaie.ase.ro/content/79/02%20-%20Ionita,%20Patriciu.pdf

[2] A-M. Suduc, M. Bîzoi, F. Gh. Filip, "Audit for Information Systems Security", https://revistaie.ase.ro/content/53/04%20Suduc,%20Bizoi,%20Filip.pdf

[3] H. Zahid, S. Hina, M. F. Hayat, G. A. Shah "Agentless Approach for Security Information and Event Management in Industrial IoT", Electronics 2023, 12(8), 1831; https://doi.org/10.3390/electronics12081831

[4] A. Saravanan, "Log collection 101: Covering the basics", [Online], Available: https://www.manageengine.com/log-management/cyber-security/log-collection-101.html

[5] E. Kidmose, M. Stevanovic, S. Brandbyge, J. Pedersen, Featureless "Discovery of Correlated and False Intrusion Alerts" IEEE Access 2020, 8, 108748–108765. [Google Scholar] [CrossRef]

[6] T. Ban, T Takahashi, S Ndichu, D Inoue, "Breaking Alert Fatigue: AI-Assisted SIEM Framework for Effective Incident Response", Available: https://doi.org/10.3390/app13116610

[7] H. Willbanks, "Legacy vs. Cloud-native SIEM: Weighing the Pros and Cons", [Online], Available: https://www.exabeam.com/siem/legacy-vs-cloud-native-siem-weighing-the-pros-and-cons/

[8] G. González-Granadillo, S. González-Zarzosa and E Diaz, Security, " Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures"

[9] J. Bradley, J. Loucks, J. Macaulay,R. Medcalf, L. Buckalew. "BYOD: A Global Perspective; Cisco Survey Report; Cisco: San Jose, CA, USA" [Google Scholar]

Cosmin MĂCĂNEAȚĂ has graduated the Faculty of Cybernetics Statistics and Economic Informatics in 2004 and he holds a Master's Degree in IT Project Management. He is the founder and the current Managing Partner of Omega Trust, one of the leading Romanian cybersecurity auditing and advisory firms, with more than 1000 projects successfully delivered at national level and other 16 countries. Cosmin has extensive experience and holds worldwide relevant certifications in cybersecurity, IT auditing, ethical hacking, and data privacy.

Between 2021 and 2023, as Project Director, Cosmin has led the cybersecurity research team of Omega Trust in successfully developing and launching to market an innovative Security Information and Event Management (SIEM) system.