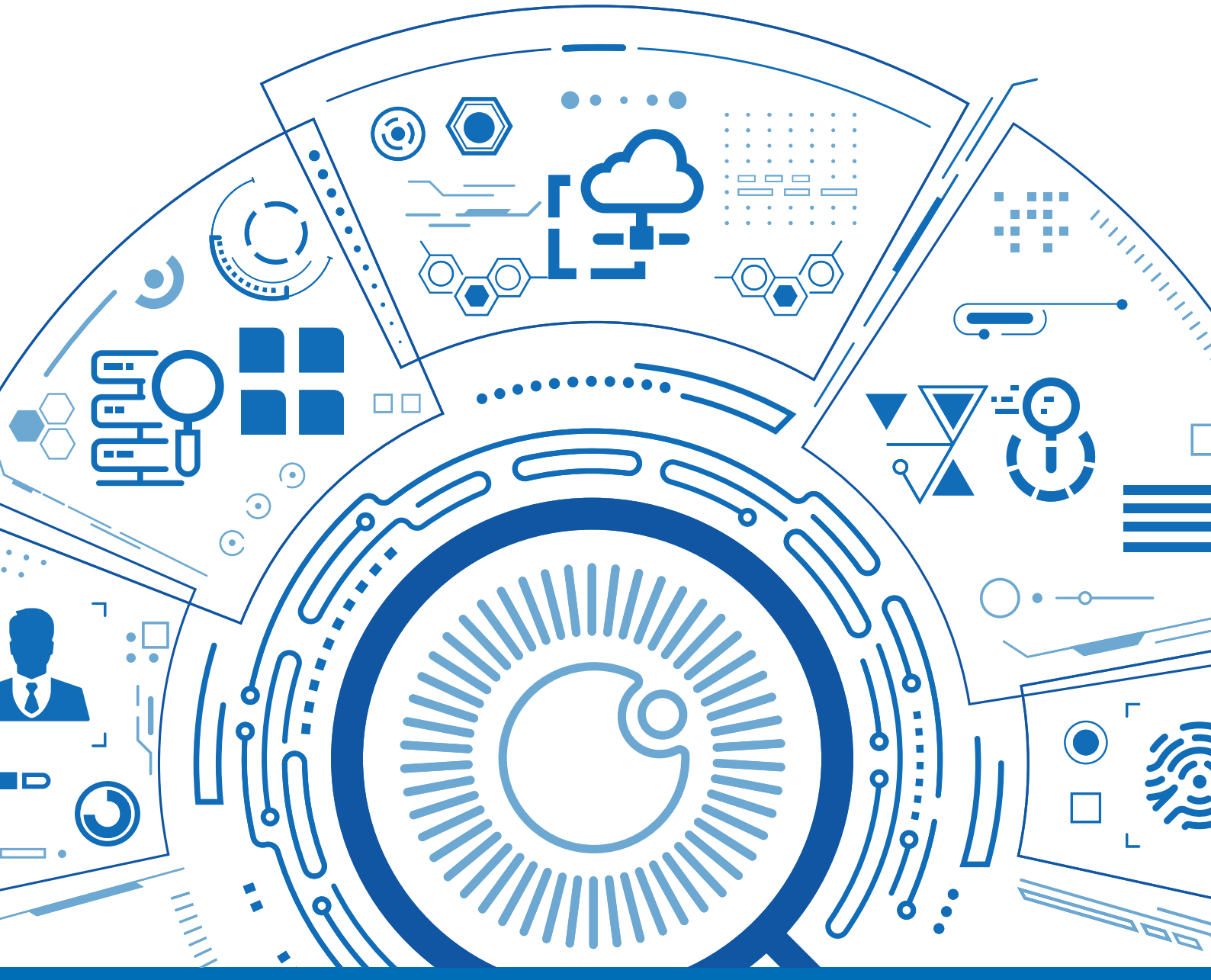
Implementing SIEM and SOAR platforms: practitioner guidance



Australian Government

Australian Signals Directorate

National Cyber Security Centre

a part of GCHQ

Communications Security Establishment

Canadian Centre for Cyber Security

Centre de la sécurité des télécommunications

Centre canadien pour la cybersécurité

Te Tira Tiaki

Government Communications Security Bureau

National Cyber Security Centre

PART OF THE GCSB

NISC

内閣サイバーセキュリティセンター

National center of Incident readiness and Strategy for Cybersecurity

JPCERT CC®

National Cyber and Information Security Agency

NÚKIB

The authors would like to acknowledge the industry partners that contributed to this publication.

Table of contents

Introduction

This publication provides high-level guidance for cyber security practitioners on Security Information and Event Management (SIEM) and Security Orchestration, Automation, and Response (SOAR) platforms. While this guidance is primarily intended for practitioners within government and critical infrastructure organisations, it can also be used by practitioners in any organisation that is interested in, or currently using, SIEM and/or SOAR platforms.

This publication has four sections that:

- defines SIEM and SOAR platforms
- outlines these technologies' potential benefits for cyber security posture
- sets out the challenges involved in properly implementing these platforms, to deliver those benefits
- provides best practice principles to which practitioners can refer throughout each stage of implementing a SIEM and/or SOAR: procurement, establishment, and maintenance.

The recommendations in this publication should be treated as generic advice; each organisation should tailor the collection, centralisation, and analysis of logs to its specific environment and risk profile.

This publication is one of three in a suite of guidance on SIEM/SOAR platforms:

Implementing SIEM and SOAR platforms: Executive guidance

This document is primarily intended for executives. It defines SIEM/SOAR platforms, outlines their benefits and challenges, and provides broad recommendations for implementation that are relevant to executives.

Implementing SIEM and SOAR platforms: Practitioner guidance

This document is intended for cyber security practitioners. In greater technical details, it defines SIEM/SOAR platforms, outlines the benefits and challenges, and provides best practice principles for implementation.

Priority logs for SIEM ingestion: Practitioner guidance

This document is intended for cyber security practitioners and provides detailed, technical guidance on the logs that should be prioritised for SIEM ingestion. It covers log sources including Endpoint Detection and Response tools, Windows/Linux operating systems, and Cloud and Network Devices.

This guidance should also be read alongside Best Practices for Event Logging and Threat Detection, which provides high-level recommendations on developing a logging strategy.

Note on terminology: as discussed below, all SIEM platforms perform log collection, centralisation, and analysis. Some SOAR platforms have an in-built SIEM and can perform these functions themselves. Many other SOAR platforms integrate with a separate SIEM and leverage the SIEM's log collection, centralisation, and analysis. Only SOAR platforms, however, perform automated response functions. In this guidance, references to SIEMs specifically are referring to log collection, centralisation, and analysis functions. If your organisation uses a SOAR with an in-built SIEM, this content will be relevant to the SOAR's log collection, centralisation, and analysis. References to SIEM and/or SOAR platforms concern guidance that is applicable to both platforms. References to SOARs specifically concern the automated response functions that only SOAR platforms can perform.

1. Defining SIEM and SOAR platforms

What is a SIEM platform?

A SIEM is a type of software platform or appliance that collects, centralises, and analyses log data from sources within a network or system for the purpose of cyber security. If properly implemented for this purpose, a SIEM platform automates the collection and centralisation of important log data that would otherwise be scattered across a network, thus making it easier for a human security team to navigate. Unlike some other log collection and centralisation tools, a well-configured SIEM then applies a predefined baseline of business-as-usual network activity, rules and filters to analyse and correlate the log data. This analysis can allow the SIEM platform to detect unusual activity on the network, which may represent a cyber security event or incident. Most SIEM products enhance their analysis by incorporating up-to-date threat intelligence.

SIEM platforms use query languages to retrieve and analyse log data. The query language used generally varies between different SIEM products.[1] Queries are run cyclically or as needed (for example, to investigate anomalous activity or conduct forensic analysis after an incident has occurred) and will return the information in the form of dashboards, reports, or alerts.

What is a SOAR platform?

A SOAR is a software platform that automates the response to anomalous activity detected on a network. The platform automates the response by applying predefined 'playbooks', which combine incident response and business continuity plans to dictate some of the actions to be taken when a specific security event occurs. These automated actions do not replace human incident responders but can streamline the response to anomalous activity. [2]

Some SOAR platforms are designed to integrate with a SIEM platform and leverage its collection, centralisation and analysis of log data. Others perform log collection, centralisation and analysis themselves. A SOAR can also be integrated with other security tools, such as firewalls, endpoint security solutions and vulnerability scanners.

1 One example way to bridge the gaps between different query languages may be by using Sigma, a YAML (Yet Another Markup Language) based format that allows detection rules to be textually presented and implemented in the SIEM platform specific query language. See: https://github.com/SigmaHQ/sigma.

2 Please note SIEMs and SOARs are just one form of detection technology. Other forms, such as canary technology, are not discussed here.

2. Potential benefits of SIEM and/or SOAR platforms

Through automation, SIEM and/or SOAR platforms can powerfully enhance a human security team's visibility of the network and their ability to detect and respond to cyber security events and incidents. However, each of their benefits will only be delivered if the SIEM and/or SOAR is properly implemented (see Section 3).

Ultimately, enhancing network visibility and the detection of, and response to, cyber security events and incidents will improve the security and integrity of information stored on the network. A properly implemented and continually maintained SIEM and/or SOAR platform can help to prevent system unavailability, theft or modification of sensitive data, and the loss of control over critical networks.

By streamlining detection and response times, a properly implemented SIEM and/or SOAR platform can help prevent a costly cyber security incident and avoid the expensive process of removing adversaries from a network. These potential, incident-based costs far outweigh the costs involved in properly implementing a SIEM and/or SOAR platform.

Enhancing visibility

By automating log collection and centralisation, analysing this data, and presenting the analysis in dashboards and reports, a SIEM makes it easier for a human security team to see and interpret what is happening across the network. This information would otherwise be extremely complex and scattered.

Another benefit of implementing a centralised log solution or SIEM platform is streamlining access to event data during an investigation into a suspicious event, eliminating the need to log into each machine to manually collect logs.

Collecting and centralising log data is also critical for any organisation's implementation of the Australian Signals Directorate's Essential Eight Maturity Model and the Cybersecurity Infrastructure Security Agency's (CISA) Cybersecurity Performance Goals (CPGs). If compliance requirements are the primary driver for your entity's implementation of a tool to collect and centralise logs, SIEM and/or SOAR platforms might not be the most appropriate or cost-effective option. Other tools are available. For example, CISA's Logging Made Easy (LME) is a no-cost, open-source platform that centralises log collection. LME is intended for small- and medium-size organisations that need a log management and threat detection system and can be downloaded directly from CISA's LME GitHub page.

Enhancing detection

A well-configured and continually maintained SIEM platform also enhances the detection of cyber security events and incidents by generating swift alerts about unusual activity on the network, which prompt the security team to investigate further. While some detections, like anomaly detection, require a large amount of data to work effectively, a well-maintained SIEM platform should assist in determining whether a detection is a false positive.

Automating log collection and securing the integrity of centralised logs also improves detection by protecting logs from unauthorised modification and deletion – a tactic some malicious cyber actors employ to maintain network or system access.[3] Organisations should also improve their level of log event standardisation by using SIEM platforms, which may enable the use of common detection rules from community or vendor sources. Using common detection rules enhances the ability to detect events and decreases the time needed to deploy these rules.

Enhancing response

SIEM and/or SOAR platforms can also improve the capacity to respond to cyber security events and incidents. By generating swift alerts about unusual activity, a SIEM platform can allow your organisation to either intervene before an event escalates into an incident or act quickly and limit the damage where an incident has occurred.

Effective log collection, centralisation, and alerting by a SIEM platform also provides incident responders with data that allows them to analyse what has happened and what needs to be done. Log centralisation may be critical where malicious cyber actors attempt to modify or delete event logs to hide their tracks.[4]

A SOAR's automated response function can also make the overall response to cyber security events and incidents more effective. A SOAR platform will never replace human incident responders; however, by automating some actions involved in responding to specific events and incidents, it can allow staff to focus on the more complex and high-value problems the event or incident has generated. To the extent that it does automate the response, a SOAR platform also matches the speed of malicious cyber actors who are using automated attack techniques.

3 See Identifying and Mitigating Living Off the Land Techniques | Cyber.gov.au.

4 This technique was observed in the Volt Typhoon campaign. For further detail, see: PRC State-Sponsored Cyber Activity | Cyber.gov.au.

Use Case: Detecting and responding to LOTL threats

Malicious cyber actors often leverage legitimate tools and features of an operating system or environment to achieve their objectives. This technique, known as Living off Off the Land (LOTL), may be difficult to detect and mitigate. However, a well-implemented SIEM or SOAR platform can enhance threat- hunting capability to successfully detect LOTL tactics, techniques, and procedures (TTPs), in the following ways:

- Collecting and centralising logs from network devices, endpoints, and other systems may assist with maintaining log integrity, prevent malicious cyber actors using LOTL TTPs from modifying/deleting logs in order to maintain their network access, and provides an overarching holistic view of all activities within the environment.
- Implementing rules that analyse collected logs greatly assists in detecting:
  - LOTL scripting actions;
  - command or tool execution; and
  - attempted system access considered unusual in the environment.
- Baselining normal user and application behaviour in a SIEM allows the platform to identify potential malicious activities, by analysing deviations from the baseline. By utilising a SOAR's machine learning models, behavioural analysis can greatly enhance response capability to detect abnormal patterns in user or system behaviours, including deltas associated with common tool execution and horizontal or vertical privilege escalation attempts.
- A SIEM or SOAR should ingest threat intelligence information, enabling an up-to-date response to LOTL TTPs. By combining logs, behavioural indicators, and threat intelligence feeds, correlation activities can analyse user and/or system activities to determine whether an action is legitimate or malicious.
- A SOAR's automation and response capability leverages playbooks to limit potential damage which may be caused by flagged or suspicious activities. Examples of automated response capabilities include, but are not limited to system quarantine, segregation or blocking of network traffic, and credential revocation.
- SIEM or SOAR platforms that are integrated with other technologies, such as endpoint detection and response (EDR) tools, can enhance the overall response to a LOTL attack by identifying and blocking tools, scripts, or other legitimate management solutions in real time. (LOTL activity is difficult to identify and prevent, often resulting in system compromise and the malicious cyber actor successfully achieving their objectives. Implementing SIEM or SOAR platforms greatly improves an organisation's ability to detect, respond to, and recover from, to LOTL techniques).

3. Challenges of implementing SIEM and/or SOAR platforms

For practitioners considering whether to procure a SIEM and/or SOAR platform, this section sets out some key challenges. Neither platform is a 'set and forget' tool. These platforms can only enhance visibility, detection and response capabilities if they are properly implemented and continually maintained by skilled personnel. The central challenge for any SIEM platform is to ensure that data is properly ingested from log sources and to set up effective detection mechanisms. The central challenge for SOAR platforms is tailoring an effective automated response to specific events.

Normalisation of ingested data

A SIEM collects, processes, and analyses data from a variety of sources within a network or system. These diverse sources may include, firewalls, intrusion detection/prevention systems (IDS/IPS), endpoints, and applications. This creates a challenge for log analysis, because each source may possess a different log format. Organisations should address the following issues to effectively implement a SIEM:

- proper parsing of data from both unstructured and structured logs.
- standardisation of terminologies across all collected logs, e.g., "src_ip" and "sourceAddress".
- time synchronisation for logs that are collected in different time zones.

Collection coverage across the environment

The ingestion of log data should also consider the coverage of the organisation's environment and risk assessment prioritisation. For example, if an organisation has only deployed log/data replication mechanisms to 10 of its 12 Active Directory (AD) servers, then the SIEM platform has a 'blind spot' to events occurring on the two AD servers not being ingested by the SIEM platform. Conversely, an organisation may accept through a risk assessment process that the SIEM platform will ingest only a certain percentage of the organisation's desktop platforms, as a representation of desktop activities.

Log centralisation versus log analysis

High-volume log consumption is more likely to occur if the SIEM platform is used primarily as a log centraliser. A SIEM should not be used for this purpose. A variety of other tools are available to organisations if their primary objective is log centralisation for the purposes of compliance or auditing. These tools, such as data lakes and alternative storage mechanisms, should be used to centralise logs that do not have value in terms of identifying potential cyber security events, threats, or incidents. This is most economic and may enable security teams to collect only the logs they require through the SIEM, without having to filter out less useful logs.

Further, log analysis is likely to be less effective when a SIEM platform is treated as a central repository for all logs, rather than logs specifically identified for security purposes (see below).

Achieving effective log analysis

To achieve effective log analysis, a SIEM platform should be carefully configured for the organisation's unique IT environment and business requirements. Analysis is effective when the SIEM produces both:

- true positives (alerts when events or incidents are occurring)
- true negatives (no alerts when there are no events or incidents occurring).

This requires human personnel to ensure that the SIEM collects and centralises the appropriate types and quantity of log data, as well as the right rules and filters. This is a significant and continuous undertaking – the SIEM should be continuously configured and tuned over time, as the environment, platform features, and threat landscape continually change.

Alternatively, if the SIEM platform uses incomplete or inconsistent data, or if the applied rules and filters are under-sensitive, it will produce false negatives: no alerts when real events or incidents are occurring. This presents significant risks in terms of detection and response, as security teams may become reliant on the platform's alerting.

Conversely, the SIEM platform may be configured to consume a high volume of logs and apply rules and filters that generate many false positives: alerts when no event or incident is occurring. Amidst a high volume of false positives, security teams may experience 'alert fatigue' and miss or delay their response to real events and incidents. This kind of imbalance is more common, with security teams preferring to collect and centralise too much data, which makes it more difficult to develop effective rules and filters.

If the SIEM is not configured to achieve effective log analysis, this may impact the functionality of any SOAR platform that is integrated with it, leading to significant consequences.

See the related publication in this series, Priority logs for SIEM ingestion: Practitioner guidance, for detailed guidance on the logs that the authoring agencies recommend prioritising for SIEM ingestion.

The risks of automating responses

SOAR platforms also should be carefully configured for the organisation's unique environment, which requires a range of staff with specialist skills. These may include:

- security professionals to identify which parts of the response should be automated
- platform engineers to design the automation
- developers to determine how response automation may affect bespoke/in-house developed products or services
- legal counsel or governance risk and compliance experts, to determine any risks and regulatory consequences arising from response automation.

If the SOAR's response functionality is not properly configured and maintained, the platform may misidentify regular user or system behaviour as an event or incident and take automated measures to isolate and respond. This can cause disruption of varying levels to service delivery. If staff are sceptical about automating responses or lack the confidence and expertise required to properly implement the SOAR over time, it may sit unused after being procured at significant expense.

For these reasons, SOAR platforms are usually not suitable for immature environments – that is, environments that lack an existing SIEM, have only a newly established SIEM capability, or lack an experienced security team. In general, investing in the proper implementation of a SIEM platform and achieving effective log analysis is a higher priority than implementing a SOAR.

Resource intensity

Properly implementing a SIEM and/or SOAR platform involves significant and ongoing costs. These may include:

- the upfront and sustained licensing and/or data use costs of the platform
- the costs of hiring and retaining staff with rare specialist skills in SIEM and/or SOAR implementation
- the upfront costs of upskilling existing staff to ensure that they have the skills to configure the platform
- the sustained costs of continually investing in staff training to ensure staff can maintain the platform and mature it over time as the technology, environment, and threat landscape continue to change
- the sustained costs involved in staff's governance and maintenance of the platform
- upfront and sustained service costs if the organisation outsources the configuration, implementation and maintenance of the platform.

It can be costly to outsource configuration and maintenance. For organisations that manage sensitive information or provide critical or unique services, it is recommended to develop an in-house capability to meet regulatory obligations or deliver reliable, bespoke capability. In-house staff may have greater understanding of the network, which is invaluable in identifying unusual or suspicious activity on it. While there are some 'always abnormal' activities that external service providers can easily identify, subtler activity may go unnoticed without in-depth knowledge of the environment. Security teams also need to have the authority to ask users about their behaviour on the network, for the purpose of preventing insider threats or investigating credential theft. Additionally, outsourcing may produce visibility gaps, work duplication and communication difficulties. If choosing to outsource, the authoring agencies recommend researching different SIEM and/or SOAR vendors that best suit the needs of your organisation.

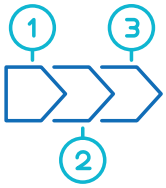
Organisations that do not outsource the configuration and maintenance of the SIEM and/or SOAR platform should expect to dedicate multiple staff to these tasks on a full-time basis. It is also worth noting that operating a SIEM can involve long periods of highly stressful work with which staff are likely to need support.

4. Best practice principles for implementing a SIEM and/or SOAR

This section provides 11 best practice principles to which practitioners can refer throughout each stage of implementing a SIEM and/or SOAR platform: procurement, establishment and maintenance.

Procurement

1. Define the scope of implementation for your organisation.
2. Consider a SIEM product with a data lake architecture.
3. Consider a SIEM product that can correlate data from multiple sources.
4. Look for the hidden costs of different products.
5. Invest in the training, not just the technology.



Establishment

6. Establish a baseline of business-as-usual activity on the network.
7. Develop a standard for the collection of logs.
8. Incorporate the SIEM into your organisation's enterprise architecture.

Maintenance

9. Evaluate threat detection.
10. Reduce log ingestion through pre-processing.
11. Test your SIEM and/or SOAR's performance.

Principles for procurement

When choosing between different SIEM and/or SOAR products, practitioners should refer to the following principles at the procurement stage, based on available resources.

1. Define the scope of implementation for your organisation

Before investing in a SIEM and/or SOAR platform, organisations should develop a proof of concept (POC) for implementing the platform/s. Different organisations will have different definitions of a POC for implementing a SIEM and/or SOAR platform. However a thorough POC can prevent your organisation from committing prematurely to a platform that presents unexpected challenges and delays during implementation.

Organisations should first identify the scope of the POC for SIEM and/or SOAR implementation. The POC may focus on a particular business area or on addressing a key risk or threat model. The POC should also consider whether available vendors meet or exceed these key technical and operational requirements. Correctly identifying the scope at first instance will assist your organisation to deliver a fully operational, end-to-end POC.

Identifying an accountable System Owner is essential to answering the questions listed below. Authoring agencies recommend that the System Owner is part of the team responsible for maintaining the platform/s. The System Owner is also responsible for governance, or overseeing and authorising changes to the technology (see Principle 8).

When designing a POC for implementation of a SIEM and/or SOAR platform, the following questions should be addressed:

- What is the primary objective of the implementation?
- Who are the relevant stakeholders in the use and outcomes of the implementation?
- What in-scope risks, threats, and use cases will the implementation address?
- Which data sources should be prioritised for ingesting?
- Will third-party or integration dependencies hinder the implementation, such as single sign-on (SSO), Application Programming Interface (API) integrations, and/or existing security tools?
- Should the platform(s) be implemented on-premises or through a Software as a Service (SaaS) provider? Does implementing a SaaS-based SIEM/SOAR reduce maintenance overhead?
- What integration with on-premises, hybrid, and cloud or multi-cloud architectures will be required?
- Which team(s) are responsible for the implementation, including ongoing maintenance of the platform(s)?
- Are human resources, including the security analysts and engineers who will operate the platform, available for several weeks of implementation?
- Do the available human resources have sufficient expertise for their role in the implementation?
- Does the vendor have sufficient support and are there other professional service support options?
- Who will be accountable for the implementation?
- Does your organisation have existing incident response processes?

In addition to the questions above, the System Owner should ensure the POC for implementation of a SIEM platform:

- addresses organisation-specific ingestion requirements, including configuring the end device to log the events required and the forwarder to look for those logs
- makes security the primary focus, with log ingestion limited to security feeds
- ensures searches and apps are run that are relevant to the potential events and incidents
- ensures new log sources are only added once existing log sources have been audited for use cases, relevant detections and usefulness
- chooses an architecture that meets compliance, security and privacy requirements

Whether the POC is for a SIEM and/or SOAR, the System Owner should also ensure it:

- includes an evaluation process to determine if the POC was successful. This process should further ensure that the feedback from the security analysts and engineers who are operating and using the platform, or the end-users, can be incorporated to determine whether the POC was successful. Elements of end-user feedback should cover topics including user interface, investigation workflows, and the automation capabilities.

The authoring agencies recommend that organisations implementing a SIEM consider the following data ingestion table for initial licensing and storage/compute requirements:

Organisation staffing	Minimum	Suggested ideal
<50 (minimal organisation)	50 GB	200 GB
50–400 (small organisation)	150 GB	300 GB
400–2000 (medium organisation)	250 GB	600 GB
2000–5000 (medium/large organisation)	500 GB	1.5 TB
>5000 (large/portfolio organisation)	1 TB	2.5 TB

2. Consider a SIEM product with a data lake architecture

When choosing between different SIEM products, the authoring agencies recommend considering a product that has incorporated – or can incorporate – a data lake architecture. Architecture patterns vary between different SIEM products. The architecture determines the SIEM's potential uses by determining how it promulgates data, centralises logs, and restricts staff access to the information it holds.

A data lake architecture within a SIEM has all security related logs replicated to a log repository. The SIEM draws logs for analysis from this repository instead of receiving its own feed. The log repository should be appropriately secured to maintain the integrity and confidentiality of its data. Organisations could also consider hosting the log repository and SIEM platform in a segregated monitoring enclave (such as one that only allows log traffic in) to improve log protection. However, hosting the SOAR platform in a segregated monitoring enclave may limit the opportunities for remediation actions to occur.

See Annex A for an explanation of common SIEM architecture patterns, including more detailed information about the data lake architecture.

3. Consider a SIEM product that can correlate data from multiple sources

The authoring agencies also recommend considering a SIEM platform that can analyse and correlate log data from multiple sources, such as legacy information technology (IT) and cloud environments within the network. Organisations should give preference to a SIEM functionality and security parameters that can incorporate other data sources into analysis, such as cyber threat intelligence feeds.

Additionally, organisations should confirm that the chosen SIEM platform can operate with the expected workloads and integrate seamlessly into existing security operations.

Note: Organisations should consider the costs regarding log management and integration, as ingesting all logs into a SIEM platform can be expensive. The authoring agencies have provided further advice regarding the selection of log sources in Priority logs for SIEM ingestion: Practitioner guidance. That guidance recommends a threat-led approach and selectively ingesting logs that balance security and cost, such as EDR alerts that leverage the same event IDs and logs.

4. Look for the hidden costs of different products

Any organisation that is considering procuring a SIEM and/or SOAR platform should plan for the upfront and sustained costs involved and look for the hidden costs involved in some products and services (see above under Resource intensity).

Some of these platforms are designed to integrate with other products made by the same vendor, making it necessary to purchase those other products to achieve full security functionality. A high concentration of IT products from a single vendor may result in a single point of failure, the consequences of which may be costly to remediate. Additionally, it may be difficult to achieve full visibility of the environment if the chosen platform struggles to integrate logs from products outside the vendor's product suite. This is a significant issue for networks containing products from multiple vendors and legacy IT. Ultimately, a misjudged selection of product may force your organisation to switch to another vendor, which can be very costly.

In some cases, competitive licensing prices may be negated by high costs in actually implementing the platform or switching from one vendor's platform to another. As detailed above, licensing costs represent only a portion of the total investment needed to effectively implement these platforms.

Most SIEM pricing models are based on the quantity of logs ingested by the platform. Some SIEM licences include a certain number of 'free' logs, while others cap log ingestion according to a pre-purchased amount. If the licensing model does not put any cap on log ingestion, organisations should be mindful of the potential to incur very significant costs if ingestion is not carefully managed. Some SIEM platform default settings and recommendations will tend to increase the quantity of logs that are ingested, resulting in greater ingestion costs.

Risks of creating a single point of failure and incurring associated costs may also arise where establishment and maintenance of the platform/s is outsourced to third parties (managed service providers or contracted external parties). Organisations should consider what outsourcing these aspects of implementation means for the geographical location of data storage and/or analysis, to which regulatory requirements may apply. Additionally, some third parties will charge organisations for the cost of running queries or training their own security staff on the platform the organisation uses.

However, if organisations are prudent in choosing an initial SIEM and/or SOAR platform or switching to a new vendor, they will make a long-term security investment that will ultimately outweigh unanticipated costs.

5. Invest in the training, not just the technology

Achieving uplift in visibility, detection, and response through SIEM and/or SOAR platforms requires especially skilled and dedicated human resources. As discussed in Section 3, for some organisations, outsourcing the establishment and maintenance of the platform/s will not be suitable. This presents a prime opportunity for upskilling and mentoring current staff.

Any organisation that intends to procure a SIEM and/or SOAR platform and develop an in-house capability should plan to dedicate significant resources to training staff in implementing the platform. This means upfront investments in training personnel, so they have the skills to establish and maintain the platform(s). Organisations should also expect to keep paying for training over time, to ensure that staff can validate their skills and maintain and mature the platform as the technology, environment and the threat landscape continually change.

Basic training topics should include:

- SIEM fundamentals, such as platform types
- logging fundamentals
- log manipulation and filtering
- querying and searching
- log analysis and investigations
- attack frameworks and tactics, techniques, and procedures (TTPs), such as MITRE ATT&CK[5]
- alerting and reporting
- building dashboards
- maintaining the health of SIEM data feeds and indexes.

Principles for establishment

The authoring agencies recommend practitioners refer to the following principles in the establishment stage of a SIEM and/or SOAR platform: initially configuring the platform(s) for their organisation's unique environment and creating standards and processes for implementation.

6. Establish a baseline of business-as-usual activity on the network

Before deploying a SIEM to collect, centralise and analyse log data, the security team should establish a baseline of business-as-usual (BAU) activity on the network. To establish a baseline, the team should assess installed tools and software, account behaviour, network traffic, service meshes, and system intercommunications. A SIEM can only effectively alert the security team to cyber security events and incidents once an accurate baseline of normal network activity has been established.

Baselines are not built in a day. Rather, depending on the organisation's complexity, they are developed over at least several weeks of collecting logs, developing queries, investigating unusual user activity and filtering and tuning alerts, until the security team has a good sense of (BAU). While a baseline is being established, there is a significant risk that if the system is already compromised, malicious activities will be accepted as normal. The authoring agencies recommend that during

5 MITRE and MITRE ATT&CK are trademarks of the MITRE Corporation.

baselining, organisations make extensive efforts to validate whether activities are actually normal, rather than malicious.

Once it is established, the baseline should then be continually maintained to reflect changes in the network. Security teams should be aware of assets that are retired or added to the network, as well as changes in user behavioural patterns. Often, a failed query will return no report. This can be mistaken for there being nothing to report, rather than the platform failing to identify events and incidents because its baseline no longer reflects the organisation's real business-as-usual activity.

The authoring agencies recommend establishing an internal process for regularly reviewing the baseline, as well as making ad hoc amendments through enterprise architecture processes (see Principle 8).

In addition to the baseline principle, organisations could establish a threat-hunt framework that both validates the baseline and introduces other hunt techniques, such as TTPs. Threat hunting is one way to test whether the baseline excludes all malicious behaviour and to determine whether new attack methods apply to the organisation's environment. An organisation's threat-hunting framework should incorporate threat intelligence in order to replicate new indicators of compromise and adversary TTPs.

7. Develop a standard for the collection of logs[6]

The SIEM System Owner should have a set of pre-defined baseline logging requirements for applications and systems. They should have an agreed approval mechanism to allow deviations from these parameters, which enable and disable additional logging if needed. High levels of logging can negatively affect device performance, log storage systems and SIEM performance, as well as increasing network traffic. Organisations should consult vendor documentation to determine whether alternative logging options can reduce these impacts.

The absence of logs from systems should be a concern to organisations. Organisations should have processes in place to identify when systems have stopped generating logs and telemetry. The SIEM platform should be configured to analyse both real-time and historical logs and other telemetry for matches against indicators of compromise (IOCs).

A log retention and archiving policy should be implemented. Various organisation-specific factors will affect the period of time for which logs must be retained and therefore remain accessible. From a security perspective, the retention period needs to accommodate the organisation's mean time to detect threats; if threats are detected quickly, less retention time is needed. Logs should also be retained for long enough to support cyber security incident investigations. The retention policy should also address how logs are aged off/deleted/purged once they no longer provide value or the retention timeframe has been met. Further, your organisation may be subject to regulatory requirements concerning the log-retention period.

The authoring agencies therefore recommend tailoring the retention period to your organisation's particular risk profile and regulatory requirements. For organisations new to detection, a starting point for retention may be a minimum of one year for logs that record administrative and security-related events and 90 days for informational logs that are used to contextualise other events.

Organisations are also encouraged to implement an event logging policy that is focused on capturing high-quality and high-fidelity cyber security events in order to aid network defenders in correctly identifying cyber security incidents. Useful event logs enrich a network defender's ability to assess security events and distinguish false positives from versus true positives.

6 https://www.cyber.gov.au/resources-business-and-government/maintaining-devices-and-systems/system-hardening-and-administration/gateway-hardening/gateway-security-guidance-package-gateway-operations-management

For more on log collection strategies, please see
Best practices for event logging and threat detection | Cyber.gov.au

8. Incorporate the SIEM into your organisation's enterprise architecture

The SIEM System Owner should be kept aware of all new data sources and changes to data sources in the network to ensure that the SIEM continues to effectively ingest all relevant log data. As such, the System Owner should be a standing member of the organisation's enterprise architecture group and/or change control board. The team(s) responsible for implementing the SIEM should also maintain regular contact with IT administrators to ensure IT systems are optimally tuned.

Optimal tuning may include asset management (for example, ensuring that the SIEM has visibility over new assets) and behavioural monitoring (for example, ensuring the SIEM is alerting activity from users that should no longer have access to the network or should not be accessing certain parts of it). IT administrator decisions to collect certain logs for compliance reasons may also impact the SIEM's functionality – the less useful the logs ingested by the SIEM, the less useful the SIEM will be. Security teams should further be aware of, and have input into, significant infrastructure decisions that may impact the SIEM's visibility, such as cloud adoption. Ensuring the security team works alongside developers so that new products generate useful security logs may save time down the line if security teams need to access those logs following an incident.

Organisations should not assume that all logs from newly installed systems ought to be sent to the SIEM. New systems and the log data they generate should be investigated for their value in identifying unusual network activity. For example, a new application may rely on existing group policies for authentication mechanisms, such that its authentication events are already sent to the SIEM.

Principles for maintenance

Practitioners should refer to the following principles throughout the maintenance phase of a SIEM and/or SOAR platform lifespan, thereby continually tuning the platform(s) and maturing detection and response capabilities over time.

9. Evaluate threat detection

After a SIEM platform is deployed, organisations should establish a procedure that regularly tests and tunes its alerting mechanisms. In line with previous advice from the authoring agencies, organisations should consider implementing a standardised naming convention for alerts with links to MITRE ATT&CK phases for faster incident response triage. All alert rules should reflect the organisation's threat model and risk profile and be tuned accordingly.

As the threat model changes, so should the detection capabilities. For example, if an organisation's key business process is migrated from a legacy to a supported platform, this should be reflected in the SIEM. Evaluating the threat detection capabilities could determine the need to both retire obsolete alert rules, as well as to test new alert rules, in order to reflect the new requirements as specified by the threat model. Organisations could also consider using a more formal evaluation framework for use case analysis, such as Management, Growth and Metrics assessment (MaGMa)[7].

7 https://www.betaalvereniging.nl/wp-content/uploads/FI-ISAC-use-case-framework-verkorte-versie.pdf

10. Reduce log ingestion through pre-processing

As discussed, excessive log ingestion is a common challenge with implementing a SIEM. Ingesting too many logs can be costly, lead to false positive alerts and strain the platform's data processing capacity, which may delay or degrade its performance. Implementing customised parsing rules that focus on information that is essential to security analysis can further reduce the amount of data ingested by extracting only the necessary fields from logs.

Organisations should pre-process logs before they are ingested into the SIEM to reduce the likelihood of these issues arising. Pre-processing can occur at one of three points: at source/host, at forwarder/replication, or at SIEM ingestion. See Annex B for detailed explanations of each of these pre-processing methods.

11. Test your SIEM and/or SOAR's performance

Organisations should conduct exercises to test the performance of their SIEM and/or SOAR platform. These exercises can be run in-house or by external service providers, such as penetration testers. To further enhance a SIEM's performance, it is essential to ensure that the log ingestion and processing pipelines are free from bottlenecks. Only by testing these platforms can the security team be confident that the required logs are being produced, those logs are being ingested into the SIEM, the SIEM is producing true positives and true negatives, and the SOAR is actioning alerts as intended. It is important that organisations continuously evaluate the performance of these platforms to help manage evolving business goals and emerging threats.

Organisations should schedule regular and repeated exercises that test well-known TTPs in order to track improvement over time. These regular exercises should be balanced against ad hoc tests with new attack vectors. Key areas for testing include:

- Active Directory and PowerShell modification
- common adversary techniques, such as a dump of LSASS.exe
- Golden Ticket, Living Off the Land and DCSync activity
- detection of command & control (C2) activity / unusual traffic to suspicious domain
- network scanning / reconnaissance from inside the network.

For more on these attacks and the importance of monitoring Active Directory modification, please see Detecting and Mitigating Active Directory Compromises.

ANNEX A: SIEM architecture patterns

The following approaches are common architectures for SIEM platforms. In the following examples, the 'log repository' refers to a generalised location for centralising logs, such as a data lake or S3 bucket. The 'source' refers to a data source within the environment. 'Current' refers to a repository or feed of recent logs, while 'old' refers to a repository or feed of older logs.

Log repositories have several uses as illustrated below. As a general point, directing the original or a replicated copy of the original log to a log repository is critical for investigations, as processing (through a SIEM or other processing tool) can damage log integrity and the originating data source may only retain logs for a limited time period

1. Data lake or repository-first approach

The authoring agencies' recommended approach is for data sources to send logs to the log repository first, rather than directly to the SIEM. In this architecture (illustrated below), all logs are replicated to a log repository and the SIEM draws recent logs for searching and processing from the current repository, instead of receiving its own feed.

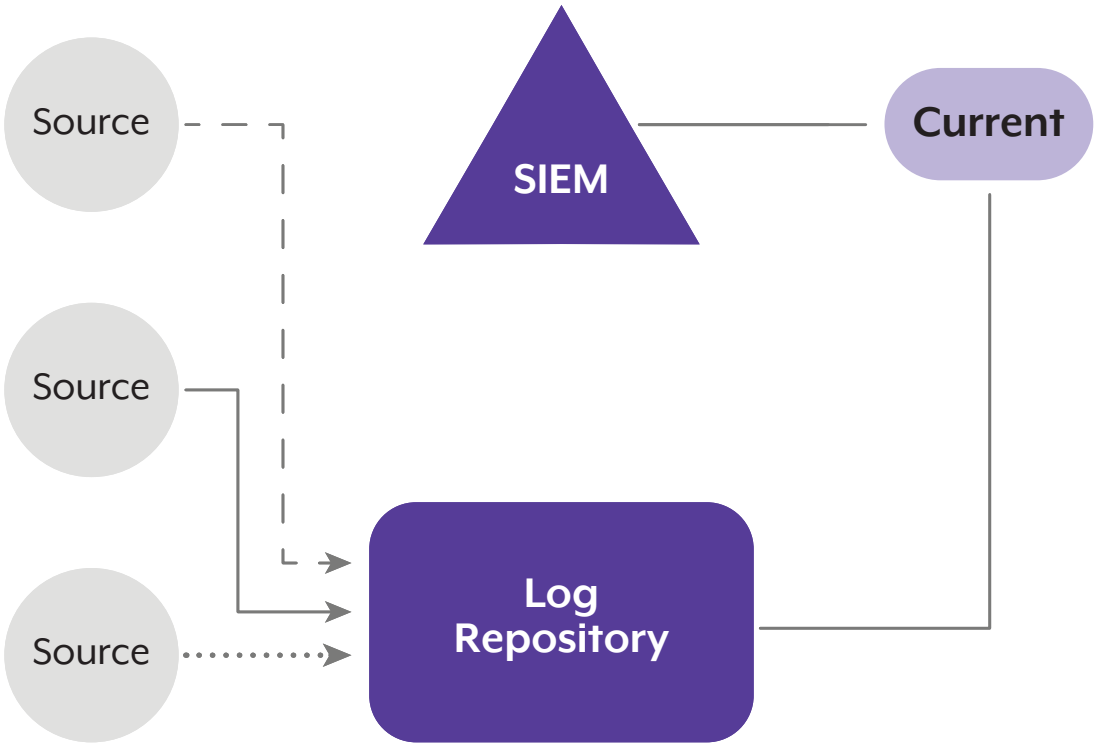


Figure 1. Data lake or repository-first approach visualisation

2. Dual log replication between log repository and SIEM

In the following architecture example, all data sources replicate and direct logs to both the SIEM and the log repository. Management of the logs ingested into the SIEM is achieved using the current and old repositories, which are themselves managed through the SIEM. This management includes any activities for retention and disposal independent of the main log repository. This architecture might be used where all data sources identified for ingest are determined to be relevant for SIEM security requirements and the feeds' utility is assisting in data volume management. Alternatively, this architecture might be useful for organisations with pre-existing SIEMs that cannot, for business reasons, easily redesign to the data lake-first approach.

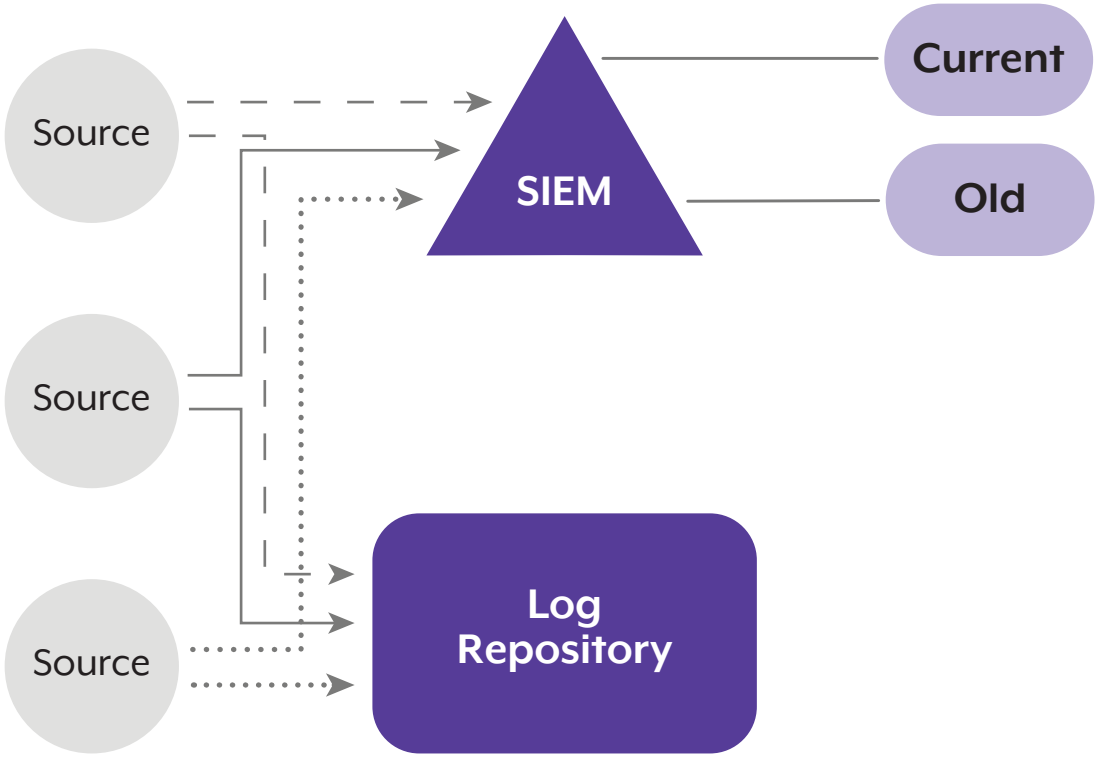


Figure 2. Dual log replication between log repository and SIEM visualisation

3. Independent log feeds based on requirements

In this example, the rationale and architecture are similar to that previous examplely mentioned, except that some data sources are configured to direct specific logs just for the log repository, while others continue to replicate logs for both the log repository and SIEM. This architecture may be useful for organisations with onerous governance or log archiving requirements that need to store logs that may not contain relevant security information.

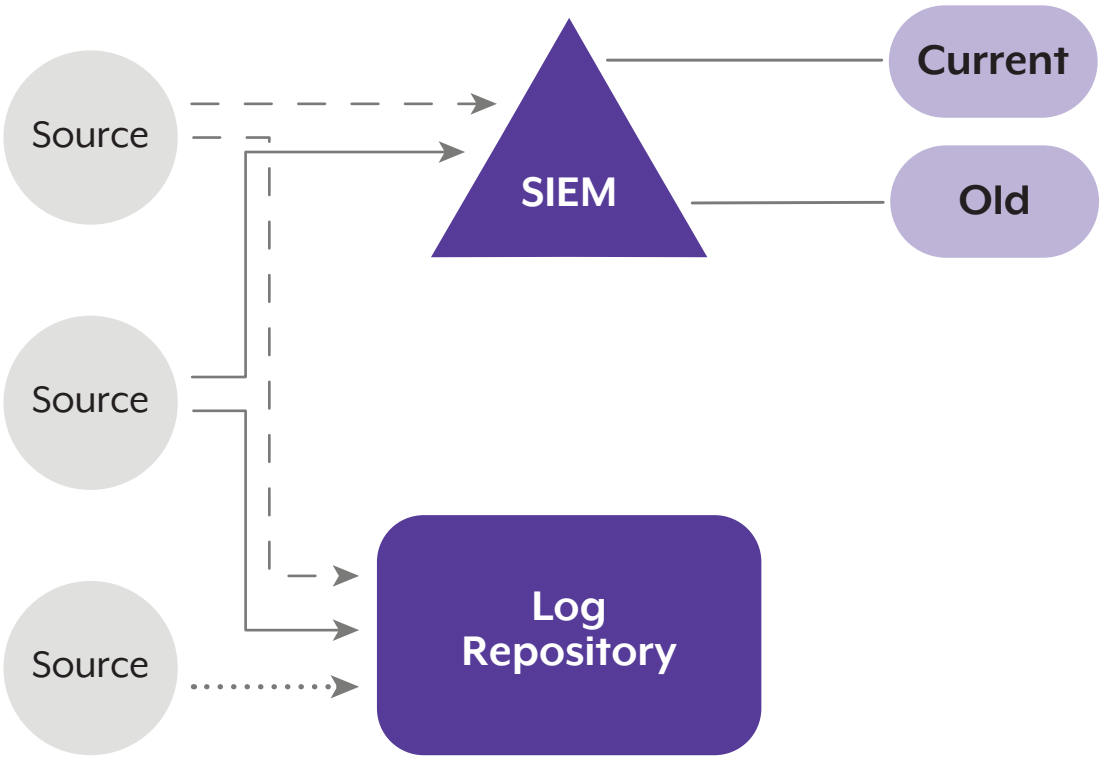


Figure 3. Independent log feeds based on requirements visualisation

4. SIEM-first architecture

The following SIEM-first architecture is sometimes used, but the authoring agencies do not recommend it. In a SIEM-first approach, logs are first collected and centralised by the SIEM, then replicated to a log repository. This approach can be inefficient from a processing performance perspective. Additionally, SIEM processing risks the logs' integrity if issues occur. This may be problematic for incident investigations, as the log repository will only contain the processed log, not the original.

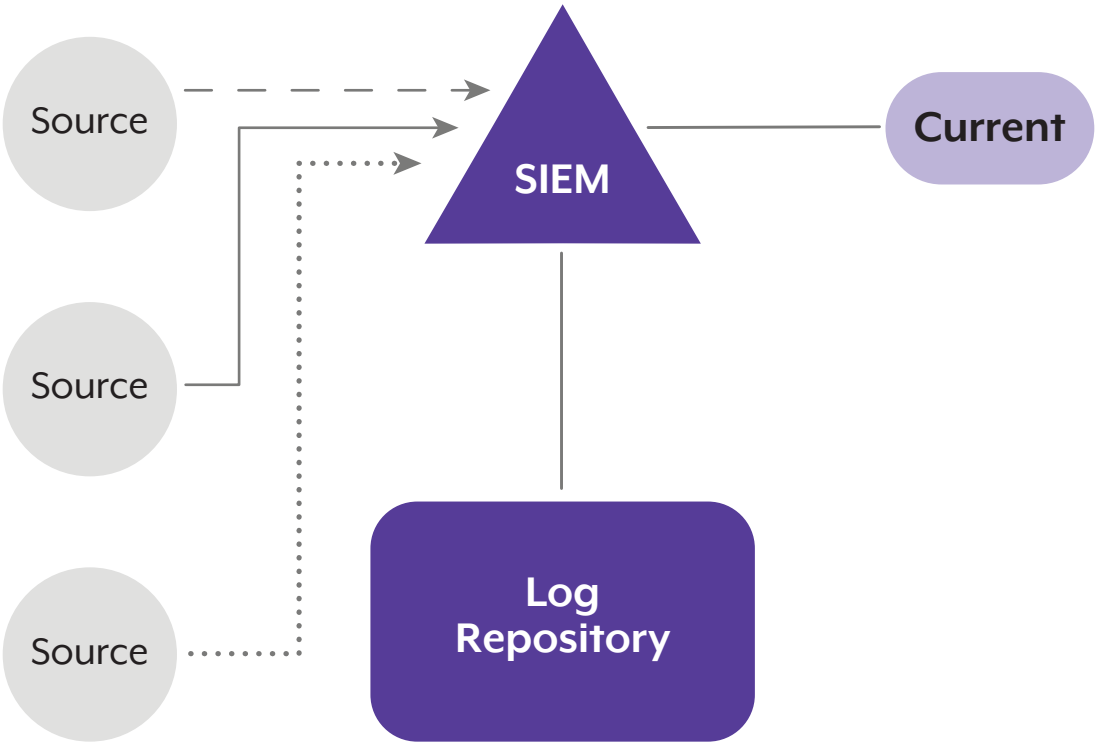


Figure 4. SIEM-first architecture

5. Protecting personally identifiable information

For organisations with specific requirements around protecting the personally identifiable information (PII) of their employees, the following diagram illustrates how the SIEM can be designed to restrict the access of SIEM users. In addition to the application of standard Role Based Access Controls (RBAC), the SIEM's data can be configured and indexed in a segmented manner to restrict SIEM user access to the data sources. In the diagram below, User A has only restricted access, while User B has full access.

Some SIEM platforms natively enable this sort of architecture, allowing log sources to be ingested into specific subsets of data in separate locations within the SIEM platform, which can be further configured with a variety of restrictive access control mechanisms to limit user access.

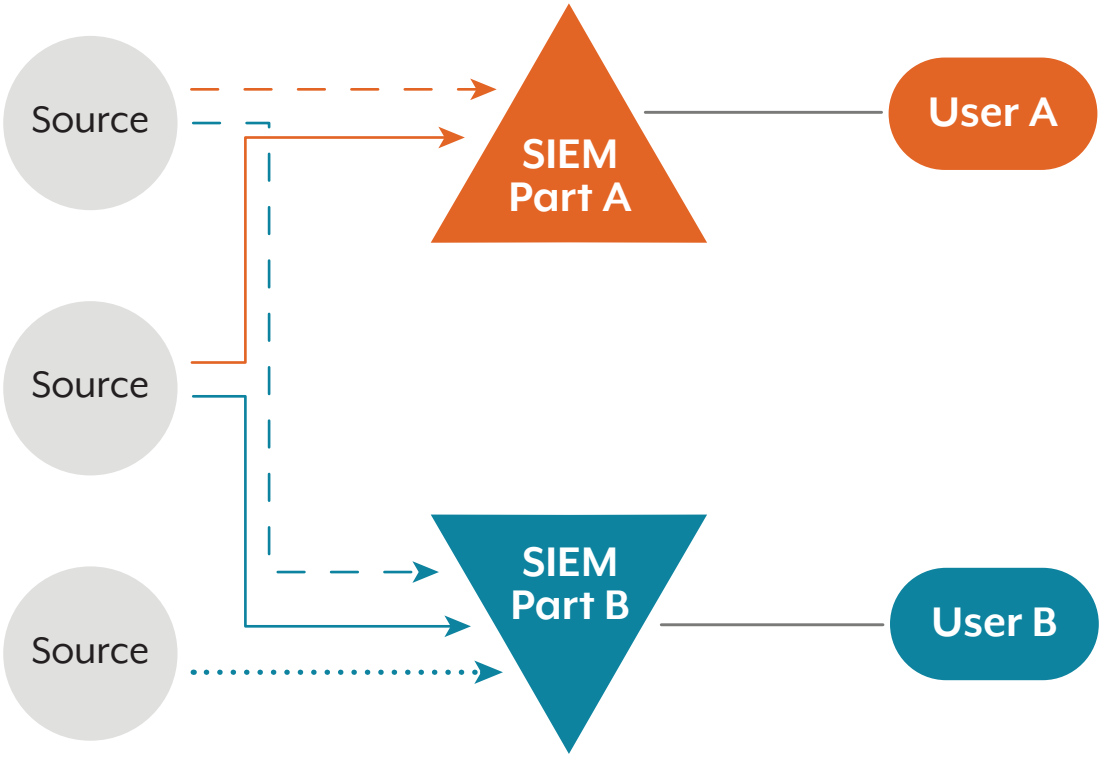


Figure 5. Protecting Personal Identifying Information (PII)

ANNEX B: Pre-processing methods

In circumstances where the SIEM is not drawing from a centralised logging solution but directly from sources, pre-processing reduces the amount of data forwarded on to the next stage of log collection and centralisation. Pre-processing can strip out unwanted information to reduce the burden on forwarders, optimise storage, achieve standardisation, and/or improve SIEM analysis. Several methods for pre-processing are outlined below.

1. Source log separation

In this method, the data source (or host device) is configured to generate and separate its logs into different types according to certain features. Only certain types of logs – in the diagram to the right, Log 1 – are then forwarded onwards for processing using a 'replication' method (such as a forwarder) to the SIEM, in addition to, or via, the log repository.

'Unwanted' logs (such as Logs 2 and 3) remain on the source for the length of time specified in log-retention policies.

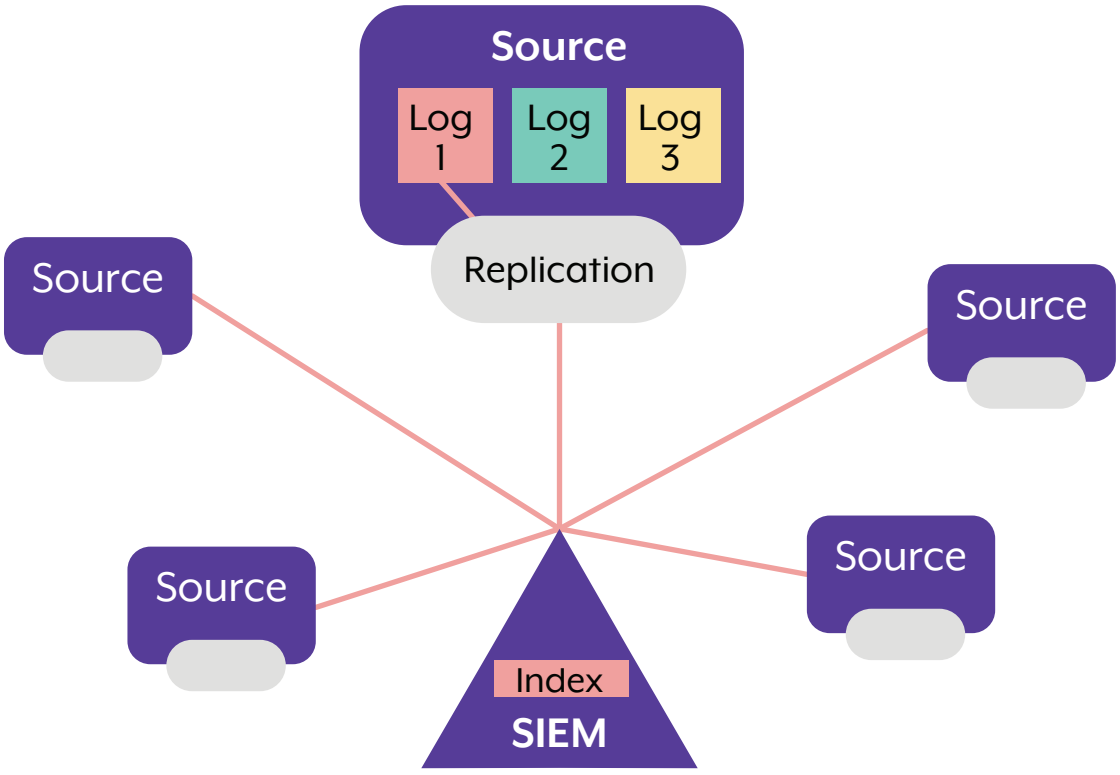


Figure 6. Separation of logs at the source

2. Replication point pre-processing

In this example, the data source (or host device) is configured to generate and separate its logs into different types according to certain features. Unlike the previous example, all three log types interact with the forwarder or replication point. Log 1 and Log 2 are pointed toward the SIEM and/or log repository.

In addition to forwarding log type 1 and 2, the replication/forwarder is used to retrieve other log sources (such as log type 3) to replicate them to a secondary location different from the SIEM. Log type 3 entries are replicated to a dedicated storage area or data lake.

This pre-processing method is commonly used for PowerShell logs, which are an important source of information for investigations, but can present an ingestion problem for SIEMs due to their size and format.

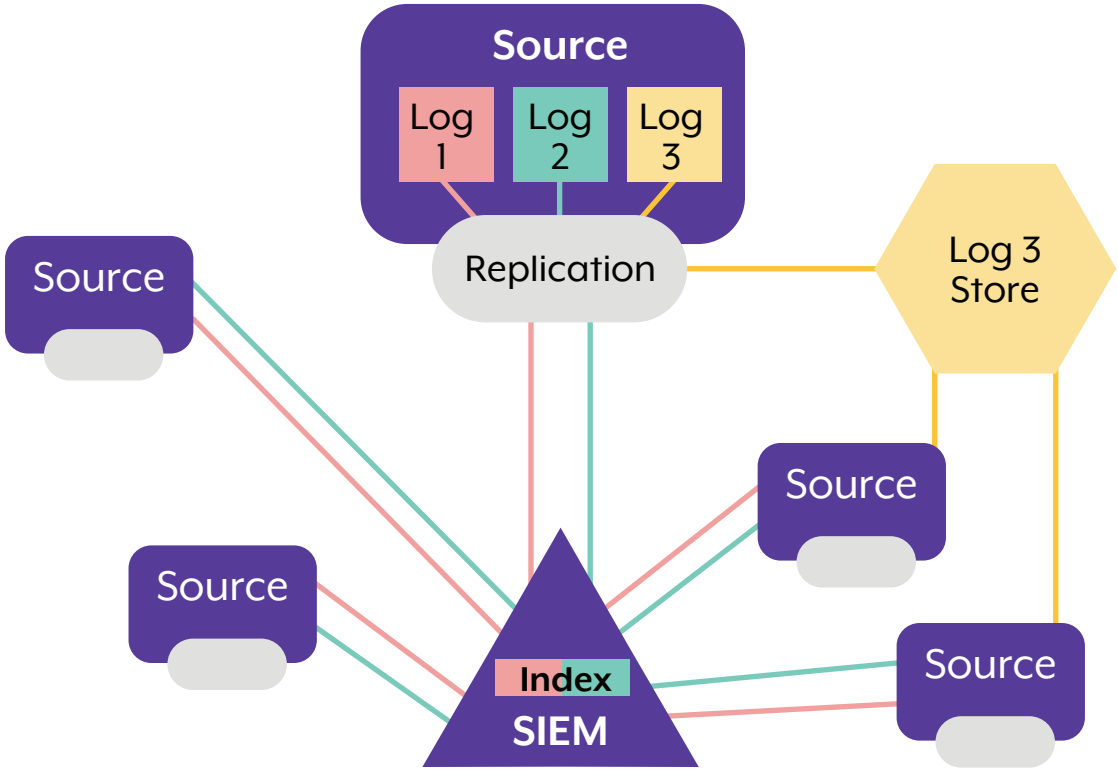


Figure 7. Pre-processing by the replication mechanism

3. SIEM ingestion

In this example, all three log sources (Log 1, Log 2 and Log 3) are replicated from the data source to the SIEM to be processed. Unlike the previous examples, where unwanted log sources are left on the source or sent to another repository, all logs are replicated to the SIEM.

The SIEM may be configured to alert only to one of these log types (e.g. log type 2), while the others (e.g. log types 1 and 3) are instead held in a storage mechanism called 'cold storage'. If a SIEM query needs further log information from the other log types, the SIEM retrieves files or events from cold storage back into the SIEM for processing.

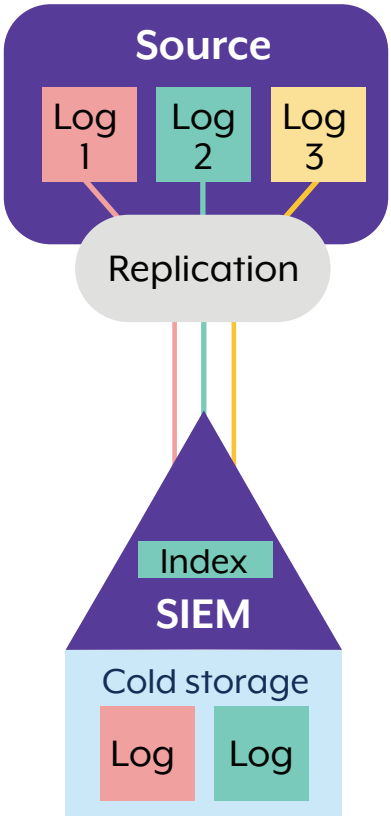


Figure 8. At SIEM separation

Disclaimer

The material in this guide is of a general nature and should not be regarded as legal advice or relied on for assistance in any particular circumstance or emergency situation. In any important matter, you should seek appropriate independent professional advice in relation to your own circumstances.

The Commonwealth accepts no responsibility or liability for any damage, loss or expense incurred as a result of the reliance on information contained in this guide.

For more information, or to report a cyber security incident, contact us:

cyber.gov.au | 1300 CYBER1 (1300 292 371)

ASD AUSTRALIAN SIGNALS DIRECTORATE

ACSC Australian **Cyber Security** Centre