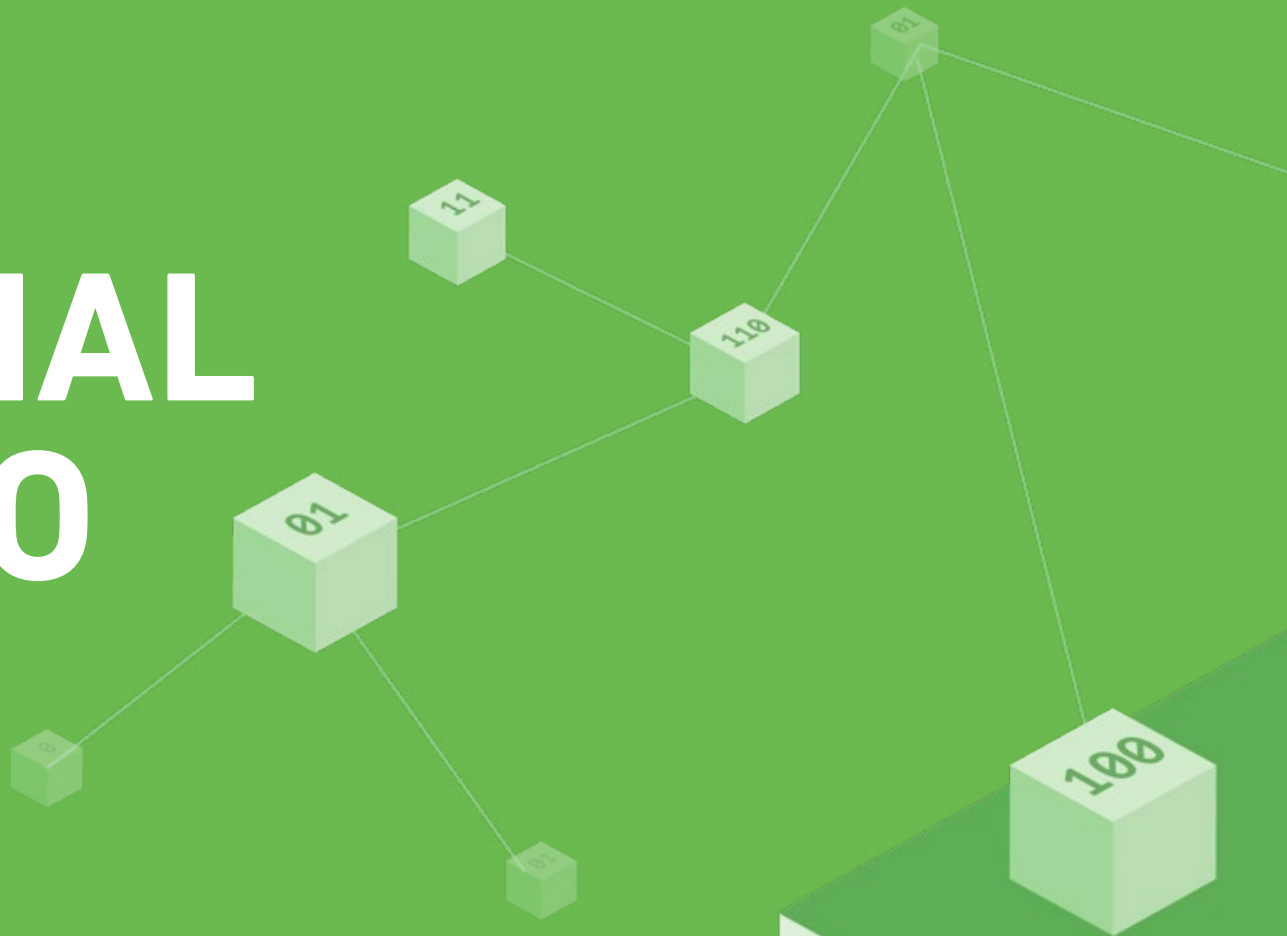


The ESSENTIAL GUIDE TO SIEM



Learn Security Information and Event Management

The SIEM is a foundational technology of the security operations center (SOC). SIEMs have been around for decades, but a new generation is emerging with new capabilities like data science-driven anomaly detection and incident response automation.

Learn everything about SIEMs, past, present and future—architecture, what's under the hood, and using SIEMs in the field to detect incidents and defend organizations.



Contents

CH01

What is SIEM?

Components, best practices, and next-gen capabilities

CH02

SIEM Architecture

How SIEMs are built, how they generate insights, and how they are changing

CH03

Events and Logs

SIEM under the hood—the anatomy of security events and system logs

CH04

UEBA

User and entity behavior analytics detects threats other tools can't see

CH05

SIEM Use Cases

Beyond alerting and compliance—SIEMs for insider threats, threat hunting and IoT

CH06

SIEM Analytics

From correlation rules and attack signatures to automated detection via machine learning

CH07

Incident Response and Automation

Security orchestration, automation and response (SOAR)—the future of incident response

CH08

The SOC, SecOps and SIEM

A comprehensive guide to the modern SOC—SecOps and next-gen tech

CH09

Evaluating and Selecting SIEM Tools - A Buyer's Guide

Evaluation criteria, build vs. buy, cost considerations and compliance

CH10

SIEM Essentials Quiz

SIEM Essentials Quiz

What is SIEM?

Security information and event management (SIEM) solutions use rules and statistical correlations to turn log entries, and events from security systems, into actionable information. This information can help security teams detect threats in real time, manage incident response, perform forensic investigation on past security incidents, and prepare audits for compliance purposes.

The term SIEM was coined in 2005 by Mark Nicolett and Amrit Williams, in Gartner's SIEM report, [Improve IT Security with Vulnerability Management](#). They proposed a new security information system, on the basis of two previous generations.

- **Security information management (SIM)** – a first generation, built on top of traditional log collection and management systems. SIM introduced long-term storage, analysis, and reporting on log data, and combined logs with threat intelligence.
- **Security event management (SEM)** – a second generation, addressing security events—aggregation, correlation and notification for events from security systems such as antivirus, firewalls and intrusion detection systems (IDS), as well as events reported directly by authentication, SNMP traps, servers, databases etc.

In the years that followed, vendors introduced systems that provided both security log management and analysis (SIM) and event management (SEM), to create SIEM solutions.

SIEM platforms can aggregate both historical log data and real-time events, and establish relationships that can help security staff identify anomalies, vulnerabilities and incidents.

The main focus is on security-related incidents and events, such as succeeded or failed logins, malware activities or escalation of privileges.

These insights can be sent as notifications or alerts, or discovered by security analysts using the SIEM platform's visualization and dashboarding tools.



Next-gen SIEM

SIEM is a mature technology, and the next generation of SIEMs provide new capabilities:

- **User and entity behavior analytics (UEBA)** – advanced SIEMs go beyond rules and correlations, leveraging AI and deep learning techniques to look at patterns of human behavior. This can help detect insider threats, targeted attacks, and fraud.
- **Security orchestration, automation and response (SOAR)** – next-gen SIEMs integrate with enterprise systems and automate incident response. For example, the SIEM might detect an alert for ransomware and perform containment steps automatically on affected systems, before the attacker can encrypt the data.

What Can a SIEM Help With?

Components and Capabilities



Data aggregation

Aggregates data from network, security, servers, databases, applications, and other security systems like firewalls, anti virus and intrusion detection systems (IDS)



Threat intelligence feeds

Combines internal data with threat intelligence feeds containing data on vulnerabilities, threat actors and attack patterns



Correlation

Links events and related data into meaningful bundles which represent a real security incident, threat, vulnerability or forensic finding



Analytics

Uses statistical models and machine learning to identify deeper relationships between data elements, and anomalies compared to known trends, and tie them to security concerns



Alerting

Analyzes events and sends out alerts to notify security staff of immediate issues, either by email, other types of messaging, or via security dashboards



Dashboards and visualizations

Creates visualizations to allow staff to review event data, see patterns and identify activity that does not conform to standard patterns



Compliance

Automates the gathering of compliance data, producing reports that adapt to security, governance and auditing processes for standards like HIPAA, PCI/DSS, HITECH, SOX and GDPR



Retention

Stores long-term historical data to enable analysis, tracking, and data for compliance requirements. Especially important in forensic investigations, which happen after the fact



Threat hunting

Allows security staff to run queries on SIEM data, filter and pivot the data, to proactively uncover threats or vulnerabilities



Incident response

Provides case management, collaboration and knowledge sharing around security incidents, allowing security teams to quickly synchronize on the essential data and respond to a threat



SOC automation

Integrates with other security solutions using APIs, and lets security staff define automated playbooks and workflows that should be executed in response to specific incidents

How SIEM Works

Present and Future

In the past, SIEMs required meticulous management at every stage of the data pipeline—data ingestion, policies, reviewing alerts and analyzing anomalies. Increasingly, SIEMs are getting smarter at pulling data together, from ever more organizational sources, and using AI techniques to understand what type of behavior constitutes a security incident.



01

Data collection

Most SIEM systems collect data by deploying collection agents on end-user devices, servers, network equipment, or other security systems like firewalls and antivirus, or via protocols syslog forwarding, SNMP or WMI. Advanced SIEMs can integrate with cloud services to obtain log data about cloud-deployed infrastructure or SaaS applications, and can easily ingest other non-standard data sources.

Pre-processing may happen at edge collectors, with only some of the events and event data passed to centralized storage.



02

Data storage

Traditionally, SIEMs relied on storage deployed in the data center, which made it difficult to store and manage large data volumes.

As a result, only some log data was retained. Next-generation SIEMs are built on top of modern data lake technology such as Amazon S3 or Hadoop, allowing nearly unlimited scalability of storage at low cost. This makes it possible to retain and analyze 100% of log data across even more platforms and systems.



03

Policies and rules

The SIEM allows security staff to define profiles, specifying how enterprise systems behave under normal conditions.

They can then set rules and thresholds to define what type of anomaly is considered a security incident. Increasingly, SIEMs leverage machine learning and automated behavioral profiling to automatically detect anomalies, and autonomously define rules on the data, to discover security events that require investigation.



04

Data consolidation and correlation

The central purpose of a SIEM is to pull together all the data and allow correlation of logs and events across all organizational systems.

An error message on a server can be correlated with a connection blocked on a firewall, and a wrong password attempted on an enterprise portal. Multiple data points are combined into meaningful security events, and delivered to analysts by notifications or dashboards. Next-gen SIEMs are getting better and better at learning what is a “real” security event that warrants attention.

What are SIEMs Used For



01

Security monitoring

SIEMs help with real-time monitoring of organizational systems for security incidents.

A SIEM has a unique perspective on security incidents, because it has access to multiple data sources – for example, it can combine alerts from an IDS with information from an antivirus product. It helps security teams identify security incidents that no individual security tool can see, and help them focus on alerts from security tools that have special significance.



02

Advanced threat detection

SIEMs can help detect, mitigate and prevent advanced threats, including:

- **Malicious insiders** – a SIEM can use browser forensics, network data, authentication and other data to identify insiders planning or carrying out an attack
- **Data exfiltration (sensitive data illicitly transferred outside the organization)** – a SIEM can pick up data transfers that are abnormal in their size, frequency or payload
- **Outside entities, including advanced persistent threats (APTs)** – a SIEM can detect early warning signals indicating that an outside entity is carrying out a focused attack or long-term campaign against the organization



03

Forensics and incident response

SIEMs can help security analysts realize that a security incident is taking place, triage the event and define immediate steps for remediation.

Even if an incident is known to security staff, it takes time to collect data to fully understand the attack and stop it – SIEM can automatically collect this data and significantly reduce response time. When security staff discover a historic breach or security incident that needs to be investigated, SIEMs provide rich forensic data to help uncover the kill chain, threat actors and mitigation.



04

Compliance reporting and auditing

SIEMs can help organizations prove to auditors and regulators that they have the proper safeguards in place and that security incidents are known and contained.

Many early adopters of SIEMs used it for this purpose – aggregating log data from across the organization and presenting it in audit-ready format. Modern SIEMs automatically provide the monitoring and reporting necessary to meet standards like HIPAA, PCI/DSS, SOX, FERPA and HITECH.

SIEM Best Practices

The [Infosec Institute](#) suggests 10 best practices for successful implementation of a SIEM platform.

Defining SIEM requirements:

- **Define requirements** for monitoring, reporting and auditing, consulting all relevant stakeholders before deploying a SIEM.
- **Determine the scope of the SIEM** – which parts of the infrastructure it will cover, necessary credentials, and log verbosity.
- **Define audit data** accessibility, retention, how to achieve data integrity, evidentiary rules, and disposal for historical or private data.

Ensure you leverage the SIEM to monitor and report on all of the following:

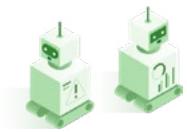
- **Access monitoring** – transgression and anomalous access to key resources
- **Perimeter defenses** – status of perimeter defenses, possible attacks and risky configuration changes
- **Resource integrity** – critical network resources – status, backups, change management, threats and vulnerabilities
- **Intrusion detection** – incidents reported by intrusion detection, or correlated/inferred using SIEM data
- **Malware defense** – violations, threats, or activity regarding malware controls
- **Application defenses** – status, configuration changes, violations and anomalies for web servers, databases and other web app resources
- **Acceptable use** – status, issues and violations regarding acceptable, mandated or metered use of system resources

SIEM Evolution

2005

2010

2017



GENERATION I Early SIEM

The first SIEMs combined security information management (SIM) and security event management (SEM). They were limited in scale of data managed and supported alerting/visualizations.

SCALABILITY

Scales vertically

HISTORIC DATA

Partial

DATA COLLECTION

Slow manual ingestion of log data

THREAT DETECTION

Manual analysis and alerts based on manual rules

INCIDENT RESPONSE

Little or no interface with downstream systems

DASHBOARDS AND VISUALIZATIONS

Very limited



GENERATION II Big Data SIEM

An integrated SIEM based on big data infrastructure, managing and correlating historical log data, real-time events and threat intelligence in one place—providing a holistic view of enterprise security data.

SCALABILITY

Scales horizontally, supporting big data

HISTORIC DATA

Full, with some filtering

DATA COLLECTION

Automated ingestion, data sources limited

THREAT DETECTION

Manual analysis, alerts and dashboards

INCIDENT RESPONSE

Limited interface with downstream systems

DASHBOARDS AND VISUALIZATIONS

Typically limited set of pre-built visualizations



GENERATION III Automation and Machine Learning

Early SIEMs had limited ability to proactively warn about and react to complex security events. New SIEMs perform automated behavioral profiling (UEBA), and can automatically interact with IT and security systems to mitigate incidents (SOAR).

SCALABILITY

Based on data lake, unlimited scale

HISTORIC DATA

Unlimited historic retention including new data sources like the cloud

DATA COLLECTION

Automated ingestion of any data source

THREAT DETECTION

Automated, based on machine learning and behavioral profiling

INCIDENT RESPONSE

Integrates with IT and security tools, full SOAR capabilities

DASHBOARDS AND VISUALIZATIONS

Full business intelligence (BI) data exploration

Next-Generation SIEMs

The Future is Here

New SIEM platforms provide advanced capabilities such as:

- **Complex threat identification** – correlation rules can't capture many complex attacks, because they lack context, or can't respond to new types of incidents. With automatic behavioral profiling, SIEMs can detect behavior that suggests a threat.
- **Entity behavior analysis** – critical assets on the network such as servers, medical equipment or machinery have unique behavioral patterns. SIEMs can learn these patterns and automatically discover anomalies that suggest a threat.
- **Automated incident response** – once a SIEM detects a certain type of security event, it can execute a pre-planned sequence of actions to contain and mitigate the incident. SIEMs are becoming full SOAR tools.
- **Lateral movement** – attackers move through a network by using IP addresses, credentials and machines, in search of key assets. By analyzing data from across the network and multiple system resources, SIEMs can detect this lateral movement.
- **Detection without rules or signatures** – many threats facing your network can't be captured with manually-defined rules or known attack signatures. SIEMs can use machine learning to detect incidents without pre-existing definitions.

An example of a next-generation SIEM is the [Exabeam Security Management Platform \(SMP\)](https://exabeam.com/product), which combines behavioral analytics based on machine learning, cloud connectors, a flexible data lake infrastructure, incident response and threat hunting capabilities. Learn more at exabeam.com/product

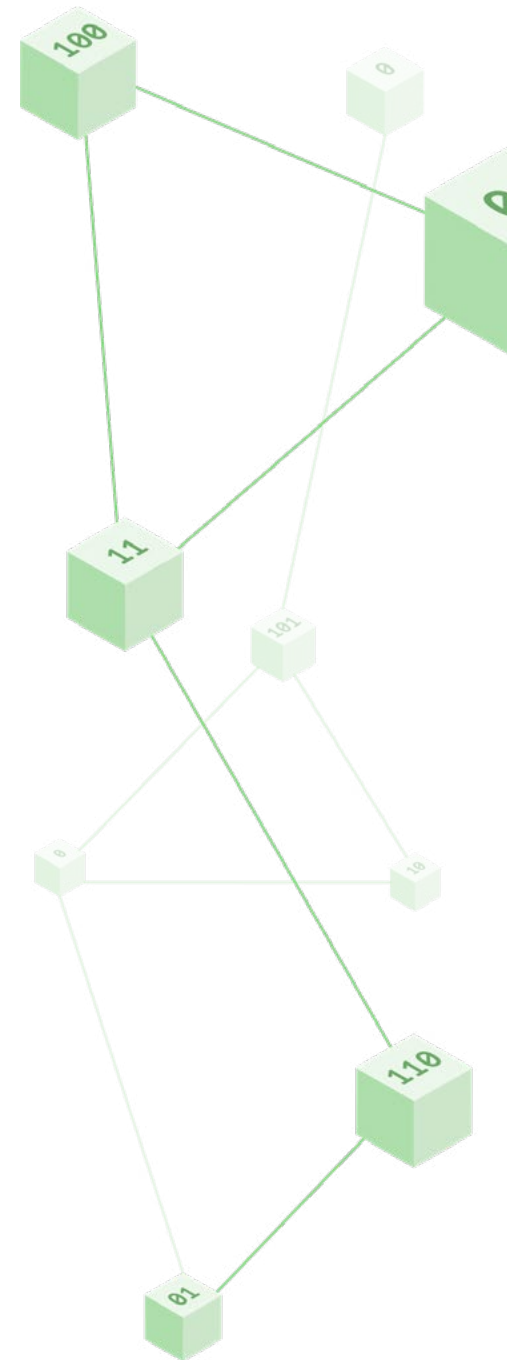
SIEM Architecture: Technology, Process and Data

In this chapter of the Essential Guide to SIEM, we explain how SIEM systems are built, how they go from raw event data to security insights, and how they manage event data on a huge scale. We cover both traditional SIEM platforms and in newer SIEM architecture based on data lake technology.

Security information and event management (SIEM) platforms collect log and event data from security systems, networks and computers, and turn it into actionable security insights. SIEM technology can help organizations detect threats that individual security systems cannot see, investigate past security incidents, perform incident response and prepare reports for regulation and compliance purposes.

In this chapter you will learn:

- **The log management process** – data collection, data management and historic log retention
- **The log flow** – from millions of events to a handful of meaningful alerts
- **SIEM log sources** – security systems, network devices, cloud systems and more
- **SIEM hosting models** – self-hosted self-managed, cloud-hosted, self-managed, hybrid-managed, and fully-managed
- **SIEM sizing** – event velocity, calculating events per second (EPS) and total event volume, hardware requirements and deployment options, including data lake
- **SIEM outputs** – reporting, dashboards, and visualizations and advanced analytics



12 Components and Capabilities in a SIEM Architecture

01

Data aggregation

Collects and aggregates data from security systems and network devices

02

Threat intelligence feeds

Combines internal data with third-party data on threats and vulnerabilities

03

Correlation and security monitoring

Links events and related data into security incidents, threats or forensic findings

04

Analytics

uses statistical models and machine learning to identify deeper relationships between data elements

05

Alerting

Analyzes events and sends alerts to notify security staff of immediate issues

06

Dashboards

Creates visualizations to let staff review event data, identify patterns and anomalies

07

Compliance

Gathers log data for standards like HIPAA, PCI/DSS, HITECH, SOX and GDPR and generates reports

08

Retention

Stores long-term historical data, useful for compliance and forensic investigations

09

Forensic analysis

Enables exploration of log and event data to discover details of a security incident

10

Threat hunting

Enables security staff to run queries on log and event data to proactively uncover threats

11

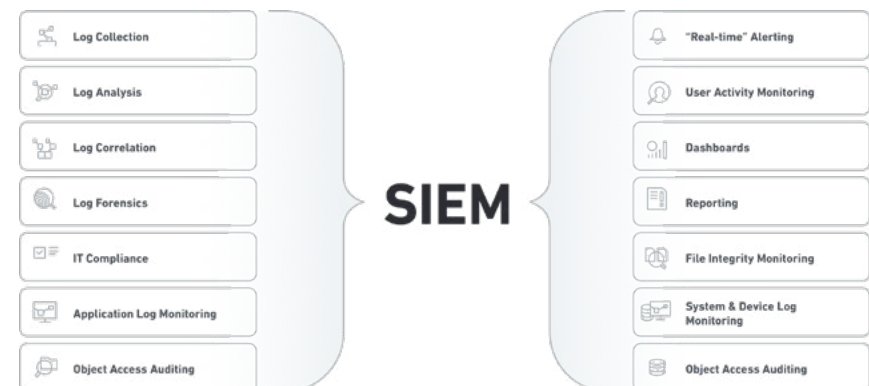
Incident response

Helps security teams identify and respond to security incidents, bringing in all relevant data rapidly

12

SOC automation

Advanced SIEMs can automatically respond to incidents but orchestrating security systems, known as security orchestration, automation and response (SOAR)



The Log Management Process

A SIEM server, at its root, is a log management platform. Log management involves collecting the data, managing it to enable analysis, and retaining historical data.

Data Collection

SIEMs collect logs and events from hundreds of organizational systems (for a partial list, see Log Sources below). Each device generates an event every time something happens, and collects the events into a flat log file or database. The SIEM can collect data in four ways:

01. Via an agent installed on the device (the most common method)
02. By directly connecting to the device using a network protocol or API call
03. By accessing log files directly from storage, typically in Syslog format
04. Via an event streaming protocol like SNMP, Netflow or IPFIX

The SIEM is tasked with collecting data from the devices, standardizing it and saving it in a format that enables analysis.

Next-gen SIEM

Next-generation SIEMs come pre-integrated with common cloud systems and data sources, allowing you to pull log data directly. Many managed cloud services and SaaS applications do not allow you to install traditional SIEM collectors, making direct integration between SIEM and cloud systems critical for visibility.

Data Management

SIEMs, especially at large organizations, can store mind-boggling amounts of data. The data needs to be:

- **Stored** – either on-premises, in the cloud or both
- **Optimized and indexed** – to enable efficient analysis and exploration
- **Tiered** – hot data necessary for live security monitoring should be on high performance storage, whereas cold data, which you may one day want to investigate, should be relegated to high-volume inexpensive storage mediums

Next-gen SIEM

Next-generation SIEMs are increasingly based on modern data lake technology such as Amazon S3, Hadoop or Elasticsearch, enabling practically unlimited data storage at low cost.

Log Retention

Industry standards like PCI DSS, HIPAA and SOX require that logs be retained for between one and seven years. Large enterprises create a very high volume of logs every day from IT systems (see SIEM Sizing below). SIEMs need to be smart about which logs they retain for compliance and forensic requirements. SIEMs use the following strategies to reduce log volumes:

- **Syslog servers** – Syslog is a standard which normalizes logs, retaining only essential information in a standardized format. Syslog lets you compress logs and retain large quantities of historical data.
- **Deletion schedules** – SIEMs automatically purge old logs that are no longer needed for compliance. By accessing log files directly from storage, typically in Syslog format.
- **Log filtering** – not all logs are really needed for the compliance requirements faced by your organization, or for forensic purposes. Logs can be filtered by source system, times, or by other rules defined by the SIEM administrator.
- **Summarization** – log data can be summarized to maintain only important data elements such as the count of events, unique IPs, etc.

Next-gen SIEM

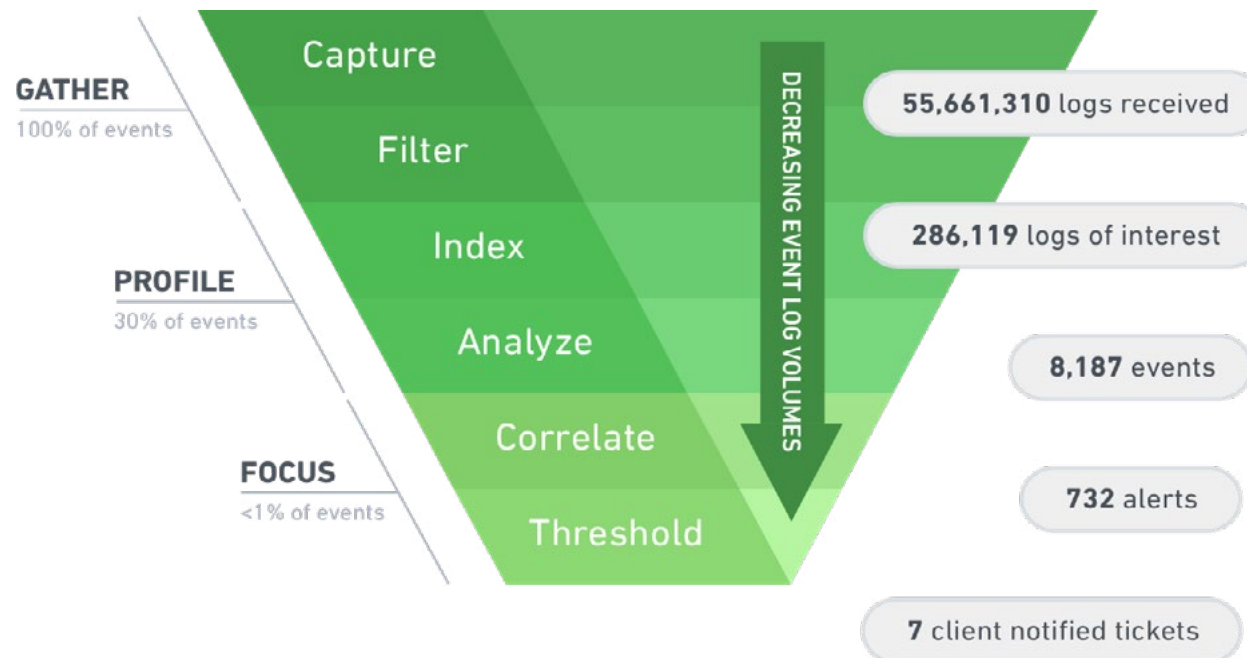
Historic logs are not only useful for compliance and forensics. They can also be used for deep behavioral analysis. Next-generation SIEMs provide user and entity behavior analytics (UEBA) technology, which uses machine learning and behavioral profiling to intelligently identify anomalies or trends, even if they weren't captured in the rules or statistical correlations of the traditional SIEMs.

Next-generation SIEMs leverage low-cost distributed storage, allowing organizations to retain full source data. This enables deep behavioral analysis of historic data, to catch a broader range of anomalies and security issues.

The Log Flow

A SIEM captures 100% of log data from across your organization. But then data starts to flow down the log funnel, and hundreds of millions of log entries can be whittled down to only a handful of actionable security alerts.

SIEMs filter out noise in logs to keep pertinent data only. Then they index and optimize the relevant data to enable analysis. Finally, around 1% of data, which is the most relevant for your security posture, is correlated and analyzed in more depth. Of those correlations, the ones which exceed security thresholds become security alerts.



SIEM Logging Sources

Which organizational systems feed their logs to the SIEM?
And which other business data is of interest to a SIEM?

Next-gen SIEM

Until recently SIEMs couldn't access log and event data from cloud infrastructure like AWS or Microsoft Azure, or SaaS applications like Salesforce and Google Apps. This created a huge blind spot in security monitoring. Some next-generation solutions come with pre-built connectors and SIEM integrations with modern cloud technology.

Security Events



- Intrusion detection systems
- Endpoint security (antivirus, anti-malware)
- Data loss prevention
- VPN concentrators
- Web filters
- Honeypots
- Firewalls

Network Logs



- Routers
- Switches
- DNS servers
- Wireless access points
- WAN
- Data transfers
- Private cloud networks

Applications and Devices



- Application servers
- Databases
- Intranet applications
- Web applications
- SaaS applications
- Cloud-hosted servers
- End-user laptops or desktops
- Mobile devices

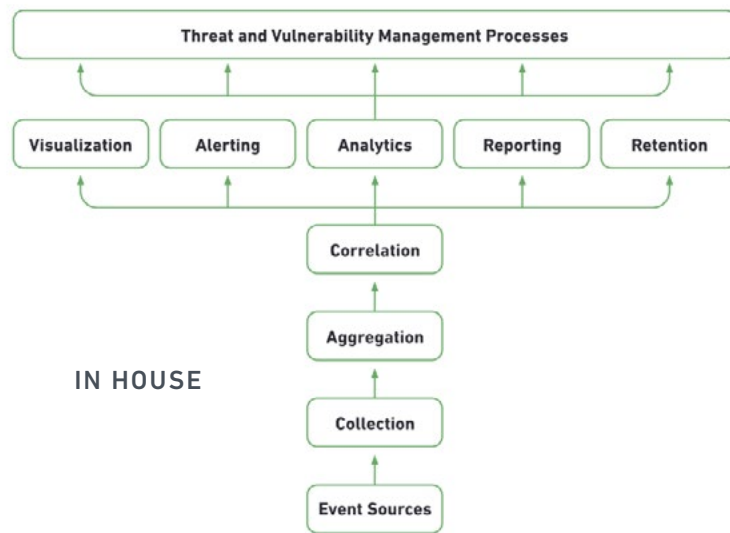
IT Infrastructure



- Configuration
- Locations
- Owners
- Network maps
- Vulnerability reports
- Software inventory

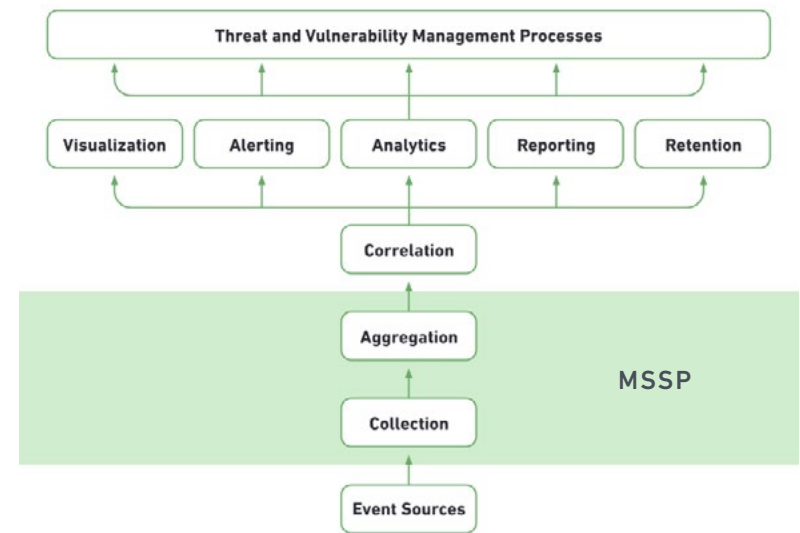
SIEM Deployment Models

There are many deployment options to consider for SIEM. Here, we look at four common ones.



Traditional SIEM

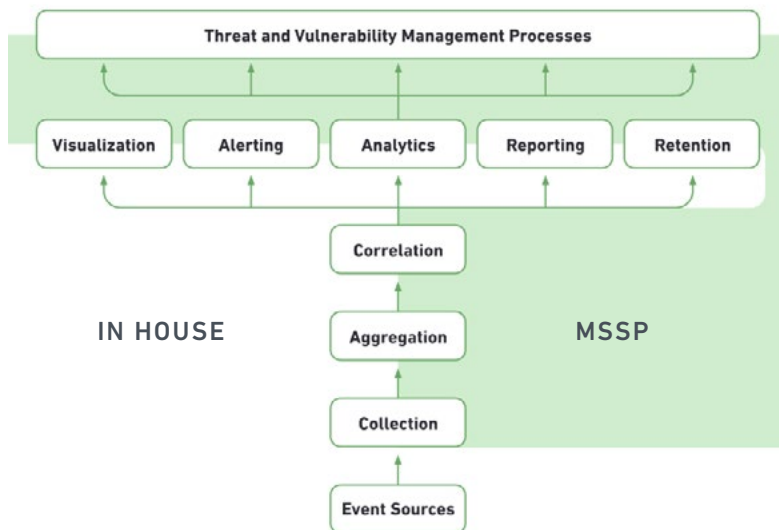
This is the traditional SIEM deployment model—host the SIEM in your data center, often with a dedicated SIEM appliance, maintain storage systems, and manage it with trained security personnel. This model made SIEM a notoriously complex and expensive infrastructure to maintain.



Cloud SIEM, Self-Managed

You handle: Correlation, analysis, alerting and dashboards, security processes leveraging SIEM data.

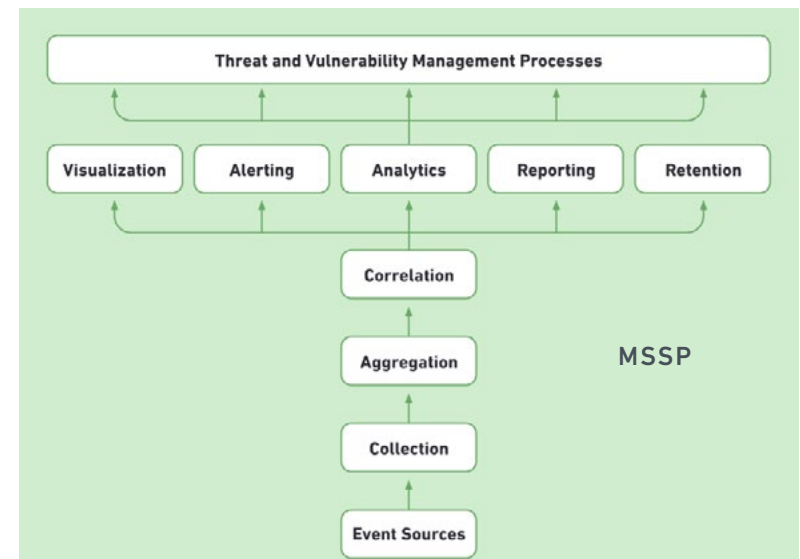
MSSP handles: Receiving events from organizational systems, collection and aggregation.



Self-Hosted, Hybrid-Managed

You handle: Purchasing software and hardware infrastructure.

MSSP together with your security staff: Deploying SIEM event collection / aggregation, correlation, analysis, alerting and dashboards.



SIEM as a Service

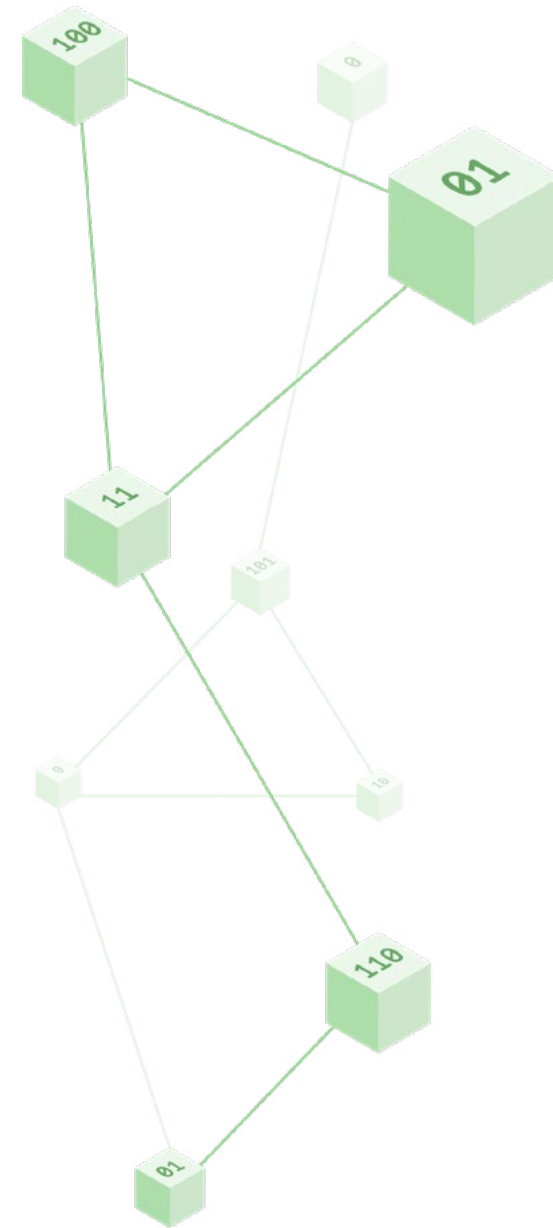
You handle: Defining program goals.

MSSP handles: Event collection, aggregation, correlation, analysis, alerting and dashboards.

Which Hosting Model is Right for You?

The following considerations can help you select a SIEM deployment model:

- **Do you have an existing SIEM infrastructure?** If you've already purchased the hardware and software, opt for self-hosted self-managed, or leverage an MSSP's expertise to jointly manage the SIEM with your local team.
- **Do you have security staff with SIEM expertise?** The human factor is crucial in getting true value from a SIEM. If you don't have trained security staff, rent the analysis services via a hybrid-managed or SIEM as a Service model.
- **Are you able to move data off-premises?** If so, a cloud-hosted or fully managed model can reduce costs and management overhead.



SIEM Sizing: Velocity, Volume and Hardware Requirements

A majority of SIEMs today are deployed on-premises. This requires organizations to carefully consider the size of log and event data they are generating, and the system resources required to manage it.

Calculating Velocity: Events Per Second (EPS)

A common measure of velocity is EPS, defined as:

$$\frac{\text{\# OF SECURITY EVENTS}}{\text{TIME PERIOD IN SECONDS}} = \text{EPS}$$

EPS can vary between normal and peak times. For example, a Cisco router might generate 0.6 EPS on average, but during peak times, such as during an attack, it can generate as many as 154 EPS.

According to the *SIEM Benchmarking Guide* by the [SANS Institute](#), organizations should strike a balance between normal and peak EPS measurements. It's not practical, or necessary, to build a SIEM to handle peak EPS for all network devices, because it's unlikely all devices will hit their peak at once. On the other hand, you must plan for crisis situations, in which the SIEM will be most needed.

A Simple Model for Predicting EPS During Normal and Peak Times

01. Measure **Normal EPS** and **Peak EPS**, by looking at 90 days of data for the target system
02. Estimate the **Number of Peaks per Day**
03. Estimate the **Duration in Seconds of a Peak**, and by extension, **Total Peak Seconds per Day**
04. Calculate **Total Peak Events per Day** = (Total Peak Seconds per Day) * Peak EPS
05. Calculate Total Normal Events per Day = (Total Seconds – **Total Peak Seconds per Day**) * Normal EPS

The sum of these two numbers is the total estimated velocity.

In addition, the SANS guide recommends adding:

- 10% for headroom
- 10% for growth

So that the final number of events per day will be:

$$(\text{Total Peak Events per Day} + \text{Total Normal Events per Day}) \\ * 110\% \text{ headroom} * 110\% \text{ growth}$$

Calculating Velocity: Events Per Second (EPS)

The following table, provided by SANS, shows typical average EPS (normal EPS) and peak EPS for selected network devices. The data is several years old but can provide ballpark figures for your initial estimates.

Table 1: Baseline Network Device EPS Averages

Qty	Type	Description	Avg EPS	Total Peak EPS	Average Peak EPS
750	Employees/Endpoints (Windows XP)	Desktops & laptops at 5 locations	Included at domain servers	Included at domain servers	Included at domain servers
7	Cisco Catalyst Switches	One at each location, one in DMZ and one in the Trusted network	5.09	51.88	26.35
7	Cisco Gateway/Routers	One at each location	0.60	380.50	154.20
5	Windows 2003 Domain Servers	One at each location	40.00	404.38	121.75
3	Windows 2003 Application Servers	In high availability cluster at data center	1.38	460.14	230.07
3	MS SQL Database Servers running on Windows 2003 Server	High availability cluster at data center	1.83	654.90	327.45
6	Microsoft Exchange Servers	One at each location with two (cluster) at the data center	3.24	1,121.50	448.60
3	MS IIS Web Servers on Windows 2003	High availability cluster at data center	1.17	2,235.10	1,117.55
2	Windows DNS Servers	At data center – failover	0.72	110.80	110.80
2	Linux Legacy Application Servers	At data center	0.12	43.60	21.80
1	Linux MySQL Database Server	One in Trusted network for legacy application	0.12	21.80	21.80
7	NitroGuard IPS	One at each location, one in DMZ and one in the Trusted network	40.53	5,627.82	1,607.95
1	Netscreen Firewall	Netscreen facing the Internet	0.58	2,414.00	2,414.00
3	Cisco Pix Firewalls	Between the data center and the other four sites, in front of Trusted network, between Trusted and the DMZ	39.00	1,734.00	1,178.00
1	Cisco VPN Concentrator	Located at data center facing the Internet	0.83	69.45	69.45
1	Squid Proxy	Located at data center	14.58	269.03	269.03
Totals:			149.79	15,598.90	8,118.80

Source: SANS Institute

Storage Needs

A rule of thumb is that an average event occupies 300 bytes. So for every 1,000 EPS (86.4 million events per day), the SIEM needs to store:

25.9GB
PER DAY

777GB
PER MONTH

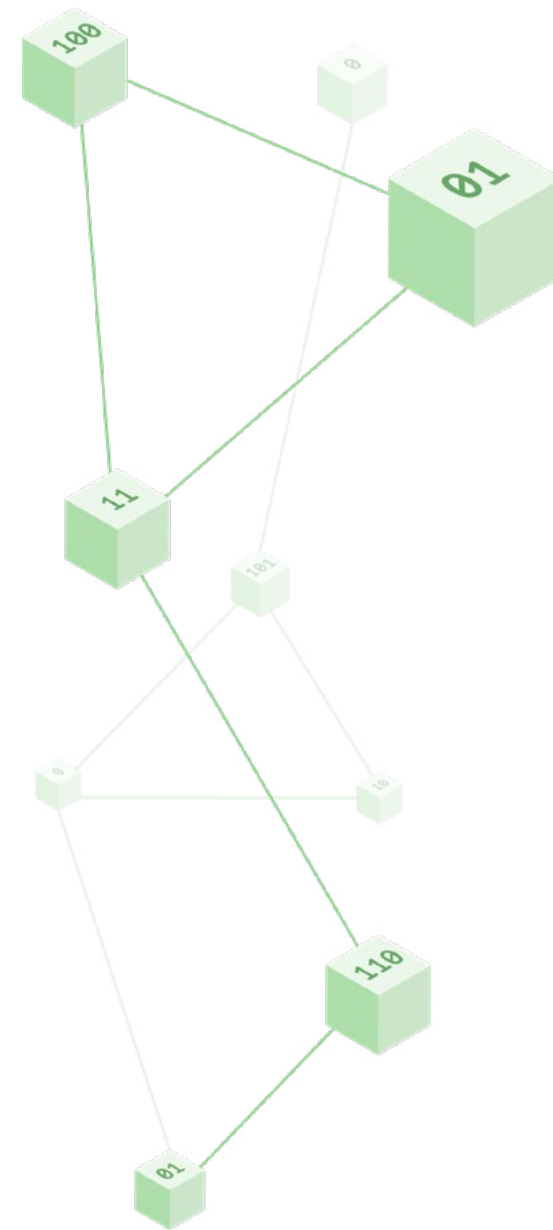
9,331GB
PER ANNUM

In order to size your SIEM, conduct an inventory of the devices you intend to collect logs from. Multiply the number of similar devices by their estimated EPS, to get a total number of Events Per Day across your network.

Hardware Sizing

After you determine your event velocity and volume, consider the following factors to size hardware for your SIEM:

- **Storage format** – how will files be stored? Using a flat file format, a relational database or an unstructured data store like Hadoop?
- **Storage deployment and hardware** – is it possible to move data to the cloud? If so, cloud services like Amazon S3 and Azure Blob Storage will be highly attractive for storing most SIEM data. If not, consider what storage resources are available locally, and whether to use commodity storage with Hadoop or NoSQL DBs, or high-performance storage appliances.
- **Log compression** – what technology is available to compress log data? Many SIEM vendors advertise compression ratios of 1:8 or more.
- **Encryption** – is there a need to encrypt data as it enters the SIEM data store? Determine software and hardware requirements.
- **Hot storage (short-term data)** – needs high performance to enable real time monitoring data analysis.
- **Long-term storage (data retention)** – needs high volume, low cost storage media to enable maximum retention of historic data.
- **Failover and backup** – as a mission critical system, the SIEM should be built with redundancy, and be backed by a clear business continuity plan.



Scalability and Data Lakes

In the past decade, networks have grown, the number of connected devices has exploded, and data volumes shot up exponentially. In addition, there is a growing need to have access to all historic data—not just a filtered, summarized version of the data—to enable deeper analysis. Modern SIEM technology can make sense of huge volumes of historic data and use it to discover new anomalies and patterns.

In 2015 O'Reilly released a report, [The Security Data Lake](#), which offered a robust approach for storing SIEM data in a Hadoop data lake. The report clarifies that data lakes do not replace SIEMs—the SIEM is still needed for its ability to parse and make sense of log data from many different systems, and later analyze and extract insights and alerts from the data.

The data lake, as a companion to a SIEM, provides:

- Nearly unlimited, low cost storage based on commodity devices.
- New ways of processing big data—tools in the Hadoop ecosystem, such as Hive and Spark—enable fast processing of huge quantities of data, while enabling traditional SIEM infrastructure to query the data via SQL.
- The possibility of retaining all data across a multitude of new data sources, like cloud applications, IoT and mobile devices.

Today additional technical options exist for implementing data lakes, besides the heavyweight Hadoop—including Elasticsearch, Cassandra and MongoDB.

Next-gen SIEM

Another benefit of data lake storage is that hardware costs become predictable. You can simply add nodes to the data lake, running on commodity or cloud hardware, to grow data storage linearly. SIEMs based on data lake technology can easily add new data sources or expand data retention at low cost.

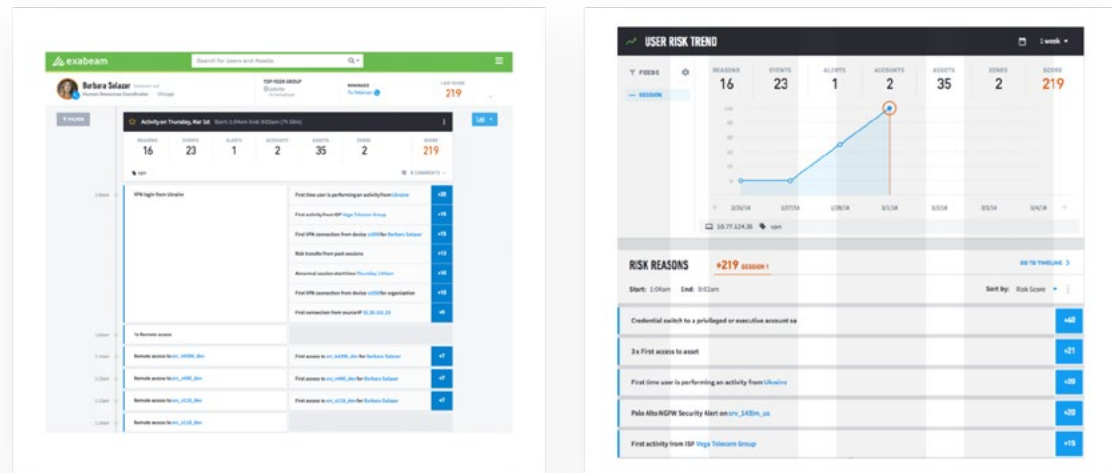
SIEM Reporting, Dashboards and Visualization

The main purpose of a SIEM is to generate actionable insights for security teams. These come in several forms:

- **Alerts and notifications** – prompt security staff to investigate an anomaly or apparent security issue
- **Data exploration** – enable security staff to freely explore data to actively hunt for threats, or investigate a known security incident
- **Dashboards** – display status of security-related systems and metrics and highlight potential security issues
- **APIs and web services** – enable the use of external systems, such as BI and behavioral analytics tools, to access SIEM data and analyze it from new perspectives

Next-gen SIEM

Next-generation SIEMs use behavioral profiling and machine learning techniques to identify security incidents and help teams collect pertinent data for the incident, across devices, user profiles and time periods.



A dashboard and automatically-created incident ticket, provided by Exabeam's next-generation SIEM platform.
Source: Exabeam

SIEM Architecture: Then and Now

Historically, SIEMs were an expensive, monolithic enterprise infrastructure, built with proprietary software and custom hardware provisioned to handle its large data volumes. Along with the software industry in general, SIEMs are evolving to become more agile and lightweight, and much smarter than they were before.

Next-generation SIEM solutions use a modern architecture that is more affordable, easier to implement, and helps security teams discover real security issues faster:

- **Modern data lake technology** – offering big data storage with unlimited scalability, low cost and improved performance.
- **New managed hosting and management options** – MSSPs are helping organizations implement SIEM by running part of the infrastructure (on-premises or on the cloud), and by providing expertise to manage security processes.
- **Dynamic scalability and predictable costs** – SIEM administrators no longer need to meticulously calculate sizing, and make architectural changes when data volumes grow. SIEM storage can now grow dynamically and predictably when volumes increase.
- **New insights with user and entity behavior analytics (UEBA)** – SIEM architectures today include advanced analytics components such as machine learning and behavioral profiling, which go beyond traditional correlations to discover new relationships and anomalies across huge data sets. Read more in our chapter on UEBA.
- **Powering incident response** – modern SIEMs leverage security orchestration, automation and response (SOAR) technology that helps identify and automatically respond to security incidents, and supports incident investigation by security operations center (SOC) staff. Read more in our chapter on incident response.

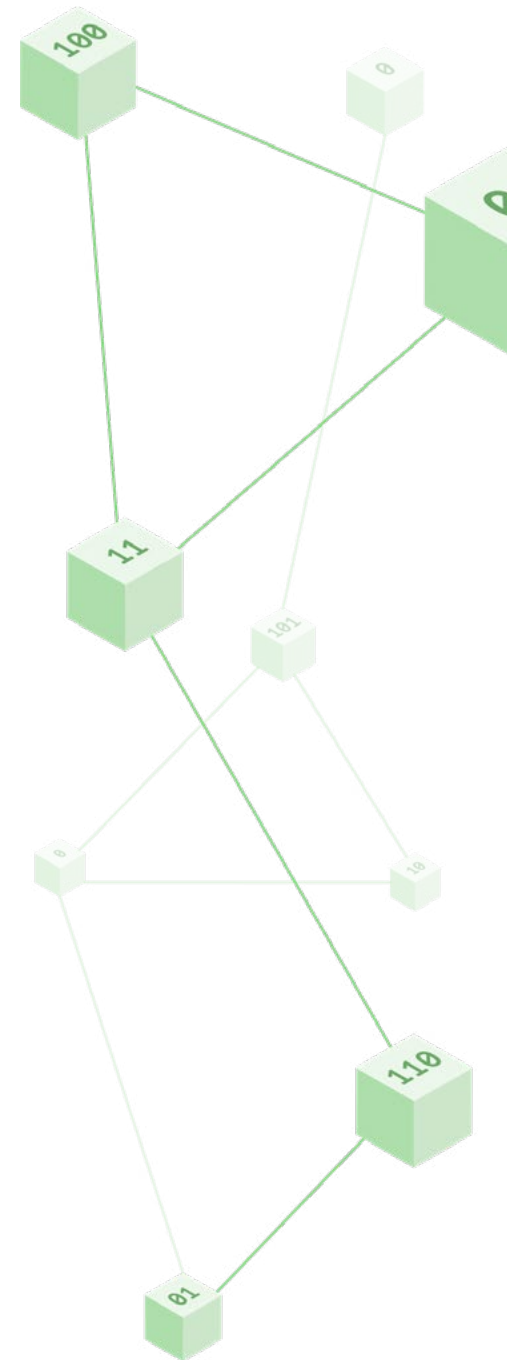
To see an example of a modern SIEM architecture, see [Exabeam's Security Management Platform](#).

Log Aggregation, Processing and Analysis for Security

Logs and events are a foundation of modern security monitoring, investigation and forensics. In this chapter you'll learn in-depth how logs are aggregated, processed and stored, and how they are used in the security operations center (SOC).

In this chapter you will learn:

- [Log aggregation](#) – four methods of log aggregation
- [Log processing](#) – how logs are parsed, normalized, enriched and indexed for fast access
- [Log types](#) – examples of logs, understanding log formats including CSV, JSON, CEF
- [Log monitoring](#) – how logs are used to identify problems and trends in production systems
- [Security event logs](#) - the basics—what are events and incidents and how they are used in security investigations
- [Log analysis for security with SIEM](#) – how SIEMs use logs to help identify and investigate security incidents
- [Using endpoint logs for security](#) – Windows event logs, Windows security logs, Linux event logs and iOS logs
- [Managing endpoint detection and response \(EDR\) logs](#) – about EDR system logs, with examples from Symantec and McAfee
- [Firewall logs](#) – firewall log analysis basics, examples from Windows Firewall, Linux Firewall, Cisco and Check Point



What is Log Aggregation?

Log aggregation is the process of collecting logs from multiple computing systems, parsing them and extracting structured data, and putting them together in a format that is easily searchable and explorable by modern data tools.

There are four common ways to aggregate logs—many log aggregation systems combine multiple methods.

Syslog



A standard logging protocol. Network administrators can set up a Syslog server that receives logs from multiple systems, storing them in an efficient, condensed format which is easily queryable.

Log aggregators can directly read and process Syslog data.

Event streaming



Protocols like SNMP, Netflow and IPFIX allow network devices to provide standard information about their operations, which can be intercepted by the log aggregator, parsed and added to central log storage.

Log collectors



Software agents that run on network devices, capture log information, parse it and send it to a centralized aggregator component for storage and analysis.

Direct access



Log aggregators can directly access network devices or computing systems, using an API or network protocol to directly receive logs. This approach requires custom integration for each data source.

What is Log Processing?

Log processing is the art of taking raw system logs from multiple sources, identifying their structure or schema, and turning them into a consistent, standardized data source.

The Log Processing Flow



01

Log parsing

Each log has a repeating data format which includes data fields and values. However, the format varies between systems, even between different logs on the same system.

A log parser is a software component that can take a specific log format and convert it to structured data. Log aggregation software includes dozens or hundreds of parsers written to process logs for common systems.

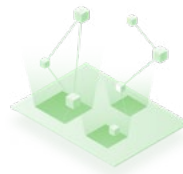


02

Log normalization and categorization

Normalization merges events containing different data into a reduced format which contains common event attributes. Most logs capture the same basic information—time, network address, operation performed, etc.

Categorization involves adding meaning to events—identifying log data related to system events, authentication, local/remote operations, etc.



03

Log enrichment

Log enrichment involves adding important information that can make the data more useful.

For example, if the original log contained IP addresses, but not actual physical locations of the users accessing a system, a log aggregator can use a geolocation data service to find out locations and add them to the data.



04

Log indexing

Modern networks generate huge volumes of log data. To effectively search and explore log data, there is need to create an index of common attributes across all log data.

Searches or data queries that use the index keys can be an order of magnitude faster, compared to a full scan of all log data.



05

Log storage

Because of the massive volumes of logs, and their exponential growth, log storage is rapidly evolving. Historically, log aggregators would store logs in a centralized repository. Today, logs are increasingly stored on data lake technology, such as Amazon S3 or Hadoop.

Data lakes can support unlimited storage volumes with low incremental storage cost, and can provide access to the data via distributed processing engines like MapReduce, or modern high performance analytics tools.

Log Types

Almost every computing system generates logs. Below are a few of the most common sources of log data.



Endpoint logs

An endpoint is a computing device within a network—such as a desktop, laptop, smartphone, server or workstation. Endpoints generate multiple logs, from different levels of their software stack—hardware, operating system, middleware and database, and applications. Endpoint logs are taken from the lower levels of the stack, and used to understand the status, activity and health of the endpoint device.



Router logs

Network devices like routers, switches and load balancers are the backbone of network infrastructure. Their logs provide critical data about traffic flows, including destinations visited by internal users, sources of external traffic, traffic volumes, protocols used, and more. Routers typically transmit data via the Syslog format, and data can be captured and analyzed via your network's Syslog servers.



Application event logs

Applications running on servers or end user devices generate and log events. The Windows operating system provides a [centralized event log](#) that collects startup, shutdown, heartbeat and run-time error events from running applications. In Linux, application log messages can be found in the /var/log folder. In addition, log aggregators can directly collect and parse logs from enterprise applications, such as email, web or database servers.



IoT logs

A new and growing source of log data is internet of things (IoT) connected devices. IoT devices may log their own activity and/or sensor data captured by the device. IoT visibility is a major challenge for most organizations, as many devices have no logging at all, or save log data to local file systems, limiting the ability to access or aggregate it. Advanced IoT deployments save log data to a central cloud service; many are adopting a new log collection protocol, syslog-ng, which focuses on portability and central log collection.



Proxy logs

Many networks maintain a transparent proxy, providing visibility over traffic of internal users. Proxy server logs contain requests made by users and applications on a local network, as well as application or service requests made over the internet, such as application updates. To be useful, proxies must be enforced across all, or at least critical segments, of user traffic, and measures must be in place to decrypt and interpret HTTPS traffic.

Examples of Logs

Common Log Formats

Common log formats: CSV, JSON, key value pair, common event format (CEF)

- **CSV log format**

5:39:55 → Time

[Fname, Lname, name@company] → User Credentials

Sign-in Failed → Authentication Event

173.0.0.0 → IP

/app/office365 → App User Signed Into

- **JSON log format**

MachineName → User's host

Message → The event is a Kerberos service ticket (user already authenticated and sending access request for specific service)

TimeGenerated → Time of event

TargetUserName → Username attempting to login

TargetDomainName → Domain user attempted to login to

ServiceName → Service user attempted to log into

- **Common event format (CEF)**

CEF is an open log management standard that makes it easier to share security-related data from different network devices and applications. It also provides a common event log format, making it easier to collect and aggregate log data. CEF uses the syslog message format.

- **Common event format**

CEF:Version|Device Vendor|Device Product|Device Version|Signature
ID|Name|Severity|Extension

Bracket enclosing Trend Micro .. 3.5.4 → Uniquely identifies the sending device. No two products may use the same vendor-product pair.

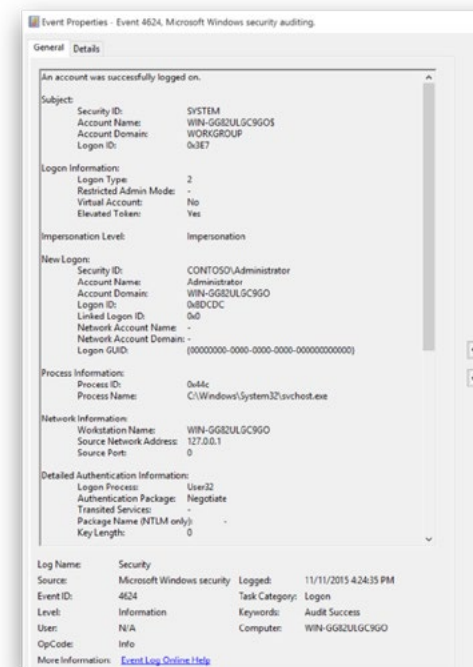
600 → Unique identifier per event type, for example in IDS systems each signature or rule has a unique Signature ID

4 → Severity of the event from 1-10

Suser=Master.. → a collection of key-value pairs which allow the log entry to contain additional info, from an extensive Extension Dictionary including events like deviceAction, ApplicationProtocol, deviceHostName, destinationAddress and DestinationPort, or custom events.

- **Sample log entry**

Jan 18 11:07:53 dsmhost CEF:0|Trend Micro|Deep Security Manager|3.5.4|600|Administrator Signed In|4|suser=Master...

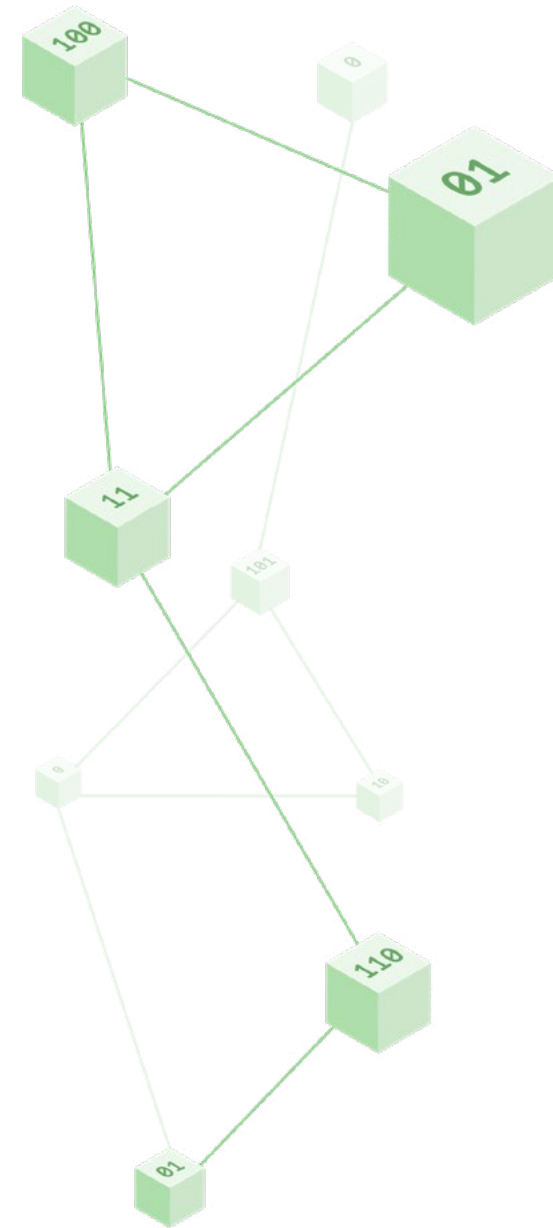


Example of Windows log data: Account successfully logged in.
Image Source: Microsoft

What is Log Monitoring?

There is a wealth of information in log files that can help identify problems and patterns in production systems. Log monitoring involves scanning log files, searching for patterns, rules or inferred behavior that indicates important events, and triggering an alert sent to operations or security staff.

Log monitoring can help identify problems before they are experienced by users. It can uncover suspicious behavior that might represent an attack on organizational systems. It can also help record baseline behavior of devices, systems or users, in order to identify anomalies that require investigation.



Security Event Logs—the Basics

Log aggregation and log monitoring is a central activity for security teams. Collecting log information from critical systems and security tools, and analyzing those logs, is the most common way to identify anomalous or suspicious events, which might represent a security incident.

The two basic concepts of security log management are **events** and **incidents**—an event is something that happens on a network on an endpoint device. One or more events can be identified as an incident—an attack, violation of security policies, unauthorized access, or change to data or systems without the owner's consent.

Common Security-Relevant Log Events

- Report from antivirus software that a device is infected by malware
- Report from firewall about traffic to/from a prohibited network address
- Attempt to access a critical system from an unknown host or IP address
- Repeated failed attempts to access a critical system
- Change in user privileges
- Usage of insecure or prohibited protocols / ports

Common Security Incidents

- Malicious email received and activated by organizational users
- Malicious website accessed by organizational users (e.g., drive by download)
- Improper or prohibited usage by an authorized user
- Unauthorized access
- An attempt to compromise, deny access to, or delete organizational systems
- Loss or theft of equipment, such as employee laptops, servers
- Data leak or malware infection via removable media

Log Analysis for Security with SIEM

In the security world, the primary system that aggregates logs, monitors them and generates alerts about possible security systems, is a SIEM solution.

SIEM platforms aggregate historical log data and real-time alerts from security solutions and IT systems like email servers, web servers and authentication systems.

They analyze the data and establish relationships that help identify anomalies, vulnerabilities and incidents. The SIEM's main focus is on security-related events such as suspicious logins, malware or escalation of privileges.

The SIEM's goal is to identify which events has security significance and should be reviewed by a human analyst, and send notifications for those events. Modern SIEMs also provide extensive dashboards and data visualization tools, allowing analysts to actively seek data points that might indicate a security incident—known as threat hunting.

Traditional SIEM Log Analysis

Traditionally, the SIEM used two techniques to generate alerts from log data: **correlation rules**, specifying a sequence of events that indicates an anomaly, which could represent a security threat, vulnerability or active security incident; and **vulnerabilities and risk assessment**, which involves scanning networks for known attack patterns and vulnerabilities.

The drawback of these older techniques is that they generate a lot of false positives, and are not successful at detecting new and unexpected event types.

Next-Generation SIEM Log Analysis

Advanced SIEMs use technology called **user and entity behavior analytics (UEBA)**. UEBA leverages machine learning to look at patterns of human behavior, automatically establish baselines, and intelligently identify suspicious or anomalous behavior.

This can help detect risks that are unknown or difficult to define with correlation rules, such as insider threats, targeted attacks, fraud, and anomalies across long periods of time or across multiple organizational systems.

Using Endpoint Logs for Security

Traditionally, monitoring and security efforts focused on network traffic to identify threats. Today, there is a growing focus on endpoints, such as desktop computers, servers and mobile devices. Endpoints are frequently targeted by threat actors who can bypass traditional security measures—for example, a laptop forgotten on a train can be stolen by an attacker and used to penetrate organizational systems. Without careful monitoring of the laptop's activity, this and similar attacks could go undetected.

Windows Event Logs

The Windows operating system provides an event logging protocol that allows applications, and the operating system itself, to log important hardware and software events. The events can be viewed directly by an administrator using the Windows Event Viewer.

Which events are logged?

Events logged in Windows event logs include application installations, security management (see Windows security logs below), initial startup operations, and problems or errors. All these event types can have security significance, and should be monitored by log aggregation and monitoring tools.

Example of Windows event log

```
Warning 5/11/2018 10:29:47 AM Kernel-Event Tracing
1 Logging
```

Windows Security Logs

The Windows security log is a part of the Windows event log framework. It contains security-related events specified by administrators using the system's audit

policy. Microsoft describes the security log as “Your Best and Last Defense” when investigating security breaches on Windows systems.

Which events are logged?

The following types of Windows log events can be defined as security events: account logon, account management, directory service access, logon, object access (for example, file access), policy change, privilege use, tracking of system processes, system events.

iOS Logs and iOS Crash Reports

Unlike Windows and Linux, the iOS operating system does not log system and application events by default, with the exception of application crash reports. iOS 10.0 onwards offers a logging API that allows specific applications to log application events and store them to a centralized location on disk. Log messages can be viewed using the console app of the log command-line tool.

Because iOS does not provide convenient remote access to logs, several third-party solutions have emerged that allow for remote collection and aggregation of iOS logs.

Linux Event Logs

Linux logs record a timeline of events that occur in the Linux operating system and applications. Central system logs are stored in the `/var/log` directory, and logs for specific applications may be stored in the application folder, for example `~/chrome/Crash Reports` for Google Chrome.

Which events are logged?

There are Linux log files for system events, kernel, package managers, boot processes, Xorg, Apache, MySQL, and other common services. As in Windows, all these events could possibly have security significance.

Which are the most critical Linux logs to monitor?

- `/var/log/syslog` or `/var/log/messages`—stores all activity data across the Linux system.
- `/var/log/auth.log` or `/var/log/secure`—stores authentication logs
- `/var/log/boot.log`—messages logged during startup
- `/var/log/maillog` or `var/log/mail.log`—events related to email servers
- `/var/log/kern`—Kernel logs
- `/var/log/dmesg`—device driver logs
- `/var/log/faillog`—failed login attempts
- `/var/log/cron`—events related to cron jobs or the cron daemon
- `/var/log/yum.log`—events related to installation of yum packages
- `/var/log/httpd/`—HTTP errors and access logs containing all HTTP requests
- `/var/log/mysqld.log` or `/var/log/mysql.log`—MySQL log files

Managing Endpoint Detection and Response (EDR) Logs

Endpoint detection and response (EDR) technology helps to detect, investigate and mitigate security incidents on organizational endpoints. EDR is complementary to traditional endpoint tools such as antivirus, data loss prevention (DLP) and SIEM. EDR technology provides visibility into events taking place on endpoints, including application access and activity, operating system operations, creation, modification, copying and movement of data, memory usage, and user access to predefined sensitive data.

EDR systems provide aggregated logs that allow security teams to analyze and explore events from across the enterprise endpoint portfolio.

Symantec Endpoint Protection Logs

Symantec Endpoint Protection is a security suite that includes intrusion prevention, firewall, and anti-malware. Endpoint Protection logs contains information about configuration changes, security-related activities such as virus detections, errors on specific endpoints, and traffic that enters and exits the endpoint.

Which events are logged?

Symantec Endpoint Protection log types include:

- Policy modifications
- Application and device control—events on endpoint devices where some behavior was blocked
- Compliance logs

- Computer status—operational status such as computer name, IP address, infection status
- Deception logs—attacker interaction with “honeypots” deployed by the security solution
- Network and host exploit mitigation
- Virus scan events
- Risk events detected by Symantec
- System log—information about operating system and services.

McAfee Endpoint Security

McAfee Endpoint Security provides centralized management for endpoint devices, anti-malware protection, application containment, web security, threat forensics and machine learning analysis for detection of unknown threats.

The solution allows you to set each endpoint device to one of three log levels: no logging, event logging, and debug logging. Logs are saved on the endpoints in the McAfee folder.

Which events are logged?

McAfee Endpoint Security saves several log files on each endpoint device:

- **myAgent.log**—aggregate log file containing historic logs
- **myNotices.log**—notices and warnings generated by the McAfee agent
- **myUninstall.log**—software uninstall events
- **myUpdate.log**—software update events
- **myInstall.log**—software installation events

Managing Firewall Logs

Firewall logs are extremely valuable for security analysis, because they contain trails of almost all traffic flowing into and out of your network. If malicious activity is occurring, even if it cannot be detected by known malware or attack signatures, it will be captured by the firewall and can probably be seen by analyzing firewall logs for unusual behavior.

For example, when a zero-day virus infects computers on your network, even if it cannot be detected yet by antivirus software, firewall logs may show unusually high numbers of denied connections, or allowed connections, with suspicious remote hosts. A routine review of firewall logs can discover trojans or rootkits trying to connect to their command and control systems via IRC, over the firewall.

Cisco Syslog and Logging Levels

Cisco routers save logs in Syslog format, and also allow logs to be viewed by the admin interface. Messages are tagged with message codes—for example, most denied connections have a message code in the 106001 to 106023 range. Most firewall devices do not have local storage space, so logs must be configured to be sent elsewhere—Cisco allows saving logs to a Syslog server on the network, via SMTP, via console port, telnet, or several other options.

What log entries are important to analyze?

- Connections allowed by firewall security policies—these can help spot “holes” in the security policies
- Connections denied by firewall security policies—might contain suspicious or attack behavior
- Using the deny rate logging feature can show DoS or brute force attacks
- IDS activity messages—show attacks identified by Cisco Intrusion Detection features
- User authentication and command usage—let you review and audit firewall policy changes

- Bandwidth usage—shows connections by duration and traffic volume—outliers could be interesting to investigate
- Protocol usage messages—show protocols and port numbers—can show unusual or insecure protocols used on the network
- NAT or PAT connections—check if you receive a report of malicious activity coming inside your network

Check Point Logging

Check Point routers can save logs in Syslog format, and also allow logs to be viewed over an admin interface. Check Point routers maintain a security log which saves events that are deemed to have security significance.

Categories of events saved to security log:

- Connection accepted
- Connection decrypted
- Connection dropped
- Connection encrypted
- Connection rejected

- Connection monitored—a security event was monitored but not blocked according to current firewall policy
- URL allowed—URL allowed for access by internal users
- URL filtered—URL disallowed for access by internal users
- Virus detected—virus detected in an email
- Potential spam stamped—email marked as potential spam
- Potential spam detected—email rejected as potential spam
- Mail allowed—non-spam email was logged
- VStream antivirus blocked a connection.

Severity levels in the Check Point security log:

- Red—connection attempts blocked by the firewall, by security policy downloaded from the service center or user-defined rules
- Orange—traffic detected as suspicious but accepted by the firewall
- Green—traffic accepted by the firewall

Log Management and Next-Generation SIEMs

Log management has always been complex, and is becoming more so with the proliferation of network devices, endpoints, microservices and cloud services, and exponentially increasing traffic and data volumes.

In a security environment, next-generation SIEM solutions can help manage and extract value from security-relevant log events:

- Next-generation SIEMs are based on data lake technology which can store unlimited data volumes of historical logs
- Next-generation SIEMs come with UEBA technology which can automatically establish baseline activity for devices and users, and identify anomalous or suspicious behavior
- Next-generation SIEMs provide advanced data exploration capabilities which can help security analysts perform threat hunting by actively searching through logs

Exabeam is an example of a next-generation SIEM platform that provides these capabilities. It can pull together logs from enterprise systems and security tools and perform the complete log management process, including log collection and aggregation, log processing, log analysis using advanced analytics and UEBA technology, and alerting about security incidents.

Learn more at exabeam.com/product

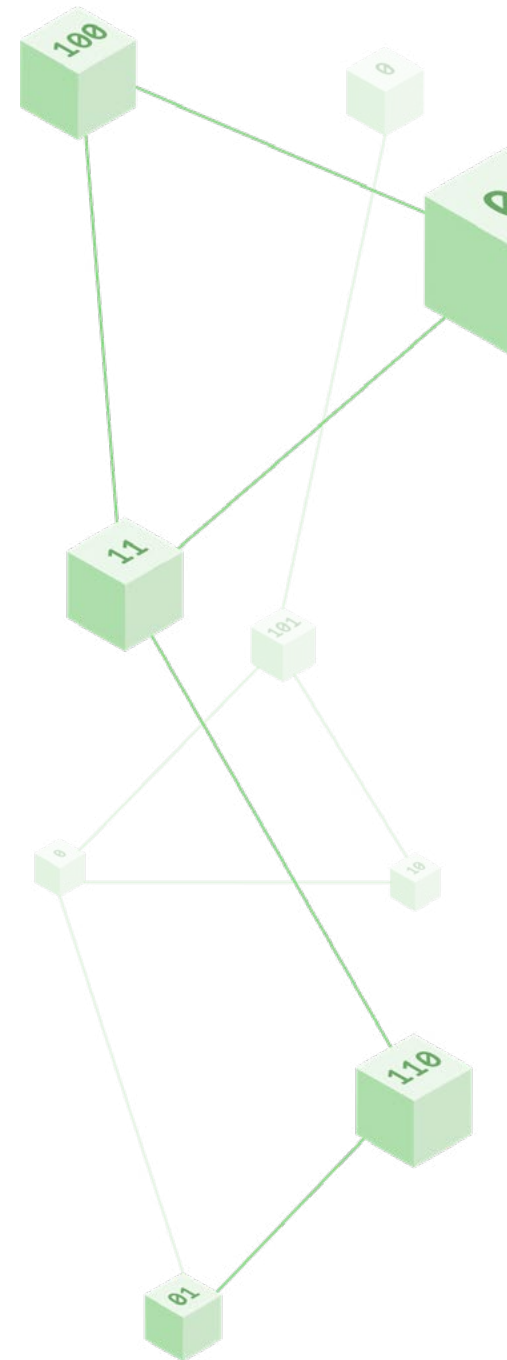
User and Entity Behavior Analytics

User and entity behavior analytics (UEBA) is a new category of security solutions that use innovative analytics technology, including machine learning and deep learning, to discover abnormal and risky behavior by users, machines and other entities on the corporate network.

UEBA can detect security incidents that traditional tools do not see, because they do not conform to predefined correlation rules or attack patterns, or because they span multiple organizational systems and data sources.

In this chapter you will learn:

- [The three pillars of UEBA solutions](#) as defined by Gartner
- How UEBA solutions are converging with [security information and event management \(SIEM\)](#)
- [Major use cases of UEBA](#) – malicious insider, incident prioritization, DLP, and more
- [How UEBA works](#) –behavioral profiling, risk modeling and timeline analysis
- An example of an [integrated SIEM and UEBA solution](#)



What is User and Entity Behavior Analytics (UEBA)?

UEBA solutions build profiles that model standard behavior for users and entities in an IT environment, such as servers, routers and data repositories. This is known as baselining. Using a variety of analytics techniques, UEBA technology can identify activity that is anomalous compared to the established baselines, discover threats and detect security incidents.

Three Pillars of UEBA

Gartner defines UEBA solutions across three dimensions:

- **Use cases** – UEBA solutions provide information on the behavior of users and other entities in the corporate network. They should perform monitoring, detection and alerting of anomalies. And they should be applicable for multiple use cases—unlike specialized tools for employee monitoring, trusted hosts monitoring, fraud, and so on.
- **Data sources** – UEBA solutions are able to ingest data from a general data repository such as a data lake or data warehouse, or through a SIEM. They should not deploy agents directly in the IT environment to collect the data.
- **Analytics** – UEBA solutions detect anomalies using a variety of analytics approaches—statistical models, machine learning, rules, threat signatures and more.

Convergence of UEBA and SIEM

There is a close relationship between UEBA and SIEM technologies, because UEBA relies on cross-organizational security data to perform its analyses, and this data is typically collected and stored by a SIEM.

In [Gartner's vision](#) of a next-generation SIEM solution, a SIEM should include built-in UEBA functionality. The report lists the following as critical capabilities of a modern SIEM:

- **User monitoring**, including baselining and advanced analytics to analyze access and authentication data, establish user context and report on suspicious behavior.
- **Advanced analytics** – applying sophisticated statistical and quantitative models, such as machine learning and deep learning, on security log and event data to detect anomalous activity. Advanced analytics should complement the traditional rule and correlation-based analytics available in traditional SIEMs.

UEBA Use Cases

Malicious insider

A malicious insider is an employee or contractor with privileged access to IT systems, who intends to perform a cyber attack against the organization. It is difficult to measure malicious intent or discover it through log files or regular security events. UEBA solutions help by establishing a baseline of a user's typical behavior and detect abnormal activity.

Compromised insider

It's common for attackers to infiltrate an organization and compromise a privileged user account or trusted host on the network, and continue the attack from there. UEBA solutions can help rapidly detect and analyze bad activities that the attacker carries on via the compromised account.

Traditional security tools find it difficult to detect a compromised insider if the attack pattern or kill chain is not currently known (such as in a zero day attack), or if the attack moves laterally through an organization by changing credentials, IP addresses, or machines. UEBA technology, however, can detect these types of attacks, because they will almost always force assets to behave differently from established baselines.

Incident prioritization

A SIEM collects events and logs from multiple security tools and critical systems, and generates a large number of alerts that must be investigated by security staff. This leads to alert fatigue, a common challenge of [security operations centers](#) (SOC).

UEBA solutions can help understand which incidents are particularly abnormal, suspicious or potentially dangerous in the context of your organization. UEBA can go beyond baselines and threat models by adding data about organizational structure—for example, the criticality of assets and the roles and access levels of specific organizational functions. A small deviation from norm for a critical protected system or a top-level administrator, might be worth a look for an investigator; for a run-of-the-mill employee only a major deviation would receive high priority.

Data loss prevention (DLP) and data leakage prevention

Data loss prevention (DLP) tools are used to prevent data exfiltration, or the illicit transfer of data outside organizational boundaries. Traditional DLP tools report on any unusual activity carried out on sensitive data—they create a high volume of alerts which can be difficult for security teams to handle.

UEBA solutions can take DLP alerts, prioritize and consolidate them by understanding which events represent anomalous behavior compared to known baselines. This saves time for investigators and helps them discover real security incidents faster.

Entity analytics (IoT)

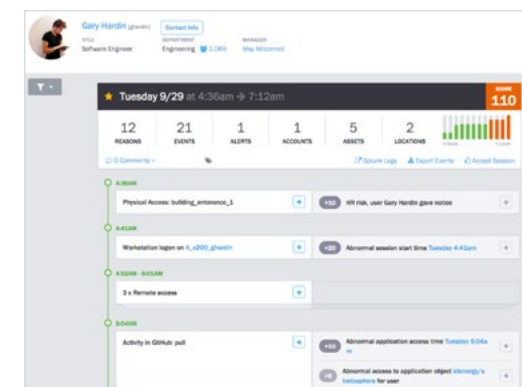
UEBA can be especially important in dealing with Internet of Things (IoT) security risks. Organizations deploy large fleets of connected devices, often with minimal or no security measures. Attackers can compromise IoT devices, use them to steal data or gain

access to other IT systems, or worse—leverage them in DDoS or other attacks against third parties.

Two sensitive categories of IoT are medical devices and manufacturing equipment. Connected medical devices may contain critical data, and may be life threatening if used directly for patient care. Manufacturing equipment can cause large financial losses if disrupted, and in some cases may threaten employee safety.

UEBA can track an unlimited number of connected devices, establish a behavioral baseline for each device or group of similar devices, and immediately detect if a device is behaving outside its regular boundaries. For example:

- Connections to or from unusual addresses or devices
- Activity at unusual times
- Device features activated which are typically not used



Unusual activity by an insider—detected by the [Exabeam UEBA solution](#) as part of its next-generation SIEM. Source: Exabeam

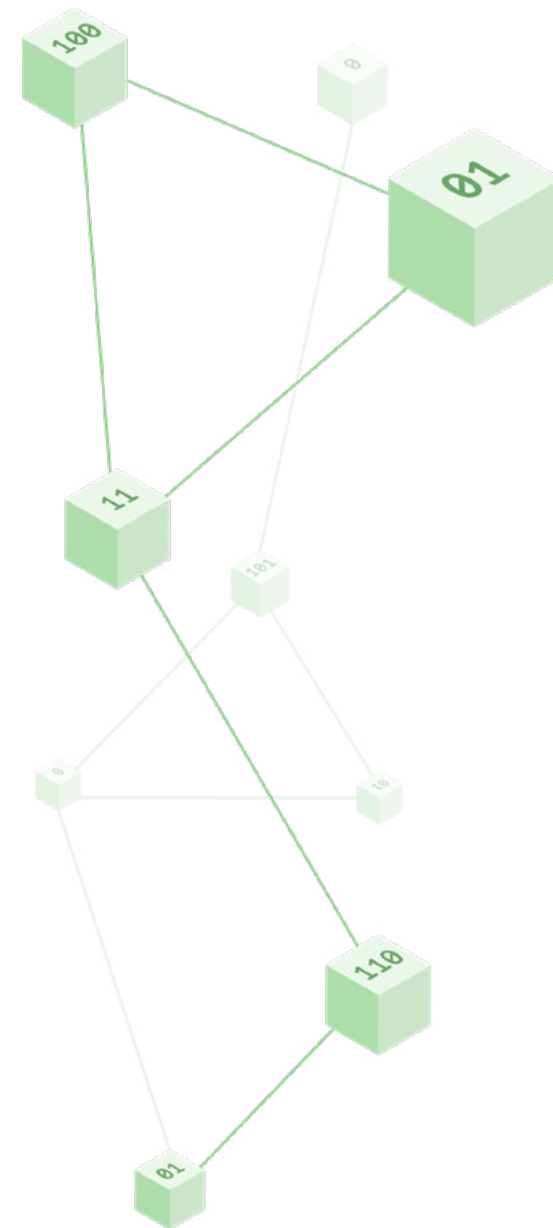
UEBA Analytics Methods

Some UEBA solutions rely on traditional methods to identify suspicious activity. These can include manually-defined rules, correlations between security events and known attack patterns. The limitation of traditional techniques is that they are only as good as the rules defined by security administrators, and cannot adapt to new types of threats or system behavior.

Advanced analytics, which is the hallmark of UEBA tools, involves several modern technologies that can help identify abnormal behavior even in the absence of known patterns:

- **Supervised machine learning** – sets of known good behavior and known bad behavior are fed into the system. The tool learns to analyze new behavior and determine if it is “similar to” the known good or known bad behavior set.
- **Bayesian networks** – can combine supervised machine learning and rules to create behavioral profiles.
- **Unsupervised learning** – the system learns normal behavior, and is able to detect and alert on abnormal behavior. It will not be able to tell if the abnormal behavior is good or bad, only that it deviates from normal.
- **Reinforced / semi-supervised machine learning** – a hybrid model where the basis is unsupervised learning, and actual alert resolutions are fed back into the system to allow fine tuning of the model and reduce the signal-to-noise ratio.
- **Deep learning** – enables virtual alert triage and investigation. The system trains on data sets representing security alerts and their triage outcomes, performs self-identification of features, and is able to predict triage outcomes for new sets of security alerts.

Traditional analytics techniques are deterministic, in the sense that if certain conditions were true, an alert was generated, and if not the system assumed “all is fine”. The advanced analytics methods listed above are different in that they are **heuristic**. They compute a risk score which is a **probability** that an event represents an anomaly or security incident. When the risk score exceeds a certain threshold, the system creates a security alert.



How UEBA Works

Holistic Analysis Across Multiple Data Sources

The true power of a UEBA solution is in its ability to cut across organizational boundaries, IT systems and data sources and analyze all the data available for a specific user or entity.

A UEBA solution should analyze as many data sources as possible. Some example data sources include:

- Authentication systems like Active Directory
- Access systems like VPN and proxies
- Configuration management databases
- Human resources data—new employees, departed employees, and any data that provides additional context on users
- Firewall, intrusion detection and prevention systems (IDPS)
- Anti-malware and antivirus systems
- Endpoint detection and response systems
- Network traffic analytics
- Threat intelligence feeds

For example, a UEBA solution should be able to identify unusual login via Active Directory, cross reference it with the criticality of the device being logged onto, the sensitiveness of the files accessed, and recent unusual network or malware activity which may have enabled a compromise.

Behavioral Baselining and Risk Scores

A UEBA solution learns normal behavior to identify abnormal behavior. It examines a broad set of data to determine a user's baseline or behavioral profile.

For example, the system monitors a user and sees how they use a VPN, at what time they arrive to work and which systems they log into, what printer they use, how often and what size of files they send by email or load to a USB drive, and many other data points that define the user's "normal behavior". The same is done for servers, databases or any significant IT system.

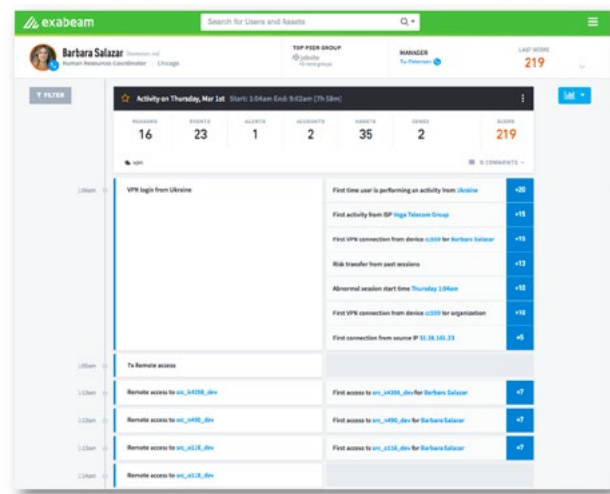
When there is deviation from the baseline, the system adds to the risk score of that user or machine. The more unusual the behavior, the higher the risk score. As more and more suspicious behavior accumulates, the risk score increases until it hits a threshold, causing it to be escalated to an analyst for investigation.

This analytical approach has several advantages:

- **Aggregation** – the risk score is made up of numerous events, so there is no need for analysts to manually review large numbers of individual alerts and mentally combine them to detect a threat.
- **Reduced false positives** – one slightly abnormal event on its own will not result in a security alert. The system requires multiple signs of abnormal behavior to create an alert, reducing the number of false positives and saving time for analysts.
- **More context** – traditional correlation rules defined by security administrators may have been correct for one set of users or systems, but not for others. For example, if a department starts employing shift workers or offshore workers, they will start logging in at unusual times, which would trigger a rule-based alert all the time. UEBA is smarter because it establishes a context-sensitive baseline for each user group. An offshore worker logging in at 3:00 a.m. local time would not be considered an abnormal event.

Timeline Analysis and Session Stitching

When analyzing security incidents, the timeline is a critical concept which can tie together seemingly unrelated activities. Modern attacks are processes, not isolated events.



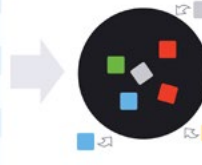
Exabeam Smart Timeline illustrating an attack chain.
Image Source: Exabeam

Advanced UEBA solutions can “stitch” together data from different systems and event streams, to construct the complete timeline of a security incident.

For example, consider a user who logged in, performed suspicious activity and then disappeared from the logs. Was the same IP used to connect to other organizational systems shortly afterwards? If so, this could be part of the same incident, with the same user continuing their attempt to penetrate the system. An additional example could be an attacker logging in to the same machine multiple times using different credentials. This also requires “stitching” together data about the various login attempts and flagging them as a single incident.

Incomplete Data

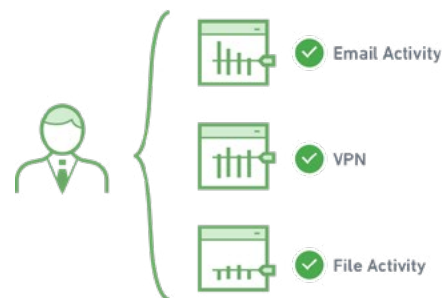
Log 1	Barry		Device_1
Log 2		142.1.12.68	Device_1
Log 3	Barry	142.1.12.68	
Log 4	Barry	37.4.1.23	
Log 5	Barry		Device_2
Log 6		37.4.1.23	Device_2



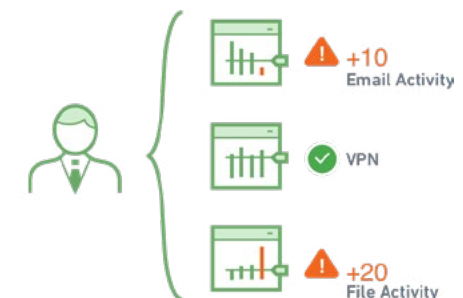
Automated Host-to-IP Mapping

TIME	USER	IP ADDRESS
12:03	Barry	142.1.12.68 37.4.1.23
12:03	Barry	Device_1
12:03	Barry	Device_2
TIME	IP ADDRESS	HOST
12:03	142.1.12.68	Device_1
12:03	37.4.1.23	Device_2

Once a UEBA solution stitches together all relevant data, it can assign risk scores to any activity along the event timeline.



Normal behavior for all users and machines is learned



Risk score is added for high risk and anomalous behavior

Example of an Integrated SIEM and UEBA Solution

Gartner's vision of an integrated SIEM and UEBA solution is today a reality. Several systems are deployed in the field which combine the breadth of data in a SIEM with the deep analytics made possible by cutting-edge UEBA engines.

One example of an integrated system is [Exabeam's Security Management Platform \(SMP\)](#). Exabeam is a full SIEM solution based on modern data lake technology. In addition, it provides the following UEBA capabilities:

- **Rule and signature-free incident detection** – Exabeam uses advanced analytics to identify abnormal and risky activity without predefined correlation rules or threat patterns. It provides meaningful alerts without requiring heavy setup and fine tuning, and with lower false positives.
- **Automatic timelines for security incidents** – Exabeam can stitch together related security events into a timeline that shows a security incident, spanning multiple users, IP addresses and IT systems.
- **Dynamic peer groupings** – Exabeam not only performs behavioral baselining of individual entities, it also dynamically groups similar entities (such as users from the same department, or IoT devices of the same class), to analyze normal collective behavior across the entire group and detect individuals who exhibit risky behavior.

- **Lateral movement detection** – Exabeam detects attackers as they move through a network using different IP addresses, credentials and machines, in search of sensitive data or key assets. It ties together data from multiple sources to connect the dots and view the attacker's journey through the network.

Learn more about Exabeam's [SIEM-integrated UEBA capabilities](#).

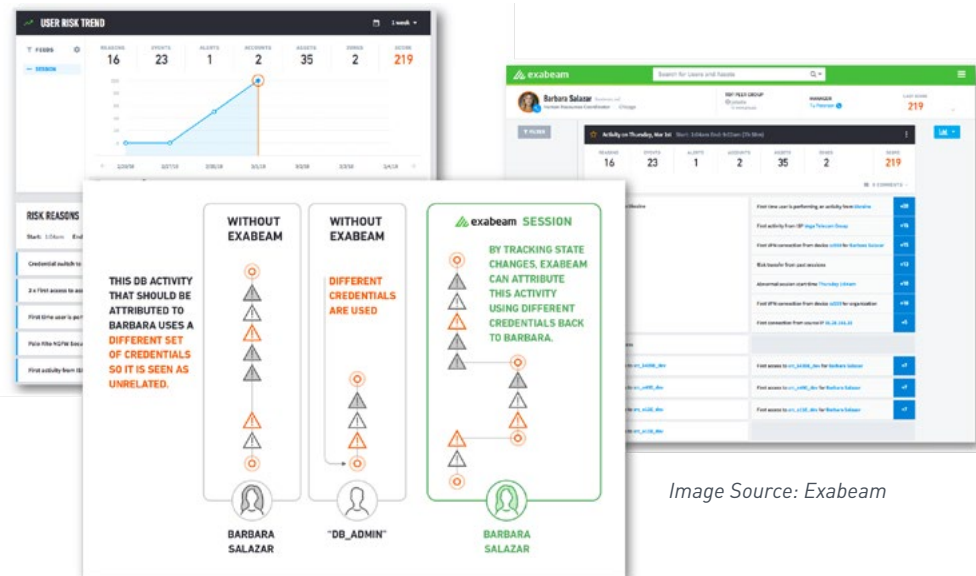


Image Source: Exabeam

10 SIEM Use Cases in a Modern Threat Landscape

Security information and event management (SIEM) systems aggregate security data from across the enterprise; help security teams detect and respond to security incidents; and create compliance and regulatory reports about security-related events. Because SIEM is a core security infrastructure with access to data from across the enterprise, there are a large variety of SIEM use cases.

Below are common SIEM use case examples, from traditional uses such as compliance, to cutting edge use cases such as insider threat detection and IoT security.

Compliance



Learn how a SIEM can help your organization comply with standards:

01. [PCI DSS compliance](#)
02. [GDPR compliance](#)
03. [HIPAA compliance](#)
04. [SOX compliance](#)

Insider Threats



Learn how a SIEM can help mitigate threats originating from trusted entities:

05. [Insider threats](#)
06. [Highly privileged access abuse](#)
07. [Trusted host and entity compromise](#)

Advanced Security



Learn how a SIEM can help with advanced security threats requiring rich data analysis:

08. [Threat hunting](#)
09. [Data exfiltration](#)
10. [IoT security](#)

SIEM for PCI Compliance

The Payment Card Industry Data Security Standard (PCI DSS) was created to secure credit cardholder data from theft and misuse. It defines 12 security areas in which companies should enhance protection for this type of data. The requirements apply to anyone involved in credit card processing, including merchants, processors, and third party service providers.

5 Ways SIEMs Can Help with PCI Compliance

01. **Perimeter security** - detecting unauthorized network connections and correlating with change management, searching for insecure protocols and services, and checking how traffic is flowing across the DMZ.
02. **User identities** - monitoring any event that results in changes to user credentials, and activity by terminated users
03. **Real time threat detection** - monitoring antivirus logs, monitoring insecure ports and services and correlating with threat intelligence.
04. **Production and data systems** - searching for dev/test or default credentials, replicas, etc. on production systems.
05. **Auditing and reporting** - collecting system and security logs, including specific PCI logging requirements, auditing them in a format suitable for PCI reporting, and generating compliance reports.

SIEM for GDPR Compliance

The General Data Protection Regulation (GDPR) is Europe's new framework for protecting security and privacy for personally identifiable information (PII), which came into force in May 2018. GDPR applies to any legal entity which stores, controls or processes personal data for EU citizens, and focuses on two categories: **personal data**, such as an IP address or username, and **sensitive personal data**, such as biometric or genetic data.

5 Ways SIEMs Can Help with PCI Compliance

01. **Data protection by design** - verifying and auditing security controls, to show that user data underwent appropriate treatment.
02. **Visibility into log data** - providing structured access to log information to enable reporting to individual data owners.
03. **GDPR logging and auditing** - monitoring critical changes to credentials, security groups, and so on; auditing databases and servers storing PII, and automatically tracking assets that store sensitive data.
04. **Breach notification** - detecting data breaches, alerting security staff, analyzing the incident to uncover full impact, and quickly generating detailed reports as required by GDPR.
05. **Record of data processing** - identifying events related to personal data, auditing any changes to the data and generating reports as required by GDPR.

Warning

The SIEM itself can represent a risk under GDPR, because log data might contain PII. GDPR permits retaining data for "legitimate interest" (Article 6), which may allow the retention of log files for security purposes. Consult with your legal council to understand what data you can or cannot retain in the SIEM under GDPR provisions.

SIEM for HIPAA Compliance

HIPAA, the Health Insurance Portability and Accountability Act, is a United States standard pertaining to organizations that transmit health information in electronic form. It applies to organizations of all sizes, from a single physician to national healthcare bodies. HIPAA's security management process standard requires organizations to perform risk analysis, risk management, have a sanction policy for data breaches, and conducts information system activity reviews—a key element of the standard which ensures all the other parts are in order.

9 Ways SIEMs Can Help with HIPAA Compliance

01. **Security management process** - discovering new IT assets, identifying systems at risk, monitoring access to system files, user activity and privileges in critical systems
02. **Employee access** - monitoring access to critical files and data, capturing login attempts and logins from terminated users.
03. **Information access management** - identifying logon success and failures, privilege escalation and modification of user accounts.
04. **Security awareness** - detecting vulnerabilities and malware, detecting systems with no antivirus, monitoring logon to security devices and critical systems.
05. **Security incidents** - automatically detecting threats, generating alerts and prioritizing them, enabling threat investigation, and orchestrating automated response to incidents.
06. **Access control** - monitoring changes to credentials and permissions, session timeouts, and changes to encryption settings.
07. **Audit controls** - monitoring changes to policies, data leakage protection (DLP) events, file integrity and log analysis for protected data.
08. **Data integrity** - monitoring modification of health information and changes to data policies.
09. **Transmission security** - identifying unauthorized communications and attempts to modify applications or storage containing health information.

SIEM for SOX Compliance

The Sarbanes-Oxley Act of 2002 (SOX) is a regulation that sets requirements for US public company boards, management and accounting firms. It was enacted as a reaction to several corporate accounting scandals, including Enron and WorldCom. Two frameworks commonly used by IT organizations to comply with SOX are [COSO](#) and [COBIT](#).

The SOX regulation focuses on making sure that an organization informs management, and is able to demonstrate, via SOX reporting procedures:

- Where sensitive data is stored
- Who has access to it
- What happened to it

A SIEM can be helpful in gathering this data and recording it for SOX audits.

5 Ways SIEM Can Help with SOX Compliance and Audits

- 01. Security policies and standards** - tracking information security policies (for example, an email security policy) and standards (for example, a standard way to secure Windows desktop machines). A SIEM can identify which IT systems are in compliance with policies and standards, and alert about violations in real time.
- 02. Access and authentication** - monitoring account creation, change requests, and activity by terminated employees.
- 03. Network security** - monitoring alerts from firewalls and other edge security devices, and identifying known attack patterns in network traffic.
- 04. Log monitoring** - aggregating security events and alerting on invalid login attempts, port scans, privilege escalations, etc.
- 05. Segregation of duties** - the SOX standard requires that no one person controls an entire data process from beginning to end. A SIEM can ensure, for example, that data entry staff only access the data they are creating, and never view or modify other data.

Insider Threats

According to insider threat statistics provided in the [Verizon 2018 Data Breach Investigation Report](#), three of the top five causes of security breaches were related to an insider threat, and insider threats go undetected for months (in 42% of cases) or even years (38% of cases).

Insider threat detection is challenging—behavior doesn't set off alerts in most security tools, because the threat actor appears to be a legitimate user. However, a SIEM can help discover insider threat indicators via behavioral analysis, helping security teams identify and mitigate attacks.

There is growing awareness of internal security threats, first and foremost insider threats:

- **Malicious insider** - a security threat originating from organization's employees, former employees, contractors or associates
- **Compromised insider** - an external entity which has obtained the credentials of an insider

6 Ways a SIEM Can Help Stop Insider Threats

01. **Detecting compromised user credentials** - SIEMs can use behavioral analysis to detect anomalous behavior by users, indicating a compromise. For example, logins at unusual hours, at unusual frequency, or accessing unusual data or systems.
02. **Anomalous privilege escalation** - SIEMs can detect users changing or escalating privileges for critical systems.
03. **Command and control communication** - SIEMs can correlate network traffic with threat intelligence, to discover malware communicating with external attackers. This is a sign of a compromised user account.
04. **Data exfiltration** - SIEMs can use behavioral analysis to combine and analyze seemingly unrelated events, such as insertion of USB thumb drives, use of personal email services, unauthorized cloud storage or excessive printing.
05. **Rapid encryption** - SIEMs can detect and stop encryption of large volumes of data. This might indicate a ransomware attack, which often originates from compromised insiders.
06. **Lateral movement** - insiders conducting an attack may attempt to switch accounts, machines and IP addresses on their way to a target. SIEMs can detect this behavior because they have a broad view of multiple IT systems.

Next-gen SIEM

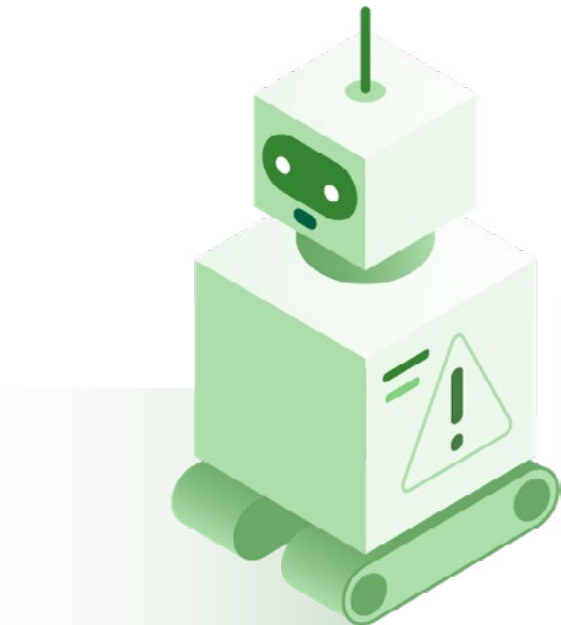
Most of the capabilities in this and the following sections are made possible by next-generation SIEMs that combine user and entity behavior analytics (UEBA). UEBA technology uses machine learning and behavioral profiling to establish baselines of IT users and systems, and intelligently identify anomalies, beyond the rules and statistical correlations used by traditional SIEMs.

Highly Privileged Access Abuse

Privileged access abuse is a complex problem stemming from gaps in access control at organizations. Users with access to IT systems are able to perform undesirable actions, because they have more access rights than they need to do their jobs. According to the [Verizon 2018 Data Breach Investigation Report](#), privileged access abuse was the third largest cause of data breaches and the second largest cause of security incidents.

5 Ways a SIEM Can Help Stop Privileged Access Abuse

01. **Unwanted activity** - monitoring and reporting on suspicious access to any sensitive data.
02. **Third-party violations** - monitoring activity by external vendors and partners who have access to organizational systems, in order to identify anomalous behavior or escalation of privileges.
03. **Departed employees** - alerting on any activity by terminated user accounts, or unexpected activity on accounts that are normally inactive.
04. **Human error** - alerting on anomalous activity that could be a disastrous human error, such as deletion of large quantities of data.
05. **Overexposure** - reporting on users who are accessing systems or data that is not within their regular usage profile.



Trusted Host and Entity Compromise

It is very common for attackers to take control of user credentials or hosts within an organizational network, and carry out attacks stealthily for months or years. According to the [Ponemon 2017 Cost of Data Breaches](#) report, the average time US companies took to detect a data breach was 206 days. So a major goal for security teams is to detect and subvert attacks quickly.

4 Ways a SIEM Can Help Detect and Stop Trusted Entity Compromise

01. **User accounts** - identifying anomalous activity, alerting about it and providing investigators the data they need to understand if a privileged user account was breached.
02. **Servers** - creating a trusted baseline of server activity, detecting deviations from this baseline and alerting security staff.
03. **Network devices** - monitoring traffic over time and detecting unusual spikes, non-trusted communication sources, insecure protocols, and other signs of malicious behavior.
04. **Anti-virus monitoring** - malware is a common entry point for host compromise. SIEMs can look at antivirus deployments broadly, reporting on events like protection disabled, antivirus removed, or status of threat updates.

Threat Hunting

Threat hunting is the practice of actively seeking out cyber threats in an organization or network. A threat hunt can be conducted on the heels of a security incident, but also proactively, to discover new and unknown attacks or breaches. According to a 2017 study by the [SANS Institute](#), 45% of organizations do threat hunting on an ad hoc or regular basis. Threat hunting requires broad access to security data from across the organization, which can be provided by a SIEM.

7 Ways SIEM Can Help with Threat Hunting

01. **Alerts from security systems** - delivering actionable alerts that provide context and data to help investigate a potential incident.
02. **Environment anomalies** - identifying anomalies in IT systems using correlations and behavioral analytics
03. **New vulnerabilities** - organizing data around a new vulnerability—timeline and systems, data and users affected.
04. **Tips from peers or the media** - searching historical data for attack patterns or signatures similar to known attacks.
05. **Threat intelligence** - combining threat intelligence with security data, to intelligently detect attacks in IT systems.
06. **Hypotheses based on known risks** - helping analysts frame a hypothesis and test it by exploring security data in the SIEM.
07. **Similar incidents** - checking if “this happened before”—searching security data for patterns similar to a current or previous security incident.

Data Exfiltration Detection

Data exfiltration happens when sensitive data is illicitly transferred outside an organization. It can happen manually, when a user transfers data over the internet or copies it to a physical device and moves it outside the premises, or automatically, as the result of malware infecting local systems.

6 Ways a SIEM Can Help Prevent Data Exfiltration

01. **Backdoors, rootkits and botnets** - detecting network traffic to command and control centers and identifying infected systems transmitting data to unauthorized parties.
02. **FTP and cloud storage** - monitoring network traffic over protocols that facilitate large data transfer, and alerting when unusual quantities or file types are being transferred, or when the target is unknown or malicious.
03. **Web applications** - monitoring usage of organizational web applications by outsiders, or inside usage of external web applications, which might involve downloads or browser access to sensitive data.
04. **Email forwarding** - detecting emails forwarded or sent to other entities other than stated recipient.
05. **Lateral movement** - data exfiltration typically involves attackers attempting to escalate privileges or accessing other IT systems, on their way to a lucrative target. SIEMs can detect lateral movement by correlating data from multiple IT systems.
06. **Mobile data security** - a SIEM can monitor data from the mobile workforce and identify anomalies that might indicate information leakage via a mobile device.

IoT Security

Many organizations are using connected devices to manage critical operations. Examples include network-connected medical equipment, industrial machinery and sensors, and power grid infrastructure. Internet of things (IoT) devices were not designed with security in mind, and many suffer from vulnerabilities. These vulnerabilities are difficult to remediate once the devices are already deployed in the field.

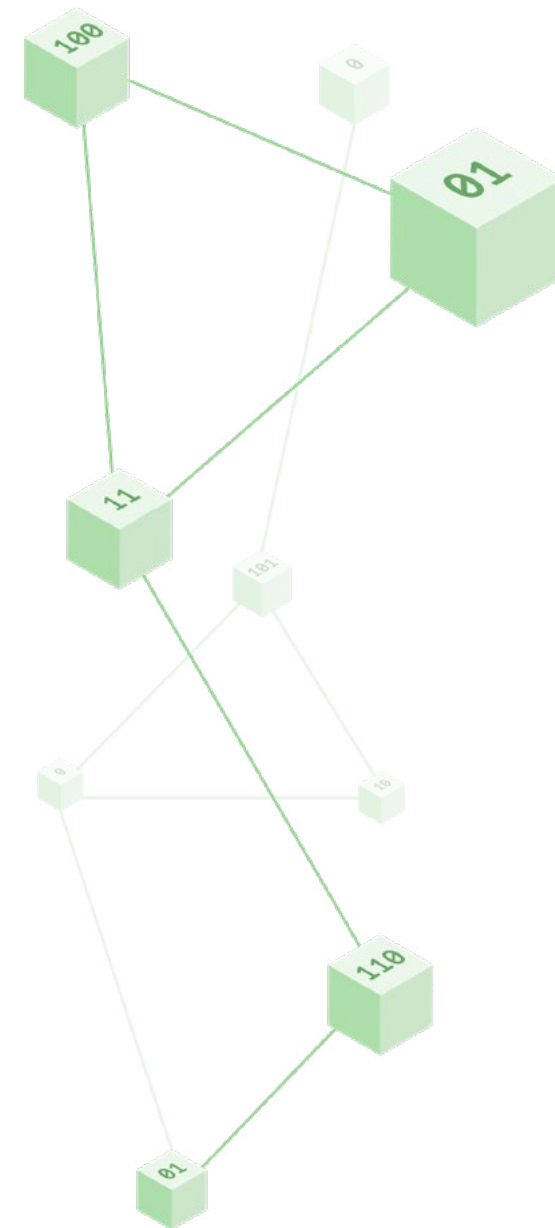
6 Ways a SIEM Can Help Mitigate IoT Threats

01. **Denial of Service (DoS) attacks** - identifying unusual traffic from organization-owned IoT devices, which might be leveraged by an attacker to perform an attack.
02. **IoT vulnerability management** - detecting old operating systems, unpatched vulnerabilities and insecure protocols on IoT devices.
03. **Access control** - monitoring who is accessing IoT devices and where they connect to, and alerting when source or target is unknown or suspicious.
04. **Data flow monitoring** - many IoT devices communicate over unencrypted protocols, most commonly NetFlow, and can be used as a vehicle to transfer sensitive data. A SIEM can monitor unusual data flows to and from IoT devices and alert security staff.
05. **Devices at risk** - identifying devices at risk due to security vulnerabilities, access to sensitive data or critical functions.
06. **Compromised devices** - identifying anomalous or suspicious behavior of IoT devices and alerting security staff that a device or fleet of devices has been compromised.

Next-Generation SIEM Technology and Advanced Use Cases

Next-generation security information and event management (SIEM) solutions, built in line with Gartner's vision of a SIEM platform integrated with advanced analytics and automation tools, can make many of these advanced use cases possible. Specifically, user and entity behavior analytics (UEBA) technology makes it possible to detect insider threats, perform more sophisticated threat hunting, prevent data exfiltration and mitigate IoT threats, even when traditional security tools don't raise a single alert. Once the devices are already deployed in the field.

Exabeam's Security Management Platform (SMP) is an example of a next-generation SIEM that comes integrated with [Advanced Analytics based on UEBA technology](#)—enabling automated detection of insider threats and mitigation of anomalous behavior that cannot be captured by traditional correlation rules. Learn more at exabeam.com/product



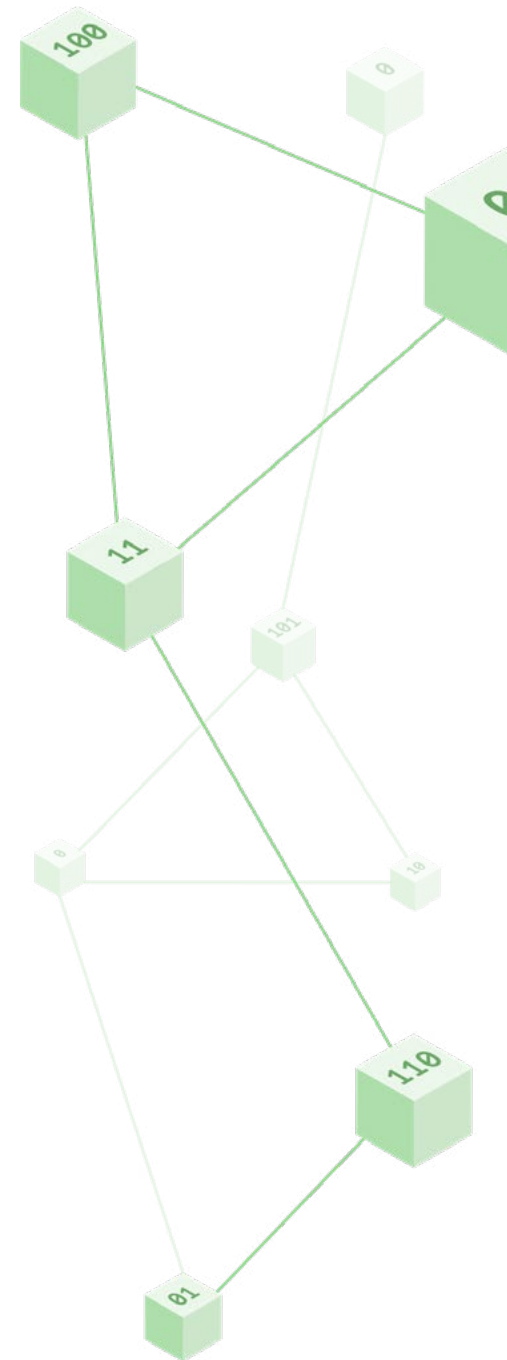
Security Big Data Analytics: Past, Present and Future

Security big data analytics (or cybersecurity analytics) is a rising force that is helping security analysts and tool vendors do much more with log and event data. In the past we were limited to manually defining correlation rules, which were brittle, hard to maintain, and resulted in many false positives.

New machine learning techniques can help security systems identify patterns and threats with no prior definitions, rules or attack signatures, and with much higher accuracy. However, to be effective, machine learning needs very big data. The challenge is storing so much more data than ever before, analyzing it in a timely manner, and extracting new insights.

In this chapter you will learn:

- [How big data analytics helps combat cyber threats](#) - both traditional and advanced analytics techniques.
- [Key concepts in big data and security](#) - including data science, machine learning, deep learning and user and entity behavior analytics (UEBA).
- [Three algorithms for detecting anomalies](#) - random forest, dimension reduction and isolation forest.
- [How SIEMs leverage big data analytics](#) - to provide new security capabilities.



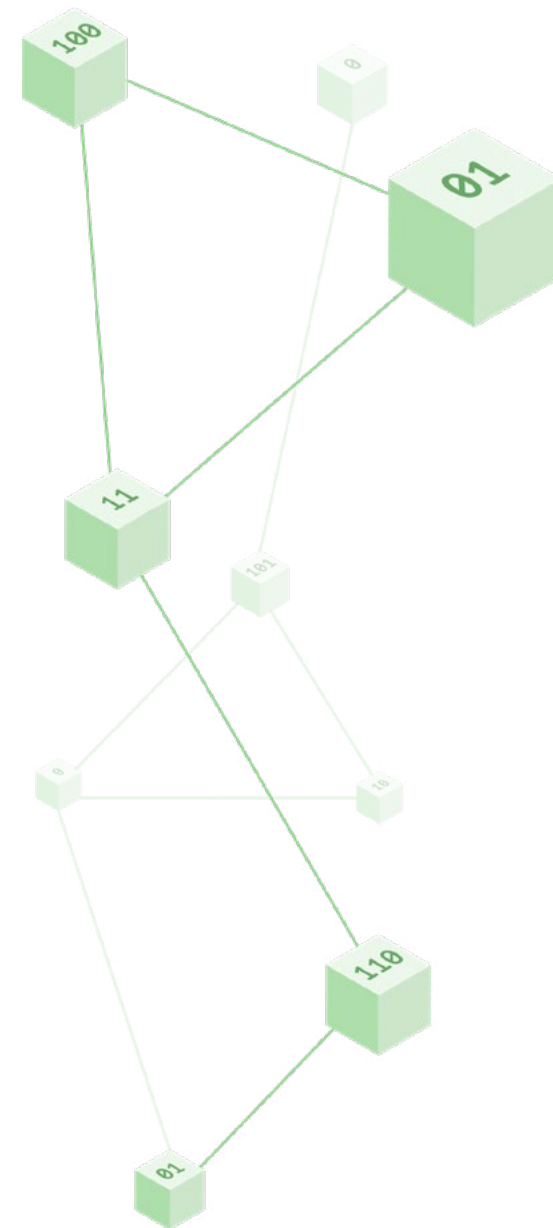
How Can Security Big Data Analytics Combat Cyber Threats?

Traditionally, security technologies used two primary analytical techniques to detect security incidents:

- **Correlation rules** - manually defined rules specifying a sequence of events that indicates an anomaly, which could represent a security threat, vulnerability or active security incident.
- **Network vulnerabilities and risk assessment** - scanning networks for known attack patterns and known vulnerabilities, such as open ports and insecure protocols.

The common denominator of these older techniques is that they are good at detecting known bad behavior. However they suffer from two key drawbacks:

- **False positives** - Because they are based on rigid, predefined rules and signatures, there is a high level of false positives, leading to alert fatigue.
- **Unexpected events** - what happens if a new type of attack is attempted that no one had created a rule for? What happens if an unknown type of malware infects your systems? Traditional systems based on correlation rules find it difficult to detect unknown threats.



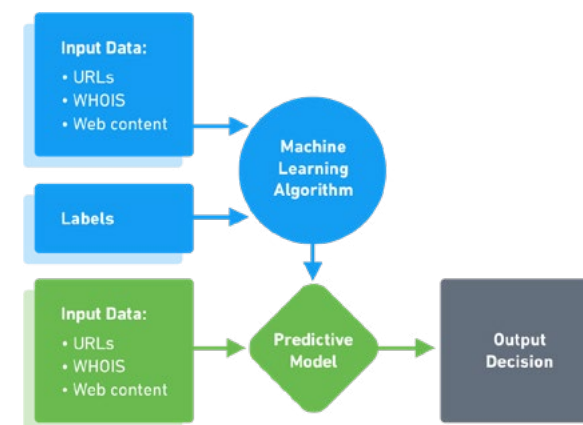
Next-gen SIEM

Advanced Threat Analytics Powered by Machine Learning

Addressing unknown risks—including insider threats, which are tricky to detect because they are users legitimately logged into corporate systems—requires advanced analytics. Advanced threat analytics technology can:

- ✓ **Identify anomalies in personnel or device behavior** - creating a model of “normal behavior” for a person, a device or group of devices on the network, and intelligently identifying anomalies, even ones that were not predefined as rules.
- ✓ **Perform machine learning-based malware detection** - intelligently analyzing binaries transmitted by email or downloaded, even if not flagged by antivirus, to understand if it is a benign program or more likely to be a malicious program.
- ✓ **Detect anomalies in the network** - creating a model of network traffic and intelligently identifying anomalies in traffic. Is something happening that is different than usual for this period or time of day?
- ✓ **Perform machine learning based intrusion detection** - identifying patterns in network traffic or access control that are similar to historic intrusions or attacks.

In order to achieve these types of analysis, new analytics methods are needed, as well as access to more data than ever before.



Supervised learning for phishing domain detection

Data Science, Machine Learning and Cybersecurity

What is Data Science?

Data science is a new discipline that leverages scientific and mathematical analysis of data sets, as well as human understanding and exploration, to derive business insights from big data.

IN THE CONTEXT OF SECURITY:

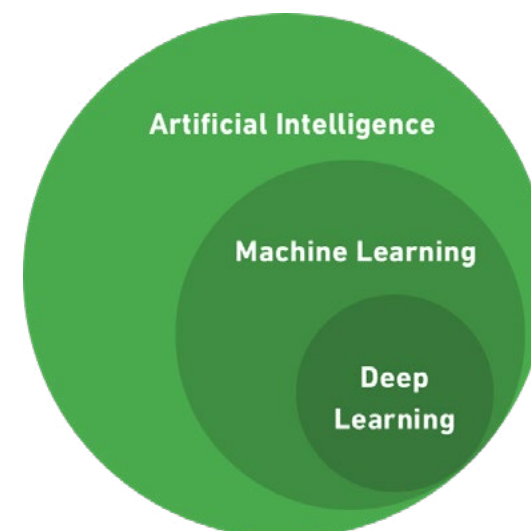
Data science is helping security analysts and security tools make better use of security data, to discover hidden patterns and better understand system behavior.

What is Machine Learning in Cyber Security?

Machine learning is part of the general field of artificial intelligence (AI). It uses statistical techniques to allow machines to learn without being explicitly programmed.

IN THE CONTEXT OF SECURITY:

Machine learning goes beyond correlation rules, to examine unknown patterns and use algorithms for prediction, classification and insight generation.

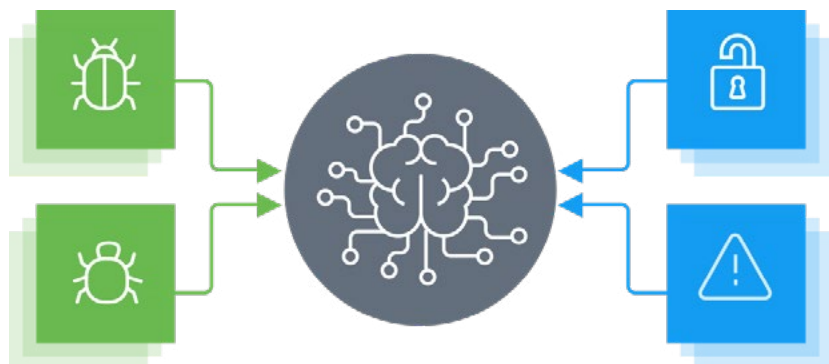


Important note

Artificial intelligence (AI) is claimed to be a part of many security analytics solutions. Don't take vendor claims for granted—check what exactly is included in the term “AI”. How are vendors building their models? Which algorithms are used? Look under the hood to understand what exactly is being offered.

Supervised vs. Unsupervised Learning

SUPERVISED MACHINE LEARNING

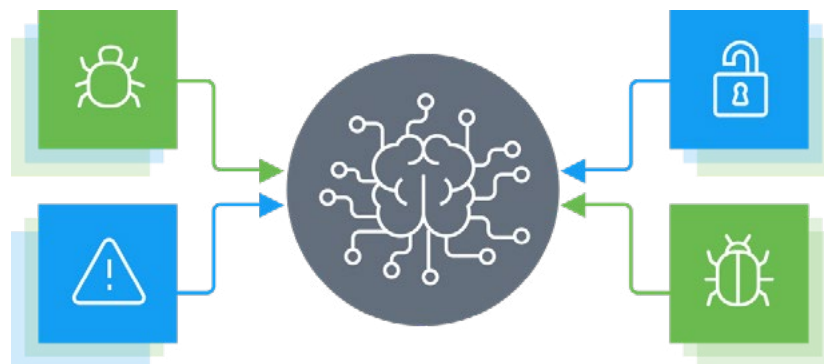


In supervised learning, the machine learns from a data set that contains inputs and known outputs. A function or model is built that makes it possible to predict what the output variables will be for new, unknown outputs.

IN THE CONTEXT OF SECURITY:

Security tools learn to analyze new behavior and determine if it is “similar to” previous known good or known bad behavior.

UNSUPERVISED MACHINE LEARNING



In unsupervised learning, the system learns from a dataset that contains only input variables. There is no correct answer, instead the algorithm is encouraged to discover new patterns in the data.

IN THE CONTEXT OF SECURITY:

Security tools use unsupervised learning to detect and act on abnormal behavior (without classifying it or understanding if it is good or bad).

What is Deep Learning in Cybersecurity?

Deep learning techniques simulate the human brain by creating networks of digital “neurons” and using them to process small pieces of data, to assemble a bigger picture. Deep learning is most commonly applied to unstructured data, and can automatically learn the significant features of data artifacts. Most modern applications of deep learning utilize supervised learning.

IN THE CONTEXT OF SECURITY:

Deep learning is primarily used in packet stream and malware binary analysis, to discover features of traffic patterns and software programs and identify malicious activity.

What is Data Mining in Cybersecurity?

Data mining is the use of analytics techniques, primarily deep learning, to uncover hidden insights in large volumes of data. For example, data mining can uncover hidden relations between entities, discover frequent sequences of events to assist prediction, and discover classification models which help group entities into useful categories.

IN THE CONTEXT OF SECURITY:

Data mining techniques is used by security tools to perform tasks like anomaly detection in very large data sets, classification of incidents or network events, and prediction of future attacks based on historic data.

What is User and Entity Behavior Analytics (UEBA)?

UEBA solutions are based on a concept called baselining. They build profiles that model standard behavior for users, hosts and devices (called entities) in an IT environment. Using primarily machine learning techniques, they identify activity that is anomalous, compared to the established baselines, and detect security incidents.

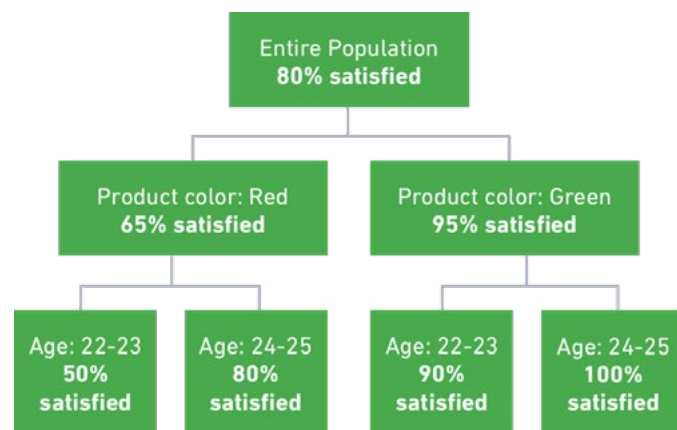
The primary advantage of UEBA over traditional security solutions is that it can detect unknown or elusive threats, such as zero day attacks and insider threats. In addition, UEBA reduces the number of false positives because it adapts and learns actual system behavior, rather than relying on predetermined rules which may not be relevant in the current context.

Algorithms for Detecting Outliers and Anomalies

Random Forest

Random forest is a powerful supervised learning algorithm that addresses the shortcomings of classic decision tree algorithms. A decision tree attempts to fit behavior into a hierarchical tree of known parameters.

For example, in the tree below customer satisfaction is distributed according to two variables, product color and customer age. A decision tree algorithm will inaccurately predict that a different color or slightly different age is a good predictor of satisfaction. This is called *overfitting*—the model uses insufficient or inaccurate data to make predictions on new data.

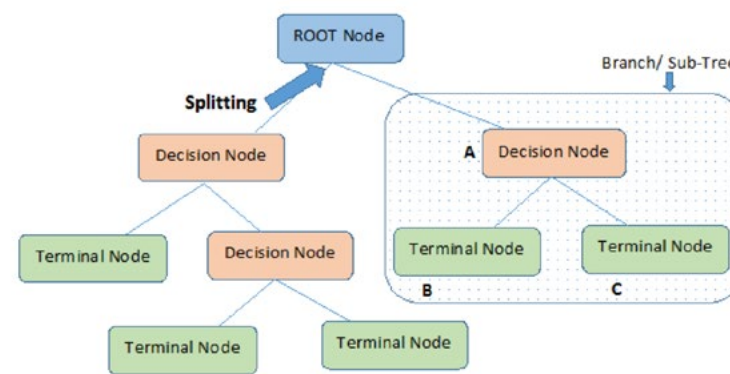


Random forest automatically breaks up decision trees into a large number of sub-trees or *stumps*. Each sub-tree emphasizes different information about the population under analysis. It then obtains the result of each sub-tree, and takes a majority vote of all the sub-trees to obtain the final result (a technique called *bagging*).

By combining all the sub-trees together, Random forest can cancel out the errors of each individual tree and dramatically improve model fitting.

IN THE CONTEXT OF SECURITY:

Random forest can help analyze sequential event paths and improve predictions about new events, even when the underlying data is insufficient or improperly structured.



Dimension Reduction

Dimension reduction is the process of converting a data set with a high number of dimensions (or parameters describing the data) to a data set with less dimensions, without losing important information.

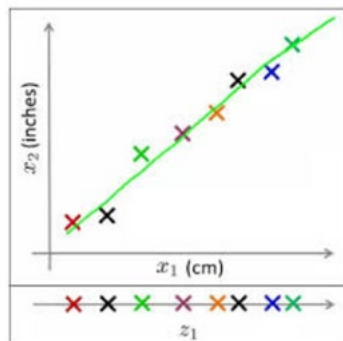
For example, if the data includes one dimension for the length of objects in centimeters and another dimension for inches, one of these dimensions is redundant and does not really add any information, as can be seen by their high correlation. Removing one of these dimensions will make the data easier to explain.

Generally speaking, a dimension reduction algorithm can determine which dimensions do not add relevant information and reduce a data set with n dimensions to k , where $k < n$.

Besides correlation analysis, other ways to remove redundant dimensions include analysis of missing values; variables with low variance across the data set; using decision trees to automatically pick the least important variables, and augmenting those trees with random forest; factor analysis; backward feature elimination (BFE); and principal component analysis (PCA).

IN THE CONTEXT OF SECURITY:

Security data typically consists of logs with a large number of data points about events in IT systems. Dimensional reduction can be used to remove the dimensions that are not necessary for answering the question at hand, helping security tools identify anomalies more accurately.



Isolation Forest

Isolation forest is a relatively new technique for detecting anomalies or outliers. It isolates data points by randomly selecting a feature of the data, then randomly selecting a value between the maximum and minimum values of that feature. The process is repeated until the feature is found to be substantially different from the rest of the data set.

The system repeats this process for a large number of features, and builds a random decision tree for each feature. An anomaly score is then computed for each feature, based on the following assumptions:

- **Features which are really anomalies** will take only a small number of isolation steps to be far off from the rest of the data set.
- **Features which are not anomalies** will take numerous isolation steps to become far off from the data set.

A threshold is defined, and features which require relatively long decision trees to become fully isolated are determined to be “normal”, with the rest determined to be “abnormal”.

IN THE CONTEXT OF SECURITY:

Isolation forest is a central technique used by UEBA and other next-gen security tools to identify data points that are anomalous compared to the surrounding data.

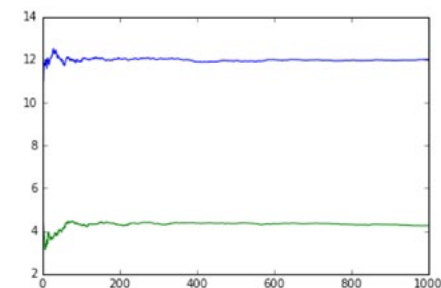


Image Source: Inside Big Data

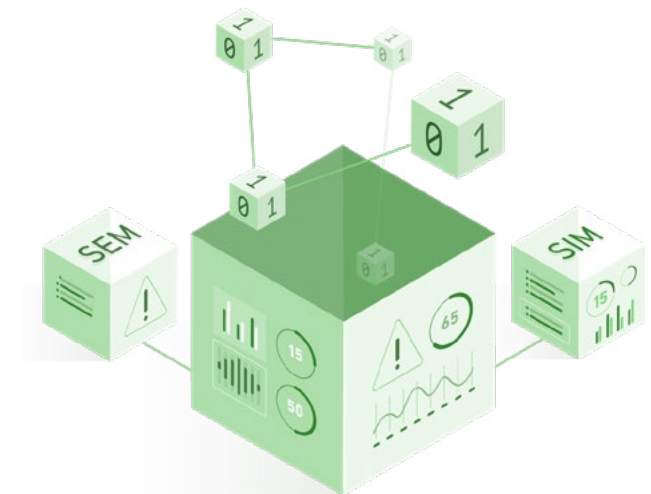
SIEM and Big Data Analytics

Security information and event management (SIEM) systems are a core component of large security organizations. They capture, organize and analyze log data and alerts from security tools across the organization. Traditionally, SIEM correlation rules were used to automatically identify and alert on security incidents.

Because SIEMs provide context on users, devices and events in virtually all IT systems across the organization, they offer ripe ground for advanced analytics techniques. Today's SIEMs either integrate with advanced analytics platforms like UEBA, or provide these capabilities as an integral part of their product.

Next-generation SIEMs can leverage machine learning, deep learning and UEBA to go beyond correlation rules and provide:

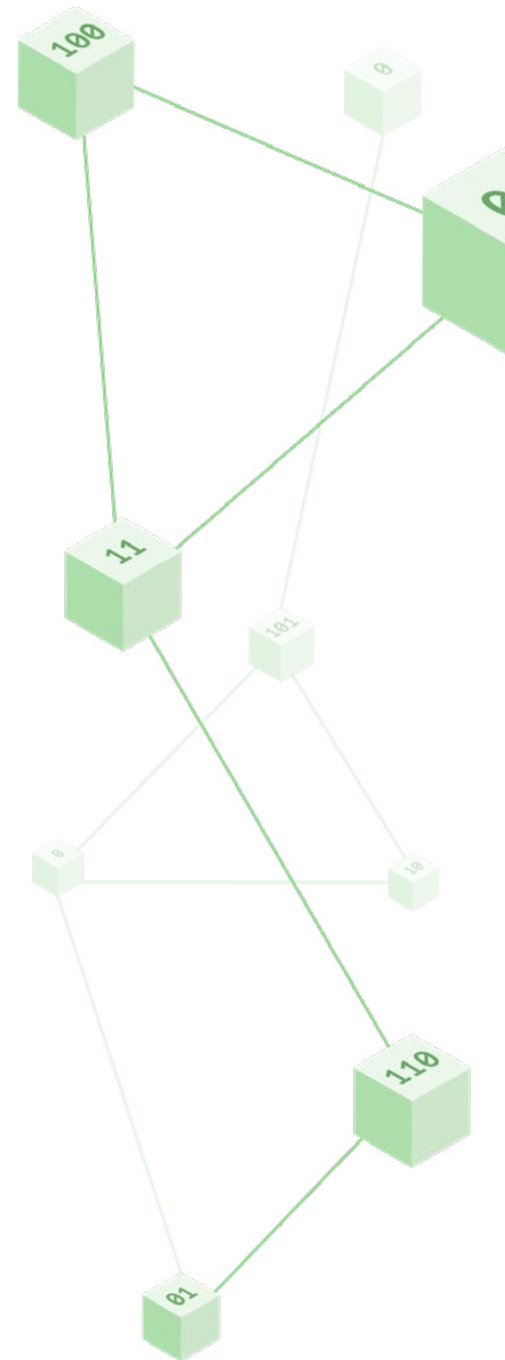
- **Complex threat identification** - modern attacks are often comprised of several types of events, each of which might appear innocuous on its own. Advanced data analytics can look at data for multiple events over a historic timeline, and capture suspicious activity.
- **Entity behavior analysis** - SIEMs can learn the normal baseline behavior of critical assets like servers, medical equipment or industrial machinery, and automatically discover anomalies that suggest a threat.
- **Lateral movement detection** - attackers who penetrate an organization typically move through a network, accessing different machines and switching credentials, to escalate their access to sensitive data. SIEMs can analyze data from across the network and multiple system resources, and use machine learning to detect lateral movement.
- **Inside threats** - SIEMs can identify that a person or system resource is behaving abnormally. They can “connect the dots” between a misbehaving user account and other data points, to discover a malicious insider, or compromise of an insider account.
- **Detection of new types of attacks** - by leveraging advanced analytics, SIEMs can capture and alert on zero day attacks, or malware which does not match a known binary pattern.



Exabeam is an example of a next-generation SIEM that comes with **advanced analytics capabilities built in**—including complex threat identification, automatic event timelines, dynamic peer grouping of similar users or entities, lateral movement detection and automatic detection of asset ownership. Learn more at exabeam.com/product

Incident Response Automation and Security Orchestration with **SOAR**

In this chapter, we explain the basics of incident response, and introduce a new category of tools—security orchestration, automation and response (SOAR)—which make incident response more efficient, more effective and more manageable at scale.



What is Incident Response?

Reactive incident response

Incident response is an organizational process that allows security teams to contain security incidents or cyber attacks, prevent or control damages. Incident response also allows teams to handle the aftermath of the attack—recovery, remediating security holes exposed by the attack, forensics, communication and auditing. This is known as reactive incident response.

Proactive incident response

Many security incidents are only discovered weeks or months after they took place—while some are never discovered. Many organizations are developing proactive incident response capabilities. This involves actively searching corporate systems for signs of a cyber attack.

Threat hunting

Threat hunting is the core activity of proactive incident response, which is carried out by skilled security analysts. It typically involves querying security data using a Security Information and Event Management (SIEM), and running vulnerability scans or penetration tests against organizational systems. The objective is to discover suspicious activity or anomalies that represent a security incident.

What is Case Management?

Case management involves collecting, distributing and analyzing data tied to specific security incidents, to allow teams to effectively respond.

Case management solutions help security staff:

- Open a case for a confirmed security incident
- Quickly aggregate all relevant data into a digital representation of the case
- Enable fast prioritization of cases for response
- Investigate and add information to the case
- Record activity in the aftermath of an attack and close the case

What is Security Orchestration, Automation and Response (SOAR)?

Security orchestration, automation and response (SOAR) is a new category of security tools defined by Gartner in their paper, [*Preparing Your Security Operations for Orchestration and Automation Tools*](#) (a departure from Gartner's previous definition the category, in 2015, as "Security Operations, Analytics and Reporting").

Gartner defines SOAR as tools that:

- Collect security threat data and alerts from different sources
- Enable incident analysis, triage and prioritization, both automatically and manually with machine assistance
- Define and enforce a standard workflow for incident response activities
- Encode incident analysis and response procedures in a digital workflow format, enabling automation of some or all incident responses

3 Key SOAR Capabilities

SOAR tools provide the following four capabilities that help security operations centers respond to incidents more effectively.



01

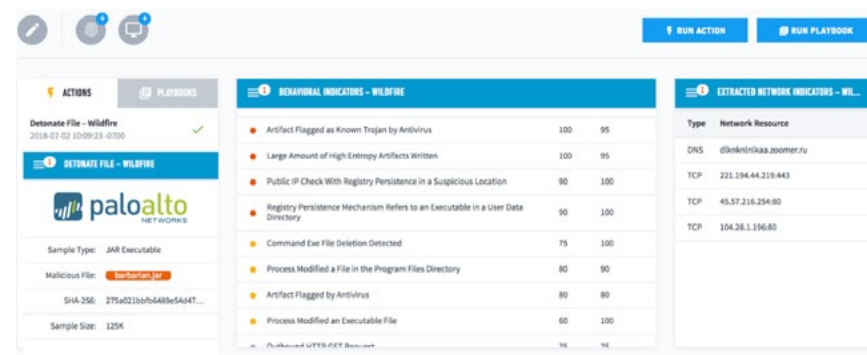
Orchestration

Orchestration is the ability to coordinate decision making, and automate responsive actions based on an assessment of risks and environment states.

SOAR tools can do this by integrating with other security solutions in a way that lets them “pull” data and also “push” proactive actions. SOAR provides a generic interface, allowing analysts to define actions on security tools and IT systems without being experts in those systems or their APIs.

AN EXAMPLE OF ORCHESTRATION: PROCESSING A SUSPICIOUS EMAIL

01. A SOAR tool can investigate whether the sender has a bad reputation, via threat intelligence, and use DNS tools to confirm the origin.
02. The tool can automatically extract hyperlinks and validate them via URL reputation, detonate the links in a secure environment, or run attachments in a sandbox.
03. Then, if an incident is confirmed, a playbook is run. The playbook looks in the email system to find all messages from the same sender or with the same links or attachments and quarantines them.



A detonating file, extracting anomalous behavioral and network indicators using Exabeam, a next-generation SIEM which includes an [Incident Responder](#) SOAR module.
Image Source: Exabeam



02

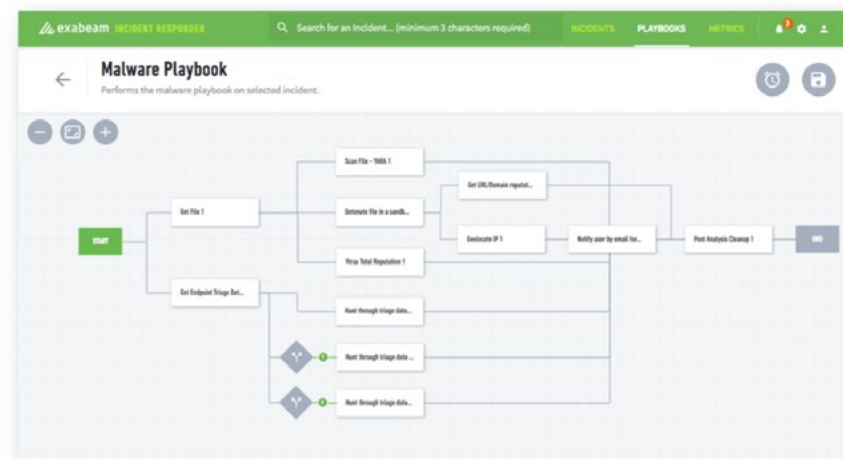
Automation

Automation is related to orchestration—it is machine-driven execution of actions on security tools and IT systems, as part of a response to an incident. SOAR tools allow security teams to define standardized automation steps and a decision-making workflow, with enforcement, status tracking and auditing capabilities.

Automation relies on security playbooks, which analysts can code using a visual UI or a programming language like Python.

AN EXAMPLE OF AN AUTOMATION PLAYBOOK: EXABEAM'S MALWARE PLAYBOOK

01. The SOAR tool scans the malware file and detonates the file in a sandbox using external services.
02. The SOAR tool checks the file against reputation services such as VirusTotal for accuracy.
03. The SOAR tool identifies the geolocation of the source or originating IP address.
04. The system notifies the user about the malware and a post-analysis cleanup is performed.



An automation playbook editor provided by Exabeam. Image Source: Exabeam



03

Incident management and collaboration

This SOAR capability helps security teams manage security incidents, collaborate and share data to resolve the incident efficiently.

Alert processing and triage

A SOAR tool gathers and analyzes security data, typically taken from the SIEM, correlates data to identify priority and criticality, and automatically generates incidents for investigation. The incident already includes relevant context information, allowing analysts to investigate further. This removes the need for a human to notice the relevant security data, identify it as a security incident and manually set up an incident in the system.

Journaling and evidentiary support

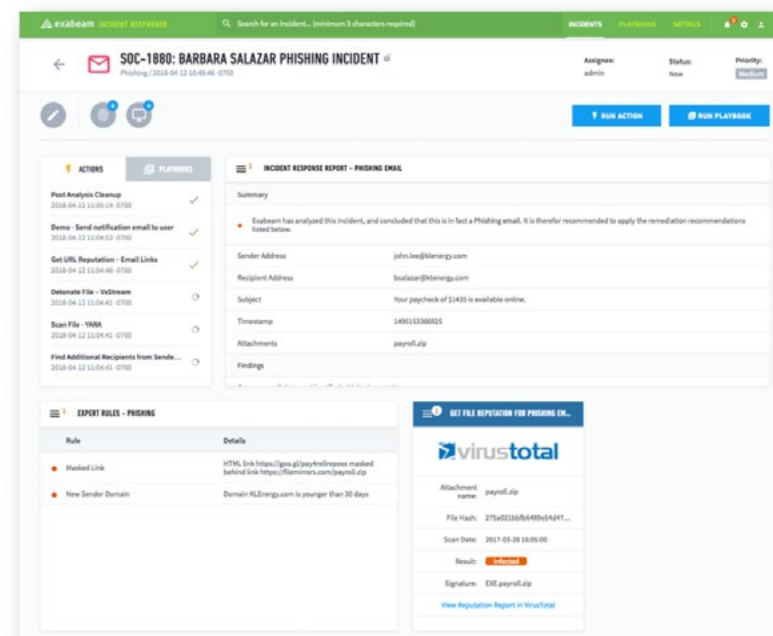
A SOAR tool provides an investigation timeline to collect and store artifacts of the security incident, for current and future analysis. Artifacts may relate to known attacker's activities, which may be carried out over an extended period. Additional artifacts can be pulled in to investigate if they are related to the ongoing incident.

Case management

The tool can record actions and decisions made by the security team, making them visible to the entire organization, as well as external auditors. Over time, the SOAR tool creates an organizational knowledge base of tribal knowledge—threats, incidents, historical responses and decisions and their outcomes.

Management of threat intelligence

A SOAR tool brings in threat data from open-source databases, industry leaders, coordinated response organizations, and commercial threat intelligence providers. The SOAR tool attaches the relevant threat information to specific incidents, and makes threat intelligence easily accessible to analysts as they are investigating an incident.



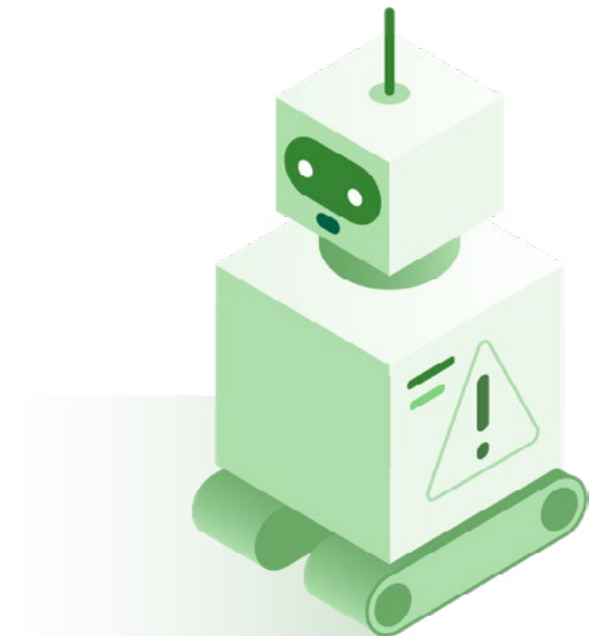
A security incident created automatically by Exabeam's Incident Responder.
Image Source: Exabeam

Dashboards and Reporting

SOAR tools are not only responsible for coordinating and automating incident response, but also for enabling central measurement of SOC activity.

SOAR tools generate reports and dashboards including:

- Analyst-level reporting on activity by each analyst, such as number and types of incidents, mean time to detect and respond per analyst, and so on.
- SOC manager reports—reporting on the number of analysts, incidents handled per analyst, and mean time for specific stages of incident response process, to identify bottlenecks.
- CISO-level reports—alignment of risks with IT metrics to see the impact of incidents on business performance and regulations; measuring efficiently by looking at MTTD and MTTR across the entire organization, and reduction of labor through automation.



How Does SOAR Fit in With SIEM?

SOAR tools work closely with SIEM, the SOC's central information system. SOAR tools leverage the integration with SIEM to:

- Receive alerts and additional security data to identify security incidents
- Draw in data required for analysts to further investigate an incident
- Assist analysts in proactive incident response and threat hunting, which relies on querying and exploring cross-organization data

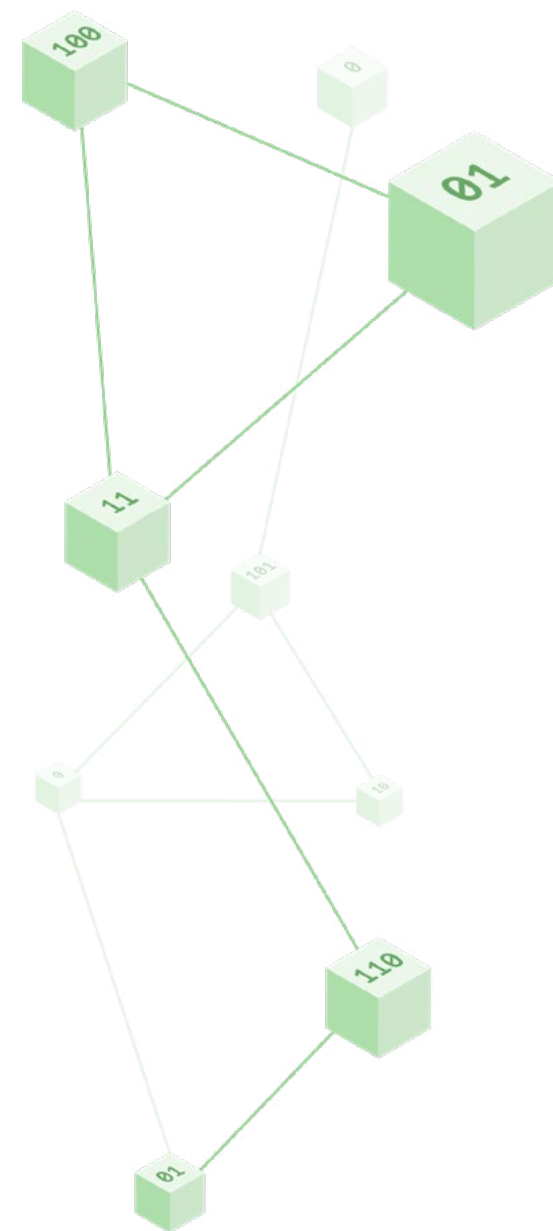
SOAR as Part of Next-Gen SIEM Solutions

According to Gartner's [Critical Capabilities for SIEM 2017](#) report, next-generation SIEM solution must include a native component that enables handling and responding to detected incidents via automated and manual case management, workflow and orchestration, as well as capabilities for advanced threat defense.

So while SOAR tools are evolving as a separate category, in Gartner's vision, SOAR should be an integrated part of the SIEM.

[Exabeam's Security Management Platform](#) is an example of this new hybrid. Exabeam is a SIEM solution based on modern data lake technology, which enables advanced analytics and user entity behavioral analytics. In addition, Exabeam comes with two components that provide full SOAR functionality:

- [Exabeam Incident Responder](#) - provides security case management, integration with third-party tools, centralized security orchestration, and automated incident response via security response playbooks.
- [Exabeam Threat Hunter](#) - a point-and-click interface that lets SOC analysts quickly perform searches to identify patterns in vast amounts of historic security data. It also provides access to complete incident timelines for past and present security incidents.

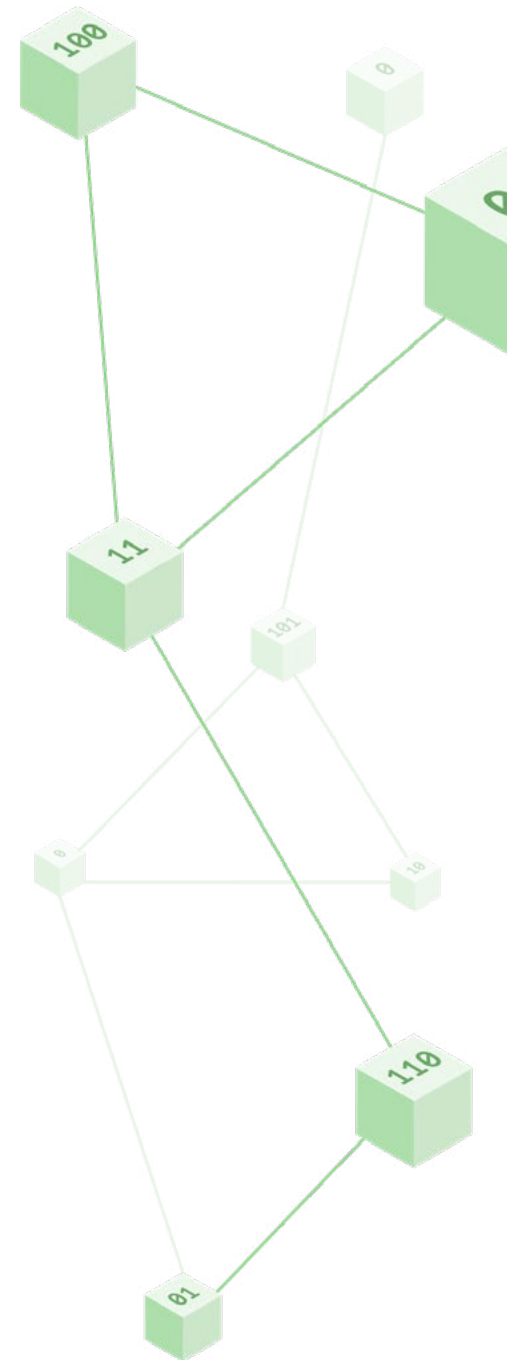


The Modern SOC, SecOps and SIEM: How They Work Together

This section is a comprehensive guide to the modern security operations center (SOC).

In this chapter you will learn:

- [What a modern SOC looks like](#) - why organizations build a SOC and their objectives
- [What is SecOps and DevSecOps](#) - how these new practices are transforming the SOC
- [SOC deployment models](#) - including new models like distributed and virtual SOC
- [SOC command hierarchy](#) - Tier 1, Tier 2, Tier 3 analysts and supporting roles
- [Technologies used in the SOC](#) - from traditional tools like SIEM, GRC and IDS, to new developments like NTA, EDR and UEBA
- [SOC processes](#) - the incident response model and how SIEMs power the basic operations of the SOC



What is a SOC?

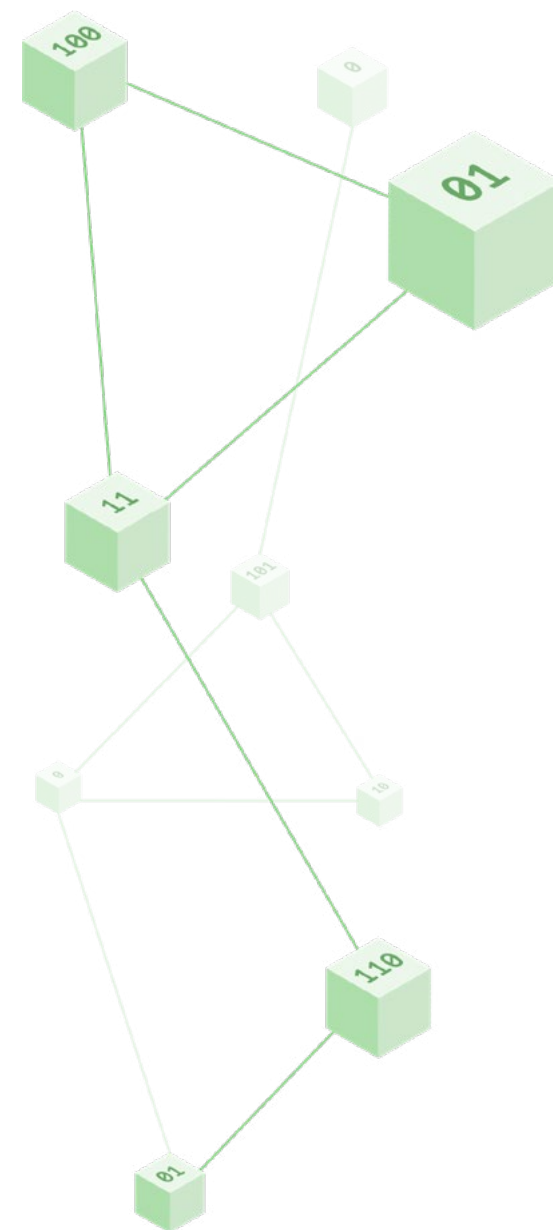
An information security operations center (ISOC or SOC) is a facility where security staff monitor enterprise systems, defend against security breaches, and proactively identify and mitigate security risks.

In the past, the SOC was considered a heavyweight infrastructure which is only within the reach of very large or security-minded organizations. Today, with new collaboration tools and security technology, many smaller organizations are setting up virtual SOC's which do not require a dedicated facility, and can use part-time staff from security, operations and development groups. Many organizations are setting up managed SOC's or hybrid SOC's which combine in-house staff with tools and expertise from managed security service providers (MSSPs).

Motivation for Building a SOC

A SOC is an advanced stage in the security maturity of an organization. The following are drivers that typically push companies to take this step:

- Requirements of standards such as the Payment Card Industry Data Security Standard (PCI DSS), government regulations, or client requirements
- The business must defend very sensitive data
- Past security breaches and/or public scrutiny
- Type of organization—for example, a government agency or Fortune 500 company will almost always have the scale and threat profile that justifies a SOC, or even multiple SOC's



Focus Areas of a SOC

A SOC can have several different functions in an organization, which can be combined. Below are SOC focus areas with the level of importance by US organizations according to the Exabeam *2018 State of the SOC* survey.

SOC Focus Area	Level of Importance in US SOC's
Control and Digital Forensics - enforcing compliance, penetration testing, vulnerability testing.	62%
Monitoring and Risk Management - capturing events from logs and security systems, identifying incidents and responding.	58%
Network and System Administration - administering security systems and processes such as identity and access management, key management, endpoint management, firewall administration, etc.	48%

SOC Facilities

The classic security operations center is a physical facility which is well protected in terms of cybersecurity and physical security. It is a large room, with security staff sitting at desks facing a wall with screens showing security stats, alerts and details of ongoing incidents. Nowadays, many SOC's look quite different. For example, a virtual SOC (VSOC) is not a physical facility, but rather a group of security professionals working together in a coordinated manner to perform the duties of a SOC.

Challenges When Building a Security Operations Center

Security teams building a SOC face several common challenges:

- **Limited visibility** - a centralized SOC does not always have access to all organizational systems. These could include endpoints, encrypted data, or systems controlled by third parties which have an impact on security.
- **White noise** - a SOC receives immense volumes of data and much of it is insignificant for security. Security information and event management (SIEM) and other tools used in the SOC are getting better at filtering out the noise, by leveraging machine learning and advanced analytics.
- **False positives and alert fatigue** - SOC systems generate large quantities of alerts, many of which turn out not to be real security incidents. False positives can consume a large part of security analysts' time, and make it more difficult to notice when real alerts occur.

All three of these challenges are addressed by a SIEM system, which powers daily operations in modern SOC's. Read more about SIEMs below in Technologies Used in the SOC.

What is SecOps?

Security operations (SecOps) is a collaboration between security and IT operations teams, where security and operations staff assume joint ownership and responsibility for security concerns. It is a set of SOC processes, practices and tools which can help organizations meet security goals more efficiently.

Before SecOps



In the past, operations and security teams had conflicting goals. Operations was responsible for setting up systems to achieve uptime and performance goals. Security was responsible for verifying a checklist of regulatory or compliance requirements, closing security holes and putting defenses in place.

In this environment, security was a burden—perceived as something that slows down operations and creates overhead. But in reality, security is part of the requirements of every IT system, just like uptime, performance or basic functionality.

Before SecOps



SecOps combines operations and security teams into one organization. Security is “shifting left”—instead of coming in at the end of the process, it is present at the beginning, when requirements are stated and systems are designed. Instead of having ops set up a system, then having security come in to secure it, systems are built from the get go with security in mind.

Towards DevSecOps

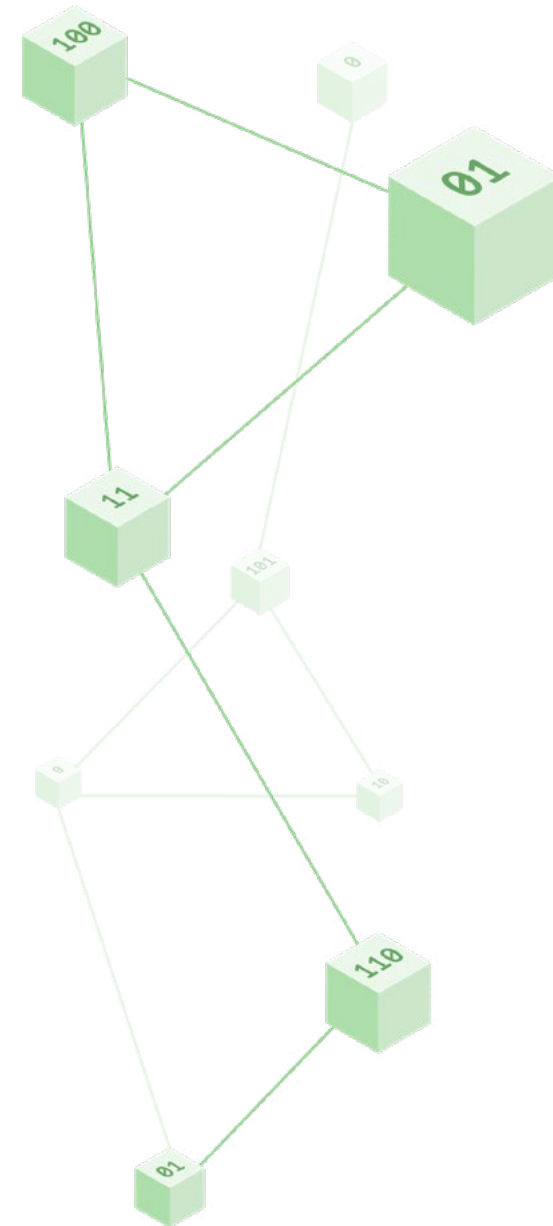


SecOps has additional implications in organizations which practice DevOps—joining development and operations teams into one group with shared responsibility for IT systems. In this environment, SecOps involves even broader cooperation—between security, ops and software development teams. This is known as DevSecOps. It shifts security even further left—baking security into systems from the first iteration of development.

SecOps in the SOC

The classic SOC is not compatible with SecOps—security analysts sit in their own room and respond to incidents, while operations are in another room, or building, running IT systems, with little or no communications between them. However, the modern SOC can foster a SecOps mentality:

- Analysts can continuously inform operations staff about threats to the organization's systems, and actual incidents
- Analysts can proactively seek out security gaps and work with operations to close them
- Operations can come to the SOC for guidance about security implications of systems, components, vendors or changes










The Security Maturity Spectrum— Are You Ready for a SOC?

Different organizations find themselves at different stages of developing their security presence. We define five stages of security maturity—in stages 4 and 5, an investment in a SOC Center becomes relevant and worthwhile.

Initial	Developing	Defined	Managed	Optimizing
1	2	3	4	5
Minimalists <i>"Security isn't our top concern. We've got AV and FWs. We're good!"</i>	Reactive <i>"We haven't explored solutions and don't believe we are at risk. We'll deal with a breach if it happens."</i>	Concerned <i>"We're at risk, but budget is a problem. We're overwhelmed by the alerts we're facing. We need help prioritizing and addressing threats."</i>	Advanced <i>"We have budget to invest in security. We have limited personnel and need to maximize them."</i>	Security Mature <i>"We're knowledgeable about security. We continuously innovate and improve our program."</i>
<ul style="list-style-type: none"> No SIEM. No logging. Basic FW at perimeter. AV in use. 	<ul style="list-style-type: none"> No SIEM. Some logging. Patch management added. Dedicated FW & DMZ. Basic identity and access management (IAM) added. 	<ul style="list-style-type: none"> Considering a SIEM or has basic SIEM deployment. Multi-FW and network segmentation added. Data classification added. Overheled by alerts and logs. Needs to prioritize them. Concerned with optimizing budget due to limited resources. 	<ul style="list-style-type: none"> SIEM is integrated with most areas. Considering analytics as a way to cut down on alert fatigue. Starting to think about tools to optimize incident investigation. Looking to increase operational efficiency and maximize personnel output. Intrigued by the idea of threat hunting. 	<ul style="list-style-type: none"> Very mature SIEM deployment. Integrated with virtually all systems. Performs threat hunting with senior analysts. Has customized security capabilities that integrate into their workflows. Capable of building their own DS algorithms. Interested in cost efficiency and reduced risk from third party solutions.

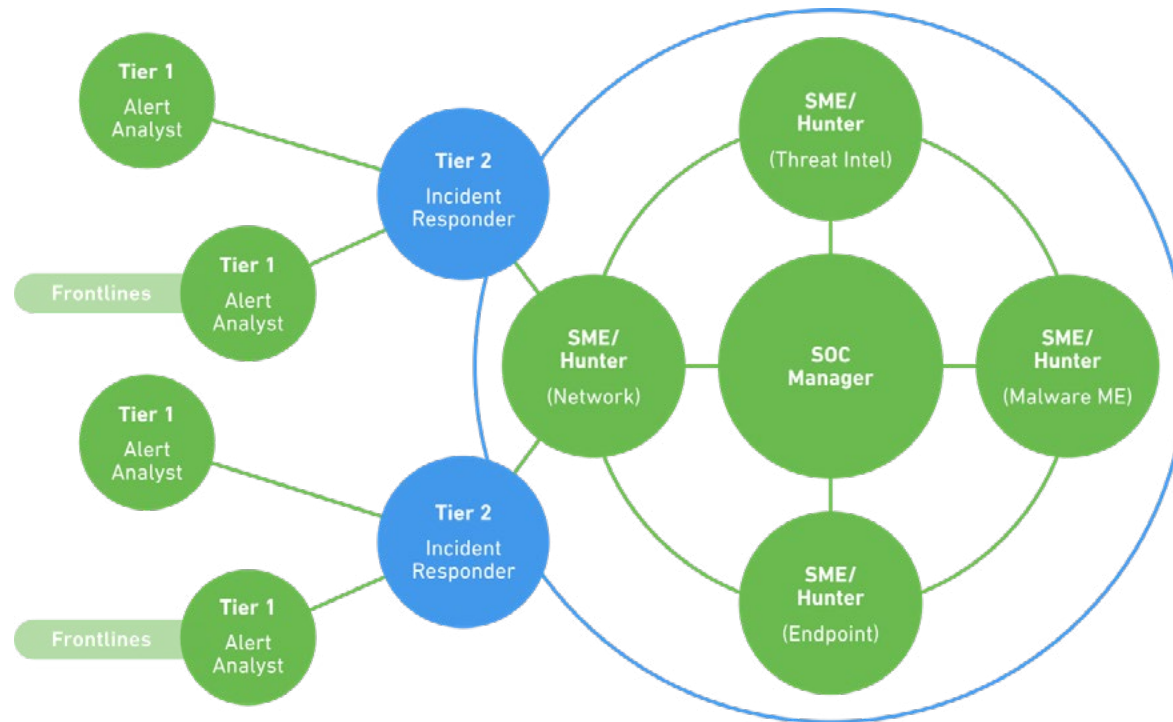
SOC Deployment Models






Following are common models for deploying a SOC within your organization:

	Reactive	Classic SOC with dedicated facility, dedicated full time staff, operated fully in house, 24x7 operations.
	Distributed SOC	Some full time staff and some part-time, typically operates 8x5 in each region.
	Multifunctional SOC / NOC	A dedicated facility with a dedicated team which performs both the functions of a network operations center (NOC) and a SOC.
	Fusion SOC	A traditional SOC combined with new functions such as threat intelligence, operational technology (OT).
	Command SOC / Global SOC	Coordinates other SOC's in a global enterprise, provides threat intelligence, situational awareness and guidance.
	Virtual SOC	No dedicated facility, part-time team members, usually reactive and activated by a high profile alert or security incident. The term virtual SOC is also sometimes used for an MSSP or managed SOC (see below).
	Managed SOC / MSSP / MDR	Many organizations are turning to managed security service providers (MSSP) to provide SOC services on an outsourced basis. Modern offerings are called managed detection and response (MDR). Managed SOC's can be outsourced completely or co-managed with in-house security staff.

Who Works in a SOC?

A SOC has a hierarchy of roles with a clear escalation path. Day-to-day alerts are received and investigated by the Tier 1 analyst; a real security incident is stepped up to a Tier 2 analyst; and business critical incidents pull in the Tier 3 analyst and if necessary, the SOC manager.



	Role	Qualifications	Duties
	Tier 1 Analyst Alert Investigator	System administration skills, web programming languages such as Python, Ruby, PHP, scripting languages, security certifications such as CISSP or SANS SEC401	Monitors SIEM alerts, manages and configures security monitoring tools. Prioritizes alerts or issues and performs triage to confirm a real security incident is taking place.
	Tier 2 Analyst Incident Responder	Similar to Tier 1 analyst but with more experience including incident response. Advanced forensics, malware assessment, threat intelligence. White hat hacker certification or training is a major advantage.	Receives incidents and performs deep analysis, correlates with threat intelligence to identify the threat actor, nature of the attack and systems or data affected. Decides on strategy for containment, remediation and recovery and acts on it.
	Tier 3 Analyst Subject Matter Expert / Threat Hunter	Similar to Tier 2 analyst but with even more experience including high-level incidents. Experience with penetration testing tools and cross-organization data visualization. Malware reverse engineering, experience identifying and developing responses to new threats and attack patterns.	Day-to-day, conducts vulnerability assessments and penetration tests, and reviews alerts, industry news, threat intelligence and security data. Actively hunts for threats that have found their way into the network, as well as unknown vulnerabilities and security gaps. When a major incident occurs, joins the Tier 2 analyst in responding and containing it.
	Tier 4 SOC Manager Commander	Similar to Tier 3 analyst, including project management skills, incident response management training, strong communication skills.	Like the commander of a military unit, responsible for hiring and training SOC staff, in charge of defensive and offensive strategy, manages resources, priorities and projects, and manages the team directly when responding to business critical security incidents. Acts as point of contact for the business for security incidents, compliance and other security
	Security Engineer Support and Infrastructure	Degree in computer science, computer engineering or information assurance, typically combined with certifications like CISSP.	A software or hardware specialist who focuses on security aspects in the design of information systems. Creates solutions and tools that help organizations deal robustly with disruption of operations or malicious attack. Sometimes employed within the SOC and sometimes supporting the SOC as part of development or operations teams.

Technologies Used in the SOC

The foundational technology of a SOC is a **security information and event management (SIEM)** system, which aggregates system logs and events from security tools from across the entire organization. The SIEM uses correlation and statistical models to identify events that might constitute a security incident, alert SOC staff about them, and provide contextual information to assist investigation. A SIEM functions as a “single pane of glass” which enables the SOC to monitor enterprise systems.

Traditional Tools Used in the SOC	Next-Gen Tools Leveraged by Advanced SOC's
<ul style="list-style-type: none">• Security information and event management (SIEM)• Governance, risk and compliance (GRC) systems• Vulnerability scanners and penetration testing tools• Intrusion detection systems (IDS), intrusion prevention systems (IPS), and wireless intrusion prevention• Firewalls and next-generation firewalls (NGFW) which can function as an IPS• Log management systems (commonly as part of the SIEM)• Cyber threat intelligence feeds and databases	<ul style="list-style-type: none">• Next-generation SIEMs which include machine learning and advanced behavioral analytics, threat hunting, built-in incident response and SOC automation• Network traffic analysis (NTA) and application performance monitoring (APM) tools• Endpoint detection and response (EDR), which helps detect and mitigate suspicious activities on hosts and user devices• User and entity behavior analytics (UEBA), which uses machine learning to identify suspicious behavioral patterns

SOC Monitoring

Monitoring is a key function of tools used in the SOC. The SOC is responsible for enterprise-wide monitoring of IT systems and user accounts, and also monitoring of the security tools themselves—for example, ensuring antivirus is installed and updated on all organizational systems. The main tool that orchestrates monitoring is the SIEM. Organizations use many dedicated monitoring tools, such as network monitoring and application performance monitoring (APM). However, for security purposes only the SIEM, with its cross-organizational view of IT and security data, can provide a complete monitoring solution.

SOC Processes Facilitated by a SIEM: Key Examples

	Malware investigation The SIEM can help security staff combine data about malware detected across the organization, correlate it with threat intelligence and help understand the systems and data affected. Next-gen SIEMs provide security orchestration capabilities, a visualization of incident timelines, and can even automatically “detonate” malware in a threat intelligence sandbox.
	Phishing prevention and detection The SIEM can use correlations and behavioral analysis to determine that a user clicked a phishing link, distributed via email or other means. When an alert is raised, analysts can search for similar patterns across the organization and across timelines to identify the full scope of the attack.
	HR investigation When an employee is suspected of direct involvement in a security incident, a SIEM can help by drawing in all data about the employee's interaction with IT systems, over long periods of time. A SIEM can uncover anomalies like logins into corporate systems at unusual hours, escalation of privileges, or moving large quantities of data.
	Departed employees risk mitigation According to a 2014 Intermedia study, The Ex-Employee Menace , 89% of employees who leave their jobs retain access to at least some corporate systems, and use those credentials to log in. A SIEM can map out the problem in a large organization, identifying which systems have unused credentials, which former employees are accessing systems, and which sensitive data is affected.

Which Tools Should You Start With?

These stages of tools adoption were proposed by Anthony Chuvakin of Gartner.

- **Greenfield SOCs** → SIEM only
- **Established SOC** → Add automated threat intelligence sandboxing, NTA and EDR.
- **Forward Leaning** → Add UEBA and a full in-house threat intelligence platform—provided as a part of next-generation SIEMs

SOC Processes

How SecOps and DevSecOps are Transforming the SOC

Security operations center processes used to be completely isolated from other parts of the organization. Developers would build systems, IT operations would run them, and security were responsible for securing them. Today it is understood that joining these three functions into one organization—with joint responsibility over security—can improve security and create major operational efficiencies.

Here are a few ways in which a SOC can integrate its processes with dev and IT:

- **Pairing threat hunters with DevOps team leaders** - instead of discovering a threat and reporting it upwards, threat hunters can work directly with dev or ops teams to close the security gap at its source.
- **Opening the SOC for guidance and advice** - anyone doing work that has a security impact should have an easy path to reach the SOC and consult with the organization's top security experts.
- **Creating a distributed SOC with DevOps members** - DevOps teams can help with incident response due to their deep knowledge of IT systems, and can learn from security staff about threats and critical vulnerabilities.
- **Creating security centers of excellence** - the SOC can work with selected dev and operations groups to implement security best practices, and then showcase these successes to the entire organization to promote SecOps practices.

A Basic Incident Response Model

While SOC's are undergoing transformation and assuming additional roles, their core activity remains incident response. The SOC is the organizational unit that is expected to detect, contain, and mitigate cyber attacks against the organization. The people responsible for incident response are Tier 1, Tier 2 and Tier 3 analysts, and the software they primarily rely on is the SOC's security information and event management (SIEM) system.

1	2	3	4	5
Event Classification Tier 1 analysts monitor user activity, network events, and signals from security tools to identify events that merit attention.	Prioritization and Investigation Tier 1 analysts prioritize, select the most important alerts, and investigate them further. Real security incidents are passed to Tier 2 analysts.	Containment and Recovery Once a security incident has been identified, the race is on to gather more data, identify the source of the attack, contain it, recover data and restore system operations.	Remediation and Mitigation SOC staff work to identify broad security gaps related to the attack and plan mitigation steps to prevent additional attacks.	Assessment and Audit SOC staff assess the attack and mitigation steps, gather additional forensic data, draw final conclusions and recommendations, and finalize auditing and documentation.
A SIEM is a foundational technology in a SOC—here is how a SIEM can help with each incident response stage:				
Alert generation and ticketing A SIEM collects security data from organizational systems and security tools, correlates it with other events or threat data, and generates alerts for suspicious or anomalous events.	Searching and exploring data A SIEM can help Tier 1 and Tier 2 analysts search, filter, slice and dice, and visualize years of security data. Analysts can easily pull and compare relevant data to better understand an incident.	Context on incidents and security orchestration When a real security incident is identified, a SIEM provides context around the incident—for example, which other systems were accessed by the same IPs or user credentials.	Reporting and dashboarding Remediation and mitigation are an ongoing activity, and they require visibility of the status and activity of critical security and IT systems. SIEMs have a cross-organization view which can provide this visibility.	Compliance reporting One of the core functions of a SIEM is to produce reports and audits for regulatory requirements and standards like PCI DSS, HIPAA and SOX—both on an ongoing basis and following an incident or breach.
<i>Next-gen SIEM</i> Next-generation SIEMs leverage machine learning and behavioral analytics to reduce false positives and alert fatigue, and discover hard-to-detect complex events like lateral movement, insider threats and data exfiltration.	<i>Next-gen SIEM</i> Next-generation SIEMs are based on data lake technology that allows organizations to store unlimited data at low cost. They also leverage machine learning and user and entity behavior analytics (UEBA) to easily identify high risk events and surface them to analysts.	<i>Next-gen SIEM</i> Next-generation SIEMs provide security orchestration, automation and response (SOAR) capabilities. They integrate with other security systems and can automatically perform containment actions.	<i>Next-gen SIEM</i> Next-generation SIEMs leverage machine learning and data science capabilities that establish smart baselines for groups of users and devices. This allows faster and more accurate detection of insecure systems or suspicious activity.	

Measuring the SOC

Here are a few important metrics that can help you understand the scale of activity in the SOC, and how effectively analysts are handling the workload.

Metric	Definition	What it Measures
Mean time to detection (MTTD)	Average time the SOC takes to detect an incident	How effective the SOC is at processing important alerts and identifying real incidents
Mean time to resolution (MTTR)	Average time that transpires until the SOC takes action and neutralizes the threat	How effective the SOC is at gathering relevant data, coordinating a response and taking action
Total cases per month	Number of security incidents detected and processed by the SOC	How busy the security environment is and the scale of action the SOC is managing
Types of cases	Number of incidents by type—web attack, attrition (brute force and destruction), email, loss or theft of equipment, etc.	The main types of activity managed by the SOC and where security preventative measures should be focused
Analyst productivity	Number of units processed per analyst—alerts for Tier 1, incidents for Tier 2, threats discovered for Tier 3	How effective analysts are at covering maximum possible alerts and threats
Case escalation breakdown	Number of events that enter the SIEM, alerts reported, suspected incidents, confirmed incidents, escalated incidents	The effective capacity of the SOC at each level and the workload expected for different analyst groups

The Future of the SOC

The SOC is undergoing an exciting transformation. It is integrating with ops and development departments, and is empowered by powerful new technologies, while retaining its traditional command structure and roles—to identify and respond to critical security incidents.

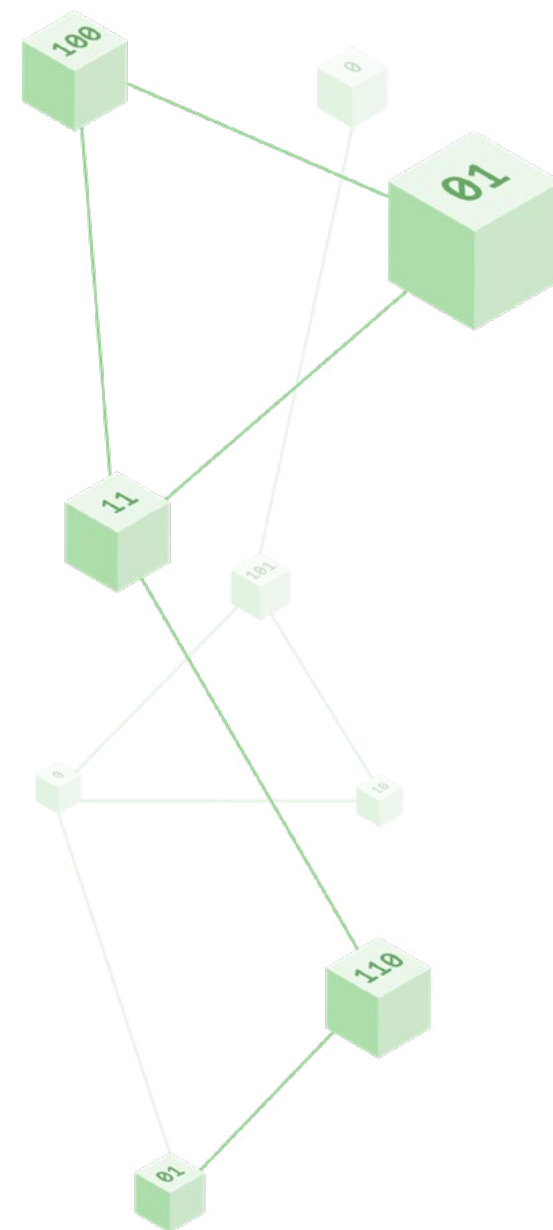
We showed how SIEM is a foundational technology of the SOC, and how next-generation SIEMs, which include new capabilities like behavioral analytics, machine learning and SOC automation, open up new possibilities for security analysts.

The impact of a next-gen SIEM on the SOC can be significant:

- Reduce alert fatigue—via user and entity behavior analytics (UEBA) that goes beyond correlation rules, helps reduce false positives and discover hidden threats.
- Improve MTTD—by helping analysts discover incidents faster and gather all relevant data.
- Improve MTTR—by integrating with security systems and leveraging security orchestration, automation and response (SOAR) technology.
- Enable threat hunting—by giving analysts fast and easy access and powerful exploration of unlimited volumes of security data.

Exabeam is an example of a [next-generation SIEM](#) which combines data lake technology, visibility into cloud infrastructure, behavioral analytics, an automated incident responder and a threat hunting module with powerful data querying and visualization.

Learn more at exabeam.com/product



Evaluating and Selecting SIEM Tools - A Buyer's Guide

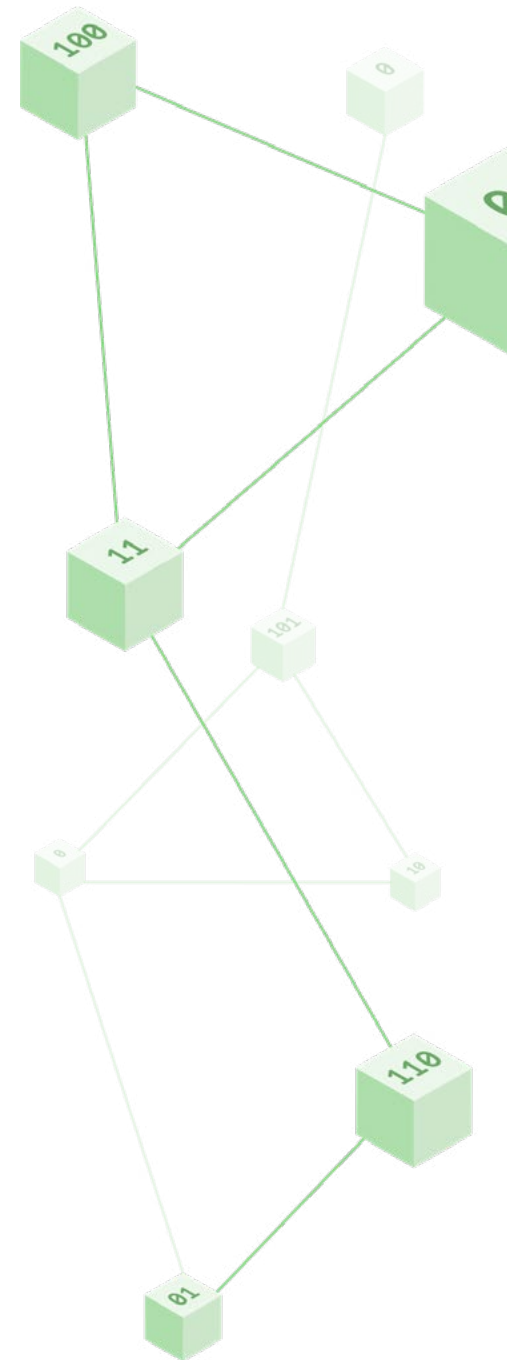
A security information and event management (SIEM) system is a foundation of the modern security operations center (SOC). It collects logs and events from security tools and IT systems across the enterprise, parses the data and uses threat intelligence, rules and analytics to identify security incidents.

Selecting, purchasing and implementing a SIEM is no small task. SIEMs were traditionally very expensive systems, both from a licensing and a hardware perspective. They also have high operating expenses, as they require trained security staff to interpret alerts and identify security incidents.

Modern SIEMs have reduced both types of costs, by offering cloud-hosted and SIEM as a service, by leveraging advanced low cost storage technology, and by making it much easier for analysts to sift through, interpret and operationalize SIEM data.

In this buyer's guide to modern SIEM solutions you will learn:

- Needs, use cases and required capabilities of SIEM products - 5 core areas in which a SIEM can help your organization
- What does a next-generation SIEM include? Next-gen components in the new, expanded SIEM model proposed by Gartner, including UEBA and SOAR
- SOC pains and required capabilities of modern SIEM solutions - top 4 issues experienced in the modern SOC and how SIEMs can alleviate them
- SIEM comparisons - open source vs. commercial vs. home grown and in-house vs. managed
- SIEM total cost of ownership - licensing models, hardware costs and sizing, storage costs, and in-house analyst costs
- Compliance and security considerations



Needs, Use Cases and Required Capabilities of SIEM Products



User Monitoring

The need: Monitoring user activity within and outside the network

Capabilities

- Monitoring user activity
- Privileged user monitoring
- Baselineing user activity and identifying anomalies



Threat Detection

The need: Specialized tools to monitor, analyze and detect threats

Capabilities

- Detecting known attack patterns, signatures, and correlations indicating an attack.
- Detect unknown attack chains via machine learning and advanced analytics.



Security Analysis

The need: Deriving more insights into security data from multiple sources

Capabilities

- Statistical analysis and correlation rules
- Machine learning to establish baselines of normal activity and detect anomalies



Incident Management

The need: An organized way to address and manage security incidents and alerts

Capabilities

- Incident prioritization —understanding which incidents are particularly abnormal or dangerous
- Automated collection of evidence for investigators
- Automated response



Compliance and Security Reporting

The need: Automatically verifying regulatory requirements and generating audit reports, managing data privacy and governance

Capabilities

- PCI DSS compliance
- HIPAA compliance
- SOX compliance
- GDPR compliance
- Other standards and regulations

What Does a Next-Generation SIEM Include?

According to Gartner's [Critical Capabilities for SIEM 2017](#) report, next-generation SIEMs must incorporate additional technologies alongside the traditional log management, statistical analysis, alerting and reporting capabilities. New SIEMs must include:

- **User and entity behavior analytics (UEBA)** - technology that models standard behavior for users, endpoints and network devices, establishing a baseline and intelligently identifying anomalies, via advanced analytics and machine learning techniques.
- **Security orchestration, automation and response (SOAR)** - technology that collects security data, prioritizes incidents, and encodes incident response in a digital workflow format, enabling automation of some or all incident response stages.

Next-gen SIEM Components

Threat intelligence

Combines internal data with third-party threat intelligence feeds on threats and vulnerabilities.

Data aggregation

Collects and aggregates data from security systems and network devices.

Correlation, security monitoring and alerts

Links events and related data into security incidents, threats or forensic findings, analyzes events and sends alerts to notify security staff of immediate issues.

Advanced analytics

Uses statistical models and machine learning to identify anomalies and detect advanced threats, detect unknown threats, detect lateral movements within a network, and enrich the context of security alerts to make it easier to investigate and detect elusive threats.

Dashboards

Creates visualizations to let staff review event data, identify patterns and anomalies

Search, data exploration and reporting

Search vast amounts of security data without reviewing raw data and without data science expertise, active explore data to discover patterns and hunt for threats, create and schedule reports on important data points.

Compliance

Gathers log data for standards like HIPAA, PCI/DSS, HITECH, SOX and GDPR and generates compliance reports. Helps to meet compliance and security regulations requirements, for example by alerting about security conditions for protected data.

Retention

Stores long-term historical data, useful for compliance and forensic investigations. Built in data lake technology facilitate unlimited, low cost, long-term storage.

Forensic analysis

Enables exploration of log and event data to discover details of a security incident, with automated attachment of additional evidence organized in a situation timeline.

Threat hunting

Enables security staff to run queries on log and event data, and freely explore data to proactively uncover threats. Once a threat is discovered, automatically pulls in relevant evidence for investigation.

Incident response support





Helps security teams identify and respond to security incidents automatically, bringing in all relevant data rapidly and providing decision support.

SOC automation

Automatically responds to incidents but automating and orchestrating security systems, known as security orchestration, automation and response (SOAR).

SOC Pains and Required Capabilities of Modern SIEM Solutions

SIEMs have been a fundamental infrastructure of the security operations center (SOC) for over two decades. However, SOC analysts experience several pains that traditional SIEMs can't solve. Below we show how next-generation SIEM technology can solve these pains.

	Security Operations Center Pain Points	How a Next-Generation SIEM Can Help
	Alert fatigue – too many alerts to review and incidents to investigate	<p>Next-generation SIEMs generate less alerts for security analysts to review [explain how facilitated by behavioral profiling]</p> <p>They can also categorize incidents based on risk reasons, incident type, and priority, enabling analysts to filter and prioritize based on risk scores and business impact.</p>
	Analyst fatigue due to mundane tasks	<p>Next-generation SIEMs provide security automation via playbooks. They integrate with IT systems and security tools to automatically attach more evidence to incidents and aid investigators in closing incidents quickly.</p>
	Takes too long to investigate incidents	<p>Next-generation SIEMs create a timeline that pulls together all the evidence related to a specific incident, across multiple users and organizational systems. This allows analysts to view the entire scope of an incident and its potential risks on one pane of glass.</p>
	Lack of skilled analysts – need to quickly train and onboard new staff	<p>Next-generation SIEMs are easier to use and provide easy to interpret incidents packaged with ancillary information. They also allow easy ways to query and explore security data, without requiring analysts to become SQL experts or data scientists. This allows even junior analysts to assess risks, prioritize incidents, and push confirmed incidents to a next-tier analyst.</p>

SIEM Comparisons

There are three main options for procuring a SIEM platform. Following are some of the pros and cons.

01

Building an open source SIEM

Open source tools such as OSSIM, OSSEC and Apache Metron can provide many SIEM capabilities including event collection, processing, correlation and alerting. Some open source solutions also provide intrusion detection system (IDS) capabilities.

Pros

No upfront expense, simpler to implement an open source SIEM than traditional SIEM solutions.

Cons

Ongoing maintenance costs can outweigh the saving in license costs. Open source SIEMs are not fully featured, mainly suitable for smaller deployments. No next-gen SIEM features.

02

Leveraging a commercial SIEM tool

Traditional SIEM tools from players like HPE, IBM and McAfee (now Intel Security) were the common choice of large organizations building a SOC to centralize security activity and incident response.

In recent years, new lightweight SIEM solutions have emerged, which are powerful, less expensive and much faster to implement. Three of these solutions have been featured in the Gartner [*SIEM Magic Quadrant 2018*](#): [Exabeam](#), Rapid7 and Securonix.

Pros

Enterprise grade, proven technology, most products have at least some next-gen SIEM capabilities.

Cons

License costs, SOC procedures are built around a specific solution's processes, leading to vendor lock in.

03

Building a SIEM platform in-house

The ELK stack – Elasticsearch, Logstash and Kibana – is a great starting point for building your own SIEM solution. In fact, most of the new contenders in the SIEM market are based on this stack. Can be suitable for very large organizations who need tailored capabilities, and want to integrate with previous investments in threat intelligence, monitoring or analytics.

Pros

Complete customization of all SIEM capabilities, easier integration with legacy systems and in-house security feeds.

Cons

Very high upfront expense and an ongoing development cost to support changes and maintenance.

In-House vs. Managed SIEM

SIEMs provided and managed by managed security service providers (MSSP) are a growing trend. Managed SIEMs are making it possible for smaller organizations, which do not have sufficient full time security staff, to enter the SIEM game.

There are four common SIEM hosting models:

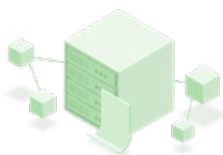
- **Self-hosted, Self-managed** - SIEM is purchased or build, then hosted in local data center and run by dedicated security staff.
- **Self-hosted, Hybrid-managed** - a SIEM is deployed in-house, typically a legacy investment, and run together by local security staff and MSSP experts.
- **Cloud SIEM, Self-managed** - a SIEM is run by an MSSP but ongoing security operations managed by in-house staff.
- **SIEM as a service** - the SIEM runs in the cloud, including data storage, with local security staff managing security processes leveraging SIEM data.

Evaluating SIEM Total Cost of Ownership

Procuring a SIEM involves several different costs, some of which are capital expenditures and some are operating expenditures.

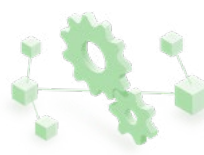
A SIEM Costing Model

SIEM CAPEX budget items



- Licenses
- Development and integration
- Training
- Hardware and storage equipment (for on-premises SIEM)
- Periodic scaling up of hardware or storage equipment (for on-premises SIEM)

SIEM OPEX budget items



- Dedicated/outsourced security analysts
- IT maintenance and resource provisioning (for on-premises SIEM)
- Ongoing integration with new organizational systems
- Cloud storage and cloud-based compute resources (for hosted SIEM)

Licensing Models

Your SIEM solution will likely use one of these three license models:

- **Volume licensing** – payment based on number of messages per second, events ingested, etc. For large organizations can drive up license costs significantly.
- **User-based pricing** – SIEM is priced based on number of “seats” without respect to the volume of data. In most organizations this will provide the lowest cost, even as data volumes grow, and without respect to the amount of historic data retained.
- **SaaS pricing** – SIEM is paid as a subscription based on actual usage.

Hardware Costs and Sizing

Organizations adopting SIEM on-premises will have to provision hardware to run the SIEM. To determine how much hardware is needed, you should first estimate the number of events the SIEM needs to handle.



The number and type of servers needed is defined by event volume, as well as the storage format, your decision whether to store data locally or in the cloud, the ratio of log compression, encryption requirements, and quantity of short-term data vs. long-term data retention.

Storage Costs and Sizing

The same calculation of events per day can be used to determine the SIEM's storage requirements. The cost of storage will depend on your SIEM deployment model:

- **For on-premises SIEM**, you will either need to setup storage infrastructure independently, and scale it up as data volumes grow. Or purchase an appliance from the SIEM vendor, but when you scale beyond the appliance's capacity, you'll have to manage storage yourself.
- **For cloud-based SIEM**, there is typically a charge per data volume, to compensate the vendor for the cost of cloud storage. Check for how many days the SIEM permits you to retain data, and if there are additional retention costs.

Number of In-House Analysts

A SIEM is not valuable without security analysts who can receive and act upon its alerts. Security analysts are needed to:

- Review alerts and decide which are actual security incidents
- Investigate incidents by pulling together relevant information, and escalating to a higher-tier analyst for action
- Analysts must be skilled and trained, preferably with a relevant security certification.

If you use a managed security service provider (MSSP), analysts will be outsourced by the service provider, or work will be divided between in-house staff and external suppliers.

If you use a SIEM with AI or machine learning capabilities, the tool's intelligence is not a substitute for human analysts. You will still need analysts, but effective AI processing of security data can substantially reduce false positives, help security analysts get the data they need faster, substantially reducing analyst labor.

Compliance and Security Considerations

When purchasing a SIEM, consider which standards or regulations your organization as a whole needs to comply with (across all departments—because SIEM is a cross-organizational infrastructure).

Important to consider:

- Review the standard and map out sections which might be related to SIEM capabilities—for example, logging and reporting on failed login attempts.
- Ensure your SIEM deployment and customization makes it easy to fulfill these requirements.
- Check which compliance and audit reports you need to submit which the SIEM can generate automatically—and whether your SIEM solution of choice can generate them.
- What are the compliance requirements on the SIEM itself, e.g., which data can be saved according to GDPR?
- How will the SIEM be secured—different depending on deployment model—on-premises, cloud, MSSP.

Exabeam - Next Generation SIEM with Unlimited Storage, Advanced Analytics and Automated Incident Response

Exabeam provides a next-generation Security Management Platform, a modern SIEM that combines end-to-end data collection, analysis, monitoring, threat detection and automated response in a single management and operations platform.

Exabeam Next-Gen SIEM Capabilities

Exabeam is a modern SIEM platform that provides all the next-gen SIEM capabilities defined in Gartner's model:

Advanced Analytics and Forensic Analysis

Exabeam provides threat identification with behavioral analysis based on machine learning. It creates behavioral baselines and intelligently identifies anomalies. Exabeam can also dynamically group peers of entities to identify suspicious individuals, and detect lateral movement across different computer systems and user accounts. To enable forensic analysis, Exabeam automatically collects all the evidence related to an incident and constructs timelines to visualize security incidents.

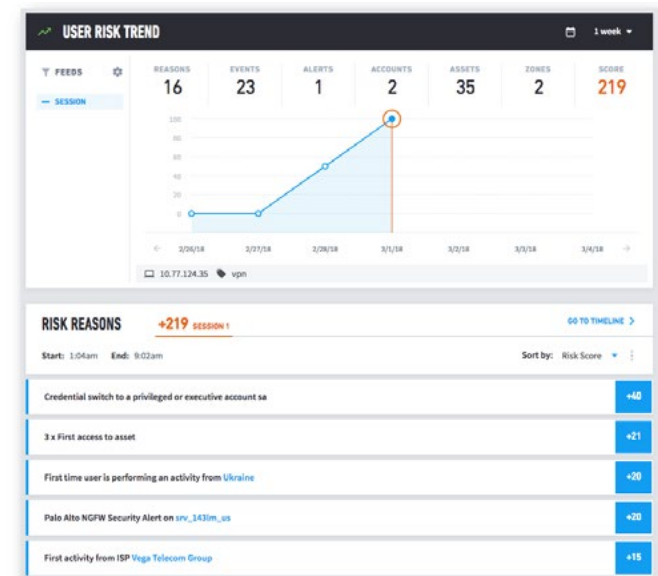


Image Source: Exabeam

Data Exploration, Reporting and Retention

Exabeam provides unlimited log data retention with flat pricing, leveraging modern data lake technology. It provides context-aware log parsing and presentation that helps security analysts quickly find what they need, and makes it possible to build rules and queries using natural language. Analysts can quickly explore, slice and dice security data without requiring expert knowledge of data science or SQL. Like traditional SIEM platforms, Exabeam also provides prebuild compliance reports for PCI-DSS, SOX, GDPR, and more.



Image Source: Exabeam

Incident Response and SOC Automation

Exabeam provides customizable case management designed for security incidents. It provides a centralized approach to incident response, gathering data from hundreds of tools and orchestrating a response to different types of incidents, via tools like email servers, active directory and firewalls, using security playbooks. Playbooks can automate investigations, containment, and mitigation.

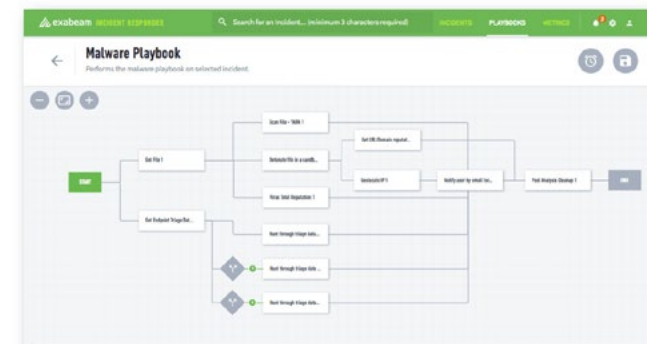


Image Source: Exabeam

Threat Hunting

Exabeam provides a point-and-click interface that lets anyone in the SOC easily create complex queries on security data. When an incident is identified, it collects evidence and organizes it into a complete incident timeline. Analysts can also enter an Alert ID from an anti malware or DLP tool, and immediately view a timeline of all related security events.

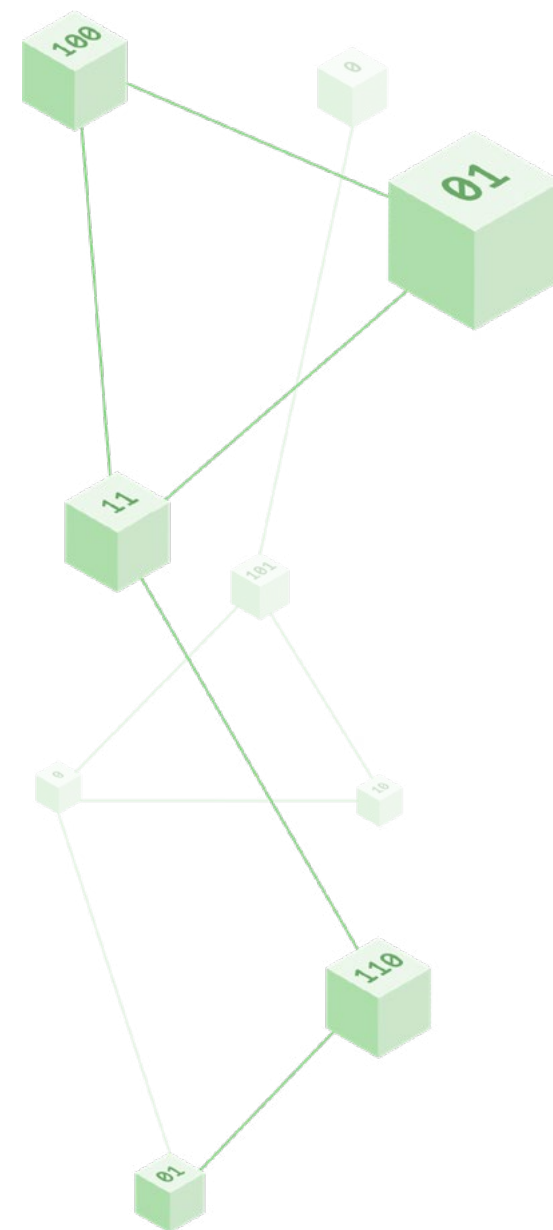


Image Source: Exabeam





Exabeam SIEM: Addressing SOC Pain Points

Exabeam's next-generation capabilities can help resolve the most common pain points in the modern SOC:

- **Alert fatigue** – Exabeam reduces fatigue by leveraging UEBA technology. It focuses analysts on alerts that represent anomalies, compared to behavioral baselines of users and network entities, and helps prioritize incidents based on organizational context.
- **Analyst fatigue due to mundane tasks** – Exabeam integrates with security tools and executes automated security playbooks when specific types of security incidents occur.
- **Takes too long to investigate incidents** – Exabeam automatically pulls in all evidence relevant to a security incident, and lays it out on an incident timeline. This provides an instant look of the incident, across multiple IT systems, users and credentials.
- **Lack of skilled analysts** – Exabeam automatically prioritizes incidents via behavioral analysis, and allows users to construct complex queries on security data using a simple drag-and-drop interface, with no need for data science or SQL expertise. This allows even junior analysts to identify important incidents and conduct in-depth investigation.



Exabeam SIEM Total Cost of Ownership

	Traditional SIEM	Exabeam
 Licensing Model	Based on event volumes	Based on seats—flat pricing for large data volumes
 Hardware Costs	Server costs should be estimated based on estimated an events per day model.	Exabeam server costs are similar to that of a traditional SIEM for the same event volume.
 Storage Costs	Storage costs grow with data volume—whether storage is on-premise or in the cloud. Usually additional costs per retention period.	Flat cost of storage included in the SIEM seat price
 In-House Analysts	Dedicated, expert-level in-house security staff needed to interpret and investigate SIEM alerts	Prioritized, friendly SIEM alerts with automated incident timelines—allows managing the SIEM with part-time or junior security analysts.

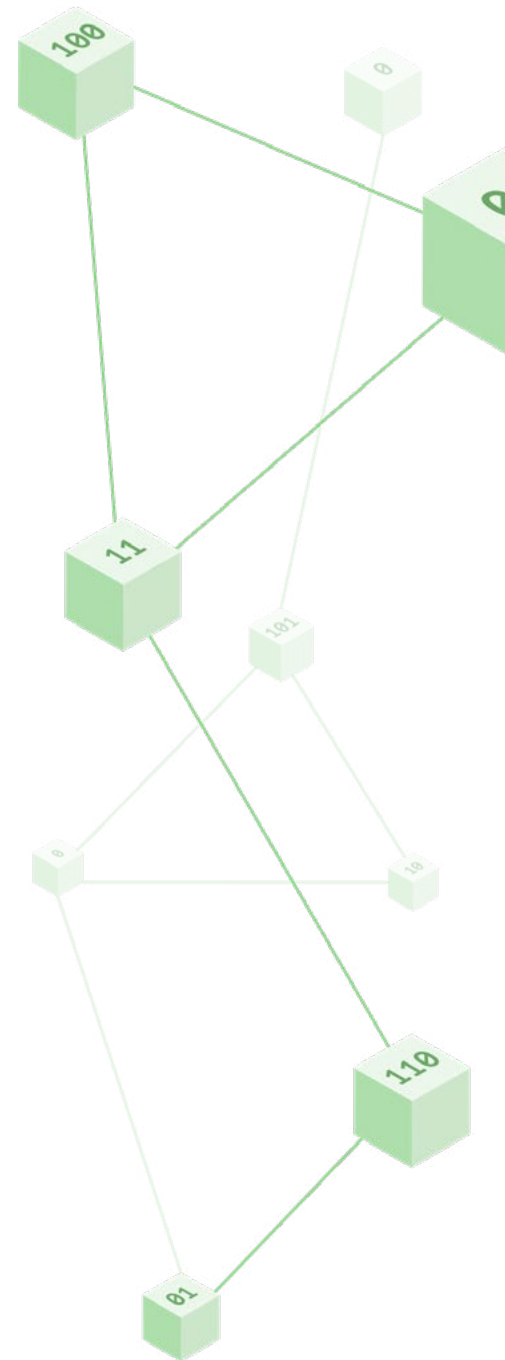
Refer to our [SIEM Cost Comparison](#) calculator to calculate exact pricing for traditional SIEM vs. Exabeam.

To test drive a next-generation SIEM, [request a demo](#) of the Exabeam Security Management Platform.

SIEM Essentials Quiz

Are you ready to show off your SIEM knowledge?
These 25 essential questions will test just how
well you know network security.

START QUIZ



About Exabeam

Exabeam is the Smarter SIEM™ company. We empower enterprises to detect, investigate and respond to cyberattacks more efficiently so their security operations and insider threat teams can work smarter. Security organizations no longer have to live with excessive logging fees, missed distributed attacks and unknown threats, or manual investigations and remediation. With the Exabeam Security Management Platform, analysts can collect unlimited log data, use behavioral analytics to detect attacks, and automate incident response, both on-premises or in the cloud. Exabeam Smart Timelines, sequences of user and device behavior created using machine learning, further reduce the time and specialization required to detect attacker tactics, techniques and procedures. For more information, visit exabeam.com.

Exabeam, Smarter SIEM, Smart Timelines and Security Management Platform are trademarks or registered trademarks of Exabeam, Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners. © 2019 Exabeam, Inc. All rights reserved.

