SUMMARY

Data breaches and security incidents occur so often that it’s easy to feel numb to the continuous threats experienced by every organization. To avert these, security controls include a spectrum of threat detection technologies and alerting systems. But since breaches continue with regularity, clearly something in the model is amiss. In evaluating what works and what doesn’t, one mature area of technology ripe for review is Security Information and Event Management (SIEM).

SIEM technology provides the heartbeat of identifying and managing threats. Its technological roots are almost two decades old and are showing their age. For the exponential rise in event data is swamping analytical capabilities of legacy SIEMs that rely primarily on static rules. For SIEM users, today is a point of crossroads – do we deliberately throttle the growing amount of event data because it’s too expensive to log and store (and risk missing undetected threats)? If we collect it, what if our legacy SIEM can’t analyze that much data? How can we leverage big data and lift SOC detect-and-defend legacy capabilities to a new level? What is the new essence of an effective SIEM?

THIS WHITE PAPER PRESENTS TEN MUST-HAVE FEATURES OF AN INNOVATIVE AND EFFECTIVE SIEM CAPABLE OF HANDLING EVOLVING THREATS. THESE NEW CAPABILITIES ENABLE FOUR CRUCIAL AREAS OF FUNCTIONALITY FOR EFFECTIVELY PROCESSING EVENTS RELATED TO SECURITY: COLLECT, DETECT, INVESTIGATE AND RESPOND.
CHALLENGES FOR LEGACY SIEM

Large organizations use a SIEM because it’s a necessary, central component of a multilayered security control infrastructure. Security alerts, system logs and other inputs can easily reach hundreds of thousands or millions of events every day. It’s impossible for humans to manually keep up with this deluge of data, so they use a SIEM to consolidate and process the work more efficiently.

The growing number of data sources pumping more events and security alerts into enterprise SIEMs is posing several operational issues. The culprit is legacy technology in many SIEMs, which is approaching two decades since early solutions appeared in the market. Four legacy characteristics of many SIEMs include:

**Volume-based pricing**
Legacy SIEMs often levy a usage fee based on the amount of data ingested and processed. Today there are more tools than ever providing a wider array of applications and bigger log files. The amount of SIEM data is soaring and volume-based pricing means organizations are subjected to unpredictable cost increases. This is a disincentive to collecting everything for analysis with modern data science. It also limits capabilities for threat detection and creates blind spots during incident investigations.

**Static correlation rules**
The typical legacy SIEM threat detection methodology is based on correlation rules. As the variety of threats – especially by unknown attack vectors – has risen, the manual effort required to create, tune and update correlation rules has left legacy SIEMs unable to detect advanced threats. This leaves SOC teams susceptible to wasting time on false positives and curbs productivity.

**Poor tracking of lateral movement**
When tracking is substandard, SOC operators get an incomplete picture of users’ activities. Without accurate tracking provided by a modern SIEM, attackers who breach defenses are often able to move laterally through the enterprise network without detection. Tracking requirements are more complex in environments with heavy use of cloud services and BYOD. Consequently, the SOC team misses threats or is sidled by false negatives, and is unable to determine the full scope of attacks without exorbitant manual effort.

**Minimal automation**
When legacy SIEM technology fails to provide adequate automation, the organization is faced with increased risk and longer durations of exposure to threats. Supporting the manual processes is a drain on operational efficiency. It also creates a scenario of steadily growing needs for more skilled operators, which is a major challenge in today’s full-employment environment.

Several of the “must have” SIEM features described below were created to address these legacy characteristics and issues.
**COLLECT**

Event data collection is the first phase of every SIEM. It begins with the technical ability to ingest, process, and use all security data no matter where it is derived, including on-premise or in the cloud.

1. **Ability to easily collect and manage data from anywhere**

Data about operating our virtual world is flooding Security Operations Centers like a tsunami. The growth in event data is about to jump into hyper-speed as billions of devices and apps in the Internet of Things weave their essence into the global web.

Event data in an organization can come from virtually anywhere, including public, private and hybrid cloud services; on-premise log sources for security controls, databases, apps, and so forth; and network telemetry data such as flows and packets. Modern SIEMs require a flexible logging infrastructure capable of handling this myriad of data types and sources.

A side effect of collecting data from such a wide array of sources is the need to manage the collectors which enable collection. A modern SIEM logging infrastructure should include centralized remote-collector management to ensure intake of all essential data. Using a central logging infrastructure will also help ensure fast, efficient performance of initial capture and subsequent analytics and operational processing of event data. This centralized management capability must make it easy to deploy, configure, update, and start and stop these collectors individually or en masse.

2. **Well-vetted big data architecture**

To facilitate effective manageability and usability, modern SIEM must enable capacity and speed with a big data architecture. Legacy SIEM architectures were developed in the early 2000’s, often using proprietary technology. The subsequent massive growth in event data (and all the data collected and processed by modern organizations) spawned new open source resources for big data such as Hadoop, MongoDB, Elasticsearch and Spark. New tools like these will enable organizations to analyze data at scale by distributing processing across nodes in a cluster for better performance and cost effectiveness.

Using a security-tuned user interface is equally important. Spraying tons of details at an operator is far less useful than having context-aware presentation of data – for example, where different pieces of information are highlighted based on the specific type of logs being viewed. Think about it this way: the details that matter most for a spam event are different than those that matter most for an unusual VPN connection.

As strange as it might sound, having all the technical pieces needed for data collection will only take you so far. Ultimately, effectiveness of a SIEM also depends on the practicality of ingesting all data – in particular from a cost standpoint.

3. **Flat pricing for logging (not consumption based)**

The biggest obstacle to leveraging large amounts of data for SIEM is cost. A typical SIEM pricing model imposes an additional cost for data ingestion so the more data you collect, the more you pay. Some call this a “SIEM tax.”

The legacy pricing model is grounded in uncertainty, for not knowing how much you will need to spend invites taking short cuts – i.e., limiting the amount of data ingested when you should be increasing it to enhance detection and response.

By contrast, flat pricing enables teams to collect all their security data without breaking the budget. A modern SIEM uses flat pricing to remove uncertainty and eliminate the need to cut another check each time you add a new data source. In other words it should be easy and affordable to collect data from each and every source designated as worthwhile by the security team.
A typical example is upgrading a firewall. Inevitably, this upgrade results in more data sent to the SIEM, which boosts ingestions and logging costs of a legacy SIEM. Multiple new inputs can dramatically expand scope of network telemetry collection. Flat-rate pricing accommodates additional event data collection without increasing the direct cost of SIEM operations.

**DETECT**

Innovative, next-generation SIEMs go well beyond the past architectural approaches of leaving data analysis as a manual exercise for the operator, or relying on correlation of static rules. The new approach enables advanced data processing and analytics – essential ingredients to automate proactive and reactive threat detection and investigation capabilities.

4. **Enriched user and asset context**

The purpose of a SIEM is transforming a multitude of collected disparate security data into useful information. Legacy SIEMs are unable to “connect all the dots” in processing the vast amount of security and operational data generated by the modern enterprise. A modern SIEM can use data science to provide context to these data by enriching and clarifying details that are automatically discovered, assigned to, and presented with their associated objects.

Bullets below show examples of useful information that can better enable effective threat detection and incident investigation. As your SIEM can enrich user and asset data with more context, accuracy of detection rules grows – providing information security analysts with the context they need to do useful investigations. In short, machine learning can help add color to available data, thus increasing your odds are of repelling attacks and preventing breaches.

- Dynamic Peer Grouping (see illustration)
- IP Association
- Asset Ownership
- User Type
- Machine Type
- Service Account Identification
- Personal Email Address Identification
Dynamic Peer Grouping automatically assigns users and machines to groups based on others in their environment who behave similarly based on day to day activity. The system can then identify anomalous activity deviations from the baseline behavior of that group.

5. Understanding of normal behavior

To accurately identify threats, an effective SIEM must understand “normal” behavior for users and other entities on the organization’s network. The industry term for this capability is User and Entity Behavior Analytics or UEBA. This functionality is used by a modern SIEM to identify unknown threats and insider threats. The system easily detects these threats by understanding how machines and humans normally behave; it then finds risky and anomalous activity that deviates from that baseline.

Behavioral baselining happens with three foundational elements of data science:

• Machine learning
• Statistical analysis
• Behavioral modeling

Determination of normal behavior is proof that behavioral learning has occurred. It’s important because legacy SIEMs often focus on correlation rules for detection, which only trigger when specified conditions are met. An alert might occur, for instance, when conditions A, B, and C happen within one hour. Huge collections of correlation rules—which are often marketed by vendors as “Advanced Analytics” —do not equal a modern UEBA solution. The reason is correlation rules do not learn behavior or identify anomalies. Rules do not look for normal behavior. This functionality is only provided by a SIEM that uses UEBA.

Identification of normal behavior is also very useful for performing investigations. Behavioral baselines contrast abnormal behavior that helps analysts to interpret an anomaly.

Note that many SIEM vendors are retooling by packaging up a large amount of advanced correlation rules into a bundle and saying this provides UEBA. But a bundle of rules does not always equate to a true UEBA solution. Especially with hundreds or thousands of correlation rules, which produce the same false positives provided by normal correlation rules. That’s not UEBA.
EXAMPLE: VPN BEHAVIOR BY COUNTRY

User and Entity and Behavior Analytics or UEBA monitors how a user interacts with an environment to understand normal behavior. A typical example is Virtual Private Network (VPN) access. UEBA learns how an individual uses VPN such as session duration, session frequency, origin IP, country of VPN connection origin, and so forth.

A modern SIEM uses UEBA to establish normal usage patterns. If a user’s normal country of VPN connection suddenly changes, this abnormal behavior signals potential danger. As additional anomalous behavior emerges during the same time frame, the user gets a higher risk score, eventually escalating to analyst review.

Modern SIEMs apply UEBA to every user and machine in every way they interact with an IT environment. VPN behavior is just one of a multitude of factors automatically weighed by a modern SIEM.

6. Automatic lateral movement tracking

Lateral movement is when cyber attackers progressively move through a network following the initial breach by changing some combination of credentials, IP addresses, or machines. Sometimes this is called east-west movement. The lateral quest is to find the high value data or assets that motivated the attack.

Lateral movement is intended to look like everyday use of the network – a masquerade that evades detection by legacy SIEMs. About 60 percent of attacks move laterally, according to a report on incident response. Popular ways attackers move through a victim’s network include using the Microsoft PowerShell Windows automation and configuration management tool, and Microsoft Windows Management Interface (WMI), according to the report.

Another challenge for legacy SIEMs is that logs do not contain all of the information necessary to follow a lateral attack. When limited by the SIEM’s lack of lateral tracking capabilities, analysts have to manually piece together the attack trail, which is time consuming, inefficient and often ineffective.

Lateral movement is identified by a modern SIEM that includes the capability to automate cataloging and analyzing changes in credential, IP address or device type and follow an attack no matter where it spreads in an environment.

Clearly, it’s important for a SIEM to have this capability. Reliance on legacy SIEM technology elevates the risk of completely missing lateral attacks.

INVESTIGATE

Investigation is the process of SOC analysts gathering evidence of a potential breach or incident by querying and pivoting in their SIEM, and building a timeline. The aim is to determine exactly what happened, what systems and users were impacted, and what they should do about it for remediation.

7. Improved security information model

The SIEM’s security information model must support modern requirements for rapid detection and response. If your organization uses a legacy SIEM, it’s important to assess its ability to meet current challenges of SOC operations.

Legacy SIEMs are based on events. And based on analysis of these events, the SIEM’s security information model must support helping analysts to build attack timelines. Only a specific play-by-play analysis can reveal exactly what happened and what must be done to remediate an incident.

The Achilles heel of legacy SIEMs is forcing analysts to manually gather evidence by querying and pivoting to create an accurate timeline. Typical elements requiring manual effort by analysts include:

- Gathering relevant security alerts and log events
- Determining how to pivot in a SIEM
- Determining asset ownership
- Enabling IP address-to-username attribution
- Assembling incident and alert timelines

Each transition from query to pivot can require hours – crucial delays when an attack may already be well under way. A modern SIEM’s security information model should automatically discover and store these data in a useful form factor for analyst consumption. It must contain the entire attack chain, without gaps or holes, so SOC managers and operators can rapidly isolate root cause and execute response playbooks. The model is a “container” that is enriched with machine learning and data science to learn asset owners, user types, normal and abnormal behavior of relevant users and everything else required for relevant action upon the incident.

A modern SIEM also performs automated host-to-IP-to-user mapping for a complete view of all activity. It preprocesses and saves information to be instantly recalled as needed – instead of waiting minutes, hours, or even days with older models. The modern security information model includes:

- **Scope** – what machines are involved and who is using them.
- **Sequence** – what happened in what order on any given day (whether or not there was a malicious event).
- **Identity** – the context and identities of all users and machines including user types, machine types, roles, groups, etc.
- **Normal and anomalous behavior** – and risk scores that show how risky or anomalous a behavior is.

---

**Security Data Should be Stored in a Useful Form Factor**

A diagram showing the process from incident scope to pre-built timelines and response playbooks.

---

Ten Must-Have Features of a Modern SIEM
8. **Prebuilt incident timelines for rapid investigation**

Timelines are the foundation of investigating security incidents. Timelines are about specific potential threat actors, graphically showing a user session as a series of events in chronological order from log on to log off. They help SOC operators to understand what behavior is normal or abnormal. Automation of timeline creation saves a huge amount of time and effort required by most SIEMs.

The modern SIEM’s information model described above is what enables automatic creation of a timeline for every event. A modern SIEM should automatically populate pre-built incident timelines to speed investigations by SOC operators.

Examples of content created by a SIEM’s timeline automation include:

- A sequence of activity surrounding an alert
- What happened before or after an alert
- Identifying if that activity was normal or abnormal for the user, the user’s peer group, the company
- Changes in devices, IP addresses, or credentials associated with lateral movement

Timeline automation radically improves productivity of SOC analysts. Faster response helps to minimize the impact of a security incident and implement remediation with alacrity.

---

**THREAT HUNTING WITH PRE-BUILT TIMELINES**

Legacy SIEMs typically require seasoned SOC analysts to handle threat hunting. Skills needed include deep security domain expertise, mastery of query languages, and the ability to interpret arcane results and determine how to proceed with an investigation. A modern SIEM with pre-built timelines can offer a better interface for threat hunting that’s easily used by a junior analyst. Instead of presenting discrete events, a pre-built timeline presents the results with context to help rapidly distill the essence of a threat – and how to fix it if needed.
9. **Incident prioritization**

Alert fatigue is a silent killer in a SOC's ability to effectively execute its responsibilities for digital security. It occurs when a blizzard of alerts overwhelms a team’s practical ability to rapidly know whether each individual alert matters – or not. A large organization may experience hundreds of thousands or millions of daily system alerts.

Legacy SIEMs typically provide minimal help with incident prioritization and may create more harm than good. They collect alerts from tools and create their own alerts – thus increasing total workflow instead of focusing results. While their data collected may help with prioritization, it is not helpful if analysts must manually parse those data to discover what matters most.

Incident prioritization can improve SOC operations exponentially by automatically telling analysts the precise order of severity.

Providing incident prioritization is a hot topic in security and the market is beginning to see a few start-ups providing this capability as a standalone solution. It may be included as a feature within some UEBA solutions. Incident prioritization is standard in a modern SIEM with UEBA.

Examples of how a modern SIEM with UEBA enables incident prioritization include:

- Ingest data from all sources and analyze it together
- Behavioral analysis identifies abnormal behavior
- Alerts associated with high risk sessions are prioritized according to their risk scores
- Machine learning and data science provide further gains (e.g. Bayesian scoring)
RESPOND

Supporting the response phase of threat management is another major weakness for legacy SIEMs. Typical help in this area is limited, which does not help lift the operation burdens of short-handed security teams by delivering meaningful degrees of automation.

EXABEAM IS THE SMARTER SIEM

Exabeam provides all ten must-have features of an innovative and effective next-generation SIEM covering the four phases of SOC operations.

10. Security orchestration and automation

Security orchestration and automation is the practical result of automating responses based on the results of a modern SIEM. The industry calls this functionality by many acronyms (SOAR, SAO, SOA) that are all about improving incident response.

SOAR replicates the expertise of your SOC team’s ninjas in three ways:

- Connects and coordinates security solutions via API-based integrations without custom scripting; this lets the SIEM collect data from or push actions to other security and IT tools
- Automates response with pre-built or customized incident workflows and playbooks of best responses to a specific event
- Reduces response times and human errors with automation of incident response allow junior analysts to run the playbooks

SOAR lets your SOC team do more with less, providing faster mean time to resolution with fewer employees required for each incident.

Orchestration capabilities establish a foundation for automation and accelerate ad hoc tasks. But playbooks are the vehicle that enables SOC teams to achieve truly meaningful levels of security automation – to respond to incidents in a thorough and consistent manner with a single click of the mouse (or in some cases zero clicks). As step-by-step response plans that combine logic and API calls, playbooks codify standard operating procedures and capture best practices for investigating, mitigating, and recovering from threats and incidents. Related, nice-to-have features include prebuilt playbook templates for common scenarios, a visual playbook editor, and the flexibility to run playbooks in either a fully or semi-automated manner.
CONCLUSION

The importance of a Security Information and Event Management solution is beyond debate. But the topic of how well a legacy SIEM addresses modern threats and operational requirements is a subject worthy of annual review in every large organization. Exabeam believes the strategy of coddling old technology and limiting the amount of operational data analyzed by a legacy SIEM is a sure recipe for successful breaches. Moving to the next generation of SIEM technology and using a solution that easily analyzes all event information for a fixed direct cost is a safer path to security threat detection and response. We invite your organization to take a closer look at Exabeam’s next-generation SIEM and quickly tap the benefits of the ten must-have features described above. For more information, please contact us at exabeam.com.

ABOUT US

Exabeam delivers next-generation security management technology that enables organizations to protect their most valuable information. The Exabeam Security Management Platform combines unlimited log data collection, advanced behavioral analytics, and automated incident response, all supported by Exabeam’s patented Smart Timelines technology that uses machine learning to track identity and behavior over time. The company’s recent industry accolades include Forbes Cloud 100, Inc. 500, and SC Awards Europe, among many other distinctions. Exabeam is privately funded by Aspect Ventures, Cisco Investments, Icon Ventures, Lightspeed Venture Partners, Norwest Venture Partners and well-known security investor Shlomo Kramer. For more information, visit https://www.exabeam.com.