

White Paper

The Complete Guide to Log and Event Management

Dr. Anton Chuvakin



Table of Contents:	2	Introduction
	3	Security Information and Event Management Defining Features
	3	Log Management Defining Features
	4	High-level Comparison: SIEM vs. Log Management
	5	SIEM and Log Management Use Cases
	6	PCI DSS
	6	FISMA
	6	HIPAA
	6	Technology Trend
	7	Example SIEM and Log Management Scenario
	7	Architecting Log Management and SIEM
	9	What to Do First? SIEM or Log Management?
	10	Do All Companies Have to Graduate from Log Management to SIEM?
	11	After Log Management and SIEM: Maturity Curve
	13	Mistakes
	16	Conclusions
	16	About the Author

Introduction

Security information and event management (SIEM) technology has existed since the late 1990s, but it has always been somewhat controversial in the security industry due to its initial promise of a “security single pane of glass” combined with slow adoption across smaller organizations. More recently, traditional SIEM has been joined by a broad-use log management technology that focuses on collecting a wide variety of logs for a multitude of purposes, from security incident response to regulatory compliance, system management and application troubleshooting. In this paper we will analyze the relationship between these two technologies—SIEM and log management—focusing not only on the technical differences and different uses for these technologies, but also on architecting their joint deployments. For example, if you need to satisfy logging requirements of PCI DSS, which one should you deploy? What technology is better suited to optimize your incident response and investigation procedures? Which one will give you real-time insight about the attacks? In addition, we will provide recommendations for companies that have deployed log management or SIEM in order for them to plot their roadmap to enhancing, optimizing and expanding their deployment. We will also recommend a roadmap for companies that have already deployed both of these technologies.

SIEM tools first appeared on the market in 1997. Their original use was for reducing network intrusion detection system (IDS) “false positives,” which plagued NIDS systems at the time. The tools were complex to deploy and use, so they were only used by the largest organizations with the most mature security programs. The market was sized at a few million dollars in the late nineties, while now, some analysts report that the market is on track to reach billions in the coming years. Today’s SIEM tools, such as

Novell® Sentinel™, are used by firms large and small, from Fortune 1000 or Global 2000 organizations to tiny SMBs—small and medium businesses.

Before beginning our analysis, it will be helpful to define “SIEM” and “log management” and explain the differences between them.

SIEM covers relevant log collection, aggregation, normalization and retention; context data collection; analysis (correlation, prioritization); presentation (reporting, visualization); security-related workflow and relevant security content. All the use cases for SIEM focus on information security, network security, data security as well as regulatory compliance.

On the other hand, **log management** includes comprehensive log collection, aggregation, original (raw, unmodified) log retention; log text analysis; presentation (mostly in the form of search, but also reporting); related workflow and content. With log management, the use cases are broad and cover all possible uses for log data across IT and even beyond.

The key difference that follows from the above definitions stems from the fact that SIEM focuses on security—the first word in “security information and event management”—and use of various IT information for security purposes. On the other hand, log management focuses on logs and wide-ranging uses for log data, both within and outside the security domain.

Security Information and Event Management Defining Features

Let's further discuss what features can be called "defining" SIEM features; most users will look for most of these features while choosing a SIEM product. The features are:

- **Log and context data collection:** This includes being able to collect logs and context data (such as identity information or vulnerability assessment results) using a combination of agentless and agent-based methods.
- **Normalization and categorization:** This covers being able to convert collected original logs into a universal format for use inside the SIEM product. The events are also categorized into useful bins such as "Configuration Change," "File Access" or "Buffer Overflow Attack."
- **Correlation:** This is used to describe rule-based correlation, statistical or algorithmic correlation, as well as other methods that include relating different events to each other and events to context data. Correlation could be in real time, but not all tools support real-time correlation and instead focus on correlating historical data from their databases. Other log analysis methods are sometimes bundled under the correlation label as well.
- **Notification/alerting:** This includes being able to trigger notifications or alerts to operators or managers. Common alerting mechanisms include e-mail, SMS, or even SNMP messages.
- **Prioritization:** This includes different features that help highlight the important events over less critical security events. This may be accomplished by correlating security events with vulnerability data or other asset information. Prioritization algorithms would often use severity information provided by the original log source as well.
- **Real-time views:** This covers security monitoring dashboards and displays, used for security operations personnel. Such displays will show collected information as

well as correlation results to the analysts in near real time; they can also be fed by historical, archived data.

- **Reporting:** Reporting and scheduled reporting covers all the historical views of data collected by the SIEM product. Some products also have a mechanism for distributing reports to security personnel or IT management, either over e-mail or using a dedicated secure Web portal.
- **Security role workflow:** This covers incident management features such as being able to open cases and perform investigative tasks, as well as automatically or semi-automatically perform typical tasks for security operations. Some products also include collaborated features that allow multiple analysts to work on the same security response effort.

The above functionality can be found in most commercial SIEM products on the market today. However, most products have strong and weak points, as well as additional "secret sauce" features.

Log Management Defining Features

Let's start by considering the defining features of a log management system. These include:

- **Log data collection:** This covers being able to collect all logs using agent-based or agent-less methods, or a combination of the two.
- **Efficient retention:** While collecting and saving log data does not sound like a big engineering challenge, being able to collect gigabytes and even terabytes of log data efficiently—and retaining it while providing fast searching and quick access to it—is not trivial. Given that many regulations mandate specific terms for log data retention (ranging all the way to multiple years), this functionality is critical to a log management system.

- Searching is the primary way to access information in all of the logs, including logs from custom applications. Search is indispensable for investigative use of logs, log forensics, and finding faults while using logs for application troubleshooting. A clean and responsive interactive search interface is thus essential for a log management system.
- Log indexing or parsing is a key component of a log management system. Indexing can speed up searches literally by a factor of a hundred. Indexing technology creates a data structure called an index that allows very fast keyword type searches and Boolean type searches across the log storage. Sometimes indexing is used to enable other full text analysis techniques. Think about this as “Google for logs.” Not all log management tools support indexing, or advertise log collection rates that don’t account for indexing, so be careful with vendor claims here.
- Reporting and scheduled reporting cover all the data collected by the log management product and are similar to SIEM reporting. The strength of reporting, whether for security, compliance or operational reasons,

can make or break the log management solution. Reporting should be fast, customizable and easy to use for a broad range of purposes. The distinction between searches and reports is pretty clear: Search goes across all available, collected logs in raw, original form (like Google goes through Web pages), while report operates on logs which are parsed into a database (like an Excel spreadsheet). Carefully evaluate how easy it is to create a custom report in a log management tool. This is where a lot of solutions fall short by requiring that their operators study the esoteric aspects of their log storage data structures before they can customize the reports.

Now let’s perform a high-level comparison between functions and features of SIEM and log management.

High-level Comparison: SIEM vs. Log Management

In the table below, we show key areas of functionality and explain how SIEM and log management are different.

Functionality Area	Security Information and Event Management (SIEM)	Log Management
Log collection	Collect security relevant logs	Collect all logs including operational logs and custom application logs
Log retention	Retain limited parsed and normalized log data	Retain raw and parsed log data for long periods of time
Reporting	Security focused reporting, real-time reporting	Broad use reporting, historical reporting
Analysis	Correlation, threat scoring, event prioritization	Full text analysis, tagging
Alerting and notification	Advanced security focused reporting	Simple alerting on all logs
Other features	Incident management, other security data analysis	High scalability for collection and searching

Now let us review how SIEM and log management technologies are used.

SIEM and Log Management Use Cases

Before discussing the joint architecture of SIEM and log management, we need to briefly present typical use cases that call for deployment of a SIEM product by a customer organization. We will start from the very high level of three main types of use cases:

1. Security, both detective and investigative: Sometimes also called threat management, this focuses on detecting and responding to attacks, malware infection, data theft and other security issues.
2. Compliance, regulatory (global) and policy (local): This focuses on satisfying the requirement of various laws, mandates and frameworks as well as local corporate policy.
3. Operational, system and network troubleshooting and normal operations: Specific mostly to log management, this use case has to do with investigating system problems as well as monitoring the availability of systems and applications.

On a more detailed level, security and compliance use cases fall under several scenarios. Let's review them in detail.

The first usage scenario is a traditional Security Operations Center (SOC). It typically makes heavy use of SIEM features such as real-time views and correlation. A SIEM customer organization will have analysts online 24x7 and have them "chase" security alerts as they "pop up." This was the original SIEM use case when SIEM technology started in the 1990s; today it is relegated to the largest organizations only.

The next use case is sometimes called the "mini-SOC" scenario. In this case, the security personnel will use non real-time, delayed views to check for security issues ("analysts come in the morning"). The analysts are online

Recently, traditional SIEM has been joined by a broad-use log management technology that focuses on collecting a wide variety of logs for a multitude of purposes, from security incident response to regulatory compliance, system management and application troubleshooting.

maybe a few hours each day and only review alerts and reports as needed and not in near-real time—unless the events happened while they were logged in to the product.

The third scenario is an "automated SOC" scenario where an organization configures their SIEM to alert based on rules and then "forgets" it until the alert. The analysts never log in unless there is a need to investigate alerts, review reports weekly/monthly or perform other rare tasks. This is the use case that many smaller organizations want and few SIEM products can deliver, at least not without extensive customization. It is worthwhile to add that a lot of SIEM products are sold with an expectation of being an automated SOC, but such expectations are rarely realized.

Log management technologies have a role in other scenarios outside of security as well. Application troubleshooting and system administration are two additional important use cases for log management systems. When the application is deployed and its logging configured, the log management system is used to quickly review errors and exception logs. It will also review summaries of normal application activity in order to determine application health and troubleshoot possible irregularities.

Another scenario is "compliance status reporting." Here analysts or security managers review reports with a focus on compliance issues. The review occurs weekly or monthly or as prescribed by a specific regulation. There is not necessarily

Today's SIEM tools, such as Novell Sentinel, are used by firms large and small, from Fortune 1000 or Global 2000 organizations to tiny SMBs—small and medium businesses.

a security or operations focus. This use case is commonly a transition phase and the organization will likely later mature to one of the aforementioned use cases. Log management tools are most often deployed for this scenario, but it is not uncommon to use a SIEM product for compliance as well. In the latter case, long-term log retention requirements often challenge the deployment.

Given that logs are very important for meeting compliance mandates, let's consider a few regulations in detail.

PCI DSS

The Payment Card Industry Data Security Standard (PCI DSS) applies to organizations that handle credit card transactions. It mandates logging specific details, log retention and daily log review procedures.

Even though logging is present in all PCI requirements, PCI DSS also contains Requirement 10, which is dedicated to logging and log management. Under this requirement, logs for all system components must be reviewed at least daily. Further, PCI DSS states that the organization must ensure the integrity of its logs by implementing file integrity monitoring and change detection software on logs. It also prescribes that logs from in-scope systems are stored for at least one year.

FISMA

Federal Information Security Management Act of 2002 (FISMA) emphasizes the need for each federal agency to develop, document and implement an organization-wide program to secure the information systems that support its operations and assets. NIST SP 800-53, "Recommended Security Controls for Federal Information Systems," describes

log management controls including the generation, review, protection and retention of audit records, plus steps to take in the event of audit failure.

NIST 800-92, "Guide to Computer Security Log Management," also created to simplify FISMA compliance, is fully devoted to log management. It describes the need for log management in federal agencies and ways to establish and maintain successful and efficient log management infrastructures—including log generation, analysis, storage and monitoring. NIST 800-92 discusses the importance of analyzing different kinds of logs from different sources and of clearly defining specific roles and responsibilities of those teams and individuals involved in log management.

HIPAA

The Health Insurance Portability and Accountability Act of 1996 (HIPAA) outlines relevant security standards for health information. NIST SP 800-66, "An Introductory Resource Guide for Implementing the Health Insurance Portability and Accountability Act Security Rule", details log management requirements for the securing of electronic protected health information. Section 4.1 of NIST 800-66 describes the need for regular review of information system activity, such as audit logs, access reports and security incident-tracking reports. Also, Section 4.22 specifies that documentation of actions and activities need to be retained for at least six years. Logs are sometimes considered part of that. Recent HITECH Act of 2009 promises to boost HIPAA implementations in the coming years.

Technology Trends

As we mentioned before, SIEM technology is more than 10 years old; it has gone through multiple phases which we could write an entirely new white paper about. We will highlight a few of the SIEM technology trends.

While SIEM started as a technology for large global companies and sensitive government agencies, it continues a march down market. Many observers predict that 2010 or 2011 will be the year of the major SIEM vendors' mid-market battle for dominance. As a result, smaller customers will get much improved tools for security management.

Another trend is acceptance of separate roles for SIEM and log management. Now, most SIEM vendors offer log management solutions as well. This also supports expanding uses for SIEM tools including IT operations, fraud analysis, application troubleshooting, going all the way up to IT GRC uses for high-level governance and risk measuring goals.

We're also witnessing the beginning of convergence between IT operations and IT management and security management. While analysts have predicted this trend for several years, it has failed to fully materialize until now. Despite that fact, many predict the trend of convergence of security management and IT operations management will continue, and security tools will have more linkage into IT operational tools such as network and system management.

Example SIEM and Log Management Scenario

This case study covers a deployment scenario of a SIEM and log management solution to satisfy PCI-DSS requirements at a large retail chain. The retailer decided to deploy a commercial log management solution when its PCI assessor suggested it would be required to pass an assessment. A log management vendor suggested that the retailer get both log management and SIEM solution at the same time. So, it progressed from not doing anything with its logs directly to running an advanced log management system and real-time correlation capability.

The project took a few months following a phased approach. The retailer's IT staff decided to implement it from the outside in, based on an initial risk assessment.

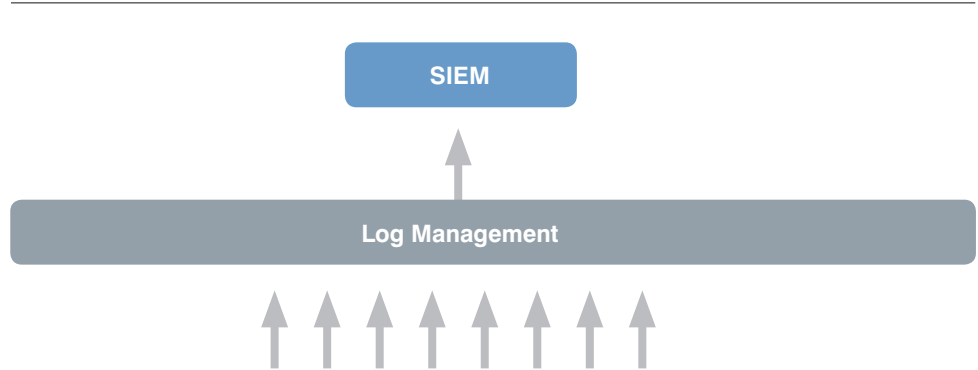
They started from their DMZ firewalls and then progressed by feeding additional logs into a log management system, while simultaneously defining correlation rules and running reports from the vendor's PCI DSS compliance package. As they learned to respond to alerts, their processes matured and they started making use of more of the SIEM functionality.

Overall, the project represented a successful implementation of PCI logging requirements. The organization passed the PCI assessment with flying colors and was commended on their comprehensive approach to logging and security monitoring. In addition, the security team built a case that their PCI SIEM implementation actually addresses additional compliance mandates since PCI DSS goes into a deeper level of details while covering essentially the same areas of IT governance. At the same time, log management tools also bolstered their operational capabilities and overall IT efficiency, while SIEM gave them the core for their future real-time detection and response capability.

Architecting Log Management and SIEM

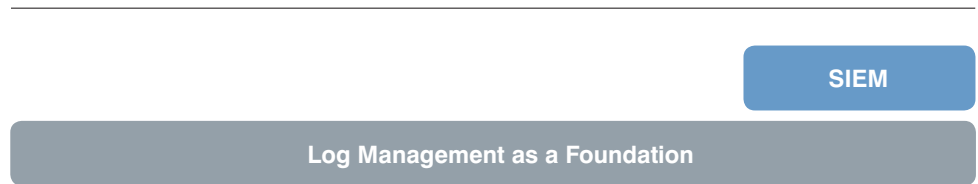
Given the differences between technologies, many organizations have deployed both SIEM and log management, or are considering enhancing an existing deployment of one of the technologies with the other. What are some of the common joint architectures of SIEM and log management?

We will refer to the most common scenario as "SIEM shield." Many of the organizations that deployed legacy SIEM solutions attempted to send too much data to their SIEM, thus overloading it and possibly losing critical data and functionality. They addressed this problem by also acquiring a log management tool and deploying it "in front" of their SIEM solution.



In this case, an inherently more scalable log management tool is deployed in front of SIEM to serve as a shield and filter to protect a less scalable SIEM tool from extreme log flows. It is not uncommon to only send every 10th event received by the “log shield” to a SIEM that is hiding behind it. At the same time, all received

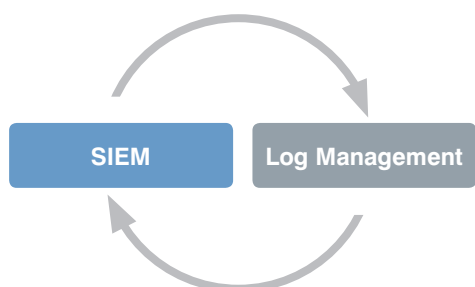
events are archived on a log management tool. For example, if a total log volume equals 40,000 log messages each second, a SIEM tool will receive only 4,000 messages a second.



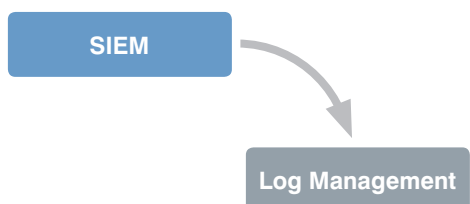
Another scenario emerges when log management is deployed first to create an enterprise logging platform. SIEM is then added as one of the applications of such a platform. This scenario can be called **“grow up to SIEM”** and accounts for up to 50 percent of SIEM deployments today.

This is the case where an organization gets a log management tool and slowly realizes a need—as well as develops an ability—for correlation, visualization, monitoring, workflows, etc. Such a scenario is the most logical for most organizations as we discuss further in this paper.

In the next case, SIEM and log management are deployed alongside each other and at the same time. This is an “emerging scenario” since more people now get both at the same time—and typically from the same vendor. Indeed, if an organization somehow realizes the need for correlation, it then needs to collect and save all the logs and have the ability to perform efficient search and raw data analytics.



Next is a SIEM deployment with log management as an archive for processed and other logs. This scenario arises when somebody buys a big SIEM for security monitoring and then, over time, realizes that something is missing. As a result, a log management tool is deployed to “dump” all logs into and to perform analysis of the raw logs that the SIEM “rejects” (i.e., doesn’t know how to parse, normalize, categorize, etc). This leads to a broadening use case from security monitoring to incident response and PCI DSS compliance.



Being able to respond better has to happen before you are forced to respond faster.

It is much easier to be prepared to respond than to monitor.

Obviously, it goes without saying there are lots of “log management only” (still growing) situations and some “SIEM only” (likely shrinking) deployment scenarios.

What to Do First? SIEM or Log Management?

Fortunately, the question of which technology needs to be deployed first has a very simple answer. If you have logs, you need log management. This equally applies to organizations with one server, all the way to organizations with 100,000 servers. Clearly, the technology they deploy to manage logs will be different, but the existence of logs leads them to log management. For example, if you have to review logs from a single machine, built-in operating system tools will usually suffice. On the other hand, if your daily log volume reaches an impressive 100 GB (not an impossible situation!), sophisticated—and thus expensive—tools needs to be deployed.

In fact, even a recent Gartner note “How to Implement SIEM Technology” (Gartner, 2009) unambiguously states, “Deploy log management functions before you attempt a wide-scale implementation of real-time event management.” Further, they clarify that when SIEM technology is driven by compliance, the same order of deployment persists: “the first phases of a SIEM deployment that is primarily driven by PCI would implement log management functions for the systems that are in scope for the PCI assessment.” The overall theme here is that being able to respond better has to happen before you are forced to respond faster.

If you have logs, you need log management.
This equally applies to organizations
with one server, all the way to organizations
with 100,000 servers.

What about those organizations that have already deployed legacy SIEM tools? For them, looking into log management as soon as possible is a smart thing to do. Being able to go through a complete collection of log records will boost their investigative capabilities and help them meet compliance mandates.

Do All Companies Have to Graduate from Log Management to SIEM?

What happens after an organization deploys a log management tool and starts using it effectively for security and compliance as well as operational purposes? The natural and logical progression is for organizations to graduate to near-real-time event management by deploying a SIEM tool.

This paper is the first document that formulates “graduation criteria” for such development. Organizations that graduate too soon will waste time and effort, and won’t realize any increased efficiency in their security operation. However, waiting too long also means that the organization will never develop the necessary capabilities to secure themselves.

In brief, the criteria are:

- **Response capability:** The organization must be ready to respond to alerts soon after they are produced.
- **Monitoring capability:** The organization must have or start to build security monitoring capability by creating a Security Operation Center (SOC) or at least a team dedicated to ongoing periodic monitoring.

- **Tuning and customization ability:**

The organization must accept the responsibility for tuning and customizing the deployed SIEM tool. Out-of-the-box SIEM deployments rarely succeed or manage to reach their full potential.

Let’s review the criteria in detail.

First, the organization must be ready to respond to alerts soon after they are produced. While the claims that “modern business works in real-time and so the security should too” are often heard from various vendors, it appears that few organizations are able to achieve that at the moment. So, before deploying SIEM ask: How real-time is your security? One might think that most of the time, security is indeed in real-time or very close to it. Network intrusion detection systems pick up attacks off the wire within microseconds, firewalls block connections as they happen, and anti-virus technology makes the best effort to catch the viruses as soon as they arrive.

Thus, few people will agree to buy a network intrusion detection system (NIDS) that will only notify of an attack after two have passed. However, those same people will have their security analysts check the IDS alarms every morning. If they discover a critical compromise, a millisecond response time of the NIDS system will not matter, but the hourly response time of the personnel will. So, if the “morning after” alert investigation results in discovering a critical system compromise, it is still deemed acceptable.

Similarly, if a virus-infected file arrives and the software can clean it “in real-time”, the problem is solved. However, in case the antivirus software detects the malicious code, but cannot automatically clean or quarantine it and issues an alert instead (which happens in the case of some backdoors and Trojans), the response falls back on the shoulders of the analysts who are likely hours behind. With today’s sophisticated threats, this is

often enough time for a serious breach to occur, which could take months to clean up. As a result, advanced alerting and stateful correlation rules will deliver sub-second responses, but you need to be prepared to respond to them.

In fact, if an organization does not have an SOC or any monitoring capability, whether security monitoring or operational monitoring with strict SLAs, many of the SIEM features will not be fully utilized. A common first step from purely responsive use of logs to full-blown security monitoring is utilizing “delayed periodic monitoring” which really means “reviewing log reports every morning.” This can be accomplished with a log management tool or with a SIEM tool.

The final graduation criteria relates to tuning and customization ability. The organization must accept the responsibility for tuning and customizing the deployed SIEM tool in order to fit its powerful and customizable features to a problem set that an organization faces. A second option is to hire a specialist consulting firm to do the tuning for them. Every business is unique, and in order to be most effective, a SIEM must take into account the unique business processes that exist. This might mean creating alerts, writing correlation rules or customizing reports in order to gain insight about the organization’s security or compliance posture. From the author’s experience, it is worthwhile to note that out-of-the-box deployments with inflated expectation of SIEM as “analyst-in-the-box” rarely succeed.

What is interesting is that organizations that have no immediate plans to migrate from, say, compliance-focused log management should still choose a logging tool that allows them to later graduate to SIEM. Even with no initial plans to move beyond compliance, many SIEM and log management deployments follow so-called “compliance+” models, which means that the tool is purchased for a particular regulatory framework, but is utilized

for many other security and IT challenges.

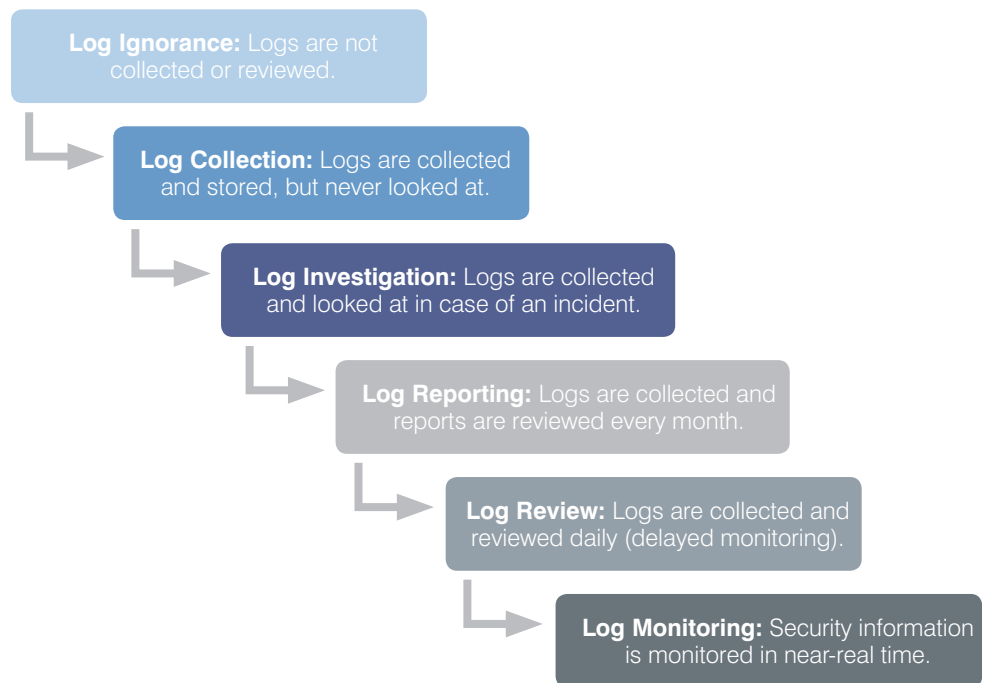
At this point, it is worthwhile to note that some of the log management tools do not offer such a “graduation path” to a SIEM. In particular, simpler tools that only allow you to collect raw logs and perform searches across them may be extremely useful; however, they might not allow you an easy way to achieve full normalization, categorization and other security-focused enrichment of log data. In general, if your tool collects and retains raw log records and cannot be paired with a SIEM solution that can make such data for security monitoring and analysis, graduation to monitoring will not be possible. Other tools will need to be purchased if your organization becomes ready for real-time monitoring.

Given that using a SIEM solution effectively gives you direct threat reduction benefits via its advanced security focused analysis (but only if your organization is ready for SIEM), the “compliance+” model makes sense. Overall, it allows the organization to move closer to that mythical “single-pane of glass” for security management.

After Log Management and SIEM: Maturity Curve

What happens next after both log management and SIEM are deployed and “operationalized” to help with compliance and deliver security benefits to an organization? There is a maturity curve that stretches from complete log ignorance, to log collection and retention, to occasional investigation, to periodic log review and then all the way to near-real-time security monitoring.

The trend here is from being ignorant, to being slowly reactive, to being quickly reactive, to eventually being proactive and aware of what is going on across your IT environment. Trying to make one jump from ignorant to proactive rarely, if ever, works!

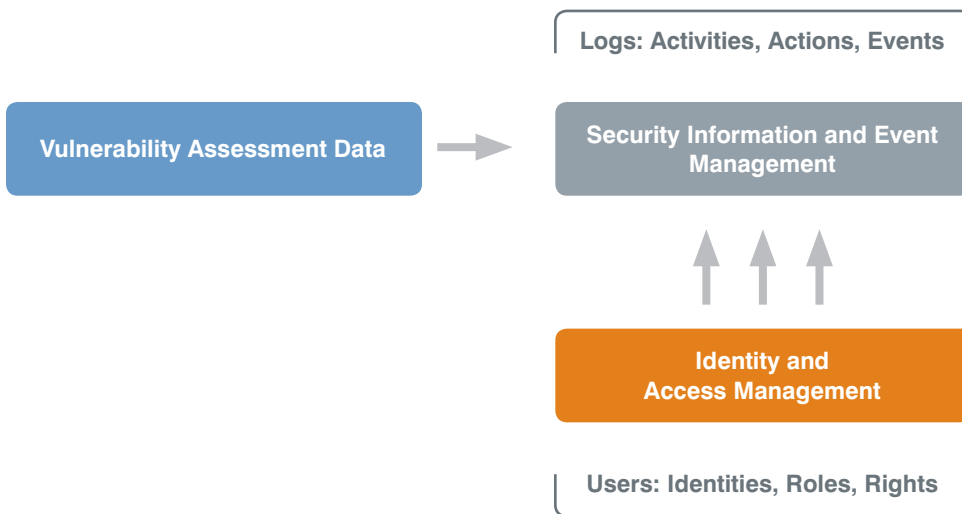


What is the next step in the evolution after that point? For starters, organizations should be continuously improving the breadth and depth of SIEM deployment by integrating it with more systems to make better use of SIEM's analytics capabilities. This gets at SIEM's core mission—security monitoring—and also solves new problems such as fraud, insider threat, application monitoring and overall user activity monitoring. SIEM starts to acquire more information and to move up the stack from network to application, from a limited number of data sources to enterprise-wide deployment. At the same time, a security organization grows with it and develops better operational procedures that allow the organization to be more agile. While expanding the deployment, it is crucial to remember that a phased approach is the only way to succeed here.

What are some of the systems that would enhance SIEM's mission and allow it to solve other problems? One of the most interesting

examples involves using information from identity management systems such as Novell Identity Manager. The information available in this system includes user identity (such as real name, work role, business unit affiliation, etc.) as well as access rights across various systems and applications. Knowing who the user is and what he is allowed to do is indispensable for security monitoring of insider activities. For example, it allows you to create "a unified identity" for each user and then use it to monitor user actions across multiple systems, even with different user names and accounts.

On top of this, identity manager integration allows a SIEM product to differentiate authorized, official logins from backdoor, unauthorized login attempts. Such integration also allows automated separation-of-duty (SoD) monitoring by making SIEM aware of which roles are not allowed to perform specific actions.



In addition, an asset management system will contain similar detailed information on all IT resources within the organization. Just like we can do with users, we can extract asset business role, business criticality, compliance relevance, administrator name and location as well as other information on what function the asset performs and who is responsible for it. Such information will dramatically improve risk computation and event prioritization functions of SIEM. Be aware that even though many vendors claim identity integration, most will only perform a simple LDAP lookup. These systems lose out on the all the rich data an identity system could provide to help a SIEM determine if activities are malicious or have regulatory relevance.

Further levels of integration—and thus increased awareness—can be provided by integrating with configuration management databases (CMDB). Such integrations allow a SIEM product to correlate detected changes across systems and applications with approved and authorized changes.

Mistakes

When planning and implementing log collection and analysis infrastructure—whether for SIEM or log management—the organizations often discover that they are not realizing the full promise of such systems. In fact, they sometimes notice that efficiency is not gained, but is lost as a result. This often happens due to the following common implementation mistakes.

We will start from the obvious—but unfortunately all too common—mistake, even in this age of Sarbanes-Oxley and PCI DSS. This mistake destroys all possible chances of benefiting from log management or SIEM.

The first mistake is **not logging at all**. Another version of the same mistake is not logging and not even knowing it until it is too late.

How can it be too late? Not having logs can lead to losing your income (PCI DSS logging requirements imply that violations might lead to your credit card processing privileges being canceled by Visa or MasterCard, thus putting you out of business), reputation (somebody stole a few credit card numbers from your database, but the media reported that all of the 40 million credit cards have been stolen since you were unable to prove otherwise) or even your freedom (see various Sarbanes-Oxley horror stories in the media).

Once both SIEM and log management have been operationalized, your organization can move up the maturity scale to comprehensive network and application visibility, user activity monitoring and other integration with different systems.

Even organizations that are well-prepared fall for this mistake. Consider this recent example. Does your Web server have logging enabled? Sure, it is a default option on both of the popular Web servers: Apache and Microsoft IIS. Does your server operating system log messages? Sure, nobody canceled `/var/log/messages`. But does your *database*? The default option in Oracle is to not perform any data access audit logging. Does Microsoft SQL fare better? Sadly, the answer is “no”, you need to dig deep in the system to even start a moderate level of audit trail generation.

Thus, to avoid this mistake one needs to sometimes go beyond the defaults and make sure that the software and hardware deployed does have some level of logging enabled. In case of Oracle, for example, it might boil down to making sure that the “audit_trail” variable is set to “db.” For other systems it might be more complicated.

Not reviewing logs is the second mistake. While making sure that logs do exist and then collecting and storing them is important, it is only a means to an end: knowing what is going on in your environment and being able to respond to it, as well as possibly predict what will happen later. As we describe above, it is a stage, but not the destination. If your company has just moved from ignoring logs to collecting logs, it is important to know that ultimately you will need to review them. If you collect logs and don’t review them, you are simply documenting your own negligence, especially if your IT security policy prescribes log reviews.

Therefore, once the technology is in place and logs are collected, there must be a process of ongoing monitoring and review that hooks into actions and possible escalations, if needed. In addition, personnel reviewing or monitoring logs should have enough information to determine what they really mean and what—if any—action is required.

It is worthwhile to note that some organizations take a half step in the right direction: They only review logs (provided they didn’t commit the first mistake and they actually have something to review) after a major incident (be it a compromise, information leak or a mysterious server crash) and avoid ongoing monitoring and log review, often by quoting the proverbial lack of resources. This gives them the reactive benefit of log analysis, which is important, but fails to realize the proactive one: knowing when bad stuff is about to happen or become worse. For example, if you review logs, you might learn that the failover was activated on a firewall, and, even though the connection stayed on, the incident is certainly worth looking into. If you don’t and your network connectivity goes away, you’d have to rely on your ever-helpful logs to investigate why both failover devices went down.

It is also critical to stress that some types of organizations *have* to look at log files and audit tracks due to regulatory pressure of some kind. As we mention previously, HIPAA regulation compels medical organizations to establish an audit record and analysis program (even though the enforcement action is notoriously lacking). Also, PCI DSS data security standard has provisions for both log collection and log monitoring and periodic review, highlighting the fact that collection of logs does not stand on its own.

The third common mistake is **storing logs for too short a time**. A SIEM system’s operational log store might retain normalized events for 30 days, but a log management system is needed for long term retention. This makes the security or IT operations team think they have all the logs needed for monitoring

and investigation or troubleshooting. This leads to the horrible realization after the incident that all logs are gone due to their shortsighted retention policy. It often happens (especially in the case of insider attacks) that the incident is discovered a long time—sometimes many months—after the crime or abuse has been committed. One might save some money on storage hardware, but lose it tenfold due to regulatory fines.

If low cost is critical, the solution is sometimes to split the retention in two parts: short-term online storage (that costs more) and long-term offline storage (that is much cheaper). A good log management tool will allow you to search through both of these stores transparently, without moving data around. A better three-tier approach is also common and resolves some of the limitations of the previous one. In this case, shorter-term online storage is complemented by a near-line storage where logs are still accessible and searchable. The oldest and the least relevant log records are offloaded to the third tier, such as tape or DVDs, where they can be stored inexpensively. However, there is no way to selectively access the needed logs. More specifically, one financial institution was storing logs online for 90 days, then in the near-line searchable storage of the log management system for two years, and then on tape for up to seven years or even more in some cases.

The fourth mistake is related to log record prioritization. While people need a sense of priority to better organize their log analysis efforts, the common mistake today is **prioritizing the log records before collection**. In fact, even some “best practice” documents recommend only collecting “the important stuff.” But what is important? This is where the above guidance documents fall short by not specifying it in any useful form. While there are some approaches to the problem, it can lead to glaring holes in security posture or even undermine the regulatory compliance efforts.

For example, many people would claim that network intrusion detection and prevention logs are inherently more important than, say, VPN concentrator logs. Well, it might be true in the world where external threats completely dominate the insider abuse and all employees and partners can simply be trusted. VPN logs, together with server and workstation logs, are what you would most likely need to conduct an internal investigation about the information leak or even a malware infection. Thus, similar claims about the elevated importance of any other log type can be similarly disputed, which would lead us to a painful realization that you do need to collect everything or most of the log records produced. But can you? Before you answer this, try to answer whether you can make the right call on which log is more important even before seeing it and this problem will stop looking unsolvable. In fact, there are cost-effective solutions to achieve just that.

The way to avoid this mistake is to deploy log management before SIEM as we prescribe earlier. This will guarantee that all needed logs are available for analysis, even if only a percentage is ever seen by a SIEM correlation engine.

The final mistake is **ignoring the logs from applications**, by only focusing on the perimeter and internal network devices, and possibly also servers, but not going higher up the stack to look at the application logging.

The realm of enterprise applications ranges from SAP and PeopleSoft to small homegrown applications, which nevertheless handle mission-critical processes for many enterprises. Legacy applications, running on mainframes and midrange systems, are out there as well, often running the core business processes too. The availability and quality of logs differ wildly across the application, ranging from missing (the case for many home-grown applications) to extremely detailed and voluminous (the case for many mainframe applications). Lack of common

logging standards and even of logging guidance for software developers leads to many challenges with application logs. Fortunately, future efforts such as MITRE CEE will remediate this problem.

Despite the challenges, you need to make sure that the application logs are collected and made available for analysis as well as for longer term retention. This can be accomplished by configuring your log management software to collect them and by establishing a log review policy, both for the on-incident review and periodic proactive log review. Look for vendors that make it easy to configure their systems to collect logs from custom applications, as these are often the most important. Later you can configure SIEM to analyze the logs for security purposes, together with network and other logs.

Conclusions

One of the paramount conclusions from this work is to remember that everybody has logs and that means that everybody ultimately needs log management. In its broadest form, log management simply means “dealing with logs.” And if you have logs, you have to deal with them—if only because many recent regulatory mandates prescribe that.

It’s also important to remember that logs are used for a very large number of situations: from traditional (incident response) to highly esoteric. Most uses of logs happen much later, after the event happens and is recorded in logs. It is much easier to be prepared to respond than to monitor.

Your organization might need to go “back to logging school” before it is ready to “graduate to SIEM.” Such graduation requires an ability to respond to alerts and customize and tune products.

Afterward, once both SIEM and log management have been operationalized, your organization can move up the maturity scale to comprehensive network and application visibility, user activity monitoring and other integration with different systems.

About the Author

Dr. Anton Chuvakin (<http://www.chuvakin.org>) is a recognized security expert in the field of log management and PCI DSS compliance. He is the author of two books “Security Warrior” and “PCI Compliance” and a contributor to “Know Your Enemy II”, “Information Security Management Handbook” and others. Anton has published dozens of papers on log management, correlation, data analysis, PCI DSS, security management (see a list at www.info-secure.org). His blog <http://www.securitywarrior.org> is one of the most popular in the industry. In addition, Anton teaches classes and presents at many security conferences across the world; he recently addressed audiences in United States, UK, Singapore, Spain, Russia and other countries. He works on emerging security standards and serves on the advisory boards of several security start-ups.

Currently, Anton is developing his security consulting practice www.securitywarriorconsulting.com, focusing on logging and PCI DSS compliance for security vendors and Fortune 500 organizations. Dr. Anton Chuvakin was formerly a Director of PCI Compliance Solutions at Qualys. Previously, Anton worked at LogLogic as a Chief Logging Evangelist, tasked with educating the world about the importance of logging for security, compliance and operations. Before LogLogic, Anton was employed by a security vendor in a strategic product management role. Anton earned his Ph.D. from Stony Brook University.