Interested in learning more about security?

SANS Institute
InfoSec Reading Room

This paper is from the SANS Institute Reading Room site. Reposting is not permitted without express written permission.

# Importance of Understanding Logs from an Information Security Standpoint

The 21st century is known as the information age. Where people, places, and devices all communicate in an endless stream of information passing by at the speed of light. With such an infinite amount of communication taking place around the world, it is important to be able to manage this information in an efficient and secure manner. Information Security has many facets and branches, but to really understand what is going on in this new world, you need the ability to read, translate, and understand the wide variety of ...

Copyright SANS Institute
Author Retains Full Rights

AD

**Title:**
*Importance of Understanding Logs from an Information Security Standpoint*

**Author:**
Stewart Allen

**Revision:**
GSEC v1.2f

## Overview

The 21st century is known as the information age. Where people, places, and devices all communicate in an endless stream of information passing by at the speed of light. With such an infinite amount of communication taking place around the world, it is important to be able to manage this information in an efficient and secure manner.

Information Security has many facets and branches, but to really understand what is going on in this new world, you need the ability to read, translate, and understand the wide variety of logs generated by the information stream.

This document will discuss the importance of logs in the 21st century, and give an idea of what problems Information Security professionals face when trying to analyze them.

We start from the beginning by defining what a log really is and what its purpose is. Then we talk about ways to improve your understanding of logs, how to decipher their cryptic formats, and how to manage logs effectively. Finally we wrap up with discussion on legalities of logs, and why it is so critical to effectively manage, maintain, and secure logs.

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

## What is a Log?

According to *Merriam-Webster's Dictionary* [1] the definition of a log is:

log (lôg, lg)
n.
A record, as of the performance of a machine or the progress of an undertaking: a computer log; a trip log.

For security professionals a log is used to record data on who, what, when, where, and why (W5) an event occurred for a particular device or application.

As events on a device or application are taking place, processes are running that generate responses based on those events, and the responses are output into a log of one form or another.

The logs can be displayed to a display monitor, a printing device, a null device, or typically any device used to display or store information. Often you will find logs stored in a log file residing on a computer.

Most logs are stored or displayed in an ASCII standard format, which is a universally understood character set for common devices and applications. This way logs generated from one device, can be read and displayed on another device.

Formatting and contents contained in the log are often left to the developer of the device or application that is generating the logs. Unfortunately, the log output is often difficult to understand and the output format will leave many people scratching their heads.

## What is the purpose of a log?

If a log has the capability to record the W5 events, then the purpose of a log is to give security professionals the ability to monitor the activities of the application or device to ensure expected or normal operations.

If you, the security professional, know that every morning at 9am log output should state a successful startup sequence for a device or application, but fail to see this, then the log has alerted you to unexpected operations.

By reviewing the log output, there is a good chance that you will be able to determine the W5 of the event, and take the necessary action to correct the problem or sponsor an investigation into an incident.

Once you and your associates have identified the issue and taken steps to resolve it, you will be able to verify the problem resolved as the correct content is found in the log the next morning at 9am.

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

**Why are logs so cryptic?**

Understanding the importance of a log, and what the purpose of the log is, then the next question to be asked is "Why are the log contents and formatting so cryptic?"

If you have every taken the time to examine a log, you will undoubtingly experienced this problem and wondered if you were ever going to figure it out.

Because a log can be generated by any device or application, the developers of that device or application will determine how the output should be formatted and exactly what content will be released to the logging processes.

If the developer is only interested in knowing "when" an application or device fails, and wants to know exactly "where" in the code the failure occurred, then the log output will most likely not show you the "who, what, or why" that caused the failure to occur. This leaves you trying to guess or piece several pieces of the log together to find those answers.

The output format could be very cryptic, and in such a format that only the developers will truly understand because they created it. They generally have no interest in generating output for the common user, nor do they have any mandate or standard that tells them to.

This creates problems for security professionals or computer administrators who are trying to investigate an incident or failure that may have occurred within their infrastructure.

In order to combat this problem, computer industry comities were formed to come up with formatting specifications that everyone could agree on and eventually understand.

As a result, it seems that two strong standards have emerged in the computer industry for the more popular UNIX and Windows environments.

Syslog is a logging system that has been standardized so that any flavor of UNIX operating system will output the same log format that can be displayed or output to standardized log files.

Windows NT operating systems support the Eventlog format, and all events output to a standardized event log format.

Some good efforts have been made of the part of Microsoft to document this format and it is one that can often be referenced against, making it that much easier to figure out what the log is telling you.

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

For example, take the *Windows 2000 Security Event Descriptions* [6] document posted on Microsoft's support site:

"This article contains descriptions of various security-related and auditing-related events, and information about how to interpret these events. These events will all appear in the Security event log and will be logged with a source of "Security." "

The document is easy to read and gives precise references to particular Event ID generated by security related events.

Although the UNIX community has standardized on the syslog protocol, and the Windows community have standardized on the NT Eventlog format, there continues to be room for improvement.

From a security standpoint for example, UNIX Syslog protocol has known security issues as far as being able to provide verification of the integrity of a syslog message from source to destination logging host.

The Internet Engineering Task Force has set out to correct this. According to *IETF Syslog Issues and Standards* [2] documents posted on the group's web site:

"Beyond documenting the Syslog protocol and its problems, the working group will work on ways to secure the Syslog protocol. At a minimum this group will address providing authenticity, integrity and confidentiality of Syslog messages as they traverse the network. The belief being that we can provide mechanisms that can be utilized in existing programs with few modifications to the protocol while providing significant security enhancements. "

Other examples exist throughout the Internet of individuals, groups and comities trying to work towards the improvement of existing logging technology and the output generated by it.

However, this still leaves the applications for each operating system, or independent devices such as printer or routers that are not required to conform to these formats.

There are many rouge or proprietary log formats that are used by these applications or devices even though they are sending logging detail to the Syslog or Eventlog service, the results are sometimes very cryptic and confusing.

This cryptic output and confusion continues to create a challenge for anyone who is trying to use the logs to understand W5.

**Understanding the Problem of Cryptic logs in Detail**

As we now understand, when an event in a device or application is logged, it generates output that anybody can view in an ASCII text format from almost any computer.

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

Now imagine that your manager comes to you and wants to know "when" the user Bob Smith ("who") last printed a report ("what") named Payroll to the printer HPGL5. ("where")

Because your application generates log data on who, what, when and where events take place it should be no problem for you to answer the question. You open up the log to find the following:

```
1241020401-%u:1032:bsmith:%eid:101:%s:e0xb3a%x:1102
1241020401-%u:1922:rjones:%eid:253:%s:e0xe2z%x:1103
1242020401-%u:1032:bsmith:%eid:021:%s:e0xa3g%x:1102
1243020401-%u:1032:mapple:%eid:221:%s:e0xe3a%x:1111
```

The above is a fictitious log, but you understand what we are trying to illustrate. A simple question doesn't mean there is a simple answer when it comes to logs.

Continue to imagine your frustration as you work your way over to another device holding more logs and find the following:

```
12:41:22:pm 02/04/01-Event 10298 successful!
12:41:38:pm 02/04/01-Error 18244 failed!
12:42:17:pm 02/04/01-Result 198282 generated from Event 10298.
12:43:54:pm 02/04/01-Event 29211 successful!
```

This log appears much easier to read. We can clearly see the date and time the various events occurred, but we still don't have an answer on Bob Smith printing a document called Payroll to a particular printer.

Using a real world example now, the Cisco Intrusion Detection System provides detailed logs on security incidents, events, and alerts.

An example taken from *Interpreting Cisco Secure IDS Log Files* [(3)] from Cisco Systems web site continues to illustrate the point of cryptic log formats:

"A sample event log file entry is provided below. An explanation of each field follows in the table below the log entry.
4,1028109,1998/12/23,22:07:00,1998/12/23,16:07:00,10008,5,100,OUT,IN, 5,8000, 51304,TCP/IP, 131.215.210.2,207.18.164.70,1034,21, 0.0.0.0, hacked in, 75736 … E0D0A"

The web site proceeds to break down the above mentioned log into an easy to read table, but strongly suggests the user read the *Cisco Secure IDS Users Guide* which dedicates well over 100 pages to trying to explain the log output generated by the device.

As the security professional, you will be required to figure out how to break down the log information into a meaningful format so the questions can be answered, you could consider it a form of translating.

Often in the case we discussed with the Managers request, the security professional will report back to the manager with "I don't know, the logs didn't tell me anything."

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

Although this is partly correct, the answers are most likely in the log if we could only figure out how to read it.

**Logs Causing Information Overload**

Another problem, especially in the security arena, are that logs can generate too much information.

If you are trying to track down when a particular event occurred, you are sometimes required to sift through a boatload of information, which on one hand is very important, but on the other is only slowing down your task.

Even if the logs are crystal clear to you but are being generated in granular detail to the $N^{th}$ level, searching for a particular event could become very cumbersome.

Thankfully almost all ASCII text tools have search capabilities to help with this issue, as long as you know what you are looking for.

Although it is important to know W5 on any event, some events may require less detail, as they are not considered critical to helping you understand the event.

Thus it becomes very important for you the security professional, to not only understand logs, but also to be able to successfully manage them in a way that will give you the detail you need without weighing you down.

**Log Management: Consolidating Logs**

Understanding the importance of logs, and that log output can be both cryptic and cumbersome, we need to come up with efficient ways to manage the information.

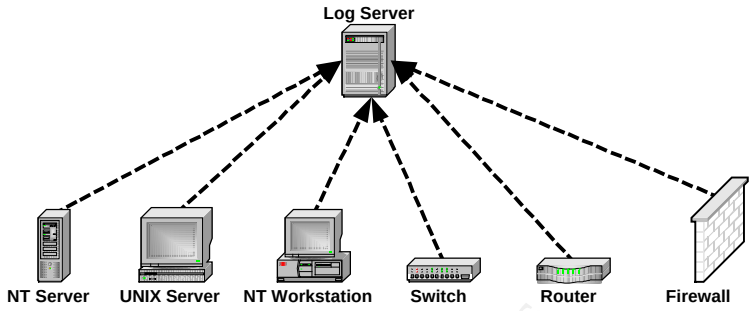
A starting point might be log consolidation.

Log consolidation will allow you to view all logs generated in a quick and hopefully organized manner without having to walk a marathon getting to and from various logs stored around your infrastructure.

There are two scenarios you might consider on how to perform consolidation. Both give you a centralized area for your logs to reside, and both will save you from doing a lot of running around. The following figure shows some examples:

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

**Scenario A** gives you a single centralized server to view and manage all your logs. The log information could be used in a 24x7 Operations Center as a centralized alerting device for console operators.



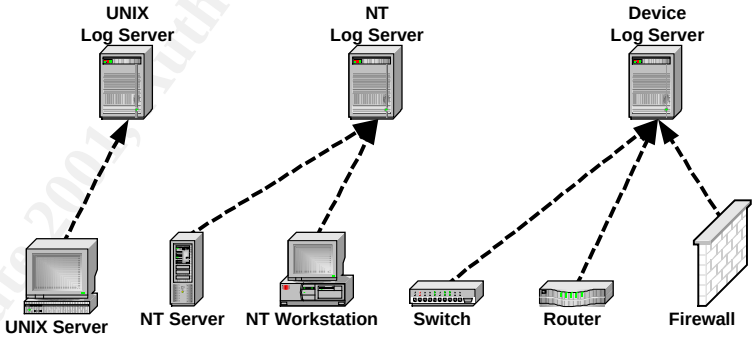

**Logging Scenario A**
All devices report back to a common central logging server.

A lot of storage space is required, and redundant hard drives or a RAID setup make for a more stable system.

The problems with Scenario A are that it creates a single point of failure or attack. If all your logs are stored on one server, then you're left in the dark if it goes offline.

A sophisticated attacker will cover their tracks by removing or corrupting any logs that might have tracked them. Having all the logs in a centralized location makes their job a lot easier, which we do not want.

**Scenario B** uses a more robust method of breaking out the logs to specified servers. One server is used for one type of log source.




**Logging Scenario B**
All similar devices report back to a designated logging server.

This means that if your NT log server is offline, you will not loose the UNIX or Device logs. To further improve this model a good practice would be to have all logs replicate on each server to create a fully redundant mesh.

The logs are still centralized, and can still be used in a 24x7 Operations Center environment. If an attacker wanted to cover their tracks, they would have to compromise all log servers making their job much harder, especially if they want to go unnoticed.

Problems associated with Scenario B are less significant than Scenario A. The biggest issue for some will be the cost for hardware to fulfill the requirement.

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

**Log Management: Log Rotation**

Having a constant consolidated stream of logs is good, but we are still trying to improve the management of the log data to make our lives easier. A common way to achieve this is to break up the log file data with log rotation.

Log rotation is an event that takes place after an interval of time has passed. Common time frames before a rotation occurs are hours, days, weeks, or months.

When a log is "rotated", the log is moved to a new location and renamed to reflect the rotation, with minimal or no interruption of the existing logging processes.

Because logs and rotations are dependant on time, it is very important that your logging hosts and devices have accurate time. To help ensure this, the Network Time Protocol (NTP) service is used.

NTP allows the logging hosts to synchronize with a centralized timeserver, which is in it self synchronized with Global Positioning Satellites (GPS). This combination ensures accurate time within your time zone and is accurate to within a few milliseconds.

By allowing NTP for each device that generates a log, you will also ensure that the data in the logs is time accurate.

CERT Coordination Center from Carnegie Mellon University talk about the importance of using NTP in their web document titled *Manage logging and other data collection mechanisms.* [(4)]

"You can then consolidate logs from different systems by matching time intervals. This will help you gain a network-wide perspective on the activities. To perform this consolidation, you will likely need to merge log files from different systems into a central log file. To avoid having to adjust the timestamps used in each, use a master clock system such NTP (Network Time Protocol) or another time synchronization protocol system. Make sure to take into account different time zones and formats for recorded time."

A successful rotation will generate a number of logs that will contain data for that particular log rotation window. Because there is a "new" log file generated from a rotation, there needs to be a way to effectively slice and name the file so that it will have meaning to the security or systems specialist.

Now that your logging hosts and devices are synchronized with NTP, you can accurately rotate the log and use time stamping techniques to store a particular log rotation window in a meaningful manner.

An example would be taking a log file named NT_Systems.log and slicing it up into files named 06-14-2001-NT-Systems.log, 06-15-2001-NT-Systems.log, 06-16-2001-NT-Systems.log, etc.

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

If an incident or event occurs, you will now have an accurate and efficient way to seek out information on the W5. Your log will be centralized, organized, and time accurate.

**Log Management: Archiving and Securing Logs**

So your now successfully logging all activity in your infrastructure, and you have broken out your logs to several logging hosts, and everything is accurately time stamped.

As you monitor your logging hosts you will notice that they are starting to fill up your storage media quickly, especially if you're in an active environment.

Archiving your logs will help prevent your logging hosts from crashing due to storage limits being reached.

More importantly however, you will have a permanent record of logged events that have occurred for that archives time period.
By having this record, you are in effect storing evidence that could be called up at a later time for use in a legal matter.

Common ways to archive you log data are a similar to those your company's systems specialist use to backup your e-mail or other application data.

Using industry standard archive software, in combination with tape storage devices, network shares, CD-R, CD-RW, or Zip/Jazz drives, you can easily archive the logs.

When you do archive your logs, you better think about the security of the logs, as you now clearly understand the importance and value of a log file.

To help prevent any form of tampering of the data contained within the logs, or to prevent virus activity from effecting or corrupting the logs, the archive should be stored in a read-only status.

A method of archiving your logs in a read-only status would be to use a storage media that can be physically write protected.

CD-R media is often chosen for archiving because the physical media is engineered to be written only once, and cannot be written over. In other words the method for transferring the data to the CD-R is a one-way protectively destructive process.

Other portable media like a Zip disk or CD-RW cannot absolutely be made read only, and therefore are not as resistant to tampering.

There are some concerns when using CD-R media for archiving purposes.

One of the concerns is that CD-R media does not provide any security to the contents on that media. The contents of the CD-R can still be read.

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

In addition, CD-R media has a limited lifespan compared to tape media, something less than 10 years. So if you are planning to store the archive for a long time you might want to make tape archive.

As you have most likely experienced, the CD-R media is easy to damage if mistreated because the surface of the storage is completely exposed. Scratches, dirt and other debris, as well as corrosive materials can all render the media useless.

A final comment on concerns with CD-R media is that media created by one device might not read in another device of the same type. Small incompatibilities, or manufactures using proprietary methods, could cause problems when trying to read or write the media.

Overall even with the concerns outlined, CD-R is still the most cost effective and preferred method to archive logs in a busy environment.
You can mitigate the concern about the contents of the media not being secured for example.

Encryption, such as that used by PGP, will scramble the contents of the logs into a format that an unauthorized on looker could not use, as the contents of the logs are readable, but mean nothing. When the logs are archived on the CD-R, they will remain in this encrypted state as the media cannot be overwritten.

Providing you use strong encryption methods, and safeguard the private keys, your log data will remain protected from prying eyes.

Taking that media and storing it off-site in a physically secured area that is properly controlled is a very good way to mitigate risks in general, and is thought to be a best practice when dealing with archived media of important value such as your logs.

**Legalities of Logs**

Previously, we mentioned that log files could be used in legal matters. In the information age, log files can be used as evidence in courts of law to show events that occurred for a certain timeframe. As long as your logs can be proven accurate and unaltered, they will stand up as written evidence in court within in the US or Canada.

As outlined in the Legal Information Institutes *2001 Federal Rules of Evidence* [(5)] web document explaining this point in legalese:

"Rule 1002. Requirement of Original
To prove the content of a writing, recording, or photograph, the original writing, recording, or photograph is required, except as otherwise provided in these rules or by Act of Congress."

Your company should have written policy that if not written by you, then Human Resources or Legal. The policy will outline what actions will be taken if evidence is required, as sometimes your company might not be the one requesting the information.

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

Law officials such as the FBI, CIA or local law enforcement may come knocking on your door looking for evidence, either directly involving your company, or not.

Your logs could be used in a their legal battle. In addition various logs from third party sources or systems may be used to support your logs or even discredit them.

Therefore having documented due diligence procedures showing the steps taken to ensure your logs are reliable, secured, and accurate, along with good company policy and regular communications with your legal council will give you peace of mind in you hear that knock.

**Conclusions**

The battle to improve our log deciphering skills will continue on well into the future of the information age as more an more devices get wired into the global networks.

Hopefully this document has enlightened you on what a log is, its true purpose, how to better understand and manage logs, and some of the legal issues you might face.

Logs are critical to Information Security professionals, as they are your forensic trail telling you the W5 of an event, managing and protecting the logs should be part of your daily procedures as they are extremely valuable.

---


Merriam-Webster's Dictionary [1]
*http://www.m-w.com/cgi-bin/dictionary*

Internet Engineering Task Force - IETF Syslog Issues and Standards [2]
*http://www.ietf.org/html.charters/syslog-charter.html*

Cisco Systems - Interpreting Cisco IDS logs [3]
*http://www.cisco.com/warp/public/707/31.html*

Carnage Mellon University - CERT Guide to Managing Log files [4]
*http://www.cert.org/security-improvement/practices/p092.html*

Cornell University Legal Information Institute - US Federal Evidence Rules (2001) [5]
*http://www.law.cornell.edu/rules/fre/overview.html#toc*

Windows 2000 Server Security Event Descriptions [6]
http://support.microsoft.com/support/kb/articles/Q299/4/75.ASP


© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

## References

**Books:**

**Incident Response: Investigating Computer Crime (2001)**
Chris Prosise, Kevin Mandia
McGraw-Hill Professional Publishing; ISBN: 0072131829

**Network Intrusion Detection: An Analyst's Handbook (2000)**
Stephen Northcutt
New Riders Publishing; ISBN: 0735710082

**Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet (2000)**
Eoghan Casey
Academic Pr; ISBN: 012162885X

**Internet:**

Interpreting NT Event Logs
http://www.microsoft.com/technet/treeview/default.asp?url=/TechNet/prodtechnol/winntas/support/usesecur.asp

Microsoft - Interpreting Windows 2000 logs for IAS
http://www.microsoft.com/windows2000/en/server/help/default.asp?url=/windows2000/en/server/help/sag_ias_log1a.htm

Understanding TCPDUMP logs
http://www.securityfocus.com/frames/?focus=ids&content=/focus/ids/articles/normaltraf3.html

O'Reily Tools of the Trade on Security Utilities and Logs
http://linux.oreillynet.com/pub/a/linux/2001/06/29/tools_two.html
http://linux.oreillynet.com/lpt/a/linux/2001/07/13/tools_trade_three.html

CERT Guide to Finding Evidence in Solaris Log files
http://www.cert.org/security-improvement/implementations/i003.01.html

Interpreting Cisco PIX syslog messages
http://www.cisco.com/univercd/cc/td/doc/product/iaabu/pix/pix_v53/syslog/pixemsgs.htm

Understanding iPlanet logs
http://docs.iplanet.com/docs/manuals/enterprise/41/ag/esmonsvr.htm#1011987

Understanding how web data is logged with Webtrends.
http://www.webtrends.com/support/tech_log_understand.htm

North American NTP Servers
http://www.eecis.udel.edu/~mills/ntp/servers.htm
All about Network Time Protocol
http://www.eecis.udel.edu/~ntp/

Canadian Evidence Act (2000)
http://laws.justice.gc.ca/en/C-5/text.html

© SANS Institute 2001, As part of the Information Security Reading Room. Author retains full rights.

Last Updated: May 19th, 2013

# Upcoming SANS Training

**Click Here for a full list of all Upcoming SANS Events by Location**

| | | | |
|---|---|---|---|
| **SANS Malaysia @ MCMC 2013** | **Cyberjaya, MY** | **Jun 03, 2013 - Jun 08, 2013** | **Live Event** |
| **SANS Pen Test Berlin 2013** | **Berlin, DE** | **Jun 03, 2013 - Jun 08, 2013** | **Live Event** |
| **Industial Control Systems Security Training - Houston** | **Houston, TXUS** | **Jun 10, 2013 - Jun 15, 2013** | **Live Event** |
| **Security Impact of IPv6 Summit 2013** | **Washington, DCUS** | **Jun 14, 2013 - Jun 16, 2013** | **Live Event** |
| **SANSFIRE 2013** | **Washington, DCUS** | **Jun 14, 2013 - Jun 22, 2013** | **Live Event** |
| **SANS Canberra 2013** | **Canberra, AU** | **Jul 01, 2013 - Jul 13, 2013** | **Live Event** |
| **SANS London Summer 2013** | **London, GB** | **Jul 09, 2013 - Jul 16, 2013** | **Live Event** |
| **Digital Forensics & Incident Response Summit 2013** | **Austin, TXUS** | **Jul 09, 2013 - Jul 16, 2013** | **Live Event** |
| **SANS Rocky Mountain 2013** | **Denver, COUS** | **Jul 14, 2013 - Jul 20, 2013** | **Live Event** |
| **SANS Mumbai 2013** | **Mumbai, IN** | **Jul 22, 2013 - Jul 27, 2013** | **Live Event** |
| **SEC528 SANS Training Program for the CompTIA® New Advanced Security Practitioner (CASP) Certification** | **Washington, DCUS** | **Jul 22, 2013 - Jul 26, 2013** | **Live Event** |
| **SANS San Francisco 2013** | **San Francisco, CAUS** | **Jul 29, 2013 - Aug 03, 2013** | **Live Event** |
| **SANS SEC 560: Network Penetration Testing @ Bangalore 2013** | **Bangalore, IN** | **Aug 05, 2013 - Aug 10, 2013** | **Live Event** |
| **SANS Boston 2013** | **Boston, MAUS** | **Aug 05, 2013 - Aug 10, 2013** | **Live Event** |
| **Critical Security Controls Summit** | **Washington, DCUS** | **Aug 12, 2013 - Aug 18, 2013** | **Live Event** |
| **Industrial Control Systems Security Training - DC** | **Washington, DCUS** | **Aug 12, 2013 - Aug 16, 2013** | **Live Event** |
| **Mobile Device Security Summit 2013** | **OnlineCAUS** | **May 30, 2013 - Jun 06, 2013** | **Live Event** |
| **SANS OnDemand** | **Books & MP3s OnlyUS** | **Anytime** | **Self Paced** |