Log Management and Retention in Corporate Environments

Brittany Wilbert and Lei Chen
Department of Computer Science, Sam Houston State University, Huntsville, Texas, United States

Abstract - *As an increase in the amount of data breaches, hacking attempts and other malicious activities have harmed corporate environments, regulations have arisen to address the concerns of the public in regards to these events. As a result, corporations have to address the need for log management and data retention within their environments. This requires an abundance of resources and policies to ensure that the log data collected meet integrity standards and can be assured to be accurate. A review of (1) why log management and retention is important, (2) the barriers of good log management policies, types of log files, and (3) the future security concerns are surveyed to better understand the needs of the corporate environment.*

Keywords: log, corporate, hacking, policies

1 Introduction

As the migration of large corporations from the physical business space to the Internet has rapidly increased over the last few decades, a greater emphasis on creating accurate audit trails of ingress and egress communication traffic from these corporate environments has become more and more vital. In addition, Federal, State, and International regulations on the transmission of data have now served as warning for corporations to enact reliable policies which follow best security practices needed in the business field in regards to computers. Until recently, log management was held in the background of security concerns. The risk of log file mismanagement was greater as the amount of log files collected increased with the expansion of hardware and software used within computer environments. However, because of law and regulations enacted over the last fifteen years, corporations now have to focus resources to provide log management solutions within their environment.

This paper will first discuss the background of log management and corporate environments. It will then discuss what log management is and why it is necessary from a corporation's point of view to implement these policies using Federal, State and International law. The paper will then transition into a discussion of the similarities and differences between Windows log files and Syslog files. Finally, the paper will discuss how cloud computing will affect log management in the future and give conclusions on the topic.

2 Background

The Syslog protocol was first used in the 1980s in a sendmail application to remotely deliver log messages to a server [1] as a part of the Berkeley Software Distribution of UNIX [2]. This protocol was created to provide the ability to report system events. Since the introduction of the protocol, the formatting used for a Syslog is extensively used within the UNIX environment, as well as other open source and many proprietary software and hardware devices. Microsoft decided to differentiate from the Syslog format by using hexadecimal outputs to store their log messages. As a result of these formatting decisions, the Windows event logs and the logs produced in the Syslog method have significant styling, formatting and log collection methods.

3 The Corporate Environment

The scope and size of the corporate environment must be defined to adequately determine how log management should be used to collect and retain log data. The National Institute of Standards and Technology (NIST) describes recommendations which may be used by computer security personnel, such as program managers, computer security incident response teams, and other individuals who are responsible for log management [3]. The four major recommendations include establishing policies and procedures for log management, prioritizing log management throughout the organization, giving support to all staff with log management responsibilities, and establishing a standard operational process for log management [3].

Because many security incidents are not discovered until after a breach has occurred, log management, data retention and log analysis is vital to determine what has occurred and what can be done to reduce the risk of similar incidents occurring in the future [4].

4 What is Log Management?

Log management is the process in which auditing logs should be parsed, protected, and used within the corporate environment. Log management contains many facets including log retention, as well as log analysis and forensics [4]. Security system logging allows for corporations to hold to confidentiality, integrity and assurance (CIA) practices.

By becoming compliant, a corporation allows itself to guarantee that resources and tools are available to provide evidence to investigators when malicious incidents occur. Also, a log management policy allows for proactive monitoring of trends within an environment [4]. With the knowledge of these trends, corporations can determine if significant changes have arisen.

A log management policy allows for support for internal investigations, data which can be used for forensic investigations, baselines for normal activity within the environment, the ability to identify operational problems, perspective for malicious activity which originates from both internal and external threats, and the ability to meet Federal, State and International laws and regulations [4].

5 Regulations and Data Retention

According to Privacy Rights Clearinghouse, between 2005 and 2012 over 560,000,000 records have been breached [5]. Some of the largest breaches have included Hartford Life Insurance Company, TJX Stores, Sony Corporation of America and the Texas Comptroller's Office [5]. As a result of these data breaches, compliance boards and organizations have been created by both private and federal authorities. Some of the largest include the Payment Card Industry Data Security Standard (PCI-DSS), Sarbanes-Oxley Compliance (SOX), and the Health Insurance Portability and Accountability Act (HIPAA). To give an example of these regulations, the Occupational Safety and Health Administration (OSHA), requires that medical and other similar types of records must be retained for 30 years [6].

These standards and regulation increasingly demand for complete and accurate audit trails to be created in order for companies to be in compliance. For corporations, ensuring that their logging and auditing practices are up to date allows for quicker incident response, forensic investigations and incident responses to occur. Also, since security incidences may not be discovered for weeks or sometimes months after the initial breach has occurred, corporations must make sure that their log records are both secure and up to date [4].

Regulatory agencies and regulations such as PCI-DSS, SOX, and The Federal information Security Management Act of 2002 (FISMA) have required companies and organizations to monitor and audit log files which are created within their infrastructures [7][8][9]. Since the fines and fees for data lost can be tremendous, corporations must give considerate attention towards how they will implement log management and data retention into their environments. Knight recommends that log management and analysis must be used together to protect an environment [10].

6 Types of Log Files

Since many corporations have a mixture of Windows log sources and Syslog sources within their environment, an understanding of the similarities and differences of these logs is vital in log management for security personnel [11][12].

6.1 Windows Logs

During the transition to Windows Vista, Microsoft decided to transition away from the previous format and to include export of the logs into the more universally used .XML format compared to using hexadecimal format of prior Windows NT operating systems. Although there are many types of Windows Event Logs, most failures typically fall into the following categories [4][7][13]:

- *Information*: Occurs when an infrequent but significant event occurs, such as when a Microsoft SQL Server successfully loads.
- *Warning*: Reports a problem that is not significant at the time, but could escalate into a large problem in the future.
- *Error*: Notifies when a significant even has occurred.
- *Success Audit*: A type of security event which notifies that a successful action has occurred. Examples include successful user account logins, or a file being successfully opened.
- *Failure Audit*: A type of security event which notifies when a failed action has occurred. Examples of a failure audit include a failed user account login attempt.

The Windows NT log structure was initiated to allow for smaller log file sizes which can be quickly transferred from one location to another. Once a log is parsed by Windows Event Viewer or other software, it provides a detailed list of information. This event types allow a security personnel to locate why an incident may have taken place.

6.2 Windows Log Example

To show how a Windows Event Log looks like, Windows Security Event ID 4624 will be used. This event is a Success Audit indicating that an account has successfully logged on. We collected this data from a Windows 7 Home Premium Edition virtual machine named ‘TEST-PC’. The parsed message appears below:

SubjectUserSid S-1-5-18
SubjectUserName TEST-PC$
SubjectDomainName WORKGROUP
SubjectLogonId 0x3e7
TargetUserSid S-1-5-21-3639450285-3132740584-3241638508-1000
TargetUserName TestUser
TargetDomainName TEST-PC
TargetLogonId 0x13adfdd
LogonType 7
LogonProcessName User32

AuthenticationPackageName Negotiate
WorkstationName TEST-PC
LogonGuid {00000000-0000-0000-0000-000000000000}
TransmittedServices -
LmPackageName -
KeyLength 0
ProcessId 0x230
ProcessName C:\Windows\System32\winlogon.exe
IpAddress 127.0.0.1
IpPort 0

This log message includes several data points which would be important for an investigation. There is a large amount of information you can obtain from this log file. First, the SubjectUserSid, S-1-5-18, is a well known security identifier (SID) which indicates that the computer used a service account for the logon [14]. The TargetUserSid is the target computer account that the logon was requested and granted. In this example, the target is the local computer.

The most important information related to this type of log is the LogonType, the LogonProcess and TargetLogonId. The LogonType is a hexadecimal code which explains the location in which the login attempt originated. In the example above, the reason for the failure was type 7, which is when the workstation is unlocked from the password protection screen [15]. Next, the User32 is the logon process which was started to initiate logon [16]. The User32 process is used for interactive logons, which is the logon screen which appears either after Windows 7 is started or when a user account is locked. Finally, TargetLogonId describes the target location in which the logon was granted. This value is a semi-unique value which is reset each time the workstation is restarted [15]. Therefore, this value can be used to track the actions of the user after they have logged into the workstation.

While log types differentiate based on event type, this method of investigation can be used to locate how and when events have occurred on a workstation.

6.3 Syslog

Most varieties of *nix and UNIX operating systems continue to use the Syslog as their primary logging type. Syslog data is a plaintext string which can then be tokenized by either a script or program to be human-readable. These logs are usually collected into several log files which are disbursed into several areas of the operating system. By default, RedHat/Fedora and Ubuntu/Debian systems generally place these files in the ~/var/log/messages directory [17]. Rainer Gerhards defines Syslog as being a non-standardized formatting system which allows log emitters to decide or configure which types of log messages will be produced [18].

Although there is not a standard format, Syslog is in general a fairly readable format. Since Syslog is free of formatting, scripts and parsers can be used to create easy to read reports. A 'Syslogger', a centralizing collector of Syslog data which is collected from configured devices within an environment, can be used to collect Syslog data [2][19]. When a corporation discusses bring in external products for their environments, each products should be compared to test the ability to provide accurate data, reliable results, and any other criteria the environment needs to provide evidence of their compliance to regulators and auditors [1]. Purcell discusses a case study which he created to compare correlation rule sets as well as other criteria which could be compared to other products [2]. Using methods such as this allows corporations to provide reliable evidences of what has occurred in their environment.

6.4 Syslog File Example

Since Syslogs are stored in plaintext, either a GUI system log viewer or accessing the /var/log/messages folder can be used to view the files. We collected this data from log messages originating from an Ubuntu 10.4 (Lucid Lynx) virtual machine using the method described by Gite [17]. The computer name for this account is ‘testtoor’. This data can then be parsed into a more human-readable format, which is shown in Table I.

TABLE I
SYSLOG AUDIT FORMATTING

Date	System Name	Failure Type	Alert Type	Reason
Apr 01 07:44:12	testtoor	exim	ALERT	/var/log/exim4/paniclog has non-zero size, mail system possibly broken
April 10 10:00:12	testtoor	passwd	password	successfully changed

Since Syslogs are stored in plaintext, either a GUI system log viewer or accessing the /var/log/messages folder can be used to view the files. We collected this data from log messages originating from an Ubuntu 10.4 (Lucid Lynx) virtual machine using the method described by Gite [17]. The computer name for this account is ‘testtoor’. This data can then be parsed into a more human-readable format, which is shown in Table I.

6.5 Comparison of Windows Log Files and Syslog

As mentioned previously, the differences between Windows Event Logs and Syslogs are extensive. To show these differences, Table II shows the differences between these log file types.

Windows event logs and Syslogs do have a few similarities. There are a few events which have similar wording, such as Alerts/Warning or password change notifications. Also, both types use files to store log events, as well as only allow administrative users to view the log directories by default.

TABLE II
DIFFERENCES BETWEEN SYSLOG AND WINDOWS LOG FILES

Differences	Log Types	
	Windows (NT Operating Systems)	**Syslog**
Location (Default)	%SystemRoot%\System32\Config	var/log/messages
Format	Hexadecimal, .xml	Plaintext
Format Standarized?	Yes	No
Used Elsewhere?	No (Microsoft Proprietary Systems)	Yes

7 Future Security Considerations

As with any management consideration within the technology environment, a need to discuss potential future obstacles which may cause additional implementation changes must be discussed. One of the more rapidly growing fields is the cloud computing. Cloud computing is the sharing of resources, hardware, and software as a means to provide services to a wider range of customers with a decrease in the amount of resources required by corporations and other organizations [20]. As a result, cloud computer log management may be implemented as a Software-as-a-Service (SaaS) instead of the traditional hardware techniques [20].

Although cloud computing will allow for greater ease in allowing for network resources to be pooled in a easier method, this change also effects more traditional businesses which are required to implement a logging solution. Cloud computing requires significant amounts of preparation before implementing a log management solution [20]. Challenges corporations may run into as they attempt to collect and retain log data which is collected from cloud environments must be discussed by a corporation's security personnel and corporate management before hardware and/or software solutions are implemented [20]. Also, corporations may need to consider a SaaS provider who may have applications which can be used to create a logging infrastructure for log file analysis.

8 Conclusion

Log management, collection, and reporting have become a significant part of business in the 21st century. As a greater number of corporations enter or expand their Internet market, rules and regulation will continue to expand the amount of auditing required to protect consumers. Both the Windows and UNIX environments have robust log creation infrastructure which allow for security personnel to locate potentially malicious activity.

As a greater emphasis for accurate log collection is pushed, corporations must strive to include better and more efficient log collection tools. These tools must include the ability to expand based on regulations, data retention requirements, and other considerations such as cloud computing. The process of constructing the framework of a log management and data retention policy cannot be ignored by organizations.

Because of the greater risk of proprietary data, employee data, and customer data being stolen, deleted or changed by both inside and outside threats, corporations must prepare themselves by thoroughly analysing the policies they have, as well as analysing the policies they may need to continue to be compliant and secure. Because of this, log management and data retention will be more and more relied upon as the secondary measure used to locate when a security incident has occurred as well as the forensic proof needed to show that a crime has been performed within a corporation's environment.

With all of these considerations to keep in mind, corporate security personnel must be knowledgeable of new standards within the field while continuing their education of log management, data retention, and log analysis practices. These personnel must also communicate these trends to management. When selecting a log management procedure, a holistic approach must be implemented. This includes employing personnel who are responsible for reviewing log data on a daily basis, regularly monitoring how this log data is stored, and making sure that log data is not lost due to changes within the organization or movement of the data to external sources.

Although log management is a difficult task, and one of many that a corporate environment must monitor, the rewards for keeping a complete log management system show that it is a necessity. If a corporation does not include a log management system, the risk of substantial financial lost due to malicious activity increases significantly. Therefore, log management is a vital component in the security practices used by corporate environments.

9 References

[1] D. Casey, "Turning log files into a security asset," *Network Security*, pp. 4-7 © 2008 Elsevier Ltd.. doi:10.1016/S1353-4858(08)70016-3.

[2] J. Purcell, "Log Analyzer for Dummies," December 2007. Available: http://www.sans.org/reading_room/whitepapers/logging/log-analyzer-dummies_2031.

[3] K. Kent and M. Souppaya, "Guide to computer security log management," September 2006. Available: http://csrc.nist.gov/ publications/PubsSPs.html.

[4] M. Gorge, "Making sense of log management for security purposes – an approach to best practice log collection, analysis and management," *Computer Fraud & Security*, May 2007, Issue 5, Vol. 7. pp. 5-10, © 2007 Elsevier Ltd.. doi:10.1016/S1361-3723(07)70047-7.

[5] Privacy Rights Clearinghouse, "Chronology of Data Breaches," April 2011. Updated May 6th, 2011. Available: http://www. privacyrights.org/data-breach.

[6] B. Wrozek, "Electronic Data Retention Policy," SANS GIAC - GSEC Security Essentials. Version 1.2e. March 2008. Available: http://www.sans.org/reading_room/whitepapers/backup/electronic-data -retention-policy_514.

[7] R. F. Smith, "Bridging the Gap Between Native Active Directory Auditing & Successful Compliance," 2011. Available: http://www.ultimatewindowssecurity.com/tools/ondemandlm/BridgingTheGap.pdf.

[8] PCI Security Standards Council, "Payment card industry (PCI) data security standard", 2006. Available: https://www.pcisecuritystandards.org/pdfs/pci_audit_procedures_v1-1.pdf.

[9] SOX-Online, "Sarbanes-Oxley Roadmaps & Basic Approaches", 2006. Available: http://www.sox-online.com/approaches.html.

[10] E. Knight., "Investigating digital fingerprints: advanced log analysis" *Network Security*, pp. 17-20, © 2010 Elsevier Ltd.. doi:10.1016/S1353-4858(10)70127-6.

[11] I. Eaton, "The ins and outs of system logging using Syslog", October 2003, Available: http://www.sans.org/reading_room/whitepapers/ logging/ins-outs-system-logging-syslog_1168.

[12] K. E. Nawyn, "EVTX and Windows Event Logging". November 2008. Available: http://www.sans.org/reading_room/whitepapers/logging/ evtx-windows-event-logging_32949.

[13] MSDN. "Types of Event Log Entries." *Microsoft Development, Subscriptions, Resources, and More.* 2012. Available: http://msdn.microsoft.com/en-us/library/zyysk5d0(v=vs.71).aspx.

[14] Microsoft Support, "Well-known security identifiers in Windows operating systems," May 9, 2011. Available: http://support.microsoft.com/kb/243330.

[15] R. F. Smith, "4624: An account was successfully logged on," 2012. Available: http://www.ultimatewindowssecurity.com/securitylog/ encyclopedia/event.aspx?eventid=4624.

[16] Microsoft Support, "Auditing User Authentication," February 23, 2007. Available: http://support.microsoft.com/kb/174073.

[17] V. Gite, "View log files in Ubuntu Linux," nixCraft: Insight into Linux Admin Work, August 2007. Available: http://www.cyberciti.biz/faq/ubuntu-linux-gnome-system-log-viewer/.

[18] R. Gerhards, Finding the Needle in the Haystack, 2005. Available: http://www.monitorware.com/en/workinprogress/Needle-in-Haystack.asp.

[19] K. Nawyn, "A Security Analysis of System Event Logging with Syslog," May 2003. Available: http://www.sans.org/reading_room/whitepapers/logging/security-analysis-system-event-logging-Syslog_1101.

[20] R. Marty, "Cloud Application Logging for Forensics," *ACM Symp. on Applied Computing* 2011 © ACM. doi: 10.1145/1982185.1982226.