ALERT LOGIC®

LOG MANAGEMENT: BEST PRACTICES

TABLE OF CONTENTS

WHY LOG MANAGEMENT?

Most organizations today are experiencing an ever-expanding onslaught of security threats. These threats are changing and multiplying exponentially, making it near impossible for IT security teams to outmaneuver threats to their environments. And alongside this, they are also facing significant fines from governmental and industry regulations, should they fail to meet compliance requirements. Yet, in spite of these new challenges, organizations' IT security budgets are still being slashed in favor of business-continuity investments.

Unfortunately, cutting IT budgets for the sake of business is inherently problematic, as the evolution of threats requires an expansion of security systems, not a reduction. For instance, setting up systems like log management was a somewhat simple process in the past, but today's complex networks and regulatory requirements demand a more complex log management setup. Log management now extends beyond simple data collection; it encompasses normalization, analysis, reporting, and disaster-proof archival processes.

Now, with the expansion of IT infrastructure into hosted and cloud deployments, there is not only more data to manage, it also resides in a variety of environments. Trying to collect and manage a continuous supply of distributed log data can quickly overwhelm an IT organization. Adding storage sounds simple in concept, yet the cost of purchasing and managing terabytes of storage can be staggering.

With these emerging challenges in mind, this paper will discuss best practices for log management, focusing on several key areas:

- Collecting the appropriate data: Consider all the sources of log data in your environment, as well as those required to meet compliance mandates, alert you to suspicious activity, and provide valuable forensic data.
- Making log data usable in a normalized, searchable format
- Reviewing and analyzing log data regularly: Log data will not help you achieve your goals if it is not examined regularly. For compliance purposes, this is a requirement.

WHICH LOGS SHOULD BE COLLECTED?

In order to meet regulatory compliance standards, all logs must be collected—not just the security logs. For example, operating system logs and application logs often contain security-related information as well as information about events that may not initially appear security related. Organizations must consider the potential value of each and every possible log source.

The following log types should be considered for collection.

ANTI-MALWARE SOFTWARE

Examples of anti-malware include anti-virus, anti-spyware, and rootkit detectors, to name a few. These logs may include information like indicators that malware was detected, disinfection attempt results, file quarantines, when file-system scans were last performed, when anti-virus signature files were last updated, and when software upgrades have taken place.

APPLICATIONS

The information logged by various applications can vary widely and may include account changes, user authentication attempts, use of privileges, usage details, client and server activity, configuration changes, major system failures, and more. Application logs can be more valuable when network communications are encrypted; however, application logs are often proprietary formats.

AUTHENTICATION SERVERS

Directory servers and single sign-on servers will typically log each and every authentication attempt showing the originating user ID, destination system or application, date and time info, and success/failure details.

FIREWALLS

Some firewalls are perimeter-focused and general in nature, while others are very application-specific or single-host (personal) focused. Firewalls cannot only block activity based on policy, they can also inspect content and ensure the state and integrity of permitted connections. Firewalls have numerous capabilities, and their logs can be very detailed and informative.

INTRUSION PREVENTION SYSTEMS

These systems record detailed information about suspicious behavior and detected attacks as well as actions taken to halt malicious activity in progress. Some intrusion protection systems, such as file integrity systems, run periodically instead of continuously, so they generate logs in batches rather than an ongoing basis.

NETWORK ACCESS CONTROL SERVERS

Network access control can operate for both internal and external hosts connecting to the internal network. At the time of connect, the hosts' security posture is determined, and hosts failing to adhere to the defined policy are quarantined onto a separate VLAN (Virtual Local Area Network) segment. NAC servers log a great deal of useful information about both successful/permitted and unsuccessful quarantined network connections.

NETWORK DEVICES (ROUTERS, SWITCHES, ETC.)

Routers can be configured to block certain types of traffic. Network devices can be configured to log very detailed connection activity but are typically configured to log very lightly. These logs can contain very informative network communication activity.

OPERATING SYSTEMS

There are many varied operating systems on servers, workstations, and assorted network devices. The host administrator typically controls logging. The types of events, as well as whether to log only successful or only failed events, or both, can be controlled.

These log entries typically contain information about service starts and stops, authentication attempts, file accesses, security policy changes, account changes, permission and privilege changes, and use of privileges. Operating system logs can also contain information from security software and system applications. They are often beneficial for identifying suspicious activity involving a particular host.

REMOTE-ACCESS SOFTWARE

Virtual Private Networks (VPNs) are the most popular type of secured remote access solutions and they log both successful and failed connection attempts. They record details such as the date and time each user connects and disconnects, as well as the types and amount of data sent and received during the connected session.

VULNERABILITY MANAGEMENT SOFTWARE

Included here are both vulnerability scanning and patch management software. These typically run on an occasional basis and log batches of log entries that include information about scanned hosts/devices including: configuration, missing software updates, vulnerabilities identified, and patch/scan currency downloads, among other things.

WEB PROXIES

Web proxies are the intermediate hosts through which websites are accessed and can be used to restrict web access as well as add a layer of protection between the user and external websites. Web proxy logs record user activity and URLs accessed by specified users.

Each and every type of log will contain varied information, and this information is in different formats. Depending on the circumstances, different log sources can be of more or less value. It should also be noted that if administrative privileges are not properly maintained and the logs secured, then the logs could be manipulated or altered. It is important to understand and limit such privileges and access to logged data as well.

LOG MANAGEMENT CHALLENGES

Many compliance mandates, such as PCI DSS 3.1, require not only that you collect all logs, but also that they be reviewed regularly, are searchable, and are stored in their original, unaltered, raw form for mandate-specific timeframes.

Logs can also be extremely useful in identifying security incidents, policy violations, fraudulent activity, and operational problems shortly after they occur. They are extremely valuable when performing audits, forensic analysis, internal investigations, establishing baselines, and identifying operational trends and long-term problems. However, the infinite variety of log data formats makes it impossible to utilize the data without data normalization.

It is reasonable to assume that the variety of log data sources and the volume of data will always increase. Compounding this challenge is the variability of data formats and distributed nature of these sources; in addition, every network infrastructure is in a constant state of change, with new systems, applications, users, and devices every day of the year.

This creates a variety of specific challenges for log management efforts. These challenges can be broken down into three areas: collection, analysis and review, and archival.

COLLECTION

When we discuss log data, we are discussing a wide and ever-changing range of data sets that must be accounted for.

- Log data is varied. Not only do systems, applications, and network devices have their own logs with varying types of specific data that are captured, but a single log source can have multiple logs to be captured. For example, applications often have multiple log files, each containing a specific type of data.
- Log data sources are distributed. Data sources may be located within internal on-premises infrastructure, collocated in a data center, hosted with a managed hosting provider, or located in the cloud. The infrastructure may be managed separately or in a hybrid environment. Log collection must span all of these environments.
- Log data sources change constantly. A new system, application, or network device may be brought online at any time, and begin generating new log data. Cloud instances may be launching for days or hours and then terminating. A log management solution must account for these changes, or else 100% log collection will not be possible. Otherwise, an organization risks discovering that a log source has not been collected after weeks or months, possibly in response to an auditor's questions.
- Log data may contain sensitive information, such as excerpts from emails, user names and passwords. This raises security and privacy concerns that necessitate proper log data security. Logs improperly secured when being transported to any centralized collection system are susceptible to intentional or unintentional alteration or destruction.

ANALYSIS AND REVIEW

Analysis and review of log data presents two significant challenges: regular review of log data, and the varying

formats of log data.

Regular log review is a valuable practice for any organization, and is a requirement of many compliance mandates. Typically, system administrators have been responsible for reviewing and analyzing log data, but this has usually been a lower priority than other activities, such as more strategic business projects. Rapid-response situations, such as performance issues, vulnerability remediation, and security incident response and investigation, also tend to take priority over log review. The result? Log management projects are never started, or at best, linger unfinished.

Log contents vary enormously. Some logs are designed for humans to read and others simply are not; some logs use standard formats, while others use proprietary formats. Some log formats are comma separated, some are space delimited, and still others use symbols or other character delimiters between the fields within a single log message.

Each log entry, or message, contains certain defined pieces of information, such as a host IP address or username. Each log source records the pieces of information that it considers important; consequently, it can be extremely difficult to link different log sources because they may or may not contain common values.

Even when two sources record the same values, they may be recorded in different and varied log messages. And even when they record the same values, they may represent them differently. For example, a date may be formatted MMDDYYYY, MM-DD-YYYY, or DD/MM/YYYY.

Deciphering dates in various formats may be simple for a human reviewer, but consider the example of an FTP (File Transfer Protocol) being recorded by one log source as "FTP" and another as "21," its well-known port number. Very few analysts can easily distinguish the full 1,024 well-known ports by port number.

One approach to dealing with this complexity is to create PERL scripts to search and produce only those log messages matching the query. This is a reasonable approach conceptually, but with the growing complexity and variety of sources, its ability to alleviate the problems of manual log review is limited.

ARCHIVAL

Log data must be treated like any other organizational data, subject to security and retention policies, as well as compliance mandates. Because it often contains sensitive data (such as customer data), breach of log data is a serious problem. As a result, protection of log data—both in transit to the log collection solution and when stored—is an important concern.

This means that access to log data must be strictly controlled, and under no circumstances should log data be alterable.

In addition, storing log data centrally from distributed sources across an organization creates a massive storage management challenge. Purchasing and deploying the required storage consumes valuable real estate and power (both for operations and cooling) and must be managed, backed up, and included in disaster-recovery planning.

AUTOMATED LOG MANAGEMENT

As the challenges of log management have grown, so too have the benefits of automated log management solutions. An appropriate log management solution provides many benefits to an organization:

- Log collection across all IT infrastructure—on premises, hosted, and in the cloud
- Sophisticated parsing of logs to enable data analysis from a variety of log sources
- Reporting tools that provide insight into your organization's security posture
- Tools to enable post-incident analysis of log data
- Reliable, regular review of log data that meets compliance mandates as well as security best practices.

The cost of log management tools and services must be weighed against the internal staff time required to attempt log management, as well as the cost of non-compliance, data loss, and security incidents.

- Log management solutions should be evaluated against the practices described in this paper.
- Does the solution provide complete log collection across all sources, in all environments?
- Is log data parsed and normalized to support the required search and analysis functions?
- Is regular log review that meets internal requirements and compliance mandates provided?
- Is data transmitted and stored securely?
- Can data be archived according to organizational retention policies, with appropriate levels of data protection?

SUMMARY

While compliance initiatives often drive the need for log management, there are a myriad of security- and availability-related benefits as well. As for compliance, there are many governing regulations and standards, most-notably PCI DSS 3.1, Sarbanes-Oxley, HIPAA, GLBA, and FISMA, which require log collection, retention, and access for forensic analysis.

Benefits from routine log analysis include improved detection of security incidents, policy violations, fraudulent activities, and operational problems. Logs are also useful for establishing performance baselines, performing auditing and forensic analysis, supporting internal investigations and identifying operational trends and long-term problems.

Log management has evolved into a complex undertaking with substantial consequences for improper implementation. With today's complex IT environments and volumes of data, most companies find that a vendor-managed solution is the ideal choice. This gives them access to security experts, continuous 24x7 active monitoring, compliance, and the latest log management implementations without costly investments or installation delays. This frees up IT professionals to focus on improving their businesses and providing superior customer service.

ABOUT ALERT LOGIC

Alert Logic, the leader in security and compliance solutions for the cloud, provides Security-as-a-Service for on-premises, cloud, and hybrid infrastructures, delivering deep security insight and continuous protection for customers at a lower cost than traditional security solutions. Fully managed by a team of experts, the Alert Logic Security-as-a-Service solution provides network, system and web application protection immediately, wherever your IT infrastructure resides. Alert Logic partners with the leading cloud platforms and hosting providers to protect over 3,300 organizations worldwide. Built for cloud scale, our patented platform stores petabytes of data, analyses over 400 million events and identifies over 50,000 security incidents each month, which are managed by our 24x7 Security Operations Center. Alert Logic, founded in 2002, is headquartered in Houston, Texas, with offices in Seattle, Dallas, Cardiff, Belfast and London. For more information, please visit www.alertlogic.com.

ALERT LOGIC®