IT Security Procedural Guide: Web Server Log Review CIO-IT Security-08-41

Revision 4

March 30, 2020

Office of the Chief Information Security Officer

VERSION HISTORY/CHANGE RECORDS

Change Number	Person Posting Change	Change	Reason for Change	Page Number of Change
		Revision 1 April 1, 2015		
1	John Sitcharing	Changes made throughout the document to reflect FISMA, NIST and GSA CIO P 2100.1 requirements.	Updated to reflect and implement various FISMA, NIST and GSA CIO P 2100.1 requirements.	Various
		Revision 2 – April 7, 2016		
1	Bo Berlas/ William Salamon	Extensive changes	Evolving threat landscape	Numerous
2	Ron Wilson	Various formatting and content changes.	Shareholder comments.	Various
		Revision 3 – March 13, 2018		
1	William Salamon	Minor changes	Scheduled update.	Various
2	Bryon Feliksa	Updated format, structure, Federal regulations, and guidance.	Incorporate most current Federal regulations, NIST guidance, and GSA requirements.	Various
3	William Salamon	Updates to clarify procedures	Feedback from stakeholders	Various
		Revision 4 – March 30, 2020		
1	Lambardo	Minor changes	Scheduled update.	Various

Approval

IT Security Procedural Guide: Web Server Log Review, CIO-IT Security-08-41, Revision 4 is hereby approved for distribution.

X DocuSigned by:
Bo Berlas
ED717926161544E...

Bo Berlas
GSA Chief Information Security Officer

Contact: GSA Office of the Chief Information Security Officer (OCISO), Security Engineering Division (ISE) at SecEng@gsa.gov.

Table of Contents

Table of Figures and Tables

1 Introduction

Web applications are critical to both the mission and the defense of the enterprise. By nature, web servers provide services directly to users over the HTTP and HTTPS services ports 80, 8080, and/or 443. Depending upon the application, servers accept requests from within General Services Administration (GSA) networks or from external internet users. Full-featured web applications are often connected to databases that may store sensitive information. The combination of accessibility and high value data present a rich target for attackers. Most users access web services using a standard browser, but attackers can use a variety of custom tools to send carefully crafted requests designed to break the server and its security defenses. Although security technology and web server configuration provide some protection for web servers, routine monitoring of these defenses is essential. This guide outlines some simple steps to help a reviewer parse through web server logs and understand what signs to look for during their review.

Regular review of web logs has many benefits to the management of both the security of web resources and system performance. When a network is experiencing slowdowns or other anomalies, log data can help provide insights into the cause of the problem. A large amount of useful data is generated in the form of web server logs. Because of the volume, format and complexity of logs, it is often difficult to obtain actionable knowledge from web server logs.

For additional information and guidance, please contact the appropriate security staff in the Office of the Chief Information Security Officer (OCISO) as identified in Appendix C.

1.1 Purpose

This guide is designed to provide an overview of how to conduct periodic web server log review that is integral to web system operation and security oversight. It does not address the specific needs of Enterprise-wide log analysis systems that aggregate logs from many servers. The guide discusses summary and detailed views of log content. It describes the common formats of Apache and Microsoft Internet Information Services (IIS) web log entries. It proposes a methodology for the task of reviewing logs for malicious or suspicious activity. Appendices point the reader to further information, software for summary log inspection, and an ASCII table to decode hexadecimal that can sometimes be found in headers.

1.2 Scope

All GSA employees and contractors with Information Technology (IT) and/or IT Security (IS) responsibilities including but not limited to GSA Operations Staff and Information System Security Officers (ISSOs) tasked with reviewing web server logs of isolated systems that may have evidence of security events, must become familiar with this guide. While the information presented may be informative for Enterprise log review, it does not address the correlation of events that is an important part of Enterprise log analysis tools. It does not address non-web logs that are generated by operating systems and database management systems.

1.3 References

Note: GSA updates its IT security policies and procedural guides on independent biennial cycles which may introduce conflicting guidance until revised guides are developed. In addition, many of the references listed are updated by external organizations which can lead to inconsistencies with GSA policies and guides. When conflicts or inconsistencies are noticed, please contact ispcompliance@gsa.gov for guidance.

The following references provide guidance, tools, and additional information on the subject of log reviews.

Federal Guidance:

- Federal Information Processing Standards (FIPS) Publication (PUB) 199, "*Standards for Security Categorization of Federal Information and Information Systems*"
- National Institute of Standards and Technology (NIST) Special Publication (SP) 800-53, Revision 4, "*Security and Privacy Controls for Federal Information Systems and Organizations*"
- NIST SP 800-92, "*Guide to Computer Security Log Management*"
- NIST SP 800-95, "*Guide to Secure Web Services*"

GSA Guidance:

- GSA Order CIO 2100.1, "*GSA Information Technology (IT) Security Policy*"

The guidance documents below are available on the GSA IT Security Procedural Guides InSite page.

- CIO-IT Security-01-02, "*Incident Response (IR)*"
- CIO-IT Security-01-08, "*Audit and Accountability*"

Additional Resources:

- Microsoft - Internet Information Services (IIS) 10
- SANS - Top 5 Essential Log Reports
- SANS - Critical Log Review Checklist for Security Incidents
- HTTP Status Code Definitions
- Open Web Application Security Project (OWASP) Guidance
- Tenable Nessus
- Splunk
- AWStats
- WebLog Expert

2 Overview

There are two aspects to web server log inspection. The first involves reviewing summary information that records general trends. This review leverages log analysis tools that can sort log data and provide various summary views. The summaries should group data from the review period in order to highlight unusual events. By providing many different views of the log data, a good analysis tool increases the likelihood anomalies will be noticed.

The second aspect of web server log review is the detailed inspection of one or more log entries. This requires detailed knowledge of the log format and the meaning of each field. Web server log entries are often cryptic and familiarity with both the log format of the specific web server and the ways in which malicious misuse presents itself is critical.

In both cases the most valuable knowledge is gained through frequent inspection which should happen at least weekly. Weekly inspections should involve the comparison of summary information with summaries from previous periods. Similarly, detailed log entries should be checked regularly with the objective of becoming familiar with normal entries, further enhancing the ability to identify inappropriate or unusual activity.

Systems that have been designated as FIPS 199, "*Standards for Security Categorization of Federal Information and Information Systems*" Moderate or High Impact are subject to continuous monitoring. Continuous monitoring does not imply true, real-time 24x7, non-stop monitoring and reporting. Instead, it means implementing monitoring and oversight processes that provide a clear picture of security state at a given time, while also providing a mirror of control effectiveness over time.

Automated log reviews can be set to occur on a reoccurring basis, such as every 5, 10 or 15 minutes; every hour or every day; and log data can be collected from the central manager at regular intervals. Information needed to monitor critical data, as well as the data processing resources and their controls, should be continuously collected. System administrators and/or security engineers must inspect the output of these tools at least weekly. However, the logs should be reviewed more frequently if the risk to the information and computing resources make it necessary. Some security events may require immediate action. For example, any unauthorized changes to system configuration must be reported in near real-time and corroborated with other system information to check for authorization.

GSA IT Security Procedural Guide: CIO-IT Security-01-08, "*Audit and Accountability*" contains specific control requirements regarding auditing/logging from NIST SP 800-53, Revision 4, "*Security and Privacy Controls for Federal Information Systems and Organizations.*" For additional information and guidance, refer to Appendix C for contact information.

2.1 Summary Views

A summary of the logs gives the reviewer a high-level view (in graphical or other condensed form) of an otherwise large, complex and confusing dataset. The most common views graphically plot some attribute over time. With enough attributes, these graphs could help to discern that a particular type of event may or may not be considered a normal/non-ominous occurrence at a particular time. An example of this might be the number of bytes sent. By plotting this against time, it could be obvious that a one-time, large transmission (e.g., 60

Mbytes) at 2:47 am on a Sunday morning might stand out as requiring an explanation, even though the same transmission might be perfectly normal at another time.

Other views may be used to find correlations that are not apparent any other way. A web site that serves a fixed set of users (AR/AP, HR, etc.) may have very consistent use by users. A change in the list of the Top 20 users or Top 20 web pages might indicate unusual activity that should be investigated.

Threats come from both inside and outside the organization. External facing web servers should identify where the users of the hosted web site are located. A large change in the country of origin of requests might have significance to GSA or other Federal agencies.

Below is a list of time slice views and Top 20 type views that might be useful.

1. Time slice views (hour/day/week)
   a. Hits
   b. Bandwidth
   c. HTTP “Method” (GET, PUT, etc.)
   d. Country of Source
   e. Average or total bytes
   f. Errors types
2. Top 20’s
   a. Users
   b. IP address of source
   c. Pages served
   d. Entry Points
3. Signature “hits” from corresponding network monitoring tools
   a. Heuristic matches
   b. Regular expression matches
   c. SNORT (or other IDS) rule matches

The above list is not exhaustive and should be adjusted depending on the type of server and the analysis software used. Appendix A lists some log analysis software tools that can be used to generate various summary views of web server log files.

When reviewing the web logs from a particular server, it is often helpful to correlate the log entries against event logs generated by network tools such as application firewalls (e.g. Palo Alto) and intrusion detection systems (e.g., Security Onion). These network-based tools often assign risk scores to certain events or collection of events. By examining high-risk network activity to/from a particular web server, the analyst can focus on specific dates and times when reviewing web logs from that server. The GSA Security Operations Division (ISO) maintains the network security tools, including our Security Information and Event Manager (SIEM) [Enterprise Logging Platform (ELP)], GSA’s Security Onion instance, and Palo Alto devices. See Appendix C for ISO Division contact information.

2.2 Detailed Views

The second way of inspecting web server logs is to look at the individual entries in the log and become familiar with what the different fields mean. Examination of the logs themselves can be tedious unless the reviewer understands what is normal and what is not. The files tend to be very large, so to help in this process the reviewer should learn the format of the logs from his/her servers and make the best use of search tools in detailed log inspection.

2.2.1 Format

Logs have many different formats and viewing the raw logs can be challenging. However, there are standard log formats for Apache and Microsoft IIS web servers that contain the same or similar information. Identifying the format can be tricky, but with a little research on the internet and an hour or so of looking at log files, the pattern will become clear. To speed up this process, an annotated description of log entry formats for both Apache and Microsoft IIS web server logs are presented below. Please note that there is no single type of Apache or IIS web server log entry, so this is useful only as an example, and not as a reference.

2.2.2 Apache Logs

Apache log configuration is located in a file called “conf/httpd.conf” in the root apache installation folder. Instructions for log configuration are embedded in the file. There are several log format samples available in the configuration file.

Below is a log entry in the Combined Log Format (CLF).

```
127.0.0.1 - eric [10/Oct/2007:13:55:36 0700] "GET /index.html HTTP/1.0" 200 2326

"http://www.example.com/eric.html" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT
5.1; .NET CLR 1.1.4322)"
```

This log entry can be parsed into the following fields:

127.0.0.1: the IP address of the client

“-“: The hyphen in the output indicates that the requested piece of information is not available. In this case, the information that is not available is the identity of the client.

eric: This is the userid of the person requesting the document as determined by HTTP authentication.

[10/Oct/2007:13:55:36 0700]: The time that the server finished processing the request.

"GET /index.html HTTP/1.0": The request line from the client.

200: This is the status code that the server sends back to the client.

2326: This entry indicates the size of the object returned to the client, not including the response headers.

"http://www.example.com/eric.html": The referring webpage.

> **"Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; .NET CLR 1.1.4322)":** The User Agent HTTP request header.

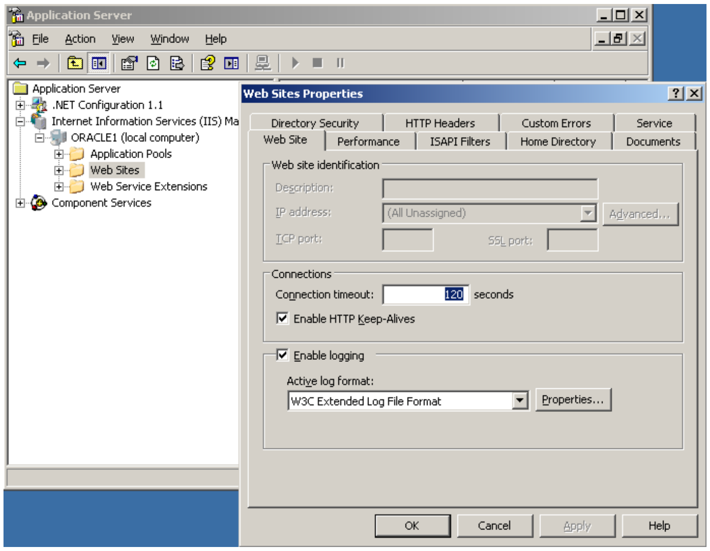
This is represented in some documentation as:

> LogFormat = "%host %other %logname %time1 %methodurl %code %bytesd %refererquot %uaquot"

There are various other formats for Apache logs as well. GSA IT recommends use of the Combined Log Format for Apache servers. There are additional capabilities that log the total number of bytes received and transmitted in each request. This however is specific to the implementation of Apache and should be modified only by the technical support engineers.

2.2.3 Microsoft IIS Logs

IIS logs are located in the “inetpub\logs\logfiles” directory. They are configured using the IIS Manager in the Administrative Tools pop-up menu via a file named “iis.msc.” The most comprehensive logging is available using the W3C Extended Log File Format. GSA IT recommends using this format for Microsoft IIS Web Servers (See Figure 2-1).



Connection timeout: 120 seconds

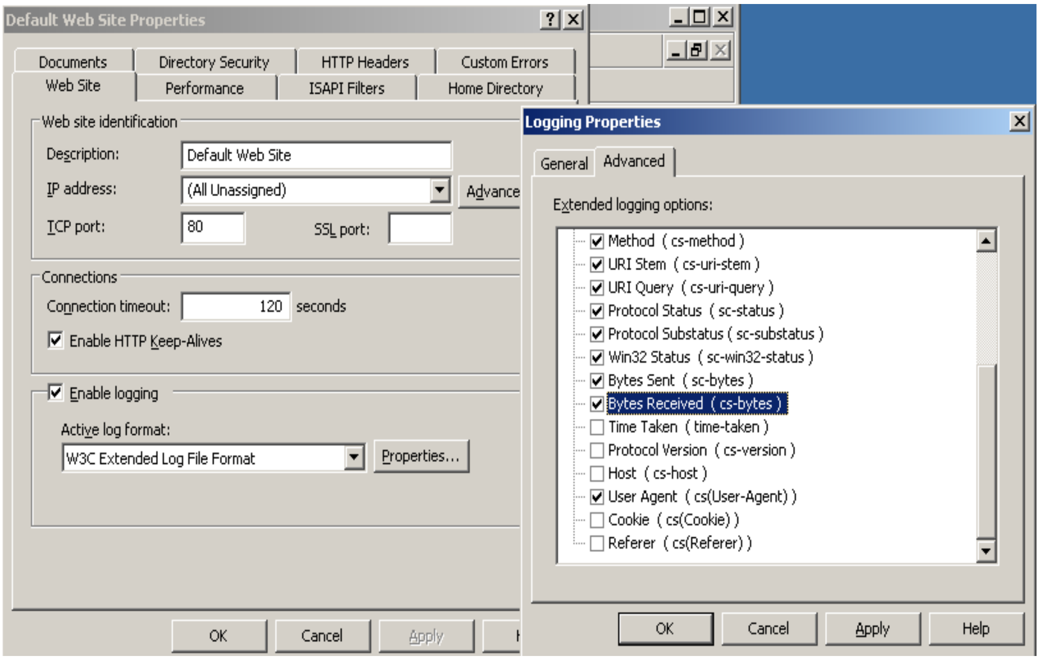
☑ Enable HTTP Keep-Alives

☑ Enable logging

Active log format: W3C Extended Log File Format

Figure 2-1: Microsoft IIS Web Server

GSA IT recommends for low, moderate, and high impact systems that IIS log the number of bytes received and sent which can be enabled via Advanced Logging Properties. (See Figure 2-2: IIS Log, to enable these fields).



Description: Default Web Site

IP address: (All Unassigned)

TCP port: 80 SSL port:

Connection timeout: 120 seconds

☑ Enable HTTP Keep-Alives

☑ Enable logging

Active log format: W3C Extended Log File Format

☑ Method (cs-method)
☑ URI Stem (cs-uri-stem)
☑ URI Query (cs-uri-query)
☑ Protocol Status (sc-status)
☑ Protocol Substatus (sc-substatus)
☑ Win32 Status (sc-win32-status)
☑ Bytes Sent (sc-bytes)
☑ Bytes Received (cs-bytes)
☐ Time Taken (time-taken)
☐ Protocol Version (cs-version)
☐ Host (cs-host)
☑ User Agent (cs(User-Agent))
☐ Cookie (cs(Cookie))
☐ Referer (cs(Referer))

Figure 2-2: IIS Log

Each daily log created by Microsoft IIS starts with a description of the fields. Figure 2-3 shows an example of the description of the log format and a corresponding log entry. As an exercise, we suggest parsing the event as shown in Figure 2-3 to find the sc-status field.

```
#Software: Microsoft Internet Information Services 6.0
#Version: 1.0
#Date: 2007-05-27 00:01:58
#Fields: date time s-sitename s-ip cs-method cs-uri-stem cs-uri-query s-port cs-username c-ip cs(User-
Agent) sc-status sc-substatus sc-win32-status
2007-05-27 00:01:58 W3SVC1 172.16.1.11 GET /OAB/1aa76d5b-11eb-422b-b914-
bb74e7fc83fc/oab.xml - 80 - 172.16.1.104 Microsoft+BITS/6.6 206 0 0
```

Figure 2-3: IIS Log Format

Note: A status code of 206 indicates success of a partial request. Status codes between 400 and 499 indicate failures. Please refer to HTTP Status Code Definitions for details.

2.2.4 Abnormal Logs

The purpose of log inspection is to identify abnormalities in system behavior. Whether the abnormality indicates a security problem or a performance problem, the log can provide information that otherwise may be overlooked. The objective is to see changes to system behavior that may have been caused by an attack, misuse, or other policy violation. The best approach to web server log review is to conduct log reviews on a regular basis, preferably once a week. This will allow gradual familiarity with the normal behavior of the system and the ability to ascribe causes to the common fluctuations in activity on the system. For example, there are daily, weekly, monthly and yearly cycles of activity that can be readily seen. Certain deviations from these should be tracked and understood. If the daily cycle has a large decrease in activity during the night, except for a spike that happens between 3:00 am and 5:00 am, there should be an explanation for it. The System Security Plan (SSP) covering the server should provide the log retention policy, in accordance with GSA record retention policies. This will allow trending analysis and referral to previous periods where similar activity has occurred.

Once familiarity of the normal behavior of the web server has been gained, entries that are unusual or do not make sense may be noticed. This is one reason logs are reviewed. However, not all abnormalities are bad. Review of logs involves following up on anomalies in the summaries or entries in the raw logs. Ideally, log entries are understandable and deviations from the norm can be explained. However, log entries may also contain many abnormal events that cannot be explained. As the reviewer gains experience with his/her system, more of the logs entries can be understood and less time is expended searching for explanations for anomalous events.

Abnormalities indicating malicious activity must be reported to the appropriate officials per GSA IT Security Procedural Guide: CIO-IT Security-01-02, "*Incident Response (IR).*"

2.2.4.1 Examples:

What types of abnormalities in the summary views should be looked for?

1. Unusual traffic patterns such as large transfers at odd times.
2. Unusual numbers of failed connections.
3. Repeated failures in SSL authentication from one source.
4. Unusual changes in request sources.
5. Unexpected changes in the Top 20 usage categories.

What does a suspicious detailed log entry look like? Below are some examples of some log entries that should arouse the suspicion of the log reviewer:

1. 2007-08-03 13:17:20 W3SVC2 172.16.1.11 GET /auktion.cgi menue=../../../../../../../../../etc/passwd 8081 - 172.16.1.100 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+5.0) 401 2 2148074254

2. 2007-08-03 13:18:46 W3SVC2 172.16.1.11 GET /index.php op=Default&Date=200607'%20UNION%20SELECT%201%2c20482%2c1%2c1%2c1%2c1%2c1%2c1%2c1%2c1%2f*&blogId=1 8081 - 172.16.1.100 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+5.0) 401 2 2148074254

3. 2007-08-03 13:16:51 W3SVC2 172.16.1.11 GET /testsite/typo3/dev/translations.php ONLY=%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/%2e%2e/etc/passwd%00 8081 - 172.16.1.100 Mozilla/4.0+

4. 2007-08-03 13:17:22 W3SVC2 172.16.1.11 GET /_vti_bin/fpcount.exe - 8081 - 172.16.1.100 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+5.0) 401 2 2148074254

The four examples above give a feel for the kinds of exploits that could show up in a web server log. The first uses an unusual pattern of dots and slashes. This request is attempting to access folders on the web server that are not intended to be served to web users. It also uses the term "passwd" which is very suspicious. The second contains SQL or Oracle keywords. These should not be present in normal web requests. The third passes an argument that includes hexadecimal values and the term "passwd". The fourth requests an executable file which is unusual and dangerous.

These examples represent a sampling of the types of abnormal entries that may be found in a web server log file. Abnormalities indicating malicious activity must be reported to the appropriate officials per CIO-IT Security-01-02.

2.3 Log Inspection Methodology

Web server log review is a tedious and time consuming process. A consistent set of inspection tasks based on best practices will provide the best results and make the job more manageable. Prepare for log review by obtaining and installing software to do the Summary Review. Select an appropriate set of summary views to use each time. Research the web server's log format. Be sure to have a reference card or resource handy that identifies each field. Please refer to NIST SP 800-95, "*Guide to Secure Web Services*."

Weekly review is the norm for most systems with moderate usage. Divide this review into summary and detail inspection tasks. Below is a list of tasks that normally constitute a weekly Web Server Log Review.

1) Collect logs:
    a. Copy logs from Web Server to analysis system.
    b. Archive raw logs as required by retention policy.
    c. Run summary reports for weekly analysis.
2) Summary Inspection:
    a. Examine the summary reports in the same order each week.
    b. Compare the previous week with the current week for each view.
    c. Identify any striking dissimilarities.
    d. If an anomaly is isolated to a particular time, note the time.
    e. If the anomaly is a particular address, entry point or is correlated with a Top 20 view, try to identify a search that will find related events during detailed review.
    f. Print and/or store the week's summaries as required.
3) Detailed Inspection:
    a. Open the logs in Notepad, WordPad or other text viewer. A very good alternative is to import the log file into a spreadsheet program such as MS Excel. Use "space" for delimiters. This allows for the use of autosort and more advanced

search features. If it is available in the environment, we recommend using graphical visualization tools as an alternative (e.g., Kibana), which can assist in identifying potential unusual behavior.

b. Bearing in mind the patterns you noted in the summary inspection, scan the contents to see if any oddities jump out at you.
c. Scroll to the time period of any anomalies identified during the summary review and see if anything can be found that might explain them.
d. Use search to locate anomalies in the Top 20 views.
e. Perform searches or visual scans for the suspicious features discussed in the next section.
f. Pay attention to the status code of entries. Repeated error codes could indicate repeated attempts to find vulnerabilities.
g. Keep records to explain the types of anomalies investigated and the explanation.

Security incidents are usually specific to one IP address. It is useful to look at all events from a particular IP address and explore the events associated with that address over the period of time being reviewed. This may explain if the event occurred over a long period of time, during or after normal working hours and/or identify its first occurrence. The starting time will help identify the source computer (internal IP addresses are often assigned temporarily) and help attribute it to a user or a network device sharing data over the network. Central to this process is finding the relationships between a requestor and some unusual logged events (e.g., if an abnormally large number of error codes are associated with one particular source IP address, this relationship should be investigated).

2.4 Suspicious Content

By following the methodology above, a reviewer may find a clue that something unusual is going on by noticing bulk changes in the summary views. By carefully considering the possible explanations for the change(s), he/she may choose to look at all log entries from a particular source or in a narrow time period. For example, a high proportion of total failed requests may lead to a particular user (or source IP address). This might lead to a detailed inspection of the user’s requests and the discovery that requests include numerous attempts to access .exe files or perpetrate other violations of policy. While the sequence of investigation in this example may be fairly obvious, there is no standard procedure. However, familiarity with summary views gives the best chance of identifying a problem. Expertise from the OCISO or other technical resources can be brought in to investigate further once an anomaly has been identified.

There are many exploits and forms of malicious logic that may be found in web server log entries. The following section discusses abnormal content that might be found during the inspection of logs.

All suspicious behavior/activity identified during a log review must be reported to the GSA Incident Response (IR) Team per CIO-IT Security-01-02. Note that if your review identifies unauthorized activity, this should be immediately report as well.

The most likely place to find malicious abnormalities in the HTTP protocol is in the request URL and other parts of the HTTP header. Often these are some form of simple obfuscation techniques used in the URL and its parameters. There are an infinite number of potential combinations of characters used to hide some aspect of the attack. The goal of these techniques is to not be detected by Intrusion Detection Systems (IDS) or Firewalls, but still get executed by the web application. Examples of this are the hexadecimal, decimal or even octal equivalents of characters. A common way to obfuscate an URL to evade IDS detection is called hexadecimal or hex encoding. For example this normal request: “GET /index.html HTTP/1.1” is equivalent to this request where “index.html” is substituted with hexadecimal values: “GET /%69%6E%64%65%78%2E%68%74%6D%6C HTTP/1.1”.

Each character could be substituted by its hexadecimal value and can take on several formats, e.g. the character “ (double quote) is equivalent to %22 or \x22 and the ; (semicolon) character is the same as %3b or\x3b. See Appendix B for a complete list of all characters.

Obfuscation can get very tricky. Not all hexadecimal content is obfuscation. It is common to see %20 (hexadecimal for a space) inserted into pathnames. Seek assistance from the OCISO for unusual log entries that appear to include obfuscation using characters such as “%”, ”\” or “/*”.

Routine attack noise in web server logs is common because people around the world are scanning everything constantly. Distinguishing between routine noise (from attempted attacks that are not worthy of notice) and unusual attacks can be difficult - knowledge and experience can help. For example, analysts can run a few commonly-available vulnerability scanners on their own system, like OWASP ZAP or Nessus, and review the logs thereafter for comparison.

2.4.1 Authentication

A URL query string should not be used for any sensitive data like session IDs, user/session information, user names, and passwords. Examples of what to look for:

- https:www.example.net/login?userid=eric&password= 4321
- https://www.example.net/asicommand?varB=321;ericid=4321

Parameter names are application specific but are easily identifiable in the web server log:

- /login\.jsp.*\?.*(userid/password)=./
- /;ericsessionid=./

2.4.2 Directory Breakout

Evidence that the requestor is attempting to get out of the root directory of the web server might look like “/../../../..”. These are frequently associated with particular files such as “/cmd.exe”. This means that the system sent a request to execute a shell command.

2.4.3 Active Code/JavaScript

Look for all the possible expressions which may trigger JavaScript or other active code. Here is a list of possible script inclusions:

- HTML tags: JavaScript, vbscript, expression, applet, meta, xml, blink, link, style, script, embed, object, iframe, frame, frame set, ilayer, layer, bgsound, title, base
- JavaScript event handlers(excerpt): onabort, onactive, onafterprint, onafterupdate, onsubmit, onunload

2.4.4 SQL Injection

SELECT, UNION, GRANT, CREATE, USER, ALTER and many other SQL terms can be inserted into URLs and used to attack the databases that support a web service. These are usually accompanied by quotes, asterisks and semicolons.

2.4.5 Miscellaneous

More advanced reviewers should be on the lookout for the following:

- URL-decoding: %xx (start with percent)
- Null byte strings termination
- Self-referencing paths: use of 1.1 and encoding equivalents
- Mixed-case characters: EriC
- Excessive use of whitespace
- Comment removal: convert DEL/*blah*/ETE FROM is DELETE FROM
- Conversion of backslash characters into forward slash characters
- Decoding HTML entities: &#99;, &quit;, &#xAA;
- Escaped characters: \+, \001, \xAA, \uAABB (start with backslash)

In the interest of brevity, the above-listed items have not been covered in greater detail in this guide; however, information on these types of abnormalities is available online and from OCISO security staff (see Appendix C for contact information).

3 Summary

The reviewer should prepare for the task by understanding the format of the web server log to be inspected. Web server log analysis software should be available to generate summary views of the log data.

First, determine "normal" behavior for the system. Then logs should be analyzed in two parts.

1. Summary inspection using different views to identify any abnormal trends or features that appear in the inspection period. Summary views will usually appear to follow consistent patterns from day to day, week to week. They will tend to show activities that occur at odd times, or from unusual sources. Any odd characteristics should be flagged and followed up in Step Two.
2. Raw logs should be searched to find the detail of unusual characteristics that are flagged in Step One. This detail should provide some explanation for the abnormal activity. As it is understood, these characteristics should be used to identify other similar abnormalities until it is clear that they represent a normal form of activity.

Using this methodology, common deviations from normal activity can be efficiently investigated and explained. In some cases, expert assistance may be required to find explanations for abnormal but non-malicious activity. As the analysts get more familiar with normal logs, automating repetitive parts of review and setting up alerts for suspicious behavior will increase the efficiency of analysis. In a mature analysis system, automated alerting can result in reviews primarily of anomalies, rather than routine logs.

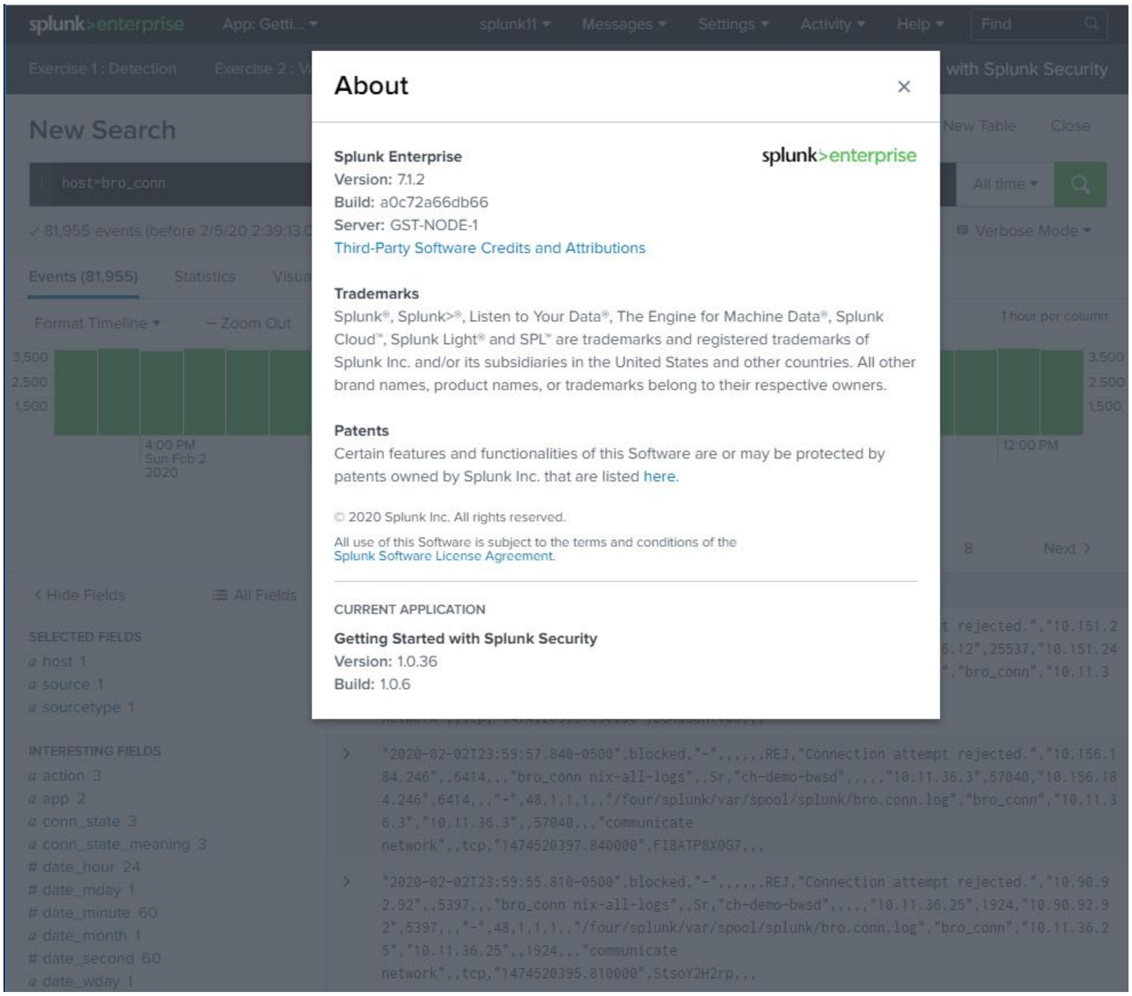
In addition, the OCISO Security Operations Division provides monthly reports of high and critical vulnerabilities on systems, including web servers. The reports provide information which can be analyzed in relation to log data, potentially identifying areas where security or performance is not satisfactory and action is required.

Appendix A: Log Analysis Tools

There are many tools to help with the analysis and review of logs in addition to the Enterprise Logging Platform. Raw log review is difficult because the information presented is not graphical and patterns are very difficult to discern. Useful tools allow the reviewer to see summarized data and statistics that illustrate patterns. The reviewer is often looking for an unusual event or sequence of events. These appear as changes in regular patterns of activity. A good log analysis tool should allow a user to quickly dig into the detail from the summary to understand why a particular change has occurred. Only a small sample of basic log analysis products is listed here. One is a high-end, enterprise scale tool; two of them are low-end commercial tools (<$100); and the other is GPL freeware. Analysis software must be configured to point to the log file to be analyzed. In the case of Web Log Explorer, this is all that needs to be done. A small amount of reading and configuration allows these tools to be extremely useful in familiarizing the reviewer with normal operation and identifying deviations from “normal.”

The difficulty with analysis tools, at times, can be their configuration. The easiest configuration involves copying the log file from the server to an analysis system (generally the reviewer’s workstation). The log file is opened from the log analysis tool and reports are generated. This is easy to set up but may become tedious week after week. More sophisticated log analyzers retrieve the logs from one or more systems using network protocols such as HTTP or HTTPS. However, server security best practices should be followed. If log transfer is to be done, the solution should be engineered to use robust authentication and secure transport protocols. These systems should be set up to support the requirements outlined within GSA Order CIO 2100.1, “*GSA Information Technology (IT) Security Policy*” and CIO-IT Security-01-08 in coordination with the GSA IT OCISO.

Splunk is an enterprise scale log review tool that enables analysts to ingest, compile, correlate, and search multiple logs for anomalous activity during the course of periodic log review as well as during incident response efforts.



About

Splunk Enterprise
Version: 7.1.2
Build: a0c72a66db66
Server: GST-NODE-1
Third-Party Software Credits and Attributions

Trademarks

Splunk®, Splunk>®, Listen to Your Data®, The Engine for Machine Data®, Splunk Cloud™, Splunk Light® and SPL™ are trademarks and registered trademarks of Splunk Inc. and/or its subsidiaries in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners.

Patents

Certain features and functionalities of this Software are or may be protected by patents owned by Splunk Inc. that are listed here.

© 2020 Splunk Inc. All rights reserved.

All use of this Software is subject to the terms and conditions of the
Splunk Software License Agreement.

CURRENT APPLICATION

Getting Started with Splunk Security
Version: 1.0.36
Build: 1.0.6

Figure A-1: Splunk

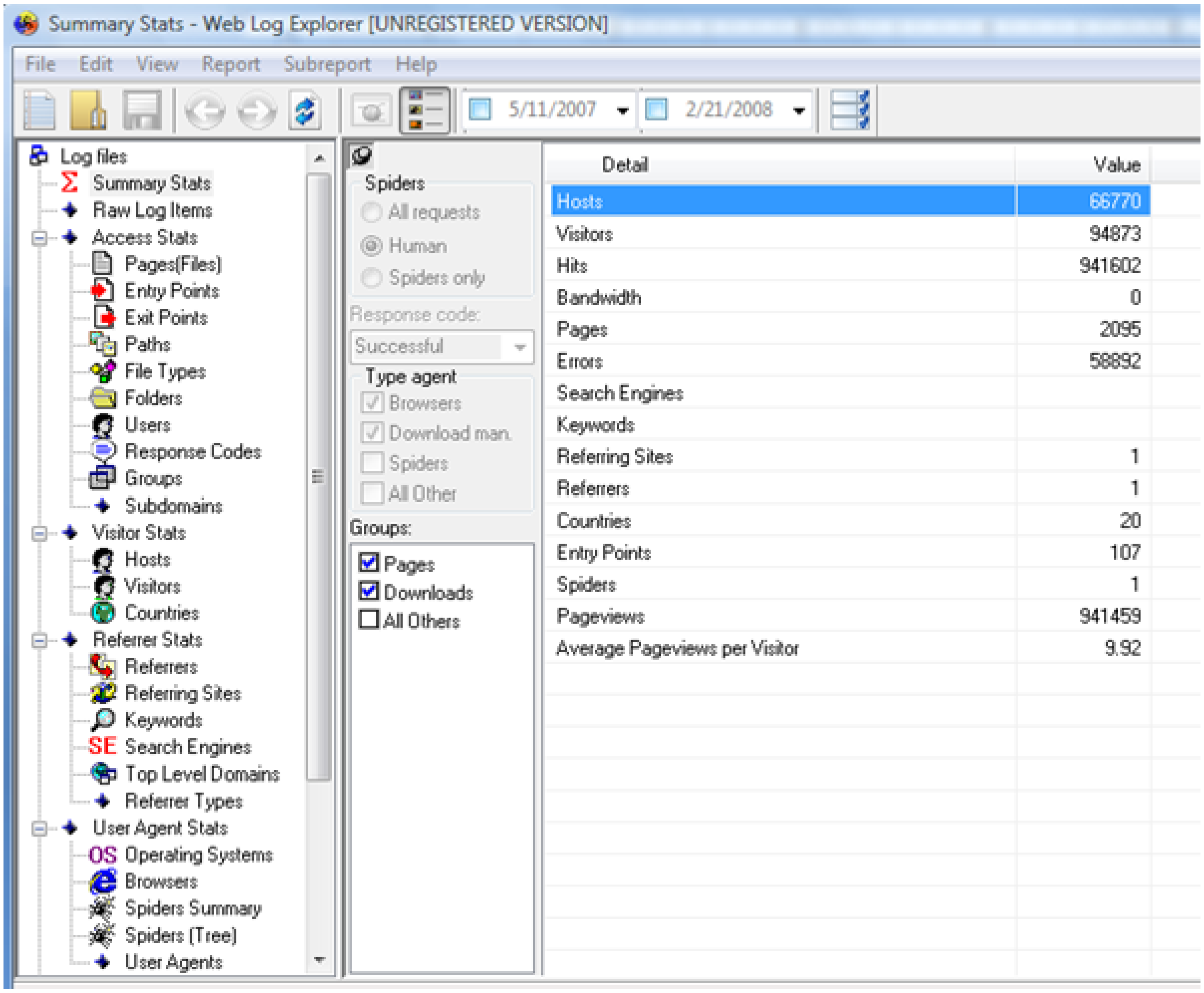
Web Log Explorer (Standard Edition)

Web Log Explorer is very easy to install and use. It has a straight forward and self-explanatory GUI. It is very useful to understand normal usage of systems.

AWStats is a powerful open source, freeware tool set. It is somewhat more complex to install than Web Log Explorer.

Also, see WebLog Expert - an access log analyzer.

There are also many enterprise log consolidation, correlation and analysis platforms that are designed to alert operators quickly in the event of malicious activity. These tools, though useful to Network and Security Operation Centers, are outside the scope of this guide.



5/11/2007 – 2/21/2008

Spiders

○ All requests

● Human

○ Spiders only

Response code: Successful

Type agent

☑ Browsers

☑ Download man.

□ Spiders

□ All Other

Groups:

☑ Pages

☑ Downloads

□ All Others

Detail	Value
Hosts	66770
Visitors	94873
Hits	941602
Bandwidth	0
Pages	2095
Errors	58892
Search Engines	
Keywords	
Referring Sites	1
Referrers	1
Countries	20
Entry Points	107
Spiders	1
Pageviews	941459
Average Pageviews per Visitor	9.92

Figure A-2: Web Log Explorer

Appendix B: ASCII Table

Table B-1: ACSII Table - Non-Printable Characters

DEC	HEX	CHARACTER (CODE)	DEC	HEX	CHARACTER (CODE)
0	0	NULL	16	10	DATA LINK ESCAPE (DLE)
1	1	START OF HEADING (SOH)	17	11	DEVICE CONTROL 1 (DC1)
2	2	START OF TEXT (STX)	18	12	DEVICE CONTROL 2 (DC2)
3	3	END OF TEXT (ETX)	19	13	DEVICE CONTROL 3 (DC3)
4	4	END OF TRANSMISSION (EOT)	20	14	DEVICE CONTROL 4 (DC4)
5	5	END OF QUERY (ENQ)	21	15	NEGATIVE ACKNOWLEDGEMENT (NAK)
6	6	ACKNOWLEDGE (ACK)	22	16	SYNCHRONIZE (SYN)
7	7	BEEP (BEL)	23	17	END OF TRANSMISSION BLOCK (ETB)
8	8	BACKSPACE (BS)	24	18	CANCEL (CAN)
9	9	HORIZONTAL TAB (HT)	25	19	END OF MEDIUM (EM)
10	A	LINE FEED (LF)	26	1A	SUBSTITUTE (SUB)
11	B	VERTICAL TAB (VT)	27	1B	ESCAPE (ESC)
12	C	FF (FORM FEED)	28	1C	FILE SEPARATOR (FS) RIGHT ARROW
13	D	CR (CARRIAGE RETURN)	29	1D	GROUP SEPARATOR (GS) LEFT ARROW
14	E	SO (SHIFT OUT)	30	1E	RECORD SEPARATOR (RS) UP ARROW
15	F	SI (SHIFT IN)	31	1F	UNIT SEPARATOR (US) DOWN ARROW

Table B-2: ASCII Table - Printable Characters

DEC	HEX	CHARACTER	DEC	HEX	CHARACTER	DEC	HEX	CHARACTER
32	0x20	<SPACE>	64	0x40	@	96	0x60	`
33	0x21	!	65	0x41	A	97	0x61	a
34	0x22	"	66	0x42	B	98	0x62	b
35	0x23	#	67	0x43	C	99	0x63	c
36	0x24	$	68	0x44	D	100	0x64	d
37	0x25	%	69	0x45	E	101	0x65	e
38	0x26	&	70	0x46	F	102	0x66	f
39	0x27	'	71	0x47	G	103	0x67	g
40	0x28	(	72	0x48	H	104	0x68	h
41	0x29	)	73	0x49	I	105	0x69	i
42	0x2A	*	74	0x4A	J	106	0x6A	j
43	0x2B	+	75	0x4B	K	107	0x6B	k
44	0x2C	,	76	0x4C	L	108	0x6C	l
45	0x2D	-	77	0x4D	M	109	0x6D	m
46	0x2E	.	78	0x4E	N	110	0x6E	n
47	0x2F	/	79	0x4F	O	111	0x6F	o
48	0x30	0	80	0x50	P	112	0x70	p
49	0x31	1	81	0x51	Q	113	0x71	q
50	0x32	2	82	0x52	R	114	0x72	r
51	0x33	3	83	0x53	S	115	0x73	s
52	0x34	4	84	0x54	T	116	0x74	t
53	0x35	5	85	0x55	U	117	0x75	u
54	0x36	6	86	0x56	V	118	0x76	v
55	0x37	7	87	0x57	W	119	0x77	w
56	0x38	8	88	0x58	X	120	0x78	x
57	0x39	9	89	0x59	Y	121	0x79	y
58	0x3A	:	90	0x5A	Z	122	0x7A	z
59	0x3B	;	91	0x5B	[	123	0x7B	{
60	0x3C	<	92	0x5C	\	124	0x7C	|
61	0x3D	=	93	0x5D	]	125	0x7D	}
62	0x3E	>	94	0x5E	^	126	0x7E	~
63	0x3F	?	95	0x5F	_	127	0x7F	<DEL>

Appendix C: Points of Contact

Enterprise Logging Platform questions should be directed to the Security Operations Division – SecOps@gsa.gov.

All other inquiries should be directed to the Security Engineering Division – SecEng@gsa.gov and gsa-ir@gsa.gov.