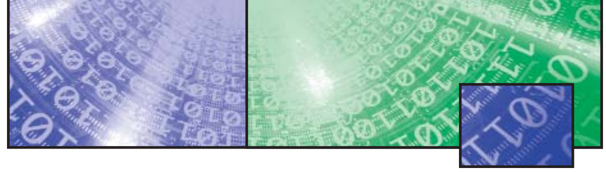

Implementing a Log Consolidation and Analysis Project

By Roy Madden
roy@informationsecurityireland.com

Introduction

Log analysis has long been regarded as one of the most effective methods for identifying and dealing with information security risk, and it is likely that most security practitioners will have used logs at one time or another in following up on a security incident. Without accurate logs, most security incidents cannot be adequately investigated.

The *reactive* use of logs to investigate incidents is commonplace. However, the use of logs in a *proactive* manner to detect possible intrusions before loss to the business occurs is typically not something that the average company does effectively. In this author's experience, 90 to 95% of security incidents can be traced through a review of logs, and the majority could also have been addressed before loss occurred if the logs in question had been reviewed at an early enough stage. Explanations for this lack of review are numerous: There are too many logs, they are widely spread out across the organization and stored on a variety of different devices, the logs themselves do not give useful information, logs can be cryptic, logs can be overly verbose. It would not be unusual for the average company to generate several gigabytes of log entries across all their systems during the course of a typical day.

It is against this background that the phrase "log consolidation and analysis" has become somewhat of an industry buzzword. While vendors will have different definitions, log consolidation and analysis can be taken to simply mean the collection and analysis of logs from disparate devices and applications, in a manner that allows a security (and/or operations) team to both understand log events and react to inappropriate patterns.

A typical log consolidation project is by no means a simple task, and a structured, methodical approach should be adopted to achieve maximum value from your investment. Here are 10 tips:

1. Generate a High-Level Inventory of Your Systems and Applications

This step is typically done in conjunction with step 2. The objective here is to familiarize yourself with the systems and applications in use. You should avoid going to a great level of detail on this step, as some of the systems may be discounted for follow-up actions after step 3. This step will include talking to data and application owners, reading technical inventories, and reviewing architecture documents if available.

2. Identify at a High Level the Logging Mechanisms Available

Again, you should not go to a great level of detail here, but should concentrate on identifying the main logging mechanisms available and gaining an understanding of their logging capabilities.

3. Perform a Risk Analysis and Map Results to Logging Mechanisms

Of all the steps in the process, this step is arguably the most critical. Following the completion of steps 1 and 2, you will have gained an understanding of the capabilities of your systems to record log data. No business can reasonably expect to have the resources to interrogate all logs identified, and you will be looking to target log review resources in the areas that will provide maximum benefit. The ideal way to tackle this step is to perform a risk analysis which has, as its focus, the identification of useful log data. Every company has their own method of performing risk analysis. The following is a simplified description of the process that I have followed with success:

a) Divide the Risk Analysis (RA) threats into groups according to the areas from which the threats originate (for example, External Untrusted (B2C), External Trusted (B2B)).

b) Using each of the headings from (a), brainstorm the risks that the business faces. A number of group workshops among technical and business stakeholders is the typical method used to accomplish this task.

c) Assign an Impact and a Likelihood score to each of the risks identified in (b). These scores should take into consideration existing control mechanisms. Remember, you are trying to perform an RA that focusses on the areas where logs can assist in reducing risk by supplementing existing controls.

d) Calculate a risk score based on (c) and agree on a cutoff level for risks to be addressed. Risks above the cutoff level are risks that the business feels are unacceptable and should be addressed.

   You have now established a basic model for the risks your business faces and the areas in which unacceptable risk lies. The next step is to assess where logs could potentially be of value.

e) Using your knowledge gained in the inventory steps 1 and 2, for each risk identified in the RA, add a score for "potential impact of log." This figure is the potential reduction in the risk score if appropriate logging was enabled. This figure should include both logging known to exist and logging levels that were identified as being potentially available. For example, if your business does not have Windows event logs enabled on servers, you should still assess the "potential impact of log" score as if the logging was enabled. The reasons will be explained in step 4.

f) Calculate the residual risk (score from (d) minus the potential impact of logs (e)). If the output score is below the cutoff level, then that risk is a candidate for addressing by means of the log consolidation project. You may also wish to incorporate certain items where logging controls greatly reduce the risk score, although not to a level below the cutoff score.

4. Perform a Gap Analysis Between Current and Required Levels of Logging

By now you should know what risks can be addressed through log review. You should also know what logs you require. The next step is to conduct a gap analysis between your current level of logging and the level of logging that is required to address the risk. Typically this involves discussions for the most part with operations personnel, who may (more likely *will*) have concerns about logging overhead and other resources required. You will have to look in detail at the logging systems available on the systems you have previously inventoried.

In many cases you will find that either logs aren't sufficiently detailed, don't capture the required information, or perhaps simply cannot be enabled. Once you have completed discussions on closing gaps identified, the Risk Analysis scores should be updated to reflect the new information. Your newly updated Risk Analysis will now show the definite candidate logs for inclusion in log consolidation.

5. Document the Forensic Quality of the Logging Mechanisms

This isn't a subject I intend to cover in detail, as it could be the contents of an article all to itself! The question is whether the logs provide sufficient forensic quality to enable you to rely on their accuracy. You will have to decide as part of this process exactly how much confidence you require from your log consolidation system, and it may be a case that you require additional safeguards to be built around log collection and storage.

Many companies will address forensic quality as an iterative process. They will try and get to a baseline level of log collection and review, which may not be appropriate for use in a criminal prosecution, and over time will improve the forensic procedures and associated quality.

6. Identify Collection Methods for Retrieving and Analyzing the Data

There are some excellent articles available on the Web discussing centralized logging servers, syslog usage and the advantages/disadvantages of commercial products. You will essentially have three choices:

a) Buy a commercial log consolidation package. The advantages of these packages is that you can buy "out of the box" collection mechanisms that are usually strong on forensic quality. The software will normalize the data collected, and pre-written reports on unusual log entries will come as standard. The disadvantages are typically cost, unacceptable technical requirements, limited systems supported, lack of support for non-standard applications, limited pre-defined reports, and sometimes limited reporting capabilities.

b) Buy multiple log analysis products for each individual system. The advantages here are that you can pick and choose what may be best-of-breed products; the disadvantages are the complexity and support requirements of multiple products.

c) Do it yourself, typically involving a scripting language such as Perl or Python, producing filtered reports from logs which have been pulled to a central syslog server. In many ways, this is the most powerful option, as it allows you to build a system that is tailored exactly to your needs, but it has a whole host of downsides, including complexity and support concerns.

7. Gain an Understanding of Expected System and User Behavior

This step requires gaining an understanding of log entries that would be considered "normal," a step that will help you considerably in understanding what constitutes the reverse (abnormal or unexpected entries in your logs). This will not be a short process and will involve gaining a detailed understanding of the entries that your logs are producing. While this work can be tedious, you will likely gain a surprising insight into the operation of your systems as part of this process. You will typically have a great deal of interaction with end users and business owners in order to gain an understanding of what is normal and abnormal behavior as reflected in the logs.

8. Filter and Report on Log Entries that Require Follow-up Investigation

This is another vital step. Filtering and reporting on log entries is extensively documented in various online sources.

With unlimited time and resources there is the possibility of reviewing every log entry and flagging of unusual entries for follow-up. Unfortunately I don't know any of us who has this luxury, and the alternative is to filter the abnormal entries that are produced. The first step in filtering should be to remove the entries identified as "normal" in step 7. From there, you will have to try and identify which of the remaining entries should be reported and followed up on. If you have bought a log analysis package, it will likely contain a reporting tool with many of these filters pre-defined. If you are doing the filtering yourself, it will be a time-consuming process to tweak the filters you have defined. At the beginning of the process, you may feel a bit overwhelmed by the task, but diligent filtering and an acceptance that you simply cannot filter and report on all events will allow the task to become more manageable.

9. Archive and Protect Logs

The forensic qualities of logs are more fully explained in step 5. You will need an archiving strategy in place to deal with old logs that have been filtered, and you will have to decide on forensic and storage requirements.

10. Investigate Exceptions and Events

The final step is to follow up on the exceptions and events that have been flagged. This is another balancing act between time consumed following up on log entries versus the benefit to be gained from understanding the entries. The skill in allocating time to investigate log events is one that will improve over time. As your understanding of the logs you produce increases and you look at the previous results of following up on exceptions, you will get a better understanding of the cost/benefit ratio that is involved.

Conclusion

There are a great number of companies out there who are struggling to get real security value from their logging systems. For most, the first step in realizing the potential value is to implement a consolidated log analysis system.

None of us likes to see security projects that don't deliver, and one of the ways you can avoid an underachieving (or failed) log consolidation project is to tackle the project in a structured, methodical manner. There can be a great temptation to simply leap into a technical-driven solution without considering the underlying security requirements. The approach in this article, while by no means perfect, does offer a means by which log consolidation and analysis can be seen to deliver value and integrate with your existing risk assessment processes.

Roy Madden is the information security officer for one of the leading financial services firms in Ireland. He is also co-founder of the Irish chapter of ISSA.