Best practices for log collection and management

In this reading, you'll examine some best practices related to log management, storage, and protection. Understanding the best practices related to log collection and management will help improve log searches and better support your efforts in identifying and resolving security incidents.

Logs

Data sources such as devices generate data in the form of events. A **log** is a record of events that occur within an organization's systems. Logs contain log entries and each entry details information corresponding to a single event that happened on a device or system. Originally, logs served the sole purpose of troubleshooting common technology issues. For example, error logs provide information about why an unexpected error occurred and help to identify the root cause of the error so that it can be fixed. Today, virtually all computing devices produce some form of logs that provide valuable insights beyond troubleshooting.

Security teams access logs from logging receivers like SIEM tools which consolidate logs to provide a central repository for log data. Security professionals use logs to perform **log analysis**, which is the process of examining logs to identify events of interest. Logs help uncover the details surrounding the 5 W's of incident investigation: *who* triggered the incident, *what* happened, *when* the incident took place, *where* the incident took place, and *why* the incident occurred.

Types of logs

Depending on the data source, different log types can be produced. Here's a list of some common log types that organizations should record:

- **Network**: Network logs are generated by network devices like firewalls, routers, or switches.
- **System**: System logs are generated by operating systems like Chrome OS™, Windows, Linux, or macOS®.
- **Application**: Application logs are generated by software applications and contain information relating to the events occurring within the application such as a smartphone app.
- **Security**: Security logs are generated by various devices or systems such as antivirus software and intrusion detection systems. Security logs contain security-related information such as file deletion.
- **Authentication**: Authentication logs are generated whenever authentication occurs such as a successful login attempt into a computer.

Log details

Generally, logs contain a date, time, location, action, and author of the action. Here is an example of an authentication log:

```
Login Event [05:45:15] User1 Authenticated successfully
```

Logs contain information and can be adjusted to contain even more information. Verbose logging records additional, detailed information beyond the default log recording. Here is an example of the same log above but logged as verbose.

```
Login Event [2022/11/16 05:45:15.892673] auth_performer.cc:470
User1 Authenticated successfully from device1 (192.168.1.2)
```

Log management

Because all devices produce logs, it can quickly become overwhelming for organizations to keep track of all the logs that are generated. To get the most value from your logs, you need to choose exactly what to log, how to access it easily, and keep it secure using log management. **Log management** is the process of collecting, storing, analyzing, and disposing of log data.

What to log

The most important aspect of log management is choosing what to log. Organizations are different, and their logging requirements can differ too. It's important to consider which log sources are most likely to contain the most useful information depending on your event of interest. This might be configuring log sources to reduce the amount of data they record, such as excluding excessive verbosity. Some information, including but not limited to phone numbers, email addresses, and names, form personally identifiable information (PII), which requires special handling and in some jurisdictions might not be possible to be logged.

The issue with overlogging

From a security perspective, it can be tempting to log everything. This is the most common mistake organizations make. Just because it can be logged, doesn't mean it *needs* to be logged. Storing excessive amounts of logs can have many disadvantages with some SIEM tools. For example, overlogging can increase storage and maintenance costs. Additionally, overlogging can increase the load on systems, which can cause performance issues and affect usability, making it difficult to search for and identify important events.

Log retention

Organizations might operate in industries with regulatory requirements. For example, some regulations require organizations to retain logs for set periods of time and organizations can implement log retention practices in their log management policy.

Organizations that operate in the following industries might need to modify their log management policy to meet regulatory requirements:

- Public sector industries, like the Federal Information Security Modernization Act (FISMA)
- Healthcare industries, like the Health Insurance Portability and Accountability Act of 1996 (HIPAA)
- Financial services industries, such as the Payment Card Industry Data Security Standard (PCI DSS), the Gramm-Leach-Bliley Act (GLBA), and the Sarbanes-Oxley Act of 2002 (SOX)

Log protection

Along with management and retention, the protection of logs is vital in maintaining log integrity. It's not unusual for malicious actors to modify logs in attempts to mislead security teams and to even hide their activity.

Storing logs in a centralized log server is a way to maintain log integrity. When logs are generated, they get sent to a dedicated server instead of getting stored on a local machine. This makes it more difficult for attackers to access logs because there is a barrier between the attacker and the log location.

Key takeaways

It's important to understand how to properly collect, store, and protect logs because they are integral to incident investigations. Having a detailed plan for log management helps improve the usefulness of logs and resource efficiency.