Thoughts on Common Log Formats and Secure Logging

R. Turner June 4, 2010

Introduction

This document discusses two potential “log-related” activities that may be of interest to the PWG IDS Working Group to consider for future work.

The original scope of this document was to consider only a Common Log Format for hardcopy devices. However, during the research phase of this document, it became clear that the scope of log-related activities should probably include both a Common Log Format, as well as a profile or recommendation for securing log files and log file transmission.

Secure Logging

There are (at least) two aspects to secure logging:

1. The integrity of log messages when transported from a device to a log "sink"

and, optionally

2. Ensuring confidentiality of log messages

It seems reasonable to consider log message confidentiality to be optional because, in many cases, the information in a log message would not be considered confidential, but "integrity" of a log message seems to be mandatory. Log records should not be tampered with, or otherwise modified by unauthorized parties, during transport from the device to a log “sink”

A log "sink" is some service somewhere that is responsible for the long term archival of log information, either for corporate information policy requirements, or regulatory requirements, or both.

It should’t be necessary to “reinvent” existing standards. It should be possible to define a "profile" of existing standards to meet both integrity and confidentiality requirements.

The IETF "syslog" charter is addressing (in some manner) the requirements proposed above. The IETF SYSLOG working group charter is included below:

Syslog has been a de-facto standard for logging system events for long time. The syslog WG recently completed standardization of the syslog protocol (RFC 5424), secure transport of the syslog protocol over TLS (RFC 5425), and non-secure transport over UDP (RFC 5426).
The WG under this charter will standardize a DTLS transport for syslog, providing a secure transport for syslog messages in cases where a connection-less transport is desired. The threats that this WG will primarily address are modification, disclosure, and masquerade. A secondary threat is message stream modification. These are consistent with those addressed in RFC 5425. Draft-feng-syslog-transport-dtls is already similar to RFC 5425 in this respect, so this draft will become the starting point for the WG document, which the WG will adjust as needed, and merge desired features from other sources, such as draft-petch-gerhards-syslog-transport-dtls, draft-hardaker-isms-dtls-tm, and draft-seggelmann-tls-dtls-heartbeat.
The WG will also complete the ongoing work to specify a standardized mechanism for signing syslog messages (draft-ietf-syslog-sign).

One point about using the work of the syslog working group would be to make sure that the security mechanisms that the group is proposing are not "bound" to the syslog protocol - For example, DTLS is a common method for securing UDP connectionless network traffic, and is applicable to many protocols, not just Syslog. I haven't read any of the SYSLOG working group drafts pertaining to signing syslog messages, but if this draft requires "normalization" or "canonicalization" of syslog protocol elements, then we may want to look at another way of signing log messages, because, as I said, I don't think we necessarily want to use SYSLOG protocol - or at least at this point (before we even start on this work), it seems rather constraining to have to operate using SYSLOG format. However, once we start considering the problem, it may turn out to be ok, I just think it's premature at this point to RELY on SYSLOG.

One other problem we may have that DTLS doesn't address with regards to confidentiality, is the fact that the log messages protected by DTLS are only protected (encrypted) during transport from device to log sink, at which point they are decrypted. I am still trying to figure out if there are any requirements that log messages must stay encrypted even when permanently archived.

Common Log Format

After considering a common log format for awhile, I decided that I would only propose the types of information that we might want to include in a log message associated with hardcopy devices.

Similar to SYSLOG, there are actually a number of categories of log messages to consider.

I'm going to constrain the discussion of hardcopy log messages to those types of event messages that are "application specific" to the hardcopy domain of applications (printing, faxing, scanning, other). Many hardcopy devices already generate log information pertaining to events regarding security, or other operation events that enable subsequent accounting (job accounting, bill-back, etc.)

Many printing devices today may be based on operating systems (Linux/Unix) that already maintain system logs. I don't want to mix-up these existing system log messages (which may be SYSLOG-based) with the types of application-specific log messages that I am referencing. I'm assuming that these existing application-inspecific log messages (OS, device, etc.) will continue to operate as they are.

Rationale for Development

In general, logs can be used for any number of purposes, including:

1. optimizing system and network performance
2. to record the actions of users
3. identifying security incidents, policy violations, fraudulent activities, and operational problems
4. performing audits and forensic analyses
5. supporting internal investigations
6. establishing baselines, and
7. identifying operational trends and long-term problems.

The primary focus of this proposed work item for the IDS is numbers 2, 3, 4, and 5 in the above list. Secondarily, the other identified uses for logs are also important, and could also be addressed within the same scope of work.

The idea behind generation of a common log format for hardcopy devices is to enable easier compliance with evolving enterprise regulatory requirements, such as:

- **Federal Information Security Management Act of 2002** (ISMA) - requires federal agencies to develop, document, and implement an organization-wide program to provide information security for the information systems that support its operations and assets.
- **Health Insurance Portability and Accountability Act of 1996** (HIPAA) - mandates safeguarding the confidentiality, integrity, and availability of electronically protected health information.
- **Sarbanes-Oxley Act of 2002 (SOX)** - applies to financial and accounting practices and the IT functions that support these practices.
- **Gramm-Leach-Bliley Act (GLBA)** - requires financial institutions to protect their customers' information against security threats.

These regulatory requirements do not stipulate that “common” log formats should be used. They only stipulate the requirement for long-term archival of device logs. However, Network and/or System administrators could really benefit from a common log format for hardcopy devices. At the moment, most large-scale system/network management applications just dump raw log information to administrators, who have to either develop their own applications for processing this information, or worse, they peruse the information manually.

NIST Involvement

In 2006, NIST’s Information Technology Laboratory issued Special Publication (SP) 800-92, Guide to Computer Security Log Management, by Karen Kent and Murugiah Souppaya, to help organizations develop, implement, and maintain effective processes for managing logs with security-related information. It contains basic information about computer security logs, the usefulness of these logs, and the challenges of managing them. Briefly mentioned are the components of the log management infrastructure; the planning processes that enable the organization to carry out consistent, reliable, and efficient log management practices; and the operational processes that aid organizations in successfully managing logs. See http://csrc.nist.gov/publications/nistpubs/index.html.

Contents of a Log Message

There are “generic” fields that are present in any log message from a device, and these will be the focus of my proposal. There are also operation-specific (fax, scan, print) information as well.

Date and Time (ISO Format or other widely-used syntax)

Device Identity - This is always an unambiguous identification of the device for which the log message is generated. One example might be the SubjectName field of an X.509 device certificate.

User Identity - This could be either the identity of a user of the device, or the device identity itself. This is the identity of an entity that is directly or indirectly responsible for generating the log message.

Severity - This is an indicator of the severity of the event, from the perspective of the administrator. It should indicate how quickly an administrator should respond to the particular event being generated. Example enumerations for severity could include “I”, “W”, or “E”, for Informational, Warning, or Error.

Subsystem Identifier - The subsystem within a hardcopy device that is emitting the particular event or log message. I believe the Printer MIB, or maybe the IPP semantic model, or other PWG work may have already defined the types of subsystems that can exist within a hardcopy device.

Event Type - Event types could map to any type of hardcopy event, such as those defined as SNMP "notifications" in hardcopy-related MIBs.

Each subsystem will define the types of events that it can generate. But for regulatory purposes, it's important to track the types of information for which subsequent forensic analysis could best utilize.

Security Events - Some entity attempted access to a resource for which the entity has not been authorized. The "resource" in question could be the device itself, in the case of a "login" or "authentication failure", or the attempt by an authenticated entity to use functionality or resources of the device that the entity is not authorized to use. This type of event could also reflect the authorized use of a resource by an entity, however the particular usage by the entity has violated a "quota" associated with the resource.

Accounting Records - These types of log messages would probably align with existing printer accounting records that are maintained by many hardcopy vendors today. Although most, if not all, vendors only generate these records for print jobs. In addition to "print job" accounting records, I am proposing the following additional accounting records for other features of an MFP:

- Fax Accounting (for each FAX either sent or received)
    - T.30 Remote System Identifier string
    - Completion Status of FAX session
    - Color/Monochrome FAX?
    - Elapsed Time
    - Inbound
        - Caller ID Information: This can either be phone number for analog fax, or for VoIP, there may be Network/ENUM/Identity information contained in this field as well
        - Where was fax routed? output print bin? Emailed to user? Other?
        - If FAX is printed, generate printer accounting record
    - Outbound
        - Dialed Number (or ENUM or other SIP-related info for VoIP call)
        - Pages successfully transmitted
        - Pages attempted
        - Source of Fax Content (inbound email? From who? Scan job? Other?)
- Scan Job Accounting
    - Completion Status of Scan Session
    - Pages Scanned
    - Destination (FAX, Remote User Directory, other)
    - Color/Monochrome Scan

As a check on print job accounting records, each print job accounting record should include as much information regarding “consumables used” as possible, whether the job was color or black/white, and how many physical pages (and impressions) were used to complete the print job.

What is described in this proposal is purely a “data model” for hardcopy device log messages. The data model could be expressed as a text file, or as an XML schema for more efficient processing by some type of automaton.

In a longer term scope of work, we should look at some type of archival capability. For instance, paper faxes are easily lost or destroyed. If an organization loses its fax communications, it can also lose irreplaceable operational knowledge. This can be costly, time-consuming and result in unnecessary delays. By integrating long-term document management archival capabilities, an organization can electronically store important fax documents to retain important corporate data and documents.

As either an archival or forensic capability, some devices may even store (probably on a centralized server) a compressed version of the first <n> pages of any printed job, or any scanned document associated with a device. A long-lived URN can be generated for this “image”, and stored in an archive, with the URN reference also included in the log message associated with the job.