Security Log Standard - Still an Open Question

Rocco Gagliardi
Defense Department, scip AG
roga@scip.ch
https://www.scip.ch

Marc Ruef (Editor)
Research Department, scip AG
maru@scip.ch
https://www.scip.ch

Abstract: A log standard is still missing. IoT urges a common format to exchange security events. Log will probably evolve towards a more structured and cryptic machine-oriented formats.

Keywords: AD, API, Cisco, Cloud, Detect, Exchange, Firewall, Framework, HTML, IBM

1. Preface

This paper was written in 2018 as part of a research project at scip AG, Switzerland. It was initially published online at *https://www.scip.ch/en/?labs.20180315* and is available in English and German. Providing our clients with innovative research for the information technology of the future is an essential part of our company culture.

2. Introduction

Even if you write a small script, you are facing the question: "What and how should I log what is going on?" The *problem of event representation, communication, and interpretation* consists in identify and transmit useful information, so that the counterpart understands context and content; in the specific case of security information, it is important to determine if a security event can be correlated by different sensors like network, host, or other. This problem is critical if we consider the number and the type of components that need to exchange such information.

3. Brief history of log formats

If multiple systems observe the same occurrence, it should be expected that their description of that event is identical. When combined with relevant event details (time, source, destination), a computer should be able to immediately determine whether two or more logs, data logs, audit logs, alerts, alarms, or audit trails refer to the same event.

In order to make this happen, we need:

- a common language to describe the event
- a common mechanism to log information
- a common mechanism to transmit information

As *NIST 800-92, Guide to Computer Security Log Management* [1] states *"there is no consensus in the security community as to the standard terms to be used to describe the composition of log entries and files."*

Many attempt to address this problem have been started, but we still miss a recognized standard. The following list is not exhaustive; I just picked up some formats I loved or hated during my career, but there are around thousand (~1000) different (syslog) message formats.

Format	Type	Proposed by	Year	Status
XDAS [2]	Open	OpenGroup	1997	Dead
CIDF [3]	Open	DARPA	1999	Dead
IDMEF [4]	Open	IETF	2002	Dead
CIEL [5]	Open	MITRE	2002	Dead
SDEE/CIDEE [6]	Proprietary	ICSA Labs	2003	Dead
CBE [7]	Proprietary	IBM, Cisco	2003	Dead
CIM [8]	Open	DMTF	2005	Alive
CEF [9]	Proprietary	ArcSight	2006	Alive
CEE [10]	Open	MITRE	2007	Killed
OLF [11]	Proprietary	eIQNetworks	2007	Dead
WELF [12]	Proprietary	WebTrends	2008	Alive
LEEF [13]	Proprietary	Q1 Labs	2013	Alive
CADF [14]	Open	DMTF	2015	Alive

4. The MITRE promise and other failures

The MITRE corporation in the early 2000 started a series of projects to address the information exchange problem, focusing on different areas of the log management. They decided to kill CIEL (Common Intrusion Event List) and created the CEE (Common Event Expression) Framework, that covers transport (Common Log Transport), syntax (Common Log Syntax), taxonomy (Common Event Expression Taxonomy), and a set of recommendations on when and what log (Common Event Log Recommendations), as part of the plan to build a national cyber information sharing ecosystem composed by *CVE* [15], *CWE* [16], *CAPEC* [17], *ATT&CK* [18], and *CAR*

[19]. But in 2014 someone decided that the CEE was no longer a priority.

As result, we have dictionaries, the words to identify the different parts of the cyber-security puzzle, but we miss the glue: we still use syslog to transmit the analysis produced by cyber security systems like IDS/IPS/EPS/++, and we still must ad-hoc *grok* the message field.

Why is so hard to define a log format? Why have so many initiatives failed?

It is not hard to define a standard! Sure, some proposal was overkill (IDMF), some other too complex to implement, but basically all standards proposed a solution for a problem not recognized by developers. Remember the `"xennet: skb rides the rocket: 19 slots"` message flooding /var/log/syslog for “some reason”! For a coder who creates such meaningful message, all the efforts to standardize content are just waste of time.

5. The future

Still in 2018 we don’t have a standard for security log messages! We have syslog somehow recognized as “universal” transport, that’s all.

The message field remain a land of conquest!

CEF, LEEF, CIM/CADF are the most used and supported formats:

Format	Transport	Encoding	Structure	Num of fields
CEF	syslog	k→v	Loose, no dictionary	91
LEEF	syslog	k→v	Loose, no dictionary	51
CIM	agnostic	agnostic	XML schema	59

At the moment, the winner is one with a loose format, transmitted over syslog, readable by human, and parsable by code. But the IoT will produce most of the messages for a code, not for a user! So, we will probably see more structured and cryptic machine-oriented formats, compressed, and composed more by IDs and less by words.

My fear is that, without a central authority, we will have a lot of git-wanna-be-master-clones.

6. Summary

Especially with the IoT knocking at the door, it is urgent to have a log standard to facilitate the correlation and identification of abnormal behaviors.

Many organization tried to define them, but without success. In the last years, support for CEE and LEEF transported over syslog established a de facto standard, but they are still archaic, big, user oriented logs, in a world demanding more compact, agile, machine centric logs.

7. External Links

[1] https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-92.pdf
[2] http://www.opengroup.org/
[3] https://tools.ietf.org/html/draft-staniford-cidf-data-formats-00
[4] https://www.ietf.org/rfc/rfc4765.txt
[5] https://cve.mitre.org/data/board/archives/2001-03/msg00013.html
[6] https://www.cisco.com/c/en/us/td/docs/security/ips/specs/CIDEE_Specification.html
[7] https://www.ibm.com/support/knowledgecenter/en/SSB23S_1.1.0.14/gtpb1/rcbefmt.html
[8] https://www.dmtf.org/standards/cim
[9] https://kc.mcafee.com/resources/sites/MCAFEE/content/live/CORP_KNOWLEDGEBASE/78000/KB78712/en_US/CEF_White_Paper_20100722.pdf
[10] https://cee.mitre.org/
[11] http://www.openlogformat.org/
[12] https://www.webtrends.com/
[13] https://www.ibm.com/developerworks/community/wikis/form/anonymous/api/wiki/9989d3d7-02c1-444e-92be-576b33d2f2be/page/3dc63f46-4a33-4e0b-98bf-4e55b74e556b/attachment/a19b9122-5940-4c89-ba3e-4b4fc25e2328/media/QRadar_LEEF_Format_Guide.pdf
[14] https://www.dmtf.org/standards/cadf
[15] https://cve.mitre.org/
[16] https://cwe.mitre.org/
[17] https://capec.mitre.org/
[18] https://attack.mitre.org/
[19] https://car.mitre.org/