Personal Information Tagging for Logs

https://tools.ietf.org/html/draft-rao-pitfol-00

Sandeep Rao, Shivan Sahib, Ryan Guest

IETF 106, PEARG

Nov 18 2019

Motiviations

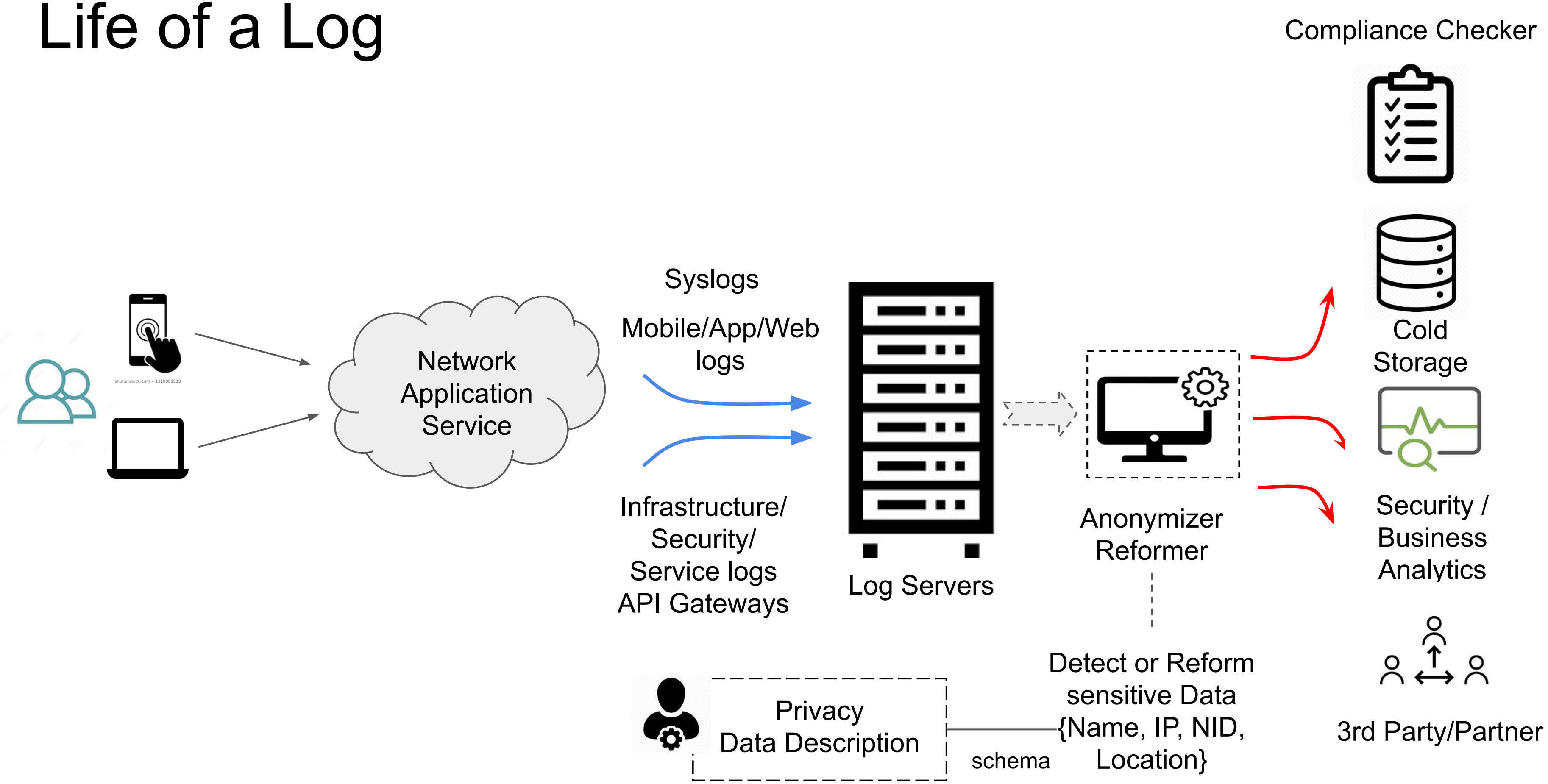
Log Data

- Information recorded by a system

Uses

- Developers - Debugging, Troubleshooting
- Operations - Performance, Maintenance
- Business - Analysis and Marketing
- Security - Profiling, Monitoring, Incident Response

Life of a Log



Compliance Checker

Network Application Service

Syslogs

Mobile/App/Web logs

Infrastructure/ Security/ Service logs API Gateways

Log Servers

Anonymizer Reformer

Cold Storage

Security / Business Analytics

3rd Party/Partner

Detect or Reform sensitive Data {Name, IP, NID, Location}

schema

Privacy Data Description

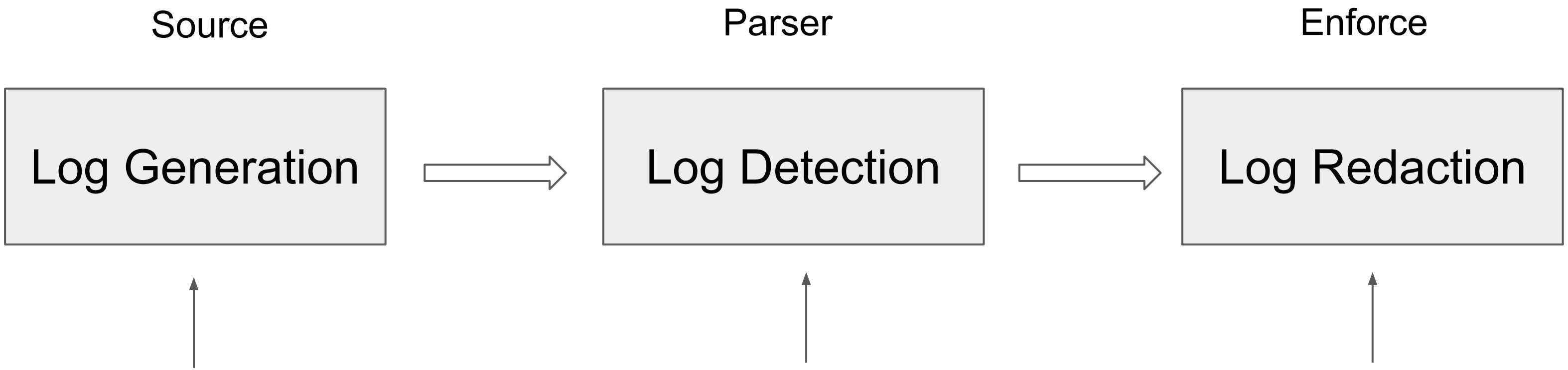
Log Evolution

- Anonymization to Personalization
- Owned to Shared
- Silos to Collaborated
- Insignificant to Vulnerable
- Monitoring to Monetization

Logs and Privacy (2)

- Privacy
  - Which, How, Who
  - Unregulated to Regulated
- Challenges
  - Subjective - many types of log data
  - Dictionary, Data-set based training model
  - Vendor specific schema / privacy policy
  - No standard - (what) data to be protected and (how) redaction

Privacy Framework for Logs



Source

Parser

Enforce

Log Generation

Log Detection

Log Redaction

- Privacy Framework for Logs
  - **Personal Data Tagging**
  - Privacy Policy Assertion

- Detect Personal Data
  - No hard-coding
  - Explicit detection methods

- Enforce Privacy Protection
  - Action based on source intents
  - Differential reduction actions

Personal Information Tagging for Logs

```
<120> Nov 16 16:00:00 10.0.1.11 ABCDEFG: [AF@0 event="AF-Authority failure" violation="A-Not authorized to object"
actual_type="AF-A” jrn_seq="1001363" timestamp="20120418163258988000" job_name="QPADEV000B" user_name="XYZZY"
job_number="256937 err_user="TESTFORAF" ip_addr="10.0.1.21" port="55875" action="Undefined(x00)" val_job="QPADEV000B"
val_user="XYZZY" val_jobno="256937" object="TEST" object_library="CUS9242" object_type="*FILE" pgm_name="" pgm_libr=""
workstation=""]
```

Insert “pii” metadata at Source (Field Level, Log Level)

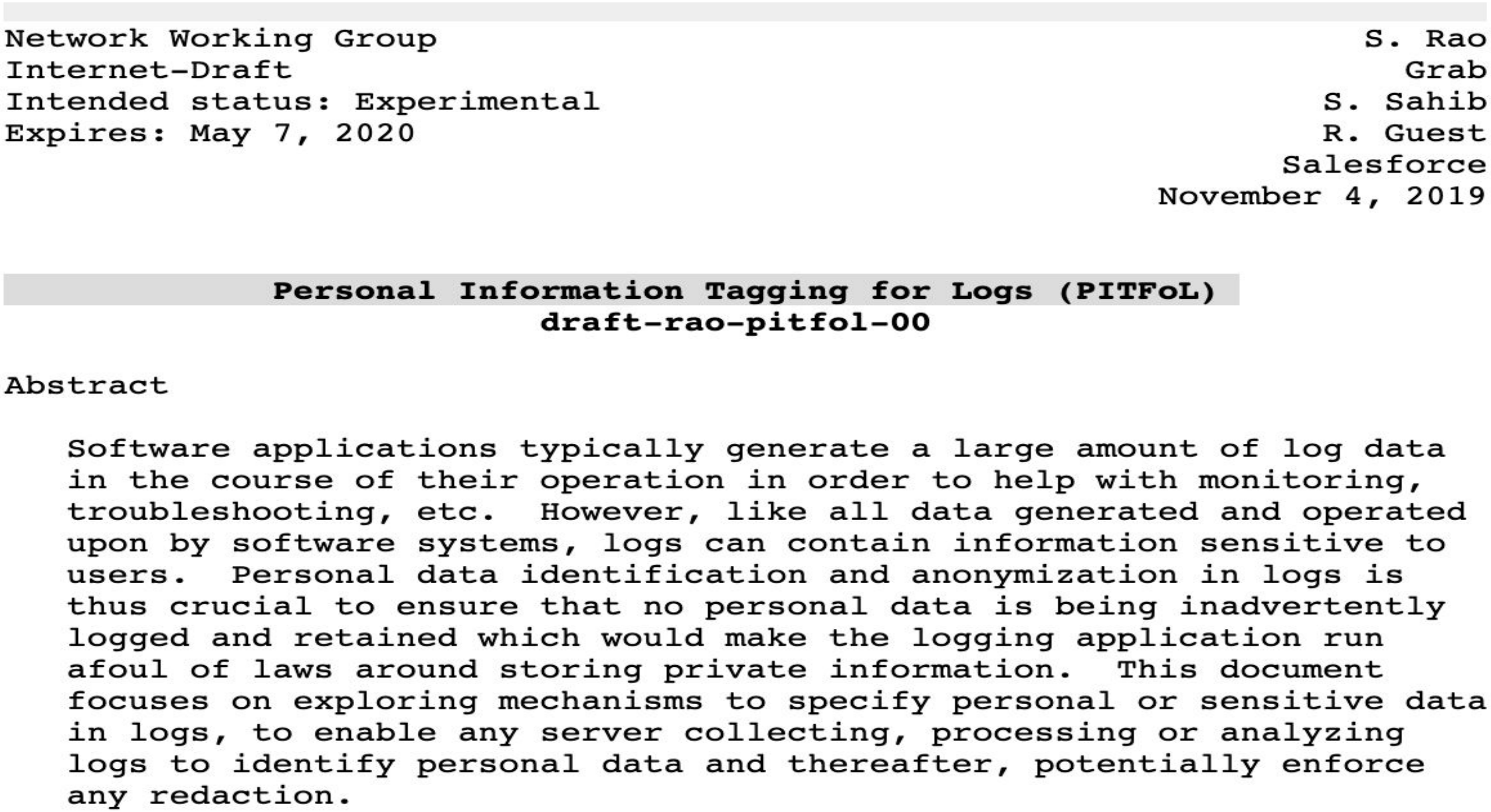
```
<120> Apr 18 16:32:58 10.0.1.11 QAUDJRN: [AF@0 event="AF-Authority failure" violation="A-Not authorized to object"
actual type="AF-A" jrn_seq="1001363" timestamp="20120418163258988000" job_name="QPADEV000B" {user_name="XYZZY"
pii_data="true"} job_number="256937" {err_user="XYZZY" pii_data="true"] [ip_addr="10.0.1.21" pii_data="true"] port="55875"
action="Undefined(x00)" val_job="QPADEV000B" val_jobno="256937" object="TEST" object_library="CUS9242" object_type="*FILE"
pgm_name="" pgm_libr="" workstation=""]
```

```
<120> Apr 18 16:32:58 10.0.1.11 QAUDJRN: [AF@0 event="AF-Authority failure" violation="A-Not authorized to object"
actual_type="AF-A" jrn_seq="1001363" timestamp="20120418163258988000" job_name="QPADEV000B" user_name="XYZZY"
job_number="256937" {err_user="XYZZY" pii_data="true"] [ip_addr="10.0.1.21" pii_data="true"] port="55875"
action="Undefined(x00)" val_job="QPADEV000B" val jobno="256937" object="TEST" object_library="CUS9242" object_type="*FILE"
pgm_name="" pgm_libr="" workstation="", pii={"user_name,err_user, ip_addr"}]
```

Current Vendor Approaches (Prior Art)

- dynatrace OneAgent configuration file
- Log4j Framework by Apache
- apigee - custom variables prefixed with “`private`” and masking configuration
- Avi Networks - Hiding PII in Logs using “`sensitive_log_profile`”

PITFoL-00



Network Working Group
Internet-Draft
Intended status: Experimental
Expires: May 7, 2020

S. Rao
Grab
S. Sahib
R. Guest
Salesforce

November 4, 2019

Personal Information Tagging for Logs (PITFoL)
draft-rao-pitfol-00

Abstract

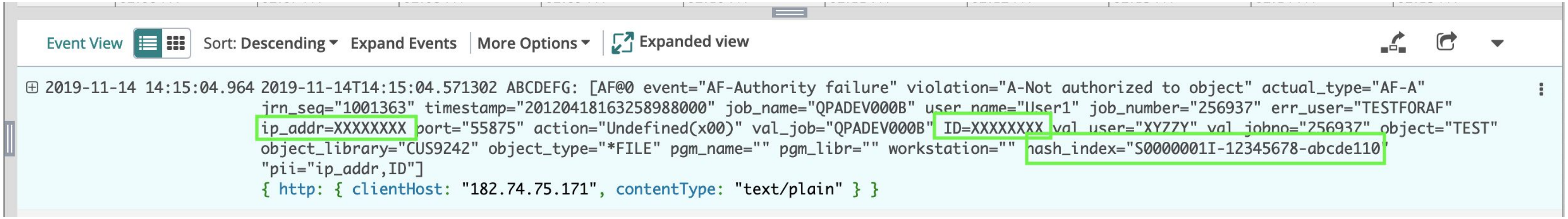
Software applications typically generate a large amount of log data in the course of their operation in order to help with monitoring, troubleshooting, etc. However, like all data generated and operated upon by software systems, logs can contain information sensitive to users. Personal data identification and anonymization in logs is thus crucial to ensure that no personal data is being inadvertently logged and retained which would make the logging application run afoul of laws around storing private information. This document focuses on exploring mechanisms to specify personal or sensitive data in logs, to enable any server collecting, processing or analyzing logs to identify personal data and thereafter, potentially enforce any redaction.

Hackathon Results

Log Input

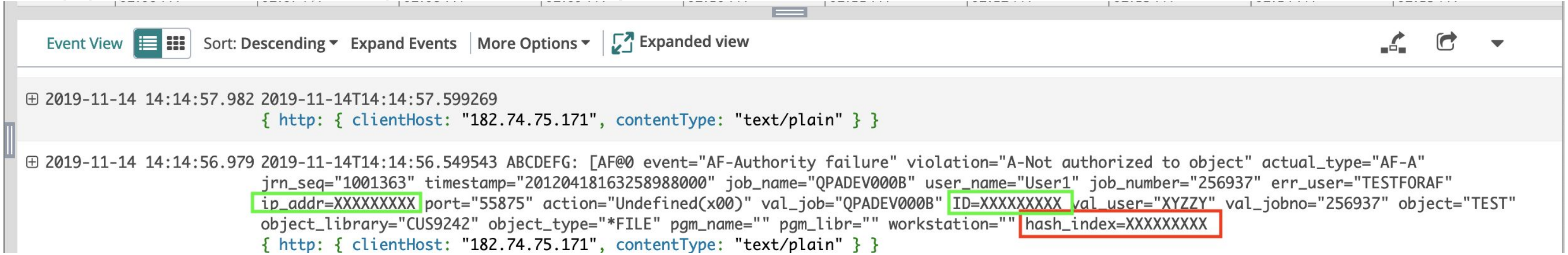
```
ABCDEFG: [AF@0 event="AF-Authority failure" violation="A-Not authorized to object" actual_type="AF-A" jrn_seq="1001363"
timestamp="20120418163258988000" job_name="QPADEV000B" user_name="User1" job_number="256937" err_user="TESTFORAF"
ip_addr="10.0.1.21" port="55875" action="Undefined(x00)" val_job="QPADEV000B" ID="S0000001I" val_user="XYZZY"
val_jobno="256937" object="TEST" object_library="CUS9242" object_type="*FILE" pgm_name="" pgm_libr="" workstation=""
hash_index="S0000001I-12345678-abcde110" "pii="ip_addr,ID"]
```

Loggly Ouput (Tagging Method)



```
⊞ 2019-11-14 14:15:04.964 2019-11-14T14:15:04.571302 ABCDEFG: [AF@0 event="AF-Authority failure" violation="A-Not authorized to object" actual_type="AF-A"
                         jrn_seq="1001363" timestamp="20120418163258988000" job_name="QPADEV000B" user_name="User1" job_number="256937" err_user="TESTFORAF"
                         ip_addr=XXXXXXXX port="55875" action="Undefined(x00)" val_job="QPADEV000B" ID=XXXXXXXX val_user="XYZZY" val_jobno="256937" object="TEST"
                         object_library="CUS9242" object_type="*FILE" pgm_name="" pgm_libr="" workstation="" hash_index="S0000001I-12345678-abcde110"
                         "pii="ip_addr,ID"]
                         { http: { clientHost: "182.74.75.171", contentType: "text/plain" } }
```

Loggly Output (Regex Method)



```
⊞ 2019-11-14 14:14:57.982 2019-11-14T14:14:57.599269
                         { http: { clientHost: "182.74.75.171", contentType: "text/plain" } }

⊞ 2019-11-14 14:14:56.979 2019-11-14T14:14:56.549543 ABCDEFG: [AF@0 event="AF-Authority failure" violation="A-Not authorized to object" actual_type="AF-A"
                         jrn_seq="1001363" timestamp="20120418163258988000" job_name="QPADEV000B" user_name="User1" job_number="256937" err_user="TESTFORAF"
                         ip_addr=XXXXXXXXX port="55875" action="Undefined(x00)" val_job="QPADEV000B" ID=XXXXXXXXX val_user="XYZZY" val_jobno="256937" object="TEST"
                         object_library="CUS9242" object_type="*FILE" pgm_name="" pgm_libr="" workstation="" hash_index=XXXXXXXXX
                         { http: { clientHost: "182.74.75.171", contentType: "text/plain" } }
```

Next Steps

- Feedback?
- Request for RG adoption?