IETF 83
CloudLog

Gene Golovinsky
ggolovinsky@qualys.com

Paris,

March 25-30, 2012

The Traditional Logging

Practically all managed hw and sw entities log their activities

- Network elements and Unix/Linux servers – Syslog
- Windows servers – Windows event facility
- Applications – proprietary files and localized string formats

Despite difference in format they all generally have enough information to identify:

- Actual entity
- Type of activities
- Time of occurrence
- Often user identity

Applications of Traditional Logging

Variety of applicable problem spaces:

- System and application management
- Network management
- SIEM
- Forensics
- Auditing
- Regulations and Compliance (PCI, SOX, HIPPA)

Applicability depends on the availability and accuracy of the data

Challenges with Shared Resources

The same applications of logs are relevant for The Shared Resource deployment or Cloud

- System management & network management
  - For service providers
- SIEM
- Forensics
  - Service providers and customers
- Auditing
  - Service providers and customers
- Regulations and Compliance (PCI, SOX, HIPPA)
  - Service providers and customers

Cloud/Shared Resources Challenges

Jul 7 00:03:15 192.168.23.1 %SEC-6-IPACCESSLOGP: list Oif-in denied tcp 66.43.204.165(4118) -> 255.255.255.255(111), 1 packet Jul 7 00:03:15 192.168.23.1 %SEC-6-IPACCESSLOGP:

- We really can't tell where 192.168.23.1 is located and which physical entity it's associated with.
- We also can't tell what else is on the same physical entity
- We also can't tell who has access to the same entity
- User identity is obfuscated

- THIS IS A BIG PROBLEM FOR *ANY APPLICATION!!!*

What We Need to Support These Applications?

Track complete user interactions with shared resource components

- All activities should have complete audit trail from the initial request to the component, from authentication, impersonation if applicable, to the modification of the resource and the success or failure of the operation

Audit Real and Effective User Identities

- All activities should be tracked with real (authenticated) and effective (impersonated) user identities.

Track Transit

- Need to track the location of the entity that is involved in the activity. These locations could be highly dynamic, sometimes even temporary or short term resources. The audit trail should include a facility to track where requests originated, and any entity locations it passed through.

What We Need to Support These Applications?

Use Syslog format

RFC 5424

- STRUCTURED-DATA : SD-ELEMETs, SD-IDs, SD-PARAMs
- Well defined, extensible, easy to understand and parse format

Define SD-ELEMENTS that are specific to Cloud Computing environment that would be mandatory for all

Allow providers and vendors to define their own SD-ELEMENTS that enable to log specifics about their implementations

- Much like enterprise MIB in SNMP

What We Need to Support These Applications? Examples

Simple non-authenticated request will produce a log

- Jul 7 09:01:40 10.0.6.94 api_aaa: AAA00000I [context@999999 aid="9BE817EB-8ACC-1004-D9DF-00000A00065E"][transit@999999 client="56.2.222.83"]
- Where
  - SD-ID – 'context@999999'
    - SD-PARAM - 'aid' – mandatory audit identifier
  - SD-ID - transit@999999
    - SD-PARAM – 'client' – mandatory IP of a client making a request (different from SD-ID 'ip' in RFC 5424 that defines IP of the entity producing the log)
    - SD-PARAM – 'gw' – optional The value is a concatenation of the string form of a UUID, identifying the gateway, a colon character (i.e. ':'), and finally the IP address on the gateway through which the request was received. It could be multiple 'gw' SD-PARAMS in the log.

The same request passing through the gateway

- Jul 7 09:01:40 10.0.6.94 api_aaa: AAA00000I [context@999999 aid="9BE817EB-8ACC-1004-D9DF-00000A00065E"][transit@999999 client="56.2.222.83" gw="37CB88DB-8AE3-1004-CBED-00007F000001:10.0.11.9"]

What We Need to Support These Applications? More Examples

Failed Authentication

- Jul 7 09:01:40 10.0.6.94 api_aaa: AAA00050I [context@999999 aid="9BE817EB-8ACC-1004-D9DF-00000A00065E"][transit@999999 client="10.0.6.94"] authentication failed, invalid password
- Where
  - SD-ID – 'context@999999'
    - SD-PARAM - 'aid' – mandatory audit identifier
    - SD-PARAM – 'rid' – optional parameter represents the real user identifier
  - SD-ID - transit@999999
    - SD-PARAM – 'client' – mandatory IP of a client making a request (different from SD-ID 'ip' in RFC 5424 that defines IP of the entity producing the log)

Successful Authentication

- Jul 7 09:01:40 10.0.6.94 api_aaa: AAA00001I [context@999999 aid="9BE817EB-8ACC-1004-D9DF-00000A00065E" rid"2:510"][transit@999999 client="10.0.6.94"] authentication successful for cid = 2 uid = 510

What We Need to Support These Applications? More Examples

Invalid Request Parameter: request log -> response log

- **Request log:**
  - Jul 7 09:01:40 10.0.6.94 inetsmgr: INM00150I [context@999999 aid="9BE817EB-8ACC-1004-D9DF-00000A00065E" rid"2:520" eid="1023:6022"][transit@999999 client="10.0.6.94"] "10.0.6.94" - "-" "GET /api/user/manager:search?customer_id=2&username=foo&bar=baz HTTP/1.1" 400 215
- **Response log:**
  - Jul 7 09:01:40 10.0.6.94 umg: UMG00000I [context@999999 aid="9BE817EB-8ACC-1004-D9DF-00000A00065E" rid"2:520" eid="1023:6022"][transit@999999 client="10.0.6.94"] invalid query parameter "baz" specified on request
- **Where**
  - SD-ID – 'context@999999'
    - SD-PARAM - 'aid' – mandatory audit identifier
    - SD-PARAM – 'rid' – optional parameter represents the real (authenticated) user identifier
    - SD-PARM – 'eid' – optional parameter represents effective (impersonated) user identifier
  - SD-ID - transit@999999
    - SD-PARAM – 'client' – mandatory IP of a client making a request (different from SD-ID 'ip' in RFC 5424 that defines IP of the entity producing the log)

Proposed Next Steps

I believe that logging and auditability of the cloud and in the cloud is crucial for its adoption

- Particularly by enterprises

These problems are very real

- Heard it from several operators and customers

IETF's job to support industry with interoperable and secure mechanisms

SO…

Proposed Next Steps

Accept this as a work item

- Cloud related work is different from the traditional static topology network:
  - Transitional and obfuscated identity
  - Managed entity location
  - Access to the shared resources
  - Resource multi-tenancy
  - Different type of security concerns – resource theft, diminished audit, detection, and incident response capabilities
  - And more…

The Scope of Cloud Log

Q: What is in scope for Cloud Log?

A: Log data format for cloud-based applications and entities

Q: What is out of scope for Cloud Log?

A: Log data transport, protocols, semantics, language bindings, toolkits.

The Status of Work

Internet Draft:

- http://tools.ietf.org/html/draft-golovinsky-cloud-services-log-format-02
  - Latest update was in March 2012
  - Expires in September 2012
  - More edits and clarifications are on the way
- Progress to OPSAWG ?

Generated some controversy

- Proponents of CEE – Common Event Expression - do not like the proposal
  - Confusion between format and content
  - CEE defines dictionary and event taxonomy
  - Cloud Log defines extensible format

And now…

QUESTIONS?

SUGGESTIONS?

COMMENTS?