Sightings Use Cases

Sebastien Tricaud
Devo Inc.

~~March 11 2020~~
May 6 2020

Section 1

Introduction

Weiner's Law of Libraries

There are no answers, only cross-references.

Sightings, explained

192.168.42.22

Sightings, explained

192.168.42.22

What is this?

Sightings, explained

192.168.42.22

- IP address
- IP version 4 address
- Problably has the mask 255.255.255.0 and CIDR 192.168.42.0/24
- RFC 1918: 192.168/16 prefix is a Private Network

A local IP? can we trust this?

A local IP? can we trust this?

Would you trust any doctor because they are wearing a mask?

Sightings, explained

192.168.42.22

- Enrichment
  - Who is talking to it?

Sightings, explained

192.168.42.22

- Enrichment
  - Who is talking to it?
  - Has it replied to requests on the HTTP port?

Sightings, explained

192.168.42.22

- Enrichment
  - Who is talking to it?
  - Has it replied to requests on the HTTP port?
  - After packet inspection, is that HTTP?

Sightings, explained

192.168.42.22

- Enrichment
  - Who is talking to it?
  - Has it replied to requests on the HTTP port?
  - After packet inspection, is that HTTP?
  - Is that HTTP traffic an HTTP server would make?

Sightings, explained

192.168.42.22

- Enrichment
  - Who is talking to it?
  - Has it replied to requests on the HTTP port?
  - After packet inspection, is that HTTP?
  - Is that HTTP traffic an HTTP server would make?
  - Is that HTTP traffic and HTTP server would make ALL THE TIME?

Sightings, explained

192.168.42.22

- Enrichment
  - Who is talking to it?
  - Has it replied to requests on the HTTP port?
  - After packet inspection, is that HTTP?
  - Is that HTTP traffic an HTTP server would make?
  - Is that HTTP traffic and HTTP server would make ALL THE TIME?
  - Are we sure this is an HTTP server?

$\Longrightarrow$ Could go in too many directions!

Sightings, explained

192.168.42.22

- Enrichment
  - Who is talking to it?
  - Has it replied to requests on the HTTP port?
  - After packet inspection, is that HTTP?
  - Is that HTTP traffic an HTTP server would make?
  - Is that HTTP traffic and HTTP server would make ALL THE TIME?
  - Are we sure this is an HTTP server?

$\Longrightarrow$ Could go in too many wrong directions!

Does that IP address mean something in my Industry?

Global Threat Intelligence

- There is no Global Threat Intelligenge
- That local IP address does not mean anything Globally

Global Threat Intelligence

- There is no Global Threat Intelligenge
- That local IP address does not mean anything Globally
  - However it could be that hardcoded IP address from an Industry Equipment
  - Or something used by your own organization, it has a meaning just for you

Global Threat Intelligence

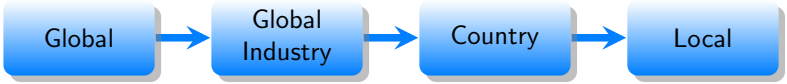
- There is no Global Threat Intelligenge
- That local IP address does not mean anything Globally
  - However it could be that hardcoded IP address from an Industry Equipment
  - Or something used by your own organization, it has a meaning just for you
- Global Threat Data, includes everything everybody knows

Global Threat Intelligence

- There is no Global Threat Intelligenge
- That local IP address does not mean anything Globally
  - However it could be that hardcoded IP address from an Industry Equipment
  - Or something used by your own organization, it has a meaning just for you
- Global Threat Data, includes everything everybody knows
  - Anti-virus signature, OSINT
  - IDS signature etc...

Context

Closer is better.



Conclusion

Threat Intelligence Community must work together to provide tools, indicators, methodologies which enable you to be **as close as possible to the place where data is created**.

Section 2

Sightings

As Pierre Légaré once said...

Winston Churchill had enough of people entering his office and leaving promptly.

As Pierre Légaré once said...

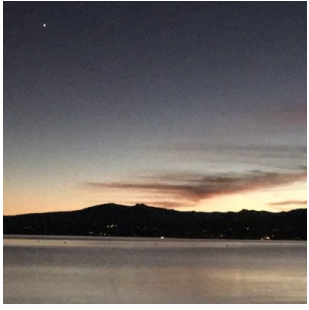
Winston Churchill had enough of people entering his office and leaving promptly. They were all apologizing!

As Pierre Légaré once said...

Winston Churchill had enough of people entering his office and leaving promptly.
They were all apologizing!
He asked his initials to be removed from his office door.

Sightings allow to share an Observation, rather than a Reputation.

Sightings is...



- How many times have we seen something?
- From when?
- Until when?

Storing is cheap!

- The entire IPv4 space is only 2^{32}
- 4.3 billion addresses
- Takes 4Gb of storage

- Faup, the URL parser (`https://github.com/stricaud/faup/`)
- From my experience with proxy logs, several hundred of thousand users
- About 15 000 unique URLs per week in average

Who is standardizing around Sightings?

- The MISP Project
  - https://www.misp-standard.org/rfc/sightingdb-format.txt
- ATT&CK
  - https://attack.mitre.org/resources/sightings/
- OASIS STIX v2
  - https://oasis-open.github.io/cti-documentation/stix/intro.html
  - https://docs.google.com/document/d/1IvkLxg_tCnICsatu2lyxKmWmh1gY2h8HUNssKIE-UIA/

Sightings in OASIS STIX v2

Definition

A Sighting denotes the belief that something in CTI (e.g., an indicator, malware, tool, threat actor, etc.) was seen.

Sightings in OASIS STIX v2

Definition

A Sighting denotes the belief that something in CTI (e.g., an indicator, malware, tool, threat actor, etc.) was seen.

Establishes relationships.

1. What was sighted: **sighting_of_ref**
2. Who/Where: **where_sighted_refs**
3. What was seen: **observed_data_refs**

Sightings in OASIS STIX v2

Definition

A Sighting denotes the belief that something in CTI (e.g., an indicator, malware, tool, threat actor, etc.) was seen.

Establishes relationships.

1. What was sighted: **sighting_of_ref**
2. Who/Where: **where_sighted_refs**
3. What was seen: **observed_data_refs**

Properties.

- **Common Properties**: type, id, created_by_ref, created, modified, revoked, labels, external_references, object_marking_refs, granular_markings
- **Sighting Specific Properties**: first_seen, last_seen, count, sighting_of_ref, observed_data_refs, where_sighted_refs, summary

TBD: Things that could be improved in OASIS STIX v2

Definition

A Sighting denotes the belief that something in CTI (e.g., an indicator, malware, tool, threat actor, etc.) was seen.

TBD: Things that could be improved in OASIS STIX v2

Definition

A Sighting denotes the belief that something in CTI (e.g., an indicator, malware, tool, threat actor, etc.) was seen.

Definition

A Sighting denotes the belief that something in CTI (e.g., an indicator, malware, tool, threat actor, etc.) was seen**, searched or accessed**.

TBD: Things that could be improved in OASIS STIX v2

"that something in CTI" can only **count** up to 999,999,999.

TBD: Things that could be improved in OASIS STIX v2

"that something in CTI" can only **count** up to 999,999,999.

count (optional)	`integer`	This **MUST** be an integer between 0 and 999,999,999 inclusive and represents the number of times the SDO referenced by the `sighting_of_ref` property was sighted.

TBD: Things that could be improved in OASIS STIX v2

"that something in CTI" can only **count** up to 999,999,999.

count (optional)	`integer`	This **MUST** be an integer between 0 and 999,999,999 inclusive and represents the number of times the SDO referenced by the `sighting_of_ref` property was sighted.

From the JSON Standard:

```
numbers that are integers and are in the
range [-(2**53)+1, (2**53)-1] are interoperable
in the sense that implementations will agree
exactly on their numeric values
```

```
>>> 2**53-1
9007199254740991
```

TBD: Things that could be improved in OASIS STIX v2

Establishes relationships.

1. What was sighted: **sighting_of_ref**
2. Who/Where: **where_sighted_refs**
3. What was seen: **observed_data_refs**

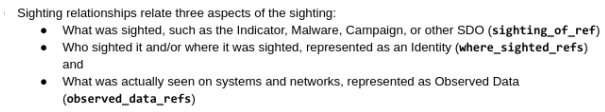
1 and 2 are the same.

TBD: Things that could be improved in OASIS STIX v2

Establishes relationships.

1. What was sighted: **sighting_of_ref**
2. Who/Where: **where_sighted_refs**
3. What was seen: **observed_data_refs**

1 and 2 are the same.



Sighting relationships relate three aspects of the sighting:

- What was sighted, such as the Indicator, Malware, Campaign, or other SDO (**`sighting_of_ref`**)
- Who sighted it and/or where it was sighted, represented as an Identity (**`where_sighted_refs`**) and
- What was actually seen on systems and networks, represented as Observed Data (**`observed_data_refs`**)

ATT&CK Sightings

Defines four types:

- direct-technique-sighting
- direct-software-sighting
- indirect-software-sighting
- technique

```
{
  "id": "32",
  "sightingType": "direct-malware-sighting",
  "startTime": "2019-01-01T08:12:00Z",
  "endTime": "2019-01-01T08:12:00Z",
  "detectionType": "raw",
  "sectors": ["healthcare"],
  "software": "MacSpy"
}
```

Very pragmatic, immediately useful.

Sightings in MISP

As usual, start with an implementation, learn, then write a specification.

- SQL backend
- SightingDB

Sightings in MISP

SightingDB format

This document describes the format used by SightingDB to give automated context to a given Attribute by counting occurrences and tracking times of observability. SightingDB was designed to provide to MISP and other tools an interoperable, scalable and fast way to store and retrieve attributes sightings.

Specification: TXT - HTML

Lead: Devo Inc

`www.misp-standard.org/rfc/sightingdb-format.txt`

Definition

Defines an JSON format to fetch and push sightings. A single one. A bulk.

Section 3

SightingDB

Doing Threat Intelligence work without Sightings is similar to not knowing root shell commands starts with a # so pasting does not hurt!

Doing Threat Intelligence work without Sightings is similar to not knowing root shell commands starts with a # so pasting does not hurt!

The objective for this section is to help you leveling up your game. Feedback and Criticisms are encouraged!

SightingDB

`https://github.com/stricaud/sightingdb/`
A Scalable Sighting Database, hybrid in-memory/on-disc whose goal is to provide an easy to use way to count attributes.

- Work sponsored by Devo Inc.
- Available under MIT license

Design

- Modeled after Zookeeper for its key-value store capability:
  - a key is a namespace, such as “foo/bar” where “bar” is a child of “foo”.
  - it allows to create as many placeholders as anyone dream
  - a value is simply a string

Why not Redis?

- Redis is not tailored for our very specific use-case
- Incrementing a value (INCR) in Redis is atomic
- Atomic means a lock on the key for writing, preventing multiple threads / resources to increment at the same time

Run a SightingDB instance

```
$ docker pull sightingdb/sightingdb
```

SightingDB Client API

$ pip3 install sightingdb

SightingDB Client API

Writing

```
import sightingdb
con = sightingdb.connection(host="localhost", apikey="changeme")
writer = sightingdb.writer(con)
writer.add("/key/namespace1", "pypi.org")
writer.add("/key/namespace1", "pypi.org")
writer.add("/key/namespace2", "example.com")
writer.commit()
```

SightingDB Client API

Reading

```
import sightingdb
con = sightingdb.connection(host="localhost", apikey="changeme")
reader = sightingdb.reader(con)
reader.add("/key/namespace1", "pypi.org")
reader.add("/key/namespace2", "example.com")
for i in reader.fetch():
    print(str(i))
```

REST API

```
$ curl -k https://localhost:9999/w/foo/bar/?val=hello
{"message":"ok"}
```

REST API

```
$ curl -k https://localhost:9999/r/foo/bar/?val=hello
{"value":"hello","first_seen":1581627580,
 "last_seen":1581627580,"count":1,"tags":"",
 "ttl":0, "consensus":1}
```

Support

- OASIS STIX v2 Sightings
  tag="oasis-stixv2:id="sighting–ee20065d-2555-424f-ad9e-0f8428623c75""
- ATT&CK Sightings
  Simply use the namespace and tags.

Key being a namespace, powerful.

- Want to be compatible with ATT&CK? **/direct-software-sighting/JCry**
- Want to store relationships with a particular IP in the finance BU? **/finance/8.8.8.8/**
- Want to store a url? **/url/**
- Want to store the url for all TLD in ch? **/url/tld/ch/**
- Want to store the ch TLD related URLs to find them faster? **/ch/tld/url/**
- Want to see how many times somebody searched for the value `https://www.stadt-zuerich.ch` from */url/*? **Shadow Sightings!**

Consensus: same value, many namespaces

```
$ curl -H "Authorization: changeme" -k https://localhost:9999/w/213.208.154.14?val=192.168.0.28
{"message":"ok"}
$ curl -H "Authorization: changeme" -k https://localhost:9999/w/1.1.1.1?val=192.168.0.28
{"message":"ok"}
```

Consensus: same value, many namespaces

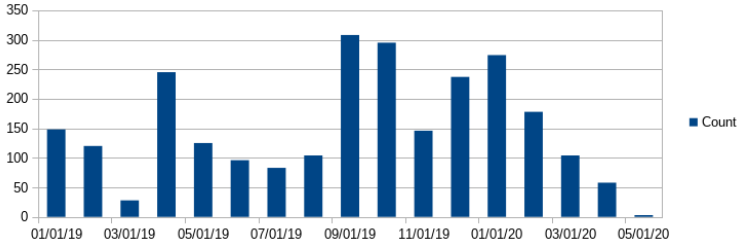
```
$ curl -H "Authorization: changeme" -k https://localhost:9999/w/213.208.154.14?val=192.168.0.28
{"message":"ok"}
$ curl -H "Authorization: changeme" -k https://localhost:9999/w/1.1.1.1?val=192.168.0.28
{"message":"ok"}

$ curl -H "Authorization: changeme" -k https://localhost:9999/r/213.208.154.14?val=192.168.0.28
{"value":"192.168.0.28","first_seen":1588741064,"last_seen":1588741064,"count":1,"tags":"",
"ttl":0,"consensus":2}
```

Get all values from a namespace

```
curl -H "Authorization: changeme" -k https://localhost:9999/r/213.208.154.14
{
    "attributes": [
        {"value":"192.168.0.56","first_seen":1588740594,"last_seen":1588742432,"count":3029,\
         "tags":"","ttl":0,"consensus":18},
        {"value":"192.168.0.28","first_seen":1588741064,"last_seen":1588741064,"count":8432,\
         "tags":"","ttl":0,"consensus":10},
        {"value":"192.168.0.42","first_seen":1588741593,"last_seen":1588742486,"count":12,\
         "tags":"","ttl":0,"consensus":57}
    ]
}
```

Statistics




```
{"value":"192.168.42.22","first_seen":1546300800,"last_seen":1588291200,"count":2552,
 "tags":"","ttl":0,
 "stats":{"1546300800":148,"1548979200":120,"1551398400":28,"1554076800":245,"1556668800":125,
          "1559347200":96,"1561939200":83,"1564617600":104,"1567296000":308,"1569888000":295,
          "1572566400":146,"1575158400":237,"1577836800":274,"1580515200":178,"1583020800":104,
          "1585699200":58,"1588291200":3},
 "consensus":1}
```

Pin-point low and slow for 192.168.42.22

- This is a total of 2552 events
- Guarantee it will never show in your SIEM dashboard
- SightingDB allows you to add a new hunting capability
- You will know it was a low and slow

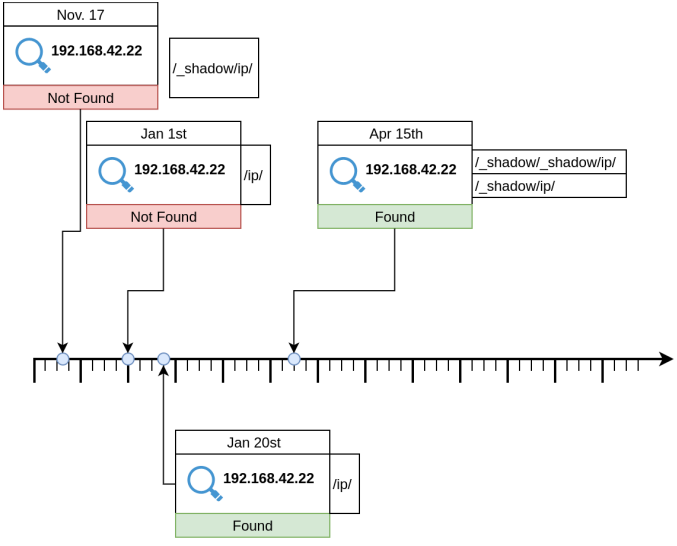
Shadow Sightings

- When we read, we write!

Shadow Sightings

- When we read, we write!
- How many time did somebody searched for a value in a namespace?
- SightingDB stores automatically into **/_shadow/**
- SightingDB also stores recursive access, enabling “one searched that one searched”

Shadow Sightings



SightingDB, a new type of database

Tailored for Sightings, just Sightings.
SightingDB is to Graph Databases what NoSQL Databases are to SQL. Very pragmatic and use-case centric.

Section 4

Conclusive Use-Cases

Behavior

Definition

Infection: Someone did lots of access since the last few hours

$\Longrightarrow$ Check when a value has its first_seen

Mapping

Definition

Mapping: All the MISP Attributes mapped to ATT&CK for Enterprise

$\Longrightarrow$ Write to SightingDB using the ATT&CK namespace

Shared values

Definition

Consensus: Knowing how many namespaces contain this value

$\Longrightarrow$ Done automatically, simply retrieve a value and get its consensus.

Ubiquitous

Definition

Browsing: Looking at all available data

$\Longrightarrow$ 1:1 check for each value

Specific Attacks

Definition

Ransomware: Show me all the machines impacted

$\Longrightarrow$ Map to the MISP Ransomware Galaxy for its namespace

Recap

- Sightings are Observations, not Reputation
- Increase the understanding of attacks
- Wayback machine at scale
- Already in MISP

Thank You!

sebastien.tricaud@devo.com
@tricaud