Carnegie Mellon
Software Engineering Institute

CERT
Situational
Awareness

The State of Standardization Efforts to support Data Exchange in the Security Domain

Roman Danyliw <rdd@cert.org>

FloCon 2004: Standards Talk

CERT® Network Situational Awareness Group
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213-3890

The CERT Network Situational Awareness Group is part of the Software Engineering Institute. The Software Engineering Institute is sponsored by the U.S. Department of Defense.

© 2004 by Carnegie Mellon University

Overview

- Flow and Packet Formats
- Alert and Event Formats
- Context-relevant Formats

© 2004 by Carnegie Mellon University

Dimensions in Representation

- Usage of representation
  - Transport vs. analysis vs. storage vs. archive
- Volume of data informs representation choice
  - Raw vs. Summaries
    - Choice often dictates a binary vs. text implementation
- Policy Scope
  - Intra-Organizational
    - Little consensus from outsiders necessary
    - Interoperation focus
  - Inter-Organizational
    - Privacy issues more acute (sanitization, filtering)
    - Common semantics are more relevant
    - Efficiency of representation is more significant

© 2004 by Carnegie Mellon University

Formats of interest

- Flow and Packet Formats
  - IPFIX
  - PSAMP
- Alert and Event Formats
  - IDWG
  - INCH
- Context-relevant Formats
  - Vulnerability Report
  - CRISP

© 2004 by Carnegie Mellon University

Flow and Packet Formats (*de facto*)

- PCAP (tcpdump)
  - http://www.tcpdump.org
- Cisco NetFlow

© 2004 by Carnegie Mellon University

IETF IP Flow Information Export (IPFIX) WG

http://www.ietf.org/html.charters/ipfix-charter.html

- Binary, extensible information model for IP flows exported from a given *observation point* (i.e., router line-card) to a *collector*
  - Based on Cisco Netflow v9
- Designates a mandatory protocol (SCTP) to use in the transport of these flows

(Note: Various text and figures were taken from the IPFIX I-Ds)

© 2004 by Carnegie Mellon University

IPFIX Flow Definition

- "… [A] set of IP packets passing an observation point … during a certain time interval. All packets belonging to a particular flow have a set of common properties [named flow keys]."
  - One or more packet header field (e.g. destination IP address), transport header field (e.g. destination port number), or application header field (e.g. RTP header fields)
  - One or more characteristics of the packet itself (e.g. number of MPLS labels)
  - One or more fields derived from packet treatment (e.g. next hop IP address, output interface)

© 2004 by Carnegie Mellon University

IPFIX Flow Definition (2)

- A flow is defined by a *flow type* function that considers the various *flow keys*

- Flexible definition provides support for:
  - Filtering
  - Sampling
  - Bi-directional and unidirectional flows

© 2004 by Carnegie Mellon University

IPFIX Information Model

- Template-based format
  - IPFIX merely specifies the possible
    - data types (e.g., IPv4 address, octet) and the
    - information items (e.g., icmpTypeCode, egressInterface)
  - Information items are unique identifiers registered with IANA or escaped via a vendor code
  - A template is merely an ordered list of pairs: <information items (i.e., fieldID), data length>
    - No static format; can be dynamically generated during the export process

© 2004 by Carnegie Mellon University

IPFIX Information Model (2)

- Two classes of records
  - Template Records
    - Describe a format
  - Data Records
    - Contain data encoded and formatted according to a Template record
- Two flavors of Data Records; those that encode the:
  - Data stream (e.g., observed flows), and
  - Control Information (e.g., selection criteria)

© 2004 by Carnegie Mellon University

IPFIX Information Model (3)

- 4-basic record types
  - Flow Data Template
    - A description for data record structure
  - Flow Data Record
    - IP flows formatted according to the Flow Data Template
  - Option Template
    - A description of the option record structure
  - Option Record
    - Control information formatted according to the Option Template Record

© 2004 by Carnegie Mellon University

IPFIX Messaging

```
+--------+-----------------------------------------------+
|        | +---------+     +---------+      +---------+ |
|Message | | Data    | ... | Data    | ...  | Data    | |
| Header | | Template| ... | Record  | ...  | Record  | |
|        | +---------+     +---------+      +---------+ |
+--------+-----------------------------------------------+
```

- Template records are sent inline with the data records
  - Frequency dictated by the quality of transport
  - Possible to send no template in an export, and reference a previously sent template in the data record
    - Collector must cache data templates

© 2004 by Carnegie Mellon University

IPFIX Message Header

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Version Number          |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Export Time                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Sequence Number                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Source ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- **128-byte preamble sent with each export**

© 2004 by Carnegie Mellon University

IPFIX Example

Src IP addr.	Dst IP addr.	Packet Number	Bytes Number
198.168.1.12	10.5.12.254	5009	5344385
192.168.1.27	10.5.12.23	748	388934

Flow Information to Export

```
+--------+-----------------------------------------------
|        | +--------------+ +----------------------+
|Message | | Data         | | Data                 |
| Header | | Template     | | Records              |
|        | | (1 Template) | | (2 Flow Data Records) |
|        | +--------------+ +----------------------+
+--------+-----------------------------------------------
```

IPFIX Encoding Format

© 2004 by Carnegie Mellon University

IPFIX Example: Template (2)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       FlowSet ID = 0          |      Length = 24 bytes        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Template ID 256         |       Field Count = 4         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     IP_SRC_ADDR = 0x0008      |       Field Length = 4        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     IP_DST_ADDR = 0x000C      |       Field Length = 4        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       IN_PKTS = 0x0002        |       Field Length = 4        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       IN_BYTES = 0x0001       |       Field Length = 4        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

© 2004 by Carnegie Mellon University

IPFIX Example: Data (3)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       FlowSet ID = 256        |          Length = 36          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          198.168.1.12                         |  #1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          10.5.12.254                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             5009                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            5344385                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          192.168.1.27                         |  #2
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           10.5.12.23                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              748                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            388934                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

© 2004 by Carnegie Mellon University

IPFIX Transport Protocol: SCTP

- Reliable service
  - TCP equivalent
- “Partially reliable” service
  - During un-congested periods, all the records marked for deletion under congestion will be reliably delivered
  - During congested periods, the exporter will drop packets to protect the network

© 2004 by Carnegie Mellon University

IPFIX I-Ds

- Requirements for IP Flow Information Export
  - draft-ietf-ipfix-reqs-16
- Architecture Model for IP Flow Information Export
  - draft-ietf-ipfix-architecture-03
- Information Model for IP Flow Information Export
  - draft-ietf-ipfix-info-03
- IPFIX Protocol Specifications
  - draft-ietf-ipfix-protocol-03

© 2004 by Carnegie Mellon University

IETF Packet Sampling (PSAMP) WG

http://www.ietf.org/html.charters/psamp-charter.html

- Binary, extensible information model for specifying
  - Selection operations (sampling and filtering) on a packet stream, and
  - Packets yielded by the selection operation
- Designates a mandatory protocol (IPFIX) to use in the transport of these packets

© 2004 by Carnegie Mellon University

Relationship between IPFIX and PSAMP

- PSAMP extends the IPFIX data model
  - A PSAMP data record is an special instance of an IPFIX flow record with different semantics
    - i.e., a flow record with only a single packet
  - Augments the IPFIX data model to support *Selection Process*
- PSAMP reuses the IPFIX transport protocol

© 2004 by Carnegie Mellon University

PSAMP Selection

- Sampling
  - "Provisioning of information about a specific characteristic of the parent population at a lower cost than a full census would demand"
- Filtering
  - Deterministic selection of packets based on the
    - packet content
    - treatment of the packet at the observation point, or
    - functions operating on the selection state.
- Possible to create schemes combing of both sampling and filtering selections

© 2004 by Carnegie Mellon University

PSAMP Sampling

- Systematic Sampling (deterministic function)
  - Count-based (spatial packet position; e.g., packet count)
  - Time-based (temporal packet position; e.g., arrival time)
- Random Sampling
  - n-out-of-N
  - Probabilistic
    - Uniform Probabilistic (same probability for each packet)
    - Non-Uniform Probabilistic (probability depends on input)
    - Flow State Probabilistic
  - Sampling probability depends on flow state

© 2004 by Carnegie Mellon University

PSAMP Filtering

- Match/Mask
  - Apply bit mask to the header or the first N-bytes
- Hashing
  - Apply a hash function to the header or first N-byte
- Packet Features
  - Properties of the packet header
- Router-state selection
  - Properties of the route or packet treatment

© 2004 by Carnegie Mellon University

PSAMP I-Ds

- A Framework for Passive Packet Measurement
  - draft-ietf-psamp-framework-05
- Sampling and Filtering Techniques for IP Packet Selection
  - draft-ietf-psamp-sample-tech-04
- Packet Sampling (PSAMP) Protocol Specifications
  - draft-ietf-psamp-protocol-01
- Information Model for Packet Sampling Exports
  - draft-ietf-psamp-info-01

© 2004 by Carnegie Mellon University

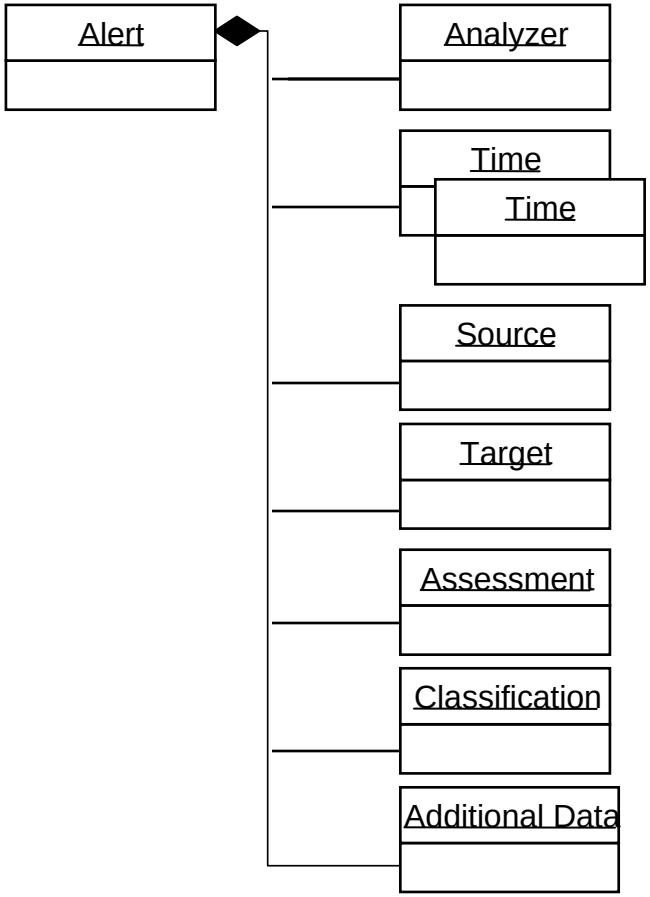
IETF Intrusion Detection WG (IDWG)

http://www.ietf.org/html.charters/idwg-charter.html

- XML information model for network and host-based Intrusion Detection System alerts
  - Intrusion Detection Message Exchange Format (IDMEF)
- Defines a protocol to exchange these alerts
  - Intrusion Detection Exchange Protocol (IDXP)
  - BEEP-based profile to exchange IDMEF

© 2004 by Carnegie Mellon University

IDMEF Data Model



- Sensor properties
- Timestamps
- Source/Target characteristics
    - IP address, ports
- Impact assessment
- Event classification
- Extension mechanism

© 2004 by Carnegie Mellon University

IDWG I-Ds

- Intrusion Detection Message Exchange Requirements
  - draft-ietf-idwg-requirements-10
- The Intrusion Detection Message Exchange Format
  - draft-ietf-idwg-idmef-mml-12
- The Intrusion Detection Exchange Protocol (IDXP)
  - draft-ietf-idwg-beep-idxp-07
- The TUNNEL Profile
  - Rfc3620

© 2004 by Carnegie Mellon University

IETF Incident Handling WG (INCH)

http://www.ietf.org/html.charters/inch-charter.html

- XML information model for exchanging "incident data" among CSIRTs
  - Incident Object Description Exchange Format (IODEF)
- No exchange protocol specified

© 2004 by Carnegie Mellon University

INCH IODEF Data Model

- Extensible framework to exchange information between CSIRTs
  - Workflow
    - incident identifiers, conveying expectations, data usage restrictions
  - Incident description and conclusions
    - Source/Destination information
    - Contact information
    - References to vulnerabilities, advisories, and artifacts
    - Classification and impact assessments
- Extensions
  - RID: DoS traceback for ISPs
  - (possible) Anti-Spam lists

© 2004 by Carnegie Mellon University

INCH I-Ds

- Requirements for Format for INcident Report Exchange (FINE)
  - draft-ietf-inch-requirements-03
- The Incident Data Exchange Format Data Model
  - draft-ietf-inch-iodef-02
- The Incident Object Description Exchange Format (IODEF) Implementation Guide
  - draft-ietf-inch-implement-00
- Real-Time Inter-Network Defense
  - draft-ietf-inch-rid-00

© 2004 by Carnegie Mellon University

IETF Cross-Registry Information Service Protocol (CRISP) WG

http://www.ietf.org/html.charters/crisp-charter.html

- XML, extensible information model for global registry information
  - i.e., Whois with structure
- Designates a mandatory protocol (BEEP) for the query/response exchange

© 2004 by Carnegie Mellon University

Vulnerability Information (*de facto*)

- Mitre CVE
  - http://cve.mitre.org/
- Mitre OVAL
  - http://oval.mitre.org/
- NIST iCAT
  - http://icat.nist.gov/icat.cfm

© 2004 by Carnegie Mellon University

Vulnerability (Report) Formats

- Common Advisory Interchange Format (CAIF)
  - RUS-CERT
  - http://cert.uni-stuttgart.de/projects/caif/
- Advisory and Notification Markup Language (ANML)
  - OpenSec
  - http://www.opensec.org/anml/
- Application Vulnerability Description Language (AVDL)
  - OASIS
  - http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=avdl

© 2004 by Carnegie Mellon University

Relevance of the Formats to Flows

- IPFIX
  - Storage and transport format for flows
- PSAMP
  - Describe acquisition process of the flows
- IDMEF
  - Describe events created from flows
- IODEF (with/without extensions)
  - Describe flow summaries, baselines, etc.

© 2004 by Carnegie Mellon University

Adoption

- Packets and Flow Formats
  - IPFIX: implementations exist (e.g., Argus)
  - PSAMP: work in progress
- Alerts and Events Formats
  - IDMEF: adoption only in Snort, Prelude, Arcsight
  - IODEF: adoption by 5-15 CSIRTs in Europe, Asia, and the US
- Context Formats
  - Vulnerability formats: work in progress, some used in closed communities
  - CRISP: work in progress

© 2004 by Carnegie Mellon University