Bidirectional Flow Measurement, IPFIX, and Security Analysis

Elisa Boschi, Hitachi Europe SAS
Brian Trammell, CERT/NetSA

Tuesday, October 10, 2006

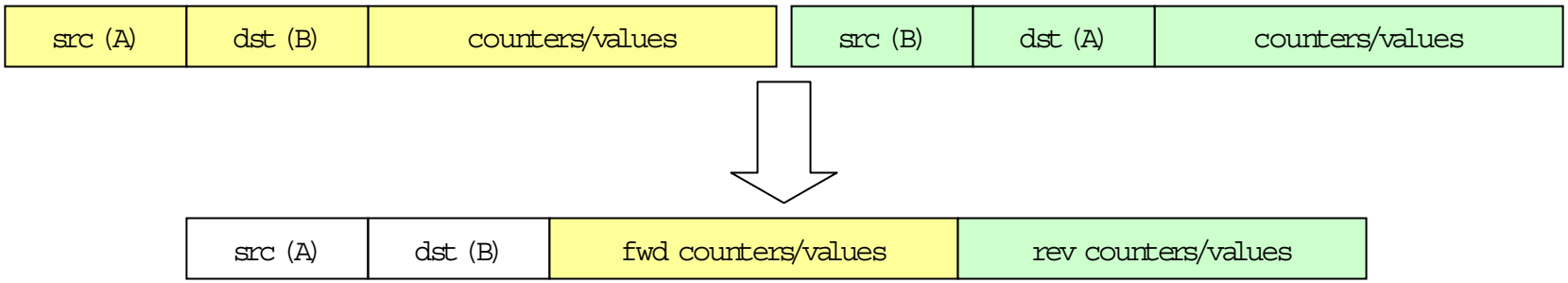
FloCon 2006 - Vancouver, WA, US

Introduction and Motivation

- Bidirectional flow (biflow) information useful for a variety of security-relevant use cases.
  - Aggregate response counting
  - Passive identification of open hosts/services
- Biflow matching easier closer to measurement interface.
  - Biflow matching is $O(n^2)$, so making n smaller is worth the effort.
- IPFIX is the emerging standard for flow export.
- Therefore, we need an efficient way to represent biflow data using IPFIX.
  - Issue identified at FloCon 2005 in Pittsburgh.

Biflows Defined

- Loosely: Association of information about both directions of a bidirectional communication.



- Practically speaking: Choose one side of the communication to be the “source”, the other the “destination”, and maintain two sets of counters and other value fields (“forward” and “reverse”).

Biflow Direction Assignment

- By Initiator: “source” is source of packet initiating the communication (active open for TCP).
  - Approximated by assuming first packet seen is the first packet sent.
  - Can be validated through use of TCP flags, application protocol analysis (e.g. UDP DNS answer count), etc.
- By Interface/Address: “source” and “destination” assigned via membership in address set or side of a given interface.
  - Useful when defining a perimeter: assign “source”

Current Work in Biflows

- QoSient Argus explicitly collects and exports biflows.
- Any technique that needs to model the network state of end systems (e.g. IDS such as Snort, Bro) must use an internal biflow data model.
- A few measurement research projects implicitly use a biflow model.
  - Most efforts are still trace-based, so no flow export or storage.
- Most flow collection is uniflow-based.
  - Legacy of NetFlow and efficiency concerns on the router.

Efficiency in Biflow Matching

- Biflow matching is $O(n^2)$ on number of concurrently matchable flows.
- *n* particularly large on asymmetrically routed networks.
  - All traffic must be centralized and matched.
  - Overhead of biflow matching may outweigh benefits.
- *n* manageable at symmetric routing points.
  - A “symmetric” routing point here need not be on the same line card or in the same router, if it’s in the same room.
- *n* trivial on smaller networks measured at layer 2.

IPFIX

- Unidirectional, multi-transport, binary flow export protocol.
  - SCTP (+DTLS): preferred
  - TCP (+TLS): supported
  - UDP (+DTLS, over dedicated link): de facto
- Uses templates to define record formats in terms of information elements.
- Defines a message format and an information model from which information elements can be chosen.
- Information model is extensible.
  - IANA registry
  - Private enterprise information elements

Applying IPFIX to Security Analysis

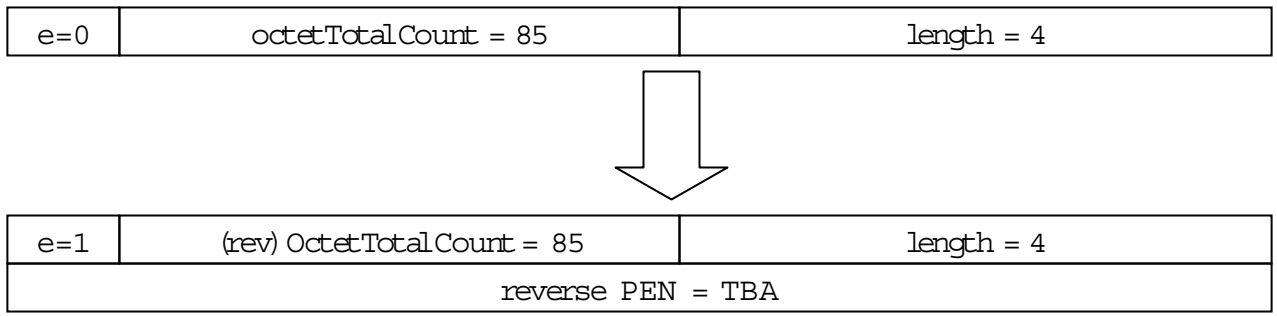
- Emerging ecosystem of interoperable IPFIX meters, collectors, and intermediate processes improves information sharing.
- Templated data format improves flexibility.
  - Interoperability on common information elements.
  - Ability to innovate with private information elements.
  - Minimal implementation effort to collect new data for specific experiments or operational measurement concerns.

Extending IPFIX for Biflow Export

- IPFIX protocol does not natively support single-record biflow export.
  - Can export uniflow halves adjacently within a record stream, but this is inefficient and semantically ill-defined.
  - No support for “reverse” direction values in information model.
- Assign direction by initiator, where possible.
  - Semantics and operational characteristics of address direction assignment are still an open issue.
- Add “reverse” values to the information model.
  - Direct allocation would add significant management overhead to the IANA information element registry.
  - So, define a new dimension.

Adding a Reverse Dimension

- Allocate an IANA private enterprise number (PEN) to the draft.



- Information elements within this PEN number space correspond to the IANA number space, except that they apply to the reverse direction of a biflow.

History

- FloCon 2005 (September 2005) identified issues with IPFIX:
  - Lack of direct support for representing biflows.
  - Use of transport protocol (SCTP) with inadequate support on commodity operating systems.
- Authors created a draft to address the biflow representation issue.
  - IETF 64 (November 2005): initial revision of biflow draft.
  - IETF 65 (March 2006): charter changed to address biflow issue.
  - IETF 66 (July 2006): biflow draft accepted as working group item.
- Identification of issue at FloCon cited as motivation for addressing it within the IETF.
  - Excellent example of community interaction in standardization

Questions and Discussion