Flow-based Accounting: Applications and Standardisation

SCAMPI Workshop
May 3, 2004

Simon Leinen, SWITCH  <simon@switch.ch>
Flow-based Accounting - Basic Idea

- Classify packets into flows (equivalence classes)
- Update per-flow account for each packet
- Export this accounting data when the flow ends
  - Failing that, periodically
The Flow Concept

Classical Flow = 5-Tuple

Defined by ("Flow key") the combination of:
  □ source IP address
  □ destination IP address
  □ protocol (TCP, UDP, ICMP, GRE etc.)
  □ source port
  □ destination port
    ○ for protocols that have port numbers (TCP, UDP, SCTP)

Optionally:
  □ TOS (type of service)/DSCP (Diffserv code point)
  □ input interface index
Metrics in per-Flow Accounting Data

- # packets in flow
- # bytes in flow
- time first packet seen ("flow start")
- time last packet seen ("flow end")

Route information for destination/source address
- destination interface
- next hop
- prefix length
- neighbour and/or origin AS

Various
- inclusive-OR of TCP flags during flow
- ...
Implementations

- **NeTraMet**
  - IETF standard

- **LFAP**
  - Cabletron (Riverstone/Enterasys)

- **NetFlow**
  - Cisco NetFlow switching (router acceleration)
  - Cisco NetFlow accounting
  - Sampled NetFlow
  - Aggregated NetFlow
  - Cisco NDE (TCAM-based)
  - Juniper "cflow"
  - Other NetFlow implementations
Properties

- Good tradeoff between
  - High level of accounting detail
  - Modest resource consumption and simple implementation

- "Classic" NetFlow vulnerable to DoS
  - Single 24-byte packet can generate 48 accounting bytes!

- Can be fixed by
  - Sampling
  - (Exporter-based) Aggregation
Common Applications

- Coarse-grained traffic analysis
  - for traffic engineering
  - for capacity planning
  - for interconnection (peering) decisions
- Detection of anomalies
  - security violations
  - faults
- Usage-based charging
- Research
  - on large-scale network/application behaviour
Example NetFlow Usage: SWITCH (I)

(Swiss Education & Research Network)
All external border (peering) routers send NetFlow
- non-sampled
- non-aggregated

-> ~20000 flows/s during normal office hours
- ~800 kb/s accounting stream
- Representing 1-2 GB/s user traffic
- 5-10% packets not counted during peak hours
  - due to contention for hardware NetFlow hash table
  - mainly concerns small packets/flows

- Flow DoS/scans can cause high accounting loss
- But we can still switch 150 Mpps...
Example NetFlow Usage: SWITCH (II)

Three consumers of accounting packet streams:

- **Fluxoscope**: system for per-site/AS statistics
  - used for billing
  - heuristics for "application" breakdown

- **DDoSVax**: ETHZ research project
  - records all flows for offline analysis (months, TBs)
  - study spread of DDoS etc.

- **SWITCH security group**
  - records full flows for forensics (days-weeks, 100s GBs)
  - locate specific infected hosts
Fluxoscope

Coarse-grain aggregation for billing/planning

Traffic with AS 3549 per Application Protocol

Typical output
DDoSVax

Flow recording and offline analysis of anomalies

Sobig.F E-Mail Worm Propagation observed in the SWITCH Network

#total bytes/hour sent to port SMTP (25/TCP)

Date and Time (CEST = UTC+2)

typical output
SWITCHcert

Flow recording for forensics

interactive "drill-down" interface to data
IETF Standardisation I: RTFM

Real-Time Flow Measurement, RFC2720-2724
- flexible flow definitions
- SNMP-based access
- not widely implemented
IETF Standardisation II: IPFIX

IP Flow Information eXport Working Group

- Established in September 2001

Workplan

- Specify requirements (done)
- Evaluate candidates (done)
  - CRANE
  - DIAMETER
  - LFAP
  - NetFlow v9  <<< selected
  - Streaming IPDR
- Refine protocol, data model etc. (ongoing)
IPFIX/NetFlow v9 Overview

- Departure from previous NetFlow versions
  - v1, v5, v6, v7: variants of 5-tuple flows
  - v8: fixed set of aggregated flows
- NetFlow v9/IPFIX is template-based
  - Template FlowSets contain descriptions of Data FlowSets
  - Data FlowSets contain actual accounting data
- Can accommodate variety of flow generalisations
- Cisco has implementations for IPv6, MPLS
  - Support available in some NetFlow data processors
  - Not widely used (or even tested) yet
Outlook

"Classical" flow accounting impossible at high speed

- You HAVE to do sampling and/or aggregation
  - The more you sample, the less you benefit from flow aggregation
  - The more you aggregate, the more information you lose

Sampling becomes an interesting alternative:

- Much simpler mechanism than NetFlow
  - lower cost
  - fewer bugs!
- Arbitrarily scalable
- Hard to cheat, given good random sampling
Prediction

For high-speed routers, the mechanism of choice will be sampling (PSAMP) rather than flow export.

NetFlow/IPFIX will still spread at the edges
- CPU-based routers
- Host stacks
- Mediation systems ("IP Detail Record")?
Thank you!

Questions?