Hardware-Accelerated Flexible Flow Measurement

Pavel Čeleda
celeda@liberouter.org

Martin Žádník
zadnik@liberouter.org

Lukáš Solanka
solanka@liberouter.org
Part I

Introduction and Related Work
Introduction

Motivation

- Networks are difficult to understand without monitoring.
- Networks are complex and prone to failures and attacks.
- Monitoring of multi-gigabit networks is a challenging problem.

What We Need?

- Real-time traffic monitoring, QoS measurement.
- Anomaly detection, security analysis and forensics.
- Capacity and topology planning, ...
Standard Flow Monitoring Solutions

**Routers** – CISCO, Juniper, Enterasys, . . .

- Busy with routing, flow monitoring addon feature.
- Flow monitoring is not implemented in all models.
- Fixed placement, possible target of attacks.
- Often mandatory sampling, no advanced features.

**Flow Probes** – nProbe, fprobe, softflowd, . . .

- Based on commodity HW – PC and standard NICs.
- Solution when flow monitoring required but not available.
- Limited performance (PCAP, PCI-X) and stability problems (packet drops, time stamps issues, . . .).
- Requires extra system tuning and system/tools hacks.
Hardware Acceleration

- PC is flexible but not fast enough to process gigabit links.
- Hardware is fast but not easy to use.

⇒ Combination of PC and programmable hardware FPGA (Field-Programmable Gate Array).
COMBO6X and COMBOv2 Card Family

- Time-critical parts of monitoring are processed in FPGA.
- New cards designed for 10+ Gb/s speeds (up to 40-100 Gb/s).
FlowMon Probe - Short Overview

Goals

- Usage of hardware acceleration for IP flow measurement.
- Implementation of advanced methods for network monitoring.

Features

- Mobile network appliance, no fixed network position.
- Independent of network infrastructure used.
- Based on Linux → "unlimited" addon smart extensions.
- Observes whole network traffic under all conditions.
- Standard compliant - NetFlow v5/9 and IPFIX.
- Secure configuration via NETCONF web interface or SSH.
FlowMon Probe - Architecture

FlowMon probe block schema.
FlowMon Probe - Summary

- Stable firmware and SW for COMBO6X HW.
- Scientific projects – flow monitoring, anomalies detection.
- Recognized by GÉANT2 as part of security toolset + NfSen.

Detailed network view with NetFlow data.
Part II

Flexible Flow Measurement
Motivation – I

New Measurement Requirements

- Application identification – statistical fingerprinting, . . .
- IDS – pushed number of bytes, number of zero window probes, sample of payload, . . .
- First N packets statistics, averages, variances, histograms, . . .

Current Flow Measurement

- Requirements not met with traditional 5-tuple NetFlow.
- IPFIX – defined and vendor-specific Information Elements.
- New vendor/user-specific Information Elements are inevitable.
Current Practice of User-Specific Measurement

- Packet sniffing with tcpdump, wireshark, ...
- Offline aggregation by arbitrary scripts.

Diagram:
- Splitter
- Tapped traffic
- tcpdump
- Offline
- Feature extraction
- Visualization
- Administrator
Challenge of Flow Monitoring Infrastructure

- Measurement and collection of ad-hoc Information Elements has not been fully addressed.
- The goal should be to specify new (non-existing) Information Element and setup exporter and collector to report it automatically.
- **Dynamic** and flexible flow measurement ➔ Tell me what you want and I will deliver.
- Steps to define new Information Elements (IE):
  1. Select packet header fields and IE to work with.
  2. Specify how to aggregate these fields into a new IE.
  3. Define triggers.

**Hardware-Accelerated Flexible Flow Measurement**

<table>
<thead>
<tr>
<th>ETH</th>
<th>IP</th>
<th>TCP/UDP</th>
<th>Application</th>
</tr>
</thead>
</table>

Aggregation ➔ Flow record

Čeleta, Žádník, Solanka
User-specific Information Elements. On-the-fly configuration of metering process.
Dynamic Flow Measurement

- Standardized definition of packet structure – NetPDL (*Network Protocol Description Language*).
- Standardized definition for flow record – IPFIX.
- Standardized definition of operation – simple C function.
Flexibility and performance of metering process.

- Possible solution: Utilization of network card with FPGA.
- Flexible, yet wired functionality.
- Line rate processing.

Collector for dynamic flow measurement.

- Sufficient performance.
- Allows not only to store flow records but also understand and visualize information encoded.
System Architecture

Bitfile synthesis
Bitfile storage

Upload XML configuration and download bitfile

Web interface

FPGA bitfile, XML configuration of metering process and exporting process.

NetCONF

Administrator

XML conf.
Visualization

Collector

IPFIX

Splitter

Tapped traffic

FlowMon probe FPGA + host

Uploader XML configuration and download bitfile

Bitfile storage

Web interface

FPGA bitfile, XML configuration of metering process and exporting process.

NetCONF

Administrator

XML conf.
Visualization

Collector

IPFIX
Probe Architecture

Firmware - FPGA

- Packet parsing engine – hardcoded Finite State Machine.
- Indexing – hash and overflow scheme.
- Fast (line-rate) flow record update engine.
- Flow cache – large SSRAM + internal memory in FPGA.

Software

- Aggregates sliced flows (if definition allows).
- Export flows.
Flexible FlowMon

Our Testbed and Deployment Network

- HW testers for line-rate (worst-case) testing.
- NREN (National Research and Education Network) backbones, university campuses and ISP networks.
- Sustained live traffic 4-5 Gb/s, 700 kpkt/s, 30 kflows/s.
- Long-time NetFlow monitoring - probes and collectors.

Performance Expectation

- Measurement of 10 Gbps without packet loss.
- Timestamp (< 60 ns) able to distinguish consequent packets.
- Cover IPFIX and allow for user-specific Information Elements.
- Variety of optional sampling methods.
Part III

Future Work and Conclusion
State of Development

- Module for assembling parsing engine – ready.
- Module for assembling flow record update engine – ready.
- NETCONF data path – ready.
- IPFIX exporter (user-defined flow record) – work in progress.
- IPFIX collector (user-defined flow record) – work in progress.

HW and SW Support

- Firmware for COMBO6X + COMBO-2XFP2 - 2x10 Gb/s.
- Linux OS - CentOS 5.