Large-Scale Geolocation for NetFlow

Pavel Čeleda, Petr Velan, Martin Rábek
Rick Hofstede, Aiko Pras

{celeda|velan|xrabek1}@ics.muni.cz, {r.j.hofstede|a.pras}@utwente.nl

UNIVERSITY OF TWENTE.

IFIP/IEEE IM 2013, 27-31 May 2013, Ghent, Belgium

Part I

Introduction

Motivation and R&D Goals – I

SURFmap - a Network Monitoring Tool Based on the Google Maps API.

Motivation and R&D Goals – II

How flow-based geolocation can be performed in a large-scale?

- exporter-based approach,
- collector-based approach.

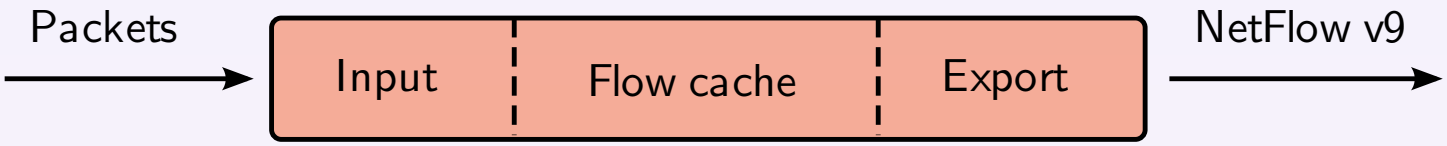
How can we benefit from geolocation data in flow records?

- traffic engineering,
- traffic profiling,
- anomaly detection.

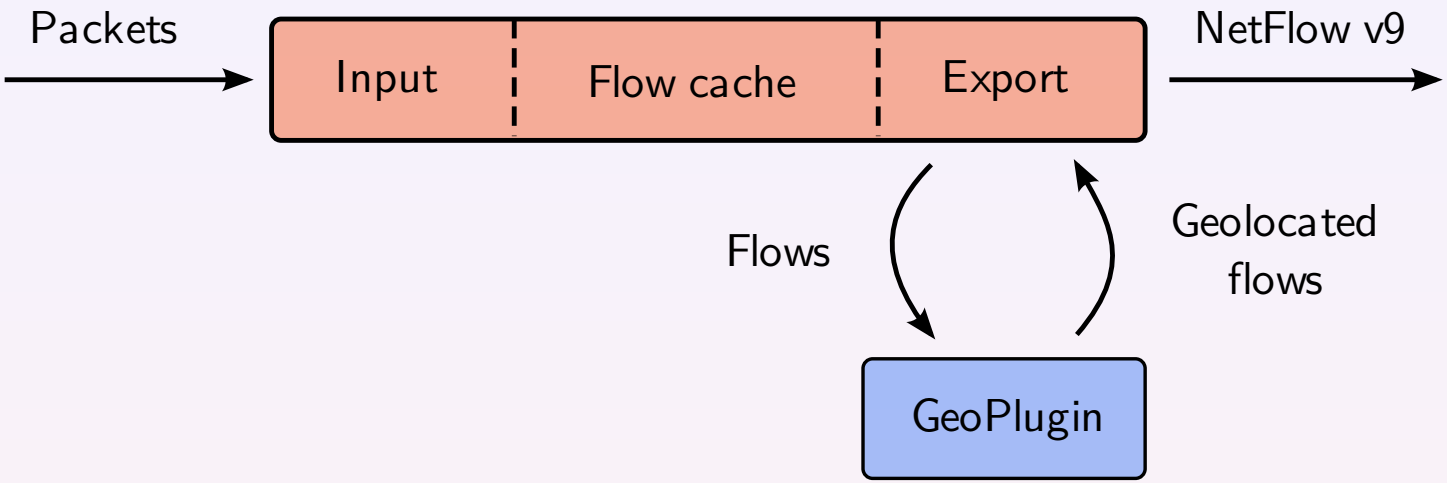
Part II

Architecture

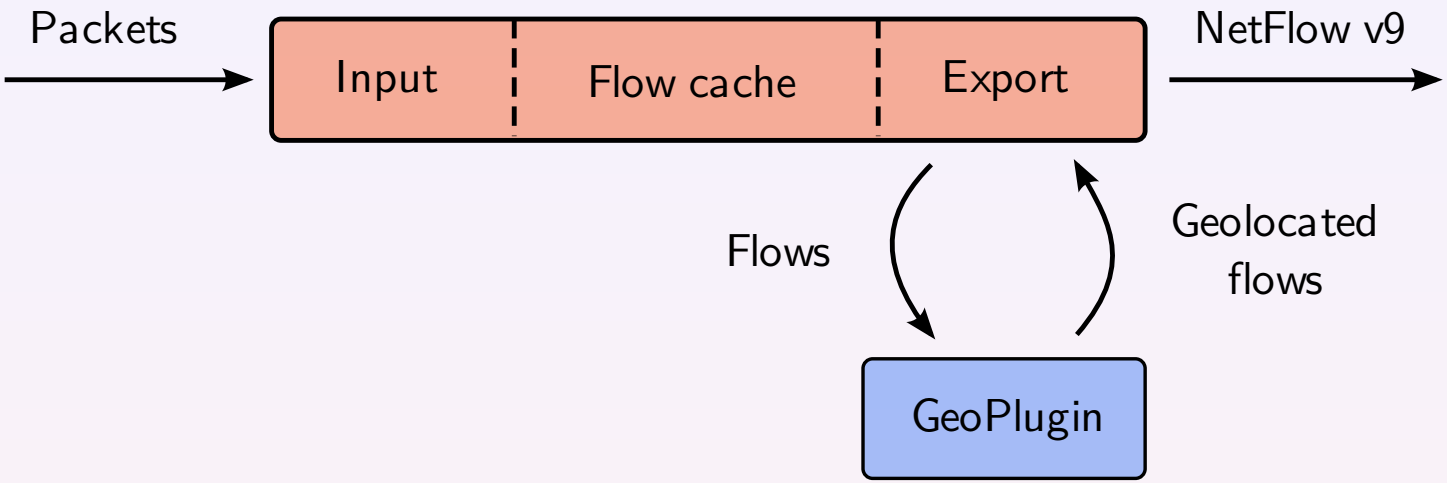
Exporter-Based Geolocation



Exporter-Based Geolocation



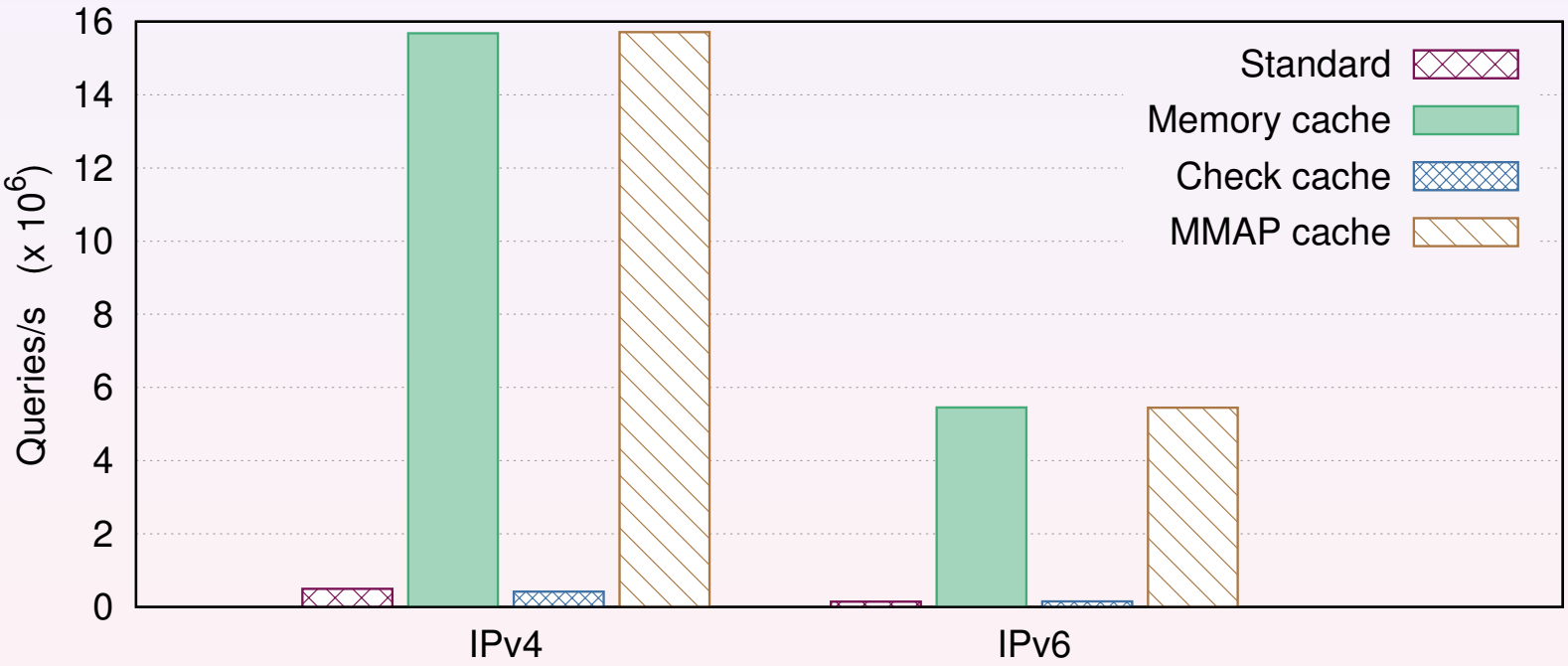
Exporter-Based Geolocation



- exporter filter plugin for IP address geolocation,
- NetFlow v9 template mapping – GEO data to AS fields `SRC_AS=*SRC_GEO, DST_AS=*DST_GEO`,
- AS mapping $\rightarrow$ transparent to any flow collector.

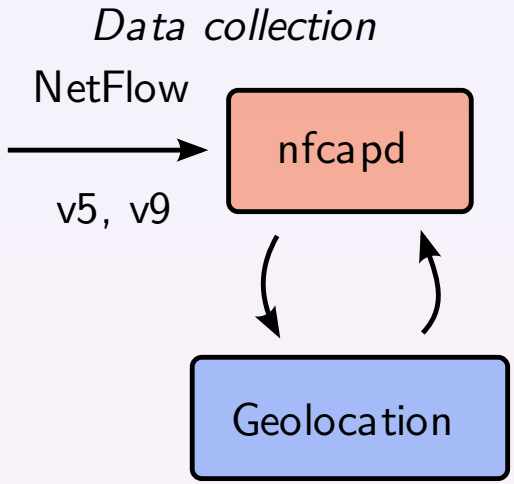
MaxMind GeoLite Country Database

- MaxMind GeoLite – free off-line country database,
- C-API for IPv4/IPv6 geolocation.



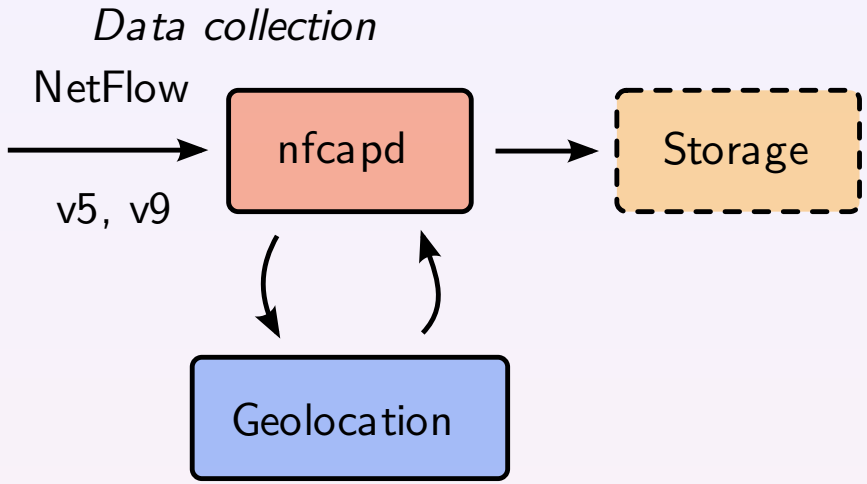
: IPv4/IPv6 geolocation database performance.

Collector-Based Geolocation



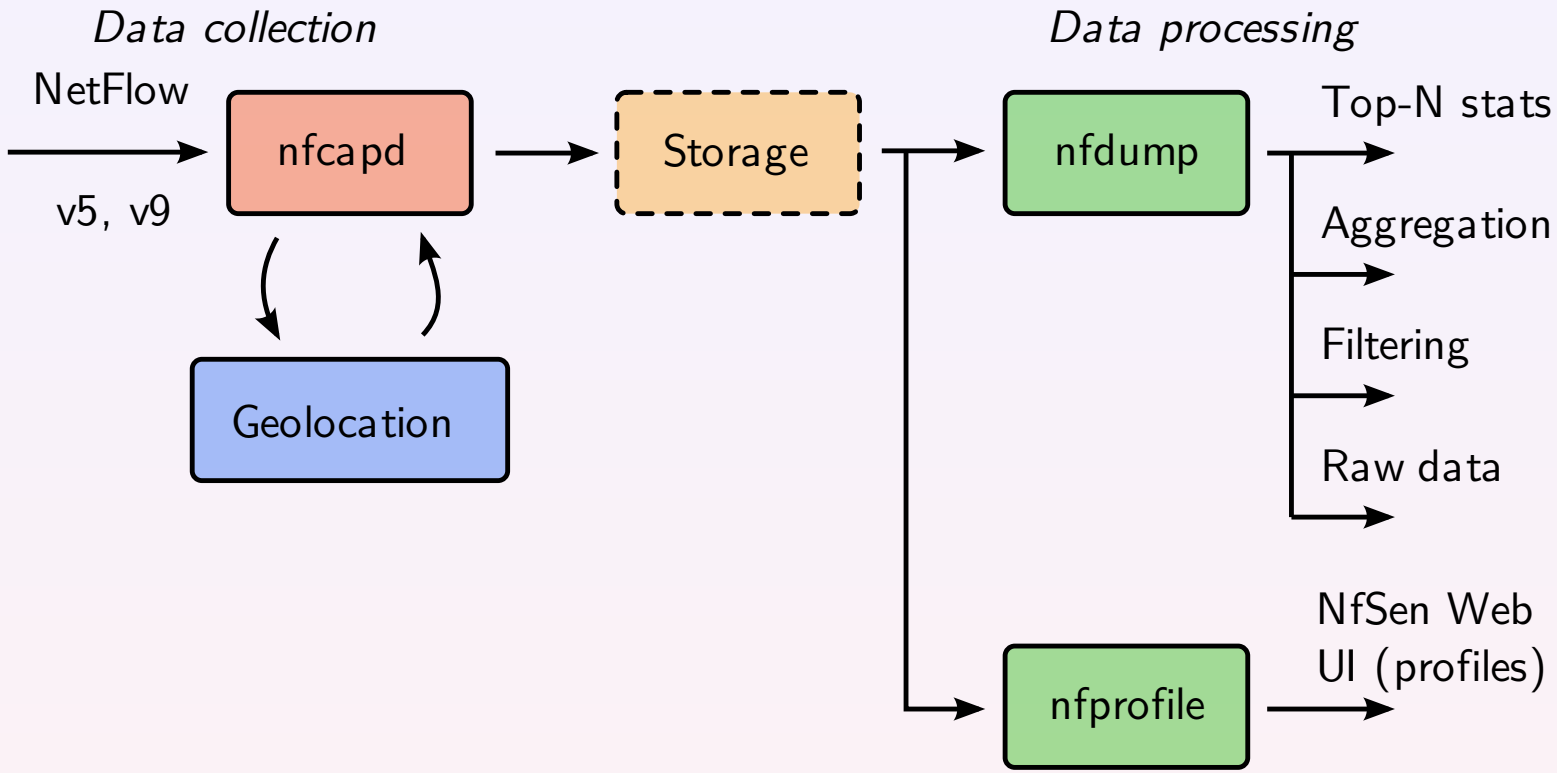
- patch for NFDUMP and NfSen toolset,
- native geolocation support for any NetFlow v5/v9, IPFIX data.

Collector-Based Geolocation



- patch for NFDUMP and NfSen toolset,
- native geolocation support for any NetFlow v5/v9, IPFIX data.

Collector-Based Geolocation



- patch for NFDUMP and NfSen toolset,
- native geolocation support for any NetFlow v5/v9, IPFIX data.

NFDUMP Database Extension #15 – Country Code

```
Flow Record:
  Flags         =                0x06 Unsampled
  size          =                  80
  first         =          1348387461 [2012-09-23 10:04:21]
  last          =          1348387462 [2012-09-23 10:04:22]
  msec_first    =                 890
  msec_last     =                 100
  src addr      =       23.63.79.144
  dst addr      =    147.251.170.165
  src port      =                  80
  dst port      =               57046
  tcp flags     =                0x1a .AP.S.
  proto         =                   6
  (in)packets   =                   4
  (in)bytes     =                 936
  input         =                   5
  src as        =               20940
  dst as        =                2852
  in src mac    = 00:0e:38:5e:30:c0
  out dst mac   = 00:1e:be:8b:26:c0
  src ctry      =                 840 ... ISO 3166-1 country code - US
  dst ctry      =                 203 ... ISO 3166-1 country code - CZ
```

NFDUMP Flow Listing

a) numeric code – %scc %dcc

Proto	Src IP Addr:Port		Dst IP Addr:Port	Src Ctry	Dst Ctry
ICMP	194.228.29.173:0	->	147.251.48.205:3.13	203	203
TCP	147.251.210.106:51885	->	69.171.227.59:443	203	840
UDP	151.40.40.243:15833	->	147.251.79.246:49159	380	203
TCP	157.55.235.165:40040	->	147.251.215.10:49464	840	203
UDP	147.251.170.77:59408	->	89.79.20.120:18973	203	616

b) alpha-2 code – %sccan %dccan

Proto	Src IP Addr:Port		Dst IP Addr:Port	Src Ctry	Dst Ctry
ICMP	194.228.29.173:0	->	147.251.48.205:3.13	CZ	CZ
TCP	147.251.210.106:51885	->	69.171.227.59:443	CZ	US
UDP	151.40.40.243:15833	->	147.251.79.246:49159	IT	CZ
TCP	157.55.235.165:40040	->	147.251.215.10:49464	US	CZ
UDP	147.251.170.77:59408	->	89.79.20.120:18973	CZ	PL

Usage example

```
nfdump -M /data/nfsen/profiles-data/live/p3000:p3001 \
       -r 2012/09/23/nfcapd.201209231005 \
       -o 'fmt:%pr %sap -> %dap %sccan %dccan' -m -c 20
```

NFDUMP Geofiltering

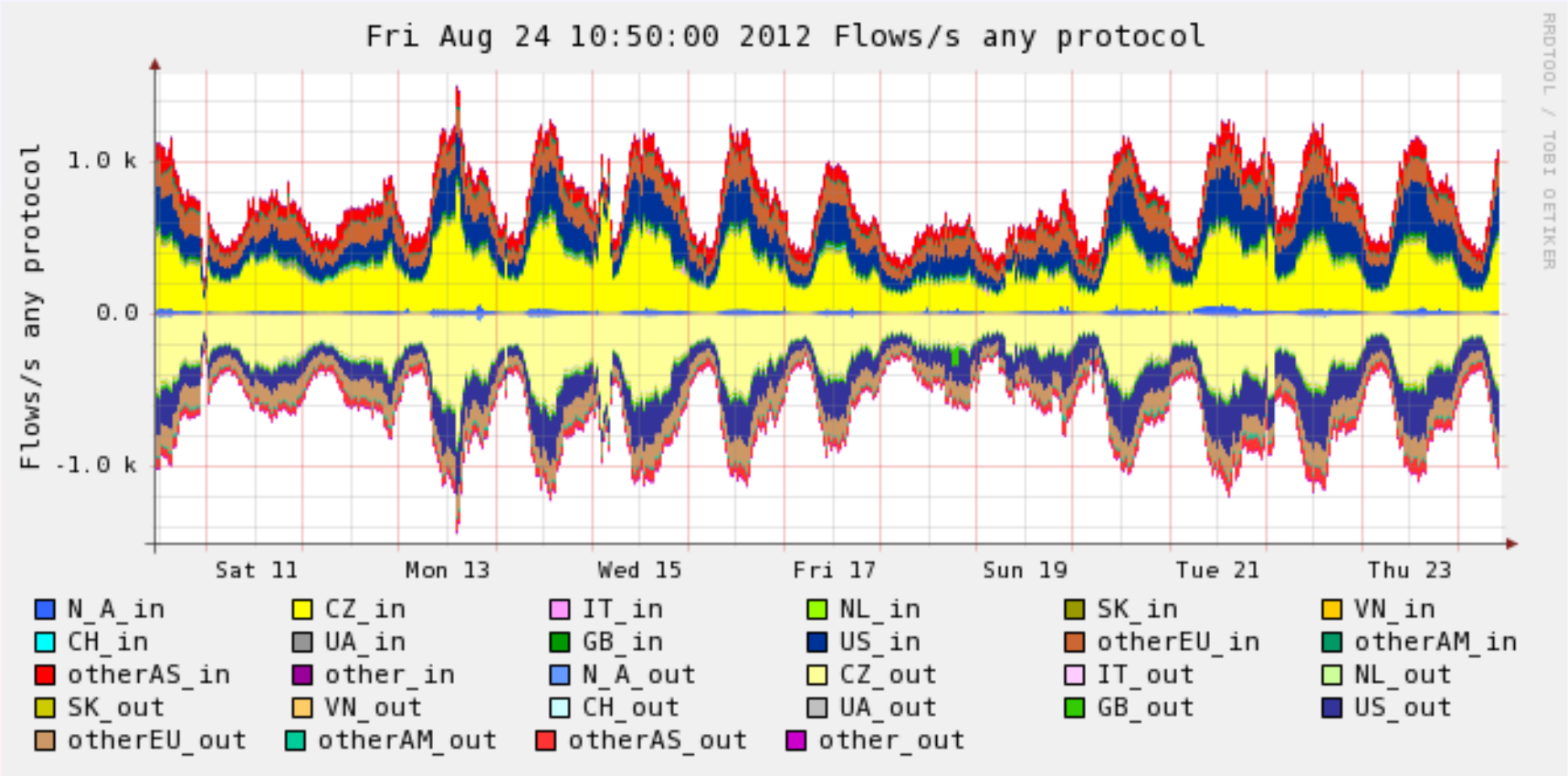
Geofiltering

- country filter syntax is similar to other NFDUMP filters syntax : `ctry [comp] <num>`,
- country can be compared to a list (red-black tree) of country codes, syntax : `ctry in [ <ctrylist> ]`,
- filters are often used for traffic profilling in NfSen.

Usage example

```
nfdump -M /data/nfsen/profiles-data/live/p3000:p3001 \
       -r 2012/09/23/nfcapd.201209232035 -c 5 \
       'src ctry 203 and not dst ctry in [ 203 840 166 ]'
```

NfSen Geoprofiling

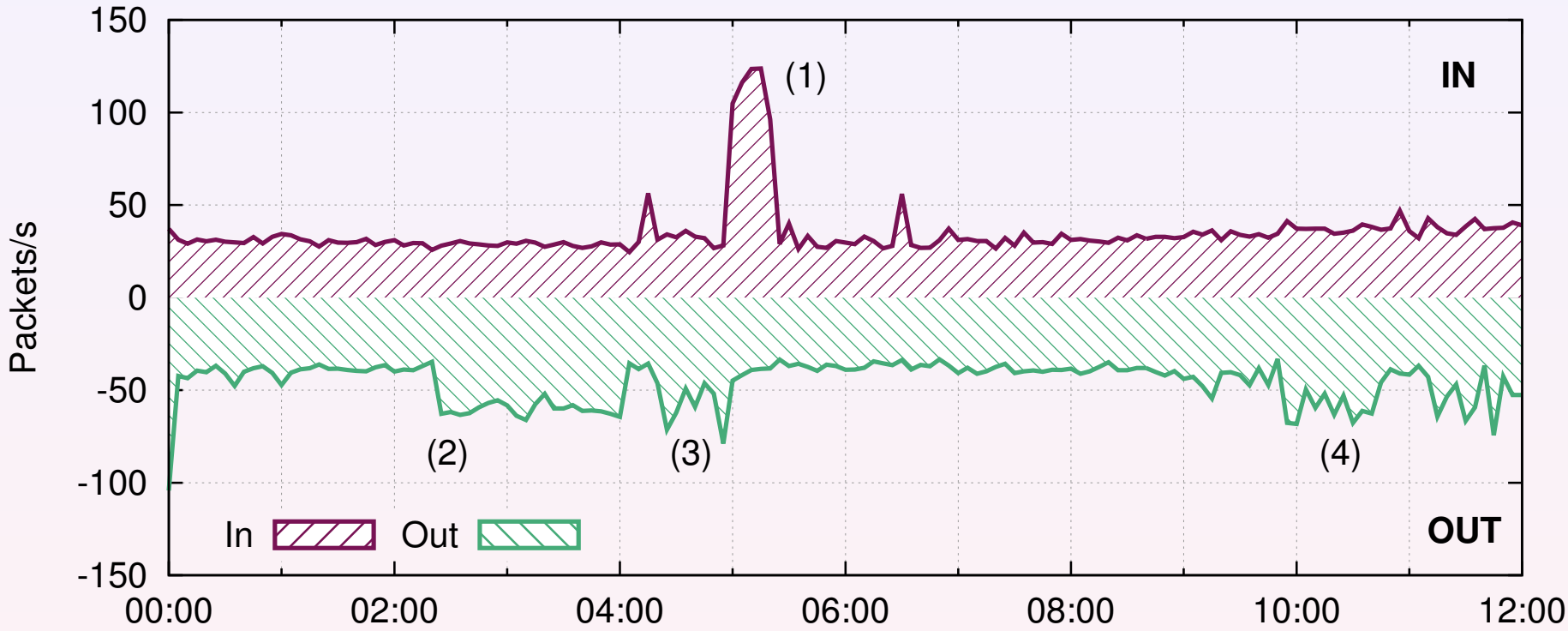


Screenshot of collector-based geolocation prototype.

Part III

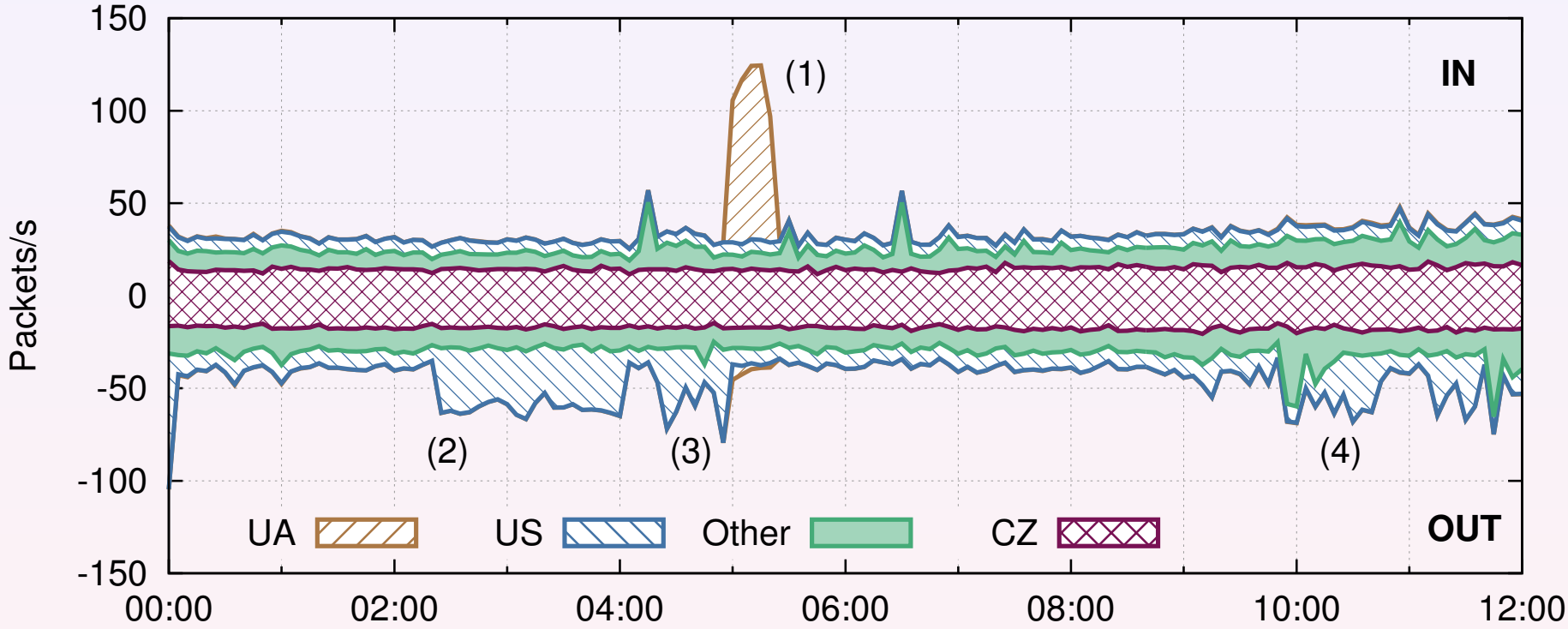
Use Case I – Traffic Profiling

Geolocated and Non-geolocated ICMP Traffic – I



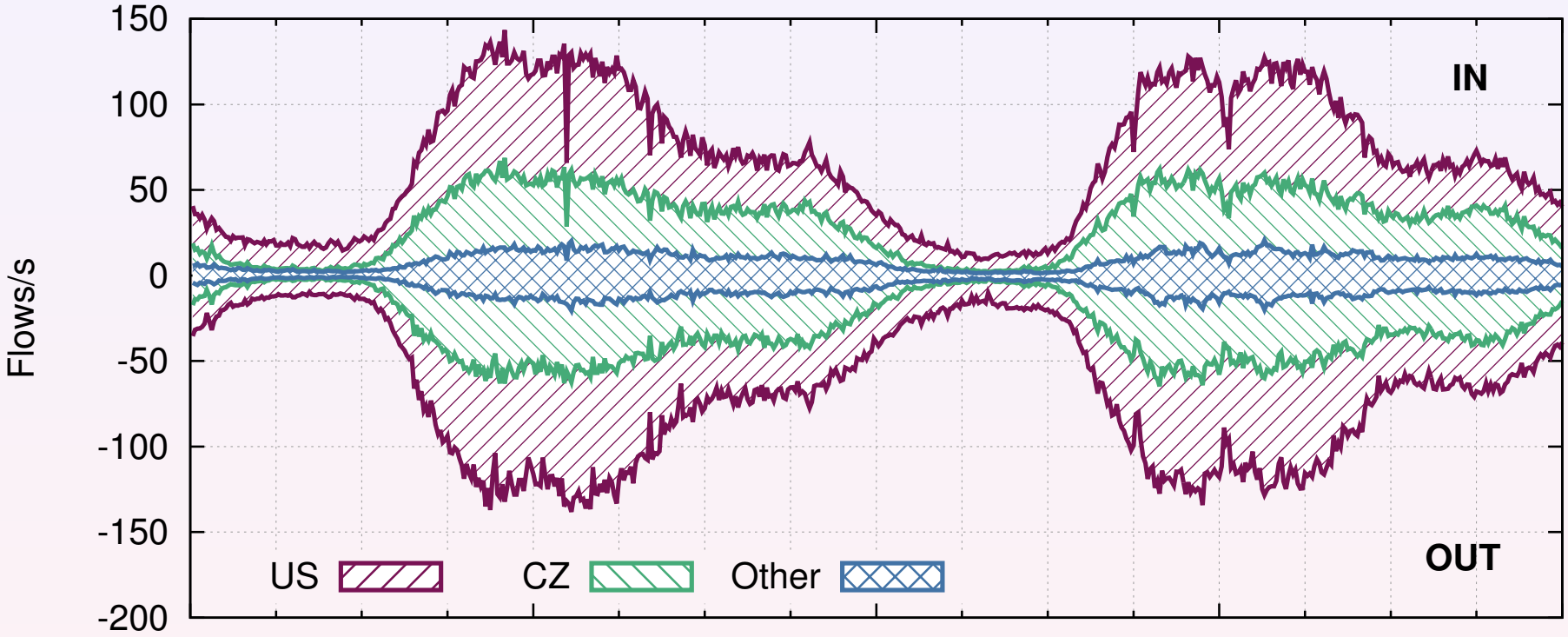
: ICMP traffic.

Geolocated and Non-geolocated ICMP Traffic – II



: Geolocated ICMP traffic.

Distribution of HTTPS Traffic over Countries – I

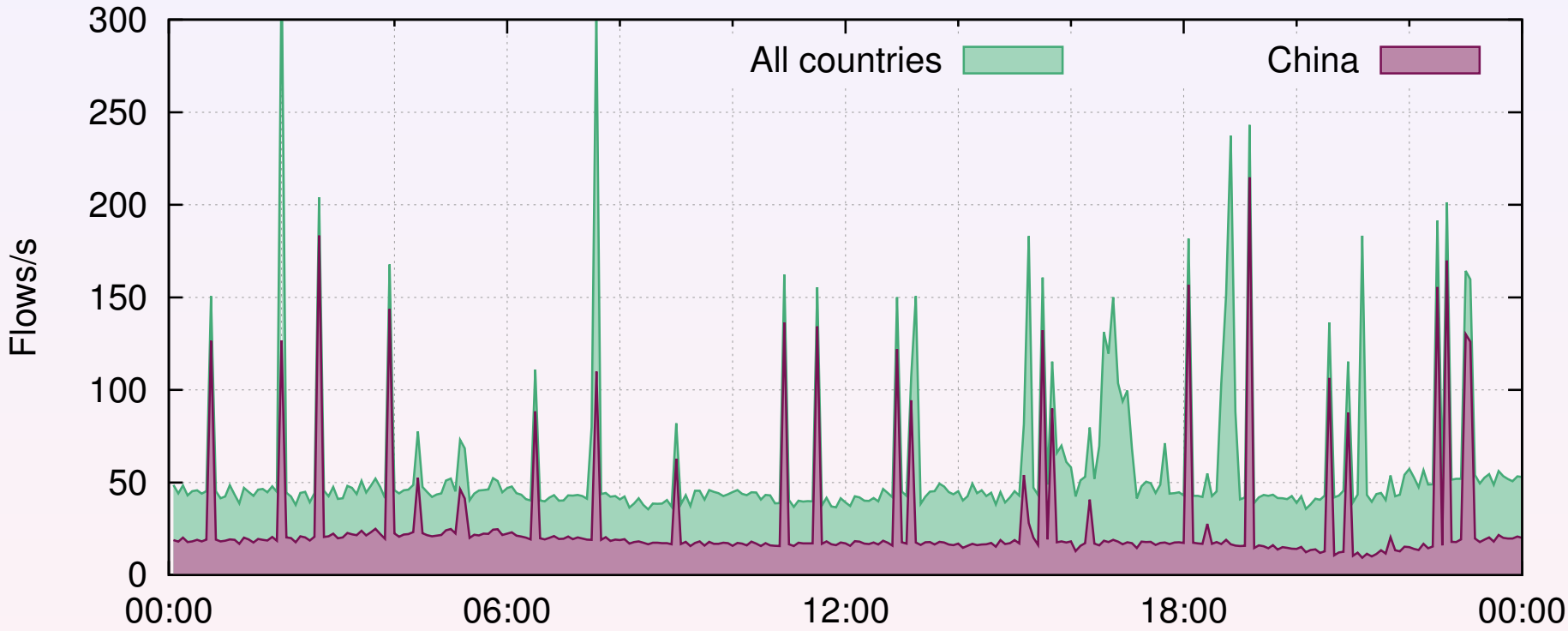


HTTPS flows/s.

Part IV

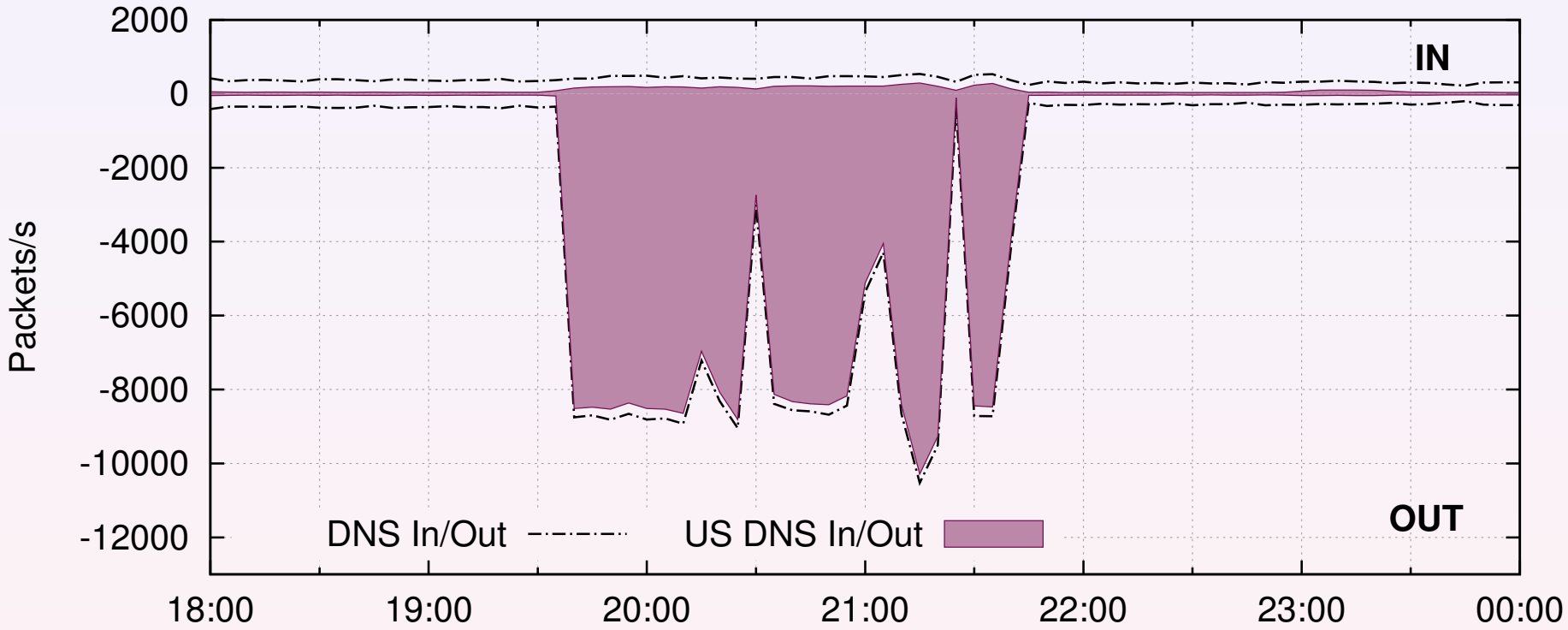
Use Case II – Anomaly Detection

Bad Neighboring Countries



: Incoming TCP SYN-only flows.

UDP DoS Attack



: UDP DoS attack from infected Linux machine.

Part V

Conclusion

Conclusion

Summary

- country-level information in flow data,
- native geolocation support for NfSen/NFDUMP,
- pilot geo-prototype deployment at MU – CESNET link.

Future Work

- IPFIX-compliant prototype for exporter-based geolocation,
- `ipfixcol` – AS and GEO support implementation,
- AS + GEO data for traffic profiling and anomaly detection.

Thank You For Your Attention!

Large-Scale Geolocation for NetFlow

UNIVERSITY OF TWENTE.

P. Čeleda, P. Velan, M. Rábek

{celeda|velan|rabek}@ics.muni.cz

R. Hofstede, A. Pras

{r.j.hofstede|a.pras}@utwente.nl

Geolocation Toolset

http://www.muni.cz/research/publications/1090804