Next Generation Flow Measurement for Application Monitoring

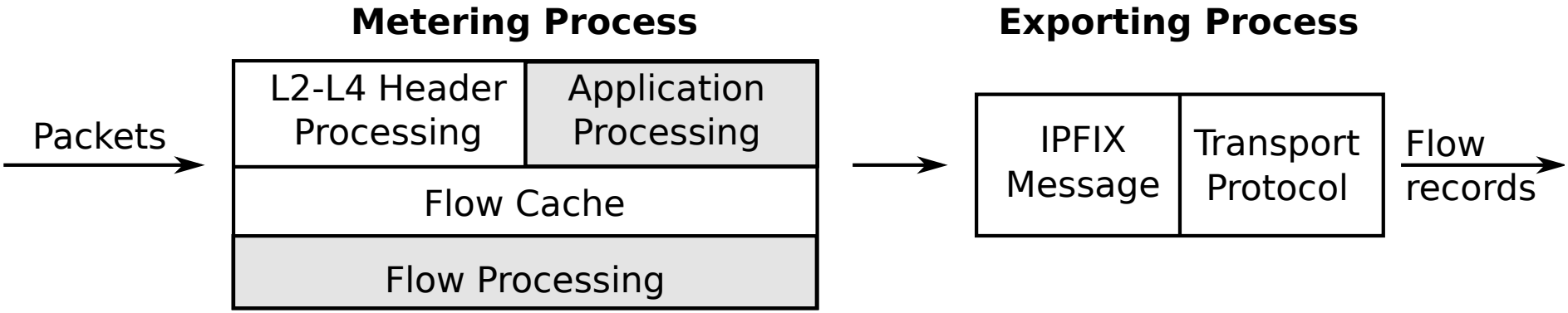
Petr Velan
velan@ics.muni.cz

May 30, 2014
Brno

Application Flow Monitoring

- Passive network monitoring
- IP flow monitoring + application protocol information
- More accurate traffic classification
- Threat detection on application level
  - Phishing
  - Invalid X.509 certificates
  - ...
- Emerging trend in network monitoring
- More work in implementation than research

Application Flow Monitoring



IP flow example

Flow start	Duration	Proto	Src IP Addr:Port		Dst IP Addr:Port	Flags	Packets	Bytes
09:41:21.763	0.101	TCP	172.16.96.48:15094	->	209.85.135.147:80	.AP.SF	4	715
09:41:21.893	0.031	TCP	209.85.135.147:80	->	172.16.96.48:15094	.AP.SF	4	1594

Application flow extension example

HTTP RT	HTTP Host	HTTP Path	HTTP Code	HTTP Type
GET	www.seznam.cz	/favicons/019/194-DBrJCJ.png	-	-
HTTP	-	-	200 OK	image/x-icon

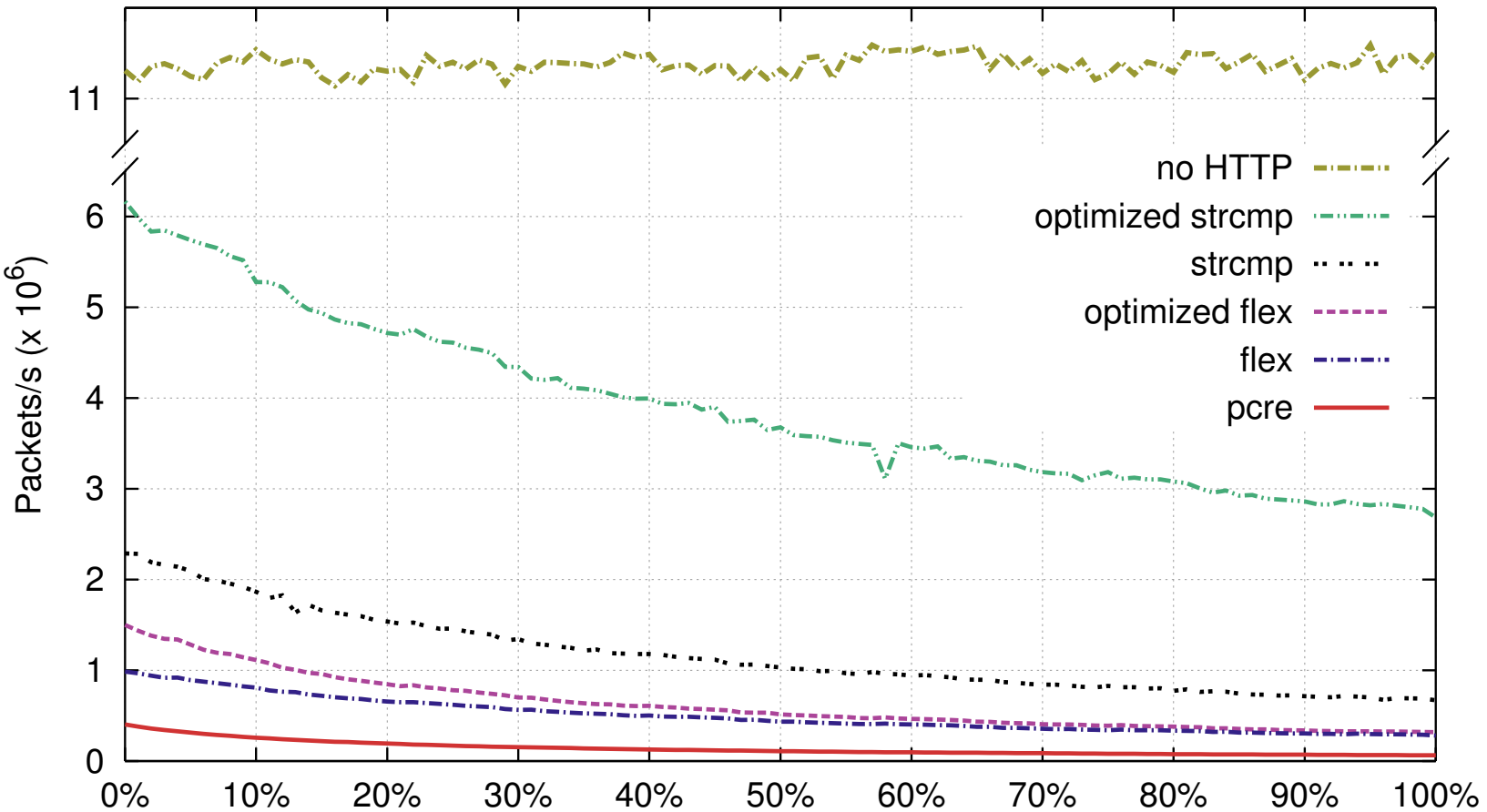
Application Flow Impacts

- R.Q. (1): **What are the impacts of application protocol measurement on flow exporters?**
  - CPU intensive processing
  - Flow cache memory requirements
  - Increasing bandwidth requirements
- Results
  - Design and Evaluation of HTTP Protocol Parsers for IPFIX Measurement[1]
  - FlowMon - Plugins for HTTP Monitoring (2012)
- Future work
  - Quantify the impacts
  - Propose solution for flow cache size
  - Specific compression of flow data stream

---

[1] Petr Velan, Tomáš Jirsík and Pavel Čeleda. **Design and Evaluation of HTTP Protocol Parsers for IPFIX Measurement.** In *Lecture Notes in Computer Science, Vol. 8115*, pages 136-147, Chemnitz, Germany, 2013.

HTTP Parsers Performance Decline



Portion of HTTP traffic in the mix (0 % - no HTTP, 100 % - only HTTP headers)

Application Flow Performance

- R.Q. (2): **What are the limits of application protocol measurement on high-speed networks?**
  - IP flow is capable of monitoring 40/100 Gbps
  - Application flow causes significant performance decline
  - No framework for performance comparison of flow measurement
  - Different results on different data sets

- Future Work
  - Create a methodology for comparison of flow measurement performance
  - Create data sets for testing application protocol parsers

Application Flow Benefits

- R.Q. (3): **How can application protocol information be used to improve flow measurement quality?**
  - Use application information to improve flow measurement
  - Better flow aggregation
- Results
  - Large-Scale Geolocation for NetFlow[1]
  - An Investigation Into Teredo and 6to4 Transition Mechanisms: Traffic Analysis[2]
- Future Work
  - Split flows based on application
  - Application protocol specific timeouts

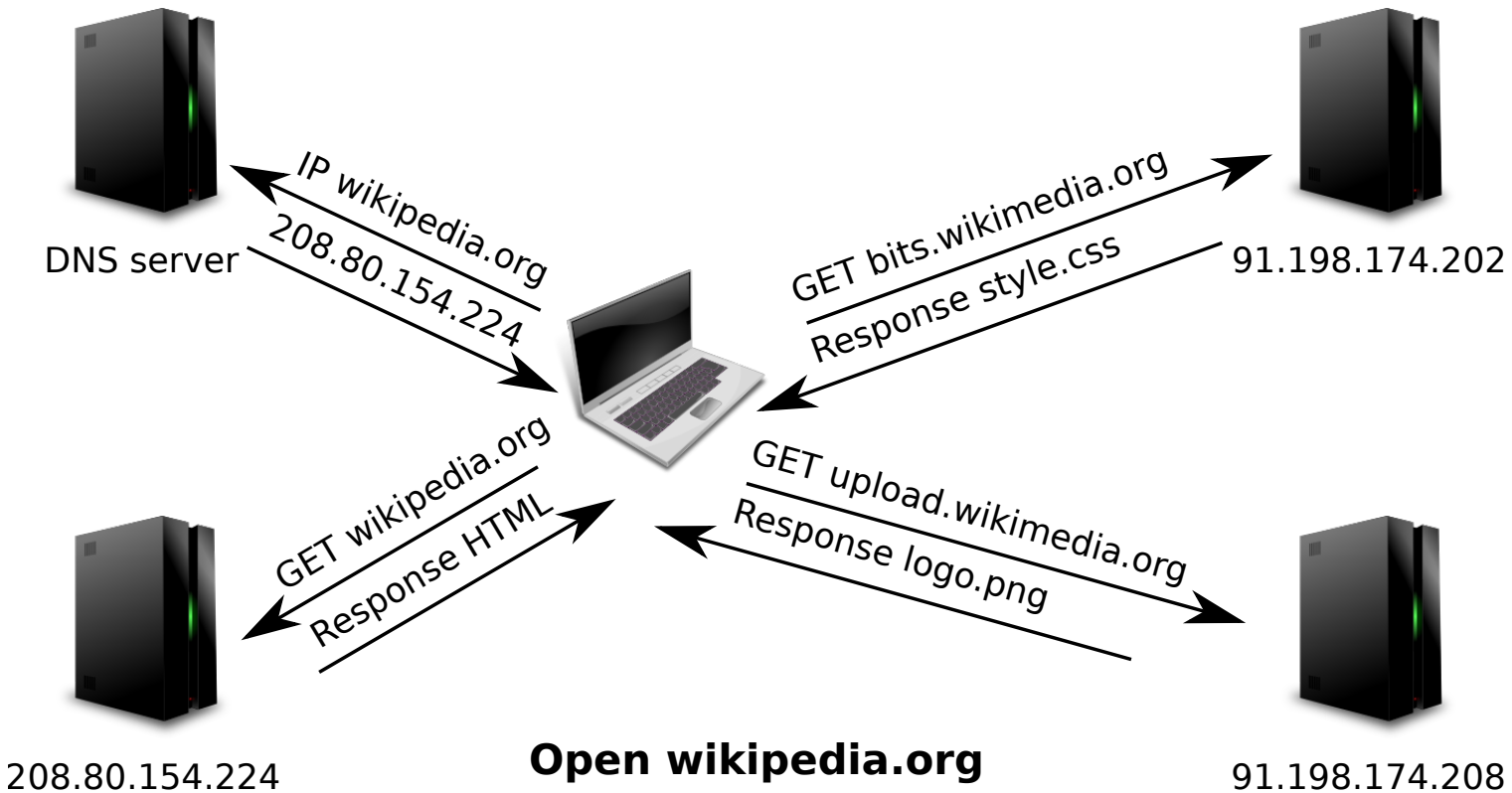
---

[1] Pavel Čeleda, Petr Velan, Martin Rábek, Rick Hofstede and Aiko Pras. **Large-Scale Geolocation for NetFlow.** In *IFIP/IEEE International Symposium on Integrated Network Management (IM 2013)*, pages 1015-1020, Ghent, Belgium, 2013.

[2] Martin Elich, Petr Velan, Tomáš Jirsík and Pavel Čeleda. **An Investigation Into Teredo and 6to4 Transition Mechanisms: Traffic Analysis.** In *38th Annual IEEE Conference on Local Computer Networks (LCN 2013)*, pages 1046-1052, Sydney, Australia, 2013.

Next Generation Flow

- R.Q. (4): **How can information from multiple packet streams be aggregated to single application event and how can we utilize application events to design the next generation flow monitoring?**



Plan of Work

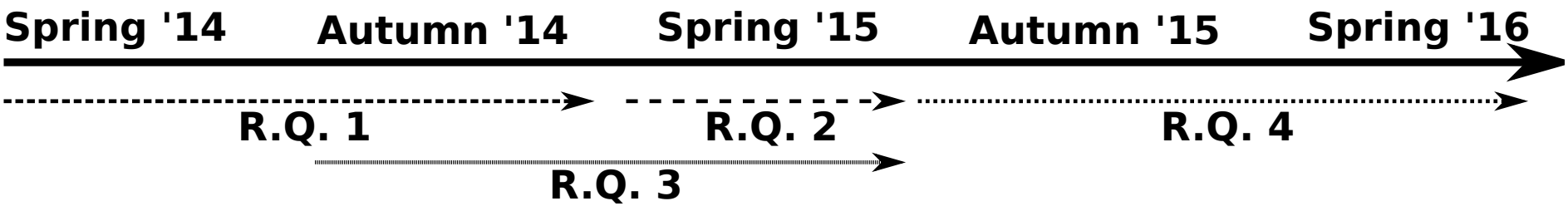
Research Questions

(1) Application Flow Impacts

(2) Application Flow Performance

(3) Application Flow Benefits

(4) Next Generation Flow



Thank You For Your Attention!

Next Generation Flow Measurement for Application Monitoring

Petr Velan

velan@ics.muni.cz

Achieved Results in Application Flow Monitoring

Papers:

- Pavel Čeleda, Petr Velan, Martin Rábek, Rick Hofstede and Aiko Pras.
  Large-Scale Geolocation for NetFlow.
  In *IFIP/IEEE International Symposium on Integrated Network Management (IM 2013)*, pages 1015-1020, Ghent, Belgium, 2013.
- Petr Velan, Tomáš Jirsík and Pavel Čeleda.
  Design and Evaluation of HTTP Protocol Parsers for IPFIX Measurement.
  In *Lecture Notes in Computer Science, Vol. 8115*, pages 136-147, Chemnitz, Germany, 2013.
- Martin Elich, Petr Velan, Tomáš Jirsík and Pavel Čeleda.
  An Investigation Into Teredo and 6to4 Transition Mechanisms: Traffic Analysis.
  In *38th Annual IEEE Conference on Local Computer Networks (LCN 2013)*, pages 1046-1052, Sydney, Australia, 2013.

Software:

- FlowMon - Plugins for HTTP Monitoring (2012)

Other results

Papers:

- Petr Velan and Radek Krejčí.
  Flow Information Storage Assessment Using IPFIXcol.
  In *Lecture Notes in Computer Science 7279*, pages 155-158, Heidelberg, 2012.
- Petr Velan.
  Practical experience with IPFIX flow collectors.
  In Filip De Turck, Yixin Diao, Choong Se on Hong, Deep Medhi, Ramin Sadre. IFIP/IEEE International Symposium on Integrated Network Management (IM 2013). Ghent, Belgium, 2013.

Software:

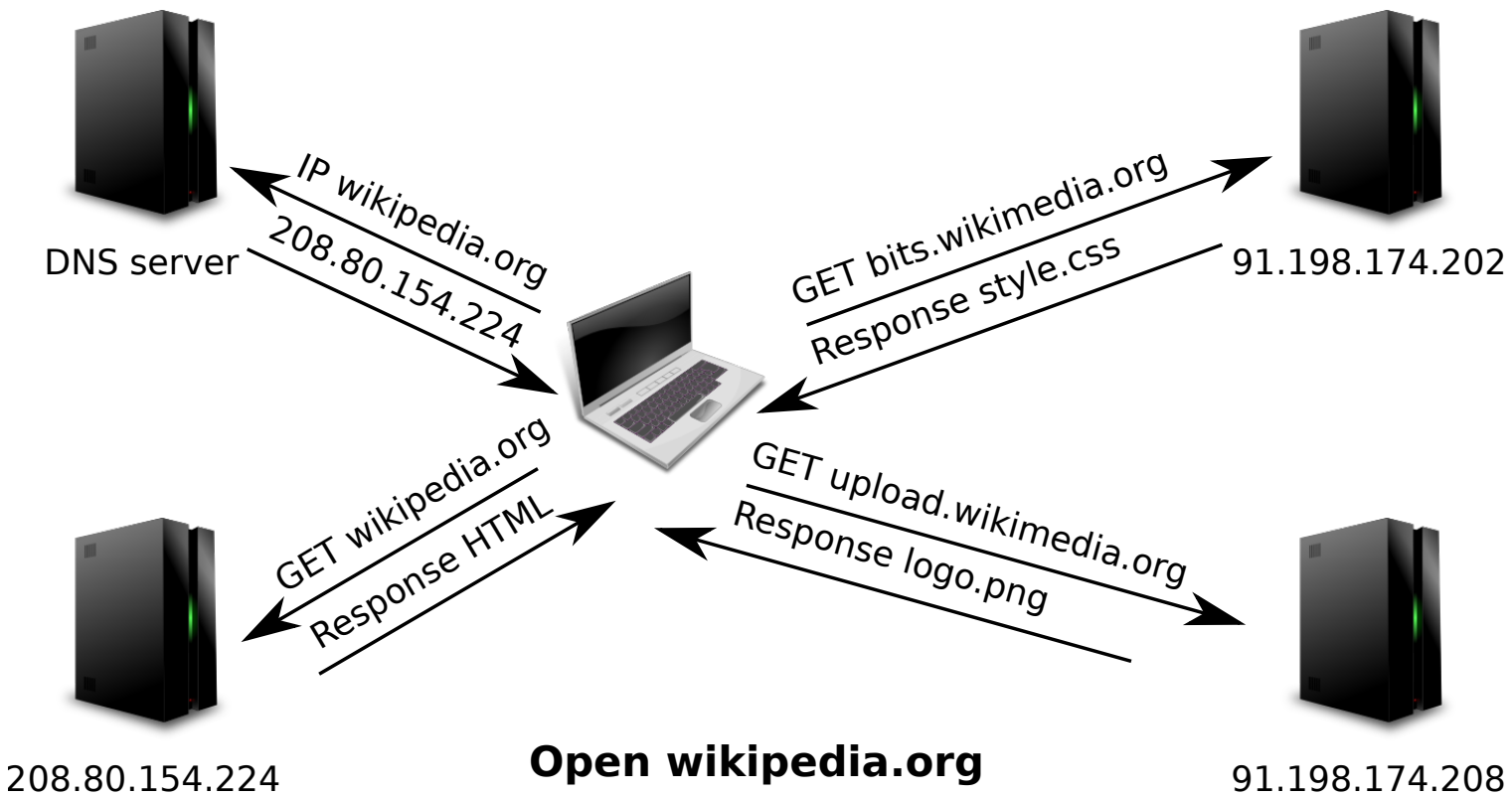
- FlowMon - IPFIX Export Plugin (2012)
- IPFIX collector (2013)
  - An implementation of IPFIX flow collector.

Current NIDS Throughput

- Flow monitoring has about 10 - 40 times the throughput of deep packet inspection
- IDS throughput depends heavily on provided features
- IDS utilize often distributed processing for speeds > 1 Gbps
- Flows allow for efficient data storage and processing of historical data

Next Generation Flow Measurement

- Prototype implementation for HTTP and DNS
- Add **Event ID** to flow records
- Group flow records by **Event ID** on collector



Amount of Processed Data

- Sampling significantly affects flow quality
- Advantage of flows is in data aggregation
- The first few packets often carry the most important information
- Amount of necessary data depends on specific protocol and analysis depth

Data Source

- Local network traffic for development
- Data from The Cybernetic Proving Ground[1]
  - Real world data
  - Experiment repeatability
- Encryption is not end-to-end, e.g. in data centers
- Privacy issues are being elevated

---

[1] http://www.muni.cz/ics/research/projects/23884/web/?lang=en