MUNI
C4E

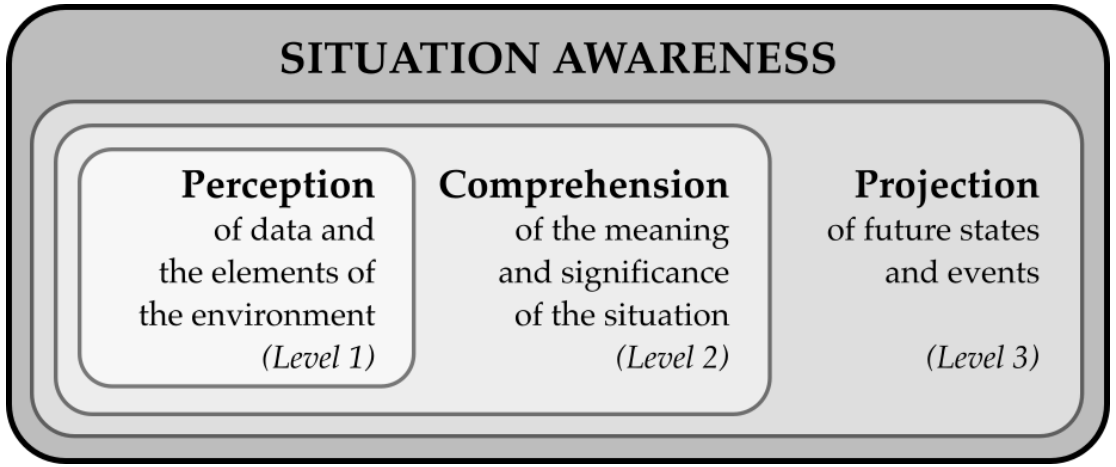
Cyber Situation Awareness via IP Flow Monitoring

NOMS 2020 – Dissertation Digest
April 23, 2020

Tomas Jirsik and Pavel Celeda

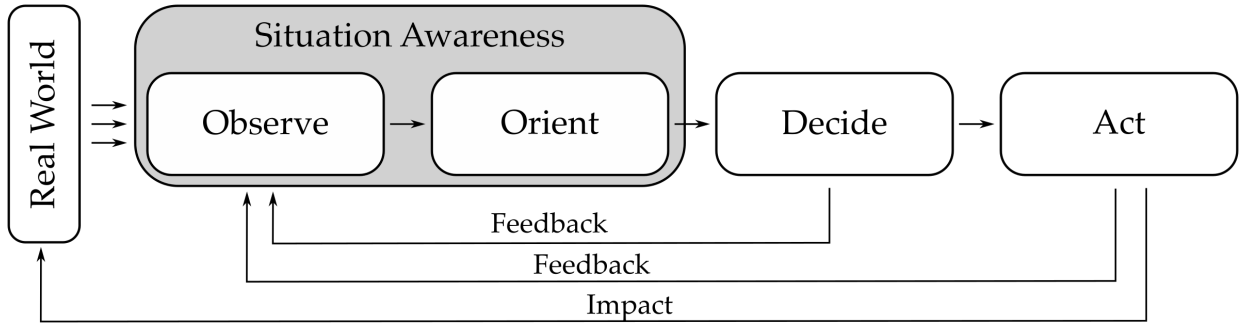
Cyber Situation Awareness

Know your network



Specifics of application in cyberspace

- **Performance** – speeds of events
- **Perception** – only by sensors



IP Flow Monitoring

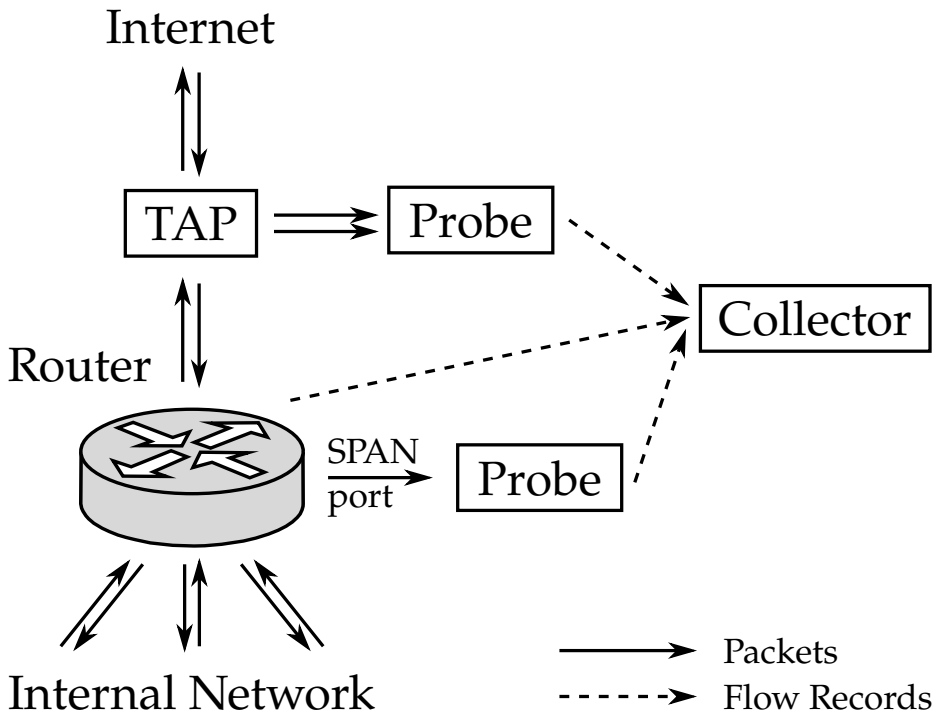
Network visibility since 1991

Connection-oriented network observation

- Aggregates packet by flow keys

Purpose

- Network reporting and analysis
- Network performance monitoring
- Attack and Anomaly Detection



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

Contemporary Challenges

Are we still aware?

Detailed view

Extended Information

Analysis Delays

Performance

Contemporary Challenges

Are we still aware?

Detailed view

Extended Information

Analysis Delays

Performance

IP Flow Monitoring needs to evolve

to maintain cyber situation awareness

Course of our Research

Step by step

Host Identification

From holistic to detailed view

Operating System Fingerprinting

Static networks

Methods

- TCP/IP parameters, User-Agent, Specific Domains, Combination

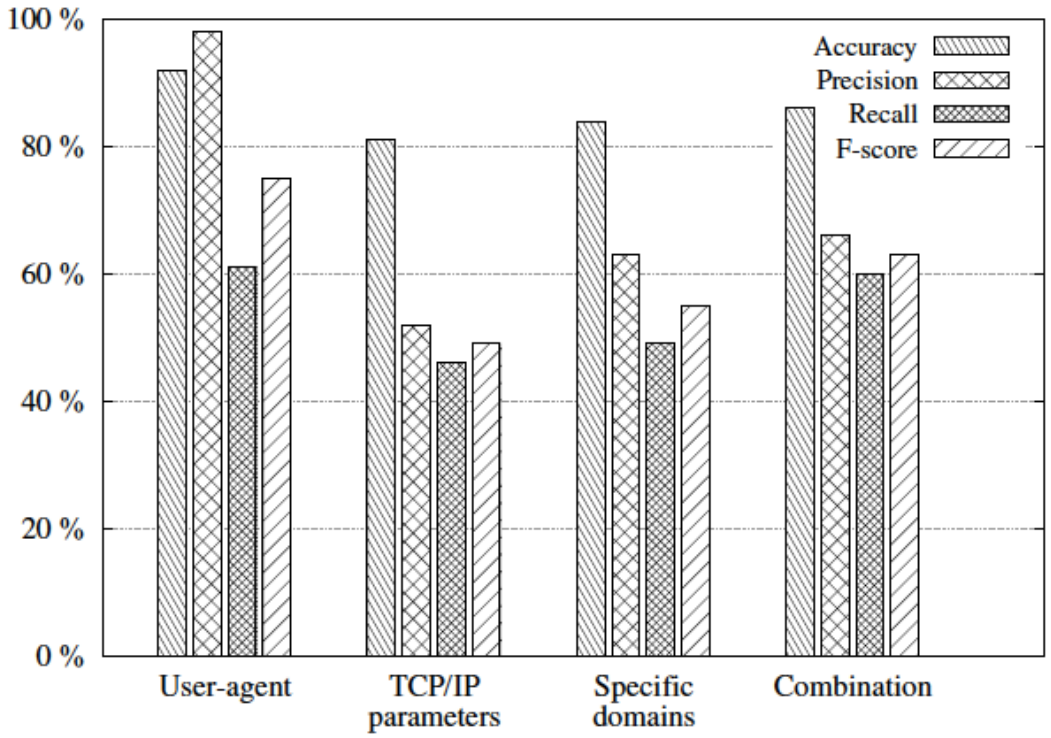
# of unique OS	# of IP in A	% of all A	# of IP in B	% of all B
1	7898	87.059	3996	95.989
2	1071	11.806	159	3.819
3	80	0.882	7	0.168
> 3	23	0.253	1	0.024
Total	9072	100	4163	100

A - whole network, **B** - dynamically addressed subnets removed

T. Jirsik and P. Celeda. "Identifying Operating System Using Flow-Based Traffic Fingerprinting". In: Advances in Communication Networking. Cham: Springer International Publishing, 2014, pp. 70–73. isbn: 978-3-319-13488-8

Operating System Fingerprinting

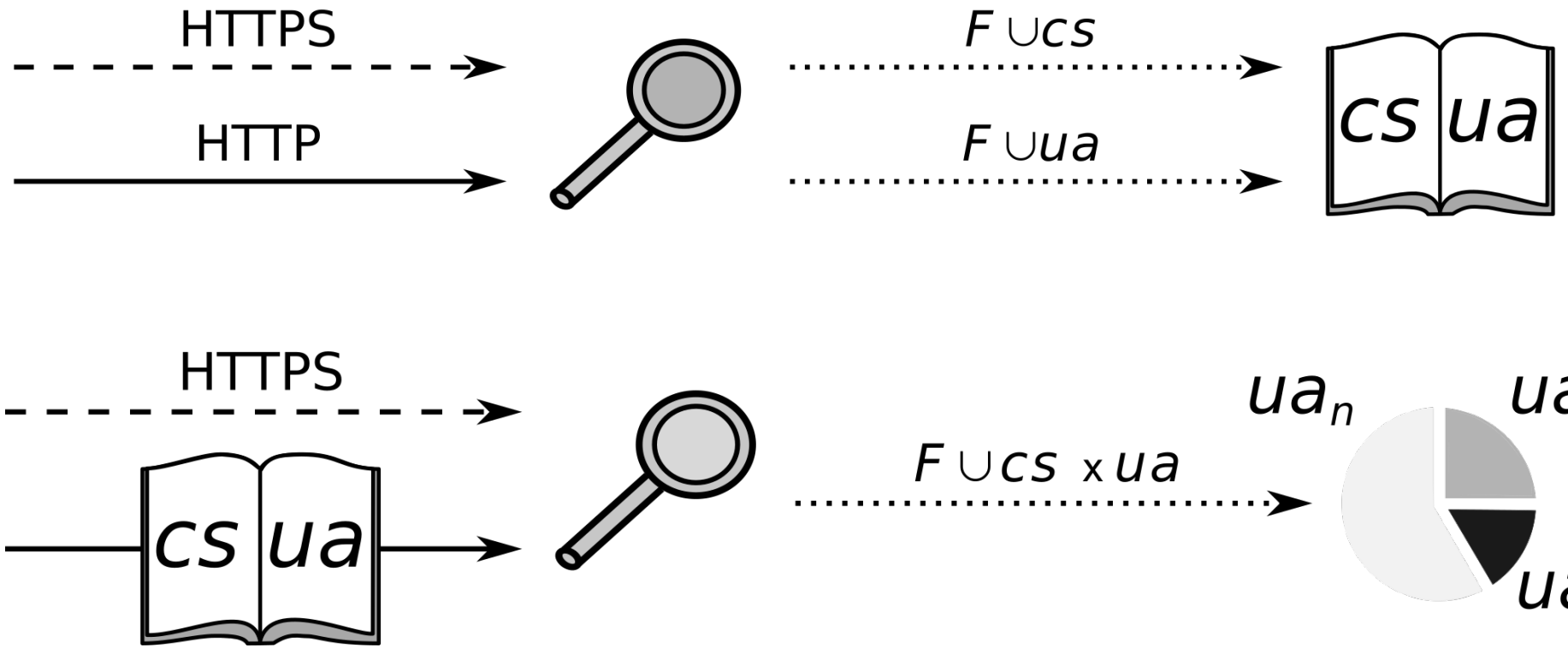
Dynamic networks



M. Lastovicka, T. Jirsik, P. Celeda, S. Spacek, and D. Filakovsky. "Passive OS Fingerprinting Methods in the Jungle of Wireless Networks". In: 2018 IEEE/IFIP Network Operations and Management Symposium. Taipei, Taiwan: IEEE, Apr. 2018

Fingerprinting in Encrypted Traffic

HTTPS traffic analysis



M. Husak, M. Cermak, T. Jirsik, and P. Celeda. "HTTPS traffic analysis and client identification using passive SSL/TLS fingerprinting". In: Eurasip Journal on Information Security 2016.1 (2016). doi: 10.1186/s13635-016-0030-7

Top N Statistics

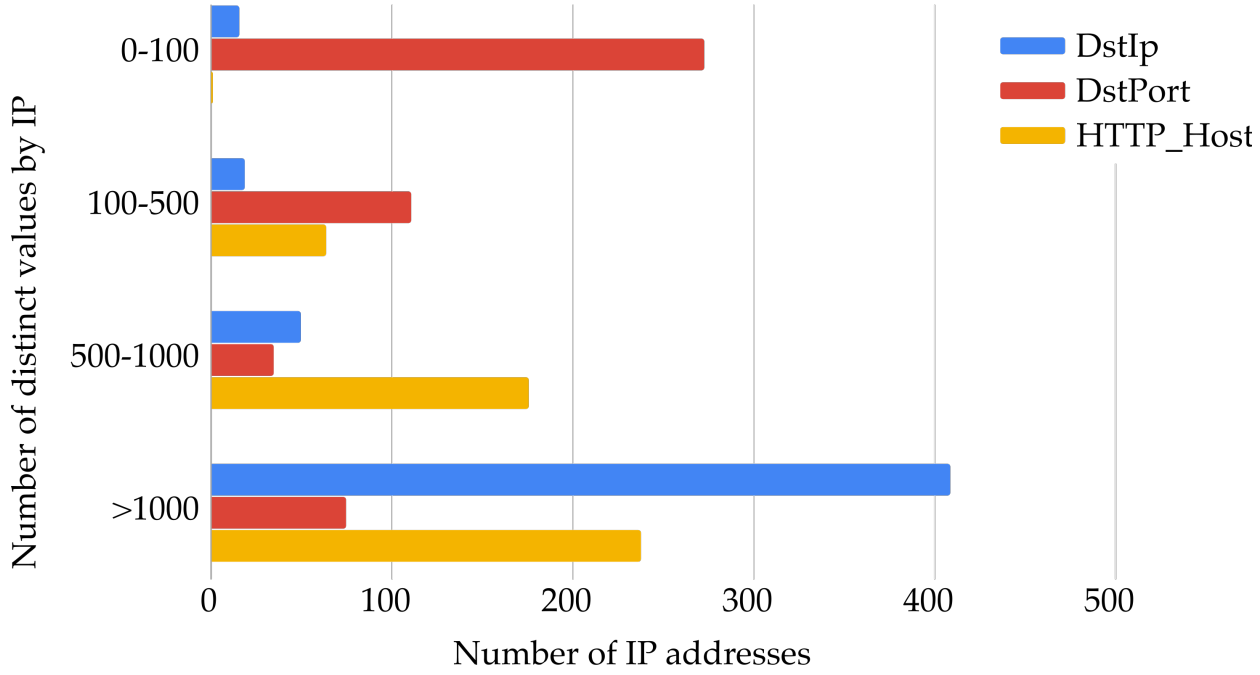
Identification based on behavioral profile

Behavioral profile

- Top 30 statistics
- Features
  - Visited peers (dst IP address)
  - Visited services (dst ports)
  - Visited domains (HTTP host)
- Different time scales

Evaluation

- Availability
- Stability
- Uniqueness
- Host identification suitability



T. Jirsik, M. Cermak, and P. Celeda. "On Information Value of Top N Statistics". In: 2016 6th International Conference on IT Convergence and Security, ICITCS 2016. 2016, pp. 1–5. isbn: 9781509037643. doi: 10.1109/ICITCS.2016.7740357

Enhanced Monitoring

More data available

HTTP Protocol Parsers

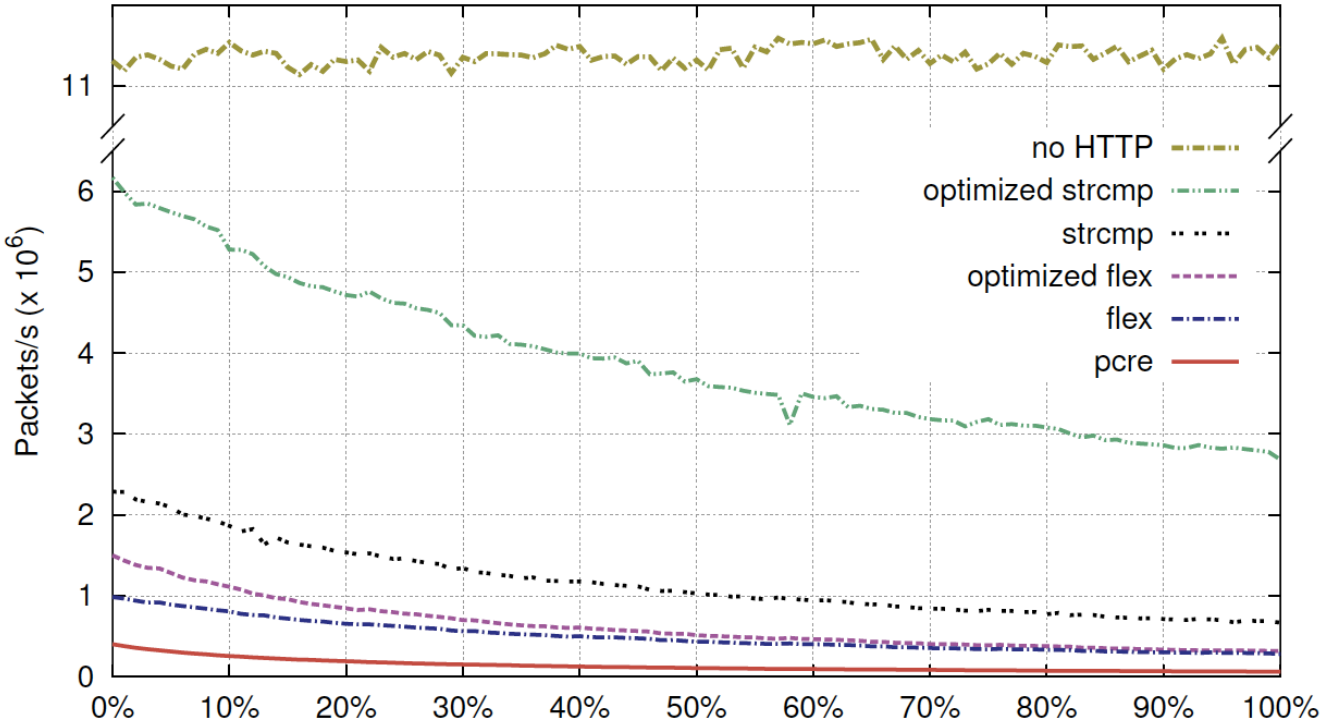
Evaluating application layer visibility

HTTP Visibility

- HTTP method, status code, host, request URI, content
- Optimization for high speed networks
- Flex-based parser

Performance comparison

- Throughput
- Number of parsed fields effect
- Packet content effect



P. Velan, T. Jirsik, and P. Celeda. "Design and Evaluation of HTTP Protocol Parsers for IPFIX Measurement". In: LectureNotes in Computer Science Vol. 8115. Springer Berlin, 2013, pp. 136–147. isbn: 978-3-642-40552-5.

Network Traffic Tunneling

Decapsulating the encapsulated

IPv6 Transition Mechanisms

- 6to4, Teredo,
- Parser to decapsulate traffic

Measurements

- Characteristics of IPv4 Tunnel Traffic
  - Frequency of TTL
  - Location of IPv4 and IPv6 Endpoints
  - Duration and Size of Flows
- Characteristics of IPv6 Tunneled Traffic
  - Distribution of HOP Limits
  - Location of Tunnel Endpoints

Server IP	Ratio	Owner	Ctry
65.55.158.118	28.33 %	Microsoft	US
94.245.121.253	27.98 %	Microsoft	GB
157.56.149.60	26.49 %	Microsoft	US
157.56.106.184	10.18 %	Microsoft	US
94.245.115.184	6.41 %	Microsoft	GB
83.170.6.76	0.04 %	B. Schmidt	DE
170.252.100.131	0.01 %	Accenture	US
94.245.127.72	0.01 %	Microsoft	GB
94.245.121.251	0.01 %	Microsoft	GB
217.31.202.10	0.01 %	CZ.NIC	CZ

Teredo endpoints

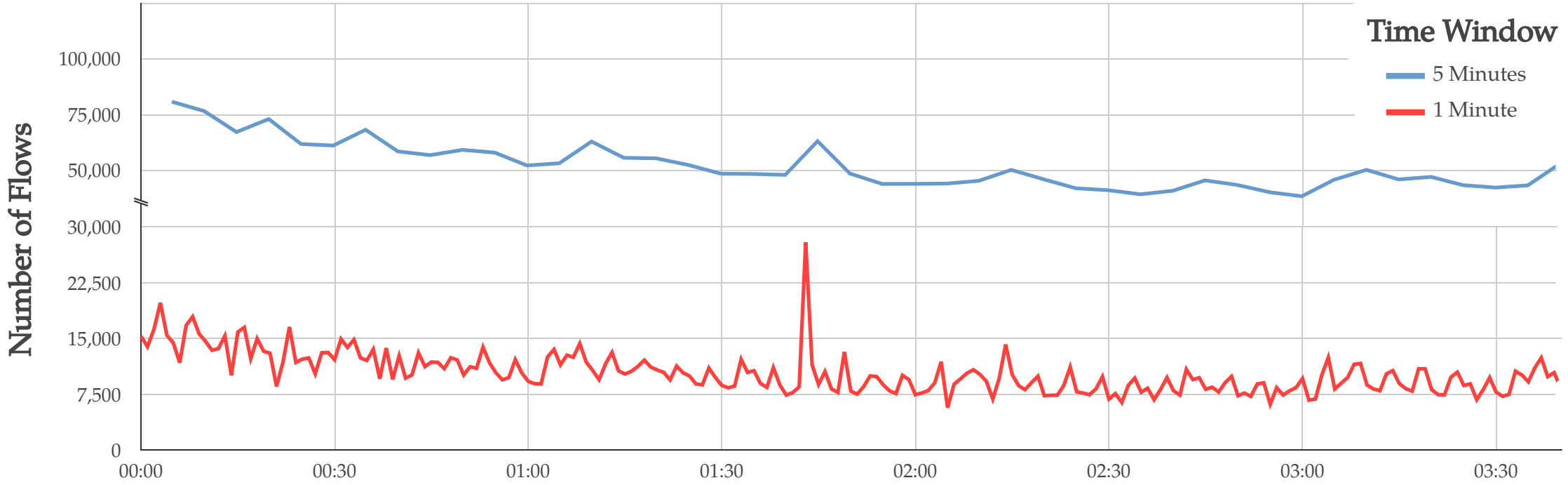
M. Elich, P. Velany, T. Jirsik, and P. Celeda. "An Investigation into Teredo and 6to4 Transition Mechanisms: Traffic Analysis". In: Proceedings of Conference on Local Computer Networks, LCN. 2013, pp. 1018–1024. doi: 10.1109/LCNW.2013.6758546

Toward Real-time Awareness

From minutes to seconds

Real-time IP Flow Analysis

Greater detail available



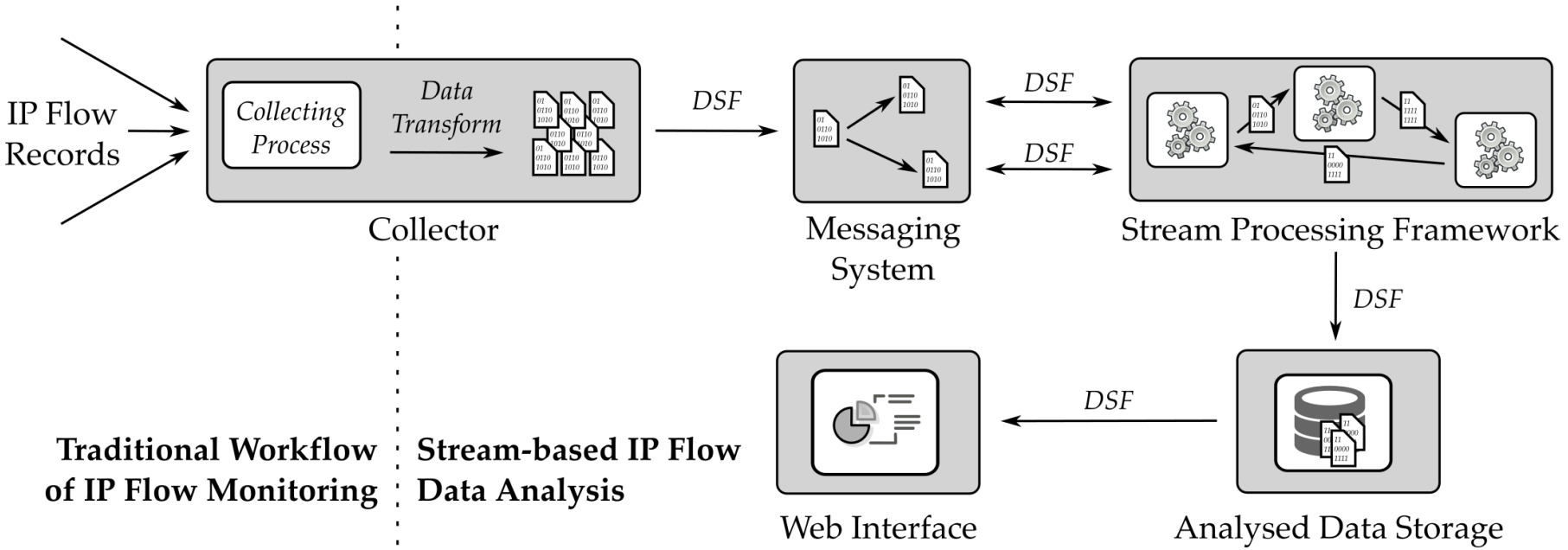
Real-time IP Flow Analysis

Taking advantage of data stream processing

Traditional processing	vs.	Stream processing
Data stored as persistent sets	**Data**	Infinite streams of individual data tuples
Large secondary memory	**Storage**	Small primary memory
Ad-hoc	**Queries**	Continuous
No real-time capabilities	**Real-time**	Real-time processing
Single-query	**Optimization**	Multi-query
Mature tools and technologies	**Maturity**	New tools and technologies

Real-time IP Flow Analysis

Distributed data stream processing



T. Jirsik, M. Cermak, D. Tovarnak, and P. Celeda. "Toward Stream-based IP Flow Analysis". In: IEEE Communications Magazine 55.7 (2017), pp. 70–76. issn: 01636804. doi: 10. 1109/MCOM.2017.1600972

Stream4Flow

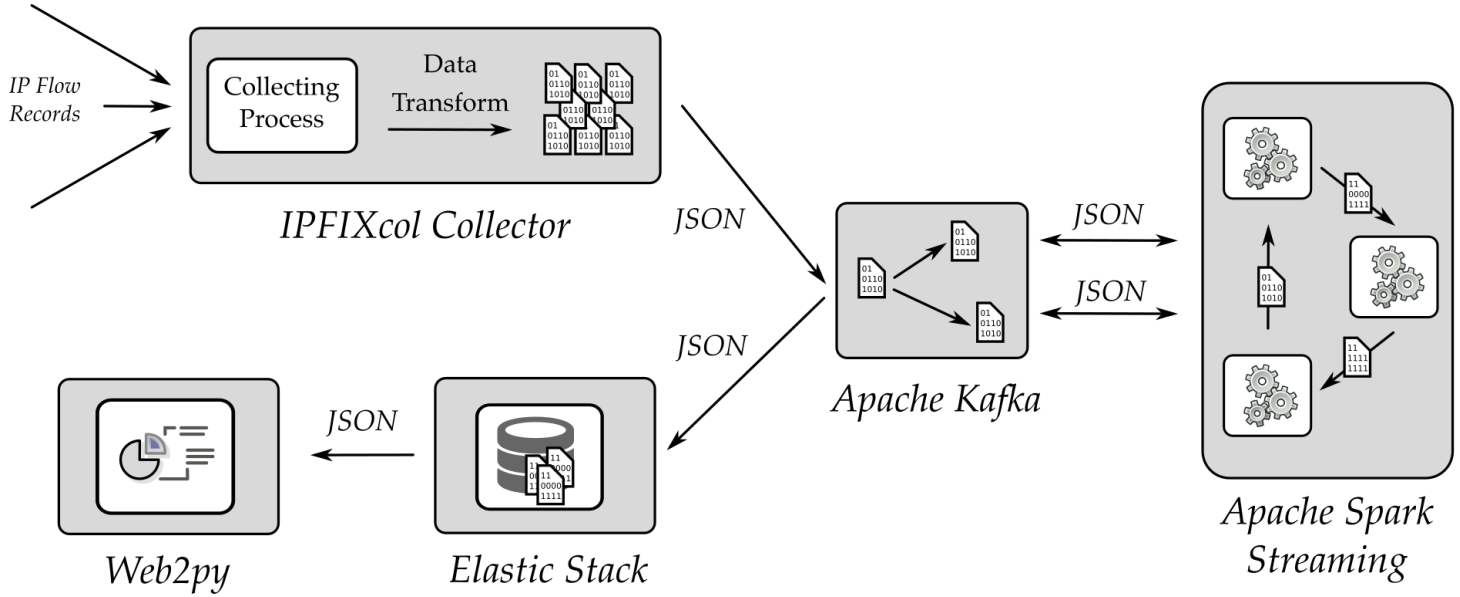
A prototype demonstrator

Features

- Full-stack solution
- High performance
- Easy deployment
- Real-time Analysis

Implemented Applications

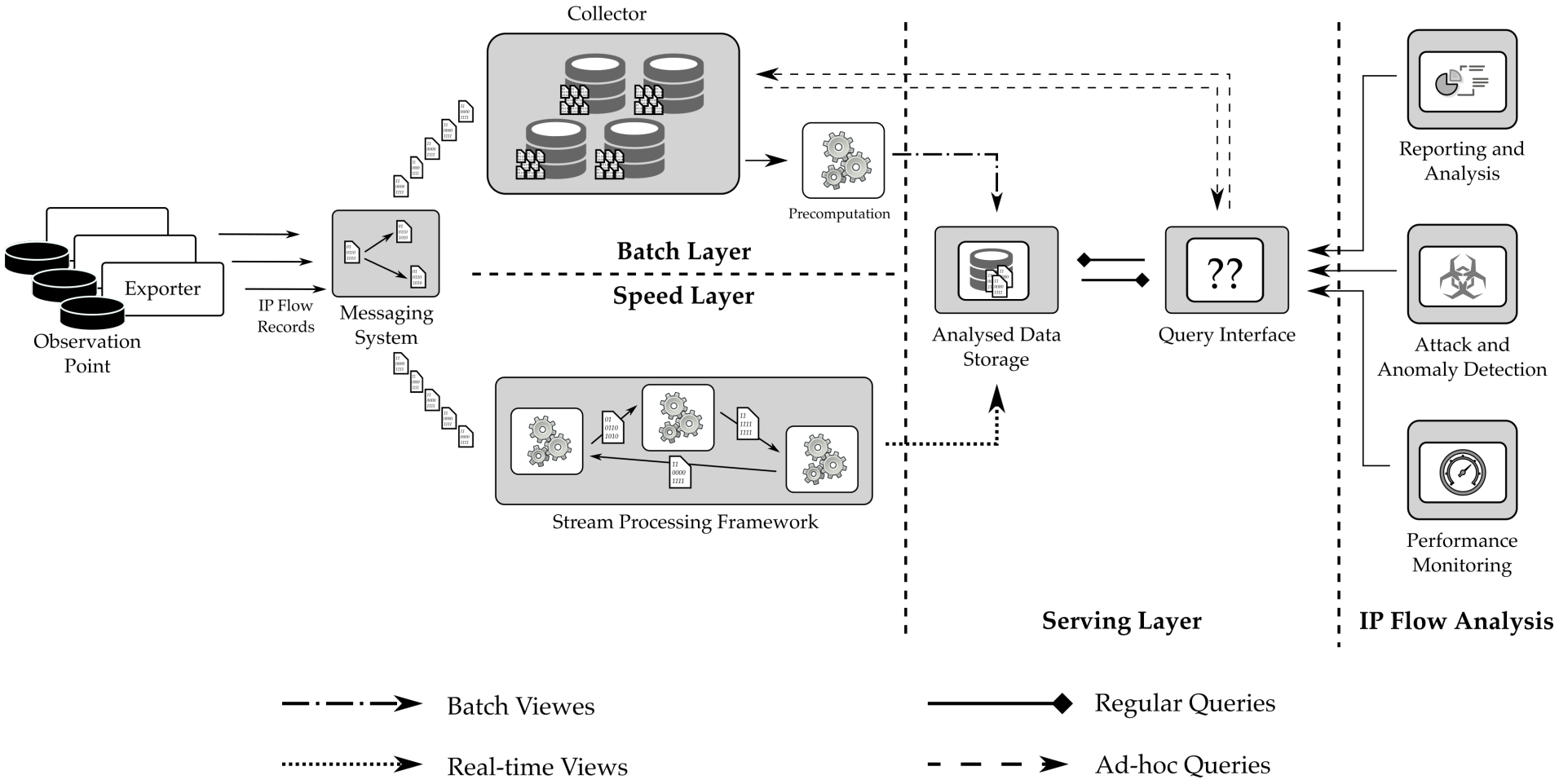
- Statistics
  DNS, Protocol, Host
- Detections
  Scans, Brute-force, DNS Resolvers



https://stream4flow.ics.muni.cz/

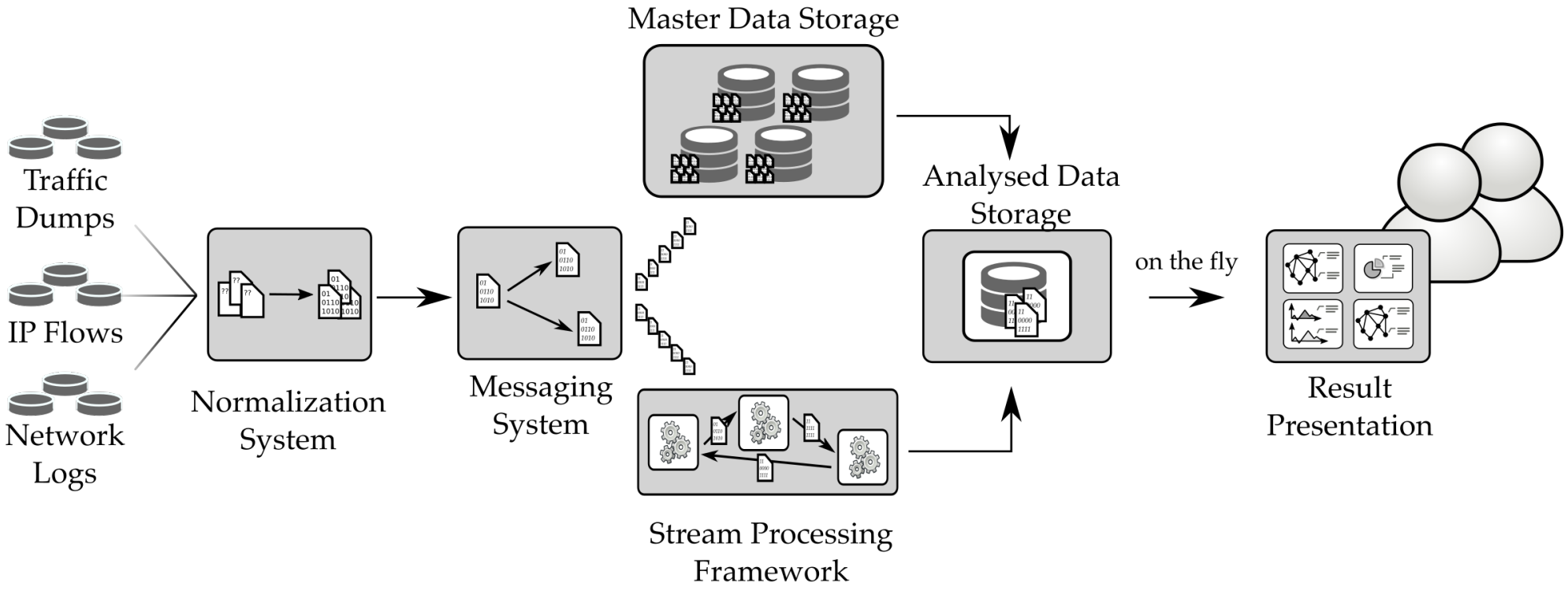
Next-generation IP Flow Monitoring

Lambda-architecture



Toward Real-time Network-wide Cyber Situation Awareness

Any data processed and analyzed



T. Jirsik and P. Celeda. "Toward Real-time Network-wide Cyber Situational Awareness". In: NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium. Taipei, Taiwan: IEEE, 2018, p. 7. doi: 10.1109/NOMS.2018.8406166

Summary

and further research directions

Main Objectives

What we have achieved

Investigate how IP flow monitoring can be improved to enhance the cyber situation awareness.

Achieved Contributions

- Developed **methods for host identification** in both unencrypted and encrypted network traffic.
- Proposed and evaluated IP flow monitoring methods **that enhance network perception and comprehension** and respond to the emerging trends in the cyber situation awareness and the IP flow monitoring.
- Provided **an option for reducing the delays** in the network IP flow monitoring workflow leading to the real-time cyber situation awareness.

Further Research

Promising research directions

- Stream-based Data Mining
- Correlation of Data Sources
- Attack Prediction
- Host Trustworthiness

MUNI
C4E

Thank you for your attention

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION,
YOUTH AND SPORTS

Full dissertation available at https://is.muni.cz/th/ejynv/thesis.pdf

C4E.CZ