WP2: Experimentation: Middlebox Modeling and Testing

David Ros, SRL

Final review, Brussels, 2019-01-30

measurement and architecture for a middleboxed internet

measurement | architecture | experimentation

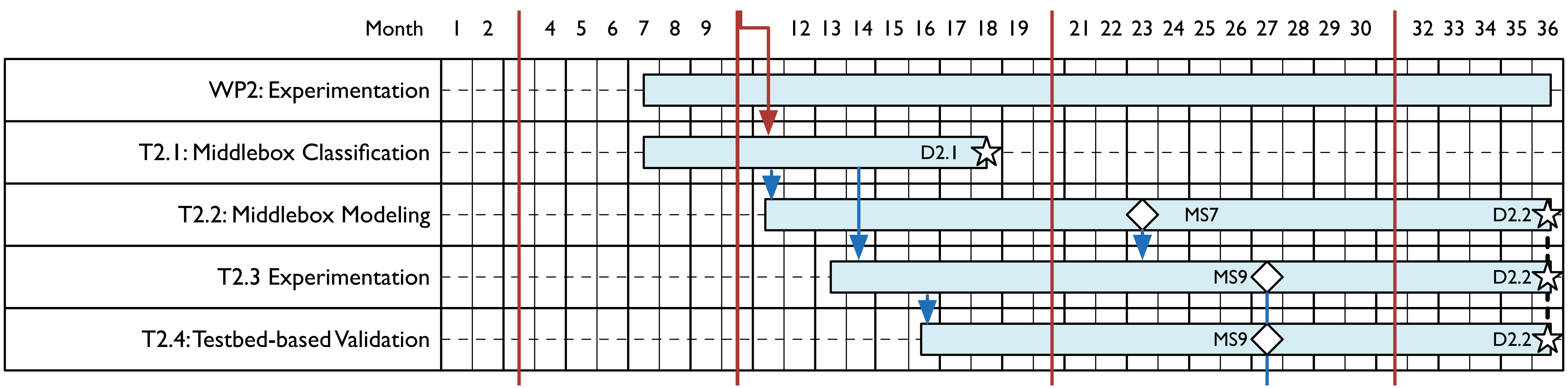
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688421.The opinions expressed and arguments employed reflect only the authors' view. The European Commission is not responsible for any use that may be made of that information.

WP2 objectives

- Development of a middlebox taxonomy, based on WP1 measurements
- Development of a **middlebox model** for testing and experimentation
- **Evaluation** of use cases applicability and deployment feasibility of protocols / protocol extensions from WP3

WP2 Tasks

- T2.1: Middlebox classification (M7 – M18)
- T2.2: Middlebox modeling (M7 – M36)
- T2.3: Model/NFV-based experimentation (M13 – M36)
- T2.4: Testbed-based validation of approach (M13 – M36)



Overview - Who did what?

Partner	Task 2.1 Middlebox classification	Task 2.2 Middlebox modeling	Task 2.3 Model/NFV-based experimentation	Task 2.4 Testbed-based validation of approach
ETH			✓	✓
TID	✓		✓	
ULg	✓	✓		
UoA	✓			✓
ZHAW			✓	
SRL	✓			✓
ALU				✓

Deliverables summary

D2.1	Middlebox Classification and Initial Model	M18
D2.2	Final Middlebox Model, Experimentation and Evaluation Report	M36

WP2 activities in the reporting period

- Modular middlebox simulator (T2.2)

- Experimental evaluation of protocols or protocol extensions, based on explicit path-to-sender or sender-to-path signaling (T2.4)
  - Datagram Packetization Layer Path MTU Discovery
  - Passive latency measurement approaches
  - Loss-Latency trade-off

- MAMI components for virtualised deployments (T2.3)
  - Extensions to the Network Modelling (NEMO) language

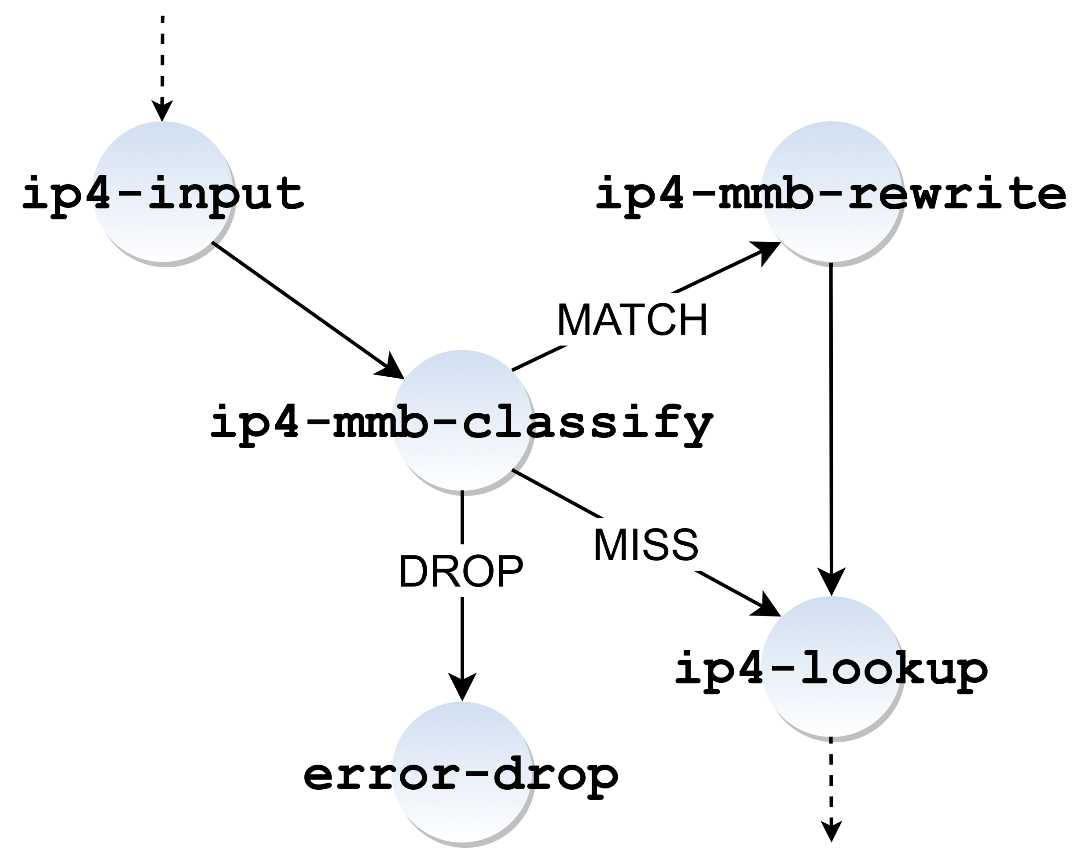
Modular middlebox simulator (mmb)

- Based on prior MAMI work on middlebox policies taxonomy (D2.1)

- Stateless and stateful packet classification and rewriting
  - Stateless packet matching based on any combination of constraints on network or transport header fields
  - Stateful TCP and UDP flow matching
  - Packet mangling or dropping
  - Bidirectional mapping

Modular middlebox simulator (mmb)

- Implemented as a Vector Packet Processor (VPP) plugin
- Two nodes added to VPP processing graph:
    - Classification
    - Rewriting



Modular middlebox simulator (mmb)

- Middlebox policies specified / configured via CLI
  - Commands for:
    - enabling / disabling mmb
    - adding stateless / stateful rules
      - mmb <add-keyword> <match> [<match> ...] <target> [<target> ...]
    - removing rules
    - displaying information

Modular middlebox simulator (mmb)

- Example
  - Drop everything but TCP:

match

- `vpp# mmb add ip-proto != tcp drop`

stateless rule

value

condition

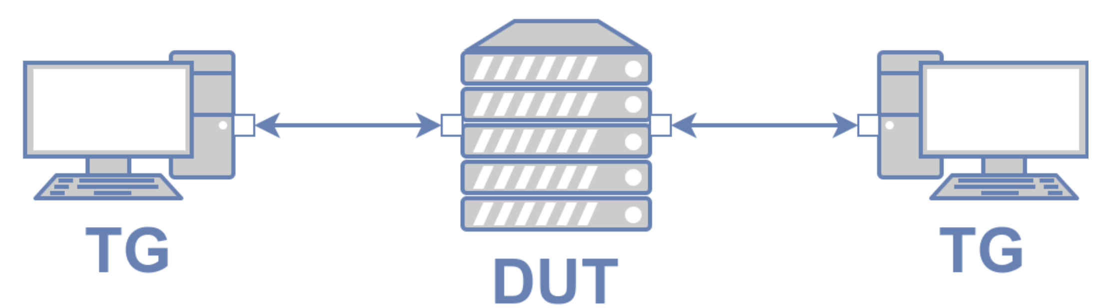
target => action to be taken if packet matches rule

protocol header field => can be IPv4, IPv6, ICMPv4, UDP, TCP (both fixed header and options)

Modular middlebox simulator (mmb)

- Performance measurements
  - mmb configured as firewall, all traffic should pass
    - Simple stateless rules, 5-tuple matching (only fast path used)
    - Parameter: # of rules that have to be evaluated
    - KPI: max throughput achieved

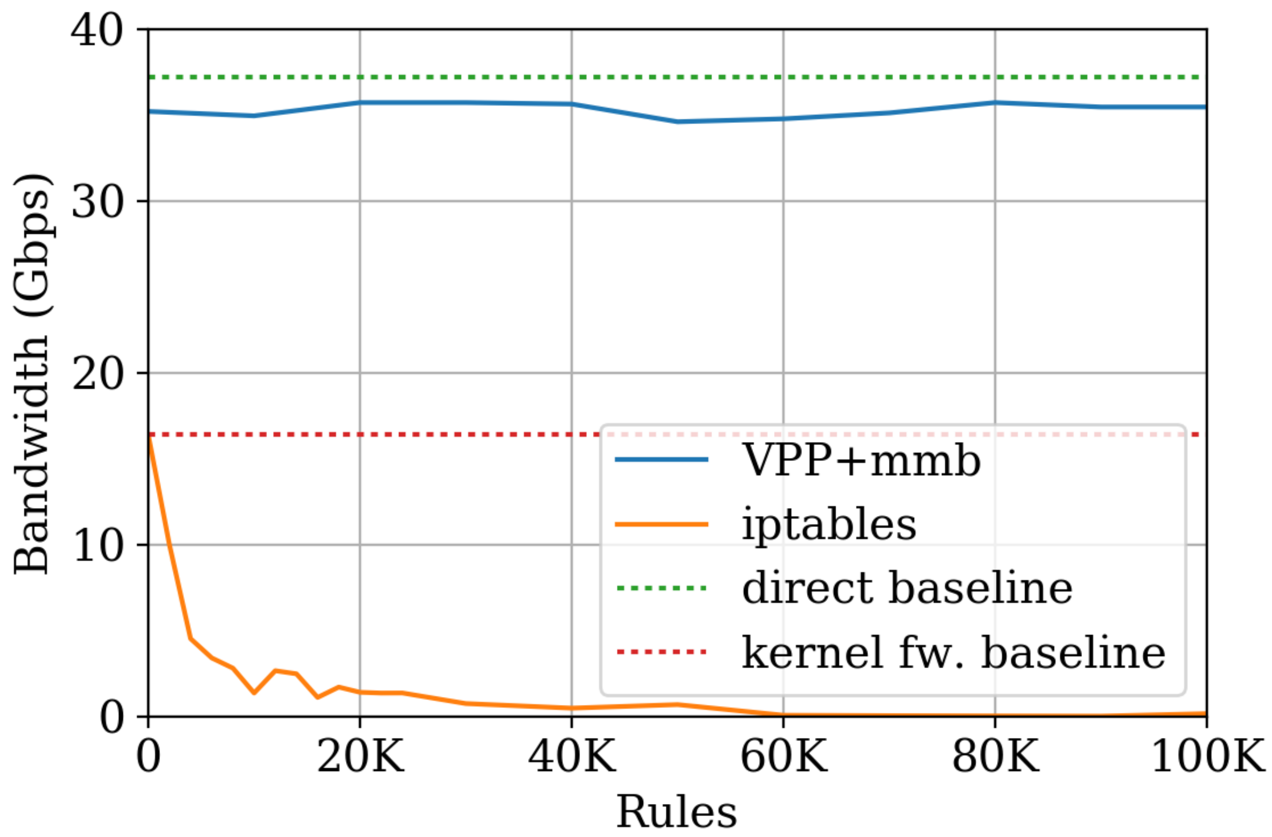
(1) calibration (direct baseline)



(2) middlebox = DUT using either mbb, iptables, or direct kernel forwarding

Modular middlebox simulator (mmb)

- Performance measurements
  - mmb configured as firewall, all traffic should pass

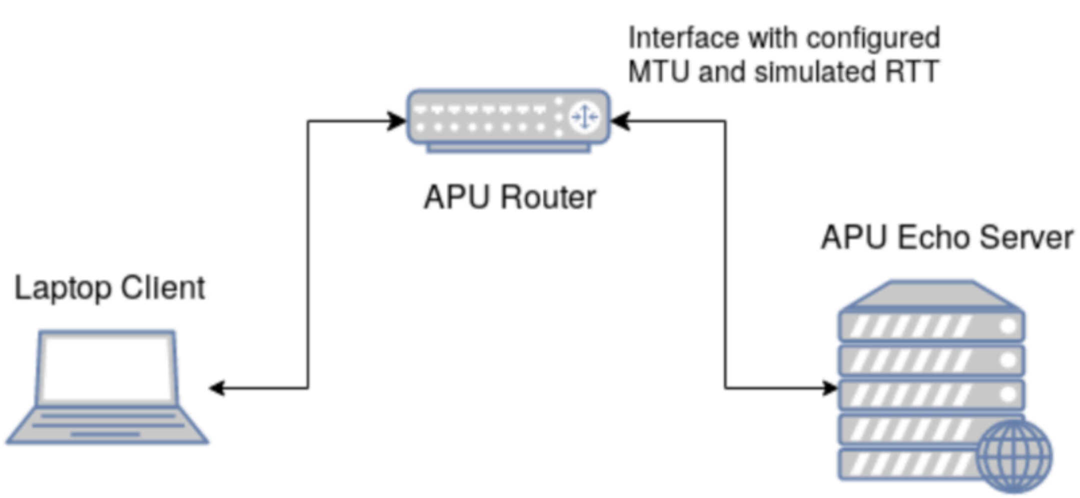


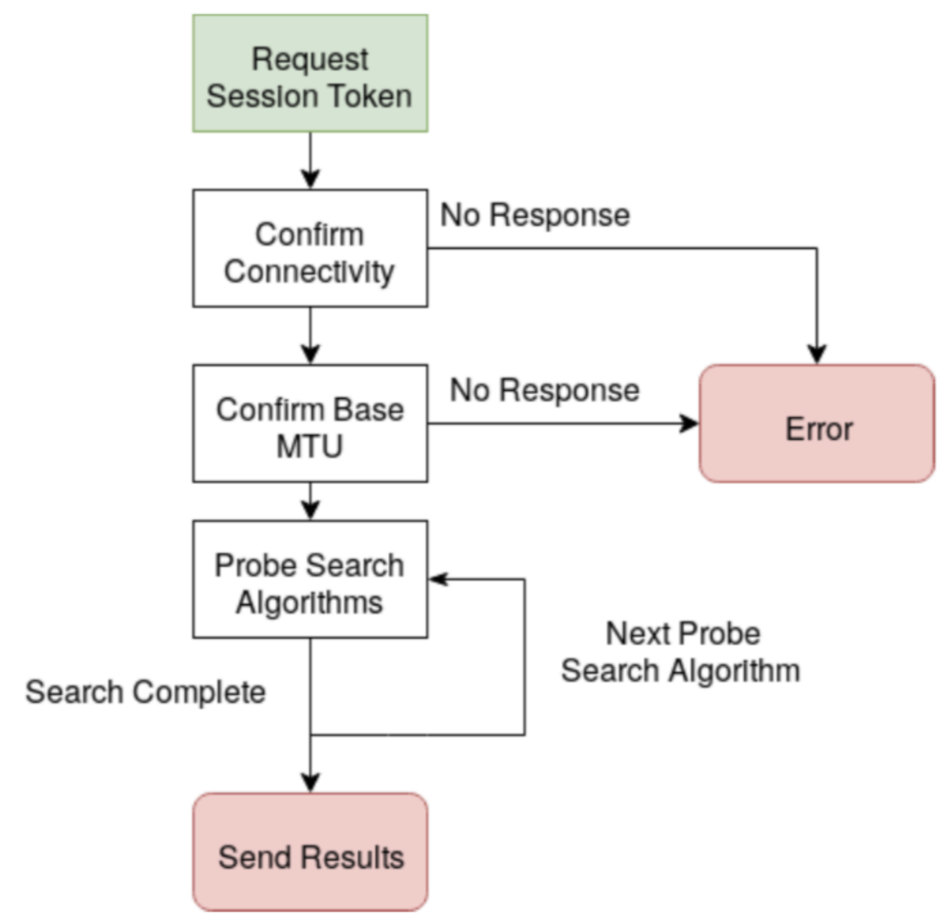
Datagram Packetization Layer Path MTU Discovery (DPLPMTUD)

- Adds PMTU discovery to datagram protocols
- Experiments to:
  - demonstrate it works as expected
  - evaluate different PMTU search methods
- Different test setups
  - Controlled lab environment and real-world paths
  - With / without support for ICMP Path Too Big (PTB) messages

Datagram Packetization Layer Path MTU Discovery (DPLPMTUD)

- Client implementing draft-ietf-tsvwg-datagram-plpmtud + either of 3 search methods:
  - step size
  - binary
  - table-based (based on common PMTU values found in WP1 experiments)





Datagram Packetization Layer Path MTU Discovery (DPLPMTUD)

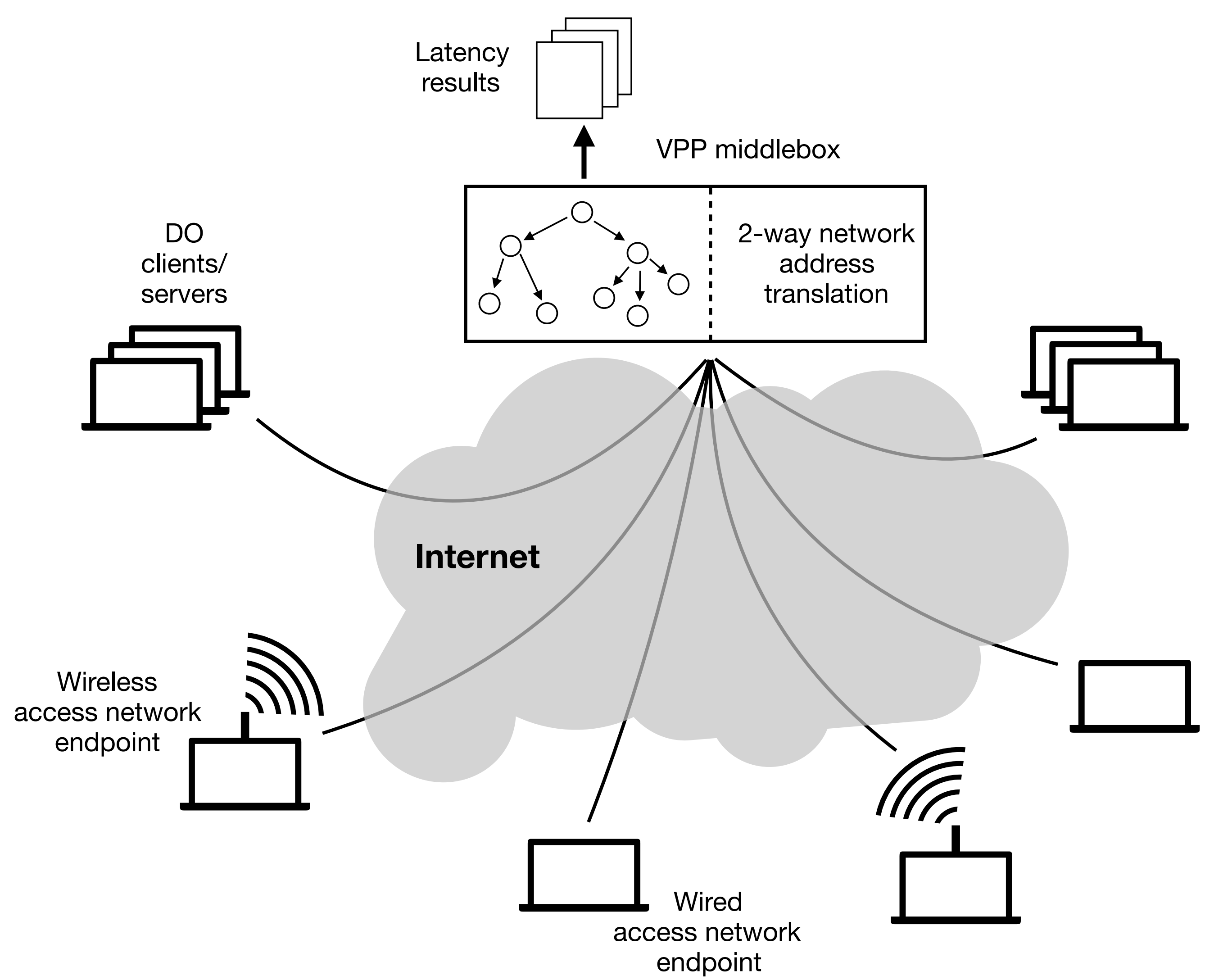
- Summary of results
  - Algorithm works without PTB support
    - PTB enabled => faster search (as expected – avoidance of a black-hole timeout)
  - No method is a clear winner, results dependent on network configuration

Passive latency measurement approaches

- VPP-based middlebox developed for latency measurements
- Experiments to:
  - Evaluate **latency spin signal**
    - Both in QUIC, TCP and PLUS
  - Compare with other approaches
    - TCP timestamps
    - PLUS Packet Serial Number (PSN) and Packet Serial Echo (PSE)
- **Latency measurement demo** to follow this presentation

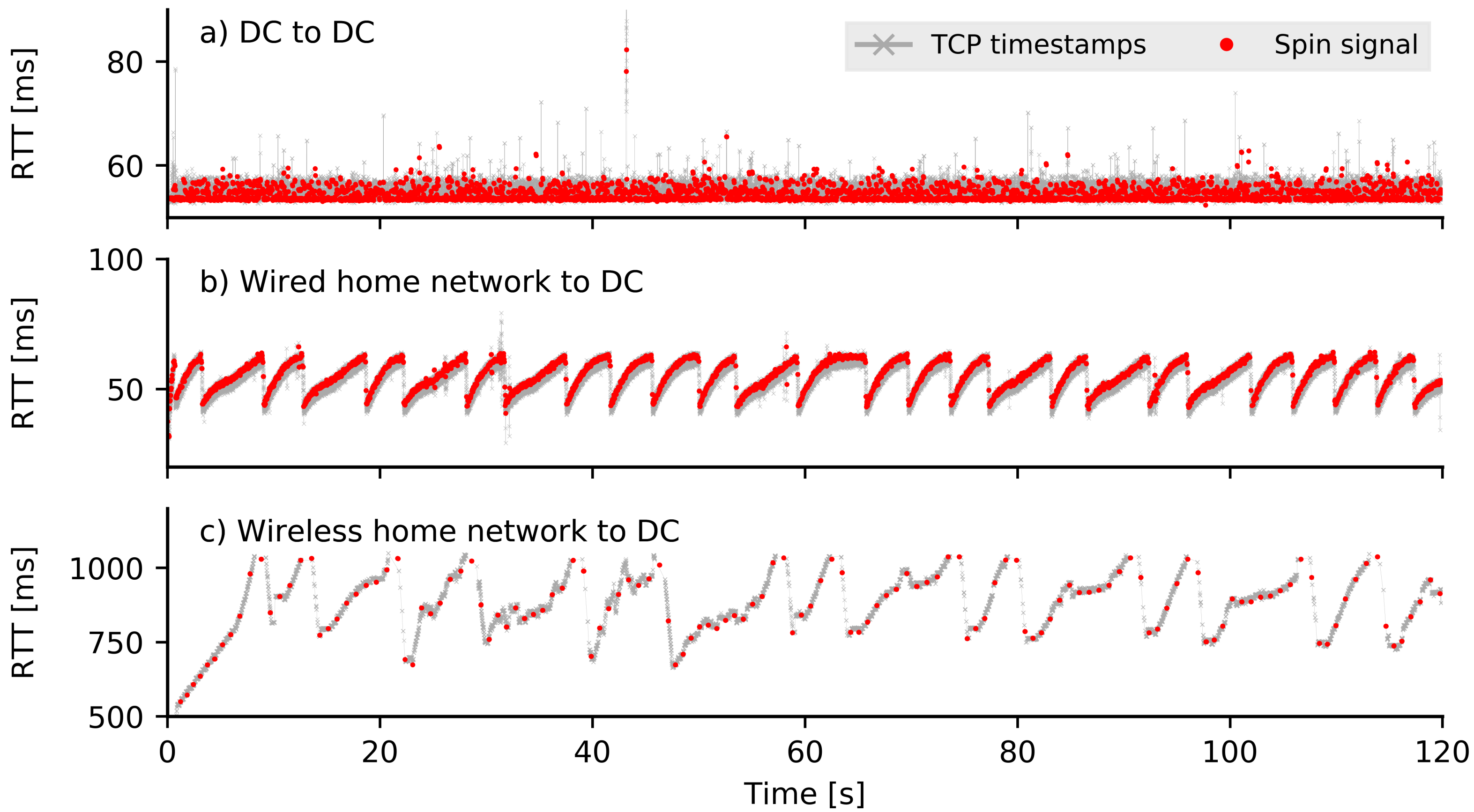
Passive latency measurement approaches

- TCP and PLUS experiments: Internet paths, hosts in:
  - MONROE testbed
  - Digital Ocean
  - wireless home
  - wired home

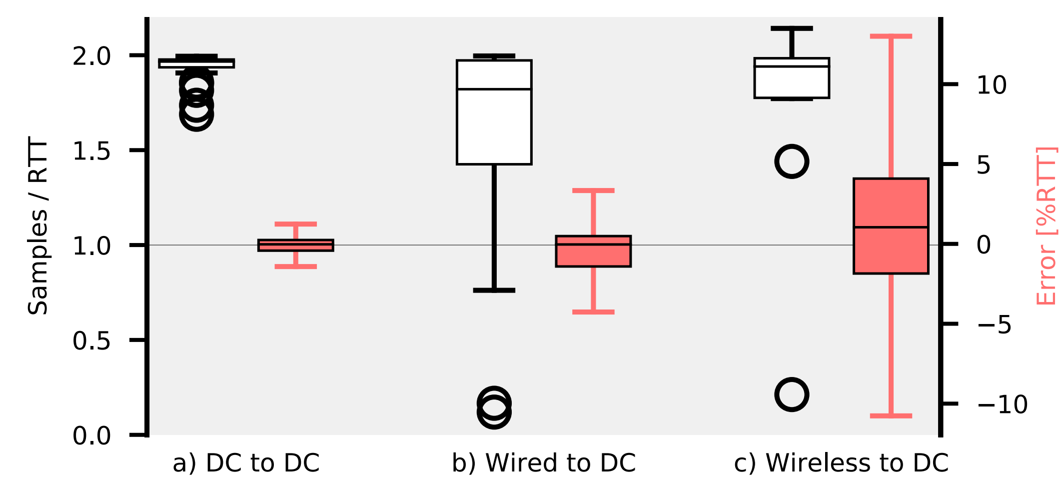


Passive latency measurement approaches

- Spin signal vs. TCP timestamps

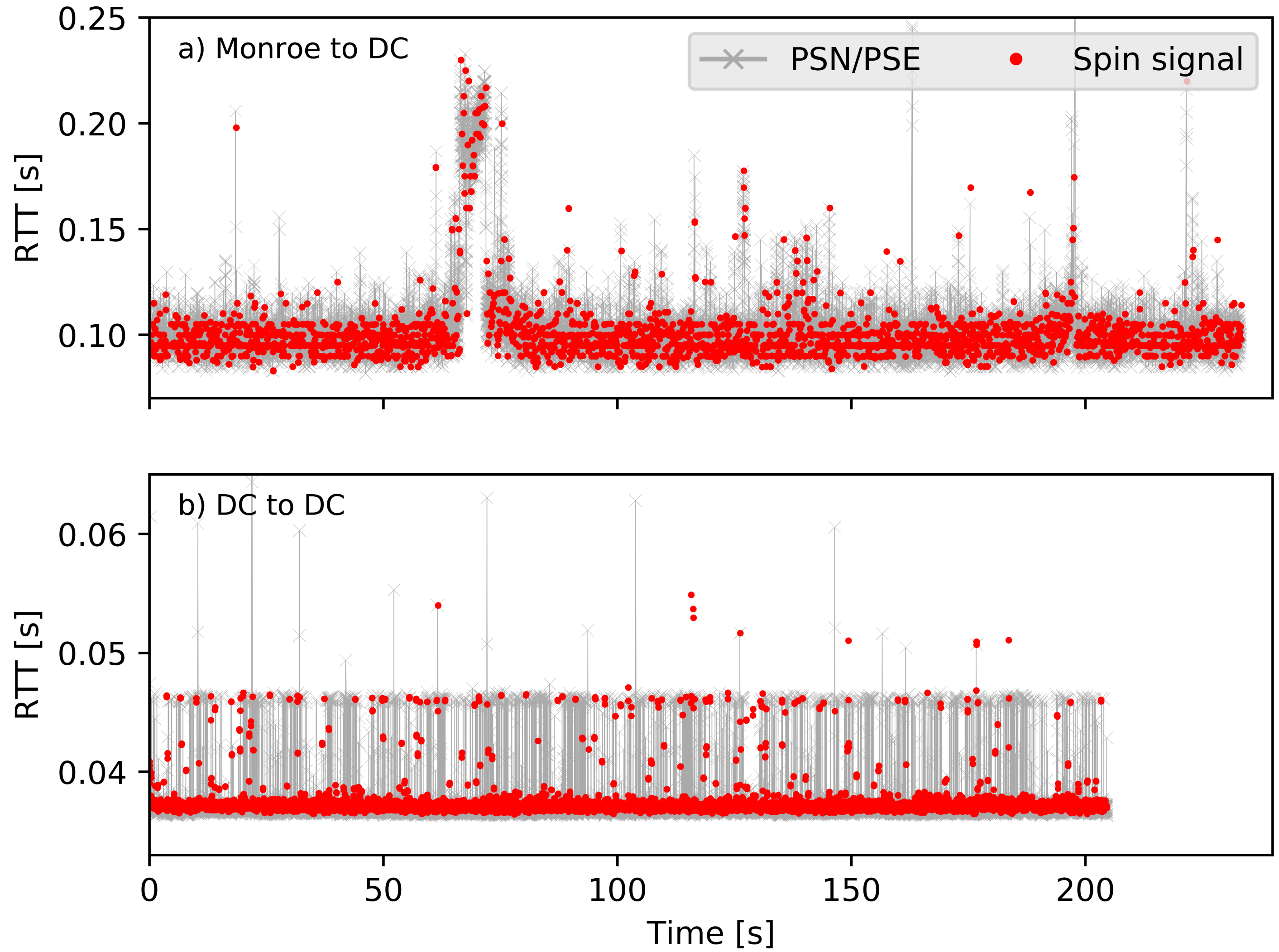


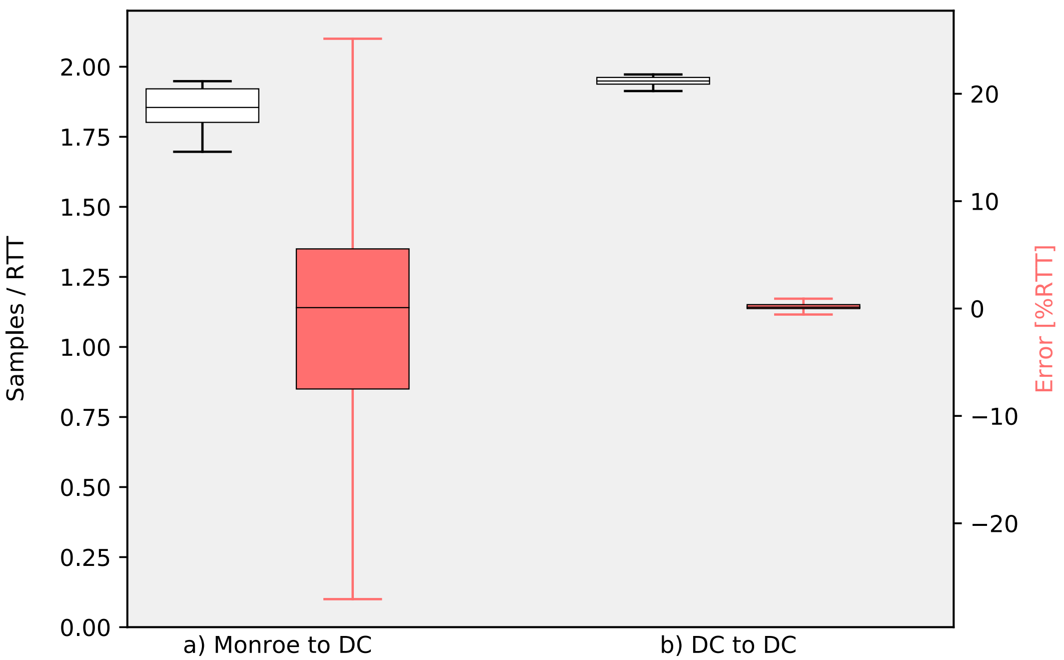
Example traces



Passive latency measurement approaches

- Spin signal vs. PLUS PSN/PSE

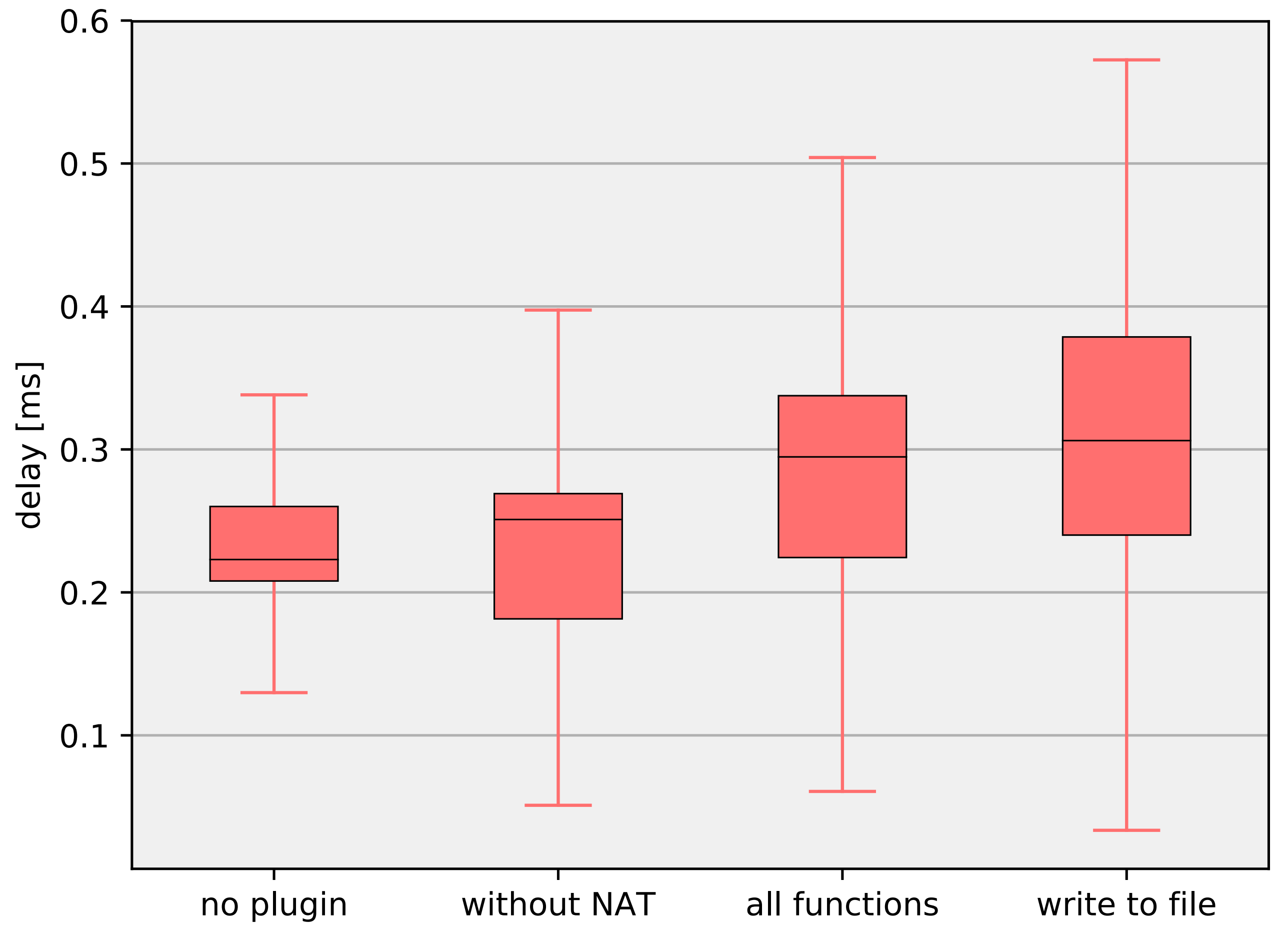




Example traces

Passive latency measurement approaches

- Performance of VPP-based latency measurement
  - Delays introduced by the plugin are small



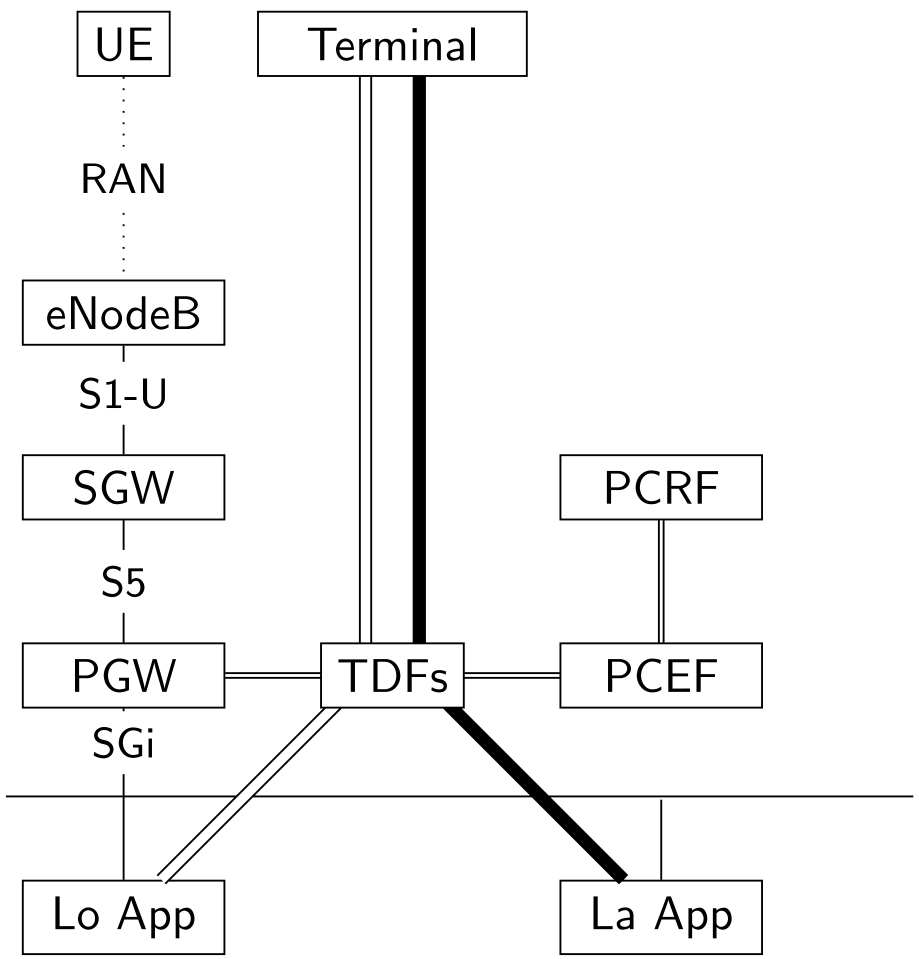
Loss-Latency Trade-off (LLT)

- Experiments to:
  - Evaluate the LLT Signal for Mobile Networks
    - Packets marked with DiffServ Code Points (DSCPs) defined in draft-you-tsvwg-latency-loss-tradeoff
    - Mapping to either of two LTE bearers: default or low-latency
  - Verify there are no incentives for cheating
  - Base comparisons with other approaches for low latency service (AQM, L4S)

- Simulation-based (ns-3) evaluation
- Testbed setup based on a set of VNFs prepared and deployed in 5TONIC (but eventually could not be used)

Loss-Latency trade-off (LoLa)

- Simple scenario: two competing flows
  - TCP large file download (prefers low loss)
  - Real-time 64 kbps audio flow (prefers low latency)
  - Baseline = no marking, i.e., both on the same (default) bearer

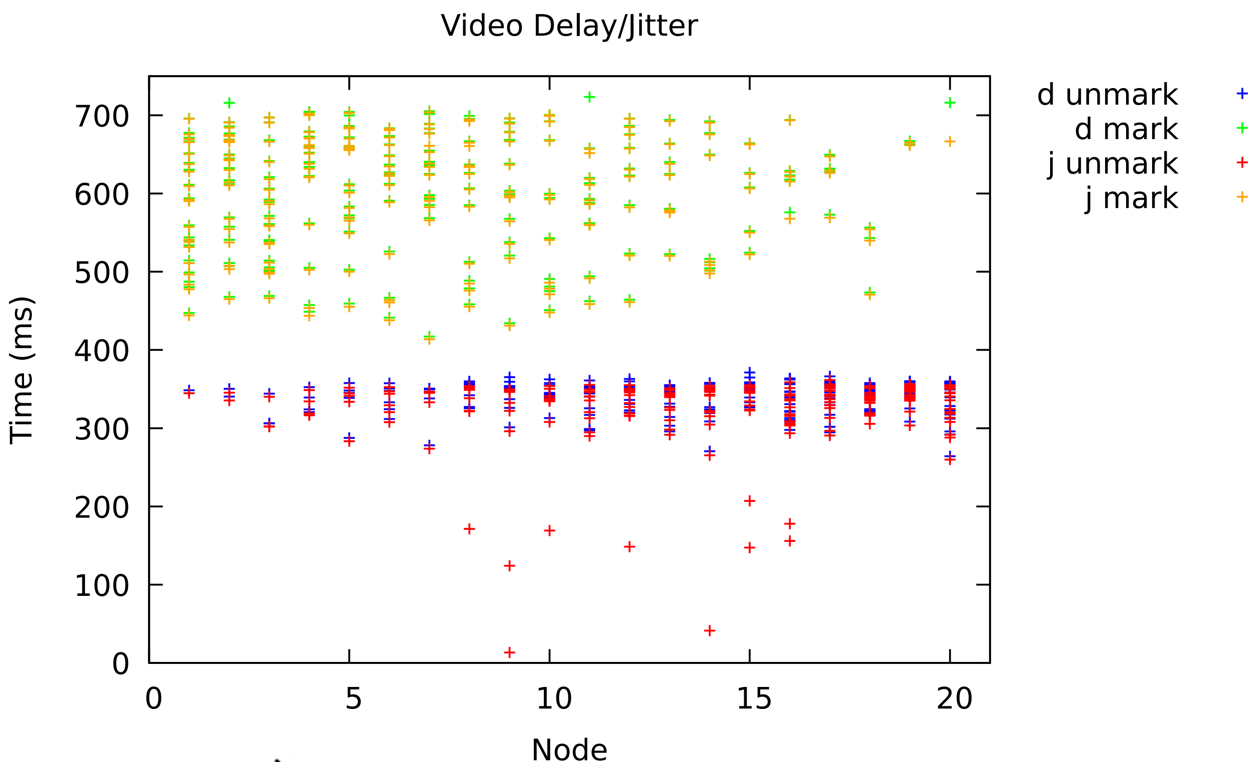


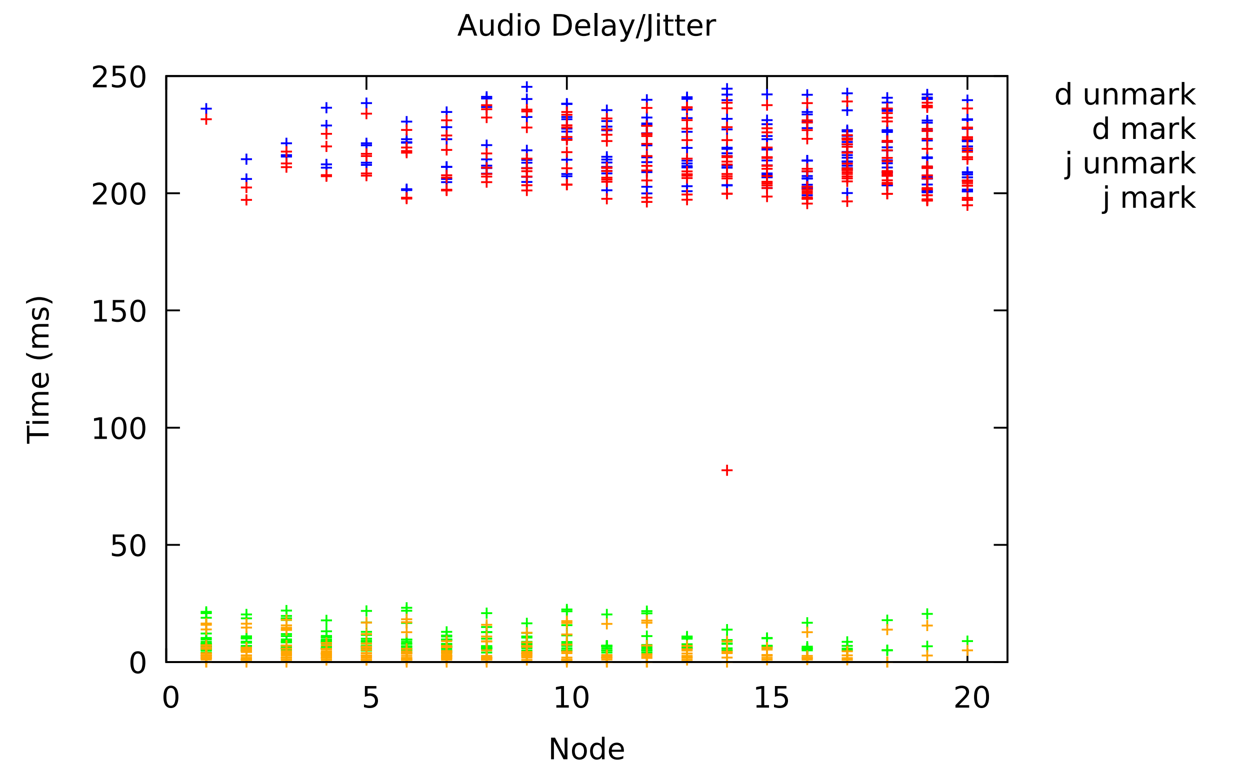
Loss-Latency trade-off (LoLa)

- Summary of results
  - Is it worth cheating? (TCP greedy flow marked as “low-latency”)
    - Answer: **no**
    - Self-inflicted higher loss & lower throughput, as expected

Loss-Latency trade-off (LoLa)

- More complex scenario: multiple UEs connected to one e-Node B
  - One audio or video CBR stream + a large file download per UE
  - In each run, some UEs use LLT marking (i.e., ≠ bearers), some don't





MAMI components for virtualised deployments

- Measurement and experimentation components prepared for virtualised deployments
  - cloud-init configuration templates for:
    - PATHspider (used for path transparency measurements in WP1)
    - VPP-latency-mb (used for passive RTT measurement experiments in WP2)
    - vpp-plus
  - VNFC and service descriptors templates for VPP-latency-mb
  - Experimental setup deployed in the 5TONIC 5G testbed

- Extensions to the Network Modeling (NEMO) language to accommodate recursive VNF Descriptors
  - Internet Draft draft-aranda-nfvrg-recursive-vnf

Summary of WP2 achievements

- Middlebox taxonomy and model **published**
- Middlebox simulator (mmb) **released as open-source software**
- **Experimental validation** of middlebox cooperation and path signaling mechanisms
  - Validation of a new Path MTU Discovery mechanism for UDP (under standardisation)
  - Validation of a new method for passive latency measurements, based on the spin signal
  - Preliminary validation (by simulation) of the LoLa approach, and preparation for experiments in a 5G testbed

WP2-related scientific publications and standardisation

- Papers
  - Under submission
    - K. Edeline, J. Iurman, C. Soldani, and B. Donnet. “mmb: Flexible High-Speed Userspace Middleboxes”. Submitted to *USENIX Annual Technical Conference (ATC)*, January 2019.
  - Published
    - P. De Vaere, T. Bühler, M. Kühlewind, and B. Trammell. “Three Bits Suffice: Explicit Support for Passive Measurement of Internet Latency in QUIC and TCP”. In *ACM IMC*, October 2018.
    - K. Edeline and B. Donnet, “An Observation-Based Middlebox Policy Taxonomy”. In *ACM CoNEXT Student Workshop*, December 2017.
    - K. Edeline and B. Donnet, “A First Look at the Prevalence and Persistence of Middleboxes in the Wild”. In *ITC29*, September 2017.

- Posters and presentations
  - T. Bühler, M. Kühlewind, and B. Trammell. “Enhancing encrypted transport protocols with passive measurement capabilities”. Poster at *ACM IMC*, November 2017.

- Internet Drafts (IRTF)
  - P. Aranda Gutierrez, D. Lopez, S. Salsano, and E. Batanero. “High-level VNF Descriptors using NEMO”. Internet draft draft-aranda-nfvrg-recursive-vnf. August 2018.

Q & A

mami

measurement and architecture for a middleboxed internet

measurement architecture experimentation

Spare slides

measurement and architecture for a middleboxed internet

measurement | architecture | experimentation

Modular middlebox simulator (mmb)

- Example
  - Source NAT:

matches

stateful rule

- `vpp# mmb add-stateful ip-saddr 10.0.0.0/24 ip-proto tcp tcp-syn shuffle tcp-sport mod ip-saddr 200.0.0.1`

target

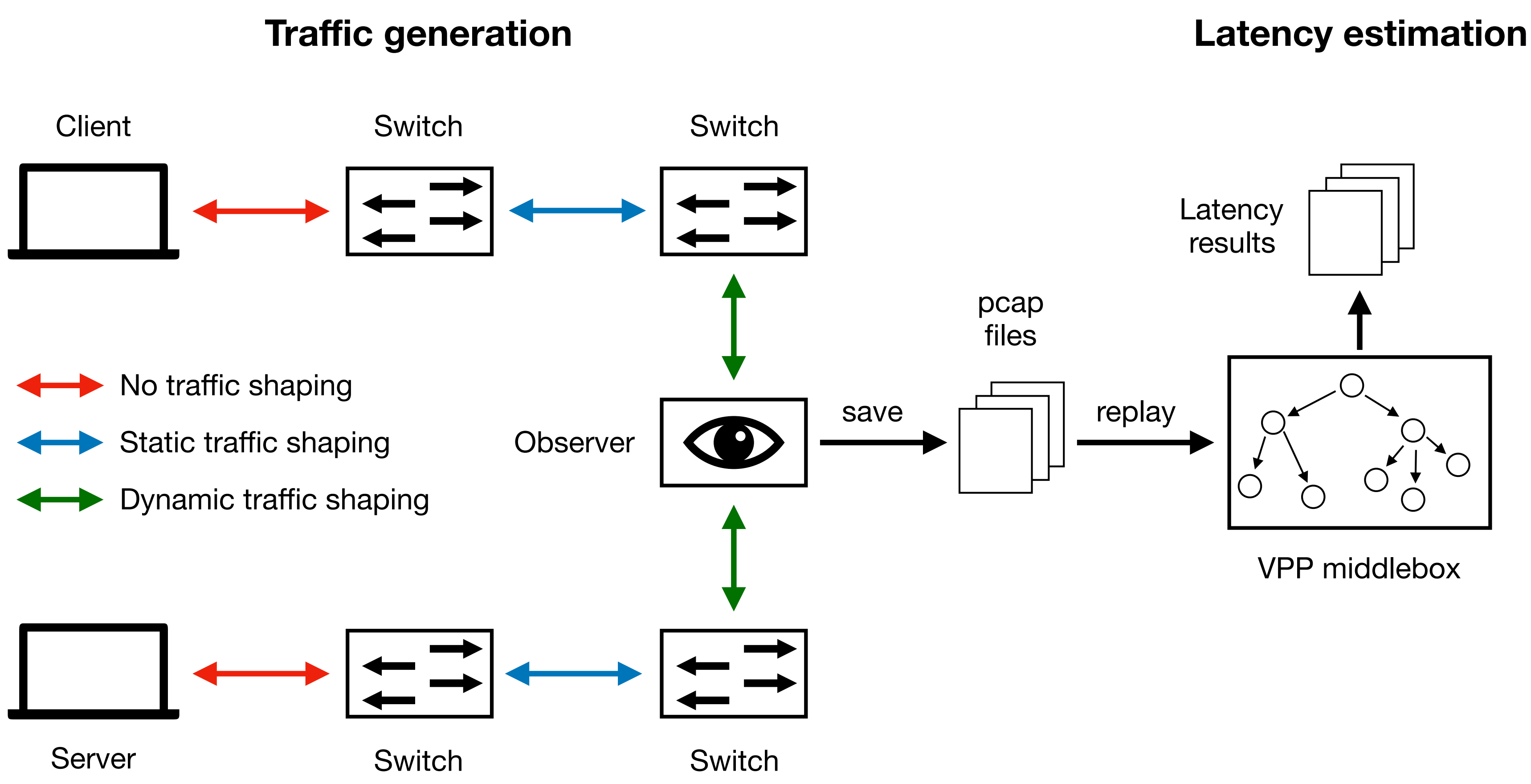
(IP and TCP protocol header fields)

(values)

(conditions: implicitly ‘==’)

Passive latency measurement approaches

- QUIC experiments: controlled lab setup (emulated network)



DPLPMTUD

Probes Received	Time to complete (s)	MPS found	Configured link size	PTB Status
273	14.3	1472	1500	Not Enabled
265	15.3	1464	1492	Not Enabled
253	14.7	1452	1480	Not Enabled
233	13.7	1432	1460	Not Enabled
227	13.3	1426	1454	Not Enabled
145	9.0	1344	1372	Not Enabled
273	14.3	1472	1500	Enabled
265	14.0	1464	1492	PTB = plpmtu
253	13.4	1452	1480	PTB = plpmtu
233	12.3	1432	1460	PTB = plpmtu
227	12.0	1426	1454	PTB = plpmtu
145	7.7	1344	1372	PTB = plpmtu

Table 1: Results of tests of the DPLPMTU with the Step Size search algorithm with a step size of 1.

Probes Received	Time to complete (s)	MPS found	Configured link size	PTB Status
6	0.3	1450	1500	Not Enabled
6	0.3	1450	1492	Not Enabled
6	0.3	1450	1480	Not Enabled
5	1.7	1400	1460	Not Enabled
5	1.7	1400	1454	Not Enabled
3	1.6	1300	1372	Not Enabled
6	0.3	1450	1500	Enabled
6	0.3	1450	1492	Enabled
6	0.3	1450	1480	Enabled
6	0.3	1432	1460	PTB > plpmtu
6	0.3	1426	1454	PTB > plpmtu
4	0.2	1344	1372	PTB > plpmtu

Table 2: Results of tests of the DPLPMTU with the Step Size search algorithm with a step size of 50.

DPLPMTUD

Probes Received	Time to complete (s)	MPS found	Configured link size	PTB Status
7	0.3	1472	1500	Not Enabled
6	1.8	1464	1492	Not Enabled
5	1.7	1412	1480	Not Enabled
5	1.7	1412	1460	Not Enabled
5	1.7	1412	1454	Not Enabled
2	1.6	1332	1372	Not Enabled
7	0.3	1472	1500	Enabled
6	0.4	1464	1492	PTB = plpmtu
6	0.4	1452	1480	PTB > plpmtu
6	0.3	1432	1460	PTB > plpmtu
6	0.3	1426	1454	PTB > plpmtu
3	0.2	1344	1372	PTB > plpmtu

Table 3: Results of tests of the DPLPMTU with the Table Derived search algorithm.

Probes Received	Time to complete (s)	MPS found	Configured link size	PTB Status
1	0.05	1472	1500	Not Enabled
9	83.3	1464	1492	Not Enabled
9	3.3	1452	1480	Not Enabled
9	3.3	1432	1460	Not Enabled
9	4.8	1426	1454	Not Enabled
9	167.4	1344	1372	Not Enabled
1	0.5	1472	1500	Enabled
2	0.3	1464	1492	PTB > plpmtu
2	0.2	1452	1480	PTB > plpmtu
2	0.2	1432	1460	PTB > plpmtu
2	0.2	1426	1454	PTB > plpmtu
2	0.1	1344	1372	PTB > plpmtu

Table 4: Results of tests of the DPLPMTU with the Binary Search algorithm.

DPLPMTUD

Search Type	Probes Received	Time to complete (s)	MPS found	Estimated RTT (ms)
Step Size 1	273	24.2	1472	88.2
Step Size 50	6	0.5	1450	88.2
Derived Table	7	0.7	1472	88.2
Binary Search	1	0.07	1472	88.2

Table 5: Results of tests of the DPLPMTU test tool in a real world environment with each of the four search algorithms. The test network did not generate any ICMP PTB messages.