WP3: Middlebox Cooperation

Gorry Fairhurst WP3 Lead

2nd Technical review

3rd October 2017

Research and Innovation Action **688421**

Call: H2020-ICT-2015: Integrating experiments and facilities in FIRE+

mami

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..

Objectives from DoW

- Definition of **use cases and requirements** for an architecture for Middlebox Cooperation Protocol (MCP) - D3.1
- **Design, implementation, and initial testing** of the **MCP** to provide an information exchange between end hosts and middleboxes
- Design of a **flexible transport stack (FTL)** to complement the MCP, restoring connectivity over the Internet
- **Threat and trust analysis** of the developed protocols, protocol extensions and transport layer mechanisms as a basis for Internet-scale deployment

WP3 Tasks Overview

T3.1: Use Case Analysis and Requirement Definition (M1 - M6)

T3.2: Design of the MCP (M7 - M24)

T3.3: Design of a flexible cooperative transport layer (M7 - M36)

T3.4: Implementation and Testing (M9 - M36)

T3.5: Threat and Trust Analysis for Middlebox Cooperation (M1 - M36)

Overview - Who did what?

Partner	Task 3.1 Use Cases	Task 3.2 MCP Design	Task 3.3 FTL Design	Task 3.4 Implementation and Testing	Task 3.5 Threat and Trust Analysis
ETH	√	√	√	√	
TID	√	√			√
UoA		√	√		
ZHAW	√			√	√
ALU (Nokia)	√		√	√	√

WP3 Objectives

- Definition of **use cases and requirements** for an architecture for Middlebox Cooperation Protocol (MCP)
- **Design, implementation, and initial testing** of the **MCP** to provide an information exchange between end hosts and middleboxes
- Design of a **flexible transport stack (FTL)** to complement the MCP, restoring connectivity over the Internet
- **Threat and trust analysis** of the developed protocols, protocol extensions and transport layer mechanisms as a basis for Internet-scale deployment

WP3 Deliverables Summary

D3.1	Use Case Analysis and Requirements
D3.2	Design of the MCP
D3.3	Design of a flexible cooperative transport layer

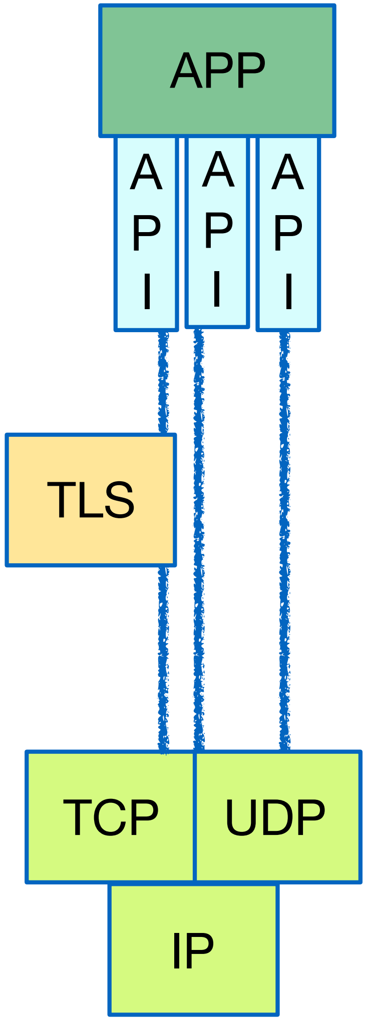
WP3 Activities in the reporting period

- Endpoint and **Middlebox Cooperation Protocol** (MCP)
  - Implementation of protocols and protocol extensions
- A New **Transport API**
- **Flexible Transport Layer** (FTL)
  - Implementation of protocols and protocol extensions
- **Security Analysis** and **Manageability**

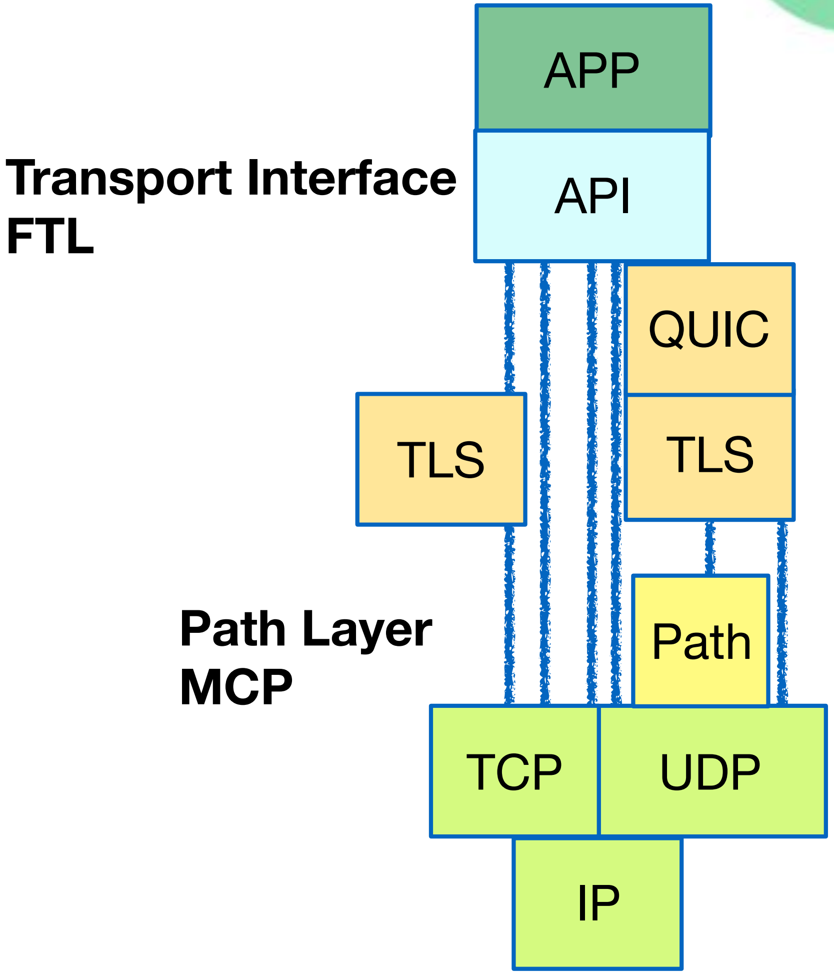
Middlebox Cooperation Concept

- An endpoint should be able to explicitly expose any signals used by on-path devices.
- An endpoint should be able to request signals from devices on the path.
- An on-path device should not be able to forge, change, or remove a signal sent by an endpoint.
- The endpoint should control signaling.
- It should be possible for an endpoint to request and receive signals from a pre- viously unknown on-path device.
- There should be no significant surface for amplification attacks.

Internet Protocol Stack



IP Stack at start of project



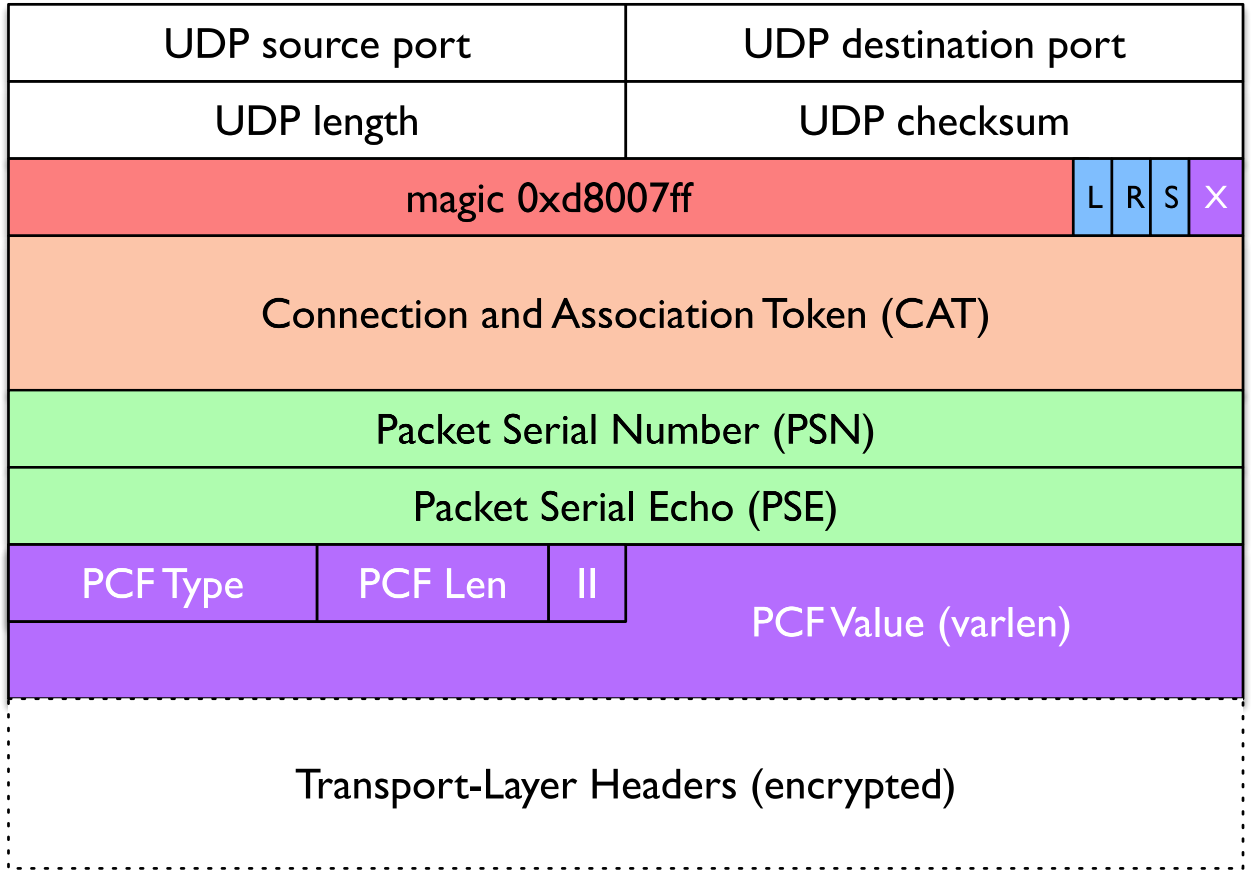
Transport Interface
FTL

Path Layer
MCP

IP Stack at end of project

MCP: Path Layer UDP Substrate (PLUS)

D3.2 includes a consistent spec for middlebox cooperation



UDP source port | UDP destination port

UDP length | UDP checksum

magic 0xd8007ff | L | R | S | X

Connection and Association Token (CAT)

Packet Serial Number (PSN)

Packet Serial Echo (PSE)

PCF Type | PCF Len | II | PCF Value (varlen)

Transport-Layer Headers (encrypted)

PLUS in IETF

Initial specification contributed to IETF PLUS design:

draft-trammell-spud-req *(Expired)*

draft-trammell-plus-abstract-mech *(Expired)*

draft-trammell-plus-statefulness *(Expired)*

draft-trammell-plus-spec *(Expired)*

PLUS work stagnated in the IETF

Concerns that a generic metadata exposure protocol could be used to force metadata injection on endpoints

We do not expect deployment of PLUS as specified in D3.2

QUIC in IETF

Google proposed a new protocol web transport (**QUIC**)

- Work adopted as an IETF activity in 2017
- All energy in transport/web space going into QUIC, which will actually deploy at scale in the near term (2019)

PLUS and QUIC in MAMI

MAMI adopted a broader focus on middlebox cooperation

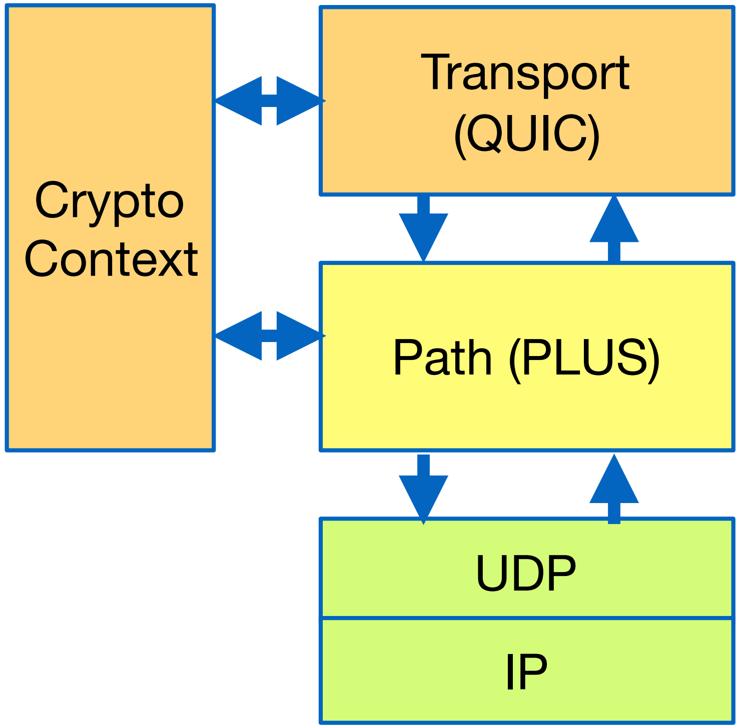
MAMI has shown concepts can be applied to other protocols

- Mechanisms using UDP
- Exploring mechanisms with QUIC

IETF applicability and manageability documents for QUIC

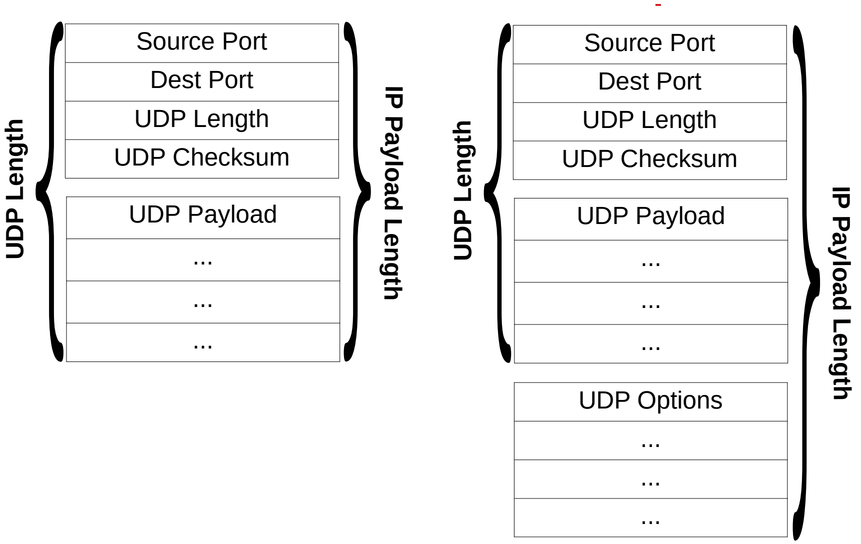
- *draft-ietf-quic-manageability (Expected to be published 2019)*
- *draft-ietf-quic-applicability (Expected to be published 2019)*
- *draft-trammell-quic-spin (Consensus to be incorporated in QUIC)*
- *draft-iab-wire-image (Approved, to be published 2019)*
- *draft-trammell-privsec-defeating-tcpip-meta (Expired)*

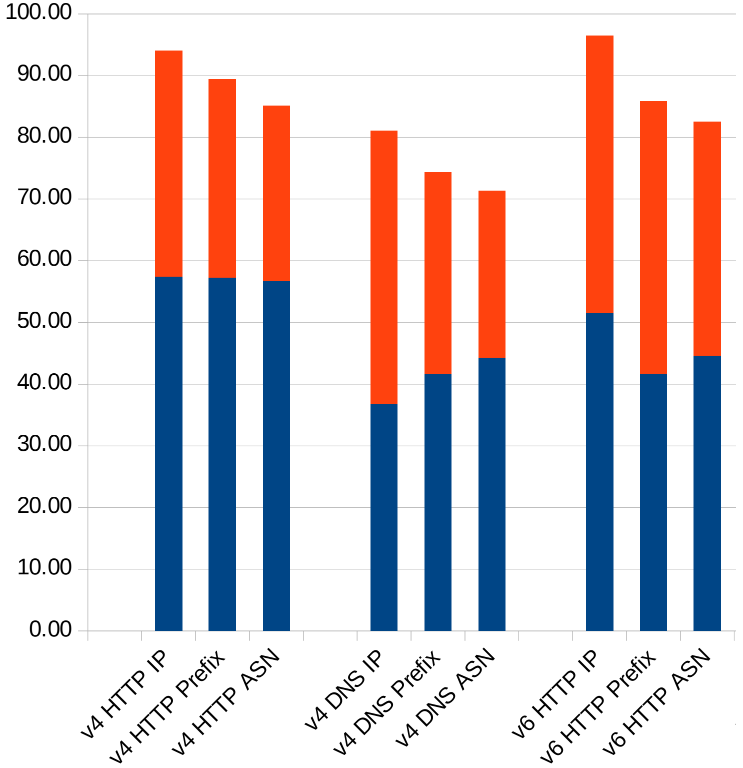
PLUS Reference Implementation using ~~QUIC~~ GUIC



- QUIC spec expected Nov 2019
- Google's QUIC (GUIC) was the best we had for experimentation
- MAMI completed a software implementation in fd.io
- MAMI fd.io testbed built
- Experimentation in WP2

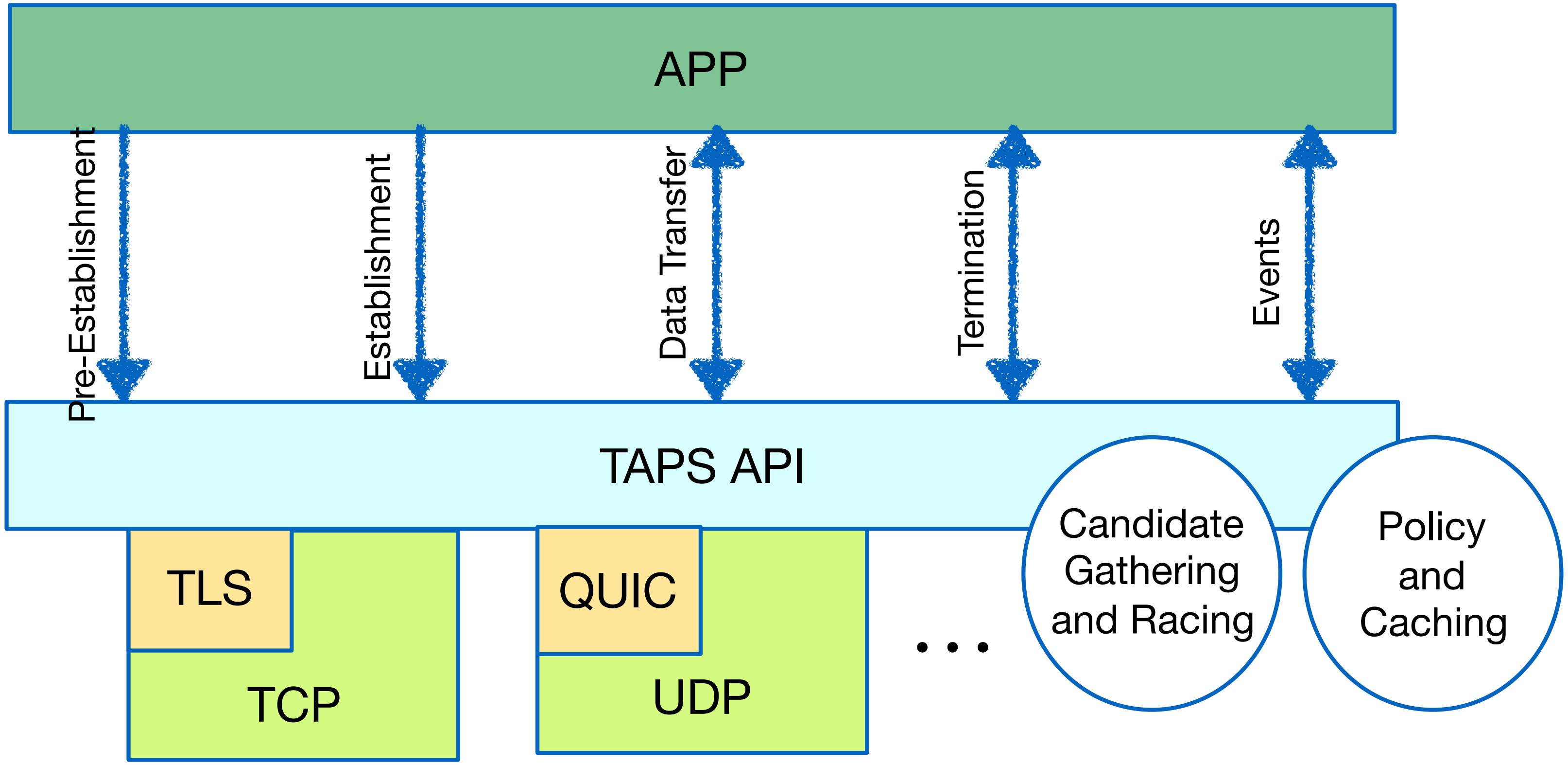
Transport Stack with UDP Options





- UDP-O provides a way to add meta-information to UDP flows
- draft-fairhurst-udp-options-cco
- Open source reference implementation on the FreeBSD

Standards-based Abstract Interface for Transport Services (TAPS)



Definition of ***unified*** (abstract) API independent of protocol

Fallback and connection racing mechanisms

API Specifications: MAMI Documents

Inputs: *Post Sockets (see D3.2; Other IETF Participants: EU NEAT Project; Apple; TU Berlin)*

API/transport state-of-the-art

- IETF Transport Services *(Published as RFC 8095)*
- *draft-ietf-taps-transports-usage-udp* *(Published as RFC 8304)*

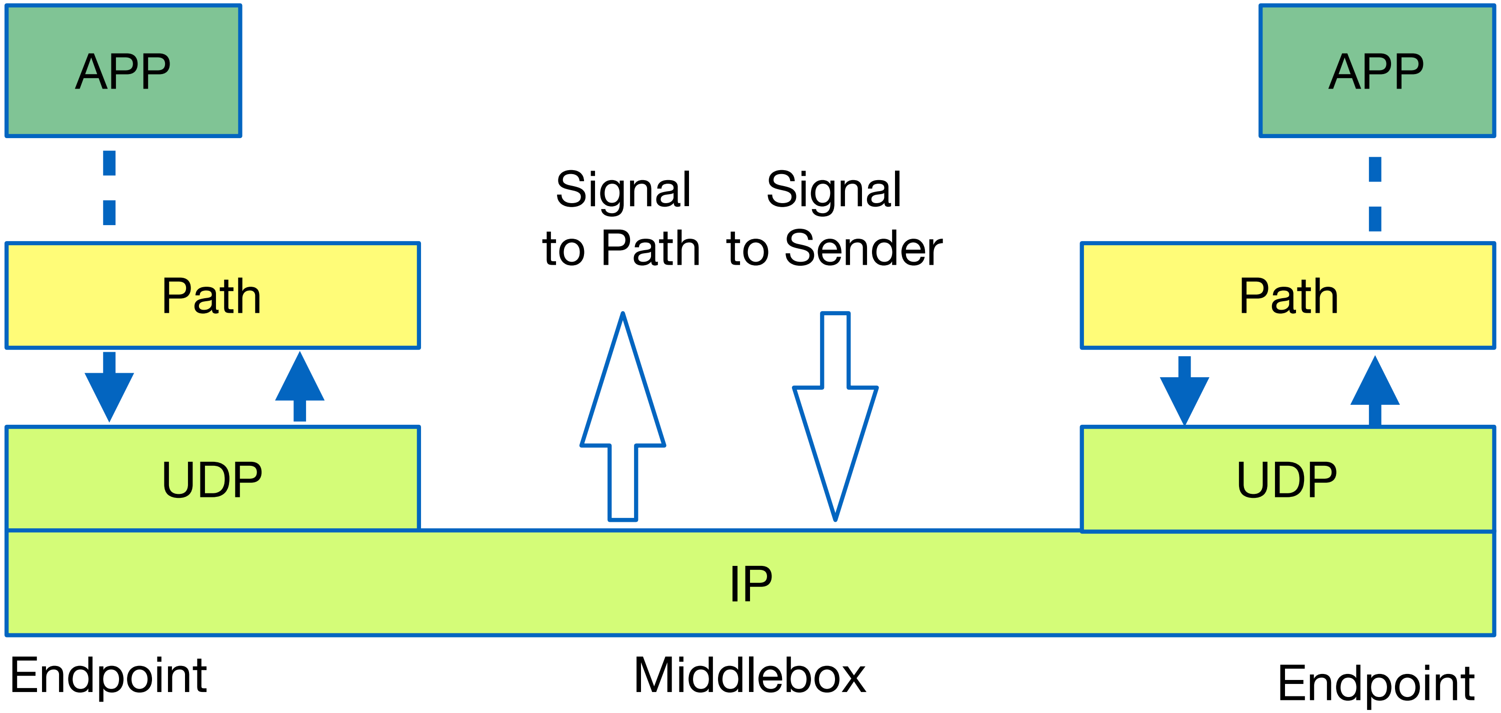
API/transport evolution contributions

- draft-kuehlewind-taps-crypto-sep *(Contribution to WG)*
- draft-trammell-taps-post-sockets *(Contribution to WG)*

API/transport evolution work items

- draft-ietf-taps-arch *(Expected to be published 2019)*
- draft-ietf-taps-impl *(Expected to be published 2019)*
- draft-ietf-taps-interface *(Expected to be published 2019)*

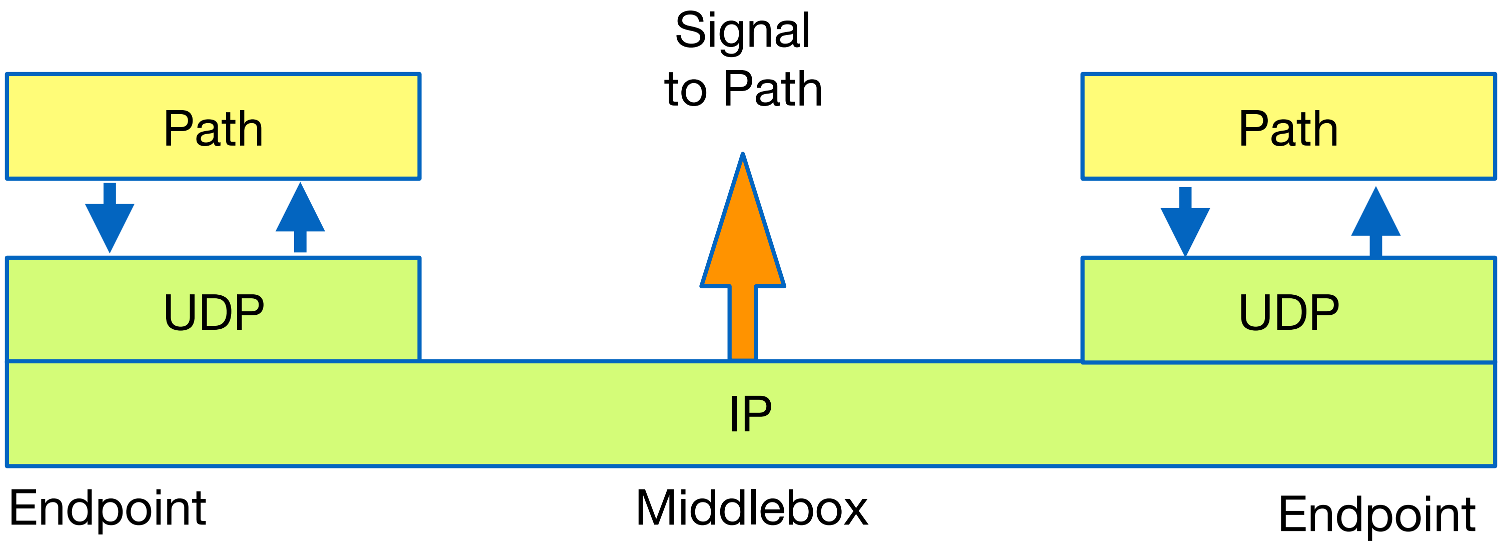
A Flexible Transport Layer (FTL)



Transport / Network Signaling Mechanisms

- Explicit Sender-to-Path Signaling
- Explicit Path-to-Sender Signaling

Transport / Network Signaling Mechanisms: Explicit Sender-to-Path Signaling



- Differentiated Services Code Point (DSCP) transparency
- LoLa Signaling Mechanism
- Short-Term, Automatically-Renewed (STAR) Certificates
- Explicit Support for Passive Latency Measurement

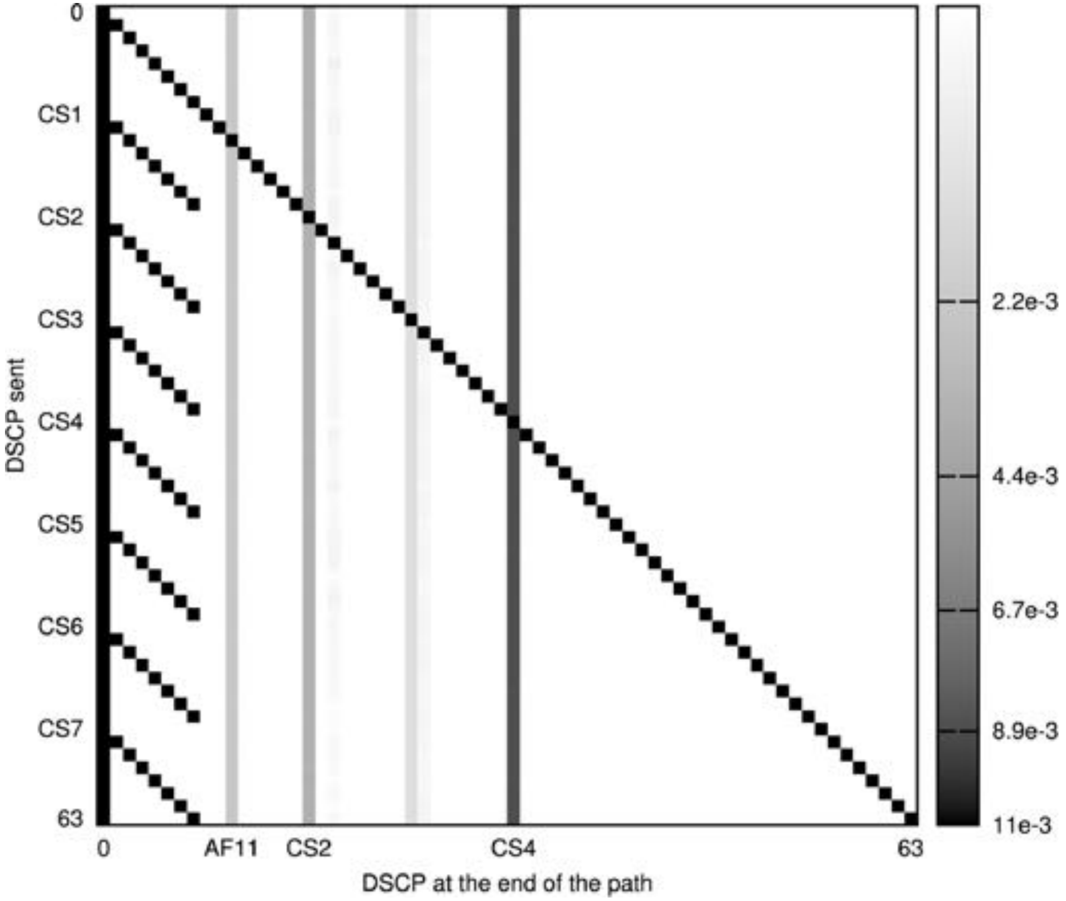
Explicit Sender-to-Path Signaling: Differentiated Services Code Point (DSCP)

Signals treatment to Path

Can DSCP's be transparent?
Measured in mobile and wired.

Lower-Effort PHB needed one…

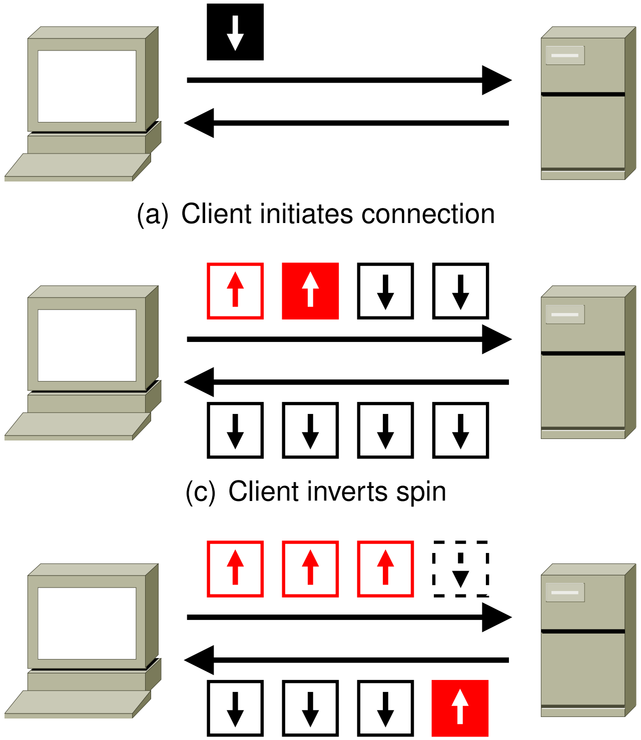
and we helped IETF find 0X01 !



Explicit Sender-to-Path Signaling: LoLa Signaling Mechanism

- Low Latency Low Loss (LoLa) Tradeoff
- Marks low-latency flows (e.g. voice / video, gaming, m2m)
- Mobile network can match to a suitable PHB (EPS Bearer)
- GSMA Technical Report
- draft-fossati-tsvwg-lola (contribution to TSVWG)

Explicit Sender-to-Path Signaling: Passive Latency Measurement in QUIC



- Extensions to QUIC wire image to support measurability
- In-network support for supporting network operations
- In-network support for managing low latency

Is RTT exposure to the path a threat to geoprivacy?

No.

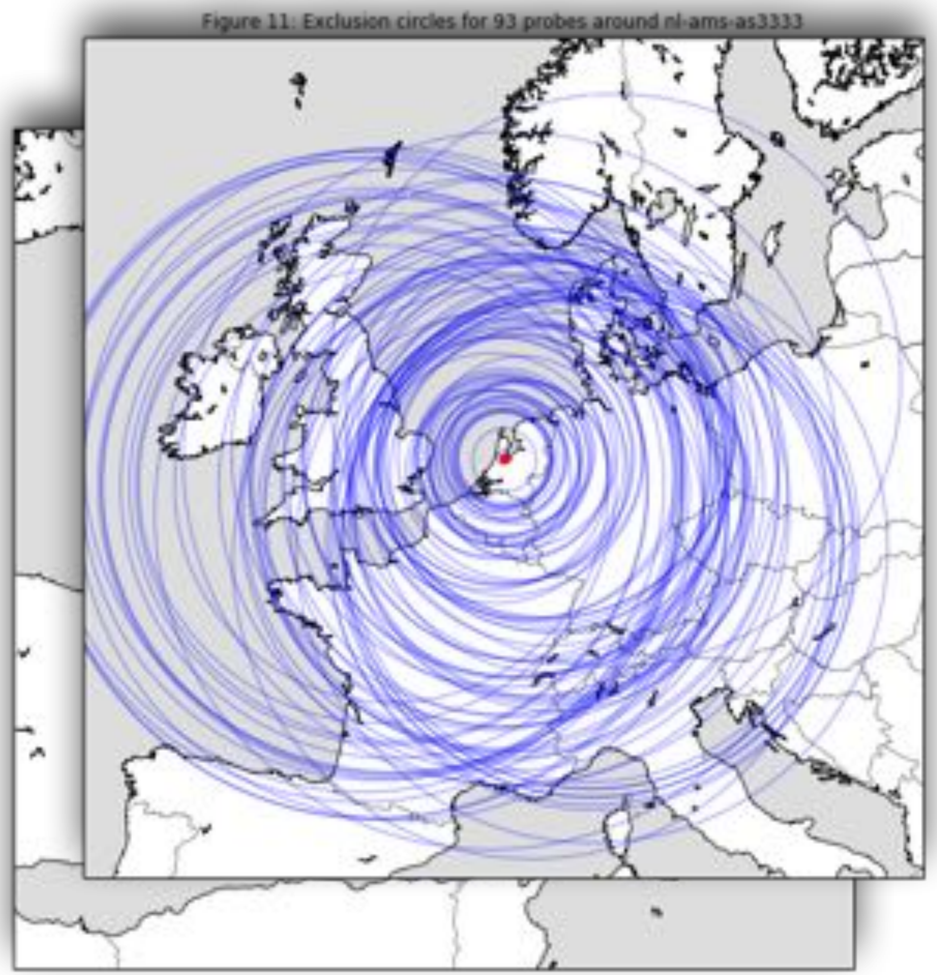


Figure 11: Exclusion circles for 93 probes around nl-ams-as3333

- min(rtt) from Atlas anchoring measurements, fiber lightspeed assumption

SPIN in QUIC

IETF Contributions

draft-trammell-privsec-defeating-tcpip-meta *(Expired)*

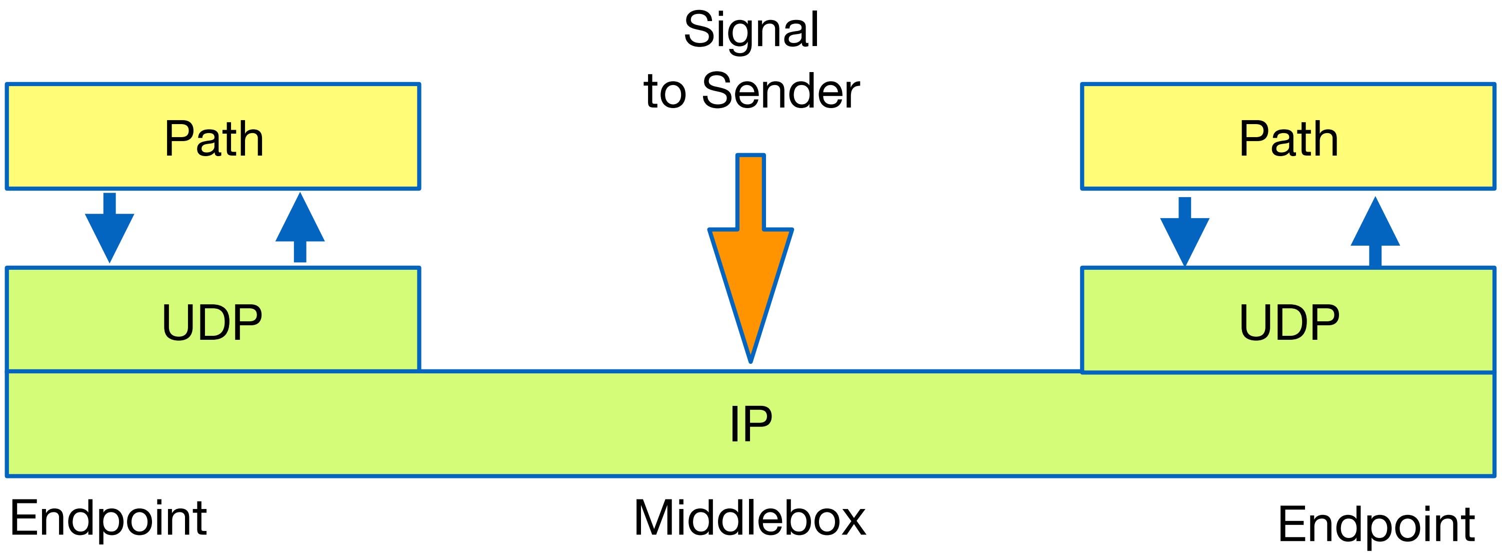
RTT exposure privacy analysis to QUIC RTT design team: github.com/britram/trilateration

draft-trammell-quic-spin-03 (see below)

draft-trammell-ippm-spin *(active)*

Simple extension to QUIC adopted for QUIC v 1

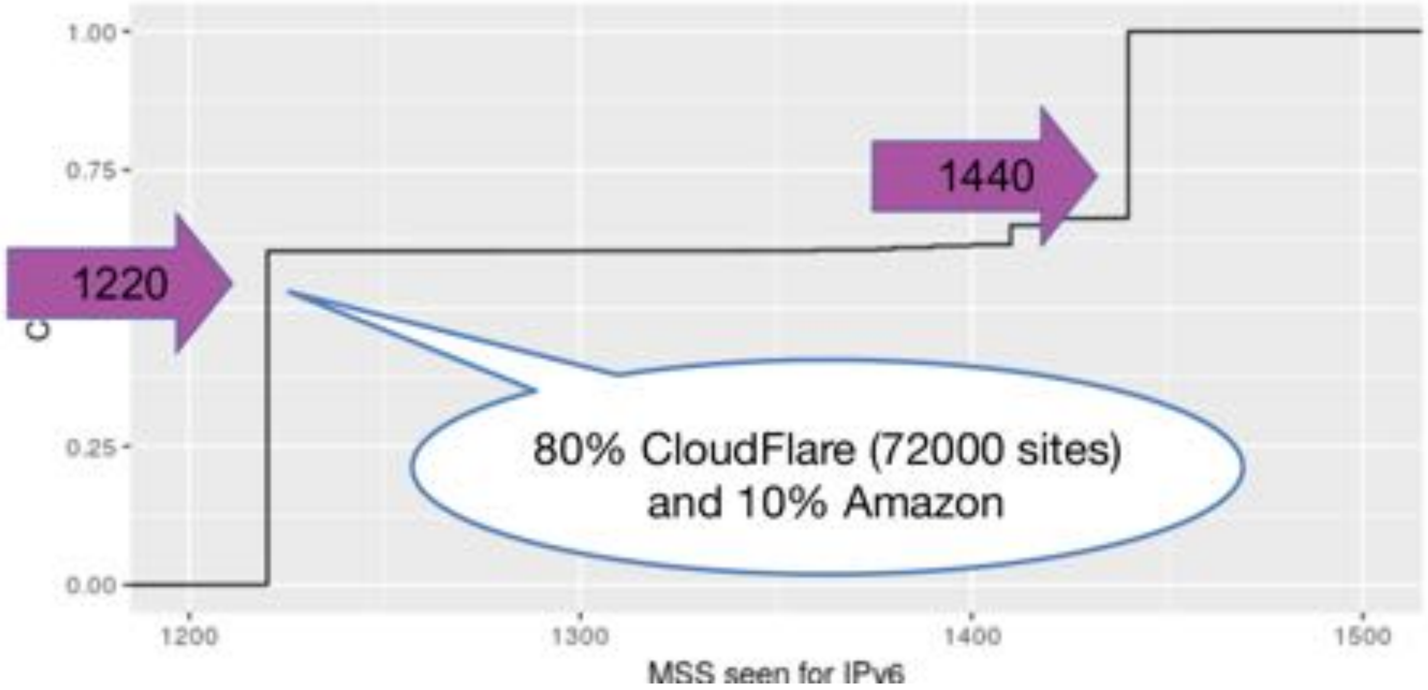
Transport / Network Signaling Mechanisms: Explicit Path-to-Sender Signaling



- TCP MSS Clamping
- The Datagram PLPMTUD Mechanism
- Explicit Congestion Signaling (ECN)
- Explicit Capacity Signals

Explicit Path-to-Sender Signaling: TCP MSS Clamping

- PMTUD has real deployment problems
- MSS being used by operators as signal (was unintended)
- WP1 found and reported issues for TCP



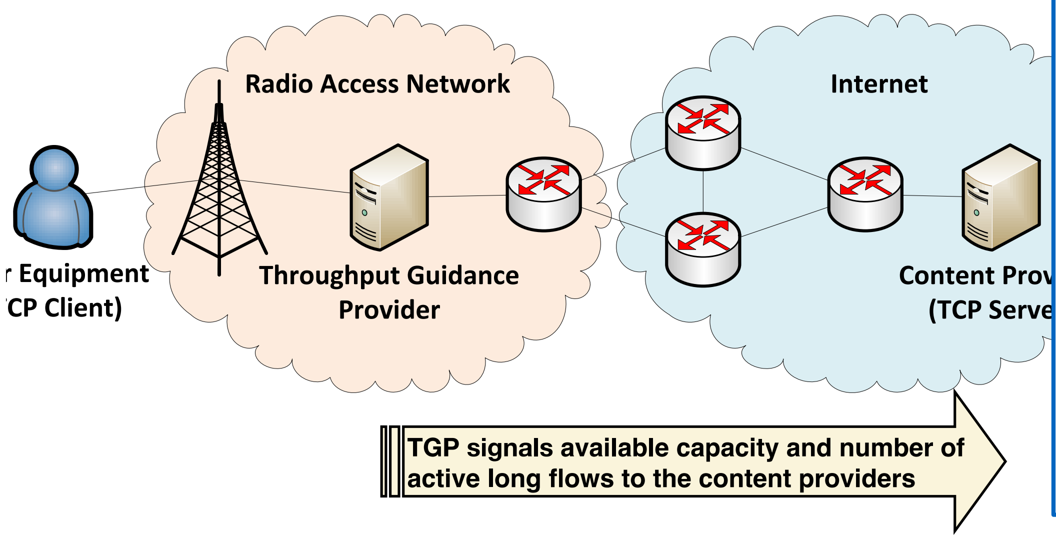
Explicit Path-to-Sender Signaling: Datagram PLPMTUD

- Adds PMTU discovery to datagram protocols
- Methods for SCTP, UDP-Apps, UDP-Options, QUIC, etc
- draft-ietf-tsvwg-datagram-plpmtud (expected to be published in 2019)
- Open Source for BSD

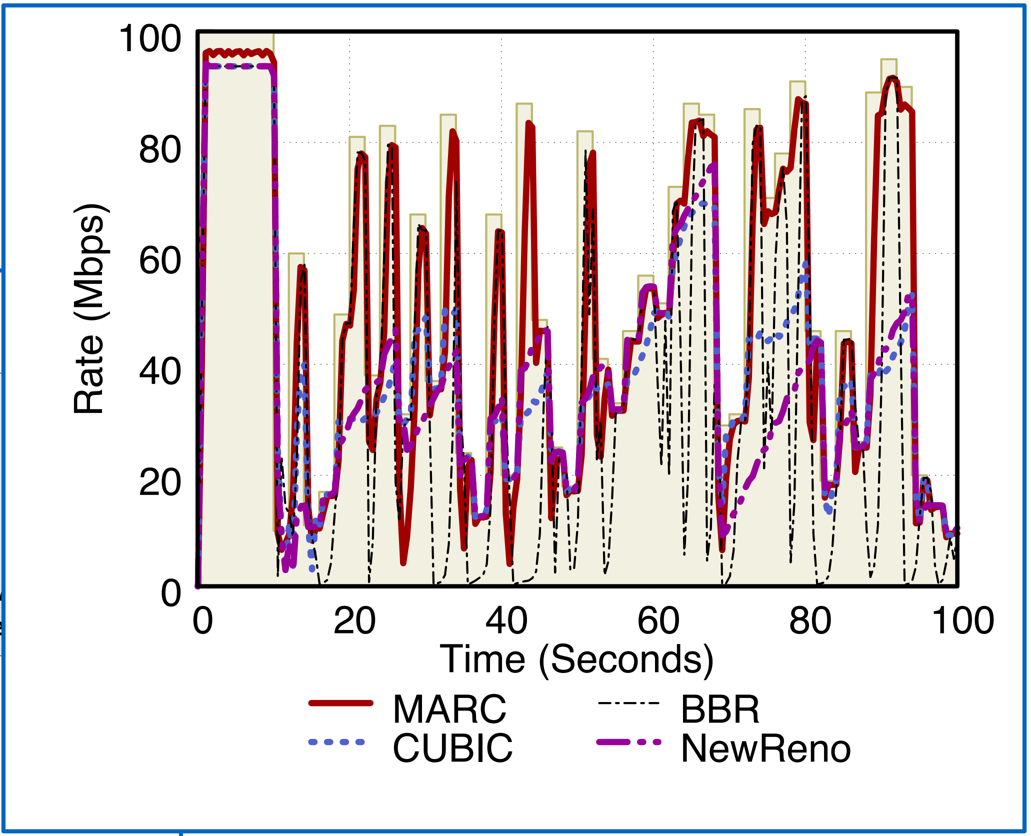
Explicit Path-to-Sender Signaling:
Explicit Capacity Signals (MARC)

Throughput Guidance signal from cellular to endpoint

Significant benefit to user



TGP signals available capacity and number of active long flows to the content providers



Threat and Trust Analysis & Manageability

- Security and Privacy Analysis for MCP (in D3.2)
- Workshops and dissemination

MAMI Management and Measurement Summit (M3S)

Invitation-only Industry workshop

Concrete examples of what is done today

Friday, March 16, 2018 in London

(https://mami-project.eu/index.php/events/mami-management-and-measurement-summit-m3s/)

MAMI Outputs

- 3 White papers (public access):
  - Challenges in Network Management with Encrypted Traffic Transport Encryption) (based on M3S)
  - Analysis and Consideration on Management of Encrypted Traffic
  - Security and Privacy Implications of Middlebox Cooperation Protocols
- IETF Informational Document
  - The Impact of Transport Header Confidentiality on Network Operation and Evolution of the Internet (draft-ietf-tsvwg-transport-expected to be published 2019)
  -

Summary of WP3 Achievements

- Story about PLUS - completed MCP Spec
-

- This slide to be replaced

Summary of WP3 Achievements

- Possible pretty drawing showing “measurement-based-design” producing real output to change architecture

- Dissemination to Industry and Academe
  - XX numbers XX Contributions

- WP3 has directly impacted standardization organizations
  - XX numbers XX Contributions
  - Efforts to continue beyond end of project

-
- This slide to be replaced

Related WP3 scientific publications during the reporting period

Papers

1. Neuhaus,, Mirja Kühlewind, Tobias Bühler, Brian Trammell, Roman Müntener, Stephan; Fairhurst, Gorry *A Path Layer for the Internet: Enabling Network Operations on Encrypted Protocols* International Conference on Network and Service Management (CNSM), IEEE, 2017.
2. B. Trammell, C. Perkins, and M. Kühlewind. *Post sockets: Toward an evolvable network transport interface*. Networking Workshop on Future Internet Transport, Stockholm, Sweden, June 2017.
3. Cui, Y.; Li, T.; Liu, C.; Wang, X.; Kühlewind, M. *Innovating Transport with QUIC: Design Approaches and Research Challenges* Journal Article IEEE Internet Computing, 21 (2), pp. 72-76, 2017, ISSN: 1089-7801.
4. A. Custura, G. Fairhurst, and I. Learmonth. *Exploring usable Path MTU in the Internet.* Network Traffic Measurement and Analysis Conference (TMA), 2018.
5. A. Custura, R. Secchi, and G. Fairhurst. *Exploring DSCP modification pathologies in the internet.* Computer Communications, 127:86–94, 9 2018.
6. M. Kühlewind, T. Bühler, B. Trammell, R. Müntener, S. Neuhaus, and G. Fairhurst. *A Path Layer for the Internet: Enabling Network Operations on Encrypted Protocols*. International Conference on Network and Service Management (CNSM). IEEE, 2017
7. P. D. Vaere, T. Bühler, M. Kühlewind, and B. Trammell. *Three bits suffice: Explicit support for passive measurement of internet latency in QUIC and TCP,* Internet Measurement Conference (IMC), 2018.

Reports

1. T. Fossati. *Content classification*. Technical Report Document No IG.01, rev 1.0, GSM Association (GSMA), 2018.
2. T. Fossati, R. Müntener, S. Neuhaus, and B. Trammell. *Security and privacy implications of middlebox cooperation protocols*. cs.NI arXiv:1812.05437 ETH TIK Technical Report 370, 2018.
3. A. Aranda, D. Lopez, and T. Fossati. *Analysis and consideration on management of encrypted* traffic. cs.NI arxiv:1812.04834, 2018.
4. Fossati, Thomas; Muentener, Roman; Neuhaus, Stephan; Trammell, Brian, Security and Privacy Implications of Middlebox Cooperation Protocols Technical Report, cs.NI, (arXiv:1812.05437), 2018, (ETH TIK Technical Report 370).
5. Aranda, Pedro A.; López, Diego; Fossati, Thomas *Analysis and Consideration on Management of Encrypted Traffic* Technical Report cs.NI, (arxiv:1812.04834), 2018.
6. Edeline, Korian; Kühlewind, Mirja; Trammell, Brian; Donnet, Emile Aben and Benoit *Using UDP for Internet Transport Evolution* Technical Report cs.NI, (arXiv:1612.07816), 2016, (ETH TIK Technical Report 366).

Presentations & Posters

1. G. Fairhurst, M. Khlewind, and D. R. Lopez. *Measurement-based protocol design*. European Conference on Networks and Communications (EuCNC), 2017.
2. Kühlewind, Mirja; Trammell, Brian; Brunstrom, Anna; Welzl, Micheal; Fairhurst, Gorry *TAPS: an abstract application interface for QUIC Presentation* 04.12.2018, (Poster at ACM CoNEXT 2018 Workshop on the Evolution, Performance, and Interoperability of QUIC (EPIQ'18)).
3. Bühler, Tobias; Kühlewind, Mirja; Trammell, Brian *Enhancing encrypted transport protocols with passive measurement capabilities* Presentation (Poster at IMC), 2017.
4. Lopez, Diego R. *Path-Aware Networking Concept* Presentation, 2018.
5. R. Secchi, A. Venné, and A. Custura. *Measurements concerning the DSCP for a LE PHB,*. IETF 99, 2017.
6. Fossati, Thomas, *1-bit Content Classification* Presentation, 2018.
7. Kühlewind, Mirja *State of ECN and improving congestion feedback with AccECN in Linux* Presentation, 2017.
8. Fairhurst G, *Encypt?*, Presentation, Networkshop, 2017.
9. Kühlewind, Mirja *QUIC und HTTP/2 – neue Internet Protokolle* Presentation, 2017.
10. B. Trammell. O*n the suitability of RTT measurements for geolocation* https://github. com/britram/trilateration/blob/master/paper.ipynb, Aug. 2017.
11. G. Fairhurst, T. Jones, and R. Zullo. *A Tale of Two Checksums*, IETF 101, 2018.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 688421.

Q&A

mami

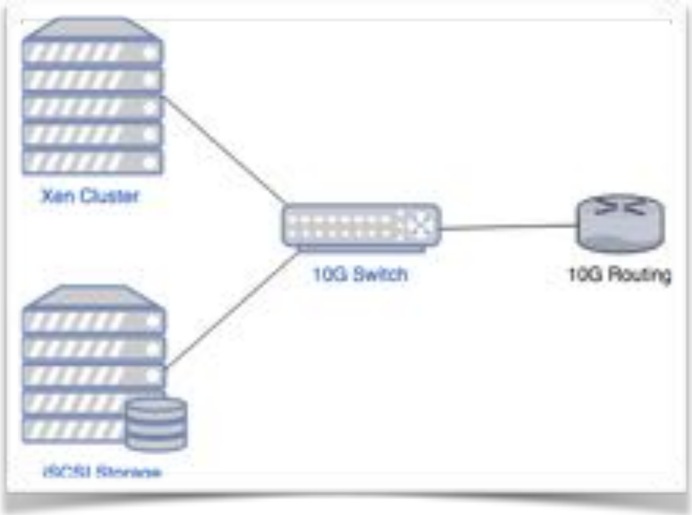
measurement and architecture for a middleboxed internet

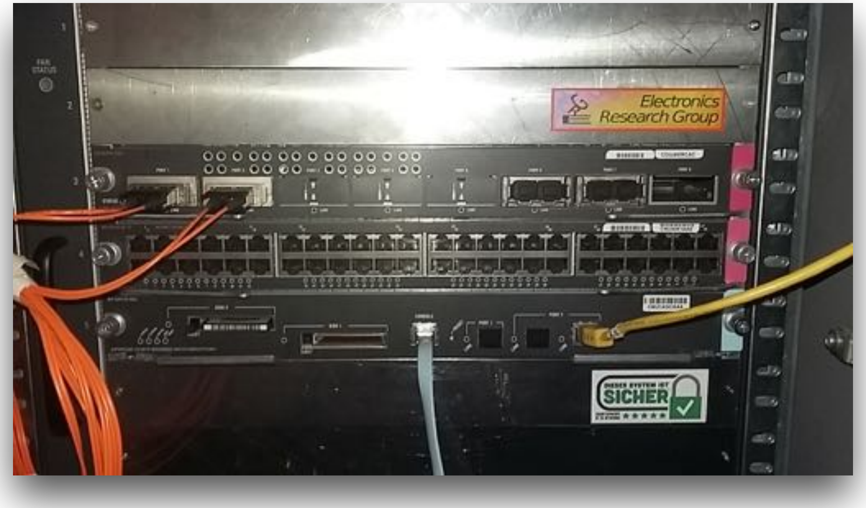
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 688421.

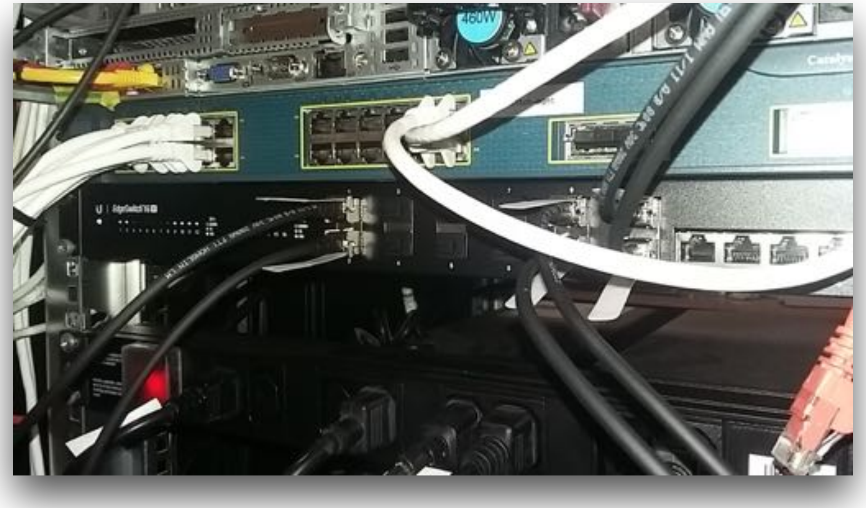
Spare Slides

measurement and architecture for a middleboxed internet

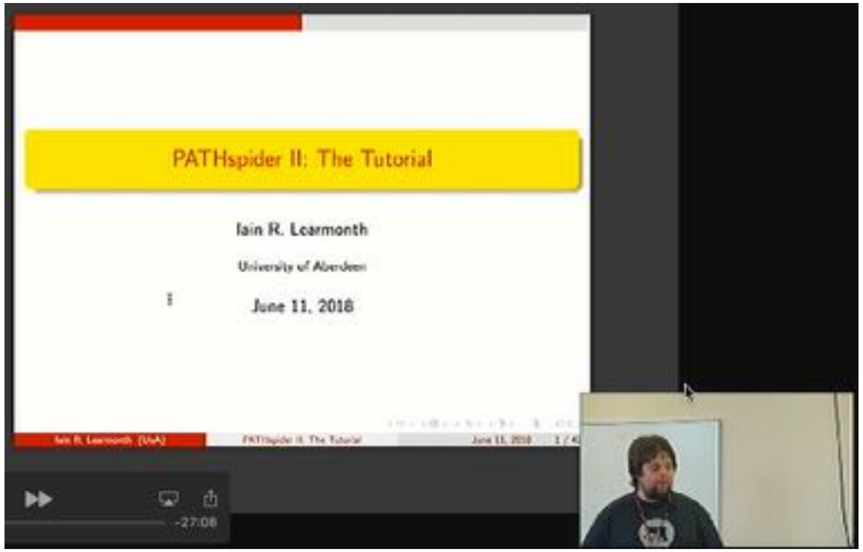
UoA MAMI Testbed Hardware







MAMI Summer School



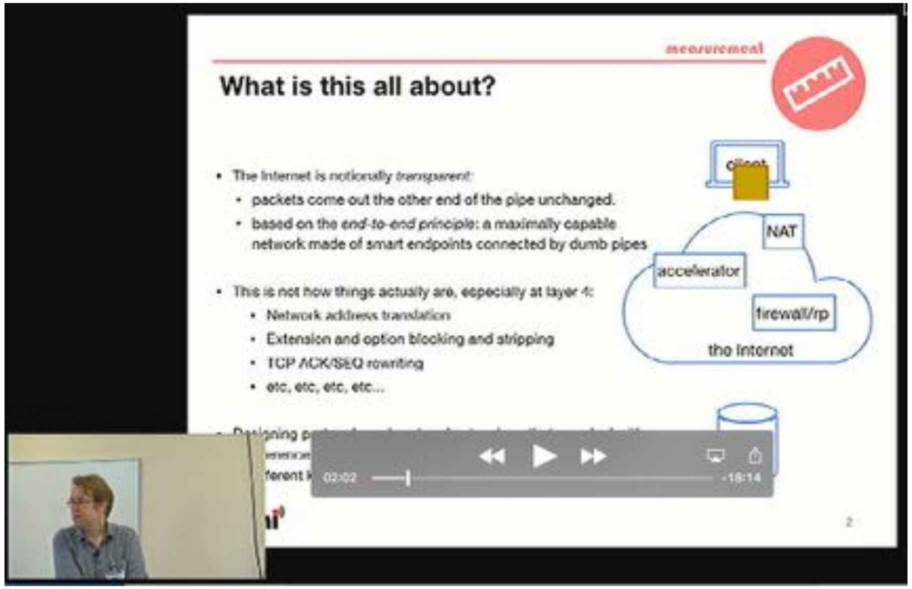
PATHspider II: The Tutorial

Iain R. Learmonth

University of Aberdeen

June 11, 2018

Iain R. Learmonth (UoA) PATHspider II: The Tutorial June 11, 2018 1 / 4

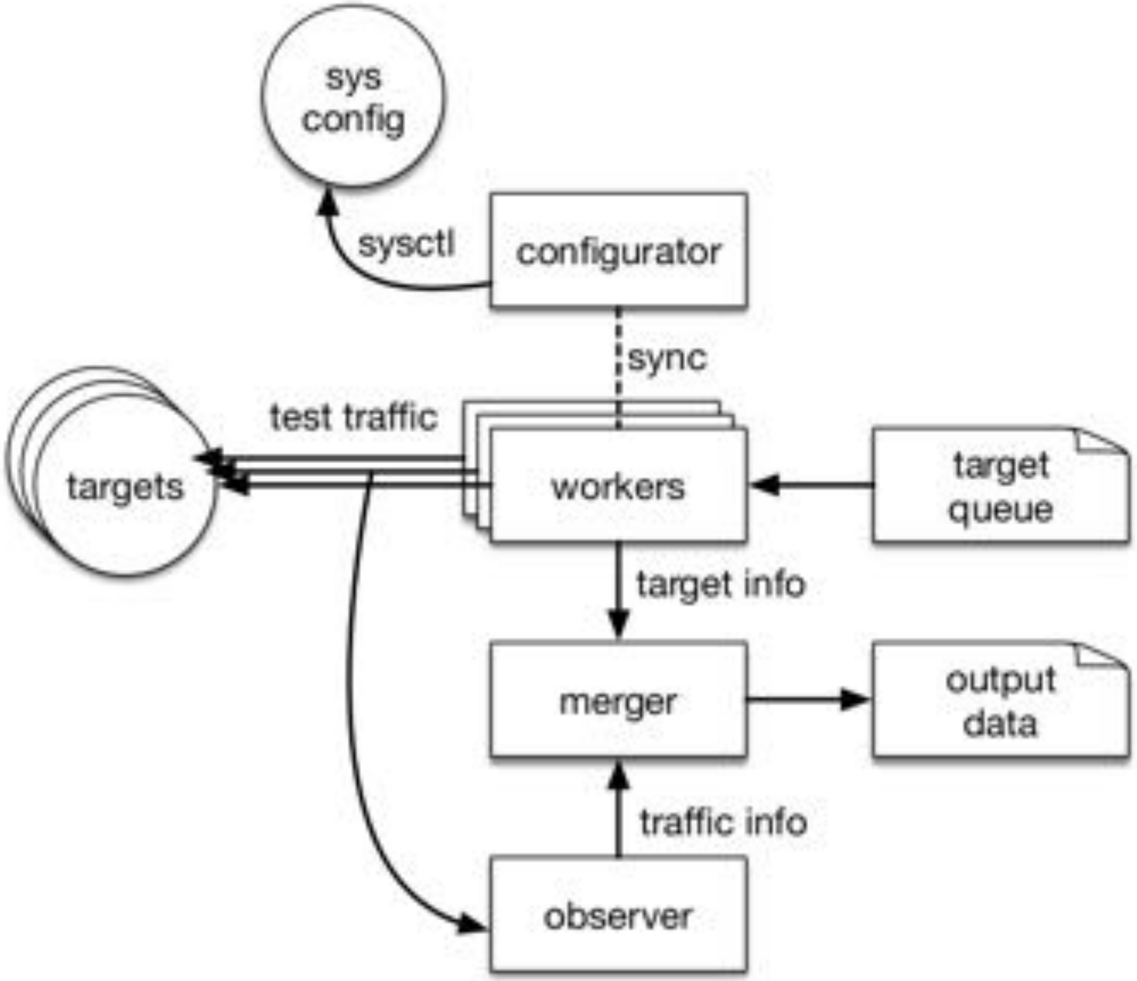


measurement

What is this all about?

- The Internet is notionally *transparent*:
  - packets come out the other end of the pipe unchanged.
  - based on the *end-to-end principle*: a maximally capable network made of smart endpoints connected by dumb pipes
- This is not how things actually are, especially at layer 4:
  - Network address translation
  - Extension and option blocking and stripping
  - TCP ACK/SEQ rewriting
  - etc, etc, etc, etc...
- Designing p

Pathspider


sys config
sysctl
configurator
sync
test traffic
targets
workers
target queue
target info
merger
output data
traffic info
observer

Explicit Path-to-Sender Signaling: Explicit Congestion Signaling

Issues own path transparency (WP1)

TCP feedback of congestion

QUIC: Contribution to QUIC design team

Explicit Path-to-Sender Signaling: Explicit Congestion Signaling

Issues own path transparency (WP1)

TCP feedback of congestion (AccECN)

Work on ECN and Manageability

QUIC: Contribution to QUIC design team