MP-DCCP progress

draft-ietf-tsvwg-multipath-dccp-04

Markus Amend on behalf of the authors, TSVWG @IETF113

LIFE IS FOR SHARING.

Main changes since IETF112 (-02 → -04)

Handshaking procedure finalized PR#40

MP_PRIO introduced for fine granular path management (enable/disable, backup mode, path prioritization) PR#42

Maximum Packet Size (i.a. PMTU) strategy **added** for multi-path specified, enhancing DCCP's mechanisms PR#68

Closing procedure first defined with **proper** MP_CLOSE (lack of description) and **instant** MP_FAST_CLOSE option PR#67, #65

Congestion Control considerations for bottleneck fairness with single-path transport added PR#54

Open Source code at GitHub updated with new handshaking procedure according to -02 and stability fixes to -03

More draft details at: https://github.com/markusa/ietf-multipath-dccp/releases

TSVWG Interim 01 (Feb 22) - Summary

NEW: Advanced MP_PRIO definition for fine-granular path management -03

```
 1                   2                   3
   0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
  +---------------+---------------+-------+-------+--------------+
  |     Type      |     Length    |Subtype| Prio  | AddrID       |
  +---------------+---------------+-------+-------+--------------+
```

The following values are available for Prio field:

* 0: Do not use. The path is not available.

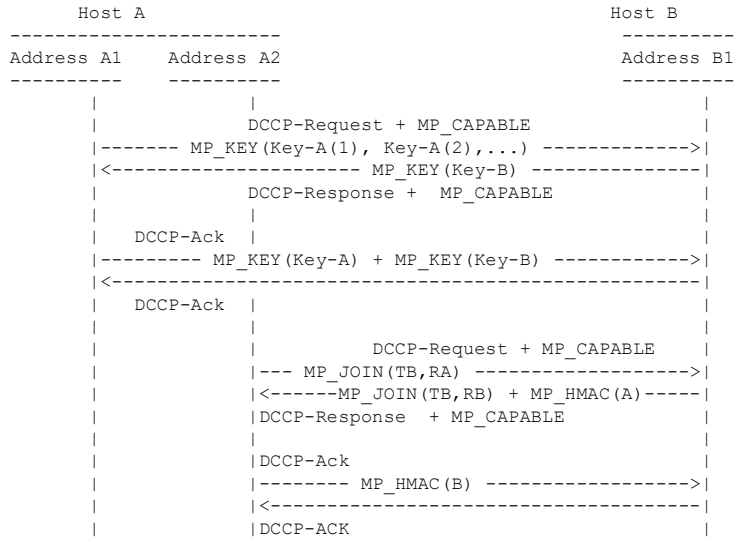
* 1: Standby: do not use this path for traffic scheduling, if another path (secondary or primary) is available.

* 2: Secondary: do not use this path for traffic scheduling, if the other paths are good enough. The path will be used occasionally, e.g. when primary paths are congested or become not available.

* 3: Primary: can use the path in any way deemed reasonable by peer. Will always be used for packet scheduling decisions.

* 4 - 15: relative priority of one path over the other to give relative path priority for primary paths. The peer should consider sending more traffic over higher priority path. Higher numbers indicate higher priority.

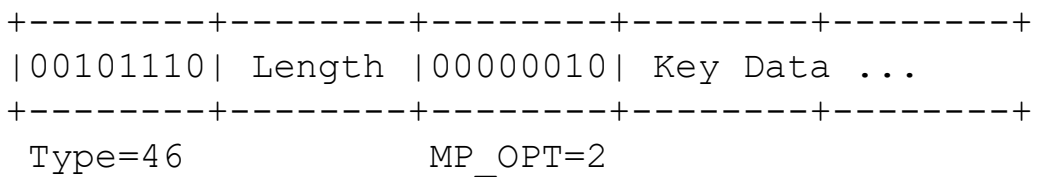
NEW: 4-way handshaking procedure resemble MPTCP logic -03



Early results of a partial P4 MP-DCCP implementation showing HW acceleration potential on SmartNICs

Authors presented relationship with 3GPP Rel. 18 roadmap

NEW: Closing procedure in -04



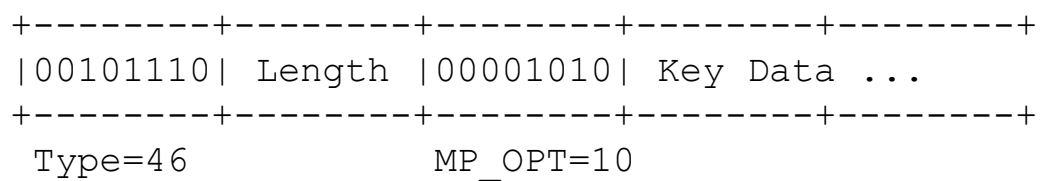
Carried on top of a *DCCP Reset* packet

MP_FASTCLOSE

Abrupt shutdown of a MP-DCCP connection, including all subflows, without awaiting confirmation from notified peer host

Key Data to protect misuse from non-eligible 3rd party.

New DCCP Reset Code 12 "Abrupt MP termination" defined



Carried on top of the regular DCCP closing packets: *DCCP CloseReq* or *DCCP Reset*

MP_CLOSE

Regular DCCP shutdown of all subflows first before MP connection is closed.

Key Data to protect misuse from non-eligible 3rd party.

Draft status – Feature set

Ready

Function/Mechanism	Draft	Open Source
Handshaking	✓	✓
MP Capable Feature	✓	✓
MP_KEY	✓	—
MP_SEQ	✓	✓
MP_HMAC	✓	✓
MP_RTT	✓	—

MP_KEY is implemented, but only “plain text “ type is supported.

MP_RTT implemented, but Type and Age is missing.

Partially ready

Function/Mechanism	Draft	Open Source
MP_CONFIRM	✓	✗
MP_JOIN	✓	—
Fallback mechanism	—	—
MP_FAST_CLOSE	✓	✗
MP_CLOSE	—	✗
MP_ADDADDR	✓	✗
MP_REMOVEADDR	✓	✗
MP_PRIO	✓	✗

Address ID is missing

New in -04

✓ Finalized, ready for review/testing

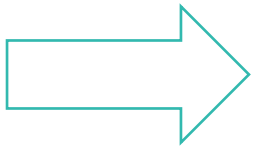
— Work on, contribution is welcome

✗ Not implemented, contribution is welcome

Draft work almost completed; Focus more on implementation now

Linux reference implementation - Status

MP-DCCP published prototype features
MP-DCCP
Encapsulation framework
Scheduling - Traffic distribution logics.
Compensate paths latency difference
As input for scheduling decisions Congestion Control CCID 2, 3, 5
(Re-)Establish/destruct flows



3GPP ATSSS requirements	
Multi-path transport	✓
Non-TCP support	✓
Steering modes	✓
Re-ordering	✓
Path measurement	✓
Path management	✓

Available for integration into Android and Linux based devices and ready for testing e2e or for ATSSS

General updates

MP-DCCP **so far only solution for non-TCP splitting support which made it into the TR after SA2#149e meeting**

- SA2 Technical Report: https://www.3gpp.org/ftp/Specs/archive/23_series/23.700-53/23700-53-010.zip

Full ATSSS compatible MP-DCCP conglomerate now published

- MP-DCCP
- Path management
- New: Scheduling schemes for all three S'
- New: Active re-ordering schemes
- New: Encapsulation framework to enable multi-path for any traffic (1)
- New: iPerf3 test tool with (MP-)DCCP support

https://github.com/telekom/mp-dccp

https://github.com/telekom/tunprox

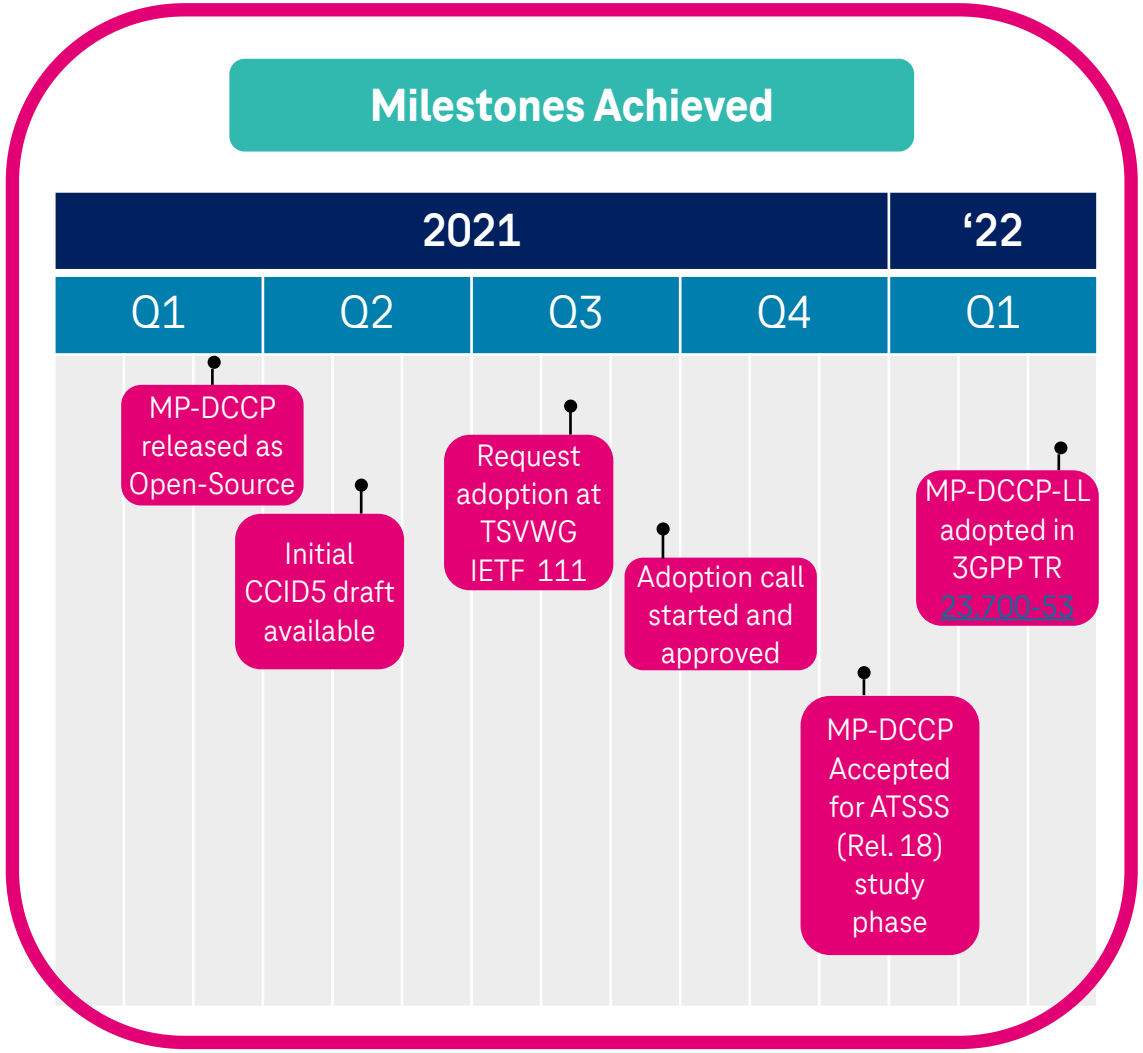
https://github.com/NathalieRM/iperf/tree/mpdccp

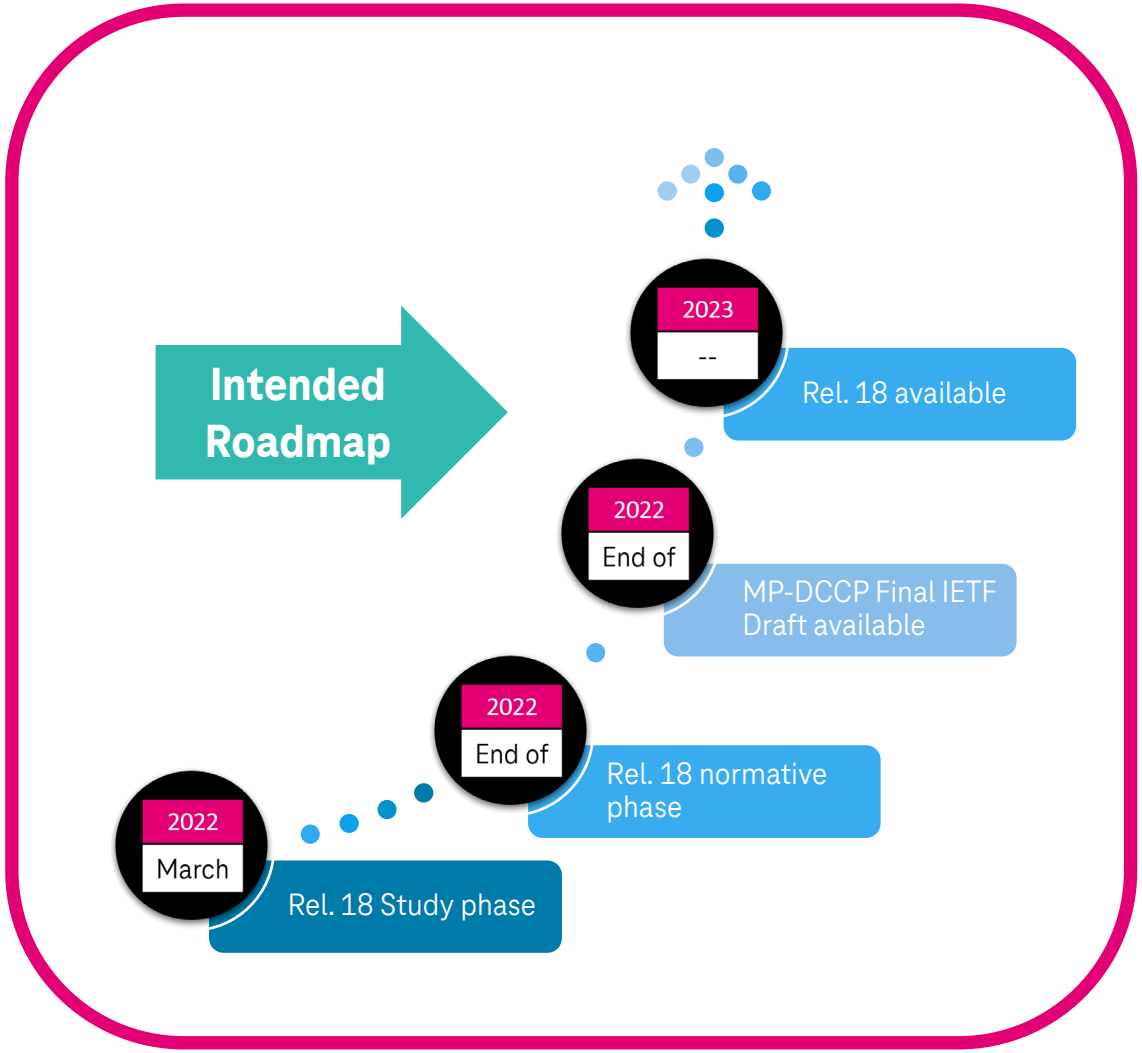
PoC with a big terminal vendor agreed using MP-DCCP public code conglomerate

Exploration started of **Random Linear Network Coding** in MP-DCCP for reducing impact of packet loss scenarios

Active draft development at GitHub with currently 9 contributors.

MP-DCCP relationship with 3GPP ATSSS work





Spontaneous Hackathon participation

Setup and presented during Hackathon to interested people

Available on-site for demoing after TSVWG session today!

Skype call between MP-DCCP enabled phone and standard phone using MP-DCCP proxy.
Commercial Wi-Fi and 4G used to demonstrate seamless handover after sudden Wi-Fi loss.

Review/testing is needed as the draft moves forward

Authors believe -04 is feature complete

All encompassing MP-DCCP Linux Kernel prototype is steadily evolving covering already most functionalities

Linux reference prototype is the base for ramping up PoCs with industry and **test interoperability**

Result generation and publication is continued with academia and in future with industry

Goal is to keep pace with 3GPP Rel. 18 timelines

→Informal sidemeeting to understand 3GPP ATSSS implications on Transport

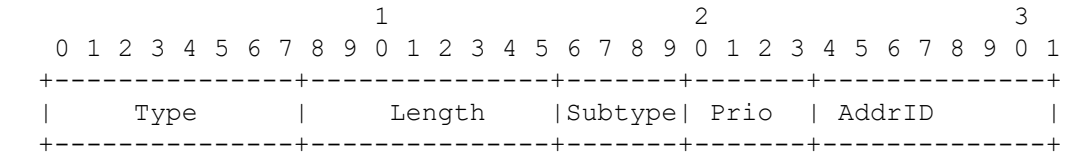
→Meet at Green Room 1 (at end of the break, 12:30 CET)

Interested people for reviewing and contribution are very welcome, contact markus.amend@telekom.de

Copied slides from TSVWG interim Feb’22

Informational backup

MP_PRIO: Fine granular path management



```
                     1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---------------+---------------+-------+-------+---------------+
|     Type      |     Length    |Subtype| Prio  | AddrID        |
+---------------+---------------+-------+-------+---------------+
```

The following values are available for Prio field:

* 0: Do not use. The path is not available.

* 1: Standby: do not use this path for traffic scheduling, if another path (secondary or primary) is available.

* 2: Secondary: do not use this path for traffic scheduling, if the other paths are good enough. The path will be used occasionally, e.g. when primary paths are congested or become not available.

* 3: Primary: can use the path in any way deemed reasonable by peer. Will always be used for packet scheduling decisions.

* 4 - 15: relative priority of one path over the other to give relative path priority for primary paths. The peer should consider sending more traffic over higher priority path. Higher numbers indicate higher priority.

Active exchange of prioritization and cost information as indicator for user plane scheduling decisions.

- Enable/disable paths (Prio=0)
- Keep paths as backup in case primary path is broken (Prio=1)
- Primary/Secondary (Prio=2-15) give 14 levels of granularity for path aggregation

Can be used to select

Steering (permanent path selection)

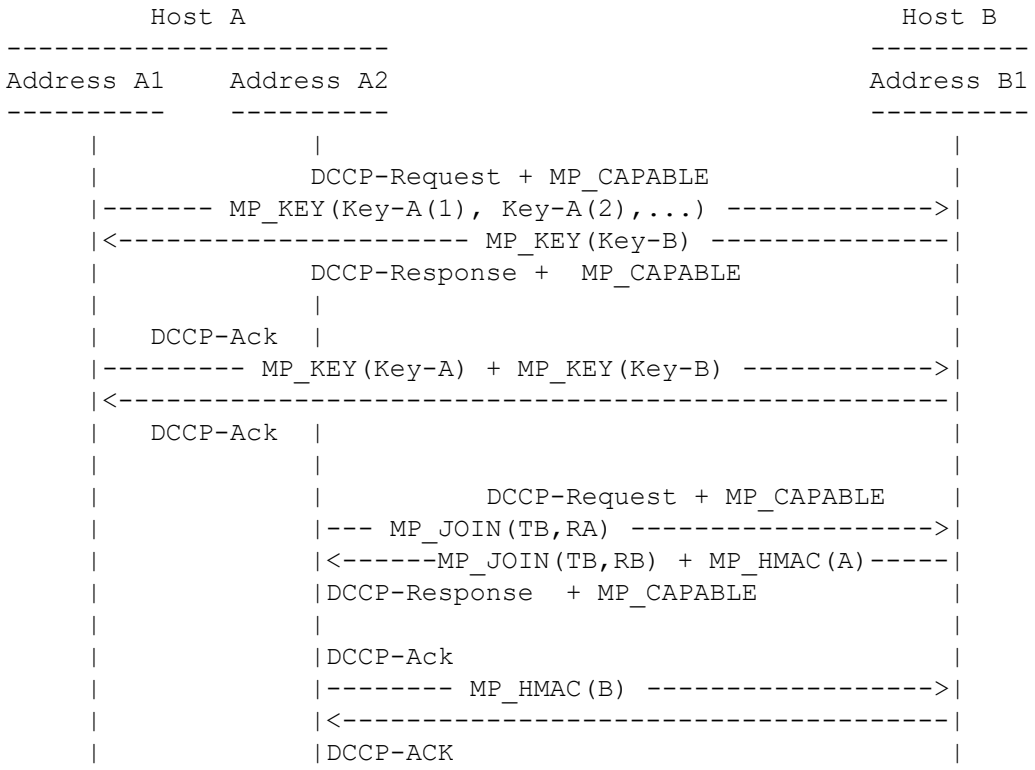
Switching (seamless handover)

Splitting (aggregation)

→ Advanced over MPTCP binary Prio parameter

→Excluded at the moment from MP-QUIC development

Handshaking procedure



Resembles MPTCP well proved handshaking mechanism including security aspects

4-way initial handshake

- **to negotiate MP support and**
- **exchange key material for setup of subsequent flows**

`MP_CAPABLE` **feature and** `MP_KEY` **option are used**

Unlimited subsequent subflow setup using `MP_JOIN` **after successful initial handshake**

Part of 3GPP Rel. 18 Study Phase for ATSSS 1/2

MP-DCCP is a selected solution to study the Key Issue (KI) on **non-TCP multi-path support for ATSSS, see S2-2200757**

MP-DCCP solution description contributed as S2-2200983 introduced as

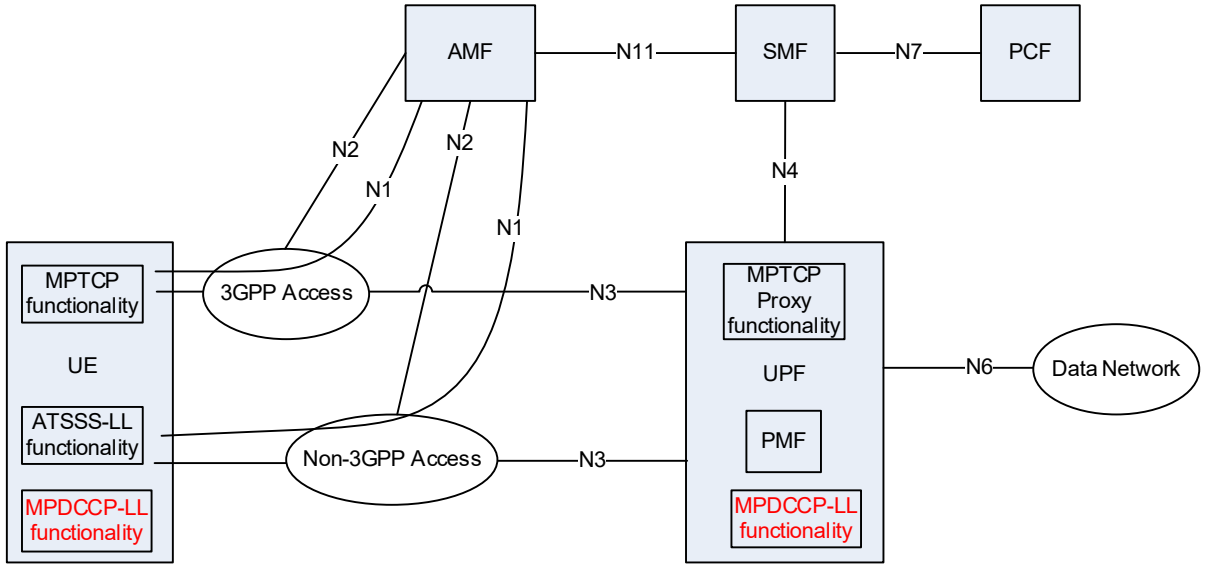
23.700-53: New Solution: MP-DCCP-LL based ATSSS steering functionality for Non-TCP traffic.

Early Supporters: Xiaomi Vodafone BT DT

Placed as Lower-Layer (-LL) solution for enabling multi-path for any traffic, at least UDP and QUIC.

Only solution which has a comprehensive set of required functionalities available as public code:

Multi-path Protocol, ATSSS defined and discussed steering modes, **re-ordering mechanisms**, Encapsulation



Part of 3GPP Rel. 18 Study Phase for ATSSS 2/2

MP-DCCP fulfils already today the requirements of **non-TCP multi-path** support for **Hybrid Access**, **end-to-end**

and in particular for

ATSSS →

3GPP ATSSS KI requirements	MP-DCCP standard/prototype capabilities	
Non-TCP support	MP-DCCP encapsulation[1]	✓
Steering modes	Scheduling[1] - Traffic distribution logics.	✓
Re-ordering	Compensate paths latency difference[1]	✓
Path measurement	As input for scheduling decisions Congestion Control CCID 2, 3, 4, 5	✓
Path management	(Re-)Establish/destruct flows	✓

[1] Prototype publication expected soon

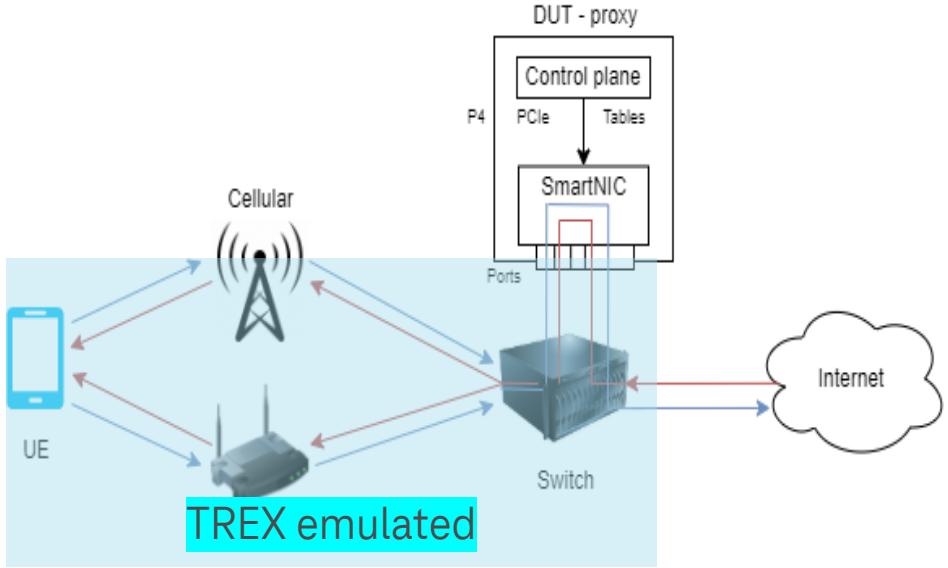
P4 based Hardware Acceleration of MP-DCCP Proxy

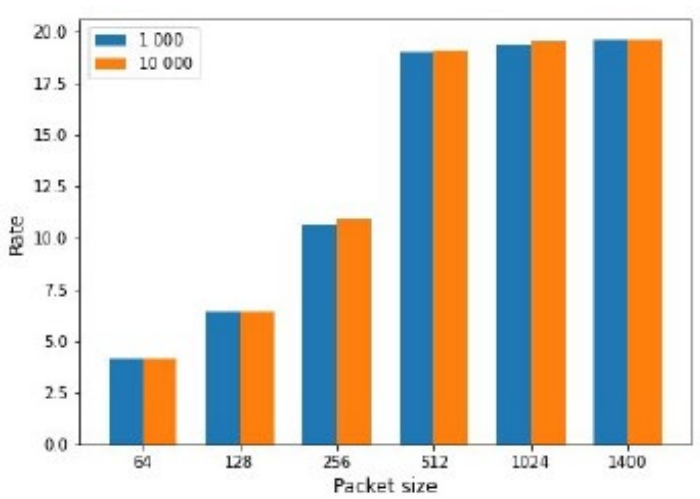
P4: Domain-specific language, specifying how data plane devices process packets - P4.org

P4 code compiles to different targets (smartNIC, Switching ASIC Tofino, ...) → hardware acceleration

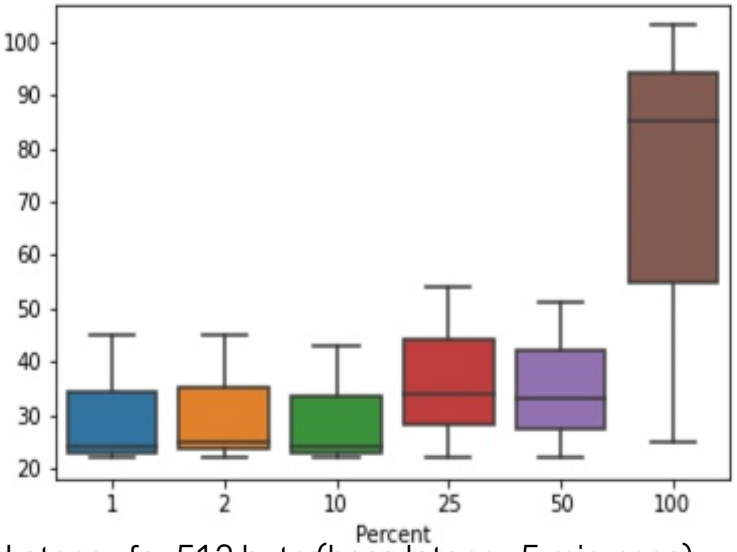
Implemented (parts of) MP-DCCP Proxy functionality in P4: DCCP header parsing and processing, MP_SEQ parsing and processing, ACK generation, keeping track of per tunnel sequence and CWND numbers per paths, Encap/Decap, RR path scheduler

Compiled to Netronome Agilo CX, TREX (v2.93 using DPDK version 21.02.0-rc1) on Intel X710 (2x10G)





Throughput over packet size for simulated 1k or 10k UEs



Latency for 512 byte (base latency 5 microsec)