MP-DCCP

for enabling transfer of UDP/IP traffic over multiple data paths in multi-connectivity networks **and the challenge of congestion control**

draft-amend-tsvwg-multipath-dccp-03

IETF 106 Meeting, ICCRG, Singapore, November 2019

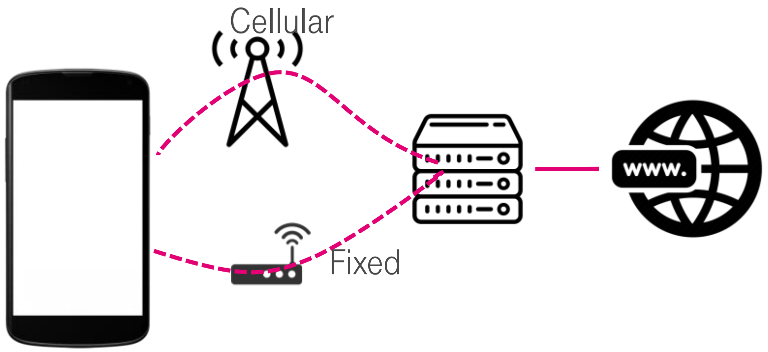
Markus Amend

KARLSTADS UNIVERSITET

CITY
UNIVERSITY OF LONDON
EST 1894

LIFE IS FOR SHARING.

MOTIVATION 1/2



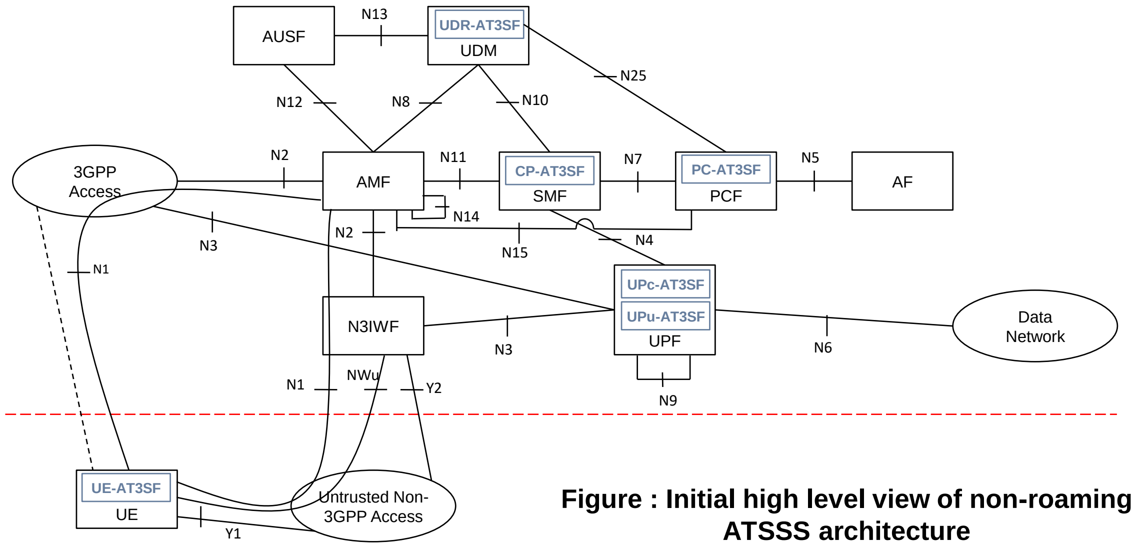
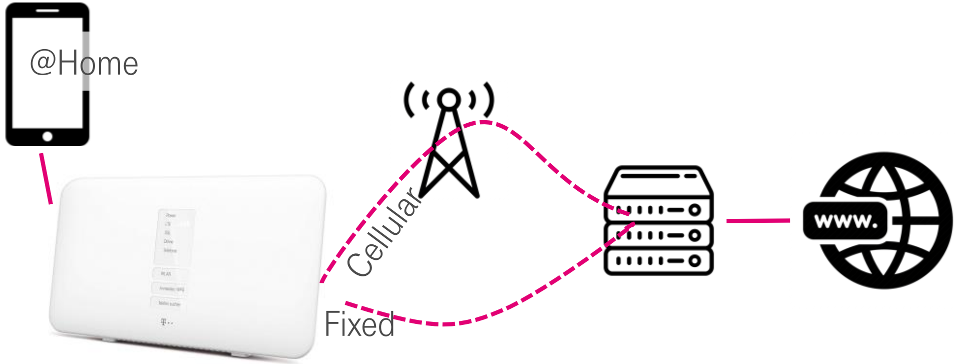


Figure : Initial high level view of non-roaming ATSSS architecture

Mobile device multi-connectivity based on expected 3GPP Rel. 16 ATSSS specification




```
+-----+                       ------
|     |                   /            \
|HOST|      +------+    |  Wireless  +-----\+-----+
|     +------+       | +-+     3G/4G   |       |      |         *****
+-----+ Wireless  +-+  \            /       |      |       **      **
                     |      |        ------          |      |     *           *
+-----+      | CPE |                        | HAG +---*  Internet *
|     |      |      |        ------         |      |     *           *
|HOST+------+      +-+  /            \       |      |       **      **
|     |Wired|      | +-+               |       |      |         *****
+-----+      +------+    |    Fixed    +-----/+-----+
                          \            /
                             ------
```

Residential multi-connectivity based on Hybrid Access at BBF

MOTIVATION 2/2

140.000 residential customer of a European ISP over one week in August 2018

Layer 3	Layer 4	>Layer 5	Share [%]	
IPv4 and IPv6	TCP		82.77	
	UDP	QUIC	11.76	16.33
		RTP	2.64	
		Other	1.93	
	Other		0.53	
Other			0.37	

HTTP/3 will push QUIC to a new order of magnitude

Demand

Multi-connectivity should cover the whole IP traffic mix in which TCP loses its dominating role because of QUIC

Findings

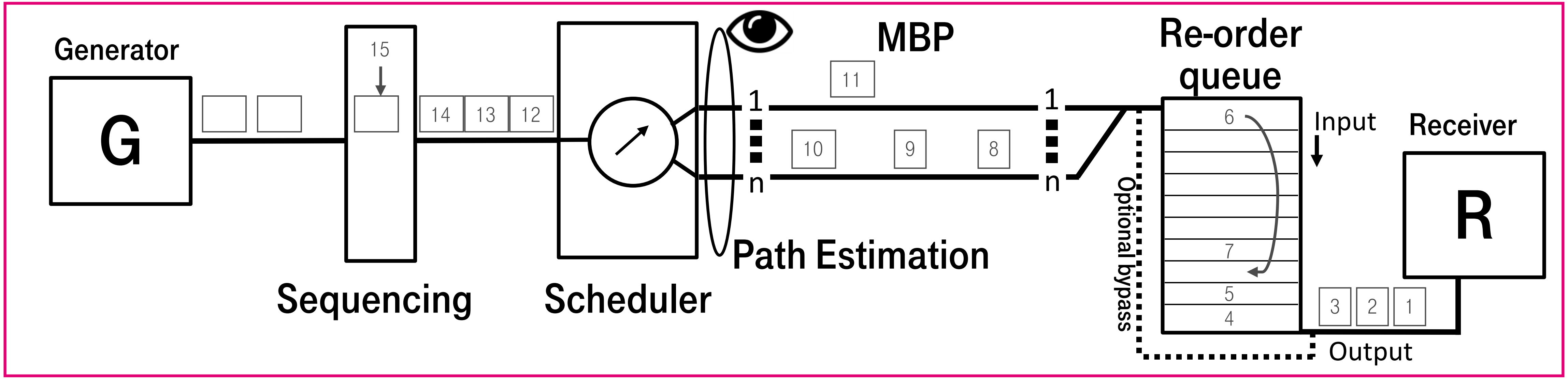
MPTCP is a good candidate to enable the TCP share for multi-connectivity. A finding from MPTCP is, that its **congestion control** empowers beneficial traffic splitting.

Multipath support for UDP or even IP does not exist.

UDP or IP encapsulation into MPTCP is not an option as it would impose reliable in-order delivery.

A potential multipath solution for UDP/IP must not impose TCP like reliability, additional high latency, packet scrambling or head-of-line blocking.Otherwise it breaks the UDP and IP principles on transportation and service expectations!

KEY COMPONENTS FOR MULTIPATH TRANSMISSION



Simultaneous path usage on PDU level comprises

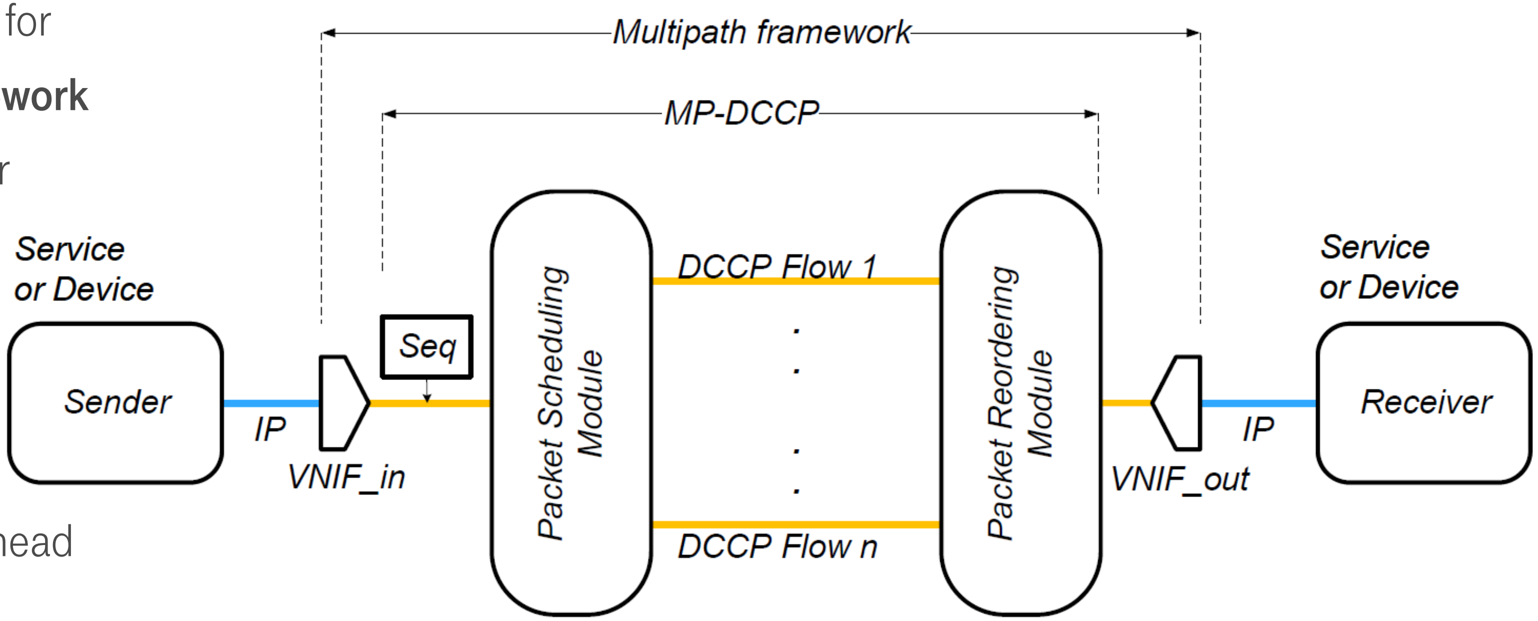
- at least on sender side a scheduling for packet based path selection
- in most real scenarios it needs additionally a re-order unit on receiver side combined with a sequencing on sender side
- possibly a special multipath bundling protocol (MBP) as transport vehicle
- for dynamic links a continuous path estimation for proper scheduling decisions and may even to optimize the re-ordering

SOLUTION: MP-DCCP FOR UDP MULTIPATH TRANSMISSION

The basic DCCP protocol is selected due to its unreliable nature, however keeping a state and employs congestion control for path estimation purposes

Making a **MP-DCCP** setup accessible for non DCCP traffic the **multipath framework** provides Virtual Network Interfaces for encapuslation purposes.

Successful DCCP transmission over uncontrolled networks and middleboxes can be ensured by overhead free **header conversion**.



https://tools.ietf.org/html/draft-amend-tsvwg-multipath-dccp-03

https://tools.ietf.org/html/draft-amend-tsvwg-multipath-framework-mpdccp-01

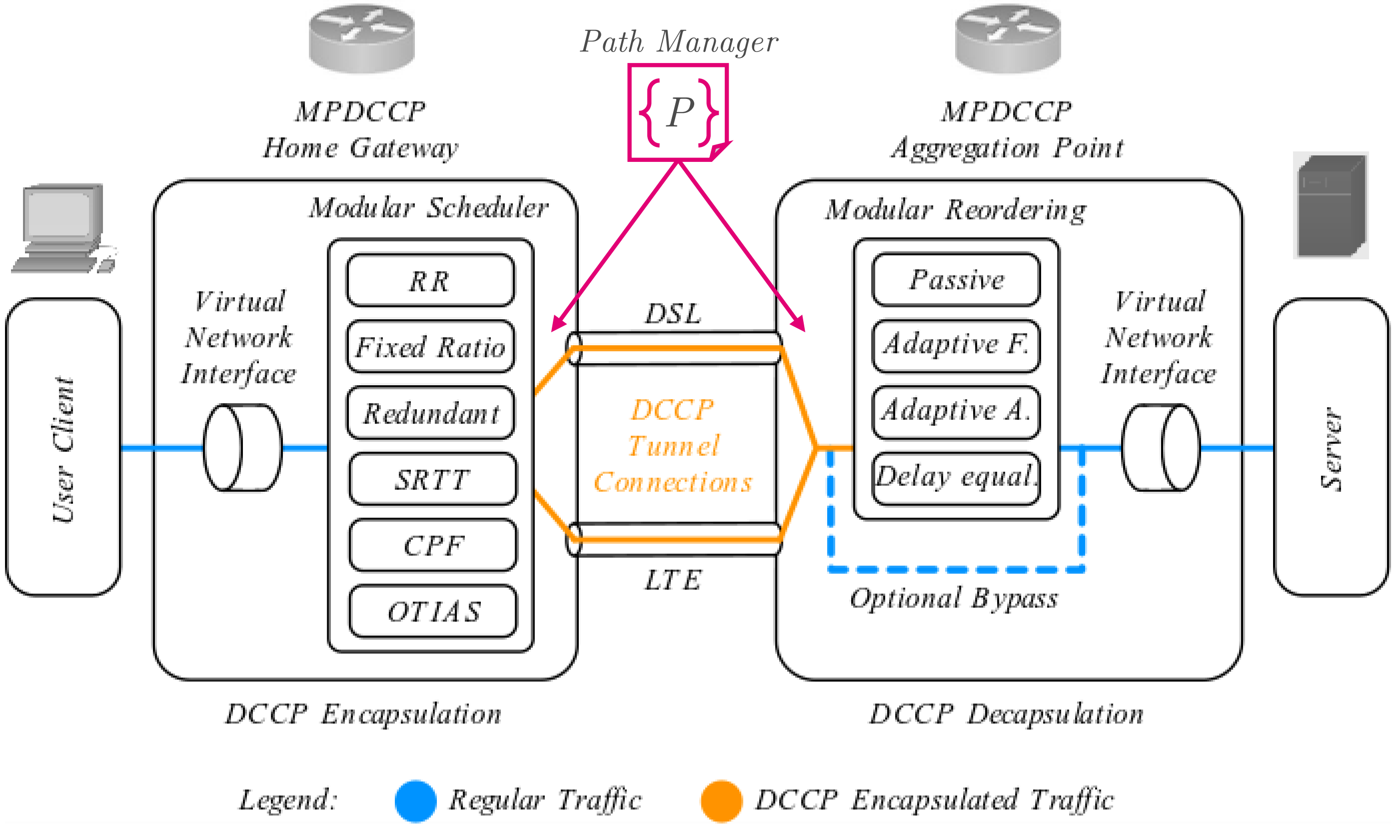
https://tools.ietf.org/html/draft-amend-tsvwg-dccp-udp-header-conversion-01

ANALYSIS AND RESULTS – TESTBED AND NS3 SIMULATIONS

Prototype is available inside Linux Kernel and ns-3 for residential and mobile use case each

- support seamless handover and path aggregation
- modular scheduler for distributing traffic
- modular re-assembly to compensate latency differences
- modular path manager to establish DCCP flows dynamically
- DCCP-UDP conversion to connect through non-DCCP aware middleboxes

→ Analysis Objective – test the ability of the framework to improve QoS/QoE on volatile paths

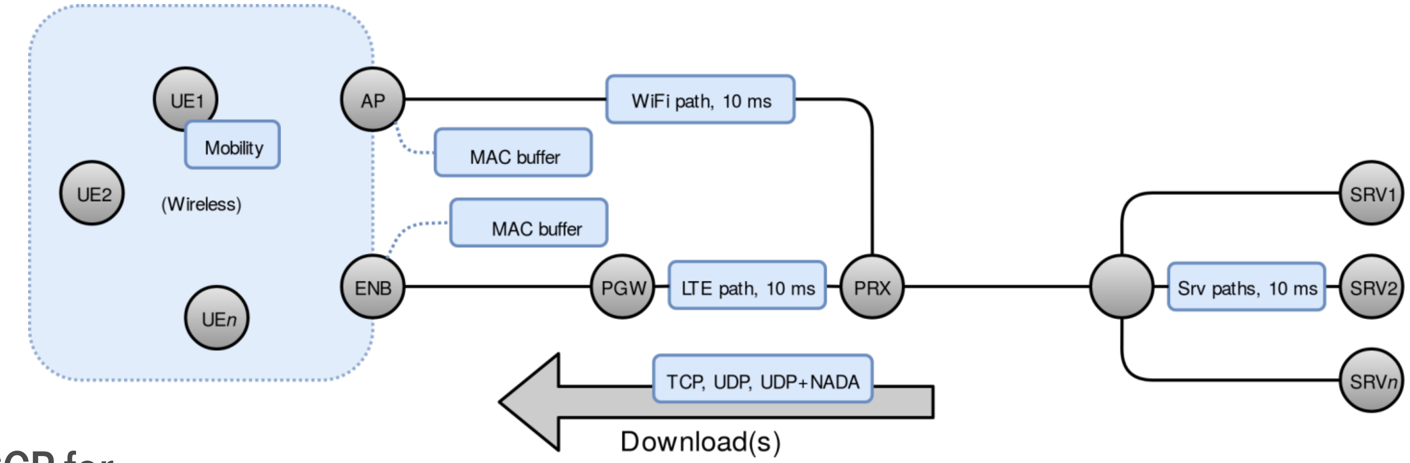


NS3 SETUP AND RESULTS

Arbitrary number of users & servers

Excess capacity on wired links (multi-Gbps)

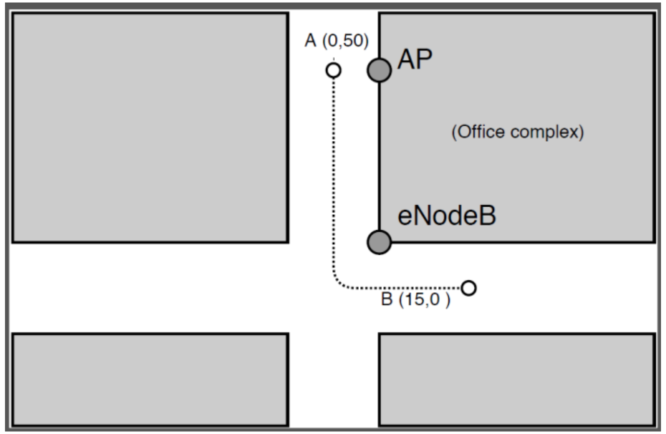
Link latency per the figure, 0 ms if not stated



Question: How fast handover is when using MPDCCP for switching and aggregation use cases?

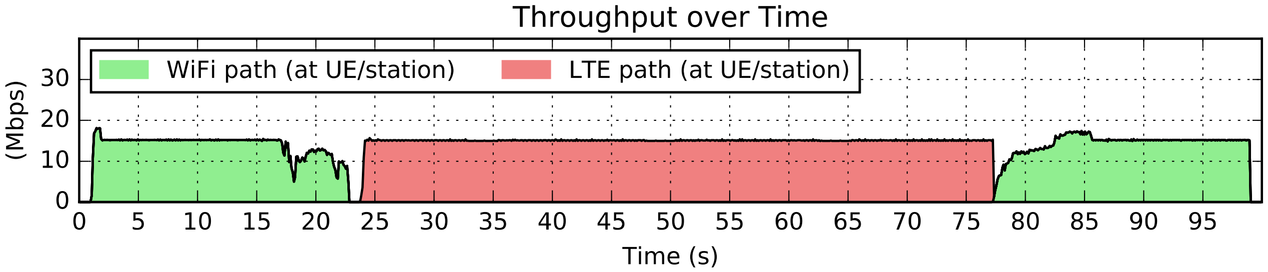
Scenario:

- Moving from A → B → A: WiFi outage near B
- Fail-over onto LTE path when approaching B
- Return to WiFi when approaching A



REMEMBER: SWITCHING AND AGGREGATION– NS3, UDP TRAFFIC

Switching in case of path failure or bad path conditions

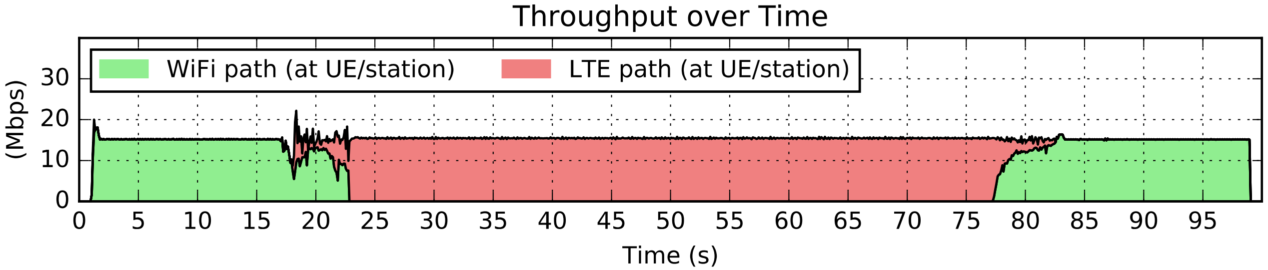


After detecting the physical loss, stream is handed over to cellular without connectivity break.

When WiFi returns, stream is handed over to Wi-Fi again

and aggregation, simultaneous path usage

is supported



Additionaly to the scenario on top, using path aggregation combined with path prioritzation on WiFi enables a smooth handover, keeping QoS stable

Compare:
https://datatracker.ietf.org/meeting/105/materials/slides-105-tsvwg-sessa-62-dccp-extensions-for-multipath-operation

SWITCHING AND AGGREGATION– NS3, UDP + NADA TRAFFIC

Evaluate support for congestion controlled UDP services

NADA –Congestion control for real-time media

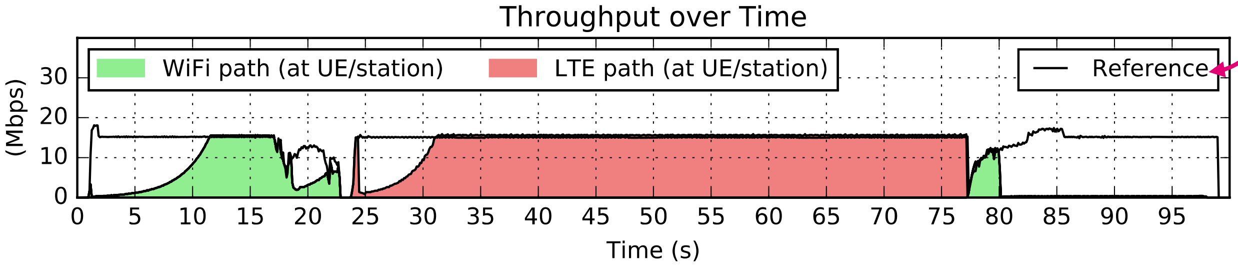
NADA behavior:

- Will increase rate when the latency is stable/low
- Will decrease the sending rate when latency increases
- Puts a cap on the sending rate (1.5 Mbps default, increased 10x for simulation)
- Targets low latency; tries to avoid buffer bloat

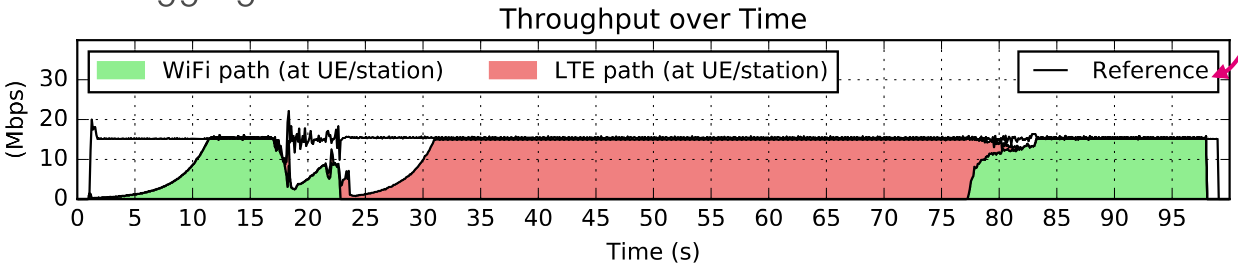
→ Aggregation advantage if the flow is large and the server congestion control is latency sensitive

Reference line indicates previous results of UDP without NADA

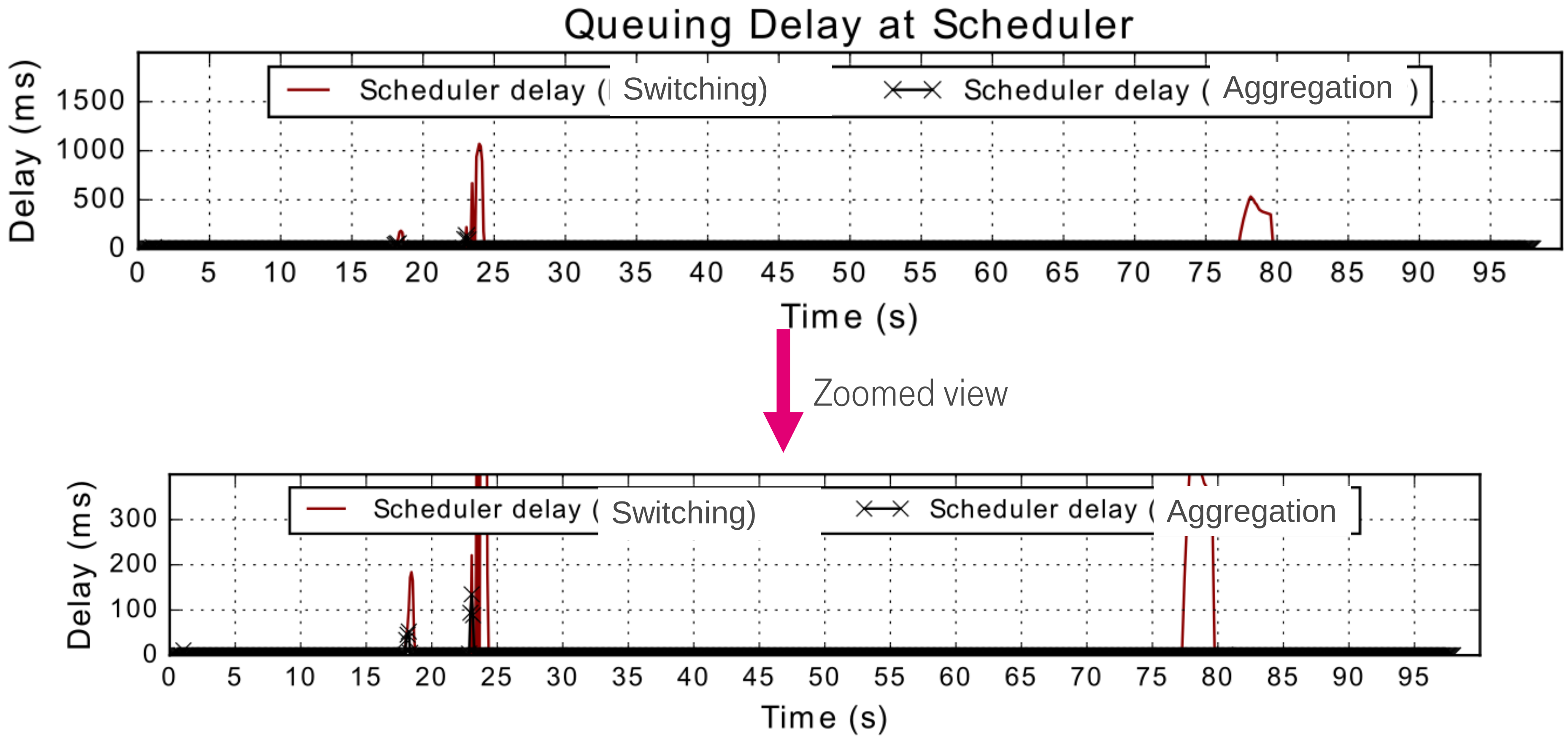
Switching



Aggregation



SCHEDULING DELAY– NS3, UDP + NADA TRAFFIC



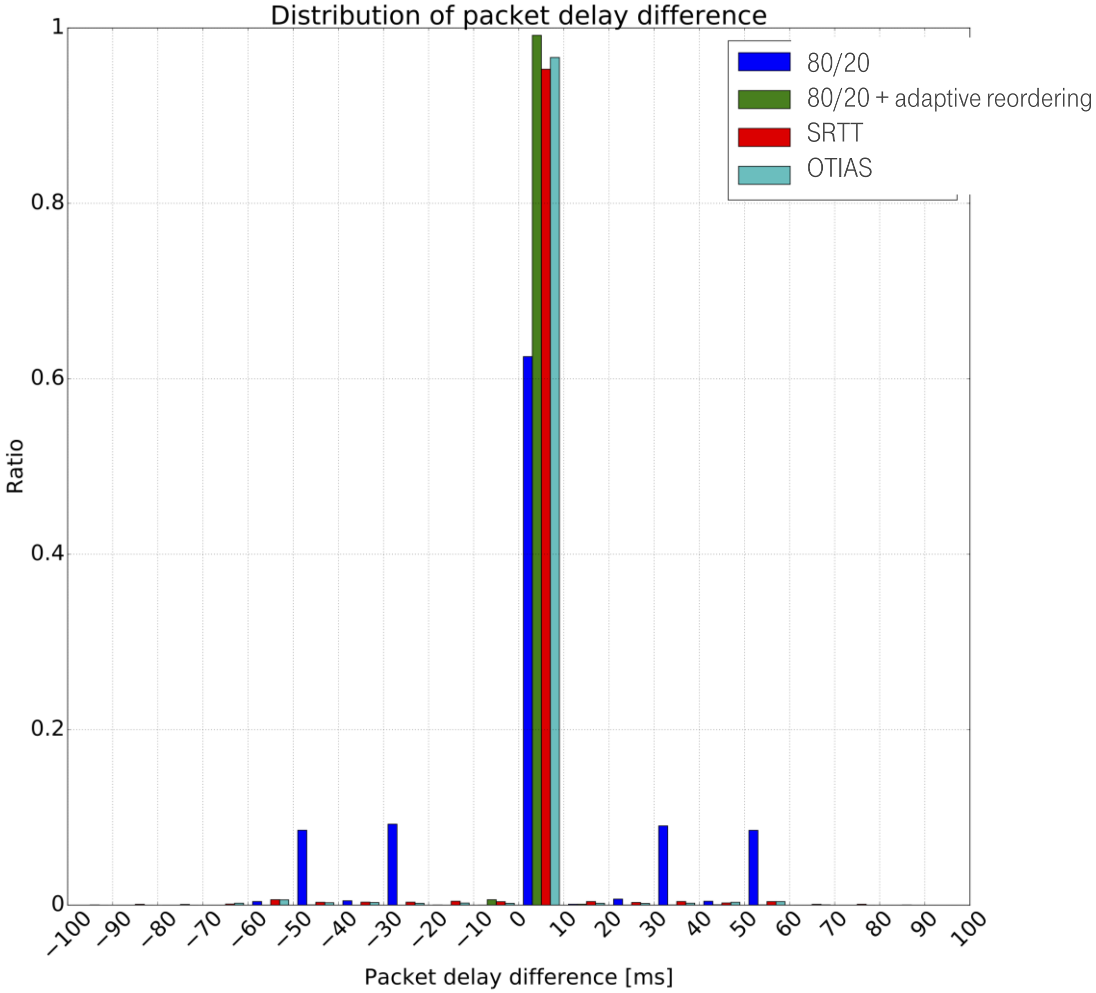
Queuing delay at scheduler over time

Switching ⟶ peek delay > 1000 ms during handover

Aggregation ⟶ peek delay < 200 ms during handover

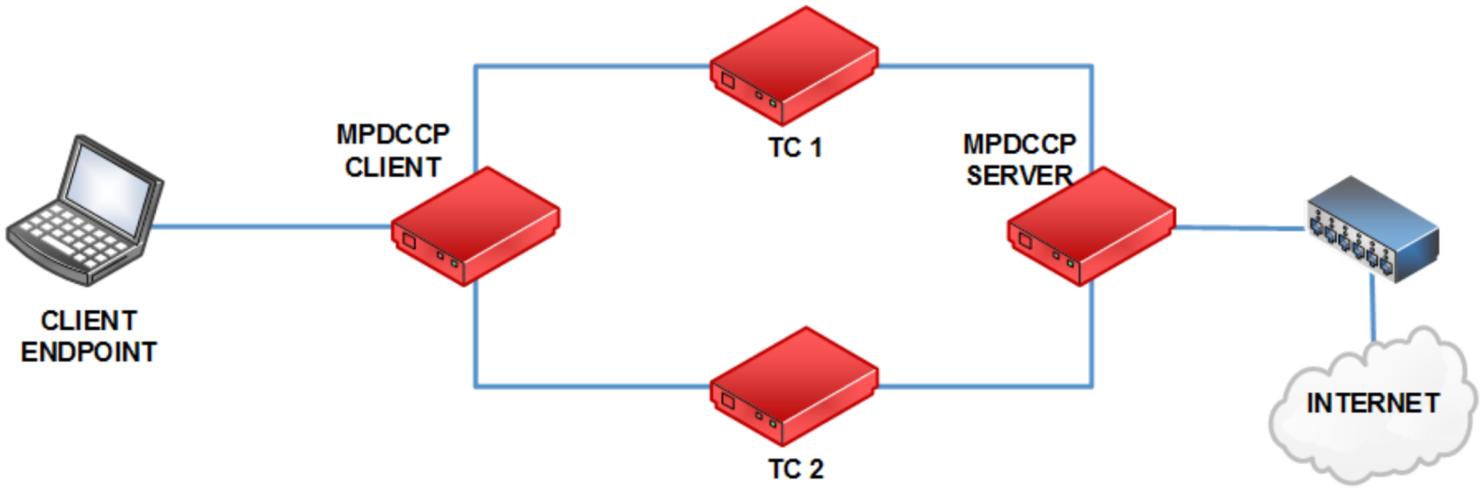
REMEMBER: MANAGING PACKET DELAY VARIATION USING SCHEDULING OR REORDERING

Path heterogeneity requires in practice a receiver side re-ordering when paths are simultaneously used to compensate latency differences.



Compare:
https://datatracker.ietf.org/meeting/105/materials/slides-105-tsvwg-sessa-62-dccp-extensions-for-multipath-operation

NEW: REAL WORLD RESULTS WITH YOUTUBE (QUIC) 1/2



- Chrome browser (QUIC enabled)
- Embedded YouTube player
- Skipping to unbuffered part of video at 10, 30, 50, 70, 90, 110s
  - Forcing buffering (stalling)
  - Always at the same unbuffered parts of the video

- Network conditions change at 60s - more on next slide
- Total duration 120s
- Playback ratio = Playing time/Total time (120s)

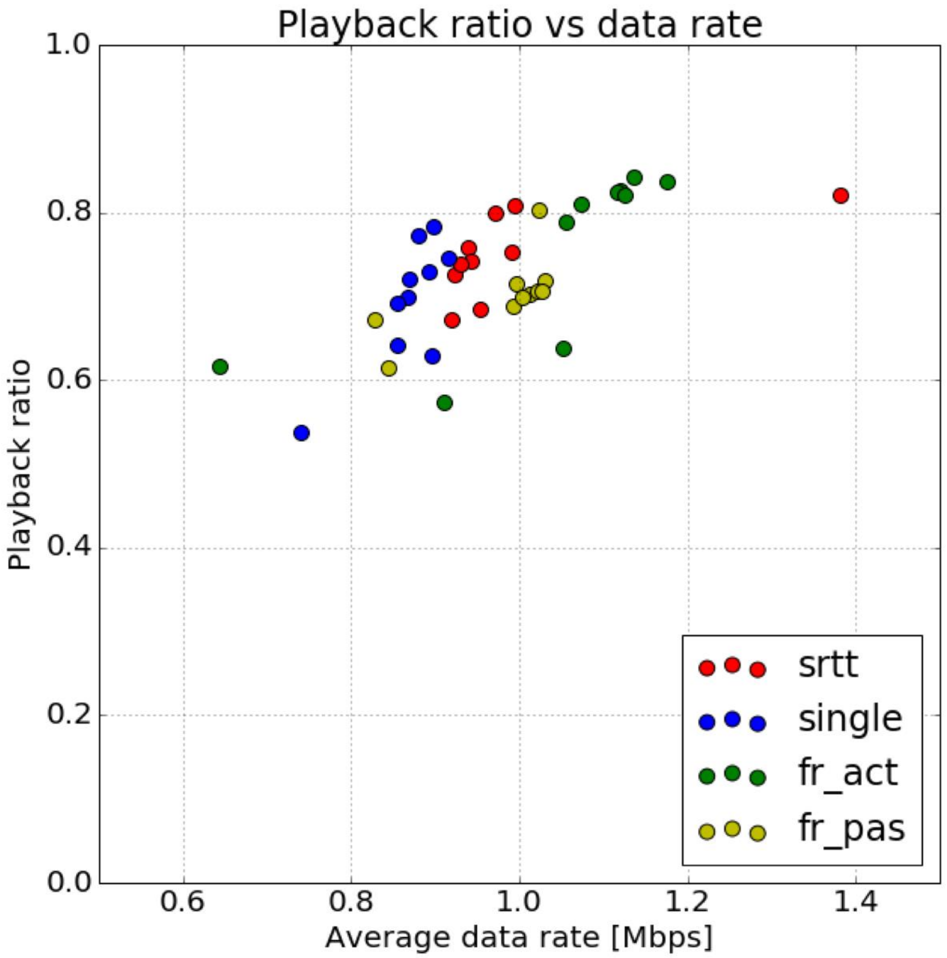
  Always < 1 due to initial loading and skipping

NEW: REAL WORLD RESULTS WITH YOUTUBE (QUIC) 2/2

Path		Bandwidth	Latency
1	$t \leq 60s$	1Mbps	10ms
	$t > 60s$	1Mbps	90ms
2		1 Mbps	50 ms

The **highest gain could be reached by path aggregation** using the fixed ratio scheduler (80:20) combined with re-ordering over the one without re-ordering and srtt scheduler **compared to no aggregation using single path**.

Detailed evaluation revealed, however, imperfect path usage even in the best performing scenario.



INTERMEDIATE CONCLUSION

Scheduling profits from DCCP‘s congestion control (CC) and can perform similar to MPTCP

Re-ordering on receiver side is proved mandatory for efficient multipath transmission of unreliable traffic. Smart algorithms to keep a traffic flow smoothly ongoing are required and can make use of CC information. Head-of-line blocking is not purposeful and it should always assessed if it is worth to wait for missing out-of-order information.

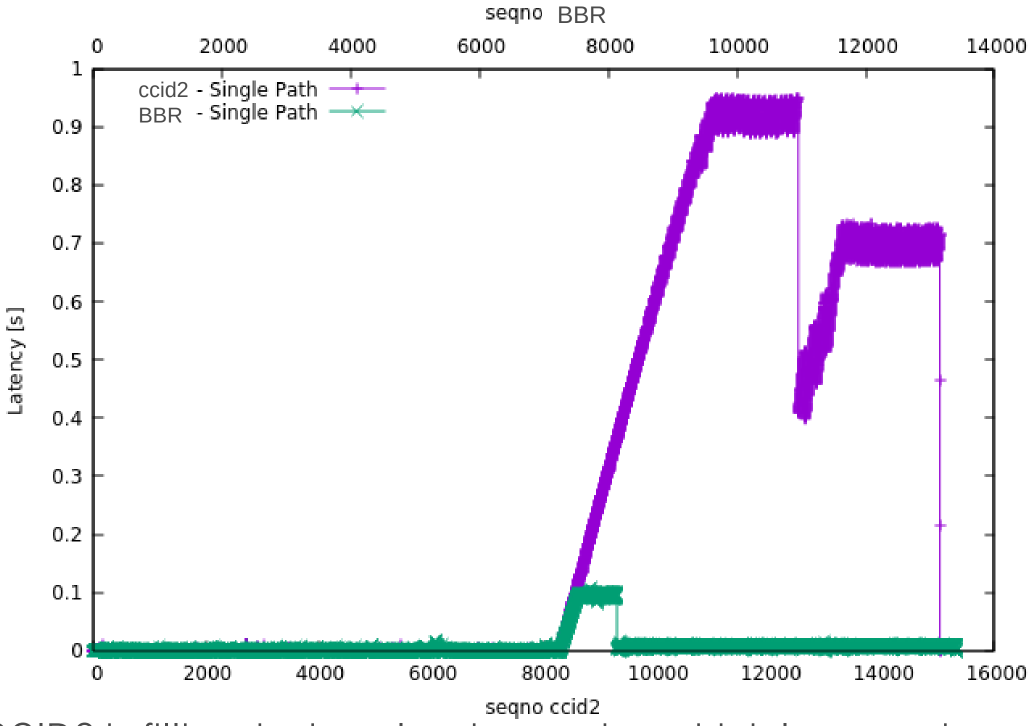
For scheduling and re-ordering purposes, the prototyp offers already several implementations. Using DCCP shipped with a CC is from this perspective a clever decision and moreover can ensure that paths becomes not overloaded.

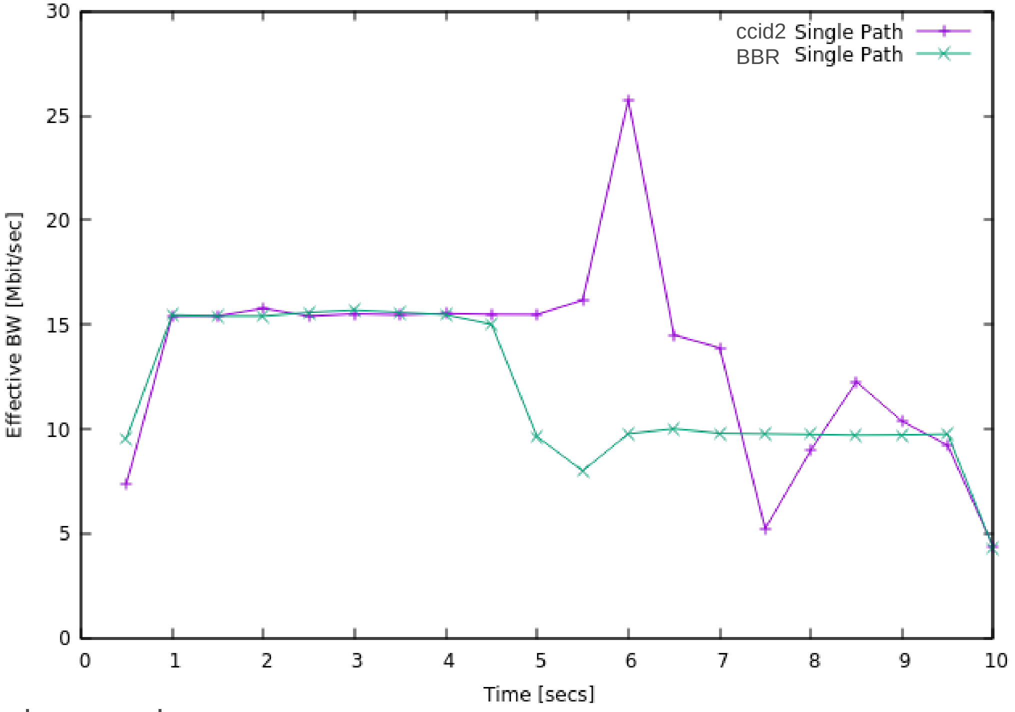
However the main target protocols transmitted over the MP-DCCP framework like QUIC employ an own CC. In combination with MP-DCCP this leads to a kind of CC over CC scenario with well known issues from literatur.

Is DCCP‘s packe loss oriented CCID2 the right CC for multipath?

NEW: IMPLEMENTING BBR FOR DCCP AND REPLACING CCID2

CCID 2 is a TCP-like congestion control for DCCP and so far used in the prototypes. It is packet-loss triggered and based on AIMD. Latency behavior in a single path setup, sending a 15Mbps UDP stream when limiting the datarate half way trough to 10 Mbps comparing CCID2 and BBR, can be seen below.



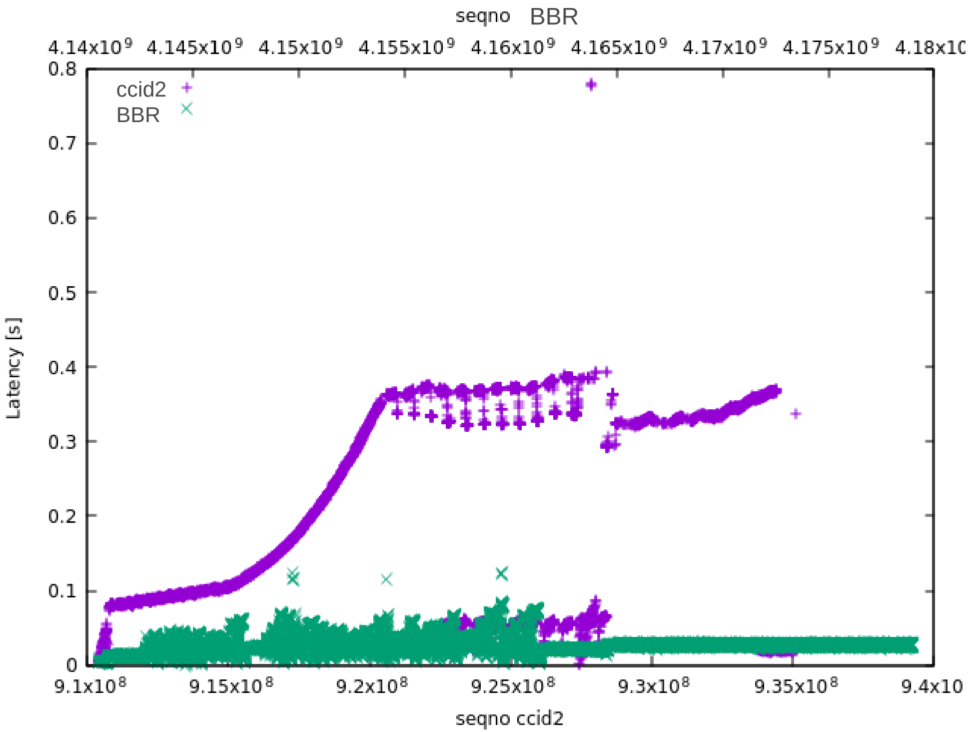


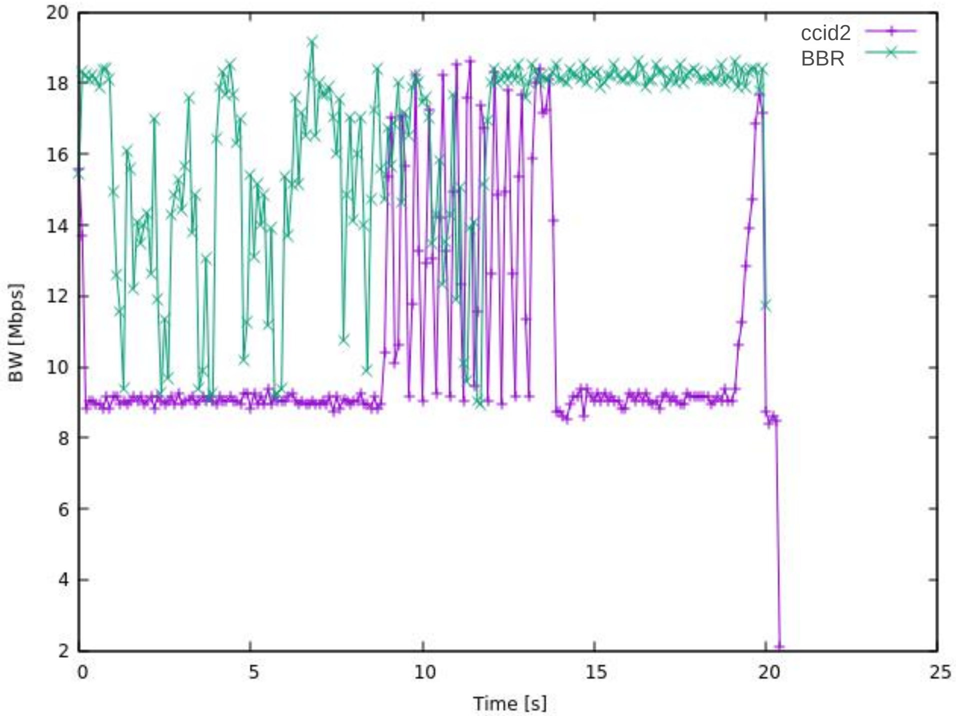
CCID2 is filling the bottelnecks causing a high latency wheres BBR keeps latency low.

For multipath environments this can lead to huge latency differences between paths and therefore a worse overall performance

NEW: IMPLEMENTING BBR FOR DCCP AND REPLACING CCID2

BBR outperforms CCID2 when a first path is saturated before overflowing in a second path. Nevertheless it is not perfect.





But there are a lot of further questions. What happens in volatile scenarios with packet losses/BER, dynamic changes of latency and datarate behavior, different mixes of inner and outer CCs …

CONCLUSION & NEXT STEPS

Further investigation of CC over CC required and could be interesting for the QUIC community as well, e.g. QUIC tunneling
https://tools.ietf.org/html/draft-kuehlewind-quic-substrate-00

Testbed is available for further testing and can be requested by interested people

Is ICCRG the right place and who is interested in elaborating this topic further?

Presumabely conclusions with minor relevance to ICCRG:

The prototype implementation and simulation show very good first results according to the demands of Steering, Switching and Splitting of 3GPP ATSSS and BBF Hybrid Access.

Scheduling and re-ordering are proved beneficial.

Congestion control for path estimation is proved

UDP/IP traffic can be transmitted in switching or aggregation scenario

Please use iccrg@irtf.org/tsvwg@ietf.org or markus.amend@telekom.de to get in touch with us.

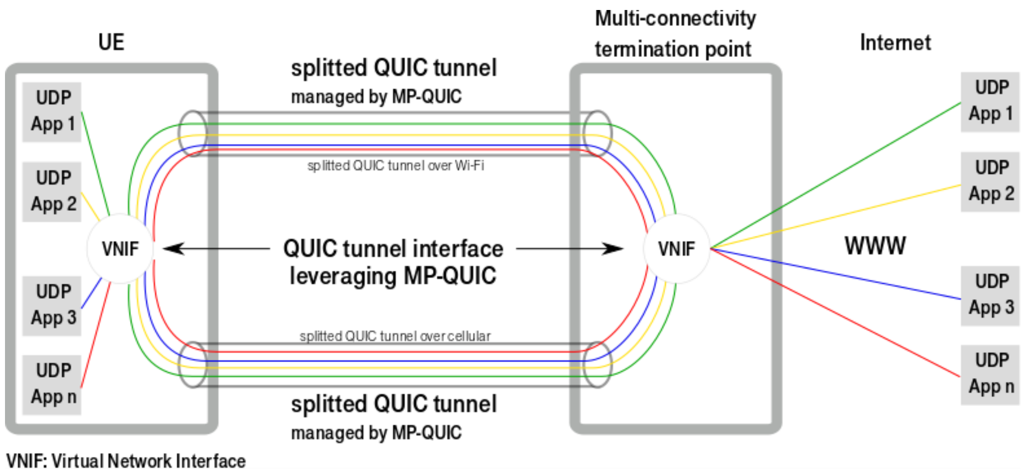
Further documents

Paper with detailed results: https://arxiv.org/pdf/1907.04567.pdf, IETF 104 presentation, IETF 105 ICCRG presentation

BACKUP

WHY NOT USE MP-QUIC INSTEAD OF MP-DCCP?

(MP-)QUIC according to https://tools.ietf.org/html/draft-ietf-quic-transport-24 is a reliable and end-to-end encrypted protocol. Its application for enabling multipath transfer for UDP/QUIC traffic only works as QUIC tunnel, managed by MP-QUIC.



- **Useless encryption is applied and requires resources**
  - UDP as guest: Turns UDP into reliable transmission ☹
    - QUIC as guest: Encryption over Encryption, otherwise like TCP below ☹☹
  - TCP as guest: TCP's CC + reliable in-order delivery over outer QUIC's CC + reliable in-order delivery ☹☹☹

REQUIRED (MP-)QUIC ADAPTATIONS

In case MP-QUIC shall become an alternative for ATSSS and Hybrid Access like network architectures, it would require a paradigm change:

→ Configurable encryption for

- reducing the useless overhead in case of QUIC over MP-QUIC (likely)
- designing a MP-QUIC ↔ QUIC converter (unlikely)

→ Deal with unreliable traffic to some extent and remove at least the reliable and in-order delivery feature

- Unreliable traffic support requires a complete re-work of current MP-QUIC framework, which bases on QUICs reliable and in-order delivery.

→ Define a QUIC tunnel protocol