A multipath framework for UDP traffic over heterogeneous access networks
              draft-amend-tsvwg-multipath-framework-mpdccp-01

Abstract

   More and more of today's devices are multi-homing capable, in
   particular 3GPP user equipment like smartphones. In the current
   standardization of the next upcoming mobile network generation 5G
   Rel.16, this is especially targeted in the study group Access Traffic
   Steering Switching Splitting [TR23.793]. ATSSS describes the
   flexible selection or combination of 3GPP untrusted access like Wi-Fi
   and cellular access, overcoming the single-access limitation of
   today's devices and services. Another multi-connectivity scenario is
   the Hybrid Access [I-D.lhwxz-hybrid-access-network-architecture][I-D.
   muley-network-based-bonding-hybrid-access], providing multiple access
   for CPEs, which extends the traditional way of single access
   connectivity at home to dual-connectivity over 3GPP and fixed access.
   A missing piece in the ATSSS and Hybrid Access is the access and path
   measurement, which is required for efficient and beneficial traffic
   steering decisions. This becomes particularly important in
   heterogeneous access networks with a multitude of volatile access
   paths. While MP-TCP has been proposed to be used within ATSSS, there
   are drawbacks when being used to encapsulate unreliable traffic as it
   blindly retransmits each lost frame leading to excessive delay and
   potential head-of-line blocking. A decision for MP-TCP though leaves
   the increasing share of UDP in today's traffic mix
   (<https://arxiv.org/abs/1801.05168>) unconsidered. In this document,
   a multi-access framework is proposed leveraging the MP-DCCP network
   protocol, which enables flexible traffic steering, switching and
   splitting also for unreliable traffic. A benefit is the support for
   pluggable congestion control which enables our framework to be used
   either independent or complementary to MP-TCP.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Multi-connectivity access networks are evolving. Ongoing
standardization at 3GPP for 5G mobile networks [TR23.793] or the so
called Hybrid Access networks [I-D.lhwxz-hybrid-access-network-archit
ecture][I-D.muley-network-based-bonding-hybrid-access] already
provides or will enable in the near future the possibility to use
multi-connectivity for a very large number of mobile users. Multi-
connectivity solutions come with many user benefits including
superior resilience against network outages, higher capacities for
user traffic and network cost optimizations. Since multi-
connectivity architectures are almost mature, new network protocols
are required to fully exploit multi-connectivity and maximise its
potential. In the simplest case, multi-connectivity is used for
load-balancing decisions in order to balance the user flows over
multiple paths. However, this has no effect on resilience or
capacity gain on those load balanced individual flows. More complex
scenarios include the dynamic shifting of traffic flows seamlessly
between multiple paths or even aggregating those paths for leveraging
the available capacity of multiple individual paths. Like [TR23.793]
this document refers to the three distribution schemes Steering (load
balancing), Switching (seamlesshandover) and Splitting (capacity
aggregation).

MP-TCP [RFC6824] is a protocol, which can be applied in the above
mentioned use cases and supports load-balancing, traffic shifting
among the multiple paths and also capacity aggregation. Further, it
leverages the inherent congestion control from TCP which adapts the
sending rate by observing congestion signals from the network. By
design, MP-TCP is limited to TCP services as it blindly re-transmits
lost packets. Consequently, when MP-TCP is used as a framework for
ATSSS, it may re-transmit packets sent from unreliable services such
as e.g. UDP unnecessary. This may lead to head-of-line blocking and
increased latency, which is detremental to real-time services. As
future multi-connectivity systems must support latency sensitive
traffic that might be transported over unreliable transport, it is
not sufficient anymore to rely on supporting only TCP. The
increasing share of UDP traffic, mainly impacted by the QUIC
introduction, may significantly reduce the share from TCP. It might
be expected that with HTTP/3 carried over QUIC [I-D.ietf-quic-http],
the previous strong dominance of TCP will be challenged by UDP.

2. Requirements

A multiaccess framework shall meet the following requirements:

o IP compatibility: A multiaccess framework shall be able to
   transport IP packets and not make any assumptions on which
   transport protocol is encapsulated.

o Support for unreliable traffic: A multiaccess framework should
   provide support for transporting unreliable traffic, such as QUIC
   or UDP based flows. Therefore, unreliable transmission should
   besupported.

o Support for flexible re-ordering: A multiaccess framework should
   support flexible re-ordering of user traffic, including no re-
   ordering at all. This requirements is important to support low
   latency traffic, where the re-creation of packet order may
   negatively impact delivery latency.

o Support for flexible congestion control: A multiaccess framework
   should support flexible congestion control, including the
   disabling of the congestion control, if the inner traffic is known
   to be congestion controlled.

o Support for flexible packet scheduling: A multiaccess framework
   should support different packet scheduling mechanisms, which
   should be configurable from the control plane. Examples are
   cheapest path first, or other more sophisticated schedulers.

o Lightweight: A multiaccess framework should be lightweight in
   computational resources and limit the encapsulaton overhead.

To use QUIC as part of a multiaccess framework, by for example
providing multipath support for QUIC, it could be beneficial if
unreliable transmission is supported as well as being able to
influence or disable QUICs congestion control. In addition, it would
be beneficial if the encryption of QUIC can be disabled. This is
because for ATSSS, it is foreseen that the underlying tunnel from the
mobile over public WLANs is baseed on IPSec.

3. IP compatible multipath framework based on MP-DCCP

We propose a new multiaccess framework, which overcomes MP-TCP's
restriction to TCP services and provides IP compatibility in
Figure 1. The framework employs MP-DCCP
[I-D.amend-tsvwg-multipath-dccp] in combination with an encapsulation
scheme. For simplification, Figure 1 assumes a traffic direction
from the left (sender) to the right (receiver) and requires
application in each direction for bi-directional transmission. The
framework in Figure 1 can replace or complement MP-TCP to reach IP
compatibility.

```
 Service          |<-             MP-DCCP           >|          Service
 or Device                                                      or Device
 +-------+          ___ +-----+ DCCP Flow 1 +------+          +--------+
 |       |    _  |Seq||Path |-------------|Re-   |    _   |        |
 | Sender|___| \__\ /_|     |      :      |order |____/ |___|Receiver|
 |       | IP|_/      |Sched|      :      |      |      \_|IP |        |
 |       |    VNIF_in |uler |-------------|engine| VNIF_out |        |
 +-------+            +-----+ DCCP Flow n +------+          +--------+
```
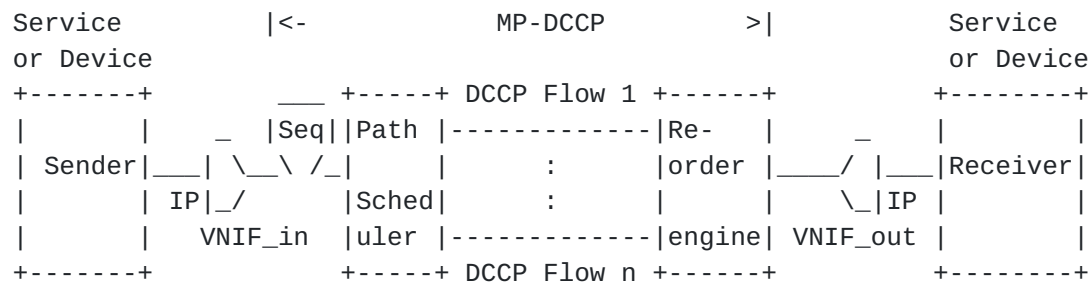
       Figure 1: IP compatible multipath framework based on MP-DCCP

   PDUs generated from the sender and travelling through the framework
   in Figure 1 pass the components in the following order:

   1. Sender: Generates any PDU based on the IP protocol.

   2. VNIF_in: IP based Virtual Network Interface as entry point to the
       multipath framework. A simple routing logic in front (between
       (1)and (2)) can act as gatekeeper and decides upon redirecting
       traffic through the VNIF or bypassing it. The VNIF adds an extra
       IP header to reach the multi-connectivity termination point.

   3. Seq: Sequencing of the PDUs passed through (2) depending on the
       incoming order. Adds an incrementing number, which is later
       added to the DCCP encapsulation in (4).

   4. Path Scheduler: Decision logic for scheduling sequenced PDUs over
       the individual connected DCCP flows for multipath transmission.
       The path scheduler can use the information from the DCCP flows
       (see (5)) inherent congestion control information like CWND,
       packet loss, RTT, Jitter, etc.. After selection of a DCCP flow,
       the PDU is encapsulated into the individual flow. Further
       information, at least the sequencing, is added on top as DCCP
       option.

   5. DCCP Flow(s): Responsible to transmit the encapsulated PDUs to
       the MP-DCCP exit point.

   6. Reorder engine: Depending on the sequencing information of (3), a
       re-assembly of the PDU stream can be applied. Different re-order
       algorithms should be supported in a configurable way, including
       no re-ordering.

   7. VNIF_out: Releases PDUs that have passed the re-ordering engine
       and strips the DCCP specific overhead. Again, routing is
       responsible to deliver the PDUs to the receiver based on the
       destination information in the PDU.

8. Receiver: Receive the PDU as generated in (1).

The simple enclosing of the MP-DCCP with Virtual Network Interface
(VNIF) provides the IP compatibility. However, a service or protocol
classifier between sender and VNIF can reduce the scope to particular
traffic, e.g. UDP, by simple routing decisions. The MP-DCCP takes
over responsibility for the multi-path transfer of the traffic, which
is directed through the VNIF_in. For possible re-assembly
operations, the IP packets may be stamped with a continuously
incremented sequence number. This is not mandatory, but assumed
required in most seamless handover and capacity aggregation use
cases. The path scheduler decides for each IP packet, which DCCP
flow it should use for encapsulation, based on a configurable
decision logic and supported by the congestion control information of
the DCCP flows available for transmission. A DCCP flow selection for
a PDU leads to its encapsulation into the respective DCCP flow and
adding extra information required for the multipath transmission,
e.g. the sequence number. Encapsulation also means, that a DCCP and
IP header is added to the original PDU to reach the multi-
connectivity end-point. When the encapsulated PDUs arrive at the
multi-path termination point, they are re-ordered depending on the
carried sequence number and a configurable logic. The re-ordering
engine may also include a logic in which packets are just forwarded
(no re-ordering). Re-ordering needs to be considered carefully since
any active intervention changes the latency responsiveness. The
multi-path termination is finally completed when the DCCP overhead is
stripped and the PDU leaves VNIF_out. Further routing depends again
on the IP layer of the original PDU.

4. Application and placement

The framework of Figure 1 is very flexible in applying multipath
support in different architectures and allows MP-DCCP to be applied
at any place between sender and receiver. Figure 2 to Figure 5
provide several architectural options for the deployment of the
framework.

```
 Device           Middlebox 1          Middlebox 2          Device
 +------+     +-------------+     +------------+     +--------+
 |Sender| -> |MP-DCCP entry| -> |MP-DCCP exit| -> |Receiver|
 +------+     +-------------+     +------------+     +--------+
```
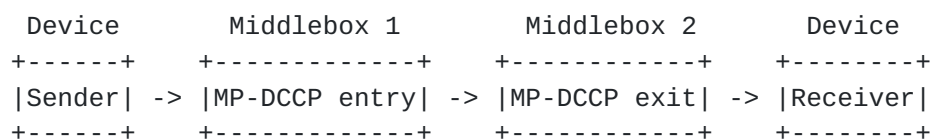
Figure 2: Sender and receiver independent MP-DCCP

```
        Device                    Middlebox        Device
  +----------------------+    +------------+    +--------+
  |Sender + MP-DCCP entry| -> |MP-DCCP exit| -> |Receiver|
  +----------------------+    +------------+    +--------+
```
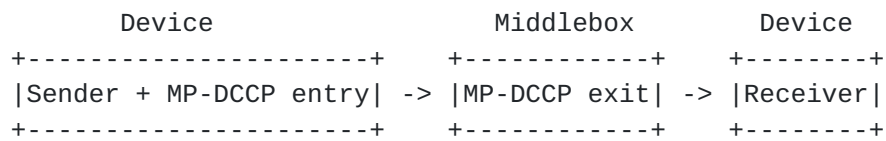
Figure 3: Sender integrated but receiver independent MP-DCCP

```
 Device          Middlebox                     Device
 +------+    +-------------+    +----------------------+
 |Sender| -> |MP-DCCP entry| -> |MP-DCCP exit + Receiver|
 +------+    +-------------+    +----------------------+
```
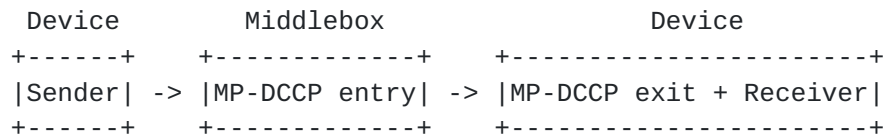
Figure 4: Sender independent and receiver integrated MP-DCCP

```
        Device                       Device
  +----------------------+    +----------------------+
  |Sender + MP-DCCP entry| -> |MP-DCCP exit + Receiver|
  +----------------------+    +----------------------+
```
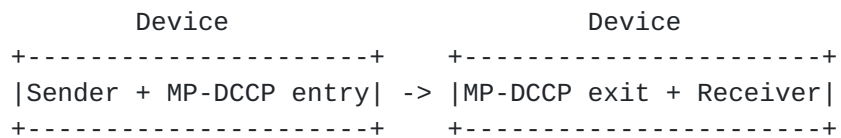
Figure 5: Sender and receiver integrated MP-DCCP

[5](#). Conclusion

The specified IP compatible multipath framework based on MP-DCCP in
this document comprises several benefits:

o Pure routing

o Inherent path estimation and measurement

o Imposes no constraints on reliability or in-order delivery of
   application PDUs

o Modular re-ordering

o Modular scheduling

o IP compatible

o Based on the standardized DCCP.

Middle-box traversing, when the framework is applied in uncontrolled
environments, is addressed in [RFC6733] and
[I-D.amend-tsvwg-dccp-udp-header-conversion].

6. Security Considerations

   [Tbd]

7. Acknowledgments

8. Informative References

   [I-D.amend-tsvwg-dccp-udp-header-conversion]
             Amend, M., Brunstrom, A., Kassler, A., and V. Rakocevic,
             "Lossless and overhead free DCCP - UDP header conversion
             (U-DCCP)", draft-amend-tsvwg-dccp-udp-header-conversion-01
             (work in progress), July 2019.

   [I-D.amend-tsvwg-multipath-dccp]
             Amend, M., Brunstrom, A., Kassler, A., and V. Rakocevic,
             "DCCP Extensions for Multipath Operation with Multiple
             Addresses", draft-amend-tsvwg-multipath-dccp-01 (work in
             progress), March 2019.

   [I-D.ietf-quic-http]
             Bishop, M., "Hypertext Transfer Protocol Version 3
             (HTTP/3)", draft-ietf-quic-http-18 (work in progress),
             January 2019.

   [I-D.lhwxz-hybrid-access-network-architecture]
             Leymann, N., Heidemann, C., Wasserman, M., Xue, L., and M.
             Zhang, "Hybrid Access Network Architecture", draft-lhwxz-
             hybrid-access-network-architecture-02 (work in progress),
             January 2015.

   [I-D.muley-network-based-bonding-hybrid-access]
             Muley, P., Henderickx, W., Geng, L., Liu, H., Cardullo,
             L., Newton, J., Seo, S., Draznin, S., and B. Patil,
             "Network based Bonding solution for Hybrid Access", draft-
             muley-network-based-bonding-hybrid-access-03 (work in
             progress), October 2018.

   [RFC6733] Fajardo, V., Ed., Arkko, J., Loughney, J., and G. Zorn,
             Ed., "Diameter Base Protocol", RFC 6733,
             DOI 10.17487/RFC6733, October 2012,
             <https://www.rfc-editor.org/info/rfc6733>.

   [RFC6824] Ford, A., Raiciu, C., Handley, M., and O. Bonaventure,
             "TCP Extensions for Multipath Operation with Multiple
             Addresses", RFC 6824, DOI 10.17487/RFC6824, January 2013,
             <https://www.rfc-editor.org/info/rfc6824>.

   [TR23.793]
             3GPP, "Study on access traffic steering, switch and
             splitting support in the 5G System (5GS) architecture",
             December 2018.

Authors' Addresses

   Markus Amend
   Deutsche Telekom
   Deutsche-Telekom-Allee 7
   64295 Darmstadt
   Germany

   Email: Markus.Amend@telekom.de

   Eckard Bogenfeld
   Deutsche Telekom
   Deutsche-Telekom-Allee 7
   64295 Darmstadt
   Germany

   Email: Eckard.Bogenfeld@telekom.de

   Anna Brunstrom
   Karlstad University
   Universitetsgatan 2
   651 88 Karlstad
   Sweden

   Email: anna.brunstrom@kau.se

   Andreas Kassler
   Karlstad University
   Universitetsgatan 2
   651 88 Karlstad
   Sweden

   Email: andreas.kassler@kau.se

Veselin Rakocevic
City University of London
Northampton Square
London
United Kingdom

Email: veselin.rakocevic.1@city.ac.uk