    Initial-Path Selection for Connection Establishment in Multipath TCP
              draft-kang-mptcp-initial-path-selection-00

Abstract

   This draft presents an optimal mechanism for the selection of initial
   path to address the problem caused by the sequential subflow
   establishment defined in current MPTCP. We propose the general
   architecture and procedures, along with the background analysis, use
   cases and results of this proposal.

Status of this Memo

Copyright and License Notice

Table of Contents

1. Introduction

   Multipath TCP enables hosts to exchange packets belonging to a single
   connection over several paths. Usually, these paths correspond to
   different networks (such as cellular and WiFi) relying on the network
   interfaces configured on a terminal. MPTCP initializes multiple
   connections in sequence, that is, when one subflow is set up in the
   manner of a three-way handshake, being as the initial subflow,
   additional subsequent subflows can be created by a special four-way
   handshake. These subflows are bound to MPTCP session. The data on the
   sender can be selected for transmission on one of these subflows or
   on all the subflows through the scheduler.

   The path or network interface for the initial subflow is configured
   by default for all connections in one system. For example, if LTE and
   WiFi are offered in mobile phone, WiFi is the default path for the
   initial subflow generally.

Practically, this design falls short in some situations where the
default path goes through a highly lousy link. For example, if the
WLAN signal is weak or broken, the establishment of the subflow will
be delayed due to continuous tries of the master flow. This, in turn,
will be translated into longer startup delay for the initial
conversations of the applications above.

A similar problem was addressed by [Kien][Markus], via robust and
simultaneous connection establishment. However, they either require
the problem changes or introduce additional burden on the server
side. To avoid such problem, we propose a new solution by introducing
a function called "Initial-path Selection" during MPTCP connection
establishment, i.e., selecting the initial path based on measurement
of available paths.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
document are to be interpreted as described in RFC 2119 [RFC2119].

3. Architecture

Figure 1 provides the architecture for "Initial-path Selection". It
can be integrated in MPTCP stack or as an isolated module in the
terminal. It obtains transmission status information for each
available path when it receives a request from an application. Then
it makes decision on initial path to MPTCP stack to establish
connection with remote host.

```
+-------------------+                        +-------------------+
|      Terminal     |                        |       Server      |
| +--------------+  |                        | +--------------+  |
| |Application n |  |                        | |Application n |  |
| +--------------+  |                        | +--------------+  |
|        |          |                        |        |          |
|        |          |                        |        |          |
| +--------------+  |                        |        |          |
| | Initial-path +--+-----------+            |        |          |
| |  Selection   |  |           |            |        |          |
| +--------------+  |     +------+----+       |        |          |
|        |          | ----|           |----   |       |          |
|        |          |     |  Internet  |      |        |          |
| +--------------+  | ----|           |----   +--------------+  |
| |  MPTCP stack |  |     +-----------+        | |  MPTCP stack | |
| +--------------+  |                        | +--------------+  |
+-------------------+                        +-------------------+
```
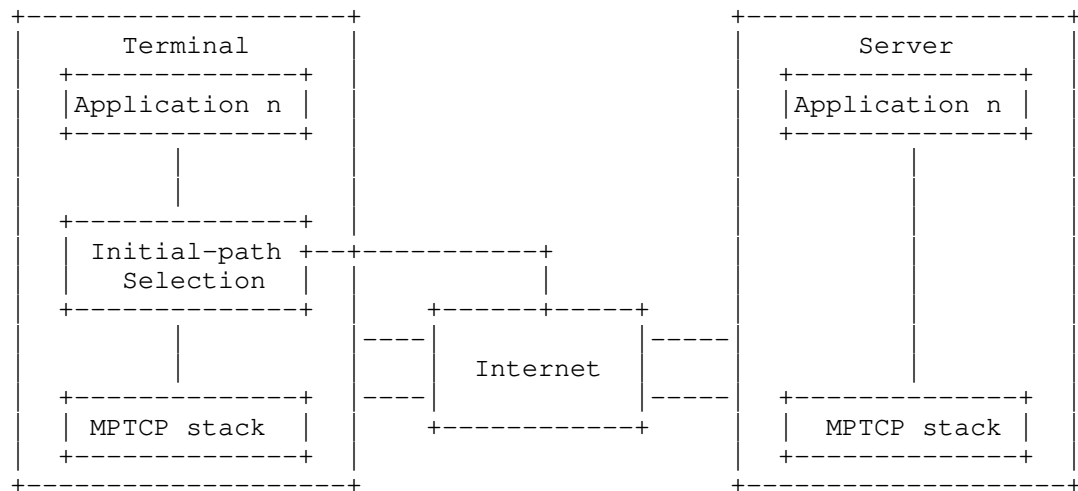
             Figure 1: Architecture for "Initial-path Selection"

4. Typical Flows

   Two typical flows are listed here. Figure 2 shows that "Initial-path
   Selection" will be executed for each MPTCP connection establishment.
   Figure 3 describes that "Initial-path Selection" can be used from the
   time that an application requests to set up a second MPTCP connection
   after the first one is completed. Considering that no heuristics is
   used for decision for the first MPTCP connection, default initial
   path can be adopted. This is just one of possible specific scenarios
   for this method.

```
                        +---------------+
                        |  Application  |
                        |   Request     |
                        +---------------+
                                |
                                V
                        +---------------+
                        | Initial-path  |
               +--->    |   Selection   |<--+
               |        +---------------+   |
               |                |           |
               |                V           |
               |        +---------------+   |
               |        | Set initial   |   |
               |        |    path       |   |
               |        |  for MPTCP    |   |Info
               |        +---------------+   |
               |                |           |
               |                V           |
               |        +---------------+   |
               |        | Establish MPTCP|---+
               |        |  connection   |
               |        +---------------+
               |                |
               |                V
               |No      +---------------+
               +----+   | Successfully? |
                        +---------------+
                                |Yes
                                V
```
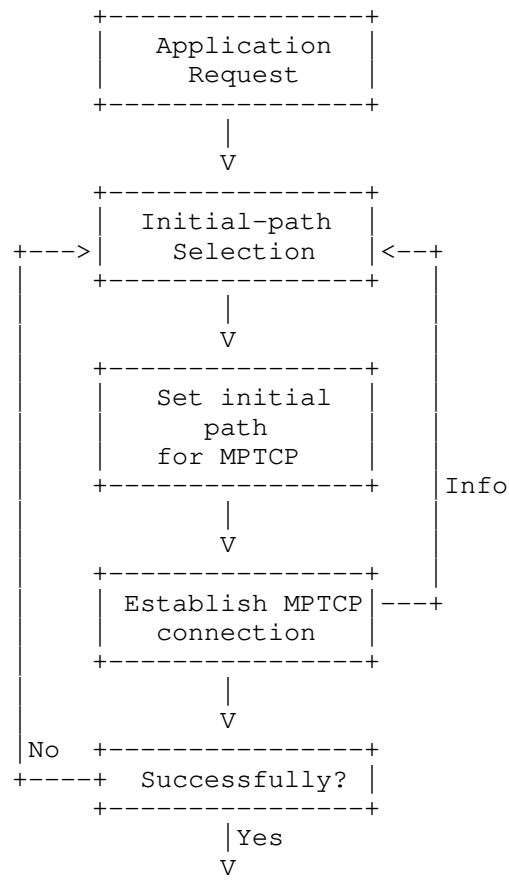
Figure 2: "Initial-path Selection" for each connection establishment

```
                    +--------------+
                    |  Application |
                    |   Request    |
                    +--------------+
                           |
                           V
                    +--------------+  Yes
                    |    First     +-------------+
                    | Connection?  |             |
                    +--------------+             |
                          |No                    |
                          V                      V
                    +--------------+      +-------------+
                    |   Initial-   |      |   Default   |
            +------>|              |<--+  |   initial   |
            |       |path Selection|   |  |    path     |
            |       +--------------+   |  +------+------+
            |              |           |         |
            |              V           |         |
            |       +--------------+   |         |
            |       | Set initial  |   |Info     |
            |       |path for MPTCP|   |         |
            |       +--------------+   |         |
            |              |           |         |
            |              V           |         |
            |       +--------------+   |         |
            |       |Establish MPTCP|---+         |
            |       |  connection  |             |
            |       +--------------+             |
            |              |                     |
            |              V                     |
            |   No  +--------------+             |
            +-------+ Successfully?|<------------+
                    +------+-------+
                          |Yes
                          V
```
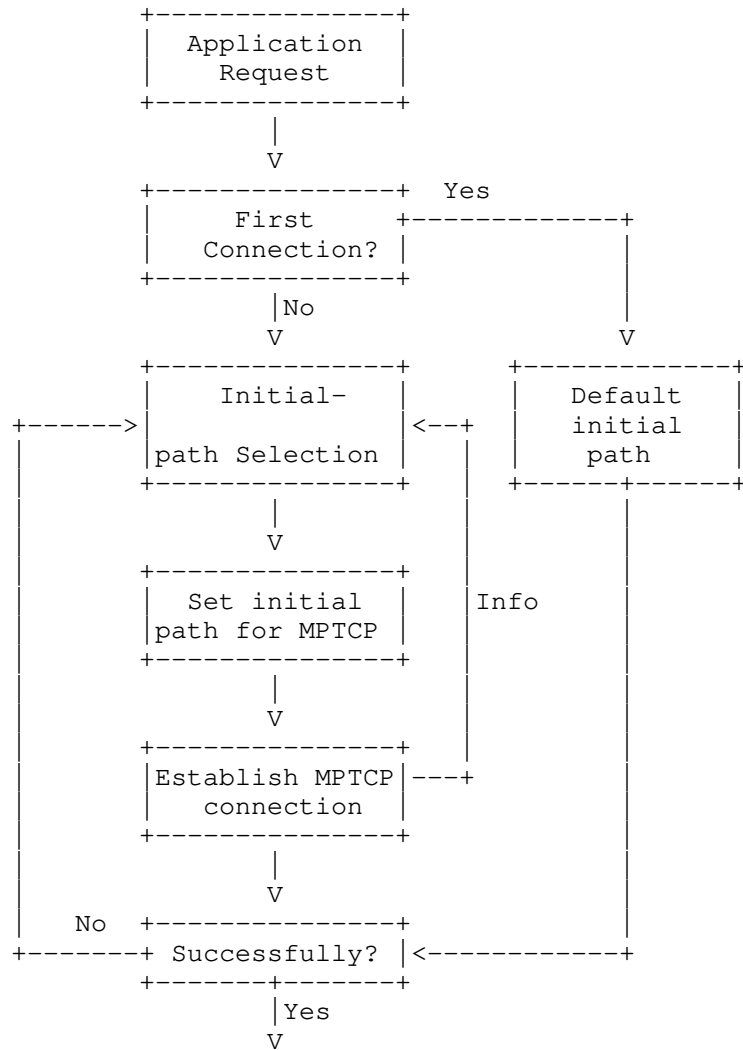
        Figure 3: Default initial path for the first connection,
     "Initial-path Selection" for subsequent connection establishment

   Figure 4 shows the abstract implementation process for "Initial-path
   Selection". Upon a request from an application, "Initial-path
   Selection" module will acquire transmission status information which
   represents the transmission performance of each available path or
   network interface and evaluate it. The transmission status
   information is characterized by at least one of the parameters:

signal strength, throughput, round-trip time (RTT) and link success
rate. In this way, we can determine a path with better transmission
performance and use the network interface to it for connection
establishment.

```
                                   |
                                   V
                  +-----------------------+
                  | Acquire transmission  |
                  |     status info for   |
                  |    available paths    |
                  +-----------------------+
                              |
                              V
                  +-----------------------+
                  | Evaluating the status |
                  |   for available paths |
                  +-----------------------+
                              |
                              V
                  +-----------------------+
                  |    Determining an     |
                  |   available path with |
                  |   better transmission |
                  |      performance      |
                  +-----------------------+
                              |
                              V
                  +-----------------------+
                  |   Using the network   |
                  |       interface       |
                  |  corresponding to the |
                  |    path with better   |
                  |      transmission     |
                  |     performance for   |
                  |       connection      |
                  |     establishment     |
                  +-----------------------+
                              |
                              V
```
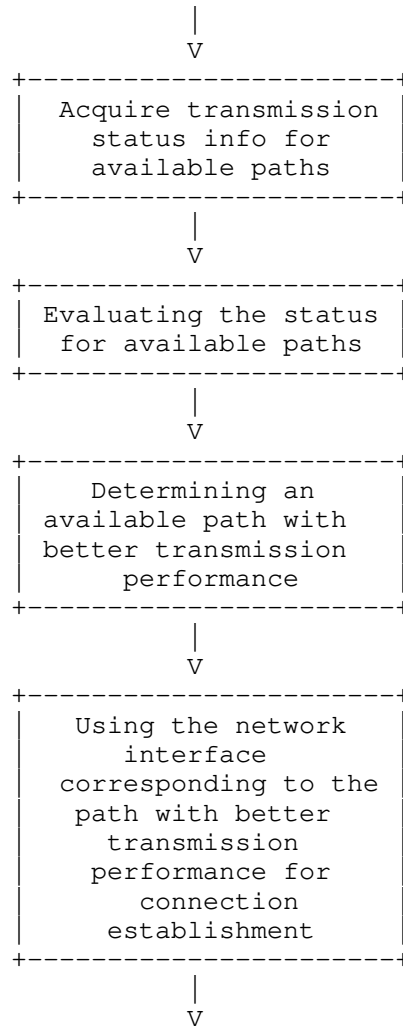
       Figure 4: Implementation process for "Initial-path Selection"

5. Possibilities Analysis

Figure 5 describes the relationship of application, MPTCP connection
and network interface at a running time. It shows that:

o Multiple applications are running on the terminal at the same
time.

o Each application can start and maintain more than one MPTCP
connection.

o Applications share network interfaces, e.g. WiFi, LTE and 5G.

So before starting a new MPTCP session, a number of heuristics can be
integrated to estimate the performance of these existing connections.
The estimation results are used to select the best one from available
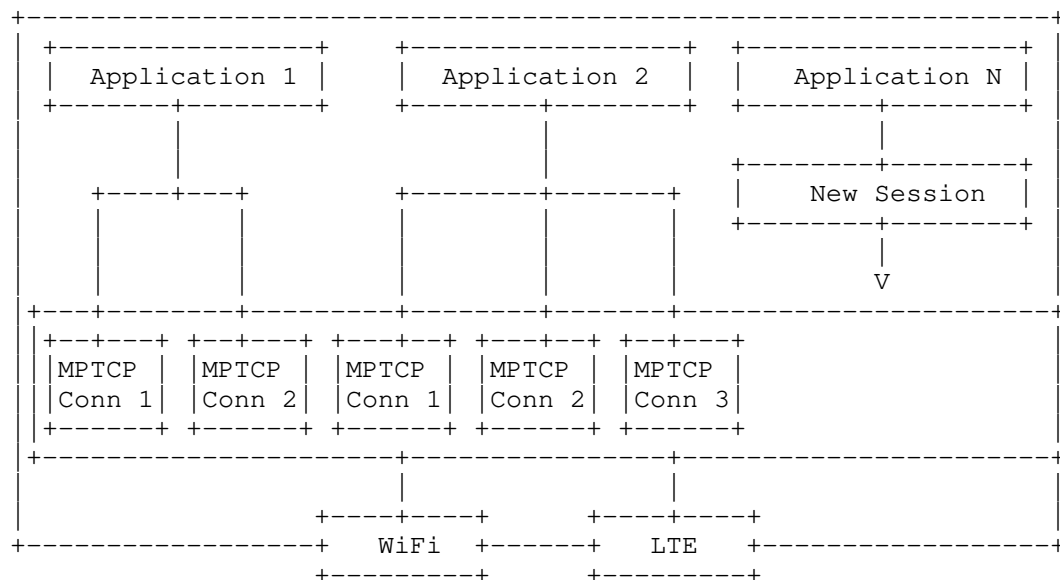paths which are bound to different network interfaces.

```
+------------------------------------------------------------------+
| +----------------+   +----------------+   +----------------+ |
| | Application 1  |   | Application 2  |   | Application N  | |
| +-------+--------+   +-------+--------+   +-------+--------+ |
|         |                    |                    |          |
|         |                    |            +-------+--------+ |
|    +----+---+           +--------+-----+  | New Session  | |
|    |        |           |        |     |  +-------+--------+ |
|    |        |           |        |     |          |          |
|    |        |           |        |     |          V          |
| +--+--------+---------+-------+------+----------------------+ |
| | +--+---+ +--+---+ +---+--+ +---+--+ +--+---+              | |
| | |MPTCP | |MPTCP | |MPTCP | |MPTCP | |MPTCP |              | |
| | |Conn 1| |Conn 2| |Conn 1| |Conn 2| |Conn 3|              | |
| | +------+ +------+ +------+ +------+ +------+              | |
| +--------------------+--------------+----------------------+ |
| |                    |              |                      |
| |              +----+----+     +----+----+                 |
| +-----------------+  WiFi +------+  LTE  +-----------------+ |
|                   +--------+     +--------+                  |
```

                Figure 5: Relationship of application, MPTCP connection and network
                                    interface

6. Path decision information

Such heuristics can be mainly divided into three levels: application
level, transport-layer level and network-interface level based on the
information acquisition method. For example, RTT is calculated for

each subflow in one MPTCP connection, so it is one of transport-layer
level. The transmission status information for each available path
SHOULD be characterized by at least one of the parameters: signal
strength, throughput, RTT and link success rate. Information in
application level is used for statistics.

o Application level: application name, domain name, port number and
location.

o Transport-layer level: RTT for each subflow within one MPTCP
connection in running applications.

Figure 6 shows a flowchart for "Initial-path Selection" based on RTT.
Normally, path RTT can be determined by a time difference between
sending a package and an ACK. For the asymmetric downloading service,
shown in figure 7, the latest RTT for these subflows is calculated by
data sender, i.e. application server (RTT = Timestamp 2 - Timestamp
1) and transformed through some option in TCP Header to data
receiver. Besides "latest RTT", algorithm for estimating the mean RTT
by more than one RTT sent from data sender is also proposed.

```
              +-------------------+
              |    New Session    |
              +-------------------+
                       |
                       V
        +-------------------+ No
        | Running connections+-------------+
        | (LTE.RTT<WiFi.RTT) |             |
        +-------------------+             |
                  | Yes                   |
                  V                       V
        +-------------------+  +-------------------+
        |     Set LTE as    |  |    Set WiFi as    |
        |    initial path   |  |    initial path   |
        +-------------------+  +-------------------+
```
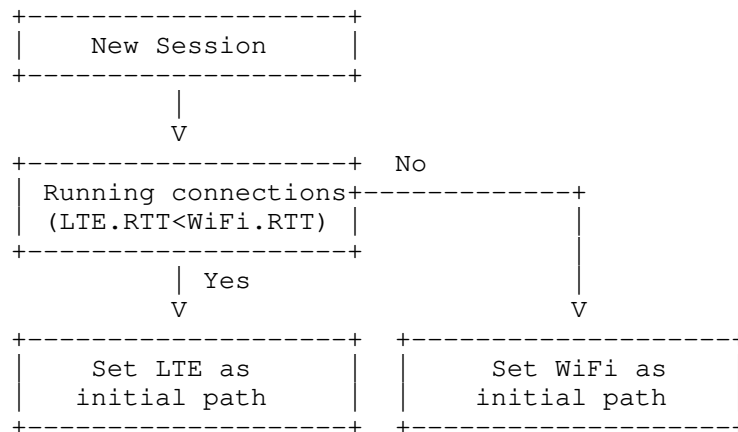
Figure 6: "Initial-path Selection" based on RTT

```
    +-----------------+                      +-----------------+
    |+---------------+|                      |+---------------+|
    ||  APP Client   ||                      ||  APP Server   ||
    |+---------------+|                      |+---------------+|
    |+---------------+|                      |+---------------+|
    ||  MPTCP Stack  ||                      ||  MPTCP Stack  ||
    |+---------------+|                      |+---------------+|
    +--------+--------+                      +--------+-------+
             |                                        |
             |  Establish a transport layer connection|
             |<-------------------------------------->|
             |                                        |
             |          Data Transmission             |
             |<---------------------------------------| Timestamp 1
             |                                        |
             |                ACK                     |
             |--------------------------------------->| Timestamp 2
             |                                        |
             |                       +-------+-------+
             |                       |Calculating RTT|
             |                       +-------+-------+
             |                                        |
             |       Data transmission + RTT          |
             |<---------------------------------------|
             |                ACK                     |
             |--------------------------------------->|
             |                                        |
```
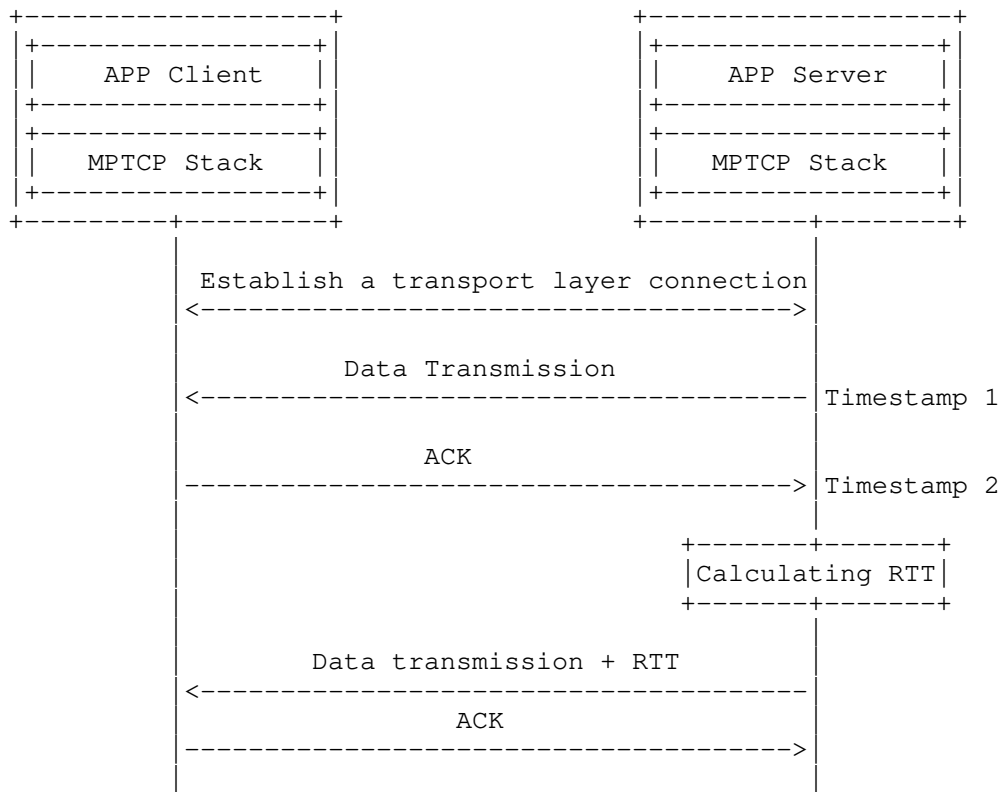
   Figure 7: RTT calculated by APP Server in asymmetric HTTP downloading

      o Network-interface level: signal strength, throughput and link
      success rate.

      Some parameters can be detected by the terminal, such as signal
      strength, throughput and RTT.

7. Examples

   Examples of actual implementation are as follows and default initial
   path is set to WiFi:

   o If WiFi.signal < LTE.signal, then set initial path = LTE.

   o If Running connections (LTE.latestRTT < WiFi.latestRTT), then set
   initial path = LTE.

o If WiFi.Throughput < LTE.Throughput, then set initial path = LTE.

Others are still in research:

o Calculating success rate by Probability Theory. For example, a geographical solution can be: if Location A (WiFi.SuccessRate < LTE.SucessRate), then set initial path = LTE.

8. Use Cases

This new proposal is effective for following actual service scenarios:

Scenario 1:

For some HTTP download service, the application downloads or plays videos through multiple short streams (<20MB per stream and 5MB for some short videos), see in Figure 8. The process of establishing a MPTCP connection frequently occurs during the use of the application. This proposal can reduce video start delay significantly in the scenario of "frequent connection establishment".

```
+-------------------------------------------------------------+
| +---------------------------------------------------------+ |
| |                      Application                        | |
| +-------+------------+-------------+------------+--------+ |
|    |         |              |           |         |         |
|    |         |              |           |    +------+-------+
|    |         |              |           |    | New Session  |
|    |         |              |           |    +------+-------+
|    |         |              |           |          |
|    |         |              |           |          V
| +-------+------------+-------------+--------------------+ |
| | +-----+-----+ +-----+-----+ +------+----+            | |
| | |   MPTCP   | |   MPTCP   | |   MPTCP   |            | |
| | |Connection1| |Connection2| |Connection3|            | |
| | +-----------+ +-----------+ +-----------+            | |
| +------------------+--------------+--------------------+ |
|    |               |              |                      |
|    |         +-----+-----+  +------+------+              |
| +------------+   WiFi    +---+     LTE    +--------------+
|              +-----------+  +------------+
```
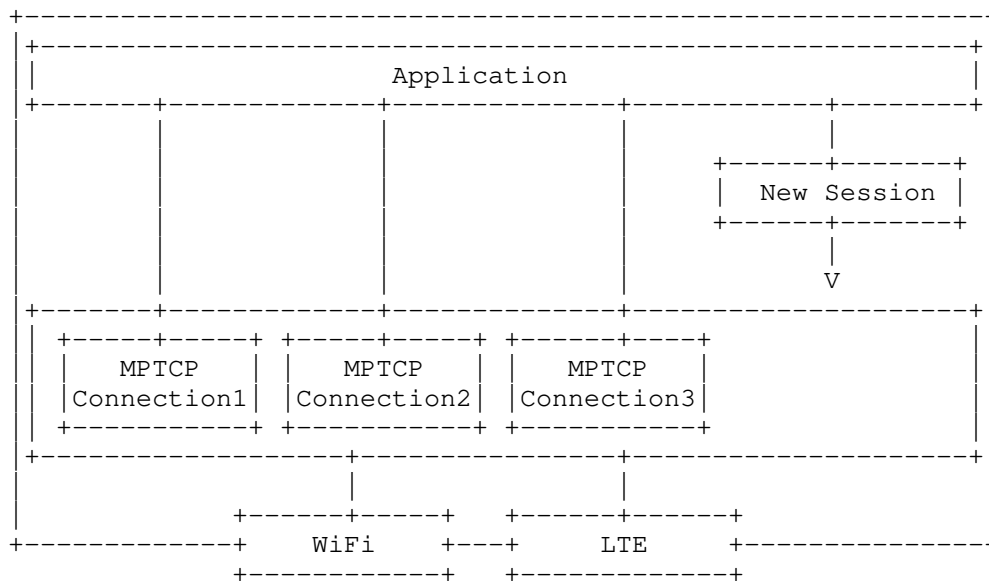
Figure 8: Scenario of "frequent connection establishment"

Scenario 2:

Multiple applications start and are running on a terminal at the same
time, see in Figure 9. Each application has the requirement to
establish an MPTCP connection. Therefore, there is an opportunity on
the terminal to use both WiFi and the LTE. This proposal can help
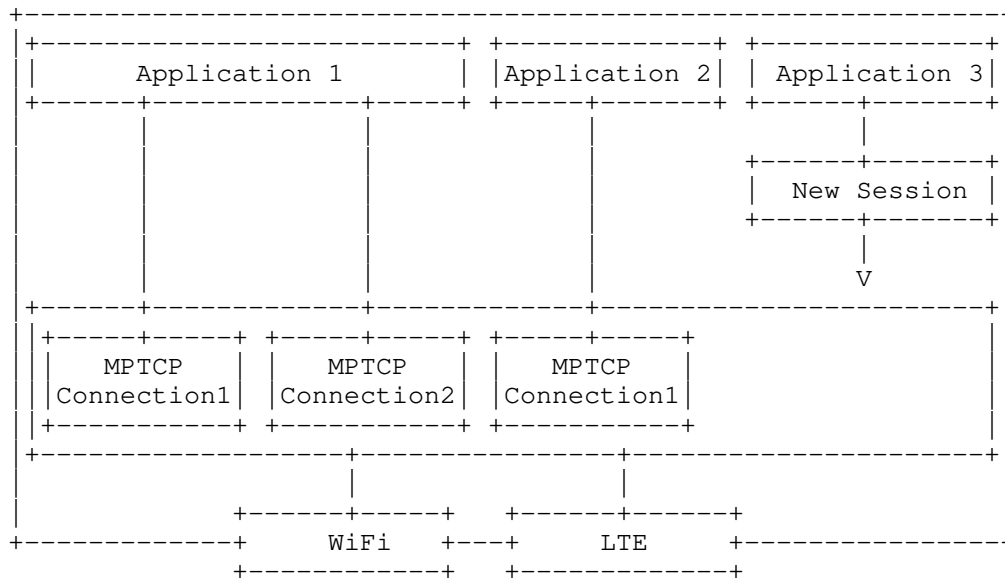load balancing between them.

```
+---------------------------------------------------------------+
| +---------------------------+ +-------------+ +-------------+ |
| |        Application 1       | |Application 2| | Application 3| |
| +------+------------+-----+ +-----+-------+ +------+-------+ |
|        |            |             |                |           |
|        |            |             |         +------+-------+    |
|        |            |             |         | New Session |    |
|        |            |             |         +------+-------+    |
|        |            |             |                |           |
|        |            |             |                V           |
| +------+------------+-------------+----------------------------+ |
| | +-----+-----+ +-----+-----+ +-----+-----+                  | |
| | |   MPTCP   | |   MPTCP   | |   MPTCP   |                  | |
| | |Connection1| |Connection2| |Connection1|                  | |
| | +-----------+ +-----------+ +-----------+                  | |
| +-----------------+---------------+----------------------+ | |
|                   |               |                        | |
|           +------+-----+   +------+------+                 | |
| +-------------+  WiFi   +---+    LTE       +---------------+ |
|              +-----------+   +-------------+                 |
```

                Figure 9: Network interface shared by multiple applications

9. Result

   For signal strength, in the case where the signal strength of default
   path (i.e. WiFi) is weak, and the LTE signal is strong, two
   conclusions can be drawn from testing:

   1. The proposed mechanism selects LTE as the initial path, instead of
   the default path over WiFi.

   2. The proposed mechanism can achieve the same effect on start-up
   delay as the case of turning on LTE only.

10. Compatibility Consideration

   This solution makes no changes to the existing flows in MPTCP base
   protocol because all its process is before connection establishment.
   But for asymmetric HTTP downloading, only data sender can provide the

accurate RTT data for reference, so it will result in the
modification in protocol parameter structure to carry it to data
receiver as above analysis.

11. Security Considerations

"Initial-path Selection" module is located at the terminal. The
design is to improve the robustness and reliability of connection
establishment, so there are no security issues.

12. IANA Considerations

Considering the asymmetric downloading service, a new kind of option
in TCP Header SHOULD be introduced for carrying RTT from data sender
to date receiver.

13. References

13.1. Normative References

    [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119, March 1997.

    [RFC6182] Ford, A., Raiciu, C., Handley, M., Barre, S., and
              J.Iyengar, "Architectural Guidelines for Multipath TCP
              Development", RFC 6182, DOI 10.17487/RFC6182, March 2011,
              <https://www.rfc-editor.org/info/rfc6182>.

    [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,
              <http://www.rfc-editor.org/info/rfc6824>.

13.2. Informative References

    [Kien]   Kien Nguyen, Kentaro Ishizu, Mirza Golam Kibria, and
              Fumihide Kojima, "A proposal for improving MPTCP
              initialization". 14 November 2016, Seoul, Korea,
              <https://www.ietf.org/proceedings/97/slides/slides-97-
              mptcp-a-proposal-for-improving-mptcp-initialization-00.pdf>

    [Markus] Markus Amend, Eckard Bogenfeld, and Andreas Matz, "A
              proposal for MPTCP Robust session Establishment (MPTCP
              RobE)", 18 July 2017,
              <https://datatracker.ietf.org/meeting/99/materials/slides-
              99-mptcp-a-proposal-for-mptcp-robust-session-establishment-
              mptcp-robe-01.pdf>

   [RobE]   Markus Amend, Veselin Rakocevic, Andreas Philipp Matz, and
               Eckard Bogenfeld, "RobE: Robust Connection Establishment
               for Multipath TCP", ANRW '18, July 16, 2018.

Author's Addresses

Jianjian Zhu
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R. China

Email: zhujianjian1@huawei.com

Jiao Kang
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R. China

EMail: kangjiao@huawei.com

Fanzhao Wang
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R. China

Email: wangfanzhao@huawei.com

Tong Li
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R. China

EMail: li.tong@huawei.com

Kai Zheng
Huawei Technologies
Information Road, Haidian District,
Beijing 100085
P.R. China

EMail: kai.zheng@huawei.com