MPTCP F. Song
Internet-Draft H. Zhang
Intended status: Informational Beijing Jiaotong University
Expires: September 14, 2017 H. Chan
A. Wei
Huawei Technologies
March 13, 2017

One Way Latency Considerations for MPTCP
draft-song-mptcp-owl-02

Abstract

This document discusses the use of One Way Latency (OWL) for enhancing multipath TCP (MPTCP). Several use cases of OWL, such as retransmission policy and crucial data scheduling are analyzed. Two kinds of OWL measurement approaches are also provided and compared. More explorations related with OWL will be helpful to the performance of MPTCP.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Both end hosts and the intermediate devices in the Internet have basically been equipping with more and more physical network interfaces. Whereas multiple interfaces had been widely used in packet forwarding, traffic engineering, etc., the importance of these interfaces at the end hosts had been confirmed and utilized [RFC6419]. Moreover, the increased capacity provided by the multiple paths created by multiple interfaces is leveraged to aggregate more bandwidths, to decrease packet delay and to provide better services. Unlike traditional TCP [RFC0793], many transport layer protocols, such as MPTCP [RFC6182] [RFC6824] enable the end hosts to concurrently transfer data on top of multiple paths to greatly increase the overall throughput.

Round-trip time (RTT) is commonly used in congestion control and loss recovery mechanism for data transmission. Yet the key issue for data transmission is simply the delay of the data transmission along a path which does not include the return. The latency for uplink and downlink between two peers may be very different. RTT, which cannot accurately reflect the delay of the data transmission along a path, can be easily influenced by the latency in the opposite direction along that path. Therefore, the use of One Way Latency (OWL) is proposed to describe the exact latency from the time that data is sent to the time data is received.

This document explains that the performance of current practices of MPTCP can be further improved by fully taking advantage of One Way Latency (OWL) during the transmission. The OWL components in the forward and reverse directions of a RTT may be asymmetric so that it can provide a better measure to the user such as for congestion control even with the regular TCP. The benefits will be more when there are multiple paths to choose from.

This document discusses the necessary considerations of OWL in MPTCP. The structure of this document is as follows: Firstly, several use cases of OWL in MPTCP are analyzed. Secondly, two kinds of OWL measurements are listed and compared. The considerations related with security and IANA are given at the end.

The potential targeted audience of this document are application programmer whose products may significantly benefit from MPTCP. This document also provides the necessary information for the developers of MPTCP to implement the new version API into the TCP/IP network stack.

2. Conventions and Terminology

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

One Way Latency (OWL): the propagation delay between a sender and a receiver from the time a signal is sent to the time the signal is received.

3. Potential Usages of OWL in MPTCP

There are a number of OWL use cases when MPTCP is enabled by the sender and receiver. . Although only 5 use cases are illustrated in this document, more explorations are still needed.

3.1. Crucial Data Scheduling

During a transmission process, there are often some crucial data that need to be immediately sent to the destination. Examples of such data include the key frame of multimedia and high priority chunk of emergency communication. One cannot guarantee the arrival sequence by using the RTTs alone of the multiple paths.

The data rate in any given link can be asymmetric. In addition, the delay in a given direction can change according to the amount of packet queue. Therefore delay in a forward direction in a path is not necessarily the same as that in the reverse direction as exemplified in Figure 1.

```
     --------   OWL(s-to-c,path1)=16ms   <--------
   /                                               \
  |     ----->  OWL(c-to-s,path1)= 5ms    -----     |
  |   /            RTT(path1)=21ms            \    |
  |  |                                          |  |
+-------+                                    +-------+
|       |----->  OWL(c-to-s,path2)= 8ms    -----|       |
|Client|                                     |Server|
|       |-----   OWL(s-to-c,path2)= 8ms   <-----|       |
+-------+           RTT(path2)=16ms          +-------+
  |  |                                          |  |
  |   \                                        /   |
  |     ----->  OWL(c-to-s,path3)=10ms    -----     |
   \                                               /
     --------   OWL(s-to-c,path3)= 8ms   <--------
                   RTT(path3)=18ms
```

Figure 1. Example with 3 paths between the client and the server with OWL as indicated in the figure. RTT information alone would indicate to the client that the fastest path to the server is path 2, followed by path 3, and then followed by path 1. path 2 is the fastest, whereas OWL indicates to the client that the fastest path to the server is path 1, followed by path 2, and then followed by path 3.

Using the results of OWL measurement, the sender can easily select the faster path, in terms of forward latency, for crucial data transmission. Moreover, the acknowledgements of these crucial data could be sent on the path with minimum reverse latency. Piggyback is also useful when duplex communication mode is adopted.

3.2. Congestion Control

Congestion in a given direction does not necessarily imply congestion also in the reverse direction.

```
        --------   No congestion (path 1)   <--------
      /                                              \
     |     ----->  Congestion    (path 1)    -----    |
     |   /                                        \   |
     |  |                                          |  |
   +-------+                                    +-------+
   |Client|                                     |Server|
   +-------+                                    +-------+
     |  |                                          |  |
     |   \                                        /   |
     |     ----->  No congestion (path 2)    -----    |
      \                                              /
        --------   Congestion    (path 2)   <--------
```

Figure 2. Example of a congestion situation with 2 paths between the client and the server. There is congestion from client to server along path 1 and also from server to client along path 2. RTT information alone will indicate congestion in both paths, whereas OWL information will show the client that path 2 is the more lightly loaded path to get to the server.

Network congestion in a given direction can be better described using OWL rather than using RTT. Especially when the congestion can be a situation in a unidirectional path, the congestion in the path from a client to a server is different from the congestion in the path from the server to the client. The RTT cannot accurately reflect the delay of interest for data transmission along a path. For MPTCP, the client needs to choose a more lightly loaded path to send packets [RFC6356]. It will then be unwise to compare the RTT among different paths, and it should instead compare the OWL among the paths.

Current version of MPTCP includes different kinds of congestion control mechanisms [RFC6356]. By reasonably utilizing OWL, the network congestion situation in a single direction could be better described.

3.3. Packet Retransmission

Continuous Multipath Transmission (CMT) increases throughput by concurrently transferring new data from a source to a destination host via multiple paths. However, when packet is identified as lost by triple duplicated acknowledgements or timeout, the sender needs to select a suitable path for retransmission. Due to the popular mechanisms of sequence control in reliable transport protocols, outstanding packets on multiple paths may reach to the destination disorderly and trigger Receive Buffer Blocking (RBB) problem (Figure 3) which will further affect the transmission performance.

```
              Packet with octets sequence #   0- 499(lost)
       -----> Packet with octets sequence #1000-1499(rcvd) ------
      /       Packet with octets sequence #2000-2499(rcvd)       \
     |                                                            |
  +-------+                                                  +--------+
  |Sender|                                                   |Receiver|
  +-------+                                                  +--------+
     |                                                            |
      \       Packet with octets sequence # 501- 999(lost)       /
       -----> Packet with octets sequence #1501-1999(lost)  -----
              Packet with octets sequence #2501-2999(lost)
```

Figure 3. Example of Receive Buffer Blocking: The packet containing octets 0-499 is lost. On the other hand the packets containing Octets 500-999, 1000-1499, 1500-1999, 2000-2499, and 2500-2999 have all been received. The octets 500-2999 are then all buffered at the receiver, and are blocked by the missing octets 0-499.

Using the results of OWL measurement, the sender can quickly determine the specific path with minimum latency in the forward direction. RBB can be relieved as soon as the receiver obtains the most needed packet(s) and submits them all to the upper layer.

3.4. Bandwidth Estimation

Understanding the bandwidth condition is beneficial for data packet scheduling, and load balancing, etc. OWL could be integrated with bandwidth estimation approaches without interrupting the regular transmission of packets.

3.5. Shared Bottleneck Detection

Fairness is critical especially when MPTCP and ordinary TCP coexist in the same network. The sender could treat OWL measurements as the sample process of shared bottleneck detection and accordingly adjust the volume of data packet on multiple paths.

4. OWL Measurements in TCP

The timestamp option in TCP [RFC7323] may be invoked to estimate latency. When sending data, the time (TSval) of sending the data is provided in the option. The receiver acknowledges the receipt of this data by echoing this time (TSecr) and also provides the time (TSval) of sending this acknowledgment. The difference of these times in the acknowledgment of data from the sender can help to estimate the OWL from the sender to the receiver. There are two problems though.

First, there may be delay from the time the receiver has received the data to the time the acknowledgment is sent. The above number may then be an upper bound of OWL.

Second, the clocks may not be synchronized between the sender and the receiver. The above measure can show the OWL in different paths only if the clocks synchronized. Without clock synchronization, the comparison of OWLs among different paths is limited to showing the OWL differences among them.

Two kinds of OWL measurement approaches are available: absolute value measurement and relative value measurement.

To obtain the absolute value of OWL, the primary condition of measurement is clock synchronization. Using Network Time Protocol (NTP) [RFC5905], end hosts can calibrate the local clock with the remote NTP server. The additional information or optional capabilities can even be added via extension fields in the standard NTP header [RFC7822]. The calibration accuracy can reach to the millisecond level in less congested situations. The obvious burden here is to persuade the end hosts to initialize the NTP option.

Obtaining the relative value of OWL is more than enough in some circumstances to establish applications on top of it. When retransmission is needed, for example, the sender may only care about which path has the minimum forward latency. When bandwidth is being estimated, the difference of forward latency, i.e. delta latency, among all available paths is needed. By exchanging with correspondent end host the local timestamps of receiving and sending the packets, both sides could obtain the relative value of OWL.

While absolute value measurement of OWL is more convenient for the applications, the overheads are the extra protocol requirement and synchronization accuracy. On the contrary, relative value measurement does not need to worry about the accuracy whereas the overhead is to add timestamps into the original protocol stack.

5. Security Considerations

This document does not contain any security considerations. However, future applications of OWL in MPTCP will definitely need to establish relevant mechanisms to improve security.

6. IANA Considerations

This document presents no IANA considerations.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

7.2. Informative References

[RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, DOI 10.17487/RFC0793, September 1981, <http://www.rfc-editor.org/info/rfc793>.

[RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, DOI 10.17487/RFC5905, June 2010, <http://www.rfc-editor.org/info/rfc5905>.

[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, <http://www.rfc-editor.org/info/rfc6182>.

[RFC6356] Raiciu, C., Handley, M., and D. Wischik, "Coupled Congestion Control for Multipath Transport Protocols", RFC 6356, DOI 10.17487/RFC6356, October 2011, <http://www.rfc-editor.org/info/rfc6356>.

[RFC6419] Wasserman, M. and P. Seite, "Current Practices for Multiple-Interface Hosts", RFC 6419, DOI 10.17487/RFC6419, November 2011, <http://www.rfc-editor.org/info/rfc6419>.

[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>.

[RFC7323] Borman, D., Braden, B., Jacobson, V., and R. Scheffenegger, Ed., "TCP Extensions for High Performance", RFC 7323, DOI 10.17487/RFC7323, September 2014, <http://www.rfc-editor.org/info/rfc7323>.

[RFC7822] Mizrahi, T. and D. Mayer, "Network Time Protocol Version 4 (NTPv4) Extension Fields", RFC 7822, DOI 10.17487/RFC7822, March 2016, <http://www.rfc-editor.org/info/rfc7822>.

Authors' Addresses

Fei Song
Beijing Jiaotong University
Beijing, 100044
P. R. China

Email: fsong@bjtu.edu.cn

Hongke Zhang
Beijing Jiaotong University
Beijing, 100044
P. R. China

Email: hkzhang@bjtu.edu.cn

H Anthony Chan
Huawei Technologies
5340 Legacy Dr. Building 3
Plano, TX 75024
USA

Email: h.a.chan@ieee.org

Anni Wei
Huawei Technologies
Xin-Xi Rd. No. 3, Haidian District
Beijing, 100095
P. R. China

Email: weiannig@huawei.com