Network Working Group
Internet Draft
Intended status: Experimental
Expires: December 2017

Changqiao Xu
BUPT
Jiuren Qin
BUPT
Hongke Zhang
BUPT
Chunshan Xiong
Huawei Technologies Co., Ltd
Lei Zhu
Huawei Technologies Co., Ltd

June 21, 2017

# A Message-Oriented Extension to Multipath Transmission Control Protocol (MPTCP) draft-xu-mptcp-momp-04.txt

## Status of this Memo

The list of Internet-Draft Shadow Directories can be accessed at
http://www.ietf.org/shadow.html
This Internet-Draft will expire on December 3, 2016.



Abstract

This memo specifies a message-oriented extension for Multipath TCP (MPTCP) which aims to serve high-bandwidth and real-time applications. By introducing a message mapping to MPTCP, Message-Oriented MPTCP (MO-MPTCP) attaches some message features like boundaries, priority and dependency to bytestream. With such message-oriented information, MPTCP senders can avoid the waste of transmission resources and improve the transmission efficiency.

Table of Contents

## 1. Introduction

With the increasingly demands for bandwidth-intensive services, e.g., high-definition (HD) video, the streaming media data which is massive, ordered and delay-sensitive is becoming the main traffic of transport layer. Usually, the streaming media data is transferred by UDP (User Datagram Protocol) which performs better than TCP (Transmission Control Protocol) in improving throughput and reducing latency. However, UDP does not have congestion control mechanism and may result in network collapses.

MPTCP which has been standardized in [RFC6824] can greatly improve the throughput of one association by concurrently transferring data on several TCP subflows. Furthermore, the congestion control mechanism provided by MPTCP can make it work without starving TCP. With these advantages, MPTCP has the potential to serve the high-bandwidth and real-time applications.

This memo introduces a Message-Oriented MPTCP (MO-MPTCP) which is a light and scalable extension to standard MPTCP [RFC6824] and more suitable for streaming media transfer. MO-MPTCP specifies a message mapping to record the information about message boundaries, priority and dependency in the connection level. Based on this mapping, MO-MPTCP offers Boundary-Based Packet Scheduling Mechanism which can avoid unnecessary transmission and Message-Oriented Transmission Optimization which can preferentially ensure the transmission of important data.

## 2. Conventions

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. New Functionalities provided by MO-MPTCP

Making the transmission of stream media as an example, the new functionalities provided by Message-Oriented MPTCP are as follows:

o Boundary-Based Packet Scheduling

In the process of stream media transmission, application layer usually delivers the data to the transport layer frame by frame. Each frame can be seen as an individual message. However, in the transport layer, limited by Maximum Segment Size (MSS) MPTCP tends to segment the big messages and splice the small messages. And it also splice the small messages into one data packets to reduce

transmission overhead. In general, all data are transferred in the size of MSS. The segmentation and splicing operation of transport layer leads to lose the original message boundaries. MO-MPTCP provides a message mapping that can record the features of application messages including boundaries, priority and dependency, etc.. This mapping can help the sender to avoid unnecessary transmissions. For example, stream media can usually tolerate the loss of partial packets, which means the sender can give transmission up and notify the receiver when a packet is considered as time out. This kind of partially reliable mechanism can refer to [PRMP]. In this situation, if a packet which contains partial data of a frame data is abandoned by the sender, as a result, this frame cannot be decoded correctly at receiver side with the absence of partial information. Current MPTCP which is based on bytestream fails to perceive this situation, and still transmits the remaining data of this frame which is a waste of transmission resources. In Message-Oriented MPTCP, thanks to the recording of message boundary, senders can abandon the remaining data simultaneously and avoid unnecessary transmission.

o Message-Oriented Transmission Optimization

Traditional transmission ignores the priority and dependency of messages and treats them equally as a bytestream, which makes the transport blindly. Using an IPMH-like [IPMH] interface, MO-MPTCP can get the priority of each message, and record the dependency between them. For instance, in the standard MPEG coding, "I" frames are essential to the recovery of the whole images and can be decoded independently, so they have the "HIGH" priority and Dependency is "NULL". Similarly, "P" frames which are decoded based on a previous frame have "MEDIUM" priority and Dependency is "PRE". "B" frames which are decoded based on both a previous frame and a latter frame have "LOW" priority and Dependency is "PRE&LAR". Through some rules, TCP packets can determine their own priorities from the messages priorities. The reliability and timeliness of high-priority packets will be guaranteed first when congestion occurs. When a duplicate acknowledgement is received in the subflow level, the sender will execute judgment for the missing packet upon their priorities and duplicate ACK numbers. The send then will retransmit the packet if needed.

## 4. Message Mapping

MO-MPTCP sets up a Message Mapping in the connection level. The Message Mapping which is similar to the Data Sequence Mapping can associates message features such as boundary and priority with stream features such as DSN. This mapping which is the foundation of MO-MPTCP can provide useful information for data scheduling in

transmission.

The Message Mapping consists of a lot of records, and each record corresponds to an application message. Its structure sketch is show in Figure 1.

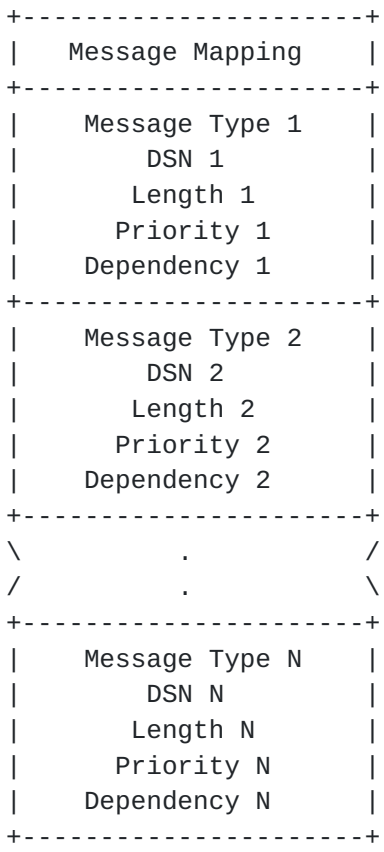


Figure 1 Message Mapping

- o Message Type is used to distinguish the classes of message. It can change its meaning depending on the application. For example, in the streaming media transmission, it represents which kind of frame this message is.

- o DSN=Data Sequence Number. DSN shows the Data Sequence Number of the first byte in an application message.

- o Length shows the number of bytes that this message contains. This parameter is usually used with DSN, and can identify the message boundaries.

- o Priority shows the importance of this message which usually be divided into three priority HIGH, MEDIUM, LOW.

- o Dependency shows the dependencies between adjacent messages. For example, "NULL" means this message is independent; "PRE" means this message depends on the previous message to be decoded; "LAT" means this message depends on the later message to be decoded. "PRE&LAT"means this message depends on both the previous and later messages to be decoded.

The DSN and Length are used to identify the boundary of an application message. And, the rest of the parameters which are unique nature of messages are used to provide information for the transmission optimization.

MO-MPTCP also provides rules for mapping establishment and removal as follows:

- o On receiving an application message, the sender SHOULD add a new record containing all necessary parameters to the Message Mapping. However these parameters may have different meaning for different applications.

- o When receiving an ACK in the MPTCP connection level, the sender SHOULD judge whether need to remove some records from the Message Mapping. Some messages are larger than the MSS, and may be partly acknowledged. MO-MPTCP provisions that the record for a message SHOULD be retained until all segments of this message are acknowledged.

## 5. Operations of MO-MPTCP

### 5.1. Boundary-Based Packet Scheduling

Boundary-Based Packet Scheduling is used in the situations where the applications can tolerate the loss of some packages to meet its requirements for timeliness. [PRMP] proposed a partially reliable extension to MPTCP called PR-MPTCP, which is designed to deal with above situations. However, PR-MPTCP is based on the bytestream and can perform better with the help of MO-MPTCP. For instance, if a TCP packet containing partial data of a message is determined to be discarded, MO-MPTCP can find and discard the remaining data that belongs to or relies on this message. The detailed operating steps are as follows:

a) MO-MPTCP offers a function to the sender. When determining to discard a packet, the sender SHOULD call thi s function and send the starting DSN and length of this packet as parameters to MO-MPTCP.

b) Every time receiving calling from the sender, MO-MPTCP SHOULD search the Message Mapping and record all the messages involved in this packet.

c) Based on the messages selected by step b) MO-MPTCP then refers to the Dependency recorded in Message Mapping and extracts some other messages which rely on them to be decoded.

d) MO-MPTCP combines all the messages selected by step b) and c) and connects them as one or more bigger messages according to their DSNs and Length. Then the new boundaries of these messages are obtained.

e) MO-MPTCP SHOULD return the starting DSN and Length of these new messages. Then, the sender can continue its original operations and discard the expanded messages according to the new boundaries.

Step b) can be classified into the following situations:

- Only one message is involved in the packet, which means this packet is just a segment of the original message. In this case, MO-MPTCP SHOULD search the Message Mapping and record this message.

- Two or more messages are involved in the packet, which means this packet contains data comes from different messages. In this case, MO-MPTCP SHOULD search the Message Mapping and record all related messages.

When executing step c) there are some notes:

- Before starting to search the Message Mapping, MO-MPTCP preferably checks the priorities of the messages provided by step b) and skips the messages which have LOW priority. Because there is usually no message relies on them.

- Although the parameter of Dependency in Message Mapping only reflects the relationship between adjacent messages, the lost a message with HIGH priority can influence several messages with lower priority. For example, if an "I" frame is decided to be discarded, the following several frames will be influenced. So, the implementation should pay attention to a chain reaction.

### 5.2. Message-Oriented Transmission Optimization

The Message Mapping records the priorities of the messages. Based on these priorities, each TCP packet can determine its own priority. The basic rules are as follows:

- If the data of a packet comes from only one message, the packet priority is the same with the message priority.

- If the data of a packet comes from several messages, the packet priority is the same with the highest message priority.

Following the above rules, senders can obtain the packet priority, which is an important reference for the transmission optimization. The main operations of the optimization are as follows:

a) Once the sender receives duplicate acknowledgement, it SHOULD obtain the priority of those corresponding TCP packets by searching the Message Mapping.

b) MO-MPTCP determines whether these packets need being retransmitted immediately based on their priorities and the number of duplicate acknowledgments. The packets with HIGH priority will not be easily discarded;The packets with LOW priority will be discarded first when congestion occurs.

c) If a TCP packet is judged to need retransmission by step b) the senders SHOULD retransmit it immediately. Meanwhile, it SHOULD also reset retransmission timer and clear the number of duplicate acknowledgment.

d) If a TCP packet does not need to be transmitted after step b) the senders can continue their original works until event in step a) happens.

## 6. Implementation Consideration

In order to achieve message-oriented control and byte-oriented transport, MO-MPTCP records the message information in the Message Mapping. In the implementations of MO-MPTCP, hosts have to reserve some memory for Message Mapping, which brings additional cost. However, with the help of Message Mapping, more intelligent and efficient transmission can be achieved. And the additional cost is reasonable and tolerable.

## 7. Interface Considerations

MO-MPTCP offers an interface to the upper layer, through which the applications can call MO-MPTCP and assign the parameters like priority and dependency. The ways in which application obtain these parameters can refer to [IPMH].

## 8. Security Considerations

This memo develops no new security scheme for MPTCP. MO-MPTCP share the same security issues discussed in [RFC6824] with MPTCP.

## 9. IANA Considerations

There is no IANA consideration for this memo.

## 10. References

### 10.1. Normative References

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

[RFC6824] Ford, A., Raiciu, C., Handley, M., and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses", RFC 6824, January 2013.

### 10.2. Informative References

[PRMP] Changqiao Xu, H. Huang, H. Zhang, C. Xiong, L. Zhu "Multipath Transmission Control Protocol (MPTCP) Partial Reliability Extension? draft-xu-mptcp-prmp-02, September 2015.

[IPMH] E, Gineste M, Dairaine L, et al. Building self-optimized communication systems based on applicative cross-layer information. Computer Standards & Interfaces, 2009, 31(2): 354-361.

## 11. Acknowledgments

This Internet Draft is the result of a great deal of constructive discussion with several people, notably Man Tang, Hui Huang, and Peng Wang.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Changqiao Xu
Beijing University of Posts and Telecommunications
Institute of Network Technology, No. 10, Xitucheng Road,
Haidian District, Beijing
P.R. China

Email: cqxu@bupt.edu.cn

Jiuren Qin
Beijing University of Posts and Telecommunications
Institute of Network Technology, No. 10, Xitucheng Road,
Haidian District, Beijing
P.R. China

Email: jrqin@bupt.edu.cn

Hongke Zhang
Beijing University of Posts and Telecommunications
Institute of Network Technology, No. 10, Xitucheng Road,
Haidian District, Beijing
P.R. China

Email: hkzhang@bupt.edu.cn

Chunshan Xiong
Huawei Technologies Co., Ltd
Science and Technology Demonstration Garden, No. 156, Zhongguancun
North Qing Road,
Haidian District, Beijing
P.R. China

Email: sam.xiongchunshan@huawei.com

Lei Zhu
Huawei Technologies Co., Ltd
Science and Technology Demonstration Garden, No. 156, Zhongguancun
North Qing Road,
Haidian District, Beijing
P.R. China

Email: lei.zhu@huawei.com