Multipath TCP K. Nguyen
Internet-Draft K. Ishizu
Intended status: Standards Track M. Kibria
Expires: May 3, 2017 F. Kojima
National Institute of Information and Communications Technology
October 30, 2016

An Improvement of MPTCP Initialization
draft-kien-mptcp-init-00

Abstract

This draft describes a new method of connection initialization for Multipath TCP (MPTCP). In the current implementation, the MPTCP's first subflow needs to be successfully initialized before an additional flow takes its turn. This yields to a degradation of MPTCP benefit in many use cases (e.g., transferring short flows). To overcome the problem, we propose to duplicate the first SYN packet and send the duplicating ones via multiple interfaces.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 3, 2017.

Copyright Notice

Table of Contents

1. Introduction

MPTCP is an evolvable and efficient tool for link aggregation, e.g., on multi-homing hosts in mobile wireless networks. The flexibility of adding and deleting subflows introduce MPTCP benefits in aggregation and soft handover in many use cases. In the current implementation, MPTCP shows several drawbacks including operations in a scenario with imbalanced paths, especially handling short flows. Since a large amount of Internet TCP traffic is short flows, MPTCP should be improved to be more suitable with the traffic pattern. In the such scenario, the problem of selection of initialization path has big impacts on the MPTCP performance. It has been proven by the theoretical analysis [analysis] and real measurements [measurement]. However, there are limited works on solving the problem.

In fact, MPTCP can not choose the initial path itself. MPTCP relies on routing information to determine the destination for the initialization. The routing is static in most cases. The host is normally configured to route all the traffic through a default gateway. As a result, the first SYN of initialization has to be sent to the gateway associated network regardless of its quality. On the other hand, the routing information is available on a host for supporting beneficial operations of MPTCP. To solve the mentioned problem, we propose to duplicate the first SYN packet. The available routing information is leveraged in sending the duplications through several networks. The first received SYN/ACK is determined the best network (i.e., the one with the smallest RTT) to initialize the MPTCP connection.

2. Problem Description

This section describes the limitation of the current implementation of MPTCP. Consider an example scenario of MPTCP communication between two host (host A and host B). MPTCP on host A with multiple addresses (i.e., two addresses A1, A2) communicates with host B via network A1, A2, respectively. In this scenario, the gateway associated with address A1 is the default. Obviously, Host A send the first SYN with MP_CAPABLE to Address B1 for MPTCP initialization. After the successful initialization, the additional subflow will be added to ongoing MPTCP transmissions following one of two methods. The later subflow is initialized a new SYNC+MP_JOIN from A2 to B1 if there is no NAT between them. In case of under NAT, the SYN+MP_JOIN will be added after sending MP_ADDADDR.

For long flows, the standard mechanism works well, even the quality of services provided by the network A1 and network A2 are different. However, if the network A1 has longer Round Trip Time (RTT) than the one of network A2. The MPTCP performance is degraded, especially in the case of short flows. Besides, the similar scenario will become popular since the different network technologies are emerging especially for the next generation of mobile networks. Therefore, it is necessary and important to solve the problem.

3. Proposal

Our proposal for solving the previous problem relies on the idea of packet duplication, specifically SYN duplication. The first SYN in initialization is duplicated. The newly created SYN packets are then sent through the multiple gateways. The proposal only requires a modifications in sending/receiving procedures of MPTCP.

We describe an beneficial use case of the proposal, which is similar to the scenario mentioned in Section 3. Note that, the network A2 has shorter RTT than the one of network A1. Initially, the key is generated on host A for the first SYN. The first SYN is constructed just like as in the standard. The SYN is also included MP_CAPABLE option. Additionally, the second SYN is newly constructed with the same content. The only difference is on the layer 3 source addresses. More specifically, the source-destination pair is (A1, B1) on the first, and (A2, B1) on the second SYN, respectively. Concurrently, the two SYNs are departed from host A to host B. This task is feasible when the routing information is available for the departures.

At the host B, the early arriving SYN (i.e., the one from A2 to B1) is received. The host B then sends an acknowledgment (SYN/ACK with MP_CAPABLE) to A2. We can observe that without modification of the

default gateway information, MPTCP has a good path selection via Network A2 for initialization. Further consideration is that, the later acknowledgment (SYN/ACK to B1) is used for an additional subflow (i.e., similar to MP_JOIN). Following the further operation, the whole period of MPTCP initialization is shortened comparing to the one in current implementation. Another obvious benefit of SYN duplications is enhancing the resilience of SYN transmission.

4. Acknowledgements

This research was conducted under a contract of Research and Development for radio resource enhancement, organized by the Ministry of Internal Affairs and Communications, Japan.

5. Security Considerations

6. References

6.1. 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>.

6.2. Informative References

[analysis] Chen, Y. and D. Towsley, "On bufferbloat and delay analysis of multipath TCP in wireless networks", IEEE/IFIP Networking, Trondheim, Norway p1-9, 2014.

[measurement] Chen, Y., Lim, Y., Gibbens, R., Nahum, E., and D. Towsley, "A measurement-based study of MultiPath TCP performance over wireless network", IEEE Internet measurement conference p110-117, 2013.

Authors' Addresses

Kien Nguyen
National Institute of Information and Communications Technology
YRP Center No.1 Building 7F, 3-4 Hikarinooka, Yokosuka
Kanagawa 239-0847
Japan

Email: kienng@nict.go.jp

Kentaro Ishizu
National Institute of Information and Communications Technology
YRP Center No.1 Building 7F, 3-4 Hikarinooka, Yokosuka
Kanagawa 239-0847
Japan

Email: ishidu@nict.go.jp

Mirza Golam Kibria
National Institute of Information and Communications Technology
YRP Center No.1 Building 7F, 3-4 Hikarinooka, Yokosuka
Kanagawa 239-0847
Japan

Email: mirza.kibria@nict.go.jp

Fumihide Kojima
National Institute of Information and Communications Technology
YRP Center No.1 Building 7F, 3-4 Hikarinooka, Yokosuka
Kanagawa 239-0847
Japan

Email: f-kojima@nict.go.jp