Multipath TCP Working Group
Internet-Draft
Intended status: Informational
Expires: June 3, 2017

K. Xu
Y. Zhang
Tsinghua University
M. Shen
Beijing Institute of Technology
Z. Ge
X. Wang
Tsinghua University

November 30, 2016

Efficient Transmission in Virtual Multi-path TCP
draft-zhang-virtual-multi-path-tcp-00

Abstract

Traditional MPTCP [RFC6824] requires that at least one node in a pair should equipped more than one NIC, and this limits the deployment [RFC6182] of MPTCP. This document proposes the VMPTCP (Virtual Multi-Path TCP) and describes how to build TCP sub-connections in VMPTCP among nodes with just one NIC.

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 June 3, 2017.

Copyright Notice

Table of Contents

1. Introduction

There are multiple pathes in Multi-path TCP [RFC6824]. To build more than one TCP sub-connections, MPTCP requires at least one node in the communication peer should be equipped with more than one NIC.

[RFC6356] describes the use of MPTCP for congestion control, [RFC6181] and [RFC6824] describe the TCP extensions for multipath operation with multiple addresses. The TCP connections, SYN and ACK messages, and the session key are carried in TCP.

However, this imposes restrictions on the deployment of MPTCP:

- A node eqiupped with just one NIC cannot benifit from MPTCP, and it is expensive to embed extra NICs to network elments
- MPTCP works as a one-to-one manner
- The I/O resources on the NIC are reserved for a particular transmission. This results in less efficient bandwidth usage

This document describes how to build virtualized TCP sub-connections among multiple single-NIC network elements. The main advantages of using VMPTCP in traditional network environment are as follows:

- No inherent restrictions on the nodes, and every node (even with one single NIC) can adopt VMPTCP to accelerate its processing
- No modification on the state-of-the-art commercial devices or network architecture
- It allows more than two participants in one VMPTCP connection, so more efficient I/O usage is possible
- Faster data transmission, more flexible connection establishment and better user experience through higher throughput
- High resilience to network failure

Taking these interface limitations into consideration like [RFC6879], [RFC6181] and [RFC7430] post the threat analysis and possible fixesfor multipath TCP.

Note that the virtual multipath TCP has no extra requirements on specific communication peers, network elements with just one NIC can also benifit from multipath. The data transmission and connection establishment are not limited to two peers any more, and multiple nodes can participate in one VMPTCP. This makes high resource utilization on each node and improves application performance.

The architecture which is described in this document can be implemented without any modifications on the state-of-the-art commercial devices or network architecture. For a single-path TCP connection between two single-NIC communication peers, if the transmission on this link cannot satisfy application requirements, VMPTCP allows other peers to join in. VMPTCP builds TCP sub-connections between the destination nodes and the newcome nodes, and these sub-connections would serve for applications together with the origin connection.

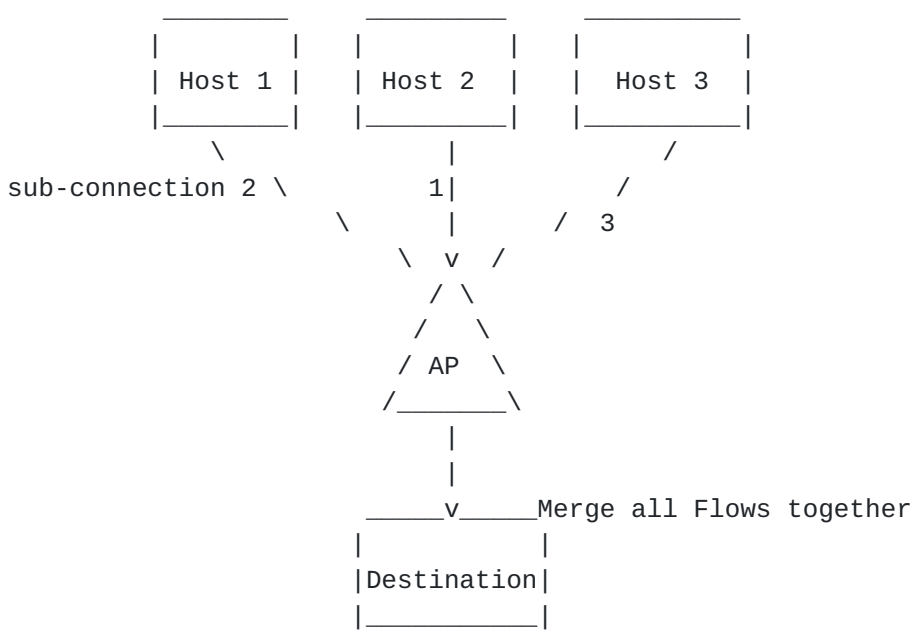
2. Terminology

Terminology defined in [RFC6181] and [RFC6824] is used extensively throughout this document.

Virtual TCP sub-connection: As there are more than one source for a particular data transmission, multiple source-destination address pairs coexist, and there is one connection between each source-destination pair. Each connection is a sub-connection and is responsible to transfer part of the required data.

3. An application of VMPTCP in wireless LAN

As shown in Figure 1, the wireless LAN consists of five nodes : Host 1, Host 2, Host 3, Destination and AP (Access Point). Note that these nodes are not necessarily required to be equipped with more than one NIC. One existing example is that when the destination node wants to download a data replica, the AP assign data sources for this download action according to VMPTCP. The advantages of VMPTCP is the multiple optional sources, and the connections establish process is also described below.



The numbers corresponding to each of the flows are described in more detail below.

Figure 1: VMPTCP with Multiple Sources

3.1. Sub-connection 1: The first connection

This is the main connection that is established between two hosts by IP address and port.

The establishment of this link should go through TCP three-way handshake and the 64-bit session key exchange. In the TCP segment, MPTCP uses MP_CAPABLE MP_JOIN and some other subtype fields under the "MPTCP Option Subtypes" [RFC6824]. Then Host 2 and the destination node generate a shared 32-bit hash key, thus, the two nodes have received an acknowledgement of this connection.

Note that Host 2 just has one NIC, so it cannot establish extra

pathes using MPTCP.

3.2. Sub-connection 2 and 3: The joint connection

In the traditional potocols, this download action can only be executed between Host 2 and the destination node. Although there are extra I/O resources in the peers, MPTCP cannot work here due to the NIC limitation.

To fully use the I/O and bandwidth resources, VMPTCP allows other hosts to join in the transmission. Note that Host 1 (or 3) has a different IP and mac from Host 2.

The sub-connection between Host 1 (or 3) still use TCP SYN segment in MPTCP potocol to generate traffic. It obtains that shared 32-bit hash key from AP and write it to the MP_JOIN segment, making it associated with the origin connection (sub-connection 1).

3.3. Merge flows from multiple sub-connections together

Once connection establishments are successfully completed, data transmission traffic would be auto adjusted among these sub-connections. Each sub-connection between host peers acts as one flow in MPTCP, and each source node (Host 1, 2, 3) acts as one NIC in the source node in the traditional MPTCP concept.

There is a data sequence number and an acknowledment number in MPTCP [RFC6824] option field, so the destinition node can keep the received data in the original order and revert to the origin data.

4. Security Considerations

Security considerations in [RFC6824] are also relevant here.

As the connection establihment process in VMPTCP is the same as that in traditional MPTCP, each host should provide the secret key to join the multiplath transmission.

In some cases, some malicious nodes would try to join the multipath transmission, so the AP in VMPTCP should deploy encryption algorithm to make node identification before sending the encrypted communication key to the node.

5. Dynamics Considerations

Dynymic considerations in [RFC6824] are also relevant here. The traffic on these multiple sub-connections can be adjusted adaptively according to the MPTCP.

5.1. Node failure

In some wireless situations, nodes could move out of the communication range, breaking the connection. In this case, the AP would recalculate the residual completion time of that transmission, if the completion time exceeds the transmission deadline and there is still unused resources on the destination node, the AP will assign other source nodes to this download action.

In sparticular, the newly assigned source-destination peer would join in the VMPTCP by establishing a new sub-connection. This process follows the above "joint connection" procedure.

The dynamic model that is described in this document brings an additional calculation requirement to AP: It should maintain state information for all established connections. This is an extra overheard than traditional MPTCP, but ensures the high transmission efficiency.

6. IANA Considerations

This document does not include an IANA request.

7. References

7.1. Normative References

[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>.

[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>.

[RFC6879] Jiang, S., Liu, B., and B. Carpenter, "IPv6 Enterprise Network Renumbering Scenarios, Considerations, and Methods", RFC 6879, DOI 10.17487/RFC6879, February 2013, <http://www.rfc-editor.org/info/rfc6879>.

[RFC7430] Bagnulo, M., Paasch, C., Gont, F., Bonaventure, O., and C. Raiciu, "Analysis of Residual Threats and Possible Fixes for Multipath TCP (MPTCP)", RFC 7430, DOI 10.17487/RFC7430, July 2015, <http://www.rfc-editor.org/info/rfc7430>.

7.2. Informative References

[RFC6181] Bagnulo, M., "Threat Analysis for TCP Extensions for Multipath Operation with Multiple Addresses", RFC 6181, DOI 10.17487/RFC6181, March 2011, <http://www.rfc-editor.org/info/rfc6181>.

[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>.

Authors' Addresses

Ke Xu
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R.China

Phone: +86-13001081658
Email: xuke@mail.tsinghua.edu.cn

Yuchao Zhang
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R.China

Phone: +86-18630500826
Email: zhangyc14@mails.tsinghua.edu.cn

Meng Shen
Beijing Institute of Technology
Department of Computer Science
Beijing 100084
P.R.China

Phone: +86-15210362001
Email: shenmeng@bit.edu.cn

Zhicheng Ge
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R.China

Phone: +86-18810543121
Email: gzc15@mails.tsinghua.edu.cn

Xiangxiang Wanf
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R.China

Phone: +86-18522247311
Email: wxx15@mails.tsinghua.edu.cn