

BANANA
Internet Draft
Intended Category: Standard Track

N. Leymann
C. Heidemann
Deutsche Telekom AG
G. Liang
China Mobile
J. Zuo
L. Zheng
Huawei
June 8, 2017

Expires: December 9, 2017

Integration of Bonding Tunnels and MPTCP
draft-zuo-banana-mptcp-integration-00.txt

Abstract

BANDwidth Aggregation for interNet Access (BANANA) based on tunnel bonding protocols and Multipath TCP (MPTCP) are two different technology suites that offer subscribers with higher bandwidth and reliability. MPTCP excels at conducting TCP traffic since its flow/congestion control scheme has been well developed while tunnel bonding protocols support not only TCP traffic but also UDP traffic. The integration of the two kinds of technologies can be considered to make use of advantages of both. Moreover, extensions made to the tunnel bonding protocols can simplify the MPTCP stack and keep it intact. This document specifies the extensions for tunnel bonding protocols to realize integrative deployment of tunnel bonding solutions and MPTCP.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

INTERNET-DRAFT

Bonding Tunnels & MPTCP Integration

June 8, 2017

Copyright and License Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Acronyms and Terminology	3
3. Integration of Tunnel Bonding and MPTCP	4
3.1. Collocated Scenarios	4
3.1.1 BANANA Endpoints Collocate with MPTCP Endpoints	4
3.1.2 HG/HAAP Collocate with MPTCP Proxy/Endpoint	4
3.1.3 HG and HAAP Collocate with Two MPTCP Proxies	5
3.2. MPTCP Traffic Recognition	6
3.3. Bypassing and Pre-coloring	6
3.4. The Anycast Mechanism	7
4. Message Types	7
4.1. Tunnel Characteristics	7
4.2. Connection Mapping	9
5. MPTCP Path Selection using Bonding Tunnels	9
5.1 Subflow Policy	10
5.2. Address Knowledge Exchange (Path Management)	10
6. Security Considerations	10
7. IANA Considerations	10
8. References	10
8.1. Normative References	10
8.2. Informative References	11
Author's Addresses	12

[1. Introduction](#)

BANdwidth Aggregation for interNet Access (BANANA) offers subscribers with higher bandwidth and resilience than they can get from a single access. MPTCP proxy based solutions and tunnel based solutions have been developed to realize BANANA.

Multipath TCP (MPTCP) enables a transport connection to operate across multiple paths simultaneously [[RFC6824](#)]. A flow/congestion control scheme has been developed by MPTCP to let subscribers with MPTCP endpoints efficiently utilize the bandwidth of the multiple paths. Recently, MPTCP WG proposed the network-based MPTCP proxy solution [[PlainProxy](#)] so that TCP endpoints can utilize the aggregated bandwidth between the two MPTCP proxies while they need not be upgraded to MPTCP endpoints.

As enabling technologies of BANANA, tunnel bonding solutions have been deployed by multiple Service Providers. In a tunnel bonding system, one tunnel is established per-connection between the two tunnel bonding boxes. These tunnels are bonded together as if there is a single IP link provided between the two boxes for the subscriber who buys the tunnel bonding box. Different from MPTCP which is a transport layer technology, tunnel bonding protocols operate under the transport layer and support both TCP and UDP.

MPTCP and a tunnel bonding protocol may coexist in one network. There are two types of such coexistence. For the first type, the tunnel endpoint and the MPTCP stack can either be collocated in a host, in a network device or on a site. For the second type, subscribers' endpoints support MPTCP. These MPTCP endpoints are distributing MPTCP traffic among the tunnels according to the MPTCP stack. The MPTCP traffic should be recognized and avoid being re-distributed by the BANANA box's load balancer. This document specifies how a tunnel bonding solution and MPTCP can be integrated in a single system in order to make use of advantages of both technologies.

In this document, message types of the bonding tunnel protocol are specified to deliver characteristics of tunnels to the MPTCP stack. MPTCP stack uses these tunnels that are setup in advance as its network paths. In this way, the time on the discovery of these network paths and the time on learning those characteristics can be saved. Hence, this kind of integration will help to accelerate the deployment of MPTCP.

2. Acronyms and Terminology

BANANA: BANDwidth Aggregation for interNet Access

HG: Home Gateway. A CPE device that is enhanced to support the simultaneous use of multiple access connections. Also used as BANANA local box.

HAAP: Hybrid Access Aggregation Point. A logical function in the network, terminating bonded connections while offering high speed Internet. Also used as BANANA remote box.

J. Zuo, et al

Expires December 9, 2017

[Page 3]

INTERNET-DRAFT

Bonding Tunnels & MPTCP Integration

June 8, 2017

CIR: Committed Information Rate [[RFC2697](#)]

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. Integration of Tunnel Bonding and MPTCP

Tunnel bonding and MPTCP can be integrated as follows. MPTCP stack and the tunnel bonding protocol endpoints can collocate in one host, one network device or one site. MPTCP stack uses the bonding tunnels as its multiple paths so that the time on path discovery can be saved. Meanwhile, the bonding tunnels do not need consider the flow/congestion control and reliability issues for the TCP traffic.

In either ways, the tunnel bonding protocol notifies the MPTCP stack about the characteristics of the tunnels, such as the priority, the available bandwidth and the Round-Trip Time (RTT). The MPTCP stack sorts the paths according to their priorities and allocates subflows according to their bandwidth demands and the available bandwidth of the tunnels or even RTT.

3.1. Collocated Scenarios

The following three scenarios explains how a tunnel bonding protocol and MPTCP are collocated in one host or one network device.

3.1.1 BANANA Endpoints Collocate with MPTCP Endpoints

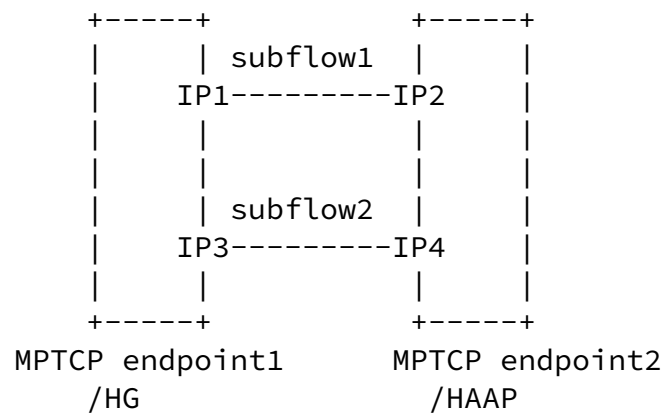


Figure 3.1: BANANA endpoints collocate with MPTCP endpoints

As shown in Figure 3.1, BANANA endpoints HG and HAAP collocate with the two MPTCP endpoints respectively. The two endpoints act as tunnel endpoints which used to be done by network based devices.

[3.1.2](#) HG/HAAP Collocate with MPTCP Proxy/Endpoint

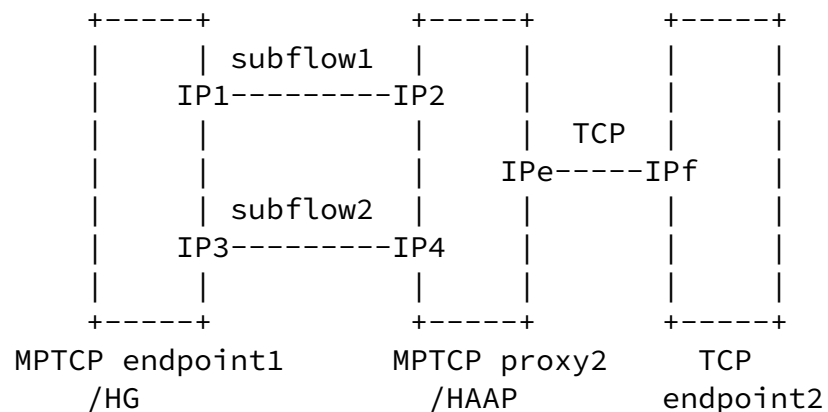


Figure 3.2: HAAP collocates with one-end MPTCP proxy
HG collocates with the MPTCP endpoint

As shown in Figure 3.2, HAAP acts as the proxy of TCP endpoint2 and speaks MPTCP with MPTCP endpoint1. MPTCP proxy2 SHOULD use IPe as the source IP address for the TCP session between itself and the TCP endpoint2. MPTCP proxy2 locally maintains the mapping between the subflows and the normal TCP session.

[3.1.3](#) HG and HAAP Collocate with Two MPTCP Proxies

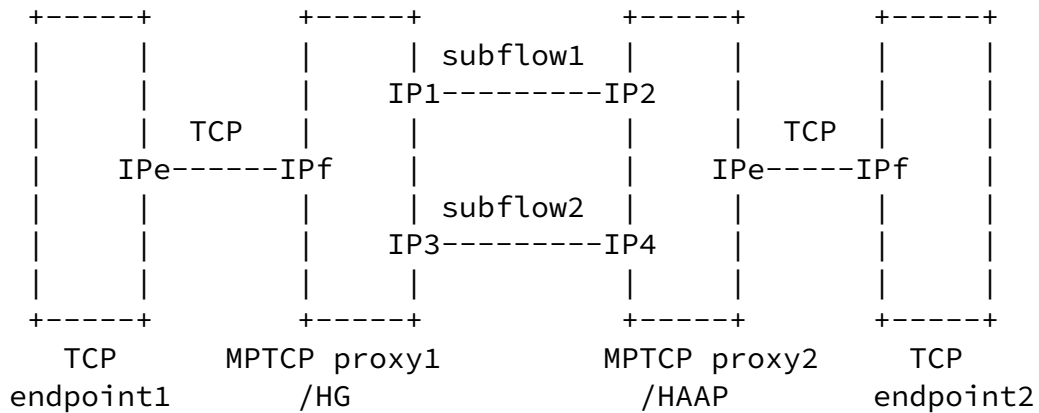


Figure 3.3: HG and HAAP collocate with two MPTCP proxies

As shown in Figure 3.3, HG and HAAP act as MPTCP proxies of TCP endpoint1 and TCP endpoint2 respectively. HG and HAAP need to learn the source and destination IP address being used by the normal TCP session and act as MPTCP proxies of the TCP endpoint1 and TCP endpoint2 from each end respectively.

The two MPTCP proxies need to exchange the mapping between the TCP connection and the MPTCP connection to ensure that a specific TCP connection is mapped to the same MPTCP connection on both proxies. This mapping is exchanged using the TLV as specified in [Section 4.2](#).

[3.2](#). MPTCP Traffic Recognition

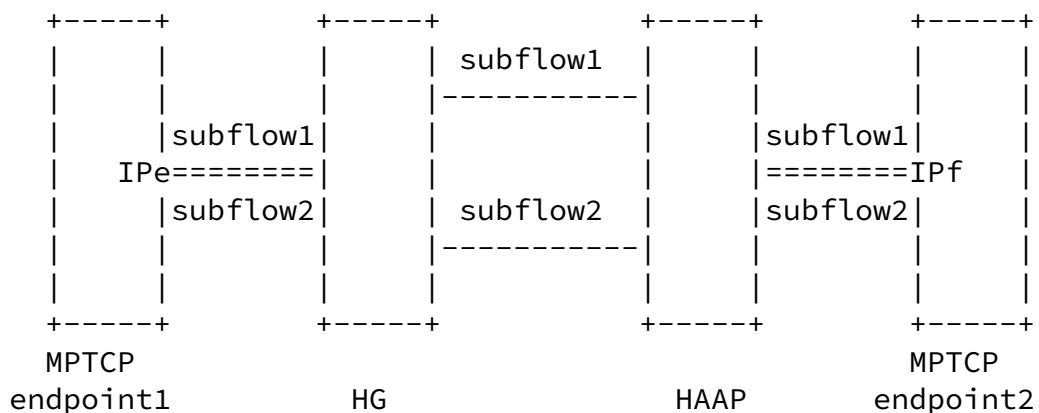


Figure 3.4: Bonding tunnels discriminate the MPTCP traffic

Suppose endpoints support the MPTCP stack. The MPTCP traffic (e.g., subflows) will be recognized by the HG/HAAP boxes and locally bypassed by the traffic distribution method of the bonding tunnel protocol. For this scenario, HG and HAAP need not to deliver the tunnel characteristics or connection mapping information of the two tunnels to the two MPTCP endpoints.

3.3. Bypassing and Pre-coloring

The bypass feature of bonding tunnels allows operators to configure traffic types not to be carried through the bonding tunnels. That is an "outer" bypass. This memo additionally enables an "inner" (inside the bonding tunnels) bypass of MPTCP traffic.

Suppose a packet of size B bytes arrives at time t.

- o If the packet is of the traffic type that is configured to bypass the bonding tunnels, it is pre-colored as green; the Single Rate Three Color Marker (srTCM) of [[RFC2697](#)] MUST operate in the Color-Aware mode with a minor change: pre-colored green packet will remain in green than be re-colored as yellow; else
- o if the packet is an MPTCP packet, this packet will be pre-colored depends on the load balancer of the MPTCP stack: it is pre-colored as green if it is to be delivered through the first tunnel or yellow if it is to be delivered through the second tunnel; the packet size B is incremented by the size of the tunnel header; the srTCM MUST operate in the Color-Aware mode with a minor change: pre-colored green packet will remain in green than be re-colored as yellow; else
- o the packet size B is incremented by the size of the tunnel header

and the srTCM MUST operate in the Color-Blind mode.

Green packets of the traffic types that are configured to bypass the bonding tunnels are carried using the first raw connection of the HG. Within the bonding tunnels, green packets are carried using the first tunnel while yellow or red packets are carried using the second tunnel.

[3.4.](#) The Anycast Mechanism

When a MPTCP proxy box and a BANANA box are collocated in a site rather than one host or one network device (Scenario 2 and Scenario 3), a box can be put in front of them to achieve anycast. Service discovery of BANANA will be handled by the front box while the tunnel endpoints will be either the BANANA box or the MPTCP proxy. Hence, two pairs of tunnels will be set up. One pair of them are terminated by the BANANA box while the other pair of tunnels are terminated by the MPTCP proxy box. Hence, MPTCP and non-MPTCP traffic is delivered in different tunnels.

The two pairs of tunnels share the same pair of paths. Different from the scenarios where BANANA box and MPTCP proxy are in the same network device, traffic for the MPTCP proxy and the BANANA box is distributed onto the paths separately. Bonding tunnels will still conduct MPTCP traffic as bypass traffic. BANANA boxes need to measure the amount of the bypass traffic periodically in order to subtract this amount from the Committed Information Rate (CIR) for the coloring mechanism [[RFC2697](#)].

[4.](#) Message Types

Two TLVs are defined by the tunnel bonding protocol to facilitate the MPTCP stack.

[4.1.](#) Tunnel Characteristics

The follow TLV carries tunnel characteristics and IP addresses of a tunnel which will be used as paths by MPTCP stack for its subflows.

Type	(1 byte)
Length	(2 bytes)
Priority	(1 byte)
Available Downstream Bandwidth	(4 bytes)
Available Upstream Bandwidth	(4 bytes)
RTT	(4 bytes)
Source IP address	(4 or 16 bytes)
Destination IP address	(4 or 16 bytes)

Type

Tunnel Characteristics (TBD), the characteristics of the tunnel.

Length

Set to 17 or 41. Source IP address and destination IP address MUST be either IPv4 or IPv6.

Priority

An unsigned integer. The bonding tunnels will determine the priority of paths for MPTCP and deliver this information through the tunnel priority. The path manager of MPTCP stack will order the paths with highest numerical priority being highest priority.

Available Downstream Bandwidth

An unsigned integer measured in kbps. The HAAP calculates the available Bandwidth by subtracting the measured bypass traffic rate, where the bypass traffic rate can be calculated by periodically counting the received packets at the HG. The HG notify the Available Downstream Bandwidth and the measured bypass traffic rate to the HAAP.

Available Upstream Bandwidth

An unsigned integer measured in kbps. There is no bypass traffic rate to be subtracted from the Committed Upstream Bandwidth. The HAAP notify the Available Upstream Bandwidth to the HG.

RTT

An unsigned integer measured in ms.

Source IP address

Set to a pre-configured IPv4/IPv6 address which is used as the source endpoint IP address of a tunnel between the two MPTCP proxies.

Destination IP address

Set to a pre-configured IPv4/IPv6 address which is used as the destination endpoint IP address of a tunnel between the two MPTCP proxies.

4.2. Connection Mapping

The following Connection Mapping TLV is specified by the tunnel bonding protocol for the purpose of exchanging between the pair of MPTCP proxies about the mapping between MPTCP connection and TCP connection.

```

+-----+
| Type                | (1 byte)
+-----+
| Length              | (2 bytes)
+-----+
| MPTCP Connection ID | (4 bytes) |
+-----+
| Source IP address of TCP | (4 or 16 bytes) |
+-----+
| Destination IP address of TCP | (4 or 16 bytes) |
+-----+

```

Type

Connection Mapping (TBD), the mapping between MPTCP connection and TCP connection.

MPTCP Connection ID

An unique identifier given to a multipath connection by the MPTCP proxy, which is the Token defined in [[RFC6824](#)].

Source IP address of TCP

The source IPv4/IPv6 address of the TCP connection.

Destination IP address of TCP

The destination IPv4/IPv6 address of the TCP connection.

5. MPTCP Path Selection using Bonding Tunnels

Based on the tunnel characteristics like bandwidth, latency, etc., of the bonding tunnels, MPTCP may select the set of paths for providing

higher throughput or resilience against path failures.

[5.1](#) Subflow Policy

As mentioned in [[RFC6824](#)], the MPTCP endpoint may use any local policy to decide how to send the traffic over the available paths. the MPTCP endpoint requires the knowledge of the path 'cost' to make effective choices, where the path 'cost' may be obtained from the TLV carried by the Tunnel Characteristics TLV. Moreover, in the event that the available set of paths changes, the MPTCP endpoint may wish to signal a change in priority of subflows to the other MPTCP endpoint, where the subflow has the same meaning of physical path. Therefore, the priority information may also be obtained from the TLV. If the multiple subflows are differentiated from port numbers while with the same IP addresses, the priority of subflows are deemed to be the same, as obtained from the TLV of the same physical path.

[5.2](#). Address Knowledge Exchange (Path Management)

As mentioned in [[RFC6824](#)], the "path management" is responsible for the exchange of additional addresses between the two MPTCP endpoints, where the Add Address (ADD_ADDR) MPTCP option is used to inform the other MPTCP endpoint of the IP addresses. If the previously informed address becomes invalid, a Remove Address (REMOVE_ADDR) option is used to announce the peer to remove the subflows related to this address. The ADD_ADDR and REMOVE_ADDR options are carried in the duplicate ACK packets and to be sent periodically. However, with the help of the IP addresses carried in Connection Mapping TLV, the MPTCP stack does not need extra signal to manage the paths.

[6](#). Security Considerations

TBD.

[7](#). IANA Considerations

Codepoints of the TLV defined in [Section 5](#) need to be registered by a specific tunnel bonding protocol.

[8](#). References

8.1. 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>>.
- [RFC2697] Heinanen, J. and Guerin R., "A Single Rate Three Color Marker", [RFC 2697](#), DOI 10.17487/RFC2697, September 1999,

J. Zuo, et al

Expires December 9, 2017

[Page 10]

INTERNET-DRAFT

Bonding Tunnels & MPTCP Integration

June 8, 2017

<<http://www.rfc-editor.org/info/rfc2697>>.

- [RFC6824] Ford, A., Raiciu, C., Handley, M., and Bonaventure, O., "TCP Extensions for Multipath Operation with Multiple Addresses", [RFC 6824](#), DOI 10.17487/RFC6824, January 2013, <<http://www.rfc-editor.org/info/rfc6824>>.
- [PlainProxy] Boucadair, M., Jacquenet, C., et al, "Extensions for Network-Assisted MPTCP Deployment Models", [draft-boucadair-mptcp-plain-mode](#), work in progress.

8.2. Informative References

- [802Type] IANA, "IEEE 802 Numbers", <<http://www.iana.org/assignments/ieee-802-numbers>>.

Author's Addresses

Nicolai Leymann
Deutsche Telekom AG
Winterfeldtstrasse 21-27
Berlin 10781
Germany

Phone: +49-170-2275345
EMail: n.leymann@telekom.de

Cornelius Heidemann
Deutsche Telekom AG
Heinrich-Hertz-Strasse 3-7
Darmstadt 64295
Germany

Phone: +4961515812721
EMail: heidemannc@telekom.de

Geng Liang
China Mobile

32 Xuanwumen West Street,
Xicheng District, Beijing, 100053,
China

EMail: gengliang@chinamobile.com

Jing Zuo
Huawei Technologies
Bantian, Longgang District,
Shenzhen 518129 P.R. China
EMail: jing.zuo@huawei.com

Lianshu Zheng
Huawei Technologies
No.156 Beiqing Rd. Haidian District,
Beijing 100095 P.R. China

EMail: vero.zheng@huawei.com