

Network Working Group
Internet-Draft
Updates: [6824](#) (if approved)
Intended status: Experimental
Expires: September 10, 2015

M. Boucadair
C. Jacquenet
France Telecom
March 9, 2015

An Extension to MPTCP for Symmetrical Sub-Flow Management
draft-boucadair-mptcp-symmetric-02

Abstract

This document specifies a MPTCP extension that allows to achieve symmetrical subflow management. In particular, this extension allows both endpoints to add new subflows whenever needed without waiting for the endpoint which initiated the first subflow to add new ones.

This document updates [RFC 6824](#).

Requirements Language

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](#)].

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 September 10, 2015.

Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

Internet-Draft

Symmetrical MPTCP

March 2015

This document is subject to [BCP 78](#) 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.	Problem Space	2
3.	Proposed Solution	3
4.	Security Considerations	5
5.	IANA Considerations	5
6.	Acknowledgements	5
7.	References	6
7.1.	Normative References	6
7.2.	Informative References	6
	Authors' Addresses	7

[1.](#) Introduction

This document specifies a MPTCP [[RFC6824](#)] extension to achieve symmetrical subflow management. The problem space is further described in [Section 2](#), while a proposed solution is discussed in [Section 3](#).

This document assumes Port Control Protocol (PCP)-enabled networks [[RFC6887](#)]. But other procedures can be used to instantiate mappings and discover the external IP address/port assigned by an upstream flow-aware device (e.g., CGN [[RFC6888](#)], firewall, etc.).

[2.](#) Problem Space

The following is extracted from[I-D.ietf-mptcp-experience]:

From a subflow viewpoint, the Multipath TCP protocol is completely symmetrical. Both the clients and the server have the capability to create subflows. However in practice the existing Multipath TCP implementations [[I-D.eardley-mptcp-implementations-survey](#)]

have opted for a strategy where only the client creates new subflows. The main motivation for this strategy is that often the client resides behind a NAT or a firewall, preventing passive subflow openings on the client.

This means that in practice only the client (that is the TCP endpoint that initiated the first subflow) can initiate new subflows. This is not optimal in situations where (1) the remote endpoints want to boost their sending rate or handover to a new IP address without waiting for the client to add new subflows, (2) or when the traffic distribution as observed by the remote endpoint does not meet its local policies. Adding new subflows should be subject to both the client's and server's local policies, not only those of the client.

[3.](#) Proposed Solution

This procedure can be activated upon bootstrap or when a network attachment change occurs (e.g., attach to a new network); it is not executed for every new MPTCP connection:

- o Each endpoint proceeds to the discovery of upstream flow-aware devices (e.g., NAT, Firewall).

This can be achieved by various means, e.g., using UPnP IGD [[IGD1](#)][[IGD2](#)], PCP server discovery [[RFC6887](#)], PCP DHCP option [[RFC7291](#)], DS-Lite AFTR [[RFC6334](#)], etc.

A NAT/firewall can be embedded in a CPE (Customer Premises Equipment) and/or hosted in the network provider's side.

- o Appropriate mappings are instantiated in those discovered flow-aware devices. In particular, external IP address(es) and port numbers are retrieved.

This can be achieved using PCP [[RFC6887](#)]. Note, mappings created by PCP MAP requests are, by definition, endpoint-independent mappings (EIMs) with endpoint-independent filtering (EIF). Filters can be associated with the PCP MAP request to restrict a mapping to be bound to specific remote peer(s).

PCP allows to dynamically control both NATs and firewalls.

Furthermore, PCP allows to retrieve the lifetime associated with an external IP address and external port number.

If the host is a UPnP IGD Control Point, [[RFC6970](#)] allows to relay UPnP IGD primitives into PCP messages. PCP can also be used to control multi-layered NATs/firewall owing to the activation of [[I-D.ietf-pcp-proxy](#)] in intermediate NATs/firewalls.

- o When initiating an MPTCP connection, external IP addresses and port numbers are communicated to the remote peer.

This can be achieved using ADD_ADDR together with a new option that will indicate that the address/port pair (identified by Address ID) enclosed in ADD_ADDR has been checked so that incoming flows can be sent to this address/port.

A second implementation flavor is to define a new option, similar to ADD_ADDR, but which will include an optional field (lifetime). The lifetime can be used as an input to the traffic management block at the remote side. This field can be derived from the lease returned in DHCP, or in PCP requests. The use of this option is an indication that appropriate actions were undertaken at the remote side to receive incoming packets. A remote peer can use the enclosed address/port to add a new subflow without any risk to experience failures at the client side. Indicating the lifetime associated with an IP address is seen as a limitation of current APIs as discussed in Section-3.2.1 of [[RFC6250](#)]. The lifetime can be used as a hint to migrate flows to another subflows.

Another implementation flavor is to update ADD_ADDR as shown in Figure 1. "IPVer" is useless since the length is sufficient to determine whether the enclosed IP address is IPv4 or IPv6.

Internet-Draft

Symmetrical MPTCP

March 2015

OLD:

1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9 0 1
-----	-----	-----
Kind	Length	ADD_ADDR IPVer Address ID
-----	-----	-----
Address (IPv4 - 4 octets / IPv6 - 16 octets)		

Port (2 octets)		

NEW:

1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9 0 1
-----	-----	-----
Kind	Length	ADD_ADDR Flags Address ID
-----	-----	-----
Address (IPv4 - 4 octets / IPv6 - 16 octets)		

Port (2 octets)		

Flags is a set of 4 flags:

+--+--+--+
C r r r
+--+--+--+

C flag MUST be set to 1 when the address/port are checked.
 "rrr" are for future assignment as additional flag bits.
 r bits MUST each be sent as zero and MUST be ignored on receipt.

Figure 1

[4. Security Considerations](#)

PCP-related security considerations are discussed in [[RFC6887](#)].
 MPTCP-related security considerations are documented in [[RFC6824](#)] and [[I-D.ietf-mptcp-attacks](#)].

[5. IANA Considerations](#)

TBC.

[6. Acknowledgements](#)

Many thank to Olivier Bonaventure who suggested the idea of updating ADD_ADDR.

[7. References](#)

[7.1. 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.

[7.2. Informative References](#)

- [I-D.eardley-mptcp-implementations-survey]
 Eardley, P., "Survey of MPTCP Implementations", [draft-](#)

[eardley-mptcp-implementations-survey-02](#) (work in progress), July 2013.

[I-D.ietf-mptcp-attacks]

Bagnulo, M., Paasch, C., Gont, F., Bonaventure, O., and C. Raiciu, "Analysis of MPTCP residual threats and possible fixes", [draft-ietf-mptcp-attacks-03](#) (work in progress), February 2015.

[I-D.ietf-mptcp-experience]

Bonaventure, O., Paasch, C., and G. Detal, "Experience with Multipath TCP", [draft-ietf-mptcp-experience-01](#) (work in progress), March 2015.

[I-D.ietf-pcp-proxy]

Perreault, S., Boucadair, M., Penno, R., Wing, D., and S. Cheshire, "Port Control Protocol (PCP) Proxy Function", [draft-ietf-pcp-proxy-06](#) (work in progress), December 2014.

[IGD1]

UPnP Forum, , "WANIPConnection:1 Service (<http://www.upnp.org/specs/gw/UPnP-gw-WANIPConnection-v1-Service.pdf>)", November 2001.

[IGD2]

UPnP Forum, , "WANIPConnection:2 Service (<http://upnp.org/specs/gw/UPnP-gw-WANIPConnection-v2-Service.pdf>)", September 2010.

[RFC6250]

Thaler, D., "Evolution of the IP Model", [RFC 6250](#), May 2011.

[RFC6334]

Hankins, D. and T. Mrugalski, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Option for Dual-Stack Lite", [RFC 6334](#), August 2011.

[RFC6887]

Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", [RFC 6887](#), April 2013.

[RFC6888]

Perreault, S., Yamagata, I., Miyakawa, S., Nakagawa, A., and H. Ashida, "Common Requirements for Carrier-Grade NATs (CGNs)", [BCP 127](#), [RFC 6888](#), April 2013.

- [RFC6970] Boucadair, M., Penno, R., and D. Wing, "Universal Plug and Play (UPnP) Internet Gateway Device - Port Control Protocol Interworking Function (IGD-PCP IWF)", [RFC 6970](#), July 2013.
- [RFC7291] Boucadair, M., Penno, R., and D. Wing, "DHCP Options for the Port Control Protocol (PCP)", [RFC 7291](#), July 2014.

Authors' Addresses

Mohamed Boucadair
France Telecom
Rennes 35000
France

Email: mohamed.boucadair@orange.com

Christian Jacquenet
France Telecom
Rennes 35000
France

Email: christian.jacquenet@orange.com