A Multi-Path Concurrent Measurement Protocol for IPPM

draft-dang-ippm-multiple-path-measurement-01

Abstract

This test method can test multi-paths concurrently between two edge nodes. This document details Multi-Path Concurrent Measurement Protocol (MPCMP).

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 https://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 11, 2019.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://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
   1.1. Requirements Language ............................ 2
   1.2. Terminology & Abbreviations ................... 2
2. Overview of Multi-Path Concurrent Measurement Protocol .... 3
3. MPCMP-Test Packet Format and Content .......................... 3
4. Expansion based on various measurement methods ............. 6
   4.1. IOAM ........................................... 6
5. Data Export ............................................ 6
6. IANA Considerations ..................................... 6
7. Security Considerations ................................... 6
8. Acknowledgements ....................................... 6
9. Normative References .................................... 6
Authors’ Addresses ...................................... 7

1. Introduction

In load-balancing scenario, there are multiple paths adopted between two edge nodes. The traffic from the Scr node to the Dst node is required to be steered into the specific path/paths basing on the SLA information of each path. In the traditional method, the paths are measured separately. If you want to ensure that the data obtained by the test is available and accurate, then the test start and end points of this set of Paths must be consistent.

The Multi-Path Concurrent Measurement Protocol (MPCMP) is required, which can be used bi-directionally to concurrently measure multi-paths metrics between two network elements. At the same time, this method also saves the number of test messages and reduces the load on the network.

1.1. 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.

1.2. Terminology & Abbreviations

- Multiple Paths
  * There are multiple paths between two nodes in the network. These paths may be equal-cost multi-path (ECMP) mode or unequal-cost multiple (UCMP) mode. In a real network, they might be one [draft-ietf-spring-segment-routing-policy] or [RFC7348] tunnel.
o Concurrent
  * In order to ensure comparability between multiple paths, the test start point and the test end point are required to be synchronized.

2. Overview of Multi-Path Concurrent Measurement Protocol

The Multi-Path Concurrent Measurement Protocol (MPCMP) is an open protocol for measurement of multi-paths metrics.

MPCMP can be embedded into a variety of transports such as NSH, Segment Routing, VxLAN, native IPv6 (via extension header), or IPv4.

3. MPCMP-Test Packet Format and Content

This section defines path header and associated data types required for MPCMP.

```
+-----+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Session ID                        |
+-----+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Path ID            |         Path-E2E-Type         |
+-----+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Flags   |          Transaction ID                         |
+-----+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: MPCMP Path header

- **Session ID**: A set of load sharing paths
- **Path ID**: One path of the session.
- **Path-E2E-Type**: A 16-bit identifier which Indicates whether the packet type is a send message or a request message.
- **Flags**: 8-bit field. Identify the query or response type. Following flags are defined:
  * Bit 0 Identify the query type
  * Bit 1 Identify the response type
  * Reserved
o Transaction ID: 16-bit identifier of one measurement transaction. The sender and receiver to identify measurement transactions based on Transaction ID.

When a measurement is for a set of paths, each query message is made for each path, but only one unified response message replies. Therefore, the message format is defined as follows.

The measurement packet format of a path is as follows.

```
+----------------------------------+
|                                  |
|                E2E PathN Option Header                        |
|                                  |
+----------------------------------+
|                                  |
|            PathN Edge-to-Edge Option Data                    |
|                                  |
+----------------------------------+
|                                  |
```

Figure 2: Query message

The field of PathN Edge-to-Edge Option Data can refer to Edge-to-Edge Option Data of [draft-ietf-ippm-ioam-data-04].

The response type message format is as follows. It suppose there are N paths between two points.
Figure 3: Response message

- Long-term measurement
  * The receiver can wait until it receives all measurement requests of a set of path and then responds.

- Short-term measurement
  * The Sender can query once t.
  * The receiver can reply once t.

The overall solution needs to consider two methods of long-period measurement and short-period measurement.
4. Expansion based on various measurement methods

The measurement message format defined by this document can be extended based on various measurement methods.

4.1. IOAM

A new type is added in IOAM-E2E-Type of IOAM Edge-to-Edge Option header [draft-ietf-ippm-ioam-data-04-section4.4] as follow.

- Bit 4: Multiple paths measurement.

This bit is set by the headend node if Multi-Path Concurrent Measurement is activated.

A common registry is maintained for IOAM-Types, see Section 6.

5. Data Export

MPCMP nodes collect information for packets traversing a domain that supports MPCMP. MPCMP process the information further and export the information using e.g., IPFIX. Raw data export of IOAM data using IPFIX is discussed in [draft-spiegel-ippm-ioam-rawexport-00].

6. IANA Considerations

This document requests the following IANA Actions.

IOAM E2E Type Registry:

Bit 4 Multiple ways measurement

7. Security Considerations

The Proof of Transit option (Section 4.3 In-situ OAM [draft-ietf-ippm-ioam-data-04-section4.4]) is used for verifying the path of data packets.

8. Acknowledgements

TBD

9. Normative References

[draft-ietf-ippm-ioam-data-04]
"A Variety of Transports",
"IOAM Edge-to-Edge Option",

[draft-ietf-spring-segment-routing-policy]
"Segment Routing Policy Architecture",

[draft-spiegel-ippm-ioam-rawexport-00]
"In-situ OAM raw data export with IPFIX",


[RFC7348] "Virtual eXtensible Local Area Network (VXLAN)",

Authors’ Addresses

Joanna Dang (editor)
Huawei
Beijing
China

Email: dangjuanna@huawei.com

Jianglong
China Telecom
Beijing
China

Email: wangjl1.bri@chinatelecom.cn