

Network Working Group
Internet-Draft
Intended status: Informational
Expires: February 14, 2021

D. Dang, Ed.
Huawei
W. Wang
China Telecom
L. LEE
LG U+
Y. Yan
Tencent
C. Cheng
Huawei
August 13, 2020

**A One-Path Congestion Metric for IPPM
draft-dang-ippm-congestion-03**

Abstract

This memo defines a metric for one path congestion across Internet paths. The traditional mode evaluates network congestion based on the bandwidth utilization of the link. However, there is a lack of E2E path congestion that is truly service oriented. So A Path Congestion Metric is required.

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 February 14, 2021.

Copyright Notice

Copyright (c) 2020 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](#) [3](#)
- [1.2. Terminology & Abbreviations](#) [3](#)
- [1.3. Motivation](#) [3](#)
- [2. A singleton definition of a Type-P-Path-Congestion Metric . .](#) [4](#)
- [2.1. Metric Name](#) [4](#)
- [2.2. Metric Parameters](#) [4](#)
- [2.3. Metric unit](#) [5](#)
- [2.4. Definition](#) [5](#)
- [2.5. Methodologies](#) [5](#)
- [2.6. Reporting the Metric](#) [7](#)
- [3. A Definition for Samples of Path Congestion](#) [7](#)
- [3.1. Metric Name](#) [7](#)
- [3.2. Metric Parameters](#) [7](#)
- [3.3. Metric Units](#) [8](#)
- [3.4. Definition](#) [8](#)
- [3.5. Methodologies](#) [8](#)
- [3.6. Reporting the Metric](#) [9](#)
- [4. Security Considerations](#) [9](#)
- [5. IANA Considerations](#) [9](#)
- [6. Acknowledgements](#) [9](#)
- [7. Normative References](#) [9](#)
- Authors' Addresses [10](#)

1. Introduction

As we know, the current network has been already being in load balancing mode, however it is partially congested. To solve the problem of unbalanced network load[[draft-liu-ican](#)], Congestion of paths requires an assessment metric. The traditional mode evaluates network congestion based on the bandwidth utilization of the link. However, there is a lack of E2E path congestion that is truly service oriented.

So this memo defines a metric for a path congestion across Internet paths. Currently there is no path congestion metric specified according to [[RFC2330](#)] framework, so the notions and conventions of Path Congestion will be introduced to extend [RFC 2330](#).

- o A 'singleton*' analytic metric, called Type-P-Path-Congestion, will be introduced to measure a single observation of A Path Congestion.
- o Using this singleton metric, a 'sample*' called Type-P-One-Path-Congestion-Poisson-Stream is introduced to measure a sequence of singleton congestions sent at times taken from a Poisson process, defined in [Section 11.1.1 of \[RFC2330\]](#).

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

- o Path Group
 - * There are multiple channels between two nodes in the network. These channels 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 Path
 - * One path of the path group
- o Path Congestion
 - * A path is congested, indicating that the traffic of the path exceeds its own throughput. The result of path congestion is that the buffer of the nodes along the path cache traffic or the buffer along the way is insufficient to cause packet loss.

1.3. Motivation

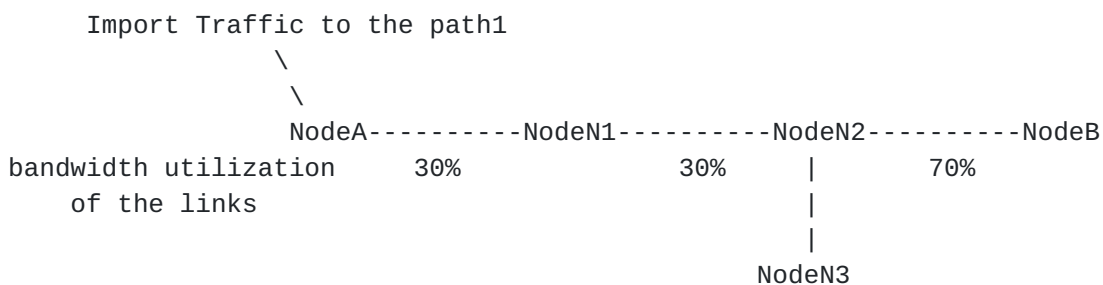


Figure 1: A Path

As Figure 1 is shown, there is a path between nodes A through B. Some traffic is imported into this path.

- o At node A, the traffic is imported to path1.
- o The bandwidth utilization of the links passed by path1 (Such as the link from Node A to NodeN1, the link from NodeN1 to NodeN2, the link from NodeN2 to NodeB) need to be checked. The bottleneck links identified where bandwidth utilization (70%) of the link between node N2 and NodeB exceeds the threshold (65%). The node where the bottleneck link is located is facing expansion and adapt some flows to another path, in order to ensure the service experiences.

The congestion of this bottleneck link may be suspected to be caused by traffic coming from NodeN1, or coming from NodeN3. There is a lack of a measure of a path congestion. So A Path Congestion Metric is required to directly report the degree of path congestion, and is accurate.

2. A singleton definition of a Type-P-Path-Congestion Metric

2.1. Metric Name

Type-P-Path-Congestion use the Type-P convention as described in[RFC2330] .

2.2. Metric Parameters

- o Src, the IP address of a host
- o Dst, the IP address of a host
- o dpt, transmission period of a measurement. For example, a measurement is triggered once dpt (such as 3.3 milliseconds)
- o dt, interval from when Src initiates the measurement message to when Dst receives it.
- o T, a time
- o SrcCountp, the number of packets at source node in one dpt period. When Src sent the first bit of a Type-P packet to Dst, This parameter starts counting. And The counting is terminated after the end of the cycle. This data can be collected locally at the entry of the traffic import the path from Src to Dst.

- o DstCount_p, the number of packets at destination node in one dpt period. When Dst receive the first bit of a Type-P packet to Dst and the counting is terminated after the end of the cycle. Because the statistics of the counting is based on the path ,the related data can be collected locally based on the PathID what may be SR label list[RFC8402] or VXLAN ID[RFC7348].

2.3. Metric unit

The value of a Type-P-Path-Congestion is either a real number or an undefined (informally, infinite) number.

2.4. Definition

For a real number c,

- o the *Type-P-Path-Congestion* from Src to Dst at T is 0, which means that Src sent the first bit of a Type-P packet to Dst at wire time* T and that there is no path congestion to Dst.
- o the *Type-P-One-way-Delay* from Src to Dst at T is greater than 0, which means that Src sent the first bit of a Type-P packet to Dst at wire time* T and there is the path congestion to Dst.

In theory, this metric is larger while the path congestion is more serious.

2.5. Methodologies

As with other Type-P-* metrics, the detailed methodology will depend on the Type-P (e.g., protocol number, UDP/TCP port number, size, Differentiated Services (DS) Field[RFC2780]).

The measurement methodology described in this section assumes the measurement and determination of Type-P-Path-Congestion in real-time.

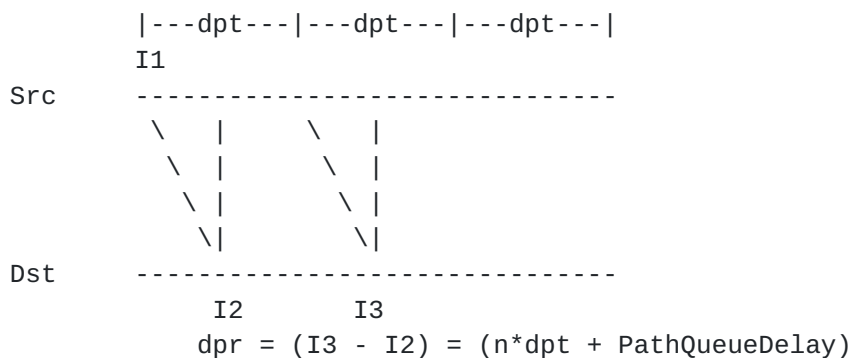


Figure 2: Measurement

As Figure 2 is shown, for a given Type-P, the methodology would proceed as follows:

- o Start after time I1. At the Src host, select Src and Dst IP addresses, and form test packets of Type-P with these addresses according to a given technique (e.g., the Poisson sampling technique). The test packet can be used to dye traffic packets or generate a packet[RFC7799][[RFC8321](#)].
- o At the Dst host, arrange to receive the packet.
- o At the Src host, place a timestamp in the prepared Type-P packet, and send it towards Dst.
- o If the packet arrives within a reasonable period of time, take a time stamp I2 as soon as possible upon the receipt of the packet. By subtracting the two time stamps, an estimate of 5.3. Type-P-One-way-Delay-Minimum[RFC7679] can be computed.
- o If the packet meets the selection function criterion for the first packet, record this first Type-P-One-way-Delay-Minimum value.
 - * Long-term measurement
 - + At the Src host, the packets continue to be generated according to the given methodology. The Src host places a time stamp in the Type-P packet, and send it towards Dst. If the packet arrives within a reasonable period of time, the Dst host take a time stamp I3 as soon as possible upon the receipt of the packet.
 - * Short-term measurement
 - + After sending the first test message, the Src host sends a measurement message once dpt. Generating the Type-P packet stream is as above.
 - + At the Dst host, when receiving the first measurement packet, it also sends a response packet to the Dst Host once dpt. The purpose of this is to improve measurement accuracy. Because when the packet is sent back the measured packet may not reach the Dst Host.
- o So
 - * In long term measurement, the congestion parameters are calculated as follows:

$$+ c = ((I3 - I2) / dpt) -1$$

- * In long-term measurement, the congestion parameters are calculated as follows:

$$+ c = (SrcCountp / DstCountp)-1$$

Therefore, the overall solution needs to consider two methods of long-period measurement and short-period measurement. The two data from the two method can also be verified against each other if necessary.

If a packet loss is detected on the path within a certain period, the congestion assessment will be terminated. The next measurement of the path can only be restarted after initialization.

2.6. Reporting the Metric

The calibration and context in which the metric is measured MUST be carefully considered and SHOULD always be reported along with metric results. We now present four items to consider: the Type-P of test packets, the threshold of infinite delay (if any), error calibration, and the path traversed by the test packets. This list is not exhaustive; any additional information that could be useful in interpreting applications of the metrics should also be reported (see [[RFC6703](#)] for extensive discussion of reporting considerations for different audiences)[[RFC6703](#)].

3. A Definition for Samples of Path Congestion

The goal of the sample definition is to make it possible to compute the statistics of sequences of Path Congestion measurements.

3.1. Metric Name

Type-P-Path-Congestion-Poisson-stream

3.2. Metric Parameters

- o Src, the IP address of a host
- o Dst, the IP address of a host
- o dpt[i], transmission period of a measurement. For example, a measurement is triggered once dpt (such as 3.3 milliseconds).
- o dt[i], interval from when Src initiates the measurement message to when Dst receives it

- o T_0 , a time
- o T_f , a time
- o λ , a rate in reciprocal seconds
- o $SrcCountp[i]$, the number of packets at source node in one dpt period. When Src sent the first bit of a Type-P packet to Dst, This parameter starts counting. And The counting is terminated after the end of the cycle. This data can be collected locally at the entry of the traffic import the path from Src to Dst.
- o $DstCountp[i]$, the number of packets at destination node in one dpt period. When Dst receive the first bit of a Type-P packet to Dst and the counting is terminated after the end of the cycle. Because the statistics of the counting is based on the path ,the related data can be collected locally based on the PathID what may be SR label list[RFC8402] or VXLAN ID[RFC7348].

3.3. Metric Units

A sequence of four parameters whose elements are:

- o T , a time, and
- o c , either a zero (signifying no congestion transmission of the packet) or greater than a zero (signifying congestion).

3.4. Definition

Given T_0 , T_f , and λ , we compute a pseudorandom Poisson process beginning at or before T_0 , with average arrival rate λ , and ending at or after T_f . Those time values greater than or equal to T_0 and less than or equal to T_f are then selected.

At each of the selected times in this process, we obtain one value of Type-P-One-Path-Congestion. The value of the sample is the sequence made up of the resulting $\langle \text{time}, c \rangle$ pair . If there are no such pair, the sequence is of length zero and the sample is said to be empty.

3.5. Methodologies

The methodologies follow directly from:

- o The selection of specific times using the specified Poisson arrival process, and

- o The methodologies discussion already given for the singleton Type-P-One-Path-Congestion metric.

$$* \text{Type-P-Path-Congestion}[i] = ((\text{SrcCountp}[i]) / (\text{DstCountp}[i])) - 1$$

If applied in a load-sharing network, Multi-Path Concurrent Measurement for IPPM is recommended[[draft-dang-ippm-multiple-path-measurement](#)].

If a packet loss is detected on the path within a certain period, the congestion assessment will be terminated. The next measurement of the path can only be restarted after initialization.

3.6. Reporting the Metric

The calibration and context for the underlying singletons MUST be reported along with the stream.

4. Security Considerations

The path congestion metric has the same security properties as [RFC7679], and thus they inherit the security considerations of that document[RFC2330].

5. IANA Considerations

TBD

6. Acknowledgements

TBD

7. Normative References

- [[draft-dang-ippm-multiple-path-measurement](#)]
"Multi-Path Concurrent Measurement for IPPM",
<<https://tools.ietf.org/html/draft-dang-ippm-multiple-path-measurement-02>>.
- [[draft-ietf-spring-segment-routing-policy](#)]
"Segment Routing Policy Architecture",
<<https://tools.ietf.org/html/draft-ietf-spring-segment-routing-policy-02>>.

[\[draft-liu-ican\]](#)

"Instant Congestion Assessment Network (iCAN) for Traffic Engineering",
<<https://tools.ietf.org/html/draft-liu-ican-02>>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997,
<<https://www.rfc-editor.org/info/rfc2119>>.

[RFC2330] "Framework for IP Performance Metrics",
<<https://tools.ietf.org/html/rfc2330>>.

[RFC2780] "[RFC2780](#)", <<https://tools.ietf.org/html/rfc2780>>.

[RFC6703] "Reporting IP Network Performance Metrics: Different Points of View",
<<https://datatracker.ietf.org/doc/rfc6703/>>.

[RFC7348] "Virtual eXtensible Local Area Network (VXLAN)",
<<https://datatracker.ietf.org/doc/rfc7348/>>.

[RFC7679] "A One-Way Delay Metric for IP Performance Metrics (IPPM)", <<https://tools.ietf.org/html/rfc7679>>.

[RFC7799] "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", <<https://tools.ietf.org/html/rfc7799>>.

[RFC8321] "Alternate-Marking Method for Passive and Hybrid Performance Monitoring",
<<https://tools.ietf.org/html/rfc8321>>.

[RFC8402] "Segment Routing Architecture",
<<https://datatracker.ietf.org/doc/rfc8402/>>.

Authors' Addresses

Joanna Dang (editor)
Huawei
Beijing
China

Email: dangjuanna@huawei.com

Jianglong
China Telecom
Beijing
China

Email: wangjl50@chinatelecom.cn

Shinyoung
LG U+
Seoul
Korea

Email: leesy@lguplus.co.kr

Hongbo
Tencent
Beijing
China

Email: nikyan@tencent.com

Liang
Huawei
Beijing
China

Email: liang.cheng@huawei.com

