

IPPM
Internet-Draft
Intended status: Standards Track
Expires: 9 August 2023

T. Mizrahi
T. Zhou
S. Belkar
R. Cohen
Huawei
5 February 2023

The Transit Measurement Option
draft-mzbc-ippm-transit-measurement-option-01

Abstract

This document specifies an IPv6 option that contains a compact set of fields which can be used for transit delay measurement and congestion detection. This option can be incorporated into data packets and updated by transit nodes along the path, enabling lightweight measurement and monitoring using constant-length data that does not depend on the number of hops in the network.

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 9 August 2023.

Copyright Notice

Copyright (c) 2023 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 Revised BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Revised BSD License.

Table of Contents

1.	Introduction	2
2.	Conventions	4
2.1.	Requirement Language	4
2.2.	Terminology	4
3.	Transit Measurement Option	4
4.	IANA Considerations	6
4.1.	IPv6 Option Type	6
5.	Security Considerations	6
6.	References	6
6.1.	Normative References	6
6.2.	Informative References	7
	Authors' Addresses	8

[1.](#) Introduction

This document introduces an IPv6 option that includes a compact set of performance-related fields. This option can be incorporated into data packets and updated by transit nodes along the path.

There is a number of in-progress documents in the IETF that define IPv6 options that can be used for tracing a path and its performance, including for example, [[I-D.ietf-ippm-ioam-ipv6-options](#)], [[I-D.filsfils-spring-path-tracing](#)], [[I-D.ali-spring-ioam-srv6](#)], [[I-D.kumar-ippm-ifa](#)], [[I-D.zhou-ippm-enhanced-alternate-marking](#)]. Some of these extensions use per-hop fields which are updated by intermediate nodes, collecting information about the performance along the path. While these extension provide detailed and fine-grained information, they incur high per-packet bandwidth and processing overhead.

The Transit Measurement option, which is defined in this document, provides coarse-grained performance information using a set of fields that have a constant length that does not depend on the number of hops along the path. These fields are defined as a new IPv6 option type, referred to as the Transit Measurement option.

The Transit Measurement option includes two main fields: Accumulated Delay and Status Bitmap. The Accumulated Delay field is used for measuring the one-way delay along the path. Each transit node incrementally adds its internal transit delay to the Accumulated Delay field, and thus at the end of the path this field includes the sum of the transit delay values of the nodes along the path. The Status Bitmap field includes a per-hop bit that indicates its congestion status. Each node along the path updates its corresponding status bit, indicating whether the node is congested. The criterion for deciding whether a node is congested is similar to the "Congestion Experienced" trigger in ECN [[RFC3168](#)].

The Transit Measurement option can be incorporated into all or a subset of the traffic that is forwarded by the source node. Notably, the Transit Measurement option adds a fixed and low overhead to data packets, which remains constant along the path.

There are several potential use cases for the Transit Measurement option, including:

- * Performance Monitoring: the Transit Measurement option can be used for continuously tracking the network and for detecting a potentially problematic state that requires further analysis. In case a potential problem is detected by the destination node, the node may take further steps to report and to analyze the problem. For example, the node can export the packet, along with additional telemetry data to a collector, or it may log the problem locally. Upon detecting such a problem, a centralized collector/analyzer may trigger a more fine-grained measurement, e.g., an IOAM trace option [[RFC9197](#)] can be enabled in order to obtain detailed information about the performance along the path and to pinpoint the potential problem. It should be noted that logging, exporting, and further analysis by the central entity are not within the scope of the current document.
- * Path Selection: in a network that uses segment routing a source can choose which of the available paths to use to each destination. By using the Transit Measurement option the source can probe each of the available paths to a given destination, and choose the path with the best performance. If there is a performance issue along one of the paths, the fine-grained status bitmap enables the source to pinpoint the location of the issue, and to try to pick an alternative path that avoids this point. Alternatively, a path can be selected by the combination of the measured delay along the path and the status bitmap; for example, the path with less bits set in the status bitmap can be the preferred path, and the measured delay can be used as a tiebreaker.

- * Congestion control: existing congestion control algorithms periodically measure the round-trip time (RTT), and can optionally use ECN indications as a criterion for determining the congestion window. It has been shown [[SIGCOMM-HPCC](#)] that congestion control can benefit from using fine-grained information about the congestion state of the routers along the path that is sent back to the source over acknowledgment packets. The Transit Measurement option provides fine-grained measurement information that has a lightweight cost in comparison to alternative per-hop measurement protocols.

2. Conventions

2.1. Requirement Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [[RFC2119](#)] [[RFC8174](#)] when, and only when, they appear in all capitals, as shown here.

2.2. Terminology

Abbreviations used in this document:

OAM: Operations, Administration, and Maintenance

ECN: Explicit Congestion Notification

3. Transit Measurement Option

This document defines a new IPv6 Option type, the Transit Measurement type, which can be included either in the Hop-by-Hop Options header or in the Destination Options header. Figure 1 presents the format of the Transit Measurement option type.

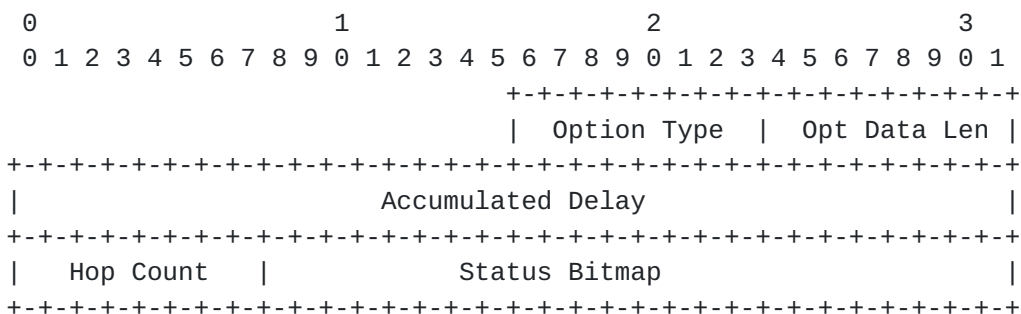


Figure 1: Transit Measurement Option Type

A node that complies to this draft MUST support the following fields, as depicted in Figure 1:

Option Type: This document assigns the value TBD-type, which indicates that this is the Transit Measurement option. The two highest order bits are "00", indicating that nodes that cannot process this option skip over it and continue processing the header, as defined in [\[RFC8200\]](#). The third-highest-order bit is set to "1", indicating that this option may change en route.

Opt Data Len: The length in octets of the two fields that follow, which is 8.

Accumulated Delay: represents the sum of the transit delay values in nanoseconds along the path of the packet, including the current node. This field is a 4-octet unsigned integer in the range 0 to $2^{31}-1$. A transit node that receives the Accumulated Delay field updates its value by adding the transit delay of the current node to the value of the Accumulated Delay field in the received packet. The 'transit delay' in this context is the time in nanoseconds the packet spent in the transit node. If the Accumulated Delay exceeds $2^{31}-1$ nanoseconds then the most significant bit is set to indicate overflow and the value is set to 0x80000000. If a transit node receives this field with the value 0x80000000 or if the node is not able to update the value of the field it SHOULD forward the packet with the unmodified field.

Hop Count/Status Bitmap: indicates the devices along the path that have experienced congestion.

Hop Count: a one-octet field that indicates the number of hops since the source node. The source node initializes this field to 0. Every transit node that supports this option increments this field by 1. A maximum of 24 hops is supported. If a transit node receives this field with the value 24 it assigns the value of all '1's (0xFF = 255), which indicates that the number of hops has exceeded the maximum.

Status Bitmap: a three octet field that represents the congestion status of each transit node along the path. The value '1' indicates that the current packet was enqueued in a queue that is congested. The criterion for whether a queue is congested or not is identical to the "Congestion Experienced" trigger in ECN. Every transit node that supports the Transit Measurement option updates the bit corresponding to the current Hop Count, after having updated the value of the Hop Count. For example,

the first transit node along the path updates the Hop Count to 1, and then updates the first (most significant bit) of the Status Bitmap. The source and destination nodes do not update the Status Bitmap. If a node detects that the Hop Count has exceeded its maximal value (24), it does not update the Status Bitmap field.

4. IANA Considerations

4.1. IPv6 Option Type

IANA is requested to allocate a value from the IPv6 Destination Options and Hop-by-Hop Options registry:

Value: TBD-type

act: 00

chg: 1

Description: Transit Measurement

5. Security Considerations

The Transit Option, and specifically the Accumulated Delay field, which is defined in this document, may be used for reconnaissance, which in turn can facilitate other types of attacks. As in other types of Operations, Administration and Maintenance (OAM) protocols, a malicious attacker can manipulate the Accumulated Delay value in order to create a false illusion of nonexistent network issues or prevent the detection of actual ones.

6. References

6.1. 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, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, [RFC 8200](https://www.rfc-editor.org/info/rfc8200), DOI 10.17487/RFC8200, July 2017, <<https://www.rfc-editor.org/info/rfc8200>>.

6.2. Informative References

[I-D.ali-spring-ioam-srv6]
Ali, Z., Gandhi, R., Filsfils, C., Brockners, F., Nainar, N. K., Pignataro, C., Li, C., Chen, M., and G. Dawra, "Segment Routing Header encapsulation for In-situ OAM Data", Work in Progress, Internet-Draft, [draft-ali-spring-ioam-srv6-06](https://www.ietf.org/archive/id/draft-ali-spring-ioam-srv6-06), 10 July 2022, <<https://www.ietf.org/archive/id/draft-ali-spring-ioam-srv6-06.txt>>.

[I-D.filsfils-spring-path-tracing]
Filsfils, C., Abdelsalam, A., Camarillo, P., Yufit, M., Graf, T., Su, Y., Matsushima, S., Valentine, M., and A. Dhamija, "Path Tracing in SRv6 networks", Work in Progress, Internet-Draft, [draft-filsfils-spring-path-tracing-02](https://www.ietf.org/archive/id/draft-filsfils-spring-path-tracing-02), 16 August 2022, <<https://www.ietf.org/archive/id/draft-filsfils-spring-path-tracing-02.txt>>.

[I-D.ietf-ippm-ioam-ipv6-options]
Bhandari, S. and F. Brockners, "In-situ OAM IPv6 Options", Work in Progress, Internet-Draft, [draft-ietf-ippm-ioam-ipv6-options-09](https://www.ietf.org/archive/id/draft-ietf-ippm-ioam-ipv6-options-09), 11 October 2022, <<https://www.ietf.org/archive/id/draft-ietf-ippm-ioam-ipv6-options-09.txt>>.

[I-D.kumar-ippm-ifa]
Kumar, J., Anubolu, S., Lemon, J., Manur, R., Holbrook, H., Ghanwani, A., Cai, D., Ou, H., Li, Y., and X. Wang, "Inband Flow Analyzer", Work in Progress, Internet-Draft, [draft-kumar-ippm-ifa-05](https://www.ietf.org/archive/id/draft-kumar-ippm-ifa-05), 12 August 2022, <<https://www.ietf.org/archive/id/draft-kumar-ippm-ifa-05.txt>>.

[I-D.zhou-ippm-enhanced-alternate-marking]
Zhou, T., Fioccola, G., Liu, Y., Cociglio, M., Lee, S., and W. Li, "Enhanced Alternate Marking Method", Work in Progress, Internet-Draft, [draft-zhou-ippm-enhanced-alternate-marking-11](https://www.ietf.org/archive/id/draft-zhou-ippm-enhanced-alternate-marking-11), 29 August 2022, <<https://www.ietf.org/archive/id/draft-zhou-ippm-enhanced-alternate-marking-11.txt>>.

- [RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", [RFC 3168](#), DOI 10.17487/RFC3168, September 2001, <<https://www.rfc-editor.org/info/rfc3168>>.
- [RFC9197] Brockners, F., Ed., Bhandari, S., Ed., and T. Mizrahi, Ed., "Data Fields for In Situ Operations, Administration, and Maintenance (IOAM)", [RFC 9197](#), DOI 10.17487/RFC9197, May 2022, <<https://www.rfc-editor.org/info/rfc9197>>.
- [SIGCOMM-HPCC]
Li, Y., Miao, R., Liu, H., Zhuang, Y., Fei Feng, F., Tang, L., Cao, Z., Zhang, M., Kelly, F., Alizadeh, M., and M. Yu, "HPCC: High Precision Congestion Control", ACM SIGCOMM Beijing, China, August 2019.

Authors' Addresses

Tal Mizrahi
Huawei
8-2 Matam
Haifa 3190501
Israel
Email: tal.mizrahi.phd@gmail.com

Tianran Zhou
Huawei
156 Beiqing Rd.
Beijing
100095
China
Email: zhoutianran@huawei.com

Shahar Belkar
Huawei
8-2 Matam
Haifa 3190501
Israel
Email: shahar.belkar@huawei.com

Reuven Cohen
Huawei
8-2 Matam
Haifa 3190501
Israel

Email: reuven.cohen@huawei.com