rtgwg Q. Xiong
Internet-Draft X. Zhu
Intended status: Standards Track ZTE Corporation
Expires: 13 April 2026 10 October 2025

Indication of Load-balancing Strategy
draft-xz-rtgwg-load-balancing-indication-00

Abstract

This document proposes the encoding of the indication for load-balancing strategies which can be encapsulated into a variety of protocols such as MPLS, IPv6 and SRv6 networks. It also provides the considerations for load-balancing configuration in control plane as well.

Table of Contents

1. Introduction

The load balancing serves as a critical mechanism for achieving high-performance transmission in wide-area networks (WANs), with diverse implementation strategies such as flow-level and packet-level load balancing, some approaches include:

*Equal-Cost Multi-Path (ECMP) and Weighted ECMP (WCMP): these mechanisms preserve the transmission order of packets within a data stream. When the network guarantees transmission and link reliability, packets typically arrive at the receiver in sequence. This eliminates the need for complex packet reordering at the receiver end, allowing simplified Go-Back-N (GBN)retransmission mechanisms to be adopted, thereby reducing device complexity and cost.

*Flowlet-based Load Balancing: this mechanism segments packets into micro-flowlets, which may be divided based on time intervals, message granularity, or fixed lengths. While these fragments are transmitted across different paths to maintain intra-flowlet ordering, they may cause inter-flowlet disorder, significantly increasing the receiver's buffer and reordering requirements.

*Packet Spraying: by randomly or load-aware proportionally distributing packets across multiple paths, this mechanism maximizes load balancing efficiency. However, it results in the highest packet disorder level, imposing the most stringent demands on the receiver's reordering capabilities.

In WANs with dynamic topologies, the load balancing strategy requires host-to-network collaboration to adapt to varying transmission scenarios. The selection of load balancing strategies should consider:

*traffic characteristics (e.g., flow size, burstiness, sensitivity to delay)

*path characteristics (e.g., latency, bandwidth, loss rate, and reliability)

*receiver’s reordering capability (e.g., buffer size, reordering algorithm efficiency)

This document proposes the encoding of the indication for load-balancing strategies which can be encapsulated into a variety of protocols such as MPLS, IPv6 and SRv6 networks. It also provides the considerations for load-balancing configuration in control plane as well.

2. Conventions Used in This Document

2.1. Abbreviations

2.2. Requirements 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.

3. Load Balancing Strategies

The reordering capabilities of incoming traffic flows may differ across servers. For example, standard RDMA NICs (Network Interface Cards) almost lack out-of-order reordering capabilities, while advanced DPU NICs achieve reordering through high-speed caches. The TCP receiving end can also perform out-of-order reordering based on caching. At network nodes, load balancing for flows is a local behavior and cannot perceive whether the server has the corresponding out-of-order recovery capability. The information from host-and-network collaboration may help to select the best load-balancing strategy.

```
          Traffic 1/2/3           Apply the load-balancing strategy
+--------+        +---------+      +-------+     +--------+      +---+---+     +-------+
|Client A|------->|Edge Node|----->|Node X |---->| Node A |----->|Node Y |<--->|Server |
+--------+        +---------+      +-------+  |  +--------+  ^   +-------+     +-------+
                 Insert the indication        |  +--------+  |
                 of load-balancing strategies +->| Node B |--+
                                              |  +--------+  ^
                                              |  +--------+  |
                                              +->| Node C |--+
                                                 +--------+
```

Figure 1: Selection of Load-balancing Strategies

As shown in Figure 1, at the edge node, the information can be carried in packets to indicate the preferred load-balancing strategies. The indication may be carried in data packets throughout the network path. And the network node can route the packets based on the indication of load-balancing strategy, allowing the network to apply different strategies for different types of flows.

4. Encapsulation for Load-balancing Indication

4.1. Indication of Load-balancing Strategies

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Flag   |F|L|P|
+---------------+
```

Figure 2: Indication Flag of Load-balancing Strategies

F: 1 bit, when F is set to 1, it indicates the flow-based load-balancing mechanism.

L: 1 bit, when L is set to 1, it indicates the flowlet-based load-balancing mechanism.

P: 1 bit, when P is set to 1, it indicates the packet-based load-balancing mechanism.

4.2. MPLS MNA Encapsulation

The indication flag could be encapsulated into MPLS MNA header as per MNA encapsulation specified in [I-D.ietf-mpls-mna-hdr].

4.3. IPv6 Header Encapsulation

The indication flag can be encapsulated into an IPv6 HbH EH as per [RFC8200] since it may be processed by transit nodes along the path in IPv6 networks.

The indication flag can also be encapsulated into an DOH EH as per [RFC8200] before an SRH since it may be processed by the forwarding nodes of the SRv6 segment list in SRv6 networks.

5. Considerations for Load-balancing Policy in Control Plane

The client or the server may notify the out-of-order recovery capability to the edge node or the controller in control plane. The controller can also sense the characteristics of a flow and decide the best load-balancing strategy and configure the indication to the edge node. And the edge node can encode the information into the encapsulation of the packets.

6. Security Considerations

To be discussed in future versions of this document.

7. IANA Considerations

TBA.

8. References

8.1. Normative References

[I-D.ietf-mpls-mna-hdr]
Rajamanickam, J., Gandhi, R., Zigler, R., Song, H., and K. Kompella, "MPLS Network Action (MNA) Sub-Stack Solution", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-hdr-16, 3 October 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-hdr-16>.

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

8.2. Informative 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/rfc/rfc2119>.

[RFC768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, DOI 10.17487/RFC0768, August 1980, <https://www.rfc-editor.org/rfc/rfc768>.

[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/rfc/rfc8174>.

Authors' Addresses

Quan Xiong
ZTE Corporation
Email: xiong.quan@zte.com.cn

Xiangyang Zhu
ZTE Corporation
Email: zhu.xiangyang@zte.com.cn