

OPSAWG
Internet-Draft
Intended status: Informational
Expires: 3 September 2026

K. Liu
R. Wang
China Mobile
2 March 2026

Alternate marking usage for loss location in per-packet load balancing
networks
draft-liu-opsawg-alt-mark-per-packet-00

Abstract

Many per-packet load balancing schemes have been proposed to mitigate network load imbalances. However, due to the randomness of packet paths, loss location is challenging in per-packet load balancing networks. An efficient solution is to leverage the alternate packet marking technique. This draft analyzes the usage and requirements of alternate packet marking for packet loss detection and location in per-packet load balancing networks.

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 3 September 2026.

Copyright Notice

Copyright (c) 2026 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
1.1. Requirements Language	3
2. Use cases in per-packet load balancing networks	3
2.1. Monitoring all packet loss on switches	4
2.2. Monitoring packet loss of certain services	4
2.3. Locating packet loss in probing systems	4
2.4. Low overhead requirements	4
2.5. Compatibility requirements	5
3. Use cases in packet-spraying networks	5
4. IANA Considerations	5
5. Security Considerations	5
6. References	6
6.1. Normative References	6
Authors' Addresses	6

1. Introduction

To mitigate network load imbalance, many per-packet load balancing schemes have been proposed. These schemes spray packets onto parallel paths to fundamentally eliminate network load imbalance. However, packet spraying brings challenges for loss detection. In flow-based networks, packets with the same 5-tuple traverse the same network path. By collecting loss packets and replaying their 5-tuples via path tracking tools such as Traceroute or INT, their network paths can be easily obtained for loss location. However, in per-packet load balancing networks, packets with the same 5-tuple may be randomly routed to different network paths, and the replayed packets may take different paths, leading to incorrect path tracking and loss location.

One possible loss location scheme in per-packet load balancing networks is to monitor packet loss on switches. However, traditional packet loss monitoring on switches cannot accurately detect all packet loss, such as silent packet loss. To accurately detect all packet loss on switches, an efficient method is to leverage the alternate packet marking technique. The core workflow of alternate packet marking is as follows: Firstly, packets are periodically and

alternately marked at the traffic entry points, such as source network interface cards (NICs) or top-of-rack (ToR) switches. Secondly, in each period, each switch calculates the difference between the ingress and egress packet counts in the previous period. At the destination point (the destination ToR switch or NIC), the marks on the packets are cleared before delivery to the service process. This draft analyzes the usage and requirements of alternate packet marking to detect and locate packet loss in per-packet load balancing networks.

1.1. 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.

2. Use cases in per-packet load balancing networks

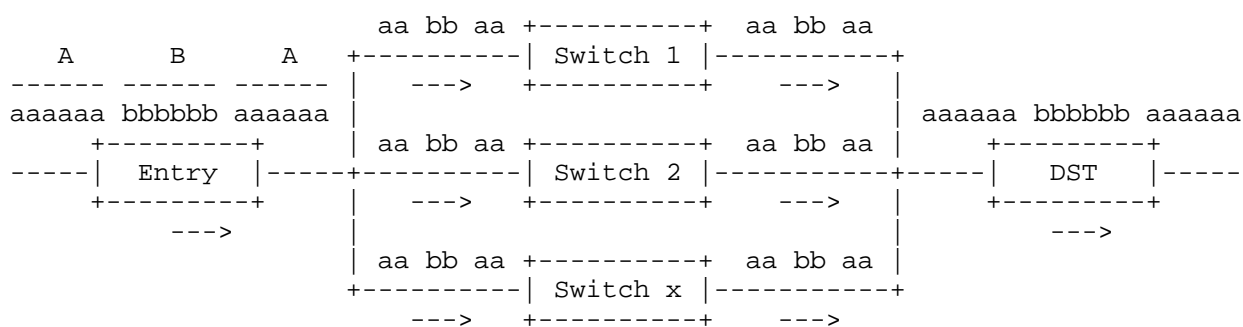


Figure 1

Figure 1 illustrates the workflow of alternate packet marking and loss counting in per-packet load balancing networks. During period A, packets are marked with flag "a" at the entry point. These marked packets are sprayed onto parallel paths, and the switches count all ingress and egress packets labeled "a". During period B, packets are marked with flag "b", and the switches count all packets labeled "b". Since no packets are labeled "a" in period B, their count remains unchanged. Each switch then calculates the difference between the number of ingress and egress packets labeled "a" to determine the total packet loss in period A. In the next period A, the switches count packets labeled "a" and calculate the difference of packets labeled "b". This process is repeated in subsequent periods. The

alternate marking ensures that the packet count in the last marking period remains unchanged, allowing for an accurate loss counting. This method can effectively detect nearly all packet loss in switching, including silent loss, which can hardly be detected by traditional packet loss monitoring on switches.

2.1. Monitoring all packet loss on switches

By marking all packets within the cluster alternately and calculating the difference between the number of ingress and egress packets on each switch, all packet loss in switching can be accurately detected. In addition, this method enables accurate loss rate monitoring for each switch, which can be used to identify abnormal switch devices.

2.2. Monitoring packet loss of certain services

Services typically have varying degrees of sensitivity to packet loss. Some services, such as distributed storage and distributed training, are highly sensitive to packet loss. For these services, it is necessary to detect and locate every packet loss. Conversely, some services, such as audio and video streaming, are less affected by packet loss. For these services, focusing only on severe packet loss events is typically sufficient. By marking packets in loss-sensitive services merely, switches can focus on packet loss event only in these services.

2.3. Locating packet loss in probing systems

Network probing systems typically proactively construct probe packets to measure network latency and packet loss rates. By replaying the anomalous probe 5-tuples (timeout or high latency) via path tracking tools, such as Traceroute or INT, these systems can further locate the anomalous device. However, in per-packet load balancing networks, the replayed probes may take different paths, resulting in an incorrect fault location. With alternate probe packet marking, the loss of probe packets can be accurately located.

2.4. Low overhead requirements

First, this method requires traffic entry points to identify and mark specific packets. Then, all switches in the cluster must recognize marked packets and determine their ingress and egress counts. Finally, at the destination point, the marks on the packets must be cleared before they are delivered to the service process. These steps introduce additional processing and latency overhead. Furthermore, if an extra header is used for packet marking, additional bandwidth overhead will be incurred. Therefore, the marking method should have minimal overhead to minimize its impact on

network performance.

2.5. Compatibility requirements

This method requires all entry/destination points to identify specific packets and add/remove packet labels. In addition, it requires all switches in the cluster to identify and count marked packets. Therefore, the scheme should be compatible with most existing switches to minimize deployment overhead.

3. Use cases in packet-spraying networks

The alternate packet marking method, a typical hybrid performance monitoring technology, has been standardized via a series of IETF RFCs to enable high-precision packet loss detection and localization.

[RFC8321] laid the foundation for alternate packet marking. It divides service flows into one-bit-marked, alternating blocks and calculates packet loss by counting the differences between adjacent measurement points. [RFC8321] supports passive and hybrid modes, and can be applied to IP, MPLS, and Ethernet networks. However, its poor anti-out-of-order performance limits its use in high-precision applications. To address this issue, [RFC9341] obsoleted [RFC8321] as an enhanced standard. The new standard (1) introduces unique block IDs to address out-of-order and retransmission interference, (2) standardizes latency and jitter measurement with D bits, and (3) unifies counting alignment. [RFC9341] greatly improves the measurement accuracy.

As a supplement to [RFC9341], [RFC9342] supports multicast scenarios with multi-receiver synchronization. [RFC9343] defines IPv6 encapsulation of alternate marking information that can be inserted into the hop-by-hop or destination options header. Therefore, it can be applied to IPv6/SRv6 networks. In terms of supporting RFCs, [RFC7799] classifies measurement methods and provides a basis for alternate marking. [RFC6374] (MPLS OAM) enables alternate marking in MPLS networks. In practice, iFIT builds on [RFC9341] and [RFC9343] and is widely used in smart metropolitan area networks and data center networks.

4. IANA Considerations

There are no IANA consideration introduced by this draft.

5. Security Considerations

There are no security issues introduced by this draft.

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>>.
- [RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", RFC 6374, DOI 10.17487/RFC6374, September 2011, <<https://www.rfc-editor.org/info/rfc6374>>.
- [RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <<https://www.rfc-editor.org/info/rfc7799>>.
- [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>>.
- [RFC8321] Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate-Marking Method for Passive and Hybrid Performance Monitoring", RFC 8321, DOI 10.17487/RFC8321, January 2018, <<https://www.rfc-editor.org/info/rfc8321>>.
- [RFC9341] Fioccola, G., Ed., Cociglio, M., Mirsky, G., Mizrahi, T., and T. Zhou, "Alternate-Marking Method", RFC 9341, DOI 10.17487/RFC9341, December 2022, <<https://www.rfc-editor.org/info/rfc9341>>.
- [RFC9342] Fioccola, G., Ed., Cociglio, M., Sapio, A., Sisto, R., and T. Zhou, "Clustered Alternate-Marking Method", RFC 9342, DOI 10.17487/RFC9342, December 2022, <<https://www.rfc-editor.org/info/rfc9342>>.
- [RFC9343] Fioccola, G., Zhou, T., Cociglio, M., Qin, F., and R. Pang, "IPv6 Application of the Alternate-Marking Method", RFC 9343, DOI 10.17487/RFC9343, December 2022, <<https://www.rfc-editor.org/info/rfc9343>>.

Authors' Addresses

Kefe Liu
China Mobile
China

Email: liukefei@chinamobile.com

Ruixue Wang

China Mobile

China

Email: wangruixue@chinamobile.com