Multi-Stage Transparent Server Load Balancing
draft-matsuhira-mslb-06

Abstract

This document specifies Multi-Stage Transparent Server Load Balancing (MSLB) specification. MSLB make server load balancing over Layer3 network without packet header change at client and server. MSLB make server load balancing with any protocol and protocol with encryption such as IPsec ESP, SSL/TLS.

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 [RFC2119].

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 http://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 December 3, 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 (http://trustee.ietf.org/license-info) in effect on the date of
1. Introduction

This document specifies Multi-Stage Transparent Server Load Balancing (MSLB) specification.

MSLB provide server load balancing function over Layer3 network without packet header change at client and server. MSLB work with any protocol and protocol with payload encryption such as IPsec ESP, SSL/TLS.

2. Traditional load balancing method

There are several load balancing technique, such as round robin DNS, IP Anycasting [RFC1546] and destination address translation. Figure 1 shows load balancing system with typical server load balancer with destination address translation technique.

![Diagram of traditional load balancing method]

It is well-known that Network address translator break internet transparency [RFC2775] and have a application dependency [RFC2993] characteristic.

Some server load balancer use application data, so with IPsec ESP, SSL/TLS, this mechanisms may not work well.

3. Architecture of MSLB

Load balancing is the technique that distribute packet to multiple server. For packet distribution, destination addresss translation technique is useful, however this technique itself break internet transparency.
After distribution, if write back to the original destination address may possible, it is possible to recover transparency. This is the basic idea and architecture of MSLB. Figure 2 shows architecture of MSLB.

```
Client ----  overwrite    +----------  write back  ----- server
destination  |                        |
address       + ----------  write back  ----- server
               |                        |
               :              :            :
               + ----------  write back  ----- server
```

Figure 2

This method process only destination address of IP header. This method can be applied to both IPv4 and IPv6.

4. configuration

4.1. basic configuration

Figure 3 shows basic server load balancing system with MSLB. This case two-stage configuration with one MSLB-F and one-stage many MSLB-Bs.

```
+-------+   +------+   +------+
|       +---+MSLB-B+---+Server|
+-------+   +------+   |       |   +------+   +------+
|       |   |      |   |       |       :          :
|      |   |      |   |       |   +------+   +------+
|Client+---+Network+---+MSLB-F+---+Network+---+MSLB-B+---+Server|
|      |   |      |   |       |       :          :
|      |   |      |   |       |   +------+   +------+
|      +---+MSLB-B+---+Server|
```

Figure 3

MSLB-F is front function of MSLB and translate destination address to one of the address of MSLB-B. BSLB-B is backend function of MSLB and
translate destination address to the original server address, i.e. address of MSLB-F. The IP address of MSLB-F and all server is the same value.

MSLB-F may multi-stage configuration. Figure 4 shows three stage configuration with two-stage MSLB-F and one-stage many MSLB-Bs.

4.2. one arm configuration

Figure 5 shows one arm configuration of server load balancing system with MSLB.
MSLB-F is front function of MSLB and translate destination address to one of the address of MSLB-B. BSLB-B s backend function of MSLB and translate destination address to the original server address, i.e. address of MSLB-F. The IP address of MSLB-F and all server is the same value.

This configuration, MSLB-F is connecting to the network with single link, that is one arm configuration. This case, return packet, i.e. packet from server to client does not pass through the MSLB-F.

5. mode

MSLB have two mode, one is address translation mode, and the other is encapsulation mode.

5.1. address translation mode

This mode using address translation technique.

Figure 6 shows packet processing with address translation mode.
In this figure, to the Client, IP address is allocated IP_C1, IP_C2, and server IP address is IP_S. This case, IP_S is also allocate to all servers and MSLB-F. And to the MSLB-B, IP_B1, IP_B2, IP_B3 is allocated. These allocation is shown in upper part of Figure 6.

Lower part of Figure 6 shows packet transfered between client and server. From Client to the Server, only destination address is translate, MSLB-F translate from IP_S to IP_B1, and MSLB-B translate from IP_B1 to IP_S. Then the destination address of packet which send client and the destination address of packet which recieve server is
same address. That mean, transparency is remained.

Return packet, i.e., from server to the client is not translate, just forwarded.

In the Internet, Client IP address and server IP address must Global IP address, however, IP address of MSLB-B may private IP address.

<table>
<thead>
<tr>
<th>Source IP address</th>
<th>net mask</th>
<th>destination IP address</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP_C1</td>
<td></td>
<td>IP_B1</td>
</tr>
<tr>
<td>IP_C2</td>
<td></td>
<td>IP_B2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7

Figure 7 shows MSLB table. MSLB have this table and translate the destination address using this table value. MSLB-F check source IP address, and translate destination address with this table.

Using IPv4-IPv6 translation may possible, i.e., IPv4 packet translated to IPv6, then translate to IPv4 or IPv6 packet translate to IPv4, then translate IPv6 may possible. Figure 8 shows possible combination of IPv4 and IPv6. These IPv4-IPv6 translation case will be defined in future.
5.2. encapsulation mode

This mode uses encapsulation technique.

Figure 9 shows packet processing with encapsulation mode.
In this figure, to the Client, IP address is allocated IP_C1, IP_C2, and server IP address is IP_S. This case, IP_S is also allocate to all servers and MSLB-F. And to the MSLB-B, IP_B1, IP_B2, IP_B3 is allocated. These allocation is shown in upper part of Figure 6.

Lower part of Figure 6 shows packet transferred between client and server. From Client to the Server, MSLB-F encapsulate original IP packet and send to MSLB-B. MSLB-B decapsulate outer IP header, and forward to the server. Inner IP packet does not change, that mean, transparency is remained.

With encapsulation mode, packet size is increase, so fragmentation is needed if encapsulated packet size exceed MTU or Path MTU. MSLB-F MUST support tunnel MTU discovery [RFC1853]. Fragmentation and Path MTU discovery [RFC1191] issue will describe in future.

Return packet, i.e., from server to the client is not encapsulate, just forwarded.

In the Internet, Client IP address and server IP address must Global IP address, however, IP address of MSLB-B may private IP address.

```
+-----------------+----------+-------------------------+
| Source IP address | net mask | destination IP address |
+-----------------+----------+-------------------------+
| IP_C1           |          | IP_B1                  |
+-----------------+----------+-------------------------+
| IP_C2           |          | IP_B2                  |
+-----------------+----------+-------------------------+
|                 |    :     |             :           |
|                 |    :     |             :           |
|                 |    :     |             :           |
+-----------------+----------+-------------------------+
```

**Figure 10**

Figure 10 shows MSLB table. MSLB have this table and encapsulate and generate outer header with destination address using this table value. MSLB-F check source IP address, and generate destination address of outer header with this table.

Using IPv4 over IPv6 encapsulation or IPv6 over IPv4 encapsulation may possible, i.e., IPv4 packet encapsulated to IPv6, then
decapsulate to IPv4 or IPv6 packet encapsulated to IPv4, then
deencapsulated IPv6 may possible. Figure 11 shows possible combination
of IPv4 and IPv6. These IPv4-IPv6 encapsulation case will be defined
in future.

<table>
<thead>
<tr>
<th>Client</th>
<th>MSLB-F</th>
<th>MSLB-B</th>
<th>Server</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td>(1)</td>
<td>&lt;-- IPv4 --&gt;</td>
<td>:</td>
<td>&lt;-- IPv4 --&gt;</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td>(2)</td>
<td>&lt;-- IPv6 --&gt;</td>
<td>:</td>
<td>&lt;-- IPv6 --&gt;</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td>(3)</td>
<td>&lt;-- IPv4 --&gt;</td>
<td>:</td>
<td>&lt;-- IPv4 --&gt;</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td>(4)</td>
<td>&lt;-- IPv6 --&gt;</td>
<td>:</td>
<td>&lt;-- IPv6 --&gt;</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
</tbody>
</table>

Figure 11

6. Ingress filtering environment

[RFC2827] describe ingress filtering for defending DoS attack which
employ IP source address spoofing.

Depend on the location of the MSLB-F and MSLB-B, it is possible that
packet from server to client is discarded by ingress filtering. In
such case, encapsulating the packet from server to client might
resolve. Figure 12 shows such solution.
7. Characteristic

MSLB has following characteristics.

- Layer 3 Load balancer
- Support NAT unfriendly application such as FTP
- work with any application layer protocol (maybe)
- work with encryption (IPsec ESP, SSL/TLS)
- work over Layer 3 network
may enforce policy with static configuration

8. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

9. Security Considerations

Security consideration does not discussed in this memo.

10. Acknowledgements

11. References

11.1. Normative References


[RFC2993] Hain, T., "Architectural Implications of NAT", RFC 2993,
DOI 10.17487/RFC2993, November 2000,

11.2. Informative References

Author’s Address

Naoki Matsuhira
Fujitsu Limited
17-25, Shinkamata 1-chome, Ota-ku
Tokyo, 144-8588
Japan

Phone: +81-3-3735-1111
Fax:
Email: matsuhira@jp.fujitsu.com