Network Working Group N. Matsuhira
Internet-Draft Fujitsu Limited
Intended status: Informational March 29, 2017
Expires: September 30, 2017

Multi-Stage Transparent Server Load Balancing
draft-matsuhira-mslb-02

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 encription 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 September 30, 2017.

Copyright Notice

Copyright (c) 2017 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 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 3
2. Traditional load balancing method 3
3. Architecture of MSLB 3
4. configuration 4
   4.1. basic configuration 4
   4.2. one arm configuration 5
5. mode 6
   5.1. address translation mode 6
   5.2. encapsulation mode 9
6. Ingress filtering environment 11
7. Characteristic 12
8. IANA Considerations 13
9. Security Considerations 13
10. Acknowledgements 13
11. References 13
   11.1. Normative References 13
   11.2. Informative References 14
Author's Address 14

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 encription 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.

```
                                         +---------+   +--------+
                                         |         +---+ Server |
              +---------+   +----------+   |         |   +--------+
              |         |   |          |   |         |       :
  +--------+  |         |   |  Server  |   |         |   +--------+
  | Client +---+ Network +---+   Load   +---+ Network +---+ Server |
  +--------+  |         |   | Balancer |   |         |   +--------+
              |         |   |          |   |         |       :
              +---------+   +----------+   |         |   +--------+
                                         |         +---+ Server |
                                         +---------+   +--------+
```

Figure 1

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 tecnique 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 possoble, 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 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.

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

```
                                            +---+  +------+  +------+
                                            |   |--+MSLB-B+--+Server|
                             +---+          |   |  +------+  +------+
                             |   |  +----+  |Net|      :         :
              +---+  +----+  |   |  |MSLB|  |   |  +------+  +------+
              |   |  |    |  |   |--+ -F +--+   |--+MSLB-B+--+Server|
    +------+  |   |  |    |  |   |  +----+  +---+  +------+  +------+
    |Client+--+Net+--+MSLB+--+Net|
    +------+  |   |  | -F |  |   |  +----+  +---+  +------+  +------+
              |   |  |    |  |   +--+MSLB+--+   |--+MSLB-B+--+Server|
              +---+  +----+  |   |  | -F |  |   |  +------+  +------+
                             |   |  +----+  |Net|      :         :
                             +---+          |   |  +------+  +------+
                                            |   |--+MSLB-B+--+Server|
                                            +---*  +------+  +------+
```

Figure 4

4.2. one arm configuration

Figure 5shows one arm configuration of server load balancing system with MSLB.

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

Figure 5

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, retuen 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 Figure 6 shows packet processing with address translation mode.

```
                                      +-------+   +------+   +------+
                                      |       +---+MSLB-B+---+Server|
    +------+                          |       |   | IP_B1|   | IP_S |
    |Client|   +-------+   +------+   |       |   +------+   +------+
    | IP_C1+---+       |   |      |   |       |
    +------+   |       |   |      |   |       |   +------+   +------+
               |Network|   |MSLB-F|---+Network+---+MSLB-B+---+Server|
               |       +---+      |   |       |   | IP_B2|   | IP_S |
    +------+   |       |   | IP_S |   |       |   +------+   +------+
    |Client+---+       |   |      |   |       |
    | IP_C2|   +-------+   +------+   |       |   +------+   +------+
    +------+                          |       +---+MSLB-B+---+Server|
                                      |       |   | IP_B3|   | IP_S |
                                      +-------+   +------+   +------+
                              :                       :
                              :                       :
      +------+----+           :   +------+----+       :+------+----+
      | data | IP |           :   | data | IP |       :| data | IP |
      +------+----+           :   +------+----+       :+------+----+
      ---------------------> : --------------------> : ------------>
        src = IP_C1           :    src = IP_C1        :   src = IP_C1
        dst = IP_S            :    dst = IP_B1        :   dst = IP_S
                              :                       :
      +------+----+           :   +------+----+       :+------+----+
      | data | IP |           :   | data | IP |       :| data | IP |
      +------+----+           :   +------+----+       :+------+----+
      <-------------------- -: <-------------------- : <------------
        src = IP_S            :    src = IP_S         :   src = IP_S
        dst = IP_C1           :    dst = IP_C1        :   dst = IP_C1
                              :                       :
```

Figure 6

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.

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

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 possibleFigure 8 shows possible combination of IPv4 and IPv6. These IPv4-IPv6 translation case will be defined in future.

```
     Client     MSLB-F         MSLB-B     Server
                  :              :
                  :              :
(1)  <-- IPv4 --> : <-- IPv4 --> : <-- IPv4 -->
                  :              :
(2)  <-- IPv6 --> : <-- IPv6 --> : <-- IPv6 -->
                  :              :
(3)  <-- IPv4 --> : <-- IPv6 --> : <-- IPv4 -->
                  :              :
(4)  <-- IPv6 --> : <-- IPv4 --> : <-- IPv6 -->
                  :              :
```

Figure 8

5.2. encapsulation mode

This mode using encapsulation technique.

Figure Figure 9 shows packet processing with encapsulation mode.

```
                                      +-------+   +------+   +------+
                                      |       +---+MSLB-B+---+Server|
                                      |       |   | IP_B1|   | IP_S |
               +-------+   +------+   |       |   +------+   +------+
               |       |   |      |   |       |
    +------+   |       |   |      |   |       |   +------+   +------+
    |Client|---+Network+---+MSLB-F|---+Network+---+MSLB-B+---+Server|
    | IP_C |   |       |   |      |   |       |   | IP_B2|   | IP_S |
    +------+   |       |   | IP_S |   |       |   +------+   +------+
               |       |   |      |   |       |
               +-------+   +------+   |       |   +------+   +------+
                                      |       +---+MSLB-B+---+Server|
                                      |       |   | IP_B3|   | IP_S |
                                      +-------+   +------+   +------+
                              :                       :
                              :                       :
      +------+----+           :   +------+----+----+  :+------+----+
      | data | IP |           :   | data | IP | IP |  :| data | IP |
      +------+----+           :   +------+----+----+  :+------+----+
      ----------------------> : --------------------> : ------------>
        src = IP_C            :   Inner header        :   src = IP_C
        dst = IP_S            :    src = IP_C         :   dst = IP_S
                              :    dst = IP_S         :
                              :   Outer header        :
                              :    src = IP_S         :
                              :    dst = IP_B1        :
                              :                       :
                              :                       :
                              :                       :
      +------+----+           :   +------+----+       :+------+----+
      | data | IP |           :   | data | IP |       :| data | IP |
      +------+----+           :   +------+----+       :+------+----+
      <--------------------- -: <-------------------- : <------------
        src = IP_S            :    src = IP_S         :   src = IP_S
        dst = IP_C            :    dst = IP_C         :   dst = IP_C
                              :                       :
```

Figure 9

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, MSLB-F encapsulate original IP packet and send to MSLB-B. MSLB-B decapsulate outer IP header, and forwarad 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 possibleFigure 11 shows possible combination of IPv4 and IPv6. These IPv4-IPv6 encapsulation case will be defined in future.

```
     Client     MSLB-F                    MSLB-B     Server
                  :                         :
                  :                         :
(1)  <-- IPv4 --> : <-- IPv4 over IPv4 --> : <-- IPv4 -->
                  :                         :
(2)  <-- IPv6 --> : <-- IPv6 over IPv6 --> : <-- IPv6 -->
                  :                         :
(3)  <-- IPv4 --> : <-- IPv4 over IPv6 --> : <-- IPv4 -->
                  :                         :
(4)  <-- IPv6 --> : <-- IPv6 over IPv4 --> : <-- IPv6 -->
                  :                         :
```

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.

```
                                      +-------+   +------+   +------+
                                      |       +---+MSLB-B+---+Server|
    +------+                          |       |   | IP_B1|   | IP_S |
    |Client|   +-------+   +------+   |       |   +------+   +------+
    | IP_C1+---+       |   |      |   |       |
    +------+   |       |   |      |   |       |   +------+   +------+
               |Network|   |MSLB-F|---+Network+---+MSLB-B+---+Server|
               |       +---+      |   |       |   | IP_B2|   | IP_S |
    +------+   |       |   | IP_S |   |       |   +------+   +------+
    |Client+---+       |   |      |   |       |
    | IP_C2|   +-------+   +------+   |       |   +------+   +------+
    +------+                          |       +---+MSLB-B+---+Server|
                                      |       |   | IP_B3|   | IP_S |
                                      +-------+   +------+   +------+

                              :                       :
      +------+----+           :   +------+----+----+  :+------+----+
      | data | IP |           :   | data | IP | IP |  :| data | IP |
      +------+----+           :   +------+----+----+  :+------+----+
      <--------------------- -: <-------------------- : <------------
        src = IP_S            :   Inner header        :   src = IP_S
        dst = IP_C            :    src = IP_S         :   dst = IP_C
                              :    dst = IP_C         :
                              : Outer header          :
                              :    src = IP_B1
                              :    dst = IP_TBD
```

Figure 12

7. Characteristic

MSLB has following characteristics.

o Layer 3 Load balancer

o Support NAT unfriendly application such as FTP

o work with any application layer protocol (maybe)

o work with encription (IPsec ESP, SSL/TLS)

o 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

[RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, DOI 10.17487/RFC1191, November 1990, <http://www.rfc-editor.org/info/rfc1191>.

[RFC1546] Partridge, C., Mendez, T., and W. Milliken, "Host Anycasting Service", RFC 1546, DOI 10.17487/RFC1546, November 1993, <http://www.rfc-editor.org/info/rfc1546>.

[RFC1853] Simpson, W., "IP in IP Tunneling", RFC 1853, DOI 10.17487/RFC1853, October 1995, <http://www.rfc-editor.org/info/rfc1853>.

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

[RFC2775] Carpenter, B., "Internet Transparency", RFC 2775, DOI 10.17487/RFC2775, February 2000, <http://www.rfc-editor.org/info/rfc2775>.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, DOI 10.17487/RFC2827, May 2000, <http://www.rfc-editor.org/info/rfc2827>.

[RFC2993] Hain, T., "Architectural Implications of NAT", RFC 2993, DOI 10.17487/RFC2993, November 2000, <http://www.rfc-editor.org/info/rfc2993>.

11.2. Informative References

[] "".

Author's Address

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

Phone: +81-3-3730-8386
Fax:
Email: matsuhira@jp.fujitsu.com