Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 22, 2010

X. Li
C. Bao
H. Zhang
CERNET Center/Tsinghua University

October 19, 2009

Address-sharing stateless double IVI
draft-xli-behave-divi-00

Status of this Memo

Copyright Notice

Abstract

This document presents the concepts and the implementations of address-sharing stateless IVI (stateless 1:N IVI) and the address-

sharing stateless double IVI (stateless 1:N dIVI).

The stateless 1:N IVI keeps the features of stateless, end-to-end address transparency and bidirectional-initiated communications of the original stateless 1:1 IVI, while it can utilize the IPv4 addresses more effectively. The stateless 1:N dIVI has above features and it does not require the DNS64/DNS46 and ALG supports.

Table of Contents

1. Introduction 3
2. Terminologies 4
3. Stateless 1:N IVI 4
  3.1. Algorithm 5
  3.2. Address format 5
  3.3. Protocol translation 7
  3.4. Routing 7
  3.5. DNS 7
  3.6. ALG 7
  3.7. The translator behavior and the IPv6 end system requirements 7
4. Stateless 1:N double IVI 8
  4.1. Port number mapping algorithm 8
  4.2. Double IVI 9
  4.3. Protocol translation 10
  4.4. Home gateway implementation 10
  4.5. End system implementation 10
5. Work flow examples 11
  5.1. IPv6 initiated communication 11
  5.2. IPv4 initiated communication 11
6. Testing prototype information 11
7. Security Considerations 11
8. IANA Considerations 11
9. Acknowledgments 12
10. References 12
  10.1. Normative References 12
  10.2. Informative References 13
Authors' Addresses 13

1. Introduction

The experiences for the IPv6 deployment in the past 10 years strongly indicate that for a successful transition, the communication between IPv4 and IPv6 address families should be supported.

Recently, the stateless and stateful IPv4/IPv6 translation methods are developed and becoming the IETF standards [I-D.ietf-behave-v6v4-framework], [I-D.ietf-behave-v6v4-xlate], [I-D.ietf-behave-v6v4-xlate-stateful]. The original stateless IPv4/IPv6 translation (stateless 1:1 IVI) is scalable, maintains the end-to-end address transparency and support both IPv6 initiated and IPv4 initiated communications [I-D.ietf-behave-v6v4-framework], [I-D.ietf-behave-v6v4-xlate], [I-D.xli-behave-ivi]. But it can not use the IPv4 addresses effectively. The IPv4 address depletion problem makes the deployment of the 1:1 IVI stateless IVI difficult. The stateful IPv4/IPv6 translation can share the IPv4 addresses among IPv6 hosts, but it only supports IPv6 initiated communication [I-D.ietf-behave-v6v4-framework], [I-D.ietf-behave-v6v4-xlate-stateful]. Rely on session initiated states, the stateful translation cannot support the end-to-end address transparency and costs more compared with the stateless translation.

In this document, we present concepts and the implementations of the address-sharing stateless IVI (stateless 1:N IVI) and the address-sharing stateless double IVI (stateless 1:N dIVI). The basic concepts of these techniques are the combination of "Address plus port addressing" (A+P) and the IPv4/IPv6 stateless translation (IVI).

The stateless 1:N IVI is the extensions of the stateless 1:1 IVI. It is the solution for the following scenarios [I-D.ietf-behave-v6v4-framework].

- o Scenario 1: An IPv6 network to the IPv4 Internet.
- o Scenario 2: The IPv4 Internet to an IPv6 network.
- o Scenario 5: An IPv6 network to an IPv4 network.
- o Scenario 6: An IPv4 network to an IPv6 network.

The stateless 1:N IVI and the stateless 1:N dIVI keep all the advantages of stateless 1:1 IVI and can use the IPv4 addresses more effectively. In addition, stateless 1:N dIVI can work without DNS64/DNS46 and ALG.

2. Terminologies

This document uses the terminologies defined in [I-D.ietf-behave-v6v4-framework].

The key words MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, MAY, and OPTIONAL, when they appear in this document, are to be interpreted as described in [RFC2119].

3. Stateless 1:N IVI

The stateless 1:N IVI is shown in the following figure.

```
              --------
            //        \\        -----------
           /            \      //          \\
          /              +----+              \        ------------
          |              |XLAT|               |------ End System 1
          |  The IPv4    +----+  An IPv6      |       ------------
          |  Internet    +----+  Network      |       ------------
          |              |DNS |  (address     |------ End System N
          \              +----+   subset)    /        ------------
           \            /      \\          //
            \\        //         ----------
              --------
                         <====>
```

Figure 1: Stateless 1:N IVI

Where the XLATE is the IPv4/IPv6 translator perform 1:N translation between IPv4 and IPv6; DNS is the DNS46 and DNS64 for providing the authoritative and resolving services; the End System 1 and End System N, etc are the IPv6-only hosts which can restrict their transport layer port range when communicating with the IPv4 Internet.

In order to share the IPv4 address among IPv6 hosts, the port number multiplexing technique is used [I-D.xli-behave-ivi]. The basic idea is similar to the ones used in NAT and A+P. This is to say that a single IPv4 address can be shared for multiple IPv6 hosts under the condition that these individual hosts can only use a subset of the 65,536 port numbers when communicating with the IPv4 Internet. For example, if the port multiplexing ratio is 128, each host with IPv4-translatable address can use 512 concurrent port numbers when communicating with IPv4 Internet. Note that there is no port number restriction when these IPv6 hosts communicate with the IPv6 Internet.

3.1. Algorithm

The stateless 1:N IVI is shown in the following figure.

```
                                                   .-------|Host0| A1/(P%N)+0
                                                  /
         ------                   -----          |
       /  The   \     ------    /  An   \        |
      |  IPv4    |--|1:N   |---|  IPv6   |------------|Host1| A1/(P%N)+1
       \Internet/   |XLATE |    \Network/        |
         ------     ------        -----          |
                                                 |\
                                                 |  -------|Host2| A1/(P%N)+2
                                                 |
                                                 |
                                                  \
                                                    -------|HostK| A1/(P%N)+K
```

Figure 2: Stateless 1:N IVI

In the above figure, the Host0, Host1, Host2, ..., HostN are sharing the same IPv4 address A1, but port number range for different hosts are not overlapped. Therefore, when these IPv6 hosts communicate with the IPv4 Internet via the translator, it looks like a single host with IPv4 address A1 communicating with the IPv4 Internet.

We use the Modulus Operator to define the port number range. If the multiplexing ratio is N, then:

- o For host K, the allowed port number (P) are P=j*N + K (j=0, 1, ..., N-1).
- o For the destination port number (P), the packets will be sent to host K=(P%N) (% is the Modulus Operator).

For example: If N=256, then host K=5 is only allowed to use port numbers of 5, 261, 517, 773, ..., 65,285 as the source port, while the packets with these port numbers as the destination port number will be send to host K=5.

3.2. Address format

In order to perform the stateless translation (IVI) between the IPv4 and IPv6, both IPv4-mapped and IPv4-translatable address are required [I-D.ietf-behave-v6v4-framework]. We use the reserved 16-bits to encode the range of the port number [I-D.ietf-behave-address-format].

The IPv4-mapped addresses are used to represent IPv4 addresses in IPv6, as shown in the following figure.

```
| 0                |32 |40             |72          |88       127|
-----------------------------------------------------------------
|    LIR           |FF | IPv4 addr     | all 0                   |
-----------------------------------------------------------------
```

Figure 3: IPv4-mapped address format

Note that we use the address format and the prefix (e.g. 2001:db8:ff00::/40) defined in [I-D.xli-behave-ivi]. There is no port number coding required for the IPv4-mapped address.

The IPv4-translatable addresses are used to represent IPv6 addresses in IPv4, we defined the extended IPv4-translatable as shown in the following figure.

```
| 0                |32 |40             |72          |88       127|
-----------------------------------------------------------------
|    LIR           |FF | IPv4 addr     |Port Coding| all 0       |
-----------------------------------------------------------------
```

Figure 4: Extended IPv4-translatable address format

Where, we use reserved 16-bits to encode the port number range based on the Modulus Operator.

The most significant 4 bits define the multiplexing ratio and the least significant 12 bits define the index of the host, as shown in the following figure.

(4 bits)	Index Range(12 bits)	Multx ratio	# of Ports
0	000-000	1	65,536
1	000-001	2	32,768
2	000-003	4	16,384
3	000-007	8	8,192
4	000-00f	16	4,096
5	000-01f	32	2,048
6	000-03f	64	1,024
7	000-07f	128	512
8	000-0ff	256	256
9	000-1ff	512	128
A	000-3ff	1,024	64
B	000-7ff	2,048	32
C	000-fff	4,096	16

Figure 5: Transport layer port number coding

3.3. Protocol translation

The protocol translation is defined in [I-D.ietf-behave-v6v4-xlate].

3.4. Routing

The routing follows the general IPv4/IPv6 routing principle, i.e. "more specifics win", same as the original stateless 1:1 IVI. [I-D.xli-behave-ivi].

3.5. DNS

The DNS handling is referring to DNS64 [I-D.ietf-behave-dns64] and DNS46 [I-D.xli-behave-ivi].

3.6. ALG

The ALG related issue is discussed in [I-D.ietf-behave-v6v4-framework].

3.7. The translator behavior and the IPv6 end system requirements

For the stateless 1:N IVI, the IPv6 end systems are required to follow the port number range defined by the extended IPv4-translatable address format when communicating with the IPv4 Internet. The behaviors of the stateless 1:N translator are:

- If the packets are from the IPv4 Internet to an IPv6 network, the IPv4 source addresses are translated to the IPv4-mapped addresses and the source port numbers are unchanged; the IPv4 destination addresses are translated to the extended IPv4-translatable addresses based on the destination port number and the destination port numbers are unchanged.

- If the packets are from an IPv6 network to the IPv4 Internet, the IPv6 source addresses and the source port numbers are checked, if the source port number matches the port number range defined by the extended IPv4-translatable address format, the IPv6 source addresses (which are the IPv4-translatable addresses) are translated to the IPv4 addresses and the source port numbers are unchanged; the destination IPv6 addresses (which are the IPv4-mapped addresses) are translated to the IPv4 destination addresses and the destination port numbers are unchanged. However, if the source port numbers do not match the port number range defined by the extended IPv4-translatable address format, the packets will be dropped.

Therefore, the IPv6 end systems must follows the port number range defined by the extended IPv4-translatable addresses. The hehavor of the IPv6 end system when communicating with the IPv4 Internet are:

- If the IPv6 end system is used as a server, different well-known ports will be served by different IPv6 hosts.

- If the IPv6 end system is used as a client, the end system must generate the source port numbers in the range defined by the extended IPv4-translatable address format. This can be done by modification of the end system, or via a port number mapping device (home gateway).

4. Stateless 1:N double IVI

In general, it is not good idea to force the modification of the end system in order to meet the IPv6 end system requirements of the stateless 1:N IVI. Alternatively, we can use the home gateway to map the randomly generated source port number to the port number range defined by extended IPv4-translatable address format.

4.1. Port number mapping algorithm

The port number mapping algorithm is straightforward. The port number mapping device maintains a database of allowed port numbers defined by the extended IPv4-translatable address format. If the packets from the end system contains the source port number which do

not match the port number range defined by the extended IPv4-translatable address format, the home gateway will translate the source port number to an allowed one and keep the record in the database for translating back the returning packets and all the packets in the same session.

The maintaining of the database can be done via the corresponding trnsport layer flags for TCP or timeout for UDP.

4.2. Double IVI

If we can use the home gateway for the port number mapping, then we can also use the home gateway (1:1 Xlate) to translate the IPv6 packets back to IPv4, as shown in the following figure.

```
        ------                  -----
      /  The   \    ------     /  An   \        -----     ----
     |  IPv4    |--|1:N   |---|  IPv6   |------|1:1  |---|Host1|
      \Internet/   |XLATE |    \Network/       |XLATE|    ----
        ------      ------      -----           -----
```

Figure 6: Double IVI (dIVI)

The advantage of double IVI is that the DNS64/DNS46 and ALG are not required.

The first IPv4/IPv6 translator (1:N XLATE) is the core network translator, the second IPv4/IPv6 translator (1:1 XLATE) is the home gateway trnaslator. The features of these translators are:

Core network translator: The core network translator (1:N XLATE) is implemended in the border between the IPv6 core network and the IPv4 Internet. It translates the packets between IPv4 and IPv6 with the 1:N stateless address mapping, same as the one used in the stateless 1:N IVI.

Home gateway translator: The home gateway translator (1:1 XLATE) is implemented between an IPv6 network and user's end syatem. It translates the packets between IPv4 and IPv6 with 1:1 stateless address mapping. In addition, the home gateway translator maps the random source port numbers to restricted port number based on the extended translatable address format and keeps the mapping table in database for the port number mapping of the retuning packets and all the packets in the same session. Note that the 1:1 XLATE is still stateless for the address mapping.

4.3. Protocol translation

The protocol translation is referring to [I-D.ietf-behave-v6v4-xlate]. Special MTU and fragmentation actions must be taken, due to the process of double translation (more details).

4.4. Home gateway implementation

The home gateway implementation is suitable for the ADSL environment, as shown in the following figure.

```
                                              ----     -----
                                           .-|hgw0|---|Host0| A1/(P%N)+0
                                          /   ----     -----
   ------                 -----          |
  /  The   \    ------     /  An   \    |    ----     -----
 |  IPv4    |--|1:N   |---|  IPv6   |------|hgw1|---|Host1| A1/(P%N)+1
  \Internet/   |XLATE |    \Network/    |    ----     -----
   ------      ------      -----        |
                                        |\   ----     -----
                                        |  -|hgw2|---|Host2| A1/(P%N)+2
                                        |    ----     -----
                                        |
                                         \   ----     -----
                                           -|hgwK|---|HostK| A1/(P%N)+K
                                             ----     -----
```

Figure 7: dIVI home gateway implementation

Where Xlate is the IPv4/IPv6 stateless 1:N IVI translator; hgw0, hgw1, ..., hgwK are the home gateways performing the port number mapping and the 1:1 IPv4/IPv6 translation function; Host0, Host1, ..., HostK are dual-stack hosts who share same IPv4 address (A1), and have different non-IPv4-translatable IPv6 addresses.

4.5. End system implementation

For the wireless mobile Internet environment, it is not difficult to modify the operating system of the mobile device, therefore it possible to integrate the port number restriction and the IPv4/IPv6 translation function in the mobile device, which is an IPv6-only host to the network and has a dual-stack socket API for the applications running on this host.

```
                                                     -----------
                                                  .-|Host0 (hgw)| A1/(P%N)+0
                                                 /   -----------
        ------                  -----           |
      /  The   \     ------     /  An   \       |    -----------
     |  IPv4    |--|1:N   |---|  IPv6   |------|Host1 (hgw)| A1/(P%N)+1
      \Internet/   |XLATE |    \Network/       |    -----------
        ------      ------       -----          |
                                                |\   -----------
                                                |  -|Host2 (hgw)| A1/(P%N)+2
                                                |    -----------
                                                |
                                                 \   -----------
                                                   -|HostK (hgw)| A1/(P%N)+K
                                                     -----------
```

Figure 8: dIVI end system implementation

5. Work flow examples

5.1. IPv6 initiated communication

Add details later.

5.2. IPv4 initiated communication

Add details later.

6. Testing prototype information

The address-sharing stateless double IVI (dIVI) with 1:N stateless translator, home gateways and the end systems have been coded in Linux and deployed in the [CERNET] (IPv4) and [CNGI-CERNET2] (IPv6). The testing homepage is at [dIVI]

7. Security Considerations

There are no security considerations in this document.

8. IANA Considerations

This memo adds no new IANA considerations.

Note to RFC Editor: This section will have served its purpose if it

correctly tells IANA that no new assignments or registries are required, or if those assignments or registries are created during the RFC publication process. From the author's perspective, it may therefore be removed upon publication as an RFC at the RFC Editor's discretion.

9. Acknowledgments

The authors would like to acknowledge the following contributors in the different phases of the address-sharing IVI and dIVI development: Maoke Chen, Yu Zhai, Wentao Shang, Weifeng Jiang and Yuncehng Zhu.

The authors would like to acknowledge the following contributors who provided helpful inputs: Dan Wing, Fred Baker, Dave Thaler, Randy Bush and Kevin Yin.

10. References

10.1. Normative References

[I-D.ietf-behave-address-format]
Huitema, C., Bao, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", draft-ietf-behave-address-format-00 (work in progress), August 2009.

[I-D.ietf-behave-dns64]
Bagnulo, M., Sullivan, A., Matthews, P., and I. Beijnum, "DNS64: DNS extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", draft-ietf-behave-dns64-00 (work in progress), July 2009.

[I-D.ietf-behave-v6v4-framework]
Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", draft-ietf-behave-v6v4-framework-01 (work in progress), September 2009.

[I-D.ietf-behave-v6v4-xlate]
Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", draft-ietf-behave-v6v4-xlate-01 (work in progress), September 2009.

[I-D.ietf-behave-v6v4-xlate-stateful]
Bagnulo, M., Matthews, P., and I. Beijnum, "NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4

Servers", draft-ietf-behave-v6v4-xlate-stateful-02 (work in progress), October 2009.

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

10.2. Informative References

[CERNET] "CERNET Homepage: http://www.edu.cn/english_1369/index.shtml".

[CNGI-CERNET2] "CNGI-CERNET2 Homepage: http://www.cernet2.edu.cn/index_en.htm".

[I-D.xli-behave-ivi] Li, X., Bao, C., Chen, M., Zhang, H., and J. Wu, "The CERNET IVI Translation Design and Deployment for the IPv4/IPv6 Coexistence and Transition", draft-xli-behave-ivi-02 (work in progress), June 2009.

[dIVI] "Test homepage for the dIVI: http://202.38.97.114:8056/test.html".

Authors' Addresses

Xing Li
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 10-62785983
Email: xing@cernet.edu.cn

Congxiao Bao
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 10-62785983
Email: congxiao@cernet.edu.cn

Hong Zhang
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 10-62785983
Email: neilzh@gmail.com