A Try and Error type approach for multihoming
draft-matsuoka-multihoming-try-and-error-00

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with
the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering
Task Force (IETF), its areas, and its working groups. Note that
other groups may also distribute working documents as Internet-
DRAFTs.

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."

The list of current Internet-Drafts can be accessed at
http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at

Abstract

[RFC5220] describes the possible problems which an end host may
experience, if the end host has multiple prefixes in a single
physical link. This document proposes a solution of so-called "try
and error" type about these problems originated in "Source Address
Selection" which is described in [RFC5220]. A new mechanism to
settle almost all of these problems is described in this document,
but actually it is not effective in some particular cases. Thus it is
necessary for every end user/host to be able to select on/off of this
mechanism.
1. Introduction

When an end host has multiple prefixes/addresses in a physical link, it is required for end node to use a specific source prefix/address which is allocated by the egress ISP network. Some types of problems in this situation are pointed out in [RFC5220]. Some kind of improvements in consumer premises equipment (CPE) router and/or an end host must be needed to fix these problems. [RFC5221] describes requirements for the mechanisms of solutions about these problems, and T.Fujisaki [draft-fujisaki-dhc-addr-select-opt-06] proposes a mechanism which ISP networks provide information to select a collect source prefix/address by using DHCPv6 option message.

Actually, a perfect solution which can fix all problems originated in "Source Address Selection" may be needed in the future. On the other hand, a "Light type" solution which cannot fix a few, but can fix almost all problems is very useful. M.Bagnulo proposed "Try and Error" type approach [draft-ietf-6man-addr-select-sol-01], which end host re-tried to communicate with another source address if the end host detected communication failure that originates in source address miss-selection. It is important that in general most end host must not have so many prefix/addresses, but 2-3 or less. Based on this assumption, an end host will find a correct source prefix/address within two times attempt.
This document proposes a new "Try and Error" type solution. This document does not include solutions to fix every kind of problems originated in "Source Address Selection". The problems and precondition that this document mentions are described in Section 2.

ICMPv6 [RFC2463] can transmit the information of network state, and many documents define ICMPv6 message, and expected behavior of an end host which detects an icmpv6 message, respectively. Both the description of proposed solution and the arrangement of relationship in these documents are in Section 3. And the importance that an end host which receives a specific ICMPv6 message retries the communication immediately is also described in this section. Requirements for end hosts and network are in section 4, and the example of implementation is in section 5. The verification from the perspective of [RFC5221] is in section 6.

2. Problems to be solved and preconditions

This document mainly describes the solution about the problem originated in "Source Address Selection" which is described in [RFC5220] section 2.1. On the other hand, this document doesn’t describe the problems about routing, for example, some routers on a link don’t cooperate each other, or a CPE router which has multiple ISP uplinks uses them in round robin manner, etc. This document is described on the assumption that correct routing information is configured manually or automatically in their routers, and end hosts will send packets to their default gateway router if an end host doesn’t have specific routing information.

3. Proposed mechanism

3.1. End host categorize source IP addresses as high/low priority

If "Try & Error" type approach is adopted, an end host has no need to recognize source IP address that can communicate to destination IP address before the communication starts. It is sufficient for an end host to select a suitable source IP address which based on a rule like [RFC3484] or others before the communication starts. When an end host recognizes communication failure because of source IP address miss-selection, an end host MUST memorize the source IP address that should not be used for communicating to the destination IP address "prohibited source address". An end host MUST reconnect to the destination IP address with using the most suitable source IP address except prohibited source addresses.

ICMPv6 error message is suitable to transmit an end host the source address miss-selection. On the other hand, if an end host receives an incoming packet including TCP Syn+Ack, the end host MUST memorize the
working source and destination address pair which is used in the incoming packet, and the end host MUST use the pair of addresses for an outgoing packet by priority.

3.2. End host detect invalid source IP address using ICMPv6 error message

[RFC2463]-Section3.1 defines ICMPv6 destination unreachable error messages "Type=1 Code=0-4" which are informed from a network node to an end host, and [RFC4443] defines informative subsets of "Type=1 Code=1" as "Type=1 Code=5,6", respectively. But there is no requirement about behavior for end node which receive these ICMPv6 error messages.

[RFC1122] defines ICMPv4 destination unreachable message "Type=3 Code=2-4" as Hard Error condition, and an end host which receives these ICMPv4 error messages must terminate the corresponding communication. F.Gont [RFC5461] describes ICMPv6 destination unreachable message "Type=1 Code=0,3" as Soft Error condition in comparison with the case in ICMPv4. According to these definition, ICMPv6 error messages "Type=1 Code=1,2,4,5,6" are classified into Hard Error condition, and an end host which receives these ICMPv6 error messages MUST terminate the corresponding communication.

[RFC4443] defines ICMPv6 error message "Type=1 Code=5" as follows. If the reason for the failure to deliver is that the packet with this source address is not allowed due to ingress or egress filtering policies, the Code field is set to 5. An end host which receives this type of error message MUST terminate the corresponding communication, because communication failure has occurred in source IP address miss-selection.

4.Requirements

4.1.End host Requirements

a. Every end host MUST have "working address pair list" which described in Section 5. This list MUST be given priority than the mechanism of source IP address selection defined in [RFC3484].

a-1. An end host which received an incoming packet including TCP Syn+Ack SHOULD memorize the source and destination IP addresses of an incoming packet as the destination and source IP addresses of "the working address pair list".

b. Every end host MUST have "prohibited address pair list" which described in Section 5.2. This list MUST be given priority than both of "working address pair list" and the source IP address selection mechanism which is defined in [RFC3484].
b-1. An end host which received an ICMPv6 "Type=1 Code=5" packet MUST obtain the source and destination IP addresses which are indicated in the invoked packet information and MUST memorize them into "the prohibited address pair list".

c. The end host which received ICMPv6 "Type=1 Code=5" MUST terminate the corresponding communication which was indicated in the invoked packet information, immediately. After memorizing the new entry into "the prohibited address pair list", the end host SHOULD retry to communicate to the same destination IP address. The most important thing is that "the prohibited address pair list" is given higher priority than both of "working address pair list" and the source IP address selection mechanism which is defined in [RFC3484]. As a result, the source IP address which has caused source address miss-selection will not be used.

c-1. In the case using TCP protocol communication, kernel in an end host MUST edit "the prohibited address pair list" before retrying the communication to the same destination IP address.

c-2. In the case using UDP protocol communication, kernel in an end host MUST edit "the prohibited address pair list" and MUST terminate the corresponding communication and inform error message to application. Afterwards, this host will try to communicate to the same destination IP address with another source IP address.

4.2. ISP network requirements

a. An ISP network which is connected to end users directly MUST reply ICMPv6 "Type=1 Code=5" to the end host that tries to communicate with different source IP prefix that the ISP allocated to the end user.

b. An ISP network which does not connect with end users directly MUST NOT reply ICMPv6 "Type=1 Code=5" to the end host. An ISP network which connects with end users directly SHOULD NOT transmit this kind of messages to end user hosts.

5. "prohibited address pair list" and "working address pair list"

5.1. Overview

Every end node MUST have "prohibited address pair list" and "working address pair list" respectively. Some system MAY have these list separately, or MAY express them in one list. The most important thing is that the prohibited address pair list is given priority than the working address pair list, and it is possible to adjust to the composition change in the network, when a working address pair is
5.2. Prohibited address pair list

In this list, every record about a specific destination IP address can have multiple source IP addresses. Records generates automatically when the system accept ICMPv6 "Type=1 Code=5" error message, and records SHOULD be configured also manually. In the above figure, the list describes as below.

+-----------------+----------------------------------------------+
|Destination      | Source IP Addresses                         |
|-----------------+----------------------------------------------+
|2001:db8:3:3::80 | 2001:db8:7000:1::abcd 2001:db8:9000:1::abcd |
+-----------------+----------------------------------------------+

These parameters in this list will change according to the network composition, so that it is necessary to keep appropriate TTL. It is preferable that TTL is less than 24 hours, and more than 60 seconds.
5.3. working address pair list

In this list, every record about a specific destination IP address can have single source IP address. Records generates automatically when the system accept incoming packets including Syn+Ack packet, and records SHOULD be configured also manually. In the above figure, the list describes as below.

<table>
<thead>
<tr>
<th>Destination</th>
<th>Source IP Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001:db8:1:1::80</td>
<td>2001:db8:9000:1::abcd</td>
</tr>
<tr>
<td>2001:db8:2:2::80</td>
<td>2001:db8:7000:1::abcd</td>
</tr>
<tr>
<td>2001:db8:3:3::80</td>
<td>2001:db8:8000:1::abcd</td>
</tr>
</tbody>
</table>

[Table 2]

The following record MUST be deleted or ignored if exists, because "prohibited address pair list" must be given priority than "working address pair list"

<table>
<thead>
<tr>
<th>Destination</th>
<th>Source IP Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001:db8:1:1::80</td>
<td>2001:db8:8000:1::abcd</td>
</tr>
</tbody>
</table>

[Table 3]

These records must keep appropriate TTL as the same with the case of prohibited address pair list.

6. Security Considerations

This document requires end hosts to abort a connection when it receives an ICMPv6 error message "Type=1, Code=5". Therefore, an attacker wishing to reset an ongoing valid connection can send a malicious ICMPv6 error message. In this point of view, ISPs which directly connects with end user should stop any kind of malicious ICMPv6 error message.
7. References


Authors’ Addresses

Hirotaka Matsuoka
NTT West
3-15 Bamba-cho Chuo-ku, Osaka 540-8511
Japan
phone: +81 6 4793 3921
Email: matsuoka.h@west.ntt.co.jp

Intellectual Property Statement

The IETF Trust takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in any IETF Document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights.

Copies of Intellectual Property disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

The definitive version of an IETF Document is that published by, or under the auspices of, the IETF. Versions of IETF Documents that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of IETF Documents. The definitive version of these Legal Provisions is that published by, or under the auspices of, the IETF. Versions of these Legal Provisions that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions.

For the avoidance of doubt, each Contributor to the IETF Standards Process licenses each Contribution that he or she makes as part of the IETF Standards Process to the IETF Trust pursuant to the provisions of RFC 5378. No language to the contrary, or terms, conditions or rights that differ from or are inconsistent with the rights and licenses granted under RFC 5378, shall have any effect and
shall be null and void, whether published or posted by such Contributer, or included with or in such Contribution.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (c) 2009 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 in effect on the date of publication of this document (http://trustee.ietf.org/license-info). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.