Network Working Group A. Kaiser
Internet-Draft A. Petrescu
Intended status: Experimental CEA
Expires: December 20, 2013 June 18, 2013

Interface Selection Mechanism for Multiple Interfaces IPv6 Hosts
draft-kaiser-if-sel-00

Abstract

This document describes an interface selection mechanism that enables multiple interfaces (multihomed) IPv6 hosts to select their most appropriate egress interface to send data over the network. The mechanism extends the Neighbor Discovery (ND) protocol [RFC4861] with two new Router Advertisement options.

Table of Contents

1. Introduction

In the context of multihomed hosts, where hosts are connected to a network with more than only one interface, the selection of a right egress interface to send data is critical. Indeed, selecting a wrong egress interface may lead to a non-optimal routing of data in the network or, in the worst case, the impossibility to reach the destination. In order to cope with the aforementioned issue, this document describes an interface selection mechanism that enables hosts to select their most appropriate egress interface, leading to an optimized end-to-end routing of data.

The proposed mechanism is based on the ND protocol. More precisely, two new options are introduced in RS and RA messages: the Link Cost Option (LCO) and the Path Cost Option (PCO).

2. Requirements

The keywords 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 [KEYWORDS].

3. Description of the mechanism

The proposed mechanism is divided into two operations, each one relying on a new ND option:

- o Gathering informations about links costs and selection of a default egress interface using the Link Cost Option

- o Gathering informations about the cost of a path to a destination and selection of an egress interface for the specific destination using the Path Cost Option (PCO)

The following subsections detail each operation as well as the corresponding option.

3.1. Link Cost Option

The LCO is used to advertise the cost of a link. This option MUST be included in all RA messages that are generated by routers. Therefore, a host that connects to a link is informed about the cost of this specific link. By extension, if a host is connected to multiple links in the network, it is informed about the cost of each of those links. Using these informations, the host selects its default interface: the interface connected to the link that has the lowest cost is selected. The host then adds in its routing table an entry that contains the default route as destination, the router that is on the corresponding link as next-hop, the selected interface as outgoing interface and the link cost as metric. Upon reception of other RA messages, the host compares the link costs included in the RAs with the one in its routing table. If the received link cost is better than the one of the default route in the routing table, the latter MUST be updated accordingly.

The following figure shows the format of the link cost option. This option is only valid in the RA messages and MUST NOT be included in the other ND messages.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |    Length     |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Link Cost                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: Identifier of the LCO (TBD by IANA).

Length: 1. Size of the option as defined in [RFC4861].

Reserved: Unused field. MUST be set to zero by sender and ignored by recipient.

Link Cost: Unsigned integer. The cost of the link.

3.2. Path Cost Option

The PCO is used by hosts to ask routers about the total cost of a path to a destination. The routers reply to the request also by using the PCO. Thus, this option MAY be included in RS and RA messages.

When a multiple interfaces host wants to send data to a destination node, it starts by checking in its routing table if a specific route to this destination exists. If no route exists, the host sends a RS message with a PCO to the all-routers multicast address on each of its interfaces. The PCO includes the IPv6 address of the destination to which the host wants to send data. Upon reception of such RS message, a router checks in its routing table which route entry it would select to forward data to this specific destination. The router then replies by sending a unicast RA message to the requesting host with a PCO that includes the total cost of the selected path from the router itself to the destination along with a lifetime that defines the validity of the route advertised. Upon reception of these informations, the host computes the total cost of the path from itself to the destination by adding to the path cost advertised by the router the corresponding link cost (which is advertised with the LCO). Once computed, the host selects as egress interface to the specific destination the interface connected to the path that has the lowest end-to-end cost. The host then updates its routing table accordingly with the new computed information.

The following figure shows the format of the path cost option. This option is only valid in the RS or RA messages and MUST NOT be included in the other ND messages.

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |    Length     |Transaction ID |  Status Code  |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                   Reserved                    | Prefix Length |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    |                          Destination                          |
    |                       Address or Prefix                       |
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                           Path Cost                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                           Lifetime                            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: Identifier of the PCO (TBD by IANA).

Length: 4. Size of the option as defined in [RFC4861].

Transaction ID: Identifier of the current RS/RA messages exchange between the host and the router.

Status Code: Code that provides additionnal informations about the results provided by the router (see Section 4 for more details). MUST be set to zero in RS messages and ignored by recipient.

Reserved: Unused field. MUST be set to zero by sender and ignored by recipient.

Prefix Length: Size of the IPv6 prefix that follows. Set to 128 if the following field contains an IPv6 address.

Dst. addr. or pref.: IPv6 prefix or address of the destination node to which the path computation is asked.

Path Cost: Unsigned integer. The total cost of the path from the advertising router to the destination.

Lifetime: The lifetime of the route advertised in the RA message. MUST be set to zero in RS messages and ignored by recipient.

4. Status Code

A Status Code is included in the PCO option. It provides additionnal informations to the host about the results of its request. The following codes are considered:

0 Success: a path to the destination is known by the router and the corresponding path cost is included in the PCO.

1 No specific route to the destination is known by the router. However, the router has a default route that can be used to forward data to the destination, but with no warranty. In this case, the corresponding path cost MUST be set to the maximum possible value (i.e. all-ones bits).

2 Failure: the packets to this destination would probably be dropped by this router because no route to reach the destination is known or because of any other reason (firewall rules, policy-related rules, ...). In this case, the path cost value is set to zero.

5. Example use case

Typical uses cases that can be considered are home networks, corporate network, ad-hoc networks, etc. This section illustrates the mechanism through a simple network topology.

5.1. Network topology

The following figure depicts the network topology used in this example.

```
                                         Link1 (4)
                                      2001:db8:1::/64
      Link2 (2)      --+-----------------------------------------+--
    2001:db8:2::/64     |                                         |
    --+--------+--      |                                         |
       |       |        |                                         |
       |       |   /----+----\                               /----+----\
       |       |   |         |                               |         |
   +---+---+   +---+ Router1 |                               | Router2 |
   |       |       |         |                               |         |
   | Host2 |       \----+----/                               \----+----/
   |       |            |                                         |
   +-------+            | Link3 (5)                     Link4 (3) |
                        | 2001:db8:3::/64         2001:db8:4::/64 |
                      --+----------+--                 --+--------+--
                                   |                     |
                                   |      +-------+      |
                                   |   I1 |       | I2   |
                                   +------+ Host1 +------+
                                          |       |
                                          +-------+
```

The figure shows two routers (Router1, Router2) and two hosts (Host1, Host2). Theses nodes are connected to each other through 4 links (Link1, Link2, Link3, Link4). The values in brackets represent the corresponding link costs. Also, Host1 is a multiple interfaces device: it is connected to Link3 via its network interface I1 and to link4 via its network interface I2.

Let us consider Host1. Its first operation consists of gathering links costs and select a default interface. To this end, Router1 and Router2 send usual periodic RA messages. These messages include a LCO that describes the cost of the link: Router1 advertises that the cost of Link3 is 5 and Router2 advertises that the cost of link4 is 3. As Link4 has a better cost that Link3, Host1 selects I2 as its default interface. Host1 then updates its routing table accordingly, as show in the following figure:

Destination Network	Next-Hop Address	Output Interface	Cost
fe80::/64	-	I1	5
fe80::/64	-	I2	3
2001:db8:3::/64	-	I1	5
2001:db8:4::/64	-	I2	3
Default	@Router2	I2	3

Let us now consider that Host1 wants to communicate with Host2. In a classical scenario, Host1 would send data for Host2 through I2, according to its routing table. Data would thus be forwarded by Router2 and Router1 through Link1 and Link2 respectively, leading to a total path cost of 9. Sending data through I1 would have been a better choice. Indeed, despite the fact that Link3 has a worse cost compared to Link4, the end-to-end path cost would be better (7). However, as Host1 is not a router, it does not have a sufficient vision of the network to make such decision. To this end, before sending data to Host2, Host1 first send a RS message that includes a PCO on all its outgoing interfaces (I1 and I2). The "Destination Address" field is filled with tha address of Host2. Upon reception of such message, the routers reply to Host1 with a PCO included in RA message: Router1 replies that its better known path to reach Host2 has a total cost of 2 (Link2) and Router2 replies that its better known path has a total cost of 6 (Link1 + Link2). As Host1 already knows the costs of Link3 and Link4, it computes that sending data through I1 would have an end-to-end cost of 7 (Link3 + Link2) whereas using I2 would lead to an end-to-end cost of 9 (Link4 + Link1 + Link2). Hence, Host1 selects I1 as its egress interface to reach Host2 and updates accordingly its routing table, as shown in the following figure:

Destination Network	Next-Hop Address	Output Interface	Cost
fe80::/64	-	I1	5
fe80::/64	-	I2	3
2001:db8:3::/64	-	I1	5
2001:db8:4::/64	-	I2	3
Default	@Router2	I2	3
2001:db8:2::/64	@Router1	I1	7

5.2. Messages exchange diagram

The following diagram shows the messages exchanges corresponding to the example described above: the first two RA messages correspond to Host1 default interface selection, the following two RS/RA messages exchanges correspond to the selection of Host1 interface to reach Host2 and the lasts messages show the final path used by data to transit from Host1 to Host2 and vice-versa.

```
   +-------+                      +---------+   +---------+     +-------+
   | Host1 |                      | Router1 |   | Router2 |     | Host2 |
   +-------+                      +---------+   +---------+     +-------+
       |                             |             |              |
       |    +----+--------------+    |             |              |
       |    | RA | LCO cost = 5 |    |             |              |
       |    +----+--------------+    |             |              |
    I1 o<----------------------------o             |              |
       |                             |             |              |
       |           +----+--------------+           |              |
       |           | RA | LCO cost = 3 |           |              |
       |           +----+--------------+           |              |
    I2 o<------------------------------------------o              |
       |                             |             |              |
       |  +----+------------------+  |             |              |
       |  | RS | PCO dst = @Host2 |  |             |              |
       |  +----+------------------+  |             |              |
    I1 o---------------------------->o             |              |
       |                             |             |              |
       |         +----+------------------+         |              |
       |         | RS | PCO dst = @Host2 |         |              |
       |         +----+------------------+         |              |
    I2 o------------------------------------------>o              |
       |                             |             |              |
       |    +----+--------------+    |             |              |
       |    | RA | PCO cost = 2 |    |             |              |
       |    +----+--------------+    |             |              |
    I1 o<----------------------------o             |              |
       |                             |             |              |
       |          +----+--------------+            |              |
       |          | RA | PCO cost = 6 |            |              |
       |          +----+--------------+            |              |
    I2 o<------------------------------------------o              |
       |                             |             |              |
       |       +----+----------+     |     +---------------+      |
       |       | DATA TO HOST2 |     |     | DATA TO HOST2 |      |
       |       +----+----------+     |     +---------------+      |
    I1 o---------------------------->o--------------------------->o
       |                             |             |              |
       |       +----+----------+     |     +---------------+      |
       |       | DATA TO HOST1 |     |     | DATA TO HOST1 |      |
       |       +----+----------+     |     +---------------+      |
    I1 o<----------------------------o<---------------------------o
       |                             |             |              |
       |                             |             |              |
```

6. Security Considerations

To be done.

7. IANA Considerations

IANA is kindly requested by the authors to allocate the following values:

- o The Link Cost Option type, which should be added to the Neighbor Discovery option type space defined in section 13 of [RFC4861]

- o The Path Cost Option type, which should be added to the Neighbor Discovery option type space defined in section 13 of [RFC4861]

8. References

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

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

Authors' Addresses

Arnaud Kaiser
Commissariat a l'Energie Atomique
8 Avenue de la Vauve
Palaiseau, Ile-de-France 91120
FR

Email: arnaud.kaiser@cea.fr

Alexandru Petrescu
Commissariat a l'Energie Atomique
8 Avenue de la Vauve
Palaiseau, Ile-de-France 91120
FR

Email: alexandru.petrescu@cea.fr