IPv6 Operations Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 5, 2007

A. Matsumoto
T. Fujisaki
NTT
R. Hiromi
K. Kanayama
Intec Netcore

Nov 2006

Requirements for distributing RFC3484 address selection policy draft-ietf-v6ops-addr-select-req-00.txt

Status of this Memo

Copyright Notice

Abstract

RFC3484 defines source and destination address selection algorithms that are commonly deployed in current popular OSs. Meanwhile, there is a possibility to provide multiple addresses in one physical network. In such a multi-prefix environment, end-hosts could encounter some troubles in the communication because of default use

of the RFC3484 mechanism.

Therefore, extending various rules beyond the default use of the RFC3484 mechanism should be considered. We propose a concept of distribution of address selection policy to an end-host as a solution to these possible problems.

In this document, we describe detailed requirements of address selection policy distribution.

Table of Contents

1. Introduction

One physical network can have multiple logical networks. In that case, an end-host has multiple IP addresses. In the IPv4-IPv6 dual stack environment or in a site connected to both ULA [RFC4193] and global scope networks, an end-host has multiple IP addresses. These are examples of the networks that we focus on in this document. In such an environment, an end-host will encounter some communication trouble documented in PS. [I-D.ietf-v6ops-addr-select-ps]

RFC 3484 [RFC3484] defines both source and destination address selection algorithms. RFC 3484 defines a default address table, and enables adding other entries to this table. Flexible address selection can be carried out.

In addition, the distribution of an address policy table is an important matter. RFC 3484 describes all the algorithms for setting the address policy table, but it makes no mention of autoconfiguration.

To make a smooth connection with the appropriate source and destination address selection inside a multi-prefix environment, nodes must be informed about routing policies of their upstream networks and possible source address selection policies. Then, those nodes must put those policies into individual policy tables.

On the other hand, the RFC3484 mechanism is commonly deployed. However, manual configuration of the policy table is not a feasible idea and some automatic mechanism is needed.

Therefore, we propose a concept of distribution of address selection policy from a network to an end-host to cooperate with the RFC3484 mechanism as a solution to these possible problems.

In this document, requirements for distribution of the address selection policy are described for promotional use of the RFC3484 mechanism. Our goal is to carry our autoconfiguration with distribution mechanism for utilization of RFC3484 more effectively.

1.1. Scope of this document

Revising address selection rules defined in RFC 3484 is out of our scope.

The routing information from an upstream network is necessary, but in this document, we are focused on how to select source and destination addresses at the RFC3484 address policy table of the end-host.

In addition, there must be some practical ways or considerations other than the RFC3484 policy table to solve the address selection problem, such as utilization of some routing protocols or operational technique with a specific route but these discussions are out of our scope. However, we select some examples of other mechanisms in Section 5 only for comparison.

2. Policy distribution model and terminology

The distribution model:

Fig. 1: (basic model)
Policies transferred from the Policy Broker to the Node over an access network.

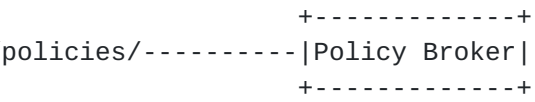
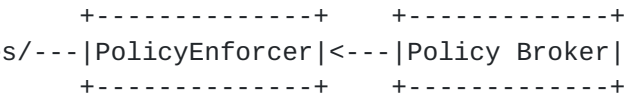


Fig. 2: (extended model)



Essentials (or Principles):

The distribution of Policy, which means that the address selection policy of RFC3484 is sent to nodes, has the following functions.

* Policy Broker, which means a Policy originator in xSP, for example, a dhcp server, originates its policy and sends it to a Node using some prefix-assignment Protocols.

* A Node receives a Policy as a client. Then the client puts those policies into its own Policy Table, which means the address selection table defined in RFC3484.

* Policy Broker should make the Policy so that it can be easily embedded into Nodes. The Policy message format should be defined based on the algorithm specified in RFC3484.

* There might be a situation in which a Policy Broker and Node are disconnected, and no direct message is exchanged. In this case, there is a middle box defined as a Policy Enforcer, for example, CPE illustrated in Fig. 2, and it relays the policy to the Node. Requirements should be considered to include the policy-relay case.

Terminology:

Node:	end-host, end-terminal
CPE:	Customer premises equipment
PE:	Provider Edge device
NAS:	Network Access Server

Policy:	address selection policy for RFC3484 rule set
Policy Enforcer:	optionally attached equipment to relay policy
Policy Broker/Server:	Policy Originator in xSP(mandate)
Prefix Delegation Protocol:	Protocols to carry prefix information data
address selection table:	address selection table defined in RFC3484
Policy Table:	address selection table defined in RFC3484

3. Requirements of Policy distribution

The purpose of the Policy distribution mechanism is to distribute a Policy Table to Nodes and configure the Policy Table on the Nodes automatically. The use of a distributed Policy Table on Nodes for other purposes (e.g, configuring routing Table on the Nodes) is out of the scope of this document.

3.1. Contents of Policy Table

A Policy Table is a set of Policies described in RFC 3484. Each Policy consists of four elements: prefix value, precedence value, label value, and zone index value. The Policy distribution mechanism should be able to distribute a Policy Table that has one or more Policies to Nodes.

3.2. Timing

The Policy Table should be distributed to Nodes by a Policy broker at any time when Nodes send a request for the Policy.

3.3. Redistribution of changed Policy Table

When a Policy broker has any change in a Policy Table that is distributed to Nodes, the Policy broker should redistribute the latest Policy Table to Nodes.

3.4. Sections

The Policy distribution mechanism should support being performed in two kinds of sections: from PE to CPE and from CPE to Node. Policy

distribution mechanisms provided in each section may or many not be the same.

3.5. Generating Policy Table per CPE/Node

The Policy distribution mechanism should allow for generating an appropriate Policy Table per Node. For example, in some cases, each Node may have a different set of assigned prefixes. In such a case, the appropriate Policy Table for each Node may also be different, and a Policy broker may be needed to generate the Policy Table according to the identity of the Node.

3.6. Security

The Policy distribution mechanism should provide for reliable, secure distribution of the Policy distribution from a Policy broker to Nodes.

4. Solutions for RFC3484 policy distribution

As described in section 4.1, the address selection policy table consists of four elements: prefix value, precedence, label, and zone-index. The policy distribution mechanism will deliver lists of these elements.

4.1. Policy distribution with router advertisement (RA) message option

The RA message can be used to deliver a policy table by adding a new ND option. Existing ND transport mechanisms (i.e., advertisements and solicitations) are used. Advantages and disadvantages are almost the same as those described in [DNS configuration RFC, RA section].

In addition, an advantage and disadvantages of distributing a policy table are as follows.

Advantages:

- The RA message is used to deliver IPv6 address prefixes. Therefore, delivering policies for selecting addresses with the address attached to the host would be natural.

Disadvantages:

- The RA message is limited in size, and the RA may not be sufficient to deliver full policies. The same compression techniques, which were adopted in RFC4191 [RFC4191] can be used to increase the number of policies delivered by RA messages.

- Currently, RA messages are not used between a PE and CPE. Other protocols may be necessary to deliver a policy table.

- Configuring a policy table in each router that advertises RA messages with an address prefix is necessary, so if a site has a lot of routers, there will be a higher management cost.

- Delivering a specific policy table to one node is impossible because RA messages are multicast.

4.2. Policy distribution in DHCPv6

By defining a new DHCPv6 option like [I-D.fujisaki-dhc-addr-select-opt], a policy table can be delivered. The advantages and disadvantages are almost the same as those described in [DNS configuration RFC, DHCPv6 section].

In addition, there are the following advantages and disadvantages.

Advantages:
- Currently, DHCPv6 prefix delegation is mainly used between a PE and CPE. Delivering a policy table with prefixes is possible.

- A DHCPv6 server can deliver a host-specific policy table.

- By using a DHCPv6 relay mechanism, managing a policy table from a central server is possible.

Disadvantages:
- The DHCPv6 message size is limited to the maximum UDP transmission size, so delivering complex policies by DHCPv6 may be impossible.

4.3. Using other protocols

Using other protocols (i.e., http and ftp) to deliver the policy table is possible.

Advantages:
- No new transport mechanisms are necessary.

Disadvantages:
- Other service discovery mechanisms will be necessary.

- The procedure to distribute information should be defined (e.g., when to distribute and where the information is stored).

- Existing protocols may not have a mechanism to inform clients about policy changes.

4.4. Defining a new protocol

Defining a new protocol to deliver a policy table will have the following advantages and disadvantages.

Advantages:
- Defining a protocol suitable for policy distribution may be possible.

Disadvantages:
- In addition to the disadvantages of 4.3, a new transport mechanism needs to be defined.

4.5. Converting routing information to policy table

In an environment in which routing information and network links are separated (e.g., between PE and CPE), converting routing information to a policy table is possible. However, when intermediate routers and nodes receive next-hop information, that is aggregated as a default route or neighbor router, and cannot generate policy table [a policy table cannot be generated].

Advantages:
- No new distribution mechanism is necessary.

Disadvantages:
- This mechanism can be used only in a limited environment.

5. Discussion

Other than this policy distribution mechanism, a few mechanisms are proposed. This section quickly reviews each proposal including a policy distribution mechanism.

5.1. Routing System Assistance for Address Selection by Fred Baker

Fred Baker proposed to us about this mechanism. A host asks the DMZ routers or the local router which is the best pair of source and destination addresses when the host has a set of addresses A and destination host has a set of addresses B. And then, the host uses the policy provided by the server/routing system as a guide in

applying the response. He also proposed a mechanism that utilizes ICMP error message to change the source address of the existing session. This point resembles 5.2 3484-update mechanism, so the following evaluation is based only on the first part of his proposal.

Advantages:

- A host can choose the best address pair that reflects the dynamic changing routing status.

- The destination address selection can be handled in this mechanisim as well as source address selection.

Disadvantages:

- A host can choose the best address pair that reflects the dynamic

- A host has to consult the routing system every time it starts a connection if the host doesn't have address selection information for the destination host or the information lifetime is expired. This could be a possible scalability problem.

- A host has to wait until the response is received from the routing system.

- The existing host/router OS implementation has to be changed a lot. In the existing TCP/IP protocol stack implementation, destination address selection is mainly the role of the application and not that of the kernel unlike source address selection. Therefore, implementing this model without causing any affects on applications is not so easy.

5.2. 3484-update

M. Bagnulo proposed a new method of address selection in his draft. [I-D.bagnulo-rfc3484-update] When the host notices that a network failure occurs or packets are dropped somewhere in the network by for example, an ingress filter, the host changes the source address of the connection to another source address. The host stores a cache of address selection information so that the host can select an appropriate source address for new connections.

Advantages:

- A host can choose the best address that reflects the dynamic changing routing status.

Disadvantages:

- A host has to learn address selection information per destination host. The number of cache entries can be too big.

- The existing host/router OS implementation has to be changed a lot. In particular, changing the source address of the existing connection is not so easy and has a big impact on the existing TCP/IP protocol stack implementation.

- There is not so much experience with this kind of address selection cache mechanism.

- The host tries every address one-by-one, so the user has to wait for a long time before the appropriate address pair is found.

5.3. shim6

shim6 is designed for site-multihoming. This mechanism introduces a new method of address selection for session initiation and session survivability, which is documented in [I-D.ietf-shim6-locator-pair-selection] and [I-D.ietf-shim6-failure-detection].

The shim6 host detects connection failures and changes the source address during the session.

Advantages:

- The shim6 host performs address selection that reflects network failures in the source and destination end-to-end link. Moreover, network failure avoidance can be achieved by end hosts themselves.

Disadvantages:

- A host has to learn address selection information per destination host. The number of cache entry can be too big.

- The existing host/router OS implementation has to be changed significantly.

- The host tries every address one-by-one, so the user has to wait for a long time before the appropriate address pair is found.

5.4. policy distribution mechanism

This mechanism takes advantages of RFC 3484 Policy Table that is widely deployed already. By distributing policies for Policy Table, you can auto-configure a host's address selection policy.

Advantages:

- A host can receive and understand address selection information before the host starts a connection. Therefore, the amount of traffic and connection overhead time can be minimized.
- A host does not need any other address-selection-related information once that host receives the address selection policy set. This can also reduce the amount of traffic.
- The existing OS implementation does not need to be changed significantly on the OS that implements the RFC 3484 policy table. Only the delivery mechanism to the table has to be prepared.
- Destination address selection can also be controlled by this mechanism.

Disadvantages:

- No other address selection rule that is beyond the RFC 3484 policy table framework can be implemented.
- The OS implementation has to be changed, and the policy distribution server, such as a gateway router, has to be prepared.
- When DHCP or RA is used for transport mechanism of policy table, frequently changing policy cannot be delivered to hosts quickly because of the nature of these protocols.

6. Security Considerations

Address false-selection can lead to serious security problem, such as session hijack. However, it should be noted that address selection is eventually up to end-hosts. We have no means to enforce one specific address selection policy to every end-host. So, a network administrator has to take countermeasures for unexpected address selection.

7. IANA Considerations

This document has no actions for IANA.

8. References

8.1. Normative References

[I-D.ietf-v6ops-addr-select-ps]
Matsumoto, A., "Problem Statement of Default Address Selection in Multi-prefix Environment: Operational Issues of RFC3484 Default Rules", draft-ietf-v6ops-addr-select-ps-00 (work in progress), November 2006.

[RFC3484] Draves, R., "Default Address Selection for Internet Protocol version 6 (IPv6)", RFC 3484, February 2003.

8.2. Informative References

[I-D.arifumi-ipv6-policy-dist]
Matsumoto, A., "Practical Usages of Address Selection Policy Distribution", draft-arifumi-ipv6-policy-dist-01 (work in progress), June 2006.

[I-D.bagnulo-rfc3484-update]
Bagnulo, M., "Updating RFC 3484 for multihoming support", draft-bagnulo-rfc3484-update-00 (work in progress), June 2006.

[I-D.fujisaki-dhc-addr-select-opt]
Fujisaki, T., "Distributing Default Address Selection Policy using DHCPv6", draft-fujisaki-dhc-addr-select-opt-02 (work in progress), June 2006.

[I-D.ietf-shim6-failure-detection]
Arkko, J. and I. Beijnum, "Failure Detection and Locator Pair Exploration Protocol for IPv6 Multihoming", draft-ietf-shim6-failure-detection-06 (work in progress), September 2006.

[I-D.ietf-shim6-locator-pair-selection]
Bagnulo, M., "Default Locator-pair selection algorithm for the SHIM6 protocol", draft-ietf-shim6-locator-pair-selection-01 (work in progress), October 2006.

[RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and More-Specific Routes", RFC 4191, November 2005.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.

Authors' Addresses

Arifumi Matsumoto
NTT PF Lab
Midori-Cho 3-9-11
Musashino-shi, Tokyo 180-8585
Japan

Phone: +81 422 59 3334
Email: arifumi@nttv6.net

Tomohiro Fujisaki
NTT PF Lab
Midori-Cho 3-9-11
Musashino-shi, Tokyo 180-8585
Japan

Phone: +81 422 59 7351
Email: fujisaki@syce.net

Ruri Hiromi
Intec Netcore, Inc.
Shinsuna 1-3-3
Koto-ku, Tokyo 136-0075
Japan

Phone: +81 3 5665 5069
Email: hiromi@inetcore.com

Ken-ichi Kanayama
Intec Netcore, Inc.
Shinsuna 1-3-3
Koto-ku, Tokyo 136-0075
Japan

Phone: +81 3 5665 5069
Email: kanayama@inetcore.com

Full Copyright Statement

Intellectual Property

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).