

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 15, 2011

A. Matsumoto
J. Kato
T. Fujisaki
NTT
March 14, 2011

Update to [RFC 3484](#) Default Address Selection for IPv6
draft-ietf-6man-rfc3484-revise-02.txt

Abstract

[RFC 3484](#) describes algorithms for source address selection and for destination address selection. The algorithms specify default behavior for all Internet Protocol version 6 (IPv6) implementations. This document specifies a set of updates that modify the algorithms and provide fixes for the identified issues.

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 15, 2011.

Copyright Notice

Copyright (c) 2011 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

Internet-Draft

[RFC3484](#) Bis

March 2011

described in the Simplified BSD License.

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.

Table of Contents

1.	Introduction	3
2.	Specification	3
2.1.	Changes related to the default policy table	3
2.1.1.	ULAs in the policy table	4
2.1.2.	Teredo in the policy table	4
2.1.3.	Deprecated addresses in the policy table	4
2.1.4.	Renewed default policy table	4
2.2.	The longest matching rule	5
2.3.	Utilize next-hop for source address selection	5
2.4.	Private IPv4 address scope	6
2.5.	Deprecation of site-local unicast address	6
3.	Security Considerations	7
4.	IANA Considerations	7
5.	References	7
5.1.	Normative References	7
5.2.	Informative References	8
Appendix A.	Acknowledgements	8
Appendix B.	Discussion	8
B.1.	Centrally assigned ULA	8
B.2.	6to4, Teredo, and IPv4 prioritization	9
B.3.	Deprecated address	9
B.4.	The longest match rule	9
Appendix C.	Revision History	10
	Authors' Addresses	11

Internet-Draft

[RFC3484](#) Bis

March 2011

1. Introduction

The IPv6 addressing architecture [[RFC4291](#)] allows multiple unicast addresses to be assigned to interfaces. Because of this IPv6 implementations need to handle multiple possible source and destination addresses when initiating communication. [RFC 3484](#) [[RFC3484](#)] specifies the default algorithms, common across all implementations, for selecting source and destination addresses so that it is easier to predict the address selection behavior.

Since [RFC 3484](#) was published, some issues have been identified with the algorithm specified there. The issues are related to the longest match algorithm used in Rule 9 of Destination address selection breaking DNS round-robin techniques, and prioritization of poor IPv6 connectivity using transition mechanisms over native IPv4 connectivity.

There have also been some significant changes to the IPv6 addressing architecture that require changes in the [RFC 3484](#) policy table. Such changes include the deprecation of site-local unicast addresses [[RFC3879](#)] and the IPv4-compatible IPv6 addresses, the introduction of Unique Local Addresses [[RFC4193](#)] etc.

This document specifies a set of updates that modify the algorithms and provide fixes for the identified issues.

2. Specification

2.1. Changes related to the default policy table

The default policy table is defined in [RFC 3484 Section 2.1](#) as follows:

Prefix	Precedence	Label
::1/128	50	0

::/0	40	1
2002::/16	30	2
::/96	20	3
::ffff:0:0/96	10	4

The changes that should be included into the default policy table are those rules that are universally useful and do no harm in every reasonable network environment. The changes we should consider for the default policy table are listed in this sub-section.

The policy table is defined to be configurable. If the local site policy needs to be different changes can be put into the policy table

manually or by using the auto-configuration mechanism proposed as a DHCP option [[I-D.ietf-6man-addr-select-opt](#)].

[2.1.1](#). ULAs in the policy table

[RFC 5220](#) [[RFC5220](#)] [Section 2.1.4](#), 2.2.2, and 2.2.3 describes address selection problems related to ULAs [[RFC4193](#)]. These problems can be solved by either changing the scope of ULAs to site-local, or by adding an entry to the default policy table entry that has its own label for ULAs.

ULAs has been specified with a global scope because the reachability of the ULAs was intended to be restricted by the routing system. Since a ULA will not be exposed outside of its reachability domain, if a ULA is available as a candidate destination address, it can be expected to be reachable. In fact, such ULA to ULA communication is often desired (in particular in sites where ULAs are intended to provide stable addresses when the global prefix may be changing) and thus needs to be prioritized.

Therefore, the scope of ULA should be kept global, and prioritization of ULA to ULA communication should be implemented in the policy table, by assigning a specific label for ULAs using `fc00::/7`.

[2.1.2](#). Teredo in the policy table

Teredo [[RFC4380](#)] is defined and has been assigned `2001::/32`. This address block should be assigned its own label in the policy table. Teredo's priority should be less than or equal to 6to4, considering

its characteristic of being a transitional tunnel mechanism. Windows already implements this.

[2.1.3.](#) Deprecated addresses in the policy table

IPv4-compatible IPv6 addresses are deprecated [[RFC4291](#)]. IPv6 site-local unicast addresses are deprecated [[RFC3879](#)]. Moreover, the 6bone testing address has also been phased out[RFC3701]. The issue is how we treat these outdated addresses.

[2.1.4.](#) Renewed default policy table

After applying these updates, the default policy table becomes:

Prefix	Precedence	Label
::1/128	60	0
fc00::/7	50	1
::/0	40	2
::ffff:0:0/96	30	3
2002::/16	20	4
2001::/32	10	5
::/96	1	10
fec::/16	1	11
3ffe::/16	1	12

[2.2.](#) The longest matching rule

This issue is related to a problem with the longest matching rule, as reported by Dave Thaler. It is a malfunction of the DNS round-robin technique. It is common for both IPv4 and IPv6.

When a destination address DA, DB, and the source address of DA Source(DA) are on the same subnet and Source(DA) == Source(DB), DNS round robin load-balancing cannot function. By considering prefix lengths that are longer than the subnet prefix, this rule establishes preference between addresses that have no substantive differences

between them. The rule functions as an arbitrary tie-breaker between the hosts in a round robin, causing a given host to always prefer a given member of the round robin.

By limiting the calculation of common prefixes to a maximum length equal to the length of the subnet prefix of the source address, rule 9 can continue to favor hosts that are nearby in the network hierarchy without arbitrarily sorting addresses within a given network. This modification could be written as follows:

Rule 9: Use longest matching prefix.

When DA and DB belong to the same address family (both are IPv6 or both are IPv4): If $\text{CommonPrefixLen}(\text{DA} \ \& \ \text{Netmask}(\text{Source}(\text{DA})), \text{Source}(\text{DA})) > \text{CommonPrefixLen}(\text{DB} \ \& \ \text{Netmask}(\text{Source}(\text{DB})), \text{Source}(\text{DB}))$, then prefer DA. Similarly, if $\text{CommonPrefixLen}(\text{DA} \ \& \ \text{Netmask}(\text{Source}(\text{DA})), \text{Source}(\text{DA})) < \text{CommonPrefixLen}(\text{DB} \ \& \ \text{Netmask}(\text{Source}(\text{DB})), \text{Source}(\text{DB}))$, then prefer DB.

[2.3.](#) Utilize next-hop for source address selection

[RFC 3484](#) source address selection rule 5 states that the address that is attached to the outgoing interface should be preferred as the source address. This rule is reasonable considering the prevalence of Ingress Filtering described in [BCP 38](#) [[RFC2827](#)]. This is because an upstream network provider usually assumes it receives those

packets from customers that will use the delegated addresses as their source addresses.

This rule, however, is not effective in an environment such as described in [RFC 5220 Section 2.1.1](#), where a host has multiple upstream routers on the same link and has addresses delegated from each upstream on single interface.

So, a new rule 5.1 that utilizes next-hop information for source address selection is inserted just after the rule 5.

Rule 5.1: Use an address assigned by the selected next-hop.

If SA is assigned by the selected next-hop that will be used to send to D and SB is assigned by a different next-hop, then prefer SA.

Similarly, if SB is assigned by the next-hop that will be used to send to D and SA is assigned by a different next-hop, then prefer SB.

[2.4.](#) Private IPv4 address scope

When a packet goes through a NAT, its source or destination address can get replaced with another address with a different scope. It follows that the result of the source address selection algorithm may be different when the original address is replaced with the NATed address.

The algorithm currently specified in [RFC 3484](#) is based on the assumption that a source address with a small scope cannot reach a destination address with a larger scope. This assumption does not hold if private IPv4 addresses and a NAT are used to reach public IPv4 addresses.

Due to this assumption, in the presence of both NATed private IPv4 address and transitional addresses (like 6to4 and Teredo), the host will choose the transitional IPv6 address to access dual-stack peers [[I-D.denis-v6ops-nat-addrsel](#)]. Choosing transitional IPv6 connectivity over native IPv4 connectivity is not desirable.

This issue can be fixed by changing the address scope of private IPv4 addresses to global. Such a change has already been implemented in some OSes.

[2.5.](#) Deprecation of site-local unicast address

[RFC 3484](#) contains a few "site-local unicast" and "fec::" descriptions. It's better to remove examples related to site-local unicast address, or change examples to use ULAs. Points that need to be re-written are:

- the 2nd paragraph in [RFC 3484 Section 3.1](#) describing the scope comparison mechanism.
- [RFC 3484 Section 10](#) containing examples for site-local address.

[3.](#) Security Considerations

No security risk is found that degrades [RFC 3484](#).

[4.](#) IANA Considerations

An address type number for the policy table may have to be assigned by IANA.

[5.](#) References

[5.1.](#) Normative References

- [RFC1794] Brisco, T., "DNS Support for Load Balancing", [RFC 1794](#), April 1995.
- [RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", [BCP 5](#), [RFC 1918](#), February 1996.
- [RFC3484] Draves, R., "Default Address Selection for Internet Protocol version 6 (IPv6)", [RFC 3484](#), February 2003.
- [RFC3701] Fink, R. and R. Hinden, "6bone (IPv6 Testing Address Allocation) Phaseout", [RFC 3701](#), March 2004.
- [RFC3879] Huitema, C. and B. Carpenter, "Deprecating Site Local Addresses", [RFC 3879](#), September 2004.
- [RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", [RFC 4193](#), October 2005.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", [RFC 4291](#), February 2006.
- [RFC4380] Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", [RFC 4380](#), February 2006.
- [RFC5220] Matsumoto, A., Fujisaki, T., Hiromi, R., and K. Kanayama, "Problem Statement for Default Address Selection in Multi-

Default Rules", [RFC 5220](#), July 2008.

[5.2.](#) Informative References

[I-D.denis-v6ops-nat-addrsel]

Denis-Courmont, R., "Problems with IPv6 source address selection and IPv4 NATs", [draft-denis-v6ops-nat-addrsel-00](#) (work in progress), February 2009.

[I-D.ietf-6man-addr-select-considerations]

Chown, T., "Considerations for IPv6 Address Selection Policy Changes", [draft-ietf-6man-addr-select-considerations-02](#) (work in progress), July 2010.

[I-D.ietf-6man-addr-select-opt]

Matsumoto, A., Fujisaki, T., and J. Kato, "Distributing Address Selection Policy using DHCPv6", [draft-ietf-6man-addr-select-opt-00](#) (work in progress), December 2010.

[I-D.ietf-ipv6-ula-central]

Hinden, R., "Centrally Assigned Unique Local IPv6 Unicast Addresses", [draft-ietf-ipv6-ula-central-02](#) (work in progress), June 2007.

[RFC2827]

Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", [BCP 38](#), [RFC 2827](#), May 2000.

[Appendix A.](#) Acknowledgements

The authors would like to thank to Dave Thaler, Pekka Savola, Remi Denis-Courmont and the members of 6man's address selection design team for their invaluable contributions to this document.

[Appendix B.](#) Discussion

[B.1.](#) Centrally assigned ULA

Discussion: Centrally assigned ULA [[I-D.ietf-ipv6-ula-central](#)] is proposed, and assigned fc00::/8. Using the different labels for fc00::/8 and fd00::/8 makes sense if we can assume the same kind of address block is assigned in the same or adjacent network.

However, the way of assignment and network adjacency may not have any relationships.

[B.2.](#) 6to4, Teredo, and IPv4 prioritization

Discussion: Regarding the prioritization between IPv4 and these transitional mechanisms, their connectivity quality is recently known to be worse than IPv4. These mechanisms are said to be the last resort access to IPv6 resources. The 6to4 should have higher precedence over Teredo, in that 6to4 host to 6to4 host communication runs over IPv4, which can result in a more optimal path, and 6to4 does not need NAT traversal.

[B.3.](#) Deprecated address

Discussion: These addresses were removed from the current specification. So, they should not be treated differently, especially if we think about future re-use of these address blocks.

Considering the inappropriate use of these address blocks, especially in outdated implementations, and bad effects caused by them, however, they should be labeled differently from the legitimate address blocks.

Or should we keep this entry for the sake of backward compatibility?

[B.4.](#) The longest match rule

[RFC 3484](#) defines that the destination address selection rule 9 should be applied to both IPv4 and IPv6, which spoils the DNS based load balancing technique that is widely used in the IPv4 Internet today.

When two or more destination addresses are acquired from one FQDN, rule 9 states that the longest matching destination and source address pair should be chosen. As stated in [RFC 1794](#), the DNS based load balancing technique is achieved by not re-ordering the destination addresses returned from the DNS server. Rule 9 defines a deterministic rule for re-ordering at hosts, hence the technique of [RFC 1794](#) is not available anymore.

Regarding this problem, there was discussion in the IETF and other places that led to some different options being suggested, as listed below.

Discussion: The possible changes to [RFC 3484](#) are as follows:

Internet-Draft

[RFC3484](#) Bis

March 2011

1. To delete Rule 9 completely.
2. To apply Rule 9 only for IPv6 and not for IPv4. In IPv6, hierarchical address assignment is a general principle, hence the longest matching rule is beneficial in many cases. In IPv4, as stated above, the DNS based load balancing technique is widely used.
3. To apply Rule 9 for IPv6 conditionally and not for IPv4. When the length of matching bits of the destination address and the source address is longer than N, rule 9 is applied. Otherwise, the order of the destination addresses do not change. The N should be configurable and it should be 32 by default. This is simply because the two sites whose matching bit length is longer than 32 are probably adjacent.

Now that IPv6 PI addressing is being assigned by some RIRs, hierarchical address assignment is not fully maintained anymore. It seems that the longest matching algorithm may not be worth the adverse effect of disabling the DNS based load balance technique.

[Appendix C](#). Revision History

02:

Suresh Krishnan's comments were incorporated.
A new source address selection rule that utilizes the next-hop information is included in [Section 2.3](#)

01:

Restructured to contain only the actual changes to [RFC 3484](#).

00:

Published as a 6man working group item.

03:

Added acknowledgements.
Added longest matching algorithm malfunction regarding local DNS round robin.
The proposed changes section was restructured.
The issue of 6to4/Teredo and IPv4 prioritization was included.

The issue of deprecated addresses was added.
The renewed default policy table was changed accordingly.

02:

Added the reference to address selection design team's proposal.

01:

Matsumoto, et al.

Expires September 15, 2011

[Page 10]

Internet-Draft

[RFC3484](#) Bis

March 2011

The issue of private IPv4 address scope was added.
The issue of ULA address scope was added.
Discussion of longest matching rule was expanded.

Authors' Addresses

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

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

Jun-ya Kato
NTT SI Lab
Midori-Cho 3-9-11
Musashino-shi, Tokyo 180-8585
Japan

Phone: +81 422 59 2939
Email: kato@syce.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

Matsumoto, et al.

Expires September 15, 2011

[Page 11]