Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 18, 2011
A. Matsumoto
J. Kato
T. Fujisaki
NTT
October 15, 2010

Update to RFC 3484 Default Address Selection for IPv6
draft-ietf-6man-rfc3484-revise-01.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 fix the known defects.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction 3
2. Specification 3
  2.1. Changes related to the default policy table 3
    2.1.1. ULA in the policy table 3
    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. Private IPv4 address scope 5
  2.4. Deprecation of site-local unicast address 6
3. Security Considerations 6
4. IANA Considerations 6
5. References 6
  5.1. Normative References 6
  5.2. Informative References 7
Appendix A. Acknowledgements 7
Appendix B. Discussion 7
  B.1. Centrally assigned ULA 7
  B.2. 6to4, Teredo, and IPv4 prioritization 8
  B.3. Deprecated address 8
  B.4. The longest match rule 8
Appendix C. Revision History 9
Authors' Addresses 9

1. Introduction

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.

RFC 3484 has several known issues to be fixed. Deprecation of IPv6 site-local unicast address and the coming of ULA brought some preferable changes to the rules. Additionally, the rule 9 of the destination address selection rules, namely the longest matching rule, is known for its adverse effect on the round robin DNS technique.

This document specifies a set of updates that modify the algorithms and fix the known defects.

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 envionment. The changes we should consider for the default policy table are listed in this sub-section.

The policy table is defined to be configurable. The changes that are useful not universally but locally can be put into the policy table manually or by using the auto-configuration mechanism proposed as a DHCP option [I-D.fujisaki-dhc-addr-select-opt].

2.1.1. ULA in the policy table

RFC 5220 Section 2.1.4, 2.2.2, and 2.2.3 describes address selection problems related to ULA. These problems can be solved by changing the scope of ULA to site-local, and/or by adding an entry for default policy table entry that has its own label for ULA.

In its nature, ULA has global scope. This is because ULA's scope is expected to be defined in routing system. It may be the case that ULA and global IPv6 address are used for source and destination addresses of communication.

On the other hand, to prioritize ULA to ULA communication is basically reasonable. ULA should not be exposed to outside of its routable routing domain, so if ULA is given from the application as a candidate destination address, it can be generally expected that the ULA is within or at least close to the source host.

Therefore, the scope of ULA should be kept global, and prioritization of ULA to ULA communication should be implemented in policy table, by assigning its own label for ULA 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 or equal to 6to4, considering its characteristic of transitional tunnel mechanism. About Windows, this is already in the implementation.

2.1.3. Deprecated addresses in the policy table

IPv4-compatible IPv6 address is deprecated. [RFC4291] IPv6 site-local unicast address is deprecated. [RFC3879] Moreover, 6bone testing address was [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 will be:

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 the longest matching rule, which was found by Dave Thaler. It is malfunction of 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 CommonPrefixLen(DA & Netmask(Source(DA)), Source(DA)) > CommonPrefixLen(DB & Netmask(Source(DB)), Source(DB)), then prefer DA. Similarly, if CommonPrefixLen(DA & Netmask(Source(DA)), Source(DA)) < CommonPrefixLen(DB & Netmask(Source(DB)), Source(DB)), then prefer DB.

2.3. Private IPv4 address scope

As detailed in Remi's draft [I-D.denis-v6ops-nat-addrsel], when a host is in NATed site, and has a private IPv4 address and transitional addresses like 6to4 and Teredo, the host chooses transitional IPv6 address to access most of the dual-stack servers.

This is because private IPv4 address is defined to be site-local scope, and as in RFC 3484, the scope matching rules (Rule 2) set lower priority for private IPv4 address.

By changing the address scope of private IPv4 address to global, this problem can be solved. Considering the widely deployed NAT with IPv4 private address model, this change works in most of the cases. If not, this behavior can be overridden by configuring policy table, or by configuring routing table on a host.

Moreover, some modern OSs have already implemented this change.

2.4. Deprecation of site-local unicast address

RFC3484 contains a few "site-local unicast" and "fec::" description. It's better to remove examples related to site-local unicast address, or change examples to use ULA. Possible points to be re-written are below.

- 2nd paragraph in RFC 3484 Section 3.1 describes scope comparison mechanism.
- RFC 3484 Section 10 contains examples for site-local address.

3. Security Considerations

No security risk is found that degrades RFC 3484.

4. IANA Considerations

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

5. References

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

[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-Prefix Environments: Operational Issues of RFC 3484 Default Rules", RFC 5220, July 2008.

5.2. Informative References

[I-D.chown-addr-select-considerations]
Chown, T., "Considerations for IPv6 Address Selection Policy Changes", draft-chown-addr-select-considerations-03 (work in progress), July 2009.

[I-D.fujisaki-dhc-addr-select-opt]
Fujisaki, T., Matsumoto, A., and R. Hiromi, "Distributing Address Selection Policy using DHCPv6", draft-fujisaki-dhc-addr-select-opt-09 (work in progress), March 2010.

Appendix A. Acknowledgements

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 inputs.

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 adjancency may not have any relationships.

B.2. 6to4, Teredo, and IPv4 prioritization

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

B.3. Deprecated address

Discussion: These addresses was removed from the current specification. So, it 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 brought by them, however, it should be labeled differently from the legitimate address blocks.
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, the rule 9 defines that the longest matching destination and source address pair should be chosen. As in RFC 1794, the DNS based load balancing technique is achived by not re-ordering the destination addresses returned from the DNS server. The Rule 9 defines deterministic rule for re-ordering at hosts, hence the technique of RFC 1794 is not available anymore.

Regarding this problem, there was discussion in IETF and other places like below.

Discussion: The possible changes to RFC 3484 are as follows:

1. To delete Rule 9 completely.
2. To apply Rule 9 only for IPv6 and not for IPv4. In IPv6, hiearchical address assignment is general principle, hence the longest matchin 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, the 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 address is admitted in some RIRs, hierachical address assignment is not maintained anymore. It seems that the longest matching algorithm may not worth the adverse effect of disalbing the DNS based load balance technique.

Appendix C. Revision History

01:
Re-structured 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 re-structured.
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:
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