

Internet Draft
February 15, 2006
Expires August 2006

D. Thaler
Microsoft

Issues With Protocols Proposing Multilink Subnets
<[draft-thaler-intarea-multilink-subnet-issues-00.txt](#)>

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of 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/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>.

Copyright Notice

Copyright (C) The Internet Society (2006). All Rights Reserved.

Abstract

There have been several proposals around the notion that a subnet may span multiple links connected by routers. This memo documents the issues and potential problems that have been raised with such an approach.

1. Introduction

The original IPv4 address definition [[RFC791](#)] consisted of a Network field, identifying a network number, and a Local Address field, identifying a host within that network. As organizations grew to want many links within their network, their choices were (from [[RFC950](#)]) to:

1. Acquire a distinct Internet network number for each cable;

subnets are not used at all.

Thaler
Draft

Expires August 2006
Multilink Subnet Issues

1
February 2006

2. Use a single network number for the entire organization, but assign host numbers without regard to which LAN a host is on ("transparent subnets").

3. Use a single network number, and partition the host address space by assigning subnet numbers to the LANs ("explicit subnets").

[RFC925] was a proposal for option 2 which defined a specific type of ARP proxy behavior, where the forwarding plane had the properties of decrementing the TTL to prevent loops when forwarding, not forwarding packets destined to 255.255.255.255, and supporting subnet broadcast by requiring that the ARP-based bridge maintain a list of recent broadcast packets. This approach was never standardized.

Instead, the IETF standardized option 3 with [[RFC950](#)], whereby hosts were required to learn a subnet mask, and this became the IPv4 model.

Over the recent past there have been several newer protocols proposing to extend the notion of a subnet to be able to span multiple links, similar to [[RFC925](#)].

Early drafts of the IPv6 scoped address architecture [[SCOPID](#)] proposed a subnet scope above the link scope, to allow for multi-link subnets. This notion was rejected by the WG due to the issues discussed in this memo, and as a result the final version [[RFC4007](#)] has no such notion.

There was also a proposal to define multi-link subnets [[MLSR](#)] for IPv6. However this notion was abandoned by the IPv6 WG due to the issues discussed in this memo, and that proposal was replaced by a different mechanism which preserves the notion that a subnet spans only one link [[RFC4389](#)].

However, other WGs continued to allow for this concept even though it had been rejected in the IPv6 WG. Mobile IPv6 [[RFC3775](#)] allows tunnels to mobile nodes to use the same subnet as a home link, with the Home Agent doing layer-3 forwarding between them.

The notion also arises in Mobile Ad-hoc NETWORKS (MANETs) with proposals that an entire MANET is a subnet, with routers doing

layer-3 forwarding within it.

In this memo we document the issues raised in the IPv6 WG which motivated the abandonment of the multi-link subnet concept, so that designers of other protocols can (and should) be aware of the issues.

Thaler
Draft

Expires August 2006
Multilink Subnet Issues

2
February 2006

[2. Issues](#)

[2.1. IP Model](#)

The term "link" is generally used to refer to a topological area bounded by routers which decrement the IPv4 TTL or IPv6 Hop Limit when forwarding the packet. A link-local address prefix is defined in both IPv4 [[RFC3927](#)] and IPv6 [[RFC3513](#)].

The term "subnet" is generally used to refer to a topological area that uses the same address prefix, where that prefix is not further subdivided except into individual addresses.

In December 1995, the original IP Version 6 Addressing Architecture [[RFC1884](#)] was published, stating: "IPv6 continues the IPv4 model that a subnet is associated with one link. Multiple subnets may be assigned to the same link."

Thus it explicitly acknowledges that the current IPv4 model has been that a subnet is associated with one link, and that IPv6 does not change this model. Furthermore, a subnet is sometimes considered to be only a subset of a link, when multiple subnets are assigned to the same link.

The IPv6 addressing architecture has since been updated twice, first in July 1998 [[RFC2373](#)] and again in April 2003 [[RFC3513](#)]. Both updates include the language: "Currently IPv6 continues the IPv4 model that a subnet prefix is associated with one link. Multiple subnet prefixes may be assigned to the same link."

Clearly the notion of a multi-link subnet would be a change to the existing IP model.

Similarly, the Mobility Related Terminology [[RFC3753](#)] defines a Foreign subnet prefix as "A bit string that consists of some number of initial bits of an IP address which identifies a node's foreign

link within the Internet topology" with a similar definition for a Home subnet prefix. These both state that the subnet prefix identifies a (singular) link.

[2.2](#). TTL/Hop Limit Issues

Since a link is bounded by routers that decrement the IPv4 TTL or IPv6 Hop Limit, there may be issues with applications and protocols that make any assumption about the relationship between TTL/Hop Limit and subnet prefix.

There are two main cases which may arise. Some applications and protocols may send packets with a TTL/Hop Limit of 1. Other applications and protocols may send packets with a TTL/Hop Limit of 255, and verify that the value is 255 on receipt. Both are ways of limiting communication to within a single link.

Thaler
Draft

Expires August 2006
Multilink Subnet Issues

3
February 2006

As for assumptions about the relationship between TTL/Hop Limit and subnet, let's look at some example references familiar to many protocol and application developers.

Stevens' "Unix Network Programming, 2nd ed." [[UNP](#)] states on page 490 "a TTL of 0 means node-local, 1 means link-local" (this of course being true by the definition of link). Then page 498 states, regarding IP_MULTICAST_TTL and IPV6_MULTICAST_HOPS, "If this is not specified, both default to 1, which restricts the datagram to the local subnet." Here, Unix programmers learn that TTL=1 packets are restricted to a subnet (as opposed to a link). This is typical of many documents which use the terms interchangeably due to the IP Model described earlier.

Similarly, "TCP/IP Illustrated, Volume 1" [[TCPILL](#)] states on page 182: "By default, multicast datagrams are sent with a TTL of 1. This restricts the datagram to the same subnet."

Steve Deering's original multicast README file [[DEERING](#)] contained the statement "multicast datagrams with initial TTL 1 are restricted to the same subnet", and similar statements now appear in many vendors' documentation, including documentation for Windows (e.g., [[TCPIP2K](#)]) and Linux (e.g., [[LINUX](#)] says a TTL of 1 is "Restricted to the same subnet. Won't be forwarded by a router.")

The above are only some examples. There is no shortage of places where application developers are being taught that a subnet is confined to a single link, and so we must expect that arbitrary applications may embed such assumptions.

Some examples of protocols today that are known to embed some assumption about the relationship between TTL and subnet prefix are:

- o Neighbor Discovery [[RFC2461](#)] uses messages with Hop Limit 255 checked on receipt, to resolve the link-layer address of any IP address in the subnet.
- o Apple's Bonjour [[MDNS](#)] uses messages with TTL 255 checked on receipt, and only responds to queries from addresses in the same subnet. (Note that multilink subnets do not necessarily break this, as this behavior is to constrain communication to within a subnet, where a subnet is only a subset of a link; however it will not work across a multi-link subnet.)

Some other examples of protocols today that are known to use a TTL 1 or 255, but do not appear to explicitly have any assumption about the relationship to subnet prefixes (other than the well-known link-local prefix) include:

- o [[LLMNR](#)] uses a TTL/Hop Limit of 1.
- o MLDv2 [[RFC3810](#)] uses a Hop Limit of 1.

Thaler
Draft

Expires August 2006
Multilink Subnet Issues

4
February 2006

- o Reverse tunneling for Mobile IPv4 [[RFC3024](#)] uses TTL 255 checked on receipt for Registration Requests sent to foreign agents.
- o [[RFC3927](#)] discusses the use of TTL=1 and TTL=255 within the IPv4 link-local address prefix.

It is unknown whether any implementations of such protocols exist that add such assumptions about the relationship to subnet prefixes for other reasons.

[2.3](#). Link-scoped multicast and broadcast

Because multicast routing is not ubiquitous, the notion of a subnet which spans multiple links tends to result in cases where multicast does not work across the subnet. Per [[RFC2644](#)], the default behavior is that routers do not forward broadcast packets either.

There are many protocols and applications today that use link-scoped multicast. The list of such applications and protocols that have been assigned their own link-scoped multicast group address (and may also have assumptions about the TTL/Hop Limit as noted above) can be found at:

<http://www.iana.org/assignments/multicast-addresses>

<http://www.iana.org/assignments/ipv6-multicast-addresses>

In addition, an arbitrarily large number of other applications may be using the all-1's broadcast address, or the all-hosts link-scoped multicast address, rather than their own group address.

The well-known examples of protocols using link-scoped multicast or broadcast generally fall into one of the following groups:

- o Routing protocols: DVMRP, OSPF, RIP, EIGRP, etc. These protocols exchange routes to subnet prefixes.
- o Addressing protocols: ND, DHCPv4, DHCPv6, Teredo, etc. By their nature this group tends to embed assumptions about the relationship between a link and a subnet prefix. For example, ND [[RFC2461](#)] uses link-scoped multicast to resolve the link-layer address of an IP address in the same subnet prefix, and to do duplicate address detection (see [section 2.4](#) below) within the subnet. DHCP uses link-scoped multicast or broadcast to obtain an address in the subnet. Teredo [[RFC4380](#)] states: "An IPv4 multicast address used to discover other Teredo clients on the same IPv4 subnet. The value of this address is 224.0.0.253", which is a link-scoped multicast address. It also says "the client MUST silently discard all local discovery bubbles [...] whose IPv4 source address does not belong to the local IPv4 subnet".

Thaler
Draft

Expires August 2006
Multilink Subnet Issues

5
February 2006

- o Service discovery protocols: SSDP, Bonjour, WS-Discovery, etc. These often do not define any explicit assumption about the relationship to subnet prefix.
- o Name resolution protocols: NetBios [[RFC1001](#)], Bonjour [[MDNS](#)], LLMNR, etc. Most often these do not define any explicit assumption about the relationship to subnet prefix, but [[MDNS](#)] only responds to queries from addresses within the same subnet prefix.

Note that protocols such as Bonjour and Teredo which drop packets which don't come from an address within the subnet are not necessarily broken by multilink subnets, as this behavior is meant to constrain the behavior to within a subnet, when a link is larger than a single subnet.

However, regardless of whether any assumption about the relationship to subnet prefixes exists, all protocols mentioned above or on the IANA assignments list will not work across a multilink subnet without protocol-specific proxying functionality in routers, and adding proxying for an arbitrary number of protocols and applications does not scale. Furthermore, it may hinder the development and use of future protocols using link-scoped multicast.

[2.4.](#) Duplicate Address Detection Issues

Duplicate Address Detection (DAD) uses link-scoped multicast in IPv6, and link-scoped broadcast in IPv4 and so has the issues mentioned in [Section 2.3](#) above.

In addition, [[RFC2461](#)] contains the statement:

"Thus, for a set of addresses formed from the same interface identifier, it is sufficient to check that the link-local address generated from the identifier is unique on the link. In such cases, the link-local address MUST be tested for uniqueness, and if no duplicate address is detected, an implementation MAY choose to skip Duplicate Address Detection for additional addresses derived from the same interface identifier."

The last possibility, sometimes referred to as Duplicate Interface Identifier Detection (DIID), has been a matter of much debate, and the current draft in progress states:

Each individual unicast address SHOULD be tested for uniqueness. Note that there are implementations deployed that only perform Duplicate Address Detection for the link-local address and skip the test for the global address using the same interface identifier as that of the link-local address. Whereas this document does not invalidate such implementations, this kind of "optimization" is NOT RECOMMENDED, and new implementations MUST NOT do that optimization.

Thaler
Draft

Expires August 2006
Multilink Subnet Issues

6
February 2006

The existence of such implementations also causes problems with multilink subnets. Specifically, a link-local address is only valid within a link, and hence is only tested for uniqueness within a single link. If the same interface identifier is then assumed to be unique across all links within a multilink subnet, address conflicts can occur.

[3.](#) Security Considerations

The notion of multilink subnets can cause problems with any security protocols which either rely on the assumption that a subnet only spans a single link, or can leave gaps in the security solution where protocols are only defined for use on a single link.

Secure Neighbor Discovery [[RFC3971](#)], in particular, is currently only defined within a single link. If a subnet were to span multiple links, SEND would not work as currently specified. This same problem also exists in cases where a subnet does not span multiple links but where Neighbor Discovery is proxied within a link. [Section 9 of \[RFC4389\]](#) discusses some possible future directions in this regard.

Furthermore, as noted above some applications and protocols (ND, Bonjour, Mobile IPv4, etc.) mitigate against off-link spoofing attempts by requiring a TTL or Hop Limit of 255 on receipt. If this restriction were removed, or if alternative protocols were used, then off-link spoofing attempts would become easier, and some alternative way to mitigate against such attacks would be needed.

[4.](#) IANA Considerations

This document has no actions for IANA.

[5.](#) Normative References

- [RFC791] Postel, J., "Internet Protocol", STD 5, [RFC 791](#), September 1981.
- [RFC950] Mogul, J. and J. Postel, "Internet Standard Subnetting Procedure", STD 5, [RFC 950](#), August 1985.
- [RFC2461] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", [RFC 2461](#), December 1998.
- [RFC2644] Senie, D., "Changing the Default for Directed Broadcasts in Routers", [BCP 34](#), [RFC 2644](#), August 1999.
- [RFC3513] Hinden, R. and S. Deering, "Internet Protocol Version 6 (IPv6) Addressing Architecture", [RFC 3513](#), April 2003.

Configuration of IPv4 Link-Local Addresses", [RFC 3927](#), May 2005.

[RFC3971] Arkko, J., Ed., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", [RFC 3971](#), March 2005.

[RFC4007] Deering, S., Haberman, B., Jinmei, T., Nordmark, E., and B. Zill. "IPv6 Scoped Address Architecture", [RFC 4007](#), March 2005.

6. Informative References

[DEERING] Deering, S., "IP Multicast Extensions for 4.3BSD UNIX and related systems (MULTICAST 1.2 Release)", June 1989.
<http://www.kohala.com/start/mcast.api.txt>

[LINUX] Juan-Mariano de Goyeneche, "Multicast over TCP/IP HOWTO", March 1998. <http://www.linux.com/howtos/Multicast-HOWTO-2.shtml>

[LLMNR] Aboba, B., Thaler, D. and L. Esibov, "Linklocal Multicast Name Resolution (LLMNR)", [draft-ietf-dnsext-mdns-45.txt](#), October 2005.

[MDNS] Cheshire, S. and M. Krochmal, "Multicast DNS", Internet Draft, June 2005. <http://files.multicastdns.org/draft-cheshire-dnsext-multicastdns.txt>

[MLSR] Thaler, D. and C. Huitema, "Multi-link Subnet Support in IPv6", [draft-ietf-ipv6-multilink-subnets-00.txt](#) (expired), June 2002. <http://www.ietf.org/proceedings/02jul/I-D/draft-ietf-ipv6-multilink-subnets-00.txt>

[RFC925] Postel, J., "Multi-LAN address resolution", [RFC 925](#), October 1984.

[RFC1001] NetBIOS Working Group in the Defense Advanced Research Projects Agency, Internet Activities Board, End-to-End Services Task Force, "Protocol standard for a NetBIOS service on a TCP/UDP transport: Concepts and methods", [RFC 1001](#), March 1987.

[RFC1884] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", [RFC 1884](#), December 1995.

[RFC2373] R. Hinden, S. Deering, "IP Version 6 Addressing Architecture", [RFC 2373](#), July 1998.

[RFC3024] G. Montenegro, Ed., "Reverse Tunneling for Mobile IP, revised", [RFC 3024](#), January 2001.

- [RFC3753] J. Manner, Ed., M. Kojo, Ed., "Mobility Related Terminology", [RFC 3753](#), June 2004.
- [RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", [RFC 3775](#), June 2004.
- [RFC3810] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", [RFC 3810](#), June 2004.
- [RFC4380] C. Huitema, "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", [RFC 4380](#), February 2006.
- [RFC4389] Thaler, D., Talwar, M., and C. Patel, "Neighbor Discovery Proxies (ND Proxy)", [RFC 4389](#), February 2006.
- [SCOPID] Deering, S., Haberman, B., Jinmei, T., Nordmark, E., Onoe, A., and B. Zill. "IPv6 Scoped Address Architecture", Internet-Draft (Obsolete), March 2005.
<http://www.ietf.org/proceedings/02jul/I-D/draft-ietf-ipngwg-scoping-arch-04.txt>
- [TCPIILL] Stevens, W. Richard, "TCP/IP Illustrated, Volume 1", Addison-Wesley, 1994.
- [TCPIP2K] MacDonald, D. and W. Barkley, "Microsoft Windows 2000 TCP/IP Implementation Details".
<http://www.microsoft.com/technet/itsolutions/network/deploy/depovg/tcpip2k.msp>
- [UNP] Stevens, W. Richard, "Unix Network Programming, Volume 1, Second Edition", Prentice Hall, 1998.

Authors' Addresses

Dave Thaler
Microsoft
One Microsoft Way
Redmond, WA 98052
Phone: +1 425 703 8835
Email: dthaler@microsoft.com

Full Copyright Statement

Copyright (C) The Internet Society (2006).

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY 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 HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF 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 this 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. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR 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 this standard. Please address the information to the IETF at ietf-

ipr@ietf.org.

Thaler

Expires August 2006

10