Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: April 15, 2011

R. Despres
RD-IPtech
B. Carpenter
Univ. of Auckland
S. Jiang
Huawei Technologies Co., Ltd

October 12, 2010

Native IPv6 Across NAT44 CPEs (6a44)
draft-despres-softwire-6a44-01

Abstract

Most CPEs should soon be dual stack, but a large installed base of IPv4-only CPEs is likely to remain for several years. Also, with the IPv4 address shortage, more and more ISPs will assign private IPv4 addresses to their customers. The need for IPv6 connectivity therefore concerns hosts behind IPv4-only CPEs, including such CPEs that are assigned private addresses. The 6a44 mechanism specified in this document addresses this need, without limitations and operational complexities of Tunnel Brokers and Teredo to do the same.

6a44 is based on an address mapping and on a mechanism whereby suitably upgraded hosts behind a NAT may obtain IPv6 connectivity via a stateless 6a44 server function operated by their Internet Service Provider. With it, IPv6 traffic between two 6a44 hosts in a single site remains within the site. Except for IANA numbers that remain to be assigned, the specification is intended to be complete enough for running codes to be independently written and interwork.

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 April 15, 2011.

Table of Contents

1. Introduction 3
2. Applicability 4
3. 6a44 IPv6 Address Format 6
4. Address Mappings and Encapsulations 8
5. MTU considerations 10
6. Host Acquisition of IPv6 Addresses and their Lifetimes 10
7. Security considerations 13
8. IANA Considerations 14
9. Acknowledgments 14
10. References 14
10.1. Normative References 14
10.2. Informative References 15
Authors' Addresses 16

1. Introduction

Most CPEs (customer premise equipments) should soon be dual stack, but a large installed base of IPv4-only CPEs is likely to remain for several years. Also, with the IPv4 address shortage, more and more Internet service providers (ISPs) will assign private IPv4 addresses of [RFC1918] to their customers. The need for IPv6 connectivity therefore includes hosts behind IPv4-only CPEs, including such CPEs that have private addresses.

At the moment, there are two traversal techniques to address this need:

1. A configured tunnel (IPv6 in IPv4 or even IPv6 in UDP), involving a managed tunnel broker, e.g. [RFC3053], with which the user must register. Well known examples include deployments of the Hexago tool, and the SixXs collaboration. However, this approach does not scale well; it requires significant support effort and is really only suitable for "hobbyist" early adopters of IPv6.

2. Teredo [RFC4380]. This is an automatic UDP-based tunneling solution that relies on a Teredo server, and on Teredo relays willing to carry the traffic. Unfortunately experience shows that this is sometimes an unreliable process in practice, with clients sometimes believing that they have Teredo connectivity when in fact they don't, or alternatively with the Teredo server and relay being very remote from the client and causing extremely long latency for IPv6 packets. This leads to user frustration and even to advice from help desks to disable IPv6.

6a44 is based on an address mapping and on a mechanism whereby suitably upgraded hosts behind a NAT may obtain IPv6 connectivity via a stateless 6a44 server function operated by their Internet Service Provider.

To address this need without the mentioned limitations, 6a44 is based on an address mapping and on a mechanism whereby suitably upgraded hosts behind a NAT may obtain IPv6 connectivity via a stateless 6a44 server function operated by their ISP. It can apply even with ISPs that, due to the IPv4 address shortage, assign private addresses of [RFC1918] to their IPv4 customers (typically with prefix 10.0.0.0/8).

6a44 is only a transition technology. It will no longer have to be used when the number of IPv4-only CPEs becomes negligible.

Except for IANA numbers that remain to be assigned, the specification is intended to be complete enough for running codes to be independently written and interwork.

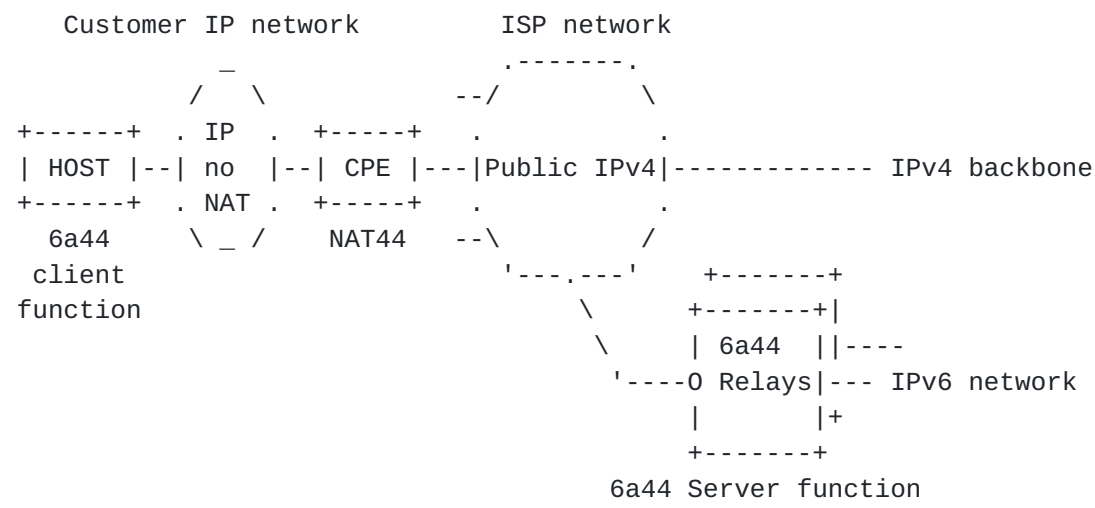
2. Applicability

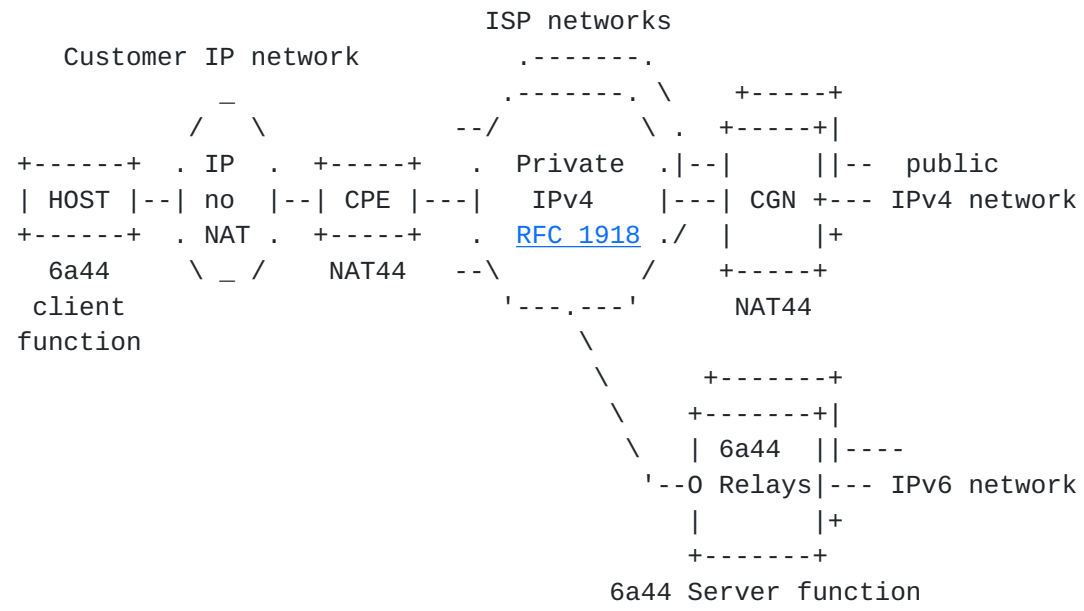
Both hosts and ISPs can be made 6a44 capable independently of each other, with 6a44 being actually used by 6a44 capable hosts where their local ISPs are 6a44-capable.

For a host to be 6a44 capable, it has to support the 6a44 Client function ("6a44-C" in some Figures). This function is placed in its TCP/IP stack at the same place as the 6to4 router function of [RFC3056]: it has an IPv4 interface in its link-layer direction and both an IPv4 interface and an IPv6 pseudo-interface in its higher layers direction.

To enable its 6a44 function, a host must have no intra-site NAT44 between itself and the site CPE. (In sites where there are intra-site NAT44s, these NATs should be configured so that hosts behind them cannot enable 6a44. In view of the specification below, it can be done with a port mapping in each of them between the well-known port of 6a44 and an internal private address that DHCP doesn't assign.) In addition, the host must have in IPv4 a link MTU of at least 1308 octets (the MTU to be guaranteed in IPv6 plus the length of an UDP/IPv4 encapsulation header).

For an IPv4 ISP network to be 6a44 capable, the ISP must operate the 6a44 Server function, ("6a44-S" in some Figures). This function is anywhere at its border between the IPv4 network and an IPv6 network in which it has a /48 prefix. Typically this prefix will be chosen from whatever shorter PA prefix has been allocated to the ISP. The 6a44 server function can be replicated in any number of routers, known as "6a44 Relays", to enhance service quality and service availability. Also, the network must have an IPv4 MTU of at least 1308 octets and, for security, must support the ingress filtering of [RFC3704] (see Section 7).





6a44 ISP CONFIGURATIONS

Figure 1

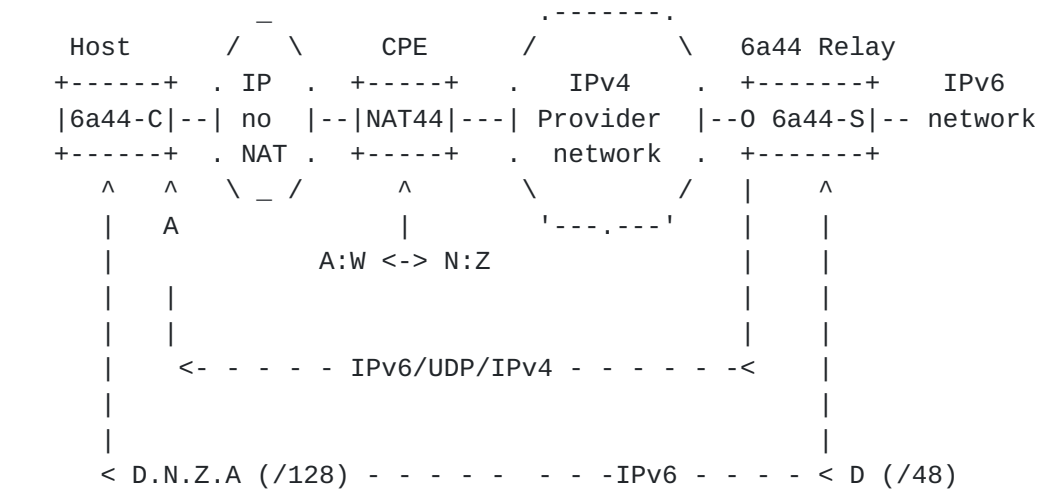
Each ISP can support one public-addressing and several private-addressing 6a44 networks.

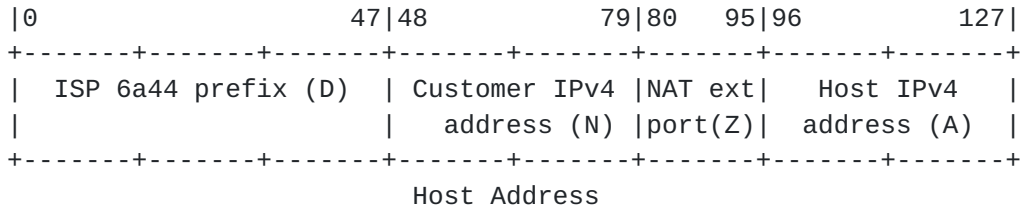
In 6a44 networks, ISPs may route IPv6 in addition to IPv4. Where this is the case, 6a44 only concerns CPEs that are IPv4-only capable. If on the contrary IPv4 is the only routed address family, 6a44 may also concerns sites where CPEs are dual-stack capable. Unable to take advantage of their IPv6 capability, they act as if they would be IPv4-only.

Figure 1 illustrates ISP-network configurations on which 6a44 can be used.

NOTE: The objective of 6a44 differs from that of Teredo ([RFC4380] and [RFC5991]). Teredo has been designed to avoid needing any ISP participation. This has permitted early deployment but didn't ensure connectivity between all Teredo addresses and all native IPv6 addresses. Also, it imposed a very significant level of complexity. On the contrary, 6a44 is designed to be explicitly supported by ISPs. As a result, connectivity between 6a44 IPv6 addresses and all native IPv6 addresses can be ensured, and implementations can remain simple.

3. 6a44 IPv6 Address Format





HOST-ADDRESS CONSTRUCTION

Figure 2

The 6a44 IPv6 address an ISP assigns to a host must first contain all what is needed to reach it from the IPv6 backbone. This includes, as illustrated in Figure 2:

- the IPv6 prefix D that the ISP has assigned border routers of its 6a44 network;
- the IPv4 address N of the customer site (external address of the NAT44 in its CPE);
- the port Z that, in the CPE NAT44 CPE, has to be used to reach the host at its address address A, and in the host the 6a44 well-known port W (to be assigned by IANA).

To ensure that two 6a44 hosts behind the same IPv4-only CPE exchange packets without entering the ISP network, the 6a44 address of each host must also contain its IPv4 address A.

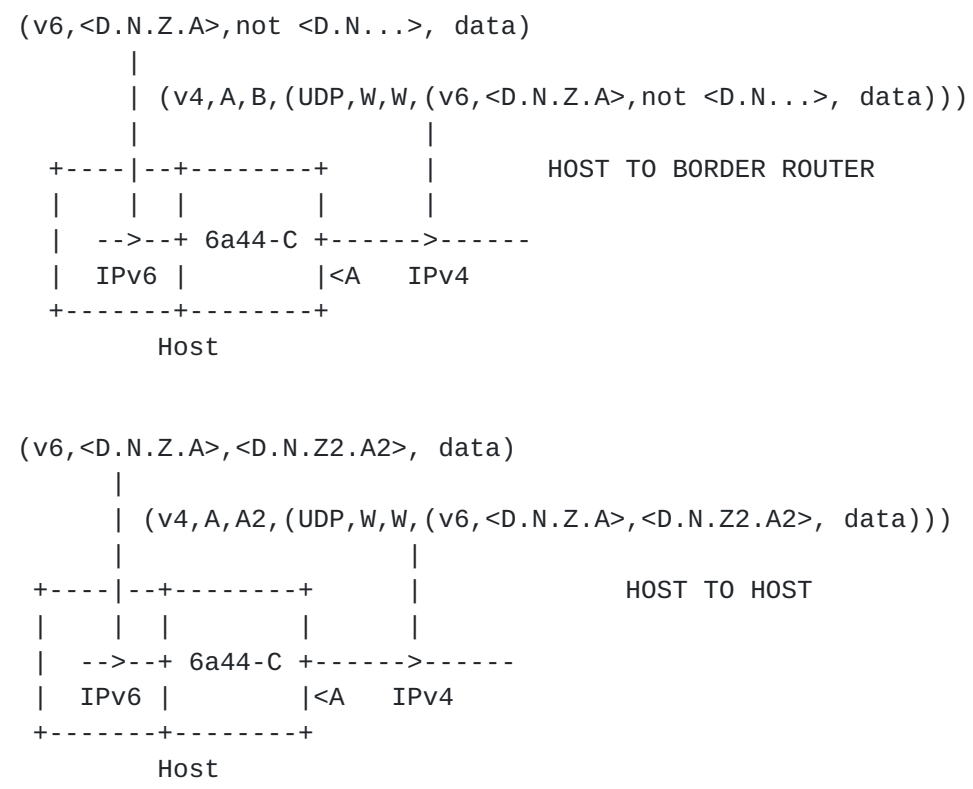
The format of 6a44 IPv6 addresses, a concatenation of D,N,Z, and A, where D has to be a /48 prefix, is detailed in Figure 2.

NOTE: Since IPv6 prefixes D assigned by ISPs to their customers always start with 001, the prefix of all IPv6 Aggregatable Global Unicast addresses specified in [RFC2374], 6a44 IPv6 addresses bend the rule of [RFC4291] that says 'for all unicast addresses, except those that start with binary value 000, Interface IDs are required to be 64 bits long and to be constructed in Modified EUI-64 format". This is however acceptable in practice because 6a44 addresses are never used on any real IPv6 link, and in particular never subject to the neighbor discovery protocol of [RFC2461] which depends on properties of interface IDs. A revision of the [RFC4291] sentence should eventually clarify this point.

4. Address Mappings and Encapsulations

Figure 3 and Figure 4 detail the address mappings and encapsulations/ decapsulations to be performed by 6a44 Client and server functions respectively, with the following notation:

- (vX,A1,A2,data): a packet of the IPvX version that has A1 as source address, A2 as destination address, and "data" as payload. (UDP,P1,P2,data): a UDP IP payload that has P1 as source port, P2 as destination port, and "data" as payload.
- B is the 6a44 well-known anycast address, that of the 6a44 Server function. X...: an address that starts with prefix X.
- not X: an address different from X
- X.Y: the concatenation of X and Y (the dot is the concatenation operator).



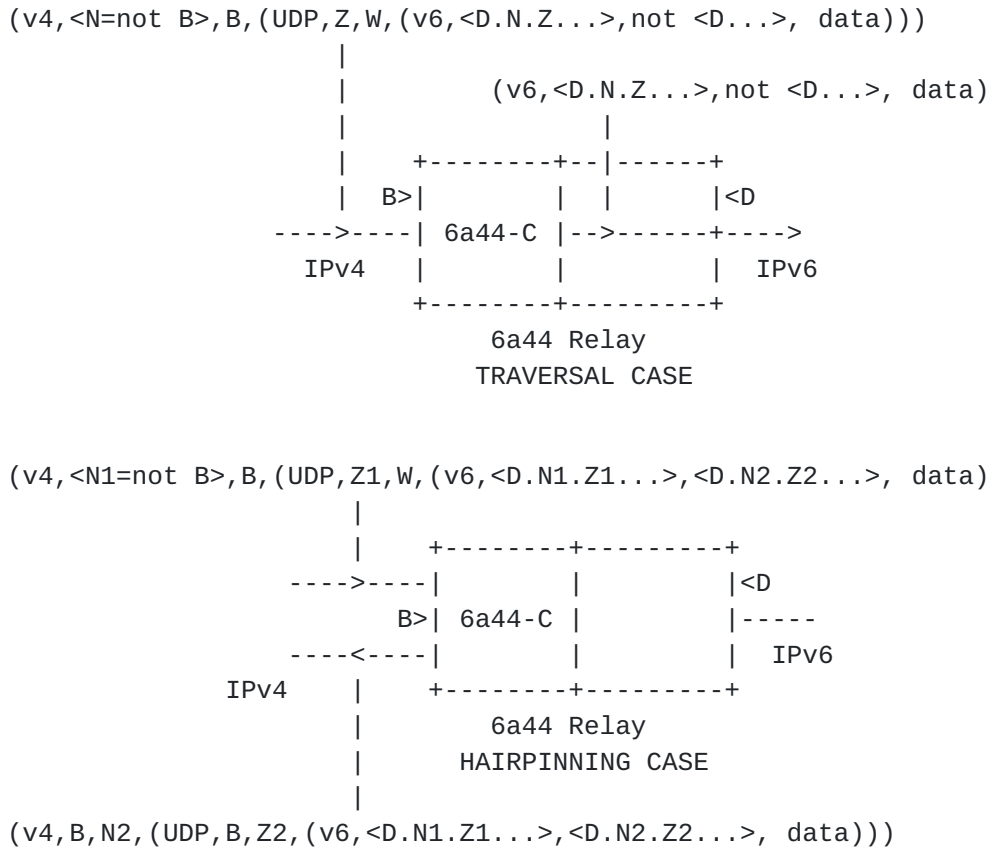
HOST MAPPINGS AND ENCAPSULATIONS

Figure 3

For protection against spoofing attacks, decapsulating functions must check consistency of IPv6 addresses fields with IPv4 addresses and UDP ports of encapsulating headers, both for source and destination addresses.

Figures present only one direction of 6a44-function traversals, but mappings that apply to the reverse direction are the same, with just a permutation of source and destination fields, for all of IPv4, IPv6, and UDP. Mappings and encapsulations/decapsulations for the reverse direction of that presented in Figures are the same, but with with source and destination permuted in IPv6, IPv4 and UDP.

Recommendations of [RFC4213] that concern these encapsulations have to be followed.



6a44-RELAY MAPPINGS AND ENCAPSULATIONS

Figure 4

5. MTU considerations

Reassembly of multi-fragment datagrams needs stateful processing, and opens the door to some denial of service attacks. To ensure a freedom of distribution of 6a44 Server functions in any number of parallel processors anywhere in 6a44 ISP networks, it has therefore to be avoided.

For this:

- 6a44 ISP networks must have internal IPv4 MTUs of at least 1308 octets (which is easy to ensure).

- 6a44 hosts must limit to 1280 octets IPv6 packets they transmit to destinations that are not neighbors on their own links. This behavior is already the normal one as long as no other IPv6 path MTU has been reliably discovered.

- 6a44 Server functions refuse packets received from their IPv6 pseudo interfaces if their sizes exceed 1280 octets, with ICMPv6 Packet Too Big messages returned to sources as required by [RFC2460].)

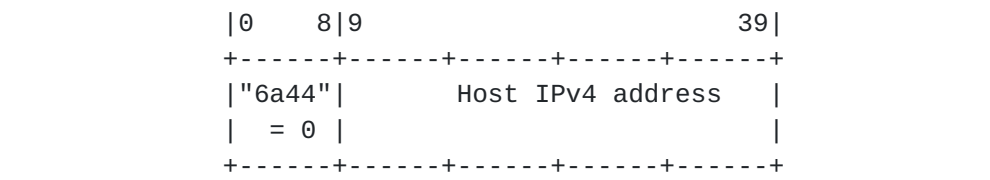
In a host, a destination is considered to be an on link neighbor if the IPv6 destination has the same bits 0-79 as the host address, and if the IPv4 destination starts with the prefix of the IPv4 link of the host. In this case, the IPv6 path MTU can be taken as that of the IPv4 link MTU minus 28 octets (a value that is typically significantly longer that 1280 octets).

6. Host Acquisition of IPv6 Addresses and their Lifetimes

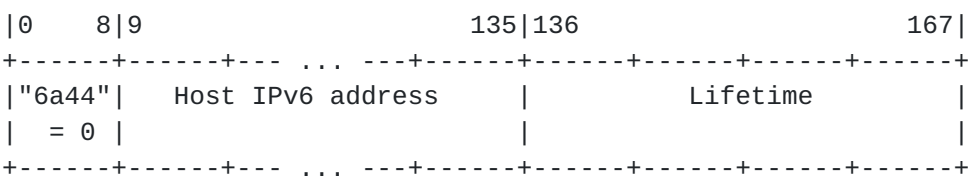
Acquisition of 6a44 addresses by hosts is independent from other mechanisms they may have to acquire other IPv6 addresses (PPP, DHCP, SLAAC, ...). It only depends on 6a44 packet exchanges between hosts and 6a44 Relays.

In order to acquire 6a44 addresses, hosts transmit IPv6 Address Request messages to 6a44 Server functions and expect IPv6 Address Indication messages in return.

Formats of these 6a44 messages are shown in Figure 5. They start with a 6a44 mark, a null octet chosen so that, in payloads of UDP datagrams received by 6a44 Client and 6a44 Server functions, 6a44 messages can be distinguished from IPv6 packets (IPv6 packets always have a non-null first octet).



IPv6 ADDRESS REQUEST (6a44 Client to 6a44 Server)

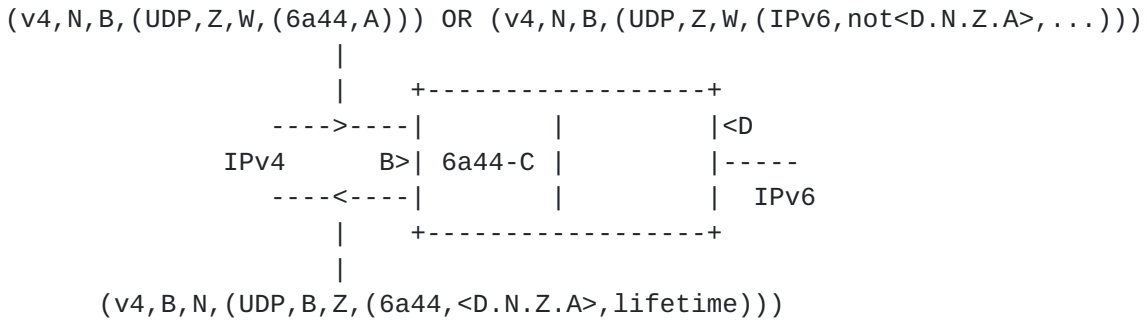


IPv6 ADDRESS INDICATION (6a44 Server to 6a44 Client)

6a44 MESSAGES

Figure 5

Message processing in 6a44 Server function is shown in Figure 6 with the same notation as in Section 4. The lifetime of returned IPv6 addresses should be the same as that of IPv4 addresses assigned by the same ISP. it is expressed in seconds.



6a44 MESSAGE PROCESSING IN BORDER ROUTERS

Figure 6

In a host, the 6a44 Client function should be activated for one of its physical interfaces only if this interface has a private IPv4 address and no other native IPv6 address. (An address is said to be native if it starts with 2000::/3 (global unicast) and neither with 2002::/16 (the 6to4 prefix) nor with 2001::/32 (the Teredo prefix).)

Message processing in a 6a44 Client function consists in transmitting from time to time IPv6 Address Requests to the 6a44 Server function, and to update the host IPv6 address and its lifetime each time an IPv6 Address Indication message is received (with due IPv4 source address verification for security).

In order to decide when to transmit such a message, the 6a44 Client function has the equivalent of the following states:

"Waiting for an IPv6 Address Indication": When this state is entered, an IPv6 Address Request is transmitted, a Response Awaited timer of 1 second is started, and a Retransmission Count is set to 0. If the timer expires with a Retransmission Count less than 10, a new IPv6 Address Request is transmitted, and the count is increased by 1. If it expires with a count equal to 10, the state is changed to "waiting before a new attempt to find a 6a44 service". If an IPv6 Address Indication is received while in this state, the timer is stopped, the state is changed to "Waiting for having to refresh the NAT-binding". This state is also re-entered each time a new IPv4 address is assigned to the link-direction interface of the 6a44 Client function.

"Waiting for having to refresh the NAT-binding": When this state is entered, a timer of 29 second is started.(This value is that chosen for SIP in [RFC5626] for the same objective, i.e. to maintain tunnel NAT bindings without particular knowledge about NAT specifics.) This timer is restarted each time an IPv6 packet is transmitted to the 6a44 Server function (not when a packet is transmitted host to host within the customer site). It is also restarted if an IPv6 Address Indication is received while in this state. (This may happen in particular if the NAT binding has changed, e.g. because CPE reset during the lifetime of the IPv6 address.) If the timer expires, the state is changed to "Waiting for an IPv6 Address Indication".

"waiting before a new attempt to find a 6a44 service": When this state is entered, a 6a44 Availability timer of 1 hour is started. When it expires, the state is changed to "Waiting for an IPv6 Address Indication".

7. Security considerations

Traffic-capture attack by a neighbor: If it would be possible to transmit from a neighboring site a bogus address indication to a 6a44 host, this host could inadvertently advertise an IPv6 address that is not his real 6a44 address. Some incoming connections that it should have received could then be redirected to a wrong address. However, because 6a44 is applicable only to ISP networks that support the ingress filtering of [RFC3704] (see Section 2), no neighbor can fake a valid Address Indication message (the IPv4 source of packets it sends cannot be the 6a44 well-known IPv4 address, the only valid source for an Address Indication message).

Spoofing attacks: With address checks of Section 4, 6a44 should introduce no spoofing vulnerabilities beyond those the underlying IPv4 networks may have. ISPs that use subscriber authentications to secure IPv4 address assignments have the effect of this authentication automatically extended to 6a44 addresses (they include the assigned IPv4 addresses).

Denial-of-service attacks: Provided 6a44 Server functions are provisioned with enough processing power, which is facilitated by their being stateless, 6a44 is expected to introduce no denial of service vulnerabilities of its own.

Subscriber authentication: This is not provided as part of 6a44, because it is assumed to have occurred when the IPv4 address assignment was made.

Routing-loop attacks: A risk of routing-loop attacks has been identified for some encapsulation/decapsulation mechanisms [draft-ietf-v6ops-tunnel-loops-00]. It doesn't exist with 6a44 because:

* IPv4 packets entering a 6a44 Server function are not forwarded if they come from another instance of the 6a44 Server function itself, i.e. if the IPv4 source is the 6a44 well-known IPv4 address Section 4.

* The encapsulation header, which is based on UDP with a specific well-known port, cannot be confused with that of other encapsulation mechanisms (in particular those of IP in IP like those of 6to4, 6rd and ISATAP).

Missing 6a44 Server function: If a 6a44-capable host is client of an ISP that doesn't support 6a44, 6a44 IPv6 Address Request messages transmitted by the host will be forwarded to the Internet backbone, with the 6a44 well-known IPv4 address as destination. Since this address doesn't start with any prefix that the backbone routes toward ISP networks, these messages will be discarded before reaching any place where a fake 6a44 Server could have been malevolently placed. There is therefore no danger that 6a44 hosts could have their IPv6 traffic routed via 6a44 Server functions that would not belong to their local ISP (i.e. where they could be observed and acted upon without control).

8. IANA Considerations

For 6a44 to be used, both its IPv4 well-known address B and its well-known port W need to be assigned by IANA.

This assignment is necessary to validate the plug-an-play operation of 6a44 with independent implementations. Having it as quickly as possible (i.e. without waiting for all details of the specification to be agreed on), would be helpful for an early validation of the 6a44 plug-and-play operation.

9. Acknowledgments

This specification results from a convergence effort of authors of [draft-despres-softwire-6rdplus-00] and [draft-carpenter-softwire-sample-00]. Useful comments have been received about these earlier proposals or later, in particular from Pascal Thubert, Dan Wing, Yu Lee, Olivier Vautrin, Fred Templin, and Ole Troan. They have to be thanked for their contributions.

10. References

10.1. Normative References

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC4213] Nordmark, E. and R. Gilligan, "Basic Transition Mechanisms for IPv6 Hosts and Routers", RFC 4213, October 2005.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

10.2. Informative References

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC2374] Hinden, R. and S. Deering, "An IPv6 Aggregatable Global Unicast Address Format", RFC 2374, July 1998.

[RFC2461] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[RFC3053] Durand, A., Fasano, P., Guardini, I., and D. Lento, "IPv6 Tunnel Broker", RFC 3053, January 2001.

[RFC3056] Carpenter, B. and K. Moore, "Connection of IPv6 Domains via IPv4 Clouds", RFC 3056, February 2001.

[RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", BCP 84, RFC 3704, March 2004.

[RFC4380] Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", RFC 4380, February 2006.

[RFC5626] Jennings, C., Mahy, R., and F. Audet, "Managing Client-Initiated Connections in the Session Initiation Protocol (SIP)", RFC 5626, October 2009.

[RFC5991] Thaler, D., Krishnan, S., and J. Hoagland, "Teredo Security Updates", RFC 5991, September 2010.

[draft-carpenter-softwire-sample-00] Carpenter, B. and S. Jiang, "Legacy NAT Traversal for IPv6: Simple Address Mapping for Premises - Legacy Equipment (SAMPLE)", June 2010.

[draft-despres-softwire-6rdplus-00] Despres, R., "Rapid Deployment of Native IPv6 Behind IPv4 NATs (6rd+)", July 2010.

[draft-ietf-v6ops-tunnel-loops-00] Nakibly, G. and F. Templin, "Routing Loop Attack using IPv6 Automatic Tunnels: Problem Statement and Proposed Mitigations - Work in progress", September 2010.

Authors' Addresses

Remi Despres
RD-IPtech
3 rue du President Wilson
Levallois,
France

Email: remi.despres@free.fr

Brian Carpenter
Department of Computer Science
University of Auckland
PB 92019
Auckland, 1142
New Zealand

Email: brian.e.carpenter@gmail.com

Sheng Jiang
Huawei Technologies Co., Ltd
KuiKe Building, No.9 Xinxi Rd.,
Shang-Di Information Industry Base, Hai-Dian District, Beijing,
P.R. China

Email: shengjiang@huawei.com