Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 9, 2010

B. Carpenter
Univ. of Auckland
S. Jiang
Huawei Technologies Co., Ltd

June 7, 2010

Legacy NAT Traversal for IPv6: Simple Address Mapping for Premises Legacy Equipment (SAMPLE)
draft-carpenter-softwire-sample-00

Abstract

IPv6 deployment is delayed by the existence of millions of subscriber network address translators (NATs) that cannot be upgraded to support IPv6. This document specifies a mechanism for traversal of such NATs. It is based on an address mapping and on a mechanism whereby suitably upgraded hosts behind a NAT may obtain IPv6 connectivity via a stateless server, known as a SAMPLE server, operated by their Internet Service Provider. SAMPLE is an alternative to the Teredo protocol.

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 December 9, 2010.

Table of Contents

1. Introduction

At the moment, there are two traversal techniques for IPv6 users who find themselves behind Customer Premises Equipments (CPEs) which are in fact Network Address Translators (NAT) supporting only IPv4:

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.

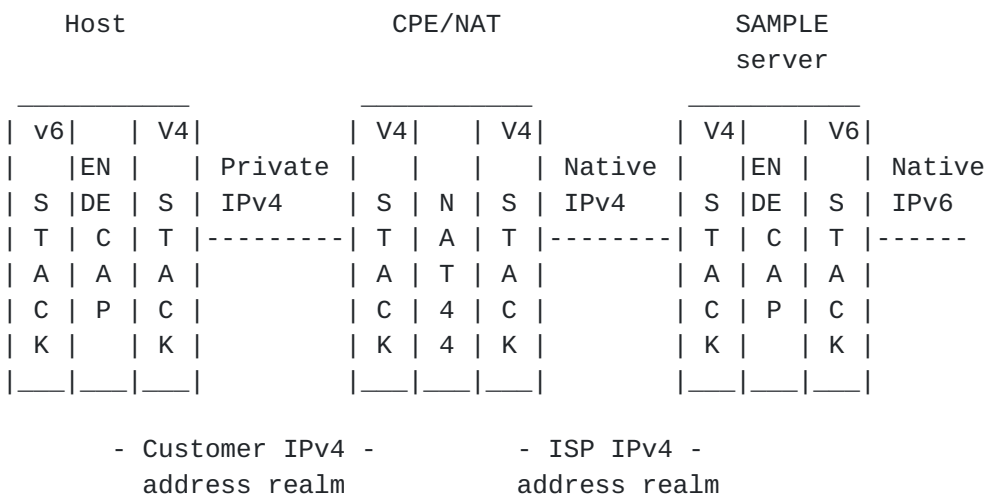
It is well established that IPv4-only CPEs are the worst product related deployment problem for IPv6 [I-D.ietf-v6ops-isp-scenarios], and it is also clear that it will be many years before such CPEs, being consumer devices sold in millions, will all be replaced. Therefore, a scaleable and reliable method for IPv6 traversal of such CPEs is desirable.

The method described here uses a subset of the stateless address mapping (SAM) mechanism proposed by [I-D.despres-softwire-sam] and elements of the Teredo method. However, it is intrinsically much simpler than Teredo, since it is designed for managed deployment by an ISP and its own clients. The authors understand that an alternative formulation of this idea, explicitly in terms of the SAM model, may also be published. The idea is also quite similar to [I-D.lee-softwire-6rd-udp] and is published in a preliminary form so that the community can evaluate the alternatives.

The method is intended for explicit adoption by an Internet Service Provider (ISP) that wishes to provide IPv6 service to customers behind IPv4-only CPE NATs, the common case today. The method is called Simple Address Mapping for Premises Legacy Equipment (SAMPLE). The ISP is required to operate a SAMPLE server and (unless operating system implementers choose to support SAMPLE directly) to provide customers with downloadable code for popular operating systems. The SAMPLE download will create a virtual IPv6 interface on top of the real IPv4 interface (just as Teredo, 6to4 and tunnel broker clients do). This is suggested to be a more practical alternative than

requiring all customers to replace their CPE. However, customers with an unsuitable operating system, or unwilling to install a download, will be advised to buy an IPv6-capable CPE.

The following figure illustrates the method symbolically:



The principle of operation is that each host that starts IPv6 communication via the SAMPLE server is assigned an IPv6 address which forms part of the ISP's regular routeable IPv6 address space. This address embeds the NAT's native IPv4 address (assigned from the ISP's IPv4 address space). It also embeds the port number assigned to the IPv6 communication stream by the NAT. Note that this applies even if the ISP is using private addressing itself; the ISP IPv4 address realm does not need to use global addresses. Needless to say, all IPv6 addresses are globally unique.

2. Normative Notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Detailed specification

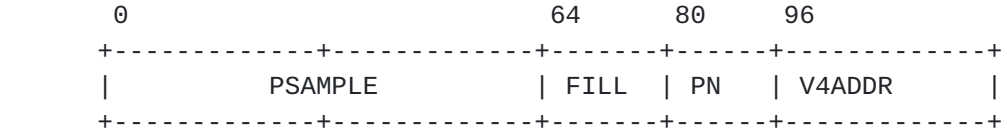
The IPv6 address assigned to a host is mapped as follows:

o Let PSAMPLE be a /64 IPv6 prefix assigned to the SAMPLE server. It may be any native IPv6 prefix chosen out of the routeable address space assigned to the ISP.

- Let V4ADDR be the /32 native IPv4 address assigned by the ISP to the CPE.
- Let PN be the external 16 bit port number assigned by the legacy CPE NAT to the host's SAMPLE interface when it first sends traffic to the SAMPLE server. (See below.)

(Note that the private IPv4 address assigned to the host behind the NAT is of no importance in the mapping.)

The IPv6 address assigned to the host is the concatenation:



There is no restriction on the 16 FILL bits except that they MUST respect [RFC4291].

IPv6 packets travelling between the host and the SAMPLE server in either direction MUST be encapsulated in UDP as described in [RFC4380]. At the host, they are decapsulated and processed by the local IPv6 stack. At the SAMPLE server, they are decapsulated and forwarded into the native IPv6 network. No state is required in the SAMPLE server; it performs blind encapsulation and decapsulation.

The SAMPLE interface in the host MUST be configured with the IPv4 address and UDP listener port number of the SAMPLE server. Apart from this, it performs blind encapsulation and decapsulation, once it has been assigned an IPv6 address.

[COMMENT: We don't want any sloppiness about reachability of the server - so an anycast address used by default seems like a really Bad Idea, judging by 6to4 experience.]

[QUESTION: Do we need a registered port number for this, or is it OK to make it configured? Or, could we re-use the Teredo listener port, 3544?]

When a host's SAMPLE interface starts up, it MUST send a Router Solicitation message to the SAMPLE server. The details are as for Teredo, except that there is no equivalent of the 'cone' bit procedure. Either the SAMPLE server will reply with a Router Advertisement within a timeout TBD, or the method will fail.

Teredo's Origin Indication mechanism is used to convey the values of PN and V4ADDR with the Router Advertisement. The SAMPLE interface can complete its own configuration upon receipt of such an RA

message.

Once an IPv6 address has been configured, the SAMPLE interface MUST send "keep alive" probes to the SAMPLE server whenever there has been no traffic through the interface for TBD seconds, in order to keep the relevant NAT state alive.

[COMMENT: Need to specify that in detail. Probably, the probe can simply be a "no next header" IPv6 packet, and the timeout will be configured to a value determined by experience.]

The SAMPLE server will act as an IPv6 router. In the simplest case, it will forward all IPv6 packets to a default route, except those whose destination address lies within the PSAMPLE prefix, which will be encapsulated and sent towards the host (CPE) and port indicated by the V4ADDR and PN values.

[QUESTION: Do we need to optimise hairpinning?]

[QUESTION: We want the server to be stateless. Is there any particular defence against DoS using bogus V4ADDR/PN values?]

4. Security Considerations

A basic assumption of SAMPLE is that it is deployed entirely within the administrative boundary of a single ISP and its customers. SAMPLE-encapsulated packets should never leave or enter that administrative boundary. Threats arising within that boundary need to be considered.

A SAMPLE server SHOULD be configured to discard (with logging if required) any incoming SAMPLE packet whose IPv4 source address does not belong to any customer of the ISP concerned. The only exception is if [RFC2827] is in use by the ISP.

[COMMENT: There is work to do here. It seems intrinsically more controlled than either 6to4 or Teredo, since the entire tunnel is confined to the ISP's IPv4 realm, but we have to look at the threats identified for those two solutions and see which apply here.

Points to consider: should the user IPv4 address be obfuscated, as in Teredo? Should some random bits be included in the FILL bits, to defeat address scanning, as in Teredo?]

5. IANA Considerations

This document requests no action by IANA.

6. Acknowledgements

This document builds on an idea extracted and simplified from [I-D.despres-softwire-sam].

Valuable comments and contributions were made by ...

This document was produced using the xml2rfc tool [RFC2629].

7. Change log

draft-carpenter-softwire-sample-00: original version, 2010-06-07

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[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.

[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.

8.2. Informative References

[I-D.despres-softwire-sam]
Despres, R., "Stateless Address Mapping (SAM) for Softwire-Lite Solutions", draft-despres-softwire-sam-00 (work in progress), March 2010.

[I-D.ietf-v6ops-isp-scenarios]
Carpenter, B. and S. Jiang, "Emerging Service Provider Scenarios for IPv6 Deployment",

draft-ietf-v6ops-isp-scenarios-00 (work in progress), April 2010.

[I-D.lee-softwire-6rd-udp] Lee, Y. and P. Kapoor, "UDP Encapsulation of 6rd", draft-lee-softwire-6rd-udp-01 (work in progress), May 2010.

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.

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

Appendix A. Main differences from Teredo

There is a critical difference from Teredo. The client address in Teredo is like this (quoting [RFC4380]):

```
+-------------+-------------+-------+------+-------------+
| Prefix      | Server IPv4 | Flags | Port | Client IPv4 |
+-------------+-------------+-------+------+-------------+
```

- Prefix: the 32-bit Teredo service prefix.
- Server IPv4: the IPv4 address of a Teredo server.
- Flags: a set of 16 bits that document type of address and NAT.
- Port: the obfuscated "mapped UDP port" of the Teredo service at the client.
- Client IPv4: the obfuscated "mapped IPv4 address" of the client.

(end quote)

Also, in Teredo, the client has to figure out which relay to use: "Teredo clients have to discover the relay that is closest to each native IPv6 or 6to4 peer. They have to perform this discovery for each native IPv6 or 6to4 peer with which they communicate."

In the SAMPLE scheme we bind IPv6 routing in both directions to the PSAMPLE /64 IPv6 locator; the client sends and receives all its IPv6 packets to and from the same SAMPLE server's IPv4 address.

This does introduce a single point of failure and a scaling bottleneck, but in exchange we get simplicity and reliability. We don't need Teredo's flag bits; any NAT that supports outbound UDP flow initiation will work.

The other major difference (and hence simplification) is that we assume that any NAT CPE is capable of straightforward port mapping for a bidirectional UDP stream, so there is no mechanism for detecting what type of NAT is in the way.

As noted above, the security threats are limited to those that can occur inside a single ISP's administrative boundary.

Authors' Addresses

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