Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 15, 2009

R. Despres
RD-IPtech

July 14, 2008

A Scalable IPv4-IPv6 Transition Architecture
Need for an address-port-borrowing-protocol (APBP)
draft-despres-v6ops-apbp-01

Status of this Memo

Abstract

This document discusses, for the IPv4-IPv6 coexistence period, the combined requirement that: (1) legacy IPv4 hosts can establish IPv4 transport connection s from customer sites having IPv6-only permanent addresses; (2) for good scalability, no network address translations (NATs), and a fortiori no application level gateways (ALGs), need to be supported within network infrastructures. To satisfy this requirement, it is concluded that an address-port-borrowing-protocol (APBP) is needed.

Table of Contents

1. Introduction

It is now well recognized that, during the transition period from IPv4 to IPv6 [RFC0791] [RFC2460], some IPv4 transport connections (TCP, UDP, etc.) will have to be established across some IPv6-only infrastructures [Bagnulo-Baker].

Various approaches have been proposed recently for a number of identified transition configurations, namely [SHANTI] [NAT64] [NATv4v6v4] [ALD] [sNAT-PT] [MNAT-PT].

SHANTI has the scalability property that no network address translator (NAT) and a fortiori no application layer gateway (ALG) is needed in internet service providers (ISP) infrastructures. It uses for this, without naming it, a type of protocol which we call here address-port-borrowing-protocol (APBP). With such a protocol, a host that has no public IPv4 address can borrow, from its ISP infrastructure, the address-port combinations it needs to establish IPv4 connections with public-IPv4 addresses.

In the client-server configuration of SHANTI, IPv6-capable client hosts, complemented to support SHANTI, establish IPv4 connections with IPv4-only servers. The complement in these hosts includes an internal IPv6-to-IPv4 NAT (with an ALG for all application protocols that need it).

This draft proposes to keep the scalability property of SHANTI (No NAT needed in ISP infrastructures), but with a scope extended to support IPv4-only hosts in sites that have no public IPv4 addresses.

It also proposes that the SHANTI complement in IPv6-capable hosts, for their support of IPv4 connections:

- be simplified, so as to not include a NAT (and a fortiori an ALG)
- be functionally more powerful, leaving no restriction on which upper layer protocols can be used (SCTP compatibility in particular)

For this, SCTP builds on concepts introduced with the Dual Stack Transition Mechanism (DSTM) which has been introduced in draft-toutain-ngtrans-dstm-00 (expired in 1999) and last documented in draft-bound-dstm-exp-04 (expired in April 2006).

A detailed description of the proposed APBP protocol is beyond the scope of this draft, but no major difficulty is expected to specify it.

The proposed architecture being new, and having not been validated on any implementation, is likely to need refinements, and possibly corrections. It is submitted for a first round of reactions.

2. A simple configuration to be supported - need for an APBP

We first consider a simple configuration among those that will need to be supported during the transition period, and analyze how the objective of no NAT in the ISP infrastructure can be satisfied Figure 1. An IPv4-only host is in a site that has an IPv6 prefix and has no public IPv4 address. It needs to establish a File Transfer Protocol connection (FTP) whith a server located somewhere on the IPv4 Internet.

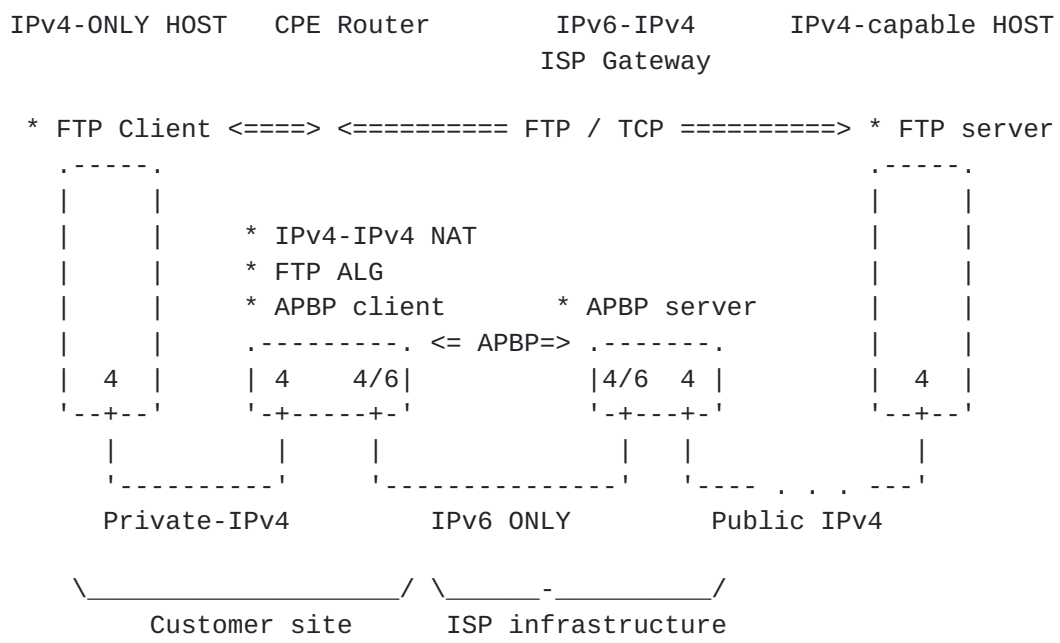
Since the client has only a PRIVATE IPv4 address, and since the server must only see a PUBLIC IPv4 address to send packets back to its clients, there must be a NAT somewhere between the two endpoints. This NAT has to include ALG for FTP.

Since we took as a requirement that no NAT be needed in the ISP infrastructure, NAT capability must be supported in the CPE router of the site. For this NAT to be able to build IPv4 packets having their public IPv4 source address, this address has to be borrowed from some ISP gateway that has interfaces to both IPv6 and public IPv4 routing domains. Between the CPE router and this ISP gateway, IPv4 packets will have to be encapsulated in IPv6 packets.

Since public IPv4 addresses have become a scarce resource, and since most customer sites need only a few of the 64K possible ports which can be associated with their IP address, a better usage of public IPv4 addresses is achievable if what is borrowed from ISP gateways is not than just an address, but rather a combination including one address an a range of ports permitted to be associated with it. Thus, the same globally unique IPv4 addresses can be used by other customer sites, for connections using different ports.

In summary, to satisfy the objective of no NAT with ALG in ISP infrastructures in the configuration of Figure 1, an address-port-borrowing-protocol (APBP) is needed.

With it, CPE routers act as APBP clients. They borrow address-port combinations from ISP gateways that act as APBP servers. Between APBP clients and APBP servers, IPv4 packets are encapsulated in IPv6 packets.



A SIMPLE CONFIGURATION TO BE SUPPORTED
(APBP = address-port-borrowing-protocol)

Figure 1

Note that the reason why the client host is IPv4-only for the considered connection can be one of the following:

- The APPLICATION is IPv4-only, and the host has no public IPv4 address (the protocol stack can be IPv4-only or dual stack).

- The PROTOCOL STACK is IPv4-only (the application can be IPv4-ony or IPv4 and IPv6).

3. Other supported transport connections

All other transport protocols than TCP and all other application-level protocols than FTP that are supported in the IPv4-IPv4 NAT of the CPE router are possible on the physical configuration of Figure 1.

Transport connections in the reverse direction are also possible, if the NAT has assigned some address-port combinations for them, typically by means of one of the existing protocols existing for this (STUN, UPnP, NAT-PMP, etc.).

More generally, any protocol combination that works today across a IPv4-IPv4 NAT of a given type will also work, with the same type of IPv4-IPv4 NAT in any site having no public IPv4 address provided that:

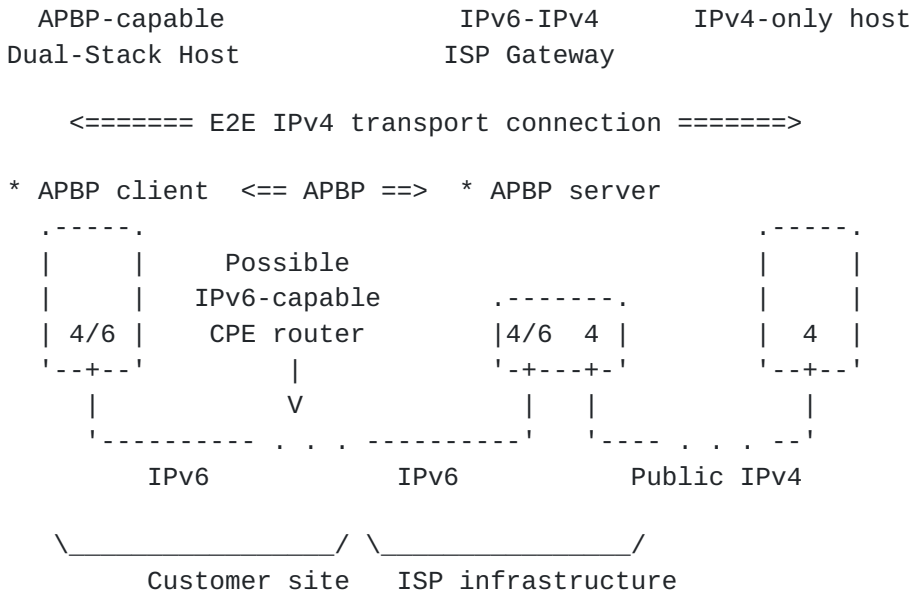
1. The site has an IPv6 prefix.

2. The local ISP supports APBP servers.

3. This IPv4-IPv4 NAT is complemented with an APBP-client function.

4. Other supported configurations

If the support of the APBP-client function is added to a dual-stack host, and if this host is attached to an ISP infrastructure where APBP is supported, this host can establish pure IPv4 end-to-end transport connections with IPv4-only remote hosts (Figure 2).



APBP BETWEEN ABPB-CAPABLE DUAL-STACK HOSTS AND ISP GATEWAYS

Figure 2

No NAT being needed on the end-to-end path, packets keep their IPv4

source and destination IPv4 addresses and ports unchanged from source to destination. All IPv4 capable applications that need public-IPv4 addresses at both ends can work across the IPv6 domain of this configuration.

In this configuration, applications that need to dynamically advertise their address-port combinations, e.g. for callbacks or referrals, obtain these combinations at their socket programming interface as usual.

With Windows Sockets, for example, local address-port combinations are obtained by applications as results of getsockname() function calls. Client applications make these calls after having made connect() function calls. Server applications call make them after their bind() function calls with INADDR_ANY as local IPv4 addresses.

5. Some protocol considerations

Since some server applications check that several related transport connections initiated by a same client do come from the same IP addresses, APBP clients should borrow only one public IP address, and manage a set of ports to be used with this address.

For simplicity, and to minimize interactions between APBP clients and servers, these ports should be obtained in significant quantity at each request. This quantity necessarily depends on policies of ISPs that operate APBP servers. Typically, a maximum quantity per customer site would help preventing denial of service attacks. Provision could be made in the protocol for APBP clients to increase and decrease from time to time the number of their borrowed ports according to fluctuations of their needs.

In APBP clients, port management may be optimized so as to increase the number of transport connections that are possible with a given number of ports. In particular, port reuse based on endpoint dependent mappings may be envisaged for ports that are assigned to outgoing transport connections known not to need endpoint independent mapping (DNS, HTTP, SMTP, POP3, NNTP, Telnet, SNMP, etc.).

To avoid that APBP client failures would cause indefinite port reservations in APBP servers, some keep-alive mechanism should be part of the protocol. Also, APBP clients should explicitly release their borrowed address-ports when they have been using none for a significant time.

For scalability of the APBP server function, the destination address used by APBP clients to request address-ports sets when they have

none yet,would advantageously be an anycast address. Responses to these requests should then indicate unicast addresses of particular APBP server instances that have responded (further exchanges of encapsulated IPv4 packets must be with these particular instances, where the particular address-ports that have been obtained are managed).

6. Security considerations

If a third party would be able to act as an ISP APBP server, it would be able to intercept all the end-to-end traffic that uses an public IPv4 address borrowed from it.

This can however be made impossible if, in infrastructures of ISPs that support APBP, precautions are taken so that addresses of APBP servers cannot be counterfeited from customer sites.

7. IANA Considerations

The protocol has still to be specified in details, but it can be expected that a UDP port number will be needed for the supervisory part of APBP.

Also, having a standardized anycast address to reach APBP servers would be simpler than depending on an ISP dependent parameter, plus maybe a DHCPv6 option to advertise it.

8. Acknowledgements

This draft is in continuity with previous works that influenced it, or at least anticipated some of its contents. In particular, the work of Laurent Toutain and his colleagues on DSTM had a significant influence. Myung-Ki Shin's work on the port option of DSTM in 2002 was unknown by the author, but anticipated the idea of address-port combination borrowed by IPv6-capable hosts. Brian Carpenter was first, as far as the author knows, to post a draft where address-port combinations were borrowed directly from ISP gateways.

9. References

9.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

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

9.2. Informative References

[ALD] "draft-woodyatt-ald-02 (work in progress)", December 2007.

[Bagnulo-Baker] "draft-bagnulo-v6ops-6man-nat64-pb-statement-01 (work in progress)", February 2008.

[DSTM] "Dual Stack Transition Mechanism - http://www.ipv6.rennes.enst-bretagne.fr/dstm/".

[MNAT-PT] "draft-v6ops-van-beijnum-mnat-pt-00 (work in progress)", February 2008.

[NAT64] "draft-bagnulo-v6ops-6man-nat64-pb-statement-00 (work in progress)", November 2007.

[NATv4v6v4] "draft-durand-v6ops-natv4v6v4-00 (work in progress)", November 2007.

[RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3489] Rosenberg, J., Weinberger, J., Huitema, C., and R. Mahy, "STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)", RFC 3489, March 2003.

[SHANTI] "draft-carpenter-shanti-01.txt (work in progress)", November 2007.

[sNAT-PT] "draft-miyata-v6ops-snatpt-00 (work in progress)", February 2008.

Author's Address

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

Email: remi.despres@free.fr

Full Copyright Statement

Intellectual Property