

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: April 21, 2011

R. Despres
RD-IPtech
October 18, 2010

IPv4 Residual Deployment across IPv6-Service networks (4rd)
A NAT-less solution
draft-despres-softwire-4rd-00

Abstract

During the long transition period from IPv4-only to IPv6-only, networks will have not only to deploy the IPv6 service but also to maintain some IPv4 connectivity for a number of customers, and this for both outgoing and incoming connections and for both customer-individual and shared IPv4 addresses. The 4rd solution (IPv4 Residual Deployment) is designed as a lightweight solution for this. It applies not only to ISPs have IPv6-only routing networks, but also to those that, during early transition stages, have IPv4-only routing, with 6rd to offer the IPv6 service, those that have dual-stack routing networks but with private IPv4 addresses assigned to customers.

In some scenarios, 4rd can dispense ISPs from supporting any NAT in their infrastructures. In some others it can be used in parallel with NAT-based solutions such as DS-lite and/or NAT64/DNS4 which achieve better IPv4-address sharing ratios (but at a price of significantly higher operational complexity).

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 21, 2011.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Definitions	4
3. Applicability	5
4. The 4rd Protocol Specification	8
4.1. Mapping Rules	8
4.2. Packet Encapsulations/Decapsulations	9
4.3. Port sets of IPv4r prefixes longer than /32	10
4.4. PMTU Considerations	12
4.5. Parameter Acquisitions by 4rd Clients	12
5. Example with IPv6-only Routing and Shared IPv4 Addresses	14
6. Security considerations	16
7. IANA Considerations	17
8. Acknowledgments	17
9. References	17
9.1. Normative References	17
9.2. Informative References	18
Author's Address	18

1. Introduction

During the long transition period from only IPv4 to IPv6, networks of Internet Service providers (ISPs) will have not only to offer IPv6 connectivity but also, for some customers, to maintain a residual IPv4 connectivity. Both outgoing and incoming connections will have to be supported. While some privileged customers will still have individual IPv4 addresses of their own, more and more others will only have shared IPv4 addresses.

All ISP routing networks will eventually be IPv6-only but, in earlier phases, some deployments of the IPv6 service can be done on ISP routing networks that only route private IPv4 of [RFC1918], the IPv6 service being offered by means of 6rd. Some others will route both IPv6 and private IPv4.

4rd is a solution for the residual support of global IPv4 connectivity across routing networks that are IPv6-only, private-IPv4-only, or IPv6-and-private-IPv4.

Depending on ISP constraints and policies, 4rd can be used across IPv6-only networks either alone, no NAT being then needed in ISP infrastructures, or in parallel with NAT based solutions that, at a price of more operational complexity, achieve better address sharing ratios such as [DS-lite] and [NAT64]/[DNS64].

This proposal is a more detailed version of what was initially described in section 3.2 of the more general Stateless Address Mapping proposal of [1]) (SAM).

At the time of writing, 4 ISPs in Japan have expressed interest for the SAM/4rd solution to offer IPv4 connectivity across IPv6-only routing networks (www.ietf.org/mail-archive/web/v6ops/current/msg05247).

2. Definitions

Locator: in a given address family, an address or a routable prefix.

IPv4r Address Family: the "residual IPv4" address family, that of IPv4r locators.

IPv4r Address: Either a global IPv4 address or the combination of a global IPv4 address and a port (an A+P address)

IPv4r prefix: Either a global IPv4 prefix (up to /32), or a global IPv4 address followed by a port-set identifier whose length is from 1 to 15.

IPv4p Address Family: That of a private address spaces of (10/8, 172.16/12, or 192.16/16, prefixes).

interior address family: in a tunnel-supporting network, the address family of encapsulating packets (in 4rd, IPv6 or IPv4p).

exterior address family: in a tunnel-supporting network, the address family of encapsulated packets (in 4rd, IPv4r).

4rd parent network: For a given 4rd network, the network that assigns to it one or several IPv4r prefixes.

4rd network: A network whose interior address family is different from global IPv4, and that supports one or several 4rd servers at its border with its 4rd parent network.

4rd server (4rd-S): A function at a border point between a 4rd network and its 4rd parent network. Via automatic tunnels, it statically shares among customers of the 4rd network IPv4r locators that have been received from the parent network.

4rd client (4rd-C): A function that obtains mapping rules from a 4rd server, derives from them its own IPv4r locator, and tunnels IPv4r packets across its 4rd network.

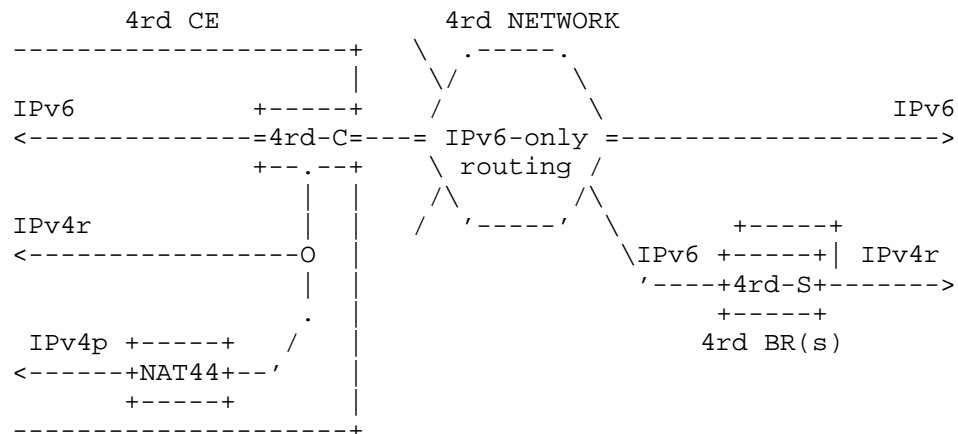
4rd BR: A router that supports one or several 4rd servers (Border Router).

4rd CE: A node supports a 4rd client and is in a customer position on a 4rd network. It may be a host, a router, or both.

network PMTU: For an identified address family, the packet size that must not be exceed to traverse the network without risk of packets being discarded (in IPv6) or fragmented (in IPv4).

3. Applicability

For 4rd to actually be used across a network, the network must be a 4rd network, and must have at least one 4rd CE.



4rd ACROSS AN IPv6-ONLY ROUTING NETWORK

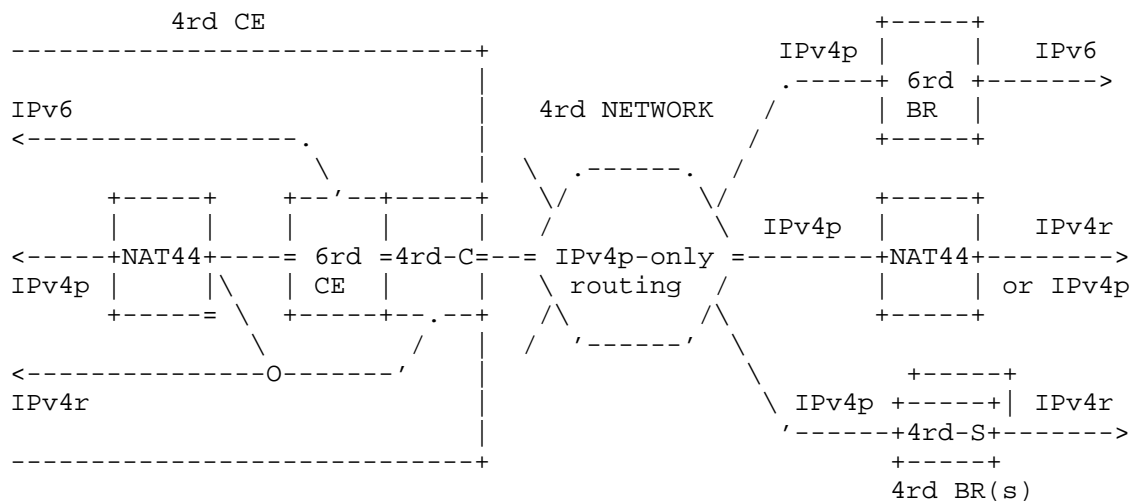
Figure 1

If the interior address family is IPv4p, the operator of must know the PMTU of its 4rd network.

Figure 1 shows a scenario where the interior address family is IPv6. In the CE, the IPv4r interface of the 4rd client can be used to provide global IPv4 addresses and reserved ports to a socket API and/or to a NAT44. This NAT can use them for its port-forwarding function, be it configured administratively or by means of UPnP or NAT-PMP. If both a socket API and a NAT44 share the set of available addresses and ports, a static switch can do split.

This scenario doesn't exclude other ways to offer IPv4 connectivity across the same IPv6-only routing network (typically DS-lite and/or NAT64/DNS64). Note however that, with each IPv4 address shared between 16 customers, each customer obtains with 4rd 3840 global-IPv4 ports (in addition to its 65 536 ports per IPv6 address), and the available IPv4 address space is multiplied by 16. Since most port-consuming applications should quickly be reachable in IPv6 (Google Maps in particular is already in this case) this should be largely sufficient in many scenarios.

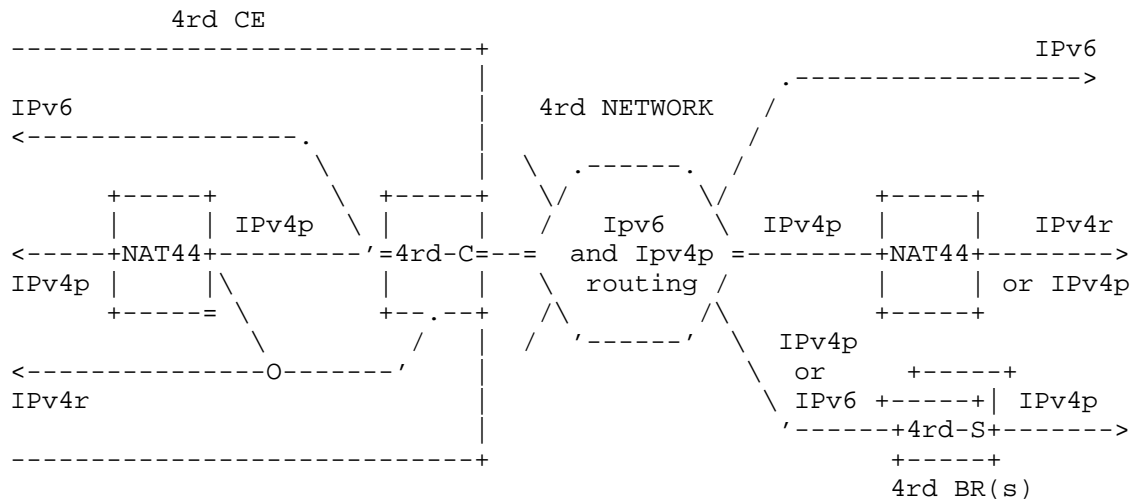
Figure 2 shows a scenario where the interior address family is IPv4p and where the IPv6 service is supported with 6rd. The 4rd CE architecture is similar to that of the previous example with two differences: IPv6, instead to be directly available at the network interface, is obtained by means of a 6rd-CE function; the NAT44, if present, can use as external addresses not only those of its IPv4r locator but also the IPv4 address assigned to the CE in the 4rd network. How the NAT44 uses this external address set is an implementation matter, but it can be noted that applications that are known to traverse cascades of NATs without problem (Web, DNS, and Mail, in particular) can use IPv4p addresses. IPv4r addresses are thus kept for IPv4 connections that may need end-to-end transparency.



4rd ACROSS AN IPv4p-ONLY ROUTING NETWORK

Figure 2

Figure 3 shows a scenario where both IPv6 and IPv4p are routed. The main difference with the IPv4p-only routing case is that 6rd is not needed. Tunnels for IPv4r packets can use IPv6 or IPv4p depending on local policies.



4rd ACROSS A DUAL-STACK ROUTING NETWORK

Figure 3

NOTE: The above scenarios can apply not only to 4rd networks operated to ISPs but also to private networks. A CPE that supports a 4rd server can, when it has an IPv4r locator, share it among hosts of its site that support 4rd clients. This is in practice a static alternative to UPnP and NAT-PMP for hosts to still have some IPv4 incoming connectivity.

4. The 4rd Protocol Specification

4.1. Mapping Rules

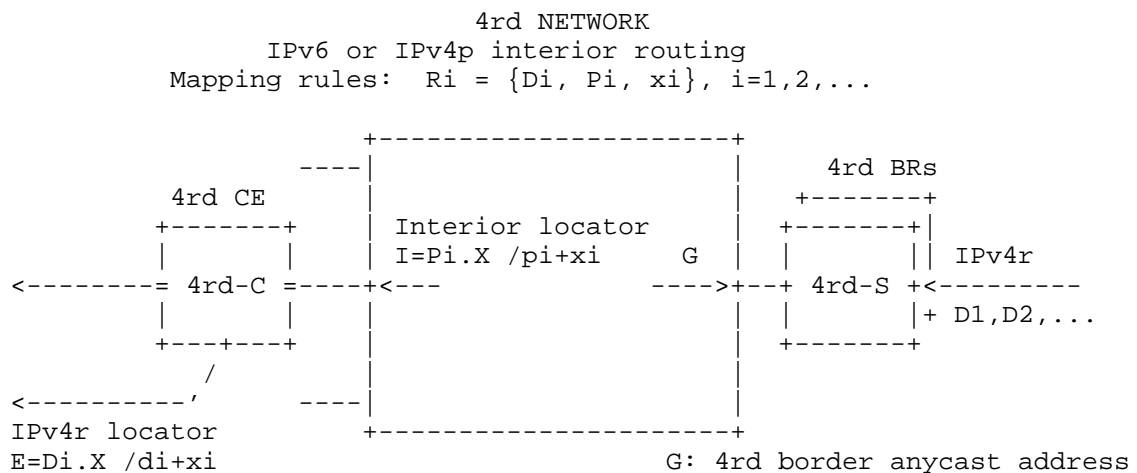
4rd mapping rules establish 1:1 mappings between interior and exterior locators. Each rule R_i comprises:

D_i : the "rule exterior prefix"

P_i : the "rule interior prefix"

x_i : the "index length", i.e. the length of the field X that, for a given 4rd client is common to its interior and exterior locators.

D_i 's of all rules of a 4rd network must be non overlapping prefixes, and the same for P_i 's.



4rd LOCATOR MAPPING RULES

Figure 4

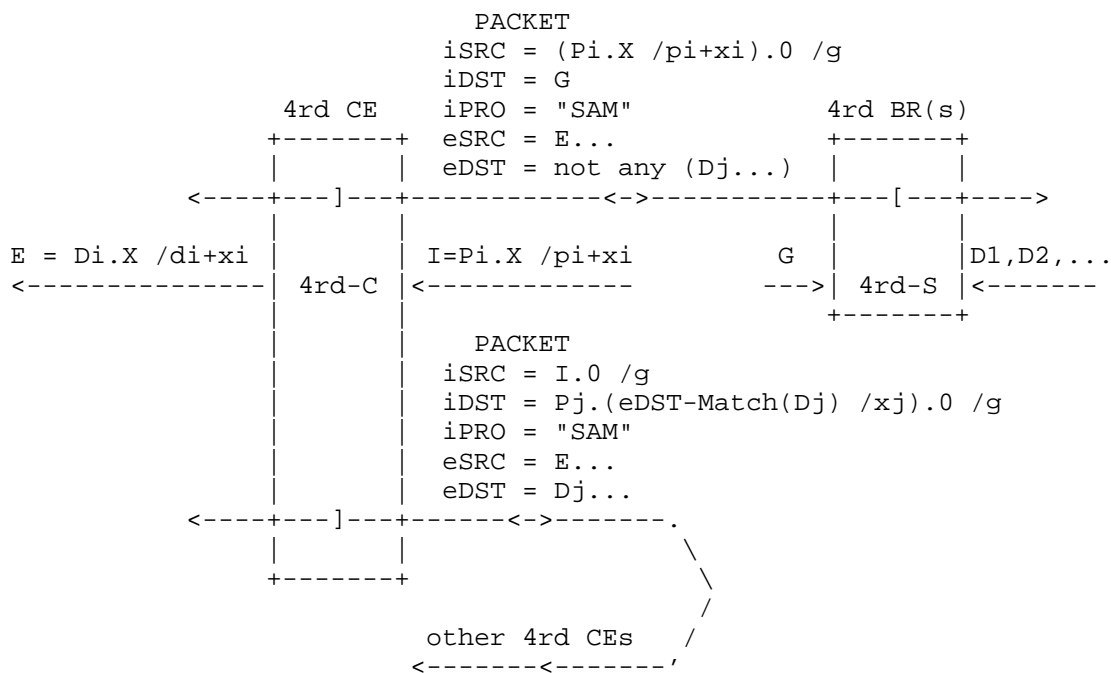
Figure 4 shows how the exterior locator "E" of a 4rd client is derived from its interior locator "I". E comprises the D_i of the rule whose P_i is recognized at the beginning of I, followed by index X whose length is the x_i of the rule, and which is copied from the I after the its P_i . In this document field acronyms are uppercase, and lengths of fields are the same letters in lower case. (Thus,

"/pi.xi" represents a locator length that is the sum of Pi's length and xi).

To derive an interior address from an exterior address, the reverse logic is used. In this document, Y... represents any address that starts with prefix Y. The interior locator I derived from exterior address E... then comprises the Pi of the rule whose Di matches the beginning of E..., followed by index X whose length is the xi of the rule and which is copied from E... after its bits used to match Di. If the obtained I is shorter than a complete interior address, it is completed with zeroes. If no rule applies (no Di found in E), the interior locator is the 4rd-server interior address G (an anycast address).

4.2. Packet Encapsulations/Decapsulations

Mapping rules: $R_i = \{D_i, P_i, x_i\}, i=1,2,\dots$



4rd PACKET ENCAPSULATIONS AND ADDRESS MAPPINGS

Figure 5

When a 4rd client or server receives a packet at its IPv4r interface (a pseudo interface in the client case), it checks the validity of its source and destination addresses. It also checks that the packet size is acceptable (see Section 4.4). If yes, it encapsulates it in an interior packet and forwards it via its interior interface.

The Next-header field, if interior addresses are IPv6, of the Protocol field if they are IPv4p, a value to be assigned by IANA for 4rd and for other applications of the SAM of [1] (SAM). A specific value for SAM is preferred to a re-use of Protocol 41, used for IP-in-IP encapsulations of 6to4, ISATAP, and 6rd, because this ensures that coexistence with these without risk of incompatibility.

Symmetrically, a 4rd client or 4rd server that receives a packet at its interior interface checks the validity of source and destination addresses in both its encapsulating and encapsulated packets. It also checks that they are mutually consistent with mapping rules of the 4rd network. If yes it decapsulates the IPv4r packet contained in the encapsulating packet, and forwards it its IPv6 interface.

Details on which addresses are acceptable in which packets are detailed in Figure 5, where SRC and DST respectively mean source and destination, PRO means protocol, where iXXX and eXXX respectively refer to interior and exterior address families.

4.3. Port sets of IPv4r prefixes longer than /32

The port-set identifiers *S* of an IPv4r prefix of length *s* in the range 33 to 47 consists in the *s*-32 bits beyond the first 32. The port set it identifies is specified with the following constraints:

"Exclusiveness" Port sets of two *S*'s must be disjoint if the *S*'s are non overlapping prefixes (10 and 1011 do overlap while 10 and 1110 don't)

"No administration" The port set of *S* must be algorithmically derived from *S* without depending on any parameter.

"Fairness-1" Port sets of two *S*'s of same lengths must contain the same number of ports.

"Fairness-2" No port-set may contain any port 0 to 4095 (these have more value than others in OS's, and are normally not used in dynamic port assignments to applications).

4.4. PMTU Considerations

To properly deal with large size IPv4 datagrams that are fragmented before entering a 4rd network, precautions have to be taken because:

- o In IPv4, intermediate nodes may have to forward packets that are longer than the MTU of next links to be traversed. For this, they fragment packets within the network.
- o In IPv6, such packets are discarded, with ICMP Packet Too Big ICMPv6 error packets returned to sources, but with all IPv6 links having to support MTUs of at least 1280 octets.

To cope with these constraints, 4rd clients and 4rd servers can reassemble multi-fragment IPv4 datagrams before processing them. (This function is stateful at the IP layer like the same function in NATs. But at the transport layer, 4rd remains stateless whereas NATs are stateful, a source of operational complexity that is avoided with 4rd.)

Each datagram, after fragment reassembly if needed, is forwarded either in a single packet, if with its encapsulation header it fits in the network PMTU, or in as many packets as needed for each one to fit in this PMTU. Optimized treatments are possible, whereby first parts of datagrams are forwarded without waiting for complete datagram reassembly, but this is an implementation matter that doesn't belong to the scope of this specification.)

4.5. Parameter Acquisitions by 4rd Clients

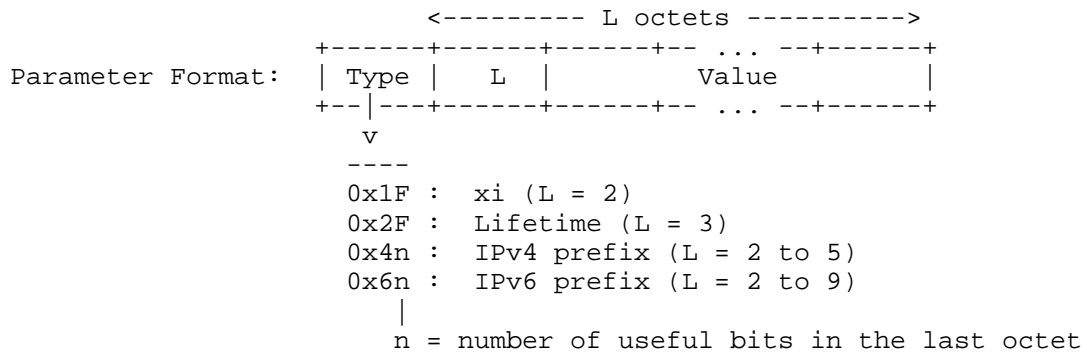
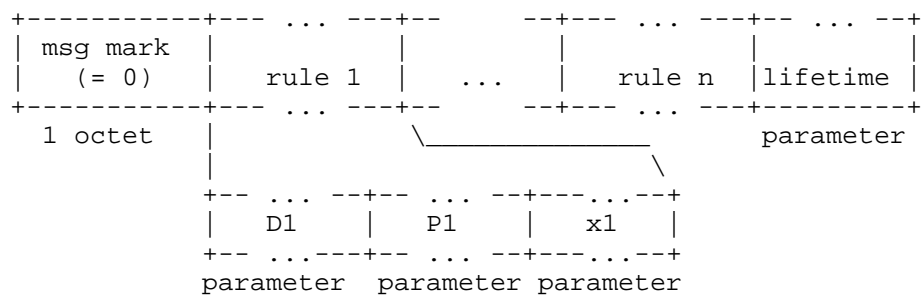
The 4rd-server address G may be obtained in various ways. It may be administratively configured (typically applicable if the 4rd network operator provides its own 4rd CEs). It can also be obtained in DHCP [RFC2131], DHCPv6 [RFC3315], Radius [RFC2865], or Diameter [RFC3588]. For these, IANA assigned numbers for 4rd remain to be chosen. In absence of all these means, G can be taken as the well-known address of SAM servers in the applicable interior address family (also to be assigned by IANA).

If a 4rd client has Gs for both IPv6 and IPv4p, it may try both and settle for either one from which it obtains responses.

To obtain its mapping rules and their common lifetime, a 4rd client sends a 4rd "Parameter Request" message to the 4rd-server anycast address G. It retransmits it until it obtains an answer, typically with longer time intervals after several unsuccessful attempts. When it receives a 4rd "Parameter Indication" message with the 4rd-server anycast address as source, it derives from the contained mapping

rules its own IPv4r locator. It also stores these rules for its future packet encapsulations/decapsulations.

4rd messages are transmitted in payloads of 4rd interior packets at the same place as encapsulated exterior packets. Their first octet is set to 0, a "Message Mark" which permits to distinguish 4rd messages from encapsulated packets (IPv4 packet headers all start with a 4 in the first 4 bits).



FORMAT OF 4rd PARAMETER INDICATIONS

Figure 7

A 4rd Parameter Request is sent with no information after the 4rd Message Mark. In order to facilitate future extensions that may prove useful, 4rd servers should ignore octets that may be received after this mark.

The following example illustrates the case of an ISP that operates an IPv6-only routing network and assigns shared global IPv4 addresses to its customers. The ISP has 2^{24} customers whose /48 prefixes start with a common prefix K/24. In IPv4, it has three global IPv4 prefixes, R1/13, R2/14, R3/14, giving a total of 2^{20} addresses. Each of these addresses must therefore be shared among 16 customers. Exterior locators E must therefore be /36s, comprising port-set identifiers S having 4 bits (each customer is thus assigned $2^{12} \times 15 / 6 = 3840$ reserved ports in global IPv4). Each interior prefix I/48 must then be composed of the common prefix K followed by the short identifier Ci of one of the three Di's. Their lengths have to be related to lengths of Di's by the formula $ci=(i-k)-(e-di)$, which gives $c1=1$, $c2=2$, and $c3=2$. Within these constraints, bit values of the Ci's may be arbitrary non overlapping prefixes, e.g. $C1 = 0b0$, $C2=0b10$, $C3 = 0b11$ (with 0bXXX being the binary number XXX). Rule are {D1/

VARIANTS:

- o It the ISP would have preferred to have only one rule, this would have been possible by using in IPv4 only the /13. Then port-set identifiers S would have had 5 bits, and each customer would have had 1920 ports in global IPv4.
- o If instead of one K/24, the ISP there would have had to use two different prefixes, K1/25 and K2/25, mapping rules could have been {D1/13, P1=K1/25, x1=23}, {D2/14, P2=K2.C2/26, x2=22}, and {D3/14, P3=K2.C2/26, x3=22}, with C2=0b0 and C3=0b1.
- o If, in a more complex scenario, the ratio between number of customers and number of IPv4 addresses would not have been a power of two, either some interior addresses or some exterior addresses would have had to be sacrificed (not assigned). For example, with K1/25, K2/26, and D1/14, D2/15, D3/15, D4/15, giving $2^{23} + 2^{22}$ customers and $2^{19} + 2^{15}$ IPv4 addresses, rules could have been {D1/14, P1=K1.C1/26, x1=22}, {D2/15, P2=K1.C2/27, x2=21}, {D3/15, P3=K1.C3/27, x3=21}, {D4/15, P4=K2.C4, x4=21}, with C1=0b0, C2=0b10, C3=0b11, C4=0.

6. Security considerations

Spoofing attacks

With address-consistency checks of Section 4.2, authentication verifications that apply interior locators also apply, indirectly, to exterior locators. Similarly, anti-spoofing protections that apply to interior addresses also apply, indirectly, to exterior locators. 4rd should therefore introduce no opportunity of its own for spoofing attacks.

Denial-of-service attacks

Reassembly of fragmented exterior datagrams introduces an opportunity for some form of DOS attacks, shared with NAT-based solutions. Note that this risk among reason to prefer native IPv6 to native IPv4 when there is the choice for a transport connection.

Risks of DOS attacks at the transport-connection layer, to which NAT-based solutions are exposed, are avoided in 4rd because of its the stateless operation of this layer.

Faked 4rd servers

If a 4rd CE uses as 4rd server address one of the two IANA assigned well-known address for this in IPv6 and IPv4, and if its ISP network has no 4rd server, packets addressed to it can be forwarded to the Internet backbone. They should however not reach any faked 4rd server because, this address starting with none of prefixes routed to other ISP networks, they will normally be discarded in the backbone. However, whether some additional protection in would be appropriate against fake 4rd servers (e.g. with a nonce in Parameter Requests and Parameter Indications), is still viewed as an open issue.

Routing-loop attacks

Routing-loop attacks that may exist in some automatic-tunneling scenarios are documented in [3]. They cannot exist with 4rd because its address checks of Section 4.2 prevent multiple traversals of a 4rd network by the same IPv4r packet, and because, 4rd using its own Protocol number, routing-loops between nodes of nodes working with two different tunnel protocols are also impossible.

7. IANA Considerations

This specification depends on the following number assignments by IANA:

- o The SAM protocol number (Section 4.2)
- o The DHCP and DHCPv6 4rd option codes (Section 4.5)
- o The Radius 4rd attribute type (Section 4.5)
- o The SAM-server well-known addresses, in IPv4 and IPv6 (Section 4.5)

8. Acknowledgments

The author has benefited from useful informal discussions with a number of IETF participants on previous SAM proposals, from which this specification is a by-product. Concerning 4rd in particular, Satoru Matsushima deserves special recognition, first for the interest in the approach he expressed from the beginning, but also for his constructive contributions, including his proposal of the 4rd acronym, and for convincing his colleagues to make actual deployment plans with this technology. Olivier Vautrin, by independently proposing the same acronym for a similar orientation, has to be thanked for the valuable encouragement this has been.

9. References

9.1. Normative References

- [RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.
- [RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, March 1997.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.
- [RFC2865] Rigney, C., Willens, S., Rubens, A., and W. Simpson, "Remote Authentication Dial In User Service (RADIUS)", RFC 2865, June 2000.
- [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.

- [RFC3513] Hinden, R. and S. Deering, "Internet Protocol Version 6 (IPv6) Addressing Architecture", RFC 3513, April 2003.
- [RFC3588] Calhoun, P., Loughney, J., Guttman, E., Zorn, G., and J. Arkko, "Diameter Base Protocol", RFC 3588, September 2003.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

9.2. Informative References

- [1] Despres, R., "Stateless Address Mapping (SAM) - a Simplified Mesh-Software Model - draft-despres-software-sam-01 - work in progress", July 2010.
- [2] Vautrin, O., "IPv4 Rapid Deployment on IPv6 Infrastructures (4rd) - draft-vautrin-software-4rd-00 - work in progress", July 2010.
- [3] Nakibly, G. and F. Templin, "Routing Loop Attack using IPv6 Automatic Tunnels: Problem Statement and Proposed Mitigations - draft-ietf-v6ops-tunnel-loops-00 - Work in progress", September 2010.
- [DNS64] Bagnulo, M., Sullivan, A., Matthews, P., and I. van Beijnum, "DNS64: DNS extensions for Network Address Translation from IPv6 Clients to IPv4 Servers [draft-ietf-behave-dns64 - work in progress]", October 2010.
- [DS-lite] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion [draft-ietf-software-dual-stack-lite - work in progress]", August 2010.
- [NAT64] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 - Clients to IPv4 Servers [draft-ietf-behave-v6v4-xlate-stateful - work in progress]", July 2010.

Author's Address

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

Email: remi.despres@free.fr

