Internet Engineering Task Force	A. Durand
Internet-Draft	Comcast
Intended status: Informational	February 24, 2008
Expires: August 27, 2008	

TOC

Distributed NAT for broadband deployments post IPv4 exhaustion
draft-durand-v6ops-natv4v6v4-01

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/ietf/1id-abstracts.txt.
The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.
This Internet-Draft will expire on August 27, 2008.

Abstract

The common thinking for the last 10+ years has been to say that dual stack was the answer to IPv6 transition and that most things would be converted to dual stack way before we ran out of IPv4. Well, it has not happened. We are going to run out of IPv4 addresses soon, way before any significant IPv6 deployment will have occured. However, the quasi totality of the Internet and most of the computers in the home are still IPv4-only. Several distributed NAT architectures, based on different possible flavors of a carrier-grade NAT, are presented as solutions to maintain some form of connectivity between those home environments and the legacy Internet.

Table of Contents

---

1. Introduction

TOC

This memo will present a service provider view on deployments post IPv4 exhaustion and some of the necessary technologies to achieve it.

1.1. Requirements Language

TOC

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

TOC

2. IPv4 exhaustion coming sooner than expected

Global public IPv4 addresses coming from the IANA free pool are running out faster than predicted a few years ago. The current model shows that exhaustion could happen as early as 2010. See http://ipv4.potaroo.net for more details. Those projection are based on the assumption that tomorrow is going to be very similar to today, ie looking at recent address consumption figures is a good indicator of future consumption patterns. This of course, does not take into account any new large scale deployment of IP technology or any human reaction when facing an upcoming shortage.
The prediction of the exact date of exhaustion of the IANA free pool is outside the scope of this document, however one conclusion must be drawn from that study: there will be in the near future a point where new global public IPv4 addresses will not be available and thus any new broadband deployment may have to consider the option of not provisioning any (global) IPv4 addresses to the WAN facing interface of edge devices. The classic IPv6 deployment model known as "dual stack" can be a non starter in such environments.

---

3. Handling the legacy

TOC

3.1. Legacy edges of the Internet for broadband customers

TOC

Broadband customers have a mix and match of IP enable devices at home. The most recent operating systems (eg Windows Vista or MacOS-X) can operate in an IPv6-only environment, however most of the legacy one can't. It has been reported, for example, that windows XP cannot process DNS requests over IPv6 transport. Expecting broadband customers to massively upgrade their software (and in most cases the corresponding hardware) to deploy IPv6 is a very tall order.

3.2. Content and Services available on the Internet

TOC

IPv6 deployment has been very long to take off, so the current situation is that almost none of the content and services available on the Internet are accessible over IPv6. This will probably change in the future, but for now, one has to make the assumption that most of the traffic generated by (and to) broadband customers will be sent to (and originated by) IPv4 nodes.

3.3. Burden on service providers

TOC

As a conclusion, broadband service providers may be faced with the situation where they have IPv4 customers willing to communicate with IPv4 servers on the Internet but may not have any IPv4 addresses left to assign to them...

4. Solution space

TOC

A number of solutions can be studied: IPv6-only, double IPv4>IPv4->IPv4 NAT, double IPv4->IPv6->IPv4 NAT, and IPv4 over IPv6 tunneling plus carrier grade IPv4->IPv4 NAT. All of them are essentially a variation on the theme of a distributed NAT where instead of provisioning each broadband customer with a unique global IPv4 address, global IPv4 addresses are share among broadband customers.

4.1. IPv6-only

TOC

The first solution that comes to mind is to simply provision new broadband customers with only IPv6 addresses. However, two immediate issues come to mind:

a. Legacy devices in the customer home will not be able to communicate with the outside.

b. New IPv6-only capable devices will not be able to communicate with legacy IPv4-only servers in the Internet.

4.2. Double IPv4->IPv4->IPv4 NAT

TOC

This solution consists of provisioning broadband customers with a private [RFC1918] address on the WAN side of the home gateway, and then translate this private IPv4 address somewhere within the service provider network by a carrier grade NAT into a global IPv4 address. Devices behind the home gateway will then be translated twice, once by the home gateway itself, and another time by the NAT within the service provider.

This solution has the advantage of being simple to understand and is the easiest to deploy in the home. It has however a number of drawbacks.
The first drawback is that some applications may have a more difficult time going through the two levels of NAT. Application relying on port mapping or port opening using UPnP may not work as expected as the carrier grade NAT may not allow those NAT traversal techniques. Note that this drawback is not specific to this solution, it is tied to the presence of a carrier-grade NAT in the architecture.
Another drawback is that this solution limits the number of customer within an access network to the size of net 10, ie somewhere between 10 and 16 million depending on address efficiency. Note that very large networks such as Comcast have already ran out of RFC1918 space a few years ago. A possible way to get around this problem is for the service provider to run several instances of net 10, one per "regional area". However, there are serious operational issues with this, especially if the service provider is running a unified backbone and a unified set of services.
A third drawback of this solution is that it can potentially create a conflict on the home gateway if the same variant of RFC1918 space is used on the WAN port and the LAN port. For example, if both the WAN port and the LAN port are configured with 10.0.0.1/24, some NAT implementations may get confused.

4.3. Double IPv4->IPv6->IPv4 NAT

TOC

When private address space is running out and/or the service provider does not want to run multiple copies of net 10, the next step is to to provision the home gateway only with an IPv6 address and associated prefix and let that home gateway translate internal RFC1918 space into global IPv6 addresses. However, as the final destination may not be configured to accept IPv6 connections, those packets will have to be translated a second time into IPv4 packets.
The first translation hapening in the home gateway can be very straightforward and in most cases stateless. This consists in header swapping and embedding the source & destination IPv4 addresses within source & destination IPv6 addresses. The prefix used to embed the source address can be any sub-prefix of the one delegated to the home gateway. The prefix used to embed the destination address is used to route the IPv6 packets to the local farm of IPv6->IPv4 translator within the service provider network. The discovery of that second prefix by the home gateway can be achive in many ways, for example through a DHCPv6 option yet to be defined.
The second translation will have to occur within the service provider network in a carrier-grade IPv6->IPv4 NAT. This translation is a

traditional NAT that requires keeping track of IP addresses and port numbers allocated.
The implications of this second level of translation are very similar to those in the model above of a double IPv4->IPv4->IPv4 translation. There will be a need for a farm of translators within the service provider network operating at line speed. Some applications may have a harder time working through the carrier-grade NAT. On top of that, some MTU adaptation will have to take place to accommodate for the longer IPv6 header.
Another issue with this approach is the role of ALGs. Although IPv4->IPv4 ALGs are now fairly well understood, there is little experience with IPv4->IPv6 or IPv6->IPv4 ALGs. One of the questions raised is, should the first home NAT, translating from IPv4 to IPv6, also use IPv4->IPv6 ALGs to translate the payload addresses to IPv6, or should it leave them in IPv4 format, knowing that the carrier-grade NAT will anyway translate them back to IPv4?

4.4. IPv6 Tunneling plus carrier-grade IPv4->IPv4 NAT

TOC

When IPv6-only connectivity is offered to the customer, one can look at IPv4 over IPv6 tunnels to re-establish connectivity for the legacy IPv4 hosts. The Softwire hub and spoke solution, based on L2TP tunnels could be the perfect candidate in that space.
The caveat is that this technique alone is not enough, the service provider still needs to assign one IPv4 address per customer. One need to collocate a carrier-grade NAT with the tunnel concentrator within the service provider. In that solution, the IPv4 private addresses generated inside of the customer network would be transported (and not translated) within IPv6 packets across the service provider network to be decapsulated and then translated IPv4->IPv4 by the combined tunnel concentrator/carrier-grade NAT.
Note that, as in the above solutions, the presence of a carrier-grade NAT will break some NAT traversal techniques

---

5. Carrier-grade NAT considerations

TOC

One constant element in the architecture of all the above solution is the presence of a carrier-grade NAT, either IPv4->IPv4 or IPv6->IPv4. As some traditional NAT traversal techniques will stop working, this will have consequences on the set of applications that can be run in IPv4 mode.
Also, because IPv4 addresses will be share among customers and potentially a large address space reduction factor may be applied, in average, only a limited number of TCP or UDP port numbers will be

available per customer. This means that applications opening a very large number of TCP ports may have a harder time to work. For example, it has been reported that a very well know web site was using AJAX techniques and was opening up to 69 TCP ports per web page... If we make the hypothesis of an address space reduction of a factor 100 (one IPv4 address per 100 customers), a home with 10 PCs, and 65k ports per IPv4 addresses available, that makes a total of 65 ports available simultaneously for each PC, which is right on the edge of the number reported above for that well known application...

---

6. Standardization considerations

TOC

Any of the above solution could work. The double NAT IPv4>IPv4->IPv4 does not require any standard effort nor any new code in order to be deployed. However, dealing with multiple copies of net 10 may be a show stopper for large service providers, as the opex associated may be high. Both the double NAT IPv4->IPv6->IPv4 and IPv6 tunneling plus carrier-grade IPv4->IPv4 NAT may require some new code in the home gateways. Thus some standardization on a framework how to use these techniques is required.

---

7. Multicast considerations

TOC

This document only describes unicast IPv4. Some multicast IPv4 considerations need to be discussed as well. This section is a placeholder.

---

8. Acknowledgements

TOC

Send the author comments if you want your name listed here.

---

9. IANA Considerations

TOC

This memo includes no request to IANA.
This draft does not request any IANA action.

TOC

10. Security Considerations

Security issues associated with NAT have long been documented. A future version of this document may include some references here to previous work.

---

11. Normative References

TOC

[RFC2119]	Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).

Author's Address

TOC

	Alain Durand
	Comcast
	1500 Market st
	Philadelphia, PA 19102
	USA
Email:	alain_durand@cable.comcast.com

Full Copyright Statement

TOC

Intellectual Property

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.