Internet Engineering Task Force	S. Perreault, Ed.
Internet-Draft	Viagénie
Intended status: Best Current Practice	I. Yamagata
Expires: September 15, 2011	S. Miyakawa
	NTT Communications
	A. Nakagawa
	Japan Internet Exchange (JPIX)
	H. Ashida
	iTSCOM
	March 14, 2011

Common requirements for IP address sharing schemes
draft-ietf-behave-lsn-requirements-01

Abstract

This document defines common requirements for Carrier-Grade NAT (CGN) devices.

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 September 15, 2011.

Copyright Notice

Table of Contents

1. Introduction

With the shortage of IPv4 addresses, it is expected that more ISPs may want to provide a service where a public IPv4 address would be shared by many subscribers (also known as NAT444 [I-D.shirasaki-nat444-isp-shared-addr]). Each subscriber is assigned a private address, and a NAT device situated in the ISPs network translates between private and public addresses.

This is not to be considered a solution to the shortage of IPv4 addresses. It is a service that can conceivably be offered alongside others, such as IPv6 services or regular, un-NATed IPv4 service. Some ISPs started offering such a service long before there was a shortage of IPv4 addresses, showing that there are driving forces other than the shortage of IPv4 addresses.

This document describes behavioural requirements that are to be expected of those ISP-controlled NAT devices. Meeting this set of requirements will greatly increase the likelihood that subscribers' applications will function properly.

Readers should be aware of potential issues that may arise when sharing public address between many subscribers. See [I-D.ford-shared-addressing-issues] for details.

This document builds upon previous works describing requirements for generic NAT devices.[RFC4787][RFC5382][RFC5508]. These documents still apply in this context. What follows are additional requirements, to be satisfied on top of previous ones.

2. Terminology

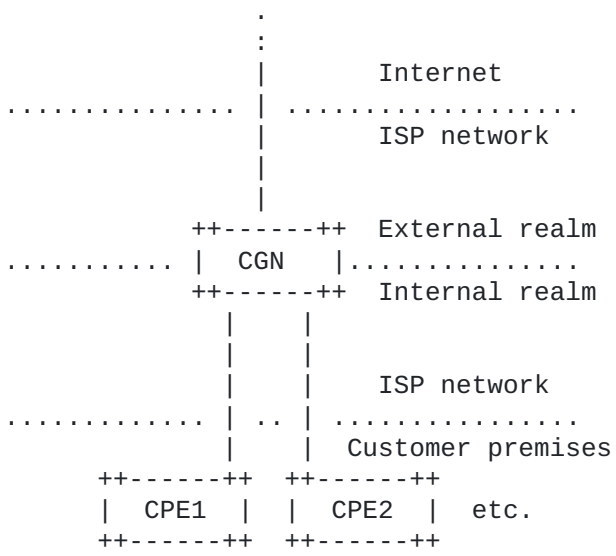
The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

Readers are expected to be familiar with [RFC4787] and the terms defined there. The following term is used in this document:

Carrier-Grade NAT (CGN): NAT device placed between a subscriber and the Internet in an ISP's network. A CGN translates IP addresses and transport-protocol port numbers in the packets that it forwards across the border between the internal and external realms.

*Note that the term "carrier-grade" has nothing to do with the quality of the NAT device; that is left to discretion of implementors. Rather, it is to be understood as a topological qualifier: the NAT device is placed in an ISP's network and translates the traffic of potentially many subscribers. Those have limited or no control over the CGN, whereas they typically have full control over a NAT placed on their premises.

Figure 1 summarises the network topology in which CGN devices operate.



3. Requirements for CGN devices

What follows is a list of requirements for CGN devices. They are in addition to those found in other documents such as [RFC4787], [RFC5382], and [RFC5508].

REQ-1:
A CGN MUST have an "IP address pooling" behaviour of "Paired".

Justification: This is a stronger form of REQ-2 from [RFC4787].

Note that this requirement applies regardless of the transport protocol. In other words, a CGN must use the same external IP address mapping for all sessions associated with the same internal IP address, be they TCP, UDP, ICMP, something else, or a mix of different protocols.

REQ-2: A CGN SHOULD limit the number of external ports (or, equivalently, "identifiers" for ICMP) that are assigned per CPE.

a. Limits SHOULD be configurable by the CGN administrator.

b. Limits MAY be configured and applied independently per transport protocol.

c. Additionally, it SHOULD be possible to limit the rate at which external ports are allocated.

Justification: A CGN can be considered a network resource that is shared by competing subscribers. Limiting the number of external ports assigned to each CPE mitigates the DoS attack that a subscriber could launch against the CGN in order to get a larger share of the resource. It ensures fairness among subscribers. Limiting the rate of allocation is intended to further help mitigate DoS attacks.

REQ-3: A CGN SHOULD limit the number of TCP sessions per CPE.

a. Limits SHOULD be configurable by the CGN administrator.

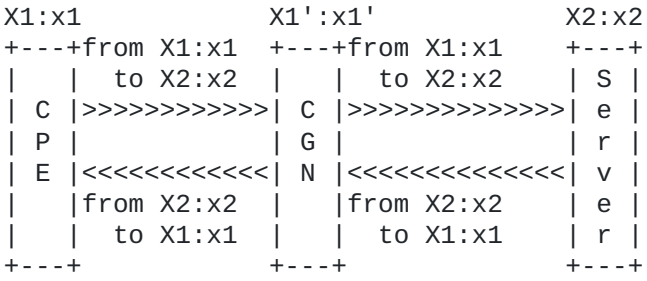
b. Additionally, it SHOULD be possible to limit the rate at which TCP sessions are instanciated.

Justification: A NAT needs to keep track of TCP sessions associated to each mapping. This state consumes resources for which, in the case of a CGN, subscribers may compete. It is necessary to ensure that each subscriber has access to a faire share of the CGN's resources. Limiting TCP sessions per CPE and per time unit is an effective mitigation against inter-subscriber DoS attacks. Limiting the rate of TCP session instanciation is intended to further help mitigate DoS attacks.

REQ-4: It SHOULD be possible to administratively turn off translation for specific destination addresses and/or ports.

Justification: It is common for a CGN administrator to provide access for subscribers to servers installed in the ISP's network, in the external realm. When such a server is able to reach the internal realm via normal routing (which is entirely controlled by the ISP), translation is unneeded. In that case, the CGN may forward packets without modification, thus acting like a plain router. This may represent an important efficiency gain.

Figure 2 illustrates this use-case.



REQ-5: It is RECOMMENDED that a CGN have an "Endpoint-Independent Filtering" behaviour.

Justification: This is a stronger form of REQ-8 from [RFC4787]. An "Address-Dependent Filtering" behaviour is NOT RECOMMENDED. This is based on the observation that some games and peer-to-peer applications require EIF for the NAT traversal to work. In the context of a CGN it is important to minimise application breakage.

REQ-6: When a CGN loses state (due to a crash, reboot, failover to a cold standby, etc.), it MUST start using a different external address pool.

Justification: This is necessary in order to prevent collisions between old and new mappings and sessions. It ensures that all established sessions are broken instead of redirected to a different peer. The previous address pool MAY of course be reused after a second loss of state.

4. Logging

It may be necessary for CGN administrators to be able to identify a subscriber based on external IPv4 address, port, and timestamp in order to deal with abuse and lawful intercept requests. When multiple subscribers share a single external address, the source address and port that are visible at the destination host have been translated from the ones originated by the CPE.

In order to be able to do this, the CGN needs to log the following information for each mapping created:

- *internal source address
- *internal source port
- *external source address
- *external source port
- *destination address (but see below)
- *destination port (but see below)
- *timestamp

A disadvantage of this is that CGNs under heavy usage may produce large amounts of logs, which may require large storage volume.

Readers should be aware of logging recommendations for Internet-facing servers [I-D.ietf-intarea-server-logging-recommendations]. With compliant servers, the destination address and port do not need to be logged by the CGN. This can help reduce the amount of logging.

5. Bulk Port Allocation

So far we have assumed that a CGN allocates one external port for every outgoing connection. In this section, the impacts of allocating multiple external ports at a time are discussed.

There is a range of things a CGN can do:

1. For every outgoing connection, allocate one external port.

2. For an outgoing connection, create a "bin" of several random external ports. Subsequent outgoing connections will use ports from the "bin". When the "bin" is full, a new connection causes a new bin to be created. A bin is smaller or equal to the user's maximum port limit.

3. Same as (2), but the ports allocated to a "bin" are consecutive instead of random.

Impacts are as follows.

Port Utilization: The mechanisms at the top of the list are very efficient in their port utilization. In that sense, they have good scaling properties (nothing is wasted). The mechanisms at the bottom of the list will waste ports. The number of wasted ports is proportional to size of the "bin".

Logging: Mechanism (1) creates a lot of log entries. Mechanisms (2) and (3) create the same number of log entries, but (3)'s log entries are smaller because a range can be expressed very compactly by indicating a range (e.g. "12000-12009"). With large "bin" sizes, the logging for mechanisms (2) and (3) can approach the logging frequency of DHCP servers.

Mechanism (1) can log destinations while mechanisms (2) and (3) cannot. This means that a CGN implementing one of the latter two will rely on the remote peer to follow the recommendations in [I-D.ietf-intarea-server-logging-recommendations]. If this is not acceptable, mechanisms (2) and (3) cannot be used.

Security: Mechanisms (1) and (2) provide very good security in that ports numbers are not easily guessed. Easily guessed port numbers put subscribers at risk of the attacks described in [RFC6056]. Mechanism (3) provides poor security to subscribers, especially if the "bin" size is small.

6. IANA Considerations

There are no IANA considerations.

7. Security Considerations

If a malicious subscriber can spoof another subscriber's CPE, it may cause a DoS to that subscriber by creating mappings up to the allowed limit. Therefore, the CGN administrator SHOULD ensure that spoofing is impossible. This can be accomplished with ingress filtering, as described in [RFC2827].

8. Acknowledgements

Thanks for the input and review by Tomohiro Nishitani, Yasuhiro Shirasaki, Takeshi Tomochika, Kousuke Shishikura, Dai Kuwabara, Tomoya Yoshida, Takanori Mizuguchi, Arifumi Matsumoto, Tomohiro Fujisaki, Dan Wing, and Dave Thaler. Dan Wing contributed much of section 5.

9. References

9.1. 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.
[RFC4787]	Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.
[RFC5382]	Guha, S., Biswas, K., Ford, B., Sivakumar, S. and P. Srisuresh, "NAT Behavioral Requirements for TCP", BCP 142, RFC 5382, October 2008.
[RFC5508]	Srisuresh, P., Ford, B., Sivakumar, S. and S. Guha, "NAT Behavioral Requirements for ICMP", BCP 148, RFC 5508, April 2009.
[RFC6056]	Larsen, M. and F. Gont, "Recommendations for Transport-Protocol Port Randomization", BCP 156, RFC 6056, January 2011.

9.2. Informative Reference

[I-D.shirasaki-nat444-isp-shared-addr]	Yamaguchi, J, Shirasaki, Y, Miyakawa, S, Nakagawa, A and H Ashida, "NAT444 addressing models", Internet-Draft draft-shirasaki-nat444-isp-shared-addr-06, July 2011.
[I-D.ford-shared-addressing-issues]	Ford, M, Boucadair, M, Durand, A, Levis, P and P Roberts, "Issues with IP Address Sharing", Internet-Draft draft-ford-shared-addressing-issues-02, March 2010.
[I-D.ietf-intarea-server-logging-recommendations]	Durand, A, Gashinsky, I, Lee, D and S Sheppard, "Logging recommendations for Internet facing servers", Internet-Draft draft-ietf-intarea-server-logging-recommendations-04, April 2011.

Appendix A. Change Log (to be removed by RFC Editor prior to publication)

Appendix A.1. Changed in -01

*Terminology: LSN is now CGN.

*Imported all requirements from RFCs 4787, 5382, and 5508. This allowed us to eliminate some duplication.

*Added references to draft-ietf-intarea-server-logging-recommendations and draft-ford-shared-addressing-issues.

*Incorporated a requirement from draft-xu-behave-stateful-nat-standby-06.

Authors' Addresses

Simon Perreault editor Perreault Viagénie 2875 boul. Laurier, suite D2-630 Québec, QC G1V 2M2 Canada Phone: +1 418 656 9254 EMail: simon.perreault@viagenie.ca URI: http://www.viagenie.ca

Ikuhei Yamagata Yamagata NTT Communications Corporation Gran Park Tower 17F, 3-4-1 Shibaura, Minato-ku Tokyo, 108-8118 Japan Phone: +81 50 3812 4704 EMail: ikuhei@nttv6.jp

Shin Miyakawa Miyakawa NTT Communications Corporation Gran Park Tower 17F, 3-4-1 Shibaura, Minato-ku Tokyo, 108-8118 Japan Phone: +81 50 3812 4695 EMail: miyakawa@nttv6.jp

Akira Nakagawa Nakagawa Japan Internet Exchange Co., Ltd. (JPIX) Otemachi Building 21F, 1-8-1 Otemachi, Chiyoda-ku Tokyo, 100-0004 Japan Phone: +81 90 9242 2717 EMail: a-nakagawa@jpix.ad.jp

Hiroyuki Ashida Ashida its communications Inc. 541-1 Ichigao-cho Aoba-ku Yokohama, 225-0024 Japan EMail: ashida@itscom.ad.jp