Internet Engineering Task Force	T. Nishitani	
Internet-Draft	I. Yamagata	
Intended status: BCP	S. Miyakawa	
Expires: June 3, 2010	NTT Communications	
	A. Nakagawa	
	KDDI CORPORATION	
	H. Ashida	
	iTSCOM	
	November 30, 2009	

TOC

Common Functions of Large Scale NAT (LSN)
draft-nishitani-cgn-03

Abstract

This document defines common functions of multiple types of Large Scale Network Address Translation (NAT) that handles Unicast UDP, TCP and ICMP.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and 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 June 3, 2010.

Table of Contents

---

1. Introduction

TOC

Global IPv4 address from the IANA pool will run out in a few years, thus network operators such as ISPs, carriers, large enterprises, universities need to shift from IPv4 services to IPv6 ones. However, IPv6 deployment seems to take a long time.
NAT [RFC3022] (Srisuresh, P. and K. Egevang, “Traditional IP Network Address Translator (Traditional NAT),” January 2001.) is a key technology to utilize IPv4 global address effectively in current practice. Operators may have to place NAT devices between end-users and the public Internet to suppress global IPv4 address consumption.
In this document, we call such a NAT device "Large Scale NAT (LSN)".

Variety of LSN (Large Scale NAT) have been proposed. Some of them are proposed for business continuity after the exhaustion, and some of them are proposed to access from IPv6 network to IPv4 Internet.

> - NAT444 [I-D.shirasaki-nat444-isp-shared-addr] (Shirasaki, Y., Miyakawa, S., Nakagawa, A., Yamaguchi, J., and H. Ashida, “NAT444 addressing models,” March 2010.)
>
> - DS-Lite (NAT464) [I-D.ietf-softwire-dual-stack-lite] (Durand, A., Droms, R., Haberman, B., Woodyatt, J., Lee, Y., and R. Bush, “Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion,” March 2010.)
>
> - NAT-64 [I-D.bagnulo-behave-nat64] (Bagnulo, M., Matthews, P., and I. Beijnum, “NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers,” March 2009.)

Each types of Large Scale NAT are shared by plural users and forward huge traffic. Because a demand is common, many of necessary functions are common.
This document recommends the common function of Large Scale NAT, so that developers and operators can easily implement these functions. Developpers of Large Scale NAT meet this set of requirements, they can consider specific functions of it. When an operator and a maker chose either implementation, the implementation has necessary functions.

---

2. Terminology

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 [RFC2119] (Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” March 1997.).
Readers are expected to be familiar with [RFC4787] (Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” January 2007.) and the terms defined there. The following term are used in this document:

> Large-Scale NAT(LSN): NAT devices placed between CPE and public Internet by an operator. LSN converts CPE IP Address, CPE Port, and CPE Identifier into LSN external IP Address, LSN external Port and LSN external Identifier in communication between CPE and GGN external.
>
> LSN external realm: The realm where IPv4 global addresses are assigned

> LSN internal realm: The realm placed between LSN and CPEs
>
> LSN external IP address: The IP address on LSN in LSN external realm mapping to CPE IP address
>
> LSN external port: The port on LSN in LSN external realm mapping to CPE port
>
> LSN external identifier: The identifier of ICMP on LSN in LSN external realm mapping to CPE identifier
>
> Customer Premises Equipment(CPE): The terminal which is placed in LSN internal realm and may establish TCP sessions to LSN external realm (e.g. a single PC or NATBox)
>
> CPE IP address: The IP address on CPE in LSN internal realm
>
> CPE port: The port on CPE in LSN internal realm
>
> CPE identifier: CPE's identifier of ICMP in LSN internal realm
>
> CPE 3-tuple: The tuple of TCP/UDP, CPE IP address, and CPE Port
> Service Server (SS) The server an operator supplies various services for CPE
>
> Service Server (SS): The server placed in external realm
>
> Service Provide Server (SPS): The server placed in external realm and controlled by LSN administrators

---

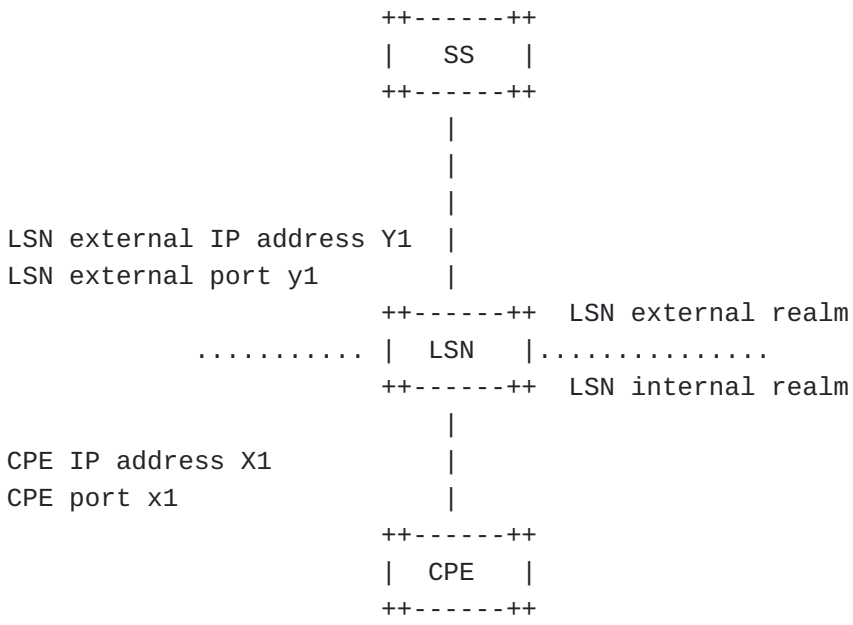

Figure 1. LSN network

---

3. The policy of assignment of LSN external IP address, port and identifier

TOC

A LSN has a pool of LSN external IP addresses, ports and identifiers. CPEs share LSN external IP addresses. Each LSN occupies combination of LSN external IP address, LSN external port and LSN external identifier exclusively. For a fair use of limited resources, LSN has a limitation for the number of the LSN external ports per CPE. LSNs need to keep high transparency to continue existing services after LSN is introduced. Requirement of high transparency for LSN leads to high scalability of LSN. High transparency means LSN basically keeps communications among CPEs except effect of limitations of the number of LSN external ports and TCP sessions.
A CPE MAY apply UDP hole punching or TCP hole punching for interactive services among CPEs like Voice over IP and P2P. LSN SHOLUD NOT interfere in services using UDP hole punching or TCP hole punching.
REQ-1: A LSN MUST allocate one external IP address to each CPE.

> a) LSN external IP address allocated to the CPE MUST be same for the UDP, TCP and ICMP.

Justification: If a LSN allocates multiple LSN external IP addresses to each CPE, some applications might not work.
REQ-2: A LSN MUST allocate LSN external ports which is mapped for CPE ports of UDP.

> a) A LSN MUST NOT overload LSN external port while a NAT UDP mapping timer does not expire.
>
> b) A LSN MAY reuse LSN external port after a NAT UDP mapping timer expires.
>
> c) A LSN SHOULD limit the number of the LSN external ports of UDP per CPE.
>
> d) The number of the LSN external ports of UDP per CPE which LSN can allocate SHOULD be configurable for the administrator of LSN.

Justification: CPEs can communicate to CPE external realm fairly by limiting the number of LSN external ports per CPE.
REQ-3: A LSN MUST allocate LSN external ports which is mapped for CPE ports of TCP.

> a) A LSN MUST NOT overload LSN external port while the port is allocated for one or more TCP sessions originated by another CPE.
>
> b) A LSN MAY reuse LSN external port while the port is allocated for no session originated by any CPE.
>
> c) A LSN SHOULD limit the number of the LSN external ports of TCP per CPE.
>
> d) The number of the LSN external ports of TCP per CPE SHOULD be an administratively configurable option.
>
> e) A LSN SHOULD limit the number of the new sessions of TCP per time unit and per CPE.

Justification: CPEs can communicate to CPE external realm fairly by limiting the number of LSN external ports per CPE. In addition, TCP LSN external port MAY have TCP sessions, and therefore the TCP session timer is necessary for every 5-Tuple. LSN can have not only the limitations of the number of LSN external ports but also TCP sessions per CPE. Thus a LSN can prevent denial of service attacks with the tons of TCP open and close by malicious CPEs.
REQ-4: A LSN MUST allocate LSN external identifiers which is mapped for CPE identifiers of ICMP.

> a) A LSN MUST NOT overload LSN external identifier before an ICMP Query session timer expires.

> b) A LSN MAY reuse LSN external identifier after an ICMP Query session timer expires.
>
> c) A LSN SHOULD limit the number of the LSN external identifier allocated per CPE.
>
> d) The number of the LSN external identifiers per CPE which LSN can allocate SHOULD be an administratively configurable option.

Justification: CPEs can communicate to CPE external realm fairly by limiting the number of LSN external identifiers every CPE.
If a CPE has already consumed many LSN external ports, the CPE might not use new ports because LSNs limit the number of ports.
REQ-5: A LSN MAY have implementations that some specific applications can work well even if each CPE's usable number of LSN external ports have already consumed.
Justification: Some specific applications don't work well due to limitation of number of number of ports by LSN, therefore other applications might be affected in the same CPE.
In Section 7 we discuss in detail.

---

4. Requirements for protocol handling

TOC

4.1. Unicast UDP Requirements

TOC

[RFC4787] (Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” January 2007.)
describes requirements of the Unicast UDP of a NAT, and the behavior of "Endpoint-Independent Filtering "is RECOMMNEDED, and a NAT MUST have an "Endpoint-Independent Mapping" behavior to ensure transparency of LSN. To have "Endpoint-Independent Filtering" and "Endpoint-Independent Mapping" behaviors for LSNs, LSNs help to establish UDP Hole Punching among CPEs. In other words, the possibility of the establishment of UDP Hole Punching among CPEs which have LSN is equal to the possibility among CPEs which don's t have LSN. If LSNs have an "Address-Dependent Mapping" or "Address and Port-Dependent Mapping" behavior, the possibility that establishment of UDP Hole Punching is less than when LSNs have an "Endpoint-Independent Mapping" behavior. If LSNs have an "Address and Port-Dependent Filtering" behavior, the possibility that establishment of UDP Hole Punching is less than when LSNs have an "Endpoint-Independent Filtering" or "Address Dependent Filtering" behavior.

If a LSN supports NAT Hairpinning, a CPE can communicate other CPEs in LSN internal realm of the same LSN.

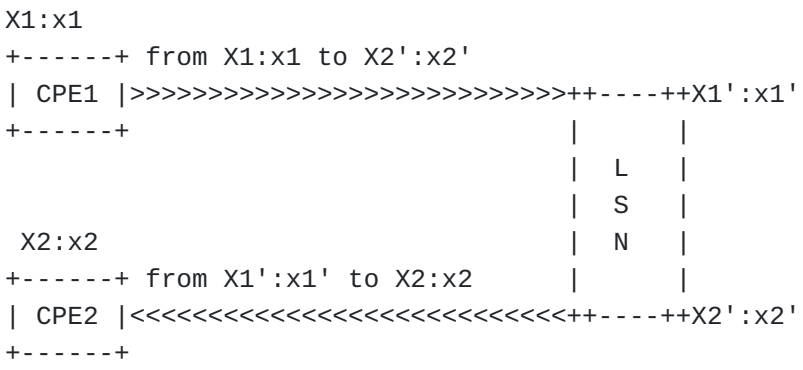

Figure 2. Hairpinning

REQ-6: A LSN SHOULD comply with [RFC4787] (Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP," January 2007.) for unicast UDP.
Justification: LSN SHOULD have to keep high transparency for unicast UDP communications. And CPE MAY use P2P and interactive services between CPEs after a LSN is introduced.

4.2. TCP Requirements

TOC

[RFC5382] (Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP," October 2008.) describes requirements of TCP of a NAT, and the behavior of "Endpoint-Independent Filtering" is RECOMMNEDED, and a NAT MUST have an "Endpoint-Independent Mapping" behavior to ensure transparency of LSN To have "Endpoint-Independent Filtering" and "Endpoint-Independent Mapping" behaviors for LSNs, LSNs help to establish TCP Hole Punching among CPEs. In other words, the possibility of the establishment of TCP Hole Punching among CPEs which have LSN is equal to the possibility among CPEs which don's t have LSN. If LSNs have an "Address-Dependent Mapping" or "Address and Port-Dependent Mapping" behavior, the possibility that establishment of TCP Hole Punching is less than when LSNs have an "Endpoint-Independent Mapping" behavior. If LSNs have an "Address and Port-Dependent Filtering" behavior, the possibility that

establishment of TCP Hole Punching is less than when LSNs have an "Endpoint-Independent Filtering" or "Address Dependent Filtering" behavior.
If a LSN supports NAT Hairpinning, a CPE can communicate other CPEs in LSN internal realm of the same LSN.
REQ-7: A LSN SHOULD comply with [RFC5382] (Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, “NAT Behavioral Requirements for TCP,” October 2008.) for TCP.
Justification: LSN SHOULD have to keep high transparency for TCP communications. And CPE MAY use P2P and interactive services between CPEs after a LSN is introduced.

4.3. ICMP Requirements

TOC

[RFC5508] (Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, “NAT Behavioral Requirements for ICMP,” April 2009.) describes requirements of ICMP of a NAT. And there MAY be a case that CPE cannot establish communication from CPEs to LSN external realm because LSN limits the number of LSN external ports, identifiers and TCP sessions per CPE. It is useful if CPE can distinguish an error to occur by the limitation of the LSN external ports, identifiers and TCP sessions from other errors.
REQ-8: A LSN SHOULD comply with [RFC5508] (Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, “NAT Behavioral Requirements for ICMP,” April 2009.) for ICMP.
Justification: LSN SHOULD have to keep high transparency for ICMP. And CPE MAY use P2P and interactive services between CPEs after a LSN is introduced.
Therefore, written in [RFC5508] (Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, “NAT Behavioral Requirements for ICMP,” April 2009.), when a LSN can't establish new session of TCP/UDP by limiting of TCP/UDP ports per user, the LSN sends an ICMP destination unreachable message, with code of 13 (Communication administratively prohibited) to the sender.

---

5. Summary of Requirements

TOC

REQ-1: A LSN MUST allocate one external IP address to each CPE.

> a) LSN external IP address allocated to the CPE MUST be same for the UDP, TCP and ICMP.

REQ-2: A LSN MUST allocate LSN external ports mapping to CPE ports of UDP.

> a) A LSN MUST NOT overload LSN external port while a NAT UDP mapping timer does not expire.
>
> b) A LSN MAY reuse LSN external port after a NAT UDP mapping timer expires.
>
> c) A LSN SHOULD limit the number of the LSN external ports of UDP per CPE.
>
> d) The number of the LSN external ports of UDP per CPE which LSN can allocate SHOULD be configurable for the administrator of LSN.

REQ-3: A LSN MUST allocate LSN external ports mapping to CPE ports of TCP.

> a) A LSN MUST NOT overload LSN external port while the port is allocated for one or more TCP sessions originated by another CPE.
>
> b) A LSN MAY reuse LSN external port while the port is allocated for no session originated by any CPE.
>
> c) A LSN SHOULD limit the number of the LSN external ports of TCP per CPE.
>
> d) The number of the LSN external ports of TCP per CPE SHOULD be an administratively configurable option.
>
> e) A LSN SHOULD limit the number of the new sessions of TCP per time unit and per CPE.

REQ-4: A LSN MUST allocate LSN external identifiers mapping to CPE identifiers.

> a) A LSN MUST NOT overload LSN external identifier before an ICMP Query session timer expires.
>
> b) A LSN MAY reuse LSN external identifier after an ICMP Query session timer expires.
>
> c) A LSN SHOULD limit the number of the LSN external identifier allocated per CPE.
>
> d) The number of the LSN external identifiers per CPE which LSN can allocate SHOULD be an administratively configurable option.

REQ-5: A LSN MAY have implementations that some specific applications can work well even if each CPE's usable number of LSN external ports have already consumed.

REQ-6: A LSN SHOULD comply with [RFC4787] (Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP," January 2007.) for unicast UDP.
REQ-7: A LSN SHOULD comply with [RFC5382] (Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP," October 2008.) for TCP.
REQ-8: A LSN SHOULD comply with [RFC5508] (Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, "NAT Behavioral Requirements for ICMP," April 2009.) for ICMP.

---

6. Identifying particular users (BOTs, spammers, etc)

TOC

It is necessary for network administrators to identify a user from an IP address and a timestamp in order to deal with abuse and lawful intercept. When multiple users share one external address at LSN, the source address and the source port that are visible at the destination host are translated ones. The following mechanisms can be used to identify the user that transmitted a certain packet.

6.1. Store Translation Log

TOC

One mechanism stores the following information at LSN.

- destination address
- destination port
- translated source address
- translated source port
- untranslated source address
- untranslated source port
- timestamp

In such environment that one LSN accommodates a lot of users or processes large amount of traffic, the amount of log will be so large and the operator has to prepare large volume of storage.

TOC

6.2. Fixed port assignment

To save costs for storage, one can adopt this port assignment mechanism at LSN. By fixing the range of external port per user/CPE, and having the mapping of internal IP address to external IP address and port, there will be no need to store per session log. Note that this mechanism is possible only if the source port is known as well as the source address, the destination address and the destination port.

---

7. Considerations about limiting the number of LSN external ports

TOC

As discussed in section 3, LSN limits the number of LSN external ports and identifier per CPE. Therefore some important applications like DNS might not work well due to applications consuming many LSN external ports.
There are two ways to solve this issue. The one is that particular applications are out of the targets for the number of port limitation for LSN. In the case, the applications should be configurable for the administrator of LSN.
The other is that LSN doesn't translate address or port for some specific applications, moreover it doesn't limit the number of LSN external ports.(we call "LSN pass-through") Therefore, LSN behave as a router. In this case, some specific applications are out of limitation for the number of LSN external ports. Some applications, which don't work well due to address translation like FTP, is effective. Reducing costs of translation is also effective. As a condition, administrators of LSN can control SPS which become a target of LSN pass-through.

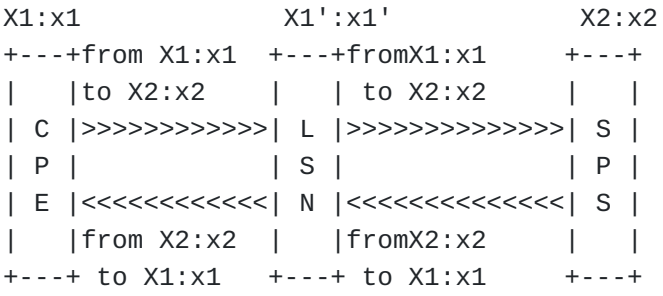

Figure 3. LSN pass-through

No matter which solutions you choose, you should consider which applications you are out of limitation target for the number of LSN external ports. When you choose too many applications, this might cause LSNs large load.

8. IANA Considerations

TOC

There are no IANA considerations.

9. Security Considerations

TOC

If malicious CPE can camouflage CPE 3-Tuple, the malicious CPE MAY prevent a normal CPE from sending data to external realm. Therefore, an operator SHOULD make policies to prevent a spoofing of CPE 3-tuple.

10. Acknowledgements

TOC

Thanks for the input and review by Yasuhiro Shirasaki, Takeshi Tomochika, Kousuke Shishikura, Dai Kuwabara, Tomoya Yoshida, Takanori Mizuguchi, Arifumi Matsumoto, Tomohiro Fujisaki

11. References

TOC

11.1. 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).
[RFC3022]	Srisuresh, P. and K. Egevang, “Traditional IP Network Address Translator (Traditional NAT),” RFC 3022, January 2001 (TXT).
[RFC4787]	Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” BCP 127, RFC 4787, January 2007 (TXT).
[RFC5382]	

	Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, “NAT Behavioral Requirements for TCP,” BCP 142, RFC 5382, October 2008 (TXT).
[RFC5508]	Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, “NAT Behavioral Requirements for ICMP,” BCP 148, RFC 5508, April 2009 (TXT).
[I-D.shirasaki-nat444-isp-shared-addr]	Shirasaki, Y., Miyakawa, S., Nakagawa, A., Yamaguchi, J., and H. Ashida, “NAT444 addressing models,” draft-shirasaki-nat444-isp-shared-addr-03 (work in progress), March 2010 (TXT).

11.2. Informative Reference

TOC

[I-D.ietf-softwire-dual-stack-lite]	Durand, A., Droms, R., Haberman, B., Woodyatt, J., Lee, Y., and R. Bush, “Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion,” draft-ietf-softwire-dual-stack-lite-04 (work in progress), March 2010 (TXT).
[I-D.bagnulo-behave-nat64]	Bagnulo, M., Matthews, P., and I. Beijnum, “NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers,” draft-bagnulo-behave-nat64-03 (work in progress), March 2009 (TXT).

Authors' Addresses

TOC

	Tomohiro Nishitani
	NTT Communications Corporation
	Gran Park Tower 17F, 3-4-1 Shibaura, Minato-ku
	Tokyo 108-8118
	Japan
Phone:	+81 50 3812 4742
Email:	tomohiro.nishitani@ntt.com
	Ikuhei 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
	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
	KDDI CORPORATION
	GARDEN AIR TOWER, 3-10-10, Iidabashi, Chiyoda-ku
	Tokyo 102-8460
	Japan
Email:	ai-nakagawa@kddi.com
	Hiroyuki Ashida
	its communications Inc.
	541-1 Ichigao-cho Aoba-ku
	Yokohama 225-0024
	Japan
Email:	ashida@itscom.ad.jp