Internet Engineering Task Force
Internet-Draft
Intended status: BCP
Expires: May 23, 2009

T. Nishitani
S. Miyakawa
NTT Communications
A. Nakagawa
KDDI CORPORATION
H. Ashida
iTSCOM

November 19, 2008

Common Functions of Large Scale NAT (LSN)
draft-nishitani-cgn-01

Status of this Memo

Abstract

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

Table of Contents

1. Introduction

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] 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 big 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]

- DS-Lite (NAT464) [I-D.durand-v6ops-natv4v6v4]

- NAT-64 [I-D.bagnulo-behave-nat64]

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.

By defining the common function in this document, developers of Large Scale NAT can put their development resource into their maker specific function. On the other hand, operator can introduce the Large Scale NAT that meets their requirement.

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 are used in this document:

Large-Scale NAT(LSN): NAT devices placed between CPE and public Internet by a 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 corresponding to CPE IP address

LSN external port: The port on LSN in LSN external realm corresponding to CPE port

LSN external identifier: The identifier of ICMP on LSN in LSN external realm corresponding 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

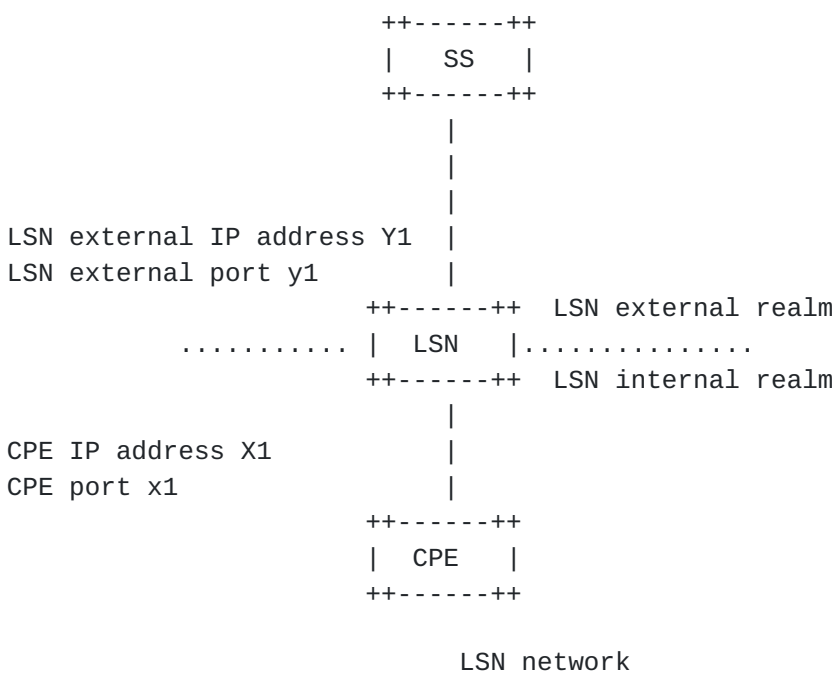
CPE IP address: The IP address on CPE in LSN internal realm

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 a operator supplies various services for CPE

Service Server (SS): The server a operator supplies various services for CPE



LSN network

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

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 and LSN external port 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 of the UDP, TCP and ICMP MUST be same.

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 corresponding 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 overload 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.

e) A LSN SHOULD NOT allocate well-known ports as LSN external ports.

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 corresponding 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.

f) A LSN SHOULD NOT allocate well-known ports as LSN external ports.

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 corresponding to CPE identifiers.

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

b) A LSN MAY overload 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.

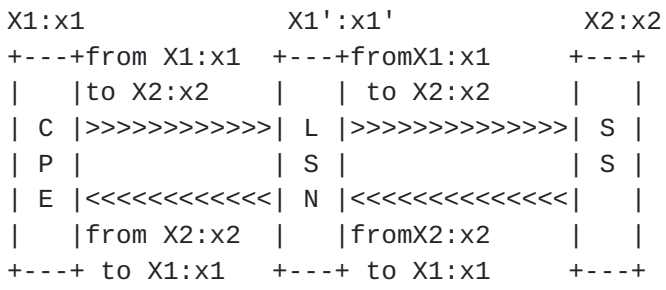
When a LSN limits the number of LSN external ports and TCP sessions, CPE may not use TCP services during using web and P2P services. For example, some services using Ajax demand few dozens of TCP sessions. P2P software like BitTorrent demands also TCP sessions more than few dozens. Some CPEs MAY use E-mail services like POP3 and SMTP even though CPE uses the services which demand many TCP sessions at the same time. Therefore it is important to reserve LSN external ports for such administratively configured services.

REQ-5: Reserving LSN external ports per CPE for the always-available services are RECOMENDED.

a) The destination port which is used for reservation of LSN external ports SHOULD be administratively configurable.

Justification: To reserve the LSN external ports for specific services, CPE can avoid the effect of the limitation of LSN external ports by LSN.

In addition, it MAY not be necessary to set a limit to the number of LSN external ports for the communications between CPEs and SS. The reason is because LSN should pass-through the communications between CPEs and SS.



pass-through

REQ-6: A LSN SHOULD pass-through the communication between CPEs and SS.

Justification: Using pass-through, LSN does not have to assign LSN external IP address, ports, and identifiers and limit to the number of ports and TCP sessions for the services that an operator manages.

4. Requirements for protocol handling

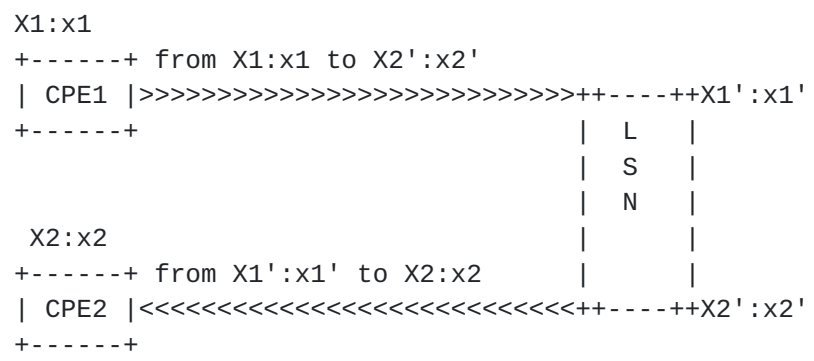
4.1. Unicast UDP Requirements

[RFC4787] 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. And 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. Because a SS is placed external LSN realm, the source IP address and port of the communication from CPE to SS is LSN external IP address and port. It is RECOMMENDED to use STUN[I-D.ietf-behave-rfc3489bis] if CPEs check the LSN external IP address and port for CPE.

An operator MAY introduce TURN [I-D.ietf-behave-turn] to support

communications among CPEs. If LSN supports "Hairpinning", LSN can hairpin the communications between CPEs in the same LSN. Therefore the requirements of Hairpinning for LSN MAY reduce requirements for the performance of TURN servers. When CPEs decide the course of UDP between CPEs, CPE MAY use [I-D.ietf-mmusic-ice] .



Hairpinning

REQ-7: A LSN SHOULD comply with [RFC4787] 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

[I-D.ietf-behave-tcp] 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. And 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. Because a SS is placed external LSN realm, the source of IP address and port of the communication from

CPE to SS is LSN external IP address and port. It is RECOMMENDED to use STUN[I-D.ietf-behave-rfc3489bis] if CPEs want to check the LSN external IP address and port for CPE.

An operator MAY introduce TURN [I-D.ietf-behave-turn] to support communications among CPEs. If LSN supports "Hairpinning", LSN can hairpin the communications between CPEs in the same LSN. Therefore the requirements of Hairpinning for LSN MAY reduce requirements for the performance of TURN servers. When CPEs decide the course of TCP between CPEs, CPE MAY use [I-D.ietf-mmusic-ice] .

REQ-8: A LSN SHOULD comply with [I-D.ietf-behave-tcp] 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

[I-D.ietf-behave-nat-icmp] 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-9: A LSN SHOULD comply with [I-D.ietf-behave-nat-icmp] for ICMP.

a) 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.

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. And it is necessary to be able to distinguish an error to occur by the limitation of the LSN external ports and TCP sessions from a network error.

4.4. Summary of Requirements

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

a) LSN external IP address of the UDP, TCP and ICMP MUST be same.

REQ-2: A LSN MUST allocate LSN external ports corresponding 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 overload 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.

e) A LSN SHOULD NOT allocate well-known ports as LSN external ports.

REQ-3: A LSN MUST allocate LSN external ports corresponding 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.

f) A LSN SHOULD NOT allocate well-known ports as LSN external ports.

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

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

b) A LSN MAY overload 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: Reserving LSN external ports per CPE for the always-available services are RECOMENDED.

a) The destination port which is used for reservation of LSN external ports SHOULD be administratively configurable.

REQ-6: A LSN SHOULD pass-through the communication between CPEs and SS.

REQ-7: A LSN SHOULD comply with [RFC4787] for unicast UDP.

REQ-8: A LSN SHOULD comply with [I-D.ietf-behave-tcp] for TCP.

REQ-9: A LSN SHOULD comply with [I-D.ietf-behave-nat-icmp] for ICMP.

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

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.

5.1. Store Translation Log

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.

5.2. 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.

6. IANA Considerations

There are no IANA considerations.

7. Security Considerations

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.

8. Acknowledgements

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

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.

[RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022,

January 2001.

[RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.

[I-D.shirasaki-nat444-isp-shared-addr] Shirasaki, Y., Miyakawa, S., Nakagawa, A., Yamaguchi, J., and H. Ashida, "NAT444 with ISP Shared Address", draft-shirasaki-nat444-isp-shared-addr-00 (work in progress), October 2008.

[I-D.ietf-behave-tcp] Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", draft-ietf-behave-tcp-08 (work in progress), September 2008.

[I-D.ietf-behave-nat-icmp] Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, "NAT Behavioral Requirements for ICMP protocol", draft-ietf-behave-nat-icmp-10 (work in progress), October 2008.

[I-D.ietf-behave-rfc3489bis] Rosenberg, J., Mahy, R., Matthews, P., and D. Wing, "Session Traversal Utilities for (NAT) (STUN)", draft-ietf-behave-rfc3489bis-18 (work in progress), July 2008.

[I-D.ietf-mmusic-ice] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", draft-ietf-mmusic-ice-19 (work in progress), October 2007.

[I-D.ietf-behave-turn] Rosenberg, J., Mahy, R., and P. Matthews, "Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)", draft-ietf-behave-turn-11 (work in progress), October 2008.

9.2. Informative Reference

[I-D.durand-v6ops-natv4v6v4] Durand, A., "Distributed NAT for broadband deployments post IPv4 exhaustion", draft-durand-v6ops-natv4v6v4-01

(work in progress), February 2008.

[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-nat64-02 (work in progress), November 2008.

Authors' Addresses

Tomohiro Nishitani
NTT Communications Corporation
Tokyo Opera City Tower 21F, 3-20-2 Nishi-Shinjuku, Shinjuku-ku
Tokyo 163-1421
Japan

Phone: +81 3 6800 3214
Email: tomohiro.nishitani@ntt.com

Shin Miyakawa
NTT Communications Corporation
Tokyo Opera City Tower 21F, 3-20-2 Nishi-Shinjuku, Shinjuku-ku
Tokyo 163-1421
Japan

Phone: +81 3 6800 3262
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.
3-5-7 Hisamoto Takatsu-ku
Kawasaki 213-0011
Japan

Email: ashida@itscom.ad.jp

Full Copyright Statement

Intellectual Property