Internet Engineering Task Force	I. Yamagata	
Internet-Draft	S. Miyakawa	
Intended status: BCP	NTT Communications	
Expires: January 13, 2011	A. Nakagawa	
	Japan Internet Exchange (JPIX)	
	H. Ashida	
	iTSCOM	
	July 12, 2010	

TOC

Common requirements for IP address sharing schemes
draft-nishitani-cgn-05

Abstract

This document defines common requirements 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 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 January 13, 2011.

Copyright Notice

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

TOC

Now there are several IPv4 address sharing schemes such as Large Scale NAT (as known as 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[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,” July 2010.), A+P[I-D.ymbk-aplusp] (Bush, R., “The A+P Approach to the IPv4 Address Shortage,” October 2009.) and so on under the discussion.
Those IPv4 address sharing schemes are intended to be used in the middle of the ISP access network against IPv4 address shortage problem by sharing one global IPv4 address by multiple users. Authors believe that there are common requirements among all IPv4 address sharing schemes to make them "transparent" as much as possible. At the BEHAVE working group of IETF, following RFCs have already defined to achieve maximum transparency at the residential CPE which has NAT function;

- RFC4787 : NAT Behavioral Requirements for Unicast UDP

- RFC5382 : NAT Behavioral Requirements for TCP

- RFC5508 : NAT Behavioral Requirements for ICMP

However so, because those RFCs are mainly aimed at residential CPE and any IPv4 address sharing schemes are a bit different from it, we believe that requirements for LSN and other schemes should be defined alternatively to those RFCs.

---

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

---

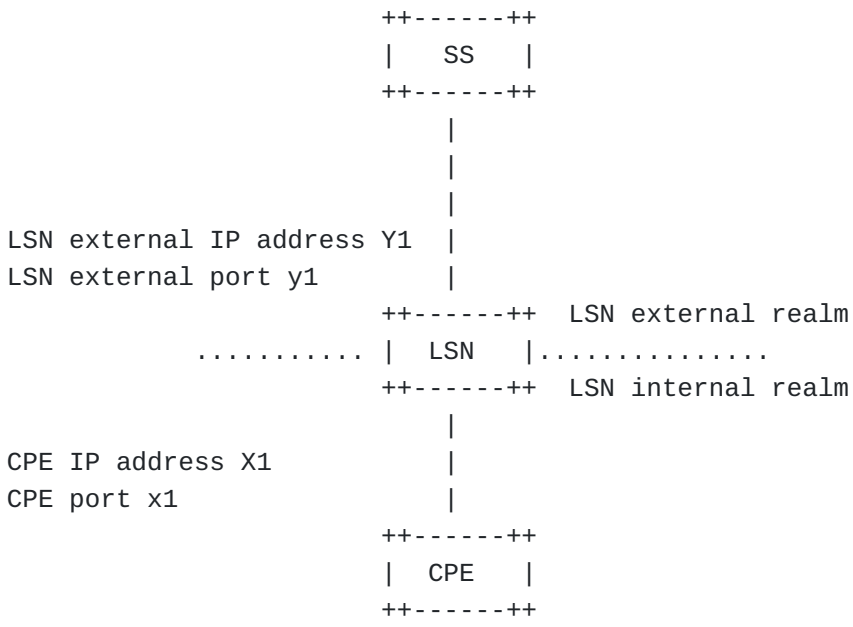

Figure 1. LSN network

---

3. Requirements for UDP

TOC

Based on RFC4787, we'd like to compile the list of the requirements as follows.
Please note that REQ-8 is slightly different for original RFC. And some of requirements have additional justification.
REQ-1: A NAT MUST have an "Endpoint-Independent Mapping" behavior.

> Status: Same as REQ-1 in RFC4787
>
> Justification: This is needed to use UNilateral Self-Address Fixing (UNSAF) which plays important role in STUN / TURN. More detailed description can be found in the original RFC. But to be more precise, in the LSN case, it may not be needed for some specific protocol such as DNS query and response.

REQ-2: It is RECOMMENDED that a NAT have an "IP address pooling" behavior of "Paired". Note that this requirement is not applicable to NATs that do not support IP address pooling.

> Status: Same as REQ-2 in RFC4787
>
> Justification: This allows applications that use multiple ports originating from the same internal IP address to also have the same external IP address. More detailed description can be found in original RFC.

REQ-3: A NAT MUST NOT have a "Port assignment" behavior of "Port overloading".

> Status: Same as REQ-3 in RFC4787
>
> Justification: This requirement must be met in order to enable two applications on the internal side of the NAT both to use the same port to try to communicate with the same destination. More detailed description can be found in original RFC.

REQ-3-a: If the host's source port was in the range 0-1023, it is RECOMMENDED the NAT's source port be in the same range. If the host's source port was in the range 1024-65535, it is RECOMMENDED that the NAT's source port be in that range.

> Status: Same as REQ-3-a in RFC4787
>
> Justification: Certain applications expect the source UDP port to be in the well-known range. More detailed description can be found in original RFC. On the other hand, almost application probably not use range 0-1023 for source port. Using ports as many as possible, it may not be needed this requirement.

REQ-4: It is RECOMMENDED that a NAT have a "Port parity preservation" behavior of "Yes".

> Status: Same as REQ-4 in RFC4787
>
> Justification: This is avoid breaking peer-to-peer applications that do not explicitly and separately specify RTP and RTCP port numbers and that follow the RFC 3550 rule to decrement an odd RTP port to

> make it even. More detailed description can be found in original RFC.

REQ-5: A NAT UDP mapping timer MUST NOT expire in less than two minutes, unless REQ-5-a applies.
REQ-5-a: For specific destination ports in the well-known port range (ports 0-1023), a NAT MAY have shorter UDP mapping timers that are specific to the IANA-registered application running over that specific destination port.
REQ-5-b: The value of the NAT UDP mapping timer SHOULD be configurable.
REQ-5-c: A default value of five minutes or more for the NAT UDP mapping timer is RECOMMENDED.

> Status: Same as REQ-5, REQ-5-a, REQ-5-b, REQ-5-c in RFC4787

REQ-6: The NAT mapping Refresh Direction MUST have a "NAT Outbound refresh behavior" of "True".

> Status: Same as REQ-6 in RFC4787
>
> Justification: Outbound refresh is necessary for allowing the client to keep the mapping alive. More detailed description can be found in original RFC.

REQ-6-a: The NAT mapping Refresh Direction MAY have a "NAT Inbound refresh behavior" of "True".

> Status: Same as REQ-6-a in RFC4787
>
> Justification: Allowing inbound refresh may allow an external attacker or misbehaving application to keep a mapping alive indefinitely. Also, it the process is repeated with different ports, over time, it could use up all the ports on the NAT. But this requirement is maybe needed for some applications occurring only incoming inbound traffic. In LSN, Making much of transparency, this requirement is more necessary.

REQ-7: A NAT device whose external IP interface can be configured dynamically MUST either

> (1) Automatically ensure that its internal network uses IP addresses that do not conflict with its external network, or
>
> (2) Be able to translate and forward traffic between all internal nodes and all external nodes whose IP addresses numerically conflict with the internal network.

> Status: Same as REQ-7 in RFC4787

REQ-8: It is RECOMMENDED that a NAT have "Endpoint-Independent Filtering" behavior.

> Status: "If application transparency is most important, it is RECOMMENDED that a NAT have Endpoint-Independent Filtering behavior. If a more stringent filtering behavior is most important, it is RECOMMENDED that a NAT have Address-Dependent Filtering behavior." is written at REQ-8 in RFC4787. In this draft, we pick up only first requirement.
>
> Justification: LSN which is placed at ISP/Carrier makes much of transparency. In particular, for applications that receive media simultaneously from multiple locations (e.g., gaming), or applications that use rendezvous techniques. But to be more precise, in the LSN case, it may not be needed for some specific protocol such as DNS query and response.

REQ-8-a: The filtering behavior MAY be an option configurable by the administrator of the NAT.

> Status: Same as REQ-8-a in RFC4787
>
> Justification: Having the filtering behavior being an option configurable by the administrator of the NAT ensures that a NAT can be used in the widest variety of deployment scenarios. More detailed description can be found in original RFC.

REQ-9: A NAT MUST support "Hairpinning".
REQ-9-a: A NAT Hairpinning behavior MUST be "External source IP address and port".

> Status: Same as REQ-9 in RFC4787
>
> Justification: These requirements are to allow communications between two endpoints behind the same NAT when they are trying each other's external IP address. More detailed description can be found in original RFC.

REQ-10: To eliminate interference with UNSAF NAT traversal mechanisms and allow integrity protection of UDP communications, NAT ALGs for UDP-based protocols SHOULD be turned off. Future standards track specifications that define an ALG can update this to recommend the ALGs on which they define default.
REQ-10-a: If a NAT includes ALGs, it is RECOMMENDED that the NAT allow the NAT administrator to enable or disable each ALG separately.

> Status: Same as REQ-10, REQ-10-a in RFC4787
>
> Justification: NAT ALGs may interfere with UNSAF methods. More detailed description can be found in original RFC.

REQ-11: A NAT MUST have deterministic behavior, i.e., it MUST NOT change the NAT translation or the Filtering Behavior at any point in time, or under any particular conditions.

> Status: Same as REQ-11 in RFC4787
>
> Justification: Non-deterministic NATs are very difficult to troubleshoot. More detailed description can be found in original RFC.

REQ-12: Receipt of any sort of ICMP message MUST NOT terminate the NAT mapping.
REQ-12-a: The NAT's default configuration SHOULD NOT filter ICMP messages based on their source IP address.
REQ-12-b: It is RECOMMENDED that a NAT support ICMP Destination Unreachable messages.

> Status: Same as REQ-12, REQ-12-a, REQ-12-b in RFC4787
>
> Justification: This is easy to do and is used for many things including MTU discovery and rapid detection of error conditions, and has no negative consequences. More detailed description can be found in original RFC.

REQ-13: If the packet received on an internal IP address has DF=1, the NAT MUST send back an ICMP message "Fragmentation needed and DF set" to the host, as described in [RFC0792] (Postel, J., “Internet Control Message Protocol,” September 1981.).
REQ-13-a: If the packet has DF=0, the NAT MUST fragment the packet and SHOULD send the fragments in order.

> Status: Same as REQ-13, REQ-13-a in RFC4787
>
> Justification: This is the same function a router performs in a similar situation. More detailed description can be found in original RFC.

REQ-14: A NAT MUST support receiving in-order and out-of-order fragments, so it MUST have "Received Fragment Out of Order" behavior.
REQ-14-a: A NAT's out-of-order fragment processing mechanism MUST be designed so that fragmentation-based DoS attacks do not compromise the NAT's ability to process in-order and unfragmented IP packets.

> Status: Same as REQ-14, REQ-14-a in RFC4787
>
> Justification: Since some networks deliver small packets ahead of large ones, there can be many out-of order fragments. NATs that are capable of delivering these out-of-order packets are possible, but they need to store the out-of-order fragments which can open up a

> Denial-of-Service (DoS) opportunity, if done incorrectly. More detailed description can be found in original RFC.

---

4. Requirements for TCP

TOC

Based on RFC5382, we'd like to compile the list of the requirements as follows.
Please note that REQ-17 is slightly different for original RFC. And some of requirements have additional justification.
REQ-15: A NAT MUST have an "Endpoint Independent Mapping" behavior for TCP.

> Status: Same as REQ-1 in RFC5382
>
> Justification: This is needed to use UNilateral Self-Address Fixing (UNSAF) which plays important role in STUN / TURN. More detailed description can be found in the original RFC. But to be more precise, in the LSN case, it may not be needed for some specific protocols.

REQ-16: A NAT MUST support all valid sequences of TCP packets for connections initiated both internally as well as externally when the connection is permitted by the NAT.
REQ-16-a: In addition to handling the TCP 3-way handshake mode of connection initiation, A NAT MUST handle the TCP simultaneous-open mode of connection initiation.

> Status: Same as REQ-2,REQ-2-a in RFC5382
>
> Justification: This is to allow standards compliant TCP stacks to traverse NATs. More detailed description can be found in original RFC.

REQ-17: It is RECOMMENDED that a NAT have an "Endpoint independent filtering" behavior for TCP.

> Status: "If application transparency is most important, it is RECOMMENDED that a NAT have an "Endpoint independent filtering" behavior for TCP. If a more stringent filtering behavior is most important, it is RECOMMENDED that a NAT have an "Address dependent filtering" behavior." is REQ-3 in RFC5382. In this draft, we pick up only first requirement.
>
> Justification: LSN which is placed at ISP/Carrier makes much of transparency. But to be more precise, in the LSN case, it may not be needed for some specific protocols.

REQ-17-a: The filtering behavior MAY be an option configurable by the administrator of the NAT.
REQ-17-b: The filtering behavior for TCP MAY be independent of the filtering behavior for UDP.

> Status: Same as REQ-3-a, REQ-3-b in RFC5382

REQ-18: A NAT MUST NOT respond to an unsolicited inbound SYN packet for at least 6 seconds after the packet is received. If during this interval the NAT receives and translates an outbound SYN for the connection the NAT MUST silently drop the original unsolicited inbound SYN packet. Otherwise the NAT SHOULD send an ICMP Port Unreachable error (Type 3, Code 3) for the original SYN, unless REQ-18-a applies.
REQ-18-a: The NAT MUST silently drop the original SYN packet if sending a response violates the security policy of the NAT.

> Status: Same as REQ-4, REQ-4-a in RFC5382
>
> Justification: This intent of this requirement is to allow simultaneous-open to work reliably in the presence of NATs. More detailed description can be found in original RFC.

REQ-19: If a NAT cannot determine whether the endpoints of a TCP connection are active, it MAY abandon the session if it has been idle for some time. In such cases, the value of the "established connection idle-timeout" MUST NOT be less than 2 hours 4 minutes. The value of the "transitory connection idle-timeout" MUST NOT be less than 4 minutes.
REQ-19-a: The value of the NAT idle-timeouts MAY be configurable.

> Status: Same as REQ-5, REQ-5-a in RFC5382
>
> Justification: The intent of this requirement is to minimize the cases where a NAT abandons session state for a live connection. More detailed description can be found in original RFC.

REQ-20: If a NAT includes ALGs that affect TCP, it is RECOMMENDED that all of those ALGs (except for FTP) be disabled by default.

> Status: Same as REQ-6 in RFC5382
>
> Justification: The intent of this requirement is to prevent ALGs from interfering with UNSAF methods. More detailed description can be found in original RFC.

REQ-21: A NAT MUST NOT have a "Port assignment" behavior of "Port overloading" for TCP.

> Status: Same as REQ-7 in RFC5382

> Justification: This requirement allows two applications on the internal side of the NAT to consistently communicate with the same destination.

REQ-22: A NAT MUST support "Hairpinning" for TCP.
REQ-22-a: A NAT's Hairpinning behavior MUST be of type "External source IP address and port".

> Status: Same as REQ-8, REQ-8-a in RFC5382
>
> Justification: This requirement allows two applications behind the same NAT that are trying to communicate with each other using their external addresses. More detailed description can be found in original RFC.

REQ-23: If a NAT translates TCP, it SHOULD translate ICMP Destination Unreachable (Type 3) messages.

> Status: Same as REQ-9 in RFC5382
>
> Justification: Translating ICMP Destination Unreachable messages avoids communication failures. More detailed description can be found in original RFC.

REQ-24: Receipt of any sort of ICMP message MUST NOT terminate the NAT mapping or TCP connection for which the ICMP was generated.

> Status: Same as REQ-10 in RFC5382
>
> Justification: This is necessary for reliably performing TCP simultaneous-open where a remote NAT may temporarily signal an ICMP error. More detailed description can be found in original RFC.

---

5. Requirements for ICMP

TOC

Based on RFC5508, we'd like to compile the list of the requirements as follows.
Some of requirements have additional justification.
REQ-25: Unless explicitly overridden by local policy, a NAT device MUST permit ICMP Queries and their associated responses, when the Query is initiated from a private host to the external hosts.
REQ-25-a: NAT mapping of ICMP Query Identifiers SHOULD be external host independent.

> Status: Same as REQ-1 in RFC5508

> Justification: ICMP Query mapping by NAT devices is necessary for current ICMP-Query-based applications to work. More detailed description can be found in original RFC.

REQ-26: An ICMP Query session timer MUST NOT expire in less than 60 seconds.
REQ-26-a: It is RECOMMENDED that the ICMP Query session timer be made configurable.

> Status: Same as REQ-2, REQ-2-a in RFC5508
>
> Justification: Setting the ICMP NAT session timeout to a very large duration (say, 240 seconds) could potentially tie up precious NAT resources for the whole duration. On the other hand, setting the timeout very low can result in premature freeing of NAT resources and applications failing to complete gracefully. A 60-second timeout is a balance between the two extremes. More detailed description can be found in original RFC.

REQ-27: When an ICMP Error packet is received, if the ICMP checksum fails to validate, the NAT SHOULD silently drop the ICMP Error packet. If the ICMP checksum is valid, do the following.

> **a.** If the IP checksum of the embedded packet fails to validate, the NAT SHOULD silently drop the Error packet; and
>
> **b.** If the embedded packet includes IP options, the NAT device MUST traverse past the IP options to locate the start of transport header for the embedded packet; and
>
> **c.** The NAT device SHOULD NOT validate the transport checksum of the embedded packet within an ICMP Error message, even when it is possible to do so; and
>
> **d.** If the ICMP Error payload contains ICMP extensions, the NAT device MUST exclude the optional zero-padding and the ICMP extensions when evaluating transport checksum for the embedded packet.

> Status: Same as REQ-3 in RFC5508
>
> Justification: An ICMP Error message checksum covers the entire ICMP message, including the payload. NAT uses the embedded IP and transport headers for forwarding and translating the ICMP Error message. More detailed description can be found in original RFC.

REQ-28: If a NAT device receives an ICMP Error packet from external realm, and the NAT device does not have an active mapping for the embedded payload, the NAT SHOULD silently drop the ICMP Error packet. If the NAT has active mapping for the embedded payload, then the NAT

MUST do the following prior to forwarding the packet, unless local policy explicitly overridden by local policy:

> **a.** Revert the IP and transport headers of the embedded IP packet to their original form, using the matching mapping; and
>
> **b.** Leave the ICMP Error type and code unchanged; and
>
> **c.** Modify the destination IP address of the outer IP header to be same as the source IP address of the embedded packet after translation.

> Status: Same as REQ-4 in RFC5508

REQ-29: If a NAT device receives an ICMP Error packet from the private realm, and the NAT does not have an active mapping for the embedded payload, the NAT SHOULD silently drop the ICMP Error packet. If the NAT has active mapping for the embedded payload, then the NAT MUST do the following prior to forwarding the packet, unless explicitly overridden by local policy.

> **a.** Revert the IP and transport headers of the embedded IP packet to their original form, using the matching mapping; and
>
> **b.** Leave the ICMP Error type and code unchanged; and
>
> **c.** If the NAT enforces Basic NAT function, and the NAT has active mapping for the IP address that sent the ICMP Error, translate the source IP address of the ICMP Error packet with the public IP address in the mapping. In all other cases, translate the source IP address of the ICMP Error packet with its own public IP address.

> Status: Same as REQ-5 in RFC5508

REQ-30: While processing an ICMP Error packet pertaining to an ICMP Query or Query response message, a NAT device MUST NOT refresh or delete the NAT Session that pertains to the embedded payload within the ICMP Error packet.

> Status: Same as REQ-6 in RFC5508
>
> Justification: This requirement ensures that the NAT Session will not be modified if someone is able to spoof ICMP Error messages for the session. More detailed description can be found in original RFC.

REQ-31: LSN devices MUST support the traversal of hairpinned ICMP Query sessions and Error messages.

> **a.**

> When forwarding a hairpinned ICMP Error message, the NAT device MUST translate the destination IP address of the outer IP header to be same as the source IP address of the embedded IP packet after the translation

> Status: "NAT devices enforcing Basic NAT MUST support the traversal of hairpinned ICMP Query sessions. All NAT devices (i.e., Basic NAT as well as NAPT devices) MUST support the traversal of hairpinned ICMP Error messages." is REQ-7 in RFC5508. LSN is kind of Basic NATs, and is enforced Basic NAT behavior, so LSN MUST support ICMP Query and Error messages.
>
> Justification: This requirement is necessary for current applications to work correctly. More detailed description can be found in original RFC.

REQ-32: When a NAT device is unable to establish a NAT Session for a new transport-layer (TCP, UDP, ICMP, etc.) flow due to resource constraints or administrative restrictions, the NAT device SHOULD send an ICMP destination unreachable message, with a code of 13 (Communication administratively prohibited) to the sender, and drop the original packet.

> Status: Same as REQ-8 in RFC5508
>
> Justification: LSN, limiting the number of the LSN external ports of UDP/TCP per CPE, often unable to establish new NAT session for a CPE, because the CPE use many sessions. In this case, LSN SHOULD send an ICMP destination unreachable message or some applications maybe not work well.

REQ-33: A NAT device MAY implement a policy control that prevents ICMP messages being generated toward certain interface(s). Implementation of such a policy control overrides the MUSTs and SHOULDs in REQ-34.
REQ-34: Unless overridden by REQ-33's policy, a NAT device needs to support ICMP messages as below, some conforming to Section 4.3 of [RFC1812] and some superseding the requirements of Section 4.3 of [RFC1812]:

> a) MUST support:
>
> 1. Destination Unreachable Message
>
> 2. Time Exceeded Message
>
> 3. Echo Request/Reply Messages
>
> b) MAY support:
>
> 1. Redirect Message

2. Timestamp and Timestamp Reply Messages

3. Source Route Options

4. Address Mask Request/Reply Message

5. Parameter Problem Message

6. Router Advertisement and Solicitations

c) SHOULD NOT support

1. Source Quench Message

2. Information Request/reply

In addition, a NAT device is RECOMMENDED to conform to the following implementation considerations:

a. d) DS Field Usage

b. e) When Not to Send ICMP Errors

c. f) Rate Limiting

Status: Same as REQ-9, REQ-10 in RFC5508

Justification: These are for conformance to RFC 1812.

REQ-35: A NAT MAY drop or appropriately handle Non-QueryError ICMP messages.

Status: Same as REQ-11 in RFC5508

Justification: NAT devices may handle of Non-QueryError ICMP messages.

---

6. LSN specified Requirements

TOC

REQ-36: 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-37: A LSN MUST allocate LSN external ports which is mapped for CPE ports of UDP.

> a) A LSN MAY reuse LSN external port after a NAT UDP mapping timer expires.
>
> b) A LSN SHOULD limit the number of the LSN external ports of UDP per CPE.
>
> c) 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-38: A LSN MUST allocate LSN external ports which is mapped for CPE ports of TCP.

> a) A LSN MAY reuse LSN external port while the port is allocated for no session originated by any CPE.
>
> b) A LSN SHOULD limit the number of the LSN external ports of TCP per CPE.
>
> c) 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-39: A LSN MUST allocate LSN external identifiers which is mapped for CPE identifiers of ICMP.

> a) A LSN MAY reuse LSN external identifier after an ICMP Query session timer expires.
>
> b) A LSN SHOULD limit the number of the LSN external identifier allocated per CPE.
>
> c) 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-40: 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.

---

7. 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.

7.1. 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

7.2. 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.

---

8. Considerations about limiting the number of LSN external ports

TOC

As discussed in section 3,4 and 5, 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.

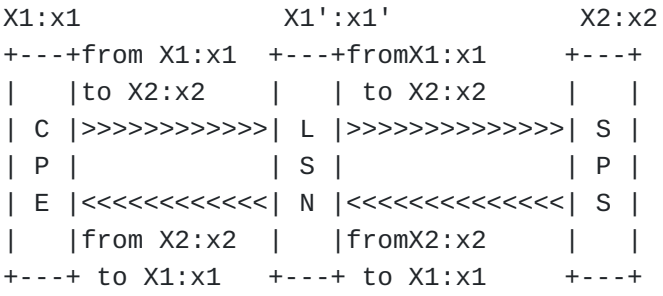

Figure 3. LSN pass-through

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.

9. IANA Considerations

TOC

There are no IANA considerations.

10. 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.

11. Acknowledgements

TOC

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 and Dan Wing.

12. References

TOC

12.1. Normative References

TOC

[RFC0792]	Postel, J., "Internet Control Message Protocol," STD 5, RFC 792, September 1981 (TXT).
[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).

12.2. 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-05 (work in progress), July 2010 (TXT).
[I-D.ymbk-aplusp]	Bush, R., "The A+P Approach to the IPv4 Address Shortage," draft-ymbk-aplusp-05 (work in progress), October 2009 (TXT).

Authors' Addresses

TOC

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