Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 2, 2012

X. Deng
M. Boucadair
France Telecom
C. Zhou
Huawei Technologies
T. Tsou
Huawei Technologies (USA)
G. Bajko
Nokia

July 01, 2011

DS-Lite AFTR NAT Bypass: Co-located B4 and NAT Model
draft-zhou-softwire-b4-nat-02

Abstract

This document describes the behavior of the B4 when co-located with a NAT while the NAT in the AFTR is disabled. The proposed solution is expected to offload the burden on the AFTR, by delegating the NAT to B4.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

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 2, 2012.

Copyright Notice

Table of Contents

1. Introduction 3
2. B4 Behavior 3
2.1. Provisioning 3
2.2. Plain IPv4 Address 3
2.3. Restricted IPv4 Address 3
2.3.1. Incoming Ports on a given restricted IPv4 address . . . 3
2.3.2. Outgoing Packets Processing 4
2.3.3. Incoming Packets Processing 4
2.4. Stateless Encapsulation 4
2.5. Fragmentation and Reassembly 4
2.6. DNS 4
3. Security Considerations 4
3.1. Port Randomization and non-contiguous port sets allocation mechanism 4
4. IANA Considerations 6
5. References 6
5.1. Normative References 6
5.2. informative References 7
Authors' Addresses 7

1. Introduction

As currently defined in [I-D.ietf-softwire-dual-stack-lite], B4 element SHOULD NOT operate a NAT function because the NAT function will be performed by the AFTR in the service provider's network. To reduce the processing requirement of NAT device at the network side, address and port translation can be made at the customer side, e.g., CPE. For convenience, we call this solution as NAT-Bypass.

This document provides descriptions on the B4 behavior when supporting NAT-Bypass.

2. B4 Behavior

2.1. Provisioning

The provisioning of the B4 element is similar to what is defined in [I-D.ietf-softwire-dual-stack-lite].

2.2. Plain IPv4 Address

A B4 MAY be assigned with a plain IPv4 address.

When a plain, IPv4 address is assigned, the NAT operations are enforced as per current legacy CPEs. The NAT in the AFTR is disabled for that user.

IPv4 datagrams are encapsulated in IPv6 as specified in [I-D.ietf-softwire-dual-stack-lite].

2.3. Restricted IPv4 Address

In the NAT-Bypass solution, the port set is provisioned to B4 through PCP option defined in [I-D.tsou-pcp-natcoord] or specific DHCP options [I-D.bajko-pripaddrassign].

The PCP Server or IPv4 DHCP server may be co-located with the AFTR.

The B4 is responsible for performing NAT and/ALG functions, as well as supporting NAT Traversal mechanisms (e.g., UPnP or NAT-PMP).

2.3.1. Incoming Ports on a given restricted IPv4 address

As described in [I-D.ietf-intarea-shared-addressing-issues], a bulk of incoming ports can be reserved as a centralized resource shared by all subscribers using a given restricted IPv4 address. In order to distribute incoming ports as fair as possible among subscribers

sharing a given restricted IPv4 address, other than allocating a continuous range of ports to each, a solution to distribute bulks of non-continuous ports among subscribers, which also takes port randomization into account, is elaborated in Section 3.1.

2.3.2. Outgoing Packets Processing

Upon receiving an IPv4 packet, the B4 performs NAT using the public IPv4 address and port set assigned to it. Then B4 encapsulates the resulting IPv4 packet into an IPv6 packet, and delivers it through IPv6 connectivity to AFTR which will then decapsulate the encapsulated packet and forward it through IPv4. The destination IPv6 address used for encapsulation should be the AFTR's address.

2.3.3. Incoming Packets Processing

Upon receipt of IPv4-in-IPv6 packet from AFTR, B4 will decapsulate the packet and translate the public IPv4 address to the private IPv4 address. Finally, it delivers the packet to the host using the translated IPv4 address. The source IPv6 address used for encapsulation at AFTR is the AFTR's address, and the destination address is set to the external address of B4.

2.4. Stateless Encapsulation

B4 may implement the stateless encapsulation specified in Section 4.4 of [I-D.ymbk-aplusp].

2.5. Fragmentation and Reassembly

No change to Section 5.3 of [I-D.ietf-softwire-dual-stack-lite].

2.6. DNS

The DNS behavior is the same as described in [I-D.ietf-softwire-dual-stack-lite].

3. Security Considerations

3.1. Port Randomization and non-contiguous port sets allocation mechanism

As port randomization is one protection among others against blind attacks, a simple non-contiguous port sets distribution mechanism is therefore proposed to distribute bulks of non-continuous ports among subscribers, and to enable subscribers operating port randomized NAT.

On every external IPv4 address, according to port set size N, log2(N) bits are randomly choosing by AFTR as subscribers identification bit (s bit) among 1st and 16th bits. Take a sharing ration 1:32 for example, Figure 1 shows an example of 5 random selected bits of s bit.

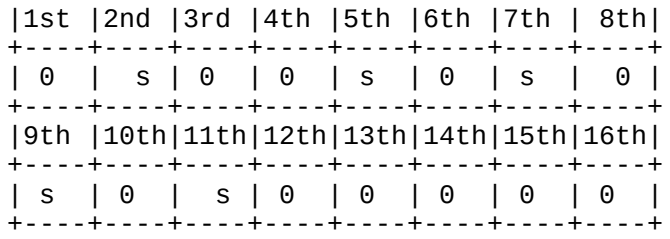

Figure 1: A s bit selection example (on a sharing ration 1:32 address).

Subscriber ID pattern is formed by setting all the s bits to 1 and other trivial bits to 0. Figure 2 illustrates an example of subscriber ID pattern on a sharing ration 1:32 address. Note that the subscriber ID pattern will be different, guaranteed by the random s bit selection, on every restricted IP address no matter whether the sharing ratio varies.

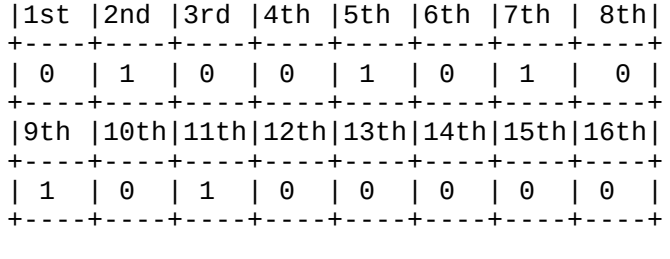

Figure 2: A subscriber ID pattern example (on a sharing ration 1:32 address).

Subscribers ID value is then assigned by setting subscriber ID pattern bits (s bits shown in the following example) according to a customer value and setting other trivial bits to 1.

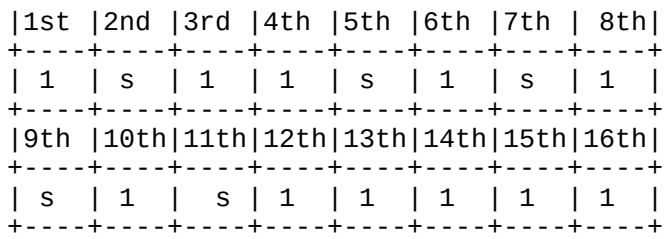

Figure 3: A subscriber ID value example (0# subscriber on this restricted address).

Subscriber ID pattern and subscriber ID value together uniquely defines a non-overlapping port set on a restricted IP address.

Pseudo-code shown in the Figure 4 describe how to use subscriber ID pattern and subscriber ID value to implement a random ephemeral port selection function in a restricted port set.

```
do{
    restricted_next_ephemeral = (random()| customer_ID_pattern)
                                & customer_ID_value;
    if(five-tuple is unique)
    return restricted_next_ephemeral;
}
```

Figure 4: Random ephemeral port selection of restricted port set algorithm.

4. IANA Considerations

None.

5. References

5.1. Normative References

[I-D.ietf-softwire-dual-stack-lite]
Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion (Work in progress)", May 2011.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

5.2. informative References

[I-D.bajko-pripaddrassign]
Bajko, G., Savolainen, T., Boucadair, M., and P. Levis, "Port Restricted IP Address Assignment(Work in progress)", September 2010.

[I-D.ietf-intarea-shared-addressing-issues]
Ford, M., Boucadair, M., Durand, A., Levis, P., and P. Roberts, "Issues with IP Address Sharing(Work in progress)", March 2011.

[I-D.tsou-pcp-natcoord]
Tsou, T., Zhou, C., Sun, Q., Boucadair, M., and G. Bajko, "Using PCP To Coordinate Between the CGN and Home Gateway Via Port Allocation (Work in progress)", March 2011.

[I-D.ymbk-aplusp]
Bush, R., "The A+P Approach to the IPv4 Address Shortage(Work in progress)", February 2011.

Authors' Addresses

Xiaohong Deng
France Telecom

Email: xiaohong.deng@orange-ftgroup.com

Mohamed Boucadair
France Telecom
Rennes, 35000
France

Email: mohamed.boucadair@orange-ftgroup.com

Cathy Zhou
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R. China

Phone:
Email: cathyzhou@huawei.com

Tina Tsou
Huawei Technologies (USA)
2330 Central Expressway
Santa Clara, CA 95050
USA

Phone: +1 408 330 4424
Email: tena@huawei.com

Gabor Bajko
Nokia

Email: gabor.bajko@nokia.com