Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: March 1, 2018

H. Xu
Q. Sun
China Telecom
Y. Fu
CNNIC

August 28, 2017

A Redundancy Mechanism for Dual-Stack Lite
draft-xu-v6ops-dslite-redundancy-01

Abstract

Dual-Stack Lite is a solution to offer both IPv4 and IPv6 connectivity to customers that are addressed only with an IPv6 prefix. This document provide a redundancy mechanism for Dual-Stack Lite.

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 March 1, 2018.

Table of Contents

1. Introduction

Dual-Stack Lite [RFC6333] is a solution to offer both IPv4 and IPv6 connectivity to customers crossing an IPv6 only infrastructure. The internet service provider no longer to provide publice IPv4 address but an IPv6 prefix to the customers as the issue of the IPv4 public address shortage. One of its key components is an IPv4-over-IPv6 tunnel, which is used to provide IPv4 connectivity across a service provider's IPv6 network. Another key component is a carrier-grade IPv4-IPv4 Network Address Translation (NAT) to share service provider IPv4 addresses among customers. As the exhaustion of the public IPv4 address, service providers have deployed DS-Lite in their network widely in nowadays, where a large number of customers are located. These customers within a network which is served by a single CGN function embedded in AFTR element may experience service degradation due to the presence of the single point of failure or loss of state information. Therefore, redundancy capabilities of the AFTR devices are strongly desired in order to deliver highly available services to customers. Failure detection and repair time should be therefore shortened.

This document describes a redundancy mechanism for DS-Lite. Some deployment consideration and recommendations for network elements are also provided.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] when they appear in ALL CAPS. When these words are not in ALL CAPS (such as "should" or "Should"), they have their usual English meanings, and are not to be interpreted as [RFC2119] key words.

3. Reliability Considerations of AFTR

As described in [RFC6908], for the robustness, reliability, and load distribution purposes, operators may deploy multiple AFTRs in their network. There are many deployment mechanism for the AFTR in ISP network, the most common type are distribution mode and centralization mode.

For the distribution mode, the CGN card is integrated into the free slot of the BRAS in a metro network. As the BRAS integrates the AFTR function of DS-lite, it provides DS-Lite connection service for a small area customers in this metro network. The service providers always integrated two CGN cards in the BRAS for reduncdancy consideration as the primary AFTR and backup AFTR.The capital cost of this mode is expensive because it always need two CGN cards for every BRAS. But 50 percent of these cards are idle most of time so that it is a big waste of money. There are various types and versions of BRAS have been deployed in the service provider’s network . Some of them have been used for over ten years and may not support the card insertion. Some of them may also don’t have free slot for the CGN card. It is not operational to replace all of them in a short period which result that it could deploy DS-Lite in some area and others can not in the same metro network.

For the centralization mode, a stand-alone AFTR device is deployed nearby the core router device at the exit of a metro network. It provides the DS-Lite connection service for the whole customers in this metro network. Service providers always deploy two stand-alone AFTR devices nearby the two core router device for the load distribution and redundancy purpose. The capital cost of this mode is more less than the distribution mode. It does not consume the slot resource of the BRAS. But it takes a big chanllenge for AFTR device for this mode in the large scale metro network because it takes performance requirements for the speed of the session creation and the maximum number of session maintenance. On the other side, it will create exta traffic when the users belong to the same BRAS are conmmunicating with eath other because it will connect to the AFTR device in the centralization mode first. It is a waste of bandwidth.

As described above, whether to use distribution mode or centralization mode depends on the trade-off between the investment and operational efficiency requirement of the service providers.

4. The Redundancy Machanism Overview

The fundamental principle of redundacncy machanism is to make the centralization mode to backup for the distribution mode. The architecture of the redundancy mechanism is illustrated as Figure 1. It deploys one AFTR card into every BRAS which surport card insertion in metro network, as to provide bassic distributed DS-Lite connection service. Moreover, it deploy two stand-alone AFTR device near the core router at the exit of the metro network. So it could provide the DS-lite connection service for the users of the BRAS which don’t surpport card insertion and don’t have free slot for the AFTR card. One advantage of this mechanism is that the stand-alone AFTR device is not only a redundancy device but also can provide DS-Lite conection service for the BRAS without AFTR card slot. Then the IGP routing would be configured on the BRAS which has the AFTR card insertion.

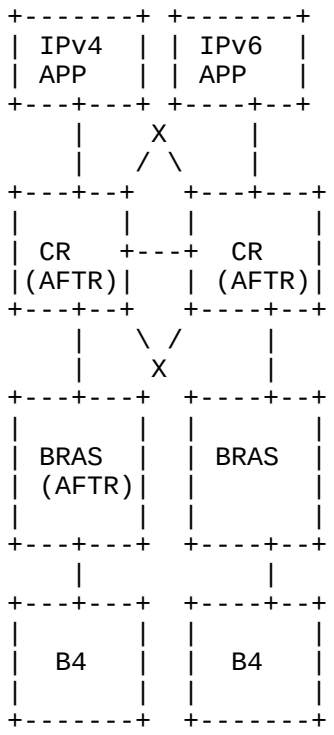


Figure 1: The architecture of the redunancy mechanism

It is made that the routing prior selected to the AFTR card on the BRAS and then selected the AFTR stand-alone device near the core router through the Metric value configuration.As the metric values of the two stand-alone AFTR device in centralization mode are the same, it ensure that the traffic of the same session would be forwarded to the same centralized AFTR device by the random selection of the hash algorithm.This mechanism is based on the IPv6 anycast function: when the AFTR card in distribution mode is breakdown,the AFTR address in router advertise message will disappear in the IGP routing table.The IP address of AFTR device in centralization mode is becoming the optimal routing. All the traffic for DS-Lite will be directed to the AFTR decive in the centralization mode as to keep the application alive.

5. The difference between the software process of the BRAS

The software process of the BRAS for distribution mode is described as Figure 2

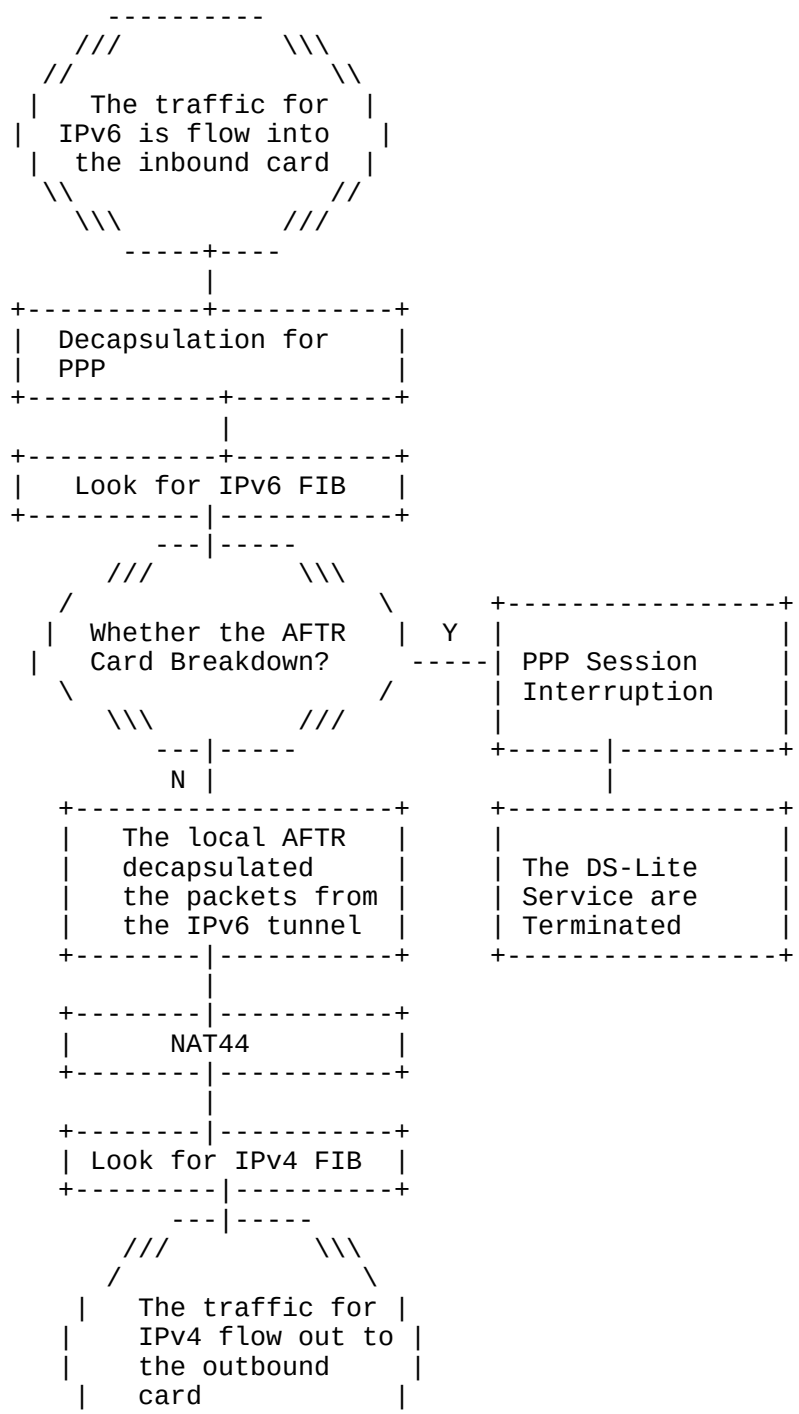
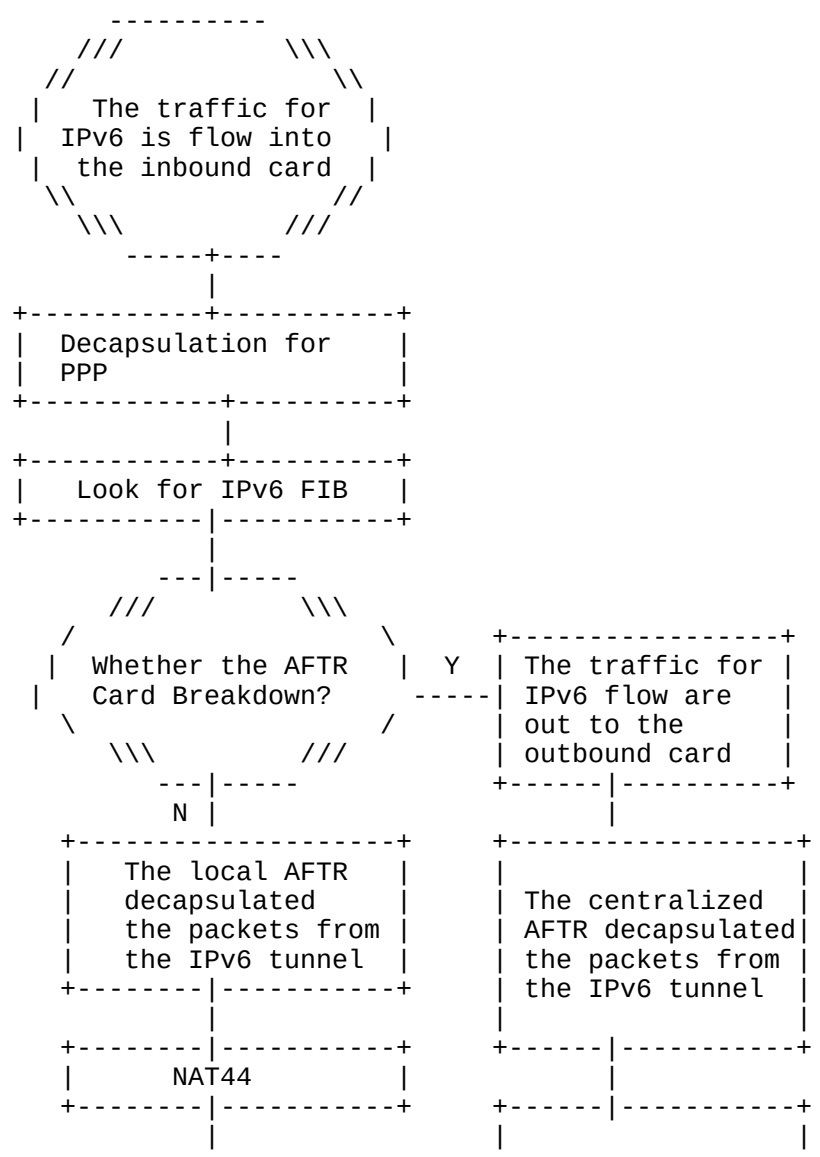

The traffic for IPv6 is flow into the inbound card
Decapsulation for PPP
Look for IPv6 FIB
Whether the AFTR Card Breakdown?
Y
PPP Session Interruption
N
The local AFTR decapsulated the packets from the IPv6 tunnel
The DS-Lite Service are Terminated
NAT44
Look for IPv4 FIB
The traffic for IPv4 flow out to the outbound card

Figure 2: The software process of the BRAS for distribution mode

And the software process of the BRAS for the new machanism is described as Figure 3:



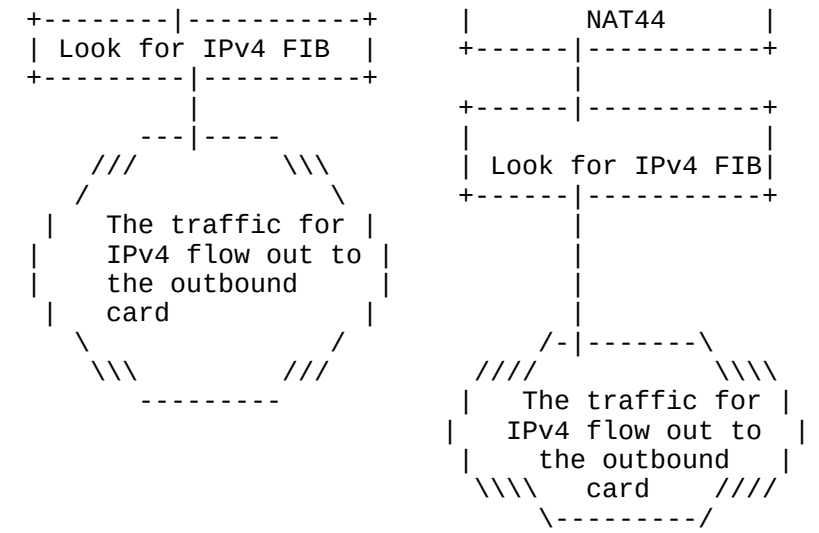


Figure 3: The software process of the BRAS for new mechanism

As compared between Figure 2 and Figure 3, the main difference for the new mechanism is that if the local AFTR card breakdown, the DS-Lite service can be maintained as the backup AFTR will take over the function to keep the application alive.

6. New requirements for the AFTR device

For this DS-Lite redundancy mechanism, there are some new requirements for the AFTR device as below:

1. If the ditribution AFTR card breakdown, the AFTR device SHOULD ensure that the traffic will not direct to the other distribution AFTR card.

2. It should use FQDN to decribe the AFTR in the DHCPv6 option as described in [RFC6334].

3. How many ditribution AFTR device could be covered by one centralization AFTR device will be different depends on the deployment by different ISPs.

4. The speed of the session creation for the centralized AFTR device could be calculated by a formula.

7. Security Considerations

The AFTR device of centralization mode will accept the tunnel request from the all DS-Lite users in the metro network. It needs aditional requirements to prevent from the spoofing attack.

1. Only the user passed the authentication could be assigned IPv6 prefix from the BRAS.

2. After assigned the IPv6 prefix to the authorized user, the BRAS will report this address to the AAA sever for recording.

3. Create a local database in the AFTR device of he centralized mode to record the IPv6 prefix of the authorized user.

4. Create an interface of the AAA sever for the AFTR device to synchrionize the IPv6 prefix of the authorized user between the AAA sever to the local database of the AFTR.

5. When the BRAS receive a new request for a new tunnel, it will compare with the source IPv6 prefix with the local database of the AFTR. If it is match, it will accept the request for tunnel. If not, it will ignore the request regarding it is from a illegal user and report the illegal address to the network management system.

6. If the authorized user offline, BRAS will ask the AAA server to delete this user from the database.

8. IANA Considerations

This draft does not request any IANA action.

9. Acknowledgements

The authors would like to thanks the valuable comments made by XXX and other members of v6ops WG.

This document was produced using the xml2rfc tool [RFC2629].

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, DOI 10.17487/RFC6333, August 2011, <https://www.rfc-editor.org/info/rfc6333>.

[RFC6334] Hankins, D. and T. Mrugalski, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Option for Dual-Stack Lite", RFC 6334, DOI 10.17487/RFC6334, August 2011, <https://www.rfc-editor.org/info/rfc6334>.

10.2. Informative References

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, DOI 10.17487/RFC2629, June 1999, <https://www.rfc-editor.org/info/rfc2629>.

[RFC6908] Lee, Y., Maglione, R., Williams, C., Jacquenet, C., and M. Boucadair, "Deployment Considerations for Dual-Stack Lite", RFC 6908, DOI 10.17487/RFC6908, March 2013, <https://www.rfc-editor.org/info/rfc6908>.

Authors' Addresses

Honglei Xu
China Telecom
No.118 Xizhimennei street, Xicheng District
Beijing, 100035
P.R. China

Email: xuhl.bri@chinatelecom.cn

Qiong Sun
China Telecom
No.118 Xizhimennei street, Xicheng District
Beijing 100035
P.R. China

Email: sunqiong.bri@chinatelecom.cn

Yu Fu
CNNIC
No.4 South 4th Street, Zhongguancun
Hai-Dian District, Beijing, 100190
P.R. China

Email: fuyu@cnnic.cn