Network Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: April 30, 2015

S. Vinapamula
Juniper Networks
M. Boucadair
France Telecom

October 27, 2014

Recommendation for Prefix Binding in the Softwire DS-Lite Context
draft-vinapamula-softwire-dslite-prefix-binding-03

Abstract

This document discusses issues induced by the change of the Basic Bridging BroadBand (B4) IPv6 address and sketches a set of recommendations to solve those issues.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

IPv6 deployment models assumes IPv6 prefixes are delegated by Service Providers to the connected CPEs (Customer Premise Equipments) or hosts, which in their turn derive IPv6 addresses out of that prefix. In the case of DS-Lite [RFC6333], the Basic Bridging BroadBand (B4) element derives an IPv6 address for the softwire setup purposes.

A B4 element might obtain a new external IPv6 address, for a variety of reasons including a reboot of the CPE, power outage, DHCP lease expiry, or other action undertaken by the Service Provider. If this occurs, traffic forwarded to a B4's previous address might be delivered to another B4 that now acquired that address. This affects all mapping types, whether implicit (e.g., by sending a TCP SYN) or explicit (e.g., using PCP [RFC6887]).

The main goal of this document is to propose recommendations to soften the impact of such renumbering issues.

Note that in some deployments, CPE renumbering may be required to accommodate some privacy-related requirements to avoid the same prefix be assigned to the same customer. It is out of scope of this document to discuss such contexts.

This document complements [RFC6908].

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].

3. The Problem

Since the network behind B4 can be overlapping across multiple CPEs, B4 address plays a key role in identifying associated resources assigned for each of the connections. These resources maintain state of Endpoint-Independent Mapping (EIM), Endpoint-Independent Filtering (EIF), preserve external IPv4 address assigned in the AFTR, and PCP mappings and flows.

However, there can be change in B4 address for any reason, may be because of change in CPE device or may be because of security extensions enabled in generating the IPv6 address. When the address change, the associated mappings created in the AFTR are no more valid. This may result in creation of new set of mappings.

Furthermore, a mis-behaving user may be tempted to change the B4's IPv6 address in order to "grab" more ports and resources at the AFTR side. This behavior can be seen as a potential DoS attack from mis-behaving users. Note that this DoS attack can be achieved whatever port assignment policy configured to the AFTR (individual ports, port sets, randomized port bulks, etc.).

Service Providers may want to limit the usage of these resources on per subscriber basis for fairness resources usage. To that aim , a subscriber is identified by the delegated IPv6 prefix and not the derived B4 address. These policies are used for dimensioning purposes and also to ensure that AFTR resources are not exhausted. However when there is a change in B4 address, this policy doesn't resolve stale mappings hanging around in the system, consuming not only system resources, but also reducing the available quota of resources per subscriber.

Clearing those mappings can be envisaged, but that will cause a lot of churn in the AFTR and could be disruptive to existing connections.

When services are hosted behind B4 element, and when there is a change in B4 address which if results in change in NAT address, these services have to advertise about their change, whenever there is a change of the B4 address. Means to discover the change of B4 address and NAT address is therefore required. Also, it doesn't address latency issues where a service has to advertise its newly assigned external IP address and port and the clients have to consume and re-initiate connections.

PCP-specific failure scenarios are discussed in [I-D.boucadair-pcp-failure].

4. Recommendations

In order to mitigate the issues discussed in Section 3, the following recommendations are made:

1. A policy SHOULD be enforced at the AFTR level to limit the number of active softwires per subscriber. The default value MUST be 1. This policy aims to prevent a misbehaving subscriber to mount several softwires to consume more resources on the AFTR side.

2. Resource contexts created at the AFTR level SHOULD be based on the delegated IPv6 prefix and not based on the B4 address. Delegated prefix may be derived from the B4 address through a configured subscriber-mask. Administrators SHOULD configure per prefix limits of resource usage, instead of per tunnel limits. These resources include, number of flows, mappings including PCP, NAT pool resources, etc.

    1. Subscriber-mask is an integer that indicates the length of significant bits to be applied on the source IPv6 address (internal side) to identify a subscriber. Subscriber-mask is an AFTR system-wide configuration parameter that is used to enforce generic per-subscriber policies. Applying these generic policies does not require to configure every subscriber prefix. Subscriber-mask must be configurable; the default value is 56.

    2. For example, suppose an IPv6 prefix 2001::/56 is delegated to a CPE. Administrator should configure resource usage limits in AFTR based on the prefix 2001::/56 and not based on any B4 address derived from the delegated prefix. AFTR will derive the prefix from B4 address through configured subscriber-mask set to 56 by the administrator.

3. In the event a new IPv6 address is assigned to B4, the AFTR SHOULD migrate existing state to be bound to the new B4's IP address. This ensures the traffic destined to the previous B4 address will be redirected to the newer B4 address. The destination address for tunneling return traffic SHOULD be the last seen as B4's address from the CPE. Doing so avoids stale mappings and minimizes the risk of service disruption.

4. In the event of change of the CPE WAN's IPv6 prefix, unsolicited PCP ANNOUNCE messages are to be sent by the B4 element to internal hosts to update their mappings. This allows internal PCP clients to update their mappings with the new B4 IPv6 address and trigger updates to rendez-vous servers (e.g., dynamic DNS).

5. When a new prefix is assigned to the CPE, stale mappings may exist in the AFTR. This will consume both implicit and explicit resources. In order to avoid such issues, stable IPv6 prefix assignment is RECOMMENDED.

6. In case for any reason an IPv6 prefix has to be reassigned, it is RECOMMENDED to reassign an IPv6 prefix only when all the resources in use associated with that prefix are cleared from the AFTR. Doing so avoids to redirect traffic, destined to the previous prefix owner, to the new one.

5. Security Considerations

Security considerations related to DS-Lite are discussed in [RFC6333].

Enforcing the recommendations in Section 4 defends against DoS attacks that would result in varying the source IPv6 address to exhaust AFTR resources.

6. IANA Considerations

This document does not require any action from IANA.

7. Acknowledgements

G. Krishna and C. Jacquenet reviewed document and provided useful comments.

Thanks to I. Farrer for the comments.

8. References

8.1. Normative references

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

[RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, August 2011.

[RFC6887] Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", RFC 6887, April 2013.

8.2. Informative references

[I-D.boucadair-pcp-failure]
Boucadair, M. and R. Penno, "Analysis of Port Control Protocol (PCP) Failure Scenarios", draft-boucadair-pcp-failure-06 (work in progress), May 2013.

[RFC6908] Lee, Y., Maglione, R., Williams, C., Jacquenet, C., and M. Boucadair, "Deployment Considerations for Dual-Stack Lite", RFC 6908, March 2013.

Authors' Addresses

Suresh Vinapamula
Juniper Networks
1194 North Mathilda Avenue
Sunnyvale, CA 94089
USA

Phone: +1 408 936 5441
EMail: sureshk@juniper.net

Mohamed Boucadair
France Telecom
Rennes 35000
France

EMail: mohamed.boucadair@orange.com