Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 12, 2012

Y. Cui
J. Wu
P. Wu
Tsinghua University
Q. Sun
C. Xie
China Telecom
C. Zhou
Huawei Technologies
July 11, 2011

Lightweight 4over6 in access network
draft-cui-softwire-b4-translated-ds-lite-01

Abstract

The dual-stack lite mechanism provide an IPv4 access method over IPv6 ISP network for end users. Dual-Stack Lite enables a broadband service provider to share IPv4 addresses among customers by combining IPv4-in-IPv6 tunnel and Carrier Grade NAT. However, in dual-stack lite. CGN has to maintain per-session address mapping, which could be a heavy burden, and produce high hardware cost as well as performance issue. This draft propose the lightweight 4over6 mechanism which moves the translation function from tunnel concentrator (AFTR) to initiators (B4s), and hence reduce the mapping scale on CGN to per-customer level. For translation usage, the mechanism will allocate port restrcted IPv4 addresses to initiators in a flexible way independent of IPv6 network in the middle.

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 12, 2012.

Copyright Notice

Table of Contents

1. Introduction

Dual-stack lite technology provides IPv4 access method over IPv6 ISP network for broadband users. To deal with the incoming IPv4 exhaustion, dual-stack lite embeds the IPv4 address sharing functionality into the mechanism, by putting an IPv4 CGN on the tunnel concentrator (AFTR). The 44CGN solves the address shortage problem at the cost of maintaining per-session address mapping.

The common estimation of AFTR capacity is that one AFTR should serve thousands of end users. The huge amount of sessions from the users would cause a fairly high hardware cost, and could easily cause performance issues. Besides, if the CGN need to support source tracing, the volume of logging data will be really huge. Therefore it’s of great significance if we can reduce the amount of address mappings effectively.

The address mapping is used directly for IPv4 private-public translation. So in order to reduce the amount of address mappings on the concentrator, the straightforward approach is moving the translation from the concentrator to initiators. Then the concentrator will be only responsible for encapsulation and forwarding, and hence get rid of the burden of translation and state maintenance. Apparently an initiator can be capable of translating its own traffic since the volume is low. However, it does need the public address and port for translation allocated from the concentrator. [I-D.cui-softwire-host-4over6] has discussed the case where full IPv4 addresses are allocated over IPv6 for IPv4-over-IPv6 usage; this draft will cover the case where port space is divided and allocated, too.

2. 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 [RFC2119].

3. Terminology

Lightweight 4over6: lightweight 4over6 is an IPv4-over-IPv6 hub and spoke mechanism proposed in this document, which supports address sharing to deal with IPv4 address exhaustion, and places the IPv4 translation on the intiator side.

Lightweight 4over6 initiator (or "initiator" for short): the tunnel initiator in lightweight 4over6 mechanism. The lightweight 4over6 initiator could be a host directly connected to IPv6, or an dual-stack CPE in front of an IPv4 local network. It is responsible for IPv4 public-private translation besides tunnel encapsulation and decapsulation.

Lightweight 4over6 concentrator (or "concentrator" for short): the tunnel concentrator in lightweight 4over6 mechanism. The lightweight 4over6 concentrator connects the ISP IPv6 network and IPv4 Internet. It provides tunnel encapsulation and decapsulation but no IPv4 public-private translation.

4. Port Restricted IPv4 Address Allocation

As is described above, a lightweight 4over6 initiator needs the public address and port for stateful translation. It's obvious that individual address + port allocation when required is not efficient due to round-trip time delay and high signaling volume caused. Besides, that way can't save the mapping amount at all. A practical manner would be allocating a group of address + port at one time beforehand, i.e., allocating IPv4 address with a port range, which is known as "restricted address".

To our knowledge there're two methods which are probably suitable for restricted address allocation from concentrator to initiator. One is extending DHCP to support address allocation with port range embedded. [I-D.bajko-pripaddrassign] discusses this DHCP usage. In this special context, we need to build the DHCPv4 procedure over IPv6 [I-D.cui-softwire-dhcp-over-tunnel]. The other is extending PCP to support port range control. See [I-D.tsou-pcp-natcoord] for details. Adopting either method, An initiator can get port restricted IPv4 addresses allocated dynamically from the concentrator. Unlike stateless 4over6 solutions like [I-D.murakami-softwire-4rd], the port restricted address allocation in lightweight 4over6 has no requirement on IPv6 address format, i.e. IPv4 and IPv6 addressing remain separated. This is more close to today's address allocation mode and provides flexibility for ISPs to manage the access network.

5. Lightweight 4over6 Initiator Behavior

An lightweight 4over6 initiator should support either extended DHCP client, or extended PCP client, as is described above. When requiring IPv4 access, the initiator would run either client to get a dynamic port restricted IPv4 address, and renew/extend it when the lease/lifetime is about to expire.

The data plane functions of the initiator includes translation and encapsulation/decapsulation. An initiator runs a standard local NAT44 with the address pool consist of the allocated port restricted address. When sending out an IPv4 packet with private source address, it performs NAT44 function and translates the source address into public. Then it encapsulate the packet with concentrator's IPv6 address as destination IPv6 address, and forward it to the concentrator. When receiving an IPv4-in-IPv6 packet from the concentrator, the initiator decapsulates the IPv6 packet to get the IPv4 packet with public destination IPv4 address. Then it performs NAT44 function and translates the destination address into private.

For host initiator case, it's also feasible that the host doesn't run the local NAT and uses the allocated public IPv4 address directly. Then the host should guarantee that every port number in the packets sent out by itself falls into the allocated port range.

6. Lightweight 4over6 Concentrator Behavior

The lightweight 4over6 concentrator should support either an extended DHCPv4 server, or an extended PCP server, to allocate port restricted addresses. When accomplishing one such allocation, the concentrator simultaneously install a mapping entry into the mapping table. This entry consists of the public IPv4 address, the port range and initiator's IPv6 address. Its lifetime is set according to the allcation. It'll be used for encapsulation of inbound packets. This mapping entry would be deleted when the lifetime expires, and refresh its lifetime when the initiator renews/extends the allocation.

The data plane functions of the concentrator is purely encapsulation and decapsulation. When receiving an IPv4-in-IPv6 packet from an initiator, the concentrator decapsulates it and forwards it to IPv4 Internet. When receiving an IPv4 packet from the Internet, it uses the destination address and port from this packet to lookup the destination initiator's IPv6 address in the mapping table. Then the concentrator encapsulates this packet using the IPv6 address found in the table as IPv6 destination address, and forwards it to the correct initiator.

7. Mechanism Analysis

Compared with original dual-stack lite, lightweight 4over6 removes the translation burden from the concentrator and distribute the job to initiators on user-side. Also it decreases the state scale on concentrator from per-session level down to per-user level. This would significantly reduce the hardware cost of the concentrator, and the probability that the concentrator become the performance bottleneck. Leveraging lightweight 4over6, one concentrator can serve a lot more customers.

Besides, since this mechanism reduces the number of simultaneous address mappings of each customer on concentrator to one, it makes concentrator logging much more feasible. Moreover, locate the translation on user side eases the ALG and NAT referral problem since it'll be no different from the situation of local NAT in today's IPv4 network. Various solutions already exsit for quite a long time.

Lightweight 4over6 allocates port restricted address independent of the IPv6 network in the middle. No specific IPv6 address format is required. IPv4 and IPv6 addressing and routing remain separated. This is close to today's address allocation mode. The ISP can provision IPv4 in a flexible, on-demand way, as well as manage the native IPv6 network without the influence of IPv4-over-IPv6 requirements.

The costs of lightweight 4over6 for achieving all these benifits are lower IPv4 address utilization ratio and extra signaling behavior. The address multiplexing manner of port restricted address is relatively more static than the CGN manner. And dual-stack lite doesn't require address and port allocation between concentrator and initiators. When compared to stateless solutions, lightweight 4over6 still keeps per-user states rather than becoming purely stateless.

Besides, ICMP ping would become a challenge in port restricted address environment. The solution is to divide ICMP id field in the same way with dividing port space, i.e. each initiator will get the IMCP id range which is identical to the allocated port range. This way the initiator uses the allocated address and restricted id range when send out a ping. Hence ping "session" intiated from a lightweight 4over6 initiator can be processed correctly by the concentrator. The inbound ping would be left unsupported on the concentrator, which is the similar behavior of NAT/CGN. See details about ICMP ping porcessing in section 4.2 of [I-D.sun-v6ops-laft6].

8. References

[I-D.bajko-pripaddrassign]
Bajko, G., Savolainen, T., Boucadair, M., and P. Levis, "Port Restricted IP Address Assignment", draft-bajko-pripaddrassign-03 (work in progress), September 2010.

[I-D.cui-softwire-dhcp-over-tunnel]
Cui, Y., Wu, P., and J. Wu, "DHCPv4 Behavior over IP-IP tunnel", draft-cui-softwire-dhcp-over-tunnel-00 (work in progress), June 2011.

[I-D.cui-softwire-host-4over6]
Cui, Y., Wu, J., Wu, P., Metz, C., Vautrin, O., and Y. Lee, "Public IPv4 over Access IPv6 Network", draft-cui-softwire-host-4over6-06 (work in progress), July 2011.

[I-D.ietf-softwire-dual-stack-lite]
Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", draft-ietf-softwire-dual-stack-lite-11 (work in progress), May 2011.

[I-D.murakami-softwire-4rd]
Murakami, T. and O. Troan, "IPv4 Residual Deployment on IPv6 infrastructure - protocol specification", draft-murakami-softwire-4rd-00 (work in progress), July 2011.

[I-D.sun-v6ops-laft6]
Sun, Q. and C. Xie, "LAFT6: Lightweight address family transition for IPv6", draft-sun-v6ops-laft6-01 (work in progress), March 2011.

[I-D.tsou-pcp-natcoord]
Zhou, C., ZOU), T., Deng, X., Boucadair, M., and Q. Sun, "Using PCP To Coordinate Between the CGN and Home Gateway Via Port Allocation", draft-tsou-pcp-natcoord-03 (work in progress), July 2011.

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

Authors' Addresses

Yong Cui
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R.China

Phone: +86-10-62603059
Email: yong@csnet1.cs.tsinghua.edu.cn

Jianping Wu
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R.China

Phone: +86-10-62785983
Email: jianping@cernet.edu.cn

Peng Wu
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R.China

Phone: +86-10-62785822
Email: weapon@csnet1.cs.tsinghua.edu.cn

Qiong Sun
China Telecom
Room 708, No.118, Xizhimennei Street
Beijing 100035
P.R.China

Phone: +86-10-58552936>
Email: sunqiong@ctbri.com.cn

Chongfeng Xie
China Telecom
Room 708, No.118, Xizhimennei Street
Beijing 100035
P.R.China

Phone: +86-10-58552116>
Email: xiechf@ctbri.com.cn

Cathy Zhou
Huawei Technologies
Section B, Huawei Industrial Base, Bantian Longgang
Shenzhen 518129
P.R.China

Phone: +86-10-58552116>
Email: cathyzhou@huawei.com