Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 17, 2013

Y. Cui
J. Wu
P. Wu
Tsinghua University
O. Vautrin
Juniper Networks
Y. Lee
Comcast

July 16, 2012

Public IPv4 over IPv6 Access Network
draft-ietf-softwire-public-4over6-02

Abstract

When the service provider networks are upgraded to IPv6, IPv4 connectivity will still be demanded by network users, at least in the recent future. This draft proposes a mechanism for end hosts or networks in IPv6 access networks to build bidirectional IPv4 communication with the IPv4 Internet. The mechanism follows the softwire hub and spoke model, and uses IPv4-over-IPv6 tunnel as basic method to traverse IPv6 network. The bi-directionality of end-to-end communication is achieved by allocating public IPv4 addresses to end hosts/networks, with no dependency on IPv6 addressing.

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 17, 2013.

Copyright Notice

Table of Contents

1. Introduction

IANA has exhausted the Global IPv4 address pool, while the RIRs are running out of IPv4 addresses. On the other hand, the size of Internet is still growing fast, as well as the demand for IP addresses. To satisfy the address demand from end users, operators have to push IPv6 to the front, by building IPv6 networks and providing IPv6 services.

When IPv6-only networks are widely deployed, users of those networks will probably still need IPv4 connectivity. This is because part of Internet will stay IPv4-only for a long time, and network users in IPv6-only networks will communicate with network users sited in the IPv4-only part of Internet. This demand could eventually decrease with the general IPv6 adoption.

Therefore, network operators should provide IPv4 services to IPv6 users, usually through tunnel. There are two types of this tunneled IPv4 services, differed in provisioned IPv4 addresses. If the operators cannot provision public IPv4 addresses, the user side can only use private IPv4 addresses, and NAT44 translation is required on the carrier side, as is described in Dual-stack Lite[RFC6333]. Otherwise the operators are capable of provisioning public IPv4 addresses. Then users can directly employ these addresses for IPv4 communication, and carrier-side translation is not needed anymore. The network users and operators can avoid all the issues raised by translation, such as ALG, NAT traversal, session state maintenance, etc.

In the second type, there have also been efforts on provisioning port-set rather than full address to individual users. Supporting port-set would require extra extensions on the provision protocol, end-user behavior and data plane function. On the other hand, from the ISPs' perspective, many of them are still capable of providing per-subscriber IPv4 addresses in their networks; in this case port-set support is not a necessary during their transition process. Therefore, this document focuses on specifying a clean IPv4-over-IPv6 solution with full IPv4 address provisioning.

Unlike stateless IPv4-over-IPv6 mechanisms, the mechanism described in this document still maintains per-subscriber binding state. The benefit is that it keeps IPv4-IPv6 addressing independent, which brings deployment and operation flexibility. The IPv6 infrastructure in the middle does not need to get involved with the IPv4-over-IPv6 mechanism at all, no special network planning is required; the deployment can be achieved in on-demand style rather than over the whole network; the IPv4 address resources are managed in a dedicated, centralized way rather than distributed to local sites with IPv6. In

general, the two types of mechanisms have different primary targets and application scenarios under the IPv4-over-IPv6 scope.

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

Public 4over6: Public 4over6 is the mechanism proposed by this draft. Public 4over6 supports bidirectional communication between IPv4 Internet and IPv4 hosts or local networks in IPv6 access network, by leveraging IPv4-in-IPv6 tunnel and public IPv4 address allocation.

4over6 initiator: in Public 4over6 mechanism, 4over6 initiator is the IPv4-in-IPv6 tunnel initiator located on the user side of IPv6 network. The 4over6 initiator can be either a dual-stack capable host, or a dual-stack CPE device. In the former case, the host has both IPv4 and IPv6 stack but is provisioned with IPv6 access only. In the latter case, the CPE has both IPv6 interface connecting to ISP network, and IPv4 interface connecting to local network; hosts in the local network can be IPv4-only.

4over6 concentrator: in Public 4over6 mechanism, 4over6 concentrator is the IPv4-in-IPv6 tunnel concentrator located in IPv6 ISP network. It is a dual-stack router which connects to both the IPv6 ISP network and IPv4 Internet.

4. Deployment Scenario

4.1. Scenario and requirements

The general scenario of Public 4over6 is shown in Figure 1. Users in an IPv6 network take IPv6 as their native service. Some users are end hosts which face the ISP network directly, while the others are local networks behind CPEs, such as a home LAN, an enterprise network, etc. The ISP network is IPv6-only rather than dual-stack, which means the ISP cannot provide native IPv4 service to users. However, it is acceptable that some router(s) on the carrier side becomes dual-stack and connects to IPv4 Internet. So if network users require IPv4 connectivity, the dual-stack router(s) will work as their "entrance".

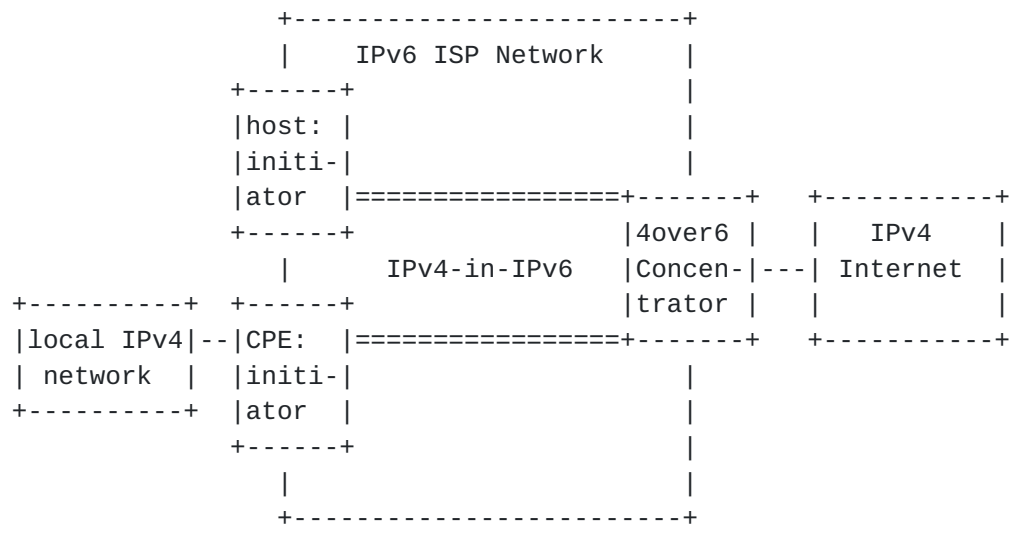

Figure 1 Public 4over6 scenario

From end user perspective, 4over6 users require IPv4-to-IPv4 communication with the IPv4 Internet. An IPv4 access service is needed rather than an IPv6-to-IPv4 translation service. Second, public IPv4 addresses will be preferred by 4over6 users. With public IPv4 address provisioning, IPv4 CGN is not required so end-to-end transparency is preserved. For special users like application servers, public address brings great convenience including straightforward access, direct DNS registration, no stateful binding maintenance on CGN, etc. For the direct-connected host case, each host should get one public IPv4 address. For the local IPv4 network case, the CPE can get a public IPv4 address and runs an IPv4 NAT for the local network. Here a local NAT is still much better than the situation that involves a CGN, since this NAT is in local network and can be configured and managed by the users.

From the operator perspective, the ISPs would like to keep their IPv4 and IPv6 addressing and routing separated when provisioning IPv4 over IPv6. Then they can manage the native IPv6 networks more easily and independently, and also provision IPv4 in a flexible, on-demand way. The cost is for the concentrator to maintain per-user address binding state. As a result, double translation is not preferred. Unlike stateless scenario, double translation in this scenario brings more complexity to IPv6 network than tunnel. Therefore this draft follows the hub and spoke softwire model.

4.2. Use cases

Public 4over6 can be applicable in several practical cases. The first one is that ISPs which still have plenty of IPv4 address resource switch to IPv6. As long as the amount of the remaining and

reclaimable IPv4 addresses can match the user amount, the ISPs can deploy public 4over6 to preserve IPv4 service for the customers.

The second case is ISPs which do not have enough IPv4 addresses switch to IPv6. For those ISPs, dual-stack lite is so far the most mature solution to provision IPv4 over IPv6. In dual-stack lite, end users use private IPv4 addresses, experience a 44CGN and some service degradation. As long as the end users use public IPv4 addresses, all CGN issues can be avoided and the IPv4 service can be full bi-directional. In other words, Public 4over6 can be deployed along with DS-lite, to provide a value-added service. Common users adopt DS-lite while high-end users adopt Public 4over6. The two mechanisms can actually get coupled easily.

There is also a special instance in the second case that the end users are IPv4 application servers. In this circumstance, public address brings significant convenience. The DNS registration can be direct, with dedicated address; the application service access can be straightforward with no translation involved for the clients; there is no need to reserve and hold session state on the CGN, and no well-known port collision will come up. So it is better to have servers take Public 4over6 for IPv4 access when they are located in IPv6 network.

Besides, the document should be explicit about the direct-connected host case and the CPE case. The host case is clear: the host is directly connected to IPv6 network, but its protocol stacks have IPv4 support as well. As to the CPE case, this document would like to only focus on the situation that the local network behind the CPE stays in private IPv4. If the local network want to run public IPv4, then it can either run IPv6 as well and enable the hosts to execute Public 4over6, or acquire address blocks from the ISP and build configured tunnel or Softwire Mesh[RFC5565] with the ISP network. The former solution is suitable for the home LAN situation while the latter solution is suitable for the enterprise network situation.

5. Public 4over6 Mechanism

5.1. IPv4 Address allocation and binding maintenance

Public 4over6 can be generally considered as IPv4-over-IPv6 hub and spoke tunnel leveraging public IPv4 address. Each 4over6 initiator uses public IPv4 address for IPv4-over-IPv6 communication. As is described above, in the host initiator case, every host gets one IPv4 address; in the CPE case, every CPE gets one IPv4 address, which is then shared by hosts behind the CPE. The key problem here is IPv4 address allocation over IPv6 network, from ISP device(s) to 4over6 initiators.

There are two possibilities here. One is DHCPv4 over IPv6, and the other is static configuration. DHCPv4 over IPv6 enables DHCPv4 message to be transported in IPv6 packet instead of IPv4 packet, so the address allocation can be achieved between 4over6 concentrator and 4over6 initiators. [I-D.ietf-dhc-dhcpv4-over-ipv6] describes the DHCP protocol format and behavior extensions to support that. As to static configuration, 4over6 users and the ISP operators MUST negotiate beforehand to authorize the IPv4 address. Application servers can falls into this case. Public 4over6 supports both address allocation manners.

Along with IPv4 address allocation, Public 4over6 MUST maintain the IPv4-IPv6 address bindings on the concentrator. In this type of address binding, the IPv4 address is the public IPv4 address allocated to a 4over6 initiator, and the IPv6 address is the initiator's IPv6 address. This binding is used to provide correct encapsulation destination address for the concentrator.

If the address is allocated through static configuration, the concentrator installs the binding manually when assigning the address, and delete the binding manually when recycling the address. Else the address is allocated by DHCPv4, the concentrator MUST participate in the DHCP procedure, either run a DHCPv4 server to dynamically allocate public addresses to 4over6 initiators, or perform the DHCP relay functions and leave the actual address allocation job to a dedicated DHCPv4 server located in IPv4. When allocating an IPv4 address (to be more precise, when sending back a DHCP ACK message to a 4over6 initiator), the concentrator installs a binding entry of the allocated IPv4 address and the initiator's IPv6 address into the address binding table. This entry MUST be deleted when receiving a DHCP RELEASE of that IPv4 address, or the lease of that IPv4 address expires.

5.2. 4over6 initiator behavior

4over6 initiator has an IPv6 interface connected to the IPv6 ISP network, and a tunnel interface to support IPv4-in-IPv6 encapsulation. In CPE case, it has at least one IPv4 interface connected to IPv4 local network.

A 4over6 initiator MUST be provisioned with IPv6 beforehand. It MUST also learn the 4over6 concentrator's IPv6 address. For example, if the initiator gets its IPv6 address by DHCPv6, it can get the 4over6 concentrator's IPv6 address through a DHCPv6 option[RFC6334]. Then it runs the DHCPv4 over IPv6 process to dynamically fetch an IPv4 address from the concentrator, or negotiate with the ISP and acquire a static IPv4 address from the concentrator. This address is assigned to the IPv4-in-IPv6 tunnel interface.

5.2.1. Host initiator

When the initiator is a direct-connected host, it assigns the allocated public IPv4 address to its tunnel interface. The host uses this address for IPv4 communication. If the host acquires this address through DHCP, it MUST support DHCPv4 over IPv6.

For IPv4 data traffic, the host performs the IPv4-in-IPv6 encapsulation and decapsulation on the tunnel interface. When sending out an IPv4 packet, it performs the encapsulation, using the IPv6 address of the 4over6 concentrator as the IPv6 destination address, and its own IPv6 address as the IPv6 source address. The encapsulated packet will be forwarded to the IPv6 network. The decapsulation on 4over6 initiator is simple. When receiving an IPv4-in-IPv6 packet, the initiator just drops the IPv6 header, and hands it to upper layer.

5.2.2. CPE initiator

The CPE case is quite similar to the host initiator case. The CPE assigns the allocated IPv4 address to the tunnel interface on the CPE. The CPE MUST support DHCPv4 over IPv6 if it acquires this address through DHCP. The local IPv4 network does not take part in the public IPv4 allocation; instead, end hosts will use private IPv4 addresses, possibly allocated by the CPE.

On data plan, the CPE can be viewed as a regular IPv4 NAT (using the tunnel interface as the NAT external interface) cascaded with a tunnel initiator. For IPv4 data packets received from the local network, the CPE translates these packets, using the tunnel interface address as the source address, and then encapsulates the translated packet into IPv6, using the concentrator's IPv6 address as the destination address, the CPE's IPv6 address as source address. For IPv6 data packet received from the IPv6 network, the CPE performs decapsulation and IPv4 public-to-private translation. As to the CPE itself, it uses the public, tunnel interface address to communicate with the IPv4 Internet, and the private, IPv4 interface address to communicate with the local network.

5.3. 4over6 concentrator behavior

4over6 concentrator represents the IPv4-IPv6 border router working as the remote tunnel endpoint for 4over6 initiators, with its IPv6 interface connected to the IPv6 network, IPv4 interface connected to the IPv4 Internet, and a tunnel interface supporting IPv4-in-IPv6 encapsulation and decapsulation. There is no CGN on the 4over6 concentrator, it will not perform any translation function; instead, 4over6 concentrator maintains an IPv4-IPv6 address binding table for

IPv4 data encapsulation.

4over6 concentrator maintains the IPv4-IPv6 address binding of 4over6 initiators. Besides manual configuration of address binding, it runs either a DHCP relay or a DHCP server which supports DHCPv4 over IPv6. When sending out a DHCP ACK, the concentrator resolves the allocated IPv4 address and the IPv6 destination address, installs the binding entry into the binding table or renews it if it already exists. When the lifetime of a binding entry/a lease of allocated address expires, or when the concentrator receives a DHCP RELEASE of allocated address, the concentrator deletes the corresponding binding entry from the table. The binding entry is used to provide correct encapsulation destination address for concentrator encapsulation. As long as the entry exists in the table, the concentrator can encapsulate inbound IPv4 packets destined to the initiator, with the initiator's IPv6 address as IPv6 destination.

On the IPv6 side, 4over6 concentrator decapsulates IPv4-in-IPv6 packets coming from 4over6 initiators. It removes the IPv6 header of every IPv4-in-IPv6 packet and forwards it to the IPv4 Internet. Before the decapsulation, the concentrator MUST check the inner IPv4 source address against the outer IPv6 source address, by matching such a binding entry in the binding table. If no binding is found, the concentrator silently drops the packet. On the IPv4 side, the concentrator encapsulates the IPv4 packets destined to 4over6 initiators. When performing the IPv4-in-IPv6 encapsulation, the concentrator uses its own IPv6 address as the IPv6 source address, uses the IPv4 destination address in the packet to look up IPv6 destination address in the address binding table. After the encapsulation, the concentrator sends the IPv6 packet on its IPv6 interface to reach an initiator. The concentrator MUST support hairpinning of traffic between two initiators, by performing de-capsulation and re-encapsulation of packets.

The 4over6 concentrator, or its upstream router SHOULD advertise the IPv4 prefix which contains the IPv4 addresses of 4over6 users to the IPv4 side, in order to make these initiators reachable on IPv4 Internet.

Since the concentrator has to maintain the IPv4-IPv6 address binding table, the concentrator is stateful in IP level. Note that this table will be much smaller than a CGN table, as there is no port information involved.

6. Security Considerations

The 4over6 concentrator SHOULD implement methods to limit service only to registered customers. The first step is to allocate IPv4

addresses only to registered customers. One simple solution is to filter on the IPv6 source addresses of incoming DHCP packets and only respond to the ones which have registered IPv6 source address. The concentrator can also perform authentication during DHCP, for example, based on the MAC address of the initiators. As to data packets, the concentrator can implement an IPv6 ingress filter on the tunnel interface to accept only the IPv6 address range defined in the filter, as well as check the IPv4-IPv6 source address binding by looking up the binding table.

7. Change Log from the -02 Version

1.Add a paragraph in section 1 to discuss the relationship with port-set-enabled solutions, remove the corresponding texts of Lightweight 4over6 draft from section 4.2;

2.Add a paragraph in section 1 to discuss the relationship with stateless IPv4-over-IPv6 solutions;

3.Add a paragraph in section 5.2 describing the provisioning steps;

4.State in Section 5.3 that the concentrator should support hairpin;

5.Add the behavior of IPv4-IPv6 source address binding verification on the concentrator before decapsulation, in Section 5.3;

6.Change the terminology of "mapping" (stateful IPv4-IPv6 address mapping) to "binding", to avoid the possible confusion with "algorithmic mapping" in stateless mechanisms.

8. Author List

The following are extended authors who contributed to the effort:

Huiling Zhao
China Telecom
Room 502, No.118, Xizhimennei Street
Beijing 100035
P.R.China

Phone: +86-10-58552002
Email: zhaohl@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

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

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

Qi Sun
Tsinghua University
Beijing 100084
P.R.China

Phone: +86-10-62785822
Email: sunqibupt@gmail.com

Chris Metz
Cisco Systems
3700 Cisco Way
San Jose, CA 95134
USA

Email: chmetz@cisco.com

9. References

9.1. Normative References

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

[RFC4925] Li, X., Dawkins, S., Ward, D., and A. Durand, "Softwire Problem

Statement", RFC 4925, July 2007.

[RFC4966] Aoun, C. and E. Davies, "Reasons to Move the Network Address Translator - Protocol Translator (NAT-PT) to Historic Status", RFC 4966, July 2007.

[RFC5549] Le Faucheur, F. and E. Rosen, "Advertising IPv4 Network Layer Reachability Information with an IPv6 Next Hop", RFC 5549, May 2009.

[RFC5565] Wu, J., Cui, Y., Metz, C., and E. Rosen, "Softwire Mesh Framework", RFC 5565, June 2009.

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

[RFC6334] Hankins, D. and T. Mrugalski, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Option for Dual-Stack Lite", RFC 6334, August 2011.

9.2. Informative References

[I-D.ietf-dhc-dhcpv4-over-ipv6] Cui, Y., Wu, P., Wu, J., and T. Lemon, "DHCPv4 over IPv6 Transport", draft-ietf-dhc-dhcpv4-over-ipv6-03 (work in progress), May 2012.

Authors' Addresses

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

Phone: +86-10-6260-3059
EMail: yong@csnet1.cs.tsinghua.edu.cn

Jianping Wu
Tsinghua University

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

Phone: +86-10-6278-5983
EMail: jianping@cernet.edu.cn

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

Phone: +86-10-6278-5822
EMail: pengwu.thu@gmail.com

Olivier Vautrin
Juniper Networks
1194 N Mathilda Avenue
Sunnyvale, CA 94089
USA

EMail: Olivier@juniper.net

Yiu L. Lee
Comcast
One Comcast Center
Philadelphia, PA 19103
USA

EMail: yiu_lee@cable.comcast.com