v6ops Working Group C. Xie
Internet-Draft C. Ma
Intended status: Standards Track China Telecom
Expires: August 12, 2022 X. Li
CERNET Center/Tsinghua University
January 29, 2022

Requirements to multi-domain IPv6-only network
draft-xie-v6ops-reqirements-multi-domain-ipv6only-00

Abstract

Dual-stack requires IPv4 and IPv6 are deployed in parallel, as it costs more to run two technologies than one. IPv6-only is considered to be the next stage of IPv6 development from dual-stack. This document specifies requirements when deploying IPv6-only in multi-domain networks.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

As the next generation of network protocol of the Internet, IPv6 has been widely deployed during the past 10 years and IPv6 is growing faster than IPv4. Document [IPv6 Deployment Status] [I-D.ietf-v6ops-ipv6-deployment] provides an overview of IPv6 transition deployment status and how the transition to IPv6 is progressing among network operators and enterprises.

Since some portion of the Internet services do not support IPv6 yet, it is easy to understand that keep IPv4 function running when IPv6 is introduced at the early stage. Which is called IPv4/IPv6 dual-stack[RFC4213]. Up to now, most of the IPv6 deployment cases are dual-stack. However, dual-stack does have a few disadvantages in the long run, like the duplication of the network resources and states, as well as other limitations for network operation. For this reason, when IPv6 increases to a certain limit, it would be better to switch to the second state, namely IPv6-only. Generally, running an IPv6-only network would reduce operational expenditures and optimize operations as compared to an IPv4/IPv6 dual-stack environment. In 2016, IAB of IETF announced that "The IAB expects that the IETF will stop requiring IPv4 compatibility in new or extended protocols. Future IETF protocol work will then optimize for and depend on IPv6."

In order to extend the service in the case of IPv4 address depletion, we need to provide IPv6 services and the still keep the ability for users to access the global IPv4 Internet. Therefore, IPv4 as a Service (IPv4aaS) is a natural consideration for IPv6-only scheme.

Several IPv6-only transition technologies have been designed in IETF during the past twenty years[comparison] [I-D.ietf-v6ops-transition-comparison]. When these schemes support the hosting of IPv4 service, different types of IPv4 and IPv6 conversion technologies are required, for example, 464XLAT[RFC6877] uses stateful NAT64 translation technology, and IVI[RFC6219] uses stateless NAT64 translation. DS-Lite[RFC6333] adopts AFTR-based 4over6 tunneling technology, while the backbone network adopts GRE tunneling or stateless translation technology, etc. This document mainly specifies the requirements for multi-domain IPv6-only network from the perspective of operators. It does not introduce any new IPv6 transition mechanisms.

2. Conventions used in this document

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

The following terms are defined in this draft:

- Multi-domain IPv6-only network: An IPv6-only network which consists of multiple ASes belonging to and operated by the same operator.

- UE: User Equipment, e.g., mobile phone.

- CPE: Customer Premise Edge device, e.g., a host, router, or switch.

- PE : Provider Edge device.

- IPv4-embedded IPv6 packet: IPv6 packet which is generated from IPv4 packet by algorithmically mapping of the source and destination IPv4 addresses to IPv6 addresses.

- o Border gateway: A PE router which run eBGP routing protocol and peering with the BGP router of other AS.

- o Conversion point: A device which provides conversion between IPv4 and IPv6.

4. The reason to consider multi-domain factor when implementing IPv6-only

In general, transition to IPv6-only from dual-stack means some or all the IPv4 protocol instances of dual-stack network will closed gradually, thereby IPv6 will become the main network-layer protocol. When IPv4 is closed at the network layer, the first question is how to make remaining IPv4 service running normally and users' experience does not deteriorate. The deployment of IPv6-only should not be based on the premise of the extinction of all IPv4-only service in short time, it is very possible that some portion of the Internet service will consistently be IPv4-based. In other words, IPv6-only network should carry not only IPv6-capable services, but also IPv4-only services.

[RFC5565]describes the IPv4-over-IPv6 scenario, where the network core is IPv6-only and the interconnected IPv4 networks are called IPv4 client networks. The P Routers (Provider Routers) in the core only support IPv6, but the AFBRs support IPv4 on interfaces facing IPv4 client networks and IPv6 on interfaces facing the core. The routing solution defined in [RFC5565] for this scenario is to run IBGP among AFBRs to exchange IPv4 routing information in the core, and the IPv4 packets are forwarded from one IPv4 client network to the other through a softwire using tunneling technology, such as MPLS, LSP, GRE, L2TPv3, etc.

[RFC6992]describes a routing scenario where IPv4 packets are transported over an IPv6 network, based on [RFC6145] and [RFC6052], along with a separate OSPFv3 routing table for IPv4-embedded IPv6 routes in the IPv6 network.

In general, the networks of large-scale operators are composed of multiple autonomous system(AS)es, different ASes may serve different scenarios, such as metro network, backbone network, 4G or 5G mobile core, data center network, and are often managed by different departments or institutions, using different routing and security policies. When introducing the IPv6-only scheme without collaboration between ASes, different ASes adopt the IPv6 transition approach independently, the result is that multiple IPv6-only islands are connected by IPv4 links between domains. As shown in figure 1, there will be more IPv4-IPv6 packet conversion gateways with different functions in the network. Under this circumstance, IPv4-embedded IPv6 packets need to be transformed back to IPv4 packets at the egress of one AS, and then back to IPv6 in the next domain, and the number of conversion points will increases along with the increasing of the number of ASes. Excessive IPv4-IPv6 conversion gateways lead to complexity of network and CAPEX increasing. Therefore, there is an urgent need for multi-domain IPv6-only

solutions to eliminate unnecessary conversion functions and improve data forwarding efficiency.

```
        +----+  +----+                           +-------+
        |UE/|--|PGW|                             | IPv4 |
        |CPE|  +----+                            |Server|
        +----+    |                              +-------+
                  |                                  |
          -----------                          -----------
         /Mobile Core\                        /           \
        |   Network   |                      |    IPv4     |
        | (IPv6-only) |                      |  Internet   |
         \           /                        \           /
          -----------                          -----------
               |                                   |
           +------+                           +---------+
           |PLAT/|                            |IPv4 BGP|
           |NAT64|                            | Router |
           +------+                           +---------+
              |                                   |
              |            -----------            |
         +---------+ IPv4/ Backbone  \IPv4 +---------+
```

```
        |Stateless|----|  Network     ----|Stateless|
        | NAT64   | link\(IPv6-only)/link | NAT64   |
        +---------+      -----------      +---------+
           XLAT-1                            XLAT-2
```

Figure 1: IPv6-only Independent Deployment in Multi-domain Network

5. Scenarios

This section describes scenarios where IPv4 packets are transported over an multi-domain IPv6-only network. A typical model of multi-domain IPv6 network is depicted in figure 2. Network 1, belonging to and operated by operator A, runs IPv6 and is composed of multiple inter-connected ASes, i.e., AS1, AS2 and AS3. In addition, network 1 provides access to different types of users, including mobile, home broadband and enterprise customers, denoted by UE1, UE2 and UE3 in figure 2. Routers that are outside the backbone but directly attached to it are known as "Customer Edge" (CE) routers.

Network 1 is open, it is interworking with the external networks. Operator 2 is one of the neighbor operators of Operator 1, AS4 of operator 2 and AS3 of operator are interconnected through BGP protocol. In order to illustrate "IPv4 As A service" , AS4 is an IPv4-only network, which means that it does not run IPv6 protocol.

In addition, cloud services are hosted in data centers and connected across multiple data centers, the edge, and public and private

clouds. The data center must be able to communicate across these multiple sites, both on-premises and in the cloud. IPv6-only network need to provide connections for cloud data center. Network 1 supports two connections modes of cloud data centers, the first one is between cloud data center and individual users, for instance, the user of CPE1 access the service hosted in DC1, the second one is the connection between cloud data centers, for instance, communications between DC1 and DC2.

The edge nodes of the Network 1 are often known as "Provider Edge" (PE) routers. The term "ingress" (or "ingress PE") refers to the router at which a packet enters the network, and the term "egress" (or "egress PE") refers to the router at which it leaves the backbone. Interior nodes are often known as "P routers". The P routers in the core only support IPv6, but the PEs support IPv4 on interfaces facing IPv4 client networks and IPv6 on interfaces facing the core.

Network 1 provides transportation services for packets that originate outside the network and whose destinations are outside the network. These packets enter the IPv6 network at one of its "edge routers". They are routed through the network to another edge router, after which they leave the network and continue on their way.


```
                      -----          -----
                     /     \        /     \
                    |  DC1  |      |  DC2  |
                     \     /        \     /
                      -----          -----
                ---------|--------------|---------
               |         |  (Operator1) |         |
               |       +---+  Network +---+       |
```

```
           |       |PE3|          |PE4|       |     (Operater2)
           |       +---+          +---+       |       +--+
           |      /    \         /     \      |      /    \
    +----+ | +---+      +--+ +--+       +---+ | +---+      +
    |UE/ |---|PE1| AS1  |R1|-|R2|       |PE5|---|BR1|  AS4 |
    |CPE1| | +---+      +--+ +--+       +---+ | +---+      +
    +----+ |      \    /        |       |     |      \    /
           |       +--+         |       |     |       +--+
           |       |R5|         |       |     |
           |       +--+         | AS3   |     |
           |        |           |       |     |
           |       +--+         |       |     |
    +----+ |       |R6|         |       |     |     (Operater3)
    |UE/ | |       +--+         |       |     |       +--+
    |CPE2|\|      /    \        |       |     |      /    \
    +----+ \ +---+      +--+ +--+       +---+ | +---+      +
           |-|PE2| AS2  |R3|-|R4|       |PE6|---|BR2| AS5  |
    +----+ / +---+      +--+ +--+       +---+ | +---+      +
    |UE/ |/|      \    /         \     /      |      \    /
    |CPE3| |       ----          -----       |       +--+
    +----+ |                                  |
            ----------------------------------
```

Figure 2. Multi-domain IPv6 Network Model

In order to illustrate the requirements of IPv6-only network, the following scenarios should be considered,

Scenario 1: IPv6 user to IPv4 server, IPv6-only user accesses IPv4 services hosted in cloud data centers.

Scenario 2: IPv4 user to IPv4 server, IPv4-only user accesses IPv4 services hosted in cloud data centers.

Scenario 3: IPv6 user to IPv6 server, IPv6-only user accesses IPv6 services hosted in cloud data centers.

Scenario 4: DC-to-DC, IPv6-only provide communications between VMs hosted cloud data centers, despite they are IPv4, IPv6 or IPv4/IPv6 dual-stack.

Scenario 5: Transit for neighbor networks, IPv6-only network serves as an interconnection between several segregated IPv4-only network, IPv4 packets are transported over the IPv6 network between IPv4 networks.

Scenario 6, 5G Transport service, SD-WAN, etc.

It should be noted that the aforementioned scenarios are only a subset of the scenarios that multi-domain IPv6-only network will support in the future.

6. Requirements from IPv6-native traffic

Since there is no IPv4-IPv6 transition issue, native-IPv6 traffic can be transported by IPv6-only network naturally, the requirements are not covered by this document.

7. Procedure

This section firstly gives a very brief overview of the procedures of the IPv4 service delivery over IPv6-only network.

When an ingress PE receives an IPv4 packet from a client-facing interface destined to a remote IPv4 network, it looks up the packet's destination IP address. In the scenario of interest, the best match will help to find another PE, the egress PE. Since this is a multi-domain IPv6-only network, the ingress and egress may belong to different ASes, for example the ingress is in AS 1 and egress is in AS 2. The ingress PE must transform the IPv4 packet into IPv6 packet and forward the packet to the egress PE. The egress PE then derives the IPv4 source and destination addresses from the IPv4-embedded IPv6 addresses, respectively [RFC6052]and restore the original IPv4 packet, and forwards it further according to the IPv4 routing table maintained on the egress. The IPv6 data-path can be shown as below,

```
                        IPv6 Data Path
                   |<------------------------>|
                   |                          |    (Operater2)
                   |   ----           -----   |       ----
                   |  /    \         /     \  |      /    \
        +----+   +---+      +--+ +--+       +---+   |      |
        |UE/ |---|PE1| AS1  |R1|-|R2|  AS3  |PE3|---| AS4  |
        |CPE1|   +---+      +--+ +--+       +---+   |      |
        +----+        \    /         \     /         \    /
                       ----           -----           ----
```

Figure 3. IPv6 Data Path from Ingress PE to Egress PE

Another case that IPv4 packets may have been transformed into IPv6 packet in UE/CPE, as done by CLAT of 464XLAT [RFC6877][RFC6877]，before they reach the edge of the network. In this case, the ingress PE receives an IPv6 packet from a client-facing interface and looks up the packet's destination IPv6 address. and forward the packet to the egress PE. The egress PE then restore the original IPv4 packet, and forwards it further by looking up its IP destination address. The IPv6 data-path can be shown as below.

```
                  IPv6 Data Path
       |<----------------------------------->|
       |                                     |    (Operater2)
       |          ----          -----        |       ----
       |         /    \        /     \       |      /    \
     +-----+   +----+      +---+ +---+      +----+  |      |
     |UE/  |---|PE1| AS1   |R1|-|R2|   AS3  |PE3|---| AS4  |
     |CPE1|    +----+      +---+ +---+      +----+  |      |
     +-----+         \    /        \     /          \    /
                      ----          -----            ----
```

Figure 4. IPv6 Data Path from UE/CPE to Egress PE

When PE of IPv6-only network UE/CPE need to implement IPv4-IPv6 conversion, a specific IPv6 address range will represent IPv4 systems (IPv4-converted addresses), and the IPv6 systems have addresses (IPv4-translatable addresses or IPv4-embedded IPv6 addresses) that can be algorithmically mapped to a subset of the service provider's IPv4 addresses. Note that IPv4-translatable addresses are a subset of IPv4-converted addresses. In this way, there is no need to concern oneself with translation tables, as the IPv4 and IPv6 counterparts are algorithmically related.

8. Requirements from IPv4 service delivery

In order to support IPv4 service delivery, the following requirements should be met by multi-domain IPv6-only network，

Requirement 1: beneficial to wider IPv6 adoption

It should largely reduce IPv4 public address consumption and accelerate the deployment of IPv6, rather than prolonging the lifecycle of IPv4 by introducing multiple layers of 44NAT.

Requirement 2: IPv4-as-a-Service

IPv6 transition mechanisms should provide IPv4 service delivery and there should be no perceived degradation of customer experience when accessing the remaining IPv4 services.

Requirement 3: end-to-end

End-to-end means, for any given IPv4 traffic flow, there should be no IPv4-IPv6 conversion point in the middle of the IPv6 data path when traversing multi-domain IPv6 network, in other words, IPv4 packet should not appear in the middle of the IPv6 data path, the maximum number of the transition point should be two. In addition, IPv6-only network should support the following two types of IPv6 data path, as mentioned in section 7.

-From UE to egress, the packets of IPv4 service can be translated into IPv6 packets within UE or CPE, and there should be no IPv4-IPv6 conversion before they reaches the egress of the network.

-From the ingress to egress, since the core of the network is IPv6-based, so all IPv4 packets which reaches the edge of the network should be transformed into IPv6 packets by the ingress and forwarded to the egress of the network. The end-to-end requirement also be valid for cloud-to-cloud communications.

Requirement 4: support of translation and encapsulation

For the data-plane, there are two approaches for traversing the IPv6 provider network: 4-6-4 translation and 4-in6 encapsulation, both of them can be supported by IPv6-only network, the core nodes do not distinguish between translation-based IPv6 packet and encapsulation-based IPv6 packet. At the egress, the PE can recover IPv4 packet by reading the next-header field of the packet. It should be noted that translation mode and encapsulation mode have the same IPv4-IPv6 address mapping algorithm.

Requirement 5: controller independent

In order to forward an IPv4 packet to the right egress point, IPv4 reachability information must be exchanged in advance between the IPv4 networks over in IPv6-only network. In general, BGP4+ is used to distribute external IPv4 routing information among AFBRs. In the scenarios of interest, the extension of BGP4+ sessions can be used to pass IPv4 routing information. This would require that IPv4-embedded

IPv6 routes be flooded throughout the entire IPv6-only network and stored on every router. It does not rely on the deployment of any centralized controller. Note that with this routing solution, the IPv4 and IPv6 header conversion performed in both directions by the PE is stateless.

Requirement 6: user stateless at the border gateway

Maintaining user status will need great volume of storage and computation power, so it is generally stored or managed at the edge of network and close to the user side. It is unsuitable to store user-related status at the inter-connection point. The border gateway with other networks should be unaware of the user-related information, it only needs to perform stateless translation or encapsulation/decapsulation.

Requirement 7: high scalability

It should achieve scalability, simplicity and high availability, especially for large-scale SPs. When PE processes IPv4-features at the edge of the network, the quantity of the IPv4-related status should not increase linearly or exponentially along with the quantity of the user or traffic. Considering this, it is better to adopt algorithm-based mapping approach to avoid excessive status storage at the edge. it would also prevent overload of the IPv6 routing table.

Requirement 8: SRv6 applicable

SRv6 can be supported by inserting SRH in translated IPv6 packet, so the network programming can be realized for IPv4 traffic flow.

Requirement 9: incremental deployment

It should deploy in an incremental fashion and the overall transition process should be stable and operational.

9. Security Considerations

There are no other special security considerations.

10. IANA Considerations

There are no other special IANA considerations.

11. Acknowledgement

This is under development by a large group of people. Those who have posted to the list during the discussion.

12. Normative References

[I-D.ietf-v6ops-ipv6-deployment]
Fioccola, G., Volpato, P., Elkins, N., Martinez, J. P., Mishra, G. S., and C. Xie, "IPv6 Deployment Status", draft-ietf-v6ops-ipv6-deployment-03 (work in progress), October 2021.

[I-D.ietf-v6ops-transition-comparison]
Lencse, G., Martinez, J. P., Howard, L., Patterson, R., and I. Farrer, "Pros and Cons of IPv6 Transition Technologies for IPv4aaS", draft-ietf-v6ops-transition-comparison-01 (work in progress), October 2021.

[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>.

[RFC4213] Nordmark, E. and R. Gilligan, "Basic Transition Mechanisms for IPv6 Hosts and Routers", RFC 4213, DOI 10.17487/RFC4213, October 2005, <https://www.rfc-editor.org/info/rfc4213>.

[RFC5565] Wu, J., Cui, Y., Metz, C., and E. Rosen, "Softwire Mesh Framework", RFC 5565, DOI 10.17487/RFC5565, June 2009, <https://www.rfc-editor.org/info/rfc5565>.

[RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, DOI 10.17487/RFC6052, October 2010, <https://www.rfc-editor.org/info/rfc6052>.

[RFC6145] Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", RFC 6145, DOI 10.17487/RFC6145, April 2011, <https://www.rfc-editor.org/info/rfc6145>.

[RFC6219] Li, X., Bao, C., Chen, M., Zhang, H., and J. Wu, "The China Education and Research Network (CERNET) IVI Translation Design and Deployment for the IPv4/IPv6 Coexistence and Transition", RFC 6219, DOI 10.17487/RFC6219, May 2011, <https://www.rfc-editor.org/info/rfc6219>.

[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>.

[RFC6877] Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT: Combination of Stateful and Stateless Translation", RFC 6877, DOI 10.17487/RFC6877, April 2013, <https://www.rfc-editor.org/info/rfc6877>.

[RFC6992] Cheng, D., Boucadair, M., and A. Retana, "Routing for IPv4-Embedded IPv6 Packets", RFC 6992, DOI 10.17487/RFC6992, July 2013, <https://www.rfc-editor.org/info/rfc6992>.

Authors' Addresses

Chongfeng Xie
China Telecom
Beiqijia Town, Changping District
Beijing, Beijing 102209
China

Email: xiechf@chinatelecom.cn

Chenhao Ma
China Telecom

Beiqijia Town, Changping District
Beijing, Beijing 102209
China

Email: machh@chinatelecom.cn

Xing Li
CERNET Center/Tsinghua University
Shuangqing Road No.30, Haidian District
Beijing, Beijing 100084
China

Email: xing@cernet.edu.cn