Workgroup: v6ops Working Group
Internet-Draft:
draft-xie-v6ops-framework-md-ipv6only-
underlay-05
Published: 22 October 2022
Intended Status: Informational
Expires: 25 April 2023
Authors: C. Xie
China Telecom
C. Ma
China Telecom
X. Li
CERNET Center/Tsinghua University
G. Mishra
Verizon Inc
M. Boucadair
Orange
T. Graf
Swisscom

Framework of Multi-domain IPv6-only Underlay Networks and IPv4 as a Service

Abstract

For the IPv6 transition, dual-stack deployments require both IPv4 and IPv6 forwarding capabilities to be deployed in parallel. IPv6-only is considered as the ultimate stage where only IPv6 transfer capabilities are used while ensuring global reachability for both IPv6 and IPv4 (usually known as IPv4aaS). This document specifies requirements and proposes a framework for deploying IPv6-only as the underlay in multi-domain networks, discusses the requirements of service traffic, major components and interfaces, IPv6 mapping prefix allocation, typical procedures for service delivery. The document also discusses related considerations with security.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

IPv6 capabilities have been widely deployed during the past decade with IPv6 traffic growing faster than IPv4. [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.

As per 2022, most IPv6 deployments rely on dual-stack[RFC4213]. Dual-stack does have a few disadvantages in the long run, like the duplication of the network resources and states and increased complexity for network operation to maintain both stacks. For those reasons, and furthermore when IPv6 usage is being the dominant, it makes more sense to consider IPv6-only to reduce network resources and operational complexity.

In 2016, the IAB announced that it "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" [IAB-statement] only. In order to provide the connectivity service after IPv4 address depletion, operators need to provide IPv6 services and preserve access to the global IPv4 Internet as a Service(IPv4aaS) is therefore a natural consideration for IPv6-only network.

Several IPv4 service continuity mechanisms have been designed within IETF during the past twenty years[I-D.ietf-v6ops-transition-comparison]. Different types of IPv4 and IPv6 conversion technologies may be considered. For instance 464XLAT[RFC6877] uses stateful NAT64 translation, MAP-E[RFC7597]and MAP-T [RFC7599] use stateless IPv4-IPv6 address translation for encapsulation and translation respectively. DS-Lite[RFC6333] adopts AFTR-based 4over6 tunneling technology.

This document specifies the requirements for multi-domain IPv6-only underlay networks and proposes a general framework for network operators. The objective of such a framework is to help large-scale operators implement the transition to IPv6-only and support cross-domain, end-to-end IPv4 service delivery over IPv6-only networks. In this document, the term of "IPv6-only network" stands for "IPv6-only underlay network", unless there is a specific statement. This document does not introduce any new IPv6 transition mechanisms nor IPv4aaS.

1.1. 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 BCP 14[RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Terminology

The following terms are defined in this document:

*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 Equipment.

*IXP: Internet Exchange Point.

*WKP: Well-Known Prefix.

*NSP: Network-Specific Prefix.

*PE : Provider Edge (Section 5.2 of [RFC4026]).

*IPv4-embedded IPv6 addresses: IPv6 addresses used to represent IPv4 nodes in an IPv6 network, 32 bits in the IPv6 address contain IPv4 address.[RFC6052]

*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.

*ASBR: A PE which runs eBGP routing protocol and peering with the BGP router of external AS.

*AFBR: A type of PE which supports both IPv4 and IPv6 address families and serves to provide transit services for the other in a backbone network (Section 1 of [RFC5565]).

*ADPT: A function entity which implements the two-way IPv4 and IPv6 packet conversion for IPv4 service delivery over IPv6-only underlay network.

*Conversion point: A function which provides conversion between IPv4 and IPv6 realms. This is, for example, the XLAT function in [RFC6144]

*GUA: IPv6 Global Unicast Address (Section 3 of [RFC3587]).

3. Focus on IPv6-only Networks

The global industry has not given a unified definition of IPv6-only network so far. This document defines such a notion as a IPv6-centric network in which data packets are forwarded upon IPv6

capability, An IPv6-only network may interconnect with external networks, including IPv4-only networks.

Generally, IPv6-only network should support the following scenarios,

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

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

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

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

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

Scenario 6: IPv6-only eBGP Edge peering in Internet Exchange Point (IXP)[I-D.ietf-bess-ipv6-only-pe-design], this serves to eliminate IPv4 provisioning at the Edge of IXP that are facing IPv4 address depletion at large peering points.

Scenario 7: 5G Transport service, SD-WAN, network slicing, etc.

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

4. Why Considering Multi-domain Factor When Implementing IPv6-only Networks?

Generally, the networks of large-scale operators comprise multiple autonomous systems (ASes). 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.

A typical model of multi-domain network is depicted in figure 1. Network 1, belonging to and operated by operator 1, is composed of multiple inter-connected ASes, AS1, AS2 and AS3. Network 1 provides access to multiple types of users, including mobile, home broadband and enterprise customers, denoted by UE1, UE2 and UE3 in figure 1. Routers that are outside the backbone but directly attached to it are known as "Customer Edge" (CE) routers. [RFC8585] specifies the

IPv4 service continuity requirements for IPv6 Customer Edge (CE) routers. Specificially, it extends the basic requirements for IPv6 CE routers to allow for support of IPv4-as-a-Service (IPv4aaS) by means of transition technologies for delivering IPv4 in IPv6-only access networks. In addition, cloud services are hosted in data centers and connected across multiple data centers, the edge, and public and private clouds. The service instances in cloud data centers must be able to communicate across these multiple sites, both on-premises and in the cloud. Multi-domain Networks need to provide connections for cloud data center. Network 1 supports two connection modes of cloud data centers, the first one is between cloud data center and individual users, for instance, the user of CPE1 accesses the service hosted in DC1, the second one is the connection between cloud data centers, for instance, communications between VMs hosted in DC1 and DC2 separately.

Network 1 is open, it is interworking with external networks. Operator 2 is one of the neighbor operators of operator 1, AS4 of operator 2 and AS3 of operator 1 are interconnected through BGP protocol. AS4 is an IPv4-only network, which means that it does not run IPv6. 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" (Provider Routers).

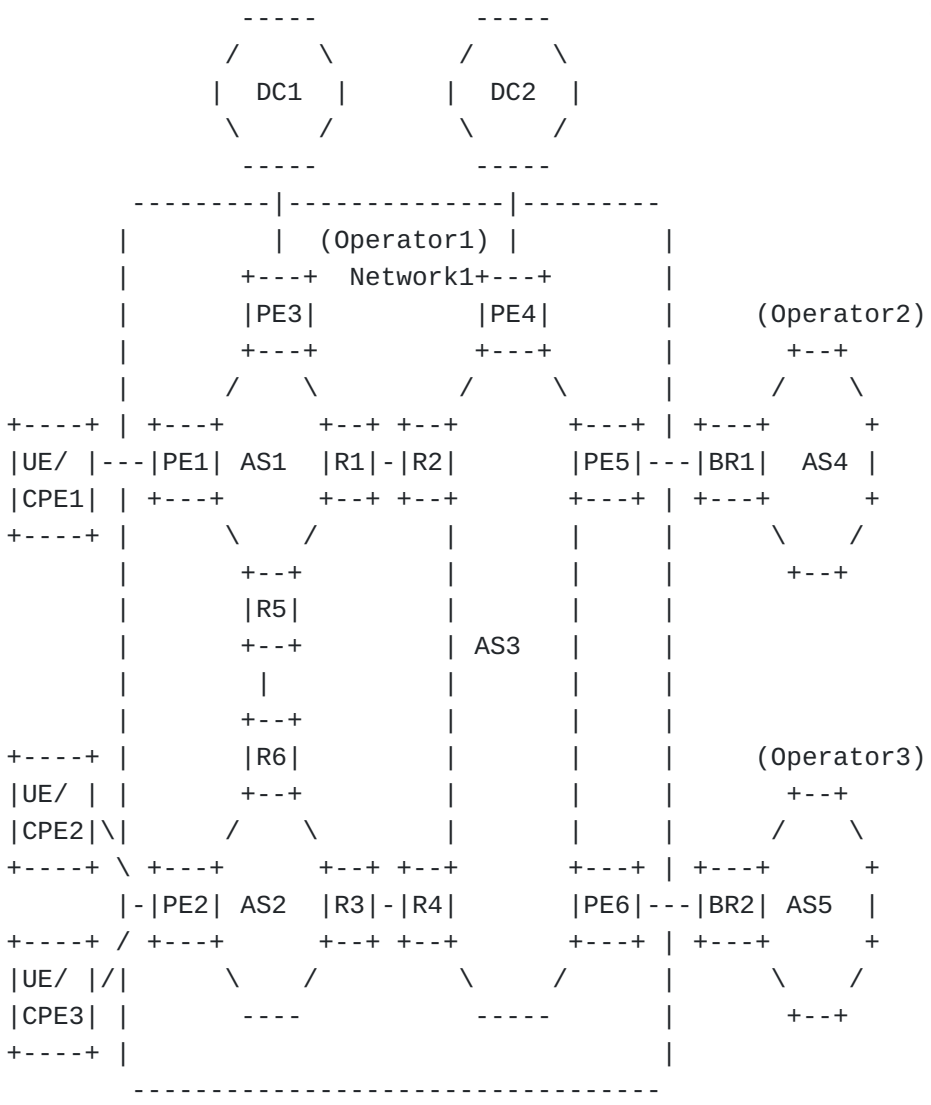

Figure 1. Multi-domain IPv6 Underlay Networks Model

For Network 1, transition to IPv6-only from dual-stack means some or all the IPv4 protocol instances of dual-stack network will be disabled gradually, thereby IPv6 will become the main network-layer protocol. To be specific, 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, in this case, the PEs need to support both address families. 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 PE routers. They are routed through the network to another PE router, after which they leave the network and continue their way.

When IPv4 capabilities are disabled, the first question is how to make remaining IPv4 services 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 services, it is very possible that some portion of the Internet service will consistently be IPv4-based. In other words, IPv6-only network should not only carry native IPv6 services, but also allow to reach 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 in the core only support IPv6, but the ASBRs support IPv4 on interfaces facing IPv4 client networks and IPv6 on interfaces facing the core. The routing solution defined in [RFC5565] 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 technologies, such as MPLS, LSP, GRE, VXLAN, L2TPv3, etc.

[RFC6992] describes a routing scenario where IPv4 packets are transported over an IPv6 network, based on [RFC7915] and [RFC6052], along with a separate OSPFv3 routing table for IPv4-embedded IPv6 routes in the IPv6 network.

For multi-domain networks, 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 2, there will be more IPv4-IPv6 packet conversion gateways with different functions in the network. Under this circumstance, IPv6 packets converted from IPv4 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 gateways will increase 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 solution to eliminate unnecessary conversion functions and improve data forwarding efficiency.

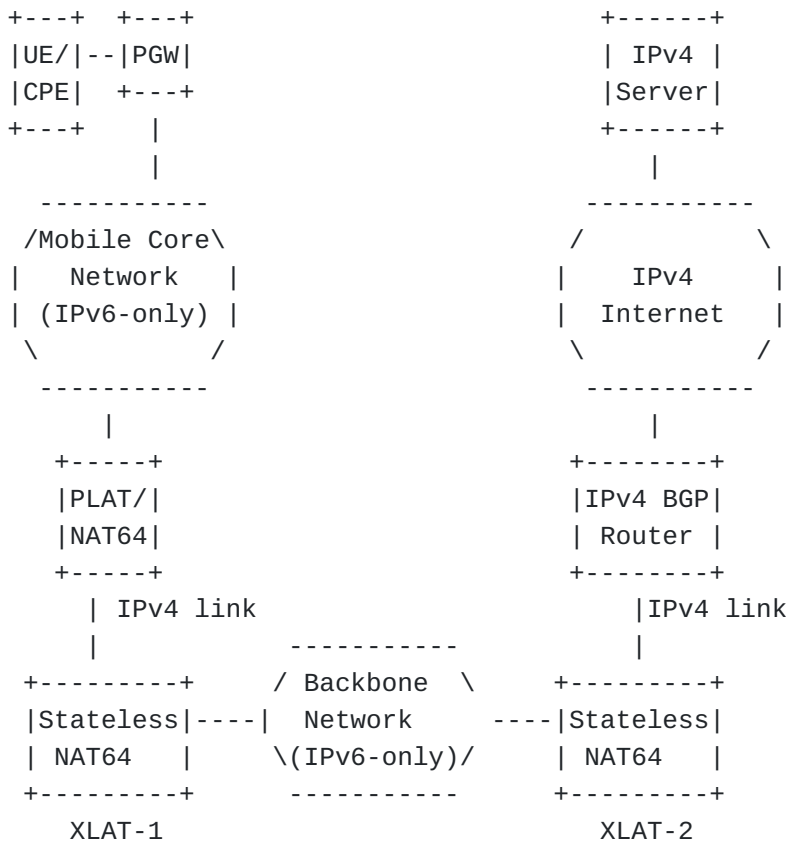

Figure 2: IPv6-only Independent Deployment in Multi-domain Networks

5. Requirements from Service Traffic

Native-IPv6 traffic can be transported over an IPv6-only network following legacy procedures.

In order to support IPv4 service continuity, the following requirements should be met by multi-domain IPv6-only networks.

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 NAT44.

Requirement 2: IPv4-as-a-Service

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

Requirement 3: optimized end-to-end

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 networks, in other words, IPv4 packet should not appear in the middle of the IPv6 data path, the quantity of the conversion points should not succeed two. In addition, IPv6-only network should support the following two types of IPv6 data path.

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

-From the ingress to egress, since the core of the network is IPv6-based, so all IPv4 packets which reach 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 should also be valid for cloud-to-cloud communications.

Requirement 4: support of double translation and encapsulation

The data-plane has two approaches for traversing the IPv6 provider network: 4-6-4 translation and 4over6 encapsulation, at least one mode should 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. Moreover, translation mode and encapsulation mode should share the same IPv4-IPv6 address mapping algorithm. Note that the double translation can reduce to single translation, while the encapsulation cannot.

Requirement 5: 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 ASBR that is interworking with external networks should be unaware of the user-related information, it only needs to perform stateless translation or encapsulation/decapsulation.

Requirement 6: high scalability

It should achieve high scalability, simplicity and availability, especially for large-scale operators. 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 avoid overloading of the IPv6 routing table.

Requirement 7: incremental deployment

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

Requirement 8: no security compromise

The technologies proposed must not introduce additional security compromise.

6. Description of the Framework

6.1. Overview

Multi-domain IPv6-only networks should support the forwarding of IPv4 service data, after transforming IPv4 packets into IPv6 ones in the UE/CPE or at the edge of the network. Take the latter case as an example, when IPv4 packets that need to traverse lPv6-only network, the ingress PE, i.e., PE1, will convert IPv4 packets into lPv6 packets by translation or encapsulation and send them into IPv6 network. After intra-domain and cross-domain transmission, the IPv6 packet reaches the egress PE, i.e., PE2, it can be restored to an IPv4 packet.

As can be seen from the above, the routing of IPv4 data in the form of IPv6 packet will follow topology of IPv6 network. With this framework, each PE will be allocated and identified by at least one IPv6 mapping prefix, denoted by Pref6(PE), it will also have one or more associated IPv4 addtess blocks which are extracted from local IPv4 routing table or address pool. The mapping relationship between IPv4 address block and IPv6 mapping prefix is called mapping rule in this context. The mapping rule announced by a given PE will have at least the following data structure,

IPv4 address block: Pref6(PE)

Since this is prefix-level mapping, there is no need to maintain user-ralated status or translation tables at the PE devices.

The mapping rule is used by the ingress to generate corresponding IPv6 source and destination addresses from its IPv4 source and destination address when its egress is the given PE, and vice versa.

-The IPv6 source address is derived by appending the IPv4 source address to the Pref6(ingress PE).

-The IPv6 destination address is derived by appending the IPv4 destination address to the Pref6(egress PE) in the mapping rule.

[RFC6052] illustrates the algorithmic translation of an IPv4 address to a corresponding IPv6 address, and vice versa, using only statically configured information. With this algorithm, IPv4-embedded IPv6 addresses are composed by concatenating the prefix, the 32 bits of the IPv4 address, and the suffix (if needed) to obtain a 128-bit address. The prefixes can only have one of the following lengths: 32, 40, 48,56, 64, or 96.

For the deployment scenario in this document, it proposed that IPv4 address is located at the last 32 bits of the IPv6 address, most significant bits first. The bits between IPv6 mapping prefix and IPv4 address are reserved for future extensions and SHOULD be set to zero. Examples of such representations are presented in Table 1.

IPv6-mapping prefix	IPv4 address	IPv4-embedded IPv6 address
2001:db8::/32	192.0.2.33	2001:db8::c000:221
2001:db8:100::/40	192.0.2.33	2001:db8::1c0:2:21
2001:db8:122::/48	192.0.2.33	2001:db8:122::c000:2:21

Table 1. Text Representation of IPv4-Embedded IPv6 Address

Using the mechanism of mapping rule exchange in IPv6-only network, an egress PE can tell other PEs that IPv4 packet whose IPv4 destination address is within the scope IPv4 address block of the mapping rule, can be forwarded in the IPv6-only network through the egress PE identified by the corresponding IPv6 mapping prefix of the mapping rule. This mapping rule can be transmitted across domains. Therefore, it gives the direction of IPv4 service data transmission in multi-domain IPv6-only networks.

It should be noted that the mapping rule contains not only the data structure above, but also other necessary information to support IPv4 service delivery over IPv6-only network, the detailed structure of the mapping rule is out of the scope of this document.

Although this document illustrates the framework of multi-domain IPv6-only networks operated by a single operator, this multi-domain model can naturally be extended to IPv6-only networks which consist of multiple ASes and are operated by different operators.

6.2. ADPT Description

This section illustrates the framework of multi-domain IPv6 network from the perspective of ADPT in PE devices. ADPT is the entity in PE which accommodates the conversion of IPv4 packets into IPv6 ones for IPv4 service delivery over IPv6-only network. ADPT comprises the following components, as shown in figure 3.

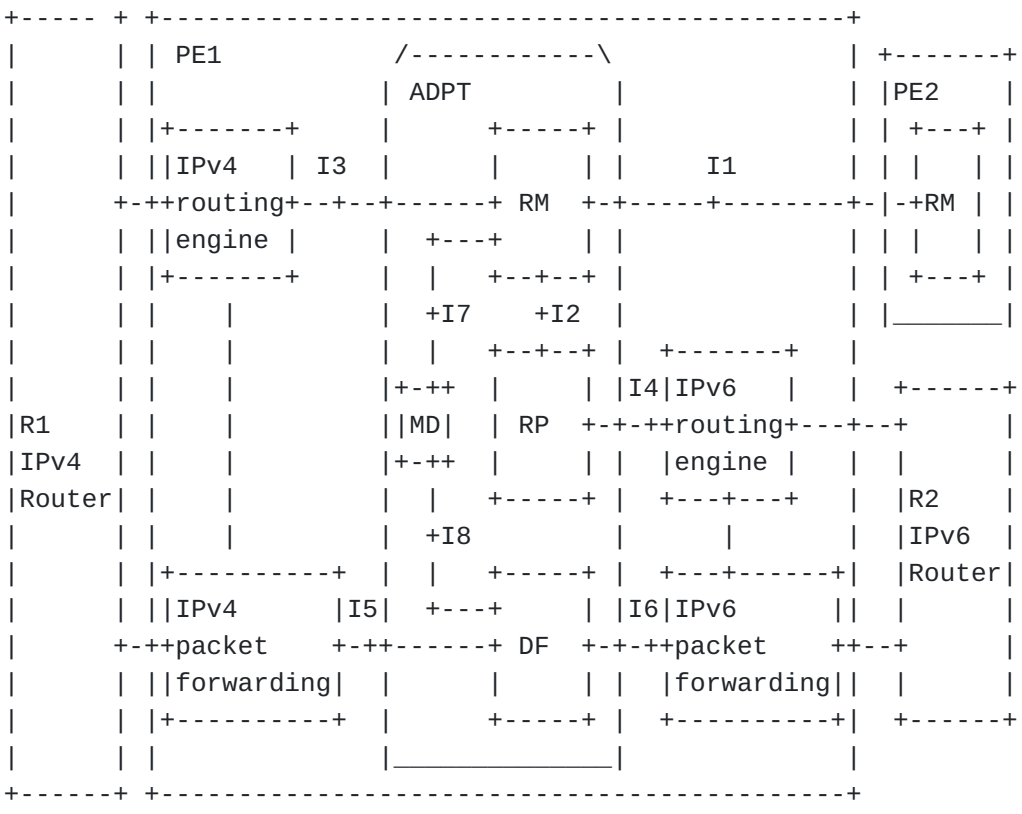

Figure 3. Framework of Multi-domain IPv6-only Networks

6.2.1. Rule Management Layer

The Rule Management Layer, i.e., RM, deals with the management of mapping relationship between IPv4 address block and IPv6 mapping prefix of PEs, as shown in figure 3.

In each PE, there is a mapping rule database, i.e., MD, to store all the mapping rule records it receive from other PEs. Rule management layer provides management functions to mapping rule database through interface I7, for example, Rule Management Layer inserts, modifies, or deletes mapping rules in the MD. The interface with the ADPT of other PE is I1, which is used for the exchanging of mapping rule with each other. The interface with Routing Processing Layer, which will be illustrated in section 6.2.2, is I2, which is used for the transmission of mapping rule through Routing Processing Layer. PE1 can extract the IPv4 address blocks from its IPv4 BGP routing instance through interface I3, and generate the mapping rules of the device in combination with its own Pref6. When the mapping rules are ready, they will be sent to Routing Processing Layer through interface I2. Correspondingly, PE1 will receive the mapping rules of other PEs through interface I2 and stores them in the local mapping rule database.

For some IPv4 address blocks which are not announced explicitly by any egress PEs to the ingress PE, there will be no corresponding mapping rule in the rule database. To solve this problem, the default egress PE is defined in this framework, which announces the default IPv6 mapping rule with the default mapping prefix to other PEs. The format of the mapping rule for default IPv4 address is as follows,

0.0.0.0/0: Pref6(PE)

6.2.2. Routing Processing Layer

Routing Processing Layer, i.e., RP, is in charge of the exchanging of mapping rule with other PEs and its related routing information at the routing layer. The exchanging of the mapping rule should precede to the process of IPv4 data transmission, otherwise, the data originated from IPv4 network will be dropped due to the absence of the IPv6 mapping prefix corresponding to its destination address.

When the request of the mapping rule from Rule Management Layer through interface I2 is being received, Routing Processing Layer will convert the mapping rule into data structure that is suitable for the transmission in the IPv6 routing system and send it to the IPv6 routing engine through interface I4. In opposite direction, when receiving the routing information from IPv6 routing engine through interface I4, Routing Processing Layer will extract mapping rule from the routing information and send it to the Rule management layer.

To support the transmission of mapping rules at the routing layer, BGP4+ protocol or other control protocols needs to be extended. However, this has been out of the scope of the draft and will be discussed in other documents. In addition, routing process layer is responsible for announcing the IPv6 route corresponding to each IPv6 mapping prefix throughout the multi-domain IPv6-only networks.

6.2.3. Data Forwarding Layer

Data Forwarding Layer, i.e., DF, provides data forwarding function to IPv6 packets, including native IPv6 packets and IPv4-embedded IPv6 packets. Multi-domain IPv6-only networks need to support both translation and encapsulation technologies for IPv4 data delivery:

1. Translation

Translation refers to the conversion of IPv4 packets into IPv6 packets or reverse conversion. When receiving an IPv4 packet through interface I5 from IPv4 packet forwarding module, the data forwarding layer will look up the mapping rule database through the interface I8, if the mapping rule corresponding to the IPv4 destination

address is found, the destination address of IPv6 header required for translation is generated by appending the IPv4 address to the Pref6 in the mapping rule. Otherwise, the default IPv6 mapping prefix is used to create the destination IPv6 address.

2. Encapsulation

Encapsulation means that PE encapsulates IPv4 packets in IPv6 packets without changing the original IPv4 packets, and then transmits them in multi-domain IPv6-only networks. Address mapping in encapsulation mode is same to that in translation method, when receiving IPv4 packet through interface I5 from IPv4 packet forwarding module, the data forwarding layer will look up the mapping rule database through the interface I8, if the mapping rule corresponding to the IPv4 destination address is found, the destination address of IPv6 header required for encapsulation is generated by appending the IPv4 address to the Pref6 in the mapping rule. If the mapping prefix corresponding to the destination IPv4 address is not found, the default IPv6 mapping prefix is used.

For an IPv4-embedded IPv6 packet, whether it is based on translation or encapsulation, the Pref6 part of the destination address can identify the egress in the network, so the forwarding of the IPv6 packet can be implemented based on the Pref6 information of the destination address.

6.3. Mapping Prefix Allocation

In order to support rule-based IPv4/IPv6 address mapping, a specific IPv6 address range will be planned to represent IPv4 address space by stateless mapping as with [RFC7915]. With this framework, there are two options to allocate IPv6 mapping prefix:

1) WKP:

A specific WKP can be allocated from the global IPv6 address prefix, e.g., 64:ff9b:: /96.

Pros:

Service providers do not need to allocate IPv6 address prefixes specially used for mapping IPv4 addresses from their own IPv6 address resources.

Cons:

After the IPv4 address is converted into IPv6 address with WKP, the IPv4 part of the IPv6 address is used for the routing of the origin of the data packet. In this way, many fine routes with prefix length greater than 96 will be introduced into the global IPv6 network. In

most networks, fine routing with long prefix length greater than 96 is not supported.

2) NSP:

Operator allocates a specific prefix from their existing IPv6 address resources for IPv4 addresses mapping.

Pros:

The specific IPv6 prefix allocated by operators can be considered as an parent prefix, and each PE can obtain IPv6 mapping prefixes allocated from the parent prefix. Within the multi-domain networks, the length of IPv6 mapping prefix can be easily tailored to meet the requirements of IPv6 network for routing length, and the routing of the packets can be based on the information of IPv6 mapping prefix part of the IPv6 address. Outside the multi-domain network, because the IPv6 mapping prefix has been included in its original IPv6 address prefix, it will not introduce any new routing items and affect the global IPv6 routing system.

Cons:

Not found yet.

As mentioned earlier, each PE will be identified by at least one IPv6 mapping prefix, which is used as the basic routing information to forward IPv4-embedded IPv6 packet to the right egress PE. For a given operator, the selection of the length of IPv6 mapping prefix should be given specific consideration. Firstly, the length of the IPv6 mapping prefix should be smaller than the maximum length of the routing prefix that the IPv6-only network specifies, so the PE can successfully announce to its peers via BGP protocol. Secondly, the length of all the IPv6 mapping prefixes should be the same, to avoid unnecessary processing cost and complexity induced by the prefix length diversity.

7. Procedure

This section gives a brief overview of the procedures of the IPv4 service delivery over IPv6-only underlay network. The requisite of IPv4 data delivery is that PEs have successfully exchanged the mapping rules with each other. The end-to-end IPv4 data delivery by IPv6-only network includes the following two cases,

1. IPv4 delivery from ingress PE to egress PE

When an ingress PE receives an IPv4 packet from a client-facing interface destined to a remote IPv4 network, it looks up in its mapping rule database to find the mapping rule which best matches

the packet's destination IP address. The IPv6 mapping prefix in the mapping rule will help to find another PE, the egress PE. Since this happens in multi-domain IPv6-only networks, the ingress and egress may belong to different ASes, as shown in figure 4, the ingress PE1 is in AS 1 and egress is PE3 in AS 3. The ingress PE must convert the IPv4 destination address into IPv6 destination address using the IPv6 mapping prefix of PE3 and forward the IPv6 packet to PE3. When PE3 receives the IPv6 packet, it derives the IPv4 source and destination addresses from the IPv4-embedded IPv6 addresses respectively and restore the original IPv4 packet[RFC6052]. Afterwards, the IPv4 packet will be further forwarded according to the IPv4 routing table maintained on the egress. The IPv6 data-path can be shown as below.

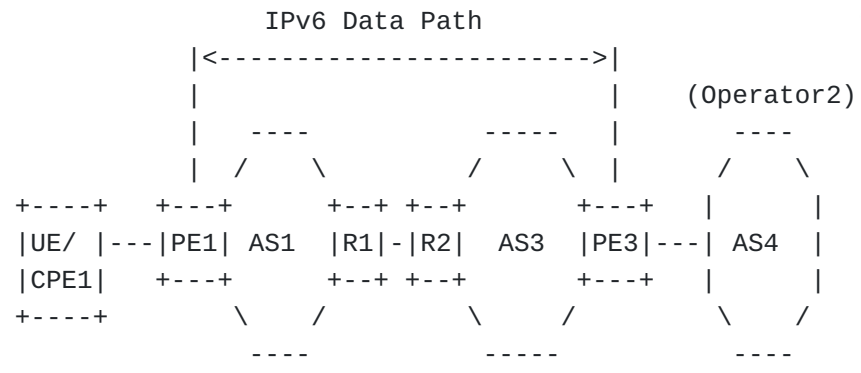

Figure 4. IPv6 Data Path from Ingress PE to Egress PE

In this case, there are only two IPv4-IPv6 conversion actions, which occur in PE1 and PE3 respectively.

2. IPv4 delivery from UE/CPE to egress PE

Another case is that IPv4 packets may have been transformed into IPv6 packet in UE/CPE, as done by CLAT of 464XLAT[RFC6877], before they reach the edge of the network.

In this case, the PLAT of 464XLAT and ADPT will converge in ingress PE, both the client-facing interface and the core-facing interface are IPv6. When IPv6 packet reaches the ingress PE, the ingress PE does not need to implement the conversion between IPv4 and IPv6 packets. For the source IPv6 address, because the address adopted by UE is generally GUA, and the source address of the IPv4-embedded IPv6 packet is IPv4-embedded address in the core of this framework, it is necessary to convert the source address from GUA to IPv4-embedded IPv6 address. In addition, because the quantity of IPv4-embedded IPv6 address is limited, it is necessary to take IPv6 address multiplexing here, one or more IPv4-embedded IPv6 addresses are shared among several IPv6-only clients with GUA addresses. For the destination address, with 464XLAT, UE synthesizes the

destination IPv4 address into IPv6 address by appending IPv4 address to the IPv6 prefix provided by DNS64 server. When the IPv6 packet reaches the edge the multi-domain IPv6 network, i.e. PE1, the destination IPv6 address is converted into IPv4-embedded IPv6 address. This process is implemented by looking for the mapping rule corresponding to the original destination IPv4 address in the MD database, and then substituting the NAT64 prefix with the IPv6 mapping prefix of the egress PE.

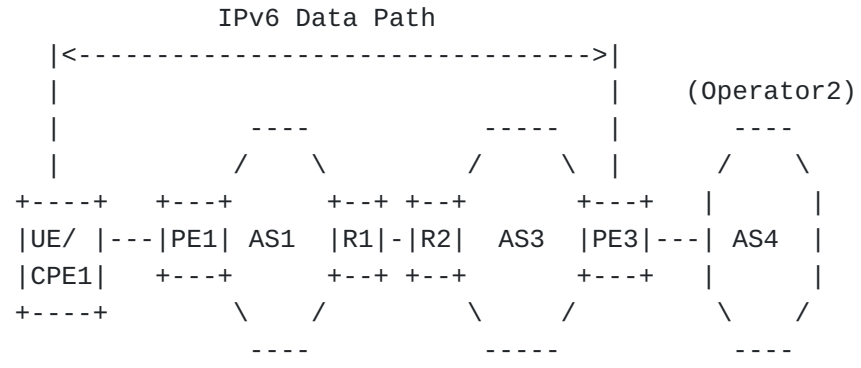

Figure 5. IPv6 Data Path from UE/CPE to Egress PE

In this case, there are only one stateless IPv4-IPv6 conversion action, which occurs in PE3. Compared with the case of independent deployment model mentioned in section 5, with the new framework the quantity of IPv4-IPv6 conversion points has been reduced from three to one.

8. Security Considerations

Besides regular security checks on configured mapping rules, the following two aspects need to be considered as well.

8.1. Authenticity and Integrity of Packets

In this framework, for each egress PE, they assume that all ingress PEs are legal and authorized to convert the received IPv4 packets into IPv6 packets and send them into IPv6-only network. If IPv6 packets cannot guarantee its authenticity or integrity, then there may be a spoofing attack. Some faked ingress PEs can send IPv6 data converted from IPv4 to attack the egress PE. After the egress PE recovers the received IPv6 packets into IPv4 packets, it routes based on the destination IPv4 address and enters the Internet. They use global IPv4 address, not private address. Therefore, these attacks cannot cause payload packets to be delivered to an address other than the one appearing in the destination address field of the IP packet. Since the PE in this framework is stateless, the effect of the attack is limited.

8.2. BGP-4 and Multiprotocol Extensions for BGP-4

The framework allows BGP to propagate mapping rule information over an IPv6-only underlay network, BGP is vulnerable to traffic diversion attacks. The ability to advertise a mapping rule adds a new means by which an attacker could cause traffic to be diverted from its normal path. Such an attack differs from pre-existing vulnerabilities in that traffic could be forwarded to a distant target across an intervening network infrastructure (e.g., an IPv6 core), allowing an attack to potentially succeed more easily since less infrastructure would have to be subverted. The security issues already exist in BGP-4 and MP-BGP for IPv6, the same security mechanisms are applicable.

9. IANA Considerations

There are no other special IANA considerations.

10. Acknowledgement

The authors would like to thank Brian E. Carpenter, Bob Harold, Fred Baker, Xipeng Xiao, Giuseppe Fioccola, Vasilenko Eduard, Zhenbin Li, Jen Linkova, Ron Bonica, Shuping Peng and Jingrong Xie for their review and comments.

11. References

11.1. Normative References

[I-D.ietf-bess-ipv6-only-pe-design] Mishra, G. S., Mishra, M. P., Tantsura, J., Madhavi, S., Yang, Q., Simpson, A., and S. Chen, "IPv6-Only PE Design for IPv4-NLRI with IPv6-NH", Work in Progress, Internet-Draft, draft-ietf-bess-ipv6-only-pe-design-03, 24 September 2022, <https://www.ietf.org/archive/id/draft-ietf-bess-ipv6-only-pe-design-03.txt>.

[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>.

[RFC3587] Hinden, R., Deering, S., and E. Nordmark, "IPv6 Global Unicast Address Format", RFC 3587, DOI 10.17487/RFC3587, August 2003, <https://www.rfc-editor.org/info/rfc3587>.

[RFC4026] Andersson, L. and T. Madsen, "Provider Provisioned Virtual Private Network (VPN) Terminology", RFC 4026, DOI 10.17487/RFC4026, March 2005, <https://www.rfc-editor.org/info/rfc4026>.

[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>.

[RFC6144] Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", RFC 6144, DOI 10.17487/RFC6144, April 2011, <https://www.rfc-editor.org/info/rfc6144>.

[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>.

[RFC7915] Bao, C., Li, X., Baker, F., Anderson, T., and F. Gont, "IP/ICMP Translation Algorithm", RFC 7915, DOI 10.17487/RFC7915, June 2016, <https://www.rfc-editor.org/info/rfc7915>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

11.2. Informative 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", Work in Progress, Internet-Draft, draft-ietf-v6ops-ipv6-deployment-08, 20 October 2022, <https://www.ietf.org/archive/id/draft-ietf-v6ops-ipv6-deployment-08.txt>.

[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", Work in Progress, Internet-Draft, draft-ietf-v6ops-transition-comparison-04, 23 May 2022, <https://www.ietf.org/archive/id/draft-ietf-v6ops-transition-comparison-04.txt>.

[IAB-statement] "IAB statement", <https://www.iab.org/2016/11/07/iab-statement-on-ipv6/>.

[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>.

[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>.

[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>.

[RFC7597] Troan, O., Ed., Dec, W., Li, X., Bao, C., Matsushima, S., Murakami, T., and T. Taylor, Ed., "Mapping of Address and Port with Encapsulation (MAP-E)", RFC 7597, DOI 10.17487/RFC7597, July 2015, <https://www.rfc-editor.org/info/rfc7597>.

[RFC7599] Li, X., Bao, C., Dec, W., Ed., Troan, O., Matsushima, S., and T. Murakami, "Mapping of Address and Port using Translation (MAP-T)", RFC 7599, DOI 10.17487/RFC7599, July 2015, <https://www.rfc-editor.org/info/rfc7599>.

[RFC8585] Palet Martinez, J., Liu, H. M.-H., and M. Kawashima, "Requirements for IPv6 Customer Edge Routers to Support IPv4-as-a-Service", RFC 8585, DOI 10.17487/RFC8585, May 2019, <https://www.rfc-editor.org/info/rfc8585>.

Authors' Addresses

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

Email: xiechf@chinatelecom.cn

Chenhao Ma
China Telecom
Beiqijia Town, Changping District
Beijing
102209
China

Email: machh@chinatelecom.cn

Xing Li

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

Email: xing@cernet.edu.cn

Gyan Mishra
Verizon Inc

Email: gyan.s.mishra@verizon.com

Mohamed Boucadair
Orange
France

Email: mohamed.boucadair@orange.com

Thomas Graf
Swisscom
Binzring 17
CH- CH-8045 Zurich
Switzerland

Email: thomas.graf@swisscom.com