v6ops
Internet-Draft
Intended status: Informational
Expires: August 1, 2017

X. Li
C. Bao
G. Han
M. Sheng
CERNET Center/Tsinghua
University

January 28, 2017

CERNET deployment of IVI/MAP-T in an IPv6-only network
draft-xli-v6ops-cernet-deployment-02

Abstract

This document presents the China Education and Research Network (CERNET)'s IPv4 as a Service (IPv4aaS) design, deployment and operation experience.

The techniques used are IPv4/IPv6 stateless translation both in the forms of single translation (IVI, for IPv6-only servers) and double translation (MAP-T, for dual-stack clients).

Table of Contents

1. Introduction

The China Education and Research Network (CERNET) is an academic network in mainland China, with the universities, institutes and schools as the customers. The student population in mainland China is about 320 million and there are no enough public IPv4 addresses available. The cloud computing, the mobile Internet and the Internet of Things make the IPv4 address exhaustion situation even worse. Ten years ago, we have deployed an IPv6-only backbone named CERNET2 and eight years ago, we have developed the IPv4/IPv6 stateless translation technology called IVI [RFC6219], which becomes the proposed IETF standard of the IPv4/IPv6 stateless translation [RFC6145], [RFC6052], etc. In order to improve the customer experience for the IPv4-only applications and application with the address literals embedded, we have developed double IPv4/IPv6 stateless translation technology called dIVI, which becomes the mapping address and port with translation (MAP-T) [RFC7599]. This document presents our experience of IPv4 as a Service (IPv4aaS), the techniques used are IVI [RFC6145], [RFC6052] and [RFC6219] and a slightly modified version of MAP-T [RFC7599].

1.1. Major Motivation

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 choice for the new campus network connected to IPv6-only CERNET2.

1.2. IPv4-as-a-Service Requirements

The design requirements for IPv4aaS in CERNET are:

1. Deploy IPv6-only single stack network as large as possible in order to reduce CAPEX and OPEX.

2. The clients should have the same user experience compared with the dual-stack network with NAT44.

3. The IPv6-only servers should be able to serve the IPv4-only clients in the Internet.

4. Make the use of the IPv4 public address as much efficient as possible.

5. Provide a path to allow IPv6-only clients to communicate to the IPv4 Internet.

6. For the scalability, resilience and security, prefer stateless technologies.

2. Architecture and Methodology

2.1. Major Design Considerations

For CERNET, we have the following situation:

- There are two nation-wide backbones, CERNET is an IPv4 backbone and CERNET2 is an IPv6-only backbone.

- Most of the campus networks are dual stack. The IPv4 interfaces of the border router of the campus networks are connected to CERNET and various commodity ISPs. The IPv6 interfaces of the border router of the campus networks are connected to CERNET2.

- For special purpose subnets which are providing service for the servers, the subnets can be dual stack, IPv4-only or IPv6-only. The IPv6-only subnets are recommended.

- For general purpose subnets which are providing services for the clients (both wired Ethernet and the WLAN with multiple SSID), the dual stack is recommended. Since it is the only way to maximize the satisfaction for the users using various operating systems (Windows XP, Windows 7/8, OSX, Linux, iOS, Android, etc.) and different applications (IPv4-only, dual stack and with address literals).

- Since we are running an academic network, we encourage users to try new technologies and make some part of the network as the testbeds.

- Ten years experience of deploying IPv6-only CERNET2, we strongly believe that the killer application of the IPv6 is the ability to communicate the global IPv4 Internet and therefore the translation technology should be the first choice. In addition, due to the scalability, security and manageability considerations, the stateless technology should be preferred. Those considerations result in the development and deployment of IVI [RFC6145], [RFC6052], [RFC6219] and MAP-T [RFC7599].

2.2. Regulatory Considerations

The government requires server operators to detect the packet sender by source IP (and port) and therefore stateless address mapping technologies are preferred. This will dramatically reduce the volume of material required to be held for logging compliance.

In addition, the stateless translation technology is preferred, since IPv6 addresses in the IPv6 packets everywhere in the network contain both the IPv6 and IPv4 address information without the requirement of decapsulation.

2.3. Security Considerations

From operation point of view, single stack (IPv6-only) is easier for ensuring the security compared with the dual stack.

The stateless mechanism can help for the trace back and identifying the source addresses (and port).

The translation mechanism can help for configuring the access list and rate-limiting without decapsulation.

2.4. Operational Considerations

The IPv4aaS in CERNET is mainly for deploying new networks. The existing IPv4 and dual stack networks will not be changed. The methods for upgrading the existing networks in order to release the public IPv4 addresses will be discussed in future documents.

2.4.1. Subnet for the Servers

For subnets for the servers, The IPv6-only subnet with stateless translation is recommended. The configuration is statically performed.

2.4.2. Subnet for the Clients

For general purpose subnets for the clients, dual stack is recommended, with DHCP for the IPv4 and SLAAC for the IPv6. There is no issue for the DNSSEC operation.

If IPv6-only subnet is required, the configuration should be done via DHCPv6 stateful mode with stateless or stateful IPv4/IPv4 translation service [RFC6145], [RFC6052], [RFC6146] [RFC6219] and DNS64 [RFC6147]. Note that there are issues for the DNSSEC operation.

We don't recommend mixing SLAAC and DHCPv6 stateful in the same subnet.

For the WLAN environment, the users can decide which subnet to use among the IPv4-only, the dual stack and the IPv6-only (with translation) by selecting different SSIDs.

2.5. End-User Experience Considerations

Due to the public IPv4 address depletion problem, the modern applications can fully support NAT44. Therefore, the end-users will be satisfied if the IPv4aaS provides the same service as NAT44 with IPv4 address sharing.

Besides TCP and UDP, concerning other protocols for example ICMP (ping), we found that the end-users will be satisfied if the IPv4aaS provides the same service as NAT44 with IPv4 address sharing.

3. Design and Deployment

Before the IPv4 address depletion, CERNET/CERNET2 has reserved a /17 public IPv4 address prefix and several /40s IPv6 prefixes for the IPv4/IPv6 stateless translation. The current implementations of both IPv6-only servers and dual-stack clients are still using these prefix. The topology is shown in Figure 1.

```
      ------                        -----           ------
    /        \       -----        /       \       /        \
   |  CERNET  |-----| IVI |------| CERNET2 |--+--|IPv6-only|
    \  IPv4  /       -----        \ IPv6  /   |   \servers /
      ------                        -----     |     ------
                                              |                ------
                                              |   ------     /        \
                                              +--|MAP-T |---|Dual-stack|
                                              |   ------     \clients /
                                              |                ------
                                              |
                                              |                ------
                                              |   ------     /        \
                                              .--|MAP-T |---|Dual-stack|
                                              |   ------     \clients /
                                              |                ------
                                              |
```

Figure 1: CERNET/CERNET2 Translation Topology

3.1. Single stateless translation for servers - IVI

The single stateless translation for servers using IVI is straightforward, as shown in Figure 2.

```
            ------                             -----           ------
          /  The   \       -----        /        \      /  The   \
         |  IPv4    |-----| IVI |------|IPv6-only|-----|  IPv6    |
          \Internet/       -----        \servers /      \Internet/
            ------                             -----           ------
```

Figure 2: IVI

The corresponding IETF documents are [RFC6145], [RFC6052] and [RFC6219].

The authoritive A record is derived from IPv6-only servers AAAA recorded and manually configured in the DNS server.

Thanks to the stateless technology, there are multiple 10G IVI translators and load balancing can be easily achieved.

3.2. Double stateless translation for clients - MAP-T

The double stateless translation for clients is a modified version of MAP-T [RFC7599]. Note that the original MAP-T is as shown in Figure 3.

```
         User N
       Private IPv4
      |  Network
      |
   O--+---------------O
   |  | MAP-T CE      |
   | +-----+--------+ |
   | NAPT44|  MAP-T | |
   | +-----+        | +-._   ,-------.                     .------.
   |      +--------+ |   ,-'         `-.                ,-'      `-.
   O-----------------O  /              \   O---------O /   Public   \
                        /   IPv6 only    \  |  MAP-T  |/     IPv4     \
                       (    Network       --+  Border +-   Network     )
                        \                /  |  Relay  |\              /
   O-----------------O  \              /   O---------O \             /
   |    MAP-T CE      |   ;".         ,-'                `-.       ,-'
   | +-----+--------+ | ,"   `----+--'                      ------'
   | NAPT44|  MAP-T | |,          |
   | +-----+        | +        IPv6 node(s)
   |   |  +--------+ |  (w/ v4-embedded-v6 address)
   O---+--------------O
       |
         User M
       Private IPv4
         Network
```

Figure 3: MAP-T Architecture

For the campus network environment, there are some modifications for the MAP-T

- o The IPv4 address sharing algorithm in the MAP-T BR is moved into MAP-T CE. Therefore the first translator (IVI in Figure 1) is with the public IPv4 sharing ratio=1, and the 1:N address sharing function is performed in the second translator (MAPT in Figure 1).

- o The IPv4-converted IPv6 addresses and IPv4-translatable IPv6 addresses have the same prefix and prefix length [RFC6052].

- o There are no additional IPv4 address sharing in NAPT44 of the MAP-T CE.

Based on the measurement statistics, the public IPv4 address sharing ratio is configured to 1024.

The multiple 10G IVI translators are shared with the single translation, the MAP-T translators are 1G equipment. We have deployed MAP-T translators in more than 100 campus networks.

4. Observations and Experiences

4.1. Effects on End-User

For the IVI/MAP-T end-users in more than 100 campus networks, they are satisfied with IPv4aaS. They did not notice that the services are provided via IPv4/IPv6 single/double translation.

4.2. Effects on Internal Staff

For IPv6-only servers provide IPv4 service via IVI, the management software is IPv6-only. Note that the IPv4 addresses managed in the system is a subset in the IPv6 address space and this will dramatically reduce the programming load, since there is no need to treat IPv4 and IPv6 differently.

For Dual-stack clients access IPv4 Internet via MAP-T, the management software is dual stack. The existing dual stack user management software can be used. The upper link of the second translator is IPv6-only, the IPv6 management software should be developed. Note that the IPv4 addresses managed in the system is a subset in the IPv6 address space and this will dramatically reduce the programming load, since there is no need to treat IPv4 and IPv6 differently.

4.3. Effects on Business

The IPv6-only server can provide the service directly for the IPv6 Internet and to the IPv4 Internet via IVI. Both CAPEX and OPEX are reduced compared with the dual stack.

The service for the clients is via MAP-T with address sharing ratio 1024. This greatly reduce the requirements of the public IPv4 addresses for deploying new networks.

5. Summary: Post-mortem Report

We are satisfied with the IVI/MAP-T deployment for IPv4aaS. More campus networks are expected to move in this direction.

5.1. Deviations from IETF Documents

The base specifications the IPv4aaS in the deployment are defined in [RFC6145], [RFC6052], [RFC6146], [RFC6219], [RFC6147] and [RFC7599]. As we discussed in this document, there are enhancements and modifications which will be presented in future documents.

5.2. The Suggestions for the IPv6 Transition

Based on ten years experience, we found that IPv4aaS is a good model for the IPv6 transition.

1. The new networks should be IPv6.

2. Use native IPv6 if possible.

3. Use single stateless translation (IVI) for the IPv6-only servers and clients.

4. Use double stateless translation (MAP-T) for the dual-stack clients.

5. The practical IPv6 transition path should be from double translation to single translation and finally to native IPv6.

6. IANA Considerations

This specification does not require any IANA actions.

7. Security Considerations

There are no other special security considerations.

8. Acknowledgements

9. References

9.1. Normative References

[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, <http://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, <http://www.rfc-editor.org/info/rfc6145>.

[RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, DOI 10.17487/RFC6146, April 2011, <http://www.rfc-editor.org/info/rfc6146>.

[RFC6147] Bagnulo, M., Sullivan, A., Matthews, P., and I. van Beijnum, "DNS64: DNS Extensions for Network Address

Translation from IPv6 Clients to IPv4 Servers", RFC 6147, DOI 10.17487/RFC6147, April 2011, <http://www.rfc-editor.org/info/rfc6147>.

[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, <http://www.rfc-editor.org/info/rfc7599>.

9.2. Informative References

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

[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, <http://www.rfc-editor.org/info/rfc6219>.

Authors' Addresses

Xing Li
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing, 100084
China

Phone: +86 10-62785983
EMail: xing@cernet.edu.cn

Congxiao Bao
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing, 100084
China

Phone: +86 10-62785983
EMail: congxiao@cernet.edu.cn

Guoliang Han
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 10-62785983
EMail: bupthgl@gmail.com

Maojia Sheng
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 10-62785983
EMail: ynsfsmj@foxmail.com