IEICE TRANS. COMMUN., VOL.E93–B, NO.5 MAY 2010

INVITED PAPER *Special Section on Technology and Architecture for Sustainable Growth of the Internet*

IPv4 to IPv6 Transformation Schemes

Shin MIYAKAWA[†a)], ***Member***

SUMMARY According to the recent observations of IPv4 (Internet Protocol version 4) address allocation status, it will be running out within few years. Consequently, to ensure the continuous extension of the Internet operation, introducing IPv6 (Internet Protocol version 6) protocol is surely needed. But at the same time, such transformation must be "smooth" for every Internet users and be compatible with today's IPv4 based practices. This paper describes several techniques and usage scenario which are discussed mainly in the IETF — Internet Engineering Task Force — and tried to be implemented as prototype products to transform today's Internet towards the IPv6 based one.

key words: *IPv6, IPv4 address exhaustion, NAT444, softwire, DS-LITE, A+P*

1. Introduction

In 1992, when the INET 92 meeting was held in Kobe, Japan, the Internet Society identified that IPv4 address space will be running out around year 2010. For example, Sect. 5.1 of an RFC1752 [1] "The Recommendation for the IP Next Generation Protocol" which was published in 1995 clearly noted that "the Internet would exhaust the IPv4 address space between 2005 to 2011" as the ALE WG [2] conclusion. According to this observation, technical standard body of the Internet — IETF (Internet Engineering Task Force) — started efforts to define a new Internet Protocol to replace existing IPv4 about 15 years ago. It must be noticed that this decision was made before WWW (World Wide Web) introduction to the world and more over, at this point, no commercial Internet service was deployed widely yet. To extend the lifetime of IP version 4, NAT (Network Address Translation) and the concept of "private IP address space" defined in RFC1918 [3] was introduced, but needs for new terminal devices such as cellular phone for example, and booming demands for emerging economy countries such as China and other countries have overcome these efforts for extension of the life of IPv4. Then, now it is considered that the global IPv4 address will be difficult to be assigned to (especially new) users around year 2010 again.

So, as we have prepared since mid 1990's, new IP protocol IPv6 [4] — IP version 6 — with its huge address space must be introduced in very near future, but of course, transition from IPv4 only network to IPv6 enabled one without troublesome is quite difficult task to be achieved. Today, TCP/IP technology is used so widely. So many implementations have been scattered into all over the world already. In many cases, it is not able to upgrade the equipments simply just because there is no technical way to replace the program code in its Read Only Memory for example, or even if it could be done technically but couldn't by financial reasons sometime. Also some implementations including commonly used today such as Microsoft's Windows XP operating system are able to use with IPv6 but still needs IPv4 support to run.

Therefore, to sustain the business over the Internet, everyone should take care about its own practices carefully and prepare for the proper modification on the design, implementation and operation of the network.

This paper describes several techniques aiming at from IPv4 to IPv4/v6 co-existing transition especially for the Internet Service Providers mainly. Because all the stake holders such as individual users, small, medium and large enterprises, organizations such as universities and governments, ASP (Application Service Providers) and more, rely on the ISP's network that provides them with the connection, the conversion of the ISP must be done firstly.

2. IPv4 Address Assignment Status Today

According to the today's observation of IPv4 address assignment [5] by Geoff Houston of APNIC (Asia Pacific Network Information Center) as of 13th of September 2009, the Projected IANA Unallocated Address Pool Exhaustion 31-Aug-2011 and Projected RIR Unallocated Address Pool Exhaustion is 04-Jul-2012.

Every ISP is assigned IP address space from RIR (Regional Internet Registry) which is under the IANA. According to the IP address assignment rule of JPNIC for example, usually, an organization (typically an ISP) that would ask new IP address space has to declare the usage plan and past consumption history for 12 months period. So, it is fair assumption that a typical ISP might have 1–2 years of buffer of IPv4 address space but 5 years at the maximum if the ISP grows its size of operation continuously. However some efforts to squeeze assigned but unused IPv4 address space like market transaction are now discussed but it could extend the last day only for few years. According to the estimation of Mr. Maemura of JPNIC is shown at the Japanese IPv6 summit conference in 2009 [6], he could expect at most 10 Class A (/8) space to be returned. As shown in the Fig. 1, today we consume about one or so Class A space per month,

Manuscript received October 16, 2009.
Manuscript revised December 28, 2009.

[†]The author is with NTT Communications, Tokyo, 108-8118 Japan.
a) E-mail: shin.miyakawa@ntt.com

DOI: 10.1587/transcom.E93.B.1078

Copyright © 2010 The Institute of Electronics, Information and Communication Engineers

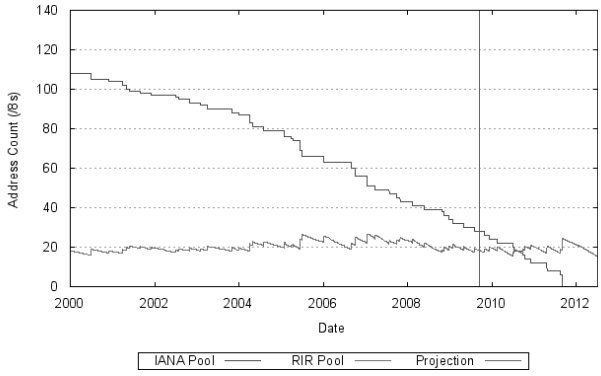

Fig. 1 Projected RIR and IANA consumption (/8s).

which means it lasts only 10 months at most. Therefore, we have to think that global IPv4 address space will running out within few years which means that new customer of the Internet will not be able to get the IPv4 address in that case.

3. Why Providing IPv4 Backward Compatibility with IPv6 is Needed

However the main object of IPv6 protocol is against IPv4 address space exhaustion problem, IPv6 itself has been newly defined and is not backward compatible with IPv4 protocol. A web site whose transport is IPv4 only cannot be accessed by the IPv6 only client computer directly. No v4 only E-mail client application can send any message to the v6 only mail servers. There are several schemes such as IPv4/v6 translator and Application Level Gateway, but there is no perfect IPv4 and IPv6 translation mechanism, because the IPv6 address space is wider than the IPv4's, so there is no 1:1 correspondence relationship between two of them. So an "application protocol independent translator" cannot translate a packet whose payload has a numeric expression of the IP address. For example, IPv6 enabled Google search engine (http://ipv6.google.com/) stores its cache data of the web pages with an IPv6 address number like;

http://[*2001:4860:b003::84*]/search?q=cache:FgEREgfx830J:ja.wikipedia.org/wiki/IPv6+ipv6&cd=1&hl=ja&ct=clnk

If a v4 only web browser accesses the link like this, the packet which is produced by this v4-only-client to access to this URL as a HTTP message must be converted a packet that contains only proper IPv4 address that corresponds to the IPv6 address in the message at the translator in between the IPv4 browser and the IPv6 server. However so, it is impossible simply because there is no way to match two types of addresses.

If this message looks like

http://[*cache.example.org*]/search?q=cache:FgEREgfx830J:ja.wikipedia.org/wiki/IPv6+ipv6&cd=1&hl=ja&ct=clnk

using only FQDN name, then, the AAAA (quad A, i.e. IPv6 address) record of this FQDN "cache.example.org" could be a clue to find corresponding IPv6 and IPv4 address with an A record (IPv4 address) of the same FQDN.

Another serious problem we have is the fact that Microsoft Windows XP operation system that still widely used in various places needs IPv4 transport to communicate the global DNS system.

Assuming the case that one client software like web browser on the Windows XP would start to communicate with some web server at "www.example.org," the client needs to communicate with DNS system first to retrieve the actual IP address which is corresponds to the FQDN www.example.org through the name resolver software on the operating system. Because name resolver on Windows XP has only IPv4 transport, even that FQDN www.example.org has IPv6 address as AAAA record and the web client can speak IPv6 TCP perfectly, the message to resolve the IPv6 address must be carried over IPv4 packets. This fact means that Windows XP needs IPv4 support to run.

Lastly, we believe that everyone can agree with the very simple fact i.e. it is quite difficult or almost impossible upgrade the entire Internet so that it works with IPv6 at one time.

Those are the strong motivations for ISPs to provide customers with some level of IPv4 backward compatibility even if we should shift our entire Internet system towards IPv6 based in near future.

Guessing how long we should keep IPv4 compatibility even if we have done the transition to IPv6 as major protocol on the Internet is the tough question to be answered. But, because Windows XP is on the market to be sold because its light weight process is suitable for UMPC (Ultra-Mobile PC) which is major part of the PC consumer market today in 2009, it means that ISP must keep support for Windows XP for a while, probably another 10 years or so at least. This can be guessed by the fact that still many Windows 2000 machines are actively used in various places today i.e. an operating system can be survived for 10 years easily.

Therefore, it could be safe to state that even if IPv4 address space will be running out around 2015, still we would keep the service by 2020 or so.

4. Toolbox

Before stepping into the technologies which is proposing, discussing, implementing and evaluating for IPv4 extension recently, author would like to introduce several existing technologies which have already known to the public. These technologies are the building blocks of the IPv4/v6 transition schemes and play great roles.

4.1 Translator and Application Level Gateway (ALG)

Translator is the machine converts an IPv4 packet to the IPv6 one and vice versa. If it will translate a packet to the other without any state, it would be ideal. But unfortunately

it is impossible to create such an ideal translator to realize one-to-one mapping. Thus, the design of a translator is quite state-full and usually with some application level gateway function to be useful and often needs to work with proper DNS system to identify especially an IPv4 packet which is corresponds to an IPv6 packet. So, however the translator technology will be effective for some limited areas of the Internet for example, hosting service provider and an enterprise network, it is not suitable in the middle of the Internet Service Providers' network. Also translator in ISP network could be security hole because it can be used as a step stone to attack somebody else hiding its original source address coming from.

4.2 6to4 and Teredo

6to4 which is defined in RFC3056 [7] is a mechanism to allow connecting IPv6 networks over the IPv4 connection. It uses so called 6to4 relay router whose address has global IPv4 address in the prefix so that any 6to4 relay router can send IPv6 packets to the other 6to4 relay without any IPv6 direct link in-between. Of course, with IPv6 links, 6to4 relay send IPv6 packets to the ordinal IPv6 network as well. 6to4 is widely implemented in Microsoft Windows operating system and many 6to4 relays are active in the Internet.

However 6to4 is widely deployed, because 6to4 uses proto41 ip-in-ip tunnel packet format, it is difficult or almost impossible to use it behind NAT within a customer premises. So, utilizing UDP, "Teredo" was proposed in RFC4380 [8] to traverse NAT mechanism to provide IPv6 over IPv4 connection. Teredo is also widely implemented in Microsoft Windows systems and is active as default setting in Vista and after.

4.3 Softwire

"Softwire" is a scheme that IPvX carries over IPvY and now standardized in IETF SOFTWIRE Working Group [9]. Among this working group items, RFC5571 [10] describes so called the "Hub and Spoke" model that is a very simple way to realize a virtual access concentrator in IPvX that terminates a tunnel connection from any device called Softwire Initiator to carry IPvY packet within the tunnel which is implemented by L2TPv2 over IPvX technology. Because L2TP especially which is able to carry IPv6 over IPv4 has implemented in various operating systems already such as Windows Vista, Linux, BSD and so on, this is quite conservative and proven tool to use in several transition method. Some ISPs like NTT Communications Corporation has been commercialized this Softwire service as "OCNIPv6" already for example.

4.4 CPE (Customer Premises Equipment)

CPE stands for Customer Premises Equipment. Often it calls as SOHO Router but not limited to. Sometime CPE just do semantic conversion from PPPoEoA (Point to Point Protocol over Ethernet over ATM) to PPPoE (Point to Point Protocol over Ethernet) that is called Bridged CPE. In this paper, a CPE router is defined as equipment to terminate a (sometime virtual) connection (which is usually PPP but not limited to especially CATV environment) with a router function. On the other hand, a bridged CPE is defined as equipment does nothing at Layer 3 of TCP/IP protocol suite. Usually, a CPE router acts as NAT (Network Address Translation) so that it can serve LAN with private address space defined in RFC1918 like 192.168.0.0/24. From the ISP point of view, a CPE router acts as a host even there are multiple hosts behind it. Sometime, a customer uses his/her own PC to terminate PPPoE session from ISP on the PC itself with only bridged CPE for economic reason. In this case, a PC will be assigned a global IPv4 address. It should not be recommended this scheme because of security issues with this, however so, at least in Japan; there are certain amounts of customers who employ this model of subscription. This fact is quite important to think about smooth v4-v6 transition schemes because we may not assume a router CPE in front of the customer some time which means we cannot expect any function in a CPE to do something in some case.

5. IPv4 Life Extension Proposals

The goal of the IPv4 life extension schemes are the same. Of course, each scheme has its own pros and cons, but all proposals are aiming at sharing one single global IPv4 address with multiple users. Especially in IETF, there are several proposals are now discussed, implemented and tested in actual environment. Some people misunderstand the purpose of these proposals just like if we have these schemes, anything else never is needed any more. It should be noted again that because of the problems all proposals have which discussed later and simply every proposals will cost a lot, transition to the IPv6 is considered as the final solution to keep continuous development of the Internet with proper cost.

5.1 Large Scale NAT (NAT444)

Most simple way to provide IPv4 backward compatibility with customers is implementing NAT device within ISP network. It should be noted that if a user does not use CPE router, this is the only choice to keep IPv4 connectivity at some level of the service because all the other proposals relies on some function on the router CPE device.

Because this scheme realizes so-called double NAT situation and usually CPE does not relay uPnP message towards its WAN interface, a mechanism to traverse NAT using uPnP IGD control message to make peer-to-peer application works will not work with this scheme. However, according to the IETF BEHAVE WG [11] recommendations, there is some security concern on usage of the uPnP to do NAT traversal, it is recommended to use STUN [12] and/or TURN [13] mechanism to realize P2P applications to communicate each other. If NAT444 device is transparent enough (it will be discussed later), there should be no

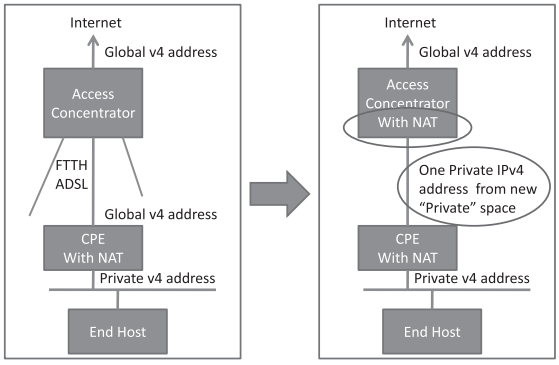

Fig. 2 NAT444.

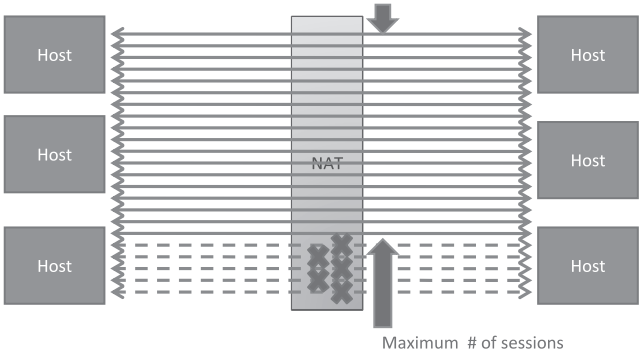

Fig. 3 Session number limitation.

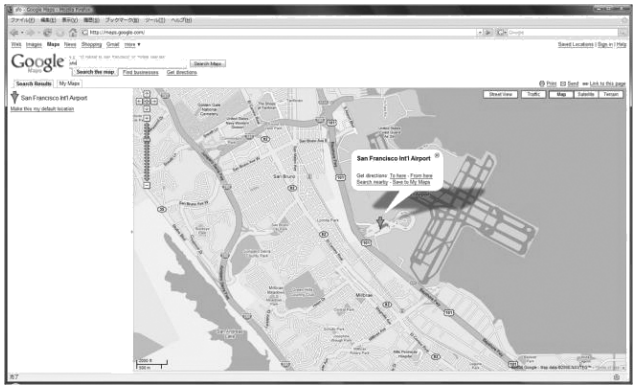

Fig. 4 MAX 30 sessions.

problem about NAT traversal.

In this paper, one step-by-step conversion scenario with this NAT444 and IPv6 over IPv4 softwire service from the IPv4 only Internet to the v4/v6 co-existing Internet will be introduced in the later section.

5.2 Dual-Stack Lite (DS-Lite)

Dual-stack lite [14] is now discussed and standardized in IETF SOFTWIRE WG. DS-lite is assuming IPv6 only access network with CPE router. IPv4 connection will be carried on the IPv4 over IPv6 softwire (L2TP or proto41 ip-in-ip) to the DS-lite concentrator called AFTR (Address Family Translation Router) at the provider edge. AFTR has NAT capability or A+P function if a DS-lite CPE is compatible with it. Now several vendors are implementing this scheme and ISC (Internet Software Consortium) producing open software based AFTR and CPE software. Theoretically, this can be terminated at the host operating system but as of year 2009, there is no implantation for Microsoft Windows or Apple Macintosh. This means that DS-lite must use compatible CPE router. Thus, deployment of this technology needs network wide upgrade with IPv6 only connection and brand-new CPE router distribution.

5.3 A+P

A+P [15] stands for "Address plus Port." Adding 2 octet port number range (which is 16 bit) of UDP and TCP into routing mechanism, simply 32 bit IPv4 address range will be extended to 48 bit as the total size of the "address." This scheme is intended to preserve end-to-end principle as much as possible rather than implementing NAT function in the ISP cloud but at the CPE router. However by definition, it work only with compatible CPE A+P router and no chance to work with existing operating system directly without upgrade, it could solve with optional NAT444 function at the concentrator.

6. Common Issues on Address Shortage Solutions

In any case, the goal to extend the life of IPv4 is to share an IPv4 address with multiple users. But by doing so, several serious problems will be occurred. This section discuss about issues among address sharing schemes.

6.1 Session Number Limitation

The most serious and inevitable problem is the session number restriction. Recalling that all proposals are aiming at sharing one single global IPv4 address by multiple users whereas one global IP address is used solely by one user today. Because port number of TCP and UDP is only 2 octet range which means 0-65535, if we assume that n users would share one address, we can assign only $65535/n$ sessions per user at the worst case (this happens, if all the sessions go to the same IP address and same port). Thus, if there is session consuming application running at the customer premise, there are certain possibilities it does not work well.

To see what happens if we limit the number of the TCP sessions, we have done an impressive experiment that to show how a web page can be showed at a web browser when we limit the number of the session that can be passing through a NAT device. This time, we selected Google Map (http://maps.google.com/) as the web page and Internet Explorer 7.0 running on top of the Windows Vista operating system. Firstly, when we limit the number of available TCP session number as 30, it seems everything works fine.

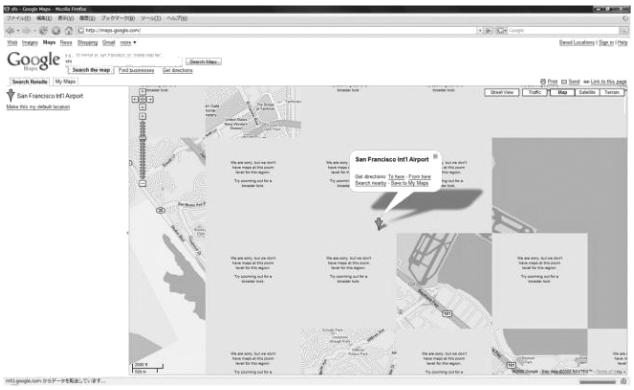

Fig. 5 MAX 15 sessions.

Table 1 Observation of the numbers of the session consumed.

Application	# of sessions
No operation	5〜10
Yahoo top page	10〜20
Google image search	30〜60
Nico Nico Doga	50〜80
OCN photo friend	170〜200+
iTunes	230〜270
iGoogle	80〜100
Rakuten	50〜60
Amazon	90
HMV	100
YouTube	90

Next, when we limit the allowable session as 15, we could see several blocks white out in the map.

This happens because Google MAP web page uses AJAX technology that allows the machine to use multiple concurrent TCP sessions to get the information faster from server to client escaping from TCP congestion control mechanism. So, when the number of TCP session has been limited, first TCP sessions can be succeeded to communicate but later TCP sessions fails. In that case, it causes that several white out blocks show up at their corresponding places on the page. According to our survey, some web site uses few hundreds of TCP sessions concurrently at one time as shown in Table 1, we have to carefully define how many sessions will be allowed per user if we deploy any address sharing scheme in the actual environment and it lefts for further study to determine how many TCP sessions should be allowed to satisfy actual customers' satisfaction.

6.2 High Availability

To secure the customers' satisfaction, any devices in the middle of the ISP network should consider about its high availability (HA). Without HA, for example, if the case some vulnerability was found in the software at the device, we have to shut down the machine to apply security patches on it and kill the customers' sessions at the same time, that is quite frustrated. On the other hand, with good HA function between a pair of machines in the network, we can shutdown one machine and upgrade that first and resume to the network once, then, we can stop the other one to apply security patches then return to the original settings without disturbing the service to the customers. But different from simple router which has no state in general to pass the packet, usually, Large Scale NAT and the other concentration box are quite state-full so it is quite difficult and/or costly to implement HA with them.

6.3 Maximum Transparency

To avoid serious drawbacks on the application when we introduce the address sharing scheme, we should any concentrator as "transparent" as possible. However A+P concentrator is quite transparent, but all A+P CPE, DS-lite and NAT444 should be no barrier for application just like SOHO router which passes many applications as much as possible to satisfy residential users enjoying VoIP, chat, p2p, on-line real-time gaming and so on. This is quite different from NAT for Enterprise that usually blocks many applications to control the behavior of the employees.

We are now proposing an internet-draft [16] to compile requirements to make those schemes to pass many applications as much as possible as long as the implementation cost allows.

7. An Example of Step by Step Transition Scenario

In this section, one possible step-by-step transition scenario from IPv4 only network to IPv4/v6 co-existing one utilizing NAT444 and softwire is introduced as an example (Original of this scenario was firstly presented by the author at the technical plenary at the IETF72 in 2008).

Talking about ipv4 to v6 transition scenarios, however there are many proposals like one is shown in [17] for example, there is no commercially reasonable scenario for ISP network which allow us to keep existing user's (v4-only) CPE devices and old equipments at the first part of the transition, as far as author knows except the following one.

7.1 IPv4 Only

The first stage is the IPv4 only case. Everything from the Internet itself, ISP backbone, access concentrator, CPE and end-host is IPv4 only. Let us start from this.

7.2 IPv4/v6 Dual Stack Backbone

It is very easy to upgrade the backbone because today's backbone routers have already implemented IPv6 function within the system. Usually the function is just not turned on. So, if operators decide to enable IPv6 in the ISP's core network, what they need to do is just turn on the functions on the equipments basically.

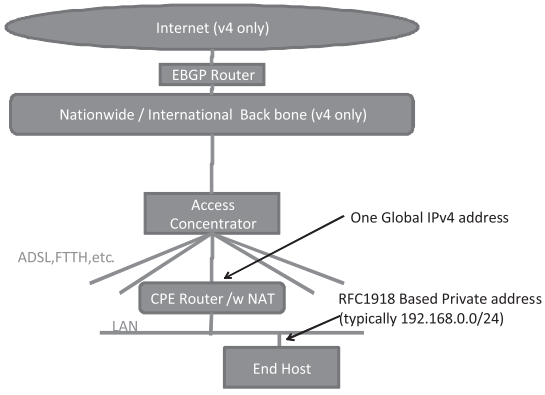

Fig. 6 IPv4 only.

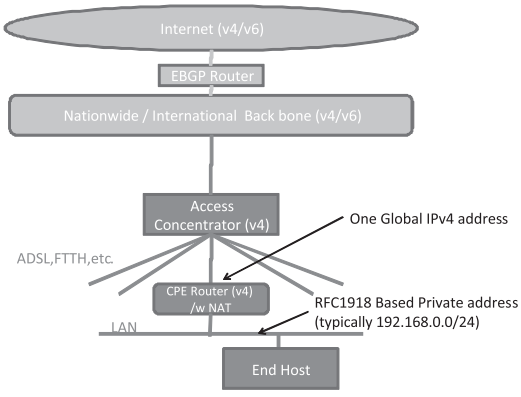

Fig. 7 Dual stack backbone.

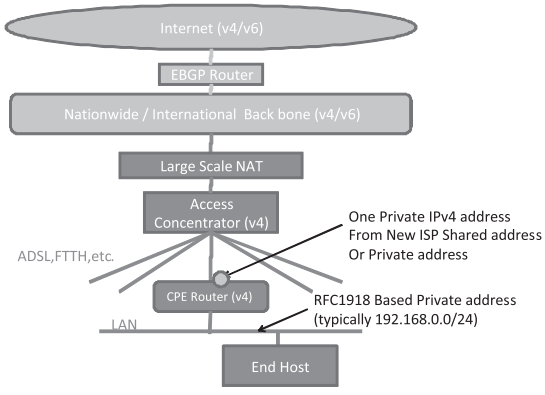

Fig. 8 NAT444.

7.3 Introduction LSN into the Network

Introduce NAT444 (Large Scale NAT) into the network. This case, a private address will be assigned to the user. Please note that nothing needs to be changed at the customer premises.

7.4 Introduction Softwire Concentrator

To avoid limitations comes from NAT444, IPv6 softwire

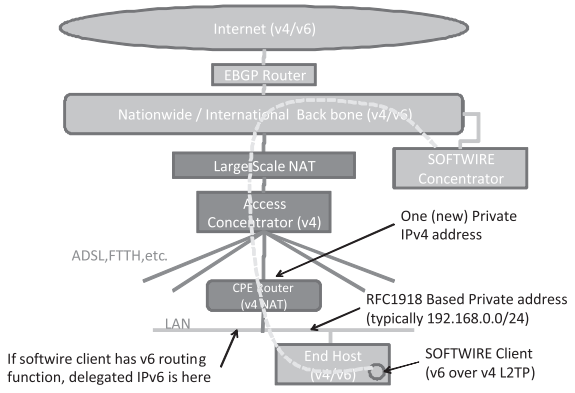

Fig. 9 Introduction softwire.

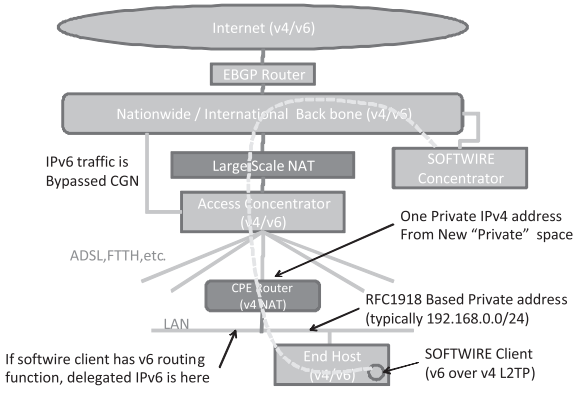

Fig. 10 Native connection but old CPE

function is introduced. As discussed before, L2TP is widely implemented at various operating systems; it is easy to terminate IPv6 softwire virtual connection at customer's machine with no additional cost. For example, Windows Vista can act as router for IPv6, it is also possible all LAN could be dual-stack.

7.5 Native Connection but CPE Remains as Originalx

Then, it is possible area to area upgrade access concentrator to implement native IPv6 access for ISP. With softwire, it is ok even if some customer remains as only IPv4 compatible CPE. ISP does not need to enforce upgrade CPE devices at one time. Rather than that, it can be left for up to the customer decision when he will change the CPE device so that native IPv6 connection can be processed at the CPE.

7.6 Dual Stack Completion

If a customer or ISP upgrades CPE device, the dual stack environment has been realized perfectly. This remains for another decade or so until we can disable IPv4 completely in some future.

8. Conclusion

Several proposals now proposing, discussing and imple-

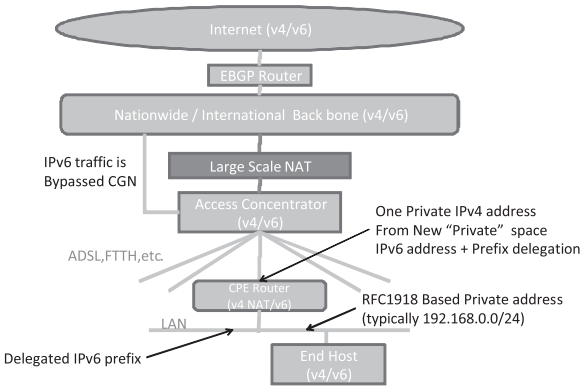

Fig. 11 Complete dual stack.

menting against the IPv4 address completion problem that is expected to be real within few years and one possible scenario for IPv4 only to IPv4/v6 transition have been introduced in this paper. Every proposals are not exclusive each other. Rather than that, we should use those place to place, time to time in the same network maybe. Although still many issues lefts for further investigation and research, to keep the healthy and continuous growth of the Internet, all the stake holders should prepare for the problem.

Acknowledgments

Author would like to thank all the members of NTT Communications' IPv4 address completion/v6 transition project who are actually working on this issues and all the people supports me in and out the company and many colleagues who allow me to have quite fruitful discussions and authoring some works like RFCs and internet drafts together in IETF and many other meetings in various places on the Earth.

References

[1] S. Bradner and S. Mankin, "The recommendation for the IP next generation protocol," RFC1752, Jan. 1995.
[2] IETF ALE WG: http://www.ietf.org/wg/concluded/ale.html
[3] Y. Rekhter, et al., "Address allocation for private Internets," RFC1918, Feb. 1996.
[4] S. Deering and R. Hinden, "Internet protocol, version 6 (IPv6) specification," Dec. 1998. (obsoletes RFC1883)
[5] G. Huston, "IPv4 address report," http://www.potaroo.net/tools/ipv4/index.html
[6] M. Maemura, "IPv4 address shortage," http://www.iajapan.org/ipv6/summit/2009/pdf/maemura.pdf
[7] B. Carpenter and K. Moore, "Connection of IPv6 domains via IPv4 clouds," RFC3056, Feb. 2001.
[8] C. Huitema, "Teredo: Tunneling IPv6 over UDP through network address translations (NATs)," RFC4370, Feb. 2006.
[9] IETF SOFTWIRE WG: http://www.ietf.org/dyn/wg/charter/softwire-charter.html
[10] B. Storer, et al., "Softwire H & S framework with L2TPv2," RFC 5571, June 2009.
[11] IETF BEHAVE WG: http://www.ietf.org/dyn/wg/charter/behave-charter.html
[12] J. Rosenberg, et al., "STUN — Simple traversal of user datagram protocol (UDP) through network address translators (NATs)," RFC 3489, March 2003.
[13] J. Rosenberg, R. Mahy, and P. Matthews, "Traversal using relays around NAT (TURN): Relay extensions to session traversal utilities for NAT (STUN)," draft-ietf-behave-turn-16, July 2009.
[14] A. Durand, ed., "Dual-stack lite broadband deployments post IPv4 exhaustion," draft-ietf-softwire-dual-stack-lite-01, July 2009.
[15] R. Bush, ed., "he A+P approach to the IPv4 address shortage," draft-ymbk-aplusp-04, July 2009.
[16] T. Nishitani, et al., "Common functions of large scale NAT (LSN)," draft-nishitani-cgn-02, May 2009.
[17] C. Huitema, et al., "Unmanaged networks IPv6 transition scenarios," RFC3750, April 2004.

Shin Miyakawa received the B.S., M.S., and Doctoral degrees in Department of Computer Science from Tokyo Institute of Technology in 1990, 1992 and 1995, respectively. After he joined to NTT Software Laboratories, Nippon Telegraph and Telephone Corporation in 1995, he had stayed in NTT Multimedia Communications Laboratories in Silicon Valley, California, USA during 1997–2002, Since 2002, he leads a research and development team for Internet Protocol Technologies as the director of IP Core Technology team, Network and Systems Laboratory in Innovative IP Architecture Center, NTT Communications Corporation in Tokyo, Japan.