Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: September 13, 2012

S. Tsuchiya, Ed.
M. Townsley
Cisco Systems
S. Ohkubo
Sakura Internet

March 12, 2012

IPv6 Rapid Deployment (6rd) in a Large Data Center
draft-sakura-6rd-datacenter-04

Abstract

IPv6 Rapid Deployment (6rd) as defined in RFC 5969 focuses on rapid deployment of IPv6 by an access service provider which has difficulty deploying native IPv6. This document describes how 6rd can be used to deliver IPv6 within a Large Data Center.

Table of Contents

1. Introduction

IPv6 Rapid Deployment (6rd) as defined in RFC 5969 focuses on rapid deployment of IPv6 by an access service provider which has difficulty deploying native IPv6. This document describes how one service provider in Japan, Sakura Interent, Inc., not for a large residential deployment, but for a large data center network.

While the protocol mechanism of 6rd is unchanged, the deployment model varies a bit from the classical "residential home access provider" model.

The motivation for using 6rd is very similar to that of the residential case where the service provider would like to offer IPv6 quickly to those users who want it, but without replacing equipment that currently does not support IPv6.

This document is provided as information to the Internet community.

2. Network Architecture

The case study presented here is based on the services provide by Sakura Internet Inc. Sakura Internet provides Internet services through Internet backbones and large data centers.

Sakura offers four types of services:

1. Housing Service, which provides Collocation and Internet Access on 5 urban datacenters (4 in Tokyo, 1 in Osaka)

2. Hosting Service, which provides shared service on the servers.

3. Dedicated Server Service, which provides customer dedicated server with variable OSs.

4. Virtual Private Server Service (VPS), which provides guest operating system on the Kernel-based Virtual Machine (KVM).

At the time of this writing, Sakura serves more than 200 Gpbs of traffic on its backbones, and around 50,000 dedicated servers, Virtual Private Servers, and collocated servers.

Figure.1 describes server-based 6rd in datacenter's network architecture.

```
                                   .--.
                                 _(.   `)
                               _(  IPv6  `)_
                              (  Internet   `)
                             ( `  .        )  )
                              `--(_______)---'
                                     |
                               +----------+
                               |  6rd BR  |
                               +----------+
                                     |
             +----------+          .--.
             |6rd CE on |        _(.   `)
             |guest OS  |      _(  IPv4  `)_
             +----------+     (  Backbone   `)     +----------+
             |   KVM    |====( `  .        )  )====|6rd CE on |
             +----------+     `--(_______)---'     |dedicated |
                                     ||            |server    |
                                     ||            +----------+
                                 +----------+
                                 |6rd CE on |
                                 |Housing GW|
                                 +----------+
                                       |
                                 -------------IPv4/IPv6 dual stack
                                   |       |
                               [Server] [Server]
```

Figure 1

Sakura has deployed commercial 6rd Border Relays, and relies on CE functionality in gateway routers or directly within the operating system of the servers.

In the latter model, there is no need for a CE gateway as the 6rd function is implemented directly in the server operating system itself.

-For Housing users, there are two options. Either the 6rd CE function is performed on the Gateway router itself, or the servers themselves can run 6rd directly.

-For Hosting users, IPv6 service can start by deploying 6rd CE function on the server OS or guest OS on the KVM.

Server administrators can start IPv6 service on demand themselves by using server-based 6rd.

There were some issues when Sakura considers IPv6 deployment on their backbone.

1.Some backbone Layer3 switches are too old.

IPv6 Switching would be software switching even if IPv4 Switching in hardware. It needs replacement.

2.Some backbone Layer3 switches required software upgrade.

IPv6 supports on hardware. But software upgraded is needed. In datacenter, there is different requirement on each server, even if the server connected to the same switch. Because the server administrator are completely different. Each server is providing different service to the different customers. So backbones maintenance time negotiation to the customer is very difficult.

3.Design Consideration

There are some scalability issues within the large datacenter.Especially backbone Layer 3 switches has huge arp and IPv4 routing entries.When the switch would upgrade and enable IPv6,NDP entry will be added,many resource such as NVRAM and TCAM will be consumed.

IPv6 datacenter also has security issues such as [draft-ietf-v6ops-v6nd-problems] and [RFC6104].

To provide native IPv6 service to the existing customer today rapidly, it needs cost,time,negotiation with customer and design consideration.

This is the reason why Sakura decided to provide server-based 6rd to the existing customer.

3. 6rd Availability in Server Operating Systems

In particular for the server-initiated case, Sakura relies on 6rd availability in Server operating systems.

Linux kernel has started to support 6rd since 2.6.33. So if Linux based Operating Systems are using 2.6.33 and the later,it can provides server-based 6rd.

FreeBSD and CentOS could not provide 6rd in default, but the patch exist.

Operating Systems	Linux Kernel	Description
Fedora14 and the later	2.6.35 and above	Server-based 6rd ready
Ubuntu 10.10 and the later	2.6.35 and above	Server-based 6rd ready
Debian6.0	2.6.32	Kernel update needs
CentOS5.6	2.6.18	needs [CentOS patch1][CentOS patch2]
FreeBSD8	N/A	needs [BSD patch]

4. Deployment Consideration

4.1. IPv4 compression address

6rd protocol specification is defined on [RFC5969]. Section 4 of [RFC5969] describes o-bit which can compress 32 bit IPv4 address in the 6rd delegated prefix. Linux Kernel also supports this feature.

So customer could get some IPv6 prefixes even if datacenter's prefix is /32.

But [BSD patch] doesn't has the feature of aggregate IPv4 address, therefore datacenter provider has to prepare /32 IPv6 prefix at least in that case.

In Sakura's case, 6rd prefix address using /32, and no compression IPv4 address.Thus the delegated 6rd address length is /64. It is enough address space for server-based 6rd.

4.2. Configuration

Section 7.1 of [RFC5969] describes 6rd CE automatic configuration method such as DHCP,TR-69 and so on.

But server-based 6rd does not needs automatic configuration because the server usually configure IPv4 address statically.

6rd is stateless tunnel technology,so if the operator once has placed/configured 6rd BR on the exit point to the IPv6 internet,they don't need additional configuration to 6rd BR even if 6rd CE(server based 6rd). will be increased.

4.3. MTU consideration

Section 9.1 of [RFC5969] describes about Maximum Transmission Unit(MTU) on 6rd tunnel. This guide also applicable for server-based 6rd.

But datacenter's IPv4 network is well-managed and is known by the server administrator. So 6rd CE's tunnel MTU could set be -20 byte from IPv4 MTU.

If the 6rd CE would be TCP server such as WWW, TCP MSS(Maximum Segment Size will be calculated automatically from tunnel MTU.

4.4. Traffic Pattern

6rd CE can communicate directly within 6rd domain.Server based 6rd also supports this feature.Therefore if the server would like to communicate anohter server in the data center,then encaplated IPv4 packets will reach 6rd server directly and de-capsulate.

4.5. Performance

6rd is tunnel technology, therefore the problem of performance is often taken up.It depends on server spec,but [ApacheBench] shows similar result with IPv4,IPv6 and 6rd.In Sakura internet server service,interface specifications are shared 100Mbps for VPS server,1Gbps for dedicated server. The result is enough to provide to customer.

5. Acknowledgements

The authors thank Hiroki Sato and Masakazu Asama,who made BSD&CentOS patch.

6. IANA Considerations

This document has no actions for IANA.

7. Security Considerations

There is same consideration as [RFC5969].

8. References

8.1. Normative References

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

[RFC5969] Townsley, W. and O. Troan, "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) -- Protocol Specification", RFC 5969, August 2010.

8.2. Informative References

[ApacheBench]
""ab - Apache HTTP server benchmarking tool"", <http://httpd.apache.org/docs/2.2/programs/ab.html>.

[BSD patch]
""BSD patch"", <http://people.allbsd.org/~hrs/FreeBSD/stf_6rd_20100923-1.diff>.

[CentOS] ""The Community ENTerprise Operating System"", <http://www.centos.org/>.

[CentOS patch1]
""CentOS Kernel patch"", <http://enog.jp/~masakazu/6rd/kernel-2.6.18-238.9.1.el5.6rd.x86_64.rpm>.

[CentOS patch2]
""CentOS iproute patch"", <http://enog.jp/~masakazu/6rd/iproute-2.6.18-11.6rd.x86_64.rpm>.

[Debian] ""Debian -- The Universal Operating System"", <http://www.debian.org/>.

[Fedora] ""Fedora Project Homepage"", <http://fedoraproject.org/>.

[FreeBSD] ""The FreeBSD Project"", <http://www.freebsd.org/>.

[I-D.ietf-armd-problem-statement]
Narten, T., "Problem Statement for ARMD", draft-ietf-armd-problem-statement-00 (work in progress), October 2011.

[I-D.ietf-v6ops-v6nd-problems]
Kumari, W., "Operational Neighbor Discovery Problems", draft-ietf-v6ops-v6nd-problems-00 (work in progress), October 2011.

[Linux 2.6.33]

""sit: 6rd (IPv6 Rapid Deployment) Support."", <http://kernelnewbies.org/Linux_2_6_33>.

[RFC3849] Huston, G., Lord, A., and P. Smith, "IPv6 Address Prefix Reserved for Documentation", RFC 3849, July 2004.

[RFC5569] Despres, R., "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)", RFC 5569, January 2010.

[RFC5737] Arkko, J., Cotton, M., and L. Vegoda, "IPv4 Address Blocks Reserved for Documentation", RFC 5737, January 2010.

[RFC5952] Kawamura, S. and M. Kawashima, "A Recommendation for IPv6 Address Text Representation", RFC 5952, August 2010.

[RFC6104] Chown, T. and S. Venaas, "Rogue IPv6 Router Advertisement Problem Statement", RFC 6104, February 2011.

[Ubuntu] ""Ubuntu Homepage"", <http://www.ubuntu.com/>.

Appendix A. Additional Stuff

A.1. OS configuration

A.1.1. Network Topology&Parameters

Describes configuration of each on OS,for reference.

```
                                .--.
                              _(.   `)
                            _(  IPv6  `)_
                           (  Internet   `)
                          ( `  .        )  )
                           `--(_______)---'
                                  |
                            +----------+
                            |  6rd BR  |
                            +----------+
                                  |    203.0.113.1
                                .--.
                              _(.   `)
                            _(  IPv4  `)_
                           (  Backbone   `)    6rd prefix:2001:db8::/32
                          ( `  .        )  )
                           `--(_______)---'
                                  |
               --------------------------------------------    203.0.113.0/24
                    |.10     |.11    |.12     |.13      |.14
                [Ubuntu] [Fedora] [Debian] [CentOS] [FreeBSD]
```

Figure 2

common parameter

BR IPv4 address	6rd prefix	IPv4MaskLen
203.0.113.1	2001:db8::/32	0

individual parameter

OS	IPv4 address	6rd delegated prefix
[Ubuntu]	203.0.113.10	2001:db8:cb00:710a::/64
[Fedora]	203.0.113.11	2001:db8:cb00:710b::/64
[Debian]	203.0.113.12	2001:db8:cb00:710c::/64
[CentOS]	203.0.113.13	2001:db8:cb00:710d::/64
[FreeBSD]	203.0.113.14	2001:db8:cb00:710e::/64

A.1.2. configuration procedure

A.1.2.1. Ubuntu

-modify "/etc/network/interfaces"

```
# vi /etc/network/interfaces
auto tun6rd
iface tun6rd inet6 v4tunnel
  address 2001:db8:cb00:710a::1
  netmask 32
  local 203.0.113.10
  endpoint any
  gateway ::203.0.113.1
  ttl 64
  up ip tunnel 6rd dev tun6rd 6rd-prefix 2001:db8::/32
  up ip link set mtu 1280 dev tun6rd
```

-reboot

A.1.2.2. Fedora

-make "/etc/sysconfig/network-scripts/ifcfg-sit1"

```
# vi /etc/sysconfig/network-scripts/ifcfg-sit1
DEVICE=sit1
IPV6INIT=yes
IPV6_MTU=1280
IPV6_DEFAULTGW=::203.0.113.1
IPV6TUNNELIPV4=any
IPV6TUNNELIPV4LOCAL=203.0.113.11
IPV6ADDR=2001:db8:cb00:710b::1/32
```

-modify "/etc/rc.local"

```
# vi /etc/rc.local
ip tunnel 6rd dev sit1 6rd-prefix 2001:db8::/32
```

-reboot

A.1.2.3. Debian

The latest version of Debian is 6.0.Debian6.0's kernel is 2.6.32.So it is required upgrade kernel.

-modify "/etc/apt/sources.list"

```
# vi /etc/apt/sources.list
deb http://ftp.jp.debian.org/debian experimental main
deb-src http://ftp.jp.debian.org/debian experimental main
```

-upgrade kernel

```
# apt-get update
# apt-get -t experimental install linux-image-2.6.38-rc6-amd64
```

-reboot

-modify "/etc/network/interfaces"

```
# vi /etc/network/interfaces
auto tun6rd
iface tun6rd inet6 v4tunnel
  address 2001:db8:cb00:710c::1
  netmask 32
  local 203.0.113.12
  endpoint any
  gateway ::203.0.113.1
  ttl 64
  up ip tunnel 6rd dev tun6rd 6rd-prefix 2001:db8::/32
  up ip link set mtu 1280 dev tun6rd
```

-reboot

A.1.2.4. CentOS

The latest version of CentOS is 5.5.CentOS5.5's kernel and iproute package does not supported 6rd.So it is required patch.

-download package

```
# wget http://enog.jp/~masakazu/6rd/kernel-2.6.18-238.9.1.el5.6rd.x86_64.rpm
# wget http://enog.jp/~masakazu/6rd/iproute-2.6.18-11.6rd.x86_64.rpm
```

-install package

```
# rpm -ivh kernel-2.6.18-238.9.1.el5.6rd.x86_64.rpm
# rpm -Uvh iproute-2.6.18-11.6rd.x86_64.rpm
```

-modify "/etc/yum.conf"

```
# vi /etc/yum.conf
exclude=kernel*,iproute
```

-modify "/etc/sysconfig/network-scripts/ifcfg-sit1"

```
# vi /etc/sysconfig/network-scripts/ifcfg-sit1
DEVICE=sit1
IPV6INIT=yes
IPV6_MTU=1280
IPV6_DEFAULTGW=::203.0.113.1
IPV6TUNNELIPV4=any
IPV6TUNNELIPV4LOCAL=203.0.113.13
IPV6ADDR=2001:db8:cb00:710d::1/32
```

modify "/etc/rc.local"

```
# vi /etc/rc.local
ip tunnel 6rd dev sit1 6rd-prefix 2001:db8::/32
```

-reboot

A.1.2.5. FreeBSD

FreeBSD does not support 6rd yet.But the patch exists.

-download patch

```
# cd /root
# fetch http://people.allbsd.org/~hrs/FreeBSD/stf_6rd_20100923-1.diff
```

-apply patch

```
# cd /usr/src
# patch -p0 < /root/stf_6rd_20100923-1.diff
```

-kernel module compile and install

```
# cd sys/modules/if_stf/
# make
# make install
```

-install manual

```
# cd /usr/src/share/man/
# make
# make install
```

-modify "/etc/rc.conf"

```
# vi /etc/rc.conf
ipv6_enable="YES"
cloned_interfaces="stf0"
ipv6_ifconfig_stf0="2001:db8:cb00:710e::1/32"
ipv6_defaultrouter="2001:db8:cb00:7101::1"
```

-reboot

A.2. OS Proportion on Sakura's VPS

The data of OS proportion on Sakura's VPS.

All of OSs could server-based 6rd.

Operating Systems	Proportion[%]
Ubuntu	31
Fedora	6
Debian	13
CentOS	39
FreeBSD	11

A.3. Apache Bench performance test result

A.3.1. Test Topology

Measurement method of http performance using [ApacheBench].

Two servers connected to each other and excute [ApacheBench] from Client

Figure 3

A.3.2. Test method

Two servers connected to each other.

Configure transport IPv4,IPv6 and 6rd.

Executes [ApacheBench] with 100 multiple and total 1000 requests from Client

A.3.3. Hardware Configuration

In this test cases,harware configuration is below.

CPU	Hyper-V	OS	NIC
Intel Xeon3060 2.4GHz	VMware ESX 4.1	Ubuntsu 10.10	E1000

A.3.4. Test Result

test item	IPv4	IPv6	6rd
Complete requests	1000	1000	1000
Failed requests	0	0	0
Requests per second[#/sec]	1.11	1.10	1.05
Time per request[ms]	900.598	912.586	951.868
Transfer rate[Kbytes/sec]	113702.5	112208.9	107578.2

Authors' Addresses

Shishio Tsuchiya (editor)
Cisco Systems
Shinjuku Mitsui Building, 2-1-1, Nishi-Shinjuku
Shinjuku-Ku, Tokyo 163-0409
Japan

Phone: +81 3 6434 6543
Email: shtsuchi@cisco.com

Mark Townsley
Cisco Systems
L'Atlantis, 11, Rue Camille Desmoulins ISSY LES MOULINEAUX
ILE DE FRANCE 92782
FRANCE

Phone: +33 15 804 3483
Email: mark@townsley.net

Shuichi Ohkubo
Sakura Internet
33F Sumitomo fudosan Nishi shinjuku Bldg.,7-20-1 Nishi shinjuku
Shinjuku-Ku, Tokyo 160-0023
Japan

Phone: +81 3 5332 7070
Email: ohkubo@sakura.ad.jp