                      4rd Applicability Statement
                 draft-sun-intarea-4rd-applicability-01

Abstract

   This document discusses the applicability of the IPv4 Residual
   Deployment (4rd) protocol, and an address sharing mechanism, NAT44 on
   4rd CE side, and routing optimization to provide IPv4 connectivity in
   4rd capable IPv6-only networks.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

The Internet continues to grow beyond the capabilities of IPv4. The
tremendous success of the Internet has strained the address space,
which is no longer sufficient to support future growth. The IANA
free pool of IPv4 address have been depleted already. With its
increase in the number of available prefixes and addresses in a
subnet, and improvement in address management, IPv6 is the only real
option on the table. During the long transition period from only
IPv4 to IPv6, the networks of Internet Service Providers (ISPs) will
have to offer more than just IPv6 connectivity. Both outgoing and
incoming connections will have to be supported.

IPv4 Residual Deployment (4rd) enables a service provider to share
IPv4 addresses among its customers by combining the well-known
technologies of stateless address mapping and tunneling, NAT44, and
IPv4 in IPv6 encapsulation.

This document discusses the applicability of the 4rd protocol, and an
address sharing mechanism, NAT44 on 4rd CE side, and routing
optimization to provide IPv4 connectivity in 4rd capable IPv6-only
networks.

2. Overview of 4rd

If 4rd [I-D.despres-intarea-4rd] is to be actually used across a
IPv6-only network to offer IPv4 connectivity
[I-D.arkko-ipv6-transition-guidelines], the network must have 4rd
Border Relay Router (BR) and 4rd Customer Edge Router (CE).

With 4rd, we need not worry that well-known (reserved) IPv4 addresses
(such as 000/8) are generated, because all 4rd CEs generating IPv4
addresses are in the range of domain 4rd prefix. (Section 3.1). The
4rd model is built on IPv4 over IPv6 stateless tunnels to reach 4rd
CEs where customers will share IPv4 addresses if the CE 4rd prefix is
longer than 32-bits (Section 3.2). 4rd differs from other solutions
such as Dual-Stack-lite [I-D.ietf-softwire-dual-stack-lite] since the
4rd CE is used to implement the NAT44 (Section 3.3). 4rd also can
implement Routing Optimization to Provide IPv4 Connectivity while the
4rd CE communicates with other 4rd CE (Section 3.4).

3. Applicability

In this section, we describe 4rd applicability with Figure 1 as an
example of 4rd enabled network. It is helpful to explain 4rd and
underdtand 4rd applicability.

```
                    +------------------------+
                   /                          \
                  |        IPv4 Internet       |
                   \                          /
                    +------------------------+
                               |
                               |
                    +----------------------+
                    |        4rd BR        |
                    +----------------------+
                               |
                               |
                 + --------------------------------+
                /                                   \
               |         ISP's IPv6 network         |
                \           2001:db8::/32           /
                 ---------------------------------
                /                |                \
               /                 |                 \
              /                  |                  \
     2001:db8:cdab::/48     2001:db8:cdef::/48     2001:db8:cda1::/48
       +-----------+          +------------+          +------------+
       | 4rd CE1   |          | 4rd CE2    |   ***    | 4rd CEn    |
       +-----------+          +------------+          + -----------+
```
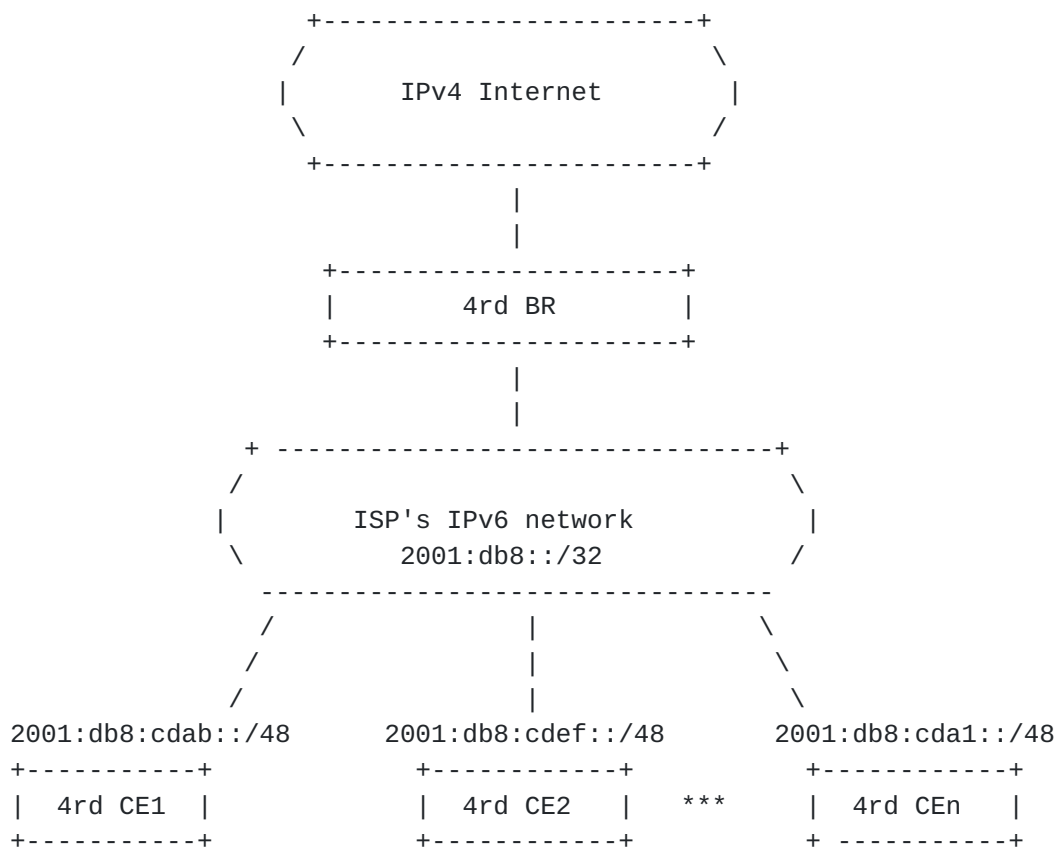
Figure 1: A 4rd network model example

3.1. IPv4 Address Assignment

[I-D.despres-intarea-4rd] specify 4rd address mapping rule. When 4rd
address mapping rule apply to the Figure 1, as Figure 2 shown, a 4rd
CE1's IPv4 address, which is in the range of domain 4rd prefix, is
generated from CE IPv6 prefix. As the prerequisite of 4rd, Domain
IPv6 prefix and Domain 4rd prefix must be given to all 4rd CEs and
4rd BRs in a specified 4rd domain. CE IPv6 prefix for all 4rd CEs in
the specified 4rd domain must be in same Domain IPv6 prefix but must
be unique for each 4rd CE.

```
     <------------ CE IPv6 prefix (max length 64) ------------>
     +------------------------------+------------------------+
     |      2001:db8::/32           |        0xcdab          |
     +------------------------------+------------------------+
     <-- Domain IPv6 prefix length --|<----CE index length---->|
                                 :           ||              :
                Domain 4rd prefix :           \/              :
                |<---------------->|<-- suffix-->|           |
                +------------------+------------------------+
   CE 4rd prefix | 192.0.2.0/24     |    0xcd    |  0xab     |
                +------------------+------------------------+
                |<----4rd CE1's IPv4 address----->|<--------->|
                           192.0.2.205(0xcd)       Port-set ID
```

              Figure 2: An example of 4rd address mapping for 4rd CE1

   All 4rd CEs IPv4 addresses are in the range fo domain of domain 4rd
   prefix. Figure 2 shows an example that 4rd CE1's IPv4 address how to
   be genarated. This means that it is not necessary for 4rd operators
   to pay attention to what IPv4 address is generated by 4rd CE. In the
   case of all 32 bits of IPv4 address is mapped to IPv6 prefix, 4rd CE
   must not generate well-known IPv4 address (such as 0/8, 127/8, etc.,
   for example) so that IPv6 prefix assignment should be taken careful
   to avoid it. 4rd is recommended to operators who do not want IPv6
   prefix assignment of which takes into account what IPv4 address
   generated from IPv6 prefix.

   As example (Figure 1) shown, the 4rd model is built on IPv4 over IPv6
   tunnels to reach 4rd CEs where up to 256 customers can share a IPv4
   addresse. In the given example, the IPv6 ISP aggregate Domain IPv6
   prefix is 2001:db8::/32. Out of this aggregate, a /48 CE IPv6 prefix
   for each 4rd CE is assumed to be issued, this allowing for a 16-bits
   CE index. Since remaining bits to mapping IPv4 address is 8 bit,
   "0xab" is a part of last octet of CE1's IPv4 address. As a result,
   192.0.2.205 is shared with 256 4rd-CEs.

3.2. IPv4 Address Sharing

   When the CE 4rd prefix is longer than 32 bits, IPv4 address will be
   shared by muliple 4rd CEs. As Figure 3 shown, the suffix of shared
   IPv4 address and Port-set ID are derived from CE index of CE IPv6
   prefix. CEs have same 4rd shared IPv4 address and unique Port-set
   ID, so CEs can be identified using IPv4 address and unique Port-set
   ID.

```
         <----------- CE IPv6 prefix length (max 64) ------------->
         +------------------------------+------------------------+
         |      2001:db8::/32           |        CE index        |
         +------------------------------+------------------------+
         <-- Domain IPv6 Prefix length --|<----CE index length---->|
                                    :            ||           :
                   Domain 4rd prefix :            \/           :
                   |<---------------->|<-- suffix-->|          |
                   +------------------+------------------------+
    CE 4rd prefix  | 192.0.2.0/24     |    0xcd     |Port-set ID|
                   +------------------+------------------------+
                   |<---4rd shared IPv4 address----->| CE1: 0xab
                          192.0.2.205(0xcd)            CE2: 0xef
                                                       ...  ...
                                                       CEn: 0xa1
```
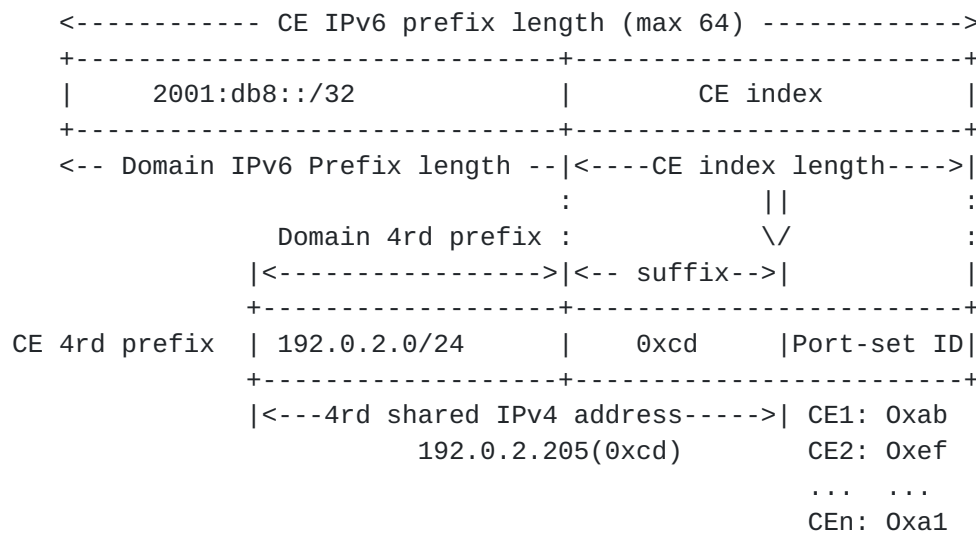
                   Figure 3: Port-set ID generation example

As shown in Figure 3 , the available port-set ID for IPv4 address
sharing is derived form the remaining 8-bits, e.g. 0xab, 0xef, 0xa1
used by 4rd CE1, 4rd CE2, 4rd CEn respectively. With this particular
scheme shown in the example, up to 256 4rd CEs can share the IPv4
address.

According to [I-D.despres-intarea-4rd],each 4rd CE has its unique
port-ranges which is calculated based on the 4rd mapping rule with
their Port-set ID. The 4rd port mapping rule uses four heads, as
shown in the Figure 4 below. This allows 4rd CEs to avoid using well
known (reserved) TCP/UDP ports 0-4095.

```
                 <------- Port (16bit) ---------->
      Port-range a +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      (head=0001)  |0|0|0|1|1 0 1 0|1 0 1 1|      |     0x1AB0 - 0x1ABF
                   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                        /  0xA  /   0xB /
      Port-range b +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      (head=001)   |0|0|1|1 0 1 0|1 0 1 1|        |     0x3560 - 0x356F
                   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                       /  0xA  /  0xB  /
      Port-range c +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      (head=01)    |0|1|1 0 1 0|1 0 1 1|          |     0x6AC0 - 0x6AFF
                   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                      /  0xA  /  0xB  /
      Port-range d +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      (head=1)     |1|1 0 1 0|1 0 1 1|            |     0xD570 - 0xD5FF
                   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                    /  0xA  /  0xB  /
      Port-set ID  +-+-+-+-+-+-+-+-+
      (CE1,0xAB)   |1 0 1 0|1 0 1 1|
                   +-+-+-+-+-+-+-+-+
```

              Figure 4: Port-range calculation example for 4rd CE1

3.3. Distributing NAT function to CEs

   Since 4rd distributes NAT44 function to the 4rd CE, the CE side NAT44
   manages user's NATed session in general of each 4rd CE unit.
   According to their own infrastructure and management systems, if ISPs
   think 4rd CE Side NAT is easier to manage than gateway side NAT, they
   can adopt the 4rd technology.

   The capex and opex costs of an ISP vary from network to network. The
   4rd solution is an ideal fit for those where a CE based NAT solution
   is acceptable, and where investment in operating and managing of a
   centralized NAPT 44 gateway is not desired.

   For an CE IPv6 prefix of a given length, e.g. a /56 PD, as more and
   more bits are used to express the shared IPv4 address, the less bits
   are available to express the port range. It is not 4rd specific
   issue, but a method that supports NAT Port Mapping needs to be
   devised. For example, 256 ports can be used by 4rd CE, but access to
   host A, B, and C requires 100 ports each. If the 4rd CE ports are
   distributed when A,B,C are accessed at the same time from the same
   4rd CE, Ports 0 - 99 are used to access host A, Ports 100 - 199 are
   used to access host B, but host C cannot be accessed because the
   remaining ports are not enough. As Figure 5, 4rd can reuse the 4rd
   CE Port by terms of recognizing the destination addresses in order to
   reduce the aspect of this problem. It is noted that the problem can

not be solved if the multiple users under same 4rd CE access the same
host (e.g. host A) exceed the number of ports.

```
                  ----------------------------------
               /                                     \
              /                 IPv4 Internet          \
           |    +--------+   +--------+   +--------+    |
            \   | host A |   | host B |   | host C |   /
             \  +--------+   +--------+   +--------+  /
               -------------------------------------
                               |
                 +--------------------+
                 |       4rd BR       |
                 +--------------------+
                               |
             +---------------------------------+
            /                                   \
           |          ISP's IPv6 network         |
            \                                   /
             +---------------------------------+
                               |
           4rd CE              |
           +--------------------------------------------------+
           |    Port number      Port number      Port number |
           |     pool for         pool for         pool for   |
           |   +--host A --+   +--host B --+   +--host C --+   |
           |   | Port0-255 |   | Port0-255 |   | Port0-255 |   |
           |   +-----------+   +-----------+   +-----------+   |
           +--------------------------------------------------+
```
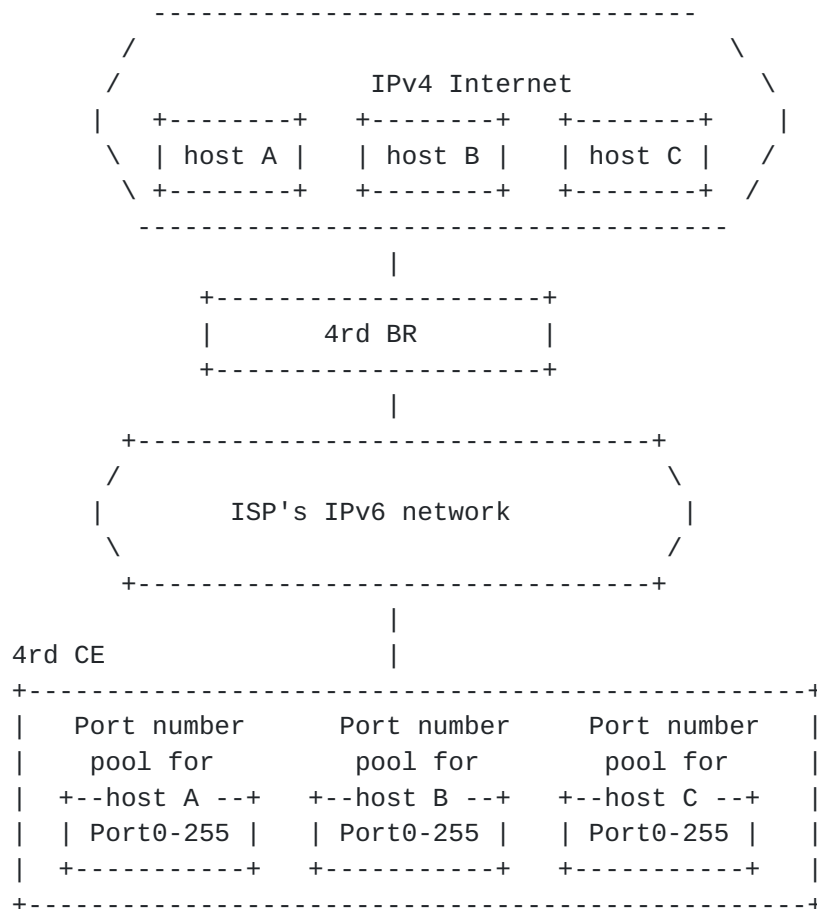
            Figure 5: Reuse port number on 4rd CE side NAT

   Figure 6 quantifies address sharing rates for different port range
   lengths, along with the overall number of ports available to a give
   CE. The 4rd allows 16 bits to be used by NAPT. In theory, one IPv4
   address can be shared by 65536 (2 to 16th power) customers. However,
   at least 1 bit is needed for Port-range Index. The greatest bit
   range that can be used to share address are 15 bits.

```
+------------+------------------------+--------------+------------+
|    Port    | User numbers that can  |  # of ports  |  Relative  |
|   prefix   |  share one v4 address   |  for a 4rd CE |   to a /8  |
|   length   |                        |              |            |
+------------+------------------------+--------------+------------+
|     1      |                1 - 2 |        30720 |     /9     |
|     2      |                1 - 4 |        15360 |     /10    |
|     3      |                1 - 8 |         7680 |     /11    |
|     4      |               1 - 16 |         3840 |     /12    |
|     5      |               1 - 32 |         1920 |     /13    |
|     6      |               1 - 64 |          960 |     /14    |
|     7      |              1 - 128 |          480 |     /15    |
|     8      |              1 - 256 |          240 |     /16    |
|     9      |              1 - 512 |          120 |     /17    |
|    10      |             1 - 1024 |           60 |     /18    |
|    11      |             1 - 2048 |           30 |     /19    |
|    12      |             1 - 4096 |           15 |     /20    |
|    13      |             1 - 8192 |            7 |     /21    |
|    14      |            1 - 16384 |            3 |     /22    |
|    15      |            1 - 32768 |            1 |     /23    |
+------------+------------------------+--------------+------------+
```

Figure 6: 4rd CE available port

3.4. Routing optimization to provide IPv4 connectivity

   Figure 7 shows that 4rd can implement IPv6 routing with direct IP
   forwarding between the two 4rd CEs. In other words, 4rd allows
   direct packet forwarding between CEs that share a 4rd domain, without
   the need of forwarding such traffic through the gateway.

   When V4a under 4rd CE1 send packets to V4b under 4rd CE2, the packets
   need not go through the gateway (Hub&Spoke) before going to 4rd CE2,
   they can go to 4rd CE2 directly (mesh). This section describes this
   in more detail from two perspectives: Network matching and Network
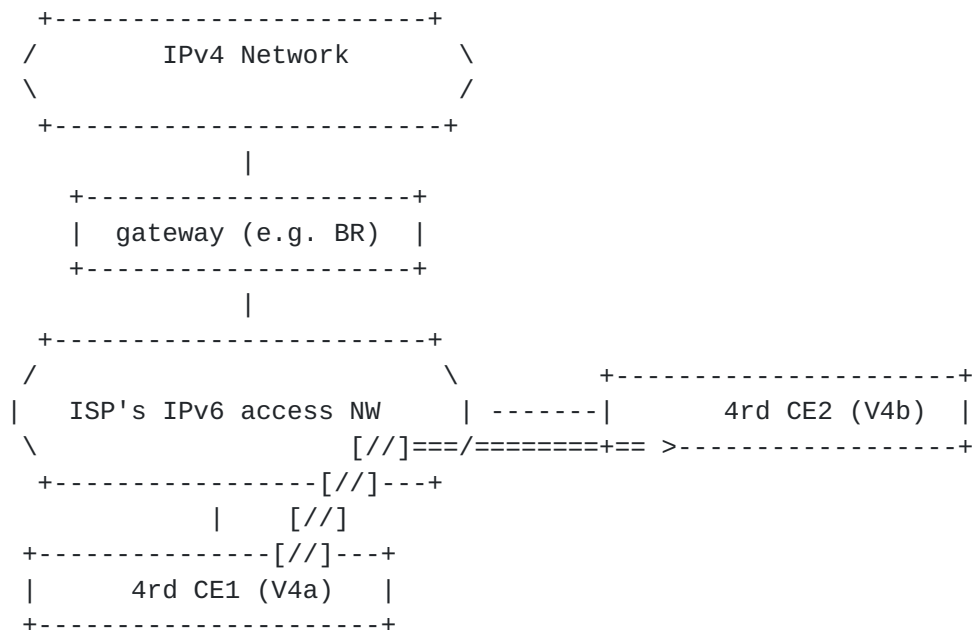   Load.

```
         +-----------------------+
        /         IPv4 Network      \
        \                           /
         +------------------------+
                    |
           +--------------------+
           |  gateway (e.g. BR)  |
           +--------------------+
                    |
         +-----------------------+
        /                        \      +----------------------+
       |   ISP's IPv6 access NW   | -------|       4rd CE2 (V4b)  |
        \                   [//]===/=======+== >------------------+
         +----------------[//]---+
                 |    [//]
        +--------------[//]---+
        |      4rd CE1 (V4a)    |
        +---------------------+
```

                             Figure 7

   Taking network matching into account, 4rd should be chosen if the
   ISP's IPv6 network allows direct CE-CE traffic forwarding (eg: IPv6
   Flat Routing Network).

   Compared to Hub & Spoke architecture solutions, mesh architecture
   solutions can achieve a smaller network load when the 4rd CE
   communicate frequently. If the ISP thinks that its network has
   insufficient bandwidth or it wants to lower the load imposed by IPv4
   connections, they should choose a mesh architecture solution such as
   4rd.

   Note: Softwire mesh [RFC5565] is also a mesh architecture solution,
   but it does not provide the IPv4 sharing function. Softwire mesh can
   be adopted if IPv4 address sharing is not needed.

3.5. Gateway Redundancy Considerations

   Since in 4rd the BR side does not implement the NAT function, the BR
   can be easily replicated. For instance, there can be N:1 redundant
   BRs in the network. Reliable network operation can be guaranteed
   because a failed BR can be immediately and easily replicated by any
   one of the N BRs.

   If the gateway side implements NAT, 1:1 gateway redundancy is the
   choice, and also upstream and downstream traffic from the same user
   must go through the same gateway. 4rd easily realizes load balancing
   because upstream and downstream traffic from the same user can go

through different gateways.

3.6. NAT Log Considerations

4rd CE uses fixed NAT rules and IPv4 users can be directly identified
by means of their IPv6 address, so the NAT log does not need to be
left. In other words, 4rd allows an operator to save on the need to
perform NAT log collection and retention. On the other hand, other
solutions may achieve higher address sharing rate than 4rd. If
operators want to reduce NAT log with these solutions, their NATs
have to be configured with [I-D.tsou-behave-natx4-log-reduction], or
allocate fixed port range for NAT rules regardless of the advantage.

It is noted that another logging necessity coming from
[I-D.ietf-intarea-shared-addressing-issues] is described in
[I-D.ietf-intarea-server-logging-recommendations] for Internet facing
servers.

4. Constraint Considerations

This section describes the constraint on user number when addresses
are shared. As Section 3.3 stated, the maximum bit range that can be
used by Port-set ID is 15 bits. In this case, 4rd CE is restricted
to one private port. In fact, it is not able to provide service. So
the customer numbers per IPv4 address that can be shared need to be
considered along with the ports needed by 4rd CE.

It is noted that deriving IPv4 addresses from CE IPv6 prefix
generated by 4rd mapping rule may lead to a low rate of IPv4 address
sharing because it does not allow for statistical multipexing gains.
ISPs who feel that this is not desirable and want to achieve high
sharing rates can select centralized gateway/NAPT solutions, which
can have a high sharing rate because allow for such statistical
multipexing gains.

5. Security Considerations

The sharing of IPv4 address reduces the port selection space per 4rd
CE, so a blind attack can be performed easily. An attack can be
performed if the attacker is able to correctly guess the source
address and source port. Address and Port Dependent Filtering (APDF)
need be implemented in order to counter blind attack. In APDF, not
only source address and source port but also destination address and
destination port need to be checked. Even so, a blind attack that
can be performed against TCP relies on the attacker's ability to
guess the 5-ruple (Protocol, Source address, Destination address,

Source Port, Destination Port). Shared address issues
[I-D.ietf-intarea-shared-addressing-issues] describes a method for
the random selection of TCP Sequence Number, that reduces the ability
of attacker to correctly guess the 5-ruple.

DNS is one of the important protocols to use UDP. In 4rd CE, all DNS
reply packets should be discarded, expect for the packets from
forward destinations. If DNS implements Port Randomization, the
attack success rate can be reduced.

We describe here a method for reducing Spoofing Attack under 4rd CE.
Using the 4rd mapping rule, the IPv4 address can be derived from IPv6
address, so we can check the IPv4 address thus derived and the IPv4
address in the header of received packet. If they are same, the
packet is forwarded, otherwise, it is dropped.

6. Conclusions

This document described the ability of 4rd to support the IPv6-only
access network. We conclude that the 4rd solution is a viable choice
when the ISP desires a stateless IPv4 address sharing option, based
on CE side NAT functionality, along with a high degree of freedom in
terms of redundancy and optimal traffic forwarding. 4rd is also
recommended to operators who do not want IPv6 prefix assignment of
which takes into account what IPv4 address generated from IPv6
prefix. When only one 4rd applicable character is needed, 4rd may be
used to only that purpose with other solutions.

7. Acknowledgements

The authors would like to thank Remi Despres for his enormous effort
to work for 4rd architecture and protocol. The authors also thank to
people who provided encouragements concerning the 4rd approach and
lead to consider 4rd in Japan. In particular, we would like to thank
Toshiya Asaba, Yoshiki Ishida, Tetsuya Innami, Ruri Hiromi, Masataka
Mawatari, Akira Nakagawa, Tomoyuki Sahara, Taisuke Sasaki, Katsuyasu
Toyama, Mark Townsley, Wojciech Dec, Yuji Yamazaki, have provided
useful discussion and comments on the document and review.

8. References

8.1. Normative References

[I-D.despres-intarea-4rd]
          Despres, R., Matsushima, S., Murakami, T., and O. Troan,

              "IPv4 Residual Deployment across IPv6-Service networks
              (4rd) ISP-NAT's made optional",
              draft-despres-intarea-4rd-00 (work in progress),
              March 2011.

   [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6
              (IPv6) Specification", RFC 2460, December 1998.

   [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing
              Architecture", RFC 4291, February 2006.

8.2. Informative References

   [I-D.arkko-ipv6-transition-guidelines]
              Arkko, J. and F. Baker, "Guidelines for Using IPv6
              Transition Mechanisms during IPv6 Deployment",
              draft-arkko-ipv6-transition-guidelines-14 (work in
              progress), December 2010.

   [I-D.ietf-intarea-server-logging-recommendations]
              Durand, A., Gashinsky, I., Lee, D., and S. Sheppard,
              "Logging recommendations for Internet facing servers",
              draft-ietf-intarea-server-logging-recommendations-03 (work
              in progress), February 2011.

   [I-D.ietf-intarea-shared-addressing-issues]
              Ford, M., Boucadair, M., Durand, A., Levis, P., and P.
              Roberts, "Issues with IP Address Sharing",
              draft-ietf-intarea-shared-addressing-issues-05 (work in
              progress), March 2011.

   [I-D.ietf-softwire-dual-stack-lite]
              Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-
              Stack Lite Broadband Deployments Following IPv4
              Exhaustion", draft-ietf-softwire-dual-stack-lite-07 (work
              in progress), March 2011.

   [I-D.tsou-behave-natx4-log-reduction]
              ZOU), T., Li, W., and T. Taylor, "Port Management To
              Reduce Logging In Large-Scale NATs",
              draft-tsou-behave-natx4-log-reduction-02 (work in
              progress), September 2010.

   [RFC2516] Mamakos, L., Lidl, K., Evarts, J., Carrel, D., Simone, D.,
              and R. Wheeler, "A Method for Transmitting PPP Over
              Ethernet (PPPoE)", RFC 2516, February 1999.

   [RFC3849] Huston, G., Lord, A., and P. Smith, "IPv6 Address Prefix

                    Reserved for Documentation", RFC 3849, July 2004.

   [RFC5565] Wu, J., Cui, Y., Metz, C., and E. Rosen, "Softwire Mesh
                    Framework", RFC 5565, June 2009.

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

Authors' Addresses

   Chunfa Sun
   Softbank BB
   Tokyo Shiodome Building. 22F
   1-9-1,Higashi-Shimbashi,Minato-Ku
   Tokyo 105-7322
   JAPAN

   Email: c-sun@bb.softbank.co.jp

   Satoru Matsushima
   Softbank Telecom
   Tokyo Shiodome Building. 22F
   1-9-1,Higashi-Shimbashi,Minato-Ku
   Tokyo 105-7322
   JAPAN

   Email: satoru.matsushima@tm.softbank.co.jp

   Jie Jiao
   Softbank BB
   Tokyo Shiodome Building. 12F
   1-9-1,Higashi-Shimbashi,Minato-Ku
   Tokyo 105-7322
   JAPAN

   Email: j-jiao@bb.softbank.co.jp