### Connection of IPv6 Domains via IPv4 Clouds

Status of this Memo

Copyright Notice

Abstract

   This memo specifies an optional interim mechanism for IPv6 sites to
   communicate with each other over the IPv4 network without explicit
   tunnel setup, and for them to communicate with native IPv6 domains
   via relay routers.  Effectively it treats the wide area IPv4 network
   as a unicast point-to-point link layer.  The mechanism is intended as
   a start-up transition tool used during the period of co-existence of
   IPv4 and IPv6.  It is not intended as a permanent solution.

   The document defines a method for assigning an interim unique IPv6
   address prefix to any site that currently has at least one globally
   unique IPv4 address, and specifies an encapsulation mechanism for
   transmitting IPv6 packets using such a prefix over the global IPv4
   network.

   The motivation for this method is to allow isolated IPv6 domains or
   hosts, attached to an IPv4 network which has no native IPv6 support,
   to communicate with other such IPv6 domains or hosts with minimal
   manual configuration, before they can obtain natuve IPv6
   connectivity.  It incidentally provides an interim globally unique
   IPv6 address prefix to any site with at least one globally unique
   IPv4 address, even if combined with an IPv4 Network Address
   Translator (NAT).

Table of Contents

## 1. Introduction

   This memo specifies an optional interim mechanism for IPv6 sites to
   communicate with each other over the IPv4 network without explicit
   tunnel setup, and for them to communicate with native IPv6 domains
   via relay routers.  Effectively it treats the wide area IPv4 network
   as a unicast point-to-point link layer.  The mechanism is intended as
   a start-up transition tool used during the period of co-existence of
   IPv4 and IPv6.  It is not intended as a permanent solution.

The document defines a method for assigning an interim unique IPv6
address prefix to any site that currently has at least one globally
unique IPv4 address, and specifies an encapsulation mechanism for
transmitting IPv6 packets using such a prefix over the global IPv4
network.  It also describes scenarios for using such prefixes during
the co-existence phase of IPv4 to IPv6 transition.  Note that these
scenarios are only part of the total picture of transition to IPv6.
Also note that this is considered to be an interim solution and that
sites should migrate when possible to native IPv6 prefixes and native
IPv6 connectivity.  This will be possible as soon as the site's ISP
offers native IPv6 connectivity.

The basic mechanism described in the present document, which applies
to sites rather than individual hosts, will scale indefinitely by
limiting the number of sites served by a given relay router (see
Section 5.2).  It will introduce no new entries in the IPv4 routing
table, and exactly one new entry in the native IPv6 routing table
(see Section 5.10).

Although the mechanism is specified for an IPv6 site, it can equally
be applied to an individual IPv6 host or very small site, as long as
it has at least one globally unique IPv4 address.  However, the
latter case raises serious scaling issues which are the subject of
further study [SCALE].

The motivation for this method is to allow isolated IPv6 sites or
hosts, attached to a wide area network which has no native IPv6
support, to communicate with other such IPv6 domains or hosts with
minimal manual configuration.

IPv6 sites or hosts connected using this method do not require IPv4-
compatible IPv6 addresses [MECH] or configured tunnels.  In this way
IPv6 gains considerable independence of the underlying wide area
network and can step over many hops of IPv4 subnets.  The abbreviated
name of this mechanism is 6to4 (not to be confused with [6OVER4]).
The 6to4 mechanism is typically implemented almost entirely in border
routers, without specific host modifications except a suggested
address selection default.  Only a modest amount of router
configuration is required.

Sections 2 to 4 of this document specify the 6to4 scheme technically.
Section 5 discusses some, but not all, usage scenarios, including
routing aspects, for 6to4 sites.  Scenarios for isolated 6to4 hosts
are not discussed in this document.  Sections 6 to 9 discuss other
general considerations.

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in [RFC2119].

## 1.1. Terminology

   The terminology of [IPV6] applies to this document.

   6to4 pseudo-interface:
        6to4 encapsulation of IPv6 packets inside IPv4 packets occurs
        at a point that is logically equivalent to an IPv6 interface,
        with the link layer being the IPv4 unicast network.  This point
        is referred to as a pseudo-interface.  Some implementors may
        treat it exactly like any other interface and others may treat
        it like a tunnel end-point.

   6to4 prefix:
        an IPv6 prefix constructed according to the rule in Section 2
        below.

   6to4 address:  an IPv6 address constructed using a 6to4 prefix.

   Native IPv6 address:  an IPv6 address constructed using another type
        of prefix than 6to4.

   6to4 router (or 6to4 border router):
        an IPv6 router supporting a 6to4 pseudo-interface.  It is
        normally the border router between an IPv6 site and a wide-area
        IPv4 network.

   6to4 host:
        an IPv6 host which happens to have at least one 6to4 address.
        In all other respects it is a standard IPv6 host.

   Note: an IPv6 node may in some cases use a 6to4 address for a
   configured tunnel.  Such a node may function as an IPv6 host using a
   6to4 address on its configured tunnel interface, and it may also
   serve as a IPv6 router for other hosts via a 6to4 pseudo-interface,
   but these are distinct functions.

   6to4 site:
        a site running IPv6 internally using 6to4 addresses, therefore
        containing at least one 6to4 host and at least one 6to4 router.

    Relay router:
          a 6to4 router configured to support transit routing between
          6to4 addresses and native IPv6 addresses.

    6to4 exterior routing domain:
          a routing domain interconnecting a set of 6to4 routers and
          relay routers.  It is distinct from an IPv6 site's interior
          routing domain, and distinct from all native IPv6 exterior
          routing domains.

## 2. IPv6 Prefix Allocation

    Suppose that a subscriber site has at least one valid, globally
    unique 32-bit IPv4 address, referred to in this document as V4ADDR.
    This address MUST be duly allocated to the site by an address
    registry (possibly via a service provider) and it MUST NOT be a
    private address [RFC 1918].

    The IANA has permanently assigned one 13-bit IPv6 Top Level
    Aggregator (TLA) identifier under the IPv6 Format Prefix 001 [AARCH,
    AGGR] for the 6to4 scheme.Its numeric value is 0x0002, i.e., it is
    2002::/16 when expressed as an IPv6 address prefix.

    The subscriber site is then deemed to have the following IPv6 address
    prefix, without any further assignment procedures being necessary:

       Prefix length: 48 bits
       Format prefix: 001
       TLA value: 0x0002
       NLA value: V4ADDR

    This is illustrated as follows:

      | 3 |  13  |    32     |   16   |             64 bits               |
      +---+------+-----------+--------+----------------------------------+
      |FP | TLA  | V4ADDR    | SLA ID |          Interface ID            |
      |001|0x0002|           |        |                                  |
      +---+------+-----------+--------+----------------------------------+

    Thus, this prefix has exactly the same format as normal /48 prefixes
    assigned according to [AGGR].  It can be abbreviated as
    2002:V4ADDR::/48.  Within the subscriber site it can be used exactly
    like any other valid IPv6 prefix, e.g., for automated address
    assignment and discovery according to the normal mechanisms such as
    [CONF, DISC], for native IPv6 routing, or for the "6over4" mechanism
    [6OVER4].

   Note that if the IPv4 address is assigned dynamically, the
   corresponding IPv6 prefix will also be dynamic in nature, with the
   same lifetime.

## 2.1 Address Selection

   To ensure the correct operation of 6to4 in complex topologies, source
   and destination address selection must be appropriately implemented.
   If the source IPv6 host sending a packet has at least one 2002::
   address assigned to it, and if the set of IPv6 addresses returned by
   the DNS for the destination host contains at least one 2002::
   address, then the source host must make an appropriate choice of the
   source and destination addresses to be used.  The mechanisms for
   address selection in general are under study at the time of writing
   [SELECT].  Subject to those general mechanisms, the principle that
   will normally allow correct operation of 6to4 is this:

   If one host has only a 6to4 address, and the other one has both a
   6to4 and a native IPv6 address, then the 6to4 address should be used
   for both.

   If both hosts have a 6to4 address and a native IPv6 address, then
   either the 6to4 address should be used for both, or the native IPv6
   address should be used for both.  The choice should be configurable.
   The default configuration should be native IPv6 for both.

## 3. Encapsulation in IPv4

   IPv6 packets from a 6to4 site are encapsulated in IPv4 packets when
   they leave the site via its external IPv4 connection.  Note that the
   IPv4 interface that is carrying the 6to4 traffic is notionally
   equivalent to an IPv6 interface, and is referred to below as a
   pseudo-interface, although this phrase is not intended to define an
   implementation technique.  V4ADDR MUST be configured on the IPv4
   interface.

   IPv6 packets are transmitted in IPv4 packets [RFC 791] with an IPv4
   protocol type of 41, the same as has been assigned [MECH] for IPv6
   packets that are tunneled inside of IPv4 frames.  The IPv4 header
   contains the Destination and Source IPv4 addresses.  One or both of
   these will be identical to the V4ADDR field of an IPv6 prefix formed
   as specified above (see section 5 for more details).  The IPv4 packet
   body contains the IPv6 header and payload.

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |Version|  IHL  |Type of Service|          Total Length         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |         Identification        |Flags|      Fragment Offset    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |  Time to Live | Protocol 41   |        Header Checksum         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                       Source Address                          |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                     Destination Address                       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                       Options                  |    Padding    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |            IPv6 header and payload ...                /
 +-------+-------+-------+-------+-------+------+------+
```

   The IPv4 Time to Live will be set as normal [RFC 791], as will the
   encapsulated IPv6 hop limit [IPv6].  Other considerations are as
   described in Section 4.1.2 of [MECH].

## 3.1. Link-Local Address and NUD

   The link-local address of a 6to4 pseudo-interface performing 6to4
   encapsulation would, if needed, be formed as described in Section 3.7
   of [MECH].  However, no scenario is known in which such an address
   would be useful, since a peer 6to4 gateway cannot determine the
   appropriate link-layer (IPv4) address to send to.

   Neighbor Unreachability Detection (NUD) is handled as described in
   Section 3.8 of [MECH].

## 4. Maximum Transmission Unit

   MTU size considerations are as described for tunnels in [MECH].

   If the IPv6 MTU size proves to be too large for some intermediate
   IPv4 subnet, IPv4 fragmentation will ensue.  While undesirable, this
   is not necessarily disastrous, unless the fragments are delivered to
   different IPv4 destinations due to some form of IPv4 anycast.  The
   IPv4 "do not fragment" bit SHOULD NOT be set in the encapsulating
   IPv4 header.

## 5. Unicast scenarios, scaling, and transition to normal prefixes

### 5.1 Simple scenario - all sites work the same

   The simplest deployment scenario for 6to4 is to use it between a
   number of sites, each of which has at least one connection to a
   shared IPv4 Internet.  This could be the global Internet, or it could
   be a corporate IP network.  In the case of the global Internet, there
   is no requirement that the sites all connect to the same Internet
   service provider.  The only requirement is that any of the sites is
   able to send IPv4 packets with protocol type 41 to any of the others.
   By definition, each site has an IPv6 prefix in the format defined in
   Section 2.  It will therefore create DNS records for these addresses.
   For example, site A which owns IPv4 address 192.1.2.3 will create DNS
   records with the IPv6 prefix {FP=001,TLA=0x0002,NLA=192.1.2.3}/48
   (i.e., 2002:c001:0203::/48).  Site B which owns address 9.254.253.252
   will create DNS records with the IPv6 prefix
   {FP=001,TLA=0x0002,NLA=9.254.253.252}/48 (i.e., 2002:09fe:fdfc::/48).

   When an IPv6 host on site B queries the DNS entry for a host on site
   A, or otherwise obtains its address, it obtains an address with the
   prefix {FP=001,TLA=0x0002,NLA=192.1.2.3}/48 and whatever SLA and
   Interface ID applies.  The converse applies when a host on site A
   queries the DNS for a host on site B.  IPv6 packets are formed and
   transmitted in the normal way within both sites.

```
                    _____
                   |                              |
                   |   Wide Area IPv4 Network      |
                   |_____|
                         /                    \
               192.1.2.3/          9.254.253.252\
    _____/_  _____
   |                        /  | |                   \          |
   |IPv4 Site A       ##########  | |IPv4 Site B       ########## |
   | _____# 6to4   #_ | | _____# 6to4   #_ |
   ||                  # router # || ||                  # router # ||
   ||IPv6 Site A       ########## || ||IPv6 Site B       ########## ||
   ||2002:c001:0203::/48          || ||2002:09fe:fdfc::/48         ||
   ||_____|| ||_____||
   |                            | |                            |
   |_____| |_____|
```
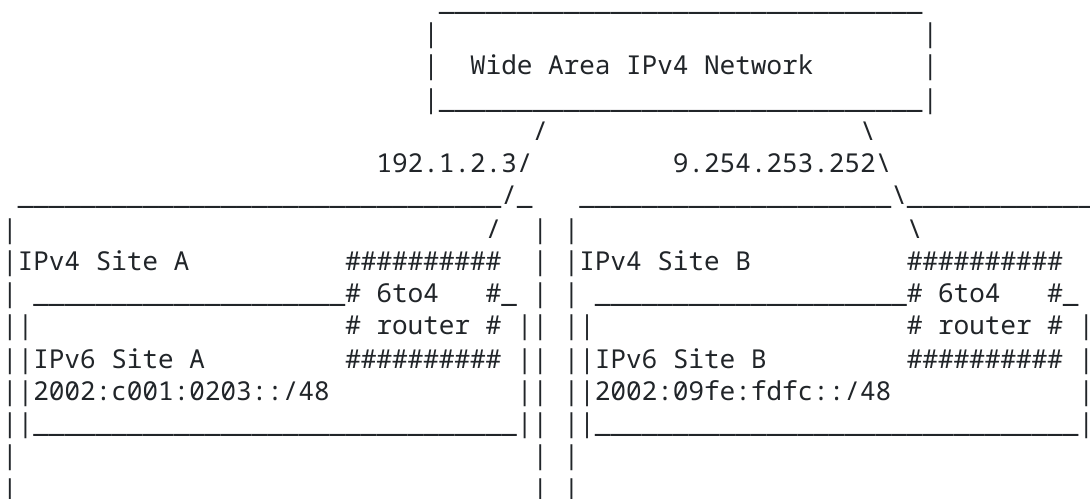
   Within a 6to4 site, addresses with the 2002::/16 prefix, apart from
   those with the local 2002:V4ADDR::/48 prefix, will be handled like
   any other non-local IPv6 address, i.e., by a default or explicit
   route towards the 6to4 border router.

When an outgoing packet reaches the 6to4 router, it is encapsulated
as defined in Section 3, according to the additional sending rule
defined in Section 5.3.  Incoming packets are decapsulated according
to the additional decapsulation rule defined in Section 5.3.  The
additional sending and decapsulation rules are the only changes to
IPv6 forwarding, and they occur only at border routers.  No IPv4
routing information is imported into IPv6 routing (nor vice versa).

In this scenario, any number of 6to4 sites can interoperate with no
tunnel configuration, and no special requirements from the IPv4
service.  All that is required is the appropriate DNS entries and the
additional sending and decapsulation rules configured in the 6to4
router.  This router SHOULD also generate the appropriate IPv6 prefix
announcements [CONF, DISC].

Although site A and site B will each need to run IPv6 routing
internally, they do not need to run an IPv6 exterior routing protocol
in this simple scenario; IPv4 exterior routing does the job for them.

It is RECOMMENDED that in any case each site should use only one IPv4
address per 6to4 router, and that should be the address assigned to
the external interface of the 6to4 router.  Single-homed sites
therefore SHOULD use only one IPv4 address for 6to4 routing.  Multi-
homed sites are discussed briefly in section 5.6.

Because of the lack of configuration, and the distributed deployment
model, there are believed to be no particular scaling issues with the
basic 6to4 mechanism apart from encapsulation overhead.
Specifically, it introduces no new entries in IPv4 routing tables.

## 5.2 Mixed scenario with relay to native IPv6

During the transition to IPv6 we can expect some sites to fit the
model just described (isolated sites whose only connectivity is the
IPv4 Internet), whereas others will be part of larger islands of
native or tunneled IPv6 using normal IPv6 TLA address space.  The
6to4 sites will need connectivity to these native IPv6 islands and
vice versa.  In the 6to4 model, this connectivity is accomplished by
IPv6 routers which possess both 6to4 and native IPv6 addresses.
Although they behave essentially as standard IPv6 routers, for the
purposes of this document they are referred to as relay routers to
distinguish them from routers supporting only 6to4, or only native
IPv6.

There must be at least one router acting as a relay between the 6to4
domain and a given native IPv6 domain.  There is nothing special
about it; it is simply a normal router which happens to have at least

one logical 6to4 pseudo-interface and at least one other IPv6
interface.  Since it is a 6to4 router, it implements the additional
sending and decapsulation rules defined in Section 5.3.

We now have three distinct classes of routing domain to consider:

1.  the internal IPv6 routing domain of each 6to4 site;
2.  an exterior IPv6 routing domain interconnecting
    a given set of 6to4 border routers, including relay routers,
    among themselves, i.e., a 6to4 exterior routing domain;
3.  the exterior IPv6 routing domain of each native IPv6 island.

1. The internal routing domain of a 6to4 site behaves as described in
section 5.1.

2. There are two deployment options for a 6to4 exterior routing
domain:

2.1 No IPv6 exterior routing protocol is used.  The 6to4 routers
using a given relay router each have a default IPv6 route pointing to
the relay router.  The relay router MAY apply source address based
filters to accept traffic only from  specific 6to4 routers.

2.2 An IPv6 exterior routing protocol is used.  The set of 6to4
routers using a given relay router obtain native IPv6 routes from the
relay router using a routing protocol such as BGP4+ [RFC 2283,
BGP4+].  The relay router will advertise whatever native IPv6 routing
prefixes are appropriate on its 6to4 pseudo-interface.  These
prefixes will indicate the regions of native IPv6 topology that the
relay router is willing to relay to.  Their choice is a matter of
routing policy.  It is necessary for network operators to carefully
consider desirable traffic patterns and topology when choosing the
scope of such routing advertisements.  The relay router will
establish BGP peering only with specific 6to4 routers whose traffic
it is willing to accept.

Although this solution is more complex, it provides effective policy
control, i.e., BGP4+ policy determines which 6to4 routers are able to
use which relay router.

3. A relay router MUST advertise a route to 2002::/16 into the native
IPv6 exterior routing domain.  It is a matter of routing policy how
far this routing advertisement of 2002::/16 is propagated in the
native IPv6 routing system.  Since there will in general be multiple
relay routers advertising it, network operators will require to
filter it in a managed way.  Incorrect policy in this area will lead
to potential unreachability or to perverse traffic patterns.

6to4 prefixes more specific than 2002::/16 must not be propagated in
native IPv6 routing, to prevent pollution of the IPv6 routing table
by elements of the IPv4 routing table.  Therefore, a 6to4 site which
also has a native IPv6 connection MUST NOT advertise its 2002::/48
routing prefix on that connection, and all native IPv6 network
operators MUST filter out and discard any 2002:: routing prefix
advertisements longer than /16.

Sites which have at least one native IPv6 connection, in addition to
a 6to4 connection, will therefore have at least one IPv6 prefix which
is not a 2002:: prefix.  Such sites' DNS entries will reflect this
and DNS lookups will return multiple addresses.  If two such sites
need to interoperate, whether the 6to4 route or the native route will
be used depends on IPv6 address selection by the individual hosts (or
even applications).

Now consider again the example of the previous section.  Suppose an
IPv6 host on site B queries the DNS entry for a host on site A, and
the DNS returns multiple IPv6 addresses with different prefixes.

```
      _____         _____
     |                           |       |                     |
     |   Wide Area IPv4 Network   |       |    Native IPv6      |
     |                           |       |   Wide Area Network  |
     |_____|       |_____|
          /                \               //
    192.1.2.3/        9.254.253.252\        // 2001:0600::/48
 _____/_  _____//_
          /   | |                   \       //   |
    ##########  | |IPv4 Site B         ##########  |
 __# 6to4   #_ | | _____# 6to4   #_ |
   # router # || ||                    # router # ||
   ########## || ||IPv6 Site B         ########## ||
           || ||2002:09fe:fdfc::/48            ||
 __Site A_____|| ||2001:0600::/48_____||
   as before   | |                              |
 _____| |_____|
```
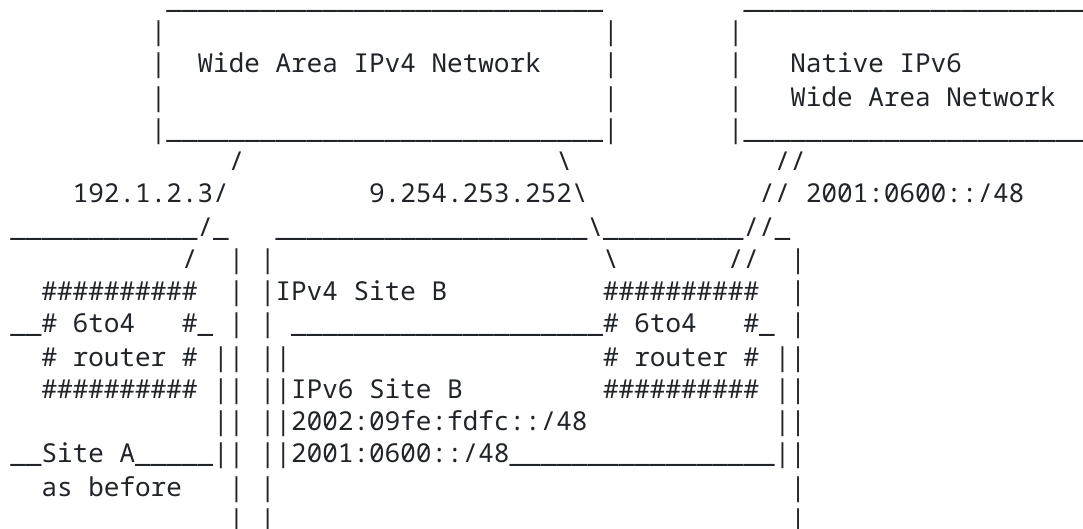
If the host picks the 6to4 prefix according to some rule for multiple
prefixes, it will simply send packets to an IPv6 address formed with
the prefix {FP=001,TLA=0x0002,NLA=192.1.2.3}/48.  It is essential
that they are sourced from the prefix
{FP=001,TLA=0x0002,NLA=9.254.253.252}/48 for two-way connectivity to
be possible.  The address selection mechanism of Section 2.1 will
ensure this.

**5.2.1 Variant scenario with ISP relay**

   The previous scenario assumes that the relay router is provided by a
   cooperative 6to4 user site.  A variant of this is for an Internet
   Service Provider, that already offers native IPv6 connectivity, to
   operate a relay router.  Technically this is no different from the
   previous scenario; site B is simply an internal 6to4 site of the ISP,
   possibly containing only one system, i.e., the relay router itself.

**5.2.2 Summary of relay router configuration**

   A relay router participates in IPv6 unicast routing protocols on its
   native IPv6 interface and may do so on its 6to4 pseudo-interface, but
   these are independent routing domains with separate policies, even if
   the same protocol, probably BGP4+, is used in both cases.

   A relay router also participates in IPv4 unicast routing protocols on
   its IPv4 interface used to support 6to4, but this is not further
   discussed here.

   On its native IPv6 interface, the relay router MUST advertise a route
   to 2002::/16.  It MUST NOT advertise a longer 2002:: routing prefix
   on that interface.  Routing policy within the native IPv6 routing
   domain determines the scope of that advertisement, thereby limiting
   the visibility of the relay router in that domain.

   IPv6 packets received by the relay router whose next hop IPv6 address
   matches 2002::/16 will be routed to its 6to4 pseudo-interface and
   treated according to the sending rule of Section 5.1.

**5.2.2.1. BGP4+ not used**

   If BGP4+ is not deployed in the 6to4 exterior routing domain (option
   2.1 of Section 5.2), the relay router will be configured to accept
   and relay all IPv6 traffic only from its client 6to4 sites.  Each
   6to4 router served by the relay router will be configured with a
   default IPv6 route to the relay router (for example, Site A's default
   IPv6 route ::/0 would point to the relay router's address under
   prefix 2002:09fe:fdfc::/48).

**5.2.2.2. BGP4+ used**

   If BGP4+ is deployed in the 6to4 exterior routing domain (option 2.2
   of Section 5.2), the relay router advertises IPv6 native routing
   prefixes on its 6to4 pseudo-interface, peering only with the 6to4
   routers that it serves.  (An alternative is that these routes could
   be advertised along with IPv4 routes using BGP4 over IPv4, rather
   than by running a separate BGP4+ session.)  The specific routes

advertised depend on applicable routing policy, but they must be
chosen from among those reachable through the relay router's native
IPv6 interface.  In the simplest case, a default route to the whole
IPv6 address space could be advertised.  When multiple relay routers
are in use, more specific routing prefixes would be advertised
according to the desired routing policy.  The usage of BGP4+ is
completely standard so is not discussed further in this document.

**5.2.2.3. Relay router scaling**

Relay routers introduce the potential for scaling issues.  In general
a relay router should not attempt to serve more sites than any other
transit router, allowing for the encapsulation overhead.

**5.2.3 Unwilling to relay**

It may arise that a site has a router with both 6to4 pseudo-
interfaces and native IPv6 interfaces, but is unwilling to act as a
relay router.  Such a site MUST NOT advertise any 2002:: routing
prefix into the native IPv6 domain and MUST NOT advertise any native
IPv6 routing prefixes or a default IPv6 route into the 6to4 domain.
Within the 6to4 domain it will behave exactly as in the basic 6to4
scenario of Section 5.1.

**5.3 Sending and decapsulation rules**

The only change to standard IPv6 forwarding is that every 6to4 router
(and only 6to4 routers) MUST implement the following additional
sending and decapsulation rules.

In the sending rule, "next hop" refers to the next IPv6 node that the
packet will be sent to, which is not necessarily the final
destination, but rather the next IPv6 neighbor indicated by normal
IPv6 routing mechanisms.  If the final destination is a 6to4 address,
it will be considered as the next hop for the purpose of this rule.
If the final destination is not a 6to4 address, and is not local, the
next hop indicated by routing will be the 6to4 address of a relay
router.

ADDITIONAL SENDING RULE for 6to4 routers

        if the next hop IPv6 address for an IPv6 packet
            does match the prefix 2002::/16, and
            does not match any prefix of the local site
                  then
                  apply any security checks (see Section 8);
                  encapsulate the packet in IPv4 as in Section 3,

                with IPv4 destination address = the NLA value V4ADDR
                extracted from the next hop IPv6 address;
                queue the packet for IPv4 forwarding.

   A simple decapsulation rule for incoming IPv4 packets with protocol
   type 41 MUST be implemented:

   ADDITIONAL DECAPSULATION RULE for 6to4 routers

           apply any security checks (see Section 8);
           remove the IPv4 header;
           submit the packet to local IPv6 routing.

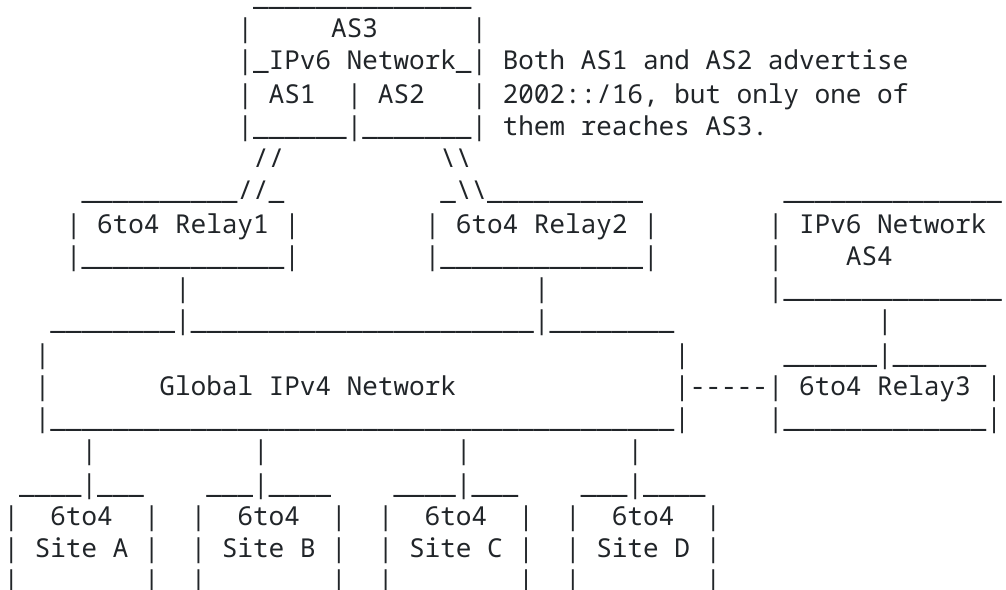## 5.4 Variant scenario with tunnel to IPv6 space

   A 6to4 site which has no IPv6 connections to the "native" IPv6
   Internet can acquire effective connectivity to the v6 Internet via a
   "configured tunnel" (using the terminology in [MECH]) to a
   cooperating router which does have IPv6 access, but which does not
   need to be a 6to4 router. Such tunnels could be autoconfigured using
   an IPv4 anycast address, but this is outside of the scope of this
   document.  Alternatively a tunnel broker can be used.  This scenario
   would be suitable for a small user-managed site.

   These mechanisms are not described in detail in this document.

## 5.5 Fragmented Scenarios

   If there are multiple relay routers between native IPv6 and the 6to4
   world, different parts of the 6to4 world will be served by different
   relays.  The only complexity that this introduces is in the scoping
   of 2002::/16 routing advertisements within the native IPv6 world.
   Like any BGP4+ advertisements, their scope must be correctly defined
   by routing policy to ensure that traffic to 2002::/16 follows the
   intended paths.

   If there are multiple IPv6 stubs all interconnected by 6to4 through
   the global IPv4 Internet, this is a simple generalization of the
   basic scenarios of sections 5.1. and 5.2 and no new issues arise.
   This is shown in the following figure.  Subject to consistent
   configuration of routing advertisements, there are no known issues
   with this scenario.

```
                 _____
                |     AS3      |
                |_IPv6 Network_|  Both AS1 and AS2 advertise
                | AS1  | AS2   |  2002::/16, but only one of
                |_____|_____|  them reaches AS3.
                 //          \\
        _____//_        _\_____                 _____
       | 6to4 Relay1 |     | 6to4 Relay2 |              | IPv6 Network |
       |_____|     |_____|              |     AS4      |
             |                   |                      |_____|
          ___|_____|_____                     |
         |                                |               _____|_____
         |      Global IPv4 Network       |-----| 6to4 Relay3 |
         |_____|             |_____|
            |         |         |         |
         ___|___   ___|___   ___|___   ___|___
        | 6to4  | | 6to4  | | 6to4  | | 6to4  |
        | Site A| | Site B| | Site C| | Site D|
        |_____| |_____| |_____| |_____|
```

If multiple IPv6 stubs are interconnected through multiple, disjoint
IPv4 networks (i.e., a fragmented IPv4 world) then the 6to4 world is
also fragmented; this is the one scenario that must be avoided.  It
is illustrated below to show why it does not work, since the
2002::/16 advertisement from Relay1 will be invisible to Relay2, and
vice versa.  Sites A and B therefore have no connectivity to sites C
and D.

```
                 _____
                |     AS3      |
                |_IPv6 Network_|  Both AS1 and AS2 advertise
                | AS1  | AS2   |  2002::/16, but sites A and B
                |_____|_____|  cannot reach C and D.
                 //          \\
        _____//_        _\_____
       | 6to4 Relay1 |     | 6to4 Relay2 |
       |_____|     |_____|
             |                   |
         ____|_____          ____|_____
        | IPv4 Network |    | IPv4 Network |
        | Segment 1    |    | Segment 2    |
        |_____|    |_____|
           |         |         |         |
        ___|___   ___|___   ___|___   ___|___
       | 6to4  | | 6to4  | | 6to4  | | 6to4  |
       | Site A| | Site B| | Site C| | Site D|
       |_____| |_____| |_____| |_____|
```
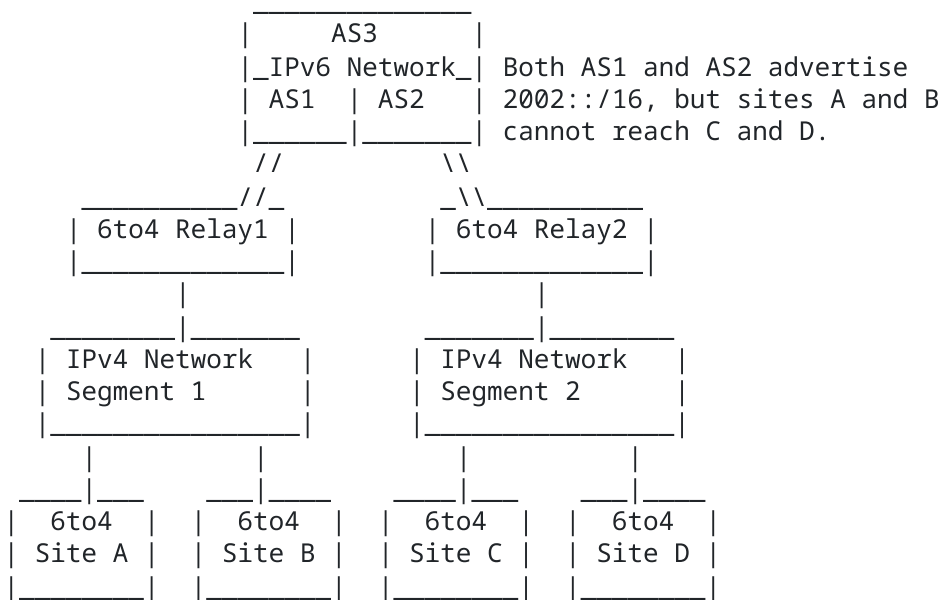
## 5.6 Multihoming

Sites which are multihomed on IPv4 MAY extend the 6to4 scenario by
using a 2002:: prefix for each IPv4 border router, thereby obtaining
a simple form of IPv6 multihoming by using multiple simultaneous IPv6
prefixes and multiple simultaneous relay routers.

## 5.7 Transition Considerations

If the above rules for routing advertisements and address selection
are followed, then a site can migrate from using 6to4 to using native
IPv6 connections over a long period of co-existence, with no need to
stop 6to4 until it has ceased to be used.  The stages involved are

1. Run IPv6 on site using any suitable implementation.  True native
IPv6, [6OVER4], or tunnels are all acceptable.

2. Configure a border router (or router plus IPv4 NAT) connected to
the external IPv4 network to support 6to4, including advertising the
appropriate 2002:: routing prefix locally.  Configure IPv6 DNS
entries using this prefix.  At this point the 6to4 mechanism is
automatically available, and the site has obtained a "free" IPv6
prefix.

3. Identify a 6to4 relay router willing to relay the site's traffic
to the native IPv6 world.  This could either be at another
cooperative 6to4 site, or an ISP service.  If no exterior routing
protocol is in use in the 6to4 exterior routing domain, the site's
6to4 router will be configured with a default IPv6 route pointing to
that relay router's 6to4 address.  If an exterior routing protocol
such as BGP4+ is in use, the site's 6to4 router will be configured to
establish appropriate BGP peerings.

4. When native external IPv6 connectivity becomes available, add a
second (native) IPv6 prefix to both the border router configuration
and the DNS configuration.  At this point, an address selection rule
will determine when 6to4 and when native IPv6 will be used.

5. When 6to4 usage is determined to have ceased (which may be several
years later), remove the 6to4 configuration.

## 5.8 Coexistence with firewall, NAT or RSIP

The 6to4 mechanisms appear to be unaffected by the presence of a
firewall at the border router.

If the site concerned has very limited global IPv4 address space, and
is running an IPv4 network address translator (NAT), all of the above
mechanisms remain valid.  The NAT box must also contain a fully
functional IPv6 router including the 6to4 mechanism.  The address
used for V4ADDR will simply be a globally unique IPv4 address
allocated to the NAT.  In the example of Section 5.1 above, the 6to4
routers would also be the sites' IPv4 NATs, which would own the
globally unique IPv4 addresses 192.1.2.3 and 9.254.253.252.

Combining a 6to4 router with an IPv4 NAT in this way offers  the site
concerned a globally unique IPv6 /48 prefix, automatically, behind
the IPv4 address of the NAT.  Thus every host behind the NAT can
become an IPv6 host with no need for additional address space
allocation, and no intervention by the Internet service provider.  No
address translation is needed by these IPv6 hosts.

A more complex situation arises if a host is more than one NAT hop
away from the globally unique IPv4 address space, since only the
outermost NAT has a unique IPv4 address.  All IPv6 hosts in this
situation must use addresses derived from the 2002: prefix
constructed from the global IPv4 address of the outermost NAT.  The
IPv4 addresses of the inner NATs are not globally unique and play no
part in the 6to4 mechanism, and 6to4 encapsulation and decapsulation
can only take place at the outermost NAT.

The Realm-Specific IP (RSIP) mechanism [RSIP] can also co-exist with
6to4.  If a 6to4 border router is combined with an RSIP border
router, it can support IPv6 hosts using 6to4 addresses, IPv4 hosts
using RSIP, or dual stack hosts using both.  The RSIP function
provides fine-grained management of dynamic global IPv4 address
allocation and the 6to4 function provides a stable IPv6 global
address to each host.  As with NAT, the IPv4 address used to
construct the site's 2002:  prefix will be one of the global
addresses of the RSIP border router.

## 5.9 Usage within Intranets

There is nothing to stop the above scenario being deployed within a
private corporate network as part of its internal transition to IPv6;
the corporate IPv4 backbone would serve as the virtual link layer for
individual corporate sites using 2002:: prefixes.  The V4ADDR MUST be
a duly allocated global IPv4 address, which MUST be unique within the
private network.  The Intranet thereby obtains globally unique IPv6
addresses even if it is internally using private IPv4 addresses [RFC
1918].

5.10 Summary of impact on routing

   IGP (site) routing will treat the local site's 2002::/48  prefix
   exactly like a native IPv6 site prefix assigned to the local site.
   There will also be an IGP route to the generic 2002::/16 prefix,
   which will be a route to the site's 6to4 router, unless this is
   handled as a default route.

   EGP (i.e., BGP) routing will include advertisements for the 2002::/16
   prefix from relay routers into the native IPv6 domain, whose scope is
   limited by routing policy.  This is the only non-native IPv6 prefix
   advertised by BGP.

   It will be necessary for 6to4 routers to obtain routes to relay
   routers in order to access the native IPv6 domain.  In the simplest
   case there will be a manually configured default IPv6 route to a
   relay router's address under the prefix
   {FP=001,TLA=0x0002,NLA=V4ADDR}/48, where V4ADDR is the IPv4 address
   of the relay router.  Such a route could be used to establish a BGP
   session for the exchange of additional IPv6 routes.

   By construction, unicast IPv6 traffic within a 6to4 domain will
   follow exactly the same path as unicast IPv4 traffic.

5.11. Routing loop prevention

   Since 6to4 has no impact on IPv4 routing, it cannot induce routing
   loops in IPv4.  Since 2002: prefixes behave exactly like standard
   IPv6 prefixes, they will not create any new mechanisms for routing
   loops in IPv6 unless misconfigured.  One very dangerous
   misconfiguration would be an announcement of the 2002::/16 prefix
   into a 6to4 exterior routing domain, since this would attract all
   6to4 traffic into the site making the announcement.  Its 6to4 router
   would then resend non-local 6to4 traffic back out, forming a loop.

   The 2002::/16 routing prefix may be legitimately advertised into the
   native IPv6 routing domain by a relay router, and into an IPv6 site's
   local IPv6 routing domain; hence there is a risk of misconfiguration
   causing it to be advertised into a 6to4 exterior routing domain.

   To summarize, the 2002::/16 prefix MUST NOT be advertised to a 6to4
   exterior routing domain.

## 6. Multicast and Anycast

It is not possible to assume the general availability of wide-area
IPv4 multicast, so (unlike [6OVER4]) the 6to4 mechanism must assume
only unicast capability in its underlying IPv4 carrier network.  An
IPv6 multicast routing protocol is needed [MULTI].

The allocated anycast address space [ANYCAST] is compatible with
2002:: prefixes, i.e., anycast addresses formed with such prefixes
may be used inside a 6to4 site.

## 7. ICMP messages

ICMP "unreachable" and other messages returned by the IPv4 routing
system will be returned to the 6to4 router that generated a
encapsulated 2002:: packet.  However, this router will often be
unable to return an ICMPv6 message to the originating IPv6 node, due
to the lack of sufficient information in the "unreachable" message.
This means that the IPv4 network will appear as an undiagnosable link
layer for IPv6 operational purposes.  Other considerations are as
described in Section 4.1.3 of [MECH].

## 8. IANA Considerations

No assignments by the IANA are required beyond the special TLA value
0x0002 already assigned.

## 9. Security Considerations

Implementors should be aware that, in addition to possible attacks
against IPv6, security attacks against IPv4 must also be considered.
Use of IP security at both IPv4 and IPv6 levels should nevertheless
be avoided, for efficiency reasons.  For example, if IPv6 is running
encrypted, encryption of IPv4 would be redundant except if traffic
analysis is felt to be a threat.  If IPv6 is running authenticated,
then authentication of IPv4 will add little.  Conversely, IPv4
security will not protect IPv6 traffic once it leaves the 6to4
domain.  Therefore, implementing IPv6 security is required even if
IPv4 security is available.

By default, 6to4 traffic will be accepted and decapsulated from any
source from which regular IPv4 traffic is accepted.  If this is for
any reason felt to be a security risk (for example, if IPv6 spoofing
is felt to be more likely than IPv4 spoofing), then additional source
address based packet filtering could be applied.  A possible
plausibility check is whether the encapsulating IPv4 address is
consistent with the encapsulated 2002:: address.  If this check is

applied, exceptions to it must be configured to admit traffic from
relay routers (Section 5).  2002:: traffic must also be excepted from
checks applied to prevent spoofing of "6 over 4" traffic [6OVER4].

In any case, any 6to4 traffic whose source or destination address
embeds a V4ADDR which is not in the format of a global unicast
address MUST be silently discarded by both encapsulators and
decapsulators.  Specifically, this means that IPv4 addresses defined
in [RFC 1918], broadcast, subnet broadcast, multicast and loopback
addresses are unacceptable.

## Acknowledgements

The basic idea presented above is probably not original, and we have
had invaluable comments from Magnus Ahltorp, Harald Alvestrand, Jim
Bound, Scott Bradner, Randy Bush, Matt Crawford, Richard Draves,
Jun-ichiro itojun Hagino, Joel Halpern, Tony Hain, Andy Hazeltine,
Bob Hinden, Geoff Huston, Perry Metzger, Thomas Narten, Erik
Nordmark, Markku Savela, Ole Troan, Sowmini Varadhan, members of the
Compaq IPv6 engineering team, and other members of the NGTRANS
working group.  Some text has been copied from [6OVER4].  George
Tsirtsis kindly drafted two of the diagrams.

## References

[AARCH]     Hinden, R. and S. Deering, "IP Version 6 Addressing
            Architecture", RFC 2373, July 1998.

[AGGR]      Hinden., R, O'Dell, M. and S. Deering, "An IPv6
            Aggregatable Global Unicast Address Format", RFC 2374,
            July 1998.

[API]       Gilligan, R., Thomson, S., Bound, J. and W. Stevens,
            "Basic Socket Interface Extensions for IPv6", RFC 2553,
            March 1999.

[BGP4+]     Marques, P. and F. Dupont, "Use of BGP-4 Multiprotocol
            Extensions for IPv6 Inter-Domain Routing", RFC 2545, March
            1999.

[CONF]      Thomson, S. and T. Narten, "IPv6 Stateless Address
            Autoconfiguration", RFC 2462, December 1998.

[DISC]      Narten, T., Nordmark, E. and W. Simpson, "Neighbor
            Discovery for IP Version 6 (IPv6)", RFC 2461, December
            1998.

   [IPV6]      Deering, S. and R. Hinden, "Internet Protocol, Version 6
               (IPv6) Specification", RFC 2460, December 1998.

   [6OVER4]    Carpenter, B. and C. Jung, "Transmission of IPv6 over IPv4
               Domains without Explicit Tunnels", RFC 2529, March 1999.

   [ANYCAST]   Johnson, D. and S. Deering, "Reserved IPv6 Subnet Anycast
               Addresses", Work in Progress.

   [MULTI]     Thaler, D., "Support for Multicast over 6to4 Networks",
               Work in Progress.

   [SCALE]     Hain, T., "6to4-relay discovery and scaling", Work in
               Progress.

   [SELECT]    Draves, R., "Default Address Selection for IPv6", Work in
               Progress.

   [RFC 791]   Postel, J., "Internet Protocol", STD 5, RFC 791, September
               1981.

   [RFC 1918]  Rekhter, Y., Moskowitz, R., Karrenberg, D., de Groot, G.
               and E. Lear, "Address Allocation for Private Internets",
               BCP 5, RFC 1918, February 1996.

   [MECH]      Gilligan, R. and E. Nordmark, "Transition Mechanisms for
               IPv6 Hosts and Routers", RFC 2893, August 2000.

   [RSIP]      Borella, M., Grabelsky, D., Lo, J. and K. Tuniguchi,
               "Realm Specific IP: Protocol Specification", Work in
               Progress.

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

   [RFC 2283]  Bates, T., Chandra, R., Katz, D. and Y. Rekhter,
               "Multiprotocol Extensions for BGP-4", RFC 2283, February
               1998.

Authors' Addresses

    Brian E. Carpenter
    IBM
    iCAIR, Suite 150
    1890 Maple Avenue
    Evanston IL 60201, USA

    EMail: brian@icair.org


    Keith Moore
    UT Computer Science Department
    1122 Volunteer Blvd, Ste 203
    Knoxville, TN  37996-3450
    USA

    EMail: moore@cs.utk.edu

Intellectual Property

    The IETF takes no position regarding the validity or scope of any
    intellectual property or other rights that might be claimed to
    pertain to the implementation or use of the technology described in
    this document or the extent to which any license under such rights
    might or might not be available; neither does it represent that it
    has made any effort to identify any such rights.  Information on the
    IETF's procedures with respect to rights in standards-track and
    standards-related documentation can be found in BCP-11.  Copies of
    claims of rights made available for publication and any assurances of
    licenses to be made available, or the result of an attempt made to
    obtain a general license or permission for the use of such
    proprietary rights by implementors or users of this specification can
    be obtained from the IETF Secretariat.

    The IETF invites any interested party to bring to its attention any
    copyrights, patents or patent applications, or other proprietary
    rights which may cover technology that may be required to practice
    this standard.  Please address the information to the IETF Executive
    Director.

Full Copyright Statement

Acknowledgement