                        6to4 Provider Managed Tunnels

Abstract

   6to4 Provider Managed Tunnels (6to4-PMT) provide a framework that can
   help manage 6to4 tunnels operating in an anycast configuration.  The
   6to4-PMT framework is intended to serve as an option for operators to
   help improve the experience of 6to4 operation when conditions of the
   network may provide sub-optimal performance or break normal 6to4
   operation. 6to4-PMT supplies a stable provider prefix and forwarding
   environment by utilizing existing 6to4 relays with an added function
   of IPv6 Prefix Translation.  This operation may be particularly
   important in NAT444 infrastructures where a customer endpoint may be
   assigned a non-RFC1918 address, thus breaking the return path for
   anycast-based 6to4 operation.  6to4-PMT has been successfully used in
   a production network, implemented as open source code, and
   implemented by a major routing vendor.

Status of This Memo

Copyright Notice

Table of Contents

1.  Introduction

   6to4 [RFC3056] tunneling, along with the anycast operation described
   in [RFC3068], is widely deployed in modern Operating Systems and
   off-the-shelf gateways sold throughout retail and Original Equipment
   Manufacturer (OEM) channels.  Anycast-based 6to4 [RFC3068] allows for
   tunneled IPv6 connectivity through IPv4 clouds without explicit
   configuration of a relay address.  Since the overall system utilizes
   anycast forwarding in both directions, flow paths are difficult to
   determine, tend to follow separate paths in either direction, and
   often change based on network conditions.  The return path is
   normally uncontrolled by the local operator and can contribute to
   poor performance for IPv6 and can also act as a breakage point.  Many
   of the challenges with 6to4 are described in [RFC6343].  A specific
   critical use case for problematic anycast 6to4 operation is related
   to conditions in which the consumer endpoints are downstream from a
   northbound Carrier-Grade NAT (CGN) [RFC6264] function when assigned
   non-RFC1918 IPv4 addresses, which are not routed on interdomain
   links.

   Operators that are actively deploying IPv6 networks and operate
   legacy IPv4 access environments may want to utilize the existing 6to4
   behavior in customer site resident hardware and software as an
   interim option to reach the IPv6 Internet in advance of being able to
   offer full native IPv6.  Operators may also need to address the
   brokenness related to 6to4 operation originating from behind a
   provider NAT function. 6to4-PMT offers an operator the opportunity to
   utilize IPv6 Prefix Translation to enable deterministic traffic flow
   and an unbroken path to and from the Internet for IPv6-based traffic
   sourced originally from these 6to4 customer endpoints.

   6to4-PMT translates the prefix portion of the IPv6 address from the
   6to4-generated prefix to a provider-assigned prefix that is used to
   represent the source.  This translation will then provide a stable
   forward and return path for the 6to4 traffic by allowing the existing
   IPv6 routing and policy environment to control the traffic. 6to4-PMT
   is primarily intended to be used in a stateless manner to maintain
   many of the elements inherent in normal 6to4 operation.
   Alternatively, 6to4-PMT can be used in a stateful translation mode
   should the operator choose this option.

2.  Motivation

   Many operators endeavor to deploy IPv6 as soon as possible so as to
   ensure uninterrupted connectivity to all Internet applications and
   content through the IPv4 to IPv6 transition process.  The IPv6
   preparations within these organizations are often faced with both
   financial challenges and timing issues related to deploying IPv6 to

the network edge and related transition technologies.  Many of the
new technologies available for IPv4 to IPv6 transition will require
the replacement of the organization's Customer Premises Equipment
(CPE) to support technologies like IPv6 Rapid Deployment (6RD)
[RFC5969], Dual-Stack Lite [RFC6333], and native dual-stack.

Operators face a number of challenges related to home equipment
replacement.  Operator-initiated replacement of this equipment will
take time due to the nature of mass equipment refresh programs or may
require the consumer to replace their own gear.  Replacing consumer
owned and operated equipment, compounded by the fact that there is
also a general unawareness of what IPv6 is, also adds to the
challenges faced by operators.  It is also important to note that
6to4 is present in much of the equipment found in networks today that
do not as of yet, or will not, support 6RD and/or native IPv6.

Operators may still be motivated to provide a form of IPv6
connectivity to customers and would want to mitigate potential issues
related to IPv6-only deployments elsewhere on the Internet.
Operators also need to mitigate issues related to the fact that 6to4
operation is often on by default, and may be subject to erroneous
behavior.  The undesired behavior may be related to the use of
non-RFC1918 addresses on CPE equipment that operate behind large
operator NATs or other conditions as described in a general advisory
as laid out in [RFC6343].

6to4-PMT allows an operator to help mitigate such challenges by
leveraging the existing 6to4 deployment base, while maintaining
operator control of access to the IPv6 Internet.  It is intended for
use when better options, such as 6RD or native IPv6, are not yet
viable.  One of the key objectives of 6to4-PMT is to also help
reverse the negative impacts of 6to4 in CGN environments.  The
6to4-PMT operation can also be used immediately with the default
parameters that are often enough to allow it to operate in a 6to4-PMT
environment.  Once native IPv6 is available to the endpoint, the
6to4-PMT operation is no longer needed and will cease to be used
based on correct address selection behaviors in end hosts [RFC6724].

6to4-PMT thus helps operators remove the impact of 6to4 in CGN
environments, deals with the fact that 6to4 is often on by default,
and allows access to IPv6-only endpoints from IPv4-only addressed
equipment.  Additionally, it provides relief from many challenges
related to mis-configurations in other networks that control return
flows via foreign relays.  Due to the simple nature of 6to4-PMT, it
can also be implemented in a cost-effective and simple manner,
allowing operators to concentrate their energy on deploying native
IPv6.

3.  6to4 Provider Managed Tunnels

3.1.  6to4 Provider Managed Tunnel Model

   The 6to4 managed tunnel model behaves like a standard 6to4 service
   between the customer IPv6 host or gateway, and the 6to4-PMT Relay
   (within the provider domain).  The 6to4-PMT Relay shares properties
   with 6RD [RFC5969] by decapsulating and forwarding encapsulated IPv6
   flows within an IPv4 packet to the IPv6 Internet.  The model provides
   an additional function that translates the source 6to4 prefix to a
   provider-assigned prefix that is not found in 6RD [RFC5969] or
   traditional 6to4 operation.

   The 6to4-PMT Relay is intended to provide a stateless (or stateful)
   mapping of the 6to4 prefix to a provider supplied prefix.

```
                         | 6to4-PMT Operation  |

            +-----+ 6to4 Tunnel +--------+  +------+  IPv6     +----+
            | CPE |-------------|6to4 BR |--| PT66 |--------- |Host|
            +-----+    IPv4     +--------+  +------+ Provider +----+
                    Network                          Prefix
                              Unified or Separate
                               Functions/Platforms
```
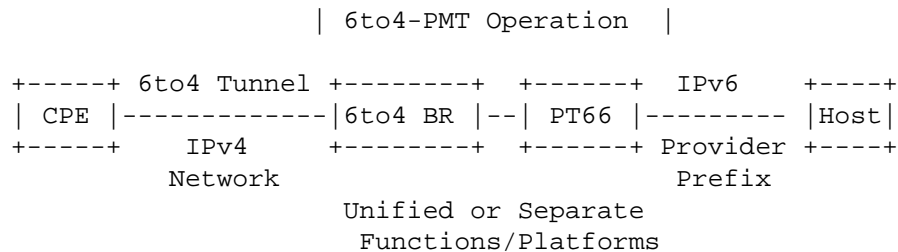
                    Figure 1: 6to4-PMT Functional Model

   This mode of operation is seen as beneficial when compared to broken
   6to4 paths and/or environments where 6to4 operation may be functional
   but highly degraded.

3.2.  Traffic Flow

   Traffic in the 6to4-PMT model is intended to be controlled by the
   operator's IPv6 peering operations.  Egress traffic is managed
   through outgoing routing policy, and incoming traffic is influenced
   by the operator-assigned prefix advertisements using normal
   interdormain routing functions.

   The routing model is as predictable as native IPv6 traffic and legacy
   IPv4-based traffic.  Figure 2 provides a view of the routing topology
   needed to support this relay environment.  The diagram references
   PrefixA as 2002::/16 and PrefixB as the example 2001:db8::/32.

```
         | 6to4 IPv4 Path      |         Native IPv6 Path          |
            -----------       -----------      -------------
           /  IPv4 Net \     /  IPv6 Net  \  / IPv6 Internet \
     +------+        +--------+        +-------+    +---------+
     | CPE  | PrefixA |6to4-PMT| PrefixB |Peering|    |IPv6 HOST|
     +------+        +--------+        +-------+    +---------+
          \          /   \            /  \            /
           ----------       ------------     --------------

            IPv4 6to4        IPv6 Provider      IPv6 Prefix
             Anycast            Prefix          Propagation
```
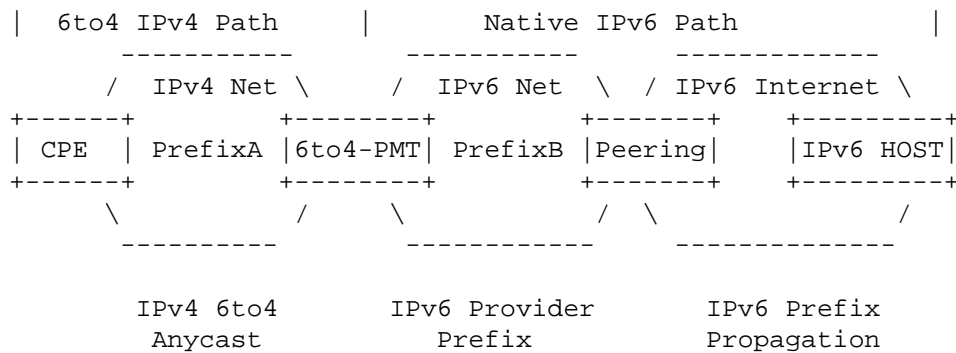
                     Figure 2: 6to4-PMT Flow Model

   Traffic between two 6to4-enabled devices would use the IPv4 path for
   communication according to [RFC3056] unless the local host still
   prefers traffic via a relay.  6to4-PMT is intended to be deployed in
   conjunction with the 6to4 relay function in an attempt to help
   simplify its deployment.  The model can also provide the ability for
   an operator to forward both 6to4-PMT (translated) and normal 6to4
   flows (untranslated) simultaneously based on configured policy.
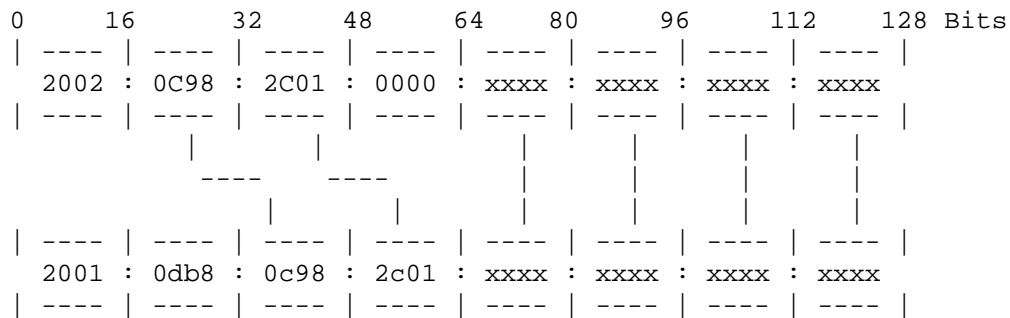
3.3.  Prefix Translation

   IPv6 Prefix Translation is a key part of the system as a whole.  The
   6to4-PMT framework is a combination of two concepts: 6to4 [RFC3056]
   and IPv6 Prefix Translation.  IPv6 Prefix Translation, as used in
   6to4-PMT, has some similarities to concepts discussed in [RFC6296].
   6to4-PMT would provide prefix translation based on specific rules
   configured on the translator that maps the 6to4 2002::/16 prefix to
   an appropriate provider assigned prefix.  In most cases, a ::/32
   prefix would work best in 6to4-PMT that matches common Regional
   Internet Registry (RIR) prefix assignments to operators.

   The provider can use any prefix mapping strategy they so choose, but
   the simpler the better.  Simple direct bitmapping can be used, or
   more advanced forms of translation should the operator want to
   achieve higher address compression.  More advanced forms of
   translation may require the use of stateful translation.

   Figure 3 shows a 6to4 Prefix with a Subnet-ID of "0000" mapped to a
   provider-assigned, globally unique prefix (2001:db8::/32).  With this
   simple form of translation, there is support for only one Subnet-ID
   per provider-assigned prefix.  In characterization of deployed OSs
   and gateways, a Subnet-ID of "0000" is the most common default case
   followed by Subnet-ID "0001".  Use of the Subnet-ID can be referenced
   in [RFC4291].  It should be noted that in normal 6to4 operation, the
   endpoint (network) has access to 65,536 (16-bits) Subnet IDs.  In the

6to4-PMT case as described above using the mapping in Figure 3, all
but the one Subnet-ID used for 6to4-PMT would still operate under
normal 6to4 operation.

Pre-Relayed Packet [Provider Access Network Side]

```
0     16      32      48      64      80      96      112     128 Bits
| ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- |
  2002 : 0C98 : 2C01 : 0000 : xxxx : xxxx : xxxx : xxxx
| ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- |
         |        |              |      |      |      |
       ----    ----              |      |      |      |
         |        |              |      |      |      |
| ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- |
  2001 : 0db8 : 0c98 : 2c01 : xxxx : xxxx : xxxx : xxxx
| ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- |
```

Post-Relayed Packet [Internet Side]

Figure 3: 6to4-PMT Prefix Mapping

3.4.  Translation State

It is preferred that the overall system use deterministic prefix
translation mappings such that stateless operation can be
implemented.  This allows the provider to place N number of relays
within the network without the need to manage translation state.
Deterministic translation also allows a customer to employ inward
services using the translated (provider prefix) address.

If stateful operation is chosen, the operator would need to validate
state and routing requirements particular to that type of deployment.
The full body of considerations for this type of deployment is not
within this scope of this document.

4.  Deployment Considerations and Requirements

4.1.  Customer Opt-Out

A provider enabling this function should offer a method to allow
customers to opt-out of such a service should the customer choose to
maintain normal 6to4 operation irrespective of degraded performance.
In cases where the customer is behind a CGN device, the customer
would not be advised to opt-out and can be assisted in turning off
6to4.

Since the 6to4-PMT system is targeted at customers who are relatively
unaware of IPv6 and IPv4, and normally run network equipment with a
default configuration, an opt-out strategy is recommended.  This
method provides 6to4-PMT operation for non-IPv6 savvy customers whose
equipment may turn on 6to4 automatically and allows savvy customers
to easily configure their way around the 6to4-PMT function.

Capable customers can also disable anycast-based 6to4 entirely and
use traditional 6to4 or other tunneling mechanisms if they are so
inclined.  This is not considered the normal case, and most endpoints
with auto-6to4 functions will be subject to 6to4-PMT operation since
most users are unaware of its existence. 6to4-PMT is targeted as an
option for stable IPv6 connectivity for average consumers.

## 4.2.  Shared CGN Space Considerations

6to4-PMT operation can also be used to mitigate a known problem with
6to4 occurring when shared address space [RFC6598] or Global Unicast
Addresses (GUA) are used behind a CGN and not routed on the Internet.
Non-RFC1918, yet unrouted (on interdomain links) address space would
cause many deployed OSs and network equipment to potentially
auto-enable 6to4 operation even without a valid return path (such as
behind a CGN function).  The operator's desire to use non-RFC1918
addresses, such as shared address space [RFC6598], is considered
highly likely based on real world deployments.

Such hosts, in normal cases, would send 6to4 traffic to the IPv6
Internet via the anycast relay, which would in fact provide broken
IPv6 connectivity, since the return path flow is built using an IPv4
address that is not routed or assigned to the source network.  The
use of 6to4-PMT would help reverse these effects by translating the
6to4 prefix to a provider-assigned prefix, masking this automatic and
undesired behavior.

## 4.3.  End-to-End Transparency

The 6to4-PMT mode of operation removes the traditional end-to-end
transparency of 6to4.  Remote hosts would connect to a 6to4-PMT-
serviced host using a translated IPv6 address versus the original
6to4 address based on the 2002::/16 well-known prefix.  This can be
seen as a disadvantage of the 6to4-PMT system.  This lack of
transparency should also be contrasted with the normal operating
state of 6to4 that provides connectivity that is uncontrolled and
often prone to high latency.  The lack of transparency is, however, a
better form of operation when extreme poor performance, broken IPv6
connectivity, or no IPv6 connectivity is considered as the
alternative.

4.4.  Path MTU Discovery Considerations

   The MTU will be subject to a reduced value due to standard 6to4
   tunneling operation.  Under normal 6to4 operation, the 6to4 service
   agent would send an ICMP Packet Too Big Message as part of Path MTU
   discovery as described in [RFC4443] and [RFC1981], respectively.  In
   6to4-PMT operation, the PMT Service agent should be aware of the
   reduced 6to4 MTU and send ICMP messages using the translated address
   accordingly.

   It is also possible to pre-constrain the MTU at the upstream router
   from the 6to4-PMT service agents that would then have the upstream
   router send the appropriate ICMP Packet Too Big Messages.

4.5.  Checksum Management

   Checksum management for 6to4-PMT can be implemented in one of two
   ways.  The first deployment model is based on the stateless 6to4-PMT
   operational mode.  In this case, checksum modifications are made
   using the method described in [RFC3022], Section 4.2.  The checksum
   is modified to match the parameters of the translated address of the
   source 6to4-PMT host.  In the second deployment model in which
   stateful 6to4-PMT translation is used, the vendor can implement
   checksum-neutral mappings as defined in [RFC6296].

4.6.  Application Layer Gateways

   Vendors can choose to deploy Application Layer Gateways (ALGs) on
   their platforms that perform 6to4-PMT if they so choose.  No ALGs
   were deployed as part of the open source and vendor product
   deployments of 6to4-PMT.  In the vendor deployment case, the same
   rules were used as with their NPTv6 [RFC6296] base code.

4.7.  Routing Requirements

   The provider would need to advertise the well-known IP address range
   used for normal anycast 6to4 [RFC3068] operation within the local
   IPv4 routing environment.  This advertisement would attract the 6to4
   upstream traffic to a local relay.  To control this environment and
   make sure all northbound traffic lands on a provider-controlled
   relay, the operator may filter the anycast range from being
   advertised from customer endpoints toward the local network (upstream
   propagation).

   The provider would not be able to control route advertisements inside
   the customer domain, but that use case is not in scope for this
   document.  In that case, it is likely that the end network/customer
   understands 6to4 and is maintaining their own relay environment and

   therefore would not be subject to the operators 6to4 and/or PMT
   operation.

   Within their own network, the provider would also likely want to
   advertise the 2002::/16 range to help bridge traditional 6to4 traffic
   within the network (native IPv6 to 6to4-PMT-based endpoint).  It
   would also be advised that the local 6to4-PMT operator not leak the
   well-known 6to4 anycast IPv4 prefix to neighboring Autonomous Systems
   to prevent PMT operation for neighboring networks.  Policy
   configuration on the local 6to4-PMT Relay can also be used to
   disallow PMT operation should the local provider service downstream
   customer networks.

4.8.  Relay Deployments

   The 6to4-PMT function can be deployed onto existing 6to4 relays (if
   desired) to help minimize network complexity and cost. 6to4-PMT has
   already been developed on Linux-based platforms that are package
   add-ons to the traditional 6to4 code.  The only additional
   considerations beyond normal 6to4 relay operation would include the
   need to route specific IPv6 provider prefix ranges used for 6to4-PMT
   operation towards peers and transit providers.

5.  Security Considerations

   6to4-PMT operation would be subject to the same security concerns as
   normal 6to4 operation as described in [RFC6169].  6to4-PMT is also
   not plainly perceptible by external hosts, and local entities appear
   as native IPv6 hosts to the external hosts.

6.  Acknowledgements

   Thanks to the following people for their textual contributions and/or
   guidance on 6to4 deployment considerations: Dan Wing, Wes George,
   Scott Beuker, JF Tremblay, John Brzozowski, Chris Metz, and Chris
   Donley.

   Additional thanks to the following for assisting with the coding and
   testing of 6to4-PMT: Marc Blanchet, John Cianfarani, Tom Jefferd, Nik
   Lavorato, Robert Hutcheon, and Ida Leung.

7.  References

7.1.  Normative References

   [RFC3056]  Carpenter, B. and K. Moore, "Connection of IPv6 Domains
              via IPv4 Clouds", RFC 3056, February 2001.

   [RFC3068]  Huitema, C., "An Anycast Prefix for 6to4 Relay Routers",
              RFC 3068, June 2001.

7.2.  Informative References

   [RFC1981]  McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery
              for IP version 6", RFC 1981, August 1996.

   [RFC3022]  Srisuresh, P. and K. Egevang, "Traditional IP Network
              Address Translator (Traditional NAT)", RFC 3022, January
              2001.

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

   [RFC4443]  Conta, A., Deering, S., and M. Gupta, Ed., "Internet
              Control Message Protocol (ICMPv6) for the Internet
              Protocol Version 6 (IPv6) Specification", RFC 4443, March
              2006.

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

   [RFC6169]  Krishnan, S., Thaler, D., and J. Hoagland, "Security
              Concerns with IP Tunneling", RFC 6169, April 2011.

   [RFC6264]  Jiang, S., Guo, D., and B. Carpenter, "An Incremental
              Carrier-Grade NAT (CGN) for IPv6 Transition", RFC 6264,
              June 2011.

   [RFC6296]  Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix
              Translation", RFC 6296, June 2011.

   [RFC6333]  Durand, A., Droms, R., Woodyatt, J., and Y. Lee,
              "Dual-Stack Lite Broadband Deployments Following IPv4
              Exhaustion", RFC 6333, August 2011.

   [RFC6343]  Carpenter, B., "Advisory Guidelines for 6to4 Deployment",
              RFC 6343, August 2011.

   [RFC6598]  Weil, J., Kuarsingh, V., Donley, C., Liljenstolpe, C., and
              M. Azinger, "IANA-Reserved IPv4 Prefix for Shared Address
              Space", BCP 153, RFC 6598, April 2012.

   [RFC6724]  Thaler, D., Ed., Draves, R., Matsumoto, A., and T. Chown,
              "Default Address Selection for Internet Protocol Version 6
              (IPv6)", RFC 6724, September 2012.

Authors' Addresses

   Victor Kuarsingh (editor)
   Rogers Communications
   8200 Dixie Road
   Brampton, Ontario L6T 0C1
   Canada

   EMail: victor.kuarsingh@gmail.com
   URI:   http://www.rogers.com


   Yiu L. Lee
   Comcast
   One Comcast Center
   Philadelphia, PA 19103
   U.S.A.

   EMail: yiu_lee@cable.comcast.com
   URI:   http://www.comcast.com


   Olivier Vautrin
   Juniper Networks
   1194 N Mathilda Avenue
   Sunnyvale, CA 94089
   U.S.A.

   EMail: olivier@juniper.net
   URI:   http://www.juniper.net