INTERNET DRAFT C. Huitema
<draft-huitema-6to4anycast-01.txt> Microsoft
Expires July 2, 2001 January 2, 2001

An anycast prefix for 6to4 relay routers

Status of this memo

Abstract

The operation of 6to4 routers require either that participation in IPv6 inter-domain routing, or that the routers be provisioned with a default route. This memo proposes a standard method to define the default route. It requires that IANA assign a "6to4 Relay anycast prefix" from which 6to4 routers can derive the static "6to4 anycast address". In order to enable efficient management of the "6to4 Relay anycast prefix" in IPv4 inter-domain routing, this memo also requests the reservation by IANA of a "6to4 Autonomous System ID." With this definition, the proposed scheme guarantees that 6to4 packets will be automatically routed to the nearest available router. It allows the managers of the 6to4 relay routers to control the sources authorized to use their resource. It makes it easy to set up a large number of 6to4 relay routers, thus enabling scalability.

1 Introduction

According to [6to4], there are two deployment options for a 6to4 routing domain, depending of whether or not the domain is using an IPv6 exterior routing protocol. If a routing protocol is used, then the 6to4 routers acquire routes to all existing IPv6 networks through the combination of EGP and IGP. If 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. This second case is typically used by small networks; for these networks, finding and configuring the default route is in practice a significant hurdle. In addition, even when the managers of these networks find an available route, this route often points to a router on the other side of the Internet, leading to very poor performances.

This memo proposes to reserve a "6to4 anycast address" in order to simplify the configuration of 6to4 routers. It also defines how this address will be used by 6to4 relay routers, how the corresponding "6to4 anycast prefix" will be advertised in the IGP and in the EGP. The memo requests the reservation by IANA of the "6to4 relay anycast prefix" and of a "6to4 Autonomous System ID."

2 Definitions

This memo uses the definitions introduced in [6to4], in particular the definition of a 6to4 router and a 6to4 Relay Router. It adds the definition of the 6to4 Relay anycast prefix,

2.1 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.

2.2 6to4 Relay Router

A 6to4 router configured to support transit routing between 6to4 addresses and native IPv6 addresses.

2.3 6to4 Relay anycast prefix

An IPv4 address prefix used to advertise an IPv4 route to an available 6to4 Relay Router, as defined in this memo.

The value of this prefix is x.x.x.0/nn [Length and value TBD IANA]

2.4 6to4 Relay anycast address

An IPv4 address used to reach the nearest 6to4 Relay Router, as defined in this memo.

The address corresponds to host number 1 in the 6to4 Relay anycast prefix, x.x.x.1. [Derived from the 6to4 Relay anycast prefix, TBD IANA]

2.5 6to4 IPv6 relay anycast address

The IPv6 address derived from the 6to4 Relay anycast address according to the rules defined in 6to4, using a null prefix and a null host identifier.

The value of the address is "2002:XXXX:XX01::". [Derived from the 6to4 Relay anycast address, TBD IANA]

2.6 6to4 Autonomous System ID

A 16-bit Autonomous system ID, for use in BGP in accordance to this memo.

The value of the 6to4 Autonomous System ID is YYYY. [TBD IANA]

3 Model, requirements

Operation of 6to4 routers in domains that don't run an IPv6 EGP requires that these routers be configured with a default route to the IPv6 Internet. This route will be expressed as a 6to4 address. The packets bound to this route will be encapsulated in IPv4 whose source will be an IPv4 address associated to the 6to4 router, and whose destination will be the IPv4 address that is extracted from the default route. We want to arrive at a model of operation in which the configuration is automatic.

It should also be easy to set up a large number of 6to4 relay routers, in order to cope with the demand. The discovery of the nearest relay router should be automatic; if a router fails, the traffic should be automatically redirected to the nearest available router. The managers of the 6to4 relay routers should be able to control the sources authorized to use their resource.

4 Description of the solution

4.1 Default route in the 6to4 routers

The 6to4 routers are configured with the default IPv6 route (::/0) pointing to the 6to4 IPv6 anycast address.

4.2 Behavior of 6to4 relay routers

The 6to4 relay routers that follow the specification of this memo shall advertise the 6to4 anycast prefix, using the IGP of their IPv4 autonomous system, as if it where a connection to an external network.

The 6to4 relay routers that advertise the 6to4 anycast prefix will receive packets bound to the 6to4 anycast address. They will relay these packets to the IPv6 Internet, as specified in [6to4].

4.3 Interaction with the EGP

If the managers of an IPv4 autonomous domain that includes 6to4 relay routers want to make these routers available to neighbor ASes, they will advertise reachability of the 6to4 anycast prefix. When this advertisement is done using BGP, the AS path leading to the

6to4 anycast prefix shall include the identifier of the local AS and the 6to4 Autonomous System ID.

The path to the 6to4 anycast prefix may be propagated using standard EGP procedures. The whole v6 network will appear to v4 as a single AS, with multiple peering points scattered over the whole Internet.

5 Discussion of the solution

The initial surfacing of the proposal in the NGTRANS working group helped is surface a number of issues, such as scaling concerns, the size of the address prefix, the need for an AS number, and concerns about risking to stay too long in a transition state.

5.1 Does it scale ?

With the proposed scheme, it is easy to first deploy a small number of relay routers, which will carry the limited 6to4 traffic during the initial phases of IPv6 deployment. The routes to these routers will be propagated according to standard peering agreements.

As the demand for IPv6 increases, we expect that more ISPs will deploy 6to4 relay routers. Standard IPv4 routing procedures will direct the traffic to the nearest relay router, assuring good performance.

5.2 Discovery and failover

The 6to4 routers send packets bound to the v6 Internet by tunneling them to the 6to4 anycast address. These packets will reach the closest 6to4 relay router provided by their ISP, or by the closest ISP according to inter-domain routing.

The routes to the relay routers will be propagated according to standard IPv4 routing rules. This ensures automatic discovery.

If a 6to4 relay router somehow breaks, or loose connectivity to the v6 Internet, it will cease to advertise reachability of the 6to4 anycast prefix. At that point, the local IGP will automatically compute a route towards the "next best" 6to4 relay router.

5.3 Access control

Only those ASes that run 6to4 relay routers and are willing to provide access to the v6 network announce a path to the 6to4 anycast prefix. They can use the existing structure of peering and transit agreements to control to whom they are willing to provide service, and possibly to charge for the service.

5.4 Why do we need a large prefix?

In theory, a single IP address, a.k.a. a /32 prefix, would be sufficient: all IGP, and even BGP, can carry routes that are arbitrarily specific. In practice, however, such routes are almost guaranteed not to work.

The size of the routing table is of great concern for the managers of Internet "default free" networks: they don't want to waste a routing entry, which is an important resource, for the sole benefit of a small number of Internet nodes. Many have put in place filters that automatically drop the routes that are too specific; most of these filters are expressed as a function of the length of the address prefix, such as "my network will not accept advertisements for a network that is smaller than a /19." The actual limit may vary from network to network, and also over time. We consider that a /16, from the old class B, would be very safe.

It could indeed be argued that using a large network is a waste of the precious addressing resource. However, this is a waste for the good cause of actually moving to IPv6, i.e. providing a real relief to the address exhaustion problem.

5.5 Why do we need a specific AS number?

Erroneous advertisements are a frequent source of errors in inter-domain routing. A misconfigured AS will advertise that it can reach some random network, divert the traffic, and effectively cut that network from some parts of the Internet. As a protection, many managers of border routers use databases to check the relation between the advertised network and the last hop in the AS path. If we use a specific AS to denote that "this is a path to IPv6", then we can enter the relation between that AS and the 6to4 access prefix in the databases used to check inter-domain routing.

5.6 Why not use Neighbor Discovery?

The ND alternative would be to use the anycast address for discovering the nearest relay, and then to build a "regular" association with this relay. It can be argued that using ND will be more a more standard way to do routing, that with a regular association failures will be easier to detect, and also that nailing the default connection to a specific 6to4 relay router will avoid any transient failure caused by the routing protocol. On the other hand, there are many arguments against this type of association.

Using ND to discover the next hop router may be in line with the generic IPv6 architecture, but is actually not in line with the specific software written for 6to4, in which the IPv4 next hop is derived algorithmically from the IPv6 destination address. In any case, maintaining ND associations for each 6to4 router that they serve will increase the load of 6to4 relay routers.

If the default connection is nailed to a specific 6to4 relay router,
the 6to4 router must actively monitor that router, detect failures,

and then try to find a secondary router, which will only be possible if routing tables have been updated. In contrast, with a pure anycast tunnelling solution, the packets will be routed to the functioning 6to4 relay router as soon as the routing tables have been updated.

After a failure of the nearest 6to4 relay router, if we use ND, the default connection will be nailed to a back-up 6to4 relay router, and will only be nailed back to the nearest 6to4 relay router if the back-up fails. In contrast, with a pure anycast tunnelling solution, the packets will always be routed to the nearest available 6to4 relay router.

Another line of argument is that the use of ND between routers is not entirely defined, and that using ND in this scenario is potentially more complex and more error-prone than just forwarding the packets to a well-known anycast address. In the absence of a clear-cut advantage for the ND solution, it is preferable to chose the simpler alternative, pure anycast tunnelling.

5.7 Will this slow down the move to IPv6 ?

Some have expressed a concern that, while the assignment of an anycast address to 6to4 access routers would make life a bit easier, it would also tend to leave things in a transition state in perpetuity. In fact, we believe that the opposite is true.

A condition for easy migration out of the "tunnelling" state is that it be easy to have connectivity to the "real" IPv6 network; this means that people trust that opting for a real IPv6 address will not somehow result in lower performances. So the anycast proposal actually ensures that we don't stay in a perpetual transition.

6 Future Work

Using a default route to reach the IPv6 Internet has a potential drawback: the chosen relay may not be on the most direct path to the target v6 address. In fact, on might argue that, in the early phase of deployment, a relay close to the 6to4 site would probably not be the site's ISP or the native destination's ISP... it would probably be some third party ISP's relay which would be used for transit and may have lousy connectivity. Using the relay closest to the native destination would more closely match the v4 route, and quite possibly provide a higher degree of reliability. A potential way to deal with this issue is to use a "redirection" procedure, by which the 6to4 router learns the most appropriate route for a specific destination. This is left for further study.

7 Security Considerations

The generic security risks of 6to4 tunneling and the appropriate protections are discussed in [6to4]. The anycast technique introduces an additional risk, that a rogue router or a rogue AS would introduce a bogus route to the 6to4 anycast prefix, and thus divert the traffic. IPv4 network manager have to guarantee the integrity of their routing to the 6to4 anycast prefix in much the same way that they guarantee the integrity of the generic v4 routing.

8 IANA Considerations

The purpose of this memo is to back a demand to IANA to allocate an IPv4 prefix dedicated to the 6to4 gateways to the native v6 Internet, and an autonomous system number dedicated to a pseudo-AS. The prefix length should be determined by the IANA; the prefix should be large enough to guarantee advertisement in the default-free BGP networks; a length of 16 will meet this requirement. This is a one time effort; there is no need for any recurring assignment after this stage.

9 Copyright

The following copyright notice is copied from RFC 2026 [Bradner, 1996], Section 10.4, and describes the applicable copyright for this document.

MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

10 Intellectual Property

11 Acknowledgements

The discussion presented here was triggered by a note that Brad Huntting sent to the NGTRANS and IPNG working groups. The note revived previous informal discussions, for which we have to acknowledge the members of the NGTRANS and IPNG working groups, in particular Scott Bradner, Randy Bush, Brian Carpenter, Steve Deering, Bob Fink, Tony Hain, Bill Manning, Keith Moore and Dave Thaler.

12 References

[6TO4] B. Carpenter, K. Moore. Connection of IPv6 Domains via IPv4 Clouds. Work in Progress.

13 Author's Addresses

Christian Huitema
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052-6399

Email: huitema@exchange.microsoft.com