Network Working Group
Request for Comments: 3068
Category: Standards Track

C. Huitema
Microsoft

June 2001

An Anycast Prefix for 6to4 Relay Routers

Status of this Memo

Copyright Notice

Abstract

This memo introduces 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 documents the reservation by IANA (Internet Assigned Numbers Authority) of the "6to4 relay anycast prefix."

1 Introduction

According to [RFC3056], there are two deployment options for a 6to4 routing domain, depending on 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 performance.

The operation of 6to4 routers requires either that the routers participate 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 introduces the IANA assigned "6to4 Relay anycast prefix" from which 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.

2 Definitions

This memo uses the definitions introduced in [RFC3056], in particular the definition of a 6to4 router and a 6to4 Relay Router. It adds the definition of the 6to4 Relay anycast prefix, 6to4 Relay anycast address, 6to4 IPv6 relay anycast address, and Equivalent IPv4 unicast address.

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 192.88.99.0/24

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, 192.88.99.1.

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:c058:6301::".

2.6 Equivalent IPv4 unicast address

A regular IPv4 address associated with a specific 6to4 Relay Router. Packets sent to that address are treated by the 6to4 Relay Router as if they had been sent to the 6to4 Relay anycast address.

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.

Anycast routing is known to cause operational issues: since the sending 6to4 router does not directly identify the specific 6to4 relay router to which it forwards the packets, it is hard to identify the responsible router in case of failure, in particular when the failure is transient or intermittent. Anycast solutions must thus include adequate monitoring of the routers performing the service, in order to promptly detect and correct failures, and also adequate fault isolation procedures, in order to find out the responsible element when needed, e.g., following a user's complaint.

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 [RFC3056].

Each 6to4 relay router that advertise the 6to4 anycast prefix MUST also provide an equivalent IPv4 unicast address. Packets sent to that unicast address will follow the same processing path as packets sent to the anycast address, i.e., be relayed to the IPv6 Internet.

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 initial AS path must contain the AS number of the announcing AS. The AS path should also include an indication of the actual router providing the service; there is a suggestion to perform this function by documenting the router's equivalent IPv4 address in the BGP aggregator attribute of the path; further work is needed on this point.

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 multi-homed network, with multiple access points scattered over the whole Internet.

4.4 Monitoring of the 6to4 relay routers

Any 6to4 relay router corresponding to this specification must include a monitoring function, to check that the 6to4 relay function is operational. The router must stop injecting the route leading to the 6to4 anycast prefix immediately if it detects that the relay function is not operational.

The equivalent IPv4 address may be used to check remotely that a specific router is operational, e.g., by tunneling a test IPv6 packet through the router's equivalent unicast IPv4 address. When a domain deploys several 6to4 relay routers, it is possible to build a centralized monitoring function by using the list of equivalent IPv4 addresses of these routers.

4.5 Fault isolation

When an error is reported, e.g., by a user, the domain manager should be able to find the specific 6to4 relay router that is causing the problem. The first step of fault isolation is to retrieve the equivalent unicast IPv4 address of the router used by the user. If the router is located within the domain, this information will have

to be retrieved from the IGP tables. If the service is obtained through a peering agreement with another domain, the information will be retrieved from the EGP data, e.g., the BGP path attributes.

The second step is obviously to perform connectivity tests using the equivalent unicast IPv4 address.

5 Discussion of the solution

The initial surfacing of the proposal in the NGTRANS working group helped us discover 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 loses 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. We expect that adequate monitoring tools will be used to guarantee timely discovery of connectivity losses.

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 IGPs, 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 /24." The actual limit may vary from network to network, and also over time.

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 Do we need a specific AS number?

A first version of this memo suggested the use of a specific AS number to designate a virtual AS containing all the 6to4 relay routers. The rationale was to facilitate the registration of the access point in databases such as the RADB routing registry [RADB]. Further analysis has shown that this was not required for practical operation.

5.6 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, one 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.

The practical operation of the 6to4 relay routers requires the development of monitoring and testing tools, and the elaboration of gradual management practices. While this document provides general guidelines for the design of tools and practice, we expect that the actual deployment will be guided by operational experience.

7 Security Considerations

The generic security risks of 6to4 tunneling and the appropriate protections are discussed in [RFC3056]. 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 managers 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 document the allocation by IANA of an IPv4 prefix dedicated to the 6to4 gateways to the native v6 Internet; there is no need for any recurring assignment.

9. Intellectual Property

The following notice is copied from RFC 2026 [Bradner, 1996], Section 10.4, and describes the position of the IETF concerning intellectual property claims made against this document.

10 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, Andrew Partan and Dave Thaler.

11 References

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

[RADB] Introducing the RADB. Merit Networks, http://www.radb.net/docs/intro.html.

12 Author's Address

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

EMail: huitema@microsoft.com

13 Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.