J. Bound
Digital Equipment Corp
P. Roque
Universidade de Lisboa

IPv6 Anycasting Service: Minimum requirements for end nodes

<draft-bound-anycast-00.txt>

Status of this memo

Abstract

This document proposes a minimum set of requirements for IPv6 hosts in order to achieve communication with nodes providing services through IPv6 anycast addresses. We present a mechanism that aims to allow TCP and UDP communication between hosts where the packet exchange is initiated through the usage of an anycast address, without requiring modifications to the general definitions of the transport protocols.

Table of Contents:

1. Introduction

IPv6 Anycast addresses [RFC-1883] allow a datagram to be addressed to a group of hosts with delivery to one and preferably only one semantic. This facility can be used to facilitate traffic path selection when a group identifies the set of routers of a particular traffic provider. Other possible uses of anycast addresses are to provide a "Host Anycasting Service" [RFC-1546], where a set of hosts can represent a particular service. While the particular mechanisms hosts can use to provide services via anycast addresses are still to be defined, this document attempts to define a minimum set of requirements that should be implemented in all IPv6 hosts in order to use those services.

The authors view a "Host Anycasting Service" as complementary, rather than orthogonal, to Service Discovery mechanisms [SVRLOC] since they can be used to provide lightweight service access without the need for previous configuration. For instance, a well-known site-local address can be used to communicate with a host that provides service discovery services.

While it is expected that the particular specifications regarding anycast address usage by application servers and routing are defined as extensions to IPv6 and companion protocols, the authors feel that mechanisms needed in every host should be defined before the massive deployment of IPv6 hosts.

The problems pertaining to the usage of anycast addresses for accessing application services can be clearly divided in three distinct components: procedures for hosts providing services via anycast addresses, routing, and procedures in hosts accessing services via anycast. This document focuses solely on the later problem, as the authors consider that it can be solved independently of the previous two.

2. Terminology and Definitions

IP

Internet Protocol Version 6 (IPv6). The terms IPv4 and IPv6 are used only in contexts where necessary to avoid ambiguity.

Anycast Address

An identifier for a set of interfaces (typically belonging to different nodes). A packet sent to an anycast address is delivered to one of the interfaces identified by that address (the "nearest" one, according to the routing protocols' measure of distance).

Communication

Any packet exchange between nodes that requires that the address of each node used in the exchange remain the same for the duration of the packet exchange. Examples are a TCP connection or UDP request/response.

3. Anycast address usage

Anycast addresses are restricted to be used as the destination address of a datagram. This requirement is imposed by necessity to determine the originating node of a datagram in error conditions. Current transport protocols [RFC-768] [RFC-793] rely, however, on the source address of the IP datagram to demultiplex incoming packets.

Independently of how the network delivers datagrams addressed to an anycast group, it's usage in normal communication depends on the ability of a host to accept a datagram originating from a distinct unicast address as a reply to a packet sent to an anycast address.

Also, as anycast addresses are syntacticly indistinguishable from unicast addresses, the client of a service provided via anycast should not need explicit knowledge of whenever a particular address is unicast or anycast, much less the particular group membership for a particular anycast address.

To fulfill the above stated goals, the authors propose the definition of a Destination Option, named Source Identification Option, to dynamically inform client hosts that a particular communication initiated through the use of an anycast address should proceed with the use of a unicast address of one of the anycast group members.

This option requires no processing from the network layer other than encoding and decoding the respective extension header and MUST be passed transparently from the network layer to the transport layer. The transport layer MAY then take in to account this information when demultiplexing datagrams. Section 5 of this document discusses in more detail the expected behavior of transport protocols when receiving a datagram with this option.

Although this document does not pretend to specify the mechanisms to be used by hosts providing a service through anycast addresses, we note that a reply to a datagram received for an anycast address will not be correctly interpreted if a Source Identification Option is not present.

4. Source Identification Option

The Source Identification Option provides a mechanism hosts can use to inform it's communications peers that datagram demultiplexing by transport protocols should be performed with respect to the identifier present in this option. This option is encoded in the Destination Options Extension Header of IP datagrams as option type TBD.

```
  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
                                 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                 |  Option Type  |  Option Len   |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                                                               |
  |                          Identifier                           |
  |                                                               |
  |                                                               |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Option Type

8-bit identifier of the type of option. The first three bits of the option are 000, indicating first that a node receiving the option may discard the option and still process the rest of the packet, and second that the option may not be modified enroute.

Option Length

8-bit unsigned integer. Length of the Option Data field of this option, in octets.

Identifier

128-bit IP address. The anycast address known to the receiver of the datagram as the destination address of a particular communication.

5. Receipt of Source Identification Option

As previously stated in section 3 of this document, a transport protocol MAY take in account the identifier received in a Source Identification Option for purposes of datagram demultiplexing.

TCP [RFC-793] communication depends on the knowledge of state information by communicating peers, initialized on a synchronization period referred to as the "three way handshake". In terms of access to a service provided via the use of anycast address, procedures must be provided to insure correct synchronization between the client and a member of the anycast group, and to maintain the same communication end-points during the duration of the connection.

An IP node performing a TCP active open sends a segment to the network addressed to the destination address of the connection. It then expects to receive a segment from this address confirming or rejecting the connection establishment request. When the destination address of the connection is an anycast address this condition cannot be met since the responding host may not send datagrams using an anycast address as source address in datagrams. In this scenario, the responding node should use the Source Identification Option, with the destination address on the received syn segment as an identifier, in order to inform the node performing the active open that the segment is related to this communication.

A TCP performing an active open MAY then use the IP address present on the Source Identification Option to demultiplex the incoming segment. If the segment causes TCP to proceed to SYN-RECEIVED or ESTABLISHED it MUST then consider that the destination address of the connection is the source address present on the received segment.

Note that for TCP, the receipt of a Source Identification Option is meaningful only when the segment refers to a connection on the SYN-SENT state. Otherwise, this option should be ignored by TCP. This will cause the received segment to be interpreted as a segment to a connection in the CLOSED state, assuming that no communication is taking place between the same address/port pairs.

Datagram exchanges using UDP [RFC-768] constitute a more problematic case. While UDP is itself connectionless, many applications using this transport protocol require state to be maintained. This implies that while some applications desire to communicate with any of the members of the anycast group, others can only tolerate anycast initiated communication requiring subsequent packets to be delivered to the same host.

Since the appropriate semantics of anycast address usage on UDP communication are application dependent, a UDP implementation should only take in to account the Source Identification Option when this behavior has been explicitly requested by the application. When such option is selected by the application incoming datagrams containing a Source Identification Option shall be demultiplexed and delivered to the application using the identifier contained in the option as the source address of the datagram. Otherwise, the Source Identification Option should be ignored by a UDP implementation.

As UDP already provides a means for determination of the originating node of a received datagram by applications, no further modifications are required to allow the use of this service with the desired semantics.

6. Issues for Further Consideration

Security considerations.

Receipt of a RST segment carried in a datagram containing a Source Identification Option.

According to [RFC-793], a segment containing a valid acknowledgement value and the RST bit on for a TCP connection in SYN-SENT state, will cause the connection to enter the CLOSED state. In the specific case of an active open to an anycast address, this abortive termination could be caused by a failure from one of the group members. The appropriate action to take in this case is an issue for further study.

Acknowledgements

We would like to thank Dan Harrington and Mike Shand who have provided comments and review of an earlier version of this work.

References

[RFC-768]
J. Postel, "User Datagram Protocol", STD-6, August 1980.

[RFC-793]
J. Postel, "Transmission Control Protocol", STD-7, September 1981.

[RFC-1546]
C. Partridge, T. Mendez, W. Milliken, "Host Anycasting Service", Informational Request for Comments, November 1993.

[RFC-1883]
S. Deering and R. Hinden, "Internet Protocol Version 6, (IPv6) Specification" Proposed Standard, December 1995.

[SVRLOC]
J. Veizades, E. Guttman, C. Perkins and S. Kaplan, "Service Location Protocol", Internet Draft, March 1996, <draft-ietf-svrloc-protocol-12.txt>

Authors' Address

Jim Bound
Digital Equipment Corporation
110 Spitbrook Road, ZKO3-3/U14
Nashua, NH 03062
Phone: (603) 881-0400
Email: bound@zk3.dec.com

Pedro Roque
Departamento de Inform'atica
Faculdade de Ciencias da Universidade de Lisboa
Campo Grande - Bloco C5
1700 Lisboa, Portugal
Email: roque@di.fc.ul.pt