A tool for checking Anycast Group Configuration (AGC)
draft-yao-opsawg-anycast-ping-00.txt

Status of This Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on May 18, 2010.

Abstract

Anycast is applied in many protocols to provide automatically fault recovery (high availability) and scalability, i.e. Anycast RP in PIM, Cooperative Home Agent and Cooperative Foreign Agent in Mobile IP.

When the Operator’s engineer completing configuring an anycast group, they need a tool to check whether each group member can be reached by the anycast address (instead of by each member’s unicast address).

When the anycast group members are changed, e.g. there are some new members added on, there are some old members deleted from the group, a tool is needed to check whether the changed group/member can be reached by the anycast address as it should be.

When an anycast group have been canceled by configuration, there need a tool to check whether the anycast group configuration have been wiped off from all group members.
The existing method is to use existent PING command to ping the anycast address. This existing method can only prove whether one member in the anycast group instead of all members of the anycast group identified by the anycast address can be reached by that anycast address. Hence it cannot satisfy above problems.

This document provides a solution to prove that each member of a given anycast group can be arrived by that anycast address.

Table of Contents

1. Introduction .................................................. 3
2. Overview ...................................................... 4
3. Protocol Operation .............................................. 4
4. IANA Considerations ........................................... 10
5. Security Considerations ....................................... 10
6. References ..................................................... 10

Authors’ Addresses .............................................. 11
Intellectual Property and Copyright Statements ................. 11
1. Introduction

Anycast is applied in many protocols to provide automatically fault recovery (high availability) and scalability, i.e. Anycast RP in PIM, Cooperative Home Agent and Cooperative Foreign Agent in Mobile IP. When operators deploying any anycast services, it has to group several IP nodes into an Anycast group first. Hence the "anycast group configuration information" has to be configured on each member of an anycast group. The "anycast group configuration information" includes at least the anycast address of the group and all member’s unicast IP addresses at least. By this "anycast group configuration information", each Anycast group member know it is a member of that Anycast group, the group’s IP address is the anycast address, and it can communicate with every other member by the unicast IP address of them.

When the network administrator or Operation & Maintaining engineer completing configuring an anycast group, they need a tool to check whether each group member can be reached by the anycast address (instead of by each member’s unicast address.)

When the anycast group members are changed, e.g. there are some new members added on, there are some old members deleted form the group, a tool is needed by the Operator’s engineer to check whether the changed group/member can be reached by the anycast address as it should be.

When an anycast group have been canceled by configuration, there need a tool to check whether the anycast group configuration have been wiped off from all group members.

The existing method is to use existent PING command to ping the anycast address, and cannot staisfy above problems.

This document provides a solution for these requirements.
2. Overview

The basic idea is as follows:

1) Anycast Ping Still works by exchanging ICMP Informational Messages (Echo Request Message and Echo Reply Message). To achieve Anycast Ping, the responding member adds the all member’s unicast IP addresses in the Responding Echo Reply Message of the invoking Echo Request message. The Ping host can get known that responding member’s unicast address by checking the source IP address of received Responding Echo Reply message.

2) For each anycast group member, find out at least one "transfer point": an IP host node or router node, anycast traffic sourced from which and destined to the anycast group will arrive at that anycast group member.

3) The Ping host sends N (N>=1) Echo Request Messages to each "transfer point" with a Type 0 Routing Header (IPv6)/ Loose Source option (IPv4) in the IP packet carrying the message. The anycast address is included in the routing header/ Loose Source option each time. Then the ping host waits to receive the Echo Reply Messages from different anycast member and decide each member’s reachability.

4) If the "transfer point" are in the administration domain of the Operator, the Operator can also remotely login the "transfer point", Ping the anycast address and observe the reachability of the anycast address by whether there is Echo Reply messages received on each "transfer point".

3. Protocol Operation

This section describes the Anycast Ping operations.

By "Anycast Ping", when a user wants to know the reachability of each member of an anycast group, it first "Anycast Ping" the anycast address from a ping host. here "Anycast Ping" is a command which is used to ping anycast address.

For a given anycast address, when a command "Anycast Ping" is launched in a ping host, the following events will happen:

1). Anycast configuration information check:

The Ping host sends N(N>=1; Can be configured according to policy. Default value is 5—the same as unicast ping) Echo Request Message to the destination anycast address. The responding member
reports "Anycast configuration information" by adding that information configured on it into "Anycast Configuration Information Option" (described in section 3.1) of the corresponding Echo Reply Message.

Hence the Ping host can know all members in the group so that it can further check whether other members can be reached by the anycast address. The Ping host can get known that responding member’s unicast address by checking the source IP address of received Responding Echo Reply message. For this situation, go to step 2).

If there is no any valid Echo Reply Message received after sending Echo Request Messages, that means there is no anycast group configured in the network domain for the given anycast address. For this situation, no further process, indication should be displayed to the user: "no such anycast group configured yet!"

If there is valid Echo Reply Message without "Anycast Configuration Information Option" received after sending Echo Request Messages, that means the given anycast group ia actually a unicast address. For this situation, no further process, indication should be displayed to the user: "Possible unicast address!"

2). Find "transfer point" for each group member

After geting known the anycast address configuration information by step 1). The ping host call the "Fingding out "transfer point"" service (i.e. can be a SNMP service provided to authorized users) to get at least a "transfer point" for each group member.

If there is no any "transfer point" found out for a given member, the ping host should display to the user: "fail to find transfer point for member X!", and continue to find out "transfer point" for other members.

If there is no any "transfer point" for each other member (means any member of the given anycast group instead of the one responding in step 1) of the given anycast group, display "all other members are unreachable!" and end "Anycast Ping"

For members there exists at least one "transfer point", go to step 3).

3). Check reachability of each member having "transfer point"

Either of the two ways can be used by the ping host:
A). Remote login

The "transfer points" should be in the administration domain of the Operator. The ping host should also remotely login the transfer points, and Ping the anycast address and display the reachability of the anycast address by whether there is Echo Reply messages received on each "transfer point".

B). Using routing option or routing header

The Ping host sends N (N>=1; Can be configured according to policy. Default value is 5—the same as unicast ping) Echo Request Message to each "transfer point" with a Type 0 Routing Header (IPv6)/ Loose Source option (IPv4) in the IP packet carrying the message. The anycast address is included in the routing header/ Loose Source option each time. Then the ping host waits to receive the Echo Reply Messages from different anycast member and decide each member’s reachability.

For IPv4, Loose Source option has not been implemented and opened to user widely in routers for security consideration and other reasons. For IPv6, there is similar problem though Type 0 Routing Header is mandatory in IPv6 specified in [RFC2460]. Hence, above way B) may be optional.

3.1. ICMP Echo Reply Message with Anycast Configuration Information Option

In "Anycast Ping", what different from the ICMP Echo Reply Message in "unicast Ping" is that an Anycast Configuration Information Option is included in this "Anycast Ping" ICMP Echo Reply Message. The Option includes: all member’s unicast IP addresses and the anycast address of the group.

For Both ICMPv4[RFC792] and ICMPv6[RFC4433], The position of Anycast Configuration Information Option in Echo Reply message is as follows:
<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Anycast Configuration Information Option

Data ...

---

For ICMPv4, the Format of Anycast Configuration Information Option:

<table>
<thead>
<tr>
<th>Type=0x01</th>
<th>Length</th>
<th>Reserved</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of Members (N) | Reserved

Member 1's IP unicast address

Member 2's IP unicast address

... 

Member N's IP unicast address

Type:

Means this option is an "Anycast Configuration Information Option"

Length:

Contains the length of the overall length of all fields in this option except the length of "Type", and "Length", as well as "Reserved", in units of 32-bit words.

Reserved:

The Reserved fields are set to zero on transmission, and ignored on reception.
Anycast Address:
The IPv4 anycast address of the anycast group.

Number of Members (N):
The Number of Group members (N) field specifies how many Group members are present in this Option.

The Member i’s IP unicast Address:
These fields are a vector of n Group Member’s IPv4 unicast addresses, where n is the value in this Option’s Number of Sources (N) field.

For ICMPv6, the Format of Anycast Configuration Information Option:
Means this option is an "Anycast Configuration Information Option"

Length:
Contains the length of the overall length of all fields in this
option except the length of "Type", and "Length", as well as
"Reserved", in units of 32-bit words.

Reserved:
The Reserved fields are set to zero on transmission, and ignored on
reception.

Anycast Address:
The IPv6 anycast address of the anycast group.

Number of Members (N):
The Number of Group members (N) field specifies how many Group
members are present in this Option.

The Member i’s IP unicast Address:
These fields are a vector of n Group Member’s IPv6 unicast addresses
, where n is the value in this Option’s Number of Sources (N) field.

3.2.  How to find out the "transfer point"

Two ways are provided in this document for finding out a "Transfer
point" for a given anycast group member.

The first way is:
Operators usually know the topology of their network. For each
anycast group member, they can find out at least one IP host node
or router node, anycast traffic sourced from which and destined to
the anycast group will arrive at that pointed anycast group member.
This can be done by using some operation and maintenance tools such
as Alcatel-Lucent’s 5650 CPAM (Control Panel Assurance Manager).

The second way is:
Based on [RFC4292], A network administrator can know current
Forwarding Information Base (FIB) of each router in his
administration domain. By accessing "inetCidrRouteTable" Object in
each Router’s MIB (Management Information Base), A network
administrator can know current FIB of each router in his domain. The
access can be achieved by SNMP (Simple Network Management Protocol).
SNMP is widely supported in current routers.
Hence we can start search the set of transfer points from the given group member. i.e. member A.

First, we find out all routers connected to each Interface of member A. For each router (Router1i, i=1,...,M) connected to interface 1 of member A, we look up its FIB with the key word (Anycast Address, IP Address of interface 1). the Anycast Address is done longest-prefix matching with all the destination prefixes of member 1’s FIB, and the IP address of interface is matched with "next-hop" of FIB. If there exists a matched FIB entry, both its destination prefix and next-hop match with the key word, that means IP packets destined to this anycast address can arrive at member A through this router, and all routers connected to this router1i should be examined likewise. If there is no any matched FIB entry, that means IP packets destined to this anycast address cannot arrive at member A through this router, next router should be tried. If there is no any router connected to member 1 can forward IP packets destined to the anycast address to member 1, that means member 1 can not be arrived by this anycast address.

By this way, at least one leaf router can be found out to be the "transfer point".

4. IANA Considerations
The "Anycast Configuration Information Option" defined in this document contains the "Type" field that carry values assigned from IANA-managed name spaces.

5. Security Considerations
This document deals with ICMP messages in ICMPv4, ICMPv6, SNMP and keeps consistent with that of [RFC792] and [RFC4433], and [RFC3411] in future evolution.

6. References


Authors’ Addresses

Chunyan Yao
Research and innovation center in Shanghai, Alcatel-Lucent.
Email: Chunyan.Yao@alcatel-sbell.com.cn
Phone: 86-21-50554550 extension 7250

Copyright Notice

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents in effect on the date of publication of this document (http://trustee.ietf.org/license-info). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.