S. Ata
Osaka City University
M. Hashimoto
Osaka Univerisity
H. Kitamura
NEC Corporation
M. Murata
Osaka University

Internet-Draft
Expires: January 19, 2005
January 20, 2006

Mobile IPv6-based Global Anycasting
draft-ata-anycast-mip6-03

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of 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/1id-abstracts.html

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

This Internet-Draft will expire on July 19, 2006.

Copyright Notice

Copyright (C) The Internet Society (2006).

Abstract

Today, the use of anycasting is quite limited. In this document we explain a new network architecture for global anycasting to solve (1) in practical use and able to realize with existing technology easily, (2) about support for stateful communications keeping a session

information at the end nodes such as TCP. The architecture is based on MIPv6 because there are many similarities between MIPv6 and global anycast.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119. Other terms in this document are defined in [ANYTERM]

1. Introduction

Today, anycast is not widely used in the IPv6 network while various uses are expected. There are several issues to be solved. One of the main reasons is because the current form of anycast cannot establish the communication keeping a session. It is because the use of anycast address as a source address of a packet. It largely limits in particular a use range of anycast that anycast cannot be used for TCP connections. In addition, there is no practical mechanism to realize global anycasting in which nodes having the same anycast address belong to different networks.

In this document we explain a new network architecture for global anycasting to solve the above-mentioned problems. The document particularly aims for following: (1) The mechanism is practical use and able to realize with existing technology easily, (2) The mechanism supports stateful communications keeping a session information at the end nodes such as TCP.

As realization technique of global anycast, there are roughly two ways; controlling anycast packets by a router or by an end node.

To support global anycast by a router a new anycast routing protocol which treats anycast packets is deployed on a router.

This technique can realize global anycasting without any extension in an end node, however, realization is not easy because many of routers in the Internet should implement the new anycast routing protocol, to receive the benefit of anycasting.

On the other hand, as for the realization by an end node, the end node which want to establish an anycast communication has a new mechanism to decide the destination of an anycast packet. This technique can be realized only by a change of an end node without adding a revision to a router, but there is a problem that an anycast initiator (i.e., mostly a client) has to collect information about anycast receivers (i.e., servers including mirrors) beforehand, to decide a direction for transferring the anycast packet.

To solve this problem, therefore, we consider the model that the third end node (called agent) except AIs and ARs manage all information about the anycast address intensively, and relay anycast packets between AI and ARs.

The merit of this model is that no revision of a router is needed, and it is able to realize global anycast easily. In addition, as for the client, there is an advantage not to have to collect information of servers beforehand: forwarding anycast packets to agent is enough.

As a result of having examined IPv6 existing communication models to be able to employ, we found that Mobile IPv6 (MIPv6), which is the protocol supporting mobile terminals for IPv6, is designed for the above model. There are many similarities between MIPv6 and global anycasting.

Therefore, in the document we show how to realize the global anycasting by applying MIPv6 mechanism. We first explain similalities and differences between MIPv6 and global anycasting. Then we describe the overview of a MIPv6-based global anycasting scheme. Through this document we refer the MIPv6-based global anycast as MGA.

2. Similarlities and Differences between MIPv6 and Global Anycast

2.1 Addressing Issue

In MIPv6, a mobile node (MN) has two type of addresses: HoA and CoA. HoA is always fixed for each MN while CoA may vary based on a place where the MN is belonging to. To continue the communication between MN and a correspondence node (CN) seamlessly, the home agent (HA) manages a binding cache (BC) to map HoA and CoA of MN. It is because the CoA of MN may change during the communication when the MN moves to another network.

Global anycast has a similar characteristic. An anycast responder (AR) has two type of addresses: anycast address (AA) and peer unicast address (PUA). AA is always fixed for all ARs while PUA is different for each AR. Anycast packets sent from the same anycast initiator (AI) may be forwarded different ARs due to the network condition. It means that a unicast address corresponding to anycast address varies with the network situation.

Therefore, it is useful to apply the mechanism of HA (in particular the function managing BC) in MIPv6 to MGA architecture.

2.2 Restriction of the Use as Source Address

In MIPv6 there is a limitation that MN cannot use its HoA for a source address of a packet when the MN does not exist on its own home link.

It is because in a network outside the home link the HoA is not valid as a subnet prefix for the network. If the MN sends a packet with specifying the HoA as the source address, the packet may be dropped by the edge router of the network due to the security reason (i.e., the router recognizes the packet is spoofed).

For this problem, MIPv6 utilizes a function of message tunneling for the communication between the HA and both MN and CN. In the case of the route optimization, the MN specifies the CoA as the source address of the packet, and adds an home address option (one of destination options in IPv6) to set the HoA of MN. The stateful communication that maintained a session is thus enabled.

Global anycast has also a similar limitation. Because it is not guaranteed that multiple anycast packet addressed to the same anycast address are delivered to the same node, the use of anycast address as the source address of the packet is also prohibited.

Therefore, the same approach on MIPv6 (utilizing tunneling and an address option) is useful for MGA.

2.3 Keeping Stateful Communication

As shown above, we can apply the mechanism of MIPv6 for the similar property on global anycasting. However, there is a crucially different point in MIPv6 and MGA. It is the factor that the binding update (BU) occurs.

In MIPv6 a binding update occurs when the MN moves to the different network (i.e., a timing when the CoA of MN has been changed). However, even if the CoA is changed the communication partner is always the same.

On the other hand, in MGA, the event of binding update means that the appropriate anycast responder has been changed. That is, after the binding update an anycast packet is delivered to a different AR from before. This is a clear difference from MIPv6. In MGA, different PUAs stand for different ARs. The communication partner has been changed by the event of binding update.

This difference gives a serious influence in stateful communications. If the binding update is happen during the communication such as TCP, following packets of TCP stream after the binding update are forwarded to the different node, the TCP connection is thus destroyed at this time.

For this reason, in MGA, it is strongly required that any binding updates for the anycast address which is used on the stateful communication should be rejected.

MIPv6 provides a route optimization as an optional function, however, from above reason MGA suggests the same function of a route optimation to be a mandatory function. Agent-relay communications should be used for non-stateful or one round of communication only. Stateful communications such as TCP must perform a route optimization prior to establishing the communication. The communication should be done with the direct connection between the AI and the selected AR (i.e., the node received the anycast packet forwarded by the agent).

3. Architecture of MIPv6-based Global Anycasting

3.1 Architectural Overview

In anycast, an anycast address assigned for a specific service may be given to multiple anycast receivers which provides the same service. An anycast initiator (a node which intends to communicate with a anycast receiver with an anycast address) can communicate with a suitable (appropriate) AR. The suitable AR may vary according to the situation.

In the anycast communication the anycast initiator does not specify the peer unicast address of the anycast receiver, but do the anycast address for the receivers. So it is necessary for correspondence between anycast address nad peer unicast address.

Therefore MGA refers to the mechanism on MIPv6. In MGA Home Anycast Responder (HAR) manages Anycast Binding Cache, which is the correspondence information of anycast address and peer unicast address.

Figure 1 shows the overview of MIPv6-based Global Anycast architecture. There are three types of nodes: one anycast initiator (AI), three anycast responders having the same anycast address AA (AR1, AR2, AR3), and one Home Anycast Responder. To receive the anycast packet addressed to AA, HAR has a unicast address AA. Peer unicast addresses of AR1, AR2, AR3 are PUA1, PUA2, PUA3, respectively.

HAR has an anycast binding cache ABC. In this figure, the correspondence unicast address for anycast address AA is PUA2.

First, AI sends an anycast packet by specifying the anycast address AA as its destination. Since an anycast address is assigned from the unicast address, the anycast packet is sent to the subnet which HAR belongs to by the unicast routing. HAR receives the anycast packet and checks its anycast binding cache. In this case, the HAR found PUA2 is suitable for forwarding the anycast packet. Then HAR relays the anycast packet to AR2 by tunneling, and AR2 receives the anycast packet.

The communication from AR2 to AI, AR2 first encapsulates the anycast packet (src: AA, dst: AI), and sends to HAR by reverse tunneling function. HAR decapsulates it and forwards to AI.

```
                                +-----+ Anycast Addr = AA
             Unicast      ......| AR1 | Peer Unicast Addr = PUA1
              Addres = AA :     +-----+
    +----+          +------+         +-----+ Anycast Addr = AA
    | AI |----------| HAR  |.........| AR2 | Peer Unicast Addr = PUA2
    +----+          +------+         +-----+
   Unicast    +-ABC-----+ :     +-----+ Anycast Addr = AA
   Addr = AI  | AA:PUA2 | :.....| AR3 | Peer Unicast Addr = PUA2
              +---------+       +-----+
```

Fig. 1 Overview of MIPv6-based Anycast Architecture
(... : tunneling)

3.2 Table Fix for Stateful Communication

As described in 2.3, above architecture works well when the communication is non-stateful or one round. However, if the HAR's anycast binding cache is updated by the anycast binding update from another AR during the stateful communication, further connection is no longer effective because the correspondence node for the anycast address is changed. Therefore, in MGA the same function as route optimization in MIPv6 is mandatory, which is called "Table Fix".

To establish the stateful communication anycast initiator must perform Table Fix immediately to communicate with anycast responder directly prior to begin the stateful communication.

The mechanism of table fix is almost similar to the route optimization in MIPv6. More specifically, when an anycast receiver receives an anycast packet from an anycast initiator, the receiver do a return routability test to the anycast initiator. After the test, the anycast receiver sends an anycast binding update message to the anycast initiator. The anycast initiator then updates its anycast binding cache based on the receiving anycast binding update. After table fix, communication is done with the peer unicast address of the anycast receiver. As in MIPv6, the anycast address is also embeded by a kind of destination option (called Anycast Address Option) in IPv6.

4. Issues to be Resolved

* Authentication mechanism to clarify anycast binding update
* Mechanism for handling multiple Anycast Receivers simulteniously in HAR
* Load distribution for HAR

5. Security Considerations

Because this model utilizes the similar mechanism in MIPv6, the model

is needed to consider security issues shown in [MIPv6].

6. References

6.1 Normative References

[RFC 2373] Hinden, R., and Deering, S., "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[ANALYSIS] Hagino, J., and Ettikan, K., "An Analysis of IPv6 Anycast", work in progress.

[MOBILEIP] Johnson, D., Perkins, C., and Arkko, J., "Mobility Support in IPv6", work in progress.

[RFC 2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997.

[ANYTERM] Doi, S., Kahlil, I., Ata, S., Kitamura, H., Murata, M., "Anycast Terminology/Functionality Definition," work in progress.

Authors' Addresses

Shingo Ata
Osaka City University
3-3-138, Sugimoto, Sumiyoshi-ku
Osaka, 558-8585
Japan

Phone: +81-6-6605-2191
Fax: +81-6-6690-5382
EMail: ata@info.eng.osaka-cu.ac.jp

Masakazu Hashimoto
Osaka University
1-5 Yamadaoka, Suita
Suita-shi, Osaka 565-0871
Japan

Phone: +81-6-6879-4543
Fax: +81-6-6879-4544
EMail: msk-hasi@ist.osaka-u.ac.jp

Hiroshi Kitamura
NEC Corporation
(Igarashi Building 4F) 11-5, Shibaura 2-Chome
Minato-ku, Tokyo 108-8557
Japan

Phone: +81-3-5476-1071
Fax: +81-3-5476-1005
EMail: kitamura@da.jp.nec.com

Masayuki Murata
Osaka University
1-5 Yamadaoka, Suita
Suita-shi, Osaka 565-0871
Japan

Phone: +81-6-6879-4543
Fax: +81-6-6879-4544
EMail: murata@ist.osaka-u.ac.jp

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (C) The Internet Society (2006). This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

Acknowledgment

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