Protocol Design for Anycast Communication in IPv6 Network

Satoshi Doi*, Shingo Ata†, Hiroshi Kitamura‡, Masayuki Murata§ and Hideo Miyahara*

*Graduate School of Information Science and Technology,
Osaka University
1–3 Machikaneyama, Toyonaka, Osaka 560–8531, Japan
Email: {s-doi,miyahara}@ist.osaka-u.ac.jp
†Graduate School of Engineering, Osaka City University
3–3–138 Sugimoto, Sumiyoshi–ku, Osaka 558–8585, Japan
Email: ata@info.eng.osaka-cu.ac.jp

‡Development Laboratories, NEC Corporation
2–11–5 Shibaura, Minato–ku, Tokyo 108–8557, Japan
Email: kitamura@da.jp.nec.com
§Cybermedia Center, Osaka University
1–30 Machikaneyama, Toyonaka, Osaka 560–0043, Japan
Email: murata@cmc.osaka-u.ac.jp

***Abstract*— Although anycast communication supports service–oriented addresses many of its current definitions in IPv6 are unclear. Furthermore, since there are no protocol standards or even consensus on routing control, inter–segment anycast communications are not yet available. In this paper, we first discuss these problems and solutions. Based on our findings, we present the Anycast Address Resolving Protocol (AARP) to establish TCP connections with a specific anycast address and propose a routing protocol for inter–segment anycasts. Our proposed architecture makes anycast addresses more useful without (or at most minimum) modifications/extensions to existing applications and/or upper–layer protocols.**

I. Introduction

A. Overview of Anycast Communications

Anycast [1] is one of the new IPv6 (IP version 6 [2]) features that supports service-oriented address assignments in IPv6 networks. An anycast address is not determined by the location of the node, but by the type of service offered at the node. In anycast communications, the client can automatically obtain the appropriate node corresponding to a specific service without knowledge of the location of the server.

Anycast is defined in Internet Protocol version 6 [2] and its addressing architecture [1] has two other types of IP addresses, unicast and multicast. Table I summarizes the communication forms for these addresses. A unicast address is a unique identifier for each network interface, and multiple interfaces must not be assigned the same unicast address. Packets with the same destination address are sent to the same node. A multicast address, on the other hand, is assigned to a group of nodes, i.e., all group members have the same multicast address and packets for this address are sent to all members simultaneously. Like a multicast address, a single anycast address is assigned to multiple nodes (called *anycast membership*), but unlike multicasting, only one member of the assigned anycast address communicates with the originator at a time. Figure 1 has an example of anycast communication. There are three nodes associated with the anycast address A_{any}. When the source node sends a packet, where the destination address is A_{any}, the packet is sent to one of three nodes (X_{uni} in this figure), not to all hosts. The advantage of anycasting is that the source node can receive a specific service without knowledge about current conditions in service nodes and/or networks. When host X_{uni} goes down, the packet for A_{any} can be sent to another host (Y_{uni} or Z_{uni}) (Fig. 1). How appropriately the destination node is chosen from anycast membership depends on the anycast routing protocol.

TABLE I
IPv6 ADDRESS TYPES.

	unicast	multicast	anycast
communication form	point to point	point to multipoint	point to point
target of address	node	group	service type
number of membership	single	multiple	multiple
roles in C/S model	both	client	server

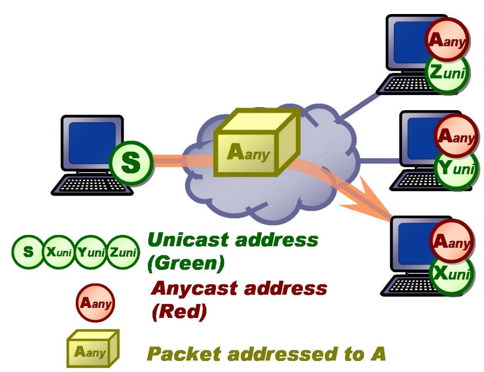

Fig. 1. Anycast communication

The main idea behind anycast communication can be found in the separation of the logical service identifier from the physical node equipment. The anycast address is assigned on a type-of-service basis and enables a service to act as a *logical node* appropriate for the service. In [3], we describe what applications are suitable for anycast communication. For example, the following applications are subject to the use of anycasting.

1) Dynamic Node Selection
By performing the routing control appropriately, the sender node can communicate with the optimal node (chosen from multiple anycast nodes) by simply specifying the anycast address. For example, if we assign the same anycast address to the WWW server and its

mirror sites, end users can access the site nearest to their location by only specifying its anycast address.

2) Well-known Anycast Address for Specific Services
By defining and assigning the well-known anycast address to widely used applications (e.g., domain name services, proxy services, etc.), the user can use these services without setting the address of the server.

Anycast communication has the interesting features noted above. However, the current use of anycast addresses is quite limited. One of the main reasons is that there are many points in the current definition of anycasting that are still unclear. Moreover, anycast communication has problems with its protocol specifications and its routing mechanism.

The first problem is that anycast communication cannot guarantee that multiple packets sent to the same anycast address will reach the same destination, while most currently available applications assume there is a single destination peer during the communication. That is why the anycast address cannot be directly used to establish a TCP connection [4]. To solve this problem, it is desirable to resolve the anycast address into the unicast address prior to beginning the communication. To realize this address resolution, we have designed a new architecture, called AARP (Anycast Address Resolving Protocol) in [5]. Its brief mechanism is described in the next section.

The second problem is that the current anycast standard does not define its routing protocol. There are several challenging issues that need to be resolved in designing the anycast routing protocols.

1) Scalability issue
A conventional IP router aggregates multiple entries that have the same prefix for the destination address (or network), and the same output interface (i.e., the same direction) to reduce the size of the routing table. However, the routing entries for anycast addresses cannot be aggregated because anycast membership locations are widespread regardless of their prefix. Hence, routing entries for anycast addresses should individually be stored on the router. It is easy to imagine explosions in routing tables as anycast addresses become more widely used.
2) Criteria for selecting anycast membership
Anycast routing is required to transfer an anycast packet to an *appropriate* anycast node, but the meaning of *appropriate* needs to differ among applications. For example, if an application requires a faster response, the propagation delay between the source node and anycast node is extremely important, i.e., the nearest node for anycast membership should be chosen. The criteria for anycast routing strongly affects anycast communication capabilities.
3) Security issues
Maintaining anycast membership is important and the easiest way for a host that intends membership is for it to simply advertise the routing entry for the associated anycast address to the router. However, such an approach can sometimes lead to serious security problems in that the anycast host can freely add or delete anycast entries in the routing table.

There have been several proposals for an anycast routing protocol (e.g., [6], [7]), but to our knowledge, none of these conforms to IPv6 anycast specifications and anycast addresses are allocated in their own address space, which is different from the unicast address space. However, the routing protocol we proposed in Section III allows the same space to be used for unicast and anycast addresses.

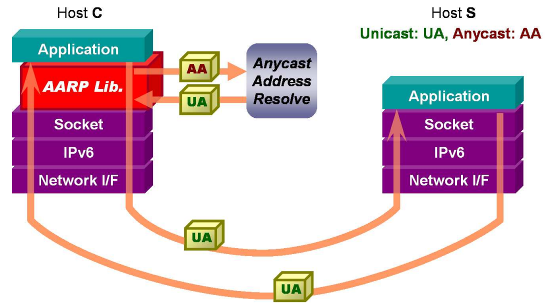

Fig. 2. Protocol stack of AARP.

The rest of this paper is organized as follows. The next section presents our approach called AARP (Anycast Address Resolving Protocol) that enables TCP communications utilizing anycast addresses. Section III discusses our routing scheme for inter–segment anycasting, where scalability and deployment are taken into account. Finally, Section IV has the concluding remarks.

II. Anycast Address Resolving Protocol

Our proposal fills a gap between anycast and upper–layer protocols like TCP and UDP without the need to modify applications or protocols. More specifically, the task of the AARP (Anycast Address Resolving Protocol) is to resolve the anycast address specified by the application into the corresponding unicast address [5].

Figure 2 shows the protocol stack for anycast communication with the AARP. The AARP is implemented as a kind of DLL (Dynamic Linkable Library) that overwrites the original (i.e., the provided operating system) APIs (Application Programming Interfaces). We call this library the AARP Library. The AARP Library (AARP Lib in Figure 2) provides the same set of APIs as the original IPv6 socket APIs, and hooks them to resolve anycast addresses. It converts an anycast address into its corresponding unicast address prior to calling the original APIs. The anycast address is only used in the application layer and the AARP Library layer. Layers below the AARP Library are not aware of the anycast address, and only handle the translated unicast address.

III. Design of Inter-Segment Anycast Routing Protocol

A. Design Choices

The design choices we made in our anycast routing protocol are as follows.

1) We allow unicast and anycast addresses within the same space and to do this we chose a *seed node* from anycast membership before assigning an anycast address. We then established the anycast address of membership to be the unicast address of the *seed node*. The anycast router forwards an anycast packet to an *appropriate* node within the anycast membership. However, the unicast router only tries to forward the anycast packet to the *seed node*. An anycast packet departing from an arbitrary node is at the very least sent to the seed node. Packet reachability is thus guaranteed.
2) We support the gradual deployment of anycasting and in our architecture the protocol works correctly and advantages can be enjoyed even if the route between

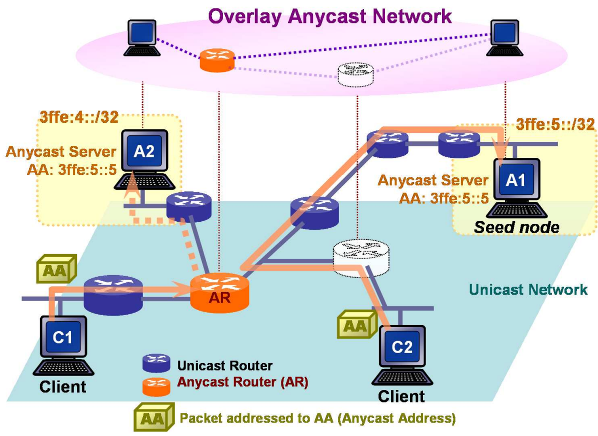

Fig. 3. Proposed architecture.

the sender and seed node has only one anycast router. Its impact will increase as more anycast routers are deployed.

3) We adopt an approach that modifies multicast routing to the anycast routing protocol to reduce the complexity of implementation, since they have many similarities. For example, they have common membership management and routing table construction procedures.

B. Proposed Architecture

Figure 3 is an overview of the routing architecture we propose and there are two types of routing topologies. The *unicast network* is the existing network topology where both unicast and anycast packets are forwarded on the basis of a unicast address. The *anycast network* is a logical overlay topology, where anycast-aware routers (called *anycast routers*) are connected to one another and only anycast packets are forwarded by treating their addresses as anycast addresses.

In an anycast network, nodes are not physically (i.e., directly) connected, but are connected via various kinds of logical peer-to-peer connections (e.g., virtual path, tunneling, or encapsulation). An anycast router is upper-compatible and does anycast routing functions and has the capabilities of unicast routers. An anycast router has an extra routing table (called an *anycast routing table*) to handle anycast addresses. An anycast routing table consists of at least (*anycast address* and *next anycast router's address*) pairs. When a packet arrives at the anycast router, it first checks the anycast routing table to find an entry regarding the destination address of the packet. If it finds this, the packet is treated as an *anycast packet* and forwarded to the next anycast router according to the anycast routing table. Otherwise, it is forwarded through the unicast routing mechanism.

Figure 3 has an example of anycast routing where we have assumed that the node selection criterion is the number of hops. A smaller count is more appropriate here. In Fig. 3, the short blue or orange cylinders represent routers and the one labelled "AR" is an anycast router. The short blue cylinder is a unicast router. There are two anycast members for the anycast address `3ffe:5::5`. Note here that `3ffe:5::5` is also the unicast address of anycast server A1. Here, node A1 is the *seed node* of anycast membership for `3ffe:5::5`. The other node A2 is in a different network (`3ffe:4::/32`). Let us now consider where two nodes (C1 and C2) send packets destined for anycast addresses `3ffe:5::5`. The difference is whether there is an anycast router on the route to seed node A1. C1 first forwards the packet to router A1 through unicast routing (solid arrow). Intermediate router A1 is anycast and can detect the packet is also anycast. According to the anycast routing (dashed arrow), anycast router R1 then forwards it to node A2, which is the node nearest to C1. However, since there is no anycast router between C2 and A1, the packet is simply forwarded to A1 through unicast routing only. Note that there is a more appropriate node (A2) in this network. For example, if we replace the router next to C2 (short white cylinder) with an anycast router, the packet could be transmitted to the more appropriate A2 node through anycast routing.

The above description reveals that our anycast routing protocol works appropriately even when there are a limited number of anycast routers. If these are increased, better routing is achieved. When all routers in the network are anycast, flexible routing adopting a control policy using various metric will be possible.

We divided the anycast routing protocol into the following two processes to define it.

1) Initiate anycast membership (Subsection III-C) The anycast router collects information on nodes that intend to join anycast memberships.
2) Construct and update routing table (Subsection III-D) According to the information collected, anycast routers construct their own routing tables and then exchange routing information with one another to reconfigure these.

Since anycast and multicast have many similar characteristics, we modified the multicast routing protocol for anycasting. Of the several multicast routing protocols that are currently available, we chose DVMRP [8], MOSPF [9] and PIM-SM [10], which are representatives of the three types of routing protocols; i.e., distance-vector, link-state and core-based-tree. Since each multicast protocol has both advantages and disadvantages, we defined the anycast routing protocol based on all of these, i.e., (1) the Distance Vector Anycast Routing Protocol (DVARP), (2) the anycast extension of OSPF (AOSPF), and (3) the Protocol Independent Anycast Sparse Mode (PIA-SM). In this paper, we present only (1) DVARP in Subsection III-D. In (2) AOSPF and (3) PIA-SM, the method of multicast routing protocol (i.e., MOSPF, PIM-SM) can be applied without any modifications.

C. Initiate Anycast Membership

Like multicasting, the host participating in (or leaving from) anycast membership must have the capability of notifying the nearest anycast router of status (joining/leaving). The method of finding a host participating in anycast membership (called *anycast host* below) is different and is based on the location of the anycast host. If the anycast host and the anycast router are on the same segment, an extended version of MLD (Multicast Listener Discovery) [11] is used [12]. We call this ARD (Anycast Receiver Discovery). An anycast host generates an MLD report message to the anycast router before joining anycast membership. However, the anycast host sends an MLD leave message prior to leaving membership. Because the destination address field of MLD packets is set to the link-local address of routers (`FF02::2`), this method can only be applied where all hosts and routers reside within the same segment.

D. Constructing and Updating Routing Table – DVARP

Since multicast membership is expected to change dynamically in DVMRP, it is hard to specify the route that multicast

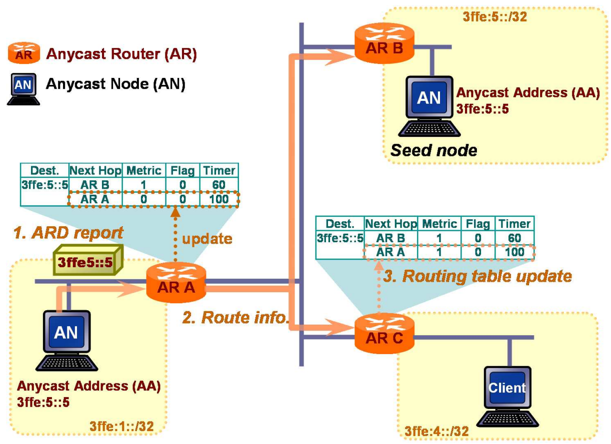

Fig. 4. DVARP.

packets will traverse before beginning transmission. Therefore, a flooding (or broadcasting) approach is effective. However, anycast membership does not change as frequently as multicast, and its routing information is more stable. Therefore, DVARP does not use a flooding method but exchanges routing information periodically like RIP.

Figure 4 has an example of updating a DVARP routing table. The DVARP operation is as follows.

1) If the anycast router detects changes in anycast membership, the anycast router updates/creates the routing entry in its own routing table.
2) Each DVARP router periodically sends its own routing information to its adjacent routers.
3) If the router receives routing information from adjacent routers, it updates entries in the routing table.

E. Comparisons of Anycast Routing Protocols

Let us now compare our proposed protocols, i.e., DVARP, AOSPF, and PIA-SM. We had the following three objectives in our comparison.

- Protocol overheads (e.g., CPU load, memory consumption)
- Convergence time due to membership changes
- Ease of implementing protocols

Table II summarizes the comparisons. Note that these are similar to obtained for multicast routing protocols. With respect to protocol overheads, both DVARP and AOSPF consume a vast amount of network resources and their traffic consumption is almost linear for the number of anycast groups and the number of nodes sharing the same anycast address. Therefore, these protocols are only applicable to small networks with high levels of available bandwidth.

PIA-SM, however, has hardly any traffic consumption because only the RP has routing information, which the other PIA-SM routers do not. Therefore, PIA-SM is more scalable than the other two protocols. However, PIA-SM has a problem in that anycast packets are not transferred through the optimal path because they are always transferred through the RP. Another problem is that traffic concentrates around the RP. These problems cause extra packet transmission delays. Because of this, PIA-SM can be applied to large networks like the Internet.

DVARP takes a long time for routes to converge although AOSPF takes less time. In PIA-SM, since all routing information is only kept by the RP, it is not necessary to exchange routing information.

TABLE II

COMPARISONS OF THREE ANYCAST ROUTING PROTOCOLS.

		DVARP	AOSPF	PIA-SM
overhead	network	O(gm)	O(gm)	RP:O(ng)
	router	O(gs)	O(gs) + O(l*log(gm))	RP:O(gs)
convergence		hop by hop		none
implementability		not available		available

n: the total number of nodes in the network, g: the number of anycast groups, m: the mean number of nodes which share the same anycast address, s: the mean number of anycast routing entries, l: the total number of links.

The implementation of PIM-SM for IPv6 can currently be done while DVMRP and MOSPF implementations cannot as far as we know. That is, PIA-SM is easier to implement than the other two.

IV. CONCLUSION AND FUTURE WORKS

The IPv6 anycast has several problems in facilitating communications with existing applications. To solve these, we proposed a new protocol called the AARP, which changes the anycast address into a corresponding unicast address, and actual communication uses the unicast address after conversion. We also proposed and designed three anycast routing protocols by focusing on and comparing the similarities between anycasting and multicasting and modifying the existing multicast routing protocol. In this paper, however, we only discussed the design of one anycast routing protocol although we are currently implementing the others and evaluating their feasibility.

REFERENCES

[1] R. Hinden and S. Deering, “IP version 6 addressing architecture,” *RFC3513*, April 2003.
[2] S. Deering and R. Hinden, “Internet protocol, version 6 (IPv6) specification,” *RFC2460*, December 1998.
[3] S. Doi, S. Ata, H. Kitamura, and M. Murata, “Ipv6 anycast functionality/terminology definition,” *Internet draft* `draft-doi-ipv6-anycast-func-term-00.txt`, February 2003.
[4] J. Hagino and K. Ettikan, “An analysis of IPv6 anycast,” *Internet draft* `draft-ietf-ipngwg-ipv6-anycast-analysis-01.txt`, July 2002.
[5] S. Ata, H. Kitamura, and M. Murata, “A protocol for anycast address resolving,” *Internet draft* `draft-ata-ipv6-anycast-resolving-00.txt`, June 2002.
[6] D. Katabi and J. Wroclawski, “A framework for scalable global IP-anycast (GIA),” in *SIGCOMM*, pp. 3–15, 2000.
[7] G. Agarwal, R. Shah, and J. Walrand, “Content distribution architecture using network layer anycast,” in *Proceedings of the Second IEEE Workshop on Internet Applications*, 2001.
[8] D. Waitzman, C. Partridge, and S. Deering, “Distance vector multicast routing protocol,” *RFC1075*, November 1988.
[9] J. Moy, “MOSPF: Analysis and experience,” *RFC1585*, March 1994.
[10] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. gung Liu, P. Sharma, and L. Wei, “Protocol independent multicast-sparse mode (PIM-SM): Protocol specification,” *RFC2117*, June 1998.
[11] S. Deering, W. Fenner, and B. Haberman, “Multicast listener discovery (MLD) for IPv6,” *RFC2710*, October 1999.
[12] B. Haberman and D. Thaler, “Host-based anycast using MLD,” *Internet draft* `draft-haberman-ipngwg-host-anycast-01.txt`, May 2002.