Internet Engineering Task Force
INTERNET DRAFT

A. Birman
D. Kandlur
J. Rubas
IBM

28 November 1995

An extension to the MARS model
draft-kandlur-ipatm-mars-directvc-00.txt

Status of This Memo

This document is an Internet-Draft. 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 not appropriate to use Internet Drafts as reference material, or to cite them other than as a ``working draft'' or ``work in progress.''

To learn the current status of any Internet-Draft, please check the ``1id-abstracts.txt'' listing contained in the internet-drafts Shadow Directories on ds.internic.net (US East Coast), nic.nordu.net (Europe), ftp.isi.edu (US West Coast), or munnari.oz.au (Pacific Rim).

Abstract

This memo describes an extension to the IP multicast architecture now being developed by the IP over ATM working group, also known as MARS (Multicast Address Resolution Server). In the extended architecture the determination whether a Multicast Server (MCS) or a point-to-multipoint VC will be used for multicasting is made by the application in the sender endpoint. Moreover, there can be senders in a multicast group using point-to-multipoint VCs while other senders use a MCS.

It is assumed that the reader is familiar with the MARS document [MARS].

1. Introduction

The MARS document [MARS] describes a mechanism for handling IP multicast over ATM. Clusters of endpoints share a MARS and use it to track and disseminate information on membership in multicast groups. This allows endpoints to establish VCCs and to transmit to their multicast group over these VCCs. Multicasting is supported using either meshes of VCs or ATM-level multicast servers (MCSs). The choice is made on a per-group basis and is transparent to the endpoints.

Yet in many important cases it is better to have the SVC management controlled directly by applications, as described in [Rekhter95]. For applications that have a relatively long lifetime and also require QoS guarantees the establishment of dedicated multicast VCs can be certainly justified. Other applications could be served by the shared service provided by MCSs and thus the overhead associated with the establishment and maintenance of separate multicast VCs can be avoided.

Moreover, there is further evidence that the MARS mechanism should be generalized to allow some endpoints in a multicast group to use separate point-to-multipoint VCs while other endpoints would rely on the MCS. Specifically, consider the scenario in which a source endpoint sends an RTP multimedia stream to a multicast group, as in an MBONE style conference. The source would use a separate point-to-multipoint VC with QoS guarantees. At the same time the source receives periodic RTP reports from the multicast receivers. These reports could use the default VCC through the MCS. In the current MARS architecture one would have to use either the MCS for all senders in the group, or a mesh of VCs for every sender in the group.

2. Summary of proposed changes to MARS

An endpoint which is a sender to a multicast group may have an application-driven requirement that it multicasts to the members of the group using a point-to-multipoint VC. We will refer to such a sender endpoint as a Direct Sender. In order to establish this point-to-multipoint VC to the members of the multicast group the endpoint has the added ability to request, and receive, from the MARS the list of ATM addresses corresponding to the members of the multicast group (the 'host map'). Furthermore, a Direct Sender needs to track membership in the multicast group, very much like an MCS. For this reason, if an MCS exists for the group then the Direct Sender is added as a leaf to ServerControlVC.

The proposed extension to MARS is intended to be backward compatible with the MARS architecture. Specifically, hosts implementing the MARS extension should interoperate with a MARS or other hosts which implement only the MARS architecture [MARS].

Two issues arise in connection with these changes, and they will be addressed in the discussion below. Since a Direct Sender may be a member of both ClusterControlVC and ServerControlVC, and since it must track sequence numbers on both ``control VCs'', it must have the ability to identify the control VC on which the message arrived. Another issue concerns applications which have a high join/leave activity, such as Distributed Interactive Simulations. These applications may cause significant non-productive traffic of MARS messages on a control VC.

3. The proposed changes to the MARS

3.1. TLVs defined

(a) get_map: When used in a MARS_REQUEST message has the following meaning: ON: endpoint is requesting a host map. OFF: undefined.

When used in a MARS_MULTI message has the following meaning: ON: the message contains the requested host map. OFF: undefined.

(b) direct_sender: When used in a MARS_LEAVE message has the following meaning: ON: endpoint is no longer a Direct Sender to the multicast group. OFF: undefined.

When used in a MARS_UNSERV message has the following meaning: ON: endpoint is no longer a Direct Sender to any multicast group. OFF: undefined.

(c) control_vc_id: When used in a message sent by the MARS it identifies the control VC on which the message was sent. It is an unsigned integer.

3.2. Sender endpoint behavior

When a Direct Sender requires a host map it sends a MARS_REQUEST with the get_map TLV ON (i.e. value set to ON). If the MARS_MULTI received in return has the get_map TLV ON then the endpoint can proceed. If the MARS_MULTI does not have the get_map TLV ON then the endpoint should assume that the MARS does not support this feature and should discontinue its Direct Sender status.

If the MARS was able to satisfy the endpoint's request then it is possible that the endpoint may receive MARS messages on multiple control VCs. The Direct Sender must be able to support MARS messages arriving on multiple control VCs. As an aid in distinguishing these messages the endpoint should support the control_vc_id TLV when attached to MARS messages. Endpoints which do not support this feature should ignore the control_vc_id TLV.

An endpoint which is a Direct Sender in a multicast group may cease to be a sender for that group due to the expiration of an inactivity timer or due to an application-initiated request. The endpoint then releases the point-to-multipoint VC to the members of the group and sends to the MARS a MARS_LEAVE message with the direct_sender TLV ON. The endpoint then determines whether it continues to be a Direct Sender for other groups. If there are no groups left for which the endpoint is a Direct Sender then it sends to the MARS a MARS_UNSERV message with the direct_sender TLV ON.

A MARS_MIGRATE message arriving from the MARS is an indicator that the endpoint will now receive updates on ServerControlVC instead of ClusterControlVC.

3.3. MARS behavior

A generic data structure, Direct Sender List, contains entries of the form {mc_address, atm.1, ..., atm.k} where mc_address is a multicast group address and ATM addresses atm.1, ..., atm.k correspond to the Direct Senders in that group.

When the MARS receives a MARS_REQUEST from an endpoint requesting a host map, the MARS will attach to the MARS_MULTI message the get_map TLV. The MARS will indicate in this TLV that the map in the MARS_MULTI is a host map.

When an endpoint is registered, the MARS (as usual) adds the endpoint to ClusterControlVC. If that endpoint then requests a host map, the MARS will also add the endpoint to ServerControlVC if there is an MCS defined for that group or if an MCS becomes defined for the group in the future. Then, it is on ServerControlVC that group membership updates will be reflected. If MARS added the endpoint to ServerControlVC, then the entry in the Direct Sender List will be updated by adding the ATM address of the endpoint to the entry.

If an MCS is not defined for the group then the endpoint is not added to ServerControlVC and group membership updates will continue to be sent on ClusterControlVC.

When the MARS receives a MARS_LEAVE message with the direct_sender TLV ON it recognizes an endpoint which ceases to be a Direct Sender in the specified multicast group. The MARS then updates the Direct Sender List.

When the MARS receives a MARS_UNSERV message with the direct_sender TLV ON it recognizes an endpoint which ceased to be a Direct Sender in any multicast group. The MARS then removes it as a leaf in ServerControlVC.

4. Non-productive traffic on ServerControlVC

There is a concern that some applications which have a high join/leave activity, such as Distributed Interactive Simulations, may cause significant non-productive traffic. For example, MARS_JOIN and MARS_LEAVE messages reflected by the MARS on ServerControlVC have to be processed by all Direct Senders. Thus, each one of these Direct Senders has to process all messages in order to select the possible few which affect its own multicast group.

In order to alleviate this problem one could go to the other extreme of creating a dedicated control VC for the Direct Senders of every multicast group. Clearly, there exists a range of additional possibilities, where the number of control VCs is somewhere in between these two extremes. Depending on the number of multicast groups and their behavior one could attempt to optimize the number of control VCs. This topic is left for future study.

5. Tracking sequence numbers

A Direct Sender may have to track sequence numbers on multiple control VCs, and thus must have the ability to identify the control VC on which the message arrived. Since the MARS sequence numbers are associated with different VCs it is sufficient that the endpoint be able to identify the VC a message arrived on. It appears that this cannot always be counted on. Therefore, additional information needs to be supplied to the endpoint which allows it to identify the control VC on which the message arrived. This information is supplied by the control_vc_id TLV.

6. Acknowledgements

The authors would like to thank Grenville Armitage (Bellcore) and Tim Smith (IBM) for their review and comments.

7. References

[MARS] G. Armitage, "Support for Multicast over UNI 3.1 based ATM Networks", Internet Draft, IP over ATM Working Group, draft-ietf-ipatm-ipmc-06.txt, August 1995.

[Rekhter95] Y. Rekhter and D. Kandlur, "IP Architecture Extensions for ATM", Internet Draft, draft-rekhter-ip-atm-architecture-01.txt, July 1995.

8. Authors' Address

Dilip Kandlur
T. J. Watson Research Center
IBM Corporation
30 Saw Mill River Rd.
Hawthorne, NY 10532

Phone: +1-914-784-7722
E-mail: kandlur@watson.ibm.com

Alex Birman
T. J. Watson Research Center
IBM Corporation
30 Saw Mill River Rd.
Hawthorne, NY 10532

Phone: +1-914-784-7275
E-mail: birman@watson.ibm.com

Jim Rubas
T. J. Watson Research Center
IBM Corporation
30 Saw Mill River Rd.
Hawthorne, NY 10532

Phone: +1-914-784-7794
E-mail: rubas@watson.ibm.com