E. Crawley
Bay Networks
February 22, 1996

Multicast Routing over ATM

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. Internet Drafts 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."

Please check the I-D abstract listing contained in each Internet Draft directory to learn the current status of this or any other Internet Draft.

Abstract

As the use of IP multicasting and Asynchronous Transfer Mode (ATM) technologies increases, it becomes important to think about ways that IP multicast routing protocols interface to ATM networks. The Multicast Address Resolution Server (MARS) [1] addresses the problem of determining group membership within a single ATM Logical IP Subnet (LIS) and provides a mapping between IP multicast addresses and ATM addresses while [2] looks at the issues of short cut routing for PIM Sparse Mode and CBT. In this document, the general problems of IP multicast routing protocols over ATM networks are explored and the issues related to the different forms of multicast routing protocols; dense vs. sparse, shared tree vs. source tree, etc. are addressed. Additionally, the issues related to short cut and QoS routing are discussed. This document is intended as a starting point for exploring the issues of IP multicast routing over ATM.

1. Introduction

Asynchronous Transfer Mode is a unique technology that provides some features and challenges to layer 3 protocols that utilize it. While ATM can be used to emulate traditional point-to-point links between network nodes, it makes sense to utilize some of the additional features of ATM if they can provide a benefit to the higher level protocols. Some of the features in ATM can provide benefit to IP multicast routing protocols. However, the benefits vary based on the multicast routing protocol and the ATM features used. The way different multicast routing protocols distribute data and control information is very important when mapping to ATM. Further, there are some additional issues that come up in ATM networks that need to be addressed if such issues as short cut ATM paths, scaleability, and Quality of Service (QoS) are considered. The issues and problems presented are by no means a complete list but they are a starting point for further discussion.

2. Data Plane and Control Plane

It is important to note the distinction between the control plane and the data plane for IP multicast services. For IP multicasting in an IP over ATM environment, the control plane is made up of an IP multicasting protocol such as Core Based Trees (CBT)[3], Distance Vector Multicast Routing Protocol (DVMRP)[4], Multicast Extensions to OSPF (MOSPF)[5], Protocol Independent Multicasting-Sparse Mode (PIM-SM)[6], Protocol Independent Multicasting-Dense Mode (PIM-DM)[7], and optionally, MARS. The Internet Group Management Protocol (IGMP) is not necessary over the ATM cloud but its presence is not precluded. The control plane is responsible for setting up the multicast tree and maintaining its state. In an IP over ATM environment, the messages used to set up the tree may possibly be forwarded over a different path than the multicast data. This issue can become very important if the control data path is used for other control protocols such as RSVP [8].

The data plane in an IPATM environment consists of the virtual circuits (VCs) that are involved in the actual forwarding multicast data. These can either be point-to-point (pt-to-pt) or point-to-multipoint (pt-to-mpt) VCs. The control plane can use the same VCs but it may use a different set. Additionally, changes in the control plane may cause changes in the VCs used in the data plane.

3. Dense Mode and Sparse Mode

The exiting multicast routing protocols are divided into two classes, dense mode and sparse mode. PIM-DM [7] and DVMRP [4] are dense mode protocols and use a “broadcast and prune” model. Broadcast and prune protocols forward data from senders to all subnets away from the sender until a router indicates that it has no clients interested in receiving the multicast by sending a control plane “prune” message toward the sending router. Broadcast and prune protocols are best suited to campus LAN environments where bandwidth is widely available and simplicity of configuration is paramount.

In contrast, sparse mode protocols such as PIM-SM [6], MOSPF [5], and CBT [3] use control plane messages for setting up multicast trees such that data is only sent to the networks that have interested receivers. Sparse mode protocols are used when the receivers are widely dispersed.

Mapping these models to the Non-Broadcast MultiAccess (NBMA) environment of ATM presents a bit of a problem. If ATM VCs are considered “expensive,” require a significant setup time, or a fully interconnected mesh does not exist, then it may be best to use a sparse mode protocol to avoid setting up VCs to destinations that may not be interested in receiving the multicast data. Sparse mode protocols can also make use of pt-to-pt VCs when transiting ATM subnetworks that do not have any receivers. In contrast, environments that consist mostly of permanent virtual circuits (PVCs) can utilize dense mode or sparse mode protocols since there is no dynamic set up cost.

4. Shared Trees and Source Trees

A further division of the exiting multicast routing protocols exists in the type of multicast distribution tree that is created. A source tree is “rooted” at the sender with a separate tree, and corresponding state, for each sending network. A shared tree is rooted at a node in the network with all senders using the same tree for distribution of multicast data. The dense mode protocols, by their very nature, create source trees. MOSPF and one mode of operation for PIM-SM also use source trees. CBT and the other mode of operation for

PIM-SM use shared trees. Source trees and shared trees present different problems when used over ATM.

Source trees more closely map the model of ATM pt-to-mpt VCs. This makes them easier to visualize on the ATM network but presents more scaleability problems regarding the use of VCs, depending on the configuration. However, for a small number of senders per multicast group, this is not an important issue. As the number of senders (and corresponding ATM network ingress points) increases, the number of pt-to-mpt VCs in the network increases. As the number of receivers (and corresponding ATM network egress points) increases, each of the pt-to-mpt VCs will require additional the additional ATM parties to be added.

Shared trees, while requiring less state in the network, are more likely to have multiple hops in the ATM network. It is quite possible for a router on a shared tree in an ATM network to have packets entering and exiting the same ATM interface if MARS or a short-cutting mechanism such as NHRP is not used.

5. ATM Hosts and MARS

Naturally, any use of IP multicast routing over ATM has to take into consideration hosts directly attached to the ATM network. MARS essentially performs the function of IGMP for an ATM LIS and is required, if there is a mix of hosts and routers on the ATM network. If the ATM LIS is made up of only routers, it is possible to operate without MARS but additional hops and packet copying may occur as packets are routed through the ATM network.

If MARS is used, the ATM LIS looks like a multicast capable subnetwork to the multicast routing protocol (see Figure 1). MARS allows the multicast routing protocol to track the nodes on the ATM LIS that are part of the multicast group and therefore can allow the nodes to set up the appropriate pt-to-mpt VCs.

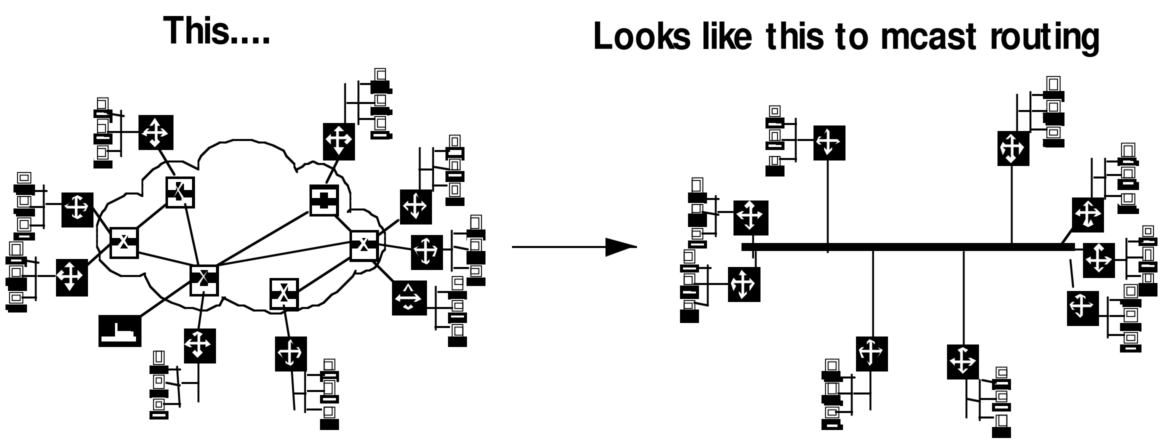

Figure 1

The MARS specification also discusses the use of MARS in conjunction with a ATM multicast server. A multicast server can be used to relieve the burden of managing pt-to-mpt VCs for a router but incurs a possible delay and bottleneck at the server. The issues and performance of these configurations should be investigated further.

6. Short Cut Routing

The ability to make connections directly from a node on one ATM LIS to a node on another ATM LIS, bypassing intermediate hops, is a very attractive feature of ATM and other NBMA protocols. NHRP [9] provides one mechanism for determining the ATM address of the egress node in an ATM cloud based on the destination IP address. Of course, this doesn't mean much if the destination address requested is an IP multicast address so short cut mechanisms have to be used in concert with a multicast routing protocol that can take advantage of the abilities of the underlying unicast routing structure. This means that PIM and CBT are best suited to the use of short cuts since they are independent of the underlying unicast routing protocol. It is unclear whether short cuts can be used with MOSPF or DVMRP but warrants additional investigation. [2] describes how NHRP can be used with PIM or CBT for building the multicast tree by using short cuts from a joining node to a Core/Rendezvous Point (RP)/Source Router

There are some possible problems that need to be thought through with short cuts and their use with multicast routing. Some of these problems can come from changes in the routing over time. These problems need to be investigated fully before making widespread use of short cuts with multicast routing.

There are also scaling issues for large multicasts when using short cuts that need to be examined. Short cuts can cause a larger amount of state to be stored in the nodes, from the additional short cut end points, and can require larger fanouts for pt-to-mpt VCs. Depending on the size of the ATM network, these can present problems.

7. Considerations for ATM VC Usage

ATM provides several different types of VCs and each has their own properties, advantages, and disadvantages. In this section, the implications of different types of VCs and their usage are discussed in relation to IP multicast routing. This is not intended to be a complete list.

7.1 Permanent Virtual Circuits (PVCs)

PVCs are the simplest form of ATM VCs and closely emulate common pt-to-pt links. Obviously, these links can be easily used by IP multicast routing and do not require any special treatment. Multicast routers are likely to have multiple PVCs running out of the same physical ATM interface and will have the extra burden copying packets for each PVC even though the packets travel out of the same physical interface. This is not the best use of resources but can be tolerated for a relatively small number of PVCs.

7.2 Switched Virtual Circuits (SVCs)

SVCs allow for more dynamic connection setup but bring along the burden of the time to setup the VCs and what to do with the data during SVC setup. Additionally, a routing overlay is needed to determine where to route SVCs.

For the moment, we can assume that an overlay of VCs exists to at least some neighbors on he same LIS so that there is connectivity to all the nodes on the LIS, albeit not necessarily one hop away. If a node determines, perhaps via incoming data rates, that an SVC needs to be established to a neighbor, it must continue to route data via an indirect path until the

SVC is setup or it will have to drop data. These policies must be understood for multicast routing since the first data arriving may be control messages to set up the tree and dropping these messages may be catastrophic to tree setup.

The unicast routing that IP multicast routing protocols depend on must be flexible enough to provide next hops in the ATM environment such that SVCs can be set up as needed to get to a particular destination. Further, it may be worthwhile to set up parallel VCs especially when QoS is considered.

7.3 Point-to-Multipoint VCs

Point to multipoint (pt-to-mpt) VCs are the only way that ATM networks can be used to replicate IP multicast data. This relieves the router from duplicating data over multiple VCs. Of course, point-to-multipoint VCs are unidirectional and come with a different set of problems relating to management and scaling.

7.4 VC Fanout

An important consideration when dealing with pt-to-mpt VCs is the number of destinations that can be supported. There may be limits in the equipment in the ATM network that limit the number of parties on a pt-to-mpt VC therefore it may be important to limit the fanout or number of parties on a pt-to-mpt VC for IP multicasting. This is especially true with large clouds where short cut routing is used.

7.5 VC Management

The MARS specification [1] discusses the issues of SVC management with respect to a single multicast group but there are some interesting issues that come up when dealing with multiple multicast groups. For the most part, VC management is a local problem that does not require standardization as long as a receiving node can accept VCs from different ATM sources and an identical hop-wise path can be made back to the ATM source node for use with protocols such as RSVP [8].

One method of managing multiple multicast groups is to assign each group to a different pt-to-mpt VC but this limits the number of groups that can be supported to the number of VCs available. A second method looks for an existing pt-to-mpt VC that goes to the same set of ATM destinations. If one is found the data is sent on the VC. If not, a new VC is set up or multiple VCs that go to the right set of ATM addresses can be used. This model can be extended for further aggregation to include policies that would allow VCs that go to a superset of the ATM destinations desired. Of course, this method can quickly hit a point of diminishing returns, but can be applied in cases where VCs are at a premium or a very large number of multicast groups need to be supported.

7.6 ATM Hard State vs. Soft-State Protocols

ATM networks use reliable messaging and hard state connections. PIM uses a soft state mechanism to periodically refresh multicast routing state. [2] discusses reducing the frequency of the periodic soft state messages when using PIM over ATM. The reasoning is that since ATM provides a reliable hard state connection, with easy detection of VC failure, there is a lesser need for soft state refresh messages to detect the same kind of failures. Note, the soft state refresh messages should not be eliminated because there are still classes of failures that cannot be detected by the loss of a VC, such as the failure of a process on an ingress or egress router.

CBT uses hard state with periodic echoing to determine neighbor status. Over ATM, the frequency of these echoes can be reduced, if necessary. MOSPF depends on link state changes as well as infrequent hellos for state management so no changes would be necessary in this respect when running over ATM.

8. ATM Futures that Impact Multicast Routing

The ATM Forum has been evolving the ATM specifications and some of the changes currently planned and those being discussed can improve the ability of IP multicast routing to run more effectively over ATM. As these features become available, multicast routing should try to take advantage of them.

8.1 Leaf Initiated Join

ATM Forum Release 4.0 contains a Leaf Initiated Join (LIJ) capability to allow an ATM end point to join an existing pt-to-mpt VC without necessarily contacting the root of the VC. This more closely matches the receiver-based model of IP multicasting. There are issues about determining the correct VC to join when a root can be managing multiple and possibly parallel pt-to-mpt VCs that must be addressed in order for this feature to be used.

8.2 Variegated Point-to-Multipoint

There have been discussions and contributions that consider the possibility of a pt-to-mpt VC that would allow different QoS parameters for each branch of the VC. This would be a benefit to RSVP over ATM and multicast routing would need to know how to set up and utilize such VCs.

8.3 Group Addressing

In order to make ATM multipoint communications more closely match the IP multicast model, schemes that allow ATM group addresses that could be directly mapped to IP multicast addresses have been discussed. This is contrary to the traditional one-to-many communications of ATM so it may take a bit of work to fit into the model. An important issue is to make whatever scheme is developed be scaleable to the needs of IP multicasting over ATM.

9. QoS Routing

Since ATM networks have QoS capabilities and these capabilities are advertised in the PNNI (Private Network-to-Network Interface) routing information, it makes sense for IP multicast routing over ATM to make use of this information in concert with RSVP. The work in this area is just starting but it will have an impact on multicast routing, particularly over ATM.

10. Conclusions

In this draft, an initial attempt has been made at gathering, and addressing, some of the issues related to IP multicast routing over ATM. There are some areas that need to be documented and investigated further. They include:

- The handling of control plane vs. data plane flows in ATM where they can possibly follow different paths.
- The use of dense mode protocols in an ATM SVC environment
- The use of IGMP in an IP over ATM environment
- The impact of source tree vs. shared tree data distribution in ATM
- The impact of short cut routing for IP multicasting in ATM
- The impact of multicast servers on IP multicast routing
- Guidelines for VC management and usage for all multicast protocols
- Tweaking of soft-state mechanisms over ATM to reduce refresh intervals
- Integrating new ATM features with multicast routing protocols
- The impact of multicast QoS routing schemes over ATM

These items should be considered as work items for the IDMR and IPATM IETF working groups.

11. References

[1] G. Armitage, "Support for Multicast over UNI 3.0/3.1 based ATM Networks", draft-ietf-ipatm-ipmc-09.txt, December 1995.
[2] Y. Rekhter, D. Farinacci, "Support for Sparse Mode PIM over ATM", draft-rekhter-pim-atm-00.txt, February 1996.
[3] A. Ballardie, S. Reeve, N. Jain, "Core Based Trees (CBT) Multicast-- Protocol Specification", draft-ietf-idmr-cbt-spec-04.txt, February 1996.
[4] D. Waitzman, C. Partridge, S. Deering. "Distance Vector Multicast Routing Protocol". RFC 1075, November 1988.
[5] J. Moy, Multicast Extensions to OSPF, RFC 1584, March 1994.
[6] S. Deering, D. Estrin, D. Farinacci, V. Jacobson, C. Liu, L.Wei, P. Sharma, S. Helmy, "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification", draft-ietf-idmr-pim-spec-02.txt, September 1995.
[7] S. Deering, D. Estrin, D. Farinacci, V. Jacobson, C. Liu, L.Wei, P. Sharma, S. Helmy, "Protocol Independent Multicast-Dense Mode (PIM-DM): Protocol Specification", draft-ietf-idmr-pim-dm-spec-01.txt, January 1996.
[8] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin. "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", draft-ietf-rsvp-spec-09.txt, February 1996.
[9] D. Katz, D. Piscitello, B. Cole, J. Luciani, "NBMA Next Hop Resolution Protocol (NHRP)", draft-ietf-rolc-nhrp-07.txt, December 1995.

12. Author's Address

Eric S. Crawley
Bay Networks, Inc.
3 Federal Street
Billerica, MA 01821
+1 508 670-8888
esc@baynetworks.com