17

# Logical IP Subnet Information Protocol for IP Routing over ATM

*Chan Park, Jin Oh Kim, Hyup Jong Kim*
*Broadband Communications Department, ETRI*
*161 Kajong-Dong, Yusong-Gu, Taejeon, 305-350, Korea*
*Tel: +82-42-860-5749*
*Fax: +82-42-861-5597*
*E-mail: cpark@winky.etri.re.kr*

**Abstract**

This paper proposes a simple IP routing solution on ATM network using IETF RFC 1577 "Classical IP and ARP over ATM", named the LIS(Logical IP Subnetwork) Information Protocol(LIP).Originally, RFC 1577 have been specified for an application of classical IP and ARP in ATM network environment configured as an LIS. Therefore, different routing protocols are required to provide an appropriate path between LISs. Currently, the NHRP(Next Hop Resolution Protocol) or the MPOA(Multi-Protocol Over ATM) are considered as candidates. However, those protocols are too heavy or complicate to implement.

LIP has been devised as an alternative, which is so easy to implement because the basic concept and procedures are identical with RFC 1577. LIP includes Server-base and PVC-base routing solutions in a Large Cloud. We introduce essential function entities and routing procedures for each case. In addition, we also describe how the IP multicast services be supported between several LISs using the LIP protocol.

**Keywords**

Resolution, routing, subnet, server, multicast, signaling

Management of Multimedia Networks and Services R. Boutaba & A. Hafid (Eds.)

# 1 INTRODUCTION

IETF RFC 1577 "Classical IP and ARP over ATM"(Laubach, 1996) proposals employ a form of Address Resolution Protocol(ARP) to handle the address translations between the IP address and ATM address. The address translations are needed due to the mismatch between two addressing schemes. The ATM address is based on E.164 or the more general NSAP(Network Service Access Point) addressing schemes. In addition to translating addresses, connections are set up and released for all source and destination pairs since ATM is a connection-oriented technology. The most commonly employed signaling protocol to set up and release connections is Q.2931(ITU-T, 1993) that adopted as a standard in the ITU-T and ATM Forum. Thus, two phases are needed to support the IP over ATM models: first, an address resolution is performed to obtain the ATM address corresponding IP address; second, a UNI(User Network Interface) signaling connection request message is sent by the host causing a hop-by-hop connection setup procedure, as specified in P-NNI(ATM Forum, 1996) or B-ISUP(ITU-T, 1996) standards, to be initiated. Upon completion of this procedure, the host can map an IP address to an ATM VPI/VCI which can be then used for subsequent data transmission.

In the classical IP model, when IP networks are overlaid over an ATM based large cloud, which means a collection of endpoints(routers or hosts), connected over a fabric such that direct communication can be established between any pair of endpoints subject to policy restrictions employed within the fabric(Woodworth, 1996), it is required that endpoints that do not share the same network prefix must communicate via an intermediate router. In other words, endpoints that do not belong to the same LIS(Logical IP Subnet) must communicate via an intermediate IP router. This means two virtual channel connections(VCC) are interconnected via intermediate router between source and destination. The router is not a connecting point such as ATM switches but an endpoint. It requires that different kind of call control functions be added to interconnect two VCCs. In some cases, several routers can be involved to reach the destination. As shown in Figure 1(a), multiple VCCs are interconnected to provide single call. This is clearly inefficient, since the ATM routers become bottlenecks. However, this also precludes the establishment of single connection with a requested QOS between the two endpoints. If the source knows the ATM address of the destination as shown in Figure 1(b), a shorter path can be established between source and destination than Figure 1(a).

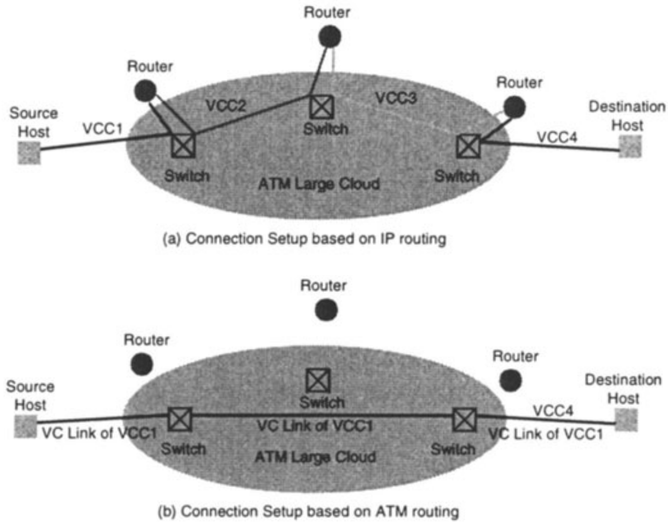



**Figure 1** IP routing versus ATM routing

Consequently, we get the following conclusion: although the router can be interconnect between two large clouds, within a large cloud, a direct VCC should be able to be set up between source and destination without help of routers. To do this, this paper shows two phases of IP over ATM model can be applied appropriately and RFC 1577 itself is suitable for this application. We propose a new approach called LIP(LIS Information Protocol) protocol that solves IP-to-ATM address translation regardless of LIS. This is made possible by addressing a certain endpoint as the Route Server of each LIS and exchanging then LIS information among Route Servers. This scheme provides performance gains by excluding the router in a large cloud, and also reduces the complexity of software needed in the endpoints to support IP routing on ATM network.

Section 2 briefly describes the existing IP over ATM model and Section 3 describes our proposal in conjunction with IP over ATM model. Section 4 shows how our proposal can be applied to the multicast IP services. Finally, in Section 5, we conclude the document.

## 2 CLASSICAL IP AND ARP OVER ATM

We begin by first describing current classical IP and ARP over ATM based on IETF RFC 1577 document. Since our proposal uses the packet format and the procedures of the ATM ARP protocol of RFC 1577, we need to understand the concept of the ATM ARP protocol.

This protocol is known as "classical IP over ATM" and introduces the notion of a LIS. Like a normal IP subnet, a LIS consists of a group of IP nodes (such as hosts or routers) that connect to a single ATM network and belong to the same IP subnet. To resolve the address of nodes within the LIS, each LIS supports a single ATM ARP server, while all nodes(LIS Clients) within the LIS are configured with the unique ATM address of the ATM ARP server. When a node comes up within the LIS, it first establishes a connection to the ATM ARP server, using the configured address. Once the ATM ARP server detects a connection from new LIS client, it transmits an *Inverse_ATM_ARP_Request* to the attaching client and requests the node's IP and ATM addresses, which it stores in its ATM ARP table. Subsequently, any node within the LIS wishing to resolve a destination IP address would send an *ATM_ARP_Request* to the server, which would then respond with an *ATM_ARP_Reply* if an address mapping is found. If not, it returns an *ATM_NAK* response to indicate the lack of a registered address mapping. The ATM ARP server ages out its address table for robustness, unless clients periodically refresh their entry with responses to the server's *Inverse_ATM_ARP_Request*. Once an LIS client has obtained the ATM address that corresponds to a particular IP address, it can then setup a connection to the address using signaling procedures. Figure 2 shows a general packet flow of classical IP over ATM model.

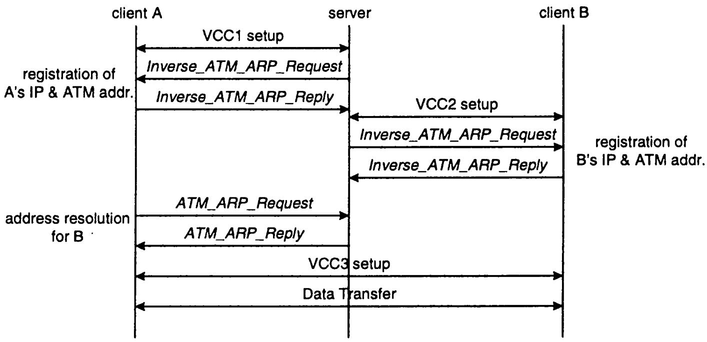



**Figure 2** Procedures of IP over ATM

The operation of classical mode is very simple. It does however, suffer from a number of limitations. One of these limitations is indicated by the phrase "classical". What this means is that the protocol does not attempt to change the IP host requirement(Braden, 1989) that any packet for a destination outside the source node's IP subnet must be sent to a default router. As shown at Figure 1(a), the communications between two nodes on two different LISs on the same ATM network must traverse each ATM router on the intermediate hops on the path between the source and destination nodes. This requirement, however, not only is clearly inefficient but also is not a good fit to the operation of IP over ATM, and a whole class of other "non-broadcast multi-access"(NBMA) networks, such as frame relay or X.25. In all such networks, it is possible to define multiple LISs, and the network itself could support direct communications between two hosts on two different LISs.(Alles, 1995)

To eliminate this limitation, the work on extensions to classical model such as "Next Hop Resolution Protocol"(NHRP)(Kats, 1995) has been proceeding by IETF Internetworking on NBMA group. However, we found that the protocol of RFC 1577 can be employed to avoid those limitations without addition of new protocols. In addition, the RFC 1577 is one of the most popular solution to provide the Internet services on the ATM network. Therefore, in order to employ such NHRP protocol, it should support the hosts which have the classical IP over ATM facilities based on RFC 1577. The more difficult problem is how the routing information is exchanged between Next Hop Servers(NHS). The draft document of IETF concerning NHRP just specifies that the routing information is exchanged between NHSs using existing IP routing protocols. However, existing routing protocols such as RIP or OSPF are not suitable mechanism for ATM. On the other hand, our proposal uses compatible protocol format and procedures with RFC 1577 as well as exchanges the routing information between route servers without any help of existing routing protocols. In next section, we describe our proposal.

## 3 LIS INFORMATION PROTOCOL(LIP)

As noted above, the classical model for IP over ATM suffers from the limitation imposed by host requirements that preclude "cut-through" routes that bypass intermediate router hops for communications between nodes on the same network, but within two different LISs. We have been working on protocols that overcome this limitation. After considering numerous different approaches(Braden, 1994) including NHRP, we are now finalizing work on a our proposal called LIS Information Protocol(LIP). In this section we describe the operation of this protocol.

### 3.1 Routing Hierarchy on NBMA Network based on LIP

Similarly to NHRP, LIP builds on the classical IP model, substituting for the concept of a LIS the notion of a logical NBMA network - that is, a network technology, such as ATM, frame relay, or X.25, which permits multiple devices to attached to the same network, but which does not easily permit the use of broadcast mechanisms, as are common on LANs. Such a network consists of set of nodes, each of which is attached to the same NBMA network, and are not physically or administratively restricted from directly communicating with each other.

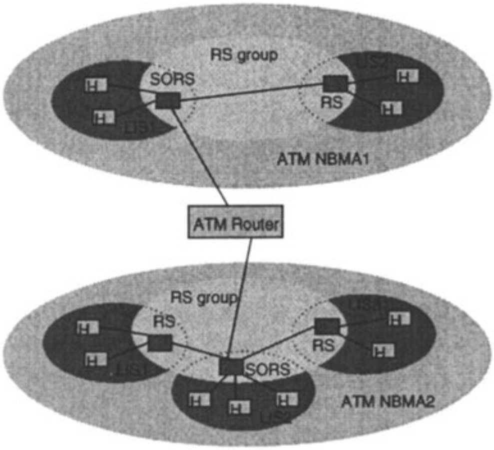



**Figure 3** The routing hierarchy of ATM NBMA

LIP uses the notion of Route Server(RS) in place of ATM ARP Server of classical IP over ATM model. All RSs in a ATM NBMA network can be divided into a couple of groups. A RS of each RS group must work as Server of RS(SORS). All SORSs can also be grouped hierarchically according to the size of the ATM NBMA network. Figure 3 shows the routing hierarchy of NBMA network based on the LIP protocol. Each SORS works as the default router of corresponding RS group, as each RS works as the default router of corresponding LIS. That is, it resolves the IP address which cannot resolve within LIS or RS group.

## 3.2 Server-base LIP Procedures

All hosts within a LIS already knows the ATM address of the RS since the Route Server coexists with the ATM ARP server. Similarly to classical IP over ATM model, all RSs(RS clients) within the RS group are configured with the unique ATM address of the SORS. When an LIS comes up within the NBMA, the RS of LIS first establishes a connection to the SORS, using the configured ATM address. Once the SORS detects a connection from new RS client, it transmits an *Inverse_ATM_LIP_Request* to the attaching client and requests the RS client's IP and ATM addresses, which it stores in its ATM LIS table. Subsequently, any RS within the RS group wishing to resolve a destination IP address would send an *ATM_LIP_Request* to the SORS, which would then forward the *ATM_LIP_Request* with a *Forward_ATM_LIP_Request* if an LIS mapping is found on ATM LIS table. If not, it returns an *ATM_LIP_NAK* response to indicate the lack of a registered LIS mapping. RS client receiving a *Forward_ATM_LIP_Request* from SORS would then response with an *ATM_LIP_Reply* to the SORS if an address mapping is found on ATM ARP table. If not, it returns an *ATM_LIP_NAK* to the SORS. When the SORS receives an *ATM_LIP_Reply* or an *ATM_LIP_NAK*, it forward with a *Forward_ATM_LIP_Reply* or a *Forward_ATM_LIP_NAK*, respectively. Here, the prefix "*Forward-*" means the packet is sent from the SORS and gives the convenience when there are multiple RS groups. The packets from RS clients such as *ATM_LIP_Request*, *ATM_LIP_Reply*, or *ATM_LIP_NAK* are sent on the connection which established to the SORS on registration. The forwarded packets using prefix "*Forward-*" are also sent on one of connections from SORS to RSs but an appropriate connection must be selected by looking up the ATM LIP table of SORS using destination IP address.

Whenever each host generates a packet with the destination IP address belonged to different LIS, it check the ATM ARP table if there is an address mapping. If there is, the host can setup a connection to the corresponding ATM address using signaling procedures. However, if not, the host transmits an *ATM_LIP_Request* to the RS of the LIS. Then, if the host which sent *ATM_LIP_Request* to its RS receives an *ATM_LIP_Reply* from the RS by above procedures, it stores IP and ATM addresses of the destination on ATM ARP Table and setup a connection to the corresponding ATM address using signaling procedures. Figure 4 shows a general packet flow of server-LIP protocol.

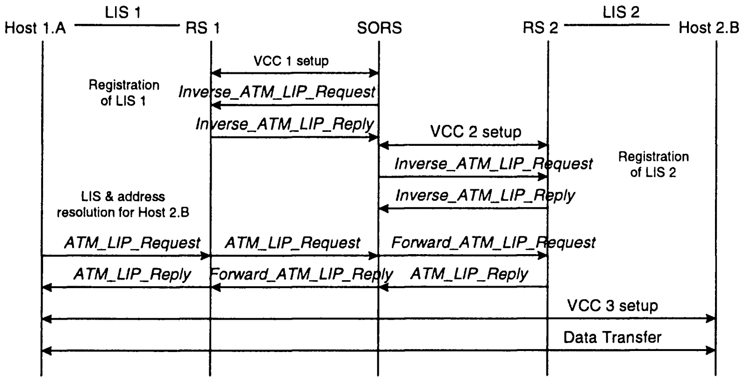



**Figure 4** Procedures of server-base LIP protocol

### 3.3 PVC-base LIP Procedures

If the number of LISs in a RS group or the number of RS groups in an NBMA network is small, each group needs not to use the automatic LIS resolution functions of the SORS. Instead, PVC(Permanent Virtual Channel) between two RSs of two different LISs wishing to communicate each other is used. Therefore, if there is not a PVC between two different LISs, it is impossible to communicate between hosts within different LISs. This means all RSs within an NBMA network need to be full-meshed by PVCs to provide communications between different LISs.

A RS must have a mechanism(e.g. manual configuration) for determining what PVCs it has. Once a PVC is set up, each RS transmit an *Inverse_ATM_LIS*

*_Request* on the established PVC and requests the peer RS's IP and ATM addresses, which it stores in its ATM LIS table with VPI/VCI of corresponding PVC.

Like the server-base model, whenever each host generates a packet with the destination IP address belonged to different LIS, it check the ATM ARP table if there is an address mapping. If there is, the host can setup a connection to the corresponding ATM address using signaling procedures. However, if not, the host transmits an *ATM_LIP_Request* to the RS of the LIS. The RS receiving an *ATM_LIS_Request* from a host check first if there is a mapping between the subnet address of the destination IP address and PVC on its ATM LIP table. If there is, the RS transmits an *ATM_LIS_Request* to peer RS at the other end of PVC. If not, it responses with an *ATM_LIP_NAK*. The RS receiving an *ATM_LIP_Request* from peer RS would then response with an *ATM_LIP_Reply* if an address mapping is found on ATM ARP table. If not, it returns an *ATM_LIP_NAK*. Then, if the host which sent *ATM_LIP_Request* to its RS receives an *ATM_LIP_Reply* from the RS by above procedures, it stores IP and ATM addresses of the destination on ATM ARP Table and setup a connection to the corresponding ATM address using signaling procedures. Figure 5 shows a general packet flow of PVC-base LIP protocol.

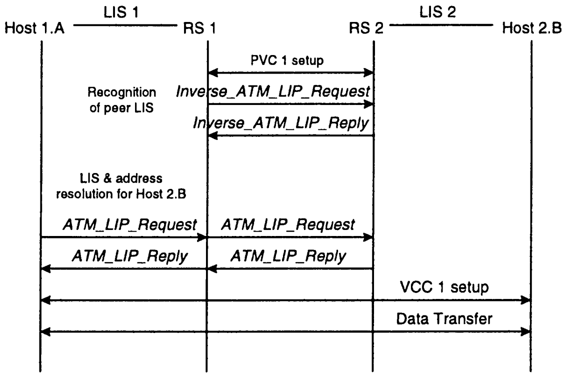



**Figure 5** Procedures of PVC-base LIP protocol

## 3.4 LIP Packet Formats

The LIP packet format is exactly equal to the ATM ARP packet format of Figure 6, except that the operation type values are added for following packets: *ATM _LIS_Request, ATM_LIS_Reply, ATM_LIS_NAK, Forward_ATM_LIS_Request, Forward_ATM_LIS_Reply, Forward_ATM_LIS_NAK, Inverse_ATM_LIS_Request, Inverse_ATM_LIS_Reply.*

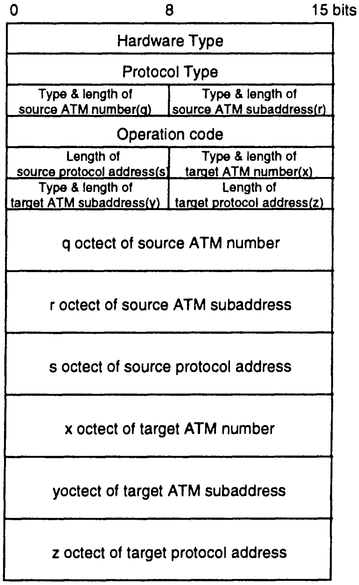



**Figure 6** LIP packet formats

One more difference is the target ATM number of *Forward_ATM_LIP_Reply* represents the ATM address of RS which responded with an *ATM_LIP_Reply*. The RS which receives an *Forward_ATM_LIP_Reply* stores the subnet address of the source IP address and the target ATM number on its ATM LIP table and setup a connection to the target ATM number using signaling procedures to reduce the delay by using a direct connection between two RSs without the intervention of the SORS and avoid the performance degradation caused by that the SORS becomes bottleneck. Figure 7 shows the RS-RS direct procedures under server-base LIP protocol.

## 4 IP MULTICAST OPERATIONS

In this section, we describe how the IP multicast services be supported by LIP protocol. IP multicast is based on IGMP(Internet Group Management Protocol)(Deering, 1989)(Fenner, 1995) and DVMRP(Distance Vector Multicast Routing Protocol)(Waitzman, 1988). IGMP allows IP hosts to join and leave multicast groups. DVMRP is used as multicast routing protocol in multicast router systems.

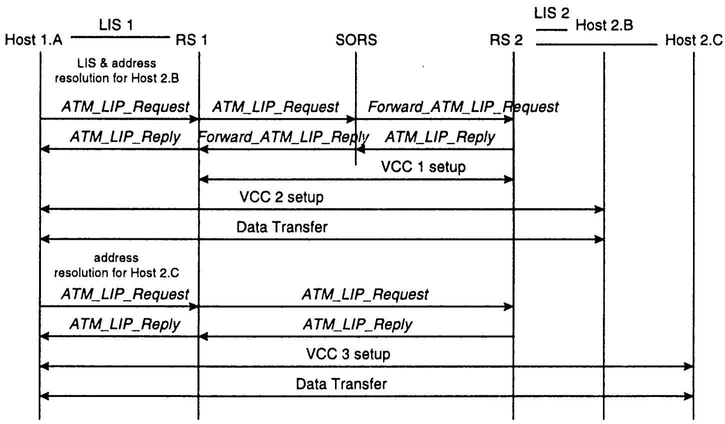



**Figure 7** RS-RS direct procedures of server-base LIP protocol

## 4.1 IP Multicast over ATM based on MARS concept

There are several possible approaches(Braun, 1996) to provide an IP multicast service over ATM but we assume that the Multicast Address Resolution Server(MARS) concept is employed within an LIS to join or to leave an IP multicast group. The MARS concept is theoretically independent of the supported network (layer 3) protocol(Armitage, 1996) and basically an extension of the ATM ARP server defined in RFC 1577. New message formats and a new protocol are defined for multicast address resolution by IETF's Internetworking Over NBMA group.

A MARS keeps a cache of {Multicast IP address, ATM address 1, ATM address 2, ... , ATM address N} mappings. A sender that wants to send a multicast packet to a certain group sends a *MARS_REQUEST* message to the MARS in order to get the corresponding set of ATM addresses for specified IP multicast address. The MARS responds with one or more *MARS_MULTI* messages containing the complete set of ATM addresses {ATM address 1, ATM address 2, ... , ATM address N} for the requested group. The sender stores the mappings in its local cache and establishes a multipoint-VC to all ATM endpoints specified in address list.

For joining and leaving a multicast group, an endpoint has to send *MARS_JOIN* and *MARS_LEAVE* messages to MARS indicating its ATM address and the multicast addresses of the groups to join or to leave. The MARS forwards information about joining and leaving endpoints over a control VC. Senders analyse the information and update their mappings in the local cache.

The MARS concept supports both an approach based on VC meshes and a server based approach as shown at Figure 8. In the VC mesh approach, a multicast sender establishes direct multipoint VCs to the local ATM endpoints. Multicast packets within the LIS are sent directly from the sender to all receivers. Depending on group membership information distributed by the MARS protocol each sender maintains its own multipoint VCs. In the server based apporach, multicast packets are sent to one or more Multicast Server(MCS) nodes. In contrast to a list of ATM endpoints addresses, the *MARS_ MULTI* message of the MARS contains a list of MCS addresses. A sender establishes a multipoint VC to the specified MCSs and sends the multicast packets to the MCSs via this multipoint VC. Then, the MCSs distribute the multicast packets to the receivers of LIS. Several Issues of current specification are still open such as the existence of more than one MCS for a certain multicast group and synchronization between several MCSs.

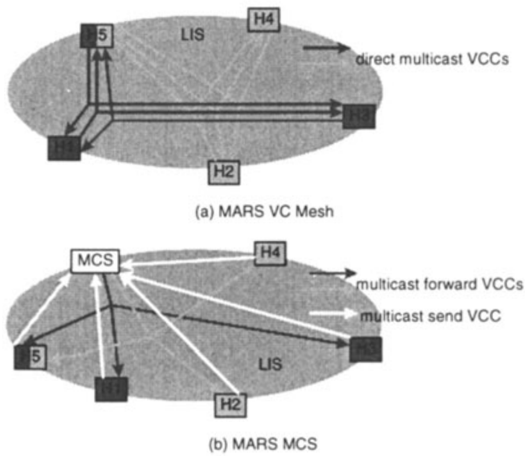
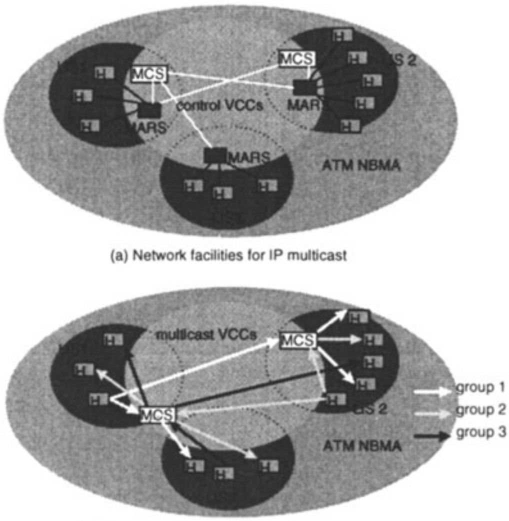



**Figure 8** IP multicast based on MARS

## 4.2 Support of IP Multicast between LISs

In this section, we describe how the LIP protocol support to IP multicast between several LISs. To do this , we assume that basically we employ the MARS MCS model approach(Talpade, 1996).

In the VC mesh approach, the increased VC usage leads to greater consumption of resources like memory for maintaining state, buffer allocation per VC, and the VCs themselves, which may be a scarce and/or expensive resource. A cluster size is currenly dictated by the LIS size, which in turn is dictated by the unicast subnetting that has been applied. An administrator might use knowledge of the mesh mode's VC consumption to pre-emptively subnet his network into small subnets, and thereby end up with more Inter-LIS devices(IP routers) between LISs. Group membership changes cause a decreased level of signaling load to be generated at the switch in the MCS approach. This is because only the MCS has to add/delete a cluster member from the point-to-multipoint VC. The other VCs are not affected. Thus signaling requests only occur at the UNI between the MCS and the switch, as opposed to occurring at the UNIs between all the sources and the switch in the VC mesh case. This is especially beneficial for large groups with several senders, or when the group is highly dynamic, or when the links between the switch and the cluster members are error-prone, which may cause group members to be temporarily dropped from the multicast group, thus making the group more dynamic than it actually is.




(a) Network facilities for IP multicast

(b) Point-to-multipoint VCCs for IP multicast groups

**Figure 9** IP multicast model between LISs

The MCS approach provides a centralized control of multicast bandwidth usage over the ATM network. This is useful in cases where policy demands a limit on the share of bandwidth available for multicast purposes. In such a case, the administrator who sets up a cluster member as the designated MCS for a group, can control the rate at which the MCS multicast data. On the contrary, data throughput and end-to-end latency may be adversely affected due to the additional level of indirection introduced by the MCS and the MCS can potentially become a bottleneck and a central point of failur. However, these problems can be solved by using multiple MCSs per group.

As shown at Figure 9, each LIS has a dedicated MARS. And each MARS has the ATM addresses of MCSs which can be connected. A point-to-point VCC is established between the MARS and the MCS to get to know of the current group membership from the MARS. A sender that wants to send a multicast packet to a certain group sends a *MARS_REQUEST* message to the MARS in order to get the set of ATM addresses for attachable MCSs with a *MARS_MULTI* message from MARS. When the MCS receives a multicast IP packet from an endpoint, it sends a *MARS_LIP_REQUEST* message to all connected MARSs in order to get the corresponnding set of ATM addresses for specified IP multicast address. The

MARS responds with one or more *MARS_LIP_MULTI* messages containing the complete set of ATM addresses {ATM address 1, ATM address 2, ... , ATM address N} for the requested group. The MCS stores the mappings in its local cache and establishes a multipoint-VC to all ATM endpoints specified in address list. Figure 10 shows conceptual message flows for IP multicast between LISs.

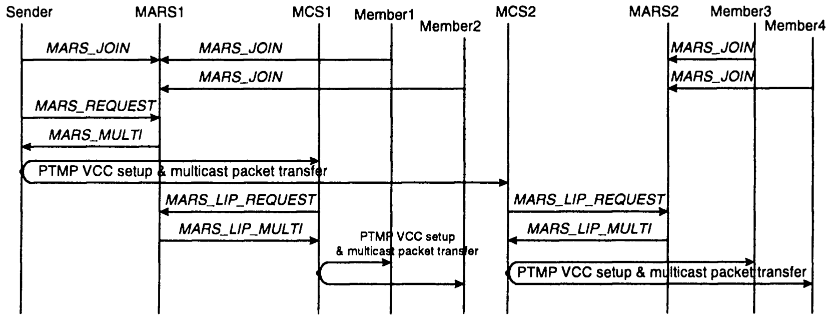



**Figure10** Conceptual message flows for IP multicast between LISs

## 5 SUMMARY

In this paper, we presented a new protocol called LIP(LIS Information Protocol) to support IP routing over ATM. Its key feature is that it gathers information about different LISs using RFC 1577 concept. This is made possible by addressing a certain endpoint as the Route Server of each LIS and by exchanging then LIS information between Route Servers. This scheme provides performance gains by excluding the router in a large cloud, and also reduces the complexity of software needed in the endpoints to support IP routing on ATM network. In addition to providing an IP routing solution on ATM, we also describe how this LIP protocol can be applied to the MASE MCS approach to support the IP multicast service between several LISs. However, the most important advantage is that it can be easily implemented and deployed since our LIP protocol is really compatible with the protocol set specified on RFC 1577.

## 6 REFERENCES

Alles, A. (1995) ATM Internetworking: Cisco Brochures.

Armitage, G. (1966) RFC 2022: Support for Multicast over UNI 3.0/3.1 based ATM Networks.

ATM Forum (1996) Private-Network Node Interface Phase 1.

Braden, P. (1989) RFC-1122: STD3:Requirements for Internet Hosts – Communication Layers.

Braden, R. (1994) RFC 1620: Internet Architecture Extensions for Shared Media.

Braun, T. Gumbrich, S. Stuttgen, H. (1996) Comparison of Concepts for IP Multicast over ATM: IEEE ATM'96 Workshop, San Francisco.

Deering, S. (1989) RFC 1112: Host Extensions for IP Multicasting.

Fenner, W. (1995) Internet Group Management Protocol, Version 2: Internet Draft.

ITU-T (1993) Draft Text for Q.2931.

ITU-T (1996) Broadband ISDN User Part (B-ISUP): Recommendations Q.2761, Q.2762, Q.2763, Q.2764.

Kats, D. Piscitello, D. (1995) NBMA Next Hop Resolution Protocol(NHRP): Internet Draft.

Laubach, M. (1996) RFC 1577: Classical IP and ARP over ATM.

Talpade, R. R. Armitage, G. J. Ammar, M. H. (1996) Experience with Architectures for Supporting IP Multicast over ATM: IEEE ATM'96 Workshop, San Francisco.

Waitzman, D. Partridge, C. Deering, S. (1988) RFC 1075: Distance Vector Multicast Routing Protocol.

Woodworth, C. Veeraraghavan, M. Shur, D.(1996) Connection Request Protocol(CRP) for IP over ATM," IEEE Globecom'96, London.