Internet Engineering Task Force
INTERNET-DRAFT
draft-salgarelli-issll-mis-00.ps
10 November 1997
Expires: 15 May 1998

L. Salgarelli, Editor / A. Corghi
CEFRIEL/Politecnico di Milano
M. Smirnow / H. Sanneck / D. Witaszek
GMD-FOKUS

Supporting IP Multicast Integrated Services in ATM Networks

Status of this memo

Abstract

This memo presents an integrated, server-based mechanism for the efficient support of the *IP Integrated Services* (IIS) model in ATM networks, namely the *Multicast Integration Server* (MIS) architecture. Instead of viewing IP-ATM multicast address resolution and QoS support separately, the approach in this memo is to consider such issues in an integrated manner. In particular, the MIS architecture defines how a layer-3 setup protocol as RSVP can be mapped to and integrated with a layer-2 multicast address resolution protocol as EARTH - *EAsy Multicast Routing THrough ATM clouds*. With the use of EARTH, several ATM point-to-multipoint connections with different QoS parameters can be associated to a single IP multicast address. An RSVP server (RSVP-S) within the MIS is used to distribute RSVP messages inside the ATM cloud and to set the corresponding QoS state in the address resolution table of EARTH (*setup protocol mapping*). In addition, this memo defines a *quantized heterogeneity* model which supports, together with the MIS, advanced IIS features as QoS heterogeneity and dynamic QoS changes in IP-ATM networks.

Editor's note

The postscript version of this memo contains figures that are not included in the text version, so it would be best to use the postscript version. Figures will be converted to ASCII in a future version.

Contents

1 Introduction

1.1 Motivation

The main problem of existing architectures for the support of IP in ATM networks is that they lack a solution for the integration of IIS with ATM and IP Multicast with ATM. Instead of trying to focus on the two problems separately, a generalized, integrated approach for providing mapping of IP multicast flows to ATM pt-to-mpt connections for both best effort and QoS multicast flows as well as RSVP control traffic should be designed. Main advantages of the integrated approach over existing architectures should include, but not be limited to, increased efficiency, reduced layer-3 processing, reduction of the amount of identical data passing the switch several times, reduced VC consumption, reduced reservation establishment delay, and finally layer-2 vs. layer-3 protocol independence.
In this memo we consider such an integrated solution, trying to merge RSVP [1] with the EARTH protocol [2]. We call this solution Multicast Integration Server (MIS). In particular, this memo focuses on the definition of how a layer-3 setup protocol like RSVP can be mapped on a layer-2 protocol as EARTH (*setup protocol mappings*), in order to design an integrated architecture. This memo bases on [3] for what is related to *service mapping issues*.
We are considering EARTH, not MARS [4], not because of differences in address resolution between the two (the multicast ARP support in EARTH mainly follows perfectly specified parts of that in MARS), but because of architectural advantages of EARTH - support for source specific and quality of service specific groups, support for multicast shortcuts (bypassing IP routers) over the entire ATM cloud which is in this case considered to be a Multicast LIS (MLIS). Benefits of EARTH are becoming even more clear when one starts considering IP multicast together with the resource reservation over ATM.

1.2 Scope

The architecture presented in this memo is open (i) to any resource reservation protocol based on receiver initiated reservations and soft state approach, and also (ii) to any IP multicast over ATM address resolution protocol based on Multicast ARP (MARP) server, supporting QoS in its MARP cache, transporting QoS objects to hosts within ATM network, and finally providing shortcuts.
However, in order to keep this memo readable and short we base the specification on RSVP and EARTH coexistence within the ATM network. Additional publications referenced below cover more details. The paper [5] describes all design alternatives for the proposed architecture, while this memo concentrates on a chosen one. Further details on the solution presented here could be found in [6], while details on EARTH could be found in [2].
We assume ATM as defined by UNI 3.1.

1.3 Architectural Overview

The Multicast Integration Server is a specialized node in an ATM cloud (MLIS in EARTH terminology). A MIS is composed by an EARTH server (EARTH-S), for IP class-D to ATM NSAP address resolution and layer-2 QoS management, and of an RSVP server (RSVP-S), for layer-3 QoS management. The ATM NSAP address of the MIS is a well-known address to each IP node on MLIS. Each participant to an IP Multicast session within MLIS, either sender or receiver, opens a control VC to the MIS. The control VCs are used to exchange EARTH and RSVP protocol messages between IP nodes and the MIS.
On one hand, EARTH protocol messages are exchanged between EARTH clients (IP end-systems) and the EARTH-S on the MIS for dynamically resolving IP Class-D addresses to ATM NSAP addresses.

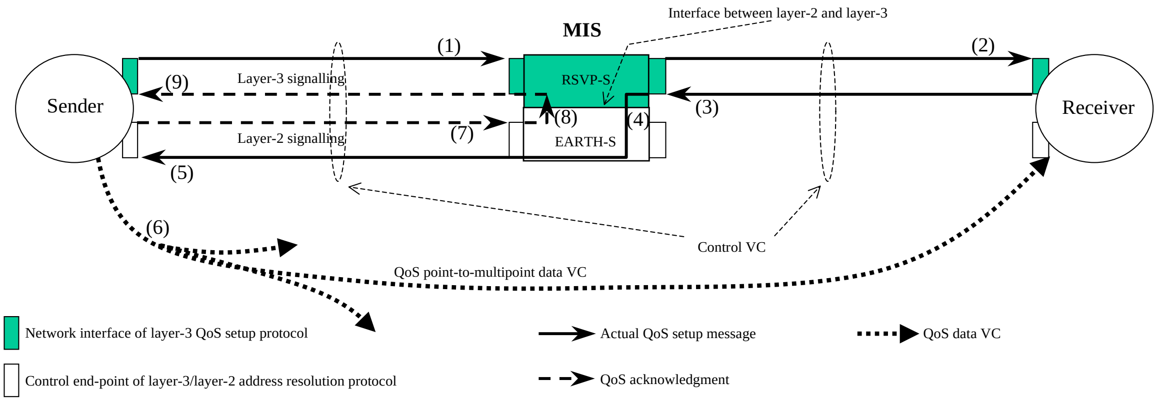

Figure 1: “Layered” view on QoS setup message distribution and setup protocol mappings in the MIS architecture.

On the other hand, RSVP protocol messages are exchanged between RSVP entities[1] (IP end-systems) and the RSVP-S on the MIS, for distributing information about the requested QoS.
In particular, `PATH` messages are sent from sender(s) to the MIS, where the RSVP-S processes, replicates and forwards them to receivers, while all reservations within the MLIS are sent to the RSVP-S, merged and forwarded to the sender(s). This server-based model contrasts to the usual RSVP distributed processing (see also [7] for a server-based approach to enforce reservations on shared/switched Ethernet). More details about the RSVP server approach followed in this memo can be found later in section 2.1.
However RSVP as well as EARTH protocol semantics remain unchanged from their original specifications [1, 2]. The MIS is the only point where the layer-3 protocol (RSVP) is mapped on the layer-2 technology (controlled by EARTH).

1.4 Operational Overview

In IP over ATM networks, we argue that capacity-based admission control is a layer-2 issue. For ATM that means: trying to setup a new VC or adding a leaf to an existing VC. Success or failure are beyond the control of the layer-3 setup protocol (RSVP). Therefore the MIS implements a mechanism for remote capacity admission control and actual setup of layer-2 data paths triggered by EARTH.
Figure 1 gives a generic overview about the layer-3 and layer-2 signalling involved for the setup of a QoS VC. In brief, the sender, using layer-3 signalling (RSVP), informs the MIS (1) about its traffic profile, while the MIS (2) forwards this information to receiver(s). Receivers request resource reservation (3). These requests are mapped from layer-3 to layer-2 within the MIS (4), and forwarded to the sender (5) as layer-2 signalling (EARTH). Sender (6) sets-up point-to-multipoint VC(s) according to the requested reservation, and (7) acknowledge layer-2 reservation(s) to the MIS. Within the MIS (8), layer-2 acknowledgment is mapped to layer-3, and finally (9) layer-3 acknowledgment is forwarded to the sender.

The shown interface between RSVP-S and EARTH-S at the MIS is the only point where QoS-relevant information passes from layer-3 to layer-2 and back, i.e. the MIS is the only location where layer-3 to layer-2 setup protocol mapping and QoS translation take place. This design feature is essential for the protocol independence and avoidance of un-coordinated operation in the MIS architecture.

[1] We call “RSVP entities” the standard implementations of RSVP residing on each IP node of MLIS, excluding the MIS, e.g. the RSVP daemons on UNIX workstations.

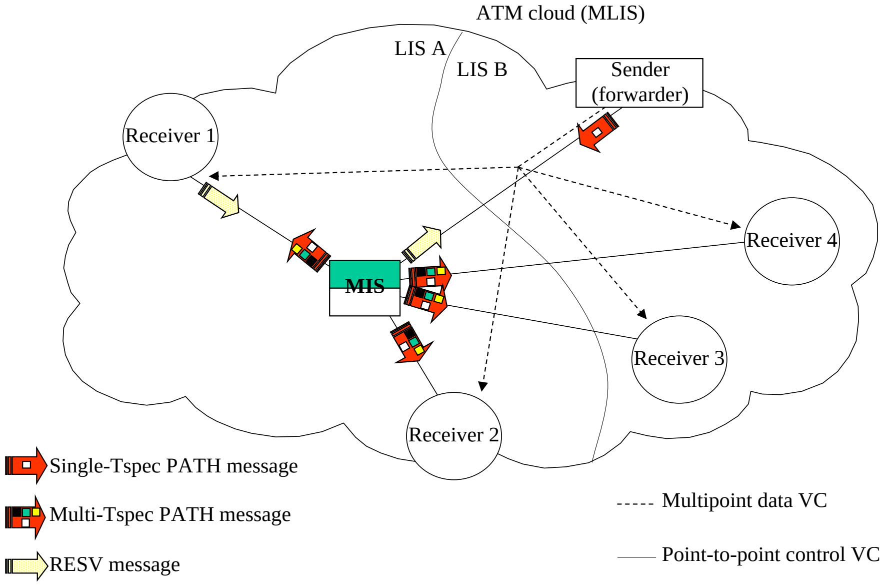

Figure 2: Distribution of `PATH` and `RESV` messages using MIS: a server-based approach to RSVP.

It should also be noted that on failure of a QoS VC setup, no reservation merging takes place and the sender needs not to be notified about the rejected QoS setup. Reservation rejection can already take place at the MIS before passing QoS information to layer-2, e.g. because of policy reasons.

The detailed functional description of the mechanism briefly introduced in this section will be presented in section 2.2.

2 Multicast Integration Server

2.1 Relation of layer-2/layer-3 protocols

The architecture presented in this memo is *server-based*, and integrates the EARTH protocol [2] (layer-2) and the RSVP protocol [1] (layer-3).
The core of the architecture is its server, the Multicast Integration Server. The MIS is composed by an EARTH server (EARTH-S) plus an RSVP server (RSVP-S). The EARTH-S implements the whole set of functionalities foreseen by the EARTH specifications [2]. On the other side, the RSVP-S is a slightly modified implementation of a normal RSVP protocol instance. RSVP-S differs from a standard RSVP implementation for the characteristics described in the following.

Server-based approach. All RSVP protocol messages related to MLIS have to pass through RSVP-S, which process and eventually modifies them before redistributing them to end-stations or Mrouters, as shown in figure 2.

Processing of `PATH` messages. RSVP-S collects all `PATH` messages coming from end-stations or Mrouters on MLIS, processes and replicates them, and redistributes them to all participants in a given RSVP session. Before redistribution, RSVP-S can enforce administrative or security policies on `PATH` messages.

Processing of `RESV` messages. RSVP-S collects all `RESV` messages coming from end-stations or Mrouters on MLIS, processes and merges them, and forwards them to the proper sender(s) of the session. Before redistribution, RSVP-S can enforce administrative or security policies on RESV messages.

Resource reservation policy enforcement. In order to limit the number of ATM VCs needed to support every multicast session, but nevertheless providing a (limited) support for reservation heterogeneity, the RSVP server approach proposed in this memo provides a mechanism for supporting a limited number of QoS levels (i.e. a limited number of point-to-multipoint VCs) per RSVP session.

Given a traffic profile (sender's Tspec) generated by a source, the RSVP-S at the MIS can associate a certain number of QoS levels to such profile, basing on policy information, statistics, or other mechanisms. As the standard PATH message arrives at the MIS, the RSVP-S decides which QoS levels are allowed for the related RSVP session, and appends an ADSPEC object containing one Tspec for each allowed QoS level, before re-distributing the message to receivers (refer to figure 2).

In addition, RSVP-S can enforce a particular policy on RESV messages before forwarding them towards senders, e.g. RSVP-S can reject reservations to a given set of receivers, or can refuse reservations of a given type.

It is out of the scope of this document (i) the definition of the algorithm for the generation of multiple Tspec from a single given Tspec[2], and (ii) the definition of the format of the ADSPEC object which contains multiple Tspecs. However the structure of such an object could be very simple, as a simple sequence of TOKEN_BUCKET objects [8], one for each allowed QoS level.

Protocol mapping of RSVP with EARTH is realized through two local interfaces provided by the EARTH-S: the *Quality of Service Support Interface* (QSSI) and the *earth Routing Support for Resource Reservation* (eRSRR) interface. QSSI provides a mean for the RSVP-S to communicate to the EARTH-S the desired QoS for a given (set of) receiver(s) and for EARTH-S to inform the RSVP-S about success or failure of layer-2 admission control (see section 2.2). eRSRR is used by RSVP-S in order to get from EARTH-S information (i.e. IP/NSAP addresses) about receivers participating in a multicast session, e.g. when RSVP-S has to replicate and re-distribute PATH messages. QSSI and eRSRR will be described in more detail in section 4.1.

EARTH and RSVP protocol messages are exchanged between clients and the MIS by way of point-to-point best-effort ATM VCs (*control VCs*) between each client and the MIS.

Clients of the MIS are end-stations and Mrouters on MLIS that want to participate in an IP over ATM multicast session. Each client must have a standard instance of the RSVP protocol running, plus an EARTH client. The EARTH client implements exactly the features foreseen by the EARTH specifications [2]. The RSVP entity on each client implements all the standard functionalities foreseen by RSVP specifications [1], plus a particular implementation of the Traffic Control Interface (TCI: [1, par. 3.11.2]). This TCI is a quasi-dummy TCI, in the sense that actual resource reservation is performed locally on senders at layer-2 by EARTH, as described later in section 2.2. On end-stations (or Mrouters) no particular communication between the local EARTH client and the local instance of RSVP is needed.

Note that MIS operations are concerned only with the management of control messages (EARTH and RSVP). The MIS itself *is not* a MultiCast Server [4], and has nothing to deal with data distribution, that is performed by the standard ATM infrastructure by means of point-to-multipoint VCs.

2.2 Functional description: setup protocol mappings

Functionally, the MIS architecture implements the RSVP protocol for layer-3 resource reservation, and EARTH as a layer-2 multicast address resolution protocol with QoS supporting features. In

[2]The algorithm could be in principle even a statically defined table mapping a set of allowed Tspecs for each "class" of possible Tspec.

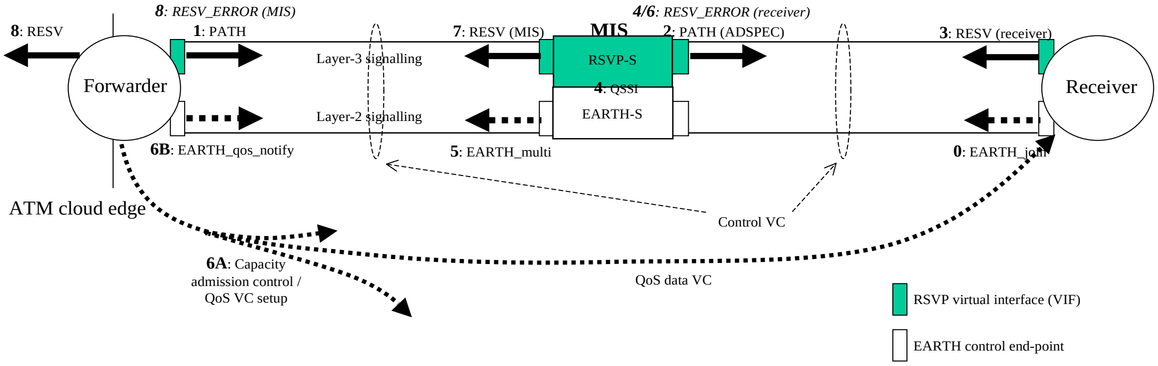

Figure 3: "Layered" view on message distribution within the MLIS.

IP-over-ATM networks, capacity-based admission control is a layer-2 issue, i.e. capacity-based admission control is the operation concerned to the capability of setting-up a new VC or adding a leaf to an existing VC. Using UNI3.1 this operation takes place at the sender(s). However, since is the MIS, not the sender(s), which implements policy admission-control, our architecture adopts a mechanism for *remote capacity admission-control*, including conversion of QoS information from layer-3 to layer-2 format and actual setup of layer-2 data paths through EARTH. All the relevant information needed by the admission-control procedure pass from the RSVP-S to the EARTH-S and back through the QSSI, as shown in the following of this section.
After the preliminary phase (not completely shown in figure), where:

- receivers join the multicast session by registering their NSAP addresses to the EARTH server in the MIS (**phase 0**);
- senders retrieve such information, querying the EARTH-S (with the `EARTH_request` message);
- eventual `PATH` messages from outside MLIS reach the forwarder(s),

operations go on as follows:

Phase 1, layer-3: `PATH` message(s) are forwarded to the MIS.
The RSVP daemon in the sender, or re-sender (MRouter), sends `PATH` messages towards the MIS.

Phase 2, layer-3: re-distribution of `PATH` messages.
At the MIS, `PATH` state is established and an ADSPEC with pre-configured ATM QoS levels is appended. RSVP-S obtains from EARTH-S the addresses of registered receivers through the eRSRR interface, then the `PATH` message is replicated and forwarded to registered receivers through each point-to-point control VC.

Phase 3, layer-3: `RESV` messages sent to the MIS.
Receivers send `RESV` messages upstream[3] to the MIS. The RSVP-S performs policy admission control and RESV state is established, but no `RESV` message is yet forwarded upstream (*pending* admission control).

Phase 4, layer-3 to layer-2: notification about requested QoS.
The translation of layer-3- to layer-2-QoS takes place in the RSVP-S and the EARTH-S is notified about the QoS requested by receivers through the QSSI.

[3]upstream - from receiver to sender.

Phase 5, layer-2: notification of QoS to sender(s).
Upon receipt of the next request (`EARTH_request` message) by the sender, a message containing address resolution information (`EARTH_multi`) including the new requested ATM QoS parameters is sent upstream by the EARTH server at the MIS.

Phase 6A, layer-2: remote capacity admission control (on sender(s)).
The EARTH client at the sender triggers ATM signalling in order to establish a (leaf to a) QoS VC. It sets the received state and gets a call-back about the success/failure in the establishment of the (leaf to) QoS VC from the ATM layer. Note that, supposing successful admission control, at this point, data might begin to flow on the QoS VC.

Phase 6B, layer-2 to Layer-3: remote capacity admission control (on MIS).
The success/failure report of the establishment of the (leaf to) QoS VC is sent back to the EARTH server (`EARTH_qos_notify`). On receipt of the `EARTH_qos_notify`, the EARTH server at MIS informs the RSVP server (through the QSSI) about success or failure of the layer-2 admission procedure.

Phase 7, layer-3: `RESV` message to sender(s).
Supposing successful policy and remote capacity admission procedures, the `RESV` message is forwarded (with next hop object set to the MIS' address) to the upstream sender/MRouter, and pending admission control is resolved. If other reservations for this sender arrive, they have to be admitted as described above, but only one `RESV` message is forwarded within the usual RSVP timer intervals (i.e. [session, sender, QoS-level]-merging takes place as defined by the usual RSVP merging procedure). The RSVP server regards the reservation as 'installed' on his 'virtual interface' (vif) towards the receiver, which is actually its control VC to this receiver. Through this vif, RSVP messages are sent and received, however the real reservation takes place at layer-2 in the (re-)sender.

Phase 8, layer-3: re-distribution of `RESV` messages outside MLIS.
After receipt of the `RESV` message at the sender, the usual state consistency check is performed. Capacity admission control is always successful (as it has already been performed at layer-2) and the `RESV` message is forwarded upstream, if needed, outside MLIS[4].

Phase 4/6/8, layer-3: handling of reservation errors.
On failure of the policy admission (**phase 4**) or capacity admission (**phase 6**) control, a `RESV_ERROR` message is sent downstream to the requesting receiver only. The sender is never notified that a `RESV` message of this receiver has arrived. Thus, it is not necessary to convey RSVP messages over two hops to finally be notified about a reservation failure, as it would be when capacity admission control is managed by RSVP at the sender.
Some RSVP protocol processing failure results in an `RESV_ERROR` message sent back to the MIS (**phase 8**), which in turn leads to deleting the EARTH server QoS state for this sender for all receivers and, after the next EARTH refresh, to teardown of the entire QoS-VC. This occurs only if the PATH state of sender and MIS somehow differ, e.g. when a new `PATH` and a `RESV` referring to an older PATH state are concurrently transmitted.

2.3 Internetworking

Internetworking between each MLIS and the rest of the Internet is concerned with two aspects: the layer-3 setup protocol and the layer-2 multicast address resolution protocol.
On one hand, the MIS architecture adopts standard RSVP protocol messages. `PATH` messages incoming in MLIS are treated by RSVP-S as `PATH` messages generated by a node on MLIS. Also

[4]This is needed if the sender is actually a MRouter/re-sender, and thus the session has senders outside MLIS.

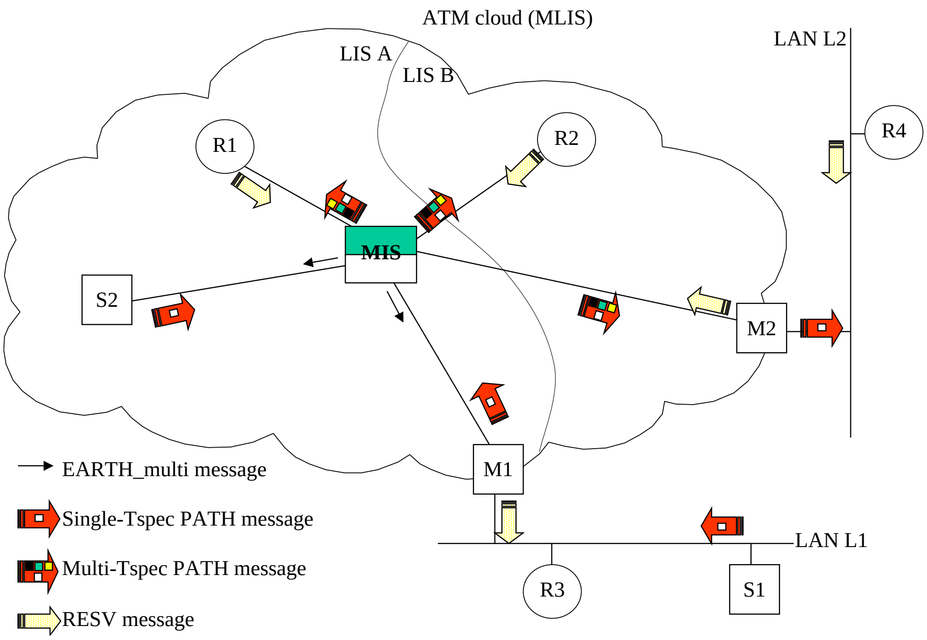

Figure 4: Internetworking: example.

`PATH` messages outgoing from MLIS are standard RSVP `PATH` messages. In this case, additional ADSPEC objects defining allowed (policed) QoS levels for a given session that are used within MLIS are converted by Mrouters to single-Tspec `PATH` messages before forwarding them outside MLIS. The single-Tspec object could be generated using the highest Tspec of the ones contained in the multi-Tspec `PATH` message.
At the same time, reservation incoming from outside MLIS are bounded by Mrouters to the nearest QoS-level available for that session, before forwarding them to the MIS. `RESV` messages forwarded by Mrouters outside MLIS are standard RSVP reservation messages. The MIS architecture foresees the use of all the other RSVP protocol messages without modifications.
On the other hand, internetworking issues related to the multicast address resolution protocol are efficiently addressed by the EARTH protocol, which specifies interworking with non-ATM clouds via DVMRP and via SCSP in case of multiple EARTH servers in large ATM clouds.

2.4 Example

In order to describe how the MIS architecture works, in this paragraph a complete and accurate description of an extensive case-of-study is presented. Let's consider the example that is shown in figure 4. For a given multicast session there are two senders and four receivers distributed on different networks. One of the two senders is on the external LAN L1 (we assume that both the LAN L1 and the LAN L2 are Ethernets), while the other one is inside the ATM cloud. The receivers are distributed on all the three networks: R1 and R2 are on the ATM cloud, while R3 and R4 are on the external LANs. Multicast router M1 is the ingress router for the flow generated by S1 and the egress router for the flow generated by S2. Multicast router M2 acts like egress router for both these data flows. All the receivers join a determined multicast session, namely $MS1$, using IGMP, on the LANs, and EARTH inside the MLIS. The inter-operation between these two protocol works according to the EARTH specification [2]. Once all the receivers have joint the multicast session and the senders have solved the multicast IP address in the proper set of NSAP addresses two point-to-multipoint best effort connections are opened inside the ATM cloud: the former originating from the sender S2

to the set <R1,R2,M1,M2> and the latter from M1 to <R1,R2,M2>. Moreover, M1 forwards the data packets from S to the LAN L1 and M2 forwards all the data packets (both from S and from M1[5]) to the LAN L2. No reservations have been still requested, so the QoS functionalities of the EARTH protocol have not yet been employed.
The EARTH table (MARPTable) of the MIS contains the information on all the receivers as best-effort receivers for the multicast session MS1. The senders, following the standard RSVP specifications [1], transmit the RSVP `PATH` messages to the multicast group. The `PATH` messages originated at S1 are received by R3 and M1. The receiver R3 receives directly the `PATH` message and it requests the desired reservation by sending a `RESV` message up to the sender. Since M1 is a Mrouter, it has to forward the `PATH` message to the ATM end-points that have joint the session and it will ask for a reservation according to the requests of the ATM receiver. According to the model described in the paragraph 2.2, the `PATH` message is sent to the MIS, where the RSVP server takes care of replicating and properly forwarding the messages to the joined receivers. The egress Mrouter M2 in turn forwards the message to the external LAN L2. The `PATH` messages originated at S2 are delivered following the same rules; in this case, both M1 and M2 acts like egress routers toward the external LANs.
Outside the ATM cloud the reservation of resources takes place as defined in [1, 9]. On the contrary, inside the ATM cloud, the MIS approach is applied. As a consequence, the external receivers (i.e. R3 and R4) send their reservation requests to the corresponding previous hops (respectively M1 and M2) and the reservation takes place between the Mrouter (it acts like a re-sender) and the receiver. On the contrary the reservation inside the MLIS follows the rules that have been presented in the paragraph 2.2. All the ATM receivers send the `RESV` message to the MIS[6], where both the RSVP tables and the EARTH tables are updated and the pending reservation state is set. The senders in the MLIS will be informed about the pending reservations through the `EARTH_multi` message and then try to set up the corresponding QoS point-to-multipoint VCs directly to receivers.

3 Service model: Quantized Heterogeneity

Although any service model for IIS over ATM can be efficiently supported by the MIS architecture, in this paper the *quantized heterogeneity* (QH) model is proposed as the advanced service model offered by the MIS.

The QH model represents a compromise between the "full" and the "limited heterogeneity" models specified in [10], by supporting a *limited number* of QoS levels, included the best-effort class, for each multicast session. Since each ATM point-to-multipoint VC provides a single QoS level, a limited number of ATM point-to-multipoint VCs per session are allowed. As shown in figure 5, each sender of a multicast session can open one best-effort point-to-multipoint connection and a limited number of QoS point-to-multipoint connections with different QoS levels. As a consequence, each receiver can choose the desired QoS level only among the allowed ones. The best-effort point-to-multipoint connection guarantees that each receiver can always join a multicast group even if no resources for a QoS connection are available, as required in [10].

The QH model does not define any guidelines about how to manage information about the allowed QoS levels in a multicast session. Moreover, it does not pose any constraint about the strategy that must be followed in the allocation of a given data stream on the ATM VCs.

The QH model can be efficiently integrated in the MIS architecture using the Multicast Integration

[5] M1 acts like a sender with respect to the ATM cloud; it is actually a re-sender of the data flow originated at S1.
[6] Note that the MIS is actually the previous hop for each ATM receiver; but no reservations take place between the MIS and receiver, as the QoS connections are directly opened from the sender to the receiver.

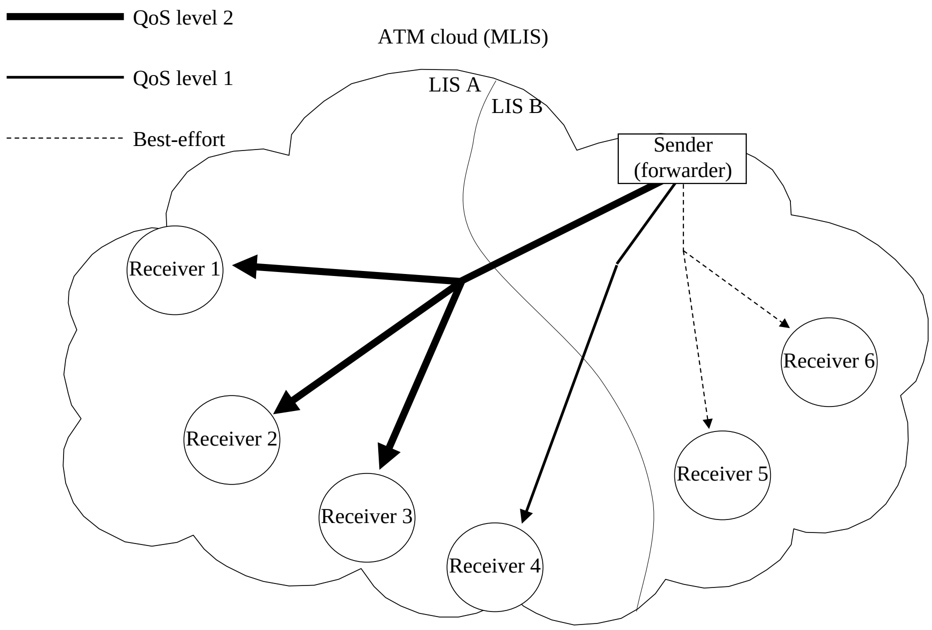

Figure 5: Quantized heterogeneity model.

Server in order to manage (i.e. to define and distribute) information about the allowed QoS levels, using multi-Tspec `PATH` messages. In this case, senders may generate single-Tspec `PATH` messages, while information about the pre-configured allowed QoS levels are appended to outgoing `PATH` messages by RSVP-S at the MIS, enforcing resource reservation policy on the ATM cloud, as described in section 2.1.

4 Specification

4.1 Interfaces

EARTH, being an ATM protocol supporting IP multicast with embedded support for the QoS settings, should work even without any reservation protocol to enable QoS settings. In any case, it would provide an address resolution for best effort flows, supporting minimal QoS functionality in order to allow setup of QoS VCs by manual configuration (UI) or management protocols.

4.1.1 QSSI and eRSRR

The EARTH-S in the MIS architecture supports resource reservation with two local interfaces: QSSI (Quality of Service Support Interface) and eRSRR (earth Routing Support for Resource Reservations). These interfaces define specific messages in order to pass information to/from the EARTH-S. The QSSI provides the `EARTH_QoS_notify` message, used to set the requested QoS information in the EARTH-S. For example, RSVP-S sends this message to the EARTH-S when accepting a reservation. With the same message, sent on QSSI in the opposite direction (from EARTH-S to e.g. RSVP-S), the EARTH-S reports a success or a failure of connection establishment at the sender side.
The routing support interface (eRSRR) gives the possibility to get locally from the EARTH-S information about receivers participating in a multicast session (needed e.g. for the RSVP-S) with two messages: `EARTH_query`, to request an address resolution, and `EARTH_response` returning to the requester addresses of receivers which joined the group.

4.1.2 User Interface

The operation of the MIS can be customized through the User Interface (UI). The UI allows the customization of both the EARTH-S and the RSVP-S.
Through the User Interface the EARTH-S could be customized to support:

- Best effort service only (zero QoS), or also non-zero values for QoS.
- Additional interfaces, like QSSI or routing support interface for interworking with RSVP or with other means for resource reservation.
- Manual configuration of the MARPTable (with the use of the primitives described in 4.2.1).

The RSVP server can be customized with the use of UI for the following topics:

- Support for hierarchically encoded traffic flows or other.
- QoS-levels Generation Algorithm.
- Local policies to handle non-conformant (with regard to traffic description) reservation requests.
- Mapping of QoS levels to ATM QoS parameters used by UNI signalling.

4.1.3 Traffic Control Interface

The generic Traffic Control Interface (TCI) is defined in RSVP specifications [1]. In the MIS architecture, TCI is implemented as follows:

- TCI at sender or forwarder (Mrouter): only a 'dummy' admission control has to be present as the VC setup has been already successful (of course local policies could be enforced, however an admission rejection at this points has a high message/latency overhead).
- TCI at *MIS*: a link-layer dependent adaptation module (LLDAL) has to be added (as for any other link-layer) which interfaces to the EARTH server's QoS Support Interface (QSSI).

4.1.4 Routing Support Interface

The Routing Support for Resource Reservation interface (RSRR) defined in [11], is implemented in the MIS architecture by simply mapping control VC endpoints to RSVP virtual interfaces [1].

4.2 MARP Table

The Multicast Address Resolution (Protocol) Table (MARPTable), hold by the EARTH-S and EARTH-clients, stores membership information, including multicast address resolution as well as QoS information.
MARPTable codes source-specific and QoS-specific information as shown below.

```
MARPTable = null | MARPTableEntry [, MARPTableEntry [, ... ]]
MARPTableEntry = <MultIPAddr, SourceSpecEntryList>
SourceSpecEntryList =  SourceSpecEntry [, SourceSpecEntry [...]]
SourceSpecEntry = <SourceIPAddr, QoSSpecMembers>
QoSSpecMembers =  QoSMembers [, QoSMembers [...]]
QoSMembers = <QoSLevel, MemberList>
MemberList = null | RecId [, RecId [...]]
RecId: = <RecNSAPA, RecIPAdd>
```

where:
`null` - represents an empty (e.g. zeroed) data structure[s];
`a | b | ...` - mutually excluding alternatives a, b, ...;
`a [, a [, ...]]` - one or any number of instances of entity of type a;
<a, b , ... > - a set of elements a, b, ...

Terms used in the description of MARPTable are:

`MultIPAddr` : Multicast IP Address (i.e IP Class D Address).

`SourceIPAddr` : is either IP Address of a source of multicast traffic or must be zero, when multicast group has only registered receivers but no specified source.

`QoSLevel` : a value representing a quality of service level - to be mapped into a quality of service supported by ATM signalling (QoS levels defined in the table `QoSMap`). The value zero means best effort service.

`RecNSAPA, RecIPAdd` : NSAP and IP Addresses of a receiver registered with EARTH.

4.2.1 MARPTable manipulation

The messages defined by the EARTH specification, as well as the ones used in QSSI, eRSRR and UI are mapped in a set of primitives which manipulate the MARPTable: `p_join`, `p_leave`, `p_get_req`, `p_get_resp`, `p_set_qos` and `p_qos_ack`.

For example the primitives `p_get_req(flag, MultIPAddr, SenderIPAddr [, SourceIPAddr])` and `p_get_resp(flag, MultIPAddr, SourceIPAddr, QoSSpecMembers)` describe request for the entry for a given group (`MultIPAddr`) and the corresponding answer. The flag indicates whether all the QoS subgroups should be returned or the best effort only. The answer (`p_get_resp`) depends on the address of the sender (`SenderIPAddr`) in the request; i.e. only sender relevant (and only for allowed sender) information will be made available. They are mapped into EARTH protocol messages `EARTH_request`, `EARTH_multi` as well as defined for eRSRR `EARTH_query` and `EARTH_response`.

The primitive concerning QoS, `p_set_qos(MultIPAddr, SourceIPAddr, prevQoSLevel, newQoSLevel, RecId)` allows to set a QoS level of a receiver. The receiver (`RecId`) is removed from one QoS specific subgroup (`prevQoSLevel`) and added to another (`newQoSLevel`) subgroup. The acknowledgment of this primitive is the primitive `p_qos_ack(MultIPAddr, SourceIPAddr, prevQoSLevel, newQoSLevel, RecId)`, with the same parameter list.

The EARTH protocol message `EARTH_QoS_notify` which arrives from the EARTH client (sender) to acknowledge a success or failure of QoS connection setup is mapped into `p_set_qos`.

The messages provided on the QSSI Interface for the QoS support (used by RSVP or by other means for resource reservations), `EARTH_RESV(MultIPAddr,SourceIPAddr, prevQoSLevel, newQoSLevel, RecId)` and `EARTH_RESV_ACK(MultIPAddr,SourceIPAddr, prevQoSLevel, newQoSLevel, RecId)`, can be mapped into two of above mentioned primitives: `p_set_qos` for the messages received on QSSI interface, and `p_qos_ack` for the message sent on the QSSI interface.

5 Discussion

We argue that the proposed architecture for the Multicast Integration Server should result in increased efficiency, reduced layer-3 processing, reduction of the amount of identical data passing the switch several times, reduced VC consumption, and finally reduced reservation establishment delay, when compared to existing architectures. All these achievements are facilitated by major key design issues collected below from the above text and commented from a more generalized perspective.

RSVP server. Besides the issues bulleted in section 2.1 it should be noted that the separation of data and control paths over ATM seems to be inevitable for any IP control protocol supporting multicast due to the unidirectional nature of ATM pt-to-mpt connections. In the MIS architecture the RSVP has additional benefits arising from the fact that it reuses control VCs being maintained by EARTH clients. To distinguish between EARTH and RSVP control messages we propose a simple dispatching module in the MIS (refer to [5]).

Why EARTH not MARS. The current IETF ION proposed standard RFC 2022 [4] does not support QoS and retains the Classical IP over ATM approach to separate the ATM cloud to a number of logically disjoint subnets (clusters) which must consist of members of a single LIS only [4, chapter 8, page 54]. However, the Classical IP over ATM [12, 13] is an IP unicast protocol over ATM. Complete mapping of IP multicast service model to the ATM multicast service model requires bypassing of IP routers (shortcuts) in order to take the advantage of connecting of all the ATM endpoints - receivers (whatever LIS they belong) with a single pt-to-mpt VC.

The reduction of VC consumption with MIS (due to shortcuts) in comparison with hypothetical MARS (even if supporting the QoS differentiation) could be estimated as O(Q*N), where Q is the number of QoS levels and N is the number of receivers.

Another advantage of EARTH (however referring to implementation only) is that it manages the VC establishment via the UNI 3.x signalling in such a way that the multicast data forwarding could be started from the [re-]sender even on non completely established pt-to-mpt VC. This fact brings additional savings to the join/leave/reservation latency, which could be estimated again as O(Q*N).

The MIS architecture is open. The MIS architecture is server-based and layered: it defines specific mapping procedures between RSVP and EARTH. The layer-3 RSVP is a consumer protocol with regard to the address resolution and data handling service provided by the layer-2 EARTH protocol.

In MIS scenario only QoS assured IP multicast requires limited and strictly specified collaboration between RSVP and EARTH protocols through the two interfaces (QSSI and eRSRR) supported by the EARTH; IP unicast could be supported by RSVP unicast over ATM which is combined with the legacy atmarp. The future development of MARS based IP/ATM networks can use the QSSI defined in this draft for both VC meshes and multicast servers.

On the other hand, any other resource reservation protocol can use EARTH's information in order to get receivers' NSAP addresses for an IP multicast session and to use the QSSI in order to set the desired QoS state for any of these receivers.

The MIS is open also for the inclusion of any policy admission control, network management or security modules.

Remote capacity admission control. Because of the separation of IP data and IP control paths over ATM the question of regulating arises between the place where the reservations are collected and merged (MIS) and the place where the capacity admission control is applied (sender(s)). This problem is solved via the modification of the original EARTH protocol (the `EARTH_qos_notify` message was added). Therefore this solution is safe with regard to the "standard" operation of RSVP and provides timely updates for the EARTH data base, which data base could be safely used by other consumer protocols.

Pre-configured ATM QoS levels. We believe that the practically feasible number of QoS levels in IP/ATM internets could not be large. Moreover, the QoS levels which could be used by IP/RSVP flows must be supported by the ATM signalling, i.e. these available levels must be known in advance and the mapping from a particular RSVP reservation to the ATM QoS should be done with some robustness in mind.

The QH model. Due to the connection oriented nature of ATM the renegotiation of the QoS consumes double resources during the transition phase. The quantized heterogeneity model provides significant savings of resources in this case. Supporting only a (configurable) limited amount of QoS levels (i.e. QoS VC) per session, the MIS architecture (i) prevents wasting of network resources unavoidable with the "full heterogeneity model" [10], and (ii) provides more flexibility than the single-QoS "limited heterogeneity" model, in full accordance with the receiver-initiated philosophy of RSVP.

Scalability. Major scalability issues depend on the MIS-centric organization of the MIS architecture. A single server both for multicast address resolution and for QoS management may become a bottleneck if the ATM cloud size grows. On one hand, some solutions to efficiently face the scalability of the EARTH protocol are proposed in [2]. On the other hand, no scalability matters are due to the RSVP protocol, as the distribution of the RSVP signalling messages is organized on the base of the EARTH control connections.

Service mappings. The MIS architecture bases on the ongoing work of the ISSLL IETF working-group for what is related to service mappings issues, in particular on what is specified in [3].

Adaptation protocols. By supporting the quantized heterogeneity model, the MIS is capable of augmenting the native capabilities of the ATM link-layer (single QoS per ATM multicast session), thus providing a set of different QoS levels to each IP Multicast session.

Statement of non-applicability. Currently, due to the inherent limitations of ATM signalling (no multipoint-to-point capability), only Fixed-Filter style reservations (explicit sender selection, distinct reservations) are supported. However a mapping scheme within the MIS, where a Shared-Explicit or Wildcard-Filter reservation is mapped to several separate layer-2 reservations (i.e. QoS-specific entries in the EARTH server's MARP table) can be realized.

Implementation. Two independent implementations of the MIS architecture are being developed at *CEFRIEL/Politecnico di Milano* and *GMD-FOKUS*.

6 Security considerations

Security and policy mechanisms can be enforced for an entire MLIS, at the MIS. However, specific security considerations should be addressed in a separate document dealing with security in IP over ATM networks, especially with the presence of shortcuts. This memo does not imply any new security issues if compared with RSVP over ATM and IP Multicast over ATM, however it amplifies security concerns, through bypassing inter LIS routers and possibly firewalls.
This draft recognizes that MIS (EARTH server) could be configured to allow or not to obtain the NSAP address of specific endpoints/LISs/etc. Security management in EARTH case will mean that it's commonly agreed security management between all LISs delegated to a future security entity within the MIS architecture, e.g. entity managing security of shortcuts.

A Glossary

In this memo the following terms are used with the following meanings:

MLIS: a [static] association of all IP multicast gateways (Mrouters) and [dynamic] association of all ATM endpoints/IP hosts willing to participate in any IP multicast session over an ATM cloud.

Shortcut: an ATM shortcut is a direct ATM connection between a sender and a (set of) receiver(s) in an MLIS, eventually bypassing inter-unicast-LIS routers.

MIS: the acronym *"MIS"* is used in general alone to indicate the *server* of the architecture proposed in this memo. The architecture itself is indicated as *"MIS architecture"*.

Mrouter: this term is used to indicate border multicast routers which provide layer-3 multicast connectivity between an ATM cloud and the rest of the internet.

Forwarder: a synonym for Mrouter.

Re-sender: another synonym for Mrouter.

MARP: Multicast Address Resolution Protocol.

IIS: Internet Integrated Services architecture [9].

EARTH-S: the server of the *EAsy multicast Routing THrough ATM clouds* protocol.

RSVP-S: the server of the RSVP protocol defined and used in the MIS architecture.

QSSI: Quality of Service Support Interface, defined in EARTH specifications [2], and recalled in this memo in section 4.1.

eRSRR: extended Routing Support for Resource Reservation Interface, defined in section 4.1.

TCI: Traffic Control Interface, defined in [1].

LLDAL: Link-Layer Dependent Adaptation Module, used in section 4.1.

References

[1] L. Zhang, R. Braden, S. Berson, S. Herzog, and S. Jamin. Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification. RFC2205, IETF, September 1997.

[2] M. Smirnov. Scalable and Efficient Multiprotocol IP Multicast over ATM. In *Proceedings of SPIE VV'97 - Broadband Networking Technologies*, Dallas - Texas, Nov 1997. SPIE.

[3] M. Garret and M. Borden. Interoperation of Controlled-Load and Guaranteed Services with ATM. Work in progress - Internet Draft, IETF, July 1997. draft-ietf-issll-atm-mapping-03.txt.

[4] G. Armitage. Support for Multicast over UNI 3.0/3.1 based ATM Networks. RFC2022, IETF, November 1996.

[5] A. Corghi, L. Salgarelli, H. Sanneck, M. Smirnov, and D. Witaszek. Design Experiences with IP Multicast Integrated Services over ATM. Unpublished, July 1997. Available at ftp://ftp.fokus.gmd.de/pub/step/papers/mis-alt.ps.gz.

[6] L. Salgarelli, A. Corghi, H. Sanneck, and D. Witaszek. Supporting IP Multicast Integrated Services in ATM Networks. In *Proceedings of SPIE VV'97 - Broadband Networking Technologies*, Dallas - Texas, Nov 1997. SPIE.

[7] R. Yavatkar, D. Hoffman, Y. Bernet, F. Baker, and R. Kennedy. SBM (Subnet Bandwidth Manager): A Proposal for Admission Control over Ethernet. Work in progress - Internet Draft, IETF, February 1997. draft-yavatkar-sbm-ethernet-03.txt.

[8] J. Wroclawski. The Use of RSVP with IETF Integrated Services. RFC2210, IETF, September 1997.

[9] D. Clark, R. Braden, and S. Shenker. Integrated Services in the Internet Architecture: an Overview. RFC1633, IETF, June 1994.

[10] S. Berson and L. Berger. IP Integrated Services with RSVP over ATM. Work in progress - Internet Draft, IETF, March 1997. draft-ietf-issll-atm-support-03.txt, .ps.

[11] D. Zappala. RSRR: A routing interface for RSVP. Internet Draft, IETF, November 1996. draft-ietf-rsvp-routing-01.ps.

[12] M. Laubach. Classical IP and ARP over ATM. RFC1577, IETF, January 1994.

[13] M. Laubach and J. Halpern. Classical IP and ARP over ATM. Work in progress - Internet Draft, IETF, October 1997. draft-ietf-ion-ipatm-classic2-03.txt.

Authors' Addresses

L. Salgarelli and A. Corghi
CEFRIEL/Politecnico di Milano
via Fucini 2, I-20133 Milano (Italy)
Voice: +39-2-23.954.1
Fax: +39-2-23.954.254
e-mail: {salga, corghi}@cefriel.it

M. Smirnow, H. Sanneck and D. Witaszek
GMD-FOKUS
Kaiserin-Augusta-Allee 31, D-10589 Berlin (Germany)
Voice: +49-30-34.637.000
Fax: +49-30-34.638.000
e-mail: {smirnov, sanneck, witaszek}@fokus.gmd.de