Internet Integrated Service over ATM: 
a solution for shortcut QoS Virtual Channels

Roberto Cocca(*), Marco Listanti(**), Stefano Salsano(*)
(* CoRiTeL - Consorzio di Ricerca sulle Telecomunicazioni - Rome - Italy
(**) INFOCOM dpt, University of Rome “La Sapienza”

TO APPEAR IN DECEMBER 1999 ISSUE
OF IEEE COMMUNICATIONS MAGAZINE

Abstract
In principle, the interaction of RSVP and ATM should allow to benefit at the IP level from some features of the ATM layer. The most interesting one is the native support of end-to-end Quality of Service provided by the ATM. On the other hand, there are issues that must be clarified to define a correct interworking: for example, the possible overlapping between the mechanisms used in the IP level and in the ATM level to support QoS or the needed IP/ATM address resolution mechanism. This paper proposes a solution to exploit ATM shortcut VCs supporting the QoS in the Internet Integrated Service model. A straightforward enrichment to the RSVP protocol is defined, which only impact the devices (hosts and routers) involved in the shortcut procedure. A mechanism for IP/ATM address resolution is provided, avoiding the use of other mechanisms such as NHRP (Next Hop Resolution Protocol). Special care has been taken to maintain compatibility with “traditional” RSVP hosts and routers.

Keywords:
RSVP, shortcut VC, IP over ATM, QoS
1. Introduction

Despite its tremendous growth, the Internet is still largely based on a very simple service model, called “best-effort”, providing no guarantees on the correct and timely delivery of data packets. Probably, this simplicity has been one of the main reasons for the success of the IP technology. The best-effort service model, combined with efficient transport layer protocol (i.e. TCP), is perfectly suited for a large class of applications, referred to as “elastic”, which can adapt (even dynamically) to different performances offered by the network, in terms of data throughput and end-to-end-delay. Web browsing and electronic mail represent a typical example of elastic applications. Real time applications, like video and audio conferencing, typically require stricter guarantees on throughput and delay. The idea of extending Internet capabilities to provide support to “real-time” applications has led the Internet community to develop the “Internet Integrated Service” (IIS) architecture [1]. The “Guaranteed Service” and the “Controlled Load Service” models have been added to the best effort service model and a signaling protocol called Resource Reservation Protocol (RSVP) [2] has been defined.

A second challenge to the Internet technology is the aggregate switching throughput that should be provided by IP routers. The advances in optical technology allow higher and higher link bandwidths at decreasing costs. In this scenario, the processing capability of the routers could constitute the most relevant bottleneck. Although gigabit IP routers will be available in the next future, currently it is widely accepted that the ATM technology offers potential advantages both for its capability in terms of aggregate switching throughput and for its native ability to support QoS in both point-to-point and point-to-multipoint VCs. In fact, ATM is currently used as an efficient network technology for transport/switching in the backbone networks.

A key point for the future of ATM is the integration with IP technology. Current IP/ATM interworking solutions (LANE [3], CLIP [4], MPOA [5]) do not natively support QoS. They focus on using ATM technology under the classical best-effort IP model.

This paper investigates the interaction of the Internet Integrated Service architecture and the ATM. The rationale for this interaction is described in the RFC 2382 [6]. According to this RFC, there are two main areas involved in supporting the IIS model, QoS translation and VC management. QoS translation concerns mapping a QoS from the IIS model to a proper ATM QoS and it has been extensively dealt with in [7]. This work concentrates on the issue of VC management, dealing in particular with shortcuts and address translation. VC management
considers how many and what kind of VCs are needed and which traffic flows are routed over which VC. Different solutions are possible, for example the first choice is between ATM Permanent and Switched Virtual Channels (PVC, SVC) or a mix of PVCs and SVCs. Another choice is related to the possibility of having a single VC for each IP flow or to aggregate several flows in a single VC. Traditional "hop-by-hop" routing or “shortcut” VCs can be used.

It is worth spending some words on the topic of ATM shortcuts. Within the Classical IP over ATM (CLIP) model, the Logical IP Sub-networks (LIS) are defined as separately administered IP subnetworks. Hosts belonging to different LISs can communicate only by going through an IP router. An end-to-end path between two hosts can be composed of several router hops even though it may be possible to open a direct VC between the two hosts over the ATM network. The possibility to use this kind of direct VCs, referred to as “shortcut” VCs, is proposed in architectures like MPOA [5]. The first problem to be solved to setup a shortcut is to know the identity and the ATM address of the remote end of the VC. Within MPOA a specific protocol, called Next Hop Resolution Protocol (NHRP) [8], allows the resolution of an IP address to the ATM address. Anyway, the basic mechanism allows the redirection of all the traffic directed to a specific IP address to a shortcut VC with no QoS requirements (e.g. using ATM UBR transport class).

The interaction of NHRP with RSVP is still unclear. In section 2 some critical issues related to this interaction are described. These considerations led us to propose a solution for the use of ATM shortcuts in the IIS architecture that is not based on NHRP. The solution avoids the use of IP/ATM address resolution mechanisms by a simple enrichment to the RSVP protocol. The fundamentals of this approach are described in section 3. The specification of the new classes and further details on the procedures to handle these classes in the routers are described in section 4. Section 5 reports about a limited test-bed implementation. The multicast case issues are listed in section 6 and finally the conclusions are given.

2. Issues in the interaction of NHRP and RSVP

The Next Hop Resolution Protocol (NHRP), defined in [8], is used to establish unicast ATM VCs that bypass IP routers. The NHRP provides for the mapping of an IP address to the corresponding ATM address.

There are some issues related to the use of NHRP: for example the “domino effect” (the generation of multiple NHRP messages for the same data packet by subsequent routers), the
possibility of stable routing loops in the router to router case, the scalability issues and the burdensome management. These issues are analyzed in [8] and [9]. Furthermore new problems arise considering a solution for the interworking of NHRP and RSVP. The interaction of these two approaches is pretty complex. The NHRP is not QoS oriented, e.g. the NHRP messages are not able to carry the traffic information to be used at each (NHRP) node as a "hint" to yield the "longest" possible QoS shortcut. Another issue is how to transport the RSVP control messages. These messages usually follow the "hop-by-hop" path according to the IP routing. When ATM shortcut VCs are available the RSVP message could use them, but side-interaction with RSVP logic must be carefully considered.

A first approach could be to use NHRP and the ATM shortcuts whenever possible, including for the transport of RSVP messages. The packet forwarding procedures should be carefully designed considering that the different packets (best effort data packets for different destinations, packets of QoS flows, RSVP control messages) should go into different VCs. A different possible solution consists in using the ATM shortcuts only for QoS flows. Best-effort packets should follow the hop-by-hop path. The NHRP procedures should start only after the reception of the first RSVP RESV message of the session. A drawback of this solution is that unneeded resources in the intermediate nodes are reserved.

In the next sections an architectural solution for the interworking between RSVP and ATM, making use of shortcut VCs, is proposed, but it does not make use of NHRP in order to overcome the problems listed above.

3. Interaction of RSVP and ATM for the support of shortcut QoS VCs

Two new information elements ("classes" in the RSVP terminology) must be added to the RSVP PATH and RESV messages. The related procedures to handle these classes will be described. Special care has been taken to maintain compatibility with "traditional" RSVP hosts and routers. No modifications are needed in the host and in the routers that do not take part in the shortcut procedure.

The procedure for the support of ATM shortcuts is shown in Figure 1. It is assumed that a set of RSVP capable routers are connected by an underlying ATM network. Classical IP over ATM is being run, therefore the packets will follow a hop by hop path. In this scenario, the goal is to determine the "longest" possible ATM shortcut, which directly connects the "Ingress" and "Egress" Router. The proposed procedure can be described as follows:
- The Ingress Router, while sending an RSVP PATH message towards the next hop internal to the ATM network, insert its own IP address using a new RSVP class called “ATM_FHOP_IP_ADDRESS” (FHOP stands for First Hop).
- This information is stored in the PATH state information in each of the subsequent IP over ATM routers in the core network. These routers forward the information unmodified. Therefore also the Egress Router receives and stores the IP address of the Ingress Router.
- The Egress Router, when forwarding the RSVP PATH message on an interface outside the ATM core network, will not include the “ATM_FHOP_IP_ADDRESS” class.
- When receiving an RSVP RESV message, a router checks if an “ATM_FHOP_IP_ADDRESS” information is stored in the relevant PATH state. In this case, the RSVP RESV message is forwarded using as IP destination the stored IP address. According to the RSVP message processing rules, such a message will not be interpreted by the intermediate routers, which will simply forward it up to the Ingress Router. In addition to the other info (e.g. Rspec), a newly defined class is added in this enriched RSVP RESV message in order to carry the Egress Router ATM address for the shortcut. This new class has been called “ATM_LHOP_ATM_ADDRESS” (LHOP stands for Last Hop). In the example in Figure 1 the Egress Router does find the Ingress Router IP address and sends the special RESV directly to it.
- The Ingress Router will receive an RSVP RESV containing the ATM address of the Egress Router in the new class and, as usual, the traffic specification for the reservation. Therefore, all the information (traffic specs and ATM address) needed to setup a QoS shortcut VC is available and the Ingress Router can send the ATM SETUP.

![Figure 1: Enrichments to RSVP for the support of ATM shortcuts](image-url)
The semantic related to the use of the two newly defined classes can be further explained. When an RSVP router (e.g. the Ingress Router in Figure 1) inserts the “ATM_FHOP_IP_ADDRESS” class in an RSVP PATH message, this means that the router is willing to setup a dedicated VC for the flow under consideration. Therefore, the router could choose to add this information only for special flows requiring a particular QoS (for example high bandwidth), which are worth to be mapped in a specific VC. The traditional “hop by hop” procedure can be supported at the same time by the same router for all the other flows. The addition of the Ingress Router IP address in an RSVP PATH is only an “offer” to the next routers, which can choose to use or to ignore this information when sending the RSVP RESV messages. If an Egress Router agrees to establish an ATM shortcut VC, it will “accept” the offer and will insert its ATM address in the “ATM_LHOP_ATM_ADDRESS” class. The information needed by the Egress Router to correlate the incoming ATM connection setup with the RSVP session will be inserted in the BHLI (Broadband Higher Layer Information) information element by the Ingress Router.

After that the procedure has been completed and the ATM QoS channel has been setup, there are two options related to the transport of next “refresh” RSVP PATH messages. The Ingress Router will continue adding the “ATM_FHOP_IP_ADDRESS” class in the refresh PATHs. A first option is to continue using the traditional IP routing as the first PATH. The main advantage of this solution is the simplicity, because nothing has to be added to the traditional RSVP procedures. Another advantage is that changes in the IP routing can be handled without any risk of “loops”. The disadvantage is that the intermediate core routers will “uselessly” process PATH messages and store PATH states for all the lifetime of the RSVP flow. If the requirements in terms of message processing and PATH state storage represent a system bottleneck, the second option is to establish ATM shortcuts for the transport of RSVP messages. In this case, either best-effort (UBR) VCs could be used, or QoS VC with minimal requirements, as the bandwidth needed to transport RSVP PATH messages is very low. If several RSVP flows between two routers are active, these flows can share the same “RSVP control” ATM VC. It is worth noting that the ATM VCs for the transport of RSVP messages cannot be torn down because of a low volume traffic, like in the classic implementations of shortcuts, but you must keep them alive as long as you need to route RSVP messages through them.
In the described scenario, a single core ATM network is represented (Figure 1), but this solution is also able to support more complex network scenarios, where multiple independent ATM networks can be crossed. In such a situation it is possible to get an end-to-end path constituted by several shortcuts (each one crossing a single ATM network), interconnected by means of IP sections. In fact, the Egress router of the first ATM network takes care of removing the ATM_FHOP_IP_ADDRESS when forwarding the RSVP PATH messages, therefore, the proposed procedure can be applied again independently in the second network.

Under particular conditions (i.e. network topologies) the sequence of shortcuts and IP sections which is obtained by the described procedure is sub-optimal. This happens when the IP route (which is followed by RSVP messages) goes out of an ATM network and then enters it again, maybe after having crossed a distinct ATM network. Figure 2 provides an example, showing the optimal ATM shortcut and the sub-optimal sequence of shortcuts (dotted lines) and IP sections (thin lines).

It is possible to enhance the solution in order to find the optimal shortcut in the most general case. The needed enhancement to the procedure is that more than one IP address must be allowed in the ATM_FHOP_IP_ADDRESS class. The first Ingress Router in each ATM network will add its IP address in the ATM_FHOP_IP_ADDRESS class. In this case, the ATM_FHOP_IP_ADDRESS class will be sent also towards IP routers without ATM functionality (as described previously they just forward the unknown classes).

Figure 2: Network topology which can lead to sub-optimal shortcuts
An event trace of the application of the enhanced solution to the network depicted in Figure 2 is shown in Figure 3, where:

- **PATH** message is an RSVP PATH message enriched with an `ATM_FHOP_IP_ADDR` class, to convey the Ingress ED IP address;
- **RESV** message is an RSVP RESV message enriched with an `ATM_LHOP_ATM_ADDR` class, to convey the Egress ED ATM address.

A PATH message without the `ATM_FHOP_IP_ADDR` class reveals the will of tearing down the QoS VC; then the Ingress Router should reroute the flow on a hop-by-hop path.

![Event trace of the proposed solution](image.png)

**Figure 3: Event trace of the proposed solution**

The price to be paid is the need to execute more checks and especially the Egress Router must be able to distinguish the IP address (if present) of the routers which belong to its own ATM network, in order to store the farthest one. Each Egress Router can select the destination where to send the RSVP RESV which allows to setup the longest shortcut.
4. New RSVP classes definition and processing rules

4.1 Definition of new RSVP classes

The definition of the RSVP classes is reported hereafter. Two new classes are proposed: ATM_FHOP_IP_ADDRESS and ATM_LHOP_ATM_ADDRESS. The maximum size of ATM_FHOP_IP_ADDRESS in the IPv4 case has been fixed to 20 bytes, which allows to carry up to 5 different IP addresses. This is needed to enhance the solution, according to the concluding remarks of section 3, while typically only one IP address is carried and the object content length is 4 bytes. The ATM_LHOP_ATM_ADDRESS has a length of 20 bytes as needed to carry the AESA ATM address format.

The choice of the class number has an important impact with respect to backward compatibility. Choosing to have the two highest order bits set to “1” implies that if an RSVP router receives the message and does not recognize the class number, it will forward the object unmodified, without generating any error message. The classes number has been temporarily fixed to the following values, but it could be modified in the future, according to IANA considerations [10].

**ATM_FHOP_IP_ADDR** Class = 208 (binary expression: 11010000)
- IPv4 ATM_FHOP_IP_ADDR Object: Class_Num= 208, Class_Type = 1

**ATM_LHOP_ATM_ADDR** Class = 209 (binary expression: 11010001)
- IPv4 AESA ATM_LHOP_ATM_ADDR Object: Class_Num= 209, Class Type = 1

With respect to performance, the proposed enrichments seem to have a very limited impact on the size of the RSVP message and on the processing and storage requirements in the RSVP routers. On the other hand, the proposed approach seems to provide a straightforward mechanism to setup the “longest” ATM shortcut, without the need of additional address resolution mechanisms like NHRP.
4.2 Processing rules in the routers

In this section the processing rules in the routers will be described. In the following IIS router stands for Internet Integrated Services router that is an RSVP capable router. An IISoA (IIS over ATM) router is an IIS router that further relies on ATM.

The processing rules performed by a router depend also on its "role": in Figure 4 an Ingress and an Egress Router are depicted in the case (a), while an Ingress/Egress router is represented in the case (b).

![Diagram of network roles](image)

**Figure 4: The role of the routers**

The algorithm performed in the routers when they get the first PATH message of each session will be described by means of a flow diagram as shown in Figure 5, where the dashed boxes border steps to be made exclusively in the enhanced solution.

Clarifying comments to the processing rules are given below.

An IISoA router must first “realize” whether it is an Ingress Router for the specified flow. If the PATH message already entered the ATM network (already passed through at least one ATM-capable router) it carries the ATM_FHOP_IP_ADDR class containing the ATM network Ingress Router IP address for the specified flow.

In the enhanced solution a router can be considered an Ingress Router for that flow only if in the PATH message it processes there are no IP addresses of IISoA routers belonging to the same ATM subnet.

In the case of branch (a) (Egress Router) only in the enhanced solution the ATM_FHOP_IP_ADDR class can contain several IP addresses belonging to different ATM
subnets. In an intermediate router the branch (b) will be followed. The branches (c) and (d) are respectively related to the Ingress router and Ingress/Egress router case.

The IIS routers should not receive messages containing new classes, but in the enhanced solution it is unavoidable. An egress router will forward the PATH message without the ATM_FHOP_IP_ADDR class as such information is no longer necessary to subsequent routers in the path.

Further forwarding the PATH message without the ATM_FHOP_IP_ADDR class will allow to repeat the same process as above in each of subsequent ATM subnets.

Figure 5: The PATH msg processing rules in the IISoA routers

The algorithm performed in the routers when they get the first RESV message of each session will be described by means of a flow diagram as shown in Figure 6.
According to the classical RSVP protocol (branch a) a message which is not addressed to the considered node will simply be forwarded towards the destination station, without any processing.

An IISoA router which receives an RSVP RESV message can be only an ingress (branch d) or an egress router (branch b), in fact according to the conceived algorithm the ATM network intermediate routers will never get RESV messages addressed to them.

The branch (c) represents an Ingress/Egress router, which is the only crossed node belonging to a specific ATM network (the PATH message did not enter again the same ATM network after having crossed that IISoA router) and which requires no actions to be performed.

"Refresh" RESV messages will follow the same path of the first RESV message for that session, obviously without creating any further RESV state.

Figure 6: The RESV msg processing rules in the IISoA routers
5. Trial implementation

The enrichments to RSVP have been implemented in a testbed running on a LAN. This section gives an overview of the architecture of the testbed, describing the used equipment and software tools used.

The testbed (see Figure 7) consists of three Pentium PCs running Linux Operating System (kernel 2.1.125) couple interconnected via 10 Mbps Ethernet links. Obviously the two external routers belong to different IP subnets.

![Figure 7: The testbed](image)

The demonstrator has been realized modifying the source code of the Rel.4.2a4 RSVP daemon [11]. The RSVP daemon is complemented in Linux Operating System by the iproute2 package, which cares about traffic control.

During the trial a simulation of both ATM-capable interfaces and just IIS interfaces was needed: in each node a configuration file contains such information for each interface and eventually its own ATM address.

<table>
<thead>
<tr>
<th>Trial Number</th>
<th>Router A</th>
<th>Role</th>
<th>ATM interface?</th>
<th>Role</th>
<th>ATM interface?</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>YES</td>
<td>D</td>
<td>YES/YES</td>
<td>R</td>
<td>YES</td>
<td>S</td>
</tr>
<tr>
<td>2</td>
<td>YES</td>
<td>S</td>
<td>YES/YES</td>
<td>R</td>
<td>YES</td>
<td>D</td>
</tr>
<tr>
<td>3</td>
<td>NO</td>
<td>D</td>
<td>NO/NO</td>
<td>R</td>
<td>NO</td>
<td>S</td>
</tr>
<tr>
<td>4</td>
<td>NO</td>
<td>S</td>
<td>NO/NO</td>
<td>R</td>
<td>NO</td>
<td>D</td>
</tr>
<tr>
<td>5</td>
<td>NO</td>
<td>D</td>
<td>NO/YES</td>
<td>R</td>
<td>YES</td>
<td>S</td>
</tr>
<tr>
<td>6</td>
<td>NO</td>
<td>S</td>
<td>NO/YES</td>
<td>R</td>
<td>YES</td>
<td>D</td>
</tr>
<tr>
<td>7</td>
<td>YES</td>
<td>D</td>
<td>YES/NO</td>
<td>R</td>
<td>NO</td>
<td>S</td>
</tr>
<tr>
<td>8</td>
<td>YES</td>
<td>S</td>
<td>YES/NO</td>
<td>R</td>
<td>NO</td>
<td>D</td>
</tr>
</tbody>
</table>

Table 1: Report of test configurations
The trials are listed in Table 1, where S, R, D stand respectively for Source, Router, Destination.

Different topology configurations have been tested changing the sequence of devices (ingress, intermediate and egress IISoA routers; legacy IIS routers) The experimentation allowed a confirmation of the feasibility of our proposal, the functional verification of the processing rules in the routers, an estimation of the software implementation complexity. In fact the software realization turned out to be very simple.

6. Open issues for the multicast case

Let us consider the extension of this approach to the multicast case. The goal is to use a similar mechanism to let the ingress ED obtain the ATM addresses of the set of egress EDs for a given IP multicast address. Then a “shortcut” ATM multicast VC can be established to support the QoS flow. This approach is basically compatible with the unicast case thanks to the fact that the ATM shortcut VC is established starting from the ingress ED which can act as root of the ATM point-to-multipoint connection. Additional leaves (i.e. egress EDs) can be successively added or removed according to RSVP requests. There are several open issues to be solved to define a working solution for the multicast case. A general problem in the interworking between RSVP and ATM multicast is that different receivers can specify different QoS for the same flow, or even no QoS at all when they want to receive the flow on a best-effort base. The interworking solution must specify the algorithm to map the IP multicast flow in one or more ATM point-to-multipoint VCs. Another problem to be considered is the interaction between the IP multicast routing and the ATM shortcuts that can be established modifying the topology. We are currently investigating on a solution where the IP multicast tree is replaced by ATM multicast subtrees whenever is possible. This means that the ATM multicast can reach the end-terminals if they are ATM capable, or it can be confined to ATM capable routers in the backbone, just like in the unicast solution.

7. Conclusions

This paper has presented an extension of the RSVP protocol to support IP/ATM address resolution and the establishment of ATM QoS Virtual Channels. The goal is to allow the IIS architecture to fully exploit the capability of an underlying ATM network. Two new classes
have been added to the RSVP PATH and RESV messages, allowing the IIS over ATM routers to signal their capability (and will) to set-up an ATM QoS VC and to exchange the ATM addresses.

The compatibility with “traditional” RSVP hosts and routers has been maintained. No modifications are needed in the host and in the routers that do not take part in the shortcut procedure. With respect to performance, the proposed enrichments seem to have a very limited impact on the size of the RSVP message and on the processing and storage requirements in the RSVP routers. On the other hand, the proposed approach seems to provide a straightforward mechanism to setup the “longest” ATM shortcut, without the need of additional address resolution mechanisms like NHRP. The IP unicast case has been considered in detail. Even if the detailed extension to the multicast case is for further study there are no conceptual obstacles.

References
