A Framework for the QoS Based Integration of IP and ATM in the DIANA Project

WERNER ALMESBERGER, SILVIA GIORDANO
Ecole Polytechnique Federale de Lausanne, ICA, 1015 Lausanne, Switzerland

PIERGIORGIO CREMONESE
Finsiel, Via Matteucci 34/B, 56124 Pisa, Italy

HANNU FLINCK, JOHN LOUGHENEY
Nokia Research Center, Itälähenkatu 22B, PO Box 422, 00045 Helsinki, Finland

MARTIN LORANG\(^1\)
University of Stuttgart, IND, Pfaffenwaldring 47, 70569 Stuttgart, Germany

\(^1\) Reference author, Email: lorang@ind.uni-stuttgart.de, Phone: +49 711 685 7991, Fax: +49 711 685 7983
Abstract. In recent years, the Internet community has been working on extensions to the current Internet protocol suite in order to enable service guarantees or service differentiation, but the standardised solutions to integrate IP and ATM still are restricted to IP to ATM on a best-effort basis. This paper deals with a framework for QoS based interworking between IP and ATM. After a brief overview on the state of the art of IP and ATM networking, particular focus is put on the integration of the IP reservation protocols RSVP and SRP with ATM. Signalling and traffic control issues as well as application requirements related with two models, often referred to as overlay and peer models respectively, are discussed. Finally, an architecture for a prototype network architecture as planned by the ACTS DIANA project is presented which will allow to investigate several approaches for the QoS based integration of IP and ATM.

1 Introduction

The integration of voice, data, and video services modified the target of networking technologies. Instead of providing a single type of service, networks now have to cope with the integration of services too and, related with that, with providing Quality of Service (QoS) as required by each of those services. ATM development was started with this new issue in mind [1], and for long was regarded as the ultimate networking technology. While this view turned out to be overly optimistic, ATM enjoys a comparably large installed base in public WANs and has also been deployed in many LANs. Furthermore, ATM is also foreseen to be used in conjunction with xDSL technologies, which may also increase its popularity.

The IETF has also been working towards support for so-called Integrated Services: A comprehensive Integrated Services framework (INTSERV, [2]) and a resource reservation protocol RSVP [24] have been developed and the latter is currently being deployed. Further work concentrates on issues like charging support and improved scalability of reservation mechanisms (DIFFSERV, [26, 27, 28]).

Besides the mainstream activities of IETF, ITU, and ATM Forum, there are also several independent developments, such as Arequipa (a mechanism designed to leverage ATM QoS mechanisms to IP in a way that allows for rapid deployment [3, 4, 10, 11] and various alternative reservation protocols to RSVP with better scalability properties (e.g. SRP [16], TSP [38], DRP [39]).

Several papers give technical overviews and a classification on the competing integrated services network solutions [6, 7], so that they are only briefly treated in a separate subsection of this introduction before the ACTS DIANA project is introduced.

In Section 2, the different signalling paradigms of the reservation protocols under investigation, namely RSVP, SRP and ATM are compared. Two models for interworking are generally considered. The use of a common layer, illustrated in Fig. 1a, where a part of the properties available at the common layer are mapped to the layer underneath, which is based on different technologies on either side, and, as shown in Fig. 1b, the separation of both technologies with a translation in a Interworking Unit (IWU). Those two models are also sometimes referred to as overlay and peer models. That section concludes with a review of existing work on a framework for RSVP and ATM.
Section 2 passes seamlessly over to section 3 with a discussion on the practical realisation and necessary changes to those architectures to be applicable to a prototype implementation as planned in the DIANA project. It is shown how signalling and traffic control issues can be solved in two concrete interworking scenarios, similar to the peer and overlay model mentioned above. A framework for both RSVP and SRP is outlined. Furthermore, the problems arising for an application while attempting to use QoS in different heterogeneous network scenarios are analysed.

Finally, section 4 surveys the important modules and their interaction to be implemented in DIANA’s so-called Integration Unit.

1.1 State of the Art of IP and ATM Networking

The dynamic and thus efficient use of ATM to support IP based data services requires protocols such as LANE [21], CLIP [22], MPOA [23] or others which resolve IP addresses to ATM addresses and which trigger the set-up of ATM connections on the arrival of IP data streams and their release after a period of inactivity. Although providing those functions, current IP over ATM specifications only offer the traditional IP best-effort service, as noted in the previous section, which is inappropriate for applications with tight QoS constraints.

Arequipa [10, 11] is one of the first solutions that was demonstrated for providing the QoS of ATM to TCP/IP applications, however in a pure ATM infrastructure only. Arequipa enabled IP applications have access to QoS and traffic control parameters of ATM, at a minimum cost considering that the necessary software only has to be deployed in the end systems and not in the entire network.

IP Switching [12] and Tag Switching [13] are other solutions which aim at supporting IP over ATM in a more efficient and scalable way without involving ATM signalling. Both solutions integrate routing and switching by means of replacing the relatively expensive look-up for IP prefixes when flow information has been cached in label based tables. They differ mainly in the way labels are allocated: IP Switching label selection is based on a traffic flow analysis whereas Tag Switching also considers network topology. Moreover, the latter is not restricted to use IP and ATM. Currently, the IETF is specifying a non-proprietary approach named Multi-Protocol Label Switching [40], which is based on Tag Switching.

Both those router/switch architectures and traditional IP routers may have to be enhanced by either implementing an Integrated Services or Differentiated Services architecture.

In Integrated Services [25], RSVP allows applications to reserve network resources for individual flows in an IP network. However, the wide scale deployment of RSVP must be approached with care because the processing of (periodically refreshed) reservation and control messages, the identification of each packet based on the IP header and the handling of per-flow reservation state becomes challenging in backbone routers passed by a huge number of individual flows.

---

**Fig. 1a: Overlay model for interworking**

**Fig. 1b: Peer model for interworking**
Conversely, the Differentiated Services (DS) architecture [14, 26, 27] achieves scalability by classifying and marking packets by means of the so-called DS field [28] in the IP header at the ingress to a DS capable IP network. The goal is to receive a particular per-hop forwarding behaviour on DS routers along their path. With this capability, Internet service providers are able to offer services besides simple best-effort to their subscribers, such as a „premium“ service that is assigned priority over the best-effort service. A DS admission control service (DACS) [29] may ensure that the resources needed for the aggregate stream are available inside the DS network. Work on realising admission mechanisms is however at an early stage of development.

Scalable Reservation Protocol (SRP) [16, 30] provides a light-weight reservation mechanism for adaptive multimedia applications [17]. The main focus is on good scalability to a very large number of individual flows. Senders and receivers actively participate in maintaining reservations, but routers can still control their conformance. Routers aggregate flows and monitor the aggregate to estimate the resources needed to support present and new reservations. As opposed to RSVP and ATM, there is no explicit signalling of flow parameters.

1.2 The DIANA Project

Since it is likely that none of the aforementioned reservation paradigms is likely to prevail over the others in the near future and thus several mechanisms will co-exist, a key issue for achieving convergence between Integrated Services, Differentiated Services and ATM to support QoS end-to-end across domain boundaries, is the integration of the Internet Integrated Services model and RSVP signalling on the one hand, DS marking and per-hop behaviour as well as ATM service categories and ATM signalling on the other hand.

In this area, DIANA started in March 1998 as a new project in the European Union 4th Framework Programme ACTS. As its main goal, the DIANA consortium will develop, integrate, validate and demonstrate resource reservation and traffic control functionality which seamlessly interoperate between ATM and IP Integrated Services networks in order to provide guaranteed QoS end-to-end. Although, as mentioned above, DIANA will mainly focus on Integrated Services and ATM, the design of the trial platform will be kept flexible enough to allow investigating different solutions for the convergence of IP and ATM, such as SRP with ATM or others.

In any case, a so-called Integration Unit will be placed at the boundary between ATM and IP domains to provide the functionality needed for the translation between the respective IP reservation protocol and ATM signalling. The control plane of this Integration Unit is assigned the key role for prototyping the signalling translation from RSVP and ATM UNI [31] signalling and vice versa, for the mapping of QoS specifications given by the flow descriptors objects with Integrated Services and ATM traffic descriptor information elements respectively, and for the allocation of ATM virtual connections for IP flows.

2 SUPPORT OF REAL-TIME TRAFFIC IN IP AND ATM

The main attributes describing real-time data streams are minimum bandwidth, minimum delay and delay variance. These attributes pose much more stringent service requirements to the underlying network than the data transfer that is based on file transfer and can tolerate random delays and best effort type of service.

There are clear indications that the number of applications using real time traffic is increasing in the Internet. Typically these applications use Real Time Protocol (RTP) [37] on top of User Datagram Protocol (UDP) to carry the real time traffic in the user plane.
However, RTP does not support any resource reservations for service guarantees but may use connections with QoS attributes set up by signalling protocols.

Therefore, this section surveys the different signalling paradigms of RSVP, SRP as well as ATM, and also reviews and evaluates latest work on a RSVP and ATM interworking framework. As already mentioned in the introduction, two models, referred to as overlay and a peer model, are considered.

2.1 Overview of Signalling Paradigms

The ITU defines signalling as the exchange of information specifically concerned with the establishment and control of connections, and with management, in a telecommunication network [32]. In this paper, the term signalling is used for any exchange of information to establish and maintain a network element state with QoS, service policy or similar attributes, associated with a connection, flow of data or aggregates of these.

The concept of signalling stems from the traditional telephony world, where the only service was a voice call with fixed bandwidth requirements and well known delay characteristics. In general, the main parameters for call establishment are: The destination (and optionally source) address and the QoS parameters stating the nature of the bearer service that the call is expecting.

The three subsequent subsections deal with the Internet reservation protocols RSVP and SRP as well as with ATM signalling. Though obeying quite different paradigms, the common idea is that the communicating parties establish the resource reservation by some sort of signalling.

Alternatively, a Service Level Agreement [27], agreed upon at subscription time or by other out-of-band communication, may specify aspects of a type of service offered in e.g. a Differentiated Services network architecture.

2.1.1 Resource ReSerVation Protocol (RSVP)

RSVP [24, 33] and Integrated Services have introduced the concept of signalling to the router based Internet network. Instead of the call, the basic element of the Integrated Services model is a flow, that is a sequence of packets from a particular source to a destination that are related in terms of routing and handling policies.

RSVP allows applications to reserve network resources in the Internet. It operates on top of IP (either IPv4 or IPv6) and it relies on standard Internet routing. It is used both in hosts and routers to reserve resources for a unidirectional flow. A RSVP reservation request contains a so called flow descriptor which characterises the reservation request and specifies the traffic profile.

RSVP is designed for both unicast and multicast communication in a heterogeneous network, where receivers may have different characteristics and multicast membership is dynamic. These requirements lead to a solution, where the receiver is responsible for initiating the resource reservation.
RSVP assumes that a multicast group already exists. As indicated in Fig. 2, the sender S1 sends a PATH message to a multicast group announcing the characteristics of the flow it is going to send. When the receivers, R1 and R2, want to make a reservation, they send a RESV message back along the reverse path previously installed by the PATH messages. While processing a RESV message, each RSVP capable router creates reservation state along the path from the receiver to the sender if the admission control check was successful. In a multicast scenario, as the one shown in Fig. 2, there are nodes that will receive two or more RESV messages from different branches of a multipoint tree. These nodes merge the received reservations and forward only one merged reservation request.

Finally, flow related parameters are set in the packet classifier and packet scheduler. The packet scheduler is responsible for negotiation with the link layer to reserve the transmission resources. It is here that mapping from the flow level QoS to the link layer QoS takes place.

With the reservation style being part of the reservation request, a receiver can signal if there should be a separate reservation for each sender of a session (Fixed-Filter), if the reservation can be shared among the named senders of the session (Shared-Explicit), or if the reservation can be shared by all the senders (Wildcard-Filter).

RSVP uses soft state for the flow reservation. This means when a reservation is made, it must be periodically refreshed. The advantage of using soft state for the reservation is that the route of the connection can be changed dynamically inside the network and the reservation will be re-established when the new PATH and RESV messages has passed through the new route. Soft state also helps to allow for dynamic multicast group membership. However, per flow state increases the complexity and scalability of routers, therefore RSVP is not recommended as a solution for backbone networks.

2.1.2 Scalable Reservation Protocol (SRP)

This section briefly describes two key aspects of SRP: The reservation mechanism and the way aggregation is accomplished. Further details can be found in [16].

Reservation Mechanism
A source that wishes to make a reservation starts by sending data packets marked as \textit{request} packets to the destination, refer to Fig. 3. Those packets are subject to packet admission control by routers, based on the following principle. Routers monitor the aggregate flows of \textit{reserved} packets and maintain a running estimate of what level of resources is required to serve them with a good quality of service.

When receiving a \textit{request} packet, see Fig. 4, a router determines whether hypothetically adding this packet to the flow of \textit{reserved} packets would yield an acceptable value of the estimator. If so, the \textit{request} packet is accepted and forwarded towards the destination. While still keeping the status of a \textit{request} packet, the router must also update the estimator as if the packet had been received as \textit{reserved}. In the opposite case, the \textit{request} packet is degraded and forwarded towards the destination, and the estimator is not updated. Degrading a \textit{request} packet means assigning a lower traffic class to it, such as best effort. Hence, a packet sent as \textit{request} will reach the destination as \textit{request} only if all routers along the path have accepted the packet as \textit{request}.

The destination periodically sends feedback to the source indicating the rate at which \textit{request} and \textit{reserved} packets have been received. This feedback does not receive any special treatment in the network. Upon reception of the feedback, the source can send packets marked as \textit{reserved} according to a profile derived from the rate indicated in the feedback. If necessary, the source may continue to send \textit{request} packets in an attempt to increase the rate that will be indicated in subsequent feedback messages.

Thus, in essence, a router accepting to forward a \textit{request} packet as \textit{request} allows the source to send more \textit{reserved} packets in the future; it is thus a form of implicit reservation.

\textit{Aggregation}

Routers aggregate flows on output ports, and possibly on any contention point as required by their internal architecture. They use estimator algorithms for each aggregated flow to determine their current reservation levels and to predict the impact of accepting \textit{request} packets. The exact definition of what constitutes an aggregated flow is local to a router.

Likewise, senders and sources treat all flows between each pair of them as a single aggregate and use estimator algorithms for characterising them. The estimator algorithms in routers and hosts do not need to be the same. In fact, hosts are supposed to implement a fairly simple algorithm, while estimator algorithms in routers may evolve independently over time.

\begin{center}
\begin{tabular}{|c|c|c|}
  \hline
  \textbf{Source} & \textbf{Application} & \textbf{Protocol stack} \\
  \hline
  Needs reservation & Reservation established ? & Yes \\
  \hline
  Doesn’t need reservation & Reservation possible ? & No (Denied) \\
  \hline
  \end{tabular}
\end{center}

\begin{center}
\begin{tabular}{|c|c|c|}
  \hline
  \textbf{Router} & Reservation established & Resources available? \\
  \hline
  Yes & Best effort & No \\
  \hline
  No & Resources available? & Yes \\
  \hline
  \end{tabular}
\end{center}

\textit{Fig. 4: SRP packet type assignment by sender and processing by routers}
ATM is a connection-oriented protocol requiring a bi-directional (virtual) connection to be established before user traffic can start flowing between the communication entities. When the virtual connection (VC) is in an active state, packets of fixed length, so-called cells, are marked with a Connection Identifier unambiguously assigned on a per-link basis. In order to support a variety of applications with different QoS requirements and traffic profiles in the packet-switched ATM, a comprehensive protocol suite with sophisticated signalling capabilities has been specified.

Each VC can choose among various service categories [34], ranging from Constant Bit Rate (CBR), Real-Time and Non-Real-Time Variable Bit Rate (VBR), Available Bit Rate and finally Unspecified Bit Rate (UBR). For each of those service categories, a QoS Parameter Information Element (IE) and a Traffic Descriptor IE have to be included in a connection SETUP message. As a result of the connection set-up, each VC gets its own set of parameters, e.g. a Minimum Cell Rate (MCR), Sustainable Cell Rate (SCR) and Peak Cell Rate (PCR).

Both ITU-T and ATM Forum have defined a subset of interfaces. This paper focuses on Private User-Network UNI 4.0 [31], which is an enhancement to the earlier UNI 3.1 and provides some kind of alignment to the recommendations for Public UNI specified by ITU-T. Besides the basic call set-up and tear down functionality, UNI 4.0 defines point-to-multipoint operation, address registration and extended QoS support. An overview of ATM signalling capabilities and procedures can be found in [36].

The concept of VCs and service categories allows for statistical multiplexing and end-to-end QoS guarantees. Furthermore, ATM supports multiple services, can provision Frame
Relay, SMDS, native ATM, voice, video and circuit emulation. However, all this features make ATM a highly complex technology, and, in particular its connection oriented paradigm, creates a lot of overhead for short-lived data flows.

In addition, when used as a layer 2 technology for IP traffic, complex address resolution mechanisms have to be implemented and executed prior to the connection set-up.

Most signalling procedures are controlled by the owner of the connection. For this reason, especially the support of multicast including dynamic and heterogeneous reservation is problematic. Furthermore, the connection attributes are negotiated at set-up and only a few of them can be altered later. Table 1 summarises the main differences in the paradigms of ATM and RSVP.

### 2.2 RSVP and ATM Interworking

Given that both, ATM and RSVP, use out-of-band signalling, the so-called control plane, which defines signalling functionality, and the user plane, which provides data transfer functionality, can be distinguished. An Interworking Unit (IWU) potentially has to intervene in both planes while preserving the end-to-end QoS constraints.

#### 2.2.1 User Plane

The user plane comprises the modules and protocols required for data addressing, data encapsulation and data transfer. In order to achieve interworking between IP and ATM networks, either a common higher layer has to be identified or a translation between a pair of layers specified.

In the first case, the overlay model (Fig. 1a), the use of IP as the common layer appears to be the most natural solution. ATM is used as a link layer and IP plays the role of the internetwork layer. The use of IP for this purpose was already proposed in the early OSI model [18]. This solution elegantly avoids interoperability problems in the user plane, because the end-to-end aspects of the communication are completely managed by the IP layer.

Another conceivable approach is to define a mapping between AAL 5 and UDP, or, which is more difficult, between AAL 5 and TCP, because each of these protocols can be viewed as acting at the transport layer in its respective stack.

However, there exists no concept for a „generic“ user plane mapping. It is therefore necessary to define the precise mapping rules for each application or set of applications. In particular, considering that the mapping of native ATM to IP fundamentally depends on the specific application, the usage of application level gateways appears to be a natural but inefficient choice.

#### 2.2.2 Control Plane

The control plane includes all functionality related to setting up and modifying communication paths, such as routing aspects, QoS negotiation (and re-negotiation), and connection management.

**Routing**

IP and ATM both provide a global addressing scheme and they also have their own routing mechanisms. Naturally, there is no compatibility in the addressing and routing schemes of both worlds. However, there are some attempts for at least partial merges [41, 51].
Routing in an interworking environment depends partially on how addresses are translated. Even if it is not directly part of signalling, it influences the latter as it defines the communication path to be set up by signalling.

The following may serve as an example: If the solution adopted at the user plane is to use end-to-end IP, and if CLIP or LANE is used to transport IP over ATM, the choice of routes is entirely determined by IP routing. However, if a mechanism is chosen that allows to identify „shortcuts“ in the ATM network (e.g. NHRP [19]), IP routing, ATM routing, and the partial routing functionality of that mechanism interact. In either case, routing is transparent to the Interworking Unit.

In a peer model (Fig. 1b), routing is performed independently in the two networks, and routing management on both networks is strictly related to the address mapping performed by the interworking unit.

**QoS Negotiation and Renegotiation**

If RSVP is used as the reservation protocol for the IP network, the work of the IETF Integrated Services over Specific Link Layers (ISSLL) approaches [42, 43, 45, 46] for mapping INTSERV traffic specifications to ATM traffic parameters can be directly applied in the IP to ATM direction.

If using the peer model, a mapping of Integrated Services to ATM parameters has to be inverted too. But the work of ISSLL may provide useful insights for implementing such a functionality.

Other reservation protocols need their specific mapping, as will be described later with SRP and ATM.

**Connection Management**

RSVP (and SRP) establish reservations in an unidirectional way. Contrary to that, ATM set-up carries both the forward and the backward traffic parameters. This fundamental incompatibility is not an issue when mapping from IP to ATM, but it may require interaction with higher-layer protocols (e.g. at the application level) in the opposite direction.

2.3 Overview on IETF Framework Drafts and Related Work for RSVP and ATM

A number of IETF drafts deal with a framework for Integrated Services and RSVP over ATM. This subsection points out in which way the outcome of that work could be applied to a baseline implementation in DIANA and discusses the advanced issues with the support of multicast and flow aggregation strategies.

2.3.1 Baseline Implementation for RSVP and ATM

A baseline implementation for RSVP over ATM as specified in [42] only requires an implementation to establish RSVP-initiated VCs to RSVP capable end points. An ATM network supporting switched VCs (SVCs) lies somewhere on the path between those RSVP sender(s) and receiver(s). This scenario, depicted in Fig. 5, is referred to as RSVP over ATM scenario in this paper.
The SVCs are set up, added to (in the case of multipoint trees), torn down, and controlled by the edge devices at the ingress to and the egress from the network, which act as both IP routers and ATM access nodes, capable of initiating and managing VCs across the ATM UNI [43].

It is important to note that, in the RSVP over ATM scenario, the set-up, modification and release of ATM connections is always triggered by RSVP control messages.

A baseline implementation of an edge device [42] sends RSVP messages over the best-effort paths set-up by one the of conventional IP over ATM protocols, e.g. CLIP, and uses independent VCs for each RSVP reservation thus doing without aggregation. As an alternative, a separate, shared RSVP control VC could be established in order to ensure a normally loss-free and hence more robust operation of the RSVP control path.

An open issue is how best-effort behaviour for non-conforming packets can be retained inside the ATM network with this architecture. In [43] cell tagging is proposed, however, since priorities and per-VC queues have been introduced to ATM, tagged traffic in a stream with guaranteed QoS might still affect lower priority traffic in an unfair manner. A solution which is both fair and conformant to the requirements of the Integrated Services is not in sight.

A convincing mapping of IETF Integrated Services (Guaranteed, Controlled Load and Best Effort Service) to ATM Forum service categories is given in [43], but RSVP provides many signalling features, such as receiver oriented reservations, heterogeneity within a multicast session, dynamic change of reservations, and multiple reservations styles that cannot easily be supported with ATM. In ATM, the first branch of a point-to-multipoint VC determines the QoS (statically) for the whole tree from the root. Hence, ATM does not allow for heterogeneity in a single point-to-multipoint VC, and separate point-to-multipoint VCs would have to be set up if ATM is supposed to support both UBR and guaranteed service(s) within a session.

The Leaf Initiated Join (LIJ) capability offered by UNI4.0 [31] enables ATM receivers to join a point-to-multipoint VC but without the option of specifying an individual QoS. Thus LIJ does not resolve the mismatch between RSVP and ATM with respect to receiver heterogeneity.

For this reason, several authors [42, 44, 45, 46] abandon the goal of full heterogeneity and introduce a limited heterogeneity or modified homogeneous model [42, 43, 46]. With limited heterogeneity, a best-effort and a single alternate QoS are offered, whereas only one VC with the maximum requested QoS is established in the modified homogeneous model.

Also multicast address resolution, which is needed to make use of the multicast capabilities of ATM, is only feasible in a close co-operation with IP group membership and multicast routing protocols. Since neither LANE nor MARS [47] operate beyond the scope of the logical IP subnetwork they are serving, and, in addition, do not support QoS, the existing standard framework is not sufficient and has (and is about) to be extended.

*Fig. 5: RSVP over ATM networking scenario*
2.3.2 Flow Aggregation with RSVP and ATM

Potentially, some of the scaling issues of RSVP can be addressed by aggregating several RSVP flows over a single VC if the destinations of the VC match for all the flows being aggregated. Similar to multicast, aggregation complicates VC management, and even worse, requires sophisticated scheduling mechanisms at the originating point of the VC. On the other hand, RSVP would become invisible inside the ATM network and normal connection classification based on the VCI and VPI would apply. Since ATM has the potential to fill the gap between the ingress and egress routers to and from the ATM network, end-to-end QoS is feasible. Nevertheless, only flows with an identical path through the ATM network can be aggregated thus limiting the positive impact on the amount of reservation state in the backbone network.

In contrast, RSVP or Differentiated Services [29, 52, 53, 54] based aggregation aims at reducing reservation state in each IP hop independently. In spite of some subtle differences, their common paradigm is that routers inside the aggregation region (the backbone) classify packets at the ingress with a limited number of different classes and do not maintain per-flow reservation state in the interior. The admission control will be done at the ingress too, but the decision will be based on congestion information within the aggregation region.

3 Issues Considered in DIANA’s Experimental Platform

The integration of RSVP and SRP with ATM requires the interaction of the respective signalling procedures and the related traffic control as well as the mapping of service classes and their parameters from the IP reservation protocols to ATM and vice versa. In this section, a networking scenario is developed based on IP as the common, user plane overlay layer. Nonetheless, still a distinction can be made - and is done for RSVP and ATM interworking - between a scenario where all applications use the same type of signalling - RSVP - and a scenario where some applications use - maybe by extending the standardised IP over ATM protocols CLIP, LANE or MPOA similar to Arequipa - UNI signalling without making the detour over IP reservation protocols. Since a convincing framework for Integrated Services to ATM traffic parameter translation already exists [43], the focus of this section is on signalling with RSVP and ATM interworking and on reservation issues for the SRP and ATM case. This section concludes with application specific issues related with the access and the handling of signalling and traffic control.

3.1 Signalling Scenarios with RSVP over ATM versus RSVP peering with ATM

3.1.1 RSVP over ATM

DIANA’s Integration Unit can be regarded as an implementation of an edge device as considered in [43] which is supposed to be fully functional in both the IP Integrated Services and RSVP protocols as well as ATM UNI signalling and traffic management. With such an edge device a RSVP over ATM scenario can be realised, see Fig. 5 above, in which both connection end-points are controlled by RSVP.

With RSVP over ATM, illustrated in Fig. 6, a RSVP sender starts sending PATH messages downstream towards the receiver. Provided that a RSVP control VC has been established before, those PATH messages can pass the ATM network without triggering the exchange of signalling messages and finally arrive at the receivers. In normal operation, the receiver returns a RESV message which is of course also processed by the Integration Unit devices which act as RSVP capable routers. The Integration Unit at the downstream ingress
to the ATM network originates a call to set-up an „unidirectional“ ATM VC, i.e. a VC with no reservation in the upstream direction. A failure in the set-up of this ATM VC should trigger the same RSVP error messages as if the local reservation had failed. But even in the case of a reservation failure somewhere in the path, a best-effort service could still be offered because, as mentioned above, a separate best-effort VC interconnects the ingress and egress Integration Unit.

The traditional method to account for changes in RSVP reservations, as described in detail by Berger [42], is to attempt to replace an existing VC with a new appropriately sized VC. During the set-up of the replacement VC, the old VC must be left in place unmodified. If the set-up of the replacement VC fails, then the old QoS VC must continue to be used. If the new reservation is greater than the old reservation, the reservation request must be answered with an error in this case. When the new reservation is less than the old reservation, the request must be treated as if the modification was successful. This behaviour is required in order to conform to RSVP error handling as defined in sections 2.5, 3.1.8 and 3.11.2 of [24].

Latest ATM signalling standards permit modifications of traffic parameters by the connection owner, i.e. the end station that initiated the connection (which is now in an active state). In [48] and [49] modification procedures for the traffic parameters Peak Cell Rate and Sustainable Cell Rate have been defined. Only parameters which have been specified at connection set-up can be altered, so a change from one ATM service category to another one is not possible.

In Fig. 6, a modification procedure could replace the normal set-up procedure without problems. No matter whether the traditional set-up and release or the modification procedures are used, the number of modifications per time should be limited and charging should give an incentive not to use this feature frequently.

RSVP can identify from either explicit messages or time-outs when a data VC is no longer needed. Therefore, data VCs set up to support RSVP controlled flows should only be released at the direction of RSVP. VCs must not be timed out due to inactivity by either the VC initiator or the VC receiver. A release at the direction of ATM should only happen when ATM is the end-point of the QoS path, as is the case in RSVP peering with ATM scenario, see Fig. 7.

3.1.2 RSVP peering with ATM

In the RSVP peering with ATM scenario, a user application directly communicating with an ATM API interworks with an application using RSVP to signal its resource demands, as
shown in Fig. 7. The Integration Unit is now the terminating entity for ATM signalling and the originating entity for RSVP PATH messages or vice versa.

![Fig. 7: RSVP peering with ATM scenario](image)

When setting up a RSVP reservation for a call originating from ATM, the Integration Unit may use the information contained in the incoming SETUP message to send a PATH message towards the RSVP controlled destination, as illustrated in Fig. 8. At the same time, an ALERT message is returned to avoid a premature expiry of signalling timers (T303, T310).

Only on the receipt of a RESV message, it should answer to the SETUP with a CONNECT message. If the original SETUP message contained one or more Alternative ATM Traffic Descriptor or a Minimum Acceptable ATM Traffic Descriptor Information Element [36], a reservation request deviating from what was specified in the PATH message may be taken into account in ATM parameter negotiation.

In any case, an ATM connection is only established if the reservation was successful on the entire RSVP path. It is important to note that the ATM application should refrain from requesting bi-directional ATM VCs and instead rely on the RSVP destination to specify the QoS of the reverse path if needed.

Alternatively, the VC set-up may be finished before PATH messages are sent towards the IP receiver. There is also no danger of a false expiry of signalling time-outs since the UNI signalling demon completes the ATM set-up without interaction with RSVP modules. However, if the receiver does not accept the traffic specification as advertised by the PATH message, a RESV message arrives at the Integration Unit which does not match the original reservation the ATM connection was based upon. Since it is not the initiator of the VC, the Integration Unit does not dispose of the standard mechanism [48] to alter the parameters specified in previous ATM Traffic Descriptor Information Elements of the already active connection.

Even worse, since the ATM connection set-up is finished before the reservations in the IP domain are established, the ATM sender and originator of the QoS path in the upper half of

![Fig. 8: RSVP peering with ATM: Signalling message flow for an ATM initiated QoS path, waiting for RSVP reservation request (Simplified)](image)
Fig. 9: Signalling message flow for RSVP peering with ATM when ATM set-up completes immediately (Simplified)

Fig. 9 may already start sending before a QoS path has been prepared throughout the network.

The opposite direction (see bottom half of Fig. 9) is less problematic because an ATM destination has no means to downgrade a reservation (By specifying an alternative traffic descriptor information element in the ATM set-up message, the call originator could give the receiver a chance of making a selection among different traffic descriptors, however, the Integration Unit, being the call originator, can simplify the procedure by not offering this option).

If both the RSVP over ATM and the RSVP peering with ATM scenario was implemented, the Integration Unit would have to run different signalling procedures as illustrated in Fig. 6, Fig. 8 and Fig. 9. A classification of incoming SETUP messages based on the ATM destination (ATM address + Service Access Point) would be required to start the right procedure.

Since specifications of IP and ATM [43, 50, 51] as well as some examples in [31] already use the Broadband Lower Layer Information Element (BLLI-IE), the Broadband Higher Layer Information Element (BHLI-IE) could be used for making this distinction.

Nevertheless, all the issues and the (in the opinion of the authors) low commercial impact with the RSVP peering with ATM scenario suggest to use RSVP also in ATM hosts when IP applications are to be supported. It is important to note that such a host may include functionality which replaces the ingress Integration Unit.

3.2 Issues of SRP over ATM

This section gives a case study of how SRP can be transported over an ATM link layer. Focus is put on the usual overlay model with switched VCs. Furthermore, the case study is restricted to the „classical“ IP over ATM model [51], thereby excluding routing issues, such as the use of „shortcuts“ [19, 23], and their implications on the estimation of resource use.

As SRP aims at providing a low delay, low loss service, any ATM traffic class can be used that guarantees a minimum bandwidth at or below which losses due to congestion are unlikely to occur, and that has either an explicit delay bound or a statistically low delay. This
includes CBR, VBR, ABR [34] and GFR [35]. For simplicity, the use of the most basic and most commonly available traffic class, CBR, is assumed.

The main difficulty in carrying SRP over ATM is that SRP builds up reservations gradually, while ATM expects all connection parameters to be known at the time the connection is set up. On reception of a request packet, a router may either use a suitable, previously established ATM connection which has excess bandwidth available, or it has to request an additional reservation to the respective neighbour node of at least the expected bandwidth increase. This can be done by setting up a new ATM VC or, if the ATM network supports this service, by re-negotiating the traffic parameters of an existing connection [48].

Since a router must know the admission decision of the ATM network prior to deciding whether a request packet can be accepted, and also since frequent set-ups or re-negotiations would put a significant load on the signalling processors in the ATM network, routers should anticipate future reservation behaviour in order to reduce the number of adjustments that need to be made.

Fig. 10 shows a simple example of how ATM bandwidth can be adjusted to accommodate SRP reservations: ATM reservations are changed in comparably large steps. An increase is initiated whenever the SRP reservation exceeds a certain threshold below the current ATM reservation. Likewise, a decrease is initiated when the SRP reservation drops below a second, lower threshold.

If the SRP reservation reaches the ATM reservation and the increase has not been accepted by the ATM network (yet), the request packet has to be degraded, thereby delaying the reservation increase. More advanced algorithms for dynamically adjusting the ATM bandwidth are for further study.

In the absence of re-negotiation capabilities in the ATM network, a similar functionality can be obtained by setting up new VCs which either replace or complement the existing VC(s). In the former case, roughly twice the required bandwidth has to be reserved in the ATM network during transitions from old to new VCs, which may lead to admission failures even if there would be enough bandwidth to accommodate the actual traffic. In the latter case, the number of active VCs in the ATM network is increased and the router needs to support load balancing over multiple parallel VCs. Furthermore, such load balancing may lead to packet reordering, which may negatively affect the performance of higher layer protocols.
Applications get access to signalling functionality via a protocol specific API, e.g. RSVP’s RAPI. Currently, DIANA is enabling some applications to work with either RSVP and SRP.

However, real RSVP and ATM interworking means to build up a scenario that supports interoperability between native ATM applications, IP applications with access to ATM control capabilities, e.g. Arequipa, and RSVP based IP applications.

Introducing a new, common layer on top of the network layer (ATM and IP both are regarded as network layers, contrary to RSVP over ATM approaches where ATM is regarded as a link layer technology), RTP for instance, forces a quite strong constraint on the application structure requiring the presence of RTP everywhere. In the case of native ATM applications, e.g. VoD with MPEG 2 over ATM, this constraint is not originally satisfied.

A peer model allows to leave the applications unchanged but needs a quite complex application-aware, intermediate system for interworking. Such a system must support a gateway providing proxy capabilities specialised for each application. It follows that for each application a new intermediate element must be developed. This element should solve all problems related to the user plane such as addressing and packet translation. It could appear attractive because it does not affect applications on the end system side.

Since the traffic profile information provided by signalling controls the network’s traffic management (traffic control, congestion control), the application’s traffic profile specification plays an important role for QoS no matter if interworking is involved at all.

The applications must be compliant to its specification and the specification has to reflect its real demands.

The way an application generates a traffic profile specification is implementation and service specific, but can in general either be complex or little precise. This is one of the reasons why SRP does without an explicit specification of traffic parameters but instead simply marks packets that need QoS.

Whenever aggregation is taking place, a scheme that allows to reallocate the network resources in an optimal way while guaranteeing QoS to the traffic flows [20] may improve efficiency. This scheme periodically adjusts the network resources based on the profile of the aggregated traffic received in the past and the aggregated traffic expected in the near future, which can be either pre-recorded or known by means of prediction or estimation. This scheme suits perfectly the dynamics of RSVP but of course comes with a significant control overhead.

### 4 Design of an Integration Unit Prototype

DIANA’s Integration Unit has to be fully functional in both the Integrated Services and RSVP protocols as well as ATM UNI signalling and traffic management. On the RSVP side, the Integration Unit adopts the role of a RSVP capable router, i.e. processes RSVP messages, reserves resources, and maintains soft state (in the control path), and classifies, polices and schedules packets (in the data path) before finally forwarding them. As an ATM end system, the Integration Unit sets up ATM connections by UNI based signalling, and accepts or refuses incoming connections.

Due to the problems with the RSVP peering with ATM scenario, the Integration Unit prototype as described here only includes the functionality required for the RSVP over ATM scenario. The key issue is that admission control for Integrated Services is now dependent on the successful set-up or modification of a connection across the ATM network, e.g.
interconnecting two Integration Unit devices. In RSVP terminology, this is part of the Link Layer Dependent Adaptation Layer (LLDAL). This layer provides, as a first step, the translation between Integrated Services traffic descriptors to the descriptors of the link layer underneath. In the case of ATM, the set-up of a new VC, the modification of an existing VC or, if aggregation is taking place, a simple look-up for a VC already active and the respective admission decisions will also be taken here.

In the Integration Unit, those functions will be provided by the so-called Interworking Control module. The RSVP demon invokes this process whenever RSVP operation and message processing requires a change in the reservation. Both the RSVP demon and the Interworking Control module as well as the UNI signalling demon manipulate kernel data structures that represent all attributes assigned to a connection and will communicate connection handles appropriately.

As QoS and traffic parameter mapping from RSVP to ATM as well as flow-to-VC management functions can be highly interdependent, especially when aggregation is taking place and a management function may decide to allocate extra resources in anticipation of further reservations, e.g. when a certain percentage of the available VC bandwidth is consumed, they deserve further studies treating QoS mapping, VC management, and CAC in an integrated manner. The outcome of this research will help to incrementally improve the Interworking Control module.

Although not shown in Fig. 11, CLIP is used to resolve the IP destination address of a flow to an ATM address if this information has not been cached from previous requests. Both the signalling demon and CLIP’s ATMARP demon (which in turn uses the UNI signalling demon when to control switched VCs) can either be called implicitly via the kernel or, as an alternative, explicitly from the Interworking Control module.
The Implementation Unit prototype is embedded in FlexTel’s Multi Purpose Switch/Router (MPSR) architecture, which offers a generic multiprocessor platform with a shared high speed bus. This system comprises Pentium based boards, ARM based ATM switching modules and various I/O modules which are connected via two either cell or frame based data busses. In combination with common operating systems (e.g. Linux) and additional, application specific software, such as the control framework shown in Fig. 11, the MPSR becomes a comprehensive and flexible environment with both switching and routing capabilities.

Hence, Fig. 11 also includes switch port cards (on the bottom left) that directly redirect native ATM VCs via a switching engine on an ARM board towards an output port. All other types of VCs are terminating in or originating from the Integration Unit control modules. Signalling messages arriving or leaving on the standard signalling VCs are handled by the UNI signalling demon whereas other switched VCs, namely the RSVP control VCs and the RSVP over ATM VCs, terminate in the IP module of the Linux kernel in the same way as the IP paths (on the bottom right of Fig. 11). From there, IP packets carrying the well-known RSVP port numbers are redirected to the RSVP demon while data packets are subject to routing and classification (to identify the flow they belong to) before they are scheduled to be sent. The settings of the classifier and scheduler are controlled by the RSVP demon by a set of Integrated Services specific data structures grouped in the Integrated Services Traffic Control Block.

![Fig. 11: Block Diagram of the Integration Unit](image-url)
Of course, new architectures have to be developed for demonstrating SRP over ATM, RSVP over SRP, RSVP peering with ATM etc. Those architectures will be considered when the RSVP over ATM implementation has proved to be working.

5 CONCLUSIONS

This paper has presented an architecture which integrates IP and ATM mainly on the basis of IETF Integrated Services. In order to accomplish QoS end-to-end, a comprehensive interworking architecture encompassing address resolution, routing, signalling and a traffic control framework is required. The outcome of the ISSLL working group of the IETF has given useful insights in this domain. Work in progress in DIANA mainly focuses on the specification, implementation and evaluation of signalling translation, and related with that, traffic and QoS parameter mapping in a so-called Integration Unit which is to be placed at the boundaries between IP and ATM networks.

DIANA’s networking model is based on IP as a common network layer and the assumption that end systems are connected to different link layers but use RSVP to provide the application with the control capabilities of the respective layer underneath and/or the next RSVP capable network element. However, also scenarios with both native ATM applications and RSVP applications have been studied carefully.

As the emerging IETF Differentiated Services framework indicates, aggregation of flows to reduce processing in core network elements is supposed to be an important issue for the Internet of the future. As has been pointed out, the Integration Unit may aggregate several flows to one VC when using a preventive scheme to trigger re-negotiation procedures. Unlike RSVP, the new IP reservation protocol SRP refrains completely from maintaining per-flow state in routers and could be an alternative in particular for adaptive applications. For sure, the integration of those and further aggregation mechanisms into existing networks and their evaluation is a challenging task for the future.

Acknowledgements

The authors would like to thank all people who are contributing to the progress in DIANA and Alexey Kuznetsov for many helpful discussions on the structure RSVPd and traffic control support in the Linux kernel. The work in DIANA would not be possible without the financing from the Commission of the European Union and the Swiss Bundesamt für Bildung und Wissenschaft. The assistance from both these organisations is therefore here also gratefully acknowledged.

REFERENCES


20


