Implementing Integrated and Differentiated Services for the Internet with ATM Networks: A Practical Approach

Gerald Eichler, Deutsche Telekom Technologiezentrum
Heinrich Hussmann, Dresden University of Technology
George Mamais and Iakovos Venieris, National Technical University of Athens
Christian Prehofer, Siemens AG
Stefano Salsano, CoRiTeL

This article reports on results of the international research project ELISA which is funded by the European Commission under contract ACTS AC310.

Series Editors:
K. Elsayed and L. Toutain.

ABSTRACT

This article reports on design, implementation, and preliminary experimentation of a network architecture that supports quality of service for Internet applications. It gives an overview of the various approaches toward communication networks that support application-specific degrees of QoS. Special emphasis is put on the integrated and differentiated services approaches and on combinations of them. A new architecture is described which aims to bring these concepts closer to practical realization in wide-area networks. The new architecture supports the integrated as well as differentiated services approaches in a smoothly integrated way, and uses the capabilities of an underlying ATM network to realize QoS. The enhancements to the existing network infrastructure are deliberately limited to the integration of a single new type of network element called an edge device. The potential benefits of such an architecture for various stakeholders are explained, and how the new architecture could be introduced smoothly in existing networks by small migration steps, also covering networks based on technologies other than ATM. It is shown that the approach can be scaled up to a very large QoS-aware overlay network for the Internet.

INTRODUCTION

The Internet gives a fascinating vision of a future global and user-friendly information infrastructure. Unfortunately, current Internet technology is technically not suitable for reliable multimedia services, since it does not provide adjustable quality of service (QoS). The Internet relies on a protocol layer (IP) which enables interoperation of various network technologies, but uses the least common denominator of the capabilities of underlying networks (best-effort quality).

Several concepts for QoS at the IP level have been proposed, in particular the Internet Engineering Task Force (IETF) integrated services (IntServ) and differentiated services (DiffServ) architectures. Moreover, there are already commercially available networks (e.g., asynchronous transfer mode, ATM) where a high potential for QoS support exists but is not exploited in the Internet. This article describes a new architecture that tightly integrates the mentioned approaches to QoS. This approach is based on a careful analysis of the main functional building blocks for QoS, and shows that rather different-looking concepts can be supported by a single homogeneous and elegant architecture. The described approach defines an open QoS architecture which provides flexibility to incorporate new concepts beyond those already supported. The close integration also achieves important synergy effects. For example, a QoS request (as defined by IntServ) may be mapped directly to an individual ATM connection or aggregated with other requests in a “controlled” DiffServ traffic class.

An overview of the main approaches toward a QoS Internet is given next, covering both IP-level concepts and ATM-based subnetworks of the Internet. We then describe the new architecture that realizes QoS in a unified way. The architecture relies on a new type of network element, the basic design of which is sketched. Practical deployment of the new architecture is made realistic by the existence of a migration strategy and a viable economic model, as outlined in the last section.

0163-6804/00/$10.00 © 2000 IEEE

Approaches Toward QoS Support in Networks

ATM

ATM provides optimal transport layer support for novel interactive multimedia applications. Application developers are interested in ATM because of its ability to set up, on demand, point-to-point virtual channels (VCs) with specified QoS. ATM provides support for different traffic classes (e.g., constant bit rate, variable bit rate, unspecified bit rate), allowing the application to specify its exact requirements (e.g., peak cell rate, sustainable cell rate). This information can be used to achieve high network utilization through statistical multiplexing.

For network operators, ATM has additional attractive features such as individual charging and billing for network usage, comparatively high security standards, and the ability to support both permanent connections (permanent virtual paths, PVPs, and permanent virtual channels, PVCs) and switched connections (switched virtual channels, SVCs).

ATM is suitable for wide area networks (WANs), where it is called broadband integrated services digital network (B-ISDN), as well as for local area networks (LANs). However, the trend of the last few years has shown that the deployment of an end-to-end ATM infrastructure is too expensive compared to competing technology (in particular, Fast Ethernet and Gigabit Ethernet in LANs). In addition, there is a lack of application programs constructed for the usage of pure ATM switched connections. For these reasons, application developers concentrate on the IP protocol stack. As a result, the current main usage of ATM is as a lower-layer technology for WANs, which in most cases carries IP-related traffic.

Internet Integrated Services

The need to support QoS-sensitive applications has led the Internet community to develop the Internet IntServ architecture. The basic concept was to foresee a set of service models to be provided in the Internet, besides the currently used best-effort model. Two service models were defined, guaranteed service (GS) and controlled load service (CLS). The former provides a strict bound on the delay and loss probability for packets of a given flow, under the condition that the flow complies with a traffic contract. The latter does not define the provided service in terms of exact values for delay or loss probability. Under the CLS model the packets of a given flow will experience delays and loss comparable to a network with no load, always assuming compliance with the traffic contract.

A key component of the IntServ model is the mechanism used for signaling QoS requests from application to network, called the Resource Reservation Protocol (RSVP) [1]. A sender host uses the RSVP PATH message to advertise the bandwidth requirements of a flow. The PATH message traverses several routers to its destination (or to its set of destinations in IP multicast). The receiving host uses the RSVP RESV message to reserve an amount of bandwidth. The RESV message traces back the path to the sender host, reserving the resources in the intermediate routers. Sophisticated support of multicast is built in, merging the reservations for the same flow coming from different receivers in the intermediate routers. RSVP is based on a soft state mechanism: reservations expire after a timeout period unless refreshed.

Once the reservation has been established, all the routers in the path have to recognize the packets belonging to a reserved flow and provide convenient handling. These actions can become an unacceptable processing burden when hundreds of thousands of different flows must be handled (e.g., in a gigabit core router). Similarly, the need to exchange and store per-flow information is another heavy burden for core routers. These scalability concerns have led the Internet community to investigate simpler solutions for the support of QoS.

Internet Differentiated Services

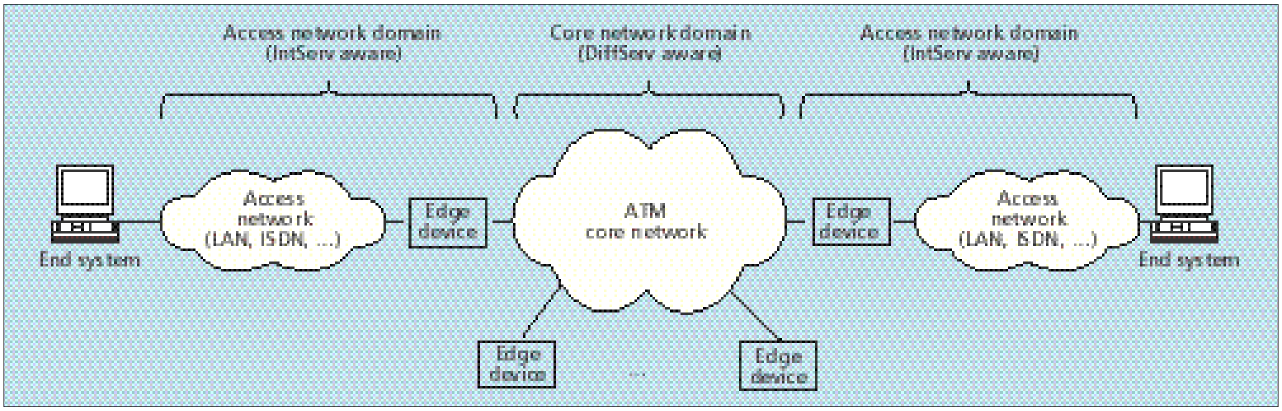
Due to the scalability problems of RSVP, a more scalable model has recently attracted much attention. With the newly introduced DiffServ model [2], user flows are only controlled at the edge of the network and then aggregated into a small set of traffic classes.

Traffic is classified and possibly conditioned while entering a network at so-called boundary nodes. Once classified, traffic will be attributed to different behavior aggregates, each of which is identified by a field in the IP header. This DiffServ code point is carried in 6 bits of what was formerly called the type of service (ToS) octet in the IP header and is now called the DS byte. Within the core network, packets are forwarded according to the per-hop behavior (PHB) associated with their DiffServ code point. Note that DiffServ is asymmetric, since only the sender controls the traffic class.

Current IETF standardization focuses on PHBs from which end-to-end services shall be constructed. The main concepts are as follows:

- Uncontrolled traffic classes, which offer qualitative, but no quantitative, service guarantees (priorities). Traffic is controlled at the ingress to the network; the admission policy usually does not consider end-to-end traffic aspects (e.g., it considers only the source and not the destination of IP flows). End-to-end aspects should be considered at a provisioning level in order to control the congestion phenomena. The definition of proper dimensioning/provisioning criteria is still an open issue. The overprovisioning of resources is of course a simple, but inefficient, solution. The basic example for this type of traffic class is the so-called IP precedence [3] set of PHBs, which comprises eight class selector code points with relative priority. Another example is the assured forwarding (AF) PHB group.
- Controlled traffic classes can provide end-to-end service guarantees based on per-flow admission control to the DiffServ core network. The most prominent example is the expedited forwarding (EF) PHB [4] for flows with known bandwidth, which is controlled at the ingress. In [4] it is proposed

The need to support QoS-sensitive applications has led the Internet community to develop the Internet IntServ architecture. The basic concept was to foresee a set of service models to be provided in the Internet, besides the currently used best-effort model.



■ Figure 1. The ELISA network architecture.

to aggregate all traffic flows into one traffic class and guarantee the bandwidth for all paths through the network.

COMBINED INTEGRATED/DIFFERENTIATED SERVICES

Since IntServ and DiffServ focus on reservations and scalable service differentiation, respectively, it is advantageous to combine both for an overall solution. IntServ can be used in the access network and DiffServ in the core network, as supported in the architecture described below. In parallel to the work described in this article, a discussion of this idea started in the IETF standardization [5]. In order to map IntServ flows to the core DiffServ network, a so-called edge device (ED) is used. The ED acts like an IntServ-capable router on the access network and like a DiffServ router in the core network. Note that it can also play the role of a boundary router, but this is not strictly necessary since the core network (operator) may employ separate boundary routers. In this combined approach, RSVP is tunneled through the core network, and its functionality is restricted to the access network. The main benefits of this model are:

- A scalable end-to-end IntServ service model with reasonable service guarantees in the core network
- Explicit reservations for access links, since bandwidth can be scarce in the access network
- Flexible access to a DiffServ core network with individual QoS for flows, in contrast to static DiffServ configuration

The remainder of the article presents an innovative architecture which extends the current state of the art described above.

THE ELISA ARCHITECTURE

OVERALL CONFIGURATION

This article reports on the EC-funded research project "European Experiment on the Linkage between Internet Integrated Services and ATM" (ELISA). The project has developed an architectural model which is referred to as the ELISA architecture. A prototype implementation of the architecture has been completed, and currently (October 1999) the integration tests have been carried out. Thus, the basic feasibility of the architecture has been proven already.

The ELISA model builds on the architectural framework discussed earlier, combining DiffServ and IntServ. As shown in Fig. 1, a clear separation is enforced between the access and core networks. This approach allows us to cope with the different evolutions of access and core network technologies. The ED represents the bridge between the two sections.

Examples for the access side, which are tried out in ELISA, are LAN (Ethernet and Fast Ethernet) access and narrowband ISDN (N-ISDN). Due to the use of IP technology, it is easy to integrate additional technologies in a seamless way. In fact, only a suitable network card has to be added to the ED together with its IP driver. Examples of potential technologies include digital subscriber line (xDSL) to support residential access with higher data rates, Gigabit Ethernet for business access, radio access for mobile and residential users, or passive optical networks for distributed services to residential users.

The core network is assumed to be ATM, providing SVCs and PVCs. Different capability levels of the ATM network can be handled by configuring the ED operations. Non-ATM core networks — for example, based on synchronous optical network (SONET) and using the DiffServ approach, can be handled as well. Multiprotocol label switching (MPLS) technology [6] is also worth mentioning, which is also proposed as a solution for the core transport section of IP networks. In a later section some comments on the relationship and possible interworking of our work and different approaches will be given.

The unifying technology for the access domain is IP. End-user hosts can use RSVP to signal their QoS requirements to the network. The EDs map these requirements into the most suitable operations in the core network. RSVP is restricted to the access section where no scalability concerns apply. Restricting RSVP to the access also makes the architecture relatively future-proof, since it can easily be adapted for different mechanisms to signal QoS requests (e.g., as they may appear in future object communication middleware). Such an upgrade only

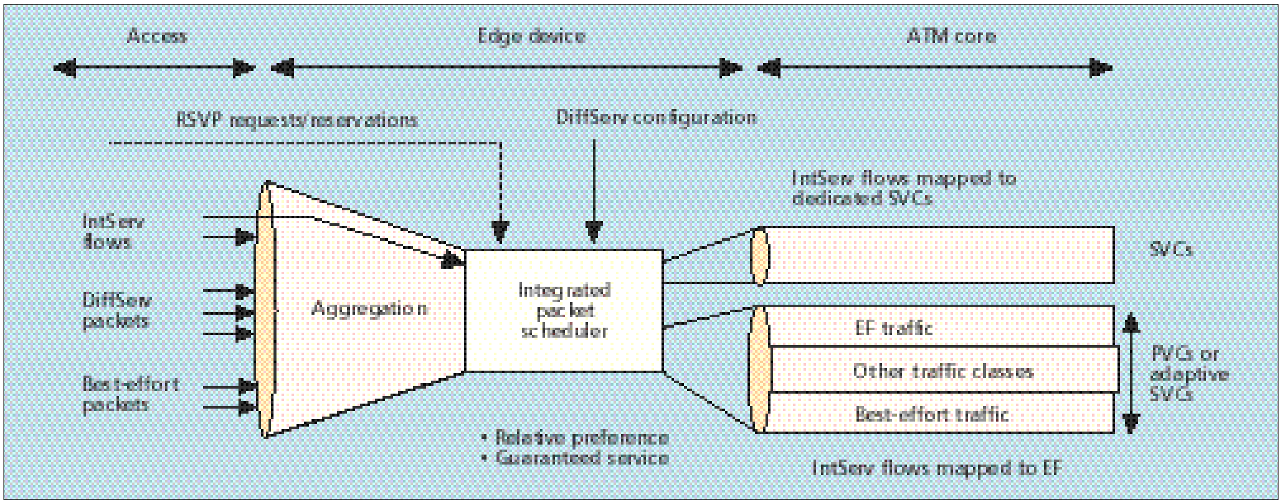


Figure 2. Edge device functionality.

involves parts of the ED, but does not affect the core network.

The role of the ED in the user plane is depicted in Fig. 2. Incoming IP packets belong to IntServ flows or DiffServ traffic classes, or they can be best-effort packets. The first task of the ED is to classify the packets, providing information to the packet scheduler. The scheduler uses the appropriate core network mechanism and performs appropriate queuing disciplines (e.g., based on traffic classes) where applicable.

The main novelty of the ELISA ED is that it provides a full spectrum of QoS realizations, ranging from a simple DiffServ model to full ATM QoS. As illustrated in Fig. 2, traffic with per-flow guarantees can be handled in the core network in several ways:

- Individual IntServ flows, requested by user applications through RSVP, can be aggregated and carried in the DiffServ EF traffic class. This makes resource reservation a local decision in the involved EDs, without any involvement of the core network. The EF traffic may be transmitted over the same ATM VC together with best-effort traffic, thanks to the admission control and queuing policy applied in the ED.
- Alternatively, IntServ flows can be mapped individually to ATM SVCs, which in addition enables precise control of delay and jitter.
- Different ways of handling DiffServ traffic can be supported. IP traffic may be marked as belonging to a specific service class by the end systems, in which case the ED is able to check the authorization of the user to do so. The ED is also able to do DiffServ marking itself, based on the knowledge of from which end system the traffic comes. Finally, the ED is able to convert IntServ requests into DiffServ traffic of the EF class, as explained above.
- Also, for the mapping of DiffServ traffic to ATM, the ED can be configured in different ways. Besides using statically preconfigured PVCs, ATM SVCs can be used to dynamically adapt the transmission capacity assigned to a traffic class. In this case, additional SVCs are added according to bandwidth demand, and the traffic is multiplexed onto several ATM VCs.

The selection of the core network mechanism is hidden from the end-user application, which just has to specify its QoS requirements (either by RSVP or using a DiffServ class). A sophisticated service management facility is used to configure the ED for the criteria for which of the above alternatives is to be used in which situation.

Service Realization

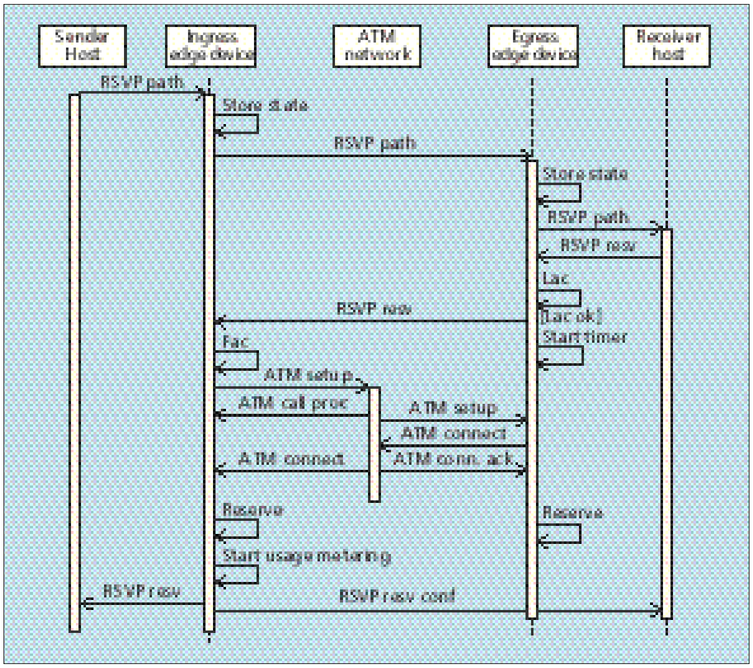
Realization of Internet Integrated Services — The ED has two options to support an IntServ flow, one of which is to map it into a dedicated ATM SVC. The procedure, which is basically aligned to the proposals under study within IETF [7], is depicted in Fig. 3. When an RSVP RESV message has reached the ingress ED, the flow admission control procedure (“fac” in the diagram) selects how to support the flow. If mapping onto an SVC is chosen, a new SVC is established starting from the ingress ED toward the egress ED.

The ingress ED also performs a QoS translation (i.e., it maps the RSVP traffic parameters into ATM parameters) [8]. When the ATM setup is completed, the ingress ED forwards the RSVP RESV message and starts the usage metering procedure.

Realization of Differentiated Services — To support the basic DiffServ functionality, a small number of code points with relative priority, access control, queuing, and scheduling functions are introduced in the ED. In analogy to [3], three different uncontrolled classes and one controlled class, shown in Table 1, are supported.

The ED has to ensure that the lower-priority classes never starve. To do so, it allocates a minimum amount of bandwidth to each class that cannot be used by any other class. Moreover, the lower-priority classes are allowed to use bandwidth that is not occupied by higher-priority traffic classes. The scheme described is realized by applying the well-known Weighted Fair Queuing (WFQ) and Class-Based Queuing (CBQ) algorithms.

In the prototype implementation, users can



■ Figure 3. Support of an IntServ flow by mapping to a dedicated ATM VC.

subscribe to one fixed traffic class, and every packet sent by a registered user is marked in the ED according to this class. It is an easy extension to also support the more flexible host marking, where the user terminal does the DiffServ marking by itself.

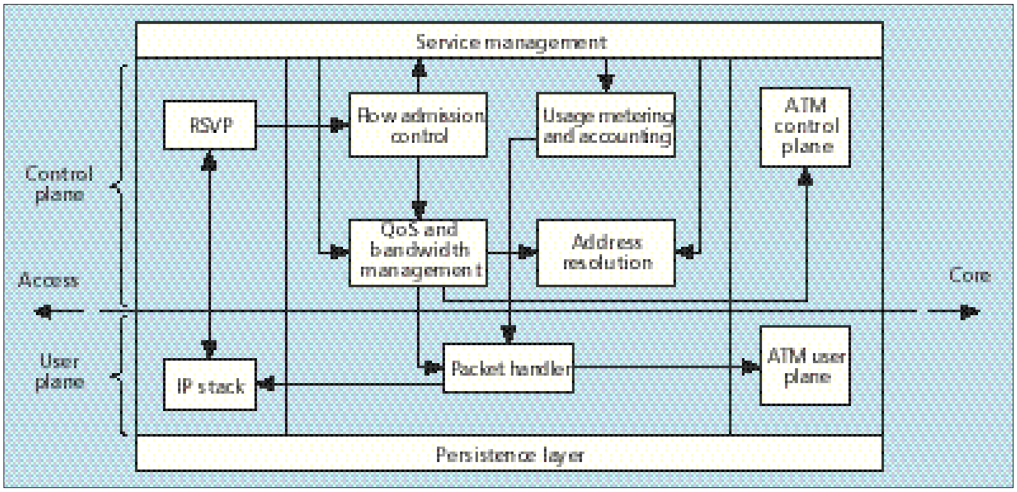
Realization of Combined Integrated/Differentiated Services — In the proposed architecture, the class-based treatment of DiffServ packet traffic is integrated with the flow-based treatment applied to IntServ traffic. The controlled EF traffic class is used for aggregated RSVP flows besides the use of dedicated ATM VCs, as discussed earlier. The EF class can be assigned to an ATM connection also used by other DiffServ classes. The exact mechanism by virtue of the CBQ algorithm is further described in the next section. With this organization, bandwidth can be commonly shared with other uncontrolled classes. Since the total bandwidth of the aggregated flows varies dynamically, one has to ensure that the bandwidth allocated for the EF class is controlled accordingly. Two options are considered: either the bandwidth allocated internally for the EF class increases in expense of the other traffic classes, or the bandwidth allocated to the ATM link increases. For the latter, ATM SVCs and link multiplexing are used, as explained in the next section.

When additional SVCs have to be established, the interaction between EDs resembles the support of dedicated ATM VCs shown in Fig. 3. In this case, ATM signaling may only be used to modify the bandwidth of an existing VC. Note, however, that this adaptation does not have to be performed for every new flow. In many cases it is sensible to overallocate ATM bandwidth, which can be used for uncontrolled classes.

Edge Device Design

The internal structure of the ED is generic and open so that several QoS architectures (not just IntServ, DiffServ, and ATM) can be supported. Therefore, the internal structure of the ED is split into several generic building blocks with well-defined internal interfaces. An internal view of the ED as a block diagram is depicted in Fig. 4. The functional block of the ED is divided in two horizontal and three vertical planes. The horizontal planes correspond to user and control plane. The vertical planes show the parts of the ED which are related to the access and ATM network as well as to the modules required for combining IntServ and DiffServ over ATM. Please note that arrows indicate usage dependencies between modules, not data or control flow.

User Plane Functionality — The user plane contains the functionality for transmission of traffic generated by user applications. The ED contains a standard IP protocol stack, which is responsible for packet routing, and the ATM related user plane protocols (ATM adaptation layer 5, AAL5) for converting IP packets into ATM cells. However, the standard IP forwarding



■ Figure 4. A block diagram of an edge device.

Service class	Short name	Code point	Controlled
Expedited forwarding	EF	101100	Yes
High priority	HP	111000	No
Priority	P	100000	No
Best effort	BE	000000	No

■ Table 1. DiffServ classes.

process cannot handle packet flows with specific QoS demands since their treatment is based only on the destination IP address. Therefore, the ED contains an additional module, the packet handler (PH), which provides all required extensions to the IP stack for QoS-based forwarding. Initially, any incoming packet is classified and policed (Fig. 5). In the second step, the standard IP process forwards the packet to one of the flow-based forwarders. Each flow-based forwarder is seen by the standard IP forwarding process as a normal IP interface. Within the flow-based forwarder, the packet experiences second-level routing based on source IP address, source and destination port number, and the protocol ID. The flow-based forwarder determines the ATM VC in which the specific packet must be sent. If the IP packet belongs to a flow that is handled according to the DiffServ model, the packet experiences third-level routing based on the value of the DS byte in the IP header. The DS classifier determines in which of the queues of the CBQ link-sharing mechanism the packet must be forwarded.

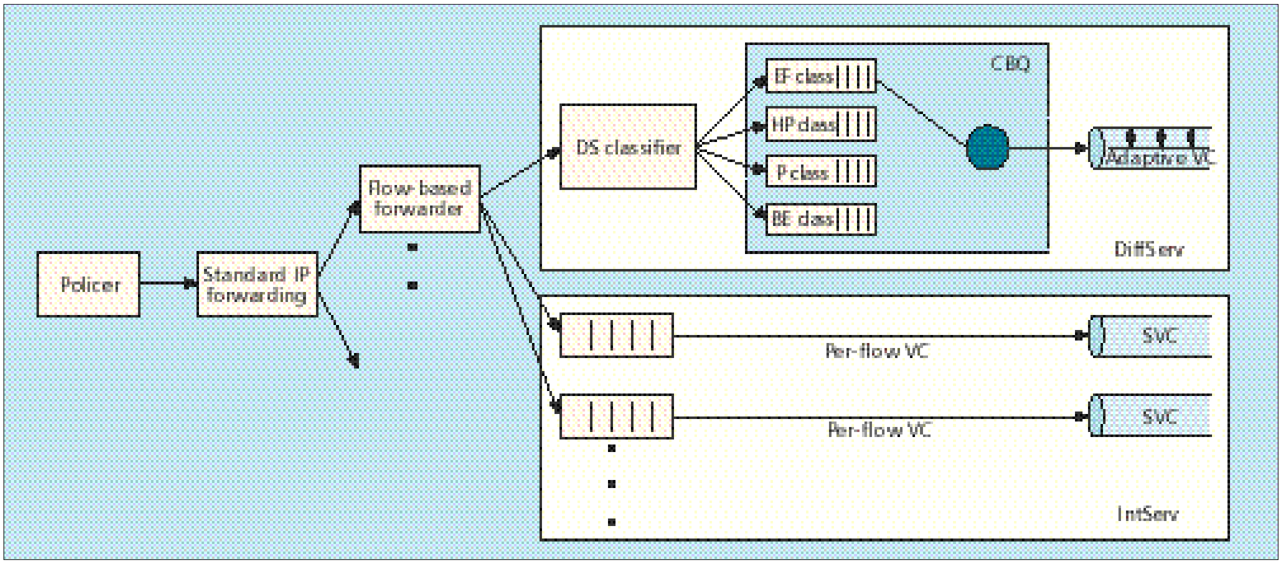
PH also implements a multiplexing layer which allows multiple ATM VCs destined to the same ED to appear as a single logical link (shown as adaptive VC in the above figure). Hence, any time a given link is congested, the ED establishes a new ATM VC which is added to the congested link. This mechanism can be replaced by explicit bandwidth modification signaling (as foreseen in ATM user–network interface, UNI, specification 4.0) if the underlying ATM network supports this feature.

Control Plane Functionality — The control plane contains the functionality for establishing and removing data paths through the network. The ED hosts a standard RSVP daemon that handles all reservation requests arriving from the access network. The ATM control plane, which in this particular implementation is UNI 3.1, provides the means to access the ATM network.

Flow admission control (FAC), and QoS and bandwidth management (QBM) extend the functionality of RSVP so that the advanced features of the PH can be utilized properly. QBM handles the internal resources of the ED. FAC has two main tasks: policy control and admission control. A supplementary functionality provided in the control plane of the ED is address resolution (AR), which resolves a destination IP address to the E.164 ATM address of the corresponding egress ED.

Moreover, the usage metering and accounting (UMA) module records all accountable events in the system, like the starting and ending times of an IP flow, the amount of bytes passed through a specific IP flow, the involved parties, and the user who made the specific reservation. Finally, service management (SM) maintains a database with all registered users and their profiles, and provides to the system administrator the means of configuring and administrating the ED.

The design of the ELISA ED has a number of advantageous features. The most important is that it can support several QoS architectures with few changes. As an example, assume that in the future the end systems do not use RSVP for signaling their QoS requests but another protocol, say, an object-oriented protocol based on Common Object Request Broker Architecture (CORBA). In this case, the RSVP module can be simply replaced by a new module. Of course,



■ Figure 5. An IntServ/DiffServ forwarding pipeline.

Applications	Service components	Service profile	ELISA mapping
Multimedia conferencing	Video, audio	Regular IP video conference.	BE, P, HP
Video telephony		Two-party ISDN-like videoconference	EF, Dedicated SVC
Premium multimedia conferencing		Regular video conference with QoS	EF, Dedicated SVC
File transfer	Data	Regular FTP	BE
On-demand retrieval		FTP/guaranteed bandwidth	EF, Dedicated SVC
Web navigation	Video, audio, data	Regular Web	BE
Premium Web navigation		Regular Web with QoS	P, HP

■ Table 2. Applications and service mapping.

if changes are needed in the interfaces of the new module with the FAC and IP stack modules, these two modules will be modified accordingly. The same applies to the ATM control and user planes. Instead of ATM, another underlying infrastructure (e.g., SONET) can be used by just replacing these two modules. Another advantage of the ELISA approach is that some currently internal modules can be put outside the ED. For instance, part of FAC can be put onto a remote bandwidth broker, which controls the available resources of a whole DiffServ cloud. By doing so, the DiffServ network can provide end-to-end services without requiring overdimensioning.

Prototype Development — The prototypical implementation of the ED carried out in the ELISA project realizes an ED on the platform of a Sun workstation with ATM and LAN interface cards. All software components shown in Fig. 4 have been implemented. Wherever adequate, publicly available software has been used and adapted; in particular, the RSVP implementation is based on the ISI implementation adapted and freely distributed by Sun Microsystems.

The nonmonolithic structure of the ED requires a standard communication for defining and realizing interfaces between the internal blocks. The Object Management Group's (OMG's) CORBA and its Interface Description Language (IDL) (in the Arachne OpenSource implementation) have been used, removing the necessity for customized protocols. Moreover, CORBA has many other significant properties such as interoperability across different platforms, abstraction, encapsulation, enhanced flexibility, and reusability. Please note that the use of CORBA does not influence the ED performance since CORBA requests are not involved in the packet forwarding process.

The design of the ED required the involvement of many people from many different companies, organizations, and countries. Precise communication between system designers was achieved by the adoption of the Unified Modeling Language (UML) for all internal analysis and design documents.

The software components of the ED are implemented in C/C++. The service management component keeps all its information in a relational database. A graphical user interface for service management has been developed as a Java applet communicating with the ED processes and the database. Thus, the system administrator can remotely configure the ED from any platform by using a Java-enabled browser.

Applications Using QoS

It is the number of attractive applications that decides the success of QoS enhancements of the Internet. For current Internet users, definitely the most attractive applications are traditional Internet services like the Web and FTP with improved performance. This can be achieved with DiffServ: The user registers and pays to get better service, and IP packets are marked at either the end system or the ED in order to actually benefit from higher priority. Besides some administrative interfaces, no application development is required in this case.

As soon as a QoS infrastructure exists, specific applications may become popular which support resource reservation. Currently, there are no commercial applications which support reservation; however, RSVP seems to have the support of major software developers. For demonstration purposes, project ELISA has implemented a set of applications that support QoS, for example, a videoconferencing environment based on MBone tools and running on UNIX (Linux) workstations. A list of the ELISA sample applications and of the mapping into QoS mechanisms is reported in Table 2.

A more detailed description of the ELISA user application and the QoS tools in the terminal can be found in [9]. The architecture described above supports various degrees of sophistication for the way applications make use of QoS. This flexibility is crucial for successful introduction of QoS into the market, since the introduction of QoS in the network is not linked to upgrades of user applications. For the combined IntServ/Diffserv case, it is worthwhile to explain the delivery process of an end-to-end service using video telephony as an example. The calling end terminal uses RSVP (PATH messages) to contact the called end terminal. The called end terminal uses RSVP to ask for the reservation of resources (RESV messages). With reference to Fig. 1, the RESV messages can enforce the reservation of resources in the two access networks. Within the core network domain the RSVP messages are not processed, and it is a task of the ED to make sure that the data flow can be supported by the DiffServ network with the required QoS.

Deployment of QoS in Existing Networks

An Economic Model for QoS

Professional users of the Internet require short response times and reliable services (i.e., QoS), and they are also willing to pay for them. On the other hand, Internet service providers (ISPs) are in high competition and are looking for new ways to create their revenue. Thus, there is an economic window of opportunity to introduce QoS-Internet as a commercial product if it can be done rather soon and in a way that gives a competitive advantage to those ISPs offering it.

An example of the current situation is the fact that many geographically distributed enterprises have invested in a corporate network infrastructure. For these customers, the QoS architecture from above is attractive to build up a virtual private network (VPN). An ED is placed at the border between each company site and the public network, and realizes QoS with the resources of the public network. Such a QoS-enabled VPN gives significant cost savings by outsourcing the administration of the corporate network. This type of customer will help public network operators achieve a reasonable return on investment for the still rather expensive long distance ATM infrastructure.

As can be seen from this example, there are several models for ownership of the ED:

- ED owned by corporate customer. In this case, the ED can be directly attached to a public ATM network, and no ISP need be involved.
- ED owned by an ISP (possibly leased to a customer). In this case, the ISP can realize synergy effects from multiple customers.
- ED owned by a public network operator. In this case, optimized solutions can be offered to the customer, since all relevant network components are under control of the VPN provider.

Each of these configurations has its own advantages, so competition among various solutions is likely, for the benefit of the customer.

MIGRATION STEPS

A significant aspect in the acceptance of a new network architecture is the migration path from existing network structures. The proposed architecture has the significant property that it can be introduced step by step without requiring extensive changes or updates in the operators' network infrastructures. Currently, many ISPs are using ATM as a lower-layer technology for carrying IP traffic. Starting from such a network structure, the migration steps are basically as follows:

- In a first step, EDs are introduced at the edges of an ATM network used in the domain of one ISP. DiffServ and to some extent also IntServ can be offered now to the customers of the involved ISP. The ED functionality makes it possible to offer IntServ (i.e., reservation by RSVP) over ATM even when only permanent ATM connections are available. This results in the first QoS islands in the Internet, but packet flows going beyond the borders of a specific island will not receive strict QoS guarantees.
- In a second step, the QoS islands resulting from the first step can be upgraded to support ATM switched connections, which enables more flexible realization of both DiffServ and IntServ. In this step, customers of the ISP have the additional choice of IntServ with precisely specified delay and jitter due to the individual mapping to ATM.
- In a third step, when ATM connectivity is available among QoS islands, the enhanced network services of the proposed architecture can be provided on a larger scale. Again, ATM PVC interconnection is sufficient for a first phase of this step, and cross-ISP switched ATM connections can be added in a later phase.

This migration strategy makes it realistic for the network infrastructure to be changed in small, controllable steps, and only where there is real demand and economic justification for QoS. The next two sections discuss special cases which appear during the migration process: interconnection of ATM islands through non-ATM networks, and the general question of scalability of the architecture to a large-scale network.

INTERWORKING WITH DIFFERENT APPROACHES

This section deals with two examples of the interworking of our solution with different approaches. First, the case of a non-ATM network portion is described; then some comments on the relationship with MPLS technology are given.

The proposed architecture can be applied even if non-ATM network portions exist within the core network domain, under the precondition that the non-ATM network portions support DiffServ. Obviously, the problematic point is where ATM VCs would have to be established between the ingress and egress EDs. To explain the basic idea of the solution, assume that the ingress and egress EDs are connected to ATM networks, but a non-ATM portion has to be crossed to connect the two EDs. In this case, two ATM VCs are established: one between the ingress ED and the ED at the (ingress) border to the non-ATM portion, and one between an ED at the (egress) border of the non-ATM portion and the egress ED. The two EDs at the border of the non-ATM portion are interconnected using DiffServ principles. Please note that the functionality of all involved EDs is still the same as in the IntServ realization proposed above. Simply, there are two ingress and two egress EDs. This idea can easily be generalized to the case where the path contains multiple non-ATM portions interleaved with ATM portions. Obviously, the dimensioning of the DiffServ parts affect the overall dimensioning of the network.

MPLS is an IETF standardization activity [6] which specifies how layer 3 traffic can be mapped to connection-oriented layer 2 transports like ATM and frame relay. The main idea is to add a label containing specific routing information to each IP packet and allow routers to assign explicit paths to various classes of traffic. Distribution of labels (i.e., signaling) can be done by different protocols, say, LDP and RSVP extensions. In our scenario with ATM VCs used as IP tunnels, it is an interesting alternative to use MPLS instead of ATM. Instead of ATM signaling, we would have to use a different protocol to set up an MPLS connection. Regarding scalability, we encounter the same problems as discussed below. An open issue is the usage of RSVP for setting up MPLS connections, since this can conflict with tunneling RSVP over the core network, as done in our approach.

There is an economic window of opportunity to introduce QoS-Internet as a commercial product if this can be done rather soon and in a way that gives competitive advantage to those ISPs offering it.

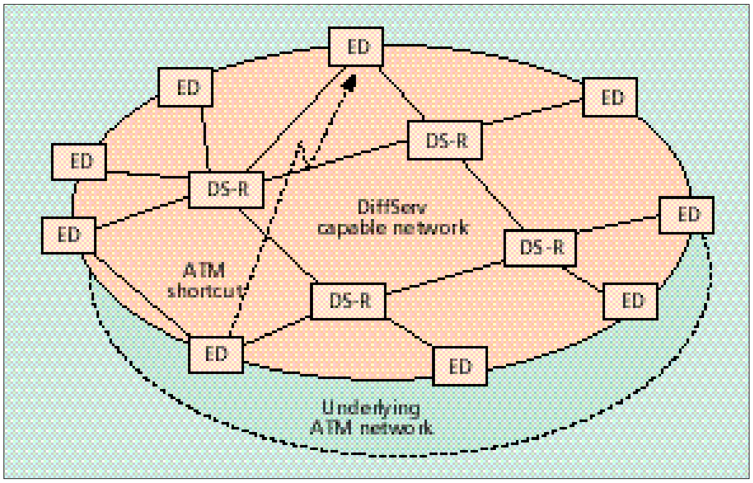


Figure 6. Target scenarios.

SCALABILITY

In a small-scale network scenario, corresponding to the first scalability steps, direct ATM connections among all the EDs can be assumed. This configuration has been deployed in the ELISA trial. Figure 6 shows a target scenario for a core network based on the proposed architecture. In such a large-scale network a fully meshed topology with direct links between all EDs is not feasible. Therefore, DiffServ-enabled routers (DS-Rs) are used in the core network. In such a scenario, good scalability is achieved since every per-flow operation is confined to the EDs and does not impact the core of the network, where only highly aggregated traffic flows are dealt with. The underlying network infrastructure (for interconnecting DS-Rs and connecting EDs) should be ATM in the ideal case, but non-ATM portions can be dealt with as described previously. Two issues must be considered when dealing with this target architecture.

The first issue is typical for the DiffServ approach: how to enforce QoS reservations in a network composed of a set of routers. The intermediate routers must be involved in providing the QoS between two EDs. There is a wide range of possible solutions, with different trade-offs between complexity and efficiency. The simplest approach is to keep local reservation/admission control in the EDs and follow a static approach. In this case a sort of advance preallocation of bandwidth is needed, which typically results in a loss of efficiency. More efficient solutions imply a dynamic exchange of information, which can involve EDs and the intermediate routers. Ad hoc devices (often called bandwidth brokers) can be introduced, with the purpose of controlling resource allocation in a DiffServ network.

The second issue is typical of wide-area IP-over-ATM architectures. Scalable mechanisms for translating IP addresses into ATM addresses are needed. These mechanisms (e.g., Next Hop Resolution Protocol, NHRP) are under study within IETF, but their interaction with RSVP is still unclear. A suitable straightforward extension to RSVP to support IP/ATM address resolution is described in [10]. If there is an ATM path between two EDs these mechanisms allow the establishment of direct ATM shortcuts.

CONCLUSION AND OUTLOOK

This article gives an overview of current approaches to supporting QoS for the Internet. An innovative approach has been described that integrates the major trends in QoS (IntServ, DiffServ, ATM) into a single scalable architecture. There is no contradiction between these trends. A “convergent network” bringing together the best of all these concepts is technically feasible and economically viable. Moreover, a generic and elegant design concept for an ED is presented which enables practical usage of the proposed architecture in well-defined migration steps, and is currently being realized prototypically.

For a future QoS-aware Internet approach, it can be expected that more sophisticated mechanisms will be used for QoS control. The concept of bandwidth brokers forms the starting point for a QoS architecture based on DiffServ principles in the core network, but adapts its resource allocations automatically to user demand. This leads to an overlay network of QoS administration agents that interact autonomously and intelligently, in order to optimize the network configuration under all possible conditions.

REFERENCES

[1] R. Braden *et al.*, “Resource Reservation Protocol (RSVP) - Version 1 Functional Specification,” IETF RFC 2205, Sept. 1997.

[2] S. Blake *et al.*, “An Architecture for Differentiated Services,” IETF RFC 2475, Dec. 1998.

[3] K. Nichols *et al.*, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers,” IETF RFC 2474, Dec. 1998.

[4] V. Jacobson *et al.*, “An Expedited Forwarding PHB,” IETF RFC 2598, June 1999; available at http://www.ietf.org/html.charters/diffserv-charter.html

[5] Y. Bernet *et al.*, “Integrated Services Operation Over Diffserv Networks,” IETF draft draft-ietf-issll-diffserv-rsvp-02.txt, June 1999; available via http://www.ietf.org/html.charters/issll-charter.html

[6] R. Callon *et al.*, “A Framework for Multiprotocol Label Switching,” IETF draft draft-ietf-mpls-framework-04.txt, June 1999.

[7] E. Crawley *et al.*, “A Framework for Integrated Services and RSVP over ATM,” IETF RFC 2382, Aug. 1998.

[8] M. Garret and M. Borden “Interoperation of Controlled-Load Service and Guaranteed Service with ATM,” IETF RFC 2381, Aug. 1998.

[9] T. Do *et al.*, “ELISA: European Linkage between Internet Integrated and Differentiated Services over ATM,” *IEEE Wksp. QoS Support for Real-Time Internet Apps.*, June 2–4, 1999, Vancouver, BC, Canada.

[10] R. Cocca, M. Listanti, and S. Salsano “Interaction of RSVP with ATM for the Support of Shortcut QoS Virtual Channels,” *Proc. 2nd Int’l. Conf. ATM*, Colmar, France, June 21–23, 1999.

BIOGRAPHIES

GERALD EICHLER graduated with a diploma from Chemnitz Technical University, Germany, Department of Information Technologies in 1993. His thesis examined parameter-driven methods for the logical design process of complex digital circuits. At the Research and Technology Centre Darmstadt he developed new satellite applications for one year. Since 1994 he has been with Deutsche Telekom’s network research department, first in the signaling and protocols section. Later he moved to network architectures with special focus on ATM and IP solutions. He has been involved in various international projects such as ACTS and EURESCOM giving support for ATM and IP aspects, network design, and multimedia applications. Since 1999 he has been working as an expert on IP QoS questions within the permanent scientific staff of T-Nova Deutsche Telekom Innovationsgesellschaft mbH preparing a new-generation German IP network.

HEINRICH HUSSMANN holds a diploma in computer science from Munich University of Technology and a doctoral degree from the University of Passau. He has done research and education at the Universities of Munich, Passau, and Dresden. For several years he was a systems engineer and team leader in the advanced development laboratory of the telecommunications division of Siemens. Since 1997 he is a full professor of computer science (software engineering) at Dresden University of Technology. He has participated in over 10 national and international projects in the area of software engineering and telecommunications, and is the author of over 50 scientific publications, including three internationally published books.

GEORGIOS K. MAMAIS received his Dipl.-Ing. degree from the Electrical and Computer Engineering Department of the National Technical University of Athens (NTUA), Greece, in 1995 and his Ph.D. from the Computer Science Division of the same department in 1999. His dissertation investigates new architectural schemes, based on the concept of intelligent networks (IN), for QoS provisioning over IP networks. He is a research associate in the Telecommunication Laboratory of NTUA, performing research in the areas of multimedia service design for broadband telecommunication networks, distributed systems, and architectures, and in IN. He has participated in the ACTS INSIGNIA project and currently is involved in the ACTS project ELISA. He is a member of the Technical Chamber of Greece and has received various academic excellence awards from NTUA and the Technical Chamber of Greece.

CHRISTIAN PREHOFER studied computer science at the Technical University in Munich and the University of Illinois at Urbana-Champaign, where he received a Master's degree in 1992. He obtained his Ph.D. in computer science in 1995 from the Technical University of Munich. At the TU of Munich, he worked on several German and European research projects on foundations of software engineering. In more than 20 publications and one book, he has contributed to formal methods, programming languages, and feature interaction. Currently, he is a systems engineer at Siemens in Munich and is also lecturing at the TU of Munich. He is working on QoS for IP networks, including architecture, signaling protocols, and traffic engineering.

STEFANO SALSANO [M] (salsano@coritel.it) received his degree with honors in electronic engineering from the University of Rome "Tor Vergata" in 1994, and his Ph.D. from the University of Rome "La Sapienza" in 1998. At the end of 1997 he joined CoRiTeL, a research institute on telecommunications, where he is now coordinating the IP-related research projects. He participated in the INSIGNIA ACTS project (on the integration of B-ISDN with IN) in the areas of architecture and protocol specification. He is participating in the ELISA ACTS project (on the support of QoS in IP/ATM networks). His research interests include architectures for broadband networks, integration of IP and ATM, and QoS support in IP networks. He is co-author of several papers on these topics.

IAKOVOS S. VENIERIS [M] received a Dipl.-Ing. degree from the University of Patras, Greece, in 1988, and a Ph.D. degree from the National Technical University of Athens (NTUA), Greece, in 1990, all in electrical and computer engineering. He is currently an assistant professor in the Electrical and Computer Engineering Department of NTUA. His research interests are in the fields of B-ISDN, high-speed switches, the Internet, internetworking, signaling, and performance evaluation. He has several publications in the above areas. He has received several national and international awards for academic achievement. He has been exposed to standardization body work, and has contributed to the European Telecommunications Standards Institute and International Telecommunication Union — Telecommunication Standardization Sector. He has participated in several European Union and national projects. He is a member of the Technical Chamber of Greece.

For a future QoS-aware Internet approach, it can be expected that more sophisticated mechanisms are used for QoS control.