A Taxonomy of Interaction Models for Internet and ATM Quality of Service Architectures

Jens Schmitt[1], *Lars Wolf*[1], *Martin Karsten*[1], *Ralf Steinmetz*[1,2]

1
Industrial Process and System Communications
Department of Electrical Engineering and Information Technology
Technical University of Darmstadt
Merckstr. 25 • D-64283 Darmstadt • Germany
2
GMD IPSI
German National Research Center
for Information Technology
Dolivostr. 15 • D-64293 Darmstadt • Germany

email: {Jens.Schmitt,Lars.Wolf,Martin.Karsten,Ralf.Steinmetz}@KOM.tu-darmstadt.de

Abstract

In communication systems there are two 'worlds' at the moment: Internet and ATM. Both possess Quality of Service (QoS) architectures which shall allow them to integrate services of data and telecommunications formerly performed by separate infrastructures. We believe that none of them will be able to oust the other. That means both will exist for at least the middle-term future. Therefore, an interaction between both appears to be necessary, especially in the field of distributed multimedia applications where both worlds 'meet' first. In order to perform gracefully, distributed multimedia applications require a certain QoS provision, in particular from the communication system. Thus, for such applications the existence of heterogeneous IP/ATM networks makes the interaction between Internet and ATM QoS architectures an important issue.

In this article a taxonomy of interaction models for the Internet and ATM QoS architectures is developed. We do not let our view be restricted by existing approaches for the interaction between ATM and Internet. Instead we will derive more unconventional models by regarding the possible communication patterns based on different topological variants for heterogeneous IP/ATM networks. The investigation is driven by applications' communication requirements. This is accomplished by examining possibly interacting applications and their communication patterns. The interaction models are contrasted and compared to each other and their assumptions and implications are shown.

The derived taxonomy of models allows us to classify proposed approaches for the interaction of Internet and ATM QoS architectures. Thereby we are able to identify the basic assumptions of these approaches and their corresponding restrictions.

Keywords: QoS, QoS in Heterogeneous Networks, IP/ATM Networks, Internet Integrated Services, ATM.

1 Introduction

Communication systems are traditionally divided into data and telecommunication. At the moment, Internet and ATM are the two major players in data respectively telecommunications. At first glance they are counterparts, at least in some respects, with different strengths and weaknesses. During the upcoming years, they can be expected to compete with each other - the Internet defending its position as a global internetwork and ATM trying to become one.

This work is supported in part by Deutsche Telekom AG, Darmstadt, Germany, and by Volkswagen-Stiftung, Hannover, Germany.

The competition between both is becoming harder because of the convergence process of data and telecommunications. This trend, i.e., that data and telecommunication systems are providing more and more similar services, is illustrated in Figure 1. If this trend continues, it means that eventually both are trying to serve the same applications, thereby getting into direct competition. Who will be the winner of that race is difficult to judge. Telecommunications people think that with ATM they have the technically more advanced technology. However, even if that is considered to be the case the Internet still has a lead with regard to that it offers worldwide connectivity already now.

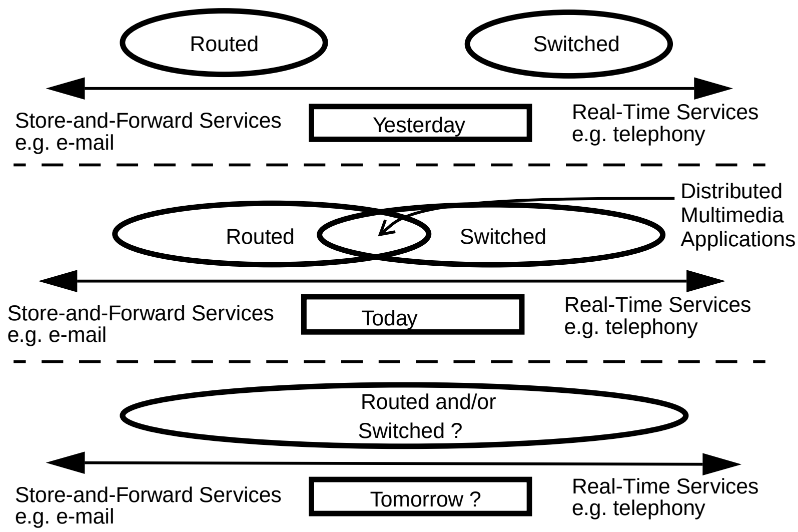

Figure 1: Convergence of the switched telecom world (e.g. PSTN, N-ISDN and now ATM) and the data communications/Internet world (based on routed datagrams).

The assumption that ATM will eventually become a global internetwork is very often doubted (by data communications people however), as Peter Newman (Ipsilon Inc.) in his keynote at the IEEE ATM Workshop ‘97 put it [29]:

> *’There may be a parallel universe in which ATM is the global internetwork - but it is not this one and it never will !’*

On the other hand, ATM will play a certain role, even if it does not become the successor of the Internet. Our opinion is that both will coexist for at least the middle-term future. This leads us to the conclusion that they must interact with each other. The question is how and on what level. Such an interaction is certainly not easy to achieve because totally different paradigms have to interwork, where one is based on connectionless, heterogeneous internetworking whereas the other one is based on connection-oriented, homogeneous networking.

The interaction of these two worlds is particularly desired in the relatively new field of distributed multimedia applications (they are in the middle of the services spectrum shown in Figure 1). These applications are dependent on the provision of QoS mechanisms by the communication system. This need has been observed in both worlds, and both have developed independently of each other QoS architectures that shall be able to provide integrated services.

Hence, one of the most important aspects of the interaction between both worlds is the seamless interworking between the QoS architectures of ATM and Internet. That means to enable the provision of QoS end-to-end regardless of what is inside the network and whether the communication endpoints are located in the same world or not, i.e. providing a homogeneous service over a heterogeneous network.

As a first step to fulfill this difficult task, the relationship of the two worlds regarding the interaction of their QoS architectures has to be analyzed. Do they have a provider-user relation or a peer-to-peer relation, and which effects would that have on applications?

As we shall see, the QoS architectures of the Internet and ATM are very different. So the question arises whether one of them has to adapt to the other, and if so, which one.

All these considerations need to be taken into account when trying to design interaction approaches for the QoS architectures of the Internet and ATM. What makes these approaches even more difficult at the moment is the fact that both architectures are still evolving at the moment and are thus very dynamic. On the other hand, this is still a chance to design both architectures with the possible interaction in mind. Indeed, some of these thoughts have already influenced the design of the architectures, as for example the Leaf-Initiated Join (LIJ) facility in ATM UNI 4.0 [6] signalling shows (although the adaptation to the Internet was surely not the only reason for the introduction of the LIJ mechanism).

In the next section we give a brief overview of existing solutions for the interaction of Internet and ATM. The Internet and ATM QoS architectures and their components are compared in Section 3. In Section 4, communication patterns in heterogeneous IP/ATM networks are examined and taken as a starting point for the derivation of a taxonomy of interaction models for Internet and ATM QoS architectures. These interaction models are presented and contrasted against each other in Section 5. In Section 6, a classification of proposed interaction approaches for the two QoS architectures is done using the taxonomy introduced in Section 5. Section 7 briefly considers the interaction between Internet and ATM if the so-called "Differentiated Services" architecture is used to provide QoS in IP networks. Section 8 summarizes and concludes from the observations being made.

2 Existing Solutions for Interaction

For the near and middle-term future it is crucial for ATM to interact with legacy networks for the protection of investments. This is especially true for asynchronous data communication. Whether ATM should try to interact on the real-time communication field is arguable from ATM's perspective, but will probably be necessary since the Internet has an existing broad fundament on which to build its Internet Integrated Services architecture (IIS) and thus will likely succeed to do so. Therefore ATM needs to interwork. On the other hand, if ATM interworks 'too well' with IIS, it might prevent its own QoS architecture to appear on the desktop.

For the interworking without defined QoS provision, several approaches have been developed over the time and have become more or less satisfying solutions to this relaxed problem. One of these solutions is the IETF's Classical IP over ATM ([26], [24]) with its extensions for multicasting, MARS (Multicast Address Resolution Server) [10], and short-cuts, NHRP (Next Hop Resolution Protocol)[27]. Another variant is ATM Forum's LAN Emulation (LANE)[3] and its successor Multi-Protocol over ATM (MPOA)[8]. IP switching [30] and similar solutions ([1], [22], [25], [31]) can be seen as representatives of a more revolutionary approach, which tries to identify data flows and build up VCs (Virtual Circuits) for them if they seem to be long-lived. The signalling protocols that build up the VCs are especially tuned for this kind of purpose and are no longer the original ATM signalling protocols. So IP Switching might be viewed not as an interaction approach with ATM, but a competing approach to ATM since essentially only the switching hardware of ATM is being used.

All of the solutions mentioned above do not support data flows requiring a predictable QoS. One could certainly use them providing the QoS based on the IIS mechanisms and use ATM SVCs (Switched Virtual Circuit) or PVCs (Permanent Virtual Circuit) as fast bit pipes. This is of course principally possible and inexpensive in terms of invested development effort, but ignores all the features provided by ATM and is very expensive in terms of usage of resources. It operates on ATM as if it were a 'dumb' point-to-point network or a leased line and does not make any use of the features provided by ATM like: the VC model (which allows for a presorting of flows on the data link layer), bandwidth management, or traffic management (traffic control and scheduling in hardware). Instead it duplicates these functions in the IIS architecture (in software, which is of course much less efficient). On the other hand, there is of course much less implementation complexity in this approach compared to other approaches that will be presented below. It should therefore despite of its obvious deficiencies be seriously analyzed with regard to the performance loss and resource wastage it incurs in comparison to the more sophisticated approaches.

The lack of QoS support of the existing solutions can be explained by the fact that they were designed for asynchronous data communication, for which QoS is not an issue. Only integrated services as they are envisioned now justify the further effort to make ATM's QoS features accessible even through higher layer communication systems protocols. The question is whether the existing approaches can just be extended to allow for QoS support or whether totally new approaches have to be developed to support QoS efficiently in heterogeneous IP/ATM environments. Most or even all of the current approaches that try to integrate IIS and ATM's QoS architecture build upon the extension of the existing solutions for asynchronous data communication. One of the main points of this paper is to reconsider whether this is sufficient. However, before regarding potential interaction models and approaches, we first shortly compare and contrast the QoS architectures of Internet and ATM.

3 Comparison of ATM and Internet QoS Architectures

In Figure 2, the most important components of both QoS architectures and their approximate semantic mapping onto each other is illustrated. We base our comparison on the latest (at the time of writing) specifications of the ATM Forum ([4], [5]) and the proposed standard RFCs of the IETF ([18], [36], [39]).

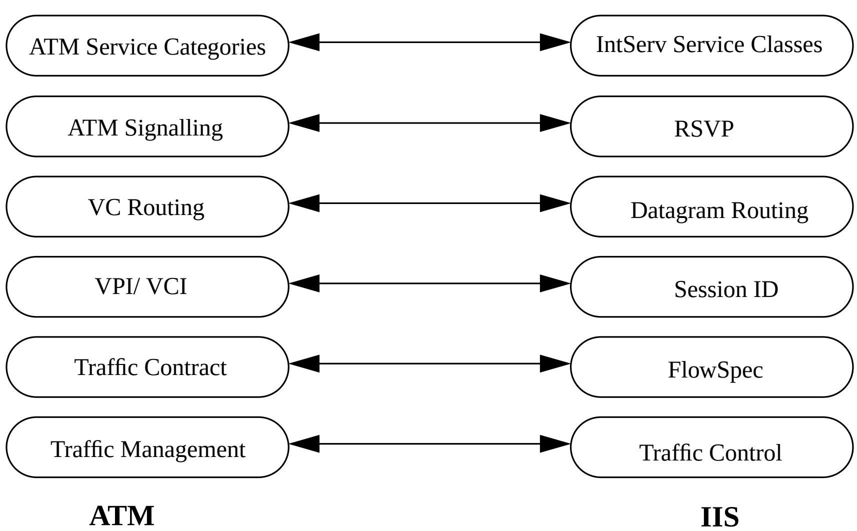

Figure 2: Mapping between ATM and IIS components (based on [2]).

Both architectures have very different capabilities and characteristics with regard to the signalling (the QoS procedures) and the QoS models (the QoS declaration or interface), but these discrepancies have to be overcome when interworking between ATM's QoS architecture and the IIS architecture. We consider the QoS models and procedures separately.

3.1 QoS Models

The most salient differences between the QoS models, i.e. the ATM Forum TM 4.0 [4] and the Integrated Services (IntServ) specifications ([36], [39]), are:

- packet-based vs. cell-based traffic parameters and performance specifications,
- the handling of excess traffic (policing): degradation to best-effort vs. tagging or dropping,
- and of course different service classes and corresponding traffic and service parameters.

While the traffic characterization of both QoS models is quite similar (token bucket rate+token bucket size/depth vs. PCR/SCR+MBS/MCR[1]), the service definitions differ substantially, such that a one-to-one mapping seems to be too 'semantic-lossy'. Thus, we think a mapping might have a dynamic or even adaptive n:m relation, i.e., the mapping is not fixed, it might adapt itself and one service class of IntServ might, depending on the actual values of the specified parameters, be mapped on different service classes in ATM and the other way around.

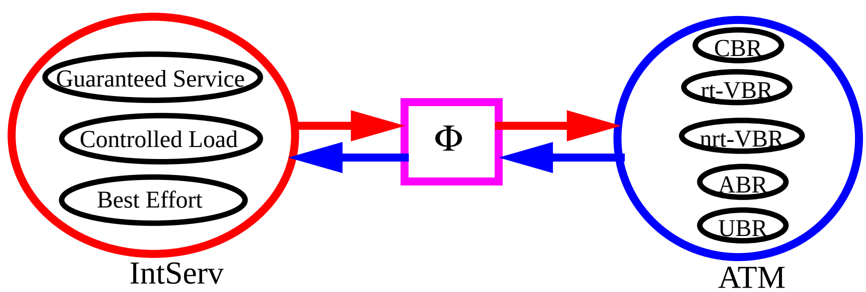

Figure 3: Mapping of Service Classes/Categories between ATM and IntServ.

The fact that the service classes of ATM and IntServ do not fit together very nicely, can be seen with IntServ's Controlled Load (CL) service [39] which seems to have no equivalent in ATM. That is due to the fact that the applications for which CL is attractive are adaptive applications (also supported by the dynamics of IIS' QoS provision, something considered but not yet implemented in ATM signalling), while in ATM's service model adaptive applications seem to be hardly recognized.

Although IntServ's Guaranteed Service (GS) [36] maps easier onto ATM's QoS model there is still no one-to-one mapping possible. While for token buckets with a small depth CBR (Constant Bit Rate) seems to be the right choice as a service category, for larger values of the token bucket depth this would lead to a wastage of bandwidth. Therefore to allow for a variable source not to waste too much bandwidth, GS should rather be mapped onto rt-VBR (real-time Variable Bit Rate) if the ratio of token bucket depth and token bucket rate exceeds a certain threshold value.

Besides the mapping of the service classes also the QoS parameters have to be mapped. While the two parameter sets certainly have an intersection, they are neither a subset nor a superset of each other, thus making an easy mapping impossible. A practical problem in this area is that the parameters are specified in different units: bytes vs. cells, and thus must be translated into each other taking into account the encapsulation and padding overhead.

Another problem is the treatment of non-conforming traffic, which in IntServ becomes best-effort traffic while it is at best being tagged (CLP (Cell Loss Priority) bit = 1) in ATM (but could also be

1. PCR = Peak Cell Rate, SCR = Sustainable Cell Rate, MBS = Maximum Burst Size, MCR = Minimum Cell Rate.

directly discarded depending on policies) and therefore is treated worse than ATM's best-effort traffic (UBR (Unspecified Bit Rate) or ABR (Available Bit Rate)). This means that traffic that is non-conforming in front of the ATM cloud would be treated better than traffic which does not conform inside the ATM cloud - an obvious mismatch.

A further QoS model mapping problem is caused by the fact that the traffic specification given by the end-systems might not represent the actually generated traffic inside the network, although applications adhered to the traffic contract. This is due to the fact that schedulers can only achieve an approximated fluid model. Therefore, reservations based on the traffic description given by the application might lead to situations where the policing functions of the ATM network might throw away data, which was conforming when entering the IP network but non-conforming when entering the ATM network. This is however not the application's fault and hence it should not be punished for it.

Another problem arises from the IIS concept OPWA (One-Pass with Advertising [32]). OPWA uses a so-called AdSpec to give the receivers an idea of which QoS they could expect from the network before they issue their reservation requests. So one of the questions arising with regard to OPWA is how to advertise an ATM cloud, which might consist of a very complex ATM network that from IIS perspective is however seen as one single hop.

3.2 QoS Procedures

While it is not easy to map the QoS models of the Internet and ATM, it is even more difficult to map their QoS procedures onto each other. This is due to the fact that they are built upon very different paradigms. While the signalling protocols of ATM are still based on the call paradigm used for telephony, the IETF viewed the support of a flexible and possibly large-scale multicast facility as a fundamental requirement. The most prominent differences between RSVP (Resource reSerVation Protocol [18]), which can be viewed as the Internet's signalling protocol, and ITU-T's Q.2931, on which all ATM signalling protocols are based, are discussed in the following:

Heterogeneous vs. Homogeneous QoS
While ATM only allows for homogeneous reservations, RSVP allows heterogeneity firstly for different QoS levels for receivers and secondly for simultaneous support of QoS and best-effort receivers. This mismatch in the semantics of RSVP and Q.2931 is a major obstacle to simple solutions for the mapping of the two.

Dynamic vs. Static QoS
RSVP supports a dynamic QoS, i.e. the possibility to change a reservation during its lifetime. ATM's signalling protocols however have been providing only static QoS so far (QoS renegotiations are currently under discussion as possible future extensions of ATM signalling protocols).

Receiver- vs. Sender-Orientation
The different designs with regard to the initiation of a QoS reservation reflect the different attitudes regarding centralized vs. distributed management, and also that the IIS architecture had large group communication in mind while the ATM model rather catered for individual and smaller group communications.

Hard State vs. Soft-State
The discrepancies between the ATM QoS architecture and the IIS architecture in how the state in intermediate systems is realized is another major obstacle to the interworking of both worlds since it leads to very different characteristics of the two QoS architectures. The soft-state of RSVP leads to a robust behavior of the protocol in case of link failures, whereas ATM's hard state is rather fragile to such situations. Yet, on the other hand hard state allows for a much more accurate and reliable QoS provision since RSVP can principally never guarantee that the QoS that was 'promised' by the network to the

application can be hold up for the whole duration of the session even if no link failure or similar situations occur. Hard-state is much more software-intensive due to the necessary fault and recovery management (which is reported to be approximately 90% of the signalling code of ATM [38]), while soft-state is much easier to implement since error recovery is built into the concept and does not explicitly have to be coded. A further advantage is that soft-state gives dynamic QoS at ’low cost’, while with hard-state dynamic QoS is difficult to implement. Soft-state on the other hand is less efficient than hard-state since it needs more work being done during ’runtime’ because of the periodic transmission of control messages even if nothing changes (pure refreshes) and because of the local processing overhead as for example timer management. Although, it must be considered that soft-state is very dependent on the quality of the dynamic routing, which, if it is poor leads to many route flaps, and thus to the fact that soft state means essentially no state.

Resource Reservation Independent or Integrated with Setup/Routing
Because RSVP is not integrated with routing flow setup and reservation are done asynchronously. This enables an independent evolution of routing and resource reservation mechanisms. Another advantage is the easy support of dynamic QoS. However a possibly major disadvantage in future may be that QoS routing is much more difficult to achieve than with ATM’s integrated connection setup/resource reservation mechanism (P-NNI [6] already supports a form of QoS routing).

Multicast Model
A further issue is the mapping of the IP multicast model on the signalling facilities in ATM for multi-party calls. While IP multicast allows for multipoint-to-multipoint communication, ATM only has point-to-multipoint VCs to emulate IP multicast by either meshed VCs or a multicast server. These are both work-arounds which can be shown to be sub-optimal for certain scenarios [37]. The proposed solution at the moment is MARS which however does not seem to be scalable enough for some applications envisioned in the Internet like DIS (Distributed Interactive Simulations), with around 10,000 group members joining and leaving rapidly.

Transmission of Control Messages
While in ATM separate control channels are used for the transmission of control messages of the signalling protocols, RSVP uses best-effort IP to send its messages (although it is proposed to give them a higher priority as soon as such facilities are available with IPv6).

Problematic is that both architectures are still changing quite rapidly, parameters are added and removed, new service categories are introduced and earlier ones are abandoned, etc. However, on the other hand these changes could also alleviate the mapping.

It shall be emphasized once more that many of the differences in signalling can be traced back to the roots of the two signalling mechanisms: RSVP is based on the observations made during the experimental MBone multicasts of the IETF meetings and therefore multicast is seen as very closely related to QoS in the IETF [15]. Q.2931 on the other hand is based on the traditional POTS (Plain Old Telephone System) signalling and its successor N-ISDN with its signalling protocol Q.931.

Among the capabilities of RSVP which are not supported by ATM are the most important: dynamic and heterogeneous QoS, and sharing and aggregation mechanisms for scalability within a session. These are characteristics which are especially useful in the multicast case. Capabilities of ATM which are not being (well) supported by IIS: the accurateness of QoS over the whole lifetime of a connection, and the scalability with regard to the number of sessions.

Besides the integration of the QoS models and signalling procedures, a practical and realizable solution needs to integrate further components as the security frameworks and the pricing/billing/accounting or policy control framework of both worlds. However these framework components have neither in ATM nor in the Internet reached a consensus, so that the interaction between those not yet existing components is difficult to anticipate.

In the next section we will enumerate all possible interaction patterns between ATM and Internet in order to identify the most important ones, which should be supported by an interaction model.

4 Enumeration of Interaction Patterns

We derive the potentially possible interaction approaches from topological observations of heterogeneous IP/ATM networks. This is illustrated in Figure 4.

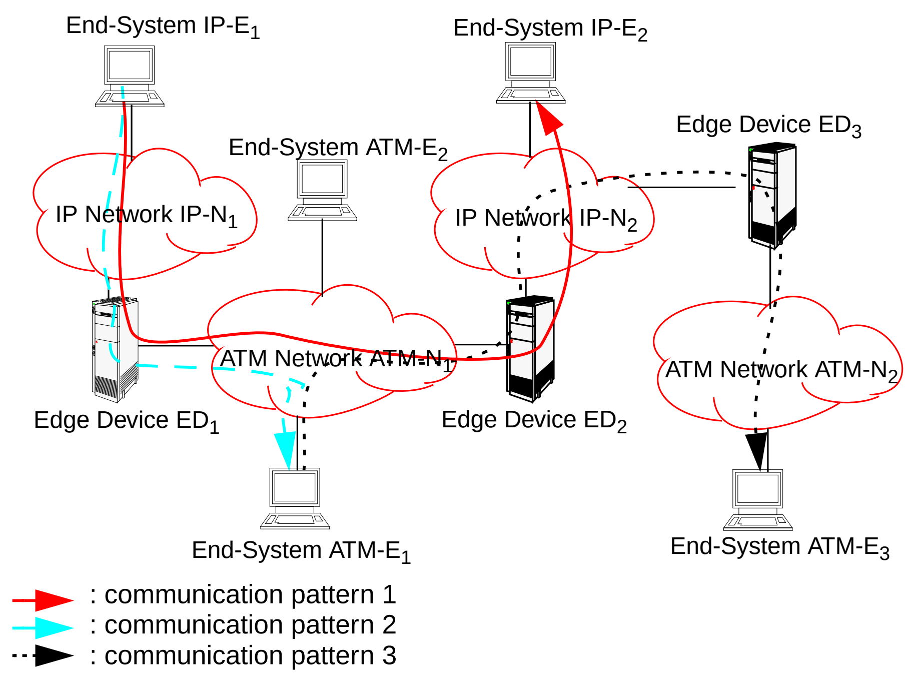

Figure 4: Possible Interactions between Internet and ATM.

From this figure the following principally possible interactions can be derived:

1. Communication between IP-connected end-systems with an ATM subnet lying on the transmission path, e.g. IP-E_1 would like to send data via ED_1 over ATM-N_1 via ED_2 into IP-N_2 to IP-E_2, symbolized by communication pattern 1 in Figure 4. An example application making use of that communication pattern could be an Internet video on demand service, where the video server would certainly like to gain from utilizing ATM's QoS abilities when transmitting over an ATM subnet to the receivers of the video stream.

2. Communication between an ATM-connected end-system and an IP-connected end-system, e.g. if IP-E_1 would like to send data via ED_1 over ATM-N_1 to ATM-E_1, symbolized by communication pattern 2 in Figure 4. An example application for this communication pattern could be a videoconference, which consists of some participants with IP-connectivity and some with ATM-connectivity.

3. Communication between ATM-connected end-systems with an IP subnet on the transmission path, e.g. if ATM-E_1 would like to send data via ED_2 over IP-N_2 via ED_3 into ATM-N_2 to ATM-E_3, symbolized by communication pattern 3 in Figure 4. An example application for this communication pattern could be the connection of two isolated ATM LANs via the Internet, e.g., for the purpose of building up a virtual private network. In order not to loose too much of the guarantees given by ATM, it would be favorable to be able to utilize IIS flows for the linking of the two ATM LANs.

Although in RFC 1821 [14] similar topological observations are made, communication pattern 1 is exclusively considered in more detail and therefore all of the IETF models for the interaction between IIS and ATM's QoS architecture are based on the support of communication pattern 1. We do not think that it is necessarily sufficient to constrain on one of the possible communication patterns.

However, which of the communication patterns are really worthwhile being investigated is in our opinion an open issue which depends on the topology of the future networking infrastructure.

We perceive two possible topological scenarios with two variants each for a future IP/ATM network:

1. ATM in the core of the network surrounded by all other network technologies to which users might be connected. If this might not seem realistic in the Internet, then this is at least quite realistic for corporate networks. This scenario however would be advocating for restricting the view on communication pattern 1 and 2. The two variants depend on whether native ATM connectivity will become an important case or not. If ATM will really play a significant role in end-systems of commercial environments or for residential users, then both communication patterns 1 and 2 have to be taken into account. If however ATM will only be a WAN solution, then a possible interaction model only needs to take into account communication pattern 1.

2. The alternative scenario is that ATM is just one of many link layer technologies. In that case all communication patterns might have some importance, even communication pattern 3. Again the two variants of this scenario depend on the question whether native ATM connectivity will be a reasonable option for end-systems. If ATM will not solely be a WAN solution then all three communication patterns will have to be taken into consideration. Otherwise, i.e. if ATM will be one of many WAN technologies, communication pattern 1 and 3 will potentially have to be supported by the interaction model.

The important point is that the future topology will heavily influence which interaction model should be chosen, i.e. the question whether ATM will play a role at all and if, will it play one solely in the WAN/MAN environment or in the LAN environment as well. Whether ATM will play an important role for end-systems depends necessarily on the question whether there will really be applications developed for it. It therefore depends very much on the fast introduction of a standardized ATM API (Application Programming Interface) and how it is accepted by application programmers which are well used to programming TCP/IP applications, but have mostly no experience with native ATM-mode applications. Another argument against the vision of ATM on the desktop has been raised: often a great gap between the services demanded by applications and the services provided by ATM is perceived. Here a mapping of cell-level guarantees and services to something more meaningful for applications like packets or frames would be needed. There seems to be at least one layer of abstraction missing - may be a perfect gap for IIS to fill out and continue IP's dominance in end-systems.

Another point which is often not considered is that the interaction approaches should also be made dependent on what the purpose of the internetwork is: global internetwork, private internetwork with centralized administration and control (of network engineering and protocol usage), or private internetwork with distributed management by independent organizations but on a scale that is still moderate.

The last two cases might be a niche for ATM because here homogeneity at least in the backbone could be achieved, especially in the centralized case. Whereas the global case, because of the heterogeneity which seems to be necessarily conjuncted to it, is always a strong argument for the use of Internet technology which accepts heterogeneity as a fact.

When considering mappings of the architecture for certain interaction approaches there are the following basic assumptions that influence heavily the design decisions for the realization of the interaction approaches and might even influence the choice of which interaction approach to use:

1. Accept the architectures as they are (passive strategy).

2. Try to change them to make interworking easier (active strategy).

Furthermore, in the following section we take a closer look at different interaction models. It turns out that these can be classified along the fact whether one of the architectures regards the other as a service provider or whether both view each other as peers that try to communicate with each other.

5 A Taxonomy of Interaction Models

Now that we have identified the possible interaction patterns between Internet and ATM, particularly with integrated services in mind, we will turn to the interaction models perceived by us that try to serve these communication patterns for QoS-dependent applications.

5.1 ATM Subordination Model

The ATM subordination model serves situations where communication pattern 1 in Figure 4 is valid. Its goal is to make as clever and efficient use of the ATM QoS facilities as possible.

In the ATM subordination model, ATM is viewed as a service provider for the IIS architecture. There are two different forms of how the interaction is designed with regard to ATM's contribution. On the one hand, ATM could be aware of the interaction and adapt itself actively, or on the other hand, ATM could remain unaltered and be passively used by the IIS architecture with all its constraints. In the latter case the IP over ATM signalling would have to be adapted, since the ATM QoS architecture would be regarded as fixed or as an external parameter which is not under control.

Since for the ATM subordination model the ATM network is viewed as a subnet, IIS is virtually overlaid over ATM. This leads to a potential duplication of functions like routing, multicasting and traffic management. Furthermore, this also bears the potential of hiding some of the good features of ATM, as for example PNNI's [6] QoS routing capabilities, and may lead to inefficient use of ATM network resources [23].

When using the ATM subordination model it must be recognized that the range of interaction patterns on the application level is strongly restricted to the case of IP-connected end-systems. Thus, there is no possibility to support communication across technological barriers, as e.g. a mixed videoconference of IP- and ATM-connected participants.

5.2 Partnership Model

The partnership model serves situations where the communication pattern 2 of Figure 4 applies, i.e. communication between ATM- and IP-connected end-systems on a peer-to-peer basis. This is why, from a more technical perspective, it may also be called peer or integrated model since it requires an integrated fashion of interworking between ATM and Internet. However, this model will only become interesting if ATM is successful enough as a complete protocol stack solution to be a serious competitor to the Internet protocols even on end-systems. If that happens the kind of interaction provided by the partnership model will probably be necessary to be supported. Hence, the partnership model has some importance, as it accepts ATM as a full-blown protocol stack that is able to operate end-to-end, and not solely as a data link technology as in the ATM subordination model. There is some justification to not reduce ATM on a link layer technology if one realizes that ATM offers facilities like: an API, routing, addressing and even session services. These are all elements which distinguish it from traditional link layer technologies and should in principle allow it to compete with IP as an end-to-end solution.

The partnership model introduces the need of a much tighter integration between ATM and IIS. The Internet is no longer just using ATM but they really need to interwork. This possibly leads to the fact that the QoS of ATM can no longer just be ordered through ATM's QoS interface, but the traffic management of both world's has to be integrated on a lower level. Since ATM's QoS architecture seems more powerful than the IIS architecture the fundamental problem is the mapping of ATM's QoS architecture on the IIS architecture, e.g. how to simulate ATM's accurateness and QoS reliability on IIS' rather crude and unreliable QoS provision.

Besides the need of a tight integration of the QoS architectures, there is also a need for an integration of very basic functions of communication systems like addressing, routing and data transfer. This is obvious for e.g. routing, since the data must be able to find its way through the combined IP/ATM network structure.

It can be seen that the ATM subordination is really only a small subset of the problems that have to be solved for the partnership model. While for example the sender-oriented reservation style of ATM versus the receiver-oriented reservation style of RSVP is not a big issue for the ATM subordination model, this discrepancy creates a really difficult problem for the partnership model. So it is far from obvious how to handle a SETUP message initiated by an ATM sender with the facilities in RSVP. In this case the receiver should initiate the corresponding reservation of which he however does not know anything yet.

5.3 Internet Subordination

The Internet subordination model serves situations where communication pattern 3 in Figure 4 is required. This is the case where an IP network acts as a transit network for communicating ATM-connected end-systems without direct connectivity. At first glance this might look exotic today, but it could have some relevance in case that there will be a scattered set of small islands of ATM networks. For example, for organizations that have geographically separated ATM LANs and which connect them via Internet to form a virtual private network (may be because its provider offers only IP connectivity), it is useful to preserve the ATM QoS as good as possible by using IIS. Nevertheless, the Internet subordination model should have exceptional character since it does not seem realistic to keep the QoS guarantees given by the ATM network over the Internet section of the transmission path, thereby causing an unpredictable QoS provision.

6 Classification of Interaction Approaches

We now use the taxonomy of interaction models derived in the previous section in order to classify existing interaction approaches between Internet and ATM with regard to their QoS architectures. We furtherly distinguish the interaction approaches with regard to whether they take a passive or active strategy.

6.1 ATM Subordination

The IETF favors the ATM subordination model since ATM is viewed as an important link layer technology, whose QoS capabilities should be utilized by IIS. However, as the IETF is definitely not ATM-fanatic as for example RFC 1821 [14] reveals when saying: 'While we believe that there is a range of capabilities in ATM networks that can be effectively used by a real-time Internet, we do not believe that just because ATM has a capability, the Internet must use it.', they do not consider a more integrated interaction model of the QoS architectures. The reason is that most people active in the IETF expect

ATM to be solely a WAN solution, and may be *the* WAN solution presenting the backbone of a future Internet, however ATM will never make it to the desktop in their view. So from their point of view a good solution could be to regard IIS and ATM as complementary techniques, where ATM is at the core, a place where its QoS routing feature is very desired, and IIS is at the edges of the network, where its ease of use is well desired.

There are two working groups in the IETF which, among others, treat the topic of interworking the IIS with ATM's QoS architecture.

The Integrated Services over Specific Link Layers (ISSLL) working group which favors a passive approach that can be seen as an extension of the Classical IP over ATM solution together with additional components like MARS and NHRP [19]. The approach is to make those components RSVP-aware and treat RSVP data flows differently from best-effort flows, i.e., setup special VCs for RSVP flows while all best-effort data share a common VC.

The other IETF working group is the Multi-Protocol Label Switching group, which encompasses several router/switch manufacturers that have built proprietary solutions for interworking ATM and IP based on ideas similar to IP Switching. The idea of MPLS is to make those proprietary solutions interoperable. There are topology-based approaches that setup VCs based on control information delivered by routing protocols ([22], [31]). A different variant of MPLS is followed by approaches in which the VC setup is triggered by some kind of identification of a significant data flow ([25], [1]). Since the VC setup is totally controlled by IP and no longer by the ATM control plane we call this an active interaction approach. The obvious extension of triggering a VC setup by RSVP control messages is the envisioned approach of MPLS to the interaction of IIS and ATM ([20], [23]).

While the work in ISSLL follows the paradigm of 'IP over ATM', the MPLS work is better described by 'ATM under IP', thereby emphasizing the active subordination of ATM in relationship to IP. In the area of passive subordination models, there has also been a considerable amount of individual research outside any standard bodies, as it is described in ([16], [17], [33], [34], [35]). Most of this work identified the fundamental design problems of the interaction and some initial implementation experiences are reported.

The ATM Forum currently investigates how its MPOA scheme can be extended to support QoS-aware network layers, which could be IIS beyond others [8]. Of course, a passive strategy is envisioned.

Between active and passive approaches there is a potential continuum of partially passive and partially active approaches. The fact that ATM signalling has been extended in order to make a better fit with IIS (among other reasons of course) could be regarded as a kind of hybrid between active and passive approaches. For example, the LIJ feature in ATM Forum's UNI TM 4.0 [5] is such a case.

6.2 Partnership

While the ATM subordination model is considered simultaneously by different groups, the partnership model has not gained much attention yet. This is certainly due to its high implementation complexity. This complexity is also the reason why it seems that only active approaches can make sense for the partnership model. Otherwise it does not seem feasible to overcome the discrepancies between the QoS architectures.

To some extent one could view the work of the ATM forum with regard to an Integrated PNNI (I-PNNI [7]), i.e. the use of a single routing algorithm, as one step into the direction of the partnership model, although in its current draft version it does not include any details with regard to QoS provision.

To cope with the complexities of the partnership model it might be a reasonable approach to interwork between the Internet and ATM not on the network layer but on higher layers. An example for the partnership model on the application layer taken by the ATM Forum at the moment is the VTOA (Voice and Telephony over ATM) Phase 2 work [9]. This tries to approach the interworking between ATM and

Internet voice transportation. However, an interworking on the network layer with ‘asymmetric’ end-systems seems to be a more fundamental answer to the problem, which certainly depends on the number of applications the two worlds are really sharing. However, as discussed at the beginning of this article, this set of application seems to be growing.

The future of interaction approaches following the partnership model very much depends on the existence of ATM end-systems. While ATM on the desktop is often seen as an unlikely scenario, it should not be neglected that there are many other devices like cameras, videophones, set-top boxes, etc., which are good candidates for pure ATM-connectivity. If they are supposed to be part of a global internetwork based on IP, the partnership model is the only possible solution.

6.3 Internet Subordination

Similar to the situation with the partnership model, the Internet subordination model has gained very few attention. With respect to the value of a solution for the Internet subordination model this is not surprising, since it has to be conceded that the situation of overlaying an ATM communication system onto an IIS system will certainly be a very special case. However as already mentioned, it could be interesting for, e.g., the setup of a virtual private network between two isolated corporate ATM networks.

From a technical point of view something distantly similar has been developed by the Cornell University and Connectware Inc.: Cells in Frames(CIF), i.e. ATM cells in Ethernet frames [21]. This concept has also gained some attention in industry which is documented by the fact that an industrial consortium, the CIF Alliance, was founded. The idea is to emulate ATM end-to-end in order to give end-systems access to all the QoS capabilities of ATM. This is the inverse situation to the ATM Forum LAN Emulation, where ATM emulates an Ethernet or Token Ring network and therefore hides all its QoS capabilities. However, what would be needed to really accomplish the kind of interaction the Internet subordination model implies, is ’Cells in Packets’, i.e. ATM cells in IP packets. By interacting on the network layer it would then be possible to cross routers.

The fundamental problem of the Internet subordination model is that the accuracy of ATM’s QoS can only be approximated by IIS but never be guaranteed, since the mapping from ATM QoS service categories and traffic and QoS parameters into IntServ terms seems very problematic. However, if QoS communication between unconnected ATM networks is required or desired, the Internet subordination model is a possible solution, and of course still better than delivering ATM traffic over the best-effort Internet.

An active strategy that approaches the problem of the mismatch between the service classes in ATM and IIS could be to introduce new service classes especially for transiting ATM traffic via the Internet.

Table 1: Classification of Proposed Interaction Approaches

	passive	**active**
ATM Subordination	**IETF ISSLL:** RSVPoATM + MARS + NHRP **ATM Forum:** (MPOA) **individual work**	**IETF MPLS:** topology-driven, dataflow-driven, request-driven
Partnership		**ATM Forum:** (I-PNNI), (VTOA)
Internet Subordination		**(CIF)**

6.4 Summary

Table 1 gives an overview of proposed interaction approaches and their classification into our taxonomy of interaction models. Once more it becomes obvious that so far only a very limited area of the whole 'interaction space' is under serious investigation (the approaches in parenthesis do not really represent solutions, but only tend into the direction of the associated interaction models).

7 Differentiated Services

We have based our discussion of interaction models on the assumption that IIS will be the QoS architecture of the Internet. Due to concerns about the scalability of IIS, a new approach called *Differentiated Services* (DiffServ) has been proposed in the IETF [11]. In the DiffServ architecture, it is planned to define standard forwarding semantics for certain types of packets, which are marked by hosts and/or edge routers. Concatenation of these forwarding semantics leads to certain traffic classes. A *Service Level Agreement* (SLA) describes a traffic profile between one or many network providers and a user or between multiple network providers. Such an SLA establishes a pipe with certain absolute or relative QoS characteristics along a data path (or parts hereof).

To assume DiffServ as the QoS architecture of the Internet certainly leads to major modifications with regard to concrete interaction approaches between Internet and ATM. There will be different service classes which have to be mapped onto ATM's service categories uni- or bidirectionally (as an example, consider the service classes that are proposed in the current drafts for Differentiated Services ([31], [13], [12]), shown in Figure 5). About the QoS procedures not much can be said at this moment, since the DiffServ working group explicitly excluded these issues so far.

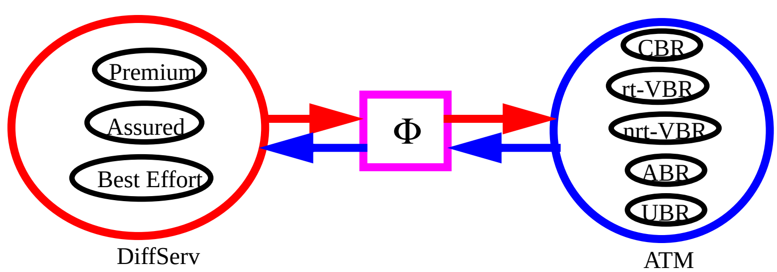

Figure 5: Mapping of Service Classes/Categories between ATM and DiffServ.

The possibility of mapping DiffServ and ATM largely depends on the dynamics of SLAs. Providing Differentiated Services over ATM links seems to be quite simple, because the following obvious mapping can be envisaged: each deterministic SLA(Premium) could potentially be assigned to a CBR or VBR VC, while in the case of statistical QoS guarantees (Assured), multiple ABR VCs with different MCRs could be used with each VC representing one service level. If DiffServ SLAs turn out not to be highly dynamic, PVCs might be an appropriate choice for providing these services.

On the other hand, mapping ATM QoS onto DiffServ SLAs imposes a number of requirements on DiffServ. To support a large number of fine-grained, highly dynamic SVCs, QoS procedures have to be defined to automatically establish appropriate SLAs. On the other hand, DiffServ inherently aggregates traffic flows by having only limited space for marking packets (currently under discussion; restricted to at most the IP TOS byte). It is not clear how this affects the QoS requirements of ATM traffic. Finally, hard end-to-end delay guarantees are currently not considered in the initial development phase of DiffServ. To this end it cannot be foreseen, whether these requirements can be met by the DiffServ architecture.

8 Summary

In our discussions of the interaction approaches of the QoS architectures of the Internet and ATM we started with a reasoning why this interaction seems necessary to us. We do not believe that one of the two will be able to totally oust the other one, but both will be with us for quite some time. However, as they tend to serve more and more the same applications due to the pertaining convergence process of data and telecommunications, they have to interwork with each other to fulfill application demands.

Since new and, also from an economic perspective, interesting applications like videoconferencing and video-on-demand services are run or will be run in both worlds, it is only natural that a seamless interworking between both worlds is demanded. For example, a videoconference should, from a user's perspective, of course neither be constrained on Internet-connected participants nor on ATM-connected participants, but should allow for mixed videoconferences with participants of both worlds.

Based on topological considerations and application scenarios we derived the required communication patterns for an interworking between ATM and Internet. Which of the communication patterns will be the prevailing ones, depends on many factors. On the one hand there are technical issues, like the fast introduction of an API to native ATM services and the existence of pure ATM end-systems such as videophones, video-servers, set-top boxes or cameras based on ATM. On the other hand, economical factors, as for example the protection of investments, have to be taken into account as well.

Based on the communication patterns we developed a taxonomy of possible interaction models, which were:

- the ATM subordination model,
- the partnership model, and
- the Internet subordination model.

We contrasted and compared these models to each other, mainly based on their applicability with regard to future topologies of internetworks combining Internet and ATM technology. Furthermore, we used the derived taxonomy to classify existing and proposed interaction approaches. Thereby we showed that only a small part of the 'interaction space' is currently under serious investigation.

In our investigation of the interaction of Internet and ATM QoS architectures we assumed IIS as the QoS architecture of the Internet. However, the currently developed Differentiated Services architecture is an alternative for providing a "better than best-effort" service over the Internet. Therefore we also had a brief look at the interaction between DiffServ and ATM.

References

[1] A.Acharya, R.Dighe, F.Ansari: *IPSOFACTO: IP Switching Over Fast ATM Cell Transport*, Internet Draft, work in progress, July 1997.

[2] A.Alles: *ATM Internetworking*, White Paper, Cisco Systems Inc., May 1995.

[3] The ATM Forum: *LAN Emulation over ATM Version 1.0 Specification*, January 1995.

[4] The ATM Forum: *Traffic Management (TM) Specification 4.0*, April 1996.

[5] The ATM Forum: *User-Network-Interface Signalling (SIG) Specification 4.0*, July 1996.

[6] The ATM Forum: *Private Network-Node Interface (PNNI) Signalling Specification*, February 1996.

[7] The ATM Forum: *Issues and Approaches for Integrated PNNI (Draft), April 1996.*

[8] The ATM Forum: *MPOA Version 1.0*, July 1997.

[9] The ATM Forum: *VTOA Desktop Baseline Text (Draft)*, January 1997.

[10] G.Armitage: *Support for Multicast over UNI 3.1 based ATM Networks*, Internet *RFC 2022*, November 1996.

[11] D. Black, S. Blake, M. Carlson, E. Davies, Z. Wang, W.Weiss: *An Architecture for Differentiated Services*, Internet Draft, work in progress, May 1998.

[12] D. Black, S. Blake, M. Carlson, E. Davies, B. Ohlmann, Dinesh Verma, Zheng Wang, Walter Weiss: *A Framework for Differentiated Services,* Internet Draft, work in progress, May 1998.

[13] F. Baker, S. Brim, T. Li, F. Kastenholtz, S. Jagannath, J. Renwick: *IP Precedence in Differentiated Services Using the Assured Service*, Internet Draft, work in progress, April 1998.

[14] M.Borden, E.Crawley, B.Davie, S.Batsell: *Integration of Real-Time Services in an IP-ATM Network Architecture*, Internet RFC 1821, August 1995.

[15] R.Braden, D.Clark, S.Shenker: *Integrated Services in the Internet Architecture*, Internet RFC 1633, 1994.

[16] A.Birman, V.Firoiu, R.Guerin, D.Kandlur: *Support for RSVP-based Services over ATM Network*, IEEE Global Internet, 1996.

[17] T.Braun, S.Giorcelli: *Quality of Service Support for IP Flows over ATM*, Proc. KIVS '97, February 1997.

[18] R.Braden, L.Zhang, S.Berson, S.Herzog, S.Jamin: *Resource Reservation Protocol (RSVP) - Version 1 Functional Specification*, Internet RFC 2205, September 1997.

[19] E.Crawley, S.Berson, L.Berger, F.Baker, M.Borden, J.Krawczyk: *A Framework for Integrated Services and RSVP over ATM*, Internet Draft, work in progress, February 1998.

[20] R. Callon, P.Doolan, N. Feldman, A. Fredette, G. Swallow, A. Viswanathan: *A Framework for Multiprotocol Label Switching*, Internet Draft, work in progress, November 1997.

[21] Cells in Frames Alliance: *ATM over Anything - Cells in Frames*, Cornell University, White Paper, 1996.

[22] V. Firoiu, J. Kurose, D. Towsley: *Performance Evaluation of ATM Shortcut Connections in Overlaid IP/ATM Networks*, University of Massachusetts, CMPSCI Technical Report 97-40, July 1997.

[23] O. Fourmaux, S.Fdida: *Multicast for RSVP Switching - An Extended Multicast Model with QoS for Label Swapping in an IP over ATM Environment*, Telecommunication Systems Journal (this issue).

[24] J.Heinanen: *Multiprotocol Encapsulation over ATM Adaptation Layer 5*, Internet RFC 1483, July 1993.

[25] Y.Katsube, K.Nagami, H.Esaki: *Toshiba's Router Architecture Extensions for ATM: Overview*, RFC 2098, February, 1997.

[26] M.Laubach: *Classical IP and ARP over ATM*, Internet RFC 1577, January 1994.

[27] J.Luciani, D.Katz, D.Piscitello, B. Cole, N. Doraswamy: *NBMA Next Hop Resolution Protocol (NHRP)*, Internet RFC 2332, April 1998.

[28] K. Nichols, S. Blake: *Definition of the Differentiated Services Field (DS Byte) in the IPv4 IPv6 Headers*, Internet Draft, work in progress, May 1998.

[29] P.Newman: *IP+ATM: If You Can't beat Them, Join Them*, IEEE ATM Workshop '97, Slide Presentation, May 1997.

[30] P.Newman, G.Minshall, T.Lyon: *Flow Labelled IP: A Connectionless Approach to ATM*, Proc. of IEEE Infocom '96, vol. 3, 1996.

[31] Y.Rekhter, B.Davie, D.Katz, E.Rosen, G.Swallow: *Cisco Systems' Tag Switching Architecture Overview*, Internet RFC 2105, February 1997.

[32] S.Shenker, L.Breslau: *Two Issues in Reservation Establishment*, Proc. of ACM SIGCOMM '95, August 1995.

[33] L.Salgarelli, A.Corghi, M. Smirnow, H. Sanneck, D. Witaszek: *Supporting IP Multicast Integrated Services in ATM Networks*, Internet Draft, work in progress November 1997.

[34] L.Sagarelli, M.DeMarco, G.Meroni, V.Trecordi: *Efficient transport of IP Flows Across ATM Networks*, IEEE ATM '97 Workshop Proceedings, May 1997.

[35] A.Schill, S.Kühn, F.Breiter: *Internetworking over ATM: Experiences with IP/IPng and RSVP*, Computer Networks and ISDN Systems, Vol. 28, 1996.

[36] S.Shenker, C.Partridge, R.Guerin: *Specification of Guaranteed Quality of Service*, Internet RFC 2211, September 1997.

[37] R.Talpade, M.Ammar: *VC Consumption in the MARS architecture for IP Multicast over ATM*, Position Paper at the IP/ATM Workshop at the University of Washington, November 1996.

[38] J.Turner: *State Management in IP+ATM networks*, Slide Presentation at the IP/ATM Workshop at the University of Washington, November 1996.

[39] A.Viswanathan, N.Feldman, R.Boivie, R.Woundy: *ARIS: Aggregate Route-Based IP-Switching*, Internet Draft, work in progress, March 1997.

[40] J.Wroczlawski: *Specification of the Controlled-Load Network Element Service*, Internet RFC 2212, September 1997.