Internetworking over ATM: Experiences with IP/ IPng and RSVP

Alexander Schill <schill@ibdr.inf.tu-dresden.de>

Sabine Kühn <kuehn@ibdr.inf.tu-dresden.de>

Frank Breiter <breiter@ibdr.inf.tu-dresden.de>

Abstract

This paper describes recent experiences with evaluating and implementing advanced internetwork communication protocols on top of ATM. First, performance results with conventional TCP/IP over ATM based on Digital Equipment's Gigaswitch /ATM are reported. It becomes obvious that current protocols must be tuned specifically in order to exploit ATM performance. In order to address advanced quality of service issues based on resource reservation, the paper describes an implementation of IPng (IP next generation) and RSVP (Resource Reservation Protocol) over ATM. Solutions for mapping quality of service and traffic parameters in an adequate way are presented. Moreover, the issue of address mapping from IPng onto ATM is discussed. Implementation results and experiences in these areas are illustrated. Finally, ongoing current work on resource reservation in advance is presented. It is outlined that longer-term resource planning and scheduling provides additional benefits for selected ATM applications.

I. Introduction

ATM components for local area networks [21] have become widely available as products. With the current UNI 3.1 (User Network Interface) specification and the emerging UNI 4.0 specification of the ATM forum [3, 4], interoperability can also be achieved on a broad basis. It now becomes more and more important to actually support applications on top of ATM with sufficient performance, and also in heterogeneous network environments. Although it is already possible to run applications directly over AAL5 (ATM Adaptation Layer 5), additional transport- and network-level protocols are required in heterogeneous settings, for example with Ethernet, FDDI, and ATM subnetworks being interconnected [1, 19] . The typical choice of many vendors is to offer the TCP/IP protocol suite [24] over ATM, based on the IP over ATM recommendation of the ATM Forum [18]. LAN Emulation ([15, 17]) presents another alternative, however with significant limitations. Most important, compatibility of existing applications with ATM is achieved by using IP over ATM.

In section 2 of this paper, we first present experiences and performance results for IP over ATM based on a local ATM network using DEC Gigaswitch/ATM. Comparisons with other standard networking technologies are also performed experimentally. It becomes obvious that current transport protocols are not ideally suited for ATM, and that tuning mechanisms and functional improvements are required.

Meanwhile, the IETF (Internet Engineering Task Force) has also specified a follow-on version of IP, the IP version 6 (or IPng - IP next generation) protocol [13, 14, 20]. The major goal is to enhance the IP address space due to the rapid growth of the Internet and to offer additional functionality. Moreover, the resource reservation protocol RSVP [5, 11, 27] has been developed. Such a protocol is crucial for guaranteeing quality of service (QoS) characteristics based on explicitly reserved network, memory and processing resources [6]. It is of particular importance in heterogeneous networks with partial ATM infrastructures.

Section 3 of the paper therefore addresses IPng and RSVP and reports concepts and first experiences with IPng and RSVP over ATM. In particular, the problem of mapping QoS and traffic parameters in an adequate way is discussed. This is of specific importance as the kind of QoS specification differs significantly between RSVP and ATM so that the mapping is non-trivial. Moreover, concepts for mapping IPng addresses onto ATM addresses are also presented, and relevant implementation-level aspects are discussed. Based on this work, early experiments with these new protocols have become possible; this way, the new functionality can readily be exploited by emerging ATM applications, especially in heterogeneous network environments.

Section 4 discusses ongoing research work on resource reservation in advance. We present new concepts for longer-term management of distributed resource requirements in ATM settings. The requirements and basic concepts of an adequate resource scheduling in conjunction with reservation protocols are discussed. This way, specific quality of service characteristics for important application scenarios can be guaranteed at a rather early stage.

Section 5 presents concluding remarks and summarizes the major findings.

According to the knowledge of the authors, only few results in these areas are found in the current literature so far. In [16], the transfer of data over ATM using available bit rate is examined. Memory management is discussed and advanced switching concepts are proposed. However, higher-level protocols are not investigated yet. [7] presents a higher-level investigation of application-level performance in ATM networks; however, like many other similar studies, it is only based on analytical and numerical models and does not include practical measurements. While many research efforts also concentrate on QoS specification and supervision [6, 25], the actual implementation of QoS-related reservation protocols especially over ATM has hardly been addressed yet, with a few exceptions such as the ST-II work described in [9] and a rough comparison of the IETF and ATM services models [8]. Although resource reservation in advance has been considered by several authors already [22, 23, 26], implementation concepts are still at a very early stage.

II. IP over ATM: Performance Evaluation

In this section, we present selected performance results of conventional TCP/IP over ATM and discuss our experiences. First, our experimental environment is briefly introduced.

II.A. Network Structure

Fig. 1 shows our network structure. Several multimedia workstations of type DECstation 3000 AXP 700 and 300 and a server are connected with our DEC Gigaswitch/ATM via fibre optic links. ATM access is implemented by adapter cards in the workstations and by line cards in the switch. The adapter cards perform cell generation from input packets of variable size and transmit the cells using SONET/SDH frames with standard 155 Mbit/s per

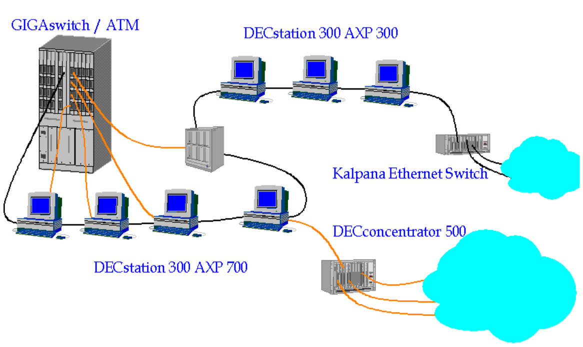

Fig. 1: Experimental environment: structural overview

channel via multimode fibre. Cell assembly and disassembly is done in hardware. Only AAL5 is currently implemented.

The switch itself offers a total performance of 10.4 Gigabit/s and is input-buffered. Possible head-of-line-blocking, a potential problem of input buffering, is reduced by a specific output port allocation algorithm (parallel iterative matching). Both PVCs (permanent virtual channels) and SVCs (switched virtual channels) are supported; signalling is based on ITU Q.93B with Q.2931 as the follow-on version. Moreover, CBR (constant bit rate), VBR (variable bit rate), and ABR (available bit rate), both with point-to-point and point-to-multipoint VCs, are basically possible. However, the driver and subsystem software currently does not support CBR yet, so that the experiments were mainly based on ABR.

The multimedia workstations are also connected via Ethernet and have access to another Ethernet switch, and also to an FDDI ring and in this way to the Internet via a concentrator. Each station is equipped with typical devices such as cameras, microphones, speakers etc., using MME (Multimedia Environment) as an internal software platform. Over all, the installation and maintenance of our environment did not create major problems, and existing applications could easily be ported to run within this infrastructure.

II.B. Results

Within our experiments, we evaluated the performance of TCP/IP over ATM. Fig. 2 shows the associated protocol structure. On top of the ATM hardware, a driver module provides the interface to the connection management module (CMM). The CMM handles connection establishment and coordinates the other modules. It

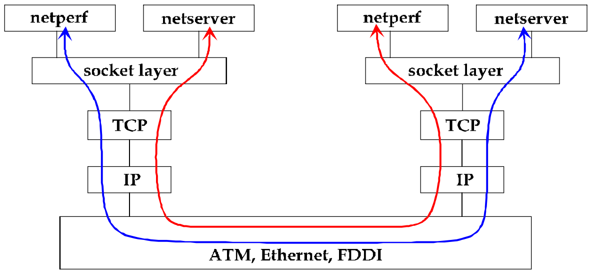

Fig. 2: Protocol structure: Overview

also has an interface to the signalling module (implementation of signalling protocol), to the ATM address resolution protocol (ATM ARP with dedicated ARP server) and to the IP convergence module. This module maps classical IP onto ATM; this includes access to the ATM ARP server, initiation of connection establishment to IP destinations, and transfer of data.

The application consists of a bidirectional client/server interaction. A component named "netserver" receives data while another component at each peer site named "netperf" sends data according to a load specification. Then on each site performance characteristics are evaluated. The components interact via a conventional socket layer offered by TCP. In alternative tests, Ethernet and FDDI were also used instead of ATM.

Fig. 3 shows a major summary diagram of the results. Both the message sizes transferred via TCP/IP and the buffer sizes at the receiver's site were varied; initial experiments have already shown a strong influence of both parameters. The major target parameter was the actual throughput that could be achieved. First, transmission was unidirectional only.

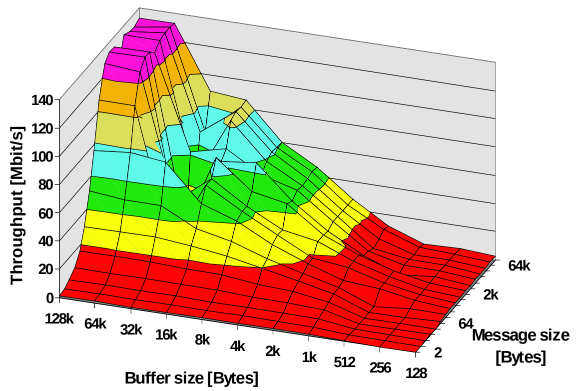

Fig. 3: TCP/IP over ATM: Throughput with varying buffer and message sizes

The maximum throughput achieved was 135 Mbit/s with optimal parameter values. Although this equals 87% of the physical bandwidth, it also is notable that the CPU load of more than 60% was significant then, caused both by the sender and receiver applications and by protocol processing. Nevertheless, the experiment has shown that the bandwidth of ATM can only be exploited based on adequate protocol parameter setting and sufficient CPU capacity. This means, that without tuning the protocol parameters, namely message sizes and buffer sizes, higher-level transport and application protocols present major performance bottlenecks in ATM applications. In any case, it becomes obvious that local protocol processing rather than physical communication causes the major performance limits with today´s hardware.

With a buffer size of less than 64 kbytes, performance dropped significantly, for example down to values between 60 and 90 Mbit/s for buffers of 32 kbytes. The higher values (90 Mbit/s) could only be achieved with significant message sizes. However, many applications often do not deliver messages that are large enough to achieve satisfying performance. For example, we also observed these problems with bulk data transfer via existing remote procedure call protocols over ATM. In these cases, the constant part of the overhead of protocol processing has increasing influence. Of course, especially very small messages resulted in very poor performance. Similarly, very small buffers lead to very poor performance, too, as buffer overflow resulted in loss of data and subsequent retransmissions.

Moreover, buffer sizes may be limited by the hardware, especially if multiple ATM applications coexist on a system. Therefore, it can be recommended to consider the possible use of ATM already during application development, for example for designing the data transfer phases and the mechanisms to be used. Automatic adaptation of protocol parameters according to the underlying network would also be a reasonable goal.

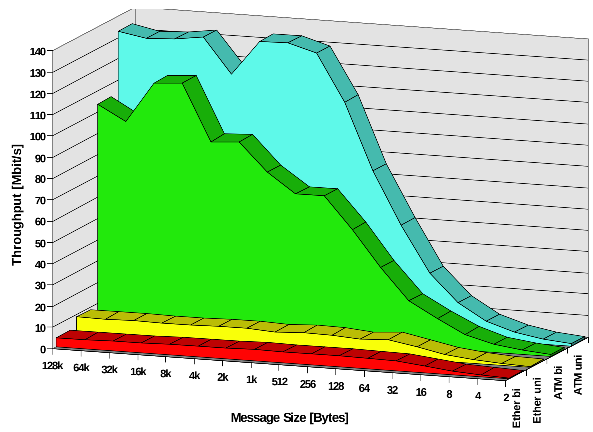

Fig. 4: Throughput comparison among heterogeneous networks

Fig. 4 shows the results of further experiments with bidirectional traffic and with Ethernet, compared with ATM, using buffer sizes of 128 kbyte. First, ATM performance drops to 110 Mbit/s and less per connection although a 155 Mbit/s VC is available for each direction. This is caused by significantly growing CPU load, because of the fact that both the sender and receiver applications and protocols have to be handled on each machine. The comparison with Ethernet shows that much lower throughput (a maximum of 8.5 Mbit/s) is achieved as expected, and that bidirectional traffic leads to even more significant reductions (down to 3 Mbit/s per connection) due to the shared medium with collisions. For FDDI, similar effects were observed, i.e. a more than 50% performance reduction per connection for bidirectional traffic. The major reason is that the FDDI bandwidth of 100 Mbit/s is to be divided among all communication partners while ATM VCs can be provided exclusively for each pair of stations, and also for each direction.

Under normal load conditions (with regular background load), we also observed that the quality of video transmission varied significantly due to the current network and local load conditions. This problem can only be addressed by offering CBR and VBR mechanisms with guaranteed bandwidth. However, this requires resource reservations in the end systems and in the active network components (i.e. switches, routers, bridges etc.). Important resources are bandwidth, memory or buffers, and CPU cycles.

For these reasons, we are currently working on the implementation of RSVP (resource reservation protocol) over ATM. Major problems and concepts are discussed below. This work is coupled with the implementation of IPng over ATM which is also described.

III. IPng and RSVP: Concepts and Implementation over ATM

The deployment of the new Internet Protocol on separate data link technologies is very important in view of a broad propagation concerning both already existent technologies such as Ethernet and new technologies such as ATM. This part of the paper describes an implemented adaptation of Internet Protocol next generation on ATM. Further, it introduces an approach to solve the problem of mapping QoS parameters required by applications onto ATM parameters to utilize the advantages of resource reservation in ATM. These considerations are especially based on the Resource ReServation Protocol (RSVP) as a future constituent of an Integrated Services Internet. The major characteristic of RSVP is an extended support of QoS for the Internet Protocol, whereby IPng is especially suitable supporting such reservation protocol. Therefore we compare the emerging integrated traffic services that are being developed by IETF and ATM Forum.

III.A. Major Concepts of IPng

IPng is a new version of the Internet Protocol which is assigned IP version number 6 and is formally called IPv6. Because IPng is an evolutionary step from the Internet Protocol (IPv4), it comprises on one hand numerous mechanisms of IP which have been expanded or kept in IPng. On the other hand, it contains new mechanisms and characteristics. The improvements of IPng in contrast to IPv4 fall primarily into the following categories:

Packet structure

The IPng header (see fig. 5) distinguishes from the IP header in such a way, that some elements have been reduced and other elements will be optional. The distinction of options into difference extension headers will reduce the bandwidth needed for the IPng header. Therefore, the IPng basic header size is only the twofold of the IP

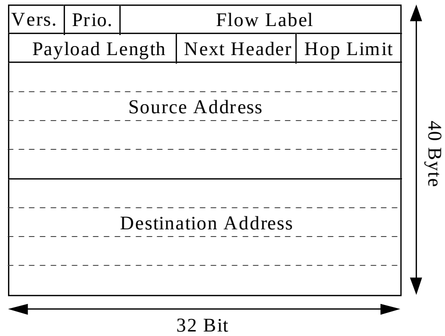

Fig. 5: IPng basic header

header even though IPng increases the IP address size from 32 to 128 bit. In addition, the time of packet processing in routers will be reduced because the extension header generally contains information concerning the endsystems. Changes in the way IPng header options are encoded allow for more efficient forwarding, less stringent limits on the length of options, and greater flexibility for considering future options.

Addressing

The extension of the IPng addresses provides more addressing hierarchy levels and a much greater number of addressable nodes. In conjunction with these aspects IPng offers new address formats. One of these is the Anycast Address identifying sets of nodes, whereby a packet sent to an Anycast Address is delivered to one of the nodes only.

Security and authentication

IPng includes additional options for the definition of extensions which provide support for authentication, data integrity, and confidentiality. These basic elements of IPng will be included in all its implementations.

Quality of service aspects

The Flow Label and the Priority fields in the IPng header may be used by a host to identify those packets for which a special handling by IPng routers is requested, such as non-default quality of service or "real-time" service. The characteristics of this special handling for the corresponding labelled packets belonging to the same flow may be conveyed by a resource reservation protocol or IPng options. This capability is important in order to support multimedia and realtime applications which require some degree of consistent throughput, delay, and/or jitter.

Transition mechanisms

To facilitate the migration from IP to IPng, IPng includes transition mechanisms, allowing an adoption and deployment of IPng in a highly diffuse fashion, and a direct interoperability between IP and IPng. Examples of such transition mechanisms are the dual IP layer, automatic and configured tunnelling, and the IP header translation as a special case of the communication of pure IPng with IP hosts.

In conjunction with the adaptation of IPng onto different link layers, IPng hosts need to resolve or determine the neighbour link layer address which is known to reside on attached links. The neighbour discovery protocol will be applied for that, using ICMP messages for information interchange. This is generally done via multicasting the addresses of neighbouring routers, the reachability of neighbour hosts, and some additional information. Moving the address resolution up to the ICMP/IP layer makes IPng more independent from the underlying link layer. However, although the neighbour discovery protocol was designed for the deployment of different link layers, it is mainly suited for broadcast media like Ethernet.

III.B. Integrating IPng and ATM

Recently, formats and methods were specified for the transmission of IPng packets over different networks like Ethernet, FDDI and Token Ring. In the following, we outline basic concepts for an initial implementation of IPng over ATM. As a general basis, the following parameters must be derivable from IPng parameters to allow support of applications via ATM networks:

- ATM address of destination,
- Quality of service and traffic parameters,
- Connection states and identifier of the ATM virtual channel.

The first three parameters represent information which are needed for the signalling protocol to establish a virtual connection. For making reservations for dedicated flows, it is required to determine the associated virtual channel identifier. Facilitating the assignment of IPng packets to ATM VCs, the IPng packet header contains the flow label, where packets with flow label zero are sent as best effort data.

To determine the ATM destination address according to the neighbour discovery protocol, the multicast capabilities of ATM must be exploited. That is, before sending multicast messages for neighbour solicitation, an address translation of a multicast IPng address into a corresponding ATM address has to be performed. To reduce the overhead which will occur in conjunction with using a central Multicast Address Resolution Server (MARS) [2], the IETF currently discusses several solutions. Currently, address resolution for our implementation of IPng over ATM is realized based on the principle of classical IP over ATM.

The specification of QoS and traffic parameters for dedicated flows has to be performed by the application. However, with the assistance of RSVP, application-level QoS and traffic parameters can be mapped onto ATM parameters explicitly as discussed below.

III.C. RSVP - Basic Concepts

Proposed as an internet draft, RSVP is a known constituent of the Integrated Services Internet. It provides especially real-time applications with guaranteed, predictable and controlled end-to-end performance across networks. It is a receiver-oriented protocol, which may be classified as a control protocol of the Internet Protocol. Therefore, it offers a flexible handling of heterogeneous receivers as well as an adaptation to dynamically changing multicast groups. The main task performed by RSVP is signaling of resource requirements at connection setup time. To be precise, RSVP does not actually reserve or allocate resources but rather indicates reservation requests to the underlying systems.

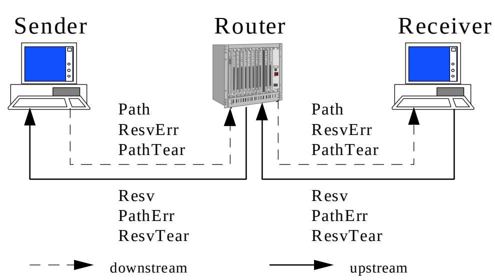

Fig. 6: Exchange of RSVP messages

The transmission of RSVP control information is implemented by encapsulating RSVP packets into IP or IPng packets. Reservations may be performed for both unicast and multicast connections. The basis of a reservation is a detailed description of the flow traffic and the QoS characteristics. In accordance with this fact, RSVP defines the so-called flow and filter specification. The flow specification specifies the traffic (TSpec) using token bucket parameters for describing bursty traffic, and also determines the required QoS parameters (RSpec). The filter specification contains an identification of the flow for which reservations have been performed. Even though the

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**