Experiments on Integrated Traffic Control for ATM in the ACTS Project EXPERT

MARTIN LORANG, THOMAS RENGER
University of Stuttgart, Inst. of Communication Networks and Computer Engineering, Pfaffenwaldring 47, 70569 Stuttgart, Germany

HARALD PETTERSEN, EGIL AARSTAD
Telenor Research & Development, P.O. Box 83, 2007 Kjeller, Norway

LAURENT JAUSSI
Swiss Federal Institute of Technology (EPFL), Telecom Lab, 1015 Lausanne, Switzerland

MARTIN POTTS
Association Swiss PTT/ASCOM (ASPA), Morgenstrasse 129, 3018 Bern, Switzerland

Abstract. In this paper first an overview of the ACTS project EXPERT is given by briefly describing the major project goals and the related technical work items, the sophisticated testbed infrastructure and the various collaborations EXPERT has with other R&D projects. Focus is then given to an important area of work within EXPERT, namely the definition and investigation of a traffic control architecture which allows different ATM service categories to share common resources and thereby to integrate them into one ATM network. For switches which employ delay priorities among service categories, a CAC method is presented which takes into account the specific requirements of each category as well as the underlying buffer architecture. This concept is validated for rt-VBR and nrt-VBR type of traffic. Extensive experiments and measurements have been made to investigate the performance of TCP over UBR with and without packet discard schemes dependent on many parameters of the involved protocol layers.

1 Introduction

EXPERT is one of more than 150 projects in the European Union 4th Framework Programme ACTS: "Advanced Communications Technologies and Services". It operates probably the largest broadband test environment in Europe. One objective of EXPERT is to provide a well-supported broadband platform for its own experimentation and trials, and for use by other projects within and outside the ACTS Programme. This platform also serves as an ACTS showcase for visitors with varied interests in ATM aspects and in EU Telecommunications R&D Programmes in general.

In addition, EXPERT is investigating several technical areas, including broadband signalling, service-related control, resource management, user Quality of Service (QoS), traffic control in a mixed traffic environment, and the development of an ATM network performance evaluation tool.

EXPERT provides a unique broadband (ATM) test environment comprising varied and sophisticated equipment both in Basel, Switzerland, and Leidschendam, The Netherlands. In association with these sites, advanced application trials are being performed to validate and demonstrate the capabilities and advantages of new broadband services and to show the benefits of the underlying ATM technology. The testbeds in Basel and Leidschendam currently comprise 13 (commercial and prototype) ATM switches, various terminals and test equipment, broadband signalling, network interworking units and broadband applications. The introduction of the Available Bit Rate (ABR) service is being prepared.

Due to the comprehensive facilities at its disposal, EXPERT has a major role within the ACTS Programme for the identification and validation of a framework of traffic control functions which enable an ATM network to support the partly conflicting QoS demands of the five service categories (CBR, rt-VBR, nrt-VBR, ABR, UBR) as defined by the ATM Forum [1]. These functions should operate efficiently for each individual category. However, in order to enable an ATM network to simultaneously support traffic from different service categories contending for common resources, the set of control functions has to be integrated into a robust and consistent framework. Partners in

EXPERT are not only working on the theoretical definition of such a framework, but are also implementing, validating and comparing the proposed solutions through extensive experiments and trials. It is especially important that solutions for guaranteeing that user-requested QoS for all types of applications can be demonstrated, before network operators can offer commercial services.

2 The EXPERT Project

2.1 Work Items

The work being undertaken in the EXPERT project covers a broad scope of ATM-oriented broadband communication activities, centred on the comprehensive, supported test environment available to the partners in the project, and to collaborating users or projects. The platform is being continually extended and upgraded with new equipment, functionalities and applications, through developments within the project, co-operation agreements with manufacturers, or purchase.

Developments in the project include an integrated services switch (for attaching both terminals with ATM-25 interfaces and N-ISDN equipment) and a performance evaluation tool for the design and planning of ATM networks. Additional items provided by partners in the project include a new ring-based gigabit campus network, and an ATM Passive Optical Network (APON) with associated VB5.1 interface functionalities. Partners are also implementing the latest ATM Forum and ITU-T signalling capabilities, and dynamic routing and resource management algorithms.

However, and of particular concern for this paper, a substantial proportion of the effort in the project is concentrated on the definition, implementation and validation of an integrated traffic control framework to enable an ATM network to support a broad range of QoS requirements. This work continues from earlier individual investigations of Connection Admission Control (CAC) and Usage Parameter Control (UPC), whilst incorporating aspects related with the newly specified ABR and Unspecified Bit Rate (UBR) services. The recent availability of long-distance (intercontinental) ATM links adds a further dimension to these topics.

2.2 Infrastructure

The EXPERT platform is illustrated in Fig. 1 below. The collection of broadband facilities offered by EXPERT (applications, terminal equipment, 13 different ATM switches, signalling capabilities and network interworking units) is unique in Europe. Terminal types include PCs, SUN workstations, TV/video equipment and N-ISDN terminals, all capable of being attached to the ATM network using

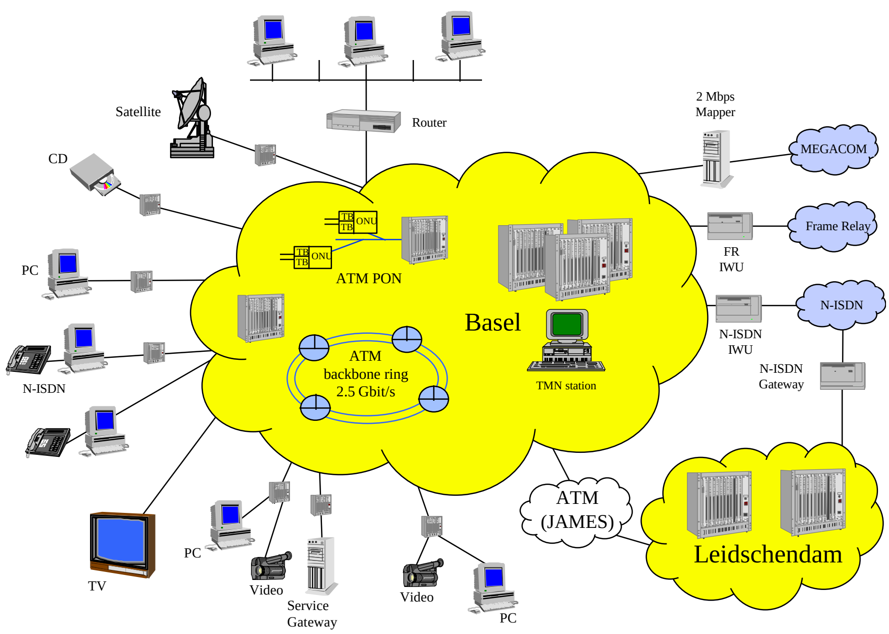

Fig. 1: The EXPERT platform

terminal adapters. Most of the ATM switches have different internal architectures, but have sufficient common external interfaces to allow them to be interconnected as required for particular trials. Applications include CSCW (Computer-Supported Collaborative Working), standard TV/video, 3-D video, CD-audio and multimedia conferencing. Network interworking is provided to/from the Frame Relay and N-ISDN networks. Permanent access is available to the European ATM network (operated by the EU project "JAMES"). Via this network, part of Deutsche Telekom Berkom´s commercial network and the trans-atlantic fibre "CANTAT-3" (operated by Teleglobe), similar testbed sites in Canada can also be reached.

2.3 Relationship to other Projects

In consideration of the facilities available to the project, EXPERT is committed to supporting other projects both within and outside the ACTS Programme. For example, in the field of UMTS (in a local hospital environment) EXPERT is providing the location for the ATM/UMTS interworking equipment. EXPERT will also collaborate with the Finnish Multimedia Programme for trialling VoD, and is a platform for ATM accounting and signalling experiments, tele-education trials and TMN (availability) aspects. The EXPERT project also has very close associations with the National Hosts of Switzerland and The Netherlands (direct physical connections), Denmark, Norway, Finland and the UK. Due to its reputation, EXPERT is also involved with collaborations from outside of the ACTS community. A particularly interesting co-operation is currently taking place with the Canadian BADLAB Testbed (from the CANARIE Programme) whereby in the framework of the G7/GIBN (Global Interoperability of Broadband Networks) Programme regular "Virtual Classroom" sessions are held between schools in Ottawa and Basel.

3 Integrated Traffic Control for ATM

ATM Service Categories (ASCs) have been defined in [1] to offer optimal support for the divergent QoS requirements of existing and future broadband applications (such as voice, video and data) and to allow for a high bandwidth utilisation. Individual ASCs are designed such that their properties complement each other with respect to criteria such as time-relation (real-time vs. non-real-time), QoS model (guaranteed vs. best effort) and control methodology (preventive vs. reactive). Currently there exist five ASCs: CBR, rt/nrt-VBR, ABR and UBR. Appropriate traffic control functions have to be associated with each ASC enabling it to offer a cost-effective service, i.e. ensure negotiated QoS with minimal network resources effort. The diversity in the properties of individual ASCs is reflected in the type and design of the required control functions. The major goal of integrated traffic control is to integrate all ASC-related control functions into a common traffic management framework which allows to share network resources efficiently among connections belonging to different ASCs.

The set of control functions together with the underlying switch buffer management is referred to as the integrated traffic control architecture. Among others, the following main criteria to assess and compare different control architectures should be considered: QoS protection, efficient use of network resources, robust operation of the control system and implementation complexity.

The basis for ASC integration is the underlying buffer management. The traditional FIFO service strategy is not applicable in integrated services scenarios and some kind of priority scheme (e.g., delay priorities, round robin or weighted fair queueing based) is mandatory. For an overview of suitable service strategies for ATM it is referred to [2, 3]. Furthermore, the buffer capacity is usually partitioned into a number of logical queues (LQ) making it possible to assign each connection (per-VC), group of connections (per-connection-group) or each ASC (per-ASC) to a separate LQ. The service sequence between LQs is controlled by the service discipline. Certainly there is a trade-off between the partitioning principle applied and the related implementation effort. While per-VC queueing offers the highest control flexibility and QoS protection, it requires to set-up and manage for each connection (VC) a separate LQ which may become unfeasible in public network scenarios where hundreds or even thousands of connections share the same link.

CAC plays a key role within the control framework since it directly controls the network load by deciding if enough resources are available to accept a new connection. Since the resource allocation for a new connection is strongly dependent on the requested ASC, it is not possible to apply a common CAC algorithm for every ASC. Hence, for each ASC a suitable CAC method has to be identified and it has to be ensured that these ASC-specific functions can co-operate in the resource allocation (bandwidth and buffer capacity). The CAC framework is closely related to the buffer management implemented in the switch. In the next section a CAC concept based on a buffer architecture employing delay priorities between ASCs will be presented.

UPC is needed to enforce the negotiated traffic parameters. The Peak Cell Rate (PCR) and additionally the Sustainable Cell Rate (SCR) with Burst Tolerance (BT) are controlled for CBR and rt/nrt-VBR connections, respectively. ABR requires a dynamic UPC to control the varying value of the allowed cell rate while UBR in principle does not need any policing.

Other important control functions are traffic shaping and packet discard schemes (e.g., Early Packet Discard (EPD), Partial Packet Discard (PPD)). While the former helps to generate a traffic pattern which allows the network to allocate less resources and, therefore, to offer a cheaper service, the latter reduces overload and avoids waste of bandwidth by discarding cells which belong to corrupted AAL-packets. In section 5 the impacts of EPD and PPD as part of the UBR control framework on the performance of TCP will be shown and discussed.

In EXPERT different solutions for integrated traffic control architectures have been defined and investigated. So far, one proposal based on delay priorities between LQs for different ASCs and a rate-based approach with explicit reservation of bandwidth and buffer capacity have been identified [4]. While the first one requires only one LQ for each considered ASC, the latter is more flexible since it can employ any grouping principle from per-VC to per-ASC queueing.

Investigations on the control mechanisms for the ABR service category and their inclusion into the integrated traffic control architecture are also performed in EXPERT [4, 5, 6]. Further studies on the interactions between the control functions and higher layer protocols are currently being extended to include ABR.

4 CAC in a Delay Priority System

Experiments have been performed within the EXPERT testbed to identify a CAC method appropriate for systems applying delay priorities between rt-VBR and nrt-VBR traffic [4, 5]. All experiments so far have been based on a Fore ASX200 switch implementing delay priority queues. The system has been loaded using synthesised traffic generators which are capable of generating artificial test traffic according to a Markov model [7]. The buffer configuration is shown in Fig. 2. The service discipline is such that the low priority queue (non-real-time traffic) is served only when the high priority queue (real-time traffic) is empty.

Through measurements a CAC boundary has been determined that can be compared with CAC boundaries calculated using different CAC algorithms. N_{rt} on/off sources are multiplexed into the rt-VBR queue and N_{nrt} on/off sources into the nrt-VBR queue. All sources have a peak bit rate of 31.1 Mbit/s and exponentially distributed on- and off-periods with a mean length equal to 20 ms and 80 ms, respectively. The rt-VBR buffer was set to 48 cells, the nrt-VBR buffer to 4000 cells.

Fig. 3 shows the measured CAC boundary giving the maximum number of rt-VBR and nrt-VBR sources that can be multiplexed with a target cell loss ratio equal to 10^{-4}. The boundaries corresponding to a convolution based CAC (bufferless model) and a linear CAC (effective bandwidth) are displayed together with the measured results. The measurements reveal the increased utilisation that may be achieved using a priority scheme between the rt-VBR and nrt-VBR classes, thereby allowing to utilise larger buffers without affecting the delay performance for the real-time traffic.

The convolution based CAC gives low utilisation since a bufferless model cannot benefit from the large buffers available. The shaded area in Fig. 3 shows the increased gain that could be obtained using a linear CAC based on the effective bandwidth for each class. The effective bandwidth for the rt-VBR traffic is determined using the rt-buffer capacity, while the calculation of the effective bandwidth for the nrt-VBR traffic is based on the nrt-buffer capacity.

However, as can be seen from Fig. 3, a significantly better link utilisation can be achieved by limiting the CAC acceptance region with a boundary (denoted ‘RT boundary’) given by rt-VBR CAC convolution calculations together with a boundary (denoted ‘NRT boundary’) obtained by computing effective bandwidths as if all connections are of the nrt-VBR type and could use the large nrt-buffer. By the latter it is meant to allocate resources according to:

$$N_{rt} \cdot EB_{rt} + N_{nrt} \cdot EB_{nrt}$$

where EB_{rt} is the effective bandwidth for the rt-VBR and EB_{nrt} is the corresponding effective bandwidth for the nrt-VBR traffic, both calculated using the nrt-buffer capacity. Fig. 3 shows that this combined approach gives an acceptance region close to the one determined by the measurements.

The following overall CAC strategy may be applied when ASCs are separated by delay priorities:

- CBR and rt-VBR: Convolution based CAC calculation allowing statistical multiplexing gain
- nrt-VBR: Allocation of capacity to nrt-VBR connections based on effective rates including rt-VBR connections in the calculation as if they are nrt-VBR connections
- ABR: Linear allocation of capacity to individual ABR connections based on Minimum Cell Rate (MCR) taking into account the effective rates used for the nrt-VBR allocation
- UBR: No allocation of capacity to individual UBR connections

To ensure the robustness of this CAC solution, it is required that the calculated effective rates define a conservative CAC boundary. One way to obtain the effective bandwidths could be to estimate them off-line using a homogeneous fluid flow model. This approach should be compared to a CAC scheme based on worst-case traffic assumptions. For the proposal of such a scheme and for more information about the use of fluid flow models for architectures based on delay priorities it is referred to [5].

5 Performance of TCP over UBR

The UBR service is intended for nrt-applications that do not require guaranteed QoS commitments. With additional, relatively inexpensive control functions such as packet discard schemes, UBR could

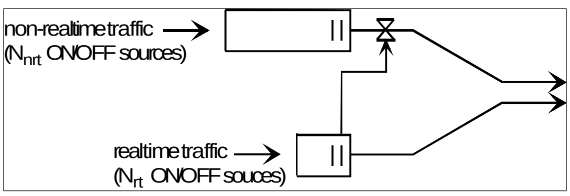

Fig. 2: Delay priority buffer configuration

Fig. 3: CAC boundaries for rt-VBR and nrt-VBR traffic

become a cost-effective alternative for the transmission of data traffic, offering a flexible solution as opposed to nrt-VBR or ABR with its sophisticated and complex rate-control protocol. Therefore, it is essential to evaluate the TCP over UBR performance, to identify how UBR can be integrated efficiently with other ASCs, and, in a final step, to compare its performance with ABR. This section surveys some key results obtained from a comprehensive set of experiments with TCP over UBR comprising measurements taken on different protocol layers.

5.1 Configuration for the TCP over UBR Experiments

Four workstations and the Fore switch build the local ATM network based on 100 Mbit/s TAXI interfaces as depicted in Fig. 4. The workstations are equipped with Fore SBA200 adapters which encapsulate IP packets in AAL5. The user application accesses the TCP/IP protocol stack and finally the ATM network interface through the BSD based socket API. The benchmarking tool *ttcp* measures the TCP memory-to-memory throughput. During the measurement, the source *ttcp* application is backlogged at any time instant. By means of VP shaping, the bottleneck rate is set to 37.44 Mbit/s, which is a rate both pairs of workstations can exploit completely.

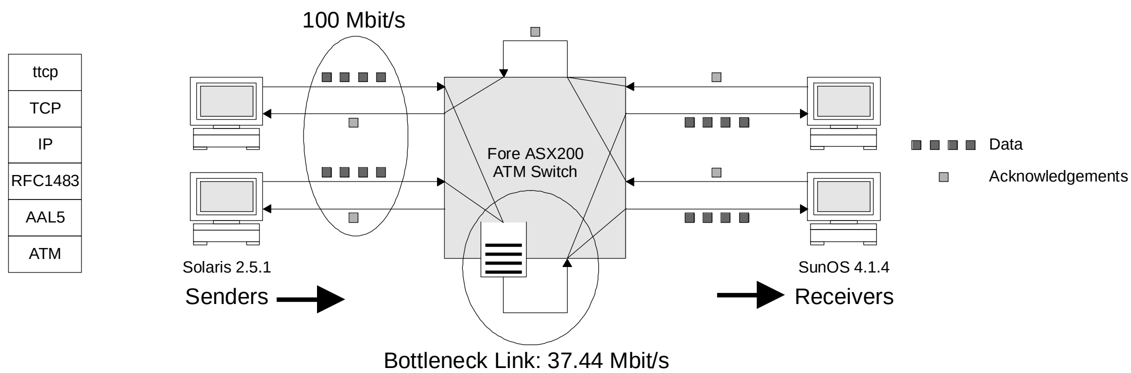

Fig. 4: Configuration for the TCP over UBR experiments

5.2 TCP over Plain UBR

TCP provides end-to-end window based flow control by means of several congestion control algorithms such as Slow Start, Congestion Avoidance, Fast Retransmit and Recovery. All these algorithms assume that packet loss indicates congestion in the network which has to be resolved by an appropriate rate decrease. Packet loss is either detected by a retransmission timeout or duplicate acknowledgements. Waiting for timeouts or duplicate acknowledgements can cause TCP to stop the transmission, and therefore, the switch buffer for UBR has to be dimensioned appropriately to ensure that packet loss can be kept within reasonable bounds.

TCP achieves the maximum throughput when no packets are lost, which means that the buffer at the multiplexing port has to be equal or larger than the sum of the receiver windows of all TCP connections involved. In this case, the TCP connections share the available bandwidth almost perfectly, as can be seen in Fig. 5 and 6 with a switch buffer size of 2176 cells. These figures also show goodput, Cell and Frame Loss Ratio (CLR and FLR) for a switch buffer size between one and two TCP windows and for different Message Transfer Units (MTU).

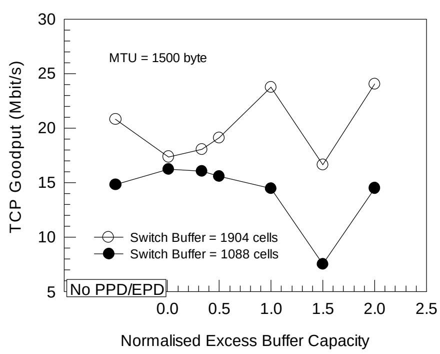

Fig. 7: TCP goodput as a function of the normalised excess capacity b

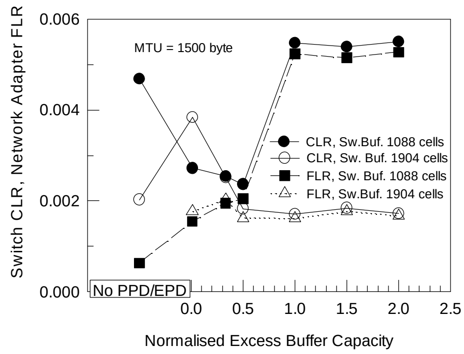

Fig. 8: FLR and CLR as a function of the normalised excess capacity b

As expected, the goodput increases and the CLR decreases with increasing buffer size up to the point where the buffer is large enough to prevent any cell loss. As opposed to simulation results [8], the goodput grows with a larger MTU. This is caused by the relatively high processing overhead for the transmission of smaller packets. This processing overhead compensates the fact that the loss of a single cell has a more negative impact on the goodput for larger packets. Fig. 6 shows as well that a larger switch buffer is required to ensure zero loss for smaller packets. This is because the effective maximum TCP window size depends on MTU [5]. As expected, the FLR is significantly larger than the CLR because the loss of a single cell inside a frame inevitably leads to the discard of the whole AAL5 frame at the destination. Only if all cells were dropped at the switch, the CLR would equal the FLR.

5.3 TCP over UBR with EPD

In its implementation of EPD, the Fore switch defines an EPD threshold (*EPDTh*) for the shared switch buffer. If the fill level of the shared buffer exceeds this threshold when the first cell of an AAL5 frame arrives, all incoming cells of this frame will be discarded except for the last one which allows to delimit the frames. If a cell in the middle of a frame is lost, the later incoming cells are discarded as well (Partial Packet Discard, PPD). Further variants of EPD are described in [8, 9].

In the following, the normalised excess capacity *b* (defined by the relation $EPDTh = SwitchBuffer - b \cdot MTU$) is applied to compare the performance of EPD for different threshold values with a fixed switch buffer. PPD is realised by setting *b* to zero.

Fig. 7 and 8 show the total goodput, the CLR and FLR for two TCP sessions, different excess capacities and without EPD for comparison. With small buffers (1088 cells), the best goodput is achieved with PPD and not with EPD, as one might expect. For a larger buffer (1904 cells), EPD improves the goodput best when the excess capacity is an integer multiple of MTU, since the excess capacity then allows to store an entire frame thus avoiding that frames are transmitted partially. As can be seen, the CLR and FLR differs for plain UBR and PPD whereas they become approximately equal when the normalised excess capacity is larger than 0.5. With a larger excess capacity, the cell loss process is driven more and more by the EPD mechanism which discards a whole frame. For

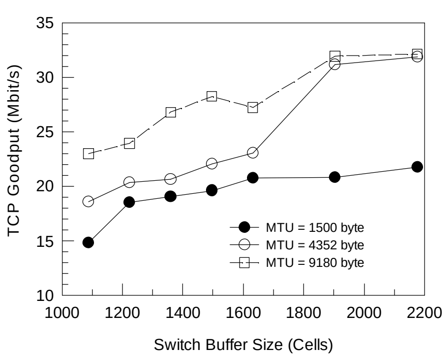

Fig. 5: TCP goodput as a function of the buffer size

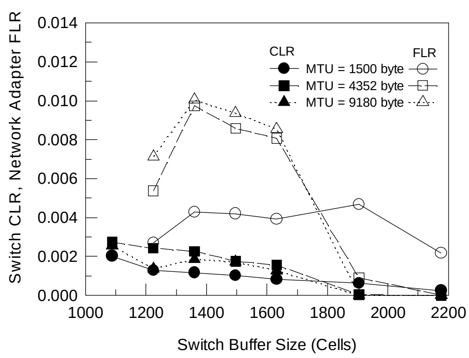

Fig. 6: FLR and CLR as a function of the buffer size

larger MTU values [5], EPD usually does not improve or even reduces the goodput achieved by PPD. The buffer space allocated as the excess capacity is seen rather as a buffer reduction than as a means to avoid the transmission of corrupted frames. For a more detailed discussion refer to [5]. The experiments have shown that the dimensioning of the EPD threshold and the excess capacity are separate issues and that the most important issue for TCP over UBR is the dimensioning of the threshold, i.e. the buffer which can be used without any restriction.

5.4 The Impact of Delayed Acknowledgements

The default SunOS TCP refrains from sending acknowledgements until the receiver window has slid more than two maximum packets or 35 % of the maximum TCP window (delayed acknowledgement, DACK). The comparison of TCP with and without DACK both for PPD and EPD in Fig. 9 and 10 reveals that TCP without DACK suffers from relatively low goodput and a significantly lower fairness index (definition taken from [10]). Obviously, TCP without DACK allows a dominant session to conserve its resource share in the network. Conversely, DACK inserts randomness to the transmission of packets. Since the ACK for every second packet is deferred, DACK prevents the synchronous sequence: „packet-ACK-new packet". If the buffer space exceeds the size of the TCP window, a connection starting first may occupy buffer corresponding to its window size thus ensuring that its TCP congestion control works very efficiently compensating the low goodput of the other connection. TCP without DACK tends to maintain synchronous, deterministic periods. Acceptable overall goodput despite unfairness is possible, as can be seen for a switch buffer of 1360 cells. Without DACK, TCP's congestion control works inefficiently when the buffer is increased further to 1632 cells and the second connection comes into play. The higher fairness for EPD suggests that EPD adds randomness to the system similar to DACK.

6 Conclusions

The broadband test environment of the EXPERT project with its extensive collaborations with other R&D projects has been described focusing on the important area of ATM traffic control. Effective control functions are required to enable an ATM network to efficiently support different ATM service categories. These functions have to co-operate efficiently with the buffer management used in the switches to achieve a robust integrated control system able to satisfy the divergent QoS requirements.

Architectures based on delay priorities between different service categories and a rate-based approach with explicit reservation of bandwidth and buffer capacity either per connection or per group of connections have been considered. A CAC strategy for delay priority architectures has been developed and shown to achieve increased resource utilisation without affecting the real-time delay performance.

Packet discard schemes proposed to enhance the UBR performance have been investigated in extensive TCP over UBR experiments. Larger goodput is observed for larger values of MTU when using plain UBR. In the configuration used, the TCP sessions benefit only little from PPD or EPD. This is contrary to what has been shown in several simulation studies and demonstrates the

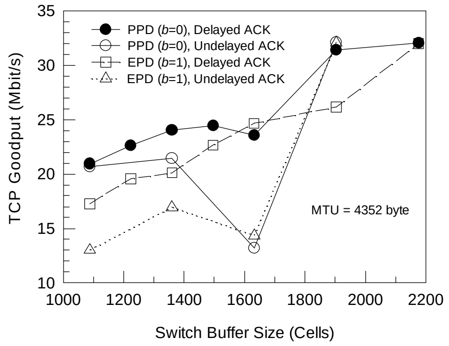

Fig. 9: TCP goodput for delayed/undelayed ACK

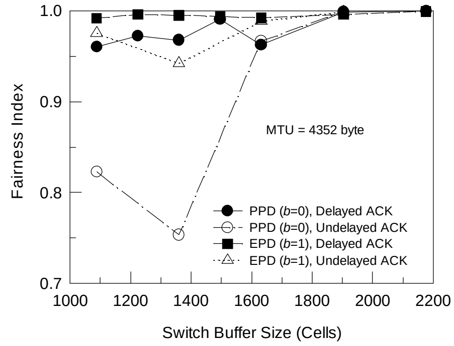

Fig. 10: TCP fairness for delayed/undelayed ACK

importance of experiments with real systems. Future work will focus on the integration of the ABR service into the traffic control framework and the related experimental activities.

Acknowledgements: The achievements being made within the project EXPERT and the comprehensive development work being undertaken is only possible with the conscientious co-operation of all partners especially of WorkPackage 4.2, together with the support of the Project Officer and funding from the Commission of the European Union, the Swiss Bundesamt für Bildung und Wissenschaft and other national bodies. The contribution of all these players is therefore hereby gratefully acknowledged.

References

[1] ATM Forum: *Traffic Management Specification Version 4.0*. AF-TM-0056, ATM Forum Technical Committee, April 96.

[2] EXPERT WP4.2: *Initial Specification of Bandwidth Management (taking into account particularly routeing strategies and the use of ABR)*. ACTS 094 Deliverable 4, March 1996.

[3] H. Zhang: *Service Disciplines for Guaranteed Performance Service in Packet Switching Networks*. Proceedings of the IEEE, Vol. 83, N-10, Oct. 1995.

[4] EXPERT WP4.2: *Specification of Integrated Traffic Control Architecture*. ACTS 094 Deliverable 6, Sept. 1996.

[5] EXPERT WP4.2: *First Results from Trials of Optimised Traffic Control Features*. ACTS 094 Deliverable 10, March 1997.

[6] C. Blondia, O. Casals: *Throughput Analysis of the Explicit Rate Congestion Control Mechanism in ATM Networks.* Proceedings of the 10th ITC Specialist Seminar, Lund, September 1996.

[7] B. Helvik, O. Melteig, L. Morland: *The Synthesized Traffic Generator: Objectives, Design and Capabilities*. Proceedings of the IBCN&S, Copenhagen, April 20-23, 1993.

[8] A. Romanov, S. Floyd: *Dynamics of TCP over ATM Networks*. IEEE JSAC, vol. 13, No 4, pp.633-641, May 1995.

[9] H. Li et al: *A Simulation Study of TCP Performance in ATM Networks with ABR and UBR Services in ATM LANs*. IEICE Trans. Comm., Vol. E79-B, No 5, pp.658-667, May 1996.

[10] R. Goyal et al.: *Further results on UBR+: Effect of Fast Retransmit and Recovery*. ATM Forum Contribution 96-1761, Dec. 96.