Industry
Canada
CRC

TRANSPORT FLOW CONTROL AND CONNECTION ADMISSION POLICIES FOR RELIABLE APPLICATIONS

by

Louise Lamont, Ilka Miloucheva
and Wolfram Stering

CRC REPORT NO. 96-007

April 1996
Ottawa

Industry Canada Industrie Canada

The work described in this document was sponsored by the Department of National Defence under Task 5CB11.

Canada

TRANSPORT FLOW CONTROL AND CONNECTION ADMISSION POLICIES FOR RELIABLE APPLICATIONS

by

**Louise Lamont, Ilka Miloucheva
and Wolfram Stering**

The work described in this document was sponsored by the Department of National Defence under Task 5CB11.

COMMUNICATIONS RESEARCH CENTRE, INDUSTRY CANADA
CRC REPORT NO. 96-007

Canada

April 1996
Ottawa

Abstract

A collaboration between the Communication Research Centre (Canada) and DeTeBerkom (Germany) was undertaken to determine transport protocol performance characteristics over high-speed trans-Atlantic ATM connections, using national High-Speed Test Networks and Teleglobe's trans-Atlantic submarine fibre CANTAT-3.

The measurements focus on TCP/IP's flow control parameters and mechanisms for Quality of Service (QoS) provision (throughput, response time) to bulk data and transaction applications. Practical issues of TCP/IP over Long Fat Networks (LFN) (trans-Atlantic ATM) are investigated and compared with TCP/IP performance on ATM Local Area Networks (LANs).

In order to study the QoS management of TCP/IP connections over ATM, we investigated the specific effects of performance factors such as: TCP flow control and send window size, network delay, system scheduling and application traffic.

In this document we discuss the mapping of applications' QoS parameters to the ATM service classes and policies for connection admissison control. We also present the performance effects of bundling or grouping transport connections over the same ATM resources. Practical considerations for mapping the traffic and QoS requirements of applications to ATM resources are discussed.

Executive Summary

A collaboration between the Communication Research Centre (Canada) and DeTeBerkom (Germany) was undertaken to study transport protocol issues for the management of QoS on long-fat networks (trans-Atlantic ATM). The focus of this work was to show how flow control and admission control mechanisms, based on TCP, affect the performance of reliable applications on ATM WAN and LAN.

A series of performance measurement tests were carried out, between Berlin and Ottawa, during the January and February 1996 time frame. Local ATM measurements were performed on the BALI network during this same time frame.

For the purpose of this joint Canada-Germany research activity, national ATM networks were interconnected via the CANTAT-3 trans-Atlantic submarine fibre cable. Use of the CANTAT-3 submarine fibre cable was provided by Teleglobe Canada. Access to CANTAT-3 via national ATM Test Network facilities was supported by the Canarie TNOC in Canada and by Deutsche Telekom and DeTeBerkom in Germany. From Pennant Point, Nova Scotia, the CANTAT-3 submarine cable lands at Sylt Germany. Permanent Virtual Circuit (PVC) links of 10Mb/s were set up between the CRC BADlab in Ottawa and DeTeBerkom in Berlin.

The research team also had access, in Berlin, to a local ATM testbed (BALI). A series of performance measurements were run in parallel. The BALI measurements provided a valuable baseline for the trans-Atlantic measurements and allowed an ATM WAN - ATM LAN performance comparison. For the LAN measurements, PVC connections of 155 Mb/sec where established in both directions.

The experiments focused on two categories of "reliable" applications: bulk applications where the delay bound is not stringent (these applications are not time sensitive on the millisecond scale but are sensitive to minimum throughput guarantee); and transaction applications which require timely delivery and response.

Some of the measurements were taken using the ttcp measurement tool and its enhancements xtcp which allowed the throughput for bulk data applications to be measured. The Swedish Institute of Science developed a benchmark tool called SPIMS [31], which allows specific QoS parameters for transaction applications (throughput, response time) to be measured.

The intent was to explore how the ATM networks can be used more efficiently. We wanted to determine if this could be achieved by reserving resources not only for individual connections but for groups of connections while maintaining the traffic and QoS requirement limitations of each connection. We demonstrated that transport connections can be grouped or bundled to improve resource reservation, QoS provisioning and cost. We also demonstrated that when connections are bundled at the end-system, the traffic of the connections as well as the network delay become important performance factors. Using an adaptive scheme, i.e. by monitoring the throughput of a given bundle over time, an appropriate set of requirements for an ATM virtual circuit (for example peak and average rate) can be obtained.

We also wanted to show that transport flow control strategies must appropriately map the QoS and traffic requirements of the applications to the ATM services in order for users to use the ATM networks efficiently and to afford the emerging ATM highways. We demonstrated that application traffic characteristics, system scheduling, the rate and window sizes and congestion control mechanisms of transport protocols are all factors that impact the performance.

With the recent deployment of ATM switches in real network infrastructures, more and more studies are now dealing with practical issues of TCP/IP, particularly on ATM LANs. We provided a first glance at TCP/IP's performance over high-speed trans-Atlantic connections using Teleglobe's CANTAT-3 submarine fibre cable. More experiments could be undertaken under a different project to determine if TCP's performance can be improved on ATM WANs. It would be interesting to determine if TCP can adapt better to ATM WANS by disabling the Slowstart and Congestion Avoidance mechanism. Reservation protocols such as RSVP could be used to allow more efficient flow control policies for connections which have to provide a certain QoS. Practical usage of the RSVP implementation for bundling connections could also be a future research topic.

Table of Contents

List of Figures

Acknowledgements

The Research Team gratefully acknowledges the support of the Canarie Test Network Operations Committee, Teleglobe Canada, Deutsche Telekom and DeTeBerkom in providing the use of CANTAT-3 and the access infrastructure.

The following individuals provided valuable assistance to the experiments and we thank them for their contributions:
M. Savoie & L.Chang (CRC/BADLab),
A. LeChasseur & Y. Poppe (Teleglobe Canada)
D. Hetzer (DeTeBerkom)
O.Bonnesz & Klaus Rebensburg (Technical University Berlin).

We would like to acknowledge the support provided by J.Robinson of CRC (Canada) and N. G. Ulamec & U. Hofmann of Techno-Z Research Austria for this project activity.

This work has been funded by the Department of National Defence, Canada.

1 Introduction

ATM technology, with its high speed and wide band capabilities, promises efficient performance for both local-area and wide-area communications by providing: greater scalability [1]; integration of multiple services and formation of virtual workgroups [2]; specification of QoS requirements for voice, video and data applications [3]; flexible service classes and adaptation layers according to specific user demands [4]; and reduced communications costs [5]. However, without solid transport flow and admission control strategies which appropriately map the QoS and traffic requirements of the applications to ATM services, there is a real risk that many users are not going to be able to afford the emerging ATM services for their applications.

The Canada-Germany Project "Support for Multimedia Protocols on International High-Speed Connections" was undertaken to study transport protocol issues for the management of QoS on long-fat networks (trans-Atlantic ATM). Use of the CANTAT-3 submarine fibre cable was kindly provided by Teleglobe Canada. Access to CANTAT-3 via national ATM Test Network facilities was kindly supported by the Canarie TNOC in Canada and by Deutsche Telekom and DeTeBerkom in Germany. The research team also had access, in Berlin, to a local ATM testbed (BALI). A series of performance measurements were run in parallel. The BALI measurements provide a valuable baseline for the trans-Atlantic measurements and allow an ATM WAN - ATM LAN performance comparison. The focus has been on two categories of "reliable" applications: bulk applications where the delay bound is not stringent (these applications are not time sensitive on the millisecond scale but are sensitive to minimum throughput guarantee); and transaction applications which require timely delivery and response. The Project has investigated performance issues of transport protocol ARQ and connection admission policies for TCP [22], [23] and XTPX [24], [25], [26].

The research has focused on the following:

- Measuring performance characteristics of reliable applications over ATM WAN and LAN while considering application traffic and QoS requirements, system scheduling, mapping of transport flow control mechanisms and ATM services;
- Defining efficient admission control policies for bundling, i.e. multiplexing, of transport connections and their mapping of ATM connections with well defined QoS parameters.

The main transport protocol flow scenarios that are addressed include:

- Impact of transport protocol parameters (window size and rate) and user application-layer framing (TSDU size) on the throughput of bulk data connections while considering the system configuration and scheduling;
- Impact of the flow and congestion control mechanisms (specifically TCP slow start) on the QoS provision for transactions.

To measure performance we have used the application-oriented measurement tool SPIMS developed at the Swedish Institute of Science [31]. The combination of this application-oriented measurement with ATM layer QoS measurements, based on protocol analyzer equipment, is addressed.

This report is organized as follows:

Section 2 gives a theoretical background and standardization framework for the ATM service architecture covering QoS and traffic management, and mapping of application classes to ATM services.

Section 3 describes the network configurations (trans-Atlantic ATM and the Berlin local ATM) that have been the test beds for the performance measurements. The characteristics of the ATM switches used in these environments are presented along with a more general discussion of ATM switch architectures.

Section 4 discusses application-level and ATM-level QoS measurement techniques and describes the tools used for this measurement program.

Section 5 covers transport-level flow control and Section 6 covers application admission control on ATM networks. This provides technical background on the issues that have been investigated in the measurements. The results of the measurement program are given in Sections 7 and 8. The measurements are presented in greater detail in the Appendix.

The XTPX study is presented in Section 9. Section 10 provides a summary of the results.

2 ATM Service Architecture

ATM has been chosen as the switching and multiplexing technique for B-ISDN. The ATM standard is designed to support diverse traffic characteristics from high speed applications along with multimedia applications.

The ATM technology can provide connection-oriented broadband services in local, metropolitan and wide-area network environment with different levels of QoS guarantees. ATM applications can specify QoS parameters such as allowable delay and error rates that determine the characteristics of a connection.

In this section, we give a brief overview of the ATM service architecture. This includes: the protocol reference model, QoS support, traffic management and service categories and how they are applied to the different application classes.

2.1 ATM Protocol Reference Model

ATM is based on a virtual-circuit-oriented fixed-sized packet (cell) switching and multiplexing technology. ATM breaks all traffic into cells where the size has been selected as a trade-off between requirements of the telephone companies (small cell size to reduce the delay for voice packet) and the data communication community (big cell to minimize the amount of segmentation and reassembly).

A cell consists of a 5-byte header information field and a 48-byte user data field. The header field contains control information for the cell (such as connection identification, cell loss priority, routing and switching).

The ATM Protocol Reference Model is based on the ITU-T standards. It is divided into three layers: physical layer, ATM cell layer and ATM adaptation layer (figure 1).

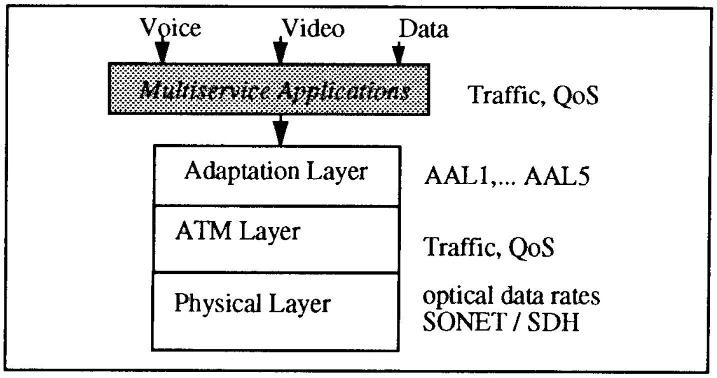


Figure 1 - ATM Protocol reference model

The ATM adaptation layer provides multiservice networking by transforming traffic from its native node, such as packets or CBR bitstreams, into ATM cells for common transport through the cell relay network. Therefore traditional data and voice information as well as video and image all have a way of incorporating their traffic into cells and hence mapping their application QoS to ATM QoS.

Since cells are associated with a particular connection, the behaviour of that connection can be tightly controlled. A specific QoS can be assigned to each connection. The ATM layer is responsible for maintaining ATM connections with assigned QoS parameters such as bandwidth, delay and cell loss. Signalling is a control procedure defined by the ATM Forum UNI / NNI to establish and release ATM connections. The VPI (Virtual Path Identifier) and VCI (Virtual Channel Identifier) are used to identify which cells belong to a particular connection with required QoS and traffic parameters. Connections which are established dynamically are called SVC (Switched Virtual Connections), in contrast to the PVC (Permanent Virtual Connections) which are set up on a permanent basis.

ATM signalling is defined in the ATM FORUM UNI Specifications 3.1 [52] and 4.0 [53]. Some significant differences between ATM UNI 3.1 and 4.0 are:

- UNI 3.1 does not allow the characteristics of a VC to be changed once it is established. UNI 4.0 allows for dynamic renegotiation of VC characteristics;
- The only approach available in UNI 3.1 for multicasting is the root-initiated point-to-multipoint ATM connection. UNI 4.0 supports Leaf Initiated Join (LIJ) to a point-to-multipoint ATM connection. This allows a leaf to directly request to join a point-to-multipoint connection without notifying or involving the root of the connection. This eliminates the scalability problem.

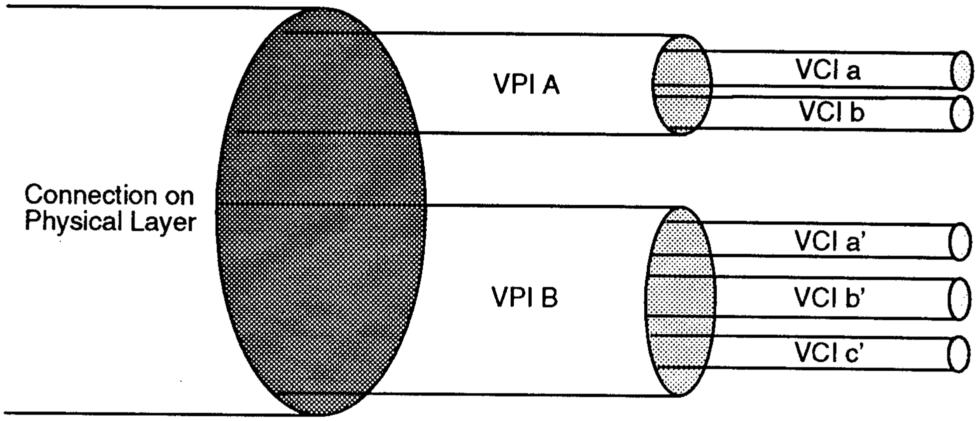


Figure 2 *- ATM Connections: Virtual Paths and Virtual Circuits*

Two end systems can use a VPI to multiplex or bundle many individual application streams together using the VCI to distinguish between the streams. The network does not interpret or modify the VCI fields of cells on VPI connections. VPIs are used to establish virtual paths on a semi-permanent basis between network endpoints while VCIs are used to establish virtual links over a given VPI.

The physical layer encodes and decodes the data into suitable electrical or optical waveforms on the communication medium used. It has a medium-dependent sublayer responsible for the correct transmission and reception of bits on the physical medium and a transmission-convergence sublayer responsible for mapping the ATM cells to the transmission medium used. The optical data rates, synchronization and framing format chosen for B-ISDN are called Synchronous Digital Hierarchy (SDH) in Europe and Synchronous Optical NETwork (SONET) in North America (see Appendix).

2.2 ATM Layer Quality of Service

Specific QoS parameters are negotiated between the networks and the end systems for each direction of the ATM connection. The network agrees to meet or exceed the negotiated QoS as long as the end system complies with the negotiated traffic contract (ITU-T Recommendation I.371). Such commitments are probabilistic in nature and are intended to be only a first-order approximation to the performance that the network expects to offer over the duration of the connection. Since there is no limit on the duration of a connection and the network can only make decisions based on the information available at the time the connection is established, the actual QoS may vary over time. In particular, transient events can cause short-term performance observations to be worse that the agreed QoS commitment. Thus, QoS commitments can only be evaluated over a long period of time and over multiple connections with similar performance commitments.

The ATM Forum Traffic Management Specification 4.0 [6] distinguishes between negotiable (CDV, MaxCDT, MeanCDT, CLR) and non-negotiable (CER, SECBR, CMR) QoS parameters. These are shown in Table 1.

ATM QoS Parameters	Description
Peak-to-peak cell delay variation (CDV)	difference between the best and worst case expectation of CDT, where the best case is equal to the fixed delay and the worst case is equal to a value likely to be exceeded with probability not greater than α
Maximum Cell Transfer Delay (Max CTD)	$1 - \alpha$ quantile of CTD (Cell Transfer Delay), i.e. elapsed time between a cell exit at measurement point 1 and the corresponding cell entry at measurement point 2 for a particular connection
Mean Cell Transfer Delay (Mean CTD)	arithmetic average taken over a sample of the cell population
Cell Loss Ratio (CLR)	defined for a connection as the ratio of the number of lost cells to the total transmitted cells
Cell Error Ratio (CER)	ratio of errored cells to the successfully transferred cells + errored cells
Severely Errored Cell Block Ratio (SECBR)	ratio of severely errored cell blocks to the total transmitted cell blocks
Cell Misinsertion Rate (CMR)	ratio of the misinserted cells to the time interval

Table 1 - ATM QoS Parameters

Figure 3 shows the delay oriented QoS parameters:

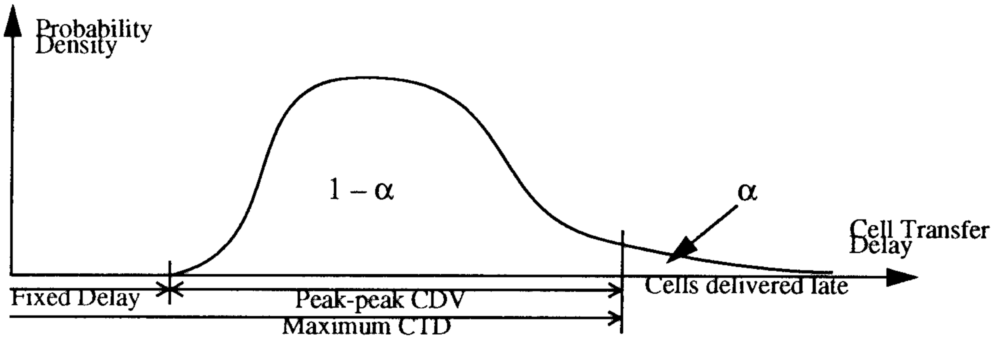


Figure 3 - Cell Transfer Delay Probability Model (for real-time service categories)

2.3 ATM Traffic Specifications

Traffic parameters describe the traffic characteristics of the connection. The traffic parameters defined in [6] are presented in the following table:

ATM Traffic Parameters	Description
Peak Cell Rate (PCR)	In cells/sec gives the maximum time interval between consecutive cells
Sustainable Cell Rate (SCR)	In cells/sec determines the average rate of the cell
Maximum Burst Size (MBS)	In cells specifies maximum allowable number of consecutive cells at PCR
Minimum Cell Rate (MCR)	In cells/sec gives the minimum time interval between consecutive cells

Table 2 - ATM Traffic Parameters

When controlling the PCR, the ATM network allows for some fluctuations in the inter-cell arrival times through the specification of a Cell Delay Variation Tolerance (CDTV)

.

The ATM traffic descriptor is the generic list of traffic parameters that can be used to capture the traffic characteristics of an ATM connection. A source traffic descriptor is a subset of the traffic parameters belonging to the ATM traffic descriptor.

It is used by a requesting source during connection establishment to capture the intrinsic traffic characteristics of the connection.

A traffic contract specifies the negotiated characteristics of the connection. The traffic contract at Public UNI consists of the connection traffic descriptor and a set of QoS parameters for each direction of the ATM layer connection. The Private UNI can optionally support the same or a different traffic contract as the Public UNI. The connection Traffic descriptor consists of the source traffic descriptor (i.e. PCR, SCR, MBS, MCR), the CDTV, and the conformance definition.

2.4 ATM Service Classes

After having looked at the QoS requirements for a comprehensive list of applications, the ATM Forum constructed a set of service categories by identifying application properties and network functionality.

The ATM service categories relate traffic characteristics to application classes and QoS requirements to network behaviour. Functions such as routing, Connection Admission Control (CAC), and resource allocation are structured specifically for each service category. Service categories are distinguished as being real-time or non real-time. Real-time categories are distinguished by the degree to which minimum spacing of cells is expected and preserved by the network.

Correspondence of application classes based on their traffic characteristics and QoS requirements to the ATM services is found in the recent ATM Forum report on Traffic Management Specification [6] and in the ATM service Architecture: From Application to Scheduling document [7]. Characterization of service categories with appropriate QoS and traffic descriptions, as well as their possible use for different application classes is given in the following table:

ATM Service Category	Specified QoS and Traffic	Application
Constant Bit Rate CBR	**PCR and CDVT** **peak-to-peak CDV** **maximum CTD** **CLR**	**- real-time applications like voice and video requiring tightly constrained delay and delay variations supports consistent availability of a fixed quantity of bandwidth** Interactive video (video conferencing) Interactive audio (telephone) video distribution (television, distributed classroom) audio distribution (radio, audio feed) video retrieval (video on demand) audio retrieval (audio library) Any data/text/image transfer application which contains smooth enough traffic or for which the end systems response time requirements justify occupying a fully reserved CBR channel
real-time Variable Bit Rate rt-VBR	**PCR and CDVT** **SCR, MBS, CDVT** **peak-to-peak CDV** **Maximum CTD** **CLR**	**- real-time applications like voice and video requiring tightly constrained delay and delay variations** **- supports rate varying with time and "bursty" sources** **- may support statistical multiplexing of real-time sources** real-time application for which the end system can benefit from statistical multiplexing by sending at a variable rate and can tolerate or recover from a small but non-zero random loss ratio real-time application for which variable rate transmission allows more efficient use of network resources

Table 3 - ATM Layer Service Categories

ATM Service Category	Specified QoS and Traffic	Application
non real-time Variable Bit Rate nrt-VBR	**PCR and CDVT** **SCR, MBS, CDVT** **Mean CTD:** - For all connections **Maximum CTD** **CLR:** - For those cells which are transferred within the traffic contract, the application expects a low CLR	**- non real-time applications which have bursty characteristics** **- may support statistical multiplexing of connections** response-time-critical transaction processing (i.e. airline reservation, banking transaction, process monitoring, frame relay interworking)
Unspecified Bit Rate UBR	**PCR and CDVT:** - does not specify traffic related service guarantee and does not include the notion of a per connection negotiated bandwidth. - It is network specific whether or not the UBR traffic parameters are subject to CAC and/or UPC. The network may choose a policy for accepting a connection regardless of PCR. In this case PCR can be used to indicate the smallest bandwidth limitation along the path of connection.	**- non real-time applications such as file transfer and email. sources are expected to be bursty** **- supports a high degree of statistical multiplexing among sources** interactive text/data/image transfer (banking transactions, credit card verification) text/data/image messaging (email, telex, fax) text/data/image distribution (news, satellite picture) text/data/image retrieval (file transfer, browsing) aggregate LAN (LAN interconnection or emulation) remote terminal (telecomputing, telnet)
Available Bit Rate ABR	**PCR and CDVT** **MCR** **CLR:** - bandwidth from the network may vary but shall not become less than MCR (which can be specified as zero) **Feedback:** - a flow control mechanisms which supports several types of feedback to control the source rate in response to changing ATM layer characteristics. - It is expected that an end-system that adapts its traffic in accordance with the feedback will experience a low CLR and obtain a fair share of the available bandwidth according to a network specific allocation policy.	**non real-time traffic** any UBR application for which the end-system requires critical guaranteed QoS data transfer (i.e. defence information) super-computer applications data communication applications requiring better delay behaviour

Table 3 - ATM Layer Service Categories

Mapping of application classes to ATM services depends on the user requirements for costs, efficient QoS support, adaptation layers and network utilization. Pragmatic considerations have been studied for MPEG-2 over ATM Adaptation Layer in the framework of Video Dial Tone Networks [3] and for efficient voice mapping to CBR or VBR services [4].

Two application traffic patterns for residential broadband services are characterized as being either periodic or bursty. If packet generation occurs at regular time intervals, it is a periodic traffic pattern. If these packet lengths are fixed in size, it is a CBR stream, otherwise it is a VBR stream. For example, compressed video streams (such as MPEG-2) are periodic traffic and can be CBR or VBR.

Bursty traffic is characterized by packets of arbitrary length generated at random times, separated by gaps of silence of random duration. Typically, the period of silence is long compared with the packet-generation period, leading to the distinctive high peak-to-average bandwidth ratios. Image browsing is a bursty application because images are transferred on demand by the user. Since this is a real-time application, delay performance guarantees are required.

For bursty data applications, allocating "fixed" bandwidth in advance might not be the best way to utilize the network resources. If fixed bandwidth is allocated, this may cause either underutilization of the allocated capacity, or limitation in the throughput of the other traffic by the allocated capacity. An alternative strategy would be to send the traffic on a best-effort basis, i.e. using UBR and ABR services. For ABR, fair allocation of bandwidth requires applications to adjust their transmission rates according to feedback from the network. ABR explicitly supports a rate-control-based feedback mechanism. This flow-control feedback loop alerts sending sessions that the network is congested, i.e the ATM network sends information to the user specifying the bit rate at which the user should be transmitting.This allows the source to adapt (i.e. increase or decrease) its transfer rate to a time-varying bandwidth. ABR service is more reliable than UBR, but it requires more detailed specification of QoS parameters (cell loss ratio and minimum cell rate).

2.5 Interaction of Internet and ATM services

The emergence of ATM as a key networking technology has triggered researchers to study the mapping of the widely used connectionless Internet services to the connection-oriented ATM architecture.

Multiprotocol encapsulation over ATM Adaptation Layer 5 with LLC / SNAP and a per-VC multiplexing option is addressed in RFC 1483 [45]. The ATM address resolution server defined in RFC 1577 [42] specifies translation of IP addresses to ATM address schemes. RFC 1626 [46] discusses issues of packet fragmentation for AAL 5 where a large IP MTU size (9 Kb) for efficient IP packet handling and a MTU discovery protocol is proposed.

The ongoing efforts of Internet Engineering Task Force (IETF) and Working Groups like IP over ATM, COLIP, Routing Over Large Clouds Working Group (ROLC) are focused on models for mapping of IP flows / routes to ATM connections on the network layer.

A large variety of IP over ATM models are discussed in "IP over ATM: A framework document" [47].

Currently, the practical ATM WAN / LAN realisations are based on the Classical IP over ATM model[43]. It defines local ATM networks (spanning one Internet address) with their IP routing and address resolutions, i.e. an ATM network environment configured as a Logical IP Subnetwork (LIS).

Signalling and negotiations of parameters such as MTU size and methods of encapsulation of IP protocols are discussed in RFC 1755 [48]. "ATM Signalling Support for IP over ATM" [48] specifies how to directly connect local nodes via ATM VCI/VPIs and how to connect to remote nodes via routers.

As the distance to the next hop router / terminal gets longer in the case of ATM WAN, the source router needs more capacity to hold incoming packets. Therefore, further models are proposed such as the WAN ATM Subnet model based on the Next Hop Resolution Protocol (NHRP) [69] defined by the ROLC IETF WG. It specifies a scalable address resolution mechanism for a non-broadcast multiaccess (NBMA) network [68].

The Conventional model [67] defined by the COLIP WG eliminates the overhead of hop-by-hop IP processing of the Classical Model by enabling the routers to splice directly at the ATM level to the VCs that are associated with the same IP flow, thus forming a "bypass pipe".

In the Peer-to-Peer or Integrated model, the interactions of the IP and ATM routing are greatly simplified, but this requires the use of a common address format and similar metrics [47].

Support for IP / ATM multicast routing based on a set of multicast group membership servers (MARS) providing multicast address resolution within ATM networks is discussed in [72].

IPnG (IPv6), and implementations in many organisations, are considered in a new framework Internet Draft "A Framework for IPv6 Over ATM"[49], [50]. Many of the problems encountered in IPv4 over ATM, must also be dealt with when running IPv6 over ATM. Among these functions are neighbour discovery and address configuration. It is desirable to logically partition a large ATM network into IPv6 subnets while maintaining IPv6 subnet semantics so that any two nodes on the ATM network can still be provided with ATM QoS services within the framework of IPv6 networking.

3 ATM Network Configurations

ATM switching provides scalability of network configurations in many dimensions: scalability in station and network attachment, in the number of nodes and in the total bandwidth. In addition, it offers the possibility for building virtual workgroups within WAN and LAN infrastructures, for building information superhighways using intercontinental WAN ATM infrastructures connecting high speed LANs and WANs. Depending on the particular application requirements (traffic, bandwidth, ATM service demand, network connectivity and requested number of VCI/VPI), the appropriate switching architecture is selected.

This section describes briefly the Regional and National Test Networks that were used in this research experiment. The BALI network was used to provide LAN measurements and is therefore also described. The main characteristics of the switches included in the Transatlantic WAN and in the BALI LAN are presented. Background Information on ATM LAN and WAN characteristics and on the classifications of ATM switches is provided in Section 3.4 and 3.5

3.1 Transatlantic ATM WAN

The CANARIE National Test Network (NTN) was built in collaboration with Stentor and Unitel, Canada's two major carriers. It connects over 60 switches in 7 regional test networks spanning over 6000 kilometres. The NTN ATM backbone consists of Newbridges's 36150 MainStreet switches that provide ATM switching capacity of up to 2.4 Gb/s.

One of the regional test network is OCRINET, the Ottawa-Carleton Regional network. Regular network operations started in January 23, 1994. OCRINET connects technology companies, telephone companies, government laboratories and educational institutions. A dozen initial nodes, including CRC, are linked using T3-speed lines. CANARIE connects across the Atlantic to Berlin via Teleglobe's CANTAT-3 submarine fibre cable.

In Canada the CANTAT-3 cable system lands at Pennant Point, Nova Scotia. CANTAT-3 is the first to offer Synchronous Digital Hierarchy (SDH)/Synchronous Optical NETwork (SONET) transmission technology. SDH is an international standard that was initiated by CCITT whereas SONET is the North American standard that was formulated by the Exchange Carriers Standards Association (ECSA) committees for ANSI and adopted by CCITT. An SDH terminal multiplexes low-speed signals into higher rates. CANTAT-3 provides high-capacity communication at 2.5 Gb/s per fibre. CANTAT-3 is a three fibre pair system. The system design capacity is 96 DS-3 which is equivalent to 16 Synchronous Transport Modules (STM-1). One STM-1, the lowest defined transmission level in SDH, has a capacity of 155.22 Mb/s. The length of the cable is just over 7,500 km. It includes 89 repeaters and 4 non-regenerative branching units.The repeaters or regenerators are spaced at intervals of about 87 km across the system.

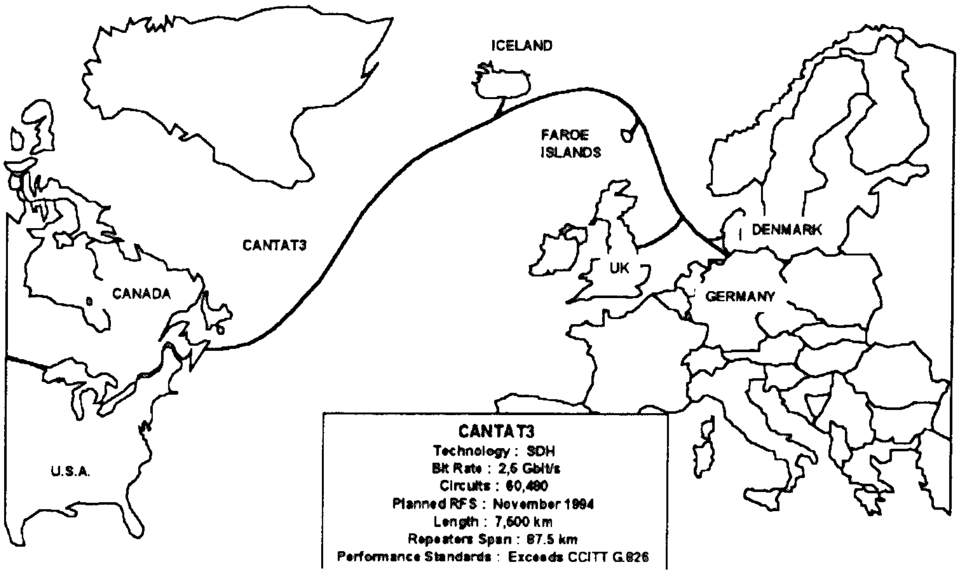


Figure 4 - CANTAT-3 Submarine Cable System

Permanent Virtual Circuit (PVC) links were set up between the CRC BADLAB in Ottawa and DeTeBerkom in Berlin. In the CRC BADLAB, a Newbridge VIVID ATM Workgroup switch is used for customer premises applications. The VIVID switch supports 12 OC3 ports and has a switching capacity of 1.6 Gb/s. It attaches to a 36150 MainStreet Access Switch. The CRC Ocrinet node is linked to Kanata's 36150 MainStreet switch which in turn links to Bank's 36150 switch. The OCRINET regional ATM network is linked to the Cisco Lightstream 100 switch in Ottawa. The Cisco Lightstream 100 switch operates at up to 2.4 Gb/s and supports from 1 to 16 ATM lines of 155 Mb/s data speed. The Cisco switch attaches to Belmont's 36150 in Quebec. A connection from Belmont's 36150 terminates at Teleglobe's facilities in Montreal. From Pennant Point, Nova Scotia, the CANTAT-3 submarine cable lands at Sylt Germany.

Figure 5 shows the ATM Transatlantic WAN link from CRC / Canada to DeTeBerkom / Berlin that was used in this research experiment for performance evaluation. Our performance scenarios are aimed at analysing the appropriate flow and admission control policies using the intercontinental ATM WAN connecting Canada and Germany.

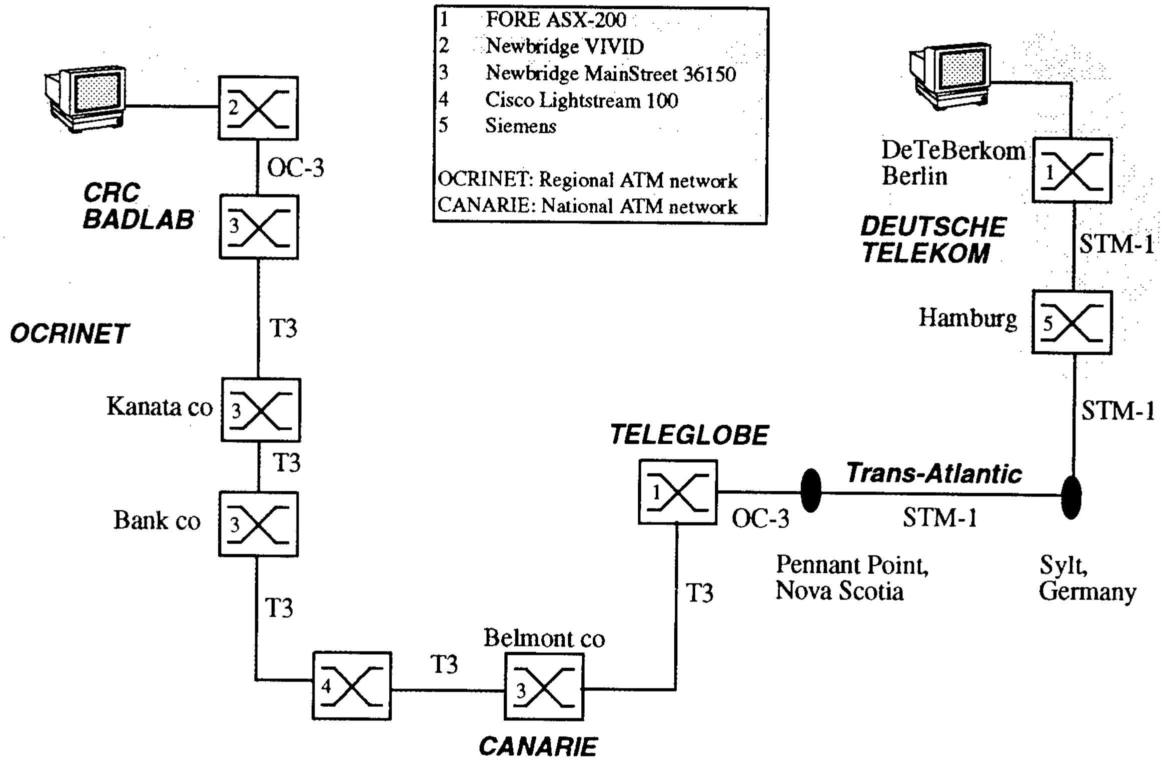


Figure 5 - Links from CRC/Ottawa to DeTeBerkom/Berlin

3.2 BALI - Berlin ATM LAN Initiative

The LAN ATM trial used in this performance study is based on BALI [36]. It provides practical results on transport protocol features and parameters used in the LAN environment and practical comparison between the LAN and WAN ATM environment.

The Berlin ATM LAN Initiative (BALI) is a pilot network established to gather technical and operational experience with the LAN ATM switching technology and to make the development of vendor-independent multimedia teleservices and applications possible. It was established by a joint project started in 1993, by DeTeBerkom, a subsidiary company of Deutsche Bundespost, by GMD-FOKUS, a research institute, and by Technical University of Berlin represented by the department FSP-PV and TUBKOM.

Three local ATM networks are connected to a joint network by means of high-speed data lines such as the BERKOM research network or leased lines. The networks are connected via 140 Mb/s PDH and 155 Mb/s SDH links. Each LAN ATM network consists of two or three switches from different providers (Fore, Netcomm, HP, Siemens, Synopsis).

Connection to conventional LANs (Ethernet, FDDI, local X.25) is done either via routers with ATM interfaces or via LAN interfaces within the switches. A number of general-purpose workstations and specific ATM terminals in each ATM LAN uses the ATM interface for direct ATM connection with the ATM switches.

Figure 6 shows the BALI configuration and the workstations and networking equipment at TUB and DeTeBerkom:

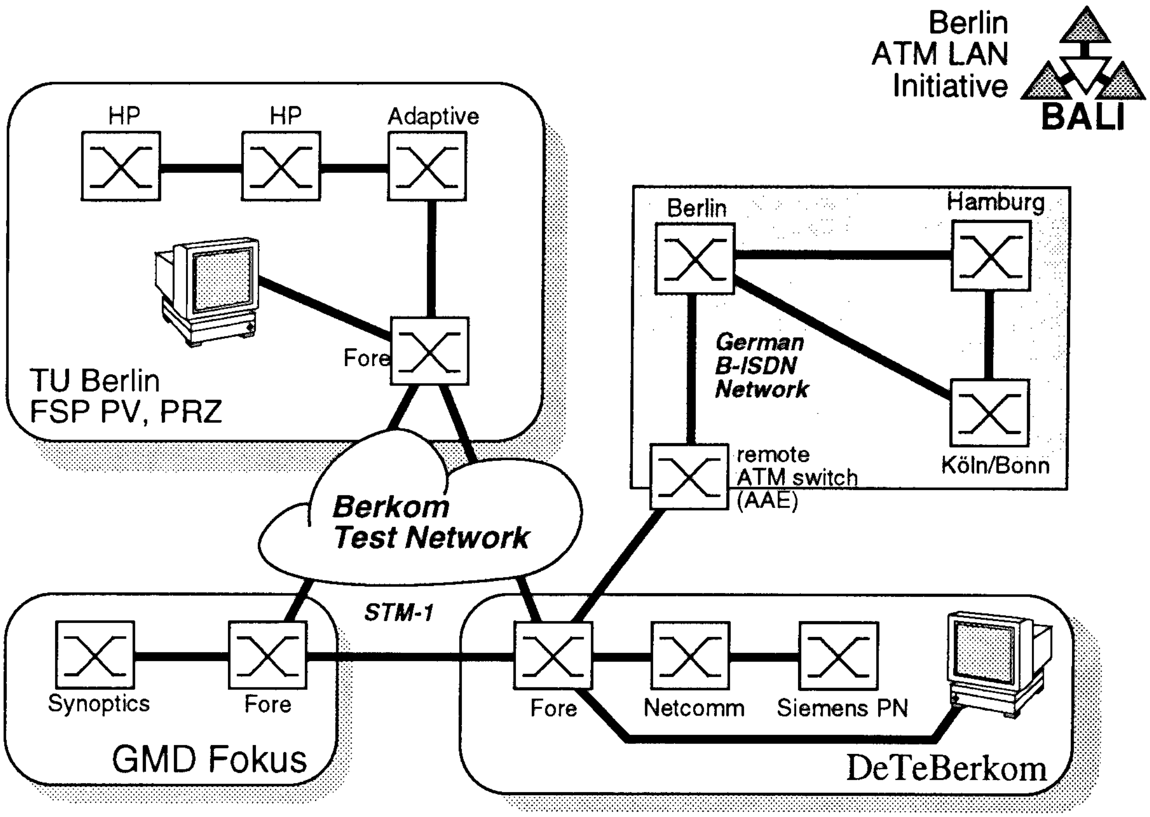


Figure 6 - Berlin Local ATM Infrastructure

3.3 ATM Switch Characteristics

The following table shows the main characteristics of the switches included in the Transatlantic WAN and in the BALI LAN.

Vendor, Model	Equip. Type	ATM Connectivity	No. of Ports	ATM Service	AAL Type	IP Encaps	Connection Types
Newbridge Networks, Inc. VIVID ATM Workgroup Switch [34]	Switch	Sonet OC-3	12	CBR, VBR, ABR, UBR[a]	N/A	—	point-to-point, point-to-multipoint
Newbridge Networks, Inc. 36150 MainStreet [35]	Switch	T3, E3, TAXI, Sonet OC-3, SDH STM-1	4, 6, 8, 16	CBR, VBR	N/A	—	point-to-point, point-to-multipoint
Cisco, Lightstream 100	Switch	T3, E3, TAXI, Sonet OC-3 SDH STM-1	1 to 16	CBR, VBR, ABR, UBR	N/A	—	point-to-point, point-to-multipoint

Table 4 - ATM Switch Characteristics

Vendor, Model	Equip. Type	ATM Connectivity	No. of Ports	ATM Service	AAL Type	IP Encaps	Connection Types
FORE Systems, Inc., ASX-200	Switch	T3, E3, TAXI, Sonet OC3, SDH STM-1	2 to 24	CBR, VBR	N/A	—	point-to-point, point-to-multipoint
FORE Systems, Inc., SBA-200	NIC	via TAXI			AAL 3/4 AAL 5	NULL, LLC	

AAL=ATM Adaptation Layer
LLC=Logical Link Control
SDH=Synchronous Digital Hierarchy
Sonet=Synchronous Optical Network
TAXI=Transparent Asynchronous Transmitter/Receiver Interface
NIC=Network Interface Card
N/A=Not applicable

Table 4 - ATM Switch Characteristics

a. Configurable , but not implemented in current version

3.4 ATM WAN and LAN

The distinction between LAN, MAN and WAN is a matter of geographic dispersion. Geographic dispersion affects the performance because of the increased propagation delay. From a pragmatic and applications point of view several contrasts can be drawn between ATM LANs and WANs. Most notably, an ATM LAN consists of a small number of ATM switches and only a few ports per switch. An ATM WAN on the other hand consists most likely of a large number of ATM packet switches with many ports per switch. Characteristics and comparison of LAN and WAN ATM are given in table 5[14]:

ATM LAN	ATM WAN
Small cheap switches (10-256 ports)	Large expensive switches (>1000 ports)
Need not be ultra-reliable and fully redundant	Reliability and redundancy are a must
Traffic policing is unnecessary, as the traffic sources are under local control	Traffic policing is required
Transmission latency is not a major issue	Transmission latency is a major issue
Can have many slower links (e.g. 155 or 622 Mbit/s); every link need not operate at gigabit speeds.	Must have gigabit links to handle aggregate traffic
The aggregate traffic is a very bursty arrival process	The aggregate traffic is a non-bursty arrival process.

Table 5 - Comparison Between ATM LAN and ATM WAN

3.5 ATM Switch Architectures

Across the spectrum of ATM networking, there are families of ATM switches that address different user needs for service environments and network scalability. Cisco's view on such hierarchy of ATM switches functionality has led to the following groups of switches:

ATM Switch Group	Characteristics
Workgroup	Optimized for deploying ATM to the desktop over low cost ATM interface. Include standardized signalling for interoperability with workstation adapters and other workgroup or campus ATM switches. Optimized for low latency and high throughput. Because there is less of a requirement for redundant components in workgroup switches they are able to be engineered with lower costs per port.

Table 6 - ATM Switch Groups

ATM Switch Group	Characteristics
Campus	for a high performance ATM backbone with multiple QoS classes, integration of LAN and ATM and standardized interswitch signalling and routing. engineered either for relatively small-scale ATM-only backbone as well as for large campus backbone. The small-scale ATM-only backbone switches can, for example, interconnect ATM-capable routers and LAN switches and/or workgroup ATM switches. requires a variety of ATM interfaces (UTP, fibre, etc.), greater traffic and congestion management control and degree of redundancy). A larger campus backbone switch allows greater network connectivity and either direct LAN or CBR ports on the ATM switch. This supports gradual user migration to ATM and multiple services on the common ATM infrastructure, i.e. more diverse multiservice capability than is needed in the ATM-only backbone.
Enterprise	includes WAN interfaces in addition to campus and workgroup LAN connections, and the system is typically used as the concentration point for enterprise communication. The reliability and availability requirements are quite stringent. A premium is placed on congestion avoidance and traffic management to optimize the use of WAN links, on diversity interfaces and services to accommodate multiple-user needs, on integration and interoperability with multiple other systems through standards or support of product specific value-added features, and on extensive network management. Being located at the core of the customer network, such switches must be highly reliable and diverse.
Carrier	Providers have a need for ATM switches with multiple media and service interfaces including CBR ATM, frame relay, LAN ports, built-in redundancy, standard-based signalling and routing, and value-added network management. ATM-capable equipment on the customer's premises to access the service or the customer connects to an ATM-based access service in the carrier s network.

Table 6 - ATM Switch Groups

Each of these ATM switch groups has its own cost, feature and performance profile. However, they all share the fundamental feature to switch ATM cells.

Some considerations about switch architecture are:
- QoS service used for traffic policing. For instance, some ATM switches treat voice-only as CBR traffic, which means they set aside bandwidth exclusively for voice calls;
- Number of VPI/VCIs which an ATM Switch can handle. If there are a lot of simultaneous end-user connections that have to be established (typically the case of corporate WAN) the switch has to be able to support a large number of VPI/VCIs on each port;
- Size of the buffer for each switch plane. For example, delay-sensitive CBR traffic (like video) would have smaller buffers because delaying would degrade the image. But UBR/ABR traffic like e-mail and Internet access is not delay-sensitive, so large buffers can be used.

4 Measurement Tools

Multimedia applications differ regarding their relative sensitivity to factors such as delay bounds, loss bounds and jitter bounds. For example, for video-conferencing, delay and jitter bounds are stringent, but some packet loss may be tolerated while for shared document editing, the delay bound may be short and the loss bound stringent, but the jitter bound unnecessary. Therefore different QoS parameters are associated with different applications and optimizing the transport protocol flow control parameters can assist in meeting the user requirements. For users, selectable QoS enables multimedia applications to be predictable and run faster, and for network managers it provides significantly better bandwidth utilization which reduces costs. Measuring the QoS performance characteristics on ATM networks is aimed at investigating QoS management mechanisms such as:

- Appropriate strategies to map application QoS requirements to ATM QoS comparing different mapping strategies;
- Efficient mapping between connections with different traffic and QoS requirements and the ATM adaptation layers and ATM VPI/VCIs, i.e. multiplexing and mapping of connections considering parameters like rate, window or TSDU length;
- Operating system scheduling and workstation configuration and their influence on the application QoS over ATM;
- ATM switch architecture and network configuration (LAN, MAN, WAN) and how it influences the application QoS parameters.

Based on such performance investigations, the user can decide on cost-effective solutions for using ATM equipment and operating and workstation environments, while considering the appropriate flow-control mechanisms and transport systems parameters.

To provide this, measurements on two layers of the application stack are required:

1. **Application layer QoS measurements** where the goal is to provide performance information on the actual application QoS provision:
    - The provision of user QoS dependent on flow control parameters of the transport system;
    - Performance study on how to multiplex (bundle) connections with the same or different application classes for the efficient use of the QoS provided by the ATM connections;
    - Obtaining statistics describing how well the QoS are provided using specific scenarios, i.e. transport mechanisms, system and network configurations.
2. **ATM layer QoS measurements.** The use of an ATM analyser tool in addition to the tools for QoS measurements at the application layer can solve many problems that deal with the efficient mapping of user QoS demands to the ATM services for specific switching architectures. For instance:
    - For a given application traffic pattern and system configuration parameters, the use of the ATM analyser and its facilities for real-time reporting of cell loss, error, delay, delay variances and misinsertion, makes it possible to provide a better explanation of performance bottlenecks and their causes. For instance, deviations in throughput using the same architecture while changing the user parameters;
    - Study of transport protocol flow and error control when ATM cells are dropped, because of congestion, using sophisticated traffic generators as part of ATM analyser tools;
    - Selection of cost-efficient mapping of user traffic to ATM service class using wire rate statistics at both the ATM and AAL layers (i.e.bandwidth utilization by user, empty and management cells, average and peak cell rate, maximum burst size);
    - For a given user traffic profile and connection bundling policy, the cell traffic profile can be analysed by using different mapping strategies to ATM VPI/VCI.

In this section we characterize tools that provide different capabilities and results. The tools that were used to make the measurements are the TTCP, XTCP and SPIMS measurement tools while the InterWATCH 95000 protocol analyser was used to clarify some of the results.

4.1 Application Layer QoS Measurement

The tools presented in this section differ in their capabilities and the results that they provide: (1) the application that they support (bulk data, transaction) and their QoS (throughput, delay, error rate); (2) their benchmark information; (3) the selection of the protocol flow control parameters (tuning of protocol parameters); (4) the specification of their traffic patterns for connections or different media; (5) their specific QoS; and (6) their ability to multiplex connections and to define complex connection scenarios describing application suites.

TTCP and XTCP

Basic measurement tools like ttcp and its enhancements xtcp allow the throughput for bulk data applications to be measured using TCP and XTP protocols.The enhancements allow the application's throughput to be analysed while varying the TSDU size as well as analysing flow-control parameters like rate and window size.

SPIMS Measurement Tool

The Swedish Institute of Science developed a benchmark tool called SPIMS [31], which allows specific QoS parameters for bulk data applications (throughput) and transaction applications (response time) to be measured. The benefits of this tool are:

- Measurements of the QoS for two kinds of applications (bulk data and request response);
- Efficient QoS benchmarks providing, in addition to its mean value, the maximum and minimum measured QoS values and its standard deviation;
- Consideration of different traffic models for bulk data and request-response applications which allows different application suites and scenarios to be analysed. For instance, request-response is a service used in distributed systems, banking transactions, conference applications and remote teaching applications.

However, a problem with this tool is that it does not consider today's QoS measurement requirements. For example:

- The variety of multimedia application QoS parameters and service considerations (for instance real-time applications and their QoS requirements);
- The different parameters and QoS supported by the transport protocols (i.e. selectable rate, window);
- Support of different network architectures (selection of different network interfaces).

The adaptation of this tool to ATM did not always provide a stable working environment while we were running our performance scenarios, even though the tool was stable in the Ethernet environment. SPIMS is based on processes that expect and use the standard Internet data paths. Transmission on Internet is slower than the data transmission paths measured on ATM. Although it is a good benchmark for some application types, this tool must be redesigned for its use on global ATM networking infrastructures.

Protocol Tuning Box

An emerging approach for measuring QoS for multimedia applications is incorporated in the Protocol Tuning Box [28]. It takes into consideration the requirements of today's protocols and networking environments. Some of the concepts that are integrated in this tool are:

- An interface to tune protocol flow control parameters (window size, rate control parameters);
- Consideration of connections for specific media and applications including graphics, pictures, video and voice transmission as well as their compression techniques;
- Consideration of a standard set of traffic models that specify connections for different application classes;
- Performance analysis of multiplexed connections specifying different traffic models or specific media;
- Data bases for storage and manipulation of required protocol parameters, traffic or media application suites and QoS results.

4.2 ATM Layer QoS Measurement

QoS measurement at the ATM layer is specified in ITU-T Recommendations I.356, I. 371, I.610 as well as in the ATM Traffic Management Specification Version 4.0 [6].

The ATM-layer QoS is measured by a set of parameters intended to characterize the performance of ATM-layer connections. These QoS parameters quantify the end-to-end network performance at the ATM-layer. ITU-T Recommendation I. 356 defines the QoS for the Bearer Service. An example for ATM QoS measurement scenario is shown in Figure 7:

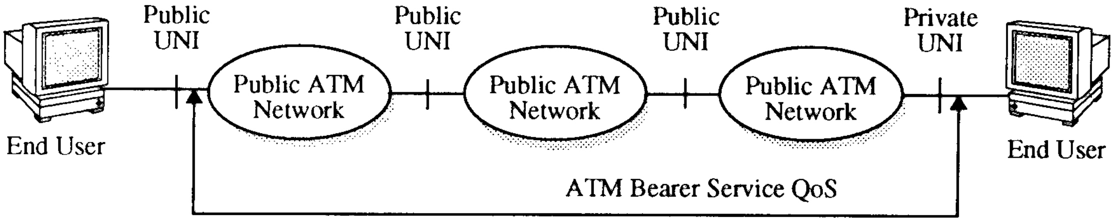


Figure 7 - ATM QoS Measurement Example

The ATM Forum Technical Committee has proposed two alternative methods for measuring the QoS performance parameters [6]. Either in-service or out-of-service methods can be used to estimate values for the ATM cell transfer performance parameters. The in-service methods are based on monitoring OAM cells that are introduced in the user cell stream at any

VPI/VCI termination or connection point. The out-of-service methods consist of establishing a test connection at any appropriate measurement point, introducing a known cell-stream content and timing and monitoring the test connection at a remote measurement point.

A variety of QoS performance parameters are defined in [6]:

- **Peak-to-Peak Cell Delay Variation (CDV)**: the term peak-to-peak refers to the difference between the best and worst case expectation of cell transfer delay.
- **Maximum Cell Transfer Delay (Max CTD):** the term cell transfer delay is defined as the elapsed time between a cell exit event at the source UNI, for instance, and the corresponding cell entry event at a destination UNI for a particular connection. The cell exit event occurs when the first bit of an ATM cell has completed transmission out of an end system. The cell entry event occurs when the last bit of an ATM cell has completed transmission into an end system. Cell transfer delay is affected by propagation, queuing, routing and switching delays.
- **Mean Cell Transfer Delay (Mean CTD)**: is defined as the arithmetic average taken over a sample of the cell population. Over LANs the MeanCTD will likely be dominated by propagation, emission, queuing and routing times. In WANs it will be mostly caused by the propagation delay over long distances.
- **Cell Loss Ratio (CLR)**: is defined as the total number of loss cells divided by the total number of transmitted cells. A cell-lost outcome occurs when no cell is received corresponding to the transmitted cell within a specified time Tmax. The cell loss ratio is expected to be caused by errors in the cell header, buffer overflows and non-ideal Usage Parameter Control (UPC) actions.
- **Cell Error Ratio (CER)**: is defined as the total number of errored cells divided by the total number of successfully transferred cells and errored cells. An errored cell outcome occurs when the payload of the received cell differs from that of the transmitted cell or when an invalid header field is detected after header control procedures are completed. An errored cell is expected to be primarily influenced by the error characteristics of the physical media.
- **Severely Errored Cell Block Ratio (SECBR)** is defined as the total number of severely errored cell blocks divided by the total transmitted cell blocks, where a cell block is a sequence of N cells transmitted consecutively on a given connection. A severely-errored cell block outcome occurs when more than M errored cells, lost cells or misinserted cells are observed in a received cell block. Is also expected to be influenced by the error characteristics of the physical media and by the buffer overflows. Operational effects such path reconfiguration may also introduce errors.
- **Cell Misinsertion Rate (CMR)** is defined as the number of misinserted cells divided by a time interval, where misinserted cells are received cells for which there is no corresponding transmitted cell. It is most often caused by undetected/ miscorrected errors in the cell header that get mapped into the wrong VPI/VCI connection. This test provides a good indication that network equipment is overloaded or is being reconfigured and is misrouting and mismulticasting cells.

The first four QoS parameters can be negotiated between the end systems and the network(s) using the UNI and network signaling protocols. The last three parameters, on the other hand, cannot be negotiated.

ATM protocol analyzers can be used for accurate QoS performance measurements and for testing ATM switches and networks for proper operation. The use of ATM protocol analyzers along with QoS measurements tools can aid the user in mapping its QoS demands to the ATM services of a particular switching architecture.

Both the Adtech AX/4000 and InterWATCH 95000 protocol analysers provide real-time cell testing which is essential for accurate QoS measurements and for ATM switch performance analysis.

The InterWATCH 95000 [30] is a portable multiport, LAN, WAN and ATM protocol analyser for high performance applications. The interWATCH 95000 can be used to generate LAN and WAN traffic, to synchronize or correlate frames from several monitor protocol traces and to monitor the data coming from a specific source. In the monitor mode, a capture buffer is automatically used as the data sink. The user can select the protocols that will be collected and analysed by the monitoring device. The selections determine what type of data is placed in the capture buffer. Statistics on the collected data can be viewed. Lists, tables, pie charts and histograms are presented as appropriate. The interWATCH decodes the ATM cell layer and the ATM adaptation layer 5. In the case of the ATM adaptation layer 5, a decode window shows the AAL5 PDUs that were received along with the time at which each AAL5 PDU was captured. The delta time between consequent AAL5 PDUs can be displayed in absolute or relative mode. The interWATCH 95000 also incorporates a unique multiport correlation feature with 1-microsecond resolution for easy identification of interoperability and LAN to ATM adaptation problems. The InterWATCH provides ATM Layer Filters which can be applied to explain QoS performance analysis results more precisely. Such filters apply to VPI, VCI, Empty Cells, Errored Cells, User Cells, High Priority Cells and Low Priority Cells.

The Adtech AX/4000 [33] provides complete physical, ATM and AAL layer generation and analysis for testing ATM equipment, switches and networks for proper operation and for testing QoS performance. Each Analyser port includes 16 programmable cell filters for separating incoming cell streams into 16 substreams. The system can provide simultaneous cell traffic statistics in real time for the entire cell stream, for each port, and each of the 16 cell substreams. An optional traffic shaper allows the user to set peak cell rate limit and sustainable cell rate limit parameters. An error injector can be used to inject errors at the cell level. In addition, the Adtech provides the capability to do ATM network emulation.
By creating an ATM network environment, the Emulator allows users to test their system's performance through a comprehensive range of ATM network characteristics and impairments. The emulated network is precisely modelled using familiar ATM QoS parameters such as cell delay, cell delay variation, cell errors, cell loss, cell misinsertion and out-of-sequence cells. Traffic policing with limiting of cell transmission rates is also incorporated. The Emulator's traffic policer monitors the foreground cell stream for traffic conformance violations based on Usage Parameter Control (UPC) procedures, and performs either cell tagging or discards the nonconforming cells according to user specified probabilities. The Emulator's traffic policer is based on the ATM Forum's Dual Leaky Bucket Generic Cell Rate Algorithm (GCRA). It uses traffic conformance test parameters such as peak cell rate, sustainable cell rate, cell-delay variation and maximum burst size.

5 Transport Level Flow Control Mechanisms on ATM

Transport protocol flow control mechanisms such as window advertisements and rate control must be used to map the user QoS requirements to the QoS provided by ATM WAN and LAN VPI/VCI channels. We focus on the efficient use of these mechanisms to support application classes, such as bulk data and transaction which require reliable transmission. Whereas the transaction application is extremely sensitive to the response time, the bulk data applications requires at least a minimum throughput guarantee.

For efficient QoS support of reliable applications on ATM WAN networks, the flow control mechanisms of transport protocols have to be adapted to the specifics of these networks. For instance:

- Integration of efficient flow and congestion control for high delay networks. A high delay network is characterized by the ratio of the transit delay of a protocol data unit (PDU) and the PDU processing time at the end systems. If this ratio increases, the basic mechanisms for classical Automatic Repeat Request (ARQ) type protocols become less efficient and can lead to severe performance problems due to the long delay in the sender-receiver control loop [20];
- Support of application-specific QoS requirements (throughput or delay) using appropriate flow control parameters. For instance, support of user throughput mapping into flow control parameters such as window size and rate;
- Consideration of the specifics of the operating system and Application Programming Interface (API) for efficient mapping of user QoS and traffic requirements to the ATM adaptation layer;
- Consideration of the specifics of the ATM Adaptation Layer (for instance AAL 5 MTU size);
- Integration of appropriate flow control mechanisms that correspond to the requested QoS and traffic parameters (i.e. service category) of the ATM connection. For instance, if mapping to ABR is required, to support rate renegotiation in order to process the feedback from this service;
- Consideration of the fragmentation problem of the ATM networks and ATM-specific schemes to deal with fragmentation.

The focus of our performance issues over ATM is the tuning of window and rate-based transport flow and congestion control mechanisms for efficient QoS support and network utilization.

The *Window (W)* defines the maximum amount of data which is allowed to be contained in the pipe and at the receiver queue, i.e. the maximum amount of data the sender can send without receiving an acknowledgment to move the window. To "fill the pipe", the window must match the *delay*bandwidth* product. The common method to select the window size is to consider the receiver's buffers. In addition, the selection of the window size can take into consideration:

- application QoS requirements (throughput);
- buffer, rate and burst requirements of the link.

Window flow control provides a burst restriction in combination with an ARQ scheme. It can define a rate, i.e. bytes per second, in an indirect way by the ratio of window and Round Trip Time (RTT).

With rate control the sender is able to restrict the burst in a given time interval. Rate control is not based on an ARQ scheme, but may use it for dynamic rate adaptation.

5.1 TCP Window Control

TCP [22], [23] was designed to provide reliable best-effort service over a connectionless network layer such as the classical IP. The protocol mechanisms of TCP are listed below along with the problems that occur when trying to ensure reasonable performance on high delay networks:

TCP Window Based Flow

TCP supports *restricted tuning of protocol parameters* to application and network demands (based primarily on the window). TCP uses window based flow and congestion control. The *maximum window* that a connection is allowed to use (i.e. for which the connection can be charged in the multimedia environment) is called *send window* and can be set (setsockopt) depending on:

- The receiver buffer;
- User throughput requirements and;
- The network bandwidth to be used by the connection.

The actual window in TCP is determined by the minimum of the send window and the *congestion window*. The congestion window is dynamically adapted to the available bandwidth*delay product by TCP congestion control (Slowstart, Congestion Avoidance). TCP sends traffic in a sliding window up to 64 KB. Over WAN ATM, TCP can cause the transmitting station to remain idle for extended periods thus robbing ATM of a significant portion of its available bandwidth. The source has to wait until it receives an acknowledgment from the address to which it has transmitted before sending further frames.

Slowstart and Congestion Avoidance

The purpose of Slowstart and Congestion Avoidance is to adapt the PDU flow to the available bandwidth since the transport agent has no idea of the available resources at connection establishment. Slowstart begins with the congestion window set to one PDU. At the arrival of an Ack, the window slides by one PDU and the congestion window is incremented by one PDU, so two PDUs are sent. This results in an exponential increase over RTT.

When the congestion window has reached the Slowstart threshold, Congestion Avoidance takes over control. The sender's packet flow is assumed to be adapted to the available bandwidth, i.e. the congestion window matches the bandwidth*delay product. A PDU is only sent in response to an Ack as the receipt of an Ack means that a PDU has left the network ("conservation of packets principle"). Sending PDUs only in response to Acks adds a "self-clocking property" to TCP. Congestion Avoidance increases the congestion window slightly by one PDU per RTT to determine if there is more bandwidth available (linear increase).

Slowstart takes RTT * ld(congestion window) seconds [23]. Performance loss increases with increasing Slowstart duration, which depend on the RTT and window size that are both high on WANs.

Slowstart can cause *burst* and retransmission problems if TCP overruns the network bandwidth when trying to reach the optimal window.

However, when a service with guaranteed bandwidth is used, Slowstart leads to significant performance loss and poor ATM bandwidth utilization.

Flow Control and QoS Guarantee

Reliable ARQ based protocols have to specify QoS requirements in two directions: in the direction of data transmission and in the reverse, the direction of their acknowledgments.

There is an indirect way to map the TCP user data flow to bandwidth requirements. The ratio of the window size and RTT, which gives the required rate, must be considered. Some ATM services, like CBR or rt-VBR, support maximum delay (maximumCDT) requirements, others the mean delay (meanCDT). However, even when the network supports the requested delay, a powerful mechanism for accurate delay (RTT) measurements is required. For instance, in [58] time stamps are used.

The bandwidth requirements on the reverse path depend on the acknowledgment steps. Acknowledgment steps are defined by the amount of sent PDUs triggering an acknowledgement at the receiver. TCP requires one acknowledgment per PDU which means a significant bandwidth utilization in the reverse direction.

TCP Enhancements for High Delay Networking and Efficient QoS Provision of Reliable Applications

1. In RFC 1323 [58], a "*window scale option*" was introduced to adapt the TCP flow for LFNs, resulting in the specification of large windows. Although it cancels the send window size restrictions (the send and congestion window can be increased up to 1 GB), it implies several drawbacks [20]:
- The credit field is right-shifted by the receiver and left-shifted by the sender for each acknowledgment. This causes additional overhead and restricts the accuracy of the send window to the nearest multiple of the scaling factor;
- Advertising a send window smaller than the scaling factor during the connection lifetime is impossible.

In addition, adapting large windows to the available bandwidth results in high variation of network utilization (e.g. TCP) leading to PDU loss and expensive retransmission as well as degradation of the QoS provision of simultaneous traffic. One solution is to shape the bursts by using rate control.

2. Selective Acknowledgments (SACK option [56]) are used to enhance the performance when errors occur on high delay networks. However, there are drawbacks to TCP selective retransmission in a high delay environment:
-The probability of many lost PDUs per window is higher and RTT is longer on LFNs. Hence, the restricted amount of gaps that can be communicated to the sender increases the time for error recovery, especially on LFNs. Taking into account several RTTs for retransmission, this is a significant performance loss;
-Dividing the congestion window by two in case of a lost PDU may be useful to quickly allow a router to recover from congestion, as is proved in the Internet. On multimedia highways, smooth adaptation of the transmission rate is preferred as the transport agent knows the connection's bandwidth and has more information about lost PDUs.

5.2 XTPX Rate Control

The Express Transport Protocol eXtension, XTPX [26] was developed for multimedia communication and can be tailored to different network qualities with its protocol tuning mechanisms [25]. XTPX is derived from the XTP 3.6 revision and incorporates its own network layer with routing facilities. XTPX supports a connection-oriented network-layer multicast service as well as a connectionless IP multicast service. The following are features of interest:

- **Separation of paradigm and policy**. XTPX provides a set of mechanisms whose functionality is orthogonal to one another, resulting in a clear separation of communication paradigm (datagram, virtual circuit, transaction, etc.) from the error control policy (fully reliable through uncorrected);
- **Separation of rate and flow control**. XTPX provides mechanisms for shaping rate control and flow control independently. Further, flow and rate control as well as error control can be tailored to the communication at hand. If desired, any of these control parameters can be turned off.

XTPX's error control mechanisms include selective repeat and go-back-n. Flow and rate control algorithms, certain protocol modes, and traffic shaping information are used to provide the requested quality of service as efficiently as possible. For a detailed description of XTP and XTPX, refer to the protocol specification [73],[74] and [75].

XTPX Rate Based Flow

Rate Control in XTPX is a flow control mechanism defining the maximum burst of PDUs the sender transmits in a given time interval. In XTPX, *burst (B)* specifies the maximum number of bytes allowed to be sent in *interval (I)*. *Rate (R)* is the result describing the maximum number of bytes sent per second.
It is obtained by: $R = B / I$.
Depending on user class and QoS, rate control can shape the traffic of the connection.

Congestion Avoidance, Traffic Shaping and Multiplexing

When rate control is not used, large bursts, that depend on the window size, are separated with long time periods between them. Such a burst structure can degrade the QoS especially for isochronous connections which require bandwidth in constant time intervals. Therefore for efficient resource utilization, rate control can be used to smooth the burst.

Due to the control of the time interval between the PDUs sent on the network, rate control can be used for traffic shaping and multiplexing of different kinds of connections including bulk data and isochronous connections, as well as for "traffic squeezing", i.e. application traffic policing [24]. With window flow control, the sender cannot explicitly control the time intervals between sent PDUs.

Flow control and QoS guarantee

Mapping to bandwidth resources of the network using the rate can be performed directly. The way to map to the network resources based on the window is indirect and depends on the delay and delay jitter. To guarantee user throughput, rate control is more efficient than the window because it does not depend on network delay.

5.3 Performance Issues of Flow Control Mechanisms on ATM Networks

The unique characteristic of ATM switching is the fragmentation of adaptation layer packets (for instance 9 KB) into small ATM cells (53 bytes) where the loss of any single cell will cause the loss and retransmission of the whole packet.
This can lead to a significant decrease of the performance, i.e QoS degradation such as throughput loss for bulk data, increase in response time for transaction applications or increased error rate for continuous and real-time applications.
Various research projects deal with the fragmentation problem on ATM networks and its performance as affected by the flow control mechanisms and parameters on the transport or ATM layer.
Using simulations, various TCP/IP over ATM performance issues have been studied:

- **Selection of transport protocol and ATM layer flow control parameters** such as buffer size, and TCP packet and window size in the case of ATM congestion [15];
- **Avoidance of performance loss due to fragmentation in congested ATM networks** [18]. In the case of congestion, the ATM network can apply intelligent cell-dropping mechanisms to improve the overall throughput. Such packet discarding strategies, proposed to alleviate the effects of fragmentation, are *Partial Packet Discard*, in which remaining cells are discarded after one cell has been dropped from a packet, and *Early Packet Discard*, a strategy that prevents fragmentation and brings throughput up to the maximal level by having the switch drop whole packets prior to buffer overflow;
- **Effect of fragmentation overhead in the form of padding bytes** [16]. When the length of the frame is not an integral number of usable payload length, the last cell for that frame is padded to the required fixed length. The amount of overhead contributed by fragmentation is equal to the number of bytes left unused in the last cell for the frame;
- **Utilization of ATM switches with large buffers and additional congestion control mechanisms** [13] in order to improve TCP/IP performance in a high speed WAN ATM;
- **Congestion-control schemes on a multi-hop ATM network and their effectiveness** such as Early Packet Discard and link level flow control at the ATM layer for TCP/IP [9].

Because of the recent introduction of ATM switches in real network infrastructures, a few studies are now dealing with practical issues of TCP/IP, especially on LAN ATM [19], and are focusing on aspects of selecting appropriate TSDU and MTU sizes, system configuration parameters, etc.

6 Application Admission Control on ATM Networks

Information highways allow applications to share resources on ATM WANs. This requires analysing the traffic and QoS objective of applications running simultaneously. Resources can be used in a cost-efficient way, when the traffic's statistical distribution of individual connections can be multiplexed over the same resources without degrading the QoS boundaries that were negotiated for each connection. To realise the cost and performance expectation from the ATM technology, research and practical experiences on connection admission control on the information highways are required.

In this context, we address the issue of translating the flow / call control specifications (traffic parameters and QoS requirements) into the corresponding ATM traffic and QoS parameters. Admission control policies for different internet classes of applications over ATM are discussed. Additional information on the Internet QoS models and on different strategies to map the application traffic into ATM resources are also included in Section 6.3 and 6.4.

6.1 Mapping of Application Traffic and QoS Specifications to ATM networks

A common way to specify application traffic characteristics is to define:
- Interarrival TSDU time distribution;
- TSDU length distribution.

The traffic characteristics of an application can be expressed as maximum rate, minimum rate and burst and then mapped to flow control and ATM traffic parameters. The traffic characteristics correspond to the ATM traffic descriptors (PCR, SCR, MBS). The mapping of the traffic parameters determines the policy of the transport flow mechanism that will be used to control the traffic.

This policy can be additionally determined by the application QoS requests such as:
- Bulk data: desired and threshold throughput;
- Transaction: desired and threshold response time.

Applications can demand minimum or maximum values for their QoS requirements. For instance, if an application specifies that the TSDUs transmitted should not exceed a maximum rate and the traffic exceeds the specified rate, TSDUs can be simply discarded.

The mapping of application traffic or user QoS to appropriate transport flow control parameters for the case of simple traffic mapping is shown below.

Application	Transport Flow -> ATM AAL PDUs	ATM Layer
max TSDU rate in TSDU/s or bytes/s	**max_R (max rate)** in bytes/s sent as ATM AAL PDU/s	**PCR** in cell/s
mean TSDU rate in TSDU / sec or bytes/s	**mean_R (mean rate)** in bytes/s sent as ATM AAL PDU/s	**SCR** in cell/s
TSDU burst in TSDUs or bytes	**B (burst)** in bytes sent as ATM AAL PDUs	**MBS** in cell

Table 7 - Mapping of Application Traffic to Transport flow Control Parameters

The table shows the mapping requirements resulting from the API (TSDU or byte oriented requirement), transport flow control parameters (byte oriented rate which is mapped in ATM AAL PDUs which are then segmented in the cell granularity of the ATM-layer).

If we assume that the application traffic requirements are mapped at a given rate of the transport flow control, and if we consider a perfect fluid model [57] for the transport flow and ATM cells, and if we ignore the segmentation overhead and ATM cell granularity, the following relation can be established:

PCR = max_R / cell_size
SCR = mean_R / cell_size
MBS = B / cell_size

However, this relation depends on the adjustment of the transport traffic parameters given in bytes/s in AAL PDUs and on the fragmentation of the AAL PDUs in cells.

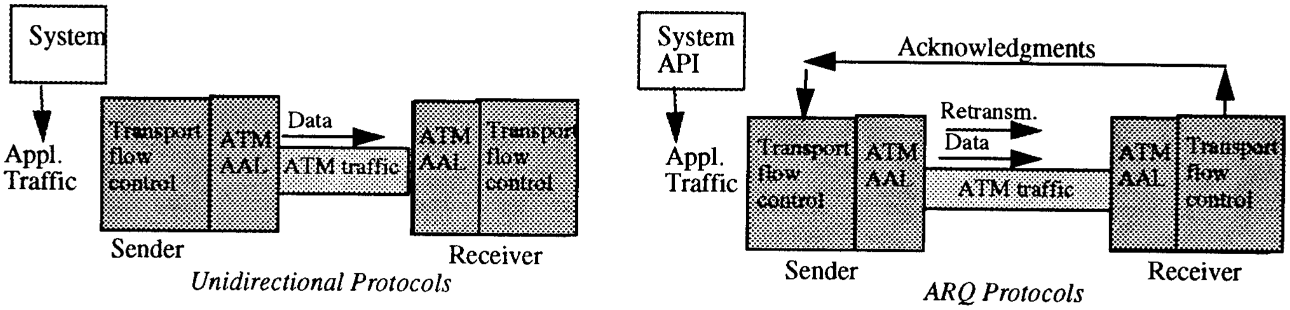
The accuracy of the mapping of application traffic to transport flow control parameters depends on the type of parameters used-window or rate.When using the rate parameter, direct traffic mapping is possible. However, when using a window flow control parameter, the resultant rate depends on the delay of the connection and can be expressed as Window Size / RTT. In addition, the "burst" of the window flow control is given with the window size (congestion or send window using TCP "Self clocking mechanism" or Slow Start mechanisms). There is no explicit burst mechanisms as in the rate control technique.

XTPX maps the rate and burst parameters while TCP maps the maximum send window size.

Additional factors which influence the traffic mapping and its policing are:

1. At the application-layer, the system configuration shapes the traffic based on the scheduling mechanisms. The specifics of the API design and its interaction with the transport flow control can also impact the traffic;
2. Traffic of reliable service, in contrast to unreliable service, can be influenced by the retransmission and the congestion policies which are considered for retransmissions. A reliable service is based on the ARQ protocol scheme which involves retransmissions that influence the resultant traffic on the network. How retransmissions in the framework of TCP or XTPX can influence the resultant traffic is discussed in [20].

These factors are shown in Figure 8.



Unidirectional Protocols

ARQ Protocols

Figure 8 *- Traffic Policing for Different Services*

The following figure gives an overview of the factors that influence traffic and QoS provision of reliable applications, i.e. those that are based on ARQ schemes.

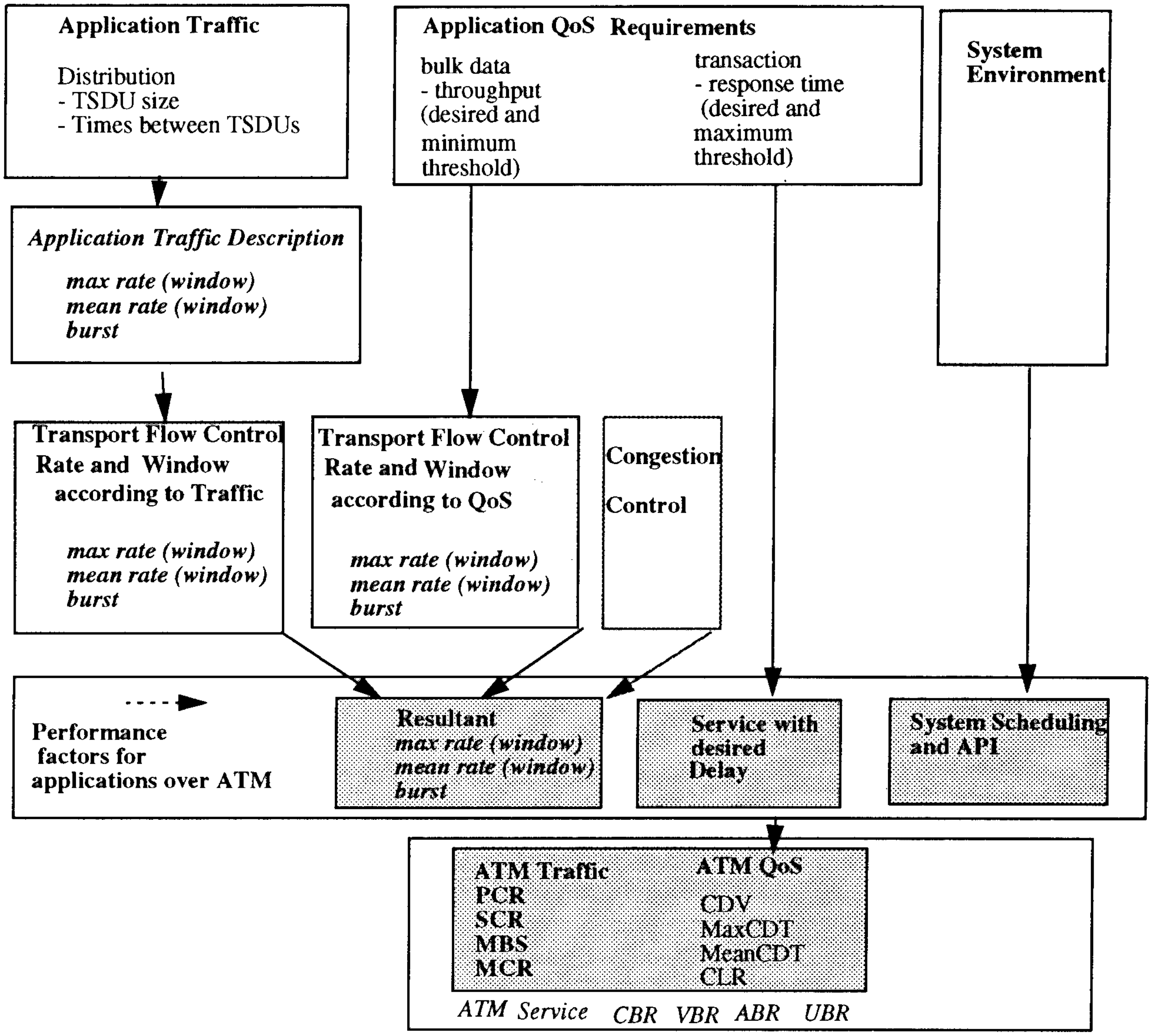


Figure 9 - Mapping of Application Traffic to Transport Protocol Flow Control and ATM Services

The measurements described in Sections 7 and 8 use different mapping strategies based on this scheme, and show the impact on the QoS of the applications and on the selected admission policies.

6.2 Admission Policies for Transport Connections on ATM Networks

In order to use the ATM networks more efficiently, resources can be reserved not only for individual connections but also for groups of connections depending on their traffic and QoS requirements. This is especially important for multimedia traffic where different transport requests are processed simultaneously. Policies for processing the transport connections with different kinds of traffic and QoS requirements are required. Admission control policies for transport connections on ATM networks define strategies for multiplexing/bundling the traffic from different connections to ATM channels.

Different multiplexing policies are proposed in [10] such as:

- ATM VPI/VCI per IP router pair;
- ATM VPI/VCI per conversation, and;
- ATM VPI/VCI per application type, for instance one VC for all FTPs passing through a pair of routers.

Considering the factors for mapping the application traffic (i.e. traffic parameters, system environments and user QoS requirements) admission policies can be based on bundling of connections. Bundles can be built using different types of transport connections:

- Unidirectional connections with similar traffic patterns and QoS requirements, for instance bulk data transfer connections;
- Unidirectional connections with different traffic and QoS requirements;
- Bidirectional connections with similar traffic patterns and QoS requirements;
- Bidirectional connections with different traffic and QoS requirements.

The requirements for bundling connections for ATM VC services are based on predicting the multiplexed traffic and QoS provision as well as evaluating of the system configuration influence. When connections are bundled at the end system, the system scheduling becomes an important performance factor especially the bundling of bidirectional connections which depend primarily on the scheduling in the operating system.

Using an adaptive scheme, i.e. by monitoring the throughput of a given bundle over time, an appropriate set of requirements for an ATM virtual circuit (for example peak and average rate) can be obtained. Section 8 demonstrates the results of bundling different types of transport connections.

6.3 Internet QoS Models

QoS models for flow guarantee, proposed in the IETF, facilitates the mapping of the QoS requirements of Internet applications to ATM [60], [61], [62], [63], [64], [65]. The Internet QoS requirements are passed on to the network elements as traffic specification (TSpec) and service specification (RSpec). QoS control is specified for traffic conforming to the TSpec given at setup time. The traffic specification is defined as a token bucket with a given bucket depth b (in bytes) specifying the maximum allowed burst size for the flow, and a bucket rate r (in bytes/second) giving the average rate of the flow. Policing of the flow can then be performed based on the traffic specification, i.e. the traffic must obey the rule that for any time unit T, the amount of data sent cannot exceed $rT + b$.

The following table summarizes the Internet QoS models. Some correspondence of the Internet services to ATM services is shown, for instance "guaranteed" corresponds to the CBR and rt-VBR services.

Internet QoS Guarantee	Description and Parameters
Guaranteed [60]	for applications which need a firm guarantee that a TSDU will arrive no later than a certain time after it was transmitted by its source. For example, some audio and video "playback" applications are intolerant of any packet arriving after their playback time. **absolute bound on maximal packet delay**: - It guarantees that packets will arrive within the guaranteed delivery time and will not be discarded due to queue overflows, provided the traffic stays within its specified parameters.
Predictive [61]	for playback applications that require a reserved rate with low packet loss and a maximum bound on end-to-end packet delay, and that can tolerate occasional dropped or late packets **fairly reliable delay bound:** The large majority of packets are delivered within the delay bound - average delays that are no worse than best effort service, and the maximum delays should be significantly better than best effort service when there is significant load on the network real-time service - Packet losses are rare as long as the offered traffic conforms to the specified traffic characterization
Controlled Delay [62]	service-adaptive and delay-adaptive applications; i.e., applications that are prepared to dynamically adapt to changing packet transmission delays and to dynamically change the level of packet delivery delay control they request from the network when their current level of service is not adequate **no quantitative assurance about maximum end-to-end delay:** - average delays are no worse than best effort service, and the maximal delays should be significantly better than best effort service when there is significant load on the network. - Packet losses are rare as long as the offered traffic conforms to the specified traffic characterization
Controlled Load [63]	applications sensitive to overloaded conditions, for instance "adaptive real-time applications" work well on unloaded nets, but degrade quickly under overloaded conditions. **no delay or loss specification:** - acceptance of a request for controlled-load service is defined as a commitment by the network element to provide the requester with service closely equivalent to that provided to uncontrolled (best-effort) traffic under lightly loaded conditions. - a network element may employ statistical approaches to decide whether adequate capacity is available to accept a service request.

Table 8 - Internet QoS models

Internet QoS models and flows are mapped to ATM resources using UNI signalling 3.1 and 4.0 [52], [53].To provide the correspondence between Internet QoS based transport flow and ATM channels, the Session Identity Notification Protocol (SINP) is proposed in the Internet community [41].

6.3.1 Internet Resource Reservation over ATM

Because of the diversity of Internet traffic and QoS requirements, different strategies are used to map the application traffic into ATM resources:

- Using flow specification and resource reservation protocols - RSVP [11], ST-2 [29], ST+ [30];
- Direct mapping for each application type, i.e. well-known port numbers, are assigned appropriate ATM QoS resources (AREQUIPA) [40].

We focus on these two approaches to establish real-time flows with QoS requirements based on flow specification with resource reservation protocols and direct ATM service requests by the applications.

RSVP[11]

RSVP was designed for the Internet Integrated Service Architecture[55] and it supports flow and filter specification at the IP level. It provides a general facility for creating and maintaining distributed reservation state across a mesh of multicast or unicast delivery paths. Different models for mapping RSVP over ATM are discussed in [57] and [44]. RSVP flow specification is mapped into appropriate traffic class and QoS parameters using the specific ATM signalling UNI 3.1 or 4.0 functions.

RSVP supports several reservation models or "styles" to fit a variety of applications - unicast or multicast - which require scalable resource reservation for unidirectional data streams. i.e. RSVP is simplex: it reserves data flow in one direction only. In order to efficiently accommodate heterogeneous receivers and dynamic group membership, RSVP makes receivers responsible for requesting resource reservations.

The following figure illustrates the RSVP PDUs and the basic protocol mechanisms for resource reservation:

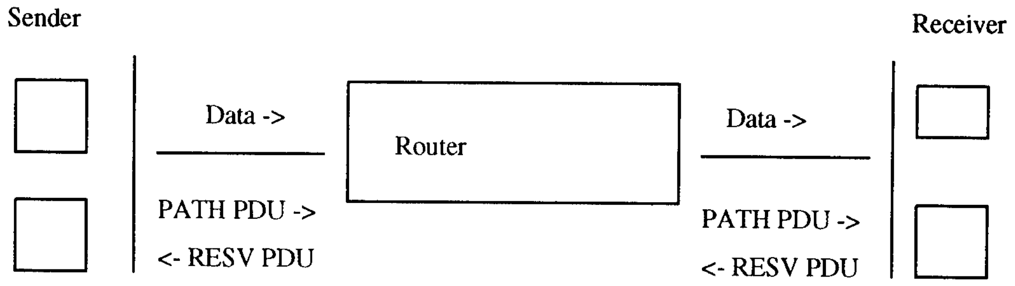


Figure 10 - RSVP Mechanisms for Resource reservation

There are two fundamental RSVP PDU types:

- PATH PDUs that carry the traffic characteristics of the data stream the sender will generate. Each RSVP sender host transmits a PATH PDU downstream along the uni-/multicast routes provided by the routing protocol(s), following the path that the data will use and storing "path state" in each router along the way. They may optionally carry a package of "One Pass With Advertising" (OPWA) information that may be used at the receivers to predict the end-to-end QoS;
- RESV PDUs carry resource reservation requests that are sent upstream towards the senders.They follow, in reverse, the routes the data packets will use, upstream to all the sender hosts included in the sender selection. RESV PDUs must be delivered to the sender hosts so that the hosts can set up appropriate traffic control parameters for the first hop.

Although designed for the integrated service environment, RSVP cannot be efficiently used for reliable services like file transfer or transaction for many reasons:

- It does not support reservations for duplex streams which is important for reliable transport services based on ARQ techniques, i.e. one channel for the data stream and reverse channel for the corresponding acknowledgments. In order for reliable applications to use the ATM networks more efficiently (i.e. cost-wise), they can multiplex their traffic using different bundling or grouping concepts.(i.e. acknowledgments and data streams);
- RSVP is a valuable resource reservation service for multicast distributions which can tolerate errors and which are delay-sensitive. PDUs can be dropped if the sender allocates more resources than the receiver required. This is extremely undesirable for reliable services.

ST-2 [29]

ST-2 is a connection-oriented protocol for establishment and maintenance of point-to-point and multipoint network connections with predefined flow control characteristics. ST-2 allows an application to specify the desired characteristics of a stream by using a Flow Specification. QoS values such as maximum end-to-end delay, smallest packet size, minimum bandwidth, lowest packet rate, accumulated mean delay, accumulated delay variance, desired PDU size, desired PDU rate are all part of the Flow Specification (FlowSpec).

Stream setup is initiated when a source ST agent generates a Connect message listing the flow specification. When an intermediate ST agent receives the Connect message, it must invoke the resource reservation, update the flow specification and send the Connect message to the next intermediate ST agent(s) leading to the target(s). Upon receiving a Connect message a receiver must determine whether it wishes to join the session and return either an Accept or a Refuse message to the source. The receiver may further reduce the resource request by updating the returned flow specification.

The stream source must examine the flow specification in a returned Accept and either reduce its QoS if the flow specification has been modified or reject the new receiver if the resources allocated are less than those currently allocated for the stream.

After stream setup the source must conform to the minimum resource allocation (i.e. minimum throughput, maximum delay) forcing all participants in the session to adapt with the least capable or demanding receiver. To satisfy the most demanding receiver the source must allocate the maximum requested resources along all links.

Reliability and robustness are incorporated into the ST-2 protocol via two separate mechanisms. First, all control messages used to create and manage a stream are transmitted reliably using hop-by-hop acknowledgments with retransmission. Second, a Hello protocol is used to query the status of neighbouring ST agents sharing active streams. Therefore ST-2's control messages are reliable.

Adjacent nodes in an ST-2 stream use a common Hop Identifier (HID). The sending agent can propose the HID to be used or it can defer the selection to the next-hop agent(s). When the sending agent proposes an HID already in use by one of the targets, the target can propose a set of free HIDs. The sending agent can then choose another HID and transmit it to the receiving agents. The process continues until all receiving agents accept the proposed HID. Data transfer in ST-2 relies on reservation; it does not contain error correction mechanisms for data exchange. Reliability can be provided by layers on top of ST-2 when required.

ST-2 supports a full-duplex option. A second stream in the reverse direction, from the targets to the origin, can be automatically created with a reverse-flow specification. If there are many targets, the resulting stream allows bidirectional communication between the origin and the targets, but not among the targets.

ST-2 requires routing decisions to be made at several points during stream setup. ST-2 assumes that an appropriate routing algorithm exits to which it has access. The setup protocol is not aware of the criteria used by the routing function to select routes.

ST-2+[30]

ST-2+ is an extension for receiver-initiated connection establishment. It allows for receivers to initiate reverse connection establishments to either the origin or routers in the multicast tree. This approach allows intermediate nodes to execute stream management functions on behalf of the sender. The protocol gains the property of scaling better with the number of receivers per stream as the origin does not need to keep track of every target. The major differences between ST-2 and ST-2+ are listed below:

- Hop Identifiers (HIDs) were eliminated because they added much complexity to the protocol and were found to be a major impediment to interoperability. HIDs have been replaced by globally unique identifiers called Stream Identifiers (SIDs);
- A number of stream options have been eliminated, they were thought to add more complexity than value. They include: point-to-point, full-duplex, reverse charge and source route;
- The concept of "subnet" was eliminated because it led to interoperability problems;
- Separation of functions between ST and supporting modules are defined;
- Allows receivers to initiate reverse-connection establishment;
- Clarification of the capability of targets to join a stream or to leave a stream.

The setup protocol for ST-2 and ST2+ is called the Stream Control Message Protocol (SCMP) and is responsible for establishing, maintaining and releasing real-time streams as ATM signalling allows connection setup within ATM networks.

With ST-2 and ST2+ a Local Resource Manager (LRM) located at intermediate systems along the path from source to destination end-systems interprets the flow specification that is carried by the setup protocol and performs admission control, analogous to ATM CAC, and allocates resources required to support the requested traffic flows. The LRM also performs packet-level traffic shaping, scheduling, policing, and so on, during the data transfer phase in the same manner that ATM switches manipulate cell streams so as to provide the guaranteed QoS. ST2 and ST2+ can hence be thought of as providing a similar traffic contract specification function with respect to packet level traffic flows that ATM UNI and NNI signalling play with respect to cell flows.

Routing protocols provide ST-2 and ST2+ with the path to reach each of the desired destinations. The routing functions are called on a hop-by-hop basis and provide next-hop information much as VC routing protocols are closely coupled with UNI and NNI signalling.

HIDs and SIDs are used much the same way as VPI/VCI are used to identify streams of ATM cells.

The flow specification associated with each stream that characterizes the traffic parameters of the flow can be related to the ATM contracts that are associated with the ATM connection. Figure 9 shows the mapping of the Internet integrated services into ATM using ST-2 as the signalling protocol.

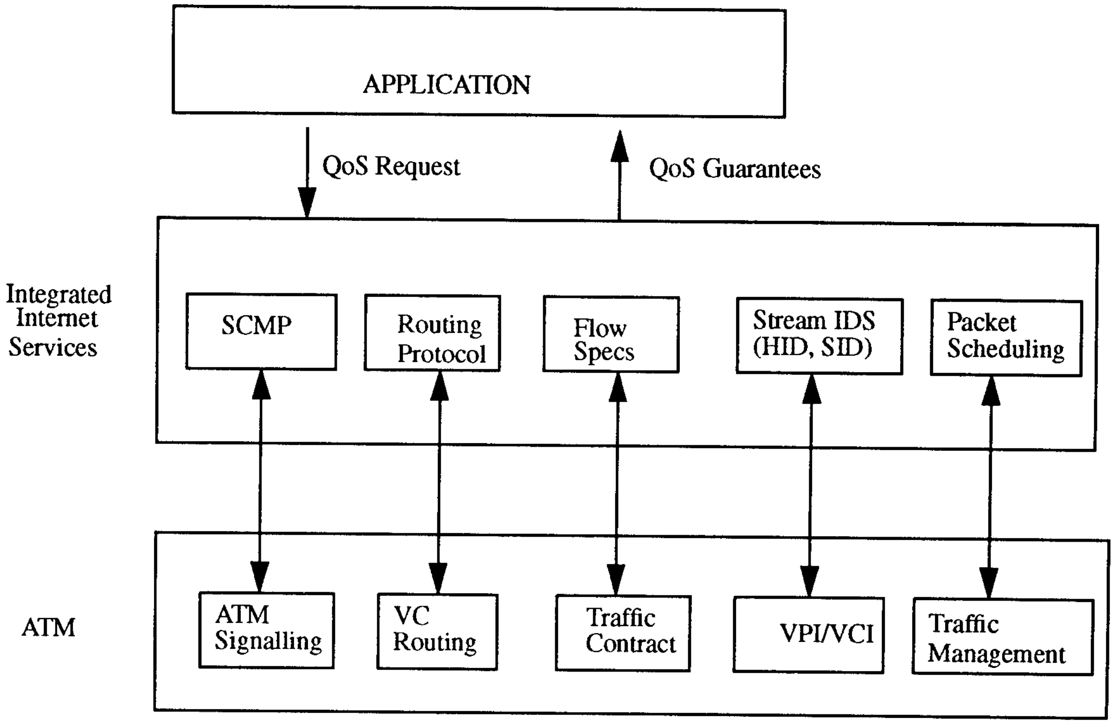


Figure 11 - Mapping of the Integrated Services Internet into ATM using ST-2

ST-2 differs from RSVP in its approach to establish resource reservations. ST-2 is based on streams from sender to the receiver. The possibility to reserve resources in both directions with the same stream-establishment step makes this protocol attractive for reliable services. However, the evolution of ST-2 to ST2+ eliminates this facility of duplex connection specification introducing the restriction of supporting only simplex streams.

Application REQuested IP over ATM (AREQUIPA) [40]

This method allows applications to request that two machines be connected by a direct ATM connection with given QoS at the link level. AREQUIPA makes sure that only data from the applications that requested the connections actually go through that connection. After setup of the AREQUIPA connection, the applications can use the standard IP protocol suite to exchange data. AREQUIPA uses an extended socket semantic for preparing the system to accept an incoming ATM connection, establishment of a new ATM connection to the given address and specified socket with the requested ATM service and QoS, and closing of the corresponding ATM connection.

7 Scenarios for Performance Analysis of Transport Flow Control on ATM Networks

The measurements presented in this section are aimed at analysing the performance and QoS provision of bulk data and transaction applications on ATM WAN and LAN. Different performance factors that influence the mapping of the traffic and the QoS requirements to the transport flow control mechanisms and to the ATM services are considered. The focus of the measurements is on the behaviour of the QoS of these applications on local and wide-area ATM networks. The possible QoS variations and the desired and threshold QoS values that are required for reliable applications are shown. The scenarios that are presented in this section address the following issues:

- Impact of rate and window size on the mapping of the application QoS requirements;
- Impact of the application traffic on the performance (small TSDU sizes, TSDU size and inter-packet delay distribution);
- Impact of system configuration (scheduling) and application programming interface (API) on ATM WAN and LAN;
- Performance issues of ATM network adaptation layer parameters (i.e. MTU size);
- Congestion control mechanisms of transport protocols (TCP Slow Start) and their influence on the ATM WAN and LAN.

7.1 Throughput Provision of Bulk Data Applications in ATM Networks

The following scenarios demonstrate the TCP flow control issues for throughput QoS provision of bulk data applications on LAN and WAN ATM. The following topics are addressed in this section:

- Comparison of the throughput QoS provision on LAN and WAN ATM for different TSDU lengths of the bulk data transmission. Different sizes of the TCP send window are considered and are evaluated as the throughput mapping mechanism for the application;
- The influence of the system configuration and scheduling (i.e. Application Programming Interface) on the throughput QoS provision is shown;
- Evaluation of the influence of the traffic parameters, i.e. TSDU length, on the throughput on LAN and WAN considering the system scheduling effects.

7.1.1 Transport Flow Control and its Impact on the Throughput in ATM WAN

Goal. These measurements show the influence of the TCP flow control parameter "send window size" on the throughput of applications in an ATM WAN environment, using different TSDU sizes. In addition to the effects of the TCP send window, we will also show the effects of the operating system environment (for instance process scheduling) and the fragmentation of the different TSDU sizes to the ATM AAL5 interface MTU on the throughput.

The comparison of these results on ATM WAN with the tests on LAN is done using identical test scenarios and workstation environments (see the descriptions).

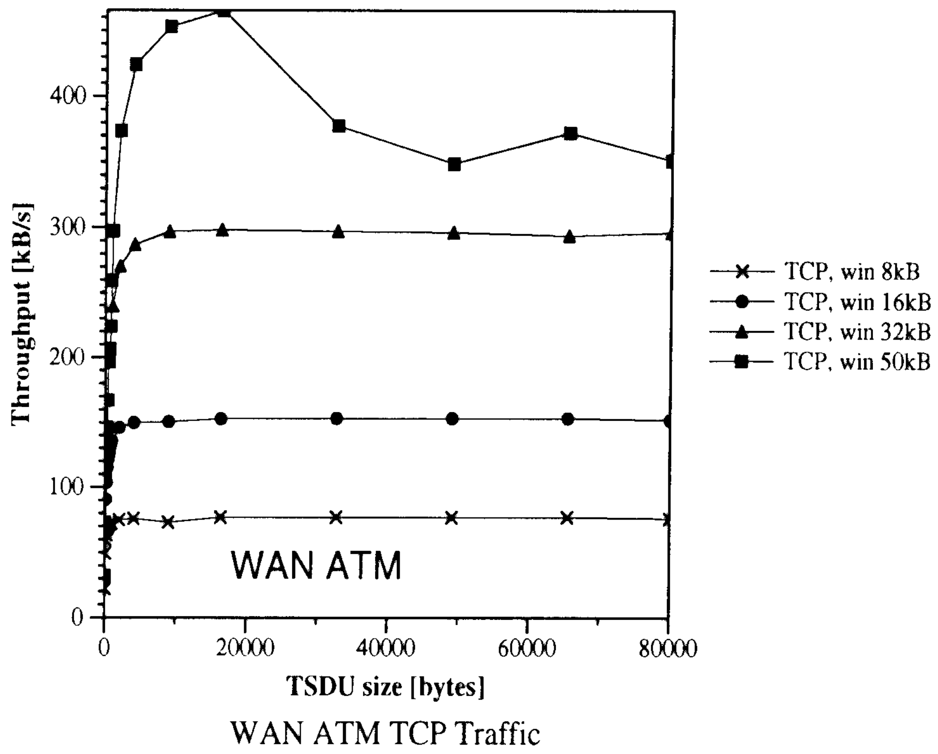


Figure 12 - TCP Send Window and its Impact on Throughput for Different TSDU Sizes on ATM WAN

Description.	Plot:	Throughput vs. TSDU size.
	Tool:	`xtcp`
	Transfer:	bulk transfer
	Protocol:	TCP/IP
	Window:	4 KB, 8 KB, 16 KB, 32 KB, 50 KB
Traffic.	TSDU size:	const
	interarrival time:	const
Network.	ATM WAN:	Transatlantic ATM Link
	ATM Link:	T3 (45 Mb/s) slowest
	ATM connection:	ATM PVC in both directions: 10 Mb/s, CBR
	Adaptation Layer:	ATM AAL5, MTU 9188, fa0
	RTT:	100 ms
Workstation.	Source host:	dune, Sparc-10, DeTeBerkom
	Source OS:	SunOS-4.1.3
	Dest. host:	multi, Sparc-10, CRC, Ottawa
	Dest. OS:	SunOS-4.1.4
Details.	Script:	shell script, `bin/tsdus.csh`
	TSDU sizes Values:	see table in the appendix.

Discussion of the results.

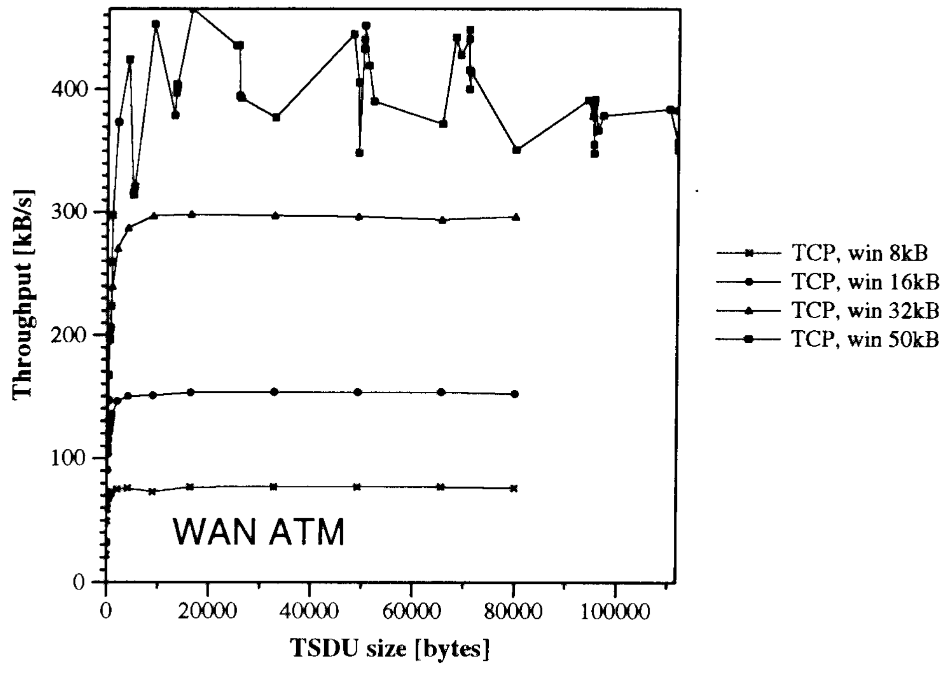
- **ATM WAN effect of high delay on the throughput provision.**
The TCP send window size can be considered a flow control parameter that provides throughput requirements to the user. With different window sizes, different levels of throughput are achieved as shown in Figure 12. The window size does not provide a direct mapping to the throughput predictions because it is based on the feedback mechanism of the ARQ protocol, i.e. the RTT for which an acknowledgment is returned: rate = window / RTT.

- **System scheduling and API effect.**
The system scheduling and the application programming interface determine the interaction of the operating system with the protocol. If this interaction is not properly tuned, the system scheduling can create a long delay before the TSDU is ready for transmission. This can cause large throughput deviations for some TSDU sizes (as is shown for the 50 KB window: up to 100 KB/s). The effect of the system scheduling and API can be improved when upgraded families of operating system, ATM interface cards and workstations are used. For instance, the corresponding table in the appendix emphasises the TSDU values for which there are significant throughput deviations from the predicted values. As shown in Figure 12, system scheduling does not represent a significant overhead on the transport service's throughput provision when small window sizes are used.

- **Throughput deviations with increasing window size**
TCP RFC 1323 implementation allows the use of larger window sizes and hence a larger portion of the bandwidth can be utilized. The use of large window sizes can lead to significant throughput deviations as experienced in this scenario for the case of TCP 50 KB send window size.

- **Future Investigations**
Further studies could be undertaken to verify if a more accurate tuning of the API, window size and the system architecture improves the segment flow.

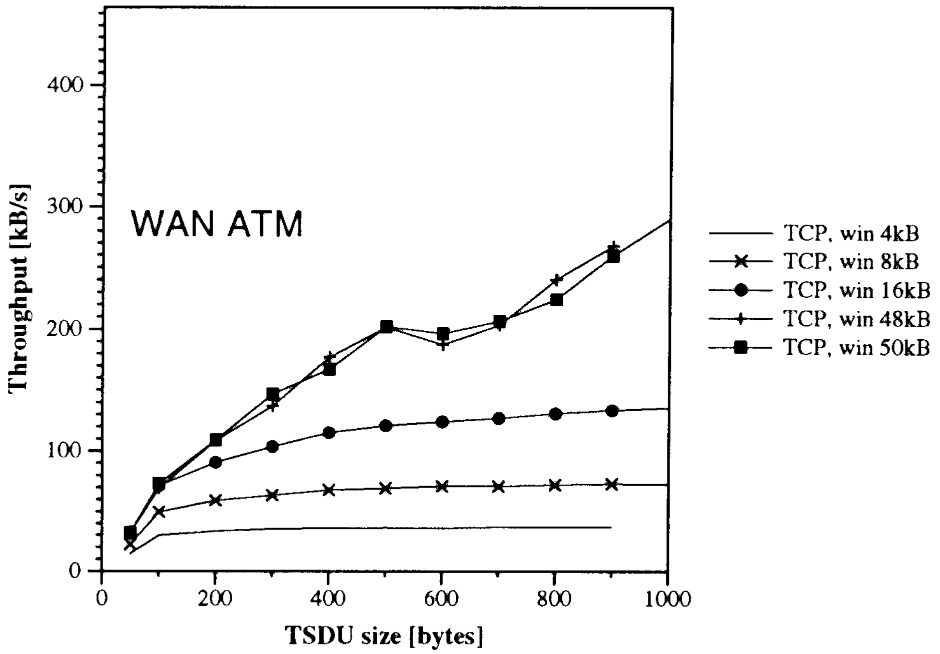
The next figure, together with the corresponding table in the Appendix, shows more precisely the throughput behaviour in the range of TSDU sizes that give significant throughput deviations. Additional measuring points were taken to show the fluctuations that occur when a window size of 50 KB is selected.



WAN ATM TCP Traffic

Figure 13 - Influence of the TCP Flow Control on the Throughput of Bulk Data Application for Specific TSDU Sizes on ATM WAN

In addition to the transport flow control mechanisms and the system architecture, certain application traffic characteristics can influence the throughput provision of applications over WAN ATM. The results of the measurements taken to investigate this are shown in Figure 14 where only small TSDU sizes have been evaluated.



WAN ATM TCP Traffic

Figure 14 - Impact of TCP Flow Control on the Throughput for Small TSDU Sizes on ATM WAN

Description.	Plot:	Throughput vs. TSDU size, small TSDU size
	Tool:	`xtcp`
	Transfer:	bulk data
	Protocol:	TCP
	Window:	4 KB, 8 KB, 16 KB, 32 KB, 48 KB, 50 KB
Traffic.	TSDU size:	const, small TSDU sizes
	interarrival time:	const
Network.	ATM WAN:	Transatlantic ATM link
	ATM Link:	T3 (45 Mb/s) slowest
	ATM Connection:	PVC in both directions: 10 Mb/s, CBR
	Adaptation Layer:	TM AAL5, MTU 9188, fa0
	RTT:	100 ms
Workstation.	Source host:	dune.bali.de, Sparc-10, DeTeBerkom
	Source OS:	SunOS-4.1.3
	Dest. host:	multi, Sparc-10, CRC/Ottawa
	Dest. OS:	SunOS-4.1.4

Discussion of the results.

- Specific traffic pattern, i.e. small TSDU sizes, on high delay ATM WAN.

The system overhead becomes more significant when small TSDU sizes are transmitted. Predictable throughput, using specific window or rate size at the transport layer, cannot be achieved for small TSDU sizes due to the system overhead and insufficient use of the ATM bandwidth (AAL 5 PDU size). If the TSDU size is small, the ATM AAL PDU sizes are not efficiently used to carry user traffic which causes throughput degradation in addition to significant system overhead. With small TSDU sizes we cannot achieve the optimal throughput for a specific window. We therefore propose that smaller TCP window sizes be used in order to use the ATM resources efficiently.

7.1.2 Transport flow control and its Impact on the Throughput in ATM LAN

Goal. The aim of these measurements is to show the influence of the TCP flow control parameter “send window size” on the throughput of applications on ATM LAN. Different TSDU sizes are considered and the throughput stability and variances are investigated.

In addition to the effect of the TCP send window, we also show the effect of the operating system environment (process scheduling) and the fragmentation of the different TSDU sizes to the ATM AAL5 MTU size.

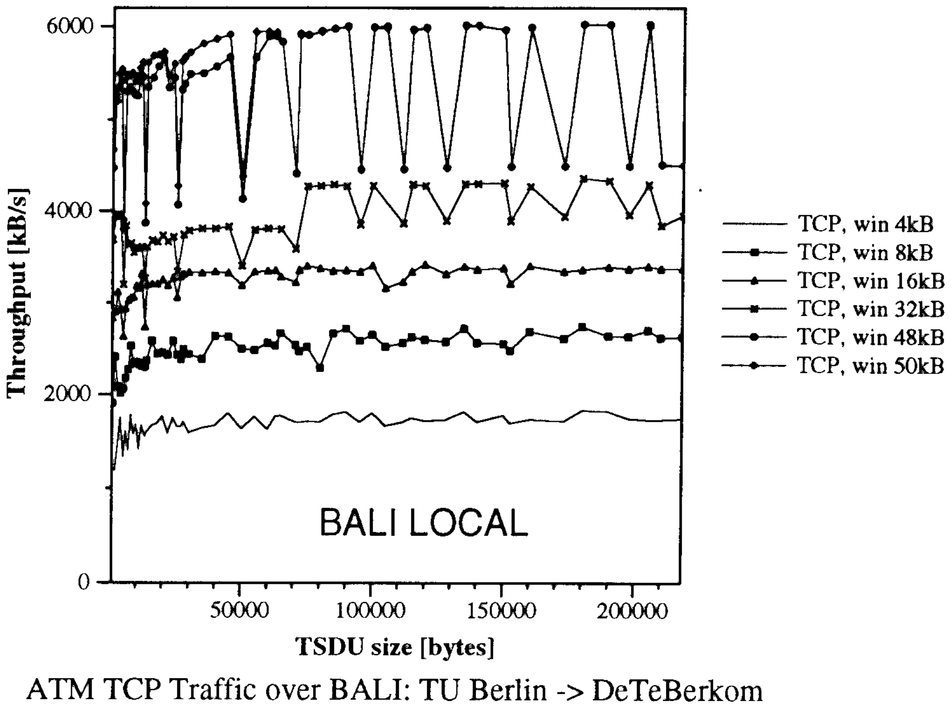


***Figure 15** - TCP Send Window and its Impact on Throughput for Different TSDU Sizes on ATM LAN*

Description.
Plot: Throughput vs. TSDU size
Tool: `xtcp`
Transfer: bulk transfer
Protocol: TCP/IP
Window: 4 KB, 8 KB, 16 KB, 32 KB, 50 KB

Traffic.
TSDU size const
interarrival time: const

Network.
ATM LAN: BALI (Berlin ATM LAN Initiative)
ATM Link: STM-1 (155 Mb/s)
ATM connection: ATM PVC in both directions: 155 Mb/s, CBR
Adaptation Layer: ATM AAL5, MTU 9180, qaa0
RTT: 2 - 4 ms

Workstation.
Source host: obelix.bali.de, Sparc-10, FSP-PV, PRZ, TU-Berlin
Source OS: SunOS-4.1.3
Dest. host: dune.bali.de, Sparc-10, DeTeBerkom
Dest. OS: SunOS-4.1.3

Details.
Script: shell script, `bin/bali_tsdus.csh`
TSDU sizes: see Appendix

Discussion of the results.

- **Provision of user throughput requirements using the send window size**

In comparison with the throughput provision on the ATM WAN, there is a significant improvement of the throughput on the ATM LAN for the same window sizes. Due to the lower delay on ATM LANs, the throughput achieved is greater than the throughput for corresponding window sizes on ATM WAN. Even though there is a difference in the bandwidth of the ATM PVC connections on LAN (155 Mb/s) and WAN (10 Mb/s), the same results would have been obtained if a larger bandwidth had been reserved on the ATM WAN since the window size is the bottleneck in the WAN measurements. On ATM LANs, the effects of certain congestion mechanisms like Slowstart and Congestion Control Avoidance seem negligible when large volumes of data are transferred that require no retransmissions.

Even when the delay provided by the ATM LAN is stable, the RTT depends on the time the acknowledgments arrives at the sender. This time in turn depends heavily on the system environment which influences the throughput prediction, as shown in Figure 15.

- **System and network configuration.**

The throughput deviations that were experienced with an older version of the Sun operating system SunOS and ATM FORE cards for the same TSDU values are reported in [27]. The deviations are more severe with the older version and therefore demonstrate that the architecture of the operating system and network interface cards are performance factors that can affect the throughput considerably.

Figure 16 and the corresponding table in the Appendix show more precisely the throughput behaviour for certain TSDU sizes giving significant throughput deviations. Additional measurement points were taken to emphasize the throughput deviations.

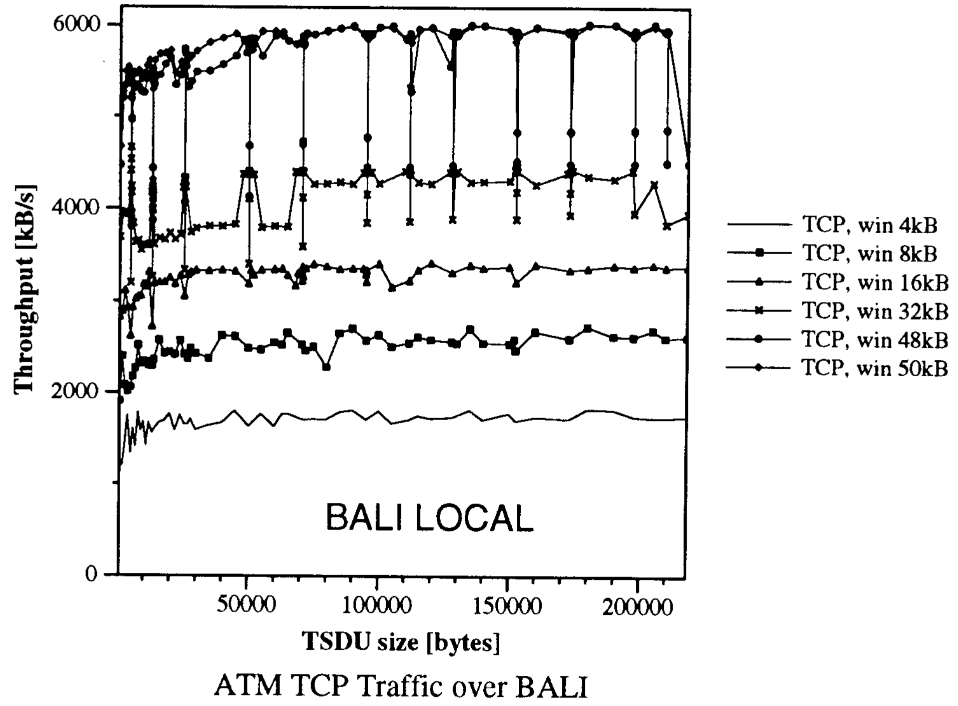


Figure 16 - Impact of TCP Flow Control on the Throughput of Bulk Data Application for Specific TSDU Sizes on ATM LAN

Again, in addition to the transport flow control mechanisms and system architecture, application traffic can influence the provision of throughput for applications over ATM LAN. This is shown in Figure 17 for small TSDU sizes.

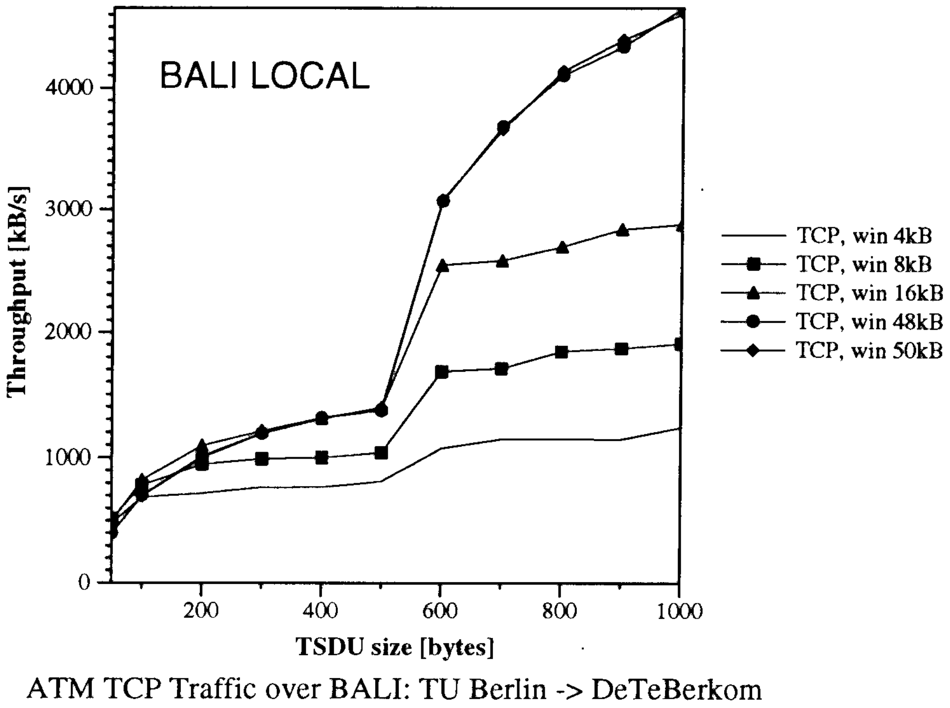


***Figure 17** - Influence of TCP Flow Control on the Throughput for Small TSDU Sizes on ATM LAN*

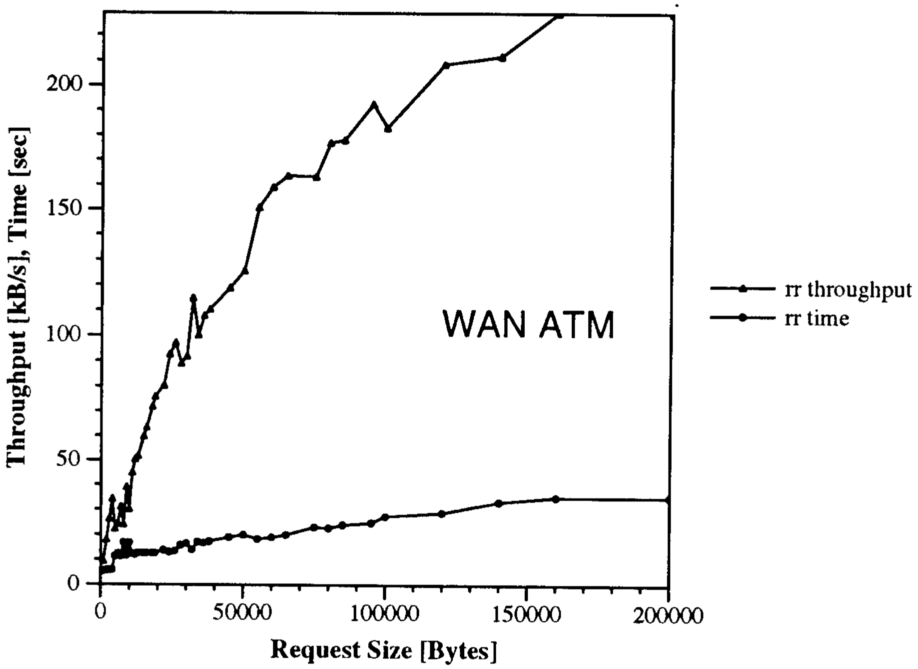
Description.	Plot:	Throughput vs. TSDU size, small TSDU sizes.
	Tool:	`xtcp`
	Transfer:	bulk transfer
	Protocol:	TCP/IP
	Window:	4 KB, 8 KB, 16 KB, 32 KB, 50 KB
Traffic.	TSDU:	const, small TSDU sizes
	interarrival time:	const
Network.	ATM LAN:	BALI (Berlin ATM LAN Initiative)
	ATM Link:	STM-1 (155 Mb/s)
	ATM connection:	ATM PVC in both directions: 155 Mb/s, CBR
	Adaptation Layer:	ATM AAL5, MTU 9180, qaa0
	RTT:	2 - 4 ms
Workstation.	Source host:	obelix.bali.de, Sparc-10, FSP-PV, PRZ, TU-Berlin
	Source OS:	SunOS-4.1.3
	Dest. host:	dune.bali.de, Sparc-10, DeTeBerkom
	Dest. OS:	SunOS-4.1.4
Details.	Script:	shell script, `bin/bali_tsdus.csh`
	TSDU sizes:	see table in the Appendix

Discussion of the results.

- As for the ATM WAN, the system overhead becomes very significant when small TSDU sizes are transmitted. In order to utilize the ATM resource efficiently, we believe that it is necessary to use appropriate flow control parameters, i.e. window size or rate, that can adapt to the traffic pattern and time requirements of the operating system.

7.2 Impact of TCP Flow and Congestion Control on Delay Sensitive Applications on ATM WAN

Goal. The aim of these measurements is to visualize the response time behaviour of transaction applications for different TSDU sizes on the WAN network. Such applications are bursty and delay sensitive. Mechanisms such as Slowstart increase the delay of transaction applications that use small TSDU sizes, e.g. banking requests, airline requests etc., especially on WAN links where the delay is the dominating factor.



Request-Response Throughput and Time

Figure 18 - Influence of TCP Specific Slowstart Mechanism on the Response Time of Transaction Applications Considering Different Request TSDU Sizes

Description.	Plot:	Request-Response throughput and time vs. TSDU size of Request
	Tool:	SPIMS
	Transfer:	request-response
	Protocol:	TCP/IP
	Window:	32 KB

Traffic.	Request TSDU:	1000 - 20000 with step 1000, 22000 - 40000 with step 2000, 45000 - 100000 with step 5000
	Response TSDU:	1024 constant for all requests

Network.	ATM WAN:	Transatlantic ATM Link
	ATM Link:	T3 (45 Mb/s) slowest
	ATM connection:	ATM PVC in both directions: 10 Mb/s, CBR
	Adaptation Layer:	ATM AAL5, MTU 9188, fa0
	RTT:	100 ms

Workstation.	Source host:	dune, Sparc-10, DeTeBerkom
	Source OS:	SunOS-4.1.3
	Dest. host:	multi, Sparc-10, CRC, Ottawa
	Dest. OS:	SunOS-4.1.4

Details. See Appendix.

Discussion of results.

The behaviour of the request throughput and response time are scaled and put as y-coordinate in the same diagram in order to show their relationships and the influence of TCP flow control and congestion avoidance mechanisms (i.e. Slowstart) on the QoS of transaction applications in the ATM WAN environment. A request-response application type was used to measure the two-way delay between the communicating entities. The request TSDUs vary in size while the response TSDU is of constant size. The results were collected in the following way: the TSDU size for one transmission or measurement step (represented by the x-axis of the plots) was specified. One measurement run was performed by iterating the measurement step a certain number of times and collecting statistical information such as the total amount of time (in seconds) it takes to transmit a packet of TSDU size a number of times. To obtain the throughput or the time (represented by the y-axis of the plots), several measurement runs were performed and the mean values calculated.

Slowstart and Request Response QoS Provision on ATM WAN.

Figure 18 shows that before the Slowstart mechanism reaches the 32 KB TCP send window, the response time increases quickly (i.e. exponentially over RTT). For instance, to transfer the first 32 KB, the response time increases from 5.3 to 16.3 sec. However for each successive 32 KB sent, the response time increases slightly, approximately two times slower, i.e. 5 sec. Transactions with smaller request sizes than 32 KB can often be found in different communication systems such as banking, data bases etc. For these type of applications there is a significant performance loss in the response time because of the long time it takes for the Slowstart to reach the optimal window size when a connection is established. This is due to the large RTT on high delay ATM WANs. We are incline to think that even if sufficient bandwidth for restricted time transaction applications is provided, the user QoS requirements for the response time will not be provided efficiently by the Slowstart mechanism.

Figure 19 shows a comparable scenario on ATM LAN.

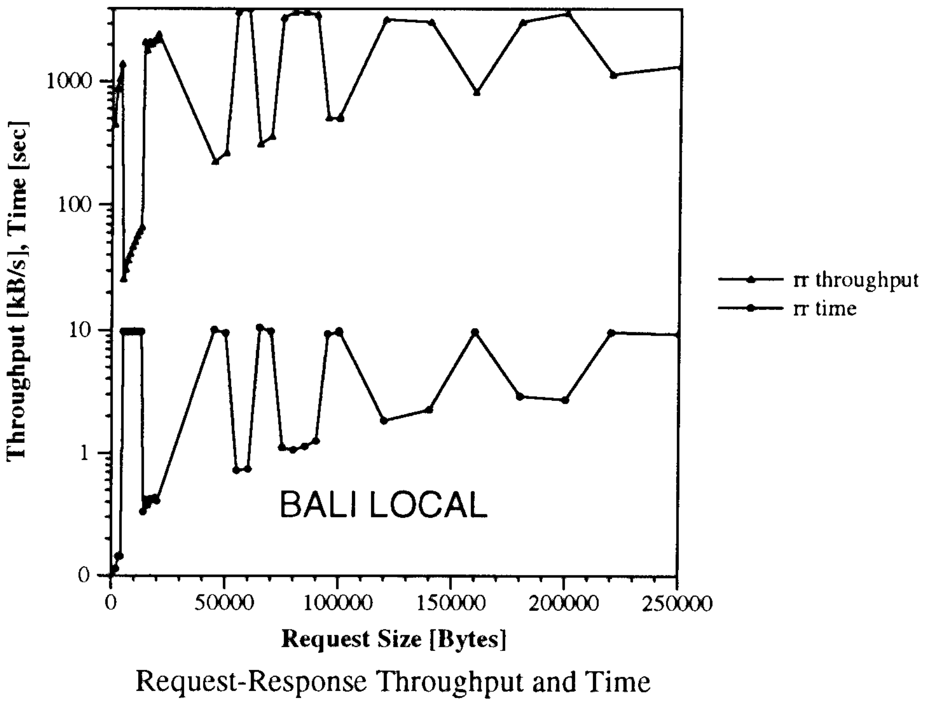


***Figure 19** - Impact of TCP Specific Slowstart Meachnisms on the Response Time of Transaction Oriented Applications on ATM LAN*

Description.	Plot:	Request-Response throughput and time vs. TSDU size of Request
	Tool:	SPIMS
	Transfer:	request-response
	Protocol:	TCP/IP

Window: 32 KB

Traffic. Request TSDU: 1000 - 20000 with step 1000,
22000 - 40000 with step 2000,
45000 - 100000 with step 5000
Response TSDU: 1024 constant for all requests

Network. ATM LAN: Berlin ATM LAN Initiative (BALI)
ATM Link: STM-1 (155 Mb/s)
ATM connection: ATM PVC in both directions: 155 Mb/s, CBR
Adaptation Layer: ATM AAL5, MTU 9180, qaa0
RTT 2 - 4 ms

Workstation. Source host: obelix.bali.de, Sparc-10, FSP-PV, PRZ, TU-Berlin
Source OS: SunOS-4.1.3
Dest. host: dune.bali.de, Sparc-10, DeTeBerkom
Dest. OS: SunOS-4.1.3

Details. See Appendix.

Discussion of results.

Scheduling and Request Response QoS Provision on ATM LAN.
Figure 19 shows that the provision of the response time QoS on ATM LAN does not seem to depend on the specific window size or Slowstart mechanisms as on ATM WAN, but seems to depend more on system scheduling. We believe that the illustrated "response time deviations" behaviour is caused by the scheduling of the transaction process and the end system overhead.

8 Performance Analysis of Application Admission Control on ATM Networks

There are different admission control policies that multiplex, i.e. bundle, connections over ATM networks for the efficient use of ATM resources. Depending on the traffic and QoS requirements selected by the user, different policies for bundling connections are used:

- Unidirectional connections of the same application class. In this case, the connections can be bundled on the same ATM channel. Performance is dependent on the ATM resource reservation and the impact of the workstation architecture;
- Bidirectional connections of the same application class. Although different ATM channels are used, the number of bundled connections and the performance still depend on resource sharing at the intermediate or end system architectures;
- Connections of different application classes with emphasis on the behaviour of QoS parameters characterizing the individual applications.

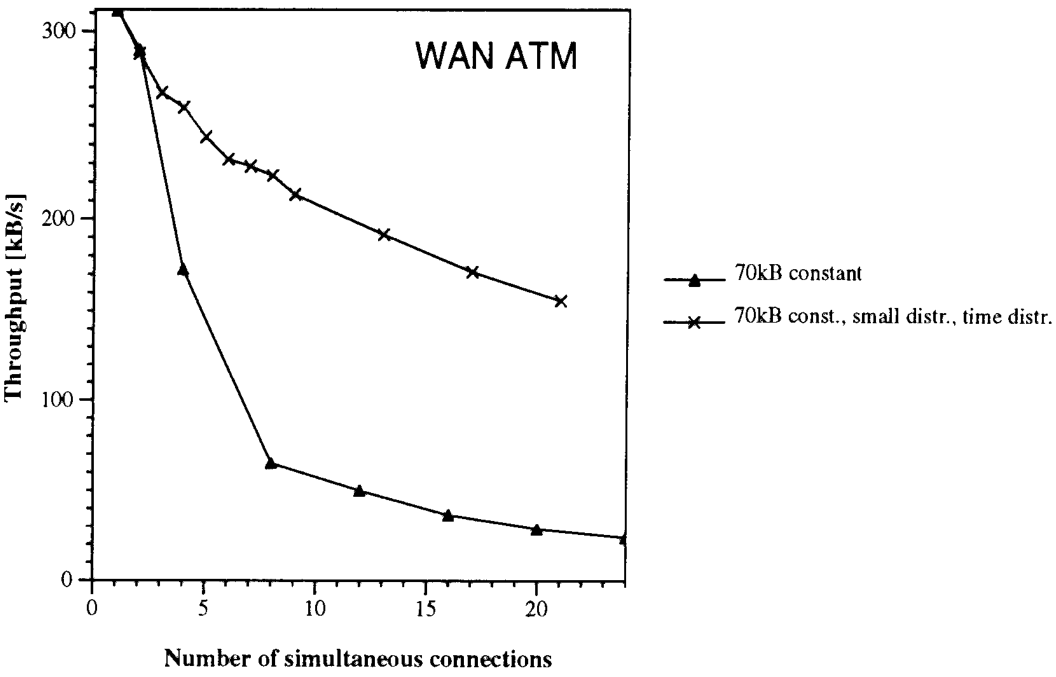
To achieve the QoS requirements (for instance thresholds for throughput and request response time) and to use the ATM resources in a cost-efficient way (i.e. to bundle the traffic over ATM connections in a optimal manner), the user must consider the impact of different factors on the performance. The focus of this section is to present practical oriented scenarios based on different bundling concepts such as unidiretional, bidirectional and bundling of connections with different traffic characterisctics and QoS requirements.

8.1 Unidirectional Bundling of Transport Connections

Goal. When unidirectional reliable connections are bundled, the data that flows in one direction and the acknowledgments that flow in the other direction are multiplexed independently, i.e. the mapping of resources can be done in a uniform way for the two directions. Performance characteristics of such bundles depend on the application traffic in one direction and the acknowledgment traffic in the other. Scheduling is also a performance factor, because each data flow is connected to its own process which is scheduled by the operating system. The specific QoS requirements impact the performance too; for instance, some connections are assigned a different window and/or rate than others. In this case the end performance depends on the interaction of system scheduling and flow control mapping.

8.1.1 Bundling of a Traffic-Intensive Connection with Connections Having Different Traffic Characteristics

Goal. These measurements show the throughput behaviour of a traffic-intensive connection (constant TSDU size and TSDU interarrival time) that is bundled with simultaneous connections that have different traffic characteristics. A traffic-intensive connection has more volume passing through it in a given time period than a less intensive connection. The traffic volume that goes through the other simultaneous connections varies according to the nature of the bivalue distribution (see the Appendix for more information on the distributions provided with SPIMS). This causes different throughput behaviour for the intensive traffic connection as the number of connections increase. The same window size is used in these measurements.



WAN TCP: Bundeled, unidirectional.

Figure 20 - Impact of the Unidirectional Bundling of Connections on a Throughput of Connections with Constant TSDU and Interarrival Times on ATM WAN

Description.	Plot:	Throughput vs. number of bundled unidirectional connections.
	Tool:	SPIMS
	Transfer:	bulk_get
	Protocol:	TCP/IP
	Window:	32 KB
Traffic.	1st curve:	All connections are simultaneous unidirectional with TSDU size of 70 KB constant, no TSDU interarrival time, homogeneous traffic. Throughput of one connection plotted.
	2nd curve:	Simultaneous unidirectional connections: 1st connection TSDU size 70 KB, all others TSDU size bivalue(10,9000,0.9) distributed and TSDU interarrival time of exp(0.2) distributed, inhomogeneous traffic. Throughput of 1st connection plotted.
Network.	ATM WAN:	Transatlantic ATM Link
	ATM Link:	T3 (45 Mb/s) slowest
	ATM connection:	ATM PVC in both directions: 10 Mb/s, CBR

Adaptation Layer: ATM AAL5, MTU 9188, fa0
RTT: 100 ms

Workstation. Source host: dune, Sparc-10, DeTeBerkom
Source OS: SunOS-4.1.3
Dest. host: multi, Sparc-10, CRC, Ottawa
Dest. OS: SunOS-4.1.4

Details. See Appendix.

Discussion of results

- Impact of traffic characteristics of simultaneous connections

The results show the impact of the traffic characteristics on the performance of unidirectional connections. When the traffic consists of a number of connections and all the connections share the same traffic characteristics (case 1: 70 KB TSDU size), the performance decreases rapidly and in the same degree for all connections. However, if one traffic-intensive connection is bundled with connections of lower traffic intensity (case 2), the performance of the traffic-intensive connection decreases much more slowly. Although the same flow control parameters are being used for both measurements, it is obvious that when we increase the number of simultaneous connections, the throughput in the case of less intensive traffic (2) decreases slower than in the case of connections with intensive traffic (1). This is due to the lower mean and peak traffic rate in case (2) which allow the simultaneous execution of more connections without modifying the QoS of the traffic intensive connection. This is not unexpected since the TSDU interarrival times are exponentially distributed and cause a lower traffic rate, leaving larger throughput for other connections.

- Impact of the high delay

The throughput behaviour in case (1) also shows the impact of the high delay which restricts the throughput on ATM WAN. For instance, the throughput when there are two simultaneously connections in case (1) is pretty similar to the throughput when there is only one connection. This behaviour can be compared with the throughput on ATM LAN where the network delay is no more a bottleneck. We can also observe that as the number of simultaneous connections increases, the load on the network increases and the effect of the high delay restricts the possible throughput.

Figure 21 illustrates a comparable case of bundling of connections on ATM LAN.

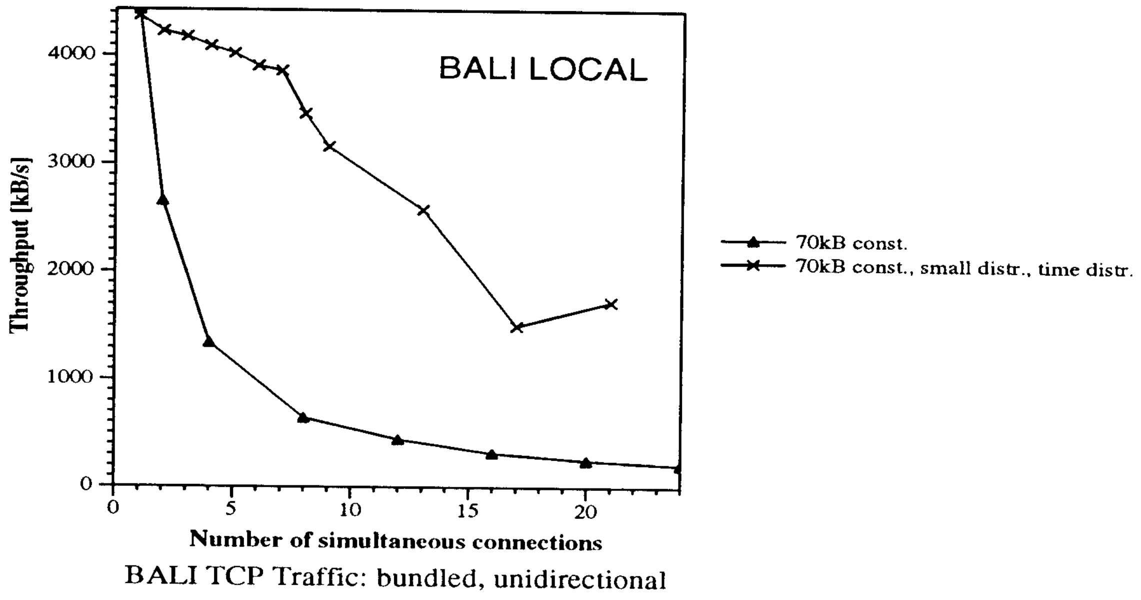


Figure 21 - Impact of Unidirectional Bundling of Connections on the Throughput of Connections with Constant TSDU and Interarrival Times on ATM LAN

Description.
Plot: Throughput vs. number of bundled unidirectional connections.
Tool: `SPIMS`
Transfer: bulk_get
Protocol: TCP/IP
Window: 32 KB

Traffic.
1st curve: All connections are simultaneous unidirectional with TSDU size of 70 KB constant, no TSDU interarrival time, homogeneous traffic.
Throughput of one connection plotted.
2nd curve: Simultaneous unidirectional connections:
1st connection TSDU size 70 KB, no TSDU interarrival time
all others TSDU size bivalue (10,9000,0.9) distributed and TSDU interarrival time of exp (0.2) distributed, inhomogeneous traffic.
Throughput of 1st connection plotted.

Network.
ATM LAN: Berlin ATM LAN Initiative (BALI)
ATM Link: STM-1 (155 Mb/s)
ATM connection: ATM PVC in both directions: 155 Mb/s, CBR
Adaptation Layer: ATM AAL5, MTU 9180, qaa0
RTT: 2 - 4 ms

Workstation.
Source host: obelix.bali.de, Sparc-10, FSP-PV, PRZ, TU-Berlin
Source OS: SunOS-4.1.3
Dest. host: dune.bali.de, Sparc-10, DeTeBerkom
Dest. OS: SunOS-4.1.3

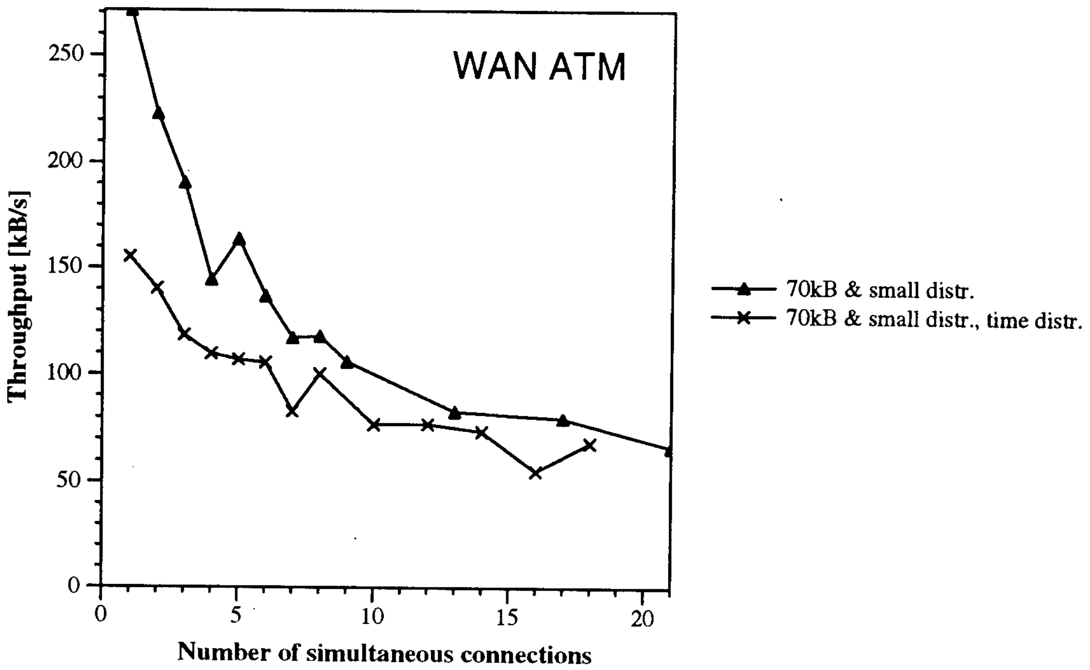
Details. See Appendix.

Discussion of results.

- The throughput of bundled connections on ATM LAN depends more significantly on the traffic than on the ATM WAN. This might be a consequence of the delay variations (2 to 4 ms) that influence the performance on the LANs. If there is low-intensity traffic on the simultaneous connections, then the throughput of a given high-intensity traffic connection is not affected as much as shown in Figure 21.

8.1.2 Bundling of Connections of Different Traffic Characteristics

Goal. This scenario shows the throughput behaviour in the case of unidirectional bundling of connections with different traffic characteristics. The prediction of the throughput of connections with varying traffic characteristics depends on the variations of the traffic intensity, i.e. on the ratio of maximum and minimum TSDU bursts and their durations.



WAN TCP: Bundeled, unidirectional.

Figure 22 - Impact of Unidirectional Bundling of Connections with Different Traffic Intensity on the Throughput in ATM WAN

Description.	Plot:	Throughput vs. number of bundled unidirectional connections.
	Tool:	SPIMS
	Transfer:	bulk_get
	Protocol:	TCP/IP
	Window:	32 KB
Traffic.	1st curve:	Simultaneous unidirectional connections: 1st connection TSDU size bivalue (100,70000, 0.75) distributed, no TSDU interarrival time, all others TSDU size bivalue (10,9000,0.75) distributed and TSDU interarrival time of exp (0.2) distributed. Throughput of 1st connection plotted.
	2nd curve:	Simultaneous unidirectional connections: 1st connection TSDU size bivalue (100,70000,0.75) distributed and TSDU interarrival time of exp (0.2) distributed, all others TSDU size bivalue (10,9000,0.75) distributed and TSDU interarrival time of exp (0.2) distributed. Throughput of 1st connection plotted.
Network.	ATM WAN:	Transatlantic ATM Link
	ATM Link:	T3 (45 Mb/s) slowest
	ATM connection:	ATM PVC in both directions: 10 Mb/s, CBR
	Adaptation Layer:	ATM AAL5, MTU 9188, fa0
	RTT:	100 ms
Workstation.	Source host:	dune, Sparc-10, DeTeBerkom
	Source OS:	SunOS-4.1.3
	Dest. host:	multi, Sparc-10, CRC, Ottawa
	Dest. OS:	SunOS-4.1.4

Details. See Appendix.

Discussion of results.

- **Deviations of the throughput due to different traffic intensity of the connections.**
  Even when the number of connections increases, the throughput remains the same or even increases if there are significant variations of the traffic load. The selected bivalue distribution can cause the TSDU size to vary from 100 to 70000 with 0.75 probability for the former value.
- **TSDU interarrival time distribution can significantly impact the traffic intensity.**
  The exponential distribution of the TSDU interarrival times, for the second plotted connections, results in lower throughput, than for the case where the TSDU interarrival times are not distributed, because the load on the network is not as great. However, as the number of simultaneous connections increases, the throughput differences resulting from the TSDU interarrival time distribution decreases. This might be due to the fact that as the number of bundled connections increases, additional connections have a lower impact on the overall performance, whereas with only a few simultaneous connections, a considerable amount of bandwidth is consumed by a new connection (i.e. if there are 2 simultaneous connections, each has half of the bandwidth; with 100 connections, each gets 1/100 of the bandwidth.)

Figure 23 shows the corresponding ATM LAN behaviour.

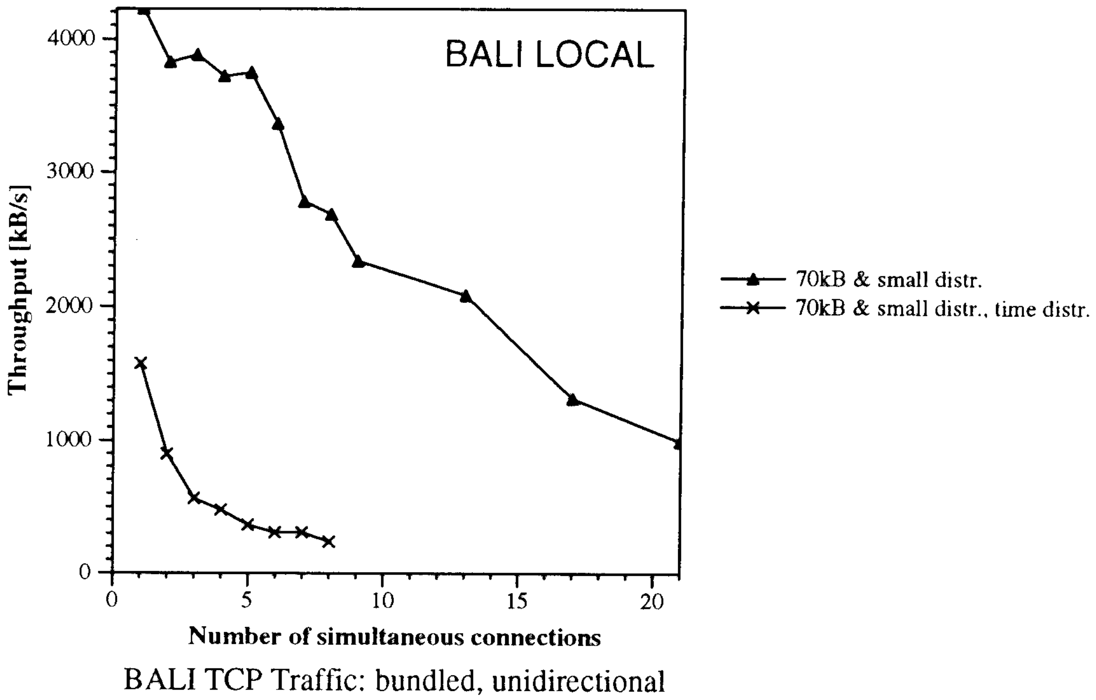


Figure 23 - Impact of Unidirectional Bundling of Connections with Different Traffic Intensity on the Throughput in ATM LAN

Description.	Plot:	Throughput vs. number of bundled unidirectional connections.
	Tool:	`SPIMS`
	Transfer:	bulk_get
	Protocol:	TCP/IP
	Window:	32 KB
Traffic.	1st curve:	Simultaneous unidirectional connections: 1st connection TSDU size bivalue (100,70000, 0.75) distributed, no TSDU interarrival time, all others bivalue (10,9000,0.75) distributed and TSDU interarrival time of exp (0.2) distributed. Throughput of 1st connection plotted.

2nd curve: Simultaneous unidirectional connections:
1st connection TSDU size bivalue (100,70000,0.75) distributed and TSDU interarrival time of exp (0.2) distributed,
all others bivalue (10,9000,0.75) distributed and TSDU interarrival time of exp (0.2) distributed.
Throughput of 1st connection plotted.

Network.

ATM LAN:	Berlin ATM LAN Initiative (BALI)
ATM Link:	STM-1 (155 Mb/s)
ATM connection:	ATM PVC in both directions: 155 Mb/s, CBR
Adaptation Layer:	ATM AAL5, MTU 9180, qaa0
RTT:	2 - 4 ms

Workstation.

Source host:	obelix.bali.de, Sparc-10, FSP-PV, PRZ, TU-Berlin
Source OS:	SunOS-4.1.3
Dest. host:	dune.bali.de, Sparc-10, DeTeBerkom
Dest. OS:	SunOS-4.1.3

Details. See Appendix.

Discussion of results.

TSDU size and traffic impact on the performance.

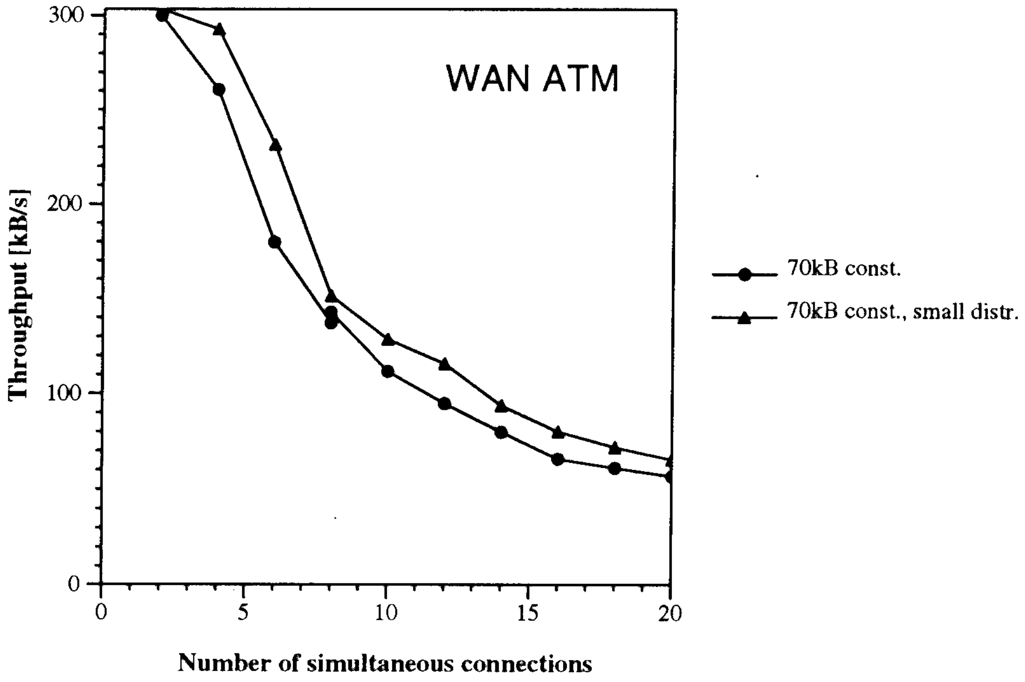
On the ATM LAN, it is shown that the TSDU size and traffic have a greater impact on the throughput of multiple unidirectional simultaneous connections than on the ATM WAN.

Throughput variations that are off the first curve are a result of the TSDU-size distribution. The inter-packet time distribution of the second curve leaves more bandwidth for other connections. A single connection doesn't impose a heavy load on the virtual circuit and takes a longer time to complete because of the inter-packet time which causes lower throughput.

8.2 Bidirectional Bundling of Transport Connections

Goal. There are a lot of practical scenarios where the bundling of bidirectional connections on the same ATM channel between two end systems is desirable for the efficient usage of ATM resources. For instance, when bidirectional connections belong to the same application or when the applications belong to the same organisation and must be charged together, etc.

This scenario demonstrates the performance characteristics when bidirectional connections are bundled and when the specifics of ATM networks are considered, i.e. the downstream and upstream ATM channels are independent in their QoS provision. Therefore the performance impact due to bidirectional connection bundling will be based on system configuration and process scheduling rather than on parameters such as traffic pattern or QoS requirement for the connection.



WAN ATM TCP Traffic

Figure 24 - Impact of Bidirectional Bundled Connections on Throughput in ATM WAN

Description.	Plot:	Throughput vs. number of inhomogeneous bundled bidirectional connections.
	Tool:	`SPIMS`
	Transfer:	bulk_get
	Protocol:	TCP/IP
	Window:	32 KB
Traffic.	1st curve:	Simultaneous homogeneous bidirectional connections are multiplexed on one VC: connection pairs of the form: 1. bulk_GET, TSDU 70 KB const., no TSDU interarrival time, 2. bulk_PUT, TSDU 70 KB const., no TSDU interarrival time. Throughput of the GET connection plotted.
	2nd curve:	Simultaneous non-homogeneous bidirectional connections are multiplexed on one VC: connection pairs of the form: 1. bulk_GET, TSDU 70 KB const., no TSDU interarrival time, 2. bulk_PUT, TSDU bivalue (10,9000,0.9) distributed, no TSDU interarrival time. Throughput of the GET connection plotted.
Network.	ATM WAN:	Transatlantic ATM Link
	ATM Link:	T3 (45 Mb/s) slowest
	ATM connection:	ATM PVC in both directions: 10 Mb/s, CBR
	Adaptation Layer:	ATM AAL5, MTU 9188, fa0
	RTT:	100 ms
Workstation.	Source host:	dune, Sparc-10, DeTeBerkom
	Source OS:	SunOS-4.1.3
	Dest. host:	multi, Sparc-10, CRC, Ottawa
	Dest. OS:	SunOS-4.1.4

Details. See Appendix.

Discussion of results.

- Impact of traffic intensity of connection pairs on ATM WAN

Even though different ATM channels are used, bidirectional connection bundling (i.e. connection pairs coming from opposite directions) affect the throughput when the number of simultaneous connection pairs increases. As Figure 24 illustrates, when the number of connection pairs increases, the incoming traffic-intensive connections consume more processing power (70 KB const.). Thus they impose a more rapid throughput decrease compared to incoming slower-traffic connections (10 or 9000 bytes with 0.9 probability).

Figure 25 shows the comparable scenario of bidirectional bundling on the ATM LAN

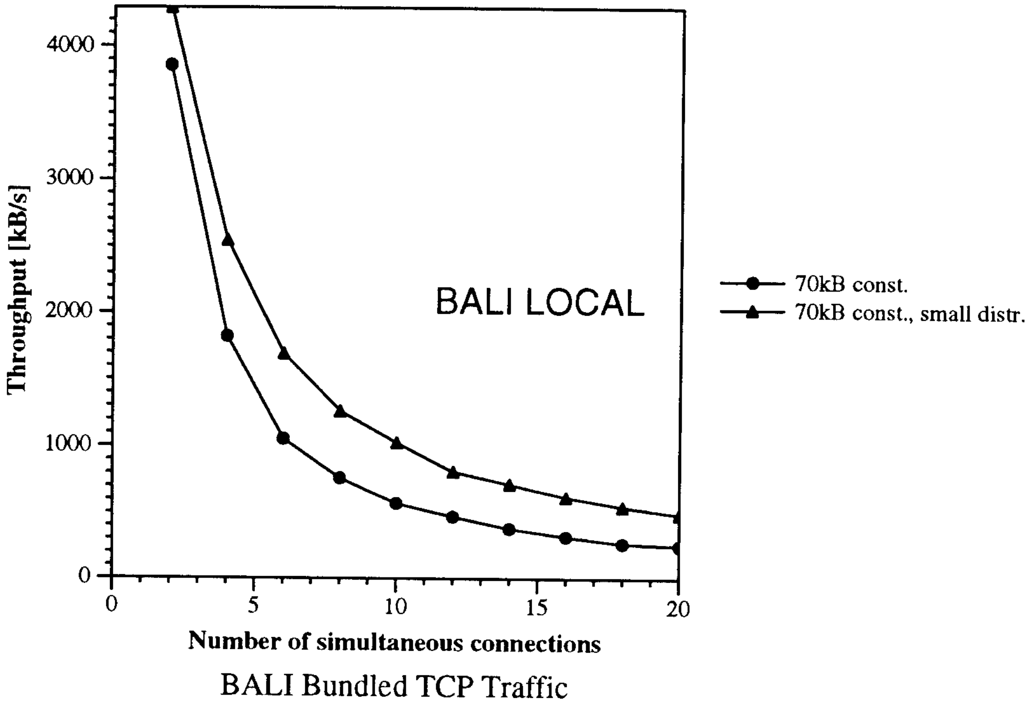


Figure 25 - Impact of Bidirectional Bundled Connections on Throughput in ATM LAN

Description.	Plot:	Throughput vs. number of inhomogeneous bundled bidirectional connections.
	Tool:	SPIMS
	Transfer:	bulk_get
	Protocol:	TCP/IP
	Window:	32 KB

Traffic. 1st curve: Simultaneous homogeneous bidirectional connections are multiplexed on one VC: connection pairs of the form:

1. bulk_GET, TSDU 70 KB const., no TSDU interarrival time,
2. bulk_PUT, TSDU 70 KB const., no TSDU interarrival time.

Throughput of the GET connection plotted.

2nd curve: Simultaneous non-homogeneous bidirectional connections are multiplexed on one VC: connection pairs of the form:

1. bulk_GET, TSDU 70 KB const., no TSDU interarrival time,
2. bulk_PUT, TSDU bivalue (10,9000,0.9) distributed, no TSDU interarrival time.

Throughput of the GET connection plotted.

Network. ATM LAN: Berlin ATM LAN Initiative (BALI)

ATM Link: STM-1 (155 Mb/s)
ATM connection: ATM PVC in both directions: 155 Mb/s, CBR
Adaptation Layer: ATM AAL5, MTU 9180, qaa0
RTT: 2 ms

Workstation. Source host: obelix.bali.de, Sparc-10, FSP-PV, PRZ, TU-Berlin
Source OS: SunOS-4.1.3
Dest. host: dune.bali.de, Sparc-10, DeTeBerkom
Dest. OS: SunOS-4.1.4

Details. See Appendix.

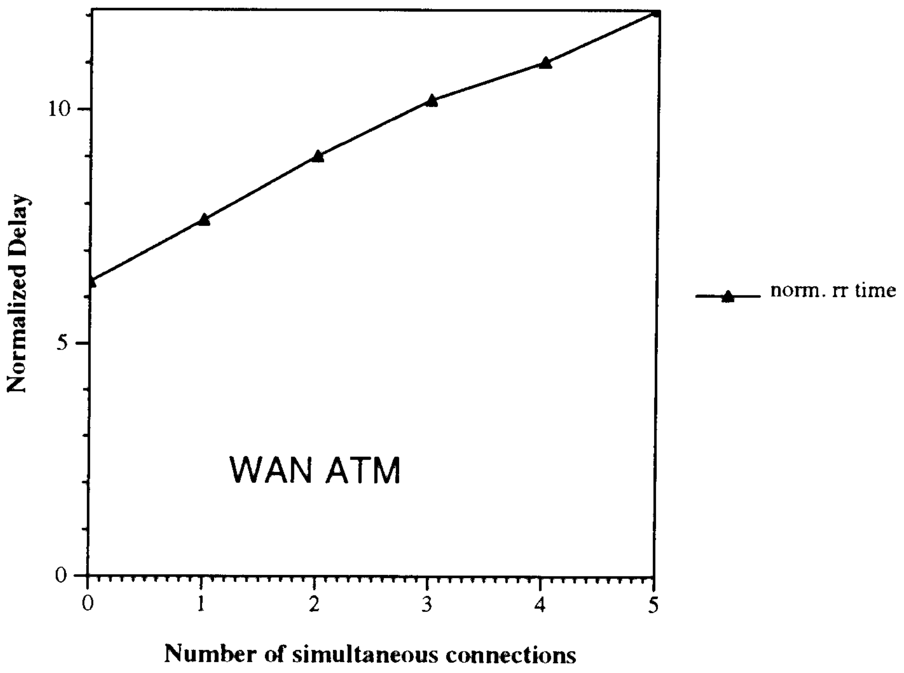
Discussion of results.

- TSDU size and traffic impact on the performance

In contrast to the ATM WAN, on the ATM LAN the impact of the same traffic, i.e. TSDU size, leads to a more rapid throughput decrease when the number of connection pairs are increased. We believe that the throughput drops faster on LANs than on WANs because the long delay is not there to smooth out the processing overhead.

8.3 Scenarios for bundling of Connections of Different Application Classes

Goal. Transport connections with different traffic characteristics and QoS objectives can be bundled. The requested throughput for each application must still be provided. This is the case when delay-sensitive applications such as transaction applications run simultaneously with bulk data applications. The response time of the transaction application depends on the characteristics and number of bundled bulk data applications.



Request-Response Times.

Figure 26 - Impact of Simultaneous Bulk Data Traffic on the Request-Response Time

Description. Plot: Normalized Request-Response time vs. number of simultaneous bulk connections
Tool: `SPIMS / xtcp`
Transfer: request-response / bulk

Protocol: TCP
Window: 50 KB

Traffic.
Request size: 70 KB constant
Response size: 1 KB constant
Simultaneous Connections: bulk transfer, TSDU size 9000, TCP window 50 KB.

Network.
ATM WAN: Transatlantic ATM link
ATM Link: T3 (45 Mb/s) slowest
ATM Connection: PVC in both directions, 10 Mb/s, CBR
Adaptation Layer: ATM AAL5, MTU 9180, fa0
RTT: 100 ms

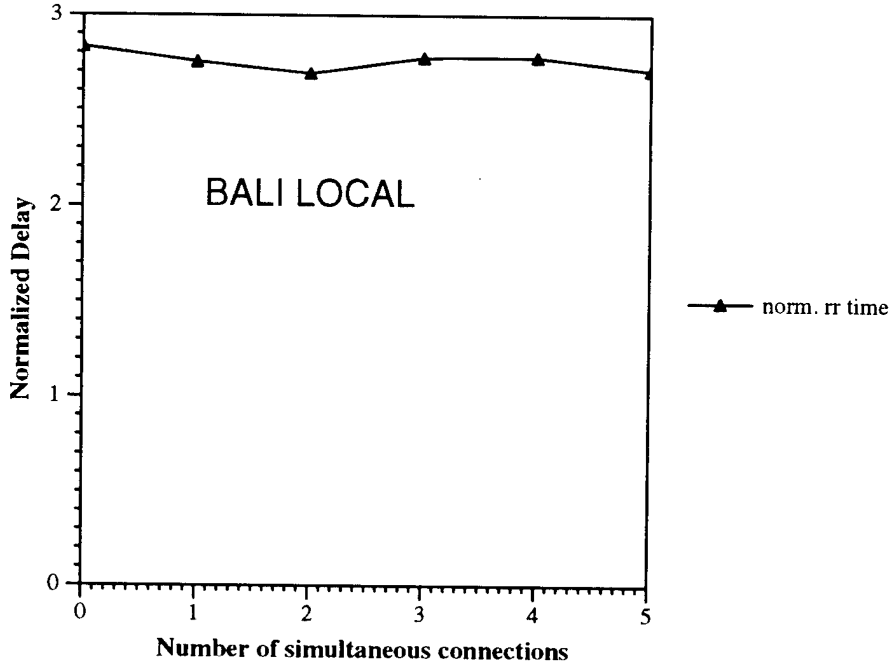
Workstation.
Source Host: dune.bali.de, Sparc-10, DeTeBerkom
Source OS: SunOS-4.1.3
Dest. Host: multi, Sparc-10, CRC/Ottawa
Dest. OS: SunOS-4.1.4

Details. See Appendix.

Discussion of results.

- **Transport flow control and high delay versus QoS provision on ATM WAN.**
  With the increasing number of bulk data connections, the response time increases accordingly due to the TCP flow control mechanism (window and Slowstart) and to the high delay (see Section 7.2 for comparison).

Figure 27 illustrates the same scenario on ATM LAN.



Request-Response Times.

Figure 27 - Impact of Simultaneous Bulk Data Traffic on the Request-Response Time, BALI

Description.
Plot: Normalized Request-Response time vs. number of simultaneous bulk connections
Tool: `SPIMS / xtcp`
Transfer: request-response / bulk

Protocol: TCP
Window: 50 KB

Traffic. Request size: 70 KB constant
Response size: 1 KB constant
Simultaneous Connections: bulk transfer, TSDU size 9000, TCP window 50 KB.

Network. ATM LAN: Berlin ATM LAN Initiative (BALI)
ATM Link: STM-1 (155 Mb/s)
ATM Connection: PVC in both directions, 155 Mb/s, CBR
Adaptation Layer: ATM AAL5, MTU 9180, qaa0
RTT: 2 - 4 ms/ xtcp

Workstation. Source Host: obelix.bali.de, Sparc-10, FSP-PV, PRZ, TU-Berlin
Source OS: SunOS-4.1.3
Dest. Host: dune.bali.de, Sparc-10, DeTeBerkom
Dest. OS: SunOS-4.1.3

Details. See Appendix.

Discussion of results.

- Operating system scheduling versus QoS provision by ATM LAN.

As the number of bulk data connections increases, there is no significant increase in the response time due to system scheduling. Only five simultaneous connections were bundled and therefore the bandwidth was not totally used.

9 Flow Control and Multiplexing Scenario Based on XTPX Protocol

The rate control mechanisms of XTP allows a direct mapping to the reserved bandwidth on the ATM connection. Figure 28 shows scenarios based on the XTPX protocol using different requirements for scalable throughput QoS. We measured the throughput of up to 12 simultaneous connections while varying the throughput QoS requirement - unrestricted rate, 500 KB/s, 200 KB/s and 20 KB/s.

Measurement Environment
Sender: Sun Sparc Station1 + (25 MHz, 15.8 MIPS)
Receiver: Sun Sparc Station 2 (40 MHz, 28.5 MIPS)
Network Interfaces: ATM FORE 2.1 Version, AAL 4
Measurement scenario
window = 29120 Bytes, recv buffer = 50 KB,
send buffer = 50 KB, TSDU length = 20 KB

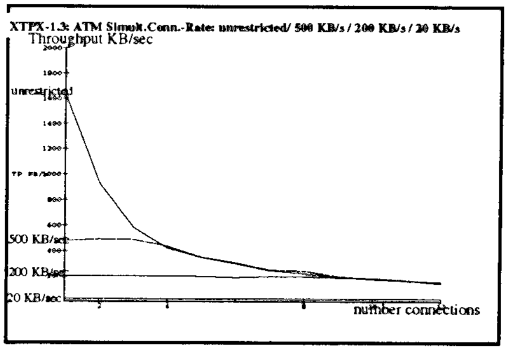


Figure 28 - Dependence of Throughput on the Number of Simultaneous Connections using XTPX

Using XTPX, the throughput QoS requirement can be scaled depending on the load, i.e. the number of simultaneous multimedia streams. For example when the maximum number of simultaneous connections is 12, the throughput of each stream is restricted to 200 KB/s. The throughput can be predicted very well because the rate control does not depend on the RTT.

Figure 28 comes from work previously done at the Technical University of Berlin on a local area ATM network with the XTPX protocol [76].

10 Conclusions

The focus of this work was to show how flow control and admission control mechanisms , based on TCP, affect the performance of reliable applications on ATM WAN and LAN.

For bulk data applications it was shown that:
- The impact of system scheduling on ATM LANs is very significant and can cause large throughput deviations. Even in the WAN environment, system scheduling can cause large throughput deviations when large window sizes are used;
- When small TSDUs are transmitted, the system overhead becomes more significant.

For delay-sensitive transaction applications it was shown that:
- Mechanisms such as Slowstart increase the delay of transaction applications that use small TSDU sizes, e.g. banking requests, airline requests, etc., especially on WAN links where the delay is the dominating factor;
- The response time on ATM LANs does not seem to depend as much on the specific window size or Slowstart mechanisms as on ATM WANs, but seems to depend more on system scheduling;
- The Slowstart mechanism takes a long time to reach the optimal window size, when a connection is established, due to the large RTT on high delay networks. Reservation protocols such as RSVP provide a way to reserve the necessary resources at connection establishment. To adapt TCP to ATM WAN it would be interesting to disable the Slowstart and Congestion Avoidance and use reservation protocols which would allow more efficient flow control policies for connections which have to provide a certain QoS.

Using practical experiments on transathlantic wide area ATM we have shown that when connections are bundled at the endsystem, the traffic of the connections as well as the network delay become an important performance factor. Using an adaptive scheme, i.e. by monitoring the throughput of a given bundle over time, an appropriate set of requirements for an ATM virtual circuit (for example peak and average rate) can be obtained. This allows for the efficient use of ATM WAN on information highway based on the prediction of the QoS provision for multiplexed traffic. It also allows an advance evaluation of the system configuration influence.

With TCP, it is more complicated to predict the behavior of the throughput when connections are bundled. The window size does not provide a direct mapping to the throughput predictions because it is based on the feedback mechanism of the ARQ protocol. With XTPX, the throughput can be predicted very well because the rate control does not depend on the RTT but rather on the burst "length" which specifies the maximum number of bytes allowed to be sent in a certain interval of time.

Practical usage of the RSVP implementation for bundling connections can be a future research topic.

11 Glossary

AAL	ATM Adaptation Layer
ABR	Available Bit Rate
ACR	Allowed Cell Rate
AIR	Additive Increase Rate
ARQ	Automatic Repeat Response
ATM	Asynchronous Transfer Mode
B-ICI	Broadband Inter-Carrier Interface
B-ISDN	Broadband Integrated Services Digital Network
BECN	Backward Explicit Congestion Notification
BT	Burst Tolerance
CAC	Connection Admission Control
CAPC	Congestion Avoidance with Proportional Control
CBR	Constant Bit Rate
CCR	Current Cell Rate
CDV	Cell Delay Variation
CDVT	CDV Tolerance
CER	Cell Error Ratio
CI	Congestion Indication (bit in RM cell)

CLP	Cell Loss Priority
CLR	Cell Loss Ratio
CMR	Cell Misinsertion Rate
CPE	Customer Premise Equipment
CTD	Cell Transfer Delay
DES	Destination End-System
DGCRA	Dynamic GCRA
EFCI	Explicit Forward Congestion Indication
ER	Explicit Rate
FECN	Forward Explicit Congestion Notification
FIFO	First In First Out Queue Service Discipline
FRS	Frame Relay Service
GCRA	Generic Cell Rate Algorithm
GFC	Generic Flow Control
IBT	Intrinsic Burst Tolerance
ICR	Initial Cell Rate
ILMI	Interim Local Management Interface
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ITT	Ideal Transmission Time
Kb	1024 bits
KB	1024 bytes
ITU	International Telecommunications Union
LAN	Local Area Network
LANE	LAN Emulation
MACR	Mean Allowed Cell Rate
MBS	Maximum Burst Size
MCR	Minimum Cell Rate
MCTD	Mean Cell Transfer Delay
MIB	Management Information Base
MP	Measurement Point
Multiprotocol	Referring to Internetworking Layer Protocols (including e.g. IP, DECNet, IPX, SNA, AppleTalk)
NNI	Network-to-Network Interface
OAM	Operations, Administration and Maintenance
PCR	Peak Cell Rate
PDU	Protocol Data Unit
PHY	Physical Layer Interface
PNNI	Private Network Node Interface
PTI	Payload Type Indicator
PVC	Permanent Virtual Connection
QoS	Quality of Service
RA	Resource Allocation (bit in RM cell)
RSVP	Resource Reservation Protocol
RM	Resource Management Cell
RTT	Round-Trip Time
SAA WG	ATM Forum Services Aspects and Applications Working Group
SAP	Service Access Point
SCR	Sustainable Cell Rate
SDU	Service Data Unit
SECBR	Severely Errored Cell Block Ratio
SES	Source End-System
SIG-WG	ATM Forum Signalling Working Group
SMDS	Switched Multimegabit Data Service
STD	Source Traffic Descriptor
ST-II	Internet Stream Protocol Version 2
ST+	Internet Stream Protocol Version 2+
SVC	Switched Virtual Connection
SW	Switch
TCP	Transport Control Protocol
TM	Traffic Management
UBR	Unspecified Bit Rate

UDP	User Datagram Protocol
UNI	User-Network Interface
UPC	Usage Parameter Control
VBR	Variable Bit Rate
VC	Virtual Connection
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VC-SW	Virtual Channel Switching Function
VPC	Virtual Path Connection
VPI	Virtual Path Identifier
VP-SW	Virtual Path Switching Function
WAN	Wide Area Network
XTP	eXpress Transfer Protocol
XTPX	eXpress Transfer Protocol eXtended
nrt-VBR	Non-Real-Time VBR
rt-VBR	Real-Time VBR

12 Appendix 1

12.1 Benchmark Descriptions

Scenario 7.1.1 - Benchmark Details

Figure 12:

TSDU sizes:	50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1024, 2048, 4096, 9000, 16384, 32768, 49152, 65536, 80000.
Dump-script:	Shell script (csh) calling xtcp with the respective parameters: tsdus.csh
Iterations:	1000
Runs:	10

Figure 13:

Same TSDU sizes as used in Figure 11 and for 50 KB window plus the following:

TSDU sizes: 5000, 5100, 5120, 5130, 5200, 5300, 13000, 13300, 13310, 13320, 13330, 13400, 13500, 25000, 25500, 25600, 25700, 26000, 48000, 49000, 50000, 50100, 50180, 50190, 50200, 51000, 52000, 68000, 69000, 70600, 70660, 70670, 70700, 71000, 94000, 95000, 95200, 95230, 95240, 95300, 96000, 97000, 110000, 111600, 111620, 111700

Figure 14:

TSDU sizes: 50, 100, 200, 300, 400, 500, 600, 700, 800, 900

Values: for TCP, window 50 KB, Figures 15, 16, 17

TSDU size	Throughput [KB/s]	Standard Deviation
50	32.2	0.84
100	73	3.43
200	108.9	3.81
300	146.7	6.09
400	167.1	36.6
500	202.1	6.04
600	196.4	3.17
700	206.4	32.31
800	224	55.04
900	259.5	58.90
1024	297.36	40.51
2048	373.6	2.27

Table 9 - WAN: Relationship of TSDU Sizes and Throughput

TSDU size	Throughput [KB/s]	Standard Deviation
4096	424.1	3.11
5000	315.9	6.28
5100	313.8	3.65
5120	315.7	5.38
5130	316.3	3.74
5200	320.5	5.25
5300	321.2	5.92
9000	452.8	1.14
13000	379.2	60.6
13300	397.4	2.31
13310	404.3	5.38
13320	401.5	2.99
13330	400.3	4.03
13400	400.6	3.41
13500	403.5	2.46
16384	465.2	0.63
25000	435.7	2.26
25500	436	2.71
25600	435.7	2.49
25700	394.8	88.48
26000	392.9	92.24
32768	377.4	104.85
48000	444.6	39.59
49000	405.7	85.58
49152	348.7	31.57
50000	433.1	52.09
50100	440.4	51.45
50180	432.2	61.22
50190	432.6	60.67
50200	451.7	1.34
51000	419.2	77.58
52000	390.5	89.21
65536	372.2	28.96
68000	442.2	45.76
69000	427.9	71.20
70600	440.8	39.24
70660	448.33	41.97
70670	415.9	67.38
70700	400.3	74.91
71000	413.9	75.78
80000	351.1	19.94
94000	391.33	52.25

Table 9 *- WAN: Relationship of TSDU Sizes and Throughput*

TSDU size	Throughput [KB/s]	Standard Deviation
95000	378.6	53.35
95200	355.5	81.57
95230	386.8	56.53
95240	348.2	52.42
95300	391.8	57.30
96000	366.9	70.03
97000	378.9	71.32
110000	384.1	60.94
111600	357	54.17
111620	383.3	64.58
111700	350.9	62.94

Table 9 - WAN: Relationship of TSDU Sizes and Throughput

Scenario 7.1.2 - Benchmark Details

Figure 15. TSDU sizes: 1000, 2000, 3000, 4000, 5125, 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13317, 14000, 16000, 18000, 20000, 22000, 24000, 25605, 27000, 28000, 30000, 35000, 40000, 45000, 50181, 55000, 60000, 63000, 65000, 70661, 72000, 75000, 80000, 85000, 90000, 95237, 100000, 105000, 111621, 115000, 120000, 128005, 135000, 140000, 150000, 152561, 160000, 173061, 180000, 190000, 197637, 205000, 209925, 218177.

Dump-script: Shell script (csh) calling `xtcp` with the respective parameters: `bali_tsdus.csh`
Iterations: 1000
Runs: 10
Values: for TCP window 50 KB

TSDU size	Throughput [KB/s]	Standard Deviation
1000	4673.74	22.33
2000	5267.94	61.81
3000	5489.03	52.18
4000	5538.70	52.06
5125	3902.50	101.85
6000	5480.66	36.48
7000	5465.48	44.61
8000	5500.06	18.42
9000	5405.17	100.22
10000	5464.89	80.86
11000	5549.82	84.20
12000	5611.24	21.031
13317	4081.26	42.63
14000	5610.53	63.76
16000	5681.88	20.12
18000	5693.70	48.40
20000	5722.65	60.27
22000	5471.35	54.32

TABLE 10. *BALI: Relationship of TSDU Sizes and Throughput*

TSDU size	Throughput [KB/s]	Standard Deviation
24000	5595.83	44.032
25605	4272.65	34.508
27000	5626.07	52.96
28000	5662.69	34.22
30000	5717.4	38.69
35000	5811.93	47.70
40000	5861.78	39.58
45000	5909.69	52.04
50181	4377.81	12.71
55000	5937.787	29.26
60000	5940.43	25.54
63000	5932.5	30.36
65000	5830.341	25.74
70661	4419.430	12.54
72000	5900.39	19.02
75000	5900.79	23.39
80000	5933.284	27.52
85000	5974.502	25.30
90000	5985.59	26.48
95237	4446.98	11.22
100000	5981.15	28.22
105000	5975.01	25.85
111621	4452.09	11.023
115000	5952.29	25.01
120000	5977.23	18.99
128005	4468.32	7.92
135000	6000.69	18.29
140000	5999.90	17.11
150000	5958.57	10.44
152561	4484.12	7.14
160000	5978.675	19.19
173061	4495.13	7.36
180000	6010.03	6.79
190000	6010.48	11.56
197637	4490.82	11.99
205000	5991.60	16.02
209925	4501.58	6.761
218177	4491.09	10.11

TABLE 10. *BALI: Relationship of TSDU Sizes and Throughput*

Figure 16: Same TSDU sizes as used in Figure 11 and:

TSDU sizes: 5000, 5100, 5120, 5130, 5200, 5300, 13000, 13300, 13310, 13320, 13330, 13400, 13500, 25000, 25500, 25600, 25700, 26000, 48000, 49000, 50000, 50100, 50180, 50190, 50200, 51000, 52000, 68000, 69000, 70600, 70660, 70670, 70700, 71000, 94000, 95000, 95200, 95230, 95240, 95300, 96000, 97000, 110000, 111600, 111620, 111700, 112000, 127000, 128000, 129000, 130000, 152000, 152560, 152570, 152600, 153000, 172000, 173000, 173060, 173070, 174000, 197000, 197600, 197630, 197640, 198000, 209000, 209920, 209930, 210000

Figure 17: TSDU sizes: 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000

Scenario 7.2 - Benchmark Details

Figure 18: Script: `spims/rr_ss.spims`, `WAN/rr_ss.sres`

```
5 rr 1000,1024*50;
5 rr 2000,1024*50;
5 rr 3000,1024*50;
5 rr 4000,1024*50;
5 rr 5000,1024*50;
5 rr 6000,1024*50;
```

Values

Request Size	Throughput [KB/s]	Time [sec]
1000	9.37	5.34
2000	17.94	5.57
3000	26.33	5.69
4000	34.20	5.85
5000	22.22	11.25
6000	24.34	12.32
7000	31.14	11.24
8000	23.98	16.68
9000	38.92	11.56
10000	30.07	16.63
11000	44.82	12.27
12000	50.34	11.92
13000	51.73	12.56
15000	59.77	12.55
16000	63.41	12.62
18000	71.73	12.55
19000	75.87	12.52
22000	80.26	13.70
24000	92.67	12.95
26000	97.123	13.39
28000	89.025	15.73
30000	91.85	16.33
32000	114.79	13.93
34000	100.11	16.98
36000	107.834	16.69
38000	110.31	17.22
45000	118.81	18.94
50000	125.54	19.91
55000	150.76	18.24

Table 11 *- WAN: Relationship of Request-Response Time/Throughput on TSDU Size, TCP Slowstart*

Request Size	Throughput [KB/s]	Time [sec]
60000	158.99	18.87
65000	163.45	19.88
75000	163.04	23
80000	176.57	22.65
85000	177.80	23.90
95000	192.32	24.70
100000	182.84	27.35
120000	208.38	28.79
140000	211.61	33.08
160000	228.59	34.99
200000	228.59	34.99

***Table 11** - WAN: Relationship of Request-Response Time/Throughput on TSDU Size, TCP Slowstart*

Figure 19: Script: `spims/rr_ss.spims,` `BALI/bali_rr_ss.sres`

```
5 rr 1000,1024*50;
5 rr 2000,1024*50;
5 rr 3000,1024*50;
5 rr 4000,1024*50;
5 rr 5000,1024*50;
5 rr 6000,1024*50;
```

Values:

Request Size	Throughput [KB/s]	Time [sec*100]
1000	448.02	11.16
2000	880.28	11.36
3000	1048.95	14.3
4000	1392.75	14.36
5000	25.71	972.1
6000	30.76	975.14
7000	36.21	966.32
8000	41.07	973.8
9000	46.55	966.5
10000	51.14	977.64
11000	56.78	968.54
12000	61.77	971.3
13000	66.73	974.06
14000	2108.43	33.2
15000	1801.15	41.64
16000	2129.92	37.56
17000	2017.08	42.14
18000	2155.17	41.76
19000	2203.15	43.12

***Table 12** - BALI: Relationship of Request-Response Time/Throughput on TSDU Size, TCP Slowstart*

Request Size	Throughput [KB/s]	Time [sec*100]
20000	2455.79	40.72
45000	222.538	1011.06
50000	263.46	948.9
55000	3795.19	72.46
60000	4038.77	74.28
65000	308.68	1052.84
70000	357.84	978.08
75000	3371.69	111.22
80000	3746.01	106.78
85000	3741.85	113.58
90000	3571.42	126
95000	507.54	935.88
100000	505.64	988.84
100000	522.95	956.1
120000	3259.80	184.06
140000	3100.08	225.8
160000	820.74	974.72
180000	3096.1	290.68
200000	3671.34	272.38
220000	1140.86	964.18
250000	1341.54	931.76

Table 12 - BALI: Relationship of Request-Response Time/Throughput on TSDU Size, TCP Slowstart

Scenario 8.1 - Benchmark Details

Figure 20: 1st script: spims/par2.spims, WAN/mpx_const_70_w32.sres

```
5 bulk_get 70000*250;
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get 70000*250;
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get 70000*250 |
            1 bulk_get 70000*250 |
            1 bulk_get 70000*250;
[...]
```

2nd script: spims/mpx_70.spims, WAN/mpx_const_70_biv_small_w32.sres

```
5 bulk_get 70000*250;
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2)|
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2)|
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2)|
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
[...]
```

Values:

No. of simult. connections	Throughput [KB/s] 70 KB const, small distr.	Throughput [KB/s] 70 KB constant
1	311.58	311.22
2	288.17	290.28
3	267.09	–
4	259.39	172.78
5	243.83	–
6	232.06	–
7	228.33	–
8	223.37	65.12
9	213.16	–
12	–	49.93
13	191.56	–
16	–	36.44
17	171.32	–
20	–	28.75
21	155.64	–
24	–	24.01

Table 13 - WAN: Influence of Bundeled Unidirectional Connections on Throughput

Figure 21: 1st script: `spims/par2.spims, BALI/bali_mpx_const_70_w32.sres`

```
5 bulk_get 70000*250;
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get 70000*250;
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get 70000*250 |
            1 bulk_get 70000*250 |
            1 bulk_get 70000*250;
[...]
```

2nd script: `spims/mpx_70.spims, BALI/bali_mpx_const_70_biv_small_w32.sres`

```
5 bulk_get 70000*250;
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2)|
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get 70000*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2)|
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2)|
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
[...]
```

Values:

No. of simult. connections	Throughput [KB/s] 70 KB const, small distr.	Throughput [KB/s] 70 KB constant
1	4372.15	4425.45
2	4231.35	2657.07
3	4180.66	–
4	4094.79	1347.10
5	4025.37	–
6	3908.76	–
7	3863.88	–
8	3466.61	643.65
9	3162.85	–
12	–	443.47
13	2581.70	–
16	–	316.24
17	1504.90	–
20	–	250.10
21	1727.14	–
24	–	207.27

Table 14 - BALI: Influence of Bundeled Unidirectional Connections on Throughput

Figure 22. 1st script: `spims/mpx_biv_all.spims,`
`WAN/mpx_biv_both_w32.sres`

```
5 bulk_get bivalue(100,70000,0.75)*250;
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
```

[...]

2nd script: `spims/multiplex.spims,`
`WAN/mpx_biv_100_70000_w32.sres`

```
5 bulk_get bivalue(100,70000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
```

```
1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
    [...]
```

Values:

No. of simultaneous connections	Throughput [KB/s] 70 KB and small distr., no inter-packet delay	Throughput [KB/s] 70 KB and small distr., inter-packet delay distr.
1	271.02	155.24
2	223.05	140.44
3	190.58	118.50
4	144.55	109.79
5	163.77	107.01
6	136.97	105.62
7	117.258928698833	82.8723766641547
8	117.99	100.33
9	106.04	–
10	–	76.77
12	–	76.92
13	82.79	–
14	–	73.51
17	79.59	–
16	–	55.13
18	–	68.27
21	66.29	–

Table 15 - WAN: Influence of Bundeled Unidirectional non-homogeneous Connections on Throughput

Figure 23: 1st script: `spims/mpx_biv_all.spims,`
`BALI/bali_mpx_biv_both_w32.sres`

```
5 bulk_get bivalue(100,70000,0.75)*250;
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
               [...]
```

2nd script: `spims/multiplex.spims,`
`BALI/bali_mpx_biv_100_70000_w32.sres`

```
5 bulk_get bivalue(100,70000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
```

```
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
5 parallel: 1 bulk_get bivalue(100,70000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2) |
            1 bulk_get bivalue(10,9000,0.75)*250 exp(0.2);
```

[...]

Values:

No. of simultaneous connections	Throughput [KB/s] 70 KB and small distr., no inter-packet delay	Throughput [KB/s] 70 KB and small distr., inter-packet delay distr.
1	4226.10	1581.11
2	3829.92	896.70
3	3884.77	562.60
4	3724.14	476.99
5	3752.16	364.47
6	3366.63	306.11
7	2781.39	307.64
8	2685.64	238.08
9	2338.72	–
13	2086.01	–
17	1319.73	–
21	995.06	–

Table 16 - BALI: Influence of Bundled Unidirectional Non-Homogeneous Connections on Throughput

Scenario 8.2 - Benchmark Details

Figure 24: 1st script: `spims/mpx_put.spims, WAN/mpx_put_70_const.sres`

```
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put 70000*100;
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put 70000*100 |
            1 bulk_get 70000*100 |
            1 bulk_put 70000*100;
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put 70000*100 |
            1 bulk_get 70000*100 |
            1 bulk_put 70000*100 |
            1 bulk_get 70000*100 |
            1 bulk_put 70000*100;
```

[...]

2nd script: `spims/mpx_put_dist.spims, WAN/mpx_put_dist.sres`

```
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100;
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100 |
```

```
            1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100;
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100 |
            1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100 |
            1 bulk_get 70000*100 |
1 bulk_put bivalue(10,9000,0.9)*100;
```

Values:

No. of simult. connections	Throughput [KB/s] GET 70 KB const., PUT 70 KB const.	Throughput [KB/s] GET 70 KB const., PUT small distr.
2	299.94	303.11
4	260.65	292.78
6	179.75	231.47
8	142.57	151.09
10	111.72	128.49
12	94.86	115.74
14	79.85	93.72
16	65.86	80.043
18	61.14	71.99
20	56.81	65.51

Table 17 - WAN: Influence of Bundled Non-Homogeneous Bidirectional Connections on Throughput

Figure 25: 1st script: `spims/mpx_put.spims,`
`BALI/mpx_put_70_const.sres`

```
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put 70000*100;
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put 70000*100 |
            1 bulk_get 70000*100 |
            1 bulk_put 70000*100;
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put 70000*100 |
            1 bulk_get 70000*100 |
            1 bulk_put 70000*100 |
            1 bulk_get 70000*100 |
            1 bulk_put 70000*100;
```

[...]

2nd script: `spims/mpx_put_dist.spims,`
`BALI/mpx_put_dist.sres`

```
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100;
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100 |
            1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100;
5 parallel: 1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100 |
            1 bulk_get 70000*100 |
            1 bulk_put bivalue(10,9000,0.9)*100 |
            1 bulk_get 70000*100 |
1 bulk_put bivalue(10,9000,0.9)*100;
```

Values:

No. of simult. connections	Throughput [KB/s] GET 70 KB const., PUT 70 KB const.	Throughput [KB/s] GET 70 KB const., PUT small distr.
2	3857.30	4291.85
4	1821.21	2545.42
6	1054.13	1692.81
8	762.22	1260.44
10	572.05	1027.11
12	471.02	812.24
14	380.12	718.49
16	318.53	621.09
18	266.21	547.81
20	249.55	489.02

Table 18 - BALI: Influence of Bundled Non-Homogeneous Bidirectional Connections on Throughput

Scenario 8.3 - Benchmark Details

Figure 26: Script: SPIMS script for request-response,
shell script for xtcp bulk transfer

Values:

No. of simult. connections	Normalized Time	Throughput [KB/s]
0	6.33925714285714	157.747189846489
1	7.67428571428571	130.305286671631
2	9.0244	110.810691015469
3	10.2209047571429	97.8386966477847
4	11.0223809428571	90.7245000135866
5	12.1275238	82.4570634938684

Table 19 - WAN: Request-Response Time, Throughput, No. of Bundled Connections

Figure 27: Script: SPIMS script for request-response,
shell script for xtcp bulk transfer

Values:

No. of simult. connections	Normalized Time	Throughput [KB/s]
0	2.83	353.17
1	2.75	363.33
2	2.69	371.70
3	2.77	360.40
4	2.77	360.07
5	2.70	369.02

Table 20 - BALI: Request-Response Time, Throughput, No. of Bundled Connections

12.2 Transmission Hierarchies

12.2.1 Digital Signal Hierarchies

In the 1960s, a worldwide effort began to upgrade public switched telephone systems from all-analog systems to systems supporting a combination of analog and digital signals. The North American effort resulted in a system in which 24 64-Kb/s voice signals are multiplexed into a single 1.544-Mb/s signal. In Europe, a system emerged that multiplexed 30 voice channels for a total rate of 32x64 Kbps = 2048 Kb/s.

Level	North America	Europe	Japan
1	1.544 Mb/s (DS 1)	2.048 Mb/s	1.544 Mb/s
2	6.312 Mb/s (DS 2)	8.448 Mb/s	6.312 Mb/s
3	44.736 Mb/s (DS 3)	34.368 Mb/s	32.064 Mb/s

Table 21 - Digital Signal Hierarchies

12.2.2 Optical Signal Level Standards

In 1984 began the efforts to support existing digital levels:
T1 rates at 1.544 in U.S., T2 = DS2, T3 =DS3 in USA,
and E1 - 2.048 in Europe, E2=DS2 Eu. E3 = DS3 Europe.

12.2.3 B-ISDN Standards

The optical data rates, synchronisation and framing format chosen for B-ISDN are called Synchronous Digital Hierarchy (SDH) in Europe and Synchronous Optical Network (SONET) in North America.

12.2.4 SONET Optical Signal Hierarchy

The following table shows the SONET STS/OC speeds in and the ITU-T STM (Synchronous Transport Module) speeds in Mb/s.

Data Rate (Mb/s)	Sonet STS/OC Designation	ITU-T STM Designation
51.84	OC-1	—
155.52	OC-3	STM-1
466.56	OC-9	—
622.08	OC-12	STM-4
933.12	OC-18	—
1244.16	OC-24	STM-8
1866.24	OC-26	—
2488.32	OC-48	STM-16

Table 22 - Sonet/SDH Transmission Rates

13 Appendix 2

SPIMS is the "*SICS Protocol Implementation Measurement System*" developed at the Swedish Institute of Computer Science. The last version is SPIMS version 3.0, released February 1991. It comes with a detailed user manual describing the usage and functions of SPIMS and covers the installation and user-written extensions as well [31].

SPIMS is a tool for measuring the performance of different protocol implementations by providing a portable tool which uses a protocol-independent specification language so that the same measurement specification can be used for measurements on different machines and operating systems. SPIMS – being simply a distributed application that executes measurement specifications and presents the measured performance to the user – uses the standard measurement facilities supplied by the operating system.

This section describes the SPIMS functions and facilities used in the performance measurements in this work. For a detailed description of how to install/run SPIMS, the SPIMS output, and how to write measurement specifications, refer to [31].

13.1 SPIMS Underlying Concepts

In order to measure the performance a real application would experience, the measurements have to be done in an environment similar to the environment the application would run in. In order to get reproducible measurements, the environment has to be controlled in some way. SPIMS requires two machines (minimum) and a network connecting them. There, SPIMS is run as a distributed application, thus measuring at the application level, including operating system overhead. Measurements are performed by having a communicating entity on each of the machines. The active party is called the *initiator*, the awaiting party the *responder* (see Figure 29).

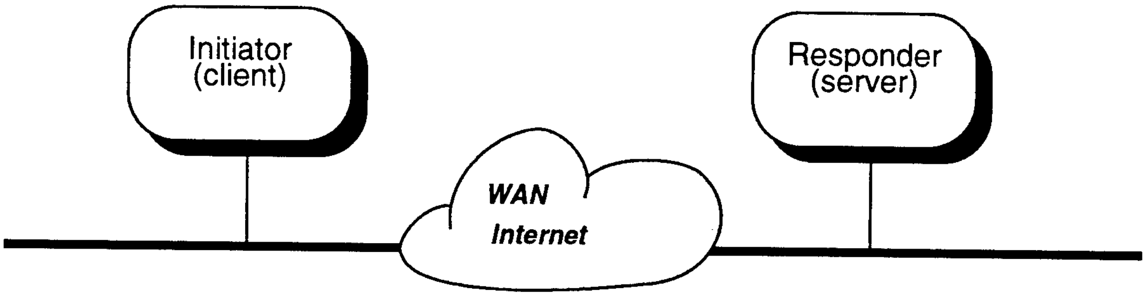


Figure 29 - SPIMS Measurement Environment.

13.2 Application Types

Each measurement is performed between an *initiator* and a *responder* and consists of a specified number of identical measurement *runs* using an *application type* with the parameters *data unit size* and *number of data units*. A simple specification for measuring the throughput is, for example:

```
10 bulk_get 1024*1000 exp(0.2);
```

This should be read as: "Run `10` runs of this measurement and collect data to be presented statistically. In each run use the **`bulk_get`** application type with a data unit size of `1024` bytes and transfer `1000` such data units for each run. Introduce an inter-packet delay that is **`exp`**onentially distributed with mean `0.2`, thus sending an average of 5 packets per second. The syntax of the above type of specification is:

```
<num runs> <application type> <data unit size>"*"<num data units> [<timing spec>] ";"
```

In this work, three application types have been used. Briefly, they are:
- **bulk_get**. Bulk data transfer read. The responder sends data as fast as possible to the initiator. See Figure 30.
- **bulk_put**. Bulk data transfer write. The initiator sends data as fast as possible to the responder.

- **request_response** or **rr**. Request-response measuring two-way delay. See Figure 30.
- There are additional application types not used in this work. Refer to [31] for an explanation of the application types *rpc, query, conn_disc, conn_disc_n.*

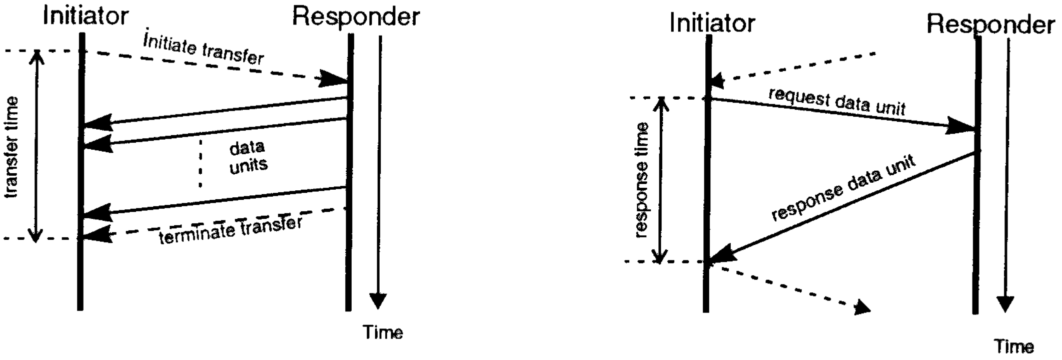


Figure 30 - The bulk_get and request_response Application Classes.

13.3 Distributions

It is possible to specify a distribution as the data unit size or as the timing specification (see the example in Section 13.2). The following distributions have been used for this work:

- **exp(mean)**. The size/time is exponentially distributed with an average of *mean*;
- **bivalue(val1,val2,p)**. The size/time is *val1* with a probability of *p* ($0<=p<=1$) and *val2* with probability ($1-p$);
- **constant(val)**. The size/time is constant and has value *val.*

13.4 Simultaneous Measurements

Measurements can be made having multiple initiator/responder pairs execute in parallel. This is achieved by specifying a measurement *composer* with the standard measurement specification. In this work, the parallel composer has been used to get the performance studies for the bundled connections. However, there are a number of other composers [31]. An example for such a specification looks like the following:

```
10 parallel 3: 1 bulk_get 1024*1000 |
               1 bulk_get 1024*1000 |
               1 bulk_get 1024*1000 ;
```

This specification says that three initiator/responder pairs are to be run, each of them executing a basic bulk_get with the given parameters. Different application types and/or parameters could be specified for each of the simultaneous measurements.

14 References

[1] A. Iwata, N.Mori, C. Ikeda, H. Suzuki. M. Ott, "ATM Connection and Traffic Management Schemes for Multimedia Interworking", Communications of the ACM, Feb. 1995, Vol. 38, No.2
[2] T. Kwok " A Vision for Residential Broadband Services: ATM-to-the-Home", IEEE Network, Vol. 9, No. 5, 1995
[3] S. Dixit, P- Skelly, "MPEG-2 over ATM for Video Dial Tone Networks: Issues and Strategies, IEEE Network, Vol. 9, No. 5, 1995
[4] K. Taylor, "The Ultimate ATM Mix: More Talk, Less Bandwidth", Data Communications International, January 1996
[5] K. Eppele, "ATM rechnet sich auch im WAN", DATACOM 9/1995
[6] ATM Forum Technical Committee, "Traffic Management Specification" Version 4.0, ATM Forum / 95-0013R8, Straw Vote, October 1995
[7] ATM Forum Technical Committee Working Group Traffic Management, "ATM Service Architecture from Applications to Scheduling, November, 1995
[8] R. Caceres, "Multiplexing Traffic at the Entrance to Wide-Area Networks", December 1992
[9] L. Kalampoukas, A. Varma, "Performance of TCP over Multi-Hop ATM Networks: A Comparative Study of ATM-Layer Congestion Control Scheme, UCSC-CRL-95-13, February, 1995
[10] B.A. Mah, "On the Use of Quality of Service in IP over ATM", The Tenet Group Computer Science Division, University of California at Berkeley, 1995

[11] R.Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification", Internet Draft draft-ietf-rsvp-spec-09.ps, February 1996

[12] J. Crowcroft, Zh. Wang, A. Smith, J. Adams, "A Rough Comparison of the IETF and ATM Service Models", IEEE Network, November/December 1995

[13] M. Perloff, K. Reiss, "Improvements to TCP Performance in High Speed ATM Networks", Communications of the ACM, February 1995

[14] R. J. Vetter, "ATM Concepts, Architectures and Protocols", Communications of the ACM, February 1995

[15] A. Romanov, "Preliminary Report of Performance Results for TCP over ATM with Congestion", July 1993

[16] R. Caceras, "Efficiency of ATM Networks in transporting Wide Area Data Traffic", Computer Networks and ISDN systems, December 1991

[17] R.Caceras,"Multiplexing Traffic at the Entrance to Wide-Area Networks, Ph.D. Dissertation, December 1992

[18] A. Romanov, S. Floyd, "Dynamics of TCP Traffic over ATM networks", IEEE Journal on Selected Areas in Communication, May 1995

[19] F. Pucher, H. A. Huy, "QoS Parameters for Virtual Connections within Local ATM Networks: A Performance Report", Intl. Workshop on Protocols for Multimedia Systems (PROMS 95), Salzburg, October, 1995

[20] I. Miloucheva, T. Ziegler, "Protocol Mechanisms for Reliable Transmission and Flow Control for Multimedia Highways", IEEE IPCCC 96, Phoenix, Arizona, 27-29 March, 1996

[21] S. Keshav, "REAL: A Network Simulator", Computer Science Department, Tech. RPT 88/472, University of California, Berkley, December, 1988

[22] J. Postel, "Transmission Control Protocol - DARPA Internet Protocol Specification", RFC 793, Sept. 1981

[23] V. Jacobson, "Congestion Avoidance and Control", ACM SIGCOMM, Aug.1988

[24] I. Miloucheva, O. Bonnesz, "XTPX Based Multimedia Transport System for Flexible QoS Support", 3rd International Conference on Telecommunication Systems, Modelling and Analysis, Nashville, March, 1995

[25] I. Miloucheva, O. Bonnesz, G. Swieckowski, "Tuning of Protocol Parameters for Enhanced QoS Guarantee and Network Utilization", IEEE ICECS, Amman, Jordan, Dec. 95

[26] I. Miloucheva, K. Rebensburg, "QoS Based Architecture Using XTP", 4th IEEE Workshop on Future Trends of Distributed Systems, Portugal, Sept. 1995

[27] M. Safaie, "Analyse der Einbettung von Transportprotokollen ueber ATM, Messungen der Performance und Diskussion der Resource Reservierungskonzepte", Master Thesis, TUB, in German, Oct. 1993

[28] I. Miloucheva, G.Swieckowski, "Protocol Tuning Box - A Tool for Performance Analysis of Multimedia Protocols", Proceedings PROMS 95, Salzburg, October, 1995

[29] C. Topolcic, "Experimental Internet Stream Protocol : Version 2 (ST-II). Internet RFC 1190, October. 1990

[30] L. Delgrossi, L. Berger, "Internet STream Protocol Version 2+, Interenet Draft , draft-ietf-st2-spec-02.txt

[31] SICS Protocol Implementation Measurement System: User Manual. Swedish Institute of Computer Science, Version 3.0, Febr. 1991

[32] interWATCH 95000 ATM-LAN-WAN Protocol Analyzer, User Guide Beta Version 0.95

[33] Adtech AX/4000 ATM Test System

[34] VIVID ATM Workgroup Switch Manual, Newbridge Networks, Inc., Kanata, Ontario, Issue 2, March 1995

[35] MainStreet 36150 ATM Switch Manual, Newbridge Networks, Inc., Kanata, Ontario, Issue 1, June 1994

[36] H. Hartentaler, D. Hetzer, B. Butscher, "BALI - Integration of Inhouse ATM and Local/Wide Area Networks". An Ideal Base for Multimedia Teleservices, A Joint Project of DeTeBerkom, GMD-FOKUS and Technische Universitaet Berlin (FSP-PV, TUBKOM), Jan. 1994

[37] B-ISDN Application "Terravision" Using an International ATM Connection between Germany and Japan, DeTeBerkom, 1995

[38] MAY - Multimedia Applications on Intercontinental Highspeed Networks, DeTeBerkom, 1995

[39] Global ATM Solutions, Deutsche Telecom, France Telecom. Sprint International, 1995

[40] W. Almesberger, J-Y Le Boudec, Ph. Oechslin "Application REQuested IP over ATM (AREQUIPA), Internet Draft , draft-almesberger-arequipa-00.txt, January, 1996

[41] Y. Goto, "Session Identity Notification Protocol (SINP), Internet Draft, draft-goto-sinp-01.txt, Jan. 1996

[42] M. Laubach," Classical IP and ARP over ATM", IETF RFC 1577, 1994

[43] M. Laubach, J. Halpern, "Classical IP and ARP over ATM Update(Part Deux)" Internet Draft, <draft-ietf-ipatm-classic2-01.txt>, 22 Febr. 1996

[44] R. O. Onvural, V. Srinivasan, "A Framework for Supporting RSVP Flows Over ATM Networks", Draft, 26 February, 1996

[45] J. Heinanen, "Multiprotocol Encapsulation over ATM Adaptation Layer 5. RFC 1483, IETF, Juy 1993

[46] R. Atkinson, "Default IP MTU for use over ATM AAL 5", RFC 1626, IETF, May 1994

[47] R. G. Cole, D. H. Shur, C. Villamizar, "IP over ATM: A Framework Document", Internet Draft, draft-ietf-ipatm-framework-doc-07, February 7, 1996

[48] M. Perez, F. Liaw, D. Grossman, A. Mankin, A.Hoffman, "ATM Signalling Support for IP over ATM", RFC 1755, IETF, Jan. 1995

[49] C. Brazdziunas, "IPng Support for ATM Services" IETF, RFC 1680, Aug. 1994

[50] P. Schulter, "A Framework for IPv6 Over ATM", Internet-Draft, draft-schulter-ipv6atm-framework-00.txt, November, 1995

[51] G. Armitage, "IPv6 and Neighbour Discovery over ATM" Internet Draft, <draft-ietf-ipatm-ipv6nd-01.txt>, February, 1996

[52] ATM Forum, "ATM User Network Interface Specification (UNI) Version 3.1" ISBN 0-13-393828-X, Prentice Hall, Englewood Cliffs, NJ, June 1995

[53] ATM Forum, "ATM User-Network Interface (UNI) Signalling Specification Version 4.0", 95-1434R6, December 1995

[54] ATM Forum, "PNNI Draft Specification", ATM Forum 94-0471R14, 1995

[55] R. Braden, D. Clark, S. Shenker, "Integrated Services in the Internet Architecture: an Overview", RFC 1633, 1994

[56] M.B. Matthis, J. Mahdavi, S. Floyd, A. Romanow, "TCP Selective Acknowledgment Option", Internet Draft, 1996

[57] A. Birman, V. Firoiu, R. Guerin, D. Kandlur, "IBM Research Report, Provisioning of RSVP-based Services over a Large ATM Network", IBM Research Division, Computer Science, RC 20250 (10/27/95), November 1995

[58] V. Jacobson, R. Braden, D. Borman, "TCP Extensions for High Performance", RFC 1323, May 1992

[59] S. Floyd, V. Jacobson, "Link-sharing and Resource Management Models for Packet Networks, IEEE/ACM Transactions on Networking, Vol. 3, No. 4, August 1995

[60] S. Shenker, C. Partridge, "Specification of Guaranteed Quality of Service", Internet Draft, draft-ietf-intserv-guaranteed-svc-03.txt, December 1995

[61] S. Shenker, C. Partridge, B. Davie, L. Breslau, "Specification of Predictive Quality of Service", Internet Draft, draft-ietf-intserv-predictive-svc-01.txt, Nov. 1995

[62] S. Shenker, C. Partridge, J. Wroclawski. "Specification of Controlled Delay Quality of Service", Internet Draft, draft-ietf-intserv-control-del-svc-02.txt , Nov. 1995

[63] J. Wroclawski, "Specification of the Controlled-Load Network Element Service", Internet-Draft, draft-ietf-intserv-ctrl-load-svc-01.txt, Nov. 1995

[64] S. Shenker, J. Wroclawski, "Network Element Service Specification Template", Internet Draft, draft-ietf-intserv-svc-template-02.txt, Nov. 1995

[65] L. Breslau, S. Shenker, " A Proposal for Accommodating Heterogeneity", Internet Draft, draft-ietf-intserv-hetero-00.txt, Nov. 1995

[66] P. Almquist, "Type of Service in the Internet Protocol Suite", RFC 1349, July 1995

[67] M. Ohta, H. Esaki, K. Nagami, "Conventional IP over ATM", Interent Draft, draft-ohta-ip-over-atm-02.txt, March 1995

[68] D. Katz, D. Piscitello, B. Cole, J.V. Luciani, "NBMA Next Hop Resolution Protocol (NHRP)", Internet Draft, draft-ietf-rolc-nhrp-07.txt, 1996

[69] D. Canserver, "NHRP Protocol Applicability Statement, Internet Draft, draft-ietf-rolc-nhrp-appl-01.txt, March 1995

[70] Y. Katsube, K-i. Nagami, H. Esaki, "Router Architecture Extensions for ATM : Overview", <draft-katsube-router-atm-overview-01.txt>, September 1995

[71] M. Borden, E. Crawley, B. Davie, S. Batsell, "Interation of Real-Time Services in an IP-ATM Network Architecture", RFC 1821, August 1995

[72] G. Armitage. "Support of Multicast over UNI 3.0 / 3.1 based Networks", Internet Draft, draft-ietf-ipatm-ipmc-12.txt, 23 February, 1996

[73] XTP Forum, "The Xpress Transport Protocol Specification", XTP Revision 4.0, 1995

[74] W.T. Strayer, B.J. Dempsey, A.C. Weaver, "XTP: The Xpress Transfer Protocol", Addison Wesley, July 1992.

[75] I. Miloucheva, B. Metzler, "XTPX Version 2.0 Protocol Specification", Interdepartmental Research Center for Process Control Applications, Technical University of Berlin, February 1993

[76] H. Buschermöhle, A. Wendt, T. Kerschat, "XTPX Version 2.0", Technical Documentation, FSP-PV, Technical University of Berlin, April 1994.

ITU-T Recommendations

E.164	Numbering Plan for the ISDN era, 08/91
G.707	Synchronous Digital Hierarchy Bit Rates, 04/91
G.708	Network Node Interfacr for the Synchronous Digital Hierarchy, 04/91
G.709	Synchronous Multiplexing Structure, 04/91
I.150	Recommendation I.150 B-ISDN ATM Functional Characteristics CCITT SG XVIII, Geneva 1990
I.211	Recommendation I.211 B-ISDN Service Aspects, CCITT Draft Recommendation
I.311	B-ISDN General Network Aspects, 04/91
I.321	B-ISDN Protocol Reference Model and its Application, 04/91
I.327	B-ISDN Functional Architecture, 04/91
I.330	ISDN Numbering and Addressing Principles, 1988
I.331	Numbering Plan for the ISDN era, 1988, contains only a pointer to E.164
I.332	Numbering Principles for Internetworking between ISDNs and Dedicated Networks with Different Numbering Plans, 1988
I.356	
I.361	B-ISDN ATM Layer Specification, 04/91
I.362	B-ISDN ATM Adaptation Layer (AAL) Functional Description, 04/91
I.363	B-ISDN ATM Adaptation Layer (AAL) Specification, 04/91
I.364	Support for Broadband Connectionless Data Service on B-ISDN
I.371	Traffic Control and Congestion Control in B-ISDN
I.413	B-ISDN User-Network Interface, 04/91
I.432	B-ISDN User-Network Interface—Physical Layer Specification, 04/91
I.610	OAM Principles of the B-ISDN Access, 04/91
Q.11bis	Numbering Plan for ISDN era, 1988
Q.11ter	Timetable for Coordinated Implementation of the Full Capability of the Numbering Plan for the ISDN era (E.164), 1988
Q.931	ISDN User-Network Interface Layer 3, Specification for Basic Call Control, 1988

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM
(highest classification of Title, Abstract, Keywords)

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.) COMMUNICATION RESEARCH CENTRE 3701 CARLING AVENUE OTTAWA, ONTARIO K2H 8S2	2. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable) UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.)

TRANSPORT FLOW CONTROL AND CONNECTION ADMISSION POLICIES FOR RELIABLE APPLICATIONS OVER ATM WAN (U)

4. AUTHORS (Last name, first name, middle initial)

LOUISE LAMONT, ILKA MILOUCHEVA, WOLFRAM STERING

5. DATE OF PUBLICATION (month and year of publication of document) APRIL 1996	6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.) 71	6b. NO. OF REFS (total cited in document)

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

CRC REPORT

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

Defence Research Establishment Ottawa
3701 Carling Avenue
Ottawa, Ontario
K1A 0Z4

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)	9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)
10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.) CRC-RP-96-007	10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availabilty (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

UNLIMITED

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

WITH THE ARRIVAL OF INFORMATION SUPERHIGHWAYS, INTERCONTINENTAL ATM WANs BECOME MORE IMPORTANT FOR THE PROVISION OF MULTIMEDIA APPLICATIONS. AN INITIATIVE BETWEEN DeTeBERKOM AND THE COMMUNICATION RESEARCH CENTRE (CRC) WAS UNDERTAKEN TO PROVIDE TRANSPORT PROTOCOL PERFORMANCE CHARACTERISTICS FOR RELIABLE APPLICATIONS AND EFFICIENT ADMISSION CONTROL POLICIES FOR BUNDLING TRANSPORT CONNECTIONS.

IN OUR WORK, WE INVESTIGATE PRACTICAL ISSUES OF TCP/IP OVER TRANS-ATLANTIC ATM AND COMPARE THEM WITH TCP/IP PERFORMANCE ON LOCAL AREA ATM NETWORKS. WE PROVIDE A FIRST GLANCE AT TCP/IP MEASUREMENTS OVER HIGH-SPEED TRANS-ATLANTIC CONNECTION, USING NATIONAL HIGH-SPEED TEST NETWORKS AND TELEGLOBE' TRANS-ATLANTIC SUBMARINE FIBRE CANTAT-3.

BASED ON SCENARIOS FOR BULK DATA AND TRANSACTION APPLICATION, WE ILLUSTRATE THE SPECIFIC EFFECTS OF DIFFERENT PERFORMANCE FACTORS SUCH AS: TCP FLOW CONTROL AND SEND WINDOW SIZE, NETWORK DELAY, SYSTEM SCHEDULING AND APPLICATION TRAFFIC ON THE QoS PROVISION IN ATM WAN AND LAN. WE ALSO PRESENT THE PERFORMANCE EFFECTS OF BUNDLING TRANSPORT CONNECTIONS OVER THE SAME ATM RESOURCES. PRACTICAL CONSIDERATIONS FOR MAPPING THE APPLICATION'S TRAFFIC AND QoS REQUIREMENTS TO ATM RESOUCRES ARE DISCUSSED

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

TCP FLOW CONTROL
PERFORMANCE MEASUREMENTS
CONNECTION ADMISSION CONTROL
ATM WAN
ATM LAN
TRANS-ATLANTIC LINKS
BUNDLING
RESOURCE RESERVATION