JOURNAL ON

COMMUNICATIONS

VOLUME XLVII JANUARY-FEBRUARY 1996

ATM NETWORKS I.

JOURNAL ON COMMUNICATIONS

A PUBLICATION OF THE SCIENTIFIC SOCIETY FOR TELECOMMUNICATIONS, HUNGARY

Editorial office
Budapest XIV. Ungvár u. 64-66.
H-1525 Budapest, P.O. Box 15.
Phone: (361) 251-1163, (361) 201-7471
Fax: (361) 251-9878, (361) 201-7471

Subscription rates
Hungarian subscribers
1 year, 12 issues 5300 HUF, single copies 650 HUF
Hungarian individual subscribers
1 year, 12 issues 860 HUF, single copies 110 HUF
Foreign subscribers
12 issues 150 USD, 6 English issues 90 USD, single copies 24 USD
Transfer should be made to the Hungarian Foreign Trade Bank, Budapest, H-1821, A/C No. MKKB 203-21411

JOURNAL ON COMMUNICATIONS is published monthly, alternately in English and Hungarian by TYPOTEX Ltd. H-1015 Bp. Retek u. 33-35. Phone/Fax: (361) 115-1759. Publisher: Zsuzsa Votisky. Type-setting by TYPOTEX Ltd. Printed by Dabasi Jegyzetnyomda, Dabas, Hungary HUISSN 0866-5583

EDITORIAL

The double ASYNCHRONOUS TRANSFER MODE (ATM) NETWORKS I. and the single ASYNCHRONOUS TRANSFER MODE (ATM) NETWORKS II. special issues are important landmarks in the history of this journal as its first broadband telecommunications related special issues. The JOURNAL ON COMMUNICATIONS of the Hungarian Society for Telecommunications is published bimonthly, sometimes more frequently, in English. Each issue is focused on a selected important subject that is discussed by both international and Hungarian authors.

According to present trends, we are facing a new "ICE" age: due to the integration of broadband Information, Communications and Entertainment (ICE) multimedia technologies and services, broadband integrated communications will become a key sector of telecommunications industry by the second half of the decade. The cell based and connection oriented ATM has been standardized in 1990 by the ITU (International Telecommunication Union) as the main carrier of the Broadband Integrated Services Digital Network (B-ISDN). The two most important characteristics of emerging high speed networks are:

- ATM will play a key role as transmission and switching technology with flexibly variable capacity, ranging from local to global networks;
- the global network will involve heterogeneous networking technologies, in which ATM is expected to interwork with other network components in several ways:
  - ATM as a carrier network for others, e.g. LAN (Local Area Networks) interconnections, up to 155 Mbit/s terminal speed presently;
  - ATM as transported by other carriers, e.g. SDH/ /SONET (Synchronous Digital Hierarchy/Synchronous Optical NETworks), up to the optical 10 Gbit/s speed presently;
  - others, e.g. management network, signalling network.

The internationational broadband community reacts adequately on the ICE age challenge: ATM and SDH/ SONET networks have already been installed in several continents. Considering Europe, the two-year long European ATM Pilot project has just been finished and is continued in a new two-year project JAMES involving 18 operator companies presently. At the same time, some European operators already provide ATM based services on tariff basis.

The Hungarian ATM market is just opening: several universities have constructed ATM trials and the Hungarian Telecommunications Company (HTC) has announced to perform a broadband multimedia application demonstration through its ATM pilot network in cooperation with the National Information Infrastructure Program (NIIP) and the Technical University of Budapest (TUB) at the 7-th Joint European Networking Conference (JENC7) organized by the TERENA (Trans-European Reserach and Education Networking Association) in Budapest, Hungary, May 13–17, 1996.

The international research community, both academic and industrial, is producing a rapidly increasing number of ATM related research results to satisfy the broadband demands. The international appreciation of Hungarian results is demonstrated by the fact that the Ericsson Company has established a Competence Centre for broadband networks related research and development in Budapest, based on the research cooperation between the Ericsson Telecommunication Systems Laboratories AB (EUA), Sweden, and the High Speed Networks Laboratory (HSN Lab) at the Department of Telecommunications and Telematics of the Technical University of Budapest, Hungary.

These special issues on ATM NETWORKS are devoted to demonstrate that Hungarian research, development, education, industry, marketing and service provision are ready to introduce the ATM technology on the opening Hungarian broadband market embedded into an international cooperation.

To maintain the highest possible quality of the special issues, as well as to provide a framework for the strict international reviewing process, an international Guest Editorial Board have been set up with the following members:

Additionally, the Hungarian members of the Guest Editorial Board have gained the valuable support of the following international broadband research communities, where the Call for Papers was also circulated and whose members participated in the reviewing process:

- IFIP (International Federation for Information Processing)
  Working Group 6.3
  "Performance of Communication Systems"
- COST (COoperation in the field of Science and Technology)
  Project 242
  "Methods for Performance Evaluation and Design of Boardband Multiservice Networks".

As a result, the ATM NETWORKS I. and ATM NETWORKS II. special issues comprise the following papers:

- 2 invited tutorial papers;
- 15 original technical research papers;
- 6 products/services papers.

The technical research articles have been selected from 23 high quality papers submitted from North America and Europe. All papers went through a strict reviewing process with at least two reviewers from the above international communities. But even those papers which could not be included into the frame of these issues, with or without Hungarian authors, received detailed comments supporting their further research work.

The Hungarian technical paper contributions come from two departments of the Technical University of Budapest:

- Department of Telecommunications;
- Department of Telecommunications and Telematics.

Both departments perform undergraduate and postgraduate education well embedded into Hungarian and international research, development, industry and service provision.

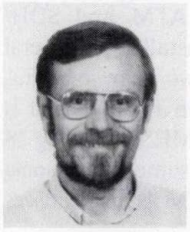
The other parts of the special issues, such as

- individual papers;
- news – events;

were edited directly by the Editor-in Chief, A. Baranyi and his Co-Editors.

The Guest Editor wishes to express his gratitude to the secretary and the members of the Guest Editorial Board and to the reviewers for their work, to the authors of the two invited tutorials for their contributions and to the IFIP WG 6.3, as well as to the COST 242 communities for their valuable support. Last, but not least, the Guest Editor wishes to thank to the Editorial Board of the Journal on Communications and, especially, the Editor-in Chief, A. Baranyi for providing the opportunity and supporting the compilation of the special issues.

TAMÁS HENK
E-mail: henk@bme-tel.ttt.bme.hu
Technical University of Budapest
H-1111 Budapest, Sztoczek u. 2., Hungary



Tamás Henk received the M.Sc. and Dr. Tech. Electronic Engineering degrees from the Technical University of Budapest, Hungary, in 1973 and 1980, respectively, and the Ph.D. degree from the Hungarian Academy of Sciences in 1985. Presently he is Associate Professor at the Department of Telecommunications and Telematics (DTT), Technical Univerrsity of Budapest, and Head of the High Speed Networks Laboratory (HSN Lab) within DTT. He was with the Research Institute for Telecommunications, Budapest, from 1973 to 1990, practising in different research and development projects and also in management. He was a Postdoctoral Research Fellow at the Univrsity College, Dublin, Ireland, from 1977 to 1979. His research interests are in signal processing and telecommunications, especially B-ISDN, multimedia, ATM networks and services. Dr. Henk is member of the IFIP Working Group 6.3 Performance on Communication Systems, secretary of Telecommunications System Committee of Hungarian Academy of Sciences, vice-president of Telecommunications Section of Scientific Society for Telecommunications, Hungary. Dr. Henk is a holder of "Tivadar Puskás" price awarded by Scientific Society for Telecommunications, Hungary – for his activity in the field of telecommunications.

AN OVERVIEW OF ATM INTERFACES

R. O. ONVURAL and G. A. MARIN

NETWORK ARCHITECTURE, IBM
RESEARCH TRIANGLE PARK, NORTH CAROLINA, USA

ATM is envisioned to provide universal networking by providing scalability in distance, scalability in speed, and integrating different services with different traffic characteristics and service requirements in the network. Scalability in distance refers to the use of ATM in the local area, campus, and wide area networks thereby providing seamless integration. Various physical interface speeds currently defined for ATM vary from 1.5 Mbps to 622 Mbps thereby providing scalability in speed. ATM provides the basic framework to integrate existing applications such as data and voice as well as emerging video and multimedia applications. Standardization is necessary to ensure interoperability. Towards providing a common architectural framework for emerging ATM networks, various standards organizations are working towards addressing the basic challenges of high speed networking. This paper presents an overview of the current status of ATM standards and specifications.

1. INTRODUCTION

The Broadband Integrated Services Digital Network (B-ISDN) architecture is envisioned to provide the basic framework to integrate video, voice, and data applications. The initial standardization work on ATM was completed in the late 1980s and Asynchronous Transfer Mode (ATM) was chosen as the technology to deliver B-ISDN. ATM is a connection oriented, packet switching technology that uses fixed size packets, referred to as cells. An ATM network provides only an ATM layer connectivity among two or more ATM layer users. There is no awareness on the specifics of applications or traffic types at the ATM layer. The ATM layer does not look inside the cell payload or needs to know what is carried in its payload.

Standardization provides users, manufacturers, and service providers the freedom of choice. Manufacturers benefit from standards as standards give them more opportunity to compete for the business of all the potential purchasers of telecommunications equipment. Service providers or organizations that build private networks take advantage of selecting their networking equipment from multiple vendors in a cost-effective manner. Users benefit the most as standard services give users a freedom in both selecting service providers and purchasing equipment that will fit best to their networking requirements in a most cost effective way based on their requirements.

ATM standards have been defined based on interfaces that isolates one or more network equipment from others while specifying the rules of interoperability among them. Currently defined ATM interfaces are illustrated in Fig. 1.

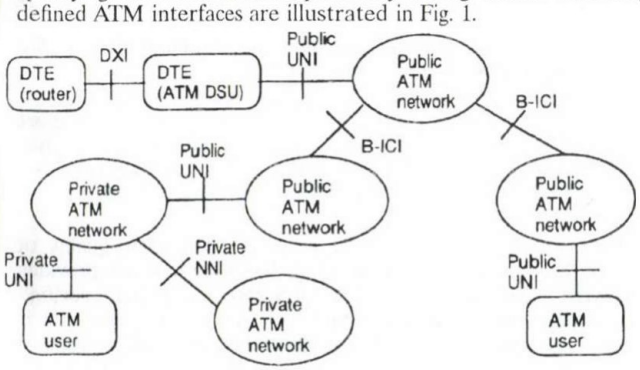


Fig. 1. ATM interfaces

User-to-network interface (UNI) is the demarcation point between an ATM end station and the network. An ATM end station in this context refers to any device that transmits ATM cells to the network. Accordingly, an ATM end station may be an interworking unit that encapsulates data into ATM cells, an ATM switch, or an ATM workstation. Depending on whether the attached ATM network is private or public, the interface is referred to as private UNI or public UNI, respectively. If two switches are connected to each other across a UNI and one switch belongs to a public network and the other to a private network, the corresponding demarcation point is always a public UNI. The UNI specification include the definitions of different physical layer interfaces, the ATM layer, a management interface, and signalling.

Network-to-network interface or network node interface (NNI) is used in different contexts. It may be the demarcation point between either two private networks or two public networks, respectively referred to as private NNI and public NNI. The same interface may also be used between two switches. In this case, NNI is a switch-to-switch interface in private networks (P-NNI) and a network node interface in public networks. Similar to UNI specification, NNI standards include the definitions of various physical layer interfaces, the ATM layer, a management interface, and signalling. Private NNI also includes the specification of P-NNI routing framework.

ATM data exchange interface (DXI) was developed to legacy equipments such as routers to interwork with ATM networks without requiring special hardware. In this specification, a DTE (a router) and a DCE (an ATM DSU) cooperates towards providing a UNI. The DXI specification includes the definitions of a data link protocol and physical layers that handle the data transfer between a DTE and a DCE as well as local management interface and management information base.

Broadband intercarrier interface (B-ICI) is a carrier-to-carrier interface (i.e., the ATM Forum version of public NNI). The B-ICI specification includes the various physical layer interfaces, ATM layer management, and higher layer functions required at the B-ICI for the interworking between ATM and services that include SMDS, Frame Relay, circuit emulation, and cell relay.

2. UNI SPECIFICATION

Connections in ATM networks are established either dynamically (switched virtual connections, i.e., SVCs) or pre-configured (permanent virtual connections, i.e., PVCs).

Physical layer interfaces currently specified for ATM is listed in Table 1.

The SVC support requires the definition of signalling protocols, procedures, and parameters for the dynamic management of ATM connections across a UNI whereas PVCs are established through switch specific procedures (i.e., network management).

Various features included in the UNI specification to address these requirements are categorized into four sections:

- physical layer interfaces;
- ATM layer;
- interim local management interface (ILMI);
- UNI signalling.

When a user wishes to communicate with another network user, it is necessary to provide the network with enough information about the connection request such as characterization of the connection traffic behavior and its service requirements. This exchange takes places across a UNI. Similarly, when the network attempts a connection establishment to an end station, the exchange between the network and an end station takes place across a UNI. The network, upon accepting the connection request, provides the user with the VPI/VCI the end station uses at the UNI for the requested connection. If the network cannot accept the connection then it informs the user that its request is rejected. When the communication is over, the user terminating the connection should inform the network to release resources associated with its connection. The dynamics of this process is referred to as UNI signalling.

Table 1. Physical layer interfaces for ATM

	Transmission rate (Mbps)	Throughput (Mbps)	System	Medium	Campus /WAN
DS-1 (T-1)	1.544	1.536	PDH	Coax	Both
E-1	2.048	1.92	PDH	Coax	Both
DS-3 (T-3)	44.736	40.704	PDH	Coax	WAN
E-3	34.368	33.984	PDH	Coax	WAN
E-4	139.264	138.24	PDH	Coax	WAN
SDH STM-1 Sonet STS-3c	155.52	149.76	SDH	Single mode fiber	WAN
SDH STM-4c Sonet STS-12c	622.08	599.04	SDH	Single mode fiber	WAN
FDDI-PMD	100	100	Block coded	Multi-mode fiber	Campus
Fiber channel	155.52	149.76	Block coded	Multi-mode fiber	Campus
Raw cells	155.52	155.52	Clear channel	Single mode fiber	WAN
Raw cells	622.08	622.08	Clear channel	Single mode fiber	WAN
Raw cells	25.6	25.6	Clear channel	UTP-3	Campus
Raw cells	51.84	49.536	Sonet	UTP-3	Campus
STS 3-C	155.52	149.76	Sonet	UTP-5	Campus

Each ATM end station requires an ATM address for end stations to uniquely identify each other. Unique ATM addresses are also required by ATM networks to locate the destination end node(s). Private and public networks use different ATM address formats. Public ATM networks use E.164 addresses (i.e., telephone numbers) whereas ATM private network addresses are based on the OSI network service access point (NSAP) format.

The ATM Forum UNI 3.1 specification includes a client registration mechanism that allows end stations to exchange their station identifiers for the ATM address information configured at the switch port. As a result of this exchange, the end station automatically acquires the ATM network address of the switch port it is attached to without any requirement for that address to be manually provisioned into the end station. Similarly, the end station part of the address is registered in the network and it is associated with its respective network part after the exchange.

B-ISDN services are categorized into eight possible service classes of which four are explicitly defined as illustrated in Table 2.

Table 2. B-ISDN service classes

Class	Timing relationship	Bit rate	Connection nature	AAL
A	Yes	Constant	Connection oriented	AAL 1
B	Yes	Variable	Connection oriented	AAL 2
C	No	Variable	Connection oriented	AAL 3/4 and AAL 5
D	No	Variable	Connectionless	AAL 3/4 and AAL 5

In addition, Class X service is defined to allow either a vendor specific AAL or as an option to bypass the adaptation layer and go directly to the ATM layer. In the former case, class X service allows proprietary ATM adaptation layer (AAL) used by "non-standard" applications. In the later case, applications sit directly on top of the ATM layer and pass the 48 byte payloads to the ATM layer by themselves. UNI 3.1 supports classes A, C, and X with AAL types 1, 3/4 and 5.

When a connection request is accepted by the network, the user and the network agrees upon a traffic contract for the duration of the connection. With this contract, the network guarantees the requested service demand of the connection as long as the source traffic stays within the specified limits. Accordingly, a traffic contract includes traffic descriptors, service requirements, and the conformance definition. Each one of these are required to be well defined and understood by both the network and the user.

A traffic parameter is a specification of a particular traffic aspect of the requested connection. The four parameters currently defined are the peak cell rate (PCR), cell delay variation tolerance (CDVT), sustainable cell rate (SCR), and burst tolerance (BT).

The peak cell rate of a connection is the inverse of the minimum time between two cells submitted to the network. When cells from different sources are multiplexed into the physical media, the cells of a particular connection may be delayed at the end station physical and/or ATM layer while cells from other connections are being transmitted. This causes peak cell rates of connections to change as seen by the switch across the UNI. CDVT specifies how much deviation from the peak cell rate is allowed at the UNI (i.e., as observed by the network).

The average rate of a connection is equal to the total number of cells transmitted divided by the duration of the connection. Based on this definition, the network can know the average rate of a connection only after the connection terminates and it can not be used by the network until the connection is terminated. Sustainable cell rate is an upper bound on the average rate of an ATM connection. SCR is used together with another metric, the burst tolerance (BT), and the PCR. BT is the duration of the period the source is allowed to submit traffic at its peak rate.

The cell loss priority (CLP) bit at the ATM cell header defines two loss priorities: high priority (CLP=0) and low priority (CLP=1). Network polices arrival cell streams based on the traffic contracts of connections to determine whether or not connections stay within the source traffic parameters agreed-upon at the call establishment phase. When a cell is detected to be non-conforming then there are two choices: either drop the cell at the interface or allow the cell enter the network, hoping that there might be enough resources to deliver the cell to its destination. In the latter case, it is necessary to make sure that non-conforming cells do not cause degradation to the services provided to conforming connections. Tagging is allowing non-conforming cells with CLP=0 to enter the network with CLP=1 (i.e., CLP bit is changed from zero to one). Cells with CLP=1 are discarded in the network whenever necessary so that the service provided to the conforming traffic (i.e., in this context CLP=0 traffic) is not affected.

Based on this framework, the allowable combinations of traffic parameters in UNI 3.1 are defined in Table 3.

In addition to the service requests in which a traffic contract is agreed upon during the connection set up, ATM networks support *best effort* service whereby no explicit guarantees are neither required nor negotiated between the user and the network. For this class of service, the only traffic parameter used is PCR with CLP=0+1.

Users at the set up phase request an ATM layer quality of service (QoS) selected from the QoS classes the network provides for ATM layer connections. A QoS class has either specified performance parameters (specified QoS class) or no parameter is specified (unspecified QoS class). The former provides a quality of service to an ATM connection in terms of a subset of performance parameters defined next. In the latter case, there is no explicitly specified QoS commitment on the cell flow. This class is intended to be used for best effort service.

Currently defined classes at the UNI are:

- unspecified QoS class 0: supports "best effort service" in which no explicit service requirement is specified;

- specified QoS class 1: supports a QoS that will meet Class A service performance requirements (i.e., circuit emulation);
- specified QoS class 2: supports a QoS that will meet Class B service performance requirements (i.e., VBR video);
- specified QoS class 3: supports a QoS that will meet Class C service performance requirements (i.e., Frame relay interworking);
- specified QoS class 4: supports a QoS that will meet Class D service performance requirements (i.e., IP over ATM).

UNI 3.1 interim local management interface (ILMI) provides an ATM user with the status and configuration information concerning (both VP and VC) connections available at its UNI. Configuration information at the ATM layer includes the size of the VPI and VCI address fields that can be used, number of configured VPCs and VCCs, and the maximum number of connections allowed at the UNI. The ILMI communication protocol is based on the SNMP network management standard. The term interim refers to the usage of this interface until related standards are completed by the standards organizations.

Table 3. Allowable combinations of traffic parameters in signalling messages

Combination	Traffic parameters
1	PCR for CLP=0 PCR for CLP=0+1
2	PCR for CLP=0 PCR for CLP=0+1 with tagging requested
3	PCR for CLP=0+1 SCR for CLP=0 BT for CLP=0
4	PCR for CLP=0+1 SCR for CLP=0 BT for CLP=0 with tagging requested
5	PCR with CLP=0+1
6	PCR for CLP=0+1 SCR for CLP=0+1 BT for CLP=0+1
Best effort service	PCR for CLP=0+1

3. DATA EXCHANGE INTERFACE (DXI)

The physical layer interfaces for ATM over wide area networks included (at the time the DXI developed) DS-3, OC-3, and OC-12. Although it is envisioned that in the near future a large portion of the physical infrastructure will be fiber, currently these interfaces are still expensive and it is necessary for the success of ATM to enable ATM services within the current infrastructure. The main objective of DXI is to provide access to ATM networks for installed equipment without costly (hardware) upgrades. Towards this goal, DXI allows a data terminal equipment (DTE), i.e. a router, and a data communications equipment (DCE) usually called ATM-data service unit (ATM-DSU) to cooperate to provide a UNI for ATM networks. The DXI framework defines the protocols for a DTE to transport a DTE-service data unit (DTE-SDU) to a corresponding peer entity via an ATM network, as illustrated in Fig. 2.

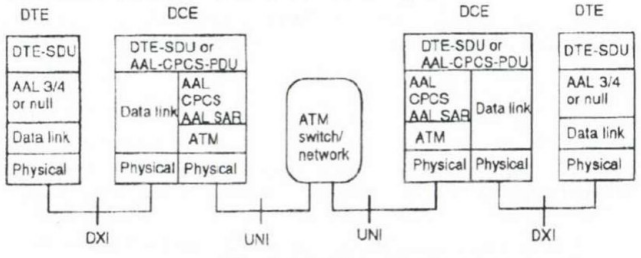


Fig. 2. DXI framework

DXI defines a data link control protocol and physical layers which handle data transfer between a DTE and a DCE. DXI local management interface (LMI) and management information base (MIB) are also specified as part of the DXI.

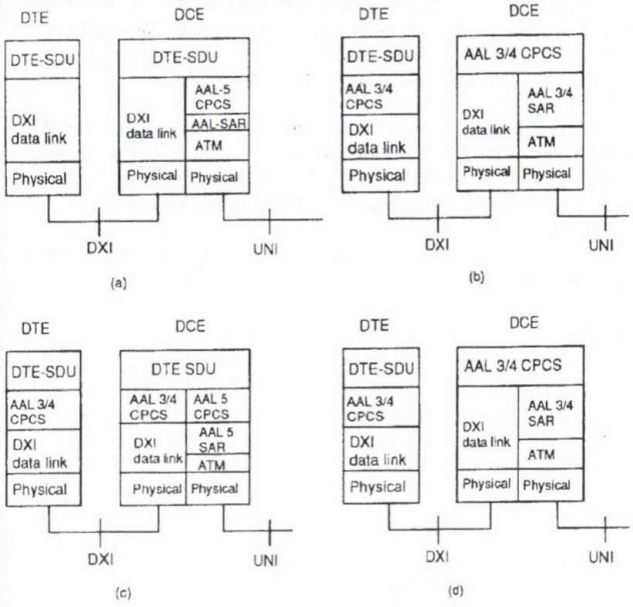


Fig. 3. Three operational modes of DXI

DXI supports V.35, RS449 and HSSI physical layer interfaces at speeds ranging from several Kbps to 50 Mbps. The physical layer interface used between the DSU and the ATM network across UNI can be any one of the physical layers specified in UNI 3.1.

The data link layer defines the method by which the DXI frames and their associated addressing are formatted for transport over the physical layer between the DTE and the DCE. The protocol is dependent on the mode selected. In particular, across a DXI, three operational modes are defined across a DXI: modes 1a, 1b, and 2.

Mode 1a

In this mode, transport of DTE-SDU is based on AAL 5 common part convergence (CPCS) and segmentation and reassembly (SAR) sublayers as illustrated in Fig. 3.

At the origination node, data link control layer receives DTE-SDU and encapsulates into DXI data link control frame (DXI-PDU) as illustrated in Fig. 4. The resulting PDU is transmitted to DCE across the DXI. The DCE strips off the DXI encapsulation and obtains the values of DXI frame address (DFA) and CLP. DCE then encapsulates the DTE-SDU into AAL 5 CPCS PDU and segments the resulting PDU into 48 byte AAL 5 SAR-SDUs. DCE also maps the DFA to the appropriate VPI/VCI of each cell. CLP bit value at the DXI header is also copied to the CLP bits of the transmitted cells.

For data transmission from the ATM network to the destination DTE, the reverse process is followed. The only exception is the use of the congestion notification (CN) bit at the DXI header which is set to one if one or more cells of the DTE-DU experienced congestion in the network as specified in their cell header, EFCI bit.

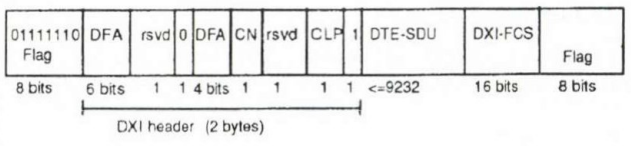


Fig. 4. DXI data link PDU format in mode 1a; where DFA: DXI frame address; CN: congestion notification; CLP: cell loss priority; rsvd: reserved; DXI-FCS: DXI frame check sequence

Mode 1b

This mode consists of mode 1a plus transport of DTE-SDU service based on AAL 3/4, as illustrated in Fig. 3. When AAL 3/4 is used, DTE first encapsulates the DTE-SDU into AAL 3/4 CPCS PDU by appending the corresponding CPCS header and trailers to the DTE-SDU. The CPCS-PDU is then encapsulated into DXI data link control frame, similar to mode 1a.

The resulting frame is transmitted to the DCE, which strips off the DXI header and DXI FCS thereby obtaining the DFA and CLP bit. The DCE segments the received frame into AAL 3/4 SAR-SDUs and translates the DFA into appropriate VPI/VCI value. CLP bit value is also copied from the DXI header to the cell header.

For data transmission from the ATM network to the destination DTE, the reverse process is followed, similar to mode 1a.

Mode 2

In this mode, DTE operations are the same as in mode 1.b with AAL 3/4. That is, DTE encapsulates the DTE-SDU into AAL 3/4 CPCS PDU. The resulting PDU is then encapsulated into DXI data link control frame and it is transmitted to the DCE.

The main difference of mode 2 compared with modes 1a and 1b are the use of CRC-32 (as opposed to CRC-16) and that the DXI header in mode 2 is 4 bytes long as illustrated in Fig. 5.

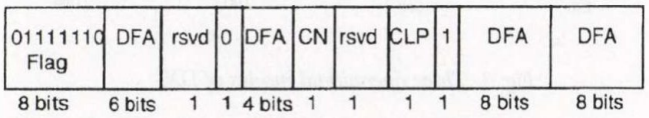


Fig. 5. DXI header in mode 2

At the DCE, both AAL 5 and AAL 3/4 connections are allowed. If AAL 5 is used then DCE strips off both the DXI and AAL 3/4 CPCS encapsulations. The remaining PDU (i.e. DXI-DSU) is then encapsulated into AAL 5 CPCS PDU, segmented into AAL 5 SAR-SDUs and transmitted to the ATM network using the services of ATM and physical layers, similar to the mode 1a (and 1b with AAL 5).

If AAL 3/4 connection is used, then DCE strips off only the DXI encapsulation and segments the AAL 3/4 CPCS PDU into AAL 3/4 SAR-SDUs, which are transmitted to the ATM network through ATM layer.

3.1. Data Exchange Interface Local Management Interface

The ATM DXI is managed by the DTE through a local management interface (LMI). Some of the ATM UNI ILMI messages are also exchanged between the DTE and the ATM switch the DCE is attached to. DCE does not pass the UNI status back to the DTE.

LMI defines the protocol for exchanging management information across a DXI. It is designed to support a management station running SNMP and/or switch running the ILMI protocol. LMI supports the exchange of DXI specific, AAL specific, and ATM UNI specific management information.

The LMI allows the DTE to set or query the operation mode of the DXI as well as the AAL assigned on a per VCC basis.

4. BROADBAND INTEREXCHANGE INTERFACE (B-ICI)

End-to-end national and international service requires networks belonging to different carriers to be interconnected. B-ICI gives ATM carrier networks the ability to interoperate towards transporting different services across each other. B-ICI version 1.1 supports permanent connections only. B-ICI specification for switched virtual connections (i.e., B-ICI signalling) is expected to be completed mid 1995.

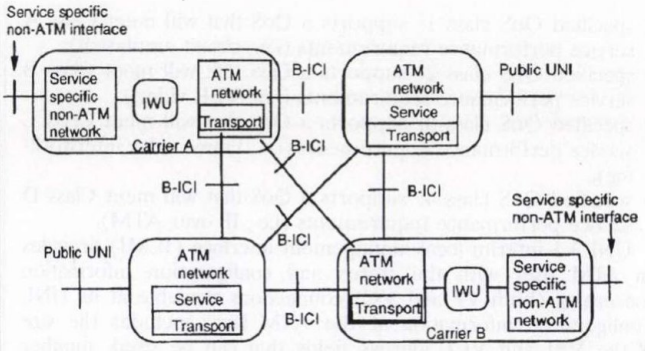


Fig. 6. A generic view of service specific networks inter-connected via ATM networks

Fig. 6 illustrates a generic example of carrier networks having service specific UNIs connected to other carrier ATM networks. B-ICI specification includes physical layer, ATM layer, and service specific functions above the ATM layer required to transport, operate, and manage a variety of intercarrier services across a B-ICI. The document also includes traffic management, network performance, and, operations and maintenance specifications.

The physical layer of the interface is based on ITU-T defined network node interface (NNI) which includes SONET/SDH physical and ATM layers with the addition of DS-3 physical layer.

The initial B-ICI is a multi-service interface that supports the following inter-carrier services:

- Cell relay service (CRS).
- Circuit emulation service (CES).
- Frame relay service (FRS).
- Switched multi-megabit data service (SMDS).

A service specific non-ATM network is connected to an ATM network via an interworking unit (IWU) that allows non-ATM service to be mapped into ATM service and vice versa. The main function performed at an IWU for each service supported across a B-ICI is as follows:

- CRS: receive ATM cells from one network and transmit to another over a permanent connection.
- CES: receive DSn frames encapsulated in AAL 1 PDUs, transmit them over PVCs across a B-ICI to another network, and reconstruct the original DSn frames at the other end.
- FRS: receive Frame Relay frames, encapsulate them in AAL 5 PDUs, transmit ATM SDUs over PVCs across a B-ICI to another network, and reconstruct the original frames at the other end.
- SMDS: receive SMDS interface protocol (SIP) L3 PDUs encapsulated in intercarrier service protocol connectionless service (ICIP-CLS) PDUs, encapsulate them in AAL 3/4, transmit ATM SDUs over PVCs across a B-ICI to another network, and reconstruct the ICIP-CLS PDUs at the other end.

These ATM cells of a particular service are multiplexed together and passed across B-ICI over one or more VPC and/or VCC that are pre-configured at subscription time. This multiplexing is done at the interface level. That is, for each service supported, there is at least one connection and cells belonging to different services are not multiplexed onto the same connection.

Traffic management and congestion control framework across B-ICI follows very closely the framework developed in UNI specification. The traffic parameters and connection traffic descriptor definitions in UNI specification also applies to B-ICI, with some simplifications. For example, source traffic descriptor across B-ICI is required to include service type, conformance definition, peak cell rate and cell delay variation tolerance while the inclusion of sustainable cell rate and burst tolerance in the traffic contract is an option. Network parameter control (NPC) function monitors and controls offered traffic and the validity of the ATM connection. Its main purpose is to protect network resources and the QoS of connections already established in the network during periods of congestion which might be caused by equipment malfunction or misoperation. Each carrier may use any implementation of NPC function as long as it does not violate the QoS objectives of a valid and compliant connection (or may not use any NPC function at all).

The interconnectivity across a B-ICI is obtained by establishing a VCC and/or a VPC at the B-ICI with a QoS suitable to meet the service requirements for end-to-end connections. Current B-ICI specification version 1.1 supports PVCs only. That is, both the source and destination points of each PVC are predefined and they are fixed for the duration of the connection. A consequence of this is that there is a fixed route between two interworking units communicating to each other across two or more public networks.

PVC based inter-carrier support for each service supported across a B-ICI is reviewed next.

Cell relay service (CRS)

CRS is a cell based information transfer service that offers its users direct access to the ATM layer at rates up to the access link rate. Both VPC and VCC connections are supported. In CRS, a point-to-point PVC denotes an ATM layer VPC or VCC from a source ATM service access point (ATM-SAP) to a destination ATM-SAP, as illustrated in Fig. 7.

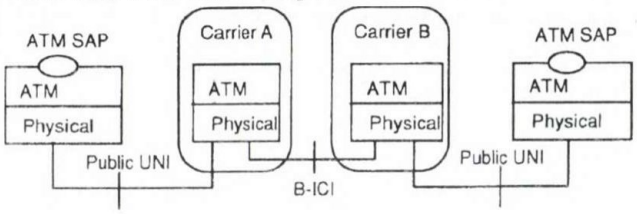


Fig. 7. CRS across a B-ICI

Circuit Emulation Service (CES)

CES supports the transport of continuous bit rate (CBR) signals using ATM technology. Fig. 8 shows examples of CES and its role in supporting CBR services.

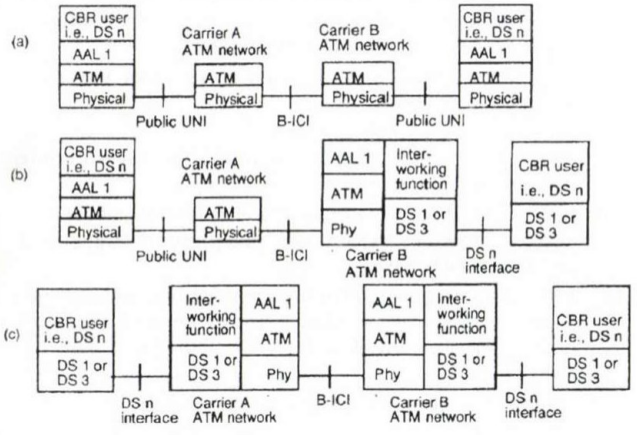


Fig. 8. Circuit Emulation Service

Two end stations support voice service over ATM using AAL 1 and generate DSn frames. The B-ICI supports the transport of DSn signals across two public networks thereby connecting two ATM end stations with UNIs at each end. In Fig. 8, one user is connected to an ATM network via a UNI whereas a DS-n interface is used at the other end. In this case, the end station that supports voice over ATM communicates with another end station possibly attached to a voice network through an interworking unit. The IWU provides interoperability between a native DS-n network and an ATM network. Finally, in Fig. 8, both users are connected with DS-n interfaces, requiring interworking functions at both ends.

Frame Relay Service (FRS)

FRS is a connection oriented data transport service. Similar to CES, two frame relay interworking scenarios are illustrated in Fig. 9. These are two network interworking scenarios originally defined by ITU-T and are adopted by the forum. In the first scenario, two FR networks/CPE are connected via an ATM network whereas scenario 2 connects a FR network/CPE with a broadband CPE emulating FR.

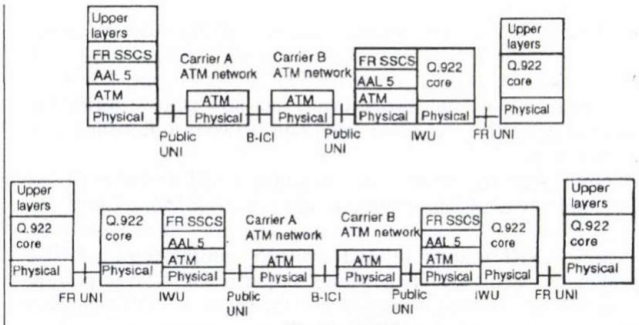


Fig. 9. FES

Switched Multimegabit Data Service (SMDS)

SMDS is a public packet switched service that provides for the transport of data packets without the need for call establishment procedures. Customer access to the SMDS will be over the SMDS subscriber network interface (SNI). The carrier network provides an internetworking function for the encapsulation of the SMDS L3_PDU within an ICIP_CLS PDU. AAL 3/4 is used to support the transfer of these PDUs within ATM cells. Fig. 10 shows an example of SMDS/ATM interworking functions.

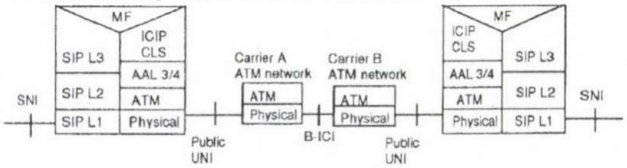


Fig. 10. SMDS/ATM interworking; where MF: mapping function

Private NNI

Private NNI (P-NNI) specification is currently being worked on at the ATM Forum. The specification includes PNNI routing and PNNI signalling.

P-NNI routing is used to find a path across a network between two end stations (point-to-point connection) or two or more end stations (point-to-multipoint connection). In P-NNI routing, the switching system that a connection request originates across its UNI is responsible for finding the end-to-end path to the destination end station. This is referred to as source routing. In determining the path, the originating switching system uses link-state routing in which each switching system advertises information about its P-NNI links to other switching systems. In this context, a P-NNI link connects one switching system to another (in a given direction) across a P-NNI.

For each connection request, the source switch finds a path based on the advertised capabilities and the desirability of other switching systems to carry connections with different characteristics. After finding the path, the originating switching system uses P-NNI signalling to request connection establishment from intermediate switching systems along the path. The sequence of switching systems visited is specified in the transit designated list (DTL) stack included in the corresponding P-NNI signalling message. Each switching system processes received connection request messages, make connection admission decisions (i.e., accept or reject), and passes the signalling message to the next switching system along the path (if accepted) or denies the connection. Connection admission decisions and path selection algorithm are vendor specific (i.e., proprietary) and are not expected to be standardized. However, the source switching system should be able to predict the outcome of a possibly unknown admission procedure at a switch (and/or a switching system) with "some" confidence so that a "large" portion of connection set up requests under normal operating conditions are successful (i.e., result in connection establishment). A generic call admission control procedure specified in P-NNI for this purpose.

The choice of what internal state information to advertise, how often, and to where requires the specification of a multi-level hierarchical routing model. In P-NNI routing, the number of hierarchical levels is not pre-set and may vary from one level (i.e., no use of hierarchy) up to 104. A very large corporate network

may be expected to vary between two to four hierarchical levels of routing. The P-NNI hierarchical model explains how each level of hierarchy works, how multiple nodes at one level can be summarized into the higher layer, and how state information among nodes within the same level and between different levels are exchanged.

P-NNI signalling defines the procedures and messages used to dynamically establish, maintain, and release ATM connections across P-NNI between two switching systems. It is also used to carry P-NNI routing related protocol information. P-NNI signalling is based on UNI signalling. In particular, a large portion of the P-NNI signalling messages are the same as UNI signalling messages. In addition, P-NNI signalling messages are used to carry topology and routing information among P-NNI nodes.

Another feature in P-NNI signalling is the incorporation of a crankback function. In particular, there is no guarantee that an intermediate network along the path will have the resources to support the requested connection. Related information when a call connection request is rejected (i.e., insufficient network resources) may be sent towards the originating network for alternate routing. An intermediate switching system along the path may attempt another route based on its view of the topology. This function is referred to as crankback.

ITU-T NNI

ITU-T standardization on NNI include the specifications of physical and ATM layers, control, user, and management planes, and NNI signalling (B-ISDN user part, i.e. B-ISUP) in the context of public networks. Unlike P-NNI, however, ITU-NNI specification does not include any routing protocol.

B-ISUP provides signalling functions required to support basic bearer services and supplementary services for B-ISDN applications. It is being developed for international applications as a network-to-network interface. However, it is also suitable for national applications as a network-node interface. In particular, coding space is reserved in order to allow national administrations and private operating agencies to introduce network specific signalling messages and elements of information within the standardized protocol structure.

The B-ISUP procedures and operation is rather involved. The elements of signalling messages and their functions used by B-ISUP are currently defined to support basic bearer services and supplementary services for capability set 1, which includes the following features.

- Point-to-point Demand (switched virtual) channel connections.
- Connections with symmetric or asymmetric bandwidth requirement.
- Single connection call.
- Basic signalling functions via protocol messages, information elements, and procedures.
- Class X, class A, class C ATM transport services.
- Request indication of signalling parameters.
- VCI negotiation.
- Out-of-band signalling for all signalling messages.
- Error recovery.
- Public UNI addressing formats for unique identification of ATM end points.
- End-to-end compatibility parameter identification.
- Signalling interworking with ISDN and provision of ISDN services.

REFERENCES

[1] R. O. Onvural: *ATM Networks: Performance Issues*, Artech House, Norwood, MA, 1994.

[2] *ATM Forum Specifications*

[2.1] UNI 3.1: ATM *forum user-network interface specification* include specifications of physical layer interfaces, the ATM layer, interim local management interface, and UNI signalling.

[2.2] B-ICI: *B-ISDN intercarrier interface* defines a carrier to carrier interface which specifies various physical layer interfaces and supports multiplexing of different services that include SMDS, frame relay, circuit emulation, and cell relay.

[2.3] DXI: *Data exchange interface* defines data link control and physical layers as well as interim local management interface specifications that allow existing routers to interwork with ATM networks without requiring special hardware.

[3] *ITU-T Recommendations*

[3.1] Q.2931: *B-ISDN signalling* defines the signalling procedures and functions used at UNI.

[3.2] Q.2100: *Q.SAAL.0 B-ISDN signalling ATM adaptation layer overview description* describes the different components that make up the AAL functions necessary to support B-ISDN signalling.

[3.3] Q.2110: *Q.SAAL.1 SSCOP specification* specifies the peer-to-peer protocol for the transfer of information and control between any pair of SSCOP entities.

[3.4] Q.2130: *Q.SAAL.2 B-ISDN signalling ATM adaptation layer- SSCF for the support of signalling at the UNI* defines a mapping between the SSCOP of the signalling AAL and the Q.2931 layer at the user-network interface.

[3.5] Q.2140: *Signalling at the NNI* defines a mapping between the SSCOP of the signalling AAL and the Q.2931 layer between network nodes and between networks.

[3.6] Q.2120: *Meta-signalling* defines the meta-signalling protocol used to establish and maintain signalling connections at the user-network interface.

[3.7] I.113: *Vocabulary of terms for broadband aspects of ISDN* provides the definitions of terms considered essential in understanding the principles of B-ISDN.

[3.8] I.121: *Broadband aspects of ISDN* presents the basic principles of B-ISDN.

[3.9] I.151: *B-ISDN asynchronous transfer mode functional characteristics* presents a high level review of the asynchronous transfer mode layer functions.

[3.10] I.211: *B-ISDN service aspects* classifies broadband services into a number of service classes and presents examples of services in each class.

[3.11] I.311: *B-ISDN general network aspects* defines and presents a high level review of ATM transport network hierarchy and the specification of each layer including signalling.

[3.12] I.321: *B-ISDN protocol reference model and its applications* defines the B-ISDN reference architecture.

[3.13] I.327: *B-ISDN functional architecture* presents a review of B-ISDN functions, where they are located, and where they are distributed in the B-ISDN.

[3.14] I.356: *Quality of service configuration and principles.*

[3.15] I.361: *B-ISDN ATM layer specification* addresses the ATM cell structure, header coding, and the ATM protocol structure.

[3.16] I.362: *B-ISDN ATM adaptation layer (AAL) functional description* presents the AAL functional organization and the basic principles of AALs.

[3.17] I.363: *B-ISDN ATM adaptation layer (AAL) specification* provides the details of various types of AALs defined to support different B-ISDN service classes.

[3.18] I.364: *Support of broadband connectionless data service* on B-ISDN describes the support of a connectionless data service based on AAL type 3/4 on B-ISDN.

[3.19] I.371: *Traffic control and congestion control in B-ISDN* presents the objectives and mechanisms of traffic control and congestion control, defines traffic descriptors, traffic parameters and traffic contract.

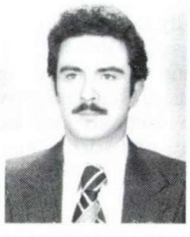
[3.20] I.374: *Network capabilities to support for multimedia.*

[3.21] I.413: *B-ISDN user-network interface* presents the reference configuration for the B-ISDN user-network interface and various examples of physical realizations.

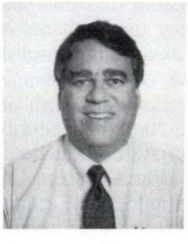
[3.22] I.432: *B-ISDN user-network interface- physical layer specification* defines the physical media, the transmission system and discusses the implementation of related OAM functions at the user-network interface.

[3.23] I.555: *Interworking* defines narrowband ISDN as well as Frame Relay interworking with an ATM network.

[3.24] I.610: *B-ISDN operation and maintenance principles and functions* identifies the minimum set of functions required to operate and maintain physical layer and ATM layer aspects of B-ISDN user-network interface as well as the individual virtual path and virtual channel connections routed through a B-ISDN.



Raif O. Onvural is a senior engineer at IBM's Research Triangle Park facilityb in the Networking Architecture organization and manages the Networking Technology Architecture department that develops network control services for ATM networks. He is IBM's venue owner to ATM Forum and has been attending the Forum meetings since February 1993. He organized several international conferences on high speed networks in general and ATM networks, in particular. He has published in various journals and conferences and has edited six books. He is also the author of the book Asynchronous Transfer Mode Networks: Performance Issues. He is a member of ACM, ORSA, IFIP TC6.3., and IEEE. Currently he is the interim chair of IEEE Technical Committee on Computer Communications.



Gerald A. Marin is the director of Networking Architecture. His organization is responsible for the design and development of IBM's high speed, multimedia architecture, Networking BroadBand Services (NBBS), enhancements to Advanced Peer-to-Peer Networking (APPN) architecture, Multiprotocol Transport Networking (MPTN), and for Advanced Program to Program Communication (APPC) and Common Programming Interface for Communications (CPI-C). In addition, his area has significant responsibility in representing IBM in standards organization (ITU-T, T1, IEEE 802) and in industry organizations (ATM Forum, Frame Relay Forum, IETF, APPN Implementors Workshop). He is also an adjunct professor, Duke University, Department of Electrical Engineering, and has published in various international conferences and journals.

INFORMATION SECURITY ASPECTS OF ASYNCHRONOUS TRANSFER MODE NETWORKING

D. STEVENSON, and F. TERHAAR-YONKERS

MCNC
RESEARCH TRIANGLE PARK, NORTH CAROLINA
P. O. BOX 12889
USA

The authors provide a basic overview of security issues associated with ATM networks. Vulnerabilities, approaches for dealing with them, standards efforts and current research efforts are included in the discussion.

1. INTRODUCTION

The emerging standardization and use of ATM, like the introduction of any other new network technology, has resulted in the re-examination of associated information security issues. ATM poses some unique security problems, in part due to the higher speed of operation that it offers and also due to the unique way that ATM systems work. We can classify network security issues into several categories. These include privacy, access control, data integrity, and denial of service attacks. Despite a perception that optical fiber is inherently more secure than copper based physical media, all these issues apply to ATM networks.

2. PRIVACY ATTACKS

It was once suggested to us that it is difficult to gain physical access to public carrier facilities and therefore there was a small danger for traffic interception. But there are a variety of privacy attacks that can be identified. It is possible to tap fiber, although it is technically more challenging than copper cables. It is possible to attack ATM switches causing them to redirect traffic via a different route that may have been compromized (we use this method at MCNC to demonstrate ATM vulnerabilities by tapping our own network). It is also possible to attack ATM traffic through the SONET infrastructure causing it to traverse a compromized portion of the network by redirecting SONET streams. Additionally, some SONET systems are designed to utilize line of sight microwave transmission and satellite technology rather than fiber optics. Such systems are relatively vulnerable to physical wiretapping.

In addition to public network vulnerabilities there are many settings where ATM LANs are being used and may be also subject to attack. Often for large organizations, multiple ATM LANs are interconnected to form private networks in a campus environment. If this infrastructure can be physically accessed it is possible to expose much or all of the ATM traffic associated with that organization.

3. PRIVACY PROTECTION

The classic defense against privacy attacks is data encryption. Data encryption can take place at many different places relative to the ISO seven layer protocol stack including the ATM layer. The richness of multiplexing possibilities provided by ATM and the need to provide separation of different communication sessions sharing the same facilities leads to the concept of key agile encryption. A key agile ATM encryptor is a system that encrypts the payload of ATM cells using a different key for each active virtual circuit. Several efforts have been undertaken to accomplish exactly this. Researchers at the University of Pennsylvania developed a network adapter for an IBM workstation with an embedded cell encryption capability [19]. The National Security Agency developed a prototype cell encryption system that was the basis for the GTE Government Systems Fastlane system [2]. These systems are both based on classified technology and target military applications. Sandia Labs [18] has developed a prototype ATM encryption system for Department of Energy security applications. Additionally MCNC has developed a prototype cell encryption system [20], [21].

The ATM Forum is in the process of defining signalling support standards for key management. It will likely take some period of time before these standards are defined, implemented and widely fielded. Until this time it will be necessary to utilize methods for key management that are transparent to ATM standards. This is the approach used in the MCNC prototype system.

A Forum contribution [8] identifies several methods that might be used for cell encryption including the use of block striping implementation (each 64 bit block within the cell is encrypted with a different encryption chip). We have found that this method is not necessary to achieve high speed cell encryption. Consequently, block striping is undesirable since it leads to an unnecessary proliferation of encryption options and greater system expense (more encryption devices and more state data is required). Excessive system expense for block striping approaches is especially significant for systems where large numbers of active keys are supported.

The choice of encryption algorithm is a topic that seems to involve as many electro-political considerations as technical. There are many methods to chose from leading to significant compatibly issues. One method that can reduce the complexity of this situation is algorithm agility, whereby the encryption system can switch algorithms as well as keys on a cell by cell basis. This functionality permits each VC at session setup time to select from a set of algorithms depending on what is available at the other end of the connection. The algorithm might be set by negotiation at call setup, or through security policies determined off-line by the system managers.

4. ACCESS CONTROL ISSUES

As in other network technologies there are significant access control issues for ATM. Most of these issues center around identification of end users seeking to gain access to an ATM connected resource. The existing ATM signalling standard does provide for a modest means of end user identification. Specifically, the SETUP message contains a field for the user to supply the associated E.164 address (essentially the phone number for the line being used). The field is optional. It is specific only to a communications line, not an individual. And it is computationally easy, if technically obscure, for a user to supply false data.

We can envision three levels of calling party identification including, anonymous caller, user supplied identity, unforgeable user identification. We assume that there are many cases where having strong assurances of the identity of the calling and called user (a person, rather than merely a line number) is desirable. Any point-to-point communications involving exchange of sensitive information would fall into this category. In addition, individuals joining existing multi-party conferences should generally be expected to provide some measure of proof of identity. The mechanisms provided for party identification in the existing specifications for call management are weak, since they are based on the phone number of the calling party. A called user has no way to determine the accuracy of this information when it is supplied. There has been some public discussion about methods to remedy this situation.

Several ATM Forum contributions are relevant to this security issue [8]–[10]. They involve proposed additions to the signalling messages to provide assurances as to their origin and integrity through the use of digital signatures. These methods may be used to provide strong proof of identity of the end parties in a connection.

5. DENIAL OF SERVICE ATTACKS

There are a great many denial of service attacks on networks. The most intuitive of these would involve application of explosives or backhoes to the physical infrastructure providing the network services. These types of physical brute force attacks are easily detected because of their obvious nature. Additionally, where there is serious threat from these attacks, redundancy and geographic dispersion of nodes and links can provide significant

protection. There are more subtle attacks that are possible however.

If an adversary has access to the ATM signalling channel (through use of the various privacy attacks mentioned above), it is possible to make a large number of subtle attacks against the network itself or subsets of users. For instance, by monitoring and analyzing the signalling information traversing a link it is possible to identify the traffic associated with a specific end node or pair of end nodes. It would then be possible to disrupt service associated with this node by corrupting the signalling information in some way, or by attacking the user cells directly (by such means as selective bit error insertion, or cell dropping). These attacks would not be prevented by use of network redundancy and not be detected by the network operators or end users.

6. PROTECTION AGAINST DENIAL OF SERVICE ATTACKS

The ATM Forum contributions mentioned above are relevant to denial of service attacks. By attaching digital signatures to the signalling messages the network can provide cryptographically strong means of detecting any signalling message tampering. The method has the potential of significantly increasing the computational burden for call setup both for the network and end user equipment. This is especially true for an exportable signature implementation like the Digital Signature Standard (DSS) where execution times are significant. DSS [1] is a US Government proposed cryptographic method for verifying the integrity and source of unclassified information. The most significant shortcoming of the ATM Forum proposed approach from a denial of service perspective is that it provides for detection of tampering but does not protect against it.

An effective way to protect against subtle denial of service attacks based on traffic analysis is to prevent the attacker from having access to traffic data [17]. To prevent traffic analysis based denial of service attacks one must prevent the attacker from being able to read the ATM signalling messages. Since the VPI/VCI field in ATM cells does not have end to end significance without being able to intercept signalling traffic, it becomes very difficult for an attacker to associate cell streams with specific end users. An effective, although expensive, method of accomplishing this is to encrypt all the data on each link through the use of bulk link encryption. In such an approach, each protected link in the network would utilize an encryption unit at each end to encrypt all data traversing it. Such a system would reasonably operate at the SONET level encrypting SONET frames, leaving the overhead information untouched. Such an approach would be transparent to SONET repeaters. A more sophisticated approach would be necessary if it is desired to be transparent for SONET multiplexing or drop add systems. With SONET link encryption systems deployed it would not be possible to intercept or even identify the signalling information so that traffic analysis based denial of service attacks would be prevented. It would also be possible to prevent certain denial of service attacks by simply encrypting the signalling channel.

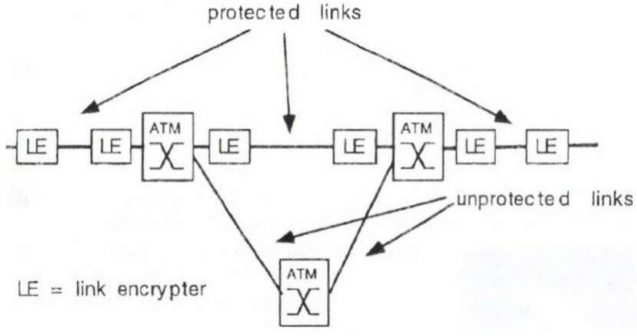


Fig. 1.

7. TRUSTED THIRD PARTY ISSUES FOR ATM

In the US there is considerable interest (and controversy) about key escrow systems, which are also referred to as trusted third party systems. The salient feature of a key escrow system is the ability of authorized entities to decrypt information through the use of Data Recovery Keys (DRK) held by trusted parties. The scenarios where data recovery keys are used classically include employers recovering essential data encrypted by an employee who is unavailable (for instance due to accidental death), or in cases of authorized wiretapping by law enforcement. Much has been written in the US about key escrow systems, especially the Clipper system, originally proposed by the National Security Agency. Denning [3] provides an enlightened description of these systems devoid of the usual historonics often accompanying other discussions. Hoffman [16] provides an even handed discussion of the cryptopolitical issues involved in key escrow systems.

In applications involving communications networks, the key escrow systems transmit an encrypted session key in the form of a Data Recovery Field (DRF). The Data Recovery Field can be used with the Data Recovery Key by authorized entities to recover the session key and access the plaintext data (see illustration below).

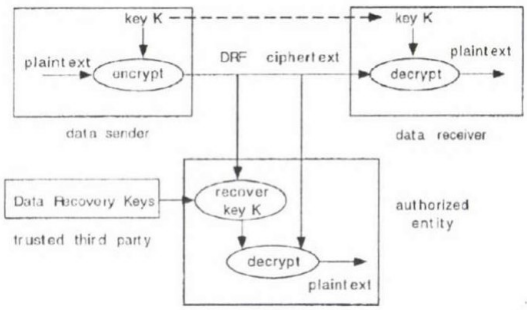


Fig. 2.

Without intending to comment on the crypto-political issues associated with key escrow systems, it is instructive to think about these systems in the context of ATM. First of all, most of the conceptual models for key escrow assume that there is a one to one correspondence between encryption devices and users, that is, the encryption devices are not shared by multiple users. It is therefore, possible for law enforcement to obtain keys from an escrow agent and capture the communications of the individual associated with a specific encryption device. It is also assumed that there is generally only a single, or at least a small number, of active sessions per encryption device. Consequently the encryption device can hold internally all the state information (including keys) needed for its operation.

For ATM this model begins to break down. Consider a cell encryption system that sits at the boundary between an ATM LAN and the ATM public network. In this system the traffic for many users will be processed by a single encryption device. Obtaining from a key escrow agent the keys associated with a single encryption device, in this case, will potentially expose the traffic of all the users. Consequently the concerns about potential misuse of key escrow information are amplified manyfold.

A second problem area is more technical in nature, specifically, the amount of state information associated with device being used for ATM encryption. For the MCNC key agile system for each encryption device up to 65,536 simultaneously sessions may be active each requiring 256 bits of state information for a total of 1 megabyte of state information. With the current state of the art for hardware design, it is impractical to have this much state memory on the encryption chip. Consequently, it becomes necessary to permit the state memory to exist off of the encryption chip. If the state memory exists off of the encryption chip then state data must be externally supplied to the device appropriately, for each session. This implies that the off chip logic supplying state information is also trusted to inform the encryption chip when it is necessary to form a Data Recovery Field (a DRF is desired only when a new session key is created, and is undesirable for cells using a previously established session key). This trust also represents a means of defeating key escrow. It may be possible to preclude such problems by utilization of multi-chip module packaging, to combine encryption and state memory into a single device incorporating physical protection.

8. SECURITY IMPLICATIONS OF EMERGING ATM STANDARDS

Recent additions to the current standards in the area of PNNI (Private NNI) and UNI-4.0 represent even more challenges to security.

PNNI routing introduces a new protocol, "HELLO" which allows dynamic discovery of hierarchical routing for complex private networks. Using HELLO, network elements exchange connectivity information in the clear to choose a PGL (Peer Group Leader). The PGL is then the switch allowed to interact with network elements outside the local peer group. The election of a PGL is based solely on a switch's address prefix. No authentication of these network elements is done. It seems likely (even in a private network) that a Certificate Authority (and necessary protocols) will be requisite to insure that switches have the authority to determine topology and hence routing. Without the confidence that the switch is allowed to participate in the network and at what level, all manner of mayhem may be perpetrated. Anything from spoofing of routes to transparently capture data, to causing the collapse of the network is possible by insertion of an illegal switch or by reprogramming one that is not physically secure.

Although PNNI routing is scaleable and deemed capable of providing routing for a global ATM Internet, the security implications appear intractable. The exchange of HELLO messages and subsequent flooding of connectivity information via PNNI Topology State Packets (PTSP) will be difficult to authenticate outside of an administrative domain. There was a recent proposal to the Forum to generate the protocols and mechanisms to "firewall" PNNI routing. This would allow "exterior routes" that are not generated by HELLO to be advertized and managed internally to the private network. Mechanisms to insure that HELLO and PTSPs do not cross PNNI boundaries will also be required.

PNNI routing manifests itself across network boundaries as PNNI signalling. PNNI signalling is subject to the same security problems and concerns as UNI call setup.

UNI 4.0 introduces a new traffic class for Available Bit Rate (ABR). The mechanism for the network to throttle ABR traffic is an end-to-end feedback loop. The source sends — in addition to the data stream a Resource Management (RM) cell as often as every 32 data cells. The RM cell is a special OA&M cell which carries bandwidth requirements and availability in it's payload. The contents of this cell may be modified by any switch in the path. In addition, switches may source RM cells, sending them directly back upstream in order to quench source traffic without waiting for the feedback loop.

The security implications here include denial of service by insertion of bogus RM cells as well as a covert channel between the SAR layers of the end systems. This scheme also breaks per VC link level encryption as the cell may and must be able to be modified in flight by any network element. It appears that an ATM firewall will have to terminate RM cells forming an additional feedback loop (virtual source and destination) between the crypto and end-system to guard against these risks.

9. CONCLUSIONS

While ATM networking holds much promise for the future, it is also true that many information security issues are inherent in the technology. Successful wide scale deployment and use of ATM will require effective solutions that address these problems. This in turn implies significant efforts in both the standards bodies and within the research and development community.

REFERENCES

[1] J. A. Adam: "Cryptography = privacy?" *IEEE Spectrum*, Special Report on Data Security, August 1992.

[2] Ch. Adams: "ATM Security Issues: Getting Down to Basics", *Infosecurity News*, July/August 1995.

[3] Critical Factors of Key Escrow Encryption Systems, *18th National Information Systems Security Conference*, Baltimore MD, Oct. 10-13, 1995.

[4] T. Smith and J. Stidd: "Xerox Corporation, Requirements and Methodology for Authenticated Signalling", *ATM Forum*/94-1213.

[5] B. Lyles: "Authenticated Signalling for ATM T1S1.5/94-118", Xerox PARC.

[6] M. Peyravian, G. Tsudik and E. Van Herreweghen: "IBM, A Certification Infrastructure for ATM", *ATM Forum*/95-0964.

[7] M. Peyravian and E. Van Herreweghen: "IBM, ATM Security Scope and Requirements", *ATM Forum*/95-0579.

[8] L. Pierson and T. Tarman: "Sandia National Laboratories, Requirements for Security Signalling", *ATM Forum*/95-0137.

[9] L. Pierson and T. Tarman: "Sandia National Laboratories, A Proposed Generic Authentication Information Element", *ATM Forum*/95-0460.

[10] L. Pierson and T. Tarman: "Sandia National Laboratories, Proposed DSS-Specific Fields for the Generic Authentication Information Element", *ATM Forum*/95-0461.

[11] R. Hilton, R. Prodan: "Cable Television Laboratories, Functional Requirements for Encryption", *ATM Forum*/95-0372.

[12] M. Lazer and Bellcore: "Framework for Developing Security-related Signalling Standards", *ATM Forum*/95-0408.

[13] Sh. Sathaye: "Fore Systems, ATM Forum Traffc Management Specification Version 4.0", *ATM Forum*/95-0013R6.

[14] R. Callony: "Bat Networks Inc., Exterior Routes in PNNI", *ATM Forum*/95-0042.

[15] D. Dykeman: "IBM Zurich, Mike Goguen, Bay Networks Inc., PNNI Draft Specification", *ATM Forum*/95-0471R10.

[16] L. J. Hoffman: "*Building in Big Brother, the Cryptographic Policy debate*", Springer-Verlag, 1995.

[17] R. M. Needham: "Denial of Service: An Example," *Comm. of the ACM*, Vol. 37, No. 11, pp. 42-46 (Nov 1994).

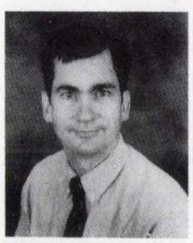
[18] Private communications with Lyndon Pierson, Sandia National Laboratories.

[19] J. Smith, B. Traw and D. Farber, "Cryptographic Support in a Gigabit Network", *Proc. INET'92*, June 15-18, 1992, Kobe Japan, pp. 229-237.

[20] D. Stevenson, G. Byrd and N. Hillery: "Information Security in Asynchronous Transfer Mode Networks", *Communications of the ACM Special Issue on ATM Networks*, Feb. 1995.

[21] D. Stevenson, N. Hillery, G. Bryd, F. Gong and D. Winkelstein: "Design of a Key Agile Cryptographic System for OC-12c Rate ATM", *Internet Society Symposium on Network and Distributed System Security*, San Diego CA, Feb. 1995.

Daniel Stevenson received a B.S. in Physics in 1974 and a M.S. in Physics in 1975. Prior to joining the staff of MCNC, he worked for Bell Telephone Labs, GTE Government Systems, Northern Telecom, and Bell Northern Research. Since 1981, his technical focus has been on exploratory design and analysis for switching systems based on emerging technologies. He is currently the Director of Advanced Network Technology at MCNC, where he oversees communications research and Advanced Network Projects. He is the system architect and coordinator for the VISTAnet project, and PI for several other networking research projects. He has numerous publications and conference presentations in the area of gigabit networking.



Frank Terhaar-Yonkers received a B.A. in Geology from SUNY at Geneseo in 1977. Prior to joining MCNC he has worked for Control Data, Cray Research, The North Carolina Supercomputing Center and as a contractor to the U.S. EPA. His career has focused on systems and infrastructure for High Performance Computing, Communications and Scientific Visualization. He is currently employed by MCNC as Manager of HPCC Research.

NETWORK AND PROTOCOL DESIGN

AN ATM NETWORK PLANNING MODEL

A. FARAGÓ, V. T. HAI, T. CINKLER, Z. FEKETE and A. ARATÓ

DEPAPRTMENT OF TELECOMMUNICATIONS AND TELEMATICS
TECHNICAL UNIVERSITY OF BUDAPEST
H-111 BUDAPEST, SZTOCZEK U. 2, HUNGARY
E-MAIL: FARAGO@TTT-ATM.TTT.BME.HU

An ATM network planning model is presented, along with algorithmic considerations. The procedure results in a physical network topology with a virtual path layout and transmission capacities assigned to the VPs. Predetermined connectivity requirements are satisfied as constraints, and the objective is to minimize an overall cost function. The global cost is the sum of local link costs and each local cost function has a fixed and a capacity dependent part. The properties of the problem and the proposed algorithmic solution are investigated and the operation is demonstrated numerically.

1. INTRODUCTION

ATM network planning is a subject that has attracted relatively little attention in the literature so far. Most papers deal with the dimensioning and logical configuration aspects, rather than designing the physical network topology itself. Even in the latest *Networks* conference there was not even a single paper focused specifically on ATM network planning, as such, that is, to the design of the physical topology, although this is the conference which is the most exclusively dedicated to the issue of network planning. The papers in the ATM Network Planning Session addressed mostly the logical configuration issues, assuming a given physical network, see e.g. [1], [2] (a further developed version of the latter was published as a journal paper [3]).

The potential reason for this lack of exploration could be that ATM is not expected to stand alone as a generic physical network. Rather than that, a large public ATM network will typically be carried by SDH or SONET high speed synchronous networks. This may result in the view that the network planning issues of ATM are shifted to SDH (SONET) planning. Thus, when it comes to ATM, then only the dimensioning, VP layout and logical configuration problems are investigated, while the physical carrier network is assumed given.

The above view, however, misses the point that the design of the carrier infrastructure network cannot be independent of the carried ATM network, at least in case one aims at a good cost/performance ratio and high utilization of resources. On the other hand, it seems not at all easy to find an ATM network planning model that captures the most relevant points, but it still avoids the overcomplications and the danger of intractability caused by the simultaneous handling of the infrastructure and the carried network.

In the present paper we attempt to develop an ATM network planning model and an algorithmic solution. Similarly to most practically relevant cases of network design, finding the *global* optimum is algorithmically intractable for large networks, as computationally hard (NP-complete) problems are involved. Therefore, we present a procedure that approaches a local optimum of the cost function. It is demonstrated numerically that the method produces good results.

The motivation and intended first application of this and related methods is to implement them in an intelligent network management support tool, called *PLASMA* [4]. The name *PLASMA* stands for PLAnning and Simulation methods in network MAnagement. The system is under development at the High Speed Networks Laboratory of the Department of Telecommunications and Telematics, Technical University of Budapest, in cooperation with Ericsson. This software tool is capable to support the work of the network traffic manager by offering and testing optimized network configuration versions, depending on the actual traffic situation, demands and many other conditions. Because of the flexible logical (re)configuration possibilities of modern high speed networks, there are quite a few tasks for the network manager that can be solved by such methods that are closer in spirit to planning tasks. This foreseen convergence of a substantial part of the design and management methodology makes it relevant to implement planning methods in the management support tool *PLASMA*.

2. THE MODEL

We propose the following model for the considered ATM network planning task. There are N network nodes, numbered by $1, 2, \ldots, N$. For each pair (p, q) of nodes a transmission *capacity demand* $T_{pq} \geq 0$ is given, expressed as an integer multiple of an elementary unit. Note that from the teletraffic viewpoint T_{pq} is not the same as the offered traffic: we assume that the offered traffic has been already mapped into capacity demands, using available teletraffic dimensioning methods. Thus, $T_{pq} \geq 0$ is the transmission capacity to be engineered to carry the traffic between p and q. (For simplicity we assume symmetry in the sense of $T_{pq} = T_{qp}$, but the model can directly be extended to the non-symmetric case, as well).

An essential feature of ATM is that the traffic in the ATM network is sent through *virtual paths (VPs)*. Network safety and load balancing considerations require that the same pair of nodes (origin-destination pair, O-D pair for short) is to be connected generally by several *independent* virtual paths that have common end-nodes but do not share internal nodes. To describe these requirements, for each O-D pair (p, q) of nodes a number k_{pq} is assumed given. The quantity k_{pq} prescribes the number of needed independent VPs that connect p and q. Again we assume symmetry in the sense of $k_{pq} = k_{qp}$, for simplicity.

The *cost structure* of the network is modeled as follows. All costs, as often done in network design, are mapped onto the links. The global cost is the sum of individual link costs. The cost c_{ij} of a link (i, j) that connect nodes i and j has the following structure:

$$c_{ij} = a_{ij} + f_{ij}(x_{ij}).$$

Here a_{ij} is the *fixed cost* of the link that is associated with creating the link. The *variable cost* $f_{ij}(x_{ij})$ is a function of the transmission capacity x_{ij} that is "put on the link" during the planning procedure. Our method allows an *arbitrary* variable cost function. For the numerical experiments we chose linear functions of the form $f_{ij}(x_{ij}) = b_{ij}x_{ij}$, which resulted in link costs of the form

$$c_{ij} = a_{ij} + b_{ij}x_{ij}.$$

The aggregated *global cost* is the sum of the individual link costs:

$$G = \sum c_{ij} = \sum (a_{ij} + f_{ij}(x_{ij})),$$

where the summation is taken over those links that are created by the planning algorithm and x_{ij} is the summed capacity of virtual paths that traverse the link (i, j). Note that the VP routes and capacities are not given in advance, they are also to be obtained by the algorithm.

Thus, the algorithmic task can be formulated as follows.

Given:

N	number of nodes
$\mathbf{T} = [T_{pq}]$	matrix of capacity demands
$\mathbf{K} = [k_{pq}]$	matrix of VP connectivity demands
$c_{ij} = a_{ij} + f_{ij}(x_{ij})$	link cost functions

Find:
- Which physical links are to be built.
- Transmission capacity of each physical link.
- k_{pq} independent VP routes to connect each O-D pair (p, q).
- Transmission capacity of each VP.

Subject to:
- The physical capacity of any link cannot be smaller than the summed capacity of the VPs that use the link.
- The summed capacity of the VPs that connect the same O-D pair is at least as large as the demand for the O-D pair.

Objective: The *aggregated global cost* of the created physical links is as small as possible.

It is not difficult to see that the problem belongs to the family of hard (so called *NP-hard* [5]) combinatorial optimization problems, as it contains the well known Traveling Salesman problem [5] as a subcase. To see this it is enough to take all link costs equal to the fixed part of the cost, while the variable part is set identically to 0. If we also set $k_{pq} = 2$ for all (p, q), then it is easy to see that the optimal solution is a minimum cost traveling salesman tour.

3. ALGORITHMIC SOLUTION

In this section an algorithmic solution is presented for the above specified ATM network planing task. First we show how to obtain an optimal system of k_{pq} independent virtual paths between a single given O-D pair (p, q), such that the capacity demand (load) T_{pq} is distributed equally among the k_{pq} VPs. Optimality means here that we want to create the new virtual paths with the smallest possible *increment* in the global cost of the existing network. That is, we want to extend the network to the smallest possible extent (in the cost sense) to satisfy the demand for a new O-D pair. Then the solution of this simpler problem is used as a subroutine for solving the main task.

3.1. VP Routing Between a Single O-D Pair for Minimum Cost Network Extension

Let us assume we already have an existing physical network in which the links have certain capacities. Let R_{ij} be the capacity of link (i, j). We wish to find k_{pq} independent virtual paths between a given O-D pair (p, q), such that the load (demand) T_{pq} is distributed uniformly among the VPs and the *additional* cost of creating the new VPs is minimum. It is assumed that the existing link capacities are already used by existing VPs because the design procedure has not put unnecessary excess capacity on any link in earlier phases.

We find the required minimum cost extension by reducing the task to a minimum cost network flow problem, as follows.

Minimum Cost Extension (MCE)

Step 1: Define the instance of a minimum cost flow problem as follows.

1. The underlying graph is a complete graph on N nodes.
2. Assign unit capacity to each edge.
3. If the physical link (i, j) exists in the network, then assign the cost $\widetilde{c_{ij}}$ to edge (i, j) by
$$\widetilde{c_{ij}} = f\left(R_{ij} + \frac{T_{pq}}{k_{pq}}\right) - f(R_{ij})\,.$$
4. If (i, j) is a non-existing physical link in the network then assign the cost to edge (i, j) by
$$\widetilde{c_{ij}} := a_{ij} + f\left(\frac{T_{pq}}{k_{pq}}\right)\,.$$
5. Assign a flow demand of k_{pq} to the O-D pair (p, q).

Step 2: Find a minimum cost integer flow in the flow problem specified in *Step 1.* (This can be done by existing network flow algorithms, see e.g. [6]).

Step 3: Due to the specially defined instance of *Step 1,* the minimum cost integer flow will consist of k_{pq} unit-flow path between p and q. These paths are the required VPs, each with capacity T_{pq}/k_{pq}.

The special cost assigment $\widetilde{c_{ij}}$ ensures here that only the *additional* cost is incurred. If (i, j) is an existing link, then $f(R_{ij} + T_{pq}/k_{pq}) - f(R_{ij})$ is the cost increment if it is used by a VP of capacity T_{pq}/k_{pq}. If the link has not yet been built, then the cost of using it contains the fixed cost, as well, this is represented by $\widetilde{c_{ij}} := a_{ij} + f(T_{pq}/k_{pq})$. Note that the procedure is not sensitive to the choice of the functions $f_{ij}(.)$, it works for arbitrary functions.

It is well known from the theory of network flows (see e.g. [6]) that for integer demand and integer edge capacities the demand can be satisfied by an integer flow (if there is a solution at all). The existence of a solution is guaranted here by the fact that we are allowed to build any new links if necessary. (Of course, as a trivial condition, it is assumed that $k_{pq} \leq N - 1$, otherwise it would be impossible to have k_{pq} independent VP routes). Thus, we see that Algorithm MCE really solves the subtask we addressed in this subsection. The solution is efficient in the sense that it runs in polynomial time (the complexity does not grow exponentially with the size of the network). This follows from the same property of network flow algorithms. Since we run a minimum cost flow algorithm for each O-D pair, therefore, the complexity is $O(N^2F)$, where F is the complexity of the minimum cost computation. The construcion of the network flow instance in *Step 1* takes $O(N^2)$ time, so it does not increase the $O(N^2F)$ term.

Remarks:

1. A simpler and faster heuristic solution is also possible by finding a minimum weight path between p and q with link weights equal to $\widetilde{c_{ij}}$. Then the edges of the path are removed from the graph and another minimum weight path is found in the remaining graph, and so on, until k_{pq} paths are found. This consecutive minimum weight path formulation does *not* guarantee, however, that the *total* cost of the extension is minimum.
2. It is unclear at the moment whether the approach can be generalized for the case when the load is non-uniformly distributed among the parallel VPs. It is not even clear whether a polynomial-time solution exists at all for the non-uniform case. The straightforward generalization of the flow approach is blocked by the fact that in the non-uniform case we cannot define the $\widetilde{c_{ij}}$ values unambigously, they would depend on which VP will use the link.

3.2. The Network Planning Algorithm

In this subsection we show how to use the MCE procedure, presented in the previous subsection, to solve the network design task. The idea is that starting with an empty network (containing no links), VPs are added step by step for each O-D pair, making a minimum cost extension of the network in each phase by the MCE algorithm.

There is, however, a remaining degree of freedom in this greedy type procedure: it which order should the O-D pairs be taken? This may affect the result seriously. There are, of course, several possibilities for chosing the order: we can try random ordering, or we can take them ordered by the size of the demand or by the cost of a minimum cost path connecting the O-D pair, where the edge costs are the fixed link costs, or we may take a combination of these.

Here we present a procedure which, in some sense, tries to optimize the greedy choice. Although it takes more computation, it still runs in polynomial time and is reasonably fast for a planning procedure where no real-time computation is needed.

VP-Based Network Planning (VPNP)

Step 0: Initialization: Take an empty network (no links).

Step 1: Run the MCE algorithm with the present network, separately for each O-D pair that has not yet been considered.

Step 2: Choose an O-D pair (p, q) that produced the smallest extension cost among the O-D pairs tried in *Step 1.*

Step 3: Do the extension with the minimum extension cost O-D pair found in *Step 2.*

Step 4: If all O-D pairs with $T_{pq} > 0$ have been connected by VPs then *STOP* else go to *Step 1.*

It is clear from the construction that this algorithm has polynomial running time because it essentially consists of $N+$ $+(N-1)+ +(N-2)+\ldots+1 = N(N+1)/2 = O(N^2)$ calls of the MCE algorithm. This yields an overall complexity of $O(N^4F)$ where F is the complexity of a minimum cost computation.

To further decrease the overall cost, we can iterate the VPNP algorithm, as follows. The idea is that we can potentially improve the VP system connecting a given O-D pair if we remove these VPs from the network and recompute them again using the MCE algorithm. Note that the new result is not necessarily identical with the previous one, since the network has changed when other VPs were added for other O-D pairs. On the other hand, the possibility of putting back the previous VP system for the O-D pair is always open. This implies that the recomputation procedure cannot degrade the already existing solution.

Iterated VP-Based Network Planning (IVPNP)

Step 0: Run the VPNP algorithm to create an initial network.

Step 1:: Order the O-D pairs in decreasing order according to the extension cost obtained when running the MCE algorithm for the O-D pair. Let the order be $(p_1, q_1), \ldots (p_m, q_m)$, where m is the number of O-D pairs with $T_{pq} > 0$.
Set i=1.

Step 2: Remove the VPs connecting p_i and q_i, decreasing the physical capacities on the links traversed by these VPs.

Step 3: Recompute the VPs for (p_i, q_i) using the MCE algorithm. Set i=i+1.

Step 4: If $i \leq m$ then go to *Step 2* else go to *Step 5.*

Step 5: If the new VP system is the same as the previous one for each O-D pair, then *STOP.* Else repeat from *Step 1.*

According to the monotone improvement in the global cost, explained above, the IVPNP algorithm will converge, since the global cost is a positive number. It is also clear, that until the halting condition is not satisfied, the global cost will strictly decrease in each round. On the other hand, we have no proof at the moment for any *guaranteed* speed of the convergence.

4. A NUMERICAL EXAMPLE

We demonstrate the algoritm on a network of 15 nodes. The nodes are randomly placed on the plane and the link fixed cost is proportional to the distance of the end nodes of the link. The variable cost is linear and it is also proportional to the distance. Let α be the cost of unit capacity over unit distance. Then the link cost in this case can be expressed as

$$c_{ij} = (1 + \alpha x_{ij}) d_{ij},$$

where d_{ij} is the distance between nodes i, j and x_{ij} is the transmission capacity "put on the link" by the VPs that traverse it.

The value of the parameter α can control the ratio of the fixed and variable cost. To make the numerical example easy to overview, we took $T_{pq} = 2$ and $k_{pq} = 2$ for each O-D pair (p, q). That is, we look for a 2-connected network in which the capacity of each VP is 1, measured in appropriate relative units.

The resulting networks for $\alpha = 0.005$ and $\alpha = 0.015$ are shown in Figs. 2 and 3, respectively. It is clearly seen that the two networks are different. If α gets larger, then it is worth building more links because the loss caused by building more links is dominated by the gain that shorter VP routes are possible, which reduces variable costs, since the capacity is put on less links. It is interesting to note that the network of Fig. 3 is not a simple extension of the one in Fig. 2, the denser network does *not* arise by just adding new links to the network of Fig. 2.

For comparison we have searched for a minimum cost *ring network* on the 15 nodes, as well. The ring topology is extremal in the sense that it can ensure 2 VPs between each O-D pair such that the total number of links is minimal. Also, the VP layout is unique in the ring, since there are only two possibilities to connect any two nodes (clokwise and counter-clockwise). Finding the minimum cost ring is, however, still difficult from the algorithmic viewpoint because it is essentially equivalent to the Traveling Salesman problem. Nevertheless, for a 15-node network we could still solve it in reasonable time by simulated annealing. The resulting networks coincide for $\alpha = 0.005$ and $\alpha = 0.015$. The obtained ring network is shown in Fig. 1.

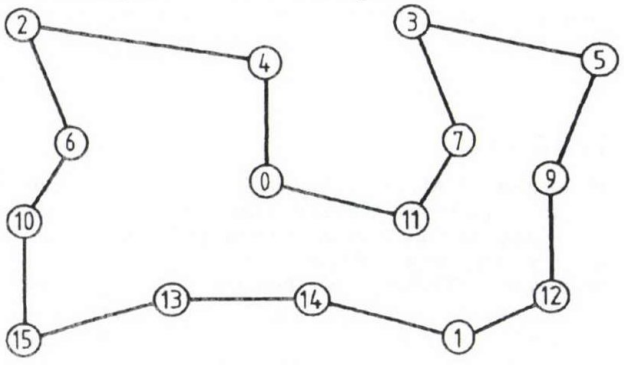


Fig. 1. Optical regular double connected network (obtained by simulated annealing)

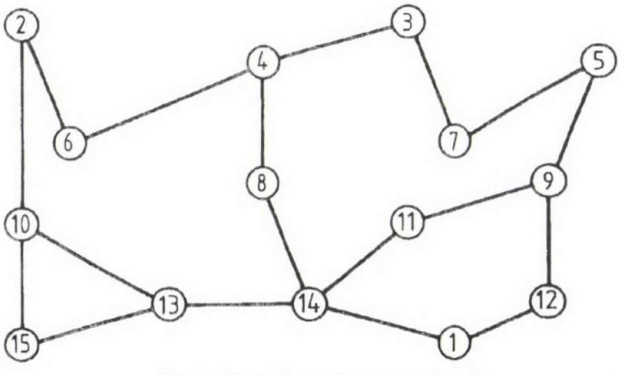


Fig. 2. Optical topology $\alpha = 0,005$

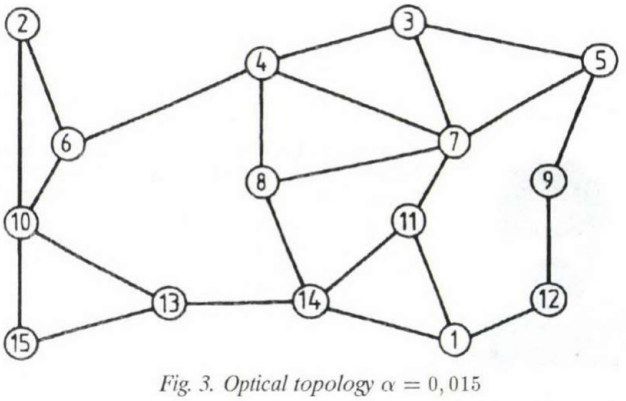


Fig. 3. Optical topology $\alpha = 0,015$

Comparing the costs of the networks obtained by our algorithm to the reference *optimal* ring network, we find the following effect. If α is small (0.005 in our case) then the cost of the ring (Fig. 1) and the cost of the network found by our algorithm agree with less than 1 % difference, although the two networks are different. If α gets three times larger (0.015), then our algorithm provides a solution with 15.6 % smaller cost than the *optimal* ring. This clearly shows that if the capacity dependent part of the cost has higher relative weigth, then it is worth deviating from the ring, which is the simplest 2-connected solution, because with "chord-links" we can achieve substantial savings in cost.

5. CONCLUSION

A virtual path based ATM network planning model has been presented with an efficient solving algorithm. The procedure results in a physical network topology with a virtual path layout and transmission capacities assigned to the VPs. Predetermined

connectivity requirements are satisfied as constraints, and the objective is to minimize an overall cost function. The global cost is the sum of local link costs and each local cost function has a fixed and a capacity dependent part. The properties of the problem and the proposed algorithmic solution are investigated and the operation is demonstrated numerically on a 15-node network.

REFERENCES

[1] M. Menozzi, U. Mocci, C. Scoglio and A. Tonietti: Traffic Integration and Virtual Path Optimization in ATM Networks, *NETWORKS'94 (Sixth International Network Planning Symposium)*, Budapest, Sept. 1994, pp. 71-76.

[2] A. Faragó, S. Blaabjerg, W. Holender, T. Henk, L. Ast, Á. Szentesi and Zs. Ziaja: "Resource Separation – an Efficient Tool for Optimizing ATM Network Configuration", Proc. *NETWORKS'94, Sixth International Network Planning Symposium*, Budapest, Hungary, Sept. 1994, pp. 83-88.

[3] A. Faragó, S. Blaabjerg, L. Ast, G. Gordos and T. Henk: "A New Degree of Freedom in ATM Network Dimensioning: Optimizing the Logical Configuration", *IEEE Journal on Selected Areas of Communications*, Spec. Issue on the Fundamentals of Networking, 13(1995/7), pp. 1199-1208.

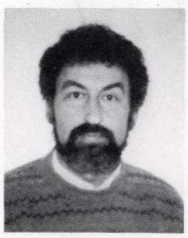
[4] Zs. Haraszti, I. Dahlquist, A. Faragó and T. Henk: PLASMA – An Integrated Tool for ATM Network Operation, *International Switching Symposium (ISS'95)*, 1995, pp. 314-318.

[5] M. R. Garey and D. S. Johnson: *Computers and Intractability – A Guide to the Theory of NP-completeness*, W.H. Freeman and Co., New York, 1979.

[6] W.-K. Chen: *Theory of Nets: Flows in Networks*, Wiley, 1990.

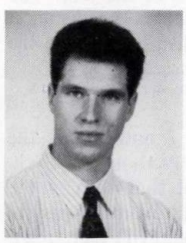
6. ACKNOWLEDGMENT

The work has been done in the joint research project of Ellemtel Telecommunication Systems Laboratories (the research and development institute of Ericsson and Telia) and the High Speed Networks Laboratory at the Department of Telecommunications and Telematics, Technical University of Budapest. The authors are grateful to Miklós Boda, Géza Gordos and Tamás Henk for their continuous support and encouragement.

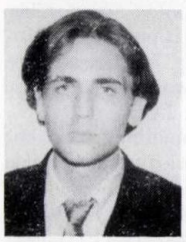


András Faragó received his M.Sc. and Ph.D. in electrical engineering from the Technical University of Budapest, in 1976 and 1981, respectively. He is associate professor at the Department of Telecommunications and Telematics, Technical University of Budapest. In the Department he is the research director of the High Speed Networks Laboratory based on a joint research project between the Technical University of Budapest and Ericsson. He worked as a visiting senior research fellow at the University of Massachusetts several times and now he is invited to Boston University. He is a member of the IFIP Working Group 6.3 Performance on Communication Systems, member of ACM and a founder member of the Hungarian Chapter of ACM. His main interest is in developing algorithms and applying mathematical methods for the design and optimization of various types of telecommunication networks and services. He published over 80 papers in the field, including 8 articles in *IEEE* journals.

Hai Vu Thanh received the M.Sc. degree in electrical engineering from the Technical University of Budapest. He is currently working toward the Ph.D. degree in the Department of Telecommunications and Telematics, Technical University of Budapest. His research interests are routing and dimensioning of ATM networks, and survivable network design.



Tibor Cinkler received his M.Sc. degree in electrical engineering from the Technical University of Budapest, Budapest, Hungary, in 1994. He is with the High-Speed Networks Laboratory at the Department of Telecommunications and Telematics, Technical University of Budapest, where he is currently working toward the Ph.D. degree in electrical engineering. His research interests are related to the optimization of broadband (ATM) networks, particularly call-level modelling for dimensioning and performance evaluation, as well as topological design and routing optimization. He is a a student member of IEEE and ACM and a member of the Scientific Society for Telecommunications, Hungary.



András Arató last year student, currenlty going for his M.S (E.E) as a member of the ATM project at the Department of Telecommunications and Telematics, has been working on the design and traffic control of ATM networks for one year. He has also been working at the Center of Information Systems of the Technical University of Budapest for 2 years as a network system manager. His current research projects focus on the Virtual Path assignment in ATM Networks.

ATM LAN NETWORK DESIGN

V. ELEK, Z. GÁL, H. L. PHAN and C. SZABÓ

DEPARTMENT OF TELECOMMUNICATIONS
TECHNICAL UNIVERSITY OF BUDAPEST
H-1111 BUDAPEST, SZTOCZEK U. 2., HUNGARY

In this paper some key issues in ATM-based local area networks are considered. An overview of the product offering that can serve as a basis for building ATM LANs is given first and basic building blocks are identified. Considerations for the topology design of ATM LANs are given. A simulation model of a typical ATM LAN scenario consisting of videoconferencing workstations and an image server is presented and simulation results are obtained to determine the basic delay and loss measures.

1. INTRODUCTION

Asynchronous Transfer Mode has been originally introduced and standardized as the multiplexing and switching technique for public B-ISDN. The main objective was to find a common transfer platform for all existing and evolving services to be provided in global area for a large number of users.

While we are witnessing ambitious but clearly medium-to long term efforts of international bodies and telecom service providers to launch B-ISDN services, in the short term the ATM deployment has shifted from public to private applications and from wide area to the LAN environment. The main reasons are at least twofold.

First, the driving forces toward ATM-based networks that are clearly foreseen in global area do already exist in local, or more general, in enterprise environment. Access to powerful data and image servers and the emerging multimedia applications require a high data rate access from desktop, together with flexibility and scalability. Conventional shared-medium LANs and MANs possess serious limitations: the scalability in terms of the aggregate speed is limited by the data rate of the medium, conventional LANs lack scalability both in speed and space and MANs' scalability in area coverage is also limited, and, finally, most of the shared medium access protocols do not or poorly support the integration of different services.

Second, the introduction of ATM in the local area seems to be significantly simpler than in a public global network. ATM LANs are less complex in terms of the number of network nodes and links, population size, number of simultaneous connections, etc. Also there is no need in employing sophisticated call and traffic control procedures that are so important for a public service and that are yet to be developed and standardized for B-ISDN applications.

Because of ATM's inherent scalability, there will be no fundamental difference between an ATM LAN and an ATM enterprise network, meaning by the latter a private network of a particular enterprise that can be a intra-building LAN, a campus network, a MAN or a national or even a global network. An ATM LAN can easily be extended to serve a large area by obviously making the network topology more complex but without changing the underlying switching and transmission techniques as opposed to the present practice when one has to move from the existing LANs' shared-medium to a switched technology available in wide area like X.25 or (narrow-band) ISDN.

ATM is also an advantageous alternative to solutions when a legacy LAN or MAN has to be connected to B-ISDN by a gateway. This way, the gateway function is shifted to the end-user equipment thus eliminating the need in a complex gateway to be used to interface an access network collecting the traffic from a large number of users.

In addition to and motivated by user needs, a significant progress has been made during the past couple of years in the available technology and standard development.

Major networking companies are now offering a wide range of devices and systems for ATM LANs ranging from hubs and routers with "added" ATM-capability to large ATM switches [1].

As for the standardization part, while the ITU-T's B-ISDN recommendations are important as guidelines, a lot of activity has been carried out in the ATM Forum to speed up the standardization process and to address the specific areas that are of primary importance for ATM LANs/enterprise networks.

As for the research aspects, there are certainly a number of interesting open issues that are necessary to look into. These issues are concentrated around the basic differences between the public/global and private/local environments. Two of the authors of this paper made an attempt to identify these differences and the related open problems [2]. A number of publications addresses mainly the architectural and protocol issues of implementing existing LAN services on an ATM basis [3]. Special issue of IEEE J. on Select. Areas in Communications is devoted to ATM LANs which demonstrates the significance of the subject [4].

Typical ATM LAN scenarios have been considered in several papers in the last years. In one of them investigations were made on the effect of multiplexing a small number of bursty data streams onto a single ATM link [5]. A simple ATM LAN configuration was simulated in [6].

An interesting research topic is to compare traditional shared-medium LANs with switch-based ones, in particular with ATM-based LANs. In [7] a performance comparison of a traditional bus-based computer architecture for a multimedia server with an architecture based on an internal LAN inside of the computer is presented.

The objective of this paper is to contribute to the network design of ATM LANs. The outline of this paper is as follows. In Section 2, basic building blocks are introduced that have been identified based on the study of product offerings. Some basic design considerations have been worked out that are addressed in Section 3. Section 4 contains the description of a typical ATM LAN scenario, the traffic models and the performance results. The conclusion of the paper outlines the further work being planned to carry out.

2. IMPACT OF STANDARDIZATION AND PRODUCT OFFERINGS

Standardization and product offering are the two interrelated key issues that influence the penetration of ATM LAN solutions. We will briefly overview the status of the standardization, in particular the relevant activities of ATM Forum, and the state-of-the art of products suitable for ATM LANs.

2.1. Standardization

One of the key issues addressed by the ATM Forum is the physical layer. In particular, the ATM Forum has approved several versions for User-Network Interface (UNI), including some of the traditional lower data rate interfaces like T1, T3 and E1, E3, the three first SONET interfaces (e.g., the STS-12c at 622.08 Mbps) as well as a 100 Mbps interface called TAXI and a 25 Mbps interface.

The Rorum has also adopted a number of specifications for LAN environments, including 100 Mbps interface using FDDI encoding, a 155 Mbps interface using Category 5 UTP and a 51 Mbit/s interface using Category 3 UTP.

Another very important issue for the introduction of ATM for LANs is the ways of providing traditional LAN services. A promising approach is LAN emulation that supports an IEEE 802 connectionless packet transfer service at the MAC layer on top of ATM. LAN emulation enables applications on legacy LANs to access ATM-attached servers, routers, and other network equipment and forwards upper-layer protocols across ATM connection without requiring any modifications to legacy software. The solu-

tion is to design a protocol layer, an ATM MAC sublayer above the AAL that emulates the service offered by an IEEE 802 LAN segment at the MAC sublayer to the IEEE 802.2 Logical Link Control Sublayer. The ATM Forum finished L-UNI (LAN Emulation User-to-Network Interface), its LAN emulation standard.

2.2. Overview of the product offering

Our analysis of the existing products suitable for ATM LANs was based on the study of some major vendors' technical information and is by no means an exhaustive one. However, analysis of the product lines of vendors including Cisco Systems, Cabletron Systems, Newbridge Networks and IBM allowed us to define some general principles of building ATM devices and their main properties. Specific information can be found in the relevant product descriptions [8]–[10].

Quite often the vendor upgrades its existing device typically called hub by ATM functionality for switching within the hub, and the interfaces remain the same, e.g., special interfaces to connect LANs such as Ethernet and token ring, serial interface, etc. If ATM appears also as an external interface, we have an ATM device which collects the traffic from non-standard subnetworks and in addition to the internal switching multiplexes them onto an ATM interface. This functionality will be called as "group access device" in the sequel.

Some vendors developed so-called ATM-cards to connect workstations to an ATM link. We will call this functionality as "individual access device".

"ATM Switch" is a natural ATM building block that is very similar to ATM switches developed for and offered to public service providers. Finally, we have also identified a building block "ATM Concentrator" with the functionality described below.

2.3. Typical building blocks for ATM LANs

1. *ATM Switch* has only Physical and ATM layer functions. This device typically has 4, 8 or 16 ATM lines. The number of cards and the types of interfaces can be customized as desired and reconfigured in the field as network requirements change. The ATM switch is a key component of an ATM LAN backbone, which connects a number of other ATM devices. All interfaces have the same data speed. The number of interfaces typically ranges from 1 to 16.
2. *ATM Concentrator* is a device with different data rates at its input ports as opposed to ATM switch. Otherwise a concentrator also provides only Physical and ATM layer functions. Another feature is that a concentrator usually cannot provide connections among its input ports, thus, the switching (at virtual circuit level) is accomplished in the ATM switch. Concentrators are used to keep the necessary number of switch ports small. The number of interfaces is typically 4, 8 or 16.
3. *ATM Group Access Device*, also commonly called "hub", has at least one IEEE 802 LAN interface. ATM hub must provide AAL functionality and support different traffic types. In addition to different IEEE 802 LAN interfaces (their typical number is up to four) these devices have several ATM interfaces as well as conventional serial interfaces.
4. *ATM Individual Access Device* represents the same functionality as the group access device but only for one user equipment. This device is normally a card plugged into the workstation's bus. It offers signaling for switched virtual connections, multiprotocol encapsulation for interoperability with ATM-connected servers and workstations, hardware support for conversion between data packets and ATM cells, traffic management functions, internetworking for virtual LANs.

As for the performance parameters of these devices, only a few published data can be found. Typical parameters for an ATM Switch are as follows. Between two ports cell delay variation ranges from 0.96 to 2.03 microseconds. Average cell delay between two ports is typically between 10.5 and 75 microseconds.

In Table 1 the functionality and features of the four basic building blocks are summarized. Fig. 1 shows an example of a simple ATM LAN built from these elements.

Table 1. Summary of properties of basic building blocks.

	Switch (1)	Concentrator (2)	Group Access Device (3)	Individual Access Device (4)
Layers covered	PHY, ATM	PHY, ATM	PHY, ATM, AALs	PHY, ATM AALs
Connectivity	1,2,3,4	1,2,3,4	1,2, legacy LANs	1,2
Typical number of ports	4,8,16	1 output 4,8,16 input	1 ATM max 16 CSMA/CD max 4 FDDI, token ring	1 ATM
Typical data rates	100 Mbps 155 Mbps	1x155 Mpbs or 100 Mbps, T1, T3, E1, E3	100 Mbps – TAXI, 155 Mbps	100 Mbps – TAXI, 25 Mbps – IBM

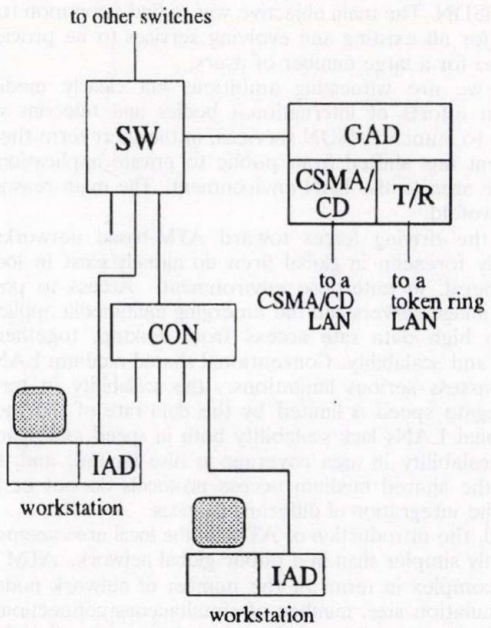


Fig. 1. An ATM LAN configuration example containing all four building blocks, where SW: ATM switch, CON: concentrator, GAD: group access device, IAD: individual access device.

3. NETWORK DESIGN CONSIDERATIONS

3.1. Topology design

In the topology design of ATM LANs the following characteristics have to be considered: the distribution of the user population and the expected traffic, its type and distribution as main components. Network flexibility and scalability, reliability and low cost also have to be taken into account.

For a small, concentrated user population even a network consisting of only one switch is usually suitable. The introduction of more complicated switching networks is motivated by practical circumstances such as remote user population, where the cost of cabling to connect users to the central switch would be high, or the case of large or growing user population to avoid the use of too complex and thus expensive switches.

While in the first case the network segmentation is determined by the "spatial distribution" of the user population (e.g., users in close proximity have to be connected to the same switch), this segmentation becomes a more complicated task in the second case.

If a concentrated but large-size user population is to be interconnected, a network topology has to be found that can

handle the expected traffic effectively, and this way the topology will depend on the traffic distribution. The traffic-independent aspect of network design in this case is the reliability.

Assuming that the number of segments (a segment is a switch with end-nodes connected to it) is relatively small, the topology can be based on star, linear (ring, bus) or meshed architecture. In a switching network with star topology, a central switch is used to interconnect network segments. The main advantages of this solution are high aggregate bandwidth and low level of multiplexing (only traffic streams from/to a given segment are multiplexed), and the disadvantage is its low reliability. Failure in the central switch can prevent any communication in the network, and even due a link failure a group of user becomes disconnected.

The ring or mesh topologies do not use central devices, thus the effects of failures can be localized. Due the higher connectivity failures cause only higher traffic in the rest of the network. On the other hand in these networks the level of multiplexing and thus its effect on the single traffic streams will be higher.

3.2. Delay considerations

If the traffic is directed to a node connected to the same switch, constant factors of the delay are the transmission and propagation delay, the processing time at the nodes and in the switch; while the waiting time in the buffer at the link to the destination node can be different, and will determine the variation of the delay. This value depends on the traffic directed only to the given node.

If the nodes of a connection belong to different switches, additional effects will modify the delay. Additional constant delay comes from the propagation delay between the switches, and the additional processing times. These values are constants, and do not depend on the traffic characteristics.

The most significant component in this case is the variable delay in the buffer at the link connecting the switches. Since on this link all the traffic between the switches is multiplexed, the quality of service parameters of the transmission to a given node can be influenced by communication between two other nodes. As the route between the users becomes longer, not only the mean delay, but also the cell delay variation increases. In ATM networks the buffer-sizes at the switches are limited, to minimize the delay variation. As the routes in our case will be significantly shorter, the CDV can be below the maximum value even if long buffers are used, decreasing the probability of cell loss.

3.3. A simple network design example

As an example, let us consider a network consisting of workstations and file/image servers (see Fig. 2). The size of the population makes the use of two switches necessary. The workstations can establish CBR connections for video-conferencing, and they frequently use the VBR file or image servers.

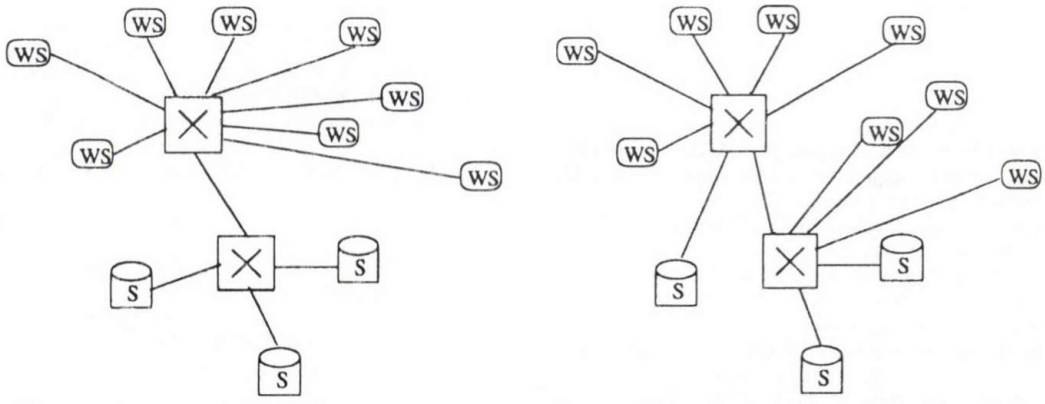


Fig. 2. Two possible network segmentation for the same environment

The optimal clustering could be to connect all the workstations to one of the switches, and the servers to the other switch. In this case the delay sensitive CBR traffic will not be affected by the bursty traffic coming from the servers before the link going to the station. But the bursty traffic can cause overflow on the inter-switch link.

If we want to decrease the level of VBR traffic on this link, a symmetric topology can be used, where workstations and servers are equally distributed between the two switches. Thus CBR traffic will appear on the inter-switch link, and will be multiplexed with the VBR traffic. This solution can work only, if the quality of the CBR service remains good enough.

Thus, to find the proper way of grouping traditional LANs, high-end workstations, data, image and video servers, etc. within an ATM LAN, we have to study the effect of multiplexing of small number of traffic streams on the quality of service parameters.

4. SIMULATION STUDY OF A TYPICAL ATM LAN SCENARIO

The main goal of the performance evaluation of a typical ATM LAN scenario was to determine the effect of multiplexing different data streams.

4.1. Simulation environment

In the first step of considering the design of an ATM LAN, we intended to examine the impact of very bursty traffic in addition to the CBR traffic on the network. In this paper, the CBR traffic is compressed video and the burst traffic is image traffic. In the section below, the network topology, the models of the building blocks, the service scenarios and traffic models with the most important networking requirements are discussed.

Configuration

We simulated a simple ATM LAN, where there are one ATM switch and N+1 stations in a star topology as shown on Fig. 3. The stations 1, 2,..., N are the normal stations, which send (and receive) video traffic to (and from) other stations, while the station 0 is an image server, from which the normal stations request image from time to time. This situation represents a video conference process with N participants and one image source that contains stored image information to be used during the videoconference. We have assumed that, in addition to point-to-point connections between fully connected videoconferencing workstations there is a point-to-multipoint (multicast) communication between the image server and the workstations. The N+1 stations are equidistantly spaced from the switch and the distance is about 2.5 km. We have examined the network of N+1 stations in 3 cases where N=3, N=5 and N=8 that means the number of participants in the conference is 3, 5 and 8, respectively.

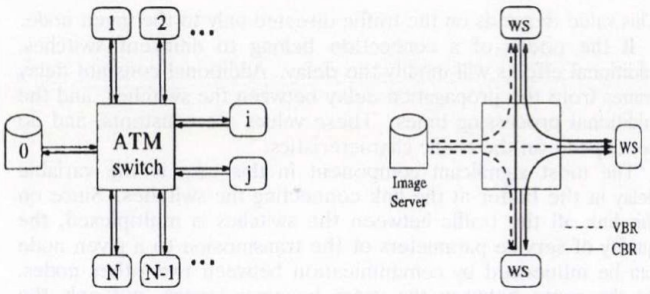


Fig. 3. Network topology and traffic distribution for the simulation.

ATM switch model

The parameters of the switch model reflect typical values. A nonblocking switch with FIFO output buffers is used, with 500-cell buffers per port. Cell loss may occur at the output buffers due the buffer overflow. Input buffers for one cell are used to store the incoming cells until switching. Assuming an internal switching speed of 8 Gbps, which is significantly higher than the speed of input links, cell loss at the input ports is not possible.

When a cell arrives at an output port, if the appropriate buffer is not full, it will be queued as last cell in the buffer; if the buffer is full, the first cell in the buffer will be dropped and the incoming one is queued as last cell. The speed of input links and output links was assumed to be the same and equal to 100 Mbps.

4.2. Traffic and service models

Video sources

Video is moving pictures. It is different from imaging and graphics mainly in the motion component. Video represents motion scenes as a rapid sequence of still frames. A PAL compatible video is $768 \cdot 516$ pixels at 25 frames per second. The smaller the window size (fewer number of lines scanned), the lower the bandwidth requirement.

The video signal can consume tremendous bandwidth. An uncompressed digital video can consume anywhere from 90 Mbps to 200 Mbps depending on the encoding. This is enough to overrun any existing network capacity. Compressed video offers a significant reduction in the bandwidth needs while maintaining similar picture quality (see Table 2, [12], [13]). The issue of compression algorithms and implementations is beyond the scope of this paper and the parameter of most interest to us is the network delay.

Table 2. Rates and bandwidth requirements for different compression techniques

Compression technique	frame rate (fps)	bandwidth
MPEG	30	1.5 Mbps video stream
MPEG II	60–75	4–10 Mbps
px64	6–15	64 kbps–2 Mbps

In the paper each video source represents a compressed video stream. An MPEG-type stream at the rate of 1.5 Mbps was used. Since the payload of an ATM cell is 48 out of 53 bytes the necessary bandwidth for a video stream is 1.65625 Mbps.

In the simulations, we have examined four cases where every normal station was assumed to transmit and receive (i) one video source, (ii) two video sources, (iii) three video sources and (iv) four video sources.

Voice sources

Sound may be classified as human speech and music. Human speech or voice is typically in the 4-5 kHz spectrum. The bandwidth of music discernible by human ear is 20-25 kHz (high fidelity systems have a bandwidth of 22 kHz). As an example, the resulting data rate for a digital stereo sound is 0.44 Mbps. Although voice transmission is normally an inherent component of videoconferencing, it causes only an insignificant increase of the data rate thus it was not taken into account when modeling the point-to-point videoconferencing connections.

However, we assumed that the image server is also a source of stored voice. in the form of a sequence of 2 kbytes packets (2028 + 20 MPEG headers), with an inter-arrival rate of 84.5 ms. These packets were typically interleaved with the video stream. Thus, the offered load for one stored voice source was of 0.218 Mbps, and the required bandwidth for a source stored voice was of 0.240852 Mbps ($= 53 \cdot 0.218/48$). Since stored voice traffic is always along with video traffic, in the simulations, we have examined the four cases described above.

Image

This traffic source was used to simulate the impact of the highly bursty load on the network performance. It was assumed that the image size distribution was uniform over 1.25–5.625 Mbytes. Each image consists of 35 slot length packets arriving at 100 Mbps (peak load) once the image transfer was initiated. The image arrival process is a Poisson process with a mean interarrival time of 6 s. In the first step of the network design, we assumed that the image server sends images one by one to workstations in turn.

4.3. Results

The simulation programs were written in SIMSCRIPT II.5 language and simulations were run on IBM RISC6000 workstations.

For the sake of saving the program runtime, we only examined the case when the image server transmits one image to one of the normal stations called "mixed" during the time duration beginning when the first packet of the image arrives at the server and ending when the last cell of the image is received by the destination station. Figs. 4 through 8 and Tables 3–5 summarize the most important results we have obtained from 12 simulation runs.

Table 3. End-to-end delays in case of 3 participants

			1CBR	2CBR	3CBR	4CBR
average delay (msec.)	video	normal st.	0.049	0.049	0.049	0.049
		mixed st.	1.516	1.748	1.756	1.761
		whole sys.	0.531	0.606	0.556	0.588
	voice	normal st.	0.049	0.049	0.049	0.049
		mixed st.	1.52	1.729	1.763	1.746
		whole sys.	0.532	0.584	0.577	0.556
	image	(mixed st.)	9.69	9.874	9.919	9.891
maximum delay (msec.)	video	normal st.	0.051	0.052	0.057	0.057
		mixed st.	2.161	2.161	2.162	2.162
	voice	normal st.	0.051	0.054	0.056	0.055
		mixed st.	2.147	2.139	2.115	2.127
	image	(mixed st.)	18.665	18.661	18.665	18.665

Table 4. End-to-end delays in case of 5 participants

			1CBR	2CBR	3CBR	4CBR
average delay (msec.)	video	normal st.	0.049	0.049	0.05	0.051
		mixed st.	1.75	1.752	1.691	1.614
		whole sys.	0.383	0.35	0.325	0.293
	voice	normal st.	0.049	0.049	0.05	0.051
		mixed st.	1.742	1.747	1.693	1.61
		whole sys.	0.376	0.351	0.326	0.293
	image	(mixed st.)	9.873	9.902	9.849	9.789
maximum delay (msec.)	video	normal st.	0.055	0.058	0.068	0.069
		mixed st.	2.159	2.162	2.158	2.167
	voice	normal st.	0.054	0.057	0.068	0.066
		mixed st.	2.126	2.098	2.094	2.11
	image	(mixed st.)	18.665	18.665	18.661	18.662

Table 5. End-to-end delays in case of 8 participants

			1CBR	2CBR	3CBR	4CBR
average	video	normal st.	0.049	0.05	0.051	0.053
delay		mixed st.	1.761	1.656	1.525	1.406
(msec.)		whole sys.	0.237	0.221	0.191	0.171
	voice	normal st.	0.049	0.05	0.051	0.053
		mixed st.	1.758	1.658	1.525	1.411
		whole sys.	0.233	0.221	0.2	0.17
	image	(mixed st.)	9.91	9.803	9.71	9.605
maximum	video	normal st.	0.062	0.066	0.072	0.089
delay		mixed st.	2.161	2.165	2.167	2.163
(msec.)	voice	normal st.	0.059	0.068	0.071	0.085
		mixed st.	2.157	2.128	2.157	2.112
	image	(mixed st.)	18.661	18.661	18.661	18.665

Effect of buffer sizes

Fig. 4 shows the average queue length in the output buffer at the switch output port, connected to the "mixed" station where, in addition to the video and stored voice traffic the image traffic is also present. Due to long image bursts, in all simulation runs the maximum queue length in the "mixed" output buffer was always 500, the output buffer size (overflow). The average queue length in this buffer proportionally increases with the number of the CBR sources (the number of the participants in the conference) and approaches to the value of 500.

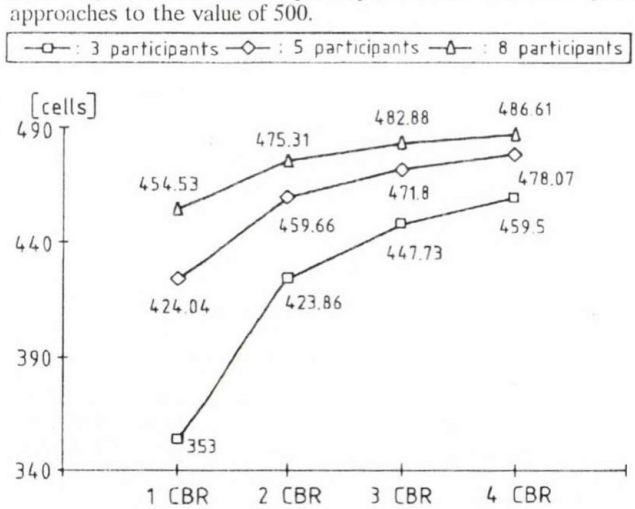


Fig. 4. The average queue length in the "mixed" output buffer

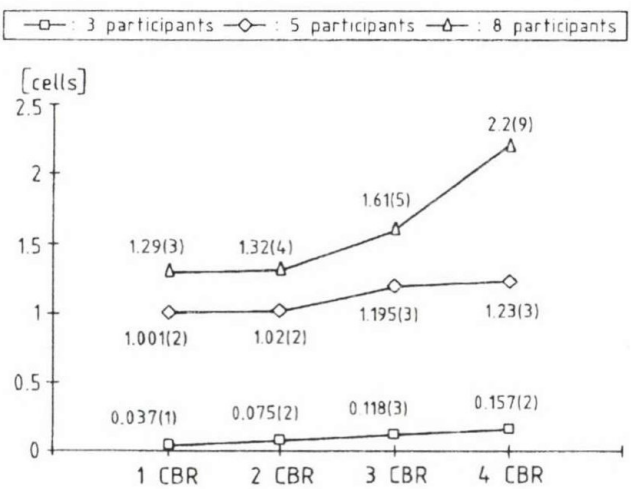


Fig. 5. The average queue length in the remaining output buffers

Fig. 5 shows the average queue length in the output buffers at the switch output ports connected to remaining "normal" stations where only the video and the stored voice traffic were present. Numbers in parentheses denote the corresponding maximum queue lengths. The average queue length also proportionally increases with the number of the CBR sources and the number of the participants, but the values are very small because the total CBR traffic passing these output ports is relatively small. For example, in the case of 8 participants and 4 TV windows (4CBR) the total CBR traffic is roughly 56 Mbps, only the half of the output link capacity (100 Mbps). The average queue length and the maximum queue length were only 2.2 and 9 cells, respectively.

Loss probabilities

A cell is considered to be lost when (1) its waiting time at the transmitting station exceeds an allowed waiting time; (2) its end-to-end delay exceeds an allowed end-to-end delay at the receiving station and (3) it is lost due to buffer overflow. In all simulation runs, there were no cell losses due to exceeding waiting time and exceeding end-to-end delay. Therefore, cell loss only occurs at the switch due to buffer overflow and only occurs at the "mixed" output port of the swtich. For this port we have calculated the loss probabilities of video traffic, stored voice traffic separately, image traffic, and an aggregate loss probability for the total traffic. As for average queue length in this "mixed" port, the loss probabilities increase as the number of CBR sources and the number of participants increase. Fig. 6 shows the aggregate cell loss probabilities for this "mixed" output port.

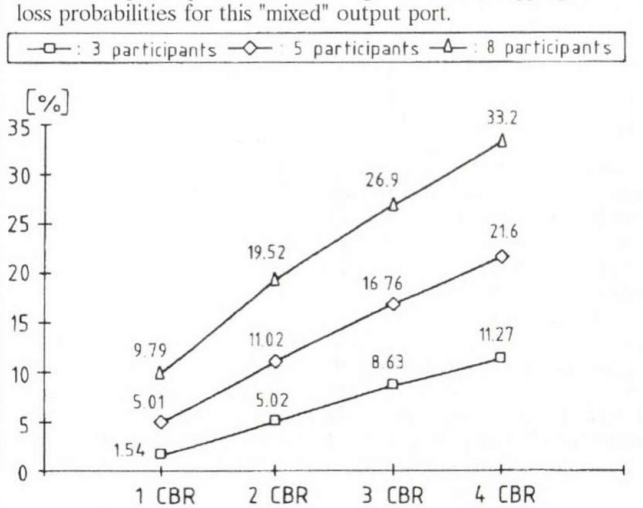


Fig. 6. The aggregate cell loss probability in the "mixed" output port

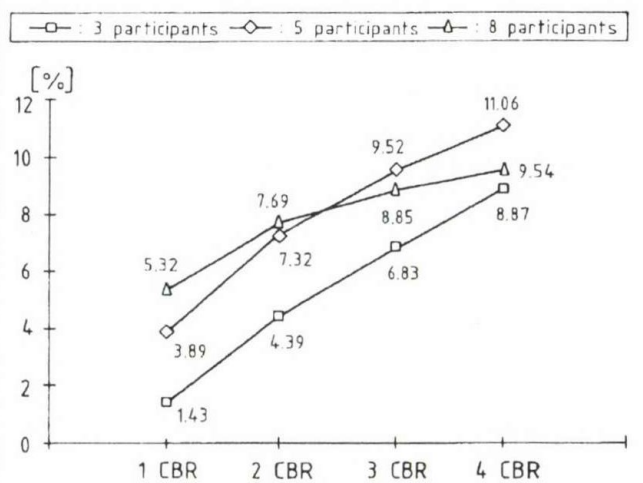


Fig. 7. The aggregate cell loss probability on the whole system

We also calculated the aggregate cell loss probability for the whole system (Fig. 7). We have also obtained an increase of the loss probability but this growth slows down when the number

of CBR sources and the participants increases. For example, for 5 participants and 3 TV windows (3CBR) the aggregate loss probability for the whole system is 9.52 %, for 4 TV windows (4CBR), it is 11.06 %, and for 8 participants the loss probabilities are 8.85 % and 9.54 %, respectively, for 3 and 4 windows. This does not mean that the system performs "better" when the number of CBR sources and the number of participants increase since in Fig. 6 the loss probabilities for the "mixed" port in these cases are 16.76 %, 21.65 % and 28 %, 33.92 %, respectively. Consequently, beyond a certain threshold when the number of CBR sources number and the number of participants increase, the growth of the number of lost cells is slower than the increase of the total load.

End-to-end delays

The end-to-end delay of a cell is the time duration between its arrival to the transmitting station and its reception at the receiving one. That is, the end-to-end delay is the sum of the waiting times at the output buffer of the transmitting station, the waiting time at the switch and the propagation time on the links. We have collected data separately for the end-to-end delay of video, stored voice and image traffic at the "mixed" station, the end-to-end delay of video, stored voice traffic at the remaining normal stations and the end-to-end delay of video, stored voice traffic on the whole system. Tables 3, 4, 5 show these delays with the average and maximum values in the case of 3, 5, 8 participants, respectively.

At the "mixed" station the average and maximum end-to-end delays of video and stored voice traffic are roughly the same and are in the range of 2 ms. For the video transmission additional delays at the codecs have to be taken into account, but the aggregate delay is still below the usual delay constraints of video applications. The average and maximum end-to-end delays of image traffic are 10 ms and 18.66 ms, respectively. Here we can observe that when the number of CBR sources and the number of participants increase, the average end-to-end delays first increase but beyond a certain threshold they start to decrease slightly. This also does not mean that the system is "better" but the increase of cell loss speed beyond a certain threshold essentially decreases the queuing time of cells crossing the switch.

At the remaining normal stations, the average and maximum end-to-end delays of video and stored voice traffic are roughly the same and are in the range of 0.05 - 0.08 ms. In the cases of 3 and 5 participants with 1 TV window (1CBR) and 2 TV windows (2CBR), the average end-to-end delays are the same and equal to 0.049 ms (Tables 3 and 4). Here the amount of CBR traffic is too small so the propagation delay is the dominant component in the end-to-end delay. In the cases of 5 participants with 3 TV windows (3CBR) and 4 TV windows (4CBR) and 8 participants, the average end-to-end delays increase accordingly to the increase in CBR sources and number of participants as expected (Tables 6 and 7).

For the whole system we can also observe that the average and maximum end-to-end delays of video and stored voice traffic are roughly the same. However, in cases of 5 and 8 participants the average end-to-end delays decrease when the number of CBR sources increases. This is because the number of cells transmitted with "small" end-to-end delay at the remaining stations plays more important part than the number of cells with "larger" end-to-end delay at the "mixed" station in the aggregate average end-to-end delay on the whole system. We do not see this decrease in the case of 3 participants, moreover there is a noticeable increment in the 1CBR and 2CBR cases (from 0.531 to 0.606 ms for video and 0.532 to 0.584 for voice – Table 3). This is because the amount of CBR traffic is still too small so that the number of lost cells of the CBR traffic depends on the time at which a cell arrives to the buffer; and the number of cells with "larger" end-to-end delay at the "mixed" station plays an important part in the aggregate average end-to-end delay, moreover in the case of 1CBR the cell loss probability is only 1.43 %.

Delay at the switch

The delay of a cell at the switch is the time duration between its arrival to the input buffer and the moment when it leaves the output buffer at the switch. Fig. 8 shows the average cell delay at the switch in all cases. The maximum delay at the switch is the same and equals to 2.12 ms, the transmitting time of 500 cells at the output link. Also as in the case of average end-to-end delay on the whole system, the average delay at the switch decreases with the increase in the number of CBR sources and participants after a certain threshold. This only means that the number of the cells transmitted with "small" delay is dominant in the number of ones with "large" delay at "mixed" port.

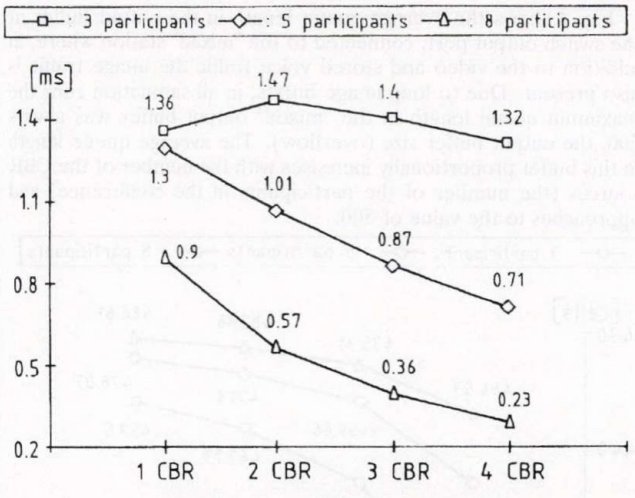


Fig. 8. Average delays at the switch

5. CONCLUSION AND FURTHER WORK

The main objective of this paper was to contribute to the solution of key network design problems of ATM LANs. We have focused on a specific issue of topology design of ATM LANs consisting of a small number of switches. This scenario reflects well one of the approaches of migration to ATM LANs which is the "ATM island" approach. Main results are the identification of basic building blocks, based on the survey of available products, some basic network design considerations and a simulation study of a practical scenario to determine the effect of multiplexing on the quality of service parameters.

Further studies may be focused on a deeper understanding of the network design issues in larger networks including the other approach of migration to ATM LANs where ATM is used as the backbone network, on the optimal combinations of shared medium and ATM techniques as well as on the comparison of performance of traditional shared medium LAN architectures and ATM LANs.

REFERENCES

[1] J. T. Johnson: "Router Vendor Makes Its ATM Vision a Reality," *Data Communications International,* pp. 57-60, September 1994.

[2] V. Elek and C. Szabo: "ATM LANs", *COST 242TD(94)07,* February 1994.

[3] Ron Jeffries: "ATM LAN Emulation: The Inside Story," *Data Communications,* September 21, 1994.

[4] "ATM LAN's: Implementations and Experiences," *Spec. Issue of IEEE J. on Selected Areas In Communications,* Vol. 13. No. 4. , May 1995.

[5] P. Camarda: "Bandwidth Management for ATM-Based LANs," *19th Conference on Local Computer Networks,* Oct. 1994.

[6] V. J. Friesen and J. W. Wong: "Migrating to ATM in the Local Area," *6th IEEE Workshop on Local and Metropolitan Area Networks,* March 1995.

[7] R. Rooholamini and V. Cherkassky: "Moving ATM Closer to Multimedia Applications," *19th Conference on Local Communication Networks,* Oct. 2-5, 1994., Minneapolis, Minnesota.

[8] "36150 Mainstreet, ATMnet Access Switch, General Information, Release 2, Generic A11112," *Newbridge Networks,* 1994.

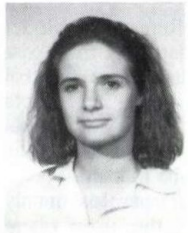
[9] "Cisco Systems Product Catalogue," January 1995.

[10] "Complete Solution Catalogue," *Cabletron Systems,* 1994.

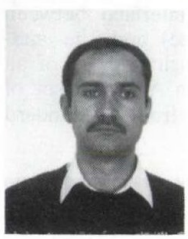
[11] "IBM Product Data Sheets," 1994.

[12] "ISO/IEC 1/SC 29 N 071; Coded representation of picture, audio and multimedia/hypermedia information; a. k. a MPEG."

[13] "CCITT Recommendation H. 261, Line Transmission on Non-telephone signals: Video Codec for Audio-visual services at $p \times 64$ kbps."



Viktória Elek received her M.Sc. degree in electrical engineering and computer sciences at Technical University of Budapest in 1993. Recently she is a Ph.D. student at the Department of Telecommunication of TUB. Her research area includes ATM local area networking issues and design of all-optical packet switching network architectures.



Zoltan Gál graduated at "Traian Vuia Technical University" (Timisoara, Roumania), where he obtained M.Sc. degree in computer and electrical engineering in 1990. Since 1993 he is Head of computer network section at Kossuth Lajos University (Debrecen, Hungary). He is working currently towards his Ph.D. degree at Technical University of Budapest, with a special interest in ATM cell switching technology in mobile network environment. His research interest includes MAC and network layer protocols for LANs, wireless and mobile networks.

H. L. Phan received the electrical engineer degree from the Technical University of Budapest in 1983. From 1983 to 1990 she was with the Department of Electrical Engineering at Technical University Hanoi, Vietnam. Currently, she is working towards the Ph.D. degree at the Technical University of Budapest. Her research interests include performance evaluation of computer communication networks, broadband ISDN and multimedia systems.

Csaba Szabó received his M.Sc. degree from St. Petersburg State University of Telecommunications, Russia, the Dr.Univ. degree from the Technical University of Budapest and the Cand.Sci. degree from the Hungarian Academy of Sciences in 1968, 1976 and 1985, respectively. Since 1970 he has been a faculty member at the Technical University of Budapest. His research interests include radio data networks, integrated services local and metropolitan area networks and high speed networks. He led several research and development projects in these areas, published over two dozen of international journal and conference papers and a large number of papers and research reports in Hungarian, he is a co-author of two textbooks and three monographs. Dr. Szabó is a member of Editorial Board of "Computer Networks and ISDN Systems", has been serving on program committees of several international conferences, is a member of ACM and a Senior Member of IEEE.

A PROPOSED ARCHITECTURE FOR AN ADVANCED ATM SIGNALLING PROTOCOL

N. DOWNEY and A. PATEL

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY COLLEGE DUBLIN
BELFIELD, DUBLIN 4, IRELAND
E-MAIL: N.DOWNEY@NET-CS.UCD.IE; APATEL@CCVAX.UCD.IE

In the rapidly moving world of high speed communications many contradictions are present. This paper examines the problem of the lack of capabilities inherent in the current ATM signalling protocols. While ATM has been designed to cater for a very broad sweep of requirements, very little has been done about an equally broad signalling protocol for ATM. The paper begins with examining ATM and progresses with giving an overview of signalling systems in general. Included in this overview are the current ATM signalling protocols and an explanation of their development and history. The protocols are then examined to show their weaknesses, which originate mainly from their origins. With these limitations in mind, the design of a new signalling protocol architecture is presented. The design concentrates on extending the functionality of the protocol. Finally the paper concludes with a discussion of the issues related to evaluating a new signalling protocol.

1. INTRODUCTION

ATM or Asynchronous Transfer Mode, is rapidly becoming one of the premiere communications technologies. ATM began life back in the late 60s and early 70s in the laboratories of the University of Cambridge (UK), CNET and in particular AT&T Bell Laboratories Inc. It was not until 1983 that silicon technology had sufficiently matured to allow practical realizations of ATM switches and other hardware to be realized. It was around this time that the telecommunications industry became interested in ATM as a possible technology which would provide additional capabilities above and beyond those available from the Plain Old Telephone System (POTS) and the Integrated Services Digital Network (ISDN). While the technology of data and telecommunications has been progressing at a rapid pace, that of the signalling required to support communications has lagged behind somewhat. Signalling refers to the control information required to initiate, terminate and in general manage connections on a communications network.

In order for the full potential of ATM to be realized a more sophisticated signalling protocol than the one currently being used will be required. A new signalling protocol should support the full capabilities of ATM both for the currently envisaged services and for future services, which may be radically different.

This paper presents an architecture for a new signalling protocol which will allow the full power of ATM to be utilized. The paper begins with a brief overview of the reasons for a new signalling protocol as well as the problems with the existing one. Following this a brief discussion of ATM, and a review of signalling systems is given. Important issues for a new signalling protocol are then discussed. Finally an architecture is presented for the new signalling protocol, and a methodology for evaluating its performance is also given.

2. NEED FOR A NEW SIGNALLING PROTOCOL

The need for a new Signalling Protocol arises out of the fact that the current ATM signalling system is based on that of N-ISDN and that many of the interesting and useful capabilities of ATM are not being utilized and will not be until the signalling protocol is updated. Of primary importance in this respect is the fine granularity of control, which is possible in ATM, over the use of Virtual Channels (VC). This granularity of control indicates the following capabilities which should be supported but which are not by the current ATM signalling protocols [1]:

a) The capability to support multiparty and multiconnection calls:
- Support for symmetric and asymmetric calls (e.g. low or zero bandwidth in one direction and high bandwidth in the other).
- Simultaneous establishment and removal of multiple connections associated with a call.
- Add and remove connections to and from an existing call.
- Add and remove a party to and from a multiparty call.
- Correlate connections composing a multiconnection call.
- Reconfigure an existing multiparty call by merging or splitting the original multiparty call into fewer/more calls.

b) The capability to control ATM virtual channel and virtual path connections (VCCs and VPCs) for information transfer:
- Establish, maintain and release ATM VCCs and VPCs for information transfer.
- Support of communication configurations on a point-to-point, multipoint, and broadcast basis.
- Negotiate the traffic characteristics of a connection at connection establishment.
- Ability to renegotiate traffic characteristics of an already established connection.

c) Others:
- Capability to reconfigure an already established connection, for instance, to pass through some intermediate processing entity such as a conference bridge.
- Support interworking between different coding schemes.
- Support interworking with other network types.

In additional the following general features have been identified as being missing from the existing signalling protocol:
- modular construction, i.e. allowing easy maintenance and updating of the protocol;
- object oriented (hides complexity by the use polymorphism, encapsulation and abstraction [2]);
- security for the signalling information;
- separation of Call Control[1] (CC) and Bearer Control[2] (BC) allowing more sophisticated call setup;
- features for dealing with anycast messages;
- ability to perform complex negotiations (e.g. multiparty negotiation);
- the ability to provide additional synchronization for multimedia data flows above and beyond that provided by a synchronous channel (Constant Bit Rate (CBR) communication).

With the above needs in mind there is clearly a need for a new ATM Signalling Protocol.

3. OVERVIEW OF ATM

The three planes of the ATM Protocol Reference Model (Fig. 1) correspond to three different types of data that may exist with an ATM transmission medium at any instant. The Control Plane represents signalling traffic, the User Plane represents user-to-user data traffic and the Management Plane represents management data relating to both of the other planes as well as the protocol layers and any resources that may be utilized.

This paper deals with ATM signalling which operates mainly in the control plane, but could operate in the user plane if user-to-user signalling is required. The interface between the signalling protocol (or the control plane) and the end-user application is usually implemented through the use of an Application Programming Interface (API). An API is a set of routines or functions that can be invoked from a standard

[1] Call Control refers to the end-to-end processing of the call and not to any physical channels required to make this happen.

[2] Bearer Control refers to the node-to-node communication channel that must be setup between every node before a call can be completed.

programming language and provide high-level control of the signalling protocol. An API for the new signalling protocol is not discussed in this paper.

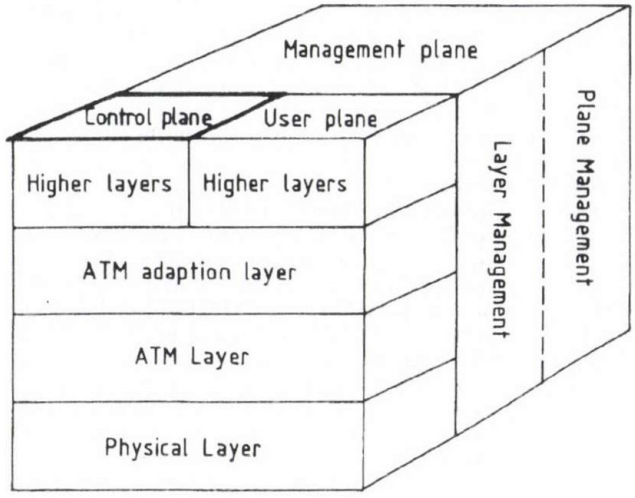


Fig. 1. ATM Protocol Reference Model

Even though ATM is "asynchronous", it is based on a variant of Time Division Multiplexing (TDM). In TDM a channel is represented by a periodic reoccurrence of a particular time slot on the transmission medium. Such a scheme (called Synchronous Transfer Mode (STM)) was proposed for use in B-ISDN but was not accepted on account of the inefficient usage of capacity that would result. ATM gets over this problem by allowing a channel to use any free slot on the transmission medium. Various types of ATM Adaptation Layers (AAL) have been defined, Table 1 describes five types which have been defined by the ATM Forum [3].

Table 1. ATM Forum Defined Service Classes

AAL Type	Characteristics
Constant Bit Rate (CBR)	• Circuit Emulation • Voice traffic • Constant bit rate video (MPEG/JPEG)
Variable Bit Rate – Real Time (VBR-RT)	Video and voice with silence removed and with tight delay bounds.
Variable Bit Rate – Non Real Time (VBR-NRT)	Transaction data which has less stringent requirements on delay bounds than VBR-RT.
Available Bit Rate (ABR)	No guarantee on cell loss or delay variation, suitable for network traffic (e.g. LANs) as it does include feedback on network congestion.
Unspecified Bit Rate (UBR)	No guarantee on cell loss or delay variation, suitable for network traffic only on a lightly loaded network due to lack of feedback on network congestion.

In [4] ATM is identified as the best candidate available for multimedia communication. The advantages identified include:

- large bandwidth;
- connection oriented and yet can exploit statistical multiplexing gains;
- small delay and delay jitter[3];
- multicast communication[4].

On the other hand ATM does suffer from some deficiencies in this respect, namely:

- it is not isochronous (there is not a direct timing relationship between the sender and the receiver due to delay variations introduced within ATM switches);
- traffic management, Connection Admission Control (CAC) and traffic policing have not been standardized yet.

Also ATM does not provide, yet, any control information which could be used to provide coordination between various multimedia sources. In essence what is required is a coordination service between several multimedia data types. Such coordination could be assisted by the signalling protocol providing special synchronization messages and the associated underlying functionality. In [5] various issues and models for representing synchronized multimedia data sources are discussed. In [6] a scheme is discussed for multimedia synchronization of data sources originating from a single point. Work can be found in [7] on the simultaneous routing of groups of Virtual Channels. Such a grouping would be very beneficial for applications such as videoconferencing, where similar types of data have to be transmitted together. [8], [9] both provide a broad review of many aspects of multimedia, including the various types of synchronization and some useful categories for synchronization. Both of these papers talk about media ropes and strands, which are ideas used in the synchronization of multimedia information. In [10] two protocols are discussed which work together to synchronize multimedia communications. The Application Synchronization Protocol (ASP) deals with the end-to-end synchronization while the Network Synchronization Protocol (NSP) which deals with trying to insure that particular multimedia objects arrive at their destination in a sequence that has been predefined. [11] develops these ideas much further and presents details of the signalling primitives necessary to provide synchronization. The new ATM signalling protocol will include features such as the above.

4. REVIEW OF SIGNALLING SYSTEMS

The idea of a signalling system has its origin in the Plain Old Telephone System (POTS). Within a POTS, signalling (at the subscriber end) is very simple. Once the hand-set is lifted the circuit to the local exchange is completed and the user dials the number required [12]. A series of pulses or different tones are used to indicate the number to be called.

More recently the Integrated Services Digital Network (ISDN) has begun to replace the POTS. The signalling within ISDN is much more sophisticated than within the POTS. Being digital rather than analogue allowed the designers of ISDN much more freedom. Instead of being restricted to analogue signals within the audio frequency range, now signalling messages could be sent in a manner independent to the data on the line.

ISDN was designed with "Out of band signalling", i.e. signalling information is carried on D-channels (16 kb/s) while data is carried on separate B (64 kb/s) channels[5]. This amounts to a separate channel for the data and for the signalling messages. With this capability ISDN can create services which operate independently of the data being transferred. The signalling protocol for N-ISDN is Q.931 [13]. One of the main problems with Q.931 is that Call Control (CC) and Bearer Control (BC) are amalgamated. This restricts the functionality of the protocol in terms of setting up a connection rather severely. CC deals with the processing of end-to-end call messages while BC deals with the setting-up of the intermediate connections which are required for the call.

Q.931 signalling messages are transferred between the end-users and various network functional entities and invoke some action or change of state of the protocol. The signalling messages of Q.931 are show in Table 2.

The sequence of messages exchanged between signalling entities are often represented in Time Sequence Diagrams (TSD). Such diagrams show, in chronological order, the order of message flow. In ISDN the TSD for a simple call from one user A to another B via the network is as follows (messages are listed in Table 2):

[3] While this statement is true at low network loads it has yet to be checked in real networks under substantial loads.

[4] The ability to multicast depends directly on the design of the ATM switch.

[5] This is the Basic Rate Interface of ISDN. For Primary Rate Interface the D channel is 64 kb/s and there are 30 B channels.

Table 2. Q.391 Signalling Messages

Call Establishment Messages	Call Information Messages	Call Clearing Messages	Miscellaneous Messages
ALERTING	USER INFORMATION	DISCONNECT	SEGMENT
CALL PROCEEDING	SUSPEND REJECT	RESTART	FACILITY
PROGRESS	RESUME REJECT	RELEASE	REGISTER
SETUP	HOLD	RELEASE COMPLETE	NOTIFY
CONNECT ACKNOWLEDGE	SUSPEND ACKNOWLEDGE	RESTART ACKNOWLEDGE	STATUS ENQUIRY
SETUP ACKNOWLEDGE	RESUME		CONGESTION CONTROL
CONNECT	HOLD ACKNOWLEDGE		INFORMATION
	SUSPEND		STATUS
	RESUME ACKNOWLEDGE		
	HOLD REJECT		
	RETRIEVE		
	RETRIEVE ACKNOWLEDGE		
	RETRIEVE REJECT		

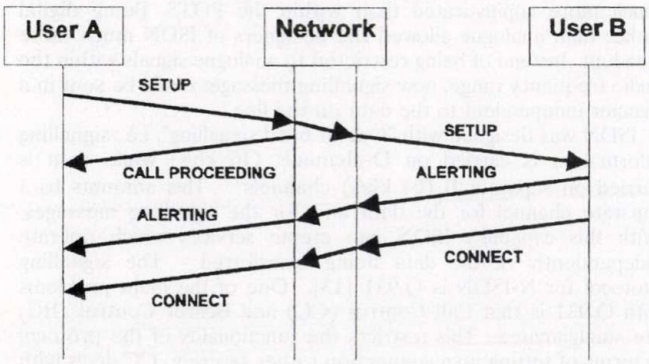


Fig. 2. Time Sequence Diagram for Basic Call Setup in N-ISDN

With the above capabilities many interesting services have been created, such as three party calls, call waiting, call forwarding etc. ISDN does however suffer from some major limitations. These are:

- lack of multimedia support (especially in terms of bandwidth and control of data flows);
- lack of scalability up to higher bandwidths;
- lack of separate call and bearer control;
- lack of flexibility.

In the 1980s when work began on ATM standardization, it was only natural that N-ISDN signalling should have formed the basis for ATM signalling. N-ISDN signalling was used within ATM for two reasons:

- to provide a quick solution to the problem of a signalling protocol for ATM;
- allow an easy integration of N-ISDN and ATM.

Only relatively few changes were made to the N-ISDN signalling protocol (Q.931). The updated version of Q.931 for ATM was initially called Q.931B. Since then it has become more stable and is now referred to as Q.2931 (by the ITU-T). While Q.2931 [14] is a different protocol it makes maximum use of the lessons learned during the Q.931 standardization procedure. The main changes were to the:

- Broadband Bearer Capability Information Element (IE);
- a new Cell Rate IE was added;
- a new AAL parameter IE;
- changes were made to the Channel Identifier IE.

In addition several messages and procedures were modified in a minor way but remain essentially the same as Q.931.

Since the initial specification of an ATM signalling protocol (by the ITU-T), it is mainly the ATM Forum (ATMF) which has been the leader in specifying ATM signalling protocols. The current version of the ATMF signalling protocol is called User-Network Interface (UNI) 3.1 [15]. The ITU-T adjusted their protocol (Q.2931) to conform to UNI 3.1. Several minor differences still exist between UNI 3.1 and Q.2931 however. The ATMF is at present progressing with UNI 4.0 which will include capabilities such as true multiparty calls etc. It is also working on Network-Node Interface (NNI) signalling protocols.

Within the Network-Node Interface (NNI) area the standardization bodies agreed to use a modification of the Signalling System No. 7 (SS#7) protocols. SS#7 is the signalling protocol used within the public network part of N-ISDN and between N-ISDNs in different countries. ATM is used as a transport mode within the public network at the lower three layers of the NNI stack. ATM cells will be encapsulated within SONET/SDH frames. The upper layers in such a scenario would correspond to those of the SS#7 network [16].

4.1. Other Related Work

This section examines the other influential work that is and has been carried out in the area of ATM signalling protocols.

4.1.1. MAGIC [17]

The most recent work on a new signalling protocol has been that of the RACE II project MAGIC – Multiservice Applications Governing Integrated Control (R2044). MAGIC is approaching this work from the point of view of services and what does the new signalling protocol for ATM need to be able to support existing as well as future services. MAGIC is basing the new protocol on the use of the OSI Layer Application Layer (7). Also MAGIC have concluded that the User Network Interface (UNI) and the Network-Node Interface (NNI) should remain different with the UNI only offering a subset of the functionality possible in the NNI[6].

4.1.2. gNET [18]

In this thesis a signalling protocol is presented as part of the work. The protocol has the following key features:

- End nodes consist of an ATM Data Link Interface (ADLI), containing the signalling entity, AAL service entities, and identified by a 64 bit address.
- AAL Users are identified by a 12 bit protocol identifier and a 4 bit AAL service specification.
- Each ADLI contains some form of traffic shaping functionality for outgoing traffic.
- Basic unicast and multicast connection establishment and removal are provided. A restricted set of traffic parameters are used for connection admission control.
- Virtual Path and Virtual Channel Connections are supported. Sufficient signalling information for the preceding services is carried within a single cell.
- Shared medium links, where multiple nodes connect to a single switch port, are supported.

[6] In particular Call Control (CC) and Bearer Control (BC) will be separated at the NNI but not so at the UNI. MAGIC argues that this is not necessary at the UNI as there is only a point-to-point connection between the end-terminal and the nearest switch.

- End node and switch node signalling entities are peers.
- All nodes use a common VPI/VCI for signalling.
- All connections are unidirectional.

Much valuable work has been done in the gNET protocol, even if it is targeted at a LAN environment and contains some (relatively minor) local requirements.

4.1.3. EXPANSE [19]

The EXPANSE protocol, which was developed in Bellcore, contains much of the concepts inherent in this work. EXPANSE is also an Object-Oriented (OO) signalling protocol as is the MAGIC protocol and has formed the basis for much of the current work on ATM signalling protocols.

5. ISSUES FOR A NEW SIGNALLING PROTOCOL FOR ATM

This section discusses important issues which must be addressed by a new ATM signalling protocol and how they can be met in practice.

5.1. Call Control (CC) and Bearer Control (BC) Servers

In the current N-ISDN signalling protocol (Q.931) [13] there are three different signalling functionalities [20]:

- Bearer Connection Control (BC) functions;
- Call Control (CC) functions;
- Service Control functions.

The Bearer Control (BC) has a link-by-link significance and refers to a "physical" connection, while the Call Control (CC) has an end-to-end significance and is a logical association. Fig. 3 shows a simple example of how the BC and CC entities interact. The CC deals with the end-to-end call processing which the BC establishes the hop-by-hop bearer (i.e. voice or data) connections. A more precise definition[7] of exactly what a BC and a CC are can be found in [1]:

"Call Control is the entire set of procedures necessary to establish, maintain, and release an association between endpoints. CC can be used for functions like compatibility and availability checking configuration control etc."

and

"Bearer Control deals with the setup, maintenance and release of connections in a telecommunications network."

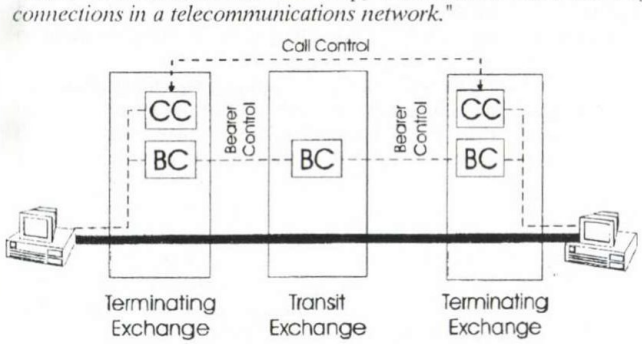


Fig. 3. Example of Call and Bearer Control

The separation of CC and BC functionalities has the following advantages [20], [1]:

1. Actions, such as negotiation of Quality-of-Services (QoS) during the setup phase, compatibility checking and provisioning may be done *before connections are setup.*
2. *Better distribution of functional entities in the network.* This is possible because not every node in the network need process Call Control (CC) messages but will have to process Bearer Control (BC) messages. This can be understood by looking at Fig. 3 above.
3. *Efficient and flexible utilization of network resources.* This is because the Call Control (CC) entities will be free to route CC messages in any manner to the next such entity on account of there not being any need for Quality-of-Service based routing at this stage of the call. Also because there is no waste of network resources (bearer bandwidth) during the CC phase (assuming the CC phase proceeds the BC phase), thus the call is more likely to be successful and the network overall will be more efficiently utilized.
4. Multimedia services can be controlled and offered in a *flexible manner.*
5. *New bearers can be easily introduced* without major modifications of the handling of the services.
6. Connections belonging to the same call can be *routed independently* of each other.

This separation of Bearer Control (BC) and Call Control (CC) allows three scenarios for basic call setup:

1. simultaneous call and bearer setup (N-ISDN scenario);
2. bearer connections setup prior to the call being made (allows very fast call setup but assumes apriori knowledge of the call);
3. call connection setup before the bearer connections are completed (allows the network resources to be efficiently utilized but leads to slow setup times).

The design of the BC and CC servers will be based on the use of the OSI Commitment, Concurrency and Recovery (CCR) protocol [21]. This protocol allows participating entities to coordinate their actions. The critical point here is that in any multi-entity communication, various entities must ensure they are available before the full communication can begin. By utilising the CCR protocol many types of sophisticated ATM calls can be setup (e.g. multiparty, multimedia etc.). It can also be used to facilitate the separation of the control of the bearer connections from that of the call itself.

The CCR protocol is described as a two-phase commit, i.e. remote entities reply if they are able to do a particular action, the initiating entity then gives the go-ahead for the action to be completed. With this type of scheme the actions of many remote entities, which may be vital if the overall action is to have any meaning, can be coordinated. The action is referred to as an atomic action. The primitives for the CCR protocol are listed in Table 3.

Table 3. OSI CCR Primitives

OSI CCR Primitive	From	Description
C-BEGIN	Master	Begin an atomic action
C-PREPARE	Master	End of phase 1; prepare to commit
C-READY	Slave	Slave is able to do its work
C-REFUSE	Slave	Slave is not able to do its work
C-COMMIT	Master	Commit the action
C-ROLLBACK	Master	Abort the action
C-RESTART	Either	Announce that a crash has occurred

All communications between CC/BC and any other entities will usually be a CCR type exchange.

5.2. Negotiation and renegotiation

Another important issue is that of negotiation of parameter values. At present, during a simple call setup, various parameters are negotiated between the end stations, these include Broadband-Lower Layer Information (B-LLI) information element. Depending on the delay (i.e. distance) between the two end-stations this negotiation may or may not be acceptable. The following formula gives an approximation to the total delay experienced by a sequence of messages between two stations.

$$D = N(L/c + d)\,,$$

where:

D Total delay involved, measured in seconds.
N Number of messages sent between the stations.
L Total distance between the stations, measured in metres.
c Speed of signals along the transmission medium, measured in m/s.
d Delay in stations between receiving a message and responding with the reply, measured in seconds.

[7] This definition is an enhanced version of the CCITT version which does not separate a call from its connections.

The calculation can be approximated by assuming that the speed of signals along the transmission medium is the same as that of light through air i.e. 3 x (10^8) m/s and that the stations process the messages in zero time. Therefore the delay caused by two end-stations 500 km apart and a negotiation requiring four messages will be at least:

$$4 \times (500 \times 10^3)/(3 \times 10^8) = 7 \text{ ms.}$$

In practice such a calculation assumes that there are no other delays due to:

- retransmissions due to errors/lost messages;
- passage through networking devices (routers/bridges etc.);
- conversion from one protocol to another;
- congestion etc.

On the INTERNET at a relatively quiet hour such a message exchange between Dublin and London could take between[8] 100 ms and 2s. The acceptability of such delays depends to a large extent on the application setting up the call. In particular delay is important for applications which deal with real-time human communications.

In a multiparty scenario the number of messages could be much bigger, thus leading to relatively large delays. Again, it would depend on the application as to whether such a delay was acceptable or not. There should be a mechanism to simplify such protracted negotiations. One possible solution is to have a "Negotiation Server" or NS. The technicalities of its implementation and as to whether it is in fact a real server are not part of this discussion. The NS would act as a broker in exchanges involving multiple party negotiations. To allow for non-negotiated discussion between parties there should be a flag in the initial message from a party that indicates if the party is prepared to negotiate or if it will either accept or reject the first offer it is made. All messages with this flag set would be sent to the nearest NS (the NS associated with the nearest switch). The NS would act to mediate between parties, offering either compromise solutions or else offering to trade one parameter against another. Following the initial setup message to the NS, all parties would be contacted by the NS to determine their requirements. The NS would then decide on suitable compromises or trade-offs and make sure that all parties are satisfied, or to exclude some, before continuing with the process. All the contacted parties would then complete the connection process to the original party.

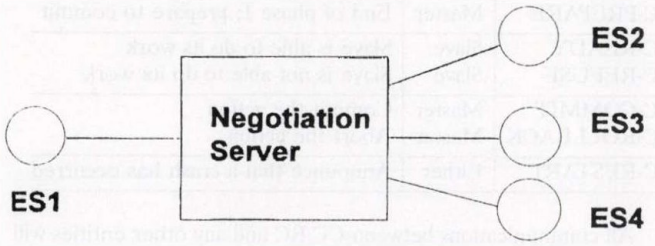


Fig. 4. Negotiation Server

This means that the NS must be able to deal with many simultaneous negotiations and dealing with many different requested services. Such a broad set of capabilities is going to make for a complicated piece of software. Much of the capabilities of the NS could be provided by a decision-based system or rule based expert system. What will be particularly difficult to provide is rules for trading off one parameter against another.

Another aspect of a Negotiation Server is to be able to renegotiate an existing connection. In a two party negotiation it would not be desirable to make use of the services provided by a NS. However in a multiparty connection this would be a desirable situation. As with the setup negotiation there should be a flag indicating whether the services of the NS are required or not.

Strategies for deciding on a common parameter value will also be investigated. Such strategies may include:

- Compromise at maximum/minimum acceptable to both parties.
- Compromise at an average value.

Simple strategies as above work fine with two parties, however with N possible parties then the negotiation process is potentially much more drawn-out. Strategies for such situations will include:

1. Every party advertise their maximum/minimum acceptable parameter values, after which the negotiation server examines all the possible values and either chooses a value or else suggests that some parties accept a parameter value that may be inconvenient or provide a worse service than would otherwise be desirable. It may be possible in some such situations for a few parties to trade less of one parameter for more of another (to be investigated).
2. Other possible strategies include one party making a suggestion on a possible value. Other parties then reply as to whether it is acceptable or not. The trouble with this approach is that the negotiation can be very drawn-out and protracted. A method such as limiting the process to several interactions and then taking a commonly acceptable value may be possible.

5.3. The Splitter

This server will deal with multicast/broadcast and anycast signalling messages. Much influential work has been done on broadcast/multicast for broadband [22], [23], [24], [25], [26]. The implementation of broadcast/multicast in ATM has been shown in these and many other papers to depend critically on the design of the ATM switch itself. On account of the relatively small number of signalling messages which will be broadcast/multicast (relative to those that would need to be sent for an application, such as Video-on-Demand), and to simplify the architecture of the switch, it has been decided to use a server approach as in [27], rather than a switching fabric approach.

There will be two possibilities for the operation of the Splitter in terms of multicasting; one will allow the adding of addresses to a list of multicast addresses held in a switch, the other will allow a list of multicast addresses to be added to a switch. The Splitter will also deal with incoming returns from a multicast message. Depending on the message it will either collect all returns and then pass on a single reply to the calling user or else pass them to the Negotiation Server for processing.

The broadcast part will operate by sending signalling messages to all known end addresses as well as one message to each known switch which it is connected to. These switches will forward the message to any other nodes they know about. A message will not be resent along its original path and if a switch encounters a broadcast message which it will not transmit it again. There will be a tag on each broadcast message identifying it. Broadcast messages will have a time-to-live value which is set such that within this period a switch will not re-transmit a previously transmitted message, but outside the period it will. The value of such a parameter will have to be carefully chosen to avoid avalanche conditions.

The anycast part will allow an End Station (ES) identify a provider of a particular service. This implies that each service would have to have a unique identifier. This could be difficult to manage in a global scenario if all services have to be notified to each server. It is probably better if only a limited number of well known services are added to the anycast service. These could include directory services, gateway services, connectionless services etc. An ES should be able to ask its local switch for information about where a particular service can be found. If the switch does not know where a particular service can be located then it should forward the message to another switch. Alternatively it could forward a multicast message to a select number of other anycast servers.

5.4. Compatibility of the signalling protocol with the OSI/RM

Having originated from within the telecommunications stan-

[8] This figure was arrived at by using the UNIX program PING and connecting between dallas.ucd.ie (source) and phoenix.doc.ic.ac.uk (destination). At 12:00 UTP on /9/15/95 ping returned the following round-trip times 81/222/1280 ms (min./avg./max.). Such figures should only be used as a guide to the possible delays involved. In essence, it gives a good indication about the magnitude of the delays involved.

dardization bodies it is no surprise that ATM does not comply with the Open Systems Interconnect/Reference Model OSI/RM (c.f. Fig. 1).

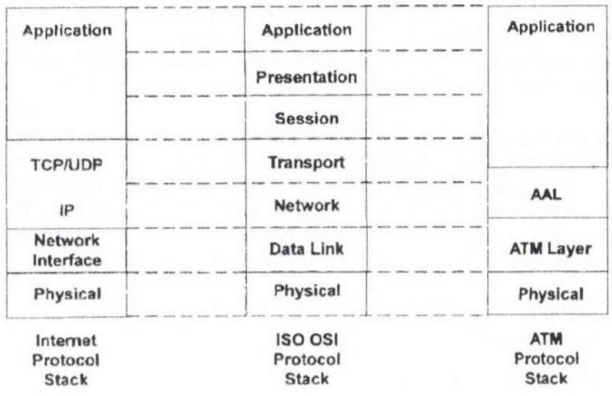


Fig. 5. Comparison of Internet, ISO OSI and ATM protocol stacks

Several good papers have been written about trying to fit ATM into the OSI/RM [28]. While certain similarities can be drawn, they are essentially different views of the world. It is therefore no surprise that ATM signalling protocols are also not based on the OSI/RM. Work is however underway to create ATM signalling protocols which fit into the OSI Layer 7 (Application Layer) scheme. Most of this work is coming from a RACE II project called MAGIC [1]. MAGIC is looking at ATM signalling protocols and has decided to adopt an OSI approach. This is due mainly to the fact that the top layer of Signalling System Number 7 (SS#7) [29] is OSI Layer 7 compatible and therefore any new NNI signalling protocol must be OSI compatible. Within MAGIC the same approach is being used at the User-Network Interface (UNI), i.e. with an OSI compatible signalling protocol. In addition within MAGIC there is no separation of the Bearer Control (BC) and Call Control (CC) at the UNI.

One of the principle aspects of this protocol is that the OSI Layer 7 view of the world is inappropriate to a protocol that is supposed to be as small and efficient as possible and to impose as small as possible a presence at a switch. One of the aims of this work is to prove that a full OSI Layer 7 set of protocols is too heavy-duty for such a system and that a more streamlined set of protocols would be better. In the architecture for the new signalling protocol the Signalling ATM Adaptation Layer (Q.SAAL) [30] will be used. The Q.SAAL is used to provide for the reliable adaptation of signalling messages to ATM cells. Fig. 6 shows the signalling protocol layers currently proposed by the ITU-T, at both the User-Network Interface (UNI) and the Network-Node Interface (NNI). The new protocol will operate on top of the Q.SAAL layer, in both the UNI and NNI configurations. Above the signalling protocol there will be the need only for an Application Programming Interface (API). This API will provide a standardized programming interface to allow applications to utilize the signalling protocol. Unlike other approaches, in particular the MAGIC project [1], which utilizes OSI Layer 7 entities to carry out the signalling, the new architecture will have more in common with the current Internet suite of protocols which do not have such a heavy-weight protocol stack. Thus the new architecture will not have any direct equivalent of the OSI Layers 4 to 6 inclusive. In particular there is no need for a Transport Layer as all its functions are handled as the functions normally associated with it are in fact provided by the signalling protocol and the protocol layers below it (Q.SAAL and ATM Layer). These are:

- addressing;
- segmentation and reassembly;
- connection establishment and termination;
- flow and rate control;
- error control;
- special data transfer services.

5.5. *Multimedia Server*

The multimedia server will act to coordinate multimedia data sources by acting on behalf of several remote users/applications. The actions of this server will include the following:

- synchronize data from different sources (in both multi-connection and multiparty);
- associate several Virtual Channels (VC) into a single meta-channel;
- alter an existing association or synchronization scenario between data sources;
- Quality-of-Service (QoS) management, which will involve coordination with the Negotiation Server.

The processing of multimedia sources will be on a per Virtual Channel which will allow the rapid processing of these channels during switching. The multimedia server will not be dealing with the actual multimedia data but rather the signalling messages necessary to control the data. Work on actual Multimedia Servers which deal with the actual data are detailed in [31], [32]. [33] presents another multimedia server which runs over TCP instead of ATM. The multimedia server will in particular be based on the work in [33], but will concentrate on multimedia signalling services rather that multimedia data services.

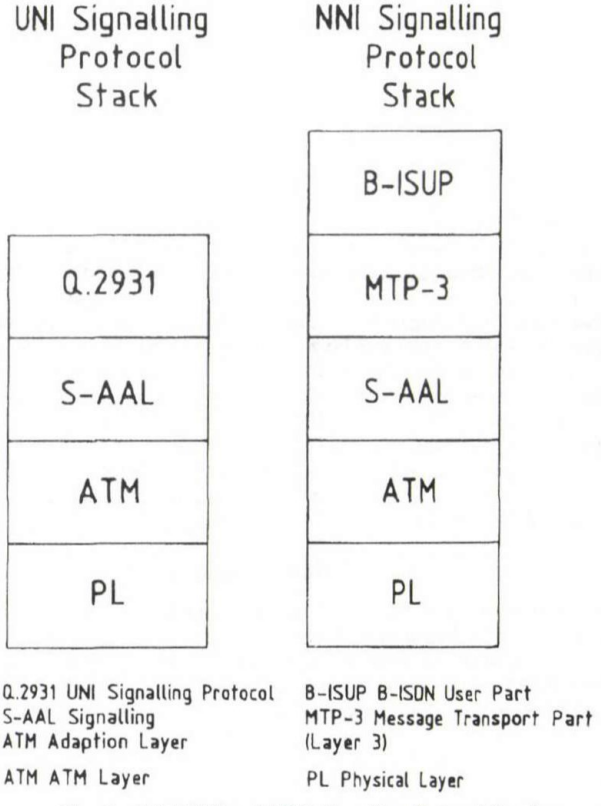


Fig. 6. ATM UNI and NNI Signalling Protocol Stacks

5.6. *ATM Security*

Security has been to a large extent overlooked within ATM. While there are many kinds of security measures and risks, for the purpose of this paper ATM security has been broken down into two poles. One is security of the data carried by ATM and the other is security of the signalling information. From the point of view of the data transfer, confidentiality is one of the more important security measures while for the signalling, access control is very important. With ATM taking some control away from the network operators, users will become empowered to carry out functions which were previously carried out by the Telecom operators. This will of course lead to many serious security risks. If the signalling messages can be made confidential then this problem would have been removed.

The different but related problem of how to make the data confidential to the sender and receiver is not dealt with by this paper in much detail. Surfice it to say that this is a difficult problem which is probably best tackled by the use of special high-speed encryption hardware.

Security both in ATM data transmission as well as in ATM signalling is of critical importance. In the papers by Katsavos, P. and Varadharajan, V. [34] and Clissmann, C. et al [35] recommendations are made to introduce an extra layer in the protocol stack. While it is understood that there are many differences between FR and ATM for the purpose of this discussion some similarities are relevant. Up to now security features have always been placed at OSI Layer 4 and above and it can be argued that ATM (and FR) is designed to be efficient by minimising the amount of processing at each network node. Both Katsavos, P. & Varadharajan, V. [34] and Clissmann, C. et al [35] suggest that security could be added below the adaptation of data bits to cells (i.e. below the AAL). It is not the job of this document to look at ATM security in general but to look at ATM signalling security in particular. Inexorably they are closely intertwined. One solution is to use a "stream" encryption scheme just above the AAL (both for data and signalling cells). Cells which are therefore not destined for a particular node would not be processed. This does however place more of the processing load on the end stations.

The alternative to security above the AAL is to have it just below. In this situation encryption would have to be provided by a hardware device in order to reduce the load on the network nodes (in terms of software). This has the advantage of fixed length data units (53 bytes). Again a "stream" cipher would be appropriate but most such ciphers require much larger units of data to be efficient. Such units of data and their associated processing would however impose timing problems on some data sources/sinks. If three bytes of "extra data" were added to each cell then DES CBC mode could be used.

Another aspect to security is that of security of the Telecommunications Management Network (TMN) [36]. TMN is the network infrastructure which will allow large telecommunications networks to be managed in an effective manner. It is normally stipulated that the TMN will be a separate network to that which it is managing. While it is still unclear how exactly TMN will operate in an ATM environment, what is obvious is that security will be of paramount importance to TMN. By implication the interface between the signalling protocol and the management plane must be secure to insure that no disruption of the operation of the TMN can occur due to unauthorized or unintentional access from the signalling protocol.

5.7. Integrated UNI/NNI signalling protocol

Due in the main part to the structure of the POTS, ATM has a slightly different flavour in both the LAN and WAN environment. On account of the use of SS#7 in the inter-exchange parts of the POTS/ISDN, ATM has started in a similar vein. Ideally ATM should be a ubiquitous network, i.e. it should appear to be the same throughout the network. Such a feature is sometimes referred to as concatativeness. While there are undoubtedly arguments in favour of a public network that can handle many different type of protocols, it would obviously be much better if ATM looked the same throughout the network.

While it is outside the scope of this paper to talk about the different UNI and NNI cell formats, it is worth briefly mentioning that the difference is only in terms of the GFC field of the cell header, whose usage is at best ambiguous and may possibly prove not of any real use at all. If this proves to be true then at some future stage it may be worthwhile to make the two of them the same.

This paper is looking at ATM signalling in particular so the recommendation in this respect is that the same signalling protocol be used for both the LAN and WAN scenarios. In practice this means the same signalling protocol for both the UNI and NNI. While at present the NNI signalling protocol runs over SS#7 and uses the B-ISUP (Broadband Integrated Services User Part) of Signalling System Number 7 (SS#7), it would be much simpler if there was only one signalling protocol for both the UNI and NNI. In the new ATM signalling protocol the UNI will be a subset of the NNI, thus simplifying the interworking between the two. Thus capabilities, which are not needed in the UNI, such as inter-domain charging can be implemented as additional capabilities in the NNI.

5.8. Interworking with Legacy Systems

A vital issue to be looked at is backward compatibility or interworking with legacy systems. While it is a commercial reality that there will always be backward compatibility issues these should not limit the scope of a signalling protocol. During this work interworking requirements will not be placed high on the list of priorities, however consideration will nevertheless be given to the need for interworking. In this respect it is planned to have a gateway server which will allow interworking between the new protocol and legacy systems. The server will process signalling messages from other systems (e.g. N-ISDN) leaving the rest of the protocol free to deal with its own signalling functions. The reason interworking is not of primary importance is to ensure that the new signalling protocol is restricted as little as possible.

6. DESIGN OF THE NEW SIGNALLING PROTOCOL

The following sections discuss the design for the new protocol. Reference is be made to the structure of the design as well as how it will be judged. The evaluation of a signalling protocol while being basically similar to that of any protocol does differ in some respects.

The design for the new signalling protocol is centred on adding many additional capabilities to the existing ATM signalling protocol. The new protocol will not be fully compatibly with the existing one. This is so as not to restrict the scope of the new protocol. Nevertheless those parts which can still be used will be kept, thus easing the interworking requirements.

An important aspect of the design will be its performance. With a large number of additional capabilities it would be easy to produce a protocol with much poorer performance, thus an analysis will be carried out on each part of the protocol to determine its effect on overall performance. It will then be possible to select those capabilities which will allow an acceptable level of performance.

6.1. Object Oriented Signalling Protocol

Unlike most of the existing ATM protocols, it is proposed that this protocol be Object-Oriented (OO). In [37] the Object Oriented technique is analysed to determine its relevance to call processing software. The results of [37] were very positive with only some drawbacks. The authors report that:

- OO is well suited to the domain of call-processing software;
- leads to a simplified design;
- faster development time due to more stable code;
- better match to the application area with OO than with conventional methods;
- less complex interfaces between software elements possible to make widespread use of inheritance;
- reduction in code size.

However the authors report that the use of OO did increase the execution time by up to 10 % on existing switching systems.

While the idea of an Object Oriented protocol is not new, much work is currently undergoing in the area of interaction and the provision of services within distributed object oriented environments. In particular the Telecommunications Information Networking Architecture – Consortium (TINA-C) [38] is very active in this area and the interactions of the signalling protocol objects will be compatible with the ideas in TINA-C. However it should be noted that most of the work in this area is concentrating on the provision of services and not on the underlying signalling protocols. The ideas from the TINA-C architecture which will be used in the new signalling protocol are the idea of Servers and their different (management interface, user interface, substance interface for dependencies with other objects and core interface which are internal to the object itself). Many of the other ideas in TINA-C are related more to Distributed Processing Environment (DPE) architecture for services.

Object orientation provides many valuable techniques which both simplify and enhance the design. The design of the new signalling protocol will broadly follow the guidelines in [39] and [40]. An initial list of object for the new protocol will include:

Table 4. Object Classes for a Signalling Protocol

Object Class	Usage
Message	Object associated with a signalling message
Bearer	Object associated with a bearer connection
Call	Object associated with a call connection
Service	Object associated with a service

The above objects are hierarchical, with those at the bottom containing those further-up. There will also be many different types of each main class of object. A request for a particular service will utilize one or more types of calls which in turn will utilize one or more types of bearer connection which in turn will utilize one or more types of message. It should however be pointed out that message objects can be used directly if required. As with any well behaved OO system, the data associated with each object is private, thus protecting parts of the system from other parts.

An important facet of Object-Orientation is that the protocol will be much more flexible and it will be easier to upgrade it or add new capabilities.

6.2. Functional Entities in the Signalling Protocol

The design of the system can also be described in terms of functionality. Each of the functional entities in Table 5 have been loosely labelled as servers:

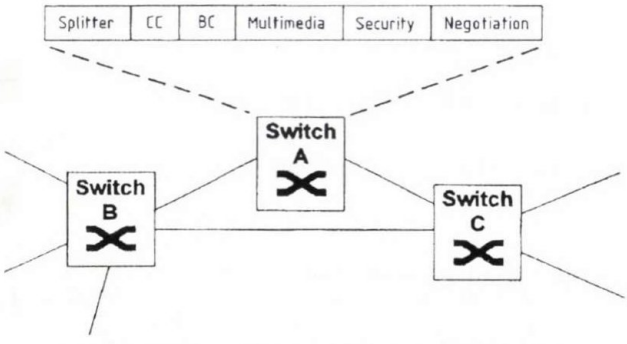


Fig. 7. Protocol Servers within an ATM Switch

Table 5. Signalling Protocol Servers

Server	Description
Anycast Server	Used to identify a given type of service on the network.
Negotiation Server	Used to mediate between hosts or other functional entities.
Multicast Server	Used to provide multicast and broadcast services in circumstances where such facilities do not exist in hardware.
Security Server	Used to provide additional security for signalling messages.
Multimedia Server	Used to deal with multimedia requirements, such as VC synchronization.
CC/BC Servers	These servers will deal with setting up CC and BC connections.

The exact details of the implementation are not being discussed here but show simply be thought of as parts of a switch which will deal with signalling messages and possibly send return messages to confirm that a particular operation has been successful/unsuccessful. The following diagram shows schematically how the various servers are located within each node of an ATM network.

7. EVALUATION OF THE SIGNALLING PROTOCOL

There are two aspects of evaluating a protocol, functionality and performance. Functionality evaluation is part of the testing and debugging phase of the development of the project. It involves evaluating the actions of the protocol to see if they perform the necessary actions in a sensible manner. Performance of the protocol will be evaluated on an ongoing basis which will be fed back into the development work to insure as high a performance as possible. At the end of the project a final performance evaluation will be given of the protocol.

A critical aspect of a signalling protocol is the number of messages necessary to carryout a particular operation. It is inevitable that in a protocol as complicated as the new ATM signalling protocol is, there is bound to be a penalty to pay in terms of the number of extra messages necessary to carry out an operation. Even in a very simple two party voice call, many different types of extra messages will be required. The following list illustrates the various types of messages which will be involved:

- call control;
- bearer control;
- negotiation;
- security.

In general the performance of a signalling protocol is little different to that of any other protocol. The main differences are in the critical nature of time related performance parameters, such as, call set-up delay, call release delay, and dependability related ones such as availability and reliability.

The ITU-T has formalized a scheme for describing Network Performance (NP) parameters [41]. The scheme is realized as a matrix of network performance measures. Each row represents one of the three basic and distinct communication functions. Each column represents one of the three mutually exclusive outcomes possible when a function is attempted. The shaded areas show entries that will be filled in during the evaluation of the signalling protocol.

Table 6. I.350 Protocol Evaluation

Performance Criterion / Function	Speed	Accuracy	Dependability
Access			
User information transfer			
Disengagement			

It is proposed in this project to evaluate the new signalling protocol using, not only the above parameters, but to also measure the effect of the new features which have been added to the protocol (i.e. negotiation, BC/CC separation etc.). This would result in three different tables, one for each of the FUNCTIONs listed in Table 7.

Table 7. Evaluation of new protocol features, where FUNCTION = {Speed, Accuracy or Dependability}

Performance Criterion / FUNCTION / Feature	Speed	Accuracy	Dependability
Security			
BC/CC Separation			
Multimedia			
Splitter (Anycast etc.)			
Negotiation			

Network performance measures can be also broken down into two categories [42]:

1. User-oriented performance such as call blocking probability, call modification request delay and call setup delay.
2. Network resource performance measures are those which are of interest to a network operator and describe particularly the requirements of the signalling protocol on network resources. Examples of this type of network performance measure include the required processor capacity or the load on the signalling transfer network.

The performance of the protocol will be derived both from simulation/analytical methods as well as from measuring the actual performance of the protocol in a real ATM network.

8. DISCUSSION AND CONCLUSIONS

The current ATM signalling protocol is based almost completely on that of N-ISDN. While this has allowed ATM to be rapidly developed and utilized, and has avoided a large investment in time and effort, it is now time to develop a new signalling protocol that will allow the full capabilities of ATM to be utilized. Since the current ATM signalling protocol is based on N-ISDN it is not unreasonable to say that it is not the best solution for ATM. ATM requires a signalling protocol that is both efficient and rich in capabilities. Multi-connection, multi-service and multimedia calls are among the important aspects missing from the current ATM signalling protocol. Security is one of the most vital aspects which is lacking. Without proper security ATM itself will not be accepted by the end-users. Of vital importance to ATM is that Call and Bearer Control be separated. Without such ATM will be little more than a fast version of N-ISDN. With such a facility many different types of call will be possible. Those where resources are allocated before the call is made as well as the more usual type where they are allocated after the request for the call has been accepted. Such capabilities will allow the full power of ATM to be utilized.

It is the assertion of this work that the UNI signalling protocol should be as small and efficient as possible but should not be restricted in any way. All the rest of the work on ATM has had a particular "flavour". Such a flavour could be described as being the provision of high speed communications with the minimum of overhead and processing necessary. Within the MAGIC project documentation it is accepted that their signalling protocol is based on assumptions made about future services and requirements. Even general assumptions can put restrictions on a system. If ATM is to reach its full potential then there should not be restrictions on how it operates. This work is therefore concentrating on a non-OSI protocol for the private-UNI and private NNI areas.

What has been presented in this paper is the draft outline of a new signalling protocol for ATM. The signalling protocol is both Object-Oriented and flexible.

9. ACKNOWLEDGEMENTS

The authors would like to acknowledge the support and assistance of TELTEC (Ireland) and Telecom Eireann.

REFERENCES

[1] Popple, G. W.: "*Requirements for Signalling – 2nd Deliverable*", RACE Project R2044, MAGIC, Oct. '92.

[2] Nicol, J. R., Wilkes, C. T. and Manola, F. A.: "Object Orientation in Heterogeneous Distributed Computing Systems", *IEEE Computer*, June '93.

[3] Rickard, N.: "ABR: Realising the Promise of ATM", *Telecommunications*, April '95.

[4] Stuttgen, H. J.: "Network Evolution and Multimedia Communication", *IEEE Multimedia*, Vol. 2, No. 3, Fall '95.

[5] Hirzalla, N., Falchuk, B. and Karmouch, A.: "A Temporal Model for Interactive Multimedia Scenarios", *IEEE Multimedia*, Vol. 2, No. 3, Fall 1995, pp. 24-31.

[6] Zhao, G. F. and Lye, K. M.: "Synchronization protocol for high quality and single-point playout multimedia applications", *Computer Communications*, Vol. 17, No. 9, Sept. 1994, pp. 677-683.

[7] Blankers, P.: "Functional Modelling for Target B-ISDN Signalling", *Proceedings of the IEEE Globcom '92* conference.

[8] Furht, B.: "Multimedia Systems: An Overview", *IEEE Multimedia*, Spring '94, pp. 47-59.

[9] Ramanathan, S., Vin, H. M. and Venkat Rangan, P.: "Towards personalized multimedia dial-up services", *Computer Networks and ISDN Systems*, Vol. 26, No. 10, July '94, pp. 1305-1322.

[10] Little, T. D. C. and Ghafoor, A.: "Spatio-Temporal Composition of Distributed Multimedia Objects for Value-Added Networks", *IEEE Computer*, October '91.

[11] Little, Thomas D. C. and Ghafoor, A.: "Multimedia Synchronization Protocols for Broadband Integrated Services", *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 9, December 1991, pp. 1368-82.

[12] Ronayne, J.: "The Integrated Services Digital Network: from concept to application", Pitman '87.

[13] "*Digital subscriber signalling system No 1 (DSS 1) – ISDN user-network interface layer 3 specification for basic call control*", ITU-T Recommendation Q.931.

[14] "*Q.2931*", ITU-T Draft Recommendation, Study Group 11/ Working Party 2, 2/11-21 C rev. 1, PL/11-24, Dec. '93.

[15] "*ATM – User-Network Interface Specification – Version 3.0*", The ATM Forum, PTR Prentice Hall. Version 3.1 is available on the Internet at URL=ftp://atmforum.com/pub/UNI/ver3.1/.

[16] Brill, A., Huber Dr. M. N. and Perti, B.: "Designing Signalling for Broadband", *TELECOMMUNICATIONS*, July '93.

[17] Gabriel, C. M. W., Heffels, J. J. M., Hou, X. and Ouibrahim, H.: "Heading towards an advanced signalling system for multimedia, multiparty services in broadband ISDN", *IEE Electronics and Communication Engineering Journal*, April 1993, pp. 103-112.

[18] Armitage, G. J.: "The Application of Asynchronous Transfer Mode in Multimedia and Local Area Networks", PhD thesis, University of Melbourne, Jan. 1994.

[19] Minzer, S.: "A Signalling Protocol for Complex Multimedia Services", *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 9, December 1991, pp. 1383-1394.

[20] Biocca, A., Paglialunga, A. and Siviero, M.: "Signalling Protocol for Broadband ISDN", *CSELT Technical reports*, Vol. XX, No. 5, Oct. 1992, pp. 395-403.

[21] "*Information Technology – OSI – Protocol for the Commitment, Concurrency and Recovery Service Element: Protocol Specification*", ITU-T Recommendation X.852, November '93.

[22] Turner, J.: "Design of a Broadband Packet Switching Network", *IEEE Transactions on Communications*, Vol. 36, No. 6, June 1988, pp. 734-745.

[23] Leam C-T: "A New Broadcast switching Network", *IEEE Transactions on Communications*, Vol. 36, No. 10, October 1988, pp. 1128-1137.

[24] Lee, T. T.: "Nonblocking Copy Networks for Multicast Packet Switching", *International Zurich Seminar on Digital*

Communications, Zurich, Switzerland, ETH, March 8-10, 1988.

[25] Byun, J. W. and Lee, T. T.: "The Design and Analysis of an ATM Multicast Switch with Adaptive Traffic Controller", *IEEE Trans. on Networking*, Vol. 2, No. 3, June '94, pp. 288-298.

[26] Doar, J. M. S.: "Multicast in the Asynchronous Transfer Mode Environment", PhD Dissertation, St. John's College, University of Cambridge, January 1993, (http://www.cl.cam.ac.uk/ftp/papers/reports/TR298-jmsd-thesis.ps.gz).

[27] Ballardie, T. P. F. and Crowcroft, J.: "Core Based Trees", *Proc. ACM SIGCOMM '93*, San Francisco, September 1993, pp. 1-11.

[28] De Prycker, M. et al.: "B-ISDN and the OSI Protocol Reference Model", *IEEE Network*, March '93.

[29] "*Introduction to CCITT Signalling System No. 7*", ITU-T Recommendation Q.700.

[30] "*Draft Recommendation Q.SAAL0 – B-ISDN Signalling ATM Adaption Layer Overview Description*", Melbourne, 8-12 Feb. '93.

[31] Rooholamini, R. and Cherkassky, V.: "ATM-Based Multimedia Servers", *IEEE Multimedia*, Vol. 2, No. 1, Spring '95, pp. 39-52.

[32] Wray, S., Glauert, T. and Hopper, A.: "Networked Multimedia: The Medusa Environment", *IEEE Multimedia*, Winter '94, pp. 54-63.

[33] Anderson, D. P. and Homsy, G.: "A Continuous Media I/O Server and Its Synchronization Mechanism", *IEEE Computer*, October 1991, pp. 51-57.

[34] Katsavos, P. and Varadharajan, V.: "A Secure Frame Relay service", *Computer Networks and ISDN Systems*, 26 (1994).

[35] Clissmann, C., O'Ciardhuain, S., Patel, A., Tschammer, V. and Wolisz, A.: "Some Notes on the Placing of Security Facilities in the ATM Architecture", *Project Report*, MESH Project, Aug. 1993.

[36] "*Principles and Architecture for the TMN*", ITU-T Recommendation M.3010.

[37] Guenther, W. and Wackerbarth, G.: "Object-Oriented Design of ISDN Call-Processing Software", *IEEE Communications Magazine*, April '93, pp. 40-45.

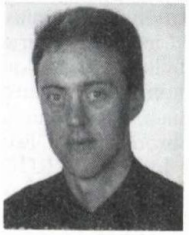
[38] Dupuy, F., Nilsson, G. C. and Inoue, Y.: "The TINA Consortium: Towards Networking Telecommunications Information Services", XV. International Switching Symposium Contribution, Berlin, Germany, 23/25 April, 1995, URL=http://www.tinac.com/market/swc/iss/fabrice_gunnar.rev2.ps.

[39] Guyer, S., Gunther, W. and Paul, S.: "Object Oriented Design Method for Switching System Software", International Computer Conference 1993, pp. 326-330.

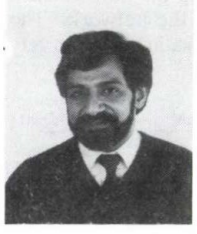
[40] Yamazaki, S., Kajihara, K., Ito, M. and Yasuhara, R.: "Object-Oriented Design of Telecommunication Software", *IEEE Software*, Jan. 1993, pp. 81-87.

[41] "*General Aspects of Quality of Service and Network Performance in Digital Networks, including ISDNs*", ITU-T Recommendation I.350, Section A.2, March 1993.

[42] "*B-ISDN Signalling Performance: Issues and Analysis*", 14th Deliverable, RACE Project R2044 (MAGIC).



Niall Downey received his B.E. (Electronic) and M.Eng.Sc. (Electronic) from University College, Dublin in 1988 and 1991 respectively. Since 1991 he has been a researcher in the Computer Networks & Distributed Research Group within the Department of Computer Science. He has been involved with several ESPRIT and RACE projects as well as many internal computer network projects. He is currently working on ATM and ATM signalling protocols.



Ahmed Patel received his M.Sc. and Ph.D. in Computer Science from Trinity College Dublin in 1977 and 1984, respectively. From 1978 to 1982 he was responsible for developing the Irish Universities Data Network and from 1982 to 1985 he developed the EuroKom computer conferencing and electronic mail service used by R&D projects in Europe. At University College Dublin he is a lecturer in Computer Science, and head of the Computer Networks and Distributed Systems Research Group, and centre director of Teltec Ireland. He is involved in various multi-national R&D projects in ESPRIT, RACE/ACTS, INFOSEC, AIM, COST and the Irish national IT and Telecommunications programmes. His main research interests include network management, security, protocols, performance evaluation, intelligent networks, CSCW and open distributed processing systems. He has published many technical papers and co-authored two books on computer network security and one book on group communications. He is a member of the Editorial Advisory Board of the Computer Communications and Collaborative Computing journals.

DESIGN AND PERFORMANCE ANALYSIS OF A DEVICE FOR THE EXECUTION OF OPERATION AND MAINTENANCE PROCEDURES IN ATM NETWORKS

A. HEMANI, T. LAZRAQ B. SVANTESSON, H. TENHUNEN

ESDLAB, DEPT. OF ELECTRONICS
ROYAL INSTITUTE OF TECHNOLOGY, ELECTRUM 229
KISTA, 164 40, KISTA, SWEDEN
E-MAIL: AHMED@ELE.KTH.SE. FAX: 46-8-752 12 40

A. POSTULA

DEPT. OF ELECTRICAL AND COMPUTER ENGINEERING
UNIVERSITY OF QUEENSLAND
BRISBANE, AUSTRALIA

Operation and Management functions (OAM) for ATM (Asynchronous Transfer Mode) perform the crucial task of detecting errors and performance degradation in the ATM network at the switch level and report it further. For 2.5 Gbit/s ATM nodes, these "housekeeping" functions are a severe bottleneck. Modelling OAM functions is necessary for evaluation of performance of an ATM system and also allows to experiment with different system solutions in early design stages. A complete model of OAM functions has been developed in VHDL simulation environment and verified by extensive simulations. Further, this model has also been used as an application for behavioural synthesis system that has been adapted for telecommunication applications. Synthesis tools helped identify the bottlenecks in the design when moving from 155 Mbit/s to 2.5 Gbit/s. In order to reach the desired throughput, a new revized model has been developed and verified by simulations. This paper presents the problems encountered in the OAM design and discusses the applied design methodology regarding simulation and synthesis. The results of simulation and synthesis of 155 Mbit/s OAM are also presented and the refined model for 2.5 Gbit/s is analyzed.

1. INTRODUCTION

ATM is a ITU-specified (previously CCITT) communications and switching technology for broadband services [1]. ATM switches available today or in the near future have port input/output speed of 155 or 622 Mbit/s [2]. Next step will be to raise the speed further with a factor of 4 up to 2.5 Gbit/s per port [3]. However, the fiber optic network being developed today allow even higher transmission rates. Therefore there is a large interest in developing ATM networks that allow speed of 2,5 Gbit/s or even beyond. The critical issue is to understand how performance trade offs and actual speed constraints affect the overall network performance.

OAM is part of ATM specifications that is responsible for detection of errors and performance degradation in the ATM network at switch level and to report it further. The principles of OAM are described in the ITU-Recommendation I.610 [4].

A part of the ATM network is depicted in Fig. 1. The OAM block is present on each physical link connected to the ATM switch. Fig. 1 shows that every ATM cell must pass the OAM block before it enters the switch, this means that the performance of the OAM block is critical for the speed of the network. The purpose of this work is to study the bottlenecks in ATM at very high speed related to OAM.

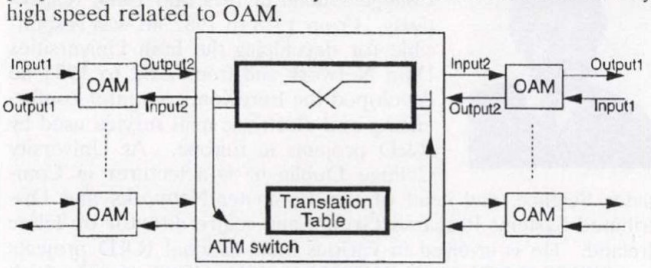


Fig. 1. The location of OAM block in the ATM network

This paper presents our design and research experience with implementing F4 functionality in hardware. Section 2 presents related work in this area. Section 3 presents OAM principles. In Section 4 we present the F4 model and discuss rationalise of its design. Section 5 is focused on the synthesis aspects of the model and presents the refined VHDL model suitable for behavioural synthesis. Problems specific to the synthesis of communication circuits are briefty discussed and our solutions are presented. Finally the synthesis results of 150 Mbit/s architecture are presented and analyzed. Section 6 covers the structure, rationalise and simulation results of the refined architecture suitable for 2.5 Gbit/s ATM cells.

2. RELATED WORK

The F4 device can be implemented in software, ASIC based hardware, or a combination of hardware and software. This work is based on a full ASIC based hardware implementation using high level synthesis techniques. After extensive search in the public domain and communicating with people who worked with OAM in industry especially Ericsson, we came to the conclusion that there is no published research which describes comparable work.

3. REVIEW OF OPERATION AND MANAGEMENT FUNCTIONS IN ATM

The different layers of the OAM functions are denoted by ITU-T [4], [5] as F1 through F5.

ATM layer	F5
	F4
Physical Layer	F3
	F2
	F1

Fig. 2. OAM levels

Fig. 2. shows the five hierarchical levels of OAM. Flow F1 is the smallest recognizable physical entity for OAM, and is located between repeaters. The level F2 can consist of many F1 levels. Level F3 performs OAM functions at the terminating endpoints of level F2. Level F4 sends OAM information and monitors the traffic at any segment or connection endpoint of a virtual path. Level F5 extends between the segment or connection points at the virtual channel level. In this work we are not addressing the F1 to F3 flows because they do not operate at ATM cell level. The size and complexity of the F5 flow are nearly the same as those in F4. The only difference is that the F4 device checks the Virtual Channel Identifier (VCI) field on the incoming ATM cells while the F5 device checks the Payload Type Identifier (PTI) field. Only the F4 flow will be considered in this paper, the results are applicable to F5 as well. According to the ITU-T Recommendation I.610, the OAM functions for the F4 flow can be classified into fault management functionality, performance monitoring functionality and activation/deactivation functionality (Fig. 3). The fault management may be grouped into two general categories: alarm surveillance functions and failure localization and testing functions. In the alarm surveillance approach, a failure signal is generated at a network element which has detected a physical layer failure e.g. loss of cell synchronization. Two alarm indications are defined, Remote Defect Indication (RDI) and Alarm Indication Signal (AIS). The AIS alarm indication is generated by the node detecting a defect to alert the downstream nodes that a defect has been detected upstream. When receiving

an AIS alarm, RDI is generated by the terminating node to alert the upstream nodes that a defect has been detected downstream. After that alarm the surveillance procedures indicate that a problem has arisen, tests my be initiated to verify the existence of the reported problem, identifying the nature of the problem, and isolating its cause. Loopback information signals may be used to verify whether the reported problem is in its jurisdiction or not.

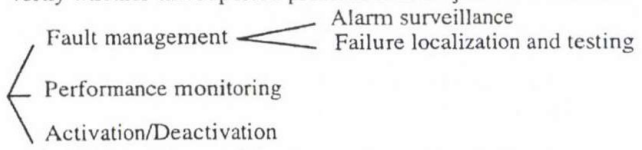
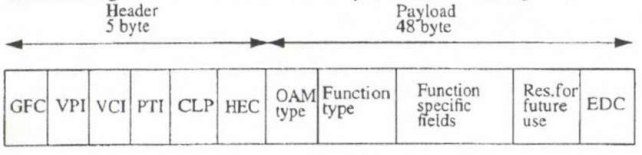


Fig. 3. OAM function for the F4 flow

Performance monitoring (PM) functions detect problems early enough to be corrected before they become more severe. OAM signals belonging to PM are inserted periodically in the user cell stream. Useful parameters e.g. cell loss ratio or cell insertion rate can be monitored thanks to the OAM signals. Performance monitoring also permits to provide the traffic performance during overloads and failures in the network by optimising the effective use of the network resources during those periods.

The activation/deactivation of performance monitoring is a system management function, initiated by the appropriate Management System or by the end user. Signals are used in the network to provide "handshaking" between the two ends of the monitoring entity. The performance monitoring activation/deactivation procedures serve mostly for establishing an agreement to start or stop monitoring and to coordinate and synchronize PM signals.



OAM Celltype	Value	Function type	Value
Fault management	0001	AIS	0000
		RDI	0001
		Loopback	1000
Performance Monitoring	0010	Forward monitoring	0000
		Backward reporting	0001
		Monitoring and reporting	0010
Activation/Deactivation	1000	Performance monitoring	0000

Fig. 4. The format of the OAM cell

The OAM information is sent in the ATM network via OAM cells. The OAM cell consists of a 5 byte header and a 48 byte payload as indicated in Fig. 4. A 24 bit label in the header identifies the logical connection to which the cell belongs. The label information consists of a 16 bit VCI and an 8 bit Virtual Path Identifier (VPI) which is used by a network of switching elements to route the cell to its destination. VCI values 3 and 4 in the header are used to carry F4 OAM flows. The payload contains the OAM type (fault management, performance monitoring, activation/deactivation), function type (AIS, RDI, loopback, etc..), function specific field (statistics, addresses, etc..), reserved area and the Error Detection Code (EDC).

4. THE F4 VHDL MODEL

The F4 block model [6] has two inputs and two outputs to connect the ATM switches, an input and output to interface the management system and an input from the physical layer. The flow coming from the upstream ATM switch goes from inputl to output2. Input2 and outputl are used for looping back the user and the OAM cells.The VHDL model (Fig. 5.) has been partitioned into four processes: an Input Handler, two Output Handlers, and an FMCellGenerator and a Timer. The following gives a brief description of the four processes:

The Timer has the task to clock the Input Handler and the FMCellGenerator by sending signals at specific times.

The Input Handler passes on the incoming user and OAM cells that are not targeted towards this node to the respective Output Handler (Fig. 5.). The cells arriving at inputl/input2 will be processed by the Input Handler and transmitted further to the Output Handler2/Output Handlerl. On receiving an AIS cell or an error signal from the physical layer, the Input Handler sends a signal to the FMCellGenerator to start generating RDI cells. The internal tables in the Input Handler contain mostly statistical information for performance monitoring. The performance management and the activation/deactivation cells are handled internally in the Input Handler. The activation/deactivation cells are generated by the Input Handler on getting a signal from the management system. The Input Handler reports also the results requested (e.g. statistics) by the management system by sending data.

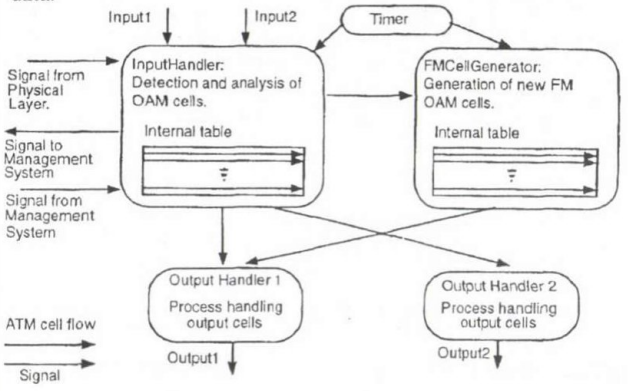


Fig. 5. The structure of the F4 model

The main task of the FMCellGenerator is to generate AIS, RDI or Loopback cells to the Output Handlerl. The internal tables in the FMCellGenerator contain information about which OAM cells will be sent on a pre-determined VPI. The FMCellGenerator is controlled by start and stop signals from the Input Handler.

After reading the incoming cells from the Input Handler and the FMCellGenerator the Output Handlers place them in a FIFO queue, as shown by arrows in Fig. 5. These cells are then sent from the F4 block to the ATM switch, giving OAM cells higher priority than the user cells.

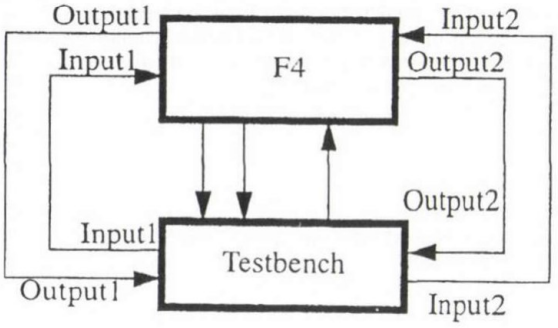


Fig. 6. Testbench

5. BEHAVIOURAL SYNTHESIS OF THE F4 MODEL

The transaction level F4 architecture shown in Fig. 5 offers several possible implementation solutions in terms of partitioning between hardware and software components. But to find the maximum performance we could obtain with the present architecture we decided to implement the F4 completely in hardware. Finding the limits for the present architecture, would give us insight into how to improve to be able to reach the 2.5 Gbit/s. target. Synthesis tools today accept design specifications at several different levels. Specifications at register transfer level and logic level are the most wide spread. However, this requires manually refining the behavioural model to make it a clock true model. In contrast,

behavioural synthesis tools allow a clock free higher level of abstraction for specification.

5.1. Synthesis aspects of the VHDL code and Partitioning

The transaction level F4 architecture shown in Fig. 5 was refined for the synthesis phase in several respects. The refined and partitioned architecture is shown in Fig. 7. This architecture was then verified to have the same functionality as the original architecture, by using the same test bench as that used for the original architecture.

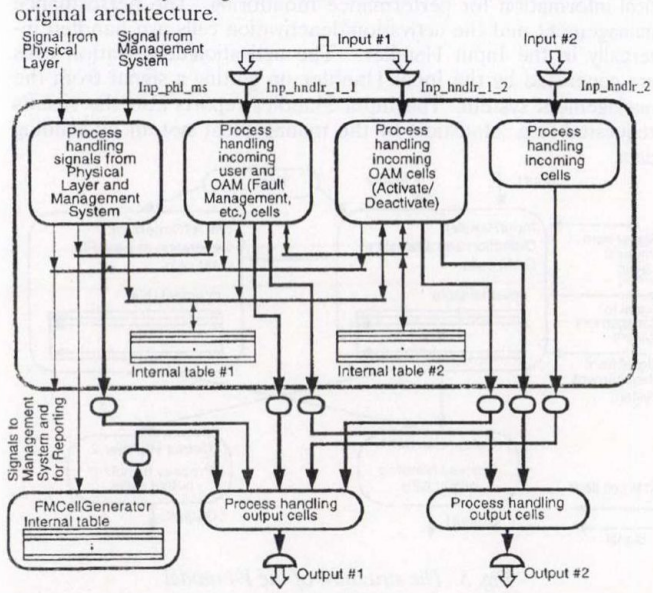


Fig. 7. Partitioned ATM F4 node

Sequentiality and Partitioning

The functionality of the Input Handler was organized as one huge monolith sequential process. This had two negative effects from the point of view of synthesis: a) The generated Register Transfer Level (RTL) code was too big for the backend logic synthesis tool to handle, and b) The algorithm describing the node's Input Handler implied an unnecessary sequential restricting processing of one cell or event at a time, and thus reducing the throughput unnecessarily.

Analysing the data dependencies between modules of the node allowed the partitioning of the original Input Handler process into four smaller processes which could be handled more flexibly by the logic synthesis tools. Although the partitioning introduced some duplicated parts and increased the total area of the node, the overall performance was improved because the partitions could work independently thus increasing the throughput of the node.

Data dependency considerations showed that the incoming cells from the second input are passed more or less directly to the first output and thus this part of the original Input Handler was extracted as an Input Handler #2. The rest of the original Input Handler was first divided between two principal inputs cells from the first input channel and signals from Physical Layer and Management System. As these two parts were data dependent, it required splitting the internal state table as well.

Wide Parallel Ports

The original model used parallel representation of ATM cells (424 bits). Although the parallel representation is very comfortable for modelling, it is impractical for synthesis with such wide ports. For this reason, these ports were realized as 8 bit wide ports. This resulted in serial transmission of cells one byte at a time. To reduce the idle time spent for transmissions by the Handlers and also to avoid possible deadlocks, the internal cell buffers were introduced because while transmitting a cell from one module to another both of them cannot perform any other operation. Although each introduced buffer can store only one ATM cell it frees the handler that is transmitting the cell, to perform other operations, for instance to receive the next cell.

Pre-defined Language features

The original model of the F4 node relied on VHDL language features to detect transactions and events on signals. For synthesis this was replaced with handshake signals to model the interface between the F4 node and the external world and between the internal modules as well.

5.2. Optimising the synthesis of F4 by adapting behavioural synthesis

The F4 model, like most telecom functionality dealing with protocol, is dominated by interaction with memory and control logic. We call such systems Control and Memory Intensive Systems (CMISTs). To maximise the throughput we could get by implementing the F4 functionality in hardware and creating a design that is efficient in terms of area, we adapted the synthesis strategies used by conventional behavioural synthesis (BS) tools to better handle characteristics of CMISTs. We enumerate these strategies and the adaptations (see [8] for details).

5.2.1 Target Architecture

As most of the current BS systems are general purpose, they do not have a pre-defined target RTL architecture. This greatly increases the complexity of the optimization problems in BS, resulting in designs that are very sub-optimal. Past research has shown that, by being application specific, it is possible to identify a generic target RTL architecture. Examples of such systems are Cathedral for Digital Signal Processor (DSP) applications and SUGAR for Microprocessor, etc.

In our strategy, when allocation detects a CMIST or when the user explicitly identifies a CMIST, the system maps the behaviour to a target architecture made up of controllers, address generators and memories.

5.2.2. Optimization strategy for CMISTs

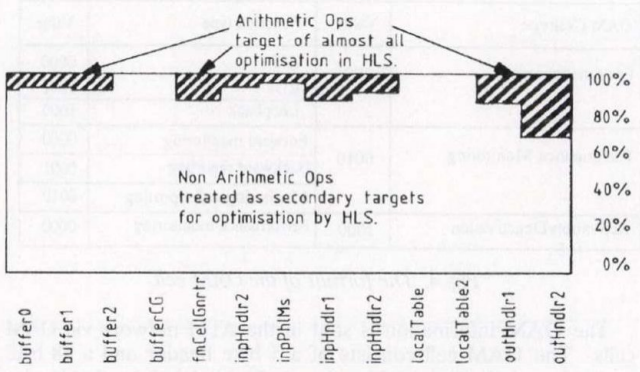


Fig. 8. CMIST characterization of F4

By experimenting with general purpose BS tools, we found out that the area optimization strategy of these tools does not work very well for F4 in particular and CMISTs in general. The principal strategy for area optimization in most BS tools is the reuse of RTL components like functional units, registers, multiplexers and buses. Further, among these RTL components, the optimization of arithmetic functional units seems to be the primary concern, relegating optimisation of registers, multiplexers and buses as secondary. However in CMIST, as a percentage of total area, arithmetic units are a small portion, whereas the dominant cost is the memory and relational functional units often with constant operands. This argument is substantiated in Fig. 8 which shows for F4 sub-modules how small is the percentage of arithmetic operations that are targeted by general purpose BS tools. In other words, the reuse strategy that is typical of BS tools today is of little use for CMISTs. Besides, the reuse logic added not only increases the overall area, it also slows down the design and thus has a negative effect on throughput.Recognising this fact, we apply an allocation strategy that retains in datapath only those operations that are candidates for the BS reuse strategy; other

operations such as relational operations with constant operands or memory indexing operations are moved to controller or allocated to the specialized address generator. This has related salutary effects: 1) The backend logic synthesis tools work better if the structured logic and random logic are segregated and optimized using different strategies, also, logic optimization techniques like product term sharing and multi-level logic minimization can reuse resources at logic level where the reuse failed at RTL level; 2) Operations moved to controller can be better optimized using the well developed Finite State Machine (FSM) optimization techniques like state assignment, minimisation etc.; 3) Memory address generation does not complicate the controller and can be optimized by specialized tools.

5.2.3. Serial parallel trade-offs for CMISTs

One of the much touted benefits of BS tools is the ability to do area/time trade-offs by using scheduling and allocation algorithms. These algorithms offer increased parallelism by using more functional units, and conversely they allow savings in area by using fewer functional units and serialising the implementation.

This strategy does not work very well for CMISTs in general and F4 in particular, because CMISTs being naturally sequential give scheduling algorithms very little freedom to do serial/ parallel trade-offs.

Though Area/Time trade-offs in terms of number of functional units may not have much potential for CMISTs, such trade-offs are nevertheless possible in terms of bandwidth to Memory and types of Memories because the CMISTs functionality is dominated by communication with memory.

5.3. Synthesis Results

The above described F4 ATM block has been synthesized using two different synthesis strategies and the results are presented in Table 1. Mietec 0.7 micron CMOS technology was used as the target technology for synthesis.

Table 1. Results for the F4 node in an ATM

Unit	Comparison of Allocation Strategies						CMIST characterisation.				
	With DP/Ctrlr trade-offs.			Without DP/Ctrlr trade-offs			CMIST ops			None CMIST	
	Area (gates)	Clock period (ns)	Clock steps	Area (gates)	Clock period (ns)	Clock steps	Rel. & Log.	Memory	Join nodes	Arithmetic	Total /% CMIST ops
buffer_0	402	9.48	9	522	6.13	13	9	4	14	2	29/93
buffer_1	456	10.25	9	546	7.26	13	9	4	17	2	32/94
buffer_2	418	10.40	9	594	7.26	13	9	4	17	2	32/94
buffer_cg	332	6.63	7	436	4.23	7	7	0	18	0	25/100
fm_cell_gnrtr	1475	14.79	51	1945	12.06	57	43	52	27	13	135/90
inp_hndlr_2	507	11.64	13	819	8.13	17	18	8	14	2	42/95
inp_phl_ms	3003	16.27	41	4780	13.47	71	186	35	195	12	428/97
inp_hndlr_1_1	6902	16.43	118	8631	11.89	185	237	138	219	63	657/90
inp_hndlr_1_2	4088	16.43	166	5109	12.24	57	247	144	221	41	653/94
local_table	646	7.12	8	1324	6.23	10	29	18	28	0	75/100
local_table_2	549	8.66	6	1188	7.10	8	27	20	23	0	70/100
out_hndlr_1	4938	16.43	33	4315	12.51	59	154	54	113	48	369/87
out_hndlr_2	2044	16.24	8	2702	10.10	28	29	14	39	22	104/79
Total	25760			31911			1004	495	945	207	2651/92

Table 1. shows two types of results: a) results with a synthesis tool that gives more emphasis on the construction of the controller part (with DP/Ctrlr trade-offs), and b) results with a synthesis tool that gives emphasis on the construction of the data path (without DP/Ctrlr trade-offs).

The third part in Table 1. gives the CMIST characterization of the synthesized units. Table 1. shows that when the percent of CMIST operations is higher than 80 % the tool with DP/Ctrlr trade-offs gives more optimal design e.g. low number of gates.

A first estimation of the delays in the different subblocks of the F4 ATM block could be calculated using the results in Table 1. This is done by multiplying the clock steps with the clock period.

At the speed of 2.5 Gbit/s the Input Handlers should have a maximum delay of 168 ns (time to process an ATM cell at 2.5 Gbit/s). Table 1. shows that the Input Handler blocks take much longer time than the required time. To reach the speed of 2.5 Gbit/s the Input Handlers have to be replaced with a faster design.

The current Output Handlers do not meet the timing constraints for a throughput of 2,5 Gbit/s because of non-optimal table search algorithm. This requires a thorough revision of the Output Handler algorithm.

6. OPTIMIZED F4 ARCHITECTURE FOR HIGH SPEED CELL TRANSFER RATE

Taking into account the bottlenecks of the design described above an improved model for OAM functions has been developed. The new model partitions the F4 algorithm into parts that can be pipelined, and removes the bottlenecks.

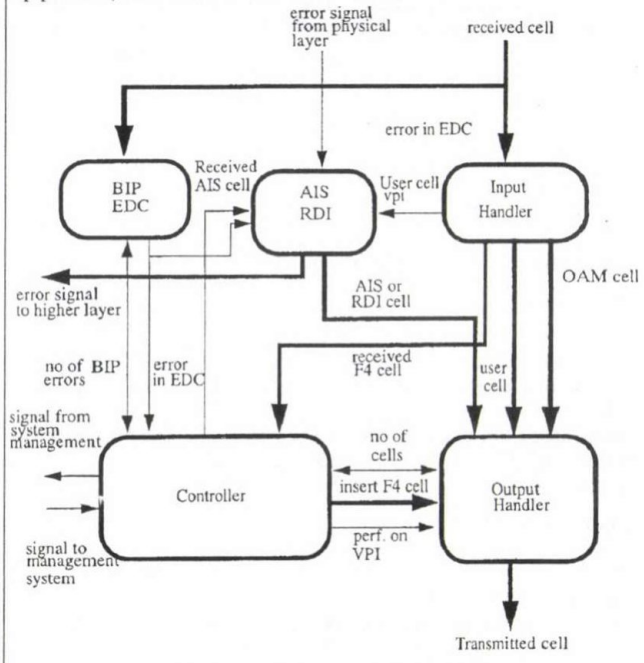


Fig. 9. Architecture of the F4 block

The improved model takes into consideration only the incoming cells from the upstream ATM network, thus having only half of the complexity of the previous module. For OAM cells that have to be sent back, the F4 block requests the management system to provide a new VPI which causes the ATM switch to reflect the cell.

The improved F4 model is composed of five independent parts (Fig. 9):

- A controller;
- An Input Handler;
- An Output Handler;
- A block generating AIS and RDI cells;
- A block calculating Bit Interleaved Parity (BIP) and EDC.

After analysing the incoming cells, the Input Handler checks whether it is a user cell or an OAM cell. The user cells are passed on to the Output Handler after the VPI information is extracted and sent to the block generating AIS and RDI cells. The OAM cells are checked if they are aimed for this node or not. If it is the case, the cells are sent to the controller, otherwise to the Output Handler. The EDC of all OAM cells is calculated by the calculating BIP and EDC block. The functionality in the Input Handler has been simplified by removing the most time consuming parts in the synthesisable F4 model. After removing the critical part the resulting Input Handler has the complexity of the inp hndlr 2 subblock in the synthesisable F4 model. With this improvement the inp hndlr 2 meets the performance requirement (168 ns).

The Output Handler reads the incoming user cells from the Input Handler, counts them and places them in a user cell queue. After a certain amount of incoming user cells have

been processed, an OAM cell is generated if the performance monitoring has been enabled. This OAM cell is also placed in the user cell buffer. If an OAM cell arrives from the controller it is placed in the OAM cell buffer. The Output Handler reads cells from the user and OAM cell queues (the OAM cell buffer has higher priority then the user cell buffer) and sends them to the ATM switch. Replacing the table search code (implemented in the synthesisable F4 model) with FIFO queues eliminates the search delays. With this improvement the Output Handler meets the performance requirement (168 ns).

The OAM cells received by the controller from the Input Handler can be fault management (AIS, RDI or loopback), performance management or activation/deactivation cells. The AIS cells are passed on to the block generating AIS and RDI cells. The RDI and the loopback cells are deleted, or updated and sent to the Output Handler. The information in the activation/deactivation cells instructs the F4 block to perform performance monitoring. If the activation has been accepted/rejected by the F4 block, the proper activation cell bearing this information will be passed on to the Output Handler.

6.1. Performance Analysis

The Output Handler in the F4 block contains a user cell queue and an OAM cell queue (Fig. 10). The OAM cell queue has a higher priority than the user cell queue. The user cell queue receives cells coming from the Input Handler and Controller (denoted as generator G1). The OAM cell queue receives cells coming from the generator G1 and from the AIS/RDI block (denoted as generator G2).

If all VPs are deteriorated, 256 AIS cells are sent in a bunch every second to the OAM cell queue and 0 user cells are sent to the user cell queue (due to the fact that all the lines are deteriorated). In this case the OAM cell queue always contains one cell and the user cell queue contains no cells.

If 255 VPs are deteriorated then 255 AIS cells are sent from the generator G2 in a bunch every second to the OAM cell queue.

- One special case is that all cells transmitted from the generator G1 are user cells (0 % OAM cells). During the time that the generator G2 sends AIS cells, a maximum number of 255 cells could arrive from the VP not deteriorated to the user cell queue. In this case the OAM cell queue and the user cell queue will contain 1 cell and a maximum of 255 cells respectively (the user cell queue has to wait until the 255 cells in the OAM cell queue has passed). The user cell queue will be reduced to 1 again due to idle cells. In this case, the OAM cell queue and the user cell queue have to be dimensioned for 1 and 255 cells respectively.
- Another special case is that all cells transmitted from the generator G1 are OAM cells (0 % user cells). During the time that the generator G2 send AIS cells, a maximum of 255 OAM cells are sent from the generator G1 to the OAM cell queue. In this case the OAM cell queue will contain a maximum of 255 cells. The OAM queue will be reduced to 1 again due to idle cells. In this case the OAM cell queue should be dimensioned for 255 cells.

More generally, if N VPs are deteriorated then N AIS cells are sent from the generator G2 in a bunch every second to the OAM cell queue. We should assume also that x % user cells and y % OAM cells are transmitted by the generator G1. During the time that the generator G2 sends AIS cells, a maximum of N cells coming from the generator G1 are spread between the user cell queue and the OAM cell queue with the values of $N * x/100$ and $N * y/100$ respectively. The user cell queue has to be dimensioned to N cells if x varies in the range 0 % to 100 %. The OAM cell queue has to be dimensioned to $N * \max(y)/100$.

We have used uC++, an extension of the object oriented language C++ supporting concurrency [10] to develop a model of the F4 block. uC++ showed reasonable performance in our applications, and required relatively small amounts of memory. Our experience shows that uC++ [11] is a very good language for behavioural simulations of large ATM systems. The simulation was done at cell level.

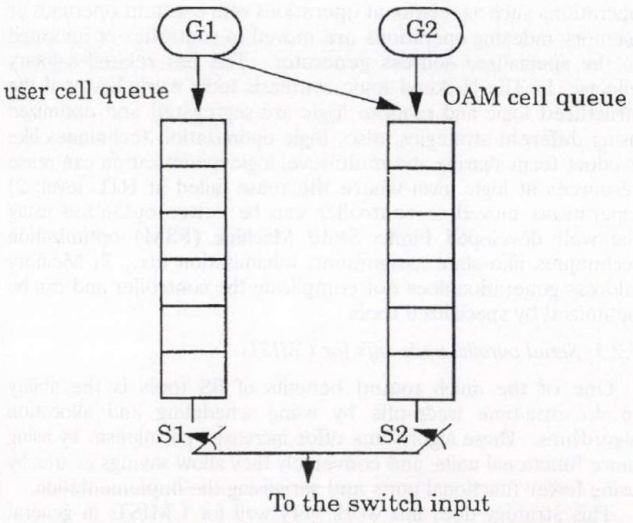


Fig. 10. Queueing model in the Output Handler in the F4 block

In the case where 10 % of OAM cells and 90 % of user cells are injected into the F4 block, theoretically the OAM cell buffer should contain a maximum of 25 cells and the user cell buffer 255 cells (in this case we assumed that 255 VPs were deteriorated). Fig. 11 shows that in ten simulation runs we obtained a user cell buffer of 255 cells and an OAM cell buffer of 25 cells (mean value). The result varies from run to run for the OAM cell buffer, depending on the random generator that generates user cells, OAM cells and idle cells.

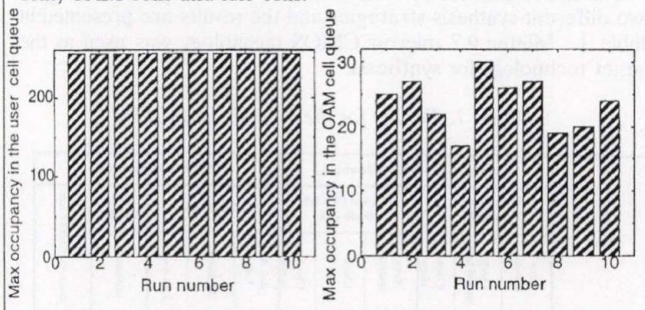


Fig. 11. Simulation results

7. CONCLUSION

We have demonstrated a successful application of Hardware Description Language based design methodology when designing a complex telecommunication hardware. Behavioural synthesis tools had to be adapted for telecommunication applications and the adaptations proved to be efficient in improving the quality of the designs. The improved tools have been also very useful in identifying the limits of the model. This was of crucial importance when improving the model for higher data transfer rates. It has also been demonstrated that with a straightforward serial implementation of the existing synthesisable F4 model, the speed of 2.5 Gbit/s couldn't be reached. The proposed architectural improvements that partitions the F4 algorithm into parts that can be pipelined removes the bottlenecks. The new architecture only takes into consideration incoming cells from the physically upstream part in the ATM network. Simulation results showed that the functionality of our architecture is correct and enables the designer to engineer the size of the buffers of the Output Handler in order to minimize the cell loss in the F4 block. Work on further improvements in the behavioural synthesis for communication circuits is under way and this will soon allow efficient synthesis of the pipelined 2.5 Gbit/s model. We also continue to work on development of algorithms for fast calculation of the EDC and BIP parts of the F4 block for 2.5 Gbit/s.

REFERENCES

[1] M. De Prycker: "*Asynchronous Transfer Mode solutions for broadband ISDN*", Series in Computer Communications and Betworking, Ellis Horwood.

[2] V. P. Kumar, J. G. Kneuer, D. Pal, B. Brunner, Phoenix: a building block for fault tolerant broadband packet switches, *Proc. of the IEEE Global Telecommunications Conference,* December 1991, Phoenix.

[3] F. Hansen: "2.5 Gbit/s ATM switch element", DTH, Electromagnetics Institute, 1993, Denmark.

[4] ITU-T Telecommunication Standardization Sector of ITU Recommendation I.150, I.211, I.311, I.321, I.327, I.361, I.362, I.363, I.413, I.432, I.610.

[5] Bellcore: Generic Requirements for Operations of Broadband Switching Systems, TA-NWT-001248 issue 2, October 1993.

[6] A. Doboli, J. Hallberg and P. Eles: A Simulation Model for the Operation and Maintenance Functionality in ATM Switches, Internal Report, Dept. of Computer Science, University of Linköping, Sweden.

[7] B. Svantesson: A Test Bench for the F4/FS block of the ATM Protocol, Master Thesis, Royal Technical University Stockholm, 1994.

[8] A. Hemani, B. Svantesson, P. Ellervee, A. Postula, J. Öberg, A. Jantsch and H. Tenhunen: "High-Level Synthesis of Control and Memory Intensive Communication Systems", *1995 ASIC Conference and Exhibit*, Austin, Texas, USA, Sept. 1995, pp. 185-191.

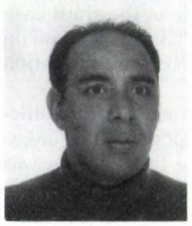
[9] G. Werner: "Memory and data-path mapping for image and video applications", Chapter 7, Application driven architecture synthesis, by Francky Cathoor, Lars Svensson. Kluwer Academic Publishers, 1993, ISBN: 0-7923-9355-4.

[10] P. A. Buhr, G. D. Ditchfield, R. A. Stroobosscher, B. M. Younger and C. R. Zarnke: "uC++: Concurrency in the Object Oriented Language C+++", *Software Practice and Experience*, Vol. 22(2) (February 1992), pp. 137-172.

[11] T. Lazraq, F. Östman, J. Öberg and H. Tenhunen: "ATM Switching System Performance Analysis via Modelling and Simulation", *Proc. of the 36th SIMS '94 Simulation Conference*, Stockholm, Sweden, pp. 326-332.

[12] A. Hemani: High-level Synthesis of Synchronous Digital Systems using Self-Organization Based Scheduling and Binding Algorithms, Ph.D. Thesis, Dec. 1992, Royal Institute Of Technology, Stockholm, Sweden.

Ahmed Hemani received the B.E. degree in Electronics Engineering from University of Indore, India in 1982 and the Ph.D. degree in Applied Electronics from the Royal Institute of Technology, Stockholm, Sweden in 1992. He is Associate Professor in the Department of Electronics, Royal Institute of Technology, where he leads research in the area of Design Methods for Electronic System Design. His research interests include VLSI Design automation, high-level synthesis for telecommunication, optimization algorithms, hardware description languages and neural networks.



Tawfik Lazraq received the M.Sc. degree in electrical engineering from Ecole Polytechnique Federale de Lausanne, Switzerland, in 1987. In 1993 he received the Techn. Lic. degree from the Royal Institute of Technology (KTH), Stockholm, Sweden. From 1987 to 1989 he worked at Brown Boveri in Baden, and Ericsson Telecom in Stockholm on electronics. From 1989 to 1993 he worked at the Department for Applied Electronics on road traffic information and communication networks. In 1993 he joined the Electronic System Design laboratory at KTH. His research interest is mainly in high performance ATM switching. He spent nearly one year at CERN, where he carried out research on evaluating ATM for event-building.

Adam Postula received MSEE and Ph.D. degrees from universities in Poland, in 1974 and in 1981, respectively, where he researched in the field of signal processing and digital system design. From 1982 he worked as a design and research engineer in ABB Automation in Sweden with multiprocessor systems and CAD tools for electronic system design. From 1987 he led the logic synthesis project within the Swedish National Microelectronic Project NMP-CAD and later continued research on high-level synthesis from VHDL at the Royal Institute of Technology and Swedish Institute of Microelectronics. Hi moved to Australia in 1992 and is a senior lecturer at the University of Queensland in Brisbane. his research is focused mainly on system design and application of high-level synthesis to telecommunication systems. Another area of his research is custom computing machines based on programmable gate arrays. He has served on the program committee of several international conferences. In 1991 he organized Euro-VHDL conference in Stockholm. In 1993 he co-oreganized and chaired the Asian Pacific Conference on Hardware Description Languages in Brisbane. He is a member of IEEE and IREE.

Bengt Svantesson is currently completing his M.Sc. at the Royal Institute of Technology in Stockholm, Sweden. In 1994, he joined the Electronic System Design laboratory as a research assistent. His research interest is high-level synthesis for telecommunications.

CELL STREAMS UNDER THE MICROSCOPE

ANALYTICAL AND SIMULATION RESULTS ON GCRA-BASED TRAFFIC SHAPING*

L. JEREB, T. V. DO and G. WOLFNER

DEPARTMENT OF TELECOMMUNICATIONS
TECHNICAL UNIVERSITY OF BUDAPEST
SZTOCZEK 2, 1111 BUDAPEST, HUNGARY
PHONE: +36 1 463-2096; FAX: +36 1 463-3266; E-MAIL: {JEREB,DO,WOLFNER}@HIT.BME.HU

This paper deals with the application of a GCRA-based traffic shaping algorithm used in the source ATM end-point. Two scenarios are considered: for the first scenario, where the traffic on all the VCs originating in one user is collectively shaped, a discrete-time analysis model is developed. In the second scenario the traffic is shaped on groups of VCs having common features (e.g. same destination node), for which a traffic shaping architecture is constructed and an algorithm is proposed for scheduling the cell flows to the ATM link. The scheduling algorithm cooperates with the shapers in the traffic shaping architecture in order to ensure the conformity of the individual cell streams with the conformance test at the UNI point. Applications of the traffic shaping architecture are demonstrated with some analytical and simulation results in the topology of a potential Hungarian ATM backbone network.

1. INTRODUCTION

Traffic shaping is a mechanism which attains desired characteristics on a cell stream emitted to ATM networks. According to ATM Forum User Network Interface Specification Version 3.0 two possible alternatives to place traffic shaping are proposed [3]: in a private ATM switch (UNI/NNI point), or in the source ATM end-point. For the former case, an example of traffic shaping used at the UNI/NNI can be seen in [4], where it is used to decrease the Cell Delay Variation of connection-oriented traffic. In this paper we consider the use of traffic shaping in the ATM end-point of a connectionless source which can logically represent either an Interworking Unit (IWU) used to connect LAN/MAN over ATM, or Customer Premises Equipment submitting connectionless traffic. Traffic shaping is applied in the ATM end-point to reduce the burstiness of connectionless traffic.

Simulation results of Ajmone Marsan et al. have shown the benefits of traffic shaping and scheduling in the ATM end-points for the network: cell and message loss probabilities can be drastically reduced even to zero at the expense of a significant increment in the end-to-end delay jitter of messages and cells [1], [2].

This contribution summarizes some new results beyond the results presented in [1], [2]. On the one hand, for a Generic Cell Rate Algorithm-based (GCRA) shaper used to shape collectively connectionless traffic streams, a discrete-time analysis model is derived, in which a Markov chain is embedded at message arrivals. On the other hand, for the considered GCRA-based shaper architecture, new results concerning the scheduling of multiple traffic streams after shaping and the QoS parameters versus the peak-to-mean ratio are presented. Various applications of the shaper architecture are demonstrated in the topology of a potential Hungarian ATM backbone network.

Since the nodes of the Hungarian ATM network are supposed at the places of the secondary exchanges of the current Hungarian optical network, the ATM links of the Hungarian topology is expected to use the Hungarian digital backbone network. We hope that this network level simulation results contribute to the study of B-ISDN introduction in Hungary.

The rest of the paper is organized as follows. In Section 2 the GCRA-based traffic shaping algorithm and the shaper architecture are described. Section 3 presents a model for the performance analysis of the GCRA-based algorithm. Simulation results are demonstrated in Section 4. Section 5 outlines some possible applications of the mathematical analysis. Finally, Section 6 concludes the paper.

2. GCRA-BASED TRAFFIC SHAPING

2.1. Traffic shaping algorithm

In this paper the GCRA-based algorithm proposed by Ajmone et al. in [1] is considered. The rationale for this solution is the fact that at present only the peak/sustainable cell rate is recommended by ITU-T as the parameter for the traffic contract and GCRA is mentioned as a possible algorithm for monitoring the connection peak/sustainable cell rate [7]. Therefore, by using the GCRA-based algorithm the source can ensure that his traffic is conforming to the GCRA conformance test performed at the UNI point.

The algorithm has two key parameters: T denotes the time increment of the algorithm, while τ is related to the allowed burs[1].

In the shaper each arriving cell is tagged with the allowed transmission time Tt_i calculated according to the algorithm based on the cell arrival time Ta_i. The allowed transmission time of a generic cell C_i is the time at which C_i is eligible for transmission and it is calculated according to the following algorithm given the theoretical arrival time TAT_i and the actual arrival time Ta_i of cell C_i ([1]):

- if $Ta_i \geq TAT_i$
  then $Tt_i = Ta_i, \quad TAT_{i+1} = Ta_i + T$
- if $TAT_i - \tau \leq Ta_i < TAT_i$
  then $Tt_i = Ta_i, \quad TAT_{i+1} = TAT_i + T$
- if $Ta_i < TAT_i - \tau$
  then $Tt_i = TAT_i - \tau, \quad TAT_{i+1} = TAT_i + T$

Note that if $\tau = 0$ the algorithm is equivalent with a simple spacer.

2.2. Traffic shaping architecture

In practice there may be many VCs submitting connectionless traffic at a CL source (e.g. a CL-server, IWU) and VCs may share a common buffer. From a point of view of traffic management some performance advantages may be gained if VCs are grouped on the basis of common characteristics (e.g. group of VCs belonging to the same VP). In this paper, three types of grouping VCs are compared. In the first one each VC is individually shaped (referred as VCG-VC Grouping). In the second one all the VCs originating from one user are collectively shaped (referred as SUG-Source User Grouping), while in the third case an intermediate solution is studied, where the VCs are grouped on the basis of the destination node (referred as DNG – Destination Node Grouping).

These situations may be modeled by using a shaping architecture depicted in Fig. 1. A more detailed description of the investigated architecture can be found in [2], where the solution was originally proposed; here only the main features are recalled.

* This work was performed with the support of the European Community through the COPERNICUS '94 contract n. 1463.

[1] In Generic Cell Rate Algorithm $1/T$ is the peak cell rate, τ denotes CDV (Cell Delay Variation) tolerance.

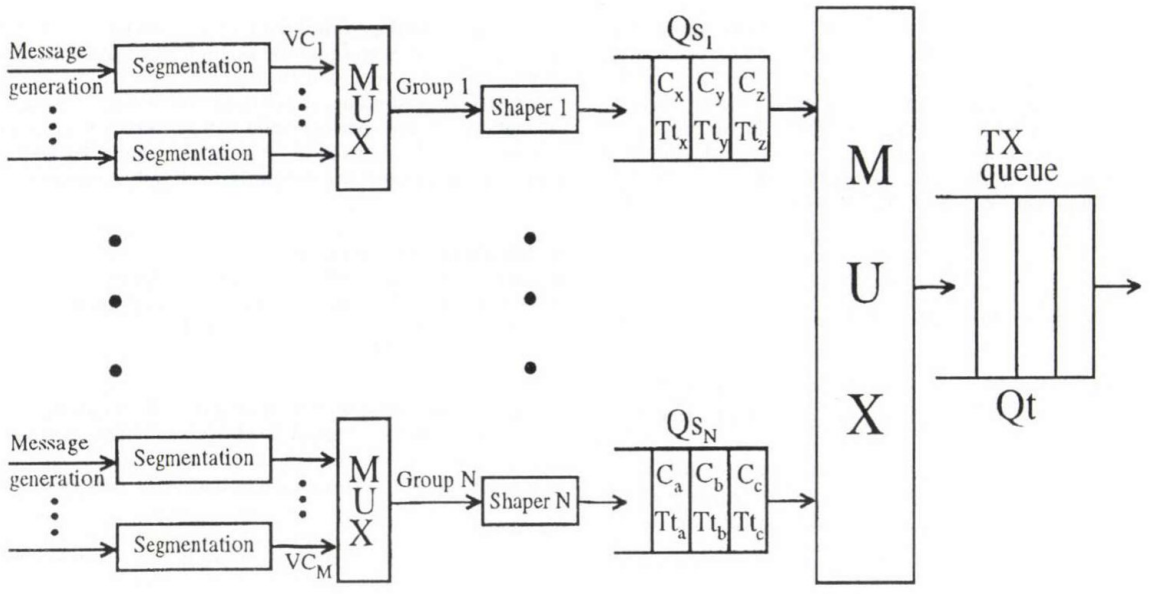


Fig. 1. Architecture of the GCRA-based traffic shaping device

Message generations correspond to the invocation of AAL service primitives from higher layer protocols. In the device messages are then segmented into cells, and batches of cells corresponding to whole messages are forwarded to the shapers, that implement the GCRA-based shaping algorithm on the traffic referring to a given *group* of VCs.

A cell after the ith shaper is waiting in queue Qs_i. If its timestamp expires, the cell can be put into the transmission buffer by the second MUX stage. Since there will be cells simultaneously contending for the transmission buffer and the link, the algorithm for scheduling cells of queues (Qs_i) is fairly critical. We will return to this question in Subsection 4.3.

3. MATHEMATICAL ANALYSIS OF THE SHAPING ALGORITHM

In this section the performance of a single shaper[2] of the proposed shaping device is analyzed in discrete-time domain. The analysis is based on the method "unfinished work at arrivals" proposed by Tran-Gia and Ahmadi in [9].

In the context of this paper the time is discretized into slots of cell duration, which corresponds with ATM environments. The time unit has exactly a length of cell duration. The parameters of the shaper and message interarrival times are scaled according to the time unit. The parameters of the algorithm T and τ are measured as integer numbers of the time unit.

3.1. Basic notations

Let us introduce the following notations:

- B: the size of the buffer at the source;
- $Ta^{(n)}$: the arrival epoch of message n, which denotes the arrival time of the message at the AAL;
- $A_n = Ta^{(n)} - Ta^{(n-1)}$: the interarrival time between the $(n-1)$th and nth message;
- $a_n(k) = Pr\{A_n = k\}$: the distribution of the discrete-valued random variable A_n. Without the loss of generality suppose that the interarrival time A_n between messages is identically distributed ($a_n(k) = a(k)\ \forall n$);
- X_n: the batch size (number of cells) of message n;
- $\forall n$: $x_n(m) = Pr\{X_n = m\} = x(m)$ the distribution of the discrete-valued random variable X_n. Similarly to the interarrival time, the message length is also supposed to be identically distributed. In the paper only such cases are considered, where the length of messages are upper bounded by the buffer size ($X_n \leq B$).

[2] corresponding to the SUG case

- $Z(t)$: the difference between the value of TAT of the shaper at t and the current time t:

$$Z(t) = max(TAT(t) - t, 0).$$

Note $TAT(t)$ is the value of TAT of the BOM (Begin Of Message) cell of the next message.

- Z_n^-: the value of $Z(t)$ in slots just before the arrival instant of message n:

$$Z_n^- = max(TAT^{(BOM_n)} - Ta^{(n)}, 0),$$

where $TAT^{(BOM_n)}$ is the theoretical arrival time of the BOM cell of message n;

- Z_n^+: the value of $Z(t)$ in slots just after the arrival instant of message n:

$$Z_n^+ = max(TAT^{(BOM_{n+1})} - Ta^{(n)}, 0),$$

where $TAT^{(BOM_{n+1})}$ is the value of TAT at $Ta^{(n)}$ after performing the shaping algorithm on the EOM (End Of Message) cell of message n, i.e. $TAT^{(BOM_{n+1})}$ is the theoretical arrival time of the BOM cell of the message following message n.

- $C(t)$: the number of cells in the buffer at time t;
- $L(t)$: the number of cells in the buffer, whose allowed transmission time is less than or equals t;
- $W(t)$ the number of cells in the buffer, whose allowed transmission time is greater than t, thus $W(t) = C(t) - L(t)$.

If there are some cells waiting in the buffer at time t, the number of slots to the TAT of the EOM cell of the last message accepted before t can be expressed as $Z(t) - T$ and the number of slots to the allowed transmission time is $Z(t) - T - \tau$. Provided that there are k cells waiting in the buffer whose allowed transmission time has not expired, the number of slots to the allowed transmission time of the first cell of them at slot t can be derived as $Z(t) - k \cdot T - \tau$ and the following inequality holds:

$$k \cdot T + \tau \leq Z(t) < (k+1) \cdot T + \tau, \qquad (1)$$

and therefore $W(t)$ can be expressed as:

$$W(t) = \lfloor \frac{max(Z(t) - \tau, 0)}{T} \rfloor, \qquad (2)$$

where $\lfloor r \rfloor$ denotes the largest integer number which is not greater than r.

Eq. (1) implies that $Z(t)$ is upper bounded by $(B+1) \cdot T + \tau$.

In case $\tau < T$, there are no cells waiting with expiring time stamp. That means the number of cells in the buffer at time t can be directly obtained from $Z(t)$.

On the other hand, in case $\tau \geq T$ there may be more than one cell waiting in the buffer with expiring time stamp because of the nature of the shaping algorithm. In this case the state of the shaper can be described by the two dimensional random variable $S(t) = (Z(t), C(t))$ or $S(t) = (Z(t), L(t))$. In the paper we only present the main steps for the $\tau < T$ case, but the detailed analysis for both cases is also described in [5], [6].

In the analysis the key task is the determination of the evolution of $Z(t)$ at message arrivals:

$$Z_n^- \Longrightarrow Z_n^+ \Longrightarrow Z_{n+1}^-.$$

Using Eq. (1) the relation between Z_n^+ and Z_n^- can be written as follows:

$$Z_n^+ = \begin{cases} Z_n^- + X_n \cdot T & Z_n^- + X_n \cdot T < (B+1) \cdot T + \tau, \\ Z_n^- & Z_n^- + X_n \cdot T \geq (B+1) \cdot T + \tau. \end{cases} \tag{3}$$

According to the above definitions, $Z(t)$ decreases by one in each slot between the arrival instants of two arbitrary messages until it reaches zero. Consequently, the following relation can be easily set up between Z_{n+1}^- and Z_n^+:

$$Z_{n+1}^- = \max(Z_n^+ - A_{n+1}, 0). \tag{4}$$

Using Eqs. (3) and (4) one can derive an iterative algorithm for calculating the distributions $Pr\{Z_n^- = k\}$ and $Pr\{Z_n^+ = l\}$, and therefore the equilibrium state distributions can be obtained as:

$$z^-(k) = \lim_{n \to \infty} Pr\{Z_n^- = k\}, \text{ and } z^+(l) = \lim_{n \to \infty} Pr\{Z_n^+ = l\}. \tag{5}$$

3.1.1. Performance measures

The performance measures can be determined by using the equilibrium distribution $z^-(k)$.

Taking into account the condition in (3) the message blocking probability P_{BM} can be expressed as follows:

$$P_{BM} = \sum_{k=0}^{(B+1)\cdot T+\tau-1} z^-(k) \cdot \Pr\{\text{message blocked } | Z^- = k\}$$

$$= \sum_{k=0}^{(B+1)\cdot T+\tau-1} z^-(k) \sum_{j=b(k)+1}^{B} x(j), \tag{6}$$

where $b(k) = \lfloor \dfrac{(B+1)\cdot T + \tau - 1 - k}{T} \rfloor$.

The delay of the ith cell of a message finding $Z(t) = k$ at arrival can be given by: $cd(i,k) = \max(k + (i-1)\cdot T - \tau, 0)$.

The mean cell delay can be calculated as:

$$MCDT = \frac{\sum_{k=0}^{(B+1)\cdot T+\tau-1} z^-(k) \sum_{j=1}^{b(k)} x(j) \cdot \sum_{i=1}^{j} cd(i,k)}{\sum_{k=0}^{(B+1)\cdot T+\tau-1} z^-(k) \sum_{j=1}^{b(k)} x(j) \cdot j}, \tag{7}$$

and the mean message delay defined as the mean delay of the EOM cells of messages can be written as follows:

$$MMDT = \frac{\sum_{k=0}^{(B+1)\cdot T+\tau-1} z^-(k) \sum_{j=1}^{b(k)} x(j) \cdot cd(j,k)}{\sum_{k=0}^{(B+1)\cdot T+\tau-1} z^-(k) \sum_{j=1}^{b(k)} x(j)}. \tag{8}$$

4. SIMULATION RESULTS

The simulation investigations were performed with the software tool CLASS (ConnectionLess ATM Services Simulator) [1]. CLASS was developed at Politecnico di Torino, and in 1994 its shaping capabilities were jointly upgraded by Politecnico di Torino and Technical University of Budapest.

CLASS is a slotted synchronous simulator, devoted to the estimation of the performance parameters of connectionless services in ATM networks. Due to this goal CLASS computes performance parameters both at the cell and at the message level such as

- cell and message loss probabilities;
- cell and message delay jitters;
- number of useless cells brought to destination.

The detailed description of the simulation method as well as the computed parameters can be found in [1].

4.1. Network topology and workload

Fig. 2 shows the investigated network which can be considered as the topology of a possible Hungarian ATM network. The nodes of the network are supposed in the places of the secondary exchanges of the current digital backbone network, while the links are expected to use the existing optical fibers (see [8] for details on the topology of the Hungarian backbone network). The network comprises 9 nodes, 21 users, and 14 bidirectional links of 150 Mbit/s capacity.

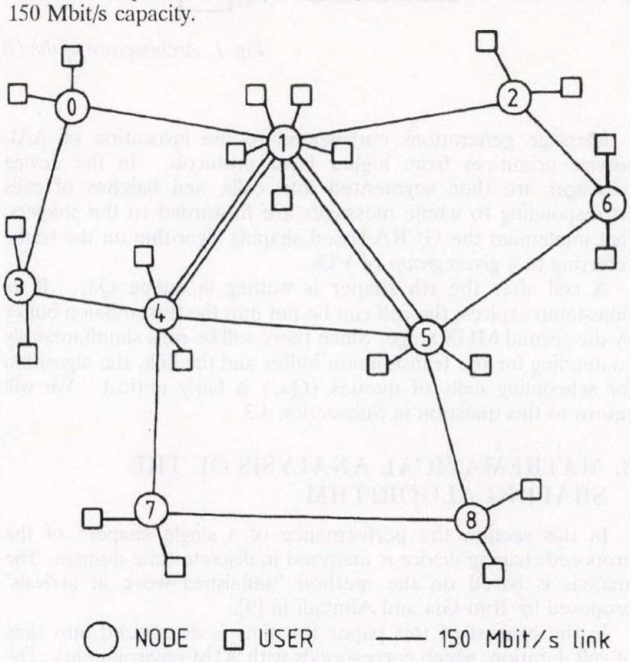


Fig. 2. Topology of the investigated Hungarian network

Table 1. Traffic matrix for the network; the relations are expressed in thousandth of the global generated traffic

Node	0	1	2	3	4	5	6	7	8
0	0	45	5	5	5	5	5	5	5
1	45	0	45	45	45	45	45	45	45
2	5	45	0	5	5	5	5	5	5
3	5	45	5	0	5	5	5	5	5
4	5	45	5	5	0	5	5	5	5
5	5	45	5	5	5	0	5	5	5
6	5	45	5	5	5	5	0	5	5
7	5	45	5	5	5	5	5	0	5
8	5	45	5	5	5	5	5	5	0
Total	80	360	80	80	80	80	80	80	80

The traffic among the nodes is uniformly distributed with the exception of the outstanding role of the center node (see Table 1). Within a node, the outgoing and the incoming traffic is also uniformly distributed among the users. Consequently, from

each of the five users in the central node 16 permanent VCs are established, while from the other eight nodes 19 permanent VCs are directed to the five central users and to the 14 others. The topology and the traffic distribution show monocentric and symmetric features, thus there are only two kinds of users and two kinds of VCs with similar characteristics in the network: the high-load VCs convey information between the users of node 1 and other users, while the low-load VCs carry traffic among the users of nodes different from node 1.

4.2. Impact of shaping on performance parameters

In this section we present simulation results for the network with the GCRA-based shaping algorithm in case of the three grouping solutions (VCG, SUG, DNG). The network parameters were chosen as follows:

- the global traffic load equals 600 Mbit/s;
- the message arrival process is Poisson;
- the message length distribution is a truncated geometric, with average 20 cells, and truncation value 200 cells;
- the size of the user buffers was chosen large enough to guarantee that no message loss occurs at the user;
- the size of all buffers inside the network was chosen to be 100 cells.

The bandwidth allocated to each group of VCs (BA_n for the n-th group) defines the normalized peak cell rate ($1/T_n$ for the n-th group). The strictness of shaping is controlled by the "bandwidth allocation factor" β, that is taken to be constant for all the groups in the network.

If B_{ni} is the average bandwidth of the i-th VC in the n-th group, $BA_n = \beta \sum_i B_{ni}$. From BA_n, T_n is obtained as

$$T_n = \max\left(1, \left\lfloor \frac{C}{BA_n} \right\rfloor\right), \tag{9}$$

where C is the speed of the link connecting the source user to the first node of the ATM network. C always equals 150 Mbit/s in these simulation experiments.

In [2] a vast number of results was presented concerning a potential Italian and this Hungarian network topology. It was learnt from the numerical results that the strictness of shaping and the way in which VCs are grouped are extremely important; the smaller the value of β and the larger the number of groups are, the better the performance improvement is in terms of cell loss probability and cell delay jitter inside the network, but the larger the price is to be paid in the user-to-user delay jitter. This behaviour is illustrated in Fig. 3 where the curves of the average user-to-user cell delay jitter and average cell delay jitter inside the network are jointly plotted with $\tau = 100$. The difference between the user-to-user cell delay jitter and the cell delay jitter inside the network is exactly the time spent by a cell in the source user buffer; i.e. inside the shaping device. As it was expected, cell and message loss probabilities inside the network behave similarly to the delay jitters inside the network, while the user-to-user message delay jitters are similar to the corresponding user-to-user cell delay jitters [2].

Now, we focus our attention on the impact of β. Since the maximum value of the normalized peak cell rate is 1, for each group a limit value $\beta_l = \frac{C}{\sum_i B_{ni}}$ of β exists. Thus, the effective controlling range of β of a specific group spreads from 1 to β_l, and obviously, there is no any change in the performance of the shaper in case $\beta > \beta_l$ compared to the case $\beta = \beta_l$. Taking into consideration the non-linearity in Expression (9), we can introduce a modified parameter β_{en} characterizing more appropriately the intensity of the traffic leaving the traffic shaping architecture:

$$\beta_{en} = \frac{C}{T_n \sum_i B_{ni}}, \tag{10}$$

where C/T_n is the peak cell rate of the nth group, and from the definition $1 \leq \beta_{en} \leq \beta_l$.

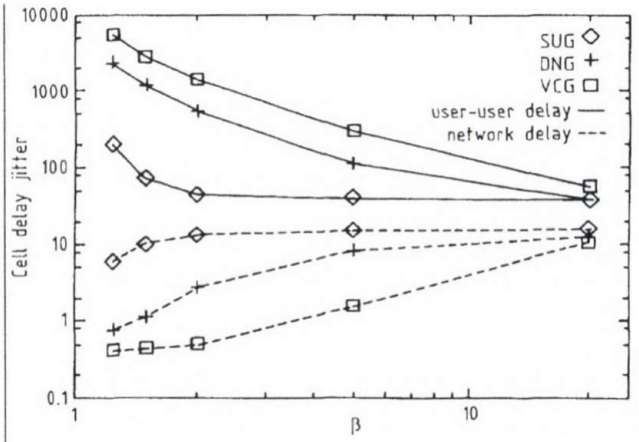


Fig. 3. Average cell delay jitter inside the network together with average user-to-user cell delay jitter versus the bandwidth allocation factor β, with $\tau = 100$

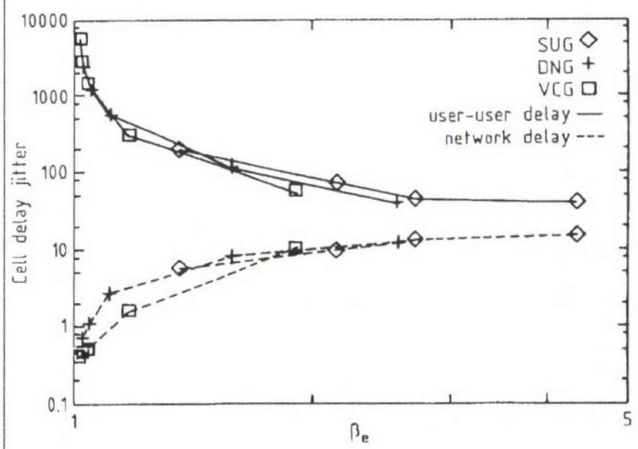


Fig. 4. Average cell delay jitter inside the network together with average user-to-user cell delay jitter versus the peak-to-mean ratio β_e, with $\tau = 100$

Furthermore, if we extend β_{en} in order to characterize or measure the strictness of shaping for a given user or for the whole network, we have to take into consideration that the value of β_m for user m can vary only between 1 and the peak/mean cell rate ratio (C/B_m), where $B_m = \sum_{n=1}^{k_m} B_n^{(m)}$ denotes the mean cell rate of the mth user comprising k_m groups with mean cell rate $B_n^{(m)} = \sum_i B_{ni}$. Consequently, for the mth user the new measure of shaping is constructed as follows:

$$\beta_{em} = \sum_{n=1}^{k_m} \frac{C}{T_n \cdot B_m} \cdot \frac{1 - B_m/C}{k_m - B_m/C} + \frac{k_m - 1}{k_m - B_m/C}, \tag{11}$$

which meets the requirement $1 \leq \beta_{em} \leq C/B_m$, and for $k_m = 1$ Eq. (10) still holds.

Finally, a new measure called peak-to-mean ratio can be introduced for the whole network as the weighted average of β_{em}:

$$\beta_e = \frac{\sum_{m=1}^{M} B_m \beta_{em}}{\sum_{m=1}^{M} B_m}, \tag{12}$$

where M is the number of users in the network.

Using β_e instead of β we can plot again the delay characteristics presented in Fig. 3. The curves in Fig. 4 show that the differences in the cell delay jitters originated mainly from different controlling range (β_e) of the different groupings.

4.3. Impact of scheduling on performance parameters

In the simulation studies published in [2] the scheduling

algorithm worked as follows: the contending cells of the shaper queues were put into the transmission buffer at the same time, immediately after their time stamp expired. The shaper queues were visited in a deterministic order.

The consequence of this scheduling scheme is that cells may be found non-compliant by the GCRA conformance test at the UNI point [1]. This effect can be explained by the fact that if backlog occurs in the transmission buffer and the total input intensity goes under the transmission link capacity, the cells even in case of very tight shaping could be placed closer to each other in the transmission buffer than it would be allowed by the shaping algorithm.

Aiming at the elimination of this disadvantage, we implemented a new scheduling algorithm which cooperates with the shapers if there are cells contending for the transmission buffer (and for the link). At each slot, when there are more than one cells waiting for the transmission link, only one cell with the smallest expiring time stamp is taken and transmitted. On the other hand for all other queues, where the head-end cell remained in the shaper with expiring time stamp, the TAT and the time stamp of all cells, except the head-end one, are increased by the time unit.

In order to compare the two scheduling schemes, we define the scheduling delay as the time difference between the time epoch when the cell leaves the transmission buffer, and the theoretical transmission time Tt_i computed upon arrival by the shaping algorithm with the original scheduler. Obviously, in case of the original scheduler, this is the delay jitter in the transmission buffer. However, in the modified scheduler its value includes not only the delay suffered in the head-end position, but also all additional waiting time introduced in order to ensure the conformity of the cell stream. Later on however, we call it simply scheduling delay. In Fig. 5 the curves for the original and modified schedulers show that in the modified case the average scheduling delay is slightly higher than in the original scheme.

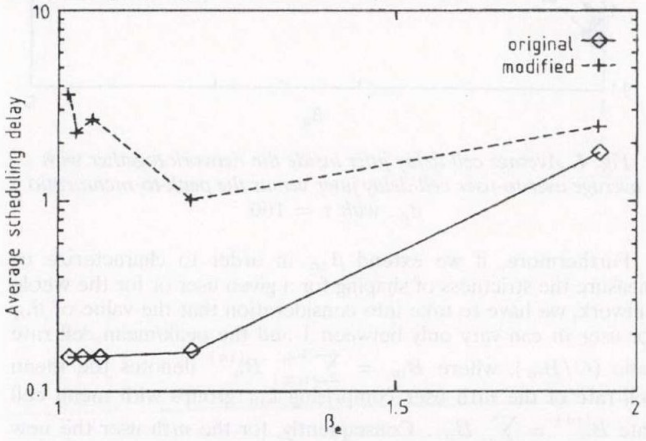


Fig. 5. Average scheduling delay versus the peak-to-mean ratio β_e, with $\tau = 0$

In Fig. 5 the modified scheduler shows interesting peaks for low β_e values. It can be explained by the fact that in the modified scheduling scheme a phasing effect can be observed. This phasing effect can be explained as follows: after the collision of cells of different groups at the second MUX stage, their shapers are synchronized until all cells are transmitted if the spacing parameter T_i of the groups have a greatest common divisor larger than 1.

This interpretation is supported by simulation runs, whose input shaping parameters are reported in Table 2. In Series A, β was chosen to be 1.25, 1.5, 2, 5 and 20, and in column GCD the greatest common divisors are provided. In Series B the shaping parameters are tuned and $T^{(1)}$ of low-load VCs is the multiple of $T^{(2)}$ of high-load VCs. In order to study the consequence of this modification in Fig. 6 the reduced scheduling delay of the original, modified, and corrected schedulers are plotted, where the reduced delay is defined as the time difference between the time epoch when the cell leaves the transmission buffer, and the theoretical transmission time Tt_i computed upon arrival by the actual shaping algorithm. The results for Series B are depicted in the intermediate curve of Fig. 6, that confirms the above phasing explanation.

Table 2. The shaping parameters T of two kinds of VCs in the Hungarian network topology for two run series

	Series A			Series B	
β	$T^{(1)}$	$T^{(2)}$	GCD	$T^{(1)}$	$T^{(2)}$
1.25	132	36	12	132	33
1.5	110	30	10	108	27
2	83	23	1	80	20
5	33	9	3	32	8
20	8	2	2	8	2

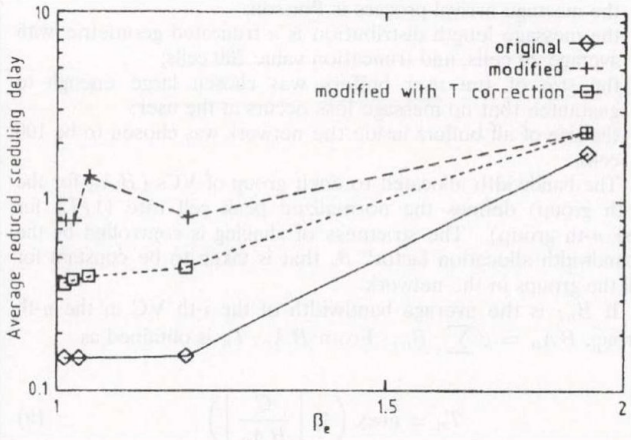


Fig. 6. Average reduced scheduling delay versus the peak-to-mean ratio β_e, with $\tau = 0$

5. APPLICATION OF MATHEMATICAL RESULTS IN DIMENSIONING SHAPERS

In Fig. 7 the analytical results on the cell delay jitter in the shaper versus the spacing parameter T are plotted, for different mean load of two kinds of VCs in the Hungarian topology.

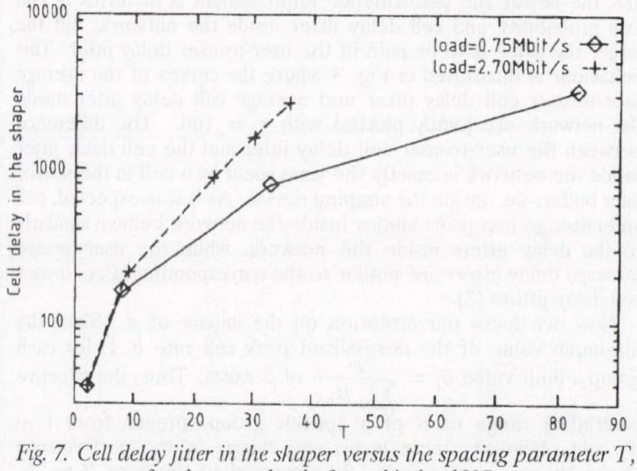


Fig. 7. Cell delay jitter in the shaper versus the spacing parameter T, for the mean load of two kinds of VCs

On the one hand, the curves in Fig. 7 show that the cell delay jitter in the user buffer can widely vary with the VC load, specially when the spacing parameter T is large: differences can grow larger than one order of magnitude. Thus, delay jitters can become extremely high for users with low bit rate. On the other hand, the availability of the results in Fig. 7 allows the shaper parameters to be dimensioned so as to provide "fair" conditions for the different users.

Consider the case of VCG shaping only, for simplicity. Up to now a "bandwidth-fair" approach was always assumed, in which all

VCs are given identical β values. Using the results of Fig. 7 it is also possible to dimension a "spacing-fair" case, where all VCs are given identical T values, or a "delay-fair" case, in which the shaper parameters are selected so as to make the cell delay jitter in the shaper similar for all VCs.

6. CONCLUSIONS

We have presented some new results on the application of a GCRA-based traffic shaping in the ATM end-point of connectionless sources. Both the mathematical analysis of the GCRA shapers and some simulation results on the effect of shaping are presented.

On the one hand, the performance measures of the GCRA-based shaper can be determined based on an iterative analysis. The analytical results can be used in dimensioning the GCRA-based shaper. Since the main part of the end-to-end delay jitter comes from the shaping stage, and only a small fraction from scheduling and queuing inside the network, the end-to-end delay jitter can be efficiently estimated.

One the other hand, a new algorithm is proposed for scheduling cells of different traffic flows after shaping. Some quantitative results show that the new scheduling scheme introduced a slightly higher delay than the scheduling proposed in [1], but it eliminates the problem that cells were found non-compliant by the GCRA conformance test.

Furthermore, a peak-to-mean ratio for measuring the controlling range of the shaper is introduced, which provides more appropriate insight into the significant differences in the impact of grouping on the QoS parameters [2]. The characterization of the relation between the peak-to-mean ratio and the burstiness of the cell flow leaving the shaper need further investigations.

REFERENCES

[1] M. Ajmone-Marsan, A. Bianco, T. V. Do, L. Jereb, R. Lo Cigno, and M. Munafó: "ATM simulation with CLASS", *Performance Evaluation*, 24 (1995), pp. 137-159.

[2] M. Ajmone-Marsan, T. V. Do, L. Jereb, R. Lo Cigno, R. Pasquali, and A. Tonietti: "Some simulation results on the performance of traffic shaping algorithms in ATM networks", In D. D. Kouvatsos, editor, *Performance Modelling and Evaluation of ATM Networks*. Chapman and Hall, London (IFIP), 1995.

[3] ATM Forum. *ATM User-Network Interface Specification Version 3.0*, September 1993. Englewood Cliffs, NJ: Prentice Hall.

[4] P. Boyer, Fabrice M. Guillemin, M. J. Servel, and J.P Coudreuse: "Spacing cells protects and enhances utilization of ATM network links", *IEEE Network Magazine*, September 1992.

[5] T. V. Do: Performance analysis of some traffic Control techniques in ATM networks. Ph.D. dissertation, Technical University of Budapest, 1995.

[6] T. V. Do, L. Jereb, and G. Wolfner: "Performance analysis of connectionless users with shaping", *COST TD(95)07*, 1995.

[7] ITU-T Recommendation I.371. Traffic Control and Congestion Control in B-ISDN, 1992, Geneve, Switzerland.

[8] L. Jereb, T. Jakab, M. Telek, A. Sipos and G. Paksy: "PLANET: A tool for telecommunication network planning and its application in Hungary", *IEEE JSAC*, 12(7), Sept. 1994.

[9] P. Tran-Gia and H. Ahmadi: "Analysis of a discrete-time $G^{[x]}/D/1 - S$ queueing system with application in packet-switching systems", in *IEEE INFOCOM'88*, pp. 861-870, New Orleans LA., 1988.

László Jereb received the M.S. degree in electrical engineering from Technical University of Budapest, in 1971, and the Candidate of Sciences in 1986. Since 1971, he has been with the Technical University of Budapest, where he has been engaged in research and education on queueing and reliability theory and telecommunication network planning. He has led several projects in developing software tools aiming at the reliability analysis of complex electronic systems and the planning and analysis of telecommunication networks. He participated in the Organizing Committee of several Relectronic Symposia and he was the technical chairman of NETWORKS'94 (6th International Network Planning Symposium). He is a member of IEEE Communication Society and of the Scientific Society for Telecommunications of Hungary.

Tien Van Do recieved his MSEE degree at the Technical Univesity of Budapest in 1991. He was a Ph.D. student working on ATM traffic control from 1991 to 1994. He is now a researcher at the Department of Telecommunications of the Technical University of Budapest and participates in the COPERNICUS-ATMIN 1463 project funded by EC. His research interests are queueing theory, ATM and telecommunication networks, performance evaluation of telecommunication networks.

György Wolfner received the M.S. degree in electrical engineering from the Technical University of Budapest in 1994. Since then he has been a Ph.D. student at the Department of Telecommunications, Technical University of Budapest. His current research interest are the performance analysis and simulation of ATM networks. He is a member of the Scientific Society for Telecommunications of Hungary.

SHAPING AND POLICING IN INTERCONNECTED ATM NETWORKS*

M. AJMONE MARSAN, A. BIANCO, R. LO CIGNO and M. MUNAFÒ

POLITECNICO DI TORINO, DIPARTIMENTO DI ELETTRONICA
CORSO DUCA DEGLI ABRUZZI, 24
I-10129, TORINO, ITALY
{AJMONE,BIANCO,LOCIGNO,MUNAFO@POLITO.IT}

The interaction between the shaping and policing functions in a scenario of interconnected ATM networks is investigated with CLASS, a simulation tool for the study of ATM networks. An abstract model of the interconnected ATM networks is first developed, in order to allow the performance study through simulation with acceptable complexity. Numerical results concerning the cell loss probability are then estimated, varying several network parameters, such as the burstiness of the sources and the cell delay variation tolerance of the policers. Two different types of policers are considered: the former discards all cells that do not conform to the parameters of the traffic contract; the latter tags such cells by setting the CLP bit in their header: cells are later discarded only if network congestion is experienced.

1. INTRODUCTION

The first deployments of national and international pilot wide area ATM networks (ATM WANs), and the availability of small, relatively inexpensive nodes for local ATM networks (ATM LANs), make the availability of real commercial long distance ATM networks in a short time an easy prediction. High-speed networks of the near future will thus feature the interconnection of ATM LANs through national and international ATM WANs.

While the hardware developments and the product offerings in the ATM field are progressing very rapidly, the definition and standardization of the traffic control and network management functions is still far from being settled. For example, one of the aspects of traffic control requiring a deeper investigation concerns the performance of the usage parameter control (UPC) functions in complex internetworking environments.

UPC functions are devoted to the control of the characteristics of the traffic on a virtual connection (VC), after the VC has been accepted and established; UPC functions can be divided into preventive functions (*traffic shaping*) and repressive functions (*traffic policing*). The goal of the traffic shaping function is the reduction of the burstiness of the flow of cells injected by the source user into the ATM network. This is achieved by adequately spacing cells belonging to a burst of activity of the source, in accordance to the traffic specifications agreed at the time of the VC setup. Traffic shaping is thus essentially a user function, performed before the ATM network ingress. Instead, traffic policing is a network function devoted to the control of the user behavior. Traffic policing enforces the traffic specifications agreed at the time of the VC setup, and acts upon cells that violate such specifications because they belong to a burst that has not been adequately shaped.

The basic algorithm for the implementation of both functions, and in general for traffic control in ATM networks, is the Generic Cell Rate Algorithm (GCRA) that was recommended by ITU-T [1] as well as by the ATM Forum [2].

Ideally, traffic policing should be enforced only at the user-network interface (UNI), in order to make sure that the flow of cells injected by the user into the network complies with the parameters of the traffic contract negotiated for the VC. However, it is generally agreed that, in a scenario featuring several interconnected ATM networks, every time a different network operator is involved, a new policer controls that the traffic crossing the broadband inter-carrier interface (B-ICI) defined between two adjacent networks is compliant with the negotiated parameters.

While the advantages offered by a careful smoothing of the burstiness of the offered traffic by means of shaping were demonstrated by several authors [3]–[7], the impact of repressive UPC has not yet been carefully investigated in the case of interconnected ATM networks, when policing is used not only at the UNI but also at every B-ICI crossed by the VC.

In this paper we study the interaction between the shaping and policing functions in a simple scenario featuring several interconnected ATM networks adopting policing both at the UNI and at every B-ICI.

The aim of our work is to gain a deeper insight on the effects of the multiplexing stages within a network on the cell loss ratios due to cell discarding by policers and buffer overflows.

The investigation is performed by simulation, varying several network parameters, such as the input traffic burstiness, the cell delay variation tolerance of the shapers and of the policers and the distance (measured in number of crossed B-ICIs) between the cell source and the policer where the cell is discarded.

The different behaviors observed for a given policing function as the distance parameter changes are fairly interesting, specially if we consider cell flows that upon generation comply with the traffic contract negotiated by the user at the time of VC setup. Indeed, in this case, the user might expect that the ATM networks do not discard cells unless congestion is experienced. Instead, the presence of several multiplexing stages between the source and the policer induces an increase in the burstiness of the cell flow, so that it becomes quite difficult for policers to discriminate between cells coming from a misbehaved source and cells whose delay jitter was increased beyond a given threshold by the network itself.

Two different types of policers are considered: the former discards all cells that do not conform to the parameters of the traffic contract; the latter tags such cells by setting the CLP (cell loss priority) bit: cells are later discarded only when network congestion is experienced.

The study is performed by simulation using CLASS [8], an ATM network simulator developed at the Electronics Department of Politecnico di Torino in cooperation with CSELT (Centro Studi E Laboratori Telecomunicazioni), the research center of Telecom Italy, and lately also in cooperation with the Technical University of Budapest.

The rest of the paper is organized as follows: Section 2 describes the simple model developed for the study of the interconnected ATM networks system. Section 3 presents the numerical results obtained from the simulation runs, and discusses the insights that can be gained from them. Finally, Section 4 ends the paper with some concluding remarks.

2. THE NETWORK MODEL

The aim of this work, as stated in the previous section, is the analysis of the interaction of the shaping and policing functions in a scenario featuring several interconnected wide area ATM networks. Given this general context, and considering a specific VC, we assume that the shaping function is performed only at the user side of the UNI, whereas several policing devices are present in the interconnected network system: one at the network side of the UNI, and one at each B-ICI that is crossed by the considered VC.

The scenario that we consider is fairly general, but its most obvious feature is that it is far too complex to be simulated in detail. We thus need to develop a simplified model that reduces the complexity of the investigation, while retaining the essential

* This work was supported in part by a research contract between Politecnico di Torino and CSELT, in part by the EC through the Copernicus project 1463 ATMIN, and in part by the Italian Ministry for University and Research.

characteristics of a system of interconnected wide area ATM networks: the number of VCs must be kept large enough to have a significant level of multiplexing, and several B-ICIs must be present.

For instance, if we imagine an ATM connection between our LAN at Politecnico di Torino and a corresponding University LAN in the USA, a reasonable internetworking scenario could comprise several B-ICIs: *Source* → (Private UNI) → *ATM LAN* → (Public UNI) → *Italian Public ATM WAN* → (B-ICI) → *Intercontinental ATM WAN* → (B-ICI) → *US national ATM WAN* → (B-ICI) → *US regional ATM WAN* → (Public UNI) → *ATM LAN* → (Private UNI) → *Destination*.

Following the guidelines deriving from these considerations, we propose a simple model where each network is represented by just one ATM switch, and where networks (switches) are connected through a B-ICI; the policing function is performed on all the VCs entering the switch, either from a UNI or from a B-ICI. Moreover, each traffic source in our model behaves as a traffic concentrator, generating traffic that refers to N VCs that are routed through the interconnected network system.

Fig. 1 illustrates the model of the interconnected network system that we used in our study. The model comprises 5 networks (represented by switches), connected to form a ring. In order to obtain a symmetrical model, we force all VCs to cross all the B-ICIs in the model. By so doing we obtain a rather simple and uniform environment, where all VCs cross the same number of policing devices, so that the number of parameters of the model (that in principle is extremely high) is reduced, and the interpretation of numerical results is simplified. Each node collects the traffic generated by 2 *source users* (the squares in the figure), and delivers the traffic to $N = 8$ *destination users* represented in the figure as a cloud of 8 users. In order to force the traffic to cross all the nodes in the system, all the traffic generated by the source users in node i is sent to the destination users in node $i - 1$ (modulo 5), and is routed clockwise to its destination, forcing it to cross all the nodes in the network. For example, the traffic generated in node 0 goes to node 4 crossing nodes 1, 2 and 3, as shown by the traffic relation represented by the dashed line in Fig. 1. The shaded squares and triangles in the figure represent the policing and shaping devices, respectively. All the links in the network have a capacity $C = 150$ Mbit/s; the size of the buffers in the nodes (both those on links going from one node to the next and those on links going to users) equals 200 cells, while the size of the transmission buffers at the users is 20,000 cells. This latter value is much bigger than the former since we assume that the transmitter should be able to store large bulks of data without losses. The link length has no influence on our results.

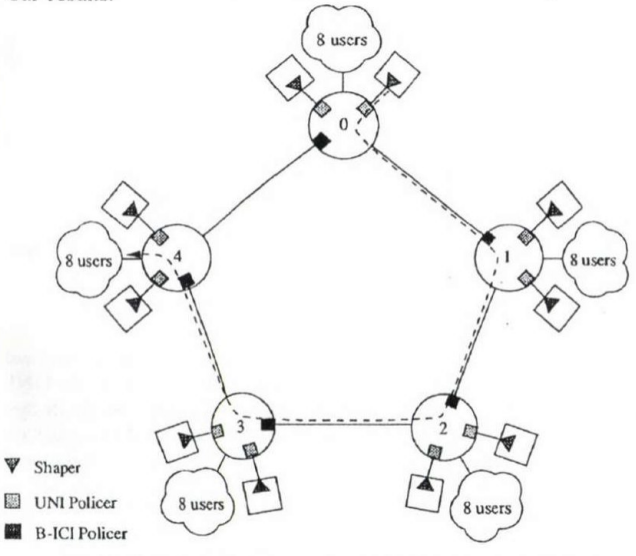


Fig. 1. Model of the interconnected ATM network system

Since the number of destination users is set to 8, each source user generates traffic on 8 different VCs. The traffic on each VC is separately shaped and then multiplexed with the traffic of the other VCs of the same source by the device described in Subsection 2.1.

The traffic on each VC is subjected to a policing function at the input of each node. The architecture and algorithms implemented by the policers are described in Subsection 2.2. Each B-ICI policer in the model controls 64 VCs; the 16 cell flows whose destinations are connected to the node are then extracted. The 16 flows generated locally are policed separately at the UNI and then multiplexed together with the remaining 48 flows coming from the preceding nodes. In the derivation of all our numerical results we assume that the source user negotiates with the network the peak VC transmission rate and that the shaper and the policers control this parameter; however, we assume that the source user also declares the mean VC transmission rate and that the network allocates the bandwidth with reference to the mean rate, in order to exploit statistical multiplexing as much as possible.

2.1. The Shaper

The shaper task is to modify the characteristics of the traffic generated by the source user in order to make them conforming with the traffic parameters negotiated between the user and the network at VC setup. The algorithm on which the shaper is based is the GCRA (Generic Cell Rate Algorithm) recommended by both the ITU-T and the ATM Forum for the policing function [1], [2]. The only difference with respect to the original specification of the algorithm consists in delaying the non-compliant cells instead of discarding or tagging them. Fig. 2 describes the GCRA algorithm with both the shaping and the policing options. A detailed description of this algorithm can be found in [5] and [6].

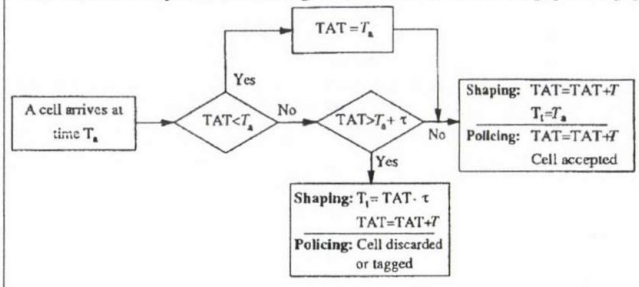


Fig. 2. Flow diagram of the Generic Cell Rate Algorithm used for shaping and policing: T_a = arrival time of a cell; T = nominal cell interarrival time; τ = cell delay variation tolerance; T_t = transmission time of a cell; TAT= theoretical arrival time of a cell

Although the basic algorithm is already a standard, the real architecture of a shaping device is still undefined, specially for what concerns the multiplexing stages before the actual cell transmission.

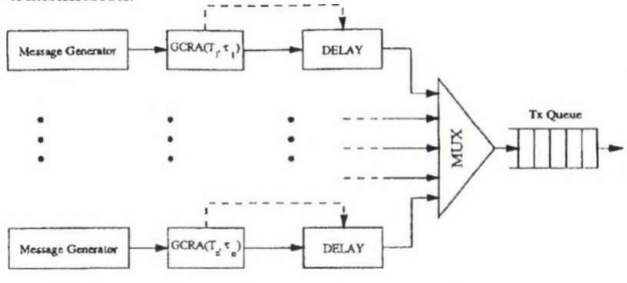


Fig. 3. Architecture of the shaper device implemented in the simulator

In [5], [6] the impact of different multiplexing strategies on the cell loss performance was investigated. We choose here the most performant among the alternatives considered in [5], [6]; the resulting shaper architecture is depicted in Fig. 3. Cells are generated by segmentation of larger data units called *messages*. As a result, cells are generated in batches and transferred to the device implementing the GCRA algorithm; cells are delayed until they become compliant with the traffic contract, specified in terms of T, the expected cell interarrival time, and τ, the cell

delay variation tolerance. After this shaping process, cells from different VCs are multiplexed and stored into a buffer, awaiting transmission. This multiplexing stage introduces some jitter in the cell flow. The implemented multiplexer algorithm is very simple: at each slot time the multiplexer visits each input queue according to a round robin schedule and transfers to the output buffer all the cells that, according to the preceding GCRA devices, are eligible for transmission. All cells are generated with the CLP bit equal to 0 and the shaper does not modify it, even when cells are delayed.

2.2. The Policer

Fig. 4 illustrates the architecture of the policers that were implemented in the simulator. The core of the device is again the GCRA depicted in Fig. 2.

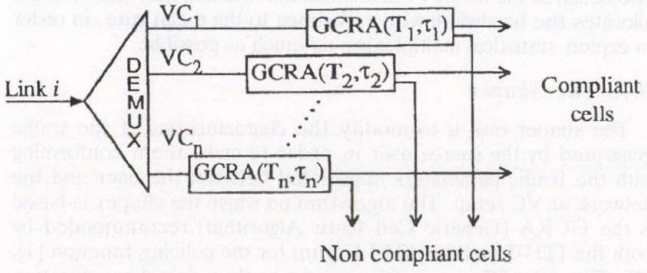


Fig. 4. Architecture of the policer device implemented in the simulator

Cells arrive at the policer from an input link of the node, and are de-multiplexed in order to separate the flows of cells relative to each VC. Each flow is then checked to control whether it still complies with its traffic contract, defined by means of T and τ. Compliant cells are forwarded to the switching matrix.

Two different actions can be envisaged for non-compliant cells, which define two different types of policer:

1. The *discarding policer*, that simply discards non-compliant cells.
2. The *tagging policer*, that tags non-compliant cells by forcing their CLP bit to 1, but still forwards them to the switching matrix, assuming that they will be discarded later on, if congestion effectively occurs. The policing algorithm acts only on untagged cells, since it is assumed that tagged cells are low-priority (they should not influence other cells) and can be discarded at any time.

The node switching fabric is assumed to be non blocking with output buffering. When discarding policers are implemented within a node, no further specification is required. When tagging policers are present, the node should be able to detect congestion, enabling the discarding of tagged cells. A simple possibility is to provide output buffers with a threshold: if the buffer is filled beyond the threshold, incoming tagged cells are not accepted, but tagged cells already in the buffer are not removed.

3. NUMERICAL RESULTS

In the previous section we illustrated the simple simulation scenario used in our investigation, and we emphasized that its simplicity and symmetry are partly due to an attempt to reduce the number of parameters. However, the parameter space of the simulation model is still very large, comprising quantities ranging from the network load to the burstiness of the source traffic and to the bandwidth allocation policy.

Before describing separately the results collected for both the discarding and the tagging policers, in Subsections 3.1 and 3.2 respectively, we briefly introduce the values of the main parameters as well as the general notation used in the discussion of results.

The traffic load

All traffic sources generate streams or groups of cells that are shaped using a cell delay variation tolerance equal to zero. Thus, the cell flow that each source submits to the network alternates CBR periods at the shaped cell rate, and silence periods. The cell stream after the shaping device is therefore largely independent of the traffic before the shaper, and in particular is very similar to the cell stream produced by an ON/OFF source.

The numerical results that are presented below refer to traffic sources that generate messages according to Poisson processes. User messages are then segmented within the ATM adaptation layer (AAL) to produce ATM cells. The message length is taken to be a geometrically distributed random variable whose mean equals 20 cells, the maximum length being truncated to 200 cells. However, we shall also present one set of numerical results derived with a different type of cell source, in order to confirm the irrelevance of the message generation process.

Each message source in the system generates 12.5 Mbit/s of data, equally partitioned among its 8 VCs; hence the total load offered to the system equals 125 Mbit/s. We assume that AAL 3/4 is used, so that the segmentation function produces 44-byte payloads to be inserted into ATM cells. Hence, the actual traffic on each link of the ring (the traffic that crosses each B-ICI) is about 120.5 Mbit/s, resulting in a link load factor $\rho \simeq 0.8$.

Actually, in our simulation experiments we considered a number of different loading factors and traffic patterns; however, the obtained numerical results always provided similar insight on the system behavior, so that only one set of simulation results is discussed here for the sake of brevity.

Traffic burstiness

The traffic burstiness is defined as the ratio between the peak and the mean rate of the cell flow after the shaping devices; it is indicated with the parameter β, and depends on the shaper parameters rather than on the source parameters.

Shaping and policing parameters

The GCRA is defined by means of the *cell interarrival time* T and the *cell delay variation tolerance* τ. We discriminate between the parameters relative to shapers and those relative to policers by means of subscripts; thus we have T_s, τ_s for the shapers and T_p, τ_p for the policers. All of these parameters refer to the peak emission rate of one specific VC, and are integer multiples of the slot time. For instance, a traffic relation with raw peak data rate $L = 3$ Mbit/s, and burstiness $\beta = 2$ will have $T_s = \lfloor \frac{C}{\beta L} \cdot \frac{53}{44} \rfloor = 30$. If not otherwise stated we let $T_p = T_s$.

Since shapers can buffer cells and actually decide when cells must be transmitted, there is no need to allow any tolerance for the cell transmission time, thus we always let $\tau_s = 0$. Instead, τ_p is one of the main parameters (together with the traffic burstiness β) under consideration in this study. Its value is crucial, since it should compensate the delay jitter introduced by the multiplexing stages within the different networks, in order not to penalize well-behaved users; but at the same time it should not relax too much the control on the traffic in order to detect misbehaved users and protect the network against congestion.

In the analysis of our simulation results we must keep in mind that in our scenario all sources are well-behaved, so that all arrivals of cells that do not conform to the traffic contract are due to multiplexing within the networks.

Before proceeding with the results description, it is worthwhile to remark the difference between the UNI policers and the B-ICI policers. Consider a traffic shaper with N VCs and denote with T_i the cell interarrival time of the i-th VC. If $\sum_{i=1}^{N} \frac{1}{T_i} \leq 1$ and $\tau_s = 0$ it can be shown that the multiplexing stage of a shaper like the one in Fig. 3 introduces a delay jitter δ smaller than N in the flows of cells coming from the GCRA devices. It follows that, under these conditions no cells can be discarded by the UNI policer if $\tau_p \geq N$. The UNI policers in our simulations are always in this condition unless $\tau_p < 8$. However, even in the case of $\tau_p = 2$ the discarding of cells by the UNI policer may happen only if 3 or more GCRA devices present cells at their output exactly at the same time instant, an event that, given the load offered by a single source, has a negligible probability. Hence we can expect that the UNI policers never discard cells, and this is confirmed by the simulations. The same cannot be said of the B-ICI policers, that handle VCs that may have already crossed a number of multiplexing stages.

Also note that the length of the simulation runs is such that an upper bound smaller than 10^{-7} on the cell loss ratio due to discarding at the policers can be derived for those simulation runs that do not produce any cell loss event.

3.1. Discarding Policers

We analyze here the results obtained when all policers in the system discard all non-compliant cells.

Figs. 5, 6 and 7 show curves of the cell loss ratio due to discarding at the policers as a function of the distance D (measured in number of crossed nodes) between the node where the cell was injected in the network and the node where it was discarded. A distance $D = 0$ means that the cell was discarded by the UNI policer, while a distance $D = 4$ means that the cell was discarded by the B-ICI policer of the destination node.

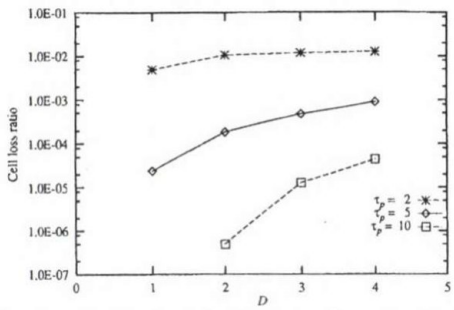


Fig. 5. Fraction of cells discarded by the policers as a function of the distance D between the node where the cell is generated and the node where it is discarded, with $\beta = 1$ and τ_p variable

The three figures refer to three different values of the traffic burstiness, namely $\beta = 1, 1.5, 2$; different values of the traffic burstiness imply different values of the parameter T for the GCRA, and different peak loads. Table 1 shows the values of the cell interarrival time T and of the peak load ρ_h, defined as the ratio between the number M of VCs present on a link and T (this definition is valid only for homogeneous traffic, otherwise the peak link load should be defined as $\sum_{i=1}^{M} \frac{1}{T_i}$). Observe that for larger values of β the peak link load grows larger than 1 (however, the average link load always remains about 0.8). Each figure contains several curves obtained with different values of τ_p.

Table 1. Values of the cell intergeneration time T and of ρ_h for the links between nodes, as a function of β

β	T	ρ_h
1	79	0.81
1.5	53	1.21
2	39	1.64

As expected, the numerical results in Figs. 5, 6, and 7 show that the cell loss ratio becomes smaller with increasing τ_p, regardless of the distance D from the source node. Simulation runs with $\beta = 1$ and $\tau_p = 20$ did not generate any cell loss event, and are thus not shown in Fig. 5.

Instead, two other phenomena that clearly emerge from these numerical results are less obvious: first, the cell loss ratio increases as D increases; second, the cell loss ratio increases as the traffic burstiness increases.

The increase of the cell loss ratio with D can be explained with the increased number of multiplexing stages that cells were subjected to. It is well-known that the delay jitter between cells of the same VC increases with the number of crossed multiplexers; however, it must be noted that every time a cell is discarded, a "gap" is created within the flow of cells, thus the cell that follows the discarded one benefits from a "delay bonus" in the sense that the next policer will expect it to arrive much earlier, i.e. at the TAT relative to the discarded cell. In spite of this, the discarding ratio increases with D. This fact is particularly annoying, since cells are discarded with higher probability when they are closer to their destination, i.e. when the network has already done most of the effort needed to transfer the cell; all this effort is thus wasted.

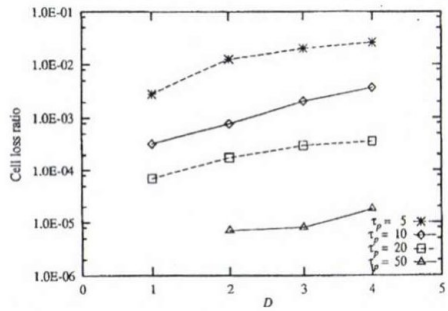


Fig. 6. Fraction of cells discarded by the policers as a function of the distance D between the node where the cell is generated and the node where it is discarded, with $\beta = 1.5$ and τ_p variable

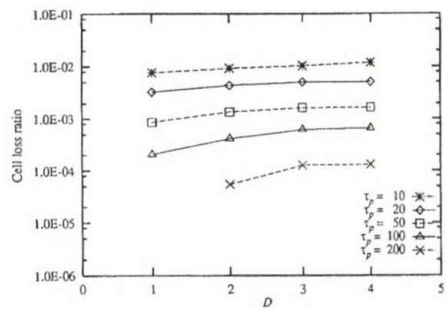


Fig. 7. Fraction of cells discarded by the policers as a function of the distance between D the node where the cell is generated and the node where it is discarded, with $\beta = 2$ and τ_p variable

Let's now consider the increase in the cell loss ratio with the traffic burstiness increases (notice that the range of variation of τ_p in the figures spans more than an order of magnitude going from $\beta = 1$ to $\beta = 2$). Although it could be imagined that an increase in the burstiness of the traffic can have a negative impact of the traffic control within the network, the magnitude of the phenomenon is greater than expected. In particular, the curves show the following behavior: when the peak cell rate (i.e. the inverse of the parameter T of the GCRA) is such that links can be temporarily overloaded (see Table 1), the cell loss ratio in the B-ICI policers remains very high even for great values of τ_p as shown in Fig. 7.

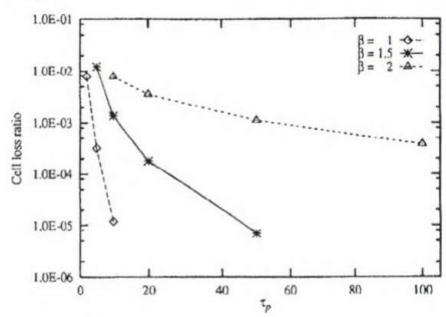


Fig. 8. Fraction of cells discarded by all the policers in the networks as a function of τ_p, for three different values of burstiness

Fig. 8 presents results obtained from the same simulation runs that originated the previous figures, but in this case the curves report the cell loss ratio averaged over all the policers in the system, as a function of the cell delay variation tolerance τ_p. The three curves refer to the three values of β we consider. In this figure the influence of the traffic burstiness is even more evident: the slope of the cell loss ratio as a function of τ_p decreases dramatically as we increase β.

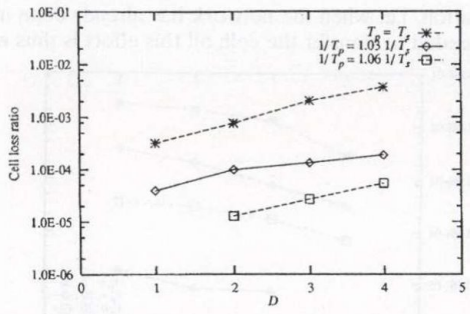


Fig. 9. Fraction of cells discarded by the policers as a function of the distance D between the node where the cell is generated and the node where it is discarded, with $\beta = 1.5$ and $\tau_p = 10$, when the bandwidth allocated by the policers is 3% or 6% higher than the one negotiated by the user

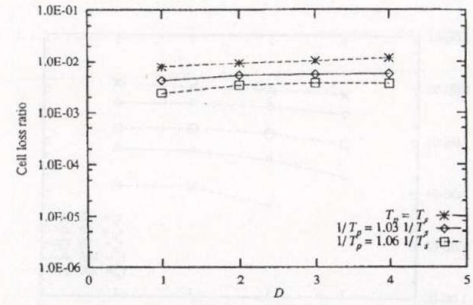


Fig. 10. Fraction of cells discarded by the policers as a function of the distance D between the node where the cell is generated and the node where it is discarded, with $\beta = 2$ and $\tau_p = 10$, when the bandwidth allocated by the policers is 3% or 6% higher than the one negotiated by the user

From the results presented so far, it is quite clear that if policers discard all non-compliant cells, and the expected interarrival time of cells at the policers is the same as the cell interdeparture time at the shapers (i.e. if $T_p = T_s$), the cell loss ratio for well-behaved sources is unacceptably high, whatever the value chosen for the cell delay variation tolerance at the B-ICI shapers.

Since the customer satisfaction requires that quality of service specifications must be met, one possible approach could be to use B-ICI policers with $T_p < T_s$, i.e. force the B-ICI policers to "over-estimate" the data rate of the VC.

Figs. 9 and 10 present cell loss ratio results in the cases where $1/T_p = 1.03 \cdot 1/T_s$ and $1/T_p = 1.06 \cdot 1/T_s$. The two figures refer to the cases $\beta = 1.5$ and $\beta = 2$, since with burstiness $\beta = 1$ no cells are discarded.

These numerical results indicate that acceptable cell loss ratios can be achieved only with a tight shaping of the source traffic before the UNI, and with a (slight) over-allocation of resources to VCs within the network. Indeed, the only good set of results is the one that we do not show because no cell loss was ever observed; it was obtained with $\beta = 1$ and with $1/T_p = 1.03 \cdot 1/T_s$ and $1/T_p = 1.06 \cdot 1/T_s$.

If we either allow an increase in the source traffic burstiness ($\beta = 1.5, 2$), or allocate to VCs in the network exactly the amount of bandwidth they require ($1/T_p = 1/T_s$), the cell loss probabilities grow to unacceptable levels.

This is due to the fact that the delay jitter introduced by the multiplexing stages within the network is very high. Unfortunately, the over-estimation of the data rate on VCs implies a waste of resources, which grows larger for larger traffic burstiness. This is quite a negative result, which seems to prevent an efficient exploitation of the statistical multiplexing gain.

Before moving to the presentation of results referring to tagging policers, we briefly discuss the irrelevance of the message generation process that was previously mentioned. Fig. 11 shows results equivalent to those previously presented in Fig. 6, but now sources inject into the ATM network cells according to ON/OFF streams. The reader can immediately observe that no qualitative difference exists between the two sets of results, as expected due to the presence of shaping devices with cell delay variation tolerance equal to zero.

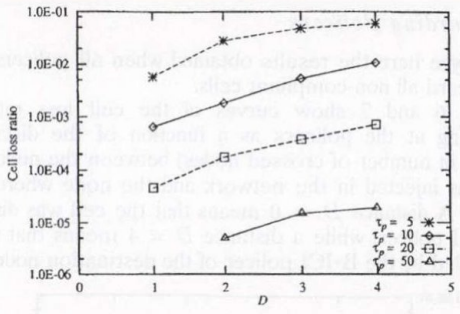


Fig. 11. Same as Fig. 6, but with ON/OFF cell flows

3.2. Tagging Policers

Up to now we considered only the case of policers that discard all non conforming cells, without bothering to control whether the network is congested or not. Especially at medium loads, when the multiplexing of different cell flows can introduce a significant delay jitter, but the chances of real network congestion are very low, this policy can lead to unnecessary cell losses and hence to a significant waste of resources.

Now we consider the case of policers that simply tag the non-conforming cells and forward them to the switching matrix: cells will be discarded only if congestion occurs somewhere, either in the output buffer of the node where the cell was tagged, or in any buffer of the subsequent nodes. The congestion in the nodes is measured by means of a threshold (denoted by S) within the buffers: when the buffer is full above this threshold, incoming tagged cells are discarded.

Fig. 12 presents the results obtained with this policy, when $\beta = 2$, for different values of the threshold as a function of τ_p.

Two groups of curves are present in this figure: the curves in the upper group (dashed lines) refer to tagged cells discarded because the buffer was filled above the threshold; the curves in the lower group (solid lines) refer to cells (mostly untagged) lost because of buffer overflow.

The dot-dashed horizontal line without marks between the two groups indicates the fraction of cells lost for buffer overflow when policers are not present, while the dotted line with the cross marks refers to the case when policers discard all non-conforming cells, and is reported for comparison purposes.

For $\beta = 1, 1.5$ no cells are discarded, so we do not present curves for those cases.

Notice that the values of τ_p considered in this case are much lower than those used to obtain the results presented in Fig. 12.

In this case we report only the overall average cell loss ratio, because with tagging policers the loss ratio does not depend significantly on D.

If we consider the case of no policers and the case in which policers discard all non-compliant cells as reference cases, it is easy to verify that the case of tagging policers offers quite remarkable advantages over both of them. Indeed, the cell loss ratio can be reduced by more than two orders of magnitude with respect to the case of policers that discard all non-compliant cells. At the same time, the fraction of cells lost for buffer overflow decreases significantly with respect to the case when no policers are present, showing an increased protection against network congestion.

Fig. 12 shows that tagging policers offer real advantages with respect to discarding policers with a small increase in the complexity of the node.

Contrary to what happened for discarding policers, acceptable cell loss ratios can now be achieved also for bursty traffic, without any overallocation of resources to VCs within the network.

This is quite a positive result, which provides confidence in the possibility of exploiting the statistical multiplexing gain. However, more accurate investigations are necessary to ensure that the

presence of misbehaving users in a network using tagging policers does not have a negative impact on well-behaved sources.

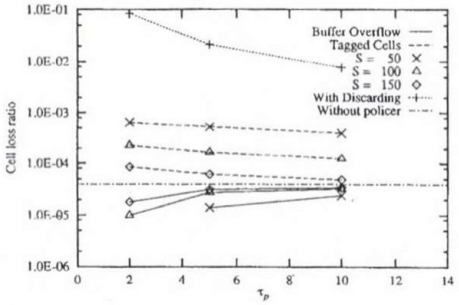


Fig. 12. Fraction of tagged cells discarded when buffers are filled beyond the threshold (dashed lines); fraction of cells lost for buffer overflow (solid lines); cell loss ratio when policers are not present (dot-dashed line); fraction of cells discarded by discarding policers (dotted line); with $\beta = 2$, for different values of the threshold S, as a function of τ_p

4. CONCLUSION AND FUTURE WORK

In this paper we investigated the interaction of the shaping and policing functions within a simplified model of a system of interconnected wide area ATM networks.

A five-node topology was used as a representative simple model of a system of interconnected ATM networks, adopting the characteristics that allow the study of the impact of policing performed at B-ICI on traffic of sources that were originally well-behaved.

The study showed that the implementation of several policing functions on long distance ATM connections presents a number of problems that are still unresolved unless resources are allocated according to the peak rate of the sources, a solution that wastes a great amount of resources and does not allow the exploitation of the statistical multiplexing gain.

Even when sources are not bursty (either tightly shaped VBR sources or CBR sources), the delay jitter introduced by the multiplexing stages within the network is high enough to create problems in controlling the traffic after a few nodes have been crossed.

Cell tagging seems to offer significant advantages with respect to the simple solution of discarding all non conforming cells; however, more investigations are necessary in order to find policing functions offering satisfactory performances both in protecting the network against congestion and against misbehaving sources.

REFERENCES

[1] ITU-TSS Study Group 13, Recommendation I.371 "Traffic Control and Congestion Control in B-ISDN", Geneve, Switzerland, July 1995.

[2] ATM Forum, "ATM User-Network Interface Specification", First Edition, Prentice-Hall.

[3] P. E. Boyer, F. P. Guillemin, M. J. Servel and J. P. Coudrose: "Spacing cells protects and enhances utilization of ATM network links," *IEEE Network*, Vol. 6, pp. 38-49, Sept. 1992.

[4] P. E. Boyer, M. J. Servel and F. P. Guillemin: "The spacer-controller: An efficient UPC/NPC for ATM networks", *ISS'92*, Vol. 2, Yokohama, Japan, p. A9.3, Oct. 1992.

[5] M. Ajmone Marsan, T. V. Do, L. Jereb, R. Lo Cigno, R. Pasquali and A. Tonietti: "Simulation of Traffic Shaping Algorithms in ATM Networks", *Second Workshop on Performance Modeling and Evaluation of ATM Networks*, Bradford, UK, July 1994.

[6] M. Ajmone Marsan, T. V. Do, L. Jereb, R. Lo Cigno, R. Pasquali and A. Tonietti: "Simulation of Traffic Shaping Algorithms in ATM Networks", in: D. Kouvatsos (editor), *Performance Modelling and Evaluation of ATM Networks*, Chapman and Hall, London, 1995.

[7] Baiju V. Patel and Chatschik C. Bisdikian: "On the Performance Behavior of ATM End-stations", *IEEE INFOCOM'95*, Boston, MA, USA, April 4-6, 1995, pp. 188-196.

[8] M. Ajmone Marsan, A. Bianco, T. Van Do, L. Jereb, R. Lo Cigno and M. Munafò: "ATM Simulation with CLASS", *Performance Evaluation*, 24 (1995), pp. 137-159.

Marco Ajmone Marsan is a Full Professor at the Electronics Department of Politecnico di Torino, in Italy. He holds a Dr.Ing. degree in Electronics Engineering from Politecnico di Torino, and a Master of Science from the University of California, Los Angeles. Since 1975 to 1987 he was at the Electronics Department of Politecnico di Torino, as an associate professor. Since 1987 to 1990 he was a Full Professor at the Computer Science Department of the University of Milan, in Italy. During the summers of 1980 and 1981 he was with the Computer Science Department, UCLA. He has coauthored over 150 journal and conference papers as well as the two books in the areas of communications and computer science. He received the best paper award at the Third International Conference on Distributed Computing Systems in Miami, Fla., in 1982. His current interests are in the fields of performance evaluation of data communication and computer systems, communication networks and queueing theory. He is a Senior Member of IEEE.

Maurizio Munafò is a research engineer in the Electronics Department of Politecnico di Torino. He holds a Dr.Ing. degree in Electronics Engineering since 1991 and a Ph.D. in Telecommunications Engineering since 1994, both from Politecnico di Torino. Since 1991 he has been with the Electronics Department of Politecnico di Torino, where he has been involved in the development of an ATM network simulator.

Andrea Bianco is a Research Assistant at the Electronics Department of Politecnico di Torino. He holds a Dr.Ing. degree in Electronics Engineering and a Ph.D. in Telecommunications Engineering both from Politecnico di Torino. Since 1991 to 1994 he has been with the Electronics Department of Politecnico di Torino as a Ph.D. student. In 1993 he spent one year at HP Labs in Palo Alto, CA, researching on ATM networks. His current research interests are in the field of access protocols for all-optical networks, performance analysis of ATM networks, simulation of communication protocols and formal description techniques.

Renato Lo Cigno is a research engineer in the Electronics Department of Politecnico di Torino. He received a Dr.Ing. degree in Electronics Engeneering from Politecnico di Torino in 1988. Since then he has been with the telecommunication research group of the department. His research interests are in communication networks simulation and performance analysis. He is the coauthor of about 20 journal and conference papers in the areas of communications.

CORRELATIONS IN ATM CELL STREAMS EXPOSED TO CELL DELAY VARIATION

S. MOLNÁR
DEPARTMENT OF TELECOMMUNICATIONS & TELEMATICS
TECHNICAL UNIVERSITY OF BUDAPEST,
H-1111 BUDAPEST, SZTOCZEK U. 2., HUNGARY
E-MAIL: MOLNAR@BME-TEL.TTT.BME.HU

S. BLAABJERG
DEPARTMENT OF TELECOMMUNICATIONS
TECHNICAL UNIVERSITY OF DENMARK
DK-2800 LYNGBY, DENMARK
E-MAIL: SBB@TELE.DTU.DK

In this paper the correlation structure of a constant bit rate (CBR) cell stream is investigated after it has passed through a FIFO queue in which it is multiplexed with a background traffic stream. The background stream arrives in the form of independent and identically distributed batches with general batch size distribution. An exact Markovian model and a computationally simpler diffusion model are used to obtain the correlation between interdeparture times of the perturbated CBR stream at the output of the queue. Numerical results are provided to evaluate the accuracy of the diffusion approximation. For the diffusion approximation it is proven that the correlation is always negative and that it converges to zero with rate $O(n^{-1/2}exp(-\alpha n))$ as the lag n tends to infinity. α is a positive constant which is determined by the drift and variance in the diffusion model, together with the interarrival time of the CBR stream.

1. INTRODUCTION

In the new standards for the Broadband Integrated Service Digital Network (B-ISDN) and its switching and transport mechanism Asynchronous Transfer Mode (ATM), the connection set up procedure includes a negotiation of a traffic contract between the user and the network [14]. At the entrance to the network (UNI) and at the border between networks owned by different operators (NNI) an enforcement, called usage parameter control (UPC) at the UNI and network parameter control (NPC) at the NNI, of the traffic parameters in the traffic contract of each connection will take place.

The *peak cell rate* of the connection is the most important traffic parameter in the traffic contract and it is the only traffic parameter which is mandatory.

The definition and the use of the peak cell rate for UPC/NPC and other traffic controls is made complicated due to the phenomenon that different cells of the same connection experiences different delays between the source and the UPC/NPC device due to switching and multiplexing with other cell streams. This phenomenon is called Cell Delay Variation (CDV), and it has forced the standardization bodies to define the peak cell rate by means of *two* parameters.

- The peak emission interval T, which is the distance between two adjacent cells of the connection transmitted at peak rate. Equivalently $1/T$ is the *peak cell rate* of the connection.
- The *CDV tolerance* τ. τ is to be interpreted as a measure of how much the instantaneous peak cell rate is allowed to deviate from the peak cell rate declared by T.

In ITU a UPC/NPC algorithm has not been standardized. However, it is standardized how to measure whether a cell stream is conforming to the traffic contract or not. The *Generic Cell Rate Algorithm* (GCRA) [14] characterized by the two parameters T and τ, is standardized for exactly this purpose. Essentially the GCRA works like a leaky bucket, that is, credits are generated at a rate of $1/T$ such that a maximum proportional to τ can be present simultaneous. Each arriving cell needs a credit to be declared conforming, and if too many cells arrive too close to each other all credits will be used, and cells arriving when no credits are present will be declared non-conforming.

Several models has been proposed for the setting of T and τ such that well behaving cell streams exposed to CDV do not suffer from a large amount of cells declared non-conforming by the GCRA.

The most frequently applied approach has modeled the CDV perturbated cell stream by a renewal stream. The variance of the time between adjacent cells of the CBR stream has been computed [11], [5] or measured/simulated [12]. A distribution matching the mean and variance has then been used and the associated GI/D/1 queue[1] has been solved yielding proper values for T and τ. However, it has turned out that this modeling approach predicts the τ value to diverge as the leak rate $1/T$ of the GCRA approaches the physical peak rate of the connection.

This important observation can be explained by the following way. The assumption of a renewal stream makes the number of credits available behave like a random walk with two barriers (0 and maximum number of credits) and with a drift equal to the difference between the credit generation rate $1/T$ and the cell arrival rate. When the credit generation rate approaches the cell arrival rate the associated random walk approaches a zero-drift random walk for which it is known that it will hit the barriers with a frequency inverse proportional to to the distance between the barriers. Therefore, in the limit where the credit generation rate equals the cell arrival rate, an enormous amount of credits (enormous τ) is needed in order to make the probability of non-conformance sufficiently small.

In reality a much smaller CDV tolerance value will be sufficient. Both simulation results [9], numerical results [4], and a simple worst case dimensioning scheme [20] which works when the peak and leak rate are equal, clearly shows a finite τ value also when the leak rate is put equal to the peak rate. The misleading behaviour of the renewal models can be explained by the fact that the CBR connection for which CDV has been introduced possesses negative correlations. Negative correlations implies that many arrivals presently will make likely few arrivals in near future which implies that as the number of credits decreases there is an increased tendency to fewer arrivals keeping the number of credits away from the zero barrier.

Thereby the correlation structure of the CDV perturbated CBR streams plays an important part in the description of the cell stream, and models which take this into account are needed. Can models which on one hand give an accurate description of the correlated cell stream and on the other hand are applicable in practice be derived? Will it be possible to compute the correlation between two arbitrary chosen interdeparture intervals (intervals between adjacent CBR cells)? Is it possible to derive something about the range of the correlations, i.e how fast do they vanish? Are they always negative, etc.? This paper is intended to answer these questions by presenting an analysis of a CBR cell stream which is multiplexed with a background stream in a FIFO queue.

The analysis is based on two models. The first one is an exact Markovian model presented in [11] and generalized by the authors in [19] which provides a solution which is numerical in nature and computational non-trivial. The second one is an approximate diffusion model presented in [3] and further developed by the authors in [19] which enables the point process characteristics to be computed in closed form and also provides a simple tool for the performance evaluation of the multiplexed CDV affected CBR cell streams [18].

2. EXACT MARKOVIAN MODEL

Consider a single server discrete time FIFO queue with infinite buffer capacity receiving the superposition of a CBR stream with interarrival time T and an interfering background stream which arrives to the queue in the form of independent and identically distributed batches with a general batch size distribution denoted by $b(k)$ (Fig. 1). Time is slotted, with the slot size being equal to the cell transmission time. Choose the time such that cell no. n arrives at time nT to the queue. Let W_n denote the waiting time of cell no. n. Then $P\{W_n = k\}$ is the probability that cell no. n

[1] which can be viewed as equivalent to the GCRA, see [13]

has a waiting time of k time units.

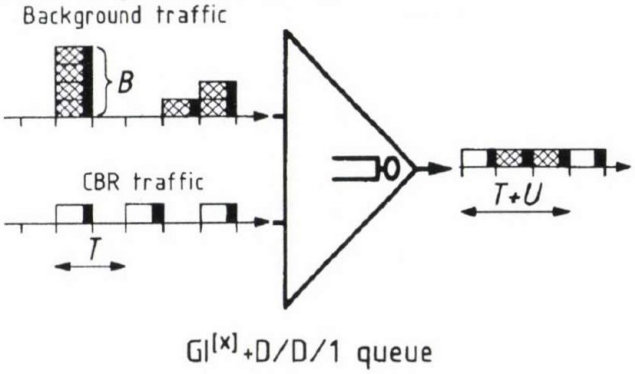


Fig. 1. The FIFO model

The sequence W_n is assumed stationary with distribution $w_k = P\{W_i = k\}$. Furthermore, due to the assumptions made, then the waiting time of successive CBR cells constitutes a discrete time Markov process in which the underlying transition matrix $\boldsymbol{Q}$ can be computed by the following way.

The element (j,k) of Q is defined as follows:

$$q_{j,k} = P\{W_i = k \mid W_{i-1} = j\} \quad j,k \geq 0 \tag{1}$$

and

$$q_{j,k} = Q(j,k-1) - Q(j,k), \tag{2}$$

where

$$Q(j,k) = P\{W_i > k \mid W_{i-1} = j\} = \sum_{n \geq 0} P_n(j,k), \tag{3}$$

with $P_n(j,k) = P\{W_i > k \mid W_{i-1} = j$ and n arrivals in $]Ti - T, Ti]\} b^{T*}(n)$, where $b^{T*}(n)$ is the T-fold convolution of the batch size distribution, and evaluated in n it represents the probability of n arrivals in $]Ti - T, Ti]$. It can be shown, see Appendix A in [19], that

$$P_n(j,k) = \begin{cases} 0 & j+1 \geq T \text{ and } n \leq T+k-j-1 \text{ or } j+1 < T \text{ and } k \geq n \\ b^{T*}(n) & n > T+k-j-1 \\ \sum_{s=1}^{n-k} b^{s*}(k+s) \times & \\ \quad \times b^{(T-s)*}(n-k-s)\frac{T-n+k}{T-s} & j+1 < T \text{ and } k < n \leq T+k-j-1 \end{cases} \tag{4}$$

The stationary waiting time distribution can be found from solving the equilibrium system $\mathrm{w} = \boldsymbol{wQ}$ where $\boldsymbol{w} = (w_0, w_1, \ldots, w_i, \ldots)$ and $Q = \{q_{j,k}\}$. In numerical applications, a truncation of the state space is in general necessary.

For the steady state queue length distribution seen by an arbitrary CBR cell, [5] provides a closed form expression for the generating function. [5] also provides the generating function of the waiting time for cell no. $n+1$ conditioning on the event that waiting time for cell n is i. From this the entries in the transition matrix can in principles be found. However, this requires not only an inversion but also the determination of $T-1$ boundary probabilities, and we have found the direct approach more appealing.

Define $\tau_n = nT + W_n$. Thus τ_n denotes the departure time [2] of cell no. n. The interdeparture time between cell no. j and cell no. $j+1$ is denoted X_j and the distribution is independent of j due to the stationary assumption. Define the shifted interdeparture time of cell no. n as: $U_n = \tau_n - nT - \tau_0$ (see Fig. 2). It is a fundamental random variable with zero mean from which all point process characteristics of interest can be derived. Let $f_n(k) = P\{U_n = k\}$ be the probability distribution of U_n for cell no. n. It can be seen [11] that

$$f_n(k) = \begin{cases} \sum_{i=0}^{\infty} w_i q_{i,i+k}^{(n)} & \text{if } k \geq -nT+1 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $q_{j,k}^{(n)}$ denotes entry (j,k) in the n'th power of the transition matrix $\boldsymbol{Q} = \{q_{j,k}\}$.

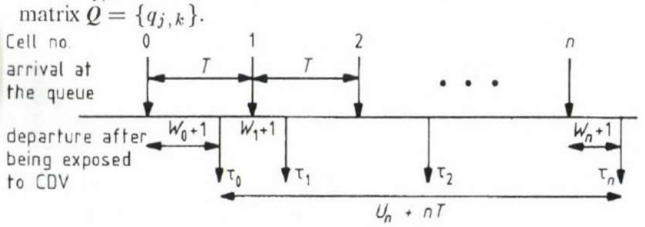


Fig. 2. The shifted interdeparture time is to be seen as the difference between the actual departure time of cell n (τ_n) and the expected departure time ($\tau_0 + nT$)

[2] Strictly speaking, τ_n denotes the time cell no. n enters service. However, this has no effect when service time is deterministic.

The *interdeparture time distribution* between cell no. 0 and n is simply obtained by the translation nT of the shifted interdeparture time distribution U_n i.e.

$$P\{\tau_n - \tau_0 = k\} = P\{U_n = k - nT\} = f_n(k - nT). \tag{6}$$

The variance of the interdeparture time between cell 0 and cell n denoted $V(n)$ is:

$$V(n) = Var(X_1 + \ldots + X_n) = Var(\tau_n - \tau_0) = \tag{7}$$
$$= \sum_{k=-nT+1}^{\infty} k^2 f_n(k).$$

From the basic formula $Var(X_1 + \ldots + X_n + X_{n+1}) = Var(X_1 + \ldots + X_n) + Var(X_{n+1}) + 2Cov(X_1 + \ldots + X_n, X_{n+1})$ it is straightforward by induction to prove that

$$C(n) := Cov(X_{n+1}, X_1) = \tag{8}$$
$$= \frac{Var(X_1 + \ldots + X_{n+1}) + Var(X_1 + \ldots + X_{n-1})}{2} -$$
$$- Var(X_1 + \ldots + X_n) = \frac{V(n+1) + V(n-1)}{2} - V(n).$$

Since the sequence $\{X_n\}$ is assumed stationary the *correlation* is given as

$$\mathrm{corr}(n) = \frac{\frac{V(n+1)+V(n-1)}{2} - V(n)}{V(1)} \tag{9}$$

i.e. in terms of the variance function alone.

3. DIFFUSION APPROXIMATION

The exact approach is in general computationally hard and the numerical complexity increases without bounds when the load on the queue approaches one. Therefore an accurate approximation which maintains the important qualities of the exact model is desirable. The following diffusion model is suggested.

The main idea of the diffusion approach is to model the evolution of the unfinished work (or virtual waiting time) between CBR arrivals by a reflected Brownian motion. Markovian dependencies are assumed between successive CBR cell waiting times in the queue as in the exact case.

Let $\tilde{W}_t$ denote the waiting time a fictitious observer would experience if he joined the diffusion queue at time t (the unfinished work at time t). The probability of $\tilde{W}_t \leq x$

conditioned on $\tilde{W}_0 = y$ is, for a Brownian motion with drift m (m assumed smaller than zero in order to ensure a stable queue), variance σ^2 and a reflection in zero, in section 2.8 of [16] derived to be:

$$P\{\tilde{W}_t \leq x \mid \tilde{W}_0 = y\} = \tag{10}$$

$$= \begin{cases} \Phi(\frac{x-y-mt}{\sigma\sqrt{t}}) - e^{2\frac{mx}{\sigma^2}}\Phi(\frac{-x-y-mt}{\sigma\sqrt{t}}), & \text{for } x \geq 0 \\ 0, & \text{for } x < 0 \end{cases}$$

where Φ denotes the standard Gaussian probability distribution. The right hand side of formula (10) is a distribution function in x for all $y \geq 0$ and all $t > 0$, and it converges towards the exponential distribution with mean $\frac{-\sigma^2}{2m}$, independent of initial condition y, when t tends to infinity.

The waiting time of CBR cell no. n, is approximated by:

$$W_n = \tilde{W}_{nT} \tag{11}$$

that is *the waiting time that CBR cell no. n experiences in the queue is approximated by the virtual waiting time in the diffusion queue at time nT.*

The diffusion approximation is used for modeling the arrivals of cells belonging to both the CBR stream and the background stream, i.e. both the CBR arrivals as well as the background arrivals and the departures are taken into account by a proper choice of drift and variance.

As in the diffusion model for the M/G/1 queue (see section 2.8 in [16]) the drift is $m = \rho - 1$. The key idea of choosing the most appropriate variance is that *we match the decay rates of the stationary waiting time distribution in the diffusion approximation with the asymptotic decay rate of the $D + GI^{[x]}/D/1$ queue.*

It is known from [6] that the queue length seen by an arriving CBR cell is asymptotically geometric i.e. $P(Q > r) \approx (1/z_\infty)^r$, where z_∞ is the dominant root of the corresponding z- transform which is found by solving the equation

$$B(z)^T - z^{T-1} = 0, \tag{12}$$

where $B(z)$ is the generating function of the batch size distribution.

The root z_∞ is the root with smallest module outside the unit disk, and it is unique and real (see Appendix A3 in [6]).

Since the diffusion decay rate is $\frac{-2m}{\sigma^2}$ we get:

$$\sigma^2 = -\frac{2m}{\ln(z_\infty)}. \tag{13}$$

3.1. The distribution of the interdeparture time

The shifted interdeparture time in the diffusion context is: $\tilde{U}_t = \tilde{W}_t - \tilde{W}_0$ (think of $t = nT$), and

$$P\{\tilde{U}_t \leq x\} =$$

$$= \int_0^\infty P\{\tau_t - \tau_0 - t \leq x \mid \tau_0 = y\} dP\{\tau_0 \leq y\} = \tag{14}$$

$$= \int_0^\infty P\{\tilde{W}_t \leq x + y \mid \tilde{W}_0 = y\} dP\{\tilde{W}_0 \leq y\}.$$

The computation carried out in [3] shows that the time dependent shifted interdeparture time distribution function is given as:

$$\tilde{F}_t(x) = P\{\tilde{U}_t \leq x\} = \tag{15}$$

$$= \begin{cases} \frac{1}{2} + \frac{1}{2}\Phi(\frac{x-mt}{\sigma\sqrt{t}}) - \frac{1}{2}e^{\frac{2mx}{\sigma^2}}\Phi(-\frac{x+mt}{\sigma\sqrt{t}}), & \text{for } x \geq 0 \\ \frac{1}{2}e^{-\frac{2mx}{\sigma^2}}\Phi(\frac{x-mt}{\sigma\sqrt{t}}) + \frac{1}{2}\Phi(\frac{x+mt}{\sigma\sqrt{t}}), & \text{for } x < 0 \end{cases}$$

From this expression it is seen that the probability distribution function $\tilde{F}_t$ for $\tilde{U}_t$ fulfills the relation: $\tilde{F}_t(x) + \tilde{F}_t(-x) = 1$, thus implying that the density function for $\tilde{U}_t$ is symmetric.

The interdeparture time distribution between cell no. k and $k+n$ is obtained from the shifted interdeparture time distribution as

$$F_n(x) = P\{\tau_{k+n} - \tau_k \leq x\} = P\{\tau_n - \tau_0 \leq x\} = \\ = P\{\tilde{U}_{nT} \leq x - nT\} = \tilde{F}_{nT}(x - nT). \tag{16}$$

3.2. The Variance Function

The variance function of $\tilde{U}_t$ is in [3] found to be:

$$Var(\tilde{U}_t) = \frac{\sigma^4}{2m^2}(2\Phi(-\frac{m}{\sigma}\sqrt{t}) - 1) + (2\sigma^2 t + m^2t^2)\Phi(\frac{m}{\sigma}\sqrt{t}) + \\ + (\frac{\sigma^3}{m}t^{1/2} + m\sigma t^{3/2})\varphi(\frac{m}{\sigma}\sqrt{t}), \tag{17}$$

where φ is the standard normal density function. By applying the expansion $\Phi(-x) = \frac{\varphi(x)}{x}(1 - \frac{1}{x^2} + O(x^{-4}))$, for $x \to \infty$, see e.g. problem 7.7.1 in [8], we obtain

$$Var(\tilde{U}_t) =$$

$$= \frac{\sigma^4}{2m^2} + (3\frac{\sigma^5}{m^3}t^{-1/2} + O(t^{-3/2}))\varphi(\frac{m}{\sigma}\sqrt{t}) \quad \text{for } t \to \infty. \tag{18}$$

3.3. The Covariance

From the approximation $W_n \approx \tilde{W}_{nT}$ we get $Var(X_1 + \ldots + X_n) \approx Var(\tilde{U}_{nT})$ from which

$$\tilde{C}(n) = \frac{Var(\tilde{U}_{(n+1)T}) + Var(\tilde{U}_{(n-1)T})}{2} - Var(\tilde{U}_{nT}). \tag{19}$$

Applying formula (17) it is seen that the covariance and therefore the correlation of any lag can be given in closed form.

Proposition 1: The correlation of any lag is negative.

Proof: From formula (8) it is seen that it suffices to prove that the function $[t \to Var(\tilde{U}_t)]$ is concave. Two times differentiations of (17) with respect to t yields

$$\frac{d^2Var(\tilde{U}_t)}{dt^2} = 2m^2\Phi(\frac{m}{\sigma}t^{1/2}) + 2\sigma m t^{-1/2}\varphi(\frac{m}{\sigma}t^{1/2}). \tag{20}$$

From the classical inequality $\Phi(-x) < \frac{\varphi(x)}{x}$ valid for $x > 0$ it is seen that $\frac{d^2Var(\tilde{U}_t)}{dt^2} < 0$ implying $Var(\tilde{U}_t)$ is concave.

Proposition 2: The correlation decays as $O(n^{-1/2}e^{-\frac{m^2}{2\sigma^2}Tn})$ as the lag n tends to infinity.

Proof: When the lag tends to infinity we may apply equation (18) which together with equation (8) yields

$$\tilde{C}(n) = \frac{3\sigma^5}{m^3(2\pi T)^{1/2}}n^{-1/2}e^{-\frac{m^2}{2\sigma^2}Tn}[\frac{1}{2}(\frac{n}{n+1})^{1/2}e^{-\frac{m^2}{2\sigma^2}T} + \\ + \frac{1}{2}(\frac{n}{n-1})^{1/2}e^{\frac{m^2}{2\sigma^2}Tn} - 1 + O(n^{-3/2})] = \\ = \frac{3\sigma^5}{m^3(2\pi T)^{1/2}}n^{-1/2}e^{-\frac{m^2}{2\sigma^2}Tn}[\frac{1}{2}e^{-\frac{m^2}{2\sigma^2}T} + \\ + \frac{1}{2}e^{\frac{m^2}{2\sigma^2}T} - 1 + O(n^{-1})] \tag{21}$$

proving the proposition.

In many cases $\frac{m^2}{2\sigma^2}T$ is smaller than 1, and a second order Taylor expansion of the exponential can be applied yielding

$$\tilde{C}(n) \approx [\frac{3}{4}\frac{\sigma m}{(2\pi)^{1/2}}T^{3/2}n^{-1/2} + O(n^{-1})]e^{-\frac{m^2}{2\sigma^2}Tn} \tag{22}$$

when $\frac{m^2}{2\sigma^2}T$ is small and n is large.

Based on the above analysis it should be noticed:

- The correlations are negative, which implies that ignoring them yield results in queueing applications which overestimates waiting times, CDV tolerance etc.
- The rate at which the correlations decay to zero in the lags has been obtained such that it is possible to set the range within which correlations should not be neglected.

4. NUMERICAL EXAMPLES

In this section some numerical examples are presented to illustrate the behaviour of both models. First we investigate the effect of the background burstiness and load on the shifted interdeparture time distribution. In the background burstiness study the interarrival time of the CBR traffic and the load on the multiplex are $T = 20$ and $\rho = 0.8$, respectively. For the analysis of the influence of the background traffic burstiness the following cases have been considered:

- *Smooth background traffic*: 2 *Bernoulli* streams each with probability $p = 0.375$ of submitting a cell to the present time slot. Thereby the peakedness of the background traffic is 0.625.
- *Poisson background traffic* with intensity 0.75.
- *Bursty background traffic*: a *negative binomial* batch size distribution with peakedness 2 and mean 0.75.

Figs. 3 and 4 depict the effect of the burstiness of the background traffic and the effect of the multiplexer load on the shifted interdeparture time distribution, respectively.

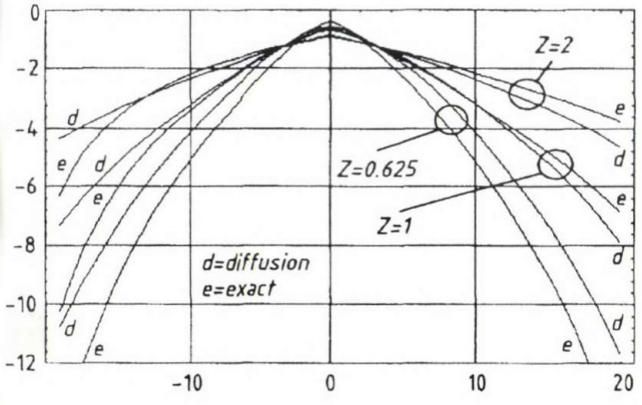


Fig. 3. Probability mass function of the shifted interdeparture time on a loglinear plot for background traffic with different burstiness ($n = 1, T = 20, \rho = 0.8$)

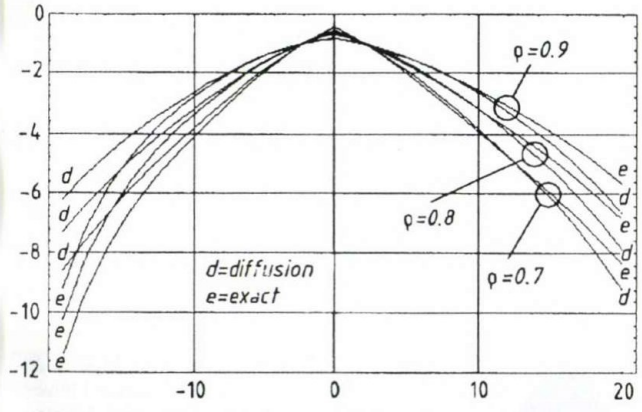


Fig. 4. Probability mass function of the shifted interdeparture time on a loglinear plot for different loads ($n = 1, T = 20, Z = 0.8$)

By comparing Figs. 3 and 4 it can be observed that the burstiness of the background traffic has a stronger influence on the shifted interdeparture time distribution than the load. Since it is the load which can be controlled or estimated, while the burstiness is highly unknown, this gives rise to practical difficulties in the dimensioning of the UPC parameters.

The diffusion approximation performs reasonable well, but for the smooth background traffic case the inaccuracy is slightly larger than for the other cases.

Finally, we compare the correlation obtained through the exact model with the results obtained by the diffusion model for lags less than 10 for two different values of the CBR peak rate in case of Poisson background traffic.

Table 1. The correlation as a function of the lag n, $T = 20$ and $\rho = 0.8$

lag n	1	2	3	4	5
exact corr.	-0.358	-0.0836	-0.0316	-0.0137	-0.0064
diff. corr.	-0.347	-0.0874	-0.0344	-0.0155	-0.0075
lag n	6	7	8	9	
exact corr.	-0.0031	-0.0016	-0.0008	-0.0004	
diff. corr.	-0.0038	-0.0020	-0.0011	-0.0006	

Table 2. The correlation as a function of the lag n, $T = 10$ and $\rho = 0.8$

lag n	1	2	3	4	5
exact corr.	-0.285	-0.0947	-0.0474	-0.0267	-0.0160
diff corr.	-0.273	-0.0964	-0.0496	-0.0286	-0.0176
lag n	6	7	8	9	
exact corr.	-0.0100	-0.0065	-0.0042	-0.0028	
diff corr	-0.0112	-0.0074	-0.0050	-0.0034	

As the examples indicate, then the diffusion approximation seem to have an overall acceptable accuracy for quick and effective computation of correlations in CDV perturbated CBR cell streams.

5. CONCLUSION

In this paper we have studied the Cell Delay Variation which a CBR cell stream is exposed to when it is multiplexed with a background cell stream in a FIFO queue. Two models, an exact Markovian model and an approximate diffusion model has been applied to obtain detailed information on the correlation structure of the CDV affected CBR stream after departure from the multiplexing queue.

Explicit computational methods for the correlations has been provided in the exact case while closed form expression are derived in the diffusion model which e.g. shows that the correlation is always negative and gives the exponential decay rate as the lag tends to infinity.

It has been found that the burstiness of the background traffic has a significant influence on CDV, and that the diffusion approximation is sufficient accurate for fast but reliable performance evaluation of quantities related to CDV.

6. ACKNOWLEDGEMENT

The research was performed as part of a joint project between Ellemtel Telecommunication System Laboratories, Sweden and the High Speed Networks Laboratory, Department of Telecommunications and Telematics, Technical University of Budapest, Hungary. The authors are grateful to Ellemtel and TeleDanmark for their support.

REFERENCES

[1] ATM Forum, *Traffic Control and Congestion Control*, Draft Baseline Document, April 1993.

[2] S. Blaabjerg: "Estimating the Effect of Cell Delay Variation by an Application of the Heavy Limit Theorem", COST 242 TD(92) 017.

[3] S. Blaabjerg: "Cell Delay Variation in a FIFO Queue: A Diffusion Approach", *Proc. of IFIP TC6 High Speed Networks and their Performance*, pp. 237-256, Raleigh, NC, USA, 26-28 October, 1993.

[4] S. Blaabjerg, S. Molnár: "Methods for UPC Dimensioning of a CDV perturbated cell stream", *RACE BRAVE Workshop*, Milano, Italy, June 14-15, 1995.

[5] C. Bisdikian, W. Matragi and K. Sohraby: "A Study of the Jitter in ATM Multiplexers", *Proc. of IFIP TC6 High Speed Networks and their Performance*, pp. 219-235, Raleigh, NC, USA, 26-28 October, 1993.

[6] P. Brown and A. Simonian: "Perturbation of a Periodic Flow in a Synchronous Server", *PERFORMANCE'87*, (eds. P. J. Courtois and G. Latouche), Elsevier Science Publishers, 1988.

[7] COST 224: *Performance Evaluation and Design of Multi-Service Networks*, (ed. J. Roberts), Final Report 1991.

[8] W. Feller: *"An Introduction to Probability Theory and Its Applications"*, Vol. I., Third edition, Wiley 1968.

[9] A. Gravey et al.: "Models for Setting ATM Parameter Values", *RACE ATM Expert Workshop*, 1995.

[10] F. Guillemin, P. Boyer, A. Dupuis and L. Romoef: "Peak Rate Enforcement in ATM Networks", *Proc. of IEEE INFOCOM'92*, Session 6A, Florence 1992.

[11] F. Guillemin and J. W. Roberts: "Jitter and Bandwidth Enforcement", *Proc. of IEEE GLOBECOM'91*, paper No. 9.3, Phoenix, AZ, USA, Dec. 2-5, 1991.

[12] F. Hübner: "Dimensioning of a Peak Cell Rate Monitor Algorithm Using Discrete-Time Analysis", *Proc. of ITC-14*, pp. 1415-1424, Antibes Juan-les-Pins, France, 6-10 June, 1994.

[13] COST 242: *Cell Delay Variation in ATM Networks*, (eds. A. Gravey and S. Blaabjerg), Interim Report, October 1994.

[14] ITU-T SG 13 Recommendation I. 371: *Traffic Control and Congestion Control in B-ISDN*.

[15] F. P. Kelly, and P. B. Key: "Dimensioning Playout Buffers for an ATM Network", 11th UK Teletraffic Symposium (Cambridge, 1994), IEE Savoy Place, London.

[16] L. Kleinrock: *"Queueing Systems"*, Vol. II., Wiley Interscience.

[17] W. Matragi, C. Bisdikian and K. Sohraby: "Jitter Calculus in ATM Networks: Single Node Case", *Proc. of INFOCOM'94*, pp. 232-241, June, 1994.

[18] S. Molnár, S. Blaabjerg and H. Christiansen: "On the Superposition of a Number of CDV Affected Cell Streams", *Proc. of International Conference on Local and Metropolitan Communication Systems LAN&MAN*, Kyoto, Japan, Dec. 7-9, 1994.

[19] S. Molnár and S. Blaabjerg: "Cell Delay Variation in an ATM Multiplex", submitted to the *European Transaction on Telecommunications and Related Technologies*.

[20] G. Niestegge: "The Leaky Bucket Policing Method in the ATM Network", *Int. Journal of Digital and Analogue Communications Systems*, Vol. 3, 1990.

Sándor Molnár graduated in 1991 at the Technical University of Budapest, Faculty of Electrical Engineering, in Telecommunications. Since September 1991 he has been working for his Ph.D. degree and from September 1995 he is a lecturer at the Department of Telecommunication and Telematics. His main interests include performance modeling and analysis of ATM networks and broadband telecommunications. Mr. Molnár is participating in the European project COST 242 on Methods for the Performance Evaluation and Design of Broadband Multiservice Networks.

Søren Blaabjerg received the M.Sc. degree in mathematics from Aarhus University in 1986. He was first with Electromagnetics Institute, Technical University of Denmark for three years in the RACE Atmospheric project. From 1990 to 1994 he worked for Ellemtel Telecommunication System Laboritories, Lund, Sweden and in 1994 he returned to the Technical University of Denmark to work at the Department of Telecommunications. His main interests include teletraffic analysis and performance evaluation of modern communication networks with special interest in B-ISDN.

QUEUEING DELAY OF AN INDIVIDUAL CELL STREAM IN AN ATM MULTIPLEXER: A GENERATING-FUNCTIONS APPROACH*

B. STEYAERT and H. BRUNEEL

UNIVERSITY OF GHENT, SMACS RESEARCH GROUP, LAB. FOR COMMUNICATIONS ENGINEERING
SINT-PIETERSNIEUWSTRAAT 41, B-9000 GENT, BELGIUM

In this paper, we present the delay analysis of a Tagged Cell Stream (TCS) in an ATM multiplexer buffer with one output link. The transmission of the tagged cells is sometimes prohibited due to the presence of background traffic, which represents a superposition of cell streams generated by various other sources, independently of the TCS arrivals. We first derive an expression for the probability generating function of the TCS delay. From this result, closed-form expressions for the mean TCS delay and the tail distribution of the TCS delay are obtained. In addition, lower and upper bounds that allow accurate estimates for these quantities and that are easily evaluated, are also given. Finally, some numerical results are considered that show the potential substantial impact of the TCS interarrival-time variance on the TCS delay quantiles.

1. SYSTEM DESCRIPTION

Let us consider a discrete-time single-server queueing model for an ATM multiplexer buffer. Time is slotted, and one slot suffices for the transmission of exactly one cell (i.e., one unit of information). Since the transmission of cells is synchronized to the slot boundaries, the transmission of any cell that arrives in an empty buffer will start at the beginning of the slot following its slot of arrival (and, consequently, terminate at the end of this slot); note that this is the so-called Please-Wait strategy, defined in [12]. We consider the buffer to have infinite storage capacity, i.e., no arriving cells are lost.

The Interarrival Time (IAT) between two successive tagged cells (say the n-th and $(n+1)$-th), i.e., the number of slots between their slots of arrival, is a random variable, in the sequel denoted by I_n. The I_n's are assumed to be i.i.d., and distributed according to the probability mass function $i(k)$, $k \geq 1$, with corresponding probability generating function $I(z)$, to be specified later on. For our purposes, it is sufficient that $I(z)$ is a rational function. Therefore, if we denote by $1/\alpha_i$, $1 \leq i \leq L$, the poles of $I(z)$, i which are assumed to have multiplicity 1, the partial fraction expansion of $I(z)$ can be written as

$$I(z) = \sum_{j=1}^{M} p(j) z^j + \sum_{i=1}^{L} \frac{C_i \alpha_i z}{1 - \alpha_i z} , \qquad (1)$$

for some constants $p(j)$ and C_i, all α_i's satisfying $|\alpha_i| < 1$.

Furthermore, there is some additional background traffic entering the system, and we consider the case that this Background Arrival Process (BAP) is described by the probability generating function $A(z)$ during any slot, i.e., the sequence of random variables describing the numbers of cell arrivals generated by the background traffic during consecutive slots are assumed to be i.i.d., and independent of the TCS arrivals. In many practical cases, $A(z)$ will be set equal to $\exp\{p(z-1)\}$, thus describing the Poisson arrival process with load p. In general, if we denote by p the load of the BAP, then the overall load of the queue, denoted by σ, is given by

$$\sigma = p + 1/I'(1) , \qquad (2)$$

(where primes denote derivatives with respect to the argument) and since at most one cell can be transmitted per slot, the equilibrium condition of the queue simply is $\sigma < 1$. All cells are served according to a FCFS (First-Come-First-Served) discipline, with the restriction that BAP cells that arrive during the same slot as a tagged cell are served first. Note that the latter assumption does not significantly affect the analysis presented here, and can be modified without great difficulty to the case where either the tagged cells are served first, or all cells arriving in the same slot are served in random order.

This work has been motivated by the need of creating efficient tools for calculating performance measures concerning the TCS delay, and in particular its tail distribution. For instance, in the case of a CBR tagged cell stream, it was shown in [1] that, in order to avoid excessive cell loss caused by the Virtual Scheduling Algorithm in the UPC function monitoring a well-behaving cell flow, a remote quantile (say the 1E-10 quantile) of the TCS delay distribution can be used as an estimate for the Cell Delay Variation τ , which is part of the traffic contract at the UNI. Consequently, in the following analysis, we devote some attention to finding accurate expressions for the TCS cell delay distribution that are both accurate and easy to calculate.

Queueing models where the overall arrival process is a superposition of two distinct streams has received some attention in the literature (e.g., [2]−[6], and the references therein), mainly for the purpose of studying the impact of the perturbing background traffic on the queueing behavior and/or the delay characteristics of a tagged arrival stream. For instance, in the context of policing in ATM networks, substantial effort has been devoted to capturing the jitter caused by background traffic to a CBR source passing through a multiplexing stage [2]. The case where the TCS generates periodic cell arrivals has been studied, either by a discrete-time [3] or a continuous-time [4] queueing model, while the continuous-time $GI + M/M/1$ queue has been studied in [5]. A slightly different model than the one described above has been analyzed in [6], leading to an expression for the probability generating function of the TCS delay, in terms of $M + L - 1$ unknowns (M and L being defined in expression (1)), that must be calculated by solving a set of linear equations. Using a totally different methodology, we manage to derive an expression for the probability generating function of the TCS delay, that does no longer require the calculation of $M + L - 1$ unknowns, as is shown in Section 3. In addition, closed-form expressions for the mean and the tail distribution of the delay, together with simple upper and lower bounds for these quantities, are also given.

2. SYSTEM EQUATIONS

Let us consider two arbitrary successive TCS arrivals (say the n-th and $(n+1)$-th). We denote by d_n (with probability generating function $D_n(z)$) the delay of the n-th tagged cell entering the buffer, and by a_k (with probability generating function $A(z)$) the number of BAP cell arrivals during the k-th slot after the n-th tagged cell arrival. As already mentioned, I_n denotes the number of slots between the arrival instants of the two tagged cells. Note that due to the arrival model described in the previous section, this random variable is independent of the value of d_n. We must consider the two following cases:

2.1. $d_n \geq I_n$

Considering this situation which is depicted in Fig. 1a, then, due to the FCFS queueing discipline, all the BAP cells that have arrived in the system between the n-th and $(n+1)$-th TCS arrival are still present in the system at the arrival instant of the $(n+1)$-th tagged cell. Since these BAP cells will be served before the $(n+1)$-th tagged cell, the relation between d_n and d_{n+1} is quite simple, and given by

$$d_{n+1} = d_n - I_n + 1 + \sum_{i=1}^{I_n} a_i . \qquad (3)$$

* This paper is based on a contribution presented at Globecom'94 (San Francisco, Ca.).

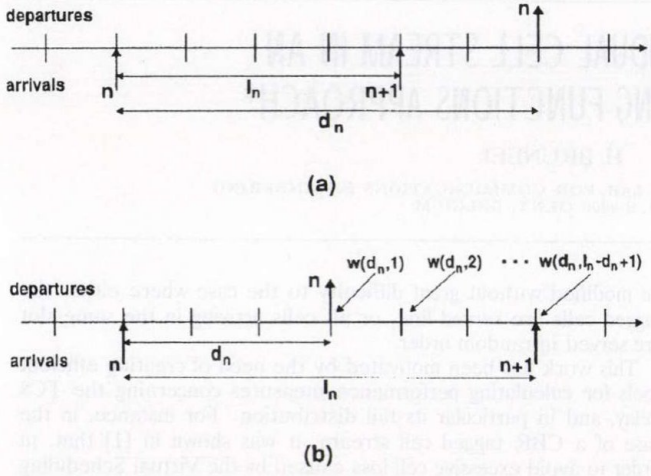


Fig. 1. Relation between d_n and d_{n+1}, for $d_n \geq I_n$ (a) and $d_n < I_n$ (b)

2.2. $d_n < I_n$

The analysis of this situation is a bit more complicated, as can be seen from Fig. 1b, since some of the BAP cells that have arrived in the system between the n-th and the $(n+1)$-th TCS arrival have already been transmitted at the moment of arrival of the $(n+1)$-th tagged cell. Therefore, when d_n and I_n take the values l and k respectively ($l < k$), let us define $w(l,m)$, $1 \leq m \leq k-l+1$ (with probability generating function $W_{l,m}(z)$), as the random variable describing the number of BAP packets in the system at the beginning of the m-th slot following the departure of the n-th tagged cell. Since the queueing discipline is FCFS, $w(l,m)$ consists of all BAP cells that have arrived in the system *after* the n-th tagged cell and *before* the m-th slot following its slot of departure, and that have not yet been transmitted at the beginning of the latter slot. The $w(l,m)$'s are related by

$$w(l,1) = \sum_{i=1}^{l} a_i \tag{4.a}$$

$$w(l,m+1) = (w(l,m)-1)^+ + a_{m+l}, \quad 1 \leq m \leq k-l, \tag{4.b}$$

where $(.)^+ \triangle \max(.,0)$. Taking into account the definition of these random variables, the delay of the $(n+1)$-th tagged cell is then simply given by

$$d_{n+1} = w(l, k-l+1)+1, \quad \text{if} \quad d_n = l \quad \text{and} \quad I_n = k. \tag{5}$$

Eqs. (3)–(5) fully describe the delay behavior of the TCS. From these equations, a closed-form expression for the probability generating function of the TCS delay can be derived.

3. PROBABILITY GENERATING FUNCTION OF THE TCS DELAY

From (3), together with the property that the delay of a tagged cell is independent of the number of BAP-arrivals during any slot following its arrival, it can be shown that

$$\mathrm{E}\left(z^{d_{n+1}} | d_n \geq I_n\right) \mathrm{Prob}[d_n \geq I_n] = zI\left(\frac{A(z)}{z}\right) D_n(z) - \\ -z\sum_{l=1}^{\infty}\sum_{k=l+1}^{\infty} z^l \left(\frac{A(z)}{z}\right)^k d_n(l) i(k), \tag{6}$$

where E[.] denotes the expected value of the quantity between brackets, and where we have defined $d_n(l)$ as $\mathrm{Prob}[d_n = l]$.

On the other hand, from (4.a,b), we obtain

$$W_{l,m+1}(z) = z^{-m} A(z)^{m+1} + \\ +(z-1)\sum_{j=1}^{m}\left(\frac{A(z)}{z}\right)^j W_{l,m+1-j}(0), \tag{7}$$

(for all m, $0 \leq m \leq k-l$) where the values of the constants $W_{l,m+1-j}(0)$ could be calculated, depending on the expression for $A(z)$. However, in the following, we will prove that the knowledge of the exact values of these constants is not really required in order to derive numerical results from the probability generating function of the TCS delay. Combining the z-transform of equation (5) with the previous result, and defining $w_{l,m}$ as $W_{l,m}(0)$, then yields

$$\mathrm{E}\left(z^{d_{n+1}} | d_n < I_n\right) \mathrm{Prob}[d_n < I_n] = \\ = z\sum_{l=1}^{\infty}\sum_{k=l+1}^{\infty}\left(z^{l-k} A(z)^k = \right. \tag{8}$$

$$\left. +(z-1)\sum_{j=1}^{k-l}\left(\frac{A(z)}{z}\right)^j w_{l,k-l-j+1}\right) d_n(l) i(k).$$

Summation of (6) and (8), together with the property that $I(z)$ is a rational function, and, therefore, can be written in the form given by expression (1), then leads to the following expression for $D_{n+1}(z)$:

$$D_{n+1}(z) = zI\left(\frac{A(z)}{z}\right) D_n(z) + \\ +z(z-1)\left(\sum_{j=1}^{M-1}\left(\frac{A(z)}{z}\right)^j \sum_{l=1}^{M-j}\sum_{k=j+l}^{M} w_{l,k-l-j+1} d_n(l) p(k) + \right. \\ \left. +\sum_{i=1}^{L}\frac{C_i\alpha_i}{z/A(z)-\alpha_i}\sum_{l=1}^{\infty}\sum_{k=0}^{\infty} w_{l,k+1} d_n(l)\alpha_i^{k+l}\right). \tag{9}$$

Assuming that the system reaches its steady-state after a sufficiently large period of time (i.e., $n \to \infty$), the distributions of d_n and d_{n+1} (and, consequently, the corresponding probability generating functions) become identical. Defining $D(z)$ as the steady-state probability generating function of the delay of an arbitrary tagged cell, and

$$Q_j \triangle \lim_{n\to\infty}\sum_{l=1}^{M-j}\sum_{k=j+l}^{M} w_{l,k-l-j+1} d_n(l) p(k), \quad 1 \leq j \leq M-1$$

$$S_i \triangle \lim_{n\to\infty}\sum_{l=1}^{\infty}\sum_{k=0}^{\infty} w_{l,k+1} d_n(l)\alpha_i^{k+l}, \quad 1 \leq i \leq L$$

$$P(z) \triangle \prod_{i=1}^{L}\frac{z/A(z)-\alpha_i}{1-\alpha_i}; \quad P_i(z) \triangle P(z)\frac{1-\alpha_i}{z/A(z)-\alpha_i}, \quad 1 \leq i \leq L$$

$$T(z) \triangle \left(\frac{z}{A(z)}\right)^{M-1}\sum_{i=1}^{L}\frac{C_i S_i \alpha_i}{1-\alpha_i} P_i(z) + \\ +P(z)\sum_{j=1}^{M-1}\left(\frac{z}{A(z)}\right)^{M-j-1} Q_j, \tag{10}$$

then, after some manipulations, we find the following expression for $D(z)$:

$$D(z) = \frac{z(z-1)A(z)^{M+L-1} T(z)}{P(z)A(z)^L z^{M-1}(1-zI)(\frac{A(z)}{z}))}. \tag{11}$$

Expression (11) is a closed-form formula for the probability generating function of the steady-state TCS delay, in terms of the $M+L-1$ unknowns S_i, $1 \leq i \leq L$ and Q_j, $1 \leq j \leq M-1$, that occur in the expression for $T(z)$. These unknowns can be calculated using the requirement that $D(z)$ must be analytic inside the unit disk. Indeed, based on the properties of the involved functions, and some simple results from complex analysis (such as Rouchés theorem; for a more complete reasoning in a different but similar case, see, for instance, Chapter 3 of [13]),

it can be shown that the denominator in the right hand side of expression (11), in the sequel denoted by $\Psi(z)$, has $M+L-1$ zeros inside the complex unit disk (which, in general, are distinct, each zero having multiplicity 1). Also, due to $\Psi(1)=0$ and $\Psi'(1)=(1-\sigma)/(\sigma-p)>1$, $z=1$ is included in this set of $M+L-1$ zeros 'inside' the unit disk with multiplicity equal to 1.

In the rest of the paper, we will denote by z_j, $1\leq j\leq M+L-2$, the zeros of the denominator in the right hand side of (11) inside the unit disk and different from $z=1$. Then the condition that $D(z)$ must be analytic inside the unit disk (and hence, $D(z_j)$, $1\leq j\leq M+L-2$, finite) together with the normalization condition $D(1)=1$, leads to a set of $M+L-1$ linear equations for the $M+L-1$ unknowns S_i, $1\leq i\leq L$ and Q_j, $1\leq j\leq M-1$, which has one unique solution, and therefore completely determines expression (11) for $D(z)$. Nevertheless, the explicit calculation of the $(M+L-1)$ unknowns is not really required here. Indeed, from a similar derivation as in [7], it can be shown that $T(z)$, being a polynomial of degree $M+L-2$ in $z/A(z)$, can be written as

$$T(z)=\frac{1-\sigma}{\sigma-p}\prod_{j=1}^{M+L-2}\frac{z/A(z)-z_j/A(z_j)}{1-z_j/A(z_j)}\,. \qquad (12)$$

Combination of (11) and (12) then finally leads to an expression for the probability generating function of the TCS cell delay, from which all the significant performance measures concerning the delay can be derived, once the $M+L-2$ zeros inside the unit disk of the denominator in the right-hand side of (11) have been obtained. In the following section, we will derive an expression for the mean delay, as well as a geometric tail approximation of the delay distribution. In addition, sufficiently close bounds for these two quantities that do not require the calculation of the $(M+L-2)$ zeros, are also established.

4. DELAY CHARACTERISTICS

4.1. Mean TCS Delay

Taking the first derivative for $z=1$ of the expression for $D(z)$, we find the following formula for the mean TCS delay

$$E[d]=p+\frac{(\sigma-p)(1-p)^2I''(1)+A''(1)}{2(1-\sigma)}+ \qquad (13)$$

$$+\frac{1-p}{2}\left(2-M-\sum_{j=1}^{M+L-2}\frac{(A(z_j)+z_j}{A(z_j)-z_j}-\sum_{i=1}^{L}\frac{1+\alpha_i}{1-\alpha_i}\right).$$

In addition, we can derive upper and lower bounds for the sum in the right-hand side that contains the zeros z_j, $1\leq j\leq M+L-2$, thus avoiding the most time-consuming part of the numerical calculations. Indeed, first of all, in a similar way as in [7], it can be shown that

$$\sum_{j=1}^{M+L-2}\frac{A(z_j)+z_j}{A(z_j)-z_j}>0\,. \qquad (14)$$

On the other hand, from the explicit expression for $T(z)$, the following upper bounds can be established after some calculations:

$$\sum_{j=1}^{M+L-2}\frac{A(z_j)+z_j}{A(z_j)-z_j}<\sum_{i=1}^{L}\frac{1+\alpha_i}{1-\alpha_i}+M-2$$

$$\sum_{j=1}^{M+L-2}\frac{A(z_j)+z_j}{A(z_j)-z_j}<\sum_{i=1}^{L}\frac{1+\alpha_i}{1-\alpha_i}+$$

$$+(M-2)\left(\frac{\sigma-p}{1-\sigma}(M-1)A(0)-1\right). \qquad (15)$$

The upper bound we will use for the sum in the left hand side is the minimum of the right hand sides of the two above expressions. As will be shown by some numerical examples, the upper and lower bounds for $\mathbf{E}[d]$ obtained by combining (14) and (15) are sufficiently close to allow a reasonably good estimate for the mean TCS delay, without requiring the calculation of the z_j's.

4.2. Tail Distribution of the TCS Delay

As has already been indicated in various papers, for a wide variety of queueing models, the tails of the buffer-occupancy distribution as well as the cell-delay distribution can be approximated very accurately by geometric forms (see e.g. [5], [8]–[11], [13]), implying in our specific case that

$$\text{Prob}[d>D]=\frac{-Cz_0^{-D-1}}{z_0-1}\,, \qquad (16)$$

where z is the pole of $D(z)$ with the smallest modulus (i.e., the zero of the denominator in the right hand side of (11) outside the unit disk with the smallest modulus), which is a real and positive quantity larger than 1. The constant C in the above expression can be calculated from the residue theorem, and is given by

$$C=\frac{1-\sigma}{\sigma-p}\frac{z_0(z_0-1)A(z_0)}{\Psi'(z_0)}\prod_{j=1}^{M+L-2}\frac{z_0-z_jA(z_0)/A(z_j)}{1-z_j/A(z_j)}\,, \qquad (17)$$

where $\Psi(z)$ represents the denominator in the right hand side of (11). Again, a close upper bound for C avoiding the calculation of the z_j's can be derived. Indeed, from the expression for $T(z)$ and using the property that z_0 is a real and positive quantity satisfying $z_0>1$, we can derive that

$$\prod_{j=1}^{M+L-2}\frac{z_0-z_jA(z_0)/A(z_j)}{1-z_j/A(z_j)}<A(z_0)^LP(z_0)z_0^{M-2}. \qquad (18)$$

Using equations (16)–(18), an upper bound for the geometric tail approximation of the TCS delay distribution is obtained, which is easily evaluated, since it merely requires the calculation of z_0. As will be shown in the next section, this upper bound is extremely close.

5. NUMERICAL EXAMPLES

Let us first concentrate on the case where the TCS generates a periodic cell stream that is perturbed by a Poisson background arrival process, a model that is of particular interest in many studies concerning the behavior of CBR traffic through an ATM network. This means that the generating functions $I(z)$ and $A(z)$ are given by

$$I(z)=z^M$$
$$A(z)=\exp\{p(z-1)\}. \qquad (19)$$

In Fig. 2, we have plotted the mean TCS delay (full line) and the corresponding upper and lower bounds (dotted lines) versus the overall load σ, for $M=8$. As one can observe, the lower bound is quite close to the actual mean delay, while the upper bound is fairly accurate for low and high values of σ; for intermediate values of σ , some improvement is required here.

In Figs. 3 and 4, for $M=4, 8, 32$ and ∞, we have plotted the probability that the delay exceeds a given integer value D, Prob$[d>D]$, versus D (full lines), together with the corresponding upper bounds (dashed lines) for this quantity, for the cases $\sigma=0.5$ and $\sigma=0.9$ respectively. These figures show that the upper bound derived in Section 4.2 is actually very tight.

Curves as the ones displayed in Figs. 3 and 4 can be used to quantify the 'cell delay variation' (CDV) or the 'delay jitter' incurred by CBR cells in ATM queues (e.g. in multiplexers or switching stages). Note that, for a given overall load σ, increasing values of M indicate increasing proportions of background traffic, which explains why longer tails are observed for the delay distribution for higher M (owing to the fact that CBR traffic is more regular than Poisson traffic and hence causes less congestion). For $M=\infty$, the results plotted here correspond to the delay of an arbitrary cell in an $M/D/1$ queue with load σ, and we observe that for increasing values of M, the TCS delay rapidly converges to the latter limit.

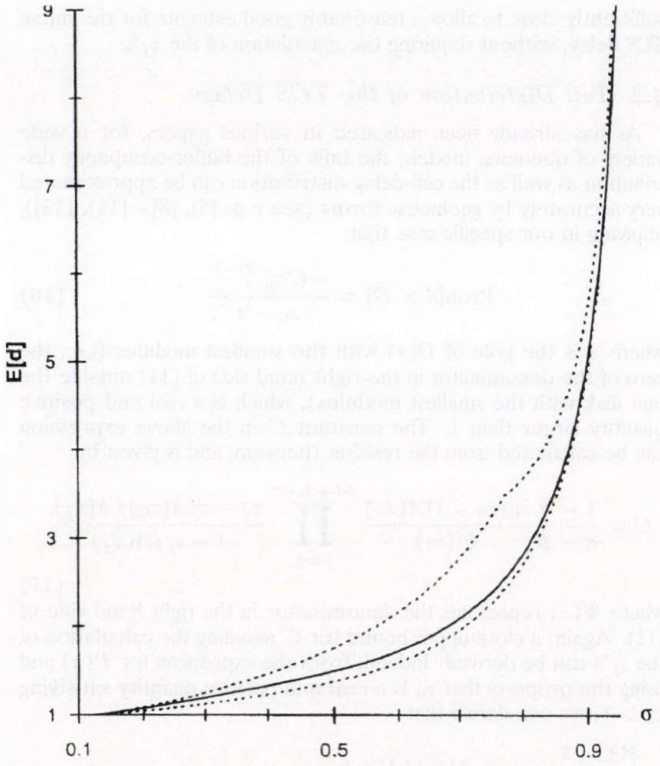


Fig. 2. $E[d]$ with lower and upper bound versus σ

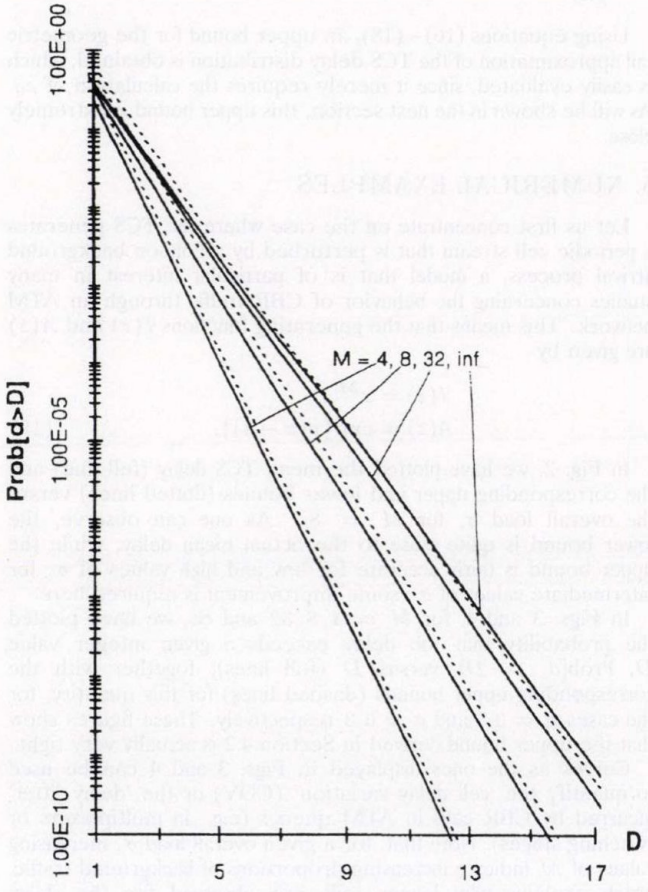


Fig. 3. $Prob[d > D]$ with upper bound versus D, $\sigma = 0.5$, $M = 4, 8, 32, \infty$

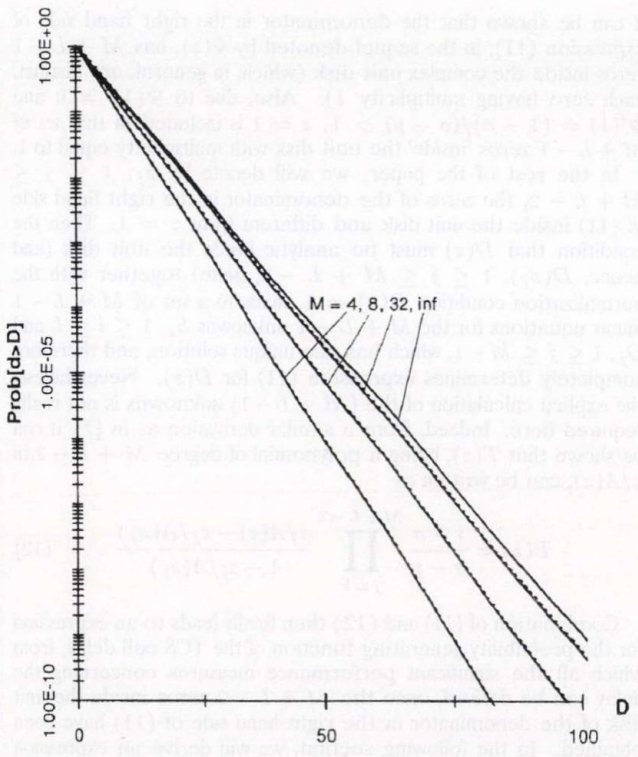


Fig. 4. $Prob[d > D]$ with upper bound versus D, $\sigma = 0.9$, $M = 4, 8, 32, \infty$

From the previous results, one might be tempted to conclude that the delay of a cell in an $M/D/1$ queue can serve as a good upper bound for estimating the TCS delay. This is true for a CBR source (possibly with some jitter), *but is no longer valid as the 'variability' of the interarrival times becomes increasingly high*. This is shown in Figs. 5 and 6. The distribution of the interarrival times is here chosen to be either a pure geometric distribution, or a weighted average of two geometric distributions. This means that $I(z)$ can be written as

$$I(z) = q\frac{(1-\alpha_1)z}{1-\alpha_1 z} + (1-q)\frac{(1-\alpha_2)z}{1-\alpha_2 z}\,, \qquad (20)$$

where $q = 1$ corresponds to a pure geometric distribution. This allows us to investigate the impact of the variance of the interarrival times on the TCS delay. For that purpose, we define R as the fraction of the variance of the IAT's versus the variance in the case of purely geometrically distributed IAT's with the same mean value (in the remainder also denoted by M). $R = 0$ thus corresponds to the case of a CBR tagged cell stream, and $R = 1$ represents the case of geometrically distributed IAT's; whenever $R > 1$, $I(z)$ is given by (20), with mean M and variance $RM(M-1)$. Choosing (α_1, α_2, q) such that the requested mean and variance are matched, this leaves us one additional parameter that can be arbitrarily fixed. Therefore we set $\alpha_2 = 0$, leading to

$$q = \frac{2(M-1)}{2(M-1)+M(R-1)}\,,$$

$$\frac{1}{1-\alpha_1} = M + \sqrt{\frac{1-q}{2q}(R-1)M(M-1)}\,. \qquad (21)$$

In Figs. 5 and 6, we have plotted the results for $M = 16$, $R = 0, 1, 5, 10$, and $\sigma = 0.5$ and 0.9 respectively, considering a Poisson background arrival process. It is readily observed when the variance remains relatively small (i.e., $R \leq 1$), the

that TCS delay is rather insensitive to these fluctuations in the arrival process, and the above mentioned $M/D/1$-limit for the TCS delay can be applied. But as the variability in the tagged arrival stream increases, the TCS delay can become dramatically high, and the $M/D/1$-results can no longer be used.

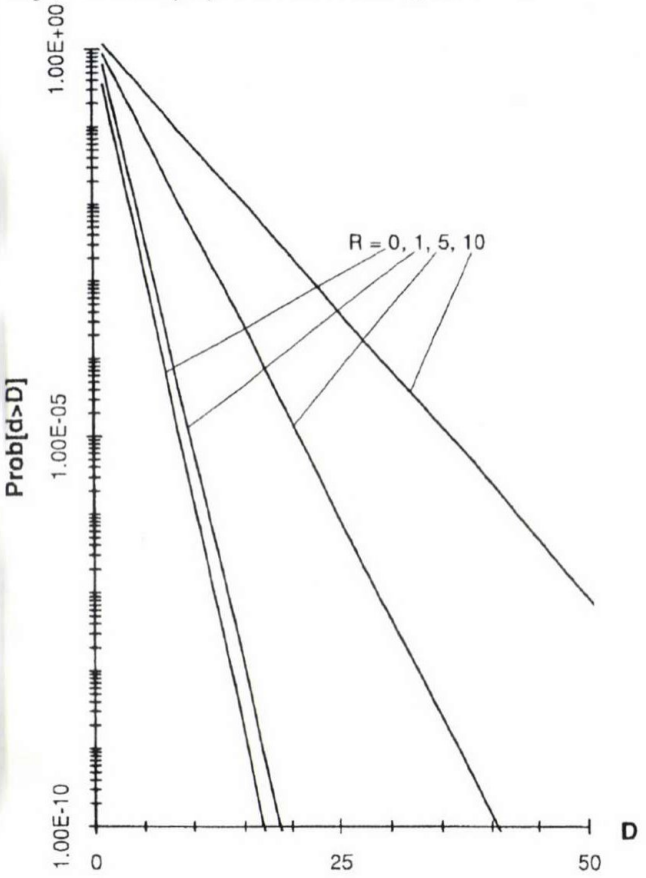


Fig. 5. Prob$[d > D]$ versus D, $\sigma = 0.5$, $R = 0, 1, 5, 10$

6. CONCLUSIONS

In this paper, we have proposed an alternative method for solving the problem of deriving the delay characteristics of a tagged cell stream perturbed by background traffic. This has led to closed-form expressions for the probability generating function, the mean and a geometric tail approximation for the TCS delay, avoiding the solution of a possibly large set of linear equations, as was suggested in previous papers. In addition, simple upper and lower bounds for the latter two performance measures that are sufficiently accurate, have also been derived. Based on some numerical examples, we conclude that fluctuations in the tagged arrival stream will have a substantial impact on the incurred delays, even when the TCS load is relatively low in comparison with the total load.

7. ACKNOWLEDGEMENT

The authors wish to acknowledge the support of Alcatel Bell Telephone Mfg. Co. (Antwerp, Belgium) and the Belgian National Fund for Scientific Research (NFWO).

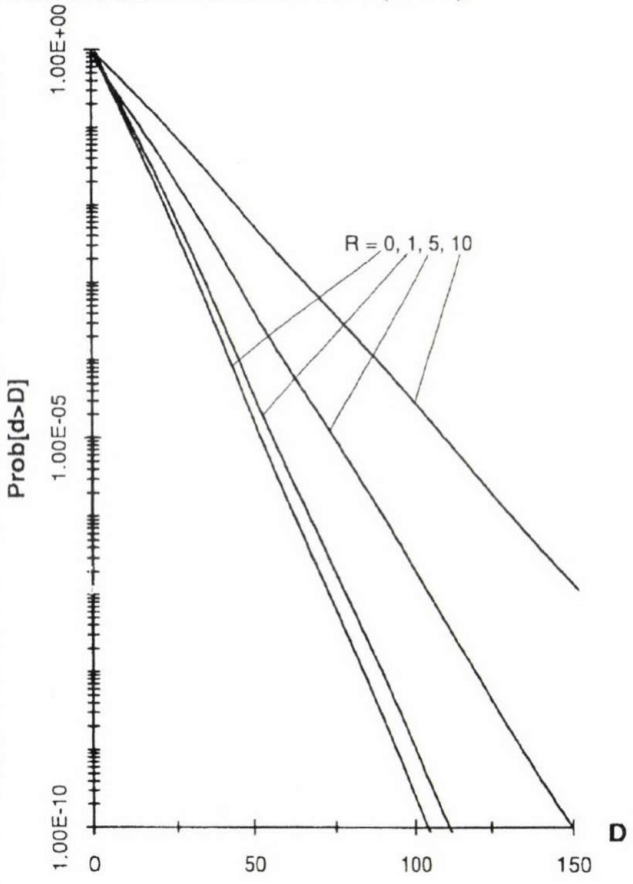


Fig. 6. Prob$[d > D]$ versus D, $\sigma = 0.9$, $R = 0, 1, 5, 10$

REFERENCES

[1] R. Wilts, J. Witters and G. H. Petit: "Throughput analysis of a usage parameter control function monitoring misbehaving constant bit rate sources", *Proceedings of the Second Workshop on Performance Modelling and Evaluation of ATM Networks* (July 1994, Bradford UK), pp. 39.1-39.17.

[2] F. Guillemin and J. W. Roberts: "Jitter and bandwidth enforcement", *Proceedings GLOBECOM'91* (December 1991, Phoenix Az.), pp. 261-265.

[3] P. Brown and S. Simonian: "Perturbation of a periodic flow in a synchronous server", *Proceedings of the 12th IFIP WG7.3 International Symposium on Computer Performance Modeling, Measurement and Evaluation* (December 1987, Brussels), pp. 89-112.

[4] J. W. Roberts: Final Report COST 224: Performance Evaluation and Design of Multiservice Networks (Commission of the European Communities, Luxembourg, 1992).

[5] A. Kuczura: "Queues with mixed renewal and Poisson inputs", *Bell Syst. Tech. Jour.*, Vol. 51.2 (1972), pp. 1305-1326.

[6] M. Murata, Y. Oie, T. Suda and H. Miyahara: "Analysis of a discrete-time single-server queue with bursty inputs for traffic control in ATM networks", *IEEE Jour. on Sel. Areas in Commun.*, Vol. 8 (1990), pp. 447-458.

[7] I. Rubin and Z. Zhang: "Message delay analysis for TDMA schemes using contiguous-slot assignments", *IEEE Trans. on Commun.*, Vol. 40 (1992), pp. 730-737.

[8] K. Sohraby: "On the asymptotic behavior of heterogeneous statistical multiplexer with applications", *Proceedings INFOCOM'92* (April 1992, Firenze), pp. 839-847.

[9] C. M. Woodside and E. D. S. Ho: "Engineering calculation of overflow probabilities in buffers with Markov interrupted service", *IEEE Trans. on Commun.*, Vol. 35 (1987), pp. 1272-1277.

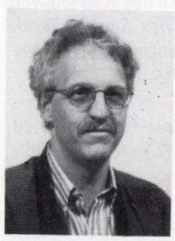
[10] E. Desmet, B. Steyaert, H. Bruneel and G. H. Petit: "Tail distributions of queue length and delay in discrete-time multiserver queueing models, applicable in ATM networks", *Proceedings ITC'13* (June 1991, Copenhagen), Vol. Queueing, Performance and Control in ATM, pp. 1-6.

[11] Y. Xiong and H. Bruneel: "Buffer contents and delay for statistical multiplexers with fixed-length packet train arrivals", *Perf. Eval.*, Vol. 17 (1993), pp. 31-42.

[12] J. F. Hayes: *Modeling and Analysis of Computer Communication Networks* (Plenum Press, New York, 1984).

[13] H. Bruneel and B. G. Kim: *Discrete-Time Models for Communication Systems Including ATM* (Kluwer Academic Publishers, 1993).

Bart Steyaert received the degrees of Licentiate in Physics and Licentiate in Computer Science in 1987 and 1989, respectively, all from the University of Ghent, Gent, Belgium. Since 1990, he has been working as a Ph.D. student at the Laboratory for Communications Engineering, University of Ghent, as a member of the SMACS Research Group. His main research interests include performance evaluation of discrete-time queueing models, applicable in B-ISDN networks.



Herwig Bruneel received the M.S. degree in Electrical Engineering, the degree of Licentiate in Computer Science, and the Ph.D. degree in Computer Science in 1978, 1979 and 1984, respectively, all from the University of Ghent, Belgium. Since 1979, he has been working as a researcher for the Belgian National Fund for Scientific Research (N.F.W.O.) at the University of Ghent, where he currently leads the SMACS Research Group of the Laboratory for Communications Engineering. He is also a Professor in the Faculty of Applied Sciences at the same university. His main research interests include stochastic modeling of digital communications sytems, discrete-time queueing theory, and the study of ARQ protocols. He has published more than 90 papers on these subjects and is coauthor of the book *H. Bruneel and B. G. Kim: Discrete-Time Models for Communication Systems Including ATM*" (Kluwer Academic Publishers, Boston, 1993).

IBM'S ATM SWTICHING TECHNOLOGY

1. INTRODUCTION

IBM's ATM strategy is focused on a total system solution. IBM is striving to develop cost-effective, leading-edge technologies to offer the best ATM product range available. IBM provides ATM solutions from the desktop, campus, and wide area environment. It offers ATM solution that let the customer keep its existing wiring and application while helping the networks to evolve and grow. IBM is helping and promoting ATM standards by participating in, and contributing to, the ATM Forum and other standard groups. Now the company provides a new, high-speed, ATM-complementary communications architecture that delivers functional and economic advantages. IBM has developed a new architecture for high-speed networking called **Networking Broad Band Services (NBBS)**. This is not "IBM's version of ATM". NBBS conforms to the public ATM standards and enhances them by specifying many aspects of ATM networking that, so far, remain outside the scope of the public standards. That is NBBS defines a working network by using the capabilities of ATM and adding the intelligence and manatgement function to make it work extremly well IBM is working in two fields:

- ATM products for the local area environment;
- ATM product for wide area environment.

2. SWITCH-ON-A-CHIP

The basis for IBM's campus and wide area ATM switch products is a chip made by VLSI technology and called Switch-on-a-chip or **PRIZMA**. It was developed at the IBM Research Laboratory in Zurich, Switzerland.

Characteristics of the Prizma switch include:

- 16 input ports;
- 16 output ports;
- 400 Mbit/s per port;
- Built in support for modular growth in number of ports;
- Built in support for modular growth in port speed;
- Built in support for modular growth in aggregate throughput;
- Built in support for automatic load sharing;
- Self-routing switch element;
- Built in multicast and broadcast;
- Aggregate date rate 6.4 Gbit/s per module;
- 2.4 Million transistor on 15x15 mm chip;
- 472 I/O pins.

The maximum internal cell length 64 bytes. As it can be seen a unique feature of the switching element is its scaleability. Conventional switches typically use a single or dual port memory (Fig. 1).

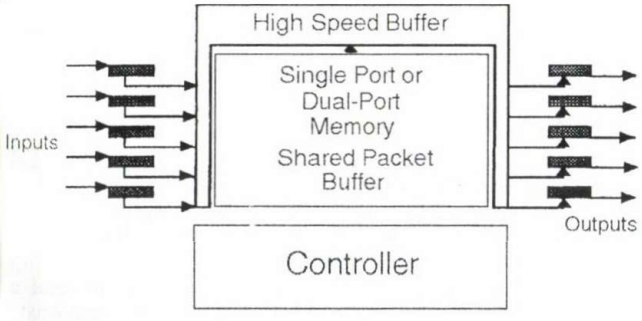


Fig. 1. Conventional shared memory implementation

Internally a time division multiplex technique is used to obtain access to the shared packet buffer. The number of ports, the number memories and the port speed determine the required width of internal high-speed bus: if the port speed is **V** and the number of ports is **2N** (in and out) the memory cycle is **Z** the required width is **2NVZ**. As the bit rate and/or the number of ports is increased, the bandwidth required increases rapidly: A switch with 16 ports, 1 Gb per port requires a memory cycle time of 13.25 nsec and an internal width of the bus of 425 bits (this is the size of an ATM cell $53 \times 8 = 425$). For ATM application, there is no relief in increasing the bus-width: there are simply no more bits in a cell which can make use of the additional bus wires.

Consequently, any increase in speed and/or ports has a dramatic effect on the required memory cycles time. The architecture of IBM's prizma switch is based on the separation of the control and data flow. The bulk of the data packet is fed through the data section only. It consists of the global shared packet buffer and parallel I/O routing trees in order to avoid the bottleneck of a shared medium for the data stream. The data section is controlled by a control section such that the required switching function is achieved (Fig. 2).

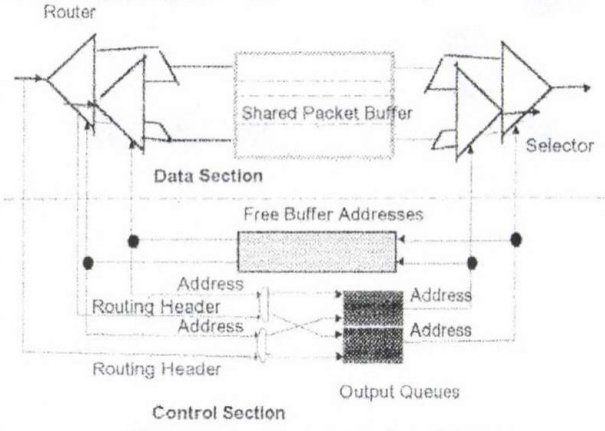


Fig. 2. Basic switch-on-a-chip architecture

The control section receives a pointer and copy of the first byte of an incoming packet: this byte is used to route this pointer into the appropriate output queue. After this the control section deletes the copy of the first byte. This mechanism works with short units: in the order of one or two bytes. This allows for a fully parallel implementation in the control section, thereby reducing the time needed to actually perform the routing function to approximately 12—15 clock cycles for a 16 by 16 switch element.

Separation of control and data completely removes any interface and both sections can be optimized for the function they carry out. As there is no multiplexing in the data section the cycle time of the internals of the data section is equal to the system-level cycle-time (e.g., 50 MHz for a byte-wide data path implementing a 400 Mbit/s switch).

2.1. Data Section

The data section contains the shared packet buffer which is built from a set of shift — registers, in which the packet is serially shifted in — and out when it arrives and leaves the chip, respectively. Each shift register holds exactly one packet. At the input side up to 16 incoming packet can be routed simultaneously into free shift registers via 16 routes, one for each input. An empty shift register address for each of this routers is provided in advance by control section. This shift register address is only renewed when an incoming packet on respective input has "consumed" it. At the output side up to 16 outgoing packet can be transmitted simultaneously from selected shift registers via 16 selectors, one for each output.

2.2. Control Section

The control section consists of 17 control queues, namely one output queue per switch output and a single free queue. From the latter, addresses pointing to empty shift registers are de-queued and fed to the input routers in the data section. As soon as a new packet has been received, the shift register address from that input route is entered into the output queue indicated by the routing-tag in the packet header and a new empty shift register address from the free queue is fetched. The output queues behave as FIFO's and contain only shift register addresses of packets ready for transmission via the corresponding output port. The addresses are sequentially de-queued and fed to output selectors in the data section. After a successful packet transmission the addresses are returned to the free queue.

2.3. Scalability

Increasing the number of ports-port expansion

Switches with a larger number of ports than the basic switch module can be realized by connecting several Switch-on-a-chip modules in parallel for a single-stage, or cascading them for a multi-stage system. The switch has built-in logic to allow address-filtering at the input and activation of an output for supporting single-statge expansion. For a multi-stage expansion every stage requires a different routing-tag in general (Fig. 3.).

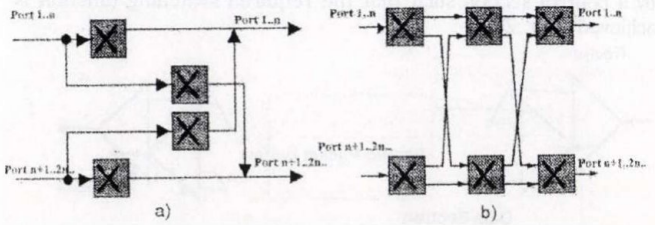


Fig. 3. Switch-on-a-chip port expansion
a) single-stage port expansion; b) multi-stage port expansion

The chip has a built in look-up table which allows routing-header bytes of the arriving packets to be shuffled. This allows multi-stage routing without customizing the individual stages. While multi-stage networks grow according to a logarithmic law, single chips Switch-on-a-chip is a greate advantage because it requires significant less chips for larger switches. (32 port single-stage switch requires 4 Switch-on-a-chip chips compared to a 64 chips if there would be only 4 input and output ports.)

Port speed expansion

To expand the actual speed of the switch ports can be reached by using multiple Switch-on-a-chip parallel: instead of 8 bit wide port, the switch ports become then 16, or more bits wide, and the doubling, tripling etc., of the port speed is achieved (Fig. 4).

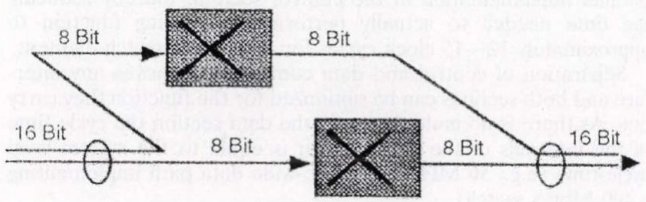


Fig. 4. Speed expansion

The chip has a built in hardware support to built such system easily. Assuming a 400 Mbit/s port speed for a single switch module, the paralleling of only two modules would be sufficient to build a 622 Mbit/s switch.

Performance expansion — increasing the aggregate throughput

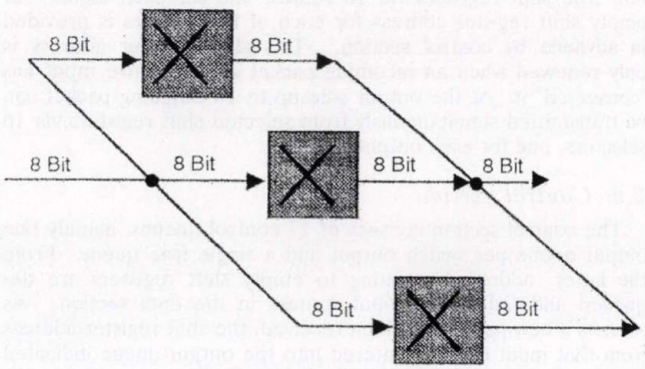


Fig. 5. Performance expansion

The aggregate throughput of a packet switch is given by the product of all ports and the port speed however the result have to be multiplied by a factor which is less than 1 because the limited number of output buffers. This factor is depend on the internal buffer memory and the traffic characteristics more bursty traffic will reduce the factor, while more internal buffer will increase the factor. The Switch-on-a-chip has the hardware control to allow cascading the internal buffer memory of multiple chips, such that the system behaves as if it were one chip with increased buffer memory. Control signal between the Switch-on-a-chip module guarantee proper packet sequence (Fig. 5).

Automatic load sharing — link paralleling

At the system level, it is often required to support access to a high-speed backbone from multiple lower-speed links. A typical example is to access a 622 Mbit/s ATM link to carry the non-local traffic of multiple (more then 4) Mbit/s ATM access links (Fig. 6).

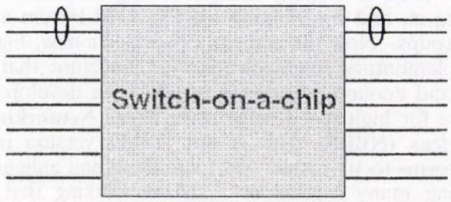


Fig. 6. Link paralleling

The usual solution this problem is that multiple lower speed ports are multiplexed together into the 622 Mbit/s link. This, however requires careful bandwidth management, and rearranging streams when large bandwidth requests must be accommodated. The Switch-on-a-chip has a built in feature, dubbed link paralleling, which manages the bandwidth on such links fully with hardware. 2 or 4 physical ports can be combined to support a double or quadruple speed link, without software to control which connection is allocated to a physical Switch-on-a-chip output port.

The four mentioned expansion methods can be combined freely to design a switch fabric. The port and performance expansion methods require the external manipulation of this Switch-on-a-chip's control signals to provide maximum flexibility infunctionality e.g., port expansion can be used to support multiple priorities. Fig. 7 shows the richness of the Switch-on-a-chip application space.

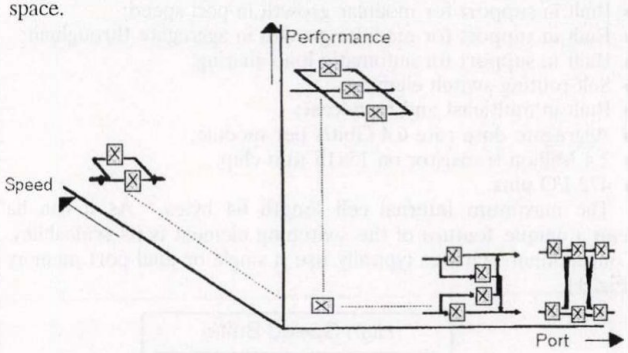


Fig. 7. Scope of the Switch-on-a-chip application space

Multicast support

Switch-on-a-chip supports to build high-speed communication networks which support applications that feature a heterogeneous mix of voice, data and video traffic. Typically such systems require the capability of handling multipoint connections for services such as video distribution and teleconferencing. Switch-on-a-chip provides a flexible multicast capability: it is possible to send a copy of a packet to all (broadcast) or only a subset (multicast) of the switch module's output ports. In order to conserve buffer memory, only one packet storage location is used, from which multiple copies are sent. The activation of a multicast connection is done through the packet routing header, and a dynamically programmable table internal to the Switch-on-a-chip module.

Finally the products which are using the Switch-on-a-chip ATM switching modules include:

- **8260** Multiprotocol Intelligent Switching Hub.
  This device is the heart of ATM Campus networking.
- **2220** Nways Broadband Switch family.
  Ranging from small branch models to large backbone switches.

REFERENCES
[1] W. E. Denzel, A. P. J. Engbersen, I. Iliadis and G. Karlsson: "A highly Modular Packet Switch for Gbit/s Rates", in *Proc. of ISS'92*, Yokohama, Japan, A8.3, 1992.
[2] W. E. Denzel, A. P. J. Engebersen and I. Iliadis: "A flexible Shared-Buffer Switch for ATM at Gbit/s Rates", accepted by *Computer Networks & ISDN Systems*.

IBM HUNGARY
Alsó-Törökvész út 2.
H-1022 Budapest, Hungary
Tel.: +36-1-212-4680; Fax: +36-1-212-4652

ATM TEST SOLUTIONS FROM HEWLETT-PACKARD

ATM or Fast Packet Switching was first developed in the late 1980s, and Hewlett-Packard quickly became involved with the testing of this emerging technology by its work in the European RACE program – HP was one of the consortium members on an early development project for ATM test equipment called Parasol.

In 1992, HP introduced the first real-time ATM analyzer based on its HP 75000 Series 90 SONET/SDH test system. This VXI-based system provided the foundation for the Broadband Series Test System (HP E4200A) which today is regarded as the industry standard test set for R&D and field trials of boardband ATM technology. The ATM community recognises that HP has helped the development of ATM technology by supplying test equipment early. The experience gained with this instrument has enabled HP to develop a whole range of dedicated ATM test solutions for different applications including installation and maintenance.

HP has also been actively involved in ATM standards development with the ATM Forum and ITU-T standards committees. HP employees have been on the Board of Directors of the ATM Forum and have chaired the Test Subworking Group of the Forum's Technical Committee. Recently, an HP employee contributed a new ITU-T standard to Study Group IV (O.191) defining an ATM test-cell standard.

Test equipment is used to resolve problems and troubleshoot equipment faults. A new technology such as ATM will create many unforeseen problems of interworking and standards compatibility. Nobody is yet sure how ATM will impact and interact with existing networks, and what type of Quality of Service issues will arise. There is still a lot to learn about the performance of real ATM networks as traffic is handled in a totally different way to conventional circuit-swtiched systems. An important ingredient to the understanding and successful operation of ATM networks is access to appropriate test equipment tailored for the specific application. By the end of 1995, HP will have 11 different solutions for testing ATM networks and services, some of which combine testing of conventional transmission technologies and LAN protocols.

Here is a selection of HP solutions:

1. HP E4200/E4210 BROADBAND SERIES TEST SYSTEM

The HP E4200 Boardand Series Test System delivers the most comprehensive coverage of ATM tests up to 622 Mbit/s – at all layers, covering all major and emerging protocols and transmission media. It has become the preeminent solution for R&D and field trials of new ATM equipment and systems.

Available as a rackmount or transportable VXI mainframe, it can be configured from a wide variety of line interface and cell protocol processor modules. The HP E4200 can handle any combination of protocols and transmission media, including video, ATM LAN decoders, signalling and so on. Plus optional software automates ATM conformance, ATM interoperability and cell-layer interoperability tests.

The HP 4219A ATM network impairment emulator allows the user to input specific network impairments such as cell loss, error mis-insertion, delay and delay variation. This capability is useful to evaluate performance under degraded conditions.

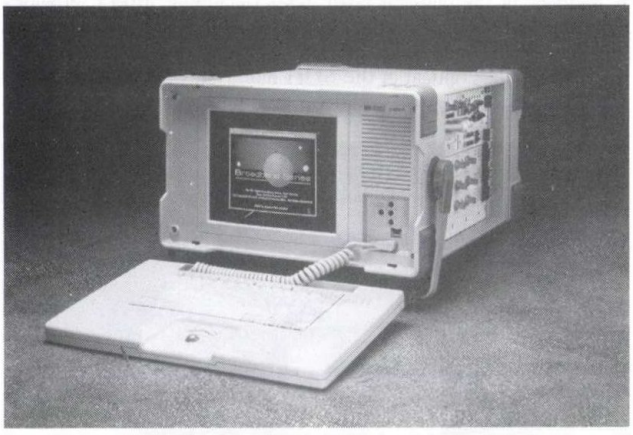


Fig. 1. HP E4200 Broadband Series Test System

2. HP E5200 BROADBAND INTERNETWORK ANALYZER

Fig. 2. HP E5200 Broadband Internetwork Analyzer

The HP E5200 Broadband Internetwork Analyzer has the tools to find and solve amost any problem experienced in installation of ATM networks. The Link Monitor highlights errors and even points towards possible causes and solutions. On-line help goes beyond the usual instructions, providing explanation and guidance through key test steps. The emphasis is on ease-of-use, however this does not mean the HP E5200 is short on performance. It is built around three parallel processors that deliver the additional computing power broadband network testing requires. Plus, it's compatible with the HP Broadband Series Test System, making it easy to share data and test programs between the lab and the field.

3. HP 37717C PDH/SDH/ATM TEST SET

For installation and maintenance of PDH/SDH and ATM networks, it's hard to beat the HP 37117C multipurpose test set. It handles all the physical layer tests (including jitter), provides cell layer and optionally service layer tests for ATM, simulates and monitors traffic and checks alarms. This one-box tester operates up to 622 Mbit/s optical in the SDH hierarchy and handles all the PDH interfaces too, with add/drop multiplexing.

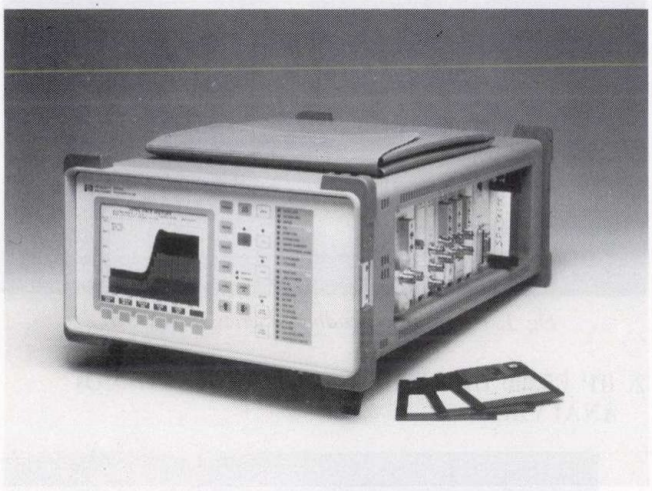


Fig. 3. HP 37717C PDH/SDH/ATM Test Set

For more information on HP's ATM test family, consult:

- *ATM test family brochure (publication number 5964-1922E);*
- *Broadband Communications Map (publications number 5963-9489E);*
- *"Emerging Test Requirements for Broadband Networks" (publication number 5964-0109E);*
- *"Building and Managing ATM Networks" (publication number 5964-1591E).*

Contact: Hugh Walker
e-mail: hugh_walker@hp1400.desk.hp.com
Hewlett-Packard Magyarország Kft.
e-mail: sándor domokos@HP-hungary-om.1.om.hp.com
H-1146 Budapest, Erzsébet királyné útja 1/c.

HEWLETT-PACKARD LTD.
South Quinsferry WEST Lothian
Scotland EH30 9TG

ALCATEL LAUNCHES SERIES 1100™ FAMILY

Alcatel Data Networks announced a new line of switches — the 1100™ HSS™. Five models will become available, targeted at both private networks and carrier central offices. All HSS run off a common architecture — dubbed Avanza — and common network management.

1. THE MODEL 400

For private networks, Alcatel is offering the Model 400. This Frame Relay switch has four slots, one of which is dedicated to a network management module. User and trunk cards are set up from the network management system, and one network management module can handle up to sixteen interconnected Model 400s. Alcatel's Ratemaster™ congestion management system enables administrators to set priority levels for different types of traffic — e.g., voice and SNA traffic can get high priority.

Interface cards support Frame Relay (up to eight ports per card, and the ports can be configured for different speeds), and there also is an Ethernet and token ring routing card that Alcatel codeveloped with Cisco. A 32-channel voice-over-Frame-Relay card is scheduled for the fourth quarter, although voice will be limited to on-net traffic and only over a network with Alcatel equipment installed on an end-to-end basis. TDM cross connect cards are also planned.

Each Frame Relay switch card has capacity for up to 8 Mbit/s, and the HSS has a 340 Mbit/s midplane design.

One application area for the 400 is a replacement for routers.

2. 700 SERIES FOR PUBLIC NETWORKS

Alcatel also announced four models in its 700 series for the public network. While the 400 is a one-shelf unit, a 700 node can have up to seven shelves. The 700s are designed a common 12-slot chassis that can be populated with a mixture of ATM, TDM and Frame Relay switch cards. The 400's midplane has been ported to the 700 but extended to include ATM capability. Cell and Frame Relay traffic can run simultaneously and internetwork within the switch and go out on respective trunk interfaces.

The Model 710 is a Frame-Relay-only switch. The Model 740 will offer Frame Relay and ATM concentration capability but no ATM switching, while Alcatel's Models 750 (2.5 Gbit/s) and 760 (10 Gbit/s) will offer full Frame Relay and ATM switching.

The Series 700 HSS switch has two sides: one performs Frame Relay and TDM interfacing, while the other side performs ATM interfacing and switching. Both sides contain interfaces which allows lines and trunks to access their switching functions, and can operate independently. It is not even necessary for both sets of interfacing and switching functions to be installed in the switch for other set of functions to operate, since each side of the switch can connect Frame Relay, ATM or future TDM traffic between its access ports. However, with both sets of functions installed in the switch, the Cell Interconnect Gateway function interworks traffic between the two sides of the switch. The Cell Interconnect Gateway acts like a Frame Relay access port to the Frame Relay side of the switch and looks like an ATM access port to the ATM side of the switch, allowing it to interwork Frame Relay and ATM traffic.

Families of line cards provide both ATM and Frame Relay access to the HSS over a wide range of speeds and protocols. High Speed Couplers (HSCs) handle high-speed ATM channels; while Medium Speed Couplers (MSCs) handle Frame Relay channels. The couplers interwork data between the access port protocols and the Frame Relay and ATM protocols used in the switch. The couplers operate at speeds between 1,2 kHz and 155 MHz and support E1, T1, DS3, G.703, G.704, STS-3c, STM-1 and OC-3 interfaces, making them suitable for private or public networking applications.

Since the switch processes traffic on the coupler cards and not on a central processor card, the performance of the HSS increases proportionally as ports are added, and a failure in a single coupler board cannot take down the whole switch. This modular architecture also allows ATM-only switch configuration to be upgraded with Frame Relay couplers and Frame Relay only switch configuration to be upgraded with ATM couplers.

The 1100 HSS switch routes traffic between Frame Relay couplers over fully meshed paths on a CRYSTAL midplane (a proven Alcatel Data Networks PBX architecture), and routes traffic between ATM couplers using an ATM switching matrix.

ATM Electrical Interface (AEI) point-to-point interconnections provide bi-directional serial connections between the ATM couplers and the switching matrix. The ATM switching matrix routes traffic between the ATM couplers using 16 AEI port x 16 AEI port x 155/622 Mbit/s point-to-point matrix, allowing it to provide point-to-point and point-to-multipoint switching at a 2,5/10 Gbit/s aggregate throughput. Its switching channels share a dynamically managed output buffer memory pool, avoiding blocking from port level contention.

The Cell Interconnection Gateway (HBCIG) board interwork between ATM, Frame Relay and future TDM traffic sources by interworking traffic between the CRYSTAL midplane and the ATM switching matrix. It provides a gateway between the medium speed Frame Relay couplers and the high-speed ATM couplers. Applique cards provide electrical and physical interfaces to the couplers and provide access ports at the rear panels of the switch. An optional bridge/router card accesses the

CRYSTAL midplane trough one of the Frame Relay couplers and its applique, and provides ports for either two Ethernet or two token ring networks and an X.25, SNA or Frame Relay wide area network.

Node Administrator (NA) software resident in the Control Unit (CU) loads, monitors and controls the HSS and collects status performance and accounting information from the HSS. Network Management Software (NMS) running on a workstation in turn controls and uploads information from the NA via an Ethernet link managed by the CU. A terminal can be used to troubleshoot the CU card and the coupler cards over V.24/V.28 (RS232E) ports accessible at the rear panel of the switch. A Local SubNetwork (LSN) bus allows software on the on-line CU to exchange data with the off-line software on an optional redundant CU, minimizing the chances of losing data if a switchover between CUs occurs.

The CU exchanges status and control signals with the Frame Relay couplers and the Cell Interconnection Gateway (HBCIG) over a point-to-multipoint Input/Output System (IOS) control bus. The CU exchanges status and control signals with the broadband ATM Switching Element (BBASE) and the ATM couplers via the IOS, HBCIG and ATM Virtual Channels. The Internal Serial Link (ISL) carries major alarms and inventory data from the couplers, the CRYSTAL interconnection gateway, the ATM switching matrix and the ATM couplers. The Input Output System (IOS) control bus and the ISL bus can be extended to up to six expansion chassis.

In case of conditions which may cause a damage to the equipment or loss of customer service the Status, Alarm and Control board (HBSAC) sends alarms to the Node Administrator via the ISL bus and the CU. It sends fan, temperature, and power supply alarms to the Node Administrator program in the CU and the HDLC ISL bus. It also drives a front panel display showing temperature, power, and fan alarms.

The 1100 HSS provides Frame Relay and ATM services to network users at its User Network Interfaces (UNI) and the Network Interfaces (NNI).

Frame Relay UNI services include:

- Bandwidth admission and traffic enforcement;
- Inbound data pipelining using frame segmentation;
- Native Frame Relay access conforming to ITU-TSS Q.922 Annex A;
- In channel signalling protocols.
- Frame Relay Assembly Disassembly (FRAD) to support devices that support an HDLC protocol.

Frame Relay NNI services include:

- Managing Permanent Virtual Circuit (PVC) over a trunk line;
  - Detecting faults on Frame Relay connections;
  - Collecting statistical information on Frame Relay connections;
  - Generating accounting data from each MSC;
  - Establishing and clearing Frame Relay Virtual Connections;
  - Performing intra- and inter-node call control and transfer procedures using the
  - Frame Relay Alcatel Network Protocol (FRANP).

At the ATM UNI the 1100 HSS provides:

- Physical layer connections;
- Bearer service functions;
- Standard ITU-TSS ATM cell structures and ATM layer functions;
- Traffic control;
- VP/VC addressing;

The 1100 HSS uses the Alcatel Network Protocol (ANP) to manage the HSS NNI to support routing and internetworking between different types of services for signalling. The 1100 HSS connects to a public network via its ATM UNI or NNI, and to a private network using its ATM broadband inter exchange carrier interface for Frame Relay Service and cell relay service.

The 1100 HSS uses the Cell Alcatel Network Protocol (CANP) to set up virtual paths (VP) and virtual channels (VC) between switches in a network based on subscriber and trunk VP and VC parameters specified by the NMS. When an operator requests a connection a CANP routing algorithm automatically chooses the best path between two network accesses.

Using an optional integrated multi-protocol router, the HSS provides LAN Data Services that support 802.3 Ethernet LAN connections for local (LAN-to-LAN) and network (LAN-to-FR) switching. The HSS interworks with the TPX^{TM}, TPF and PSX^{TM} series of AND products, adapts non ATM traffic using its CIG card, and encapsulates IP protocol in Frame Relay frames. the HSS synchronizes its public network transmission interfaces (i.e. T1, E1, E3, DS3, STM1 and STS3c) with clocks traceable to a primary reference source like a Stratum 1, 2 or 3 clock.

3. NODE MANAGEMENT

The 1100 Node Management (NM) is the next generation state-of-the-art Network Management (NM) System developed by Alcatel Data Networks the 1100 NM provides basic node management capabilities for individual Alcatel products. It also manages homogeneous network containing equipment in the HSS product line that support Frame Relay, LAN, SNMP and ATM protocols. The NM architecture distributes management intelligence between the NM platform and the agent that resides on the 1100 HSS.

Alcatel has based the 1100 NM platform architecture upon the HP OpenView® (Distributed Management Environment (DME). The open architecture of the 1100 NM supports SNMP and FTP standards and uses these standards to manage many of the aspects of the 1100 HSS product line.

NMS features include:

- User access to the NMS via a graphical user interface.
- Management Protocols handle communications between the HSS and the NMS.
- The ability to manage the HSS via Local Terminal or via a Remote Terminal.
- Open to third party NMSs or to other Alcatel NMSs.
- Installation program that automatically decompresses files during installation, making it unnecessary for the user to manually load files.
- The Management Information Base (MIB) represents network configuration and status information in the form of a language between the agent and the manager.
- Network management support for Frame Relay/ATM network interworking including configuration management for internetworking Scenario 1 and interworking Scenario 2.
- An optional backup workstation contains the NM software. If the primary NM fail, it takes less than fifteen minutes to start this backup system.

The Performance Management function gathers statistical data on the NE and relays it to the NM where it can be used to analyze trends in the operation of the NE. The statistics include real time statistics, interval statistics, historical statistics, and accounting statistics.

ATM VP/VC service management includes the following features:

- Management of VPs via manual set-up.
- Management and establishment of VCs over existing VPs.
- Manual routing at VP level.
- Manual reconfiguration.
- Manual point-to-multipoint support.
- VP/VC management support when VPs and VCs are established by signalling.
- Resource management.
- Synchronization.

The configuration management function specifies the types of code to be loaded into the NE, creates the corresponding configuration parameters, and transfers the code and configuration parameters to the NE. Configuration parameters set the configuration of the hardware (nodes, boards, ports, etc.) and the associated protocols, accesses and services. Both the NE and the NM perform semantic checking, while syntactic checking is done on the NM. Configuration management provides configuration parameters, multiple code and configuration versions, and industry standard file transfer protocols that initially load and the update software.

The Security Management function provides authentication for users, and secures NM accesses to the NE via the SNMP and the FTP. The Fault Management function reports alarms, obtains diagnostics, correct faults, and graphically displays the status of

network resources. SNMP passes fault management information between the NM and the NE.

The NM provides network management support for the integrated LAN Bridge/Router (LBROUTER). For configuration Management, this support includes integration at the user level, with a separate application used to configure the LBROUTER.

The 700 products can replace TDM muxes in private networks, while offering carriers a scalable system ranging from the central office to edge devices and within the customer premises.

The Avanza architecture, which is shared by the Model 400 and the Model 700 series, could appeal to local and interexchange carriers who want to offer customers a turn-key, end-to-end broadband solution. The carriers have had success marketing this approach for LAN interconnection and Frame Relay, and by having a single-vendor solution available for both their internal and customer networks, they could reduce inventory, training and implementation costs.

The 1100 HSS™, TPX™, PSX™ RateMaster™ are trademarks of Alcatel Data Networks Inc. The HP OpenView is registered trademark of Hewlett-Packard Company.

Table 1. Technical specifications of Alcatel 1100™ HSS™ Series 700 multiservice network switch

	ATM	Frame Relay	HDLC/SDLC	LAN	TDM
Number of ports (max):	120	528	528	116	128
Line speeds (max):	155 Mbit/s, 622 Mbit/s (in the future)	4 Mbit/s	4 Mbit/s	16 Mbit/s	2 Mbit/ps
Line interfaces:	155 Mbit/s: SMF,MMF,coax/G.703	V.24,V.28,V.35,V.36	V.24,V28,V.35,V.36	STP,UTP,AUI	G.703 (Nx64 Kb/ps)
	34/35 Mbit/s: E3/DS3	V.11/X.21/X.24	V.11/X.21/		
	1,5/2 Mbit/s: T1/E1	T1/E1	X.24, T1/E1		
Throughput:	ATM Switching Matrix: 10 Gbit/s, bi-directional(16x622 Mbit/s)				
Frame Relay:	220.000 f/ps, bi-directional (280 byte-frames)				
Protocol support:	ATM, Frame Relay, transparent HDLC/SDLC (X.25, SNA), Ethernet 802.3, Token Ring 802.5				
Virtual channels (max):	ATM:8.192, Frame Relay: 122.200				
Standards compliance:	ANSI, ETSI, ITU_TSS, Frame Relay Forum, ATM Forum				

ALCATEL

MEASURING ATM BY WANDEL & GOLTERMANN

1. INTRODUCTION

Wandel & Goltermann offers two kinds of solutions for measuring ATM:

- The OC-3/STM-1 ATM Analysis Package for its DA-30® C Family of Internetwork Analyzers for monitoring, troubleshooting, simulation, and interconnect-device testing on ATM networks, and
- The ATM Module for STM-1/STS-3c for its ANT-20 Advanced Network Tester, for installing, maintaining, and testing ATM systems.

2. DA-30C PRODUCT DESCRIPTION

The DA-30C analyzer (Fig. 1) is a system platform for multiple analyzers, line interface modules, and software packages. It gets its remarkable power from multiple independent RISC and CISC processors, providing the DA-30C analyzer with the power to capture, monitor and generate traffic at full line rates.

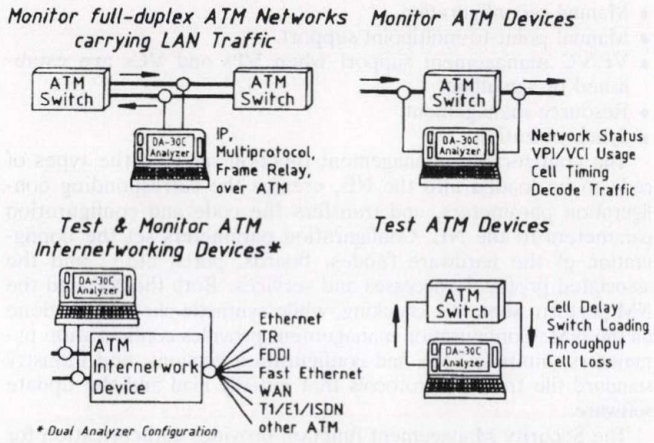


Fig. 1. Applications of the OC-3/STM-1 ATM Interface

The DA-30C analyzer equipped with the OC-3/STM-1 ATM Analysis Package utilizes the dual analyzer, real-time reassembly, and full protocol decode capabilities of the DA-30C analyzer, lab advantages include the ability to verify the functionality of any LAN traffic running on an ATM link.

The OC-3/STM-1 ATM Analysis System is comprised of a plug-in board that slides into the DA-30C analyzer mainframe and an ATM Analysis Software Package running under the DA-30 for Windows interface.

The OC-3/STM-1 ATM Analysis System has four modes of operation:

- *Filter (Trigger) Capture*: with 10 parallel hardware filters and 16 Mbytes of fast-capture RAM, the DA-30C analyzer can pull in exactly the desired ATM cells or reassembled network frames for capture and/or real-time analysis.
- *Monitor* providing:
  - real-time network status, error, and traffic statistics;
  - histographs displaying any available network statistics as they occur over time;
  - ATM layer tests;
  - VP/VC bandwidth discovery, automatically determining the first 1024 active VPI/VCIs on an ATM link;
  - cell interarrival time distribution;
  - cell delay distribution and cell loss utilizing cell transmit to create a measurement of the delay from the DA-30C analyzer transmitter back to its own receiver;
  - real-time AAL5 reassembly.
- *Off-Line Examine and Real-Time Decode*
- *Transmit*: trasmitting user-defined sequences of cells or segmented AAL PDUs at up to full line rate. Several single error and continuous alarm insertions are available for the transmit signal.

3. ANT-20 ATM FUNCTIONALITY

The ANT-20 Advanced Network Tester (Fig. 2) has been developed for use by operators and manufacturers of modern communications networks. Its modular concept for SDH, SONET, PDH, ATM, TMN, etc. offers flexibility with sure future viability, while it is one of the most compact instruments in its class well suited to field applications. Built around a standard Microsoft® Windows™ graphical user interface and a large screen, application-oriented instrument settings are provided with simultaneous display of significant parameters and results.

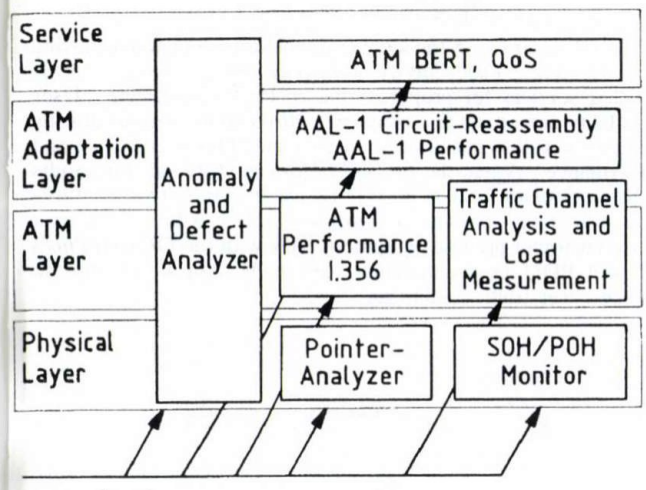


Fig. 2. Analyzers of ANT-20

The ATM options provide the possibility for ATM Traffic Generation and Analysis up to 155 Mbit/s. As the ATM module for STM-1/STS-3C is built into the ANT-20, the extension slot of the instrument remains free for other modules, thus enabling combination of the ATM options with jitter or STM-16 at the same time.

Frame structures to ITU-T G.804/832/709 and ANSI T1.105/107 from 1.5 Mbit/s to 155 Mbit/s are provided for generating and analyzing ATM traffic. Users can freely configure a test channel for error measurements when checking ATM signal paths. The bandwidth of the test channel can be varied continuously up to 100 % in order to test traffic control functions.

In addition to physical layer measurements, the ATM module detects correctable/non-correctable header errors, cell delay and the bit error ratio in the cell payload. The cell distribution and the average cell rate are analyzed. Cell loss events and the used/unused transmission bandwidth are depicted in a chart of load vs. time.

Wandel & Goltermann offers the following applications and special features in terms of ATM:

- ATM Mappings: E1, E3, E4;
- background load generator;
- ATM performance analysis;
- AAL-1 segmentation and reassembly;
- AAL-1 error measurements and statistics;
- AAL-1 error insertion;
- AAL-1 circuit emulation/reassembly;
- load profile: burst;
- F4/F5 alarm flow generation and analysis;
- peak cell rate measurement;
- cell classification user channel;
- real-time measurement of the following performance parameters: errored cells, cell loss, cell misinsertion rate and mean cell transfer delay. The measurement conforms to the ITU I.356 and O.191 standards.

Customer requirements with respect to ATM functions for the ANT-20 will determine further development.

For further information, contact:

Wandel & Goltermann
Electronic Measurement Technology
Krisztina Kapitánffy
Tel/Fax: (36-1) 203-8172
H-1113 Budapest
Bartók Béla út 120. room 330.
P.O.B. 161, H-1502 Budapest

Krisztina Kapitánffy
WANDEL & GOLTERMANN

HUNGARIAN TELECOMMUNICATIONS COMPANY LTD. (HTC) PLANS ATM

HTC, the Hungarian public network operator, is planning and preparing for operation of broadband integrated services networks (B-ISDN). HTC decided to provide B-ISDN services on ATM platform. ATM technology by its features — flexible, service independent and efficient — ensures great opportunities to its users.

Some of the primary broadband services considered are:

- LAN/WAN networking;
- high-speed medical imaging and telemedicine;
- world-wide videoconferencing;
- interactive multimedia services;
- visualization aids to specialists, including medical imaging, engineering design and scientific research;
- multimedia information tools and resource discovery from remote databases;
- access to supercomputers, load sharing and disaster recovery;
- fast turn around on advertising agency ads.

HTC starts with an ATM trial network in Budapest area, which provides permanent and switched VC connections in the first and the second phase, respectively. Meanwhile the trial network will be extended to the larger cities in Hungary. The ATM trial network will be connected to the paneuropean pilot network, too. So that international connection will be ensured, as well. After the field trial period, services will be commercially available by the end of 1997.

The services, which HTC will support up to 155 Mbit/s over the ATM network are the followings:

- CBR services;
- data applications;
- video services;
- multimedia services

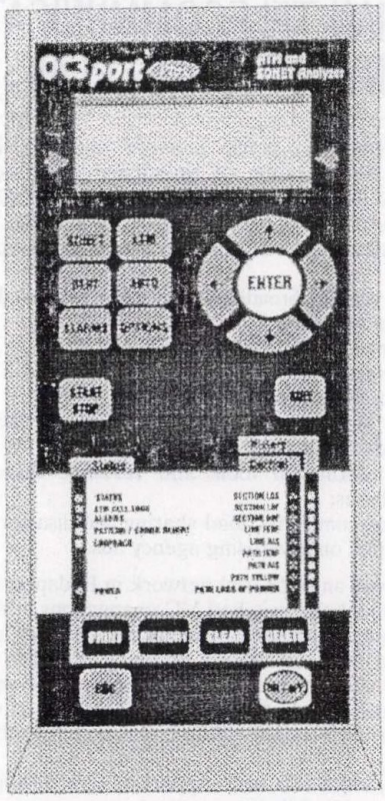
on permanent and afterwards on switched connections.

Ágnes Szente
HTC
H-1013 Budapest, Krisztina krt. 6-8., Hungary

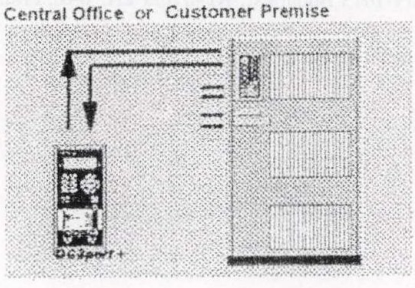
OC3PORT PLUS IS THE PERFECT TESTER

1. ATM EQUIPMENT INSTALLATION AND PRODUCTION TESTING

The OC3*port Plus* contains a complete set of ATM features necessary for performing ATM equipment installation testing. In addition, all test results can be stored for the later creation of installation tests reports. Each physical interface can be tested for proper Virtual Circuit availability, proper service category performance and overall Quality of Service performance.

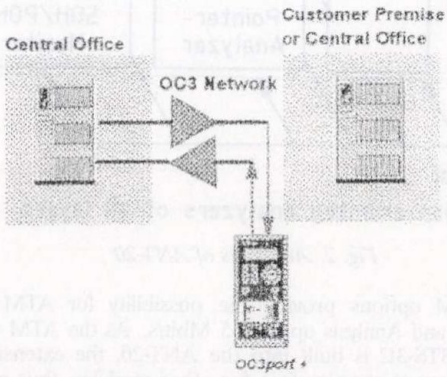


The OC3*port Plus* is specifically designed to allow testing of any type of OC3/ATM equipment for proper operation, configuration, compliance and interoperability. The OC3*port Plus* can also test the equipment's Operation and Management (OAM) channel and Interim Local Management (ILMI) channel to allow identification of network or equipment problems. In addition when SVC's are being used the OC3*port Plus* and be used to setup, test and tear down switched circuits.



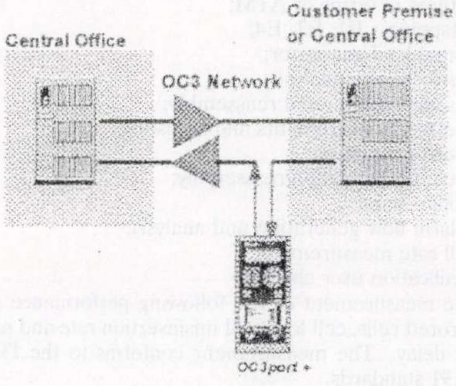
2. ATM SERVICE TURN-UP TESTING

The OC3*port Plus* is the ideal test tool for testing and qualifying ATM services. Field Service personnel can quickly test newly installed services for proper traffic contract compliance. User availability and access to particular Virtual Circuits can be quickly verified, as well as qualification of guaranteed level of service and performance. Again, the OC3*port Plus* can store all test results so that customer installation reports can be generated. Features such as IP Ping transmission and monitoring can be used to test access customer premise equipment. Also with the OC3*port Plus*'s cell and PDU programmability, test traffic can be created to simulate user traffic conditions.



3. ATM NETWORK MAINTENANCE AND IN-SERVICE MONITORING

The OC3*port Plus* contains many features for monitoring intermittent network problems and determining customer traffic parameters. The OC3*port Plus* can be placed in a passive monitoring mode where it can monitor all OC3 and ATM network conditions while allowing all user data to pass through nonintrusively. In this mode all OC3 error and alarm conditions are monitored and Data and Time stamped so that intermittent problems can be diagnosed. In addition, ATM virtual circuits can be monitored to determine usage patterns, also ATM level errors and alarms can be monitored, time stamped and saved for future investigation.



Interfaces: OC3 SONET or STM-1 (both with singlenode or multinode fiber or with UTP5 copper physical interfaces).

DeskNet Systems Inc. (New York)
Hungarian distributor:
Schoeller Network design Hungary Ltd.
1122 Budapest, Abos u. 16.
Tel.: 361 156 0043; Fax: 361 202 6086
Contact person: Mr Tibor Szendrényi

DEUTSCHE TELEKOM'S DEVELOPMENTS FOR THE EUROPEAN INFORMATION INFRASTRUCTURE*

The paper presents an overview on the implementation of ATM-based information transfer services in Germany, investigating the various scenarios meeting different network structures. Author comes to the conclusion that the main requirements are more or less the same for all the scenarios. Tariff policy applied by Deutsche Telekom is also presented in the paper.

1. INTRODUCTION

With regard to the implementation of infrastructures supporting the European Information Infrastructure, Deutsche Telekom has already made great progress. The telephone network is nearly 100 % digital and will be a full digital network in 1997. ISDN is available at every location in Germany, and more than 2 Mio. B-channels are already sold to customers. Former east Germany's telecommunications network has been upgraded with leading edge technology, including 1.2 Mio. fiber optic subscriber lines.

Deutsche Telekom's online service "T-Online"Trademark has attracted 850,000 users and is growing by about 50,000 customers per month. T-Online is now available in about 3 % of all German households. Due to its nationwide access on the basis of local tariff via ISDN using 64,000 bits and the analogue telephone network with up to 28,000 bit/s, T-Online experiences great success. It offers all kinds of information including access to the Internet and a worldwide e-mail service.

With increasing processing power of state-of-the-art personal computers, applications increasingly use graphical human interfaces or even video animation for better performance. Multimedia applications using CD-ROM as a cheap local media become more and more popular. This creates increasing demand for data transfer rates with 64 kbit/s as the lower boundary when adequate quality of service shall be assured. However, Multimedia applications have an increasing demand for higher bandwidth and also variable telecommmunication transfer rates, adaptable to the currently required audio and video performance.

The factors described above have caused many network operators and service providers to study the commercial offer of broadband telecommunication services.

Currently most of the initiatives are trials to get a better understanding with this new technology. However, experts forecast commercial availability of broadband telecommunication services for the middle of the nineties. Some network operators offer commercial broadband services already today, e.g. Deutsche Telekom offers direct dialling broadband service covering some 80 cities in Germany based on 140 Mbit/s circuit switching since 1988 (see below).

2. ASYNCHRONOUS TRANSFER MODE – ATM

Recognized Standards Organizations (regional and worldwide) and industry groups have provided initial standards for broadband services. The principle of "asynchronous transfer mode, ATM" has been standardized as a unique mechanism for information transport and switching. ATM has become a promising concept, which meets many different requirements. For the first time in the history of telecommunication it will be possible to fully separate user requirements and related services and applications from network design. Cells of uniform length and structure are used to carry the information. Information can be of any kind, slowest data, speech, image, video, high-speed data or even HDTV-signals. The application determines the amount of data to be transmitted; low information rate results in only few cells per second to be transferred, high information rate results in many cells per second to be transferred.

* The paper is based on a part of the lecture given at the Scientific Days of MATÁV-PKI, Budapest, November 1995.

"ATM, the key for multimedia telecommunication"

The advantage of the ATM principle can be seen in its flexibility. Depending on the amount of information to be transported more or less cells may be filed with information. In periods of no information the channel capacity can dynamically be used for other information sources. The ATM principle enables "bitrate on demand" service limited only by the maximum physical channel rate. When the related mechanisms are available, the user will only pay for used cells, i.e., the tariff will be volume based.

2.1. Implementation of Asynchronous Transfer Mode

Deutsche Telekom has implemented an ATM-based pilot network interconnecting the cities of Hamburg, Berlin and Cologne (with remote access to Bonn). Its objective is to gain experience in the new technology of ATM and to test all kinds of broadband applications. It is in operation since CeBIT 1994.

Initially it offers permanent and reserved communication services based on the "Virtual Path" technique. The information transfer rate will be up to 155 Mbit/s (including the control overhead).

While signalling procedures are internationally standardized — this has been achieved by 1994 — on demand services will be offered in 1996.

Services which are offered include interconnection of Local Area Networks (LAN) and Metropolitan Area Networks (MAN) with the help of connectionless servers as well as monomedia and multimedia services based on connection oriented transport classes. Subsequently one specific application trial is described in more detail.

Currently this pilot network has been expanded to cover major German cities. Today ATM services are available in the regions Hamburg, Hannover, Cologne, Bonn, Wiesbaden, Darmstadt, Heidelberg, Karlsruhe, Stuttgart, Ulm, Nuremberg, Munich, Leipzig and Berlin. Commercial services have also been offered in 1995. The largest contract was signed with the DFN-Verein, an organisation which serves state universities and public research facilities with a countrywide ATM service.

Pilot Conditions and Tariffs

Initial tariffs have been filed early in February 1995. The prices are oriented to future commercial offers and — during the pilot phase — will be gradually applied, i.e., during the first few weeks of connections (test phase) the access is free of charge, until some two thirds of the full price is applied after completion of the test phase. This pricing policy shall give potential users a sound and reliable basis for their own cost/benefit calculations.

The pricing principles are based on

- a unique connection fee, depending on the maximum channel rate (2, 34, or 155 Mbit/s);
- a recurrent monthly fixed charge (differentiated on bandwidth requirements);
- usage based charges taking into account distance (local/long distance) and volume.

European ATM Trial

BT, Deutsche Telekom, France Telecom, STET, and Telefonica have prepared a Memorandum of Understanding for the early provision of an ATM based broadband ISDN in Europe. Following these five founding operators mentioned above, there are now seventeen operators from fifteen European countries who have finally signed this MoU. They have formed the "ATM Pilot Coordination Group, APCG" which coordinates all necessary

steps of the pilot. Each network operator participates with one network node. All nodes are interconnected via either 155 Mbit/s or 34 Mbit/s. During the pilot phase the services will be based on "ATM-Virtual Path Connections" a kind of "ATM-Leased Bandwidth". The services includes "Connectionless Broadband Data Services, CBDS" to interconnect local area networks as well as "Connection-Oriented Broadband Services" for real-time audio-visual or multimedia communications.

The pilot has begun mid 1994, starting with permanent and reserved services based on "Virtual Path" technique. As with METRAN, it was again EURESCOM who has provided the network operators with the necessary specifications.

Franco-German Alliance

In the framework of traditional close cooperation of Deutsche Telekom with our colleagues from France Telecom, both companies have agreed to provide international ATM-services in 1995. To support this plan, experts from Telekom together with experts from France Telecom have specified equipment and network capabilities to support compatible services on the basis of standards as prepared by the Standardization Sector of the ITU and by ETSI[1], and based on specifications prepared by EURESCOM. The specifications are largely complete.

With the interconnected network nodes installed in Karlsruhe (with a multiplexor in Heidelberg) and Stuttgart (with a multiplexor in Ulm) by Deutsche Telekom and in Paris and Lyon by France Telecom the agreed pilot phase is now ready for operation. Negotiations with customers are under way. The national implementations of ATM infrastructures by France Telecom and Deutsche Telekom are considered as access networks to the joint ATM network. Full Commercial Service will be offered at the beginning of 1996.

In the initial commercial phase the services offered will be permanent service and reserved service based on ATM virtual path connections, i.e. connections are to be established by means of network management. Specifications for on demand services using switched virtual circuits are in progress. Initial offering of on demand services is expected in 1996, when the signalling protocols will be implemented.

2.2. Can ATM Solve Problems?

With the ATM-principle we get a technology, which is promising to solve many of today's problems. Nevertheless, a number of questions remain open today, or can finally be answered only when sufficient experience with ATM networks will be available. Especially answers to the following questions need to be given:

- What will be the traffc characteristics of future broadband applications, especially of those applications, which may generate variable information rate (bursty) traffic?
- How will a large national or international ATM network behave in case of overload?
- Which measures need to be taken to compensate for signal delays and its variations, which are inherent characteristics of the ATM technique?
- What are the tariffing principles to be applied in an ATM network?

In addition answers should be available on the costs of an ATM network, the local availability of such an ATM network, the availability and stability of international standards, the availability of in-house systems and terminals complying with ATM standard interfaces, and last but not least the applications which solve end user's communications problems.

Most of these questions cannot yet be finally answered. Nevertheless, I am convinced that ATM gives us the technology which at least has the power to meet the requirements to a network of tomorrow, namely to provide flexible, intelligent, costsaving information transfer.

[1] ETSI is the European Telecommunications Standards Institute, located at Sofia Antipolis, France.

3. FIBER NETWORKS, HYBRID NETWORKS, ASYMMETRIC DIGITAL SUBSCRIBER LOOP, HIGH BITRATE DIGITAL SUBSCRIBER LOOP

Deployment of fiber optical cables in the local loop (fiber to the hub, fiber to the kerb, fiber to the home) will take some time and will request enormous amounts of investment. However, digital technologies have even made it possible to use existing cable networks (telephone networks and cable-TV networks) for high bandwidth services. Currently, a number of different scenarios can be observed. Their individual application is determined by the actual market position of the individual network operator as well as by the availability of related technologies and their costs.

4. SCENARIOS FOR BROADBAND NETWORKS

In all cases the requirements to be met with each scenario are more or less the same. They can be described as follows:

User requirements

- application adequate information transfer rate (bandwidth on demand);
- short response time;
- individual service profile;
- low tariffs;
- high service quality.

Network operator requirements

- flexible to changing user requirements;
- market demand oriented investment;
- quick installation;
- high network performance;
- low investment costs;
- low maintenance costs;
- integration into existing network.

Regulator's requirements (where applicable)

- non-discriminating access (ONP, ONA);
- standardized access interfaces;
- transparent tariff system;
- nationwide network infrastructure.

There exist two scenarios which can be considered as the "traditional scenarios", i.e., scenario 1 is a typical telecom network operator scenario, whereas scenario 2 is typical for cable-TV operators. Scenario 1 can be further subdivided into sub-cases, where existing copper cables for telephony (twisted pair based) are used to provide broadband services.

4.1. ATM-based Broadband ISDN Scenario

A typical scenario for the implementation of broadband capabilities into an existing telecommunication network is given in Figs. 1–3. In a first phase broadband capabilities are implemented in an overlay architecture as shown in Fig. 1. With increasing demand for broadband services this overlay structure will partly be integrated with existing network elements into a uniform network structure, Fig. 2. In a final phase a fully integrated structure will be achieved, Fig. 3. All kinds of services will then be supported by a uniform digital network, including interactive communication services and radio- and TV-distribution services.

4.2. Scenario based on Asymmetrical High-speed Digital Subscriber Loop Technology

Advanced digital technology enables traditional telephone networks based on twisted copper pairs to be used for broadband services. The key enablers are digital transmission systems which allow bitrates in the range of 1.5–6 Mbit/s over ordinary telephone cables. The European Telecommunications Standards Institute (ETSI) has standardized a High-speed Digital Subscriber Loop system (HDSL), which can be used for symmetric transmission of bitrates of up to 6 Mbit/s over 1–3 ordinary copper pairs.

This transmission system would enable network operators to extend their services offered in the telephone network from ordinary telephony or narrowband ISDN to broadband services including such as high quality video conference and interactive video services.

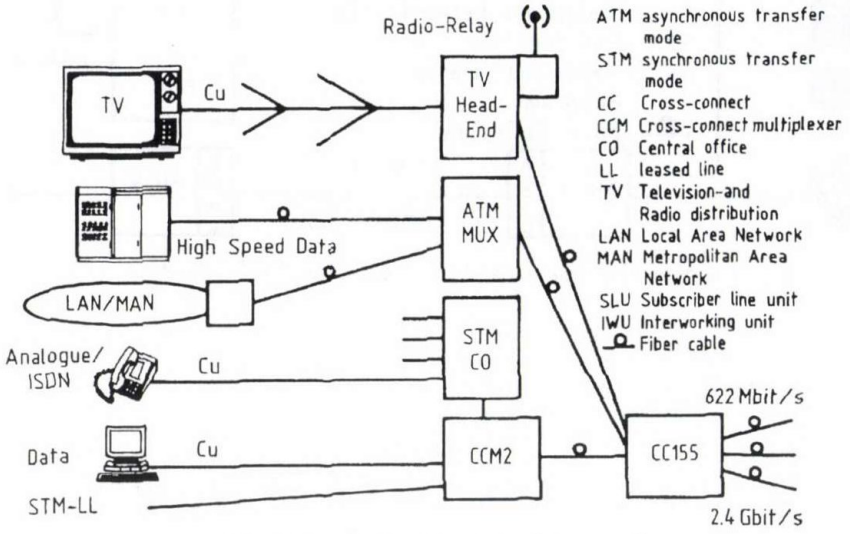


Fig. 1. Scenario based on an overlay network

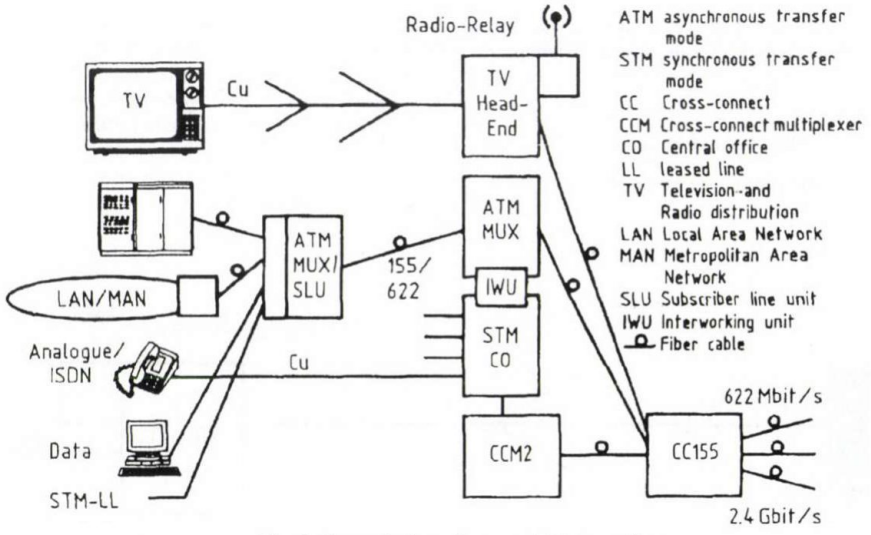


Fig. 2. Scenario based on partial integration

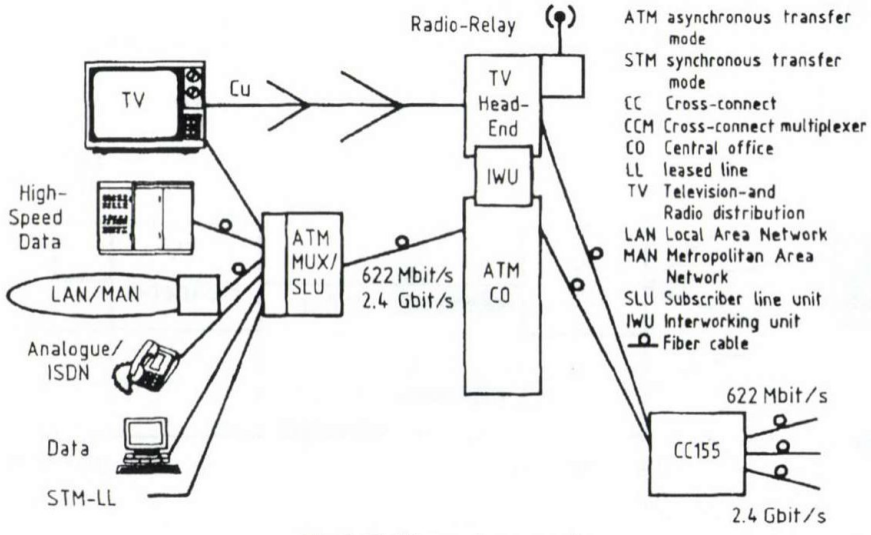


Fig. 3. Full integration scenario

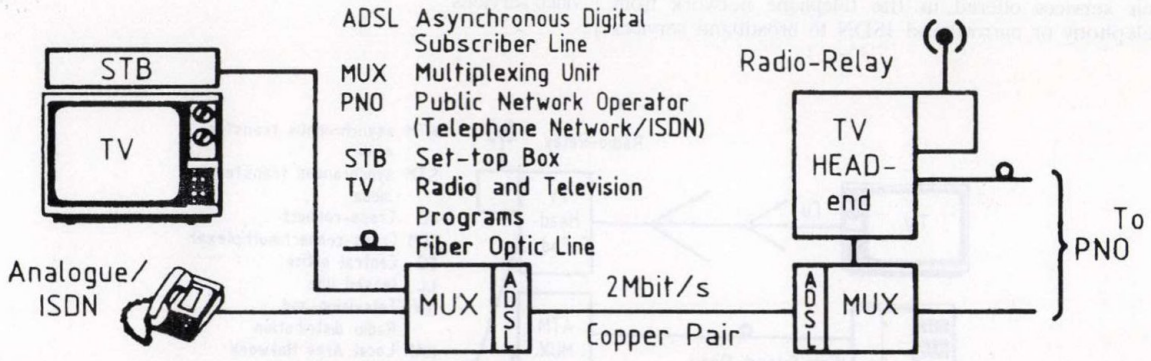


Fig. 4. Network Scenario with Asynchronous Digital Subscriber Line

Especially for interactive video services such as video-on-demand the Asymmetric Digital Subscriber Loop system (ADSL) has been developed. It supports the combination of traditional telephony or narrowband ISDN (symmetric channel) with a unidirectional video channel of 1.5–2 Mbit/s plus a bidirectional control channel. ADSL systems are currently implemented in some interactive video trials (Bell Atlantic, BT, Deutsche Telekom). A typical network scenario is shown in Fig. 4.

However, both systems, ADSL and HDSL are expected to play only an intermediate role until fiber in the loop systems are available. Their major drawback is, that they can only provide one video channel at one time. Therefore they can only be seen as an add on to e.g. existing cable-TV networks or satellite networks for broadcasting services.

4.3. Cable-TV based scenario, hybrid scenario

Cable-TV networks have typically a tree structure, i.e. their architecture is completely different from traditional telecommunication (telephone) networks. Consequently there are some restrictions imposed when such networks will be converted into multi-service networks including interactive communication services such as telephony, data, etc.

However new fiber-technologies combined with digital add-drop multiplexers allow cable-TV networks to be upgraded for the provision of interactive services. Fig. 5 shows an example of a network evolution scenario initially based on a coaxial cabel-TV network. In a second phase provisions are made to provide interactive telecommunication services as shown in Fig. 6 (initially limited to narrowband services such as e.g. telephony). In a last phase the network is completely converted into a fiber network (Fig. 7) offering more or less the same services as a broadband ISDN approach described in section 4.1. above.

Hybrid networks (combinations of fiber in the loop systems with either coaxial cable, systems or twisted pair systems may play an important role in the near future. In such hybrid scenarios fiber cables serve a small local area (200–500 homes) where the link between the central office or the head-end station and the distribution cabinet in the street is provided via a fiber optical cable and the last 50–300 meters are equipped with coaxial cables or twisted pair cables.

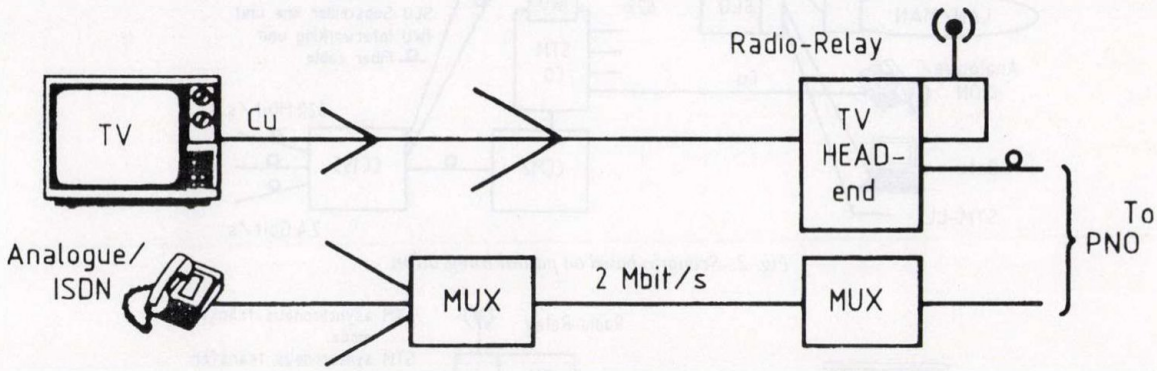


Fig. 5. Network scenario based on a TV-Distribution network

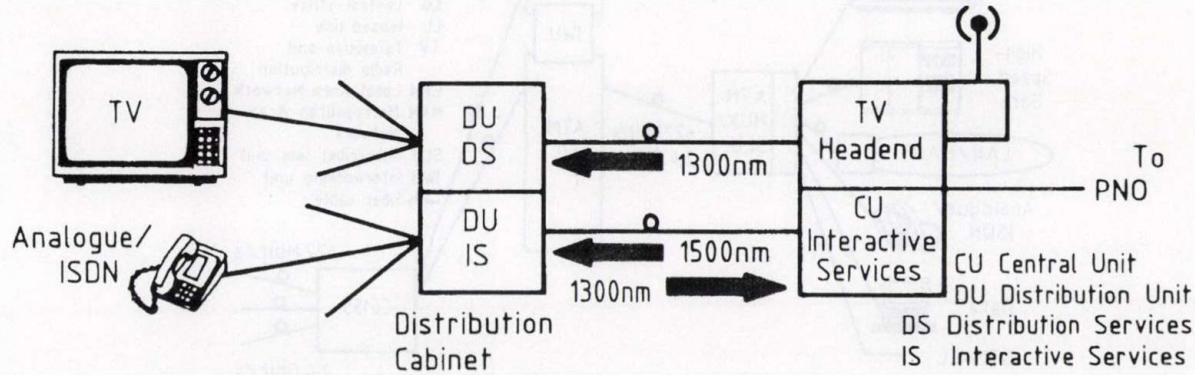


Fig. 6. Network scenario with partial integration of distribution services and interactive services

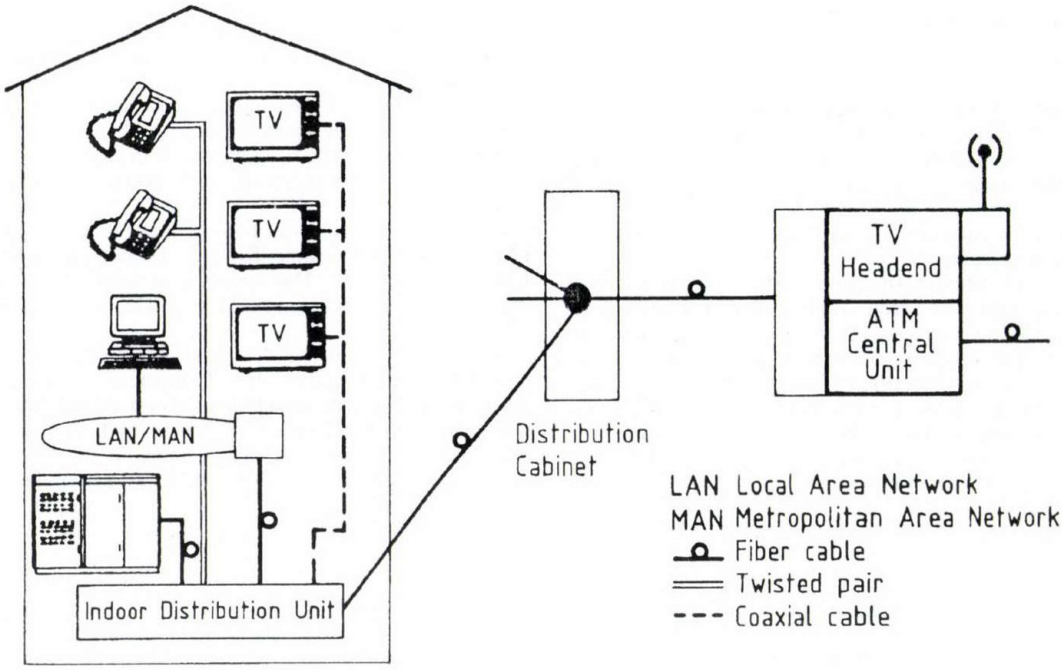


Fig. 7. Scenario based on a fully interated fiber network

5. INTERACTIVE VIDEO TRIALS OF DEUTSCHE TELEKOM

In order to gain experience with the new technologies Deutsche Telekom currently establishes a number of technical trials and market trials. In these trials different technical platforms for interactive video services will be tested. The different technologies for these platforms include different transmission systems (fiber, coaxial cable, twisted pair with ADSL transmission system), different server technologies, different set-top box functions, different architectures.

6. CONCLUSION

With the clear trend towards an information and communication society the telecommunications and information industry will play an increasing role in the coming years. Future telecommunications networks must be extremely flexible. Variable information transfer rate will be standard. As, on the other side of the medal, all industries and the whole human society will depend upon efficient information processing and transmission, these future telecommunications networks must provide an extraordinary performance and reliability. This requires that the network operators increasingly consider questions concerning fault tolerant network architecture, data integrity, safeguarding against unauthorized access to information, privacy. With today's technologies and their expected improvements in the coming years, the above described requirements can be met. Networks can be implemented which serve our society in an effective and efficient way. ATM based services will play a key-role in this direction.

JOACHIM CLAUS
DEUTSCHE TELEKOM
Generaldirektion
Friedrich-Ebert-Allee 140, Bonn, Germany

■ NEW RESEARCH LAB FOR ATM

To define the place and the role for research work in the society is a very sophisticated task. If we try to find the proper answer to the above problem I think the only possible generic solution is that the development and evolution of mankind is the target of every researcher. The Science of History would like to help avoiding mistakes made by our ancestors, scrutinizing medical problems reduces the danger of diseases and epidemics. They are supporting the whole humanity. Therefore they are financed by government or foundations.

There is a quite different situation in the field of technical sciences. In this field a research result earlier or later can be realized in a new product or service. The research combined with development and production enhances the success of a company. Therefore it is the primary interest of every company to own new technical solutions. To possess as many new methods, components or systems, as possible motivates the management to maintain R & D Labs. The companies are hoping that the staff of the Lab will achieve a great amount of novelties ensuring a good position of the company on the world market.

In this century these Laboratories have demonstrated that their results are used not only to enhance the profit of the company but sometimes to improve our world concept too. The ambitious leader and the research oriented environment helped the talented research fellows to become world famous scientists. There are well-known examples in telecommunications, the Bell Lab, the Post Office Research Station (Dollis Hill) or the industrial Labs of Bell Antwerpen, Siemens, Nippon Electric, Corning Glass, etc. In Hungary we also had a prominent research place the Bródy Lab supported by the Tungsram Company and hallmarked by Zoltán Bay and Károly Simonyi. Later in the 1960-s in the BHG there was a research and development team producing the world's first program controlled rural switching system put into operation on the north shore of lake Balaton.

In the last 30 years there was no place in Hungary being able to produce new results of worldwide significance. Appreciating the knowledge and talent of the Hungarian engineers we must confess that the environment did not facilitate the high-level research. There were some trials at the University, and at the Post Office Research Station that have nice traditions but the embargo and the financial difficulties suppressed them. The political changes made it possible that in cooperation with the most advanced Universities, Laboratories and industrial research teams we could join the world leading research projects of telecom companies.

Several places began to make use of this process. One of them was the cooperation between the Department of Telecommunication and Telematics (BME-TTT) of the Technical University of Budapest and the Swedish Ellemtel. Later Ellemtel was bought by LM Ericsson and became Ericsson Telecommunication Systems Laboratories (EUA). The cooperation resulted in new methods in planning and managing ATM networks, which deserved merit all-over the world.

These experiences led the management of Ericsson to decide to establish in Hungary a new Competence Center, the Traffic Analysis and Network Simulation Laboratory (TRANSLAB) whih will be part of the Hungarian Ericsson Távközlési Kft. The task of TRANSLAB, beginning its operation on the first of February is to scrutinize the traffic load of the ATM, traffic simulations, and network management and traffic throughput optimization. TRANSLAB will serve as a central place for the whole Ericsson world and will be responsible for traffic research on the ATM field. ATM makes the transmission of all kind of information possible including speech, computer signals, video and sound programs. It will be the key to the information high-way. So the results can influence the structure of the world telecom network.

The research fellows will be canvassed from the BME-TTT High Speed Network Laboratory where several students will acquire their master's and Ph.D. degree. These young experts have published their results at different international symposia and in dignified well established journals. The reception of their work was very good. We wish further good success for the new laboratory and for the young experts both in TRANSLAB and HSNL. We hope that their names will be well-known and they become appreciated members of the society of the telecommunication R & D workers.

GYÖRGY LAJTHA
MATÁV Hungarian Telecommunications Company
H-1013 Budapest, Krisztina krt. 55., Hungary

JOURNAL ON COMMUNICATIONS IN 1996

In 1996 we are continuing the practice of publishing English and Hungarian issues alternately. English issues will be devoted to significant technical events in Hungary and to Hungarian activities of international interest. The Hungarian issues are surveying important topics related to our domestic developments.

Topics of the planned issues are the followings:

English issues:

- ATM networks I-II
- Sensors
- Access networks
- Military electronics
- Broadcasting in Hungary

Hungarian issues:

- SDH
- ISDN
- Intelligent networks
- Internet
- Informatic infrastructure
- Telecom operators

Subscription rates:

1 year 12 issues	180 USD
6 English issues	100 USD
single issues	24 USD

Subscription orders should be sent to the circulation manager:
Klára Gerencsér
Fenyő u. 1. H-1016 Budapest
Tel: (361) 175-7147; Fax: (361) 201-7471

CECILIA UEBEL BLOMGREN, Manager, Business Development ATM Broadband, Ericsson Telecom, Sweden.

"Broadband technology makes it even more important to be focused."

The new broadband technology is changing the shape of communication between people. Soon we will all have free access to voic data, text, image and video via public networks. 24 hours a day. At home and at work - anywhere in the world. How we use this limited only by our imagination.

Ericsson recognised the need to harness this vast potential early on. We invested heavily in research. And out of our vision of global information society has come technology that serves the individual's needs, while helping our customers open up new bus ness opportunities. Ericsson's approach to broadband communication is comprehensive. At its core is an ATM broadband system for all sizes switching, transport and access nodes. This system permits the rapid, economic development and introduction of new variable bandwidth, mult media, interactive broadband services for business and residential users.

Ericsson's 80,000 employees are active in more than 100 countries. Their combined expertise in switching, radio and networking makes Ericsson world leader in telecommunications.

It's about communication between people. The rest is technology.

Telefonaktiebolaget LM Ericsson, S-126 25 Stockholm, SWEDEN.

ERICSSON