Analytical Models for the Multiplexing of Worst Case Traffic Sources and their Application to ATM Traffic Control.

José M. Barceló Ordinas

March 5, 1998

Analytical Models for the Multiplexing of Worst Case Traffic Sources and their Application to ATM Traffic Control.

José M. Barceló Ordinas

This thesis has been written at Polytechnic University of Catalonia, Computer Architecture Department. c/ Gran Capitan, Modulo C6-E105, Barcelona E-08071, Spain.
tel : + 34 3 4016798, fax : + 34 3 4017055, e-mail : josebac.upc.es

This thesis has been promoted by:

PhD advisor:	Dr. Jorge García Vidal	Universidad Politecnica de Cataluña. Spai
PhD Tutor:	Prof. Dr. Olga Casals Torres	Universidad Politecnica de Cataluña. Spai

Members of the PhD thesis tribunal:

Prof. Dr. Chris Blondia	University of Antwerp, Belgium
Prof. Dr. Jorma Virtamo	Telekniikan Laboratorio, Finland
Prof. Dr. Vicente Casares Giner	Universidad Politecnica de Valencia, Spain
Dr. Josep Paradells Aspas	Universidad Politecnica de Cataluña, Spain
Prof. Dr. Ramon Puigjaner Trepat	Universidad de les Illes Balears, Spain

Contents

List of Figures

List of Tables

Preface.

This document gathers the work of my PhD during the last years. In this time I have collaborated in fourth European projects: EXPLOIT, BAF, EXPERT and NETPERF and I have been introduced to give lectures at University in the areas of basics on operating systems and basics on network computers.

With the defence of my PhD I end my fourth life cycle as a student. In every one of these cycles: primary school, secondary school, University and PhD, I have known many people and learnt from their experience and friendship. Now, It is time to acknowledge those ones that have belonged to this last cycle of my life.

Of course, first of all, I must thank my family. In my Mediterranean education family is a concept that involves not only my parents, brothers and sisters but also a set of relatives with whom I have a close relation. I would like to thank them all for supporting me during all the years of my life. They make myself feel sure.

A very special acknowledge to Olga Casals and Jorge García for not only introducing me in the ATM world, but also for all the knowledge, help, advise, support, guidance, etc. they have given me in these years. It has been a pleasure to work with them.

With Fernando Cerdán and Llorenc Cerdá I have shared a lot of interesting moments at University. With Llorenc I have shared the work-room for the last two years and a lot of interesting discussions. The same with Fernando. Furthermore with Fernando and Mikel Uria, I have visited a good percentage of restaurants, pubs, discos, bars and other "tugurios" in Barcelona. Well, Fernando and Mikel, there are still a lot of places to "visit".

A special regard to Frank Kloster and Aimee Figueroa for their friendship and all the wonderful moments we have had in Barcelona. I hope that when they are far from Barcelona they will remember our friendship, although if they definitively move out to Mexico I will go from time to time. Also a kind regard to very different people that I have met or with whom I have lived in the last years: Nuria, Rosa, Cristina, Elena, Ingmar, Stephan, Katy, Masae, Gemma.

When I arrived to the department, I was confined in a big room with several students. All of them were working in projects. We call ourselves the "008 room", and spent a lot of wonderful moments. Regards to Maite, Luz, Merce, Susana, Cris, Enric, Quique, Luis, Gerard, Toni, Andres, Joan Manel, Pepe and "afines". Also those ones that did not belong to the "008" but that joined us as Josep Ramon or Jordi G. Sorry If I have missed someone.

I would like to thank all adjuncts and associates that have helped me in these years I have been giving lectures, specially to Enric Morancho who help me in a subject so alien

to me in that time as OS. Also to those ones who help me in CBXC. To Marga Grimalt who programmed in her PFC the scheduling simulators.

I spent my first two years working in the EXPLOIT project. My work consisted in a set of trials on ATM Traffic Control. We had to travel a week per month in average to Basel. I would like to regard all the people I met in this project: all people of ASPA (Manuela, Gabi, Max, George), people of other WPs as Thomas, Chris, Anne-Marie, Martin, Harald, Stephane, Dani, Steven, Laurie, and others that I can not remember. There were so many people in that project. I don't forget the people of my own WP: Vassilis Nellas, Antonio Martinez, Nicolas Mitrou and specially to Laurent Jaussi with whom I spent a lot time working, eating and playing sometimes squash in Basel. We also spent a lot of weekends working together, Laurent at EPFL and I at UPC. Hard days trying to get something from that platform, but we also had a lot of fun.

Other project in which I knew very nice people was BAF. We had very interesting days plenty of work and fun in Milan with Frans Panken, Babul Miah and Steven Winstanley. Thanks also to Hans Boekhorst in ATT Huzen and Luigi Capprioti in Milan for the technical and human support in the experimental work in this project.

In projects as EXPERT and NETPERF I have also met very interesting people to whom I acknowledge, although my involvement in these projects have allowed me only to met them from time to time.

I thank to all referees that have given me some input when I have sent a paper to a conference and to the referees that have given me the different input and comments before the presentation of this thesis.

And finally, thanks to Professors J. Virtamo, C. Blondia, V. Casares, R. Puigjaner and Dr. J. Paradells, members of the PhD tribunal, for coming to the presentation.

Chapter 1

Introduction.

1.1 Moving towards B-ISDN.

Broadband digital networks (B-ISDN) are intended to transport all kind of communications services. As we move towards integration in telecommunication networks, Asynchronous Transfer Mode (ATM) has been proposed as the mechanism to transport data, video and voice in B-ISDN. The ATM technique offers switching based on statistical multiplexing to any kind of digital communication services. Typical rates in ATM are 155,52 and 622,08 Mbps, although lower and higher rates can be utilized. In ATM, information is transmitted in the form of fixed packets called cells. The different communication services are distinguished in the way their sources produced cells (e.g. at constant bit rate as those produced by circuit-switched channels and some video coders. at variable bit rate as those ones as variable video coders or bursty data sources).

In [84] (see also [4]), some guidelines about the nature of broadband nature are given in the form of a classification in three service categories: interactive communications as telephony, video-conferences, transmission of live events or circuit emulation: transfer of stored information for temporary storage as bulk data transfer, off-line transfer of movies, transfer of text or pictures in databases (WWW) and finally transfer of audio and video for immediate playback. This classification is based on the timeless requirements of user applications and does not cover all kind of applications as distributed computing. Furthermore, there is not a clear mapping between these service categories and the ATM transfer capabilities proposed in the ITU-T 371. [1].

B-ISDN standards define the ATM layer above the physical layer and below the ATM adaptation layer (AAL) that relates upper level layers with the ATM layer. ATM makes use of statistical multiplexing techniques in a slotted medium. Since connections may share links and buffers in a given route, cells may suffer delays or losses that may affect the way Quality of Service (QoS) objectives requested by the connections are met. QoS objectives play an important role in ATM. Somebody may think that integration is only achieved through statistical multiplexing techniques. However, due to the heterogeneous nature of the traffic, it seems logical that different communication services ask for different QoS requirements. For instance, real-time interactive communications as voice and video are more sensitive to cell delay than to cell losses while transfer of data information may be more sensitive to cell losses than to cell delay. Achieving integration is then

linked with the idea of keeping at the same time QoS objectives that are very different from one service to other. Furthermore, a network operator would like a network able to transport the maximum number of connections possible, it is to say, maximum utilization factors in each network node. Here, we have three ideas that make one to think: How is it possible to integrate services with so different traffic characteristics and different QoS requirements obtaining at the same time maximum network efficiency ? A short old question not still fully answered. That is one of the reasons of so many ATM studies in the last years.

There are so many aspects to be studied in ATM networks that it is difficult to begin to describe what has been or is being investigated in ATM: ATM switches, multiple-access protocols, ATM LAN emulation, ATM Adaptation layer protocols, Traffic Control and Congestion Control Functions, multicast, IP over ATM, etc. This thesis is centered in concepts used in Traffic Control and Congestion Control Functions. Therefore, we will focus in this area assuming that the reader knows the basics ideas behind ATM. For those ones who begin in ATM we recommend ATM publications as [47], [53] or [60].

1.2 Controlling the ATM traffic.

Traffic parameters, tolerances and Quality of Service requirements are essential in the specification and design of Traffic Control and Congestion Control Functions such as Connection Admission Control (CAC), Usage/ Network Parameter Control (UPC/ NPC), Network Resource Management (NRM) or traffic shaping among others. For instance, the CAC must be able to accept or reject connections meeting the QoS objectives (e.g. a Cell Loss Ratio) required in the ATM Traffic Contract or the UPC function has to monitor a connection based on the specified traffic parameters and tolerances (e.g. a peak cell rate and a cell delay variation tolerance). On one hand, these parameters have to be easily understandable by the end user. On the other hand, they must be enforceable by the Control Functions and participate in resource allocation schemes. The ITU-T 371 specifies a set of traffic parameters and tolerances in order to describe the source behavior and in order to be easily used by the Control Functions: the peak cell rate, the cell delay variation tolerance, the sustainable cell rate and the burst tolerance are among them.

An important question is how the traffic characteristics of the sources can be fit in these parameters. As an example, we can think in a Constant Bit Rate connection (e.g. a periodic source) that defines in its Traffic Contract a Peak Cell Rate (PCR). The CAC will admit this connection based for instance on a peak allocation scheme (the sum of the peak cell rates of all the connections sharing the same link has to be less than the link channel capacity). Since the cells of several connections can arrive at the same time to the access multiplexor, a buffer able to cope with the cell level congestion is needed to meet a QoS based on cell loss ratio. For that purpose the size of the buffer could be designed using queueing models as the $M/D/1$ queue system. Therefore, this CAC would use a source traffic parameter and a QoS objective as input data to decide the acceptance of the connection. In this case the source is characterized by its peak cell rate.

As another example we can think on the UPC mechanism and the Constant Bit Rate connection. The function of the UPC is to monitor that the traffic sent by a connection is according to its Traffic Contract, the PCR used in the CAC algorithm. For that purpose. the UPC function must enforce the defined Traffic Contract parameters and may use cell tagging or cell discarding mechanisms to mark or drop those cells that does not conform with the declared Traffic Contract. However, there are many factors that may affect the Constant Bit Rate connection making that the ingress CBR traffic is not a pure periodic traffic: AAL mechanisms, physical layer interfaces as SDH or E1, access networks or previous multiplexing stages are some of the factors that may introduce jitter (called in ATM Cell Delay Variation) in a source. To avoid the tagging or dropping of cells of a connection that has suffered some jitter, the source declares in the Traffic Contract a Cell Delay Variation tolerance (CDV-T) that will be taken into account by the UPC mechanism. But now, due to this tolerance, there are many kind of traffic patterns that conform with the UPC definition and that have different congestion effects in the switches: pure CBR periodic traffic, CBR jittered traffic or a burst of back-to-back periodic cells can be conforming to the UPC mechanism. In fact, a CBR connection is not necessary a pure periodic source, and can be a source that transmits as a maximum cell rate its PCR. The larger the CDV tolerance is defined the higher the number of cells that will be conforming. even if the traffic pattern is very different to that expected. For instance, assuming a periodic CBR source with declared peak emission interval (the inverse of the peak cell rate) equal to 10 cell slots and a declared CDV tolerance of 100 cell slots, a leaky bucket as a UPC function might allow pass as conforming either a periodic traffic of 1 cell every 10 cell slots or a periodic traffic of 12 back-to-back cells every 120 cell slots. The congestion that would produce both kind of sources in a buffer that multiplex a given number of connections (being the system stable) is very different.

1.3 Outline.

In this thesis, we investigate the effect of multiplexing in a buffer CBR sources that have declared a peak cell rate and a CDV tolerance but that send different traffic patterns. For that purpose we have chosen a worst case traffic definition that fulfils the Traffic Contract and that it is enforceable by the UPC mechanism.

Of course, assuming traffic sources that emit back-to-back cells, we are considering a pessimistic picture of the network. We remit to the ITU-T 371 recommendation that literally says: "When allocating resources, the network should take into account the worst case traffic passing through UPC/NPC in order to avoid impairments to other ATM connections. The trade-offs between UPC/NPC complexity, worst case traffic and optimization of network resources are made at the discretion of network operators". Assuming that all connections can send worst case traffic at the same time can be taken as too pessimistic. However. the study of a worst case situation can give us an insight of how expensive in terms of networks resources is to have a safe CAC and the complexity of resource allocation even for the multiplexing of CBR connections.

The thesis has the following structure:

- Chapter 2 begins with an overview of which parameters and procedures defined in the Traffic Contract are necessary to operate resource allocation schemes and Control Functions. The ITU-T 371 and the ATM Forum, [1] and [3], specify these parameters. In this chapter the ATM transfer capabilities, the Connection Traffic Descriptor and the set of QoS parameters that an ATM connection may negotiate at call set up are outlined. An algorithm called Generic Cell Rate Algorithm (GCRA) gives the cell conformance of a connection respect to the cell rate and a tolerance. The chapter ends with a description of models proposed in the COST 242, [4], to quantify the Cell Delay Variation of a CBR connection crossing either a single queue system or a tandem of queues.

- Chapter 3 describes how Cell Delay Variation is introduced in real networks. For that purpose. a set of experiments performed in two ATM Testbeds where delay distributions were measured are presented. The first Testbed is an ATM optical access network belonging to the BAF Testbed (BAF: Broadband Access Facilities, RACE project 2024). The experiments obtained resulted in the writing of a deliverable. [9]. where all the partners participated. The results presented in this chapter are an extract of [9] that was published in [72]. The second example is a set of experiments performed in the EXPLOIT Testbed in Basel (EXPLOITation of an ATM Technology Testbed for Broadband Experiments and Applications, RACE project 2061). together with the Swiss ATM pilot network as a part of the Pan European ATM pilot Network (PEAN) and the Telecommunications Laboratory of the Ecole Polytechnique Fedérale de Lausanne (EPFL). Results on experiments performed in the EXPLOIT Testbed centered in CDV can be found in [14], [57] and [58]. The experimental results are compared with some of the models to quantify CDV described in chapter 2. The contributions to the thesis in this chapter are the experimental results together with the experience gained in how to set up a complex network configuration in a real network.

- Chapter 4 introduces the Benes approach to the Virtual Waiting Time as a tool that can be used in a wide range of queuing models. This approach expresses the *Virtual Waiting Time*. V_s, of a G/G/1 queue and gives an upper bound of the queue length. In this chapter some of the most used models in ATM networks using this approach at cell level are described: the M/D/1, the ND/D/1, the modulated ND/D/1 and the M+D/D/1. The chapter finishes with other analytical tool presented by *P. Humblet et al* in [55] based on the Ballot Theorem of *Takacs*.

- Chapter 5 defines the concept of *Worst Case Traffic* (WCT) as that traffic pattern that being conforming to the UPC requires the biggest amount of network resources. An exact analytical model based on the Benes approach that calculates the Complementary Probability Distribution Function of the Virtual Waiting Time for the multiplexing of periodic WCT is presented. The mathematical model solves the discrete-time queue system and the fluid-flow approximation system. A set of results is provided. First the WCT model is evaluated and compared with upper bounds proposed in [78]: a batch model and a bound proposed by Ramamurthy and Dighe. The allocation of buffer resources and admissible load taking into account

WCT conditions is also studied. Finally, the analytical model is used to analyze the impact of Cell Delay Variation on resource management in a real network. Taking the EXPLOIT testbed experiments and using a worst case situation, a bound on the maximum number of cells in a burst can be computed. An interpretation of the impact of worst case traffic in the buffers is given, concluding that the high peak bit rate sources are the most dangerous. The analytical model and some of the examples can be found in [44] and [45], the last example can be found in [14], [57] and [58]. The contributions to the thesis in this chapter are the discrete-time and the fluid-flow model to analyse the superposition of WCT sources. I also give a set of examples of the impact of WCT sources in a multiplexor and I compared the exact model with two bounds that up to this moment were in the literature. I also comment the limitations of the exact model.

- Chapter 6 compares some WCT models using several traffic inputs: WCT bursts Poisson distributed, CBR multiplexed with periodic WCT, or Poisson traffic multiplexed with WCT bursts Poisson distributed are some of the models studied. Some of these analytical models are exact while others are approximations that behave better or worst depending on the traffic parameters. The models are reported in [19]. The contribution of this chapter consists of a set of models to overcome some of the limitations of the periodic WCT model: homogeneity and computational complexity when the number of sources and the burst size grows.

- Chapter 7 studies what happens when periodic WCT or CBR traffic is multiplexed in networks with tree topology of rooted queues. To solve this kind of configurations, we derive closed-form formulas for the queue length distributions in a discrete-time M-stage tree queueing network loaded with periodic traffic sources. We obtain expressions for the CPDF of the Virtual Waiting Time in any queue of the tree. We also calculate the average queueing delay and average waiting time in any queue of the tree network. Results of this chapter can be found in the BC'98, [21], and as interim reports in [22] and [23]. The contribution of this chapter consists of the solution of a two stage network with a tree topology. The extension to a Q stage network is based on previous works found in the literature. However, this extension is based on the two and one stage systems. Therefore these are the important cases to study.

- Chapter 8 is dedicated to resource allocation. There are several solutions to cope with WCT sources. The first is to design the network resources in terms of buffers to face with a worst case situation. However, this solution could introduce high delays in sources that have heavy delay constrains as real-time services. Other solution could be to use traffic shaping techniques. We briefly describe some of these techniques and study the trade-offs in the design of a CAC combined with a traffic shaping mechanism. Finally, scheduling algorithms can be used to smooth the traffic and provide QoS guarantees. Parts of this chapter can be found in [15] and [20]. The contributions of this chapter to the thesis consist of an extension of the $WCT/nD/1$ queue system to study the spacing of WCT sources on a VP. We also propose a study combining CAC and traffic shaping. The part of scheduling algorithms have been taken from the literature, simulating those algorithms that we have considered more important. The idea is to complete the study of WCT

sources and their impact in other connections such as CBRs.

- Chapter 9 is devoted to general conclusions and comments.

As can be seen, the thesis has a mixture of simulations, experimental work and analytical models. The advantage of simulating is the freedom in changing variables, design parameters and the scalability of a simulator. The simulators have been programmed in C and executed in workstations running UNIX OS. The input traffic is modeled as statistical sources that generate cells following a probabilistic function. These functions make use of random numbers generated by the computer. A major problem in simulating is that the simulation has to run a long number of cell slots to obtain small probabilities. For instance to get a cell loss probability of $10^{-\epsilon}$ in a Virtual Connection (VC) is recommended to generate a multiple of 10^{ϵ} cells on that VC (e.g. 100 times 10^{ϵ}). Furthermore, one has to generate cells belonging to other VCs. That leads to run long simulations if low probabilities are required. Simulations have been performed taking confidence intervals of 95 %.

The experimental work has been performed in two Testbeds: BAF and EXPLOIT. Some drawbacks were encounter when the experiments were carried out. The first was the newness of the experimental platforms. At that time most of the equipment was not still full commercial and came from prototypes, being sometimes unstable. Therefore, the experiments were performed at the same time that we learned how to fix bugs. Our experience in the platforms make us conclude after some time that changing dramatically the configuration of the network meant a lot of software crashes in some equipments as the traffic generators/analyzers. Hopefully, we had full support from the manufacters. An experimental session in EXPLOIT consisted of a week's work every month for each Work Package. During the two years I worked in the project, my Work Package dedicated the last half of the second year to perform experiments related with jitter and its impact on resource management. Thus, in this thesis I will only refer to these specific experiments.

As I mentioned, we tried as much as possible to keep always the same equipment and lines configured. Since the equipment available to perform experiments were limited, mainly the traffic generators, we had to make up some ideas: obtain phase-moving sources, recycling traffic, duplication of traffic using optical splitting devices and load balancing when multiplexing in the buffers were some of the solutions that we applied to go on with the experiments session after session. Of course, the network configuration become rather complex and the number of traffic assumptions were quite high: correlated traffic was a result on these assumptions. However, I think that we minimize as much as possible the correlation effects achieving a fair network configuration with the definition of an experimental criteria or rules that we followed rigorously. These conditions are outlined in chapter 3. In the experiments performed in the BAF Testbed, we again encountered many of these problems. For instance, we had only a traffic generator and we had to use again optical splitters. To solve the mentioned correlation effects, the traffic at T_b interfaces was policed at different values. In this way, the VPs lost some background traffic due to cell dropping. Of course, incorrelations were not fully eliminated but at least minimized.

Finally, we have developed analytical tools to evaluate the impact of worst case traffic in

buffers. Once the conditions and assumptions of the model are defined, a mathematical tool is applied to obtain the analytical model. The advantage of an exact analytical model is the degree of accuracy obtained respect of a simulation. The disadvantage is the difficulty to modify the analytical model if we want to change some condition or assumption. In this case a simulator provides more manouvrebility and scalability.

As a disadvantage, analytical models also have their range of applicability and computational complexity. For instance as it is stated in [91], see chapter 4, the $M/D/1$ queue system obtained with the Benes approach suffers of numerical loss accuracy due to cancellation of terms under certain traffic conditions. Virtamo applied a mathematical method to solve this problem based on a polynomial representation. However, not always it is possible to solve this kind of problems. In our case, when we have applied Poisson processes as input traffic, we have had to cut summatories. In some cases we could apply the same method as Virtamo. In the cases that it was not possible we have compared the analytical model with simulations.

Chapter 2

Traffic Contract and Cell Delay Variation in ATM networks.

This chapter describes functions and parameters for traffic and congestion control in ATM. The objective is to introduce the ATM Traffic terminology that will be used through the thesis. The ATM Traffic Contract specifies the connection traffic descriptor together with the ATM transfer capabilities and Quality of Service objectives. The main goal of Control and Congestion Functions in ATM is to protect the user and the network in order to achieve network Performance objectives. The different Control Functions are briefly outlined. Cell conformance of a cell stream to negotiated parameters is obtained making use of a reference algorithm. Finally the chapter summarizes different models to account for Cell Delay Variation.

The chapter is organized as follows: section 2.1 is dedicated to the Traffic Contract and ATM Transfer Capabilities and Quality of Service parameters. Section 2.2 describes the ATM Traffic Control and Traffic Congestion Functions. Section 2.3 defines the Generic Cell Rate Algorithm for cell conformance and how can be modeled as a $G/D/1$ queue system. The concept of Cell Delay Variation is introduced together with the clumping and dispersion effects. Finally, section 2.4 is referred to models to quantify the CDV in ATM networks (single queue systems and tandem of queues) and the parameters that influence the choosing of CDV tolerances are outlined.

2.1 Traffic Contract and ATM Transfer Capabilities.

A traffic contract is specified at connection set-up. This contract specifies the negotiated characteristics of the connection. At the public UNI (User Network Interface), the traffic contract consists of:

- The selected ATM Transfer Capability.
- A Connection Traffic Descriptor: Source Traffic Descriptor (e.g. Peak Cell Rate, PCR, Sustainable Cell Rate, SCR, Maximum Burst Size, MBS, Minimum Cell Rate, MCR), the CDV tolerance and the conformance definition.
- A set of Quality of Service (QoS) parameters (e.g. Cell Transfer Delay, CTD, Cell

Loss Ratio. CLR, 2-point CDV, ...) classified in classes by ITU-T 356, see [2] (e.g. stringent class, tolerant class, bi-level class and unspecified class).

- The setting of the tagging option.

Note that the CAC (Connection Admission Control) and UPC/NPC (Usage/Network Parameter Control) may use the parameters described by the traffic contract to operate their procedures and achieve the best network resource utilization (CAC and UPC/NPC procedures are network operator specific). Therefore the interest in describing the most important parameters than can be used in these procedures.

ATM Transfer Capabilities.

ATM transfer capabilities (ATC) are defined by the ITU-T 371, [1], as a way to specify a combination of QoS commitments and ATM traffic parameters into a set of traffic procedures suitable for some ATM applications and that allows efficient resource allocation. It is mandatory that the ATC used by an ATM connection be implicitly or explicitly declared at connection set-up.

ATM transfer capabilities are called service categories in ATM Forum terminology, [3]. Some of the ATC are equivalent to the ATM Forum service categories but with other names. However a few discrepancies stay, see [3] for a mapping between ATC and ATM Forum services categories.

- **Deterministic Bit Rate (DBR):** is used by connections that request a static amount of bandwidth characterized by the peak cell rate (e.g. voice. video, emulating circuit switching facilities). At connection set-up, a user can negotiate one of the following three source traffic descriptors (PCR for user data CLP=0+1 and PCR for end-to-end user OAM cells, PCR for user data CLP=0+1 or the service type (e.g. telephony or video-phone)). For each PCR is mandatory to declare a CDV tolerance. Note that the DBR capability is not only defined for CBR applications but also for those ones that may not continuously transmit at the negotiated PCR but wants a negotiated contract (e.g. in terms of QoS) compatible with the DBR capability.

- **Statistical Bit Rate (SBR):** is used for connections that want to send at a variable rate obtaining multiplexing gain. SBR uses the SCR/IBT parameters together with the PCR and CDV tolerance to describe in great detail the traffic flow being transmitted by the source. The SBR capability can also specify as source traffic descriptor the service type.

  The ITU-T defines three kinds of SBR capabilities, SBR1, SBR2 and SBR3, depending on the Cell Loss Priority (CLP) bit and the possibility of using tagging options.

- **ATM Block Transfer (ABT):** In this capability the transfer characteristics are negotiated on an ATM block basis. Once an ATM block is accepted, the network allocates resources according to the QoS received to that block, as it would do with a DBR connection with the same peak cell rate and QoS. The user negotiates for each direction the PCR and CDV tolerance and the pair SCR/IBT.

  Two ABT capabilities are specified: the ATM Block Transfer with Delayed Transmission (ABT/DT) and ATM Block Transfer with Immediate Transmission (ABT/IT). With DT the peak cell rate of successive blocks is negotiated dynamically with the network making use of Resource Management (RM) cells. In the IT definition, the user sends an ATM block without the positive acknowledge of the network. If there are not enough resources available in the network when transmitting a block, this one can be discarded.

- **Available Bit Rate (ABR):** With this capability it is intended that the user adapts to the network characteristics upon receiving feedback from the network (reactive control). The feedback is conveyed to the source through special cells called Resource Management (RM) cells. With this idea in mind it is intended that all the available resources of the network are used providing moreover a fairly bandwidth share among users.

  At connection set-up the user specifies to the network a maximum cell rate (its PCR) and a minimum usable bandwidth (the Minimum Cell Rate, MCR, that may be 0). Then, the bandwidth available varies from the PCR to the MCR depending on the network conditions. The RM cells are in charge to inform the end user of the state of the network in order the user can adapt its cell rate. The QoS requirements are very loose, there are no Cell Transfer Delay or Cell Delay Variation commitments in ABR, and as far as the user adapts its rate as the RM cells specify it, the QoS on Cell Loss Ratio remains.

DBR capability is called CBR service category in the ATM Forum, SBR is called real-time Variable Bit Rate (rt-VBR) and non-real-time Variable Bit Rate (nrt-VBR), ABR has the same name, ABT has not equivalence, and finally, the ATM Forum defines two more service categories: Unspecified Bit Rate (UBR) and Guaranteed Frame Rate (GFR) recently proposed. The UBR service category is intended for non-real-time applications (e.g. file transfer or e-mail). The network does not make any commitment to guarantee QoS parameters as Cell Loss Ratio or Cell Transfer Delay to UBR connections. The network may or may not apply PCR to the CAC or UPC functions. To keep the simplicity of the UBR category, but at the same time to add some features as the possibility to reserve a minimum bandwidth per VC considering that the flow of cells are AAL5-PDUs, a new service category called Guaranteed Frame Rate has been proposed. This minimum bandwidth allows to GFR to provide certain guaranteed bandwidth, see [52] for implementations that support GFR service category in ATM switches.

Connection Traffic Descriptor.

The Connection Traffic Descriptor specifies the traffic characteristics of an ATM connection. It consists on the source traffic descriptor, the Cell Delay Variation Tolerance (CDV-T) and the conformance definition.

The source traffic descriptor is the set of parameters of an ATM source and is used to capture the intrinsic traffic characteristics of the connection requested by a particular source. Any of these traffic parameters may participate in resource allocation schemes and have to be enforceable by the UPC/NPC. As main parameters, we can mention the following:

- The Peak Cell Rate (PCR) of an ATM connection is the inverse of the minimum inter-arrival time (T_{PCR}) of two ATM cells, where T_{PCR} is called the Peak Emission Interval (PEI).
- The Sustainable Cell Rate (SCR) gives an upper bound on the conforming average rate of an ATM connection. SCR can be expressed in terms of its inverse T_{SCR}.
- The Intrinsic Burst Tolerance (τ_{IBT}) is defined together with the SCR and it is a parameter used by a traffic source within the customer equipment that accounts for the Maximum Burst Size (MBS) that can be transmitted at peak rate and still be in conformance with the GCRA, being MBS=$\lfloor 1 + \frac{\tau_{IBT}}{T_{SCR}-T_{PCR}} \rfloor$ cells.

The Cell Delay Variation (CDV) tolerance is intended to capture the cell delay variation or jitter caused on an ATM connection. The CDV tolerance, τ, is defined respect to the GCRA(T_{PCR},τ). The CDV tolerance defined with the T_{SCR} in the GCRA is the τ_{SCR}. In fact τ_{SCR} is composed by two terms: τ_{IBT} and an additional tolerance added at the customer equipment that accounts for the CDV introduced by multiplexing stages at the cell level and at the burst level called τ'_{SCR}. Thus $\tau_{SCR} = \tau_{IBT} + \tau'_{SCR}$.

The Generic Cell Rate Algorithm (GCRA) is used to define conformance with respect to the traffic contract. For each cell arrival, the GCRA specifies if the cell is conforming or not. In the case the cell is not conforming several actions can be carry out (e.g. cell tagging, cell discarding). See chapter 2 for a definition of the GCRA.

Quality of Service parameters.

At connection set-up, a user can specify its preferred QoS class (see ITU-T 356 [2] for QoS class definitions and network performance objectives). If the network can not support the requested QoS class it will clear the connection request. If not, the network operator commits to support the QoS bounds specified for each QoS class always that the end system complies with the negotiated Traffic Contract. There are four QoS classes: stringent class, tolerant class, bi-level class and unspecified class. Each class is defined as a combination of bounds on Network Performance parameters such as Cell Transfer Delay (CTD), Cell Loss Ratios (CLR), 2-point CDV, Cell Error Ratio (CER), Severely Errored Block Ratio (SECBR) and Cell Misinsertion Rate (CMR). Some of these bounds are unspecified for some classes. The ITU also studies the possibility to specify individual QoS parameters via signalling.

2.2 ATM Traffic Control and Traffic Congestion Functions.

ATM traffic congestion is defined by the ITU-T 371, [1], as the state of Network Elements in which the network is not able to meet the negotiated Network Performance objectives for a given established connection. As an example, we can assume a connection crossing an ATM switch and therefore being multiplexed with other connections in the switch shared buffers. Since the buffer is finite, we can find multiplexing situations due to the nature of the multiplexed streams (e.g. peak rates, burst lengths, activity factors, etc) for which the number of cells lost in our connection is higher than the negotiated QoS Cell Loss Ratio parameter. It is clear that there must be Control Functions able to distinguish situations where there may be cell losses due to buffer overflows and check up whether the negotiated QoS are met or not.

The ITU-T defines two types of Control Functions:

- **Traffic Control Functions**: defined as the actions taken by the network to avoid congestion. As Traffic Control Functions the ITU-T mentions the following:
  - Network Resource Management (NRM): makes use of Virtual Paths (VPs) e.g. to simplify CAC or facilitate the segregation of traffic types requiring different QoS.
  - Connection Admission Control (CAC): allows to determine if a Virtual Path/ Channel connection can be accepted or rejected given that the QoS requirements have to be met.
  - Usage/Network Parameter Control (UPC/NPC): allows to monitor and control the accepted connections in such a way that the parameters negotiated are not violated.
  - Priority Control and Selective Cell Discarding: the network may selectively discard cells of the lower priority or if allowed change the Cell Loss Priority bit to a lower priority (tagging option).
  - Traffic Shaping: alters the traffic characteristics of a connection to achieve a better performance: e.g. spacing cells or using scheduling mechanisms.
  - Fast Resource Management (FRM): is used to dynamically allocate resources to connections.
- **Congestion Control Functions**: defined as the actions taken by the network to minimize the intensity, spread and duration of the congestion.
  - Selective Cell Discarding.
  - Explicit Forward Congestion: is a mechanism that allows the network to recover/avoid of congestion states.

These functions operate for very different time-scales that can vary from the cell insertion time to the call/connection inter-arrival time.

2.3 Ruled based traffic parameters.

The most usual example of a DBR ATM source is one which transmits one cell at periodic intervals given by its Peak Emission Interval - PEI - (conversely its Peak Cell Rate - PCR), which is a parameter declared by the source at connection set-up within the Traffic Contract. In order to control congestion, the traffic that the DBR sources emit is monitored and controlled by a mechanism called Usage Parameter Control (UPC). The point at which the UPC mechanism controls the traffic should be at the network ingress. When the traffic reaches the said point it may have been distorted due to the use of ATM Adaptation Layers, due to multiplexing with traffic from other connections or by the use of medium access protocols in the network which give access to the ATM network. These phenomena produce a jitter in the cells that is known as Cell Delay Variation (CDV). In order to avoid cells from sources which emit traffic in accordance with the declared Traffic Contract being erroneously declared non-conforming by the UPC, the sources must include in its Traffic Contract a parameter known as CDV Tolerance (CDV-T) which quantifies the said phenomenon. The UPC mechanism must therefore take into account the CDV-T in order to avoid declaring cells from sources that fulfil the Traffic Contract to be non-conforming.

The cell conformance of a connection to the negotiated cell rate $1/T$ and the tolerance allocated τ is defined in respect of a reference algorithm called the Generic Cell Rate Algorithm - GCRA (e.g. the continuous-state leaky bucket algorithm or the virtual scheduling algorithm). A cell k is conforming to $GCRA(T, \tau)$, if $y_k = c_k - a_k$ is smaller than τ, where a_k is the arrival time of cell k, and c_k is a "theoretical time" calculated recursively as:

$$c_{k+1} = \begin{cases} c_k + T & if \;\; 0 \leq y_k \leq \tau \\ a_k + T & if \;\; y_k < 0 \\ c_k & if \;\; y_k > \tau \end{cases} \tag{2.1}$$

In the case of DBR connections, T is the PEI ($T = 1/PCR$) and τ is the CDV-T. For SBR connections T is the reciprocal of the SCR and τ is the sum of the IBT and the CDV-T at sustainable cell rate level.

The GCRA can be modeled as a $G/D/1$ queue system, [4]:

- The set of arrivals times $\{a_k\}$ defines a general arrival process to a virtual deterministic queue of service duration T,
- The set of theoretical times $\{c_k\}$ represents the virtual departure times from the virtual single server queue, (c_k represents the departure of the previous arrived cell).
- If $y_k = c_k - a_k$ is negative, cell k initiates a busy period and $a_k - c_k$ is the duration of the previous idle period, see figure 2.1.
- If $y_k = c_k - a_k$ is non-negative, y_k represents the queue waiting time of cell k in the $G/D/1$ queue system.

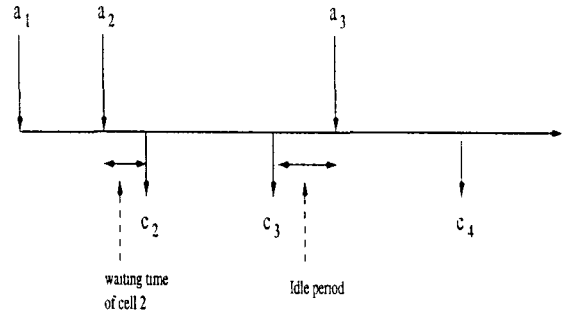


Figure 2.1: Generic Cell Rate Algorithm.

From the last point, a cell is conforming if $y_k \leq \tau$, which means that τ gives a bound to the queue waiting time a cell has in the $G/D/1$ queue system. If τ is a multiple of T, the $G/D/1$ queue capacity would be of τ/T. Therefore, a bound in the number of conforming cells accessing the conformance algorithm can be established, see [69] for a proof:

$$B_s = 1 + \lceil \frac{\tau}{T - \Delta} \rceil \tag{2.2}$$

where, B_s is the maximum burst size allowed by the conformance algorithm, $1/\Delta$ is the access rate of the line on which the cell is observed, and $\lceil x \rceil$ is the upper integer part of x. In the particular case in which the access line rate is the same as the link rate, $1/\Delta = 1$.

Note, that if τ is smaller than $T - \Delta$, that is, only a negligible amount of CDV can be tolerated, the burst size is 2. A lower bound on the minimum CDV that any network must tolerate is given in [1]:

$$\tau_{min} = max\{T, f(T)\} \qquad where \qquad f(T) = \alpha \cdot \Delta \cdot (1 - \frac{\Delta}{T})$$

where α is a dimensionless coefficient whose suggested value is $\alpha = 80$ (this quantity comes as a quantile of the waiting time distribution for the cells of a deterministic source with period T multiplexed in a FIFO queue with Poisson background traffic with load 0.85, it is to say, the $M + D/D/1$ queue system). The function $f(T)$ gives an upper bound on the CDV in a multiplexing stage performing Peak Cell Rate spacing on a link loaded at 0.85.

Due to the different traffic conditions, a cell belonging to a cell stream can experience shorter delays than other cells belonging to the same stream. That means that this cell has an emission interval shorter than the source Peak Emission Interval during a period of time. This effect is called the clumping effect, see [1] or [3]. Furthermore, other cells experience larger delays during periods of time that previous cells. have This effect is called the dispersion effect.

As we said, several effects, e.g. an access network, a multiplexing stage, AAL layers, etc. could cause the jitter introduced in the network. We can see that if the CDV-T, τ, is

not well defined, a clumping effect due to the jitter can produce the discarding of the cell in the GCRA(T,τ). If a cell is clumped too much, it will move away of the theoretical time, producing a higher y_k (remember that $y_k = c_k - a_k$). So, when $y_k > \tau$ the cell will be non-conforming. Here comes the importance to evaluate the cell delay variation a cell suffers in its route.

The Leaky Bucket is one of the most studied and used mechanisms to provide policing in ATM networks. The Leaky Bucket is closely related to the GCRA, and it is also called token bucket filter. We can find a description of this mechanism in numerous papers, e.g. [4], [34], [35], [79] are some references to find some information about the token bucket filter or the Leaky Bucket. Following the nomenclature of [34], the token bucket filter is characterized by a leak rate, r, and a depth or token pool, b. A token bucket is filling up with tokens at a rate of r up to the depth b (its maximum). When a packet is generated, p tokens are removed from the bucket (p is the size of the packet).

If p_i is the size of packet i that is generated at time t_i, a traffic stream conforms to a token bucket filter (r, b) if there are always tokens in the bucket each time a packet is generated. That means that the traffic stream conforms if the sequence n_i defined as following is higher than 0 for all i ($n_i \geq 0$):

$$\begin{aligned} n_0 &= b \\ n_i &= min\{b, n_{i-1} + (t_i - t_{i-1})r - p_i\} \quad i > 0 \end{aligned} \tag{2.3}$$

In ATM networks, where packets are of fixed size, the token bucket filter corresponds to a $GCRA(\frac{1}{r}, \frac{b-1}{r})$. If $v(t,s)$ is the amount of data that enters the network in the interval (t,s), the token bucket filter guarantees a bound on the input traffic to the network:

$$v(t,s) \leq b + (t-s)r \tag{2.4}$$

2.4 Models to quantify the CDV in ATM networks.

We now know the importance of setting traffic parameters and specially the CDV tolerance, which appears as a parameter that bounds the number of conforming cells passing through the UPC function. To set this value, a good idea would be to quantify the CDV introduced by the different elements in an ATM network. One of the most important elements that introduces jitter in ATM is the queueing delay. We can find in [4] a survey of models to account for Cell Delay Variation and methods for setting traffic parameters values. We here summarize some of them.

2.4.1 Single queues.

The simplest way to approximate the CDV introduced in an ATM network is through the GCRA. If T is the Peak Emission Interval (PEI) of a CBR connection, we said that a cell is conforming if $y_k = c_k - a_k \leq \tau$, where a_k is the arrival time of cell k and c_k is the theoretical time arrival. Let $W_j - W_i$ be the difference between cell transfer times from the source to the observation point experienced by two cells of the same connection. Let t_0 be the source emission time for cell 0. Thus: $c_k = t_0 + W_0 + kT$ and $a_k = t_0 + kT + W_k$

$$y_k = c_k - a_k = (t_0 + W_0 + kT) - (t_0 + kT + W_k) = W_0 - W_k < \tau \qquad (2.5)$$

Therefore, τ can be chosen as an upper bound for $W_0 - W_k$ for all cells conforming the GCRA. We can use W_{max} defined as a quantile of the distribution of the observed transfer time for a single cell: $W_{max} = sup\{w \mid P\{W \leq w\} \leq 1 - \epsilon\}$, where ϵ is the target value for the probability of transmitting a non-conforming cell. This model applies when the negotiated PEI. T, is equal to the original time period of the connection (there is no over-allocation of resources to the connection). If T is less than the original time period this approximation is a conservative worst case for estimating τ. The draw back of this approximation is that the cell transfer time distribution has to be known.

During the COST 242 several methods based on a GI approximation for the case the cell transfer time distribution is not known were proposed. All the methods are fed by a renewal input and makes use of the $GI/D/1$ queue model described in the former subsection. Note that if T_{CBR} is the mean value of the GI inter-cell time, and T is the negotiated PCR. then $T < T_{CBR}$ to be the $GI/D/1$ queue stable. Then, the CDV tolerance τ is taken as the level in the infinite queue for which the probability of overflow is less than the target non-conforming value. Examples of this methods are the heavy traffic assumption and the Kingman's bound that provide upper bounds for τ. Kingman's bound gives closer results than the heavy traffic assumption, see [4] for more information about the study of these approximations.

2.4.2 Tandem queues.

The CDV is accumulated when the connection is multiplexed with interfering traffic across several multiplexing stages. In such situation two models are proposed to evaluate the CDV tolerance: the convolution approach and the recursive approach, see [4].

The convolution approach considers independence in the delays in successive queues. Given the delay distribution in single stages, the end-to-end delay distribution can be computed as the convolution of the distributions in successive individual queues. This approach has the draw back of needing the delay distribution at each queue. Since in general this is unknown, some approaches are proposed. The first one considers a single queue whose input is Poisson traffic (note that Poisson is considered as the worst case traffic of the multiplexing of smooth sources). Thus, the $M/D/1$ queue system is used as a model if the rate of the connection is small with respect to the capacity channel. Now, using the heavy traffic approximation for the delay distribution for a single stage and being x the transmission time unit:

$$W(x) = 1 - e^{-\alpha x} \qquad where \quad \alpha = \frac{2(1-\rho)}{\rho} \qquad (2.6)$$

If we assume that there are M queues in tandem, the convolution results in the Erlang-M density function:

$$w_M(x) = \frac{\alpha^M}{(M-1)!} x^{M-1} e^{-\alpha x} \tag{2.7}$$

Now, as in the former subsection, we can approximate the CDV tolerance as an upper bound to $W_M(x)$ (taking a quantile to a target value ϵ). As we will see when we present experimental results, it is expected that this approximation gives pessimistic results if the queue is not heavily loaded or if the rate is high with respect the capacity channel.

Another possibility is to compute statistical characteristics such as mean and variance of the queueing time in a single queue and assuming independence at each queue match the end-to-end queueing delay distribution with a Gaussian or Gamma distribution. The effect of increasing the number of stages M in τ/T has been illustrated in [4]. Simulations show that if M increases, τ increases moderately, however using the convolution approach the end-to-end delay increases steeply. This shows that the convolution approach is accurate only for a limiting number of multiplexing stages (in the example they studied $\rho = 0.85$, T=T_{cbr}=5, $\epsilon = 10^{-3}$, then $M \leq 10$).

Finally, the recursive approach can be used when the number of stages is high. This approach applies if the load at each queue is 1 and the interfering traffic is batch Bernoulli. and it allows to derive recursively the successive inter-cell time distributions along the tandem networks, see [4] for a deeper study.

2.4.3 Impact of CDV on resource allocation.

The traffic characteristics of a connection depend on several parameters as the statistical parameters of the source, however also on the negotiated traffic contract. Among the parameters that influence the choosing of the CDV tolerance are, see [4]:

- *Load of the access network:* the CDV tolerance is an increasing function of the access network load.
- *Original statistical characteristics:* For a CBR source multiplexed with Poisson background traffic, the CDV tolerance is an increasing function of the PEI, and the maximum burst size $B_s = 1 + \lceil \frac{\tau}{T-1} \rceil$ decreases as the PEI increases.
- *Influence of the background traffic:* the CDV tolerance depends on the statistical characteristics of the background traffic. If the background traffic is not fully characterized, it can be difficult to choose a CDV tolerance value.
- *Over-allocation factor:* decreasing the over-allocation factor ($T < T_{cbr}$), where T is the negotiated rate and T_{cbr} is the original rate, the CDV decreases.

From all these points we can conclude that there are a lot of ways of choosing a pair (T,τ). The CDV tolerance has an important impact on resource allocation, since the larger the allocated CDV tolerance, the longer the burst of conforming cells that can pass through a policing mechanism. Resource allocation has been extensively studied in many works (e.g. [4], [26], [27], [28], [40], [41], [50], [51], [54], [63], [68], [85]), but it still is an open issue. The reason is that a Connection Admission Control mechanism has to

guarantee that the QoS of the connections already established would not be damaged by the new connections that are accepted and at the same time has to try to achieve the maximum utilization of the network. However, among the resource allocation proposals, some presuppose the knowledge of parameters that are difficult to obtain and control, others require a great deal of calculations while others cannot guarantee the complete absence of network congestion. Furthermore, in order to have a CAC method that achieves the aforementioned requirements, it is necessary to use mechanisms of Traffic Shaping ([28] and [80]) or Scheduling ([34], [48] and [74]) that would bring additional cost to the designing of the network.

CBR service category supports real-time traffic (e.g. video and voice) and circuit emulation services. This service requires tight CDV bounds (e.g. for dimensioning the elastic buffer required at the terminating end of the connection for absorbing the accumulated CDV). ITU-T, [1], recommends to take into account worst case patterns passing through UPC/NPC. As we will see in the following chapters, worst case traffic can conduct to a decrease of network performance or to low allocation of resources. Therefore, engineering rules for DBR and SBR categories are of main concern in ATM network studies (e.g. how resources should be accommodated to a kind of traffic to meet Quality of Service requirements ?).

Chapter 3

Cell Delay Variation introduced in a real network. EXPLOIT and BAF Testbeds.

Although a lot of theoretical research has been done and is available in the literature, few experimental works can be found. The main reason is the high cost of building and maintenance of an ATM platform. European projects have in the recent years funded ATM Testbeds for experimental research. In this section we want to show two examples of how the jitter is introduced in a real network. For that purpose, we present a series of experiments performed in ATM Testbeds where delay distributions were measured.

The first one was an ATM optical access network belonging to the BAF Testbed (Broadband Access Facilities, RACE project 2024). The experiments were performed in Milano during two weeks. The aim of the experiments were to validate the access network performance, matching the results of a set of experiments with the results of simulators developed during the design of the access network by several partners of the project.

The specifications and description of the BAF experiments can be found in [8]. The experiments were performed by partners from the University of Nijmegen (Netherlands), Queen Mary & Westfield College (Great Britain) and Polytechnic University of Catalonia (Spain). The experiments obtained resulted in the writing of a deliverable, [9], where all the partners participated. The results presented in this chapter are an extract of [9] that was presented in [72].

The second example is the effect of CDV in a network composed by a set of ATM switches where interfering traffic was multiplexed with a tagged source. The CDV experiments were performed during six sessions of one week (each one) in the EXPLOIT Testbed in Basel (EXPLOITation of an ATM Technology Testbed for Broadband Experiments and Applications, RACE project 2061), together with the Swiss ATM pilot network as a part of the Pan European ATM pilot Network (PEAN) and the Telecommunications Laboratory of the Ecole Polytechnique Fedérale de Lausanne (TCOMM-EPFL). The total time dedicated to experiments in this project was 2 years, but only 6 month were dedicated to CDV experiments.

The specifications and description of the experiments can be found in [10] and [14]. The experiments were performed by partners from the Ecole Polytechnique Fedérale de Lausanne (EPFL: Switzerland), Alcatel STR AG (Switzerland), Telefónica I+D (TID: Spain), National Technical University of Athens (NTUA: Greece), Laboratoires d Electronique Philips (LEP: France) and Polytechnic University of Catalonia (UPC: Spain). Results on experiments in the EXPLOIT Testbed by our work-package group can be found in deliverables [11], [12], [14] and published in [24], [56], [57] and [58]. Among the experiments carry out during the project, we present those ones centered in CDV performed by partners from EPFL. Alcatel STR and UPC.

The chapter is organized as follows: section 3.1 describes the ATM Access Network. the BAF Testbed and the results obtained in the BAF demonstrator. Section 3.2 describes the network configuration used in the EXPLOIT experiments. The total EXPLOIT Testbed is not shown since involve part of the equipment not used during this phase of the experiments as some ATM switches, TV coders, mappers, ISDN over ATM, or PC adapters. Section 3.3 is devoted to conclusions and comments about the experiments.

3.1 Delay and CDV in an ATM Access Network. The BAF Testbed.

The BAF project main task was the development and study of an ATM access network. The access network objectives are to support the sharing of access resources efficiently among small business and residential users who can not afford the high costs of a unique interface in a single integrated service network. The ATM Passive Optical Network (APON) with a tree topology seems a good candidate to provide access to public networks to these users by providing a high degree of resource sharing and a flexible and low cost access.

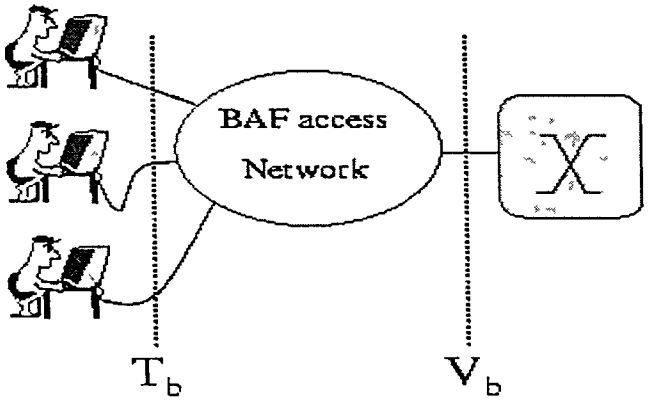


Figure 3.1: BAF access network.

The main objectives of the experiments were the validation of the BAF architecture and the evaluation of the MAC protocol. The measurements were performed between the T_b that give access to users and the V_b interface that give access to a local exchange, see figure 3.1.

Here, we describe briefly the APON architecture of the BAF system, the Medium Access Control (MAC) protocol that provides access to the shared medium, the ATM access network Testbed and results on delay distributions and 1-point CDV, see also [9], [17], [32], [33], [43], [60], [71] and [72].

3.1.1 ATM-PON system architecture.

The architecture of the access network under consideration consists of 3 main subsystems, namely the OLT (Optical Line Terminator), the ODN (Optical Distribution Network) and the ONU (Optical Network Unit) as is depicted in Figure 3.2. The OLT is directly connected to the local exchange and forms the entrance to the access network, with respect to the backbone (or core) network. The OLT also contains the algorithm that distributes the available data rate over all terminations connected. This algorithm is the *engine* of the MAC protocol, whose primary goal is to prevent collisions in the ODN between ATM cells originating from different network terminations.

Either an Optical Network Unit (ONU) or a Network Termination (NT), supporting respectively Fiber To The Curb (FTTC) and Fiber To The Home (FTTH), terminates the access network. In the case of FTTC an ONU serves up to eight NT1s and the path to the home can be completed by either copper or wireless communications. In the case of FTTH, the NT represents a home. Each NT1 or NT provides one T_b-interface. The total access network is limited to a maximum of 81 T_b interfaces (or customers). To provide each customer a reasonable data rate, the transmission rate at T_b and V_b interfaces are assumed to be 149.76 Mbit/s (STM1 in SDH) and 599.04 Mbit/s (STM4 in SDH) respectively.

The ODN is situated between the ONU and the OLT. The ODN is a Passive Optical Network (PON) with a tree topology and a splitting ratio of 1:32. The maximum distance between the OLT and ONU is assumed to be 10 km, covered by fibers. Among many transmission technologies applicable to the access network (ADSL, CATV systems, Active nets, etc.), the PON represents a very promising solution. Using this approach, multiple users may share photonic equipment and fibers in the local loop. In the downstream direction (i.e. network to user) the splitting point of the PON provides a passive instrument that broadcasts ATM cells to all terminations connected. In upstream direction, the access network performs a multiplexing function. Access to the shared medium is arbitrated by the MAC protocol, operating on the ATM layer.

3.1.2 The MAC protocol.

The aim of the MAC protocol is to prevent collisions at the splitting point of the ODN between ATM cells originating from different network terminations. Since the ATM layer does not support retransmissions of lost cells, the MAC protocol should guarantee

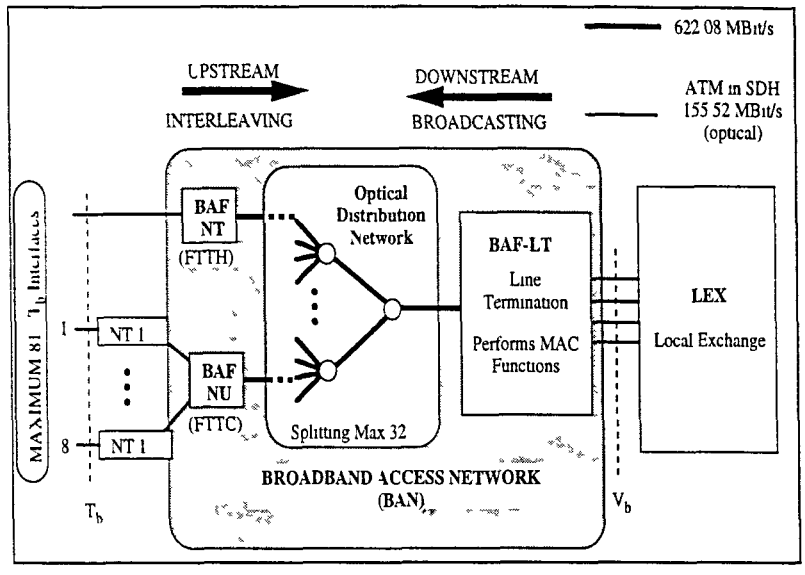


Figure 3.2: Architecture of the access network.

collision-free access.

The MAC protocol implemented in the access network under consideration was proposed by partners from the University of Nijmegen and Polytechnic University of Catalonia and was chosen for the implementation of the prototype used during the experiments. The MAC protocol implemented was called *Global FIFO* and has been described in detail in [32]. Furthermore, the performance of the Global FIFO and other MAC protocols for the APON technology were compared in [4, Part II] and [60]. These performance comparisons were based on simulation results.

This Global FIFO MAC protocol operates on a request/permit basis. Requests are sent upstream to inform the *Permit Distribution Algorithm* (PDA) in the OLT about the number of ATM cells waiting in the buffer of the network termination (NT). The PDA then issues *Permits* to allow a termination to send a single ATM cell, each permit allows a single cell to be transmitted. Although the network termination may be either an ONU or an NT1, for reasons of simplicity we use the acronym NT.

Requests

An NT1 advertises its transmission requirement by sending requests to the MAC controller, which is the *master* of the protocol and is located in the OLT. A request contains information about the state of the buffer in the termination. Two types of requests are distinguished:

1. *requests coupled with upstream cells* An upstream cell originating from an NT1 is preceded by a request field that contains the queue length of the NT1 buffer. The format of an APON slot is described in the following point.

2. *request blocks (RB).* Requests coupled with upstream cells have the disadvantage

that a termination can only reveal its state when it is allowed to send a cell in the upstream direction. To overcome from this drawback, RBs are introduced. An RB contains the request of nine consecutive NT1s. It has the same length as an upstream cell.

Permits

A permit is the output of the PDA. As soon as a network termination receives a permit, it is allowed to send one ATM cell in the upstream direction. Two types of permits are distinguished:

1. *permit for an ATM cell.* When the central controller decides that an NT1 can send an ATM cell, it forwards a permit containing the address of that NT1 to the *Global FIFO* buffer. This permit can consequently be added to a downstream ATM cell. Remark that since the downstream cell is broadcasted no coupling between the address of permits and ATM cells is required. The permit class bit CL=1 indicates that this is a permit for an ATM cell.

2. *permits for request blocks.* When the Global FIFO buffer is empty, a permit for an RB is issued. Such a permit contains the address of the NT1 that is the first one to contribute to a new RB. A total of nine NT1s can contribute to one RB. As the OLT only issues a permit for an RB when it has no permits for ATM cells, none of the available data rate is wasted by sending permits for RBs. Besides giving all NT1s the possibility to advertise their needs for transmission requirements since their last request, the introduction of RBs increases the reaction speed of the protocol.

Permits are issued by the Permit Distribution Algorithm, which uses two counters and a special rule to determine when a permit for an ATM cell can be issued. A bundle-spacer is used to space the cells generated by all the connections that share a T_b since the PDA only knows the data rate requirements of a whole T_b and not of each individual connection of the T_b. For more specific information regarding the implementation of the counters and the used rule we refer to [17] and [43], respectively.

3.1.3 APON slot format.

Transmissions in both the upstream and the downstream directions are organized in TDM slots of 448 bits each. These slots are referred to as APON slots.

In the upstream direction an APON slot can either contain a request block or an ATM cell. In the case a slot contains an ATM cell, a physical layer preamble and a REQ field are added. The latter informs the OLT about the queue length of the buffer, from which the PDA in the OLT calculates the number of transmission rights required. The preamble consists of bits for fine ranging purposes, and bits to detect changings in fiber propagation parameters which may occur as a result of variations in temperature. Furthermore, the preamble includes a bit timing extraction pattern, a byte alignment pattern, a gap and 3 bits for tuning purposes. Due to non-equal delays between OLT and terminations, power levels may change for information flows originating from different terminations, such that tuning becomes necessary. As depicted in Figure 3.3, a preamble is attached to

both an upstream ATM cell and to each request field that contributes to a RB. Although theoretically a total of 14 NTs can send their buffer status to the OLT simultaneously. for implementation reasons it was assumed that only 9 NTs can contribute to a single RB, prevailing a relative gap of 20 bits.

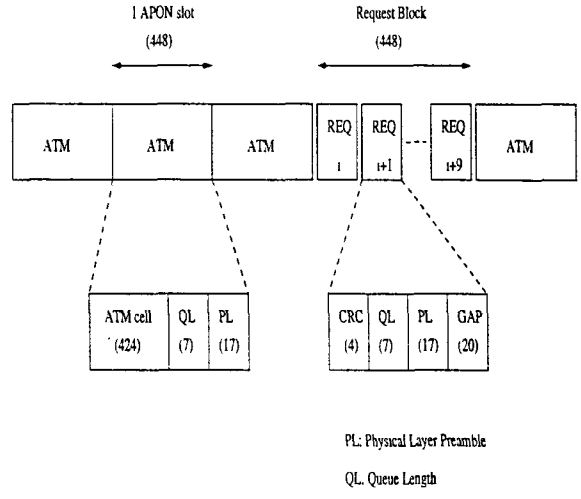


Figure 3.3: Upstream transmission format.

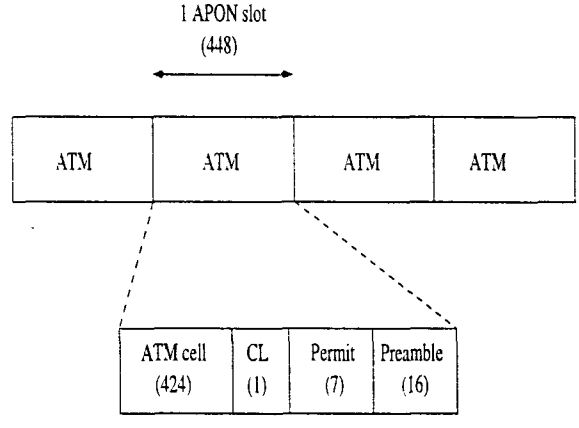


Figure 3.4: Downstream transmission format.

In the downstream direction, an APON slot contains an ATM cell, a permit, a CL field and a preamble. The latter is, again, inherited from the physical layer. Since the access network allows supporting up to a maximum of 81 T_b interfaces, the permit requires 7 bits. The CL-field of 1 bit allows discriminating between permits for RBs and permits for ATM cells. In conclusion, the APON slot format chosen includes a total overhead of 24 bits per slot, for both the upstream and the downstream transmission format.

3.1.4 The BAF Testbed.

The RACE project 2024 provides a powerful platform to experiment and validate in a real environment the architecture and topology of the optical access network and the performance of the MAC protocol.

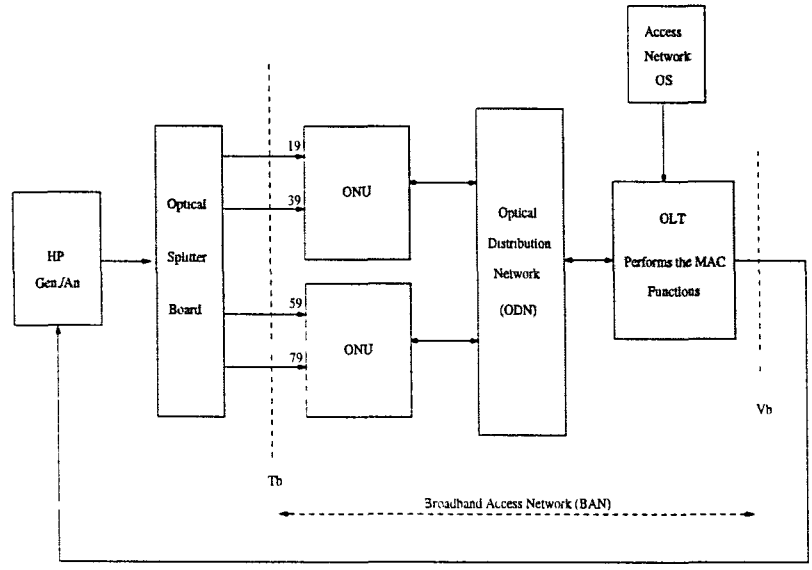


Figure 3.5: Architecture of the access network Testbed.

Figure 3.5 depicts the configuration of the ATM Access Network used for the experiments. The platform consists of a traffic Generator/Analyzer (HP E4210) and the Access Network demonstrator. The Broadband Access Network was controlled by software from the Access Network OS. This software mainly performs operations such as start-up control, VP configuration, alarms and events handling, encryption control and status information in the Broadband Access Network. These allow us to operate the access network manually (e.g. install and monitor VP connections, VP policing, ranging procedures, T_b address definitions, etc).

The main problem with experimentally evaluating the access network performance is the limited number of equipment available in a Testbed of these characteristics (mainly due to the high equipment costs). The Testbed contained only 2 ONU's and 4 T_b's interfaces were available for trial purposes. To load these interfaces a HP 4210 Generator/Analyzer was used to generate a tagged/reference source and the background traffic. Since only one traffic generator was available, an optical splitter was used to increase the load. Although this allowed us to increase the load, it also introduces highly correlated traffic. In an attempt to solve this, more background traffic than necessary to load the system was generated, and then the connections at different T_b interfaces were policed to different levels. Making use of this technique, cells are dropped by the police function, located at the UNI, and thus less correlated background traffic is obtained. The establishment of different VP connections for the T_b's complemented this policing. We must say that correlations introduced by making use of the optical splitter were not

fully eliminated, but they were minimized as much as possible. To build a fairly loaded Access Network configuration, the four T_b's were given equi-spaced addresses, that is, the addresses assigned to the T_b interfaces were respectively 19, 39, 59 and 79 (see figure 3.5).

The HP test equipment allowed us to timestamp the CBR tagged source to measure off-line network performance parameters as defined in [2] (e.g. Cell Transfer Delay and Cell Delay Variation). However, since the analyzer module was able to capture a maximum of 130K reference cells with calendar time-stamps (resolution of 100 ns), statistics could be calculated for this amount of cells only. Different tagged sources (in terms of data rates) were multiplexed with background traffic (CBR and Poisson traffic) to obtain results for three different load conditions. It must be stated that since the generator only offered an STM-1 line the traffic is multiplexed inside the generator.

3.1.5 Performance of the Access Network

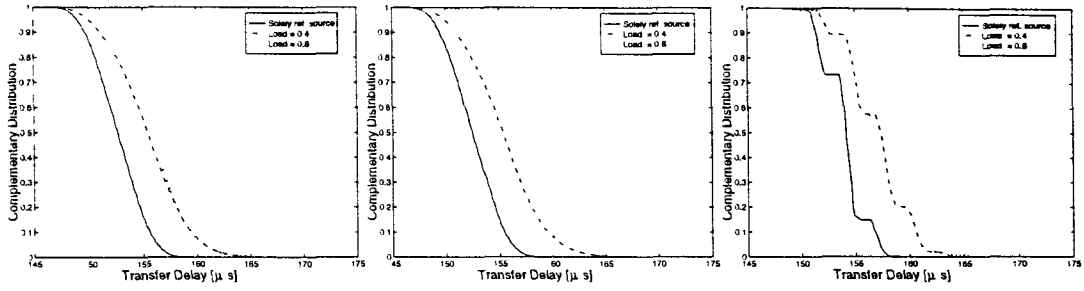
This section reports performance results of the ATM layer experiments, [72] and [9], carried out on the access network demonstrator. Performance evaluation is achieved by studying the complementary distribution function (i.e. $\Pr\{X > x\}$) of two important performance measures, namely the transfer delay and the 1-Point CDV. ATM cells of one particular *reference* (or *tagged*) CBR connection were monitored, under different load conditions, at the borders of the access network.

As we saw in subsection 2.4.3, the impact on the performance measures chosen depend on the load conditions of the access network, the data rate of the reference CBR source and the characteristics of the background traffic. Therefore, sources of 64 K bit/s, 2 M bit/s and 34 M bit/s were chosen as reference connections while a combination of CBR and Poisson connections were chosen as background traffic. Experiments were performed in cases where the access network was loaded up to 0.4 and 0.8. The background traffic was composed with a mixed of 4Mbit/s CBR sources and Poisson traffic, see [9] for a description of the background traffic profile. We must mention that the Poisson traffic is generated inside the HP test equipment. Since we have a unique link interface, the Poisson traffic behaves as Bernoulli. In addition, experiments were performed in the case where the reference source was the only input the policing unit allowed, referred to as *solely ref. source*. Results of these three different load conditions lead to a better understanding of both the behavior of the access network and the impact of background traffic.

Transfer delay

The *transfer delay* of the access network is defined as the time difference between the sending of a cell at the output port of the traffic generator ($\approx$ time arriving at T_b interface) and the reception of this cell at the input port of the traffic analyzer ($\approx$ time arriving at V_b interface).

In [72] is shown that the results obtained from the experimental platform match those obtained by computer simulations. The small differences between the experimental and simulation results is due to small fluctuations in the hardware, SDH frame for-



(a) 64 kbit/s (b) 2 Mbit/s (c) 34 Mbit/s

Figure 3.6: Experimental results for transfer delay: CBR reference source for 3 different bit rates.

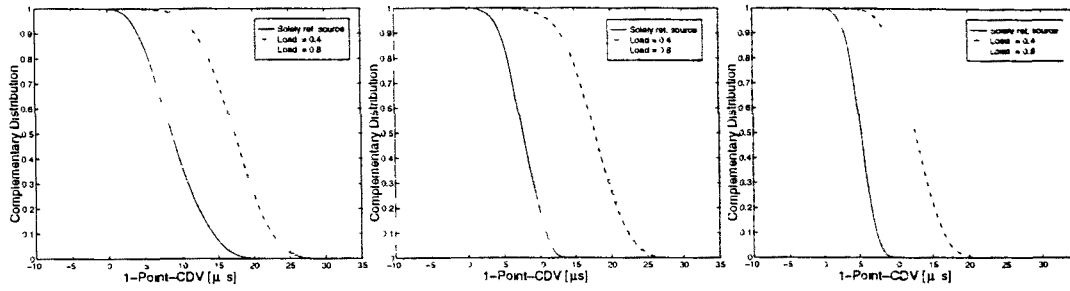
mation/deformation and the presence of the OAM cells which are issued every 512 cells in the demonstrator, but were not modeled within the computer simulator. The 'steps' shown in Figure 3.6.(c) can be ascribed to the formation/deformation of SDH containers and the granularity effects introduced by the traffic generator resulting from the fact that the service rate used is not divisible by the maximum rate provided. Clearly, this effect becomes larger for higher service rates while its impact on low bit rates was overruled by the variation in delay introduced by the MAC protocol.

The total delay introduced by the access network described comprises the following components:

- the time needed until a request can be sent (variable).
- the request propagation delay (fixed).
- the delay introduced by the PDA (variable).
- the waiting time in the Global FIFO queue (variable).
- the permit propagation delay (fixed).
- the cell propagation delay (fixed).
- the delay introduced by the intermediate hardware, such as T_b, U_b and V_b interfaces.

In addition, there is a fixed delay and negligible delay that is introduced in the fibers from and to the traffic generator/analyzer.

We can observe that the delay introduced in the reference source has two components: a fixed delay component mainly due to propagation delays of the transmissions of requests, permits and cells (around 145 μs), and a variable component due to the MAC protocol. The last component is the one that contributes to the CDV. If we take a quantile of the variable delay and take this value as the CDV, we observe a CDV of the order of 20-25 μs. This CDV is mainly due to the delay introduced by the time needed until a request can be sent, by the PDA and by the global FIFO queue that is the core of the MAC protocol.



(a) 64kbit/s (b) 2 Mbit/s (c) 34 Mbit/s

Figure 3.7: Experimental results for 1-point CDV: CBR reference source for 3 different bit rates.

The 1-Point CDV

The most important reason why the access network causes *Cell Delay Variation* (CDV) is that variable queuing, VP-shaping and multiplexing of VCs disturb the profile of each connection. An accurate performance characterization of CDV is a network performance parameter known as *1-Point CDV*. The 1-Point CDV describes the variation of the arrival times pattern with respect to the negotiated peak cell rate. It is measured by observing successive upstream cell arrivals at the V_b-interface and only considers cell clumping, i.e. the effect of cell inter arrival distances which are shorter than T, the reciprocal of the peak cell rate. The characterization of CDV by means of 1-Point CDV was given in [2] and is recommended for CDV assessment by ITU-T.

As with the transfer delay, in [72] the results obtained from the experimental platform are compared with those obtained by means of computer simulations and match rather well. We can first check that taking a quantile of the transfer delay is a good approximation of the CDV. The conclusions of taking a direct measurement of the CDV will be the same as those ones taken by the transfer delay distribution. However, we could add some comments. In the cases where the reference source has a higher bit rate, the relative contribution to the total load of the T_b, which is possibly shared with other connections, is subsequently higher. The bundle spacing function of the PDA therefore spaces the individual cell stream more accurately, leading to less CDV. If the reference source is the only active source, CDV is mainly introduced because an NT must wait for a request block before it can inform the LT about new arrivals. That means that the sources without background traffic (and more for low bit rate connections) rely exclusively on the RB mechanism for sending requests. In the case the reference source is multiplexed with background traffic, the CDV introduced is due to the mechanism of sending permits together with the CDV introduced by the Global FIFO.

In [9] can be found a description and evaluation of all experiments performed in the BAF Testbed (e.g. evaluation of the MAC protocol reaction time or of the MAC queue length at T_b interface).

3.2 Experimental evaluation of CDV impact on ATM resource management. The EXPLOIT Testbed.

The RACE project 2061 EXPLOIT together with the Pan European ATM Network (PEAN) provide a powerful platform to measure and experiment different aspects of ATM traffic engineering. The main objectives of the EXPLOIT Testbed were to provide an ATM platform for pilot applications, broadband experiments and terminal testing: development of interworking units to link to other broadband platforms and existing networks; implement enhancements to the basic testbed functionality; perform traffic experiments in areas of traffic control, resource management and performance assessment and provide input to standardization and guidelines to network operators.

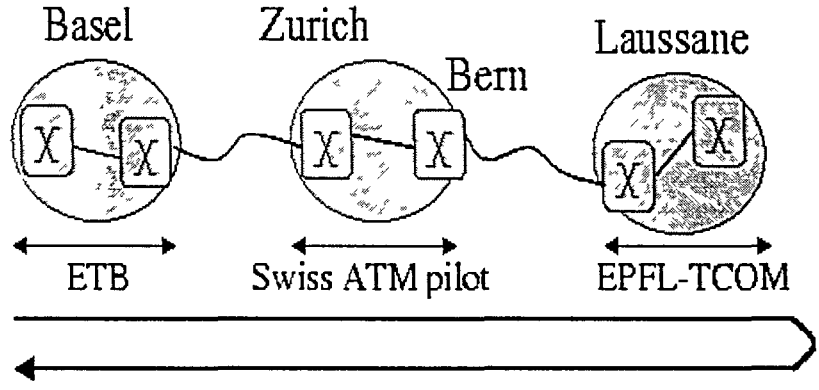


Figure 3 8: Network configuration.

We set up a network configuration between the EXPLOIT Testbed (ETB), the Swiss ATM pilot network and the EPFL Telecommunication Laboratory in Lausanne, see [14], [57] and [58] and see figure 3.8. The objective of such configuration was the set-up of a network involving several ATM segments able of introducing CDV. This set-up was intended to be as real as possible in the sense of introducing switches with different architectures and switching facilities (e.g. ALCATEL, FORE, Philips switches were involved in the experiments), different network links (e.g. E3 PDH 34 Mbit/s, pure ATM and STM-1 interfaces at 155 Mbit/s). The route involved switches across Switzerland (Basel and Lausanne as end Testbeds with measurement points and Zurich and Bern as cross switches with allocated VP-bandwidth and policing facilities).

With such configuration, we can cross up to eight switches, three on the ETB, three on the Swiss ATM pilot network and two at the EPFL Telecommunication Laboratory. The Traffic Under Test (TUT) is multiplexed with controlled Background Traffic (BT) at the ETB and at the EPFL Telecommunication Laboratory, and is multiplexed with unknown background traffic in the Swiss ATM pilot network. The load in the ATM pilot depended on the applications that other projects were testing. The Swiss ATM pilot network authorities allocated us a bandwidth associated with a CDV tolerance (for their UPC function). It was our responsibility to remain within the traffic contract. In both

ends of our network configuration, Test Equipments were available to measure different traffic parameters, such as Cell Transfer Delay, Inter Arrival Time, 1-point CDV and 2-point CDV distributions. The results obtained have been compared with different simple and tractable models proposed in the literature ([4], see also section 2.4) to estimate the amount of CDV accumulation within the network. Other models are available but much more difficult to use for practical purposes.

Being the final goal the measurement of CDV accumulation, a network configuration that consists of a large number of multiplexing stages was needed. Moreover, the multiplexing stages had to offer the possibility to insert controlled background traffic (BT) in order to be able to monitor a reference cell stream, referred as TUT (Traffic Under Test) in the following, between different multiplexing stages. The configuration had to be flexible in the sense that test equipments would have to be attached to different ports and generate different kinds of background traffic directed to different output lines and thus multiplexing stages.

3.2.1 Network configuration.

In [77] can be found a complete description with all the infrastructure in terms of ATM switches, terminal adapters, inter-working equipment and traffic Generation/Analysis equipment of the EXPLOIT Testbed (ETB), see figure 3.9. In the following we will only refer to that part of the EXPLOIT testbed that was relevant for the CDV experiments.

- **The ETB sub-network.**

  On the ETB, three main switching systems are available for such purposes, namely the ALEX (Alcatel-BELL), LaTEX (Philips PKI) and the ASX-200 (FORE). The ALEX system is a prototype of a public ATM switch based on a 16x16 switching elements using shared buffer memory. LaTEX is also a prototype of a public ATM switch whose internal structure is a combination of cross-point and output buffered switch. Both switching elements had STM-1 and pure ATM at 155.52 Mbit/s. The ASX-200 is a bus-based output buffered ATM-LAN switch. In our version, the switching capacity is restricted to 155 Mbit/s per input port with a maximum of 16 inputs.

  We define on the ETB a network configuration that consists of five multiplexing stages within the three switching systems, see figure 3.10. By multiplexing stages we refer to those stages where the TUT was multiplexed with background traffic (these stages are numbered in figure 3.10 from 1 to 5: two in the LaTEX switch and three in the ALEX switch). These multiplexing stages are fully controlled in the sense that the traffic flowing through the switching systems is perfectly known and reproducible.

  Two test equipments are used to generate the background traffic as well as the TUT traffic, namely the ATM-100, [5], and the A8640, [7], as well as to perform

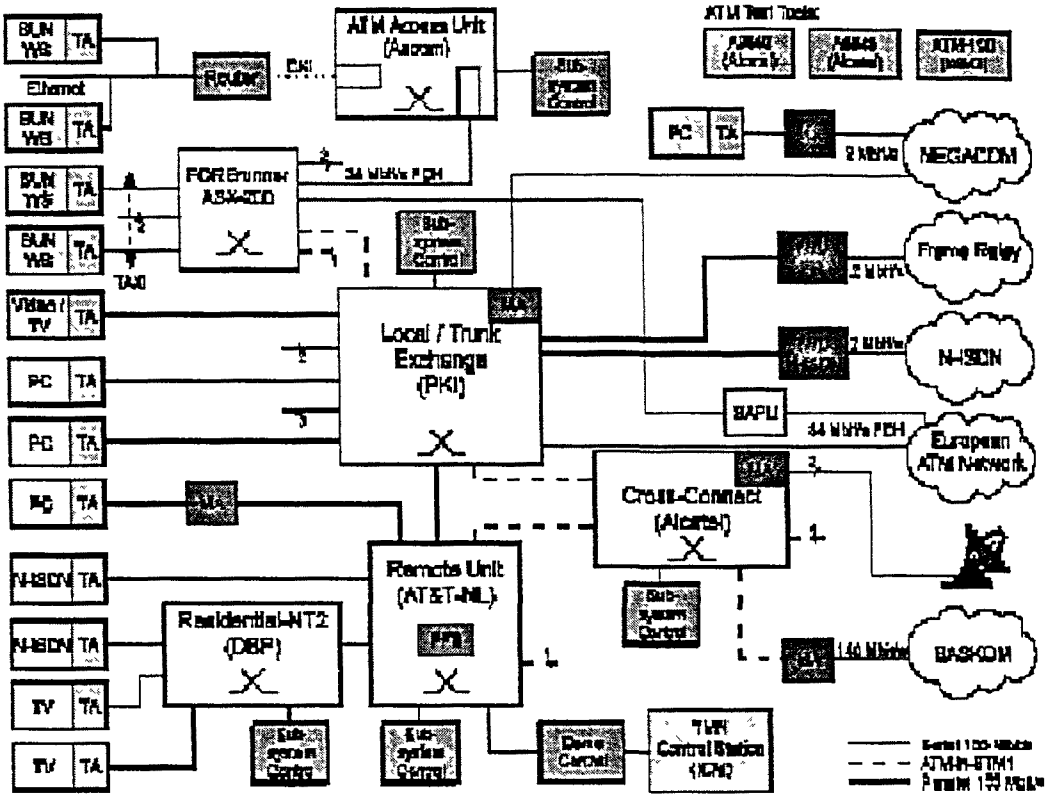


Figure 3.9: EXPLOIT Testbed map.

a part of the analysis. The ATM-100 provides two traffic generator modules able to generate CBR sources and General Modulated Deterministic Process (GMDP) sources. The ATM-100 offers pure ATM interfaces and STM-1 (electrical or optical) at 155.52 Mbit/s. As testing facilities the ATM-100 can save up to 32 kcells for off-line analysis and allows on-line delay and inter-arrival time histograms. The A8640 also offers two generation modules with electrical or optical STM-1 interfaces at 155.52 Mbit/s. The A8640 generates only CBR sources without phase-moving properties (see next sections about this topic) and has two test modules to analyze both off-line (up to 2 Mcells) and on-line records of delay and inter-arrival time histograms.

- **The ATM pilot network.**

  The TUT is going through 4 links between the ETB sub-network and the EPFL Telecommunications Laboratory (TCOM) crossing 3 switches situated in the cities of Zurich, Bern and Lausanne. The Swiss TELECOM, operator of the Swiss ATM pilot, commits to offer a VP path from Basel to Lausanne with a CBR traffic contract consisting in the PCR of the VP connection and a CDV tolerance. The QoS part of the contract is omitted in the pilot phase, but they committed to offer the best VP service using PCR allocation to all VPs. The contract consisted on a VP

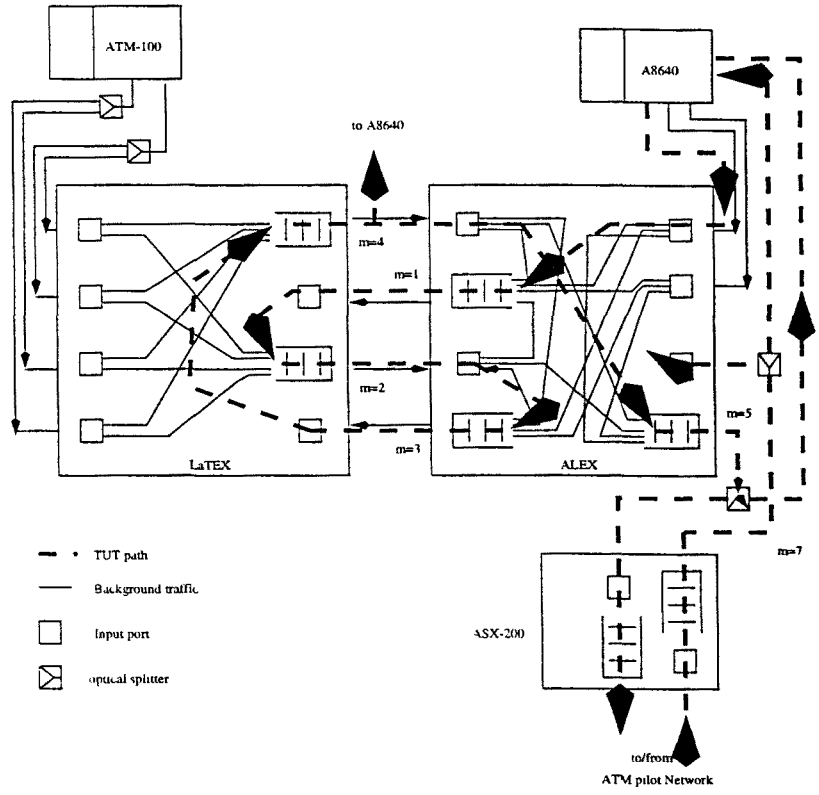


Figure 3.10: ETB configuration network.

PCR of around 24 Mb/s and a CDV-T of 367 μs. The background traffic in these segments is thus not known and will depend on the traffic that other projects sends through the Swiss ATM pilot. The TUT is policed at both entry points ETB-ATM pilot and EPFL-ATM pilot.

- **The EPFL Telecommunications Laboratory (TCOM).**

  At EPFL, an ATM local area network has been set up for research projects. It consists of two Fore ASX- 200 with various attached workstations. A HP 75000. [6], was used to generate TUT as well as BT traffic, see figure 3.11. The HP75000 allows the generation of CBR traffic without phase-moving property and Poisson Traffic. There has a test analyzer able to capture 30 Kcells with timestamps for off-line analysis.

All multiplexing points where the TUT was multiplexed with Background Traffic (BT) are listed on figures 3.10 and 3.11 with numbers which range from 1 to 7. Points from 1 to 5 are in the ETB, point 6 stays in the EPFL Testbed, and point 7 is the point where TUT comes back from the ATM pilot network to the ETB. In all these points is possible to introduce a measurement point connected to any of the traffic analyzers mentioned.

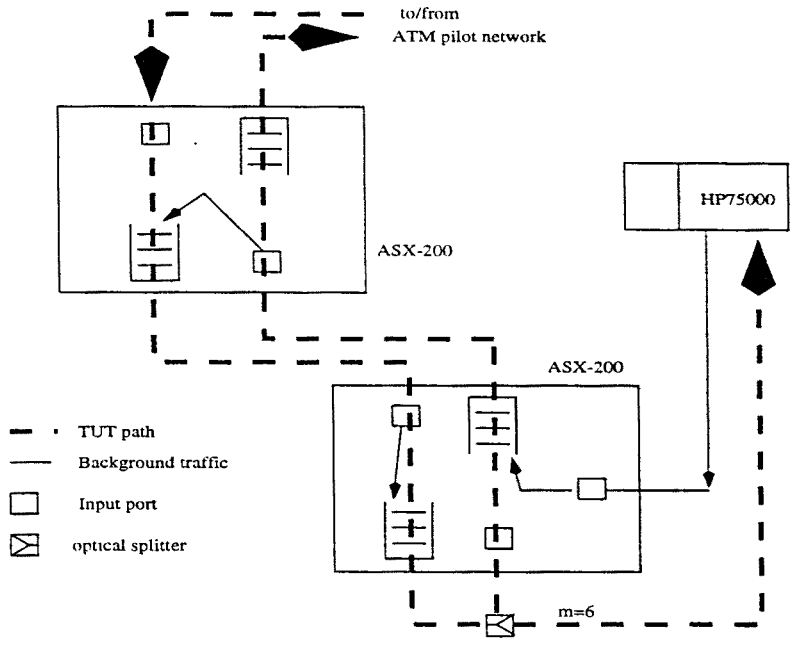


Figure 3.11: EPFL configuration network.

3.2.2 Traffic sources.

Having a small number of generation ports is a limitation typical of many ATM test platforms due to the high costs of such equipments. As can be seen from figures 3.10 and 3.11, the number of test equipments ports with BT generation capabilities is restricted to five in our configurations: the ATM-100 and A8640 (both with two generating modules in the ETB and the HP75000 at EPFL. Our concern was to determine the best way to use the generation capabilities at our disposal.

The problem is to design experiments from which results can be extracted and interpreted according to scenarios as realistic as possible in respect to real operational conditions. We assume that each output lines from an operational switch might receive traffic from several physical inputs of the switch realizing thus an Nx1 multiplexer where N is restricted by the number of input lines but can be less. The switching systems used had a restricted number of input lines that could be used to input BT and TUT traffic.

- **Traffic on the ETB segment.**

  The traffic sources on the ETB are all of the CBR type. The traffic generators available produce periodic sources with the same phase. It is well known, [4], that the measurements might be heavily dependent on the particular phase relationship between the sources. To avoid this problem, we use phase-moving CBR sources. These sources emit cells with a constant time interval during a given time period and then change their phase for the next period, and so on. If the time between phase changes is long enough the source can be considered as CBR and the phase

ATM Test equipment	PCR (cell/s)	MTB-PC mean time between phase changes
ATM-100	4585	50 msec
A8640	4596	no phase change

Table 3.1: CBR sources parameters for the background traffic.

change effect can be neglected. However, not all Test equipments provide facilities to change the phase. Since the ATM-100 allows the generation of GMDP sources, a phase-moving CBR source can be modeled as a two-state source with the same PCR in both states and a defined mean time between phase changes. The A8640 and the HP75000 do not provide facilities to change phases.

In all the experiments presented here, the background traffic at each multiplexing stage on the ETB consists of an homogeneous set of CBR sources whose parameters are given in table 3.1. The total number of background sources depends on the total load in the multiplexing stages. All the experiments have been performed with a total load of 90 % including the TUT and the background traffic. Thus the number of background sources varies depending on the TUT PCR being the load of the background traffic $\rho_{bt} = \rho_{tot} - \rho_{TUT}$.

A maximum of 4 physical input lines were available at each multiplexing stage allowing experiments with several load balances (e.g. from 2 to 4 input lines). In order to load 5 multiplexing stages with 4 input lines each one with four traffic generators, traffic recycling and traffic duplication techniques were used. Traffic duplication is obtained by using optical splitters and VP switching. Note that this means to have correlated traffic in the experiments. By traffic recycling it is meant that the same background traffic will be multiplexed with the TUT in 2 multiplexing stages. However, we try to minimize the effects produced by these techniques by imposing a set of conditions to avoid that the same duplicated traffic was always multiplexed in the same stages. Thus, in the design of the experiments the following conditions were applied:

- At least 50 % of the CBR connections at each stage should have phase moving.
- Traffic recycling will only involve phase moving sources.
- No more than 50 % of the total background load at each stage should consist in recycled traffic.
- The background traffic at each stage should preferably come from 4 input lines.
- Duplicated traffic is multiplexed in different multiplexing stages.

As can be observed, we introduce three traffic conditions to perform the experiments with the equipment available: background traffic recycling, load balancing and traffic duplication. Background traffic is split on different VP channels. These

VP channels are de-multiplexed at the input of the switches and redirected to different output ports. On the LaTEX, the traffic from each ATM-100 generator is duplicated by means of the optical splitters and inserted in 4 input lines. At each input line only two VP channels are allowed. From the input lines the VPs are redirected to output ports by VP switching. With this configuration there is no recycling in both LaTEX multiplexing stages and the duplicated traffic inputs different multiplexing stages, allowing increasing the number of input lines.

On the ALEX, we dispose of 3 multiplexing stages. Two of them as output ports connected to LaTEX and another one connected to the ASX-200. The 4 input lines of these output ports are 2 input lines coming from the LaTEX and 2 input lines coming from the A8640. The thin lines on figure 3.10 represent the route of the background VP channels. When not used anymore a VP channel is blocked at the next input port, see [58] for a deeper discussion on the source traffic set-up.

We used to kind of traffic configurations:

- Configuration 1: the background traffic is inserted in 2 input lines at all multiplexing stages in Basel. The optical splitters are not used in this case.
- Configuration 2: the background traffic is inserted in 4 input lines at all multiplexing stages in Basel. The optical splitters are used. This configuration allows a better mixing of CBR sources at each stage.

The reason of having two configurations comes from the fact that the switches at the EXPLOIT Testbed were shared with other European projects. The experiments were performed during 6 weeks but distributed in one week per month. That means that in different weeks we had the interaction from different projects. When some European project used the Basel Testbed, we could not make use of all the switch lines. So we were limited to use configuration 1: only 3 input lines per multiplexing stage in Basel (two input lines for the background traffic and other for the TUT). This situation occurred the first three experimental weeks. Finally, the administrator of the Basel Testbed communicated us that we could have the whole Testbed for ourselves until the end of the project, the remaining three weeks. Having 5 input lines per stage (4 lines for the background traffic and other for the TUT) meant to introduce more contention. From our point of view, we had a more fairly configuration and the opportunity of introducing more CDV. However, we did not have enough time to repeat the experiments performed previously. We decided to go on with the new set up, and differentiate the results from both configurations.

Finally as a summary, it should be noticed that there were never less than 60 background sources active at each stage for all experiments. At least 30 of them were with phase-moving (criteria 1) with mean time between phase changing of 50 msec. That means 600 phase change every second and approximately 10^6 in total in an experiment duration of 30 minutes.

- **Traffic on the ATM pilot.**

  The traffic in the Swiss ATM pilot was unknown but controlled by PCR allocation (a VP of 24 Mb/s and a CDV tolerance of 367 μs). The Network Management Center (NMC) of the pilot network gave us access to the load statistics for all segments of the VP route through Basel-Zurich-Bern-Lausanne during the experiments. The statistics were collected every 15 minutes in terms of the total cells passing each multiplexing stage. Although these statistics do not allow us to use queueing models to analyze the behavior of our TUT source, it allows us to estimate the load conditions in which our source was multiplexed. That gives us a rough estimation of the impact of the background traffic in our source. We accounted that most of the time the total load on all the segments of the pilot was very low (including our traffic). In the cases for which the background traffic intensity was high the total load on the E3 line was below 40 %. Following this estimates we have consider that the only multiplexing stages where the background traffic could introduce CDV are those ones on the ETB and the EPFL laboratory where we controlled totally the background traffic. However, as we will see in the results, from time to time, we measured higher delays that we expected with zero load conditions in the ATM pilot. In these cases, we have assumed that during very short amounts of time there was some traffic in the ATM pilot network multiplexed with ours.

- **Traffic on the EPFL segment.**

  At EPFL only one port was available to generate traffic from the HP75000. The background traffic consisted of a Bernoulli source. In order to load the two multiplexing stages at EPFL, the same background traffic is recycled at the second stage and then blocked when it returns to the first one.

3.2.3 Point Measurements on the Testbeds.

All measurements were taken on the TUT source. The TUT was generated by the A8640 on the ETB and goes through the following route, see the thick line on figures 3.10 and 3.11: is multiplexed in 5 stages on the ETB (two on the ALEX and three on the LaTEX) with controlled background traffic following the criteria specified in the former subsection. At each of the multiplexing stages a measurement point was introduced (numbered from 1 to 5) to take statistics on delay distributions. Going out from the ETB, the TUT crosses the ATM pilot network to EPFL where is again multiplexed with background traffic from the HP75000 (measurement point 6 is here introduced). The TUT returns from EPFL through the ATM pilot network to the ETB. At this point two strategies were considered: the first one was to finish the experiment taking statistics from the TUT (we are at point measurement 7). The second one is consider a larger number of multiplexing stages re-introducing the TUT again in the same path. At this point we are assuming that the TUT in the second loop is enough uncorrelated respect to the first loop TUT since has crossed three ATM segments in different traffic conditions. We call this technique TUT recirculation and we assume that for low TUT PCRs, the effects of having the TUT circulating two times in the same queues can be neglected crossing three ATM segments before being again multiplexed with the first one. Anyway, we have

always kept the total TUT load (including recirculation) below 12 % of the link load (STM-1) in order not to introduce undesirable effects due to recirculation.

The measurements were performed on the points numbered from 1 to 7 in figures 3.10 and 3.11. From most of the experiments only delay and interarrival time statistics were taken. Since the ETB and EPFL platforms are not synchronized (e.g. with a GPS module), delay measurements could only be taken when the traffic returned to the ETB. In the case of the HP75000, 1-point CDV statistics on off-line processing were also taken. For this measurements, the TUT has to be generated by the HP75000 at EPFL.

Summarizing these points we can consider the following points:

- m means number of multiplexing stages.
- $m = 1$ to $m = 5$ are the output of multiplexing stages in the ETB. We copied with an optical splitter the TUT and measured the delay distributions and parameters as mean and variance of the delay distribution on-line.
- $m = 6$ is situated in the EPFL Testbed.
- $m = 7$ is the point that connects the Swiss ATM pilot with the ETB.
- We have considered that since in the ATM pilot the load was small, $m = 7$ really only represents 7 multiplexing stages: 5 belonging to the ETB platform and two belonging to the EPFL platform.
- $m = 12$ is the same point as $m = 5$. However the TUT has circulated twice in the ETB.
- $m = 14$ is the same point as $m = 7$. However the TUT has circulated twice in the ETB, the Swiss ATM pilot and EPFL platform.

The main goal of these experiments was to investigate experimentally the CDV at UPC/NPC and its impact on bandwidth management. For that purpose, the best way to evaluate the CDV tolerance is to perform a direct policing experiment or to measure the 1-point CDV. To perform a direct policing experiment at all multiplexing points was a difficult task due to the available equipment at that time and the fact that we needed a huge amount of experiments to obtain results. This is because a policing experiment only returns the non-conforming ratio as a function of the PCR and CDV tolerance GCRA parameters. Several experiments have to be performed to reach the required non-conforming ratio for a given PCR-CDV tolerance. Furthermore, measure the 1-point CDV was only possible on the EPFL Testbed (only one measurement point). To solve this problem, we decided to take statistics on the delay distributions and apply the models proposed in [4] and mentioned in subsection 2.4.1 and 2.4.2. Note that one of the drawbacks of some of these models is the knowledge of the delay distribution. In our case, we have the advantage of having the measured delay distribution. Therefore, we have applied the simple approximation and we have chosen as measured CDV the peak-to-peak CDV: the α-quantile of the Complementary Probability Distribution Function (CPDF) of the Cell Transfer Delay (CTD) minus the fixed delay, see [3].

To get a picture of this approximation L. Jaussi from EPFL performed before starting the set of experiments a simulation to compare approximation of taking the quantile of the CTD as a bound of the CDV-T. The simulation model a CBR TUT source crossing 5 multiplexing stages and being multiplexed with background interfering traffic, see table 3.2. The background interfering traffic is the superposition of Bernoulli sources. The system was simulated such that emulated configuration 1 of our experiments: 2 input lines. On each input line a given number of Bernoulli sources were pre-multiplexed.

TUT IAT (cell slots)	CDV-T	Delay Quantile [95%]
5	52	59...63
10	63	67...72
20	68	68...83
40	72	71...84
80	77	69...86

Table 3.2: Comparison of exact CDV-T with the simple delay approximation.

As it can be observed in table 3.2. the results based on the quantile of the CTD upper bound the CDV-T and give a good approximation. The second column gives the simulated CDV-T required to achieved a non-conforming ratio of 10^{-4}, while in column 3 is given the 10^{-4} quantile of the CPDF CTD.

As a summary. knowing the CPDF CTD we have approximated the CDV with the following approaches:

- Simple approximation: taking a quantile of the CPDF CTD.
- Compute the mean and variance of the CTD and fit the end-to-end CTD with a Gamma distribution.
- Apply the convolution approach taking each network segment.
- Assume a worst case dimensioning with the $M/D/1$ queue system and assuming independence use the Erlang-M distribution.

3.2.4 Fixed delays and CDV without background traffic.

A first set of experiments was performed to evaluate the fixed delay introduced between the generation of the TUT source in the A8640 and the measurements points. These experiments were performed without any background traffic on the ETB or the EPFL platforms. Table 3.3 presents the fixed Cell Transfer Delay (CTD) and the CDV on point measurements 5 and 7. Point 5 is situated in the output of the ETB (before entering the ATM pilot, see figure 3.10). The delay here is due to the processing and transmission times of the cells through ALEX and LaTEX switches. Point 7 includes the CTD introduced in the loop between the ETB and the EPFL Testbed. This fixed delay is equal to 4.4 msec.

Measur. point	Fixed CTD (μsec)	CDV (μsec)
5	354.4	27.3
7	4452.1	81.8

Table 3.3: CTD and CDV without background traffic on the ETB and EPFL.

The CDV measured is calculated as the maximum CTD minus the minimum CTD. This CDV is given by the physical layer overhead, multiplexing with OAM cells, switch processing and switch architecture. In our configuration the TUT is transmitted in SDH lines and E3 PDH lines. SDH and PDH cell framing introduce CDV at ATM streams. We must think that from TUT generation in the A8640 to point measurement 5 there are 5 switching nodes and 6 network segments while from the ETB to the EPFL the TUT crosses up to 11 switch nodes with 12 network segments. The E3 PDH lines were situated in the ATM pilot network while on the ETB all the lines were STM-1. This explains the higher CDV value at point measurement 7.

3.2.5 CDV with background traffic.

Figures 3.12, 3.13, 3.14 and 3.15 give the CTD CPDF for 4 different TUT: 21.2 Mbit/s, 8.98 Mbit/s, 4.55 Mbit/s and 7.93 Mbit/s. The total load is always 90 % and the fixed delay has been removed. Figures 3.12 and 3.13 have been measured with configuration 1 (2 input lines for background traffic at each stage) and figures 3.14 and 3.15 with configuration 2 (4 input lines for background traffic at each stage). Now the CDV will come from the conditions encounter by the TUT without background traffic (e.g. physical layer overhead, switch architecture, etc) and by the queueing delay introduced by being multiplexed with background traffic. Therefore, note that the theoretical queueing models do not take into account the first class of delays.

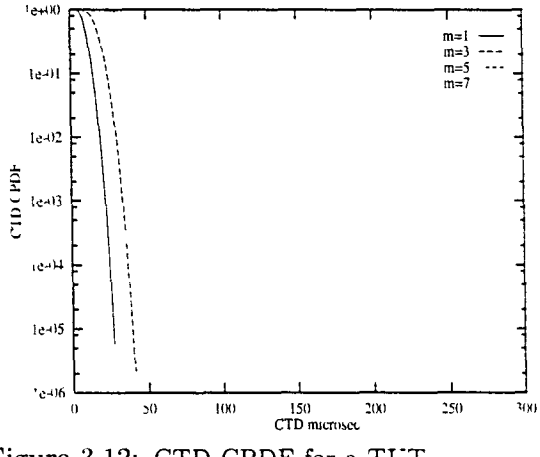


Figure 3.12: CTD CPDF for a TUT PCR of 21.2 Mbit/s (50'000 cells/sec). Configuration 1. m is the number of multiplexing stages.

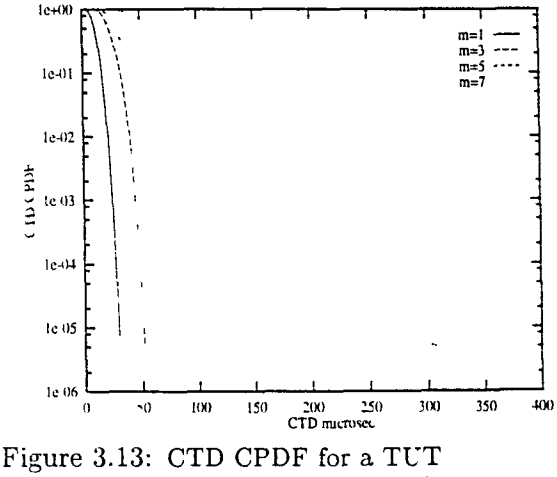


Figure 3.13: CTD CPDF for a TUT PCR of 8.98 Mbit/s (21'180 cells/sec). Configuration 1. m is the number of multiplexing stages.

The curves show that the ATM pilot and the EPFL segment introduce the most part of the CDV (e.g. in curve 3.12, the CDV value of a 10^{-6} quantile introduced in the 5 multiplexers on the ETB is around 50-60 μsec, while for the ATM-pilot-EPFL segments is 250 μsec). To justify this values we first see that the CDV introduced in the ATM pilot without background traffic, see table 3.3, is around 55 μsec (the double than in the ETB segment). This is mainly due to the E3 PDH links and the possibility of some delay due to the multiplexing with some traffic on the ATM pilot segment (remember that there always be a load less than 40 % on the ATM pilot segment). Secondly, the background traffic at the ETB segment is the superposition of CBR sources while at the EPFL segment is Bernoulli. While a queue on the ETB can be modeled as an $ND/D/1$ queue system, an EPFL queue can be modeled as an $N * Geo/D/1$ queue system. It is well known that an $N * Geo/D/1$ queue system upper bounds a $ND/D/1$ queue system and both are upper bound by a $M + D/D/1$ queue system. These facts can explain that the ATM pilot-EPFL segments introduce more CDV than the ETB segments. The different tail behavior in curves 3.12 and 3.13 respect to curves 3.14 and 3.15 are difficult to justify. A possible explanation could be that in the ATM pilot segments, although the load was very low, our traffic could be multiplexed with short high peak rate bursts (e.g. IP-over-ATM applications that were tested from time to time on the ATM pilot network).

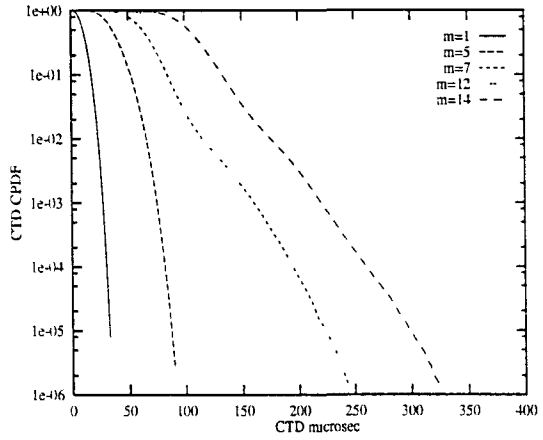


Figure 3.14: CTD CPDF for a TUT PCR of 4.55 Mbit/s (10'731 cells/sec). Configuration 2.

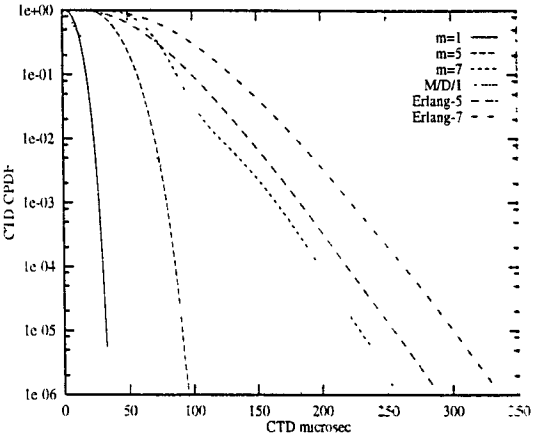


Figure 3.15: CTD CPDF for a TUT PCR of 7.93 Mbit/s (18'703 cells/sec). Configuration 2. Comparison with $M/D/1$ and Erlang-M models.

In curves 3.14 the recirculation technique is used in order the TUT crosses more than 7 multiplexing stages (e.g. m=12 and m=14 multiplexing stages). Since the Traffic Contract on the ATM pilot gives as a VP bandwidth limited to around 52000 cells/s (approximately 22 Mbit/s) we have to use TUT's which PCR is less than 5 Mbit/s. Point 12 is the same that point 5 after recirculating the TUT. We can observe that configuration 2 which has 4 input lines, introduces much more CDV that configuration 1 that has only 2 input lines for multiplexing stages. Thus, balancing more the traffic among more input lines is worst. This is logic, since the contention of 2 input lines is less than the contention that produces 4 lines.

Assuming a worst case dimensioning with the $M/D/1$ queue system and heavy traffic approximation, the delay distribution of stage m can be approximated with the Erlang-M distribution, equation (2.7). In figure 3.15 the delay distributions for m=1, 5 and 7 are compared with an $M/D/1$, Erlang-5 and Erlang-7 models for the same load conditions. As can be observed, the $M/D/1$ approximation is a rather pessimistic bound. Table 3.5 gives a summary of this approximation for values of a quantile of 10^{-6} that are compared with the measurements values shown in table 3.4.

TUT PCR (cell/s)	m=1	m=5	m=7	m=12	m=14
Configuration 1					
50'000	28.3	59.5	254.8	na	na
35'307	28.3	53.8	294.4	na	na
21'180	31.1	70.8	359.6	na	na
Configuration 2					
18'703	33.9	99.5	263.3	na	na
10'731	33.9	93.4	254.8	291.6	339.7
5'448	33.9	96.3	215.2	254.8	280.3

Table 3.4: Measured CDV in μsec (10^{-6} quantile of the CTD CPDF). "na" means non available, the recirculation technique was not used due to high TUT PCR.

m=1	m=5	m=7	m=12	m=14
186.9	297.3	348.2	458.7	501.1

Table 3.5: Estimated CDV in μsec (10^{-6} quantile of the CTD CPDF) using the Erlang-M approximation.

As It can be seen in table 3.4, the CTD quantile measured both for configuration 1 and 2 are closed to each other for all TUTs at measurements points from m=1 to m=5. At point m=7, when the TUT comes back from the ATM pilot, all measured values lie within the interval 250± 50 μsec, except one result (that one plotted in figure 3.13) that is much higher. This effect can be observed in the difference of the tail of figures 3.12 and 3.13 respect to figures 3.14 and 3.15. We interpret this result as some interaction in the ATM pilot with some background traffic. The results with m=12 and m=14 have been measured with TUT recirculation and only for PCR of 4.55 Mbit/s and 2.3 Mbit/s. In stages from m=1 to m=5 where the traffic was fully controlled we can observed that the CDV is an increasing function of the PEI. This result was stated on the studies performed on the COST project. For stages higher than m=5 it is difficult to observed this effect This is due to the fact that the results for m=7, 12 or 14 depend on the ATM pilot multiplexing conditions such as PDH framing, multiplexing with any bursty traffic coming from another projects or recirculation techniques.

Table 3.5 shows the Erlang-M approximation. If we compare tables 3.4 and 3.5 we can see what it was said, the Erlang-M approximation gives pessimistic values for CDV values in the multiplexing cases m=1 to m=5. The reason is that the background traffic is 2 Mb/s CBR connections. Furthermore, we must take in mind that the measured values include

other interactions than only queueing delays as was indicated with the CBR without background traffic. The only ones that approximate better the Erlang-M approximation was due to possible multiplexing in the ATM pilot with bursty traffic (stages higher than m=5). We must mention the case $m = 7$. In this measurement point, the TUT has been multiplexed in 5 stages in the ETB, in 2 stages in the Swiss ATM pilot, in 2 stages in the EPFL platform and in 2 stages in the Swiss ATM pilot. The result is 11 multiplexing stages. However, the 4 multiplexing stages of the Swiss ATM pilot represent loads lower than 0.4. Since the Erlang-M applies under the assumption of heavy traffic, we compare the measurements of point $m = 7$ with an Erlang-7 and in the interpretation of the results we must remember that from time to time there were bursty traffic in these stages. I want to note that even if I consider Erlang-11 or Erlang-7, it is really difficult to compare and get conclusions of this comparison.

TUT PCR (cell/s)	m=1	m=5	m=7	m=12	m=14
50'000	36.8	79.3	314.3	na	na
35'307	39.6	73.6	220.8	na	na
21'180	45.3	90.6	229.3	na	na
18'703	50.9	124.6	218.0	na	na
10'731	53.8	127.4	192.5	226.5	311.4
5'448	53.8	121.7	223.7	240.7	317.1

Table 3.6: Gamma approximation in μsec (10^{-6} quantile of the CTD CPDF). "na" means non available

Another model to approximate the delay distribution and furthermore the CDV is to compute the mean and variance of the end-to-end queueing delay assuming independence and match it with a Gamma distribution. Thus, the mean and variance of the delay at each measurement point have been used to fit the Gamma distribution. The ATM analysers gave us the transfer delay distribution and at the same time the mean and variance of this distribution. Therefore, the use of a Gamma distribution is an alternative method to compute the CDV. I do not present the parameters that fit the Gamma distribution since this values were analysed by the other partner of the project. However I explain the technic used to get this alternative computation of the CDV. The CDV due to other factors than queueing delay are not small (especially after crossing the ATM pilot). The mean and variance of the delay under zero load condition have been removed from the measured mean and variance of the delay. The resulted mean and variance have been matched to a Gamma distribution with the assumption that the delay introduced by queueing is independent of the delay introduced by other factors. To compare with the measurements, the resulting Gamma distribution has been combined by convolution with the delay distribution due to the other factors. So, at the end all the sources of CDV are added to the final distribution. In our case, this provides better results than directly fitting the Gamma distribution to the end-to-end delay distribution.

The results of table 3.6 fits much better the measurements than the Erlang-M approximation. However, they are sometimes optimistic as the case of the TUT with PCR of 21'180 cells/s. This result corresponds to that one of figure 3.13 for which the tail of the CPDF at m=7 is different of what would be expected for a Gamma CPDF which has a

negative exponential tail. Matching a Gamma distribution is a good approximation to compute CDVs.

TUT PCR (cell/s)	m=12	m=14
10'731	300.1	353.9
5'448	257,6	288,8

Table 3.7: Estimation of the CDV from a delay measurement.

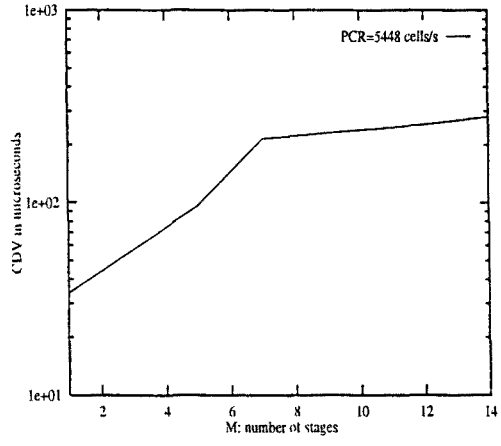


Figure 3.16: CDV as a function of the number of stages for a PCR of 5448 cells/s.

We can see what happens if we take the mean and variance of a single queue and assuming independence, we match the end-to-end delay distribution with a Gamma or a Gaussian distribution. Thinking in what we know about the experiments, we can say that this method can give good results for $m \leq 5$ since the only traffic that was multiplexed in these stages was CBR background traffic. The mean and variance of the first queue will not have into account the interfering traffic of the ATM pilot network or the kind of background traffic used on the EPFL segment network. Therefore, this method will not give good results for those multiplexing stages situated at measurement points higher than 5. However, we could consider the ETB network segment, and the ATM pilot network together with the EPFL network segment as independents and compute an estimation of the delay CPDF quantiles at measurement points $m = 12$ and $m = 14$ by convolution of the measured delay PDF at $m = 5$ and $m = 7$. Results of this method can be found in table 3.7. As we can observe, these results fit well the measured ones. We note that we have only fitted values for two TUTs since the other ones were not available due to bandwidth VP constrains in the ATM pilot network.

Finally, we have plotted the CDV as a function of the number of stages for a PCR of 5448 cells/s in figure 3.16. We can see what it is stated in COST 242: the CDV does not grow unbounded as the number of stages increases. In the graph we observe that as the number of stages increase the CDV grows more moderately.

3.3 Conclusions and comments about the experiments.

In this chapter the results provided by two ATM platforms have been presented. These Testbeds were used to perform traffic experiments in the context of ATM layer and more specifically in the case of the BAF Testbed to validate a MAC protocol and in the case of the EXPLOIT Testbed in the context of traffic control and traffic congestion. Here, my intention is to present some examples of how CDV is introduced in a real network.

In the case of the BAF Testbed the results obtained to validate the good behavior of the MAC protocol in terms of transfer delay and 1-point CDV allow me to show how CDV is introduced. In this case, the 1-point CDV curves are the ones that show the CDV behavior of the MAC protocol. Most of the CDV is introduced by the Permit Distribution Algorithm (PDA), the multiplexing with background traffic and the Global FIFO queue. In the case of a zero background load, the CDV is introduced by the PDA since an NT must wait for a request block before it can inform the LT about new arrivals. In the case of presence of background traffic, the CDV is introduced by the PDA together with the multiplexing of background traffic and the Global FIFO.

In the EXPLOIT Testbed (ETB) the Cell Transfer Delay (CTD) of a reference connection crossing a tandem of switches has been measured. The reference connection has crossed three ATM segments, the ETB, the Swiss ATM pilot network and the EPFL Communications Laboratory. Each one of these segments have different switch architectures (FORE systems, PKI or ALCATEL BELL) and traffic conditions. We have used models proposed in [4] to compute the CDV as a quantile of the CTD CPDF, see subsections 2.4.1 and 2.4.2. To load the network several traffic techniques have been used: background traffic recycling, load balancing and traffic duplication. We have tried to minimize the possible effects (e.g. correlated traffic) that these techniques could have in the reference source.

The CDV has been compared with values derived theoretically, e.g. the Erlang-M approximation that overestimates the CDV measured. Having the measured mean and variance of the CTD, we have matched the CDV using the convolution approach and assuming independence at each queue, e.g. fitting a Gamma distribution. We have seen that these methods approximate much better the CDV. However they have the drawback that the delay distribution must be known. That was our case since the Testbed give us the opportunity of measuring delay through the entire network configuration. Furthermore, in the Testbed we also have the CDV introduced not only by queueing but also by the physical layer, multiplexing with OAM cells, switch processing and switch architecture. We have shown, see the experiments without background traffic, that the PDH links introduced non negligible amount of CDV.

The experiments, both in the BAF Testbed and the EXPLOIT Testbed, were limited by the limiting equipment, mainly in Test Generator/Analyzers. The limiting of this kind of equipment made that the experiments were influenced by correlated traffic as a consequence of using optical splitters in order to increase the load in the multiplexors. To decrease the effect of correlated traffic we defined a set of rules to be followed: use of phase-moving sources, less than 50% of background recycled traffic at each stage, bal-

anced loads, duplicated traffic only at different stages were some of the solutions applied to our network configuration. However, I think that we achieved a relative fair configuration in spite of the limitations. Other drawback we had to face and that to my knowledge was shared by other European projects was the lack of real users who could make more satisfactory our results. Now, this is one of the main goals in European Testbeds.

We must mention that the results do not present confidence intervals. This is due to the small number of experiments performed in the Testbed. Only one experiment represents a huge amount of time. For example to fulfil the memory of the HP analyzer (130 Kcells) in the BAF Testbed with a 64 Kbit source can represent around 15 minutes, however. loading this memory in a computer, represents around 40 minutes.

The contribution of this chapter to the thesis is: a experimental study of the CDV accumulated on real ATM Testbeds. The methods proposed to compute the CDV come from the literature (e.g. use of the Erlang-M distribution or to fit delays to a Gamma distribution assuming independence). The results depend on the specific set up configuration of the Testbeds. With both groups, BAF group and EXPLOIT group, we had to learn and develop a set of skills to apply the theoretical background in a real environment. Therefore, I consider very profitable all the experience gained in the set up of the experiments, not only in terms of results but also in terms of technology. We were in constant contact with most of the manufactures of the prototypes that gained with the feedback of our experiments.

Chapter 4

The Benes approach applied to ATM Networks.

This chapter is devoted to the Benes approach to the Virtual Waiting Time as a tool to calculate several ATM queue systems. The Benes approach has been used in many systems either at cell level (the $M/D/1$, the $ND/D/1$ or the $M+D/D/1$ queue systems) and at burst level (bursts Poisson distributed or a finite number of On/Off sources with On and Off periods exponentially distributed).

The chapter begins in section 4.1 describing the Benes Approach to the Virtual Waiting Time of a $G/G/1$ queue system. Different systems to which the Benes approach can be applied are also given: fluid arrival processes or the $G/D/1$ queue system are among them. Mathematical proof of the approach for the $G/D/1$ queue system is given together with an intuitive way of understanding the approach. Section 4.2 describes several applications of the Benes approach in the ATM cell level as the $M/D/1$ queue system, $ND/D/1$ or the $M+D/D/1$ queue systems, while 4.3 describes how to apply it to the burst level. Finally, in section 4.4, the Ballot theorem applied to networks may be also used to solve ATM systems at cell level.

4.1 The Benes approach to the virtual Waiting Time.

The Benes approach has been recently applied to study constant service time servers in ATM networks, see [4], [70] and [87]. This approach expresses the *Virtual Waiting Time*, V_s, of a G/G/1 queue and gives an upper bound of the queue length distribution. In constant service time single server systems we can relate the number of customers in the server at time s, L_s, with the unfinished work in the system, V_s, in the following way (being the unit of work the service time of one customer):

$$L_s = \lceil V_s \rceil \ \textit{and} \ P\{L_s > x\} = P\{V_s > x\} \ \textit{being} \ x \ \textit{an integer} \tag{4.1}$$

We assume a system with infinite buffer and service capacity a unit of work per unit of time. In general, let be $N(t)$, $t \geq 0$, the amount of work arriving into the system in the interval [-t,0). Let $\phi(t) = N(t) - t$ be the excess of work arriving in that interval and V_{-t} is the work still in the system at time -t. It can be shown, [4] and [70], that for $t \geq 0$:

$$if \quad V_0 = sup_{t\geq 0}\{\phi(t)\} \quad then \quad P\{V_0 > x\} \geq P\{\phi(t) > x\} \tag{4.2}$$

For stability we require $\phi(\infty) = -\infty$. Giving a partition of the event $\{V_0 > x\}$ according to the last exit time u=sup $\{t \geq 0 : \phi(t) = x\}$ of the excess work function and assuming that the system is stable, the Benes principle concludes that:

$$\{V_0 > x\} \quad \Leftrightarrow \quad \{\exists \ unique \ u \ s.t: \ \phi(u) = x \ and \ \phi(w) < x \ for \ w > u\} \tag{4.3}$$

In those cases in which $\phi(t)$ is continuous and differentiable (see 4.1 for a realization of the process $\phi(t)$):

$$P\{V_0 > x\} = \int_{u>0} P\{\phi(u + du) < x \leq \phi(u) \ and \ \phi(w) < x \ for \ w > u\} \tag{4.4}$$

We can deduce the following equivalence (this deduction is done for the G/D/1 queue system in the following subsections):

$$\{V_{-u} = 0\} \quad \Leftrightarrow \quad \{\phi(w) < \phi(u) \ for \ w > u\} \tag{4.5}$$

Now, we can write equation (4.4) as:

$$P\{V_0 > x\} = \int_{u>0} P\{\phi(u + du) < x \leq \phi(u) \ and \ V_{-u} = 0\} \tag{4.6}$$

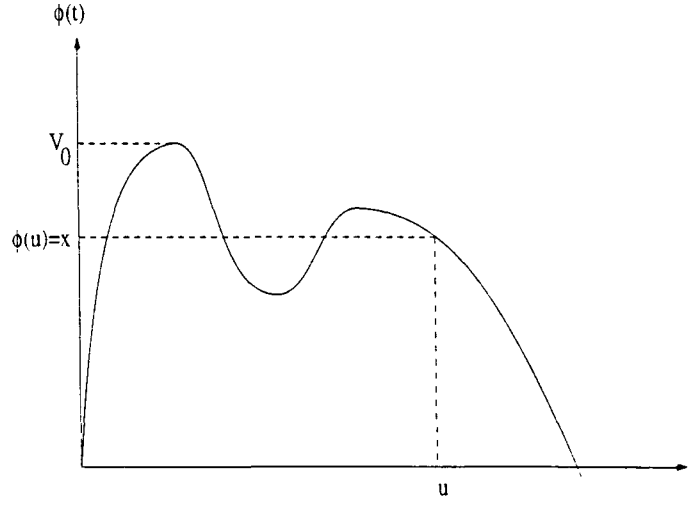


Figure 4.1: Realization of the process $\phi(t)$

Except in a few cases, it is difficult to obtain simple expressions for this integral. However, expressing the probability $P\{\phi(u+du) < x \leq \phi(u) \ \ and \ \ V_{-u} = 0\}$ as the product $P\{\phi(u+du) < x \leq \phi(u) \ \mid \ V_{-u} = 0\} \cdot P\{V_{-u} = 0\}$ bounds can be given by approximation of the term $P\{V_{-u} = 0\}$. For instance, a simple upper bound is obtained by approximating $P\{V_{-u} = 0\}$ by 1.

In those cases in which $\phi(t)$ is not differentiable (e.g. Brownian motion) we refer to [4] for an application to the Benes approach.

The Benes approach can be applied to some interesting systems that appear in the study of ATM networks. Among the most interesting we can summarize:

- **Fluid arrival process.**

  If $N(t)$ is a continuous function, the system behaves as a reservoir with constant output c, see [70]. Let Γ_t be the arrival rate at time -t. Then,

  $$\Gamma_t = \frac{dN(t)}{dt} = 1 + \frac{d\phi(t)}{dt}$$

  Writing $\phi(u+du) = \phi(u) + (\Gamma_u - 1)du$, and summing over all values of Γ_u, see [70], equation (4.6) can be expressed as:

  $$P\{V_s > x\} = \int_{u>0} \int_{\lambda<c} (c-\lambda) \frac{\delta}{\delta x} P\{N(t) \leq x + cu, \ \ \Gamma_u = \lambda, \ \ V_{s-u} = 0\} \ d\lambda \ du \tag{4.7}$$

- **The $G/D/1$ system.**

  Considering the discrete-time system in which work arrives in packets of fixed length called cells, as for example happens in ATM networks, we define $N(s-t, s)$ as the number of cell (client) arrivals in the interval [s-t,s), and the function $\phi(t)$ as $\phi(t) = N(s-t, s) - t$. Thus, the integral in equation (4.4) is substituted by a summatory. Defining L_s as the queue length at time s and applying equivalence (4.1), we arrive at:

  $$P\{L_s > x\} = \sum_{t=1}^{\infty} P\{\phi(t) = x \quad and \quad L_{s-t} = 0\} \tag{4.8}$$

  To arrive to this result, Roberts and Virtamo in [87] give proof of this formula through 4 properties that have function $\phi(t)$. We here give a summary of these properties and their proof for a better understanding of the Benes approach:

  Proof: (see [87])

  - If the system is stable, there is a positive value t such that $N(s-u, s) < u$ for $u > t$. In other words, it has to be some point in a sufficient large interval of time where the arriving work to the system is less than the work served in the system.

- We know that if in an interval of time u the arriving work $N(s-u,s) < u$ the system is empty at time s. Now, It can be shown that at any arbitrary positive time u: $\{\phi(w) < \phi(u) \ \ for \ \ w > u\} \ \Leftrightarrow \ \{L_{s-u} = 0\}$

  We can rewrite $\{\phi(w) < \phi(u) \ \ for \ \ w > u\}$ as:

$$\begin{aligned}\{\phi(w) < \phi(u) \ \ for \ \ w > u\} = \\ \{N(s-w,s) - w < N(s-u,s) - u \ \ for \ \ w > u\} = \\ \{N(s-w,s) - N(s-u,s) < w - u \ \ for \ \ w > u\} = \\ \{N(s-w,s-u) < w - u \ \ for \ \ w > u\} \Rightarrow \ \{L_{s-u} = 0\}\end{aligned}$$

  In the other hand, $L_{s-u} = 0$ means that there is a value $w > u$ for which the arrival work in the interval $(s-w, s-u)$ is less than $w-u$. Then: $N(s-w,s-u) < w-u$. Now it is easy to derive the left-hand inequality.

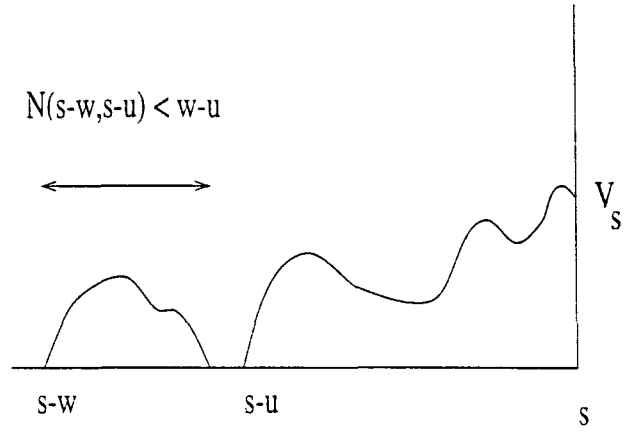


Figure 4.2: Property 2.

- We want to show that if $L_s > x$ there exists a $t > 0$ such that $\phi(t) \geq x$. For stability, there is a t such that the system is empty ($L_{s-t} = 0$). We call t_{min} the smallest $t > 0$ such the last condition holds. For any $t < t_{min}$, the system is in a busy period since there is work to serve. In the interval $[s - t_{min}, s)$, $N(s - t_{min}, s)$ work has arrived and t_{min} work has been served. Therefore, since at time $s - t_{min}$ the system is empty, $L_s = N(s - t_{min}, s) - t_{min} + 1$. From here, we conclude that if $L_s > x$ then $\phi(t_{min}) \geq x$.

- If $L_s > x \Rightarrow$ for any $0 \leq q \leq x$, there exists a $t > 0$ such that $\phi(t) = q$. This comes from the fact that $\phi(t+1) \geq \phi(t)$ if some work arrives in the interval $[s-t-1, s-t)$, and $\phi(t+1) = \phi(t) - 1$ if no work arrives in that interval. If $L_s > x$ and taking into account that $\phi(t)$ is continuous, $\phi(t)$ will take every possible value between $L_s - 1 \geq x$ and 0 as t increases beyond t_{min}.

From these points, we can conclude that if $L_s > x$ there exists an unique point t such that $\{\phi(t) = x \ \ and \ \ \phi(w) < x \ \ for \ \ w > t\}$ or equivalently from the second property $\{\phi(t) = x \ \ and \ \ L_{s-t} = 0\}$.

Another view of the proof.

We have shown the Benes formula following the proof given in [87]. An alternative way of understanding the proof could be the following: We define ξ as any realization of the queue length. We have to show the following relations:

$$\begin{cases} \bigcup_{t=1}^{\infty} \{\phi(t) = x \ \ and \ \ L_{-t} = 0\} \Leftrightarrow \{L_0 > x\} \\ \\ \{\phi(t) = x \ \ and \ \ L_{-t} = 0\} \cap \{\phi(t') = x \ \ and \ \ L_{-t'} = 0\} \ = \varnothing \ \ for \ t \neq t' \end{cases} \quad (4.9)$$

- To show the first relation, $\bigcup_{t=1}^{\infty} \{\phi(t) = x \ \ and \ \ L_{-t} = 0\} \Leftrightarrow \{L_0 > x\}$, we have to show that:

  - $\forall \ t > 0 \ \mid \ \xi \in \{\phi(t) = x \ \ and \ \ L_{-t} = 0\} \ \Rightarrow \ \xi \in \{L_0 > x\}$
    This relation is easy to see since the event $\{\phi(t) = x \ \ and \ \ L_{-t} = 0, \ \ t \geq 0\}$ implies that as a maximum $t-1$ units of work have been served at time 0. Since there has been $N(t, 0) = t + x$ arrivals in the interval [-t,0), at time 0 the queue length is $L_0 > x$ for any $t > 0$.

  - $\forall \ \xi \in \{L_0 > x\} \ \Rightarrow \ \exists \ t > 0 \ \mid \ \xi \in \{\phi(t) = x \ \ and \ \ L_{-t} = 0\}$
    This is like proposition 3 and 4 of Roberts and Virtamo. We assume that $L_0 > x$. We define as t_n with $n \geq 0$ those values of t such the system is empty, being $t_{n+1} < t_n$. Without loss of generality, we say that at t_0 the system is empty and $\phi(t_0) = x'$ with $x' \geq x$. Since, $\phi(t_n) = \phi(t_{n-1}) - 1$ if there are no arrivals, and the system is stable ($\phi(t_\infty) = -\infty$), $\phi(t)$ takes all the values from x' to $-\infty$. Therefore, there are a set of points in which $\{\phi(t) = x \ \ and \ \ L_{-t} = 0\}$.

- $\{\phi(t) = x \ \ and \ \ L_{-t} = 0\} \cap \{\phi(t') = x \ \ and \ \ L_{-t'} = 0\} = \varnothing \ \ for \ t \neq t'$
  We could wonder if the event $\{\phi(t) = x \ \ and \ \ L_{-t} = 0\}$ is included in the event $\{\phi(t') = x \ \ and \ \ L_{-t'} = 0\}$, it is to say, that there is another point where the Benes approach applies. To see that these events are disjoint we make the following reasoning: assume an ξ such that $\xi \in \{\phi(t) = x \ \ and \ \ L_{-t} = 0\}$ and $\xi \in \{\phi(t') = x \ \ and \ \ L_{-t'} = 0\}$ for $t' > t$.

  Being $t' + x$ the number of arrivals over the interval [-t',0), to be empty the system at time -t, the following inequation must holds: $N(t', t) < t' - t$. Under the conditions given before, in the interval $[-t', t)$ have arrived: $t' + x - (t + x) = t' - t$ cells. However, this is contradictory with the fact that we have assumed an empty system at time -t.

In the particular case in which we have a stable queue with periodic input of period T and arrivals in the interval [s-T,s), we can rewrite (4.8) as (see [87]):

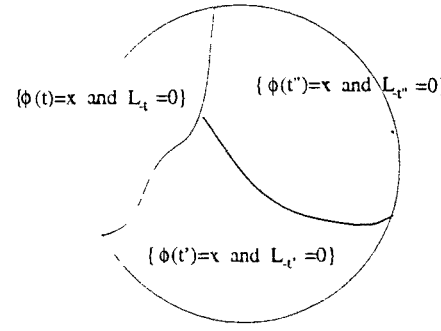


Figure 4.3: Disjoint events.

$$P\{L_s > x\} = \sum_{t=1}^{T} P\{\phi(t) = x \quad and \quad \phi(u) < x, \quad t < u \leq T\} \tag{4.10}$$

To show this formula, we can make the same reasonings as for the general formulation, see [87]. We consider a periodic system of period T handling $N \leq T$ sources. Each source emits a unit of work uniformly distributed in the period. We can now rewrite the properties presented by Roberts and Virtamo in the interval $0 \leq t \leq T$ (where we consider T as an integer).

- Stability: $\phi(T) \leq 0$. This means that $N(s - T.s) \leq N \leq T$.
- $L_s > x \Rightarrow$ there exists a $1 \leq u \leq T$ such that $\phi(u) \geq x$.
- $L_s > x \Rightarrow$ for any q $(0 \leq q \leq x)$ there exists a $1 \leq u \leq T$ such that $\phi(u) = q$.

From these points it is concluded that there exists a unique point $1 \leq u \leq T$ such that $\{\phi(t) = x \quad and \quad \phi(u) < x, \quad t < u \leq T\}$.

4.2 Queueing models at cell level using the Benes approach

The Benes approach to the Virtual Waiting time can be applied to a wide range of queuing models. Here, some of the most used models in ATM networks at cell level are outlined.

4.2.1 The $M/D/1$ system

Poisson arrival processes have been extensively used for queue modeling (e.g. buffer dimensioning). We assume that the cells arrive as a Poisson process. If we choose the service time as the time unit, the arrival intensity is equal to the server load ρ. The $M/D/1$ is a good approximation to the superposition of CBR sources when non of them dominates (e.g. the period of the sources is large enough) if the multiplex load is not very high. This model ignores the correlations at cell level and overestimates congestion

effects in short time (see [87]).

Thus, to calculate the complementary unfinished work distribution of the M/D/1 system applying the Benes approach, we express equation (4.8) at time 0 as $Q(x) = P\{L_0 > x\}$:

$$Q(x) = \sum_{t=1}^{\infty} P\{\phi(t) = x \text{ and } L_{-t} = 0\} = \sum_{t=1}^{\infty} P\{\phi(t) = x \mid L_{-t} = 0\} \cdot \\ \cdot P\{L_{-t} = 0\} = (1-\rho) \sum_{t=1}^{\infty} \frac{(\rho t)^{t+x}}{(t+x)!} e^{-\rho t} \tag{4.11}$$

Roberts and Virtamo in [87] propose an alternative representation of $Q(x)$ changing variables and applying Jensen's theorem (see [89]):

$$\sum_{i=0}^{\infty} \frac{(a+ib)^i}{i!} e^{-(a+ib)} = \frac{1}{1-b} \tag{4.12}$$

provided that $|b\, e^{-b}| < e^{-1}$.

We have to change variables in equation (4.12): $t \to t - x$ and:

$$Q(x) = (1-\rho) \sum_{t=x+1}^{\infty} \frac{(\rho(t-x))^t}{t!} e^{-\rho(t-x)} \tag{4.13}$$

Summing and resting from $t = 0$ to $t = x$:

$$Q(x) = (1-\rho) \sum_{t=0}^{\infty} \frac{(\rho(t-x))^t}{t!} e^{-\rho(t-x)} - (1-\rho) \sum_{t=0}^{x} \frac{(\rho(t-x))^t}{t!} e^{-\rho(t-x)} \tag{4.14}$$

Now, we can apply Jensen's theorem to the first summatory taking into account that: $a = -\rho x$, $b = \rho$ and $i = t$:

$$Q(x) = 1 - (1-\rho) \sum_{t=0}^{x} \frac{(\rho(t-x))^t}{t!} e^{-\rho(t-x)} \tag{4.15}$$

However, loss of accuracy due to the cancelation of terms, appears in the calculation of the complementary unfinished work distribution of the M/D/1 system applying the Benes approach, formula (4.15). This problem appears for large values of x when ρ is closed to 1. Virtamo, in [91], solves this numerical problem through a polynomial representation of the unfinished work distribution.

We must express $Q(x) = 1 - (1-\rho)\, e^{\rho x}\ P_{\lfloor x \rfloor}(x - \lfloor x \rfloor)$, where $P_n(x)$ is the polynomial $P_n(x) = \sum_i a_i x^i$ with coefficients $P_n = (a_0, a_1, ..., a_n)$. Now, beginning with $P_0 = (1$ and defining $\beta = \rho\, e^{-\rho}$, polynomial P_{n+1} can be calculated recursively as:

$$P_{n+1} = (\sum_i a_i, -\beta a_0/1, -\beta a_1/2, ..., -\beta a_n/(n+1))$$

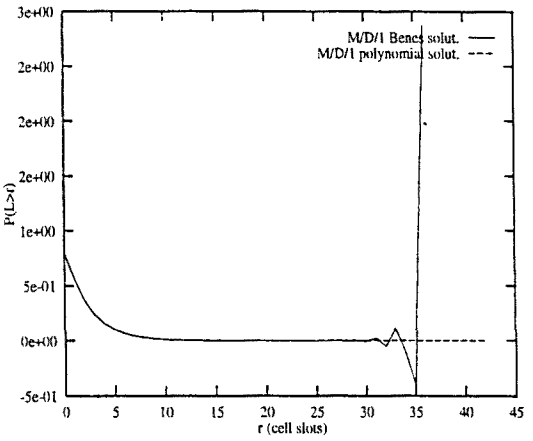


Figure 4.4: Loss of accuracy due to cancelation of terms in the M/D/1 model for $\rho = 0.8$

We can observe in figure 4.4 the loss of accuracy produced in the M/D/1 final equation (4.15). For a load of $\rho = 0.8$ and values of $x > 30$ the cancelation of terms produce probabilities which oscillate between negative values and positive values higher than 1. With the polynomial representation the numerical problem is solved.

4.2.2 The $N \cdot D/D/1$ system

This model takes into account the periodic nature of the sources, and thus, the correlation at cell level. although has the limitation of assuming identical periods for all the sources (homogeneous traffic). The input process consists, see [87], of a superposition of N independent sources of identical period T, uniformly distributed over the period. In contrast with the $M/D/1$ system, where the arrivals were uncorrelated, in the $N \cdot D/D/1$ successive cell arrivals are negatively correlated expecting shorter queue lengths. Thus, given the same load, the $M/D/1$ is commonly used as an upper bound to the $N \cdot D/D/1$.

We consider the unfinished work in the system at time 0 to be V_0. Since we assume an stable system ($N < T$), for any arrival phase, there has to be some instant in the interval [-T,0) where the queue is empty. That means that V_0 only depends on the arrivals in this interval and the arrivals in the previous period can be "switch-off" (it is as if the system begins empty). Considering as the only arrivals the N independent sources over the period T. we apply equation (4.10):

$$P\{L_0 > x\} = \sum_{t=1}^{T} P\{N(0,t) = t+x\} \cdot P\{\phi(u) < x,\ t < u \leq T \mid N(0,t) = t+x\} \tag{4.16}$$

$P\{N(0,t) = t-x\}$ represents the probability that arrives exactly t+x cells in the interval [-t,0). Since the sources are uniformly distributed, this probability follows a binomial distribution:

$$P\{N(0,t) = t+x\} = \binom{N}{t+x} (\frac{t}{T})^{t+x} \cdot (1-\frac{t}{T})^{N-t-x} \tag{4.17}$$

$P\{\phi(u) < x,\ t < u \leq T \mid N(0,t) = t + x\}$ represents the probability that the system is empty at time -t, giving that $t + x$ cells have arrived in the interval [-t,0). We can now construct an auxiliary system, where arrives $N - t - x$ cells uniformly distributed over a period of length $T - t$. This system has to be empty at time 0. Therefore, this probability is $1 - \rho'$, being ρ' the load of the auxiliary system: $\rho' = (N - t - x)/(T - t)$.

Then, equation (4.16) corresponds to:

$$P\{L_0 > x\} = \sum_{t=1}^{N-x} = \binom{N}{t+x} (\frac{t}{T})^{t+x} \cdot (1 - \frac{t}{T})^{N-t-x} \cdot (\frac{T - N + x}{T - t}) \tag{4.18}$$

Average waiting time and average number in the queue in an $N \cdot D/D/1$ system queue.

Dron et al show in [38] that the average waiting time in the queue, $\overline{W}$, of a queue system multiplexing N periodic sources of period T_{cbr} (with load $\rho = \frac{N}{T_{cbr}}$) is the following:

$$\overline{W} = \frac{(N-1)!}{2} \sum_{k=1}^{(N-1)} \frac{1}{T_{cbr}^k} \frac{1}{(N-1-k)!} \tag{4.19}$$

This result is obtained using the Laplace-Stieltjes Transform (LST) of the virtual waiting time distribution of the $nD/G/1$ queue system found in [88].

The number of cells in the queue $\overline{Q}$ can be calculated using Little's formula: in periodic systems this can be interpreted as following, see [55]:

$$\begin{gathered} P\{W = x\} = P\{a\ departure\ leaves\ x-1\ cells\ in\ the\ queue\} = \\ \frac{P\{Q=x\}}{P\{Q>0\}} = \frac{T_{cbr}}{N} P\{Q = x\}\ \ with\ \ x > 0 \end{gathered} \tag{4.20}$$

Therefore applying this relation:

$$\overline{Q} = \rho \cdot \overline{W} = \frac{N!}{2T_{cbr}} \sum_{k=1}^{(N-1)} \frac{1}{T_{cbr}^k \cdot (N-1-k)!} \tag{4.21}$$

Idle and Busy periods in an $N \cdot D/D/1$ system queue.

In [92], Virtamo gives expressions for the busy and idle periods of an $N \cdot D/D/1$ system queue. Let T be the period of the sources and N the number of sources. To calculate the idle period distribution, the joint event that characterizes an idle period is defined by the following events (A, B and C):

A: there is an arrival in the interval [0,dt) (for periodicity is equivalent to an arrival in the interval [T,T+dt)).

B: there are non arrivals in the interval $[T-l, T)$ (is equivalent to all the N-1 arrivals occur uniformly distributed on the interval $[0, T-l)$).

C: the system is empty at time $T-l$.

If $Q_{idle}(l) = P\{length\ of\ idle\ period \geq l\}$ is the complementary idle period distribution. Virtamo shows that $Q_{idle}(l) = \frac{d\bar{Q}_{idle}(l)}{d\bar{Q}_{idle}(0)}$, being $d\bar{Q}_{idle}(l) = P\{A \cap B \cap C\} = P\{A \cap B\} \cdot P\{C \mid A \cap B\}$. Following these definitions:

$$Q_{idle}(l) = (1 - \frac{l}{T})^{N-2} \cdot (1 - \frac{l}{T-N+1}) \tag{4.22}$$

Note that the maximum idle length can be $l = T-N$ with non-zero probability. Virtamo also give expressions for the average number of idle and busy periods n_{idle} in the interval period T and for the average idle length $\bar{l}_{idle}$:

$$n_{idle} = \frac{N}{T} \cdot (T-N+1)$$

$$\bar{l}_{idle} = \frac{T-N}{n_{idle}} = \frac{T}{N} \cdot \frac{T-N}{T-N+1}$$

For the busy periods distribution $q_{busy}(l) = P\{length\ of\ busy\ period = l\}$, the following events are defined:

A: the system is empty before time 0.

B: there is an arrival at time 0.

C: there are $l-1$ uniformly distributed arrivals in the interval $[0, l)$.

D: the system is empty at time l.

Now, the busy period distribution can be expressed as: $q_{busy}(l) = P\{C \cap D \mid A \cap B\} = \frac{1}{P\{A \mid B\}} \cdot P\{C \mid B\} \cdot P\{A \cap D \mid B \cap C\}$, and it can be shown that:

$$q_{busy}(l) = \binom{N-1}{l-1} (\frac{T-N}{T-N+1}) \cdot (\frac{l^{l-2}(T-l)^{N-l-1}}{T^{N-2}}) \tag{4.23}$$

To calculate the average busy length $\bar{l}_{busy}$:

$$\bar{l}_{busy} = \frac{N}{n_{idle}} = \frac{T}{N} \cdot \frac{T}{T-N+1}$$

4.2.3 The modulated $N \cdot D/D/1$ system.

A model that can be applied to shaped SBR WCT is the modulated $N \cdot D/D/1$ system. A SBR WCT source is an On/Off source with bursts at two levels if $\tau > (T_{PCR} -$

1): the On/Off periods duration are controlled by the GCRA(T_{SCR},τ_{IBT}), and in the On period there are bursts of back-to-back cells followed by silences governed by the GCRA(T_{PCR},τ). In the case of shaped sources, $\tau < (T_{PCR} - 1)$, the SBR WCT source is modeled as an On/Off source, emitting cells spaced at the declared PCR, T_{PCR}. The number of cells on the *On* period is given by the Maximum Burst Size ($MBS = 1 + \lfloor \frac{\tau_{IBT}}{T_{SCR} - T_{PCR}} \rfloor$).

If $P\{N > T_{PCR}\} \leq \epsilon$ (with N the number of sources), the burst scale queueing is almost inexistent and the $P\{L_0 > x\}$ can be approximated by the modulated $N \cdot D/D/1$ system.

Consider the superposition of N independent On/Off sources with intercell spacing T_{PCR}, see figure 4.5. The burst length is multiple of T_{PCR} and the silence length can be of any length. We consider that the burst begins with a cell arrival and ends with a silence of T_{PCR} slots. The activity factor is determined by the expression: $\alpha_i = \frac{T_{on}}{T_{on}+T_{off}}$.

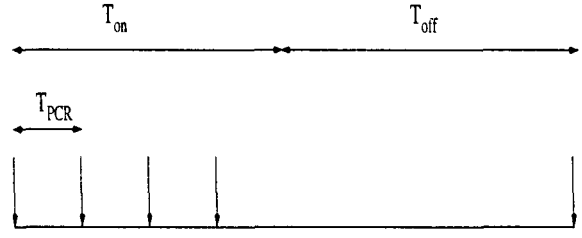


Figure 4.5: Modulated process.

Note that we can apply the Benes approach since always there is an empty slot between a period of T_{PCR} slots (e.g. the system is stable: $N < T_{PCR}$). If $P_n = P\{N(0, T_{PCR}) = n\}$ is the number of cell arrivals in a period T_{PCR}, it is to say, the number of sources active will be given by the convolution of the N Bernoulli variables α_i. Conditioning in the number of arrivals $N(0, T_{PCR})$:

$$P\{L_0 > x\} = \sum_{n=0}^{N} P\{N(0, T_{PCR}) = n\} \cdot \sum_{x<k\leq n} P\{N(T', T_{PCR}) = k \ \ and \ \ L_{T'} = 0 \mid N(0, T_{PCR}) = n\} \tag{4.24}$$

where $T' = T_{PCR} - k + x$. Note that $P\{N(T', T_{PCR}) = k \ \ and \ \ L_{T'} = 0 \mid N(0, T_{PCR}) = n\}$ is the virtual waiting time of a $n \cdot D/D/1$ system that we represent as $Q^n_{T_{PCR}}(x)$. Therefore, equation (4.24) comes to:

$$P\{L_0 > x\} = \sum_{n=0}^{N} P_n \cdot Q^n_{T_{PCR}}(x) \tag{4.25}$$

If all the sources are equal, the activity factors are the same ($\alpha_i = \alpha$) and P_n is given by a binomial and we can conclude that:

$$P\{L_0 > x\} = Q^n_{T_{PCR}/\alpha}(x) \tag{4.26}$$

4.2.4 The $M + D/D/1$ queue system.

A model usually applied to the study of the Cell Delay Variation of a CBR source in an ATM multiplexer is the $M+D/D/1$ queue system, see [86]. In this system a CBR source is multiplexed with Poisson background traffic of load ρ_p. Note that the Poisson traffic is the Worst Case situation for the multiplexing of shaped CBR sources (e.g. those ones with $\tau < (T-1)$).

We assume that the CBR period is T and the Poisson arrival intensity is equal to the Poisson load ρ_p since the service time is the time unit. Let $q(x,y) = P\{L_i = y \mid L_{i-1} = x\}$ be the transition probabilities for the queue length at the arrival of cell *i-th* to the queue server. Therefore: $Q(x,y) = P\{L_i > y \mid L_{i-1} = x\}$ and we can express: $q(x,y) = Q(x,y-1) - Q(x,y)$.

We define the conditional probability $Q_k(x,y)$ as:

$$Q_k(x,y) = P\{L_i > y \mid L_{i-1} = x \; and \; k \; Poisson \; arrivals \; in[(i-1)T, iT)\}$$

and:

$$Q(x,y) = \sum_{k=0}^{\infty} Q_k(x,y) \frac{(\rho_p T)^k}{k!} e^{-\rho_p T} \tag{4.27}$$

Since L_i is a Markov chain we have to solve the state equations to calculate the queue length distributions $p(y) = P\{L_i = y\}$:

$$p(y) = P\{L_i = y\} = \sum p(y) q(x,y)$$

The only term that remains to be calculated is $Q_k(x,y)$. We have three cases:

- If $x+1 \geq T$ and the number of Poisson arrival is $k \leq y-(x+1-T)$ or if $x+1 < T$ and $k \leq y$ there is no possibility to reach y and $Q_k(x,y) = 0$.
- If the number of Poisson arrivals is $k > y-(x+1-T)$ the value y will be always reached and $Q_k(x,y) = 1$.
- The rest of the situations corresponds to those ones where the queue length at arrival of cell (i-1)-*th*, $L_{i-1} = x$ plus the cell arrived at that time are consumed before the period ends: it is to say $x+1 < T$. Then for $y < k \leq y-(x+1-T)$. $Q_k(x,y)$ is the queue distribution $P\{L_0 > y\}$ of an $N \cdot D/D/1$ queue system with k arrivals in a period T.

$$Q_k(x,y) = \sum_{t=1}^{k-y} \binom{k}{t+y} (\frac{t}{T})^{t+y} \cdot (1-\frac{t}{T})^{k-t-y} \cdot (\frac{T-k+y}{T-t})$$

In [86]. there are given expressions for the inter-exit time distributions: $P\{U_i = u\}$, where $U_i = W_i - W_0$. being W_i the waiting time of cell i.

4.2.5 Other queue systems at cell level.

There are more queue systems that can be modeled at cell level with the Benes approach. These models can be found in COST 242, [4]. For instance the $\sum D_k/D/1$ to model the superposition of heterogeneous periodic sources or the $\sum N_k D_k^{X_k}/D/1$ to model batch arrivals. We can also find in [4] bounds and asymptotic results for some of the different queueing systems described in this chapter.

4.3 Queueing models at burst level using the Benes approach

At burst level, J.W. Roberts *et al*, in [26] and [85], apply the Benes bound for fluid systems to the superposition of independent on/off sources, both for the finite source model and for the infinite source model.

For the finite source model, they consider N on/off independent sources. Source i is defined as follows: successive burst and silence durations are independent with densities $b_i(.)$ and $a_i(.)$ respectively. The expected silence duration is $1/\lambda_i$ and the expected burst duration is $1/\mu_i$. During a burst, work is generated at rate h_i.

Applying equation (4.7) to the finite source model, they arrive to:

$$P\{V_0 > x\} = \int_{u>0} \sum_{\mathbf{d}\cdot\mathbf{h}<c} (c - \mathbf{d}\cdot\mathbf{h}) \frac{\delta}{\delta x} P\{N(t) \leq x + cu,\ \mathbf{D_u} = \mathbf{d},\ V_{-u} = 0\}\ du \tag{4.28}$$

Where $D_t(i) = \mathbf{1}\{source\ i\ is\ on\ at\ -t\ \}$, and the scalar $\mathbf{d}\cdot\mathbf{h}$ gives the arrival rate. For the infinite source model, assuming there are $K_{-u} = k$ bursts in progress at time $-u$:

$$P\{V_0 > x\} = \int_{u>0} \sum_{k\cdot\mathbf{h}<c} (c - k\cdot h) \frac{\delta}{\delta x} P\{N(t) \leq x + cu,\ \mathbf{K_u} = \mathbf{k},\ V_{-u} = 0\}\ du \tag{4.29}$$

See [26] and [85] for further information if the interarrivals densities are exponentially distributed and whether the burst rate is greater than or less than the output rate c.

4.4 The ballot theorem applied to periodic queues.

The Benes approach is not the only tool to analyze periodic queue systems, although it has been proved a very powerful tool. *P. Humblet et al* present in [55] an alternative solution to the $nD/D/1$ queue system based on the ballot Theorem of *Takacs*. Here we present a summary of the method.

4.4.1 The ballot theorem.

The classic ballot theorem introduces the following situation. There are two candidates to a voting. Candidate A obtains n votes and candidate B obtains m votes ($n > m$).

What is the probability that candidate A goes ahead always during the voting ? With this idea in mind. Takacs rewrites the ballot theorem to be applied in networks and Humblet *et al* make use of Takacs work to solve periodic systems.

To present the ballot theorem *P. Humblet et al* start with the following problem: water is provided to a tub with an instantaneous rate that is either 0 or at least S liters/hour. The tub empties at a rate of S liters/hour. If the total amount of water delivered in an hour is V liters ($V \leq S$), the fraction of time during which water flows is V/S. To proof this statement they define an integrable real periodic function $a(\cdot)$ with period T and that takes values that are either 0 or greater than or equal than $S > 0$.

Defining $V = \int_0^T a(u)du$ and assuming $V \leq ST$, they proof that the measure of time in which there are no overflow is:

$$A = \left\{ \; t \in [0,T) \;\; | \;\; \int_t^{t+r} a(u)du \leq Sr, \;\; \forall r \in [0,T) \; \right\} \;\; is \;\; T - V/S$$

From this start. they deal with integer-valued functions on the integers. We present $\frac{1}{2}$ theorems used later to solve the $nD/D/1$ queue system. We here skip the proof. and recommend [55] for a deeper study of the method and proof of the theorems:

Theorem 1 : Let $a_1, a_2, ..., a_N$ be nonnegative integers with sum equal to $K \leq N$. Among the N cyclic permutations of $(a_1, a_2, ..., a_N)$, there are exactly $N - K$ for which the sum of the first r elements never reaches r for all $r \in [1, N]$.

Theorem 2 : If $(A_1, A_2, ..., A_N)$ are cyclically interchangeable random variables with sum equal to K, then:

$$Pr\left[\sum_{j=1}^{i} A_j < i, \;\; \forall i \in [1,N] \right] = \left\{ \begin{array}{ll} 1 - K/N & if \;\; K < N \\ 0 & otherwise \end{array} \right.$$

Theorem 3 : If $(A_1, A_2, ..., A_N)$ are cyclically interchangeable random variables taking on nonnegative integral values, then for $u > 0$:

$$Pr\left[\sum_{j=1}^{i} A_j > i - u, \;\; \forall i \in [1,N] \right] = 1 - \sum_{r=u}^{N} \frac{u}{r} Pr\left[\sum_{j=1}^{r} A_j = r - u \right]$$

Theorem 4 : If $u \geq 0$ is an integer and $(A_1, A_2, ..., A_N)$ are cyclically interchangeable random variables taking on nonnegative integral values summing to $K < N - u$, then:

$$Pr\left[\sum_{j=1}^{i} A_j < i + u, \;\; \forall i \in [1,N] \right] = 1 - \sum_{r=1}^{K-u} \frac{N - K + u}{N - r} Pr\left[\sum_{j=1}^{r} A_j = r + u \right]$$

4.4.2 The ballot theorem applied to the $ND/D/1$ queue system.

We consider a slotted system axis with slot length the unit time. K sources of the same period M send packets to the system independently and uniformly distributed of each other ($K < M$ to be the system stable). The queue serves packets in a FIFO discipline.

Defining as A_t the number of arrivals in the t-th slot and noting that the pattern is periodic ($A_t = A_{t\ mod\ M}$), we can express the queue length at the end of the t-th slot Q_t as:

$$Q_t = \begin{cases} 0 & t = 0 \\ max\{(Q_{t-1} - 1), 0\} + A_t & t > 0 \end{cases}$$

and iterating on t:

$$Q_t = max_{0 \le i < t}(\sum_{j=t-i}^{t} A_j - i) \qquad t \ge 0$$

taking into account that the system is periodic of period M, and $Q_0 = Q_M$:

$$Q_0 = max_{0 \le i < M}(\sum_{j=-i}^{0} A_j - i) \tag{4.30}$$

Now, we can use this equation to calculate $P\{Q_0 > x\}$:

$$P\{Q_0 > x\} = 1 - P\{Q_0 \le x\} = 1 - P\{\sum_{j=-i}^{0} A_j - i \le x,\ \forall i \in [0, M]\}$$
$$= 1 - P\{\sum_{j=1}^{i'} A_{1-j} < x + i',\ \forall i' \in [1, M]\} = \sum_{r=1}^{K-x} \frac{M-K+x}{M-r} P\{\sum_{j=1}^{r} A_{1-j} = x + r\} \tag{4.31}$$

where we have used theorem 4. If we consider that the packets are uniformly and independently distributed over the period, we can express $P\{\sum_{j=1}^{r} A_{1-j} = x + r\}$ as a binomial distribution and we conclude equation (4.31) as:

$$P\{Q_0 > x\} = \sum_{r=1}^{K-x} \frac{M-K+x}{M-r} \binom{M}{r+x} (\frac{r}{M})^{r+x} (1 - \frac{r}{M})^{M-r-x} \tag{4.32}$$

As can be seen, the ballot theorem applied to a periodic system obtain the same kind of formula as the Benes approach. I have considered more intuitive the Benes approach. Therefore, I have chosen this approach to study the superposition of WCT sources.

Chapter 5

Multiplexing Worst Case Traffic (WCT) Sources.

Due to the jitter captured in the ATM connection, a source must declare tolerances in order to conform the UPC mechanism. In the case of a DBR connection the tolerance is called Cell Delay Variation (CDV) tolerance, τ, and in the case of a SBR connection its CDV tolerance. τ_{SCR}. Given a traffic contract there are many traffic patterns that can conform to that contract. In this chapter a worst case traffic (WCT) definition is given conforming to the UPC mechanism. An analytical model based on the Benes approach is developed to evaluate the worst case pattern in a DBR connection.

The chapter is organized as follows. section 5.1 defines the worst case traffic compatible with a Leaky Bucket control. A model that calculates the CPDF of the Virtual Waiting Time in a single slotted queue system that multiplexes WCT sources is developed in section 5.2. A fluid version is solved in section 5.3. A batch model that can be used as an upper bound is described in section 5.4. Finally, in section 5.5 several examples are given to evaluate the behavior of WCT sources in terms of network resources.

5.1 Worst Case Traffic compatible with Leaky Bucket control.

Interest in the study of WCT models proceeds from problems raised by traffic control for Deterministic Bit Rate (DBR) and Statistical Bit Rate (SBR) traffic sources in ATM networks. Due to the jitter introduced in the network, a DBR source declares a Cell Delay Variation Tolerance (CDV-T) together with other traffic parameters as the PEI. For a SBR source we saw that two more traffic parameters were declared: the Sustainable Cell Rate (SCR) and Intrinsic Burst Tolerance (IBT), see chapter 2. The Generic Cell Rate Algorithm (GCRA) was used to set these traffic parameters to achieve the so called negligible non-conformance rate.

However, given these conditions we can see how based on ruled traffic parameters we can define the *worst case* conditions for the network. The tolerance introduced in the GCRA mechanism has a collateral effect which complicates the congestion control in ATM networks. Due to the tolerance declared, the GCRA mechanism allows to pass

many different traffic patterns which can have very different effects on network congestion. For example, the GCRA can declare as conforming a traffic pattern consisting of back-to-back cells in periodic bursts, i.e. *Worst Case Traffic* (WCT). As we shall see, this type of traffic can produce much greater levels of congestion than those obtained by the superposition of cells emitted periodically and affected by a certain random jitter (hence the name). If the network is dimensioned for DBR traffic without bursts, but the users actually emit WCT, the Quality of Service (QoS) of the network can be seriously degraded. Hence the concern to quantify the possible effect of WCT on the network QoS.

5.1.1 Worst Case traffics based on traffic parameters.

A Worst Case Traffic connection is defined as that one that all its cells are conforming to the negotiated traffic, but that requires the greatest amount of resources.

For CBR connections, conformance with the Traffic Contract is defined by means of PEI (T) and CDV-T (τ). There are many traffic patterns which are compatible with the same values of PEI and CDV-T. For example:

- The source originates cells in a periodic form with an inter-cell distance of T slots. By the time these cells reach the network they have been affected by CDV and suffer a jitter picked up by τ.
- The source emits periodic bursts of back-to-back cells, see figure 5.1. The bursts have a Burst Size (B_s) length which must fulfill $B_s \leq \lceil 1 + \frac{\tau}{T-1} \rceil$ and the silences have a maximum length of $SL = B_s \cdot T - B_s$. When $B_s = \lceil 1 + \frac{\tau}{T-1} \rceil$ we say that the source emits a Worst Case Traffic (WCT).

The first type of traffic could correspond to that generated by a CBR coder for voice or video. The second type of traffic could be generated by certain coders or by "tricky" users.

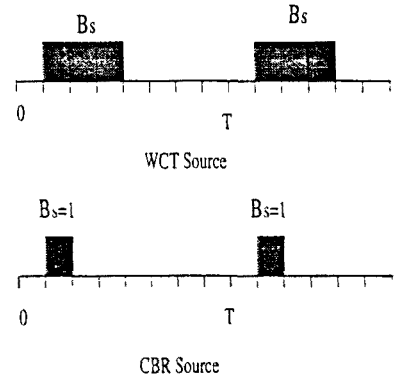


Figure 5.1: CBR and WCT periodic sources.

Although both types of traffic fulfill the conditions of the contract, they produce very different effects as regards network congestion. Specifically, as we will see, if we want to keep network congestion under control, the possible presence of WCT causes grave

problems. In fact we can consider that WCT is the worst situation we can have for traffic which has declared given values of PEI and CDV-T. One could devise other traffic patterns that produce slightly worst effects on network congestion, but their construction is much more artificial and less tractable (e.g. see [16]).

The study of WCT is important for several factors. First, as the ITU-T 371 states, the network can not relay in the fact that the CBR application is going to send periodic traffic. The only thing that the network knows is that the CBR application can send traffic according to the traffic contract (idem for VBR applications). A misbehaving customer can try to take advantage of the UPC function and send bursts of cells fulfilling the traffic contract but demanding most resources that expected. Adaptation ATM Layer (AAL) schemes can also introduce WCT in a natural way. See as an example [28], where a multimedia workstation is assumed to send several AALs each of them generating CBR traffic at different peak rates and accumulating data in their own queue. The queues are periodically emptied at a rate of 150 Mb/s with scanning period set to the lowest involved peak emission period. Therefore, if the network has to ensure QoS objectives, these possible network conditions must be taken into account (e.g. the CAC might take into account WCT, the buffers might be greater or traffic shaping might be used to smooth the WCT bursts).

Related works on Worst Case Traffic are the following: *Mignault et al* give in [63] a survey on WCT models both for CBR and VBR sources. The same team study in [27] and [50] resource allocation for worst case traffic models taking into account both call and transfer level QoS. They identify three frameworks to allocate resources: the PCR allocation, the burst loss and the burst delay framework. As WCT models they propose the $M/D^{[k]}/1/N$ queue model and a fluid model presented by *Roberts et al* in [85] using the Benes bound to study the burst scale behavior in a queue system fed by WCT on/off sources in which traffic in the on period has been shaped.

Elwalid et al give in [41] an approach based on the Chernoff bound and and Large Deviation Approximations for leaky-bucket regulated traffic. They consider that the output of a leaky-bucket regulator is extremal and on/off periodic with independent and uniformly distributed random phases. In [64], the work of *Elwalid et al* is compared with the fluid model presented in [85].

All of them solve systems with shaped traffic. For systems with non-shaped traffic applied to DBR connections, we can use a batch arrival system, see [4], or an approximation based on the ND/D/1 queue system presented by Ramamurthy and Dighe in [78]. These models are summarized in section 5.4.

In this chapter, we solve an exact model for a multiplexer fed by homogeneous periodic WCT (with back-to-back cells) sources using the Benes approach to the Virtual Waiting time. We give the CPDF of the Virtual Waiting time for the discrete-time model and for a fluid-flow approximation, see sections 5.2 and 5.3 respectively. We will compare the WCT result with models presented in the literature as the bounds proposed in [78] based on the ND/D/1 queue system and a batch queue system model, see section 5.5. Finally in this section, we evaluate the performance of a multiplexor studying the impact that

have the WCT sources in the buffer occupancy. The content of this chapter can be found in [44] and [45].

5.2 Multiplexing WCT sources in a slotted queue.

We consider a multiplexer with a service time equal to an ATM slot loaded with N identical and independent WCT sources, [44] and [45]. Each WCT source produces a periodic stream of cells, of period T, with the following pattern: It emits a constant number, b, of back-to-back cells and then it remains silent during a constant time $T-b$, (b is the Burst Size, B_s). The time slots in which each source becomes active are uniformly and independently distributed within the period, see figure 5.1. In order to have a stable queue, we assume that $\frac{N \cdot b}{T} < 1$. We consider that arrivals take precedence over departures (i.e. first we have cell arrivals (if any), we observe the system, and finally then we have the service (if any)).

5.2.1 Waiting time Distribution.

Having periodic input traffic of period T we can use equation (4.10) to express the complementary distribution of the queue length. If we call B_{-t} the number of sources which become active in the interval $[-t, 0)$, equation (4.10) can be rewritten as:

$$\begin{aligned} P\{L_0 > x\} &= \sum_{t=1}^{T} \sum_{i=1}^{N} P\{N(t,0) = t + x | N(u,0) < u + x, t < u \leq T, B_{-t} = i\} \cdot \\ &\quad \cdot P\{N(u,0) < u + x, t < u \leq T | B_{-t} = i\} \cdot P\{B_{-t} = i\} \\ &= \sum_{t=1}^{T} \sum_{i=1}^{N} A(t, t+x, i) \cdot L(t, i) \cdot B(t, i) \end{aligned} \tag{5.1}$$

The term A(t,t+x,i):

The term A(t,t+x,i) will be given by the convolution of the i bursts that become active in the interval $[-t, 0)$. Depending on the value of t, we consider three regions (see figure 5.2 where 4 burst arrivals of size b cells are considered):

- Region I: If $t < b$, the bursts will contribute in part to the next period. We must find by convolution, the number of cells that contribute between 0 and $-t$.

  We define $q_m(t)$ as the discrete-time unitary pulse in the interval [1,m], and $q_m^{(i)}(t)$ is its i-th discrete-time convolution. For this region, the contribution of the i active sources to the term $A(t, t+x, i)$ will be:

  $$A(t, t+x, i) = \frac{1}{t^i} q_t^{(i)}(t+x)$$

- Region II: If $b \leq t < T - b$, some bursts will contribute with all their cells (those ones that become active at time $t \geq b$), while some ones will contribute with part

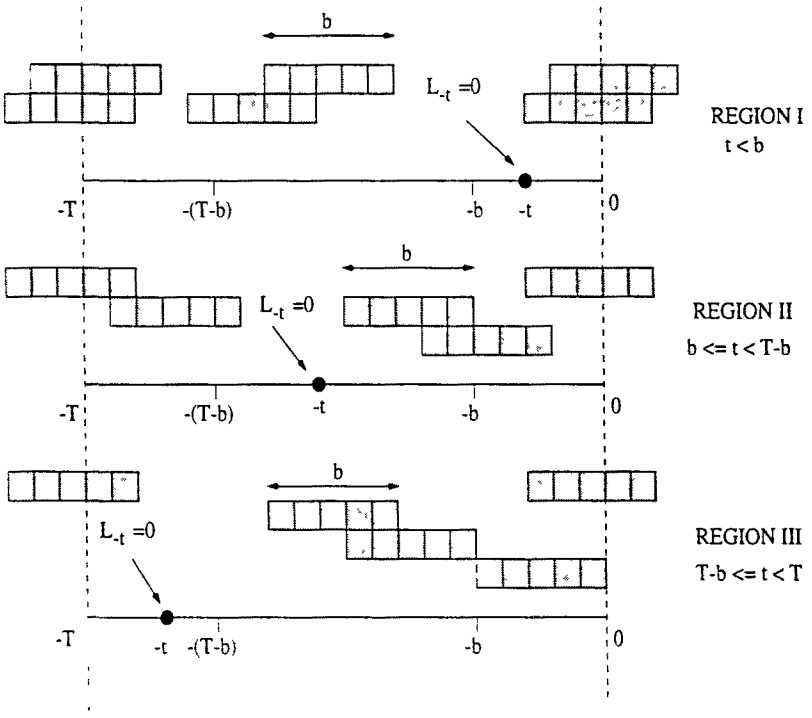


Figure 5.2: Regions to calculate the term A(t.t+x,i).

of the burst as in region I (those ones that become active at time $t < b$).

For a given time t in this region, a source can be active at time $t < b$, contributing with a pulse such as $\frac{1}{t} q_{b-1}(t+x)$, or a source can be active at time $b \leq t < T - b$. contributing with an impulse: $\frac{t-(b-1)}{t} \delta(t+x-b)$, being $\delta(y)$ the discrete-time impulse function.

Therefore. the contribution of i bursts being active at any time in the interval [-t,0) will be:

$$
\begin{aligned}
&A(t.t-x,i) = \left(\tfrac{1}{t}\, q_{b-1}(t+x) + \left(\tfrac{t-b-1}{t}\right) \delta(t+x-b) \right)^{(i)} = \\
&= \sum_{j=0}^{i} \binom{i}{j} \left(\tfrac{1}{t^{i-j}} q_{(b-1)}^{(i-j)}(t+x) \right) * \left(\tfrac{(t-b+1)^j}{t^j} \delta(t+x-bj) \right) = \\
&= \sum_{j=0}^{i} \binom{i}{j} \tfrac{(t-b-1)^j}{t^i} q_{(b-1)}^{(i-j)}(t+x-bj)
\end{aligned}
$$

- Region III: $T-b \leq t < T$, the arrival of a burst at any time $t > T-b$ is incompatible with a zero queue length at time -t. Therefore all the burst arrivals, $i = N$, must be at any time $t \leq T - b$. That means to particularize the formula in region II to $i = N$.

From the above we obtain:

$$A(t,t+x,i) = P\{N(t,0) = t+x | N(u,0) < u+x, t < u \leq T, B_{-t} = i\}$$

$$\begin{cases} \frac{1}{(t)^i} \quad q_t^{(i)}(t+x) & if \quad t < b, \quad 0 \leq i \leq N \\ \sum_{j=0}^{i} \binom{i}{j} \frac{(t-b+1)^j}{t^i} q_{(b-1)}^{(i-j)}(t+x-bj) & if \quad b \leq t < T-b, \quad 0 \leq i \leq N \\ \sum_{j=0}^{N} \binom{N}{j} \frac{(t-b+1)^j}{t^N} q_{(b-1)}^{(N-j)}(t+x-bj) & if \quad T-b \leq t < T, \quad i = N \end{cases} \tag{5.2}$$

A simple expression for $q_m^{(i)}(t)$ is derived in [44] (see also [42]): Let $q_m(t)$ be a discrete-time pulse of amplitude 1 in $[1,m]$ and let $q_m^*(z)$ be the z-transform of such pulse. We are interested in finding a simple expression for the i-th discrete-time convolution of $q_m(t)$. This is equivalent to find a simple expression for the coefficients of the polynomial:

$$(q_m^*(z))^i = (\sum_{n=1}^{m} z^n)^i \tag{5.3}$$

We can express $q_m^*(z)$ as:

$$q_m^*(z) = (1-z^m) q_\infty^*(z) \tag{5.4}$$

We can easily derive (for example, using the convolution algorithm, [31]) that:

$$(q_\infty^*(z))^i = z^i \sum_{n=0}^{\infty} \binom{i+n-1}{n} z^n \tag{5.5}$$

Hence, we obtain ($i > 0$):

$$(q_m^*(z))^i = z^i \sum_{n=0}^{i(m-1)} \sum_{s=0}^{[\frac{n}{m}]} (-1)^s \binom{i+n-sm-1}{n-sm} \binom{i}{s} z^n \tag{5.6}$$

From that we derive a formula for $q_m^{(i)}(t)$ ($i > 0$ and $t = i, \ldots, im$):

$$q_m^{(i)}(t) = \sum_{s=0}^{[\frac{t-i}{m}]} (-1)^s \binom{i}{s} \binom{t-sm-1}{i-1} \tag{5.7}$$

and $q_m^{(i)}(t) = 0$ for other values of t. For $i = 0$ we define

$$q_m^{(0)}(t) = \delta(t) \tag{5.8}$$

where $\delta(y)$ is the discrete-time impulse function.

The term L(t,i):

To calculate $P\{N(u,0) < u+x, t < u \leq T | B_{-t} = i\}$ we make use of similar arguments as those ones to solve the $N \cdot D/D/1$ (see [87]). The event $\{N(u,0) < u+x, t < u \leq T | B_{-t} = i\}$ corresponds to the situation in which an auxiliary queue loaded with periodic arrivals of period T-t in an interval [-(T-t),0) is empty at time 0. The arrivals to the said auxiliary queue are of two classes, see figure 5.3.b:

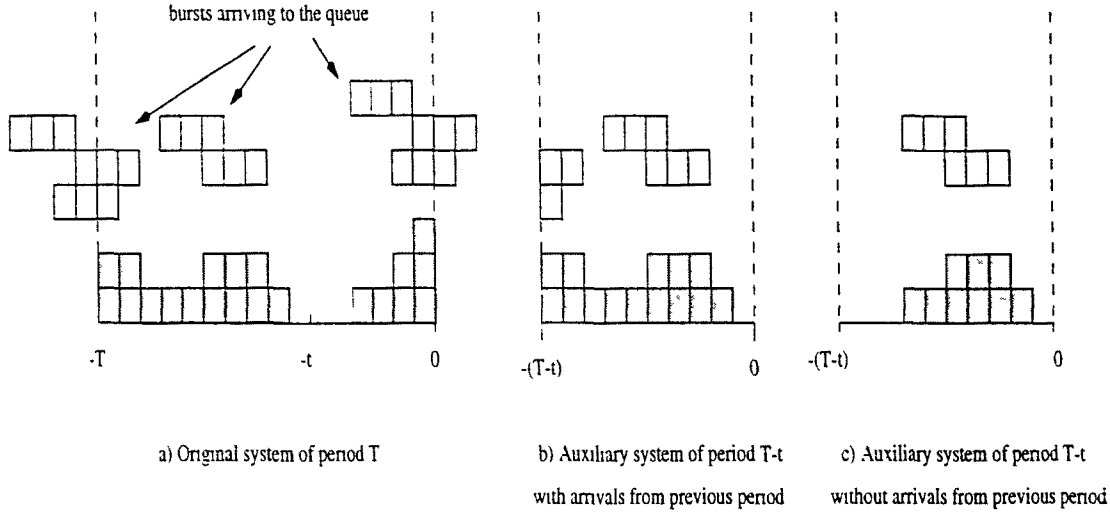


Figure 5.3: Auxiliary system to calculate term L(t.i)

- N-i independent bursts of b cells distributed uniformly over the period [-(T-t),0)
- A batch arrival at time -(T-t) consisting of ib-t-x cells. This batch arrival models the contribution of the bursts which (in the original system) began in the previous period and end partly in the actual period.

However, if the original system is stable (i.e. $Nb < T$), the auxiliary system will be also (i.e. $(N-\imath)b + (ib - t - x) < T - t$). Under these conditions the batch arrival will be served before the end of the period and we do not need to consider the batch arrival to compute the probability that the auxiliary queue is empty at the end of its period.

Therefore in order to calculate $L(t, \imath)$ we need only take into account the N-i independent bursts, see figure 5.3.c:

$$L(t,i) = P\{N(u) < u+x, t < u \leq T | B_{-t} = i\} = (1 - \frac{(N-i)\cdot b}{T-t}) \qquad (5.9$$

The term B(t,i):

Being the bursts uniformly and independently of each other distributed within a period, the probability that i sources become active in the interval $[-t, 0)$ follows a binomial distribution:

$$B(t,i) = P\{B_{-t} = i\} = \binom{N}{i} (\frac{t}{T})^i \cdot (1 - \frac{t}{T})^{N-i} \tag{5.10}$$

5.2.2 Busy periods.

We can apply the study of the idle and busy period distributions of the $N \cdot D/D/1$ system queue, see subsection 4.2.2, to the WCT system queue. Now, the WCT busy period distribution will be B_s cells the CBR busy period distribution: $q_{busy}^{B_s}(l) = q_{busy}^{ND/D/1}(l/B_s)$ taking into account that busy period lengths are multiple of the Burst size B_s.

5.3 The fluid WCT model.

A similar study can be done for a system using a fluid-flow approximation, [44] and [45]. In this case the WCT is defined as a periodic source which produces h information units during a constant time b and which remains silent during a constant time $T - b$. The multiplexer is able to serve at rate c. We assume $h \geq c$. In this case the server is saturated whenever at least one burst is in progress, [26], and equation (4.28) can be expressed as:

$$P\{L_0 > x\} = c \int_0^T P\{\phi(t) = x, D_t = 0\} \cdot P\{L_t = 0 \mid \phi(t) = x, D_t = 0\} dt, \tag{5.11}$$

where $N(t)$ is the work arriving in the interval $[-t, 0)$, and $\phi(t) = N(t) - ct$, and D_t is the arrival rate at time -t. Introducing B_t as the number of sources which become active in the interval $[-t, 0)$, equation (5.11) can be rewritten as:

$$\begin{aligned} &P\{L_0 > x\} = \\ &c \int_0^T \sum_{i=1}^{N} P\{\phi(t) = x, D_t = 0, B_t = i\} P\{L_t = 0 | \phi(t) = x, D_t = 0, B_t = i\} dt \end{aligned} \tag{5.12}$$

The derivation of the final formula in the case of WCT fluid sources follows the same steps as in the discrete-time case.

We also have to distinguish between three regions depending on the values of t. Denoting the density function $f_t(x + ct) = P\{N(t, 0) = x + ct \mid L_t = 0, D_t = 0, B_t = 1\}$ as the work produced by an individual source which becomes active in the interval [-t,0), and $p_y(x)$ as a pulse with amplitude 1 in the interval $[0, y)$. Then, $f_t(x) = \frac{1}{ht}\, p_{ht}(x)$ and:

- Region I: If $t < b$: Probability $P\{N(t, 0) = x + ct \mid D_t = 0. B_t = i\}$ will be the i-th convolution of function $f_t(x + ct)$:

  $$f_t(x + ct)^{(i)} = \frac{1}{(ht)^i}\, p_{ht}^{(i)}(x + ct)$$

  Note that condition D_t implies that all the work generated in the interval [-t,0) comes from the i active sources.

- Region II: If $b \leq t < T-b$: As in the discrete-time case, a source can become active in [-t,-b) or in [-b,0). Therefore:

$$f_t(x+ct) = \left(\frac{1}{ht}\, p_{hb}(x+ct) + \left(\frac{t-b}{t}\right) \delta(x+ct-hb) \right)$$

and its i-th convolution:

$$f_t(x+ct)^{(i)} = \left(\tfrac{1}{ht}\, p_{hb}(x+ct) + \left(\tfrac{t-b}{t}\right) \delta(x+ct-hb) \right)^{(i)} =$$
$$\sum_{j=0}^{i} \binom{i}{j} \left(\tfrac{1}{(ht)^{i-j}} q_{(hb)}^{(i-j)}(x+ct) \right) * \left(\tfrac{(t-b)^j}{t^j} \delta(x+ct-bhj) \right) =$$
$$\sum_{j=0}^{i} \binom{i}{j} \tfrac{1}{h^{i-j}} \tfrac{(t-b)^j}{t^i}\, q_{(hb)}^{(i-j)}(x+ct-bhj)$$

- Region III: $T-b \leq t < T$, the arrival of a burst at any time $t > T-b$ is incompatible with a zero queue length at time -t as in the discrete-time case. Therefore that means to particularize the formula in region II to $i = N$.

The derivation of probability $P\{L_t = 0 | \phi(t) = x, D_t = 0, B_t = i\}$ is similar to the discrete-time case:

$$P\{L_t = 0 | \phi(t) = x, D_t = 0, B_t = i\} = \frac{P\{L_t = 0 | \phi(t) = x, B_t = i\}}{P\{D_t = 0 | \phi(t) = x, B_t = i\}} \tag{5.13}$$

Where the two probabilities of the last term can be calculated as:

$$P\{L_t = 0 | N(t.0) = x + ct, B_t = i\} = (1 - \frac{(N-i) \cdot bh}{T-t})$$

$$P\{D_t = 0 | N(t,0) = x + ct, B_t = i\} = (\frac{T-t-b}{T-t})^{(N-i)}$$

And finally, the term $P\{D_t = 0, B_t = i\}$ is:

$$p\{D_t = 0. B_t = i\} = \binom{N}{i} \frac{t^i (T-b-t)^{N-i}}{T^N} \tag{5.14}$$

We must now derive the i-th convolution of a pulse in continuous time. Let $p_y(x)$ be a pulse with unitary amplitude in $(0, y)$. In order to get an explicit formula for its i-th convolution, we obtain first the Fourier transform of $p_y(x)$:

$$P_y(\omega) = \int_{-\infty}^{+\infty} e^{-j\omega x} p_y(x) dx = j \frac{e^{-j\omega y} - 1}{\omega} \tag{5.15}$$

Hence. the transform of $p_y^{(i)}(x)$ is:

$$P_y^i(\omega) = j^i (\frac{e^{-j\omega y} - 1}{\omega})^i \tag{5.16}$$

and $p_y^{(i)}(x)$ can be expressed as $(i > 0)$:

$$p_y^{(i)}(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{j\omega x} P_y^i(\omega) d\omega = \frac{j^i}{2\pi}\sum_{k=0}^{i}\left(\begin{array}{c} i \\ k \end{array}\right)(-1)^{i-k}\int_{-\infty}^{+\infty}\frac{e^{-j\omega(ky-x)}}{\omega^i}d\omega \tag{5.17}$$

We now consider the integral:

$$\int_{-\infty}^{+\infty}\frac{e^{j\omega(x-ky)}}{\omega^i}d\omega = \int_{-\infty}^{+\infty}\frac{cos(\omega(x-ky))}{\omega^i}d\omega + j\int_{-\infty}^{+\infty}\frac{sin(\omega(x-ky))}{\omega^i}d\omega \tag{5.18}$$

If i is even the imaginary part of the right hand side of equation (20) vanishes. For the real part we have ([73], [76]):

$$\begin{aligned}\int_{-\infty}^{+\infty}\frac{cos(\omega(x-ky))}{\omega^i}d\omega &= \frac{(x-ky)^{i-1}}{(i-1)!}cos(i\frac{\pi}{2})\int_{-\infty}^{+\infty}\frac{sin(\omega(x-ky))}{\omega}d\omega = \\ &= \pi\frac{|x-ky|^{i-1}}{(i-1)!}cos(i\frac{\pi}{2})\end{aligned}$$

If i is odd the real part vanishes while for the imaginary part we have:

$$\begin{aligned}\int_{-\infty}^{+\infty}\frac{sin(\omega(x-ky))}{\omega^i}d\omega &= \frac{(x-ky)^{i-1}}{(i-1)!}cos((i-1)\frac{\pi}{2})\int_{-\infty}^{+\infty}\frac{sin(\omega(x-ky))}{\omega}d\omega = \\ &= \pi\frac{(x-ky)^{i-1}}{(i-1)!}cos((i-1)\frac{\pi}{2})sgn(x-ky)\end{aligned}$$

Finally we obtain the following expression for $p_y^{(i)}(x)$ $(i > 0)$:

$$p_y^{(i)}(x) = \frac{1}{2}\sum_{k=0}^{i}\left(\begin{array}{c} i \\ k \end{array}\right)(-1)^k\frac{(x-ky)^{i-1}sgn(x-ky)}{(i-1)!} \tag{5.19}$$

For $i = 0$ we define:

$$p_y^{(0)}(x) = \delta(x), \tag{5.20}$$

where $\delta(x)$ denotes the Dirac delta in continuous-time.

The resulting integrals can be evaluated numerically, for example, by means of the Gauss method or by Simpson's method.

5.4 Upper bound models to the WCT model.

G. Ramamurthy and R. Dighe, in [78], use a batch arrival model as an upper bound. We can note that since the cells of the same batch arrive in the same slot the queue length will be greater than if a WCT burst arrive. That means that a batch model will upper bound the WCT model. Furthermore, Ramamurthy and R. Dighe propose the following approximation to the waiting time distribution:

Let $L_0^{(B_s)}$ be the queue length at time 0 of a multiplexer loaded with N WCT sources of period $T = B_s \cdot T_{CBR}$. Then:

$$P\{L_0^{(B_s)} > x\} \approx P\{L_0^{(1)} > x/B_s\} \tag{5.21}$$

Note, that they are making use of the observation that the queue length grows roughly by a factor of B_s when the burst size increases.

We can find in [4] the $\sum N_k D_k^{X_k}/D/1$ queue system, where batches of fixed size are considered in a single multiplexer. In this system there are K source classes with cells of class k arriving in constant batches of size b_k and constant interarrival time of $D_k \cdot b_k$. There are N_k sources of class k, and the class load is: $\rho_k = N_k/D_k$

In order to compare with the WCT system, we particularize the system to homogeneous batch sources of size b and period T. See figure 5.4 for a reference between a WCT and a batch periodic source. Applying equation (4.10) and considering $B_t = i$ sources active in the interval [-t,0), we arrive to:

$$\begin{aligned} P\{L_0 > x\} &= \sum_{t=1}^{T} \sum_{i=1}^{N} P\{N(t,0) = t + x | N(u,0) < u + x, t < u \leq T, B_{-t} = i\} \cdot \\ &\cdot P\{N(u,0) < u + x, t < u \leq T | B_{-t} = i\} \cdot P\{B_{-t} = i\} \\ &= \sum_{i=\lceil x/b \rceil}^{N} \binom{N}{i} \left(\tfrac{ib-x}{T}\right)^i \cdot \left(1 - \tfrac{ib-x}{T}\right)^{N-i} \cdot \left(1 - \tfrac{(N-i)b}{T-ib-x}\right) \end{aligned} \tag{5.22}$$

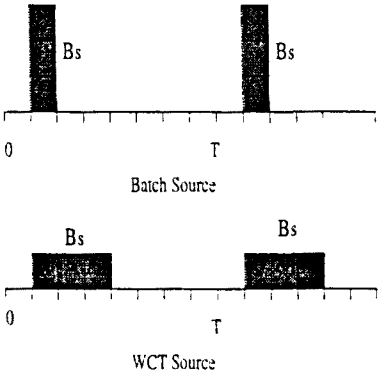


Figure 5.4: Batch and WCT periodic sources.

5.5 Performance results.

The fluid model.

As a first result we show in figure 5.5 how well fits the fluid approximation with the discrete model. We draw a curve with the following parameters: a CBR period $T_{CBR} = 10$, number of sources $N = 8$ (that means a load of $\rho = N/T_{CBR} = 0.80$), and burst sizes of B_s=10,5 and 2. As It can be observed, the fluid model fits quite well the discrete one. We must take into account that to calculate the integral of the fluid model we have used numerical methods (e.g. the Gaussian quadrature or the Simpson method).
I have chosen the discrete time model since it is solved exactly.

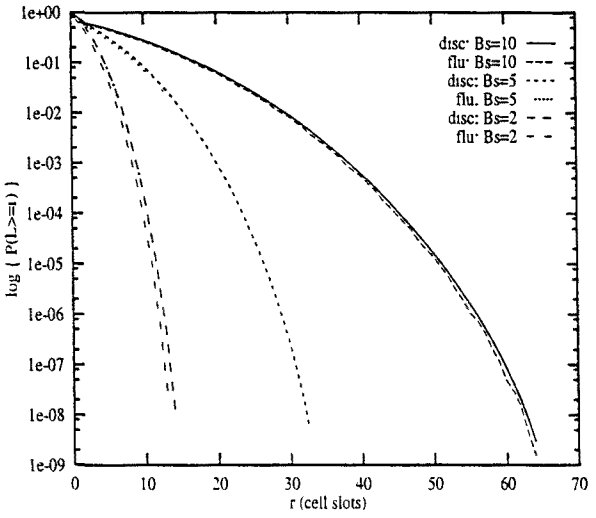


Figure 5.5: Comparison between the fluid and discrete time models.

WCT sources performance in a multiplexer.

We are interested in evaluate the queue length distribution of a multiplexer loaded with WCT sources as a function of the several parameters involved in the system: the period of the sources and the burst size, and we are also interested in figures as the admissible load, see [44].

We consider the multiplexing of homogeneous ATM connections which declare in the Traffic Contract a pair Peak Emission Interval and Cell Delay Variation Tolerance, (T,τ). Under these circumstances, as it was explained in subsection 2.3, the UPC mechanism might allow bursts of size $B_s \leq 1 + \lfloor \frac{\tau}{T-1} \rfloor$, where the PEI and CDVT are expressed in ATM cell slots.

Figure 5.6 shows the complementary queue length distribution when 12 sources with a T of 15 cell slots are multiplexed and transmit bursts of size 1, 2, 5 and 10 cells. These sources correspond to 10 Mb/s CBR connections multiplexed in a 150 Mb/s link. In figure 5.7, we vary the number of sources and T while maintaining constant B_s. The queue length for batches of B_s cells Poisson distributed with the same load ($\rho = 0.8$) is

also drawn.

In table 5.1, we show the buffer size needed to accommodate the multiplexing of WCT sources with different periods (T) for several burst sizes (B_s). This buffer size is calculated as a quantile of the distribution of the queue length: $X_{max} = sup\{x : P\{L_0 \leq x\} \leq 1-\epsilon\}$, where we have taken $\epsilon = 10^{-8}$.

$rho = 0.8$	B_s=1	B_s=2	B_s=5	B_s=10
T=15	9	17	42	82
T=30	13	25	60	120
T=60	17	33	82	163

Table 5.1: Buffer size for a quantile of the queue length distribution with $\epsilon = 10^{-8}$ for several periods T and bursts sizes (B_s).

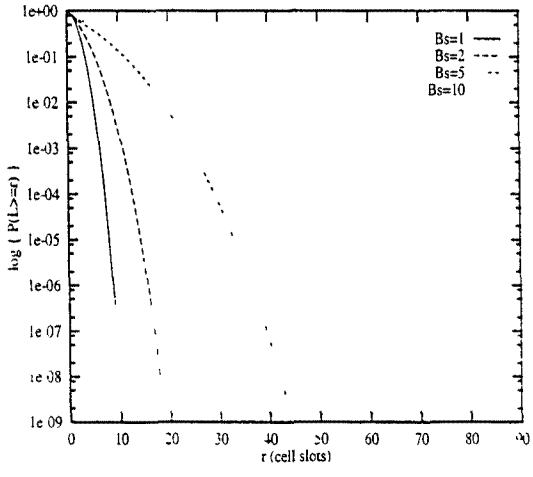


Figure 5.6: Queue Length CPDF for N=12 connections, T=15 cell slots.

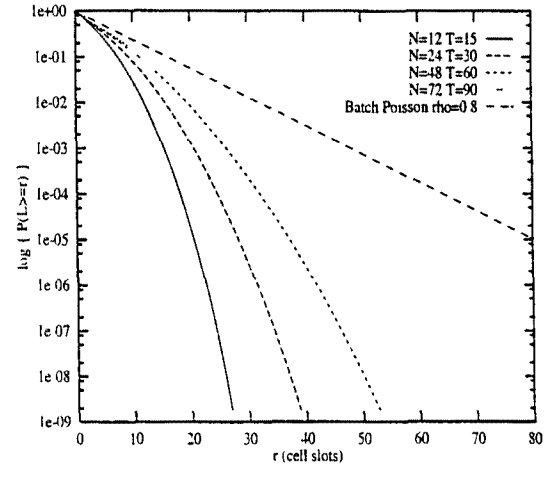


Figure 5.7: Queue Length CPDF for WCT connections (B_s=3), batch Poisson (B_s=3).

From figure 5.6 and table 5.1 we observe that increasing the burst size, the queue length increments considerably. From a buffer size of 9 cells for sources with T=15 and B_s=1, we pass to a buffer size of 82 cells for sources with T=15 and B_s=10. We can also conclude that the queue length increases by a factor approximately equal to B_s to maintain the same queue length probability than an $ND/D/1$ queue system.

In figure 5.7 we can observe that multiplexing low bit rate sources emitting bursts of size $B_s = 3$ we can use a Poisson process to bound the queue length probability. This fact is studied and commented in the following chapter.

Models that upper bounds the WCT multiplexing model.

We study these approximations in figures 5.8, 5.9 and 5.10. We first compare with the batch model and secondly we will see how well Ramamurthy and Dighe's bound fits the

$rho = 0.8$ T=15	B_s=2	B_s=5	B_s=10	B_s=15
wct	17	42	82	123
batch	18	45	90	134

$rho = 0.8$ B_s=5	T=15	T=30	T=60
wct	42	60	82
batch	45	64	86

Table 5.2: Buffer size for a quantile of the queue length distribution with $\epsilon = 10^{-8}$

WCT model.

In figure 5.8, we compare a multiplexer loaded with periodic WCT sources sending bursts of B_s back-to-back cells with a system sending periodic batches of size B_s. The traffic parameters in both models are the same: sources with a declared T of 15, and the burst size B_s is varied. In figure 5.9, we plot the same curves but fixing the burst size to B_s=5 and varying the periods declared by the sources. We can observe that the batch approximation does not depend on the value T. However, the higher the burst size is the worst behaves the approximation. To get a better understanding of these curves, we have written a table (see table 5.2) for the quantile of the queue length for $\epsilon = 10^{-8}$.

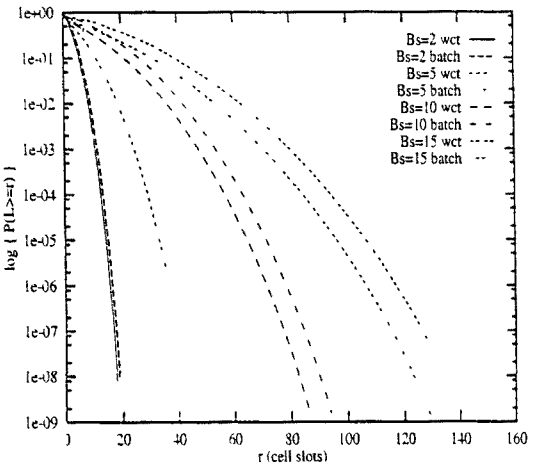


Figure 5.8: Queue length CPDF for several B_s, given a T=15, for WCT sources and batch sources.

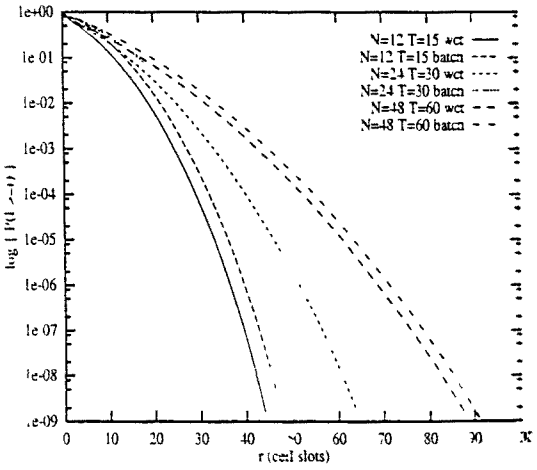


Figure 5.9: Queue length CPDF for several periods, given a B_s=5, for WCT sources and batch sources.

We apply in figure 5.10 the approximation proposed by Ramamurthy and Dighe for WCT sources. We label as "wct" that curves calculated with the WCT model, and as "appox" those ones using equation (5.21). We observe the same behavior as with the batch bound. The bound does not depend on the period, but strongly on the burst size. The higher

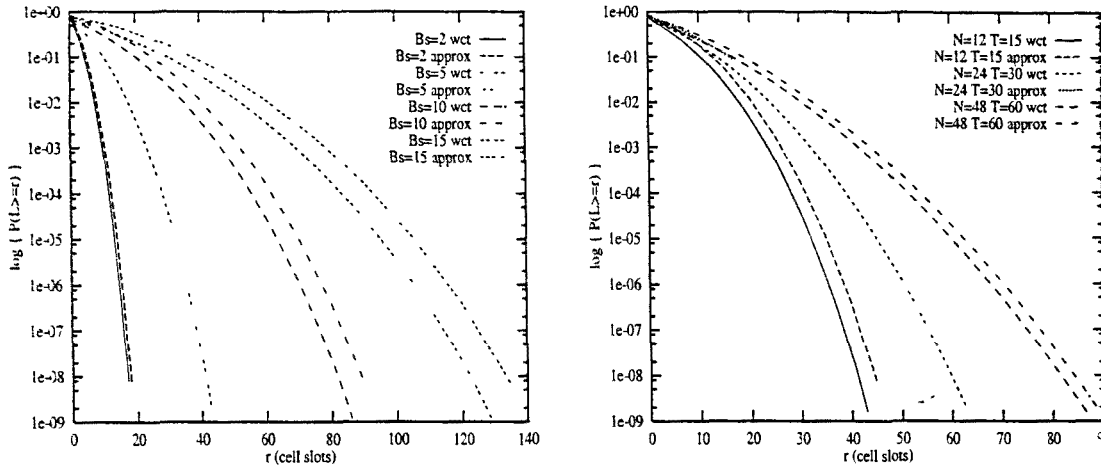


Figure 5.10: Approximating WCT system by a bound.

the B_s the worst is the bound.

Admissible Load in the multiplexer.

In figure 5.11, the admissible load as a function of the B_s is plotted. We can observe that to keep the same QoS given a buffer size, the admissible load decreases dramatically. As QoS has been chosen the value 10^{-9}. The buffer size is the queue length that corresponds to the quantile for a value of 10^{-9} when we multiplex 12 CBR sources of 10 Mb/s on a link of 150 Mb/s: 10 cells.

Thus. for CBR sources, the admissible load is 0.8. With the same queue length and a burst of 2 cells. we have to drop the admissible load in the multiplexer to 0.4, for a burst of 3 to 0.33 and for a burst of 4 to 0.2. That means that with the same buffer size to keep the same QoS, we can only admit 3 connections if they send bursts of size 4 cells instead of the 12 connections that we could admit if they send 1 cell periodically. This is an important effect, since although increasing the buffer size to absorve WCT may seem a solution, time constrains for real-time DBR sources cannot be guaranteed with large buffers.

WCT produced by AAL buffers.

Now we study an example in which the user produces WCT. Let us assume that we scan each 480 time slots the AAL buffers of a multimedia workstation with different CBR connections. The user generates traffic at a total rate of 10 Mbps. and the physical link rate is 150 Mbps. This means that clumps of 32 back-to-back cells will enter the ATM network. Figure 5.12 shows the complementary buffer queue length distribution of a multiplexer handling such traffic sources (load 0.80). To have a quantile of the buffer length probability lower than 1×10^{-10} we need a buffer value of 288 cells. Using a buffer

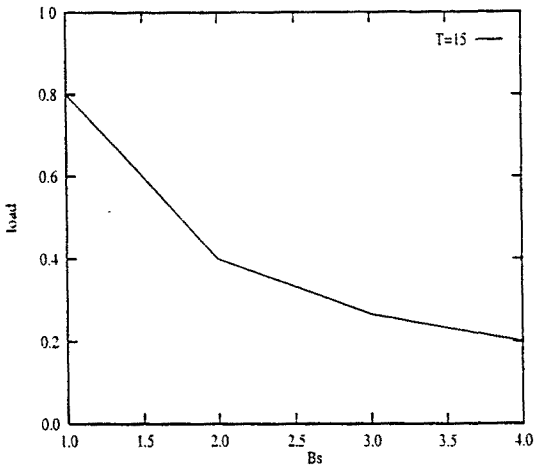


Figure 5.11: Admissible load for connections with T=15

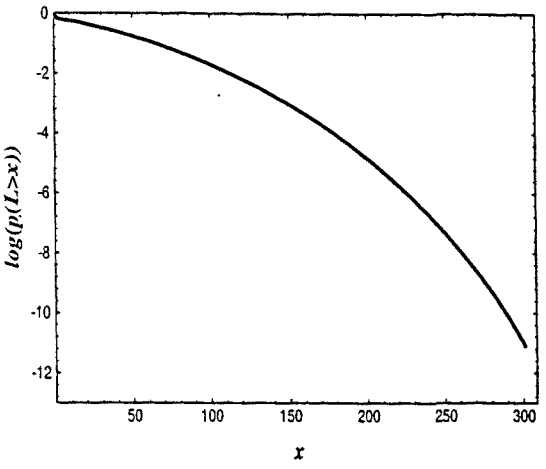


Figure 5.12: Scanning AAL buffers of a multimedia workstation.

of 128 means to reduce the load to 0.33 while for a buffer of 64 cells the admissible load goes down to 0.13.

WCT impact on resource management in an ATM network: the EXPLOIT Testbed.

As a last example, we come back to the experiments performed in the EXPLOIT Testbed, and analyze the impact on resource management of the CDV measured. Using a Worst Case Traffic situation, we can compute a bound on the size of the maximum burst size (Bs) of back-to-back cells that a multiplexer can absorb for a given load, queue length (QL) and a Cell Lost Ratio (CLR) target. In figure 5.13, see also [58], the acceptable CDV tolerance has been computed as a function of the PCR for a CLR target of 10^{-6}, a load of 90 % and queue lengths of 70 and 100. We have also plotted the CDV computed by the Erlang-7 and Erlang-12 formulas and the measured CDV for m=7 multiplexing stages.

The interpretation of these curves is the following: any point that is below the curves QL (either with 70 or 100 cell places) can be absorbed by the correspondent buffer independently of the pair (PCR,τ). As can be seen, a queue length of 70 can absorb clumps of cells equivalent to less than 270 μsec for any PCR while a queue length of 100 cells can absorb the equivalent 400 μsec for any PCR. This situation, as we will extend in chapter 8, corresponds with the idea of negligible CDV: we have enough resources to allocate connections following the $M/D/1$ queue system. If we would like to give a CDV-T of 400 μs, we can observe that PCRs higher than 25 Kcell/s could not be allocated for a buffer of 70 places. It would be necessary to have a buffer of 100 places to allocate any PCR with CDV-T of 400 μs. Finally, for a CDV-T of 1 ms, any PCR lower than 5 Kcell/s would be above the acceptance curve of 70 buffer places.

At UNI (User Network Interface) the CDV tolerance allocated must be below the WCT curve in order to keep the QoS required by the source. At least what we can assure is

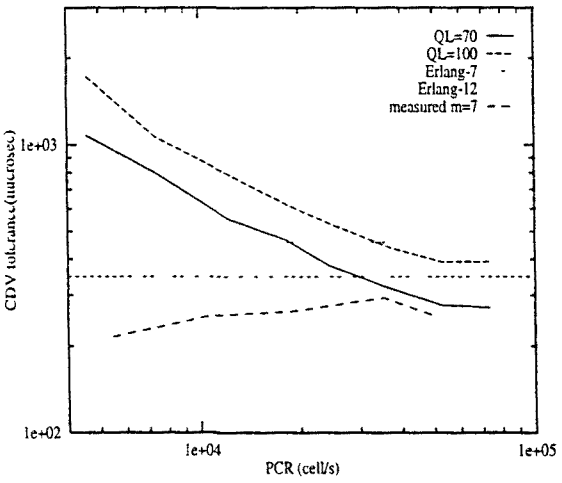


Figure 5.13: Acceptable CDV tolerance as a function of the PCR.

that if the CDV is below the WCT curve a simple PCR allocation can be performed (as if we consider negligible CDV). If the CDV introduced in the access network is above the curve, then traffic shaping or PCR over-allocation has to be used.

At NNI (Network to Node Interface) and assuming that the traffic has been policed at UNI and at the entrance of the network a CDV tolerance below the acceptance WCT curve corresponding to the buffer resources was allocated, we can observe that after m=7 multiplexing stages we have accumulated a maximum CDV of 360 μs for a PCR of 21.180 cell/s and around 250-300 μs for the rest of the PCRs. That means that for almost all the PCR we are below the acceptance curve for a buffer of QL=70. The case of PCR=21.180 cell/s corresponds to a case where bursty traffic was multiplexed with that connection.

As a general conclusion taking into account the configuration and set up of our experiments we can see that:

- experimentally high peak rate connections accumulate the same order of magnitude of CDV than low bit rate connections.
- the acceptance curve given a queue length for low bit rate connections accepts high CDVs (e.g. up to 1 ms for QL=70 and PCR $<$ 5 Kcell/s)
- while acceptance curve given a queue length for high bit rate connections accepts low CDVs (e.g. up to 260 μs for QL=70 and PCR $>$ 80 Kcell/s)

Therefore. high peak rate connections can be a problem since the WCT acceptance curve given a queue length can be below the CDV allocated to that connections. For low peak rate connections the accumulated CDV is always below the curve. This can give us the idea that not for all connections traffic shaping must be performed. High PCR connections are the ones that can be above the WCT curve and then must be shaped. However the low PCR connections do not accumulate enough CDV to be above the WCT

sources. That means that they are not so dangerous as the high PCR ones. We will come back on this idea in chapter 8, where we analytically study the trade-offs of CAC and traffic shaping.

Limitations of the WCT model.

Finally, we must discuss the limitation of this model. As a first limitation, I would mention the hypothesis of homogeneity. We have considered all the sources as the same type (WCT). We would consider interesting to analyze other models in which not all the traffic is WCT or in which not all the WCT are of the same characteristics.

A second limitation of the model resides on the computational parameters. We would like to know for which parameters the model becomes difficult to compute. Since the model depends on too many parameters to calculate the order of computation, we have draw a curve of the CPU time spent in calculating $P\{L_0 > x\}$ as a function of the parameters burst size, b, period, T and number of sources, N. I have chosen x=10. We have found that the most limiting parameters are the burst size and the number of sources, see figures 5.14 and 5.15. For different values of x, the CPU time remains almost constant. and for an increasing value of T the CPU time increases slightly, but not so abruptly as with the number of sources or the burst size. The computer utilized is a Silicon Graphics with OS IRIX Release 6.3 IP32. Thus, We would consider interesting to analyze other models in which we can multiplex a high number of sources.

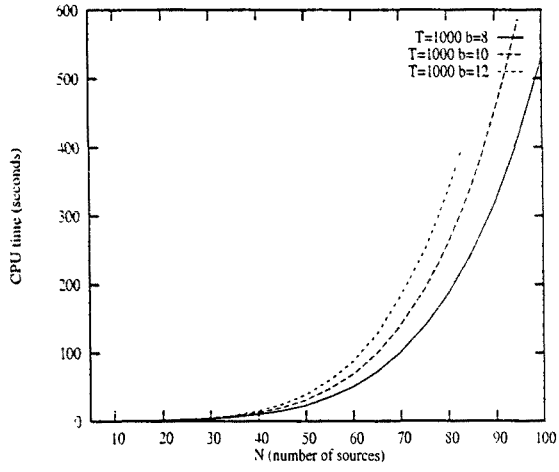


Figure 5.14: CPU time as a function of the number of sources.

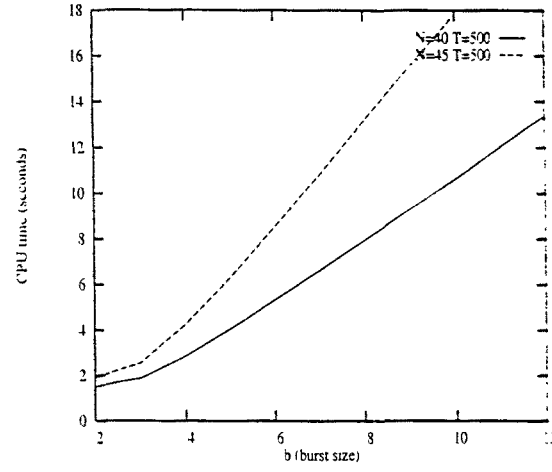


Figure 5.15: CPU time as a function of the burst size.

5.6 Conclusions and comments.

We have analyzed a multiplexor fed by periodic WCT sources, and we have studied the performance and impact of these WCT sources in terms of buffer size. We have considered the discrete-time case and a fluid approximation that fits the discrete-time model very well when the burst size is high. We have also compared the discrete-time model with bounds proposed by some authors. As we can observe in the different examples.

worst case traffic sources decrease the admissible load in the multiplexor given a defined buffer size and a QoS parameter. In other words, we should increase the buffer size to cope with bursts of back-to-back cells. In fact, the queue length increases by a factor approximately equal to the burst size respect to the $ND/D/1$ queue system to keep the QoS in terms of a quantile of the queue length distribution.

However, increasing the buffer size is not a good solution if we multiplex real time connections with hard delay constrains. For instance, we must think in how to dimension the receiver play black buffers. The size of these buffers depends on the end-to-end delay (see [4] on how a play-back buffer can be dimensioned). If we increase the switching buffers across the network the end-to-end delay can grow considerably.

We could discuss if it is realistic a scenario with all sources transmitting WCT. My opinion is that this kind of scenario would be really pessimistic and that it is very improbable that all users misbehave at the same time. However in the other hand, the network as it is mentioned in the ATM FORUM and ITU-T 371. [3] and [1], should possible take into account worst case traffic passing through the UPC to avoid impairments with other connections.

Therefore, it seems clear that some techniques should be implemented to avoid the impact of WCT in the network. Chapter 8 is dedicated to some of these techniques, but we have already deduced in the example dedicated to resource management in the EXPLOIT Testbed that not all the sources are so dangerous to the network. Protecting the network from high bit rate connections and dimension the buffers taking into account WCT for low bit rate connections could be enough to guarantee QoS parameters. We will come back to this idea in chapter 8.

The contributions to the thesis in this chapter are: analytical models in discrete-time and fluid flow approximation to study the superposition of periodic homogeneous WCT sources, and a set of examples. The upper bounds used (the one based on the ND/D/1 queue system and the batch bound) were taken from the literature.

Chapter 6

Other WCT models for DBR connections.

As an extension of the former chapter, we present here a set of models that describe the multiplexing of WCT sources with several distributions. The objective of this chapter is to find some models able to overcome the limitations of the periodic WCT model. These limitations are on one hand the assumption of homogeneity and on the other hand the high CPU computation times when the number of sources or the burst size grows.

As it was mentioned in chapter 3, Poisson traffic is used as an upper bound on the superposition of deterministic sources when non of them dominates (e.g. large number of periodic sources). The $M/D/1$ does not take into account the correlation effects at cell level. The bound overestimates the congestion effects (the queue length distribution) for heavy loaded systems and it improves if the load is low.

For WCT sources. we can also find a bound considering that the bursts follow a Poisson arrival process. In this case, exact solutions to the queue length distributions can be calculated using again the Benes approach to the Virtual Waiting Time. With this assumption we want to find a model that overcomes the computational limitation when the number of sources is high.

On the other hand. we also are interested in knowing how is the queue distribution if not all the traffic is WCT. For instance. the input traffic are sources whose B_s is lower than the maximum burst size (they behave as CBR sources), multiplexed with sources whose B_s is equal to the maximum burst size (they behave as WCT sources). With this assumption we want to overcome the limitation of only having WCT sources.

If we mix both assumptions we can find bounds for the multiplexing of bursts of cells and CBR sources when the number of sources grows.

For that purpose. we have calculated a set of multiplexing models that contemplate several scenarios and their range of applicability. see [19]:

- Multiplexing WCT Poisson distributed: bursts of size B_s arrive following a Poisson process of parameter λ. The solution is exact but numerically difficult to compute

for high loads due to infinity summatories. The model can be found in section 6.1.

- Multiplexing WCT sources with CBR sources uniformly distributed: WCT sources of period T_{wct} uniformly distributed are multiplexed with CBR sources of period T_{cbr}. For this case the solution is approximated. The model can be found in section 6.2.1.
- Multiplexing WCT sources with burst size equal to B_s uniformly distributed over a period T_{wct} with Poisson background traffic with burst size 1 cell. The solution is approximated and better for low Poisson loads and high WCT periods. The model can be found in section 6.2.2.
- Multiplexing bursts of B_s cells Poisson distributed with Poisson traffic. The solution is exact. The bursts can be modeled as WCT bursts or as batches. In the case of having batches a polynomial solution can be found. The models are described in section 6.3.1 and 6.3.2.

6.1 WCT models Poisson distributed

We study a single server system where bursts of b back-to-back cells arrive following a Poisson process of parameter λ and load ρ.

Introducing B_t, equation (4.8) can be written as:

$$P\{L_0 > x\} = \sum_{t=1}^{\infty}\sum_{i=0}^{\infty} P\{\phi(t) = x, B_t = i\} P\{L_t = 0 \mid \phi(t) = x, B_t = i\} \tag{6.1}$$

The bursts follow a Poisson distribution: $P\{B_t = i\} = \frac{(\lambda t)^i}{i!} e^{-\lambda t}$. Since the system is memoryless $P\{L_t = 0 \mid \phi(t) = x, B_t = i\} = 1 - \rho$.

To derive expressions for probability $P\{\phi(t) = x \mid B_t = i\}$, we distinguish between two cases depending on the values of t: for $t < b$ we have a similar situation with region I of the periodic model, for $t \geq b$, it behaves as region II. Following these indications:

$$P\{L_0 > x\} = (1-\rho)\sum_{t=1}^{b-1}\sum_{i=1}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda t}\, q_t^i(t+x) +$$
$$+ (1-\rho)\sum_{t=b}^{\infty}\sum_{i=1}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda t} \sum_{j=0}^{i} \binom{i}{j} (t-b+1)^j\, q_{b-1}^{(i-j)}(t+x-bj) \tag{6.2}$$

It is known, see [4] or [87], that the $M/D/1$ queue system is a good approximation to the multiplexing of CBR sources if non of them dominates and the load is not too high.

This model is exact and fast to calculate for low Poisson loads (~ 0.7), however it is difficult to compute for high loads (e.g. high factorial values and high powers). Remember that the $M/D/1$ system solved with the Benes approach presents loss of numerical

accuracy. Virtamo, in [91], solves this problem through a polynomial representation. see subsection 4.2.1. The same problem appears in the WCT Poisson distributed system. Summatories in infinite have to be cut out at some point conducting to numerical loss. This is accentuated for high loads and high burst sizes as can be seen in figures 6.1 and 6.2, where the WCT model is compared with simulations for loads of 0.6 and 0.8 respectively. The simulations have been performed with a confidence interval of 95%. We can observe that for high loads the models suffers of a lack of accuracy. For loads around $\sim$ 0.7 the analytical model behaves well.

Since Poisson models provide good approximation when the periods are large, we can assume that the CDV tolerance declared will not be high enough to allow pass large bursts. Therefore, a batch model can also give a good estimate of the Virtual Waiting time distribution as we saw in chapter 4. We have to mention the difficulty in applying Jensen's theorem to equation (6.2) to apply later a polynomial representation. However, this theorem can be applied to a batch model. Note that the batch model gives an upper bound on the queue length distribution.

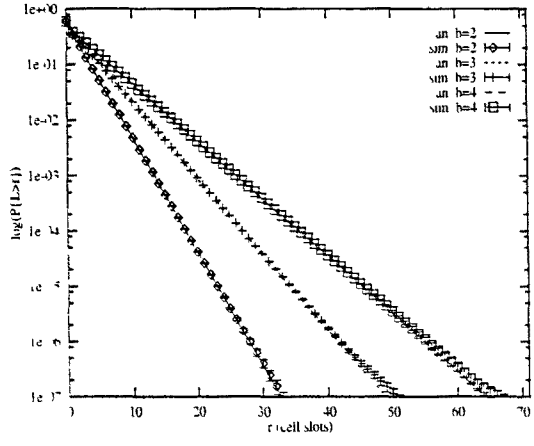


Figure 6.1: Comparison between WCT analytical and WCT simulation: $\rho = 0.6$

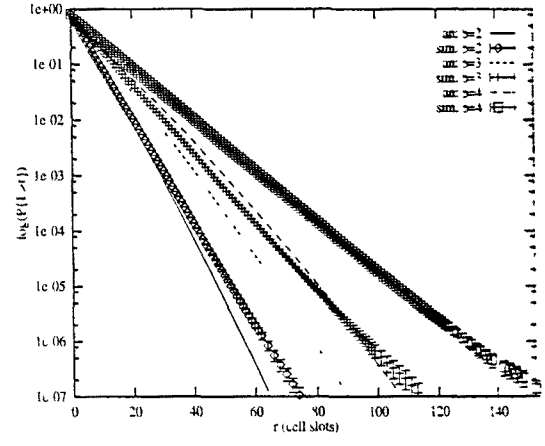


Figure 6.2: Comparison between WCT analytical and WCT simulation: $\rho = 0.8$

6.2 Multiplexing WCT sources with CBR traffic or Poisson Traffic.

6.2.1 Periodic WCT traffic with periodic CBR traffic.

In this case, we multiplex the N_1 periodic WCT sources with period T_{wct} and burst size b and N_2 CBR sources with period T_{cbr}. For simplicity, we assume that the WCT period is a multiple of the CBR period. We define B_t as the number of bursts which arrive in the interval $[-t, 0)$, and K_t as the number of CBR cells which arrive in the interval $-t, \ldots, -(\lceil \frac{t}{T_{cbr}} \rceil - 1) \cdot T_{cbr}$. Introducing B_t and K_t, $P\{L_0 > x\}$ can be expressed as:

$$P\{L_0 > x\} = \sum_{t=1}^{T_{wct}} \sum_{i=0}^{N_1} \sum_{m=0}^{N_2} P\{\phi(t) = x, B_t = i, K_t = m\}$$
$$P\{L_0 = 0 | \phi(t) = x, B_t = i, K_t = m\} \quad (6.3)$$

At time t, a CBR source will emit $\lfloor \frac{t}{T_{cbr}} \rfloor$ cells with probability 1 and one cell with probability:

$$\alpha_t = \frac{t}{T_{cbr}} - \lfloor \frac{t}{T_{cbr}} \rfloor \quad (6.4)$$

In general, we are unable to derive an exact solution for the conditional probability $P\{L_0 = 0 \mid \phi(t) = x, B_t = i, K_t = m\}$, but we can determine bounds. We can derive an approximation for this probability considering that: $(\lfloor \frac{T_{wct}}{T_{cbr}} \rfloor - \lfloor \frac{t}{T_{cbr}} \rfloor) \cdot N_2 - m$ CBR cells and $N_1 - i$ bursts uniformly distributed over an interval of length $(T_{wct} - t)$ have to arrive. This is not true, since in fact, we know that in each T_{cbr} period exactly N_2 cells will arrive, and $T_{wct} - t$ can be longer than one T_{cbr} period. The event $\{L_0 = 0 \mid \phi(t) = x, B_t = i, K_t = m\}$ corresponds to the arrival patterns that would result in an auxiliary queue, loaded with those periodic arrivals over the period $T_{wct} - t$, and being empty at time $-t$. Therefore:

$$P\{L_0 = 0 \mid \phi(t) = x, B_t = i, K_t = m\} = (1 - \frac{(N_1 - i)b + (\lfloor \frac{T_{wct}}{T_{cbr}} \rfloor - \lfloor \frac{t}{T_{cbr}} \rfloor) \cdot N_2 - m}{T_{wct} - t}) \quad (6.5)$$

As before we have three cases to derive the probabilities of equation (6.3):

$$P\{\phi(t) = x, B_t = i, K_t = m\} =$$

$$\begin{cases} \binom{N_1}{i} \binom{N_2}{m} \frac{(T_{wct}-t)^{N_1-i}}{T_{wct}^{N_1}} \alpha_t^m (1-\alpha_t)^{N_2-m} q_t^{(i)}(t + x - N_2 \lfloor \frac{t}{T_{cbr}} \rfloor - m) \\ \qquad for \quad 1 \le t < b, \quad i = 0..N_1, m = 0..N_2 \\ \binom{N_1}{i} \binom{N_2}{m} \frac{(T_{wct}-t)^{N_1-i}}{T_{wct}^{N_1}} \alpha_t^m (1-\alpha_t)^{N_2-m} \sum_{j=0}^{i} \binom{i}{j} (t-b+1)^j \\ \qquad q_{b-1}^{(i-j)}(t + x - bj - N_2 \cdot \lfloor \frac{t}{T_{cbr}} \rfloor - m) \\ \qquad for\, b \le t < T_{wct} - b, \; i = 0..N_1, m = 0..N_2 \\ \binom{N_2}{m} \frac{\alpha_t^m (1-\alpha_t)^{N_2-m}}{T_{wct}^{N_1}} \sum_{j=0}^{N_1} \binom{N_1}{j} (t-b+1)^j \\ \qquad q_{b-1}^{(N_1-j)}(t + x - bj - N_2 \lfloor \frac{t}{T_{cbr}} \rfloor - m) \\ \qquad for \quad T_{wct} - b \le t < T_{wct}, \quad i = N_1, m = 0..N_2 \\ 0 \qquad for \quad T_{wct} - b \le t < T_{wct}, \quad i < N_1, \quad m = 0..N_2 \end{cases} \quad (6.6)$$

We first compare how well fits the analytical model varying the burst size of the WCT sources with a simulation. The simulation has a confidence interval of 95%. We fix the

CBR load to $\rho_{cbr} = 0.4$ and T_{cbr}=10. The WCT sources has a period of $T_{wct} = B_s * 15$ and a load of $\rho_{wct} = 0.4$. We can observe in figure 6.3 that the model behaves well even for low burst sizes.

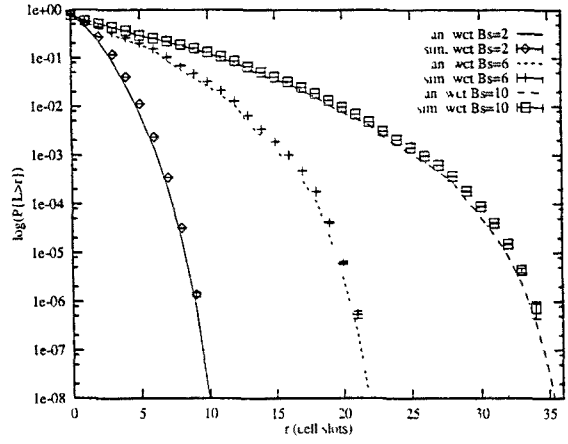


Figure 6.3: Comparison between analytical and simulation: $\rho = 0.8$, $T_{wct} = B_s * 15$, T_{cbr}=10.

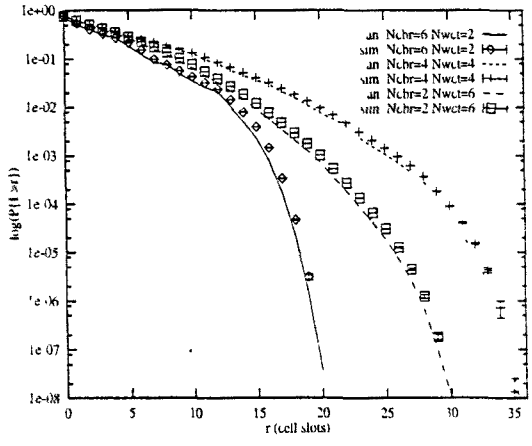


Figure 6.4: Comparison between analytical and simulation: $\rho = 0.8$, T_{wct}=100, B_s=10, T_{cbr}=10. ρ_{wct} and ρ_{cbr} varies.

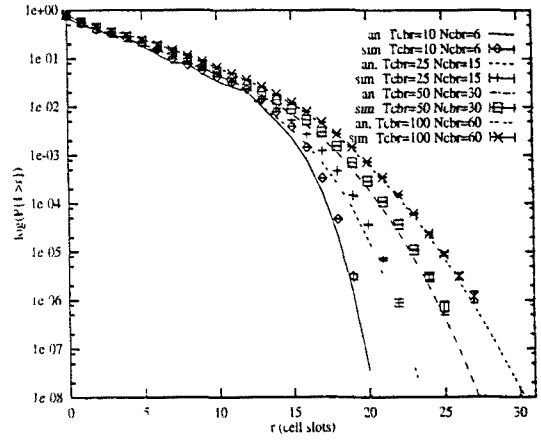


Figure 6.5: Comparison between analytical and simulation: ρ_{wct} and ρ_{cbr} fixed, T_{wct}=100, B_s=10 and N_{wct}=2, T_{cbr} varies.

In figure 6.4, the total load is 0.8, and we vary the "wct" and "cbr" load, keeping fixed the periods. The T_{wct} is 100 and the burst size is B_s=10. while the T_{cbr} is 10. We can see that the analytical results fit quite well the simulations. Observe that T_{wct} is multiple of T_{cbr}. In figure 6.5, we can see how fits the analytical results in simulations varying the T_{cbr} and keeping T_{wct}=100, N_{wct}=2 and B_s=10. The total load is maintained to $\rho = 0.8$.

We have performed similar comparison between the analytical method and simulations for different scenarios. The model behaves well if the T_{wct} is a multiple of T_{cbr}. If this is not the case, we would have to consider the last "cbr" period overlapping a second "wct" period. This would increase the degree of the approximations done, decreasing the accu-

racy of the approach. We have to mention that if the quotient $\frac{T_{wct}}{T_{cbr}}$ is high the accuracy of the model also decreases due to the approximation of having $(\lfloor \frac{T_{wct}}{T_{cbr}} \rfloor - \lfloor \frac{t}{T_{cbr}} \rfloor) \cdot N_2 - m$ CBR cells uniformly distributed over the interval $[T_{wct}, t)$.

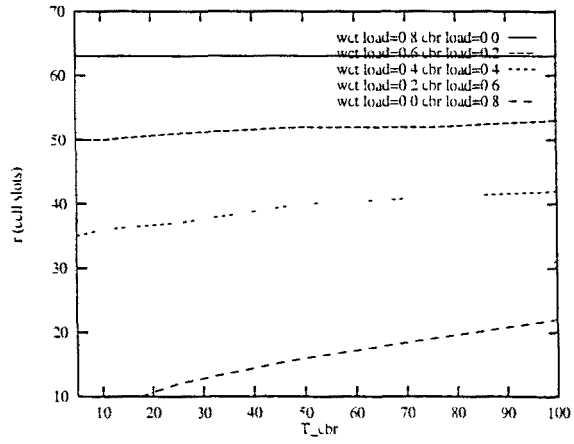


Figure 6.6: Buffer size as a function of the T_{cbr}. T_{wct}=100 and B_s=10.

Finally, we have plotted the buffer size as a function of the "cbr" period. T_{cbr}, for a quantile of the queue length distribution of $\epsilon = 10^{-9}$, where we have "wct" sources of period T_{wct}=100 and B_s=10. The total load has been fixed to $\rho = 0.8$. while the ρ_{wct} and ρ_{cbr} varies. We can see as the WCT is the most demanding in allocating buffer resources. while the "cbr" is the less demanding. In the middle lie the rest of the cases.

6.2.2 Periodic WCT traffic with Poisson traffic.

We multiplex periodic WCT sources with Poisson background traffic into a single server system with infinite capacity. The N WCT sources behave as in figure 5.1. a periodic on/off source with "on" period of b cells and silence of $T_{wct} - b$. The WCT load is $\rho_{wct} = B_s \cdot N/T_{wct}$. The background traffic is a Poisson cell stream of parameter λ_p. The total load is $\rho = \rho_{wct} + \rho_p$.

We will approximate the model considering the Poisson traffic that arrives in a period of T_{wct} slots. This approximation will only be valid for those cases when the Poisson load is low and the period is high. The approximation consists of conditioning in the number of Poisson arrivals in one period. If the Poisson load is high and the period is small. the number of Poisson arrivals in the period can be significative enough to loss accuracy in the model. We must take care that the probability of having $n \geq T_{wct} + x - Nb$ Poisson arrivals in an interval of length T_{wct} is less than an ϵ to use this model, being ϵ a small value.

$$P\{L_0 > x\} = \sum_{n=0}^{\infty} e^{-\lambda_p T_{wct}} \frac{(\lambda_p T_{wct})^n}{n!} P\{L_0 > x | K_{T_{wct}} = n\} \qquad (6.7)$$

To calculate $P\{L_0 > x | K_{T_{wct}} = n\}$, we define again B_t as the number of bursts which arrive in the interval [-t.0) and K_t the number of Poisson cells which arrive in the interval

[-t,0).

$$P\{L_0 > x | K_{T_{wct}} = n\} =$$

$$\begin{cases} \sum_{t=1}^{T_{wct}} \sum_{i=0}^{N} \sum_{m=0}^{n} P\{\phi(t) = x, B_t = i, K_t = m\} \\ \qquad P\{L_t = 0 | \phi(t) = x, B_t = i, K_t = m\} \\ \qquad\qquad if \;\; Nb + n < T_{wct} + x \\ 1 \qquad\qquad Otherwise \end{cases} \tag{6.8}$$

Now we can solve these probabilities as:
$P\{\phi(t) = x, B_t = i, K_t = m\} \cdot P\{L_t = 0 | \phi(t) = x, B_t = i, K_t = m\} \;=$

$$\begin{cases} \binom{N}{i}\binom{n}{m} \frac{t^m (T_{wct}-t)^{N+n-i-m}}{T_{wct}^{N+n}} \left(1 - \frac{(N-i)b+(n-m)}{T_{wct}-t}\right) q_t^{(i)}(t+x-m) \\ \qquad for \quad 1 \le t < b, \quad i = 0..N. \quad m = 0..n \\ \binom{N}{i}\binom{n}{m} \frac{t^m (T_{wct}-t)^{N+n-i-m}}{T_{wct}^{N+n}} \left(1 - \frac{(N-i)b+(n-m)}{T_{wct}-t}\right) \sum_{j=0}^{i} \binom{i}{j} (t-b+1)^j \\ \quad q_{b-1}^{(i-j)}(t+x-bj-m) \quad for \;\; b \le t < T_{wct} - b, \; i = 0..N. \; m = 0..n \\ \binom{n}{m} \frac{t^m (T_{wct}-t)^{n-m}(1-\frac{n-m}{T_{wct}-t})}{T_{wct}^{N+n}} \sum_{j=0}^{N} \binom{N}{j} (t-b+1)^j q_{b-1}^{(N-j)}(t+x-bj-m) \\ \qquad for \quad T_{wct} - b \le t < T_{wct}, \quad i = N, \quad m = 0..n \\ 0 \qquad\qquad for \quad T_{wct} - b \le t < T_{wct}, \quad i < N. \quad m = 0..n \end{cases} \tag{6.9}$$

We compare again in figure 6.7 and 6.8 the analytical model with simulations. In figure 6.7 we vary the burst size and WCT period to see how these parameters affect the results. We can observe that for low periods the model is inaccurate (e.g. $T_{wct} = 20$ or $T_{wct} = 40$). When the period grows the model improves its accuracy. In figure 6.8 we fix the period to $T_{wct} = 100$ and the burst size to $B_s = 10$. The number of WCT sources and the Poisson load varies but keeping a total load of $\rho = 0.8$. We see that if the period is high even for relative high Poisson loads (~ 0.6) the model offers a good accuracy. For higher Poisson loads we could use other models as a Poisson burst model.

6.3 Multiplexing WCT or Batch sources Poisson distributed with Poisson Traffic.

6.3.1 Multiplexing WCT sources Poisson distributed with Poisson Traffic.

We proceed as above to evaluate the queue length distribution. In this case we multiplex bursts of b back-to-back cells Poisson distributed with parameter λ_{wct} and load ρ_{wct}. with Poisson background traffic of parameter λ_p and load ρ_p. The total load is $\rho = \rho_{wct} + \rho_p$.

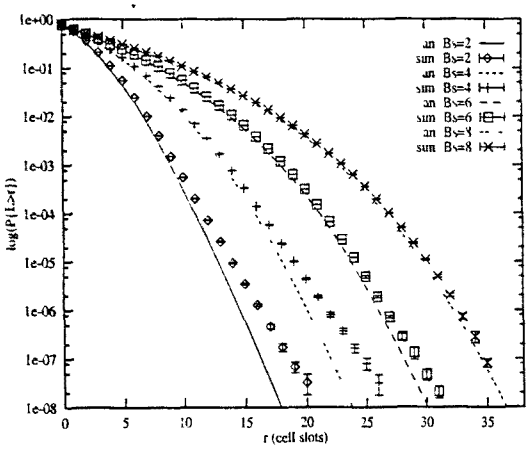


Figure 6.7: Comparison between the analytical model with simulations. $\rho_p = 0.4$, $\rho_{wct} = 0.4$. B_s varies, $T_{wct} = 10B_s$.

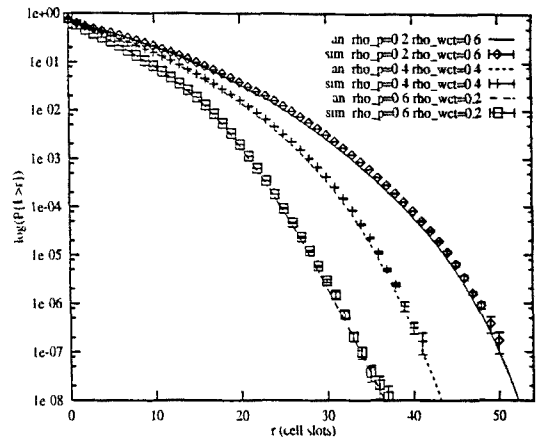


Figure 6.8: Comparison between the analytical model with simulations. ρ_p and ρ_{wct} varies. $T_{wct} = 100$, $B_s = 10$

$$P\{L_0 > x\} = \sum_{t=1}^{\infty}\sum_{i=0}^{\infty}\sum_{n=0}^{\infty} P\{\phi(t) = x, B_t = i, K_t = n\} \\ P\{L_0 = 0|\phi(t) = x, B_t = i, K_t = n\} \quad (6.10)$$

As in the other models we can calculate the former probabilities:

$$P\{\phi(t) = x, B_t = i, K_t = n\} \cdot P\{L_0 = 0|\phi(t) = x, B_t = i, K_t = n\} \; = \\ \begin{cases} (1-\rho)\frac{(\lambda_{wct}t)^i(\lambda_p t)^n}{i!\;\;n!}e^{-(\lambda_{wct}-\lambda_p)t}\frac{1}{t^i}q_t^i(t+x-n) & for \quad 1 \le t < b \\ (1-\rho)\frac{(\lambda_{wct}t)^i(\lambda_p t)^n}{i!\;\;n!}e^{-(\lambda_{wct}-\lambda_p)t}\sum_{j=0}^{i}\binom{i}{j}\frac{(t-b+1)^j}{t^i}q_{b-1}^{(i-j)}(t+x-bj-n) & \\ \qquad for \quad b \le t < \infty \end{cases} \quad (6.11)$$

As we mentioned in previous subsections, for low burst sizes a Poisson batch model fits quite well a WCT Poisson model. Moreover, for high loads the Poisson model loss accuracy due to cancelation of terms in the calculation of the complementary unfinished work distribution applying the Benes approach. Virtamo, in [91], solves this numerical problem through a polynomial representation of the unfinished work distribution. Application of the Virtamo's polynomial representation to the batch model is possible since it is easiest to apply the Jensen's equality (formula (4.12)), while to apply this relation to formula (6.11) is more difficult due to the terms $q_s^i(y)$. Thus in the next subsection we calculate batches Poisson distributed multiplexed with cells Poisson distributed applying a polynomial representation.

6.3.2 Multiplexing Batch sources Poisson distributed with Poisson Traffic.

Suppose an infinite buffer loaded with Poisson distributed cells with arrival rate λ_1 and bursts of cells in batches of size b also Poisson distributed with arrival rate λ_2. The total load of the system is $\rho = \rho_1 + \rho_2 < 1$. Using the Benes approach, [87] and [4], we can express the complementary unfinished work distribution as:

$$\begin{aligned} P\{L_0 > x\} &= (1-\rho) \sum_{n=x}^{\infty} \sum_{ib+n>x}^{\infty} \frac{\lambda_1^n}{n!} \frac{\lambda_2^i}{i!} e^{-(\lambda_1+\lambda_2)(ib+n-x)} (ib+n-x)^{n+i} \\ &= (1-\rho)\, e^{(\lambda_1+\lambda_2)x} \sum_{n=0}^{\infty} \sum_{ib+n>x}^{\infty} \frac{\beta_1^n}{n!} \frac{\beta_2^i}{i!} (ib+n-x)^{n+i} \end{aligned} \tag{6.12}$$

Where β_1 is $\lambda_1 e^{-(\lambda_1+\lambda_2)}$ and β_2 is $\lambda_2 e^{-(\lambda_1+\lambda_2)b}$.

$$P\{L_0 > x\} = 1 - (1-\rho)\, e^{(\lambda_1+\lambda_2)x} \sum_{n=0}^{\lfloor x \rfloor} \sum_{i=0}^{\lfloor \frac{x-n}{b} \rfloor} \frac{\beta_1^n}{n!} \frac{\beta_2^i}{i!} (ib+n-x)^{n+i} \tag{6.13}$$

For ρ larger than 0.7, numerical problems appear in the use of formula (6.12) or (6.13). To avoid these problems we can use similar arguments as in [91].

Let F(x) be defined by $F(x) = 1 - P\{L_0 > x\}$. It can be shown that its derivative is:

$$F'(x) = \lambda_1(F(x) - F(x-1)) + \lambda_2(F(x) - F(x-b)) \tag{6.14}$$

with initial conditions in $F(0) = 1-\rho$ and $F(x) = 0$ for $x < 0$. Introducing the function $G(x) = e^{-(\lambda_1+\lambda_2)x} F(x)$, with initial conditions $G(0) = 1-\rho$ and $G(x) = 0$ for $x < 0$, it is easy to derive:

$$G'(x) = -\beta_1 G(x-1) - \beta_2 G(x-b) \tag{6.15}$$

The function G(x) is a polynomial of degree n, $P_n(y_n)$, in the local coordinates $y_n = x - n$ (where: $x \in [n, n+1)$), with continuity $P_n(0) = P_{n-1}(1)$. Therefore, integrating equation (6.15):

$$P_n(y) = P_{n-1}(1) - \beta_1 \int_0^y P_{n-1}(z)dz - \beta_2 \int_0^y P_{n-b}(z)dz \tag{6.16}$$

Since $P_n(y) = \sum_{i=0}^{n} a_i^{(n)} y^i$, we can calculate the polynomial coefficients recursively. In this representation, coefficient $a_i^{(n)}$ is the *i-th* coefficient of polynomy P_n. If $P_0 = 1-\rho$ and we express a vector of coefficients of polynomy n as $P_n = (a_0^{(n)}, a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)})$, thus for $n > 0$:

$$P_n = (\sum_i a_i^{(n-1)}, -\beta_1 a_0^{(n-1)}/1 - \beta_2 a_0^{(n-b)}/1, -\beta_1 a_1^{(n-1)}/2 - \beta_2 a_1^{(n-b)}/2, \\ , -\beta_1 a_2^{(n-1)}/3 - \beta_2 a_2^{(n-b)}/3, \dots) \quad (6.17)$$

taking into account that the coefficients $a_i^{(n-b)}$ are only applied if $n \geq b$. Now, F(x) is determined by the equation: $F(x) = e^{(\lambda_1+\lambda_2)x} P_{\lfloor x \rfloor}(x - \lfloor x \rfloor)$.

6.4 Conclusions and comments.

In this chapter we have developed several approximations to WCT models. Some of these models allow us to evaluate the performance of a multiplexor when is fed by a mix of WCT sources and CBR sources. This case is an intermediate multiplexing situation where not all the sources behave as worst case traffic.

Other models consider that the peak bit rate is low enough to approximate the periodic sources by a Poisson process. In this case bursts of the same size arrive following a Poisson distribution. We have also analyzed the case in which part of the load represent bursts of cells and part of the load represent single cells.

As we have seen, some of these models provide good approximations under some traffic conditions. Mainly, the periodic mix of WCT and CBR sources behaves well if the WCT period is multiple of the CBR, while if we multiplex periodic WCT sources with Poisson traffic the model give accurate results if the Poisson load is low. The higher the Poisson load the less accurate the model is. However, for moderate and low Poisson loads the model behaves well. Finally, those models that multiplex only Poisson processes (bursts and cells Poisson distributed) are exact. However, for high loads these models as the $M/D/1$ queue system present loss of accuracy due to the cancelation of terms. A Batch Poisson model solved with the Benes approach can be easier computed through a polynomial representation than the corresponding WCT model. This is because in the WCT model we have to calculate the convolution of the *i-th* pulses. This term difficulties the application of Jensen's equality. However, in the Batch model is easier to apply this formula.

As a final conclusion I would say that the last model (WCT as batch Poisson multiplexed with Poisson traffic) offers reasonable CPU computation times, although it has the drawback of being the more pessimistic assumption in resource allocation terms.

The contributions of this chapter to the thesis are: a set of WCT models that overcome the limitations of the WCT periodic model.

Chapter 7

WCT in a tree network of ATM multiplexers.

We are interested in studying how WCT affects the queue length in a tree topology of rooted queues. These topologies may appear for instance in an access network. However. The main motivation of this chapter was to see how WCT affected next stages. The most interesting situation could be to consider two queues in tandem with WCT cross traffic. However, this configuration become quite complex. To gain knowledge about the topic of WCT in queues in tandem I have studied a topology simpler as can be a tree topology without cross traffic.

To solve this kind of configurations, we derive closed-form formulas for the queue length distributions in a discrete-time M-stage tree queueing network loaded with periodic traffic sources. In this type of network, the queues can be grouped in M groups or stages. Every queue of a stage is fed by all the exit traffic of any given number (which could be 0) of queues from the previous stage as well as by certain number (which could be 0) of external sources of traffic. All the entrance traffic in the network is routed to the root queue, which occupies the first stage (see in figure 7.7 an example for the case $M = 3$). We consider the discrete-time case in which all the queues have a single server with constant service time.

As is shown in [66] only two configurations are relevant to solve this system: the case $M = 1$, which corresponds to a single server queue with a constant service time (solved in chapter 5) and the case $M = 2$. The solution of the case $M > 2$ can be found solving systems with $M = 1$ and $M = 2$. Therefore, the main contribution in this chapter is to solve the two stage queueing system.

Before beggining this analysis I will summarize the work done by Morrison in [66]. This work. section 7.1, describes some lemmas to solve concentrating trees of discrete-time queues and will allow us later to extend the two stage network to a more general network. In section 7.2 we use again the Benes approach to solve the two stage network and calculate the CPDF of the Virtual Waiting Time in the root queue of the tree. At the end of this section, we make use of the lemmas introduced by Morrison to extend the study to a M stage network. Section 7.3 is devoted to results. Finally in section 7.4, to

complete the study, the average queueing delay and average number of cells in the root queue is calculated using a method based on priorities given in [65] by Modiano et al. Sections 7.2, 7.3 and 7.4 can be found in [21], [22] and [23].

7.1 Concentrating trees of discrete-time queues.

In this section we reproduce the property presented by J. A. Morrison in [66] to study concentrating rooted tree networks of discrete-time single server queues.

7.1.1 Pooling of data from M buffers into a single buffer.

Morrison establishes a combinatorial lemma which shows that there is an equivalent buffer with prescribed input that has the same output as the buffer in which data is pooled.

Consider the pooling of M buffers into a single buffer. The server transmits one packet in a unit interval. The buffer capacity is unlimited. Defining the following variables:

- $b_n^{(j)}$, $j = 1, ... M$: denote the contents of buffer j at time n.
- $x_n^{(j)}$, $j = 1, ..., M$: denote the number of packets entering buffer j in the time interval (n.n+1].
- c_n: denote the contents of the buffer of the root at time n.
- y_n: denote the number of packets entering directly the root queue in the time interval (n.n+1].
- Define $U(l)$ as:

$$U(l) = \begin{cases} 1 & l = 1, 2, ... \\ 0 & l = 0 \end{cases}$$

Therefore, we can express the contents of each buffer at time $n-1$ as:

$$b_{n+1}^{(j)} = b_n^{(j)} - U(b_n^{(j)}) + x_n^{(j)} \tag{7.1}$$

$$c_{n+1} = c_n - U(c_n) + \sum_{j=1}^{M} U(b_n^{(j)}) + y_n \tag{7.2}$$

Defining the inputs:

$$\begin{cases} w_0 = \sum_{j=1}^{M} b_0^{(j)} + y_0 \\ w_n = \sum_{j=1}^{M} x_{n-1}^{(j)} + y_n \qquad n = 1, 2, ... \end{cases} \tag{7.3}$$

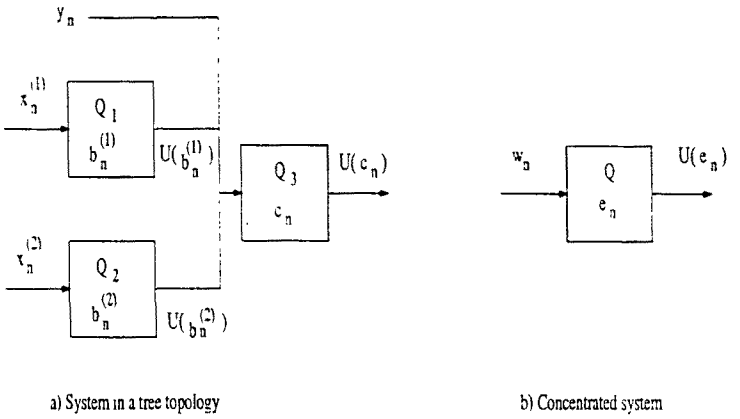


Figure 7.1: Equivalent queue in concentrating trees.

Morrison proofs that there is an equivalent buffer e_n with prescribed input which has the same output:

$$\begin{cases} U(e_n) = U(c_n) \\ e_{n+1} = e_n - U(e_n) + w_n \qquad n = 0, 1, 2, ... \end{cases} \tag{7.4}$$

Having as an example figure 7.1.1. Q_3 is fed by the outputs of two queues Q_1 and Q_2. The equation (7.4) says that queue Q_3 has the same output as queue Q with the same input as Q_1 and Q_2.

7.1.2 Buffer length in the root queue.

Consider now a system with I tandem queues, with unit service time and where the output of buffer i enters the buffer $i+1$. We define the following variables:

- $d_n^{(i)}$, $i = 1, ..., I$: denote the contents of buffer i at time n.
- $z_n^{(j)}$, $i = 1, ... I$: denote the number of packets entering buffer i from external sources in the time interval (n.n+1].

Then assuming $d_0^{(i)}$ and $z_0^{(i)}$ are non-negative integers. the content of buffer i at time $n+1$ is:

$$d_{n+1}^{(i)} = d_n^{(i)} - U(d_n^{(i)}) + U(d_n^{(i-1)}) + z_n^{(i)} \qquad n = 0.1.2, ... \quad i = 1, .., I \tag{7.5}$$

We now define the following inputs:

$$v_n^{(i)} = \begin{cases} \sum_{k=1}^{i} z_{n-i+k}^{(k)} & n = i-1, i. ... \\ d_0^{(i-n-1)} + \sum_{k=i-n}^{i} z_{n-i+k}^{(k)} & n = 0.1, ...i-2 \end{cases} \tag{7.6}$$

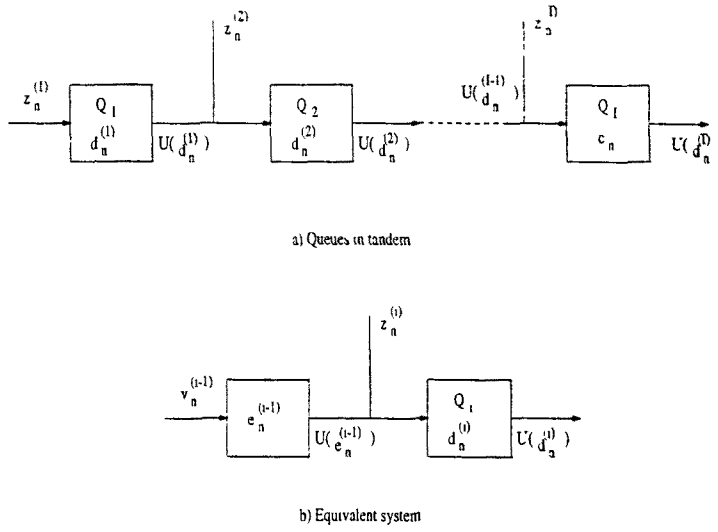


Figure 7.2: Equivalent queue in a tandem queue system.

Corollary: There is an equivalent buffer for which holds that:

$$\begin{cases} U(e_n^i) = U(d_n^i) \\ e_{n+1}^{(i)} = e_n^{(i)} - U(e_n^{(i)}) + v_n^{(i)} \end{cases} \qquad n = 0, 1, 2, \ldots \quad i = 1, .., I \tag{7.7}$$

and the buffer length at queue i will be:

$$d_{n+1}^{(i+1)} = d_n^{(i+1)} - U(d_n^{(i+1)}) + U(e_n^{(i)}) + z_n^{(i+1)} \qquad n = 0, 1, 2, \ldots \quad i = 1, \ldots I - 1 \tag{7.8}$$

That means that the content of buffer $i+1$ at time $n+1$ depends on the content of the buffer at time n, the external arrivals at time n and the output of the equivalent buffer at time n. Then, by replacing the first i queues by a single equivalent queue, with the same output as the *i-th* queue, we have reduced the problem of several queues in tandem to that of just two in tandem.

7.2 WCT multiplexing in an M-stage tree queueing network.

7.2.1 Two-stage tree network.

In this section we study the impact of WCT sources in two-stage tree networks. With this system we can later easily using the lemmas proposed by Morrison to extend the study to more complex systems. Therefore, the main contribution of this chapter to the thesis is to solve the two stage system. We consider the pooling of K buffers into a single queue, the root queue, which also receives N_r WCT sources as input (called exogenous traffic). Each one of the i (i=1...K) buffers multiplexes N_i WCT sources. We consider every WCT source in the network to have identical characteristics and every queue to

have a constant service time of one ATM slot. In fact the hypothesis that every source have the same characteristics can be relaxed to the hypothesis that the external sources which feed a given queue have the same characteristics and all the sources which enter the network are of equal period.

We are interested in the distribution of the queue length at the root queue (the queue length distribution of the queues of the second stage can be obtained using section 3). For simplicity's sake we substitute the burst arrivals of the sources which feed the second stage queues by batch arrivals of size b. This does not suppose any change in the results for the root queue, see figure 7.4. Therefore when we calculate the queue length distribution in the root queue we assume that each source which enters the queues of the second stage generates the following pattern: a batch of b cells followed by a silence of $T-1$ slots, see figure 7.3. The batch is uniformly distributed over the period T.

The WCT sources that enter directly in the root queue have the same characteristics as the sources described in section 2: independent WCT which emit a constant number, b, of back-to-back cells followed by a silence of $T-b$ slots.

In order to have a stable system we assume $\frac{(\sum_{i=1}^{K} N_i \cdot b + N_r \cdot b)}{T} < 1$.

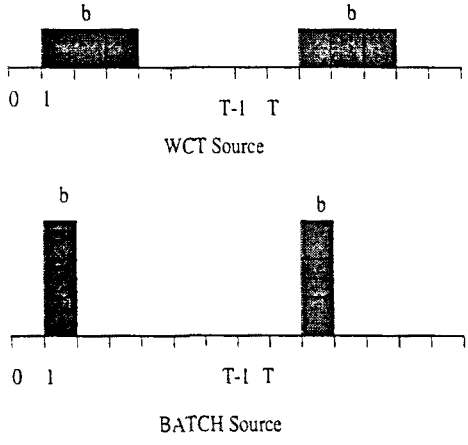


Figure 7.3: WCT and Batch periodic sources.

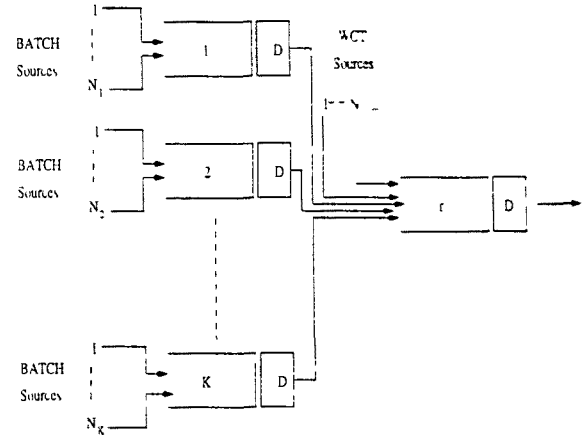


Figure 7.4: Two-stage tree network.

We define $N^r(t,0)$ as the number of cell arrivals to the root queue r in the interval -t,0). and L^r_{-t} as its queue length at time -t. We can express $P\{L^r_0 > x\}$ as:

$$P\{L^r_0 > x\} = \sum_{t=1}^{T} P\{N^r(t,0) = t+x \ \ and \ \ N^r(u,0) < u+x, t < u \leq T\} \qquad (7.9)$$

Note that events that occur in the queues of the second stage at any time will influence the root queue one slot later. We now introduce the following notation:

- $B^r(t_1, t_2)$ is the number of exogenous WCT sources that become active in an interval $[-t_1, -t_2)$ in the root queue.
- $B^i(t_1, t_2)$ is the number of batch arrivals to the *i-th* queue in an interval $[-t_1, -t_2)$, with $i = 1, \cdots, K$.
- $G(t, t+x, n_1, ..., n_K, n_r) = P\{N^r(t,0) = t+x \mid N^r(u,0) < u+x, t < u \leq T, B^1(t+1,1) = n_1 ..., B^K(t+1,1) = n_K, B^r(t,0) = n_r\}$
- $L(t, n_1, ..., n_K, n_r) = P\{N^r(u,0) < u+x, t < u \leq T \mid B^1(t+1,1) = n_1 ..., B^K(t+1,1) = n_K, B^r(t,0) = n_r\}$
- $B(t, n_1, ..., n_K) = P\{B^1(t+1,1) = n_1, ..., B^K(t+1,1) = n_K\}$.

Equation (7.9) can be transformed into:

$$P\{L_0^r > x\} = \sum_{t=1}^{T} \sum_{n_1=0}^{N_1} .. \sum_{n_K=0}^{N_K} \sum_{n_r=0}^{N_r} G(t, t+x, n_1, ..., n_K, n_r) \cdot \cdot L(t, n_1, ..., n_K, n_r) \cdot B(t, n_1, ... n_K) \cdot P\{B^r(t,0) = n_r\} \quad (7.10)$$

We now determine the expression for each one of these terms.

Term $G(t, t+x, n_1, ..., n_K, n_r)$:

In order to obtain this term we should bear in mind that the event $\{N^r(u,0) < u+x, t < u \leq T\}$ implies that the K queues are empty at time -(t+1). Therefore the contribution which each queue of the second stage makes to the root queue during the interval $[-t, 0)$ can only have as its origin the cells which arrive at the said queue during the interval [-(t+1),-1). We introduce now the following notation:

- $N^{ri}(t,0)$ is the number of cells that arrive to the root queue r from the *i-th* queue in the interval $[-t, 0)$.
- $R^{ri}(t, k_i, n_i)$ is the probability $P\{N^{ri}(t,0) = k_i \mid$ *queue i is empty at -(t+1),* $B^i(t+1,1) = n_i\}$.
- $N^{rr}(t,0)$ is the number of cells that arrive to queue r from the exogenous WCT sources.
- $R^{rr}(t, k_r, n_r)$ is $P\{N^{rr}(t,0) = k_r \mid$ *queue r is empty at -t,* $B^r(t,0) = n_r\}$.

Probability $G(t, t+x, n_1, ..., n_K, n_r)$ will be the convolution of R^{ri} $(i = 1, ...K)$ and R^{rr}:

$$G(t, t+x, n_1, ..., n_K, n_r) = \sum_{\sum_{j=1}^{K} k_i + k_r = t+x} \prod_{i=1}^{K} R^{ri}(t, k_i, n_i) \cdot R^{rr}(t, k_r, n_r) \quad (7.11)$$

Terms $R^{rr}(t, k_r, n_r)$ **and** $R^{ri}(t, k_i, n_i)$:

The term $R^{rr}(t, k_r, n_r)$ has the same expression as equation (5.2) of section 2: $R^{rr}(t, k_r, n_r) = A(t, k_r, n_r)$.

To calculate the term $R^{ri}(t, k_i, n_i)$, we must take into account the fact that the temporal axis in the queues of the second stage is shifted in one time slot (the arrivals which are relevant to the root queue in the time instant $t = 0$ are produced in the queues of the second stage during the interval [-(T+1),-1)). Since each queue begins empty, we can substitute queue i for a queue of period t in the interval [-(t+1),-1), to which there arrive n_i batches of size b uniformly and independently distributed. It follows that, if $n_i b$ cells arrive at queue i and k_i cells are emitted, at time instant $t = -1$ there will be produced either $n_i b - k_i + 1$ queue (in the case $n_i b > k_i$, see figure 7.5.c) or a queue of length 0 or 1 (in the case $n_i b = k_i$, see figures 7.5.a,b). In consequence we will have the following cases:

$$R^{ri}(t, k_i, n_i) = \begin{cases} 1 & if \quad k_i = 0 \quad and \quad n_i = 0 \\ 0 & if \quad k_i = 0 \quad and \quad n_i > 0 \\ 0 & if \quad k_i > 0 \quad and \quad n_i \cdot b < k_i \\ q_b(t, n_i, 0) + q_b(t, n_i, 1) & if \quad k_i > 0 \quad and \quad n_i \cdot b = k_i \\ q_b(t, n_i, n_i b - k_i + 1) & if \quad k_i > 0 \quad and \quad n_i \cdot b > k_i \end{cases} \tag{7.12}$$

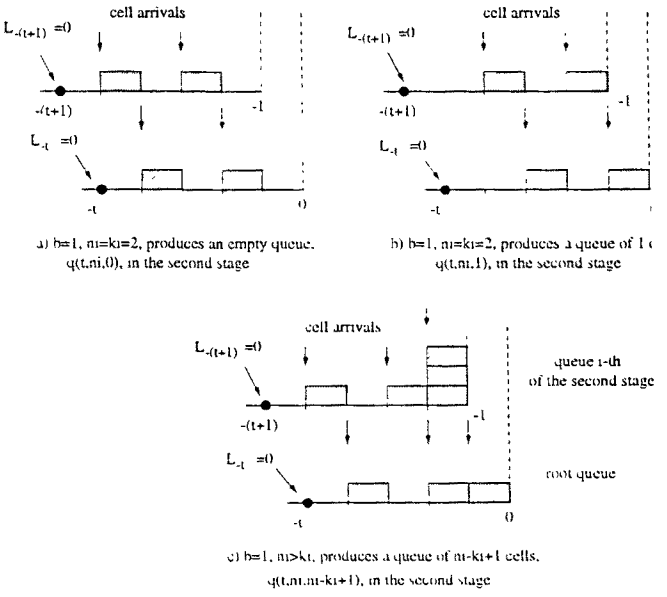


Figure 7.5: Cases to calculate the term $R^{ri}(t, k_i, n_i)$, (with b=1).

Where $q_b(T, N, r)$ is the probability that the queue length of a $ND/D/1$ with period T and N batch arrivals of length b is r. As is easily demonstrated, the expression of $q_b(T, N, r)$ is:

$$\begin{cases} q_b(T,N,r) = Q_b(T,N,r-1) - Q_b(T,N,r), \\ Q_b(T,N,r) = \sum_{i=\lceil \frac{r}{b} \rceil}^{N} \binom{N}{i} (\frac{ib-r}{T})^i (1-\frac{ib-r}{T})^{N-i} (1-\frac{(N-i)b}{T-ib+r}) \quad T+r > Nb \end{cases} \quad (7.13)$$

Term $L(t, n_1, ..., n_K, n_r)$:

To calculate $L(t, n_1, ..., n_K, n_r) = P\{N^r(u,0) < u+x, t < u \leq T \mid B^1(t+1,1) = n_1, ... B^K(t+1,1) = n_K, B^r(t,0) = n_r\}$ we use similar arguments as in [87] or [44]. The event $P\{N^r(u,0) < u+x, t < u \leq T \mid B^1(t+1,1) = n_1, ..., B^K(t+1,1) = n_K, B^r(t,0) = n_r\}$ corresponds to the situation in which the root queue of an auxiliary network of queues loaded with periodic arrivals of period $T-t$ in the interval [-(T-t),0) is empty at time 0.

During a period, the root queue of the said auxiliary network receives cells with different origins, see figure 7.6.b:

- cells which proceed from the exits of the second stage queues,
- cells which proceed from bursts generated by the external sources which directly feed the root queue and that began during the actual period, and
- cells which proceed from bursts generated by the external sources which directly feed the root queue and that began in a previous period but that in part contribute to the actual period.

The cells which enter the root queue from the second stage queues in the auxiliary system are those which arrived at the second stage queues in $N_i - n_i$ batches of b cells uniformly distributed in the interval [-(T-t+1),-1). At this point it is important to note that for the auxiliary root queue to be empty in the time instant $t = 0$ it is necessary for the second stage queues to be empty too. This means that all cells which arrive at the second stage queues between [-(T-t+1),-1) must be served beforehand and routed to the root queue at time instant $t = 0$.

Arrivals which proceed directly from external sources and which feed the root queue in the auxiliary system consist of $N_r - n_r$ bursts of b cells uniformly distributed in the interval [-(T-t).0).

Lastly, in the root queue of the auxiliary system we have an entrance in batch at the beginning of interval [-(T-t).0) with which we model the contribution of the bursts which (in the original system) began to be emitted in the previous period. This batch consists of $\sum_{i=1}^{K} n_i b + n_r b - t - x$ cells.

Since the original root systems is stable, $\sum N_i b + N_r b < T$, the auxiliary system will be so too: $\sum_{i=1}^{K}(N_i - n_i)b + (N_r - n_r)b + \sum_{i=1}^{K} n_i b + n_r b - t - x < T - t$. Therefore the batch arrival at the beginning of the interval [-(T-t),0) will be served before the end of the interval, and will not contribute to the length of the queue in the instance $t = 0$.

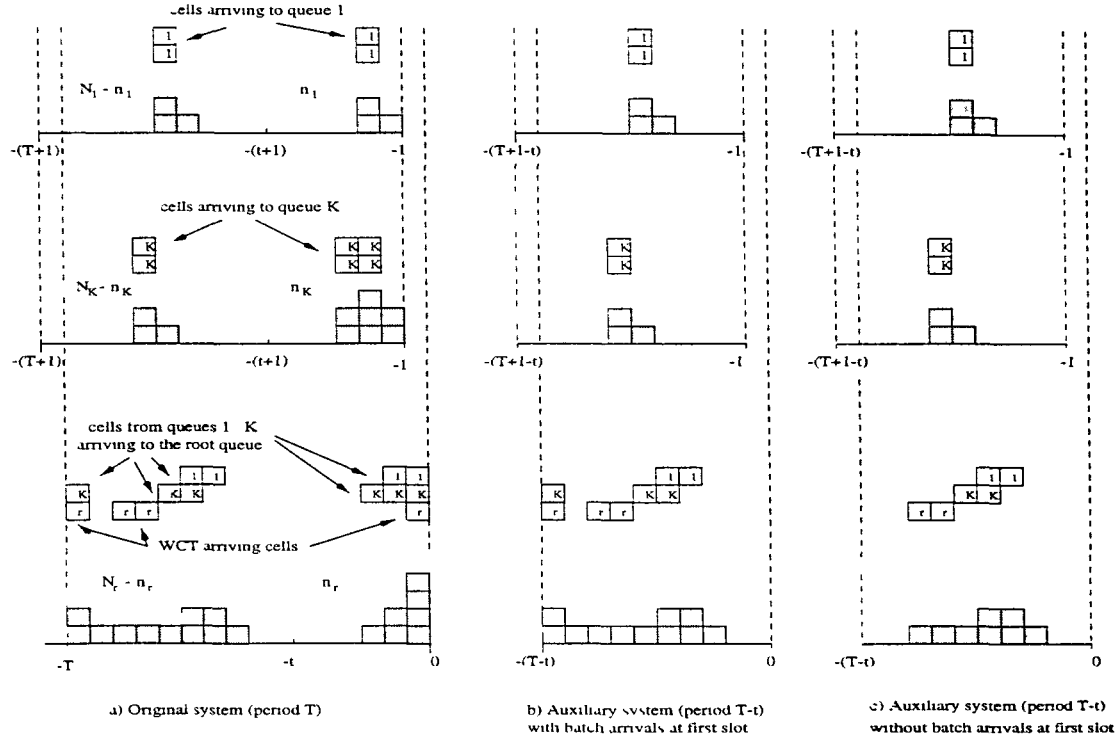


Figure 7.6: Equivalent system to calculate term $L(t, n_1, ..., n_K, n_r)$. a) Original system. b) Auxiliary system. where bursts from the auxiliary second stage queues and external WCT bursts plus cells belonging to bursts that began in the previous period arrive to the auxiliary root queue. c) Auxiliary system, where only bursts from the auxiliary second stage queues and external WCT bursts arrive to the auxiliary root queue

We have then a situation in which the only arrivals which have to be counted in order to obtain an empty queue in the auxiliary root queue at time 0 are those which proceed from the exit of the $N_i - n_i$ batches and the $N_r - n_r$ WCT sources (see figure 7.6.c). Therefore:

$$L(t, n_1, ..., n_K, n_r) = 1 - \frac{(\sum_{i=1}^{K} (N_i - n_i) \cdot b + (N_r - n_r) \cdot b)}{T - t} \quad (7.14)$$

Terms $P\{B^r(t,0) = n_r\}$ and $B(t, n_1, ..., n_K)$:

$P\{B^r(t,0) = n_r\}$ has the same expression as equation (5.10) in section 2: a binomial distribution. Since each queue is independent and the batches are uniformly distributed the term $B(t, n_1, ..., n_K)$ can be calculated as:

$$B(t, n_1, ..., n_K) = \prod_{i=1}^{K} P\{B^i(t+1, 1) = n_i\} = \prod_{i=1}^{K} \binom{N_i}{n_i} (\frac{t}{T})^{n_i} \cdot (1 - \frac{t}{T})^{N_i - n_i} \quad (7.15)$$

7.2.2 M-stage tree network.

Once the two stage system is solved, we can generalize the results for an M-stage tree network. Multiplexing stage i is pooled by K_{i-1} queues and N_{pq} WCT sources (with q=1...K). We can make use of the property of the rooted tree networks with discrete-time single server queues presented by J.A. Morrison in [66] an summarized in subsection 7.1, where it is showed that in such networks, the rooted tree network can be reduced to a two-stage network with a prescribed input. Specifically, to calculate the queue length distribution at the root queue, we can assume that the external sources that enter the queues which hang from each branch of a second stage queue, enter the second stage queue directly.

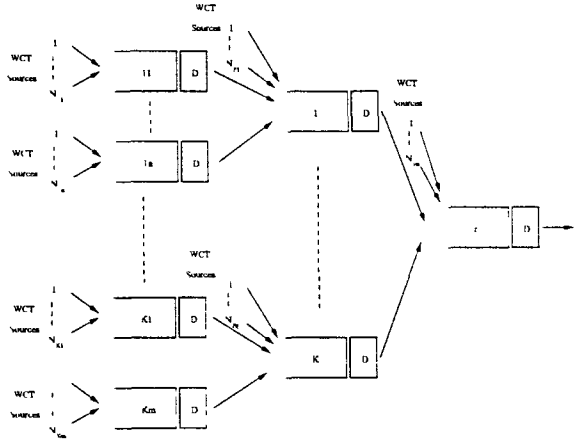


Figure 7.7: Three stage network.

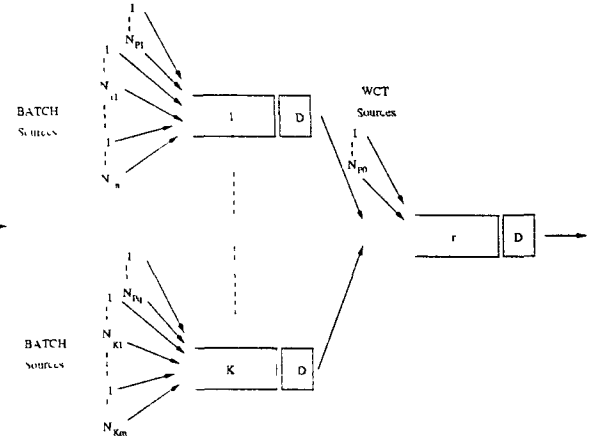


Figure 7.8: Equivalent network.

Using this argument, and the fact that each queue of the network is a root queue of a corresponding sub-tree, the analysis of the network with M-stages can be done from the analysis of equivalent systems with $M = 1$ and $M = 2$.

For example, considering a three multiplexing stage network, see figure 7.7. Each queue is fed by WCT sources. In order to obtain the queue length distribution at queue 0 (the root queue), we can substitute this system for an equivalent two-stage network (figure 7.8). In this equivalent system, only the sources that enter the root queue directly are WCT. The rest behave as Batch sources.

7.3 Results.

We consider a scenario in which we multiplex connections in networks with a tree topology. This topology could be present in an ATM access network. Since the connections have declared a pair (PEI, CDVT), all the sources might send WCT.

Figure 7.9 and 7.10 show the impact of WCT in the buffer of the root queue. In this example, we have considered that two buffers pool the root stage and therefore the system only has two stages. Each of these buffers multiplex four WCT sources of period 15 cell slots. Four more WCT sources directly enter the second stage. The total load is 0.8.

In figure 7.10 the burst size is kept to 4 cells, while the period T varies. Although the increase in queue length on increasing Bs is slightly absorbed, we can still observe the same effect as in a single stage network: The buffer occupancy increases linearly with Bs when WCT enters the network.

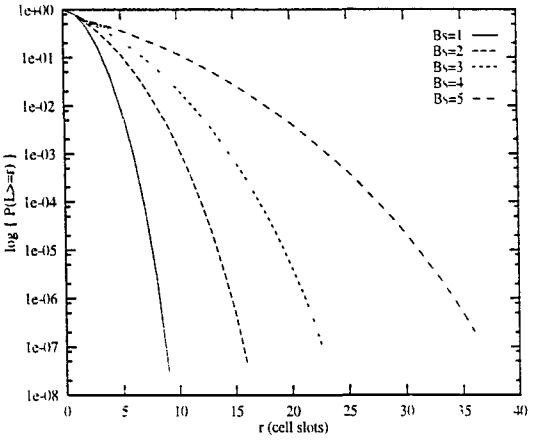


Figure 7.9: Queue Length CPDF for the root queue. Two stages. $T = 15$, $\rho = 0.8$ and $B_s = 4$ varies

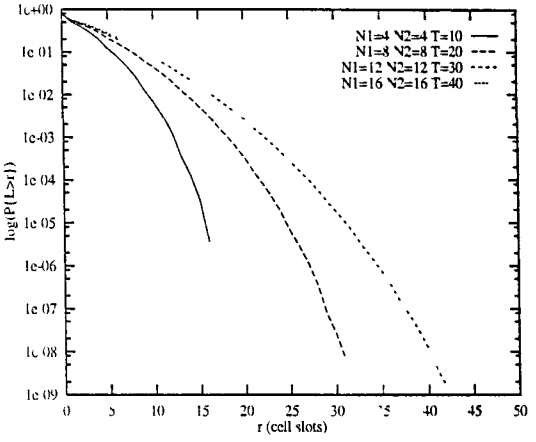


Figure 7.10: Queue Length CPDF for the root queue. Two stages. $B_s = 4$, $\rho = 0.8$ and Period varies.

In order to see how the number of second-stage queues influences the root queue, we have studied various configurations with different loads. In figure 7.11 we have drawn the required buffer length in a root queue so that the probability of cell losses is below 10^{-10} in function of Bs when the second stage has two queues. We call the number of sources that enter the second-stage queues N_1 and N_2 respectively. As can be seen in figure 7.11, the worst situation occurs with balanced loads (e.g. $N_1 = N_2 = 6$ over a period $T = 15$). In figure 7.12, we have plotted the total buffer size required by both queues in the second stage. As can be seen, the number of cells required for both second-stage queues remains constant for a single Bs independent of the load balancing.

Finally in figure 7.13, we show the queue length distribution when the second stage has $K = 2, 3, 4, 12$ queues. All the curves are calculated with a total load of $\rho = 0.8$ and a $Bs = 6$. The period of the sources is $T = 15$ and the load is balanced among the K queues ($N_k = \frac{12}{K}$ sources in each queue). The worst situation is when each queue is fed by only one source ($K = 12$). In that case, the same results would be obtained as if 12 WCT sources were multiplexed in one stage.

7.4 Average queueing delay in the root queue.

In [22], we obtained expressions for the Virtual Waiting time distribution for any queue in a tree network with Worst Case Traffic (WCT) or Constant Bit Rate (CBR) as input traffic. We are interested in calculate the average queueing delay and average waiting time in any queue of the tree network. For that purpose we will make use of an equivalent network based on priorities proposed by *Modiano et al* in [65]. In this proposal, they first

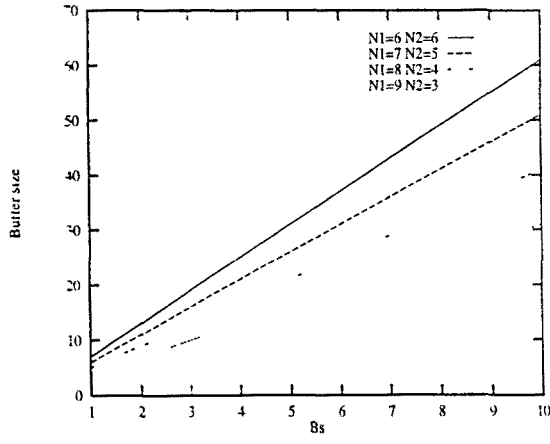


Figure 7.11: Buffer size in the root queue versus Bs. Two stages.

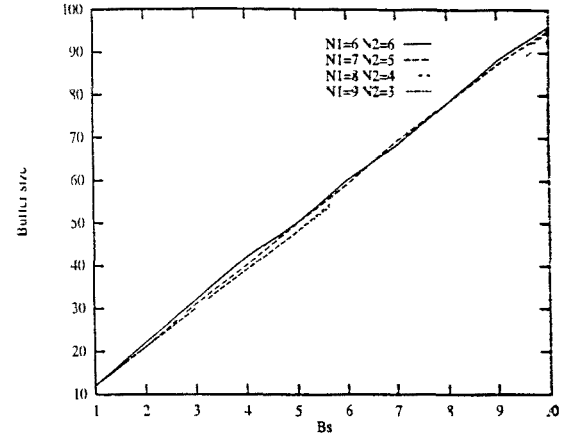


Figure 7.12: Sum of buffer sizes required by the two queues of the second stage versus Bs.

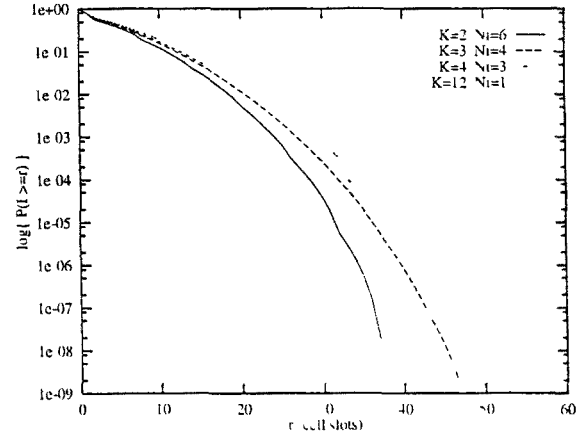


Figure 7.13: Queue length distribution for 2,3,4,12 queues in the second stage.

consider a simple case with two queues in tandem and Poisson traffic as input. High-priority is given to customers in transit from one queue to the other, while low-priority is given to that traffic that enters directly the second queue (called exogenous traffic). Solving this system an extension to larger systems can be easily derived. *Modiano et al* solve the tree system with Poisson traffic. As we will see, the solution of waiting times for any queue of the tree network is easy to compute.

Here, we first extend the model to a $G/D/1$ queue system where the average waiting time in the queue is known from the first and second moments of the interarrival-time and idle-time distributions, see [61]. This formula applies if the interarrival-times are independents and the idle-time distribution is known (e.g Poisson traffic as input). If we have a queue fed with periodic sources as in ATM (CBR sources), we can not apply this formula since the interarrival-times are not independent due to the periodic nature of the traffic. However, in [38] is given the average waiting time for an $ND/D/1$ queue system. Given this formula, we can apply the model proposed in [65] and calculate the average waiting time and the average number of cells in the queue (or in the system) for

homogeneous periodic sources in a tree network of ATM multiplexers. We compare the waiting times in the root queue system of the tree network whose input is Poisson Traffic with the same tree network whose input are CBR or WCT sources for different loads and periods.

7.4.1 Average delays in two queues in tandem and its extension to a tree network.

Two queues in tandem.

We consider the following system, figure 7.14: two queues in tandem. To the first queue system enters traffic with arrival rate λ_1. The output of this queue enters the second queue and is multiplexed with a second stream traffic (which we call exogenous traffic) with arrival rate λ_2.

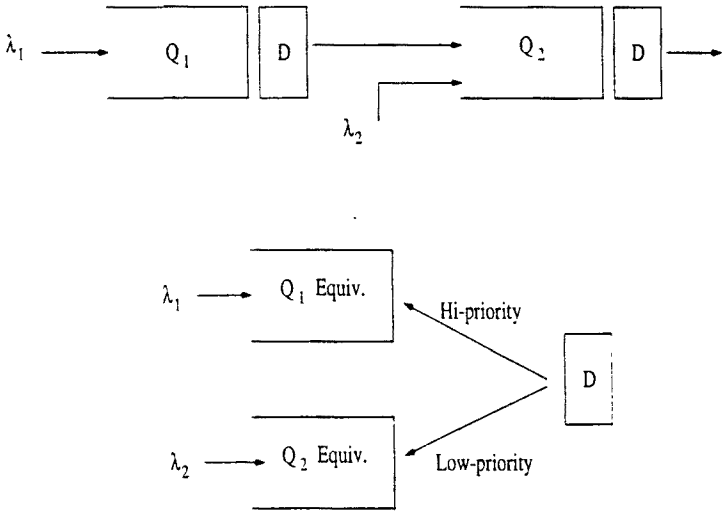


Figure 7.14: Two queues in tandem and its equivalent queue system.

To calculate the mean number of customers in the second queue, we can study an equivalent system with priorities as propose *Modiano et al* in [65]. We assume that the customers departing from the first queue are served with priority without any delay at the second queue, and that the exogenous traffic at the second queue are served with lower priority (e.g. when non of the customers from the first queue are at the second queue). see figure 7.14.

Although the delay in the equivalent system for each class would be not the same as in the original system, the average delay in the overall system is not affected by the priorities. Note that customers from the first queue do not experience any delay in the second queue since they are served immediately (those customers experience all their delay in the first queue). Customers from exogenous traffic are served when there is no priority traffic in the system (e.g. there are not customers departing from the first queue).

Thus, in the equivalent system, see figure 7.14, two exogenous traffic streams with arrival rates of λ_1 and λ_2 enter the system and λ_1 traffic is served with priority to λ_2. We can calculate the delay experienced by the exogenous traffic at the second queue since it is the same as the delay experienced by the low-priority traffic in the equivalent system.

We define $\overline{Q}$ as the average overall queue size (low and high priority traffic), $\overline{Q_1}$ as the average queue size at the high priority queue and $\overline{Q_2}$ as the average queue size at the low priority queue. Therefore:

$$\overline{Q_2} = \overline{Q} - \overline{Q_1} \tag{7.16}$$

As an example, consider that Poisson traffic with arrival rate λ_1 enters the first queue and that exogenous Poisson traffic with arrival rate λ_2 enters the second queue. We know that the mean number of customers for an $M/D/1$ queue (service-time is the unit time) with arrival rate of λ is: $\overline{Q_{M/D/1}} = \frac{\lambda^2}{2(1-\lambda)}$. Therefore, (7.16) becomes:

$$\overline{Q_2} = \frac{(\lambda_1+\lambda_2)^2}{2(1-\lambda_1-\lambda_2)} - \frac{\lambda_1^2}{2(1-\lambda_1)} \tag{7.17}$$

Applying Little's formula:

$$\overline{W_2} = \frac{\overline{Q_2}}{\lambda_1+\lambda_2} \tag{7.18}$$

and taking into account the customer in service, we have delays in the system:

$$\overline{W_{s_2}} = \frac{\overline{Q_2}}{\lambda_1+\lambda_2} + 1 \tag{7.19}$$

$$\overline{Q_{s_2}} = \overline{Q_2} + \lambda_1 + \lambda_2 \tag{7.20}$$

Extension to a tree network.

We now consider a two level tree network, figure 7.15. The first stage has K queues whose input has arbitrary distribution with arrival rate λ_i. The service is deterministic with service-time the unit time. We define $Q_i(t)$ as the number of customers in queue $i = 1, ..., K$ including queue room and service room at slot t, and $Q_r(t)$ the number of customers in the root queue including only queue room at slot time t. $Q(t)$ is the number of customers in the total system.

Therefore, there holds:

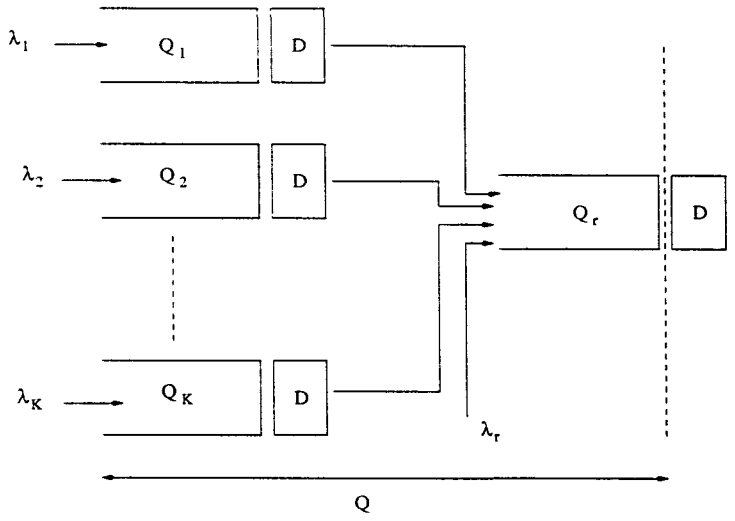


Figure 7.15: Tree network topology.

$$Q(t) = \sum_i Q_i(t) + Q_r(t) \tag{7.21}$$

and we define $A_i(t)$ as the number of arrivals in slot t to the queue Q_i, and $A_r(t)$ as the exogenous traffic entering directly the root queue. Let be the function $I(x) = 1$ if $x > 0$ and $I(x) = 0$ otherwise. Thus, we express the following relations:

$$Q_i(t) = Q_i(t-1) + A(t) - I(Q_i(t-1)) \;\; with \;\; i = 1,..K \tag{7.22}$$

$$Q_r(t) = Q_r(t-1) + \sum_{i=1}^{K} I(Q_i(t-1)) + A_r(t) - I(\sum_{i=1}^{K} Q_i(t-1) + Q_r(t-1)) \tag{7.23}$$

That is. the number of customers at queue i at slot t is the number of customers at queue i at slot $t-1$ plus the new arrivals in this slot less the customer being served if any. The same holds for queue Q_r but we must take into account the cells arriving from any queue of the previous stage and the arriving exogenous traffic.

It follows that we can apply $Q(t) = \sum_i Q_i(t) + Q_r(t)$ and:

$$\begin{aligned} Q(t) = Q(t-1) + \sum_{i=1}^{K} A_i(t) + A_r(t) - I(Q(t-1)) = \\ = Q(t-1) + A(t) - I(Q(t-1)) \end{aligned} \tag{7.24}$$

Note that the input to queue Q is $A(t) = \sum_{i=1}^{K} A_i(t) + A_r(t)$, and we can use priorities as we did with the two queues to calculate $\overline{Q_r}$ as:

$$\overline{Q_r} = \overline{Q} - \sum_{i=1}^{K} \overline{Q_i} \tag{7.25}$$

We come back to the example with Poisson traffic. Let λ_i be the mean arrival rate to queue $i = 1, ... K$ and λ_r the mean arrival rate of the exogenous traffic. The number of customers in the root queue will be:

$$\overline{Q_r} = \frac{(\sum_{i=1}^{K} \lambda_i + \lambda_r)^2}{2(1 - \sum_{i=1}^{K} \lambda_i - \lambda_r)} - \sum_{i=1}^{K} \frac{\lambda_i^2}{2(1-\lambda_i)} \tag{7.26}$$

Note that we can make use of the Lemma's presented by Morrison in [66] to calculate the average delay in any queue of a tree network with M stages, each stage pooled by M_k buffers and exogenous traffic in any queue.

Average delays in a tree network with WCT sources.

Now we are interested in calculate average delays in any queue system of the tree whose input is homogeneous CBR traffic of period T_{cbr} or homogeneous WCT with period T_{wct}. Note that equations (7.24) and (7.25) are still valid. We first have to know the average waiting time in an $ND/D/1$ queue system.

However, *Dron et al* show in [38] that the average waiting time in the queue, $\overline{W}$, of a queue system multiplexing N periodic sources of period T_{cbr} (with load $\rho = \frac{N}{T_{cbr}}$) is the following:

$$\overline{W} = \frac{(N-1)!}{2} \sum_{k=1}^{(N-1)} \frac{1}{T_{cbr}^k} \frac{1}{(N-1-k)!} \tag{7.27}$$

The number of cells in the queue $\overline{Q}$ can be calculated using Little's formula: in periodic systems this can be interpreted as following, see [55]:

$$P\{W = x\} = P\{a\ departure\ leaves\ x-1\ cells\ in\ the\ queue\} = \frac{P\{Q=x\}}{P\{Q>0\}} = \frac{T_{cbr}}{N} P\{Q = x\} \quad with \quad x > 0 \tag{7.28}$$

Therefore applying this relation:

$$\overline{Q} = \rho \cdot \overline{W} = \frac{N!}{2T_{cbr}} \sum_{k=1}^{(N-1)} \frac{1}{T_{cbr}^k \cdot (N-1-k)!} \tag{7.29}$$

In the general case of multiplexing N WCT sources with burst size B_s and period $T_{wct} = B_s T_{cbr}$, the average waiting time in the queue can be expressed as B_s times the average

waiting time in the queue for a $ND/D/1$ system with N CBR sources of period T_{cbr}. If we call $\overline{W_{wct}}$ the average waiting time in the queue for the WCT and $\overline{W_{cbr}}$ the average waiting time in the queue for the corresponding CBR traffic of period $T_{cbr} = \frac{T_{wct}}{B_s}$

$$\overline{W_{wct}} = B_s \overline{W_{cbr}} = \frac{(N-1)!\, B_s}{2} \sum_{k=1}^{(N-1)} \frac{1}{T_{cbr}^k} \frac{1}{(N-1-k)!} \tag{7.30}$$

$$\overline{Q_{wct}} = B_s \overline{Q_{cbr}} = \frac{N!\, B_s}{2T_{cbr}} \sum_{k=1}^{(N-1)} \frac{1}{T_{cbr}^k} \frac{1}{(N-1-k)!} \tag{7.31}$$

We consider the tree topology as in the former section. The first stage has K queues fed each one by N_i WCT sources of period $T_{wct} = B_s T_{cbr}$ being $i = 1, ..., K$ and B_s the burst size. Let ρ_i be the load at each queue ($\rho_i = \frac{N_i B_s}{T_{wct}}$). The output of each of these queue systems is multiplexed with N_r exogenous WCT sources with load $\rho_r = \frac{N_r B_s}{T_{wct}}$. see figure 7.16. The total load in the root queue is $\rho = \sum_i \rho_i + \rho_r$. We define N as the total number of WCT sources in the system: $N = \sum_i N_i + N_r$. Therefore, we can express the average number of cells in the root queue as:

$$\overline{Q_r} = \frac{N!\, B_s}{2T_{cbr}} \sum_{k=1}^{(N-1)} \frac{1}{T_{cbr}^k (N-1-k)!} - \sum_{i=1}^{K} \frac{N_i!\, B_s}{2T_{cbr}} \sum_{k=1}^{(N_i-1)} \frac{1}{T_{cbr}^k (N_i-1-k)!} \tag{7.32}$$

The average waiting time in the queue is given by Little's formula as:

$$\overline{W_r} = \frac{\overline{Q_r}}{\rho} = T_{cbr} \frac{\overline{Q_r}}{\sum_{i=1}^{K} N_i + N_r} \tag{7.33}$$

And the average waiting time and average number of cells in the system:

$$\overline{W_{r_s}} = \frac{\overline{Q_r}}{\rho} + 1 \tag{7.34}$$

$$\overline{Q_{r_s}} = \overline{Q_r} + \rho \tag{7.35}$$

7.4.2 Results

In table 7.1 and 7.2 we compare the average waiting time in the queue, $\overline{W}$, for a single stage for the multiplexing of CBR sources with period $T_{cbr} = 100$, $T_{cbr} = 1000$ and Poisson. We compare the same scenarios when the root queue is pooled by $K = 2$ and $K = 4$ buffers. In this case the load is balanced among the buffers. Periods are the same as in the single system. The average waiting time in the root queue is compared for the

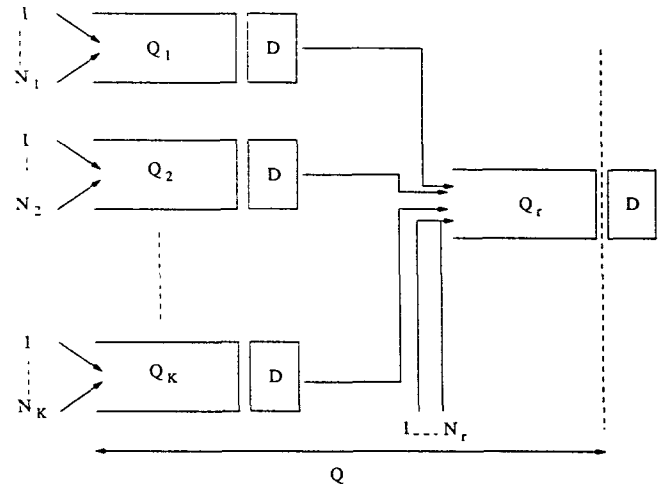


Figure 7.16: Tree network topology with WCT sources as input traffic.

	one stage		
ρ	$T_{cbr} = 100$	$T_{cbr} = 1000$	$M/D/1$
0.98	4.766	12.482	24.5
0.94	3.557	6.420	7.833
0.90	2.747	4.101	4.5
0.84	1.956	2.514	2.625
0.80	1.598	1.941	2.0
0.70	1.02	1.1486	1.166
0.50	0.463	0.496	0.5

Table 7.1: Comparison between CBR sources and Poisson sources in a single stage.

different periods and for Poisson traffic. We can observe what it is known: Poisson is pessimistic for high loads in a single queue system. For $K = 2$ the average delay decreases slightly, but not too much. We also observe that if the number of buffers is increased it looks like more to a single system.

In table 7.3, we have considered that the load is unbalanced between two buffers that pool the root queue. 80% of the total load enters the first queue, while 20% enters the second queue. The average delay in the root queue also decreases slightly. However we must take into account that the average delay of the 80% traffic will be higher in the previous stage.

Finally, in table 7.4 we consider WCT sources with burst size $B_s = 10$ and period $T_{wct} = 200$. As we stated, the average waiting time is B_s times the average waiting time in an $ND/D/1$ of period $T_{cbr} = 20$.

	two stages, K=2			two stages, K=4		
ρ	$T_{cbr} = 100$	$T_{cbr} = 1000$	$M/D/1$	$T_{cbr} = 100$	$T_{cbr} = 1000$	$M/D/1$
0.98	4.321	12.005	24.01	4.615	12.321	24.33
0.94	3.145	5.980	7.389	3.414	6.268	7.679
0.90	2.365	3.695	4.090	2.612	3.957	4.354
0.84	1.618	2.154	2.262	1.833	2.382	2.492
0.80	1.286	1.610	1.666	1.482	1.817	1.875
0.70	0.768	0.881	0.897	0.922	1.043	1.060
0.50	0.333	0.330	0.333	0.399	0.425	0.428

Table 7.2: Two stage tree network with K=2 buffers and K=4 buffers that pool the root queue, balanced load in each buffer. Comparison between CBR sources and Poisson sources.

	two stages, K=2		
ρ	$T_{cbr} = 100$	$T_{cbr} = 1000$	$M/D/1$
0.98	3.5869	11.043	23.023
0.94	2.5250	5.2098	6.5972
0.90	1.8385	3.0687	3.4494
0.84	1.2187	1.6859	1.7852
0.80	0.9422	1.2198	1.2698
0.70	0.5379	0.628	0.6412
0.50	0.2044	0.2202	0.2222

Table 7.3: Two stage tree with K=2 buffers that pool the root queue, load in Q_1 is $\rho_1 = 0.2\rho$ and load in Q_2 is $\rho_2 = 0.8\rho$.

	two stages, K=2		two stages, K=4	
ρ	$T_{wct} = 20$	$T_{wct} = 200$	$T_{wct} = 20$	$T_{wct} = 200$
ρ	$B_s = 1$	$B_s = 10$	$B_s = 1$	$B_s = 10$
0.9	1.1578	11.578	1.353	13.53
0.8	0.7714	7.7148	0.9322	9.322
0.7	0.5231	5.2312	0.6495	6.495
0.6	0.3570	3.5704	0.4586	4.586
0.5	0.2422	2.4228	0.3173	3.173

Table 7.4: Comparison between two stage tree network with K=2 buffers and K=4 buffers that pool the root queue, balanced load in each buffer, WCT sources.

7.5 Conclusions and comments.

In this chapter we have studied a tree network of ATM multiplexors. We think that although this kind of topology is limited, its solution can lead to give us more insight to more general queue systems such as queues in tandem with cross traffic. We are interested in knowing how the delay distribution evolves through the multiplexing stages. We have obtained the delay distribution in any queue of the tree, and the average delay and average number of cells in the system.

We have observed that the impact of WCT on the queue length distribution is still important. although slightly absorbed. We can say that the buffer occupancy is an increasing function of the burst size but not so linearly as with the single stage.

Studying the buffer occupancy in all the buffers under several load conditions, we can say that the most demanding situation in terms of buffer units needed in the root queue is that in which the load is balanced among the second stage queues. However, the total buffer units needed in all the second stages remain constant independent on the load balancing. This situation can also been observed in the average delay computed from different cases.

However, the results also depend on how many second stage queues are connected to the root queue. The best situation is when there are only two queues in the second stage. Increasing the number of second stages mean to increase the buffer occupancy in the root queue. The extreme case would be when each WCT source is attached to one queue. In that case the tree system would be equivalent to have a single queue system.

The contributions of this chapter to the thesis are: the analysis of a tree topology network whose input is periodic traffic. The relevant analysis is the two stage network since the general case of M stages depend on the two stage case. This analysis consists of two parts: the queue length distribution of the root queue and the average queueing times. The calcul of the queue length distribution of the root queue is totally new. To calculate the average queueing time, I have used a method found in the literature.

Chapter 8

Resource management.

In previous chapters we have seen how CDV is accumulated, e.g. in an access network and after crossing tandem of queues. The ITU-T 371, [1], recommends that the UPC/NPC should accommodate the effect of maximum CDV allowed on ATM connections. It also recommends that when allocating resources, the network should take into account the worst case traffic passing through the UPC/NPC in order to avoid impairments with other connections.

However the ITU-T371 gives freedom to the network operators to search for trade-offs between UPC/NPC complexity, worst case traffic and optimization of resources in the network. We have shown that for a DBR connection the worst case traffic is a periodic on/off source whose on period consists on a burst of back-to-back cells. This traffic pattern produces much greater levels of congestion than those obtained by the superposition of cells emitted periodically and affected by a certain random jitter. If the network is dimensioned for DBR traffic without bursts, but the users actually emit WCT, the Quality of Service of the network can be seriously degraded. This situation would be really pessimist since accounting for WCT can lead to low utilization factors, see subsection 5.5. Furthermore, anybody can argue that not all connections will produce WCT patterns. On the other hand, if we dimension the buffers to account only for periodic traffic (e.g. with the $ND/D/1$ or with the $M/D/1$ queue system) and some sources (not all) send WCT, we will have some degree of congestion. We saw in chapter 6, figure 6.6, the buffer size demanded by intermediate cases of WCT sources multiplexed with CBR sources.

The ITU-T 371 proposes traffic shaping mechanisms as a solution to ensure traffic conformance maintaining QoS objectives. The idea behind these mechanisms is to alter the traffic characteristics of VCCs or VPCs but at the same time meeting QoS objectives and conformance to the UPC. Examples of these mechanisms are cell spacing, leaky bucket shaping, dual-leaky bucket or scheduling algorithms.

In this chapter we study how it is possible to allocate DBR sources taking into account the CDV tolerances required in the traffic contract. If we do not take into account CDV or better said if we can consider that the CDV is negligible, simple resource allocation schemes can be used. In section 8.1 the concept of negligible CDV is defined as is intro-

duced in the ITU-T 371 and COST 242. Section 8.2 is dedicated to describe the several traffic shaping mechanisms.

One of the conclusions of the experiments performed in the EXPLOIT Testbed about CDV, was that high PCR connections were the ones that with a high CDV tolerance defined. could cause more congestion if they sent WCT, see chapter 4 section 5.5. Thus, one could conclude that the sources that have to be shaped are the high PCR connections. while the buffer is designed to absorb the low PCR connections independently of the CDV tolerances they have declared. We study the trade-offs in the design of a CAC combined with a Traffic-Shaping mechanism in section 8.4 considering only the shaping of high PCR connections.

Several scheduling algorithms proposed in the literature are compared through simulation in section 8.6. These schemes provide isolation among connections at the same time that provide guaranteed services in terms of end-to-end delay, end-to-end jitter and throughput. The schemes studied are Weight-Fair-Queueing and Virtual Spacing that are compared with the FIFO discipline used up to now.

8.1 The concept of negligible CDV.

A definition of negligible CDV can be found in [1]. [4] and [29]. A cell stream is considered to have negligible CDV if it can be assumed that cells arrive following a Poisson process of the same intensity. It is thus a conservative assumption. Then multiplexer dimensioning and admission control can be based on the M/D/1 queue system. As it is stated in [4] and [29]:

> an ATM multiplexer with FIFO discipline. service rate c and buffer capacity B, is assigned a nominal capacity of $c' = \rho c$ such that if the sum of the rates multiplexed is less than c' and have negligible CDV, then the cell loss ratio will be less than the probability of exceeding queue length B in the M/D/1 queue system with load ρ.

The advantage of considering the M/D/1 queue stays in that if a stream is "better" than another is it will have better multiplexing performance. The criteria to have a traffic better than Poisson are given in [29]. Summarizing these conditions we can say that the superposition of traffics that individually are better than Poisson remains better than Poisson. If a stream better than Poisson is multiplexed with Poisson interfering traffic, it remains better than Poisson. Finally, replacing the Poisson interfering traffic for a traffic that is better than Poisson improves the quality of the output.

Therefore. if a connection has negligible CDV at the ingress of the network, a simple admission control strategy can be applied at each node. Now. it remains to specify how to ensure negligible CDV on a connection. To do that, shaping mechanisms are proposed to restore the nominal cell interval.

8.2 Traffic shaping mechanisms.

Traffic shaping is a mechanism that alters the traffic characteristics to achieve better network efficiency at the same time that conformance of cells is ensured. The ITU-T371, [1], specifies that the network operator has the following options to choose:

- *Not shaping*: then dimension the network to accommodate any flow of conforming cells at the network ingress and at the same time ensure conformance at the network egress without any shaping function.
- *Shaping*:
  - Dimension the network to accommodate any flow of conforming cells at the network ingress and apply output shaping to meet conformance at the egress network.
  - Shape the traffic at the ingress network, allocate resources according to this traffic and insure conformance tests at the egress network.

Mechanisms to shape traffic could be peak cell rate reduction, burst length limiting, reduction of CDV by spacing cells or queue service disciplines.

As examples of shaping mechanisms that can be found in the literature we mention:

- *Leaky bucket shaping*: a user implements a leaky bucket on his connection to ensure that cells are not discarded by the GCRA. A study on the impact of leaky bucket controlled sources can be found in [85] and [30].
- *The Spacer-controller*: proposed by Boyer (e.g. see [28]), the spacer-controller behaves as a leaky bucket with the difference that provides a buffer to queue cells. Thus, the spacer-controller enforces a minimal spacing of cells according to the negotiated peak bit rate.
- *Dual Cell Spacing*: proposed by M. Ritter in [80], the Dual Cell Spacer takes into account the PCR as well as the SCR/BT. It has also been used to study the aggregation of VCCs in a VPC, see [81]. The dual spacer is equivalent to two shaping devices operating in series. The first one enforces the cell stream to the SCR and the IBT and the second one to the minimum inter-cell time.
- *Scheduling algorithms: Weight Fair Queueing mechanisms*: these mechanism appeared firstly as an attempt to provide fairness in the amount of resources provided to the users in packet networks. A fair queueing mechanism serves sessions in proportion to an established service share. Therefore in conjunction with a leaky bucket policy a fair queueing mechanism provides maximum delay guarantees in packet networks. In a following subsection we will see some of these mechanisms: General Processor Sharing (GPS) and Packet-by-Packet General Processor Sharing (PGPS), [74], the Virtual Clock, [94], the Virtual Spacing (VS), [82] also called Self-Clock Fair Queueing (SCFQ) in [48].

Resource allocation methods taking into account leaky bucket regulators have in the last years been paid attention. *Elwalid et al* in [41] proposes a framework based on leaky bucket regulated traffic. To each regulated traffic is allocated network resources (e.g. bandwidth and buffer resources) independently of other sources. They consider that the departing traffic of the regulator is an On/Off periodic process and find the effective bandwidth for the lossless multiplexing. In the case of having a loss system, the statistical multiplexing gain can be computed from the effective bandwidth and the lossless effective bandwidth. For the computing of the effective bandwidth they use the Chernoff Bound and large deviation approximations.

A. Gravey et al. in [50] and [63], also address the problem of resource allocation using worst case traffic. They identify a burst loss framework where the buffer size is too small to accommodate burst of cells (cells are lost when the instantaneous load exceeds 1) and a burst delay framework where the burst size is smaller than the buffer size (cells are buffered when the instantaneous load exceeds 1). In the burst loss framework, the traffic must be shaped at PCR since there are only enough buffer to cope with the cell scale congestion. In the case of the burst delay framework, the assumption of shaping traffic is less restrictive since there are buffer space to cope with the burst scale, however the models proposed, by simplicity, assume shaped traffic.

J. Mignault et al in [64] compare the method proposed by *Elwalid et al* with an allocation method proposed earlier in [85]. This method is based on the Benes bound and gives an upper bound on the probability of exceeding a queue level r in an infinite buffer system and using a fluid approximation. They show that Elwalid method is conservative and in a worst case resource allocation framework does not reach the maximum multiplexing gain given a set of network resources.

8.3 A simple model to study the spacing of WCT sources.

We are interested in modelize a traffic shaper able to handle periodic sources. For simplicity we will consider the case of homogeneous periodic sources that may send periodic bursts of back-to-back cells. The model consists on modifying the $N\ WCT/D/1$ to a queue system able to release cells after a constant time n times the service unit: the $N\ WCT/nD/1$. Bazanowski already studies in [25] the spacing of periodic cells through the $N\ D/nD/1$.

First of all, we consider the $N\ WCT/nD/1$ queue system. The system works as follows: periodic homogeneous bursts of back-to-back cells arrive to a queue system with infinite buffer and constant service time equal to D units of time, being the time unit the correspondent to transmit a cell. The phase of the bursts are uniformly distributed over the period. For stability: $\frac{NbD}{T} < 1$.

Taking into account these consideration we can use equation (5.1) of section 5.2 to calculate the Virtual Waiting Time distribution:

$$P\{V_0 > x\} = \sum_{t=1}^{T} \sum_{i=1}^{N} P\{\phi(t) = x \mid \phi(u) < x,\ t < u \leq T, B_{-t} = i\} \cdot$$
$$\cdot P\{\phi(u) < x,\ t < u \leq T \mid B_{-t} = i\} \cdot P\{B_{-t} = i\} \qquad (8.1)$$
$$= \sum_{t=1}^{T} \sum_{i=1}^{N} A(t, t+x, i) \cdot L(t, i) \cdot B(t, i)$$

To solve this formula we must take into account that each cell of the bursts spends D units of time in the server. Therefore:

- A(t,t+x,i) will be for the three cases:

$$\begin{cases} \frac{1}{(tD)^i} q_{tD}^{(i)}(t+x) & if\ t < b, \quad i \leq N \\ \sum_{j=0}^{i} \binom{i}{j} \frac{(t-b+1)^j}{t^i D^{i-j}} q_{(b-1)D}^{(i-j)}(t+x-bjD) & if\ b \leq t < T-b, \quad i \leq N \\ \sum_{j=0}^{N} \binom{N}{j} \frac{(t-b+1)^j}{t^N D^{N-j}} q_{(b-1)D}^{(N-j)}(t+x-bjD) & if\ T-b \leq t < T, \quad i = N \end{cases} \qquad (8.2$$

- Term B(t,i) has a binomial distribution: $\binom{N}{i} (\frac{t}{T})^i \cdot (1 - \frac{t}{T})^{N-i}$

- Term L(t,i) would be: $(1 - \frac{(N-i) \cdot bD}{T-t})$

To calculate the number in the system, we have to calculate $L_t = \lceil \frac{V_t}{D} \rceil$ and the number in the queue $B_t = [L_t - 1]^+$. For instance the distribution of the number in the queue would be:

$$\begin{cases} Pr\{B_0 = 0\} = \sum_{i=0}^{D} Pr\{V_0 = i\} \\ \\ Pr\{B_0 = x\} = \sum_{i=1}^{D} Pr\{V_0 = Dx + i\} \quad for\ x > 0 \end{cases} \qquad (8.3$$

We want to apply this model to calculate the delay in a shaper able to shape homogeneous WCT. The shaper works as follows: homogeneous periodic sources send bursts of back-to-back cells that pass through the shaper. Each connection is characterized by its period T and its burst size b. The shaper enforces a minimum distance of D slot units per cell. Therefore we can apply this model to the spacing of WCT connections on a VP. We can observe that the first cell from a burst that arrives will be served immediately only if the idle period is equal or greater than the spacing time and will be delayed a time equal to the remaining of the spacing time. We can see that the delay of these cells will be the Virtual Time of the $N-1$ bursts. So the delay in the spacer will be given by the Virtual Waiting Time distribution of the $(N-1)$ $WCT/nD/1$ queue system. The special case of $b = 1$ is studied in [25], where the cell delay experimented in the shaper

is calculated making use of the $N\ D/nD/1$.

Garcia and Ritter make use in [81] of the $N\ D/nD/1$ queue system to determine traffic parameters for the aggregation of VCCs carrying CBR streams on a VPC. They propose as a first solution to enforce a declared PCR larger than the aggregated cell rate $\frac{1}{T_{VP}}$. They chose the spacing period to fulfil delay constrains in the shaper and compute traffic parameters as the Burst Tolerance (BT) on the VPC. Since this method does not provide good BT tolerances for low rate connections, they propose a second method based on the dual spacer. The dual spacer, [80], is equivalent to two shaping devices operating in series. The first one enforces the cell stream to the SCR and BT and the second one to the PCR. With this method they provide a BT reduced respect to the first solution to form a VPC.

8.4 Trade-offs in the design of a CAC combined with a Traffic-Shaping mechanism.

We have mentioned that a possible solution to the problems that may cause the multiplexing of WCT sources in the buffer occupancy is the use of traffic shaping techniques. The Spacer-Controller, [28], the Leaky Bucket shaping, [4], or the Dual Spacer, [80] or scheduling algorithms are different proposals to shape traffic. These mechanisms shape the traffic so that they can guarantee the absence of bursts in CBR traffic that enters the network. In this case, the CAC would accept connections according to the number of sources that can be allocated to the channel capacity, without taking into account WCT conditions. These connections would be considered as having negligible CDV.

Most of the actually traffic shaping mechanisms in ATM switches perform leaky bucket shaping or spacing. What a switch offers is the possibility of shaping to a limited number of connections per port. Depending on the switch technology some offers more traffic shaping facilities than others do. This makes us think that due to the large number of connections that are possible in ATM interfaces these shapers are usually complex and expensive devices. Ideally all the connections should be shaped to conform their traffic contract. Therefore if we can dimension the network to handle some part of the traffic without being shaped and we shape the most dangerous connections (those ones that might pass larger bursts without being punished by the UPC function) the problem would be solved.

We can simplify the shaping device by limiting the number of VCs/VPs that can be managed, N, to a small number. If we assume that the CDV-T introduced by the access network can be bounded by a certain value τ independently of the values PEI, we observe that the connections that can produce larger bursts are high-speed connections. This idea was also observed in the experiments performed in the EXPLOIT Testbed. If we come back to figure 5.13, we can see that low bit rate connections do not accumulate enough CDV to be above the WCT curve that defines a negligible CDV considering worst case assumptions. However the high rate connections accumulated enough CDV so that the CDV tolerance defined might be above the WCT curve.

Then, we consider a simplified shaping device which only shapes those VCs/VPs whose PCR is larger than a certain value PCR_{min} (or conversely, those connections whose Peak Emission Interval, is smaller than a certain value T_{max}), see figure 8.1. By fixing $T_{max} = N$ we shall be able to shape all the connections with $PEI \leq T_{max}$, where N is the maximum number of sources that can be shaped.

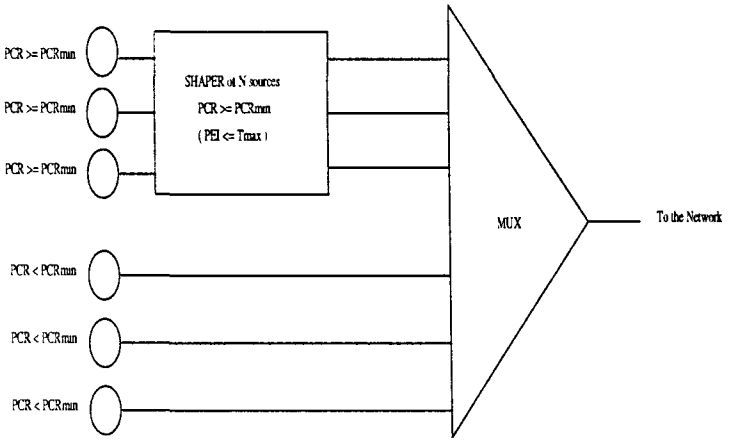


Figure 8.1: Shaper in the multiplexor only applied to high bit rate connections.

In such shaping device, connections with $PEI \leq T_{max}$ (or $PCR \geq PCR_{min}$) pass through the shaper, and when they enter in the network they behave as CBR sources (which can be analyzed with the nD/D/1 queue, see [87]). Connections with $PEI > T_{max}$ would not be shaped and WCT might enter the network.

If we use this shaping device. a worst case situation regarding network congestion would be that none of the sources connected to the network are shaped and all of them have declared a traffic contract given by the pair $(PEI, CDV\text{-}T) = (T_{max}+1, CDV\text{-}T)$. That means that all the sources could send WCT according to $B_s = \lfloor 1 + \frac{CDV-T}{T_{max}} \rfloor$.

In figure 8.2. we plot the admissible load of this worst case situation compatible with a CLR 10^{-9} in a switch with buffer length of 128 for different values of CDV-T, (τ_s) and load ρ. To obtain this figure we have used the formulas developed in section 5.2, based on the Benes approach. assuming that the traffic consists of periodic bursts of back-to-back cells of length $B_s = \lfloor 1 + \frac{\tau_s}{T_{max}} \rfloor$ uniformly distributed.

From this figure. we see that even for small values of T_{max} the admissible load can be high for reasonable values of τ. For example. if $N = T_{max} = 40$ and $\tau_s = 750\ \mu s$ the admissible load is over 0.85. Of course, for τ_s higher, the admissible load decreases for the same N since larger non shaped bursts can be accepted by the CAC. We have to remark that this would be the admissible load under worst case conditions. This means that this would be the worst admissible load that could be achieved.

For further illustration we can construct table 8.1. In this table we show the required

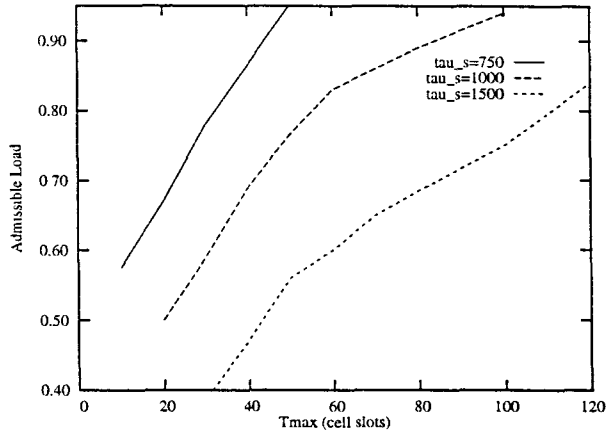


Figure 8.2: Admissible load for a buffer size of 128 and CLR=10^{-9}.

	τ_s=750 μs	τ_s=1000 μs	τ_s=1500 μs
$\rho = 0.7$	21	42	83
$\rho = 0.8$	30	55	110
$\rho = 0.85$	37	65	123

Table 8.1: Values of T_{max} for several values of ρ and τ_s.

values of T_{max} for achieving an admissible load ρ (ρ = 0.7, 0.8 and 0.85) for different values of τ_s (τ_s = 750, 1000 and 1500 μs).

The former example assumes an extreme worst case condition since none of the sources passes the shaper. If we consider that part of the traffic passes the shaper, we can bound the network congestion probability considering the traffic that passes through the shaper as Poisson traffic with $B_s = 1$ and the rest of the traffic as Poisson traffic with $B_s > 1$ (given by (T_{max}, τ_s)).

8.5 Performance evaluation of the CAC combined with a traffic-shaper.

From results obtained in subsection 6.3.2, we can derive a CAC procedure working in conjunction with a shaping device which can obtain high network use and can guarantee the absence of network congestion. The considered shaping mechanism is able to handle $N = T_{max}$ connections. For simplicity, we assume that the access network introduces a CDV-T which can be bounded by τ_s for all values of PEI. Under this assumption, the shaper should shape the connections with $PEI \leq T_{max}$.

Given a buffer length and CLR requirements, the CAC computes off line the acceptance region given by the pairs $(\rho_{shp}, \rho_{n-shp})$. The connections which pass through the shaper ($PEI \leq T_{max}$) are modeled as Poisson traffic with load ρ_{shp}. For the connections which

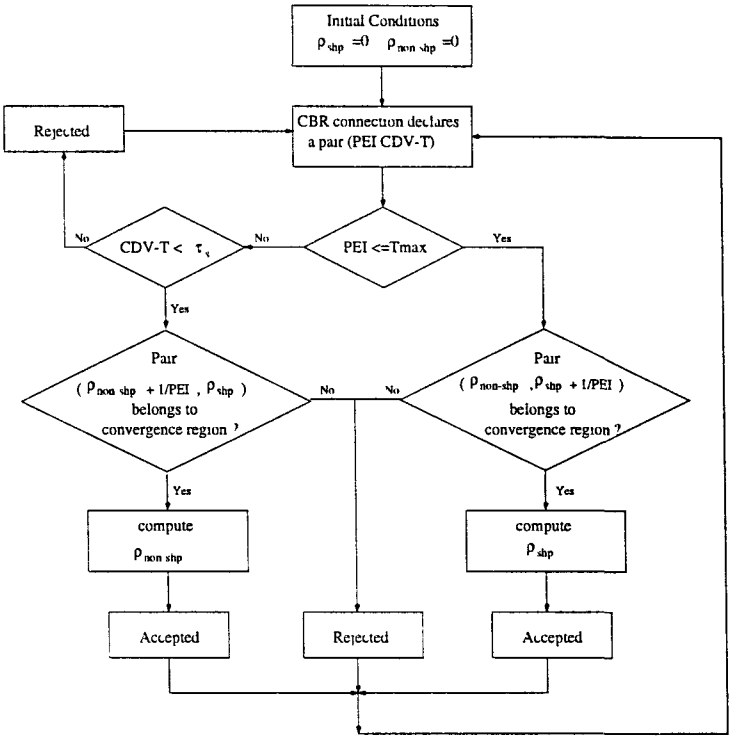


Figure 8.3: CAC algorithm flowchart.

are not shaped, those ones that have declared a $PEI > T_{max}$ and a CDV-T $< \tau_s$, we assume a worst case model consisting of batches of size $B_s = \lfloor 1 + \frac{\tau_s}{T_{max}} \rfloor$ following a Poisson distribution with load ρ_{n-shp}. An analitycal model for the multiplexing of these two kinds of traffic is given in subsection 6.3.2. In this models we use Poisson traffic of parameter ρ_{shp} and batches Poisson distributed of parameter ρ_{n-shp}.

At connection set-up of new CBR connections, the CAC procedure is the following (see figure 8.3):

- A new CBR connection declares the pair (PEI,CDV-T).
- If the new connection declares a $PEI \leq T_{max}$ and the pair $(\rho_{shp} + \frac{1}{PEI}, \rho_{n-shp})$ belongs to the acceptance region, the new connection is accepted. Otherwise the new connection is rejected. In case that the new connection is accepted, we update the value of ρ_{shp}
- If the new connection declares a $PEI > T_{max}$ and CDV-T $< \tau_s$ and the pair $(\rho_{shp}, \rho_{n-shp} + \frac{1}{PEI})$ is in the acceptance region, the new connection is accepted. Otherwise it is rejected. In case that the new connection is accepted, we update the value of ρ_{n-shp}

In figure 8.4, the CAC curve is calculated for a multiplexer with a buffer size of 128 cell slots and a shaper of 80 connections. The Traffic Contract specifies a cell loss ratio of 10^{-9}. Using the M/D/1 queue as a model for multiplexing dimensioning, (see [4]), we

observe that the CAC can allocate CBR connections whose total load, ρ_{shp}, rounds 0.93 for the CLR and buffer length required.

If connections with $PEI > 80$ are also multiplexed, the CDV-T declared by these connections would allow them to send bursts of size B_s. Fixing the maximum CDV-T, τ_s, figure 8.4 shows the allowance region for τ_s varying from 200 to 2000 μsec. The maximum burst size (B_s) has been chosen as the maximum allowed by the network: $B_s = \lfloor 1 + \frac{\tau_s}{T_{max}} \rfloor$. We observe that as τ_s grows, the admissible load of the non-shaped connections, ρ_{n-shp}, decreases considerably.

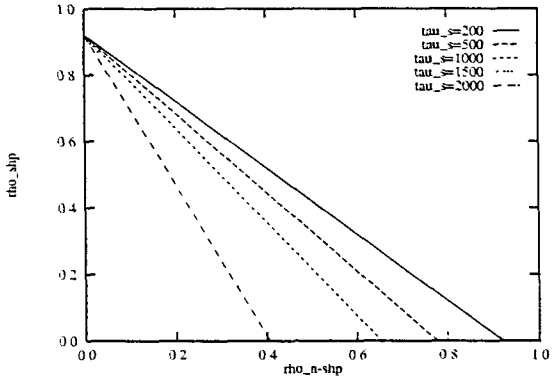


Figure 8.4: CAC curve for T_{max}=80. Buffer Size=128 cell slots, CLR=10^{-9}

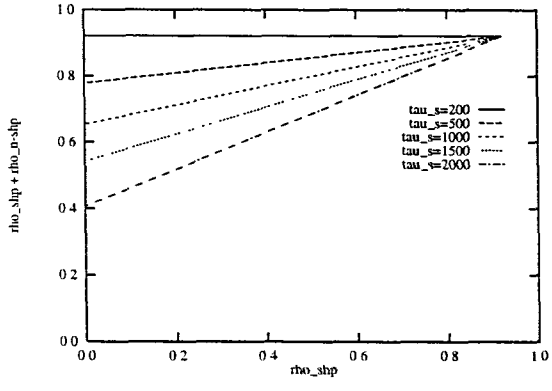


Figure 8.5: Shaped traffic versus all the traffic with the former QoS conditions

In figure 8.5, we show the total admissible load in the network. We have plotted $\rho_{shp}-\rho_{n-shp}$ versus ρ_{shp} for several τ_s in the same network conditions as in figure 8.4. As an example, if τ_s is 1000 μsec, we can guarantee a total load between 0.65 and 0.93 if we use a shaper able to handle 80 connections. Shaping only 150 connections, see figure 8.6, we can guarantee in worst case conditions that for a CDV declared of 1 msec the admissible load will almost be higher than 0.8.

8.6 Scheduling algorithms.

Different queueing disciplines have been proposed to provide guaranteed service to packet networks in terms of end-to-end delay and end-to-end delay jitter bounds, throughput and cell loss guarantees. The main objectives in a scheduling algorithm providing guaranteed service are: isolating sessions from each other, avoiding the influence of misbehavior of one traffic session over the others, sharing the resources available (statistical multiplexing) in the network, protecting the network from load fluctuations and protecting the network from best-effort traffic (ATM capabilities with preventive and reactive control mechanisms can share the resources in an integrated network).

One of the most well known mechanisms is the first-in-first-out (FIFO) discipline. FIFO is a simple discipline that provides acceptable cell loss and delay performance and exploits

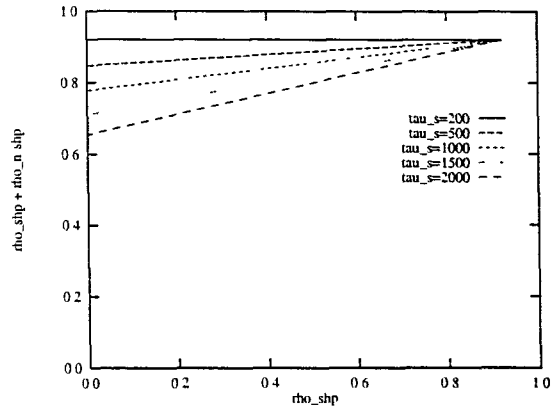


Figure 8.6: Shaped traffic versus all the traffic with T_{max}=150 connections.

multiplexing gain among sessions. However, FIFO does not provide isolation among the sessions. The effect of different scheduling disciplines in an Integrated Services Packet Network for real-time applications is studied in [34]. When a burst of packets passes through a queue shared with others sessions, the FIFO algorithm splits the delay evenly among the sessions producing an increasing delay in the sessions that conform to their traffic contract.

Real-time applications, where the receiver has to play-back the data, are very sensitive to delay. Packets that arrive before the play-back point must be buffered and used later to recover the signal. Packets with higher delay are useless in the recovering of the signal. As H. Zhang explains in [96], a service discipline must be efficient, protective, flexible and simple. Efficiency is defined as that discipline that can meet the same end-to-end performance guarantees under a heavier load of guaranteed service traffic. A service discipline must protect well behaving connections from other connections or from load fluctuations. The service discipline must be able to support all kind of applications with different traffic characteristics and performance requirements, and finally it must be easily implemented.

L. Zhang, in [94], describes a new traffic control algorithm called Virtual Clock that provides both a guaranteed throughput and average delay bound for each session, but as [82] and [74] show, Virtual Clock fails when sessions produce large burst of data, since the algorithm may punish the session a long time after it became active.

Other scheduling algorithms for providing guaranteed traffic and bounded delay have been proposed in [82], [37], [74], and [48]. All of them are work-conserving and are based in a sorted priority time stamped mechanism. We here describe briefly two of them, Packet-by-packet General Processor Sharing and Self-Clock Fair Queueing and compare them with FIFO.

Parekh and Gallager propose a fluid flow fair queueing scheme called General Processor Sharing (GPS) and derive bounds for the end-to-end delay when the input traffic conforms to a leaky bucket algorithm. This algorithm serves sessions according to es-

tablished service shares, independently of the load of the sessions. They also describe a packet-based version, first proposed by [37] as Weighted Fair Queueing (WFQ), called Packet-by-packet General Processor Sharing (PGPS). PGPS timestamps every packet evaluating the behavior of the fluid scheme (GPS). For that, Parekh and Gallager define a time function named *virtual time* that represents the progress of work in the GPS system and that has an increase rate of time equal to the rate at which backlogged sessions receive service. Therefore, the order can only be determined for the packets presents at the end of the service time. Based on this function, the PGPS scheme timestamps each packet and serves them in increasing order of timestamps.

But, as Golestani states in [48], PGPS may be rather complex to apply to a high rate network since it is based in the fluid GPS system and it is necessary to keep track of the sessions backlogged at any time to evaluate the *virtual time* in the equivalent GPS system. In ATM the number of connections active can be of hundreds, being the time provided to calculate the timestamp of the order of the packet transmission (on a link of 150 Mbs it would be around 2.73 μs). To avoid this problem, Roberts and Golestani propose in [82] and [48] an alternative algorithm simpler to implement in an ATM network and that behaves almost as the PGPS scheme. Roberts calls this scheme Virtual Spacing (VS) and Golestani Self-Clock Fair Queueing (SCFQ). In fact VS is the ATM version of SCFQ. They propose to use the service tag of the packet receiving service at that time instead of a virtual time function that simulates a fluid scheme. With this scheme, keeping track of sessions backlogged at any instant is avoided, and so the calculation of the timestamp is faster.

Although, PGPS or WFQ and VS or SCFQ have been proposed for high-speed networks as queueing schemes, an exhaustive study of these schemes in an ATM network has not been performed yet. Also, both schemes have been studied using deterministic traffic constrains (hard guarantees), [74], [49]. Works using stochastic traffic constrains (soft guarantees) have just begun to appear, ([93] and [95]).

8.6.1 Weighted Fair Queueing - Packet by Packet Generalized Processor Sharing

Weighted Fair Queueing (WFQ) or Packet by Packet Generalized Processor Sharing (PGPS) are rate-based control schemes where a Quality of Service is characterized by means of average or worst-case end-to-end delay and guaranteed throughput given source's traffic parameters such as mean bit rate, peak bit rate or burstiness. PGPS is a time stamped-based algorithm calculated through its fluid version, the Generalized Processor Sharing (GPS). Therefore, we will first explain the main characteristics of the GPS scheme, [74]. The GPS server is work conserving and operates at fixed rate C. The algorithm is defined by positive real numbers ϕ_1, ϕ_2, ..., ϕ_N. If $S_i(\tau, t)$ is the amount of session i traffic served in an interval $(\tau, t]$, then a GPS is defined as:

$$\frac{S_i(\tau,t)}{S_j(\tau,t)} \geq \frac{\phi_i}{\phi_j} \qquad j = 1, 2, \ldots N \tag{8.4}$$

for any session i that is backlogged in the interval $(\tau, t]$. Summing over all sessions we

can get:

$$S_i(\tau,t) \geq \frac{\phi_i}{\sum_j \phi_j}(t-\tau)C \tag{8.5}$$

If this equation holds then session i has a guaranteed throughput g_i:

$$g_i = \frac{\phi_i}{\sum_j \phi_j}C \tag{8.6}$$

With these properties the GPS scheme can guarantee a throughput r_i which does not depend on the demands of the other traffic sessions if $r_i \leq g_i$. Since GPS guarantees a clearing backlog rate g_i, the delay will be bounded.

For the implementation of the PGPS, Parekh and Gallager define a *Virtual Time* function, $V(t)$, that can be interpreted as a measure of the system progress. Considering that the set of backlogged sessions is B_j in an interval (t_{j-1}, t_j), the rate of change of $V(t)$ is the following:

$$\begin{cases} V(0) = 0 \\ V(t_j) = V(t_{j-1}) + \frac{\tau}{\sum_{j \in B_j} \phi_j} & \tau \leq t_j - t_{j-1}, \quad j = 2,3,.. \end{cases} \tag{8.7}$$

Now PGPS will stamp each packet with the following algorithm: if the k^{th} packet of session i arrives at time a_i^k and its length is L_i^k, we define the virtual finishing time. F_i^k as:

$$\begin{cases} F_i^0 = 0 \\ F_i^k = max(F_i^{k-1}, V(a_i^k)) + \frac{L_i^k}{\phi_i} \end{cases} \tag{8.8}$$

Packets are time stamped with the virtual finishing time and served in increasing order of timestamp. When a busy period ends (the server is free of packets), the algorithm reinitializes. A special case of the PGPS or WFQ is the Rate-Proportional Processor Sharing (RPPS), where the weight ϕ_i is allocated proportionally to the bandwidth required by the connection. Under these assumptions, in [75], deterministic bounds for the end-to-end delay are calculated. If K is the number of hops traversed by a connection. L_{max} is the maximum packet size, and the traffic conforms to a (σ, ρ) model, where σ and ρ represent respectively the burstiness and the long term bounding rate of the connection, the worst case end-to-end delay (D_i) and the worst case end-to-end delay jitter (J_i) are bounded by the following formulas:

$$D_i \leq \frac{\sigma_i + (K-1)\cdot L_{max}}{\rho_i} + \sum_{k=1}^{K} \frac{L_{max}}{C_k} \tag{8.9}$$

$$J_i \leq \frac{\sigma_i + (K-1)\cdot L_{max}}{\rho_i} \tag{8.10}$$

8.6.2 Virtual Spacing - Self-Clock Fair Queueing

In PGPS whenever a session i is backlogged or is cleared from the queue the slope of the *Virtual Time* function $V(t)$ changes, making necessary to recalculate the function. In ATM where the number of connections on a link may be high there could be a breakpoint to recalculate the virtual time every few slots. Another problem arisen by the PGPS scheme would be the necessity to store and up-date the number of backlogged sessions at any time. Self-Clock Fair Queueing (SCFQ) ([48], [49]) also called Virtual Spacing ([82]) tries to avoid the complexity presented by these problems defining a virtual time not depending on a simulated system (e.g. as the GPS). For this scheme, the *Virtual Time* is defined as the service tag of the packet being served at that time. The algorithm is the following:

$$\begin{cases} F_i^0 = 0 \\ F_i^k = max(F_i^{k-1}, F_l^j) + \frac{L_i^k}{\phi_i} \end{cases} \tag{8.11}$$

where F_l^j is the virtual finishing time of the connection being served at that time. The packets are served in increasing order of timestamp as in PGPS. Now with the same notation as before, assuming RPPS (Rate-Proportional Processor Sharing) networks and introducing N_k as the number of connections on a link, the worst case end-to-end delay (D_i) and the worst case end-to-end delay jitter (J_i) are the following, ([49]):

$$D_i \leq \frac{\sigma_i + (K-1) \cdot L_{max}}{\rho_i} + \sum_{k=1}^{K} (N_k - 1) \cdot \frac{L_{max}}{C_k} \tag{8.12}$$

$$J_i \leq \frac{\sigma_i + (K-1) \cdot L_{max}}{\rho_i} \tag{8.13}$$

An important property of Virtual Spacing is that the term F_l^j (time stamp of the current cell being served) is only involved in the calculation of the time stamp of a cell of a non-back-logged connection. For the cells of back-logged connections always the time stamp is calculated as its last time stamp plus $\frac{L_i^k}{\phi_i}$. This property eases the implementation of the sorting algorithm.

8.6.3 Scheduling Algorithms in ATM

Implementing PGPS in ATM where high speeds prevail means high calculation complexity to derive the *Virtual Time* function and store the active sessions at any time. PGPS means that not only a sorting queue algorithm is needed to schedule waiting times, but also to recognize which connections are active at any time. This can mean a very time consuming task if the number of sources is very high. However a Virtual Spacing algorithm has only to sort arriving cells to a memory as it is proposed in [4]. Therefore, a Virtual Spacing algorithm seems more promising as a scheduling algorithm for ATM networks.

Considering that the weight ϕ_i is allocated proportionally to the bandwidth, ρ_i, required by the connection (RPPS) and that the packet size is constant, the WFQ and VS algorithms would be as follows:

$$WFQ \quad \begin{cases} F_i^0 = 0 \\ F_i^k = max(F_i^{k-1}, Virtual\ \ Time) + \frac{1}{\rho_i} \end{cases} \tag{8.14}$$

$$VS \quad \begin{cases} F_i^0 = 0 \\ F_i^k = max(F_i^{k-1}, Spacing\ \ Time) + \frac{1}{\rho_i} \end{cases} \tag{8.15}$$

Studies of the behavior of deterministic traffic constrain show that the previous bounds can be rather pessimistic. For only one node and CBR sources, equations (8.9) and (8.12), have the following expressions: $D_i \leq (T_i+1)$ for the PGPS discipline and $D_i \leq (T_i+N-1)$ for the SCFQ discipline, where T_i is the period of the target connections and N the number of connections on the link. These bounds are very tight as can be seen in the following example. Assume a 150 Mb/s link loaded with two types of connections of 2 Mb/s and 25Mb/s. If we have one source of 25 Mb/s (load of 0.166) and 52 sources of 2Mb/s (load of 0.693) to achieve a total load of 0.86, the worst-case delay calculated with the PGPS bound is 7 slots (19 μsec) for the 25 Mb/s source and 76 slots (207 μsec) for a 2 Mb/s source. The worst-case delay calculated with the SCFQ scheme would be 58 slots (158 μsec) for the 25 Mb/s source and 127 slots (346 μsec) for a 2 Mb/s source. The fact that the SCFQ delay bound depends on the number of sources implies high differences with the PGPS bounds when low speed sources are multiplexed.

We have plotted in figures 8.7 and 8.8 the delay observed in a multiplexer simulating the above examples. As we can observe, the delays are no so tight. The curves have been simulated with a 95 % confidence interval. In table 8.2, we compare the delay introduced by three scheduling algorithms in the simulations for a quantile of 10^{-6} with the deterministic constrains. As it can be observed, the low bit rate connections accumulate more delay in PGPS and SCFQ than the high bit rate ones, however, we can also observe that the deterministic constrains are very tight. The FIFO constrain has been calculated using a formula derived in [35]: $D_i \leq \frac{\sum \sigma_i}{1-\sum \rho_i}$.

For medium-high rate connections as could be video streams these schemes can work well. The problem may be for low bit rate streams as telephone connections. Solutions proposed for this problem are to group the telephone streams in a VP and find traffic descriptors for this stream. Then a scheduling as VS may be applied to the VP.

	25 Mb/s connection			2 Mb/s connection		
	FIFO	PGPS	SCFQ	FIFO	PGPS	SCFQ
Deterministic Constrains μsec	1014	19	207	1014	158	346
Simulations μsec	46.3	19	35.4	46.3	51.8	54.5

Table 8.2: Comparison between simulated delays and deterministic delay constrains.

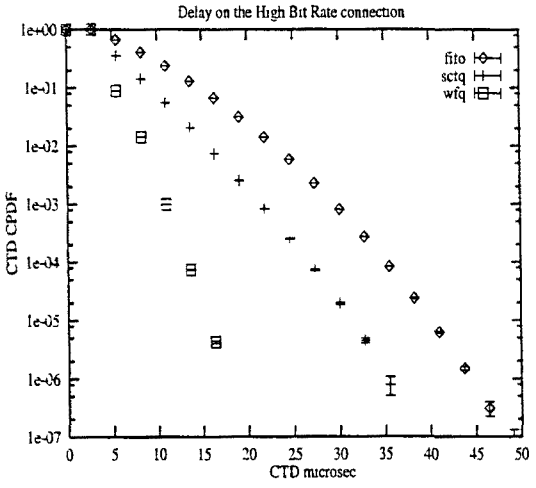


Figure 8.7: Delay on the 25 Mb/s connection.

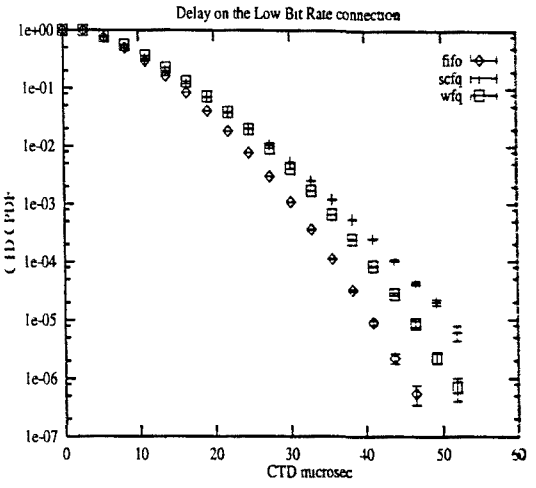


Figure 8.8: Delay on the 2 Mb/s connection.

A similar analysis to study the multiplexing of CBR sources with "*misbehaving*" sources modeled as WCT (a periodic On/Off source transmitting a burst of back-to-back cells in the On period and remaining silent in the Off period) can be performed. As an example we multiplex 6 10 Mb/s CBR sources (load of 0.4) with 6 WCT sources with a burst of 8 cells and a total period of 120 slots (load of 0.4). The worst-case delay calculated with the PGPS and SCFP bounds respectively give 16 slots (43.6 μsec) and 26 slots (71 μsec) for a CBR source, while for a WCT source it gives 121 slots (330 μsec) and 132 slots (360 μsec).

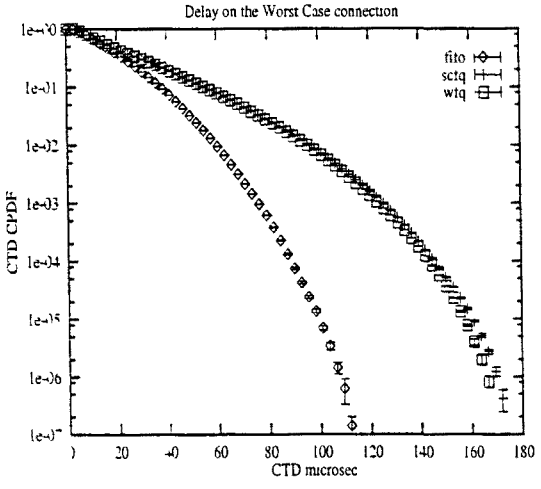


Figure 8.9: Delay on the WCT source.

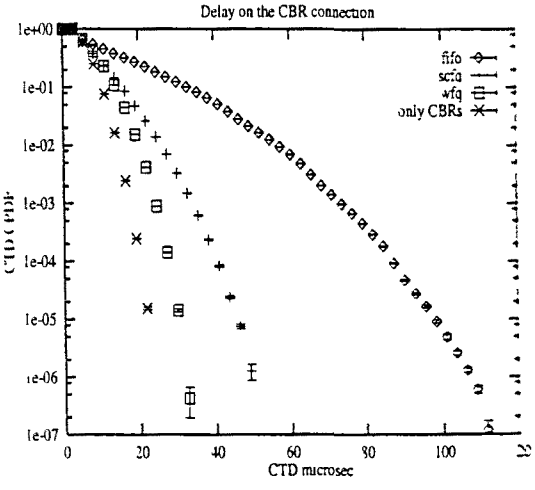


Figure 8.10: Delay on the 10 Mb/s CBR source.

With these examples we can observe two phenomenons: first that the SCFQ bounds are tighter than the PGPS ones. The reason is that the SCFQ bound depends on the number of sources at each link. If this number is high, the bound foresees high delay values. The second is that the delay of the low bit rate sources can be high. This effect has to be into

account for low bit rate sources with real time constrains (e.g. telephone connections) in order to dimension the play-back buffers. We have to note that the size of the buffer as it is stated in [4] depends on the end-to-end delay.

From the simulations we conclude that FIFO behaves as expected, it does not isolate connections from misbehaving sources. As can be seen, the delay introduced by the WCT sources is shared among all the connections including the CBR ones. WFQ and SCFQ isolate CBR from WCT connections, punishing the WCT sources while keeping the delay constrains in the CBR sources. However, WFQ and SCFQ introduce more delay in low rate connections that in high rate connections. As can be observed in the delay formulas, the end-to-end delay depends on the rate of the source. The lower the rate the higher the delay constrain is. WFQ behaves and produces lower delays than SCFQ, however, the implementation of WFQ can be cumbersome. WFQ needs to keep track of backlogged connections to calculate the Virtual Time function, while SCFQ needs only the timestamp of the current served connection. We can also observe in figure 8.10 the difference between multiplexing only 12 CBRs at 10 Mb/s (FIFO discipline) with the multiplexing of 6 CBRs and 6 WCT. The curve "only CBRs" represents the delay produced by the other CBRs connections. We see that WFQ and SCFQ isolate quite well while FIFO introduces high delays in the CBR due to the interaction of the WCT sources.

The problem of multiplexing bursty traffic (e.g. WCT) with real time traffic (e.g. CBR traffic) is discussed in [4]. The situation is the following: consider that bursty traffic is multiplexed with CBR streams and the scheduling is Virtual Spacing (the same example would apply to WFQ) with the same rate parameter r. Since the CBR streams have small delays, their cells will arrive to a non-backlogged queue and their timestamp will be derived using the Spacing Time (the timestamp of the actual served cell plus $\frac{1}{r}$). The bursty cells, however, that in general will belong to backlogged streams will have their timestamps spaced $\frac{1}{r}$ respect their last timestamp. Since all these timestamps will be greater than actual Spacing Time, a CBR arriving cell will have to wait for service behind one cell of every backlogged stream.

The proposed solution to this situation is to give priority to non-backlogged streams modifying the algorithm (8.15) as follows:

$$VS \qquad \begin{cases} F_i^0 = 0 \\ F_i^k = max(F_i^{k-1} + \frac{1}{\rho_i}, Spacing\ \ Time) \end{cases} \tag{8.16}$$

We have simulated the profits of this improvement on the VS algorithm in figures 8.11 and 8.12. The simulations consist on the same set up that figures 8.9 and 8.10. As can be observed, the CBR source improves its delay distribution while the WCT source is punished more.

8.7 Conclusions and comments.

In order to be able to apply a CAC based on the concept of negligible CDV, we must smooth the traffic. Solutions proposed in the literature are traffic shaping techniques as

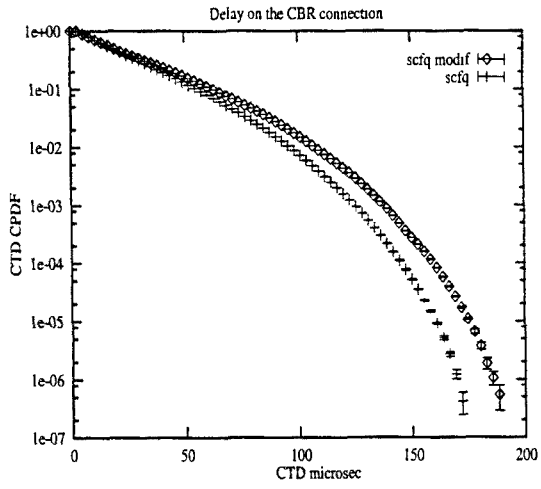


Figure 8.11: Delay on the WCT source for the VS and its modification.

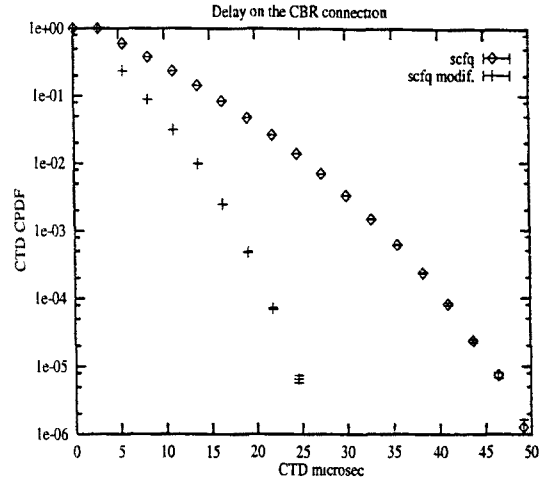


Figure 8.12: Delay on the 10 Mb/s CBR source for the VS and its modification.

leaky bucket shaping, the spacer-controller or scheduling algorithms.

We have first studied a simple method to space WCT sources in a VP. The model consists of the N $WCT/nD/1$ queue system. In this system, cells that arrive to the queue bring a work of D units of time. Therefore, all the cells of the VP will be spaced D units of time.

Actually most of the switches in the market provide only leaky bucket shaping facilities for a limited number of connections. Therefore, it is not so unrealistic to make use of a CAC procedure combined with a simplified shaper (i.e. the number of connections which are shaped is limited) if it is possible to bound the length of the bursts (the related CDV-T) in the network. The shaper might only shape the connections that have a higher impact in buffer dimensioning (e.g. high bit rate connections). For those that do not pass through the shaper, a CDV bound can be computed by using WCT models. We have shown that for reasonable CDV tolerances the admissible load, which can be achieved, is high. If the CAC takes into account that certain part of the traffic is shaped and the CDV-T is bounded, high network utilization can be achieved and at the same time network congestion can be avoided.

Other solutions to smooth the traffic are scheduling algorithms. Several proposals for packet networks have recently appeared in the literature. The PGPS proposed by Parekh and Gallager (also called Weight Fair Queueing, WFQ) and the SCFQ proposed by Golestani provide guaranteed traffic and bounded delay. These techniques are work-conserving and are based on a sorted priority time stamped mechanism. Roberts calls the ATM version of the SCFQ Virtual Spacing. We have compared PGPS and VS respect to a FIFO scheduling to study how these disciplines behave in an ATM environment, and specifically in an scenario with WCT connections. We have observed that PGPS and VS are able to isolate CBR sources from WCT offering them bounded delay constrains, being the best PGPS. However, in an ATM environment PGPS may be difficult to implement

since is based on a GPS simulation that must keep track of all the backlogged connections at any time. In the other side, as Roberts states, VS may be easier to implement, since its sorting algorithm is most efficient: it only depends on the time stamp of the current cell being served. Moreover, this time stamp only applies in the algorithm when a cell of a connection that is not backlogged arrives to the queue. The time stamp of the rest of the cells belonging to backlogged connections only depends on the time stamp of the last cell of that connection plus its cell rate factor.

The contributions of this chapter to the thesis are: an extension of the homogeneous WCT periodic model to study the spacing of WCT in a VP. A study of the trade offs of a CAC combined with a traffic shaping mechanism. This part comes from the observation that although the ideal situation could be to shape all the traffic, most of the actual switches provide a limited shaping capacity. Therefore, we could gain from the knowledge of how to dimension the resources of the network with the tools provided. Furthermore, I have used one of the analytical models studied in chapter 6. Finally, the study through simulation of the scheduling mechanisms was carried out to complete the thesis. For instance, we can observe how the WCT affects the CBR connections in terms of delays and how these mechanisms allow to isolate CBR from WCT connections.

Chapter 9

General conclusions and comments.

ATM layer Traffic Control Functions require the knowledge of the parameters declared on the Traffic Contract to operate efficiently. For a DBR connection two of the most important parameters are the Peak Cell Rate (PCR) and the Cell Delay Variation (CDV) tolerance. The PCR accounts for the maximum cell rate the connection will emit while the CDV tolerance accounts for the variation that the traffic suffers respect to negotiated PCR.

The importance of these parameters stands for resource allocation schemes and consequently for Connection Admission Control. The impact of the declared pair PCR - CDV tolerance in these schemes resides in how the cell conformance algorithm works. For instance, given a PCR - CDV tolerance pair, there are many traffic patterns that fulfil this declaration, and conversely, given a traffic pattern there are a set of pairs PCR - CDV tolerances that can be declared. Thus, the importance first in studying how the jitter or CDV affects a connection through its route and secondly studying how the pair of parameters PCR - CDV tolerance affect resource allocation schemes.

Following the studies done in the COST 242, we performed a set of experiments using real Testbeds in order to see what components and how these components could introduced jitter in a connection. These Testbeds correspond to an ATM access network (the BAF Testbed) and an experimental network (the EXPLOIT Testbed) that was connected to the ATM Pan European pilot network.

In the BAF access network we observed how the MAC protocol introduced CDV in the source under test. The architecture of the BAF access network makes use of an ATM-based Passive Optical Network with a tree topology and the access to the shared medium in the APON is arbitrated by a Global FIFO multiple access protocol operating on a request/permit basis. Most of the CDV introduced in the access network was due to the operating of the protocol: the algorithm that issues the permits. The order of the

measured CDV taking into account that only 4 of the 81 T_b interfaces were loaded was around 25 μs.

The project EXPLOIT allowed us to measure the accumulation of CDV on a tandem of queues. The experiment set-up consisted of three network segments: the EXPLOIT Testbed segment, the Swiss ATM pilot segment and the EPFL-TCOM Laboratory network segment. Each segment was composed by several switch architectures (sometimes from different manufactures) and different interfaces (e.g. STM-1 or E3 line interfaces). The work conditions at each network segment also were different. In the EXPLOIT segment and in the EPFL-TCOM segment the traffic conditions were fully controlled by ourselves, while we had a PCR allocation CAC at the Swiss ATM pilot consisting of an ATM contract based on a maximum VP PCR of 24 Mb/s and a CDV-T of 367 μs. Our traffic crossed two nodes in this segment (Zurich and Bern) and was multiplexed with traffic coming from other projects. During our experiments, we checked that the total load in this segment was always lower than 0.4 and seeing the shape of our curves we concluded that consisted of bursty traffic. To compute the CDV, we used the approaches described in the COST 242, see [4]. The results show that the accumulated CDV was of the order of hundreds of μs. The CDV was an increasing function of the number M of multiplexing stages although this increase was moderate when M increased.

From my experience in the experimental Testbeds, I could conclude that it is difficult to get definitive conclusions due to the different assumptions defined during the experiments and due to the Testbed limitations. I must say that in my work in the projects I had all the help from all the partners from these projects and that the limitations were due to the high costs that suppose to buy equipment and the fact that most of the equipment of the platforms at that time were prototypes.

My intention including these results were to show two examples of how CDV could be introduced in a real ATM network and the magnitude of these values. Of course, this values and the conclusions that can be taken are totally dependent on the set up configuration. For instance, from the simulations of the BAF protocol we know that loading the system with the maximum number of T_b interfaces, 81, the CDV introduced is higher than using the 4 T_b interfaces of the BAF demonstrator. Therefore, I would mention that the order of CDV measured taking into account our network configuration was of hundreds of ms. but that it could be more.

But more than the results from themselves, I would like to emphasize the knowledge that supposes to work with real equipment. On one hand one has all the knowledge provided in the literature (mainly from analytical models and simulations) and has to develop a set of skills to apply this knowledge in a real environment. In our case, most of the knowledge came in the BAF project from the work realized during the definition and design of the access network by the several partners of the project, and in the case of the EXPLOIT project from analytical models and simulations that appeared in the literature, most of them recorded in [4]. As an example I could mention the difficulty in loading a whole network with the available equipment in the Testbed. To my knowledge, this situation is common to all the projects due to the high costs of these equipments. Therefore, in this field, my conclusions more than the results, are all the experience and

techniques that we used to really obtain the results: duplicated traffic, recirculation techniques, phase-moving CBR properties, recycling techniques or VP policing were some of the techniques utilized. We had to define a set of rules when using all these techniques to have a fair load configuration in terms of minimizing secondary effects, as could be the presence of correlated traffic. Furthermore, some of the equipment in the Testbeds were prototypes. We were in contact with some of the manufactures of these equipment to report them failures and giving them feedback about their equipments during the two years working in EXPLOIT. I think that manufactures can gain from the experience of experimental platforms.

The different experimental results presented in this thesis together with some previous results and works that were not included in this thesis can be found in: [9], [10], [11], [12], [13], [14], [24], [33], [56] [57], [58] and [72].

One of the main objectives of this thesis is to study the impact of WCT sources in a multiplexor through the development of a set of analytical models. Due to the tolerance declared, the GCRA mechanism allows to pass many different traffic patterns that can have very different effects on network congestion. For example, the GCRA can declare as conforming a traffic pattern consisting of cells emitted at their contracted PCR that have pick up certain jitter due to the network conditions or even can declare as conforming a connection that emits back-to-back cells in periodic bursts. The last kind of pattern is called *Worst Case Traffic* (WCT) since requires the greatest amount of resources from the network.

To study the impact of these sources, we have make used of the Benes approach that expresses the *Virtual Waiting Time* of a G/G/1 queue system. This approach has been extensively used in ATM to analyze queue systems at cell level (e.g. M/D/1, ND/D/1, M+D/D/1 or $\sum D_k$/D/1 queue systems) and at burst level (e.g. superposition of on/off sources exponentially distributed). I chose the Benes approach instead of other tools as could be the ballot theorem since the Benes approach seemed me more intuitive to solve WCT systems.

Up to this moment, there were several approaches to study the superposition of periodic WCT sources based on bounds as for example to use batch arrivals or an approximation based on the ND/D/1 presented in [78]. In this thesis we have studied the superposition of N independent WCT sources through an exact analytical model. Each WCT source produces a periodic stream of cells, of period T, with the following pattern: It emits a constant number, b, of back-to-back cells and then it remains silent during a constant time $T - b$. The time slots in which each source becomes active are uniformly and independently distributed within the period. The WCT periodic model calculated at discrete-time is exact. A fluid flow version has also been presented. However, the resulting integrals of this model must be solved numerically, for example, using the Simpson's method.

The results show the great impact that have this sources in the multiplexor. To keep the same QoS, the multiplexor must increase roughly by b its buffer size. If we keep the same buffer size, the admissible load in the multiplexer drops to keep the same QoS. Several

examples showing the effects of WCT are described in chapter 5.

The model presents two limitations: the assumption of homogeneity and the high computation times when the number of sources and the burst size are high. To overcome these problems we have developed in this thesis a set of new models. From the literature, we know that a M/D/1 queue model can be used as an upper bound of the queue length in the case of having sources whose rate is small respect with the channel capacity and the load is not high. We can use the same arguments for WCT and make use of a Poisson process with burst arrivals to upper bound periodic WCT sources with low rates. Since not always all the traffic has to be WCT connections, we have studied a set of intermediate cases as the multiplexing of WCT sources with CBR sources, or Poisson background traffic.

As a main conclusion of this chapter I would say that the models in which all the traffic (WCT or WCT+CBR) are modeled as Poisson processes are exact but they have the drawback of having loss of numerical accuracy due of cancelation of terms when the load and the buffer size are high. In the case of upper bounding the WCT by a batch arrival the problem can be solved using a polynomial representation. This solution was first utilized and presented by Virtamo in [91] for the M/D/1 queue system. We have here extended this method for the cases of superposition of Poisson traffic with batch Poisson arrivals.

We have solved approximations for the intermediate cases in which we have periodic WCT and periodic CBR traffic, or we have periodic WCT and Poisson traffic. These models are not exact, but may provide a good estimation of the queue length distribution under certain traffic circumstances explained in chapter 6.

We have also derived closed-form formulas for the queue length distributions in a discrete-time M-stage tree queueing network loaded with periodic traffic sources (for instance WCT sources, or batch sources or CBR sources). In this type of network, the queues can be grouped in M groups or stages. I thought that although this kind of topology is limited, its solution could lead to give us more insight to more general queue systems such as queues in tandem with cross traffic. This was my first intention. However, even for only two queues in tandem with cross traffic the system become really complex to solve. Tandem queues have been solved in the literature through restrictions such as independence assumptions or product-form solutions. Here, I have tried to give a study of the WCT considering a topology more general than a single queue system. A topology that can be analysed and is still tractable is a tree topology without interfering traffic. In any case, I still think that the tree topology can help us to understand several traffic problems arisen in queue systems. For instance, we can study how the balance of loads and the number of queues in a previous stage affects the root queue. We can see how the root queue is smooth if the traffic comes from a previous stage. The contribution in this thesis is the solution of the tree with two stages. The extension to M stages is easily done following a previous work of Morrison presented in [66]. I have to mention that the tree topology of the BAF system was based on a passive optical tree and there were no buffering in the splitting points. That was the reason to implement a MAC protocol to share the medium.

To complete this chapter, we thought interesting to find the average number of cells in the queue and the average delay in the queue. To calculate this values, we applied an equivalent priority model proposed by Modiano *et al* in [65]. Having this equivalent model in mind is possible to find average values in any queue of the tree topology.

Using periodic sources we have observed that the WCT is slightly smooth having two queueing stages. The root queue length increases if the load is balanced among the previous stages and is and increasing function of the number of queues in the previous stage. In the results of this chapter, we present through several configurations how to study a network with balance-unbalance loads.

With all these models our intention was two-folded: study the impact on a multiplexor in terms of buffer occupancy or admissible load and to have a set of models that we could apply in studies concerning WCT sources. The different WCT models presented in this thesis can be found in: [18], [19], [21], [22] [23], [44] and [45].

The study of WCT is important for several factors. First, as the ITU-T 371 states, the network can not relay in the fact that the CBR application is going to send periodic traffic. The only thing that the network knows is that the CBR application can send traffic according to the traffic contract (idem for VBR applications). A misbehaving customer can try to take advantage of the UPC function and send bursts of cells fulfilling the traffic contract but demanding most resources that expected. Adaptation ATM Layer (AAL) schemes can also introduce WCT in a natural way.

Therefore, if the network has to ensure QoS objectives, these possible network conditions must be taken into account (e.g. the CAC might take into account WCT, the buffers might be greater or traffic shaping might be used to smooth the WCT bursts). Since the DBR transfer capability is intended to carry real time CBR connections, to increase the buffer size can lead to high cell delays that can affect for instance the play-back buffers dimensioning. Mechanisms to ensure a negligible CDV definition could be traffic shaping mechanisms: for instance a spacer or scheduling mechanisms.

In the last chapter of this thesis we study some of these mechanisms. A simple method to space WCT cells in a VP could be studied through the $NWCT/nD/1$ queue system. In this model, we assume that each cell of the burst brings D units of work. Thus, each cell will be spaced D units at the output of the spacer. Of course, the spacing period is done per VP.

Other solution studied in this chapter consists of shaping only part of the traffic while at the same time we dimension the multiplexor taking into account a WCT assumption for the rest of the traffic. The idea is the following: most of the actual switches can shape a limited number of connections. Therefore, we could limit the number of connections that pass through the shaper to N (those VCs/VPs whose PCR is larger than a certain value PCR_{min}). The rest of the connections would enter the multiplexor directly and thus, could send bursts at peak cell rate. However, to send high bursts of cells, these low PCR connections would have to declare high CDV tolerances. To study such mechanism we need a model in which the traffic that passes the shaper is considered as CBR and

the traffic that does not pass the shaper is considered as WCT. I have chosen the Poisson traffic multiplexed with batches Poisson distributed to analyze this problem since this model assumes the worst of the situations, is exact and can be computed easily (allows a polynomial representation). We show how with a limited number of shaped connections and in conjunction with the CAC, the admissible load achieved can be high.

Finally, we have compared through simulation several of the scheduling mechanisms presented in the literature. The main objectives in a scheduling algorithm providing guaranteed service are to isolate sessions from each other, avoiding the influence of misbehavior of one traffic session over the others, to share the resources available (statistical multiplexing) in the network, to protect the network from load fluctuations and to protect the network from best-effort traffic. Among the algorithms proposed, we have compared the FIFO system with the Packet-by-packet General Processor Sharing (PGPS also called Weight Fair Queueing, WFQ) and with the Virtual Spacing (VS, also called Self-Clock Fair Queueing, SCFQ), see [74] and [82].

PGPS, [74], serves sessions according to established service shares, independently of the load of the other sessions making use of a Virtual time function that represents the progress of work in the GPS system and that has an increase rate of time equal to the rate at which backlogged sessions receive service. Based on this function, the PGPS scheme timestamps each packet and serves them in increasing order of timestamps.

A second method to timestamp sessions is the Virtual Spacing algorithm, [82]. VS uses the service tag of the packet receiving service at that time instead of a virtual time function that simulates a fluid scheme. With this scheme, keeping track of sessions backlogged at any instant is avoided, and so the calculation of the timestamp is faster.

From the simulations we conclude that FIFO behaves as expected, it does not isolate connections from misbehaving sources. As can be seen, the delay introduced by the WCT sources is shared among all the connections including the CBR ones. PGPS and VS isolate CBR from WCT connections, punishing the WCT sources while keeping the delay constrains in the CBR sources. However, PGPS and VS introduce more delay in low rate connections that in high rate connections. As can be observed in the delay formulas, the end-to-end delay depends on the rate of the source. The lower the rate the higher the delay constrain is. PGPS behaves and produces lower delays than VS, however, the implementation of PGPS can be cumbersome. PGPS needs to keep track of backlogged connections to calculate the Virtual Time function, while VS needs only the timestamp of the current served connection. Therefore, the conclusion is that both methods produce good isolation on CBR streams. However, VS seems easier to implement in a switch.

Finally I would like to justify some decisions about the writing of the thesis. The first one is that the mathematical analysis is included in the chapters and not in appendices. The reason is that the analytical models form a big part of the thesis. I think that part of a PhD is to acquire a methodology. Explaining step by step the analytical reasoning may show that I have get this methodology.

The second comment allude to the fact that I have referenced related work appeared in

the literature. The reason is that the thesis is formulated in the context of Traffic Control in ATM. Therefore, I first assume that the reader is familiar with this framework e.g. CAC, policing, GCRA, definition of PCR or CDV. I have tried to define those parameters that I have considered more important. I have also considered the related work of other authors. For instance analytical tools that I have used but that I have not proposed. In this way I have tried to clarify that I know how to work with these tools. On the other way I can not summarize all the work done in Traffic Control. I thought that a nice way of presenting the thesis was to brevely comment those works and referencing them for further consult if the reader was interested.

Acronyms

AAL	Adaptation ATM Layer
ABR	Available Bit Rate
ABT/DT	ATM Block Transfer with Delay Transmission
ABT/IT	ATM Block Transfer with Immediate Transmission
ADSL	Asymmetric Digital Subscriber Line
APON	ATM Passive Optical Network
ATC	ATM Transfer Capability
ATM	Asynchronous Transfer Mode
BAF	Broadband Access Facilities
B-ISDN	Broadband Integrated Service Digital Network
BT	Burst Tolerance
BT	Background Traffic
CAC	Connection Admission Control
CATV	Cable Antenna TeleVision
CBR	Constant Bit Rate
CDV	Cell Delay Variation
CDV-T	Cell Delay Variation tolerance
CL	Permit Class bit
CLP	Cell Loss Priority
CLR	Cell Loss Ratio
CPDF	Complementary Probability Distribution function
CTD	Cell Transfer Delay
DBR	Deterministic Bit Rate
ETB	EXPLOIT TestBed
EXPLOIT	EXPLOITation of an ATM Technology Testbed for Broadband Experiments and Applications
FIFO	First In First Out
FTTC	Fiber To The Curb
FTTH	Fiber To The Home
GCRA	Generic Cell Rate Algorithm
GFR	Guaranteed Frame Rate
GI	General Independent
GMDP	General Modulated Deterministic Process
GPS	General Processor Sharing
IBT	Intrinsic Burst Tolerance
ITU	International Standards Organization
MAC	Medium Access Control
MBS	Maximum Burst Size
MCR	Minimum Cell Rate
NRM	Network Resource Management
NNI	Network to Node Interface
NPC	Network Parameter Control
NT	Network Termination

OAM	Operation and Maintenance
ODN	Optical Distribution Network
OLT	Optical Line Terminator
ONU	Optical Network Unit
PCR	Peak Cell Rate
PDA	Permit Distribution Algorithm
PDH	Plesiochronous Digital Hierarchy
PEAN	Pan European ATM Pilot
PEI	Peak Emission Interval
PGPS	Packet by Packet General Processor Sharing
PON	Passive Optical Network
QL	Queue Length
QoS	Quality of Service
RB	Request Block
RM	Resource Management
RPPS	Rate Proportional Processor Sharing
SBR	Statistical Bit Rate
SCFQ	Self-clock Fair Queueing
SCR	Sustainable Cell Rate
SDH	Synchronous Digital Hierarchy
STM	Synchronous Transport Mode
TDM	Time Division Multiplexing
TUT	Traffic Under Test
UBR	Unspecified Bit Rate
UPC	Usage Parameter Control
UNI	User Network Interface
VBR	Variable Bit Rate
VC	Virtual Channel
VCC	Virtual Channel Connection
VP	Virtual Path
VPC	Virtual Path Connection
VS	Virtual Spacing
WCT	Worst Case Traffic
WFQ	Weight Fair Queueing

Bibliography

[1] ITU-T Recommendation I-371, "Traffic Control and Congestion Control in B-ISDN", Draft issue, Geneva, 1995.

[2] ITU-T Recommendation I-356, "B-ISDN ATM Layer Cell Transfer Performance", Draft issue, Geneva, 1995.

[3] ATM Forum, "ATM Forum Traffic Management Working Specification", Version 4.0, 1996

[4] J. Roberts, U. Mocci, and J. Virtamo, Eds., "Broadband Network Teletraffic - Final Report of Action COST 242". Springer Verlag, 1996.

[5] Wandel & Golterman, "ATM-100, ATM test tool. BN 3027 prototype, description and operating manual". 1994

[6] HP, "HP 75000 Broadband Series Test System Operation Manuals", Hewlett-Packard Company, IDACOM Telecommunications Operation, Edmonton, Alberta, Canada 1994

[7] Alcatel STR: A8640 operation manual. Switzerland. January 1995

[8] BAF deliverable *R2024_QMW_SW2_DS_P_019_b1*. "Description of Experiments, final version" June 1994

[9] BAF deliverable *R2024_CST_W42_DS_R_026_b1*. "Evaluation of Experiments", June 1994

[10] EXPLOIT Deliverable: R2061/EXP/SW3/DS/R/008/b1, "Interim Specification of the Experiments on Resource Management"

[11] EXPLOIT deliverable *R2061/EXPLOIT/SW3/DS/R/019/b1*, L. Jaussi, N. Mitrou, V. Nellas, A. Martinez, J.M. Barcelo, O.Casals, "Results on experiments on Resource Management using Test equipment", June 1994

[12] EXPLOIT deliverable *R2061/EXPLOIT/SW3/DS/R/029/b1*, L. Jaussi, N. Mitrou, V. Nellas, A. Martinez, J.M. Barcelo, O.Casals, "Results on experiments on Resource Management using Real Applications", December 1994

[13] EXPLOIT deliverable *R2061/EXPLOIT/SW3/DS/R/041/b1* L. Jaussi, N. Mitrou, V. Nellas, J.M. Barcelo. O.Casals. F. Aglan. "Specifications of experiments on Routing Strategies and Bandwidth Management", March 1995

[14] EXPLOIT deliverable *R2061/EXPLOIT/SW3/DS/R/047/b1*, L. Jaussi, N. Mitrou, V. Nellas, J.M. Barcelo, O.Casals, F. Aglan, "Final Report on Routing Strategies and Bandwidth Performance". December 1995

[15] EXPERT deliverable *AC094/EXPERT/WPG4/DS/R/P/004/A4*, "Initial Specification of Bandwidth Management". W. Meurisse et al, March 1996

[16] E. Aarstad. "A comment on Worst Case Traffic". COST 242 TD (93)

[17] T. Apel, C. Blondia, O. Casals, J. Garcia and K. Udhe, " Implementation and Performance Analysis of a Medium Access Control Protocol for an ATM Network", Proceedings Architecture and Protocols for High Speed Networks, pages 1-20, Dachstuhl Seminar 1993

[18] J.M. Barceló, J. García, O. Casals, "The Benes approach for studying the multiplexing of ATM traffic", Joint International Meeting Euro/Informs XV-XXXIV. OR/MS for the new Millennium, Barcelona, Spain, July 1997, Invited paper.

[19] J.M. Barceló, J. García, O. Casals, "Comparison of Models for the Multiplexing of Worst Case Traffic Sources", First Workshop on ATM Traffic Management. WATM'95 IFIP WG 6.2, Paris, December 1995.

[20] J.M. Barceló, J. García, O. Casals, "Worst Case Traffic models and their application to CAC", Interim Report UPC-DAC-1996-44, October 1996,

[21] J.M. Barceló, J. García, "Multiplexing periodic sources in a tree network of ATM multiplexers", Fourth International Conference on Broadband Communications (BC'98), Stuttgart. Germany, April 1998

[22] J.M. Barceló, J. García, O. Casals, "Worst Case Traffic in ATM tandem queues in a tree topology", Interim Report UPC-DAC-1997-13. March 1997, submitted to a Journal

[23] J.M. Barceló, J. García, "Average Waiting times in a tree network of ATM multiplexers" Interim Report UPC-DAC-1997-42, March 1997

[24] J.M. Barcelo, O.Casals, L. Jaussi, V. Nellas, N. Mitrou "Results of experiments on resource Management performed by EXPLOIT WP3.5 performed on the EXPLOIT Testbed" ATM Hot Topics on Traffic and Performance: from RACE to ACTS, Milan, Italy, June 1995

[25] Z. Bazanowski, U. Killat, J. Pitts, "The ND/nD/1 Model and its Application to the Spacing of ATM Cell", Procedings of the International Conference on Information Networks and Systems'94, Saint Petersburg, 1994

[26] B Bensaou, J. Guibert, J.W. Roberts, A. Simonian, "Performance of an ATM multiplexer queue in the fluid approximation using the Benes approach", Annals of the opr. Res., 49:137-160, 1994

[27] J. Boyer, A. Gravey, K. Sevilla "Resource Allocation for Worst Case Traffic in ATM Networks". ATM Traffic Management WATM'95. Paris, December 1995.

[28] P. Boyer, F.M. Guillemin, M.J. Servel and J. P. Coudreuse, "Spacing Cells at the ATM Network Entry Points". Special Issue of IEEE Network Magazine on Switching and Congestion Control in ATM Networks. September 1992.

[29] F. Brichet, L. Massoulie, J.W. Roberts, "Stochastic ordering and the notion of negligible CDV", ITC 15, Editors: V. Ramaswami and P.E. Wirth, Elsewier Science, 1997

[30] F. Brichet, J.W. Roberts, A. Simonian, "Performance of an ATM multiplexer fed by leaky bucket controlled sources", IFIP workshop on ATM Traffic Management, WATM'95. Paris. Dec. 1995

[31] J. P. Buzen. "Computational Algorithm for Closed Networks with Exponential Servers". Comm. Assoc. Comput. Mach., 1973, 16, 527-531.

[32] O. Casals, J. Garcia, C. Blondia, "A Medium Access Control Protocol for an ATM Access Network", Proc. of the 5-th Int. Conf. on Data Comm. Systems and their Performance. High Speed Networks, Ed. H. Perros and Y. Viniotis. North Carolina, USA, October 1993

[33] F. Cerdán, J.M. Barceló, J. García, O. Casals. "Protocolo de Control de Acceso al Medio para una Red de Acceso ATM", IV Jornadas Telecom I+D, Madrid, Nov 1994

[34] D.D Clark, S. Shenker, L. Zhang, "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanism". In ACM SIGCOMM'92, 1992

[35] R. L. Cruz, "A calculus for network delay, part I: network elements in isolation", IEEE Transactions on Information Theory, Vol 37, No 1: 114-131, January 1991

[36] R. L. Cruz, "A calculus for network delay, part II: network analysis", IEEE Transactions on Information Theory, Vol 37, No 1: 132-141, January 1991

[37] A. Demers, S. Keshav and S. Shenker, "Analysis and simulation of a Fair queueing algorithm", Journal of Internetworking: research and experience, 1, pp 3-26, 1990

[38] L. Dron, G. Ramamurthy, B. Sengupta, "Delay analysis of continuous Bit Rate Traffic over an ATM network", IEEE JSAC, Vol. 9, No 3, April 1991

[39] A. E. Eckberg, "The single server queue with periodic arrival process and deterministic service times", IEEE transactions on Communications, Vol COM-27, No 3, March 1979

[40] A.I. Elwalid, D. Mitra, "Effective Bandwidth of General Markovian Traffic Sources and Admission Control of High Speed Networks", IEEE/ACM Transactions on Networking, Vol 1, N-3, June 1993

[41] A.I. Elwalid, D. Mitra, R.H. Wentworth, "A new approach for allocating buffers and bandwidth to heterogeneous, regulated traffic in an ATM node", IEEE JSAC, Vol 13, No 6, August 1995

[42] W. Feller, "An Introduction to Probability theory and its Applications", Vol 2, Ed. John Willey.

[43] J. García, O. Casals, C. Blondia, F. Panken, "The Bundle-Spacer: a Cost Effective alternative for Traffic Shaping in ATM Networks", Proceedings of IFIP Conference on Local and Metropolitan Communication Systems, 4:55-62, 1993

[44] J. García, J.M. Barceló, O. Casals, "An Exact Model for the Multiplexing of Worst Case Traffic Sources", IFIP Conference on Performance on Computer Networks, PCN'95, Instambul, Turkey, October 1995.

[45] J. García, J.M. Barceló, O. Casals, "Multiplexing Worst Case Traffic sources in a slotted queue", COST 242 TD(94)

[46] R.J. Gibbens, F.P. Kelly, P.B. Key "A Decision-Theoretic Approach to Call Admission Control in ATM Networks", IEEE JSAC, Vol-13, N-6, August, 1995

[47] D. Ginsburg, "ATM solutions for enterprises internetworking", Ed. Addison-Wesley, 1996

[48] S.J. Golestani, "A Self-Clocked Fair Queueing Scheme for Broadband Applications", Proc. IEEE INFOCOM'94, Toronto, CA, June 1994, pp 636-646

[49] S.J. Golestani, "Network Delay Analysis of a Class of Fair Queueing Algorithms", IEEE JSAC, Vol 13, N-6, August 1995

[50] A. Gravey, J. Boyer, K. Sevilla, J. Mignault, "Resource allocation for worst case traffic in ATM networks", Performance Evaluation, 30(1997), 19-43

[51] R. Guerin, H. Ahmadi, M. Naghashineh, "Equivalent Capacity and its Application to Bandwidth Allocation in High Speed Networks", IEEE JSAC, Vol-9, N-7, pp 17.28, 1991

[52] R. Guerin, J. Heinanen, "UBR+ service Category Definition", ATM Forum contribution, ATM96-1598, December 1996

[53] Handell, Huber, Schroder, "ATM Networks, concepts, protocols and applications", Ed. Addison-Wesley, 2nd Ed., 1994

[54] J.Y. Hui, "Resource Allocation for Broadband Networks", IEEE JSAC, vol 6, N-9, December 1988, pp. 1598-1608

[55] P. Humblet, A. Bhargava, M. Hluichyi, "Ballot Theorem applied to the Transient Analysis of $nD/D/1$ queues", IEEE/ACM Transactions on Networking, Vol 1, No 1, February 1993.

[56] L. Jaussi, N. Mitrou, V. Nellas, A. Martinez, J.M. Barcelo, O.Casals, "First Results and analysis of Traffic Experiments performed by EXPLOIT WP3.2 on the ETB", Open Workshop on ATM Broadband Experiments, Basel, Switzerland, September 1994

[57] L. Jaussi, J.M. Barcelo, S. Louis, O. Casals, "CDV Impact on ATM Resource Management", Proceedings of the Opnet conference, Paris, February 1996.

[58] L. Jaussi, J.M. Barcelo, S. Louis, O. Casals, "Experimental Evaluation of CDV Impact on ATM Resource Management", European Transactions on Telecommunications, ETT Vol. 7, No 5, Sept./Oct. 1996

[59] F.P Kelly, "Effective Bandwidth at Multiplexing Queues" Queueing Systems, 9 (1991), 5-16

[60] U. Killat, editor, "Access to B-ISDN via PONs; ATM Communication in Practice", John Wiley & Sons Ltd & B.G. Teubner, 1996

[61] L. Kleinrock, "Queueing systems", Vol I, John Wiley and Sons, 1976

[62] H. Kroner, T. Renger, R. Knobling, "Performance Modeling of an Adaptative CAC Strategy for ATM Networks", In J. Labetouille and J.W. Roberts, editors, Proc. ITC-14, pp. 1077-1088, Elsevier Science, 1994.

[63] J. Mignault, A. Gravey, C. Rosemberg, "A survey of straightforward statistical multiplexing models for ATM networks" ATM Experts Workshop, Basel (1995), Telecommunications Systems J.5 (1996), 177-208

[64] J. Mignault, C. Rosemberg, A. Gravey, "A Reference Resource allocation Method for ATM Variable Bit Rate Services" Proceedings on the ICCC'97, pp 187-198, Cannes, France, Nov. 1997

[65] E. Modiano, J.E. Wieselthier, A. Ephemerides, "A simple analysis of average queueing delay in tree networks", IEEE Transactions on Information Theory, Vol. 42, No 2, March 1996

[66] J.A. Morrison, "A Combinatorial Lemma and its Application to Concentrating Trees of Discrete-Time Queues", The Bell System Technical Journal, Vol. 57, No 5, May-June 1978.

[67] J.A. Morrison, "Two Discrete-Time Queues in Tandem", IEEE Transactions on Communications, Vol COM-27, No 3, March 1979

[68] T. Murase, H. Suzuki, S. Sato, T. Takeuchi, "A Call Admission Control Scheme for ATM Networks using a simple Quality Estimate", IEEE JSAC, Vol-9, N-9, December 1991

[69] G. Niestegge, "The Leaky Bucket policing method in the ATM network", International Journal of Digital and Analog Communication Systems, Vol3, 187-197, 1990

[70] I. Norros, J W. Roberts, A. Simonian, J.T. Virtamo "The superposition of Variable Bit Rate Sources in an ATM multiplexer", IEEE JSAC., Vol. 9, No. 3, April. 1991.

[71] F. Panken, "Design and performance evaluation of multiple-access protocols for ATM-based passive optical networks", Doctoral Thesis. Thesis Publishers, Amsterdam, 1997

[72] F. Panken, J.M. Barcelo, B. Miah, S. Winstanley, "Investigations on delay and CDV in an ATM-Based optical Network", IEEE ATM'97 Workshop, Lisboa, Portugal May, 26-28, 1997

[73] A. Papoulis, "The Fourier Integral and its Applications", McGraw-Hill, 1962, New York.

[74] A. Parekh, R. Gallager, "A Generalizer Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single Node Case", Proc. IEEE INFOCOM'92, 1992

[75] A. Parekh, R. Gallager, "A generalizer processor sharing approach to flow control in integrated services networks: the multiple node case", INFOCOM '92, 1992

[76] A. P. Prudnikov, Y. A. Brychkov and O. I. Marichev, "Integrals and Series", Volume 1. Gordon and Breaach Science Publishers, New York.

[77] M Potts, "The EXPLOIT Testbed", European Transactions on Telecommunications, ETT Vol. 7, No 5. Sept./Oct. 1996

[78] G. Ramamurthy. R. S. Dighe, "Distributed Source Control: a Network Access Control for Integrated Broadband Packet Networks", IEEE JSAC, Vol 9, No 7, September 1991

[79] E.P. Rathgeb. "Modeling and performance comparison of policing mechanisms for ATM networks". IEEE JSAC, 9(3), April 1991

[80] M. Ritter. "Performance analysis of the dual cell spacer in ATM systems", IFIP 6th International Conference on High Performance Networking, Palma de Mallorca, Spain, Sep. 1995

[81] M. Ritter. J. Garcia, "Determination of Traffic Descriptors for VPs carrying Delay sensitive Traffic", Broadband Communications'96,

[82] J.W. Roberts. "Virtual Spacing for integrated high speed data and real time services", COST 242 TD(93). June 1993

[83] J.W. Roberts, "Virtual Spacing for flexible traffic control", Int. J. of Communications Systems, 7:307-318. 1994

[84] J.W. Roberts. "What ATM Transfer Capabilities for the B-ISDN ?". First Workshop on ATM Traffic Management WATM'95, Paris. December 1995.

[85] J.W. Roberts. B Bensaou, Y. Canetti, "A traffic control framework for high speed data transmission". IFIP Transactions C-15, Modeling and Performance Evaluation of ATM technology. North Holland, January 1993

[86] J.W. Roberts. F. Guillemin, "Jitter in ATM networks and its impact on peak rate enforcement", Performance Evaluation, 16:35-68, 1992

[87] J.W. Roberts. J. Virtamo, "The Superposition of Periodic Cell Arrival Streams in an ATM Multiplexer". IEEE Trans. on Comm., Vol. 39. No. 2, Feb. 1991.

[88] B. Sengupta. "A queue with superposition of arrival streams with an application to packet voice technology". PERFORMANCE'90, P.J.B. King, I. Mitrani and P.J. Pooley. Eds. The Netherlands. 1990

[89] R. Syski. "Introduction to Congestion Theory in Telephone Systems", Amsterdam, The Netherlands: Norh Holland, 1986, 2nd ed.

[90] O. Rose, M. Frater, "A comparison of models for VBR video traffic sources in B-ISDN". IFIP Transactions C-24. Broadband Communications II, North-Holland 1994

[91] J. Virtamo. "On the numerical evaluation of the distribution of unfinished work in an M/D 1 system", Electronics Letters, 31(7): 531-532. March 1995

[92] J. Virtamo. "Idle and Busy period distributions of an infinite capacity $N \cdot D/D/1$ queue". In Proceedings of the 14th International Teletraffic Congress- ITC 14. Antibes. Juan-les-Pins. France. 6-10 June. 1994. Volume 1, Elsevier Science B.V., 1994

[93] O. Yaron. M. Sidi "Generalized Processor Sharing Networks with Exponentially Bounded Burstiness Arrivals", Proc. IEEE INFOCOM'94. June 1994, June 1994, pp 628-634

[94] L. Zhang, "Virtual Clock: a new traffic control algorithm for packet switching", ACM Transactions on Computer Systems, 9(2):101-124. May 1991

[95] Z. Zhang. D. Towsley. J. Kurose. "Statistical Analysis of the Generalized Processor Sharing Scheduling Discipline". IEEE JSAC. Vol 13. N-6. August 1995

[96] Hui Zhang. "Service Disciplines for Guaranteed Performance Service in Packet Switching Networks". Proceedings of the IEEE, Vol. 83. No 10, October 1995