Studies on
Performance Analyses of Protocols
for ATM Networks

by Kenji KAWAHARA

Submitted in partial fulfillment of
the requirement for the degree of
DOCTOR of ENGINEERING
(Information and Computer Science)

at

OSAKA UNIVERSITY
Toyonaka, JAPAN

May 1996

Preface

Up to the present, most networks are dedicated to specific purposes like telephony, TV distribution, circuit-switched or packetized data transfer. The Broadband Integrated Services Digital Network (B-ISDN) concept allows us to handle such a kind of transfer services in a common network interface. The International Telecommunication Union Telecommunication (ITU-T) standardization sector has recommended the Asynchronous Transfer Mode (ATM) as the ground on which B-ISDN is to be built.

ATM technique contributes to traffic transmission of a wide range of bit rates, from the order of kbit/s up to about 155Mbit/s or 622Mbit/s by dividing such traffic into small fixed-size cells of 53 octets (bytes) in length, and those cells may be statistically multiplexed onto one virtual channel (VC) / virtual path (VP). Although the statistical multiplexing may achieve efficient utilization of ATM network resource, it should cause cell loss due to congestion in any network elements, e.g., ATM switches.

The first objective of this dissertation is to investigate some control schemes to reactively avoid congestion for Available Bit Rate (ABR) traffic, which is sensitive to cell loss. In Chapter 2, we focus on reactive feedback congestion control for the system which models the relation between sources and congested node. By using the network model, we provide an approximate analysis to evaluate the performance of the cell loss probability and the average waiting time in the system. Numerical results show that the reactive congestion control are very effective in avoiding congestion and in achieving the statistical gain. Furthermore, the binary congestion control algorithm with cell push-out scheme is shown to provide the best performance among the control algorithms treated in Chapter 2. In Chapter 3, we analyze the link-by-link backpressure congestion control by extending the approach, developed in Chapter 2, where we deal with the relation between the congested node and its upstream nodes instead of sources. It has been noted that, at the upstream node, cells not forwarded to the congested node are adversely affected due to head-of-line (HOL) blocking by backpressure. From numerical results, we observe that backpressure control with the cell push-out scheme can overcome this drawback and further improve the loss performance in terms of cells transmitted to the congested node than by plain backpressure

control, without increasing delay so much.

Even if the above congestion control is applied to the system, cell loss will occur. In a higher layer over ATM, a packet is discarded when one or mode cells are lost in that packet, and the destination node then requires its source to retransmit the corrupted packet. Therefore, once one of cells constituting a packet is lost, its subsequent cells of the corrupted packet result in wasting network resources. Thus, selectively discarding those cells will enable us to efficiently utilize network resources.

The second objective is to investigate the effectiveness of selective cell discard schemes as another congestion control in Chapter 4. To do so, we propose the exact analysis of the packet loss probability in the system applying these schemes without the retransmission of lost packets. We find from numerical results that Tail Dropping (TD) and Early Packet Discard (EPD) as selective cell discard schemes can improve packet loss probability, and TD achieves performance comparative to that of EPD with less complexity of implementation.

Supposing the case that congestion causes the loss of cells from Constant Bit Rate (CBR) traffic. This traffic is sensitive to delay as well as cell loss. As a way of recovering from loss, Forward Error Correction (FEC), which does not require any retransmission of lost cells, may be used for such delay-sensitive traffic.

The third objective of this dissertation is to evaluate the performance of FEC as error control in Chapter 5. FEC adds redundant cells to data ones, all of which are organized into a group called a block, and allows receivers to correct up to some lost cells without retransmission. We exactly analyze the distribution of the number of cells lost within a block and examine the benefits due to FEC by means of the analysis. Numerical results show that using a block of large size improves the block loss probability significantly, however, it can cause the receiver to wait longer until the last cell within a block comes up, and that the possible block size thus depends on some requirement in terms of delay time in each application.

Finally, the author would like to hope that the research in this dissertation will be helpful for further study in this field.

Acknowledgements

Completing my dissertation was more difficult than I imagined, but I have gotten much more help than I ever could have hoped for. Now I get to thank every one for their help.

First and foremost, I am deeply indebted to Professor Hideo Miyahara of Osaka University for his encouragement and assistance to award me Ph. D. in Osaka University.

Next, I wish to express my sincere appreciation to Professor Yuji Oie of Information Technology Center (ITC) at Nara Institute of Science and Technology (NAIST), my adviser, mentor and teacher when I was in both Sasebo National College of Technology and Kyushu Institute of Technology (K.I.T.), for his untiring support and great encouragement. He introduced the idea of this research to me. His suggestions and criticism have guided my study. In addition, he spent much time reading all drafts of this dissertation to ennoble it.

I would like to thank Professor Akihiro Hashimoto and Professor Tohru Kikuno of Osaka University for their support.

I would like to express my gratitude thank Associate Professor Masayuki Murata of Osaka University concerning the work in Chapters 2 and 3, and also thank Associate Professor Tetsuya Takine of Osaka University concerning the work in Chapters 4 and 5 for their encouragement and beneficial suggestions.

I wish to thank Miss Kazumi Kumazoe, who is now working for NEC Corporation, for her help and support of the work in Chapter 5, thank Mr. Kouichirou Kitajima, who is now working for Fuji XEROX, for his supporting the work in Chapter 4, and also thank students in Oie's Laboratory in K.I.T. and all members of Oie's Laboratory in NAIST for their encouragement.

I also would like to thank Associate Professor Mariko Gotoh, who is an English teacher of K.I.T., for her giving me the skill of English conversation and hearing, and for her supporting the practice of presentation at international conferences.

I often interrupted Professor Oie's precious and even private time to write and discuss papers. Special thanks go to his family for their understanding.

I wish to thank Professor Kunihiro Chihara, who is the director of ITC at NAIST, and other staffs for their encouragement and support.

Finally, my greatest appreciation goes to my father and my grandmother for their constant support and great encouragement. Further, I would like to express my gratitude to my mother for giving birth to me.

Kenji Kawahara
Information Technology Center, Nara Institute of Science and Technology
Ikoma-shi, Nara, in May, 1996

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 B-ISDN

Up to the present, most networks are dedicated to specific purposes like telephony, TV distribution, circuit-switched or packetized data transfer [HÄND 95]. Using pre-existing networks for new applications may lead to characteristic drawbacks, as such networks are not usually tailored to the needs of services that were unknown when networks were implemented. So data transfer over the telephone network is limited by a lack of bandwidth, flexibility and quality of analogue voice transmission equipment. Telephone networks were engineered for a constant bandwidth service, and using them for variable bit rate data traffic requires costly adaptation.

In 1984, the Plenary Assembly of the *CCITT* (International Telegraph and Telephone Consultative Committee, which has become the International Telecommunication Union Telecommunication standardization sector, *ITU-T*, as of the beginning of 1993) adopted the I series recommendations dealing with Integrated Services Digital Network (*ISDN*) matters. The CCITT stated that an ISDN is a network, in general evolving from a telephony IDN, that provides end-to-end digital connectivity to support a wide range of services, including voice and non-voice services, to which users have access by a limit set of standard multi-purpose user-network interfaces [ITUC 84]. One of ISDN standard interfaces was defined and called **basic access**, comprising two 64 kbit/s B channel and a 16 kbit/s signalling D channel. Another type of interface, the **primary rate access**, with a gross bit rate of about 1.5 Mbit/s or 2 Mbit/s, offers the flexibility to allocate high-speed H channels or mixture of B channels, H channels and a 64 kbit/s signalling channel.

This original ISDN is based on the digitized telephone network [ITUC 84] which is characterized by the 64 kbit/s channel. The 64 kbit/s ISDN is basically a circuit switched network, but it can offer access to packet-switched services. ISDNs are being implemented

in this decade. Their benefits for the user and network provider include:

- common user-network interface for access to a variety of services
- enhanced (out-of-band) signalling capabilities
- service integration
- provision of new and improved services

ISDNs can offer users an interface with capacity of up to about 2 Mbit/s. However, connection of Local Area Networks (LANs) or transmission of moving images with high resolution may require considerably higher bit rate. Consequently, the conception and realization of a broadband ISDN (*B-ISDN*) was desirable.

ITU-T Recommendation I.113 [ITUI 113] defines 'broadband' as 'a service or system requiring transmission channels capable of supporting rates greater than the primary rate'. Armbrüster and Rothamel [ARMB 90] tried to compile technical characteristics for major B-ISDN applications. The results are given in Table 1.1.

Table 1.1: Characteristics of broadband services

Service	Bit rate (Mbit/s)	Burstiness
Data transmission	1.5 to 130	1 - 50
Document transfer / retrieval	1.5 to 45	1 - 20
Video-conference / video telephony	1.5 to 130	1 - 5
TV distribution	30 to 130	1
HDTV distribution	130	1

Burstiness = peak bit rate/average bit rate

From this table, it is found that the bit rate available to a broadband user is up to hundreds of Mbit/s. The higher bit rate on 64 kbit/s based ISDN was realized by mixing B channel, H channel and D channel. However, This could not manage a bit rate gradually existing from 64 kbit/s telephony to hundreds of Mbit/s HDTV distribution. Furthermore, channel structures are fixed at subscription time and not dynamically changed. This deadlock was finally overcome by adopting an interface model based on a complete breakdown of its payload capacity into small pieces called *cells*, each of which can serve any purpose. They may be employed to carry information relating to any type of connections: this technique is referred to as the **Asynchronous Transfer Mode** (*ATM*).

1.2 ATM

1.2.1 B-ISDN based on ATM

The asynchronous transfer mode (ATM) is considered the ground on which B-ISDN is to be built [ITUI 121]. The term *transfer* comprises both transmission and switching aspects, and a transfer mode is thus a specific way of transmitting and switching information in a network.

In ATM, all information to be transferred is packed into fixed-size slots called cells. These cells have a 48 octet information field and a 5 octet header (see Fig. 1.1). Whereas the information field is available for the user, the header field carries information that pertains to the ATM layer functionality itself, mainly the identification of cells by means of a label. ATM uses a label field inside each cell header to define and recognize individual communications. In this respect, ATM resembles conventional packet transfer modes. Like packet switching techniques, ATM can provide a communication with a bit rate individually tailored to the actual need, including time-variant bit rates. In ATM based B-ISDN, a bit rate of about 155 Mbit/s or 622 Mbit/s will be typically offered to the user across the broadband user-network interface.

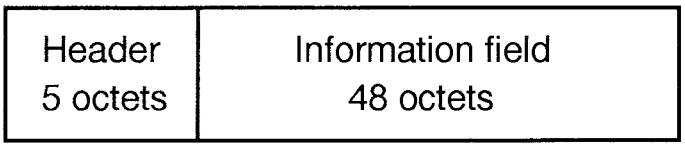


Figure 1.1: ATM cell structure

The term *asynchronous* refers to the fact that, in the context of multiplexed transmission, cells allocated to the same connection may exhibit an irregular recurrence pattern as they are filled according to the actual demand. The multiplexing and switching of cells are independent of the actual application. Thus, the same piece of equipments can handle a low bit rate connection as well as high bit rate connection, do it of stream or burst nature. Dynamic bandwidth allocation on demand with a fine degree of granularity is provided. The flexibility of the ATM-based B-ISDN network access resulting from the cell transport concept strongly supports the idea of a unique interface which can be employed by a variety of customers with quite different service needs.

ATM combines advantageous features of both connection- and packet-oriented techniques. The former requires only low overhead and processing, and, once a connection

is established, the transfer delay of the information is low and constant. The latter is much more flexible in terms of the bit rate assigned to individual (virtual) connections. ATM is a connection-oriented, hardware-controlled, low-overhead concept of virtual channels which have no flow control or error recovery.

1.2.2 ATM Layer Function

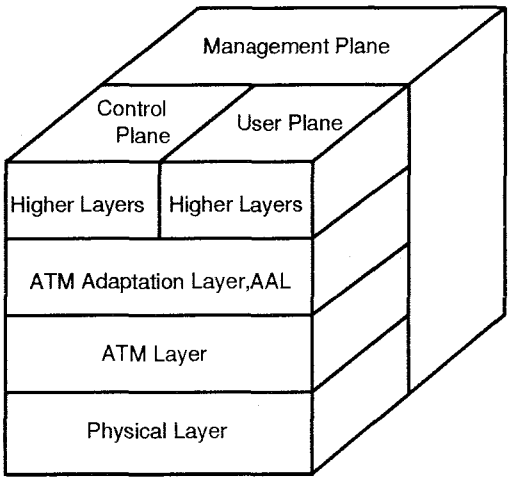


Figure 1.2: B-ISDN protocol reference model

Figure 1.2 shows the B-ISDN protocol reference model. Functions of each plane and layer in Fig. 1.2 are defined in [ITUI 321]. I will only explain the function of ATM layer based on cell header description in the remainder of this section and also show that of ATM Adaptation Layer (AAL) in Section 1.3 in order to clarify the motivation of this dissertation.

The ATM layer is the one above the physical layer. Its characteristic features are independent of the physical medium. Figure 1.3 depicts the cell header structure. Four functions of this layer have been identified as follows [HÄND 95].

- In the transmit direction, cells from individual Virtual Channel (VC) and Virtual Path (VP) are multiplexed into one resulting cell stream by the **cell multiplexing** function. The composite stream is normally a non-continuous cell flow. At the receiving side, the **cell demultiplexing** function splits the arriving cell stream into individual cell flows appropriate to the VC or VP.

The ATM layer has two hierarchical levels; virtual channel level and virtual path level, and both are defined in [ITUI 113]:

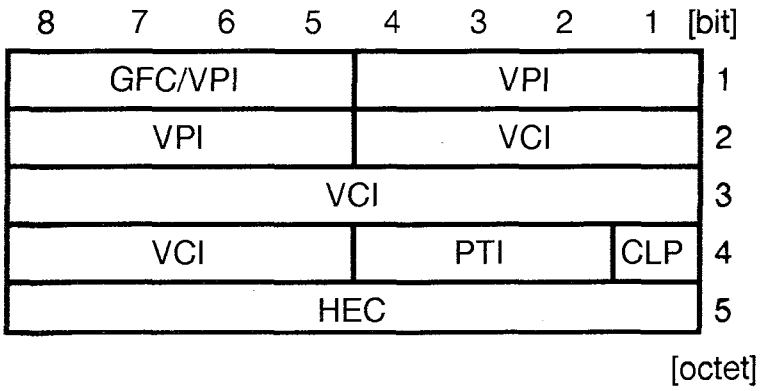


Figure 1.3: Cell header structure

Virtual channel: 'A concept used to describe unidirectional transport of ATM cells associated by a common unique identifier value.' This identifier is called the **virtual channel identifier (VCI)** and is part of the cell header. The VCI is assigned in a field of 16 bits of the cell header for routing.
Virtual path: 'A concept used to describe unidirectional transport of cells belonging to virtual channels that are associated by a common identifier value.' This identifier is called the **virtual path identifier (VPI)** and is part of the cell header. The VPI field at the B-ISDN UNI(User Network Interface) consists of 8 bits while that at NNI(Network Network Interface) comprises the first 12 bits of the cell header, thus providing enhanced routing capabilities.

A transmission path may comprise several VPs, and each VP may carry several VCs. The VP concept allows the grouping of several VCs.

- **VPI and VCI translation** are performed at ATM switching nodes and/or at cross-connect nodes. Within a VP node, the value of the VPI field of each incoming cell is translated into a new VPI value for the outgoing cell. The value of the VPI and VCI are translated into a new value at a VC switch.

- The **cell header generation/extraction** function is applied at the termination points of the ATM layer. In the transmit direction, after receiving the cell information field from the AAL, the cell header generation adds the appropriate ATM cell header except for the **Header Error Control (HEC)**. In the opposite direction, the cell header extraction function removes the cell header. Only the cell information field is passed to the AAL.

- The **GFC(Generic Flow Control)** function is only defined at the B-ISDN UNI. GFC supports control of the ATM traffic flow in a customer network. It can be used to alleviate short-term overload conditions at the UNI. The GFC field in the cell header consists of 4 bits.

Three header bits are used for the **payload type identification (PTI)**. The payload of user information cells contains service adaptation functions. The left bit of PTI is used to distinguish between user cells and F5 cells which support OAM(Operation and Management) of VCCs(Virtual Channel Connection). The center bit is allocated to the **ATM-layer-user-to-ATM-layer-user (AUU)** indication which is used by AAL type 5 (see Section 1.3). The right bit is used for the **congestion indication (CI)** bit which may be modified by any congested network element to inform the end-user about its state (see Section 1.4).

The **cell loss priority (CLP)** field consists of one bit which is used explicitly to indicate the cell loss priority. If the value of the CLP bit is '1', the cell is subject to discard, depending on the network conditions. However, the agreed quality of service (**QOS**) parameters will not be violated. In the other case (CLP = '0'), the cell has high priority and therefore sufficient network resources have to be allocated to it. The CLP bit may be set by the user or the service provider (also see Section 1.4).

The HEC field is part of the cell header, but it is not used by the ATM layer. The HEC sequence is processed by the physical layer and is specified in ITU-T Recommendation I.432 [ITUI 432].

1.3 AAL : ATM Adaptation Layer

The AAL is between the ATM layer and higher layers. Its basic function is the enhanced adaptation of services provided by the ATM layers to the requirements of the higher layer [ITUI 362]. AAL functions are organized in two sublayers: **segmentation and reassembly (SAR)** sublayer and **convergence sublayer (CS)**. The essential functions of the SAR sublayer are, at the transmitting side, segmentation of higher layer PDUs (Protocol Data Units) into a suitable size for the information field of the ATM cell and, at the receiving side, reassembly of the particular information fields into higher layer PDUs. The CS is service dependent.

In order to minimize the number of AAL protocols, ITU-T proposed a service classification specific to the AAL. This classification was made with respect to the following parameters:

- timing relation between source and destination (required or not required)
- bit rate (constant or variable)
- connection mode (connection-oriented or connectionless)

	Class A	Class B	Class C	Class D
Timing relation	Required		Not required	
Bit rate	Constant	Variable		
Connection mode	Connection oriented			Connec-tionless

Figure 1.4: Service classification for AAL

Figure 1.4 depicts the AAL classes, and several AAL protocol types are defined in following subsections. Each type consists of a specific SAR sublayer and CS. This classification fits the AAL service classes. However, no strict relationship between the AAL service classes and the AAL protocol types is requested. Other combination of the described SAR and CS protocols may be used to support specific services.

1.3.1 AAL Type 1

Constant Bit Rate (CBR) services (class A) use AAL type 1 because it receives/delivers data with a constant bit rate from/to the layer above. **Timing information** is also transferred between source and destination. Indication of lost or errored information is sent to the higher layer if these failures can not be recovered within the AAL.

SAR sublayer

The SAR-PDU consists of 48 octets and its format is shown in Fig. 1.5. The first octet includes the Protocol Control Information (PCI) and all other octets are available for the SAR-PDU payload. The PCI is subdivided into a 4 bit sequence number (SN) and a 4 bit sequence number protection (SNP) field. The SN consists of a convergence sublayer indication (C) bit and a 3 bit sequence count (SC) field. The SNP field contains a 3 bit CRC which protects the SN field and an even parity bit. The SC value of the SN makes it possible to detect the loss or misinsertion of cells.

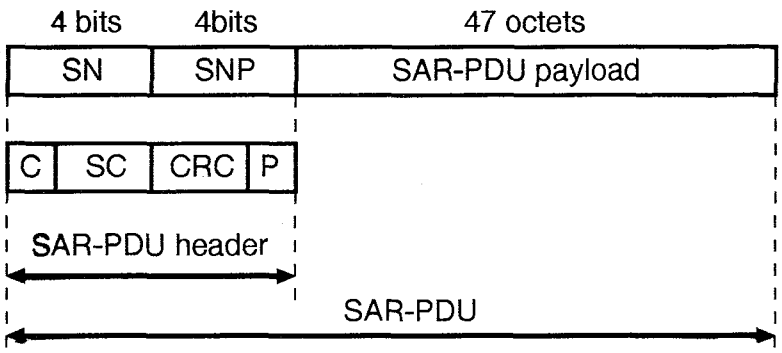


Figure 1.5: SAR-PDU format for AAL type 1

Convergence Sublayer

The function of the CS depends strongly on the service to be supported. Some of these functions are listed in the following.

- Source clock frequency recovery using the synchronous residual time stamp method (SRTS). For details of this method see [ITUI 363].
- Transfer of structure information between source and destination.
- **Forward Error Correction (FEC)** may be used to ensure high quality for some video and audio applications. This may be combined with octet interleaving to give more secure protection against errors. I will introduce this method described in [ITUI 363] at Section 1.5.

1.3.2 AAL Type 2

AAL type 2 is proposed for **Variable Bit Rate (VBR)** services with a timing relation between sources and destination (class B, for example, VBR audio or video). This type is not well defined yet and the services and functions can be anticipated. It might also be possible that AAL type 1 is enhanced to provide AAL type 2 functions.

1.3.3 AAL Type 3/4

The name of this AAL reflects its development: as the service classes were specified, separate AALs were allocated for class C and class D services, namely AAL type 3 and AAL type 4. The AAL 3 was intended to provide framing service for connection-oriented data

protocols like X. 25, and the AAL 4 was intended to provide that for connectionless protocols like IP. The AAL is used for connection oriented as well as for connectionless data communication. However, the AAL itself does not perform all functions required by a connectionless service, since functions like routing and network addressing are performed on the higher layer or the network one. Therefore, both types have merged, thereby supporting both service classes.

SAR sublayer

Figure 1.6 illustrates the SAR-PDU format. In general, CS-PDUs are of variable length. When accepting such a PDU, the SAR sublayer generates SAR-PDUs containing up to 44 octets of CS-PDU data. The CS-PDU is preserved by the SAR sublayer. This requires a **segment type (ST)** indication. The ST indication identifies a SAR-PDU as being **beginning of message (BOM)**, **continuation of message (COM)**, **end of message (EOM)** or **single-segment message (SSM)**.

Multiplexing of multiple CS-PDUs on a single VCI/VPI is supported by a 10 bit **multiplexing identifier (MID)**. The use of the MID field allows the multiplexing of 2^{10} AAL-user-to-AAL-user connection on a single user-to-user ATM layer connection for connection-oriented data communication. For connectionless data communication, the MID field allows interleaving SAR-PDUs up to 2^{10} CS-PDUs on the same semi-permanent ATM layer VC.

Last segment and single one SAR-PDUs may contain less payload octets than the maximum of 44, and thus also need an indication of the number of valid octets in a 6 bit **length indicator (LI)** field.

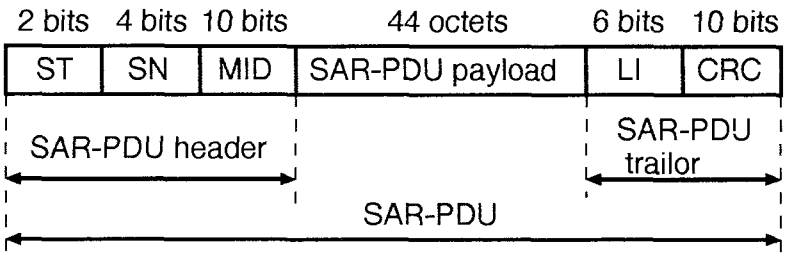


Figure 1.6: SAR-PDU format for AAL type 3/4

1.3.4 AAL Type 5

AAL type 5 will be applied to VBR sources without a timing relation between source and destination. It provides services similar to AAL type 3/4 and will mainly be used for data applications. The reason for defining this additional AAL type was its reduced overhead (i.e., SAR-PDU payload field contains only 44 octets out of a 48 octet SAR-PDU for AAL type 3/4 traffic). Its function and service mode are identical to those for AAL type 3/4. However, one essential difference is that AAL type 5 dose not support a multiplexing function, and thus there is no MID field.

SAR sublayer

The SAR sublayer accepts **service data unit (SDU)** which is of an integer multiple of 48 octets from the CS. No additional overhead is added to the received SDUs at the SAR sublayer. Only segmentation and, in the reverse direction, reassembling functions are performed. For the recognition of the beginning and end of a SAR-SDU (which is correspond to CS-PDU), AAL type 5 makes use of the AUU parameter, which is part of the PT field in the ATM header (cf. section 1.2). An AUU parameter value of '1' indicates the end of a SAR-SDU, while a value of '0' indicates the beginning or continuation of a SAR-SDU. Thus a ST field provided in AAL type 3/4 is not required.

1.4 Traffic Control and Congestion Control in ATM Networks

According to ITU-T Recommendation I.371 [ITUI 371], the primary role of traffic control in B-ISDN is to protect the network and the user in order to achieve predefined network performance objectives in terms of **quality of services (QOS)** such as Cell Loss Ratio (CLR), Cell Transfer Delay (CTD) and Cell Delay Variation (CDV). Traffic control refers to the set of actions taken by the network to avoid congestion. However, congestion may occur because of misfunctioning of traffic control functions caused by unpredictable statistical fluctuations of traffic flows or of network failures, possibly leading to excessive cell losses or unacceptable end-to-end cell transfer delays. Therefore, functions referred to as congestion control ones are intended to react network congestion in order to minimize its intensity, spread and duration. Furthermore, an additional role of traffic control is to utilize network resources efficiently.

1.4.1 Traffic Control Functions

In [ITUI 371], the following six functions are defined in detail as traffic control functions in ATM networks. Here, I will briefly introduce each function.

Connection Admission Control, CAC

CAC represents the set of actions taken by the network at call set-up phase in order to accept or reject an ATM connection. A connection request for a given call is accepted only when sufficient resources are available to carry the new connection through the whole network at its requested QOS while maintaining the agreed QOS of already established connections in the network.

Usage/Network Parameter Control, UPC/NPC

UPC/NPC are performed at the user-network interface (UNI) and the network-network interface (NNI), respectively, and represent the set of actions taken by the network to monitor and control traffic on an ATM connection in terms of cell traffic volume and cell routing validity. This function is also called 'police function'. The main purpose is to enforce the compliance of every ATM connection to its negotiated traffic contract.

Network Resource Management

One tool of network resource management which can be employed for traffic control is the virtual path technique. By grouping several virtual channels together into a virtual path, CAC and UPC/NPC can be simplified as only aggregated traffic of an entire virtual path has to be handled.

Priority Control

ATM cells have an explicit cell loss priority (CLP) bit in the header as previously mentioned in Section 1.2, and at least two different ATM priority classes can be distinguished. Kröner [KRÖN 90] describes different buffering mechanisms for switching/multiplexing systems with two cell loss priorities:

- Common buffer with pushout mechanism: Cells of both priorities share a common buffer. If the buffer is full and a high-priority cell arrives, a cell with low priority (if available) will be pushed out and lost.

- Partial buffer sharing: Low-priority cells can only access the buffer if the queue length of the buffer is less than a given threshold (< total buffer capacity). High-priority cells can access the whole buffer.

- Buffer separation: Different buffers are used for the two priorities. This mechanism is simple to implement but cell sequence integrity can only be maintained if a single priority is assigned to each connection.

Traffic Shaping

Traffic shaping alters the traffic characteristics of a stream of cells on a VPC or VCC in order to reduce the peak rate, limit the burst length or reduce the cell delay variation by suitably spacing cells in time. It must maintain the cell sequence integrity of an ATM connection. It is also used in conjunction with suitable UPC functions, provided the additional delay remains within the acceptable QOS negotiated at call set-up.

Fast Resource Management

Fast resource management is a tool that enables the immediate allocation of necessary capacity, such as bit rate or buffer space, to individual burst-type connections for the duration of a burst. Indication of a burst by the user and allocation acknowledgement by the network could be signalled in-band via specific ATM layer messages.

1.4.2 Congestion Control Functions

The following two functions are simply described as congestion control functions in ATM networks, and however, these specifications may be for further study [ITUI 363].

Selective Cell Discard

A congested network elements may selectively discard cells explicitly identified as belonging to a non-compliant ATM connection and/or those cells with CLP = 1 cell loss priority. This is to primarily protect, as long as possible, high priority CLP = 0 flows.

Explicit Forward Congestion Indication

A network element in a congested state may set an explicit forward congestion indication (CI) in the cell header as described in Section 1.2. At the receiving end, the customer equipment may use this indication to implement protocols which adaptively lower the cell

rate of the connection during congestion. I will refer to this scheme as **Forward Explicit Congestion Notification, (FECN)** in contrast to **Backward - ECN (BECN)**, which will be presented in the following subsection.

1.4.3 ABR Traffic Control Mechanism

The **Constant Bit Rate (CBR)** and **Variable Bit Rate (VBR)** bearer capabilities have been defined by the ATM Forum for connections requiring specified ATM layer Quality of Service (QOS) commitments as regards CLR, CTD and CDV. These QOS guarantees can be delivered by the network by means of a bandwidth resource allocation process, executed at connection set-up time and denoted as the CAC function.

However, the traffic generated by data applications is highly unpredictable and of an extremely bursty nature (see Table 1.1), i.e., a highly variable packet generation rate and with varying packet sizes. It follows that the framework of a fixed static traffic contract as defined for connections using the CBR or VBR bearer capability is not ideally suited for this type of applications. Since data applications are rather sensitive to cell loss but can tolerate substantial variations in delay, there is a need for the network to inform the user via feedback of an impending exhaustion of resources which could ultimately result in cell loss. Therefore, additional ATM layer traffic management facilities are necessary in order to effectively transport traffic from such sources so that network operators can make the most of any unused link capacity without affecting the performance of the CBR and VBR connections. To this end, the *ATM Forum* and also the ITU-T are defining a novel ATM layer bearer capability, referred to as the **Available Bit Rate (ABR)** one [PRYC 95]. On the other hand, the bearer capability, whose connection is not sensitive to cell loss as well as delay, is specified as **Unspecified Bit Rate (UBR)** one. ABR or UBR are usually specified in the traffic contract when the ATM network is providing **best-effort** service. Thus, these two classes of traffic are referred to as best-effort traffic [SIU 95]. Therefore, AAL type 3/4 and type 5 traffic may correspond with best-effort one.

The ATM Forum adopted a **rate based** approach [BONO 95], as opposed to a **credit based** approach with buffer allocation [KUNG 95], with **closed loop** traffic control mechanism for providing an end-to-end ABR bearer capability. This ABR traffic control mechanism will allow the **Source End System (SES)** to dynamically adjust its cell sending rate based on feedback control information received from the network, indicating its availability status of bandwidth resources.

To obtain this information, the ABR connection's SES injects **Resource Management (RM)** cells into its information cell stream on a regular basis, in order to probe the network,

i.e., the RM cells will actually fulfil the role of network bandwidth scouts. For an ABR traffic control mechanism that operates on a pure **end-to-end** basis, these RM cells are then returned by the **Destination End System (DES)**. Thus, for the ABR information flow from SES to DES, there exists two RM cell flows: one in the forward direction from SES to DES and another in the backward direction from DES to SES.

Moreover, the adopted rate based policy based on these RM cell flows also provides the option to safely create separately controlled ABR segments. The introduction of the concept of a **virtual destination** and of a **virtual source** allows the information control loop to be segmented at any convenient point along the end-to-end path. For any ABR connection, an intermediate network may therefore create its own independently and internally controlled segment by terminating the loop. Furthermore, it is possible to create a **link-by-link** closed loop rate control mechanism.

Some kinds of ABR traffic control schemes are listed as follows.

Binary feedback information

If the SES receives the backward RM cell with **Congestion Indication (CI)** = 0, namely, no congestion experienced, then it increases its sending cell rate. Otherwise, i.e., when it receives the RM cell with CI = 1 (congestion experienced in the forward path) or no receiving of that, it reduces the sending rate; it is a kind of positive feedback control [YANG 95].

Explicit rate feedback information

In interleaving the RM cell at the SES, it is supposed to initially write the cell rate into the **Explicit Rate (ER)** field in the RM cell. The value in ER field could be rewritten to reduce (not increase) by intermediate switches in accordance with their available rate. When the backward RM cell is received by originating SES, the cell rate value advertised in the ER field specifies an explicit upper bound at which the SES may submit its cells to the network, since this value represents the bottleneck rate of the ABR connection on its end-to-end path.

Intelligent marking

CI bits and the ER field are in the RM cell, and thus intelligent marking selectively employs binary feedback or explicit rate feedback for achieving an ABR **fairness** requirement.

Enhanced proportional rate control algorithm (EPRCA)

EPRCA may be similar to binary feedback, except for the followings:

- The RM cell insertion frequency proportional to the number of data cells, N_{rm}, sent to the network.
- Additive increase approach when the SES increases its cell rate
- Multiplicative decrease approach if the SES reduces its cell rate

Backward explicit congestion notification (BECN)

The intermediate switch that receives a forward RM cell with a current cell rate larger than the rate it can support, may reduce the ER parameter. Beside continuing the RM cell forward, the switch may at the moment also generate an RM cell back towards the SES with CI = 1. Thus, BECN based feedback control operates on the time scale of the round-trip propagation delay between SES and congested point, not DES as in using FECN.

1.5 Error Control for AAL Type 1 Traffic

As described in Section 1.2, although ATM deals with no error control and it relies on higher layer protocol based on end-to-end control, **Forward Error Correction (FEC)** may be used to ensure high quality for some video and audio application [ITUI 363].

This correction method combines FEC and octet interleaving, from which CS-PDU structure is defined. FEC uses the **Reed-Solomon** (124,128) code which is able to correct up to 2 errored symbols (octets) or 4 erasures in the block of 128 octets. An erasure represents an errored octet whose location in the block is known. The specific polynomials to be used for Reed-Solomon code are for further study. In the transmitting CS, 4 octet Read-Solomon code is appended to 124 octets of incoming data from the upper layer. The resulting 128 octet long blocks are then forwarded to the octet interleaver. Figure 1.7 shows a format of the interleaving matrix.

The octet interleaver is organized as a matrix of 128 columns and 47 rows. The interleaver is used as follows; at the input, incoming 128 octets long blocks are stored row by row (one block corresponding to one row); at the output, octets are read out column by column. The matrix has $128 \times 47 = 6016$ octets, corresponding to 128 SAR-PDU payload. These 128 SAR-PDU payloads constitute one CS-PDU.

In this process, the loss of one SAR-PDU payload in the matrix implies one erasure in each row of the matrix. Erasures correspond to dummy cell payloads inserted in the cell

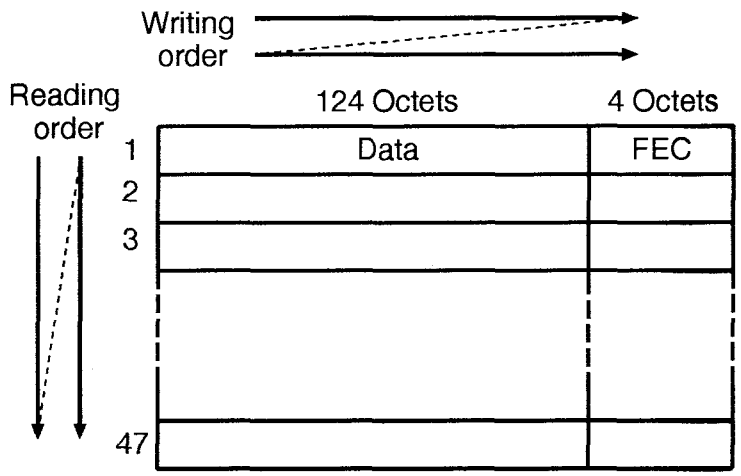


Figure 1.7: Format of the interleave matrix

flow when a cell loss has been detected. Misinserted cells which have been detected are merely discarded in the CS.

For the synchronous of the CS-PDU, the CS indicator bit of the SAR-PDU header (see Section 1.3) is set to 1 for the first SAR-PDU payload of the CS-PDU.

Within any CS-PDU matrix, this method can perform the following corrections:

- 4 cell losses ; or
- 2 cell losses and 1 errored octet in each row ; or
- 2 errored octets in each row if there is no cell loss.

The overhead of this method is 3.1%, and the delay is 128 cells.

1.6 Overview of the Dissertation

So far, I have introduced a general view of ATM technique, for realizing broadband-ISDN(B-ISDN) in Section 1.1, especially ATM layer functions (in Section 1.2) and AAL ones which are dependent on each service type such as voice on circuit emulation mode, constant and variable bit rate audio, video, data transfer and so on in Section 1.3.

Again, ATM-based B-ISDN (hereafter, I will call this **ATM network** in brief) handles the transfer of small fixed-size cells on ATM layer. ATM enables us to transmit traffic of a wide range of bit rates, from the order of kbit/s to hundreds of Mbits/s, in unique network interface. Therefore, in ATM networks, it is expected that the statistical multiplexing of cells from different service classes of higher layer may achieve efficiently utilization of network resources (i.e., bandwidth and the buffer of switch). On the other hand, it may cause cell loss even at any places such as UNI, NNI, intermediate ATM switch or cross-connect point along end-to-end connection. Cell loss is considered as one of main problems or drawbacks in ATM networks. Therefore, traffic control function is absolutely required in ATM network to avoid congestion as presented in ITU-T Recommendation I.371 [ITUI 371], and I have described the summary of traffic control functions in Section 1.4.

It is considered that CAC and UPC are fundamental traffic control functions to *preventively* avoid congestion. These control functions are based on network contracts and declared parameters of traffic from users at the call set-up time. However, user traffic has highly correlated nature (namely high burstiness, see Table 1.1) and may have unpredictable statistical fluctuation during established connection phase, that is, short-term congestion may still occur. In order to relieve such congestion and to minimize its intensity, spread and duration, congestion control function is also recommended in [ITUI 371]. Its specification is not clarified for whole AAL service class, but only for available bit rate (ABR) class, and it has been summarized in Section 1.4, in which it is required to apply not only **preventive** (or, open-loop [YANG 95, OHSA 95]) control such as CAC and UPC, but also rate-based, closed-loop **reactive** control.

To this end, first in Chapter 2, I will focus on reactive congestion control scheme based on **BECN** as shown in Section 1.4 for ABR traffic in ATM networks. For this purpose, I treat the two-stage queueing network where the output from the multiple *sources* in the first stage is multiplexed at the switching node, called *server*, in the second stage with some propagation delay between two stages. It thus can be assumed that the server becomes the bottle-neck or congested node in ATM networks. In this model, the cell transmission rate at each source can be adjusted to the queue length in the server because each sources can be equipped with buffers. For the cell arrival processes at each source, I employ the

two-state Markov Modulated Bernoulli Process (**MMBP**) which enables us to treat bursty traffic such as data. By using this model, I provide an approximate analysis to examine the impact of propagation delay, buffer size, and traffic characteristics such as average arrival rate and burstiness on the **cell loss probability** and the **average waiting time** of reactive congestion control. Furthermore, I will propose the **cell push-out** scheme, in which the cell transmission also depends on its own queue length. This scheme forces each source to *push out* its **head-of-line (HOL)** cell to the server when there is no space in its buffer for newly arriving cell to improve the overall cell loss performance in the system. It is noted that this cell push-out scheme must be distinguished from pushout mechanism as priority control defined in Section 1.4.

ECN based reactive congestion control operates on the time scale of the round-trip propagation delay of ATM connection. Therefore, in ATM **wide area network (WAN)**, the feedback delay of congestion information or the RM cell becomes so large that it may degrade the performance in terms of the cell loss probability as well as the average waiting time in intermediate switches. Thus, **link-by-link backpressure** (or BP in short) congestion control is expected to improve the performance in ATM WAN environment. Link-by-link control would be optionally realized for controlling ABR traffic as mentioned in Section 1.4. Therefore, in Chapter 3, I will investigate the performance of link-by-link BP control. I analyze the BP control by applying the analytical approach, developed in Chapter 2, where I deal with the relation between the congested server and its upstream nodes instead of sources. The multiplexed cell stream of congested server and other congestion free servers must be dealt with at each upstream node, MMBP is also intended to model the superposition of these traffic [HEFF 86]. As mentioned in [JAIN 92], the BP control would make cells not traveling the congested server adversely affected in terms of cell loss performance by the HOL-blocking due to BP. Therefore, I will also examine how the cell push-out scheme efficiently relieves the HOL-blocking and how it successfully enables cell not transmitting the congested server to travel its upstream node without blocked.

In Chapter 4, I will focus on another congestion control, **selective cell discard** control based on the burst or packet level. The packet of the ATM upper layer is divided into one or more cells, and which are transmitted in ATM networks. If a cell gets lost within ATM networks, the AAL layer of destination can not reassemble the packet to which the cell belongs, and the transport-layer requires the source to retransmit the packet. Therefore, once one of cells constituting a packet is lost, its subsequent cells in the packet are not worth relaying within the ATM network. Thus, the network resources could be utilized efficiently in a way to selectively discard such cells at the multiplexing node or the ATM switch. In applying this scheme to ABR traffic, the intermediate ATM switch or cross-connect point

should identify the end cell of the upper layer packet. As mentioned in Section 1.4, AAL 3/4 protocol has this information in the header of AAL layer PDU whereas it is filled in the ATM layer header in case of AAL 5 traffic. Therefore, this scheme could be applied to AAL 5 traffic. In order to evaluate the effectiveness of this scheme, I should keep track of the packets from which cells have been lost. Therefore, I will propose the exact analysis of the **packet loss probability** on the system which is applied selective cell discard schemes by the following assumptions: the distribution of the packet length from AAL 5 traffic is geometric one and sources are homogeneous in terms of traffic characteristics.

In conventional networks, as a way to recover from loss, **ARQs (Automatic Repeat Requests)** have prevailed, which rely on sources to retransmit the packet when they are informed of cell loss happening within the networks by the receiver. However, as mentioned in Section 1.5, **FEC (Forward Error Correction)** may be used to ensure high quality for some video and audio application on AAL type 1 connection. This FEC scheme is based on octet interleaving with Reed-Solomon coding. On the other hand, the **Reed-Solomon erasure (RSE)** code developed in [ANTH 90] has been implemented by the device operating at 1 Gbit/s because its less complexity compared with the conventional Reed-Solomon code. This makes FEC combined with **cell interleaving** possible, and enhances the capability of recovery from cell loss. Thus, in Chapter 5, I will evaluate the performance of FEC with cell interleaving. In FEC scheme based on RSE code, the FEC encoder produces some redundant cells, say r cells, every predetermined number of data cells, say p cells. Those redundant cells together with data cells form a group called a *block*. In this case, the receiver can correct up to r cells among cells lost in a block without retransmission. Accordingly, the **block loss probability** is a primary performance measure instead of the individual cell loss probability. Further, the **cell loss distribution in a block** would be highly correlated since cell generation processes from some kind of sources have bursty nature. Therefore, we will provide an exact analysis of the distribution of the number of cells lost in a block, and obtain the block loss probability from the distribution. It is assumed the **interrupted Bernoulli process (IBP)**, which is a subset of MMBP, as cell arrival process of sources in order to deal with bursty traffic.

In Chapter 6, some concluding remarks and suggestions for future research are given in terms of studies in this dissertation.

The results discussed in Chapter 2 are mainly taken from [KAWA 95a], Chapter 3 from [KAWA 95b, KAWA 95c], Chapter 4 from [KAWA 96] and Chapter 5 from [KAWA 94].

Chapter 2

Performance Analysis of Reactive Congestion Control

2.1 Introduction

In packet switching networks, congestion control methods are mainly divided into two categories: preventive control and reactive control[HABI 91]. In conventional packet switching networks, reactive controls are generally adopted in which the source node recognizes network congestion by means of the feedback information sent from the destination node or intermediate ones. Then, the source node limits the packet injection into the network by some appropriate method. However, in high speed networks of today, propagation delay of such congestion information are no longer negligible compared to the packet transmission times. This sort of congestion control policy results in severe performance degradation for real-time applications during the congestion period, which is greater than round-trip propagation delay. Therefore, preventive controls are now being recognized as an attractive way of treating congestion problems in high speed networks. For example, in ATM networks, congestion is prevented by limiting the number of connections on each link in connection admission control (CAC) and usage parameter control (UPC), which ensure required quality of service (QOS) for the established connections (see, e.g., [GILB 91]).

However, it often happens that users do not have enough information related to the characteristics of their traffic before transmitting it, as mentioned in [GERS 91]. It is therefore desirable that the CAC is performed based upon simple parameters associated with incoming traffic. Gersht and Lee [GERS 91] have proposed congestion control at two levels, VC level and cell level. In their mechanism, the simple CAC at VC level successfully prevents congestion in combination with reactive congestion control at cell level, which is applied only to nonreal-time traffic. Abould-Magd and Henry [ABOU 92]

and Newman [NEWM 93] have shown by means of simulations that reactive congestion control, referred to as *Explicit Congestion Notification (ECN)* mechanism, is effective in ATM networks in allowing an efficient use of network resources while maintaining adequate QOS. In particular, Newman [NEWM 93] has focused on ATM LANs and shown that a backward ECN mechanism is very effective in achieving high link utilization in ATM LANs of up to at least 50 km. The analytical approach for reactive congestion control was developed by Wang and Sengupta [WANG 91], in which propagation delay between sources and a congested node is taken into account, and packets are assumed to be always transmitted in available slots from each source according to a Poisson distribution.

The main objective of this Chapter is to provide insights into the benefits gained by using the reactive control algorithm and its limitations. For this purpose, we treat the two-stage queueing network where the output from the multiple *sources* in the first stage is multiplexed at the switching node, called *server*, in the second stage with some propagation delay between two stages. In this model, the cell transmission rate at each source can be adjusted according to the queue length in the server because each source is equipped with buffers. For the cell arrival processes at each source, we employ the two-state Markov Modulated Bernoulli Process (MMBP) which is intended to model the superposition of bursty traffic [HEFF 86]. By using this model, we provide an approximate analysis to examine the impact of propagation delay, buffer size, and traffic characteristics such as average arrival rate and burstiness on the performance of some reactive congestion control algorithms. In particular, we are interested in ATM LANs in which propagation delay is very low, and reactive congestion controls are thus expected to be very effective.

Our modeling intends to treat the following case of practical interest; the ECN is sent along the forward path to the destination, referred to as the *forward* ECN, and each source stops/begins to transmit its cells according to the ECN. The *backward* ECN mechanism can be also treated in which the latency experienced by the ECN is smaller than in the forward ECN mechanism. [1] Another type of congestion control which we will treat is that each source controls its cell transmission depending on its own queue length as well as the ECN. More specifically, the latter mechanism forces each source to *push out* its head-of-line cell to the server when there is no space in its own buffer for a newly arriving cell. If at most one cell arrives at a source per slot, cell loss does not occur at the source in this case. Furthermore, the pushed-out cells from some sources can enter the server buffer if there is any space in the buffer, whereas these cells, in turn, increase the server queue, causing the sources to stop their cell transmission for a longer duration. Accordingly, this should lead to

[1] Our analytical approach can be applied to a more general case where the ECN contains the queue length of the congested node so that each source varies its cell transmission rate according to such information.

performance improvement with respect to cell loss at the cost of some additional waiting time incurred in the buffer, along with an appropriate ECN-based congestion control. Within an acceptable waiting time, performance improvement can actually be obtained, which will be shown in our numerical examples.

The remainder of this Chapter is organized as follows. In Section 2.2, we describe our analytical model. Section 2.3 provides an approximate analysis for deriving the cell loss probability and the average waiting time under reactive congestion controls considered in this Chapter. Numerical results and discussions are given in Section 2.4. Section 2.5 presents concluding remarks.

2.2 Analytical Model

In this section, we describe our model for analyzing the reactive congestion control algorithms. Our queueing system consists of two-stage queues with finite capacity in series, as illustrated in Fig. 2.1. In the first stage, there are R sources in parallel. S_{1j} denotes the j-th source in the first stage, which has a cell buffer of K_j in length. Cells transmitted from these sources are multiplexed at the server S_2 in the second stage, which is equipped with a buffer of B in length.

We consider a discrete time queueing system with time slot being equal to a cell transmission time on the output link. It is assumed that cells arrive at each source S_{1j} according to a two-state Markov-Modulated Bernoulli Process (MMBP), which is characterized by a set of parameters, $(\Lambda_{1j}, \Lambda_{2j}, r_{1j}, r_{2j})$. It has two states, state 1 and state 2. At each slot, the process changes its state from state 1 to state 2 with probability r_{1j}, and from state 2 to state 1 with probability r_{2j}. The mean number of cells arriving at the j-th source per slot is Λ_{1j} during state 1 and Λ_{2j} during state 2, respectively. Thus, the average arrival rate at the j-th source queue is given by

$$\Lambda_j = \frac{r_{2j}}{r_{1j}+r_{2j}}\Lambda_{1j} + \frac{r_{1j}}{r_{1j}+r_{2j}}\Lambda_{2j} \quad (1 \le j \le R). \tag{2.1}$$

Suppose that $\Lambda_{1j} \ge \Lambda_{2j}$. In addition to the average arrival rate, we define the burstiness, β_j, which is commonly used to represent how the arrival traffic is bursty, and the burst-length, L_{b_j}, as follows,

$$\beta_j \triangleq \frac{\Lambda_{1j}}{\Lambda_j}, \qquad L_{b_j} \triangleq \frac{1}{r_{1j}} \quad (1 \le j \le R). \tag{2.2}$$

We introduce the probability matrix of state transition and that of cell arrival for each

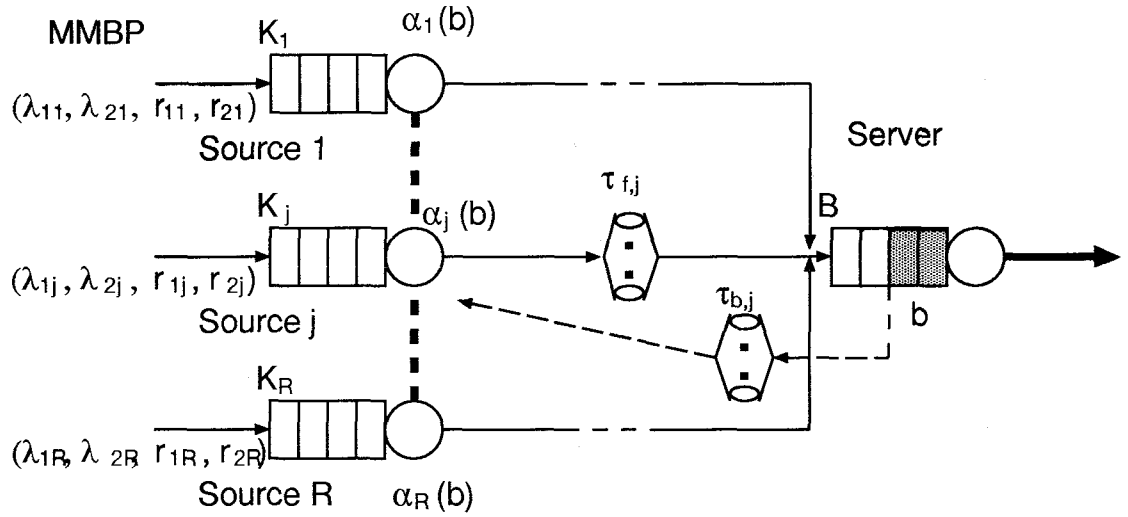


Figure 2.1: Analytical model for reactive congestion control

MMBP at source queues, $\boldsymbol{T}_j$ and $\boldsymbol{a}_{i,j}$, respectively, as follows,

$$\boldsymbol{T}_j = \begin{pmatrix} 1-r_{1j} & r_{1j} \\ r_{2j} & 1-r_{2j} \end{pmatrix}, \quad \boldsymbol{a}_{i,j} = \begin{pmatrix} a_{1j}(i) & 0 \\ 0 & a_{2j}(i) \end{pmatrix} \qquad (1 \leq j \leq R). \tag{2.3}$$

Each element, $a_{lj}(i)$ $(l = 1, 2)$, in the above matrix $\mathbf{a}_{i,j}$ corresponds to the probability that the number of i cells arrive at the j-th source queue at each slot when MMBP is in state l, and is given by

$$a_{lj}(i) = \binom{N_{lj}}{i} \left(\frac{\Lambda_{lj}}{N_{lj}}\right)^i \left(1 - \frac{\Lambda_{lj}}{N_{lj}}\right)^{N_{lj}-i} \qquad (1 \leq j \leq R), \tag{2.4}$$

where N_{lj} is the maximum number of cells arriving at each slot at the j-th source when it is in state l.

We introduce a control function $\alpha_j(b)$ for each source S_{1j}, which is a function of the queue length, b, of the server, and takes values ranging from 0 to 1. At each slot, by using $\alpha_j(b)$ as a probability of transmitting a cell, the j-th source determines whether it transfers the cell in its buffer to the server or not. Apparently, as the number of cells at the server becomes large, i.e., when the server is congested, the cells should be held momentarily in the sources to relieve congestion at the server quickly. In this case, the transmission rate at each one of the sources should be decreased. However, if the transmission rate is kept low at sources even after congestion is relieved at the server, cell loss at sources would unnecessarily happen.

Furthermore, we take into account propagation delay between the source and the server; let $\tau_{f,j}$ be cell propagation delay from the j-th source to the server, and $\tau_{b,j}$ be delay experienced by the congestion notification from the server to the j-th source, as illustrated in Fig. 2.2-(a). Namely, the j-th source transmits its cell to the server according to the state of the server queue at the $\tau_{b,j}$-th slot preceding the current slot. The state of the server at the current slot thus depends on the state of the j-th source at the $\tau_{f,j}$-th slot behind the current slot. As a result, it is very hard to carry out an exact analysis of this system on a slot basis. [2]

We set $\tau_{f,j}$ and $\tau_{b,j}$ to the same value, τ, for all the sources. A block of consecutive $2\tau + 1$ slots will be called a *super-slot* (see Fig.2.2-(b)). Furthermore, we use the following model which makes the analysis tractable, and then obtain the cell loss and the average

[2]Wang and Sengupta [WANG 91] proposed a sophisticated approach for analyzing a similar system by taking into account propagation delay between sources and the server as mentioned in Section 2.1. It is also of interest to investigate a manner to apply their approach to the system considered here. Since we are in particular interested in ATM LANs with very low propagation delay, we derive here an exact analysis of the system under no propagation delay, and then approximately analyze the system based upon the analysis under some propagation delay.

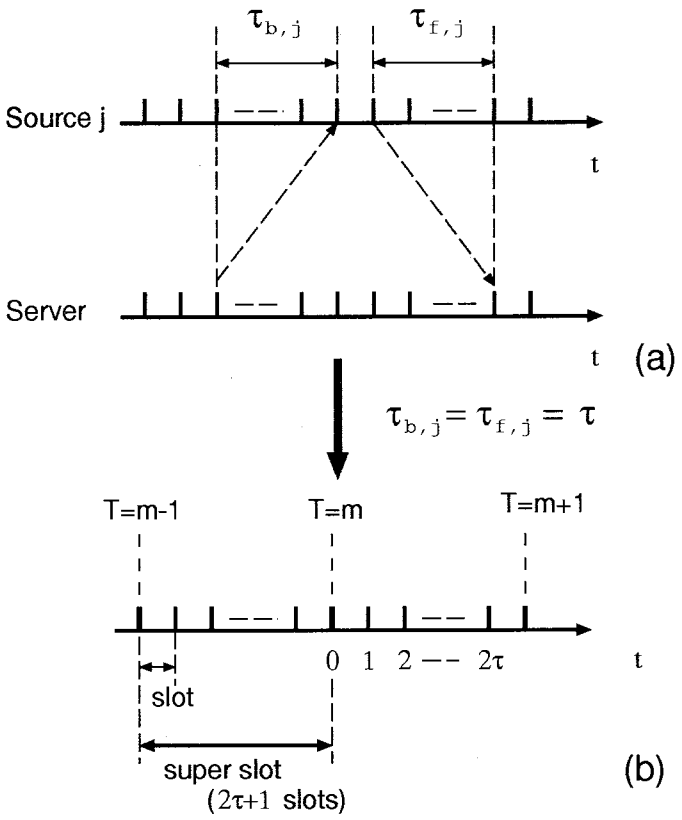


Figure 2.2: Behavior of cell transmission and outline of *super-slot*

waiting time characteristics at an arbitrary slot by using the steady state probability at the beginning of the super-slot. It is worth noting that the following model is the exact one when τ is equal to 0.

1) The server receives cells transmitted during a super-slot from each source only at the end of the super-slot. Since cells transmitted at the i-th slot ($i = 0, 1 \ldots, 2\tau$) of a super-slot arrive at the server in $2\tau - i$ slots, the average delay time of cells transmitted from sources to the server is equal to $\frac{\sum_{i=0}^{2\tau}(2\tau - i)}{2\tau+1}$, resulting in τ.

2) Each source is assumed to receive the information of congestion notification from the server only at the beginning of a super-slot with no delay. The value of the control function $\alpha_j(b)$ is changed only at the beginning of a super-slot, and is kept fixed all over the super-slot. Since the feedback delay at the i-th slot of a super-slot is actually equal to i slots, the average feedback delay time per super-slot from the server to sources becomes $\frac{\sum_{i=0}^{2\tau} i}{2\tau+1}$. This is also equal to τ.

2.3 Derivation of Steady State Probability of Buffer Contents

The state of our system can be completely described by a set of 1) the queue length, D_2, of the server, 2) the queue length, D_{1j}, of each source, and 3) the MMBP state, S_j, of each source. Since it forms a Markov chain on a super-slot basis, we can derive the steady state probability at the beginning of a super-slot, $P(b, k_1, s_1, \cdots, k_R, s_R) \stackrel{\triangle}{=} \Pr(D_2 = b, D_{11} = k_1, S_1 = s_1, \cdots, D_{1R} = k_R, S_R = s_R)$, by using the transition probability matrix $\boldsymbol{Q}$ obtained below for this process, and then also derive the overall system cell loss probability.

2.3.1 Transition Probability Matrix for System State, Q

In order to derive $\boldsymbol{Q}$, we below obtain the following matrices: 1) the transition probability matrix per slot, $\boldsymbol{G}_{b,j}(x)$, of the j-th source, 2) the transition probability matrix per slot, $\boldsymbol{G}_b(y)$, of all the sources, and 3) the transition probability matrix per super-slot, $\boldsymbol{L}_{b,b'}$, concerning the server.

Transition Probability Matrix for Each Source

To represent the transition probability matrix per slot for each source, we assume that the events on each source at a slot occur in the following sequence,

1. At the beginning of a slot, each source decides whether it transmits its head-of-line (HOL) cell according to $\alpha_j(b)$. The value of $\alpha_j(b)$ at the beginning of a super-slot is assumed to be used throughout that super-slot as mentioned earlier. If the j-th source can transmit a cell at the slot, it starts to send the HOL cell to the server.

2. New cells arrive at the sources and are stored in their buffers. If the j-th source contains k_j cells at the beginning of a slot, it can store at most $K_j - k_j + 1$ new cells. When more than $K_j - k_j + 1$ cells arrive, some of them will be lost. If there is no cell in the buffer at the beginning of a slot, the source immediately starts to send one of the newly arriving cells. [3]

3. The source finishes sending the HOL cell at the end of a slot.

Let $\boldsymbol{G}_{b,j}(x)$ denote the transition matrix associated with the j-th source on the condition that the server contains b cells in the buffer. Its (k,l)-block element corresponds to the probability that the number of x cells are transmitted from the j-th source to the server, and that the number of cells, D_{1j}, at the j-th source is changed from k to l. As mentioned in Section 2.2, each source transmits its HOL cell to the server with probability $\alpha_j(b)$, which will be specified for each control algorithm in Section 2.4.

As mentioned in Section 1, we deal with two congestion control mechanisms, *discipline A* and *discipline B*, which are different in the buffer management at the sources. $\mathbf{G}_{b,j}(x)$ is obtained below for these two disciplines.

1) discipline A (basic one)
Whether the HOL cell is transmitted or not depends only on $\alpha_j(b)$. $\boldsymbol{G}_{b,j}(x)$ takes the following form under this discipline.

$$\boldsymbol{G}_{b,j}(0) = (1-\alpha_j(b)) \times \begin{pmatrix} \frac{\boldsymbol{a}_{0,j}\boldsymbol{T}_j}{(1-\alpha_j(b))} & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-1,j}\boldsymbol{T}_j & \sum_{i=K_j}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \mathbf{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-2,j}\boldsymbol{T}_j & \sum_{i=K_j-1}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \cdots & \ddots & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \sum_{i=1}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \mathbf{0}[2] & \cdots & \cdots & \mathbf{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j \end{pmatrix},$$

[3] Some assumptions other than this one (cut-through assumption) can be used. This one is chosen only as one of available assumptions.

$$\boldsymbol{G}_{b,j}(1) = \alpha_j(b) \times \begin{pmatrix} \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \boldsymbol{a}_{2,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j,j}\boldsymbol{T}_j & \sum_{i=K_j+1}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-1,j}\boldsymbol{T}_j & \sum_{i=K_j}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \boldsymbol{0}[2] & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \sum_{i=2}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \boldsymbol{0}[2] & \cdots & \boldsymbol{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \sum_{i=1}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \end{pmatrix}$$
$$(0 \le b \le B, \quad 1 \le j \le R), \tag{2.5}$$

where $\boldsymbol{0}[m]$ represents a zero matrix of dimension $m \times m$.

2) discipline B (*pushout* discipline)

When cell arrivals cause buffer overflow at the j-th source, the HOL cell is pushed out to the server regardless of the value of $\alpha_j(b)$, thereby enabling one of the new cells to enter the source buffer. Thus, the transition matrix $\boldsymbol{G}_{b,j}(x)$ under discipline B is different from that under discipline A only in the rightmost column, which is related to the case where the buffer of the j-th source becomes fully occupied. Under this discipline, we obtain $\boldsymbol{G}_{b,j}(x)$ as

$$\boldsymbol{G}_{b,j}(0) = (1-\alpha_j(b)) \times \begin{pmatrix} \frac{\boldsymbol{a}_{0,j}\boldsymbol{T}_j}{(1-\alpha_j(b))} & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-1,j}\boldsymbol{T}_j & \boldsymbol{a}_{K_j,j}\boldsymbol{T}_j \\ \boldsymbol{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-2,j}\boldsymbol{T}_j & \boldsymbol{a}_{K_j-1,j}\boldsymbol{T}_j \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \cdots & \ddots & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \boldsymbol{a}_{1,j}\boldsymbol{T}_j \\ \boldsymbol{0}[2] & \cdots & \cdots & \boldsymbol{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j \end{pmatrix},$$

$$\boldsymbol{G}_{b,j}(1) = \alpha_j(b) \times \begin{pmatrix} \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \boldsymbol{a}_{2,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j,j}\boldsymbol{T}_j & \sum_{i=K_j+1}^{\infty} \frac{\boldsymbol{a}_{i,j}\boldsymbol{T}_j}{\alpha_j(b)} \\ \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-1,j}\boldsymbol{T}_j & \sum_{i=K_j}^{\infty} \frac{\boldsymbol{a}_{i,j}\boldsymbol{T}_j}{\alpha_j(b)} \\ \boldsymbol{0}[2] & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \sum_{i=2}^{\infty} \frac{\boldsymbol{a}_{i,j}\boldsymbol{T}_j}{\alpha_j(b)} \\ \boldsymbol{0}[2] & \cdots & \boldsymbol{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \sum_{i=1}^{\infty} \frac{\boldsymbol{a}_{i,j}\boldsymbol{T}_j}{\alpha_j(b)} \end{pmatrix}$$
$$(0 \le b \le B, \quad 1 \le j \le R). \tag{2.6}$$

Transition Probability Matrix of Cells Transmission of All Sources

We introduce $\boldsymbol{G}_b(y)$ which represents the probability that the number of y cells are transmitted at a slot from the sources to the server in total, and it can be expressed by the Kronecker product of cell arrival matrices from the sources, i.e.,

$$\boldsymbol{G}_b(y) = \sum_{\sum_{j=1}^{R} x_j = y} \boldsymbol{G}_{b,1}(x_1) \otimes \cdots \otimes \boldsymbol{G}_{b,R}(x_R) \quad (0 \le y \le R). \tag{2.7}$$

Transition Probability Matrix Concerning the Queue Length of Server per Super-slot

To obtain the probability matrix which denotes the transition per super-slot with respect to the queue length of the server, we assume the events on the server at a slot occur in the following sequence,

1. The server begins to forward its HOL cell.
2. Cells arrive from each source. If the server contains b cells at the beginning of the slot, it can store at most $B - b + 1$ cells of arriving cells. If the server contains no cells at the beginning of a slot, it immediately begins to forward one of the arriving cells.
3. The server finishes forwarding the cell at the end of a slot.

In what follows, we obtain the transition matrix $\boldsymbol{L}_{b,b'}$, which represents the transition of the number of cells stored at the server from b to b' at the beginning of two consecutive super-slots.

For this purpose, we temporarily define the matrix $\boldsymbol{M}_{b,t}(x)$, whose elements represent the probabilities that the server queue length is equal to x at the end of the t-th slot in a super-slot, given that it is equal to b at the beginning of the super-slot. $\boldsymbol{L}_{b,b'}$ is related to this matrix as follows.

$$\boldsymbol{L}_{b,b'} \triangleq \boldsymbol{M}_{b,2\tau}(b'). \tag{2.8}$$

Suppose that $x + 1 - b$ cells arrive at the server at the 0-th slot, which includes b cells at the beginning of the 0-th slot. One of the b cells will be transmitted, and x cells will be kept at the server at the end of the 0-th slot. Therefore, we have $\boldsymbol{M}_{b,0}(x)$ at the end of the 0-th slot in a super-slot as follows.

$$\boldsymbol{M}_{b,0}(x) = \begin{cases} \sum_{j=0}^{1} \boldsymbol{G}_b(j-b), & \text{if } x = 0 \text{ and } b \le 1 \\ \boldsymbol{G}_b(x+1-b), & \text{if } 1 \le x \le B-1 \\ \sum_{j=1}^{R-B+b} \boldsymbol{G}_b(B+j-b), & \text{if } x = B \end{cases}. \tag{2.9}$$

In a similar manner, we obtain the following relation between $\mathbf{M}_{b,t}(x)$ and $\boldsymbol{M}_{b,t-1}(y)$.

$$\boldsymbol{M}_{b,t}(x) = \begin{cases} \sum_{j=0}^{1} \sum_{i=0}^{\min(j,R)} \boldsymbol{M}_{b,t-1}(j-i)\boldsymbol{G}_b(i), & \text{if } x = 0 \\ \sum_{i=0}^{\min(x+1,R)} \boldsymbol{M}_{b,t-1}(x+1-i)\boldsymbol{G}_b(i), & \text{if } 1 \le x \le B-1 \\ \sum_{j=1}^{R} \sum_{i=j}^{\min(B+j,R)} \boldsymbol{M}_{b,t-1}(B+j-i)\boldsymbol{G}_b(i), & \text{if } x = B \end{cases}$$
$$(1 \le t \le 2\tau). \tag{2.10}$$

We can obtain $\boldsymbol{M}_{b,2\tau}(b')$ by using Eq.(2.9) and Eq.(2.10) iteratively.

2.3.2 Derivation of Steady State Probability

Using the matrix $\boldsymbol{L}_{b,b'}$, the whole transition probability matrix $\boldsymbol{Q}$ can be built as follows;

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{L}_{0,0} & \boldsymbol{L}_{0,1} & \cdots & \boldsymbol{L}_{0,B} \\ \boldsymbol{L}_{1,0} & \boldsymbol{L}_{1,1} & \cdots & \boldsymbol{L}_{1,B} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{L}_{B,0} & \boldsymbol{L}_{B,1} & \cdots & \boldsymbol{L}_{B,B} \end{pmatrix}. \tag{2.11}$$

Note that the dimension of $\boldsymbol{Q}$ is given by

$$(B+1)q \times (B+1)q, \quad (q \triangleq 2^R(K_1+1)\cdots(K_R+1)). \tag{2.12}$$

Consequently, we can obtain the steady state probability vector by solving the following equation;

$$\boldsymbol{PQ} = \boldsymbol{P}, \quad \boldsymbol{Pe}\,((B+1)q) = 1, \tag{2.13}$$

where $\boldsymbol{e}(m)$ represents a column vector of dimension m, whose elements are all equal to 1, and

$$\begin{aligned} \boldsymbol{P} &= [\boldsymbol{P}_0, \cdots, \boldsymbol{P}_b, \cdots, \boldsymbol{P}_B], \\ \boldsymbol{P}_b &= [\cdots, P(b, k_1, s_1, \cdots, k_R, s_R), \cdots], \\ & \quad 0 \le k_j \le K_j, \quad s_j = 1,2 \quad (0 \le j \le R), \end{aligned} \tag{2.14}$$

whose elements are ordered appropriately in accordance with the configuration of the transition matrices.

It may be difficult to directly solve Eq. (2.13) since the number of states grows explosively as the size of buffer capacity at sources and/or the server becomes large (See Eq.(2.12)). We thus employ the computational algorithm proposed in [STEW 89], by which the computational cost is determined by the buffer size of the source queue and independent of that of the server queue. [4]

[4]The dimension of the sub-matrices $\boldsymbol{L}_{i,j}$ is a function of the source buffer size, but is not a function of the server buffer size. In addition, if propagation delay and the number of sources are relatively small, many of the sub-matrices $\boldsymbol{L}_{i,j}$ become $\boldsymbol{0}$ ones when $|i-j|$ is relatively large. Furthermore, the computational scheme in [STEW 89] enables us to obtain the solution to Eq.(2.13) only by means of non-zero sub-matrices. Therefore, the computational cost becomes less expensive as propagation delay becomes lower; that is, this computational scheme is very effective in ATM LANs we have focused on here.

2.3.3 Derivation of Cell Loss Probabilities

In this subsection, we will derive the cell loss probability at an arbitrary slot by using the steady state probability at the beginning of a super-slot, which has been obtained in the previous section. It is noted that no cell loss occurs at sources under discipline B, if at most one cell arrives at each source at a slot.

Average Number of Lost Cells at the Source Queue

Using the steady state probability vector $\boldsymbol{P}$ derived in the previous subsection, we first obtain the average number of lost cells, $R_{loss1,j,t}$, at the t-th slot of a super-slot at the j-th source, and then get the average number of lost cells, $R_{loss1,j}$, at an arbitrary slot.

In order to obtain $R_{loss1,j,t}$, we define the state probability at the beginning of a super-slot on the condition that the queue length D_2 of the server is equal to b, as follows;

$$\begin{aligned}\pi_{b,0}(k_1, s_1, \cdots, k_R, s_R) &\triangleq \Pr(D_{11} = k_1, S_1 = s_1, \cdots, D_{1R} = k_R, S_R = s_R | D_2 = b) \\ &= P(b, k_1, s_1, \cdots, k_R, s_R) / \Pr(D_2 = b), \end{aligned} \tag{2.15}$$

and introduce its vector representation,

$$\boldsymbol{\Pi}_{b,0} \triangleq [\cdots, \pi_{b,0}(k_1, s_1, \cdots, k_R, s_R), \cdots]. \tag{2.16}$$

Note that $\Pr(D_2 = b) = \boldsymbol{P}_b \boldsymbol{e}(q)$ in which q indicates the dimension of vector $\boldsymbol{P}_b$. The conditional state probability vector, at the beginning of the t-th slot of a super-slot, $\boldsymbol{\Pi}_{b,t}$, can be given in terms of $\boldsymbol{\Pi}_{b,0}$ as

$$\boldsymbol{\Pi}_{b,t} = \boldsymbol{\Pi}_{b,0} \boldsymbol{G}_b^t, \qquad (1 \le t \le 2\tau), \tag{2.17}$$

where $\boldsymbol{G}_b \triangleq \sum_{i=0}^{R} \boldsymbol{G}_b(i)$, in which $\boldsymbol{G}_b(i)$ is given in Eq.(2.7).

Noting that $\alpha_j(b)$ at the beginning of a super-slot is used over it, we can derive $R_{loss1,j,t}$ for the j-th source controlling its traffic under discipline A as follows;

$$\begin{aligned} R_{loss1,j,t} &= \sum_{\sigma} \Pr(D_2 = b) \pi_{b,t}(k_1, s_1, \cdots, k_R, s_R) \\ &\times \left\{ \alpha_j(b) \sum_{i=K_j+2-k_j}^{N_{s_j,j}} a_{s_j,j}(i) \left[i + k_j - (K_j + 1) \right] \right. \\ &\left. + [1 - \alpha_j(b)] \sum_{i=K_j+1-k_j}^{N_{s_j,j}} a_{s_j,j}(i)(i + k_j - K_j) \right\}, \end{aligned} \tag{2.18}$$

where σ represents all combinations of $(b, k_1, s_1, \cdots, k_R, s_R)$ for b, k_i and s_i that satisfy $0 \le b \le B$, $0 \le k_i \le K_i$ and $s_i = 1, 2$ $(1 \le i \le R)$. In the above equation, the first (the

second) term in the braces corresponds to the case where the j-th source is (not) permitted to transmit its cell.

On the other hand, under discipline B (namely, HOL cell pushout scheme), the j-th source transmits its HOL cell regardless of the value of $\alpha_j(b)$ when newly arriving cells cause buffer overflow. Thus $R_{loss1,j,t}$ is simply given under discipline B by

$$R_{loss1,j,t} = \sum_{\sigma} \Pr(D_2 = b)\pi_{b,t}(k_1, s_1, \cdots, k_R, s_R) \sum_{i=K_j+2-k_j}^{N_{s_j,j}} a_{s_j,j}(i)\left[i + k_j - (K_j + 1)\right]. \tag{2.19}$$

Finally, we can obtain $R_{loss1,j}$ as follows;

$$R_{loss1,j} = \frac{\sum_{t=0}^{2\tau} R_{loss1,j,t}}{2\tau + 1}. \tag{2.20}$$

Average Number of Lost Cells at the Server Queue

We obtain below the average number of lost cells, $R_{loss2,t}$, at the t-th slot of a super-slot at the server queue. Suppose that the queue length of the server is equal to x at the end of the $t-1$-st slot; its associated probability matrix is $\boldsymbol{M}_{b,t-1}(x)$. At the t-th slot, y cells arrive at the server from all the sources with probability $\boldsymbol{G}_b(y)$. In this case, one of the x cells is transmitted, and $B+1-x$ of the y cells are then stored in the server buffer. However, other $x+y-(B+1)$ cells are lost. Accordingly, we have $R_{loss2,t}$ given by

$$R_{loss2,t} = \begin{cases} \sum_{b=B+2-R}^{B} \sum_{y=B+2-b}^{R} \boldsymbol{P}_b \boldsymbol{G}_b(y) \boldsymbol{e}(q) \{b + y - (B+1)\} \\ \qquad \text{if } t = 0, \\ \sum_{b=0}^{B} \sum_{x=B+2-R}^{B} \sum_{y=B+2-x}^{R} \boldsymbol{P}_b \boldsymbol{M}_{b,t-1}(x) \boldsymbol{G}_b(y) \boldsymbol{e}(q) \{x + y - (B+1)\} \\ \qquad \text{if } 1 \le t \le 2\tau. \end{cases} \tag{2.21}$$

The average number of lost cells at each slot at the server, R_{loss2}, is given by

$$R_{loss2} = \frac{\sum_{t=0}^{2\tau} R_{loss2,t}}{2\tau + 1}. \tag{2.22}$$

Overall system cell loss probability

By using Eqs. (2.20) and (2.22), we finally obtain the overall system cell loss probability as

$$P_{loss} = \frac{\sum_{j=1}^{R} R_{loss1,j} + R_{loss2}}{\sum_{j=1}^{R} \Lambda_j}, \tag{2.23}$$

where Λ_j is the arrival rate of the arriving traffic at the j-th source, which has been defined in Eq.(2.1).

2.3.4 Average Waiting Time

Let $P_{s_{1j},t}(k_j)$ and $P_{s_2,t}(b')$ denote the probability that the queue length at the j-th source and the server are equal to k_j and b', respectively, at the t-th slot of a super-slot. By using $\boldsymbol{M}_{b,t}(b')$ given by Eq.(2.10) and $\pi_{b,t+1}(k_1, s_1, \cdots, k_R, s_R)$ given by Eq.(2.17), we obtain these probabilities as follows.

$$\begin{aligned} P_{s_{1j},t}(k_j) &= \sum_{\sigma_{k_j}} \Pr(D_2 = b)\pi_{b,t+1}(k_1, s_1, \cdots, k_j, s_j, \cdots, k_R, s_R) \qquad (1 \le j \le R), \\ P_{s_2,t}(b') &= \sum_{b=0}^{B} \boldsymbol{P}_b \boldsymbol{M}_{b,t}(b')\boldsymbol{e}(q), \end{aligned} \tag{2.24}$$

where σ_{k_j} represents all combinations of $(b, k_1, s_1, \cdots, k_{j-1}, s_{j-1}, s_j, k_{j+1}, s_{j+1}, \cdots, k_R, s_R)$ for b, k_i and s_i that satisfy $0 \le b \le B$, $0 \le k_i \le K_i$ $(i = 1, \cdots, j-1, j+1, \cdots, R)$ and $s_i = 1, 2$ $(1 \le i \le R)$. Note that $\pi_{b,t+1}(k_1, s_1, \cdots, k_R, s_R)$ is the state probability at the beginning of the $t+1$-st slot, and that it can be regarded as the state probability at the end of the t-th slot as well. Hence, we have the average queue length at the j-th source and the server, denoted by $L_{q,1j}$ and $L_{q,2}$, as follows.

$$\begin{aligned} L_{q,1j} &= \frac{\sum_{t=0}^{2\tau}\sum_{k_j=0}^{K_j} k_j P_{s_{1j},t}(k_j)}{2\tau+1} \quad (1 \le j \le R), \\ L_{q,2} &= \frac{\sum_{t=0}^{2\tau}\sum_{b'=0}^{B} b' P_{s_2,t}(b')}{2\tau+1}. \end{aligned} \tag{2.25}$$

By using Little's formula, we obtain the total average waiting time at the j-th source and the server.

$$W_j = \frac{L_{q,1j}}{\Lambda_j - R_{loss1,j}} + \frac{L_{q,2}}{\sum_{i=1}^{R}(\Lambda_i - R_{loss1,i}) - R_{loss2}}. \tag{2.26}$$

2.4 Numerical Results and Discussions

In this section, we investigate the effectiveness of reactive congestion control schemes by means of numerical results, which can be obtained by using our analytical method presented in the previous section. We can study various types of reactive congestion control schemes in a way to determine $\alpha_j(b)$ appropriately according to their schemes.

2.4.1 Numerical Parameters

The numerical results presented here are obtained under the following conditions.

- source nodes
  Sources are homogeneous in terms of the input traffic characteristics, buffer size and the function $\alpha_j(b)$. Namely, for any j such that $1 \leq j \leq R$,

$$(\Lambda_{1j}, \Lambda_{2j}, r_{1j}, r_{2j}) = (\lambda_1, \lambda_2, r_1, r_2), \quad \Lambda_j = \lambda,$$
$$\beta_j = \beta, \quad L_{b_j} = L_b, \quad N_{lj} = \mathcal{N}_l, \ (l = 1, 2),$$
$$K_j = K, \quad \alpha_j(b) = \alpha(b).$$

  Now, we define the total arrival rate at the server, λ_{all}, and that becomes equal to $R\lambda$.

- cell arrival model
  The Interrupted Bernoulli Process (IBP) is used as a cell arrival process to sources for simplicity of discussion although the analytical model can treat more general processes, MMBP. Namely, $\mathcal{N}_1 = 1$, $\mathcal{N}_2 = 0$ and $\lambda_2 = 0$, i.e., a cell arrives with probability λ_1 during state 1, and no cell arrives during state 2.

We will show the sensitivity of the cell loss probability to the reactive congestion control algorithms. For this purpose, we consider a rather simple algorithm, in which cases $\alpha(b)$ takes only two values, i.e., 0 and 1. Namely, each source can emit its HOL cell until the queue length of the server, b, exceeds some threshold denoted by H. This type of control is below referred to as the *binary control*, BC for short. The binary control is of practical interest from an implementation point of view because the ECN sent by the server is a simple binary information, i.e., choking or relieving signal (or cell).

BC1: $\alpha(b) = \begin{cases} 1 & b \leq H-1 \\ 0 & H \leq b \leq B \end{cases}$

BC2: a combination of BC1 and pushout discipline (discipline B, see Section II)

Numerical results in the case without any congestion control are also provided just for comparison; this case is denoted by NC (No control).

2.4.2 Numerical Results

First, we verify the accuracy of our analysis by comparing numerical results with simulation ones, and next discuss the optimum threshold H_{opt} at which the overall cell loss probability, P_{loss}, is minimized under BC1 and BC2. Finally, we investigate the effectiveness of the reactive congestion control schemes.

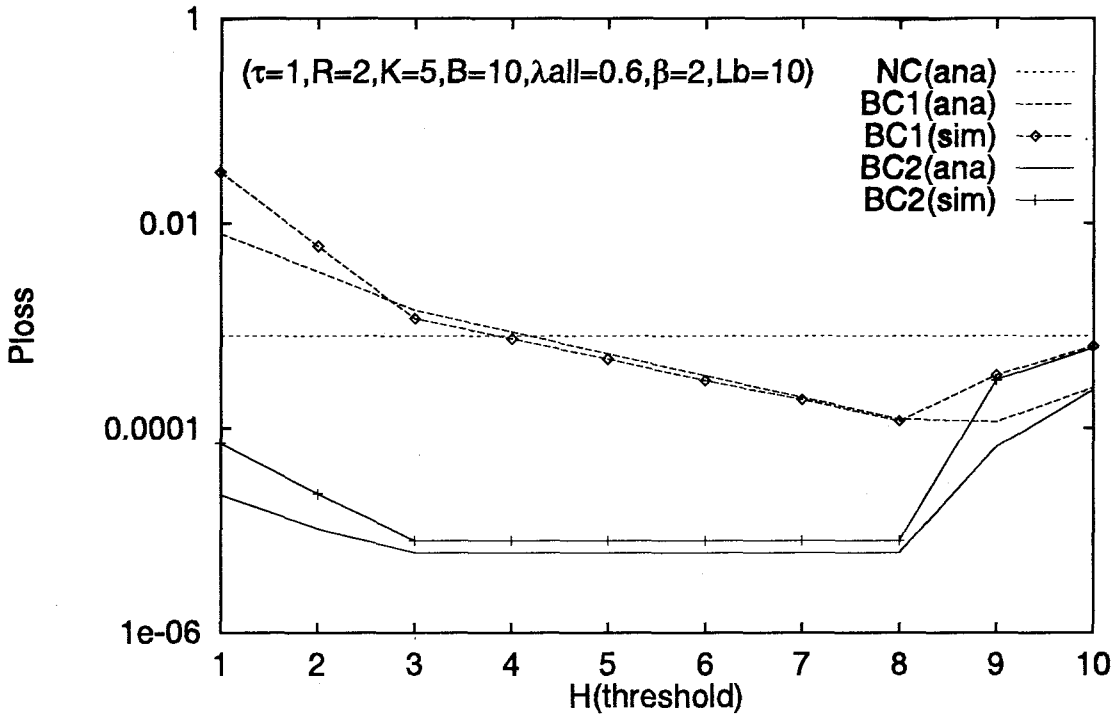


Figure 2.3: Cell loss probability as a function of H: comparison between analysis and simulation

Optimum threshold H_{opt}

Figure 2.3 illustrates the cell loss probability as a function of the threshold H. We fix the buffer size of the server $B = 10$, that of each source $K = 5$, the number of sources $R = 2$, and propagation delay $\tau = 1$. With regard to input traffic characteristics of each source, we fix the input rate $\lambda = 0.3$, the burstiness $\beta = 2$, and the burst-length $L_b = 10$; the total arrival rate at the server λ_{all} is thus set to 0.6. In this figure, a comparison between analytical results and simulation ones is also shown.

We first see from this figure that analytical results agree with simulation ones very well over a wide range of H. Furthermore, the optimum threshold which minimizes P_{loss} is equal to 8 in BC1 from the analytical result. If H is set to a value smaller than 8, it may happen that sources are not allowed to transmit their HOL cells to the server in spite of the fact that the server queue has some space to store arriving cells. This results in increasing cell loss at sources. On the other hand, if H is larger than 8, more cells from each source will be dropped at the server queue. First, focusing on cell loss in the server, we will consider the condition on the optimum threshold. Since one-way propagation delay is τ slots, sources keep sending cells to the server during $(2\tau + 1)$ slots even after the server sends a choking signal (or cell (see Fig.2.2)). As a result, at most $(2\tau + 1) \times R$ cells arrive at the server in this period, while the server can forward $2\tau + 1$ cells in the same period. The server should thus send the ECN signal to each source when its queue length exceeds $B - (2\tau + 1)(R - 1)$ in order to prevent cell loss at the server. Finally, the optimum threshold for the server is $B + 1 - (2\tau + 1)(R - 1)$ or less. On the contrary, larger H is better for the sources because they are less frequently prevented from sending cells. The optimum threshold thus depends on the cell loss performance at both the server and the sources. When cell loss at the server is dominant in the total cell loss, the optimum threshold can be given as follows:

$$H^*_{opt} = B + 1 - (2\tau + 1)(R - 1), \quad \text{in case of BC1.} \tag{2.27}$$

By using this H^*_{opt}, P_{loss} is improved about one order of magnitude compared with that under NC.

Suppose that, under BC2, the binary control with pushout scheme, the server sends the choking signal when its queue length exceeds H^*_{opt} given in the above equation. For example, cell loss will occur if all R sources send their cells during $2\tau + 1$ slots until they receive the choking signal and two cells or more are further pushed out in the following slot. In addition, it is noted that the sources push out their HOL cells to the server only when i) their buffers are filled up, and ii) newly cells arrive at them. Thus, a sequence of these events will happen with very small probability. In fact, BC2 drastically reduces the cell loss probability about two orders of magnitude compared with NC. Moreover, BC2

successfully minimizes P_{loss} over a wide range of H. If H is set to a value smaller than H^*_{opt} for BC1, the pushout occurs more often on each source, and it will be more likely that pushed-out cells enter the server buffer because of the larger space remained at the server. Therefore, we can expect that P_{loss} is almost the same value as that in the case of $H = H^*_{opt}$. However, when a value smaller than $2\tau + 1$ is chosen as H, the server queue can become empty after relieving cells are transferred to sources. As a result, the server utilization decreases, thereby increasing the cell loss probability as shown in Fig. 2.3(see the cases of $H = 1,\ 2$). Therefore, the range of optimum threshold of BC2 can be given as follows;

$$2\tau + 1 \leq H_{opt} \leq B + 1 - (2\tau + 1)(R - 1), \quad \text{in case of BC2} \tag{2.28}$$

It is worth noting that a wider range of the optimum threshold H_{opt} is available under BC2 than under BC1. Therefore, the threshold can be determined based upon a rather rough estimation on propagation delay between the server and the sources. This is a very desirable feature because it is often very difficult for the server to know exactly the propagation delay.

We investigate the impact of the source buffer size K in Fig. 2.4, where we set the parameters as in Fig. 2.3 except for λ_{all}, which is set to 0.5. It is shown that both BC1 and BC2 reduce the cell loss probability with larger K, and that BC2 is further superior to BC1 in terms of cell loss performance. We observe that the range of the optimum threshold in BC2 does not vary with K, while the optimum threshold H_{opt} in BC1 is larger than H^*_{opt} given in Eq. (2.27) when K is small. This implies that H_{opt} in BC1 can be affected by the buffer size K.

The impact of the propagation delay

Figure 2.5 shows the cell loss probability as a function of one-way propagation delay τ. In this figure, H is chosen in a manner to minimize P_{loss} in BC1 and BC2; the resulting P_{loss} is denoted by $P_{loss}(H_{opt})$, in what follows. Other parameters are set as in Fig. 2.3 except for τ. The comparison between analytical and simulation results are also shown. The analytical results there agree with the simulation ones very well, as in Fig. 2.3. When τ is larger than 2, there exists no H satisfying Eq. (2.27) and inequality (2.28). In such cases, BC1 is not effective in reducing the cell loss probability, while BC2 is still effective there.

Figure 2.6 illustrates the impact of the server buffer size B in NC and BC2. We see from this figure that we can gain almost the same improvement in terms of cell loss reduction for various values of B, when H_{opt} given in inequality (2.28) is available.

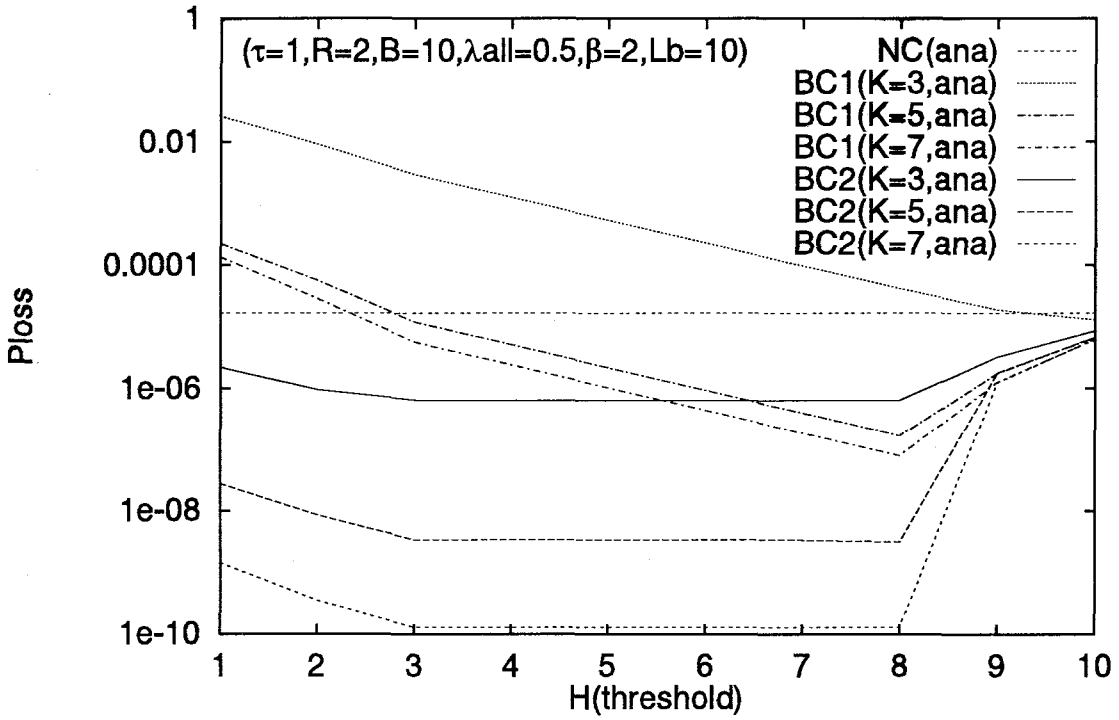


Figure 2.4: Impact of the buffer size of sources, K

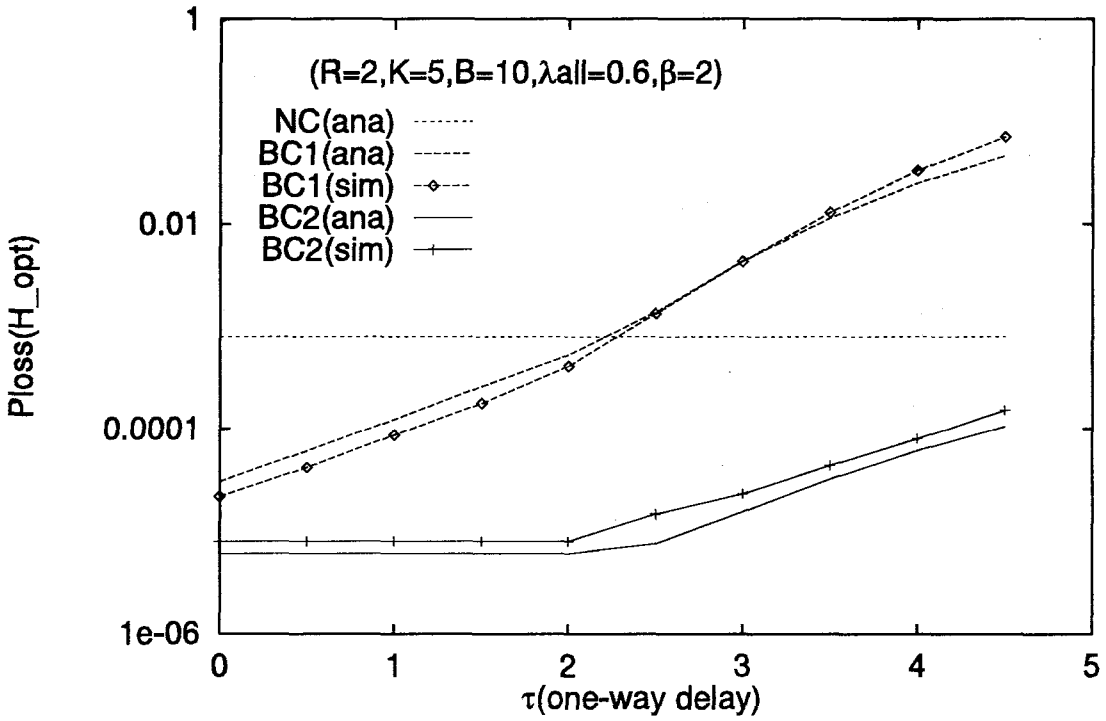


Figure 2.5: Cell loss probability as a function of τ: comparison between analysis and simulation

The impact of the traffic characteristics

Figure 2.7 shows the sensitivity of $P_{loss}(H_{opt})$ to the total arrival rate at the server λ_{all}. Under BC1, the optimum threshold H_{opt} is larger than H^*_{opt} given in Eq. (2.27) when λ_{all} is large. The improvement due to the congestion control such as BC1 and BC2 becomes larger as λ_{all} decreases. Supposing that the quality of service (QOS) in terms of the cell loss probability is less than 10^{-6}, we see that, under BC2, the maximum amount of traffic that the server can accommodate increases up to about 0.57 from 0.44, while that under BC1 increases up to 0.52.

We show the sensitivity of $P_{loss}(H_{opt})$ to the input traffic burstiness β in Fig. 2.8. From Eqs. (2.1) and (2.2), β is obtained as follows;

$$\beta = \frac{\lambda_1}{\lambda} = \frac{r_1 + r_2}{r_1}. \tag{2.29}$$

Keeping λ_{all} constant, we vary β by changing λ_1 and either r_1 or r_2, both of which denote the phase transition probabilities of the input traffic. In this figure, *ON-fixed (OFF-fixed)* denotes the case where r_2 (r_1) is varied. In the case of ON-fixed, the cell loss probability increases monotonously as β becomes larger. On the other hand, in the case of OFF-fixed, there is a plateau in terms of P_{loss}; it is noted that the average ON period decreases here as β increases.

We also present the sensitivity of $P_{loss}(H_{opt})$ to the burst-length L_b in Fig. 2.9. Again, from this figure, it is shown that BC2 improves the performance over a wide range of L_b. Suppose here that the QOS with regard to P_{loss} is less than 10^{-6}. BC2 can accept the input traffic of the burst-length which is four times as long as that allowed by NC.

So far, we have focused on the situation where there are only two sources in the system. Figure 2.10 depicts the cell loss probability in cases with three and four sources. This figure demonstrates that the reactive control will work well even in a system consisting of a larger number of sources; i.e., the effectiveness of the control algorithms is less sensitive to the number of sources given a total traffic rate. However, it is worth noting that, from inequality (2.28), the range of the optimum threshold H_{opt} depends on the number of sources R.

Average waiting time

Figure 2.11 shows the total average waiting time spent at the source and the server. Cells are kept waiting in the buffer for a longer duration under congestion control than under no congestion control. The average waiting time is larger in BC2 than in BC1 because pushed-out cells entering the server increases the server queue length; this additional waiting time

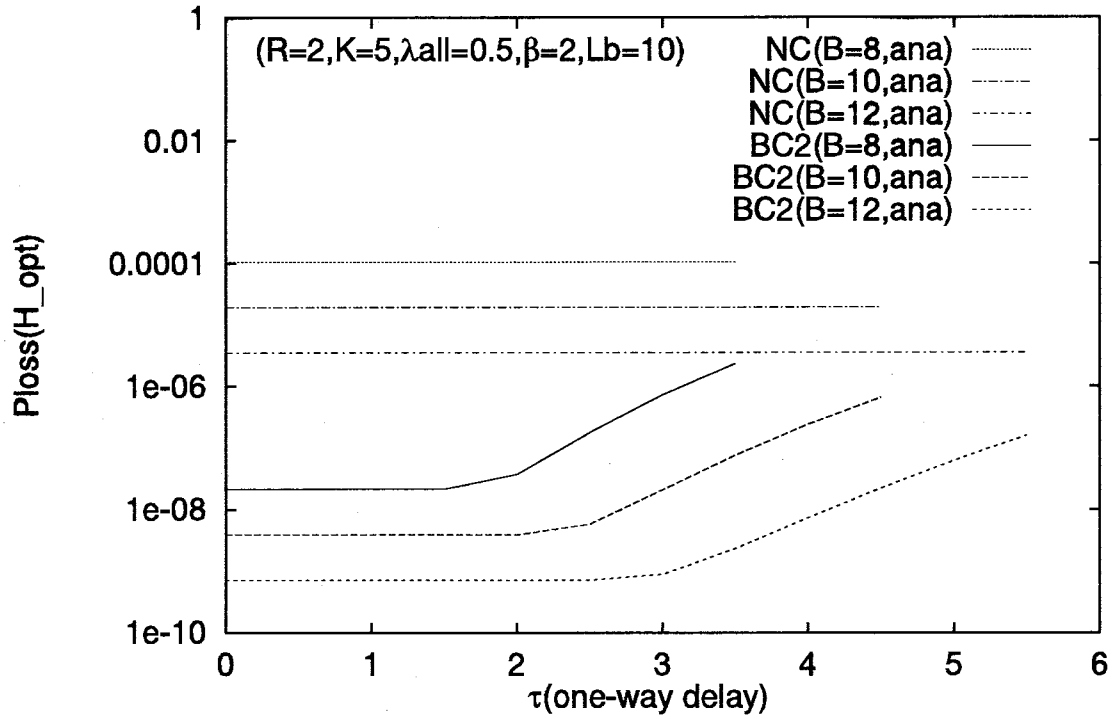


Figure 2.6: Impact of the buffer size of server, B

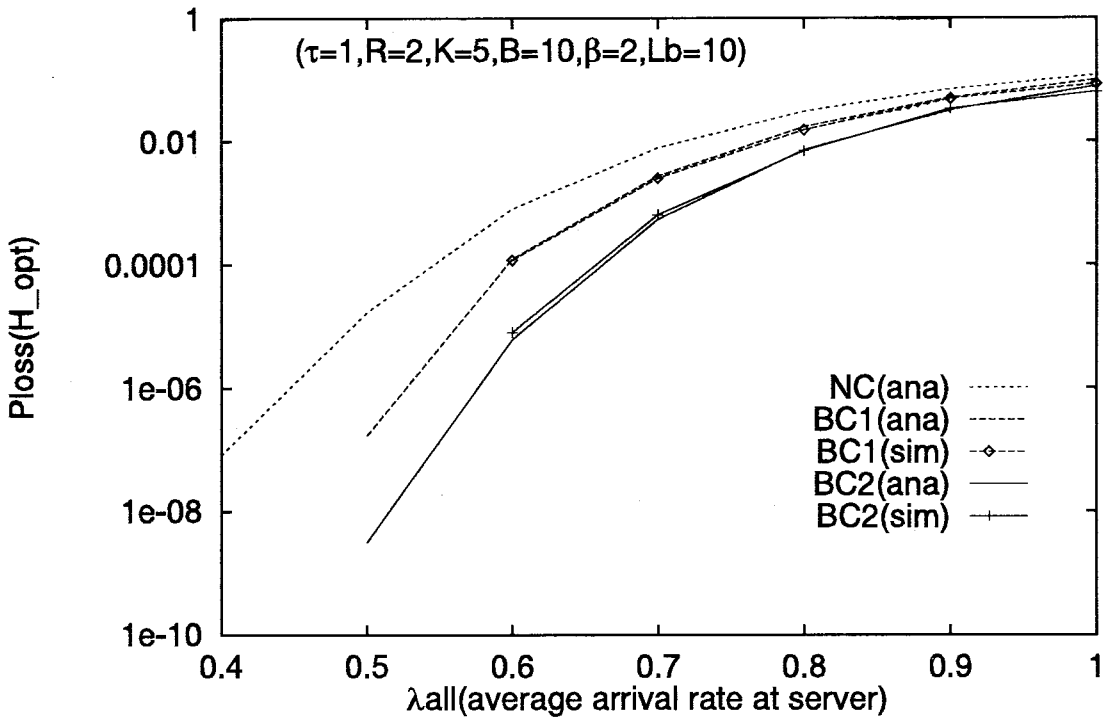


Figure 2.7: Cell loss probability versus total offered traffic rate at server

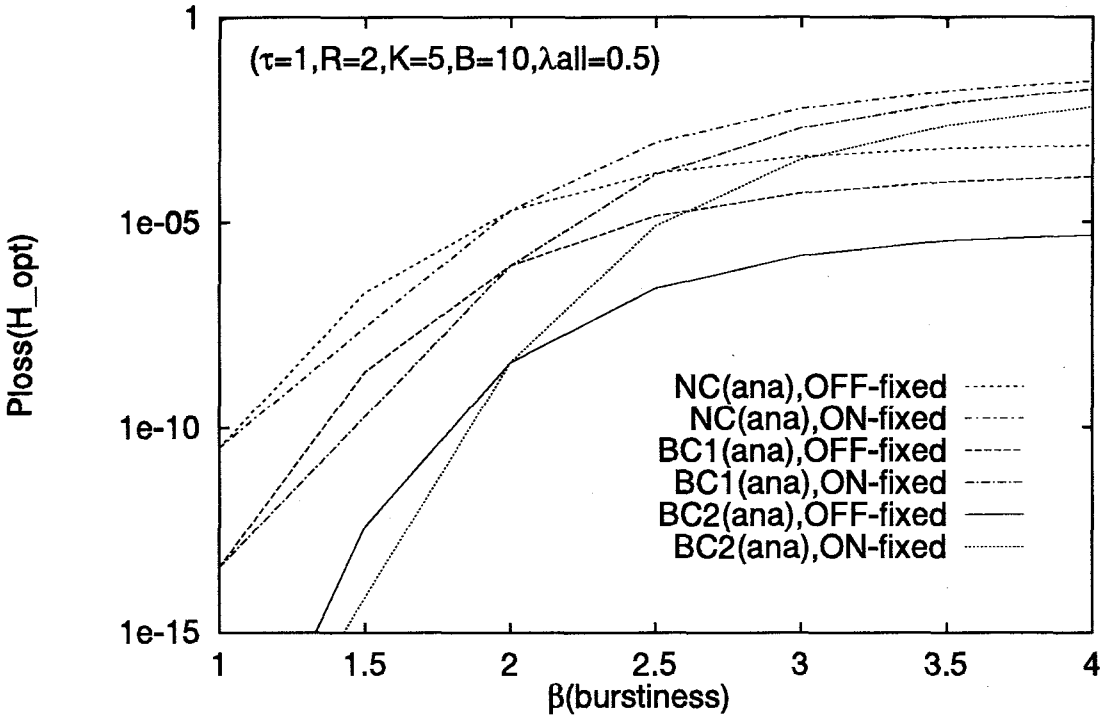


Figure 2.8: Cell loss probability versus traffic burstiness

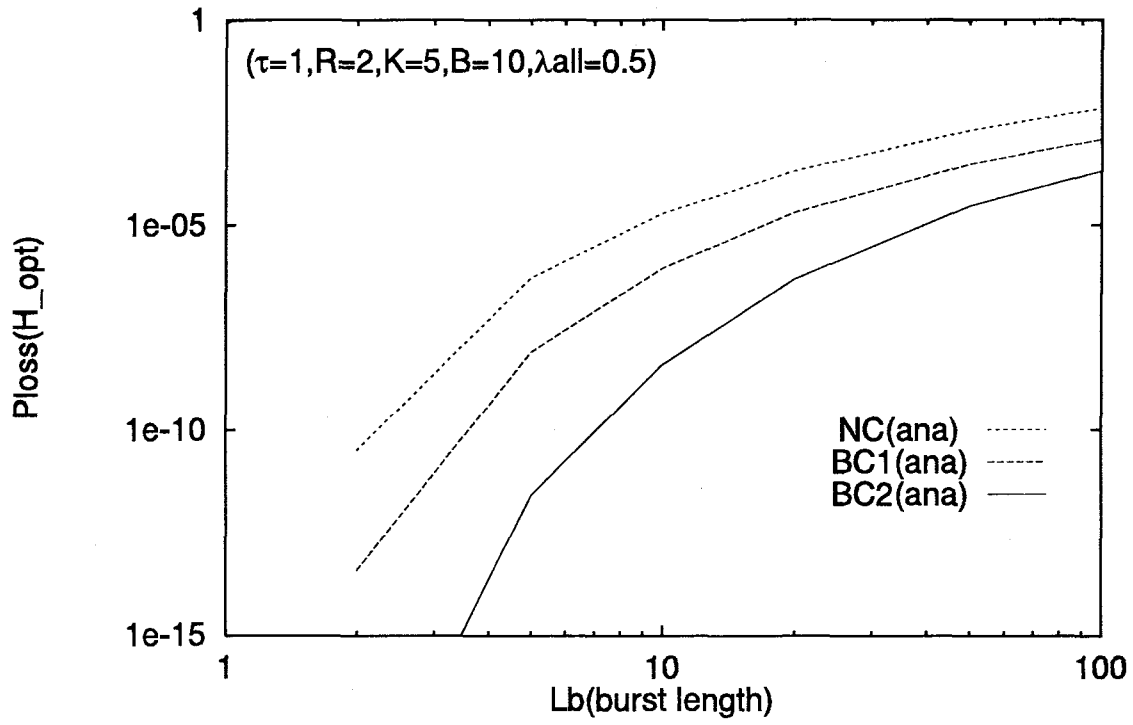


Figure 2.9: Cell loss probability versus traffic burst-length

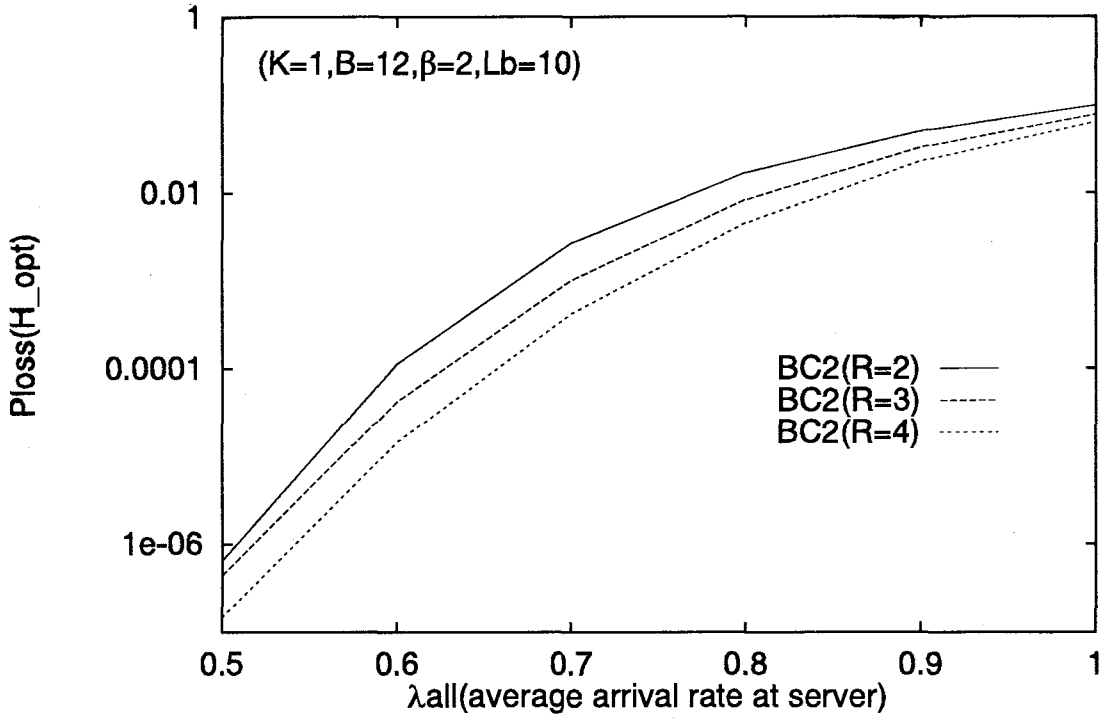


Figure 2.10: Cell loss probability versus total offered traffic rate: cases of 2,3 and 4 sources in case of BC2

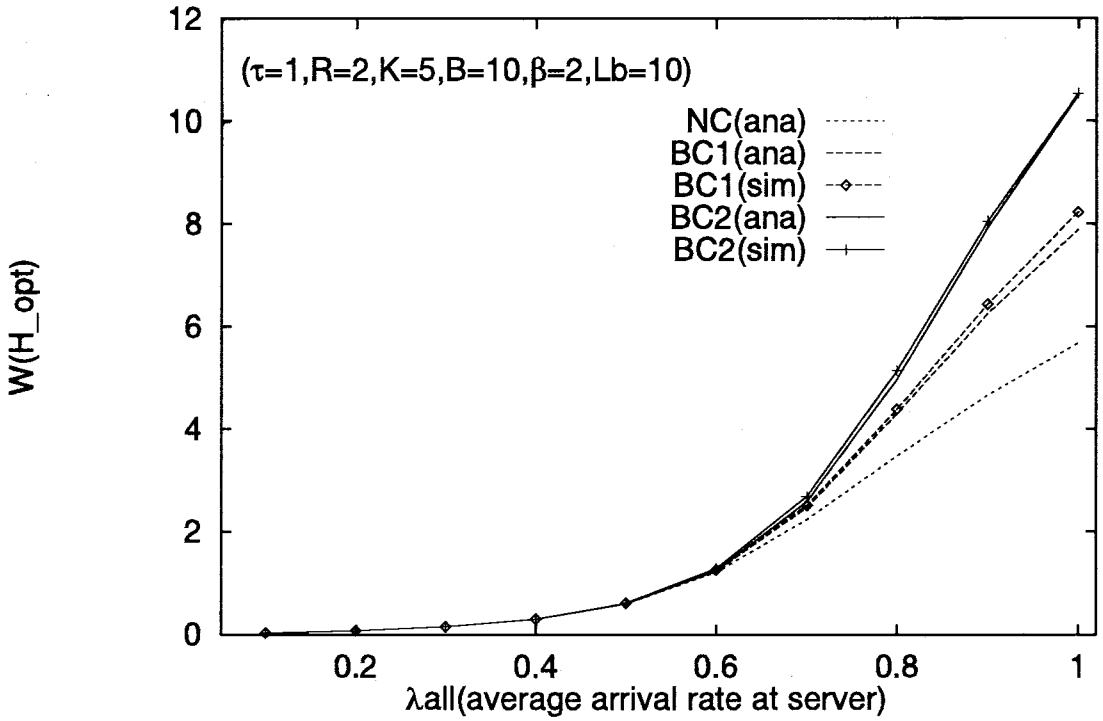


Figure 2.11: Average waiting time versus total offered traffic rate

in BC2 is traded for cell loss improvement, as shown in several figures. We see that the average waiting time is kept within an acceptable level for nonreal-time applications over a wide range of λ_{all}.

2.5 Conclusions

An analytical model has been developed to provide insights into the benefits obtained through reactive congestion control and its limitations in ATM LANs with relatively low propagation delay. Our model is exact in cases without propagation delay, and is approximate in cases of non-zero propagation delay. As illustrated in Fig. 2.1, our analytical model consists of several sources and the server, which represents a congested node. The MMBP is employed as a cell arrival process at the source, which enables us to handle bursty traffic.

By the use of our model, we carried out an approximate analysis of two types of reactive congestion control algorithms. Numerical results agree with simulation ones, and indicate that a significant reduction in cell loss probability can be achieved with the reactive congestion control, in particular, when propagation delay and utilization are rather low, like in ATM LANs. The results obtained here are summarized below.

1. The binary congestion control algorithm combined with pushout discipline (BC2) provides better performance than the binary control (BC1).

2. We can determine the optimum threshold, H_{opt}, of the queue length in a congested node that minimizes the overall cell loss probability, P_{loss}, quantitatively, under both BC1 and BC2. Under BC2, it is a function of the number of sources R and propagation delay τ, and can be determined based upon a rather rough estimation of R and τ. As a result, we see that BC2 is better than BC1 in that it is easier to determine H_{opt} under BC2 than under BC1.

3. The statistical gain can be increased by the reactive congestion control algorithms over a wide range of the amount of offered traffic, the burstiness and the burst-length. For example, in the case treated in Fig. 2.7, BC2 successfully increases the maximum amount of traffic that can be accommodated by the server up to about 0.57 from 0.44 under no congestion control, keeping the cell loss probability less than 10^{-6}.

4. The reactive control is effective even in a system with a large number of sources.

Chapter 3

Performance Analysis of Backpressure Congestion Control

3.1 Introduction

Recently, ATM switching technology has been studied and developed as a way of realizing B-ISDN which can support multimedia traffic including voice, video and data. In the ATM network, information is divided into short, fixed length *cells*. Within the network, cells are switched in a statistical multiplexing manner like a conventional packet switched network to facilitate network resources efficiently. Therefore, cell loss occurs due to congestion at intermediate nodes in the network unless some appropriate congestion control scheme is employed.

In the ATM network, the preventive congestion control is an effective way to prevent congestion, which is implemented by Connection Admission Control (CAC) and Usage Parameter Control (UPC). By CAC, the number of connections on each link is limited based upon the traffic descriptors of the incoming traffic and the required QOS (Quality of Service) parameters. Once the connection is admitted by CAC, the cell stream from the user is monitored by UPC to decide whether the incoming traffic is conformed to the declared traffic parameters or not. However, it is often the case that users cannot have enough information with regard to its traffic descriptors before transmitting its own traffic. Thus, it would be insufficient to only apply preventive congestion control for relieving congestion, and another device for congestion control is required; that is the reactive congestion control [GERS 91, NEWM 93, IKE 94]. The reactive control together with preventive one can avoid short-term congestion at the cell level as shown in [GERS 91]. Examples of reactive control schemes proposed for ATM networks are the Explicit Congestion Notification (ECN) schemes and the link-by-link backpressure (or BP in short) control [IKE 94].

In this paper, we pay our attention to BP control because of its following desirable property. In BP method, the congested node informs the upstream nodes of its congestion status to choke/relieve cell transmission. The feedback delay of congestion information is therefore ameliorated, and BP control method becomes useful in ATM WAN [VICK 94]. On the contrary, as mentioned in [JAIN 92], the BP control can bring the whole network to congestion and make cells not traveling the congested node adversely affected if the overload is of long duration whereas it works well in short-lived congestion. Furthermore, in [KOLA 94], Kolarov and Ramamurthy have compared the performance of the combined method of FECN/BECN and BP with that in each case of FECN, BECN and BP. They showed that the combination of ECN and BP can offer the superior performance, but also suggested that the Head-Of-Line (HOL) blocking caused by BP leads to significant performance degradation in WAN.

We will analyze the BP control by applying the approach developed in Chapter 2 consisting of nodes in two stages. Nevertheless, we can examine the fundamental properties related with the BP control through the analysis as mentioned later. In particular, the cell push-out scheme is focused on as a possible way to overcome the drawback of the BP control mentioned above. Even if a node is prohibited from transmitting a cell by the BP signal from its downstream congested node, the cell push-out scheme allows the (upstream) node to send its cell when its queue length reaches some predetermined threshold. By means of the analysis, we examine how the cell push-out scheme efficiently relieves the HOL-blocking effect due to the BP control, and how it successfully enables cells not transmitting a congested node to travel its upstream nodes without blocked.

The remainder of the paper is organized as follows. In Section 3.2, we describe our analytical model and explain the cell push-out scheme considered in this Chapter. Section 3.3 provides an analysis for deriving the cell loss probability and the average waiting time. We also give numerical results and discussions in Section 3.4, and Section 3.5 presents concluding remarks.

3.2 Analytical Model

In this section, our queueing model for analyzing the backpressure (BP) congestion control method is described. We consider a discrete time queueing system with time slot being equal to a ATM cell transmission time. Our model consists of two-stage queues with finite capacity in series, as illustrated in Fig. 3.1. We intend to model the upstream nodes by queues in the first stage and the congested node by one of queues in the second stage. Other queues in the second stage are downstream nodes which do not experience congestion. In

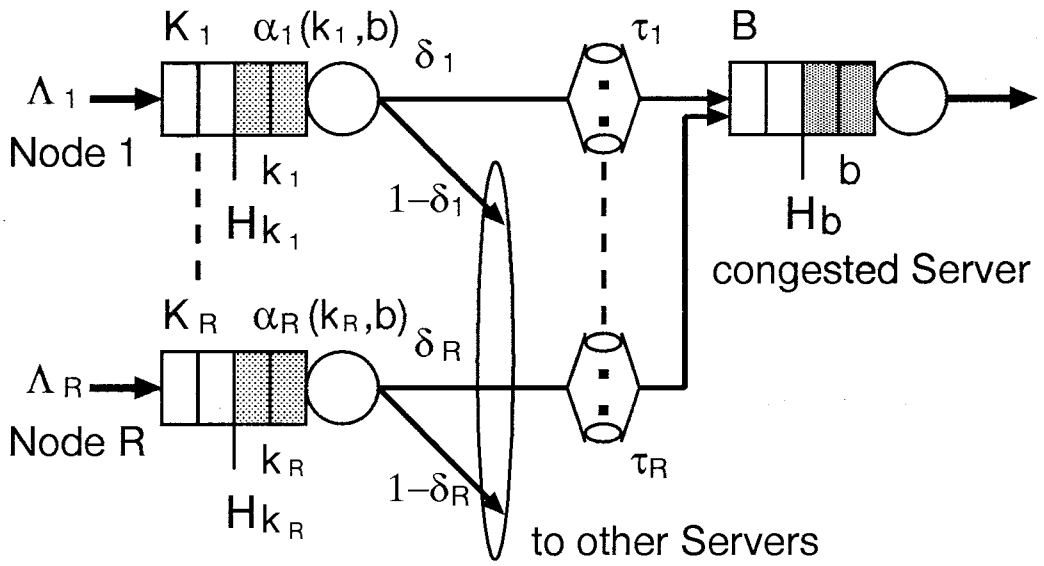


Figure 3.1: Analytical model for back-pressure congestion control

the first stage, there are R nodes in parallel, which will be called *nodes* for simplicity. An incoming cell at each node is transmitted to one of the nodes in the second stage, which will be referred to as the *server*. We will introduce our analytical model followed by our BP control methods in Subsection 3.2.1. In Subsection 3.2.2, the cell push-out scheme is also taken into account, which is our main subject of the current paper.

3.2.1 Assumption

We consider that the upstream nodes are requested to stop cell transmission when congestion occurs at the congested server (see Fig. 3.1). In what follows, we will refer to this node as a *congested server*, which is equipped with a buffer of B in length. On the other hand, at the *other servers*, the congestion does not take place. Therefore, the congestion notification is never received from those servers.

It is assumed that the j-th ($1 \leq j \leq R$) node has a cell buffer of K_j in length and cells arrive at each node according to a two-state Markov-Modulated Bernoulli Process (MMBP), which is characterized by a set of parameters, $(\Lambda_{1j}, \Lambda_{2j}, r_{1j}, r_{2j})$. In the l-th ($l = 1, 2$) state of MMBP, N_{lj} cells or less arrive according to a Bernoulli Process at each slot. Thus, the average arrival rate at the j-th node queue is given by

$$\Lambda_j = \frac{r_{2j}}{r_{1j} + r_{2j}}\Lambda_{1j} + \frac{r_{1j}}{r_{1j} + r_{2j}}\Lambda_{2j} \quad (1 \leq j \leq R). \tag{3.1}$$

The cell stored at the head of the j-th node queue is transmitted to the congested server with probability δ_j and other servers with probability $1 - \delta_j$. We take into account propagation delays, denoted by τ, between the j-th node and the congested server.

3.2.2 Definition of Control Scheme

As shown in Fig. 3.1, we introduce a control function $\alpha_j(k_j, b)$ for the j-th node, which is a function of both the queue length, b, of the congested server and that, k_j, of its own buffer length, and takes values ranging from 0 to 1. At each slot, by using $\alpha_j(k_j, b)$ as a probability of transmitting a cell, the j-th node determines whether it transfers the cell in its buffer or not. We define $\alpha_j(k_j, b)$ in compliance with some control schemes as follows.

Back-Pressure Control (BP1)

Each node can emit its HOL cell until the queue length of the congested server, b, exceeds some threshold value (denoted by H_b), i.e.,

$$\alpha_j(k_j, b) = \alpha_j(b) = \begin{cases} 1, & \text{if } 0 \leq b < H_b \\ 0, & \text{if } H_b \leq b \leq B \end{cases} .$$

The value of H_b can be determined to completely avoid cell loss at the congested server as shown in Chapter 2, and is given by the following equation:

$$H_b = B + 1 - (2\tau + 1)(R - 1), \tag{3.2}$$

where B, R and τ are the buffer size of the congested server, the number of the nodes sharing the congested server, the propagation delay between each node and the congested server, respectively.

Back-Pressure with Push-out (BP2)

If the congestion lasts a rather long duration at the node, the BP causes following problems [JAIN 92]:

1. It may bring the whole network to a standstill.

2. Cells which do not go through the congested servers are adversely affected.

In order to overcome the above weaknesses of the BP control, we introduce the cell push-out mechanism. The j-th node, which cannot transfer its HOL cell due to the backpressure signal from the congested node, can *push* it *out* whenever the queue length of itself exceeds some threshold value, H_{k_j}, and accept newly arriving cells. See Fig. 3.2. We should note here that the previously proposed push-out schemes (see, e.g., [HEBU 90, CHAN 94, LEME 94]) are a sort of space priority mechanisms regarding a queue management method for two priority classes. On the other hand, the push-out mechanism, in this paper, is used as a means of avoiding the performance degradation caused by the HOL blocking of cells in the upstream nodes. In this context, the control function $\alpha_j(k_j, b)$ is defined as follows;

$$\alpha_j(k_j, b) = \begin{cases} 1, & \text{if } 0 \leq b < H_b \text{ or } (H_{k_j} \leq k_j \leq K_j \text{ and cells arrive}) \\ 0, & \text{if } H_b \leq b \leq B \text{ and } \{(0 \leq k_j < H_{k_j}) \text{ or} \\ & \quad (H_{k_j} \leq k_j \leq K_j \text{ and no cells arrive})\} \end{cases} .$$

No Control (NC)

Numerical results in the case without any congestion control are also provided for comparison purposes.

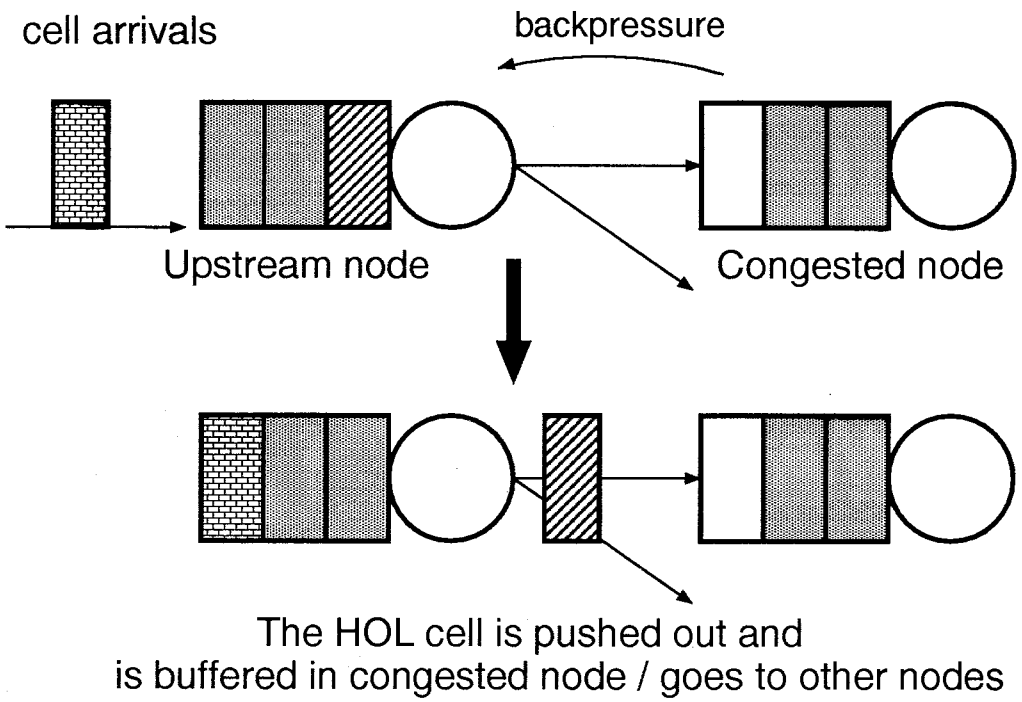


Figure 3.2: Push-out scheme in the analysis

3.3 Steady State Probability

3.3.1 Definition of State Probability

The state of our system can be completely described by a set of 1) the queue length, D_2, of the congested server, 2) the queue length, D_{1j}, of each node, and 3) the MMBP state, S_j, of each node. Therefore, we can derive the steady state probability $P(b, k_1, s_1, \cdots, k_R, s_R) = \Pr(D_2 = b, D_{11} = k_1, S_1 = s_1, \cdots, D_{1R} = k_R, S_R = s_R)$ by using the transition probability matrix $\boldsymbol{Q}$. The matrix can be obtained by following the analytic approach presented in Chapter 2, which is summarized in the below.

1. Define the transition probability matrix per slot, $\boldsymbol{G}_{b,j}(x)$, of the j-th node, when the j-th node transmits x cells to the congested server on the condition that the congested server contains b cells in its buffer.

2. Derive the transition probability matrix $\boldsymbol{G}_b(y)$, when y cells are forwarded from whole sources to the congested server, by the convolution of $\boldsymbol{G}_{b,1}(x)$, $\boldsymbol{G}_{b,2}(x)$, ..., $\boldsymbol{G}_{b,R}(x)$.

3. Construct the matrix $\boldsymbol{L}_{b,b'}$, which is concerning the transition of the queue length of the congested server.

4. Finally obtain $\boldsymbol{Q}$ by using $\boldsymbol{L}_{b,b'}$ as submatrices.

In this Chapter, since we assume that some of cells arriving at each node are not transmitted to the congested server but to other servers, derivation of $\boldsymbol{G}_{b,j}(x)$ in Step 1 is different from the one in Chapter 2 while others are all the same. Therefore, in the remainder of this subsection, we will present the matrix $\boldsymbol{G}_{b,j}(x)$.

First, we obtain the transition matrix $\boldsymbol{G}^*_{b,j}(x)$, when the j-th node transmits x cells, on the condition that the congested server contains b cells in the buffer. Here, we introduce the probability matrix, $\boldsymbol{T}_j$, of state transition and that, $\boldsymbol{a}_{i,j}$, of cell arrival for each MMBP at the j-th node queue, respectively, as follows,

$$\boldsymbol{T}_j = \begin{pmatrix} 1 - r_{1j} & r_{1j} \\ r_{2j} & 1 - r_{2j} \end{pmatrix}, \quad \boldsymbol{a}_{i,j} = \begin{pmatrix} a_{1j}(i) & 0 \\ 0 & a_{2j}(i) \end{pmatrix}, \quad (1 \leq j \leq R). \tag{3.3}$$

Each element, $a_{lj}(i)$ $(l = 1, 2)$ in the above matrix $\boldsymbol{a}_{i,j}$, corresponds to the probability that the number of i cells arrive at the j-th node at each slot when MMBP is in state l, and is given by

$$a_{lj}(i) = \binom{N_{lj}}{i} \left(\frac{\Lambda_{lj}}{N_{lj}}\right)^i \left(1 - \frac{\Lambda_{lj}}{N_{lj}}\right)^{N_{lj}-i} \quad (1 \le j \le R). \tag{3.4}$$

Thus $\boldsymbol{G}^*_{b,j}(x)$ is given by the following equation.

$$\begin{aligned}
\boldsymbol{G}^*_{b,j}(0) &= \overline{\mathcal{A}_{b,j}} \cdot \begin{pmatrix} \frac{\boldsymbol{a}_{0,j}\boldsymbol{T}_j}{(1-\alpha_j(0,b))} & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-1,j}\boldsymbol{T}_j & \sum_{i=K_j}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \boldsymbol{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-2,j}\boldsymbol{T}_j & \sum_{i=K_j-1}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \cdots & \ddots & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \sum_{i=1}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \boldsymbol{0}[2] & \cdots & \cdots & \boldsymbol{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j \end{pmatrix}, \\
\boldsymbol{G}^*_{b,j}(1) &= \mathcal{A}_{b,j} \cdot \begin{pmatrix} \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \boldsymbol{a}_{2,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j,j}\boldsymbol{T}_j & \sum_{i=K_j+1}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \cdots & \boldsymbol{a}_{K_j-1,j}\boldsymbol{T}_j & \sum_{i=K_j}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \boldsymbol{0}[2] & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \boldsymbol{a}_{1,j}\boldsymbol{T}_j & \sum_{i=2}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \\ \boldsymbol{0}[2] & \cdots & \boldsymbol{0}[2] & \boldsymbol{a}_{0,j}\boldsymbol{T}_j & \sum_{i=1}^{\infty} \boldsymbol{a}_{i,j}\boldsymbol{T}_j \end{pmatrix} \\
& \qquad (0 \le b \le B,\ 1 \le j \le R)
\end{aligned} \tag{3.5}$$

$\mathcal{A}_{b,j}$ denotes the cell transmission matrix of the j-th node, and $\overline{\mathcal{A}_{b,j}}$ is the matrix associated with the event that no cell is transmitted at the j-th node. Those are derived in the following form,

$$\begin{aligned}
\mathcal{A}_{b,j} &= \begin{pmatrix} \alpha_j(0,b)\boldsymbol{I}[2] & \boldsymbol{0}[2] & \cdots & \boldsymbol{0}[2] \\ \boldsymbol{0}[2] & \alpha_j(1,b)\boldsymbol{I}[2] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0}[2] \\ \boldsymbol{0}[2] & \cdots & \boldsymbol{0}[2] & \alpha_j(K_j,b)\boldsymbol{I}[2] \end{pmatrix}, \\
\overline{\mathcal{A}_{b,j}} &= \boldsymbol{I}[2(K_j+1)] - \mathcal{A}_{b,j}, \qquad (0 \le b \le B,\ 1 \le j \le R)
\end{aligned} \tag{3.6}$$

where, $\boldsymbol{I}[m]$ and $\boldsymbol{0}[m]$ represent a unit and zero matrix, respectively.
Therefore, we can obtain $\boldsymbol{G}_{b,j}(x)$ as follows,

$$\begin{aligned}
\boldsymbol{G}_{b,j}(0) &= \boldsymbol{G}^*_{b,j}(0) + (1-\delta_j)\boldsymbol{G}^*_{b,j}(1), \\
\boldsymbol{G}_{b,j}(1) &= \delta_j \boldsymbol{G}^*_{b,j}(1) \qquad (0 \le b \le B,\ 1 \le j \le R).
\end{aligned} \tag{3.7}$$

Then, the steady state probability can be obtained by applying an analytical approach in Chapter 2.

3.3.2 Derivation of Performance Measure

As shown in Chapter 2, performance measures can be derived by means of the steady state probability; 1) the average number of lost cells per slot at the j-th node, $R_{loss1,j}$, and at the congested server, R_{loss2}, and 2) the average queue length at the j-th node, $L_{q,1j}$, and at the congested server, $L_{q,2}$. Then we can define the following performance measures of interest.

1. The cell loss probability
   ~ In terms of cells transmitted to the congested server via the j-th node:

$$P_{loss,j,c} = \frac{R_{loss1,j}}{\Lambda_j} + \frac{R_{loss2}}{\sum_{i=1}^{R} \delta_i \Lambda_i}. \tag{3.8}$$

   ~ In terms of cells sent toward other servers via the j-th node:

$$P_{loss,j,o} = \frac{R_{loss1,j}}{\Lambda_j}. \tag{3.9}$$

2. The average waiting time
   ~ In terms of cells transmitted to the congested server via the j-th node:

$$W_{j,c} = \frac{L_{q,1j}}{\Lambda_j - R_{loss1,j}} + \frac{L_{q,2}}{\sum_{i=1}^{R} \delta_i (\Lambda_i - R_{loss1,i}) - R_{loss2}}. \tag{3.10}$$

   ~ In terms of cells sent toward other servers via the j-th node:

$$W_{j,o} = \frac{L_{q,1j}}{\Lambda_j - R_{loss1,j}}. \tag{3.11}$$

3.4 Numerical Results and Discussions

In this section, we evaluate the performance of backpressure control and investigate the effectiveness of introduction of the cell push-out scheme using the analysis presented in the previous section.

3.4.1 System Configuration Considered for Numerical Results

We examine a system consisting of two nodes, i.e., $R = 2$, and both nodes are homogeneous with respect to the following parameters; for j such that $j = 1, 2$,

- the buffer size, $K_j = K$, and the threshold for the cell push-out scheme, $H_{k_j} = H_k$,
- the control function, $\alpha_j(k_j, b) = \alpha(k, b)$,
- the MMBP parameters, $(\Lambda_{1j}, \Lambda_{2j}, r_{1j}, r_{2j}) = (\Lambda_1, \Lambda_2, r_1, r_2)$ and the maximum number of cells arriving at each slot in the l-th ($l = 1, 2$) state of MMBP, $N_{lj} = \mathcal{N}_l$.

More specifically, as a stochastic model for an input traffic at the nodes, we exploit an interrupted Bernoulli process (IBP), in which there exist two states, i.e., ON state and OFF state, and cells are generated only during ON state. In numerical examples, the averages of ON and OFF duration are equally set at 10[cells], and N cells or less are generated with probability λ_1 at each slot during ON state. Namely, we will use $(\Lambda_1 = \lambda_1, \Lambda_2 = 0, r_1 = 0.1, r_2 = 0.1)$, $\mathcal{N}_1 = N$ and $\mathcal{N}_2 = 0$. We also set the average arrival rate at each node queue Λ_j to λ, which is equal to $\lambda_1/2$ from Eq. (3.1).

3.4.2 Impact of the Threshold, H_k, for the Cell Push-out : the Case where $\delta_1 = \delta_2 = \delta$

In this subsection, we investigate the impact of the threshold H_k for the cell push-out scheme on the cell loss probability in the following system under **BP2**; that is, each node and the congested server are equipped with a buffer of 8 cells, $K = 8$, and 10 cells, $B = 10$, respectively, and the propagation delay between them is equal to one cell, $\tau = 1$. Therefore, from Eq. (3.2), the threshold for the backpressure, H_b, becomes equal to 8. Furthermore, it is here assumed that the transfer rate, δ_j, is the same at both nodes, in which δ_j refers to the ratio of the amount of the traffic transmitted to the congested server to the total amount of the traffic arriving at the j-th node. Since $\delta_1 = \delta_2 = \delta$, we rewrite the cell loss probabilities considered here as follows;

$$P_{loss,1,c} = P_{loss,2,c} = P_{loss,c},$$
$$P_{loss,1,o} = P_{loss,2,o} = P_{loss,o}.$$

Figure 3.3 shows the cell loss probability in case of **BP2** as a function of the threshold H_k for the cell push-out, and also shows both of the cell loss probabilities under **NC** and

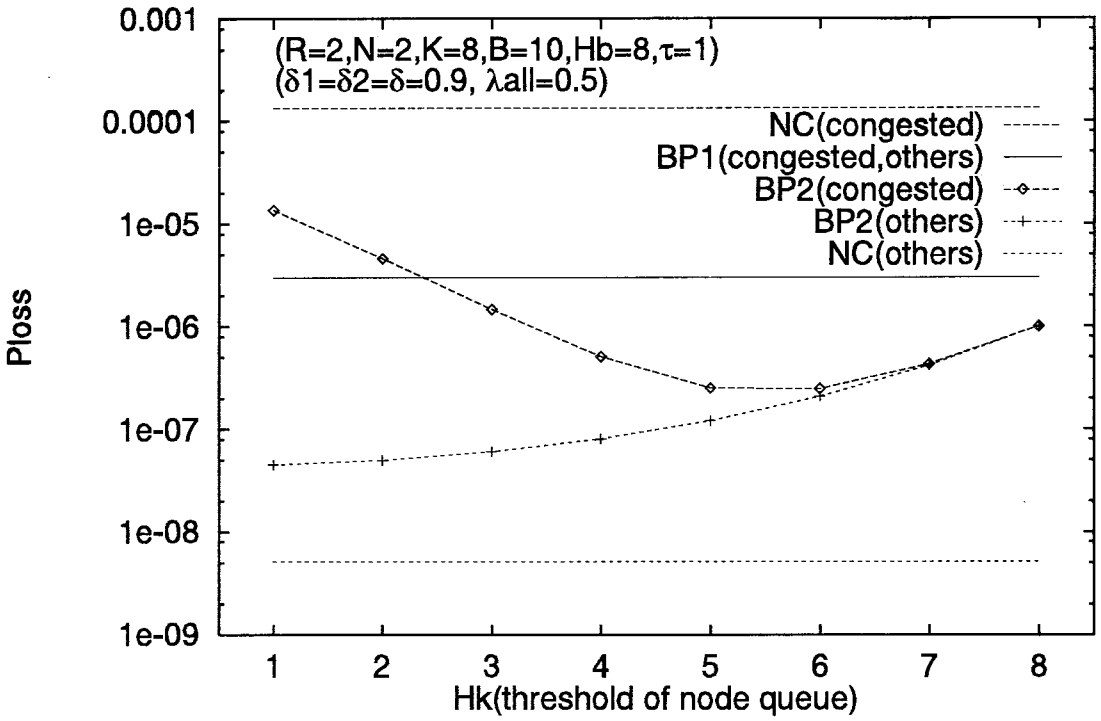


Figure 3.3: Cell loss probability as a function of H_k

BP1 for comparison. We fix the transfer rate δ at 0.9, and the cell arrival rate at the congested server at 0.5, which is denoted by λ_{all} and is given by $2\delta\lambda$. Furthermore, we set the maximum number N of cells arriving at the node to 2.

In case of **NC**, the loss probability of cells transmitted to the congested server, $P_{loss,c}$, is about 10^{-4}, and that toward other servers, $P_{loss,o}$, is about 10^{-8}. Under **BP1**, both $P_{loss,c}$ and $P_{loss,o}$ are equal to 3×10^{-6}; namely, **BP1** degrades the performance in terms of cells going to other servers whereas it improves $P_{loss,c}$. Comparing with the performance of **BP1**, we see from Fig. 3.3 that the use of the cell push-out scheme (**BP2**) is very effective in improving both $P_{loss,c}$ and $P_{loss,o}$ over a wide range of H_k. Therefore, it is seen from this figure that the cell push-out scheme 1) avoids congestion efficiently, and 2) has less influence on traffic not traveling the congested intermediate node than the plain backpressure control.

In Fig. 3.4, we examine the impact of the threshold H_b for the backpressure control under **BP2**. For this purpose, the cell loss probabilities for $H_b = 6, 7, 8$ and 9 are shown there. It is noted that $H_b = 8$ is given by Eq. (3.2). As shown in the figure, $P_{loss,o}$ decreases as H_b increases. This is due to the fact that a larger value of H_b less frequently affects cells stored in the upstream nodes. On the other hand, $P_{loss,c}$ significantly degrades when $H_b = 9$ while there are not remarkable differences among $P_{loss,c}$ for other values of H_b. We observe that Eq. (3.2) gives an appropriate threshold for the backpressure control even under **BP2**. Thus, in what follows, we will exploit H_b given by Eq. (3.2) as the threshold for the backpressure control.

Table 3.1 shows the impact of the maximum number N of cells arriving at each slot in the node on $P_{loss,c}$ in case of **BP2**. N can be regarded as the number of connections going through the node. It is shown in this table that, when $K = 8$, H_k minimizing $P_{loss,c}$ is equal to 6, 5 and 5 for $N = 2, 3$ and 5, respectively. Furthermore, when K is increased from 8 to 10, H_k minimizing $P_{loss,c}$ becomes 6 for $N = 5$. Although it is rather difficult to obtain the optimum threshold minimizing $P_{loss,c}$ as a function of K and N, we see from these results that H_k^* given by the following equation offers a good performance.

$$H_k^* = K - N. \tag{3.12}$$

This equation indicates that the cell push-out should be performed when the remaining buffer capacity becomes N or less. The impact of δ on $P_{loss,c}$ is shown in Table 3.2. As δ decreases, H_k minimizing $P_{loss,c}$ also decreases and becomes more different from H_k^* given by Eq. (3.12). We see, however, that H_k^* given by Eq. (3.12) can achieve a good performance similar to the best one.

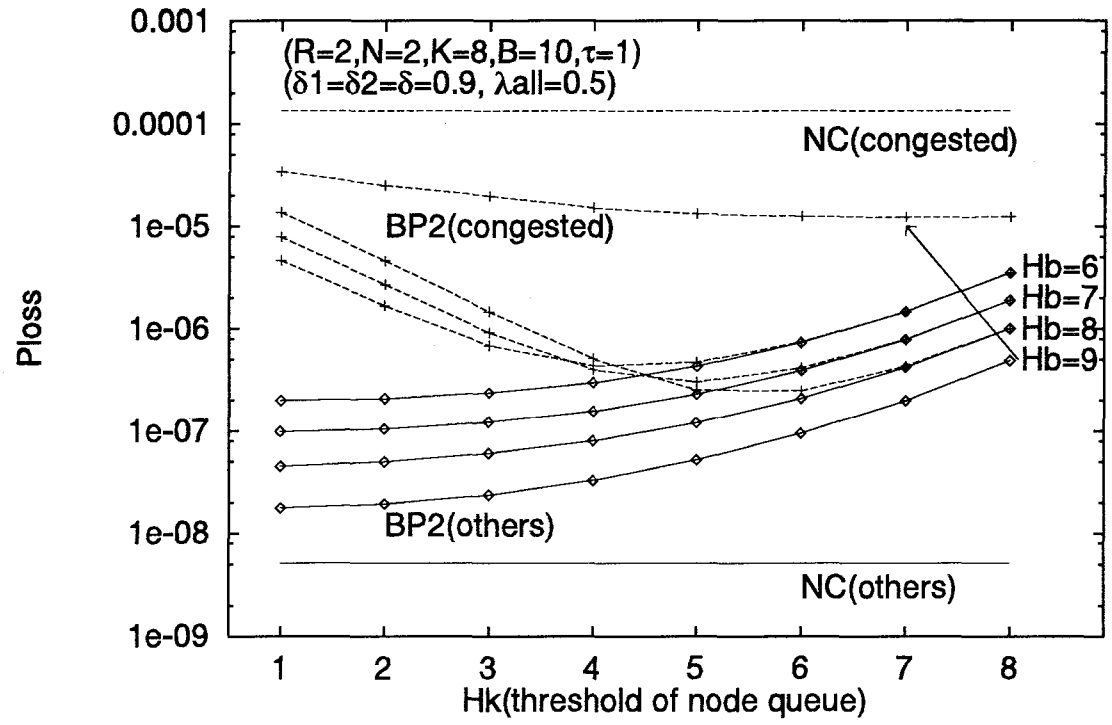


Figure 3.4: Impact of the threshold for back-pressure control, H_b

Table 3.1: Impact of the maximum number N of cells arriving at the node

H_k	$P_{loss,c}$ in case of BC2			
	$N = 2$	$N = 3$	$N = 5$	
1	1.37e-05	1.68e-05	2.09e-05	5.72e-05
2	4.58e-06	6.81e-06	1.05e-05	2.41e-05
3	1.46e-06	2.79e-06	5.82e-06	9.37e-06
4	5.08e-07	1.45e-06	4.21e-06	3.87e-06
5	2.51e-07	**1.15e-06**	**3.98e-06**	1.99e-06
6	**2.47e-07**	1.34e-06	4.54e-06	**1.50e-06**
7	4.29e-07	2.00e-06	5.89e-06	1.62e-06
8	1.01e-06	3.40e-06	8.37e-06	2.16e-06
9	$(K = 8, B = 10)$			3.21e-06
10	$(K = 10, B = 8) \rightarrow$			5.07e-06

$(R = 2,\ \delta = 0.9,\ \lambda_{all} = 0.5,\ \tau = 1)$

Table 3.2: Impact of the transfer rate δ

H_k	$P_{loss,c}$ in case of BC2			
	$\delta = 0.6$	$\delta = 0.7$	$\delta = 0.8$	$\delta = 0.9$
3	3.74e-05	3.68e-06	1.87e-06	1.46e-06
4	**3.70e-05**	2.74e-06	8.52e-07	5.08e-07
5	3.76e-05	**2.63e-06**	**5.93e-07**	2.51e-07
6	3.91e-05	3.04e-06	6.80e-07	**2.47e-07**

$(R = 2,\ N = 2,\ \lambda_{all} = 0.5,\ \tau = 1,\ K = 8, B = 10)$

From theses discussions, we will employ H_k^* given by Eq. (3.12) as the threshold for the cell push-out scheme from now on.

3.4.3 Impact of Arrival Rate and the Buffer Size on the Effectiveness of BP2

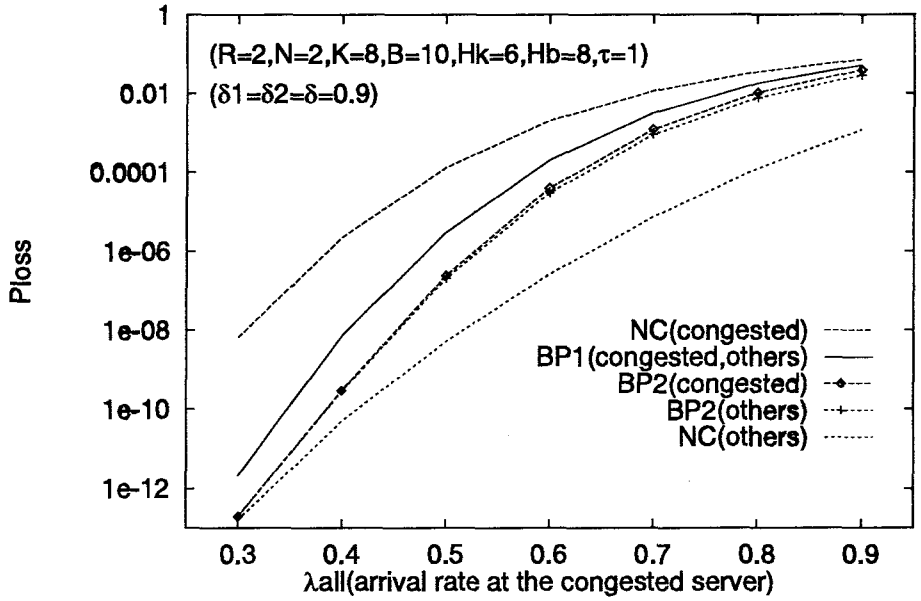
Figure 3.5 shows the sensitivity of both the cell loss probability and the average waiting time to the total arrival rate, λ_{all}, at the congested server. From Fig. 3.5-(a), P_{loss} in **BP2** is improved about one order of magnitude over a wide rage of λ_{all} compared with that in **BP1**. Both $P_{loss,c}$ and $P_{loss,o}$ in **BP2** approach $P_{loss,o}$ in **NC** as λ_{all} decreases. This is due to the fact that cells pushed out from the nodes can enter the congested server more often for a smaller value of λ_{all}. Under **BP2**, cells pushed out result in the increase of the queue length of the buffer at the congested server, while they contribute to the decrease of cell loss as we have seen in several figures so far. Namely, the improvement on the cell loss performance is achieved at the cost of additional delay in the buffer under **BP2**. However, we see from Fig. 3.5-(b) that the average waiting times, W_o and W_c, in **BP2** do not degrade so much compared with that in **BP1**.

We investigate the impact of the buffer size K of the nodes on the cell loss probability in Fig. 3.6. It is seen that, under **NC**, $P_{loss,c}$ is almost insensitive to K. Under both **BP1** and **BP2**, $P_{loss,c}$ is almost the same as $P_{loss,o}$. In addition, the figure shows that **BP2** offers a better performance than **BP1**. For example, suppose that the required quality of service (QOS) in terms of $P_{loss,c}$ must be less than 10^{-6}. Under **BP2** each node equipped with a buffer of only 7 cells guarantees this QOS, whereas each node needs a buffer of 10 cells under **BP1**.

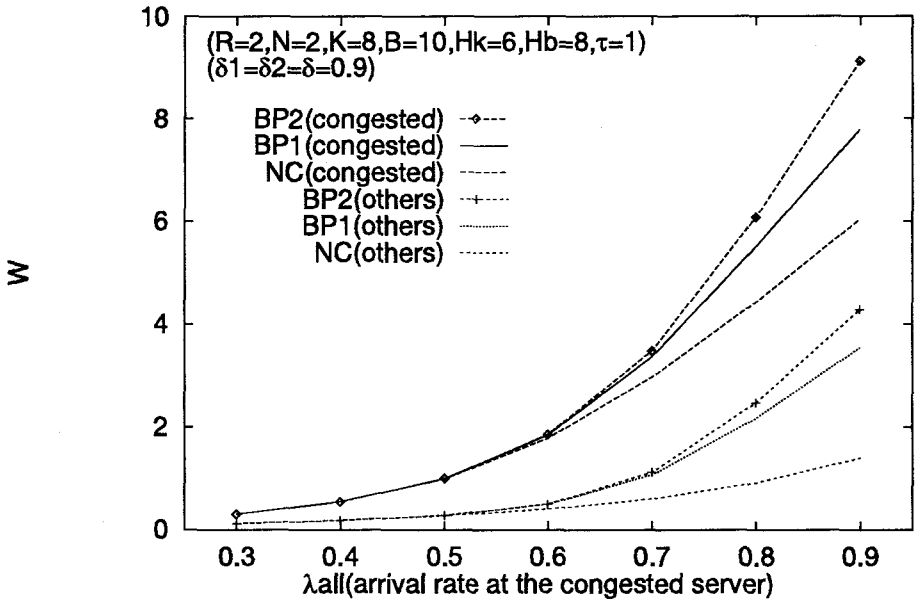
3.4.4 The Effectiveness of BP1 and BP2 in Case where $\delta_1 \neq \delta_2$

In this subsection, we consider the situation where $\delta_1 \neq \delta_2$; δ_2 is relatively small while δ_1 is always set to 1. In this context, the loss probability of cells arriving at node 2 is focused on to investigate the effectiveness of both the plain backpressure control (**BP1**) and that with the cell push-out scheme (**BP2**).

Figure 3.7 represents the cell loss probability as a function of the transfer rate δ_2 at node 2. Both the buffer size of the nodes and that of the congested node are 5 cells: $K = 5$ and $B = 5$. We set the average arrival rate at each node λ to 0.2; λ_{all} at the congested server thus becomes equal to $0.2(1 + \delta_2)$. In case of **NC**, as δ_2 increases, $P_{loss,c}$ becomes larger while $P_{loss,o}$ is kept constant. Under **BP1**, $P_{loss,c}$ can be larger than that in **NC** when δ_2 is small. This is due to a way to determine the threshold for the backpressure control, H_b.



-(a)cell loss probability



-(b)average waiting time

Figure 3.5: Impact of the total offered traffic rate at the congested server, λ_{all}

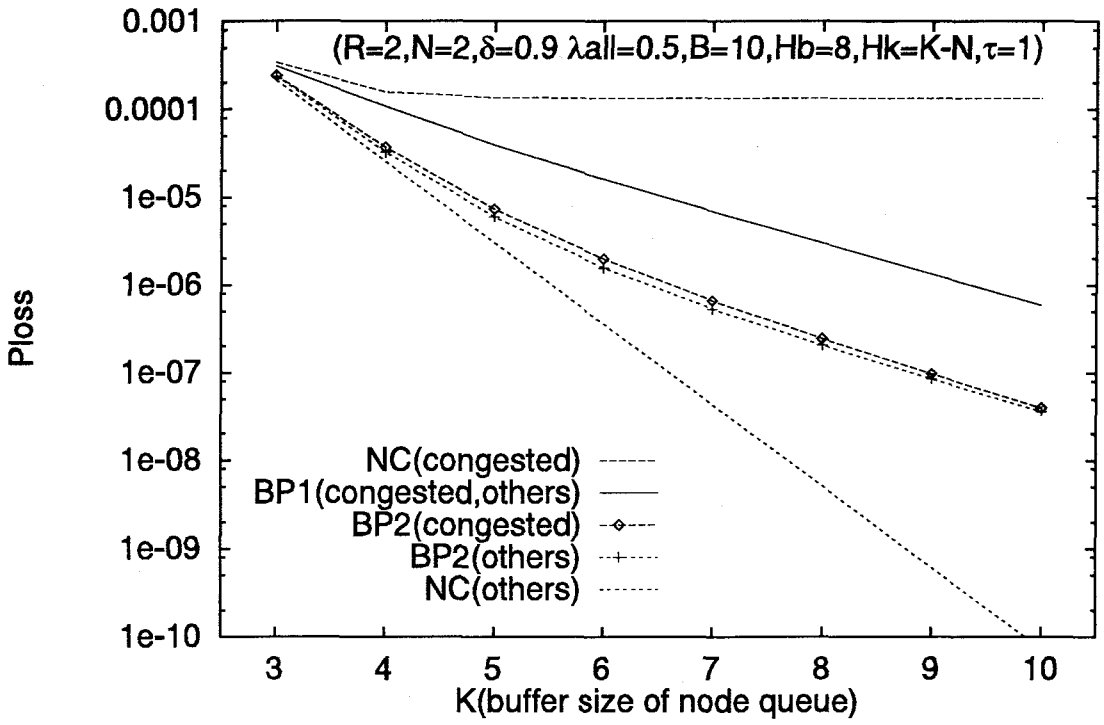


Figure 3.6: Cell loss probability as a function of the buffer size of adjacent nodes, K

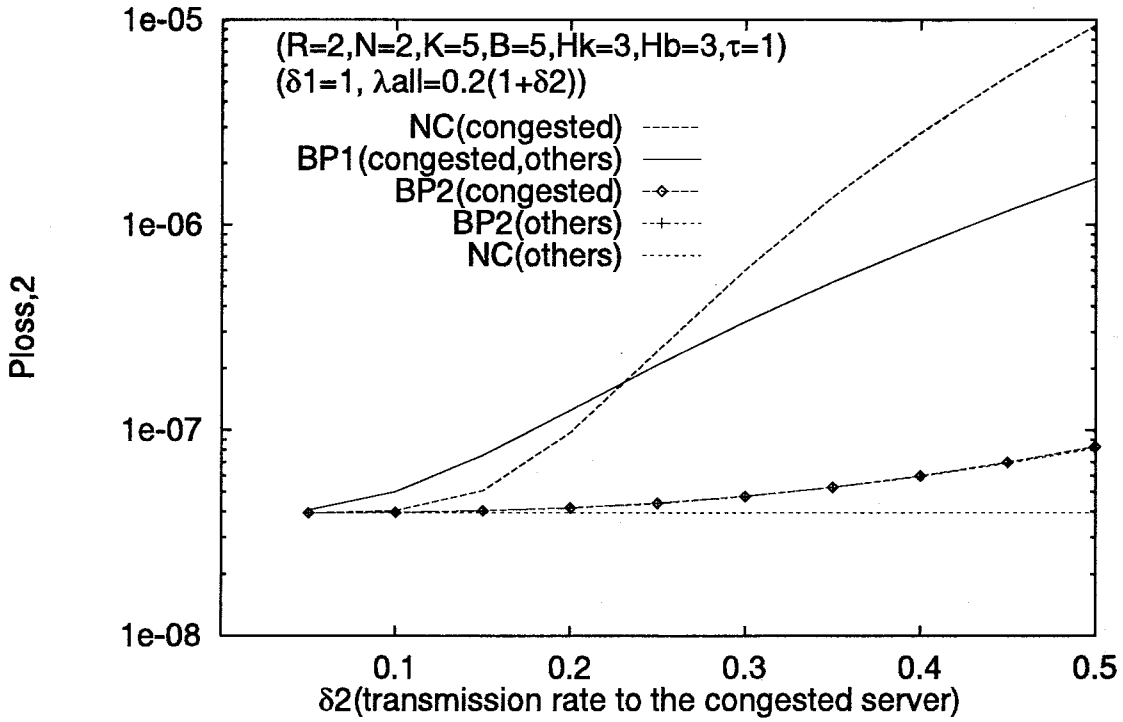


Figure 3.7: Cell loss probability versus δ_2

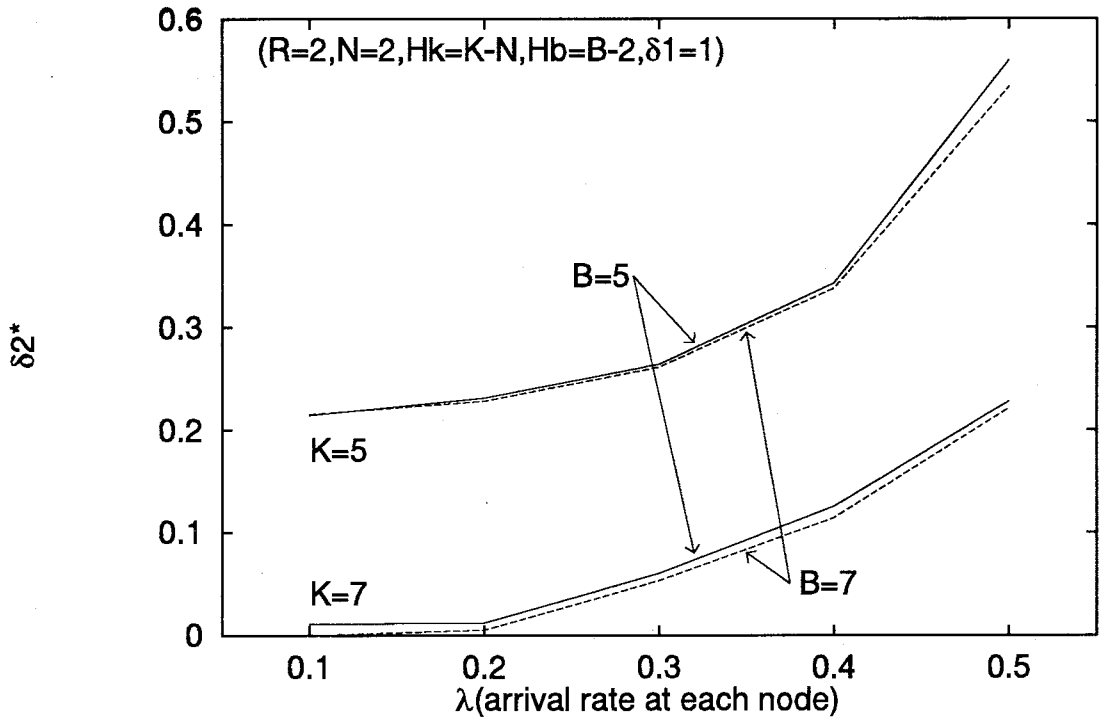


Figure 3.8: δ_2^* versus arrival rate at each node, λ

As shown in Eq. (3.2), H_b is determined in such a way to completely avoid cell loss at the congested buffer even in the worst case where R cells arrive there in each slot during $2\tau+1$ slots of the roundtrip propagation delay. Therefore, when δ_2 is small, H_k given by Eq. (3.2) is very pessimistic; i.e., it happens very often that the buffer at the congested node is not filled with cells. Moreover, most of cells at the node are destined for other servers in this case, but can not be forwarded due to the backpressure from the congested server. This causes these nodes to be congested as well. On the other hand, in case of **BP2**, since the cell push-out successfully relieves the HOL blocking due to the backpressure, the cell loss probability can be kept close to $P_{loss,o}$ in **NC** over a wide range of δ_2 as shown in Fig. 3.7.

Define δ_2^* as the maximum of transfer rate for which **BP1** does not work well to improve $P_{loss,c}$. For example, in Fig. 3.7, δ_2^* is approximately equal to 0.23. We show δ_2^* of **BP1** as a function of the arrival rate at the node λ for a few values of K and B in Fig. 3.8. The figure shows that δ_2^* becomes larger with λ. Namely, the plain backpressure control can not avoid congestion when the amount of traffic arriving at the node is large, and it will in turn cause congestion to rapidly spread all over the network. In addition, we can observe that δ_2^* can be reduced by increasing the buffer size K of the node whereas δ_2^* does not depend on B. Furthermore, it should be noted that δ_2^* of **BP2** is always equal to 0. Thus, the backpressure control with the cell push-out scheme can improves the cell loss performance regardless of any parameters considered here.

3.5 Conclusions

We have evaluated the performance of backpressure control by means of analytical approach in the two-stage queueing model consisting of the congested node and its upstream nodes. Numerical results show that the plain backpressure control scheme (**BP1**) adversely influences the cells not traveling the congested server. Through the analysis, it is shown that the backpressure control combined with the cell push-out scheme (**BP2**) can overcome this drawback and further improve the loss performance in terms of cells transmitted to the congested server. We also suggested an appropriate threshold of the queue in order to *push out* cells. In addition, **BP2** is considered to be effective in WANs because of its ability to efficiently relieve the HOL blocking due to **BP1**.

The case considered here is a limited one in that a system consists of only two-stage nodes. The investigation on the property of the backpressure in a multistage network is of practical interest, and remains as future work.

Chapter 4

Performance Evaluation of Selective Cell Discard Schemes

4.1 Introduction

The Asynchronous Transfer Mode (ATM) has been adopted as a way to accommodate various traffic from multimedia in B-ISDN. In ATM networks, user information is transmitted on cells that consist of 48-byte payload and 5-byte header. On the other hand, the transport-layer located above the ATM-layer handles the transmission based on the packet which is composed of one or more cells. The function of the transport-layer is to manage end-to-end communication and to perform the error control and traffic control using packets. If a cell gets lost within the network, the destination can not reassemble the packet to which the cell belongs, and requires the source to retransmit the packet. Therefore, once one of cells constituting a packet is lost, its subsequent cells in the packet are not worth relaying within the network. Thus, the network resources could be utilized efficiently in a way to selectively discard such cells at the multiplexing node or the ATM switch [ARMI 95, ROMA 95, NORI 94, KAMA 95].

The ATM Adaptation layer (AAL) which is the upper layer of ATM transforms a packet received from the transport-layer into cells at the source node and vice versa at the destination node. AAL5, one of AAL protocols, appends the trailer behind the packet and directly divides the packet into each 48-byte payload of ATM-layer. Furthermore, the ATM-layer-user-to-ATM-later-user (AUU) parameter is provided in the payload type indication (PTI) field of the ATM cell header, and it indicates whether the cell is at the end of a packet or not; if AUU is set to 1, then the cell is so. As a result, the node can recognize the last cell of a packet only by observing the header of cells in case of AAL5, without checking its payload.

There is mainly two types of selective cell discard schemes: tail dropping (**TD**, or partial packet discard (PPD) in [ROMA 95]) and early packet discard (**EPD**). **TD** requires the node to discard all subsequent cells from a packet if one cell has been dropped. On the other hand, in applying **EPD**, the node drops entire packets when its queue length reaches a threshold level. We will call this as **EPD1**. Romanow and Floyd [ROMA 95] compare the performance of **TD** with that of **EPD1** in regard to throughput, and suggest that **EPD1** achieves higher throughput than **TD**. Furthermore, in [NORI 94], by extending **EPD** method, in which the node sets up variable thresholds depending on the number of accepted packets, the better performance is obtained than by **EPD1**. It will be referred to as **EPD2** in this paper.

The above studies were carried out by simulation and it was supposed that the packet length is fixed. In order to analytically evaluate the performance, we must keep track of the packets from which cells have been lost. Such a kind of analyses have been proposed (e.g. see [CIDO 93]). However, it becomes hard to treat the case where the packet size is relatively large since the number of states kept in the analysis grows explosively. Kamal [KAMA 95] deals with a tractable model for the tagged-source to only analyze **TD** by supposing that the packet length follows geometric distribution in terms of cells, i.e. the number of states is independent of the packet length. However, this analysis is an approximate one in terms of modeling other sources. In this Chapter, we propose the exact analysis of **TD** by the following assumptions: the distribution of the packet length from AAL5 traffic is geometric one and sources are homogeneous in terms of traffic characteristics. Furthermore, we extend the analysis to treat the cases of both **EPD1** and **EPD2**. Thus, we extensively investigate the impact of the buffer size, the number of sources and traffic characteristics on the effectiveness of selective cell discard schemes.

The remainder of this paper is organized as follows. In section 4.2, we describe our analytical model and explain the selective cell discard scheme in detail. In section 4.3, we define the states of each sources, and derive the steady state probability and the packet loss probability in **TD**, **EPD1** and **EPD2**, respectively. Section 4.4 provides the numerical results to investigate the influence of traffic characteristics on the packet loss probability. Finally, section 4.5 provides concluding remarks.

4.2 Model and Selective Cell Discard Schemes

4.2.1 Model Description

We consider a discrete-time queueing system with time slot being equal to a cell trans-

mission time. In our model, there are R sources transmitting their cells to a multiplexing node or a switch with an output buffer of B cells as illustrated in Fig. 4.1. We suppose that sources are homogeneous in terms of traffic characteristics. Each source generates cells according to an interrupted Bernoulli process (IBP), which is characterized by a parameter set of $(\lambda_{\text{on}}, \alpha, \beta)$. IBP has two states: on-state and off-state. In each slot, the process changes its state from on to off with probability α, and from off to on with probability β. In on-state, one cell arrives in each slot with probability λ_{on}, while no cell arrives in off-state. Thus, the average arrival rate of each source is given by

$$\lambda = \frac{\beta}{\alpha + \beta}\lambda_{\text{on}}. \tag{4.1}$$

As shown in [KUMA 96], the burst length of IBP, which indicates the number of cells generating in an on-period, follows a geometric distribution with parameter $(1 - q)$, where

$$q = \frac{\alpha}{1 - (1 - \alpha)(1 - \lambda_{\text{on}})}. \tag{4.2}$$

It follows that the average burst length is equal to $1/q$. Furthermore, we assume that a burst consists of one packet exactly; the distribution of the packet length is a geometric one with parameter $(1 - q)$ and the average packet length L_p is given by $1/q$.

4.2.2 Selective Cell Discard Schemes

Tail Dropping, TD

As previously mentioned, once one of cells constituting a packet is lost, its subsequent cells of the corrupted packet result in wasting network resources. Thus, under **TD**, the switch discards such cells instead of relaying them even if the buffer can accommodate them [ARMI 95, ROMA 95, NORI 94, KAMA 95]. This prevents those cells from wasting network resources, and will contribute to reducing the number of corrupted packets.

Early Packet Discard, EPD

EPD enforces a switch to drop entire packets prior to its buffer overflow. If the first cell of a packet arrives at a switch when the number of cells in the buffer is more than or equal to some threshold, the first cell is discarded as well as its subsequent cells. Consequently, that packet will be dropped entirely. In addition, once a cell of an accepted packet is lost due to buffer overflow, its subsequent cells are also dropped as in **TD**. Two kinds of **EPDs** are considered. One is **EPD** based upon a fixed threshold (**EPD1**) [ROMA 95], and the other (**EPD2**) is based upon a variable threshold which is a function of the number of ongoing packets which were already accepted [NORI 94].

4.3 Analysis

In this section, we will derive the steady state probability and the packet loss probability in the system where **TD**, **EPD1** or **EPD2** is applied in the node.

First, for this purpose, we define the states of each source as shown in Fig. 4.2 :

- off : the source does not generate a packet.
- on : the source is transmitting a packet and no cells have been lost.
- on* : the source is transmitting a packet, and one or more cells have already discarded in the packet.

Since sources are homogeneous, we can completely describe the state of our system at the tth slot by a set of 1) the queue length $D(t)$ of the multiplexing node, 2) the number $S_{\text{on}}(t)$ of sources in on-state and 3) the number $S_{\text{on}^*}(t)$ of sources in on*-state. The system state $\{D(t), S_{\text{on}}(t), S_{\text{on}^*}(t); t = 0, 1, 2, \ldots\}$ forms a discrete-time Markov chain with finite states. Note that when $S_{\text{on}}(t) = j$ and $S_{\text{on}^*}(t) = k$, the number of sources in off-state at the tth slot is given by $R - j - k$, where R denotes the number of sources in our system. Here, we define S as the state space for $(S_{\text{on}}(t), S_{\text{on}^*}(t))$, i.e., $S = \{(j,k)|0 \leq j, k, j+k \leq R\}$. A state (j,k) is labeled using the following one-to-one mapping function $M(j,k)$ [KANG 93]:

$$M(j,k) = \frac{1}{2}j(2R - j + 3) + k. \tag{4.3}$$

Therefore, S can be equivalently represented by the state space $S^* = \{0, \cdots, M_s - 1\}$, where $M_s = (R+1)(R+2)/2$.

4.3.1 Analysis of TD

We assume that phase state transitions of IBP occur just after cell generations from each source. We first define the cell arrival probability at the multiplexing node and the transition probability of the number of sources in each state, in turn.

Cell arrival probability at the node

We define $a_{M(j,k)}(x)$ as the probability that x cells arrive at the multiplexing node from sources being in on-state when $S_{\text{on}}(t) = j$ and $S_{\text{on}^*}(t) = k$. Note that cells generating at the source being in on*-state are dropped and can not enter the buffer of the node. Thus,

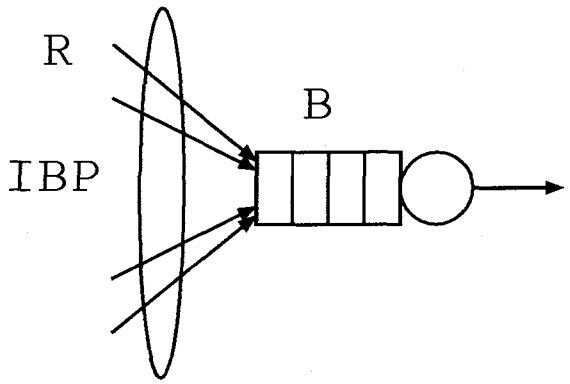


Figure 4.1: Multiplexer model to discard corrupted packets

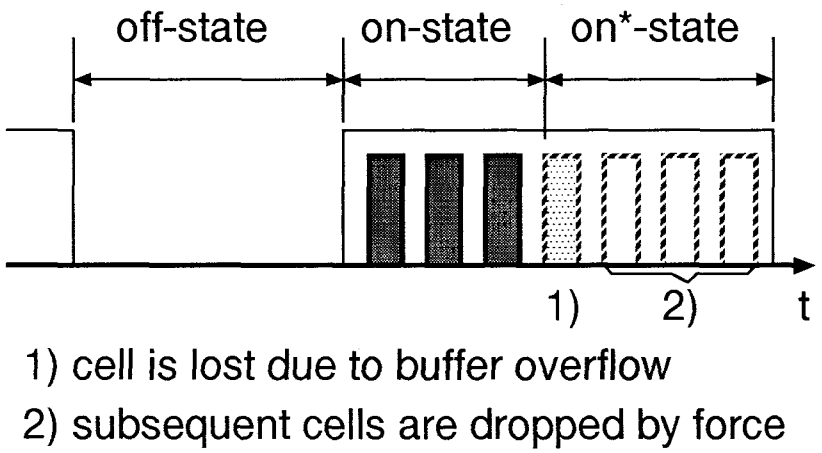


Figure 4.2: Definition of states in source

the probability is independent of $S_{\text{on}^*}(t)$, and is given as follows.

$$\begin{aligned} a_{M(j,k)}(x) &\triangleq \Pr\{N(t) = x | S_{\text{on}}(t) = j, S_{\text{on}^*}(t) = k\} \\ &= \binom{j}{x} \lambda_{on}^{x} (1 - \lambda_{on})^{j-x}. \end{aligned} \tag{4.4}$$

The related matrix representation, $\boldsymbol{a}(x)$, takes the following form.

$$\boldsymbol{a}(x) = \text{diag}[a_0(x), \cdots, a_{M(j,k)}(x), \cdots, a_{M_s-1}(x)]. \tag{4.5}$$

Transition probability of the number of sources being in each state

We obtain below the transition probability of the number of sources being in each state under the condition that y cells are lost in a slot. Let $L(t)$ denote a random variable which represents the number of cells lost due to buffer overflow at the tth slot. Figure 4.3 illustrates the transition diagram in the case of $L(t) = y$. Defining the following probability,

$$c_{M(j,k),M(m,n)}(y) \triangleq \Pr\{S_{\text{on}}(t+1) = m, S_{\text{on}^*}(t+1) = n | L(t) = y, S_{\text{on}}(t) = j, S_{\text{on}^*}(t) = k\}, \tag{4.6}$$

we derive it by means of the transition diagram shown in Fig. 4.3 as follows.

Suppose that cells from y sources are lost due to buffer overflow, z sources among those y sources change their states to on*-state, and, on the other hand, $y - z$ sources change their states to off-state because lost cells are the last ones of packets. Furthermore, p sources being in on-state are assumed to keep their states. Table 4.1 shows the transition of the number of sources in this case. The rightmost column (the bottom row) there represents the number of sources in each state before (after) the transition. We thus have $c_{M(j,k),M(m,n)}(y)$ as

$$c_{M(j,k),M(m,n)}(y) = \begin{cases} \sum_{z=\max(0,n-k)}^{\min(y,n)} \binom{k}{k-n+z} \alpha^{k-n+z} (1-\alpha)^{n-z} \\ \quad \cdot \sum_{p=\max(0,j+k+m-R)}^{\min(m,j-y)} \binom{j-y}{p} \alpha^{j-y-p} (1-\alpha)^{p} \binom{y}{z} \alpha^{y-z} (1-\alpha)^{z} \\ \quad \cdot \binom{R-j-k}{m-p} \beta^{m-p} (1-\beta)^{R-j-k-m+p} \qquad \text{if } y \le j, \\ 0 \qquad \text{if } y > j. \end{cases} \tag{4.7}$$

Furthermore, we define $\boldsymbol{c}(y)$ as

$$\boldsymbol{c}(y) = \begin{pmatrix} c_{0,0}(y) & \cdots & c_{0,M_{(m,n)}}(y) & \cdots & c_{0,M_s}(y) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ c_{M_{(j,k)},0}(y) & \cdots & c_{M_{(j,k)},M_{(m,n)}}(y) & \cdots & c_{M_{(j,k)},M_s}(y) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ c_{M_s,0}(y) & \cdots & c_{M_s,M_{(m,n)}}(y) & \cdots & c_{M_s,M_s}(y) \end{pmatrix}. \tag{4.8}$$

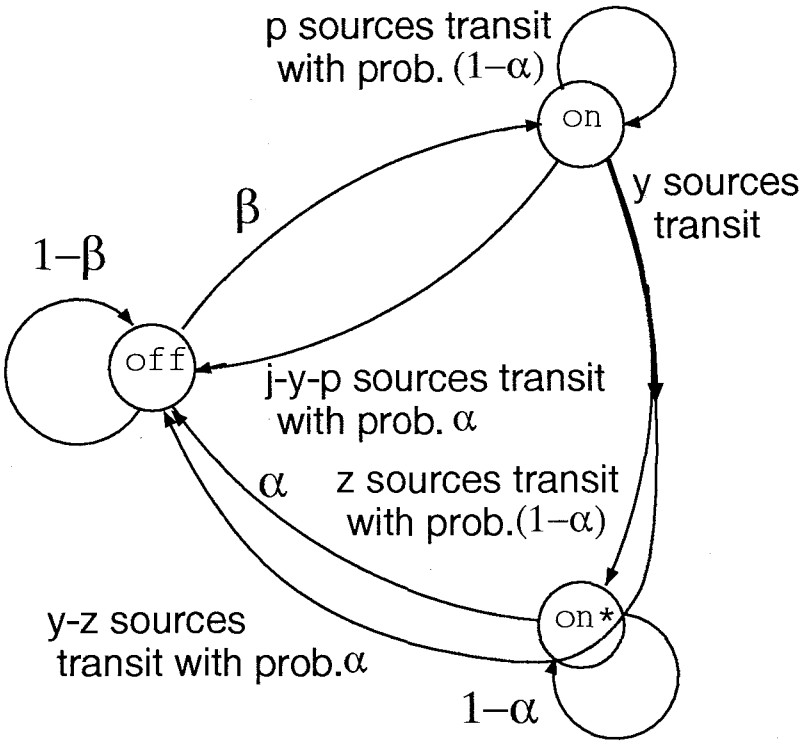


Figure 4.3: Transition diagram of the number of sources in each state, focused on that in on-state for **TD**

Table 4.1: Transition of the number of sources in each state for **TD**

	$S_{\text{on}}(t+1)$	$S_{\text{on}^*}(t+1)$	$S_{\text{off}}(t+1)$	
$S_{\text{on}}(t)$	p	z	j-z-p	j
$S_{\text{on}^*}(t)$	-	n-z	k-n+z	k
$S_{\text{off}}(t)$	m-p	-	R-j-k-m+p	R-j-k
	m	n	R-m-n	R

Table 4.2: Transition of the number of sources in case where $b \geq H$ for **EPD1**

	$S_{\text{on}}(t+1)$	$S_{\text{on}^*}(t+1)$	$S_{\text{off}}(t+1)$	
$S_{\text{on}}(t)$	m	z	j-m-z	j
$S_{\text{on}^*}(t)$	-	p	k-p	k
$S_{\text{off}}(t)$	-	n-z-p	R-j-k-n+z+p	R-j-k
	m	n	R-m-n	R

Steady state probability of our system

We define the system state probability at the tth slot as

$$P_t(b, M(j,k)) = \Pr\{D(t) = b, S_{\text{on}}(t) = j, S_{\text{on}^*}(t) = k\}, \tag{4.9}$$

and introduce its vector representation,

$$\begin{aligned} \boldsymbol{P}_t &= [\boldsymbol{P}_{t,0}, \cdots, \boldsymbol{P}_{t,b}, \cdots, \boldsymbol{P}_{t,B}], \\ \boldsymbol{P}_{t,b} &= [\cdots, P_t(b, M(j,k)), \cdots] \qquad 0 \le M(j,k) \le M_s - 1. \end{aligned}$$

Letting $\boldsymbol{A}(x,y)$ denote the transition matrix whose elements represent the probabilities that x cells arrive given that y cells out of those x ones get lost at the node, we have

$$\boldsymbol{A}(x,y) = \boldsymbol{a}(x) \cdot \boldsymbol{c}(y). \tag{4.10}$$

Further, we define the followings.

$$\boldsymbol{A}_x = \boldsymbol{A}(x,0), \qquad \overline{\boldsymbol{A}_x} = \sum_{i=x}^{R} \boldsymbol{A}(i, i-x). \tag{4.11}$$

As a result, $\boldsymbol{P}_{t,b}$ is related to $\boldsymbol{P}_{t+1,b'}$ as follows.

$$\boldsymbol{P}_{t+1,b'} = \begin{cases} \boldsymbol{P}_{t,0}\boldsymbol{A}_{b'} + \sum_{b=1}^{b'+1} \boldsymbol{P}_{t,b}\boldsymbol{A}_{b'-b+1} & \text{if} \quad 0 \le b' \le B-2, \\ \boldsymbol{P}_{t,0}\boldsymbol{A}_{B-1} + \sum_{b=1}^{B} \boldsymbol{P}_{t,b}\overline{\boldsymbol{A}_{B-b}} & \text{if} \quad b' = B-1, \\ \boldsymbol{P}_{t,0}\overline{\boldsymbol{A}_B} & \text{if} \quad b' = B. \end{cases} \tag{4.12}$$

The transition probability matrix, $\boldsymbol{Q}$, of the **TD** system thus becomes

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_{B-3} & \boldsymbol{A}_{B-2} & \boldsymbol{A}_{B-1} & \overline{\boldsymbol{A}_B} \\ \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_{B-3} & \boldsymbol{A}_{B-2} & \overline{\boldsymbol{A}_{B-1}} & \boldsymbol{0}[M_s] \\ \boldsymbol{0}[M_s] & \boldsymbol{A}_0 & \cdots & \boldsymbol{A}_{B-4} & \boldsymbol{A}_{B-3} & \overline{\boldsymbol{A}_{B-2}} & \boldsymbol{0}[M_s] \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \ddots & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \overline{\boldsymbol{A}_2} & \boldsymbol{0}[M_s] \\ \vdots & \cdots & \cdots & \ddots & \boldsymbol{A}_0 & \overline{\boldsymbol{A}_1} & \boldsymbol{0}[M_s] \\ \boldsymbol{0}[M_s] & \cdots & \cdots & \cdots & \boldsymbol{0}[M_s] & \overline{\boldsymbol{A}_0} & \boldsymbol{0}[M_s] \end{pmatrix}, \tag{4.13}$$

where $\boldsymbol{0}[M]$ represents a zero matrix of dimension $M \times M$. Furthermore, the steady state probability of the system is given by

$$x(b, M(j,k)) = \lim_{t \to \infty} P_t(b, M(j,k)), \tag{4.14}$$

and its vector is represented by

$$\begin{aligned} \boldsymbol{x} &= [\boldsymbol{x}_0, \cdots, \boldsymbol{x}_b, \cdots, \boldsymbol{x}_B], \\ \boldsymbol{x}_b &= \lim_{t\to\infty} \boldsymbol{P}_{t,b}. \end{aligned}$$

Consequently, we can obtain $\boldsymbol{x}$ by solving the following equation:

$$\boldsymbol{x} = \boldsymbol{x}\boldsymbol{Q}, \quad \boldsymbol{x} \cdot \boldsymbol{e}\,(M_s(B+1)) = 1, \tag{4.15}$$

where $\boldsymbol{e}(m)$ represents a column vector of dimension m, whose elements are all equal to 1.

It may be difficult to directly solve Eq. (4.15) since the number of states grows explosively as the buffer size B becomes large. The algorithm proposed in [TAKI 93] enables us to calculate $\boldsymbol{x}$ even for a large B.

Derivation of Packet loss probability

Figure 4.3 is very helpful in deriving the packet loss probability. First, a packet generation corresponds to a transition of the state from off-state to on-state in Fig. 4.3. Next, when the packet is (not) successfully relayed, the state of the source finally returns to off-state from on-state (on*-state). Consequently, the packet loss probability is given by a ratio of the mean number of sources which change their states from on*-state to off-state to the mean number of sources which do from off-state to on-state in a slot.

First, let K denote the mean number of sources which changes their states to off-state with packet loss in a slot. There exist two related transitions from on*-state to off-state in Fig. 4.3. One of them in fact comes to off-state directly from on-state. This transition corresponds to a phenomenon that only the last cell of a packet is discarded and the state changes into off-state at the following slot. The mean number of sources which take this transition in a slot is denoted by J'. We also denote the mean number of sources which change their states from on*-state to off-state in a slot by K'. Thus, K' and J' are given as follows:

$$\begin{aligned} K' &= \alpha \sum_{b=0}^{B} \sum_{k=1}^{R} \sum_{j=0}^{R-k} k x(b, M(j,k)), \\ J' &= \alpha \sum_{b=0}^{B} \sum_{j=1}^{R} \sum_{k=0}^{R-j} \sum_{l=B-b+1}^{j} (b+l-B) x(b, M(j,k)) a_{M(j,k)}(l). \end{aligned}$$

As a result, K becomes

$$K = K' + J'. \tag{4.16}$$

Next, let I denote the mean number of packets generating in a slot. Note that there is a possibility that a source generates no packet during on-state. Let P_0 denote the probability of no packet generation in an on-period of IBP. Thus, letting I' denote the mean number of sources which change their states from off-state, we have

$$I = (1 - P_0)I', \tag{4.17}$$

where I' and P_0 are given by

$$\begin{aligned} I' &= \beta \sum_{b=0}^{B} \sum_{i=1}^{R} \sum_{j=0}^{R-i} ix(b, M(j, R-i-j)), \\ P_0 &= \sum_{i=0}^{\infty} \alpha(1-\alpha)^i (1-\lambda_{on})^{i+1} = \frac{\alpha(1-\lambda_{on})}{1-(1-\alpha)(1-\lambda_{on})}. \end{aligned}$$

Finally, by using K and I, we obtain the packet loss probability, P_{loss}:

$$P_{\text{loss}} = \frac{K}{I}. \tag{4.18}$$

4.3.2 Analysis of EPD1

Next, we provide the analysis of **EPD1**. Under this scheme, the node discards an entire packet if the first coming cell of the packet arrives there when the number of cells in its buffer is more than or equal to a threshold H.

The transition probability associated with the number of sources being in each state, derived in 4.3.1 under **TD**, is the same as that given by Eq. (4.7) except in the case of $D(t) = b \geq H$: the transition diagram is presented in Fig. 4.4. In that case, the transition probability (corresponding to $c_{M(j,k),M(m,n)}(y)$) is denoted by $c^*_{M(j,k),M(m,n)}(y)$. Transitions from off-state to on-state there do not occur, but transitions from off-state to on*-state do, differently from them in the case of **TD**. Noting this feature, we obtain the transitions related to $c^*_{M(j,k),M(m,n)}(y)$ as shown in Table 4.2. We hence have

$$c^*_{M(j,k),M(m,n)}(y) = \begin{cases} \binom{j-y}{m}(1-\alpha)^m \alpha^{j-m-y} \sum_{z=0}^{\min(y,n)} \binom{y}{z}(1-\alpha)^z \alpha^{y-z} \\ \quad \cdot \sum_{p=\max(0,j+k+n-R-z)}^{\min(k,n-z)} \binom{k}{p} \alpha^{k-p}(1-\alpha)^p \\ \quad \cdot \binom{R-j-k}{n-z-p} \beta^{n-z-p}(1-\beta)^{R-j-k-n+z+p} & \text{if } y \leq j-m, \\ 0 \qquad \text{if } y > j-m. \end{cases} \tag{4.19}$$

and its matrix representation as follows:

$$\boldsymbol{c}^*(y) = \begin{pmatrix} c^*_{0,0}(y) & \cdots & c^*_{0,M_s}(y) \\ \vdots & \ddots & \vdots \\ c^*_{M_s,0}(y) & \cdots & c^*_{M_s,M_s}(y) \end{pmatrix}. \tag{4.20}$$

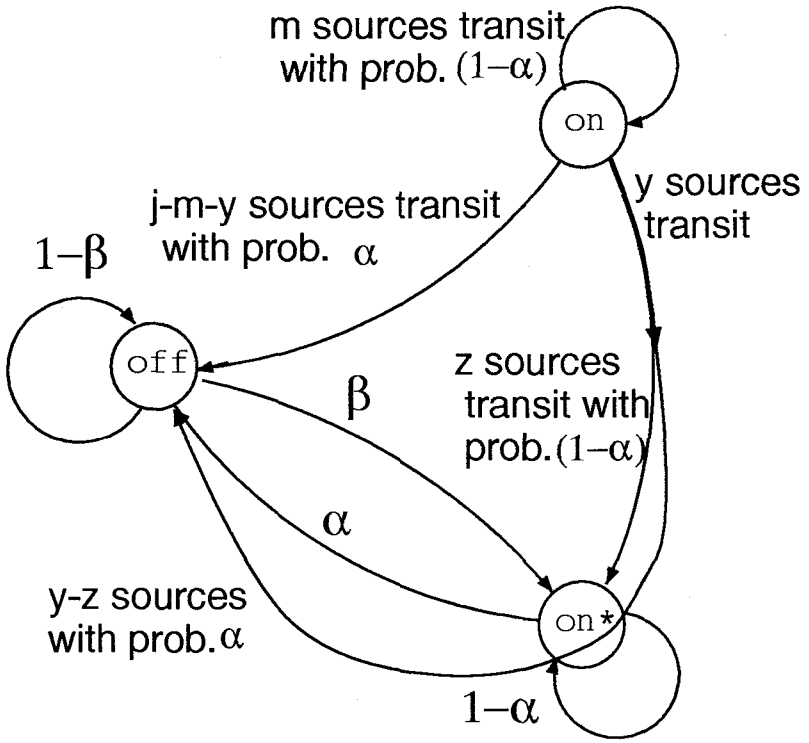


Figure 4.4: Transition diagram of the number of sources for **EPD**

Derivation of packet loss probability

When $D(t) = b \geq H$, the probability matrix corresponding to $\boldsymbol{A}(x, y)$ becomes

$$\boldsymbol{A}^*(x, y) \triangleq \boldsymbol{a}(x) \cdot \boldsymbol{c}^*(y). \tag{4.21}$$

The relation between $\boldsymbol{P}_{t,b}$ and $\boldsymbol{P}_{t+1,b'}$ can be obtained in a similar way of 4.3.1 as follows.

$$\mathbf{P}_{t+1,b'} = \begin{cases} \mathbf{P}_{t,0}\mathbf{A}_{b'} + \sum_{b=1}^{H-1} \mathbf{P}_{t,b}\mathbf{A}_{b'-b+1} + \sum_{b=H}^{b'+1} \mathbf{P}_{t,b}\mathbf{A}^*_{b'-b+1} \\ \qquad \text{if} \quad 0 \leq b' \leq B-2, \\ \mathbf{P}_{t,0}\mathbf{A}_{B-1} + \sum_{b=1}^{H-1} \mathbf{P}_{t,b}\overline{\mathbf{A}_{B-b}} + \sum_{b=H}^{B} \mathbf{P}_{t,b}\overline{\mathbf{A}^*_{B-b}} \\ \qquad \text{if} \quad b' = B-1, \\ \mathbf{P}_{t,0}\overline{\mathbf{A}_B} \\ \qquad \text{if} \quad b' = B, \end{cases} \tag{4.22}$$

where $\boldsymbol{A}_x$ and $\overline{\boldsymbol{A}_x}$ are given in Eq. (4.11), and

$$\boldsymbol{A}^*_x = \boldsymbol{A}^*(x, 0), \qquad \overline{\boldsymbol{A}^*_x} = \sum_{i=x}^{R} \boldsymbol{A}^*(i, i-x). \tag{4.23}$$

Hence, the transition matrix $\boldsymbol{Q}$ for **EPD1** is given by

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_{B-3} & \boldsymbol{A}_{B-2} & \overline{\boldsymbol{A}_{B-1}} & \overline{\boldsymbol{A}_B} \\ \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_{B-3} & \boldsymbol{A}_{B-2} & \overline{\boldsymbol{A}_{B-1}} & \mathbf{0}[M_s] \\ \mathbf{0}[M_s] & \boldsymbol{A}_0 & \cdots & \boldsymbol{A}_{B-4} & \boldsymbol{A}_{B-3} & \overline{\boldsymbol{A}_{B-2}} & \mathbf{0}[M_s] \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \cdots & \boldsymbol{A}_{B-(H-3)} & \boldsymbol{A}_{B-(H-2)} & \overline{\boldsymbol{A}_{B-(H-1)}} & \mathbf{0}[M_s] \\ \vdots & \vdots & \cdots & \boldsymbol{A}^*_{B-(H-2)} & \boldsymbol{A}^*_{B-(H-1)} & \overline{\boldsymbol{A}^*_{B-H}} & \mathbf{0}[M_s] \\ \vdots & \vdots & \cdots & \cdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \ddots & \boldsymbol{A}^*_0 & \boldsymbol{A}^*_1 & \overline{\boldsymbol{A}^*_2} & \mathbf{0}[M_s] \\ \vdots & \cdots & \cdots & \ddots & \boldsymbol{A}^*_0 & \overline{\boldsymbol{A}^*_1} & \mathbf{0}[M_s] \\ \mathbf{0}[M_s] & \cdots & \cdots & \cdots & \mathbf{0}[M_s] & \overline{\boldsymbol{A}^*_0} & \mathbf{0}[M_s] \end{pmatrix}, \tag{4.24}$$

and the steady state probability vector, $\boldsymbol{x}$, is obtained from Eq. (4.15).

We derive the packet loss probability, P_{loss}, for **EPD1** in the same way as that for **TD** (namely, by using Eq. (4.18)) except for the following: when $D(t) = b \geq H$, a source may generate no packet after its transition from off-state to on*-state until it returns to off-state. The mean number of sources which take such a transition in a slot is denoted by I''. We then have

$$I'' = \beta P_0 \sum_{b=H}^{B} \sum_{i=1}^{R} \sum_{j=0}^{R-i} ix(b, M(j, R-i-j)). \tag{4.25}$$

Note that on average, I'' sources does not generate no packets among sources which change their states from on*-state to off-state in a slot. Thus, in **EPD1**, K in Eq. (4.18) should be rewritten as

$$K = (K' - I'') + J', \tag{4.26}$$

while I in Eq. (4.18) is the same as that in Eq. (4.17).

4.3.3 Analysis of EPD2

In **EPD2**, the threshold at the tth slot depends on the number $S_{\rm on}(t)$ of accepted packets, and is determined in a way to prevent cells of already accepted packets from getting lost due to buffer overflow. Since a number of cells in an accepted packet have already left the node, each accepted packet may actually require the buffer capacity of less than the average packet length L_p (cells).

From these reasons, Noritake et al. in [NORI 94] have introduced a parameter δ ($0 \leq \delta \leq 1$), which will be called as the acceptance coefficient, and the threshold H_j ($0 \leq j \leq R$) depending on $S_{\rm on} = j$ as follows

$$H_j \triangleq B - \delta j L_p. \tag{4.27}$$

Thus a newly arriving packet is more likely to be discarded as more packets have been already accepted, since $0 \leq H_R \leq \cdots \leq H_1 \leq B$. Note also that δ should be determined in such a way to provide low packet loss probability; this will be discussed later based on numerical results.

The transition probability matrix corresponding to Eq.(4.20) depends on the current value of the threshold. Let $\boldsymbol{c}_r(y)$ denote the conditional transition probability matrix, given newly arriving packets are discarded if r packets or more are already accepted. We then have

$$\boldsymbol{c}_r(y) = \begin{pmatrix} \tilde{c}_{0,0}(y) & \cdots & \tilde{c}_{0,M_s}(y) \\ \vdots & \ddots & \vdots \\ \tilde{c}_{M_s,0}(y) & \cdots & \tilde{c}_{M_s,M_s}(y) \end{pmatrix}, \tag{4.28}$$

where

$$\tilde{c}_{M_{(j,k)},M_{(m,n)}}(y) = \begin{cases} c_{M_{(j,k)},M_{(m,n)}}(y) & \text{(given in Eq. (4.7))} \quad \text{if } j < r \\ c^*_{M_{(j,k)},M_{(m,n)}}(y) & \text{(given in Eq. (4.19))} \quad \text{if } j \geq r \end{cases}. \tag{4.29}$$

Let $\boldsymbol{A}_r(x,y)$ denote the transition probability matrix when x cells arrive and y cells get lost due to buffer overflow on the same condition as that of $\boldsymbol{c}_r(y)$. $\boldsymbol{A}_r(x,y)$ is given by

$$\boldsymbol{A}_r(x,y) = \boldsymbol{a}(x) \cdot \boldsymbol{c}_r(y), \tag{4.30}$$

and we define the followings.

$$\boldsymbol{A}_{r,x} = \boldsymbol{A}_r(x,0), \qquad \overline{\boldsymbol{A}_{r,x}} = \sum_{i=x}^{R} \boldsymbol{A}_r(i, i-x).$$

Note that $\boldsymbol{A}_{r,x}$ and $\overline{\boldsymbol{A}_{r,x}}$ $(0 \le r \le R)$ are constructed to represent the transition when the queue length $D(t)$ of the node satisfies $H_r \le D(t) < H_{r-1}$ $(1 \le r \le R)$ or $D(t) \ge H_0$ $(r = 0)$. For example, all the transition probabilities from states with $D(t) = b$ $(H_r \le b < H_{r-1})$ to states with $D(t+1) = b' \le B-2$ is represented by $\boldsymbol{A}_{r,b'-b+1}$. Note also that when the queue length $D(t)$ is less than H_R, all the transitions are governed by $\boldsymbol{A}_{R,x}$ or $\overline{\boldsymbol{A}_{R,x}}$. We shall explain this fact. First of all, when $D(t) < H_R$, any of newly arriving packets are not discarded due to **EPD**. On the other hand, $\boldsymbol{A}_{R,x}$ and $\overline{\boldsymbol{A}_{R,x}}$ represents transitions with **EPD** which is effective only if $S_{\mathrm{on}} = R$. However, since the number of sources in the system is assumed to be R, $S_{\mathrm{on}} = R$ implies that no sources changes its state from off-state to on-state; no packets are newly generated. Thus, $\boldsymbol{A}_{R,x}$ and $\overline{\boldsymbol{A}_{R,x}}$ are considered to represent transitions without **EPD**, and therefore we can apply these matrices to the case $D(t) < H_R$ as well as $H_R \le D(t) < H_{R-1}$

Accordingly, by using $\boldsymbol{A}_{r,x}$ or $\overline{\boldsymbol{A}_{r,x}}$ for b such that satisfies $H_r \le b < H_{r-1}$ $(0 \le r \le R)$, we can relate $\boldsymbol{P}_{t+1,b'}$ with $\boldsymbol{P}_{t,b}$ as follows:

$$\boldsymbol{P}_{t+1,b'} = \begin{cases} \boldsymbol{P}_{t,0}\boldsymbol{A}_{R,b'} + \sum_{b=1}^{H_{R-1}-1} \boldsymbol{P}_{t,b}\boldsymbol{A}_{R,b'-b+1} + \sum_{b=H_{R-1}}^{H_{R-2}-1} \boldsymbol{P}_{t,b}\boldsymbol{A}_{R-1,b'-b+1} \\ \quad + \cdots + \sum_{b=H_r}^{H_{r-1}-1} \boldsymbol{P}_{t,b}\boldsymbol{A}_{r,b'-b+1} + \cdots \\ \quad + \sum_{b=H_2}^{H_1-1} \boldsymbol{P}_{t,b}\boldsymbol{A}_{2,b'-b+1} + \sum_{b=H_1}^{b'+1} \boldsymbol{P}_{t,b}\boldsymbol{A}_{1,b'-b+1} \\ \qquad\qquad \text{if} \quad 0 \le b' \le B-2, \\ \boldsymbol{P}_{t,0}\boldsymbol{A}_{R,B-1} + \sum_{b=1}^{H_{R-1}-1} \boldsymbol{P}_{t,b}\overline{\boldsymbol{A}_{R,B-b}} + \sum_{b=H_{R-1}}^{H_{R-2}-1} \boldsymbol{P}_{t,b}\overline{\boldsymbol{A}_{R-1,B-b}} \\ \quad + \cdots + \sum_{b=H_2}^{H_1-1} \boldsymbol{P}_{t,b}\overline{\boldsymbol{A}_{2,B-b}} + \sum_{b=H_1}^{B-1} \boldsymbol{P}_{t,b}\overline{\boldsymbol{A}_{1,B-b}} + \boldsymbol{P}_{t,B}\overline{\boldsymbol{A}_{0,0}} \\ \qquad\qquad \text{if} \quad b' = B-1, \\ \boldsymbol{P}_{t,0}\overline{\boldsymbol{A}_{R,B}} \qquad \text{if} \quad b' = B. \end{cases} \tag{4.31}$$

The transition matrix $\boldsymbol{Q}$ for **EPD2** becomes

$$\boldsymbol{Q} = \begin{pmatrix}
\boldsymbol{A}_{R,0} & \boldsymbol{A}_{R,1} & \cdots & \boldsymbol{A}_{R,B-3} & \boldsymbol{A}_{R,B-2} & \overline{\boldsymbol{A}_{R,B-1}} & \overline{\boldsymbol{A}_{R,B}} \\
\boldsymbol{A}_{R,0} & \boldsymbol{A}_{R,1} & \cdots & \boldsymbol{A}_{R,B-3} & \boldsymbol{A}_{R,B-2} & \overline{\boldsymbol{A}_{R,B-1}} & \boldsymbol{0}[M_s] \\
\boldsymbol{0}[M_s] & \boldsymbol{A}_{R,0} & \cdots & \boldsymbol{A}_{R,B-4} & \boldsymbol{A}_{R,B-3} & \overline{\boldsymbol{A}_{R,B-2}} & \boldsymbol{0}[M_s] \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\
\vdots & \ddots & \cdots & \boldsymbol{A}_{r,B-(H_b-3)} & \boldsymbol{A}_{r,B-(H_b-2)} & \overline{\boldsymbol{A}_{r,B-(H_b-1)}} & \boldsymbol{0}[M_s] \\
\vdots & \vdots & \cdots & \boldsymbol{A}_{r,B-(H_b-2)} & \boldsymbol{A}_{r,B-(H_b-1)} & \overline{\boldsymbol{A}_{r,B-H_b}} & \boldsymbol{0}[M_s] \\
\vdots & \vdots & \cdots & \cdots & \vdots & \vdots & \vdots \\
\vdots & \cdots & \ddots & \boldsymbol{A}_{1,0} & \boldsymbol{A}_{1,1} & \overline{\boldsymbol{A}_{1,2}} & \boldsymbol{0}[M_s] \\
\vdots & \cdots & \cdots & \ddots & \boldsymbol{A}_{1,0} & \overline{\boldsymbol{A}_{1,1}} & \boldsymbol{0}[M_s] \\
\boldsymbol{0}[M_s] & \cdots & \cdots & \cdots & \boldsymbol{0}[M_s] & \overline{\boldsymbol{A}_{0,0}} & \boldsymbol{0}[M_s]
\end{pmatrix}, \quad (4.32)$$

and the steady state probability vector, $\boldsymbol{x}$, is obtained from Eq. (4.15).

We derive I'' in **EPD2**, which is given by Eq.(4.25) in the case of **EPD1**, as

$$I'' = \beta P_0 \left\{ \sum_{e=2}^{R} \sum_{b=H_e}^{H_{e-1}-1} \sum_{i=1}^{R} \sum_{j=e}^{R-i} ix(b, M(j, R-i-j)) + \sum_{b=H_1}^{B-1} \sum_{i=1}^{R} \sum_{j=1}^{R-i} ix(b, M(j, R-i-j)) \right.$$
$$\left. + \sum_{i=1}^{R} \sum_{j=0}^{R-i} ix(B, M(j, R-i-j)) \right\}.$$

Thus, by using Eqs. (4.17), (4.18) and (4.26), we finally obtain the packet loss probability for **EPD2**.

4.4 Numerical Results and Discussions

In this section, we evaluate the performance of selective cell discard schemes using the analysis presented in the preceding section.

4.4.1 Effectiveness of TD

We compare the performance of **TD** with that of the system with no-control, which is denoted by **NC** for short, and investigate the impact of the buffer size and the number of sources on the performance. In order to estimate the improvement of the packet loss probability P_{loss} by **TD**, we first represent the distribution of the queue length and the average number of sources in each state.

Figure 4.5 shows the distribution of the queue length D in the steady state, which is given by

$$\Pr\{D = b\} = \boldsymbol{x}_b \cdot \boldsymbol{e}(M_s), \quad 0 \le b \le B,$$

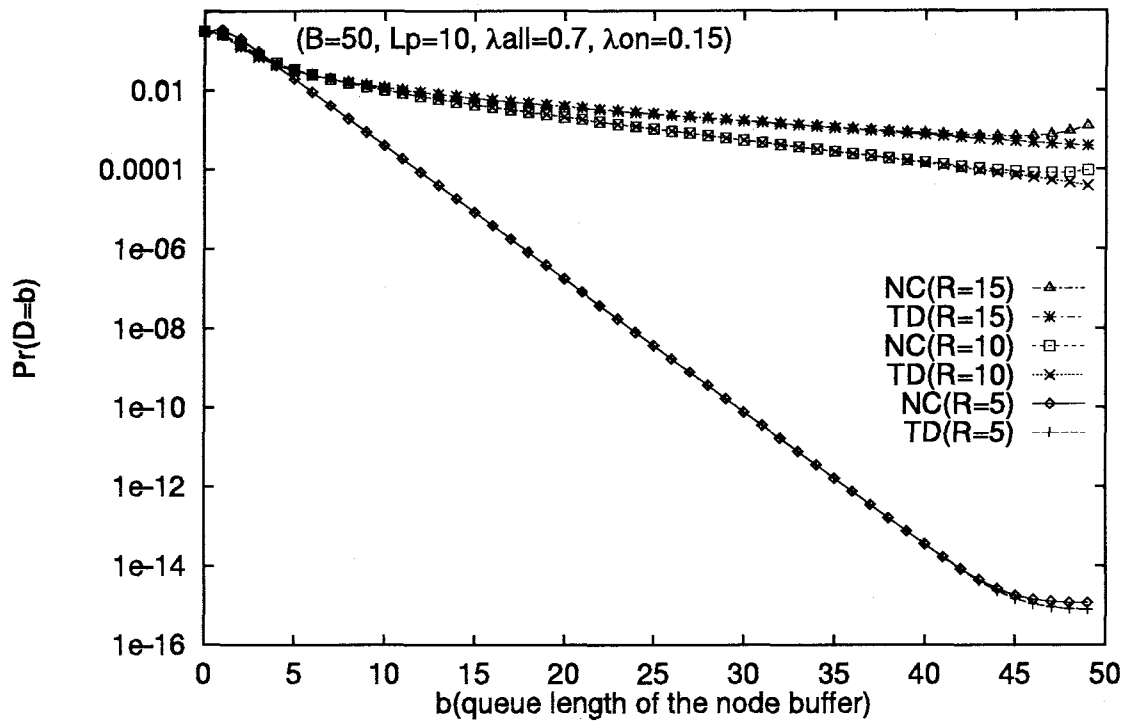


Figure 4.5: Distribution of the queue length

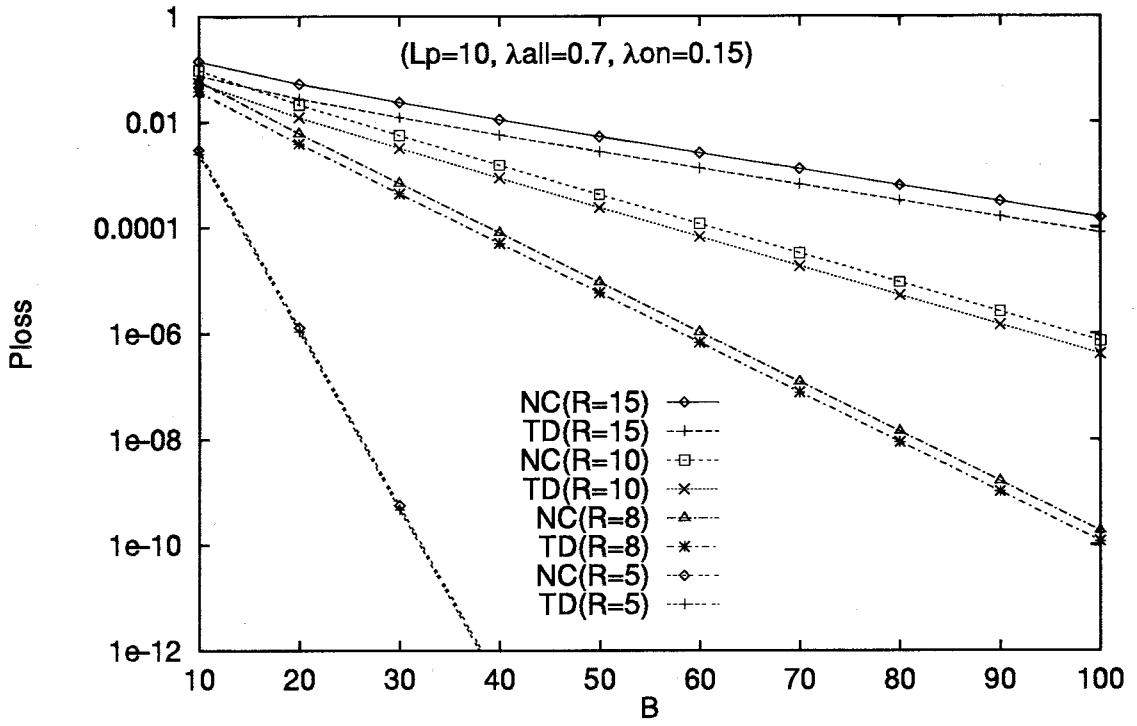


Figure 4.6: Packet loss probability as a function of the buffer size B of the node

where M_s indicates a dimension of the vector $\boldsymbol{x}_b$. The mean packet size L_p is fixed to 10 (cells) in this figure. We also set the buffer size B of the node to 50 (cells), the arrival rate λ_{on} in on-period (namely, peak rate) of IBP to 0.15, and the total arrival rate λ_{all} $(= R\lambda)$ at the node to 0.7. In this setting, we change the number R of sources to 5, 10 and 15. We observe a little difference between the distributions of **TD** and **NC** only when the queue length is greater than about 45 for all R.

Table 4.3 shows the average number of sources in each state in the same setting as in Fig 4.5. Let N_{on}, N_{on^*} and N_{off} denote the average numbers of sources in on-state, on*-state and off-state, respectively. We then have

$$\begin{aligned} N_{\text{on}} &= \sum_{b=0}^{B}\sum_{j=1}^{R}\sum_{k=0}^{R-j} jx(b, M(j,k)), \\ N_{\text{on}^*} &= \sum_{b=0}^{B}\sum_{k=1}^{R}\sum_{j=0}^{R-k} kx(b, M(j,k)), \\ N_{\text{off}} &= \sum_{b=0}^{B}\sum_{i=1}^{R}\sum_{j=0}^{R-i} ix(b, M(j, R-i-j)). \end{aligned}$$

By applying **TD**, N_{on} becomes larger and N_{on^*} gets smaller than those of **NC** while N_{off} does not change. The probability that each source is in off-state is given by $\frac{1/\beta}{1/\alpha+1/\beta} = \frac{\alpha}{\alpha+\beta}$. The probability of the number of sources being in off-state follows the binomial distribution with parameters $(R, \frac{\alpha}{\alpha+\beta})$; its average N_{off} thus becomes equal to $R\frac{\alpha}{\alpha+\beta}$. From Eq. (4.1) and the definition of $\lambda_{\text{all}} = R\lambda$, we can simply rewrite N_{off} as

$$N_{\text{off}} = R - \frac{\lambda_{\text{all}}}{\lambda_{\text{on}}}. \tag{4.33}$$

Note also that the sum of N_{on} and N_{on^*} is kept constant even when R changes, since it is equal to $\lambda_{\text{all}}/\lambda_{\text{on}}$ (see Eq. (4.33)), which is independent of R.

We expect from Fig. 4.5 and Table 4.3 that **TD** improves the packet loss probability P_{loss}. Figure 4.6 illustrates P_{loss} as a function of the buffer size B of the node. When $R = 5$, P_{loss} drastically decreases as B increases. We compare the ratio of P_{loss} in **NC** to that in **TD** when $B = 30$. For $R = 5, 8, 10$ and 15, the ratio is equal to 1.185, 1.575, 1.752 and 1.927, respectively; namely, **TD** is more effective for a larger R. On the other hand, Fig. 4.6 shows that the ratio is kept almost the same value for each R while B varies. Thus, we conclude that the effectiveness of **TD** is almost insensitive to the buffer size.

Table 4.3: Average number of sources in each state for some values of R

	$R = 5$		$R = 10$		$R = 15$	
	NC	**TD**	**NC**	**TD**	**NC**	**TD**
N_{off}	0.33	0.33	5.33	5.33	10.33	10.33
N_{on}	4.67	4.67	4.66	4.67	4.64	4.66
N_{on^*}	7.57e-14	5.83e-14	1.78e-3	1.01e-3	2.24e-2	1.14e-2

$(B = 50,\ L_p = 10, \lambda = 0.7,\ \lambda_{\text{on}} = 0.15)$

Table 4.4: Average number of sources in each state: cases of $L_p = 10$ and 100

	$L_p = 10$		$L_p = 100$	
	NC	**TD**	**NC**	**TD**
N_{off}	5.33	5.33	5.33	5.33
N_{on}	4.66	4.67	4.17	4.47
N_{on^*}	1.78e-3	1.01e-3	4.96e-1	1.97e-1

$(R = 10,\ B = 50,\ \lambda = 0.7,\ \lambda_{\text{on}} = 0.15)$

4.4.2 Impact of Traffic Parameters on Packet Loss Probability of TD

In this subsection, we investigate the impact of the peak rate λ_{on}, the total arrival rate λ_{all} and the mean packet length L_p on the packet loss probability P_{loss}. Figure 4.7 shows P_{loss} as a function of λ_{all}. In this figure, we fix $R = 10$, $B = 50$ and $L_p = 10$ while λ_{on} takes $0.1, 0.15, 0.3$ and 0.5. It is seen from the figure that the difference between P_{loss} in **TD** and that in **NC** gets larger as either λ_{all} or λ_{on} becomes larger, although it is not significant.

Figure 4.8 shows the impact of L_p on P_{loss}. The difference between P_{loss} in **TD** and that in **NC** gets larger as L_p becomes larger. In addition, Table 4.4 gives the average number of sources in each state. Note that the sum of N_{on} and N_{on^*} is constant for any value of L_p because it is given by $\lambda_{\text{all}}/\lambda_{\text{on}}$. We see that **TD** is more effective in decreasing N_{on^*} in the case of $L_p = 100$ than in the case of $L_p = 10$. We shall show below what leads to this feature. Focus on cells belonging to a corrupted packet. As the average packet length becomes larger, it is very likely that there would be more cells following the first lost cell. Thus, since **TD** can gets rid of such cells that will waste the buffer capacity, **TD** will be more effective in using the buffer capacity efficiently for a longer packet. In fact, as shown in Fig. 4.9 illustrating the distribution of the queue length in cases of $L_p = 10$ and 100, **TD** successfully prevents the buffer from being filled with cells; **TD** reduces the probability of the buffer being full of cells from about 0.01 to 0.001 when $L_p = 100$.

4.4.3 Effectiveness of EPD1

Figure 4.10 shows the packet loss probability, P_{loss}, as a function of the threshold H of **EPD1**. P_{loss} of both **NC** and **TD** are also given there for comparison. The related parameters are set as follows: $R = 10$, $B = 50$, $L_p = 10$, $\lambda_{\text{all}} = 0.7$ and $\lambda_{\text{on}} = 0.15$. The figure shows that P_{loss} of **EPD1** gets smaller as H becomes larger, but it eventually can not become smaller than that of **TD**. On the other hand, Fig. 4.11 illustrates the distribution of the queue length in the cases of $H = 20$ and 40. The buffer is less congested in the case of $H = 20$ than in the case of $H = 40$, whereas P_{loss} is larger in the former case than in the latter case. **EPD1** with a smaller threshold successfully makes the buffer less congested at the cost of discarding newly arriving packets earlier. Namely, a smaller threshold can decrease the loss probability of already accepted packets, while it can increase that of newly arriving packets. Hence, we expected that there would exist some threshold that provided the better packet loss probability than that in **TD**. However, these figures show that **EPD1** does not succeed in improving the packet loss probability in the system

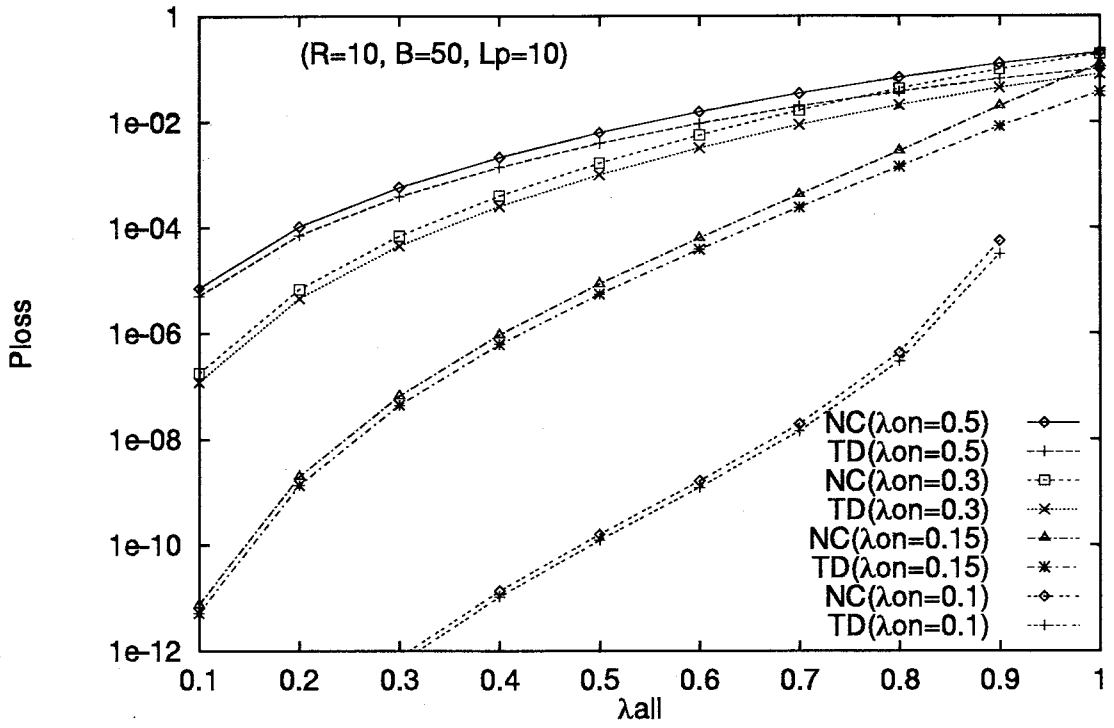


Figure 4.7: Packet loss probability as a function of the total arrival rate λ_{all} at the node

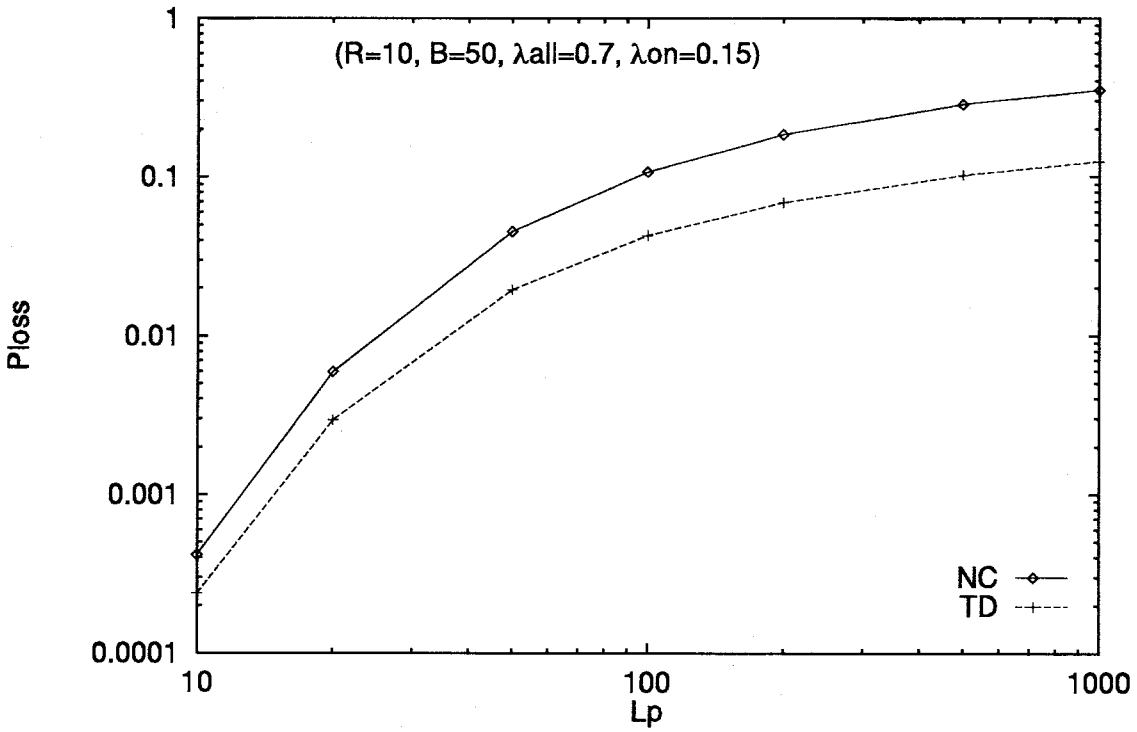


Figure 4.8: Impact of the packet length, L_p

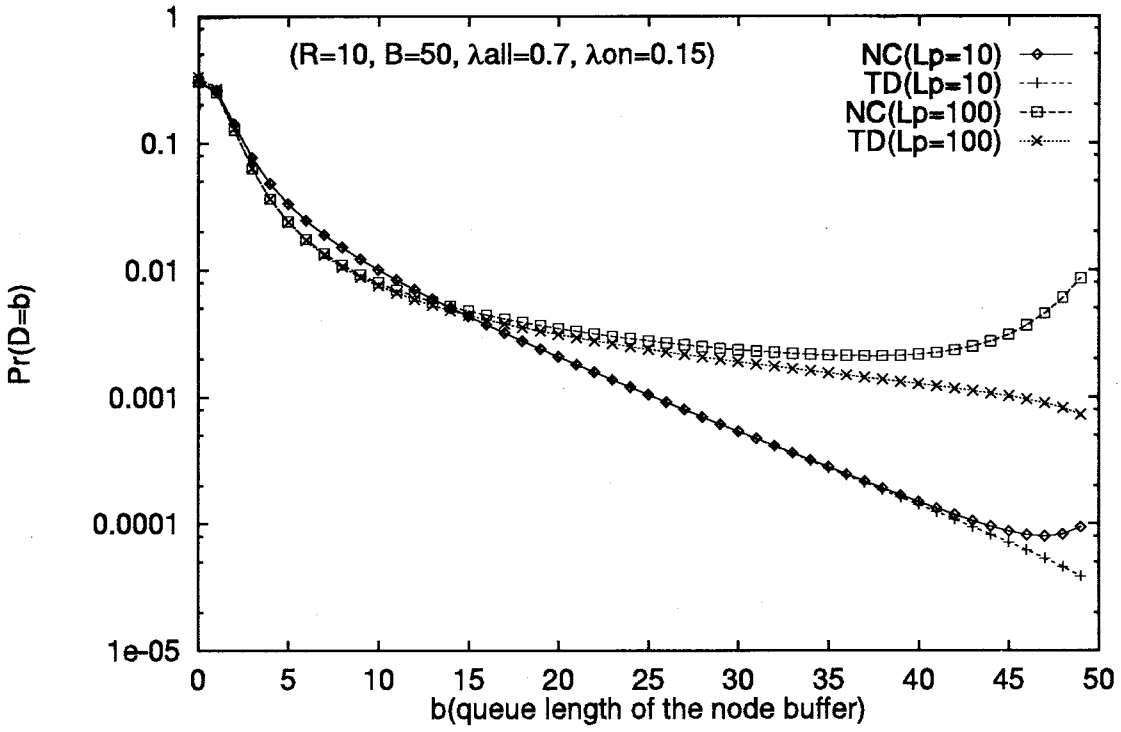


Figure 4.9: Distribution of the queue length: cases of $L_p = 10$ and 100

considered here, in which the packet length follows a geometric distribution.

4.4.4 Effectiveness of EPD2

Figure 4.12 shows P_{loss} as a function of δ for **EPD2**. The related parameters is set as in Fig. 4.10 of **EPD1**. It is shown in this figure that a larger δ causes more packets to be lost; when δ exceeds 0.2, **EPD2** can not improve P_{loss} at all compared with **NC**. We shall compare P_{loss} of **EPD2** with that of **EPD1**, for example, in the case where $\delta = 0.1$. When $R = 10$, $B = 50$ and $L_p = 10$, a set of thresholds $(H_{10}, H_9, \cdots, H_2, H_1)$ becomes $(40, 41, \cdots, 48, 49)$ from Eq. (4.27), i.e., no newly arriving packets are discarded until the queue length of the node reaches 40. In Fig. 4.10, when $H = 40$, P_{loss} of **EPD1** is about 5×10^{-4} whereas P_{loss} of **EPD2** is about 3×10^{-4} if $\delta = 0.1$. This suggests that the threshold should be determined depending on the number of accepted packets, unlike in **EPD1**. A similar result was obtained in terms of throughput performance by means of simulations in [NORI 94], which further has shown that **EPD2** provides better throughput performance than **TD** when packets are of fixed size. On the other hand, we assume packets of variable length which is geometrically distributed. Furthermore, **EPD2** considered here determines thresholds according to the average packet length L_p instead of each actual packet length.

From the above discussion, we see that **EPD** does not always outperform **TD**; in particular, **EPD** does not outperform **TD** when the length of packets are geometrically distributed as in our study.

4.5 Conclusions

Selective cell discard schemes, such as Tail dropping (**TD**) and Early Packet Discard (**EPD**), have been considered as a way to efficiently utilize the network resources by selectively discarding cells or packets that are not worth relaying within the network. In this Chapter, we carried out an exact analysis of the packet loss probability in cases of **TD**, **EPD** based upon a fixed threshold (**EPD1**), and that based upon variable ones (**EPD2**), given that the packet length follows a geometric distribution, and that sources are homogeneous in terms of traffic characteristics. So far, the related works have been based upon simulations or based upon an approximate analysis.

We have obtained the following results by means of numerical results. In comparison with the packet loss probability under no-control, **TD** reduces the packet loss probability slightly, and **TD** is more effective in doing so as the total arrival rate at the node, the peak

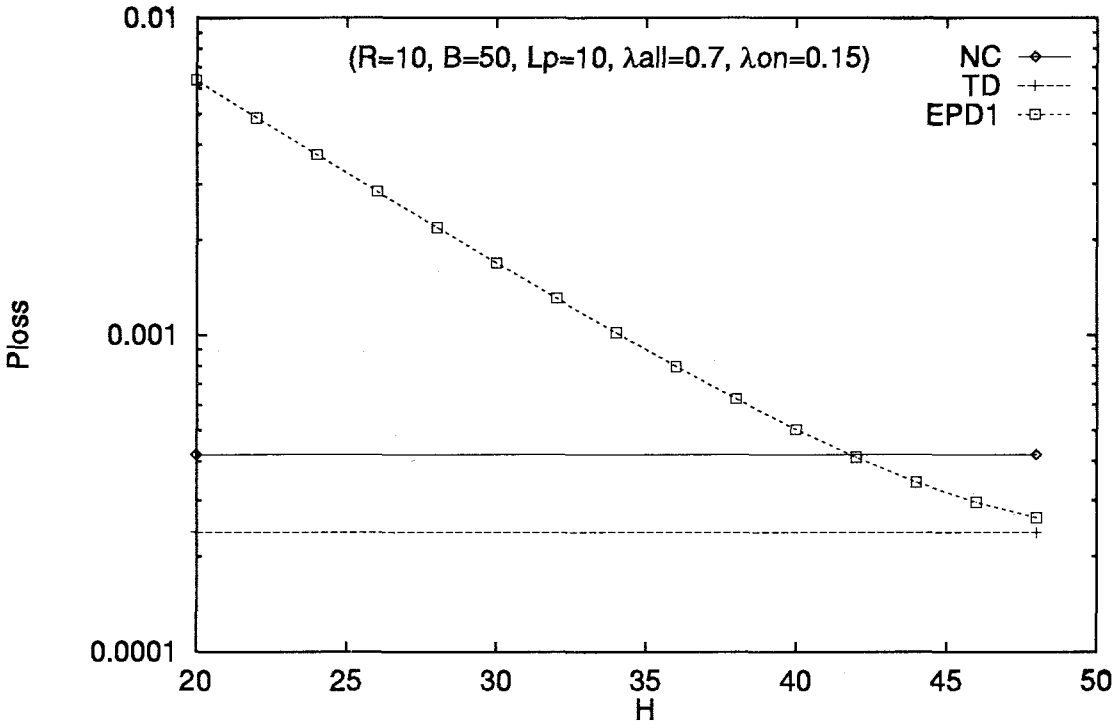


Figure 4.10: Packet loss probability as a function of the threshold H of **EPD1**

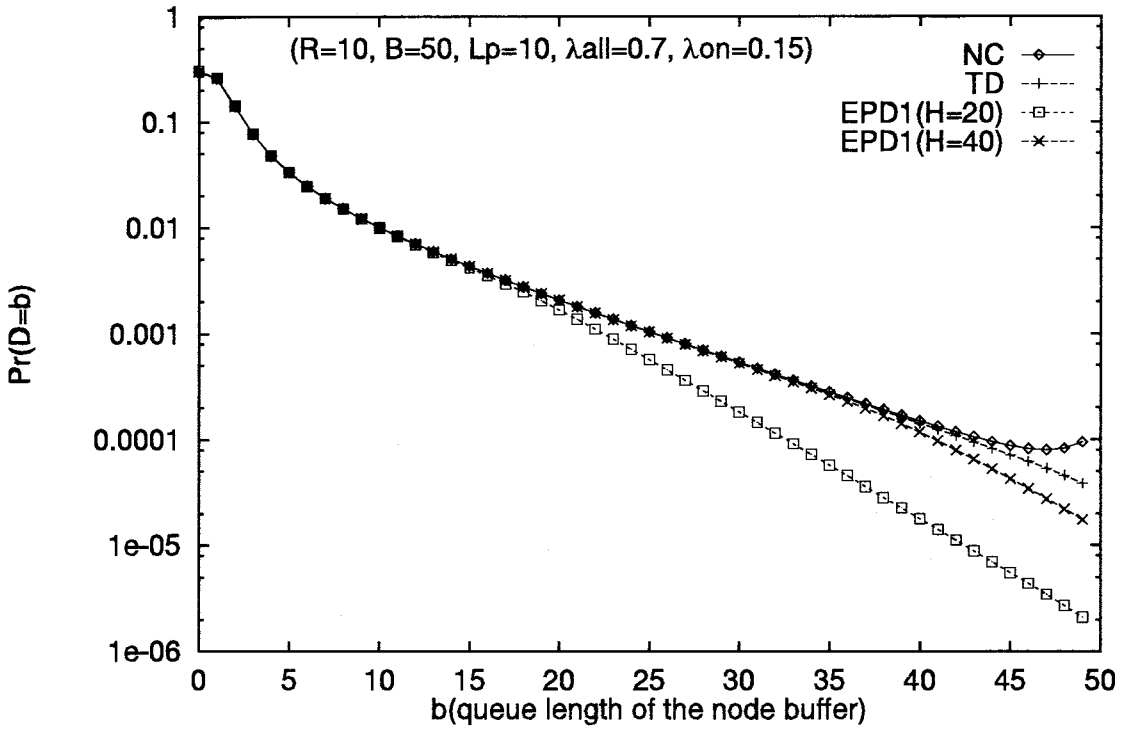


Figure 4.11: Distribution of the queue length: cases of $H_b = 20$ and 40

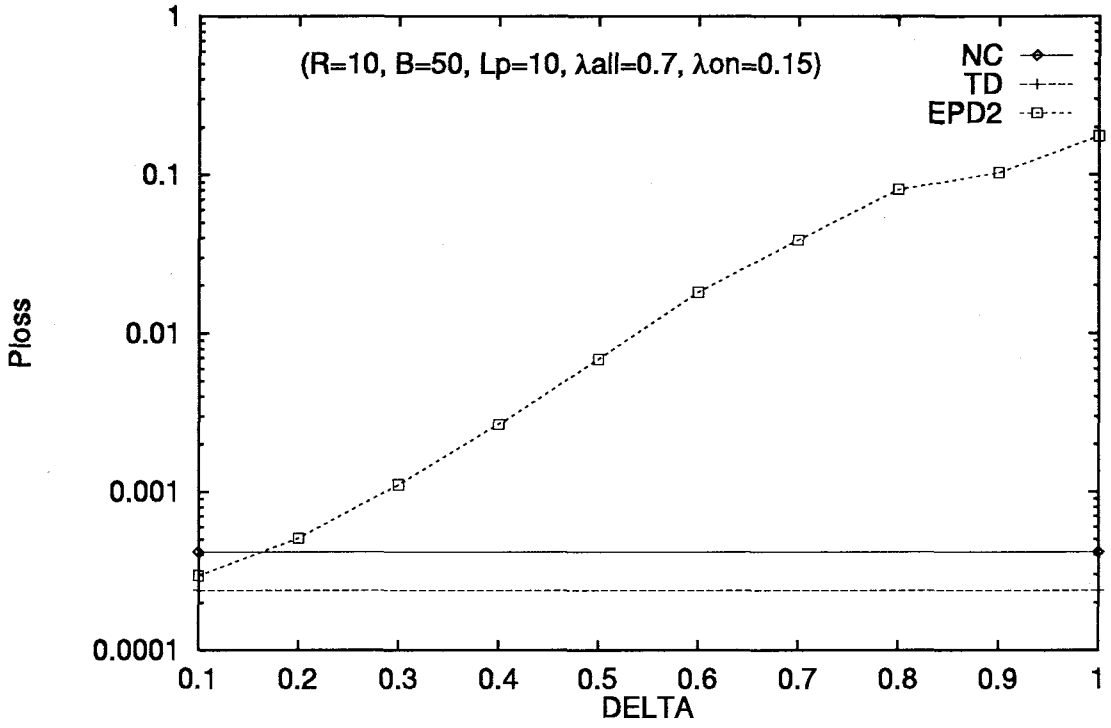


Figure 4.12: Packet loss probability as a function of δ for **EPD2**

rate of each source and the number of sources becomes larger, while the performance of **TD** is insensitive to the buffer size of the node.

On the other hand, the packet loss probability becomes smaller as the threshold is set to a larger value in **EPD1** and as the acceptance coefficient δ is set to a smaller value in **EPD2**. Contrary to the previous works with fixed packet lengths, however, our study shows that both **EPD1** and **EPD2** can not provide better performance than **TD** in terms of the packet loss probability. We expect that this fact comes from the assumption of the geometrically distributed packet length. When the packet length is fixed, newly arriving packets have more cells than the remaining cells which will arrive from accepted packets. Thus, **EPD**s make sense. However, when the packet length follows a geometric distribution, the number of remaining cells which will arrive from an accepted packet has the same distribution as the number of cells in a newly arriving packet, since the residual life of the geometric distribution has the same distribution as itself. Thus, we conclude that the effectiveness of **EPD**s strongly depend on the packet length distribution, and when the packet length has high variation, early packet discard schemes cannot have better performance than **TD**.

Chapter 5

Forward Error Correction: Cell Loss Distribution in a Block

5.1 Introduction

In high speed networks using optical fibers as a transmission media, like ATM networks, cell loss due to congestion becomes dominant as transmission errors. As a way to recover from loss, ARQs (Automatic Repeat reQuest) have prevailed, which rely on sources to retransmit lost cells when they are informed of cell loss happening within the networks by the receivers [BURT 72]. In these retransmission based error recovery schemes, the transmission delay can reach an unacceptable level for applications with stringent delay requirements in high speed networks because their propagation delay is very large relative to packet transmission times. This can result in ARQs being ineffective for such applications.

On the other hand, FEC (Forward Error Correction) has been studied as an error control alternative to ARQ for such applications as mentioned above in the transport protocol TP++ tailored for high speed communications [BIER 92]. This is because FEC allows receivers to correct up to some number of lost cells without retransmissions. In addition, the Reed-Solomon erasure (RSE) code developed in [ANTH 90] has been implemented by the device operating at 1 Gb/s because of its less complexity compared with the conventional Reed-Solomon code. This makes FEC more suitable for high speed communications. Even in the ATM LAN with relatively small propagation, FEC will contribute toward improving QOS in connectionless multicast communications for video conferences which do not employ any ARQs. In FEC scheme based on the RSE code, the FEC encoder produces some redundant cells, say r cells, every predetermined number of data cells, say p cells. Those redundant cells together with data cells form a group called a *block*. In this case, the receiver can correct up to r cells among cells lost in a block.

Even though adding more redundant cells enhances the capability of correcting lost cells, the amount of offered traffic to the network also increases. This can cause in turn cell loss to happen more often. Hence, there is the tradeoff between the enhancement of the capability of correcting lost cells and the increase of cell loss. In the transport layer located above the ATM layer, a block of cells, often referred to as a packet, is processed as a data unit instead of a cell; if some cells are lost in a block, it is very likely that all the cells in the block are eventually discarded at destinations. Accordingly, the block loss probability is a primary performance measure instead of the individual cell loss probability [ERNS 93]. Further, the cell loss distribution in a block is known to be highly correlated, so that we cannot deduce the block loss probability from the individual cell loss probability [CIDO 93].

The main goal of this Chapter is to investigate the effectiveness of FEC in ATM networks which are expected to support diverse applications such as voice, video and data transfer. It is known that many of traffic sources in ATM networks exhibit a fair amount of correlation [HEFF 86]. Thus, the key issue to be considered is to investigate how the correlation of cell arrivals affects the cell loss distribution in a block. To this end, we develop a discrete-time finite-buffer queueing model for a system composed of multiple sources with correlated traffic, where some sources employ FEC. Further we provide an exact analysis of the distribution of the number of cells lost in a block generated by a particular source. Our analytical approach allows us to investigate how the cell loss distribution in a block is affected by various parameters related to both traffic characteristics and FEC based on the RSE code. In particular, we consider the peak arrival rate and burst length as the parameters with regard to traffic characteristics. Also we consider the number of redundant cells in a block and the block size as the parameters with regard to the FEC.

There are several related works [ERNS 93, CIDO 93, SCHA 90, ZANG 92]. In particular, Biersack [ERNS 93] has extensively studied FEC based on the RSE code in some specific contexts of practical interest by means of simulations, and also addressed the related results obtained by approximate analysis in [SCHA 90, ZANG 92]. Contrary to these works, we take an analytical approach to investigate the effectiveness of FEC in ATM networks and show how traffic characteristics have impacts on the effectiveness of FEC. Cidon *et al.* [CIDO 93] have analyzed the number of packets lost in a block, assuming exponential service times (in the continuous-time model) or independent arrivals (in the discrete-time model). Further, in the discrete-time model, they assumed all cells arrive from a single source. Even though they mentioned a possibility to extend their model to a model with multiple sources, they did not provide any specific analysis in fact. Besides, their focus was not on FEC based on the RSE code. As shown in the following sections, we develop and analyze a queueing model with several sources, where some sources employ FEC based on

the RSE code and other sources do not.

The remainder of the Chapter is organized as follows. In section 5.2, we develop a mathematical model for a system composed of multiple sources with correlated traffic, where some of sources employ FEC based on the RSE code. Section 5.3 provides an exact analysis of the distribution of the number of cells lost in a block generated by a particular source. Section 5.4 provides numerical results to investigate how traffic characteristics have impacts on the effectiveness of FEC. Finally, section 5.5 presents concluding remarks.

5.2 Model Description

In this section, we describe our analytical model involving correlated cell arrival processes. By means of this model, we will obtain the distribution of the number of cells lost in a block.

We consider a discrete time queueing system with time slot being equal to a cell transmission time. In our model, there are many sources transmitting their cells to the same multiplexer or switch with an output buffer with capacity of B cells (see Fig. 5.1). In particular, we focus our attention only on one source among them, and call it the *tagged* source to obtain its performance measures such as cell loss and block loss probabilities. Other sources are referred to as *external* sources.

At the tagged source, cells arrive according to an interrupted Bernoulli process(IBP), which is characterized by a parameter set of $(\lambda_{ON}, \alpha, \beta)$. IBP has two states, ON-state and OFF-state. In each slot, the process changes its state from ON to OFF with probability α, and from OFF to ON with probability β. During ON-state, one cell arrives in each slot with probability λ_{ON}, while no cell arrives during OFF-state. Thus, the average arrival rate of the tagged source is given by

$$\lambda = \frac{\beta}{\alpha + \beta}\lambda_{ON}. \tag{5.1}$$

On the other hand, we assume that the aggregated cell arrivals from all external sources are governed by a finite-state Markov chain. Let M denote the number of phase states in the Markov chain. With probability $u_{i,j}$ $(i, j = 1, \ldots, M)$, the phase state changes from state i in the previous slot to state j in the current slot. Furthermore, $a_{i,j}(x)$ denotes the probability that x cells are transmitted to the switch when the phase state changes from state i to state j. The arrival process of external sources is then characterized by the $M \times M$ matrix $\boldsymbol{C}_x$ whose (i,j)th element represents the conditional joint probability that the Markov chain is in state j in the current slot and x cells arrive from external sources given the Markov chain being in state i in the previous slot:

$$\boldsymbol{C}_x = \begin{pmatrix} u_{1,1}a_{1,1}(x) & \cdots & u_{1,M}a_{1,M}(x) \\ \vdots & \ddots & \vdots \\ u_{M,1}a_{M,1}(x) & \cdots & u_{M,M}a_{M,M}(x) \end{pmatrix}. \tag{5.2}$$

This arrival process is called the discrete MAP (DMAP) [BLON 92] and has been used to model bursty traffic [TAKI 93, TAKI 94].

In the remaining of this section, we address the definition of *block* and the function of FEC. When using FEC, after every p cells (referred to as *data cells*) arrive at sources, r cells (referred to as *redundant cells*) are produced by the Reed-Solomon erasure (RSE) coding[ANTH 90] and sent out to the switch. These $p+r$ cells are organized into a group which is referred to as a *block*. Without using FEC, a block would be corrupted when at least one cell is lost in the block. On the other hand, the RSE code producing r redundant cells per p data cells enables us to correct up to r cells lost in the block. This results in reducing the block loss probability. With the increase of r, the receiver could correct more lost cells. However, the actual traffic rate offered to the switch increases as redundant cells increase. This can cause more lost cells. For these reasons, there will be some optimum ratio of r to p. In this paper, in order to take the influence of increasing traffic due to FEC code into consideration, we define the effective cell arrival rate λ'_{ON} in ON state as

$$\lambda'_{ON} = \left(1 + \frac{r}{p}\right) \lambda_{ON}. \tag{5.3}$$

It is noted that λ'_{ON} is not greater than one, so that r should be less than or equal to $(1/\lambda_{ON} - 1)p$.

5.3 Analysis

5.3.1 Definition of State Probability

The state of our system in the mth slot ($m = 1, 2, \ldots$) can be completely described by a set of 1) the queue length $D_r(m)$ of output buffer in the ATM switch, 2) the IBP state $S_s(m)$ of the tagged source, 3) the number $D(m)$ of data cells generated in the current block up to now, and 4) the phase state $S_e(m)$ of external sources. Note that $\{D_r(m), S_s(m), D(m), S_e(m)\}$ forms a discrete-time Markov chain with finite states. We define $x(b, s_s, d, s_e)$ as the steady state probabilities of the Markov chain $\{D_r(m), S_s(m), D(m), S_e(m)\}$:

$$x(b, s_s, d, s_e) \triangleq \lim_{m\to\infty} \Pr\{D_r(m) = b, S_s(m) = s_s, D(m) = d, S_e(m) = s_e\}. \tag{5.4}$$

In Appendix, we provide the analysis of the steady state probability $x(b, s_s, d, s_e)$.

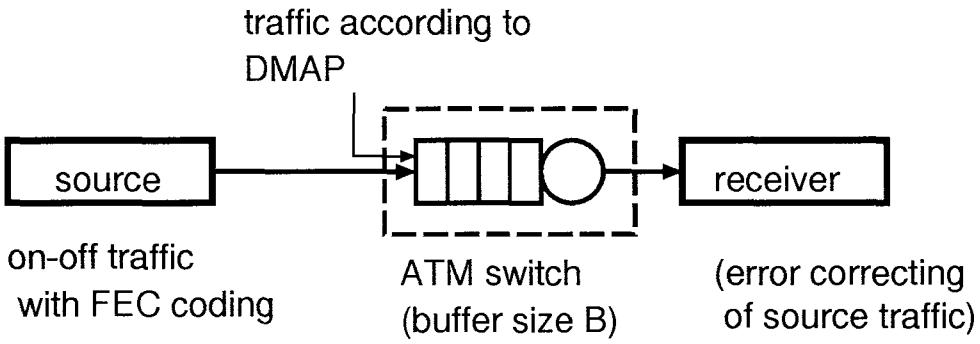


Figure 5.1: Analytical model for FEC

5.3.2 Derivation of Cell Loss and Block Loss Probabilities

In order to obtain the distribution of the number of cells lost in a block, we have to find the slot in which the tagged source has finished transmitting all the cells in the previous block. In that slot, the tagged source is in state $(S_s(m), D(m)) = (1, p+r)$. Hence, first we obtain the steady state probability denoted by $y(b, s_e)$ with respect to random variables D_r and S_e in the slot given that the tagged source first visits state $(1, p+r)$. The vector having those probabilities as elements is defined as

$$\boldsymbol{y} \triangleq [\cdots, y(b, s_e), \cdots], \ 0 \le b \le B, \ 1 \le s_e \le M. \tag{5.5}$$

The state $(1, p+r)$ can be visited only from the states $(0, p+r-1)$ and $(1, p+r-1)$. The steady state probabilities of these states are denoted by $\boldsymbol{x}(0, p+r-1)$ and $\boldsymbol{x}(1, p+r-1)$:

$$\boldsymbol{x}(n, p+r-1) \triangleq [\cdots, x(b, n, p+r-1, s_e), \cdots] \qquad n = 0, 1, \ 0 \le b \le B, \ 1 \le s_e \le M. \tag{5.6}$$

Using $\boldsymbol{x}(0, p+r-1)$ and $\boldsymbol{x}(1, p+r-1)$, we have

$$\boldsymbol{y} = \frac{\{\lambda_{ON}(1-\alpha)\boldsymbol{x}(1, p+r-1) + \lambda_{ON}\beta\boldsymbol{x}(0, p+r-1)\}\left(\tilde{\boldsymbol{P}}(1) + \tilde{\boldsymbol{P}}_L(1)\right)}{\{\lambda_{ON}(1-\alpha)\boldsymbol{x}(1, p+r-1) + \lambda_{ON}\beta\boldsymbol{x}(0, p+r-1)\}\, \boldsymbol{e}[(B+1)M]}, \tag{5.7}$$

where $\boldsymbol{e}[n]$ denotes an $n \times 1$ vector whose all elements are equal to one, $\tilde{\boldsymbol{P}}_L(1)$ ($\tilde{\boldsymbol{P}}(1)$) represents the transition probability matrix with respect to random variables D_r and S_e given that a cell transmitted from the tagged source is (not) lost in the output buffer:

$$\tilde{\boldsymbol{P}}_L(1) = \begin{pmatrix} \boldsymbol{0}[M] \cdots \boldsymbol{0}[M] & \boldsymbol{0}[M] & \overline{\boldsymbol{C}_{B-1}} - \hat{\boldsymbol{C}}_{B-1}(1) \\ \vdots \ \vdots \ \vdots & \overline{\boldsymbol{C}_{B-2}} - \hat{\boldsymbol{C}}_{B-2}(1) & \boldsymbol{0}[M] \\ \vdots \ \vdots \ \vdots & \vdots & \vdots \\ \vdots \ \vdots \ \vdots & \overline{\boldsymbol{C}_1} - \hat{\boldsymbol{C}}_1(1) & \boldsymbol{0}[M] \\ \vdots \ \vdots \ \vdots & \overline{\boldsymbol{C}_0} - \hat{\boldsymbol{C}}_0(1) & \boldsymbol{0}[M] \\ \underbrace{\boldsymbol{0}[M] \cdots \boldsymbol{0}[M]}_{B-1 \ \ matrices} & \overline{\boldsymbol{C}_0} & \boldsymbol{0}[M] \end{pmatrix}, \tag{5.8}$$

$$\tilde{\boldsymbol{P}}(1) = \begin{pmatrix} \boldsymbol{0}[M] & \boldsymbol{C}_0 & \boldsymbol{C}_1 & \cdots & \boldsymbol{C}_{B-3} & \boldsymbol{C}_{B-2} & \hat{\boldsymbol{C}}_{B-1}(1) \\ \boldsymbol{0}[M] & \boldsymbol{C}_0 & \boldsymbol{C}_1 & \cdots & \boldsymbol{C}_{B-3} & \hat{\boldsymbol{C}}_{B-2}(1) & \boldsymbol{0}[M] \\ \vdots & \ddots & \boldsymbol{C}_0 & \cdots & \boldsymbol{C}_{B-4} & \hat{\boldsymbol{C}}_{B-3}(1) & \boldsymbol{0}[M] \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \boldsymbol{C}_0 & \hat{\boldsymbol{C}}_1(1) & \boldsymbol{0}[M] \\ \vdots & \vdots & \vdots & \vdots & \ddots & \hat{\boldsymbol{C}}_0(1) & \boldsymbol{0}[M] \\ \boldsymbol{0}[M] & \cdots & \cdots & \cdots & \cdots & \boldsymbol{0}[M] & \boldsymbol{0}[M] \end{pmatrix}, \tag{5.9}$$

where $\mathbf{0}[M]$ denotes a zero matrix of dimension $M \times M$, $\boldsymbol{C}_x$ is given in (5.2), $\overline{\boldsymbol{C}_x}$ represents the transitions when x or more cells are transmitted:

$$\overline{\boldsymbol{C}_x} \triangleq \sum_{k=x}^{\infty} \boldsymbol{C}_k, \tag{5.10}$$

and

$$\hat{\boldsymbol{C}}_i(1) = \sum_{k=i}^{\infty} \boldsymbol{C}_k \frac{\binom{k}{i}}{\binom{k+1}{i+1}} \tag{5.11}$$

represents the transition probability matrix given that the buffer has still capacity enough to accommodate up to $i+1$ cells and the cell transmitted from the tagged source is successfully buffered.

To obtain the distribution of the number of cells lost in the block, we define $y_{b,s_e}^{(n)}(i)$ as

$$y_{b,s_e}^{(n)}(i) \triangleq \Pr\{L^{(n)} = i, D_r = b, S_e = s_e | S_s = 1, D = n\}, \tag{5.12}$$

where $L^{(n)}$ denotes the number of lost cells among first n cells in a block. We also define $\boldsymbol{y}^{(n)}(i)$ as its probability vector:

$$\boldsymbol{y}^{(n)}(i) \triangleq [\cdots, y_{b,s_e}^{(n)}(i), \cdots], \qquad 0 \le b \le B,\ 1 \le s_e \le M. \tag{5.13}$$

Moreover, let $\boldsymbol{P}$ denote the transition probability matrix with regard to random variables D_r and S_e when no cell is transmitted from the tagged source:

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{C}_0 & \boldsymbol{C}_1 & \cdots & \boldsymbol{C}_{B-2} & \boldsymbol{C}_{B-1} & \overline{\boldsymbol{C}_B} \\ \boldsymbol{C}_0 & \boldsymbol{C}_1 & \cdots & \boldsymbol{C}_{B-2} & \overline{\boldsymbol{C}_{B-1}} & \mathbf{0}[M] \\ \mathbf{0}[M] & \boldsymbol{C}_0 & \cdots & \boldsymbol{C}_{B-3} & \overline{\boldsymbol{C}_{B-2}} & \mathbf{0}[M] \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \boldsymbol{C}_0 & \overline{\boldsymbol{C}_1} & \mathbf{0}[M] \\ \mathbf{0}[M] & \cdots & \cdots & \mathbf{0}[M] & \overline{\boldsymbol{C}_0} & \mathbf{0}[M] \end{pmatrix}. \tag{5.14}$$

Therefore, we obtain $\boldsymbol{y}^{(n)}(i)$:

$$\boldsymbol{y}^{(n)}(i) = \begin{cases} \eta_1 \boldsymbol{y}\tilde{\boldsymbol{P}}(1) + \boldsymbol{y}\boldsymbol{P}\sum_{j=0}^{\infty}\eta_{j+2}\boldsymbol{P}^j\tilde{\boldsymbol{P}}(1), & n = 1, i = 0, \\ \eta_1 \boldsymbol{y}\tilde{\boldsymbol{P}}_L(1) + \boldsymbol{y}\boldsymbol{P}\sum_{j=0}^{\infty}\eta_{j+2}\boldsymbol{P}^j\tilde{\boldsymbol{P}}_L(1), & n = 1, i = 1, \\ \eta_1 \boldsymbol{y}^{(n-1)}(0)\tilde{\boldsymbol{P}}(1) + \boldsymbol{y}^{(n-1)}(0)\boldsymbol{P}\sum_{j=0}^{\infty}\eta_{j+2}\boldsymbol{P}^j\tilde{\boldsymbol{P}}(1), & 2 \le n \le p+r, i = 0, \\ \eta_1 \boldsymbol{y}^{(n-1)}(i)\tilde{\boldsymbol{P}}(1) + \boldsymbol{y}^{(n-1)}(i)\boldsymbol{P}\sum_{j=0}^{\infty}\eta_{j+2}\boldsymbol{P}^j\tilde{\boldsymbol{P}}(1) & \\ \quad + \eta_1 \boldsymbol{y}^{(n-1)}(i-1)\tilde{\boldsymbol{P}}_L(1) + \boldsymbol{y}^{(n-1)}(i-1)\boldsymbol{P}\sum_{j=0}^{\infty}\eta_{j+2}\boldsymbol{P}^j\tilde{\boldsymbol{P}}_L(1), & \\ \qquad 2 \le n \le p+r, \quad 1 \le i \le n-1, & \\ \eta_1 \boldsymbol{y}^{(n-1)}(n-1)\tilde{\boldsymbol{P}}_L(1) + \boldsymbol{y}^{(n-1)}(n-1)\boldsymbol{P}\sum_{j=0}^{\infty}\eta_{j+2}\boldsymbol{P}^j\tilde{\boldsymbol{P}}_L(1), & \\ \qquad 2 \le n \le p+r, i = n, & \end{cases} \tag{5.15}$$

where η_i denotes the probability that the length of the first passage time from state $(1, p+r)$ to state $(1,1)$ is equal to i slots. (Note that all the first passage times from state $(1,n)$ to state $(1,n+1)$ $(1 \leq n \leq p+r-1)$ have the same distribution as that from state $(1,p+r)$ to state (1,1).) To obtain η_i, we need the transition probabilities among three states : $(1,1)$, $(0,p+r)$, and $(1,p+r)$, since the state $(1,p+r)$ can transits only these states.

$$\begin{array}{c} \\ (1,1) \\ (1,p+r) \\ (0,p+r) \end{array} \begin{array}{c} \begin{array}{ccc} (1,1) & (1,p+r) & (0,p+r) \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ \lambda_{ON}(1-\alpha) & (1-\lambda_{ON})(1-\alpha) & \alpha \\ \lambda_{ON}\beta & (1-\lambda_{ON})\beta & 1-\beta \end{pmatrix} \end{array}. \tag{5.16}$$

Using these transition probabilities, η_i is given by

$$\eta_i = \begin{cases} \lambda_{ON}(1-\alpha), & \text{if } i=1 \\ ((1-\lambda_{ON})(1-\alpha), \alpha) \begin{pmatrix} (1-\lambda_{ON})(1-\alpha) & \alpha \\ (1-\lambda_{ON})\beta & 1-\beta \end{pmatrix}^{i-2} \begin{pmatrix} \lambda_{ON}(1-\alpha) \\ \lambda_{ON}\beta \end{pmatrix} & \text{if } i \geq 2 \end{cases}. \tag{5.17}$$

From the definition of $\boldsymbol{y}^{(n)}(i)$, we see that $\boldsymbol{y}^{(p+r)}(i)$ $(i = 0, 1, \cdots, p+r)$ denotes the probability that i cells are lost among $p+r$ cells constituting a block. Thus, we have

$$\Pr\{L = i\} = \boldsymbol{y}^{(p+r)}(i)\boldsymbol{e}[(B+1)M], \tag{5.18}$$

where L denotes the number of cells lost in a block. By means of this probability, it is straightforward to obtain the cell loss probability distribution $F_{loss}(x)$, and the block loss probability B_{loss}:

$$F_{loss}(x) = \sum_{i=0}^{x} \Pr\{L = i\}, \tag{5.19}$$

$$B_{loss} = \sum_{i=r+1}^{p+r} \Pr\{L = i\}. \tag{5.20}$$

5.4 Numerical Results

As mentioned in Section 2, we assume a general stochastic model (DMAP) to represent the behavior of the external sources. In obtaining numerical results, it is convenient to exploit a more specific model which is mostly used to represent a correlated cell arrival process. In this section, we assume that cells arrive at each of external sources according to IBP as well as the tagged source. Further we assume that there exist N_e external sources and they are divided into E groups of sources, where each of $N_e^{(j)}$ sources in the jth group has the same parameters $(\lambda_j, \alpha_j, \beta_j)$ in terms of IBP. Note that $\sum_{j=1}^{E} N_e^{(j)} = N_e$. This model

is convenient to handle a heterogeneous system of sources, e.g., some of them use FEC whereas others do not, or some of them have bursty traffic while others do not, and so on.

First, defining the random variable $S_e^{(j)}$ that indicates the number of IBP sources being in ON state among the jth group of sources, we have the following transient probability that $S_e^{(j)}$ will make the transition from state k in the mth slot to state l in the $(m+1)$st slot:

$$\begin{aligned} p_{k,l}^{(j)} &\triangleq \Pr\{S_e^{(j)}(m+1) = l | S_e^{(j)}(m) = k\} \\ &= \sum_{i=\max(0,l-N_e^{(j)}+k)}^{\min(k,l)} \binom{k}{i} (1-\alpha_j)^i \alpha_j^{k-i} \binom{N_e^{(j)}-k}{N_e^{(j)}-k-l+i} (1-\beta_j)^{N_e^{(j)}-k-l+i} \beta_j^{l-i}, \\ & \qquad 1 \le j \le E. \end{aligned} \tag{5.21}$$

We also define $a_l^{(j)}(x)$ as the probability that x cells arrive when l external sources of the jth group are in ON state:

$$a_l^{(j)}(x) = \binom{l}{x} (\lambda_j)^x (1-\lambda_j)^{l-x}, \qquad 0 \le l \le N_e^{(j)}, 1 \le j \le E. \tag{5.22}$$

The transition probability matrix $\boldsymbol{C}_x^{(j)}$ for the jth group hence becomes

$$\boldsymbol{C}_x^{(j)} = \begin{pmatrix} p_{0,0}^{(j)} a_0^{(j)}(x) & p_{0,1}^{(j)} a_1^{(j)}(x) & \cdots & p_{0,N_e^{(j)}}^{(j)} a_{N_e^{(j)}}^{(j)}(x) \\ p_{1,0}^{(j)} a_0^{(j)}(x) & p_{1,1}^{(j)} a_1^{(j)}(x) & \cdots & p_{1,N_e^{(j)}}^{(j)} a_{N_e^{(j)}}^{(j)}(x) \\ \vdots & \vdots & \ddots & \vdots \\ p_{N_e^{(j)},0}^{(j)} a_0^{(j)}(x) & p_{N_e^{(j)},1}^{(j)} a_1^{(j)}(x) & \cdots & p_{N_e^{(j)},N_e^{(j)}}^{(j)} a_{N_e^{(j)}}^{(j)}(x) \end{pmatrix}. \tag{5.23}$$

Thus, we can rewrite $\boldsymbol{C}_x$ given by eq. (5.2) as follows.

$$\boldsymbol{C}_x = \sum_{\sum_{j=1}^{E} x_j = x} \boldsymbol{C}_{x_1}^{(1)} \otimes \cdots \otimes \boldsymbol{C}_{x_E}^{(E)}. \tag{5.24}$$

In particular, we suppose that the system considered below consists of 16 sources; there are one tagged source and 15 external sources ($N_e = 15$), and focus on the following two cases: homogeneous and heterogeneous cases. In the homogeneous case, all of external sources have the same parameters as those of the tagged one, and thus $E = 1$. We then investigate the impact of 1) the buffer size, 2) traffic parameters, and 3) block size on the performance. On the other hand, in the heterogeneous case, external sources are divided into two groups; one includes sources applying FEC and the other consists of sources not applying FEC, and thus $E = 2$. With this setting, we examine the performance degradation of sources that do not apply FEC.

With respect to the performance of FEC, the following two features are worth noting. First, increasing the number r of redundant cells per block enhances the capability of regenerating lost cells at destination nodes. Second, this causes, at the same time, the effective arrival rate of incoming traffic to the network to increase as shown in eq.(5.3), thereby increasing cell loss probability. Therefore, in the following subsections, we present some numerical results on these features along with the block loss probability. That is, 1) block loss probability when sources do not apply FEC for various values of the total arrival rate λ_{all} multiplexed at one output port, 2) the distribution of the number of cells lost in a block, and 3) the block loss probability for various values of the number r of redundant cells per block. The first two performance measures help us to understand how the block loss probability can be improved or not by means of FEC.

5.4.1 The Impact of the Buffer Size

In this subsection, we investigate the impact of the buffer size B at the output port on ATM switch on the performance. In Fig. 5.2, we fix the average ON duration (hereafter, we call it *burst length*) $1/\alpha = 400$, the peak rate of sources $\lambda_{ON} = 0.05$ and the block size $p = 20$. Thus p is equal to the average number λ_{ON}/α of cells per one burst.

Figure 5.2-(a) shows the impact of B on the block loss probability $B_{loss}(r = 0)$ without using FEC for various values of the total arrival rate λ_{all}. Since λ_{all} is given by

$$\lambda_{all} = 16\lambda = 16\frac{\beta}{\alpha + \beta}\lambda_{ON}, \tag{5.25}$$

the increase of λ_{all} implies the increase of β (that is, the decrease of the average OFF duration $1/\beta$). As expected, $B_{loss}(r = 0)$ for a specific value of λ_{all} decreases as the buffer size B increases.

Figure 5.2-(b) shows the block loss probability B_{loss} as a function of the number r of redundant cells when λ_{all} is chosen in a way to keep $B_{loss}(r = 0)$ equal to 10^{-6}. We set the traffic parameters as in Fig. 5.2-(a) except for r. This figure shows the benefits due to FEC and its related characteristics. We observe that FEC is more effective in improving B_{loss} when the buffer size B is smaller.

In order to show the effectiveness of applying FEC on the overall performance of the system, we define the gain G as

$$G \triangleq \log \frac{B_{loss}(r = 0)}{B_{loss}(r = r_{opt})}, \tag{5.26}$$

where r_{opt} denotes the optimum number of redundant cells per block that makes B_{loss} smallest. Table 5.1 shows the gain for various values of B in such a case that $B_{loss}(r =$

$0) = 10^{-6}$ and 10^{-9}. We see that r_{opt} and G become smaller as B is getting larger. Furthermore, by comparing two cases of $B_{loss}(r = 0) = 10^{-6}$ and 10^{-9}, we observe that FEC is more effective for a smaller block loss probability B_{loss}, e.g., when $B = 16$ and $B_{loss}(r = 0) = 10^{-9}$, FEC succeeds in decreasing B_{loss} by three orders of magnitude.

5.4.2 The Impact of the Traffic Parameters

In this subsection, we show the impact of the traffic parameters such as the peak rate λ_{ON} and the burst length $1/\alpha$.

The impact of the peak rate λ_{ON}

In Fig. 5.3, we investigate the impact of the peak rate λ_{ON}. We fix the buffer size $B = 16$ and the block size $p = 20$. In this setting, we vary the peak rate λ_{ON}, while keeping $p = \lambda_{ON}/\alpha$ constant. Thus varying the peak rate λ_{ON} also implies varying the average ON duration $1/\alpha$.

Figure 5.3-(a) illustrates $B_{loss}(r = 0)$ as a function of the total arrival rate λ_{all} for various values of the peak rate λ_{ON}. We observe that $B_{loss}(r = 0)$ for a given λ_{all} increases as the peak rate λ_{ON} increases.

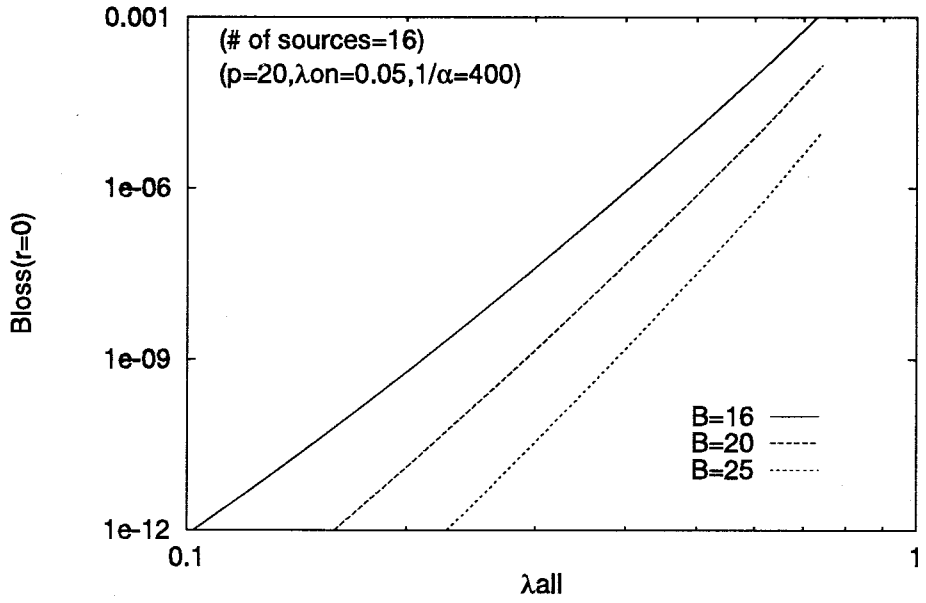
Figure 5.3-(b) shows B_{loss} as a function of the number r of redundant cells. We observe that FEC is effective for a small value of the peak rate λ_{ON}.

Table 5.2 shows that FEC is ineffective when the peak rate $\lambda_{ON} = 0.1$ and 0.2. Note that, when $\lambda_{ON} = 0.05, 0.1$ and 0.2, the sums of the peak rates of all sources is given by $16\lambda_{ON} = 0.8, 1.6$ and 2.4, respectively. Thus we conclude that FEC is ineffective when the sum of the peak rates of all sources exceeds 1.0.

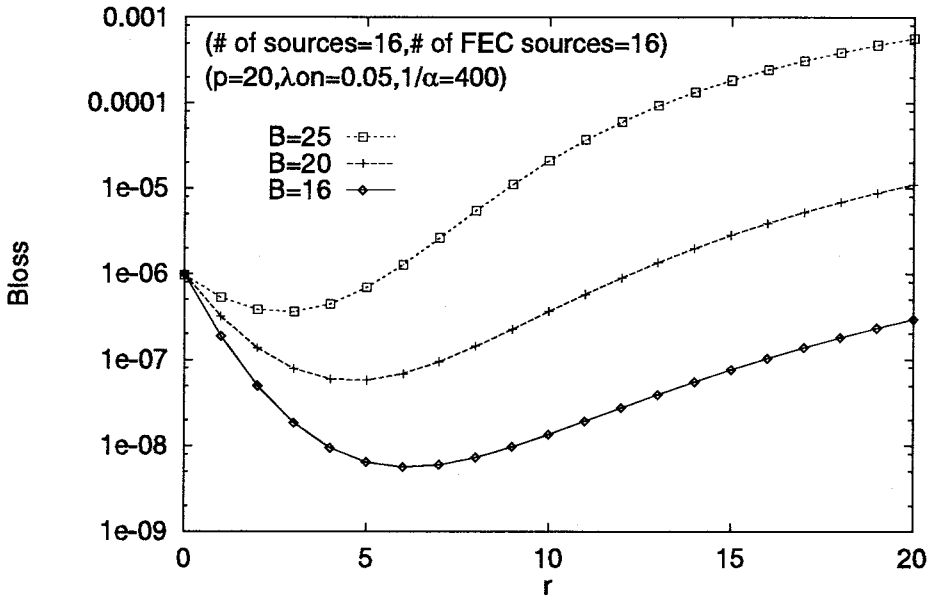
The impact of the burst length $1/\alpha$

In Fig. 5.4, we investigate the impact of the average burst length $1/\alpha$. The peak rate λ_{ON} and the block size p are fixed to 0.05 and 20, respectively. In this setting, we change the burst length $1/\alpha$ into 400, 2000, 4000, and 40000 slots which are of 1, 5, 10, and 100 blocks, respectively.

In Fig. 5.4-(a), there are not significant differences in those performance measures for different burst lengths. However, from Fig. 5.4-(b), we observe that FEC is more effective when $1/\alpha$ is smaller. In order to understand this fact, we show the survivor function $1 - F_{loss}(x)$ when $r = 0$ (that is, the whole sources do not apply FEC) in Fig. 5.4-(c). $1 - F_{loss}(x)$ denotes the probability of $x + 1$ or more cells lost in a block, and would indicate the block loss probability in using FEC of x redundant cells if we do not take into



(a)block loss probability $B_{loss}(r = 0)$ versus λ_{all}



(b)$B_{loss}(r)$ versus the number r of redundant cells

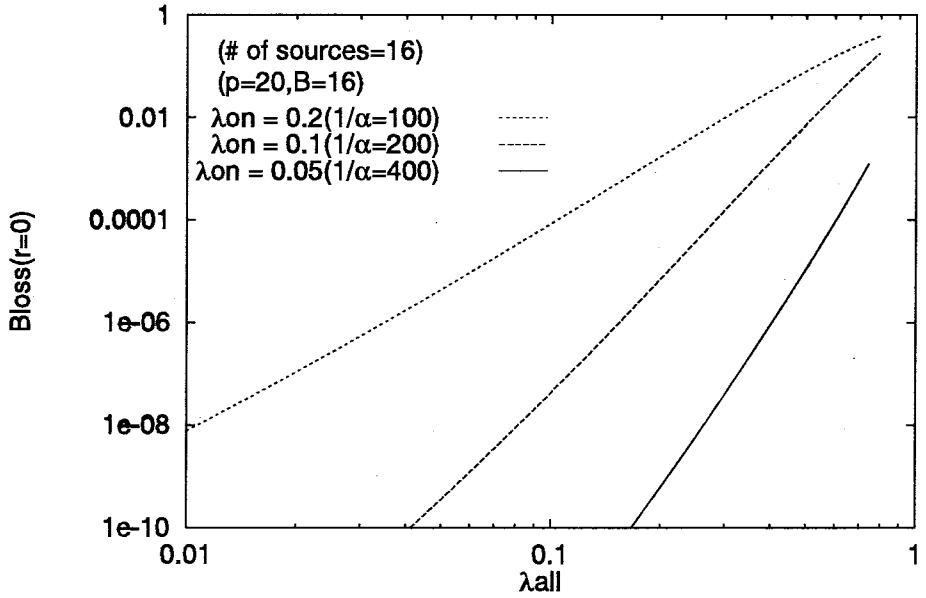
Figure 5.2: Impact of the buffer size on cell loss characteristics

Table 5.1: Impact of B on r_{opt} and the gain G

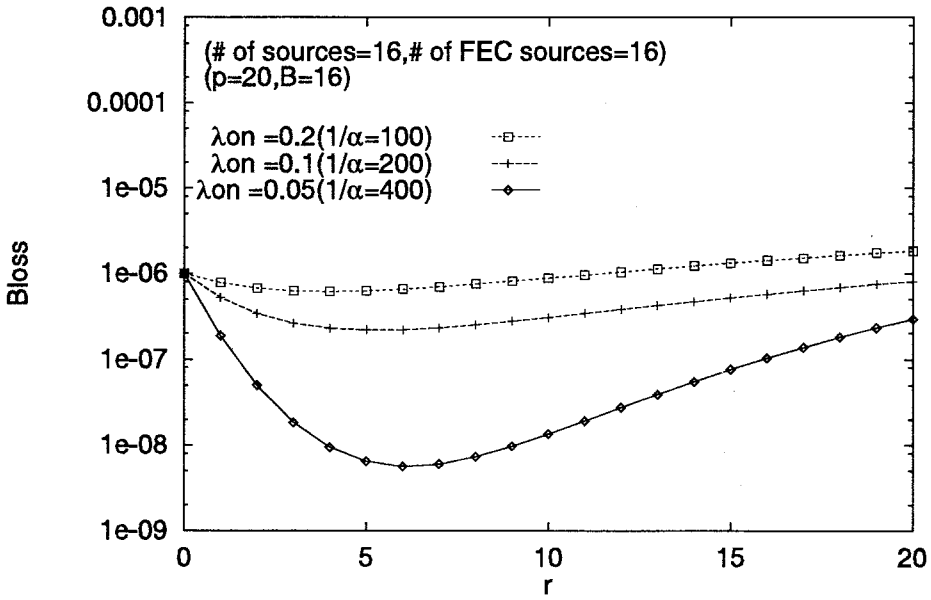
	$B_{loss}(r=0)=10^{-6}$		$B_{loss}(r=0)=10^{-9}$	
B	r_{opt}	gain G	r_{opt}	gain G
16	6	2.25	7	3.06
20	5	1.24	6	1.68
25	3	0.44	4	0.68

Table 5.2: Impact of λ_{ON} on r_{opt} and the gain G

	$B_{loss}(r=0)=10^{-6}$		$B_{loss}(r=0)=10^{-9}$	
λ_{on}	r_{opt}	gain G	r_{opt}	gain G
0.05	6	2.25	7	3.06
0.1	5	0.66	5	0.91
0.2	4	0.21	3	0.17



(a)block loss probability $B_{loss}(r = 0)$ versus λ_{all}



(b)$B_{loss}(r)$ versus the number r of redundant cells

Figure 5.3: Impact of the peak rate λ_{ON} on loss characteristics

account the increase of the block size due to the redundant cells. In this figure, $1 - F_{loss}(x)$ decreases more slowly with x when the burst length $1/\alpha$ is larger. This implies that more cells in a block can be lost when $1/\alpha$ becomes larger as long as the block loss probability is fixed. This is why FEC is more effective when $1/\alpha$ is smaller.

Table 5.3 shows r_{opt} and the gain G for various values of $1/\alpha$. As the burst length becomes larger, both r_{opt} and G are getting smaller, whereas they are not very sensitive to the burst length when the burst length is large.

5.4.3 The Impact of the Block Size

In Fig. 5.5, we investigate the impact of the number of data cells p in a block. We fix the buffer size $B = 16$ and the peak rate $\lambda_{ON} = 0.05$.

We set p to $10, 20, 30$ and 50, while keeping $\lambda_{ON} = p\alpha$ constant. Thus the burst length $1/\alpha$ is set to $200, 400, 600$ and 1000, respectively.

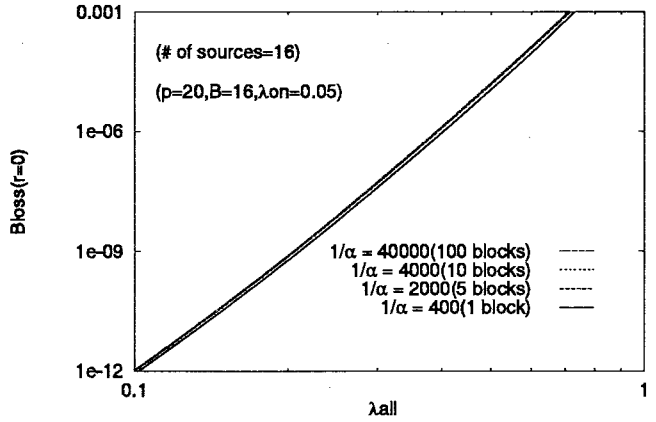
Figure 5.5-(a) shows the block loss probability B_{loss} as a function of the ratio of the number r of redundant cells to the number p of data cells, where λ_{all} is chosen in such a way to keep $B_{loss}(r = 0)$ equal to 10^{-6} [1]. As can be seen from this figure, FEC is remarkably effective when p is relatively large although $1/\alpha$ is also large in such a case. We shall show what leads to this feature. Fig. 5.5-(b) shows the survivor function $1 - F_{loss}(x)$ when the amount of traffic is fixed. We observe that $1 - F_{loss}(x)$ decreases very quickly with x/p when p is relatively large. Thus, for a fixed value of r/p, this figure shows how effectively a large p can enhance the capability of the destination to correct lost cells.

Furthermore, as shown in Table 5.4, we have a significant improvement on the block loss probability with a large p in which the ratio of the optimum number r_{opt} of redundant cells to p is around 0.3.

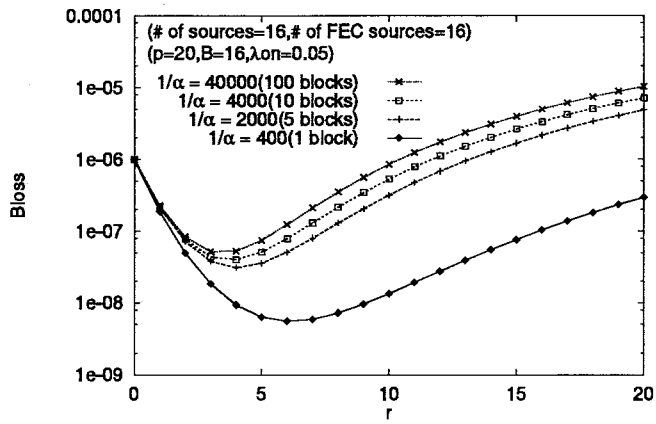
5.4.4 The Impact of the Number of FEC Sources

Figure 5.6 illustrates the impact of the number of FEC sources. Figure 5.6-(a) and 5.6-(b) show the block loss probabilities of the tagged source that applies and does not apply FEC, which are denoted by $B_{loss-FEC}$ and $B_{loss-nonFEC}$, respectively. We fix the buffer size $B = 16$, the peak rate $\lambda_{ON} = 0.05$, the burst length $1/\alpha = 200$ and the block size $p = 10$. The number of FEC sources are set to $0, 4, 8, 12$ and 16 while the total number of sources is kept equal to 16.

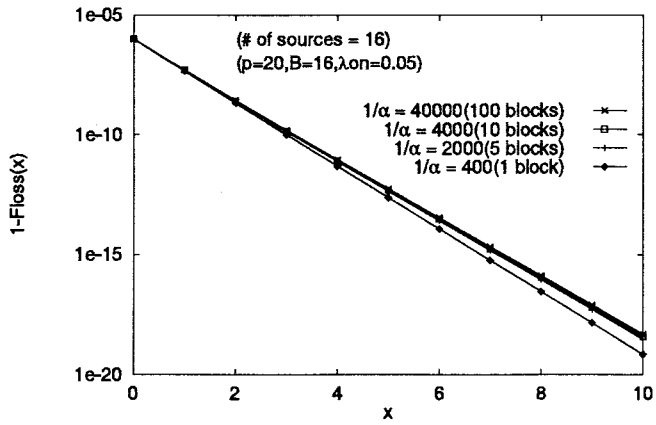
[1] Since there are not significant differences in $B_{loss}(r = 0)$ even when p is changed, we here omit the figure corresponding to Fig. 5.4-(a).



(a)block loss probability $B_{loss}(r = 0)$ versus λ_{all}



(b)$B_{loss}(r)$ versus the number r of redundant cells



(c)survivor function $1 - F_{loss}(x)$ without using FEC

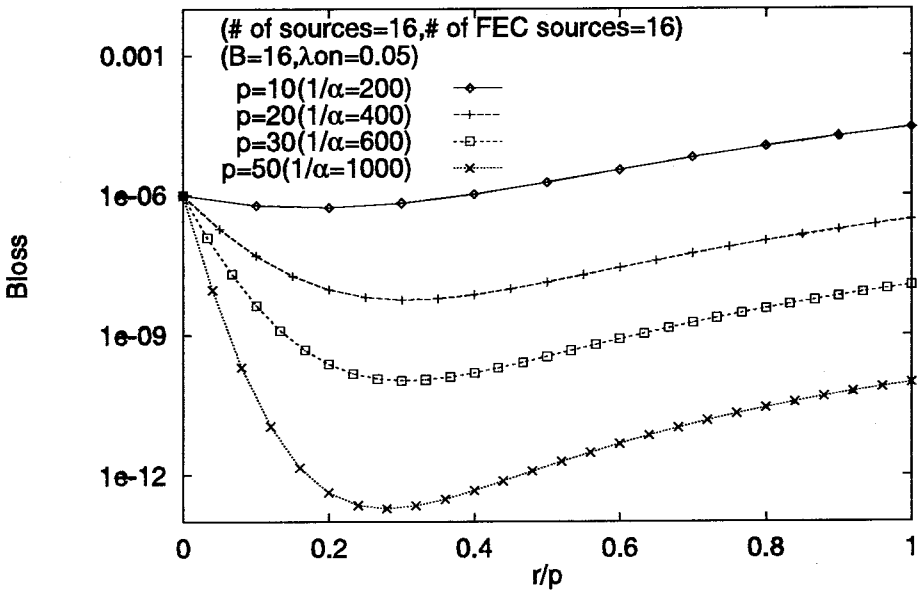
Figure 5.4: Impact of the burst length $1/\alpha$ on loss characteristics

Table 5.3: Impact of $1/\alpha$ on r_{opt} and the gain G

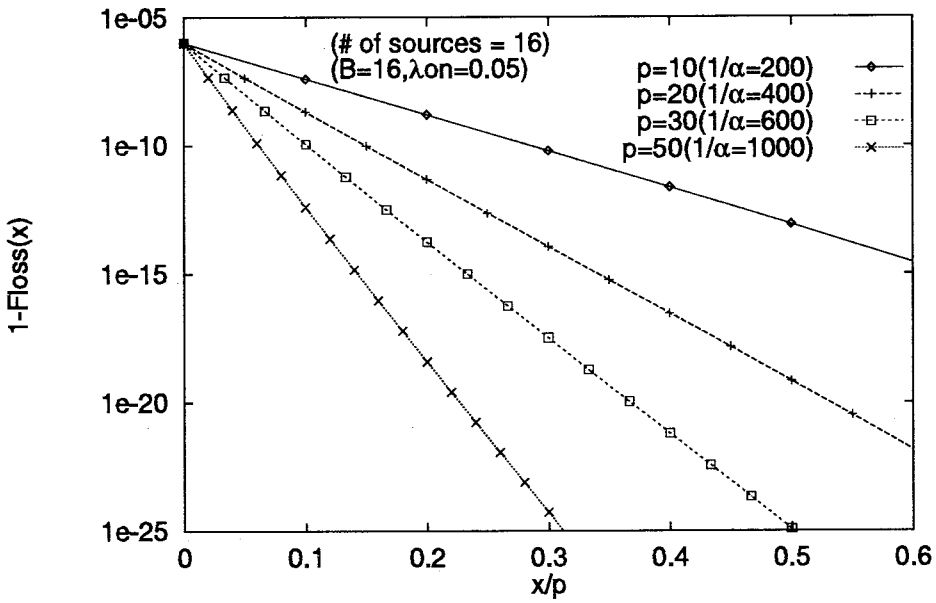
	$B_{loss}(r=0)=10^{-6}$		$B_{loss}(r=0)=10^{-9}$	
$1/\alpha$	r_{opt}	gain G	r_{opt}	gain G
400	6	2.25	7	3.06
2000	4	1.51	5	2.20
4000	4	1.40	4	2.09
40000	3	1.29	4	2.07

Table 5.4: Impact of p on r_{opt} and the gain G

	$B_{loss}(r=0)=10^{-6}$		$B_{loss}(r=0)=10^{-9}$	
p	r_{opt}/p	gain G	r_{opt}/p	gain G
10	0.1	0.26	0.15	0.52
20	0.3	2.25	0.35	3.06
30	0.3	3.97	0.33	5.06
50	0.28	6.72	0.28	8.04



(a)$B_{loss}(r)$ versus the number r of redundant cells



(b)survivor function $1 - F_{loss}(x)$ without using FEC

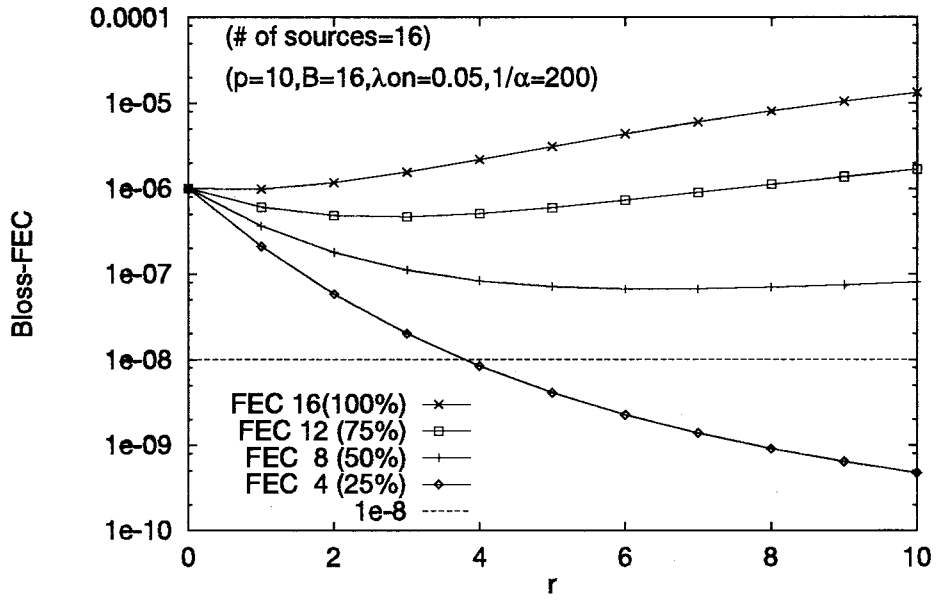
Figure 5.5: Impact of the block size p on loss characteristics

As shown in Fig. 5.5-(a), when p is set to 10 and the whole sources apply FEC, FEC is slightly effective. Nevertheless, when only 4 sources apply FEC, $B_{loss-FEC}$ decreases drastically, whereas $B_{loss-nonFEC}$ does not degrade so much. However, non-FEC sources may suffer from severe performance degradation when the number of FEC sources is large, as can be seen from Fig. 5.5-(b). Therefore, we conclude that FEC should be applied only to a small part of sources whose required QOS is comparatively strict in a way not to damage non-FEC sources seriously.

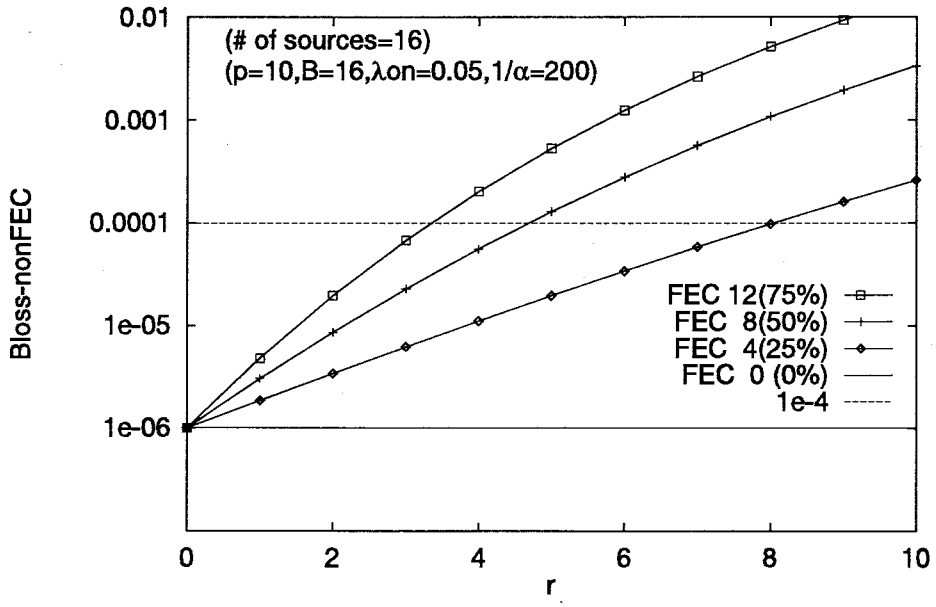
Suppose, for example, that there are two types of sources; in some sources, the block loss probability must be less than 10^{-8}, and, in other sources, that must be less than 10^{-4}. In order to meet these requirements, FEC is used in the former type of sources, say FEC sources. It is shown from Fig. 5.6 that the required QOS can not be guaranteed when there are $8, 12$ and 16 FEC sources. On the other hand, when only 4 sources apply FEC, the QOS can be achieved with r redundant cells such that $4 \leq r \leq 8$.

5.5 Conclusions

FEC has been considered as an effective error recovery scheme for applications with stringent delay requirements in high speed networks like ATM networks. This is due to its capability of recovering from loss without retransmissions. Furthermore, the development of the FEC encoder and decoder operating at 1 Gb/s makes FEC practical in high speed communication. In this Chapter, we carried out an exact analysis of the distribution of the number of cells lost within a block in ATM networks with correlated cell arrivals in order to investigate the characteristics of FEC. We have shown numerical results on its distribution and block loss probability in many cases. In conclusion, our analytical approach developed here is useful in evaluating the advantage of FEC and also its disadvantage to damage non-FEC traffic in a system composed of both FEC sources and non-FEC sources. By numerical results, FEC has shown to be very effective in reducing the block loss probability when the peak rate is not high. The optimum ratio of the number of redundant cells to that of data cells in a block depends on many parameters such as the peak arrival rate and burst length associated with transmitted traffic as well as the number of FEC sources. In addition, using a block of large size improves the block loss probability significantly. However, it is noted that a block of larger size can cause the receiver to wait longer until the last cell within a block comes up, and the possible block size thus depends on some requirement in terms of delay time in each application.



(a)In case of FEC sources: $B_{loss-FEC}$



(b)In case of non-FEC sources: $B_{loss-nonFEC}$

Figure 5.6: Impact of the number of FEC sources on block loss probability

5.A Appendix: Analysis of Steady State Probability

To simplify notations, we define $\boldsymbol{p}_m^{(d)}$ as the state probability vector at the mth slot:

$$\begin{aligned}\boldsymbol{p}_m^{(d)} &\triangleq [\boldsymbol{p}_{m,0}^{(d)}, \cdots, \boldsymbol{p}_{m,b}^{(d)}, \cdots, \boldsymbol{p}_{m,B}^{(d)}],\\ \boldsymbol{p}_{m,b}^{(d)} &\triangleq [\cdots, p_m^{(d)}(b, s_s, s_e), \cdots], \quad s_s = 0, 1,\ 1 \le d \le p+r,\ 1 \le s_e \le M,\end{aligned} \tag{5A-1}$$

where

$$p_m^{(d)}(b, s_s, s_e) \triangleq \Pr\{D_r(m) = b, S_s(m) = s_s, D(m) = d, S_e(m) = s_e\}, \tag{5A-2}$$

M denotes the number of phases in external sources, and B denotes the buffer size of the output buffer in the switch. Further $s_s = 0$ and 1 correspond to the OFF- and ON-states, respectively, in the IBP state of the tagged source. In what follows, we analyze the steady state probability of the Markov chain $\{D_r(m), S_s(m), D(m), S_e(m)\}$.

5.A.1 State Transition Probabilities of the Tagged Source

We assume that cell generations from the tagged source follow the transitions of the phase state of IBP. Define the transition probability matrix $\boldsymbol{T}_{n_m}^{(d)}$, whose elements are the probabilities of the transition with respect to the random variable S_s and of n_m cells being transmitted from the tagged source at the mth slot, given that $D = d$ at the mth slot. Namely, its (f, g) element $\boldsymbol{T}_{n_m(f,g)}^{(d)}$ is given by

$$\begin{aligned}\boldsymbol{T}_{0(k+1,l+1)}^{(d)} &\triangleq \Pr\{S_s(m+1) = l, n_m = 0 | S_s(m) = k, D(m) = d\}\\ &= \begin{cases} 1-\beta & \text{if } k = l = 0,\\ \alpha & \text{if } k = 1, l = 0,\\ (1-\lambda_{ON})\beta & \text{if } k = 0, l = 1,\\ (1-\lambda_{ON})(1-\alpha) & \text{if } k = l = 1,\end{cases}\end{aligned} \tag{5A-3}$$

$$\begin{aligned}\boldsymbol{T}_{1(k+1,l+1)}^{(d)} &\triangleq \Pr\{S_s(m+1) = l, n_m = 1 | S_s(m) = k, D(m) = d\}\\ &= \begin{cases} \lambda_{ON}\beta & \text{if } k = 0, l = 1,\\ \lambda_{ON}(1-\alpha) & \text{if } k = l = 1,\\ 0 & \text{otherwise.}\end{cases}\end{aligned} \tag{5A-4}$$

Each of the above matrices has the dimension 2×2.

Note that when a transition driven by $\boldsymbol{T}_0^{(d)}$ happens, $D(m+1) = d$, and when a transition driven by $\boldsymbol{T}_1^{(d)}$ happens, $D(m+1) = d+1$ $(d < p+r)$ or $D(m+1) = 1$ $(d = p+r)$. Furthermore, the matrices $\boldsymbol{T}_0^{(d)}$ and $\mathbf{T}_1^{(d)}$ are irrelevant to a specific value of d. Hereafter, for all d $(1 \le d \le p+r)$, we shall denote $\boldsymbol{T}_0^{(d)}$ and $\boldsymbol{T}_1^{(d)}$ by $\boldsymbol{T}_0$ and $\boldsymbol{T}_1$, respectively.

5.A.2 Transition Probability Matrix

We assume that the events on the output buffer occur in the following order.

1. At the beginning of a slot, the transmitter starts to send the cell in the head of line of the output buffer.

2. Then, the output buffer starts to accommodate cells transmitted from the tagged and/or external sources.

3. At the end of the slot, the transmitter finishes sending the cell.

From this assumption, when k cells are waiting in the output buffer and l cells are arriving there from the tagged and/or external sources, only $B-k$ cells among the l cells are eligible to enter the buffer and the remaining $l-(B-k)$ cells are dropped out.

Therefore focusing on the transition with respect to the queue length, we have the following relations:

$$\boldsymbol{p}^{(1)}_{m+1,b'} = \begin{cases} \boldsymbol{p}^{(p+r)}_{m,0}\boldsymbol{A}_{1,b'} + \boldsymbol{p}^{(1)}_{m,0}\boldsymbol{A}_{0,b'}, & \text{if } b=0 \text{ and } b' \le B-1, \\ \boldsymbol{p}^{(p+r)}_{m,0}\overline{\boldsymbol{A}_{1,B}} + \boldsymbol{p}^{(1)}_{m,0}\overline{\boldsymbol{A}_{0,B}}, & \text{if } b=0 \text{ and } b'=B, \\ \boldsymbol{p}^{(p+r)}_{m,b}\boldsymbol{A}_{1,b'-b+1} + \boldsymbol{p}^{(1)}_{m,b}\boldsymbol{A}_{0,b'-b+1}, & \text{if } 1 \le b \le B \text{ and } b-1 \le b' \le B-2, \\ \boldsymbol{p}^{(p+r)}_{m,b}\overline{\boldsymbol{A}_{1,b'-b+1}} + \boldsymbol{p}^{(1)}_{m,b}\overline{\boldsymbol{A}_{0,b'-b+1}}, & \text{if } 1 \le b \le B \text{ and } b'=B-1, \end{cases} \tag{5A-5}$$

and for $2 \le d \le p+r$

$$\boldsymbol{p}^{(d)}_{m+1,b'} = \begin{cases} \boldsymbol{p}^{(d-1)}_{m,0}\boldsymbol{A}_{1,b'} + \boldsymbol{p}^{(d)}_{m,0}\boldsymbol{A}_{0,b'}, & \text{if } b=0 \text{ and } b' \le B-1, \\ \boldsymbol{p}^{(d-1)}_{m,0}\overline{\boldsymbol{A}_{1,B}} + \boldsymbol{p}^{(d)}_{m,0}\overline{\boldsymbol{A}_{0,B}}, & \text{if } b=0 \text{ and } b'=B, \\ \boldsymbol{p}^{(d-1)}_{m,b}\boldsymbol{A}_{1,b'-b+1} + \boldsymbol{p}^{(d)}_{m,b}\boldsymbol{A}_{0,b'-b+1}, & \text{if } 1 \le b \le B \text{ and } b-1 \le b' \le B-2, \\ \boldsymbol{p}^{(d-1)}_{m,b}\overline{\boldsymbol{A}_{1,b'-b+1}} + \boldsymbol{p}^{(d)}_{m,b}\overline{\boldsymbol{A}_{0,b'-b+1}}, & \text{if } 1 \le b \le B \text{ and } b'=B-1, \end{cases} \tag{5A-6}$$

where

$$\begin{aligned} \boldsymbol{A}_{0,i} &= \boldsymbol{T}_0 \otimes \boldsymbol{C}_i, \quad 0 \le i \le B-1, \\ \boldsymbol{A}_{1,i} &= \boldsymbol{T}_1 \otimes \boldsymbol{C}_{i-1}, \quad 1 \le i \le B-1, \\ \boldsymbol{A}_{1,0} &= \boldsymbol{0}[2M], \\ \overline{\boldsymbol{A}_{0,i}} &= \boldsymbol{T}_0 \otimes \overline{\boldsymbol{C}_i}, \quad 0 \le i \le B, \\ \overline{\boldsymbol{A}_{1,i}} &= \boldsymbol{T}_1 \otimes \overline{\boldsymbol{C}_{i-1}}, \quad 1 \le i \le B, \\ \overline{\boldsymbol{A}_{1,0}} &= \boldsymbol{T}_1 \otimes \overline{\boldsymbol{C}_0}, \end{aligned} \tag{5A-7}$$

$\mathbf{0}[m]$ denotes a zero matrix of dimensions $m \times m$, and $\otimes$ denotes the Kronecker product. Note that $\boldsymbol{C}_x$ and $\overline{\boldsymbol{C}_x}$ are given in Eqs. (5.2) and (5.10), respectively. The matrices $\boldsymbol{A}_{i,j}$ and $\overline{\boldsymbol{A}_{i,j}}$ have the dimension $q \times q$, where $q \triangleq 2M$.

Note here that the structure of the transitions to $\{D_r(m+1) = b', S_s(m+1), D(m+1) = d, S_e(m+1)\}$ (whose probabilities are represented by $\boldsymbol{p}^{(d)}_{m+1,b'}$) is independent of a specific value of d $(1 \le d \le p+r)$. This implies that the Markov chain $\{D_r(m), S_s(m), D(m), S_e(m)\}$ is exactly lumpable with respect to a random variable D [SCHW 84]. Namely, the steady state probabilities of $\{D_r, S_s, D = d, S_e\}$ $(1 \le d \le p+r)$ are all the same, i.e.,

$$\begin{aligned}&\lim_{m\to\infty} \Pr\{D_r(m) = b, S_s(m) = s_s, D(m) = d, S_e(m) = s_e\} \\ &\quad = \lim_{m\to\infty} \Pr\{D_r(m) = b, S_s(m) = s_s, S_e(m) = s_e\}/(p+r). \end{aligned} \tag{5A-8}$$

In what follows, we shall derive the steady state probability of the aggregated Markov chain $\{D_r(m), S_s(m), S_e(m)\}$, and using Eq. (5A-8), we obtain the steady state probability of the original Markov chain.

Let $p_m(b, s_s, s_e)$ denote the probability of $\{D_r(m) = b, S_s(m) = s_s, S_e(m) = s_e\}$ and $\boldsymbol{p}_{m,b}$ denote $[\cdots, p_m(b, s_s, s_e), \cdots]$, where $0 \le b \le B$, $s_s = 0, 1$ and $1 \le s_e \le M$. By summing up both sides of eqs.(5A-5) and (5A-6) for all d $(1 \le d \le p+r)$, we have

$$\boldsymbol{p}_{m+1,b'} = \begin{cases} \boldsymbol{p}_{m,0}\boldsymbol{A}_{b'} + \sum_{b=1}^{b'+1} \boldsymbol{p}_{m,b}\boldsymbol{A}_{b'-b+1} & \text{if} \quad 0 \le b' \le B-2, \\ \boldsymbol{p}_{m,0}\boldsymbol{A}_{B-1} + \sum_{b=1}^{B} \boldsymbol{p}_{m,b}\overline{\boldsymbol{A}_{B-b}} & \text{if} \quad b' = B-1, \\ \boldsymbol{p}_{m,0}\overline{\boldsymbol{A}_B} & \text{if} \quad b' = B. \end{cases} \tag{5A-9}$$

where for $0 \le b \le B$

$$\boldsymbol{A}_b = \boldsymbol{A}_{0,b} + \boldsymbol{A}_{1,b}, \quad \overline{\boldsymbol{A}_b} = \overline{\boldsymbol{A}_{0,b}} + \overline{\boldsymbol{A}_{1,b}}. \tag{5A-10}$$

Therefore, it follows from Eq. (5A-9) that the transition probability matrix $\boldsymbol{Q}$ related to the state probabilities $\boldsymbol{p}_{m,b}$ takes the following form.

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_{B-3} & \boldsymbol{A}_{B-2} & \boldsymbol{A}_{B-1} & \overline{\boldsymbol{A}_B} \\ \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_{B-3} & \boldsymbol{A}_{B-2} & \overline{\boldsymbol{A}_{B-1}} & \mathbf{0}[q] \\ \mathbf{0}[q] & \boldsymbol{A}_0 & \cdots & \boldsymbol{A}_{B-2} & \boldsymbol{A}_{B-1} & \overline{\boldsymbol{A}_{B-2}} & \mathbf{0}[q] \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \ddots & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \overline{\boldsymbol{A}_2} & \mathbf{0}[q] \\ \vdots & \cdots & \cdots & \ddots & \boldsymbol{A}_0 & \overline{\boldsymbol{A}_1} & \mathbf{0}[q] \\ \mathbf{0}[q] & \cdots & \cdots & \cdots & \mathbf{0}[q] & \overline{\boldsymbol{A}_0} & \mathbf{0}[q] \end{pmatrix}. \tag{5A-11}$$

Note that the dimension of $\boldsymbol{Q}$ is equal to $q(B+1) \times q(B+1)$.

We define the steady state probability vector as follows.

$$\begin{aligned} \boldsymbol{x}^* &\triangleq \lim_{m\to\infty} \boldsymbol{p}_m = [\boldsymbol{x}_0^*, \cdots, \boldsymbol{x}_b^*, \cdots, \boldsymbol{x}_B^*], \\ \boldsymbol{x}_b^* &\triangleq \lim_{m\to\infty} \boldsymbol{p}_{m,b}. \end{aligned} \tag{5A-12}$$

Consequently, we obtain the steady state probability vector $\boldsymbol{x}^*$ by solving the following equation:

$$\boldsymbol{x}^* = \boldsymbol{x}^* \boldsymbol{Q}, \quad \boldsymbol{x}^* \boldsymbol{e}[q(B+1)] = 1, \tag{5A-13}$$

where $\boldsymbol{e}$[m] represents a column vector of dimension m, whose all elements are equal to 1.

It may be difficult to directly solve Eq. (5A-13) since the number of states grows explosively as the buffer capacity B becomes large. The algorithm proposed in [TAKI 93] enables us to calculate $\boldsymbol{x}^*$ independently of the buffer size. Furthermore, using Eq. (5A-8), we obtain the steady state probabilities $x(b, s_s, d, s_e)$ in terms of $\boldsymbol{x}^*$.

Chapter 6

Concluding Remarks

6.1 Summary of This Dissertation

In the dissertation, I have considered both congestion control and error control in ATM networks. In particular, I have focused on reactive feedback control and selective cell discard scheme for ABR traffic as congestion control, and Forward Error Correction (FEC) as error control. In order to quantitatively evaluate those control schemes, some analytical approaches based on the queueing theory have been proposed.

In the remainder of this section, I would like to just review this dissertation based on the chapter organization.

In Chapter 2, I dealt with the **reactive feedback control** based on **BECN** (Backward - Explicit Congestion Notification). Our analytical model consists of several sources and the server with some propagation delay between them. The model is exact in cases without propagation delay, and is approximate in case of non-zero propagation delay. By using the model, I carried out an approximate analysis of two types of reactive control algorithms: the binary control (BC1) and that with cell push-out scheme at each source (BC2). The binary control is that sources choke/relieve their cell transmission when the queue length of the server buffer goes over/falls below an adequate threshold, H. Numerical results agreed with simulation ones very well, and BC2 provides better performance in terms of cell loss probability than BC1 while the average waiting time is not degraded so much. Further, BC2 is better than BC1 in that it is easier to determine the optimum value of H, which minimizes the overall cell loss probability under both BC1 and BC2. We also found that the statistical gain can be increased by the reactive congestion control algorithms over a wide range of the total rate, the burstiness and the burst-length of offered traffic.

In Chapter 3, I evaluated the performance of **link-by-link backpressure congestion control** by means of analytical approach, developed in Chapter 2, in the two-stage queueing

model consisting of the congested node and its upstream nodes. Numerical results have shown that the plain backpressure (BP1) adversely influences the cells which do not go through the congested server. The backpressure combined with the cell push-out scheme (BP2), proposed in Chapter 2, can overcome this drawback and further improves the loss performance in terms of cells transmitted to the congested server. In addition, BP2 is considered to be effective in ATM WANs because of its ability to efficiently relieve the HOL blocking due to BP1.

In Chapter 4, I carried out an exact analysis of the packet loss probability in the system applying **selective cell discard** scheme, such as TD, EPD based upon a fixed threshold (EPD1) of the buffer on the ATM switch, and that based upon variable ones (EPD2), given that the packet length follows a geometric distribution, and that sources are homogeneous in terms of traffic characteristics without the retransmission of lost packets. From numerical results, the packet loss probability becomes smaller as the threshold is set to larger value in EPD1 and as the acceptance coefficient δ is set to smaller value in EPD2. Contrary to the previous works [ROMA 95], however, my study shows that both EPD1 and EPD2 can not provide better performance than TD. We expect that this fact comes from the assumption of the geometrically distributed packet length. In this case, the number of cells which will arrive from an accepted packet at the switch has the same distribution as the number of cells in a newly arriving packet. Thus, I concluded that the effectiveness of EPDs strongly depends on the packet length distribution.

In Chapter 5, I developed an exact analysis of the distribution of lost cells within a block in order to investigate the characteristics of **FEC combined with cell interleaving**. By numerical results, FEC has shown to be very effective in reducing the block loss probability when the peak rate of each source is not high. The optimum ratio, which makes the block loss probability minimized, of the number of redundant cells to that of data cells in a block depends on many parameters such as the peak arrival rate and the burst length as well as the number of FEC sources. In addition, we found that using a block of larger size improves the block loss probability significantly. However, it is noted that a block of larger size can cause the receiver to wait longer until the last cell within a block comes up, and the possible block size thus depends on some requirement in terms of delay time in each application.

6.2 Issues for Future Research

In the analysis of reactive congestion control proposed in Chapter 2, I treated only the binary control although it enables us to evaluate the case where sources gradually increase/decrease those transmission rate according to the reduced/raised queue length of the congested server, namely the analysis would be able to investigate such control mechanism like as EPRCA (studied, e.g., in [OHSA 95]).

On link-by-link backpressure congestion control, I have analyzed the limited system consisting of nodes in two stages, which represents the relation between the congested node and its upstream nodes. In the system, I could not clarify how the congested state is spreading over networks. Thus, the investigation on the property of the backpressure in a multistage network is of practical interest, and remains as future work. Further, by considering such system, it is expected that the congested state would finally reach source end system in a network if the overload duration gets long. Thus, end-to-end feedback control is still required, although link-by-link control is performed. From this point of view, in [RAMA 95], the integration of end-to-end rate based feedback control and link-by-link credit based control scheme has been proposed for ABR flow control. It is also interesting topic.

Concerning the analysis of some selective cell discard schemes, I assumed the system where lost packets are not retransmitted by any sources. From numerical results obtained in Chapter 4, EPD more reduces the queue length of the buffer in the intermediate ATM switch than the case of TD whereas EPD achieves less effectiveness than by TD. It implies that, in the system with lost packet retransmission, retransmitted packet can easily stored in the buffer in case of EPD, and the performance of EPD would be superior to that of TD as already mentioned in [ROMA 95]. In order to exactly show this assumption, we may have to take into account the packet loss performance with lost packet retransmission or throughput performance in the system. We would need to clarify the behavior of higher layer protocol such as transport one over ATM networks.

The performance of FEC for CBR traffic was shown in Chapter 5, in which the number of cells in a unit block is set to the larger value, the block loss probability is considerably improved. However, the block reconstructing time at the receiver becomes larger. Therefore, we should investigate the relationship between the loss performance and delay one. Furthermore, FEC may be applied to VBR audio and video as mentioned in [ITUI 363]. In this case, if the block size is fixed, then the block constructing time at sources is relatively large when cells generated at lower rate, thereby causing the unacceptable delay for VBR stream on ATM connection. One of ideas for overcoming this problem is that the block

size is changed according to cell generation rate of VBR stream, and it would be for further study.

Bibliography

[ABOU 92] O. Aboul-Magd and H. Gilbert, "Incorporating Congestion Feedback in B-ISDN Traffic Management Strategy," *Proc. ISS'92*, paper A5.3, Oct. 1992.

[ANTH 90] A. J. McAuley, "Reliable Broadband Communication Using a Burst Erasure Correcting Code," *Proc. ACM SIGCOMM '90*, pp. 297–306, August 1990.

[ARMI 95] G. J. Armitage and K. M. Adams, "Packet Reassembly During Cell Loss," *IEEE Network*, vol. 7, no. 5, pp. 26–34, Sep. 1993.

[ARMB 90] H. Armbrüster and H. J. Rothamel, "Breitbandanwendungen unddienste - Qualitative und quantitative Anforderungen an künftige Netze," *Nachrichtentechnische Zeitschrift*, vol. 43, no. 3, pp. 150–159, March 1990.

[BIER 92] E. W. Biersack et al., "Gigabit Networking Research at Bellcore," *IEEE Network*, vol. 6, no. 2, pp. 42–48, March 1992.

[BLON 92] C. Blondia and O. Casals, "Statistical Multiplexing of VBR Sources: A Matrix Analytic Approach," *Performance Evaluation*, vol. 16, pp. 5–20, 1992.

[BONO 95] F. Bonomi and K. W. Fendick, "The Rate-Based Flow Control Framework for the Available Bit Rate ATM Service," *IEEE Network*, vol. 9, no. 2, pp. 25–39, 1995.

[BURT 72] H. O. Burton and D. Sullivan, "Errors and Error Control," *Proceedings of the IEEE*, pp.1293–1301, November 1972.

[CHAN 94] C. G. Chang and H. H. Tan, "Queueing Analysis of Explicit Policy Assignment Push-Out Buffer Sharing Schemes for ATM Networks," *IEEE Infocom'94*, pp. 500–509, June 1994.

[CIDO 93] I. Cidon, A. Khamisy and M. Sidi, "Analysis of Packet Loss Processes in High-Speed Networks," *IEEE Trans. Information Theory*, vol 39, no. 1, pp.98–108, Jan. 1993.

[ERNS 93] E. W. Biersack, "Performance Evaluation of Forward Error Correction in an ATM Environment," *IEEE J. Select. Areas in Commun.*, vol. 11, no. 4, pp. 631–640, May 1993.

[GERS 91] A. Gersht and K. J. Lee, "A Congestion Control Framework for ATM Networks," *IEEE J. Select. Areas Commun.*, vol.9, no.2, pp.1119–1130, 1991.

[GILB 91] H. Gilbert, O. Aboul-Magd, and V. Phung, "Developing a Cohesive Traffic Management Strategy for ATM Networks," *IEEE Network*, vol.29, no.10, pp.36–45, October 1991.

[HABI 91] I. W. Habib and T. N. Saadawi, "Controlling Flow and Avoiding Congestion in Broadband Networks," *IEEE Commun. Magazine*, vol.29, no.10, pp.46–53, October 1991.

[HÄND 95] R. Händel, M. N. Huber and S. schröder, *ATM Networks (Second edition): Concepts, Protocols, Applications*, Addition-Wesley publishing company, 1995.

[HEBU 90] G. Hebuterne and A. Gravey, "A Space Priority Queueing Mechanism for Multiplexing ATM Channels," *Computer Networks and ISDN System*, No. 20, pp.37–43, 1990.

[HEFF 86] H. Heffes and D. M. Lucantoni, "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Multiplexer Performance," *IEEE J.Selected Areas in Commun.*, vol. SAC-4, no. 6, pp.856–868, September 1986.

[IKE 94] C. Ikeda and H. Suzuki, "Adaptive Congestion Control Schemes for ATM-LANs," *IEEE Infocom'94*, pp. 829–838, June 1994.

[ITUC 84] ITU-T: COM XVIII-228-E, Geneva, March, 1984.

[ITUI 113] ITU-T: Recommendation I.113. "Vocabulary of Terms for Broadband Aspects of ISDN", Rev. 1, Geneva, 1991.

[ITUI 121] ITU-T: Recommendation I.121. "Broadband Aspects of ISDN", Rev. 1, Geneva, 1991.

[ITUI 321] ITU-T: Recommendation I.321. "B-ISDN Protocol Reference Model and its Application", Geneva, 1993.

[ITUI 362] ITU-T: Recommendation I.362. "B-ISDN ATM Adaptation Layer (AAL) Functional Description", Rev. 1, Geneva, 1993.

[ITUI 363] ITU-T: Recommendation I.363. "B-ISDN ATM Adaptation Layer Specification", Rev. 1, Geneva, 1993.

[ITUI 371] ITU-T: Recommendation I.371. "Traffic Control and Congestion Control in B-ISDN", Rev. 1, Geneva, 1993.

[ITUI 432] ITU-T: Recommendation I.432. "B-ISDN USer-Network Interface - Physical Layer Specification", Rev. 1, Geneva, 1993.

[JAIN 92] R. Jain, "Myth About Congestion Management in High-Speed Networks, " *Information Networks and Data Communication, IV, Elsevier Science Publishers B. V. (M. Tienari and D. Khakhar Editors)*, pp. 55–70, 1992.

[KAMA 95] A. E. Kamal, "A Performance Study on the Effect of the End-of-Message Indicator in AAL Type 5," in *Proc. IEEE Infocom'95*, pp. 1264–1272, Apr. 1995.

[KANG 93] C. G. Kang and H. H. Tan, "Queueing Analysis of Explicit Priority Assignment Partial Buffer Sharing Schemes for ATM Networks," *Proc. IEEE Infocom'93*, pp.810–819, Mar.-Apr. 1993.

[KAWA 94] K. Kawahara, K. Kumazoe, T. Takine, Y. Oie, "Forward Error Correction in ATM Networks: An Analysis of Cell Loss Distribution in a Block," *Proc. IEEE Infocom '94*, pp.1150–1159, Toronto, June 1994.

[KAWA 95a] K. Kawahara, Y. Oie, M. Murata, H. Miyahara, "Performance Analysis of Reactive Congestion Control for ATM Networks," *IEEE J. Select. Areas in Commun.*, vol. 13, no. 4, pp.651–661, May 1995.

[KAWA 95b] K. Kawahara, Y. Oie, M. Murata, H. Miyahara, "Performance Analysis of Backpressure Congestion Control: Preliminary Case," *Proc. IEEE GLOBECOM '95*, pp.304–309, Singapore, Nov. 1995.

[KAWA 95c] K. Kawahara, Y. Oie, M. Murata, H. Miyahara, "Performance Analysis of Backpressure Congestion Control: Preliminary Case," *The Trans. of IEICE*, vol. J78-B-I, no. 12, pp.837–845, Dec. 1995. (in Japanese)

[KAWA 96] K. Kawahara, K. Kitajima, T. Takine, Y. Oie, "Performance Evaluation of Selective Cell Discard Schemes in ATM Networks," *IEEE Infocom'96*, pp.1054–1061, San Francisco, Mar. 1996.

[KOLA 94] A. Kolarov and G. Ramamurthy, "Comparison of Congestion Control Schemes for ABR Service in ATM Local Area Networks," *IEEE Globecom'94*, pp. 913–918, 1994.

[KRÖN 90] H. Kröner, "Comparative Performance Study of Space Priority Mechanisms for ATM Networks," *Proc. Infocom '90*, pp. 1136–1143, San Francisco, 1990.

[KUNG 95] H. T. Kung and R. Morris, "Credit-Based Flow Control for ATM Networks," *IEEE Network*, vol. 9, no. 2, pp. 40–48, 1995.

[KUMA 96] K. Kumazoe, K. Kawahara, T. Takine and Y. Oie, "Analysis of Packet Loss in Transport Layer over ATM Networks," *submitted for publication*, Mar. 1996.

[LEME 94] M. Lemercier and G. Pujolle, "A Performance Study of a New Congestion Management Scheme in ATM Bloadband Networks :The Multiple Push-out," *ITC 14, (J. Labetoulle and J. W. Roberts Editors) Elsevier Science B. V.*, pp. 1261–1270, 1994.

[NEWM 93] P. Newman, "Backward Explicit Congestion Notification for ATM Local Area Networks," *Proc. IEEE Globecom'93*, pp.719–723, Nov.-Dec. 1993.

[NORI 94] K. Noritake, S. Ushijima and M. Hirano, "A Study on Selective Cell Discard Control in ATM-CL Network," *IEICE Tech. Report*, SSE94-77, Jun. 1994 (in Japanese).

[OHSA 95] H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda and H. Miyahara, "Rate-Based Congestion Control for ATM Networks," *ACM SIGCOMM Computer Communication Review*, vol. 25, pp. 60–72, April 1995.

[PRYC 95] Martin De Prycker, *Asynchronous Transfer Mode (Third edition): Solution for Broadband ISDN*, Prentice Hall, 1995.

[RAMA 95] K. K. Ramakrishnan and P. Newman, "Integration of Rate and Credit Schemes for ATM Flow Control," *IEEE Network*, vol. 9, no. 2, pp. 49–56, 1995.

[ROMA 95] A. Romanow and S. Floyd, "Dynamics of TCP Traffic over ATM Networks," *IEEE J. Select. Areas Commun.*, vol. 13, no. 4, pp. 633–641, May 1995.

[SCHA 90] N. Schacham and P. McKenney, "Packet Recovery in High-Speed Networks Using Coding and Buffer Management," *Proc. IEEE INFOCOM'90*, pp.124–131, 1990.

[SCHW 84] P.J. Schweizer, *Aggregation Methods for Large Markov Chains,* Mathematical Computer Performance and Reliability, G. Iazeolla et al. (eds.), pp.275–286, North-Holland, Amsterdam, 1984.

[SIU 95] K. Y. Siu and R. Jain, "A Brief Overview of ATM: Protocol Layers, LAN Emulation, adn Traffic Management," *ACM SIGCOMM Computer Communication Review*, vol. 25, pp. 6–20, April 1995.

[STEW 89] W. J. Stewart, "Recursive Procedures for the Numerical Solution of Markov Chains," *Queueing Networks with Blocking (H.G.Perros and T.Altiok Editors), Elsevier Science Publishers B.V.*, pp.229-247, 1989

[TAKI 93] T. Takine, T. Suda and T. Hasegawa, "Cell Loss and Output Process Analyses of a Finite-Buffer Discrete-Time ATM Queueing System with Correlated Arrivals," *Proc. IEEE INFOCOM'93*, pp.1259–1269, Mar.-Apr., 1993.

[TAKI 94] T. Takine, B. Sengupta and T. Hasegawa, "An Analysis of a Discrete-Time Queue for Broadband ISDN with Priorities among Traffic Classes," *IEEE Transactions on Communications*, vol. 42, nos. 2–4, pp.1837-1845, February/March/April 1994.

[VICK 94] B. J. Vickers and T. Suda, "Connection Service for Public ATM Networks," *IEEE Commun. Mag.*, vol. 32, no. 8, pp. 34–42, Aug. 1994.

[WANG 91] Y. T. Wang and B. Sengupta, "Performance Analysis of a Feedback Congestion Control Policy under Non-Negligible Propagation Delay," *Proc. Sigcomm'91*, pp.149–157, 1991.

[YANG 95] C. Q. Yang and A. V. S. Reddy, "A Taxonomy for Congestion Control Algorithms in Packet Switching Networks," *IEEE Network*, vol. 9, no. 4, pp. 34–45, 1995.

[ZANG 92] L. Zang, "Statistics of Cell Loss and Its Application for Forward Error Recovery in ATM Network," *Proc. IEEE ICC'92*, pp.694–698, 1992.