Interoperability of EFCI and ER Switches for ABR Services in ATM Networks

Yuan-Cheng Lai Ying-Dar Lin
National Chiao Tung University

Abstract-

With the advances in switching technologies, ER switches are becoming popular since they perform better than EFCI switches. In the transitional period, the EFCI and ER switches may coexist in the same ATM network. Hence the efficiency of various ER schemes should be re-evaluated in the mixed EFCI-ER environment, not only in the homogeneous ER environment. Also some important implications are observed in the heterogeneous environments. Because the location of an ER switch in the topology is critical to its performance, some placement rules, which describe how to place the ER switches for the network operators, are developed in order to achieve a better performance of the network. In summary, the results presented here should serve as an operational reference for the service providers.

1 Introduction

ATM (Asynchronous transfer mode) is the most promising transfer technology for implementing B-ISDN. It supports applications with distinct QoS requirements such as delay, jitter, and cell loss and distinct demands such as bandwidth or throughput, so it can enable multimedia communication over the networked environment. However, in order to support old applications or some applications that do not choose to use performance-guaranteed services, ATM network also provide best-effort services such as LAN emulation and IP-over-ATM. To provide these services for a wide variety of applications, ATM Forum has defined a family of service categories including CBR (Constant Bit Rate) service, rt-VBR (Real-time variable bit rate), nrt-VBR (Non-real-time VBR), UBR (Unspecified bit Rate) services, and ABR (Available Bit Rate) service [1].

As we know, the ABR service is suited for a wide variety of applications since it can support the data traffic economically [2]. However, when these applications, such as IP-over-

ATM, transmit data packets in ATM networks, they segment each packet into ATM cells. The loss of any cell causes the retransmission of the entire packet. Under the condition of congestion, the collapse of throughput happens since too many packet retransmission resulted from cell loss [3]. Hence, it is necessary to adopt some congestion methods to decrease the cell loss and increase the throughput.

Recently some feedback flow control schemes have been presented in order to provide the ABR services [4-11]. Two strategies have been proposed: credit-based [4,5] and rate-based [2,6-11]. The credit-based scheme is a link-by-link window flow control scheme. Each link in the network independently runs the flow control mechanism. The rate-based scheme use feedback signal from the network to control the rate at which each source transmit cells into the network. In late 1994, the ATM Forum has voted for rate-based control as the ABR service.

In the Traffic Management Specification TM 4.0, the detailed and concise operations of the rate-based flow control are described. The behavior of traffic sources and destinations is clearly defined in order to provide the base line for the vendors to follow. However, the methods that the switches should use to control the source rate are up to the vendors. Currently most ATM vendors have already provided ATM switches equipped with Explicit Forward Congestion Indication (EFCI) functions recommend by the ITU-T [12]. These switches, which use EFCI marking, are called the first-generation switches [13]. With the advances in switching technologies, the second-generation switches, which own the Explicit Rate (ER) setting capacity, is becoming popular.

In the transitional period from first to second-generation switches, the interoperating EFCI and ER switches becomes unavoidable. Because of the differences in their basic operations, whether they can cooperate efficiently in the same VC path is questionable. In this article, the performance of interoperating various ER schemes with EFCI is presented and compared. Although we know some properties of these ER schemes from many papers [2,6-11], the knowledge is obtained from the environment that all switches use the same flow control algorithm. However, from our simulation, the behavior of these ER schemes in the mixed EFCI-ER environment is sometimes different with that in the pure ER switches. Hence these ER schemes must be re-evaluated.

In the mixed EFCI-ER environment, the location of ER is a critical issue for performance. When a customer buys a switch equipped with ER setting, which EFCI switch should be replaced first to obtain the best performance? We performed some simulations with various network configurations. The simulation results provide important implications on which EFCI switches to be replaced. Therefore, from our results, the network managers can have better insight to help them purchasing proper ER switches to fit their needs. Also when they are about to install their ER switches, they can choose the appropriate EFCI switches to be replaced.

The remainder of this article is organized as follows. In the second section, we describe the operations of EFCI switches and various ER switches. Simulation environment are presented in the third section. In the forth section, we present and discuss simulation results obtained in the homogeneous environment and in the heterogeneous EFCI-ER environment. Some guidelines to place ER switches are given in the fifth section. The final section points out some future works.

2 ABR Flow Control

First we briefly introduce the basic operation of the closed-loop rate-based control mechanism [1]. When a virtual circuit (VC) is established, the source end system (SES) sends the cells at the allowed cell rate (ACR) which is set as initial cell rate (ICR). In order to probe the congestion status of the network, the SES sends a forward Resource Management (RM) cell (Table 1) every Nrm-1 data cells. Each switch may set certain fields of the RM cell to indicate its own congestion status, or the bandwidth the VC source should use. The destination end system (DES) returns the forward RM cell as a backward RM cell to the SES. According to the received backward RM cell, the SES adjusts its allowed cell rate,

Field	Description	Octet(s)	Bit(s)
Header	ATM Header	1-5	all
ID	Protocol Identifier	6	all
DIR	Message type: Direction	7	8
BIN	Message type: BECN Cell	7	7
CI	Message type: Congestion Indication	7	6
NI	Message type: No Increase	7	5
RA	Message type: Request/Acknowledge	7	4
Reserved	Message type: Reserved	7	3-1
ER	Explicit Cell Rate	8-9	all
CCR	Current Cell Rate	10-11	all
MCR	Minimum Cell Rate	12-13	all
QL	Queue Length Protocol Identifier	14-17	all
SN	Sequence Number	18-21	all
Reserved	Reserved	22-51	all
Reserved	Reserved	52	8-3
CRC-10	CRC-10	52	2-1
CRC-10	CRC-10	53	all

Table 1: Fields and their positions in RM-cells.

which is bounded between Peak cell rate (PCR) and Minimum cell rate (MCR).

The RM cell contains a 1-bit congestion indication (CI) which is set to zero, and an explicit rate (ER) field which is set to PCR initially by the SES. When the SES receives a backward RM cell, it modifies its ACR using additive increase and multiplicative decrease. The new ACR is computed as follows, depending on CI and ER fields in RM cells:

$$\text{ACR}=\max(\min(\text{ACR} + \text{RIF} \cdot \text{PCR},\text{ER}),\text{MCR}), \text{ if CI}=0,$$

$$\text{ACR}=\max(\min(\text{ACR} \cdot (1\text{-RDF}),\text{ER}),\text{MCR}), \text{ if CI}=1,$$

where RIF is the rate increase factor and RDF is the rate decrease factor.

According to the way of congestion monitoring and feedback mechanism, various switch mechanisms can be classified into two types. One is the Explicit Forward Congestion Indication (EFCI) switch, the other is the Explicit Rate (ER) switch. The components of the rate-based flow control are shown in Figure 1.

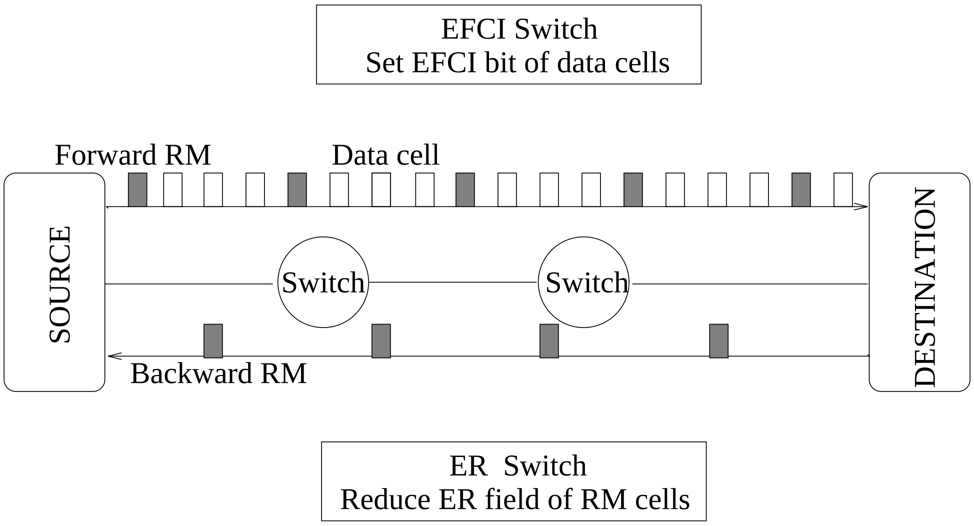


Figure 1: ABR rate-based flow control

2.1 EFCI scheme

In this scheme [14], when congestion occurs, the switch sets the EFCI bit to one (EFCI=1) in the header of each passing data cells. The DES, if a cell with EFCI=1 has been received, marks the CI bit (CI=1) to indicate congestion in each backward RM cells.

In most cases, the queue length is used to decide whether congestion occurs or not. As the queue length exceeds a threshold, denoted by Q_t, congestion is claimed. When the queue length falls below the threshold, congestion is relieved.

2.2 Explicit Rate Feedback Schemes

In the ER schemes, the switch computes the fair share of bandwidth that a VC can be supported with, determines the load and determines the actual explicit rate. When each RM cell passes, the switch sets the ER field to the determined explicit rate. Note each switch is not allowed to increase the ER field. Thus a source shall receive the minimum allowed cell rate of all the switches along the path. Examples of ER switch mechanisms are the EPRCA, ERICA, CAPC, Charny Max-Min, Tsang Max-Min schemes [15-19].

• Enhanced Proportional Control Algorithm (EPRCA) [15]

Each switch maintains a mean allowed cell rate (MACR) using a running exponential weighted average. When a switch receives a forward RM cell during the congestion period, MACR is updated as

$$\text{MACR}=(1\text{-}\alpha)\text{MACR} + \alpha\text{CCR},$$

where α is the exponential averaging factor generally set to be 1/16 and CCR is the current cell rate of the VC recorded in the RM cell. The fair bandwidth share is computed as a fraction of the MACR:

$$\text{Fair share} = \text{DPF} \cdot \text{MACR},$$

where DPF is a switch down pressure factor set close to but below 1. When a switch receives a backward RM cell, it reduces the ER field to the fair share if its queue length is larger than Q_t.

- Explicit Rate Indication for Congestion Avoidance (ERICA) [16]

The ERICA uses a load factor, z, to indicate the overload or underload state of the switch. The load factor is defined as

$$\text{z} = \frac{\text{Input rate}}{\text{Target rate}}.$$

The input rate is measured over a fixed averaging interval and the target rate is usually set slightly below the link bandwidth. Because the goal of this algorithm is to maintain the load factor close to one, the sources ought to change their current sending rates inversely proportional to the calculated load factor. The VC share and fair share are as follows:

$$\text{VCshare} = \frac{\text{CCR}}{\text{z}},$$

$$\text{Fair share} = \frac{\text{Target rate}}{\text{Number of active connections}}.$$

A switch updates the ER field, in the backward RM cell it received, to be the maximum value of the fair share and VCshare.

- Congestion Avoidance using Proportional Control (CAPC) [17]

Again as in the ERICA scheme, the switches set a target utilization slightly below 1 and compute the load factor. The main difference lies in the way the fair share is computed, which depends on whether z<1 or z>1. Thus, we have

$$\text{Fair share} = \text{Fair share} \cdot \min(\text{ERU}, 1+(1\text{-}z) \cdot \text{Rup}), \text{ if } z<1, \text{ and}$$

$$\text{Fair share} = \text{Fair share} \cdot \max(\text{ERF}, 1\text{-}(z\text{-}1) \cdot \text{Rdn}), \text{ if } z>1,$$

where Rup is a slope parameter between 0.025 and 0.1, and Rdn is between 0.2 and 0.8. ERU and ERF determine the maximum allowed increase and minimum allowed decrease, respectively. Usually ERU is set to 1.5 and ERF is set to 0.5. When a returning RM cell arrives at the switch, the ER field is updated to be the fair share.

•Charny Max-Min Scheme [18]

The fair share is computed using an iterative procedure in this scheme. Initially, the fair share is set to the link bandwidth divided by the number of active VCs. Some VCs can not achieve the fair share at a switch because of the constraints imposed by the limited amount of bandwidth available at other switches along its path. For this switch, these VCs are called "constrained VCs". The switch can determine whether a VC is constrained or not by comparing the fair share with the CCR field in the received forward RM cell. If the CCR field is less than the fair share, the VC is a constrained VC. Otherwise, it is an unconstrained VC.

For high throughput, the available bandwidth which the constrained VCs can not use should be utilized by the unconstrained VCs. Hence the fair share is computed as follows:

$$\text{Fair share} = \frac{\text{Link bandwidth} - \sum \text{Bandwidth of constrained VCs}}{\text{Number of VCs} - \text{Number of constrained VCs}}.$$

As a forward RM cell traverses the network, the switch determines whether the VC is constrained or not, recomputes the fair share and reduces the ER field and CCR field of the RM cell down to their fair share. The ER field and CCR field of a backward RM cell may be further reduced down to the most current fair share on the forward path.

•Tsang Max-Min Scheme [19]

This scheme is similar to Charny Max-Min method, except for three differences:

1. The switch does not update the CCR field of the RM cell.

2. The switch determines a VC state depending on ER field, instead of CCR field, of the RM cell.

3. The switch determines the VC state and computes the fair share on forward and backward RM cells, not just the forward RM cell.

The following parameters are set for the above schemes in our example and simulation: Nrm=32, PCR=155Mbps, MCR=0bps, ICR=PCR/16, RIF=PCR/256, and RDF=1/16. For EFCI, we used Q_t=1000 cells. For EPRCA, we set Q_t=1000 cells, α=1/16, and DPF=7/8. Target rate is set to be 95% of the link bandwidth for EFICA and CAPC. Also in the CAPC scheme, we use the following parameters: ERU=1.5, Rup=0.05, ERF=0.5, and Rdn=0.5.

2.3 An Example

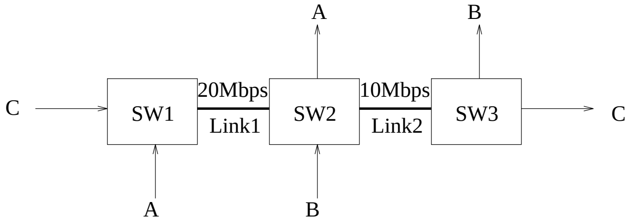
We show a simple example to illustrate the operation of five ER schemes described above. The example is shown in Figure 2. The network configuration consist 3 switches and three connections (A,B,C), as shown in Fig. 2(a). Connection A and B traverse a link, and C traverses two links. The operation of these schemes are exhibited in Fig. 2(b).

3 Simulation Environments

We show the numerical results of these schemes in the homogeneous environment in which all switches utilize the same control scheme and in the heterogeneous environment in which the switches utilize the different control schemes. We first examine the network configuration and performance metrics.

3.1 Network Configuration

We use a simple 3-switch configuration as shown in Figure 3 to be our network topology. It is sufficient to exhibit the characteristics of various switches [20,21]. It consists of 3 switches and some connections grouped into three groups (G1, G2, G3). Group 1 (G1) is the VCs traveling through link 1 only and Group 2 (G2) is the VCs traveling through link 2 only. Group 3 (G3) is the VCs passing both links 1 and 2. There are N1, N2, and N3 connections in the G1, G2, and G3, respectively. In our simulation, there are two VCs in each group, i.e. N1=N2=N3=2.



(a)

At some time instant : CCR(A)=20M, CCR(B)=5M, CCR(C)=10M

Scheme	Initial status	Operation
EPRCA (DPF=7/8)	Congestion MACR1=16 MACR1=8	MACR1 adjustment MACR2 adjustment ER(A) ≈(7/8)*16=14 ER(B)≈(7/8)*8=7 ER(C) ≈ (7/8)*8=7 ....continue....
ERICA (Traget rate=0.95)		Z1=(20+10)/(20*0.95) Z2=(5+10)/(10*0.95) ER(A)=max(20/Z1,20*0.95/2)=12.6 ER(B)=max(5/Z2, 10*0.95/2)=4.8 ER(C)=max(10/Z2,10*0.95/2)=6.4 ....continue....
CAPC (ERU=1.5 Rup=0.05 ERF=0.5 Rdn=0.5)	FS1=15 FS2=10	Z1=(20+10)/(20*0.95) Z2=(5+10)/(10*0.95) ER(A)=FS1*(1-(Z1-1)*0.5)=10.7 ER(B)=FS2*(1-(Z2-1)*0.5)=7.1 ER(C)=FS2*(1-(Z2-1)*0.5)=7.1 ....continue....
CMM		Switch 1: A unconstrained C unconstrained Switch2: B unconstrained C unconstrained ER(A)=10 ER(B)=5 ER(C)=5 → Switch 1: A unconstrained C constrained Switch2: B unconstrained C unconstrained ER(A)=15 ER(B)=5 ER(C)=5 → Stable
TMM		Switch 1: A unconstrained C constrained Switch2: B unconstrained C unconstrained ER(A)=15 ER(B)=5 ER(C)=5 → Stable

(b)

Figure 2: An Example to explain five ER schemes.

The link between two neighboring switches is 100 km in length and 155 Mbps in capacity. For G3, the distance between a source and a switch is 1km. On the other hand, the distance between a source and a switch is 51km for G1 and G2. Hence the propagation delay is the same for all groups. The reason is because we want to eliminate any unfairness caused by the different propagation delays. Also all sources considered in the simulations are persistent.

3.2 Performance Metrics

The following performance measures are reported in the simulation.

Maximum queue length (MQL)– The maximum queue length is directly related to the

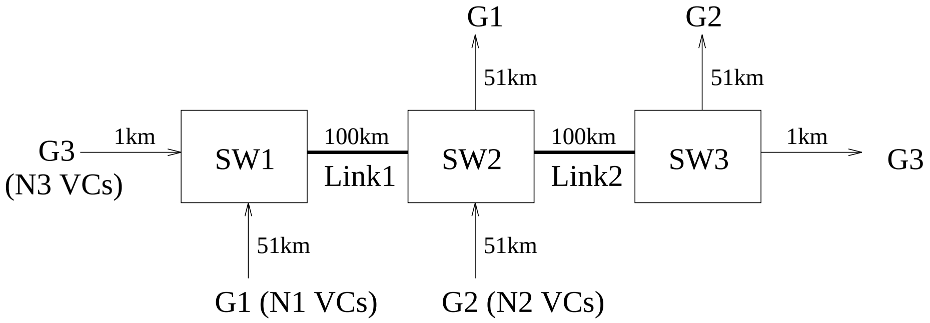


Figure 3: Simulation Model.

cell loss probability when the buffer at the switch is finite. Also we can observe the situation about the oscillation of ACR with this value. When this value is high, the ACR usually has larger oscillation.

Utilization(U)– From the value of utilization, we can know how much bandwidth is wasted at the switch. The utilization is a clear measure of the switch capacity.

Fairness(F)– Unfair behavior is observed from the value of the fairness, which is defined as

$$\mathrm{F} = \max(1 - \max(\text{all } |x_i - 1|), 0)$$

where x_i is the ratio of the actual throughput to the fair throughput for source i. This definition is the maximum ratio difference between ideal and achieved rate. We do not use the commonly used definition [2,6,8,22], $\mathrm{F} = \frac{(\sum x_i)^2}{n \sum x_i^2}$, because only a few differences are observed. This is the major drawback of this metric.

As an example, we consider the case that 25 connections sharing the same bottleneck link which bandwidth is 100Mbps. The fair solution is simply 4Mbps. If a scheme starves one connection completely and distributes its share evenly among other connections, the ratio value (x_1, x_2,...,x_{24},x_{25}) is (100/96,100/96,...,100/96,0). The commonly used definition gives the fairness value 96%, whereas the new metric gives 0, which appears more appropriate for the case of starvation.

4 Numerical Results

4.1 Homogeneous Environment

Simulation results in the homogeneous environment are shown in Figure 4. MQL1 and U1 represent the maximum queue length and utilization at the switch 1. Similarly MQL2 and U2 represent the maximum queue length and utilization at the switch 2. Note the results about switch 3 are not shown because it does not become a bottleneck at any time.

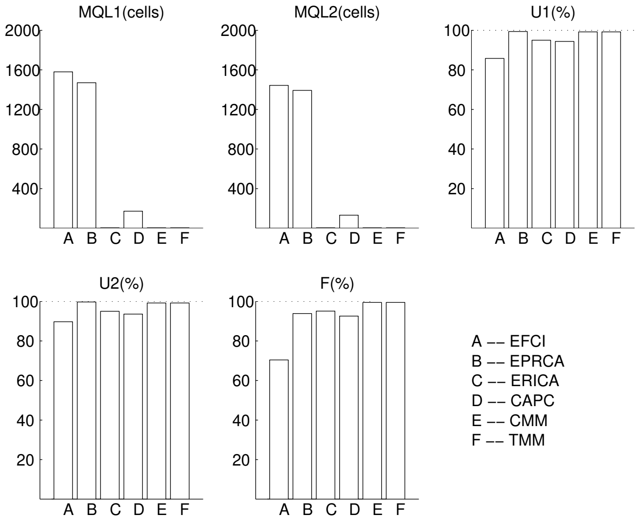


Figure 4: Comparison of the various switch schemes in the homogeneous environments.

From the figure, it is obviously observed that the resulting behavior is more unfair in the homogeneous EFCI environment than those in the homogeneous ER environments. The main reason is that the beat down problem occurs when we use the pure EFCI switches. The beat down problem is that VCs passing through a larger number of switches get less bandwidth than VCs passing through a smaller number of switches [2,8,22]. This effect is caused that the VCs traveling more hops have a higher probability of getting their cells marked than those traveling fewer hops. As a result, it is likely that the long-hop VCs can not increase their rates, and consequently they are beaten down by the short-hop VCs.

Also we observe that the maximum queue length of EFCI and EPRCA is large. This is caused by the large oscillation of ACR. Actually, in the homogeneous environments, the ACR of EFCI and EPRCA has large oscillation, the ACR of ERICA and CAPC has small

Type	Scheme	Setting	Advantage	Drawback
EFCI	EFCI	using queue length	simplicity	high oscillation of ACR high maximum queue length seriously beat down problem
ER	EPRCA	using MACR	simplicity	high oscillation of ACR high maximum queue length beat down problem parameter tuned problem
ER	ERICA	using load factor and CCR	do not need per-connection information	achieve target load sensitive to CCR errors
ER	CAPC	using load factor and some parameters	do not need per-connection information	achieve target load parameter tuned problem
ER	CMM	Max-Min method using CCR	oscillation free high utilization fairness	need per-connection information sensitive to CCR errors
ER	TMM	Max-Min method using ER	oscillation free high utilization fairness	need per-connection information sensitive to CCR errors

Table 2: Comparison of the various switch schemes in the homogeneous environments.

oscillation, and the ACR of CMM and TMM is oscillation free [9-11].

Regarding utilization, link capacity is not fully utilized because queue length threshold, Q_t, is set too low in the homogeneous EFCI environment. However, if Q_t is set high, the maximum queue length will be raised dramatically [23,24]. Hence we sacrifice some bandwidth to keep the maximum queue length in the reasonable range. On the other hand, high utilization is achieved in the homogeneous ER environments. EPRCA has high utilization. ERICA and CPAC achieve the target rate. For TMM and CMM, they utilize almost the complete link bandwidth. We summary these results which are observed in the homogeneous environments as Table 2.

The noticeable thing is that CMM and TMM are sensitive to CCR errors. This sensitivity does not depend on whether CCR value is actually used by the algorithm or not, but rather on the fact that both schemes are purely computational algorithms with no rate measurements involved. It should be noted that this issue may be overcome by using rate measurements at the source.

4.2 Heterogeneous Environment

In this heterogeneous simulation we use an EFCI scheme on switch 1 and an ER scheme on switch 2. The results of interoperating EFCI with various ER schemes are in Figure 5. In order to compare the performance in the heterogeneous environment, we also plot the performance in the pure EFCI environment and pure CMM environment, which is denoted by A and G, respectively. Also Figure 6 shows a typical ACR dynamics in the EFCI-ER environments. From the results, we observe some important implications which are not shown in the homogeneous environments.

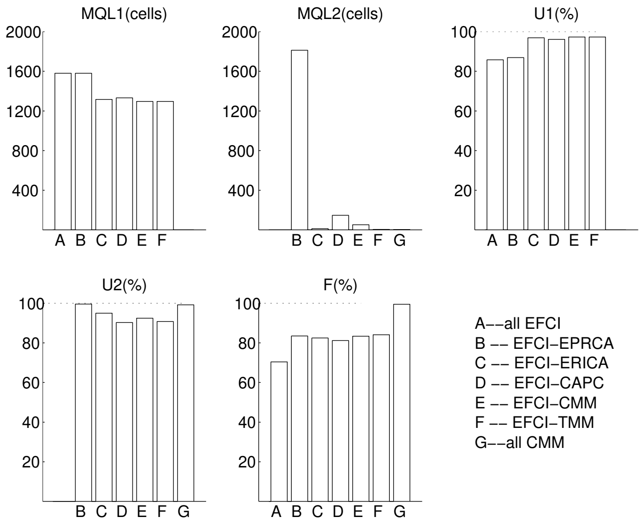


Figure 5: Comparison of the various ER schemes in the mixed EFCI-ER environments (RIF=1/256).

- Beat down problem exists when EFCI switch appears.

In the mixed EFCI-ER environment, the beat down problem still exists. This is due to the existence of EFCI switch, as observed in Figure 6. The ACR of VCs which pass EFCI and ER switches (G3) is not increased to a high level because it is restricted by ER switches. On the other hand, it must be decreased to a low level when the congestion occurs in the EFCI switches. As a result, beat down problem happens.

Actually, more EFCI switches cause more severe beat down problem in the ATM networks. On the other hand, more ER switches relieve the severity of beat down problem.

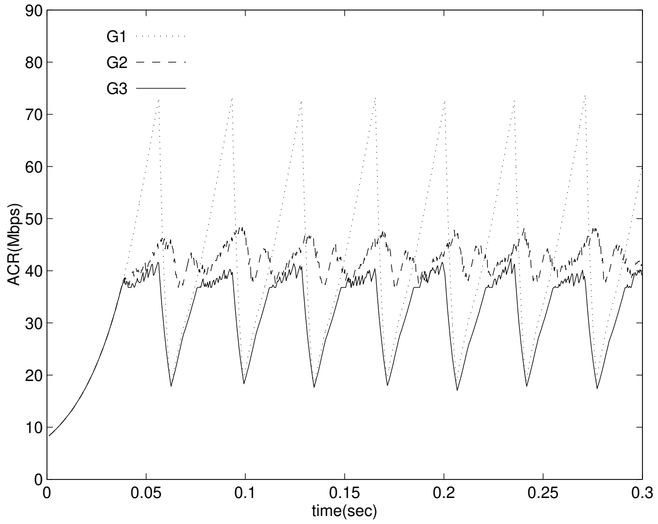


Figure 6: A typical ACR dynamics in the EFCI-ER environment.

- Utilization of EFCI switches raises and the utilization of ER switches drops.

From Figure 5, this phenomenon is easily observed. Simply speaking, because EFCI switches own the high oscillation of ACR, it can easily grab the unused bandwidth. Hence the utilization of EFCI switches raises. On the other hand, the low oscillation of ACR in the ER switches causes the slow response to unused bandwidth. Therefore much bandwidth is wasted, and the utilization of ER switches drops.

- CCR field in the RM cells may be belated.

In the homogeneous ER environments, this condition happens rarely. However, in the mixed EFCI-ER environment, it occurs often. When the congestion of the EFCI switch occurs, i.e. there are at least Q_t cells in the EFCI switch, the algorithm using the CCR field in the RM cells does not obtain the current cell rate because the RM cells are queued at EFCI switch. Hence the CCR fields of delayed RM cells are belated. As a result, ER switch may use inaccurate CCR field.

- Do not use ER field in the RM cells.

TMM is a representative algorithm using the ER field in the RM cells. We can observe that the utilization of the TMM switch is low in the mixed EFCI-ER environments. As shown in Figure 7, TMM does not obtain any ER information because the EFCI switch does not set the ER field. Hence the scheme loses the responsiveness. Hence we suggest that all schemes do not use the ER information in the RM cells if EFCI switches may co-exist.

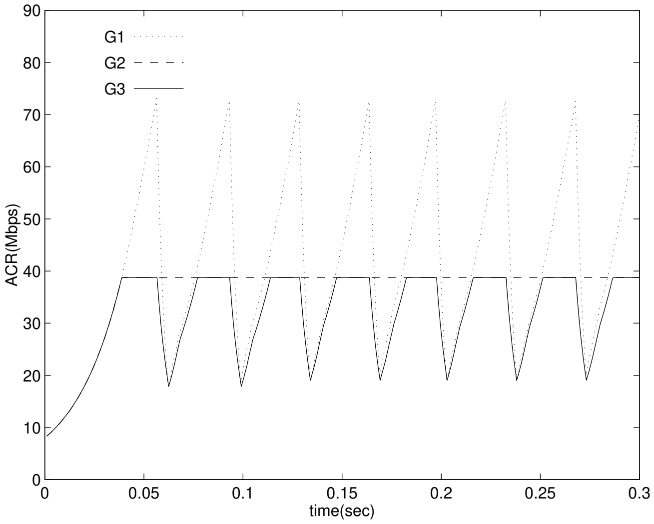


Figure 7: ACR dynamics in the EFCI-TMM environments.

• The value of RIF should not be set to a large value.

Some papers use the high RIF to reduce the length of the transient period in the homogeneous ER environment [18,19,20]. When using the high RIF, the ACR of source can quickly achieve the level of ER field, regardless of what value of RIF. Hence the response time is reduced, especially for CMM or TMM.

In order to understand the effect of RIF in the mixed EFCI-ER environment, we conducted the same simulation except with RIF=1. The results are shown in Figure 8. We observe that all EFCI switches must endure very large maximum queue length. Also the EFCI switches are often in the congestion state. We do not like to see this situation. On the other hand, ER switches do not obtain any benefit from that. All ER switches must endure larger maximum queue length. The utilization and fairness of all ER switches drop.

• The performance of CMM is better.

Finally the performance comparisons in the homogeneous and mixed environments are tabulated in Table 3. To clearly display the advantages and weaknesses of various switches, ranking is used, instead of the real value. From the table, we observe that the EPRCA is

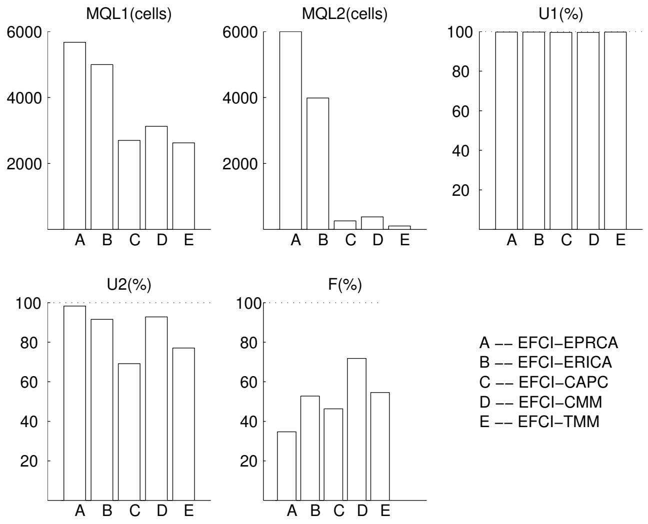


Figure 8: Comparison of the various ER schemes in the mixed EFCI-ER environments (RIF=1).

Scheme	Utilization			Max queue length			Fairness		
	Homo	Mixed (RIF= $\frac{1}{256}$)	Mixed (RIF=1)	Homo	Mixed (RIF= $\frac{1}{256}$)	Mixed (RIF=1)	Homo	Mixed (RIF= $\frac{1}{256}$)	Mixed (RIF=1)
EPRCA	1	1	1	5	5	5	5	5	5
ERICA	4	2	3	3	2	4	3	3	3
CAPC	5	5	5	4	4	2	3	3	4
CMM	2	3	2	1	3	3	1	1	2
TMM	2	4	4	1	1	1	1	1	2

Table 3: Performance ranking of the various ER schemes in the homogeneous and mixed EFCI-ER environments.

the best in utilization and the worst in the maximum queue length and fairness. The Max-Min schemes can achieve the better fairness. The performance of the ERICA and CAPC is medium. In summary, from the ranking of utilization, maximum queue length, and fairness, we feel that CMM is a better scheme.

5 Location of ER switch

In this section, we want to investigate the location where ER switches should be placed in the mixed EFCI-ER environments. To save space, CMM scheme is used to represent for the various ER switches because of its good performance. Some simulations are done on various configurations. From the simulation results, a few rules are given to help us place the ER switch in the EFCI-ER environments.

• Bottleneck vs. non-bottleneck : N1=0, N2=2, N3=2

Two VCs travel through the link 2 only, and the other two VCs travel through both links 1 and 2. In this case, switch 2 is the bottleneck switch. The simulation results are shown in Figure 9. We observe that the performance is better when ER scheme is placed in the bottleneck switch. This placement can produce the same results as the case that all switches adopt the CMM scheme. On the other hand, if the ER switches are placed at the non-bottleneck points, their functions are not fully utilized. We can see that the placement of the ER switch at a non-bottleneck point has no effect in our simulation.

There is another important reason for this arrangement. Under this configuration, if the algorithm at the bottleneck switch is substantially more oscillatory than the algorithm at the non-bottleneck switch, the rate mismatch problem occurs, which causes unfairness [20].

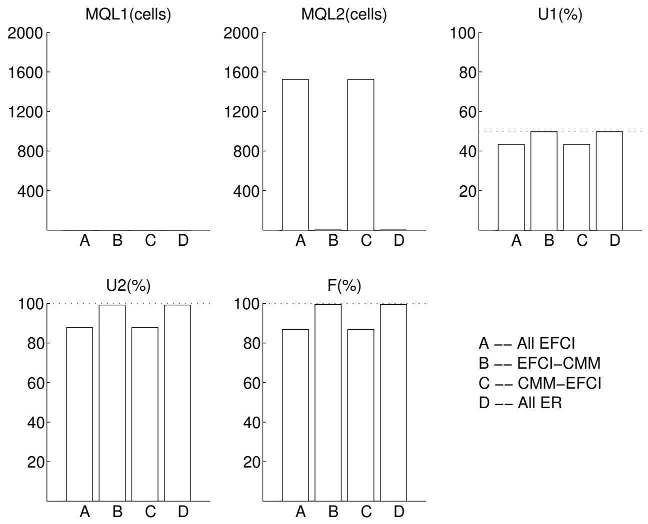


Figure 9: Comparison of the various switch schemes in the configuration (N1=0, N2=2, N3=2).

• Most critical vs. least critical : N1=2, N2=4, N3=2

Under this configuration, the number of VCs passing through the link 1 is 4, and the number of VCs passing through the link 2 is 6. First we define the concept of the term 'critical'. The most critical switch for a VC can be described as a switch that gives the source the lowest fair share based on the max-min fairness criteria. Similarly, the least critical switch is the switch that gives the highest fair share in the VC's path. In our experiment, switch 1 is the least critical switch and switch 2 is the most critical switch.

The simulation results are shown is Figure 10. The total utilization of both switch (U1+U2) of EFCI-CMM placement is higher than that of CMM-EFCI placement. Also the fairness degree of EFCI-CMM is obviously larger than that of CMM-EFCI placement. Hence we suggest that an ER switch ought to be placed at the most critical point first. This conclusion is completely opposite to another study in [21]. It suggested that an ER switch should be placed at the least critical point first because that can relieve the severe beat down problem more.

From Table 4, the throughput of G3 is actually higher when the ER scheme is imposed on the least critical switch, compared with that when imposed on the most critical switch. Therefore, we should place the ER switch at the least critical point from the view of the beat down problem [21]. However, the extra bandwidth of G3 is mainly obtained from the bandwidth released by G1. In fact, much bandwidth is wasted while just a little bandwidth is used by G3.

Therefore, it is our opinion that an ER switch ought to be placed at the most critical point first. This placement can produce smaller maximum queue length, larger fairness degree, and higher throughput, although the beat down problem remains serious.

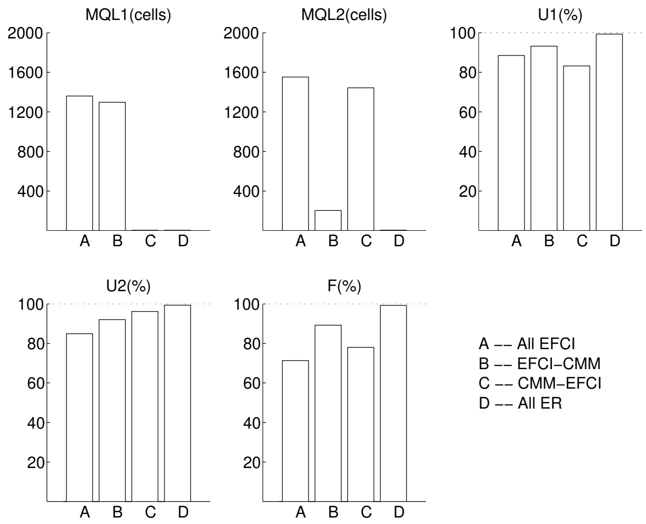


Figure 10: Comparison of the various switch schemes in the configuration (N1=2, N2=4, N3=2).

Scheme	**G1**	**G2**	**G3**	**F(%)**
Ideal	51.83	25.92	25.92	100
all EFCI	50.35	29.10	18.47	71.3
EFCI-CMM (most critical switch is ER)	49.39	25.33	23.13	89.2
CMM-EFCI (least critical switch is ER)	40.43	26.14	24.25	78.0
all CMM	51.63	25.72	25.72	99.2

Table 4: Throughput of all groups under the various combination of switch schemes.

- Near source vs. near destination : N1=2, N2=2, N3=2

In this case, the performance is better when the ER switch is located near the source, from Figure 11. This is because lots of ER schemes (except for CMM and TMM) record the information from the forward RM cell, and modify the ER field on the backward path. Hence the newer information is received at the ER switch and quicker feedback is sent to the source when the ER switch is near the source. On the other hand, the EFCI switch marks each forward data cell when congestion occurs, and does not do anything on the backward path. Hence when it is placed near the source, the propagation delay increases, which causes the slower response at the source.

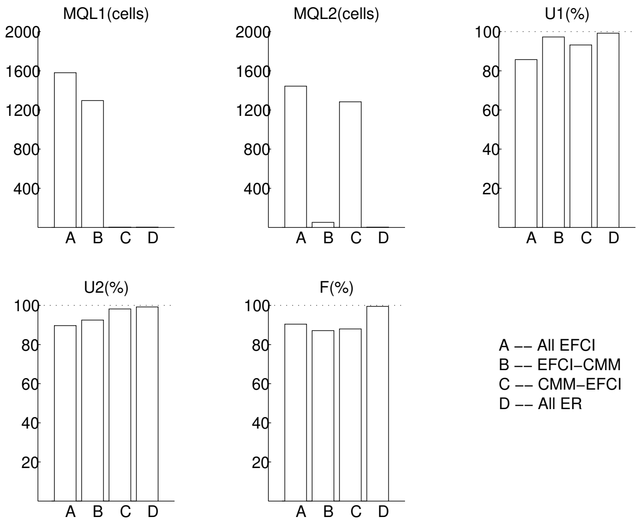


Figure 11: Comparison of the various switch schemes in the configuration (N1=2, N2=2, N3=2).

Under the configurations described as above, we can determine how to place the ER switch. However, there may be some conflicts among them. For example, a backbone switch is generally the bottleneck and most critical point, and is also far from the source. According to the simulation results (Fig. 9, 10, 11), we set the priority order as: bottleneck > critical > distance. That is, we feel that a backbone switch should be considered first to have ER capacity. Nevertheless, ER switches should replace EFCI switches whenever possible everywhere in the network.

6 Future works

There are a number of ways to extend further research work on the interoperability issues discussed in this article. The first one is to conduct more simulation experiments in a larger variety of configurations and under different traffic patterns. In particular, a performance comparison in the presence of intermittent and greedy flows might be useful. Another one is the interoperability of the newly developed ER schemes and EFCI switches. Currently some new algorithms to set ER field have been developed [25-27]. These algorithms should be verified in the heterogeneous environments. Last, the interoperability among the various ER switches is also worth of research. After the transitional from the first to second generation switches, all switches may be equipped with ER function. The cooperation or conflict among different ER schemes should be studied.

References

[1] ATM Forum, "ATM Forum Traffic Forum Traffic Management Specification Version 4.0," ATM Forum af-tm-0056.000, April 1996.

[2] B. Flavio and W. F. Kerry, "The Rate-Based Flow Control Framework for the Available Bit Rate ATM service," IEEE Network, March 1995, vol. 9, no. 2, pp. 25-39.

[3] S. Floyd and A. Romanow, "Dynamics of TCP traffic over ATM networks," Proc. ACM Sigcomm'94, Sept. 1994, pp. 79-88.

[4] C. M. Ozveren, R. Simcoe, and G. Varghese, "Reliable and Efficient Hop-by-Hop Flow Control," IEEE JSAC, vol. 13, no. 4, May 1995, pp. 642-650.

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

[6] R. Jain, "Congestion Control and Traffic Management in ATM Networks: Recent Advances and A Survey," Computer Networks and ISDN Systems, vol. 28, no. 13, Oct. 1996, pp. 1723-1738.

[7] P. Newman, "Backward Explicit Congestion Notification for ATM Local Area Networks," Proc. IEEE Globecom, Dec. 1993, pp. 719-723.

[8] H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda, and H. Miyahara, "Rate-Based Congestion Control for ATM Networks," Computer Communication Review, April 1995.

[9] D. Sisalem and H. Schulzrinne, "Switch Mechanisms for the ABR Service: A Comparison Study," http://www.lirmm.fr/atm/articles.html, June 1996.

[10] A. Kolarov and G. Ramamurhy, "Comparison of ER and EFCI Flow Control Scheme for ABR Service in Wide Area Networks," Proc. IFIP Broadband Communication, April 1995.

[11] L. Gerda and O. Casals, "A Simulation Study of Switching Mechanisms for ABR Service in ATM Networks," Lloren Cerd Olga Casals. Interim Report UPC-DAC-1996-21, Feb. 1996.

[12] ITU-T SG 13, "Traffic Control and Congestion Control in B-ISDN," Draft Rec. I.371, Nov. 1995.

[13] H. Saito, K. Kawashima, H. Kitazume, A. Koike, M. lshizuka, and A. Abe, "Performance Issues in Public ABR Service," IEEE Communication, vol. 34, no. 11, Nov. 1996, pp. 40-48.

[14] N. Yin and M. G. Hluchy, "On Closed-Loop Rate Control for ATM Cell Relay Networks," Proc. IEEE Infocom'94, 1994, pp. 99-108.

[15] L. Roberts, "Enhanced Proportional Rate Control Algorithm," ATM Forum 94-0735R1, Aug. 1994.

[16] R. Jain, S. Kalyanaraman, R. Goyal, S. Fahmy, and R. Viswanathan, "The ERICA Switch Algorithm: A Complete Description," ATM Forum 96-1172, Aug. 1996.

[17] A. W. Barnhart, "Explicit Rate Performance Evaluations," ATM Forum 94-0983, Oct. 1994.

[18] A. Charny, K. K. Ramakrishnan, and A. Lauck, "Time Scale Analysis and Scalability Issues for Explicit Rate Allocation in ATM Networks," IEEE/ACM Trans. on Networking, vol. 4, no. 4, Aug. 1996, pp. 569-581.

[19] Danny H. K. Tsang and Wales K. F. Wong, "A New Rate-based Switch Algorithm for ABR Traffic to Achieve Max-Min Fairness with Analytical Approximation and Delay Adjustment," Proc. IEEE Infocom'96, 1996, pp. 1174-1181.

[20] N. T. Plotkin and J. J. Sydir, "The Rate Mismatch Problem in Hetergeneous ABR Flow Control", Proc. IEEE Infocom'97, 1997.

[21] Y. Chang, N. Golmie, and David Su, "Study of Interoperability between EFCI and ER Switch Mechanisms for ABR Traffic in an ATM Network," Proc. 4th Int'l Conf. Computer Communication and Networks, 1995, pp. 310-315.

[22] A. Arulambalam, X. Chen, and N. Ansari, "Allocating Fair Rates for Available Bit Rate Service in ATM Networks," IEEE Communication, vol. 34, no. 11, Nov. 1996, pp. 92-100.

[23] Y. C. Lai and Y. D. Lin, "Performance Analysis of Rate-based Congestion Control Scheme and Choice of High and Low Thresholds," Proc. 6th Int'l Conf. Computer Communication and Networks, Sept. 1997, pp. 70-75.

[24] Y. C. Lai and Y. D. Lin, "Performance Analysis of Rate-based Flow Control under a Variable Number of Sources," Computer Networks and ISDN systems, to appear.

[25] F. M. Chiussi, A. Arulambalam, Y. Xia, and X. Chen, "Explicit Rate ABR Schemes Using Traffic Load as Congestion Indicator," Proc. 6th Int'l Conf. Computer Communication and Networks, Sept. 1997, pp. 76-84.

[26] L. Kalampoukas, A. Varma, and K. K. Ramakrishnan, "An Efficient Rate Allocation Algorithm for ATM Networks Providing Max-Min Fairness," Proc. 6th IFIP Int'l Conf. High Performance Networking, Sept. 1995. pp. 143-154.

[27] A. Arulambalam, X. Chen, and N. Ansari, "An Intelligent Explicit Rate Control Algorithm for Available Bit Rate Services in ATM Networks," Proc. ICC'97, June 1997, pp. 200-204.