PAPER

Congestion Control for ABR Service Based on Dynamic UPC/NPC

Katsumi Yamato†, Hiroshi Esaki † *Members*

SUMMARY A novel reactive congestion control scheme based on Dynamic UPC/NPC (Usage/Network Parameter Control) in ATM networks is proposed. In this scheme, policing parameters at the UPC/NPC are dynamically modified in response to the reception of RM (Resource Management) cells. In a congested state, traffic volume submitted to the network is regulated by Dynamic UPC/NPC, while providing negotiated QoS (Quality of Service) for each ATM connection.

When end-stations (or edge-entities between network segments) operate according to ER-based (Explicit Rate based) behavior, a UPC/NPC function will indicate (send) an ER value toward each source end-station using backward RM cells. In this case, the policing parameter at the UPC/NPC should take the same value as the ER value. When end-stations (or edge-entities) operate according to EFCI-based (Explicit Forward Congestion Indication based) behavior, the modified policing parameter at the UPC/NPC point must be harmonized with the modified cell transmission rate at the source end-stations (or at the edge-entities). In order to improve the control performance for the long distance connections, backward RM cells will be generated by the NPC function (UPC function will be optional) at the egress of a congested network in response to the reception of EFCI marked cells (or forward RM cells) as a *proxy* destination end-station, and they will be sent back toward the UPC/NPC function at the ingress of the network.

As a result, the proposed control scheme enables the network to recover from the congested state securely and provide the negotiated service quality, even if cooperation of (rate-based) flow control at each source end-station (and at edge-entities between network segments) is not expected.

key words: ATM Network, Congestion Control, UPC/NPC, ABR Service, RM Cell, Policing Parameter

1. Introduction

ATM technology is recognized as a solution for B-ISDN as well as high speed LANs (Local Area Networks). Progress has been made in the standardization and specification of related technologies in ITU-T, the ATM Forum, and other related organizations.

ATM networks are expected to provide several service classes[1]. For CBR (Constant Bit Rate) or VBR (Variable Bit Rate) service classes, ATM networks may allocate certain network resources through CAC (Connection Admission Control) in order to achieve the required CLR (Cell Loss Rate) and CTD (Cell Transfer Delay) performance objectives, and networks do not expect flow control to be performed by end-stations according to the congestion indication issued by the network. In the case of ABR (Available Bit Rate) service classes, however, the networks achieve CLR performance objectives, but do not achieve CTD performance objectives. In order to minimize the CLR, the networks indicate their congestion levels by using ER (Explicit Rate) values in forward/backward RM (Resource Management) cells, or by using EFCI (Explicit Forward Congestion Indication) bits in cell headers. When backward RM cells are received by source end-stations, the source end-stations will change their cell transmission rates for the ABR connection according to the congestion level indicated in the RM cells. Therefore, such networks do expect the cooperation of source end-stations' flow control to regulate their cell submission rate to the network.

In the case of ABR services, each connection will request an MCR (Minimum Cell Rate) value as a bandwidth to be provided by the network at the connection set-up phase. When MCR=0, the connection will not be blocked due to the lack of bandwidth or buffer resource[1]. The source end-station can submit cells at a rate of at least the MCR value (but never exceeding PCR (Peak Cell Rate) value), if the network bandwidth is available. The number of cells submitted to the network will be regulated through the cooperation of the source end-station's flow control according to the congestion indication (ER value or EFCI bit) of the network. This means that congestion will last in the network if the source end-station does not cooperate with it††. The ATM network *must* provide a certain cell loss quality for the ABR service[1], and this is especially true for public networks. Many reactive congestion control schemes for ABR services have been discussed and proposed in the ATM Forum[1, 2, 3, 4, 5], but none of them provide sufficient control. For example, these control schemes expect the cooperation of end-stations in order to prevent the network from falling into a congested state[3, 4]. However, if some end-stations do not cooperate with the schemes correctly (e.g., due to the misbehavior of these end-stations), the network may

Manuscript received April 18, 1995.
Manuscript revised September 18, 1995.
†The authors are with R&D Center, TOSHIBA CORPORATION, Kawasaki-shi, 210 Japan.

††In this paper, congestion is defined as a state of network elements (e.g., switches) in which the network will not be able to achieve the negotiated network performance objectives for already established connections or new connections.

fall into a congested state.

In this paper, we propose a novel reactive congestion control scheme based on *Dynamic* UPC/NPC[6, 7] (Usage/Network Parameter Control[8, 9]) which monitors and controls the cell flow entering the ATM network according to the indications provided by backward RM cells. By applying the proposed Dynamic UPC/NPC policy, the network can provide sufficient CLR quality without expecting flow control cooperation at the source end-stations. The features of the proposed scheme are secure operation, quick response to congestion, and control flexibility.

2. Conventional Traffic Control and Congestion Control Frameworks

2.1 Conventional ABR Congestion Control

In this subsection, ABR congestion control schemes proposed in the ATM Forum[1, 4] will be outlined.

In an ABR service, the source (end-station) sends one forward RM cell to the destination (end-station) for every N cells (Nrm), and the destination changes the direction bit in the received RM cell from "Forward" to "Backward" and returns the backward RM cell to the source. The source sets the ER (Explicit Rate) value in the RM cell to PCR, and sets the CI (Congestion Indication) bit in the RM cell to zero (not congested). To indicate whether the network is in a congested state or not, the switches in the network will implement either EFCI-based (Explicit Forward Congestion Indication based) or ER-based (Explicit Rate based) behavior. In **EFCI-based behavior**, the switch in a congested state marks the EFCI bit (sets EFCI=1) in each cell sent by the source in order to indicate congestion in the forward direction (from the source to the destination). And at the destination, if a cell with EFCI=1 has been received, it marks the CI bit (sets CI=1; congested) in the backward RM cell to return it to the source. In **ER-based behavior**, the switch reduces the ER value in a forward and/or backward RM cell to indicate its congestion level when the available cell transmission rate at the switch is less than the ER value written in the received RM cell. Here, in both of the above behaviors, the switch optionally sets CI=1 in the backward RM cells to ensure that the source does not increase (or decrease) its cell transmission rate. At the source, when a backward RM cell is received with CI=1, it decreases the ACR (Allowed Cell Rate : a cell transmission rate at which cells are allowed to be transmitted) by the amount negotiated at the connection set-up phase. On the other hand, when a backward RM cell is received with CI=0, it increases the ACR by the negotiated amount[1, 4]. When any backward RM cell is received, the source sets the ACR to the minimum of the ER value in the backward RM cell and the ACR value calculated by itself. The source shall never send cells at a rate exceeding the ACR.

The issue which we would like to raise first in this paper is that the above scheme expects the cooperation of end-stations (both source and destination) in order to prevent the network from falling into a congested state. In other words, if non-compliant end-stations, which ignore the control signals issued by the network, are connected to the network, the network may fall into a congested state and may not be able to provide the negotiated QoS (Quality of Service) for every ABR cell flow. As examples, the following problems at the non-compliant end-stations are considered;

- Ignoring the contents in the received backward RM cell at the source end-stations :
  A non-compliant source end-station may not reduce its cell transmission rate (i.e., ACR) when a backward RM cell is received with CI=1. In addition, a non-compliant source end-station may modify its cell transmission rate to a value larger than the ER value in the received backward RM cell.
- Insufficient forward RM cell generation at the source end-stations :
  A non-compliant source end-station may not generate sufficient forward RM cells. The network (and destination end-station) may therefore not be able to indicate sufficient congestion and control indications to the source end-station.
- Insufficient backward RM cell transmission at the destination end-stations :
  A non-compliant destination end-station may not return backward RM cells correctly to the source end-station, and these may include control and congestion signals written in the received cells (i.e., forward RM cells and EFCI marked cells).

Another important issue is that the performance of the above scheme depends on the propagation delay between the source end-station and the destination end-station. If the scheme is applied to long distance VCCs (Virtual Channel Connections) (e.g., VCCs across WANs (Wide Area Networks)), it will take a long time for the source end-station to receive the RM cell returned by the destination end-station. Therefore, the network may recover from the congested state more slowly. In the ATM Forum, to reduce the propagation delay, each connection can be divided into two or more ABR flow control loops, each of which is originated by a virtual source (except the first) and is terminated by a virtual destination (except the last)[1, 4]. The virtual source/destination will perform the same functions as an ABR source/destination end-station. Here, if the virtual source and virtual destination are applied to a long distance VCC (Virtual Channel Connection), the virtual source and virtual destination of each ABR control segment may need a huge buffer space to provide sufficient CLR quality.

Therefore, a new control mechanism is needed to avoid deterioration of the QoS for well-behaved end-stations that may be a result of the problems outlined above including long distance VCCs without virtual sources/destinations.

2.2 Conventional UPC/NPC Functions

UPC/NPC is defined as the set of actions taken by the network to monitor and control individual ATM connections in terms of their traffic volume and validity of their VPI/VCI (Virtual Path/Channel Identifier) values at the user access or the network access respectively. The main purpose of UPC/NPC is to protect network resources from malicious as well as unintentional misbehavior which will degrade the QoS of other already established connections. Appropriate actions are taken when a violation of policing parameters for individual connections, which are negotiated as traffic contracts with the network at the connection set-up phase, is detected. To enforce the negotiated traffic contract, discarding, tagging, or rescheduling (i.e., shaping) shall be performed on cells which are identified as not conforming to the contract[8, 9].

Originally, UPC/NPC was performed to monitor and control traffic offered by CBR or VBR connections. For these connections, the network would allocate certain network resources because these service classes would clearly require certain CTD (and CDV (Cell Delay Variation)). However, for the ABR service class, the network may allocate only its resources corresponding to the MCR values required at the connection set-up phase. Each ABR connection can use resources corresponding to its MCR value at any time and can use some of remaining resources which will not be allocated to any connections. Sharing the remaining resources by multiple ABR connections would cause network congestion unless some reactive congestion control is carried out in the network. Even if conventional UPC/NPC monitors the ABR cell flow using PCR as a policing parameter, it will not be sufficient to protect the network resource. This is because sharing the remaining resources by ABR connections would cause network congestion if some of the end-stations did not carry out the reactive congestion control correctly.

Dynamic GCRA (Generic Cell Rate Algorithm) is proposed in the ATM Forum[1] as an algorithm to monitor the ABR cell flow. The Dynamic GCRA specifies cell conformance to the ER value in the backward RM cell. When switches in the ATM network operate according to ER-based behavior, the UPC/NPC which executes the Dynamic GCRA will modify the policing parameter according to the ER value in the received RM cell. However, when switches in the network operate according to EFCI-based behavior (i.e., when the network can indicate congestion by marking the CI bit in the RM cell), the UPC/NPC which executes the Dynamic GCRA may not modify the policing parameter because the network cannot modify the ER value. Therefore, even if the UPC/NPC executes the Dynamic GCRA, it will not be sufficient to protect the resource of the EFCI-based network.

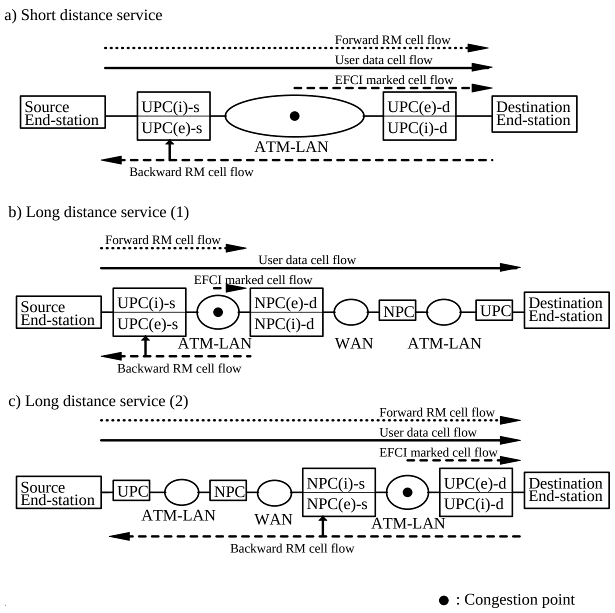
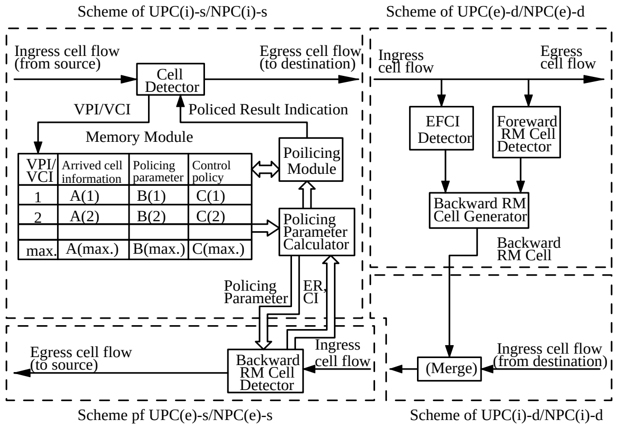


Fig. 1 Reactive Congestion Control using UPC/NPC



VPI/VCI	Arrived cell information	Policing parameter	Control policy
1	A(1)	B(1)	C(1)
2	A(2)	B(2)	C(2)
max.	A(max.)	B(max.)	C(max.)

Fig. 2 Actual Implementation of the UPC/NPC

3. Congestion Control based on Dynamic UPC/NPC

In this section, the proposed congestion control scheme based on Dynamic UPC/NPC[6, 7] is described.

3.1 Functional Model of Dynamic UPC/NPC

In order to avoid network congestion caused by misbehaving end-stations and/or neighboring network segments on ABR connections, we propose that some new functions should be added to conventional UPC/NPC (in this paper, we call this "Dynamic UPC/NPC"). Figure 1 shows an example of the proposed reactive congestion control scheme based on Dynamic UPC/NPC, and Figure 2 shows an example of actual implementation of the proposed UPC/NPC. For convenience, we classify the UPC/NPC according to its location as follows;

- **UPC(e)/NPC(e)**:
  UPC(e)/NPC(e) is located at the egress point of an ATM network between the congested point (switch) and the destination end-station. The following functions are added to conventional UPC/NPC;
  - **EFCI Detector**, which detects EFCI marked cells,
  - **Forward RM Cell Detector**, which detects forward RM cells,
  - **Backward RM Cell Detector**, which detects backward RM cells and, if necessary, overwrites the ER values in the cells according to the policing parameter modified by the Policing Parameter Calculator in UPC(i)/NPC(i),
  - **Backward RM Cell Generator**, which generates backward RM cells sent back toward the source end-station when forward RM cells (or EFCI marked cells) are received.

  The function "Backward RM Cell Generator" may be needed only if the propagation delay between the UPC(e)/NPC(e) and the destination end-station is very large. In this case, the backward RM cell should be generated according to the backward RM cell generation rule defined for switches[1].
- **UPC(i)/NPC(i)**:
  UPC(i)/NPC(i) is located at the ingress point of an ATM network between the source end-station and the congested point. It monitors and regulates the cell flow between the source end-station and the network. The following functions are added to the conventional UPC/NPC;
  - **Policing Parameter Calculator**, which calculates a policing parameter according to the contents (either CI bit or ER value) in the backward RM cells received by the Backward RM Cell Detector in UPC(e)/NPC(e).

As described in section 2.2, when switches in the ATM network operate according to EFCI-based behavior that cannot modify the ER value, Dynamic GCRA[1] at the conventional UPC/NPC cannot operate correctly. On the other hand, the Dynamic UPC/NPC can modify the policing parameter according to the CI bit in the backward RM cell, and replace the ER value in the backward RM cell with the modified policing parameter. Therefore, the network which applies the Dynamic UPC/NPC can inform the source end-station and/or the neighboring network of the congestion level by using the ER value even if the switches in the network operate according to EFCI-based behavior.

To explain the functions of the UPC/NPC located at the egress and ingress points of congested networks, the denotations "UPC(i)", "NPC(i)", "UPC(e)", and "NPC(e)" are used in Fig. 1. In other words, UPC(i)/NPC(i) and UPC(e)/NPC(e) are subsets of the proposed Dynamic UPC/NPC mechanism. Actually, both these functions will be included in the single UPC/NPC entity shown in Fig. 2.

3.2 Detailed Action of Dynamic UPC/NPC

The detailed action of the Dynamic UPC/NPC is shown in Fig. 1 a) and Fig. 2. Although only "Dynamic UPC" is explained in this subsection, almost the same would apply to "Dynamic NPC". Here, for convenience, the denotations "UPC(i)-s", "UPC(e)-s", "UPC(i)-d", and "UPC(e)-d" are used in order to explain the proposed UPC/NPC mechanism precisely. UPC(i)-s and UPC(e)-s are subsets of the proposed UPC located between the source end-station and the network, and UPC(i)-d and UPC(e)-d are subsets of the UPC located between the destination end-station and the network. The cells transferred from the source end-station are policed at UPC(i)-s.

3.2.1 In the Case of Normal States

In a normal state (i.e., not a congested state), the UPC(i)-s monitors and controls the cell flow like a conventional UPC (e.g., PCR policing). As shown in Fig. 2, when a cell sent by a source end-station arrives at the UPC(i)-s, a Cell Detector detects it and extracts its VPI/VCI value. When a Memory Module receives the VPI/VCI value from the Cell Detector, arrived cell information ((the set of) time(s) of the latest cell arrival(s)) and a policing parameter for the received VPI/VCI is transmitted from the Memory Module to the Policing Module. The Policing Module determines whether or not the arrived cell conforms to the policing parameter, and informs the Cell Detector of the result. In this paper, the maximum cell transmission rate allowed by the UPC(i)-s is used as a policing parameter. When the arrived cell conforms to the policing parameter, the Cell Detector transmits the cell toward a destination end-station and the Memory Module updates arrived cell information to include the arrival time of the latest cell. When the arrived cell does not conform to the policing parameter, the Cell Detector discards, tags, or reschedules the cell to enable the cell flow to conform with the policing parameter.

3.2.2 In the Case of Congested States

In a congested state, the Dynamic UPC may modify the policing parameter to enable the recovery of switches in the network from the congested state by reducing the incoming cell flow. The UPC scheme is applicable to networks whose switches operate according to EFCI-based or ER-based behavior[1, 4].

3.2.2.1 Operation for EFCI-based Network

When a switch is going to fall into (or is already in) a congested state, it marks EFCI bits for all cells passing through the switch. The EFCI marked cells are conveyed along the forward path toward the destination end-station. When the EFCI marked cell is received by the UPC(e)-d, the EFCI Detector in Fig. 2 detects it and notices that the switch along the path of the received cell is in a congested state. Here, the EFCI marked cell received at the UPC(e)-d should be transparently transmitted to the destination end-station.

When the UPC(i)-d receives the backward RM cell that is returned by the destination end-station, the UPC(i)-d sets CI=1 in the received backward RM cell if the UPC(e)-d notices the congestion of the switch. According to the above behavior of the UPC(i)-d, the network would inform the source end-station of the existence of the congestion even if the destination end-station did not mark the CI bit in the backward RM cell correctly because of system failure.

If the destination end-station does not generate a backward RM cell correctly (e.g., due to the non-compliant behavior of the destination end-station), the source end-station cannot reduce the cell transmission rate according to the network information conveyed in the backward RM cell (e.g., ER and CI). In this case, the network would not be able to provide sufficient control performance. In order to avoid this, the network can implement the following functionality in UPC(i)-d and UPC(e)-d as an *optional function*. When the EFCI Detector receives an EFCI marked cell and notices that the network is in a congested state, if the Forward RM Cell Detector detects a forward RM cell, it requests the Backward RM Cell Generator to generate a backward RM cell. The RM Cell Generator then generates a backward RM cell indicating the existence of congestion and sends it toward the source end-station via UPC(i)-d. The frequency of backward RM cell generation should follow the rule defined for switches that generate backward RM cells [1]. Here, if the UPC(i)-d detects a backward RM cell returned by the destination end-station, it would remove it from the connection.

When a backward RM cell with CI=1 passes through the UPC(e)-s, the Backward RM Cell Detector detects it and notices that the route of the ATM connection specified by the RM cell is going to fall into (or is in) a congested state. To avert congestion of the network, the RM Cell Detector informs the Policing Parameter Calculator in UPC(i)-s of the existence of the congestion, and the Calculator modifies the policing parameter with some defined (i.e., predefined or established by user-network negotiation) back-off interval. The back-off interval is important in preventing the mis-policing of incoming cell flow by modification of the policing parameter at the Calculator earlier than the completion of congestion control at the source end-station. When the RM Cell Detector informs the Calculator of the detection of the RM cell, the Calculator calculates the new policing parameter in order to regulate the allowable traffic volume submitted to the network according to the congestion level. The modified policing parameter is transmitted to the Policing Module in UPC(i)-s, after the back-off interval has expired. In order to inform the source end-station of the new policing parameter, the Backward RM Cell Detector replaces the ER value in the backward RM cell with the new policing parameter in the Policing Module and sends it to the source end-station. The source end-station will modify its cell transmission rate according to the ER value in the received backward RM cell. When the back-off interval expires after the backward RM cell is sent by the Dynamic UPC, the new policing parameter is used to monitor the incoming cell flow.

When the switch recovers from the congested state as a result of the above control, the EFCI is not marked and the backward RM cell with CI=1 is not sent back from the destination end-station (or UPC(e)-d). Therefore the UPC(e)-s knows that the network is not in a congested state any more. Then, the UPC(i)-s changes the policing parameter for the related connection to relax the restrictions on the amount of cells submitted to the network, and the UPC(e)-s indicates the modified policing parameter by using the ER value in the backward RM cell sent to the source end-station. The Dynamic UPC monitors the incoming cell flow based on this modified policing parameter.

3.2.2.2 Operation for ER-based Network

When a switch is going to fall into (or is in) a congested state, it reduces the ER value in the forward and/or the backward RM cells passing through it. When a backward RM cell passes through the UPC(e)-s, the Backward RM Cell Detector detects the ER value in the received RM cell and sends the cell to the source end-station. The RM Cell Detector informs the Policing Parameter Calculator in UPC(i)-s of the ER value, and when the back-off interval (described above) expires after the backward RM cell is sent by the UPC(e)-s, the Calculator modifies the policing parameter to the ER value. Even if the backward RM cell is not transmitted to the source end-station transparently, the source end-station can harmonize its cell transmission rate with the policing parameter at UPC(i)-s using the ER value in the backward RM cell.

3.2.3 Additional Action of Dynamic NPC

In the case of long distance connections without VDs/VSs (Virtual Destinations/Sources) (Fig. 1 b) and c)), the propagation delay of control messages and forward/backward RM cells are crucial factors in congestion control. With the proposed congestion control scheme, however, the backward RM cells are generated more promptly by the NPC(e) attached to the congested network according to the reception of forward RM cells or EFCI marked cells. Therefore, the congestion control scheme would respond to network congestion more quickly for long distance ATM connections, even if VDs/VSs were present along the path.

3.2.4 Advantages of Dynamic UPC/NPC

According to the current Dynamic GCRA specification proposed in the ATM Forum[1], the Dynamic GCRA may not be applied to the UPC/NPC in EFCI-based network. According to the proposed Dynamic UPC/NPC scheme, on the other hand, the network can regulate input traffic by itself, regardless of the switch behavior. Since the proposed Dynamic UPC/NPC can modify the policing parameter according to the CI bit and replace the ER value with it, the network can regulate input traffic with arbitrary rate control policy even when the switches in the network operate according to EFCI-based behavior. The network segment applying the proposed Dynamic NPC can inform the source end-station and/or the neighboring network segment(s) of the congestion level by using the ER value, even when the EFCI-based network segment is in a congested state. As a result, regardless of the switch behavior, the network which applies the Dynamic UPC/NPC can recover from the congested state securely and control the input traffic flexibly.

The proposed Dynamic UPC/NPC scheme can decide the control policy, which is memorized in Memory Module in Fig. 2, for each ATM connection individually, if necessary. For example, the emergent traffic and/or the traffic with high QoS priority may not be regulated in the case of light congestion. Finally, the proposed Dynamic UPC/NPC scheme would respond to the network congestion more quickly for long distance ATM connections, since the Dynamic UPC/NPC could immediately regulate the input traffic when it receives the backward RM cell indicating the congestion.

4. Performance Evaluation

In this section, the performance of the proposed congestion control scheme is evaluated. We evaluated the following two cases, representing two of main objectives of the proposed mechanism.

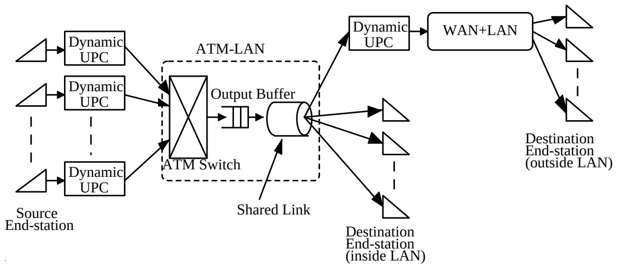


Fig. 3 Performance Evaluation Model (1)

- **Secure operation**, even if non-compliant source end-stations are connected to the network, and
- **Fast operation with long distance connections**, even if VDs/VSs do not exist and switches do not generate backward RM cells along any given long-distance VCC.

In this paper, we adopted the cell loss ratio and the throughput (at the network layer level) as measures of network performance for ABR connections. This is because it is considered necessary for the ABR service to provide a sufficiently small cell loss ratio for those connections. Here we assumed that switches in the network operate according to EFCI-based behavior, which implies that they can only set EFCI=1 of arrived cells to indicate congestion, and that the switches cannot inform the source end-stations of congestion directly according to the CI bit in backward RM cells. In other words, we chose switches having only basic functions to evaluate the worst case scenario of the network operation. We also assumed that no VSs and VDs were present in any of the connections.

4.1 Simulation Model

Figure 3 shows a simulation model. In this model, ten source end-stations are connected to an ATM-LAN with 150Mbps transmission lines. These source end-stations generate frames, which are service data units (SDUs, i.e., IP packets) at higher layer (i.e., network layer) and are divided into ATM cells. It is assumed that the length of each frame is equivalent to that of 170 cells, which corresponds to about 8KByte (i.e., the size of a page in modern computer systems or close to the default MTU (Maximum Transmission Unit) size defined in [10]). It is also assumed that the frame generation in each end-station follows the Poisson arrival process with an average arrival rate equal to 9Mbps, which corresponds to about 140 frames/sec. When there is any frame in a transmitting buffer at the source end-station, it divides the frame into cells and sends them to the ATM-LAN. The transmitting buffer at the end-station is assumed to store infinite frames. It is assumed that each end-station uses an ABR service connection to transmit these cells, and that the PCR value of each connection is set to 150Mbps. Therefore in this simulation, unless the ATM network is going to fall into (or is in) a congested state, UPC sets a policing parameter to the PCR value and monitors cell flow based on the parameter. It is also assumed that each source end-station sends a forward RM cell to the destination end-station for every 16 cells (i.e., Nrm=16).

An ATM switch has one output buffer which applies a simple FIFO (First In First Out) queuing policy, and determines whether or not it is going to fall into (or is in) a congested state by monitoring the queue length of its buffer. The buffer size of the switch is 256 cells, and the threshold value of the queue length to determine that the switch is going to fall into a congested state is 96 cells (excluding the evaluation in Fig. 6 which parameterized the threshold value). If the queue length is equal to or larger than the threshold value, the switch judges that it is going to fall into (or is in) a congested state. If the queue length is less than the threshold value, the switch judges that it is in a normal state. When the queue length is equal to the buffer size, arriving cells are discarded due to buffer overflow.

As described in section 2.1, when an EFCI-based switch detects congestion, it sets the EFCIs of arrived cells, and when the destination end-station receives a cell with EFCI=1, it sets CI=1 in the backward RM cell and returns it to the source end-station. When the backward RM cell with CI=1 is received by the Dynamic UPC attached to the source end-station, it reduces the policing parameter (normally set to PCR) according to the predefined procedure after the back-off interval has expired. When a backward RM cell with CI=0 is received by the UPC, it increases the policing parameter after the back-off interval has expired. In this simulation, the policing parameter policed by the UPC is reduced to half of its previous value whenever the UPC receives a backward RM cell with CI=1. Here, the policing parameter will never be smaller than MCR (in this simulation, MCR=0). In a normal state, the policing parameter is doubled whenever the UPC receives a backward RM cell with CI=0, until the parameter for UPC equals PCR. Before the UPC modifies the policing parameter (i.e., during the back-off interval), the UPC informs the source end-station of the modified policing parameter by overwriting the ER value in the backward RM cell, and if the source end-station is compliant, it modifies the cell transmission rate according to the ER value.

4.2 Simulation Results

4.2.1 Evaluation of Secured Operation

Figure 4 shows the cell loss ratio for the cell flow of compliant and non-compliant end-stations (including cell loss either at the UPC or at the output buffer in a switch) versus percentage of non-compliant end-

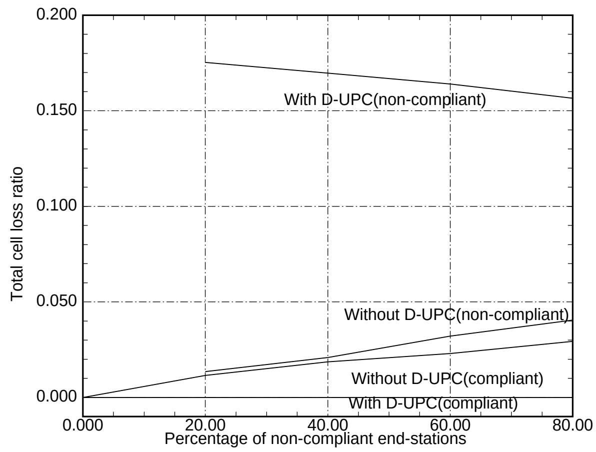


Fig. 4 Total Cell Loss Ratio versus Percentage of Non-compliant End-stations

stations. Figure 5 shows the throughput of these end-stations versus the percentage of non-compliant end-stations. For example, when 20% of the end-stations are non-compliant, one fifth of all the end-stations do not modify their cell transmission rates even when they receive ER values in backward RM cells which indicate the appropriate cell transmission rate. In this simulation model, it is assumed that there are only local ATM connections, whose destination end-stations are within the same ATM-LAN segment as the source end-station. It is also assumed that the physical length between the source end-station and the destination end-station (inside-LAN) is 0.56km and the propagation delay is about $2.8\mu s$, which corresponds to the time required for transmitting one cell to the 150Mbps line. In this paper, we define the *throughput* as a ratio of cells received by an higher layer at the destination end-station to cells transmitted by an higher layer at the source end-station.

Without the proposed Dynamic UPC, the cell loss ratio becomes large and the throughput becomes small when the percentage of non-compliant end-stations becomes large. Here, cell loss occurs only at the output buffer in the switch. Since non-compliant end-stations do not reduce the cell transmission rate even in a congested state, the cell loss ratio for the cell flow of compliant end-stations becomes large and the throughput for compliant as well as non-compliant end-station becomes small.

In the case of the proposed Dynamic UPC, the cell loss ratio for compliant end-stations becomes sufficiently small (as can be seen in Fig. 4, the cell loss ratio becomes zero in this simulation), and the throughput for compliant end-stations becomes large (as can be seen in Fig. 5, the throughput becomes 1.0 in this simulation). This is because the proposed Dynamic UPC can reduce the amount of cells entering ATM networks

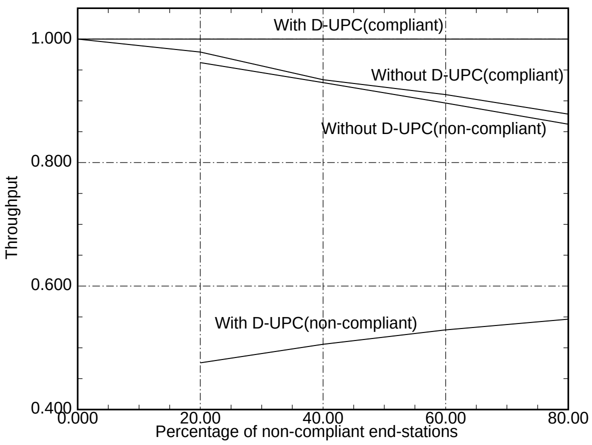


Fig. 5 Throughput versus Percentage of Non-compliant End-stations

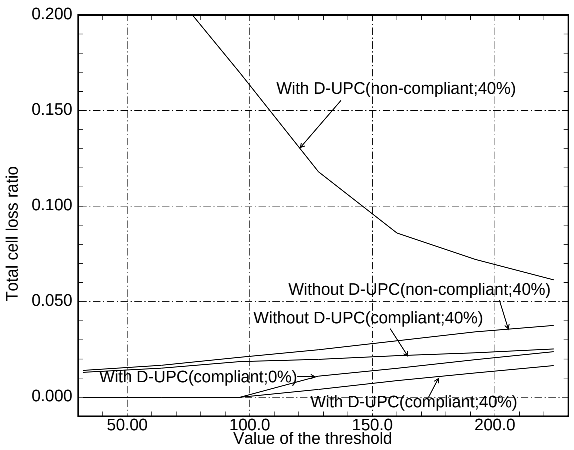


Fig. 6 Total Cell Loss Ratio versus Threshold of Congestion Judgement (buffer size = 256 cells)

in the congested state in a secure manner. The number of cells which overflow at the output buffer also becomes sufficiently small. In Fig. 4, we can see that the cell loss ratio for the non-compliant end-stations becomes large for the proposed UPC, and in Fig. 5, we can see that the throughput for the non-compliant end-stations becomes small. Cell loss occurs only at the UPC point, since non-compliant end-stations do not reduce the cell transmission rate. In addition, the output buffer will be protected from cells from the non-compliant end-stations. As emphasized earlier, the proposed congestion control scheme can provide the desired services for compliant end-stations in a secure manner.

Now, the number of frames stored really in a frame level buffer at each compliant end-station will be considered. According to this simulation, the maximum number of frames stored in the frame level buffer at a compliant end-station was 4 frames, while that at a non-compliant end-station was 2 frames (these values are the same as in the case without Dynamic UPC). This difference (2 frames in this simulation) is thought to be due to the restriction of the cell transmission rate at the compliant end-station reacting to the backward RM cells. We could therefore estimate that, even when flow control is performed at the end-station, the increase in the required frame level buffer size would be 2 frames, and the increase in the queuing delay (i.e., latency) would be a maximum of $1ms$ ($\simeq 2.8\mu s$/cell $\times$ 170 cells $\times$ 2 frames). According to the above estimate, the buffer size ($\simeq$ 4frames $\simeq$ 32KB) and the queuing delay ($\simeq 1ms$) would not become too large even if the compliant end-station reduces the cell transmission rate.

Finally, the threshold value at the output buffer in the switch which determines whether or not the switch is in a congested state will be discussed. Figure 6 shows the total cell loss ratio versus the threshold value. The percentage shown in Fig. 6 indicates the percentage of non-compliant end-stations. Since the congestion control scheme regulates traffic in an overly stringent manner when the threshold value is small, the network cannot provide enough throughput. This is because the cell transmission rate of the source end-station becomes small compared with the transmission line rate between the source end-station and the network. Therefore, when the network applies the proposed control scheme, the cell loss ratio for non-compliant end-stations becomes large due to the discarding of cells at UPC(i). Since the congestion control scheme does not regulate traffic well when the threshold value is too large and close to the actual buffer size, the number of cells discarded at the output buffer in the switch will become large. Therefore, when the network applies the proposed control scheme, the cell loss ratio for the compliant end-stations becomes large (the ratio is, however, smaller than that without the proposed Dynamic UPC), and the scheme will not be able to provide the desired service quality. It is important to establish a threshold value which enables a high performance level to be achieved by the proposed control scheme. The calculation method for the optimum threshold value needs further study. It could, however, also just be an engineering decision taken during operation. As shown in Fig. 6, the suitable threshold value for compliant end-stations in this evaluation model seems to be around 100 cells under the condition of no cell loss.

4.2.2 Evaluation of Fast Operation with Long Distance Connections

In this subsection, the performance of the proposed congestion control scheme is evaluated by applying it to long distance connections (Fig. 1 b) and c)). Here, we assumed the network operation to be as follows;

- EFCI-based behavior (not ER-based behavior),

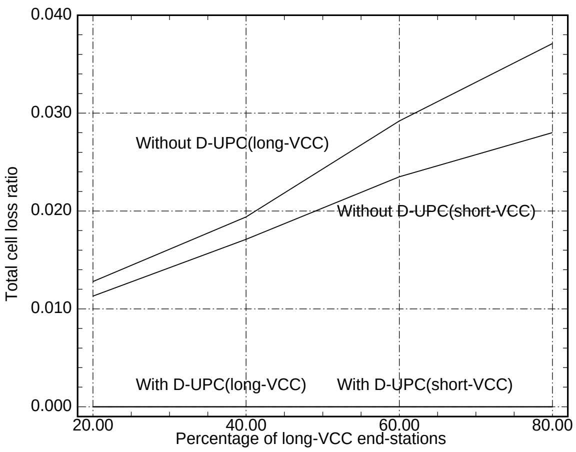


Fig. 7 Total Cell Loss Ratio versus Percentage of Long Distance VCCs (in case of Fig. 1-b)

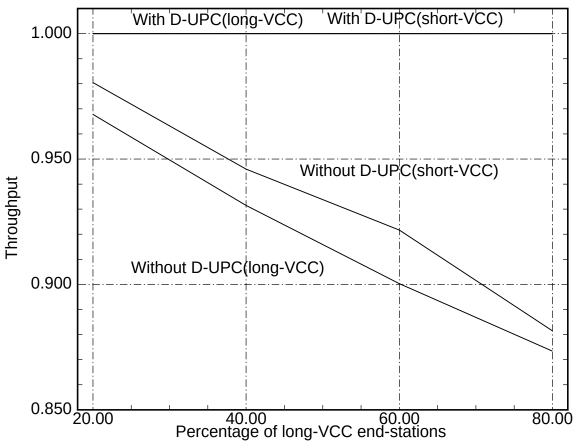


Fig. 8 Throughput versus Percentage of Long Distance VCCs (in case of Fig. 1-b)

- no VDs/VSs along the connections,
- no switch that can generate backward RM cells.

In other words, to evaluate the worst case scenario for control performance, we assume that the network had only basic functions.

Firstly, we evaluated the performance of the proposed congestion control scheme shown in Fig. 1 b). The performance evaluation model is shown in Fig. 3. Figure 7 shows the total cell loss ratio versus the percentage of long distance VCCs, and Figure 8 shows the throughput of each end-station versus the percentage of long distance VCCs. Here, a long distance VCC means that the destination end-station of the VCC is located outside of the ATM-LAN segment that accommodates the source end-stations. A short distance VCC means that the destination end-station of the VCC is located at the ATM-LAN segment where the source end-station exists. In this simulation model, it is assumed that the propagation delay of short distance VCCs is $2.8\mu s$,

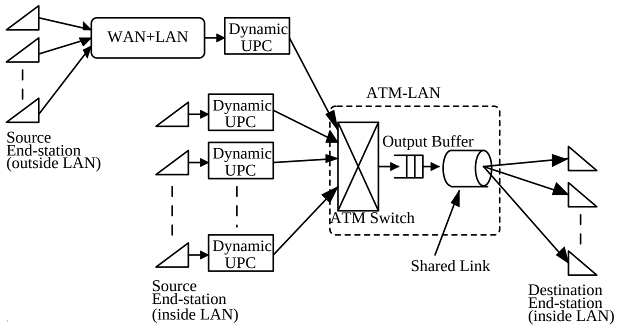


Fig. 9 Performance Evaluation Model (2)

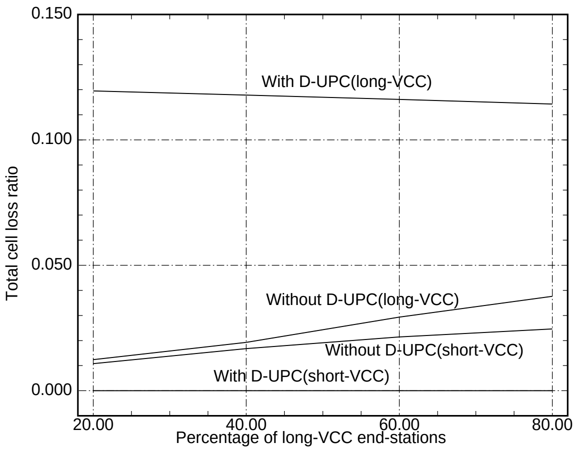


Fig. 10 Total Cell Loss Ratio versus Percentage of Long Distance VCCs (in case of Fig. 1-c)

while that of long distance VCCs is $420\mu s$. This means that the length of a long distance VCC corresponds to about 84km. Here, in this simulation, only compliant end-stations are assumed to be connected to the network.

By applying the proposed Dynamic UPC/NPC, we can attain better characteristics with regard to both the cell loss ratio and the throughput regardless of the percentage of long distance VCCs. The cell loss ratio for both short and long distance VCCs remains zero, and the throughput for both VCCs remains 1.0. The decreased cell loss ratio was expected as it is a result of the decrease in the effective propagation delays of the forward and backward RM cell transmission loops, which is due to the operation of the proposed Dynamic NPC, located at the egress point of the ATM-LAN.

Secondly, performance of the proposed congestion control scheme shown in Fig. 1 c) is evaluated. A performance evaluation model is shown in Fig. 9. Figure 10 and Figure 11 show the total cell loss ratio versus percentage of long distance VCCs, and the throughput of each end-station versus the percentage of long distance VCCs respectively. Here, a long distance VCC means that the source end-station of the VCC is lo-

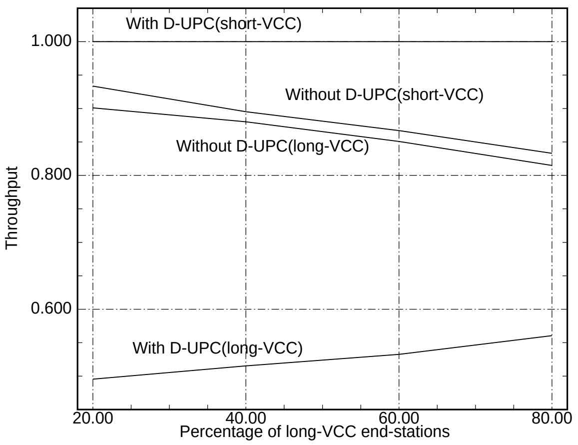


Fig. 11 Throughput versus Percentage of Long Distance VCCs (in case of Fig. 1-c)

cated outside the ATM-LAN segment that accommodates the destination end-stations. Likewise the simulation model shown in Fig. 3, in this simulation model, it is assumed that the propagation delay of short distance VCCs is $2.8\mu s$, while that of long distance VCCs is $420\mu s$. Only compliant end-stations are assumed to be connected to the network. Here, for long distance VCCs, it reduces the policing parameter in a moment (not waiting for the expiration of some back-off interval) when the backward RM cell is received by the Dynamic NPC.

By applying the proposed Dynamic UPC/NPC, the cell loss ratio for short distance VCCs remains zero, regardless of the percentage of long distance VCCs, and the throughput for short distance VCCs remains 1.0. For long distance VCCs, however, we cannot attain better characteristics with regard to both the cell loss ratio and the throughput. This is because the Dynamic NPC reduces the policing parameter as soon as it receives a backward RM cell with CI=1. Since the Dynamic NPC reduces the amount of cells entering the network in the congested state before the source end-station of the long distance VCC reduces its cell transmission rate, cell loss occurs at the NPC point (not at the output buffer in the switch). Therefore, the proposed congestion control scheme can provide the desired services only for short distance VCCs.

In order to reduce the cell loss ratio for long distance VCCs, the Dynamic NPC point will need a buffer space enough to store the cells that arrive after the reduction of the policing parameter until the source end-station (or neighboring Dynamic NPC point) reduces its cell transmission rate. To achieve sufficiently small cell loss ratio for the long distance VCCs, the required buffer space will be the same order as that for VS/VD point. When switches in the network operate according to EFCI-based behavior that cannot modify the ER value in the backward RM cell, the network segment which applies VS/VD cannot inform outside the segment of the congestion level. Therefore, the cell transmission rate is calculated by only the VS according to VS behavior. However, the network segment which applies the Dynamic NPC can inform outside the segment of the congestion level by replacing the ER value with the new policing parameter which is calculated based on CI bit, even if the switches cannot modify the ER value. In addition, the Dynamic NPC can decide the control policy for ATM VCs individually if necessary. Therefore, while the required buffer space for the Dynamic NPC will be equivalent to it for the VS/VD, the Dynamic NPC scheme may provide higher control flexibility independently of the switch behavior.

5. Conclusion

We have proposed a novel reactive congestion control scheme using Dynamic UPC/NPC. The UPC/NPC policing parameter(s) is modified according to the state of the ATM network such that the traffic volume submitted to a congested network is regulated securely. The proposed control scheme prevents the network from falling into (or enables it to recover from) a congested state irrespective of the cooperation of end-stations. We verified this for the case of non-compliant end-stations using computer simulations. The simulations also showed that proposed scheme works well even if long distance ATM connections exist in the network.

Acknowledgement

The authors would like to thank Dr. T. Kamitake and Mr. Y. Katsube of Toshiba Corporation for their helpful comments.

References

[1] Sathaye, S.S., "ATM Forum Traffic Management Specification Version 4.0," ATM Forum 95-0013R6, Jun. 1995.
[2] Newman, P., "Update on BECN Congestion Control," ATM Forum 93-855, Jun. 1993.
[3] Hsiaw, H., Fendick, K., Bonomi, F., Roberts, L., Jain, A., Makrucki, B., Tofigh, T., Barnhart, A.W., Gun, L., Holden, B., Fedorkow, G., Daigle, J.N., Kline, D., Suzuki, H., Ramamurthy, G., Newman, P., Giroux, N., Kositpaiboon, R., Hughes, D., Hluchyj, M.G., Garg, G., Siu, K.Y., Tzeng, H.Y. and Yin, N., "Closed-Loop Rate-Based Traffic Management," ATM Forum 94-438R2, Sep. 1994.
[4] Bonomi, F. and Fendric, K.W., "The Rate-Based Flow Control Framework for the Available Bit Rate ATM Service," *IEEE Network,* vol.9, no.2, pp.25–39, Mar./Apr. 1995.
[5] Kung, H.T. and Morris, R., "Credit-Based Flow Control for ATM Networks," *IEEE Network,* vol. 9, no. 2, pp.40–48, Mar./Apr. 1995.
[6] Yamato, K. and Esaki, H., "Congestion Control using UPC/NPC Function," ATM Forum 94-338, May 1994.
[7] Yamato, K. and Esaki, H., "Reactive Congestion Control by Dynamic UPC/NPC," ATM Forum 94-561, Jul. 1994.

[8] ITU-T Recommendation I.371, *Traffic Control and Congestion Control in B-ISDN,* Nov. 1994.
[9] "ATM User-Network Interface Specification Version 3.1," ATM Forum, Sep. 1994.
[10] Laubach, M., "Classical IP and ARP over ATM," RFC 1577, Jan. 1994.

Appendix A: List of Abbreviations

ABR	Available Bit Rate
ACR	Allowed Cell Rate
BECN	Backward Explicit Congestion Notification
CAC	Connection Admission Control
CBR	Constant Bit Rate
CDV	Cell Delay Variation
CI	Congestion Indication
CLR	Cell Loss Rate
CTD	Cell Transfer Delay
EFCI	Explicit Forward Congestion Indication
ER	Explicit Rate
MCR	Minimum Cell Rate
NPC	Network Parameter Control
PCR	Peak Cell Rate
QoS	Quality of Service
RM cell	Resource Management cell
UPC	Usage Parameter Control
VBR	Variable Bit Rate
VCC/VPC	Virtual Channel/Path Connection
VCI/VPI	Virtual Channel/Path Identifier
VD/VS	Virtual Destination/Source

Katsumi Yamato received the B.E. and M.E. degrees from Osaka University (Japan) in 1988 and 1990, respectively. He joined R&D Center of Toshiba Corporation in 1990, and has been engaged in research on traffic control for ATM network. Mr. Yamato received the paper award from IEICE (Japan) in 1990.

Hiroshi Esaki received the B.S. and M.S. degrees from Kyushu University, Fukuoka, Japan, in 1985 and 1987, respectively. In 1987, he joined Research and Development Center, Toshiba Corporation, where he engaged in the research of ATM systems. He has been at Bellcore in New Jersey (USA) as a residential researcher from 1990 to 1991, and has engaged in the research on high speed computer communications. From 1994, he has been at CTR (Center for Telecommunications Research) of Columbia University in New York (USA) as a visiting scholar. He is currently interested in an internet architecture using ATM technology and a mobile computing.