

ATM Congestion Control

By Fang Lu (flu@cis.ohio-state.edu)

ABSTRACT: The concepts in congestion control for ATM networks are explained. The specifications for ATM traffic control proposed by ATM Forum are presented. Some representative schemes are described and compared.

Table of Contents

- [Introduction](#)
 - [How ATM Works](#)
 - [Why Need Congestion Control](#)
- [Connection Parameters](#)
 - [Quality of Services](#)
 - [Usage Parameters](#)
- [Service Categories](#)
- [What is Expected from Congestion Control](#)
 - [Objectives](#)
 - [Selection Criteria](#)
- [Generic Functions](#)
 - [Connection Admission Control](#)
 - [Usage Parameter Control](#)
 - [Priority Control](#)
 - [Traffic Shaping](#)
 - [Network Resource Management](#)
 - [Frame Discard](#)
 - [Feedback Control](#)
 - [ABR Flow Control](#)

Introduction

The future telecommunication should have such characteristics: broadband, multimedia, economical implementation for diversity of services [15] Broadband integrated services digital networks (B-ISDN) provides what we need. Asynchronous Transfer Mode (ATM) is a target technology for meeting these

requirements.

In ATM networks, the information is transmitted using short fixed-length cells, which reduces the delay variance, making it suitable for integrated traffic consisting of voice, video and data. By proper traffic management, ATM can also ensure efficient operation to meet different quality of service (QoS) desired by different types of traffic.

How ATM Works

1. ATM network uses fixed-length cells to transmit information. The cell consists of 48 bytes of payload and 5 bytes of header. The flexibility needed to support variable transmission rates is provided by transmitting the necessary number of cells per unit time.
2. ATM network is connection-oriented. It sets up virtual channel connection (VCC) going through one or more virtual paths (VP) and virtual channels (VC) before transmitting information. The cells is switched according to the VP or VC identifier (VPI/VCI) value in the cell head, which is originally set at the connection setup and is translated into new VPI/VCI value while the cell passes each switch.
3. ATM resources such as bandwidth and buffers are shared among users, they are allocated to the user only when they have something to transmit. So the network uses statistical multiplexing to improve the effective throughput.

Why Need Congestion Control

The assumption that statistical multiplexing can be used to improve the link utilization is that the users do not take their peak rate values simultaneously. But since the traffic demands are stochastic and cannot be predicted, congestion is unavoidable. Whenever the total input rate is greater than the output link capacity, congestion happens. Under a congestion situation, the queue length may become very large in a short time, resulting in buffer overflow and cell loss. So congestion control is necessary to ensure that users get the negotiated QoS.

There are several misunderstandings about the cause and the solutions of congestion control[6]

1. Congestion is caused by the shortage of buffer space. The problem will be solved when the cost of memory becomes cheap enough to allow very large memory.

Larger buffer is useful only for very short term congestions and will cause undesirable long delays. Suppose the total input rate of a switch is 1Mbps and the capacity of the output link is 0.5Mbps, the buffer will overflow after 16 second with 1Mbyte memory and will also overflow after 1 hour with 225Mbyte memory if the situation persists. Thus larger buffer size can only postpone the discarding of cells but cannot prevent it. The long queue and long delay introduced by large memory is undesirable for some applications.

2. Congestion is caused by slow links. The problem will be solved when high-speed links become available.

It is not always the case, sometimes increases in link bandwidth can aggravate the congestion problem because higher speed links may make the network more unbalanced. For the configuration showed in the Figure 1, if both of the two sources begin to send to destination 1 at

their peak rate, congestion will occur at the switch. Higher speed links can make the congestion condition in the switch worse.

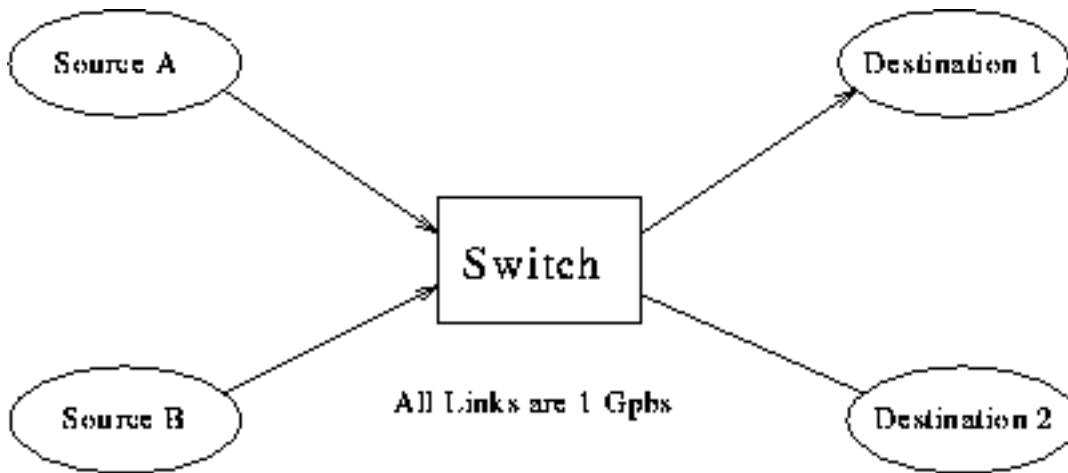


Figure 1: A Network with High Speed Links

3. Congestion is caused by slow processors. The problem will be solved when processor speed is improved.

This statement can be explained to be wrong similarly to the second one. Faster processors will transmit more data in unit time. If several nodes begin to transmit to one destination simultaneously at their peak rate, the target will be overwhelmed soon.

Congestion is a dynamic problem, any static solutions are not sufficient to solve the problem. All the issues presented above: buffer shortage, slow link, slow processor are symptoms not the causes of congestion. Proper congestion management mechanisms is more important than ever.

[Return to the Table of Contents](#)

Connection Parameters

Quality of Service

A set of parameters are negotiated when a connection is set up on ATM networks. These parameters are used to measure the Quality of Service (QoS) of a connection and quantify end-to-end network performance at ATM layer. The network should guarantee the QoS by meet certain values of these parameters.

- Cell Transfer Delay (CTD)}:

The delay experienced by a cell between the first bit of the cell is transmitted by the source and the last bit of the cell is received by the destination. Maximum Cell Transfer Delay (Max CTD) and Mean Cell Transfer Delay (Mean CTD) are used.

- Peak-to-peak Cell Delay Variation (CDV)}:

The difference of the maximum and minumum CTD experienced during the connection.

Peak-to-peak CDV and Instantaneous CDV are used.

- Cell Loss Ratio (CLR)}:

The percentage of cells that are lost in the network due to error or congestion and are not received by the destination.

Usage Parameters

Another set of parameters are also negotiated when a connection is set up. These parameters discipline the behavior of the user. The network only provides the QoS for the cells that do not violate these specifications.

- Peak Cell Rate (PCR):

The maximum instantaneous rate at which the user will transmit.

- Sustained Cell Rate (SCR):

The average rate as measured over a long interval.

- Burst Tolerance (BT):

The maximum burst size that can be sent at the peak rate.

- Maximum Burst Size (MBS):

The maximum number of back-to-back cells that can be sent at the peak cell rate.

BT and MBS are related as follows: $\text{Burst Tolerance} = (\text{MBS} - 1) \left(\frac{1}{\text{SCR}} - \frac{1}{\text{PCR}} \right)$

- Minimum Cell Rate (MCR):

The minimum cell rate desired by a user.

[Return to the Table of Contents](#)

Service Categories

Providing desired QoS for different applications is very complex. For example, voice is delay-sensitive but not loss-sensitive, data is loss-sensitive but not delay-sensitive, while some other applications may be both delay-sensitive and loss-sensitive.

To make it easier to manage, the traffic in ATM is divided into five service classes [\[16\]](#):

- CBR: Constant Bit Rate

Quality requirements: constant cell rate, i.e. CTD and CDV are tightly constrained; low CLR.

Example applications: interactive video and audio.

- rt-VBR: Real-Time Variable Bit Rate

Quality requirements: variable cell rate, with CTD and CDV are tightly constrained; a small nonzero random cell loss is possible as the result of using statistical multiplexing.

Example applications: interactive compressed video.

- nrt-VBR: Non-Real-Time Variable Bit Rate

Quality requirements: variable cell rate, with only CTD are tightly constrained; a small nonzero random cell loss is possible as the result of using statistical multiplexing.

Example applications: response time critical transaction processing.

- UBR: Unspecified Bit Rate

Quality requirements: using any left-over capacity, no CTD or CDV or CLR constrained.

Example applications: email and news feed.

- ABR: Available Bit Rate

Quality requirements: using the capacity of the network when available and controlling the source rate by feedback to minimize CTD , CDV and CLR.

Example applications: critical data transfer, remote procedure call and distributed file service.

These service categories relate traffic characteristics and QoS requirements to network behavior. The QoS requirement for each class is different. The traffic management policy for them are different, too.

The QoS and Usage parameters for these classes are summarized in Table 1:

Parameters	CBR	rt-VBR	nrt-VBR	UBR	ABR
PCR and CDVT(5)	specified	specified	specified	specified(3)	specified(4)
SCR, MBS, CDVT(5,6)	n/a	specified	specified	n/a	n/a
MCR(5)	n/a	n/a	n/a	n/a	specified
Peak-to-peak CDV	specified	specified	unspecified	unspecified	unspecified
Mean CTD	unspecified	unspecified	specified	unspecified	unspecified
Max CTD	specified	specified	unspecified	unspecified	unspecified
CLR(5)	specified(1)	specified(1)	specified(1)	unspecified	specified(2)

Table 1: ATM Service Categories

Notes:

1. The CLR may be unspecified for CLP=1.
2. Minimized for sources that adjust cell flow in response to control information.
3. May not be subject to CAC and UPC procedures.
4. Represents the maximum rate at which the source can send as controlled by the control information.
5. These parameters are either explicitly or implicitly specified for PVCs or SVCs.
6. Different values of CDVT may be specified for SCR and PCR.

Among these service classes, ABR is commonly used for data transmissions which require a guaranteed QoS, such as low probability of loss and error. Small delay is also required for some application, but is not as strict as the requirement of loss and error. Due to the burstiness, unpredictability and huge amount of the data traffic, congestion control of this class is the most needed and is also the most studied.

ATM Forum Technical Committee specified the feedback mechanism for ABR flow control. We will discuss it in more detail later.

What is Expected from Congestion Control

Objectives

The objectives of traffic control and congestion control for ATM are: Support a set of QoS parameters and classes for all ATM services and minimize network and end-system complexity while maximizing network utilization.

Selection Criteria

To design a congestion control scheme is appropriate for ATM network and non-ATM networks as well, the following guidances are of general interest [\[5\]](#).

- **Scalability**

The scheme should not be limited to a particular range of speed, distance, number of switches, or number of VCs. The scheme should be applicable for both local area networks (LAN) and wide area networks (WAN).

- **Fairness**

In a shared environment, the throughput for a source depends upon the demands by other sources. There are several proposed criterion for what is the correct share of bandwidth for a source in a network environment. And there are ways to evaluate a bandwidth allocation scheme by comparing its results with a optimal result.

- **Fairness Criteria** [\[16\]](#)

1. Max-Min

The available bandwidth is equally shared among connections.

2. MCR plus equal share

The bandwidth allocation for a connection is its MCR plus equal share of the available bandwidth with used MCR removed.

3. Maximum of MCR or Max-Min share

The bandwidth allocation for a connection is its MCR or Max-Min share, which ever is larger.

4. Allocation proportional to MCR

The bandwidth allocation for a connection is weighted proportional to its MCR.

5. Weighted allocation

The bandwidth allocation for a connection is proportional to its pre-determined weight.

- **Fairness Index**

The share of bandwidth for each source should be equal to or converge to the optimal value

according to some optimality criterion. We can estimate the fairness of a certain scheme numerically as follows. Suppose a scheme allocates x_1, x_2, \dots, x_n , while the optimal allocation is y_1, y_2, \dots, y_n . The normalized allocation is $z_i = x_i / y_i$ for each source and the fairness index is defined as following:

$$\text{Fairness} = \frac{\sum(z_i) * \sum(z_i)}{\sum(z_i * z_i)}$$

- **Robustness**

The scheme should be insensitive to minor deviations such as slight mistuning of parameters or loss of control messages. It should also isolate misbehaving users and protect other users from them.

- **Implementability**

The scheme should not dictate a particular switch architecture. It also should not be too complex both in term of time or space it uses.

[Return to the Table of Contents](#)

Generic Functions

It is observed that events responsible for congestion in broadband networks have time constants that differ by orders of magnitude, and multiple controls with appropriate time constants are necessary to manage network congestion[14].

We can classify the congestion control schemes by the time scale they operate upon[12]: network design, connection admission control (CAC), routing (static or dynamic), traffic shaping, end-to-end feedback control, hop-by-hop feedback control, buffering. The different schemes are functions on different severity of congestion as well as different duration of congestion[5].

Another classification of congestion control schemes is by the stage that the operation is performed: congestion prevention, congestion avoidance and congestion recovery. Congestion prevention is the method that make congestion impossible. Congestion avoidance is that the congestion may happen, but the method avoid it by get the network state always in balance. Congestion recovery is the remedy steps to take to pull the system out of the congestion state as soon as possible and make it less damaging when the congestion already happened.

No matter what kind of scheme is used, the following outstanding problems are the main difficulties that need to be treated carefully: the burstiness of the data traffic, the unpredictability of the resource demand and the large propagation delay versus the large bandwidth.

To meet the objectives of traffic control and congestion control in ATM networks, the following functions and procedures are suggested by the ATM Forum Technical Committee[16].

Connection Admission Control

Connection Admission Control (CAC) is defined as the set of actions taken by the network during the call set-up phase in order to determine whether a connection request can be accepted or should be rejected.

Based on the CAC algorithm, a connection request is progressed only when sufficient resources such as bandwidth and buffer space are available along the path of a connection. The decision is made based on the service category, QoS desired and the state of the network which means that the number and conditions of existing connections.

Routing and resource allocation are part of CAC when a call is accepted.

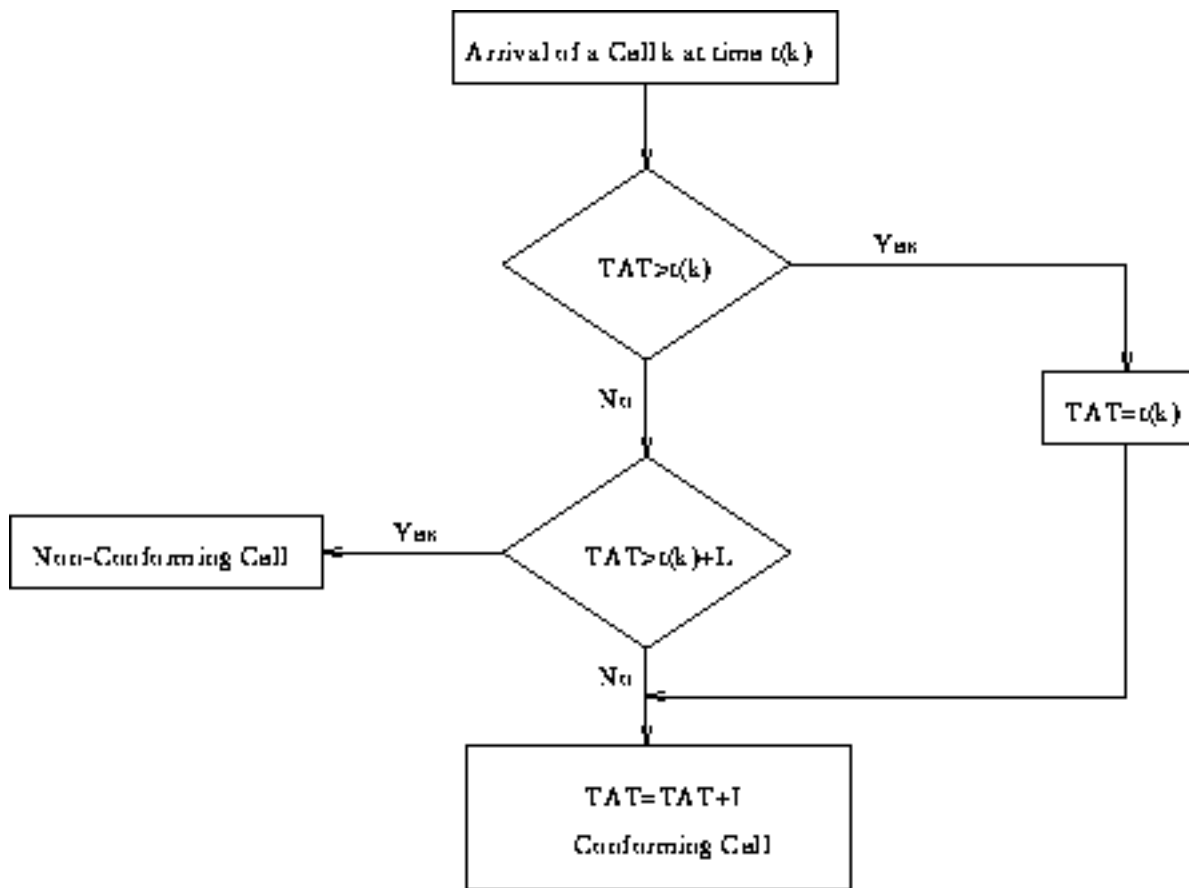
Usage Parameter Control

Usage Parameter Control (UPC) is defined as the set of actions taken by the network to monitor and control traffic at the end-system access. Its main purpose is to protect network resources from user misbehavior, which can affect the QoS of other connections, by detecting violations of negotiated parameters and taking appropriate actions.

Generic Cell Rate Algorithm

The Generic Cell Rate Algorithm (GCRA) is used to define conformance with respect to the traffic contract. For each cell arrival, the GCRA determines whether the cells conform to traffic contract of the connection. The UPC function may implement GCRA, or one or more equivalent algorithms to enforce conformance.

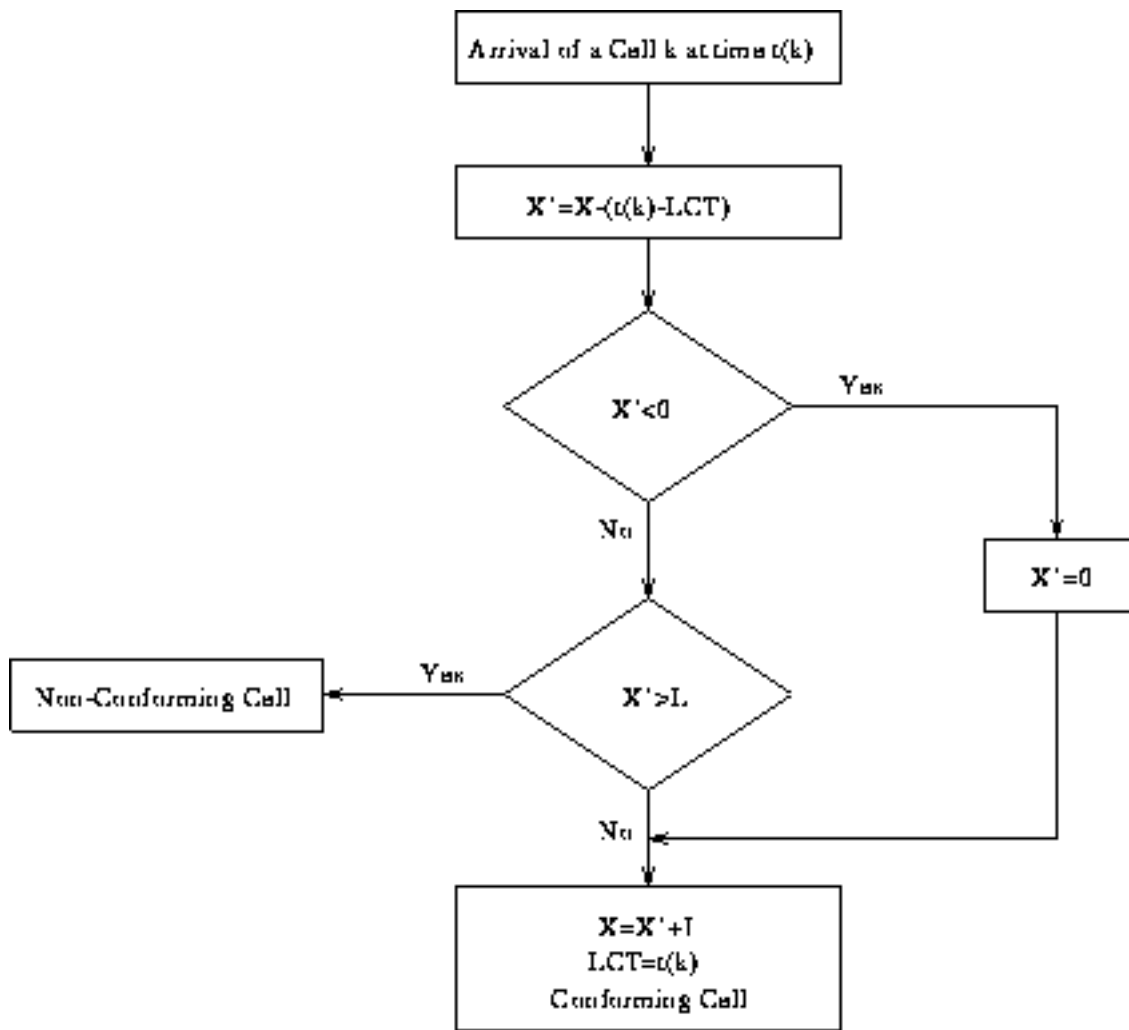
GCRA is a virtual scheduling algorithm or a continuous-state Leaky Bucket Algorithm as defined by the flowchart in Figure 2 and Figure 3. It is defined with two parameters: the Increment (I) and the Limit (L). The notation GCRA(I,L) is often used.



TAT - Theoretical Arrival Time

At the time of arrival t of the first cell of the connection, $TAT = t(1)$

Figure 2: Virtual Scheduling Algorithm



X - Value of the Leaky Bucket counter

X' - auxiliary variable

At the time of arrival of the first cell of the connection, $TAT = t(1)$

Figure 3: Continuous-State Leaky Bucket Algorithm

The GCRA is used to define the relationship between PCR and CDVT, and relationship between SCR and BT. The GCRA is also used to specify the conformance of the declared values of and the above parameters.

Priority Control

The end-system may generate traffic flows of different priority using the Cell Loss Priority (CLP) bit. The network may selectively discard cells with low priority if necessary such as in congestion to protect, as far as possible, the network performance for cells with high priority.

Traffic Shaping

Traffic shaping is a mechanism that alters the traffic characteristics of a stream of cells on a connection to achieve better network efficiency whilst meeting the QoS objectives, or to ensure conformance at a subsequent interface.

Examples of traffic shaping are peak cell rate reduction, burst length limiting, reduction of CDV by suitably spacing cells in time, and queue service schemes.

Traffic shaping may be performed in conjunction with suitable UPC functions.

Leaky Bucket Algorithm

The most famous algorithm for traffic shaping is leaky bucket algorithm. This method provides a pseudo-buffer (Figure 4). Whenever a user sends a cell, the queue in the pseudo-buffer is increased by one. The pseudo-server serves the queue and the service-time distribution is constant. Thus there are two control parameters in the algorithm: the service rate of the pseudo-server and the pseudo-buffer size.

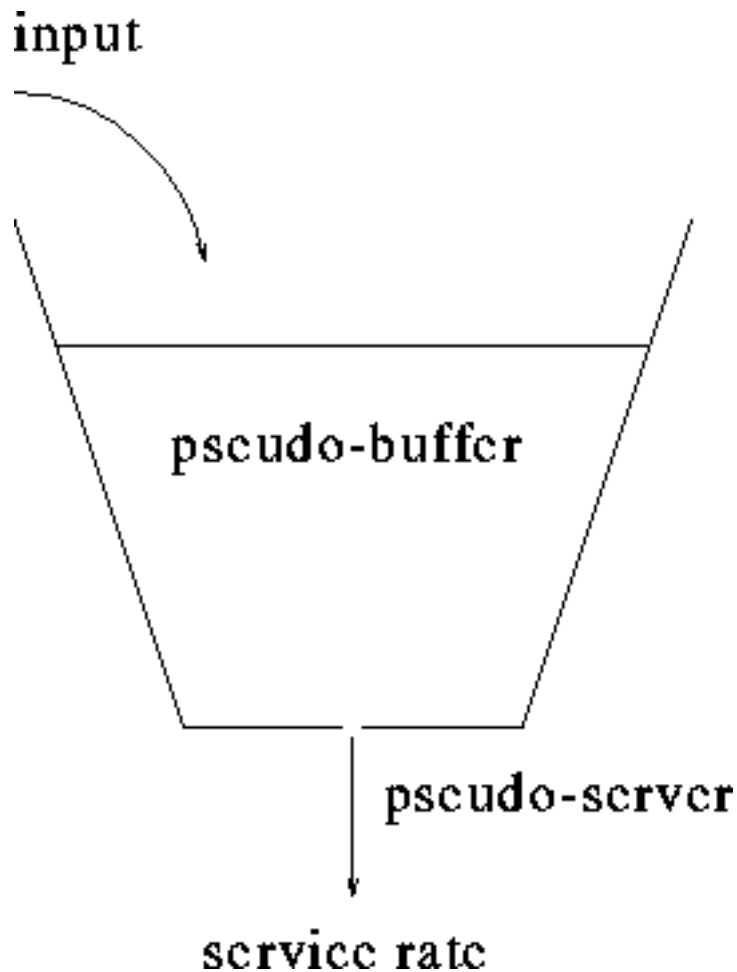


Figure 4: Leaky Bucket Method

As long as the queue is not empty, the cells are transmitted with the constant rate of the service rate. So the algorithm can receive a bursty traffic and control the output rate. If excess traffic makes the pseudo-buffer overflow, the algorithm can choose discarding the cells or tagging them with CLP=1 and

transmitting them.

PCR or SCR can be controlled by choosing appropriate values of service rate and buffer size. In addition, PCR and SCR can both be controlled by combining two buckets with one for each of the parameters. And there are many variances of the original scheme.

Network Resource Management

In Network Resource Management (NRM) is responsible for the allocation of network resources in order to separate traffic flows according to different service characteristics, to maintain network performance and to optimise resource utilisation. This function is mainly concerned with the management of virtual paths in order to meet QoS requirements.

Frame Discard

If a congested network needs to discard cells, it may be better to drop all cells of one frame than to randomly drop cells belonging to different frames, because one cell loss may cause the retransmission of the whole frame, which may cause more traffic when congestion already happened. Thus, frame discard may help avoid congestion collapse and can increase throughput. If done selectively, frame discard may also improve fairness.

Feedback Control

Feedback controls are defined as the set of actions taken by the network and by the end-systems to regulate the traffic submitted on ATM connections according to the state of network elements.

Feedback mechanisms are specified for ABR service class by ATM Forum Technical Committee. We will discuss it in detail later.

ABR Flow Control

As we have discussed before, the ABR service category uses the link capacity that is left over and is applied to transmit critical data that is sensitive to cell loss. That makes traffic management for this class the most challenging by the fluctuation of the network load condition, the burstiness of the data traffic itself, and the CLR requirement.

The ATM Forum Technical Committee Traffic Management Working Group have worked hard on this topic, and here are some of the main issues and the current progress of this area.

Some Early Debates

Congestion management in ATM is a hotly debated topic, many contradictory beliefs exist on most issues. These beliefs lead to different approaches in the congestion control schemes. Some of the issues have been closed after a long debate and the ATM Forum Technical Committee finally adopted one of them, and others are still open and the debates are continuing.

1. Open-Loop vs Close-Loop

Open-loop approaches do not need end-to-end feedback, one of the examples of this type are prior-reservation and hop-to-hop flow control. In close-loop approaches, the source adjust its cell rate in responding to the feedback information received from the network.

It has been argued that close-loop congestion control schemes are too slow in today's high-speed, large range network, by the time a source gets the feedback and reacts to it, several thousand cells may have been lost. But on the other hand, if the congestion has already happened and the overload is of long duration, the condition cannot be released unless the source causing the congestion is asked to reduce its rate. Furthermore, ABR service is designed to use any bandwidth that is left over the source must have some knowledge of what is available when it is sending cells.

The ATM Forum Technical Committee Traffic Management Working Group specified that feedback is necessary for ABR flow control.

2. Credit-Based vs Rate-Based

Credit-Based approaches consist of per-link, per-VC window flow control. The receiver monitors queue lengths of each VC and determines the number of cells the sender can transmit on that VC, which is called "credit". The sender transmits only as many cells as allowed by the credit.

Rate-Based approaches control the rate by which the source can transmit. If the network is light loaded, the source is allowed to increase its cell rate. If the network is congested, the source should decrease its rate.

After a long debate, ATM Forum finally adopted the rate-based approach and rejected the credit-based approach [5]. The main reason for the credit-based approach not being adopted is that it requires per-VC queueing, which will cause considerable complexity in the large switches which support millions of VCs. It is not scalable. Rate-Based approaches can work with or without per-VC queueing.

3. Binary Feedback vs Explicit Feedback

Binary Feedback uses one bit in the cell to indicate the elements along the flow path is congested or not. The source will increase or decrease its rate by some pre-decided rule upon receiving the feedback. In Explicit Feedback, the network tells the source exactly what rate is allowed for it to send.

Explicit Rate (ER) feedback approach is preferred, because ER schemes have several advantages over single-bit binary feedback [5]. First, ATM networks are connection oriented and the switches know more information along the flow path, the increased information can only be used by explicit rate feedback. Secondly, the explicit rate feedback is faster to get the source to the optimal operating point. Third, policing is straightforward. The entry switches can monitor the returning message and use the rate directly. Fourth, with fast convergence time, the initial rate has less impact. Fifth, the schemes are robust against errors in or loss of a single message. The next correct message will bring the system to the correct operating point.

There are two ways for explicit rate feedback: forward feedback and backward feedback. With forward feedback, the messages are sent forward along the path and are returned to the source by the destination upon receiving the message. With backward feedback, the messages are sent directly back to the source by the switches whenever congestion condition happens or is pending

in any of the switches along the flow path.

4. **Congestion Detection: Queue Length vs Queue Growth Rate**

Actually this issue does not cause too much debate. In earlier schemes, large queue length is often used as the indication of congestion. But there some problems with this method.

First, it is a static measurement. For example, a switch with a 10k cells waiting in queue is not necessarily more congested than a switch with a 10 cell queue if the former one is draining out its queue with 10k cell per second rate and the queue in the latter is building up quickly. Secondly, using queue length as the method of congestion detection was shown to result in unfairness[5]. Sources that start up late were found to get lower throughput than those which start early.

Queue growth rate is more appropriate as the parameter to monitor the congestion state because it shows the direction that the network state is going. It is natural and direct to use queue growth rate in a rate-based scheme, with the controlled parameter and the input parameter have the same unit.

RM-cell Structure

In the ABR service, the source adapts its rate to changing network conditions. Information about the state of the network like bandwidth availability, state of congestion, and impending congestion, is conveyed to the source through special control cells called Resource Management Cells (RM-cells).

ATM Forum Technical Committee specifies the format of the RM-cell. The already defined fields in a RM-cell that is used in ABR service is explained in this section.

1. **Header**

The first five bytes of an RM-cell are the standard ATM header with PTI=110 for a VCC and VCI=6 for a VPC.

2. **ID**

The protocol ID. The ITU has assigned this field to be set to 1 for ABR service.

3. **DIR**

Direction of the RM-cell with respect to the data flow which it is associated with. It is set to 0 for forward RM-cells and 1 for backward RM-cells.

4. **BN**

Backward Notification. It is set to 1 for switch generated (BECN) RM-cells and 0 for source generated RM-cells.

5. **CI**

Congestion Indication. It is set to 1 to indicate congestion and 0 otherwise.

6. **NI**

No Increase. It is set to 1 to indicate no additive increase of rate allowed when a switch senses impending congestion and 0 otherwise.

7. **ER**

Explicit rate. It is used to limit the source rate to a specific value.

8. **CCR**

Current Cell Rate. It is used to indicate to current cell rate of the source.

9. **MCR**

Minimum Cell Rate. The minimum cell rate desired by the source.

Service Parameters

ATM Forum Technical Committee defined a set of flow control parameters for ABR service.

1. **PCR**

Peak Cell Rate, it is the source desired but the maximum rate the network can support. It is negotiated when the connection is set up.

2. **MCR**

Minimum Cell Rate, the source need not to reduce its rate below it under any condition. It is negotiated when the connection is set up.

3. **ICR**

Initial Cell Rate, the startup rate after idle periods. It is negotiated when the connection is set up.

4. **AIR**

Additive Increase Rate, the highest rate increase possible. It is negotiated when the connection is set up.

5. **Nrm**

The number of cells transmitted per RM-cell sent. It is negotiated when the connection is set up.

6. **Mrm**

Used by the destination to control allocation of bandwidth between forward RM-cells, backward RM-cells, and data cells. It is negotiated when the connection is set up.

7. **RDF**

Rate Decrease Factor, to control the number of cells sent upon idle startup before the network can establish control in one Round Trip Time (RTT). It is negotiated when the connection is set up.

8. **ACR**

Allowed Cell Rate, the source can not transmit with rate higher than it.

9. **Xrm**

The maximum RM-cells sent without feedback before the source need to reduce its rate. It is negotiated when the connection is set up.

10. **TOF**

Time Out Factor, to control the maximum time permitted between sending forward RM-cells before a rate decrease is required. It is negotiated when the connection is set up.

11. **Trm**

The inter-RM time interval used in the source behavior. It is negotiated when the connection is set up.

12. **RTT**

Round Trip Time between the source and the destination. It is computed during call setup.

13. **XDF**

Xrm Decrease Factor, specify how much of the reduction of the source rate when XRM is triggered. It is negotiated when the connection is set up.

These parameters are used to implement ABR flow-control on a per-connection basis, and the source, switch and destination must behave within the rules that defined by these parameters.

The function and usage of these parameters are still under study.

Source, Destination and Switch Behavior

ATM Forum Technical Committee also specifies the source, destination, and switch behavior for the service.

There are two notations that need to be explained before we discuss the network behavior.

In-Rate Cells : The cells that counted in the user's rate with $CLP=0$. In-rate cells include data cells and in-rate RM-cells.

Out-of-Rate Cells : These cells are RM-cells and are not counted in the user's rate. They are used when $ACR=0$ and in-rate RM cells can not be send. The CLP is set to 1 for them.

In this section, we discuss some highlights of the specification.

● Source Behavior

1. The value of ACR shall never exceed PCR, nor shall it ever be less than MCR.
The source shall never send in-rate cells at a rate exceeding ACR.
2. The source shall start with ACR at ICR and the first in-rate cell sent shall be a forward RM-cell.
3. The source shall send one RM-cell after every N_{rm} data cells.
4. If the source does not receive any feedback since it sends the last RM-cell, it shall reduce its rate by at least $ACR * T * TDF$ after $TOF * N_{rm}$ cell intervals.
5. If at least X_{rm} in-rate forward RM-cells have been sent since the last backward RM-cell with $BN=0$ was received, ACR shall be reduced by at least $ACR * XDF$.
6. When a backward RM-cell is received with $CI=1$, ACR shall be reduced by at least $ACR * N_{rm} / RDF$.
If the backward RM-cell has both $CI=0$ and $NI=0$, the ACR may be increased by no more than $AIR * N_{rm}$.
7. Out-of-rate forward RM-cells shall not be sent at a rate greater than TCR.

● Destination Behavior

1. When a data cell is received, the destination shall save the EFCI state.
2. When returning an RM-cell, it shall set CI if saved EFCI is 1.
Congested destination may set both CI and NI, or reduce ER.
3. If an RM-cell has not been returned while the next one arrives, throw away the old one.
4. The destination can generate a backward RM-cell without having received a forward RM-cell.

● Switch Behavior

1. The switch may set the EFCI flag in the data cell headers.
2. The switch may set CI or NI in the RM-cells, or may reduce the ER field.
3. The switch may generate backward RM-cells with CI or NI set.

Representative Schemes

The following is a brief description of congestion control schemes that are proposed to the ATM Forum. The various mechanisms can be classified broadly depending upon the congestion monitoring criteria used and the feedback mechanism employed.

● EFCI Control Schemes

These class of feedback mechanisms use binary feedback involving the setting of the EFCI bit in the cell header. The simplest example of a binary feedback mechanism is based on the old DECbit scheme. In this scheme all the VCs in a switch share a common FIFO queue and the queue length is monitored. When the queue length exceeds a threshold congestion is declared and the cells passing the switch have their EFCI bit set. When the queue length falls below the threshold the cells are passed without their EFCI bit set. The source will adjust its rate accordingly when it sees the feedback cells with the EFCI bit set or not. Variations of this scheme include using two thresholds for the indication and removing congestion respectively.

Binary feedback mechanisms can sometimes be fair because long hop VCs have higher possibility to have their cell EFCI bit set and get fewer opportunities to increase their rate. It is called the "beat down problem". This problem can be alleviated by some enhancements to the basic scheme such provide separate queues for each VCs.

But a coherent problem with binary feedback mechanisms are that they are too slow for rate-based control in high-speed networks [5].

● Explicit Rate Feedback Schemes

As we have discussed before, explicit rate feedback control would not only be faster but would offer more flexibility to switch designers [5]. Many explicit rate feedback control schemes has been proposed, the following are some that is documented by the ATM Forum.

1. Enhanced Proportional Rate Control Algorithm (EPRCA)

In EPRCA, the source sends data cells with EFCI set to 0 and sends RM-cells every n data cells. The RM-cells contain desired explicit rate (ER), current cell rate (CCR) and congestion indication (CI). The source usually initializes CCR to the allowed cell rate (ACR) and CI to zero.

The switch computes a mean allowed cell rate (MACR) for all VCs using exponential weighted average:

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

and the fair share as a fraction of this average, where alpha and the fraction are chosen to be 1/16 and 7/8 respectively. The ER field in the returning RM-cells are reduced to fair share if necessary. The switch may also set the CI bit in the cells passing when it is congested which is sensed by monitoring its queue length.

The destination monitors the EFCI bits in data cells and mark the CI bit in the RM-cell if the last seen data cell had EFCI bit set.

The source decreases its rate continuously after every cell by a fixed factor and increases its rate by a fixed amount if the CI bit is not set. Another rule is that the new increased rate must never exceed either the ER in the returned cell or the PCR of the connection.

The main problem of this scheme is that the congestion detection is based on the queue length and this method is shown to result in unfairness [5]. Sources that start up late may get lower throughput than those that start early.

2. Target Utilization Band (TUB) Congestion Avoidance Scheme [5]

In each switch, a target rate is defined as slightly below the link bandwidth, such as 85-90% of the full capacity. The input rate of the switch is measured over a fixed averaging interval. The load factor z is then computed as:

$$\text{Load Factor } z = \text{Input Rate} / \text{Target Rate}$$

When the load factor is far from z , which means the switch is either highly overloaded or highly underloaded, all VCs are asked to change their load by this factor z . When the load factor is close to 1, between $1-\delta$ and $1+\delta$ for a small δ , the switch gives different feedback to underloading sources and overloading sources.

A fairshare is computed, and all sources whose rates are more than the fair share are asked to divide their rates by $z/(1+\delta)$, while those below the fair share are asked to divide their rates by $z/(1-\delta)$.

3. Explicit Rate Indication for Congestion Avoidance (ERICA)

This scheme tries to achieve efficiency and fairness concurrently by allowing underloaded VCs to increase their rate to fair share in spite of the conditions of the network and the sources already equal or greater than fair share may increase their rate if the link is underused. And the target capacity of the switch is set higher, 90-95% of the full bandwidth.

The switch calculates fair share as:

$$\text{Fairshare} = \text{Target capacity} / \text{Number of active VCs}$$

And the remaining capacity that a source can use is:

$$\text{VCshare} = \text{CCR} / \text{Load Factor } z$$

Then the switch sets the source's rate to the maximum of the two.

The information used to compute the quantities comes from the forward RM-cells and the feedback is given in the backward RM-cells. This ensures that the most current information is used to provide the fastest feedback.

Another advantage of this scheme is that it has few parameters which can be tuned easily.

4. Congestion Avoidance Using Proportional Control (CAPC)

In this scheme, the switches also set a target utilization slightly below 1 and measure the input rate to compute load factor z .

During underload (z is less than 1), fair share is increased as:

$$\text{Fairshare} = \text{Fairshare} * \text{Min}(\text{ERU}, 1+(1-z)*\text{Rup})$$

where Rup is a slope parameter between 0.025 and 0.1, and ERU is the maximum increase allowed.

During overload (z is greater than 1), fair share is decreased as:

$$\text{Fairshare} = \text{Fairshare} * \text{Max}(\text{ERF}, 1-(z-1)*\text{Rdn})$$

where Rdn is a slope parameter between 0.2 and 0.8, and ERF is the minimum decrease required.

The source should never allowed to transmit at a rate higher than the fair share.

The distinguishing feature of this scheme is that it is oscillation-free in steady state.

5. ER Based on Bandwidth Demand Estimate Algorithm

The switch calculates the Mean Allowed Cell rate (MACR) basing on a running exponential average of the ACR value from each VC's forward RM-cells as:

$$\text{MACR} = \text{MACR} + (\text{ACR} - \text{MACR}) * \text{AVF}$$

where AVF (ACR Variation Factor) is set to 1/16.

If the load factor is less than 1, the left-over bandwidth is reallocated according to:

$$\text{MACR} = \text{MACR} + \text{MAIR}$$

where MAIR is the MACR Additive Increase Rate.

The ER value is computed as:

$$\text{ER} = \text{MACR} * \text{MRF}$$

where MRF is the MACR Reduction Factor if congestion is detected,

$$\text{ER} = \text{MACR}$$

if no congestion is detected.

The congestion condition is detected by observing that the queue derivative is positive.

[Return to the Table of Contents](#)

Summary

Traffic management for ATM networks encompasses a number of interrelated elements operating over different levels and time scales. Among these, the congestion control is one of the most important component for the performance of the network and also is the most challenging one.

This report introduced the concepts in congestion control for ATM networks, and the specifications for

ATM traffic control proposed by ATM Forum are explained. The work done so far by the members of ATM Forum is presented.

[Return to the Table of Contents](#)

References

1. Amir Atai and Joseph Hui

``A Rate-Based Feedback Traffic Controller for ATM Networks''

Proceedings of the 1994 IEEE International Conference on Communications

Vol 3, 1994

Summary

This scheme also uses explicit feedback to adjust the source rate. The most important feature of it is that arrival rate rather than the queue length is used as the measure of congestion (MOC). The arrival rate is more accurate and may detect the onset of congestion sooner so that it can account for the propagation delay.

The preceding several schemes all use rate as the control objective. There is long debate over using rate or using credit as the control approach[5]. The ATM forum finally select rate mainly because that credit control need per-VC queueing, which is hard to implement.

2. Saewoong Bahk and Magda El Zarki

``Preventive Congestion Control based Routing in ATM Networks''

Proceedings of the 1994 IEEE International Conference on Communications

Vol 3, 1994

Summary

This is a routing scheme which in cooperation with congestion control. The basic idea is that shortest path is always used under light load, alternative path is used if the shortest one becomes congested. In selecting the alternative path, different classes of traffic such as delay-sensitive or loss-sensitive is treated differently, considering the different QoS required for them.

This scheme works best when the traffic in the network is unbalanced, with some part of it heavily loaded and other part is not. One shortcoming of this scheme is that it decides the path when the connection is set up and does not change the path if the network state changes. It should work better if it uses dynamic routing. But the advantage of static routing is that the implementation is simple.

3. Andreas Pitsillides and Jim Lambert

``Adaptive Connection Admission and Flow Control: Quality of Service with High Utilisation''

Proceedings of the 1994 IEEE INFOCOM'94

Vol 3, 1994

Summary

The scheme uses system identification and adaptive control theory. First it uses k-step ahead prediction and recursive least squares to estimate the network state and the control parameters and uses adaptive feedback and adaptive feedforward techniques to control the behavior of the network.

Similar to the first scheme, it applies some approach borrowed from control theory and it also combines CAC with lower level flow control. The idea is perfect if the characteristics of the network system can be obtained accurately by the estimation, which is still to be studied.

4. Parviz Yegani and Marwan Krunz and Herman Hughes
 ``Congestion Control Schemes in Prioritized ATM Networks"
 Proceedings of the 1994 IEEE International Conference on Communications
 Vol 2, 1994

Summary

The scheme uses priority queueing to meet the quality of service (QoS) requirements when the switch is congested.

When congestion happened at a node, the total input rate is greater than the output link capacity, the queue will build up in the node. Thus enough buffering is important in congestion management. As the congestion condition persists, the storage space in the node may run out, then dropping cells is inevitable. So selective discarding is also an important decision to be made in congestion management.

In this scheme, priority is used to achieve the desired performance in congestion conditions: delay-sensitive cells are given higher priority to be transmitted and loss-sensitive cells are given higher priority to get a buffer space so that they are less likely to be dropped.

Another selective discarding scheme drops all cells of one packet because one cell loss may cause the retransmission of the whole packet, which may cause more traffic when congestion already happened [5].

5. Raj Jain
 ``Congestion Control and Traffic Management in ATM Networks: Recent Advances and A Survey"
<ftp://ftp.netlab.ohio-state.edu/pub/jain/papers/cnis.ps>

Summary

The basic concept in ATM is introduced, especially the importance, mechanism and criteria of congestion control in ATM networks. Different congestion schemes are described and compared. The debate between rate-based and credit-based approach is presented.

6. Raj Jain
 ``Myths about Congestion Management in High Speed Networks"
ftp://ftp.netlab.ohio-state.edu/pub/jain/papers/cong_myths.ps

Summary

Some wrong ideas about congestion management in high-speed networks are pointed out and explained. Several competing approaches about congestion control and avoidance are presented, and the advantages and weakness of both sides are discussed. The conclusion is that a combination of several schemes is needed to achieve a desired performance.

7. Lampros Kalampoukas
 ``Performance of TCP over Multi-Hop ATM Networks: A Comparative Study of ATM-Layer Congestion Control Schemes"

Summary

Several congestion control schemes are studied by simulation, including ATM-Early Packet Discard (ATM-EPD) and ATM flow-controlled virtual channels (ATM-FCVC), which support TCP/IP on a multi-hop ATM network. The performance of these schemes are compared in terms of effective throughput, the number of retransmissions, degree of fairness, mean packet delivery time and packet delivery time standard deviation for different buffer size and network configurations.

8. Lin, T.Y. and Chen, Y.C.

``Congestion control approach for LAN/MAN interconnection via ATM''

Proceedings - IEEE INFOCOM

Vol 2, 1994

Summary

A scheme to solve the speed mismatch at ATM LAN/MAN gateways. A source quench message is sent to the host to slow down the traffic if the output buffer occupancy of the source gateway reaches a certain threshold. The distance gateway sends a turn-off signal to the source gateway for the same reason.

9. Chang, Chung-Ju and Cheng, Ray-Guang

``Traffic control in an ATM network using fuzzy set theory''

Proceedings - IEEE INFOCOM

Vol 3, 1994

Summary

The scheme is realized by using some fuzzy logic controllers. A fuzzy traffic controller is designed, whose major components are a fuzzy congestion controller, a traffic negotiator and a fuzzy admission controller.

Connection admission control is a open-loop, preventive congestion scheme. It is a higher level method and is necessary because the network cannot satisfy all the QoS negotiated using lower level schemes if there are too many users are there. But CAC alone is likely to be too conservative and not sufficient in regard of resource usage, because the peak rate bandwidth must be allocated to the source in case all of the source trying to transmit at their peak rate simultaneously. Thus lower level management must be used to take the advantage of the gain from statistical multiplexing. The merit of this scheme is that it uses both CAC and lower level congestion control approach to get the network work efficiently.

The problem is that it over simplifies the characteristics of the network. There are so many factors that influence the state of the traffic condition, the membership functions and control rules given cannot represent all the complex problems. On the other hand, if the membership functions and control rules are designed too complex, the algorithm will be impossible to run in real-time.

10. Ikeda, Chinatsu; Suzuki, Hiroshi

``Adaptive congestion control schemes for ATM LANs''

Proceedings - IEEE INFOCOM

Vol 2, 1994

Summary

The scheme provides two different ways to control two different types of traffic.

For the best effort class, the scheme combines forward explicit congestion notification (FECN) and backpressure to reduce the rate. The FECN is for the congestion condition which last longer than a round trip time and backpressure is for the immediate release to prevent the buffer overflow of the congested node.

For the guaranteed burst class, the scheme applies a modification of fast reservation protocol (FRP), which reduces the peak rate to be received after each NACK.

The merits of this scheme is that it treat different types of traffic according to their particular patterns and special needs, and that it combines two levels of control efforts to achieve the desired performance.

Another issue here is that it uses FECN, the notification of congestion is sent forward and then returned by the destination. Some control schemes uses backward explicit congestion notification (BECN) instead. Some believe that BECN can reach the source quickly and release the congestion faster. But others argue that BECN is not necessarily faster and may not reflect the whole state of the network.

11. Sanchez, Pedro-Ivan; Mazumdar, Ravi
 ``Definition of congestion: applications to ATM bandwidth management''
 Proceedings - IEEE INFOCOM
 Vol 2, 1994

Summary

The paper introduced the notion of Congestion Process (CP) and applies the concepts to police the source traffic.

Traffic policing is used to make sure that the user is sending its cells within the negotiated parameter and to isolate the misbehaving source to prevent it from causing harm to the network and other users. It is called usage parameter control (UPC) in ATM networks. Traffic shaping is used to shape the traffic sent to the network to prevent the switch from being overwhelmed. Both of the approaches help to prevent or lessen the congestion situation.

One of the widely used scheme for policing and shaping is leaky bucket algorithm. It has two control parameters, one of which controls the average rate (or the peak rate) the source can transmit, and the other controls how much the source can temporarily exceed that rate. There are many variances of the original scheme.

12. Yin, Nanyang; Hluchyj, Michael G.
 ``On closed-loop rate control for ATM cell relay networks''
 Proceedings - IEEE INFOCOM
 Vol 1, 1994

Summary

The scheme uses explicit forward congestion indication (EFCI) as feedback. This is an improvement compared with the previous scheme in that the source has more information as to how to adjust its transmission rate.

Another merit of this scheme is that the source can transmit at a rate greater than its sustainable cell rate (SCR) if the load of the network is light. Hence the resources of the network are fully

utilized.

Third, only the sources that exceed their SCR are required to decrease their rate when congestion is experienced. Thus the negotiated quality of service (QoS) is guaranteed.

The weakness of this scheme is that it uses the average queue length to determine the congestion condition, which is inaccurate and misleading in some cases[5].

13. Gong, Yu; Akyildiz, Ian F.

``Dynamic traffic control using feedback and traffic prediction in ATM networks"
Proceedings - IEEE INFOCOM
Vol 1, 1994

Summary

The scheme uses a feedback cell (FC) to get the information of the network along the path. FC is issued periodically and is transmitted with highest priority. The source adjusts its rate upon receiving each returning FC.

The merit of this scheme is that it considers the feedback delay when uses the information in the FC. The problem is that binary feedback is usually not sufficient for the algorithm to converge to the desired rate[5].

14. Ramamurthy, G.; Dighe, R.S.

``Performance analysis of multilevel congestion controls in BISDN"
Proceedings - IEEE INFOCOM
Vol 1, 1994

Summary

Using a hierarchical model to analysis the performance of the multilevel control of congestion in ATM networks. It is observed that allowing a small probability of burst level block can significantly decrease the probability of call blocking and increase the link utilization. Thus, to achieve an optimal performance rather than get improvement on a certain point.

15. Saito, Hiroshi

Teletraffic technologies in ATM networks
Artech House, Boston
1994

16. ATM Forum Traffic Management Specification Version 4.0

[Return to the Table of Contents](#)

This paper is available on-line at http://www.cis.ohio-state.edu/~jain/cis788-95/atm_cong/index.html
Last Modified, August 21st, 1995