ATM FORUM: Technical Working Group
ATM_FORUM/93-771

****

SOURCE: Bellcore
Mark W. Garrett, Deh-phone K. Hsing and Ljiljana Trajkovic
445 South Street
Morristown, NJ

****

ISSUES ADDRESSED: Proposed Traffic Management Work Item - Study of Scheduling Schemes

****

DATE: July 27-29, 1993

****

DISTRIBUTION TO: ATM Forum Technical Working Group Members

****

ABSTRACT

This contribution proposes to initiate work on scheduling schemes within the Traffic Management group. The goal is to encourage use of scheduling schemes with desirable properties which enable useful quality of service (QOS) commitments. This would constitute an integral part of the overall traffic control and congestion control architecture of the ATM network in order to fulfill the traffic contract to all ATM connections established. Some rationale and issues to be addressed as a result of this activity are provided.

NOTICE

1. Introduction

This contribution proposes that we initiate a new work item within the Traffic Management group to study scheduling schemes.

With simple scheduling schemes (such as FIFO), an ATM multiplexer may introduce significant distortion of the input traffic flow on some connections due to temporary contention at the ATM multiplexer. For instance, it is possible for a smoothly spaced cell stream to become quite clustered at the ATM multiplexer resulting in a bursty stream arriving at the next node. To maintain a QOS commitment this would require a decrease in link utilization or an increase in available buffers. A more viable and robust solution will be for the Forum to establish some guidelines on the scheduling mechanism in terms of the scheduler's effect on the traffic (not in terms of the algorithm itself). Some investigation of appropriate scheduling schemes is necessary to develop a better understanding of feasible bounds on the ''distortion'' introduced at a multiplexing point, and the related effects on link utilization and buffer requirements. Specifically, we propose that the Traffic Management group consider a class of ''fair'' scheduling schemes.[1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13]. [14] [15] [16] The existing literature indicates that there is a substantial and maturing knowledge of such schemes and that it should be possible for ATM technology to reflect these technical advances.

This type of ''fair'' scheduling scheme provides a mechanism to give each connection access to its allocated bandwidth. An idealized model is one where cells from each connection are ordered for service as if each VC were processed at its allocated rate *plus* its fair share of the unused bandwidth. The server then supplies cells to the output link in the determined order. To some extent these schemes are practical approximations to this ideal. This makes service commitments for CBR sources with low-delay tolerance possible. It also forms an architectural structure providing access to bandwidth resources within a bounded and calculable time. The schemes which are work-conserving (a few in the given references are not) are most attractive because they provide the mechanism for connections to share available bandwidth in an underloaded network without affecting the service guarantees for any other user. Finally, as the word ''fair'' implies, these schemes can ensure equitable use of any unused or unallocated bandwidth. These properties provide a rich structure within which many useful QOS classes can be implemented.

2. Objective

The output of this study could provide a general guideline on the amount of ''distortion'' allowed between the input traffic at one interface point and the output traffic at another interface point (e.g., guideline on a set of values of τ that a network equipment could accommodate). The following are some desirable characteristics of a scheduling scheme:

- reasonably simple implementation;
- reasonably small buffer requirement (due to cost, delay and complexity);
- minimal distortion introduced to the traffic streams (ideally zero).

3. Issues Addressed

The following are some open issues identified in current Forum UNI specification which will be addressed by the studies on scheduling schemes.

1. Connection Traffic Parameter Specification at Public UNI

   It is recognized that although the same algorithm (i.e., the GCRA) is used to specify the traffic of an ATM connection at the source and at the entry point of the public network (the public UNI), some modifications to the values of the corresponding GCRAs are needed. Currently, the Forum spec assumes that the same CDV tolerance (τ) will be added to determine conformance at public UNI for both PCR and SCR of a connection. That is, if the traffic is conforming at the source PHY-SAP to GCRA(T,0) and GCRA(T_s,τ_s), it is tested for conformance at the public UNI according to GCRA(T, τ) and GCRA(T_s,$\tau_s + \tau$). It can be shown that this is always true as long as τ is chosen to be equal to or greater than the difference between the maximum delay and the minimum delay introduced by the customer premise equipment between the source and the public UNI. Unless shaping is performed at the public UNI, the value (or values) of τ supported by the network at the public UNI has implications on the delay requirement imposed on the customer premise equipment which in turn is closely related to the scheduling schemes implemented. Specifically, the following issues could be addressed by the studies of ‘‘fair’’ scheduling schemes.

a. How the value of τ could be related to the PCR of a connection.

b. Whether τ could be bounded by a value smaller than the ‘‘distortion’’ on τ_s through the use of ‘‘fair’’ scheduling.

2. Support of Best Effort Service

Although further study is needed to better characterize the best effort service being discussed within the group, the following are some basic characteristics of this service:

- The source is adaptive to the network load. That is, it will slow down when the network load is high and would like to use the available network resources as much as they can when needed.
- This service is more loss sensitive but less delay sensitive.

In order to minimize the interference to other services (e.g., those services with non-adaptive sources) and provide fair sharing of available network resources among the best effort service users, the scheduling scheme is an important element that needs to be investigated.

4. Proposal

We believe that the freedom to innovate in the process of equipment design and implementation is crucial for the introduction of continually improved products. Hence we do not recommend to standardize any specific scheduling scheme. However, it seems appropriate at this point to investigate scheduling algorithm implementations in order to provide some added value to the general industry knowledge. This exercise will provide an existence proof for a scheduling algorithm with a low traffic distortion bound. This will not prohibit any other approach as long as its performance is at least as good. No additional standardized communication is needed among the equipment to support whatever scheduling schemes are implemented. The issue of interoperability is greatly aided by this approach, as long as each implementation conforms to a ‘‘small’’ bound on distortion (especially if a bound of zero turns out to be feasible).

This contribution proposes that we initiate a new work item to study scheduling schemes. It is proposed that we invite contributions and tutorial sessions on various scheduling schemes that will discuss not only the algorithms but also their performance and implementation complexities.

Bellcore intends to support this proposal with simulation and analysis over the course of the next 6-12 months, and will report results as they become available. We wish to ascertain a number of specific algorithms, their implementability and traffic properties upon which QOS commitments can be made.

REFERENCES

1. J. Nagle, ‘‘On Packet Switches with Infinite Storage’’, IEEE Trans. Comm., Vol. 35, pp 435-438, 1987.
2. L. Zhang, ‘‘A New Architecture for Packet Switching Network Protocols,’’, Ph.D. Dissertation, Massachusetts Institute of Technology, August, 1989.
3. A. Demers and S. Keshav and S. Shenker, ‘‘Analysis and Simulation of a Fair-Queueing Algorithm,’’ Proc. ACM SigComm Symp., September, 1989, pp 1-12.
4. D. Ferrari, D. Verma, ‘‘A Scheme for Real-Time Channel Establishment in Wide-Area Networks,’’ IEEE J. Sel Areas on Commun., Vol. 8, No. 3, pp 368-79, April 1990.
5. S. J. Golestani, ‘‘Congestion-Free Transmission of Real-Time Traffic in Packet Networks,’’ Proc. IEEE Infocom, pp. 527--36, June 1990.
6. S. J. Golestani, ‘‘A Stop and Go Queueing Framework for Congestion Management,’’ Proc. ACM SigComm, pp. 8--18, September 1990.
7. L. Zhang, ‘‘VirtualClock: A New Traffic Control Algorithm for Packet Switching Networks’’, Proc. ACM SIGComm Symp., pp. 19--29, September 1990.
8. C. R. Kalmanek, H. Kanakia, S. Keshav, ‘‘Rate Controlled Servers for Very High-Speed Networks,’’ Proc. IEEE Globecom, December 1990.
9. S. Keshav, ‘‘Congestion Control in Computer Networks,’’ Ph.D. Thesis, Univ. California, Berkeley, June 1991.
10. S. Keshav, ‘‘A Control Theoretic Approach to Flow Control,’’ Proc. ACM SigComm, September 1991.
11. H. Zhang, S. Keshav, ‘‘Comparison of Rate-Based Service Disciplines,’’ Proc. ACM SigComm, September 1991.
12. S. Keshav, ‘‘On the Efficient Implementation of Fair Queueing,’’ J. Internetworking Research and Experience, 1991.
13. A. K. J. Parekh, ‘‘A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks,’’ Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, Mass, February, 1992.
14. A. K. Parekh, R. G. Gallager, ‘‘A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks---The Single Node Case,’’ Proc. IEEE Infocom, pp. 915--924, May 1992.
15. D. Clark, S. Shenker, and L. Zhang, ‘‘Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanisms,’’ Proc. ACM SigComm, pp. 14--26, August 1992.
16. A. K. Parekh and R. G. Gallager, ‘‘A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks---The Multiple Node Case,’’ Proc. IEEE INFOCOM, San Francisco, pp. 521--530, May 1993.