

Research Report

A Critique of ATM from a Data Communications Perspective

Israel Cidon

IBM Research Division
T. J. Watson Research Center
Yorktown Heights, NY 10598

Jeff Derby

IBM Communication Systems
Research Triangle Park, NC 27709

Inder Gopal and Bharath Kadaba

IBM Research Division
T. J. Watson Research Center
Yorktown Heights, NY 10598

LIMITED DISTRIBUTION NOTICE

This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents and will be distributed outside of IBM up to one year after the date indicated at the top of this page. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties).

A CRITIQUE OF ATM FROM A DATA COMMUNICATIONS PERSPECTIVE

Israel Cidon

*IBM T. J. Watson Research Center
Yorktown Heights, NY 10598*

Jeff Derby

*IBM Communication Systems, Research Triangle Park,
NC27709*

Inder Gopal, Bharath Kadaba

*IBM T. J. Watson Research Center
Yorktown Heights, NY 10598*

ABSTRACT

Fast Packet switching(FPS) is emerging as the preferred technology for future high speed, integrated networks. Asynchronous Transfer Mode (ATM) is an aspect of the Broadband ISDN Standard that is in very early stages of development. The standards activities are restricted thus far to choosing a 48 byte fixed "cell" and cell label swapping for routing. Even at this early stage, there are several concerns regarding ATM relating its suitability for data communications. These concerns are brought to focus in this paper by comparing it to an alternative approach to FPS developed at IBM called PARIS. PARIS uses variable length packets and source routing headers. By using LAN traffic data, we show that the fixed length packets in ATM can result in significantly worse transmission efficiency over variable size in many real traffic scenarios; considerably more processing power (requiring VLSI implementation) is needed to handle segmentation and reassembly overhead associated with ATM small cells, and statistical multiplexing present some unique problems. Also, we present some qualitative arguments to show that the label swapping approach for routing is more complex to implement, potentially slower in processing call setup and more difficult to support datagrams when compared to the source routing technique.

1.0 INTRODUCTION

In the current communications network technology revolution, amidst a number of ongoing debates and disagreements, there appears to be a general consensus on the following points. First, it is both feasible and desirable that all types of traffic (voice, data and video) be carried on a common backbone network; second, Fast Packet Switching (FPS) is the most suitable method to accomplish this. Here, FPS simply means that information transfer in the network is done in packets that carry additional control bits (headers) and processing in the intermediate nodes is limited to using these headers to route and schedule packets on the appropriate outgoing links. This restricted processing permits simple hardware implementation allowing one to build switches capable of handling millions of packets per second - a requirement of future networks supporting integrated services. The format of the packets and the routing and scheduling techniques are of crucial importance and forms the focus of this paper.

The CCITT has defined the asynchronous transfer mode (ATM) as the FPS mechanism to be employed in broadband ISDN (B-ISDN) [ATMA90]. ATM is based on the use of short, fixed-length packets called "cells". The IEEE 802.6 draft standard for metropolitan-area networks (MANs) employs the same cell structure as B-ISDN [IEEE88]. These standards are primarily driven by public-network providers with the goal of building integrated services public networks. The standards should permit these independently constructed networks to inter-operate easily, thereby resulting in ubiquitous worldwide service offerings. In the United States, AT&T and the RBOCs are driving the B-ISDN standard through the ANSI T1S1 committee and the MAN standard through the IEEE 802.6 committee. Initial use in US may be via SMDS (Switched Multi-megabit Data Services), a datagram service offering planned by RBOCs with interconnection of customer LANs as the primary application [SMDS89].

While the B-ISDN standard is still evolving, there is agreement on packet format and routing method [ATMB90]. As noted above, ATM packets are short, fixed-length "cells", with the overall cell length equal to 53 bytes in the current proposal. An "adaptation layer" is defined above the

Additional components of the network function such as end point protocols interfacing with different traffic sources as well as network control and management aspects (such as bandwidth allocation) while critical to the total network system are not pertinent to this paper.

ATM transport to map user traffic on the ATM transport, segmenting user packets into cells and reassembling cells into user packets. User packets are transported on a integral number of cells using padding bits to complete unfilled cells. This approach is very different from the traditional packet switches that use variable length and unpadded packets.

In a B-ISDN network, ATM cells are routed based on the contents of a "label" in the header of each cell. The labels are used in intermediate nodes in conjunction with routing tables to determine the outgoing link on which the cell should be transmitted. The label is valid only for the current hop and is replaced by a new label that will be interpreted at the next hop. The routing table in any intermediate node contains an entry for each label assigned on each incoming link, with the entry providing a mapping to the appropriate outgoing link and the new label to be used on that link. The assignment of labels and construction of the routing-table entries are carried out as part of a connection-setup procedure. This type of routing method is well known and used by many popular data networks (for example, SNA/APPN [BGGJP85])

While there good reasons to aim for a ubiquitous worldwide broadband network capable of providing advanced services, there are several concerns regarding the above approach. These include:

- The design is significantly sub-optimal for data. Specifically, the short cell size seems to have been motivated primarily by voice-traffic considerations, with the intent of keeping packetization delay and queuing delay small. It appears that this design choice has an adverse impact on data traffic. This issue is critical given the expectation of a dramatic increase in the volume of data traffic to be handled by broadband networks. Indeed, the first network offering to use ATM cells is likely to be one that at least initially will carry only data traffic, namely SMDS.

Scalability to networks with gigabit/sec links. The routing mechanism as well the short cells may prove to be bottlenecks in building high end switching nodes.

These problems are more closely examined in this paper. The concerns regarding ATM approach are brought to a sharper focus by comparing it to an alternative approach to FPS developed at IBM called PARIS. We provide a brief overview of PARIS below.

The PARIS [CG88] high speed wide area networking project has been in progress since 1986 in IBM Research. A significant amount of work has been done in defining a complete network architecture for supporting high speed integrated (voice, video, and data) networks.

PARIS uses variable-length packets that can range in size from a few bytes to several thousand bytes determined by network implementation considerations such as buffer size. This variable length packet approach is referred to in this paper as Packet Transfer Mode (PTM).² Variable-length packets are employed in LANs and in classical packet networks. They are also used in fast-packet, wide-area network structures such as ISDN frame relay. Here, it is possible to transport user packets in integral form with minimal "adaptation layer" processing.

PARIS makes use of a source routing (referred to as Automatic Network Routing or ANR) scheme which lends itself to easy hardware implementation and extremely fast call setup and take-down procedures. It also enables the implementation of additional routing functions such as copy, broadcast and transparent route switching. High packet switching performance is achieved at low implementation cost through the use of simple intermediate node algorithms which are optimized for hardware implementation. Most processing intensive tasks such as flow control, error recovery and adaptation protocols for voice and video are done on an end-to-end basis. A 4-node prototype supporting at 100 Mbits/sec links has been built to demonstrate some of the key network concepts.

Our approach to evaluating ATM is to compare it to the PARIS technology; specifically, the aspects we consider are:

Fixed slotted cell structure vs. variable unslotted packet structure

2. Label routing vs ANR

In examining the first aspect we focus on data traffic as the predominant traffic to be carried by these FPS networks. Specifically, we focus on LAN-LAN traffic since we expect this to be a main source of traffic in the network. For example, the SMDS services are aimed at carrying this type of traffic initially. We consider real traffic measurement data on such LANs as well as analytical techniques to quantify the performance differences. The following issues are studied in detail:

- Data transmission efficiency
 - FIFO queuing behavior
 - Adaptation layer processing
 - Statistical Multiplexing

² The term PTM was originated by Chuck Davin and David Tannenhouse at MIT

In examining the second main point of comparison, namely label routing vs ANR, we present qualitative arguments on the following aspects:

Nodal hardware/software complexity

- Speed and efficiency of connection setup and takedown
- Datagram support

It is important to note at this point that FPS mechanisms other than PARIS have been defined that are not based on the use of cells. One of these, usually referred to as Frame Relay, has in fact been standardized by the CCITT for use in "narrowband" ISDN. Frame Relay employs variable-length frames just as PARIS does. Unlike PARIS, Frame Relay employs label routing rather than source routing. However, in Frame Relay the use of the label is defined only at the interface between the user and the network. It is in fact possible for a network that provides a Frame Relay service to employ PARIS-style source routing internally. Given these considerations, it is possible to conclude that the results developed on the sequel can be generalized beyond the comparison of PARIS with ATM. That is, the results associated with the use of variable-length frames are applicable to Frame Relay with virtually no modification, while the results associated with the use of source routing could be applied to a network that provides a Frame Relay service while employing source routing internally.

Following is a brief summary of this paper. In the next section, by using LAN traffic data, we show that the fixed length packets in ATM can result in significantly worse transmission efficiency over variable size in many real traffic scenarios; considerably more processing power (requiring VLSI implementation) is needed to handle segmentation and reassembly overhead associated with ATM small cells. Overall, this supports our belief that, particularly for data, the offered traffic in networks will be very heterogeneous in nature with wide variations in terms of rate, packet sizes, variance, etc.. The ATM approach is to convert this heterogeneous traffic stream into a homogeneous cell stream of fixed sized small cells. The PTM approach is based on the claim that the conversion of a set of heterogeneous user traffic streams into a homogeneous cell structure introduces more overhead (in utilization and complexity) than it saves in design complexity and delay.

In the last section of the paper, we present arguments to show that the label swapping approach for routing is more complex to implement both from the processing and storage at the intermediate node when compared to the source routing technique. The connection setup/takedown may be

slower in ATM and the datagram support more cumbersome to implement. These are only qualitative in nature at this point and additional detailed comparison of operational prototypes are necessary to understand these issues further.

2.0 FIXED CELL VERSUS VARIABLE PACKET

In this section, we present and discuss various issues that arise because of the restriction to small fixed sized cells in ATM and compare it to PTM.

2.1 ATM assumptions

The basic cell structure defined by the CCITT for ATM in B-ISDN has a nominal payload of 48 bytes with a header of 5 bytes. This same basic cell structure has also been adopted for the metropolitan-area network (MAN) standard being developed by IEEE 802.6. Additionally, 4 bytes have been extracted from the nominal payload and added to the overhead. It should be noted here that the addressing information contained in the user frame (such as IEEE MAC addresses or ISDN addresses) as well as any frame check sequence contained in the user frame are considered to be part of the payload by the adaptation layer. In terms of the cell format, the best case in terms of minimum overhead would be to assume that 48 bytes will be user payload with only 5 bytes for overhead.

Two fixed-length alternatives will be considered below, namely:

A53- This is the current CCITT and IEEE 802.6 proposal

A69- This represents an earlier CCITT and IEEE 802.6 proposal. We have included this just to contrast it with the current agreed upon standard although the differences turn out to be not very significant.

Finally, with regard to frame delimiting, it is assumed that cell synchronization is obtained by using "synch cells" containing a unique header and that these cells are sent so rarely that their contribution to the total overhead is negligible. Since the cells are all the same length, identifying the start of each cell is then a simple matter of counting bytes.

2.2 PTM assumptions

The PARIS architecture is an example of PTM in that it employs variable-length network frames that explicitly includes adaptation-layer functions. Any user frame whose length is less than or equal to $(N_P)_{\max}$ is sent in a network frame whose payload size is equal to the user-frame size. Any

user frame whose length is greater than $(N_P)_{\max}$ is split into segments of length $(N_P)_{\max}$. The last segment, which in general will be shorter than $(N_P)_{\max}$, is sent in a network frame with payload size equal to its length. The importance of including the adaptation layer function is that implementation choices for the network can be at least partially decoupled from considerations relating to the nature of attachments to the network. For example, an $(N_P)_{\max}$ of 2K bytes can be chosen based on network buffer-size considerations, without limiting the network's ability to carry bridged traffic from an FDDI ring with a maximum user-frame size of about 4500 bytes.

The variable-length alternatives that will be considered below are oriented towards the PARIS architecture in that they include the adaptation-layer mechanism outlined above as well as the notion of variable-length headers to support source routing. The specific alternatives to be considered, both with an average overhead of 12 bytes, are:

PAR2K- PTM with $(N_P)_{\max} = 2K$

PAR4K- PTM with $(N_P)_{\max} = 4K$

Finally, it is assumed for the variable-length cases that frame delimiting is realized using HDLC flags, and that the HDLC "zero-bit-stuffing" mechanism is employed to prevent data from mimicking a flag. This will add slightly to the net overhead for any given frame depending on the bit-pattern it contains. For random data, the incremental overhead due to this bit-stuffing is negligibly small and is ignored.

2.3 Offered traffic models

For the various comparison points we want to consider, the results vary widely based on the offered traffic model. Thus, we performed studies on a wide spectrum of offered traffic models chosen from published results, extrapolations, or specific application characteristics.

In order to compare the fixed-length and variable-length alternatives with regard to data-transmission efficiency, overhead, and related issues, it is necessary to characterize the traffic offered to the network. We note that providing connectivity to LANs and MANs is expected to be a major application of fast-packet networks, with SMDS, which is targeted at precisely this application, likely to be the first available service offering in the US to employ the ATM cell concept. As a result, we have based the traffic models used in our study on reports of measured traffic patterns on operating LANs. We focus on sets of published measurements from M.I.T. [MIT86], Univer-

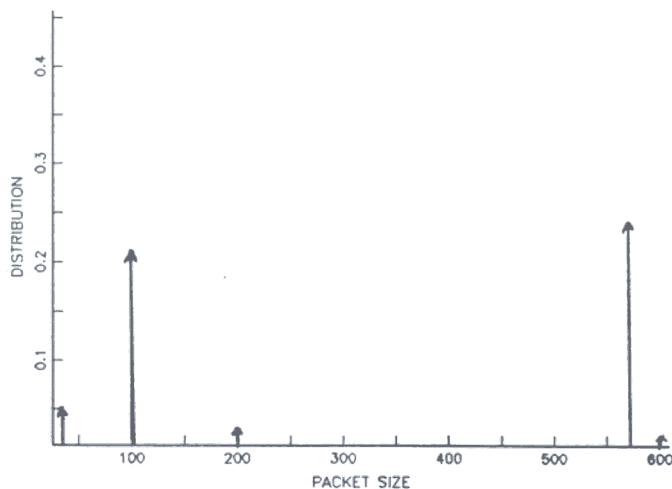
sity of Delaware [UD87] and Berkeley [BER87]. In addition, we have attempted to construct models representative of the traffic one may expect to be generated as users begin to take full advantage of wide-area, fast-packet networks. We have two sets of such models:

The packet-size distributions in the three published reports have been "stretched" by scaling their abscissas up while keeping the ordinates fixed. The result is a set of distributions in which the packets are generally longer but the relative numbers of long packets and short packets remains approximately the same.

A distribution has been created that may be representative of an application that would generally send only very large blocks of data; an image application might be one example here. In this case, the large majority of packets sent would be very long.

Our study thus considers the following traffic models:

MIT: This is the MIT distribution, shown in figure below. It displays behavior typical of several other measured packet-length distributions. This distribution is bimodal, with almost 45% of the packets sent being less than 50 bytes in length and most of the remainder being between 530 and 570 bytes in length.

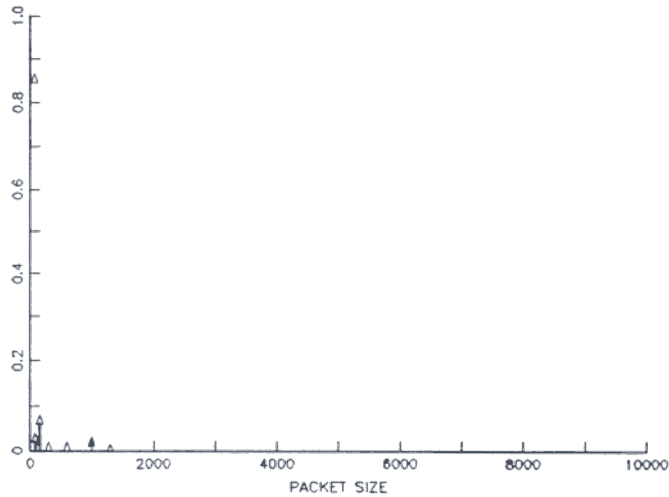


MIT distribution

We also used at the Berkeley distribution although the results are not included here for space reasons. It is bimodal, with 21.4% of the packets having a length of 46 bytes and 40.8% of

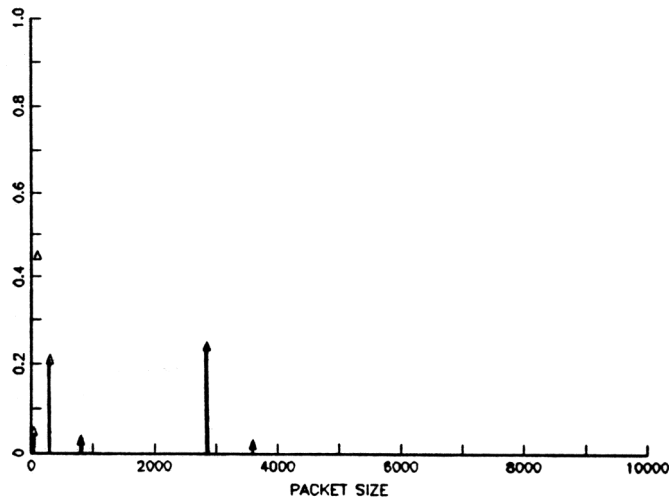
the packets having a length of 1072 bytes. In addition, almost 90% of all the bytes transmitted were contained in packets whose length was greater than 1000 bytes.

2. **DEL:** This is the University of Delaware distribution also shown below. This distribution differs somewhat from the above distributions in that in this case a large majority (87%) of the packets sent were short control packets containing 64 bytes each.

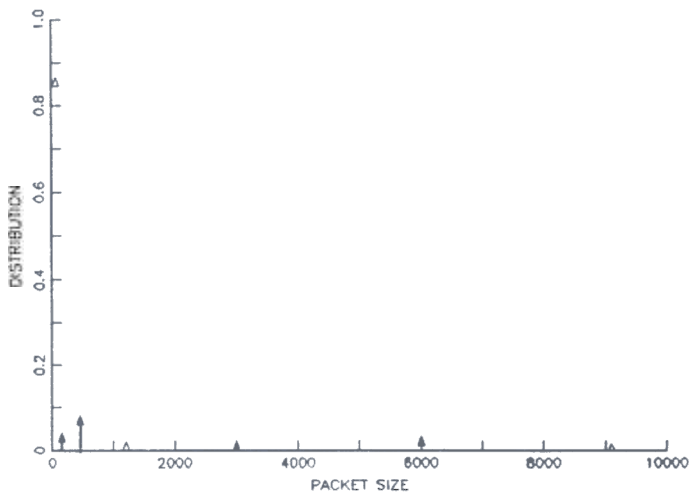


Delaware distribution

3. **MIT/and DEL/:** This are “stretched” version of the MIT and Delaware distributions and they are shown below.



MIT stretched distribution



Delaware stretched distribution

4. *IMAGE*: This is the distribution representing transmission of very large blocks of data. As shown in the figure, a large majority (80%) of the packets sent are about 9000 bytes in length, with the remainder assumed to be short supervisory packets. The length of the large packets was chosen to be essentially equal to the maximum packet length supported by SMDS and IEEE 802.6.

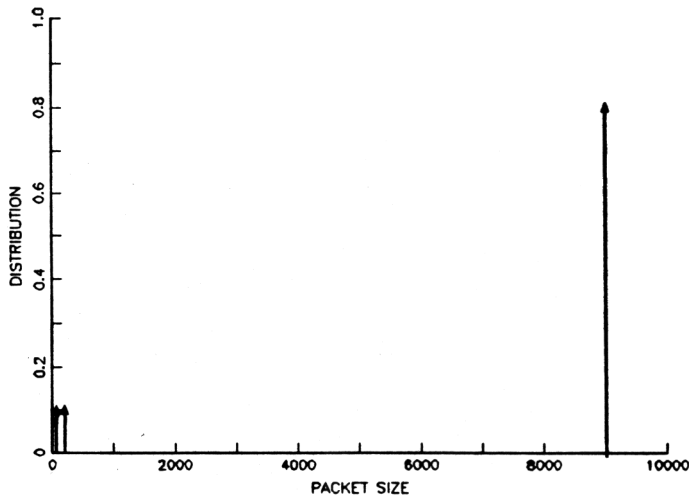


Image distribution

In using these distributions in our study, we have assumed that they specify directly the packet-size distribution in the fast-packet wide-area network. For example, in using the MIT distribution we consider a scenario in which the LAN is connected to a fast-packet WAN, with the probability that a packet from the LAN is forwarded across the WAN being independent of the packet length.

2.4 Data Transmission efficiency

The parameter of interest in this section is the data transmission efficiency. This is defined as the ratio of useful user data transmitted on a communication link to the actual bit rate of that link. The sources of inefficiency that we model are the header overhead (one header per ATM cell/PTM packet) and the cell padding caused by integrality constraints of fixed length ATM cells. We ignore additional overheads caused by packet framing, bit stuffing, cell synchronization, or SONET framing. The assumption is that these additional overheads are very specific to the underlying transmission medium and cannot easily be captured in a generally meaningful manner. Since there is only a single header to be used per user packet, PTM has intrinsically higher transmission efficiency. The use of source routing in PTM does cause a somewhat larger header size than the 9 bytes of ATM (approx. 12 bytes).

2.4.1 Analysis

Let the user data size be given by U (All sizes are assumed to be in bytes). Further let the payload in an ATM cell size is denoted by A and the header size per cell denoted by H_{ATM} . Similarly, denote the maximum PTM packet payload by P and the average PTM header by H_{PTM} . The ATM transmission efficiency, E_{ATM} , is given by the expression:

$$E_{ATM} = \frac{U}{(\text{ceiling}(\frac{U}{A}) \times (A + H_{ATM}))}$$

Similarly, If all packets were of fixed size, U , the PTM transmission efficiency, E_{PTM} , is given by the expression:

$$E_{PTM} = \frac{U}{U + (\text{ceiling}(\frac{U}{P}) \times (H_{PTM}))}$$

