The Impact of Header Compression and Rate Control on Typical TCP/IP traffic over ATM connections.

G. J. Armitage [gja@ee.mu.OZ.AU, (03) 344-7976]
K.M. Adams [kma@ee.mu.OZ.AU, (03) 344-7974]

Dept. of Electrical & Electronic Engineering
University of Melbourne
Parkville 3052,
Victoria,
AUSTRALIA

FAX +61 3 344 6678

Proc. Australian Broadband Switching and Services Symposium, Wollongong, July 1993. This pdf reconstructed Dec 2005.

ABSTRACT

This paper looks at the variation of cell use efficiency with IP packet size, and argues that it is a misleading measure of ATM's usefulness. We focus on the number of cells needed to carry 'typical' IP traffic, and show that the existence of large IP packets renders the inefficient cell use of small IP packets less important. We also show that the effect of cell rate control on IP packet rates will not disrupt applications generating 'typical' IP traffic. Simple plots of IP bit rate versus packet size on a rate limited connection are shown to be misleading.

1. INTRODUCTION.

The ITU-TSS (ITU Telecommunication Standardisation Sector, CCITT until this year) has adopted a cell based Asynchronous Transfer Mode (ATM) to provide flexible and uniform physical layer transport for a variety of communication services [1, 2] . Various ATM Adaptation Layers (AAL) have been defined to support a range of services. The AAL "adapts" the cell based ATM physical layer to packet, datagram, or bitstream oriented higher layers [3,4]. It exists at the endpoints of virtual connections, where higher layers wish to establish communication.

The ability to carry packetised traffic has prompted the Network Working Group of the Internet Engineering Task Force (IETF) to develop a draft standard for using ATM links to Internet Protocol (IP) traffic [5]. ATM virtual connections will initially support long-haul IP links, with IP-over-ATM extending closer to the desktop as ATM itself evolves and expands.

One item of discussion in IP circles is the potential inefficiency of carrying IP traffic over ATM's fixed size cells. IP traffic traces by Caceres [6] have been used to cast poor light on the ATM cell size, arguing that the size of most IP packets results in underutilisation of ATM cell payloads [7]. We believe that these conclusions are unnecessarily pessimistic. A further issue is cell based rate control, imposed either by the local AAL or at some distant point along a virtual connection, and its impact on packet rates as the IP packet size changes.

Section 2 provides a basic introduction to ATM and the AAL services that can support IP traffic. In Section 3 the IETF approach of [5] is discussed and compared with some alternatives. The match between ATM and "typical" IP packet size is examined in section 4, with some comments on the conclusions reached in [7]. Section 5 discusses the interaction between IP packet size and cell rate limiting, and its effect on IP throughput. Conclusions are given in section 6.

2. ASYNCHRONOUS TRANSFER MODE.

ATM is a packet switched system based on short, fixed length cells. Each cell consists of 48 bytes of user data and a 5 byte header (Figure 1), carrying a Virtual Channel Indicator (VCI) and Virtual Path Indicator (VPI) in its header. Cells are routed through switching nodes, using the combination of VPI and VCI as a label to associate each cell with established virtual connections. (The cell header is described in more detail in recommendation I.361 [8]).

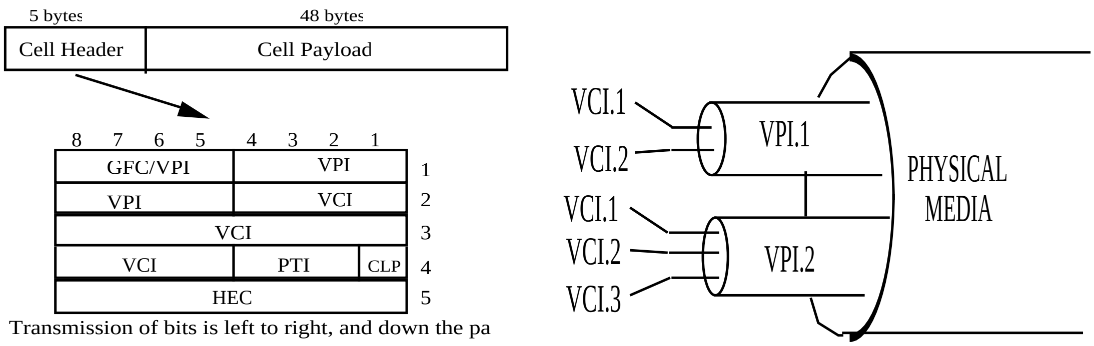

Figure 1

Recommendation I.150 [1] describes ATM as "*a connection-oriented technique. Connection identifiers* [the VPI and VCI] *are assigned to each link of a connection when required, and released when no longer needed.*". An ATM connection consists "*..of the concatenation of ATM layer links in order to provide an end-to-end transfer capability to endpoints.*". Virtual channels are considered to exist within virtual paths, which are themselves unique only within a given physical path [9], also shown in Figure 1. Virtual Channel Connections provide end-to-end cell paths between users of the ATM layer (typically the AALs at either end).

2.1 The ATM Adaptation Layer (AAL).

The ATM layer demultiplexes received cells before passing them up to the AAL (Figure 2). The AAL is split into the Convergence Sublayer (CS) and the Segmentation and Reassembly sublayer (SAR) (described more fully in recommendations I.362 and I.363 [3,4]). SAR functions provide transmission and error detection facilities on a cell by cell basis. CS functions provide transmission and error detection facilities over the 'natural' unit of data utilised by the user service (bytes, bitstreams, or variable length packets).

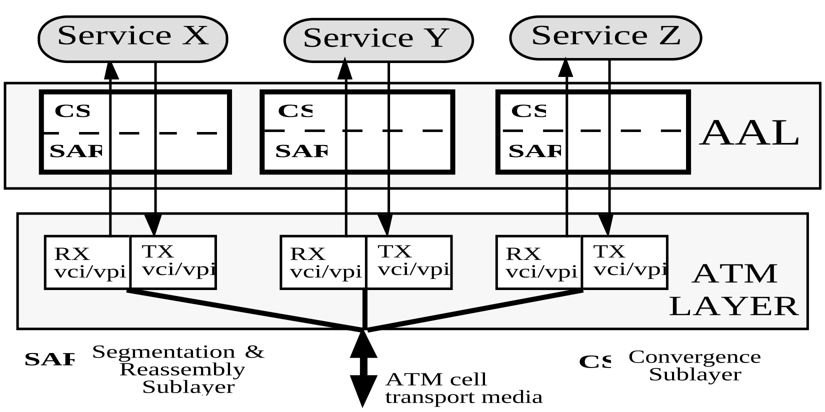

Figure 2

Proc. Australian Broadband Switching and Services Symposium, Wollongong, July 1993. This pdf reconstructed Dec 2005.

Data communication protocols (such as the DARPA Internetworking Protocol , IP [10,11]) are based around arbitrary and unpredictable packet transmission. Support for such services is presently provided by two 'standard' AALs, Type 3/4 (AAL3/4, section 4 of I.363), and Type 5 (AAL5, section 6 of I.363). AAL5 is a product of the datacomms industry work [12, 13], and has been chosen by the IETF to support IP over ATM [5]. We will look at both AAL5 and AAL3/4 for carrying IP, and note some of the reasons for the IETF's decision. In both cases only the basic, or Common Part, of each AAL is used. Thus user data is carried in CPCS_PDUs (Common Part CS Protocol Data Units).

2.2 The AAL5 format.

Figure 3 shows the CS encapsulation used by AAL5. The resultant CPCS_PDU is then segmented into 48 byte blocks and transmitted. IP packets may be transported directly as CPCS_PDU payloads. However, the IETF proposal aims to support multiple protocols across a single virtual connection - a function that AAL5 doesn't support.

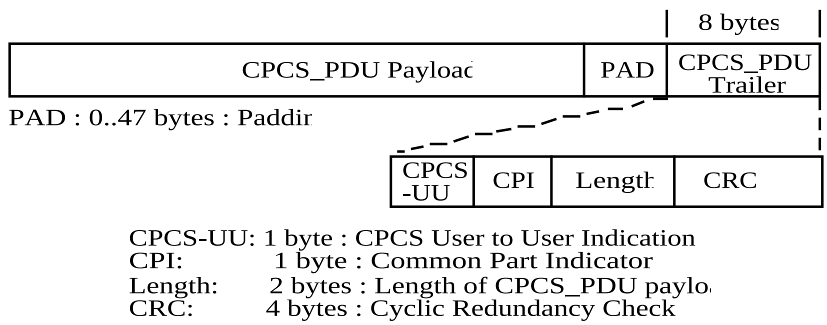

Figure 3

Where dynamic virtual connection establishment is possible, each protocol may open its own 'on demand' - VC based multiplexing. Alternatively, a form of 802.3 Logical Link Control (LLC) encapsulation is proposed, to carry multiple protocols over a single AAL5 connection. Figure 4 shows how an IP packet will be encapsulated as a Routed Non-ISO PDU, EtherType 0x800, to form a CPCS_PDU.

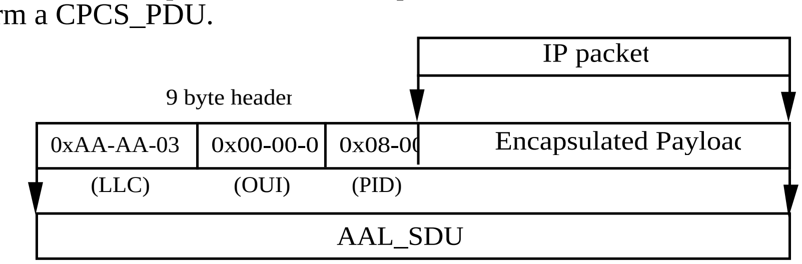

Figure 4

2.3 Comparing AAL3/4 and AAL5 for carrying IP.

Figure 5 shows the CS encapsulation applied by AAL3/4, adding 8 bytes of header and trailer, and upto 3 bytes of padding. The CPCS_PDU is then segmented into 44 byte blocks and further encapsulated by the SAR layer. The complexity of AAL3/4 is due in part to its ability to multiplex different sources of CPCS_PDUs over a single virtual connection, using a 10 bit Multiplex ID (MID) field in each SAR_PDU. Running IETF-encapsulated IP over an AAL3/4 connection would always be less efficient than AAL5. Stripping the 9 byte encapsulation of IP packets, and using AAL3/4's MID field to multiplex different protocols (up to 2^{10}), does not compensate for the AAL3/4 overhead in each cell.

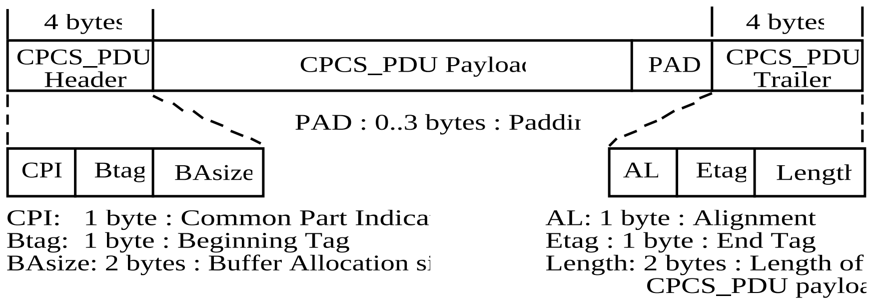

Figure 5

Instead of comparing the '% efficiency' of cell utilisation of each scheme, with their familiar saw-tooth shapes, Figure 6 shows the number of cells needed to transmit an IP packet of a given size. This is a more useful plot because it shows that for small packet sizes, both schemes will still send roughly the same number of cells. However for packets above 200-250 bytes, AAL3/4 begins to use roughly 1 cell more per packet than AAL5. The 'jump point', where cell use goes from 1 to 2 cells, and from 2 to 3 cells, occurs at the following points:

Service	1 to 2 cells	2 to 3 cells
AAL 5 + encapsulation	32 bytes	80 bytes
AAL 3/4 (no encapsulation)	37 bytes	81 bytes

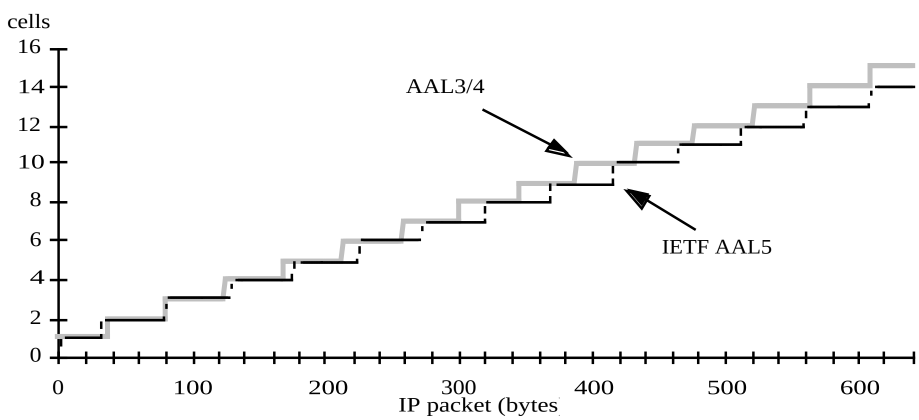

Figure 6

Caceres noted in [6,7] that 'typical' wide area network IP traffic virtually never saw IP and TCP headers longer than 20 bytes each. Thus a zero length TCP payload will generate a 40 byte IP packet - requiring 2 cells under either scheme. The TCP payload may be up to 40 bytes using AAL5, or 41 bytes using AAL3/4, before a third cell needs to be transmitted.

3. TYPICAL IP TRAFFIC AND ATM CELL UTILISATION.

Although taken over 3 years ago the packet traces discussed by Ramon Caceres are still interesting as a guide to 'typical' TCP payload sizes. In [7] the following major characteristics of wide area IP traffic flowing between some major sites and 'the Internet' were identified:

Proc. Australian Broadband Switching and Services Symposium, Wollongong, July 1993. This pdf reconstructed Dec 2005.

> TCP packets account for over 83% of total IP packets, UDP accounts for less than 15%.
>
> Almost 40% of TCP payloads are of zero length., and 30% of all TCP payloads are between 1 and 10 bytes long. 84% of all TCP payloads are under 256 bytes.
>
> 16% of all TCP payloads are around 512 and 536 bytes long.(These are attributed to file transfer applications limited by the default wide area networking Maximum Segment Size (MSS) values configured into most Internet hosts. IP fragments accounted for less than 0.05% of total packets at all three sites tested.)

Given a collection of possible AAL and encapsulation schemes, and the predominance of small TCP payloads, Caceres argues that the 53 bytes ATM cell is a very poor choice of size. His focus is on the ratio of TCP payload bytes to ATM physical layer bytes. As Caceres found that UDP traffic was between 10 and 100 times less significant than TCP traffic (both in terms of packet counts and byte counts), we shall focus only on TCP traffic.

4.1 The issue - How many cells are being used?

This is important because virtual connection costs will be related to the number of cells transmitted. Using Caceres' results we can see that at least 70% of the traffic on an ATM virtual connection will be bursts of 2 cells, regardless of the AAL scheme used. The other major peak around 512 and 536 byte payloads corresponds to IP packets of 552 and 576 bytes respectively. Using AAL5 this leads to 16% of the ATM traffic being 12 and 13 cell bursts respectively. (Using AAL3/4 gives us 13 and 14 cell bursts respectively).

The abundance of small IP packets suggests that header compression, such as RFC 1144 [14], might be explored. The only limitation is implementing suitably fast compression and decompression engines at either end of the virtual connection. The claimed compression is impressive - the combined IP and TCP headers should collapse to 5 bytes. TCP payloads of up to 26 bytes (using AAL5 + encapsulation), and 31 bytes (using AAL 3/4 and no encapsulation) could be carried within a single cell.

The effect on the figures for tiny TCP payloads is dramatic. Now 70% of the traffic on the virtual connection will be single cell messages, half the previous number of cells. The effect on the large TCP payloads is not so dramatic. Both 512 and 536 byte payloads will need 12 cells to be carried on an AAL5 connection, and 13 cells on an AAL3/4 connection.

Around 14% of TCP payloads have lengths between 11 and 256 bytes. Caceres shows a histogram from which a rough visual estimate may be made that the mean payload length in this range is about 40 bytes. Without header compression this corresponds to 80 bytes of IP packet. This is on the edge of 2 or 3 cells for both AAL3/4 and AAL5 techniques, so the pessimistic value of 3 will be used to estimate the number of cells used by this traffic. With header compression only 2 cells are needed. Assuming that all other payload sizes are insignificant, we can roughly estimate the overall impact of header compression.

Over 100,000 TCP payloads without header compression we can expect a total of

100,000 * 0.7 * 2 = 140,000 cells due to tiny payloads, and
100,000 * 0.14 * 3 = 42,000 cells due to 'medium' payloads, and
100,000 * 0.16 * 13 = 208,000 cells due to large payloads.

A total of <u>390,000</u> cells on the virtual connection.

Over 100,000 TCP payloads with header compression we can expect a total of

100,000 * 0.7 * 1 = 70,000 cells due to tiny payloads, and
100,000 * 0.14 * 2 = 20,800 cells due to 'medium' payloads, and
100,000 * 0.16 * 12 = 192,000 cells due to large payloads.

A total of 282,800 cells on the virtual connection.

So RFC 1144 header compression can potentially reduce the total cell traffic to 72.5% of its uncompressed value.

4.2 How important is "inefficient" cell usage?

One concern that has been raised is that IP over ATM will be highly inefficient simply because the majority of packets create wasted space in ATM cells. That 70% of TCP payloads need fractionally more than 1 cell, yet must use 2, has led to overly pessimistic views being aired [15].

The important point to note is that although large TCP payloads account for only 16% of all packets, they account for over 50% of all cells generated on a standard connection. With header compression the difference is even more dramatic, with 68% of cell traffic due to these large payloads. The cell streams generated by these payloads are generally completely full, consequently the cell usage is high. It should be obvious that it is wrong to extract efficiency estimates from the inefficient 2 cell messages. Rather the significant impact of over 50% of cells originating from large TCP payloads must be taken into account.

4.3 Some local experience.

Over almost 7 days, between 12pm 24/5/93 and 2am 31/5/93 we traced network traffic on the main Ethernet of our department. Whilst Caceres focussed on WAN traffic, we monitored server-workstation, workstation-workstation, and workstation-world traffic. Briefly the results were:

Category	TCP	UDP	Compression	Total IP packets
Server-workstation	38.5%	57%	95.2%	20.3 million
Workstation-workstation	44.6%	55.4%	95.9%	11.3 million
Workstation-world	98.4%	1.6%	90.1%	4.6 million

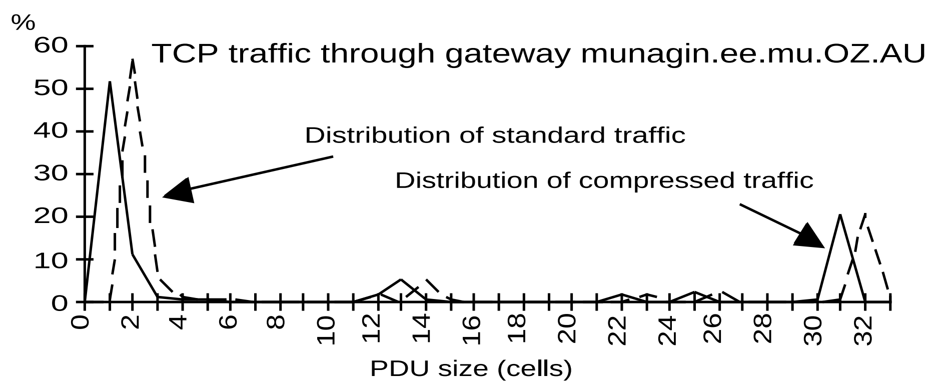

Figure 7

In comparision to Caceres results, our 'WAN' traffic shows similar TCP/UDP ratio but the TCP traffic has a higher component of large packet traffic. Thus whilst the benefits of IP header compression are less dramatic, the overall cell use efficiency improves. Figure 7 shows the percentage of IP packets transmitted vs their size in ATM cells for one of our gateways. Our

Proc. Australian Broadband Switching and Services Symposium, Wollongong, July 1993. This pdf reconstructed Dec 2005.

internal traces show vastly more UDP traffic, which does not benefit from RFC1144 compression. More detailed analysis will be provided during the presentation.

5. RATE CONTROL, AND IT'S IMPACT ON PACKET RATE.

The cell is the smallest unit on which rate limiting will occur [16], whether it occurs within a switch or within the ATM layer of an end node. The achievable packet rate (and perceived bit rate) will simply depend on the number of cells needed per packet. These two criteria can be misused if analysis is careless. Consider an AAL5 connection carrying IETF encapsulated IP, with a rate limit of 10,000 cells per second.

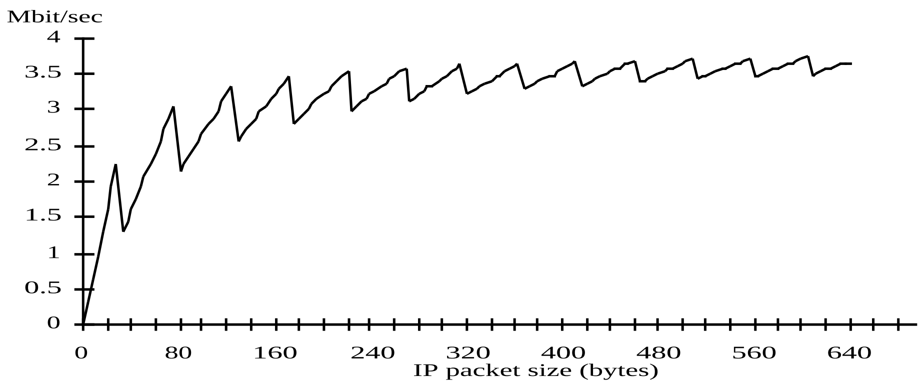

Figure 8

Figure 8 shows the 'achievable throughput' at the IP level. Careful analysis should recognise, however, that whilst throughput appears to drop significantly for small IP packets, this is partly an artefact. Small IP packets carry less information bytes as a percentage of overhead. However, file transfers, etc, that rely on returning ACKs need not be worried about rate limits. Figure 9 shows an alternative view of rate limits - the actual number of packets per second that may be sent.

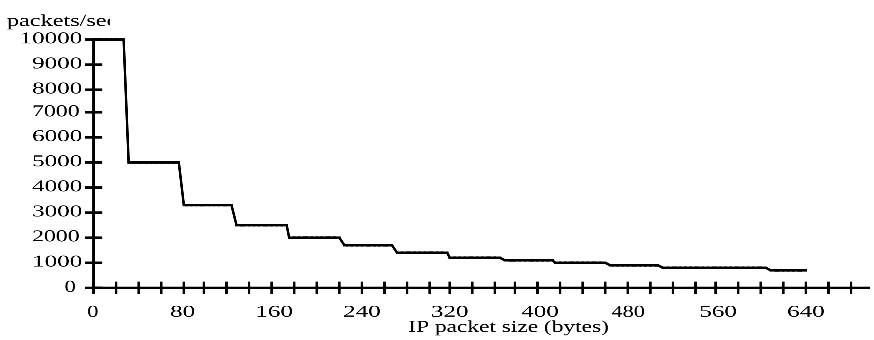

Figure 9

ACKs, and other small packets, may be sent at much higher rates than larger packets. During TCP transfers, the ACK flow will be limited by the large packet flow rate in the reverse direction. Header compression can also help here. TCP payloads between 0 and 10 bytes normally become IP packets between 40 and 50 bytes - limited to 5,000 per second. With compression, they become IP packets between 5 and 15 bytes long - raising the limit to 10,000 per second. This may be an important issue for IP nodes multiplexing hundreds of 'telnet' style connections over one ATM link.

Proc. Australian Broadband Switching and Services Symposium, Wollongong, July 1993. This pdf reconstructed Dec 2005.

6. CONCLUSION.

We have provided an overview of Asynchronous Transfer Mode, and described its basic architecture of Virtual Paths, Virtual Channels, and Adaptation Layers. The IETF's intended use of AAL5 with encapsulation has been compared with the use of AAL3/4 to carry multiprotocol traffic. The benefits of using the MID field of an AAL3/4 connection as an alternative toAAL5+IETF encapsulation is shown to be marginal, and limited to small IP packet sizes.

Plots of cell use efficiency versus IP packet size are not as important as the number of cells transmitted for a given IP packet size. "Efficiency" figures lead to pessimistic analysis of a virtual connection's ability to carry IP traffic. RFC 1144 header compression can potentially reduce the number of cells transmitted over a virtual connection to almost 72.5% of pre-compression values - an important point when connection usage charges are on a per cell rather than per bit basis.

We suggest that small IP packets, and their originating applications, are not seriously affected by cell rate limiting, despite what a "bitrate vs packet size" plot may suggest. Many small IP packets are from remote logins and associated ACKs. The ACKs from file transfer applications will be limited to packet transmission rate of the data packets, much lower than the rate limit theoretically imposed on small IP packets. RFC1144 header compression is again seen to significantly reduce the effect of cell rate limits on 'small' IP packets.

REFERENCES.

[1] CCITT SGXVIII - Report R109, Draft recommendation I.150, "B-ISDN Asynchronous Transfer Mode Functional Characteristics", July 1992.

[2] Martin de Prycker, "Asynchronous Transfer Mode - Solution for Broadband ISDN", Ellis Horwood, 1991.

[3] CCITT SGXVIII Draft recommendation I.362, "B-ISDN ATM Adaptation Layer (AAL) functional description", June 1992.

[4] CCITT SGXVIII Draft recommendation I.363, "B-ISDN ATM Adaptation Layer (AAL) specification", January 1993.

[5] Juha Heinanen, "Multiprotocol Encapsulation over ATM Adaptation Layer 5", IETF Network Working Group, DRAFT Request For Comment, February 24, 1992.

[6] Ramon Caceres, "Measurements of Wide Area Internet Traffic", University of California Berkeley, Technical Report No. UCB/CSD 89/550, December 1989.

[7] Ramon Caceres, "Efficiency of ATM Networks in Transporting Wide-area Data Traffic", submission to Computer Networks and ISDN systems, December 1991.

[8] CCITT SGXVIII Revised recommendation I.361, "B-ISDN ATM Layer specification", June 1992.

[9] CCITT SGXVIII - Report R111, Draft recommendation I.311, "B-ISDN General Network Aspects", July 1992.

[10] Jon Postel, "Internet Protocol - DARPA Internet Program Protocol Specification", RFC 791, September 1981, Section 1.4.

[11] D. Comer, "Internetworking with TCP/IP - Principals, Protocols, and Architecture", Prentice Hall, 1988.

[12] ANSI T1S1.5/91-449, "AAL5 - A New High Speed Data Transfer AAL.", ANSI T1S1, November 1991.

[13] Apple Computer, Bellcore, Sun Microsystems, Xerox, "Network Compatible ATM for Local Network Applications, Phase 1", Version 1.0, April 1992.

[14] Van Jacobson, "Compressing TCP/IP Headers for Low-Speed Serial Links", RFC 1144, February 1990.

[15] Open discussion, "IP-over-ATM", Usenet newsgroup comp.dcom.cell-relay, February/March 1993.

[16] CCITT SGXVIII Draft recommendation I.371, "Traffic Control and Congestion Control in B-ISDN", June 1992.

ACKNOWLEDGEMENT.

The assistance of AOTC through Telecom Research Laboratories research contract 7064 is hereby acknowledged.

Proc. Australian Broadband Switching and Services Symposium, Wollongong, July 1993. This pdf reconstructed Dec 2005.