Analysis of Traffic Measurements in the VISTAnet Gigabit Networking Testbed

D. S. Holtsinger
H. G. Perros
A. A. Nilsson

Center for Communications and Signal Processing
Department of Electrical and Computer Engineering
North Carolina State University

TR-93/12
July 1993

Analysis of Traffic Measurements in the VISTAnet Gigabit Networking Testbed[1]

Douglas S. Holtsinger†, Harry G. Perros‡, Arne A. Nilsson†

†Department of Electrical and Computer Engineering, and
Center for Communications and Signal Processing
North Carolina State University
Raleigh, NC 27695

‡Department of Computer Science, and
Center for Communications and Signal Processing
North Carolina State University
Raleigh, NC 27695

Abstract

The characterization of ATM traffic is particularly important for developing and validating traffic models for the analysis of congestion control mechanisms, communication network protocols, and architectures for ATM networks. Unfortunately, little is known about the actual behavior of ATM traffic, since few ATM networks have been built and instrumented for measuring traffic. In this paper, we present an analysis of traffic mea-

[1]Supported in part by BellSouth, GTE Corporation, and NSF and DARPA under cooperative agreement NCR-8919038 with the Corporation for National Research Initiative.

surements from a medical application called Dynamic Radiation Therapy Planning which runs in a distributed fashion on the VISTAnet gigabit networking testbed. We believe that ours is one of the first measurements of traffic from a real application operating in a high-speed networking environment employing ATM technology. We show that our traffic measurements fit neatly into a proposed hierarchical traffic model which describes the traffic source behavior at different time-scales.

1 Introduction

VISTAnet [11] is a three-year research project focused on gigabit communications, meta-computing, advanced visualization techniques, and medical treatment planning. The overall goal is to advance the state-of-the-art in gigabit networking and its application. The project is being carried out jointly by BellSouth, GTE, MCNC, North Carolina State University (NCSU), and the University of North Carolina at Chapel Hill (UNC-CH).

As part of the VISTAnet project, an experimental gigabit network has been deployed for the purposes of experimenting with and evaluating gigabit networking techniques. A diagram of this gigabit network and the terminal equipment is shown in figure 1. The network implements Asynchronous Transfer Mode (ATM) technology, which is the underlying transport mode of choice for the Broadband Integrated Services Digital Network (B-ISDN) [14]. Computing platforms are interconnected in a physical star topology via optical fiber links operating at 622.08 Mb/s.

The traffic measurements which we have obtained from the VISTAnet testbed were taken during the execution of an interactive application called Dynamic Radiation Therapy Planning (DRTP), which is the driving application of VISTAnet [2]. The goal of the research underlying the development of the DRTP application is to improve the planning techniques used for radiation therapy on cancer patients.

The application runs in a distributed fashion on the heterogeneous computing platforms connected to the gigabit network. To meet the computational goals, the machine characteristics are matched with the computational requirements of the tasks which compose the DRTP application. None of the computing platforms individually can solve the problem while simultaneously meeting the interactive response times which are required from the application. In effect, the gigabit network acts as a computer "backplane" for a "metacomputer" consisting of several heterogeneous machines which at times can place sufficient demands on the network to saturate the capacity of the 622.08 Mb/s optical fiber links [11].

The importance of characterizing traffic in high-speed computer communication networks is critical, since many of the proposed congestion control mechanisms, network protocols, and architectures for high-speed networks contain broad assumptions about the behavior of traffic sources. These assumptions need to be validated through actual traffic measurements in high-speed networks. Very little is known about the actual behavior of traffic in high- speed communication networks, mostly because there are so few experimental testbeds available to perform measurements. We believe that ours is one of the first measurements of traffic from a real application operating in a high-speed networking environment employing ATM technology.

The outline of this paper is as follows. In section 2, we discuss the VISTAnet network topology, architecture, and interfaces. In section 3, we discuss the Dynamic Radiation Therapy Planning application and the computational role of each computing platform, and in section 4 we present the traffic measurements from the application. Finally in section 5, we present a summary and our conclusions.

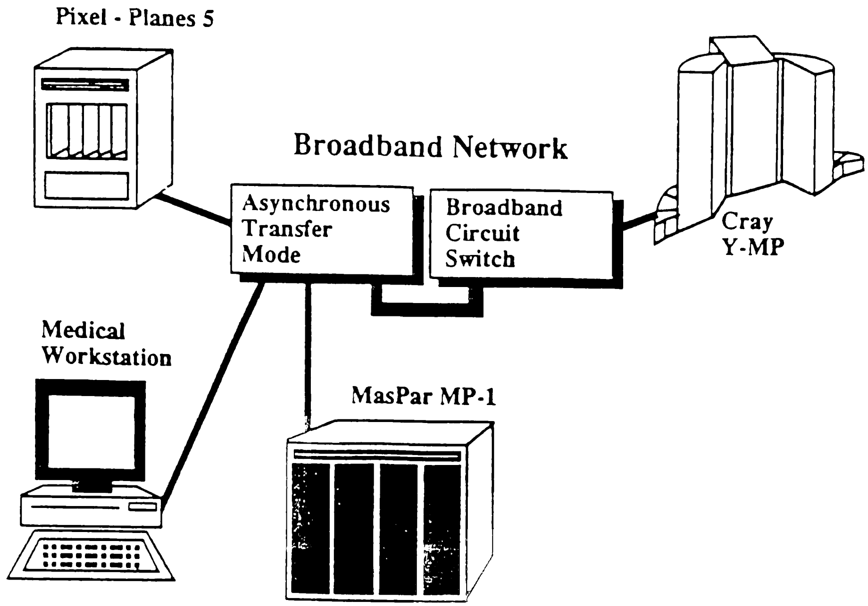

Figure 1: VISTAnet Gigabit Testbed Diagram

2 VISTAnet Network Facilities

2.1 VISTAnet Network Topology

In this section, we briefly discuss the network topology of the VISTAnet testbed, which is illustrated in figure 2. A Fujitsu FETEX-150 multiservice switching platform [4, 7, 12] provides ATM switching services in the network. The host ATM switch is implemented using self-routing modules in a multi-stage network and provides switching functions and serves as the center for call processing and OAM&P. Two broadband remote switching units (BRSU) in the network contain the customer interfaces and perform line concentration functions. A broadband cross-connect switch developed by GTE is located at the Parkwood CO. Three customer sites connect to the ATM network in a physical star configuration via SONET fiber links operating at 622.08 Mb/s, and they are geographically distributed over a maximum distance of approximately 18 km.

At the MCNC site, a HIPPI interface connects a Cray Y-MP 8/432 supercomputer to a local HIPPI switch. The HIPPI interface is a simplex point-to-point link which can transmit digital data at either 800 Mb/s or 1600 Mb/s peak data rates over multiple twisted-pair copper cabling at distances up to 25 m [8]. The VISTAnet testbed uses two simplex 800 Mb/s HIPPI links to provide a duplex connection to the Cray. A Terminal Adapter developed jointly by Fujitsu and BellSouth converts the HIPPI interface into a SONET/ATM interface [2].

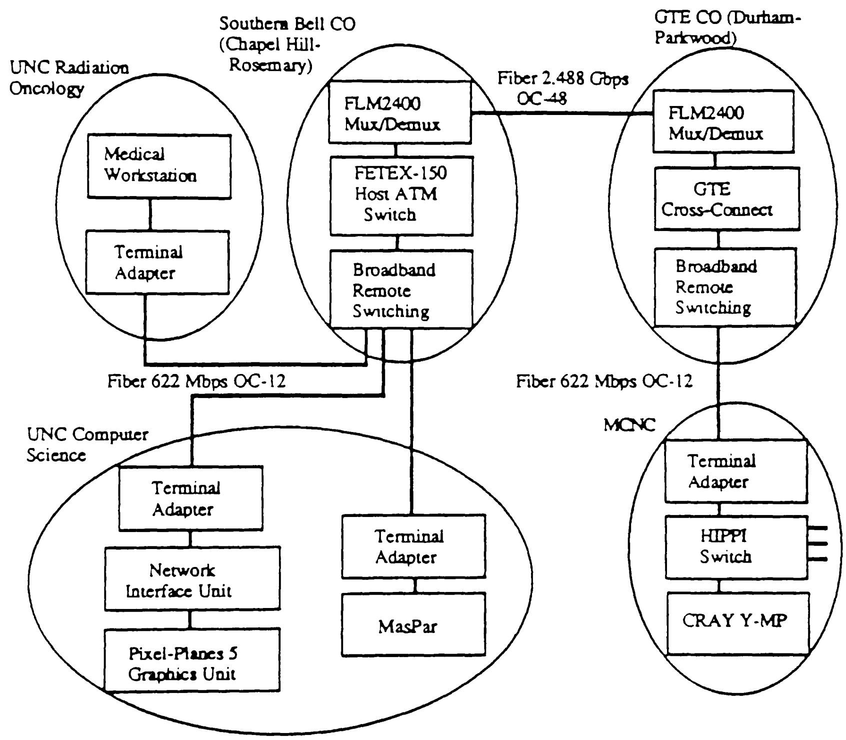

Figure 2: VISTAnet Network Topology

At the UNC Radiation Oncology site, a Silicon Graphics 340 VGX workstation provides the user interface for the Dynamic Radiation Therapy Planning application, and connects to the ATM network through a Terminal Adapter. A board developed by MCNC plugs into the workstation and provides a HIPPI interface to the Terminal Adapter.

At UNC-CH in the Department of Computer Science, a Pixel-Planes 5 graphics unit [1] developed by UNC–CH, and a MasPar MP-1 massively parallel computer each connect to the ATM network via their respective interfaces. The Network Interface Unit [10], developed by UNC–CH, attaches to the communications ring internal to the Pixel-Planes and communicates with the Terminal Adapter via a 800 Mb/s HIPPI interface. In the following section, we discuss the HIPPI interface and the Terminal Adapter in more detail.

2.2 VISTAnet Network Interfaces

Figure 3 shows a simplified diagram of the HIPPI interface and the conversion of HIPPI packets into ATM cells by the Terminal Adapter at the MCNC site. All of our measurements were taken at the MCNC site. The HIPPI link is used by the Cray to communicate with the Terminal Adapter, which provides a customer interface to the ATM network. At the two other customer sites, additional Terminal Adapters connected to the ATM network communicate with terminal equipment through a HIPPI interface.

The Terminal Adapters near the two ends of a virtual connection each manage a duplex HIPPI interface. The TAs communicate between themselves to provide a single transparent HIPPI interface to the attached terminal equipment. The existence of the Terminal Adapters and the ATM network is not known by the terminal equipment, which only sees a single interface.

Before a packet may be transmitted across the interface, the HIPPI source must set up a HIPPI connection. It does this by asserting a REQUEST indicator, and a connection becomes established when the destination returns a CONNECT indicator to the source. After the connection is established, the source asserts the I-field onto the interface, which indicates the packet destination. In figure 3, the Cray is shown as the source and the Terminal Adapter is shown as the destination for data flowing from the Cray to the ATM network.

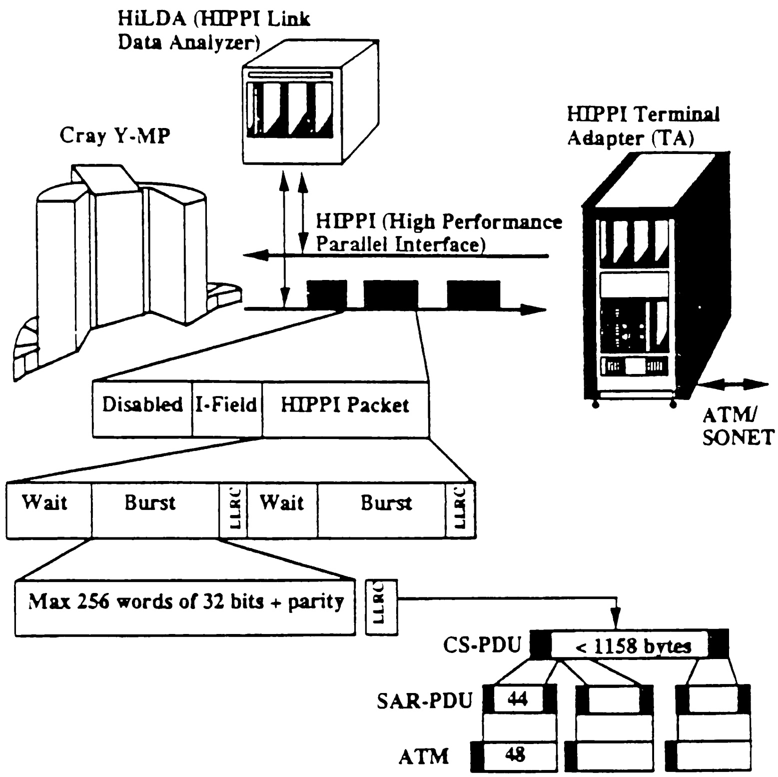

Figure 3: HIPPI Interface at MCNC

End-to-end virtual connections in VISTAnet are currently established on a semi-permanent basis. The VISTAnet project implements a restriction of one packet transmission per HIPPI connection. (This latter restriction does not hinder performance as the source Terminal Adapter does not wait for the destination terminal equipment on the other side of the ATM network to respond before establishing a connection).

Packets are transmitted by the HIPPI interface as a series of bursts, and each burst may contain up to 256 words of 32 bits plus 4 parity bits. Only one burst in a packet may contain less than 256 words. Words are always transmitted consecutively at 40ns clock cycle intervals during a burst. At the end of every burst, a 32 bit Length-Longitudinal Redundancy Check (LLRC) is transmitted which implements even parity across the individual bits of multiple words in a burst. Between the transmission of bursts within a packet there may be arbitrarily long waiting periods where no data is transmitted.

The Terminal Adapter (TA) forms a CS-PDU from each burst plus the LLRC. The TA segments the CS-PDU into Type-4 AAL cells, and transmits the cells through the ATM network to the destination TA, which then reverses the segmentation process and reassembles the cells into bursts. A full burst of 1Kbytes plus parity and LLRC translates into 27 ATM cells. Special Type-4 AAL cells are used for indicating the beginning of a packet, flow control, and connection establishment.

Each Terminal Adapter contains a 32 KByte buffer for holding a maximum of 31 bursts awaiting transmission into the ATM network. The TAs also contain two 32 KByte buffers for holding cells which arrive from the ATM network.

Flow control is implemented across the interface by requiring the destination to transmit a single READY indication to the source for each burst that it is prepared to accept. For each READY indication received, the source has permission to send one burst. After the connection is first established, the destination transmits READY indicators to the source indicating the amount of its available buffer space.

A window flow control algorithm operates between the source and destination TAs. The MID field of the AAL SAR header is used to identify individual bursts. The source TA increments the MID field after the transmission of every burst, modulo 1024. After the destination TA receives N1 bursts from the source TA (where N1 is an adjustable parameter), it sends to the source TA the MID field of the last burst number received and the amount of available buffer space at the destination TA. The source TA subtracts the last MID field received from the MID of the last burst which the source transmitted (modulo 1024). The source TA then subtracts the result from the amount of available buffer space at the destination TA to obtain the number of bursts which the source TA may transmit. The destination TA's transmission of the last MID received and buffer space information resets a timeout timer. The timer triggers the transmission of this information after T1 seconds (where T1 is adjustable).

Traffic measurement facilities are contained in the HIPPI Link Data Analyzer (HILDA) developed by MCNC, which has a variety of capabilities and modes [13]. We used HILDA to measure the packet lengths in bytes and the packet interarrival times in HIPPI clock cycles. In addition, the Terminal Adapter contains measurement facilities for capturing the pattern of ATM cell transmission and reception at the interface to the ATM network [1]. We used both of these facilities to measure the radiation dose traffic which is transmitted by the Cray Y-MP to the Pixel-Planes processor, which constitutes much of the traffic generated by the application.

3 Dynamic Radiation Therapy Planning

Radiation treatment involves delivering a high and uniform radiation dose to a tumor while also limiting the dose intensity to nontumor-bearing tissue. The restriction of treatment planning to two dimensions can result in a suboptimal treatment plan. Medical researchers at UNC believe that treatment planning can be greatly improved by allowing the radiation oncologist to interactively optimize the treatment plan by viewing the radiation dosage distribution overlayed on a three dimensional representation of the patient's anatomy [2].

Figure 4 illustrates the flow of information between computing platforms during the execution of the Dynamic Radiation Therapy Planning application. A three dimensional representation of the patient's anatomy is created offline with multiple CAT scan or MRI images. The radiation treatment plan is entered at the workstation, where it may be modified interactively during the execution of the application. These modifications may either result from a single user change to the treatment plan, or a series of closely-spaced changes to the plan resulting from the user dragging an interface control tool at the medical workstation.

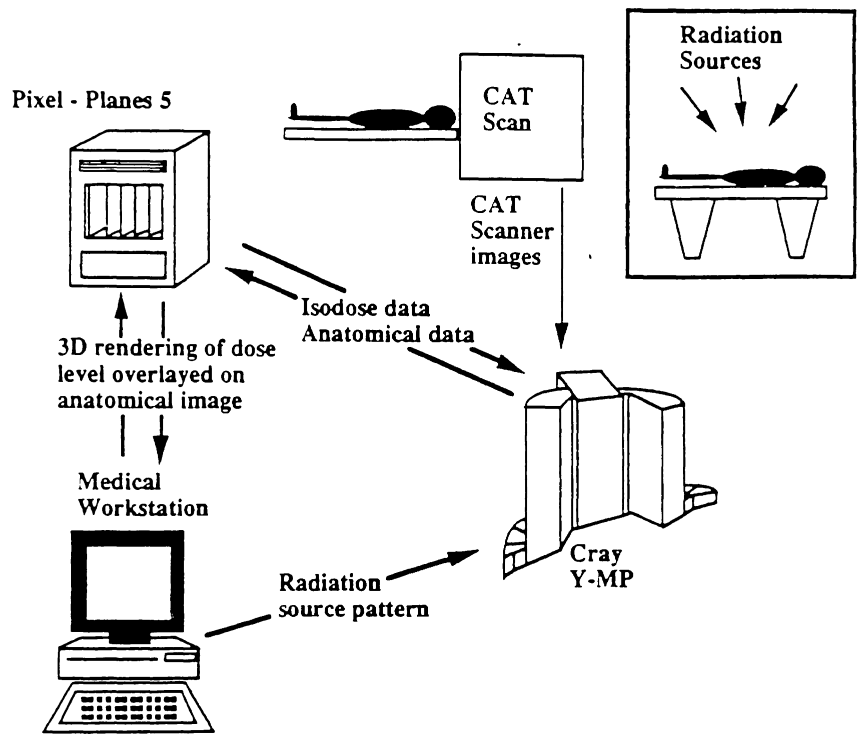

Figure 4: Dynamic Radiation Therapy Planning

Modifications to the treatment plan are sent to the Cray Y-MP, which re-calculates the radiation dose distribution and sends it to the Pixel-Planes for processing. The Pixel-Planes overlays the three dimensional radiation dose distribution on the patient's anatomical image, and then transmits the resulting two-dimensional perspective image to the medical workstation for graphical display. Viewpoint changes to the image are transmitted to the Pixel-Planes processor, which determines the new perspective image and sends it to the medical workstation for display. While not shown in figure 4, the MasPar will be used to compute the absorbed radiation dose for various anatomical structures, and it will send its output to the medical workstation for numerical display.

The application demands high throughput from the network in order to achieve low latency in delivering the entire dose dataset, and long silence periods occur during which no traffic is generated, which results in very low link utilizations. As a result, we expect this application to place severe demands on network resource allocation and congestion control mechanisms.

At present, there is no higher-level protocol which operates between the Cray and the Network Interface Unit (NIU), other than the HIPPI-FP layer [9], and the RingP delivery service layer which runs on top of the HIPPI-FP layer. The RingP layer [10] encapsulates one or more variable-length ring messages into a single packet to be transmitted by the HIPPI-FP layer, which is an unreliable connectionless datagram service. These ring messages are injected directly into the Pixel-Planes communications ring by the attached NIU. Essentially, the Cray uses the Pixel Planes as a dumb I/O device. Experimentation with a full protocol stack is planned [10].

No per-packet acknowledgement scheme is done at any protocol level, except for a single acknowledgement sent by the Pixel-Planes to the Cray indicating that it is ready to receive another dose. A window flow control algorithm operates between the two TAs, and HIPPI flow control is performed between the Cray and the source terminal adapter, and between the destination terminal adapter and the Network Interface Unit attached to the Pixel-Planes. Corrupted packets are simply dropped by the NIU.

In the next section, we present our measurements of the radiation dose traffic which is transmitted by the Cray Y-MP to the Pixel-Planes processor.

4 Analysis of Traffic Measurements

4.1 Hierarchical Traffic Description

We have found that the ATM traffic measurements from VISTAnet fit neatly into a hierarchical traffic description, and that each level of the hierarchy corresponds with a particular aspect of the operation of the Dynamic Radiation Therapy Planning application in VISTAnet. Hierarchical traffic descriptions have been explored by a number of authors [3, 6].

Figure 5 shows our proposed hierarchical traffic description of the measurements taken from the DRTP application. At lowest level of the hierarchy (cell-level), the ATM traffic behavior is largely dependent upon the conversion of HIPPI bursts into ATM cells by the Terminal Adapter. The time-scale is measured in terms of ATM slot durations at the OC-12 transmission rate (microseconds).

At the next higher level (burst-level), the ATM traffic behavior is dependent upon the software algorithm used to transmit the dose dataset at the Cray, the operation of the Cray's I/O system in transmitting HIPPI packets, and the packet size as selected by the application programmer. The time-scale is measured in terms of packet interarrival times (hundreds of microseconds to several milliseconds).

At the next higher layer (dose-level), the traffic behavior depends upon on the user interaction with the application, which determines the interarrival time of dose datasets to the Terminal Adapter (hundreds of milliseconds up to several minutes). At the highest layer (session-level), the traffic is described in terms of the dose dataset size and the duration of a session. There could be additional higher layers which describe the utilization of the application by medical personnel (hours, days). In the next section, we describe the

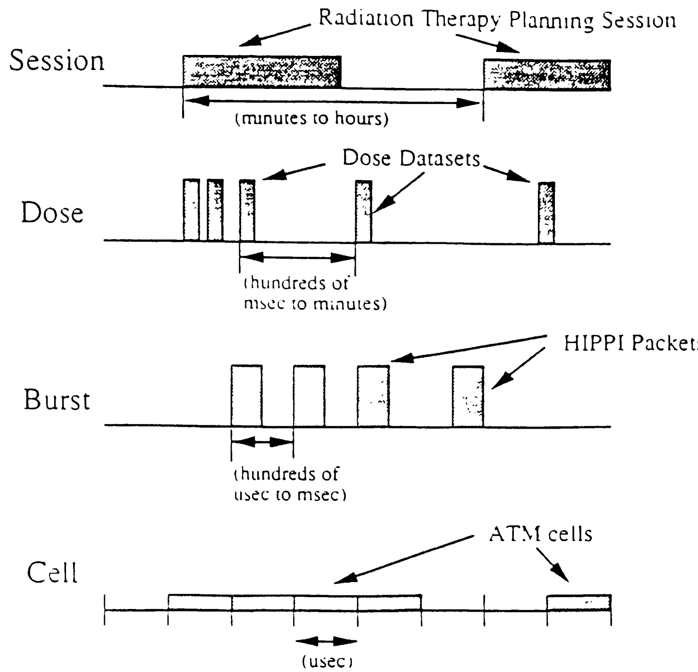

Figure 5: Hierarchical Traffic Description

traffic characteristics at the highest level shown in figure 5, and the sections which follow contain analyses of traffic behavior at the lower levels.

4.2 Session-level Traffic Characteristics

The radiation dose distribution calculated by the Cray Y-MP has a fixed size per session. The dose dataset may change in size after each time the application is invoked by the user, which marks the beginning of a new session.

Table 1 gives the measured dose dataset sizes for sessions occurring on two different days (further measurements can be found in [5]). The sessions are listed in chronological order. The first column shows the number of bytes transported across the HIPPI interface in a single dose (including overhead), and the number of ATM cells necessary to transport a single dose. Also shown in the table is the ATM link utilization during the session, which is defined as the total transmission time of the ATM traffic during a session divided by the duration of the session.

The rate at which the dose datasets are sent from the Cray to the Pixel Planes processor varies from approximately one every minute to one every four seconds. Even at this latter rate, the link utilization barely exceeds 10^{-3} despite dose dataset sizes of almost 1 Mbyte.

Day	Dose Size	Doses per Session	Transmission Time (sec.)	Session Duration (sec.)	Link Utilization
3/11	919 Kb (24238 cells)	194	3.33	9582	3.48×10^{-4}
	435 Kb (11477)	115	0.93	3474	2.68×10^{-4}
	435 Kb (11470)	75	0.61	343	1.78×10^{-3}
	919 Kb (24238)	149	2.57	2178	1.18×10^{-3}
3/18	981 Kb (25876)	316	5.58	7680	7.26×10^{-4}

Table 1: Link Utilization and Dose Dataset Sizes

4.3 Dose-level Traffic Characteristics

Figure 6 shows the first few seconds of the ATM traffic rate as a function of time at the MCNC Terminal Adapter for measurements collected on 3/18/93. The traffic rate is defined here as the total amount of ATM traffic (number of bits, including cell header) arriving over a fixed interval of N ATM cell slots at the OC-12 rate, divided by the duration of the interval in seconds. In the figure, N is 400,000 ATM slots, or approximately 280 ms.

Each thin vertical dark line in the figure represents the traffic rate averaged over a duration of 400,000 ATM slots. The thicker dark areas represent periods in which the traffic rate is non-zero over several consecutive 400,000 slot periods. A single "dose period", or interval of time in which an entire dose dataset is transmitted by the Cray to the Pixel Planes processor, can be observed in figure 6 as a single thin vertical dark line.

Figure 7 shows a histogram of the silence period duration in ATM slots between dose the dataset arrivals. The silence period duration is the number of ATM slots between the transmission of the last ATM cell in a dose dataset and the first ATM cell in the next dose dataset. The X-axis labels are centered directly under the tick which corresponds to the beginning of a bin. The number "N" in each histogram represents the total number of samples. The far right-most bin and the far left-most bin represent outlier bins. The shaded area represents the fraction of interarrival times which are outliers.

Between the transmission of doses, a small number of control packets are transmitted by the Terminal Adapter. The silence period duration tends to cluster at around 1.3E6 ATM slots, or about 900 ms. Several silence periods exceeding 15 minutes duration have been observed in the measurements.

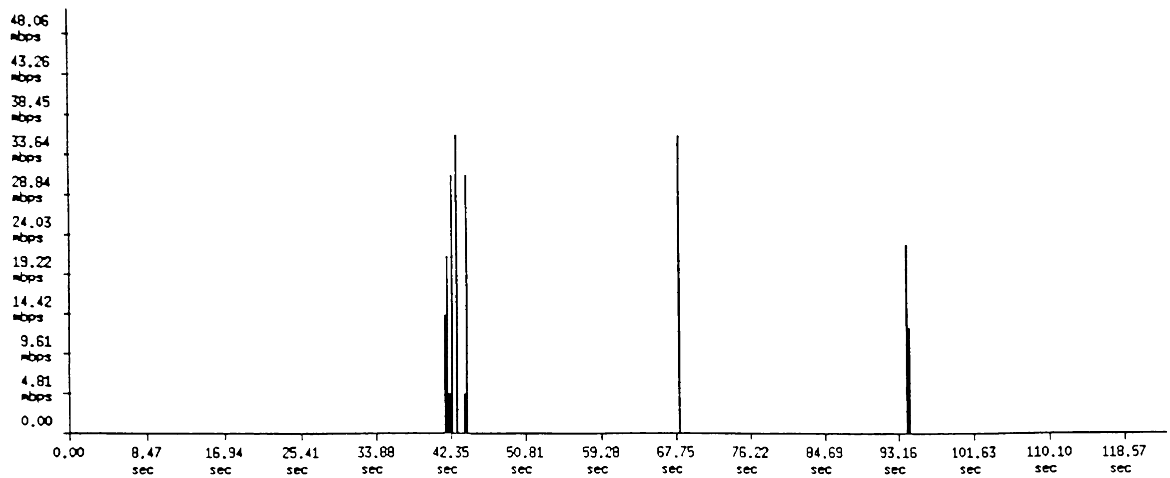

Figure 6: Traffic rate, 3/18/93, N=400,000 ATM slots

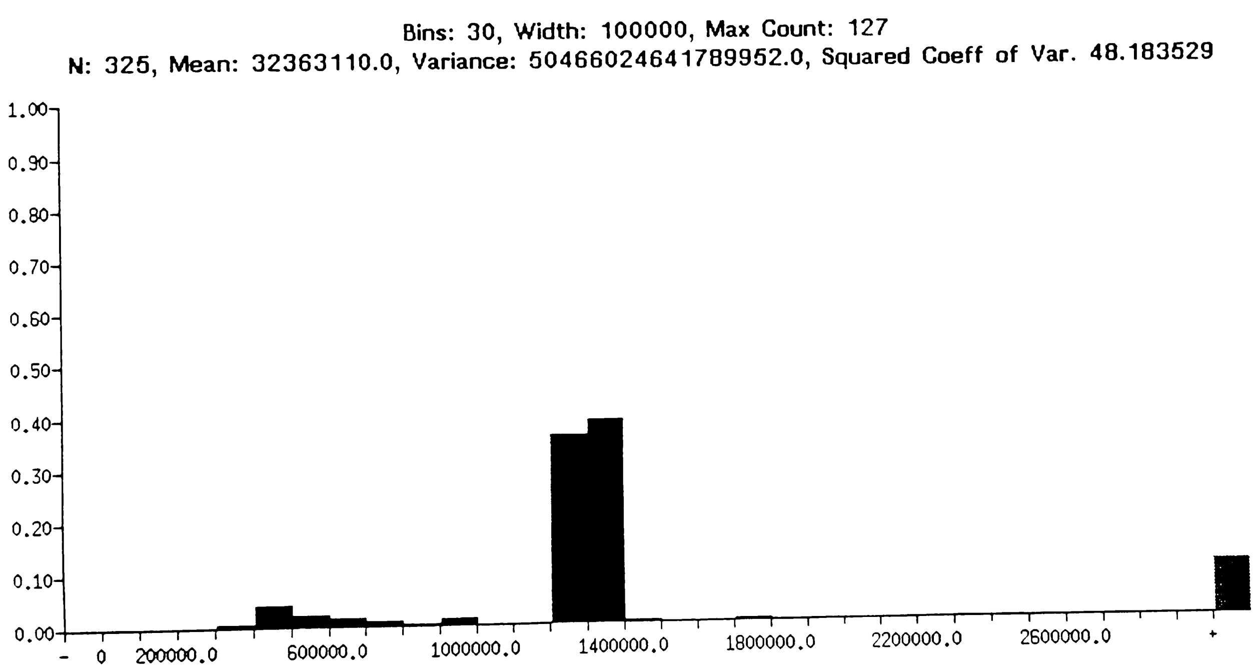

Figure 7: Inter-Dose Silence Period Distribution in ATM slots, 3/18/93

4.4 Burst and Cell-level Traffic Characteristics

Figure 8 shows a closeup view of the measured ATM traffic rate during a typical dose period, with the traffic rate averaged over $N = 125$ ATM slots, or roughly 88 usec. Preceding the transmission of the dose are 8 HIPPI control packets of size 24 bytes. During this session, a dose dataset is transmitted by the Cray as a series of 15 packets of size 61576 bytes each (or 1624 ATM cells) plus one packet of size 57480 bytes (1516 ATM cells) as the last packet transmitted in the dose. The Terminal Adapter always converts packets into continguous bursts of ATM cells. For example, a burst of ATM cells corresponding to a single HIPPI burst can be seen in figure 8 at time 17.67 msec relative from the origin.

The pattern of ATM cell transmission shown in figure 8 is a consequence of the generally periodic packet arrivals from the Cray supercomputer. After the radiation dose distribution is calculated by the Cray, the dataset is divided up into fixed size packets whose size is selected by the application programmer. The application copies a packet's worth of dose data into a temporary packet buffer during each iteration of a 'while' loop, and submits the packet to the Cray I/O subsystem for transmission across the ATM network.. The packet size has been demonstrated to have a significant influence on the Cray's packet throughput [11], with larger packet sizes producing greater throughput.

Figure 9 illustrates a scatter plot of the 64 Kb packet interarrival time in ATM slots within a dose, which clearly shows the periodic nature of packet arrivals. Figure 10 shows a histogram of the 64 Kb packet interarrival time distribution in ATM slots within a dose for the traffic session on 3/18/93. Roughly 10% of the packet interarrival times exceed 5200 ATM slots (3.67 ms) in this session, and the remainder are smoothly distributed between 1900 ATM slots (1.34 ms) and 5200 ATM slots.

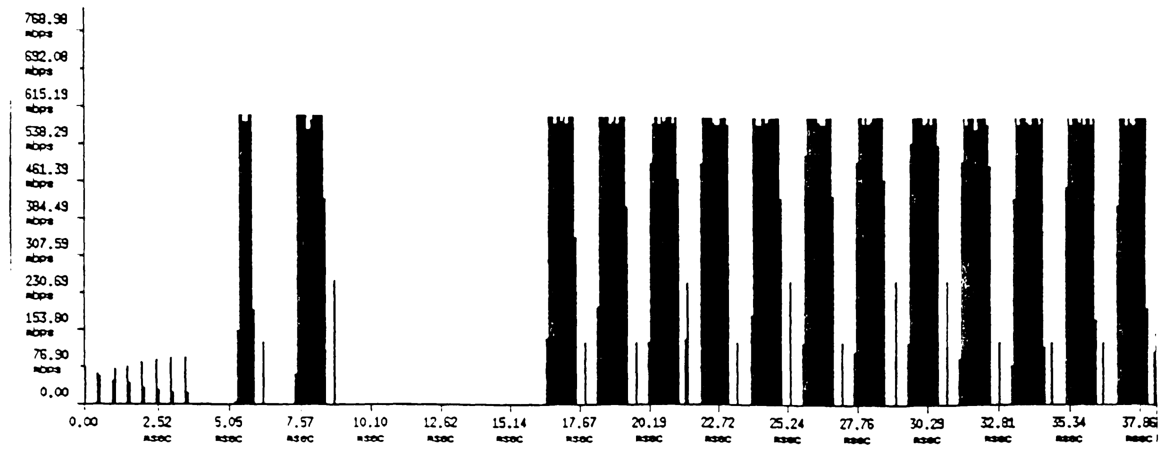

Figure 8: Traffic rate, 3/18/93, N=125 ATM slots

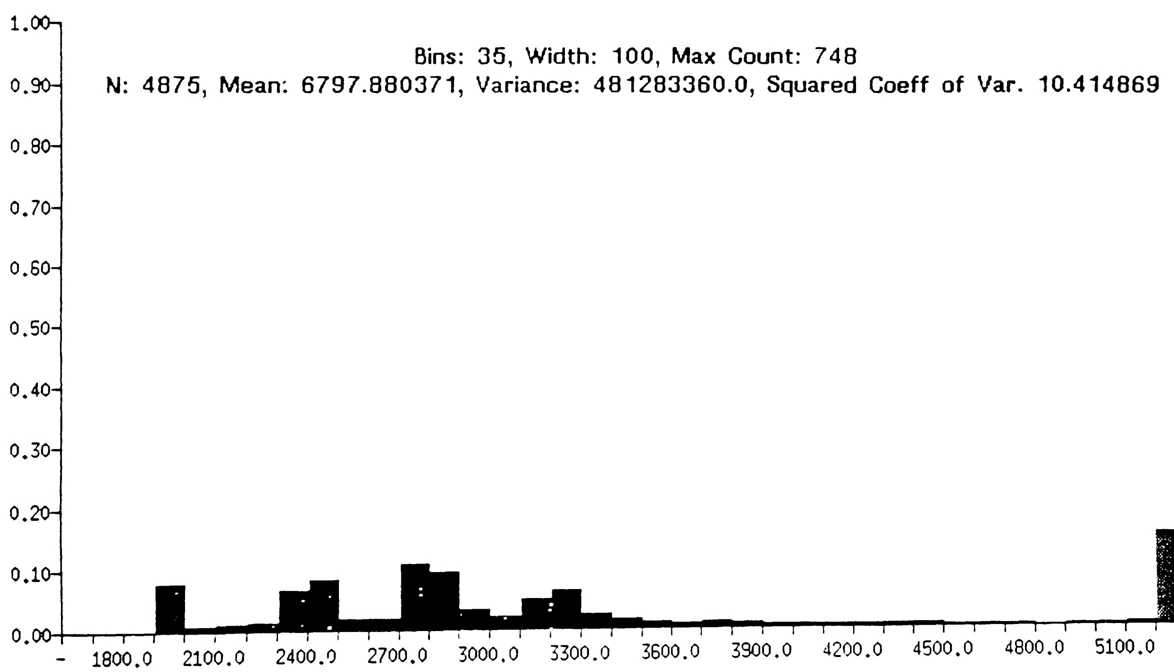

Figure 9: HIPPI Packet Interarrival Time in ATM slots, 3/18/93

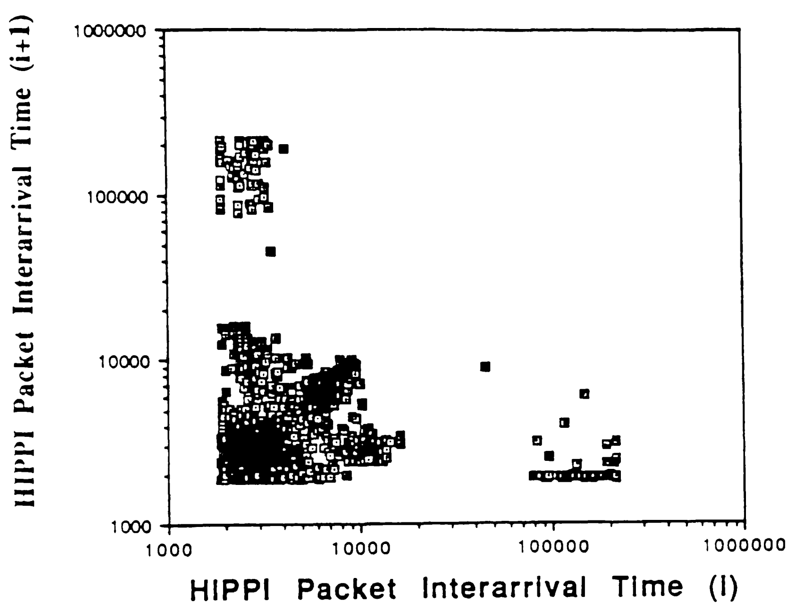

Figure 10: HIPPI Packet Interarrival Time Distribution in ATM slots, 3/18/93

5 Summary of Results and Conclusion

Our results indicate that the traffic characteristics are dependent on the time-scale at which they are observed. At the time-scale of microseconds, little traffic variation can be observed, since the HIPPI bursts are of fixed length, which produce ATM cell bursts of fixed length. At the next larger time-scale, we observe generally periodic arrivals of HIPPI packets arriving at intervals on the order of several milliseconds.

At the dose-level, which is on the time-scale of hundreds of milliseconds to several minutes, there is more traffic variation (squared coefficient of variation of silence duration of about 48), and dose arrivals and dose dataset sizes are mostly a function of the user's interaction with the application.

We have proposed a hierarchical traffic description and have attempted to account for the behavior of the source in terms of the operational characteristics of the Dynamic Radiation Therapy Planning application in the VISTAnet gigabit testbed.

References

[1] B.E. Basch, W.A. Bruwer, D.D. Casey, W.L. Smith, and D.R. Spears. VISTAnet: A BISDN field trial. *IEEE LTS*, pages 22–30, Aug. 1991.

[2] K. Chipman, P. Holzworth, J. Loop, N. Ransom, D. Spears, and B. Thompson. Medical applications in a B-ISDN field trial. *IEEE Journal on Selected Areas in Communications*, 7:1173–1187, September 1992.

[3] O. Gihr and P. Tran-Gia. A layered description of ATM cell traffic streams and correlation analysis. In *IEEE Globecom*, pages 137–144, 1991.

[4] K. Hajikano, K. Murakami, E. Iwabuchi, O. Osono, and T. Kobayashi. Asynchronous transfer mode switching architecture for broadband ISDN. In *Proc. of the ICC*, page 29.3, June 1988.

[5] D.S. Holtsinger, H.G. Perros, and A.A. Nilsson. Analysis of VISTAnet traffic measurements. Technical report, North Carolina State University, Center for Communications and Signal Processing.

[6] R. Jain and S.A. Routhier. Packet trains – measurements and a new model for computer network traffic. *IEEE Journal on Selected Areas in Communications*, 4(6):986–995, September 1986.

[7] Y. Kato, H. Hayami, J. Kamoi, K. Murakami, M. Takeno, and S. Hattori. A VLSIC for the ATM switching system. In *Proc. of the ISS*, May 1990.

[8] ANSI Preliminary Draft Proposal. High-performance parallel interface, mechanical, electrical, and signalling protocol specification, X3.183-199x. June 1991.

[9] ANSI Preliminary Draft Proposal. High-performance parallel interface framing protocol (HIPPI-FP), X3.210-199x. February 1992.

[10] R. Singh, S. Tell, and D. Becker. VISTAnet network interface unit: prototype system specifications. Technical Report TR91-017, University of North Carolina, Chapel Hill, Department of Computer Science.

[11] D.S. Stevenson and J.G. Rosenman. Vistanet gigabit testbed. *IEEE Journal on Selected Areas in Communications*, 10:1413–1420, December 1992.

[12] A. Takahashi, T. Nishino, K. Murakami, and S.C. Dunning. A broadband switching system for public network. In *Proc. of the ISS*, May 1990.

[13] D. Winkelstein and D. Stevenson. HiPPI link data analysis system: Test equipment for high speed network analysis. In *TriComm 91*, Chapel Hill, NC, April 1991.

[14] CCITT Study Group XVIII. Recommendations on B-ISDN. June 1992.