Local Area Network Traffic Characteristics, with Implications for Broadband Network Congestion Management

Henry J. Fowler
Bellcore
331 Newman Springs Road
Red Bank, NJ 07701

Will E. Leland
Bellcore
445 South Street
Morristown, NJ 07962

Abstract

This paper examines the phenomenon of congestion in order to better understand the congestion management techniques that will be needed in high-speed, cell-based networks. The first step of this study is to use high time-resolution local area network (LAN) traffic data to explore the nature of LAN traffic variability. Then, we use the data for a trace-driven simulation of a connectionless service that provides LAN interconnection. The simulation allows us to characterize what congestion might look like in a high-speed, cell-based network.

The most striking aspect of the LAN data is the extreme traffic variability on time scales ranging from milliseconds to months. Conventional traffic models do not capture this behavior, which has a profound impact on the nature of traffic congestion. When our realistic data is applied to simple models of LAN interconnection, we observe that:

- During congested periods, congestion persists and losses can be significant.
- Congestion losses cannot be avoided by modest increases in buffer capacity.
- The consequences of misengineering can be serious.
- Fortunately, most congested periods are preceded by signs of impending danger.

1. Introduction

Technological advances and customer demands are rapidly ushering in public high-speed packet switched networks. The target network for high-speed, high-bandwidth communication is the Broadband Integrated Services Digital Network (BISDN).[1] BISDN ultimately will offer a wide range of new services exhibiting many varied traffic characteristics that complicate congestion management. To address these concerns, one must first understand the nature of congestion itself. This paper examines some forms of congestion that may arise in BISDN.

Considerable effort has been, and continues to be, devoted to understanding the congestion management needs of BISDN. Some efforts focused on particular aspects of the problem, such as algorithms for discarding Asynchronous Transfer Mode (ATM) cells. [2] [3] Others examined performance issues, such as the effects of non-random cell arrivals on switch design.[4] Still other efforts have been devoted to developing an overall structure for studying congestion management in BISDNs.[5] [6]

While defining congestion management approaches is extremely important work, the nature of the congestion to be managed remains almost completely unknown. It is difficult to forecast how congestion will arise in a BISDN, because we do not yet know which services will be offered and which will contribute a significant portion of the BISDN traffic. In addition, the traffic characteristics of many new services, such as video, are poorly known. While we cannot describe the congestion behavior of all possible future services, this paper characterizes the congestion behavior that may be induced by interconnecting existing local area networks (LANs). We believe that such insight provides a clearer understanding of the congestion management issues that will be encountered in the emerging integrated high-speed networks.

The literature contains many different definitions of congestion. In this paper, we define congestion to be a user-visible degradation of network performance. Thus, if the network traffic load is heavier than usual but performance is acceptable, congestion has not occurred (although the network may be in a congestion-prone condition). For example, even if a buffer is more full than usual, the queue is not congested under this user-oriented definition as long as delays are acceptable and no losses occur.

This paper examines networks that provide interconnection for local area networks through a connectionless service. LAN interconnection is expected to rapidly become an important BISDN service, both because the need already exists and because the high peak-to-mean ratios in LAN traffic volume allow high multiplexing gain and, hence, the economies of operation needed to offer a public service. However, the burstiness of LAN traffic will be an important test of the congestion management capabilities of any high-speed integrated-service network. Thus, this paper examines the characteristics of data traffic and the implications for congestion management in such a network. The paper does not advocate particular congestion management algorithms.

We use very high quality, high time-resolution LAN traffic data, gathered by Leland and Wilson[7], that allow us to study traffic characteristics – most significantly, the burstiness of data traffic. However, it is hard to draw specific conclusions about congestion by examining only the collected measurements. We therefore simulated some aspects of a possible LAN interconnection service for a more quantitative look at congestion. Although the results cannot predict exact traffic behavior in future networks, they provide significant qualitative conclusions about the nature of congestion and the task of congestion management.

1.1 Paradigms of Traffic Behavior

The appropriate congestion management techniques for a network supporting a LAN interconnection service depend heavily on the characteristics of data traffic. The ability to react to congestion depends on the magnitude and frequency of variability in this traffic.

An important theme of this paper concerns merging two *paradigms*, or models of the world, regarding the impact of traffic characteristics on the congestion management strategy. The first paradigm is based on many decades of experience with voice networks that provide Plain Old Telephone Service (POTS). Much literature exists on the theory of how humans initiate telephone calls, how long we talk, and so on. (References [8] and [9] present these theoretical foundations.) The POTS network is a connection-oriented network, where calls are assigned constant bandwidth. The fundamental assumptions about this traffic – verified by years of experience – are:

- Within an hour, traffic fluctuations can be predicted reasonably well with traffic models based on a Poisson arrival process.
- Variations from hour to hour are significant, but follow a pattern that tends to repeat. There is one hour of the day (the ‘‘busy hour’’) that tends to be higher than the others.
- There are predictable seasonal variations; for example, there is a ‘‘busy season’’ that repeats year after year.
- Traffic variation from year to year is generally slow and steady, and the use of linear forecasting techniques is well established.[10]
- When excessive traffic is offered to the POTS network, new calls are denied access.

In this network, the ‘‘low-frequency’’ traffic variations are either repetitive (such as daily variations) or of low magnitude (year-to-year variations). Higher-frequency variability caused by individual call attempts is well understood. Therefore, in a POTS network a major way to manage congestion is to engineer the network so congestion rarely occurs.

However, when discussing data traffic, another paradigm underlies much of our thinking. In this paradigm, the dominant traffic variability is in the range of milliseconds, and packets and ‘‘packet trains’’ (see [11] and [4]) enter the network at random times. (A packet train is a closely-spaced sequence of packets between the same source and destination.) The variability at this time scale is enormous, with relatively low mean utilization and high burstiness. However, because typical data traffic models provide little variability at time scales longer than the durations of packet trains, the implicit assumptions are that the low frequency variations are either low magnitude or highly predictable, and therefore have little effect on congestion. This study shows that neither assumption is accurate.

After exploring some consequences of actual data traffic behavior, we assert that neither the POTS paradigm nor the packet-train paradigm is sufficient as a basis for congestion management in a network that provides LAN interconnection. Both paradigms provide valuable insight, but neither adequately accounts for the high variability, across many time domains, of actual data traffic. Thus, the world of data traffic is new and quite different from the POTS world. Integrated high-speed data networks offer a far greater congestion management challenge.

1.2 Congestion Time Signatures

A concept used in this paper is the ‘‘time signature’’ of congestion, which refers to the evolution across time of some critical measurement (such as demand for a resource or performance of a network component) before and after each episode of congestion. This paper examines the concrete example of the variation of the buffer occupancy and packet loss of an interoffice link. The time signature heavily influences which responses to congestion are possible or appropriate. Two important aspects of a time signature are the speed of *congestion onset* and the speed of *congestion abatement.* Assuming congestion is not currently visible, *congestion onset* describes the phenomenon in which the demand for the resources of some network component increases with time until performance degradation becomes visible to the user. *Congestion abatement* describes the phenomenon in which the demand for critical resources decreases over time until performance degradation ends. Abatement applies only to the traffic *offered* to the resource, and is not affected by congestion management actions taken for that resource.

The speed of congestion onset and congestion abatement are major determinants of what congestion management approaches are feasible. If onset and abatement are on the order of a few milliseconds, then automatic switch-to-switch controls cannot be effective. If onset and abatement are far slower (on the order of many seconds or minutes), then manual controls that require human intervention through an Operations System (OS) can be used. Between these two time scales is a range in which automatic controls (which might involve OS participation) may be employed effectively.

In the many arguments about what congestion management methods are appropriate, attitudes are heavily influenced by assumptions about the traffic time signatures. A central question is whether congestion arises suddenly or evolves gradually. The answer depends on the relative magnitude of the low and high frequencies of traffic variation. If low frequencies are significant, then human or switch-to-switch based actions such as rerouting can help avoid congestion. But if the high frequencies dominate the low, then a switch is largely on its own to take actions to relieve congestion (such as discarding low priority packets from buffers). For a high-speed network, the time required for one switch to alert another can be long compared to the time for congestion to build and dissipate. Thus, by the time a switch detects congestion and alerts another switch, the congestion episode may have ended.

To decide what congestion management capabilities should be designed into high-speed packet networks, it is necessary to understand what time signatures are reasonable to expect. What is the nature of the traffic variability, and which variation frequencies are dominant?

2. The Data

The study uses high time-resolution data that were collected between March 1989 and January 1990 on several Ethernet® LANs at the Bellcore Morristown Research and Engineering Center, using a custom-built hardware monitor. The traffic monitoring techniques are described in more detail in Leland and Wilson[7]; the important features for the purposes of this paper are that the data are uniquely complete and accurate: all packets were recorded, regardless of traffic levels, and the recorded timestamps are correct to within 100 microseconds.

Although there have been many studies of LAN traffic since the early Ethernet measurements of Shoch and Hupp,[12] the emphasis has been on intermediate time scale behavior[13] [14] [15] or on user-oriented measures of behavior such as LAN throughput and delay[16] [17]. Little published data are available on the packet loss and timestamp errors introduced by the data collection techniques employed. The early studies typically used time measurements with low resolution (on the order of one second), with high loss rates (as much as 5% to 9% of all packets). Recent studies, such as Feldmeier[18], Gusella[19], and Jain and Routhier[11], have used higher time-resolution measurements, but also lose timing accuracy and clip traffic peaks: the reported systems failed to record between 1% and 5% of the LAN packets. Many of these studies, moreover, have employed experimental configurations that actively generated traffic on the network being monitored. These constraints on the accuracy of the characterization of rapid packet sequences make earlier studies inadequate for studying the shortest time scales relevant to this investigation. At the other extreme of time scales, these studies often used small samples – often only a few hundreds of thousands of packets collected within a few weeks. LAN traffic is highly variable over all time scales, so distinguishing temporary aberrations from normal behavior can be difficult when only small sample periods are recorded. By collecting detailed records on hundreds of millions of packets, on different intra-company networks, at intervals over the course of a year, this study can consider features of the observed traffic that persist across time, and can illuminate some of the patterns of variation.

Nonetheless, the data used in this study have some significant limitations:

- Although data were collected from different LANs, the LANs were all in the same company, and so cannot be considered representative of all LAN traffic. (However, as will be discussed below, some characteristics may be universal.)
- At any one time, only one LAN could be monitored. Therefore, correlations in the activity on different LANs could not be measured.
- When the data are used in a trace-driven simulation, the simulated traffic does not reflect the protocol adaptations that would occur in a real network.

2.1 Characteristics of the Observed Traffic

In this section, we briefly characterize the observed arrival behavior for several hundred million packets, with particular attention to the wide range of time domains made accessible by our high time-resolution, long-term data collection. The availability of this data allows a more careful consideration of the time scales of traffic variability than previously has been possible. For brevity, our examples are drawn largely from aggregate LAN traffic recorded during October 1989; where appropriate, we will compare these results with other studies in the literature, with our observations from other months, and with synthetic traffic models.

Ethernet is a registered trademark of Xerox Corporation

The most familiar characterizations of packet traffic are arrival rates over medium to long time scales: the number of bytes or packets seen per second, hour, or day. The classic 1979 study of a 2.94-megabit per second Ethernet by Shoch and Hupp[12] reported seeing about 2.2 million packets per day, representing less than 1% utilization of the network capacity, with a pronounced diurnal cycle. The diurnal cycle is still readily apparent in our 1989 measurements (as illustrated by Figure 1 for external[1] traffic in October 1989), but the value of ‘‘less than 1% utilization’’ is no longer realistic, even given the higher 10-megabit per second Ethernet speeds now in use: for internal traffic in October 1989, mean monthly utilization exceeded 15%, with daily peak hours generally exceeding 30% utilization on weekdays and peak minutes exceeding 50%.

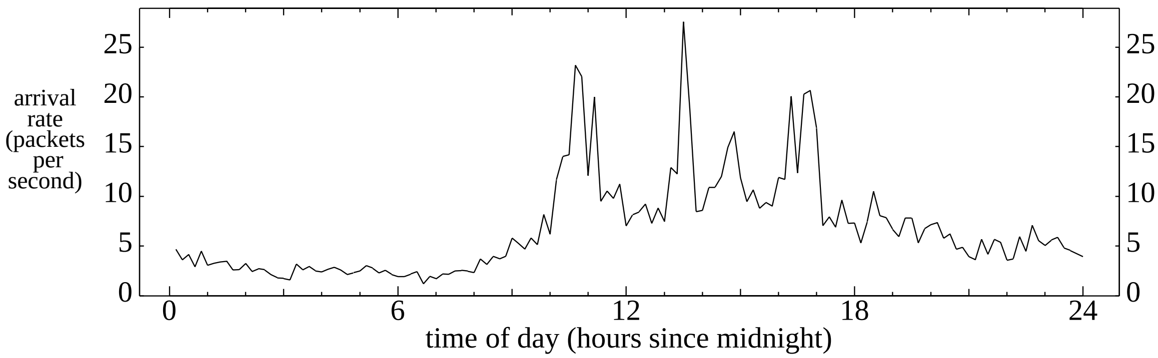

Figure 1. Mean External Traffic by Time of Day

Despite these differences, a comparison of the external traffic behavior observed in our study and the rate statistics available for 1979 suggests some striking similarities. In our data, the mean daily utilization of the Ethernet that can be attributed to external traffic is now in the 1% range characteristic of all traffic in 1979. The peak utilizations observed over shorter times for external traffic are also similar for the 1979 traffic: for example, the peak external minute in October 1989 represented an Ethernet utilization of 14.7%, while Shoch and Hupp reported a peak minute of 17%. This limited comparison suggests that highly time-varying traffic behavior might have been similar in 1979 to the behavior now observed; unfortunately, more detailed information is not available in the 1979 report.

Variability in the offered traffic often is described using the ratio of peak traffic to mean traffic. As Figure 2 illustrates, the observed value of peak bandwidth depends critically on the time interval over which the peak bandwidth is determined. For this particular two-week sample of external traffic, the peak bit rate observed in any 5 second interval is 152 times the mean arrival rate, while the peak rate observed in any 5 millisecond interval is 715 times the mean. Such behavior differs profoundly from the ‘‘burstiness’’ of simple arrival models, as is illustrated by the peak-to-mean ratios shown in Figure 2 for a comparable Poisson arrival process (described in more detail in section 3.2).

---

1. *Internal* traffic consists of all packets on a LAN, regardless of source or destination; *external* traffic consists of packets originating on one LAN but routed to another.

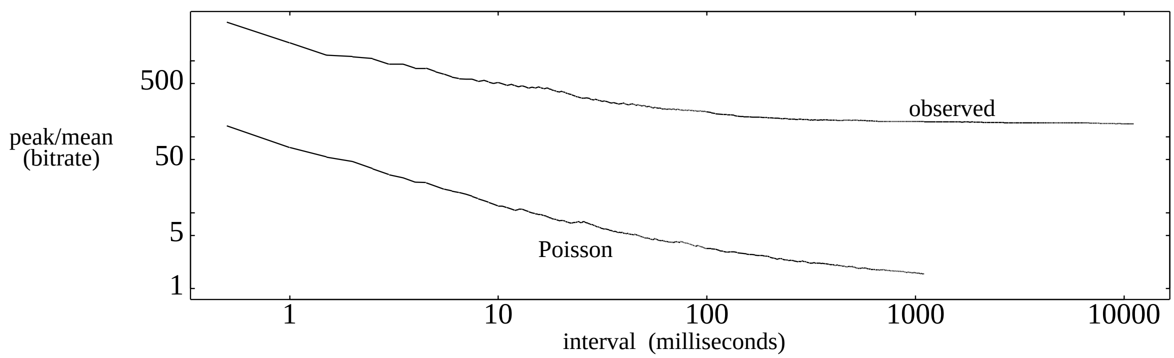

Figure 2. Peak to Mean Ratio vs Sample Duration

A more formal measure of the variability of the offered data traffic over different time scales is provided by the *index of dispersion of arrivals*[20]: the ratio of the variance in the observed number of arrivals in a time interval L to the observed mean number of arrivals in that interval. Figure 3 compares the index of dispersion calculated for the nine multiplexed LAN streams used in the congestion examples of Section 4 with the index of dispersion for a comparable batch Poisson process (described in more detail in Section 3.2 below).

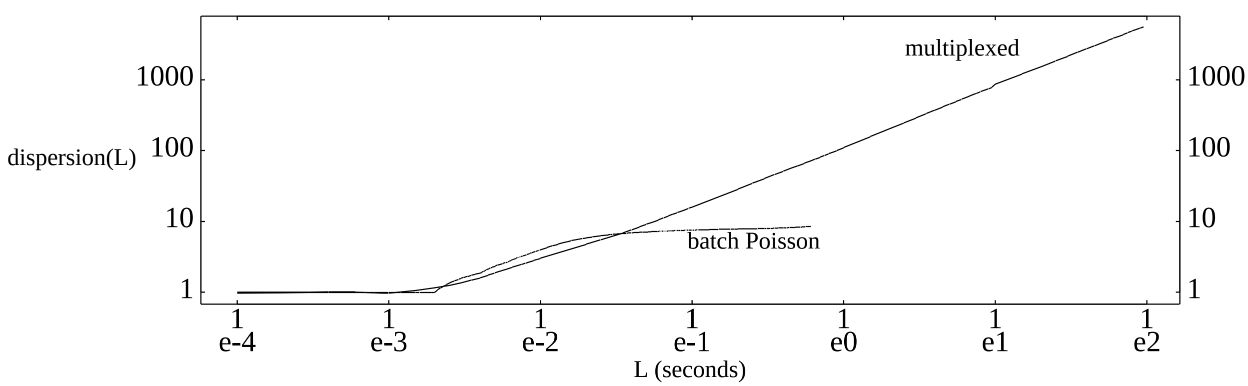

Figure 3. Index of Dispersion: Observed vs Batch Poisson

It is immediately apparent that multiplexing these highly variable streams has not produced a smooth, or even a Poisson-like, aggregate. It is especially significant for congestion studies that the dispersion monotonically increases throughout a time span of 5 orders of magnitude. In contrast, pure Poisson processes have an index of dispersion equal to one; other simple arrival models, such as batch Poisson, deterministic batch, hyperexponential and MMPP (Markov-modulated Poisson process) have indices of dispersion that converge to fixed values over a time scale on the order of the time constants of the models.

The observed variability extends throughout the range of available time domains. Examining the data traffic over two weeks, as shown in Figure 4, reveals dramatic variations in offered traffic. On still longer time scales, we note that between March 1989 and August 1989, the traffic on the studied LANs doubled without an increase in the user population, and nearly doubled again between August 1989 and January 1990.

> *The LAN traffic measurements show a high level of variability on every time scale that was measured: the index of dispersion shows that the variability remains extreme as the time scales increase from milliseconds to minutes, while informal comparison of traffic across weeks and months reveals qualitatively the same phenomena. Intuitively, both low frequency and high frequency components must be considered when studying congestion.*

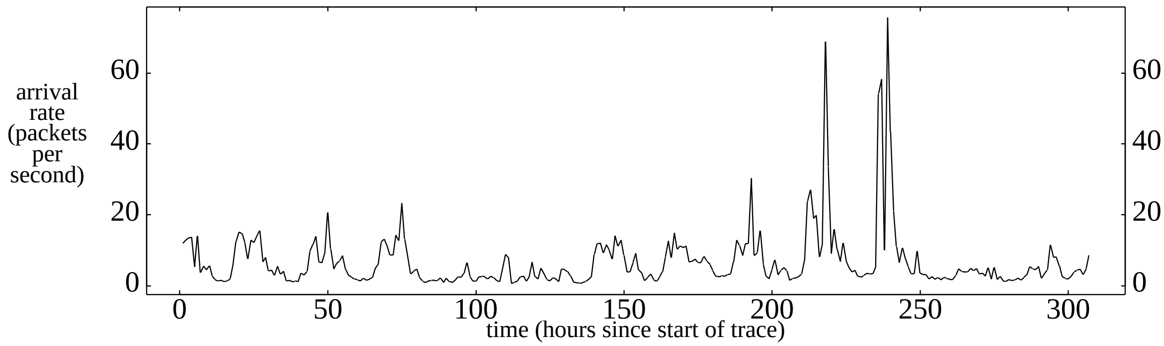

Figure 4. External Traffic, 10 October to 23 October, 1989

Because data traffic has profound variability at all time scales, several critical questions about congestion management must be examined more closely.

2.2 Questions About Congestion Management in a High-Speed Network

Congestion management has three aspects: *prevention, avoidance,* and *recovery.* Congestion prevention involves designing and building a network that minimizes the probability that congestion will occur. Prevention includes judiciously engineered (or ‘‘sized’’) components, a well-designed routing algorithm, traffic enforcement to ensure that a user’s access line does not exceed its subscribed traffic rate, and queueing policies that protect critical classes of traffic.[21] Congestion avoidance is action that is taken by the network before performance degradation occurs, to reduce the chance of congestion. An example is changing the packet routing tables to route traffic around a heavily loaded network component. Congestion recovery is action taken by the network after performance degradation is detected to limit the effects of congestion. An example is discarding lower-priority packets when buffers are full. Specific questions regarding the implications for congestion management strategies in general are listed below.

Congestion Prevention: In POTS today, we can predict with reasonable precision what the peak demands will be, and then size the network to work well under most circumstances. In such a situation, congestion prevention is the most essential aspect of congestion management. Because the analysis in Section 2 shows that data traffic does not have a predictable level of busy period traffic, if we try to size the network so that congestion is rare, we lose the economies of scale that the service must exploit. Although accurate sizing of the network remains an essential part of congestion management, it cannot suffice to reduce congestion adequately. This study considers three specific questions that illuminate the degree to which congestion may be minimized in advance:

Q1. What is the variation of data traffic on time scales of hours, days, and months?

Q2. Because we cannot predict peak usage, are the consequences of underengineering significant? In other words, if the prediction of peak usage is too low, will it have a significant impact on performance?

Q3. Is it possible to solve queue-induced congestion by increasing the capacity of the buffers?

Congestion Avoidance: Congestion avoidance involves detecting when congestion is imminent, and taking action designed to prevent it. In POTS, congestion is predicted by monitoring traffic patterns, or by monitoring the events that stimulate traffic.[2] To understand how effective avoidance may be in future

---

2. POTS network traffic managers generally watch television news so they can be aware of the impact of outside events on voice traffic. For example, a plane crash at an airport will generate many calls to that area.

networks, we consider:

Q4. Are there patterns in data traffic that can be used to predict congestion?

Congestion Recovery: In POTS, when congestion occurs, actions are taken to help the network to recover. However, with data network congestion, it is not clear when actions are warranted. The congestion may spontaneously dissipate quickly, or may have little serious consequence. Before any realistic proposal can be made for recovery, one needs to know:

Q5. How long does congestion persist?

Q6. What patterns of loss occur during congestion?

The limitations of the available data prevent us from modeling the effects of correlations among traffic on independent LANs and of correlations in protocol reactions. In particular, over times comparable to multiple roundtrips through the interconnected networks, adaptive higher-level protocols can reduce their loss rates significantly. Over short time frames, however, client-level protocols cannot react sufficiently quickly to avoid the characteristic congestion signatures observed in these trace-driven simulations.

3. The Simulation Study

This study used our LAN traffic data as input to a trace-driven simulation that models the traffic that might be offered to an interoffice link connecting two connectionless packet switches. Of course, the specific results in this paper will not be realized in any actual network: in addition to the differences in LAN traffic in other environments, effects would be seen from the integration of LAN traffic with other services on a multiservice network and from the operation of congestion management techniques (such as admission control) outside the scope of this study. Nonetheless, we can make general qualitative conclusions about the simulation results, and check their reasonableness with the characteristics of data traffic that seem to be invariant.

3.1 The LAN Interconnection Service Model

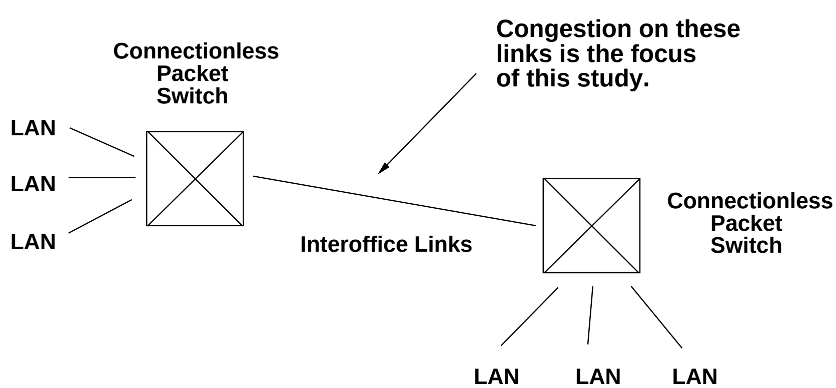

Figure 5. Network Supporting the LAN Interconnection Service

Figure 5 shows a simple network that provides LAN interconnection. The packet switches shown provide connectionless transport for packets of varying length; these packets are in turn composed of fixed-length cells. In this study, the connectionless switch is assumed to be dedicated to data traffic, but the same functionality also could be offered using an Asynchronous Transfer Mode (ATM) broadband switch with a service module that performs packet-level functions.

Each LAN is assumed to have a separate access connection to the switch, and the switches are connected by one or more interoffice links. Congestion due to traffic access contention for these interoffice links is the major focus of the simulation.

Congestion Management Limitations for the Network Supporting this Service: As compared to a POTS network, in which the network can control congestion well, the modeled LAN interconnection service offers challenges beyond the burstiness of data.

When a packet is sent from one LAN to another via a connectionless service, the network sees no single underlying virtual connection.[3] An implication is that the network cannot impose any end-to-end flow control that affects only one application-to-application dialog. Thus, we assume that end-to-end flow control cannot be provided by the network service, but might be provided by the applications using the connectionless service.

We also assume that the network cannot control the effective data rate of the LANs. There are two reasons for this assumption in our study. First, in any near-term time scale of interest, the embedded base of LANs will not be able to interpret and respond to such network requests (although in the future it may be possible in principle for the network to reduce the offered traffic rates of LANs that use the newly-proposed Medium Access Protocols). Second, because customers are unlikely to permit their intra-LAN traffic to suffer because of an interoffice network problem, reduction in the traffic offered by a LAN to the network will require traffic shaping at the network access point. Such shaping is beyond the scope of this study, although the burstiness of individual LAN traffic suggests that the loss and delay behavior described below for our switch simulation also may appear if queueing is done at the access point.

Thus, we assume that the network will not provide the congestion management options of end-to-end flow control, or network-based control of the rate of traffic entering the network. In a worst case scenario, all subscribers could simultaneously send traffic into the network at the LANs' maximum rates, creating extreme congestion. However, this study assumes that in practice, LANs will continue to be used over the near term as they are already being used, and that their joint traffic demand can be represented by independent streams of existing LAN traffic.

3.2 Details of the Simulation

In the simulation, the following assumptions were made:

1. The fundamental multiplexing unit of the network is a fixed-length, 53-octet ATM-like cell.
2. The unit of interest in this study is a variable-length packet (application protocol data unit) of up to 210 cells.
3. Although the service could be offered over an ATM network, the simulation models only packet switches.
4. The access and interoffice links operate at 45 Mbits/second.
5. All LANs are 10 Mbits/second Ethernets, and each LAN accesses its home switch through its own separate interface.

Each interoffice link has a queue associated with it, and this queue can receive traffic from any LAN that homes on the switch, as shown in the figure below.

---

3. There is no paradox in offering a connectionless service over the connection-oriented BISDN, because BISDN is connection-oriented only at the ATM level. Over a BISDN platform, there could be one connection between each LAN customer and its serving packet switch, and connections between packet switches. At the packet level, service is connectionless.

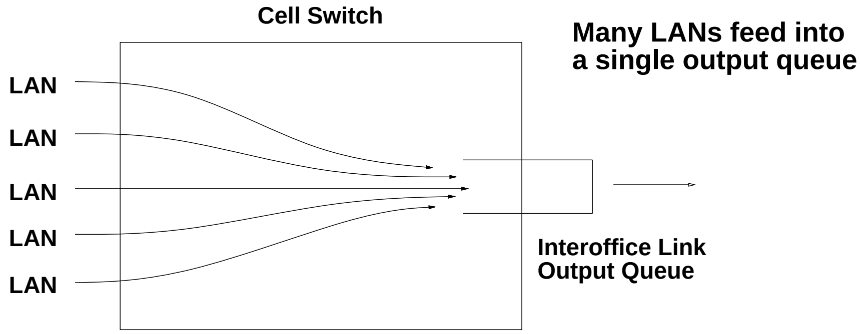

Figure 6. LAN Inputs to a Single Interoffice Link

Modeling the Link Buffer: In order to choose a size of the link buffer for the simulation, it was necessary to make assumptions regarding the performance requirements of the service. For a realistic estimate of these requirements, we consulted the performance objectives of a planned connectionless data service, Switched Multi-megabit Data Service (SMDS).[22] SMDS performance objectives include a packet loss rate of 0.01% (10^{-4}), and an end-to-end delay of no more than 20 milliseconds for 95% of the packets. We adopted these figures as guidelines for the performance objectives of the simulated interoffice link. The end-to-end delay criterion provided the most serious constraint, as it implies that no single link may introduce delays of more than a few milliseconds.[4] As a reference model, we selected a 960-cell link buffer (corresponding to a maximum buffer delay of 10 milliseconds) and a combination of LAN traffic streams that produced 0.01% packet loss with this buffer capacity. Some consequences of varying buffer capacity and traffic load are explored below.

The link buffer discard policy was chosen to be unprioritized: when a packet arrives at the link buffer, it either fits entirely in the buffer, or is immediately discarded without affecting packets that are already buffered. (Recall that the basic unit of transport and switching is the cell, but that the users are interested in packets. Therefore we assumed for this study that transmitting partial packets was not an acceptable policy. If more elaborate higher-level error-recovery schemes were modeled, a more subtle discard policy might be appropriate.)

Modeling the Switch: The simulated switch was assumed to be a perfect output-queueing switch, having no effect on the arrival of packets at the interoffice link. A major motivation for ignoring switch effects is to avoid linking the analysis to any particular switch architecture. The assumption is reasonable for switches that operate faster than the transport facilities that connect them. In addition, the effects of the switch on the traffic would be most noticeable on time scales shorter than the ones of interest here.

Modeling the Offered Traffic: We employed replicated data from our measured LAN traces to model the combined traffic offered to an interoffice link by multiple LANs. Each modeled LAN traffic stream was generated from an independent hour of traced data, to reduce spurious correlations. This approach gives a far more realistic model of actual data traffic than is offered by standard formal models; nonetheless, as noted above, it may well underestimate the peak demands on an actual link. The model introduces an additional simplification by assuming that all traffic from each source stream goes to the same interoffice link; our analysis suggests, however, that a more sophisticated model in which individual users were simulated would show the same dispersion behavior: the multiplexed traffic stream therefore

---

4. Although we simulate only one link, we require that its performance conforms to the restrictions of a system involving multiple links.

may be viewed as approximating the traffic generated by a larger pool of independent users.

For comparison with the results based on multiplexing actual recorded traffic streams, we also explored several simple formal models. In the figures below, the results labeled ‘‘batch Poisson’’ were derived by offering the same simulated network a Poisson stream of packet trains modeled on the observed traffic. Each packet train consisted of eight 589-byte packets separated by two milliseconds. The arrival time of the first packet in each train was determined by a pure Poisson process, so cells belonging to one train may be interspersed among cells of other trains. In figures where only one traffic load is considered (such as Figure 10), the arrival parameter λ of the Poisson process was selected to create the same mean packet loss (0.01%) observed for the multiplexed traffic traces at the reference buffer capacity of 960 cells.

4. Results of the Study

This section is organized to address the questions raised earlier in Section 2.2, and to consider the implications of our results for congestion prevention, avoidance, and recovery. The figures discussed below are based on trace-driven simulations representing at least one simulated hour of traffic through a 45 Mbits/second interoffice link for each combination of traffic streams used.

For the purposes of the congestion signature figures shown below, our simulation defined the end of a congestion episode as the first time after congestion onset that the buffer occupancy returned to zero, and defined the starting time of a congestion episode to be the time of the first packet loss following the end of the previous congestion episode. Each plotted signature represents an average over a few thousand to a few tens of thousands of congestion episodes, depending on the simulation parameters. The signatures make evident that under this narrow definition of a congestion episode, the episodes are strongly clustered: an intuitive interpretation, based on the dispersion curves, is that traffic ‘‘spikes’’ (which cause actual losses) ride on longer-term ‘‘ripples’’, that in turn ride on still longer-term ‘‘swells’’. Because of this fractal-like traffic behavior, we obtained similar results when using a variety of alternative definitions of a congestion episode.

4.1 How Predictable or Stable are Long-Term Traffic Variations? (Q1)

In POTS today, we predict with reasonable precision what the peak demands will be, and then size the network appropriately. Data network traffic also has well-defined peak periods; for example, Figure 1 (in Section 2.1) displayed the external data traffic by time of day, averaged over two weeks in October 1989. However, the actual amount of busy period data traffic is not as readily predictable, as was discussed in Section 2. Very low frequency variability (on the scale of months) is significant, which seems to be an intrinsic quality of data traffic due to LAN and application changes such as software upgrades. The magnitude of low frequency traffic variability implies that precise engineering of components such as interoffice links will be difficult. As a result, networks may need to be engineered conservatively to tolerate higher load than expected, network control may need to be designed so that existing capacity utilization can be quickly reconfigured,[23] and provisioning operations should be planned so that the time required to add network capacity is relatively small.

4.2 The Sensitivity of Performance to Misengineering (Q2)

Question Q2 concerns the magnitude of performance degradation when the interoffice link is offered heavier traffic than expected. Although only a modest number of combinations of input traffic streams were examined in this study, the results suggest that sensitivity can be severe. Figure 7 compares the loss rates for the simulated system, as a function of the aggregate offered traffic in Mbits/second. In interpreting this figure, we should emphasize that y-axis is logarithmic, spanning a factor of over 10,000 in the loss rates against a factor of three in offered load. Looking more closely at some particular points, we observe that when 9 streams from the October traces were multiplexed together, the aggregate load was 17.5 Mbits/second and 0.01% of the submitted packets were lost. Adding a tenth LAN stream to this particular traffic mix increased the load by only 6% (to 18.5 Mbits/second), but doubled the loss rate. Using a different combination of 10 streams increased the load by 17% (to 20.5 Mbits/second), while increasing the loss rate by a factor of eight.

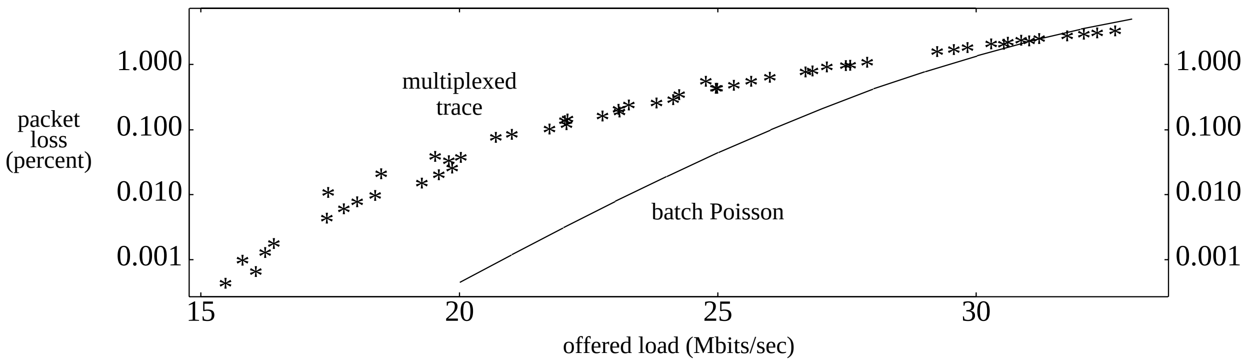

Figure 7. Packet Loss vs Offered Traffic Load

A similar story is told by the patterns of delay in regimes of acceptable loss, as shown in Figure 8 for the 95th percentile of packet delay with a buffer capacity of 960 cells. Here, for example, going from 9 LAN streams totaling 17 Mbits/second to 10 streams at 20 Mbits/second nearly triples the delay: from 1.27 milliseconds to 3.53 milliseconds.

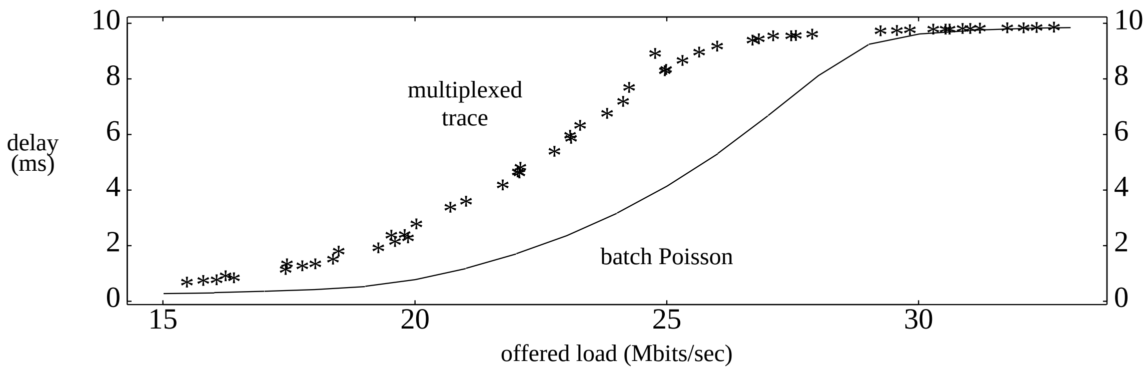

Figure 8. Packet Delay (95th percentile) vs Offered Traffic Load

The first lesson must be that determining the traffic level to engineer for, given some number of client LANs, is extremely difficult: in our study, different selections of the same number of LANs with essentially the same number of users produced quite different aggregate traffic rates, and even more disparate patterns of loss. But even if it were easy to determine what traffic level to engineer for, the performance drops off extremely quickly as the engineered capacity is exceeded.

4.3 The Effect of Increased Buffer Capacity on Performance (Q3)

One natural approach to reducing the sensitivity of the system to small changes in offered traffic is to increase the capacity of the link buffer. Figure 9 shows the loss rates for various buffer capacities: the lower set of data points corresponds to our running example of 9 LAN streams giving an offered load of 17 Mbits/second, while the upper curve comes from a set of 10 LAN streams totaling 20 Mbits/second. The initial impression created by the 17 Mbits/second results in Figure 9 is that small increases in buffer capacity might be sufficient to avoid most congestion losses, as it appears that linear increases in buffer size produce nearly exponential decreases in losses. This impression is encouraged by simple formal traffic models, such as batch Poisson arrivals, which predict an exponential decrease in loss rates with increasing buffer capacity.

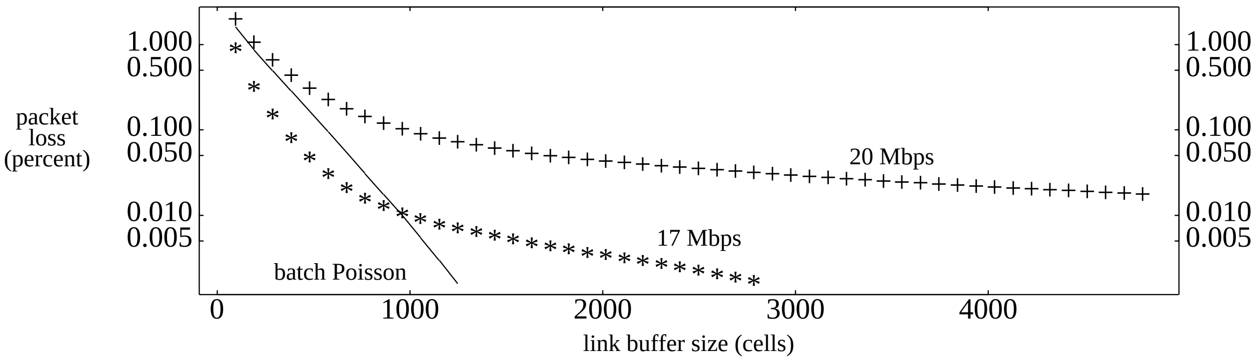

Figure 9. Packet Loss vs Link Buffer Capacity

Unfortunately, a more quantitative examination of the results reveals two difficulties. First, the decrease in loss with buffer size is far slower than the increase in loss with traffic load. To compensate for the increase in loss caused by the 17% increase in traffic represented by the upper curve would require that the buffer capacity be more than quadrupled. Second, and more seriously, the curves flatten with increasing capacity: further expansion of the buffers achieves less and less loss reduction. Of course, one might look at the figure and postulate that increasing the buffer size by a factor of, say, 100 might entirely avoid the traffic sensitivity problem. The difficulty with this naive approach is not so much the higher cost of such large buffers, but the upper limit on buffer size imposed by the delay requirement.

The conclusion suggested is that increased buffering will not eliminate the need for congestion management in the network, and the incremental load carried by larger buffers may be insignificant.

4.4 Indications of Approaching Congestion (Q4)

Question Q4 concerns whether one can predict that congestion is imminent. For Figures 10, 11, and 12, we have chosen the offered traffic so that the link queueing meets the loss requirement and is well within the delay requirement. Figure 10 shows how the occupancy of the link buffer changes before and after the onset of the current congestion episode. It is important to note that the points used to plot Figure 10 are averaged over one second, and that the peak buffer occupancy during any given second is far higher. Thus, although Figure 10 shows a maximum one-second buffer occupancy of about 170 cells, the 960-cell buffer experienced overflow. (See Figures 11 and 12.)

In sharp contrast to a batch Poisson model, a low-frequency pattern is clearly present in the trace-driven results. The instantaneous occupancy of the buffer is so volatile that this raw measurement is of little value, but a properly filtered variable could be used to extract the low-frequency pattern. On a time scale of minutes, traffic correlations arising from exogenous events also might be expected to show onset behavior similar to the signatures of normal congestion seen here. Thus, for both indigenous and exogenous congestion, some warning could be provided before actual packet losses become significant. However, this warning will not always be enough in advance of congestion to allow human involvement, and sometimes may provide insufficient time for machine-to-machine actions. The applications of filtered statistics for congestion management are an area of future work.

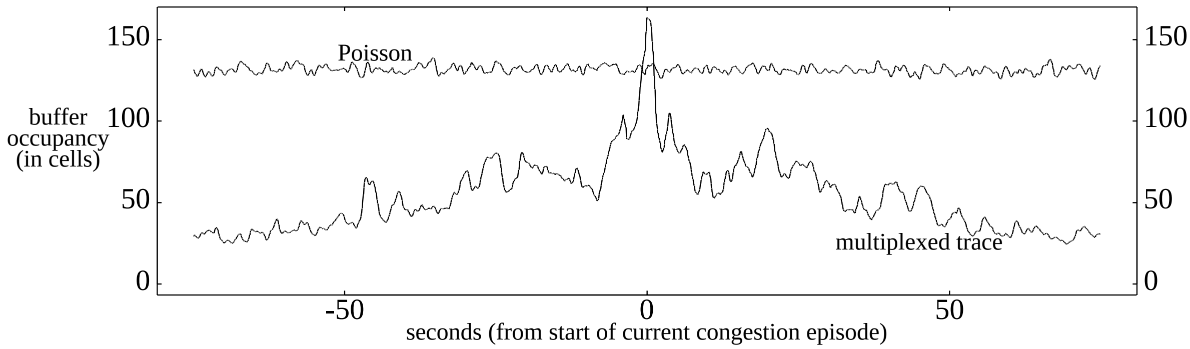

Figure 10. Buffer Occupancy Congestion Signature

4.5 Speed of Congestion Abatement (Q5)

Figure 11 shows that, although strictly speaking the time intervals in which there are losses may be short and may come and go quickly, these intervals of loss tend to recur over periods of seconds to minutes. Broadly speaking, congestion-prone conditions persist for extended periods during which the network repeatedly enters into episodes of actual packet loss. Although the overall loss rate in this simulation run is 0.01%, the loss rates are orders of magnitude greater for substantial intervals. The situation for underengineered networks, in which the offered traffic causes long-term mean loss rates above the design goal of 0.01% is far more serious in absolute losses during congested periods, but displays similar patterns.

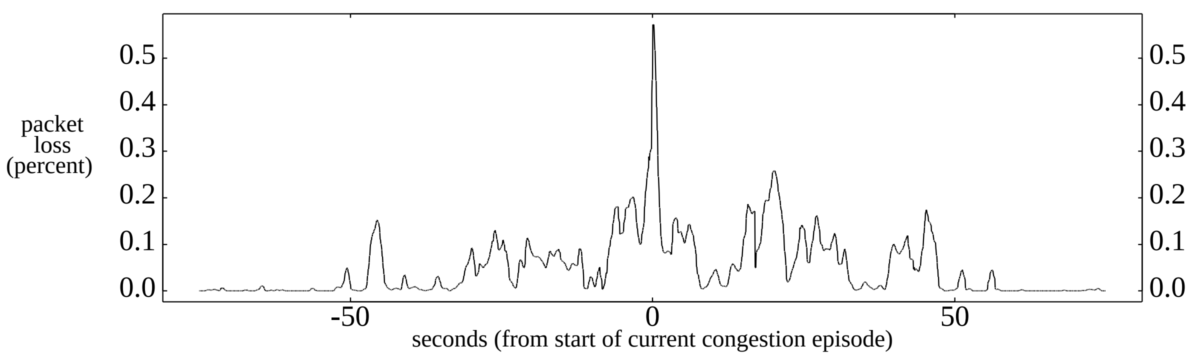

Figure 11. Packet Loss Signature, Long Range

These results, while providing hope for the reactive management of congestion, also limit the validity of our trace-driven simulation, which cannot incorporate the effects of automatic and human responses to congestion.

4.6 Loss Rates Immediately After a Congestion Episode Begins (Q6)

Figure 12 shows that, even assuming a properly sized network, losses may exceed the long-term average loss rate by an order of magnitude in the first second following the onset of congestion, while the loss rate is elevated by over two orders of magnitude during the first 100 milliseconds. This severe concentration of losses is caused by the strong high-frequency component of the traffic data; the low long-term loss rate reflects the overall rarity of congestion episodes rather than a low loss rate during congestion.

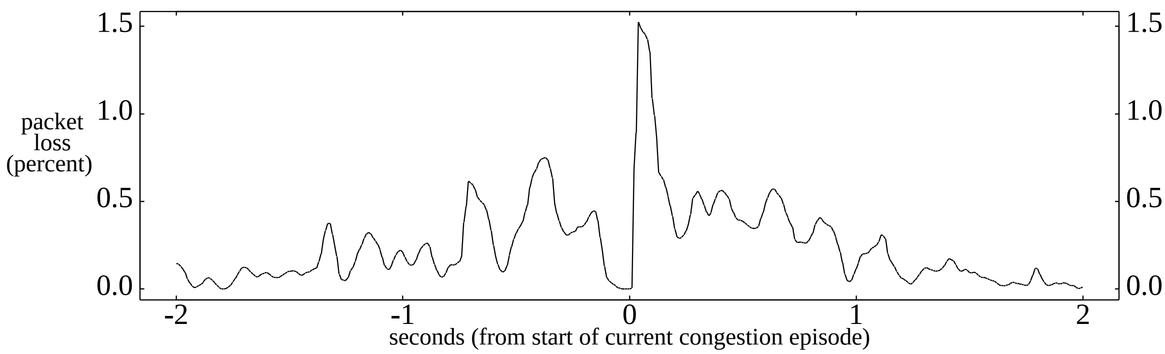

Figure 12. Packet Loss Signature, Near Term

We reemphasize that these conclusions regarding losses during congestion are quite conservative. In this case, the queue buffer was precisely sized to meet the performance objective, and the LANs were uncorrelated. No exogenous event (such a network failure or simultaneous traffic load from different LANs) caused congestion to occur. Under the impact of outside events, congestion and the resulting losses could be far larger.

4.7 Implications for Congestion Management

The following is a survey of the implications for congestion management of the above observations.

Congestion Prevention: It is not reasonable to expect that a data network can be engineered precisely to the size that is required. In contrast to POTS, data network traffic does not exhibit a predictable level of busy period traffic. The overall traffic rate can change dramatically in a matter of a few months, as was observed in the monitored LANs. In a network where the user population is growing, the increase in traffic could be even more rapid.

If more traffic is offered to the network than it is engineered to accommodate, the performance degradation can be significant. As a result, networks should be engineered conservatively to handle higher load than expected, and the lead-time required to augment the network should be short. A flexible bandwidth[24] or rapid reconfiguration[23] capability may be essential to exploit the global resources effectively.

While it is important for queue buffers to be reasonably large, their sizing need not be precise. Although some traffic models suggest that loss rates drop off sharply with increases in buffer capacity, real data traffic (Figure 9) shows that such behavior cannot be expected. Larger buffers will not prevent congestion from occurring.

Congestion Avoidance: It may be possible to compare a smoothed variable against thresholds to detect congestion or to activate congestion avoidance responses such as traffic rerouting. This possibility exists because the underlying low-frequency arrival pattern appears to be a reliable predictor of congestion onset. However, the low-frequency traffic pattern itself may be hard to determine, and so may not always give an unambiguous indication of congestion far enough in advance of the actual episode. The use of thresholds requires further study, but this approach is reasonable to pursue.

Congestion Recovery: Losses may be considerable during congestion, and any congestion management strategy must take this possibility into account. A broad implication of this conservative study is that critical traffic (such as network control or routing messages) must be protected by some priority mechanism or queueing policy[21] to ensure that messages needed to manage the congestion are not lost. In addition, during congestion episodes it may be desirable for the network to be able to distinguish which packets are the most critical to the end applications.

5. Summary and Conclusions

A necessary first step in designing congestion management for high-speed integrated-service networks is to understand the nature of congestion in such networks. This paper examined congestion for the specific example of a possible LAN interconnection service. LAN interconnection not only provides an important test case for the study of burstiness, but also is expected to be an important early service offered by high-speed public networks. We employed highly accurate traces of actual LAN traffic, instead of simplified models, in order to capture the extreme traffic variability that will confront practical congestion management mechanisms. Although the exact characteristics of LAN traffic undoubtedly will change in the coming years, the data show certain underlying similarities across samples. We hope that this work will point to traffic models that are representative of data traffic on longer time scales, and to methods of congestion management that are appropriate in such an environment.

The important overall finding of this study is the enormous time range over which data traffic is bursty. This point is graphically illustrated by the dispersion curve (Figure 3) as well as by other measures of traffic correlation, such as power spectra and Palm densities, discussed elsewhere [7] [25]. Paradigms of data traffic that ignore either high or low frequency variations are incomplete: *all* frequency components are important.

More specific observations are:

1. Data traffic does not have a simple recurring pattern over a period of days, weeks, or months.

2. Small errors in engineering can incur dramatic penalties in loss or delay.

3. Larger buffers provide limited protection against ‘‘spikes’’ in the offered load. Therefore, it is important to design a data network with the understanding that congestion will occur.

4. There are times when a network appears to be more congestion-prone than at other times. There is often a low-frequency ‘‘swell’’ in the aggregate traffic rate that precedes an episode of user-visible congestion. A statistic that properly smooths the short term fluctuations may be useful for detecting the approach of congested periods.

5. Although episodes of actual loss may come and go quickly, congestion-prone conditions appear to persist long enough for the effects of user and protocol responses to be felt. Properly understanding the way that a congestion episode abates requires a study that models these effects.

6. The examination of congestion episodes shows that when congestion occurs, losses are far greater than the background loss rate. This extreme variability in loss rates is caused by the strong high-frequency component of data traffic. We expect that in actual operating networks, more serious forms of congestion may arise that cause even greater losses.

An important future research topic is to develop a traffic model that displays characteristics similar to those displayed by real LAN traffic. A good test of such a model is that it must exhibit variation on many time scales and display performance characteristics similar to those documented in this study.

6. Acknowledgements

This study would not have been possible without the expertise of Daniel Wilson. We also are grateful to

Cedric Druce, Dave Feldmeier, and Mario Vecchi for their helpful comments on this manuscript.

REFERENCES

1. R. Sinha (Chief Editor), T1S1 Technical Subcommittee, *Broadband Aspects of ISDN - Baseline Document*, T1S1.5/90-001 R1, April 1990.
2. D. W. Petr and V. S. Frost, ‘‘Priority Cell Discarding for Overload Control in BISDN/ATM Networks: An Analysis Framework,’’ *International Journal of Digital and Analog Communications Systems,* Vol. 3, No. 2, April-June 1990, pp. 219-227.
3. A. E. Eckberg, Jr., D. T. Luan and D. M. Lucantoni, ‘‘Bandwidth Management: A Congestion Control Strategy for Broadband Packet Networks- Characterizing the Throughput-Burstiness Filter,’’ *International Teletraffic Congress Specialist Seminar,* Adelaide, 1989, Paper No. 4.4.
4. T. Eliazov, V. Ramaswami, W. Willinger and G. Latouche, ‘‘Performance of an ATM Switch: Simulation Study,’’ *Proceedings IEEE Infocom ’90,* San Francisco, June 1990, pp. 644-659.
5. J. Y. Hui, *Switching and Traffic Theory for Integrated Broadband Networks,* Kluwer Academic Publishers, 1990.
6. C. A. Cooper and K. I. Park, ‘‘A Reasonable Solution to the Broadband Congestion Control Problem,’’ *International Journal of Digital and Analog Communications Systems,* Vol. 3, No. 2, April-June 1990, pp. 103-115.
7. W. E. Leland and D. V. Wilson, ‘‘High Time-Resolution Measurement and Analysis of LAN Traffic: Implications for LAN Interconnection,’’ *Proceedings IEEE Infocom ’91*, Bal Harbour, Florida, April 1991, pp. 1360-1366.
8. R. I. Wilkinson, ‘‘Theories for Toll Traffic Engineering in the U.S.A.,’’ *Bell System Technical Journal*, Vol. 35, No. 2, March 1956, pp. 421-514.
9. W. S. Hayward and J. P. Moreland, ‘‘Total Network Data System: Theoretical and Engineering Foundations,’’ *Bell System Technical Journal,* Vol. 62, No. 7, September 1983, pp. 2183-2207.
10. A. J. David and C. D. Pack, ‘‘The Sequential Projection Algorithm: A New and Improved Traffic Forecasting Procedure,’’ *Ninth International Teletraffic Congress,* Torremolinos, Spain, 1979.
11. R. Jain and S. A. Routhier, ‘‘Packet Trains: Measurements and a New Model for Computer Network Traffic’’, *IEEE Journal on Selected Areas in Communications*, Vol. SAC-4, No. 6, September 1986, pp. 986-995.
12. J. F. Shoch and J. A. Hupp, ‘‘Measured Performance of an Ethernet Local Network’’, *Communications of the ACM*, Vol. 23, No. 12, December 1980, pp. 711-721.
13. D. R. Boggs, J. C. Mogul and C. A. Kent, ‘‘Measured Capacity of an Ethernet: Myths and Reality’’, *Proceedings SIGCOMM ’88*, Stanford, *Computer Communications Review*, Vol. 18, No. 4, August 1988, pp. 222-234.
14. P. D. Amer and R. N. Kumar, *Local Area Network Measurement*, Technical Reports 85-3 and 85-12, Department of Computer and Information Sciences, University of Delaware, April 1985.
15. R. Caceres, *Measurements of Wide Area Internet Traffic*, Report UCB/CSD 89/550, Computer Sciences Division, University of California, Berkeley, December 1989.

16. L. F. Cabrera, E. Hunter, M. Karels and D. Mosher, *A User-Process Oriented Performance Study of Ethernet Networking Under Berkeley UNIX 4.2BSD*, Report UCB/CSD 84/217, Computer Sciences Division, University of California, Berkeley, December 1984.

17. E. D. Lazowska, J. Zahorjan, D. R. Cheriton and W. Zwaenepoel, *File Access Performance of Diskless Workstations*, Technical Report 84-06-01, Department of Computer Science, University of Washington, June 1984.

18. D. C. Feldmeier, ‘‘Traffic Measurements of a Token Ring Network’’, *Proceedings IEEE Computer Network Symposium*, Washington, D.C., November 1986, pp. 236-243.

19. R. Gusella, *The Analysis of Diskless Workstation Traffic on an Ethernet*, Report UCB/CSD 87/379, Computer Sciences Division, University of California, Berkeley, November 1987.

20. D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events*, Methuen, London, 1966.

21. S. J. Golestani, ‘‘A Stop-and-Go Queueing Framework for Congestion Management,’’ *Proceedings SIGCOMM ’90*, Philadelphia, *Computer Communications Review*, Vol. 20, No. 4, September 1990, pp. 8-18.

22. Bellcore Technical Advisory TA-TSY-000772, *Generic System Requirements in Support of Switched Multi-megabit Data Service*, Issue 3, October 1989.

23. G. Gopal, C.-K. Kim, and A. Weinrib, ‘‘Dynamic Network Configuration Management,’’ *Proceedings IEEE ICC ’90*, Atlanta, April 1990, pp. 295-301.

24. G. R. Ash and S. D. Schwartz, ‘‘Traffic Control Architectures for Integrated Broadband Networks,’’ *International Journal of Digital and Analog Communications Systems*, Vol. 3, No. 2, April-June 1990, pp. 167-176.

25. W. E. Leland, ‘‘LAN Traffic Behavior from Milliseconds to Days,’’ *Proceedings Seventh International Teletraffic Conference Specialist Seminar*, Morristown, NJ, October 1990, pp. 6.1.1-6.1.6.

CONTENTS

LIST OF FIGURES