Traffic Pattern Reproduction and Manipulation for LAN-to-LAN SMDS Experiments

Ying-Dar Lin[1]
Steve Holmgren
Cedric Druce
Franklin James

Bell Communications Research
331 Newman Springs Road
Red Bank, NJ 07701

Abstract

It is recognized that LAN traffic contains highly correlated packet streams and is extremely bursty over a wide range of time scales. No formal stochastic models, like Poisson, batch Poisson, and Markov Modulated Poisson processes, can capture well the burstiness and correlation of LAN traffic. A hierarchical model is proposed in this paper to capture and reproduce LAN traffic. A packet trace of simulated LAN NFS™ traffic is fed into a SMDS access path simulator to evaluate the performance in terms of loss and delay measures. By manipulating a set of parameters, the traffic simulator constructed from this model is able to reproduce various traffic patterns. A burstiness analysis shows that the model can capture well the degree of burstiness over different time scales. The ability of DS1 and DS3 in supporting NFS traffic and the relationship between NFS read/write patterns and performance are also studied using this traffic pattern simulator. An "interval performance criterion" is described as an alternative for performance specification.

1. Introduction

As distributed computing systems become more tightly coupled by providing a higher degree of resource access transparency, another trend is to extend the scope of the systems from a local area to a wider area. A sophisticated distributed system needs strong support from the underlying network system. SMDS (Switched Multi-megabit Data Service) [2], which serves as an initial step toward public Broadband ISDN and will be one of the many services offered by B-ISDN in the future, is being deployed to achieve the above goal. It is anticipated that LAN interconnection for distributed systems over a metropolitan area will be the primary application of early SMDS subscribers. Unlike a private interconnection scheme where sometimes many lines are leased over a long term, SMDS provides a connectionless, packet-switched, public service with low cost, (relative to a mesh of leased lines).

To replace a mesh of point-to-point lines, SMDS needs to provide comparable performance in terms of delay and loss under the bursty traffic environment. Before SMDS can be deployed, analysis and experiments need to be conducted to predict the performance and foresee the applications that can be supported by SMDS. LAN traffic is known to contain highly correlated packet streams and very bursty arrivals. [3][4] LANs are subject to congestion and their performance depends heavily on the patterns of the offered traffic. An accurate performance study thus demands a realistic description of the workload and needs to explore the relationship between the performance and the workload patterns resulting from various applications.

Stochastic processes like Poisson, batch Poisson, and Markov Modulated Poisson processes are commonly used models to describe the packet arrival process in analytical and experimental studies. These models cannot capture well the characteristics of correlation and burstiness. Poisson-based processes lose variability when the time scale is over one second. However, the real traffic has burstiness at all time scales ranging from milliseconds to thousands of seconds. [4] Batch processes can only exhibit deterministic, simple correlation patterns, where complicated protocol behavior can not be captured. More sophisticated models like Markov Modulated Poisson processes, which have different rates at different states to increase the coefficient of variation, have indices of dispersion that converge to fixed values while the actual traffic has a monotonically increasing dispersion. [4] Moreover, none of these models accounts for packet stream correlation.

In fact, LAN traffic patterns are influenced by protocols, applications, and user behaviors. A hierarchical traffic model is proposed here to characterize traffic patterns by synthesizing the components within the traffic profile. This model is based on the observations of a burst hierarchy at the user, application, and protocol layers and the client-server interactions over the network. A traffic simulator for LAN interconnection over SMDS is con-

™ NFS is a trademark of Sun Microsystems Corporation.

92CH3132-8/0000-1879 $3.00 © 1992 IEEE

structed from the model. Measurements of NFS (Network File System) read/write transaction profiles are collected on an SMDS testbed and used to calibrate the traffic simulator. The simulated packet trace which represents LAN external NFS traffic across SMDS is fed into a SMDS access path simulator to evaluate the delay and loss performance. A burstiness analysis is conducted on the packet trace to calculate the degree of burstiness in terms of peak/mean ratio and coefficient of variation over various time intervals.

Section 2 gives a brief description of SMDS for LAN interconnection. In section 3, we characterize traffic patterns in term of locality, correlation, burstiness, cyclic repetition, and predictability. The hierarchical traffic model is presented in section 4. The implementation details of the NFS experiment and the results are shown in section 5 and 6, respectively.

2. SMDS for High-Speed LAN Interconnection

In contrast to high-speed networking in the intra-premises environment, wide-area, networks are bottlenecks to extending distributed systems to an inter-premises domain. Under this bandwidth constraint, only primitive applications like remote file transfer (eg. "ftp", "smtp") and remote login (eg. "telnet", "rlogin") can be supported with acceptable performance. Thus, the bottleneck between local area networks and wide area networks has to be eliminated. A typical current solution for high-speed LAN interconnection over a metropolitan area is a private T1 (1.544 Mbps) network. Customers lease T1 links to have a point-to-point connection between two separately located LANs. Extra lines have to be leased if multiple locations are involved.

SMDS is a public, packet-switched service for high-speed data sharing. A subscriber accesses SMDS via a DS1 or DS3 link. DS1 is a 1.544 Mbps link with approximately 1.173 Mbps maximum effective throughput after the physical framing and SIP (SMDS Interface Protocol) overhead has been removed, while DS3 is a 45 Mps link with approximately 34 Mps maximum effective throughput available to the subscriber. A logical private network can be created by the source address and destination address screening features. Inter-enterprise data sharing is achieved by adjusting the screening features. Unlike the T1 solution, the switching tasks are shifted to the SMDS operator. The service access protocol is SIP, which is based on DQDB as specified in the IEEE 802.6 standard.

SMDS is targeted to achieve high performance LAN interconnection. A list of performance objectives are documented in [2]. Among them, delay criteria for individually addressed packets measured from first-bit-out at one SNI (Subscriber Network Interface) to last-bit-in at another SNI should be less than 20 msec, 80 msec, and 140 msec for DS3/DS3, DS1/DS3, and DS1/DS1 configurations, respectively, for 95% of all packets delivered. DS3/DS3 here means two CPEs on both sides of SMDS are using DS3-based access paths. The loss ratio for individually addressed packets (of length 1200 octets) should be less than 0.01%. However, this connectionless packet-switched service is subject to congestion and performance degradation just like any other connectionless packet-switched networking service, especially when the network is designed to interconnect LANs, where traffic is extremely bursty. Some applications can tolerate this degraded delay and loss performance, while the others can not. Similarly, some applications tend to generate very bursty traffic patterns and congest the network, while the others do not. Thus, analysis and experiments are desired to foresee the applications that are suitable to run across SMDS and evaluate SMDS's ability to support bursty applications.

3. Characterization of Traffic Patterns

Characterizing traffic patterns for different types of applications is important for performance studies. To understand network traffic behavior, we characterize "traffic patterns" into five aspects, namely, locality, correlation, burstiness, cyclic repetition, and predictability.

Locality:

On a LAN, more than 80% of the traffic is contributed by less than 20% of communicating pairs, ie. traffic is not uniformly distributed. It is essential to capture this distribution in order to optimize the network configuration. A good indicator for the strength of locality is the pair (C%, P%) where C% of traffic is contributed by P% of communicating pairs.

Correlation:

The packet arrival process is not Poisson (memoryless). Packets tend to arrive in bursts within which packets are highly correlated. A burst hierarchy even exists with various levels. These correlations are determined by protocols, applications, and users where short-term correlation is determined by protocols and long-term correlation is determined by the user behaviors. The degree of correlation within an observed packet stream depends on the above three factors and the multiplexing effect. Multiplexing two packet streams will reduce the degree of correlation within the resulting stream.

Burstiness:

Different degrees of burstiness exist on different time scales, from milliseconds to thousands of seconds. It is common to have peak/mean ratio over 250 for 1-millisecond intervals and 10 for 1-second intervals. Moreover, highly correlated streams tend to generate more bursty patterns and in turn have inferior delay and loss performance. That is, the results of performance studies using constant or Poisson arrival process will be too optimistic. Three measures are usually used as indicators of burstiness degree: peak/mean, coefficient of variation, and index of dispersion. [5]

Cyclic repetition:

In most of the networks, a temporal cycle exists in the traffic distribution. A cycle hierarchy may even exist. For example, one week is a cycle and one day is a subcycle within that cycle. Being able to keep track of the distribution cycle would enable dynamic configuration management which tunes the network dynamically according to the cycle. The index of dispersion measure for arrivals does converge to an upper bound when the size of interval is on the order of days. [4] Study is needed to see how well this measure can detect a cycle and the strength of the cycle.

Predictability:

A good estimation of traffic demand can improve the resource allocation strategy. A dynamic resource pre-allocation scheme could be used if the system were able to capture the repetitive cycle behavior and predict for the next cycle. However, traffic in many networks is difficult to estimate. Based on the Law of Large Numbers, the predictability increases as the size of the user population increases. Since a small set of users is supported in a LAN, LAN traffic is very bursty with a high degree of fluctuation. If we take a snapshot of a LAN, the number of current communicating pairs is very small. Thus, LAN traffic can only be predicted over a long time scale, on the order of days.

4. Hierarchical Traffic Model

A good traffic model should be able to capture the traffic patterns characterized in the previous section. Motivated by the needs for modeling the complex traffic dynamics and understanding the relationship between patterns and performance, the proposed traffic model incorporates a hierarchical structure and a client-server interaction platform. The model is client-oriented where clients are activity initiators and servers are just passive respondents. In general, a system is decomposed into several subsystems where the subsystems are created and deleted dynamically by the system. A system represents a currently active communication session between a client and a server at this level. Note that the dynamic concept of a session is embedded in this model.

A set of parameters are available to calibrate the model to reflect different networking scenarios: entity configuration, transaction profiles, and transaction request arrival distributions. Their functionality is detailed in the following subsection.

4.1 Burst Hierarchy: Users, Applications, and Protocols

Packets arrive as bursts. A set of packets may be correlated at the protocol level (packets in the same swapped page), at the application level (packets in the same file transfer), and at the user level (packets caused by the same user). The level at which they are correlated determines at what level the burst is identified. Thus, a packet stream can be divided into a set of bursts where each burst belongs to a protocol entity. Bursts of the protocol entities owned by the same application can be grouped into a super-burst. A super-super-burst can be formed by grouping super-bursts of the application entities owned by the same user. Thus, a burst hierarchy can be constructed from a raw packet stream.

Figure 1 shows the hierarchical traffic model represented by a hierarchy of owner/member sets. An owner/member set is a mapping between members and the owner of that set. There are five levels in the hierarchy: network, user, application, protocol, and packet. An instance of the hierarchy of sets is shown in Figure 2. A packet at the lowest level can always trace upward via the owner/member sets to find out which protocol, application, user, and network it belongs to. Similarly, a higher level hierarchy entity can always travel downward to find all of the packets owned by it.

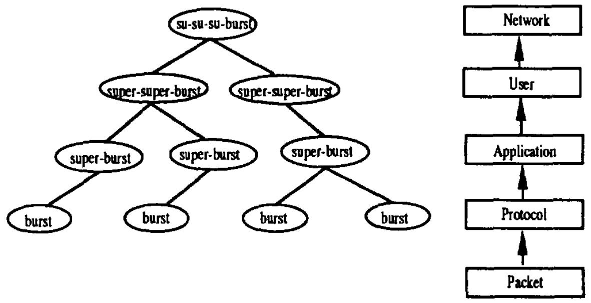


Figure 1. Levels in the heirarchical traffic model.

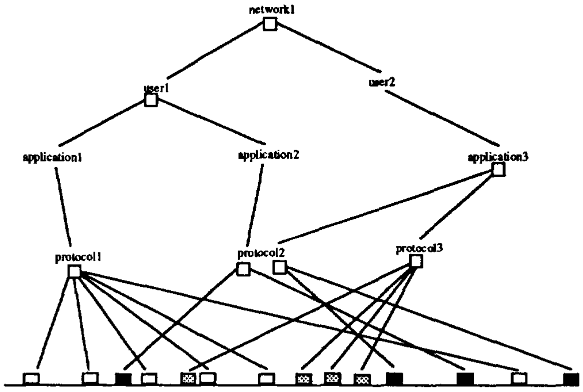
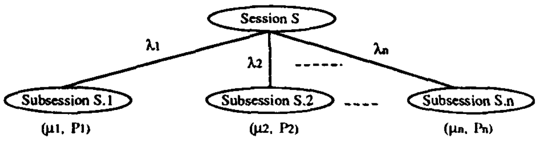


Figure 2. An instance of a burst hierarchy.

The hierarchy of sets is just a framework of this model. A schema for each entity in the hierarchy is required to make the framework "alive". This schema is a detailed description of how the entity behaves. It includes the following profiles:

- Entity/session configuration
- Basic subsession profiles
- Subsession request arrival distributions

A session configuration describes what kinds of subsessions this session can generate. The behavior of each type of subsession is plotted in the subsession profile. The third element determines the arrival process of each type of subsession. Each session in the hierarchy is thus well-defined with these three profiles. Figure 3 shows three representations for a session S where λ_i is the request interarrival time distribution, μ_i is the request size distribution and P_i is the basic subsession profile.



Session S [S.1(λ_1, μ_1, P_1), S.2(λ_2, μ_2, P_2), ..., S.n(λ_n, μ_n, P_n)]

$$S = \begin{bmatrix} \lambda_1, \mu_1, P_1 \\ \lambda_2, \mu_2, P_2 \\ ---- \\ \lambda_n, \mu_n, P_n \end{bmatrix}$$

Figure 3. Alternate representations for a session S.

4.2 Client-Server Interaction

A transaction which is initiated by a client has its client on one site and its server on the other site. At each site, there may be many clients of different transactions with servers on the other sites as well as many servers of different transactions with clients on the other sites. The traffic patterns to be modeled in our study are the input traffic from one site ***into*** the SMDS network. This offered traffic includes packet streams from both clients and servers over an access line at this site.

5. Simulation Study of NFS Traffic

A SMDS subscriber can choose DS1 or DS3 as the SMDS access line. This decision needs to be made based on the traffic characteristics and performance requirements of the applications to be run over SMDS. This simulation study is aimed at testing a tool to provide the above information to the subscribers. In the meantime, we need to validate the degree of burstiness in the reproduced traffic. Figure 4 decomposes LAN traffic into a protocol hierarchy. In this measurement of a 3-hour period, NFS traffic contributes 71% of transmitted packets, which is 90% of transmitted bytes. It appears that NFS traffic is a good candidate for this performance study. Based on the proposed traffic model, a simulator was written to simulate NFS traffic across an SMDS access link at the application level.

5.1 NFS Traffic Simulator

As shown in Figure 4, NFS invokes RPC (Remote Procedure Call) to transfer the requests and responses between clients and servers. RPC, in turn, uses UDP which uses the IP protocol to transfer packets. [6] When a read file request is generated at the client according to the distributions of read request interarrival time and request size, several NFS "lookup" requests are used to traverse the remote directory and a NFS "getattr" request to get the file location and attributes. Several block operations are then performed to retrieve the file one block at a time. The procedure to handle a write file request is similar except that blocks of data are transferred from a client to its server. The block size can range from 8K-byte to 512-byte which is adjustable according to a window flow control protocol. A NFS request/response is encapsulated by a RPC header and then by a UDP header. The limitation of UDP packet length is 64K-bytes, so that a block can fit in one UDP packet. However, the maximum frame size on Ethernets is 1518-bytes. Thus, a block of 8K-bytes is then fragmented into 6 Ethernet frames and a block of 2K-bytes is fragmented into 2 Ethernet frames.

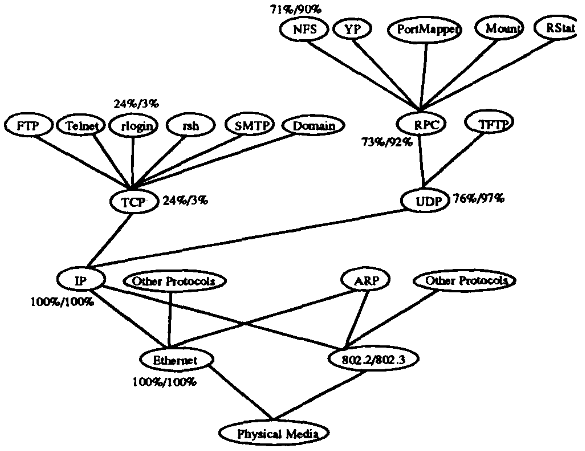


Figure 4. Protocol hierarchy and typical traffic distribution on LANs, showing percentage of total packets/bytes.

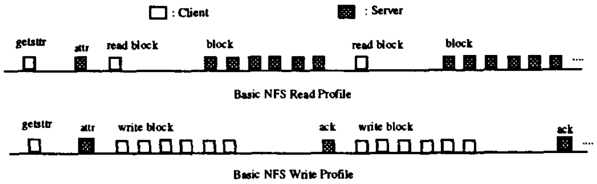


Figure 5. Basic NFS read/write profiles.

In order to calibrate the profiles, data was gathered from two LANs connected by DS1 access links to a prototype SMDS switch and is shown in Figure 5 as a basic NFS read/write profile. This profile, along with a similar NFS/SMDS read/write profile, is combined with application configurations (the number and location of mainframes, diskless workstations, and diskful workstations on the network). Estimated NFS request distributions (including inter-arrival time and file size) for each type of station are fed into the NFS traffic simulator. For clients on this site, the packets from the clients to their remote servers are extracted into a packet trace. Similarly, for servers on this site, the packets from the servers to their clients are dumped into a trace. These packet traces produced by different stations are aggregated and sorted by packet timestamp. A simple CSMA/CD backoff is applied to adjust the packet timestamps when Ethernet access contention occurs.

5.2 SMDS Access Path Simulator and Analysis Modules

A single queue is used to simulate a SMDS access path which is a DS1 or DS3 link. The buffer size throughout this experiment is assumed to be fixed at 250 packets. The access delay is defined to be from the time the packet arrives at the Ethernet transmitter queue to the time its transmission into the SMDS access line is finished. It includes possible Ethernet queueing, backoff, and transmission delay, DS1/DS3 queueing and transmission delay. For the purpose of comparing DS1 and DS3 performance, we have set the delay and loss objectives for the access link to be the same as the SMDS network switching delay and loss specified in [2], namely 20 msec for DS3/DS3 access, and 140 msec for DS1/DS1 access.

6. Simulation Results

The reproduced packet trace from the simulation does show embedded correlations within the packet stream. However, another characteristic to verify is how bursty the reproduced traffic is. We need to compare the reproduction of our model with real traffic and evaluate the perspective of this approach.

NFS is designed to integrate file systems in a distributed computing system and make them appear as one file system. As a result, a tighter performance requirement than for file transfer is desired. Also, due to its block operation, NFS traffic contains highly correlated packet streams and has strong burstiness. The results here should provide SMDS subscribers information about how much NFS traffic, with its high performance requirement, high correlation, and strong burstiness, can be supported using DS1 and DS3 access links.

6.1 Burstiness Analysis

In this analysis, a DS1 link with the configuration of 21 active stations (3 mainframes, 6 diskless workstations, and 12 diskful workstation) is studied (see Table 2, row 3). The reproduced traffic has 5.91% Ethernet utilization, 50.41% DS1 utilization, and average values of 96 packets/sec, 73913 bytes/sec, 10358 μsec interarrival time, 28.73 queue length upon arrival, 43236 μsec packet delay, 8.19% delay violation, and 3.47% packet loss. Each packet with a delay greater than the delay objective is counted as a "delay violation", and each packet that arrives at the queue when the buffer is full is counted as a lost packet.

Each packet trace simulates NFS traffic for a duration of 1200,000,000 μsec (20 minutes). For a burstiness analysis, we need to compute the values at different time scales. Peak/Mean ratio and coefficient of variance of inter-arrival time for different interval sizes are plotted in Figure 6. (Note that some axes are logarithmic.) This shows that the reproduced traffic has a high degree of burstiness in time scales ranging from milliseconds to 10 seconds, compared to Poisson and batch Poisson which converge in 1-sec intervals [4].

Figure 7 displays packet delay and delay violation ratio averaged over 1-sec intervals for part of a packet trace representing a 20 minute simulation. Between the 30th second and 90th second, only one spike at about the 44th second has average delay over the requirement, 140 ms. However, there are 5 intervals with individual packet delay violation ratios over 25%. This indicates that average delay, as a performance measure, is not as precise as delay violation ratio. Furthermore, it was found that during the simulated 20 minutes there are 1-ms, 10-ms, 100-ms, and 1-sec intervals where all the packets violate the delay requirement. That is, we can still find examples of the busiest 1-sec intervals during the 20 minutes where all the packets have a delay over 140 ms. Similarly for packet loss, it is shown in Figure 8 that many 1-sec intervals have 100% packet loss.

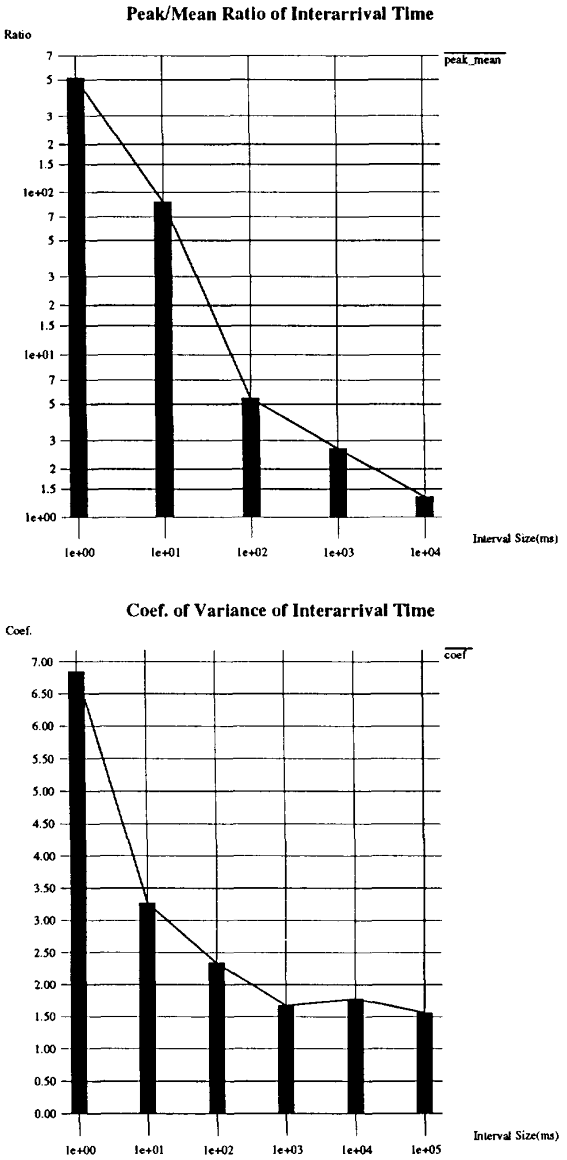
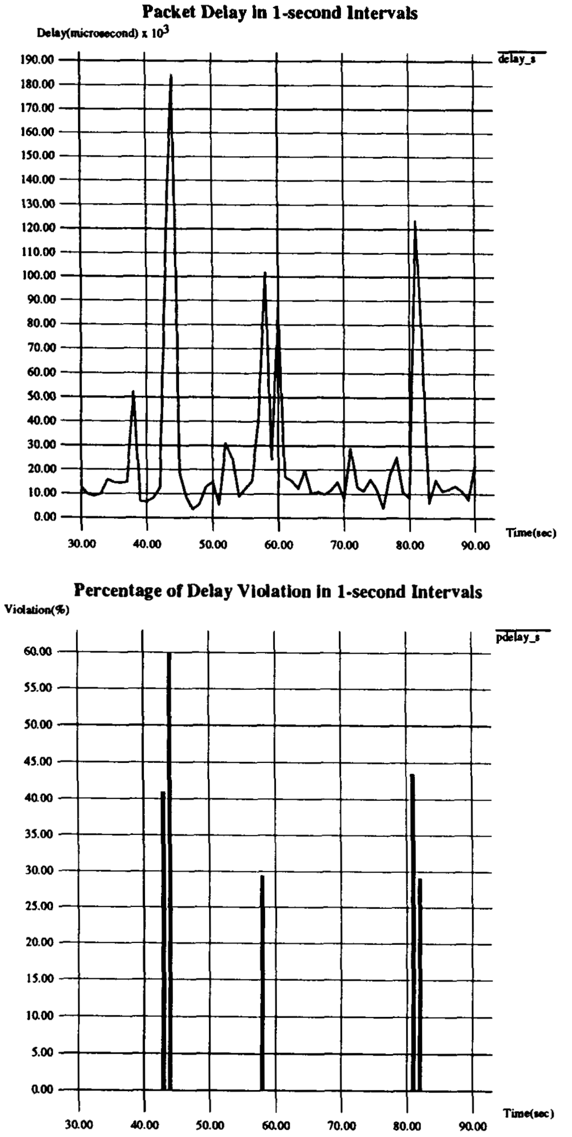


Figure 6. Two measures of burstiness of packet interarrival times from simulation results.



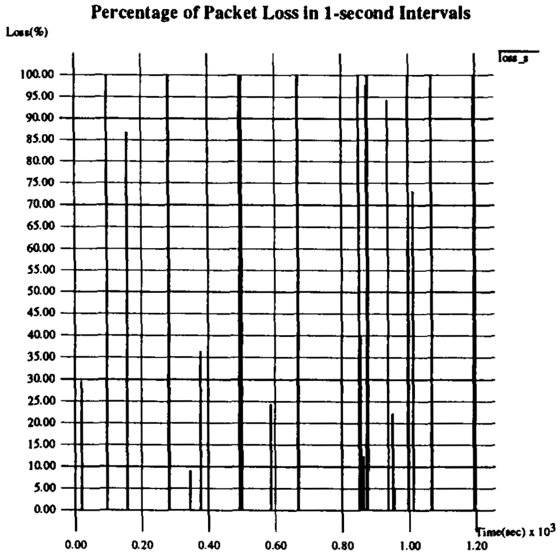
Packet Delay in 1-second Intervals

Percentage of Delay Violation in 1-second Intervals

Figure 7. Averaged delay and delay objective violation ratio during a portion of a 20-minute simulation.

This burstiness analysis has two implications:

- The performance violation ratio is a more precise measure than the averaged value when a QOS (Quality Of Service) is to be specified.
- Data traffic is extremely bursty, and long averaging intervals cannot capture its short-term behavior and performance.



Percentage of Packet Loss in 1-second Intervals

Figure 8. Packet loss ratio for a 20-minute simulation run.

Combining the above observations, Table 1 represents an example of how alternative "interval performance objectives" for delay and loss might be specified as a function of the averaging interval size. Smaller interval sizes are given less stringent objectives to accommodate stronger burstiness.

Table 1: Example of alternative performance requirements expressed as a function of the averaging interval.

Interval Size	Delay (95%)	Loss
1 hour	140 ms	0.01%
1 minute	180 ms	0.5%
1 second	250 ms	10%

6.2 Performance Analysis

Table 2 summarizes the simulation results for DS1 and DS3 access lines. The entry DS1(2,4,8), for example, means both source site and destination site are using DS1 access paths and there are 14 active stations (2 mainframes, 4 diskless workstations, 8 diskful workstation) on both sites. DS1 can support NFS traffic of 4.05% Ethernet utilization, which is 34.54% DS1 utilization, with overall delay violation ratio at 1.29% and a loss ratio at 0.21%, while DS3 can support 41.54% Ethernet utilization, which is 11.81% DS3 utilization, with delay violation ratio at 0.839% and zero packet loss.

7. Conclusion and Future Work

Motivated by the observations in the characterization of traffic patterns, a network traffic model is proposed. A simulation study for NFS traffic across SMDS is conducted to apply the traffic model to SMDS access links in supporting high-performance applications in distributed systems over a wide area. The results show the high applicability of our model to traffic pattern reproduction and manipulation for performance studies. An alternative interval performance criteria is described to better specify the system performance at various time scales.

The traffic simulator has been shown to be a very useful tool which can capture correlation up to the application level in the hierarchical model and burstiness from 1-msec to 10-second intervals. Increasing the simulation level to include higher burst levels in the hierarchy should increase the degree of correlation and burstiness. The simulator may be used to explore the traffic pattern/performance relationship and the packet trace could be played back on the SMDS testbed in Bellcore to evaluate SMDS switching network performance. Interval performance analysis modules will be added to study DS1/DS3's short-term behavior.

References

[1] This work was completed while the author was participating in the Technical Summer Internship Program at Bellcore; his current address is the Computer Science Department, University of California, Los Angeles, Los Angeles, CA 90024.

[2] Bell Communications Research, *Generic System Requirements in Support of Switched Multi-Megabit Data Service*, Bellcore Technical Requirements, TR-TSY-00772, Issue 1, May 1991.

[3] K. M. Khalil, K. Q. Luc, D. V. Wilson, *LAN Traffic Analysis and Workload Characterization*, proceedings of the IEEE 15th Conference on Local Computer Networks, 1990.

[4] W. E. Leland and D. V. Wilson, *High Time-Resolution Measurement and Analysis of LAN Traffic: Implication for LAN Interconnection*, INFOCOM 91.

[5] D. R. Cox and P. A. Lewis, *The Statistical Analysis of Series of Events*, Methuen, London, 1966.

[6] W. R. Stevens, *UNIX Network Programming*, p.204-209, Prentice Hall, 1990.

Table 2: Simulation results for various configurations of access lines and active stations.

Configuration	Ethernet (% util)	Access Link (% util)	Packets/ sec	Mbits/sec	Avg inter-packet arrival time (ms)	Avg queue length (pkts)	Average delay (ms)	% of pkts exceeding delay objective	Loss (%)
DS1(1,2,4)	2.17	18.5	34	0.216	29.0	0.9	10.4	0	0
DS1(2,4,8)	4.05	34.5	65	0.405	15.3	6.8	17.4	1.3	0.2
DS1(3,6,12)	5.91	50.4	96	0.591	10.3	28.7	43.2	8.2	3.5
DS1(4,8,16)	7.99	68.1	127	0.798	7.8	82.0	145.8	28.8	20.1
DS3(1,2,4)	2.18	0.62	22	0.217	43.7	0.0	1.2	0	0
DS3(2,4,8)	4.02	1.14	42	0.401	23.6	0.0	1.2	0	0
DS3(3,6,12)	6.35	1.81	66	0.635	14.9	0.0	1.3	0	0
DS3(4,8,16)	8.53	2.42	90	0.852	11.1	0.1	1.3	0	0
DS3(10,20,40)	20.7	5.9	218	2.07	4.5	0.2	1.7	0	0
DS3(20,40,80)	41.5	11.8	438	4.15	2.2	0.3	3.0	0.8	0
DS3(30, 60,120)	62.6	17.8	659	6.26	1.5	0.4	9.1	10.9	0