Multicast Network Traffic and Long-Range Dependence

Oznur Ozkasap
oozkasap@ku.edu.tr
Department of Computer Engineering
Koc University, Istanbul
Turkey

Mine Caglar
mcaglar@ku.edu.tr
Department of Mathematics
Koc University, Istanbul
Turkey

ABSTRACT

It is now well known that telecommunication traffic in high-speed data networks exhibits long-term correlations, which characteristic is simply called long-range dependence. Such correlations imply significant queuing delays and cannot be predicted by classical Markovian models. As a result, the architecture and protocol research has shifted towards the consideration of long-range dependence. A number of studies have shown that considerable degree of correlation exists in many attributes of high-speed data traffic. In contrast to the previous studies, this paper examines the existence of long-range dependence in multicast message traffic over a high-speed network and studies its properties. We accomplish a detailed analysis of Bimodal Multicast protocol in comparison with Scalable Reliable Multicast.

KEY WORDS: long-range dependence, multicast network traffic, Bimodal multicast (Pbcast), Scalable Reliable Multicast (SRM), multicast latency.

1. INTRODUCTION

The availability of high-speed networks and the growth of the Internet have triggered the use of multicast communication in large-scale settings. Furthermore, the widespread availability of IP multicast [DC90] and the Mbone [K95] have important consequences in terms of the use of large-scale multicast communication. These developments have considerably increased both the geographic extent and the size of communication groups. Several distributed applications exploiting multicast communication require reliable delivery of messages to all destinations. Examples include Internet media distribution, computer supported collaborative work, electronic stock exchanges and reliable information dissemination. As the size and geographic extent of such applications increase, scalable reliable multicast protocols become an essential underlying communication structure. One of the key properties that should be offered by a scalable reliable multicast protocol is predictable and stable delivery latency of multicast messages. Failure scenarios such as router overload and system-wide noise that are known to be common in Internet protocols can cause existing best-effort scalable reliable multicast protocols to behave pathologically [LMJ97,P97]. This indicates poor queuing performance, which is both implied and characterized by the long-term correlations present in the network data traffic.

It is now well known that telecommunication traffic in high-speed data networks exhibits long-term correlations, which characteristic is simply called long-range dependence (LRD). Such correlations imply significant queuing delays and cannot be predicted by classical Markovian models [N95]. As a result, the architecture and protocol research has shifted towards the consideration of LRD.

A number of studies have shown that considerable degree of correlation exists in many attributes of high-speed data traffic. The well-known study of [LTWW94] has established that network packet traffic exhibits long-range dependence and could be modeled by statistically self-similar processes. This work has analyzed Ethernet packet traces taken at Bellcore. Previously, network traffic was modeled with Poisson processes and all analysis of networks was based on that assumption. The significance of the work is the demonstration of the fact that the network traffic patterns are too bursty to be reasonably modeled using traditional Poisson arrivals. Another study by [PF95] demonstrated that these results are valid for wide-area TCP traffic. Several other studies focus on a special type of network traffic and its long-range dependent behavior. For instance, the work of [CB95] has considered WWW traffic and showed evidence that WWW traffic exhibits characteristics that are consistent with self-similarity. Likewise, [GW94] considers variable-bit rate video traffic and presents a statistical analysis. It shows that video sequence is long-range dependent and can be modeled using a self-similar process.

In this paper, we examine the existence of long-range dependence in multicast message traffic over a high-speed network and study its properties. We accomplish a detailed analysis of message latency behavior of Bimodal Multicast protocol in comparison with Scalable Reliable Multicast (SRM). A single parameter, called Hurst parameter, can be used to measure long-range dependence in the data traffic. The message latency distributions are used to characterize LRD behavior in this study.

The outline of the paper is as follows. In Section 2, we describe the multicast protocols under investigation. In Section 3.1, for long-range dependence, we first define the Hurst parameter as a measure, and discuss its characterization through message latency. In Section 3.2, the wavelet estimation method for the Hurst parameter is described. Section 4 discusses our results on the LRD behavior in multicast traffic. In Section 4.1, simulations focus on randomized message loss in a large-scale network. In Section 4.2, we consider a clustered network configuration, which is quite common in today's networks. The corresponding multicast latency distributions are investigated in Section 4.3. We evaluate the impact of a member's distance from data source for both LRD and latency distributions in Section 4.4. Finally, our conclusions are stated in Section 5.

2. PROTOCOLS

2.1 Bimodal Multicast Protocol

Bimodal Multicast (or Pbcast in short) [BHO99] is a new option in scalable reliable multicast protocols. The protocol is inspired by prior work on epidemic protocols [DGH87], Muse protocol for network news distribution [LOM94], the SRM protocol [FJL97], and the lazy transactional replication method of [LLS92]. The important aspects of Pbcast are an epidemic loss recovery mechanism, stable throughput property and a bimodal delivery guarantee. The protocol is constructed using a novel gossip based transport layer. The transport layer employs random behavior to overcome scalability problems. Higher-level mechanisms implementing stronger protocol properties such as message ordering and security can be layered over the gossip mechanisms. Pbcast has a very high probability of providing steady throughput even if message loss occurs or some group members fail. Stable throughput property entails predictable and small variance in the data delivery rate where data is generated at a steady rate. Unlike SRM (Scalable Reliable Multicast), which provides best-effort reliability, Pbcast provides a form of reliability that can be rigorously quantified [BHO99]. This is called bimodal delivery guarantee: with very high probability a Pbcast message will reach almost all destinations, and with very small probability, it may reach very few of its destinations. The protocol provides FIFO ordered delivery of messages.

Pbcast consists of two sub-protocols: an optimistic dissemination protocol and a two-phase anti-entropy protocol. The former is a best-effort, hierarchical multicast used to efficiently deliver a multicast message to its destinations. This phase is unreliable and does not attempt to recover a possible message loss. If IP multicast is available in the underlying system, it can be used for this purpose. Otherwise, a randomized dissemination protocol can play this role. The latter is an anti-entropy protocol that operates in a series of unsynchronized rounds. Each round is composed of two phases. The first phase is responsible for message loss detection. The second phase runs only if a message loss is detected, and corrects such losses. Detailed information on Pbcast is given in [BHO99].

2.2 Scalable Reliable Multicast

Scalable Reliable Multicast [FJL97] is a reliable multicast protocol which is designed according to the models of IP multicast group delivery, application level framing (ALF) principle, and the adaptivity and robustness in the TCP/IP architecture design. IP multicast [DC90] allows data sources to send to a group without needing any knowledge of the group membership. Basically, IP multicast is a best-effort delivery model and provides no reliability guarantees. ALF [CT90] is an architectural design principle for data communication. It introduces the integration of the protocol levels from the transport level to the application level. The goal is to provide flexibility and efficiency in the use of the network. However, this leaves the application to include most part of the transport functionality. SRM follows the core design principles of TCP/IP: It requires only the basic IP delivery model and builds reliability on an end-to-end basis. No change or special support is required from the underlying IP network. In a fashion similar to TCP adaptively setting timers or congestion control windows, SRM algorithms dynamically adjust their control parameters based on the observed performance within a session. SRM does not provide ordered delivery of messages. The protocol aims to scale well both to large networks and sessions. It exploits a receiver-based reliability mechanism.

As discussed in [FJL97], there is not a single setting for the timer parameters that gives optimal performance for all topologies, session memberships, and loss patterns. For applications where it is desirable to optimize the tradeoff between delay and the number of duplicate requests and repairs, an adaptive algorithm can be used. Adaptive SRM adjusts the timer parameters in response to the past behavior of the loss recovery algorithms.

3. CHARACTERIZATION OF LONG-RANGE DEPENDENCE

In this section, first, the Hurst parameter is introduced as a measure of LRD, which is then characterized through message latency distributions. Second, wavelet estimation of the Hurst parameter is described.

3.1 Hurst Parameter of Message Latency

Precisely, LRD is defined as the slow, power-law like decrease of the autocovariance function γ of a stationary sequence at large lags k, given by $\gamma(k) \sim k^{2H-2}$, with $0.5<H<1$. The parameter H is called the Hurst parameter, whose value represents the magnitude of the correlation. The value H=0.5 corresponds to an independent sequence as in Gaussian white noise, and the larger the H, the slower the decay of the function γ at large lags. So, we say that there exists more LRD as H increases.

The long-term correlations in the traffic can be characterized through latency process among others [AV98, BB98]. There exist traffic models for the workload, which corresponds to the message arrival process in multicast traffic studied in the present paper. It is shown in [AV98] that the Hurst parameter computed from the workload process and the latency process are in agreement. Along these lines, [BB98] study LRD through packet delay traces in the Internet traffic. Our approach in this paper will be similar. We will concentrate on the latency data obtained from the simulations of multicast message traffic and compute the Hurst parameter H from these data. Since the simulation is performed until it reaches the stationary state, the latency data forms a stationary sequence to facilitate the proper estimation of H. Here, the lag k of the function γ will have the unit of number of messages. We estimate H using wavelet estimation method as will be described next.

3.2 Wavelet Estimation Method

The wavelet estimation method is known to have very good properties for estimating the Hurst parameter H as opposed to variance-time estimation and other heuristic methods [AV98, VA99, C2000]. It is unbiased, consistent, and also a computationally efficient method of estimation.

We apply the wavelet estimation method as given in [VA99], using Daubechies wavelets with two vanishing moments. Let $d(j,k)$, $k=1,\ldots, n_j$, $j=1,\ldots,J$ denote the 'details' obtained by the discrete wavelet transform of the sequence of message latencies x_k , $k=1,\ldots,N$, where J is such that $2^{J+1} \le N \le 2^{J+2}$, and n_j is the number of coefficients available at octave j. The statistic central to the method is given by

$$m_j = \frac{1}{n_j} \sum_{k=1}^{n_j} d^2(j,k) \qquad j = 1,\ldots,J$$

Let c_f denote the coefficient in the spectrum of the latency sequence. That is, it is the counterpart of cγ of the autocovariance function in the spectrum. Then, the Hurst parameter H and the coefficient c_f are estimated through a weighted linear regression of

$$y_j = \log_2(m_j) - g_j$$

over $j=j_1,\ldots,j_2$, where j_1 and j_2 are the scales relevant for long-range dependence. Typically, these are the larger scales. The constant

$$g_j = E(\log_2(m_j)) - j(2H-1) - \log_2(c_f C)$$

is introduced to ensure that the fundamental hypothesis of regression holds (with C a constant that depends on H). Then, the slope a of the regression line is ($2H$-1) and the estimate of H is given by

$$\hat{H} = \frac{a+1}{2}$$

which is unbiased and consistent.

4. ANALYSIS AND RESULTS

By employing the simulation model that we have developed for Pbcast [O2000], and the available model for SRM and adaptive SRM, a simulation study for evaluating the message latency behavior of multicast protocols under various network settings has been accomplished. The underlying platform for our simulation model is ns-2 network simulator [BBE99]. In this section, we give and discuss results and quantitative analysis of our simulation study.

We performed this study by using latency data of multicast protocols under different topologies and network noise settings, which will be described next. To compute H, we analyzed latency data sets of length 2^{12}=4096 multicast messages. The simulations of different protocols and their versions are conducted with the same sequence of random numbers for proper comparison.

The latency of a data message at a process is defined as the delay between the time that a message is initially multicast to the group by data source and the time the message is first delivered by the process. There are basically two cases:

- The message is not exposed to a failure and delivered at the end of best-effort transmission,
- The message drops because of a failure in the network, and error recovery mechanism takes part to recover the message and makes sure it reaches to the intended destination processes.

In any case, a process can receive duplicate copies of a message, but in our analysis, we do not consider duplicate receipts, and just use first receipt time of a message to calculate its latency. Since Pbcast protocol provides FIFO ordered delivery, we analyze its latency distribution in two forms: Latency distribution at node level and latency distribution after FIFO ordering. In contrast, since SRM doesn't guarantee ordered delivery, we just analyze its latency distribution at node level.

4.1 Randomized Message Loss

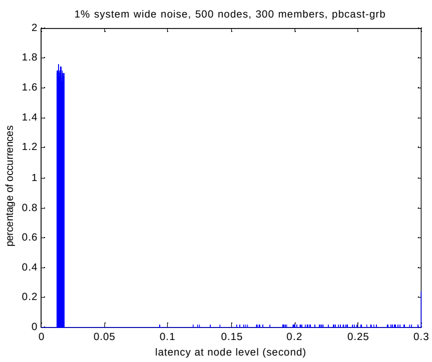
In this part of the study, data were gathered on a 500-node tree topology where randomly selected 300 nodes are group members. The sender located at the root node sends with rate 0.01 (100 multicast messages per second), and on all network links there is a system-wide noise with rate 1%.

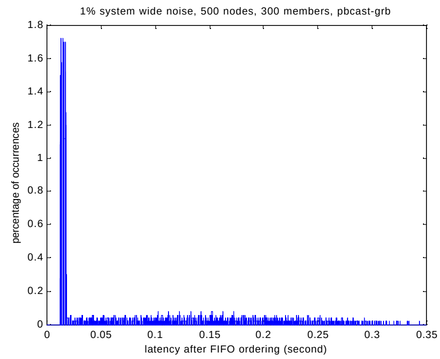
Figure 1 shows latency histograms of the protocols for this scenario where x-axis is latency in seconds (in increments of 0.1msec intervals) and y-axis is percentage of occurrences. Figure 1(a) and (c) are the node level latency distribution of Pbcast, and SRM respectively. As it is shown, a typical receiver delivers messages with lower latencies when Pbcast protocol is used for group communication. As pointed in Figure 1(c), SRM has a long tail with a maximum observed latency of nearly 800ms, and a group of packets delivered at around 400ms. Overall, SRM has a significant number of packets delivered during the first 100ms and a second broad distribution containing almost 5% of packets, which arrive with latencies between 300ms and 800ms. Notice that the basic SRM distribution is not as tight as the unordered Pbcast distribution, which has more than 90% of its packets arriving at the lowest possible latencies. In the case of Pbcast, around 2% of packets are delayed and arrive in the period between 200ms and 300ms, with no larger latencies observed.

We also investigated message latencies of Pbcast after FIFO ordering is accomplished. In that case, depending on the message loss rate experienced by the receiver, some percentage of messages are delivered with higher latencies since messages not in order are buffered prior to delivery in order to guarantee FIFO ordering property (Figure 1(b)). These higher latencies reflect the cost of waiting for messages to be retransmitted and placing them into the correct delivery order.

These results are important at least in settings where steady delivery of data is required by the application. We observe that as SRM is scaled to larger groups, steadiness of throughput can be expected to degrade. We experimented with a variety of noise levels, and obtained similar results, although the actual number of delayed packets obviously depends on the level of noise in the system.

We have found the Hurst parameter *H* for Pbcast to be 0.54 whereas for SRM to be 0.65. In this case *H* being close to 0.5, Pbcast performs very well with no LRD implication. On the other hand, the latencies for SRM are LRD although *H* is not very high. Hence, Pbcast performs better. Our estimate of *H* for Pbcast after FIFO is 0.72, which qualifies the latency in this case to be LRD. But, this is a moderate value like the *H* for SRM, in terms of its implications on the network performance.





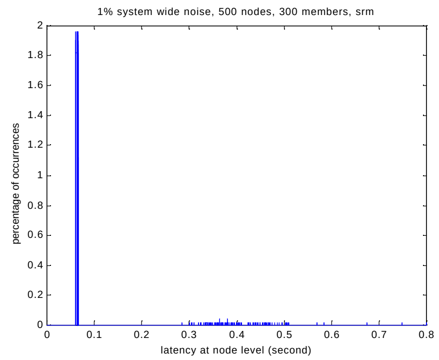


Figure 1. Latency histograms. (a) Pbcast at node level, (b)Pbcast after FIFO ordering, (c)SRM

4.2 Clusters connected by a noisy link

In the previous section, we focused on the impact of randomized message loss on the performance of Pbcast and SRM protocols. Other scenarios might be local area networks connected by long distance links and networks where routers with limited bandwidth connect group members. Such configurations are common in today's networks.

In this scenario, we simulate a clustered network with 80 nodes. The network consists of two 40-node fully connected clusters, and a single link connects those clusters where all nodes form an 80-member process group. Sender is located on the first cluster, and it generates 100 multicast messages per second. There is 1% intra-cluster noise formed in both clusters, and a high noise, of rate 20%, 40% or 50%, is injected on the link connecting the clusters. This inter-cluster noise rate corresponds to the probability that a message transmitted from the first cluster to the second will drop and hence get lost. We then explore the latency characteristics of a receiver on the second cluster.

In this configuration, both SRM and adaptive SRM deliver some messages with very long delays of many seconds. Particularly, in the adaptive case about 5% of all data messages are delayed by 5 seconds or more before delivery. On the other hand, Pbcast delivers all data messages within 1 second and hence can be seen as offering relatively steady data throughput in networks with this configuration.

Table 1. Hurst parameter of latency for clustered network

Noise Rate	Pbcast		SRM	
	before FIFO	after FIFO	non-adaptive	adaptive
20%	0.54	0.63	0.52	0.55
40%	0.49	0.61	0.66	0.59
50%	0.54	0.61	0.55	0.65

Table 1 gives the estimates of the Hurst parameter *H* for the three different noise rates and the two protocols. We see that *H* is very low for Pbcast before ordering at around 0.5 for all levels of noise rate. In these cases, the latencies have very low dependence among each other. As one would expect, FIFO ordering has an implication towards longer and more correlated delays over the network. This fact is demonstrated with higher values of *H*, around 0.6, for Pbcast after the ordering. Although this would classify the traffic as LRD, it is not as high as the values around 0.7 to 0.8 that are typical in the other high-speed data networks [AV98,BB98]. On the other hand, the value of *H* for both adaptive and non-adaptive SRM protocol varies from around 0.5 to 0.65, again with no specific pattern with the noise level. These values leading to only moderate LRD characterization as in the case of Pbcast after FIFO, do not have adverse implications on the network performance. However, even non-adaptive SRM could have long-range dependence, as a protocol, while Pbcast before FIFO does not.

4.3 Distribution functions for multicast latencies

The histograms obtained from latency data for clusters connected by a noisy link support the LRD analysis of Section 4.2. The Pbcast latencies are concentrated around low values (all less than 1 second) and the histograms look like normal and exponential distributions and/or their mixtures. Hence, the tails of the histograms decay exponentially. As an example, the latency distributions with 50% noise rate are given in Figure 2. For Pbcast before FIFO, an exponential distribution fits well, with mean 0.18-second, where *H* was found to be 0.54. For Pbcast after FIFO, a normal distribution fits well with mean 0.43-second, where *H* was found to be 0.61. In contrast, for the distributions of SRM latency the tails are prominent. Although an exponential distribution fits well for non-adaptive SRM latency as in Figure 2 (c), the mean delay, 1.22-second, is higher than that for Pbcast even after FIFO, and the distribution has long right tail. For this case *H* was found to be 0.55. The adaptive SRM was found to be LRD with *H* equal to 0.65. This is in agreement with the corresponding distribution in Figure 2 (d), which has a long right tail with very large observations and mean 1-second. A Pareto distribution, which is quite common in LRD, fits well in this case.

4.4 Impact of a member's distance from data source

In this part of the study, our interest is in the impact of receiver's distance from the data source on latency. Network settings are as follows: The network consists of 20 nodes with linear (chain) topology where first node is the sender spreading 100 messages per second to the process group, and there is 1% noise on the outgoing link from the sender. Remaining members are receivers. Each link has a transmission delay of 5msec. We analyzed application level latency distributions of all receivers and observed that the distribution is basically the same for the receivers.

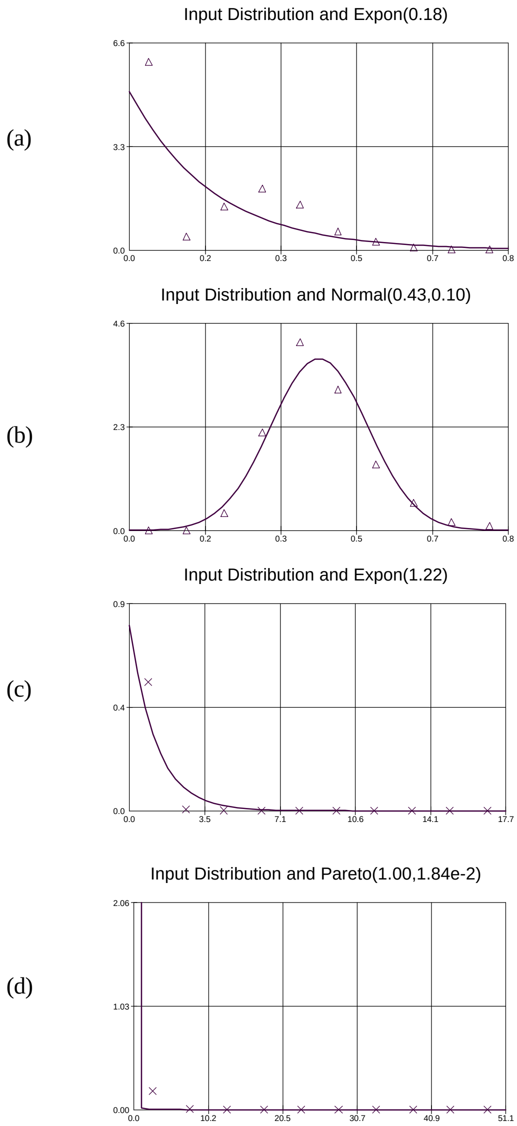


Figure 2. Latency distributions with 50% noise rate (a) Pbcast before FIFO, (b) Pbcast after FIFO, (c) non-adaptive SRM, (d) adaptive SRM.

In the linear topology, the traffic becomes strongly long-range dependent as observed from *H* values given in

Table 2. However, as suggested by the theoretical analysis of Pbcast [BHO99], the protocol is quite robust. The parameter H does not increase significantly with the distance of receiver from data source. Also, the histograms of the latencies indicate that the distribution does not change. The only effect is to introduce a small offset to the distribution, corresponding closely to the network delay itself. We obtain the same results for the other network topologies.

Table 2. Hurst parameter of latency for linear topology

	Distance from data source (in hops)			
	2	8	14	18
H	0.85	0.86	0.87	0.87

5. CONCLUSIONS

In this paper, the existence of long-range dependence in multicast message traffic over a high-speed network has been investigated. We have characterized LRD behavior through message latency distribution of Bimodal Multicast protocol in comparison with Scalable Reliable Multicast protocol (SRM). A single parameter, namely the Hurst parameter, is estimated as a representative measure of LRD.

We have found that Bimodal Multicast protocol does not intrinsically lead to long-range dependent traffic when the network topology is of tree or cluster type. After FIFO ordering, the correlations become more important. Yet, the traffic becomes only moderately LRD even in that case. On the other hand, SRM protocol shows LRD although at moderate level. Since LRD behavior of the traffic leads to adverse implications for network performance, we conclude that Bimodal Multicast is a superior protocol in this sense.

The traffic in the linear topology is found to be strongly LRD. In this case, the good properties of the Bimodal Multicast protocol cannot overcome the effect of the topology. However, the protocol is robust in the sense that LRD does not change with the distance of receiver from data source.

An area for further study would be modeling of the traffic in presence of LRD for network engineering such as prediction of buffer sizes and link capacity. Along the same line, one could consider optimizations for increasing the efficiency of the multicast transmission.

REFERENCES

[AV98] Abry, P., Veitch, D., Wavelet analysis of long-range-dependent traffic, *IEEE Transactions on Information Theory 44*: (1), 1998, 2-15.

[BB98] Borella, M.S., Brewster G.B., Measurement and Analysis of Long-Range Dependent Behavior of Internet Packet Delay, *Proceedings of IEEE INFOCOM 1998*, v. 2, 497-504.

[BBE99] Bajaj, S., Breslau, L., Estrin, D., et al., Improving Simulation for Network Research, 1999, USC Computer Science Dept. Technical Report 99-702.

[BHO99] Birman, K.P., Hayden, M., Ozkasap, O., Xiao, Z., Budiu, M. and Minsky, Y., Bimodal Multicast, *ACM Transactions on Computer Systems, 17*(2), 1999, 41-88.

[C2000] Caglar, M., 2000, Simulation of Fractional Brownian Motion with Micropulses, *Advances in Performance Analysis*, 3: 43-69.

[CB95] Crovella, M.E. and Bestavros, A, 1997, Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes in *IEEE/ACM Transactions on Networking, 5*(6), 835-846.

[CT90] Clark, D. and Tennenhouse, D.L., Architectural Considerations for a new Generation of Protocols, *Proceedings of the '90 Symposium on Communications Architectures and Protocols*, Philadelphia, PA, 1990, 200-208.

[DC90] Deering, S. and Cheriton, D., Multicast Routing in Datagram Internetworks and Extended LANs, *ACM Transactions on Computer Systems, 8*(2), 1990, 85-110.

[DGH87] Demers, A., Greene, D., Hauser, C., Irish, W., Larson, J., Shenker, S., Sturgis, H., Swinehart, D. and Terry, D., Epidemic Algorithms for Replicated Database Maintenance, *Proceedings of the Sixth ACM Symposium on Principles of Distributed Computing*, Vancouver, British Columbia, 1987, 1-12.

[FJL97] Floyd, S., Jacobson, V., Liu, C., McCanne, S. and Zhang, L., A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing, *IEEE/ACM Transactions on Networking, 5*(6), 1997, 784-803.

[GW94] Garrett, M. and Willinger, W., Analysis, Modeling and Generation of Self-Similar VBR Video Traffic, *In Proceedings of ACM SIGCOMM*, 1994.

[K95] Kumar, V., 1995, *Mbone: Interactive Multimedia on the Internet*, New Riders Publishing, Indianapolis, Indiana, USA.

[LLS92] Ladin, R., Lishov, B., Shrira, L. and Ghemawat, S., Providing Availability using Lazy Replication, *ACM Transactions on Computer Systems, 10*(4), 1992, 360-391.

[LMJ97] Labovitz, C., Malan, G.R. and Jahanian, F., Internet Routing Instability, *Proceed. of SIGCOMM`97*, 1997, 115-126.

[LOM94] Lidl, K., Osborne, J. and Malcome, J., Drinking from the Firehose: Multicast USENET News, *USENIX Winter* 1994, 1994, 33-45.

[LTWW] Leland, W.E., Taqqu, M.S., Willinger,W., Wilson, D.V., On the Self-Similar Nature of Ethernet Traffic (Extended version), IEEE/ACM Transactions on Networking, 2: (1), 1-15, 1994.

[N95] Norros, I., On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks, IEEE Jour. Selec. Areas in Comm., 13, 1995, 953-962.

[O2000] Ozkasap, O., Scalability, Throughput Stability and Efficient Buffering in Reliable Multicast Protocols, PhD. Dissertation, Ege University, Department of Computer Engineering, Izmir, Turkey, May 2000. (Available as Technical Report, TR2000-1827, Department of Computer Science, Cornell University)

[P97] Paxson, V., 1997, End-to-End Internet Packet Dynamics, Proceedings of SIGCOMM `97, 139-154.

[PF95] Paxson, V. and Floyd, S., Wide-Area Traffic: The Failure of Poisson Modeling, IEEE/ACM Transactions on Networking, 2: (3), 226-244, 1995.

[VA99] Veitch, D. and Abry, P., A Wavelet Based Joint Estimator of the Parameters of Long-Range Dependence, IEEE Trans. on Information Theory, 45(3), 1999, 878-897.