1 Future Directions and Open Problems in Performance Evaluation and Control of Self-Similar Network Traffic

Kihong Park

Network Systems Lab, Department of Computer Sciences, Purdue University, West Lafayette, IN 47907

1.1 INTRODUCTION

Since the seminal study of Leland *et al.* [41] on the self-similar nature of network traffic, significant advances have been made in understanding the statistical properties of measured network traffic—in particular, Internet workloads—why self-similar burstiness is an ubiquitous phenomenon present in diverse networking contexts, mathematical models for their description and performance analysis based on queueing, and traffic control and resource management under self-similar traffic conditions. Chapter 1 gives a comprehensive overview including a summary of previous works, and the individual chapters give a detailed account of a cross section of relevant works in the area. Chapter 20 provides a discussion of traffic and workload modeling, with focus on long vs. short time scales and nonuniform scaling observed in wide area IP traffic [23, 24].

This chapter presents a broad outlook into the future in terms of possible research avenues and open problems in self-similar network traffic research. The specific items described in the chapter are but a subset of interesting research issues, and is meant to highlight topics that can benefit from concerted efforts by researchers in the community due to their scope and depth. The research problems are organized around recent developments and the landscape of previous accomplishments, grouped into three areas—workload characterization, performance analysis, and traffic control. Physical modeling, which can be viewed as a fourth category, is grouped with workload characterization.

Workload Characterization The original focus of self-similar burstiness in local area and wide area network traffic has expanded into the generalized framework of *workload modeling* which captures source behavior and structural properties of network systems, not necessarily restricted to network and link layers. This stems, in part, from the realization that network performance—as measured by packet drop, queueing delay, and jitter at multiplexing points in the network—is affected by a multitude of factors including variability of streamed real-time VBR video, connection arrival patterns, their durations, the make-up of files being transported, control actions in the protocol stack, and user behavior that drives network applications. Increasingly, these activities transpire under the umbrella of the World Wide Web (WWW), and characterizing the structural properties—static and dynamic—of the global wired/wireless Internet that impact network performance has become an important goal. The research challenge lies in identifying, quantifying, and modeling *invariant*—or "slowly changing"—system traits in the midst of a rapidly growing network infrastructure that are relevant to network performance.

Performance Analysis Performance analysis of queueing systems with self-similar input has yielded the fundamental insight that queue length distribution decays polynomially vis-à-vis the more accustomed case of exponential decay with Markovian input. In the resource provisioning context, this is interpreted to mean that resource dimensioning using buffer sizing is an ineffective policy relative to bandwidth allocation. There remain a number of challenges. First, the queueing results are asymptotic in nature where buffer capacity—in some form—is taken to infinity to achieve tractability. Little is known about finite buffer systems except for observations on the dependence of packet loss rate on the "effective" time scale induced by buffer size, and its delimiting impact on correlation structure at larger time scales with respect to its influence on queueing [28, 60]. Second, performance evaluation with self-similar traffic has concentrated on first-order performance measures—i.e., packet loss rate (captured in various forms)—which is but one, albeit important, yardstick. In the modern network environment with multimedia and other QoS-sensitive traffic streams comprising a growing fraction of network traffic, second-order performance measures in the form of "jitter" such as delay variation and packet loss variation are of import to provisioning user-specified QoS. Self-similar burstiness is expected to exert a negative influence on second-order performance measures and multimedia traffic controls—e.g., packet-level FEC—that are susceptible to concentrated packet loss. Third, performance analysis is carried out in equilibrium which may be problematic for self-similar workloads given their slow convergence properties. As a related point, the bulk of TCP connections is known to be short-lived, and there is a disconnect between steady-state techniques and performance evaluation of short- and medium-duration flows. The same problem exists when using simulation as the principal performance evaluation tool.

Traffic Control Traffic control for self-similar traffic has been explored on two fronts: one, as an extension of performance analysis in the resource provisioning context, and two, from the multiple time scale traffic control perspective where correlation structure at large time scales is actively exploited to improve network performance. The resource provisioning aspect advocates a small buffer/large bandwidth resource dimensioning policy which—when coupled with the central limit theorem—yields predictable multiplexing gains when a large number of independent flows are aggregated. Whereas resource provisioning is open-loop in nature, multiple time scale traffic control seeks to achieve performance gains by exploiting correlation structure in self-similar traffic at time scales exceeding the time horizon of the feedback loop to impart proactivity to reactive controls (e.g., TCP). This is relevant in broadband wide area networks where the delay-bandwidth product problem is especially severe, and mitigating the performance degradation due to outdated feedback critical to facilitating scalable, adaptive traffic control. The initial success of this approach [67, 68, 62, 69] (see Chapter 18 for an application to rate-based congestion control) leads to a generalization to *workload-sensitive traffic control* where facilitation of workload-sensitivity is expanded along several traffic control dimensions including the two core features for harnessing predictability at large time scales: long-range correlation in network traffic and heavy-tailedness of connection durations. Workload-sensitive traffic control is a broad area that can benefit from concerted efforts at several fronts spanning novel mechanisms for detecting and exploiting large time scale predictability structure, short-duration connection management, packet scheduling, end system support, and dynamic admission control with self-similar call arrivals and/or heavy-tailed connection durations.

1.2 OPEN PROBLEMS IN WORKLOAD CHARACTERIZATION

1.2.1 Physical Modeling

Unlike in many systems of study including economic, social, and certain physical sciences (e.g., astronomy, earth and atmospheric science), network systems admit to design, implementation, and controlled experimentation of the underlying physical system at nontrivial scales—e.g., protocol deployment in autonomous sytems belonging to a single service provider—which facilitates an intimate, mechanistic understanding of the system at hand. Model selection is not bound by "black box" evaluations, and physical models that can explicate traffic characteristics in terms of elementary, verifiable system properties and network mechanics, in addition to data fitting, provide an opportunity to be exploited. The challenge lies in combining relevant features from workload modeling, network architecture—protocols and transmission technology—user behavior, and analytical modeling into a consistent, effective description of network systems, in particular, the Internet. As such, physical modeling is a research program that transcends workload modeling, encompassing both performance analysis and traffic control.

1.2.2 Multifractal Traffic Characterization

Since the collection and analysis of the Bellcore Ethernet LAN data [41], follow-up works [1, 15, 27, 57] have shown the robustness of self-similar burstiness in network traffic. This has led to the heuristic description: *Poisson connection arrivals with heavy-tailed connection duration times leads to self-similar burstiness in the multiplexed network traffic*. This is a rough, "first-order" description of the empirical facts—e.g., TCP connection arrivals exhibit self-similarity (see Chapter 15 on TCP workload modeling)—which serves to point toward the principal causal attribute of self-similarity: heavy-tailed activity durations. Recent analysis of WAN IP traffic [24, 23] has revealed multifractal structure in the form of nonuniform scaling across short and long time scales (see Chapter 20 for a comprehensive discussion). That is, on top of the monofractal picture captured by the heuristic statement above, there exists further variability within each connection—in particular, heavy-tailed TCP connection lifetimes—that fall outside the scope of monofractal self-similarity which principally concerns large time scale structure in network traffic. The refined, short time scale structure can be described by cascade constructions—also used in the generation of deterministic fractals such as 2-dimensional grey-scale fractal images [5]—where variability (within a connection) is obtained by recursive application of "measure redistribution" according to some fixed rule (cf. Chapter 1, Figure 2 (middle)). Several problems remain unsolved:

Multiplicative Scaling and Causality What causes multiplicative scaling observed for short-range correlation structure? Is it related to fragmentation in the TCP/IP protocol stack (including the MAC layer)? TCP's feedback control (ARQ and window-based congestion control)? ACK compression? Topological considerations? If a combination, are there dominant factors? Cascades—although suggestive of certain physical causes—are ultimately a data modeling construct, and fall short of establishing a mechanistic description of the underlying workload. From a workload generation or synthesis perspective, given the possible dependence of multiplicative scaling in short time scale traffic structure on feedback control, an open-loop generation of traffic is unsatisfactory for closed-loop traffic control and its performance evaluation purposes.

Impact of Refined Short Time Scale Modeling Is multiplicative scaling a robust, invariant phenomenon as self-similarity is for large time scale structure? Can modeling of short time scale structure lead to a better understanding of dynamical properties of network protocols? Does a refined model of short-range structure lead to a more accurate prediction of network performance? In other words, is refined modeling of short time scale structure in network traffic a "relevant" research activity? It is clear that in some contexts (see, e.g., Chapter 12 for a discussion of short-range vs. long-range dependence issues) short-range structure can dominate performance. Refined traffic modeling, in general, if not checked with respect to its potential to advance fundamental understanding, can become a "data fitting" activity—the subject of time series analysis—yielding limited new networking insights. The standards required of refined traffic modeling work must therefore be evermore stringent.

1.2.3 Spatial Workload Characterization

Physical modeling carried out in [15, 51] which reduces the root cause of self-similarity in network traffic to heavy-tailed file size distributions on file systems and Web servers is a form of *spatial* workload modeling. That is, the *temporal* property of network traffic—which is a primary factor determining performance—is related to the spatial or structural property of networked distributed systems. Following are a number of extensions to the spatial workload modeling theme that may exhibit features related to "correlation at a distance," characteristic of self-similarity in network traffic.

Mobility Model In an integrated wired/wireless network environment, understanding the movement pattern of mobiles is relevant for effective resource management and performance evaluation. Current models are derived from transportation studies [19, 34, 40] which possess a coarse measurement resolution, or more commonly, make a range of user mobility assumptions including random walk, Poisson number of base stations/cells visited as a function of time, and exponential stay durations whose validity is insufficiently justified. It would not be too surprising to find correlation structure at large time and/or space scales—a user, after the morning commute, may stay at her office for the remainder of the day except for brief excursions, students on a campus move from class to class at regular intervals and in clusters, users congregate in small regions (e.g., to take in a baseball game at a stadium) in numbers significantly exceeding the average density, traffic commute obeys predictable flow patterns—which, in turn, can impact performance due to sustained load on base stations connected to wireline networks. A measurement-based mobility model (and tools for effective tracing [48]) that accurately characterizes user mobility is an important component of future workload modeling.

Logical Information Access Pattern With the Internet and the World Wide Web becoming interwoven in the socio-economic fabric of everyday life, it becomes relevant to characterize information access pattern by information content (in addition to geographical location) so as to facilitate efficient access and dissemination. Popular Web sites—i.e., URLs—may be accessed more frequently than less popular URLs in a statistically regular fashion, for example, with access frequency obeying power laws as a function of some popularity index (e.g., ranking). Hypertext documents and hyperlinks can be viewed as forming a directed graph, and the resulting graph structure of the World Wide Web can be analyzed with respect to its connectivity in an analogous manner as has been carried out recently for Internet network topology [22]. An *information topology* project that parallels efforts in Internet topology and distance map discovery (e.g., IDMaps [26, 37]), and identifies how logical information is organized on the World Wide Web—including possible invariant scaling features in its connectivity structure and access pattern—may have bearing on network load/temporal traffic properties, and consequently, network performance.

User Behavior Most network applications are driven by users—e.g., via interaction with a Web browser GUI—and thus the connection, session, or call arrival process is intimately tied with user behavior, in particular, as it relates to network state. Starting with the time-of-day, user behavior may be a function of network congestion leading to self-regulation (a user may choose to continue his Web surfing activities at a later time if overall response time is exceedingly high, a form of backoff), congestion pricing may assign costs above and beyond those exacted by performance degradation, users may switch between different service classes in a multiservice network [10, 20], users may perform network access and control decisions cooperatively or selfishly leading to a noncooperative network environment characteristic of the Internet, users may observe behavioral patterns when navigating the Web, and so forth. The challenge lies in identifying robust, invariant behavioral traits—possibly exhibiting scaling phenomena—and quantify their influence on network performance.

Scaling Phenomena in Network Architecture The recent discovery of power law scaling in network topology [22] points toward the fact that scaling may not be limited to network traffic and system workloads. On the other hand, power law scaling in the connectivity structure of the Internet stretches the meaning of "workload characterization" if it is to be included under the same umbrella. More importantly, it is unclear whether the diffusive connectivity structure implied by power laws affects temporal traffic properties and network performance in unexpected, nontrivial ways. For example, routing in graphs with exponential scaling in their connectivity structure is different from routing in graphs with power law scaling, but that is not to say that this has implications to traffic characterization and performance above and beyond its immediate scope of influence—number of paths between a pair of nodes, their make-up, and generation of "realistic" network topologies for benchmarking. If the distribution of link capacities were to obey a power law, then it is conceivable that this may exert a *traffic shaping* effect in the form of variable stretching-in-time of a transmission which can inject heavy-tailedness in transmission or connection duration that is not present in the original workload. The challenge in architectural characterization lies in identifying robust, invariant properties exhibiting scaling behavior, and relating these properties to network traffic, load, and performance where a novel and robust relationship is established.

1.2.4 Synthetic Workload Generation

An integral component of workload modeling is synthetic workload generation. In many instances, in particular, those where the workload model is constructive in nature, the process of generating network traffic is suggested by the model under consideration. There are two issues of special interest to self-similar traffic and workload generation that can benefit from further investigation.

Closed-loop Workload Generation Many traffic generation models are *time series* models which output a sequence of values—interpreted as packets or bytes per time unit—which are then fed to a network system (simulation or physical). These "open-loop" synthetic traffic generation models can be used to evaluate queueing behavior of finite/infinite buffer systems with self-similar input, they can be used to generate background or cross traffic impinging on bottleneck routers, and they can serve as traffic flows that are controlled in an open-loop fashion including traffic shaping. In network systems governed by feedback traffic controls where source output behavior is a function of network state, open-loop traffic generation is, in general, ill-suited due to its a priori fixed nature which does not incorporate dependence on network state. Traffic emitted from a source is influenced by specific control actions in the protocol stack—e.g., TCP Tahoe, Reno, Vegas, and flow-controlled UDP—and capturing network state and protocol dependence falls outside the scope of open-loop traffic generation models. A closed-loop traffic generation model for self-similar traffic that captures both network and protocol dependence—based on physical modeling—works by generating file transmission events with heavy-tailed file size distribution at the application layer, lets each file transmission event pass through the protocol stack (e.g., TCP in the transport layer), which then results in packet transmission events at the network/link layer. The consequent traffic flow reflects both the control actions such as reliability, congestion control, fragmentation, and buffering undertaken in the protocol stack as well as feedback from the network. The closed-loop workload generation framework allows the effect of different control actions on network traffic and performance to be discerned and evaluated. Several issues remain: Which connection arrival model should be used at the application layer (e.g., exponential vs. heavy-tailed inter-connection arrival times), and for what purpose? Should the arrival time of the next connection be counted from the time of completion of the previous connection, or independently/concurrently? How sensitive are the induced traffic properties and network performance to details in the application layer workload model (cf. Chapter 14 for related results)? Are there conditions under which traffic generated from closed-loop workload models can be approximated by open-loop traffic synthesis models? For example, the use of independent loss process models for tractable analysis of TCP dynamics [50] is an instance of open-loop approximation. It is important to delineate the conditions under which open-loop approximation is valid as it is possible to "throw out the baby with the bathwater."

Sampling from Heavy-tailed Distributions The essential role played by heavy-tailedness in self-similar traffic models renders sampling from heavy-tailed distributions a key component of synthetic workload generation models (e.g., on/off, M/G/∞, physical models). As discussed in Chapters 3 and 1, sampling from heavy-tailed distribution suffers from convergence and underestimation problems where the sample mean $\bar{X}_n$ of a heavy-tailed random variable X converges *very slowly* to the population mean (cf. Figure 2 of Chapter 3). For researchers accustomed to light-tailed distributions—e.g., exponential (Markovian models), Gaussian (white noise generation)—where convergence is exponentially fast, it is possible to use heavy-tailed distributions in performance evaluation studies without explicitly considering its idiosyncracies and potentially detrimental consequences on the conclusions advanced. For example, a common mistake arises when comparing short-range and long-range dependent traffic models with respect to their impact on queueing where self-similar input is generated using heavy-tailed random variables. The traffic rate is assumed equal by virtue of the configured distribution parameters. As a case in point, short-range and long-range dependent traffic may be generated from an on/off source where the off-periods are exponential and i.i.d., but the on-period is exponential with parameter $\lambda > 0$ for short-range dependent input and Pareto with shape parameter $1 < \alpha < 2$, location parameter $k > 0$ for long-range dependent input. For α close to 1, k and λ may be chosen such that the population mean of on-periods in the two cases are equal, i.e., choose k and λ such that

$$\frac{1}{\lambda} = \frac{k\alpha}{\alpha - 1}.$$

Unless the number of samples is "extremely" large—and the traffic series correspondingly long—the actual traffic rate of sample paths in long-range dependent traffic will be nonnegligibly smaller than the corresponding traffic rate of short-range dependent traffic. Thus observations on packet loss and other performance measures may stem from sampling errors—in particular, smaller traffic intensity due to insufficient samples in the self-similar case—than differences in correlation structure of the inputs. How to remedy the problem? Cognizance of the problem is necessary, but not sufficient, to address the potential pitfalls of the problem. We can consider three approaches. One, perform *sufficient* sampling such that statistics from sample paths approach that of population statistics. This is the most straightforward approach. The main drawback is that events (e.g., lifetime of connections) and performance measurements of interest may occur at time scales significantly smaller than that required to reach steady state. Also, the shere sample size and corresponding time requirement puts a heavy computational burden on simulation and experimental studies. Two, perform various forms of sample path normalization. For example, the traffic intensity λ_{on} (bps) during on-periods can be varied such that the actual traffic rate matches that of a pre-specified target. This is most suited for open-loop workload generation (e.g., CBR or VBR traffic over UDP). The main justification of this approach is that λ_{on} does not affect the correlation structure of the generated traffic series. For closed-loop workload generation, one may vary k such that sample path normalization with respect to first-order properties is achieved. Again, correlation structure or second-order properties are not affected by k. This is a heuristic approach and the values of λ_{on} and k depend on the sample size, and must be empirically calibrated. Since first-order performance measures such as packet loss rate and average queueing delay are heavily impacted by offered load—in some instances dominating the influence of second-order structure—it is pertinent to perform sample path normalization if the effect of correlation structure on performance is to be discerned. The fundamental soundness of this approach, however, requires further investigation. When sample sizes are insufficient to yield matching sample and population statistics, second-order properties of the generated traffic may be impacted as well. How severe is the sampling problem with respect to second-order structure? Are "corrections" viable? If a certain number of samples is needed to achieve sample paths with statistics approaching that of the population distribution, what fundamental justification is there to allow short-cutting the required sampling process? Perhaps, long stretches of time where the sample mean of long-range dependent traffic is significantly smaller than that of short-range dependent traffic is the natural state-of-affairs (i.e., with respect to network traffic), during which first-order properties dominate second-order properties in impacting performance. In the long run, there are bound to be stretches of time where the opposite is true. This is an intrinsic problem with no simple answers. Three, as a continuation of the second approach, the investigation of speed-up methodologies is the subject of rare event simulation [4, 61] where various techniques including extreme value theory, large deviations theory and importance sampling are employed to establish conditions under which simulation speed-up is possible. In the case of light-tailed distributions, simulation speed-up using importance sampling is well-understood, however, the heavy-tailed case is in its infancy and remains a challenge [4].

1.2.5 Workload Monitoring and Measurement

Systematic, careful monitoring of Internet workloads is a practically important problem. It would be desirable to have a measurement infrastructure that filters, records, and processes workload features at sufficient accuracy which, in turn, is essential to reliably identifying invariant features and trends in Internet workloads. It is unclear whether there are open research problems related to workload monitoring and measurement instrumentation above and beyond a range of expected engineering issues—placement of instrumentation, what to log, efficient probing (resource overhead, minimally disturb Schrödinger's cat), efficient storage, synchronization, and so forth. It is possible that there are hidden subtleties but, if so, they await to be uncovered. Given the recent interest in Internet topology, distance map, and "weather map" discovery (see, e.g., [26]), integration and coordination of various measurement and estimation related activities may deserve serious consideration. A laisser faire approach without coordinated efforts may be impeded by protective walls set up by service providers with respect to autonomous systems under private administrative control, which can render certain measurement efforts difficult or infeasible.

1.3 OPEN PROBLEMS IN PERFORMANCE ANALYSIS

1.3.1 Finite Buffer Systems and Effective Analysis

Queueing Analysis of Finite Buffer Systems Most queueing results with self-similar input are asymptotic in nature where either buffer capacity is assumed infinite and the tail probability of queue length distribution in steady state $\Pr\{Q_\infty > x\}$ estimated as $x \to \infty$, or buffer capacity b is assumed finite but buffer overflow probability is computed as b becomes unbounded. Little is known about the finitary case, and the works in [28, 60] provide approximate, heuristic arguments regarding the impact of finite time scale implied by bounded x and b. Large deviation techniques [64] are too coarse to be effectively applied to finite x and b, and not surprisingly, the unbounded case or bufferless queueing case (i.e., $b = 0$) are more easily amenable to tractable analysis. The bufferless case can provide indirect insight on performance with "small" buffer capacities, and complements the conclusions advanced in the asymptotic case (cf. Chapter 17 for a discussion of bufferless queueing with self-similar input). The divide between our understanding of unbounded and zero memory systems, on the one hand, and finitary systems of interest, on the other, limits the applicability of these techniques both quantitatively and qualitatively—above and beyond polynomial decay of queue length distribution and its broad interpretation as amplified buffering cost—to resource provisioning and control. The difficulty underlying analysis of finite buffer systems with non-Markovian, in particular, self-similar input is a fundamental problem at the heart of probability theory, and, perhaps, beyond the scope of applied probability. Fundamental advancement in understanding, proof techniques and tools are needed to overcome the challenges—a longer term venture. For networking applications, this points toward the need for experimental queueing analysis to fill the void in the interim. As discussed in Section 1.2.4, there are a number of problems and issues associated with performance evaluation under workloads involving sampling from heavy-tailed distributions due to slow convergence of sample statistics to population statistics. When empirical performance evaluation is carried out with synthetic traffic—in addition to measurement traces—which are then used to support generalizations and comparative evaluations, extreme care needs to be exercised to check the influence of sampling. This is a highly nontrivial problem unto its own, and provides an opportunity for theoretical advances in rare event simulation with heavy-tailed workloads [4] to facilitate experimental queueing analysis and performance evaluation.

Tight Buffer/Packet Loss Asymptotics Significant effort has been directed at deriving tight upper and lower bounds for the tail of the queue length distribution of various queueing systems (e.g., on/off, M/G/∞, or FBM input and constant service rate server) with long-range dependent input [11, 17, 18, 38, 42, 43, 49, 56, 66]. Most of the approaches can be viewed in the framework of large deviation analysis where the queue length process is shown to obey a large deviation principle (LDP) with specific rate function and time scale, assuming the arrival process satisfies LDP. Irrespective of the limited applicability and impact to network design and control, refined characterization of the large buffer asymptotics is of independent interest, and relevant for advancing the foundations of queueing theory. We refer the reader to the queueing analysis chapters in the book (see, e.g., Chapters 8, 9, and 10) for a detailed discussion of related research issues. Chapter 9, in addition, provides an excellent overview of recent results.

1.3.2 Second-order Performance Measures

Impact of Self-similarity on Jitter Performance evaluation under self-similar traffic conditions has focused on first-order performance measures such as packet loss rate and mean queueing delay. Second-order performance measures that relate to jitter—e.g., delay variance and packet loss variance—are of import to multimedia data transport and QoS provisioning. Even under conditions where self-similarity has limited impact on performance with respect to first-order measures, persistent periods of high- and low-contention implied by self-similar burstiness can exert a negative effect on second-order performance measures. Figure 1.3.1 shows packet drop traces at a bottleneck router where 32 TCP Reno connections are multiplexed on a common output port. Each connection transports heavy-tailed files, and the four traces stem from the same set-up except that the shape parameter (or tail index) of the heavy-tailed distribution (Pareto) is varied in the range 1.05, 1.35, 1.65, 1.95. The plots depict packet drop traces at 100 second time aggregation. We observe significant variation in packet drops across different 100 second time segments for the $\alpha = 1.05$ tail index case, which diminishes as α approaches 2. This is even more

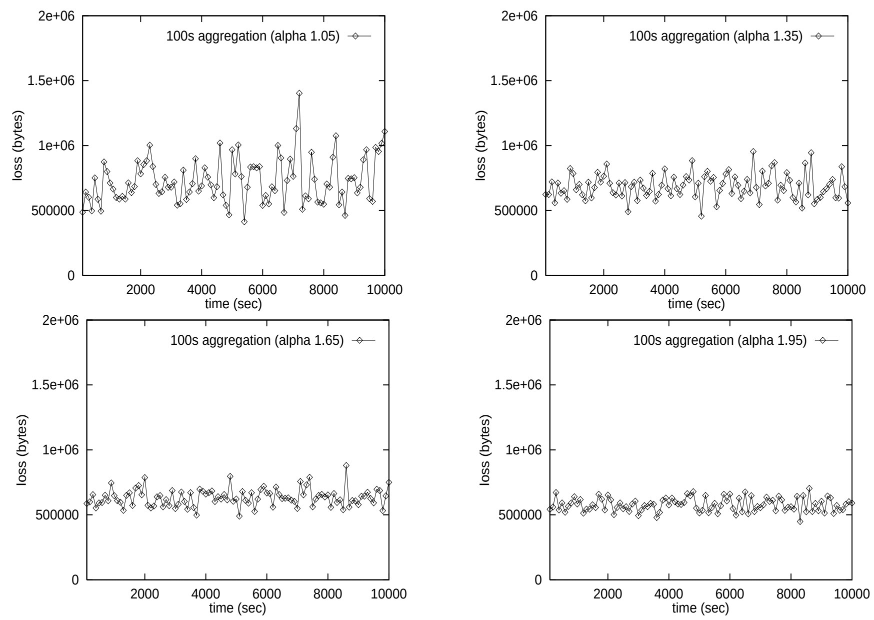

Fig. 1.3.1 Packet drop trace at bottleneck router multiplexing 32 TCP connections that transport heavy-tailed files, at 100 second time aggregation. Top row: file size distribution with tail index $\alpha = 1.05$ (left) and tail index 1.35 (right). Bottom row: Corresponding traces for shape parameter 1.65 (left) and $\alpha = 1.95$ (right).

pronounced when the flows are open-loop controlled over UDP. Given the difficulties underlying queueing analysis of finite buffer systems with self-similar input with respect to first-order performance measures, the challenges facing queueing analysis for second-order measures are even greater. There is, however, a special case—bufferless queueing—which is more amenable to tractable analysis (see Chapter 17 for a discussion of bufferless queueing in the context of predictive control). Bufferless queueing, which can be viewed as an extreme form of the small buffer/large bandwidth resource provisioning strategy, derives its justification from (1) the high delay penalty associated with long-range dependent traffic, and (2) the observation that a large number of independent flows—correlated in time or not—when aggregated over space (i.e., flows) is approximately Gaussian by the central limit theorem. Thus, for second-order stationary processes (most traffic models fall under this class) this yields a simple handle on the marginal distribution which can be used to compute deviation probabilities—the aggregate traffic exceeds a given link capacity—using the tail of the Gaussian. Real-time multimedia traffic such as video and audio can tolerate some packet loss (described by first-order measures), but for the same packet loss rate, concentrated packet drops exert a more detrimental impact on QoS than dispersed losses. (Similarly for delay variations in buffered systems.) A comprehensive understanding of the packet loss process and, in general, delay variation in buffered systems is needed to complement the hereto one-sided focus on first-order performance measures when evaluating the effect of self-similar burstiness on network performance.

Impact of Self-similar Burstiness on Packet-level FEC As a continuation of the second-order performance measure issue, we draw attention to the impact of self-similar burstiness on packet-level forward error correction (FEC). Packet-level FEC is a form of error control which injects redundancy into a packet stream such that reliable transmission is achieved in the presence of packet drops or erasures. That is, given $k \geq 1$ data packets (for simplicity assume fixed size), encoding results in $n = k + h$ packets ($h \geq 0$) which satisfy the property that from any k subset—i.e., not more than h packets are lost—the original k packets can be recovered through decoding. When transporting multimedia traffic with real-time constraints over wide area networks where end-to-end latency renders ARQ infeasible, FEC facilitates QoS control where reliability is proactively affected. The effectiveness of FEC hinges on the "k-out-of-n" property being realized over an end-to-end path which is impeded by concentrated packet drops. As pointed out in [7, 46], burstiness resident in VBR video can significantly impact error recovery. Packet-level FEC injects further complexity into the dependency structure of loss processes due to queueing. In real-time data transport (e.g., MPEG video at a specified frame rate), packets belonging to a frame must, in addition, arrive in a timely manner [53, 54]. Otherwise, they are considered as useless as if they had been dropped by the network. It has been shown that correlated erasures stemming from queueing can significantly impede the efficacy of packet-level FEC when compared to independent packet drops [2, 12, 13, 14]. Little is known, however, about performance analysis of packet-level FEC under self-similar burstiness—either from the source traffic itself or interference from cross traffic—and looms as an important challenge. As in the general second-order performance measure case, a starting point is bufferless queueing under self-similar input where the packet loss process with respect to a block of n consecutive packets—the packets belonging to a self-similar traffic stream are viewed as totally ordered—is analyzed. For simplicity, n can be considered fixed although, in general, n is variable.

1.3.3 Short-range vs. Long-range Correlation

Related to refined traffic modeling (see Section 1.2) is the issue of short-range vs. long-range correlation in determining queueing behavior and performance in network systems. This book is meant, in part, to clarify some of the surrounding issues given the mixed—and, sometimes, conflicting—messages and conclusions advanced in various works [16, 21, 28, 32, 52, 60]. It has been shown that buffer capacity and time scale, traffic intensity and bandwidth, marginals, payload type (e.g., aggregated data traffic, VBR video), and the performance measure of interest—first-order vs. second-order statistics—collectively determine what the relative import of short time scale and long time scale structure is on performance. Recent works along the line "here is a traffic model with controllable short-range and long-range structure, and short-range, or alternatively, long-range structure is dominant in impacting performance" oftentimes provide insufficient comparative evaluation of related works yielding one-sided and, on the surface, contradictory conclusions. The reader may take away the message that there are few unconditional truths in performance evaluation with self-similar traffic, modeling of Internet traffic admits a large degree of freedom in choosing models (i.e., assumptions) and parameters, and queueing with self-similar input is but one—albeit important—facet of self-similar network traffic research. On the other hand, as a science with engineering applications to network design, resource provisioning and control, further clarification efforts that focus on carefully qualified comparative evaluations are needed to distill the facts, assumptions, and derived conclusions into a mutually consistent and coherent description—unless the works contain technical errors, by definition, this is possible—where assumptions and opinions are delineated from scientific facts. Without these efforts, ambiguities and resulting confusion may put forth avoidable barriers to effectively applying the lessons and knowledge learned from self-similar traffic research to networking practice.

1.3.4 Queueing Analysis of Feedback Control Systems

Analysis of feedback controlled queueing systems is a difficult problem. Tractability is achieved by considering queueing systems with state-dependent arrival rates and admission control systems where arrivals are admitted/rejected according to a decision rule [33, 65]. In both cases, injection of control is carefully administered such that the Markov property is preserved. Since the bulk of current Internet traffic is governed by TCP—a complicated feedback congestion control—and its state-dependent actions may influence the very traffic being measured and analyzed, it is important that the influence of feedback control on traffic characteristics and performance be ascertained. For example, in the multifractal characterization of network traffic [23], it is conjectured that multiplicative scaling observed in short-range structure is influenced by TCP's control actions. If so, why is this the case and what are the underlying mechanisms? In [51] (see also Chapter 14), it is shown that multiplexing of concurrent TCP connections at a bottleneck router transporting heavy-tailed files leads to self-similar burstiness, but the empirical slope of the Hurst parameter curve as a function of the tail index is less than $-1/2$, the slope value implied by the relation $H = (3 - \alpha)/2$ stemming from the on/off model where connections are assumed independent [70]. Does coupling among feedback controlled connections sharing common resources lead to changes in traffic properties? Does it matter whether the feedback congestion control is TCP, rate-based control, or adaptive FEC? TCP, because of its idiosyncracies and historical evolution, is not an easy protocol to analyze. Tractable analysis of its dynamics, to date, is only achieved by making an independence assumption on the loss process [45, 50] which, in general, is a heavy price to pay for tractability. A more fundamental avenue of exploration is the analysis of feedback congestion controls including linear increase/exponential decrease controls—a tractable exercise without injecting decoupling by assumption—which have been investigated, principally, for infinite source models [9, 25, 47, 55, 63, 71]. For heavy-tailed workloads where most file transfers are small and a few very large, little is known about the consequent system behavior. For large file transfers, steady-state may be reached due to its approximation of an infinite source. For small transfers—the bulk of connections—analysis remains a challenge. In a similar vein, the impact of feedback on traffic properties when multiplexing a number of heavy-tailed sources has not been sufficiently investigated. The dynamics and effect of feedback congestion control has been studied with respect to fairness, stability, and synchronization issues; traffic properties represent another dimension to its multifaceted influence on network performance.

1.3.5 Impact of Packet Scheduling

In modern routers, flow protection and multi-service support is provided in the form of configurable service classes—per-flow or aggregate-flow—over various scheduling disciplines including GPS, priority queues, and RED. From a performance analysis perspective, the question arises as to what impact self-similar burstiness has on packet scheduling, from subtle effects such as the influence on resource sharing and performance across service classes due to work conservation, to more active design questions such as optimal scheduling with respect to target objective functions. The comparative evaluation of FCFS, PS, and LCFS-PR under heavy traffic conditions (see Chapter 6) provides insight into the role of scheduling with self-similar input. Performance analysis of EDF, fixed priority schemes, and related scheduling algorithms under self-similar input remain to be investigated. Refined analysis of work conservation and its impact on GPS with respect to providing inter-flow (in general, inter-service class) protection under self-similar workloads looms as a challenging problem. At a broader level, influence of self-similarity and heavy-tailedness on scheduling need not be restricted to routers. Empirical evidence of heavy-tailedness across UNIX process lifetime distribution [29, 30], UNIX file size distribution [35, 51], and Web document size distribution [3, 15], points toward CPU scheduling policies that make active use of the heavy-tailed property. For example, given the empirical observation that most tasks require short service times whereas a few require very long service times, a shortest job first (SJF) scheduling policy—known to be optimal with respect to average waiting time—is expected to yield amplified performance gain vis-à-vis FCFS and other workload-insensitive schedulers. The technical challenge lies in achieving tractable analysis of relevant performance measures such as waiting time under heavy-tailed workloads when using SJF or other workload-sensitive schedulers. The service time of a task may be known a priori—e.g., if related to the size of documents at Web servers—or, it may be estimated on-line. Heavy-tailedness implies predictability—if a task has been active for some time, then it is likely to persist into the future (see Chapter 1, Section 4)—which can be used to perform on-line identification of long-running tasks.

1.4 OPEN RESEARCH PROBLEMS IN TRAFFIC CONTROL

1.4.1 Closed-loop Traffic Control

Multiple Time Scale Traffic Control Self-similarity and long-range dependence implies predictability structure at large time scales that may be exploitable for traffic control purposes. Most feedback traffic controls are impervious to this information, principally, due to the fact that the time scale of the feedback loop—i.e., round-trip time—is an order of magnitude (or more) smaller than the time scale at which "long-range" correlation structure, in practice, manifests itself: millisecond vs. second range. The multiple time scale traffic control framework was introduced in [67], and shown to be effective at yielding significant performance improvement when large time scale correlation is exploited for traffic control. In [67] (see also Chapter 18), large time scale correlation structure was on-line estimated and utilized to modulate the bandwidth consumption behavior of linear increase/exponential decrease rate-based congestion control for throughput maximization. In [68], the approach was adopted to window-based congestion control and reliable transport using TCP (e.g., Reno and Vegas) with similar performance gains. In [69], multiple time scale traffic control was extended to adaptive redundancy control where adaptive packet-level FEC is used for end-to-end QoS control of real-time traffic [8, 54]. An important benefit of multiple time scale traffic control is the mitigation of performance cost of reactive controls in broadband wide area networks with a large delay-bandwidth product due to outdated feedback. The mechanism for exploiting large time scale structure employed in [67, 68, 69] couples the feedback traffic control (i.e., short time scale module) with the large time scale module via a well-defined modular interface, and is called selective aggressiveness control (SAC). Two methods are distinguished: selective slope control and selective level control. In selective slope control—used in rate-based and window-based TCP congestion control—the slope of the linear increase phase is varied as a function of expected contention level during large time intervals (e.g., 2 second), increasing the slope if the contention level is predicted to be

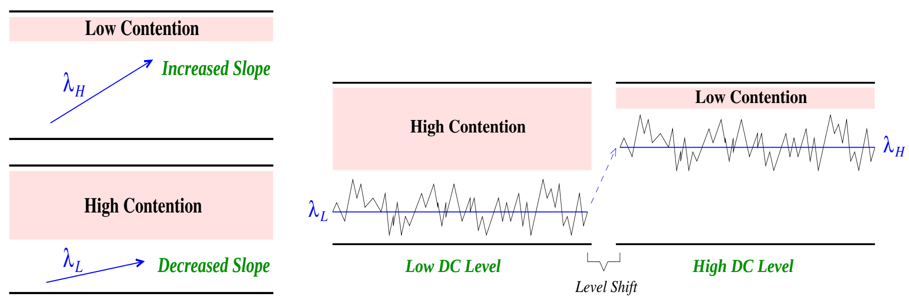

Fig. 1.4.1 Left: Selective slope control—i.e., slope shift—during linear increase phase for high- and low-contention periods. Right: Selective level control—i.e., "DC" level shift—between high- and low-contention periods.

low (thereby amplifying aggressiveness), and vice versa if the oppositive is true. In selective level control—used in adaptive redundancy control—a "DC" level, which is held constant during a large time scale interval, is shifted from high to low (and vice versa) across successive time intervals as a function of predicted contention level. This is depicted in Figure 1.4.1. Selective aggressiveness control is but one approach to engaging large time scale predictability structure for traffic control. There may be other approaches and mechanisms, equally or perhaps even more effective, which can be used to harness long-term predictability for traffic control. Their identification and evaluation is a subject for continued exploration. We note that there are four-fold challenges to be overcome in this endeavor: (a) correlation structure is dispersed over large distances in time, (b) information is probabilistic, (c) the mechanism should be implementable over existing traffic controls, and (d) it should not "harm" the underlying traffic control (if not help it) with respect to performance.

Multi-layered Feedback Control Multiple time scale traffic control, by coupling short time scale and large time scale control modules, leads to a multi-layered feedback control. The large time scale module dynamically estimates the optimal slope and level values to use for particular network contention levels which are then used to modulate the small time scale feedback control module. Even if the underlying small time scale feedback control (e.g., linear increase/exponential decrease rate-based control, TCP, or AFEC) is stable, this does not imply that the coupled system is stable. It is expected, however, that by exploiting time scale separation between the two control modules stability of the overall system can be achieved under fairly weak conditions. In particular, if T_L, T_S denote the time scales of the large and small time scale modules, respectively, $T_S \ll T_L$ (in a suitable sense), and the small time scale feedback control is assumed to converge fast (e.g., within a small factor of T_S by exponential convergence), then the overall system may be shown to be stable (but not asymptotically stable) by the quasi-stationarity argument that the small time scale control converges "well within" T_L—a locally stationary regime from its perspective—leading to a concatenation of locally well-behaved trajectories (short transient followed by convergence), assuming the large time scale control itself is stable. The most challenging problem arises when sufficient separation between small and large time scales does not hold. Complicated dynamics can ensue and identification of sufficiency conditions for stability with relevant counterparts in networking looms as a challenging problem.

Workload-sensitive Traffic Control Multiple time scale traffic control can be extended to workload-sensitive traffic control where traffic controls are made, to varying degrees, cognizant of workload properties in the broadest sense when this is deemed beneficial to do so. For example, when TCP is invoked by HTTP in the context of Web client/server interactions, the size of the file being transported—known at the server—may be conveyed or made accessible to protocols in the transport layer, including the selection of alternative protocols, for more effective data transport (see, e.g., [31] for a discussion of a specific mechanism in the context of congestion control). For short files, which constitute the bulk of connection requests in heavy-tailed file size distributions of Web servers, elaborate feedback control geared toward steady-state efficiency may be by-passed in favour of lightweight mechanisms in the spirit of optimistic control which can, in some instances [39], result in improved "effective bandwidth." Recently, the heavy-tailed characteristics of IP flow durations [23, 24] has been used to selectively perform routing table updates based on connection lifetime classification, where it is shown that desensitizing route updates triggered by short-lived flows can enhance routing stability [62]. Endowing workload-sensitivity to traffic control need not be restricted to congestion control, error control, and routing (for that matter, closed-loop control), and represents a broad area for future exploration facilitated by the recent discoveries and advancement in traffic characterization.

Optimal Prediction of Long-range Correlation Structure Predicting the future traffic level from past observations is an important component to affecting traffic control under self-similar traffic conditions. Chapters 17 and 18 provide heuristic approaches to estimating the future traffic level based on conditional expectation—optimal with respect to mean square error—but due to its nonlinearity effective optimal prediction remains a technical challenge [6]. With long-range dependent time series and their slow convergence properties, it becomes difficult to devise effective predictors that can be rigorously shown to have desirable properties. Traditional estimation theory achieves tractability through assumption of Markovian input (Kalman filters) or restriction on observation models (Wiener filters). On the positive side, the conditional expectation predictor $E[X^{(m)}(i) \mid X^{(m)}(i-1)]$ has proved effective in empirical evaluations with respect to yielding traffic level estimates that are close to the true traffic level. Furthermore, the performance gain obtained by engaging approximate information has been shown to approach the gain achievable when perfect information is available [67]. Thus there is room for further performance gain due to improved prediction, but its magnitude is expected to be incremental. On a related front, inference of network state through minimally intrusive actions is relevant for effectively incorporating the prediction mechanisms in network protocols. For example, in [68] a TCP connection's interaction with other traffic flows at bottleneck routers and the consequent impact on its output behavior—observable at the sender—is used to infer the contention level at bottleneck routers. That is, no separate probing mechanism is engaged to estimate network state. The effectiveness of this minimally intrusive scheme is shown to be dependent on the tracking ability of the underlying feedback control [68]; state estimation suffers most heavily during the linear increase phase after backoff which results in better tracking performance for TCP Vegas over TCP Reno. Inference can be performed by other means including arrival behavior of ACK packets, and further exploration of effective measurement schemes is of interest. When explicit probing is employed, the question arises as to how to perform accurate sampling and estimation of network state without significantly affecting Schrödinger's cat in the process. Accurate sampling is made difficult by the slow convergence properties of self-similar traffic, and trade-offs between accuracy and probing duration can benefit from further investigation (also relevant to measurement based admission control [36]).

1.4.2 Open-loop Control and Resource Provisioning

Resource Reservation and Admission Control The performance analysis techniques and issues discussed in Section 1.3 have direct bearing on the ability to compute performance bounds for routers fed with self-similar input, estimation of effective bandwidth and statistical multiplexing gain, bandwidth and buffer capacity dimensioning/provisioning, and admission control. Chapter 19 presents a specific open-loop architecture based on per-VC framing, and Chapter 16 discusses design and performance evaluation considerations of open-loop architectures under self-similar traffic conditions. The same observations regarding the need for finitary analysis, incorporation of second-order performance measures, relative impact of short-range and long-range correlation structure, and influence of scheduling advanced in Section 1.3 hold for the resource provisioning context. Similarly, the small buffer/large bandwidth resource provisioning strategy is expected to play a dominant role in facilitating guaranteed services—deterministic or statistical—in the context of open-loop control.

Dynamic Admission Control In resource provisioning, admission control is exercised in a static manner where the set of connections requesting service is known beforehand. In dynamic admission control, a connection i has an arrival time s_i and duration τ_i associated with it. The server has finite capacity C. Assuming Poisson call arrivals and exponential holding times, this defines a Markov Decision Process [59] which can be solved using standard techniques. In general, there may be multiple service classes and a reward (or utility) function r_i for servicing connection i with a particular service level. We can distinguish two ways by which self-similarity is introduced. One, the arrival process is allowed to be self-similar (see Chapter 15 for a discussion of TCP session arrivals), and two, the connection duration is allowed to be heavy-tailed. The latter is consistent with a canonical source model introduced by Likhanov *et al.* [42] (see also Chapter 8 and 1) where connection arrivals are Poisson but connection durations are heavy-tailed. The optimal decision rule, under heavy-tailed call durations, is a function of r_i which, in turn, may depend on τ_i. By exploiting the memoryless property of connection arrivals, it is expected that optimal decision procedures can be derived for a range of specific admission control models in the framework of Markov renewal processes. The most challenging problem arises when connection arrivals are self-similar. The correlated nature of arrivals renders computation of expected reward conditioned on current system state difficult. For non-Markovian arrival processes, it is possible to achieve tractability by adopting a worst-case approach—i.e., stochastic structure in the arrival process is ignored—based on competitive analysis [44, 58]. The goodness of an on-line algorithm is evaluated with respect to its performance vis-à-vis an optimal off-line algorithm where the ratio of their respective solutions—the competitive ratio—is evaluated over all input instances, i.e., sample paths. As far as evaluating the impact of different arrival processes, in particular, self-similar processes is concerned, competitive analysis yields limited insight due to its worst-case property.

Optimistic Traffic Control for Short-lived Connections Long-range dependence and its exploitation for traffic control, in particular, feedback traffic control is, by definition, best suited for flows or connections whose life time or connection duration is long lasting. Trivially, if the connection is too short, then reliable network state estimation is bound to fail or yield highly variable results. Empirical workload measurements have shown that connection durations and file sizes are heavy-tailed, which implies that the majority of connections are short-lived although a few connections contribute the bulk of traffic in terms of volume. Thus, in spite of the import of effectively managing long-lived connections due to their disproportionate contribution, if performance measures such as transmission completion time for short-lived connections are considered, then performance improvement may be achievable by utilizing workload information that is presently ignored. For example, if a server knows the size of the file to be transmitted, then conditioned on the file size, either an optimistic control in the spirit of open-loop control may be invoked for small files to reduce overhead with reactive actions undertaken only in the fraction of instances where needed (e.g., retransmissions for reliable transport), or full-fledged feedback congestion control with multiple time scale extension engaged when the requested file is large. If the file size distribution is known to be heavy-tailed but the size of each transfer instance not known, then optimistic control may be exercised by default until such time when the connection in question—by connection duration prediction—is concluded to be long-lived (see, e.g., [39] for a comparison of aggressive open-loop congestion control vs. linear increase/exponential decrease feedback congestion control). From a multiple time scale traffic control perspective, the fact that a connection is short-lived need not deter from utilizing a priori information about long-term network state if available. One possible extension is the use of network state information that is not tied to any particular connection but shared across many. The effectiveness of this scheme hinges on maintaining persistent, shared information at a host (e.g., Web server or client) where multiple connections are multiplexed in space or time. Since "network state" may depend on the location of each individual destination host and the route taken, a challenge lies in maintaining such information in a compact, easily accessible, and updated form. The conditional expectation predictor, for long-range dependent traffic, can be taken to say that the past observation should be used as the future prediction—i.e., $E[X^{(m)}(i) \,|\, X^{(m)}(i-1)] = X^{(m)}(i-1)$—which facilitates a slowly changing, invariant prediction table that can be indexed if the past observation $X^{(m)}(i-1)$ is known.

References

1. R. Addie, M. Zukerman, and T. Neame. Fractal traffic: measurements, modelling and performance evaluation. In *Proc. IEEE INFOCOM '95*, pages 977–984, 1995.
2. O. Ait-Hellal, E. Altman, A. Jean-Marie, and I. Kurkova. On loss probabilities in presence of redundant packets and several traffic sources. *Performance Evaluation*, 36:485–518, 1999.
3. M. F. Arlitt and C. L. Williamson. Web server workload characterization: The search for invariants. In *Proceedings of SIGMETRICS '96*, pages 126–137, May 1996.
4. Soren Asmussen. Rare events simulation for heavy-tailed distributions. In *Proc. RESIM '99 Workshop*, pages 139–157. University of Twente, Enschede, The Netherlands, March 11–12 1999.
5. Michael Barnsley. *Fractals Everywhere*. Academic Press, 1988.
6. Jan Beran. *Statistics for Long-Memory Processes*. Monographs on Statistics and Applied Probability. Chapman and Hall, New York, NY, 1994.
7. Ernst Biersack. Performance evaluation of forward error correction in ATM networks. In *Proc. ACM SIGCOMM '92*, pages 248–257, 1992.
8. J. Bolot, S. Fosse-Parisis, and D. Towsley. Adaptive FEC-based error control for Internet telephony. In *Proc. IEEE INFOCOM '99*, pages 1453–1460, 1999.
9. J. Bolot and A. Shankar. Analysis of a fluid approximation to flow control dynamics. In *Proc. IEEE INFOCOM '92*, pages 2398–2407, 1992.
10. S. Chen and K. Park. An architecture for noncooperative QoS provision in many-switch systems. In *Proc. IEEE INFOCOM '99*, pages 864–872, 1999.
11. J. Choe and N. Shroff. A central-limit-theorem based approach for analyzing queue behavior in ATM networks. *IEEE/ACM Trans. Networking*, 6(5):659–671, 1998.
12. I. Cidon, A. Khamisy, and M. Sidi. Analysis of packet loss processes in high-speed networks. *IEEE Trans. Information Theory*, 39(1):98–108, 1993.
13. I. Cidon, A. Khamisy, and M. Sidi. Dispersed messages in discrete-time queues: delay, jitter, and threshold crossing. In *Proc. IEEE INFOCOM '94*, pages 218–223, 1994.
14. I. Cidon, A. Khamisy, and M. Sidi. Delay, jitter and threshold crossing in ATM systems with dispersed messages. *Performance Evaluation*, 29(2):85–104, 1997.
15. M. Crovella and A. Bestavros. Self-similarity in world wide web traffic: Evidence and possible causes. *IEEE/ACM Trans. Networking*, 5:835–846, 1997.
16. N. G. Duffield, J. T. Lewis, N. O'Connel, R. Russell, and F. Toomey. Statistical issues raised by the bellcore data. In *Proc. 11th IEE Teletraffic Symposium*, 1994.
17. N. G. Duffield and N. O'Connel. Large deviations and overflow probabilities for the general single server queue, with applications. Technical Report DIAS-STP-93-30, DIAS Technical Report, 1993.

18. Nick Duffield. On the relevance of long-tailed durations for the statistical multiplexing of large aggregations. In *Proc. 25th Allerton Conference on Communication, Control and Computing*, pages 741–750, 1997.

19. ECMT. European Transport Trends and Infrastructural Needs. OECD Publications, 1995.

20. R. Edell and P. Varaiya. Providing internet access: What we learn from the INDEX trial. To appear in *IEEE Network*, 1999.

21. A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Trans. Networking*, 4:209–223, 1996.

22. M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology. In *Proc. ACM SIGCOMM '99*, 1999.

23. A. Feldmann, A. C. Gilbert, P. Huang, and W. Willinger. Dynamics of IP traffic: A study of the role of variability and the impact of control. In *Proc. ACM SIGCOMM '99*, 1999.

24. A. Feldmann, A. C. Gilbert, and W. Willinger. Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic. In *Proc. ACM SIGCOMM '98*, pages 42–55, 1998.

25. K. Fendick, M. Rodrigues, and A. Weiss. Analysis of a rate-based control strategy with delayed feedback. In *Proc. ACM SIGCOMM '92*, pages 136–148, 1992.

26. P. Francis, S. Jamin, V. Paxson, L. Zhang, D. Gryniewicz, and Y. Jin. An architecture for a global Internet host distance estimation service. In *Proc. IEEE INFOCOM '99*, pages 210–217, 1999.

27. M. Garret and W. Willinger. Analysis, modeling and generation of self-similar VBR video traffic. In *Proc. ACM SIGCOMM '94*, pages 269–280, 1994.

28. M. Grossglauser and J-C. Bolot. On the relevance of long-range dependence in network traffic. In *Proc. ACM SIGCOMM '96*, pages 15–24, 1996.

29. M. Harchol-Balter. Process lifetimes are not exponential, more like $1/t$: implications on dynamic load balancing. Technical report, EECS, University of California, Berkeley, 1996. CSD-94-826.

30. M. Harchol-Balter and A. Downey. Exploiting process lifetime distributions for dynamic load balancing. In *Proceedings of SIGMETRICS '96*, pages 13–24, 1996.

31. A. Heddaya and K. Park. Congestion control for asynchronous parallel computing on workstation networks. *Parallel Computing*, 23:1855–1875, 1997.

32. D. Heyman and T. Lakshman. What are the implications of long-range dependence for VBR-video traffic engineering? *IEEE/ACM Transactions on Networking*, 4(3):301–317, June 1996.

33. A. Hordijk and F. Spieksma. Constrained admission control to a queueing system. *Adv. Appl. Prob.*, 21:409–431, 1989.

34. P. Hu and J. Young. 1990 NPTS databook: national personal transportation survey. Federal Highway Administration, 1993.

35. Gordon Irlam. Unix file size survey - 1993. Available at `http://www.base.com/gordoni/ufs93.html`, September 1994.

36. S. Jamin, P. Danzig, S. Shenker, and L. Zhang. A measurement-based admission control algorithm for integrated services packet networks. In *Proc. ACM SIGCOMM '95*, pages 2–13, 1995.

37. S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, and L. Zhang. On the placement of Internet instrumentation. To appear in *Proc. IEEE INFOCOM '00*, 2000.

38. P. R. Jelenlovic and A. A. Lazar. The effect of multiple time scales and subexponentiality in mpeg video streams on queueing behavior. *IEEE J. Select. Areas Commun.*, 15(6):1052–1071, 1997.

39. Hyogon Kim. *A Non-Feedback Congestion Control Framework for High-Speed Data Networks*. PhD thesis, University of Pennsylvania, 1995.

40. R. Kitamura. Time-of-day characteristics of travel: An analysis of 1990 NPTS data. Federal Highway Administration, Chapter 4, 1995.
41. W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of ethernet traffic. In *Proc. ACM SIGCOMM '93*, pages 183–193, 1993.
42. N. Likhanov, B. Tsybakov, and N. Georganas. Analysis of an ATM buffer with self-similar ("fractal") input traffic. In *Proc. IEEE INFOCOM '95*, pages 985–992, 1995.
43. Z. Liu, P. Nain, D. Towsley, and Z. Zhang. Asymptotic behavior of a multiplexer fed by a long-range dependent process. To appear in *Journal of Applied Probability*, 1999.
44. M. Manasse, L. McGeoch, and D. Sleator. Competitive algorithms for online problems. In *Proc. ACM STOC '88*, pages 322–332, 1988.
45. M. Mathis, J. Semke, J. Mahdavi, and T. Ott. The macroscopic behavior of the TCP congestion avoidance algorithms. *Computer Communication Review*, 27, 1997.
46. Anthony McAuley. Reliable broadband communication using a burst erasure correcting code. In *Proc. ACM SIGCOMM '90*, pages 197–306, 1990.
47. A. Mukherjee and J. Strikwerda. Analysis of dynamic congestion control protocols - a Fokker-Planck approximation. In *Proc. ACM SIGCOMM '91*, pages 159–169, 1991.
48. B. Noble, G. Nguyen, M. Satyanarayanan, and R. Katz. Mobile network tracing. RFC 2041, 1996.
49. I. Norros. A storage model with self-similar input. *Queueing Systems*, 16:387–396, 1994.
50. J. Padhye, V. Firoiu, D. Towsley, and J. Kurose. Modeling TCP throughput: A simple model and its empirical validation. In *Proc. ACM SIGCOMM '98*, 1998.
51. K. Park, G. Kim, and M. Crovella. On the relationship between file sizes, transport protocols, and self-similar network traffic. In *Proc. IEEE International Conference on Network Protocols*, pages 171–180, 1996.
52. K. Park, G. Kim, and M. Crovella. On the effect of traffic self-similarity on network performance. In *Proc. SPIE International Conference on Performance and Control of Network Systems*, pages 296–310, 1997.
53. K. Park and W. Wang. AFEC: an adaptive forward error correction protocol for end-to-end transport of real-time traffic. In *Proc. IEEE IC3N*, pages 196–205, 1998.
54. K. Park and W. Wang. QoS-sensitive transport of real-time MPEG video using adaptive forward error correction. In *Proc. IEEE Multimedia Systems '99*, pages 426–432, 1999.
55. Kihong Park. Warp control: a dynamically stable congestion protocol and its analysis. In *Proc. ACM SIGCOMM '93*, pages 137–147, 1993.
56. M. Parulekar and A. Makowski. Tail probabilities for a multiplexer with self-similar traffic. In *Proc. IEEE INFOCOM '96*, pages 1452–1459, 1996.
57. V. Paxson and S. Floyd. Wide-area traffic: the failure of Poisson modeling. In *Proc. ACM SIGCOMM '94*, pages 257–268, 1994.
58. Serge Plotkin. Competitive routing of virtual circuits in ATM networks. *IEEE J. Select. Areas Commun.*, 13(6):1128–1136, 1995.
59. M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.
60. B. Ryu and A. Elwalid. The importance of long-range dependence of VBR video traffic in ATM traffic engineering: myths and realities. In *Proc. ACM SIGCOMM '96*, pages 3–14, 1996.
61. P. Shahabuddin. Rare event simulation in stochastic models. In *Proc. 1995 Winter Simulation Conference*, pages 178–185, 1995.

62. A. Shaikh, J. Rexford, and K. Shin. Load-sensitive routing of long-lived IP flows. In *Proc. ACM SIGCOMM '99*, 1999.

63. Scott Shenker. A theoretical analysis of feedback flow control. In *Proc. ACM SIGCOMM '90*, pages 156–165, 1990.

64. A. Shwartz and A. Weiss. *Large Deviations for Performance Analysis*. Chapman&Hall, 1995.

65. Shaler Stidham. Optimal control of admission to a queueing system. *IEEE Trans. Automat. Contr.*, AC-30(8):705–713, 1985.

66. B. Tsybakov and N. D. Georganas. Self-similar traffic and upper bounds to buffer overflow in an ATM queue. *Performance Evaluation*, 36(1):57–80, 1998.

67. T. Tuan and K. Park. Multiple time scale congestion control for self-similar network traffic. *Performance Evaluation*, 36:359–386, 1999.

68. T. Tuan and K. Park. Performance evaluation of multiple time scale TCP under self-similar traffic conditions. Technical Report CSD-TR-99-040, Department of Computer Sciences, Purdue University, 1999.

69. T. Tuan and K. Park. Multiple time scale redundancy control for QoS-sensitive transport of real-time traffic. To appear in *Proc. IEEE INFOCOM '00*, 2000.

70. W. Willinger, M. Taqqu, R. Sherman, and D. Wilson. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. In *Proc. ACM SIGCOMM '95*, pages 100–113, 1995.

71. C. Yang and A. Reddy. A taxonomy for congestion control algorithms in packet switching networks. *IEEE Network*, 9:34–45, July/August 1995.