Behavior of TCP-like elastic traffic at a buffered bottleneck router

Hans-Peter Schwefel

Abstract—**A major challenge in traffic modeling and performance analysis for the *Transmission Control Protocol* (TCP) stems from the fact that the incoming traffic is not independent of the congestion level in the network. This paper investigates a queueing model where the traffic essentially shows ON/OFF characteristics, i.e. the number of active TCP connections of finite (probabilistic) duration varies as described by a stochastic process. The essential behavior of TCP-like flow-control mechanisms is captured in the analytic model by the feature that the packet-rate of active connections can be throttled in order to avoid that the overall packet-stream exceeds the output-bandwidth of the bottleneck router. By appropriate adjustment of the connection duration, the number of packets in the connections remains unaffected. However, since TCP reacts to existing congestion, the throttling mechanism is only activated when the buffer-occupancy at the bottleneck router exceeds a certain threshold. The impact of such a flow-control mechanism on the characteristics of the incoming traffic as well as on the performance behavior at the bottleneck router is discussed and illustrated by numerical results of the analytic model. In particular, the use of (truncated) Power-Tail distributions for the ON periods leads to conclusions about the behavior of long-range dependent traffic under the influence of TCP's flow-control mechanism.**

Keywords— **TCP flow-control, ON/OFF models, Markov Modulated Poisson Processes, Long-Range Dependence, Truncated Power-Tail Distributions**

I. INTRODUCTION

ADEQUATE network design and capacity planning requires some knowledge (a model) of the traffic that is sent into the network. Traffic modeling for data networks is already difficult even if the incoming traffic is assumed to be independent of the events that happen in the network, see e.g. [1]. Nevertheless some progress has been made: For instance, the use of *Markov Modulated Poisson Processes* (MMPPs) allows to capture the inherently bursty property of network traffic while the resulting queueing models still remain tractable, i.e. performance parameters such as mean delay and loss probabilities can be computed. A special case of MMPPs are so-called ON/OFF models with exponential ON and OFF times: During an ON period packets are transmitted with Poisson Rate λ_p, while in the subsequent OFF period, no packet transmissions occur, see [2]. In order to account for the now widely acknowledged *Long-Range Dependent* (LRD) properties of data traffic (see e.g. [3]), such ON/OFF models can be extended by using non-exponential ON-time distributions with large or infinite variance, see [4], [5].

The results of the performance analysis of queueing models with such an arrival process can be used to predict the performance in different network design scenarios. In most of these queueing models, the arrival process is defined to be independent of the congestion level at the queue, i.e. the offered traffic remains unchanged if congestion occurs somewhere in the transmission path.

However, most of today's Internet traffic is transmitted using TCP (*Transmission Control Protocol*), whose built-in flow-control mechanism introduces a dependence between the network parameters (its congestion level) and the offered packet traffic, see [6]. As a consequence, large delays and high buffer-overflow probabilities in the network components themselves can be avoided, but at the cost of slowed down transmission rates at the source.

This work was supported by a fellowship from the German Academic Exchange Agency (DAAD).

This paper introduces a queueing model which captures the essence of the TCP flow-control mechanism while still remaining tractable. The performance results for this model and their practical implications are discussed. In particular, the use of Power-Tail like distributions for the ON periods leads to conclusions about the behavior of LRD traffic under the influence of the TCP flow-control mechanism.

Frequently, the behavior of TCP traffic is investigated in the scenario of TCP connections with infinite amount of data to transmit. For instance, [7] derives an estimate for the throughput for such an ever-lasting TCP connection, if its packets are subject to some loss probability p along their transmission path.

Reference [8] introduces and discusses a model that includes some kind of ON/OFF behavior of the users. Their model works on flow-level, i.e. individual packets are not considered. One of the consequences is that the throughput is insensitive to the actual distribution of the size of the connections; only their mean size matters. The model in this paper is based on the same idea, but it is packet-based, which allows to include a queueing model and a more sophisticated feedback mechanism. One of the implications is that the insensitivity towards the distribution of the connection size does not hold any more, if some buffer-space is available in the bottleneck router.

II. A MODEL FOR DYNAMIC TCP TRAFFIC

A. *Base Model: N-Burst*

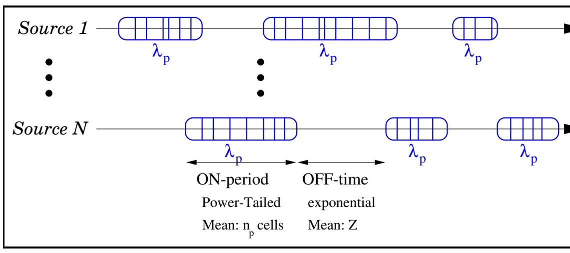

Fig. 1. The N-Burst Arrival Process: Packets from N ON/OFF sources are multiplexed together.

In this paper, the interest is particular in dynamic TCP connections, i.e. connections start and end as described by some stochastic process. When neglecting the flow control mechanism, those dynamics can be captured by multiplexing several ON/OFF packet streams, as shown in Fig. 1: each of the N sources emits packets at a Poisson-rate λ_p (peak-rate) during its

ON-time (a *burst*), and then transmits nothing during its OFF-time. In our scenario λ_p is the packet rate that is determined by the speed of the access line. This aggregated ON/OFF model is called N-Burst in correspondence to previous work [4], [5].

Let κ be the mean rate of the individual source (the average for the ON- and OFF-times together), then the N sources collectively generate packets at the mean rate $\lambda = N\kappa$. Let Z be the mean duration of the OFF periods and $\overline{n_p}$ the mean number of packets in an ON period. Then the mean duration of the ON period is $\overline{x_p} = \overline{n_p}/\lambda_p$ and

$$\kappa = \frac{\overline{n_p}}{\overline{x_p} + Z}.$$

If both ON times and OFF times are exponentially distributed, the model can be easily described in the MMPP framework, which goes back to work in [2]. The duration of the OFF period is expected to be less critical for the performance behavior. For simplicity they are assumed to be exponential with mean Z. In terms of the ON periods however, the analysis of many measurements of network traffic ([3], [9]) indicate that so-called Power-Tail distributions are a more adequate choice for the duration of the ON periods. The probabilities that long ON periods with duration longer than x are observed only drops off slowly with a Power-Law, $R(x) \sim x^{-\alpha}$. When working with truncated tails[1] – which are practically more meaningful in any case – a Phase-type representation of that class of distributions can be used, see [10] and Appendix B. That way, the packet stream of the multiplexed ON/OFF sources can still be described as an MMPP, yet with much more complicated structure. A description of the MMPP representation of the N-Burst model is given in Appendix B and C.

If such traffic is the input to a bottleneck router (here: exponential server) whose output capacity only allows for a packet-rate of $\nu < N\lambda_p$, it is shown in [4], [5] that the Power-Tail properties of the ON time distributions can cause poor performance behavior as measured by mean delay and overflow probabilities.

B. Modification 1: Shared Bottleneck Bandwidth

The assumption of ON/OFF traffic with constant Poisson rate λ_p during the ON periods is reasonable for many real-time applications and for protocols without flow-control mechanisms. However, TCP works differently. After a certain number of packets (the so-called *congestion window*) is sent out, the sender waits for acknowledgment packets from the receiver. The size of the congestion window is dynamically adjusted when congestion is detected. We omit the details here, since they will not show up in the model in any case. Also, they depend on the actual TCP implementation. The interested reader be referred to [11]. The important feature of the flow-control mechanism is that through the adjustment of the size of the congestion window, the effective sending rate of packets can be throttled from its maximum λ_p (which is determined by the speed of the access line) to a sufficiently low rate such that congestion is (hopefully) avoided.

The first modification of the N-Burst model is made along those lines of throttling the packet-rate of each individual source, if the bottleneck router in the transmission path cannot handle all the active connections any more. If we assume that the output bandwidth of the bottleneck router corresponds to a packet-rate ν, and i sources (connections) are active, the maximal sending rate λ_p is only used if $i\lambda_p < \nu$, i.e. when no overload situation at the bottleneck is created. If $i\lambda_p > \nu$, all sources equally throttle down their packet-rate by a factor

$$\beta_i = \frac{\nu}{i\lambda_p},$$

so that the i sources collectively generate packets at rate[2] ν. We call the resulting ON/OFF model the SHARED model. The approach for sharing the bottleneck capacity between the active connections was also used in [8] for a TCP model on flow-level. However, modeling on packet-level is necessary to do queueing analysis and also to implement more realistic throttling approaches as in II.C.

Since the actual number of packets in the connection must not be changed, a throttling of the packet-rate also requires a change in the duration of the connection: In the unthrottled N-Burst model with exponential ON time distribution, the rate of a transition to the $(i-1)$ active state is $i/\overline{x_p}$. This rate is changed in the SHARED model to $\beta_i\, i/\overline{x_p}$, i.e. the state holding time is extended. See Appendix C for the general case of Matrix-Exponential ON times.

The rates for transitions that correspond to starting connections remain unchanged however, since idle sources are not affected by the throttling. The modification of the MMPP to obtain the SHARED model from the N-Burst model are far from trivial. It has to be investigated whether the distribution of the number of packets per connection is not affected by that modification. While an individual source is active, the throttling that is caused by other active sources only slows down the scaling of the time (the *local clock*) for the individual throttled sources. The state transitions in the MMPP representation of the single source remain in the same order. Numerical computations of the distribution of the number of packets in the SHARED model also show that that distribution is not affected by the throttling.

Note that the extension of the ON periods is not made up for by a reduction of the length of the subsequent OFF period of the same source. Therefore, the throttling not only decreases the observed packet rates during the connection, but also the long-term average packet rate of the individual source.

C. Modification 2: React to existing congestion

In contrast to the assumptions in the SHARED model, the individual real TCP source does not have the knowledge about any other, newly starting TCP connections, but it only reacts to existing congestion situations. In that sense, the control mechanism of the SHARED model is too good, since it adjusts the sending rates of the sources instantaneously when new connections become active.

A second modification of the traffic model accounts for that behavior: As long as no congestion is present, the sources generate packets during ON-periods with Poisson rate λ_p as in the

[1] i.e., Power-Law drop-off of the complementary distribution function is only observed for some finite, but potentially large range of x

[2] Slightly smaller β_i can be used to capture the fact that the regulation is never optimal. However, here we use the approach that the sources share exactly the bandwidth ν.

basic N-Burst model. Whenever the buffer-occupancy at the bottleneck router reaches B_1 or more packets, TCP is assumed to recognize the congestion situation and the arrival process switches to the SHARED model. As an approximation, the buffer itself is assumed to be infinite, so that we do not have to worry about retransmissions of lost packets (they are stored in the 'backup' buffer beyond level B_1). This model will be called TCP$_{B_1}$. It is easy to see that the N-Burst/M/1 queue is the limit $B_1 \to \infty$, where the throttling of packet-rates is never performed. At the other end, SHARED is the other limiting model for $B_1 = 0$.

Note that a single source of the TCP$_{B_1}$ model can generate bursty traffic even within the same ON period: the source starts transmitting packets with maximum Poisson rate λ_p, but later on it typically becomes throttled when the buffer-occupancy reaches the threshold B_1.

The computational methods to solve for the steady-state solution of the TCP$_{B_1}$ queueing model are described in Appendix D. They follow in spirit the model in [12].

The remainder of this paper discusses the impact of the throttling on the Packet-stream of a single source as well as the performance behavior of the TCP$_{B_1}$ model.

III. Impact on incoming traffic

In this paper the numerical examples look at the scenario that N fast (10Mb/s) LANs are connected via the bottleneck router to a slow (1Mb/s) access line. Each LAN is assumed to be used by only one ON/OFF source. If we assume an average packet-size of 1kB and average connection sizes of 50kB, the parameters for the TCP$_{B_1}$ model follow as: $\overline{n_p} = 50$ packets, $\lambda_p = 1250$ packets/s, $\nu = 125$ packets/s, $\overline{x_p} = 40$ ms in the unthrottled N-Burst. Furthermore, we assume that the exponentially distributed OFF periods have mean $Z = 5$ s. Consequently, the average packet rate in the unthrottled N-Burst comes out to be $\kappa = 9.92$ packets/s and somewhat lower in the TCP$_{B_1}$ model when $B_1 < \infty$, see next section. Whenever the truncated Power-Tail distributions of Appendix B are used for the ON-periods, the tail-exponent $\alpha = 1.4$ is used. Note that since $\lambda_p > \nu$, even a single TCP connection by itself gets throttled as soon as the buffer-occupancy reaches B_1 packets.

Before we investigate the performance of the TCP$_{B_1}$ model, we look at some properties of the traffic from a single source in the SHARED and N-Burst models which represent the limits of the TCP$_{B_1}$ model for $B_1 = 0$ respectively $B_1 = \infty$.

Connection Duration: As already mentioned in the previous section, the distribution of the number of packets per connection is identical in the N-Burst and in the SHARED model. However, since the packet-rate is at times (in our scenario always) reduced in the SHARED model, the duration of the connection is prolonged. The distribution of the connection duration in the SHARED model is a complicated Matrix-Exponential distribution even when the original ON-time distribution in the N-Burst model is exponential. Figure 2 plots the numerical values of the tail-probabilities of the connection duration in the N-Burst model and the SHARED model. The throttling does not affect the shape of the tails, exponential remains exponential, and so do Power-Tails. For the given parameters and $N = 4$ ON/OFF sources, the throttling in the SHARED model reduces the utilization of the bottleneck router from 31.75% in the N-Burst model to 29.08%.

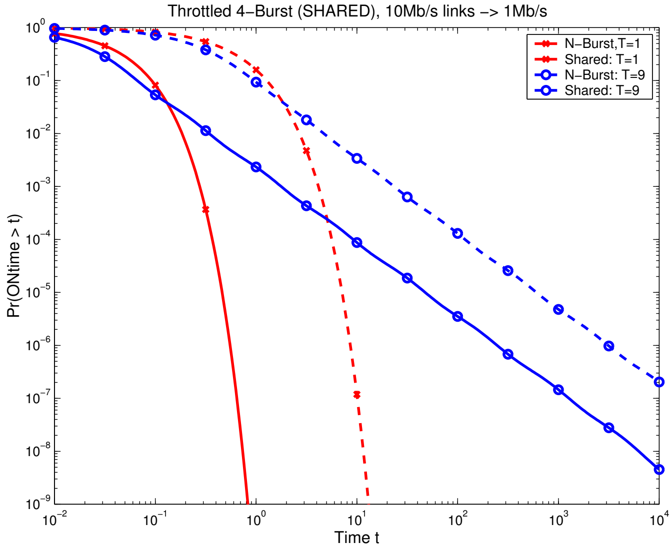

Fig. 2. The complementary distribution function of the duration of the ON periods in the N-Burst and SHARED model: Power-Tail behavior (appearing as straight line) is observed for both models when $T = 9$ (i.e. when more phases are used in the truncated Power-Tail distribution, see Appendix B).

Autocorrelation: Since LRD properties[3] of network traffic are currently widely discussed, it is interesting to investigate whether such correlation structure could be a result of the throttling of the individual connections. Therefore, we look at the coefficient of autocorrelation for both, inter-packet times and number of packets in an interval of size $\Delta = 1$ s, for the traffic of a single ON/OFF source, which is subject to various degrees of throttling due to other TCP connections in the SHARED model. Figure 3 shows the numerical values of the coefficient of correlation on log-log scale: A single ON/OFF source with exponential ($T = 1$) ON and OFF periods generates inter-packets times that are in fact a renewal process. Thus the correlation is zero in the N-Burst model with $T = 1$. The throttling in the SHARED model introduces some correlation (dashed line with crosses) yet the values of $r(k)$ are so small that they would not be noticeable in real measurements (due to the unavoidable noise). However, this is different in the counting process (dashed-dotted curve for N-Burst and dotted curve for SHARED) in the case $T = 1$. There the value of $r(k)$ is raised substantially, but still no LRD properties can be observed.

If truncated Power-Tail distributions are used for the ON periods, the correlation of the inter-packet times shows the LRD properties already for the N-Burst model (solid line). The throttling in the SHARED model increases those values, but only marginally so that the curves cannot be distinguished in Figure 3. In the counting process, the effects of throttling can be seen much more clearly.

IV. Performance Results

When talking about performance of the TCP$_{B_1}$ model, it is important to keep in mind that now the delay is not only caused

[3] i.e., the autocorrelation of counts or inter-packet times drops off very slowly as $r(k) \sim k^{1-\alpha}$, where $1 < \alpha \le 2$.

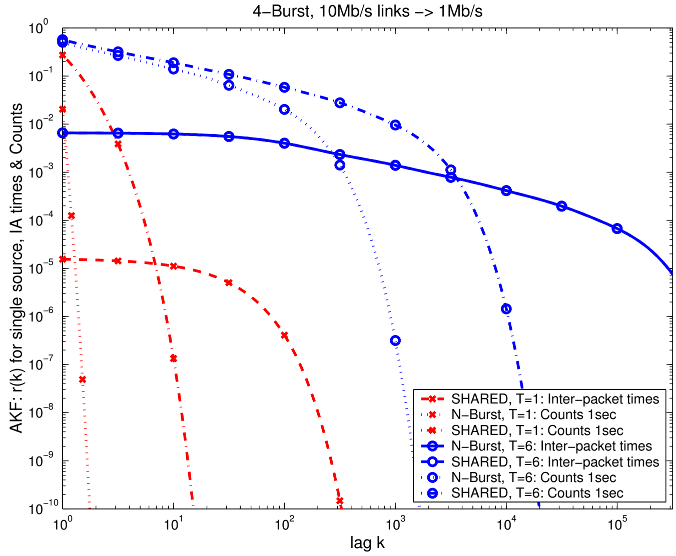

Fig. 3. The autocorrelation function of the inter-packet times and the counting process: The correlation of the inter-packet times of the SHARED model is only marginally larger than in the N-Burst model when $T = 6$ (circles). Therefore the two curves are indistinguishable. For $T = 1$ (crosses), an exponential ON/OFF source in the N-Burst model shows no correlation in its inter-packet times, that curve does not appear on the logarithmic scale.

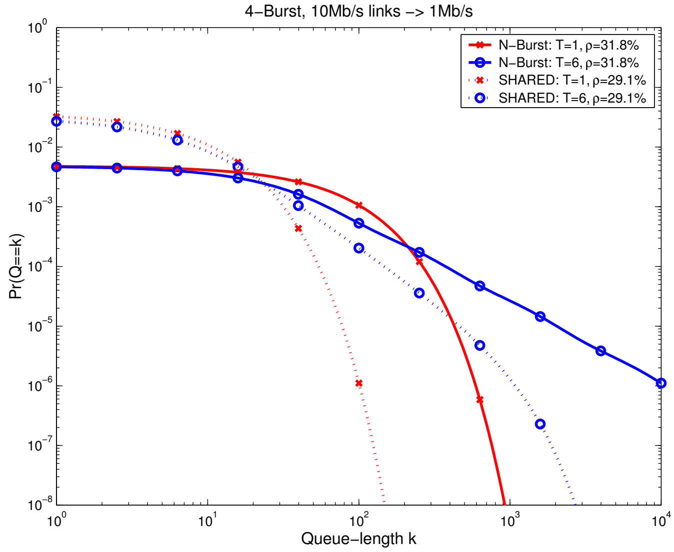

Fig. 4. The queue-length distribution of the N-Burst and SHARED model: The throttling in the SHARED model (dotted lines) reduces the probabilities of very long queues. This is true for exponential connection sizes (crosses) as well as for truncated Power-Tails (circles).

at the bottleneck router (Sect. IV-A), but packets can be delayed already at the source (Sect. IV-B).

A. *Buffer-Occupancy*

It is known ([13], [5]) that the queue-length probabilities for an unthrottled N-Burst model with Power-Tailed ON periods also show a Power-Tail, whose exponent depends on the original tail-exponent α and the number of sources that are sufficient to over-saturate the router when they are simultaneously in a long ON period. In our scenario, $\lambda_p > \nu$, a single source already creates an over-saturation period for the router, so the queue-length distribution decays with a Power-Law with exponent α, see Fig. 4. If we now turn to the SHARED model, then the arrival rate never exceeds the service-rate due to the throttling. Instead those two rates are temporarily identical, so for a limited time interval Δ, the model behaves like an M/M/1 queue with $\rho = 1$. From results of the transient analysis of M/M/1 queues (see e.g. [14]), it is known that the queue-length of an M/M/1 queue with utilization $\rho = 1$ grows to values of approximately $\sqrt{2\nu\Delta}$ in time Δ. In the original N-Burst model, the queue grows to $(\Lambda - \nu)\Delta$ during the over-saturation period of length Δ with average arrival rate Λ. Therefore, the throttling leads to much slower queue-growth [4], as demonstrated by the dotted lines in Fig. 4. This corresponds to the observations in [6] that overflow-probabilities in simulation experiments of TCP traffic are not nearly as dramatic for large buffers, as conventional ON/OFF models without flow-control predict.

Let us now turn to the more realistic TCP$_{B_1}$ model, which takes into account the buffer-space at the router. Figure 5 illustrates, that the queue-length distributions of the two limiting models, N-Burst and SHARED, provide an excellent description of the behavior of the more complicated TCP$_{B_1}$ model: For buffer-occupancies below the threshold B_1, the unthrottled N-Burst arrival process is active and the queue-length probabilities in the TCP$_{B_1}$ model follow the ones of the limiting N-Burst/M/1 model. The analog observation holds for buffer-occupancies above B_1, when the SHARED model takes over. The probability-mass that is taken away from the N-Burst curve for large buffers is now mainly concentrated in a peak around the buffer threshold B_1. Thus, Fig. 5 illustrates nicely the effectiveness of the flow-control mechanism.

[4]The model would also allow to throttle the sources more strongly, in which case the growth of the queue-length would be limited further.

B. *Average Packet-Rate*

The results in the previous section showed, that the flow-control mechanism prevents the built-up of huge queues. After all this is not too surprising, since that is its goal in the first place. However, there is of course a price to be payed for the improvement of the performance at the bottleneck router: The packet-rates during the connections are reduced, so packets are held back at the source instead of at the router. This additional delay must not be neglected in a fair discussion of the effectiveness of TCP flow control.

The steady-state solution for the TCP$_{B_1}$ queueing model in Appendix D allows to determine an average packet-rate $\mu_p(B_1)$ that each connection achieves. Obviously, in the limit $B_1 \to \infty$, the N-Burst model is obtained, so $\mu_p(B_1)$ converges monotonically from below to λ_p. The numerical computation in Fig. 6 show clearly that at the other end $B_1 = 0$, the average packet-rate in the connection is independent of the actual type of burst-length distribution. That is in principle the insensitivity result that was already pointed out in [8] for the flow-level model.

However, Fig. 6 also shows that the insensitivity is abolished, as soon as buffer-space $B_1 > 0$ is available. Exponential connection sizes (uppermost curve in Fig. 6) are better off in those scenarios, because most of the time, the buffer can absorb a large part of the connection without throttling the source. In case of

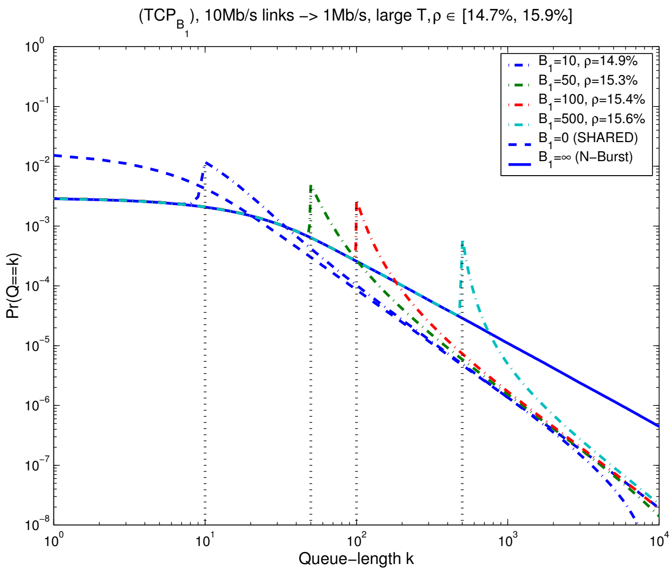

Fig. 5. The queue-length distribution of the TCP$_{B_1}$ model for $N = 2$ sources and long Power-Tails: Below the peak at $k = B_1$, the distributions is close to the N-Burst model. After the peak, the queue-length of the TCP$_{B_1}$ model approaches the distribution of the SHARED model. The model here uses TPT distributed connection sizes with a large number of phases, such that the impact of the truncation is not visible in the plotted range of queue-lengths.

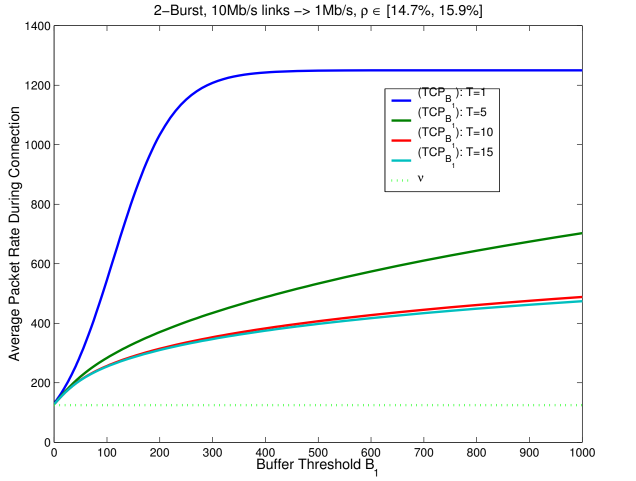

Fig. 6. The average packet-rate during connections of the TCP$_{B_1}$ model for $N = 2$ sources: An increased buffer-size B_1 also leads to less throttling. However, the gain is less pronounced for Power-Tailed connection sizes (with large T). The uppermost solid curve corresponds to $T = 1$, while the lowest solid line corresponds to the highest $T = 15$ (the same order as indicated in the legend).

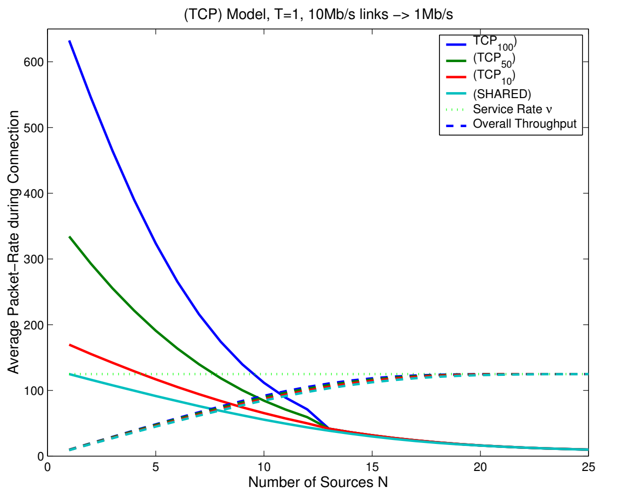

Fig. 7. The decay of the average packet-rate during connections due to throttling for an increasing number of sources N: For $N > 12$ sources, the router is almost saturated, so that buffer-size B_1 is almost always exceeded and throttling is hardly ever turned off. The uppermost solid curve corresponds to $B_1 = 100$, while the lowest solid curve corresponds to the SHARED model where $B_1 = 0$ (same order of the solid curves as in the legend). All models use exponential connection sizes ($T = 1$).

Power-tails on the other hand, the occasional huge connections increase the probability of throttling taking place.

Finally, we investigate the impact of adding additional sources (LANs) at the access of the router in Fig. 7. With increasing number of sources N, the overall utilization of the TCP$_{B_1}$ model increases, shown by the ratio of the throughput (dashed lines) to the constant service rate ν (dotted line). Note that due to the throttling, the TCP$_{B_1}$ queueing model always remains stable ($\rho < 1$), yet the utilization is very close to 1 for more than 15 sources. The average-packet rate $\mu_p(B_1)$ during the connections benefits most from larger buffer-space, when the utilization is low (small N). For very high utilization ($N \geq 13$), the buffer is almost always filled up to level B_1 regardless of the actual value of B_1. Therefore, the connections are almost always throttled. The curves in Fig. 7 show the scenario of exponential connection sizes. If we use Power-Tails instead, the curve for the SHARED model is not affected at all (insensitivity!), while the benefit of $\mu_p(B_1)$ for small utilization values becomes less pronounced.

In conclusion: For bottlenecks with high overall utilization, buffer-space and also the actual distribution of the connection sizes has little impact. But in on average lowly utilized routers (e.g. $\rho < 0.5$), buffer-space can lead to substantial improvement of the packet-rates during the connections. However, then there is a strong impact of the actual distribution of the connection sizes.

V. Conclusions and Future Directions

When modeling performance for TCP traffic it is important to capture the feedback between the network and the offered input traffic. Also, the dynamics of newly starting or ending TCP connections should not be neglected.

The model that is introduced in this paper presents an extension of the existing ON/OFF models in [4], [5]. The extension captures the essential features of TCP flow control. These are:

- The sending rate of packets is adjusted in order to avoid congestion. Optimally, all active connections share the output bandwidth of a bottleneck router. This is implemented in the SHARED model.
- Rather than keeping track of the bandwidth requirements of the active connections, TCP only reacts to an existing congestion situation. Therefore, its flow-control mechanism does not come into play, before actual congestion has occurred. Model TCP$_{B_1}$ takes this feature into account by conditioning the throttling of the individual packet rates not only on the number of

active sources but also on buffer-occupancy of the bottleneck switch.
The developed TCP_{B_1} model still remains tractable, so it provides exact numerical results for several performance parameters. Here, we investigate the buffer-occupancy at the bottleneck router and the average packet-rate that is obtained by the individual connections. Note that the use of truncated Power-Tail distributions (see Appendix B) for the connection size[5] in the TCP_{B_1} model allows to mimic LRD properties in the traffic.

The results from the analysis of the model are the following:
• The autocorrelation of inter-packet times is increased by the throttling, yet no LRD like behavior could be observed that is caused purely by the throttling mechanism. Furthermore, the dominating part of the autocorrelation function of the inter-packet times is contributed by non-exponential ON period distributions rather than by the shaping due to the flow control. The autocorrelation function of the counting process is more strongly affected.
• The throttling prevents the build-up of large queues at the router. However, the prize is that the packets are delayed already at the sources. The numerical results for the average packet rate during a TCP connection illustrate that the delay shifts from the router to the source.
• In low utilized routers, additional buffer-space can dramatically improve the packet-rates during connections, and thus it reduces the time that is necessary to send out all the packets in a connection. The improvement is most pronounced if the distribution of the connection size is well-behaved (exponential). If Power-Tail distributions are involved, the improvements are still observed but only to a smaller extent.

The results from this paper provide quite a few insights about the behavior of such regulated traffic in networks. However, it also raises a number of questions which have to be considered in future work
• Theoretical Issues: It is mentioned in Sect. II-B that the distribution of the number of packets in a connection is not affected by the way, the throttling is implemented. This is intuitively reasonable and was validated numerically for several different parameter sets. However, a mathematical proof is still missing.
• Queueing Delay: The investigation of average packet rates during connections can be used to derive conclusion about the time, it takes to send out the packets of a connection. However, when $B_1 > 0$, the last packet of the connection can still be stuck in the buffer. That particular queueing delay of the last packet of a connection is of interest for more detailed performance evaluation.
• Model modification: The TCP_{B_1} model is still only an approximation of the true TCP behavior in the following features:
– Delayed Reaction to Congestion: TCP needs at least a round-trip time or a timeout interval to recognize a packet loss and react to it. In the model, a buffer-occupancy of at least B_1 results in instantaneous action (replacement of the N-Burst arrival process by the SHARED process).
– Slow-Start: Current TCP implementations start off with a congestion window of size 1 which is only gradually increased. As a consequence, short connections only achieve a packet-rate which is lower than λ_p even when no congestion arises during the connection. That feature is currently not implemented in the model.
The question is of course, in which scenarios do the current model simplifications matter? A feeling for what the answer probably looks like can be derived by simulation runs. However, the experiments have to be evaluated carefully, since recent investigations in [15] have shown that generalizations from TCP simulation results can be questionable.
• Applicability of steady-state analysis: It is known that steady-state analysis for LRD traffic can be misleading, since the performance parameters rarely reflect the large fluctuations that will be observed in practice for such traffic. Transient analysis as performed for the N-Burst model in [16] provides a better description of such behavior. In this paper, we only present steady-state analysis. It has to be investigated, whether transient analysis provides additional insight in the flow-control mechanisms.

[5] Strictly speaking, the TPT distributions of Appendix B are used for the duration of the ON-time in the unthrottled N-Burst model. However, it can be shown that the distribution of the number of packets per ON-time shows also a truncated Power-Tail.

Appendix

A. Matrix Exponential (ME) Distributions

As defined in [14], the vector-matrix pair $\langle \mathbf{p}, \mathbf{B} \rangle$ represents a ME distribution with complementary distribution function $R(x)$, and density function $f(x)$ in the following way:

$$R(x) = \mathbf{p} \exp(-x\mathbf{B})\boldsymbol{\epsilon}',$$
$$f(x) = -\frac{R(x)}{dx} = \mathbf{pB} \exp(-x\mathbf{B})\boldsymbol{\epsilon}',$$

where $\boldsymbol{\epsilon}'$ is a column-vector with all components being 1.

The concept follows from *Phase-Type* distributions (see [17]), but it is broader in the sense that the elements of $\mathbf{p}$ and $\mathbf{B}$ do not need to have a phase interpretation: they could be negative or complex, as long as the resulting $R(x)$ is a well-defined reliability function ($R(x) \geq 0$, monotonic, $R(0) = 1$, $R(\infty) = 0$). Despite the fact that the elements of $\mathbf{B}$ can be very different from common transition rates, $\mathbf{B}$ is called the *rate matrix*.

The moments of the distribution come out by integration:

$$\mathrm{E}\left(X^k\right) = k!\,\mathbf{p}\,\mathbf{B}^{-k}\boldsymbol{\epsilon}' \tag{1}$$

B. Truncated Power-Tail (TPT) Distributions

It is shown in [10] that a family $\langle \mathbf{p}_T, \mathbf{B}_T \rangle$ of ME distributions with increasing number T of phases can be constructed, such that the individual distributions show Power-Law behavior for some range before they drop off exponentially. By using more phases, the location (so-called PT Range) of the drop-off can be systematically controlled. See [10] and [5] for more details.

Let

$$0 < \theta < 1, \quad \text{and} \quad \gamma := \left(\frac{1}{\theta}\right)^{1/\alpha} > 1 \cdot$$

Then, let $< \mathbf{p_T}, \mathbf{B_T} >$ be the ME representation of a T-phase Hyperexponential distribution with

$$\mathbf{p}_T = \frac{1-\theta}{1-\theta^T}[1,\, \theta,\, \theta^2,\, \cdots .\, \theta^{T-1}], \quad \text{and}$$

$$\mathbf{B}_T = \mu\ \mathbf{diag}\left(1, \gamma^{-1}, ..., \gamma^{-(T-1)}\right).$$

The parameter, μ, is a positive constant that can be chosen to set the mean of the distribution by solving (2) for μ.

The expected value follows directly from (1)

$$\mathrm{E}(X_T) = \mathbf{p_T}\,\mathbf{B_T^{-1}}\,\epsilon'_\mathbf{T} = \frac{1}{\mu}\frac{1-\theta}{1-\theta^T}\frac{1-(\gamma\theta)^T}{1-\gamma\theta} \qquad (2)$$

C. MMPP representation of the TCP model

Markov modulated Poisson Processes are described by a generator matrix $\mathcal{Q}$ of the modulating process, and a diagonal matrix $\mathcal{L}$ that contains the Poisson packet rates for each state. A single source ($N = 1$) ON/OFF process with exponential OFF periods with mean Z and general Matrix-Exponential ON periods with representation $< \mathbf{p}, \mathbf{B} >$ has the representation:

$$\mathcal{Q}_1 = \left[\begin{array}{c|c} -1/Z & 1/Z\,\mathbf{p} \\ \hline -\mathbf{B}\varepsilon' & -\mathbf{B} \end{array}\right], \quad \mathcal{L}_1 = \left[\begin{array}{c|c} 0 & \\ \hline & \lambda_p\mathbf{I} \end{array}\right].$$

The aggregation of the traffic from N identical ON/OFF sources could be represented using N Kronecker sums of $\mathcal{Q}_1$ respectively $\mathcal{L}_1$. However, for our purpose it is easier to express $\mathcal{Q}_N$ in a Quasi-Birth-Death structure where the levels are defined by the number of active sources. Here, a general notation which includes the N-Burst and the SHARED model is given:

$$\mathcal{Q}_N = \left[\begin{array}{c|c|c|c|c} \mathbf{Z_0} & \mathbf{X_0} & & & \\ \hline \mathbf{Y_1} & \mathbf{Z_1} & \mathbf{X_1} & & \\ \hline & \ddots & \ddots & \ddots & \\ \hline & & \mathbf{Y_{N-1}} & \mathbf{Z_{N-1}} & \mathbf{X_{N-1}} \\ \hline & & & \mathbf{Y_N} & \mathbf{Z_N} \end{array}\right],$$

$$\begin{aligned} \text{where:}\quad \mathbf{X_i} &= \tfrac{N-i}{Z}\,\mathbf{I}^{\otimes i}\otimes\mathbf{p}, & i &= 0, ..., N-1, \\ \mathbf{Y_i} &= -\beta_i\,(\mathbf{B}\varepsilon')^{\oplus i}, & i &= 1, ..., N, \\ \mathbf{Z_i} &= -\beta_i\mathbf{B}^{\oplus i} - \tfrac{N-i}{Z}, & i &= 0, ..., N. \end{aligned}$$

The notation $\mathbf{A}^{\otimes i}$ expresses the Kronecker product with i factors $\mathbf{A}\otimes ... \otimes\mathbf{A}$. Analogously, $\mathbf{A}^{\oplus i}$ stands for the Kronecker sum with i summands, where $\mathbf{A}\oplus\mathbf{A} = \mathbf{A}\otimes\mathbf{I} + \mathbf{I}\otimes\mathbf{A}$.

Note that the main diagonal blocks $\mathbf{Z_i}$ of $\mathcal{Q}_N$ are quadratic but with growing dimension T^i for i active sources. Consequently, the other two diagonals contain non-quadratic matrices $\mathbf{X_i}$ and $\mathbf{Y_i}$.

$$\mathcal{L}_N = \left[\begin{array}{c|c|c|c|c} 0 & & & & \\ \hline & \beta_1\,\lambda_p\mathbf{I} & & & \\ \hline & & \beta_2\,2\lambda_p\mathbf{I}^{\otimes 2} & & \\ \hline & & & \ddots & \\ \hline & & & & \beta_N\,N\lambda_p\mathbf{I}^{\otimes N} \end{array}\right]$$

For the numerical computations the state-space size is even further reduced, see [18].

If all $\beta_i \equiv 1$, the N-Burst model without throttling is obtained. For the SHARED model:

$$\beta_i = \min\left\{\nu/(i\lambda_p),\, 1\right\}, \qquad (3)$$

i.e. no throttling is performed as long as $i\lambda_p < \nu$. Otherwise, the Poisson packet-rates in $\mathcal{L}_N$ as well as the burst-ends (described by $\mathbf{Y_i}$) and the internal state-changes of the ON-time distributions (in $\mathbf{Z_i}$) are slowed down by the factor β_i.

D. Mixed Matrix-geometric Solution of TCP Queueing Model

The infinite generator matrix of the TCP$_{B_1}$ queueing model has the following Quasi-Birth-Death Structure:

$$\hat{\mathbf{Q}} = \left[\begin{array}{ccccccc} \overline{\mathbf{A}_1} & \mathbf{A}_0 & & & & & \\ \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 & & \\ \hline & & & \mathbf{C}_2 & \mathbf{C}_1 & \mathbf{C}_0 & \\ & & & & \ddots & \ddots & \ddots \end{array}\right],$$

where $\overline{\mathbf{A}_1} = \mathcal{Q}_N - \mathcal{L}_N$, $\mathbf{A}_0 = \mathcal{L}_N$, $\mathbf{A}_1 = \mathcal{Q}_N - \mathcal{L}_N - \nu\mathcal{I}$ refer to the unthrottled N-Burst base model with $\beta_i = 1$. From level B_1 on those matrices are replaced by the shared-bandwidth model SHARED: $\mathbf{C}_0 = \tilde{\mathcal{L}}_N$, $\mathbf{C}_1 = \tilde{\mathcal{Q}}_N - \tilde{\mathcal{L}}_N - \nu\mathcal{I}$, with the throttling factors $\tilde{\beta}_i$ as in 3. The service process is always exponential with rate ν: $\mathbf{A}_2 = \mathbf{C}_2 = \nu\mathcal{I}$.

It can be shown (see [12] for the necessary proofs in a comparable scenario) that the block-partitioned steady-state probability distribution of such a process, $\boldsymbol{\pi} = [\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, ...]$ with $\boldsymbol{\pi}\hat{\mathbf{Q}} = \mathbf{0}$, can be expressed in the following mixed matrix-geometric form

$$\begin{aligned} \boldsymbol{\pi}_i &= \mathbf{a}\,\mathbf{R}^i + \mathbf{b}\,\mathbf{S}^{B_1-1-i}, & i &= 0, ..., B_1 - 1 \\ \boldsymbol{\pi}_k &= \boldsymbol{\pi}_{B_1}\,\mathbf{T}^{k-B_1}, & k &= B_1, B_1+1, ... \end{aligned}$$

where the matrix factors $\mathbf{R}, \mathbf{S}, \mathbf{T}$ are the minimal solutions of the following quadratic matrix equations:

$$\mathbf{A}_0 + \mathbf{R}\mathbf{A}_1 + \mathbf{R}^2\mathbf{A}_2 = 0, \quad \mathbf{A}_2 + \mathbf{S}\mathbf{A}_1 + \mathbf{S}^2\mathbf{A}_0 = 0,$$

$$\mathbf{C}_0 + \mathbf{T}\mathbf{C}_1 + \mathbf{T}^2\mathbf{C}_2 = 0.$$

The vectors $\mathbf{a}$, $\mathbf{b}$, and $\boldsymbol{\pi}_{B_1}$ follow from the boundary equations at level 0, $B_1 - 1$, B_1 and from normalization $\boldsymbol{\pi}\varepsilon' = 1$ as the solution of the following system of linear equations:

$$[\mathbf{a},\ \mathbf{b}]\left[\begin{array}{c|c|c} \mathbf{L_a} & \mathbf{R_a} & \mathbf{d_1} \\ \hline \mathbf{L_b} & \mathbf{R_b} & \mathbf{d_2} \end{array}\right] = [\mathbf{0},\ \mathbf{0},\ 1],$$

where

$$\begin{aligned} \mathbf{L_a} &= \overline{\mathbf{A}_1} + \mathbf{R}\,\mathbf{A}_2, \\ \mathbf{R_a} &= \mathbf{R}^{B_1-2}\left(\mathbf{R}\mathbf{A}_0 - (\mathbf{A}_0 + \mathbf{R}\mathbf{A}_1)\,\mathbf{C}_2^{-1}\cdot\right. \\ &\qquad \left.\cdot(\mathbf{C}_1 + \mathbf{T}\mathbf{C}_2)\right), \\ \mathbf{L_b} &= \mathbf{S}^{B_1-2}\left(\mathbf{S}\overline{\mathbf{A}_1} + \mathbf{A}_2\right), \end{aligned}$$

$$\mathbf{R_b} = \mathbf{A}_0 - (\mathbf{SA}_0 + \mathbf{A}_1)\ \mathbf{C}_2^{-1}\ (\mathbf{C}_1 + \mathbf{TC}_2)\,,$$

$$\mathbf{d_1} = \sum_{k=0}^{B_1-1} \mathbf{R}^k \varepsilon' - $$

$$-\mathbf{R}^{B_1-2}\,(\mathbf{A}_0 + \mathbf{RA}_1)\,\mathbf{C}_2^{-1}\ \left(\sum_{k=0}^{\infty} \mathbf{T}^k \varepsilon'\right),$$

$$\mathbf{d_2} = \sum_{k=0}^{B_1-1} \mathbf{S}^k \varepsilon' - (\mathbf{SA}_0 + \mathbf{A}_1)\ \mathbf{C}_2^{-1}\ \left(\sum_{k=0}^{\infty} \mathbf{T}^k \varepsilon'\right).$$

Finally,

$$\boldsymbol{\pi}_{B_1} = -\mathbf{a}\,\mathbf{R}^{B_1-2}\ (\mathbf{A}_0 + \mathbf{RA}_1) + \mathbf{b}\ (\mathbf{SA}_0 + \mathbf{A}_1)\ \mathbf{C}_2^{-1}\,.$$

The equations for $B_1 = 0$ (SHARED), $B_1 = 1$, and $B_1 = \infty$ (N-Burst) are somewhat simpler, but they have to be treated separately.

ACKNOWLEDGMENTS

The research for this paper was performed while the author was enjoying the hospitality of the Network Design and Performance Analysis Department at AT&T Labs, NJ, USA. In particular, D. Heyman and V. Ramaswami gave valuable input to this work. Furthermore, the author wants to thank the 'German Academic Exchange Agency' (DAAD), whose fellowship provided the funding for this research.

REFERENCES

[1] V. Paxson and S. Floyd, "Wide area traffic: The failure of poisson modeling," *IEEE/ACM Transactions on Networking*, vol. 3, pp. 226–244, 1995.

[2] H. Heffes, "A class of data traffic processes – covariance function characterization and related queueing results," *Bell Syst. Tech. J.*, vol. 59, pp. 897–930, 1980.

[3] Will Leland, Murad Taqqu, Walter Willinger, and Daniel V. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, February 1994.

[4] Hans-Peter Schwefel and Lester Lipsky, "Performance results for analytic models of traffic in telecommunication systems, based on multiple on-off sources with self-similar behavior," in *Teletraffic Engineering in a Competitive World, Vol 3A*, P. Key and D. Smith, Eds. 1999, pp. 55–66, Elsevier Science B.V.

[5] Hans-Peter Schwefel and Lester Lipsky, "Impact of aggregated, self-similar on/off traffic on delay in stationary queueing models (extended version)," *To appear in Performance Evaluation*, 2000.

[6] Ake Arvidson and Pär Karlsson, "On traffic models for tcp/ip," in *Teletraffic Engineering in a Competitive World, Vol 3A*, P. Key and D. Smith, Eds. 1999, pp. 457–466, Elsevier Science B.V.

[7] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling tcp throughput: A simple model and its empirical validation," in *ACM SIGCOMM*, 1998.

[8] D. Heyman, T. Lakshman, and A. Neidhardt, "A new method for analysing feedback-based protocols with applications to engineering web traffic over the internet," in *ACM Sigmetrics*, 1997, pp. 24–38.

[9] Mark Crovella and Azer Bestavros, "Self-similarity in world wide web traffic: Evidence and possible causes," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835–846, December 1997.

[10] Michael Greiner, Manfred Jobmann, and Lester Lipsky, "The importance of power-tail distributions for telecommunication traffic models," *Operations Research*, vol. 41, no. 6, pp. 313–326, March 1999.

[11] W. Stevens, *TCP/IP Illustrated, Vol.1 The Protocols*, Addison-Wesley, 1994.

[12] Udo Krieger and Valery Naumov, "Analysis of a delay-loss system with a superimposed markovian arrival process and state-dependent service times," in *Proceedings of the MMB conference*. September 1999, Universität Trier.

[13] V. Dumas and A. Simonian, "Asymptotic bounds for the fluid queue fed by sub-exponential on/off sources," *Advances in Applied Probability*, vol. 32, no. 1, March 2000.

[14] Lester Lipsky, *QUEUEING THEORY: A Linear Algebraic Approach*, MacMillan Publishing Company, New York, 1992.

[15] Andras Veres and Miklos Boda, "The chaotic nature of tcp congestion control," in *Proceedings of the IEEE Infocom*, 2000.

[16] Soeren Asmussen, Manfred Jobmann, and Hans-Peter Schwefel, "Exact buffer overflow calculations for queues via martingales," *Submitted to Queueing Systems*, 2000.

[17] Marcel Neuts, *MATRIX-GEOMETRIC SOLUTIONS IN STOCHASTIC MODELS*, John Hopkins University Press, London, 1981.

[18] Hans-Peter Schwefel, "Modeling of packet arrivals using markov modulated poisson processes with power-tail bursts," M.S. thesis, Institut für Informatik, Technische Universität München, August 1997.