Forty-Eighth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 29 - October 1, 2010

Throughput Optimal Scheduling in the presence of Heavy-Tailed Traffic

Krishna Jagannathan, Mihalis Markakis, Eytan Modiano, and John N. Tsitsiklis
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: {krishnaj, mihalis, modiano, jnt}@mit.edu

***Abstract*—We investigate the tail behavior of the steady-state queue occupancies under throughput optimal scheduling in the presence of heavy-tailed traffic. We consider a system consisting of two parallel queues, served by a single server. One of the queues receives traffic that is heavy-tailed (the "heavy queue"), and the other receives light-tailed traffic (the "light queue"). The queues are connected to the server through time-varying ON/OFF links. We study a generalized version of max-weight scheduling, called the max-weight-α policy, and show that the light queue occupancy distribution is heavy-tailed for arrival rates above a threshold value. We also obtain the exact 'tail coefficient' of the light queue occupancy distribution under max-weight-alpha scheduling. Next, we show that the policy that gives complete priority to the light queue guarantees the best possible tail behavior of both queue occupancy distributions. However, the priority policy is not throughput optimal, and can cause undesirable instability effects in the heavy queue.**

Finally, we propose a log-max-weight (LMW) scheduling policy. We show that in addition to being throughput optimal, the LMW policy guarantees that the light queue occupancy distribution is light-tailed, for all arrival rates that the priority policy can stabilize. Thus, the LMW scheduling policy has desirable performance on both fronts, namely throughput optimality, and the tail behavior of the light queue occupancy distribution.

I. Introduction

Traditionally, traffic in telecommunication networks has been modeled using Poisson and Markov-modulated processes. These simple traffic models exhibit 'local randomness', in the sense that much of the variability occurs in short time scales, and only an average behavior is perceived at longer time scales. With the spectacular growth of packet-switched networks such as the internet during the last couple of decades, these traditional traffic models have been shown to be inadequate. This is because the traffic in packetized data networks is intrinsically more 'bursty', and exhibits correlations over longer time scales than can be modeled by any Markovian point process. Empirical evidence, such as the famous Bellcore study on self-similarity and long-range dependence in ethernet traffic [15] lead to increased interest in traffic models with high variability.

Heavy-tailed distributions, which have long been used to model high variability and risk in finance and insurance, were considered as viable candidates to model traffic in data networks. Further, theoretical work such as [13], linking heavy-tails to long-range dependence (LRD) lent weight to the belief that extreme variability in the internet file sizes is ultimately responsible for the LRD traffic patterns reported in [15] and elsewhere.

Many of the early queueing theoretic results for heavy-tailed traffic were obtained for the single server queue; see [3], [5], [20] for surveys of these results. In [6], the authors study the tail behavior of the waiting time in an M/G/2 system, when one of the service time distributions is heavy-tailed, and the other is exponential.

It turns out that the service discipline plays an important role in the delay experienced in a queue, when the traffic is heavy-tailed. For example, it was shown in [1] that any non-preemptive service discipline leads to infinite expected delay, when the traffic is sufficiently heavy-tailed. Further, the asymptotic behavior of delay under various service disciplines such as first-come-first-served (FCFS) and processor sharing (PS), is markedly different under light-tailed and heavy-tailed scenarios [3], [23]. This is important, for example, in the context of scheduling jobs in server farms [12].

In the context of communication networks, a subset of the traffic flows may be well modeled using heavy-tailed processes, and the rest better modeled as light-tailed processes. In such a scenario, there are relatively few studies on the problem of scheduling *between* the different flows, and the ensuing nature of interaction between the heavy-tailed and light-tailed traffic. An important paper in this category is [4], where the interaction between light and heavy-tailed traffic flows under generalized processor sharing (GPS) is studied. In that paper, the authors derive the asymptotic workload behavior of the light-tailed flow, when its GPS weight is greater than its traffic intensity. In a related paper [2], the authors obtain the asymptotic work-load behavior under a general coupled-queues framework, which includes GPS as a special case.

One of the key considerations in the design of a scheduling policy for a queueing network is *throughput optimality*, which is the ability to support the largest set of traffic rates that is supportable by a given queueing network. Queue length based scheduling policies, such as max-weight scheduling [21], [22] and its many variants, are known to be throughput optimal in a general queueing network. For this reason, the max-weight family of scheduling policies has received much attention in various networking contexts, including switches [17], satellites [18], wireless [19], and optical networks [7], [8].

This work was supported by NSF grants CNS-0626781, CNS-0915988, and CCF-0728554, and by ARO Muri grant number W911NF-08-1-0238.

In spite of a large and varied body of literature related to max-weight scheduling, it is somewhat surprising that the policy has not been adequately studied in the context of heavy-tailed traffic. Specifically, a question arises as to what behavior we can expect due to the interaction of heavy and light-tailed flows, when a throughput optimal max-weight-like scheduling policy is employed. Our present work is aimed at addressing this basic question.

We study a system consisting of two parallel queues, served by a single server. One of the queues is fed by a heavy-tailed arrival process, while the other is fed by light-tailed traffic. We refer to these queues as the 'heavy' and 'light' queues, respectively. The queues are connected to the server through time-varying ON/OFF links. In this setting, we analyze the tail behavior of the queue occupancy distributions under max-weight-α scheduling, which is a generalized version of max-weight scheduling. Specifically, while max-weight scheduling makes scheduling decisions by comparing the queue lengths in the system, the max-weight-α policy uses different powers of the queue lengths to make scheduling decisions.

In a recent paper [16], a special case of the problem considered here is studied. Specifically, it was shown that when the heavy-tailed traffic has an infinite variance, the light-tailed traffic experiences an infinite expected delay under max-weight scheduling. It was also shown that by a choice of parameters in the max-weight-α policy that increases the preference afforded to the light queue, it is possible to make the expected delay of the light-tailed traffic finite. In the present paper, we considerably extend these results by characterizing the tail distribution of the queue occupancies under very general heavy-tailed arrival processes, while also allowing for randomly time-varying channels.

Under max-weight-α scheduling, we show that the light queue occupancy distribution is light-tailed if the arrival rate to the light queue is below a certain threshold value, and heavy-tailed if the arrival rate is above the threshold value. Further, when the arrival rate is above the threshold value, we obtain the exact 'tail coefficient' of the queue occupancy distributions, which helps us identify all the bounded moments of the queue lengths.

Intuitively, the reason max-weight-α scheduling induces heavy-tailed asymptotics on the light queue distribution is that the light queue has to compete for service with the heavy queue, which is occasionally very large. The simplest way to guarantee a good tail behavior for the light queue distribution is to give the light queue complete priority over the heavy queue, so that it does not have to compete with the heavy queue for service. However, giving priority to the light queue has an important shortcoming – it is not a throughput optimal scheduling policy for the system, and can cause undesirable instability effects in the heavy queue.

We therefore find ourselves in a situation where on the one hand, the throughput optimal max-weight-α scheduling policy can lead to heavy-tailed asymptotics for the light queue. On the other hand, giving priority to the light queue leads to good tail behavior for the light queue, but is not throughput

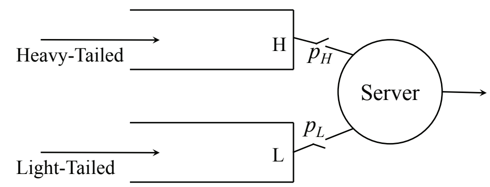

Fig. 1. A system of two parallel queues, with one of them fed with heavy-tailed traffic. The channels connecting the queues to the server are unreliable ON/OFF links.

optimal. To remedy this situation, we propose a log-max-weight (LMW) scheduling policy, which gives significantly more importance to the light queue compared to max-weight-α scheduling. We show that under LMW scheduling, the light queue occupancy distribution is light-tailed for all arrival rates that are stably supportable under priority scheduling for the light queue. Additionally, we show that the LMW policy is throughput optimal, and can therefore stabilize traffic rates that are not supportable under priority scheduling. Thus, the LMW policy has both desirable attributes – it is throughput optimal, and ensures good tail behavior for the light queue distribution.

The remainder of this paper is organized as follows. In Section II, we introduce the system model and specify the necessary definitions and assumptions. In Section III, we study priority scheduling. Section IV deals with queue occupancy behavior under max-weight-α scheduling. In Section V, we analyze the queue occupancy behavior under log-max-weight scheduling. Section VI concludes the paper. Due to space constraints, we omit several proofs, and refer the reader to [14, Chapter 5].

II. System Description and Preliminaries

In this section, we describe the system model, and specify our assumptions about the traffic statistics. Our system consists of two parallel queues, H and L, served by a single server, as depicted in Fig. 1. Time is slotted, and stochastic arrivals of packet bursts occur to each queue in each slot. The server is capable of serving one packet per time slot, from only one of the queues according to a scheduling policy. Let $H(t)$ and $L(t)$ denote the number of packets that arrive during slot t to H and L respectively. Although we postpone the precise assumptions on the traffic statistics to Section II-B, let us loosely say that the input $L(t)$ is light-tailed, and $H(t)$ is heavy-tailed. We will refer to the queues H and L as the heavy and light queues, respectively.

The queues are connected to the server through time-varying links. Let $S_H(t) \in \{0, 1\}$ and $S_L(t) \in \{0, 1\}$ respectively denote the states of the channels connecting the H and L queues to the server. When a channel is in state 0, it is OFF, and no packets can be served from the corresponding queue in that slot. When a channel is in state 1, it is ON, and a packet can be served from the corresponding queue if the server is assigned to that queue. This channel model can be used to represent fading wireless links in a two-user up-link or down-link system. We assume that the scheduler can observe the current channel states as well as the queue lengths before making a scheduling decision in a slot.

The processes $S_H(t)$ and $S_L(t)$ are independent of each other, and independent of the arrival processes. We assume that $S_H(t)$ and $S_L(t)$ are i.i.d. from slot to slot, distributed according to Bernoulli processes with positive means p_H and p_L respectively. That is, $\mathbb{P}\{S_i(\cdot)=1\}=p_i,\ i\in\{H,L\}$. We say that a particular time slot t is *exclusive* to H, if $S_H(t)=1$ and $S_L(t)=0$, and similarly for L.

Before we specify the precise assumptions on the arrival processes, we pause to make some relevant definitions.

A. Heavy-tailed and light-tailed random variables

Definition 1: A non-negative random variable X is said to be *light-tailed* if there exists $\theta>0$ for which $\mathbb{E}[\exp(\theta X)]<\infty$. A random variable is *heavy-tailed* if it is not light-tailed. In other words, a light-tailed random variable is one that has a well defined moment generating function in a neighborhood of the origin. The complementary distribution function of a light-tailed random variable decays at least exponentially fast. Heavy-tailed random variables are those which have complementary distribution functions that decay slower than any exponential. We now define the tail-coefficient of a random variable.

Definition 2: The *tail coefficient* of a random variable X is defined by

$$C_X=\sup\{c\geq 0 \mid \mathbb{E}[X^c]<\infty\}.$$

In words, the tail coefficient is the threshold where the power moment of a random variable starts to blow up. Note that the tail coefficient of a light-tailed random variable is infinite. On the other hand, the tail coefficient of a heavy-tailed random variable may be infinite (e.g., log-normal) or finite (e.g., Pareto). In this paper, we restrict our attention to the class of heavy-tailed random variables which have a finite tail coefficient.

We now state the precise assumptions on the arrival processes.

B. Assumptions on the arrival processes

1) The arrival processes $H(t)$ and $L(t)$ are independent of each other.
2) $H(t)$ is independent and identically distributed (i.i.d.) from slot-to-slot.
3) $L(t)$ is i.i.d. from slot-to-slot.
4) $L(\cdot)$ is light-tailed with $\mathbb{E}[L(t)]=\lambda_L$.
5) $H(\cdot)$ is heavy-tailed with tail coefficient C_H ($1<C_H<\infty$), and $\mathbb{E}[H(t)]=\lambda_H$.

The conditions for a rate pair (λ_H,λ_L) to be stably[1] supportable in this system are well known. Specifically, it follows from the results in [22] that the rate region of the system is given by

$$\{(\lambda_H,\lambda_L)\ \mid 0\leq\lambda_L<p_L,\ 0\leq\lambda_H<p_H,$$
$$\lambda_H+\lambda_L<p_H+p_L-p_Hp_L\}\,. \tag{1}$$

Thus, the rate region is pentagonal, as illustrated by the solid line in Fig. 2.

Let $q_H(t)$ and $q_L(t)$, respectively, denote the number of packets in H and L during slot t under a particular scheduling policy, and let q_H and q_L denote the corresponding steady-state queue occupancies when they exist. Our aim is to characterize the distributions of q_H and q_L under various scheduling policies. We now proceed to analyze the behavior of the steady-state queue occupancies in this system under three scheduling policies, namely, priority scheduling, max-weight-α, and LMW.

III. The Priority Policies

In this section, we study the two 'extreme' scheduling policies, namely priority for L and priority for H. Our analysis helps us arrive at the important conclusion that the tail of the heavy queue is inevitably heavy-tailed under any scheduling policy.

A. Priority for the heavy-tailed traffic

Under priority for H, the heavy queue receives service whenever it is non-empty and connected to the server. L receives service during its exclusive slots, and when both queues are connected, but H is empty. It should be intuitively clear at the outset that this policy is bound to have an undesirable impact on the light queue. The reason we analyze this policy is that it gives us a best case scenario for the heavy queue. The following result shows that the heavy queue occupancy distribution is one order heavier than its input distribution under this policy.

Proposition 1: Under priority for H, the steady-state queue occupancy distribution of the heavy queue is a heavy-tailed random variable with tail coefficient equal to C_H-1. That is, for each $\epsilon>0$, we have

$$\mathbb{E}\left[q_H^{C_H-1-\epsilon}\right]<\infty, \tag{2}$$

and

$$\mathbb{E}\left[q_H^{C_H-1+\epsilon}\right]=\infty. \tag{3}$$

Since priority for H affords the most favorable treatment to the heavy queue, it follows that the tail behavior of H can be no better than the above under *any* policy.

Proposition 2: Under any scheduling policy, q_H is heavy-tailed with tail coefficient at most C_H-1. That is, Equation (3) holds for *all* scheduling policies.

B. Priority for the light-tailed traffic

Under priority for L, the light queue is served whenever its channel is ON, and L is non-empty. The heavy queue is served during the exclusive slots of H, and in the slots when both channels are ON, but L is empty. This policy ensures that the light queue does not have to compete with the heavy queue for service, and hence guarantees the best possible behavior of the light queue occupancy distribution. However, we show that this policy is not throughput optimal, and that it fails to

[1]The notion of stability we use is the positive recurrence of the system occupancy Markov chain.

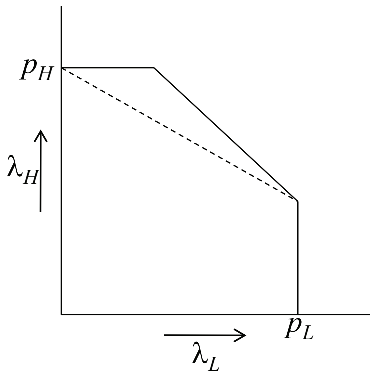

Fig. 2. The rate region of the system is shown in solid line, and the set of stabilizable rates under priority for L is the region under the dashed line.

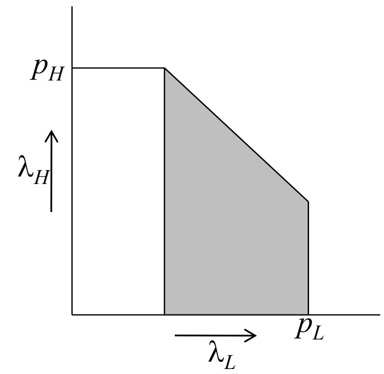

Fig. 3. Under max-weight-α scheduling, q_L is light-tailed for arrival rates in the unshaded region, and heavy-tailed in the shaded region.

stabilize the heavy queue for some arrival rates within the rate region in (1). The following theorem characterizes the behavior of both queues under priority for L.

Theorem 1: The following statements hold under priority scheduling for L.

(i) If $\lambda_H > p_H(1 - \lambda_L)$, the heavy queue is *unstable*, and no steady-state exists.

(ii) If $\lambda_H < p_H(1 - \lambda_L)$, the heavy queue is stable, and its steady-state occupancy q_H is heavy-tailed with tail coefficient $C_H - 1$.

(iii) q_L is light-tailed, and satisfies a large deviation principle [11].

In Fig. 2, the line $\lambda_H = p_H(1 - \lambda_L)$ is shown using a dashed segment. The above theorem asserts that H is stable under priority for L only in the trapezoidal region under the dashed line, while the rate region of the system is clearly larger. Therefore, priority for L is *not* throughput optimal in this setting. To summarize, priority for L can lead to instability of the heavy queue, but for all arrival rates that it *can* stabilize, the tail behavior of both queues is as good as it can possibly be.

The special case in which the queues are always connected to the server, i.e., $p_H = p_L = 1$, is interesting. In this case, the set of arrival rates stabilizable under priority for L coincides with the stability region of the system, which is given by

$$\{(\lambda_H, \lambda_L) \mid \lambda_H + \lambda_L < 1\}.$$

Therefore, when the queues are reliably connected to the server, priority scheduling for the light-tailed traffic is throughput optimal, and also ensures the best possible tail behavior for both queues.

IV. MAX-WEIGHT-α SCHEDULING

In this section, we analyze the tail behavior of the light queue distribution under max-weight-α scheduling. For fixed parameters $\alpha_H > 0$ and $\alpha_L > 0$, the max-weight-α policy operates as follows. During each slot t, compare

$$q_L(t)^{\alpha_L} S_L(t) \gtrless q_H(t)^{\alpha_H} S_H(t),$$

and serve one packet from the queue that wins the comparison. Note that $\alpha_L = \alpha_H$ corresponds to the usual max-weight policy, which serves the longest connected queue in each slot. $\alpha_L/\alpha_H > 1$ corresponds to emphasizing the light queue over the heavy queue, and vice-versa.

It can be shown using standard Lyapunov arguments that max-weight-α scheduling is throughput optimal for all $\alpha_H > 0$ and $\alpha_L > 0$. That is, it can stably support all arrival rates within the rate region (1). This throughput optimality result follows, for example, from [9, Theorem 1].

We show that under max-weight-α scheduling, the tail behavior of the steady-state light queue occupancy distribution is strongly dependent on λ_L, the arrival rate to the light queue. Specifically, we show that q_L is light-tailed when λ_L is below a threshold value, and heavy-tailed with a finite tail coefficient for λ_L above the threshold value.

The following result shows that the light queue distribution is light-tailed under any 'reasonable' policy, as long as the rate λ_L is smaller than a threshold value.

Proposition 3: Suppose that $\lambda_L < p_L(1 - p_H)$. Then q_L is light-tailed under *any* policy that serves L during its exclusive slots.

Proof: The proof is straightforward once we note that the exclusive slots of L occur independently during each slot with probability $p_L(1-p_H)$. Indeed, consider the L queue under a policy that serves L *only* during its exclusive slots. Under this policy, the L queue behaves like a G/M/1 queue with light-tailed inputs at rate λ_L, and service rate $p_L(1 - p_H)$. It can be shown using standard large deviation arguments that q_L is light-tailed under the policy that serves L only during its exclusive slots. Therefore, q_L is light-tailed under any policy that serves L during its exclusive slots. $\square$

The above proposition implies that for $\lambda_L < p_L(1 - p_H)$, the light queue distribution is light-tailed under max-weight-α scheduling. The region $\lambda_L < p_L(1-p_H)$ is shown unshaded in Fig. 3. Thus, q_L is light-tailed under max-weight-α scheduling for arrival rates in the unshaded region.

In the remainder of this section, we investigate the tail behavior of the light queue under max-weight-α scheduling when the arrival rate is above the threshold, i.e., for $\lambda_L > p_L(1-p_H)$. In this case, the light queue receives traffic at a higher rate than can be supported by the exclusive slots of L alone. Therefore, the light queue has to compete for service with the heavy queue during the slots that both channels are ON. Since the heavy queue is very large with positive probability, it seems intuitively reasonable that the light queue will suffer from this competition, and also take on a heavy-tailed behavior. This intuition is indeed correct, although proving the result is non-trivial.

We prove that the light queue distribution is heavy-tailed when $\lambda_L > p_L(1-p_H)$ for all values of the scheduling parameters α_L and α_H. We also obtain the exact tail coefficient of the light queue distribution for 'plain' max-weight scheduling ($\alpha_L/\alpha_H = 1$), and for the regime where the light queue is given more importance ($\alpha_L/\alpha_H > 1$).

A. Max-weight scheduling

Let us first characterize the tail coefficient of the steady-state light queue occupancy under the max-weight policy, which serves the longest connected queue in each slot. Since q_L is light-tailed for $\lambda_L < p_L(1-p_H)$ according to Proposition 3, we will focus on the case $\lambda_L > p_L(1-p_H)$.

Theorem 2: Suppose that $\lambda_L > p_L(1-p_H)$. Then, under max-weight scheduling, q_L is heavy-tailed with tail coefficient $C_H - 1$.

In terms of Fig. 3, the theorem asserts that q_L is heavy-tailed with tail coefficient $C_H - 1$ for all arrival rates in the shaded region. Proving the above result involves showing (i) an upper bound: $\mathbb{E}\left[q_H^{C_H-1-\epsilon}\right] < \infty$, and (ii) a lower bound: $\mathbb{E}\left[q_H^{C_H-1+\epsilon}\right] = \infty$, for all $\epsilon > 0$. We deal with each of them below.

1) Upper Bound for max-weight scheduling:

Proposition 4: Under max-weight scheduling, we have

$$\mathbb{E}\left[q_L^{C_H-1-\epsilon}\right] < \infty, \ \forall\ \epsilon > 0.$$

Proof: This is a special case of Proposition 6, in the next section. ☐

2) Lower Bound for max-weight scheduling:

Proposition 5: Suppose that $\lambda_L > p_L(1-p_H)$. Then, under max-weight scheduling, we have

$$\mathbb{E}\left[q_L^{C_H-1+\epsilon}\right] = \infty, \ \forall\ \epsilon > 0.$$

The proof of this result is quite involved, so we informally describe the idea behind its construction, and refer the reader to [14, Proposition 5.4] for the formal proof. In our intuitive argument, we will 'show' that

$$\lim_{t\to\infty} \mathbb{E}\left[q_L(t)^{C_H-1+\epsilon}\right] = \infty. \tag{4}$$

The above is the limit of the expectation of a sequence of random variables, whereas what we really want in Proposition 5 is the expectation of the limiting random variable q_L. Although it is by no means obvious that the limit and the expectation can be interchanged here, we will ignore this as a technical point in our informal argument.

The main idea behind the proof is to consider the renewal intervals that commence at the beginning of each busy period of the system. Without loss of generality, let us consider a busy period that commences at time 0, and define the renewal reward process $R(t) = q_L(t)^{C_H-1+\epsilon}$. By the key renewal theorem [10],

$$\lim_{t\to\infty} \mathbb{E}[R(t)] = \frac{\mathbb{E}[R]}{\mathbb{E}[T]},$$

where $\mathbb{E}[R]$ denotes the expected reward accumulated over a renewal interval, and $\mathbb{E}[T] < \infty$ is the mean renewal interval. It is therefore enough to show that $\mathbb{E}\left[\sum_{i=0}^{T} q_L(i)^{C_H-1+\epsilon}\right] = \infty$.

To see intuitively why the above expectation is infinite, let us condition on the busy period commencing at time 0 with a burst of size b to the heavy queue[2]. After this instant, the heavy queue drains at rate p_H, assuming for the sake of a lower bound that there are no further bursts arriving at H. In the mean time, the light queue receives traffic at rate λ_L, and gets served only during the exclusive slots of L, which occur at rate $p_L(1-p_H)$. With high probability therefore, the light queue will steadily build up at rate $\lambda_L - p_L(1-p_H)$, until it eventually catches up with the draining heavy queue. It can be shown that the light queue will build up to an $O(b)$ level before it catches up with the heavy queue. Further, the light queue occupancy stays at $O(b)$ for a time interval of length $O(b)$. Therefore, with high probability, the reward is at least $O(b^{C_H-1+\epsilon})$ for $O(b)$ time slots. Thus, for some constant K,

$$\mathbb{E}\left[\sum_{i=0}^{T} q_L(i)^{C_H-1+\epsilon}\right] \geq \mathbb{E}\left[Kb \cdot b^{C_H-1+\epsilon}\right] = \mathbb{E}\left[Kb^{C_H+\epsilon}\right].$$

The last expectation is infinite because the initial burst size has tail coefficient equal to C_H.

In words, the light queue not only grows to a level proportionate to the initial burst size, but also stays large for a period of time that is proportional to the burst size. This leads to a light queue distribution that is one order heavier than the burst size distribution.

B. Max-weight-α scheduling with $\alpha_L > \alpha_H$

In this subsection, we characterize the exact tail coefficient of the light queue distribution under max-weight-α scheduling, with $\alpha_L > \alpha_H$. We only treat the case $\lambda_L > p_L(1-p_H)$, since q_L is known to be light-tailed otherwise. Our main result for this regime is the following.

Theorem 3: Suppose that $\lambda_L > p_L(1-p_H)$. Then, under max-weight-α scheduling with $\alpha_L > \alpha_H$, q_L is heavy-tailed with tail coefficient

$$\gamma = \frac{\alpha_L}{\alpha_H}(C_H - 1). \tag{5}$$

In terms of Fig. 3, the above theorem asserts that q_L is heavy-tailed with tail coefficient γ for all arrival rates in the shaded region. As before, proving this result involves showing (i) an upper bound of the form $\mathbb{E}\left[q_H^{\gamma-\epsilon}\right] < \infty$, and (ii) a lower

[2]It is easy to show that this event has positive probability.

bound of the form $\mathbb{E}\left[q_H^{\gamma+\epsilon}\right] = \infty$, for all $\epsilon > 0$. We deal with each of them separately.

1) Upper Bound for max-weight-α scheduling:

Proposition 6: Under max-weight-α scheduling, we have

$$\mathbb{E}\left[q_L^{\gamma-\epsilon}\right] < \infty, \ \forall \ \epsilon > 0.$$

Proof: The result is a consequence of a theorem in [9]. Indeed, max-weight-α scheduling in our context is equivalent to comparing $q_L(t)^{\beta\alpha_L} S_L(t)$ versus $q_H(t)^{\beta\alpha_H} S_H(t)$, where $\beta > 0$ is arbitrary, and scheduling the winning queue in each slot. In particular, if we choose $\beta = (C_H - 1)/\alpha_H - \epsilon/\alpha_L$, the conditions imposed in [9, Theorem 1] are satisfied for any $\epsilon > 0$, so that the steady-state queue occupancies satisfy

$$\mathbb{E}\left[q_L^{\gamma-\epsilon}\right] < \infty,$$

and

$$\mathbb{E}\left[q_H^{C_H-1-\frac{\alpha_H}{\alpha_L}\epsilon}\right] < \infty. \tag{6}$$

□

Remark 1: (i) Proposition 6 is valid for any parameters α_L and α_H, and not just for $\alpha_L > \alpha_H$.

(ii) Equation (6) and Proposition 2 together imply that the tail coefficient of q_H is equal to $C_H - 1$ under max-weight-α scheduling, for any parameters α_L and α_H.

2) Lower Bound for max-weight-α scheduling with $\alpha_L > \alpha_H$:

Proposition 7: Suppose that $\lambda_L > p_L(1-p_H)$. Then, under max-weight-α scheduling with $\alpha_L > \alpha_H$, we have

$$\mathbb{E}\left[q_L^{\gamma+\epsilon}\right] = \infty, \ \forall \ \epsilon > 0.$$

The proof is lengthy and intricate, but conceptually similar to the proof of Proposition 5. We present an informal sketch, and refer the reader to [14, Proposition 5.6] for the complete proof. We consider the renewal process that commences at the beginning of each busy period of the system, and define the reward process $R_\gamma(t) = q_L(t)^{\gamma+\epsilon}$. We will show that the expected reward accumulated over a renewal interval is infinite. The key renewal theorem would then imply that $\lim_{t\to\infty} \mathbb{E}\left[q_L(t)^{\gamma+\epsilon}\right] = \infty$. Finally, the result we want can be obtained by invoking a truncation argument to interchange the limit and the expectation.

To intuitively see why the expected reward over a renewal interval is finite, let us condition on the busy period commencing with a burst of size b at the heavy queue. Starting at this instant, the light queue will build up at the rate $\lambda_L - p_L(1-p_H)$ with high probability. However, the light queue only builds up to an $O(b^{\alpha_H/\alpha_L})$ level before it 'catches up' with the heavy queue and wins back the service preference. It can also be shown that the light queue catches up in a time interval of length $O(b^{\alpha_H/\alpha_L})$. It might therefore be tempting to argue that the light queue stays above $O(b^{\alpha_H/\alpha_L})$ for an interval of duration $O(b^{\alpha_H/\alpha_L})$. Although this argument is not incorrect as such, it fails to capture what typically happens in the system. Let us briefly follow through with this argument, and conclude that it does not give us the lower bound we want.

Indeed, following the above argument, the reward is at least $O(b^{(\gamma+\epsilon)\alpha_H/\alpha_L}) = O(b^{C_H-1+\epsilon\alpha_H/\alpha_L})$ for $O(b^{\alpha_H/\alpha_L})$ time slots, so that the expected reward over the renewal interval is lower bounded by $\mathbb{E}_b\left[O(b^{C_H-1+\alpha_H/\alpha_L+\epsilon\alpha_H/\alpha_L})\right]$. However, the above expectation turns out to be *finite* for $\alpha_L/\alpha_H > 1$. Therefore, the above simple bound fails to give the result we are after.

The problem with the above argument is that it looks at the time scale at which the light queue *catches up*, whereas the event that decides the tail coefficient happens *after* the light queue catches up. In particular, the light queue catches up relatively quickly, in a time scale of $O(b^{\alpha_H/\alpha_L})$. However, after the light queue catches up with the heavy queue, the two queues drain together, with most of the slots being used to serve the heavy queue. In fact, as we show, before the light queue occupancy can drain by a constant factor after catch-up, the heavy queue drains by $O(b)$. As such, the light queue remains at an $O(b^{\alpha_H/\alpha_L})$ level for $O(b)$ time slots. Therefore, the expected reward can be lower bounded by

$$\mathbb{E}_b\left[O(b)O(b^{C_H-1+\epsilon\alpha_H/\alpha_L})\right] = \mathbb{E}_b\left[O(b^{C_H+\epsilon\alpha_H/\alpha_L})\right] = \infty,$$

which is what we want. In sum, the light queue builds up relatively quickly until catch-up, but takes a long time to drain out after catch-up.

C. Max-weight-α scheduling with $\alpha_L < \alpha_H$

We finally consider the case $\alpha_L < \alpha_H$, and study the tail behavior of q_L. Recall that max-weight-α scheduling with $\alpha_L < \alpha_H$ corresponds to giving the heavy queue preference over the light queue. In this regime, we show that q_L is heavy-tailed with a finite tail coefficient, for arrival rates in the shaded region of Fig. 3. However, we are unable to determine the exact tail coefficient of q_L for some arrival rate pairs in this regime.

Our first result for this case is an upper bound on the tail coefficient of q_L. Intuitively, we would expect that the tail behavior of q_L in this regime cannot be better than it is under max-weight scheduling. In other words, the tail coefficient of q_L in this regime cannot be larger than $C_H - 1$. This intuition is indeed correct.

Proposition 8: Suppose that $\lambda_L > p_L(1-p_H)$. Then, under max-weight-α scheduling with $\alpha_L < \alpha_H$, the tail coefficient of q_L is at most $C_H - 1$.

Proof: Follows similarly to the proof of Proposition 5. Specifically, conditioning on an initial burst of size b arriving to the heavy queue, it can be shown that with high probability, q_L will be $O(b)$ in size for at least $O(b)$ time slots. □

Next, to obtain a lower bound on the tail coefficient of q_L, recall that Proposition 6 holds for the present regime as well. Thus, γ (defined in (5)) is a lower bound[3] on the tail coefficient of q_L. In sum, we have shown that for $\lambda_L > p_L(1-p_H)$, the light queue occupancy distribution is heavy-tailed, with a tail coefficient that lies in the interval $[\gamma, C_H - 1]$.

It turns out that we can obtain the exact tail coefficient of q_L for arrival rates in a *subset* of the shaded region in Fig. 3.

[3]Note that γ is *smaller* than $C_H - 1$ in this regime.

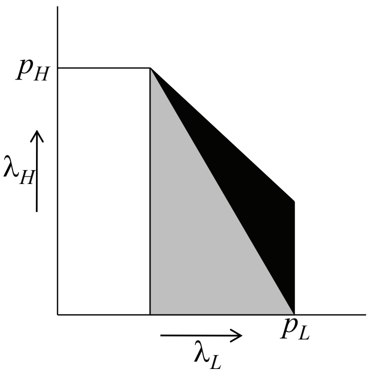

Fig. 4. Under max-weight-α scheduling with $\alpha_L < \alpha_H$, q_L is light-tailed for arrival rates in the unshaded region, and heavy-tailed with tail coefficient equal to $C_H - 1$ in for arrival rates in the gray region. For arrival rates in the region colored black, the tail coefficient lies in $[\gamma, C_H - 1]$.

Specifically, consider the region represented by $p_L(1-p_H) < \lambda_L < p_L(1-\lambda_H)$. In Fig. 4, this region is shown in gray. It can be shown that all arrival rates in the region shaded gray can be stabilized under priority for H. Furthermore, under priority for H, it can be shown that q_L is heavy-tailed with tail coefficient equal to $C_H - 1$, when $p_L(1-p_H) < \lambda_L < p_L(1-\lambda_H)$.

Since the tail of q_L under max-weight-α scheduling with any parameters is no worse than under priority for H, we can conclude that the tail coefficient of q_L is at least $C_H - 1$ when $p_L(1-p_H) < \lambda_L < p_L(1-\lambda_H)$. Combining this with Proposition 8, we conclude that the tail coefficient q_L is *equal* to $C_H - 1$, when the arrival rate pair lies in the gray region of Fig. 4.

Proposition 9: Suppose that $p_L(1-p_H) < \lambda_L < p_L(1-\lambda_H)$. Then, under max-weight-α scheduling with $\alpha_L < \alpha_H$, the tail coefficient of q_L is equal to $C_H - 1$.

The region shaded black in Fig. 4 ($\lambda_L > p_L(1-\lambda_H)$) corresponds to the arrival rates for which priority for H is *not* stabilizing[4]. Under max-weight-α scheduling with $\alpha_L < \alpha_H$, we are unable to determine the exact tail coefficient of q_L for arrival rates in the black region of Fig. 4. However, we have shown earlier that the tail coefficient lies in the interval $[\gamma, C_H - 1]$. We conjecture that the tail coefficient of q_L equals γ, for arrival rates in the region shaded black.

D. Special case of reliable links

The tail behavior of q_L under max-weight-α scheduling, in the special case of reliably connected links ($p_H = p_L = 1$) is interesting. Specifically, it follows from the results above that the light queue occupancy distribution is heavy-tailed under max-weight-α scheduling, for any values of the scheduling parameters and non-zero arrival rates. The tail coefficient of q_L in this special case is given by the following proposition, which follows from our earlier analysis.

Proposition 10: Suppose that the queues are always connected to the server. Then, under max-weight-α scheduling, q_L is heavy-tailed for all non-zero arrival rates. Further, the tail coefficient of q_L is given by

(i) $C_H - 1$ for $\frac{\alpha_L}{\alpha_H} \le 1$, and
(ii) $\gamma = \frac{\alpha_L}{\alpha_H}(C_H - 1)$ for $\frac{\alpha_L}{\alpha_H} > 1$.

[4]This case is symmetric to the case in Theorem 1(i).

V. Log-Max-Weight Scheduling

In this section, we study the performance of log-max-weight scheduling policy. During each time slot t, the log-max-weight policy compares

$$q_L(t)S_L(t) \gtreqless \log(1 + q_H(t))S_H(t),$$

and serves one packet from the queue that wins the comparison.

The main idea in the LMW policy is to give preference to the light queue to a far greater extent than any max-weight-α policy. Specifically, for $\alpha_L/\alpha_H > 1$, the max-weight-α policy compares q_L to a power of q_H that is smaller than 1. On the other hand, LMW scheduling compares q_L to a logarithmic function of q_H, leading to a significant preference for the light queue. We will show that this significant de-emphasis of the heavy queue with respect to the light queue ensures a better tail behavior for q_L compared to max-weight-α scheduling.

Furthermore, the LMW policy has another useful property when the heavy queue gets overwhelmingly large. Although the LMW policy significantly de-emphasizes the heavy queue, it does not *ignore* it, unlike priority for L. That is, if the H queue occupancy gets overwhelmingly large compared to L, the LMW policy will serve the heavy queue. In contrast, priority for L will ignore any build-up in H, as long as L is non-empty. This property ensures that the LMW policy stabilizes all arrival rates within the rate region in (1).

We show that LMW scheduling has desirable performance on both fronts, namely throughput optimality, and the tail behavior of the light queue occupancy. The LMW policy can be shown to be throughput optimal, using the results in [9]. In terms of the tail, we show that the LMW policy guarantees that the light queue occupancy distribution is light-tailed, for all arrival rates that can be stabilized by priority for L. For arrival rates that are not stabilizable under priority for L, the LMW policy will still stabilize the system, although we are not able to guarantee that q_L is light-tailed for these arrival rates.

Let us now state the main result regarding LMW scheduling.

Theorem 4: Under LMW scheduling, q_L is light-tailed if at least one of the following conditions hold:

(i) $\lambda_L < p_L(1-p_H)$, or
(ii) $\lambda_H < p_H(1-\lambda_L)$.

Note that for $\lambda_L < p_L(1-p_H)$, q_L is easily seen to be light-tailed under LMW scheduling, since the arrival rate is small enough to be supported by the exclusive slots of L. The second condition in Theorem 4 states that for all arrival rates that can be stabilized under priority for L (i.e., the trapezoidal region in Fig. 2), q_L is light-tailed under LMW scheduling.

The union of the two regions in which q_L is light-tailed according to Theorem 4 is shown unshaded in Fig. 5. As can

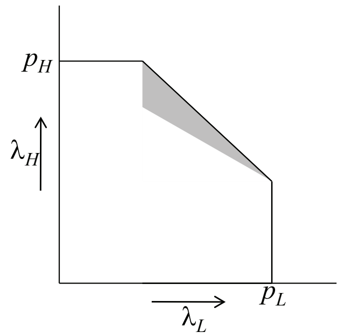

Fig. 5. Under LMW scheduling, q_L is light-tailed for arrival rates in the unshaded region. In the shaded triangle, the tail behavior of q_L is unknown.

be seen, the unshaded region occupies most of the rate region, except for the shaded triangle. For arrival rates in the shaded triangle, the LMW policy still stabilizes the system. However, we are unable to determine the tail behavior of q_L for arrival rates in the shaded triangle.

VI. Conclusions

We considered a system of parallel queues fed by a mix of heavy-tailed and light-tailed traffic, and served by a single server. We studied the tail behavior of the queue occupancy distributions under various scheduling policies. We showed that the occupancy distribution of the heavy queue is inevitably heavy-tailed. In contrast, the light queue occupancy distribution can be heavy-tailed or light-tailed, depending on the arrival rates and the scheduling policy. A major contribution of this paper is in the tail characterization of the queue occupancy distributions under max-weight-α scheduling. We showed that the light queue occupancy distribution under max-weight-α scheduling is light-tailed for arrivals rates below a certain threshold, and heavy-tailed for arrival rates above the threshold.

Another important contribution of the paper is the log-max-weight policy, and the corresponding asymptotic analysis. We showed that the light queue occupancy is light-tailed under LMW scheduling, for all arrival rates that are stabilizable under priority for the light queue. Additionally, the LMW policy also has the desirable property of being throughput optimal, unlike priority scheduling.

Although we study a very simple queueing network in this paper, we believe that the insights obtained from this study are valuable in much more general settings. For instance, in a general queueing network with a mix of light-tailed and heavy-tailed traffic flows, we expect that the celebrated max-weight policy has the tendency to 'infect' competing light-tailed flows with heavy-tailed asymptotics. A similar effect was also noted in [16], in the context of expected delay.

We also believe that the LMW policy occupies a unique 'sweet spot' in the context of scheduling light-tailed traffic in the presence of heavy-tailed traffic. This is because the LMW policy de-emphasizes the heavy-tailed flow sufficiently to maintain good light queue asymptotics, while also ensuring network-wide stability.

For future work, we propose the extension of the results in this paper to more general single-hop and multi-hop network models.

References

[1] V. Anantharam, "Scheduling strategies and long-range dependence," *Queueing Systems*, vol. 33, no. 1, pp. 73–89, 1999.

[2] S. Borst, O. Boxma, and M. Van Uitert, "The asymptotic workload behavior of two coupled queues," *Queueing Systems*, vol. 43, no. 1, pp. 81–102, 2003.

[3] S. Borst, O. Boxma, R. Núñez-Queija, and A. Zwart, "The impact of the service discipline on delay asymptotics," *Performance Evaluation*, vol. 54, no. 2, pp. 175–206, 2003.

[4] S. Borst, M. Mandjes, and M. van Uitert, "Generalized processor sharing with light-tailed and heavy-tailed input," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, p. 834, 2003.

[5] O. Boxma and B. Zwart, "Tails in scheduling," *ACM SIGMETRICS Performance Evaluation Review*, vol. 34, no. 4, p. 20, 2007.

[6] O. Boxma, Q. Deng, and A. Zwart, "Waiting-time asymptotics for the M/G/2 queue with heterogeneous servers," *Queueing Systems*, vol. 40, no. 1, pp. 5–31, 2002.

[7] A. Brzezinski and E. Modiano, "Dynamic reconfiguration and routing algorithms for IP-over-WDM networks with stochastic traffic," *Journal of Lightwave Technology*, vol. 23, no. 10, p. 3188, 2005.

[8] ——, "Achieving 100% Throughput in Reconfigurable Optical Networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 4, pp. 970 – 983, 2008.

[9] A. Eryilmaz, R. Srikant, and J. Perkins, "Stable scheduling policies for fading wireless channels," *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, pp. 411–424, 2005.

[10] R. Gallager, *Discrete stochastic processes*. Kluwer, 1996.

[11] A. Ganesh, N. O'Connell, and D. Wischik, *Big queues*. Springer Verlag, 2004.

[12] M. Harchol-Balter, A. Scheller-Wolf, and A. Young, "Surprising results on task assignment in server farms with high-variability workloads," in *ACM SIGMETRICS 2009*. ACM, 2009, pp. 287–298.

[13] D. Heath, S. Resnick, and G. Samorodnitsky, "Heavy tails and long range dependence in on/off processes and associated fluid models," *Mathematics of Operations Research*, vol. 23, no. 1, pp. 145–165, 1998.

[14] K. Jagannathan, "Asymptotic performance of queue-length-based network control policies," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2010.

[15] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on networking*, vol. 2, no. 1, pp. 1–15, 1994.

[16] M. Markakis, E. Modiano, and J. Tsitsiklis, "Scheduling policies for single-hop networks with heavy-tailed traffic," in *Allerton Conference, Monticello IL*, 2009.

[17] N. McKeown, A. Mekkittikul, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch," *IEEE Transactions on Communications*, vol. 47, no. 8, pp. 1260–1267, 1999.

[18] M. Neely, E. Modiano, and C. Rohrs, "Power and server allocation in a multi-beam satellite with time varying channels," in *IEEE INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, vol. 3, 2002.

[19] ——, "Dynamic power allocation and routing for time varying wireless networks," in *IEEE INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 1, 2003.

[20] K. Park and W. Willinger, *Self-similar network traffic and performance evaluation*. John Wiley & Sons, Inc. New York, NY, USA, 2000.

[21] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, 1992.

[22] ——, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 466–478, 1993.

[23] A. Wierman and B. Zwart, "Is tail-optimal scheduling possible?" Under Submission.