Two-Level Processor-Sharing Scheduling Disciplines: Mean Delay Analysis

Samuli Aalto
Helsinki University of Technology
samuli.aalto@hut.fi

Urtzi Ayesta*
France Telecom R&D
urtzi.ayesta@francetelecom.com

Eeva Nyberg-Oksanen
Helsinki University of Technology
eeva.nyberg@hut.fi

ABSTRACT

Inspired by several recent papers that focus on scheduling disciplines for network flows, we present a mean delay analysis of Multilevel Processor Sharing (MLPS) scheduling disciplines in the context of M/G/1 queues. Such disciplines have been proposed to model the effect of the differentiation between short and long TCP flows in the Internet. Under MLPS, jobs are classified into classes depending on their attained service. We consider scheduling disciplines where jobs within the same class are served either with Processor Sharing (PS) or Foreground Background (FB) policy, and the class that contains jobs with the smallest attained service is served first. It is known that the FB policy minimizes (maximizes) the mean delay when the hazard rate of the job size distribution is decreasing (increasing). Our analysis, based on pathwise and meanwise arguments of the unfinished truncated work, shows that Two-Level Processor Sharing (TLPS) disciplines, e.g., FB+PS and PS+PS, are better than PS scheduling when the hazard rate of the job size distribution is decreasing. If the hazard rate is increasing and bounded, we show that PS outperforms PS+PS and FB+PS. We further extend our analysis to study local optimality within a level of an MLPS scheduling discipline.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: Performance Attributes; F.2.2 [**Nonnumerical Algorithms and Problems**]: Sequencing and Scheduling; G.3 [**Probability and Statistics**]: Queueing Theory

General Terms

Performance, Algorithms

*Urtzi Ayesta is also with INRIA Sophia Antipolis, email: uayesta@sophia.inria.fr

Keywords

Scheduling, MLPS, PS, FB, LAS, M/G/1, mean delay, unfinished truncated work

1. INTRODUCTION

Seminal work on one server scheduling disciplines was made in the 1960's and 1970's. Schrage [1] proved that the Shortest Remaining Processing Time (SRPT) discipline minimizes the mean delay, i.e., the expected time in the system. The SRPT discipline, however, requires the knowledge of the remaining service times of all jobs, but these remaining service times are not always known, e.g., the remaining length of an Internet session is not known by the scheduler. Kleinrock [2] gives a thorough overview of scheduling in such systems where only the attained service times are known together with the service time distribution. Yashkov [3] mentions that he has proven (in [4]) that the Foreground Background (FB)[1] discipline is optimal among this set of scheduling disciplines if the hazard rate of the service time distribution is decreasing (DHR). Righter and Shantikumar [5] showed the optimality of FB in the stochastic sense, i.e., the tail of the queue-length distribution, $P[N^{FB} > k]$, is minimal for all $k \geq 0$. Righter et al. [6] show that FB minimizes the mean delay when the job size distribution is of type New Worse than Used in Expectation (NWUE), which is a weaker condition than DHR. In a recent paper, Feng and Misra [9] prove the optimality of FB when the hazard rate is decreasing.[2] The performance of FB under job size distributions with high variability and simulations of FB on the packet level has been studied by Rai et al. [10].

These job size aware scheduling disciplines have lately received new attention [7, 8, 9, 10, 11, 12] mainly due to the measurement findings that Internet traffic is heavy tailed [13, 14]. This means that data transfers are composed of many very short flows, but the majority of traffic is due to few very long flows. In such a case, it seems reasonable to favor the short flows by a suitable scheduling algorithm. For example, if a router is able to determine, for each flow, the number of bytes or packets already transmitted, the flows could be divided into different priority classes with the highest priority given to the flows with the least number of transmitted bytes or packets. Such implementations are discussed

[1]Also called Feedback Processor Sharing (FBPS) and Least Attained Service (LAS).
[2]There seems, however, to be a subtle deficiency in [6, 9] regarding the proof of FB's optimality. We discuss this further in Section 3.

in [15, 7, 12]. Another method, based on the packet arrival rates within each flow, is described in [16].

On the flow level, these packet level scheduling algorithms can be modelled using Multilevel Processor Sharing (MLPS) scheduling disciplines.[3] An MLPS discipline is defined by a finite set of thresholds

$$0 = a_0 < a_1 < \cdots < a_N < a_{N+1} = \infty,$$

which define $N + 1$ levels. Jobs are classified into levels based on their attained service. Between these levels, a strict priority discipline is applied with the lowest level having the highest priority. Thus, those jobs that have attained service less than a_1 time units are served first. In this paper, the MLPS disciplines with two levels are called Two-Level Processor Sharing (TLPS) disciplines.

With respect to the performance analysis of these MLPS disciplines, the ordinary Processor Sharing (PS) discipline is a natural point of comparison, since it has been proposed as an ideal model for the bandwidth sharing among TCP flows in a bottleneck router [17, 18]. Using a similar reasoning, the packet level priority discipline for TCP flows described above could be modelled as an MLPS discipline with PS used as the internal discipline within each level. Another natural point of comparison is FB due to its optimality property in the case of job size distributions with decreasing hazard rate, examples of which are the Pareto and hyper-exponential distributions that have been used to model flow sizes in the Internet.

In addition to the FB discipline, Feng and Misra [9] consider a TLPS discipline (called FLIPS) where those jobs with attained service time less than threshold a are served according to the FB discipline. If such jobs do not exist, the ordinary PS discipline is used. However, they do not compare the FLIPS discipline to the other disciplines in the mean delay sense. Guo and Matta [7] study, by simulations, another TLPS discipline (called ML-PRIO) where those jobs with attained service time less than threshold a are given priority and served according to the PS discipline. If such jobs do not exist, the ordinary PS discipline is used. The same discipline is considered, by analytical and numerical techniques, in [12]. Their numerical experimentations suggest that ML-PRIO is better than PS when the hazard rate is decreasing.

In this paper we compare the mean delays of TLPS disciplines for which the scheduling discipline within one level is either FB or PS, more specifically FB+PS (FLIPS) and PS+PS (ML-PRIO), in the context of M/G/1 queues. We show that (i) if the hazard rate is decreasing, then

$$E[T^{\mathrm{FB}}] \leq E[T^{\mathrm{FB+PS}}] \leq E[T^{\mathrm{PS+PS}}] \leq E[T^{\mathrm{PS}}],$$

but (ii) if the hazard rate is increasing and bounded, then

$$E[T^{\mathrm{FB}}] \geq E[T^{\mathrm{FB+PS}}] \geq E[T^{\mathrm{PS+PS}}] \geq E[T^{\mathrm{PS}}].$$

For more general MLPS disciplines, we show that, within a level, FB is always better than PS if the hazard rate is decreasing and vice versa for a job size distribution with increasing and bounded hazard rate. The results follow from a thorough analysis of the unfinished truncated work (defined in Section 2.1) including both meanwise and pathwise arguments. In particular, we provide a framework to prove that the mean unfinished truncated work is smaller under a given scheduling discipline compared to some other scheduling discipline.

[3]Introduced by L. Kleinrock in the early 1970's, see [2].

The paper is organized as follows. Section 2 derives how, for different scheduling disciplines, differences in the mean delay can be deduced from differences in the mean unfinished truncated work, given a job size distribution and its hazard rate. In particular, we show that if the mean unfinished truncated work in one scheduling discipline is smaller than in another, the scheduling discipline has a smaller mean delay when the hazard rate of the job size distribution is decreasing. In Section 2 we also provide a framework that gives a sufficient condition for the comparison of the mean unfinished truncated work for different scheduling disciplines. Section 3 relaxes the Poisson arrival assumption and scheduling disciplines are compared using pathwise arguments on the unfinished truncated work for TLPS and MLPS scheduling disciplines. Section 4 summarizes the results in form of theorems, which order several TLPS disciplines according to the mean delay and within a level of an MLPS scheduling discipline, given assumptions on the hazard rate and the job size distribution. Section 5 concludes the paper.

2. MEAN DELAY

Consider any work-conserving scheduling discipline of an M/G/1 queue. Let λ denote the arrival rate and S the service time of a job. We assume that $E[S] < \infty$ and the system be stable, i.e., $\rho = \lambda E[S] < 1$. Furthermore, we assume that the service time distribution is continuous with the corresponding density function denoted by $f(x)$. Let

$$F(x) = \int_0^x f(y)\,dy$$

and $\overline{F}(x) = 1 - F(x)$. The corresponding hazard rate function is denoted by

$$h(x) = \frac{f(x)}{\overline{F}(x)}.$$

2.1 Mean delay in terms of mean unfinished truncated work

Let N denote the number of jobs in the system at an arbitrary time, and denote the service times of these jobs by S_i and their attained services by X_i, $i \in \{1, \ldots, N\}$. The random variable U_x refers to the *unfinished truncated work* with truncation threshold x, i.e., the sum of remaining truncated service times of those jobs with attained service less than x time units,

$$U_x = \sum_{i=1}^{N} ((S_i \wedge x) - (X_i \wedge x)),$$

where we have used notation $a \wedge b = \min\{a, b\}$. Note that $U_0 = 0$.

Let $T(y)$ denote the delay of a job whose service time is y time units, and $N(y)$ the number of such jobs who has attained service at most y time units. According to [2, equation (4.11)],

$$\overline{N}'(y) = \lambda \overline{F}(y) \overline{T}'(y), \tag{1}$$

where we have used notation $\overline{N}'(y) = \frac{d}{dy}E[N(y)]$ and $\overline{T}'(y) = \frac{d}{dy}E[T(y)]$. In addition to the continuous part, for which $\overline{N}(y_i^+) - \overline{N}(y_i^-) = 0$, there may be a countable number of points y_i such that $\overline{N}(y_i^+) - \overline{N}(y_i^-) > 0$. For example, for the FIFO discipline, point 0 is such, and for the MLPS disciplines, the thresholds $a_n > 0$ are such. By Little's result, we conclude that, for all such points y_i,

$$\overline{N}(y_i^+) - \overline{N}(y_i^-) = \lambda \overline{F}(y_i)(\overline{T}(y_i^+) - \overline{T}(y_i^-)). \qquad (2)$$

Formulas (1) and (2) can be combined as follows:

$$d\overline{N}(y) = \lambda \overline{F}(y) d\overline{T}(y), \qquad (3)$$

Let $R_x(y)$ denote the remaining truncated service time of a job that has attained service y time units,

$$R_x(y) = (S(y) \wedge x) - y, \qquad y < x,$$

where $S(y)$ refers to the service time of a job whose attained service is y time units. For the mean value $\overline{R}_x(y) = E[R_x(y)]$ we have

$$\overline{R}_x(y) = E[(S \wedge x) - y | S > y] = \frac{1}{\overline{F}(y)} \int_y^x \overline{F}(t)\, dt. \qquad (4)$$

In particular, we have

$$\overline{R}_x(0) = E[(S \wedge x)] = \int_0^x \overline{F}(t)\, dt. \qquad (5)$$

From here on we restrict ourselves to such scheduling disciplines that do not use any information about the remaining service times of the jobs, but only information on the attained services. The family of such disciplines is denoted by Π. For such disciplines, $N'(y)$ and $R_x(y)$ (as well as $N(y_i^+) - N(y_i^-)$ and $R_x(y_i)$) are independent and we have the following formula for the mean unfinished truncated work $\overline{U}_x = E[U_x]$:

$$\overline{U}_x = \int_{0^-}^{x^-} \overline{R}_x(y) d\overline{N}(y).$$

By applying (3) and (4), we get

$$\begin{aligned} \overline{U}_x &= \lambda \int_{0^-}^{x^-} \int_y^x \overline{F}(t)\, dt\, d\overline{T}(y) \\ &= \lambda \int_0^x \int_{0^-}^{t^-} d\overline{T}(y)\, \overline{F}(t)\, dt \\ &= \lambda \int_0^x \overline{T}(t^-)\, \overline{F}(t)\, dt. \end{aligned}$$

Since almost everywhere in the interval $[0, x]$, we have $\overline{T}(t^-) = \overline{T}(t)$, the integral remains unchanged if we replace $\overline{T}(t^-)$ by $\overline{T}(t)$. Thus, we have the following result (cf. [2, equation (4.60)]):

$$\overline{U}_x = \lambda \int_0^x \overline{F}(t) \overline{T}(t)\, dt. \qquad (6)$$

We note that the limiting value $\overline{U}_\infty$ with $x \to \infty$ refers to the mean unfinished work that is known to be the same for all work-conserving disciplines. Furthermore, in the limit $x \to \infty$ formula (6) is valid even for more general ergodic arrival processes than Poisson process [19].

Differentiating (6) yields

$$(\overline{U}_x)' = \lambda \overline{F}(x) \overline{T}(x), \qquad (7)$$

where we have used notation $(\overline{U}_x)' = \frac{d}{dx}\overline{U}_x$. Thus,

$$\overline{T}(x) = \frac{(\overline{U}_x)'}{\lambda \overline{F}(x)}. \qquad (8)$$

By averaging over x, we get the following result for the mean delay $E[T]$:

$$E[T] = \int_0^\infty \overline{T}(x) f(x)\, dx = \frac{1}{\lambda} \int_0^\infty (\overline{U}_x)' h(x)\, dx. \qquad (9)$$

2.2 Difference of mean delays

Formula (9) is our main tool in the mean delay comparisons, as the following two propositions reveal.

PROPOSITION 1. *Let $\pi_1, \pi_2 \in \Pi$. If $\overline{U}_x^{\pi_1} \le \overline{U}_x^{\pi_2}$ for all $x \ge 0$ and the hazard rate $h(x)$ is decreasing, then*

$$E[T^{\pi_1}] \le E[T^{\pi_2}].$$

PROOF. If the hazard rate is bounded so that $h(0)$ and $h(\infty)$ exist, then the result can be derived easily as follows. By (9), we have

$$E[T^{\pi_1}] - E[T^{\pi_2}] = \frac{1}{\lambda} \int_0^\infty [(\overline{U}_x^{\pi_1})' - (\overline{U}_x^{\pi_2})'] h(x) dx.$$

After integrating by parts [20, pp. 118-119], we get

$$\begin{aligned} E[T^{\pi_1}] - E[T^{\pi_2}] &= \frac{1}{\lambda}(\overline{U}_\infty^{\pi_1} - \overline{U}_\infty^{\pi_2}) h(\infty) \\ &- \frac{1}{\lambda}(\overline{U}_0^{\pi_1} - \overline{U}_0^{\pi_2}) h(0) - \frac{1}{\lambda} \int_0^\infty (\overline{U}_x^{\pi_1} - \overline{U}_x^{\pi_2})\, dh(x), \end{aligned}$$

where the integral is defined as a Stieltjes integral with respect to the decreasing function $h(x)$. By observing that $\overline{U}_0^{\pi_1} = \overline{U}_0^{\pi_2} = 0$ and $\overline{U}_\infty^{\pi_1} = \overline{U}_\infty^{\pi_2}$, we obtain

$$E[T^{\pi_1}] - E[T^{\pi_2}] = -\frac{1}{\lambda} \int_0^\infty (\overline{U}_x^{\pi_1} - \overline{U}_x^{\pi_2})\, dh(x),$$

from which the result straightly follows.

But if the hazard rate is unbounded near 0, we have to apply Lemma 1 (in Appendix) to $\phi(x) = \overline{U}_x^{\pi_1} - \overline{U}_x^{\pi_2}$ implying that, for all $x \ge 0$,

$$\int_0^x ((\bar{U}_y^{\pi_1})' - (\bar{U}_y^{\pi_2})') h(y)\, dy \le 0.$$

Thus, in the limit $x \to \infty$,

$$\int_0^\infty ((\bar{U}_y^{\pi_1})' - (\bar{U}_y^{\pi_2})') h(y)\, dy \le 0.$$

The claim follows now from (9). □

PROPOSITION 2. *Let $\pi_1, \pi_2 \in \Pi$. If $\overline{U}_x^{\pi_1} \le \overline{U}_x^{\pi_2}$ for all $x \ge 0$ and the hazard rate $h(x)$ is increasing and bounded, then*

$$E[T^{\pi_1}] \ge E[T^{\pi_2}].$$

PROOF. Since the hazard rate is assumed to be bounded, we deduce, as in the first part of the proof of Proposition 1, that

$$E[T^{\pi_1}] - E[T^{\pi_2}] = -\frac{1}{\lambda} \int_0^\infty (\overline{U}_x^{\pi_1} - \overline{U}_x^{\pi_2})\, dh(x),$$

where the integral is defined as a Stieltjes integral with respect to the increasing function $h(x)$. The result follows straightly from this. □

To compare two scheduling disciplines π_1 and π_2 in the mean delay sense, it is, thus, sufficient to prove that $\overline{U}_x^{\pi_1} \le \overline{U}_x^{\pi_2}$ for all $x \ge 0$.

2.3 Framework for comparisons between scheduling disciplines

In this subsection, we develop a mathematical framework that gives a sufficient condition that $\overline{U}_x^{\pi_1} \le \overline{U}_x^{\pi_2}$ for all $x \ge 0$.

Proposition 3. *Let $\pi_1, \pi_2 \in \Pi$. If there exists some $x^* \ge 0$ such that $\overline{T}^{\pi_1}(x) \le \overline{T}^{\pi_2}(x)$ for all $x < x^*$ and $\overline{T}^{\pi_1}(x) \ge \overline{T}^{\pi_2}(x)$ for all $x > x^*$, then $\overline{U}_x^{\pi_1} \le \overline{U}_x^{\pi_2}$ for all $x \ge 0$.*

Proof. It follows from (6) that, for all $x < x^*$,

$$\overline{U}_x^{\pi_1} \le \overline{U}_x^{\pi_2}. \qquad (10)$$

On the other hand, for all $x > x^*$,

$$(\overline{U}_x^{\pi_1})' = \lambda \overline{F}(x)\overline{T}^{\pi_1}(x) \ge \lambda \overline{F}(x)\overline{T}^{\pi_2}(x) = (\overline{U}_x^{\pi_2})'.$$

Since both π_1 and π_2 are work-conserving disciplines, for which the mean unfinished work is equal, we have

$$\overline{U}_\infty^{\pi_1} = \overline{U}_\infty^{\pi_2}.$$

The formulas above together with the fact that $\overline{U}_x$ is a continuous function of x guarantee that, for all $x \ge 0$,

$$\overline{U}_x^{\pi_1} \le \overline{U}_x^{\pi_2},$$

which completes the proof. □

Note that the condition of Proposition 3 seems natural in the context of MLPS disciplines. For instance, let $\pi_1 \in \Pi$ be a scheduling discipline that gives priority to short jobs. If one compares the conditional mean delay of π_1 with respect to PS, one expects there will exist some x^* such that $\overline{T}^{\pi_1}(x) \le \overline{T}^{\mathrm{PS}}(x)$ for all $x < x^*$ and that $\overline{T}^{\pi_1}(x) \ge \overline{T}^{\mathrm{PS}}(x)$ for all $x > x^*$. In Section 2.4 we show that this is indeed the case for $\pi_1 = \mathrm{PS} + \mathrm{PS}$.

2.4 Comparison between PS+PS and PS disciplines

In this subsection we show that the PS+PS(a) discipline is better than PS in reducing the value of $\overline{U}_x$.

It is well known that the mean delay of a job with service time $x > 0$ in a PS system reads as

$$\overline{T}^{\mathrm{PS}}(x) = \frac{x}{1-\rho}.$$

According to [2, equations (4.27), (4.36), (4.39)], the corresponding mean delay in a PS+PS(a) system can be expressed as follows:

$$\overline{T}^{\mathrm{PS+PS}(a)}(x) = \begin{cases} \dfrac{x}{1-\rho_a} & \text{if } x \le a, \\ \overline{T}^{\mathrm{FB}}(a) + \dfrac{\alpha(x-a)}{1-\rho_a} & \text{if } x > a. \end{cases} \qquad (11)$$

Here $\rho_a = \lambda E[S \wedge a]$ refers to the truncated load, and $\alpha(x)$ is such that $\alpha'(x) = \frac{d}{dx}\alpha(x)$ satisfies the following integral equation:

$$\begin{aligned} \alpha'(x) &= \frac{\lambda}{1-\rho_a}\int_0^x \alpha'(y)\overline{F}(a+x-y)\,dy \\ &\quad + \frac{\lambda}{1-\rho_a}\int_0^\infty \alpha'(y)\overline{F}(a+x+y)\,dy + c(x) + 1 \end{aligned}$$

with $c(x) \ge 0$. Note that $\overline{T}^{\mathrm{PS+PS}(a)}(x)$ is differentiable, at least, for all $x > a$.

A numerical example of $\overline{T}^{\pi}(x)$ is depicted in Figure 1 for a bounded Pareto service time distribution BP(k, p, α). A random variable X has a bounded Pareto distribution if its distribution is given by $F(x) = (1-(k/x)^\alpha)/(1-(k/p)^\alpha)$ for all $k \le x \le p$. The hazard rate of a bounded Pareto distribution is decreasing for all $k \le x \le p$.

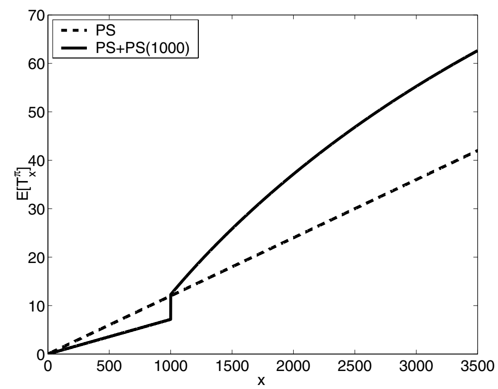

Figure 1: Mean delay $\overline{T}^{\pi}(x)$ of a job with service time x, for $\pi = \{\mathrm{PS}, \mathrm{PS}+\mathrm{PS}(1000)\}$, under service time distribution BP$(13, 3500, 1.1)$ and load $\rho = 0.9$

Denote now

$$\alpha^* = \inf_{x>0} \alpha'(x).$$

The following inequality follows straightforwardly from the integral equation:

$$\begin{aligned} \alpha^* &\ge \inf_{x>0}\Big\{\frac{\lambda}{1-\rho_a}\int_0^x \alpha'(y)\overline{F}(a+x-y)\,dy \\ &\quad + \frac{\lambda}{1-\rho_a}\int_0^\infty \alpha'(y)\overline{F}(a+x+y)\,dy + 1\Big\} \\ &\ge \inf_{x>0}\Big\{\frac{\lambda}{1-\rho_a}\int_0^x \alpha^*\overline{F}(a+x-y)\,dy \\ &\quad + \frac{\lambda}{1-\rho_a}\int_0^\infty \alpha^*\overline{F}(a+x+y)\,dy + 1\Big\} \\ &= \frac{\lambda\alpha^*}{1-\rho_a}\int_0^\infty \overline{F}(a+z)\,dz + 1. \end{aligned}$$

Now, taking into account that

$$\int_0^\infty \overline{F}(a+z)\,dz = E[S] - E[S \wedge a],$$

we deduce that

$$\alpha^* \ge \alpha^* \frac{\rho-\rho_a}{1-\rho_a} + 1.$$

Thus,

$$\alpha^* \ge \frac{1-\rho_a}{1-\rho}. \qquad (12)$$

By combining (11) and (12), we get

$$\begin{cases} (\overline{T}^{\mathrm{PS+PS}(a)})'(x) = \dfrac{1}{1-\rho_a} \le \dfrac{1}{1-\rho} & \text{if } x < a, \\ (\overline{T}^{\mathrm{PS+PS}(a)})'(x) \ge \dfrac{\alpha^*}{1-\rho_a} \ge \dfrac{1}{1-\rho} & \text{if } x > a. \end{cases} \qquad (13)$$

Note further that

$$(\overline{T}^{\mathrm{PS}})'(x) = \frac{1}{1-\rho}. \qquad (14)$$

PROPOSITION 4. *For all $a \geq 0$ and $x \geq 0$,*

$$\overline{U}_x^{\mathrm{PS+PS}(a)} \leq \overline{U}_x^{\mathrm{PS}}.$$

PROOF. We verify that $\overline{T}^{\mathrm{PS+PS}(a)}(x)$ and $\overline{T}^{\mathrm{PS}}(x)$ satisfy the assumptions stated in Proposition 3.

Let $a \geq 0$. Since $\overline{T}^{\mathrm{PS}}(0) = \overline{T}^{\mathrm{PS+PS}(a)}(0) = 0$, it follows from (13) and (14) that, for all $x < a$,

$$\overline{T}^{\mathrm{PS+PS}(a)}(x) \leq \overline{T}^{\mathrm{PS}}(x).$$

Since PS+PS is a work-conserving discipline, there exists a value

$$x^* = \inf\{x \geq a \mid \overline{T}^{\mathrm{PS+PS}(a)}(x) \geq \overline{T}^{\mathrm{PS}}(x)\}.$$

Thus, for all $x < x^*$.

$$\overline{T}^{\mathrm{PS+PS}(a)}(x) \leq \overline{T}^{\mathrm{PS}}(x).$$

On the other hand by equations (13) and (14) we have, for all $x > x*$

$$\overline{T}^{\mathrm{PS+PS}(a)}(x) \geq \overline{T}^{\mathrm{PS}}(x),$$

which completes the proof. □

A numerical example of $\overline{U}_x^{\pi}$ is depicted in Figure 2 for a bounded Pareto service time distribution $\mathrm{BP}(13, 3500, 1.1)$.

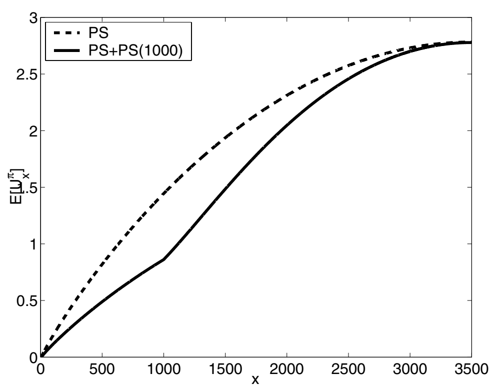

Figure 2: Mean unfinished truncated work $\overline{U}_x^{\pi}$ of those jobs with attained service less than x, for $\pi = \{\mathrm{PS}, \mathrm{PS+PS}(1000)\}$, under service time distribution $\mathrm{BP}(13, 3500, 1.1)$ and load $\rho = 0.9$

Combining Propositions 1 and 4, we deduce that PS+PS is better than PS in the mean delay sense under condition DHR. This result gives a clue that the two-level packet scheduling algorithm referred to in Section 1 is a reasonable method to favor short flows in the Internet.

It is even possible to consider the pathwise difference $U_x^{\pi_1}(t) - U_x^{\pi_2}(t)$, from which relationships for the difference $\overline{U}_x^{\pi_1} - \overline{U}_x^{\pi_2}$ follow. The pathwise arguments in the following section are based on the rate $\sigma_x^{\pi}(t)$ at which the unfinished truncated work is reduced.

3. PATHWISE COMPARISON OF UNFINISHED TRUNCATED WORK

In this section we relax the assumption of Poisson arrivals and concentrate on the pathwise comparison of the unfinished truncated work with a fixed truncation threshold $x > 0$. Therefore, we add the time dimension into our considerations.

Consider any work-conserving discipline π. Let $A(t)$ denote the number jobs who have arrived up to time t, $N^{\pi}(t)$ the number of jobs at time t, and $\mathcal{N}^{\pi}(t)$ the set of jobs in the system at time t, where $|\mathcal{N}^{\pi}(t)| = N^{\pi}(t)$. Furthermore, let $N_x^{\pi}(t)$ denote the number of such jobs with attained service less than x and $U_x^{\pi}(t)$ the unfinished truncated work at time t. Note that

$$N_x^{\pi}(t) = 0 \quad \Leftrightarrow \quad U_x^{\pi}(t) = 0.$$

Finally, let A_i denote the arrival time of job i, S_i its service time, and $X_i^{\pi}(t)$ its attained service at time t. Then we have the following relationship:

$$i \in \mathcal{N}^{\pi}(t) \quad \Leftrightarrow \quad A_i \leq t \text{ and } X_i^{\pi}(t) < S_i.$$

Furthermore,

$$U_x^{\pi}(t) = \sum_{i=1}^{A(t)} ((S_i \wedge x) - (X_i^{\pi}(t) \wedge x)),$$

or, equivalently,

$$\begin{aligned} U_x^{\pi}(t) &= \sum_{i \in \mathcal{N}^{\pi}(t)} ((S_i \wedge x) - (X_i^{\pi}(t) \wedge x)) \\ &= \sum_{i \in \mathcal{N}_x^{\pi}(t)} ((S_i \wedge x) - X_i^{\pi}(t)). \end{aligned}$$

On the other hand, we have

$$U_x^{\pi}(t) = \sum_{i=1}^{A(t)} (S_i \wedge x) - \int_0^t \sigma_x^{\pi}(u)\, du, \tag{15}$$

where $\sigma_x^{\pi}(t)$ refers to the rate at which such jobs that have attained service less than x units are served at time t. Certainly, we have, for any t,

$$\begin{cases} \sigma_x^{\pi}(t) = 0, & \text{if } N_x^{\pi}(t) = 0, \\ \sigma_x^{\pi}(t) \leq 1, & \text{if } N_x^{\pi}(t) > 0. \end{cases} \tag{16}$$

For example, for the PS discipline this rate is as follows:

$$\sigma_x^{\mathrm{PS}}(t) = \begin{cases} 0, & \text{if } N_x^{\mathrm{PS}}(t) = 0, \\ \dfrac{N_x^{\mathrm{PS}}(t)}{N^{\mathrm{PS}}(t)} \leq 1, & \text{if } N_x^{\mathrm{PS}}(t) > 0. \end{cases} \tag{17}$$

Note further that the first term in (15) is independent of the scheduling discipline.

3.1 Optimal discipline to minimize unfinished truncated work

Let Π_x^* denote the family of work-conserving scheduling disciplines that give full priority to jobs that have attained service less than x units. Note that

$$\mathrm{FB},\ \mathrm{FB+PS}(x),\ \mathrm{PS+PS}(x) \in \Pi_x^*.$$

Furthermore, $\mathrm{FB+PS}(a) \in \Pi_x^*$ for all $a \geq x$. For any $\pi^* \in \Pi_x^*$, we have

$$\sigma_x^{\pi^*}(t) = \begin{cases} 0, & \text{if } N_x^{\pi^*}(t) = 0, \\ 1, & \text{if } N_x^{\pi^*}(t) > 0. \end{cases} \tag{18}$$

It follows that $U_x^{\pi^*}(t)$ is the same for all $\pi^* \in \Pi_x^*$. It is also easy to show that any $\pi^* \in \Pi_x^*$ is optimal in this sense.

PROPOSITION 5. *For any $\pi^* \in \Pi_x^*$ and $t \geq 0$,*

$$U_x^{\pi^*}(t) = \min_{\pi \in \Pi} U_x^{\pi}(t).$$

PROOF. This follows immediately from equations (15), (16) and (18) combined with the observation that if $\sigma_x^{\pi}(t) > \sigma_x^{\pi^*}(t)$, then necessarily $N_x^{\pi^*}(t) = 0$, which implies that $U_x^{\pi^*}(t) = 0 \leq U_x^{\pi}(t)$. □

Instead of $U_x^{\pi}(t)$, both Righter et al. [6] and Feng and Misra [9] consider the corresponding untruncated random variable

$$V_x^{\pi}(t) = \sum_{i \in \mathcal{N}_x^{\pi}(t)} (S_i - X_i^{\pi}(t)).$$

Under the condition that $\overline{V}_x^{\pi 1} \leq \overline{V}_x^{\pi 2}$ for all $x \geq 0$, Righter et al. [6] prove that $E[T^{\pi_1}] \leq E[T^{\pi_2}]$ when the service time distribution is of type NWUE; under the same condition, Feng and Misra [9] obtain the same result when the hazard rate is decreasing.

In order to prove the optimality of FB, both Righter et al. [6] and Feng and Misra [9], use a similar pathwise argument as above to demonstrate that FB minimizes $V_x^{\pi}(t)$ in the pathwise sense.

However, if the remaining service times are not truncated, there is a vertical downward jump in $V_x^{\pi}(t)$ amounting to $S_i - x$ time units at time t when job i jumps out of $\mathcal{N}_x^{\pi}(t)$. This fact is not taken into account in the proofs of FB's optimality in [6, Lemma (3.5)] and [9, Lemma 2.4][4]. Therefore, while FB minimizes $U_x^{\pi}(t)$ (as seen above in Proposition 5), FB does not minimize $V_x^{\pi}(t)$ nor $\overline{V}_x^{\pi}$.

Numerical examples of $\overline{V}_x^{\pi}$ and $\overline{U}_x^{\pi}$ for a bounded Pareto service time distribution $BP(13, 3500, 1.1)$ are depicted in Figures 3 and 4, respectively. The former one demonstrates the non-optimality of FB regarding the unfinished untruncated work, while the latter one illustrates the optimality regarding the unfinished truncated work.

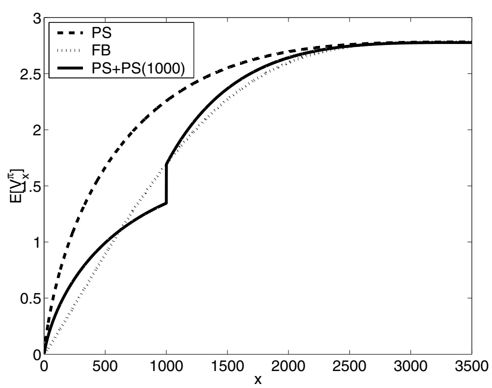

Figure 3: Mean unfinished untruncated work $\overline{V}_x^{\pi}$ for $\pi = \{\mathrm{PS}, \mathrm{FB}, \mathrm{PS} + \mathrm{PS}(1000)\}$ under service time distribution $\mathrm{BP}(13, 3500, 1.1)$ and load $\rho = 0.9$

[4]After a private communication Feng and Misra corrected this deficiency independently utilizing the same definition as here of truncated remaining service time, see *http://www1.cs.columbia.edu/ misra/pubs/mama.pdf*

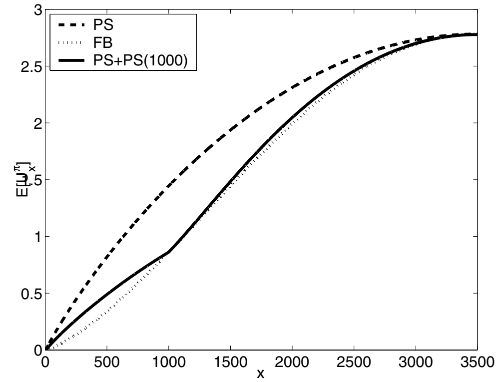

Figure 4: Mean unfinished truncated work $\overline{U}_x^{\pi}$ for $\pi = \{\mathrm{PS}, \mathrm{FB}, \mathrm{PS} + \mathrm{PS}(1000)\}$ under service time distribution $\mathrm{BP}(13, 3500, 1.1)$ and load $\rho = 0.9$

3.2 Comparison between TLPS disciplines

Consider now the TLPS disciplines FB+PS(a) and PS+PS(a). We show that the former one is better than the latter one in minimizing U_x.

First we consider the TLPS disciplines FB+PS(a) and PS+PS(a) for which $a \geq x$.

PROPOSITION 6. *For any $a \geq x$ and $t \geq 0$,*

$$U_x^{\mathrm{FB}}(t) = U_x^{\mathrm{FB+PS}(a)}(t) \leq U_x^{\mathrm{PS+PS}(a)}(t).$$

PROOF. The result is due to Proposition 5 and the fact that FB and FB+PS(a) belong to π_x^* in this case. □

Next we consider the TLPS disciplines FB+PS(a) and PS+PS(a) for which $a \leq x$. For the former one, we have

$$\sigma_x^{\mathrm{FB+PS}(a)}(t) = \begin{cases} \dfrac{N_x^{\mathrm{FB+PS}(a)}(t)}{N^{\mathrm{FB+PS}(a)}(t)}, & \text{if } N_a^{\mathrm{FB+PS}(a)}(t) = 0, \\ 1, & \text{if } N_a^{\mathrm{FB+PS}(a)}(t) > 0, \end{cases} \tag{19}$$

and for the latter one

$$\sigma_x^{\mathrm{PS+PS}(a)}(t) = \begin{cases} \dfrac{N_x^{\mathrm{PS+PS}(a)}(t)}{N^{\mathrm{PS+PS}(a)}(t)}, & \text{if } N_a^{\mathrm{PS+PS}(a)}(t) = 0, \\ 1, & \text{if } N_a^{\mathrm{PS+PS}(a)}(t) > 0. \end{cases} \tag{20}$$

It follows that in this case we have, for any $t \geq 0$,

$$\sigma_x^{\mathrm{FB+PS}(a)}(t) = \sigma_x^{\mathrm{PS+PS}(a)}(t). \tag{21}$$

PROPOSITION 7. *For any $a \leq x$ and $t \geq 0$,*

$$U_x^{\mathrm{FB}}(t) \leq U_x^{\mathrm{FB+PS}(a)}(t) = U_x^{\mathrm{PS+PS}(a)}(t).$$

PROOF. The first inequality follows from Proposition 5 and the fact that FB $\in \Pi_x^*$. The second equality is due to equations (15) and (21). □

We note that in the particular case $x = a$, there is no difference between the three disciplines,

$$U_a^{\mathrm{FB}}(t) = U_a^{\mathrm{FB+PS}(a)}(t) = U_a^{\mathrm{PS+PS}(a)}(t).$$

We further note that, while PS+PS is better than PS regarding the mean value of the unfinished truncated work (see Proposition 4), the pathwise version of this result is not true, which we verified by a counter-example.

3.3 Comparison inside MLPS disciplines

In this subsection, we consider the multilevel PS disciplines $\mathrm{MLPS}(a_1, \ldots, a_N)$, where

$$0 = a_0 < a_1 < \ldots < a_N < a_{N+1} = \infty.$$

For any $\pi \in \mathrm{MLPS}(a_1, \ldots, a_N)$, let $\pi_n \in \{\mathrm{FB}, \mathrm{PS}\}$ denote the scheduling discipline used at level n, where $n \in \{1, \ldots, N+1\}$.

First we observe that $\pi \in \mathrm{MLPS}(a_1, \ldots, a_N)$ is (locally) optimal, i.e., $\pi \in \Pi_x^*$, if there is $n \in \{1, \ldots, N+1\}$ such that

(i) $a_n = x$, or

(ii) $a_{n-1} < x < a_n$ and $\pi_n = \mathrm{FB}$.

The reason is that, for all these disciplines π and for all $t \geq 0$, we have

$$\sigma_x^\pi(t) = \sigma_x^{\mathrm{FB}}(t).$$

Next we compare two MLPS disciplines. For that, define the following order relation among the disciplines $\{\mathrm{FB}, \mathrm{PS}\}$:

$$\mathrm{FB} \preceq \mathrm{FB}, \quad \mathrm{FB} \preceq \mathrm{PS}, \quad \mathrm{PS} \not\preceq \mathrm{FB}, \quad \mathrm{PS} \preceq \mathrm{PS}.$$

Proposition 8. *Assume that* $\pi \in \mathrm{MLPS}(a_1, \ldots, a_N)$, $\pi' \in \mathrm{MLPS}(a'_1, \ldots, a'_{N'})$, $n \in \{1, \ldots, N+1\}$ *and* $n' \in \{1, \ldots, N'+1\}$ *such that*

$$a_{n-1} = a'_{n'-1} \leq x \leq a_n = a'_{n'}.$$

(i) If $\pi_n = \pi'_{n'}$, *then* $U_x^\pi(t) = U_x^{\pi'}(t)$ *for all* $t \geq 0$.

(ii) If $\pi_n \preceq \pi'_{n'}$, *then* $U_x^\pi(t) \leq U_x^{\pi'}(t)$ *for all* $t \geq 0$.

Proof. (i) Assume that $\pi_n = \pi'_{n'}$. It is easy to see that $\sigma_x^\pi(t) = \sigma_x^{\pi'}(t)$ for all $t \geq 0$ in this case. Thus, by (15), we have $U_x^\pi(t) = U_x^{\pi'}(t)$ for all $t \geq 0$.

(ii) Due to (i), it is sufficient to consider the case $\pi_n = \mathrm{FB}$ and $\pi'_{n'} = \mathrm{PS}$. In this case,

$$\sigma_x^\pi(t) = \begin{cases} 0, & \text{if } N_x^\pi(t) = 0, \\ 1, & \text{if } N_x^\pi(t) > 0, \end{cases}$$

while

$$\begin{cases} \sigma_x^{\pi'}(t) = 0, & \text{if } N_x^{\pi'}(t) = 0, \\ \sigma_x^{\pi'}(t) \leq 1, & \text{if } N_x^{\pi'}(t) > 0. \end{cases}$$

for all $t \geq 0$. Thus, by (15), we have $U_x^\pi(t) \leq U_x^{\pi'}(t)$ for all $t \geq 0$. □

Note that the results related to the comparison of the TLPS disciplines FB+PS(a) and PS+PS(a) presented in Propositions 6 and 7 follow from this one.

4. MEAN DELAY COMPARISONS OF MLPS DISCIPLINES

As mentioned in section 2, in cases where the hazard rate is monotonous, the order of the scheduling disciplines is determined by the meanwise difference $E[U_x^{\pi_1} - U_x^{\pi_2}]$. The following theorems related to the mean delay comparisons between scheduling disciplines follow from the propositions in the previous sections.

Theorem 1. *Assume that the hazard rate* $h(x)$ *is decreasing.*

(i) $E[T^{\mathrm{FB}}] = \min_{\pi \in \Pi} E[T^\pi]$.

(ii) For all $a \geq 0$,

$$E[T^{\mathrm{FB}}] \leq E[T^{\mathrm{FB+PS}(a)}] \leq E[T^{\mathrm{PS+PS}(a)}] \leq E[T^{\mathrm{PS}}].$$

(iii) For any $\pi, \pi' \in \mathrm{MLPS}(a_1, \ldots, a_N)$ *such that* $\pi_n \preceq \pi'_n$ *for all* $n \in \{1, \ldots, N+1\}$,

$$E[T^\pi] \leq E[T^{\pi'}].$$

Proof. (i) The claim follows immediately from Propositions 1 and 5.[5]

(ii) The claim follows immediately from Propositions 1, 4, 6 and 7.

(iii) The claim follows immediately from Propositions 1 and 8. □

This theorem strengthens the intuition that the multilevel packet scheduling algorithm referred to in Section 1 is a reasonable method to favor short flows in the Internet.

Theorem 2. *Assume that the hazard rate* $h(x)$ *is increasing and bounded.*

(i) $E[T^{\mathrm{FB}}] = \max_{\pi \in \Pi} E[T^\pi]$.

(ii) For all $a \geq 0$,

$$E[T^{\mathrm{FB}}] \geq E[T^{\mathrm{FB+PS}(a)}] \geq E[T^{\mathrm{PS+PS}(a)}] \geq E[T^{\mathrm{PS}}].$$

(iii) For any $\pi, \pi' \in \mathrm{MLPS}(a_1, \ldots, a_N)$ *such that* $\pi_n \preceq \pi'_n$ *for all* $n \in \{1, \ldots, N+1\}$,

$$E[T^\pi] \geq E[T^{\pi'}].$$

Proof. (i) The claim follows immediately from Propositions 2 and 5.

(ii) The claim follows immediately from Propositions 2, 4, 6 and 7.

(iii) The claim follows immediately from Propositions 2 and 8. □

4.1 Numerical comparison

In this subsection we present some numerical examples that illustrate the benefits of PS+PS(a) and FB+PS(a) in reducing the mean delay with respect to PS.

In Figure 5, we depict the mean delay of the disciplines PS+PS(a) and FB+PS(a) as a function of the threshold a. In addition the mean delay values for the PS and FB disciplines are also depicted. The service time distribution is bounded Pareto $\mathrm{BP}(13, 3500, 1.1)$.

We first note that since the hazard rate of a bounded Pareto distribution is decreasing, the inequalities stated in Theorem 1 are indeed satisfied. As expected, we observe that

$$E[T^{\mathrm{FB+PS}(0)}] = E[T^{\mathrm{PS+PS}(0)}] = E[T^{\mathrm{PS}}]$$

and that

$$\lim_{a \to \infty} E[T^{\mathrm{FB+PS}(a)}] = E[T^{\mathrm{FB}}]$$

whereas

$$\lim_{a \to \infty} E[T^{\mathrm{PS+PS}(a)}] = E[T^{\mathrm{PS}}].$$

[5]Optimality of FB is not a new result [6]. We present it here for the paper to be self-contained.

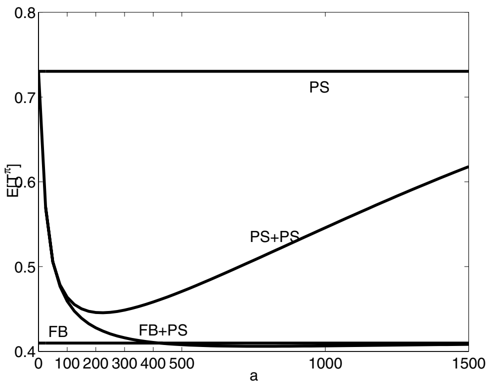

Figure 5: Mean delay $E[T^\pi]$, for $\pi = \{\mathrm{PS}, \mathrm{PS}+\mathrm{PS}(a), \mathrm{FB}+\mathrm{PS}(a), \mathrm{FB}\}$, as a function of a under service time distribution $\mathrm{BP}(13, 3500, 1.1)$ and load $\rho = 0.9$

Interestingly, we observe that choosing appropriately the value of the threshold a, $\mathrm{PS}+\mathrm{PS}(a)$ can significantly reduce the mean delay value with respect to PS (up to 40% of reduction). Similarly, with the same optimal choice of the threshold a, $\mathrm{PS}+\mathrm{PS}(a)$ can compete with FB in respect to the mean delay.

Another positive performance characteristic of $\mathrm{PS}+\mathrm{PS}(a)$ stems from the apparent insensitivity properties with respect to the value of the threshold. For instance, the mean delay of $\mathrm{PS}+\mathrm{PS}(a)$ does not vary significantly in the range $100 \le a \le 300$ ($F(100) = 0.9$ and $F(300) = 0.97$). This effect is even more pronounced when more heavy-tailed distributions are considered (for example unbounded Pareto).

5. CONCLUSIONS

We have shown, in the mean delay sense, the order between PS, FB and TLPS scheduling disciplines, FB+PS and PS+PS, under given job size distributions. The analysis considered general distributions, based on the characteristics of their hazard rate. The given proofs are based on pathwise and meanwise arguments. Using similar techniques, we showed the order between disciplines within any level of an MLPS discipline.

Our results imply that, under the current Internet traffic mix, the better the scheduling policy takes into account the attained service of the job, the smaller the mean delay of the system is. Regarding implementations one must consider how this is done on the packet level. As proposed by [7] and [12], this can be done by assuming that the attained service can be inferred from the packet count or byte count (sequence number) of the flow. However, implementing a packet level mechanism based on FB, would require a strict ordering of each byte or packet of each flow at the core router's, e.g., infinite number of priority classes or full flow level information at all routers. On the other hand for two-level scheduling, such as PS+PS, only division of packets or bytes into two priority classes at the edge of the network is required.

Topics of further research include studying more general MLPS disciplines including the quantitative gain between changing a discipline within a level of an MLPS discipline. Since the optimality of FB and SRPT holds under general job arrival processes, one would expect that PS+PS would be better than PS even when the Poisson assumption is removed. Other further research includes studying the order of these MLPS scheduling disciplines in terms of other metrics than mean delay, e.g., slowdown or throughput, and generalizing the results to distributions with weaker conditions than DHR, such as NWUE.

APPENDIX

LEMMA 1. *Let ϕ be a real-valued function on $[0,\infty)$ and h a non-negative function on $(0,\infty)$. Assume that ϕ is differentiable, $\phi(0) = 0$, and $\phi(x) \le 0$ for all $x \ge 0$. Assume further that h is decreasing. Then, for any $x \ge 0$,*

$$\int_0^x \phi'(y)h(y)dy \le 0.$$

PROOF. Note first that, for all $x \ge 0$,

$$\int_0^x \phi'(y)dy = \phi(x) \le 0. \tag{22}$$

Furthermore, since $\phi(0) = 0$ and $\phi(x) \le 0$ for all $x \ge 0$, there exists x_0 such that $\phi'(x) \le 0$ for all $x < x_0$. Define now $A_1^{\mathrm{b}} = 0$, and then recursively, for all $k = 1, 2, \ldots$,

$$A_k^{\mathrm{e}} = B_k^{\mathrm{b}} = \inf\{x > A_k^{\mathrm{b}} \mid \phi'(x) > 0\}$$
$$B_k^{\mathrm{e}} = A_{k+1}^{\mathrm{b}} = \inf\{x > B_k^{\mathrm{b}} \mid \phi'(x) \le 0\}.$$

The x-axis is thus divided into intervals $A_1, B_1, A_2, B_2, \ldots$, where $A_k = [A_k^{\mathrm{b}}, A_k^{\mathrm{e}})$ and $B_k = [B_k^{\mathrm{b}}, B_k^{\mathrm{e}})$. Note further that, for all $x \in (A_k^{\mathrm{b}}, A_k^{\mathrm{e}})$,

$$\phi'(x) \le 0, \tag{23}$$

and, for all $x \in (B_k^{\mathrm{b}}, B_k^{\mathrm{e}})$,

$$\phi'(x) > 0. \tag{24}$$

Furthermore, denote, for all k,

$$\Delta_k = \int_0^{B_k^{\mathrm{e}}} \phi'(y)dy, \tag{25}$$

$$I_k = \int_0^{B_k^{\mathrm{e}}} \phi'(y)h(y)dy. \tag{26}$$

1° Consider first the intervals A_1 and B_1. By (23), we have, for all $x \in [A_1^{\mathrm{b}}, A_1^{\mathrm{e}}]$,

$$\int_0^x \phi'(y)h(y)dy \le 0.$$

Let then $x \in [B_1^{\mathrm{b}}, B_1^{\mathrm{e}}]$. Now, since h is decreasing and $A_1^{\mathrm{e}} = B_1^{\mathrm{b}}$, we have

$$\begin{aligned}\int_0^x \phi'(y)h(y)dy &= \int_0^{A_1^{\mathrm{e}}} \phi'(y)h(y)dy + \int_{B_1^{\mathrm{b}}}^x \phi'(y)h(y)dy \\ &\le h(B_1^{\mathrm{b}}) \int_0^x \phi'(y)dy \le 0.\end{aligned}$$

The last inequality is due to (22). Thus, the claim is true in these first intervals A_1 and B_1. In particular, we have

$$I_1 \le h(B_1^{\mathrm{b}})\Delta_1 \le 0.$$

2° Let then $k > 1$ and assume that

$$I_{k-1} \le h(B_{k-1}^{\mathrm{b}})\Delta_{k-1} \le 0. \tag{27}$$

By (27) and (23), we have, for all $x \in [A_k^{\mathrm{b}}, A_k^{\mathrm{e}}]$,

$$\int_0^x \phi'(y)h(y)dy = I_{k-1} + \int_{A_k^{\mathrm{b}}}^x \phi'(y)h(y)dy \le 0.$$

Let then $x \in [B_k^{\mathrm{b}}, B_k^{\mathrm{e}}]$. Similarly as in 1°, since h is decreasing and $A_k^{\mathrm{e}} = B_k^{\mathrm{b}}$, we have

$$\begin{aligned}\int_{A_k^{\mathrm{b}}}^x \phi'(y)h(y)dy &= \int_{A_k^{\mathrm{b}}}^{A_k^{\mathrm{e}}} \phi'(y)h(y)dy + \int_{B_k^{\mathrm{b}}}^x \phi'(y)h(y)dy \\ &\le h(B_k^{\mathrm{b}}) \int_{A_k^{\mathrm{b}}}^x \phi'(y)dy.\end{aligned}$$

Thus, by (27), (22) and the fact that $h(B_{k-1}^{\mathrm{b}}) \ge h(B_k^{\mathrm{b}})$, we have

$$\begin{aligned}\int_0^x \phi'(y)h(y)dy &= I_{k-1} + \int_{A_k^{\mathrm{b}}}^x \phi'(y)h(y)dy \\ &\le h(B_{k-1}^{\mathrm{b}})\Delta_{k-1} + h(B_k^{\mathrm{b}}) \int_{A_k^{\mathrm{b}}}^x \phi'(y)dy \\ &\le h(B_k^{\mathrm{b}}) \int_0^x \phi'(y)dy \le 0.\end{aligned}$$

Thus, the claim is true in these intervals A_k and B_k. In particular, we have

$$I_k \le h(B_k^{\mathrm{b}})\Delta_k \le 0,$$

which completes the proof. □

A. REFERENCES

[1] L.E. Schrage, "A proof of the optimality of the shortest remaining processing time discipline," *Operations Research*, vol. 16, no. 3, pp. 687–690, 1968.

[2] L. Kleinrock, *Queueing Systems, vol. 2*, John Wiley and Sons, 1976.

[3] S.F. Yashkov, "Processor-sharing queues: Some progress in analysis," *Queueing Systems*, vol. 2, pp. 1–17, 1987.

[4] S.F. Yashkov, "On feedback sharing a processor among jobs with minimal serviced length (in Russian)," *Technika Sredstv Svyazi. Ser. ASU*, no. 2, pp. 51–62, 1978.

[5] R. Righter and J.G. Shanthikumar, "Scheduling multiclass single server queueing systems to stochastically maximize the number of successful departures," *Probability in the Engineering and Informational Sciences*, vol. 3, pp. 323–333, 1989.

[6] R. Righter, J.G. Shanthikumar, and G. Yamazaki, "On external service disciplines in single-stage queueing systems," *Journal of Applied Probability*, vol. 27, pp. 409–416, 1990.

[7] L. Guo and I. Matta, "Differentiated control of web traffic: A numerical analysis," in *Proceedings of SPIE ITCOM'2002: Scalability and Traffic Control in IP Networks*, Boston, MA, 2002.

[8] A. Wierman, N. Bansal, and M. Harchol-Balter, "A note on comparing response times in the M/G/1/FB and M/G/1/PS queues," *Operations Research Letters*, vol. 32, no. 1, pp. 73–76, 2004.

[9] H. Feng and V. Misra, "Mixed scheduling disciplines for network flows," *ACM SIGMETRICS Perform. Eval. Rev. Special issue on MAMA 2003*, vol. 31, no. 2, 2003, pp. 36–39.

[10] I. Rai, G. Urvoy-Keller, and E. Biersack, "Analysis of LAS scheduling for job size distributions with high variance," in *Proceedings of ACM SIGMETRICS*, San Diego, CA, 2003, pp. 218–228.

[11] A. Wierman and M. Harchol-Balter, "Classifying scheduling policies with respect to unfairness in an M/GI/1," in *Proceedings of ACM SIGMETRICS*, San Diego, CA, 2003, pp. 238–249.

[12] K. Avrachenkov, U. Ayesta, P. Brown, and E. Nyberg, "Differentiation between short and long TCP flows: predictability of the response time," in *Proceedings of IEEE Infocom 2004*, Hong Kong, 2004.

[13] K. Thompson, G.J. Miller, and R. Wilder, "Wide-area Internet traffic patterns and characteristics," *IEEE Network*, vol. 11, no. 6, pp. 10–23, 1997.

[14] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, and C. Diot, "Packet-level traffic measurement from the Sprint IP backbone," *IEEE Network*, vol. 17, no. 6, pp. 6–16, 2003.

[15] W. Noureddine and F. Tobagi, "Improving the performance of interactive TCP applications using service differentiation," *Computer Networks Journal*, vol. 40, no. 1, pp. 19–43, Sept. 2002.

[16] M. Alzati, M. Bottigliengo, C. Casetti, and M. Mellia, "Fair marking of web flows using partial state information," in *Proceedings of International Teletraffic Congress (ITC-18)*, Berlin, 2003, pp. 851–860.

[17] D.P. Heyman, T.V. Lakshman, and A.L. Neidhardt, "A new method for analysing feedback-based protocols with applications to engineering Web traffic over the Internet," in *Proceedings of ACM SIGMETRICS*, Seattle, WA, 1997, pp. 24–38.

[18] S. Ben Fredj, T. Bonald, A. Proutiere, G. Regnie, and J. Roberts, "Statistical bandwidth sharing: A study of congestion at flow level," in *Proceedings of ACM SIGCOMM*, 2001, pp. 111–122.

[19] T.M. O'Donovan, "Distribution of attained service and residual service in general queueing systems," *Operations Research*, vol. 22, pp. 570–575, 1974.

[20] F. Riesz and B. Szökefalvi-Nagy, *Functional Analysis*, Ungar, 1955.