

Comparison of PS, SRPT, and FB Scheduling Policies under Exponential and Heavy Tail Job Sizes

Idris A. Rai, Guillaume Urvoy-Keller, Ernst W. Biersack
Technical Report, TR-01.11.1
Institut Eurecom
2229, route des Crêtes
06904 Sophia-Antipolis, France
{rai,urvoy,erbi}@eurecom.fr

Abstract

We present numerical analysis that compares the performances of PS, SRPT, and FB scheduling policies in terms of their conditional mean response times in an $M/G/1$ model, where G represents an exponential job size distribution or a job size distribution with highly varying sizes such as heavy tailed distribution. The results show that for heavy tailed job size distribution, FB offers mean slowdown close to SRPT and a small percentage of the largest jobs experiences a higher slowdown under FB than under PS.

1 Definitions

Assume a job size distribution with probability mass function (pmf) $f(x)$, and let the average arrival rate be λ . The second moment of the distribution is given by $m_2(x) = \int_0^x t^2 f(t) dt$ and the load of jobs with sizes less than or equal to x is given as $\rho(x) = \lambda \int_0^x t f(t) dt$. Mean response time for a job of size x , $E[T(x)]$, is the sum of its average residence time, $E[R(x)]$, and average waiting time, $E[W(x)]$, i.e.,

$$E[T(x)] = E[W(x)] + E[R(x)]. \quad (1)$$

The formulas of average waiting time and average residence time of a job size x for the shortest remaining processing time (SRPT) and foreground-background (FB) a.k.a shortest elapsed time (FB) scheduling policies are given by [4, 3] below. The terms FB and FB are used interchangeably in the report.

- Average waiting time for a job of size x

$$E[W(x)]_{FB} = \frac{\lambda(m_2(x) + x^2(1 - F(x)))}{2(1 - \rho(x) - \lambda x(1 - F(x)))^2} \quad (2)$$

$$E[W(x)]_{SRPT} = \frac{\lambda(m_2(x) + x^2(1 - F(x)))}{2(1 - \rho(x))^2} \quad (3)$$

- Average residence time for a job of size x

$$E[R(x)]_{FB} = \frac{x}{1 - \rho(x) - \lambda x(1 - F(x))} \quad (4)$$

$$E[R(x)]_{SRPT} = \int_0^x \frac{1}{1 - \rho(t)} d(t) \quad (5)$$

The mean response time for the processor sharing policy is given as:

$$E[T(x)]_{PS} = \frac{x}{1 - \rho} \quad (6)$$

Mean slowdown of a job of size x denoted as $E[S(x)]$ is defined as $E[S(x)] = \frac{E[T(x)]}{x}$. Unless otherwise indicated, whenever we say response time, $E[T(x)]$, or slowdown, $E[S(x)]$, we mean the expected (mean) response time or slowdown conditioned to a job of size x .

Finally, x is called the 100β percentile for a given random variable X with cumulative distribution function $F(x)$ if the function $Q(\beta)$ is such that

$$Q(\beta) = \inf_{x \in X} \{F(x) \geq \beta\}.$$

In this report, we evaluate and compare the PS, FB and SRPT scheduling policies in terms of their slowdown as a function of job size, system load, and percentiles of job sizes for jobs with exponential and heavy tail job size distributions.

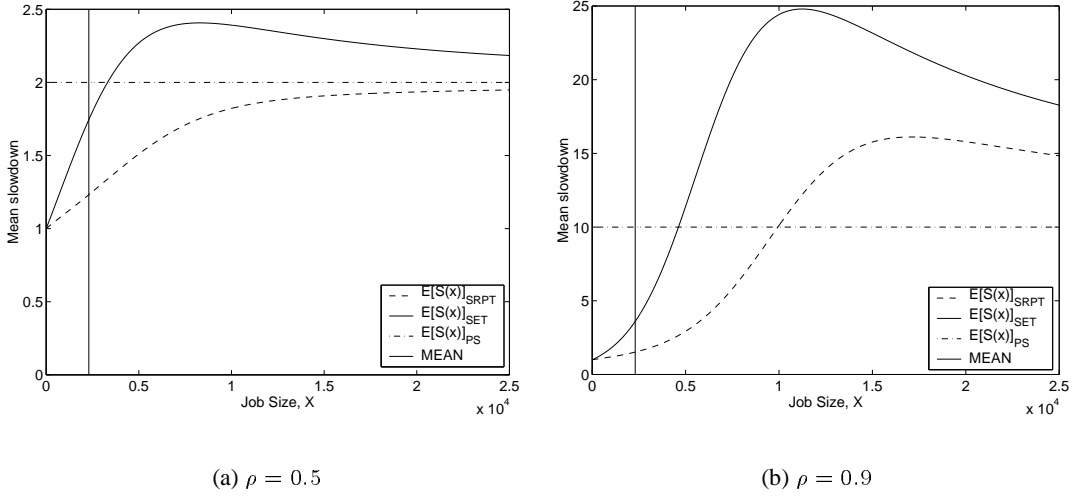


Figure 1: Expected slowdown for $Exp(\frac{1}{\mu} = 2.3 * 10^3)$ as a function of job size

2 Evaluation of scheduling policies under exponentially distributed job sizes

The probability density function of exponential distribution with mean $1/\mu$ is given as:

$$f(x) = \mu e^{-\mu x}, \text{ for } 0 \leq x \leq \infty \quad (7)$$

In this section, we compare performances of the scheduling policies in terms of slowdown for exponentially distributed job sizes. Figures 1(a) and 1(b) show the slowdown of the policies as a function of job size when load is $\rho = 0.5$ and $\rho = 0.9$ respectively. We observe from the figures that the slowdown is not a monotonic function. Short jobs have a lower slowdown under FB and SRPT than under PS. This is in accordance with the definitions of SRPT and FB policies that they favor short jobs. PS has the smallest slowdown for large jobs at high system load $\rho = 0.9$. In general, FB yields higher slowdown than SRPT for all job sizes. In particular, the slowdown is higher under FB than under SRPT by up to a factor of four as seen in Figure 2(b).

At load $\rho = 0.5$, SRPT has the lowest slowdown as seen in Figures 1(a) and 2(a). Similarly, large jobs have lower slowdown under PS than under FB. However, their slowdown ratio under these large jobs at load $\rho = 0.5$ is close to their ratio at load $\rho = 0.9$. FB is quite close to SRPT at load $\rho = 0.5$ except for a few jobs, just above the mean, which have maximum slowdown under FB about 50% higher than under SRPT as seen in Figure 2(a).

Figures 3(a) and 3(b) show the slowdown as a function of load for the 50th and 99th percentile jobs of the distribution respectively. For the 50th percentile job, SRPT and FB have much lower slowdown values than PS, particularly at load $\rho = 0.9$, whereas the 99th percentile job has higher slowdown under FB

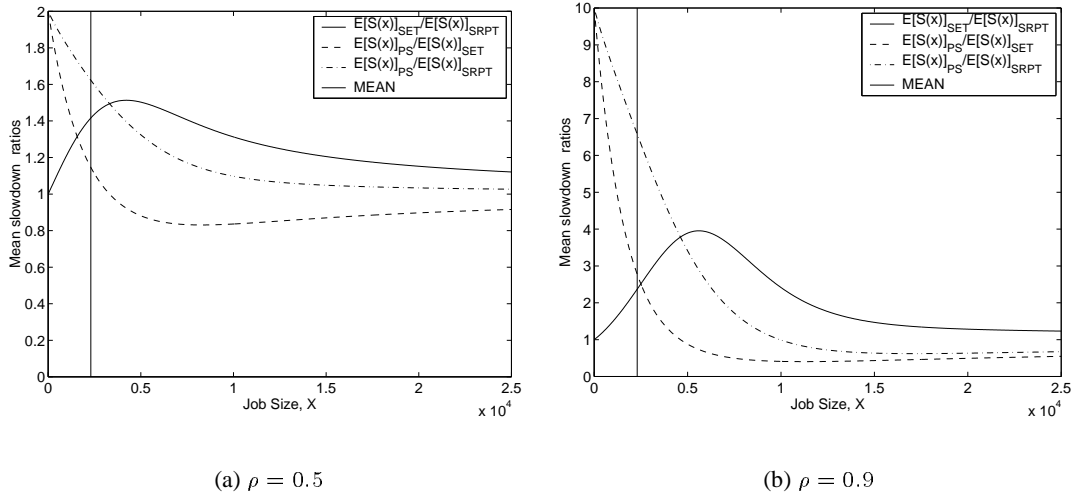


Figure 2: Ratios of expected slowdown for $Exp(\frac{1}{\mu} = 2.3 * 10^3)$ as a function of job size

than under SRPT and PS. It can also be observed from Figures 3(b) and 4(a) that PS and SRPT perform about the same for the 99th percentile job except when the system load approaches 1. At that point the slowdown of PS goes to infinity. This is in accordance with the definition of SRPT that it is stable for all jobs as long as the system load is less than 1.0 as mentioned in [1]. It can also be observed from Figure 3(b) that there exist system load values at which the 99th percentile job has lower slowdown under PS than under SRPT. This observation is also noted in [1].

For the 50th percentile job of the distribution, we observe that the slowdown of FB does not grow to infinity when load approaches 1 as was the case for the 99th percentile job. Under PS, both jobs have their slowdown values increase to infinity when the system load approaches 1. This may be explained by the fact that PS uses very short quantum values to impose fairness among jobs, which increases the average residence time of the 50th percentile job. On the other hand, the same job completes service in a few quanta under FB. The relation between the slowdown of the policies can further be seen in Figures 4(a) and 4(b).

In Figures 5(a) and 5(b), the slowdown is plotted against the percentile of the job size distribution for load $\rho = 0.5$ and $\rho = 0.9$ respectively. At load $\rho = 0.9$, more than 10% of the largest jobs are penalized under FB, i.e., have higher slowdown than under PS, as compared to about 2% under SRPT. At load $\rho = 0.5$, no job is penalized under SRPT whereas about 20% of the largest jobs have higher slowdown under FB as compared to PS. However, the slowdown values at load $\rho = 0.9$ and $\rho = 0.5$ differ greatly. The maximum slowdown of FB, for example, at load $\rho = 0.9$ is about 10 times its maximum value at load $\rho = 0.5$. Finally, looking at Figure 6, we again see that SRPT outperforms FB by up to a factor of four when load $\rho = 0.9$. The difference between FB and SRPT is much less pronounced at load $\rho = 0.5$.

In summary, at load $\rho = 0.9$, FB and SRPT outperform PS in terms of slowdown for small jobs, whereas

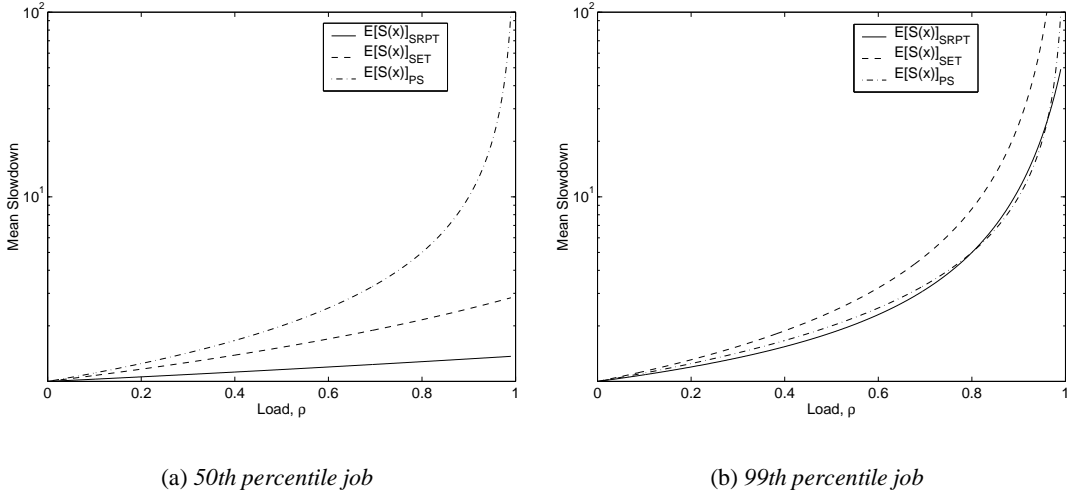


Figure 3: Expected slowdown for $Exp(\frac{1}{\mu} = 2.3 * 10^3)$ as a function of load

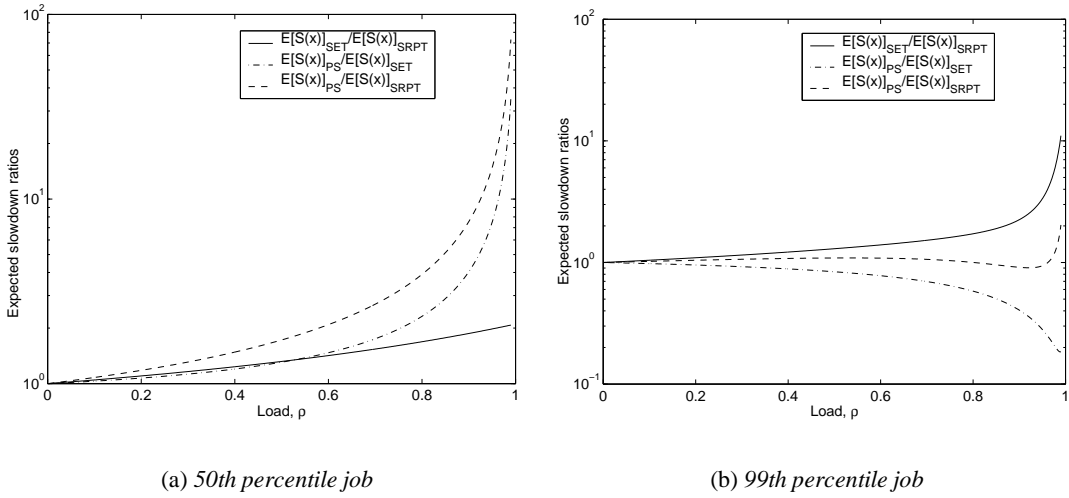


Figure 4: Ratios of expected slowdown for $Exp(\frac{1}{\mu} = 2.3 * 10^3)$ as a function of load

large jobs have the smallest slowdown under PS. At load $\rho = 0.5$, all jobs maintain the smallest slowdown under SRPT.

Large jobs in the 99th percentile have the highest slowdown under FB at any system load. Like PS, the slowdown of these jobs under FB seems to grow to infinity as load reaches 1. In general, FB and SRPT have similar slowdowns for small jobs at any load, while their slowdowns differ greatly for large jobs as load approaches 1.

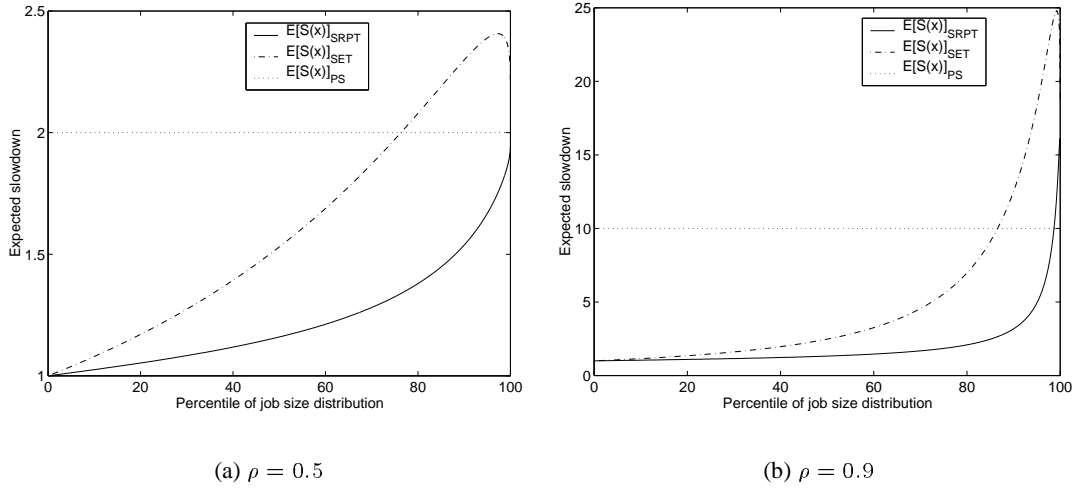


Figure 5: Expected slowdown for $Exp(\frac{1}{\mu} = 2.3 * 10^3)$ as a function of percentiles of job size distribution

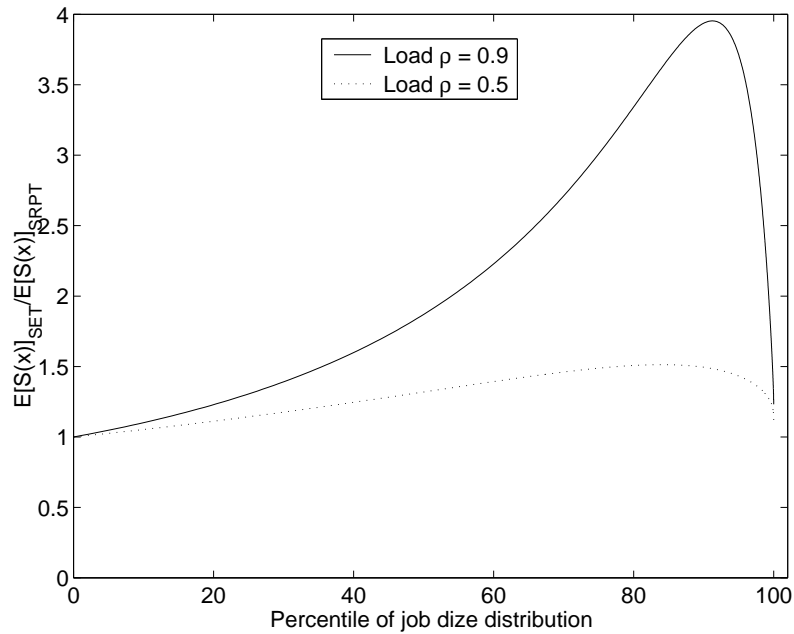


Figure 6: The $\frac{E[S(x)]_{FB}}{E[S(x)]_{SRPT}}$ for $Exp(\frac{1}{\mu} = 2.3 * 10^3)$ as a function of percentiles of job size distribution

3 Heavy Tail Distribution

Job sizes are said to follow heavy tail distribution when more than half of the load is constituted in the 1% largest jobs. The tails of heavy tail distributions exhibit power law. In contrast to the exponential distribution, the tail of the heavy tail distribution diminishes very slowly. The Pareto distribution is an example of a heavy tail distribution. Given a random variable X , the probability mass function of Pareto

distribution is given in Equation (8) below;

$$f(x) = a^\alpha \alpha x^{-\alpha-1} \text{ for } x \geq a \text{ and } 0 \leq \alpha \leq 2. \quad (8)$$

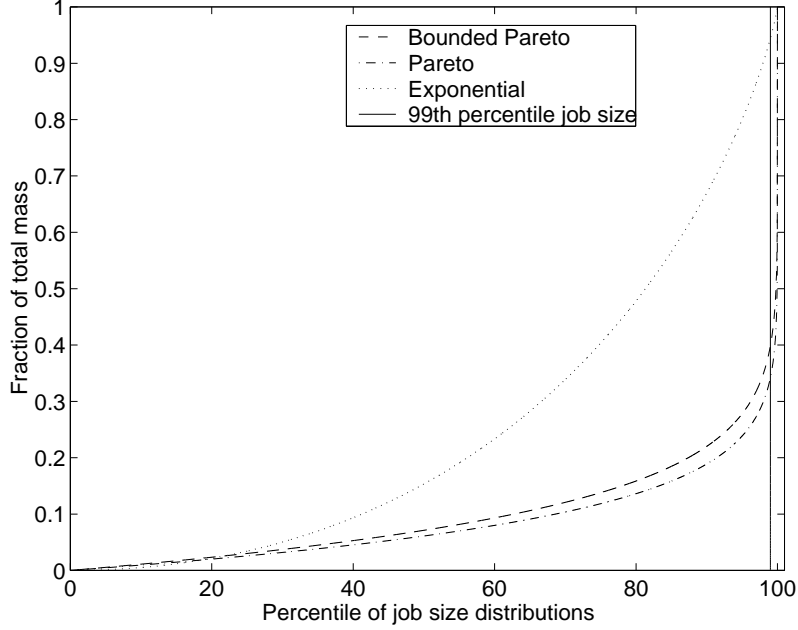


Figure 7: The fraction of total mass for $BP(332, 10^{10}, 1.1)$, *pareto* with $\alpha = 1.1$, and *exponential* distribution as a function of percentiles of the distributions

In practice however, there is an upper bound on the job sizes. Therefore, a *Bounded Pareto Distribution* (BP) that is commonly represented as $BP(a, p, \alpha)$ (where a and p are the minimum and maximum job sizes respectively) is used in this report to evaluate the policies. The pmf of the BP is given as,

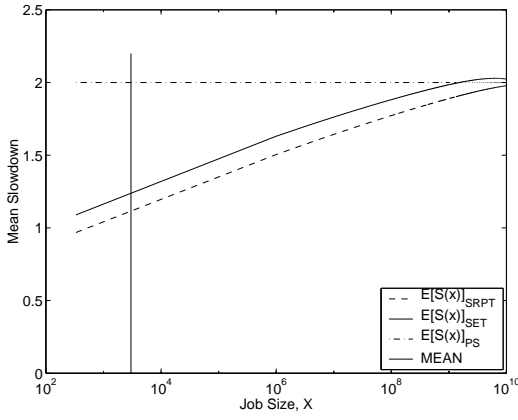
$$f(x) = \frac{a^\alpha \alpha}{(1 - (a/p)^\alpha)} x^{-\alpha-1} \text{ for } a \leq x \leq p \text{ and } 0 \leq \alpha \leq 2. \quad (9)$$

Observe that the bounded pareto is an approximation of a heavy tail distribution. The approximation is more valid for large values of P . This can be seen by using the mass-weighted distribution, $F_w(x)$, defined in [2] as:

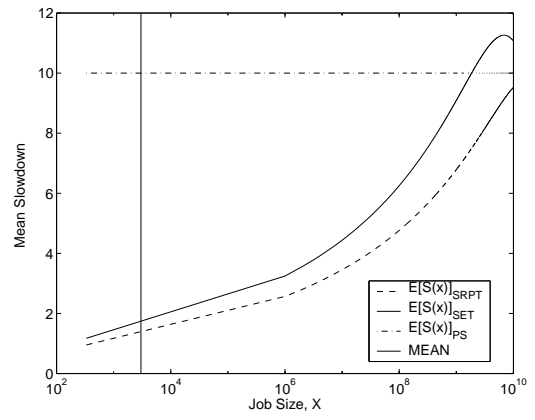
$$F_w(x) = \frac{\int_{-\infty}^x u dF(u)}{\int_{-\infty}^{\infty} v dF(v)} \quad (10)$$

If plotted against $F(x)$, this function yields a fraction of total mass contained in the fraction of jobs with size less than or equal to x :

Figure 7 shows the mass-weighted distribution function, $F_w(x)$, plotted against cumulative distribution function, $F(x)$. We observe that the fraction of total mass comprised by 1% the largest jobs for $BP(332, 10^{10}, 1.1)$ is a bit less than that 1% largest jobs of Pareto distribution comprises. As can be seen, $BP(332, 10^{10}, 1.1)$ approximates heavy tail distribution well since about 60% of the total mass is

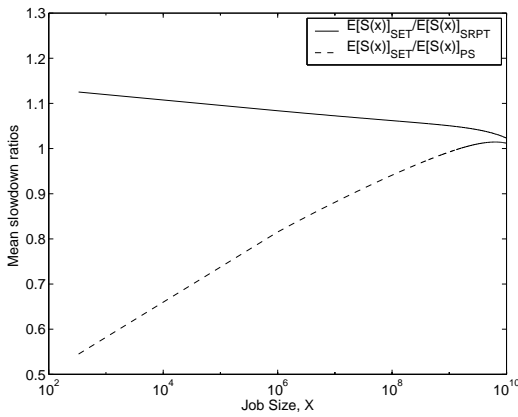


(a) $\rho = 0.5$

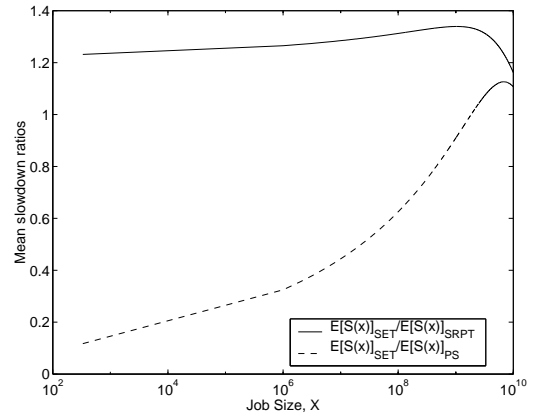


(b) $\rho = 0.9$

Figure 8: Expected slowdown for $BP(332, 10^{10}, 1.1)$ as a function of job size



(a) $\rho = 0.5$



(b) $\rho = 0.9$

Figure 9: Ratios of expected slowdown of the policies for $BP(332, 10^{10}, 1.1)$ as a function of job size

due to the 1% largest jobs. The 1% largest jobs for exponential distribution constitute less than 10% of the load.

In the next section, we present and comment on the evaluation of the PS, FB, and SRPT scheduling policies for $BP(332, 10^{10}, 1.1)$ with mean job size of 3000.

3.1 Evaluation of scheduling policies under heavy tail job size distribution

As for the exponential distribution, we use the slowdown and the slowdown ratio between scheduling policies as a performance metric to analyze and compare the scheduling policies under heavy tail job sizes.

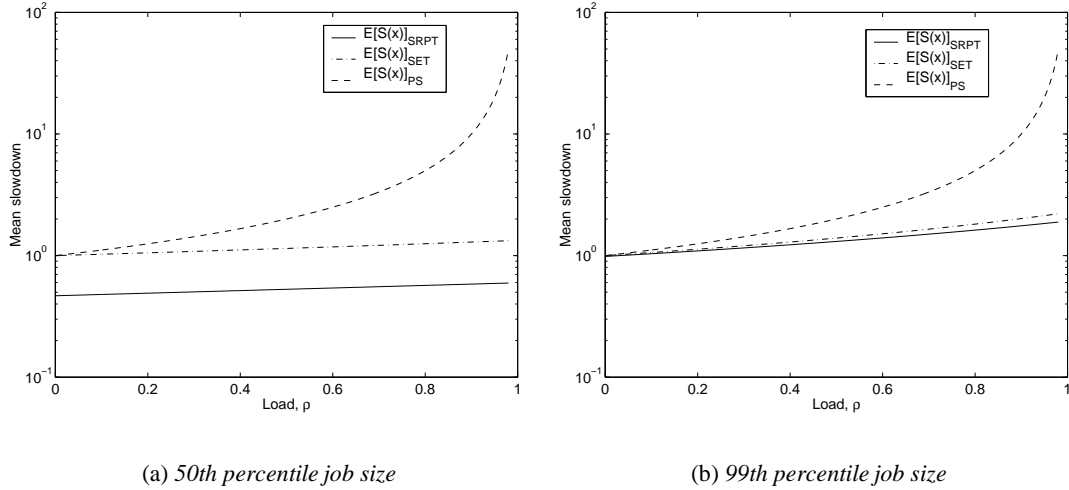


Figure 10: Mean slowdown for $BP(332, 10^{10}, 1.1)$ as a function of load

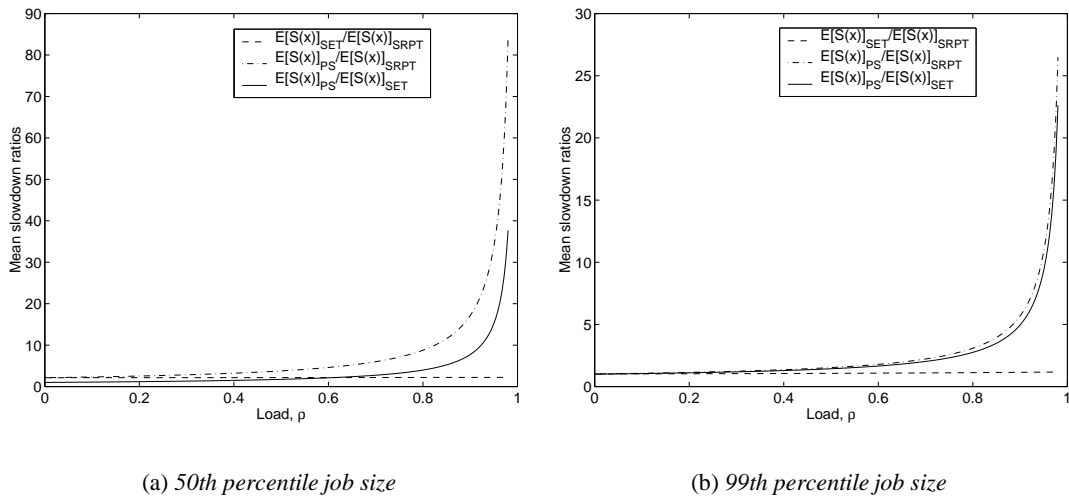


Figure 11: Ratios of mean slowdown for $BP(332, 10^{10}, 1.1)$ as a function of load

In all considered cases, it can be noted from the figures that SRPT offers the smallest slowdown to most job sizes, see Figures 8(a), 8(b), 12(a), and 12(b). Observe from the figures that the slowdowns of large

jobs under SRPT asymptotically approach to that of PS irrespective of the system load.

FB outperforms PS in terms of slowdown except for a few largest jobs as can be seen in Figures 8(a), 8(b), and 12(b). Looking at these figures, one notes that small jobs experience very low slowdowns under FB whereas a few large jobs have a higher slowdown under FB than under PS. This is due to the fact that FB favors short jobs. We see in Figures 10(a) and 10(b) that FB offers, independent of the load, a lower slowdown than PS to jobs in the 99th percentile and the 50th percentile.

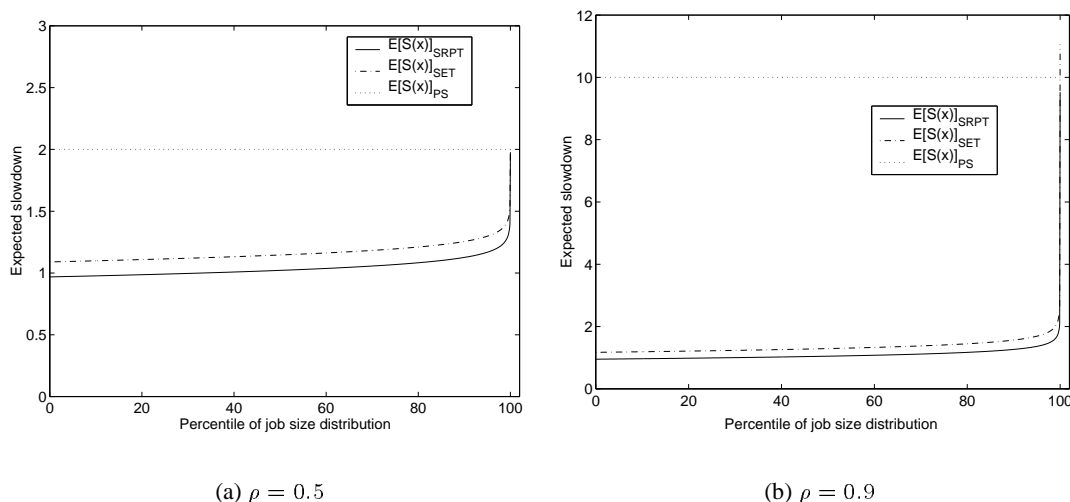


Figure 12: Mean slowdown for $BP(332, 10^{10}, 1.1)$ as a function of percentiles of job size distribution

The general conclusion is that the performance of FB in terms of slowdown is always quite close to that of SRPT. As can be seen in Figures 9(b), 11(a), 11(b), and 13, the slowdown ratio between FB and SRPT is close to 1. In particular, the slowdown of job sizes under FB is no more than 40% higher than under SRPT at load $\rho = 0.9$ and no more than 14% higher at load $\rho = 0.5$. Furthermore, the slowdown ratio between FB and SRPT is almost constant for any load and percentiles.

Finally, observe that SRPT yields lower slowdown than PS to all job sizes and system load values. FB also shows similar results except for the largest 1% of the jobs.

4 Conclusion

The evaluation of the scheduling policies reveals that FB and SRPT have better performance in terms of their slowdown under heavy tail jobs than under exponential job sizes. The figures show that the worst slowdown values of FB and SRPT are higher than that of PS for exponential job sizes. In addition, at load $\rho = 0.9$, FB penalizes a much larger percentage of the largest jobs in case of exponential jobs sizes; more than 12%, compared to 1% of the largest jobs in case of heavy tail job sizes as seen in Figures 5(b) and 12(b). Furthermore, we note from Figure 5(a) that at load of $\rho = 0.5$, more large jobs, about 25%,

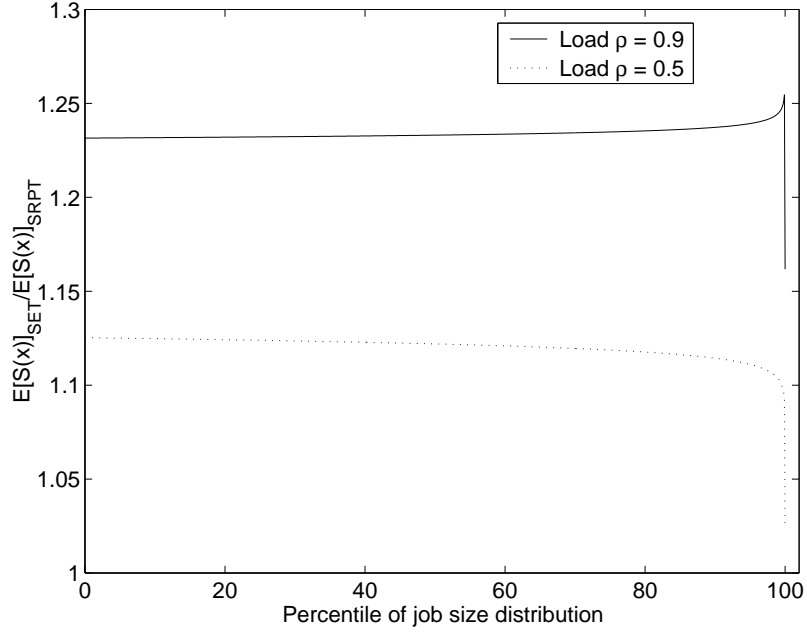


Figure 13: Mean slowdown for $BP(332, 10^{10}, 1.1)$ as a function of percentiles job of size distribution

have higher slowdown under FB than under PS. This is not the case under heavy tail job sizes at the same load as seen in Figure 12(b).

Large heavy tail jobs have lower slowdown under SRPT than under PS regardless of the system load whereas for exponential job sizes and load $\rho = 0.9$, about 2% of the largest jobs have higher slowdown under SRPT than under PS as seen in Figure 5(b).

The slowdown ratio between FB and SRPT is quite small for heavy tail jobs, less than 1.5, whereas it is close to 4 for exponential job sizes for all evaluation parameters considered. In particular, at load $\rho = 0.9$, the worst slowdown value of FB is less than 40% larger than the worst slowdown of SRPT for heavy tail jobs whereas it is close to 400% larger for exponential jobs, see Figures 2(b), 6, 9(b), and 13.

In addition to the mentioned performance benefits FB and SRPT offer to heavy tail job sizes, it is worth noting small jobs that have lower slowdown under these policies than under PS are more for $BP(332, 10^{10}, 1.1)$ job sizes than for exponential job sizes. This supports the facts that heavy tail distribution constitutes many short jobs and that these policies indeed favor those short jobs.

Finally, the results show that FB and SRPT perform quite close in terms of slowdown of heavy tail job sizes. Therefore using FB to give preference to short jobs results to similar performance improvements as for SRPT if jobs sizes are heavy tailed. That is, the response time of short jobs under FB is greatly reduced without severe penalty for the largest jobs. Moreover, the slowdown of large heavy tail jobs under FB does not grow to infinity very fast as load $\rho \rightarrow 1$ as opposed to large exponential jobs as seen in Figures 3(b) and 10(b). In contrary, observe in the figures that the slowdowns of large jobs of both

distributions under PS grow to infinity as load $\rho \rightarrow 1$.

References

- [1] N. Bansal and M. Harchol-Balter, “Analysis of SRPT Scheduling: Investigating Unfairness”, In *Sigmetrics 2001 / Performance 2001*, pp. 279–290, June 2001.
- [2] M. Crovella, “Performance Evaluation of Heavy-Tailed Distributions”, In *7th Ann. Workshop on Job Scheduling Strategies for Parallel Processing*, pp. 1–9, 2001.
- [3] L. E. Schrage, “The queue M/G/1 with feedback to lower priority queues”, *Management Science*, 13(7):466–474, 1967.
- [4] L. E. Schrage and L. W. Miller, “The queue M/G/1 with the shortest processing remaining time discipline”, *Operations Research*, 14:670–684, 1966.