Asymptotic Performance of the Non-Forced Idle Time Scheduling Policies in the Presence of Variable Demand for Resources

Ana Radovanović and Cliff Stein

Abstract

A fundamental scheduling problem is to schedule a set of jobs on a set of machines so that as many jobs as possible can be scheduled while respecting the resource constraints on each machine. A further complication in many computer systems is that scheduling decisions must be made on-line, that is, as soon as job arrives it must either be scheduled or rejected. Many practical systems and algorithms schedule jobs in such a way that loads on machines tend to be balanced. However, in case when job requests are highly variable, such scheduling systems do not necessarily result in a good performance. If loads are balanced then when large jobs arrive, there may not be any "holes" large enough to hold them, thereby resulting in increased queueing delays and overall job sojourn times in the system. In this paper, we focus on addressing the job assignment problem in an on-line loss system framework, i.e., when a job arrives, it gets scheduled to a particular machine using a certain scheduling policy or, if there are not enough available resources on any of the machines, the incoming job is lost. This type of scheduling is often called Non-Forced Idle Time Scheduling and includes many common policies, including Best Fit, First Fit and Worst Fit.

In this paper, we use compound point processes to capture stochastic variability in the request process, we derive simple asymptotic estimates for the job loss rate in these scheduling systems. Furthermore, we derive the asymptotic lower bound for the job loss rate by comparing it to the performance of a single large machine and derive an asympotic average case competitive ratio for the class of analyzed Non-forced Idle Time Scheduling policies. Although our proofs are asymptotic, we perform experiments which show an excellent match between simulation results and theoretical performance bounds, even for relatively small resource capacities and large values of measured loss rates.

Index Terms

Machine scheduling, Loss networks, Subexponential distributions.

I. Introduction

In order to effectively run jobs on multiple machines, such as a cluster of workstations, it is important to manage the resources well. Jobs must be assigned to machines in an intelligent manner so as to efficiently manage resources, and to leave free resources for jobs that arrive in the future. Poor job assignment results in an inability of machines to provide enough resources to newly arriving jobs, which leads to increased queueing delays and a smaller thoughput. In this paper, we model such a system as a loss system, where jobs are either scheduled or lost. We analyze online scheduling policies that make scheduling decisions based on the current amount of available resources on machines and the size of an incoming job. Each job arrives with a certain processing requirement and it is scheduled as long as there is at least one machine that has sufficient available capacity to satisfy this demand. These type of scheduling policies are known as Non-Forced Idle Time (NFIT) Scheduling. Furthermore, we assume that requests for resources arrive according to some point process in time. In case a job is scheduled on a machine, it stays in the system for some random amount of time after which it leaves the system and frees the resources that it had used. We evaluate our policies by considering the job loss rate, i.e., the long run proportion of lost jobs due to inability of the system to satisfy their requirements at the moment of their arrivals.

The online M-machine scheduling problem is one of the most widely-studied problems in computer science (see, for example, [12] and references therein). For a more recent survey on on-line scheduling policies an interested reader is refered to [6]. Many variants of this problem have been addressed in both worst-case and average-case frameworks, with many different objectives. Much of the previous work focusses on deterministic settings. In off-line deterministic settings, all job parameters are known in advance, whereas in on-line settings, job parameters

Google, Inc.
Cliff Stein is in the Operations Research Department, Columbia University.

are known only when jobs arrive. In the stochastic framework, users of resources arrive according to some random process in time, and stay in the system for unknown, random time. Much of this analysis arose in connection-level models for circuit-switched networks and Internet congestion control, where randomly varying number of flows share links of a network. The analytic approaches mostly focus on estimating fluid and diffusion approximations of various performance variables of interest in specific regimes such as under, critically and over loaded (utilized) systems. An interested reader is referred to a more recent work in [11], as well as one of foundations in [7] and references therein. Many on-line scheduling algorithms tend to balance loads, and there are explicit policies that, in the case where loads on machines get unbalanced, use job migration to keep the system balanced (see [1] and references therein). Our goal in this paper is to show that in the case when job resource requirements exhibit large stochastic variability, and when the goal is to minimize the probability that a newly arrived job does not find enough available resources on any of the machines to satisfy its requirements, load balancing do not necessarily perform better than any of the NFIT scheduling policy.

In this paper we analyze the above described system in the presence of renewal arrivals and random resource requirements. In particular, we assume that the request arrivals follow a compound renewal process, with the corresponding holding times being arbitrarily distributed with finite mean. In addition, we assume that the holding times corresponding to different arrivals are mutually independent and independent of arrival points as well. In order to cope with variability in resource requirements, we model them as truncated subexponential random variables where job sizes can not exceed a certain large value. We develop a simple estimate for the asymptotic job loss rate of NFIT scheduling policies when the maximum job requirement and machine sizes grow large. In particular, we relate the job loss rate to the stationary distribution of the largest jobs in a single machine system with unlimited capacity and a specific constraint on the number of largest jobs it can store. This result is stated in Theorem 1. Due to space limitations of this extended abstract, we omit some of the details of its proof. Furthermore, in Theorem 2, we develop an asymptotic lower bound for the performance (job loss rate) of NFIT scheduling policies and show that it can be reached under certain assumptions on machine capacities and job requirement distributions. A direct implication of our results is the explicit asympotic formula for the performance ratio between the M-machine scheduling system and a large single machine loss system with the same total capacity. One possible and somewhat surprising implication of our results is that they can be used to estimate the stationary distribution for specific cases of circuit switched (loss) networks with renewal arrivals and generally distributed holding times. Although our analysis is asymptotic, our numerical experiment illustrates the applicability of our results even for relatively small capacities and relatively large values of the loss rates.

II. MODEL DESCRIPTION

Let requests for resources of the same type arrive at time points $\{\tau_n, -\infty < n < \infty\}$ that represent a renewal process with rate $0 < \lambda < \infty$, i.e., $\mathbb{E}[\tau_n - \tau_{n-1}] = 1/\lambda$. At each point τ_n, B_n amount of resources is requested and then scheduled to one of $M < \infty$ machines. Scheduling decisions are made based on the current system's state, i.e., amount of free resources on each machine, and an incoming job's resource requirement. If the available capacity on each machine is less than B_n, this job is lost; otherwise, B_n amount of resources will be occupied for a random amount of time θ_n, $\mathbb{E}\theta_n < \infty$, after which they are released and become available again for future requests. Sequences $\{B_n\}$, $\{\theta_n\}$ of i.i.d. random variables are assumed to be independent of arrival points $\{\tau_n\}$ and from each other.

In order to model highly variable, but still bounded, resource requirements of incoming jobs, the main results in this paper hold under certain assumptions on the distribution function of B_n. We assume that B_n, $-\infty < n < \infty$, can not exceed C (maximum resource requirement of a job), and that it can be represented as $B_n = X_n \wedge C$, where $x \wedge y \triangleq \min(x, y)$ and X_n belongs to the class of subexponential distribution functions, defined as follows:

Definition 1: Let $\{X_i\}$ be a sequence of positive i.i.d. random variables with distribution function F such that $F(x) < 1$ for all $x > 0$. Denote $\bar{F}(x) = 1 - F(x)$, $x \geq 1$, the tail of F and $\bar{F}^{n*} = 1 - F^{n*}(x) = \mathbb{P}[X_1 + \cdots + X_n > x]$, the tail of the n-fold convolution of F. F is subexponential distribution function $F \in \mathcal{S}$ if one of the following equivalent conditions holds:

- $\lim_{x\to\infty} \frac{\bar{F}^{n*}(x)}{\bar{F}(x)} = n$ for all $n \geq 2$,
- $\lim_{x\to\infty} \frac{\mathbb{P}[X_1+\cdots+X_n>x]}{\mathbb{P}[\max(X_1,\ldots,X_n)>x]} = 1$ for all $n \geq 2$.

Subexponential distributions belong to the class of long-tailed distributions, by definiton introduced by Chistyakov in [3],

$$\lim_{x\to\infty} \frac{1-F(x-y)}{1-F(x)} = 1, \text{ for all } y \in \mathbb{R}. \tag{1}$$

For a brief introduction to subexponential distributions the reader is referred to a recent survey [5]. This class of distributions is fairly large and a well known examples incorporate regularly varying (in particular Pareto), some Weibull, Log-normal and "almost" exponential distributions.

Next, let $\mathcal{N}_{\pi,n}^{(m)}$, $1 \le m \le M$, be a set of indices $i < n$ of resource requirements B_i that are requested prior to τ_n, are scheduled by scheduling policy π on machine $1 \le m \le M < \infty$ (found enough available resources on machine m), and are *still active* at time τ_n. Furthermore, define $N_{\pi,n}^{(m)} \triangleq |\mathcal{N}_{\pi,n}^{(m)}|$, $1 \le m \le M$, to be the number of elements in the set $\mathcal{N}_{\pi,n}^{(m)}$. Thus, the total number of resources that a request at time τ_n finds engaged on machine m is $Q_{\pi,n}^{(m)}$, and can be expressed as $Q_{\pi,n}^{(m)} = \sum_{i\in\mathcal{N}_{\pi,n}^{(m)}} B_i$. Furthermore, we assume that machine m has resource capacity $C_m = k_m C + f_m(C)$, $1 \le m \le M$, where $k_m \in \mathbb{N}_0$, and $\limsup_{C\to\infty} f_m(C)/C \to 0$ as $C \to \infty$, and that there exists some large enough constant $K < \infty$ such that $|\max_m f_m(C) - \min_m f_m(C)| < K$ (we comment on possible relaxations of this condition later). Descriptively, the previous assumption on machine capacities states that each machine can store at most some nonegative integer number of jobs of the largest size C and that the leftover capacities $f_m(C)$, $1 \le m \le M$, increase in C in a sub-linear fashion and do not differ much from each other.

Our goal in this paper is to estimate the stationary loss rate of a scheduling policy π, belonging to a class of policies that are called Non-Forced Idle Time Policies (Non-FITP). A policy π is Non-FITP if scheduling decisions at the moment of a job arrival τ_n are based on the state of all machines $(Q_{\pi,n}^{(1)}, \dots, Q_{\pi,n}^{(M)})$, an incoming demand B_n and a loss can happen only in the case when there is no machine with enough available resources to satisfy requirements of an incoming job, i.e.,

$$L_\pi^{(C)} \triangleq \mathbb{P}\left[\cap_{m=1}^{M}\{Q_{\pi,n}^{(m)} + B_n > C_m\}\right]. \tag{2}$$

The proof of existence of the stationary loss rate $L_\pi^{(C)}$ is technically involved and does not provide any new insights on system's performance and, for that reason, we omit it in this extended abstract. The basic idea is to construct a Markov process with general state space of which $(Q_{\pi,n}^{(1)}, \dots, Q_{\pi,n}^{(M)})$ is a functional, and prove existence of its stationary distribution.

In this paper we are using the following standard notation. For any two real functions $a(t)$ and $b(t)$ and fixed $t_0 \in \mathbb{R}\cup\{\infty\}$ we will use $a(t) \sim b(t)$ as $t \to t_0$ to denote $\lim_{t\to t_0}[a(t)/b(t)] = 1$. Similarly, we say that $a(t) \gtrsim b(t)$ as $t \to t_0$ if $\liminf_{t\to t_0} a(t)/b(t) \ge 1$; $a(t) \lesssim b(t)$ has a complementary definition.

III. Asymptotic Job Loss Rate of Non-forced Idle Time Policies

In this section we estimate the stationary loss rate $L_\pi^{(C)}$ when resource capacity C grows large. Let $N_{\pi,n}^{(m)}(C) \triangleq \sum_{i\in\mathcal{N}_{\pi,n}^{(m)}} 1[B_i = C]$ represent the number of jobs of size C on machine m at the moment of nth arrival τ_n. Our main results in this paper relate the stationary distribution of the number of the largest jobs in the system $N_{\pi,n}^{(m)}(C)$ to the stationary distribution of the number of the largest jobs in system with unlimited capacity, say $N^\infty(C)$, that accepts all jobs with requirements smaller than C, but maintains a limit of maximum $\sum_m k_m$ for the number of the largest jobs (size C) it stores.

Then we state our main result.

Theorem 1: Let $\{X_n, -\infty < n < \infty\}$ be subexponential random variables with finite mean. Then, if $B_n = X_n \wedge C$, $-\infty \le n \le \infty$, the stationary loss rate $L_\pi^{(C)}$ satisfies

$$L_\pi^{(C)} \sim \mathbb{P}\left[\sum_{m=1}^{M} N_{\pi,n}^{(m)}(C) = \sum_m k_m\right] \mathbb{P}\left[X_n > \max_m f_m(C)\right] \text{ as } C \to \infty, \tag{3}$$

where

$$\mathbb{P}\left[\sum_{m=1}^{M} N_{\pi,n}^{(m)}(C) = \sum_m k_m\right] \sim \mathbb{P}\left[N_n^\infty(C) = \sum_m k_m\right] \text{ as } C \to \infty. \tag{4}$$

Proof: First, to prove the asymptotic upper bound for $L_\pi^{(C)}$, we start by conditioning on the number of the largest size jobs on machines. Then, loss probability $L_\pi^{(C)}$ can be upper bounded as

$$\begin{aligned} L_\pi^{(C)} &\leq \mathbb{P}\left[\cap_{m=1}^{M}\{Q_{\pi,n}^{(m)} + B_n > C_m\}, \sum_m N_{\pi,n}^{(m)}(C) < \sum_m k_m\right] \\ &\quad + \mathbb{P}\left[\cap_{m=1}^{M}\{Q_{\pi,n}^{(m)} + B_n > C_m\}, \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m\right] \\ &\triangleq I_1 + I_2. \end{aligned} \tag{5}$$

Next, we focus on proving an asymptotic estimate for I_1 as $C \to \infty$. Note that $B_i, i \in \mathcal{N}_{\pi,n}^{(m)}$, are mutually dependent, negatively associated random variables. Therefore, one of the main challenges is to come up with a way to handle this dependency. By the assumption on machine resource capacities, $N_{\pi,n}^{(m)}(C) \leq k_m$, $1 \leq m \leq M$, and, therefore, $\{\sum_m N_{\pi,n}^{(m)}(C) < \sum_m k_m\} = \cup_m\{N_{\pi,n}^{(m)}(C) < k_m\}$. Thus, using the union bound we obtain

$$\begin{aligned} I_1 &= \mathbb{P}\left[\cap_{m=1}^{M}\{Q_{\pi,n}^{(m)} + B_n > C_m\}, \cup_m\{N_{\pi,n}^{(m)}(C) < k_m\}\right] \\ &\leq \sum_{m=1}^{M} \mathbb{P}[Q_{\pi,n}^{(m)} + B_n > C_m, N_{\pi,n}^{(m)}(C) < k_m] \\ &\leq \sum_{m=1}^{M} \mathbb{P}\left[\sum_{i\in\mathcal{N}_{\pi,n}^{(m)}} B_i + (X_n \wedge C) > C_m, N_{\pi,n}^{(m)}(C) < k_m\right] \\ &\triangleq \sum_{m=1}^{M} I_{1m}. \end{aligned}$$

Now, since $C_m = k_m C + f_m(C)$ and $B_i = X_i \wedge C$, we further upper bound the previously obtained expression as

$$I_{1m} \leq \mathbb{P}\left[\sum_{i\in\mathcal{N}_{\pi,n}^{(m)}\cap\mathcal{N}_n^{(<C)}} X_i + (X_n \wedge C) > C + f_m(C)\right], \tag{6}$$

where we use $\mathcal{N}_{\pi,n}^{(<C)}$ to denote set

$$\mathcal{N}_n^{(<C)} \triangleq \{i < n | X_i < C, \theta_i > \tau_n - \tau_i\}.$$

In words, $\mathcal{N}_n^{(<C)}$ represents a set of jobs that arrived prior to τ_n, requesting less than C resources and, assuming that all of them are accepted, these jobs would still be on one of machines, i.e., their duration is longer than the time that elapsed since their arrival. In other words, the previously defined set of jobs would correspond to the number of jobs in the system of M machines with unlimited capacity, that arrived prior to time τ_n, requested less than C resources and are still in process at the instance of nth arrival. Therefore, (6) can be further upper bounded as

$$I_{1m} \leq \mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(<C)}} X_i + (X_n \wedge C) > C + f_m(C)\right].$$

Note that X_i, $i \in \mathcal{N}_n^{(<C)}$, are mutually independent subexponential random variables. Next, after conditioning on

the size of X_n, we obtain

$$I_1 \leq \mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(<C)}} X_i > f_m(C), X_n \geq C\right] + \mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(<C)}} X_i + X_n > C + f_m(C), X_n < C\right]$$
$$\leq \mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(<C)}} X_i > f_m(C)\right]\mathbb{P}[X_n \geq C]$$
$$+\mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(<C)}} X_i + X_n > C + f_m(C), X_n < C, X_i < C \text{ for all } i \in \mathcal{N}_n^{(<C)}\right], \quad (7)$$

where in the last equality we use independence of X_n and X_i, $i \in \mathcal{N}_n^{(<C)}$.

Now in view of the results derived in [4], for every integer n and for i.i.d. subexponential random variables $X_1, \dots, X_n$,

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i > C\right] \sim \mathbb{P}[\max(X_1, \dots, X_n) > C] \text{ as } C \to \infty,$$

implying the asymptotic relation

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i > C, X_i \leq C \text{ for every } 1 \leq i \leq n\right] = o(\mathbb{P}[X_1 > C]) \text{ as } C \to \infty.$$

Now, in order to show that n can be replaced by $N_n^{(<C)}$ in the previous expression, we need to integrate it with respect to density of $N_n^{(<C)}$, i.e.,

$$\mathbb{P}\left[\sum_{i\in\mathcal{N}^{(<C)}\cup\{n\}} X_i > C, X_i < C \text{ for every } i \in \mathcal{N}_n^{(<C)} \cup \{n\}\right]$$
$$= \sum_{k=0}^{\infty} \mathbb{P}[N_n^{(<C)} = k]\mathbb{P}\left[\sum_{i=1}^{k+1} X_i > C, X_i \leq C \text{ for every } i = 1, 2, ..., k+1\right].$$

Now, due to the lemma stated by Kesten (see Lemma 7 on page 149 of [2]), for any $\epsilon > 0$, there exists a positive constant $K(\epsilon)$ such that

$$\frac{\mathbb{P}[\sum_{i=1}^{k} X_i > C,\ X_i \leq C \text{ for every } 1 \leq i \leq k]}{\mathbb{P}[X > C]} \leq \frac{\mathbb{P}[\sum_{i=1}^{k} X_i > C]}{\mathbb{P}[X > C]} \leq K(\epsilon)(1+\epsilon)^k$$

for any integer k and for all capacity values $C < \infty$. Then, since the probability generating function $\mathbb{E}z^{N_n^{(<C)}}$ is finite for any $z \in \mathbb{C}$ (see [10] and [8] for the detailed proof), we have

$$\sum_{k=0}^{\infty} \mathbb{P}[N_n^{(<C)} = k](1+\epsilon)^k < \infty,$$

and, therefore, by applying the dominated convergence theorem we conclude that

$$\lim_{C\to\infty} \frac{\mathbb{P}[\sum_{i\in\mathcal{N}_n^{(<C)}\cup\{n\}} X_i > C, X_i < C \text{ for all } i \in \mathcal{N}_n^{(<C)} \cup \{n\}]}{\mathbb{P}[X_i > C]}$$
$$= \lim_{C\to\infty} \sum_{k=0}^{\infty} \frac{\mathbb{P}[N_n^{(<C)} = k]\mathbb{P}\left[\sum_{i=1}^{k+1} X_i > C,\ X_i \leq C \text{ for every } 1 \leq i \leq k+1\right]}{\mathbb{P}[X_i > C]}$$
$$= 0. \quad (8)$$

Now, using the result obtained in (8), one can directly derive that the second term of (7) can be asymptotically upper bounded by

$$\mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(<C)}\cup\{n\}} X_i > C + f_m(C), X_i < C \text{ for all } i\in\mathcal{N}_n^{(<C)}\cup\{n\}\right]$$
$$= o(\mathbb{P}[X_i > C]) \text{ as } C\to\infty,$$

which in conjunction with (7) implies that for all $1\le m\le M$

$$I_{1m} \sim \mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(<C)}} X_i > f_m(C)\right]\mathbb{P}[X_n\ge C] \text{ as } C\to\infty. \tag{9}$$

Next, we estimate the asymptotic upper bound for I_2 in (5). This case addresses the scenario where a newly arriving job finds the system with a maximum possible number of jobs with the largest size ($\sum_m k_m$ jobs of size C) in process. Then, since in this case the rest of active jobs on machine m can not exceed capacity $f_m(C)$, after conditioning on the size of $B_n = X_n\wedge C$, one can upper bound I_2 as

$$I_2 = \mathbb{P}\left[\cap_{m=1}^M\left\{\sum_{i\in\mathcal{N}_n^{(m)}} X_i + (X_n\wedge C) > f_m(C)\right\}, \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m\right] \tag{10}$$

$$\le \mathbb{P}\left[\cap_{m=1}^M\left\{\sum_{i\in\mathcal{N}_n^{(m)}} X_i + (X_n\wedge C) > f_m(C)\right\}, \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m, \cup_m\{X_n\le f_m(C)\}\right]$$
$$+\,\mathbb{P}\left[\cap_{m=1}^M\left\{\sum_{i\in\mathcal{N}_n^{(<f_m(C))}} X_i + (X_n\wedge C) > f_m(C)\right\}, \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m, X_n > \max_m f_m(C)\right], \tag{11}$$

where $\mathcal{N}_n^{(<f_m(C))}$ is defined similarly as before and represents a set of all jobs that arrived prior to τ_n, requested less than $f_m(C)$ resources, and are still active at τ_n in an unlimited capacity system (system without rejections).

Next, the second expression of (11) can be trivially upper bounded as

$$\mathbb{P}\left[\cap_{m=1}^M\left\{\sum_{i\in\mathcal{N}_n^{(<f_m(C))}} X_i + (X_n\wedge C) > f_m(C)\right\}, \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m, X_n > \max_m f_m(C)\right]$$
$$\le \mathbb{P}\left[\sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m\right]\mathbb{P}\left[X_n > \max_m f_m(C)\right], \tag{12}$$

where in the last expression we used independence between newly arriving job size $B_n = X_n\wedge C$ and $\sum_m N_{\pi,n}^{(m)}(C)$.

Now, we focus on estimating the first term in (11). First, after applying the union bound, we obtain

$$\mathbb{P}\left[\cap_{m=1}^M\left\{\sum_{i\in\mathcal{N}_n^{(m)}} X_i + (X_n\wedge C) > f_m(C)\right\}, \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m, \cup_m\{X_n\le f_m(C)\}\right]$$
$$\le \sum_{m=1}^M \mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(m)}} X_i + (X_n\wedge C) > f_m(C), X_n\le f_m(C), \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m\right]. \tag{13}$$

The key observation in estimating each summand in (13) is that

$$\mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(m)}} X_i + (X_n \wedge C) > f_m(C), X_n \le f_m(C), \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m\right]$$
$$\sim \mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{\infty}} X_i + (X_n \wedge C) > f_m(C), X_i \le f_m(C), i \in \mathcal{N}_n^{\infty}\cup\{n\}\right]\mathbb{P}\left[\sum_m N_n^{\infty}(C) = \sum_m k_m\right], \quad (14)$$

where $\mathcal{N}_n^{\infty}$ is a set of all jobs in the system of unlimited capacity that are still in process at the moment of nth arrival and $N_n^{\infty}(C)$ is the number of the largest size jobs in this system, where we assume that the system does not allow more than $\sum_m k_m$ jobs of these jobs. We state this result in Lemma 1 after this proof. The proof of this lemma is very technical and due to space limitations, we leave its details for the full version of this paper. Next, in view of the completely analogous arguments that led to (8), we conclude that

$$\mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(m)}} X_i + (X_n \wedge C) > f_m(C)\}, X_n \le f_m(C), \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m\right]$$
$$\sim o\left(\mathbb{P}[X_n > f_m(C)]\mathbb{P}\left[\sum_m N_n^{\infty}(C) = \sum_m k_m\right]\right) \text{ as } C \to \infty.$$

Now, by the assumption $|\max_m f_m(C) - \min_m f_m(C)| < K$ for all C large and (1), we conclude that the left hand side of inequality (13), in conjunction with (12) satisfies

$$I_2 \lesssim \mathbb{P}[X_n > f_m(C)]\mathbb{P}\left[\sum_m N_n^{\infty}(C) = \sum_m k_m\right] \text{ as } C \to \infty. \quad (15)$$

Finally, in order to prove asymptotic lower bound for (3), we can simply use

$$L_{\pi}^{(C)} \ge \mathbb{P}\left[\sum_{m=1}^{M} N_{\pi,n}^{(m)}(C) = \sum_m k_m, X_n > \max_m f_m(C)\right]$$

and Lemma 1, which in conjuction with (15), (9), (5) and assuming $\sum_{m=1}^{M}\mathbb{P}[\sum_{i\in\mathcal{N}_n^{(<C)}} X_i > f_m(C)]\mathbb{P}[X_n \ge C] = o(\mathbb{P}\left[\sum_{m=1}^{M} N_{\pi,n}^{(m)}(C) = \sum_m k_m, X_n > \max_m f_m(C)\right])$ as $C \to \infty$, concludes the proof of the theorem. $\diamond$

Remark: (i) It may appear surprising that the performance of the loss system from above does not depend on engagement durations, as long as they have finite mean. In addition, the result is quite general and provides the asymptotic result for a huge (subexponential) class of possible resource requirement distributions. (ii) Another surprising conclusion that follows from the result of Theorem 1 is that the job loss rate does not depend on the number of machines M, but only of the maximum number of the largest jobs one can store on them, i.e., $\sum_m k_m$.

Lemma 1: Under the modeling assumptions from the beginning of Section II, the stationary job loss probability satisfies

$$\mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{(m)}} X_i + X_n \wedge C > f_m(C)\}, X_n \le f_m(C), \sum_m N_{\pi,n}^{(m)}(C) = \sum_m k_m\right]$$
$$\sim \mathbb{P}\left[\sum_{i\in\mathcal{N}_n^{\infty}} X_i + X_n \wedge C > f_m(C)\}, X_i \le f_m(C), i \in \mathcal{N}_n^{\infty}\cup\{n\}\right]\mathbb{P}\left[\sum_m N_n^{\infty}(C) = \sum_m k_m\right],$$

as $C \to \infty$, where $\mathcal{N}_n^{\infty}$ is a set of all jobs in an unlimited capacity system which does not allow more than $N_n^{\infty}(C) = \sum_m k_m$ jobs of size C.

A. A Single Large Machine and Performance Comparisons

In this section we relate the asymptotic loss rate of the M-machine system that runs a Non-Forced Idle Time Scheduling policy to the loss rate of a single large machine with the same total capacity $\sum_m C_m$. In a single machine case, we assume that an arriving job is lost only if there is no available resource capacity to satisfy its processing requirement. It is not hard to show that the loss rate of such a "non-fragmented" single machine case is less then the loss rate of the M-machine system. Thus, if we use $L_s^{(C)}$ to denote the loss rate in case of a single machine, we conclude

$$\min_{\pi \in Non-FITP} L_\pi^{(C)} \geq L_s^{(C)}. \tag{16}$$

Next, it can be shown that $L_s^{(C)} \sim \mathbb{P}[N_n^\infty = \sum_m k_m]\mathbb{P}[X_n > \sum_m f_m(C)]$ as $C \to \infty$. The proof of this claim follows from the identical arguments used in the proof of Lemma 1, and we omit it due to the space limitations. Here, we just state the theorem:

Theorem 2: Let $\{X_n, -\infty < n < \infty\}$ be subexponential random variables with finite mean. Then, if $B_n = X_n \wedge C$, $-\infty \leq n \leq \infty$, the stationary loss rate of a single machine with capacity $\sum_m C_m$, $L_s^{(C)}$, satisfies

$$L_s^{(C)} \sim \mathbb{P}\left[N_n^\infty(C) = \sum_m k_m\right] \mathbb{P}\left[X_n > \sum_m f_m(C)\right] \text{ as } C \to \infty. \tag{17}$$

Remark: An immediate implication of Theorems 1 and 2 is that in the case one can estimate job request distribution with significant accuracy, we can directly obtain the asymptotic average-case competitive ratio for a scheduling policy π, i.e., from (3), (4) and (17), it follows that for C large enough

$$L_\pi^{(C)} \approx \frac{\mathbb{P}[X_n > \max_m f_m(C)]}{\mathbb{P}[X_n > \sum_m f_m(C)]} L_s^{(C)}. \tag{18}$$

In view of the previous expression, one can show that for certain subclasses of subexponential distributions, e.g., subexponential concave (see Definition 3.1 and Lemma 3.1 in [9]), and specific subsets of functions $f_m(C)$, $1 \leq m \leq M$, it can be directly shown that $\mathbb{P}[X_n > \max_m f_m(C)] \sim \mathbb{P}[X_n > \sum_m f_m(C)]$ as $C \to \infty$, implying asymptotic optimality of Non-FITP policies.

IV. Numerical Example

In this section with a described simulation experiment we had supported the accuracy of our asymptotic formulas, proved in Theorems 1 and 2. Our goal is to show that even though our results are asymptotic, the derived estimates for the performance ratio between chosen Non-FITP policy (Best Fit on two machines) and a single machine case match experiments with high accuracy even for systems with moderately large capacities and in regimes of "large" loss rates.

We simulate two systems: (i) a two-machine case with the Best Fit scheduling of incoming jobs, and (ii) a single, large machine with the resource capacity equal to the sum of capacities of the two machines in case (i). Best Fit policy schedules an incoming job to a machine with the smallest amount of available resources that can still satisfy job's requirements.

We let the first 10^8 requests to arrive before conducting measurements in order for the systems to reach stationarity. We intentionally choose the two machine case in order to reduce the time our simulated systems need to converge to stationarity. Then, we count the request loss rate ($Loss_{BF}$ and $Loss_{SingleM}$) among 10^9 arrivals.

When simulating the Best Fit policy, we assume that the two machines have the same capacities $C_1 = C_2 = C + \sqrt{C}$. Measurements are conducted for maximum job resource requirement values of $C = 200 + 50j, 0 \leq j \leq 9$. We let jobs arrive at Poisson time points with rate $\lambda = 1$. In addition, we assume that job durations are exponentially distributed with mean $\mu = 1$. Next, let job requirements B_i be drawn from a finite support distribution, where $\mathbb{P}[B_i = i] = \frac{1}{i^{1.2}}$, $1 \leq i \leq C-1$, and $\mathbb{P}[B_i = C] = 1 - \mathbb{P}[B_i < C]$ (power law distribution).

As commented in the Remark at the end of the previous section, our main results in Theorems 1 and 2 imply that

$$Ratio_{BF} \triangleq \frac{Loss_{BF}}{\mathbb{P}[B_i > \sqrt{C}]} \approx \frac{Loss_{SingleM}}{\mathbb{P}[B_i > 2\sqrt{C}]} \triangleq Ratio_{SingleM}.$$

The experimentally obtained ratios, $Ratio_{BF}$ and $Ratio_{SingleM}$ we present on Figure IV. Even though considered capacities in our experiments are relatively small, we obtain amazingly accurate match between measured $Ratio_{BF}$ and $Ratio_{SingleM}$, showing that the asymptotically computed performance ratio might hold even in cases of relatively small capacities C.

V. Conclusion

In this paper we analyze the performance of a class of machine scheduling policies known as Non-Forced Idle Time policies. We consider a stochastic framework where jobs arrive according to a compound renewal process in time, require random amount of resources and take arbitrarily distributed processing time (holding time) with finite mean. Our main objective is to estimate the performance of the analyzed class of scheduling policies in case when job resource requirements exhibit significant statistical variability. In order to cope with this objective, we model them as truncated subexponential random variables where job requirements can not exceed a certain large value. We develop a simple estimate for the asymptotic job loss rate of Non-Forced Idle Time Scheduling policies when the maximum job requirement and machine sizes grow large. We develop an asymptotic lower bound for the performance (job loss rate) of Non-Forced Idle Time Policies and show that it can be reached under certain assumptions on machine capacities and job requirement distributions. Our results have many interesting implications related to estimating distributional properties of the specific loss systems in the presence of general probabilistic assumptions such as renewal arrivals and generally distributed processing durations.

References

[1] A. Barak R. S. Borgstrom A. Yair B. Awerbuch and A. Keren, An opportunity cost approach for job assignment in a scalable computing cluster, *IEEE Transactions on Parallel and Distributed Systems*, 7:760–768, July 2000.

[2] K. B. Athreya and P. E. Ney, Branching Processes, *Springer-Verlag*, 1972.

[3] V. P. Chistyakov, A theorem on sums of independent positive random variables and its application to branching random processes, *Theor. Probab. Appl.*, 9:640–648, 1964.

[4] P. Embrechts and C. M. Goldie, On closure and factorization properties of subexponential and related distributions, *J. Austral. Math. Soc.*, Series(A) 29:243–256, 1980.

[5] C. M. Goldie and C. Klüppelberg, Subexponential distributions, In M.S. Taqqu R. Adler, R. Feldman, editor, *A Practical Guide to Heavy Tails: Statistical Techniques for Analysing Heavy Tailed Distributions*, pages 435–459. Birkhäuser, Boston, 1998.

[6] J. Sgall K. Pruhs and E. Torng, Online Scheduling, In J. Leung, editor, *Handbook of Scheduling: Algorithms, Modeling and Performance Analysis*, CRC Press, 2004.

[7] F. P. Kelly, Loss networks, *Annals of Applied Probability*, 1:319–378, 1991.

[8] L. Liu, B. R. K. Kashyap, and J. G. C. Templeton, On the $GI^X/G/\infty$ system, *Journal of Applied Probability*, 27:671–683, 1990.

[9] Petar Momcilović P.R. Jelenković and Bert Zwart, Large deviation analysis of subexponential waiting times in a processor sharing queue, 2001.

[10] L. Takacs, Queues with infinitely many servers, *R.A.I.R.O. Recherche Operationnelle*, 14:109–113, 1980.

[11] N. H. Lee W. N. Kang, F. P. Kelly and R. J. Williams, State space collapse and diffusion approximation for a network operating under a fair bandwidth sharing policy, *Preprint*, 2007.

[12] H. Karloff Y. Bartal, A. Fiat and R. Vohra, New algorithms for an ancient scheduling problem, *Proceedings of ACM Symposium on Theory of Algorithms*, 1992.