Journal of the Operations Research
Society of Japan
2004, Vol. 47, No. 4, 304-313

STABILITY OF A SIMPLE RE-ENTRANT LINE WITH INFINITE SUPPLY OF WORK — THE CASE OF EXPONENTIAL PROCESSING TIMES

Gideon Weiss*
The University of Haifa

(Received August 13, 2003; Revised March 15, 2004)

Abstract We consider a two machine 3 step re-entrant line, with an infinite supply of work. We assume that processing times are exponentially distributed. We show that this system is stable under LBFS priority policy.

Keywords: Queue, manufacturing, priority scheduling policies, stability, virtual infinite buffers

1. Introduction

We consider a production system with two machines, and a 3 step production process, where each part is processed first by machine one for the first step, then by machine two for the second step, and finally again by machine one for the third step, before leaving the system. The processing times for each of the 3 steps are independent sequences of independent identically distributed random variables, with means m_i and rates $\mu_i = 1/m_i$, $i = 1, 2, 3$. This system is the simplest example of a re-entrant line (as defined by Kumar [10]), which in turn is a special case of a multi-class queueing network (as described by Harrison [8]). This particular system has previously been studied in [3, 5, 9, 14].

It is known that if parts arrive at this system in a renewal stream, at rate α, then under the condition $\rho_1 = \alpha(m_1 + m_3) < 1, \rho_2 = \alpha m_2 < 1$ the queues of parts waiting for each step are stable, and in fact the system is positive Harris recurrent, for any work conserving policy (Dai and Weiss [5]). It is also known that any re-entrant line with $\rho_i = \alpha \sum_{k \in C_i} m_i < 1$, $i = 1, \dots, I$ (where the consecutive processing steps are $k = 1, \dots, K$, and steps $k \in C_i$ are performed at machine i) has stable queues, and is positive Harris recurrent, under the LBFS (Last Buffer First Served) policy (Kumar and Kumar [11] and Dai and Weiss [5]).

If however the arrival rate α is high enough to equal the bottleneck processing rate, i.e. $\max\{\alpha(m_1 + m_3), \alpha m_2\} = 1$, then the system is not stable: As time increases, the queue length at some of the buffers will $\Rightarrow \infty$. Thus such a system cannot work at a rate $\max\{\rho_1, \rho_2\} = 1$, without accumulating unbounded queues.

In this note we consider a different situation, which is typical of manufacturing systems. We assume that there is an infinite supply of work available, so that there are always parts ready for processing step 1. In that case machine 1 will always be busy. We investigate the stability of this system under LBFS policy. In particular we show that if $m_1 + m_3 > m_2$ then under LBFS policy machine 1 will work all the time (that is we will have $\rho_1 = 1$), but the queues for steps 2,3 will be stable, and the system will be positive recurrent. This result

*Research supported in part by Israel Science Foundation Grant 249/02.

has some practical applications in job-shop scheduling heuristics, for further explanations and numerical experiments see [12, 13, 15].

For simplicity we assume in this note that all the processing times are exponentially distributed.

2. The Two Machine 3 Buffer Re-entrant Line with Virtual Infinite Queue in Buffer 1, and with Exponential Processing Times

Our re-entrant line manufacturing system is described schematically in Figure 1. Processing

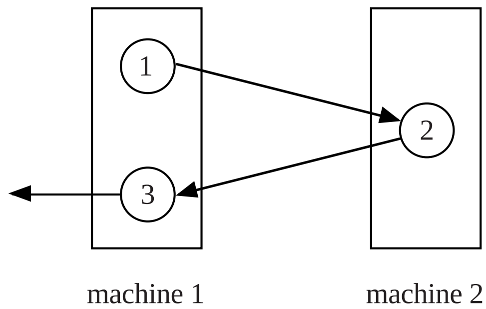

Figure 1: A 2 machine 3 step system, with virtual infinite buffer

times at step i are i.i.d exponentially distributed with mean m_i, rate $\mu_i = 1/m_i$, for $i = 1, 2, 3$ and the three sequences are independent. Without loss of generality we scale time so that $\mu_1 + \mu_2 + \mu_3 = 1$.

There are always parts available for processing of step 1. When parts finish processing step 1 by machine 1, they queue in buffer 2 where they remain until they are processed by machine 2 for step 2, and then they move to buffer 3, where they remain until they are processed by machine 1 for step 3, at which time they leave the system. Each buffer is processed in FIFO order. Processing is non-idling, that is a machine will always process a part when there is work. We assume that machine 1 gives preemptive priority to buffer 3: Whenever there are parts in buffer 3, machine 1 will work on the first of them. When buffer 3 empties, machine 1 will immediately resume processing of a part in step 1. This is possible by the assumption that there is an infinite supply of work. We can think of it as if buffer 1 has an infinite queue of parts waiting for step 1. We call such a buffer *a virtual infinite queue*. The queue is virtual because in practice buffer 1 need not contain many parts, but it needs to be monitored so it will never be empty. If during the processing of step 1 a part arrives from buffer 2 into buffer 3, machine 1 will preempt its work at buffer 1, and immediately start processing buffer 3.

Since the processing times are exponential we can describe this system as a discrete state continuous time Markov jump process, with the state given by the number of parts in buffers 2,3, denoted n_2, n_3. The state of the system at time t is $X(t) = (n_2, n_3), t \geq 0$. The transition rates of $X(t)$ are presented in Figure 2. They are:

$$
\begin{aligned}
(n_2, n_3) &\rightarrow (n_2 - 1, n_3 + 1) \text{ at rate } \mu_2, \quad n_2 > 0 \\
(n_2, n_3) &\rightarrow (n_2, n_3 - 1) \text{ at rate } \mu_3, \quad n_3 > 0 \\
(n_2, 0) &\rightarrow (n_2 + 1, 0) \text{ at rate } \mu_1, \quad n_2 \geq 0
\end{aligned}
\tag{2.1}
$$

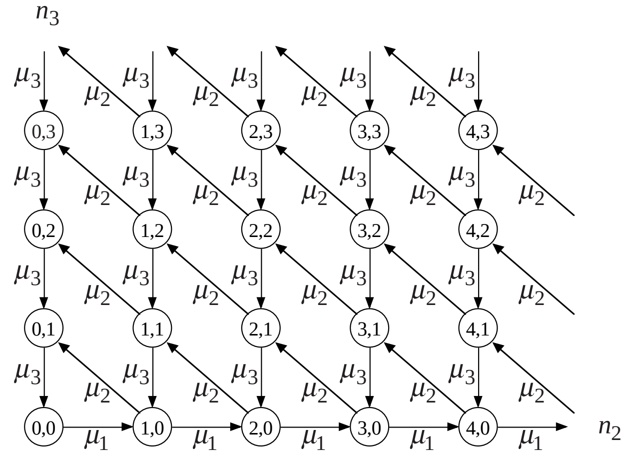

Figure 2: Transition rates for the Markovian states of the re-entrant system

The sample paths of this Markov system evolve as follows: Whenever $n_3 = 0$ parts are processed out of buffer 1 and n_2 increases until machine 2 completes the processing of a part out of buffer 2, at which time the state is $(n_2 - 1, 1)$, and machine 1 switches to buffer 3. While buffer 3 is not empty, parts arrive at buffer 3 from buffer 2 at rate μ_2 and depart out of buffer 3 at rate μ_3, so buffer 3 behaves like an M/M/1 queue, except that the total number of arrivals into buffer 3 cannot exceed n_2. Thus machine 1 stays at buffer 3 for a truncated busy period of an M/M/1 queue: either the busy period ends before buffer 2 runs out of parts, or buffer 2 empties first, and then machine 2 will idle while machine 1 will complete the processing of all remaining parts in buffer 3. Note that in the latter case the system will arrive at state $(0, 0)$, and the next transition will be to $(1, 0)$.

The dependence of the likely behavior of sample paths on the parameters is of interest:

Consider first the case $m_2 > m_1 + m_3$. Jobs can enter buffer 3 out of buffer 2 at a rate which must be $\leq \mu_2 = 1/m_2$. Hence, machine 1 will under LBFS be processing parts out of buffer 3 at a rate not exceeding $\mu_2 = 1/m_2$, and so the long term fraction of time that machine 1 will be processing buffer 3 will not exceed $\mu_2 m_3 = \frac{m_3}{m_2}$. Therefore machine 1 will be processing parts out of buffer 1 a fraction of time $\geq 1 - \frac{m_3}{m_2}$, and so parts will be leaving buffer 1 and entering buffer 2 at a rate $\geq \mu_1(1 - \frac{m_3}{m_2})$. But $\mu_1(1 - \frac{m_3}{m_2}) - \mu_2 = \frac{m_2 - m_1 - m_3}{m_1 m_2} > 0$ if $m_2 > m_1 + m_3$. Hence buffer 2 will fill up at a rate > 0. Hence almost surely for each sample path of the process there will be a finite time after which buffer 2 will never be empty again. Following that finite time, both buffer 1 and buffer 2 will have unlimited supply of work, and so both machines 1 and 2 will work all the time. Buffer 3 will behave line an M/M/1 queue with arrival rate μ_2 and service rate μ_3, and machine 1 will work on buffer 1 in the idle periods of buffer 3. In practice, if $m_2 > m_1 + m_3$ we will not let the queue at buffer 2 grow indefinitely. In fact, it will be reasonable to replenish buffer 2 only when it falls below some threshold B_2. In that case there will be some steady state probability that buffer 2 will fall to 0 before it is replenished, but by increasing B_2 we can make this probability arbitrarily small and so make the throughput of the system arbitrarily close to μ_2.

We will show in this note that the system is stable if $m_2 < m_1 + m_3$. We distinguish the case when $m_3 > m_2$, the case when $m_1 + m_3 > m_2 > m_3$, and the case when $m_2 = m_3$.

Consider first the case $m_3 > m_2$. In that case an M/M/1 queue with input rate μ_2 and service rate μ_3 is unstable, hence there is a positive probability for each busy period to be infinite. This probability is a lower bound on the probability that a truncated busy period starting with n_2 will deplete all the n_2 parts out of buffer 2 before emptying buffer 3. Hence, there is a positive lower bound on the probability of returning to $(0, 0)$ at the end of each busy period, no matter how many parts are in buffer 2 at the beginning of the busy period. Hence we have a bound on the expected number of busy periods between visits to $(0, 0)$. On the other hand, because the M/M/1 queue is unstable, the length of a busy period increases without bound as the initial content of buffer 2, the value of n_2 at the beginning of the truncated busy period, increases. So the main issue is to show that the average truncated busy period has finite expected duration.

Consider next the case $m_1 + m_3 > m_2 > m_3$. In that case an M/M/1 queue with input rate μ_2 and service rate μ_3 is stable. Hence we have a uniform bound on the expected duration of each truncated busy period. On the other hand, the probability that a busy period ends in state $(0, 0)$ tends to 0 as the initial level of buffer 2, $n_2 \to \infty$. So now the main issue is to show that the expected number of busy periods between visits to $(0, 0)$ is finite.

Finally, in the case $m_2 = m_3$ the M/M/1 queue is unstable, with busy periods which are finite with probability 1, but the expected length of each busy period is infinite. Hence in this case we do not have a uniform bound on the expected lengths of truncated busy periods, and we also do not have a lower bound on the probability that a truncated busy period will end with $n_2 = n_3 = 0$.

In the next Section 3 we give a formal proof of the stability. The proof uses a Lyapunov function approach for the case $m_1 + m_3 > m_2 > m_3$, and uses stochastic domination for the cases $m_3 \geq m_2$.

3. Stability of LBFS when $m_1 + m_3 > m_2$.

Our main result in this note is:

Theorem 3.1 *Assume that $m_1 + m_3 > m_2$. Then the Markov jump process $X(t)$ is positive recurrent.*

In the remainder of this Section we prove Theorem 3.1. We first make some preliminary observations, define several processes and quantities related to $X(t)$, and introduce some notation.

To prove that a Markov jump process is positive recurrent we need first to know that the process is irreducible. That $X(t)$ is irreducible (i.e. it is possible to go from any state to any other state in a finite number of steps) is obvious from a quick observation of the transition rates (2.1) and Figure 2.

We can uniformize our chain to have Poisson events at rate 1. Such an event will with probability μ_i be either a completion of the processing of a part in buffer i, if a machine is working on buffer i at the time, or it will be a null event. Denote by $N(t)$ the rate 1 Poisson process which counts these uniformized transitions of $X(t)$.

Denote by X_s, $s = 0, 1, \ldots$ the discrete time Markov chain of the states after each jump of the Poisson process $N(t)$. Then $X_0 = X(0), X_{N(t)} = X(t), t > 0$. Note that Figure 2 describes the transition probabilities of the Markov chain X_s (since we assumed $\mu_1 + \mu_2 + \mu_3 = 1$). The process X_s has positive probability for a null transition in every state. Hence it is a-periodic (in addition to being irreducible).

Clearly, $X(t)$ is positive recurrent if and only if X_s is positive recurrent.

Consider the visits of X_s to the set of states $R = \{(n_2, 0), n_2 = 0, 1, 2, \ldots\}$, the states in which buffer 3 is empty. Let $0 \leq S_0 < S_1 < S_2 < \ldots$ be the times at which $X_s \in R$. Let Y_k be the embedded Markov chain defined by $Y_k = n_2$ if $X_{S_k} = (n_2, 0), k = 0, 1, 2, \ldots$.

The transition probabilities of the embedded Markov chain Y_k, are given by:

$$\begin{aligned} n_2 &\to n_2 && \text{with probability } \mu_3, \\ n_2 &\to n_2 + 1 && \text{with probability } \mu_1, \\ n_2 &\to n_2 - L && \text{with probability } \mu_2, \end{aligned} \tag{3.1}$$

where L is the random number of parts processed in a busy period of an M/M/1 queue, with arrival rate μ_2 and service rate μ_3, truncated by n_2 (the total number processed is $\leq n_2$).

There are several ways in which one can prove stability, or more precisely positive recurrence of a discrete state discrete time Markov process; for a recent survey on the topic see [6].

The first is to solve the balance equations, and check that the solution is positive and converges to 1. Unfortunately the balance equations of our 2 machine 3 buffers re-entrant line (as in most multi-class queueing networks) do not seem tractable [1].

The second method is to prove that one of the states of the Markov chain is positive recurrent: An a-periodic irreducible Markov chain is positive recurrent if we can find a single state such that the expected time to return to it is bounded. More generally, one may be able to show that a regenerative event happens at time intervals with a finite expectation.

The third method is to use the Foster-Lyapunov criterion (see for example [2, 6, 7])

Theorem 3.2 (Foster Lyapunov Criterion) *Let Z_n be a discrete state discrete time a-periodic irreducible Markov chain. The following condition is sufficient for the chain to be positive recurrent. There exist a finite set of states S_0, a non-negative function of the states h, and positive constants B, ϵ such that, conditional on $Z_0 = z$:*

(i) *For all $z \in S_0$, $\mathbb{E}_z(h(Z_1)) < B$*

(ii) *For all $z \notin S_0$, $\mathbb{E}_z(h(Z_1)) - h(z) < -\epsilon$*

The fourth method is to use stochastic domination: If we can show that the Markov chain under discussion is dominated in the sense of stochastic ordering by a process which is positive recurrent then the chain under discussion must also be recurrent.

A very powerful fifth method has been developed recently by Dai [4], to prove positive Harris recurrence of multi-class queueing networks under given policy. Dai's theorem states that a multi-class queueing network is stable under a given policy (i.e. the general state-apace Markov process which describes the behavior of the network under the policy is positive Harris recurrent) if the corresponding fluid limit model of the multi-class network under the policy is stable. Unfortunately, Dai's theorem does not apply directly to our network: Dai's theorem assumes i.i.d. interarrival times, whereas in our network arrivals occur only when buffer 3 is empty.

In our proof now we will use the Foster-Lyapunov criterion directly on X_s in the case $m_1 > m_2$, we will use the Foster-Lyapunov criterion on the embedded chain Y_k in the case $m_1 + m_3 > m_2 > m_3$. Finally, we will use stochastic domination to show that the state $X(t) = (0, 0)$ is positive recurrent for $m_3 \geq m_2$. We also give a more direct proof that the state $X(t) = (0, 0)$ is positive recurrent for $m_3 > m_2$.

3.1. Proof of positive recurrence in the case $m_1 > m_2$

This case is in fact covered by the subsequent two cases, but we discuss it separately because it is much simpler. In this case we show that X_s is positive recurrent by the use of the Foster-

[1] Recently Adan and Weiss [1] have obtained a closed form formula for the steady state distribution

Lyapunov criterion, where we take the set of states $S_0 = \{(0,0)\}$ and the linear Lyapunov function $h(n_2, n_3) = 2m_1 n_2 + (m_1 - m_2)n_3$.

Clearly, $\mathbb{E}_{(0,0)}(h(X_1)) = 2$ is finite.

Conditional on the initial state $x_0 = (n_2, n_3) \neq (0,0)$ we calculate:

$$\mathbb{E}_{x_0}(h(X_1)) - h(x_0) = \begin{cases} -\frac{m_1+m_2}{m_2} - \frac{m_1-m_2}{m_3} & n_3 > 0 \\ -\left(\frac{m_1}{m_2} - 1\right) & n_3 = 0, n_2 > 0 \end{cases}$$

Since this is less than a negative constant, the chain is positive recurrent.

It is easy to see that if $m_1 \leq m_2$ no function of the form $h(n_2, n_3) = an_2 + bn_3$ can serve as a Lyapunov function for the process.

3.2. Proof of positive recurrence in the case $m_1 + m_3 > m_2 > m_3$

In this case the M/M/1 queue with arrival rates μ_2 and processing rate μ_3 is stable. Hence a busy period of this queue will serve a finite number of parts and last for a finite time. In fact the expected number of parts to be served in a busy period is $\frac{m_2}{m_2-m_3}$, and its expected duration is $\frac{m_2 m_3}{m_2-m_3}$ (see for example textbook of Wolff [16]). An actual excursion of $X(t)$ away from R, starting from $(n_2 - 1, 1)$, consists of such a busy period truncated by the total number of arrivals $\leq n_2$. Hence the expected length of time for $X(t)$ or X_r to return to R is bounded by $\frac{m_2 m_3}{m_2-m_3}$.

Since the excursions away from the set of states R have expected duration bounded by a constant, it follows that X_r and $X(t)$ are positive recurrent if and only if Y_s is positive recurrent.

We now show that Y_s is positive recurrent. We use the Lyapunov function $h(n_2) = n_2$. We have, for $y_0 = n_2$:

$$\mathbb{E}_{y_0}(h(Y_1)) - h(y_0) = \mu_1 1 - \mu_2 \mathbb{E}(\text{number served in the truncated busy period})$$

But if n_2 is taken large enough then the truncated busy period will with high probability equal an ordinary busy period, and the expected number of parts processed in the truncated busy period will be arbitrarily close to $\frac{m_2}{m_2-m_3}$. We now choose $\delta > 0$ small enough, we then choose n_2 large enough, and we define $S_0 = \{0, 1, \ldots, n_2\}$, so that for any $y_0 \notin S_0$ we have:

$$\mathbb{E}_{y_0}(h(Y_1)) - h(y_0) \leq \mu_1 1 - \mu_2 \frac{m_2}{m_2 - m_3} + \delta = \frac{m_2 - m_1 - m_3}{m_1(m_2 - m_3)} + \delta < 0.$$

On the other hand, for all $y_0 \in S_0$

$$\mathbb{E}_{y_0}(h(Y_1)) \leq y_0 + 1 \leq n_2 + 1.$$

This shows that Y_s is positive recurrent.

3.3. Dominance as m_2 increases, and proof for the case $m_3 \geq m_2$

Define the departure processes from the three buffers, $D_i(t)$ is the number of departures from buffer i, for $i = 1, 2, 3$ over the time period $(0, t]$, so that the buffer levels are $X_i(t) = X_i(0) + D_{i-1}(t) - D_i(t), i = 2, 3$. Let $\tau_x = \inf\{t > 0 : X(t) = (0,0) | X(0) = x\}$, be the time to return to an empty system, state $(0,0)$, from initial state x, and in particular let $\tau_0 = \tau_{(1,0)}$. Note that in state $(0,0)$ only buffer 1 is being processed, and the system always transits to state $(1,0)$ after a time $\sim \exp(\mu_1)$. Hence τ_0 is the time to return to state $(0,0)$ measured from the instant that the process leaves state $(0,0)$. To show that the system is positive recurrent it is enough to show that $\mathbb{E}(\tau_0) < \infty$.

We consider two systems as above, with parameters $m_1, \tilde{m}_2, m_3$ and $m_1, \tilde{\tilde{m}}_2, m_3$ respectively. We shall distinguish quantities related to the two different system by superscribing them with $\tilde{\cdot}$ or $\tilde{\tilde{\cdot}}$. We denote by $\geq_{ST}$ stochastic dominance ($X \geq_{ST} Y$ if $\mathbb{P}(X > x) \geq \mathbb{P}(Y > x)$ for all x). The following Proposition examines the effect of a change in m_2 on the behavior of the process.

Proposition 3.1 *Consider the two systems with* $\tilde{\tilde{X}}_2(0) = \tilde{X}_2(0) = x_2 > 0$ *and* $\tilde{\tilde{X}}_3(0) = \tilde{X}_2(0) = x_3$. *If* $\tilde{\tilde{m}}_2 > \tilde{m}_2$ *then* $\tilde{\tilde{\tau}}_{x_2,x_3} \geq_{ST} \tilde{\tau}_{x_2,x_3}$

Proof. We do the proof by constructing coupled sample paths for both systems, in which the required inequalities hold almost surely. This implies stochastic dominance.

Both systems start in the same state $x = (x_2, x_3)$ with $x_2 > 0$. We generate the sequences of processing times for successive parts at buffers 1,2,3 as follows: The sequences of processing times at buffers 1 and 3 are the same for both systems. Processing times for buffer 2 at both systems are generated as a sequence of exponential processing times of rate $\tilde{\mu_2}$. Each of these processing times completes a job and results in a departure from buffer 2 in system $\tilde{\cdot}$, while in the system $\tilde{\tilde{\cdot}}$ it only results in a departure with probability $\tilde{\tilde{\mu}}_2/\tilde{\mu}_2$.

We consider the times before buffer 2 empties in either of the two systems, i.e. $\{t : \tilde{\tilde{X}}_2(s) > 0, \tilde{X}_2(s) > 0, 0 \leq s < t\}$. Since buffer 2 in both systems is not empty until time t, processing at buffer 2 proceeds for the full duration t at both systems. Because buffer 2 processing times are longer ($\geq$) in system $\tilde{\tilde{\cdot}}$ we have $\tilde{\tilde{D}}_2(t) \leq \tilde{D}_2(t)$ for every sample path. Since departures out of buffer 2 occur earlier in system $\tilde{\cdot}$ than in system $\tilde{\tilde{\cdot}}$, each part is available for processing in buffer 3 in system $\tilde{\cdot}$ earlier than in system $\tilde{\tilde{\cdot}}$. Hence (recalling that buffer 3 has preemptive priority over buffer 1) parts start their processing earlier in system $\tilde{\cdot}$. This implies that $\tilde{\tilde{D}}_3(t) \leq \tilde{D}_3(t)$ for every sample path. Finally, since system $\tilde{\cdot}$ devotes more time to buffer 3 than system $\tilde{\tilde{\cdot}}$, all the remaining time, which is devoted to buffer 1, is less in system $\tilde{\cdot}$ than in system $\tilde{\tilde{\cdot}}$. Hence: $\tilde{\tilde{D}}_1(t) \geq \tilde{D}_1(t)$ for every sample path.

We now have: $\tilde{\tilde{X}}_2(t) = x_2 + \tilde{\tilde{D}}_1(t) - \tilde{\tilde{D}}_2(t) \geq x_2(0) + \tilde{D}_1(t) - \tilde{D}_2(t) = \tilde{X}_2(t)$, and $\tilde{\tilde{X}}_2(t) + \tilde{\tilde{X}}_3(t) = x_2 + x_3 + \tilde{\tilde{D}}_1(t) - \tilde{\tilde{D}}_3(t) \geq x_2 + x_3 + \tilde{D}_1(t) - \tilde{D}_3(t) = \tilde{X}_2(t) + \tilde{X}_3(t)$ for every sample path, as long as $\tilde{X}_2(s) > 0, 0 \leq s < t$. Of course for such t we also have that $\tilde{\tilde{X}}_2(s) > 0, 0 \leq s < t$, since $\tilde{\tilde{X}}_2(s) \geq \tilde{X}_2(s)$. But this implies that $\tilde{\tilde{\tau}}_x \geq \tilde{\tau}_x$ for every sample path. ■

We have shown that our system is stable for average processing times $m_1 + m_3 > \tilde{\tilde{m}}_2 > m_3$. Hence for such a system $\mathbb{E}(\tilde{\tilde{\tau}}_0) < \infty$. Consider now $m_3 \geq \tilde{m}_2$. By Proposition 3.1 we have $\tilde{\tilde{\tau}}_0 \geq_{ST} \tilde{\tau}_0$. Hence, $\mathbb{E}(\tilde{\tau}_0) \leq \mathbb{E}(\tilde{\tilde{\tau}}_0) < \infty$. This proves positive recurrence for $m_3 \geq \tilde{m}_2$.

3.4. Alternative proof of positive recurrence in the case $m_3 > m_2$

We provide an alternative direct proof of the positive recurrence in the case $m_3 > m_2$. This proof is less elegant but more direct. It is of interest because it provides some estimates on the recurrence time. In the case that $m_3 > m_2$ the M/M/1 queue with arrival rate μ_2 and service rate μ_3 is unstable. For such an unstable M/M/1 queue, any busy period has a positive probability $q_\infty = 1 - \frac{m_2}{m_3}$ of serving an infinite number of customers, and of being non-ending (see for example textbook of Wolff [16]). Since an excursion of the process $X(t)$ away from R, which starts from a first state $(n_2 - 1, 1)$, will consist of a truncated busy period, the probability that all n_2 jobs will be served in the truncated busy period is $> q_\infty$. But in that case, the state in R to which $X(t)$ will return will be $(0, 0)$.

This proves that 0 is a positive recurrent state of the chain Y_s, since it can be reached from every state n_2 in a single step, with a probability $> \mu_2 q_\infty$. Hence Y_s is a positive recurrent chain. It also follows that X_r and $X(t)$ are recurrent.

To show that X_r and $X(t)$ are positive recurrent, we need however to show that $(0,0)$ is positive recurrent for X_r and $X(t)$. Let $T = \min\{t : t > 0, X(t-) \neq (0,0), X(t) = (0,0)\}$ be the first return time to $(0,0)$. We need to find an upper bound on the expected return time conditional on starting from $(0,0)$, $\mathbb{E}_{0,0}(T) = \mathbb{E}(T|X(0) = (0,0))$. Following $X(0) = (0,0)$ the process jumps to $X(T_0) = (1,0)$ after time $T_0 \sim \exp(\mu_1)$. Hence $\mathbb{E}_{0,0}(T) = \mathbb{E}(T_0) + \mathbb{E}(T|X(0) = (1,0)) = m_1 + \mathbb{E}_{1,0}(T)$. Hence, we wish to find an upper bound on $\mathbb{E}_{1,0}(T)$.

Starting from $(1,0)$ consider the joint sample paths of the three processes, $X(t), X_s, Y_k$ over $0 \le t \le T$. They start at $X(0) = (1,0), X_0 = (1,0), Y_0 = 1$. Thereafter $X(t) = X_{N(t)} \neq (0,0)$, $0 < t < T$ while $X(T) = X_{N(T)} = (0,0)$. Out of $s = 0,1,\dots,N(T)$, let $0 = S_0 < S_1 < \cdots < S_K = N(T)$ be the discrete times at which $X_s \in R$, that is $X_{S_k} = (Y_k, 0), k = 1,\dots,K$, with $Y_0 = 1,\ Y_k > 0, k = 2,\dots,K-1,\ Y_K = 0$. Define $\tau_1, \tau_2, \dots, \tau_K$ as the durations which satisfy $N(\tau_1 + \cdots + \tau_k -) < S_k,\ N(\tau_1 + \cdots + \tau_k) = S_k,\ k = 1,\dots,K$. Then $\tau_1, \tau_2, \dots, \tau_K$ are the durations between the times of jumps of $N(t)$ which take $X(t)$ into R or which leave $X(t)$ in R. Clearly, $\tau_1, \tau_2, \dots, \tau_K$ sum up to the return time T (conditional on $X(0) = (1,0)$).

Each random variables τ_k consists of a single Poisson interval, plus with probability μ_2 a busy period of the M/M/1 queue truncated by Y_{k-1}.

We are now ready to estimate $\mathbb{E}_{1,0}(T)$:

$$
\begin{aligned}
& \mathbb{E}_{1,0}(T) = \mathbb{E}(\tau_1 + \tau_2 + \cdots + \tau_K \,|\, Y_0 = 1) \\
= \ & \mathbb{E}\left(\sum_{k=1}^{\infty} 1_{K \ge k} \tau_k \,|\, Y_0 = 1\right) \\
= \ & \sum_{k=1}^{\infty} \mathbb{P}(K \ge k \,|\, Y_0 = 1) \mathbb{E}(\tau_k \,|\, K \ge k,\ Y_0 = 1) \\
= \ & \sum_{k=1}^{\infty} \mathbb{P}(K \ge k \,|\, Y_0 = 1) \sum_{y>0} \mathbb{P}(Y_{k-1} = y \,|\, K \ge k,\ Y_0 = 1) \mathbb{E}(\tau_k \,|\, Y_{k-1} = y,\ K \ge k,\ Y_0 = 1)
\end{aligned}
$$

but $K \ge k \Leftrightarrow Y_1 > 0, \dots, Y_{k-1} > 0$, and $\tau_k | Y_{k-1}$ is independent of $Y_1, \dots, Y_{k-2}$, hence:

$$
= \sum_{k=1}^{\infty} \mathbb{P}(K \ge k \,|\, Y_0 = 1) \sum_{y>0} \mathbb{P}(Y_{k-1} = y \,|\, K \ge k,\ Y_0 = 1) \mathbb{E}(\tau_k \,|\, Y_{k-1} = y).
$$

We now use the following facts:

(i) The probability that τ_k includes a busy period is μ_2. If a busy period starts then it will deplete buffer 2 and lead to state $(0,0)$ with a probability which exceeds $1 - \frac{m_2}{m_3}$. Therefore, $\mathbb{P}(Y_k > 0 \,|\, Y_{k-1} > 0) \le 1 - \mu_2(1 - \frac{m_2}{m_3}) = 1 - (\mu_2 - \mu_3)$. Hence:

$$
\mathbb{P}(K \ge k) \le (1 - (\mu_2 - \mu_3))^{k-1}.
$$

(ii) It is always the case that $Y_k \le Y_{k-1} + 1$, with equality only if a part completes processing at buffer 1 (which happens with probability μ_1). Therefore $Y_k \le k + 1$.

(iii) The expected length of a busy period of the M/M/1 queue truncated at y input parts is no more than the expected sum of the processing times of all the parts by both machines, that is: $m_2(y-1) + m_3 y$. τ_k includes a busy period with probability μ_2. Hence:

$$
\mathbb{E}(\tau_k \,|\, Y_{k-1} = y) \le 1 + \mu_2(m_2(y-1) + m_3 y) = y \frac{m_2 + m_3}{m_2}.
$$

we therefore get:

$$\begin{aligned}
&\mathbb{E}_{1,0}(T) \\
= \;& \sum_{k=1}^{\infty} \mathbb{P}(K \geq k \,|\, Y_0 = 1) \sum_{y>0} \mathbb{P}(Y_{k-1} = y \,|\, K \geq k,\, Y_0 = 1) \mathbb{E}(\tau_k \,|\, Y_{k-1} = y) \\
\leq \;& \sum_{k=1}^{\infty} (1 - (\mu_2 - \mu_3))^{k-1} \sum_{y=1}^{k} \mathbb{P}(Y_{k-1} = y \,|\, K \geq k,\, Y_0 = 1) \;\; y \frac{m_2 + m_3}{m_2} \\
\leq \;& \sum_{k=1}^{\infty} (1 - (\mu_2 - \mu_3))^{k-1} \;\; k \frac{m_2 + m_3}{m_2} \\
= \;& \frac{\mu_2 + \mu_3}{(\mu_2 - \mu_3)^2 \; \mu_3} < \infty.
\end{aligned}$$

This completes the proof.

Acknowledgements

I am grateful to an anonymous referee for suggesting the use of dominance in Section 3.3.

References

[1] I. Adan and G. Weiss: Analysis of a simple Markovian re-entrant line with infinite supply of work under the LBFS policy. Preprint (2004).

[2] P. Bremaud: *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues* (Springer, New York, 1999).

[3] R-R. Chen and S.P. Meyn: Value iteration and optimization of multiclass queueing networks. *Queueing Systems Theory and Applications,* **32** (1999), 65–97.

[4] J.G. Dai: On positive Harris recurrence of multi-class queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability,* **5** (1995), 49–77.

[5] J.G. Dai and G. Weiss: Stability and instability of fluid models for re-entrant lines. *Mathematics of Operations Research,* **21** (1994), 115–134.

[6] S. Foss and T. Konstatopoulos: An overview of some stochastic stability methods. *Journal of the Operations Research Society of Japan,* (2004), this issue.

[7] P. Glynn: *Stochastic Systems* (Stanford University Graduate Lecture Notes, 2000).

[8] J.M. Harrison: Brownian models of queueing networks with heterogeneous customer populations. In W. Fleming and P.L. Lions (eds.): *Proceedings of the IMA Workshop on Stochastic Differential Systems* (Springer, New York, 1988).

[9] G. Koole and A. Pot: Workload minimization in re-entrant lines. *European Journal of Operations Research* (Submitted, 2004).

[10] P.R. Kumar: Re-entrant lines. *Queueing Systems: Theory and Applications,* **13** (1993), 87–110.

[11] S. Kumar and P.R. Kumar: Performance bounds for queueing networks and scheduling policies. *IEEE Transactions on Automatic Control,* **38** (1994), 1600–1611.

[12] Y. Nazarathy: *Evaluation of on-line scheduling rules for high volume job shop problems, a simulation study* (M.A. Thesis, University of Haifa, 2001).

[13] Y. Nazarathy and G. Weiss: A simulation study of fluid imitation multi-class job-shop scheduling heuristics. (In preparation 2004).

[14] G. Weiss: On optimal draining of a fluid re-entrant line. In F.P. Kelly and R.J. Williams (eds.): *Proceedings of the IMA Workshop on Stochastic Networks* (Springer, New York, 1995), 93–105.

[15] G. Weiss: Scheduling and control of manufacturing systems — a fluid approach. *Proceedings of the 37 Allerton Conference* (Monticello, Illinois, 21–24 September 1999), 577–586.

[16] R.W. Wolff: *Stochastic Modeling and the Theory of Queues* (Prentice Hall, New Jersey, 1989).

Gideon Weiss
Department of Statistics
The University of Haifa
Mount Carmel 31905, Israel
E-mail: gweiss@stat.haifa.ac.il