Acta Universitatis Apulensis No. 28/2011
ISSN: 1582-5329 pp. 9-26

STUDY OF THE QUEUEING MODEL OF STORAGE AND TRANSMISSION BANDWIDTH ALLOCATION

Amina Angelika Bouchentouf and Faiza Belarbi

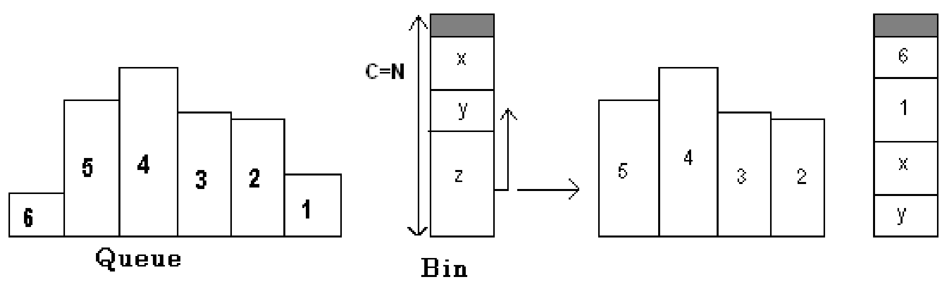
Abstract. In this paper, we study a queueing model of storage and transmission bandwidth allocation in computer and communication systems, so we are interested to the behavior of a network understanding a queue which contains a number infinite of items of size 1, 2, 3 and a bin of capacity $C = N$ (N is finite). An analysis of the stability properties of the bandwidth allocation algorithm First Fit for the distributions concentrated is made on these three sizes.
In this paper fluid limits are used to obtain the conditions of ergodicity.

2000 *Mathematics Subject Classification*: 37A25, 60K25.

1. Introduction

The traffics which cross the communication networks are of an extreme heterogeneity: data, voice, videos, etc, the requests in bandwidth are consequently highly variable. The impact of this extreme variability is, for the moment, rather little analyzed on the total behavior of a network.
Until the beginning of the Nineties, it was usually allowed that this type of situation was not in rupture with the framework of the traditional networks where the requests of the traffics to a given node are statistically close. It was however well-known that the state of the stability of these networks is much more difficult to characterize than that of the traditional networks. Studies of Rybko and Stolyar (1992), of Lu and KUMAR (1991) and of Bramson (1994), thereafter, completely changed this point of view. They show that statistical heterogeneity only can destabilize a network: even if, for each node of the network, the average charge of work which arrive is strictly smaller than its capacity, the network can oscillate so that the total number of requests in the network diverges. For these counterexamples, each node of the network empties an infinity of time but overall the network diverges. This situation is impossible in the traditional networks. These networks with heterogeneous traffics are gathered under name networks multi-class. Processes of Markov associated

these networks multi-class are very delicate to study, even with regard to the behavior macroscopic.
In this paper we consider a model of processing bandwidth requests. Requests have different bandwidth sizes. The arriving messages are of different nature, to be transmitted they require different throughput. In this paper we consider the First Fit algorithm: A message in the queue is allocated if its throughput is less than the available bandwidth at that time and none of the other messages arrived before it in the queue can be transmitted.
For convenience we shall use the bin packing terminology: the network is a bin of size $C = N$, messages are items and the bandwidth required by a message is the size of the item. Items have the same distribution as some random variable S_1. A stream of such items arrives at rate λ at the Bin and each item requires a service of mean 1. In this setting the First Fit algorithm can be described as follows: The sum of the items in the Bin is less than C the size of the Bin. Following every event (arrival or departure), the queue is scanned from the beginning in search of an item whose size is smaller than the empty space left in the bin. This procedure is repeated until the end of the queue is reached. An item in the bin is served at speed 1. As we shall see, the probabilistic description of this model is not easy to handle; it involves an infinite dimensional vector space (a space of stings).



A departure for the First Fit algorithm

In this paper we prove that under some hypothesis, the Markov process which defines a queueing system containing an infinite number of items of size 1 2, 3 and of a network of a finite capacity of size N describing the First Fit algorithm is ergodic if the "natural" condition is satisfied, i.e. if the load of the system is less than 1, see [1].
At first we give the estimation of the wasted space in the network when it works only with two sizes, this result help us to give the cindition of ergodicity for the

Markov process in the general case.
Our result concerning ergodicity use the formalism of fluid limits, so the next section recalls some of the results in this domain.

2. Fluid limits

A fluid limit scaling is a particular important way of scaling a Markov process. It is related to the first order behavior of the process, roughly speaking, it amounts to a functional law of large numbers for the system considered. It is in general quite difficult to have a satisfactory description of an ergodic Markov process describing a stochastic network. When the dimension of the state space is greater than 1, the geometry complicates a lot any investigation: Analytical tools such as Wiener-Hopf techniques for dimension 1 cannot be easily generalized to higher dimensions. It is possible nevertheless to get some insight on the behavior of these processes through some limit theorems. The limiting procedure investigated consists in speeding up time and scaling appropriately the process itself with some parameter. The behavior of such rescaled stochastic processes is analyzed when the scaling parameter goes to infinity. In the limit, one gets a sort of caricature of the initial stochastic process which is defined as a fluid limit. A fluid limit keeps the main characteristics of the initial stochastic process while some stochastic fluctuations of second order vanish with this procedure. In " good cases ", a fluid limit is a deterministic function, solution of some ordinary differential equation. As it can be expected, the general situation is somewhat more complicated. These ideas of rescaling stochastic processes have emerged recently in the analysis of stochastic networks, to study their ergodicity properties in particular. See Rybko and Stolyar [3] for example.
In this section $(X(t))$ is an irreducible Markov process on some countable space $\mathcal{S}$ embedded in a normed space. We assume that the bounded subsets of $\mathcal{S}$ are finite. The rescaled process is defined by

$$X_x(t) = \frac{\|X(t\|x\|)\|}{\|x\|}, \tag{1}$$

Since $X(0) = x,\ X_x(0) = 1$. The time variable are scaled by factor $\|x\|$. The following theorem is the combination of two results, one due to Filonov [2] and the other due to Rybko and Stolyar [3], it gives an ergodicity criterion.

Theorem 1 *If there exist an integrable stopping time U, constants K and $\epsilon > 0$ such that*

$$\limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(\|X(U)\|)}{\|x\|} \leq 1 - \epsilon, \tag{2}$$

$$\limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(U)}{\|x\|} \leq K, \tag{3}$$

the Markov process $(X(t))$ *is ergodic. If the variable* $X(t)$ *has a second moment for all* $t \geq 0$, *for a fixed* $K \geq 0$ *sufficiently large, the hitting time*

$$H = \inf\{t \geq 0 / \|X(t)\| \leq K\}$$

has a second moment of order $\|x\|^2$, *i.e.*

$$\limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(H^2)}{\|x\|^2} < +\infty \tag{4}$$

3. The string valued Markov process

In this paper we admit that the items arrive according to a Poisson process $\mathcal{N}_\lambda$ with parameter λ; for $t \geq 0$, the quantity $\mathcal{N}_\lambda(]0,t])$ denotes the number of arrivals between 0 and t. The capacity of the bin C is equal to N.

The size (S_i) of the items form an i.i.d. sequence with a common distribution $F(dx)$ given by

$$F(dx) = p\delta_1 + q\delta_2 + r\delta_3,$$

where δ_x is the Dirac measure in x and $p,\ q,\ r$ are non negative numbers such that $p + q + r = 1$. An item of size s will also be called an item s.

The set of the possible sizes is denoted by $\mathcal{T} = \{1,2,3\}$ and $\mathcal{T}^{(\mathbb{N})}$ is the set of finite vectors with coordinates in $\mathcal{T}$, if $x \in \mathcal{T}^{(\mathbb{N})}$, $\|x\|$ denotes the number of coordinates of x and $\emptyset$ is the empty vector.

The sojourn times of the items in the bin is an i.i.d. sequence with an exponential distribution with parameter 1. An element X of the state space $\mathcal{S}$ of the Markov process describing the storage process can be written as $X = (B, L)$, where L and B are elements of $\mathcal{T}^{(\mathbb{N})}$, the set of finite vectors with coordinates in $\mathcal{T}$. The vector $B = (b_j;\ j = 1, ..., \|B\|)$ describes the sizes of the items in the bin, since these items fit in the bin,

$$\sum_{j=1}^{\|B\|} b_j \leq C$$

and the vector $L = (l_i, i = 1, ..., \|L\|)$ represents the state of the queue. Since the First Fit algorithm scans the queue from the beginning in search of an item that may fit in the bin. Any item in the queue cannot fit in the bin, i.e. for any $i = 1, ..., \|L\|$ the following inequality holds

$$l_i + \sum_{j=1}^{\|B\|} b_j > C$$

The following notation will be employed in the remainder of this paper:
"$\overbrace{(a,a,...,a)}^{x\ items\ a}$"=[$x$ items a]", $\overbrace{(a,a,..,a,b,..,b)}^{x\ items\ a, y\ items\ b}$ " =[x items a, y items b].
Important point: The possible states which can be taken by the bin depend on its capacity.
The values which can be taken by B are

$\emptyset,\ (1),\ (1,1),\ (1,1,1), ..., [N\,\text{items of size}\,1].$
$(1,2),\ (1,1,2), ..., [(N-2)\,\text{items of size}\,1,2].$
$(1,3),\ (1,1,3), ..., [(N-3)\,\text{items of size}\,1,3].$
$(2),\ (2,2),\ (2,2,2), ..., [\frac{N}{2}\,\text{items of size}\,2]$, "If N is even" the bin is full
$(2),\ (2,2),\ (2,2,2), ..., [\frac{N-1}{2}\,\text{items of size}\,2]$, the bin is not full.
$(2,2,1),\ (2,2,1,1), ..., [2,2,(N-4)\,\text{items}\,1],\ C \geq 5.$
$(2,2,2,1,1), ..., [2,2,2,(N-6)\,\text{items}\,1],\ C \geq 9.$
$(2,2,2,1),\ (2,2,2,2,1), ..., [\frac{N-1}{2}\,\text{items}\,2,1],\ C \geq 7$ and the bin is full.
$\vdots$
$\vdots$
$(2,2,2,2,1,1),\ (2,2,2,2,2,1,1,1), ..., [\frac{l}{2}\,\text{items}\,2,(N-l)\,\text{items}\,1],\ l \geq 4,$
$C \geq 5$, and the bin is full.
$(3),\ (3,3),\ (3,3,3), ..., [\frac{N}{3}\,\text{items of size}\,3]$, if N is multiple of 3 and bin is full.
$(3),\ (3,3),\ (3,3,3), ..., [\frac{N-1}{3}\,\text{items of size}\,3]$, the bin is not full "there is an empty space of size 1".
$(3,3,1),\ (3,3,3,1), ..., [\frac{N-3}{3}\,\text{items}\,3,1],\ N$ is odd and the bin is full
$(3,3,3,1,1),\ (3,3,3,1,1,1),\ (3,3,1,1,1), ..., [x_1\,\text{items}\,3, y_1\,\text{items}\,1]$, such that the bin is full with x_1 items 3 and y_1 items 1.
$(3,2),\ (3,2,2),\ (3,2,2,2), ..., [3,\ \frac{N-3}{2}\,\text{items}\,2],\ C \geq 5,\ N$ is odd and the bin is full.
$(2,3,3),\ (2,3,3,3),\ (2,3,3,3,3), ..., [2,\ \frac{N-2}{3}\,\text{items}\,2]$, the bin is full.
$(2,2,3,3),\ (2,2,3,3,3),\ (2,2,2,3,3,3,3), ..., [x_2\,\text{items}\,2, y_2\,\text{items}\,3]$, such that the bin is full with x_2 items 2 and y_2 items 3.
$(3,2,1),\ (3,2,1,1), ..., [3,2,(N-4)\,\text{items}\,1]$, "$N$ is even", the bin is full.
$(3,3,2,1),\ (3,3,3,2,1), ..., [\frac{N-3}{3}\ \text{items}\,3,2,1]$, "$N$ is multiple of 3", the bin is full.
$(2,2,3,1),\ (2,2,2,3,1), ..., [\frac{N-4}{2}\,items\,2,3,1]$, the bin is full.
$(3,3,2,2,1,1),\ (3,3,2,2,2,1,1),\ (3,3,3,2,2,1,1,1), ..., [x_3\,\text{items}\,1, y_3\,\text{items}\,2,$
$z_3\,items\,3]$, such that the bin is full with x_3 items 1, y_3 items 2 and z_3 items 3.

Notice that the order the components in B has no importance for the dynamic of the system, for this reason we shall consider B as a set. The order is important for the vector L since the First Fit discipline checks if the first coordinate l_1 fits in the bin, then the coordinates l_2, l_3, and so on. The vector L is a string of 1, 2 and

3.
We shall say that the model is stable when $(X(t))$ is an ergodic Markov process on $\mathcal{S}$. In Dantzer [1] it has been proved that the condition $\lambda \mathbb{E}(\mathrm{S}_1) \leq C$ is necessary for the stability of the system, i.e. the Markov process $(X(t))$ is transient if $\lambda \mathbb{E}(\mathrm{S}_1) > C$.

Definition 1 *The norm $\|X\|$ of the state $X = (B, L) \in \mathcal{S}$ is the sum of the $\|B\|$ and $\|L\|$. The load $W(X(t))$ of $(X(t)) = (B(t), L(t)) = ((b_i(t)), (l_j(t)))$ is defined as*

$$W(X(t)) = \sum_{i=1}^{\|B(t)\|} b_i \sigma_i^0(t) + \sum_{j=1}^{\|L(t)\|} l_j \sigma_j(t)$$

where, for $i \in \{1, ..., \|B\|\}$ and $j \in \{1, ..., \|L\|\}$, $\sigma_i^0(t)$ is the residual service time of the item $b_i(t)$ and $(\sigma_j(t))$ the service time of the item $l_j(t)$.

Notice that the load of the system increases at rate $\lambda \mathbb{E}(\mathrm{S}_1)$ in average and decreases at most at rate N.

4. Ergodicity of the system under the natural condition

In this case we are not interested too much by the structure of L-component of the initial state. The following lemma gives an estimation of the wasted space when there are only two possible sizes: 1 and 2.

Lemma 1 *Under the conditions $\lambda \mathbb{E}(\mathrm{S}_1) < N$, if $r = 0$ (only items 1 and 2 arrive) and*

$$\tau = inf\{t \leq 0 / \|L(t)\| = 0\} \text{ and } D = \int_0^\tau \mathbb{1}_{\{b_1(t)+...+b_{\|B(t)\|}(t)<N\}} dt,$$

then there exist some constants K_1 and K_2 such that

$$\mathbb{E}_x(D) \leq K_1 \log(1 + \|x\|) + K_2.$$

for any $x = (l, b) \in \mathcal{S}$.

Proof. We denote by the variable D : duration of time during which the bin is not full during a busy period. Notice first that there is no waste of space as long as there are items 1 in the L-component of $(X(t))$. We can therefore assume that l is a string of items 2.
In this context the only possibility to waste space with a non empty queue is when the state $(B(t))$ of the bin is $[(N-1)$ items of size 1], or [2, $(N-3)$ items of size 1], [$\frac{N-2}{2}$ items of size 2, 1]" with N even", [$\frac{N-1}{2}$ items of size 2]" with N is odd", or [q_1 items 1,q_2 items 2] such that $q_1 \neq q_2$ $q_1 \geq 2$, $q_2 \geq 2$, and if $q_1 = q_2$ [$\frac{N-1}{3}$ items 1, $\frac{N-1}{3}$ items 2]; (3 is necessarily divisor of $N-1$).

(In all these cases there is an empty space of size 1).
We set $A_0 = l$ and $T_0 = 0$ and by induction we define
$T_{n+1} = \inf\{t > T_n : C(t-) = N, C(t) < N,$
and all the items 2 present at time T_n are served at time t$\}$
with $C(s) = b_1(s) + ... + b_{\|B(s)\|}(s)$ and $A_n = \|L(T_n)\|$ for $n \geq 1$. Notice that $L(T_n)$ is necessarily a (possibly empty) string of items 2. The sequence $(B(T_n), A_n)$ is clearly a Markov chain.
• If b the initial state of the bin is $[(N-1)$ items of size 1] (there is an empty space of size 1). As long there is at least an item 1 in the queue, because of the First Fit discipline, the items 2 are ignored. Since $\lambda p \leq \lambda \mathbb{E}(\mathrm{S}_1) < N$, after an integrable amount of time not depending on $\|l\|$, at least Two places will be vacant in the bin and consequently an item 2 will enter the bin.
In this situation the number of items 1 is the number of customers of an $M/M/N$ queue (N servers)with parameter λp for the input rate and 1 for the service rate.
• If b is $[2, (N-3)$ items of size 1]. We have two cases to discuss.

1. $\lambda p < 2$. This condition clearly implies that, with probability 1, at some time there will no item 1 in the system and, consequently, an item 2 will enter the bin, then all the other items 2 will enter consecutively . The expected value of this duration of time is easily seen to be bounded with respect to $\|l\|$. Starting from that time, only items 2 are served as long as the initial items 2 are present (since these items are located at the beginning of the queue, the First Fit algorithm selects them). When the initial items 2 have been served the queue are an i.i.d. strings of items 2 and 1. At that moment an item 1 will enter the bin then two,..., until $(N-2)$ items 1 "$[(N-2)$ items 1, 2]. Later, when the number of items 1 in the system is $(N-3)$ the system will waste some space, this is precisely the definition of time T_1, A_1 is the number of items at that time.

2. $\lambda p > 2$. This condition implies that, if the state of the bin does not change, the arriving items 1 will saturate $(N-2)$ places in the bin. In this case, the number of items 1 is the number of customers of transient $M/M/(N-2)$ queue starting with $(N-3)$ customers (in the bin at time 0). A change in the state of the bin may occur only if this transient queue is empty.

   a) The $M/M/(N-2)$ queue never reaches the empty state. After some small amount of time (i.e. its expected value is bounded with respect to $\|l\|$), the bin will be full with an item 2 and $(N-2)$ items 1. The condition $\lambda \mathbb{E}(\mathrm{S}_1) < N$ implies that $\lambda q < 1$ $(\lambda p > 2)$, therefore, with probability 1 after some period of time the system will not contain any item 2. At that moment the state of the bin will be [N items 1]. (Recall that $\lambda p \leq$

$\lambda \mathbb{E}(\mathrm{S}_1) < N$). It is easily seen that, with probability 1, the total number of items 1 will be less than $(N-2)$. An item 2 will be in the bin at that time, this is the starting situation.

b) The $M/M/(N-2)$ queue reaches the empty state: we have two cases to discuss:
• If N is even the queue reaches the empty state; $\frac{N}{2}$ items 2 occupy the bin. The initial items 2 are served. In this situation T_1 is the next time where there is some wasted space.
• If N is odd the queue does not reaches the empty space; After an integrable amount of time we will have "[2, $(N-3)$ items 1]⇒ [2,$(N-2)$ items 1]". Then after a moment the bin will be [$\frac{N-1}{2}$ items 2], and T_1 is the time there is some wasted space, after that all items 2 were served.

Notice that the case a) occurs only a geometrically distributed number of times. Hence, the duration of time between time 0 and T_1 when the bin is not full has a bounded expected value (with respect to $\|l\|$).

Using the proposition 5 of Dantzer [1], it is easy to check that there exists some constant $c > 0$ such that the following convergence holds in L_1 and almost surely

$$\lim_{\|x\| \to +\infty} \frac{\mathbb{E}_x(T_1)}{\|x\|} = c.$$

If I is the duration of time between 0 and T_1 when the bin is not full, the expected value of the load at time T_1 satisfies the following inequality (all the service times are i.i.d. exponentially distributed random variables with parameter 1),

$$\mathbb{E}_x(W(X(T_1))) \leq \mathbb{E}_x(W(x)) + \mathbb{E}\left(\sum_{i=1}^{\mathcal{N}_\lambda([0,T_1])} \mathrm{S}_i\right) - N\mathbb{E}_x(T_1 - I)$$

using Wald's formula (T_1 is a stopping time), we get

$$\mathbb{E}_x(W(X(T_1))) \leq \mathbb{E}_x(W(x)) + (\lambda \mathbb{E}(\mathrm{S}_1) - N)\mathbb{E}_x(T_1) + N\mathbb{E}_x(I),$$

since there are no items 1 in the queue at 0 and T_1 (let us return to the case to a), we have
$\mathbb{E}_x(W(X(0))) = (N-2)\|x\|$ and $\mathbb{E}_x(W(X(T_1))) = (N-1) + (N-2)\mathbb{E}_x(A_1)$.
The quantity $\mathbb{E}_x(I)$ being bounded with respect to $\|x\|$, it follows that

$$(N-2) \limsup_{\|x\| \to +\infty} \frac{\mathbb{E}_x(A_1)}{\|x\|} = \limsup_{\|x\| \to +\infty} \frac{\mathbb{E}_x(W(X(T_1)))}{\|x\|} \leq (N-2) + c(\lambda \mathbb{E}(\mathrm{S}_1) - N).$$

Consequently, there exist a_0 and $\alpha_1 < 1$ such that for $\|x\| > a_0$,

$$\mathbb{E}_x(A_1) \leq \alpha_1 \|x\|, \tag{5}$$

Where $\alpha_1 = 1 + \frac{c}{(N-2)}(\lambda \mathbb{E}(\mathrm{S}_1) - N)$

$$\gamma_1 = -\log\left(\frac{1 + \alpha_1\|x\|}{1 + \|x\|}\right) > 0. \tag{6}$$

If we set

$$\nu_1 = \inf\{n \geq 1/A_n \leq a_0\},$$

the sequence

$$(Z_n) = (\log(1 + A_{n\wedge\nu_1}) + \gamma_1(n \wedge \nu_1)).$$

is a super-martingale. Indeed, if $(\mathcal{F}_n)$ is the natural filtration associated to the sequence (A_n), on the event $\{\nu_1 > n\}$ the Markov property gives the equality

$$\begin{aligned} \mathbb{E}(Z_{n+1}/\mathcal{F}_n) - Z_n &= \mathbb{E}_{(B(T_n),A_n)}(\log(1 + A_1)) - \log(1 + A_n) + \gamma_1 \\ &\leq \log(1 + \mathbb{E}(A_1/A_n)) - \log(1 + A_n) + \gamma_1 \leq 0, \end{aligned}$$

by Jensen's inequality and the relations (5) and (6). Consequently $\mathbb{E}(Z_n) \leq Z_0$, hence $\gamma_1\mathbb{E}(n \wedge \nu_1) \leq \mathbb{E}(Z_n) \leq Z_0$, by letting n go to infinity, we get

$$\mathbb{E}_x(\nu_1) \leq \frac{\log(1 + \|x\|)}{\gamma_1} \tag{7}$$

For $n \geq 1$, the bin is always full between T_n and T_{n+1}, except during some integrable period whose expected value is bounded with respect to size of the initial state. By Wald's formula, the contribution of the ν_1 cycles in the integral defining D is bounded by $K\mathbb{E}_x(\nu_1) \leq K\log(1 + \|x\|)/\gamma_1$, for some constant K.
Since $\lambda\mathbb{E}(\mathrm{S}_1) < 1$, the proposition 6 of Dantzer [1] shows that the system is ergodic. Consequently, starting from the state $(B(T_{\nu_1}), A_{\nu_1})(\leq a_0))$, the hitting time of the empty state $\emptyset$ is integrable and with an expected value bounded with respect to $\|x\|$. Therefore, the expected value of the contribution of this period in the integral defining D is bounded with respect to $\|x\|$.
• the state of the bin is $[\frac{N-2}{2}$ items 2, 1] "N is even" .
We have two cases to discuss.

1) $\lambda p < 2$. This condition implies that with probability 1, after an integrable amount of time,there will no items 1 in the system, an item 2 will enter the bin then all the other items 2 will be served consecutively. The expected value of this duration of time is easily seen to be bounded with respect to $\|l\|$. Starting from that time, the queue will be an i.i.d. string of items 2 and 1. At that moment an item 1 will enter the bin, then two items 1, at that moment

the bin will be full with two items 1 and $\frac{N-2}{2}$ items 2. Later, with probability 1, when the number of the items 1 in the system is 1 the system will waste some space, this is precisely the definition of T_1, A_1 is the number of items at that moment.

2) $\lambda p > 2$. This condition implies that, if the state of the bin does not change, the arriving items 1 will saturate 2 places in the bin. In this case, the number of items 1 is the number of customers of transient $M/M/2$ queue starting with one customer (in the bin at time 0). A change in the state of the bin may occur only if this transient queue is empty.

   a) The $M/M/2$ queue never reaches the empty state. After some small time (i.e its expected value is bounded with respect to $\|l\|$), the bin will be full with $\frac{N-2}{2}$ items 2 and two items 1. The condition $\lambda\mathbb{E}(\mathrm{S}_1) < N$ implies that $\lambda q < 1$, therefore, with probability 1 after some period of time the system will not contain any item 2. At that moment the state of the bin will be $[N$ items $1]$. Recall that $\lambda p \leq \lambda\mathbb{E}(\mathrm{S}_1) < N$. It is easily seen that, with probability 1, the total number of items 1 will be less than 2. $\frac{N-2}{2}$ items 2 will be in the bin at that time, and this is the starting situation.

   b) The $M/M/2$ reaches the empty state. $\frac{N}{2}$ items 2 occupy the bin. The initial items 2 are served. In this situation, T_1 is the next time there is some wasted space.

Then we continue the demonstration in the same way as previously except that there are small changes:
since there are no items 1 in the queue at 0 and T_1 (let us return to the case to a), we have
$\mathbb{E}_x(W(X(0))) = 2\|x\|$ and $\mathbb{E}_x(W(X(T_1))) = (N-1) + 2\mathbb{E}_x(A_1)$.
The quantity $\mathbb{E}_x(I)$ being bounded with respect to $\|x\|$, it follows that

$$2 \limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(A_1)}{\|x\|} = \limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(W(X(T_1)))}{\|x\|} \leq 2 + c(\lambda\mathbb{E}(\mathrm{S}_1) - N),$$

Consequently, there exist a_0 and $\alpha_2 < 1$ such that for $\|x\| > a_0$,

$$\mathbb{E}_x(A_1) \leq \alpha_2\|x\|, \tag{8}$$

Where $\alpha_2 = 1 + \frac{c}{2}(\lambda\mathbb{E}(\mathrm{S}_1) - N)$

$$\gamma_2 = -\log\left(\frac{1+\alpha_2\|x\|}{1+\|x\|}\right) > 0. \tag{9}$$

Then we will not change anything in the remainder demonstration.

•If the state of the bin is $[\frac{N-1}{2}$ items 2]" N is odd". We have two cases to discuss.

1) $\lambda p < 2$. After an integrable amount of time, with probability 1, an item 2 will enter the bin then all the other items 2 will be served consecutively. The expected value of this duration of time is easily seen to be bounded with respect to $\|l\|$. Starting from that time, the queue will be an i.i.d. string of items 2 and 1. At that moment an item 1 will enter the bin with probability 1, the bin will be full with item 1 and $\frac{N-1}{2}$ items 2. Later when the system does not contain any item 1, this later waste some space, this is precisely the definition of T_1, A_1 is the number of items at that time.

2) $\lambda p > 2$. This condition implies that, if the state of the bin does not change, the arriving items 1 will saturate one place in the bin "$[\frac{N-1}{2}$ items 2, 1]. In this case, the number of items 1 is the number of customers of transient $M/M/1$ queue starting with 0 costumer (in the bin at time 0). A change in the state of the bin may occur only if this transient queue is empty.

   a) The $M/M/1$ queue never reaches the empty state. After some small time (i.e its expected value is bounded with respect to $\|l\|$), the bin will be full with $\frac{N-1}{2}$ items 2 and one item 1. The condition $\lambda \mathbb{E}(\mathrm{S}_1) < N$ implies that $\lambda q < 1$, therefore, with probability 1 after some period of time the system will not contain any item 2. At that moment the state of the bin will be [N items 1]. Recall that $\lambda p \leq \lambda \mathbb{E}(\mathrm{S}_1) < N$. It is easily seen that, with probability 1, the total number of items 1 will be less than 2. $\frac{N-1}{2}$ items 2 will be in the bin at that time. This is the starting situation.

   b) The $M/M/1$ queue reaches the empty state. In this case we can not say that the bin reaches the empty space because, the items 1 will be served and $\frac{N-1}{2}$ items 2 occupy the bin. In this situation, T_1 is the next time there is some wasted space.

Then we continue the demonstration in the same way as previously except that there are small changes which are the following:
since there are no items 1 in the queue at 0 and T_1, we have
$\mathbb{E}_x(W(X(0))) = 1\|x\|$ and $\mathbb{E}_x(W(X(T_1))) = (N-1) + 1\mathbb{E}_x(A_1)$.
The quantity $\mathbb{E}_x(I)$ being bounded with respect to $\|x\|$, it follows that

$$1 \limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(A_1)}{\|x\|} = \limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(W(X(T_1)))}{\|x\|} \leq 1 + c(\lambda \mathbb{E}(\mathrm{S}_1) - N),$$

Consequently, there exist a_0 and $\alpha_3 < 1$ such that for $\|x\| > a_0$,

$$\mathbb{E}_x(A_1) \leq \alpha_3 \|x\|, \tag{10}$$

Where $\alpha_3 = 1 + c(\lambda \mathbb{E}(\mathrm{S}_1) - N)$

$$\gamma_3 = -\log\left(\frac{1 + \alpha_3\|x\|}{1 + \|x\|}\right) > 0. \tag{11}$$

The remainder of the demonstration is the same one as in the previously cases.
• the state of the bin is $[\frac{N-1}{3}$ items $1, \frac{N-1}{3}$ items 2].

$*N$ **is odd such that** 3 **is divisor of** $N - 1$.

1) $\lambda p < 2$. This condition implies that from a certain moment with a probability 1 we will not have any more items of size 1 in the system, at that time the initial items of size 2 will insert in the bin. Later the state of the bin will reach $[\frac{N-1}{2}$ items 2]. Let us notice that we have already discussed this point, therefore we will follow the same discussion.

2) $\lambda p > 2$. This condition implies that, if the state of the bin does not change, the arriving items 1 will saturate $\frac{N+2}{3}$ places in the bin "the bin is full". In this case, the number of items 1 is the number of customers of transient $M/M/\frac{N+2}{3}$ queue starting with $\frac{N-1}{3}$ customers (in the bin at time 0). A change in the state of the bin may occur only if this transient queue is empty.

   a) The $M/M/\frac{N+2}{3}$ queue never reaches the empty state. After some small of time (i.e its expected value is bounded with respect to $\|l\|$), the bin will be full with $\frac{N+2}{3}$ items 1 and $\frac{N-1}{3}$ item 2. The condition $\lambda \mathbb{E}(\mathrm{S}_1) < N$ implies that $\lambda q < 1$, therefore, with probability 1 after some period of time the system will not contain any item 2. At that moment the state of the bin will be $[N$ items 1]. Recall that $\lambda p \leq \lambda \mathbb{E}(\mathrm{S}_1) < N$. It is easily seen that, with probability 1, the total number of items 1 will be less than $\frac{N+2}{3}$. At that moment $\frac{N-1}{3}$ items 2 will be in the bin at that time. This is the starting situation.

   b) The $M/M/\frac{N+2}{3}$ queue reaches the empty state. As N is odd this transient queue does not reach the empty space, after amount of time all the items 1 will be served consecutively, $\frac{N-1}{2}$ items 2 occupy the bin. The initial items 2 are served. T_1 is this moment when there is some wasted space.

since there are no items 1 in the queue at 0 and T_1, we have
$\mathbb{E}_x(W(X(0))) = \frac{N+2}{3}\|x\|$ and $\mathbb{E}_x(W(X(T_1))) = (N-1) + \frac{N+2}{3}\mathbb{E}_x(A_1)$.
The quantity $\mathbb{E}_x(I)$ being bounded with respect to $\|x\|$, it follows that

$$\frac{N+2}{3}\limsup_{\|x\|\to+\infty}\frac{\mathbb{E}_x(A_1)}{\|x\|} = \limsup_{\|x\|\to+\infty}\frac{\mathbb{E}_x(W(X(T_1)))}{\|x\|} \leq \frac{N+2}{3} + c(\lambda\mathbb{E}(\mathrm{S}_1) - N),$$

Consequently, there exist a_0 and $\alpha_4 < 1$ such that for $\|x\| > a_0$,

$$\mathbb{E}_x(A_1) \leq \alpha_4\|x\|, \tag{12}$$

Where $\alpha_4 = 1 + \frac{3c}{N+2}(\lambda\mathbb{E}(\mathrm{S}_1) - N)$

$$\gamma_4 = -\log\left(\frac{1+\alpha_4\|x\|}{1+\|x\|}\right) > 0. \tag{13}$$

$*N$ **is even such that** 3 **is divisor of** $N-1$ **.**
we have two points to discuss:

1) $\lambda p < 2$. This implies that at some time with a probability 1 we will not have items of size 1 in the system, at that time the initial items of size 2 will insert in the bin. Later the state of the bin will reach [$\frac{N}{2}$ items 2]"the bin is full". The expected value of this duration of time is bounded with respect to $\|l\|$. And starting from that time only items 2 are served. When the initial items 2 are served the queue is an i.i.d string of items 1 and 2. Later, when the number of items of items 1 in the system is $\frac{N-1}{3}$, the system will waste some space, this is precisely the definition of T_1, A_1 is the number of items at that time.

1) $\lambda p > 2$. This condition implies that, if the state of the does not change, the arriving items 1 will saturate $\frac{N+2}{3}$ places in the bin, In this case, the number of items 1 is the number of customers of transient $M/M/\frac{N+2}{3}$ queue starting with $\frac{N-1}{3}$ customers (in the bin at time 0). A change in the state of the bin may occur only if this transient queue is empty. Thus we have two cases to discuss, "$M/M/\frac{N+2}{3}$ reaches the empty state or not"And for this reason we will follow the same discussion as we have already fact it in the case "N is odd",except that in this case, the transient queue "$M/M/\frac{N+2}{3}$ can reach the empty state: $\frac{N}{2}$ items 2 occupy the bin. The initial items 2 are served. In this situation T_1 is the next time where there is some wasted space.

- the state of the bin is [q_1 items 1, q_2 items 2]"$q_1 \neq q_2$"

1) $\lambda p < 2$. This condition implies that from a certain moment with a probability 1 we will not have any more items of size 1 in the system, at that time some items of size 2 will insert in the bin. The expected value of this duration of time is easily seen to be bounded with respect to $\|l\|$. Starting from that time the First Fit algorithm selects only items 2, when all the initial items 2 have been served the queue is an i.i.d string of items 1 and 2. Then with probability 1, at least $q_1 + 1$ items 1 enter in the bin. Later when the number of items 1 in the system is q_1 the system will waste some space, this is precisely the definition of time T_1, A_1 is the number of items at that time.

2) $\lambda p > 2$. This condition implies that, if the state of the bin does not change, the arriving items 1 will saturate $q_1 + 1$ places in the bin "the bin is full". In this case, the number of items 1 is the number of customers of transient $M/M/q_1+1$ queue starting with q_1 customers (in the bin at time 0). A change in the state of the bin may occur only if this transient queue is empty.

   a) The $M/M/q_1 + 1$ queue never reaches the empty state. After some small time (i.e its expected value is bounded with respect to $\|l\|$), the bin will be full with $q_1 + 1$ items 1 and q_2 item 2. The condition $\lambda \mathbb{E}(\mathrm{S}_1) < N$ implies that $\lambda q < 1$, therefore, with probability 1 after some period of time the system will not contain any item 2. At that moment the state of the bin will be [N items 1]. Recall that $\lambda p \leq \lambda \mathbb{E}(\mathrm{S}_1) < N$. It is easily seen that, with probability 1, the total number of items 1 will be less than $q_1 + 1$. q_2 items 2 will be in the bin at that time. This is the starting situation.

   b) The $M/M/q_1 + 1$ queue reaches the empty state. "there are only few cases where this transient queue reaches the empty space". The cases where the initial items 2 occupy the bin such that there is no empty space. In this situation T_1 is the time there is some wasted space, after that all items 2 were served.

since there are no items 1 in the queue at 0 and T_1, we have
$\mathbb{E}_x(W(X(0))) = (q_1 + 1)\|x\|$ and $\mathbb{E}_x(W(X(T_1))) = (N - 1) + (q_1 + 1)\mathbb{E}_x(A_1)$.
The quantity $\mathbb{E}_x(I)$ being bounded with respect to $\|x\|$, it follows that

$$(q_1 + 1) \limsup_{\|x\| \to +\infty} \frac{\mathbb{E}_x(A_1)}{\|x\|} = \limsup_{\|x\| \to +\infty} \frac{\mathbb{E}_x(W(X(T_1)))}{\|x\|} \leq (q_1 + 1) + c(\lambda \mathbb{E}(\mathrm{S}_1) - N),$$

where $W(.)$ is the load (see Definition 1). Consequently, there exist a_0 and $\alpha_5 < 1$ such that for $\|x\| > a_0$,

$$\mathbb{E}_x(A_1) \leq \alpha_5 \|x\|, \tag{14}$$

Where $\alpha_5 = 1 + \frac{c}{(q_1+1)}(\lambda \mathbb{E}(\mathrm{S}_1) - N)$

$$\gamma_5 = -\log\left(\frac{1 + \alpha_5\|x\|}{1 + \|x\|}\right) > 0. \tag{15}$$

the rest of the proof is the same as in previous cases. $\square$
Now we consider the general case with three sizes. The condition $\lambda \mathbb{E}(\mathrm{S}_1) < N$ turns out to be sufficient for ergodicity when $\lambda p > (N-3)$.

Proposition 1 *If $\lambda \mathbb{E}(\mathrm{S}_1) < N$ and $\lambda p > (N-3)$, then $(X(t))$ is an ergodic Markov process.*

Proof. Let $(x_n) = (b_n, l_n)$ a sequence of $\mathcal{S}$ whose norm converges to infinity.
Since the number of configurations in the bin is finite, by taking subsequences we can suppose that the sequence of the initial states in the bin (b_n) is constant, hence $(x_n) = (b, l_n)$ using Proposition 4 of Dantzer [1], we can assume that for the states (x_n) the bin is not full. Consequently (l_n) does not contain any item 1, it is a sequence of strings of items 2 and 3.
We denote by τ the first time when the bin is not full after all the initial items 2 and 3 have left the system;
$\tau = \inf\{t / |B(t)| < N, \quad \text{after all the initial items 2 and 3 have left the system}\}$
τ is clearly a (possibly infinite) stopping time. If D is the duration of time between time 0 and τ during which the bin is not full, we claim that D is integrable and, moreover,

$$\lim_{n \to +\infty} \frac{\mathbb{E}_{x_n}(D)}{\|x_n\|} = 0.$$

If our assertion is true, between 0 and τ the load of the system is decreased at rate N, except during some periods of total duration D, i.e. for $t \geq 0$ we have

$$\sum_{i=1}^{\|b\|} b_i \sigma_i^0 + \sum_{i=\|b\|+1}^{\|l_n\|+\|b\|} l_{n,i} \sigma_i^0 + \sum_{i=1}^{\mathcal{N}_\lambda(]0, t\wedge\tau])} \mathrm{S}_i \sigma_i - N(t \wedge \tau - D) \geq 0,$$

the sequences (σ_i) and (σ_i^0) are the respective service times of the arriving items and of the initial items. These variables are independent and exponentially distributed with parameter 1. Taking the expectation of the two members of this inequality, we get the relation

$$\mathbb{E}_{x_n}(t \wedge \tau)(N - \lambda \mathbb{E}(\mathrm{S}_1)) \leq \|x_n\| + N \mathbb{E}_{x_n}(D).$$

By letting t go to infinity, according to our assumption on $(\mathbb{E}_{x_n}(D))$ we obtain the inequality

$$\limsup_{n\to+\infty} \frac{\mathbb{E}_{x_n}(\tau)}{\|x_n\|} \leq \frac{1}{N-\lambda\mathbb{E}(\mathrm{S}_1)}. \tag{16}$$

In the same manner, the following inequality holds,

$$\mathbb{E}_{x_n}\left(W(X(t\wedge\tau))\right) \leq W(x_n) + (\lambda\mathbb{E}(\mathrm{S}_1)-N)\mathbb{E}_{x_n}(\tau\wedge t) + N\mathbb{E}_{x_n}(D);$$

Fatou's Lemma and Lebesgue's Theorem give when t goes to infinity

$$\mathbb{E}_{x_n}\left(W(X(\tau))\right) \leq W(x_n) + (\lambda\mathbb{E}(\mathrm{S}_1)-N)\mathbb{E}_{x_n}(\tau) + N\mathbb{E}_{x_n}(D),$$

since all the initial items 2 and 3 are served at time τ, consequently

$$\limsup_{n\to+\infty} \frac{\mathbb{E}_{x_n}(W(X(\tau)))}{W(x_n)} \leq \frac{\lambda\mathbb{E}(\mathrm{S}_1)}{N} < 1. \tag{17}$$

By using the fact that $W(.)$ is an equivalent norm to $\|.\|$ on $\mathcal{S}$, relations (16) and (17) and Theorem 1 show that the Markov process $(X(t))$ is ergodic.
All we have to prove now that $(\mathbb{E}_{x_n}(D))$ is negligible with respect to $\|x_n\|$ when n is large. There are several possibilities for b, the common content of the bin for the initial states (x_n). We discuss the different cases, throughout this discussion, we shall say that random variable H is a "bounded integrable variable" if the sequence $(\mathbb{E}_{x_n}(H))$ is bounded with respect to $\|x\|$.

(1) If b is [$(N-1)$ items 1] or [N items 1]. As long as there is at least an item 1 in the queue, all the other items are ignored. The condition $\lambda\mathbb{E}(\mathrm{S}_1) < N$ implies that $\lambda p < N$. From the point of view of the items 1, the system is a stable $M/M/N$ queue. Hence the first time there will be at least two empty places is a bounded integrable variable $(\lambda p > N-3)$. At that time, an item 2 will be inserted in the bin b=[2,$(N-2)$ items 1]. Notice that for this period, the duration of time during which the bin is not full is a bounded integrable variable.

(2) If b has some items 2

a) If **by adding an item of size 3 one will exceed the maximum capacity of the bin which is N then:**
The items 3 are not taking in account, consequently a string of items 3 builds up at the beginning of the queue. Since the condition $\lambda\mathbb{E}(\mathrm{S}_1) < N$ implies $\lambda p + 2q < N$, the system with the items 1 and 2 is stable (see Dantzer [1]. Lemma 1 shows that until an item 3 enters the bin the wasted space is negligible compared to the number of initial items 2.

b) **else** : All the items 3 are selected by First Fit algorithm. Since the condition $\lambda \mathbb{E}(\mathrm{S}_1) < N$ implies $\lambda q + 3r < N$, the system with the items 2 and 3 is stable, after amount of time, the condition $\lambda p > N - 3$ implies that the residual space in the bin left by items 2 or 3 is saturated by 1. Consequently the duration of time the bin is not full is a bounded integrable variable.

(3) If b contains some items 3.

a) If **by adding an item of size** 3 **one will exceed the maximum capacity of the bin which is** N **then:**
All the other items 3 are ignored. At that time the items 2 are selected by First Fit algorithm. Then when this is finished, we note that the condition $\lambda p > N - 3$ implies that the duration of time the bin is not full is a bounded integrable variable.

b) **If the bin contains items** 3 **such that by adding an item** 2 **or** 3 **one can exceed** N :
Clearly one can assume that l_n is string of items 3, otherwise, if at some moment 2 enters the bin, the cases considered above some show that items 1 or 2 will be cleared from the system until some items 3 are in the bin. The condition $\lambda p > (N - 3)$ implies that the residual space in the bin left by the items 3 is saturated by the items 1. Finally we will end to the same result.

This discussion shows that the assertion is proved and consequently, the proposition. □

The result of this proposition is fairly easy to understand: under the condition $\lambda p > N - 3$, basically there is no waste of space so that the natural condition $\lambda \mathbb{E}(\mathrm{S}_1) < N$ is sufficient for the ergodicity of $(X(t))$. Notice however that the proof of this intuitive result (Lemma 1 and Proposition 1) has required the detailed analysis of the possible evolution starting from a given initial state.

References

[1] Jean-François Dantzer, Mostapha Haddani and Philippe Robert, *On the stability of a bandwidth packing algorithm*, Probability in the Engineering and Informational Sciences, 41, 1, (2000), 1, 57-79.

[2] Y.Filonov, *A criterion for the ergodicity of discrete homogeneous Markov chains,* akademiya Nauk Ukrainskoi SSR. Institut Matematiki. Ukrainskii Matematicheskii Zhurnal, 41, 10, (1989), 1421-1422.

[3] A. N. Rybko and A. L. Stolyar, *On the ergodicity of random processes that describe the functioning of open queueing networks,* Problems on Information Transmission, 28, 3, (1992), 3-26.

Amina Angelika BOUCHENTOUF
Department of Mathematics
University Djillali LIABES of Sidi Bel Abbes,
B. P. 89, Sidi Bel Abbes 22000, Algeria.
email:*bouchentouf_amina@yahoo.fr*

Faiza BELARBI
Department of Mathematics
University Djillali LIABES of Sidi Bel Abbes,
B. P. 89, Sidi Bel Abbes 22000, Algeria.
email:*faiza_belarbi@yahoo.fr*