ITC19/ Performance Challenges for Efficient Next Generation Networks 1365-1374
LIANG X.J. and XIN Z.H.(Editors)
V.B. IVERSEN and KUO G.S.(Editors)
Beijing University of Posts and Telecommunications Press

Does repacking improve performance of multiclass loss networks with overflow routing?

Lise George [1], Matthieu Jonckheere [1], and Lasse Leskelä [2]

[1] France Telecom R&D,
38-40 rue du general Leclerc, 92794 Issy-les-Moulineaux Cedex, France
email: {lise.george,matthieu.jonckheere}@francetelecom.com,
[2] Institute of Mathematics, Helsinki University of Technology,
P.O. Box 1100, FI-02015 TKK, Finland
email: lasse.leskela@iki.fi

Abstract. This paper considers multiclass loss networks with monoskill and multiskill servers and overflow routing. Accordingly, an arriving job is assigned to a corresponding vacant monoskill server, if possible. Otherwise the job is routed to a multiskill server, or rejected. We derive an efficiently computable upper bound for the system utilization by studying a modified routing where jobs can be redirected from multiskill to monoskill servers. This repacking policy improves performance also in terms of blocking probability when the service requirements of different job classes are statistically identical. Numerical simulations illustrate that our bounds provide good approximations for the performance of the original system.
Keywords: multiclass loss network, stochastic comparison, overflow routing, repacking

1 INTRODUCTION

In many communication systems, different services can be provided by a combination of monoskill servers, assigned to serving a certain class of jobs, and multiskill servers capable of dealing with all type of jobs. The routing of jobs is usually based on an overflow policy where an arriving job is preferentially routed to a corresponding monoskill server. If all monoskill servers corresponding to the class of the arriving job are busy, the job is routed to a multiskill server, if available, or rejected otherwise. This model fits several telecommunications applications such as call centers, streaming media, grid computing and wireless networks. Dimensioning has become an important economical issue in large call centers with specialized agents dedicated to diverse services [1–3]. As video-on-demand becomes more popular, new optimization methods are needed for streaming media servers that can be specialized in terms of clip size or delay constraints [4, 5]. In the field of grid computing, sharing and specialization of computational tasks between a set of distant workstations and supercomputers may produce considerable capacity gains. In wireless networks, a call can be served by different carriers; network operators may share a chunk of bandwidth for dealing with traffic surges. Dimensioning this type of systems is challenging because the analysis of the overflow processes is very complex, and exact numerical solution becomes intractable as soon as the number of servers in the system is greater than a few units [6].

Numerous papers study blocking in loss networks (see [7] for an overview). The simplest way to estimate the blocking probability is the exponential approximation where the overflow process is modeled by a Poisson process. Some other methods include the Hayward–Fredericks method [8] and the hyperexponential decomposition [9]; the latter provides accurate estimates but is computationally demanding. Our approach is to consider a slightly modified system, where jobs from the multiskill servers are instantaneously redirected to monoskill servers as soon as places become vacant. Applying the Markov reward approach [10], we analytically prove that repacking increases the mean number of jobs in the system. As a consequence, a computationally efficient upper bound for the carried load of the original system is derived. Moreover, repacking improves system performance also in terms of blocking probability when the service times are statistically identical. Numerical simulations illustrate that the blocking probability of the modified system is a good approximation for the blocking probability of the original system

The rest of the paper is structured as follows. In Section 2, we introduce the model and notation. In Section 3, we prove a dynamical stability property, which is later used to prove the main theorem. In Section 4, we illustrate by numerical simulations the tightness of our bounds. Section 5 concludes the paper.

2 OVERFLOW ROUTING AND REPACKING

2.1 Overflow routing

Consider a loss network serving K classes of jobs. The network consists of M_k monoskill servers (or resources) assigned to serving jobs of class k, and N multiskill servers capable of serving all classes of jobs. Customers of class k arrive according to a Poisson process of intensity λ_k and require exponential service times with mean $1/\mu_k$. We assume that the arrival processes and the service times are independent. Arriving service requests are routed according to an overflow policy, where a job of class k is always routed to a corresponding monoskill server, if there is one available. If all monoskill servers for class k are busy, the job is routed to a multiskill server. If all multiskill servers are also busy, then the service request is rejected, see Figure 1. We assume that the rejected requests leave the system without retrials.

Let $X_{1,k}$, $X_{2,k}$ be the number of jobs of class k being served by the monoskill and multiskill servers, respectively. Then the system state can be described by the random vector $X = (X_{i,k}) \in Z^{2K}$ indexed by $i = 1, 2$ and $k = 1, \dots, K$. For $x \in Z^{2K}$, denote $|x| = \sum_i \sum_k |x_{i,k}|$, and define the seminorms $|x|_i = \sum_k |x_{i,k}|$, $i = 1, 2$. The state space of the system is denoted by $S = \{(x_{i,k}) \in Z^{2K} : \ 0 \le x_{1,k} \le M_k \ \forall k, \ |x|_2 \le N\}$, with unit vectors $e_{i,k}$, $i = 1, 2$, $k = 1, \dots, K$. By construction, X is a multi-dimensional birth–death process on S, and the generator of X has the nontrivial entries

$$q(x, y) = \begin{cases} \lambda_k 1(x_{1,k} < M_k), & y = x + e_{1,k}, \\ \lambda_k 1(x_{1,k} = M_k, |x|_2 < N), & y = x + e_{2,k}, \\ \mu_k x_{i,k}, & y = x - e_{i,k}, \quad i = 1, 2. \end{cases} \tag{1}$$

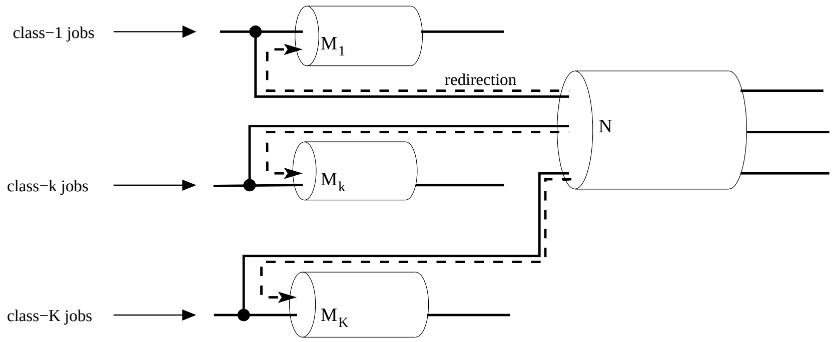

Fig. 1. Overflow system and repacking.

2.2 Repacking

Consider the same network with slightly modified routing, where jobs are redirected from multiskill to monoskill servers as soon as places become vacant. Denoting the corresponding state vector by $X' = (X'_{i,k})$, it follows that X' is also a birth–death process on S. The generator of X' has the nontrivial entries

$$q'(x,y) = \begin{cases} \lambda_k 1(x_{1,k} < M_k), & y = x + e_{1,k}, \\ \lambda_k 1(x_{1,k} = M_k, |x|_2 < N), & y = x + e_{2,k}, \\ \mu_k x_{1,k} 1(x_{2,k} = 0), & y = x - e_{1,k}, \\ \mu_k x_{1,k} 1(x_{2,k} > 0) + \mu_k x_{2,k} & y = x - e_{2,k}. \end{cases} \tag{2}$$

The repacking policy guarantees that all states x with $x_{1,k} < M_k$ and $x_{2,k} > 0$ are transient for X' and thus have zero stationary probability. As a consequence of this special feature, the aggregate process $\tilde{X}' = (X'_{1,k} + X'_{2,k})_{k=1}^K$ describing the net amount of jobs of class k is a birth–death process on

$$\tilde{S} = \{x \in Z^K : \; 0 \le x_k \le M_k + N \quad \forall k, \quad |x| \le \textstyle\sum_k M_k + N\}.$$

The process $\tilde{X}'$ has the transitions $x \mapsto x + e_k \in \tilde{S}$ and $x \mapsto x - e_k \in \tilde{S}$, occurring at rates λ_k and $\mu_k x_k$, respectively. It is well-known (see for example [11]) that $\tilde{X}'$ has the product form stationary distribution given by

$$\frac{\mathrm{P}(\tilde{X}' = x)}{\mathrm{P}(\tilde{X}' = 0)} = \prod_k \frac{(\lambda_k/\mu_k)^{x_k}}{x_k!}, \quad x \in \tilde{S}. \tag{3}$$

3 PERFORMANCE ANALYSIS

3.1 Performance measures

In this section we will compare the performance of the multiclass network with and without repacking. Performance is measured in terms of *carried load* a, defined as the mean rate of work (in erlangs) served by the network, and *overall blocking probability* b, which is the probability that an arbitrary arriving job is rejected. Denote the set of blocking states for class k by $B_k = \{x \in S : x_{1,k} = M_k, |x|_2 = N\}$. Then by the Little law, the stationary blocking probabilities $b_k = \mathrm{P}(X \in B_k)$ for class k satisfy

$$(1 - b_k)\frac{\lambda_k}{\mu_k} = \mathrm{E}(X_{1,k} + X_{2,k}).$$

Since $a = \sum_k (1 - b_k)\lambda_k/\mu_k$, and $b = (\sum_k \lambda_k b_k)/(\sum_k \lambda_k)$, this implies

$$a = \mathrm{E}\,|X|, \quad b = 1 - \frac{\sum_k \mu_k\, \mathrm{E}(X_{1,k} + X_{2,k})}{\sum_k \lambda_k}. \tag{4}$$

Analogous reasoning for X' shows that

$$a' = \mathrm{E}\,|X'|, \quad b' = 1 - \frac{\sum_k \mu_k\, \mathrm{E}(X'_{1,k} + X'_{2,k})}{\sum_k \lambda_k}. \tag{5}$$

Observe that for the system with repacking,

$$\mathrm{E}\,|X'| = \mathrm{E}\,|\tilde{X}'| \quad \text{and} \quad \mathrm{P}(X' \in B_k) = \mathrm{P}(\tilde{X}' \in \tilde{B}_k),$$

where $\tilde{X}'$ is the aggregate process defined in Section 2.2, and $\tilde{B}_k = \{x \in \tilde{S} : x_k = M_k \text{ or } |x| = \sum_j M_j + N\}$. Thus, numerical evaluation of a' and $b' = (\sum_k \lambda_k b'_k)/(\sum_k \lambda_k)$ can be efficiently carried out using the product form structure (3) for the stationary distribution of $\tilde{X}'$.

On the other hand, there are no simple closed form expressions for a or b, since neither X nor its aggregated version are reversible. Moreover, brute force numerical solution of the stationary distribution of X is not feasible, because the size of the state space is of the order $N^K M_1 \cdots M_K$. For example, with $K = 2$, $M_1 = M_2 = N = 9$, the rate matrix $q(x, y)$ has over 30 million entries. However, in Section 3.3 we will show how a' and b' can provide bounds for a and b.

3.2 Stability of overflow routing with respect to initial states

We will next prove a key continuity property of the original system: small perturbations of the initial state do not significantly change the future behavior of the system. Smallness will here be characterized in terms of

$$\Delta = \{0,\ \pm e_{2,k}, \quad e_{2,j} - e_{2,k}, \quad j, k = 1, \ldots, K\} \subset Z^{2K}.$$

Let us first focus on the discrete-event dynamics of the system. Let $A_k, D^m_{1,k}, D^n_{2,k}$ be independent Poisson processes on the positive real line with intensities λ_k, μ_k, μ_k, respectively, where $m = 1, \ldots, M_k$, $n = 1, \ldots, N$, and $k = 1, \ldots, K$. Further, fix $\tau_0 = 0$, and denote by τ_n the n-th point of the aggregate point process $\sum_k (A_k + \sum_{m=1}^{M_k} D^m_{1,k} + \sum_{n=1}^{N} D^n_{2,k})$. Without loss of generality we may assume that all τ_n are distinct.

Given an initial state $x^0 \in S$, define the sequence x^n by

$$x^{n+1} - x^n = \begin{cases} e_{1,k}, & \text{if } \tau_{n+1} \in A_k,\ x^n_{1,k} < M_k, \\ e_{2,k}, & \text{if } \tau_{n+1} \in A_k,\ x^n_{1,k} = M_k,\ |x^n|_2 < N, \\ -e_{i,k}, & \text{if } \tau_{n+1} \in \sum_{m=1}^{x^n_{i,k}} D^m_{i,k}, \quad i = 1, 2, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Then a realization $x(t)$ of the process X is given by setting

$$x(t) = x^n, \quad \text{for} \;\; t \in [\tau_n, \tau_{n+1}). \tag{7}$$

Lemma 1. *Let $x(t)$ and $y(t)$ be two solutions of the evolution equations (6) and (7) driven by the point processes $A_k, D^m_{1,k}, D^n_{2,k}$. Then*

$$y(0) - x(0) \in \Delta \quad \Longrightarrow \quad y(t) - x(t) \in \Delta \quad \textit{for all} \;\; t \geq 0.$$

Proof. By (7), it is enough to show that $y^n - x^n \in \Delta$ for all n. To proceed by induction, assume $y^n - x^n \in \Delta$. We will consider separately the different cases according to the type of the next event τ_{n+1}.

First, assume $\tau_{n+1} \in A_k$. If $x^n_{1,k} = y^n_{1,k} < M_k$, then $y^{n+1} - x^{n+1} \in \Delta$, so assume that $x^n_{1,k} = y^n_{1,k} = M_k$. If $|x^n|_2$ and $|y^n|_2$ are both equal to N or both strictly less that N, then also $y^{n+1} - x^{n+1} \in \Delta$. On the other hand, if $|x^n|_2 < N$ and $|y^n|_2 = N$, then $y^n - x^n \in \Delta$ implies $y^n = x^n + e_{2,j}$ for some j. Thus, $y^{n+1} - x^{n+1} = y^n - x^n - e_{2,k} = e_{2,j} - e_{2,k} \in \Delta$. By symmetry, the same conclusion holds when $|x^n|_2 = N$ and $|y^n|_2 < N$.

Next, let us consider the case where $\tau_{n+1} \in D^m_{1,k}$ for some m. Because $x^n_{1,k} = y^n_{1,k}$, it follows that $y^{n+1} - x^{n+1} = y^n - x^n \in \Delta$.

Finally, assume $\tau_{n+1} \in D^m_{2,k}$ for some m. Now, if $x^n_{2,k}$ and $y^n_{2,k}$ are both strictly less than m or both at least m, then $y^{n+1} - x^{n+1} = y^n - x^n \in \Delta$. Alternatively, if $x^n_{2,k} < m \leq y^n_{2,k}$, then $y^{n+1} - x^{n+1} = y^n - x^n - e_{2,k}$. Moreover, $y^n - x^n \in \Delta$ now implies that either $y^n = x^n + e_{2,k}$ or $y^n = x^n - e_{2,j} + e_{2,k}$ for some j. Thus, $y^{n+1} - x^{n+1} \in \Delta$ also in this case. The remaining case with $y^n_{2,k} < m \leq x^n_{2,k}$ is analogous.

3.3 Stochastic comparison of the system performance

In Section 3.1 we saw that the carried load and the overall blocking probability are hard to compute for the original system, while for the system with repacking these quantities have nice analytical formulas. In this section we show how the system with repacking can provide bounds for the original system. We will first give the results, then the proofs.

Theorem 1. *In the stationary regime, the mean net amount of jobs in the system with repacking is always greater than or equal to the corresponding quantity in the original system, that is,* $\mathrm{E}\,|X'| \geq \mathrm{E}\,|X|$.

Combining Theorem 1 with the balance formulas (4) and (5) gives the following results as corollaries.

Corollary 1. *The carried load in the system with repacking is always greater than or equal to the carried load in the original system.*

Corollary 2. *When $\mu_k = \mu$ for all k, then the overall blocking probability for the system with repacking is less than or equal to the overall blocking probability in the original system.*

The next example shows that the requirement for statistically identical service times is necessary in Corollary 2.

Example 1. Consider the system with two job classes and $(M_1, M_2, N) = (1, 0, 1)$. Assume $(\lambda_1, \mu_1) = (1, 1)$ and $(\lambda_2, \mu_2) = (1, \mu)$. Then by solving the equilibrium distribution one can check that

$$(b_1, b_2) = \left(\frac{10 + 9\mu + 2\mu^2}{20 + 36\mu + 10\mu^2}, \frac{20 + 16\mu + 3\mu^2}{20 + 36\mu + 10\mu^2} \right),$$

while

$$(b_1', b_2') = \left(\frac{2 + \mu}{4 + 5\mu}, \frac{4 + \mu}{4 + 5\mu} \right).$$

Comparison of the above expressions shows that $b_1' \geq b_1$ and $b_2' \leq b_2$ for all μ. Further, $b' \leq b$ for $\mu \geq \frac{2}{5}$, while $b' > b$ for $\mu < \frac{2}{5}$. This shows that repacking improves the performance experienced by the jobs of class 2 at the cost of bigger loss for class 1.

The proof of Theorem 1 is based* on the Markov reward approach [10]. Consider the uniformized Markov chain with the transition matrix $Q_\Lambda(x, y) = \delta_{x,y} + \Lambda^{-1} q(x, y)$, where $\delta_{x,y}$ is the Kronecker delta, and the scalar Λ is large enough to guarantee that Q_Λ is a stochastic matrix. Let $r(x) = |x|$, and define $V^n = \sum_{j=0}^{n-1} Q_\Lambda^j r$, where V^n and r are regarded as vectors indexed by $x \in S$. Then as n grows to infinity, $\frac{1}{n} V^n(x)$ converges to the stationary mean $\mathrm{E}\,|X|$ for all $x \in S$, allowing us to study $\mathrm{E}\,|X|$ using induction for V^n. Let us first prove a monotonicity property of the original system.

Lemma 2. *For all $x \in S$ with $x - e_{2,j} \in S$, we have*

$$V^n(x - e_{2,j}) \leq V^n(x) \quad \text{for all } n.$$

Proof. Because $V^0 = 0$, the inequality holds for $n = 0$. To proceed by induction, assume that the claim is true for index n, and let $x \in S$ such that $x_{2,j} > 0$. Then

$$\begin{aligned}
&\Lambda V^{n+1}(x) - \Lambda V^{n+1}(x - e_{2,j}) \\
&= \Lambda + \left(\Lambda - \textstyle\sum_k (\lambda_k + \mu_k(x_{1,k} + x_{2,k}))\right) [V^n(x) - V^n(x - e_{2,j})] \\
&+ \textstyle\sum_k \lambda_k 1(x_{1,k} < M_k)[V^n(x + e_{1,k}) - V^n(x + e_{1,k} - e_{2,j})] \\
&+ \textstyle\sum_k \lambda_k 1(x_{1,k} = M_k, |x|_2 < N)[V^n(x + e_{2,k}) - V^n(x + e_{2,k} - e_{2,j})] \\
&+ \textstyle\sum_k \mu_k x_{1,k}[V^n(x - e_{1,k}) - V^n(x - e_{1,k} - e_{2,j})] \\
&+ \textstyle\sum_k \mu_k (x_{2,k} - \delta_{j,k})[V^n(x - e_{2,k}) - V^n(x - e_{2,k} - e_{2,j})] \\
&+ \textstyle\sum_k \lambda_k 1(x \in B_k)[V^n(x) - V^n(x - e_{2,j} + e_{2,k})].
\end{aligned}$$

As a consequence of Lemma 1, $V^n(x) - V^n(x - e_{2,j} + e_{2,k}) \geq -1$ for all k, so the last term in the sum above is greater than or equal to $-\Lambda$. This proves the lemma, because the other terms involving V^n are nonnegative by the induction assumption.

Proof (Proof of Theorem 1). In order to prove that $\mathrm{E}\,|X'| \geq \mathrm{E}\,|X|$ applying [10, Theorem 2.1], it is enough to show that for all $x \in S$ and all n,

$$\sum_{y \neq x} (q'(x, y) - q(x, y))(V^n(y) - V^n(x)) \geq 0. \tag{8}$$

* Due to multidimensional blocking, the comparison techniques based on integral stochastic orderings [12, 13] are not directly applicable here.

Because the left-hand side of (8) equals

$$\sum_k \mu_k x_{1,k} 1(x_{2,k} > 0)(V^n(x - e_{2,k}) - V^n(x - e_{1,k})),$$

it suffices to show that for all n,

$$V^n(x - e_{2,j}) - V^n(x - e_{1,j}) \geq 0 \quad \text{for all } x \in S \ \text{ such that } \ x_{1,j}, x_{2,j} > 0. \tag{9}$$

Assume that (9) holds for index n (it is trivial for $n = 0$). Let $x \in S$ with $x_{1,j}, x_{2,j} > 0$. Then a direct calculation shows that

$$\begin{aligned}
&\Lambda V^{n+1}(x - e_{2,j}) - \Lambda V^{n+1}(x - e_{1,j}) \\
&= (\Lambda - q(x - e_{1,j}))[V^n(x - e_{2,j}) - V^n(x - e_{1,j})] \\
&+ \textstyle\sum_k \lambda_k 1(x_{1,k} < M_k)[V^n(x + e_{1,k} - e_{2,j}) - V^n(x + e_{1,k} - e_{1,j})] \\
&+ \textstyle\sum_{k \neq j} \lambda_k 1(x_{1,k} = M_k, |x|_2 < N)[V^n(x + e_{2,k} - e_{2,j}) - V^n(x + e_{2,k} - e_{1,j})] \\
&+ \textstyle\sum_k \mu_k (x_{1,k} - \delta_{j,k})[V^n(x - e_{1,k} - e_{2,j}) - V^n(x - e_{1,k} - e_{1,j})] \\
&+ \textstyle\sum_k \mu_k (x_{2,k} - \delta_{j,k})[V^n(x - e_{2,k} - e_{2,j}) - V^n(x - e_{2,k} - e_{1,j})] \\
&+ \textstyle\sum_{k \neq j} \lambda_k 1(x \in B_k)[V^n(x - e_{2,j} + e_{2,k}) - V^n(x - e_{2,j})],
\end{aligned}$$

where $q(x) = \sum_{y \neq x} q(x, y)$, as usual. The last term in the above sum is nonnegative by Lemma 2, and so are also the other terms by the induction assumption.

4 NUMERICAL EXAMPLES

4.1 Carried load and blocking probability

Increasing the system utilization by repacking may lead to worse performance in terms of overall blocking probability, if the mean service times of job classes differ significantly. This is illustrated in Figure 2 where the carried load and the overall blocking probability for the original system (using simulation) and for the system with repacking are plotted for a network with two customer classes and $(M_1, M_2, N) = (5, 0, 5)$. The traffic characteristics are (i) $(\lambda_1, \mu_1) = (\lambda_2, \mu_2) = (\frac{5}{2}\rho, 1)$; and (ii) $(\lambda_1, \mu_1) = (50\rho, 20)$, $(\lambda_2, \mu_2) = (\frac{5}{2}\rho, 1)$. Here ρ is the offered load normalized by the number of servers. In case (i), Corollaries 1 and 2 imply that repacking improves both performance quantities. On the other hand, in case (ii) the overall blocking probability can be larger for the system with repacking. Note that by (3), the stationary distribution of $\tilde{X}'$ and the carried load in the system with repacking remain unchanged when altering the service times so that the per-class loads $a_k = \lambda_k / \mu_k$ remain constant**. This insensitivity property suggests that the results of Section 3 could be extended to nonexponential service times.

4.2 Efficient approximation of the blocking probability

The plots in Figure 2 show that the performance differences of the original and the modified system remain quite small for a wide range of loads. Hence the system with

** The overall blocking probability depends explicitly on λ_k and hence does not remain unchanged.

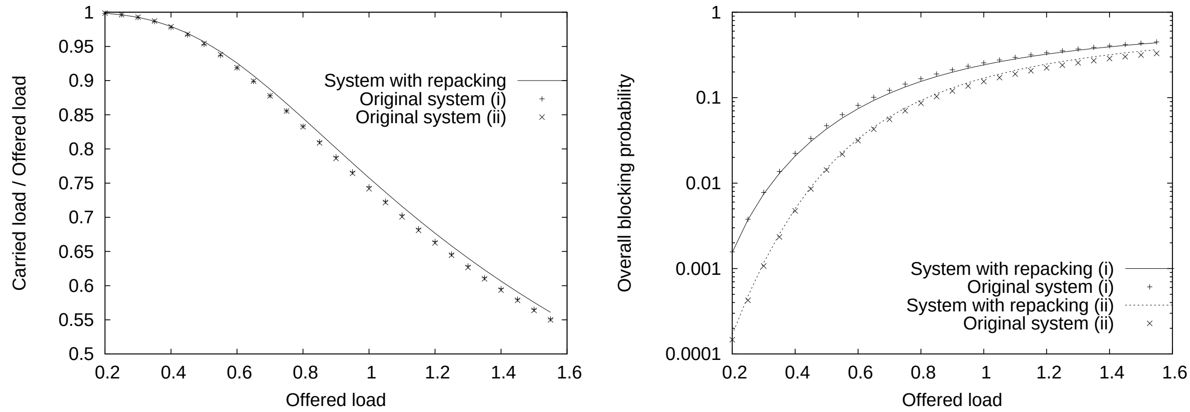

Fig. 2. Comparison of the carried load (left) and the overall blocking probability (right) in the system with and without repacking.

repacking constitutes a good approximation for the carried load and the overall blocking probability of the original system. In Figure 3 we compare this approximation with the simple exponential approximation where the arrivals to the multiskill server group are modeled as Poisson processes, so that $b_{exp} = \mathrm{Erl}(\sum_k \mathrm{Erl}(a_k, M_k)a_k, N)$ where $\mathrm{Erl}(\rho, n)$ is the Erlang blocking formula for an $M/M/n/n$ queue of intensity ρ. In the left plot of Figure 3, the network consists of two job classes with $(M_1, M_2, N) = (5, 0, 5)$ and traffic characteristics $(\lambda_1, \mu_1) = (\lambda_2, \mu_2) = (5\rho, 1)$. In the right-hand side we have $(M_1, M_2, N) = (5, 5, 4)$ and $(\lambda_1, \mu_1) = (\lambda_2, \mu_2) = (7\rho, 1)$. It is clear that the approximation given by the system with repacking behaves much better than the exponential one.

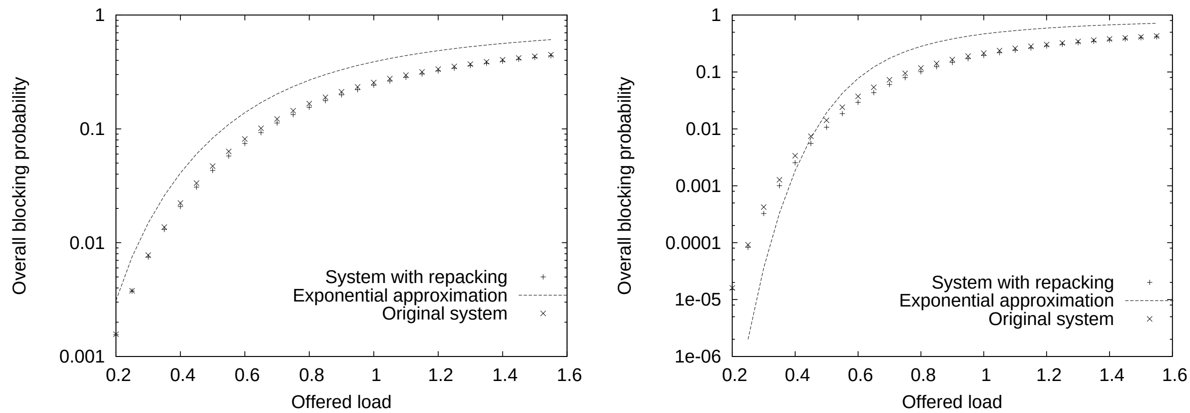

Fig. 3. Approximation of the blocking probability using repacking and the exponential approximation for $(M_1, M_2, N) = (5, 0, 5)$ (left) and $(M_1, M_2, N) = (5, 5, 4)$ (right).

5 CONCLUSION

We presented an efficiently computable upper bound for the carried load and an approximation of the overall blocking probability for multiclass loss networks with overflow routing. The insensitivity properties of the bounding process suggest that some of the results might be generalized to systems with nonexponential service times. Finding bounds for the per-class blocking probabilities remains an open problem. Other interesting directions for future work are extensions to more complex network topologies and service policies, especially processor sharing.

ACKNOWLEDGEMENTS

Part of this work was done while M. Jonckheere and L. Leskelä were visiting the Mittag–Leffler Institute. They wish to thank the people of the institute for their support.

References

1. Gans, N., Koole, G., Mandelbaum, A.: Telephone call centers: Tutorial, review and research prospects. Manufacturing & Service Operations Managment **5** (2003) 79–141
2. Chevalier, P., Shumsky, R., Tabordon, N.: Routing and staffing in large call centers with specialized and fully flexible servers. Submitted (2004)
3. Örmeci, E.L.: Dynamic admission control in a call center with one shared and two dedicated service facilities. IEEE Trans. Aut. Contr. **49** (2004) 1157–1161
4. Eager, D., Vernon, M., Zahorjan, J.: Optimal and efficient merging schedules for video-on-demand servers. In: ACM Multimedia 99. (2003)
5. Sarhan, N., Das, C.: Providing time of service guarantees in video-on-demand servers. In: The Twelfth International World Wide Web Conference. (2003)
6. Louth, G., Mitzenmacher, M., Kelly, F.P.: Computational complexity of loss networks. Theor. Comp. Sc. **125** (1994) 45–59
7. Kelly, F.P.: Loss networks. Ann. Appl. Prob. **1** (1991) 319–378
8. Fredericks, A.: Congestion in blocking systems – a simple approximation technique. The Bell System Technical Journal **59** (1980) 805–827
9. Franx, G.J., Koole, G., Pot, A.: Approximating multi-skill blocking systems by hyperexponential decomposition. Submitted (2004)
10. van Dijk, N.M.: Bounds and error bounds for queueing networks. Ann. Oper. Res. **79** (1998) 295–319
11. Kaufman, J.S.: Blocking in a shared resource environment. IEEE Trans. Comm. **29** (1981) 1474–1481
12. Massey, W.A.: Stochastic orderings for Markov processes on partially ordered spaces. Math. Oper. Res. **12** (1987) 350–367
13. Whitt, W.: Stochastic comparisons for non-Markov processes. Math. Oper. Res. **11** (1986) 608–618