A Tandem Queueing Network with Feedback Admission Control

L. Leskelä and J. Resing

REPORT No. 09, 2004/2005, fall

ISSN 1103-467X
ISRN IML-R- -09-04/05- -SE+fall

INSTITUT MITTAG-LEFFLER
THE ROYAL SWEDISH ACADEMY OF SCIENCES

A tandem queueing network with feedback admission control

Lasse Leskelä* Jacques Resing†

December 31, 2004

Keywords: queueing network, admission control, feedback signaling, censoring

MSC: 60K25, 68M20, 90B15, 90B22

Abstract

We study a tandem queueing network with feedback admission control. The steady-state behavior of the network is analyzed by studying a related censored process with a $G/M/1$ type generator. Our numerical results show how the presence of feedback signaling may break the monotonicity of typical performance characteristics.

1 Introduction

Admission control can be employed to avoid congestion in queueing networks subject to heavy load. In distributed networks the admission controller typically operates on the edge of a subnetwork, involving a feedback signaling loop. In this paper we will analyze the performance of this type of control schemes. For mathematical tractability, we will restrict ourselves to the simplest non-trivial instance of such a system, the Markovian two-node tandem queueing network with unlimited buffers.

The input to the system is modeled as a Poisson process with rate λ, and the service times of jobs in nodes 1 and 2 are assumed to be independent exponentially distributed random variables with parameters μ_1 and μ_2, respectively. The admission control scheme is based on X_2, the number of jobs present in node 2 (see Figure 1). As long as X_2 is smaller than or equal to a certain threshold level K, new jobs arriving to the network are accepted. However, whenever X_2 exceeds K, new jobs arriving to the system are rejected.

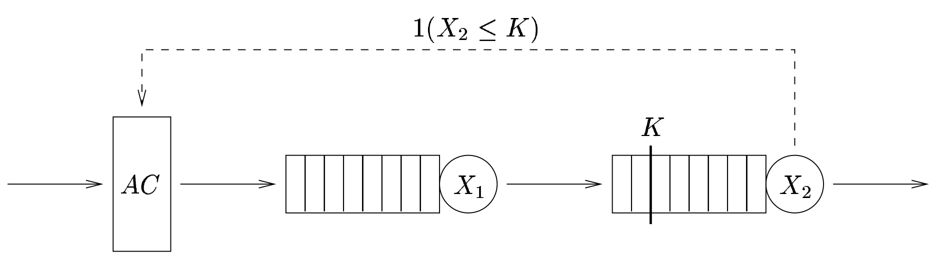

Figure 1: The admission control mechanism.

Other types of feedback signaling have been studied in the literature. In [5, 7, 3, 12] the first server stops processing when the number of jobs at the second station becomes too high. In [6], new jobs are rejected whenever the number of jobs at the first station reaches a certain threshold.

*Helsinki University of Technology, P.O. Box 1100, 02015 HUT, Finland, e-mail: lasse.leskela@iki.fi

†Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, e-mail: j.a.c.resing@tue.nl

In all these models, the content of one of the buffers always remains bounded. This is clearly not the case in our model where, in spite of the admission control mechanism, the number of jobs in both stations can grow arbitrarily large. In the special case with $K = 0$, our model coincides with the one studied in [1].

The stability of the queueing network of Figure 1 was recently studied in [8], where it was shown that the system with admission threshold K is stable if and only if the triple (λ, μ_1, μ_2) satisfies the relation

$$\lambda\left(1 - (\mu_1/\mu_2)^{K+1}\right) < \mu_1. \tag{1}$$

This condition partitions the parameter domain for (λ, μ_1, μ_2) into four distinct regions. Figure 2 illustrates these regions when we fix $\lambda = 1$. A_1 represents the parameter values for which the uncontrolled system is stable, A_2 is the region where the system is stable for any value of K, A_3 the region where the system is stable for K small enough, and A_4 the region where the system can not be stabilized. The regions A_3 and A_4 are separated by the curve $\mu_2 = \mu_1/(1 - \mu_1)$.

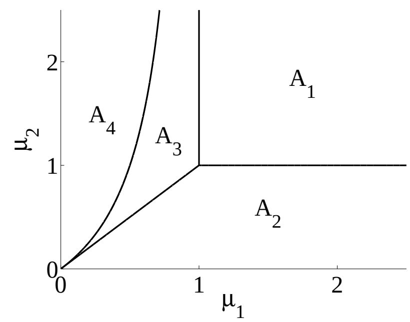

Figure 2: The stability regions of the system.

In this paper we will study the steady-state performance of the system. In Section 2 we will apply censoring (cf. e.g. [4]) to reduce the problem of solving the equilibrium distribution of the queue lengths to a corresponding problem for a simpler Markov process whose equilibrium distribution has a matrix-geometric form. Further, we will present in Section 3 an approach for efficient evaluation of the steady-state performance of the network. We will conclude the paper by giving numerical examples that illustrate how the presence of feedback signaling affects the system.

2 Queue length analysis

Let us denote by $X = (X_1(t), X_2(t))_{t\geq 0}$ the Markov process where $X_i(t)$ is the number of jobs at server i at time t. Unless explicitly stated otherwise, we will assume from now on that (1) holds, so that X is ergodic. In the sequel we will first study the process X sampled at periods of time during which $X_2 \leq K$. This censored process has a generator with a special block structure of the so-called $G/M/1$ type [9] which we will exploit to evaluate its equilibrium distribution. Afterwards, we will find the steady-state probabilities of the process X itself.

2.1 The censored process

Let us denote the state space of X by $S = Z_+^2$. The censored process $Y = (Y_1(t), Y_2(t))_{t\geq 0}$ is defined by $Y_i(t) = X_i(\gamma(t))$ where $\gamma(t) = \inf\{\tau \geq 0 : \int_0^\tau 1_{\{X_2(s)\leq K\}}\, ds > t\}$. It follows from the strong Markov property (cf. [10, Section III.21]) that Y is a Markov process on $S^- = \{(n, k) \in S : k \leq K\}$.

To conveniently describe the infinitesimal generator of Y, we will employ the following notation for $(K+1)$-dimensional square matrices. We will denote by I the identity matrix, while T_L and T_R will stand for the left and right shift matrices given by $(T_L)_{i,j} = \delta_{i-1,j}$ and $(T_R)_{i,j} = \delta_{i+1,j}$

for $0 \le i,j \le K$ where $\delta_{i,j}$ denotes the Kronecker delta. Further, we will denote the projection matrices onto 0-th and K-th coordinate by U_0 and U_K, that is, $(U_0)_{i,j} = \delta_{i,0}\delta_{j,0}$ and $(U_K)_{i,j} = \delta_{i,K}\delta_{j,K}$. If we order the states in S^- lexicographically, the generator of Y can be written in the form

$$Q = \begin{pmatrix} B_0 & A_0 & 0 & 0 & \cdots \\ B_1 & A_1 & A_0 & 0 & \cdots \\ B_2 & A_2 & A_1 & A_0 & \cdots \\ B_3 & A_3 & A_2 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix},$$

where the matrices A_n and B_n are given by

$$\begin{aligned} A_0 &= \lambda I, \\ A_1 &= \mu_2 T_L - (\lambda + \mu_1 + \mu_2)I + \mu_2 U_0, \\ A_2 &= \mu_1 (T_R + q_1 U_K), \\ A_{n+1} &= \mu_1 q_n U_K, \quad n \ge 2, \end{aligned}$$

and

$$\begin{aligned} B_0 &= \mu_2 T_L - (\lambda + \mu_2)I + \mu_2 U_0, \\ B_1 &= \mu_1 (T_R + U_K), \\ B_{n+1} &= \mu_1 (1 - q_1 - \cdots - q_n) U_K, \quad n \ge 1. \end{aligned}$$

The numbers q_n represent the probabilities that, if the process X leaves the set S^- in some state $(m+n, K)$, it enters S^- again in state (m, K), where $m \ge 1$. It is not hard to check that q_n is equal to the probability that a random walk on the integers starting at state 0 and with probabilities $\mu_1/(\mu_1+\mu_2)$ and $\mu_2/(\mu_1+\mu_2)$ of going to the right and left, respectively, reaches state -1 for the first time in exactly $2n-1$ steps. It is well-known [2] that this quantity equals

$$q_n = C_{n-1} \left(\frac{\mu_1}{\mu_1+\mu_2}\right)^{n-1} \left(\frac{\mu_2}{\mu_1+\mu_2}\right)^{n},$$

where $C_n = \frac{1}{n+1}\binom{2n}{n}$ are the Catalan numbers.

For $k = 0, \ldots, K$, denote by e_k the k-th basis vector of the $(K+1)$-dimensional euclidean space, and let $e = \sum_{k=0}^{K} e_k$. By convention, all vectors are treated as row vectors. One can check that Neuts' mean drift condition [9, Formula (1.7.11)] for the stability of Y will be equivalent to (1). Furthermore, the steady-state probabilities of the censored process are given [9] in the matrix-geometric form

$$P(Y = (n,k)) = x_0 R^n e_k^T, \quad (n,k) \in S^-, \tag{2}$$

where the matrix R is the unique minimal non-negative solution of

$$\sum_{n=0}^{\infty} R^n A_n = 0, \tag{3}$$

and x_0 is the unique positive row vector satisfying

$$x_0 \sum_{n=0}^{\infty} R^n B_n = 0, \quad \text{and} \quad x_0 (I - R)^{-1} e^T = 1. \tag{4}$$

2.2 The steady-state queue length distribution

Once we have found the equilibrium distribution of the censored process Y, we will next show how this distribution can be used to obtain the steady-state probabilities for the queue length process X. First, we note that

$$P(X = (n,k)) = P(X_2 \le K)\, P(Y = (n,k)), \quad (n,k) \in S^-. \tag{5}$$

Secondly, we know that, in steady-state, the mean rate at which jobs are accepted into the system must be equal to the mean rate of jobs coming out of the system, so that

$$\lambda P(X_2 \le K) = \mu_2(1 - P(X_2 = 0)).$$

Note that (5) implies $P(X_2 = 0) = P(X_2 \le K)P(Y_2 = 0)$. Substituting this into the above equation we get

$$P(X_2 \le K) = \frac{\mu_2}{\lambda + \mu_2 P(Y_2 = 0)}. \tag{6}$$

Thus, we see that (5) and (6) yield the steady-state probabilities of X for all states $(n, k) \in S^-$, and what remains is to find the corresponding quantities on $S^+ = S \setminus S^-$.

For the states in S^+, we will first find out the probabilities $P(X = (n, K + k))$ for $n, k > 0$ by inspecting the excursions X makes in S^+. Note that if X visits the state $(n, K + k)$, the time it spends there has mean $1/(\mu_1 + \mu_2)$, and that the mean rate of transitions from $(n + m, K)$ to S^+ equals $\mu_1 P(X = (n + m, K))$. Now we see by conditioning on the state in S^- from where X enters S^+ that for all $n, k > 0$,

$$P(X = (n, K + k)) = \frac{\mu_1}{\mu_1 + \mu_2} \sum_{m=k}^{\infty} q_{k,m}\, P(X = (n + m, K)), \tag{7}$$

where $q_{k,m}$ is the probability that X will visit state $(n, K + k)$ during an excursion in S^+ which was initiated from state $(n + m, K)$. It is not hard to see that $q_{k,m}$ does not depend on the values of n and K, and is equal to the probability that a random walk on the integers starting from state 1 at time 0 and with probabilities $\mu_1/(\mu_1 + \mu_2)$ and $\mu_2/(\mu_1 + \mu_2)$ of going to the right and to the left, respectively, will visit state k at time $2m - k - 1$ without visiting state 0 in any time inbetween. Using the ballot theorem (cf. [11]) one can verify that this quantity equals

$$q_{k,m} = \frac{k}{m}\binom{2m - k - 1}{m - 1}\left(\frac{\mu_1}{\mu_1 + \mu_2}\right)^{m-1}\left(\frac{\mu_2}{\mu_1 + \mu_2}\right)^{m-k}. \tag{8}$$

Finally, the probabilities $P(X = (0, k))$ with $k > K$ can be found by observing that the steady-state mean rate of transitions out of the set $\{(n, k') : n = 0, k' \ge k\}$ equals the corresponding rate into that set, so that

$$\mu_2 P(X = (0, k)) = \mu_1 P(X_1 = 1, X_2 \ge k - 1), \quad k > K. \tag{9}$$

Remark 2.1. *Our censoring approach can also be applied in more general situations. As an example, consider the system with multiple thresholds $K_1 < \cdots < K_N$ and acceptance probabilities $p_1 > \cdots > p_N$. Jobs arriving to the system are accepted with probability p_n when $K_{n-1} < X_2 \le K_n$ for some $n = 1, \ldots, N$, where $K_0 = -1$. Arriving jobs are always rejected whenever $X_2 > K_N$.*

2.3 The special case with the strictest admission threshold

In the special case with $K = 0$, (3) degenerates into a scalar equation. Using the fact that the generating function for the Catalan numbers satisfies

$$\sum_{n=0}^{\infty} C_n x^n = \frac{1 - \sqrt{1 - 4x}}{2x} \quad \text{for} \quad |x| < 1/4,$$

one can verify that for $|R| < 1$, equation (3) is equivalent to

$$\frac{\lambda}{\mu_1 + \mu_2} - \frac{\lambda + \mu_1}{\mu_1 + \mu_2} R + \frac{1}{2} R \left(1 - \sqrt{1 - 4\frac{\mu_1}{\mu_1 + \mu_2}\frac{\mu_2}{\mu_1 + \mu_2} R}\right) = 0, \tag{10}$$

while equations (4) reduce to $x_0 = 1 - R$. Assuming that the stability condition (1) holds, the minimal non-negative solution of (10) is given by

$$R = \frac{\lambda}{2\mu_1\mu_2}\left(\sqrt{(\lambda+\mu_1-\mu_2)^2+4\mu_1\mu_2}-(\lambda+\mu_1-\mu_2)\right), \tag{11}$$

and satisfies $0 < R < 1$. Thus, the conditional distribution of X_1 given $X_2 = 0$ is geometrical with parameter R. In fact, by substitution into the balance equations one can verify that the steady-state distribution (which has also been independently derived in [1]) of the system equals

$$P(X_1 = n, X_2 = k) = \begin{cases} \dfrac{\lambda}{\lambda+\mu_2}(1-R)\left(1-\dfrac{\mu_1}{\lambda}R\right)^{k-1}, & n = 0,\, k \geq 1, \\ \dfrac{\mu_2}{\lambda+\mu_2}(1-R)\,R^n\left(1-\dfrac{\mu_1}{\lambda}R\right)^{k}, & \text{otherwise.} \end{cases}$$

3 Performance analysis

3.1 Performance metrics

Tuning the admission control parameters for a network is a trade-off between the amount of accepted jobs and the performance experienced by them. For the amount of accepted jobs, a natural thing is to investigate the throughput θ, the steady-state number of accepted jobs per unit time, which in our case is given by

$$\theta = \lambda\, P(X_2 \leq K). \tag{12}$$

For the performance of accepted customers, a natural quantity of interest is the mean steady-state sojourn time $E(D)$. Under stability, we see using Little's law that

$$E(D) = \frac{E|X|}{\theta}, \tag{13}$$

where $|X| = X_1 + X_2$ denotes the total amount of jobs in the network.

To evaluate θ and $E(D)$, we could use formulas (5) – (9) derived in Section 2.2. However, this approach is not very appealing from a computational point of view, since it involves multiple infinite summations over the state space. To overcome this issue, we will now derive expressions for θ and $E(D)$ directly in terms of the steady-state distribution of the censored process Y.

Theorem 3.1. *Under stability, the steady-state throughput θ and the steady-state mean sojourn time $E(D)$ of jobs accepted to the network are given in terms of the equilibrium distribution of Y by*

$$\frac{1}{\theta} = \frac{1}{\lambda}P(Y_2 = 0) + \frac{1}{\mu_2}, \tag{14}$$

and

$$E(D) = \frac{1}{\lambda}E(Y_1 1_{\{Y_2=0\}}) + \frac{1}{\mu_2}E(|Y|+1). \tag{15}$$

Proof. The validity of (14) is clear after substitution of (6) into (12). To prove the second claim, let us consider the level transitions for the total amount of jobs $|X|$. Under stability, the mean rate of events where the value of $|X|$ changes from n to $n+1$ is given by $\lambda P(|X| = n, X_2 \leq K)$, while the corresponding rate backwards from $n+1$ to n equals $\mu_2 P(|X| = n+1, X_2 > 0)$. Thus,

$$\lambda P(|X| = n, X_2 \leq K) = \mu_2 P(|X| = n+1) - \mu_2 P(X_1 = n+1, X_2 = 0)$$

for all $n \geq 0$. Multiplying both sides of the above equality by $n+1$ and then summing over n we see that

$$\lambda E(\,(|X|+1)1_{\{X_2 \leq K\}}\,) = \mu_2 E|X| - \mu_2 E(X_1 1_{\{X_2=0\}}). \tag{16}$$

Taking θ from (12) and solving (16) for $E|X|$, and substituting these into (13) now yields (15). □

3.2 Long-term behavior of the unstable system

When studying the performance of the system as a function of its parameters, it is also interesting to see what happens in the unstable parameter region. Recall from (1) that instability of the system implies $\mu_1 < \mu_2$, so that the rate at which work is fed into the second server is strictly less than its service capacity. Thus, intuition suggests that only the first queue will grow to infinity. The next theorem verifies the validity of these heuristics.

Theorem 3.2. *Assume $\lambda(1-(\mu_1/\mu_2)^{K+1}) > \mu_1$. Then the process X started from an arbitrary initial state satisfies as $t \to \infty$,*

$$X_1(t) \to \infty \quad \textit{almost surely},$$
$$X_2(t) \to Z \quad \textit{in distribution},$$

where Z is a geometrically distributed random variable with parameter μ_1/μ_2.

Proof. Let $N_\lambda, N_{\mu_1}, N_{\mu_2}$ be independent Poisson processes with rates λ, μ_1, μ_2, respectively. Then X can be represented as the unique solution of

$$X_1(t) = X_1(0) + \int_{(0,t]} 1_{\{X_2(s)\le K\}}\, N_\lambda(ds) - \int_{(0,t]} 1_{\{X_1(s)>0\}}\, N_{\mu_1}(ds), \tag{17}$$

$$X_2(t) = X_2(0) + \int_{(0,t]} 1_{\{X_1(s)>0\}}\, N_{\mu_1}(ds) - \int_{(0,t]} 1_{\{X_2(s)>0\}}\, N_{\mu_2}(ds). \tag{18}$$

Let $\tilde{X_2}(t)$ be the solution of

$$\tilde{X_2}(t) = X_2(0) + N_{\mu_1}(t) - \int_{(0,t]} 1_{\{\tilde{X_2}(s)>0\}}\, N_{\mu_2}(ds). \tag{19}$$

Then a pathwise comparison of (18) and (19) shows that $\tilde{X_2}(t) \ge X_2(t)$ for all t almost surely. This implies that $X_1(t) \ge U(t)$ for all t a.s., where

$$U(t) = X_1(0) + \int_{(0,t]} 1_{\{\tilde{X_2}(s)\le K\}}\, N_\lambda(ds) - N_{\mu_1}(t).$$

Note that $\tilde{X_2}$ equals the number of customers in a stable $M/M/1$ queue with arrival rate μ_1 and mean service time $1/\mu_2$. Thus, $\tilde{X_2}(t) \to Z$ in distribution, where Z is geometric with parameter μ_1/μ_2. Since N_λ is independent of $\tilde{X_2}$, it is not hard to see that $\lim_{t\to\infty} \frac{1}{t}\int_{(0,t]} 1_{\{\tilde{X_2}(s)\le K\}}\, N_\lambda(ds) = \lambda P(Z \le K)$ a.s. Now using $\lim_{t\to\infty} \frac{1}{t} N_{\mu_1}(t) = \mu_1$ a.s., we see that with probability one,

$$\lim_{t\to\infty} U(t)/t = \lambda(1-(\mu_1/\mu_2)^{K+1}) - \mu_1.$$

Since the above limit is strictly positive, $U(t) \to \infty$ and thus $X_1(t) \to \infty$ a.s.

To verify that $X_2(t) \to Z$ in distribution, it is enough to show that X_2 and $\tilde{X_2}$ will couple in finite time. Let $T_0 = \sup\{t : X_1(t) = 0\}$. Since $X_1(t) \to \infty$, T_0 is a.s. finite. Define $T_1 = \inf\{t \ge T_0 : \tilde{X_2}(t) = 0\}$. Since $\tilde{X_2}$ represents the state of a stable $M/M/1$ queue, T_1 is finite a.s. Further, since X_2 is dominated by $\tilde{X_2}$, we see that $X_2(T_1) = \tilde{X_2}(T_1) = 0$. Since the pathwise dynamics of X_2 and $\tilde{X_2}$ coincide for $t \ge T_1$, we conclude that $X_2(t) = \tilde{X_2}(t)$ for all $t \ge T_1$. □

Theorem 3.2 implies that the long-term average rate of rejected jobs in the unstable system tends to $\lambda(\mu_1/\mu_2)^{K+1}$. This expression for the asymptotic loss rate seems somewhat paradoxical at the first sight, since the relation

$$\text{throughput} = \text{input rate} - \text{loss rate}$$

now suggests that for large values of K, the throughput should be close to the input rate. However, we need to bear in mind that X_1 grows to infinity, making the above conservation law no more valid. The above equation can be corrected by subtracting from the right-hand side the term $\lambda(1-(\mu_1/\mu_2)^{K+1}) - \mu_1$, which is the long-term growth rate of the number of jobs in the system.

3.3 Numerical results

To evaluate numerically the performance quantities θ and $E(D)$ as functions of the system parameters λ, μ_1, μ_2, K, we will combine the expressions of Theorem 3.1 with formula (2). Since $\sum_{n=0}^{\infty} nR^n = (I-R)^{-2}R$, we have the following

$$\begin{aligned} P(Y_2 = 0) &= x_0(I-R)^{-1}e_0^T, \\ E(Y_1 1_{\{Y_2=0\}}) &= x_0(I-R)^{-2}Re_0^T, \\ E|Y| &= x_0(I-R)^{-2}Re^T + x_0(I-R)^{-1}f^T, \end{aligned}$$

where $f = \sum_{k=0}^{K} ke_k$. The matrix R can be numerically solved from equation (3) using the method of successive substitutions [9]. When R is calculated, the vector x_0 will be obtained from (4).

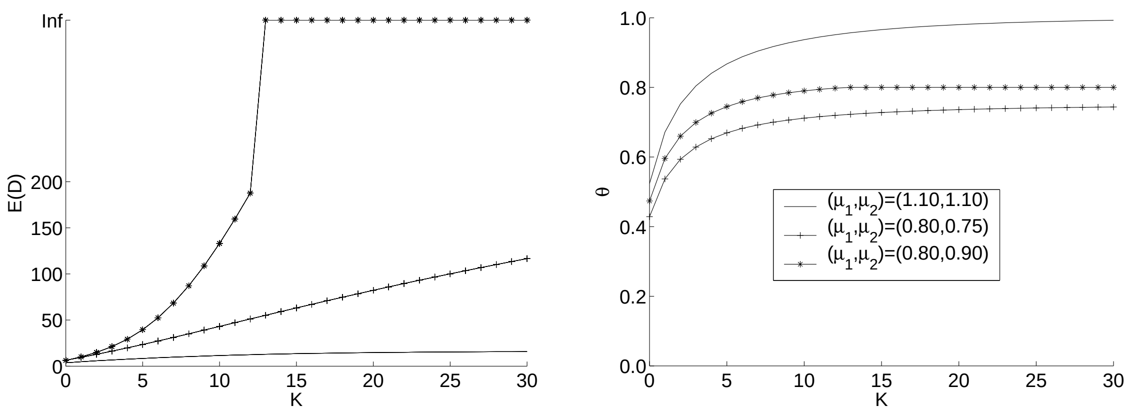

Figure 3: $E(D)$ and θ as functions of K.

Figure 3 illustrates numerically computed values for $E(D)$ and θ as functions of K, with $\lambda = 1$. Here we have chosen three pairs of (μ_1, μ_2), one from each of the regions A_1, A_2 and A_3 in Figure 2. Note that with $(\mu_1, \mu_2) = (0.80, 0.90)$, the system is stable for $K \le 12$ and unstable otherwise. Intuition suggests that for large K, the behavior of the admission-controlled system should be close to the standard two-server tandem queue, for which we know that the throughput equals $\min(\lambda, \mu_1, \mu_2)$, while the mean sojourn time is $\frac{1}{\mu_1-\lambda} + \frac{1}{\mu_2-\lambda}$ for $\lambda < \min(\mu_1, \mu_2)$ and infinite otherwise. Comparing these formulas with the curves in Figure 3, we see that the numerical computations are in agreement with this intuition.

Figure 4 illustrates numerically computed contours of $E(D)$ and θ for varying μ_1 and μ_2 with $\lambda = 1$ and $K = 5$. The contours in the upper right corner represent values of (μ_1, μ_2) with the lowest mean sojourn time (highest throughput) while those in the lower left corner correspond to (μ_1, μ_2) with the highest mean sojourn time (lowest throughput). Figure 4 shows that adding more capacity to the second server may either increase or decrease the steady-state mean sojourn time of the accepted customers. This effect, due to the presence of the feedback signaling loop, is more clearly visible in Figure 5 where θ and $E(D)$ are plotted against μ_2, fixing $\mu_1 = 0.6$.

Acknowledgements

The work presented in this article was done while the authors were visiting the Mittag–Leffler Institute. The authors would like to thank the organizers and visitors of the Fall 2004 program as well as the institute staff for providing a superb research milieu.

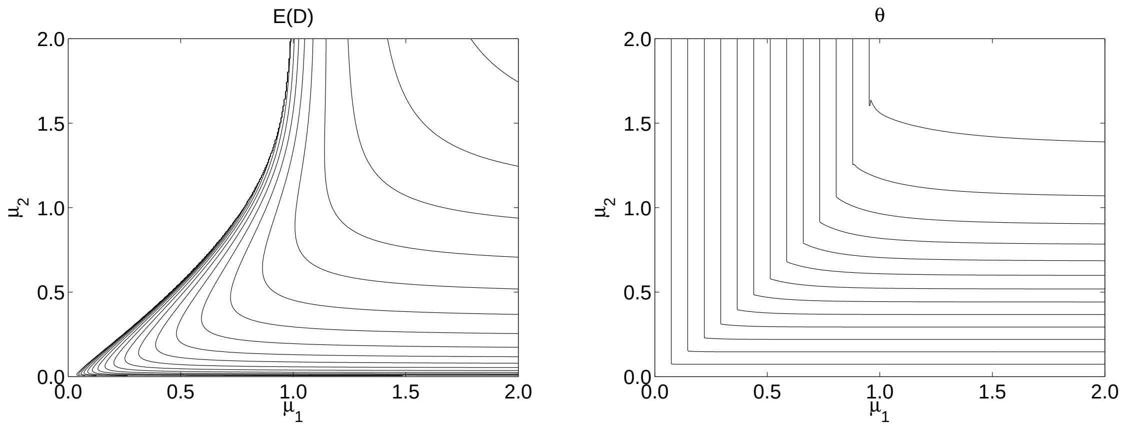

Figure 4: Contours of $E(D)$ and θ as functions of (μ_1, μ_2) with $\lambda = 1$ and $K = 5$.

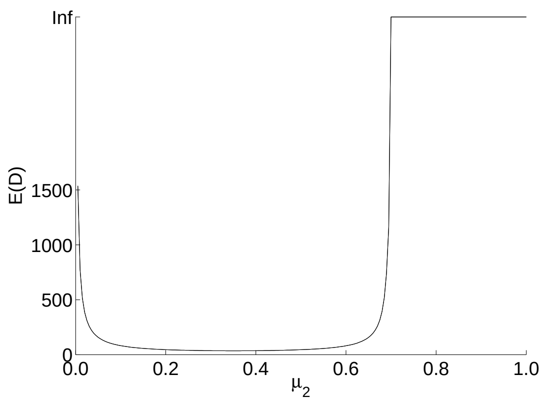

Figure 5: $E(D)$ as a function of μ_2 with $\lambda = 1$, $\mu_1 = 0.6$, and $K = 5$.

References

[1] I. J. B. F. Adan and G. Weiss. Analysis of a simple Markovian re-entrant line with infinite supply of work under the LBFS policy. Technical Report SPOR-Report 2004-14, Eindhoven University of Technology, 2004.

[2] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume I. Wiley, third edition, 1966.

[3] W. K. Grassmann and S. Drekic. An analytical solution for a tandem queue with blocking. *Queueing Systems*, 36(1):221–235, 2000.

[4] W. K. Grassmann and D. A. Stanford. Matrix analytic methods. In *Computational Probability*, pages 153–203. Kluwer, 2000.

[5] A. G. Konheim and M. Reiser. A queueing model with finite waiting room and blocking. *Journal of the ACM*, 23(2):328–341, 1976.

[6] D. Kroese, W. R. Scheinhardt, and P. G. Taylor. Spectral properties of the tandem Jackson network, seen as a quasi-birth-and-death process. *Ann. Appl. Prob.*, 14(4):2057–2089, 2004.

[7] G. Latouche and M. F. Neuts. Efficient algorithmic solutions to exponential tandem queues with blocking. *SIAM J. Alg. Discr. Meth.*, 1:93–106, 1980.

[8] L. Leskelä. Stabilization of an overloaded queueing network using measurement-based admission control. Technical Report A470, Helsinki University of Technology, Institute of Mathematics, 2004.

[9] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models.* John Hopkins University Press, 1981.

[10] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales*, volume 1. Wiley, second edition, 1994.

[11] L. Takács. *Combinatorial Methods in the Theory of Stochastic Processes.* Wiley, 1967.

[12] N. D. van Foreest, M. Mandjes, J. C. W. van Ommeren, and W. R. W. Scheinhardt. A tandem network with server slow-down and blocking. Submitted, 2004.