Robust link dimensioning in multi-rate loss networks

L. R. Bercho, I. S. Bonatti, P. L. D. Peres and A. Lopes
University of Campinas, Campinas SP, Brazil.

Abstract— **Traditional formulation of resource access often leads to unfair blocking probabilities among distinct classes of services in multi-rate loss networks where full sharing polices are adopted. Classes demanding large amounts of bandwidth have blocking probabilities higher than those with low requirements. Moreover, wideband classes are also particularly sensitive to overload in the traffic from other narrowband classes. This paper discusses a strategy of bandwidth reservation in order to improve blocking characteristics and overload robustness in multirate loss networks.**

Keywords— **Multirate loss networks, bandwidth reservation, call blocking probabilities.**

I. Introduction

THE analysis of call blocking at a link with distinct capacity requirement classes as well as distinct arrival rates and holding times is presented in this paper. The determination of trunk reservation parameters which provide an important mechanism to control the relative call blockings to desired levels is based on asymptotic overload conditions. An unbuffered link having capacity C which is shared by several distinct service classes is considered. Calls of each service arrive in a Poisson stream and request a fixed integral amount of capacity, which may depend on the service class. An arriving call is blocked and lost if enough capacity is not available. Otherwise, some resource amount (equal to the effective bandwidth of the call class) is allocated for a period equal to the duration of the call.

Quality of service (QoS) is rapidly becoming a feature of fundamental importance in the increasing wideband networks. Some of the recently developed services rely upon the ability to establish communication channels with characteristics that satisfy a given number of parameters (e.g. bandwidth, maximum jitter, maximum latency) during the data transmission. An efficient approach is to establish a *virtual connection* between end-to-end terminals, as native ATM networks do.

From the point of view of dimensioning, if the traffic belonging to an end-to-end connection can be accurately modeled by an effective bandwidth then a virtual connection behaves as a circuit-switched connection (like the plain old telephone line calling) [1], [2], [3], [4]. Therefore, similar procedures can be used to calculate both the types of networks if one can take into account the distinct requirements of resources (bandwidth or number of circuits) for a given requisition. The incoming traffic can be grouped into a finite number of classes sharing some common characteristics. The notion of effective bandwidth is a useful and practical tool for connection admission control (CAC) and capacity planning in telecommunication networks.

L. R. Bercho, I. S. Bonatti, P. L. D. Peres and A. Lopes are with the School of Electrical and Computer Engineering, University of Campinas, P.O. Box 6101, 13081-970, Campinas SP, Brazil. Phone: (19) 37883759, Fax: (19) 32891395, E-mails: {bercho@dt,ivanil@dt,peres@dt,amauri@decom}.fee.unicamp.br

For a full shared resource environment, exact blocking probabilities calculation can be determined as the classical product form solution or by the recursion formula proposed by Kaufman [5] that greatly enhances calculation performance. Approximations for heavy traffic networks were proposed in [6], [7], [8]. Asymptotic approximations for large networks have been obtained in [9].

All these methods are suitable for network dimensioning where nominal blocking probabilities are the only parameters under analysis. However, despite the lower computational requirements of many approximate methods, the accuracy may not be adequate when used as a part of other network algorithms, such as *Fixed Point* equation system, with noticeable errors. In addition, the accurate analysis of blocking caused by variations in offered traffic or overload situations is better achieved through exact methods, provided that the numerical complexity remains tractable.

In this paper, a strategy of resources reservation is proposed in order to influence call blocking probabilities and their dependence on traffic variation. Trunk reservation is a technique commonly used in circuit-switched networks to prevent performance degradation under overload conditions. The usage of trunk reservation in multi-rate systems has been often discussed in literature [10],[11], [12], [13], [14], [15], but two correlated main problems are not well solved yet: exact blocking probabilities computation (there are only approximations); and obtaining a simple procedure for dimensioning the reservation capacity.

As shown in this paper, a simple strategy based on Kaufman's formula, and a reservation heuristic provide an effective method for network link dimensioning. The quality of the protection against overload can be validated by the product form calculation in order to check the efficacy of the heuristic proposed.

The remaining sections of the paper are organized as follows: In section 2 the full sharing results are refreshed; then the dimensioning strategy is proposed in section 3 and its performance is illustrated by some examples; the conclusion section closes the paper.

II. Full Sharing Strategy

A *multirate link* contains a finite number C of unbuffered resources (circuits or basic bandwidth units) available, which are shared among k distinct classes. Calls of each class arrive in a Poissonian stream and require a fixed, integer amount of resources, which in turn may depend on the service class. An arriving call is blocked and lost if there is not enough available resources at the arrival time. Otherwise, the call requirements is held for the duration of the call. Therefore, the parameters defining each class are: the mean arrival rate λ_i; the mean holding time $1/\mu_i$ and the resource requirements c_i, for $i = 1, \ldots, k$. The nominal traffic load of each class is defined by the quantity

$$\bar{a}_i = \frac{\lambda_i}{\mu_i} \; ; \quad i = 1, \ldots, k \qquad (1)$$

The link states can be described by the vector $n = [n_1, \cdots, n_k]$, where n_i denotes the number of requisitions under service for the class i. Also, a k-dimensional set denominated Ω will contain all possible vector states, provided that the following assumptions are made:

- the arrivals are independent of the state of the system and the arrival process is Poissonian;
- blocked calls are cleared;
- the Laplace transform of the service time distribution for each class exists;
- the departures are never blocked, i.e. the space of allowed states Ω is coordinate convex (if $n \in \Omega$ and $n_i > 0$, then $n - e_i \in \Omega$, where e_i contains 1 in position i and 0 elsewhere).

Kaufman s formula

An efficient recursive procedure was proposed by Kaufman [5] to calculate the blocking on a multi-rate environment where distinct service classes share a common link without any reservation. The solution of the model is based on a mapping from the k-dimensional state space Ω into a unidimensional state space producing the state probabilities

$$P_j = \tilde{P}_j \left(\sum_{m=0}^{C} \tilde{P}_m \right)^{-1} \qquad (2)$$

where $\tilde{P}_j$ are the unnormalized probabilities satisfying the recursive algorithm

$$\tilde{P}_m = \begin{cases} 0 & ; \; m < 0 \\ 1 & ; \; m = 0 \\ \dfrac{1}{m} \displaystyle\sum_{i=1}^{k} \bar{a}_i \, c_i \, \tilde{P}_{m-c_i} & ; \; 0 < m \leq C \end{cases} \qquad (3)$$

The blocking probabilities for the service classes are

$$B_i = \sum_{m=C-c_i+1}^{C} P_m \; ; \quad i = 1, \ldots, k \qquad (4)$$

The equations (2), (3) and (4) are summarized by the Kaufman's formula (sometimes named it Multirate Erlang-B formula)

$$B = \mathcal{K}(a, c, C) \qquad (5)$$

where a is the load traffic vector, c is the effective band units vector, C is the link capacity and B is the blocking probability vector.

The complexity of the Kaufman's formula computation is $O(kC)$ [16] and its computation efficiency has been improved by some papers in the literature [17]. Note that the Kaufman's formula is the very known Erlang's B formula if one single class of traffic is offered to the link.

For the ordered service classes it can be shown from equation (4) that the blocking probabilities satisfy

$$c_1 > \cdots > c_k \Rightarrow B_1 > \cdots > B_k \qquad (6)$$

The blocking experienced by the low bandwidth classes are smaller than that experienced by the high bandwidth classes. Figure 1 illustrates the full sharing policy unfairness for an example with three service classes, whose parameters are shown in Table I. Note in Figure 1 that the class with the highest blocking probability is the one with the largest bandwidth requirement.

By means of the Kaufman's formula it is possible to calculate a link capacity where every class blocking probability remains bellow a given threshold. If a superior limit of 2% blocking is specified for all the three classes, then it will take a capacity of at least $C = 58$ bandwidth units. The resulting nominal blocking probabilities for this capacity are $B_1 = 1.75\%$, $B_2 = 0.71\%$, and $B_3 = 0.32\%$

TABLE I

THE PARAMETERS FOR AN EXEMPLE WITH THREE SERVICE CLASSES.

Class	c_i	a_i
1	4	3.0
2	2	6.0
3	1	12.0

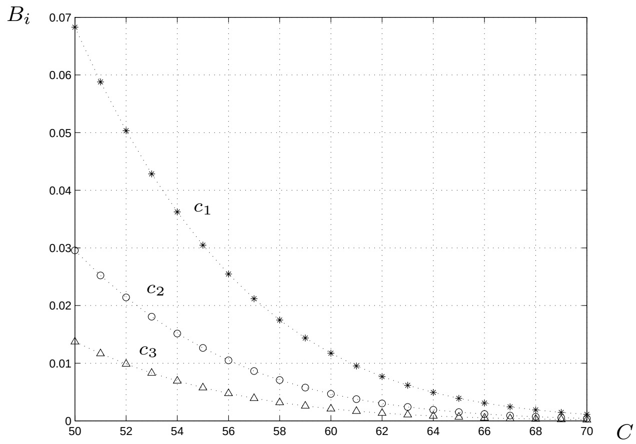

Fig. 1. Blocking probabilities as function of the link capacity.

Additionally, a full sharing policy for link access disturbs more severely the largest bandwidth class when the system is under overload conditions. To analyse the link performance under overload conditions, consider the traffic for each class a_i defined as:

$$a_i = \bar{a}_i(1 + \alpha v_i) \qquad (7)$$

where $\bar{a}_i$ is the nominal value of the traffic; v_i is equal to 1 if the i-class is overloaded and is zero otherwise; and α is a variable ranging from 0 to the maximum expected overload.

Figure 2 shows the class blocking probabilities as a function of overload in the traffic of the classes 2 and 3, that is, $v = [0\ 1\ 1]$. The total number of available resources is $C = 58$. Note that B_1 (upper curve) grows faster than the others, even though its own traffic remains stationary at the nominal value.

An increase of 50%, for instance, in offered traffics 2 and 3 would result $B_1 = 12.0\%$, $B_2 = 5.5\%$, and $B_3 = 2.6\%$.

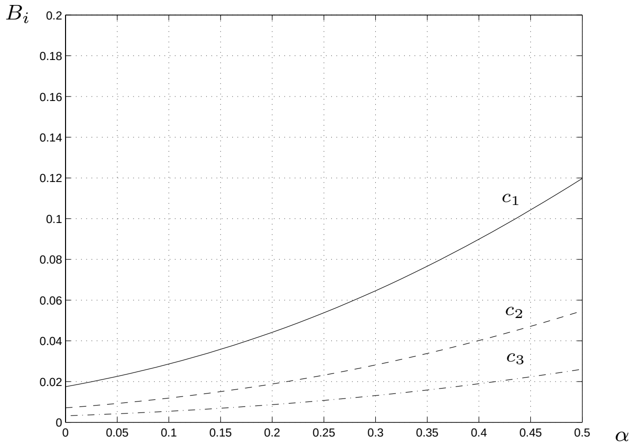

Fig. 2. Blocking probabilities as function of the traffic overload $v = [0\ 1\ 1]$ (link capacity $C = 58$).

Overload impact on class 1 (the largest bandwidth class) is even worse if all traffics are overloaded, as suggested in Figure 3 for the same number of resources $C = 58$. In this case, an overload of 50% on every class would produce $B_1 = 20\%$, $B_2 = 9.7\%$, and $B_3 = 4.7\%$. To overcome this problem a simple reservation strategy is proposed to protect the largest bandwidth class.

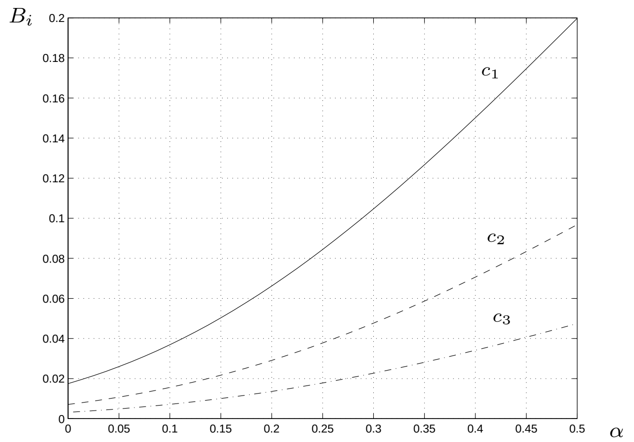

Fig. 3. Blocking probabilities as function of the traffic overload $v = [1\ 1\ 1]$ (All classes).

III. Reservation Strategy

In order to influence the call blocking probabilities and their dependence on traffic variation, the capacity C and the reservation R_1 are computed by the simple algorithm described below.

First step: For a given specified (nominal) bound vector $\bar{B}$, compute C by

$$\begin{cases} \min\ C \\ \mathcal{K}(a, c, C) \leq \bar{B} \end{cases} \tag{8}$$

where the inequality applies simultaneously for all the entries of the vectors.

Second step: For a given asymptotic specified bound $\bar{b}_1$ over the blocking probability B_1 on extremely overloading conditions for all classes, except class 1, compute R_1 by

$$\begin{cases} \min\ R_1 \\ \mathcal{K}(a_1, c_1, R_1) \leq \bar{b}_1 \end{cases} \tag{9}$$

The following procedure solves both the steps of the proposed algorithm with minor modifications in the procedure for the second step.

```
procedure unidimensional_search;
  {
     let
          C_M ← upper bound for C
          C_m ← lower bound for C

     while  C_M − C_m > 1
       {
          let  C ← ⌊(C_M + C_m)/2⌋

          if    K(a, c, C) > B̄
            then  C_M ← C
            else  C_m ← C
       }
  }
```

where $\lfloor x \rfloor$ denotes the largest integer lower than x.

The initial lower bound C_m can be evaluated by the sum over all classes of the carried traffic estimate of each class multiplied by its bandwidth unit

$$C_m = \sum_{i=1}^{k} a_i\ c_i\ (1 - \bar{B}_i) \tag{10}$$

and the initial upper bound C_M by [18]

$$C_M = C_m + \max_{i=1,\ldots,k} \bar{B}_i^{-1} \tag{11}$$

A reservation R_1 of 20 circuits is obtained when the algorithm proposed is applied to the example shown in Table I for $\bar{b}_1 = 12\%$.

Note that the numerical complexity of the algorithm proposed for the determination of the capacity C and reservation R_1 is quite low. However, only the asymptotic behavior of the blocking probabilities are determined when computing the two steps of the algorithm. In a synthesis context (planning or dimensioning a network) these estimates are generally sufficient.

For a more detailed analysis the resulting blocking probabilities are computed for several distinct load conditions in next section.

Product Form Solution

Kaufman's formula does not work properly when circuit reservation is used; actually, the exact computation of the blocking probabilities becomes a very hard problem that can be solved only for a small number of particular reservation strategies.

If the access strategy to the link resources (including reservation or not) has bidirectional transitions between all neighboring states in Ω then the Markovian equilibrium equations have the local balance property [19] (see the appendix).

This condition excludes several reservation policies in which symmetry of transition is violated for some states. Access strategies that preserve such symmetry include static reservation where *at least* R_i resources are segregated for a class i and the remaining C_0 resources are shared among all classes. The *at least* reservation strategies can be defined by

$$\Omega = \begin{cases} 0 \leq \sum_{i=1}^{k} n_i c_i \leq C = \sum_{i=1}^{k} R_i + C_0 \\ \\ 0 \leq n_i c_i \leq R_i + C_0 \quad ; \quad i = 1, \cdots, k \end{cases} \tag{12}$$

Note that the algorithm proposed is a particular case of the *at least* strategy and thus it has the local balance property. The local balance property allows the determination of the states probabilities via the Product Theorem ([20], [21]) as

$$P_n = \prod_{i=1}^{k} \frac{a_i^{n_i}}{n_i!} \cdot G^{-1}(\Omega) \tag{13}$$

with

$$G(\Omega) = \sum_{n \in \Omega} \left(\prod_{i=1}^{k} \frac{a_i^{n_i}}{n_i!} \right) \tag{14}$$

and

$$B_i = \frac{G(\Omega_i^B)}{G(\Omega)} \tag{15}$$

where $\Omega_i^B \subset \Omega$ denotes the set of blocking states for each class.

The numerical complexity of the Product Theorem method is $O(C^k)$ [16], which limits its use to systems containing only few classes.

Figure 4 shows the class blocking probabilities as a function of overload traffic $v = [0\ 1\ 1]$ of the classes 2 and 3 in the same link with capacity $C = 58$ resource units. The nominal blocking probabilities become $B_1 = 1.6\%$, $B_2 = 1.8\%$, and $B_3 = 0.75\%$ which means that reservation produces a slight change in the original full sharing condition. In counterpart, under overload conditions the B_1 curve grows only moderately compared with the values obtained in Figure 2. For instance, an increase of 50% in both traffics 2 and 3 produces $B_1 = 5.5\%$, $B_2 = 15\%$, and $B_3 = 5.5\%$.

This reservation strategy has proper behavior when all traffics are simultaneously overloaded (that is $v = [1\ 1\ 1]$), as shown in Figure 5. Note that all curves increase similarly.

The reservation strategy does not impose severe penalties on the remaining traffics if only the largest bandwidth traffic is overloaded, as shown in Figure 6.

IV. Conclusion

A methodology for the computation of both the capacity and the reservation in a link receiving multi-rate service classes using simple numerical operations has been presented. The exact blocking probabilities under several traffic overload scenarios was computed via the Product Theorem methodology. Experiences with case studies have shown that the proposed strategy is a robust and computationally efficient technique for link dimensioning.

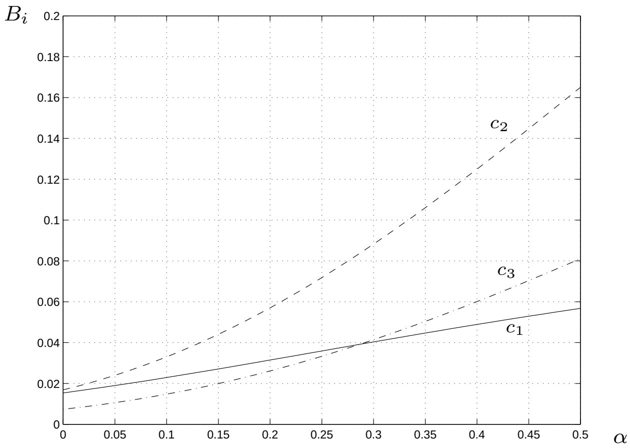

Fig. 4. Blocking probabilities as function of the traffic overload $v = [0\ 1\ 1]$ ($C = 58$, $R_1 = 20$).

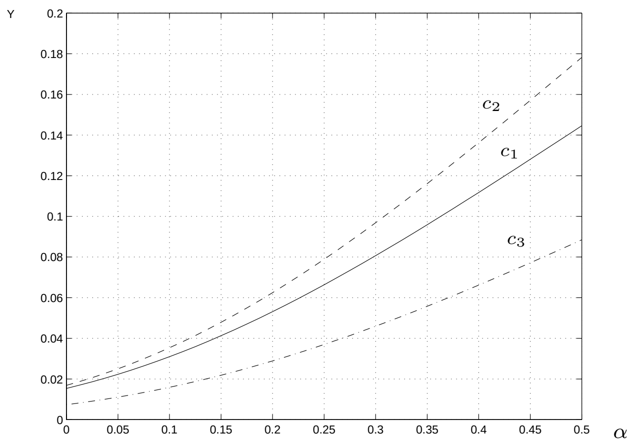

Fig. 5. Blocking probabilities as function of the traffic overload $v = [1\ 1\ 1]$ ($C = 58$, $R_1 = 20$).

V. Acknowledgment

This research has been supported by grants from "Conselho Nacional de Desenvolvimento Científico e Tecnológico" – CNPq, Brazil.

Appendix

This appendix presents a brief analysis of two reservation strategies, *at least* and *the last* circuit reservation, to emphasize the differences between them and to show that it is not possible to apply the Product Theorem to the last circuit reservation strategy.

Consider a link with two circuits shared by two Poissonian traffics of arrival rates λ and β. The λ is the protected traffic and the β is the intruder one. An accepted call holds a circuit during a random exponential distributed time with mean μ^{-1}. Blocked arriving calls are lost and cleared.

The *at least one in two* strategy means that at least one circuit must be reserved to the λ traffic. An arriving call β can seize a

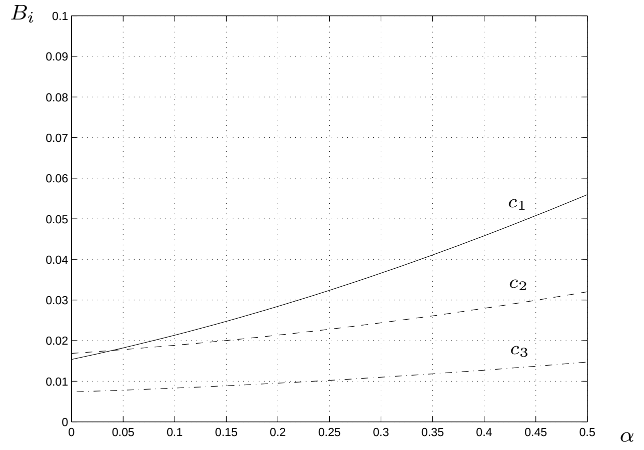

Fig. 6. Blocking probabilities as function of the traffic overload $v = [1\ 0\ 0]$ ($C = 58$, $R_1 = 20$).

circuit only if one or both the circuits are free. This situation is characterized by

$$C = 2\ ;\ R_\lambda = 1\ ;\ R_\beta = 0 \Longrightarrow C_0 = 1 \qquad (16)$$

The set Ω is defined as

$$\Omega = \{n_\lambda + n_\beta \leq C = 2\ ;\ n_\lambda \leq C_o + R_\lambda = 2\ ;\ n_\beta \leq C_o + R_\beta = 1\} = \{00, 10, 20, 01, 11\} \qquad (17)$$

The sets Ω^λ (blocking states for the traffic λ) and Ω^β (blocking states for the traffic β) are giving by

$$\Omega^\lambda = \{20, 11\}\ ;\ \Omega^\beta = \Omega^\lambda + \{01\} \qquad (18)$$

Figure 7 shows the state diagram (the dashed transition applies) for the *at least one circuit in two* strategy. Note that the transi-

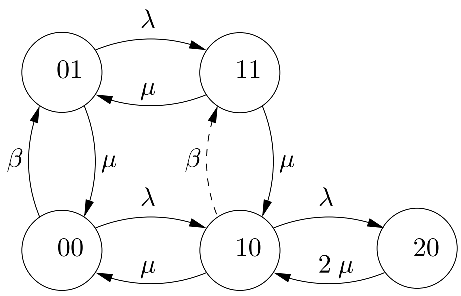

Fig. 7. State transition diagram for a link with two circuits: λ is the protected traffic and β is the intruder traffic.

tions between any two adjacent states exist in both directions.

The Product Theorem (unnormalized probabilities) yields

$$\tilde{P}_{00} = 1\ ;\ \tilde{P}_{10} = \rho\ ;\ \tilde{P}_{01} = \rho_i\ ;\ \tilde{P}_{11} = \rho\rho_i\ ;\ \tilde{P}_{20} = \frac{\rho^2}{2} \qquad (19)$$

where $\rho = \lambda/\mu$ and $\rho_i = \beta/\mu$. These values can be validated by solving the Markov equilibrium equations.

The limiting case ($\beta \to +\infty$) yields

$$\lim_{\rho_i \to +\infty} B_\lambda = \frac{\rho}{1+\rho} \qquad (20)$$

Note that equation (20) is the Erlang-B formula for one circuit.

The *last one circuit in two* strategy means that β traffic never take the last free circuit no matter how many circuits were taken by the λ traffic. Figure 7 shows the states transitions for this strategy (the dashed transition no more applies).

The Ω and Ω^λ sets are identical in both strategies. However, the blocking states for β traffic are now given by

$$\Omega^\beta = \{01, 10, 11, 20\} \qquad (21)$$

Due to the nonexistence of bidirectional transitions between any two adjacent states in Ω, the Product Theorem no longer applies. The only way to calculate the probabilities is by solving the Markov equilibrium equations. This solution, in the limiting case ($\beta \to +\infty$), yields

$$\lim_{\rho_i \to +\infty} B_\lambda = \frac{\rho}{2+\rho} \qquad (22)$$

It is important to underline that for the same reservation threshold the *the last* strategy produces a more efficient protection to the λ traffic than that the *at least* strategy, as seen in Figure 8.

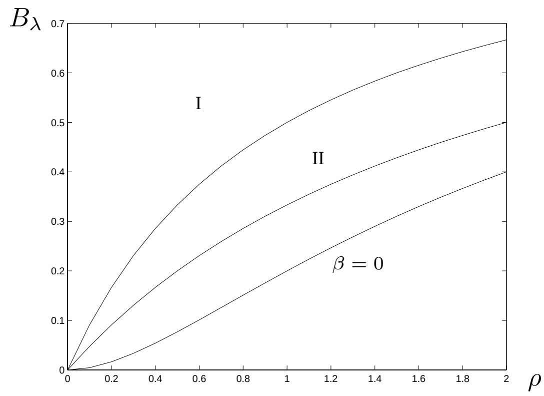

Fig. 8. The λ traffic blocking probabilities for $\beta \to +\infty$ (I *at least*, II *the last*) and for $\beta = 0$.

REFERENCES

[1] A. W. Berger and W. Whitt W, "Effective bandwidths with priorities," *IEEE-ACM Transactions on Networking*, vol. VOL. 6, no. NO. 4, pp. 447–460, August 1998.

[2] W.-Y. Choi and C.-H. Jun, "A Bernoulli process approximation with buffer size adjustment for the analysis of an ATM multiplexer," *IEEE Letters on Communications*, vol. 1, no. 5, pp. 149–151, September 1997.

[3] F. P. Kelly, "Notes on effective bandwidths," in *Stochastic Networks: Theory and Applications*, F. P. Kelly, S. Zachary, and I. B. Ziedins, Eds., pp. 141–168. Oxford University Press, 1996.

[4] G. Kesidis, J. Walrand, and C.-S. Chang, "Effective bandwidths for multiclass Markov fluids and other ATM sources," *IEEE Transactions on Networking*, vol. 1, no. 4, pp. 424–428, August 1993.

[5] J. S. Kaufman, "Blocking in a shared resource environment," *IEEE Transactions on Communications*, vol. 29, no. 10, pp. 1474–1481, 1981.

[6] N. G. Bean, R. J. Gibbens, and S. Zachary, "Asymptotic analysis of single resource loss systems in heavy traffic, with applications to integrated networks," *Advances in Applied Probability*, vol. VOL. 27, no. NO. 1, pp. 273–292, March 1995.

[7] J. A. Morrison, K. G. Ramakrishnan, and D. Mitra, "Refined asymptotic approximations to loss probabilities and their sensitivities in shared unbuffered resources," *SIAM Journal on Applied Mathematics*, vol. VOL. 59, no. NO. 2, pp. 494–513, December 1998.

[8] A. Simonian, J. W. Roberts, F. Theberge, and R. Mazumdar, "Asymptotic estimates for blocking probabilities in a large multi-rate loss network,"

Advances in Applied Probability, vol. VOL. 29, no. NO. 3, pp. 806–829, September 1997.

[9] S. Chung K. W. Ross, “Reduced load approximations for multirate loss networks,” *IEEE Transactions on Communications*, vol. VOL. 41, no. NO. 8, pp. 1222–1231, August 1993.

[10] S. Oueslati-Boulahia and J. W. Roberts, “Impact of ”trunk reservation” on elastic flow routing,” *Lectures Notes on Computing Science - NETWORKING 2000*, vol. 1815, pp. 823–834, 2000.

[11] K. Lindberger, “Dimensioning and design methods for inegrated atm networks,” in *Proceedings of 14th International Teletraf˘c Congress - ITC 14*, 1994, vol. 1b, pp. 897–906.

[12] A. Gersht and K. J. Lee, “Virtual-circuit load control in fast packet switched broadband networks,” in *Proceedings of IEEE Globecom 89*, 1989, pp. 214–220.

[13] S. A. Johnson, “A performance analysis of integrated communication systems,” *British Telecom Technical Journal*, vol. VOL. 3, no. NO. 4, October 1985.

[14] K. Lindberger, “Simple approximations of overflow system quantities for additional demands in the optimization,” in *Proceedings of 10th International Teletraf˘c Congress - ITC 10* , 1983, pp. 5.3–3.

[15] J. W. Roberts, “Teletraffic models for the telecommunication integrated service networks,” in *Proceedings of 10th International Teletraf˘c Congress - ITC 10*, 1983, pp. 1.1–2.

[16] J. Roberts, U. Mocci, and J. Virtamo, *Broadband network teletraf˘c - performance evaluation and design of broadband multiservice networks - ˘nal report of action COST 242* , Springer-Verlarg, Berlin, Heidelberg, New York, 1996.

[17] A. A. Nilsson and M. J. Perry, “Multirate blocking probabilities: Numerically stable computations,” in *Proceedings of 15th International Teletraf˘c Congress - ITC 15*, Elsevier Science, Ed., Washington, DC, USA, 1997, pp. 1359–1368.

[18] S. A. Berezner, A. E. Krzesinski, and P. G. Taylor, “On the inverse of Erlang’s function,” *Journal of Applied Probability*, vol. VOL. 35, pp. 246–252, 1998.

[19] K. M. Chandy, “The analysis and solutions for general queueing networks,” in *Proceedings of the Sixth Annual Princeton Conference on Information Sciences and Systems*, Priceton University, USA, March 1972, pp. 224–228.

[20] Leonard Kleinrock, *Queueing Systems Volume II: Computer Applications*, John Wiley & Sons, Inc., New York, 1976.

[21] J. T. Virtamo, “Reciprocity of blocking probabilities in multiservice systems,” *IEEE Transactions on Communications*, vol. 36, no. 10, pp. 1174–1175, 1988.