FACTA UNIVERSITATIS (NIŠ)
SER.: ELEC. ENERG. vol. 24, no. 1, April 2011, 1-9

One Approach in Evaluating the Overflow Probability Using the Infinite Fluid-Flow Queue

Branimir Trenkić and Mihajlo Stefanović

Abstract: If it is going to have practical significance, the evaluation of overflow probabilities must be (1) precise, (2) computationaly stable and (3) real time compatible. The existing approximation expressions are gotten by using the traditional fluid-flow analytic techniques, which are mostly based on spectral analysis. The limitation of this approach comes from numerical difficulties caries by the spectral approach.

In this paper, we suggested a fluid-flow approach in evaluating the overflow probability, which fully satisfies the above stated criteria, and removes the numerical difficulties of existing methods. The authors approach is based on the renewal argument and exploiting the similarity between fluid queues and Quasi-Birth-and-Death (QBD) processes.

Keywords: Fluid queues, Quasi-birth-and-death processes, Approximating method, ON/OFF model.

1 Introduction

In this paper, the description of the procedure for getting the approximate formula is given, with which we can execute a precise evaluation of overflow probability in infinite buffer fluid queue driven by a Markovian environment. Generally, the asymptotic approximation of overflow probability is based on the asymptotic decay rate of the buffer content (r.v. X) tail probabilities, and has the following exponential form:

$$G(x) = P(X > x) \sim \beta^{-\eta x}, \qquad \text{as} \quad x \to \infty \tag{1}$$

Manuscript received on December 3, 2010.

B. Trenkić is with IRITEL, Batajnički put 23, 11000 Belgrade, Serbia (e-mail: trenkic@iritel.com). M. Stefanović is with University of Niš, Faculty of Elektronic Engineering, Aleksandra Medvedeva 14, 18115 Niš, Serbia, (e-mail: mihajlo.stefanovic@elfak.ni.ac.rs).

where η is a positive constant called *the asymptotic decay rate*, β is a positive constant called *the asymptotic constant*, and $f(x) \sim g(x)$ as $x \to \infty$ means that $f(x)/g(x) \to 1$ when $x \to \infty$; see [1].

From the above stated, the sequence of activities in the procedure of getting the asymptotic approximation of overflow probabilities is naturally imposed. The first step is the finding of a suitable expression for the exact calculation of the buffer content tail probabilities $G(x)$. After that, we examine the behavior of $G(x)$, the probability of overflow beyond x, for large values of x. The explained procedure, using the traditional fluid-flow analytic techniques (spectral analysis) is given in [2]. In general, when the input rate is characterized by an N-state Markov chain, the distribution of the buffer content probabilities is of the form:

$$F(x) = P(X \leq x) = \sum_{i=1}^{N} a_i \Phi_i e^{z_i x} \tag{2}$$

where the z_i and Φ_i are, respectively, generalized eigenvalues and eigenvectors associated with the solution of the differential equation satisfied by the steady state probabilities of the system, and the a_i's are coefficients determined from boundary conditions. The next step is consisted in the asymptotic analysis of overflow probability given from expression (2), $(G(x) = 1 - F(x))$. The key result of this analysis is that this quantity can be further approximated by only considering the contribution of the term corresponding to the largest negative, or dominant eigenvalue in (2). Under such assumptions, the buffer overflow probability, $G(x)$, is known [2], [3] to be of the form

$$G(x) \sim \beta e^{z_0 x} \tag{3}$$

where the z_0 is the largest negative eigenvalue. As for the asymptotic constant β, the case is not as simple. Namely, getting the expression which precisely approximates the value of constant β demands the determination of all real eigenvalues (see [2]).

The limitation of this approach comes from the fact that such eigenvalues are of both signs, and therefore numerical errors may lead to solutions that are unstable (computed probabilities become negative).

To exceed the aroused problems we often resort to one-parameter approximation (approximate β by 1). Logically as we shall see this approximation weaken the precision of the evaluation of overflow probabilities.

The noticed numerical difficulties carried by the spectral approach, urged the authors of this paper to the idea to use analytic techniques, that are much more computationaly stable, in the finding of an approximation formula for the overflow probability. The natural choice of the authors is the matrix-geometric analytic techniques.

The results originated from the idea of exploiting the similarity between fluid queues and QBD processes, and which are published in multiple papers [4], [5], [6] allow the usage of this technique.

The organization of the paper is as follows. In Section 2, we give the precise definition of a fluid queue and we summarize some basic results from the literature [4] and [6], among else, the expression for the stationary density vector of fluid buffer content using Markov-renewal approach. That is the starting point of our work, which is in detail explained in Section 3. Section 4 through a numeric example gives graphical result, comparing our expression with the exact results from the 'standard' formula of Anick, Mitra and Sondhi, [2]. These show that our new formula provides excellent accuracy for the load values at which queuing becomes important.

2 The renewal approach to fluid queues

Markov-modulated fluid queues $\{(X(t),\varphi(t)) : t \in R^+\}$ are two-dimensional Markov processes of which the first component $X(t)$ is called the *level* at time t and the second component $\varphi(t)$ is called the *phase*. The level represents the content of a buffer containing fluid; the phase is the state of an irreducible Markov process on a finite state space, evolving in the background. In the simplest case, the content of the buffer varies linearly in time, according to the state of $\{\varphi(t)\}$: in the intervals of time during which the phase remains equal to $i \in S$, say, the level varies linearly at the rate r_i. We assume that the rates r_i take any real value, except zero because this assumption is not restrictive (see [6]). We decompose the set of phases S into two disjoint subsets S_+ and S_- , where $S_+ = \{i \in S | r_i > 0\}$ and $S_- = \{i \in S | r_i < 0\}$; its infinitesimal transition generator is denoted by Q and is decomposed in a conformant manner:

$$Q = \begin{pmatrix} Q_{++} & Q_{+-} \\ Q_{-+} & Q_{--} \end{pmatrix} \tag{4}$$

We denote by ξ the steady state probability vector corresponding to the generator Q; it is the unique solution of the system $\xi Q = \mathbf{0}, \xi \mathbf{1} = 1$. Throughout the paper, we use $\mathbf{0}$ and $\mathbf{1}$ respectively for vectors of zeros and ones, of the appropriate dimension. Also, parts of vectors and matrixes that arise in the decomposition as in (4), we shall note in the same way.

The stationary mean drift of the fluid queue (X,φ) is defined by $\mu = \xi_+ \mathbf{r}_+ + \xi_- \mathbf{r}_-$, where $\mathbf{r} = \{\mathrm{r_j} : \mathrm{j} \in \mathrm{S}\}$.

Define the matrix $T = C^{-1}Q$ (where $C = \mathrm{diag}(|r_i| : i \in S)$), actually is the generator of a fluid queue with net rates equal to +1 or -1, obtained from (X,φ) by changing the time scale ant the input rates but the fluid level changing by the same amount overall.

Denote by $p = (0, p_-)$ the steady state probability mass vector of the empty buffer (as the fluid queue instantaneously leaves the level zero if the phase is in S_+, $p_+ = 0$).

The density $\pi(x)$ and the probability vector p_- are expressed in terms of a matrix, denoted as Ψ and defined as follows: for i in S_+ and j in S_-, Ψ_{ij} is the probability that, starting from $(0,i)$ at time 0, the fluid queue returns to the level zero in a finite amount of time and does so in phase j. The results that we will further use in our paper, we can sum up in form of one theorem (Theorem 2.1 in [5])

Theorem1. If $\mu < 0$, then the stationary density of the buffer content of the process (X, φ) is given by

$$\pi(x) = p_- Q_{-+} e^{Kx} [C_+^{-1}, \Psi C_-^{-1}] \tag{5}$$

for $x > 0$, where

$$K = T_{++} + T_{-+} \tag{6}$$

The vector p_- is the unique solution of the system

$$p_-(Q_{--} + Q_{-+}\Psi) = 0$$

$$p_-(1 - Q_{-+}K^{-1}[C_+^{-1}, \Psi C_-^{-1}]e) = 1,$$

The matrix Ψ is the minimal non-negative solution of the equation

$$T_{+-} + T_{++}\Psi + \Psi T_{--} + \Psi T_{-+}\Psi = 0. \tag{7}$$

Very efficient algorithms exist to solve; see [4] and [6].

3 Proposed method

The theory presented in the previous Section, and summed up in Theorem 1, does not give an explicate expression for the queue-length tail probabilities which is suitable for asymptotic analysis. The first step will be the acquiring of such an expression. By applying the result of the presented theorem and with simple analytic manipulations, we can get an expression for the exact calculation of the queue-length tail probabilities, $G(x) = P[X > x]$ (i.e. overflow probabilities). From (5) by definition we have that,

$$\begin{aligned} G(x) = \int_x^\infty \pi(t)dt = \int_x^\infty p_- T_{-+} e^{Kt} [C_+^{-1}, \Psi C_-^{-1}] \mathbf{1} dt = \\ = p_- T_{-+} (-K)^{-1} e^{Kx} [C_+^{-1}, \Psi C_-^{-1}] \mathbf{1}. \end{aligned} \tag{8}$$

Equation (8) is the starting point of our evaluation of overflow probabilities which is based on the asymptotic decay rate.

Matrix K has the key role in asymptotic analysis of (8). The definition and the detailed properties of this matrix are given in the papers [5] and [6]. Here we will mention only one property which is important for our future work - all the eigenvalues of this matrix have strictly negative real parts and among them there is one, denote by ζ, which is real and has the geometrical and algebraic multiplicities equal to one. It is maximal in the sense that every other eigenvalue of matrix K has a real part which is strictly less than ζ. Furthermore, there exist real, strictly positive, left and right eigenvectors of K for the eigenvalue ζ, which we will denote by w and z respectively, and which are normalized by $w \cdot \mathbf{1} = w \cdot z = 1$. Using this feature of the matrix K and result given in [5], that the same is valid for matrix e^{Kx} (for some level x), we may write

$$e^{Kx} \sim \; e^{\zeta x} zw, \qquad \text{as} \quad x \to \infty \tag{9}$$

On the basis of (8) and (9) a natural conclusion imposes about the asymptomic decay rate of tail probabilities $G(x)$. Namely, the value of the asymptomic decay rate parameter from the equation (1) determines identified eigenvalue of the matrix K, i.e. $\eta = -\zeta$. On substituting in (8) the expression in (9), the general form of that expression will be,

$$G(x) = P[X > x] \sim \beta e^{\zeta x} \tag{10}$$

Here we will give only one remark that is concerned with the comparison of the result gotten by using the traditional approach and the explained (our) approach, which concern the asymptomic decay rate. Although the results are identical (with opposite signs), the computationally effort is significantly reduced by the suggested method because matrix K (defined on states of S_+) is by definition surely smaller dimension that N. The traditional approach, let's remind our selves, is based on the calculation of the dominant eigenvalue of matrix which dimension is $N \times N$.

We are left with finding the asymptotic constant β in (1), for which there is and undivided opinion that it is a very complex operation. The start point of our evaluation of the asymptotic constant is the expression (10).

We shall, in this paper, show that with simple mathematical operations, using the results gotten in this Section, we can get an expression that enables the evaluation of the asymptotic constant in (1). From (10) follows directly,

$$e^{-\zeta x} G(x) \sim \beta \qquad (x \to \infty) \tag{11}$$

And now in (11) if we include the results from (8) and (9), when $x \to \infty$,

$$\begin{aligned} G(x) = P[X > x] \sim \beta e^{\zeta x} \\ e^{-\zeta x} G(x) \sim \beta \\ \beta \sim p_- T_{-+} (-K)^{-1} z w [C_+^{-1}, \Psi C_-^{-1}] \mathbf{1} \end{aligned} \tag{12}$$

Furthermore, from (9) follows the identity $-\frac{1}{\zeta} = w(-K)^{-1} z$. Using that identity by simple algebraic manipulations from (12) we get,

$$\beta \sim -\zeta p_- T_{-+} (-K)^{-2} [C_+^{-1}, \Psi C_-^{-1}] \mathbf{1} \tag{13}$$

And as a part of the gotten expression on the right side $(p_- T_{-+} (-K)^{-2} [C_+^{-1}, \Psi C_-^{-1}] \mathbf{1})$, it stands for in fact a formula for the calculation of mean queue-length (see [6]), asymptotic constant β we can approximate with $-\zeta E[X]$, i.e. the final expression for the asymptotic approximation of overflow probabilities is given by

$$G(x) = P[X > x] \sim -\zeta E[X] e^{\zeta x} \qquad (x \to \infty) \tag{14}$$

So the parameters of asymptotic approximation (1) are given with $\beta \sim -\zeta E[X]$ and $\eta = -\zeta$. The quality of the gotten asymptotic approximation, we shall show through a numerical example.

4 Numerical examples

With the explained numerical example we make possible the direct comparison our expression with the exact results given by 'standard' formula of Anick, Mitra and Sondhi, [2]. We take the well-known example of ON/OFF sources feeding data to the buffer of a common communication channel (see Fig. 1). The durations

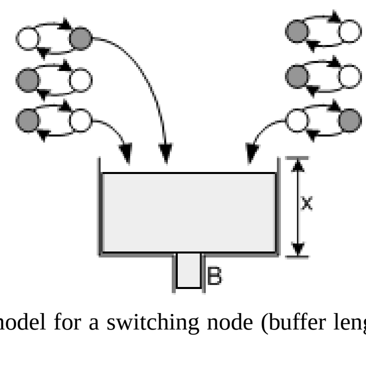

Fig. 1. Fluid-flow model for a switching node (buffer length x) under ON/OFF traffic sources.

of the ON and the OFF-periods are exponentially distributed, respectively with rates α and λ. Assume that the output channel capacity is B and that each source continuously feeds in data at the rate R during its ON-periods. Then, if i is the number of ON-sources, the net input rate to the buffer is $r_i = i \cdot R - B$, negative for small values of i, positive for large values.

We can also see that the traffic intensity ρ of the system can be represented by

$$\rho = \frac{NR\lambda}{B(\alpha+\lambda)} \tag{15}$$

For exponentially distributed ON- and OFF- periods, the source is furthermore completely characterized by three parameters, α, λ and R. The added parameter which we shall use in calculations is the ratio of the output channel capacity to an ON source's transmission rate, B/R (equation (13)).

The lead parameters of the system are sufficient for the exact calculation of values of overflow probabilities ([2]), so to get the approximated results of the same probabilities, by using our expression for the calculation of asymptotic parameters.

For the numerical examples, we fixed the parameters as $\alpha = 1$ and $\lambda = 0.4$ so we can get comparable results more simply.

Fig. 2 displays the exact tail probabilities $G(x)$ and the approximation (1), using our expressions for η and β, for three different values of traffic intensity $(\rho = 0.78,\ 0.85,\ 0.94)$. The gotten results show the significant precision of ap-

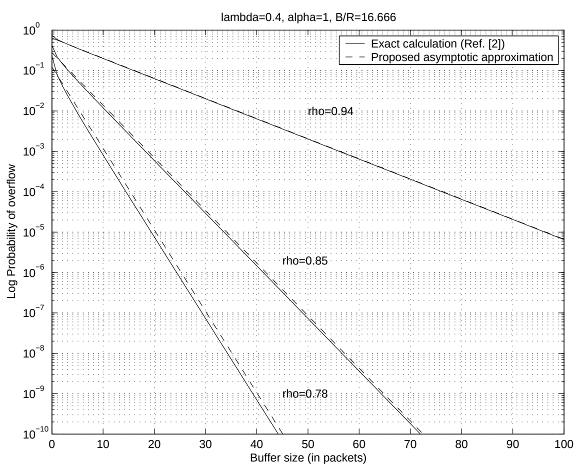

Fig. 2. Probability of overflow vs. buffer size, exact and asymptotic approximation. For $N = 55$, $N = 50$ and $N = 46$.

proximated values with the asymptotic parameters gotten by the expressions proposed in this paper. We can also notice that with the lowering of traffic intensity the precision of the suggested asymptotic approximation is weakened. As we already remarked, this does not lower the practical meaning of the suggested method of asymptotic approximation because the regions of high load those load values at which queuing becomes important.

Fig. 3. shows that the usage of one-parameter approximation $(\beta = 1)$, although it is often in literature that it is suggested as a accepted solution because of it's simplicity (see [1] and [3]), it is not always a good solution. Namely, it is given that one comparison of one-parameter approximation and two-parameter asymptotic approximation with the exact calculation of overflow probabilities. The values of the parameters in two-parameter approximations are gotten by using our method of evaluation. The gotten results confirm the importance of an adequate estimation of asymptotic constant in asymptotic approximation of the overflow probability.

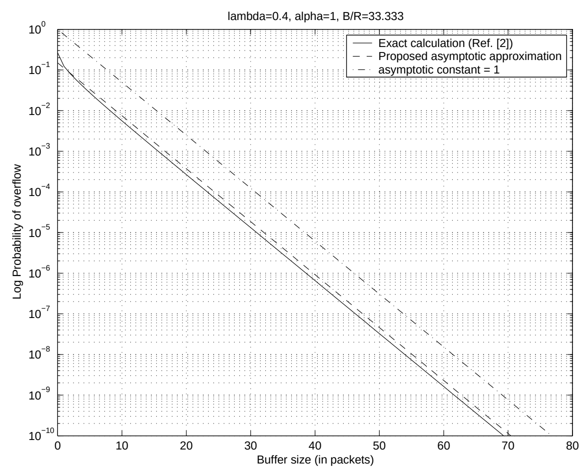

Fig. 3. Probability of overflow vs. buffer size, exact, proposed asymptotic approximation and asymptotic approximation with $\beta = 1$ (Ref. [3]). For $N = 100$ $(\rho = 0.85)$.

5 Conclusions

The contribution of this paper can be viewed from two aspects. Firstly, the claim that the asymptotic constant, if fed with the asymptomic decay rate, can simply be approximated with the multipliacation of that decay rate and the mean of queue-length - is not new (in some form taken in [1]). In this paper we, in a mathematicaly corect way, have showen, that in a fluid-flow enviroment, the presented claim is correct.

Secondly, the proposed method of approximation of overflow probability fully satisfies the criteria set in the begining of this paper. The procedure is precise and computationaly very simple. It gives a simple way to approximate both of the asimptomic parameters (β, η). Even the calculation of asymptotic decay rate parameter is less computationaly demanding than the traditional way - Reason: the dimension of the matrix K, as mentioned earlier. We specially underline that the equation for getting the aproximational value of the asimptomic constant can be directly mathematicaly derived from queue-length distribution.

References

[1] J. Abate, G. Choudhury, and W. Whitt, "Exponential approximations for tail probabilities in queues, i: Waiting times," vol. 43, no. 5, pp. 885–901, 1995.

[2] D. Anick, D. Mitra, and M. Sondhi, "Stoshastic theory of a data handling system with multiple sources," *Bell System Tech. Journal*, vol. 61, no. 8, pp. 1871–1894, 1982.

[3] R. Guerin, H. Ahmadi, and M. Naghshineh, "Equivalent capacity and its application to bandwidth allocationin high-speed networks," *IEEE Journal of Selected Areas in Communications*, vol. 9, no. 7, pp. 968–981, 1991.

[4] V. Ramaswami, "Matrix analytic methods for stochastic fluid flows," in *Teletraffic Engineering in a Competitive World (Proceedings of the 16th International Teletraffic Congress)*, Edinburgh, UK, 1999, pp. 1019–1030.

[5] S. Soares and G. Latouche, "Level-phase independence for fluid queues," *Stochastic Models*, vol. 21, no. 2, pp. 327–341, 2005.

[6] ——, "Further results on the similarity between fluid queues and QBDs," in *Proceedings of the 4th International Conference on Matrix-Analytic Methods*. River Edge, NJ: World Scientific, 2002, pp. 89–106.