LETTERS TO THE EDITOR

Dear Editor,

Overflow probability upper bound in fluid queues with general on–off sources

1. Introduction

In high-speed packet switched networks, a critical problem is to evaluate the probability of buffer overflow when the superposition of a finite number of variable bit rate sources is offered to a multiplexer. It is well known that fluid models, where discrete packet arrivals are assimilated to fluid flows, provide a useful description for studying such queues (see for example Kosten (1986), Norros et al. (1991), Simonian and Virtamo (1991), Bensaou et al. (1993)). Consider then a fluid reservoir with unlimited capacity which fills at instantaneous input rate Λ_t and empties at output rate c. In this letter, we focus on input processes (Λ_t) resulting from the superposition of N heterogeneous on–off fluid sources, when successive silence ('off') and activity ('on') periods constitute independent stationary sequences of i.i.d. random variables. In the following, the input rate of source i when active is denoted by h_i. The incoming workload on time interval $(0, t]$ is $W_t = \int_0^t \Lambda_s ds$ and variable V_0 represents the stationary distribution of the reservoir content.

The complementary distribution function $Q(x) = \boldsymbol{P}(V_0 > x)$ has been studied in various papers (Kosten (1986), Anick et al. (1982), Stern and Elwalid (1991)) for specific on and off period distributions. More general distributions have been considered in Bensaou et al. (1990), (1993) for evaluating the so-called Beneš upper bound $Q(x) \leqq q(x)$, where $q(x)$ is defined by

$$q(x) = \int_0^\infty \varphi(t, x + ct)dt \tag{1}$$

with

$$\varphi(t, w) = \sum_{\boldsymbol{\varepsilon}:\boldsymbol{\varepsilon}\cdot\boldsymbol{h}<c} (c - \boldsymbol{\varepsilon}\cdot\boldsymbol{h})\boldsymbol{P}(\Lambda_t = \boldsymbol{\varepsilon}\cdot\boldsymbol{h};\, w < W_t \leqq w + dw)/dw$$

where $\boldsymbol{\varepsilon}$ is any N-dimensional vector with 0–1 coordinates and $\boldsymbol{h} = (h_i)_{i \leqq N}$. The bound is tight when c is large or when traffic is light (i.e. when $\boldsymbol{P}(V_t = 0) \approx 1$). It is equivalent, though not identical, to the bound suggested in Stern and Elwalid (1991).

The contribution of the present letter is to provide an alternative expression for bound $q(x)$ as defined in (1). This new expression is written as a contour integral in the complex plane involving the double Laplace transform of the function φ. This formula is then applied, in particular, to Cox-type densities and enables us to write the bound $q(x)$ as an explicit sum of exponentially decreasing factors. The decreasing rates of the latter are defined as roots of equations involving the Laplace transforms of the densities of on and off periods.

2. Contour integral formula

Let v_i be the average stationary activity rate of source i and $\rho = \Sigma_{i \leqq N} h_i v_i / c$ the server load. We assume that the distributions of on and off periods have piecewise smooth densities and that their Laplace transform has a negative abscissa for absolute convergence. Let $\varphi^{**}[s, z]$ denote the double Laplace transform of $\varphi(t, w)$ with respect to variables t and w. Our main result can then be stated as follows.

Theorem 1. *Assuming that the stability condition $\rho < 1$ holds, we have*

$$q(x) = \frac{1}{2i\pi} \text{ p.v. } \int_{\mathscr{L}} \varphi^{**}[-cz, z] e^{zx} dz \tag{2}$$

where $\mathscr{L}$ is a vertical line in the complex plane strictly to the left of the imaginary axis and to the right of all the singularities of $\varphi^{**}[-cz, z]$ with negative real part (p.v. *stands for principal value*).

The proof of this theorem relies on the following preliminary proposition.

Proposition 1. *$q(x)$ is piecewise smooth and the abscissa for absolute convergence of its Laplace transform q^* is $-\sigma < 0$.*

The proof is given in Section 5. We apply the Laplace transform inversion formula (Churchill (1972)) to q^*, which reads

$$q(x) = \frac{1}{2i\pi} \text{ p.v. } \int_{\mathscr{L}} q^*[z] e^{zx} dz \tag{3}$$

where $\mathscr{L}$ is a vertical line in the complex plane such that $\mathbb{R}(\mathscr{L}) > -\sigma$. To calculate $q^*[z]$ we take the Laplace transform of (1) with respect to variable x and use Fubini's theorem. We obtain for $\mathbb{R}(z) > -\sigma$

$$q^*[z] = \int_0^\infty \left(\int_{ct}^\infty \varphi(t, w) e^{-zw} dw \right) e^{zct} dt. \tag{4}$$

As $\Lambda_s \leqq \Sigma_{i \leqq N} h_i$ for all s, W_t is bounded for fixed t and the Laplace transform $\varphi^*[t, \cdot]$ is defined over $\mathbb{C}$. Splitting the integral yields

$$\forall z \in \mathbb{C}, \quad \int_{ct}^\infty \varphi(t, w) e^{-zw} dw = \varphi^*[t, z] - \int_0^{ct} \varphi(t, w) e^{-zw} dw. \tag{5}$$

The substitution of (5) in (4) gives $q^*[z] = \varphi^{**}[-cz, z] - h(z)$ for $\mathbb{R}(z) > -\sigma$, where

$$h(z) = \int_0^\infty \left(\int_0^{ct} \varphi(t, w) e^{z(ct-w)} dw \right) dt.$$

Proposition 2. *$h(z)$ is an analytic function defined on the half plane $\mathbb{R}(z) < 0$ and bounded over $\mathbb{R}(z) < -\eta$ for all $\eta > 0$.*

The proof is given in Section 5. By Jordan's lemma (Churchill (1972)), we then have $(1/2i\pi) \int_{\mathscr{L}} h(z) e^{zx} dz = 0$, for any vertical line $\mathscr{L}$ such that $\mathbb{R}(\mathscr{L}) < 0$. The latter relation used in (3) yields formula (2).

3. Cox-type distributions

In this section, we assume silence durations (off periods) and burst volumes (on periods) have Cox-type distributions, i.e. with rational Laplace transforms. Let k_{ai} and k_{bi} be the degrees of the denominators of these transforms. A *double vector* $\bar{\boldsymbol{u}}$ is an N-dimensional vector $(\bar{u}_i)_{i \leqq N}$ where each $\bar{u}_i$ is a $(k_{ai} + k_{bi})$-dimensional vector. By convention, $\bar{\boldsymbol{u}} \cdot \bar{\boldsymbol{v}} = \Sigma_{i \leqq N} \bar{u}_i \cdot \bar{v}_i$ and $\bar{\boldsymbol{u}}^{\bar{\boldsymbol{v}}} = \Pi_{i,j} u_{i,j}^{v_{i,j}}$. Denote by $\bar{\boldsymbol{\theta}}_j$ any double vector such that each $\bar{\theta}_{i,j}$ has one non-zero coordinate equal to 1 and let $\mathscr{U}(\boldsymbol{\varepsilon})$ be the set of double vectors $\bar{\boldsymbol{u}}$ such that each $\bar{u}_i$ has one non-zero coordinate equal to 1 if $\varepsilon_i = 1$ and is a null vector otherwise. We can then obtain the following result.

Theorem 2. *When on and off periods have Cox-type distributions, the Beneš upper bound can be expressed as the finite sum*

$$q(x) = \sum_{j \in \mathscr{J}} k_j e^{z_j \cdot x} \tag{6}$$

where $(z_j)_{j \in \mathscr{J}}$ is the finite set of solutions with negative real part of the algebraic equations

$$\bar{\boldsymbol{\theta}}_j \cdot \bar{\boldsymbol{s}}(z) + cz = 0 \tag{7}$$

with $\bar{\boldsymbol{s}}(z) = (\bar{s}_i(z))_{i \leqq N}$ and where $\bar{s}_i(z)$ is the vector of solutions of $a_i^*[s] b_i^*[z + s/h_i] = 1$. The $(k_j)_{j \in \mathscr{J}}$ are complex coefficients defined by

$$k_j = \frac{-1}{[\bar{\boldsymbol{\theta}}_j \cdot \bar{\boldsymbol{s}}(z_j) + c]} \cdot \sum_{\boldsymbol{\varepsilon} : \boldsymbol{\varepsilon} \cdot \boldsymbol{h} < c; \bar{\boldsymbol{u}} \in \mathscr{U}(\boldsymbol{\varepsilon}) : \bar{\boldsymbol{u}} \leqq \bar{\boldsymbol{\theta}}_j} (c - \boldsymbol{\varepsilon} \cdot \boldsymbol{h}) \bar{\boldsymbol{r}}_B(z_j)^{\bar{\boldsymbol{u}}} \bar{\boldsymbol{r}}_4(z_j)^{\bar{\boldsymbol{\theta}}_j - \bar{\boldsymbol{u}}}$$

where $\bar{r}_{Ai}(z)$, $\bar{r}_{Bi}(z)$, $i \leqq N$ are the residues of functions $p_{Ai}^{**}[s, z]$, $p_{Bi}^{**}[s, z]$ (defined in Section 5) at $s = \bar{s}_i(z)$, respectively.

Cox-type distributions are within the scope of application of Theorem 1. Applying the theorem of residues to contour integral (2) then enables us to derive the above formulae. Details of the proof can be found in Guibert (1994). Sorting the roots $(z_j)_{j \in \mathscr{J}}$ by decreasing real part, we readily derive from (6) the asymptotic form $q(x) \sim k_0 \exp(z_0 x)$ for large x.

4. Conclusion

We derived an expression for the Beneš upper bound as a general contour integral formula. When on and off periods have Cox-type distributions, it readily leads to a finite spectral expansion whose coefficients are straightforward to obtain via the theorem of residues. The practical accuracy of this expression has been validated in Guibert (1994). We also believe that our formula can be applied to other types of distributions for on and off periods.

5. Annex

Denote by $p_{Ai}(t,\cdot)$ (or $p_{Bi}(t,\cdot)$) the density of the incoming workload over time interval $(0, t]$ due to source i, given it is 'off' (or 'on') at epoch 0. These densities are Dirac distributions at epochs 0 and $h_i t$. Let $\varphi_{\boldsymbol{\varepsilon}}(t, w) = \boldsymbol{P}(\Lambda_t = \boldsymbol{\varepsilon}\cdot\boldsymbol{h};\ w < W_t \leqq w + dw)/dw$. For the superposition of independent sources, we can write

$$\varphi_{\boldsymbol{\varepsilon}}^*[t,z] = \prod_{i:\varepsilon_i=1} v_i p_{Bi}^*[t,z] \prod_{i:\varepsilon_i=0} (1-v_i) p_{Ai}^*[t,z] \tag{8}$$

where $v_i = \alpha_i/(\alpha_i + h_i\beta_i)$. It is known from [7] that, for general probability densities of on and off periods, we can write

$$\begin{cases} p_{Ai}^{**}[s,z] = \dfrac{1}{s} + \dfrac{-\alpha_i z}{s^2(z+s/h_i)} \dfrac{(1-a_i^*[s])(1-b_i^*[z+s/h_i])}{1-a^*[s]b^*[z+s/h_i]} \\ p_{Bi}^{**}[s,z] = \dfrac{1}{s+h_i z} + \dfrac{h_i\beta_i z}{s(z+s/h_i)^2} \dfrac{(1-a_i^*[s])(1-b_i^*[z+s/h_i])}{1-a_i^*[s]b_i^*[z+s/h_i]} \end{cases} . \tag{9}$$

Now, as detailed below, the following lemma is essential for the justification of Propositions 1 and 2.

Lemma 1. *Given the stability condition $\rho < 1$, there exists $\sigma > 0$ and $c' < c$ such that $\varphi^*[t, -\sigma] = O(e^{c'\sigma t})$ for large t.*

Proof. Since $\varphi^*[t,z]$ is a finite positive combination of $\varphi_{\boldsymbol{\varepsilon}}^*[t,z]$ it suffices to prove this assertion for each $\varphi_{\boldsymbol{\varepsilon}}^*[t,z]$. The behaviour of $\varphi_{\boldsymbol{\varepsilon}}^*[t,z]$ for large t is related to the rightmost singularities of its Laplace transform with respect to variable t. Formulae (8) and (9) show that the singularities in s of $\varphi_{\boldsymbol{\varepsilon}}^{**}[s,z]$ come from the factors $1 - a_i^*[s]b_i^*[z+s/h_i]$, $a_i^*[s]$ and $b_i^*[z+s/h_i]$. Assume $z = -\sigma < 0$ is small. The root of equation $1 - a_i^* s b_i^*[-\sigma + s/h_i] = 0$ with largest real part is asymptotic to $v_i h_i \sigma$, hence bounded from above by $(v_i h_i + \eta)\sigma$ for some $\eta > 0$. As a_i^* and b_i^* have a negative convergence abscissa, the singularities induced by $a_i^*[s]$ and $b_i^*[z+s/h_i]$ are smaller than $v_i h_i \sigma$ for small σ. It can then be derived from formula (8) that $\varphi_{\boldsymbol{\varepsilon}}^*[t, -\sigma] = O(\exp[\Sigma_{i\leqq N}(v_i h_i + \eta)\sigma t])$ uniformly in large t. Given the stability assumption, it is now sufficient to choose η and c' such that $\Sigma_{i\leqq N}(v_i h_i + \eta) < c' < c$.

Proof of Proposition 1. $\varphi(t, w)$ has Dirac components for $w = (\boldsymbol{\varepsilon}\cdot\boldsymbol{h})t$ only. Since on and off periods are piecewise smooth, this first implies that $\varphi(t, w)$ is piecewise smooth in w for $w \geqq ct$ so that $q(x)$ is piecewise smooth by formula (1).

Now, for σ defined as in Lemma 1, we have

$$\int_{ct}^{\infty} \varphi(t, w) e^{\sigma(w-ct)} dw = e^{-\sigma ct} \int_{ct}^{\infty} \varphi(t, w) e^{\sigma w} dw$$
$$\leqq e^{-\sigma ct} \varphi^*[t, -\sigma]$$

which is $O(e^{-(c-c')\sigma t})$. We thus conclude that $q^*[z]$ expressed as in (4) exists for $\mathbb{R}(z) > -\sigma$.

Proof of Proposition 2. Using the Chernoff bound (Kleinrock (1975)) $\boldsymbol{P}(X \geqq x) \leqq e^{\ln(\boldsymbol{E}[e^{uX}]) - ux}$ for all $u \geqq 0$, we obtain for arbitrary $\gamma > 0$

$$\forall u \geqq 0, \qquad \int_{\gamma t}^{\infty} \varphi_{\boldsymbol{\varepsilon}}(t, w) dw = \boldsymbol{P}(\Lambda_t = \boldsymbol{\varepsilon} \cdot \boldsymbol{h};\, W_t > \gamma t) \leqq e^{\ln \varphi_{\boldsymbol{\varepsilon}}^*(t, -u] - u\gamma t}.$$

Choosing $c' < \gamma < c$ as in Lemma 1, we then deduce

$$\int_0^{\infty} \left(\int_{\gamma t}^{+\infty} \varphi(t, w) dw \right) dt < +\infty. \tag{10}$$

We can now write

$$\int_0^{ct} |\varphi(t, w) e^{z(ct-w)}| \, dw = \int_0^{ct} \varphi(t, w) e^{\mathbb{R}(z)(ct-w)} dw$$
$$= \int_0^{\gamma t} \varphi(t, w) e^{\mathbb{R}(z)(ct-w)} dw + \int_{\gamma t}^{ct} \varphi(t, w) e^{\mathbb{R}(z)(ct-w)} dw$$
$$\leqq e^{\mathbb{R}(z)(c-\gamma)t} \int_0^{\gamma t} \varphi(t, w) dw + \int_{\gamma t}^{\infty} \varphi(t, w) dw$$

with $\mathbb{R}(z) \leqq 0$. Now, $\int_0^{\gamma t} \varphi(t, w) dw$ is $O(1)$ uniformly in t and in view of (10), we conclude that $h(z)$ is well defined for $\mathbb{R}(z) < 0$ and bounded for $\mathbb{R}(z) < -\eta$ for all $\eta > 0$.

Similar arguments show that $\int_0^{\infty} (\int_0^{ct} \varphi(t, w)(ct - w) e^{\mathbb{R}(z)(ct-w)} dw) dt < +\infty$, hence the function $h(z)$ is analytic for $\mathbb{R}(z) < 0$.

References

Anick, D., Mitra, D. and Sondhi, M. M. (1982) Stochastic theory of a data handling system with multiple sources. *Bell Syst. Tech. J.* **61**, 1871–1894.

Bensaou, B., Guibert, J. and Roberts, J. (1990) Fluid queuing models for a superposition of on/off sources. ITC Broadband Seminar, Morristown.

Bensaou, B., Guibert, J., Roberts, J. and Simonian, A. (1993) Performance of an ATM multiplexer queue in the fluid approximation using the Beneš approach. *Ann. Operat. Res.* To appear.

Churchill, R. V. (1972) *Operational Mathematics*. McGraw-Hill, New York.

Guibert, J. (1994) Overflow probability upper bound for heterogeneous fluid queues handling general on-off sources. *ITC 14, Antibes*.

Kleinrock, L. (1975) *Queueing Systems Vol. I, Theory*. Wiley, New York

Kosten, L. (1986) Liquid models for a type of information buffer problem. *Delft Progr. Rep.* **11**, 71–86.

NORROS, L., ROBERTS, J., SIMONIAN, A. AND VIRTAMO, T. (1991) The superposition of variable bit rate sources in an ATM multiplexer. *IEEE JSAC* **9**, 378–387.

SIMONIAN, A. AND VIRTAMO, J. (1991) Transient and stationary distributions for fluid queues and input processes with a density. *SIAM J. Appl. Math.* **51**, 1732–1739.

STERN, T. E. AND ELWALID, A. I. (1991) Analysis of separable Markov-modulated rate models for information-handling systems. *Adv. Appl. Prob.* **23**, 105–139.

Yours sincerely,

JACKY GUIBERT
CNET/France Telecom
38–40 rue du Général Leclerc
92 131 Issy-les-Moulineaux Cedex
France