Performance analysis of a heterogeneous fluid FIFO multiplexer

Ernst Nordström

October 30, 2009

1 Introduction

The original fluid flow model for the FIFO queue with constant service rate was proposed by Kosten [2] and further developed by Anick, Mitra and Sondhi [1]. The tradional assumption in fluid flow models is exponentially distributed acivity periods. Kosten and AMS presented formulas for the fluid overflow probability of the infinite FIFO queue. Tucker [4] and Jacobsen and Dittman [5] derived a formula for the fluid loss probability of the finite homegeneous and heterogeneous FIFO queue, respectively. Tucker also presents a formula for the delay distribution for a homogeneous FIFO queue. This note derives the mean delay in a heterogeneous finite FIFO queue. Tne fluid flow literature does not seem to contain this result.

2 Traffic assumption

We consider a FIFO system that is offered traffic from K classes. The capacity of the system is denoted C [Mbps]. The buffer capacity is denoted B [Mbit]. The j-th class, $j \in J = \{1, \ldots, K\}$, is characterized by the following:

- Number of calls: N_j,
- Peak bit rate requirement: p_j [Mbps],
- OFF to ON state transition rate: α_j [s^{-1}],
- ON to OFF state transition rate: β_j [s^{-1}],

3 Model

The buffer is modeled as a fluid reservoir with a hole in the bottom and arriving information is modeled as a fluid running into the reservoir. Let $j \in J$ denote the index of an arbitrary class. Let the stochastic variables Σ and Q denote the stationary state of the system fluid process, and the queue length, respectively. Let $\mathbf{k} = (k_j)_{j \in J}$ denote the fluid state vector, where $k_j = k_j(t)$ denotes the number of sources in their ON state at time t. Vectors and matrices will appear in bold letter such that they can be distinguished from numbers, functions etc. Let $\mathbf{p} = (p_1, ..., p_K)$ denotes the peak rate vector for the sources and by the scalar product of $\mathbf{k}$ and $\mathbf{p}$ we mean:

$$a(\mathbf{k}) = \mathbf{k} \cdot \mathbf{p} = \sum_{i=1}^{K} k_i p_i \tag{1}$$

which is the total input rate in state $\mathbf{k}$.

The fluid state space is defined by the set:

$$\mathbf{S} = \{\mathbf{k} = (k_j)_{j \in J} : 0 \leq k_j \leq N_j, j \in J\} \tag{2}$$

The average load on the system is

$$\rho = \frac{\sum_{i=1}^{K} N_i p_i \frac{\alpha_i}{\alpha_i + \beta_i}}{C} \tag{3}$$

We assume the average input rate to be smaller than the output capacity, i.e. $\rho < 1$.

The sample path description of the buffer dynamics is:

$$dQ/dt = \begin{cases} a(t) - C & \text{if} \quad Q(t) > 0 \\ (a(t) - C))^{+} & \text{if} \quad Q(t) = 0 \\ (a(t) - C)^{-} & \text{if} \quad Q(t) = B \end{cases} \tag{4}$$

where generally $(x)^+ = \max(x, 0)$ and $(x)^- = \min(x, 0)$ denotes the positive and negative part of x, respectively, and $a(t)$ is the total arrival rate at time t.

Let $\mathbf{F}(x) = \{F_{\mathbf{k}}(x)\}_{\mathbf{k} \in \mathbf{S}}$ denote the stationary buffer distribution column vector, where $F_{\mathbf{k}}(x) = \Pr(\Sigma = \mathbf{k}, Q \leq x), \mathbf{k} \in \mathbf{S}, 0 \leq x \leq B$. The system can be described by the following Kolmogorov differential equation [1]:

$$\mathbf{D} \frac{\mathrm{d}}{\mathrm{dx}} \mathbf{F}(q) = \mathbf{M} \mathbf{F}(x), 0 < x < B, \tag{5}$$

where $\mathbf{D}$ is a diagonal matrix with entry $(\mathbf{k}, \mathbf{k})$ equal to

$$d_{\mathbf{k}} = (\sum_{i=1}^{K} p_i k_i - C) \tag{6}$$

and where entry $(\mathbf{k}, \mathbf{n})$ in $\mathbf{M}$ looks as follows:

$$\begin{aligned} m_{(\mathbf{k},\mathbf{k})} &= -\sum_{i=1}^{K} (N_i - k_i)\alpha_i + k_i \beta_i, \quad \text{for } \mathbf{k} \in S \\ m_{(\mathbf{k}, k_1, \ldots, k_i - 1, \ldots, k_K)} &= (N_i - \beta_i + 1)\alpha_i, \quad \text{for } \mathbf{k} \in S \\ m_{(\mathbf{k}, k_1, \ldots, k_i + 1, \ldots, k_K)} &= (k_i + 1)\beta_i, \quad \text{for } \mathbf{k} \in S \end{aligned} \tag{7}$$

4 Solution

The solution $\mathbf{F}(x) = (\mathbf{F}_{\mathbf{k}}(x))_{\mathbf{k} \in \mathbf{S}}$ is obtained from a spectral expansion:

$$\mathbf{F}_{\mathbf{k}}(x) = \sum_{\mathbf{n} \in \mathbf{S}} a_{\mathbf{n}} \exp(z(\mathbf{n})x)(\varphi_{\mathbf{n}})_{\mathbf{k}} \tag{8}$$

where $\{z(\mathbf{n})\}$ are the solution to the generalized eigenvalue problem $z\mathbf{D}\varphi = \mathbf{M}\varphi$. In practice, the eigenvalue for state $\mathbf{k}$ is found by solving a non-linear algebraic equation. The eigenvector $\varphi_{\mathbf{n}}$ corresponding to the eigenvalue $z(\mathbf{n})$ is given by the coefficients of certain polynomial in K variables.

The coefficients $a_{\mathbf{k}}$ are found by means of the initial condition [5]:

$$\begin{aligned} F_{\mathbf{k}}(0) = 0, \quad a(\mathbf{k}) > C \\ 0 = u_{\mathbf{k}} = \pi(\mathbf{k}) - \lim_{x \to B} F_{\mathbf{k}}(x), \quad a(\mathbf{k}) < C \end{aligned} \tag{9}$$

where $\pi(\mathbf{k})$ denotes the overall probability of sources being in state $\mathbf{k}$. This probability is found from the multi-binomial distribution:

$$\pi(\mathbf{k}) = \prod_{j \in J} \binom{N_j}{k_j} \left(\frac{\alpha_j}{\alpha_j + \beta_j} \right)^{k_j} \left(\frac{\beta_j}{\alpha_j + \beta_j} \right)^{x_j - k_j} \tag{10}$$

5 Performance measures

In this subsection we present formulas for the buffer overflow probability, fluid loss probability, and mean waiting time in the queue.

The buffer overflow probability $G(x) = \Pr(Q > x)$ can be written as

$$G(x) = 1 - \sum_{\mathbf{k} \in S} F_{\mathbf{k}}(x) \tag{11}$$

The overall fluid loss probability p_{loss} is defined as the fraction lost information to offered information. Loss can only take place when the buffer is at maximum and the input rate is larger than the output rate. Therefore we get:

$$p_{loss} = \frac{\sum_{\{\mathbf{k} | \mathbf{k} \cdot \mathbf{p} > C\}} (\mathbf{k} \cdot \mathbf{p} - C) u_{\mathbf{k}}}{\sum_{i=1}^{K} N_i p_i \frac{\alpha_i}{\alpha_i + \beta_i}} \tag{12}$$

The buffer overflow probability is an upper bound of the fluid loss probability. The fluid loss probability $p_{loss}(j)$ for sources in class j is the fraction between lost class j information to the offered class j information, and is therefore given as

$$p_{loss}(j) = \frac{\sum_{\{\mathbf{k} | \mathbf{k} \cdot \mathbf{p} > C\}} \frac{k_j p_j}{\mathbf{k} \cdot \mathbf{p}} (\mathbf{k} \cdot \mathbf{p} - C) u_{\mathbf{k}}}{\sum_{i=1}^{K} N_i p_i \frac{\alpha_i}{\alpha_i + \beta_i}} \tag{13}$$

The distribution of queue length is given by

$$\Pr(Q \le x) = \sum_{\mathbf{k} \in \mathbf{S}} F_{\mathbf{k}}(x) \tag{14}$$

The mean queue length is given by

$$\overline{Q} = \int_0^B x \frac{\mathrm{dPr}(Q \le x)}{\mathrm{d}x} \mathrm{d}x = \sum_{\mathbf{k} \in \mathbf{S}} \left\{ [xF_{\mathbf{k}}(x)]_0^{B-} - \int_0^{B-} F_{\mathbf{k}}(x)\mathrm{d}x + Bu_{\mathbf{k}} \right\} \tag{15}$$

where $F_{\mathbf{k}}(B-)$ is defind as $\lim_{x \to B} F_{\mathbf{k}}(x)$. The mean queueing delay for class j, $\overline{W}_j$, is obtained from Little's formula:

$$\overline{W}_j = \frac{\overline{Q}}{N_j m_j (1 - p_{loss}(j))} \tag{16}$$

where m_j denotes the mean arrival rate of a source from class j:

$$m_j = p_j \frac{\alpha_j}{\alpha_j + \beta_j} \tag{17}$$

6 Eigenvalues and eigenvectors

The eigenvalue for state $\mathbf{k}$ can be found by solving the algebraic equation $f(z(\mathbf{k})) = g(z(\mathbf{k}))$ where

$$f(z(\mathbf{k})) = z(\mathbf{k})(C - \sum_{i=1}^{K} \frac{N_i}{2} p_i) - \sum_{i=1}^{K} \frac{N_i}{2} (\alpha_i + \beta_i) \tag{18}$$

$$g(z(\mathbf{k})) = \sum_{i=1}^{K} (k_i - \frac{N_i}{2}) \sqrt{(z(\mathbf{k})p_i + \beta_i - \alpha_i)^2 + 4\alpha_i\beta_i} \tag{19}$$

The eigenvector $\varphi_{\mathbf{k}}$ corresponding to the eigenvalue $z(\mathbf{k})$ is given as the coefficients in the following polynomial in K variables:

$$p_{\mathbf{k}} = \prod_{i=1}^{K} (N_i - r_i(z))_i^k (N_i - s_i(z))^{N_i - k_i} \tag{20}$$

where

$$r_i(z) = \frac{-(zp_i + \beta_i - \alpha_i) + \sqrt{(zp_i + \beta_i - \alpha_i)^2 + 4\alpha_i\beta_i}}{2\alpha_i} \tag{21}$$

$$s_i(z) = \frac{-(zp_i + \beta_i - \alpha_i) - \sqrt{(zp_i + \beta_i - \alpha_i)^2 + 4\alpha_i\beta_i}}{2\alpha_i} \tag{22}$$

7 Discussion

The size of the state space for the fluid flow model is $N_s = |\mathbf{S}| = \prod_{i=1}^{K} (N_i + 1)$. As far as computational complexity is concerned, is is dominated by the calculation of the coefficients $a_{\mathbf{k}}$ from the boundary conditions. The coefficients are found by solving a set of N_s linear equations. This system of equations can be solved by Gauss elimination or, which can be more efficient, by some iterative method. In any case, the associated complexity is in the order of $O(N_s^3)$. Hence, the complexity of the fluid flow model increases very fast with increasing number of classes K and increasing class sizes N_i. Therefore, the fluid flow model is only considered to be useful as a reference model, and not as a basis for implementation of call admission control in real multi-service networks.

References

[1] D. Anick, D. Mitra, M. Sondhi, "Stochastic theory of a data handling system with multiple sources", *Bell System Technical Journal*, Vol. 61, No. 8, pp. 1871-1894, 1982.

[2] L. Kosten, "Stochastic theory of a multi-entry buffer", *Delft Progress Report Series F*, Vol. 1, pp. 10-18, 1974.

[3] K. Kvols, S. Blaabjerg, "Bounds and approximations for the periodic on/off queue with application to ATM traffic control", In *Proc. IEEE Infocom*, pp. 487-494, 1992.

[4] R.C.F Tucker, "Accurate method for analysis ofa packet- speech multiplexer with limited delay", *IEEE Transactions on Communications*, Vol. 36, No. 4, 1998.

[5] S. Jacobsen, L. Dittman, "A fluid flow queueing model for heterogeneous on/off traffic", RACE BLNT Workshop, Munchen, 1990.