On the Asymptotic Behavior of a Fluid Queue With a Heavy-Tailed M/G/∞ Arrival Process

Predrag R. Jelenković
Department of Electrical Engineering
Columbia University, New York, NY 10027
email: predrag@ee.columbia.edu

June 10, 2000

Abstract

We consider an infinite buffer fluid queue with a constant capacity and an M/G/∞ arrival process. M/G/∞ process consists of sessions with Poisson starting times, independent intermediately regularly varying session durations and constant arrival rates. The session duration distributions and arrival rates are selected from a finite set of distinct choices. For this queue we derive the asymptotic behavior of the stationary queue length distribution under the assumption that exactly one long session is enough to cause a large queue build-up.

Increased utilization in communication networks (e.g. the Internet) can be achieved through sharing of bandwidth and buffer resources among different user sessions. This sharing may result in increased congestion and reduced quality of service. Therefore, it is important to have efficient computational methods for predicting the congestion. A fluid queue with M/G/∞ arrivals is often considered as a baseline model of congestion. In this queueing system sessions arrive at Poisson times and have independent duration during which they produce fluid at a constant rate. Our investigation focuses on the asymptotic behavior of the stationary queue length distribution of this system when session lengths are heavy-tailed. The heavy-tailed nature of Internet traffic has been observed in numerous statistical experiments. Further motivation and additional references on this problem can be found in [1, 4, 6, 9, 2, 8]. Our main result is presented in Theorem 1.

More formally, consider a Poisson process on the positive real axes with rate λ and jump points $\{T_n, n \geq 1\}$ representing session initiation times. Let $\{\tau, \tau_n, n \geq 1\}$ be a sequence of i.i.d. session lengths independent of $\{T_n\}$. Then, for any $r > 0$

$$A^0(t) \stackrel{\text{def}}{=} r \sum_{n=1}^{\infty} 1(T_n \leq t < T_n + \tau_n) \tag{1}$$

represents an M/G/∞ process with starting point at 0 and peak rate r; the distribution of this process is completely specified with (λ, r, τ). Next, assume that $\mathbf{E}\tau < \infty$ and define τ^e to be the excess time of τ with its distribution equal to $\mathbf{P}[\tau^e \leq x] = \int_0^x \mathbf{P}[\tau > u]du/\mathbf{E}\tau$. Let $\{\tau_n^e, n \geq 1\}$ be independent copies of τ^e and N a Poisson random variable with mean $\mathbf{E}N = \lambda \mathbf{E}\tau$. Assume that all random variables are mutually independent. Now, if we define

$$A^e(t) \stackrel{\text{def}}{=} r \sum_{n=1}^{N} 1(0 \leq t < \tau_n^e), \tag{2}$$

then the stationary M/G/∞ process $A(t)$ has the following representation

$$A(t) = A^e(t) + A^0(t);$$

note that $\rho \stackrel{\text{def}}{=} \mathbf{E}A(t) = r\lambda\mathbf{E}\tau$. Next, let $A_i(t), 1 \le i \le L$, be a collection of independent M/G/∞ processes with parameters (λ_i, r_i, τ^i) and, with a small violation of notation, define an M/G/∞ process with L session classes as

$$A(t) \equiv A(t, L) = \sum_{i=1}^{L} A_i(t);$$

here $\rho = \mathbf{E}A(t) = \sum_{i=1}^{L} \rho_i, \rho_i = \mathbf{E}A_i(t)$.

Now, consider an infinite buffer fluid queue with arrival rate $A(t)$ and constant capacity $c, c > \rho$. Then, since $A(t)$ is stationary and reversible, standard queueing arguments show that the stationary queue length is equal in distribution to

$$Q \stackrel{\text{d}}{=} M \stackrel{\text{def}}{=} \sup_{t \ge 0}(A(t) - ct).$$

In order to state our main result we need to introduce the class of intermediately regularly varying distributions. A nonincreasing positive function $f(t)$ is said to be intermediately regularly varying ($f \in \mathcal{IR}$) if

$$\lim_{\delta \uparrow 1} \limsup_{t \to \infty} \frac{f(\delta t)}{f(t)} = 1.$$

For a random variable τ we say that it is intermediately regularly varying ($\tau \in \mathcal{IR}$) if $\mathbf{P}[\tau > t] \in \mathcal{IR}$. It is well known that $\tau \in \mathcal{IR}, \mathbf{E}\tau < \infty$, implies $\tau^e \in \mathcal{IR}$. Pareto distributions are well known examples from $\mathcal{IR}$. Also, for $\tau^i \in \mathcal{IR}, 1 \le i \le L$ there exist $\alpha \ge 0$ and a finite constant C such that for all $x > 0$

$$\max_{1 \le i \le L} \mathbf{P}[\tau^i > x] \le \frac{C}{x^\alpha};$$

see equation (1.6) of [7]. In this paper we assume that $\alpha > 1$. This technical condition is necessary because the proof of our main theorem uses Theorem 1 of [7] which requires it.

Throughout the paper we use the customary notation $f(x) \sim g(x)$ as $x \to \infty$ to denote $\lim_{x \to \infty} f(x)/g(x) = 1$.

Theorem 1 *If $\rho < c < r_i + \rho, \tau^i \in \mathcal{IR}, 1 \le i \le L, \alpha > 1$, then*

$$\mathbf{P}[Q > x] \sim \sum_{i=1}^{L} \frac{\rho_i}{c - \rho} \mathbf{P}\left[\tau^{i,e} > \frac{x}{r_i + \rho - c}\right] \quad \textit{as} \quad x \to \infty.$$

Remarks: For a single class case ($L = 1$) this theorem was first conjectured in [4], where a tight lower bound was proved. An informal sketch of a proof of this conjecture was presented in [3]. The result was proved rigorously (for $L = 1$) in [8] under the assumption of τ being regularly varying. Here, we provide a generalization and an alternative proof to Theorem 1 from [8]. This theorem suggests that related results for the discrete time M/G/∞ model, obtained in [5] for Pareto distributions, should hold for a larger intermediately regularly varying class.

The proof of Theorem 1 will be based on the following two auxiliary lemmas. Let

$$M^0(c) \stackrel{\text{def}}{=} \sup_{t \ge 0}(A^0(t) - ct) \tag{3}$$

and

$$M^e(c) \stackrel{\mathrm{def}}{=} \sup_{t\geq 0}(A^e(t) - ct), \tag{4}$$

where A^0 and A^e are defined in (1) and (2).

Lemma 1 *If $r = 1, \rho < c < 1$ and $\tau \in \mathcal{IR}, \alpha > 1$, then*

$$\mathbf{P}[M^0(c) > x] \sim \rho \frac{1+\rho-c}{c-\rho} \mathbf{P}\left[\tau^e > \frac{x}{1+\rho-c}\right] \quad as \quad x \to \infty.$$

Proof: Let $U_0 = 0$ and $U_n, n \geq 1$, be the ends of activity periods in $A^0(t)$. Then, since $c < 1$, it is clear that

$$M^0(c) = \sup_{n\geq 0}(A^0(U_n) - cU_n).$$

Hence, this result is the same as Theorem 4.6 of [4], which was proved for τ regularly varying with non-integer exponents. This more general result is immediate consequence of the proof of Theorem 4.6 of [4] and Theorem 1 of [7]. $\diamond$

Lemma 2 *If $r = 1, 0 \leq c < 1$ and $\tau^e \in \mathcal{IR}$, then*

$$\mathbf{P}[M^e(c) > x] \sim \rho \mathbf{P}\left[\tau^e > \frac{x}{1-c}\right] \quad as \quad x \to \infty.$$

Proof: The proof follows easily from the established theory of heavy-tailed distributions. For completeness a detailed proof is given in the appendix. $\diamond$

Proof of Theorem 1: We start with the case $L = 1$. The lower bound was proved in Theorem 4.5 of [4]. For the upper bound, since $\sup(A(t) - c) = r_1 \sup(A(t)/r_1 - c/r_1)$, it is enough to prove the result for $r_1 = 1$. For simplicity we suppress the index 1 in $(\rho_1, \tau^{1,e})$. Assume first that $c < 1$. Then, for $c - \rho - \epsilon \geq 0$, by elementary algebra

$$\begin{aligned} M &= \sup_{t\geq 0}(A^0(t) - (\rho+\epsilon)t + (A^e(t) - (c-\rho-\epsilon)t)) \\ &\leq \sup_{t\geq 0}(A^0(t) - (\rho+\epsilon)t) + \sup_{t\geq 0}(A^e(t) - (c-\rho-\epsilon)t)) \\ &= M^0(\rho+\epsilon) + M^e(c-\rho-\epsilon). \end{aligned} \tag{5}$$

Similarly,

$$M \leq M^0(c) + M^e(0),$$

which together with (5) yields

$$\begin{aligned} M &\leq \min(M^0(\rho+\epsilon) + M^e(c-\rho-\epsilon), M^0(c) + M^e(0)) \\ &\leq \max(M^0(c), M^e(c-\rho-\epsilon)) + \min(M^0(\rho+\epsilon), M^e(0)). \end{aligned}$$

The last inequality and independence of $M^0(\rho+\epsilon)$, $M^e(0)$ implies

$$\begin{aligned} \mathbf{P}[M > x] \leq \mathbf{P}[M^0(c) > \delta x] &+ \mathbf{P}[M^e(c-\rho-\epsilon) > \delta x] \\ &+ \mathbf{P}[M^0(\rho+\epsilon) > (1-\delta)x]\mathbf{P}[M^e(0) > (1-\delta)x]. \end{aligned} \tag{6}$$

Observe that the last term in the preceding inequality is equal to $o(\mathbf{P}[\tau^e > x])$ as $x \to \infty$ by Lemmas 1 and 2. Now, by using this observation and Lemmas 1 and 2 in (6) we derive

$$\begin{aligned}\limsup_{x\to\infty} \frac{\mathbf{P}[M > x]}{\mathbf{P}[\tau^e > x/(1+\rho-c)]} &\le \rho \limsup_{x\to\infty} \frac{\mathbf{P}[\tau^e > \delta x/(1+\rho+\epsilon-c)]}{\mathbf{P}[\tau^e > x/(1+\rho-c)]} \\ &\quad + \rho\frac{1+\rho-c}{c-\rho} \limsup_{x\to\infty} \frac{\mathbf{P}[\tau^e > \delta x/(1+\rho-c)]}{\mathbf{P}[\tau^e > x/(1+\rho-c)]}\end{aligned}$$

which, since $\tau^e \in \mathcal{IR}$, by passing $\delta \uparrow 1$ and $\epsilon \downarrow 0$ yields the upper bound for $c < 1$.

When $c \ge 1$, for any $h > 0$ we decompose the arrival process $A(t)$ into two stationary independent M/G/∞ processes $A_h(t)$ and $a_h(t)$ containing all the sessions greater and smaller/equal than h, respectively. Note that $\rho_h \overset{\text{def}}{=} \mathbf{E}A_h(t) = \lambda \mathbf{E}\tau 1[\tau > h]$ and $\mathbf{E}a_h(t) = \lambda \mathbf{E}\tau 1[\tau \le h] < \rho$ (for more details see the proof of Theorem 4.5 in [4]). Now, for all h large enough such that $\rho_h < c - \rho$

$$\begin{aligned}M &\le \sup_{t\ge 0}(A_h - (c-\rho)t) + \sup_{t\ge 0}(a_h - \rho t) \\ &\overset{\text{def}}{=} M_h + m_h,\end{aligned}$$

with both m_h and M_h being almost surely finite. Hence, for any $0 < \delta < 1$

$$\mathbf{P}[M > x] \le \mathbf{P}[M_h > \delta x] + \mathbf{P}[m_h > (1-\delta)x].$$

Next, using standard Chernoff type bounds, it is easy to show that m_h is exponentially bounded. Therefore, $\mathbf{P}[m_h > x] = o(\mathbf{P}[\tau^e > x])$ and

$$\limsup_{x\to\infty} \frac{\mathbf{P}[M > x]}{\mathbf{P}[\tau^e > x/(1+\rho-c)]} \le \frac{\rho_h}{c-\rho-\rho_h} \frac{\mathbf{E}\tau}{\mathbf{E}\tau 1[\tau > h]} \limsup_{x\to\infty} \frac{\mathbf{P}[\tau^e > \delta x/(1+\rho_h+\rho-c)]}{\mathbf{P}[\tau^e > x/(1+\rho-c)]}.$$

Thus, by passing $h \to \infty$ and $\delta \uparrow 1$ we obtain the desired upper bound and conclude the proof of the case $L = 1$. Now, we proceed to prove the case $L = 2$.

($L = 2$, *Upper bound.*) For any $\epsilon > 0$, such that $\rho + \epsilon < c$

$$\begin{aligned}M &\le \sup_{t\ge 0}(A_i(t) - (c - \rho_{3-i} - \epsilon)t) + \sup_{t\ge 0}(A_{3-i}(t) - (\rho_{3-i} + \epsilon)t) \\ &= M_i + \underline{M}_{3-i},\end{aligned}$$

where $M_i \overset{\text{def}}{=} \sup_{t\ge 0}(A_i(t) - (c - \rho_{3-i} - \epsilon)t)$ and $\underline{M}_i \overset{\text{def}}{=} \sup_{t\ge 0}(A_i(t) - (\rho_i + \epsilon)t)$ are both almost surely finite random variables. Thus,

$$M \le \min(M_1 + \underline{M}_2, M_2 + \underline{M}_1) \le \max(M_1, M_2) + \min(\underline{M}_1, \underline{M}_2),$$

which, for any $0 < \delta < 1$, implies

$$\mathbf{P}[M > x] \le \mathbf{P}[M_1 > \delta x] + \mathbf{P}[M_2 > \delta x] + \mathbf{P}[\underline{M}_1 > (1-\delta)x]\mathbf{P}[\underline{M}_2 > (1-\delta)x].$$

The preceding inequality and the already proved case $L = 1$ yields for all sufficiently large x

$$\mathbf{P}[M > x] \le \frac{(1+\epsilon)}{c-\rho-\epsilon} \sum_{i=1}^{2} \rho_i \mathbf{P}[\tau^{i,e} > \delta x/(r_i + \rho + \epsilon - c)] + O(\mathbf{P}[\tau^{1,e} > x]\mathbf{P}[\tau^{2,e} > x]). \quad (7)$$

Thus, by using the fact that $\tau^{i,e} \in \mathcal{IR}$ implies $\sum \mathbf{P}[\tau^{i,e} > x] \in \mathcal{IR}$, dividing (7) with $\sum \mathbf{P}[\tau^{i,e} > x/(r_i + \rho - c)]$, and taking $\limsup_{x\to\infty}$ and $\lim_{\delta\uparrow 1, \epsilon\downarrow 0}$ we derive the desired bound.

($L=2$, *Lower bound.*) First, observe that for any $\epsilon > 0$ such that $c - \rho_{3-i} - \epsilon > \rho_i + \epsilon$ we have

$$M_i \leq \underline{M}_i. \tag{8}$$

Then,

$$\begin{aligned}\mathbf{P}[M > x] &\geq \mathbf{P}[M > x, \underline{M}_1 > y, \underline{M}_2 \leq y] + \mathbf{P}[M > x, \underline{M}_1 \leq y, \underline{M}_2 > y] \\ &\stackrel{\text{def}}{=} P_1 + P_2. \end{aligned}\tag{9}$$

Now, for $\epsilon > 0$ define $Z_i \stackrel{\text{def}}{=} \sup_{t\geq 0}((\rho_i - \epsilon)t - A_i(t))$; note that these variables are almost surely finite. Next, we derive the following sequence of elementary inequalities

$$\begin{aligned}P_1 &\geq \mathbf{P}[M > x, \underline{M}_1 > y, \underline{M}_2 \leq y, Z_2 < y] \\ &= \mathbf{P}[\sup_{t\geq 0}(A_1(t) + A_2(t) - ct) > x, \underline{M}_1 > y, \underline{M}_2 \leq y, \sup_{t\geq 0}((\rho_2 - \epsilon)t - A_2(t)) \leq y] \\ &\geq \mathbf{P}[\sup_{t\geq 0}(A_1(t) + [(\rho_2 - \epsilon)t - y] - ct) > x, \underline{M}_1 > y, \underline{M}_2 \leq y, \sup_{t\geq 0}((\rho_2 - \epsilon)t - A_2(t)) \leq y] \\ &= \mathbf{P}[\sup_{t\geq 0}(A_1(t) - (c - \rho_2 + \epsilon)t) > x + y, \underline{M}_1 > y, \underline{M}_2 \leq y, Z_2 \leq y] \\ &= \mathbf{P}[\underline{M}_2 \leq y, Z_2 \leq y]\mathbf{P}[M_1 > x + y, \underline{M}_1 > y] \\ &= \mathbf{P}[\underline{M}_2 \leq y, Z_2 \leq y]\mathbf{P}[M_1 > x + y],\end{aligned}$$

where the last equality follows from (8). Similarly, we derive an analogous bound for P_2, which, when substituted in (9) renders

$$\mathbf{P}[M > x] \geq \prod_{i=1,2} \mathbf{P}[\underline{M}_i \leq y, Z_i \leq y](\mathbf{P}[M_1 > x + y] + \mathbf{P}[M_2 > x + y]).$$

Now, by applying the result for $L = 1$ and then passing $y \to \infty$ yields the lower bound and completes the proof for $L = 2$.

The general case $L > 2$ follows easily by induction. Assume that the result holds for $L \geq 2$ classes. Now, if the arrival process has $L+1$ classes, define $A_1'(t) \stackrel{\text{def}}{=} \sum_{i=1}^{L} A_i(t)$ and $A_2'(t) = A_{L+1}(t)$ and complete the proof by using the induction hypothesis and exactly the same arguments as for the case $L = 2$. We omit the details. $\diamond$

Appendix

In the following proof we use some known results from the theory of subexponential distributions $\mathcal{S}$; it is well known that $\tau \in \mathcal{IR}$ implies $\tau \in \mathcal{S}$ (see [4, 10]).

Proof of Lemma 2: Let $S(n) = \sum_{i=1}^{n} \tau_i^e$ and $I(n) = \max_{1\leq i\leq n} \tau_i^e$, then clearly $M^e(c) \equiv S(N) - cI(N)$. First, we show that

$$\mathbf{P}[S(n) - cI(n) > x] \sim n\mathbf{P}[\tau^e > x/(1-c)] \quad \text{as} \quad x \to \infty. \tag{10}$$

The lower bound is immediate from $S(n) - cI(n) \geq (1-c)I(n)$ and τ^e being subexponential. For the upper bound, Theorem 2.1 of [10] yields

$$\mathbf{P}[S(n) - I(n) > x] \sim \binom{n}{2}\mathbf{P}^2[\tau^e > x] \quad \text{as} \quad x \to \infty.$$

This and

$$\mathbf{P}[S(n) - cI(n) > x] \le \mathbf{P}[(1-c)I(n) > \delta x] + \mathbf{P}[S(n) - I(n) > (1-\delta)x]$$

yields

$$\limsup_{x\to\infty} \frac{\mathbf{P}[S(n) - cI(n) > x]}{\mathbf{P}[\tau^e > x/(1-c)]} \le n \limsup_{x\to\infty} \frac{\mathbf{P}[\tau^e > \delta x/(1-c)]}{\mathbf{P}[\tau^e > x/(1-c)]},$$

which by using $\tau^e \in \mathcal{IR}$ and passing $\delta \uparrow 1$ yields the upper bound and finishes the proof of (10). Next, since $\tau^e \in \mathcal{S}$, for any $\epsilon > 0$ we can choose a finite constant K_ϵ, such that for all n and $x \ge 0$ (see Lemma A.4 (ii) of [4])

$$\mathbf{P}[S(n) - cI(n) > x] \le \mathbf{P}[S(n) > x] \le K_\epsilon (1+\epsilon)^n \mathbf{P}[\tau^e > x].$$

Hence,

$$\begin{aligned} \frac{\mathbf{P}[S(N) - cI(N) > x]}{\mathbf{P}[\tau^e > x/(1-c)]} &\le \frac{1}{\mathbf{P}[\tau^e > x/(1-c)]} \sum_{n=1}^{\infty} \mathbf{P}[N = n] \mathbf{P}[S(n) > x] \\ &\le K_\epsilon \mathbf{E}(1+\epsilon)^N \sup_{x \ge 0} \frac{\mathbf{P}[\tau^e > x]}{\mathbf{P}[\tau^e > x/(1-c)]} < \infty. \end{aligned} \tag{11}$$

Thus, using (10), (11) and dominated convergence we finish the proof

$$\begin{aligned} \lim_{x\to\infty} \frac{\mathbf{P}[S(N) - cI(N) > x]}{\mathbf{P}[\tau^e > x/(1-c)]} &= \lim_{x\to\infty} \frac{1}{\mathbf{P}[\tau^e > x/(1-c)]} \sum_{n=1}^{\infty} \mathbf{P}[N = n] \mathbf{P}[S(n) - cI(n) > x] \\ &= \sum_{n=1}^{\infty} \mathbf{P}[N = n] \lim_{x\to\infty} \frac{\mathbf{P}[S(n) - cI(n) > x]}{\mathbf{P}[\tau^e > x/(1-c)]} = \rho. \end{aligned}$$

References

[1] O. J. Boxma. Fluid queues and regular variation. *Performance Evaluation*, 27&28:699–712, 1996.

[2] N. G. Duffield. Queueing at large resources driven by long-tailed M/G/∞-modulated processes. *Queueing Systems*, pages 245–266, 1998.

[3] P. R. Jelenković. Subexponential teletraffic models. In *24th Conference on The Mathematics of Operations Reseach*, Lunteren, The Netherlands, January 1999.

[4] P. R. Jelenković and A. A. Lazar. Asymptotic results for multiplexing subexponential on-off processes. Advances in Applied Probability, 31(2), June 1999.

[5] N. Likhanov and R. R. Mazumdar. Cell loss asymptotics in buffers fed by heterogeneous long-tailed sources. preprint, 1999.

[6] M. Parulekar and A. M. Makowski. Tail probabilitites for M/G/∞ input processes (I): preliminary asymptotices. *Queueing Systems*, 27:271–296, 1997.

[7] S. Resnick and G. Samorodnitsky. Activity periods of an infinite server queue and performance of certain heavy tailed fluid queues. *Queueing Systems*, 33, 1999.

[8] S. Resnick and G. Samorodnitsky. Steady state distribution of the buffer content for M/G/∞ input fluid queues. preprint, 1999.

[9] S. Vamvakos and V. Anantharam. On the depature process of a leaky bucket system with long-range dependent input traffic. *Queueing Systems*, 28:191–214, 1998.

[10] E. Willekens. The structure of the class of subexponential distributions. *Probab. Th. Rel. Fields*, pages 567–581, 1988.