J. Appl. Prob. **28**, 245–250 (1991)
Printed in Israel

QUEUES WITH NEGATIVE ARRIVALS

EROL GELENBE,* *University René Descartes*
PETER GLYNN,** *Stanford University*
KARL SIGMAN,*** *Columbia University*

Abstract

We study single-server queueing models where in addition to regular arriving customers, there are negative arrivals. A negative arrival has the effect of removing a customer from the queue. The way in which this removal is specified gives rise to several different models. Unlike the standard FIFO $GI/GI/1$ model, the stability conditions for these new models may depend upon more than just the arrival and service rates; the entire distributions of interarrival and service times may be involved.

STABILITY; REMOVALS

1. Introduction

Consider a single-server FIFO queue with two types of arrivals, *regular* and *negative*. Regular arrivals correspond to customers who upon arrival, join the queue with the intention of getting served and then leaving the system. At a negative arrival epoch, the system is affected if and only if customers are present; in which case a customer is removed from the system. Intuitively, the introduction of negative arrivals makes the system less congested than if they were not present. In particular, one would expect that a steady-state distribution for queue length can exist even when the regular arrival rate is greater than the service rate. In the present paper, we consider several single-server models of this type and give necessary and sufficient conditions for the existence of a unique steady-state distribution of queue length. As we show, the stability conditions may depend upon more than just arrival and service rates.

The notion of negative arrivals was introduced in [3] in the case of a Markovian network generalising a Jackson network [4] to include negative as well as regular

Received 25 August 1989; revision received 2 March 1990.
* Postal address: Ecole des Hautes Etudes en Informatique, Université René Descartes (Paris V), 45 rue des Saints-Peres, 45006 Paris, France.
Research supported in part by CNRS-CS (French National Program on Parallel Computation).
** Postal address: Department of Operations Research, Stanford University, Stanford, CA 94305–4022, USA.
Research supported by the U.S. Army Research Office under Contract DAAL-03-88-K-0063.
*** Postal address: Department of Industrial Engineering and Operations Research, Columbia University, New York NY 10027, USA.
Research supported by NSF Grant DDM 895–7825.

(positive) arrivals. In particular, it was shown [3] that such a network has product form, though its customer flow equations are non-linear. Several practical applications motivate these kinds of models. Negative arrivals can represent commands to delete some transaction, as in distributed computer systems or databases [1] in which certain operations become impossible because of locking of data or because of inconsistency. Negative and positive customers may also represent inhibitory and excitatory signals, respectively, in mathematical models of neural networks [5], [6], while queue length in this case represents the input potential to a neuron. Other types of models that allow customers to leave early before service (but not due to external causes as in the present paper) have been considered in the literature and are called queues with *impatient* customers or queues with *reneging* (see for example, Section 2.9, p. 122 of [5]).

2. Removal of the customer in service

In this section we consider the case when a negative arrival removes the customer (if any) in service (RCS). We study two different models of this kind. The first model assumes that regular arrivals form a marked point process $\psi = \{(t_n, S_n): n \geqq 0\}$, where (t_n, S_n) denotes the nth customer's arrival and service time. Negative arrivals form a Poisson process (at rate α) assumed independent of ψ. Let $L(t)$ denote the number of customers in the system at time t. Let $\{Z_n\}$ and Z be i.i.d $\sim \exp(\alpha)$ independent of ψ. Our first model is very elementary, as the following lemma shows.

Lemma 2.1. *$\{L(t)\}$ has the same distribution as if the model was a FIFO single-server queue with marked point process $\hat{\psi} = \{(t_n, \hat{S}_n): n \geqq 0\}$ with $\hat{S}_n = \min(S_n, Z_n)$.*

Proof. Because of the properties of the Poisson process, when the nth customer begins service, the amount of time spent with the server is distributed as $\min(S_n, Z_n)$. It follows that $\{L(t)\}$ is stochastically identical to that for a FIFO single-server queue having customer arrival times $\{t_n\}$ and service times $\{\hat{S}_n\}$.

The importance of the above proposition is that it allows us to analyze our model using known methods. We give one such example.

Proposition 2.1. *If regular interarrival times are i.i.d. (general distribution A) with rate $0 < \lambda < \infty$, and service times are i.i.d. (general distribution G) at rate $\mu > 0$ and are independent of the arrival process then $\{L(t)\}$ is a positive recurrent regenerative process if and only if $\lambda < \hat{\mu}$ where*

$$\hat{\mu}^{-1} \overset{\text{def}}{=} E\{\min(S, Z)\} = \int_0^\infty e^{-s\alpha}(1 - G(s))ds.$$

Proof. From Lemma 2.1 we see that $\{L(t)\}$ is stochastically the same as the number in system process of a FIFO $GI/GI/1$ queue with interarrival time distribution A and service time distribution that of $\min(S_n, Z_n)$. The result thus follows from well-known results on such queues (see for example, Chapter 9 of [11]).

We point out (more generally) that if ψ is either stationary ergodic or governed by a Harris recurrent Markov process (HRMP) then so is $\hat{\psi}$. In this case let S denote a generic service time from a *Palm version* of ψ. Then $\{L(t)\}$ has a stationary ergodic version as long as $\lambda < E\{\min(S, Z)\}$ ([2], [8], [9]).

Our second model assumes that the superposition of negative and regular arrivals forms a renewal process with general interarrival time distribution F assumed to have non-zero finite first moment; T will denote a generic random variable $\sim F$. Independent of the past, with probability $0 < p < 1$ an arrival is regular and with probability $q = 1 - p$ an arrival is negative. We assume that customer service times $\{S_n : n \geqq 0\}$ are i.i.d. $\sim G$, where G is a general distribution with non-zero finite first moment. $\mu \overset{\text{def}}{=} (E(S))^{-1}$ is the service rate (S denotes a generic service time). We assume that the service time sequence is independent of the exogenous renewal process. We refer to this as the *coin flipping* model (CF). Let $N(t)$ denote the counting process of a non-delayed renewal process having cycle lengths i.i.d. $\sim G$. Let T_- denote a generic negative interarrival time. T_- has the distribution of a geometric sum (parameter q) of i.i.d. F distributed interarrival times; $E(T_-) = E(T)/p$.

$$E(N(T_-)) = \int_0^\infty E(N(t)) \left\{ \sum_{n=1}^{\infty} p^{n-1} q F^{*n}(dt) \right\}$$

denotes the expected number of potential service completions during T_-. Here F^{*n} denotes the nth-fold convolution of F.

Proposition 2.2. *For the coin flipping RCS model,* $\{L(t)\}$ *is positive recurrent regenerative if and only if* $\delta \overset{\text{def}}{=} (p/q) - E(N(T_-)) - 1 < 0$.

Proof. First we prove sufficiency. Let t_n denote the time of the nth negative arrival. Let $X_n = L(t_n +)$. It is easily seen that X is an aperiodic irreducible Markov chain with state space the non-negative integers. We show that X is ergodic by use of *Foster's negative drift criterion* (see for example, Theorem 6.1 of [10]). To this end, it suffices to show that for some k there exists an $\varepsilon > 0$ such that for $l > k$, $E_l(X_1) < l - \varepsilon$. Observe that p/q is precisely the expected number of regular arrivals between two consecutive negative arrivals. For l sufficiently large we thus have $E_l(X_1) - l \approx \delta < 0$ by our hypothesis. It follows that for any $0 < \varepsilon < |p/q - E(N(T_-)) - 1|$ a k can be found giving us the desired negative drift. Thus X is ergodic. Assume that $X_0 = 0$, let τ_n denote the consecutive times at which X hits 0 and embed them into continuous time via $s_n = t_{\tau_n}$. By ergodicity $E(\tau_1) < \infty$ and hence by Wald's equation $E(s_1) = E(\tau_1)E(T_-) = E(\tau_1)E(T)/p < \infty$. Thus $L(t)$ is positive recurrent regenerative with respect to the renewal process (s_n). For necessity, we apply (iii) of Theorem 11.3 of [10]. Assume that $p/q - E(N(T_-)) - 1 = \delta > 0$. We must find a bounded non-negative function g such that $E_l g(X_1) \geqq g(l)$, $x > k$ and $g(l) > \sup\{g(y) : y \leqq k\}$, $l > k$. For k large, we will use $g(x) = \min(x, k + \delta)$. Let

A be distributed as the number of regular arrivals during T_-. If $l > k$, $g(l) = \min(l, k+\delta)$ and $E_l g(X_1) = E\{\min((l + A - N(T_-) - 1)^+, k+\delta)\} \approx k + \delta$ as $k \to \infty$. Thus X_n is transient. In particular only finitely many regular arrivals ever find an empty system. But visits of X_n to any fixed state i also can occur only finitely many times, thus, $L(t) \to \infty$ as $t \to \infty$ and $L(t)$ can not be positive recurrent regenerative with respect to any embedded renewal process. For the case $\delta = 0$, we can apply Theorem 8.1 of [10] with $g(x) = x$, $A = [0, k]$ and $\beta = p/q$ to show that the set $[0, k]$ is null recurrent for all sufficiently large k. This implies that X is null recurrent and hence $P(s_1 < \infty) = 1$ and $E(s_1) = \infty$ (via Wald's equation). In this case, let p_k denote the limit as $t \to \infty$ of $1/t \int_0^t P(L(s) = k)ds$. From renewal theory $p_k = 0$ if $E \int_0^{s_1} I(L(s) = k)ds < \infty$. To establish this, consider an alternative system (designated by $\tilde{L}$, etc.) with the same input except each service time has been scaled by $\alpha > 0$. For $\alpha < 1$, $\tilde{L}(t) \leqq L(t)$ for all t. Moreover, for α sufficiently small $\tilde{\delta} < 0$ so that $\tilde{L}(t)$ is positive recurrent. Hence

$$E \int_0^{s_1} I(L(s) \leqq k)ds \leqq E \int_0^{\tilde{s}_1} I(\tilde{L}(s) = k)ds < \infty.$$

Thus $p_k = 0$ for all k implying that $L(t)$ is not stochastically tight and hence can not be positive recurrent regenerative with respect to any embedded renewal process.

Remark 2.1. In the case of Poisson regular arrivals (rate λ), negative interarrival times i.i.d. $\sim A$ (general distribution), and i.i.d. service times (general distribution G) one can use the method used in proving Proposition 2.2 to obtain the following result: $L(t)$ is positive recurrent regenerative if and only if $\lambda E(T_-) - E(N(T_-)) - 1 < 0$ where now $T_- \sim A$.

3. Removal of the customer at the tail of the queue

In this section we consider a model where at a negative arrival epoch, the customer in the system who arrived most recently is removed; if the system is empty then nothing is done. Observe that the only time a customer is removed from service is when he is the only one in the system at a negative arrival epoch. This model amounts to removing the customer at the tail of the queue; hence, we refer to it as the RCT model. We assume that the superposition of regular and negative arrival times, $\{t_n\}$, forms a renewal process at rate $0 < \lambda < \infty$ with interarrival times T_n i.i.d. $\sim F$. Let $K_n = 1$ if the nth arrival is positive and 0 if negative. We shall assume that $\{K_n\}$ is a recurrent Markov chain independent of $\{t_n\}$ and having the property that $P(K_1 = 1 \mid K_0 = 0) = 1$, that is, two or more negative arrivals cannot occur in a row. Let $p = P(K_1 = 1 \mid K_0 = 1)$, $q = P(K_1 = 0 \mid K_0 = 1) = 1 - p$; $0 < p < 1$. Service times S_n are assumed i.i.d. $\sim G$, independent of $\{t_n, K_n\}$ and we assume that $0 < \mu^{-1} \stackrel{\text{def}}{=} E(S) < \infty$.

Proposition 3.1. *For the RCT model, $\{L(t)\}$ is positive recurrent regenerative if and only if $\lambda < \mu(1 + p)/(1 - p)$.*

Proof. For mathematical convenience, we imagine that every arrival (regular or negative) brings a service time with it; S_n denotes the service time of the nth arrival. Let $V(t)$ denote the total amount of work in system at time t. Define $W_n = V(t_n -)$ if $K_n = 1$;

$W_n = V(t_n+)$ if $K_n = 0$. Define $\Delta_n = S_n K_n - T_n - S_n(1 - K_{n+1})$. It follows (with a little thought) that

$$W_{n+1} = \{W_n + \Delta_n\}^+, \tag{3.1}$$

and hence we have a reflected random walk with increments Δ_n governed by the Harris ergodic Markov chain $X_n = (S_n, T_n, K_n, K_{n+1})$. From known results (see [8]) we obtain Harris ergodicity of the Markov chain $Z_n = (W_n, X_n)$ if and only if $E_\nu(\Delta_n) < 0$ where ν is the stationary distribution for X_n. It is easily shown that $P_\nu(K_n = 0) = p/(1+p)$ and hence (since $\{K_n\}$ is assumed independent of $\{(S_n, T_n)\}$) that

$$E_\nu(\Delta_n) = E(S)\frac{(1-p)}{(1+p)} - E(T).$$

Thus (W_n, X_n) is positive Harris recurrent if and only if $\lambda < \mu(1+p)/(1-p)$; in this case let π be the stationary distribution for (W_n, X_n). $P_\pi(W_n = 0) > 0$ and hence either $P_\pi(W_n = 0, K_n = 1) > 0$ or $P_\pi(W_n = 0, K_n = 0) > 0$. But if the event $\{W_n = 0, K_n = 0\}$ occurs (a negative arrival leaves behind an empty system) then so will $\{W_{n+1} = 0, K_{n+1} = 1\}$ (the next arrival, necessarily regular, finds the system empty) since by assumption, two negative arrivals can not occur in a row. It follows that $P_\pi(W_n = 0, K_n = 1) > 0$ and hence by Wald's equation $L(t)$ is positive recurrent with regeneration points the consecutive times at which a regular arrival finds an empty system. Analogous to the proof of Proposition 2.2, $\{L(t)\}$ is not stochastically tight when $E_\nu(\Delta_n) \geqq 0$, and hence cannot be positive recurrent regenerative with respect to any embedded renewal process.

Remark 3.1. Results like Proposition 3.1 can be obtained in a more general set-up: when (S_n, T_n, K_n) is assumed governed by a positive Harris recurrent Markov chain Θ (and $P(K_{n+1} = 0 \mid K_n = 0) = 0$). In this case, however, the regenerative structure of $L(t)$ is more complicated.

Remark 3.2. When the input to RCT is a general marked point process $\psi = (t_n, S_n, K_n)$, the stability conditions can be quite varied as the following example shows. Start with an initially empty $M/M/1$ queue with arrival rate λ and service rate μ; $\rho \stackrel{\text{def}}{=} \lambda/\mu < 1$. Let $\hat{\psi} = (t_n, S_n)$ denote the corresponding point process of arrival and service times. Let I_n denote the length of the nth idle period and choose U_n such that $(U_n \mid I_n) \sim \text{Uniform}(0, I_n)$ and (U_n, I_n) are i.i.d. As soon as the nth idle period begins, let a negative arrival occur U_n time units later. Construct a new marked point process $\psi = (\bar{t}_n, \bar{S}_n, K_n)$ where $\bar{t}_n$ denotes the nth arrival (regular or negative), $\bar{S}_n$ the nth arrival's service time (we imagine that negative arrivals bring one with them), and K_n is defined as before. By construction, every negative arrival finds an empty system and hence has no effect on the system; $\rho < 1$ remains the correct stability condition. Moreover, the proportion of arrivals that are negative is strictly greater than zero. Finally observe that the point process regenerates at each negative arrival epoch.

References

[1] Cellary, W., Gelenbe, E. and Morzy, J. (1988) *Concurrency Control in Distributed Databases*. North-Holland, Amsterdam.

[2] Franken, P., Koenig, D., Arndt, U. and Schmidt, V. (1981) *Queues and Point Processes*. Akademie Verlag, Berlin.

[3] Gelenbe, E. (1991) Product form networks with negative and positive customers. *J. Appl. Prob.* **28**(3).

[4] Gelenbe, E. and Pujolle, G. (1986) *Introduction to Networks of Queues*. Wiley, London.

[5] Gross, D. and Harris, C. M. (1985) *Fundamentals of Queueing*, 2nd edn. Wiley, New York.

[6] Kandel, E. C. and Schwartz, J. H. (1985) *Principles of Neural Science*. Elsevier, Amsterdam.

[7] Rumelhart, D. E., McClelland, J. L. and the PDP Research Group (1986) *Parallel Distributed Processing*, Vols. I and II. Bradford Books and MIT Press, Cambridge, Mass.

[8] Sigman, K. (1988) Queues as Harris recurrent Markov chains. *Queueing Systems* **3**, 179–198.

[9] Sigman, K. (1989) One-dependent regenerative processes and queues in continuous time. *Math. Operat. Res.* **15**, 175–189.

[10] Tweedie, R. L. (1976) Criteria for classifying general Markov chains. *Adv. Appl. Prob.* **8**, 737–771.

[11] Wolff, R. W. (1989) *Stochastic Modeling and the Theory of Queues*. Prentice-Hall, Englewood Cliffs, NJ.