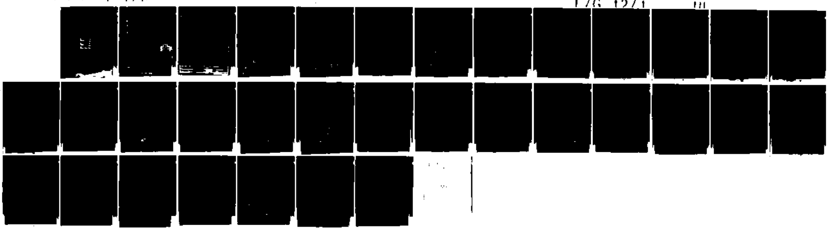
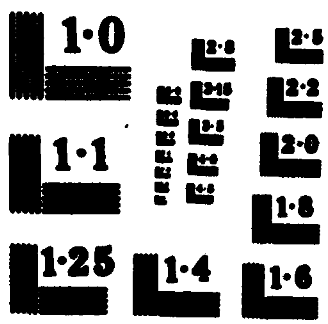
AD A148 538   A SINGLE SERVER QUEUE WITH MIXED TYPES OF INTERRUPTIONS   1/1
(U) DUKE UNIV DURHAM NC DEPT OF COMPUTER SCIENCE
V F NICOLA NOV 84 AFOSR TR 84 1097 DAAG29 84 K 0045

UNCLASSIFIED   F/G 12/1   NL




1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5

AFOSR-TR- 84-1097

(4)

AD-A148 538

A Single Server Queue With Mixed Types of Interruptions *

Victor F. Nicola

Department of Computer Science
Duke University
Durham, NC 27706

ABSTRACT

This document considered ~~We consider~~ an M/G/1 queue with mixed types of Poisson interruptions. A derivation of the Laplace Steiltjes Transform of the completion time associated with a customer's service is presented. We introduce the definition of the effective service time and give probabilistic arguments to the derivation of the first and the second moments of the completion time. The average number of customers in the system is obtained and the relation to the Pollaczek-Khintchine formula is noted. The results are applied to the modeling of checkpointing and recovery in a transactional system.

AFOSR-84-0132

* This work was supported in part by Air Force Office of Scientific Research, by the Army Research Office under contract DAAG[illegible] and by the National Science Foundation under grant [illegible].

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 84-1097
6a. NAME OF PERFORMING ORGANIZATION Duke University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code) Department of Computer Science Durham NC 27706	7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-84-0132
8c. ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F PROJECT NO.: 2304 TASK NO.: A5 WORK UNIT NO.:

11. TITLE (Include Security Classification)
A SINGLE SERVER QUEUE WITH MIXED TYPES OF INTERRUPTIONS

12. PERSONAL AUTHOR(S)
Victor F. Nicola

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	NOV 84	30

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The investigators consider an M/G/1 queue with mixed types of Poisson interruptions. A derivation of the Laplace Steiltjes Transform of the completion time associated with a customer's service is presented. The investigators introduce the definition of the effective service time and give probabilistic arguments to the erivation of the first and the second moments of the completion time. The average number of customers in the system is obtained and the relation to the Pollacsek-Khintchine formula is noted. The results are applied to the modeling of checkpointing and recovery in a transactional system.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL MAJ Brian W. Woodruff	22b. TELEPHONE NUMBER [illegible] 22c. OFFICE SYMBOL [illegible]

DD FORM 1473, 83 APR [illegible]

TABLE OF CONTENTS

1. Introduction

The analysis of queueing systems with mixed types of interruptions is important in many computer systems modeling applications such as systems operating in different modes, systems subject to breakdowns and repairs and priority queueing systems.

The M/G/1 queue with a single type of Poisson interruptions was dealt with extensively by Gaver [4] for a variety of service-interruption interactions. The analysis was based on the definition of the completion time. He derived the Laplace Steiltjes transform(LST) of the completion time and used the method of imbedded Markov chain and the renewal theory to obtain the generating function of the distribution of the number of customers in the system. In this paper we extend the results of Gaver to allow the simultaneous presence of different types of interruptions. The analysis is also based on the completion time. General Markovian reward models for the analysis of the completion time has been studied recently [10,11] for various types of service-interruption interaction. Yet due to the assumption of exponential holding times the model presented here is not included as a special case.

We introduce the definition of the effective service time associated with a customer's service. This is meaningful and useful for the demonstation of an alternative probabilistic arguments to the derivation of the first and the second moments of the completion time. The steady-state average number of customers in the system is obtained and the relation to the Pollaczek-Khintchine formula is noted.

Section 2 contains a description of the system and the different types of the service-interruption interaction and introduces basic definitions. In section 3 we derive the probability distribution function of the completion time and give probabilistic arguments to the derivation of its first and second moments. The steady-state average number of customers in the system is obtained in section 4. In section 5, applicaton to the modeling of checkpointing and recovery in a transactional database system is considered.

2. Basic Definitions

Consider the M/G/1 queue subject to different sources of Poisson interruptions of different types. Customers receive service according to the FCFS discipline. It is necessary to distinguish different types of interruptions. Independent interruptions may arrive when the system is idle or when the system is servicing a customer. Active interruptions may arrive only when the system is servicing a customer. No interruptions may arrive when the system is servicing an interruption.

The following is a classification of the different types of service-interruption interactions considered in this paper (see figure 1).

I) Preemptive interruption (*pmu*):

Customer's service is preempted immediately on arrival of an interruption. After servicing the interruption there are two possibilities; namely:

a) preemptive-resume (*prs*): the customer's service is resumed from the point at which it was preempted.

b) preemptive-repeat (*prt*): the customer's service is repeated from its beginning. In preemptive-repeat-identical(*pri*) interruption, the same identical customer's service is repeated. In preemptive-repeat-different(*prd*) interruption, a corresponding customer's service time of the same distribution is repeated.

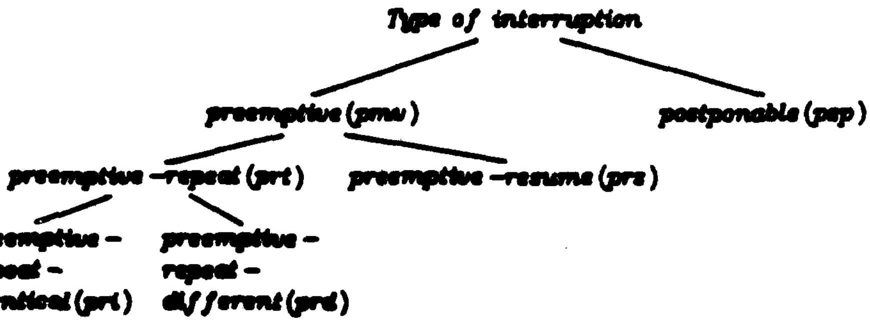


Figure 1. Classification of different types of service-interruption interaction

II) Postponable interruption (pop):

Customer's service continues upon the arrival of an interruption. The interruptions accumulated during the customer's service are serviced immediately after servicing the customer.

Any of the interruptions classified above may be active(*act*) or independent(*ind*). We define the subsets of interruption sources: *aprd*, *apri*, *aprt*, *aprs*, *apmu*, *apop* and *act*, corresponding to the different types of active interruption, and the subsets: *iprd*, *ipri*, *iprt*, *iprs*, *ipmu*, *ipop* and *ind*, corresponding to the different types of independent interruption. It follows for the subsets of active interruption sources

$aprt = apri \cup aprd$,

$apmu = aprt \cup aprs$,

$act = apmu \cup apop$,

and for the subsets of independent interruption sources

$iprt = ipri \cup iprd$,

$ipmv = iprt \cup iprs$,

$ind = ipmv \cup ipsp$.

Define also the subsets: pri, prd, prt, prs, pmv and psp such that

$pri = apri \cup ipri$,

$prd = aprd \cup iprd$,

$prt = aprt \cup iprt$,

$prs = aprs \cup iprs$,

$pmv = apmv \cup ipmv$,

$psp = apsp \cup ipsp$.

The total set of interruption sources, T, that may be present in the system is given by

$T = pmv \cup psp = act \cup ind$.

In subsequent discussion the index $t\,(t \in T)$ indicates the source of interruption (note that there may be more than one source of the same type).

The following notations describe the system

λ is the customer's arrival rate.

S is the customer's service time; a random variable with a probability distribution function $G(x)=P(S \le x)$ and $LST\ \tilde{S}(s)$.

u_t is the arrival rate of interruptions from source t.

I_t is the time duration of source t interruption; a random variable with probability distribution function $G_t(x)=P(I_t \le x)$ and $LST\ \tilde{I}_t(s)$.

We define the effective service time, S_e, to be the random interval of time spent by the system in servicing a customer, including the repetitions due to prt interruptions during the customer's service and excluding the time duration of interruptions. Thus

$$S_e = \begin{cases} S, & \text{if no } prt \text{ interruptions are present} \\ \sum_{d=1}^{N_D} \left\{ \sum_{i=1}^{N_I(d)} S(i,d) + S(d) \right\} + \sum_{i=1}^{N_I'} S'(i) + S', & \\ & \text{if } prt \text{ interruptions are present} \end{cases} \tag{2.1}$$

$N_D \geq 0$ is the total random number of *prd* interruptions (possibly from different sources) that arrived during the customer's service.

$N_I(d) \geq 0$, $1 \leq d \leq N_D$, is the total random number of *pri* interruptions (possibly from different sources) that arrived between the $(d-1)-th$ and the $d-th$ *prd* interruptions ($d-1=0$ corresponds to the beginning of the customer's service).

$S(i,d)$, $1 \leq i \leq N_I(d)$, $1 \leq d \leq N_D$, is the random interval of time between the $(i-1)-th$ and the $i-th$ *pri* interruptions that arrived between the $(d-1)-th$ and the $d-th$ *prd* interruptions ($i-1=0$ corresponds to the beginning of service after the $(d-1)-th$ *prd* interruption).

$S(d)$, $1 \leq d \leq N_D$, is the random interval of time between the $d-th$ *prd* interruption and the preceding *pri* interruption. Note that $S(d)$ and $S(i,d)$, $1 \leq i \leq N_I(d)$, are dependent random variables.

$N_I' \geq 0$ is the total random number of *pri* interruptions that arrived between the last *prd* interruption and service completion.

S' is the random interval of time between the last *pri* interruption and service completion; it is the customer's service time which is restarted following the last *prd* interruption.

$S'(i) \leq S'$, $1 \leq i \leq N'$, is the random interval of time between the $(i-1)-th$ and the $i-th$ *pri* interruptions that arrived after the last *prd* interruption. Note that S' and $S'(i)$, $1 \leq i \leq N_I'$, are dependent random variables.

It is important to note that when different types of interruptions are present, S_e is merely determined by the *prt* (*pri* and *prd*) interruptions. If there are no *prt* interruptions, then S_e is identical to the customer's service time S. Let $Q_e(x)=P(S_e \leq x)$ be the probability distribution function of the effective service time and denote by $S_e^{\sim}(s)$ its LST.

The completion time, C (as defined by Gaver), is the random interval of time between the instant at which the customer's service begins and the instant at which the service of the next customer may begin (does begin provided that a customer is present). It follows that

$$C = S_e + \sum_{t \in T} \sum_{k=1}^{N_t} I_t(k) \tag{2.2}$$

N_t is the random number of source t interruptions that arrived during the customer's service, and $I_t(k)$ is the random time duration of the $k-th$ interruption of source t. S_e is as given by equation (2.1). Let $G_e(x)=P(C \le x)$ be the probability distribution function of the completion time and denote by $C^{\sim}(s)$ its LST. It is important to notice that the completion times of successive customers are independent and identically distributed random variables.

The following notations are presented for the completion time and will be used for all random variables under consideration.

The $i-th$ moment $E(C^i)$ is given by the following relation

$$E(C^i) = (-1)^i \left[\frac{d^i C^{\sim}(s)}{ds^i} \right]_{s=0}, \quad i=1,2,... \tag{2.3}$$

The expected residual time, $R(C)$, is given by the following relation

$$R(C) = \frac{E(C^2)}{2\,E(C)} . \tag{2.4}$$

In the following section we consider the analysis of the completion time.

3. The Completion Time

This section is devoted to the analysis of the completion time associated with a customer's service. Clearly the analysis is independent of queueing aspects. We will derive the probability distribution of the completion time and give an alternative probabilistic arguments to obtain its first and second moments.

Consider a single server subject to different types of Poisson interruptions. Interruptions may arrive only during the customer's service. The interruptions from source $t \in T$ have durations that are independent and identically distributed.

It is important to remark that the *psp* and the *prs* types of interruption have exactly the same effect on the completion time (but not the same queueing effect). Therefore we can group these two types of interruption into one type, say *prs*, in the analysis of the completion time. Without loss of generality we will consider the simultaneous presence of a single source from each type of interruption; namely, *prs*, *pri* and *prd*. The results are similar in the case where one or more sources from each type are present. Thus in the analysis of this section we will consider the mixture of the following three sources of interruption:

i) *prs* source, with interruption rate v_s and duration I_s with *LST* given by $\Gamma_s(s)$.
ii) *pri* source, with interruption rate v_i and duration I_i with *LST* given by $\Gamma_i(s)$.
iii) *prd* source, with interruption rate v_d and duration I_d with *LST* given by $\Gamma_d(s)$.

The effective service time and the completion time associated with a customer's service were defined in section 2. The corresponding *LST*s are defined as follows

$$\tilde{S}_e(s) = \begin{cases} \int_0^\infty \tilde{S}_e(s|S=x)\, dG(x), & \textit{if no prd interruptions are present} \\ E(e^{-sS_e}), & \textit{otherwise} \end{cases} \tag{3.1}$$

where

$$\tilde{S}_e(s|S=x) = E(e^{-sS_e}|S=x) \ . \tag{3.2}$$

Similarly,

$$\tilde{C}(s) = \begin{cases} \int_0^\infty \tilde{C}(s|S=x)\, dG(x), & \textit{if no prd interruptions are present} \\ E(e^{-sC}), & \textit{otherwise} \end{cases} \tag{3.3}$$

where

$$\tilde{C}(s|S=x) = E(e^{-sC}|S=x) \ . \tag{3.4}$$

We are particularly interested in the first and the second moments of the above random variables. The expected residual time is given by the following relations

$$R(S_e) = \begin{cases} \int_0^\infty R(S_e|S=x)\, dG(x), & \textit{if no prd interruptions are present} \\ \dfrac{E(S_e^2)}{2\,E(S_e)}, & \textit{otherwise} \end{cases} \tag{3.5}$$

where

$$R(S_e|S=x) = \frac{E(S_e^2|S=x)}{2\,E(S_e|S=x)} \ . \tag{3.6}$$

Similarly,

$$R(C) = \begin{cases} \int_0^\infty R(C|S=x)\, dG(x), & \textit{if no prd interruptions are present} \\ \dfrac{E(C^2)}{2\,E(C)}, & \textit{otherwise} \end{cases} \tag{3.7}$$

where

$$R(C|S=x) = \frac{E(C^2|S=x)}{2\,E(C|S=x)} \ . \tag{3.8}$$

The following remark will be used in subsequent analysis. Let H be the holding time in the operating state (servicing the customer) between any two interruptions, and denote by $G_h(x)=P(H\leq x)$ its probability distribution function. From the Poisson property of all interruptions it follows that

$$dG_h(x) = v\, e^{-vx}\, dx \tag{3.9}$$

where $v = v_s + v_t + v_d$ is the total interruption rate.

In the next section we proceed to derive the *LST* of the completion time.

3.1. The Laplace Steiltjes Transform (*LST*)

Consider a customer that starts being serviced with initial service time $S_0 = x$ (note that the customer's service time changes after any *prd* interruption). Upon the arrival of *prs* interruption the customer's service is preempted for the duration of the interruption. The same customer's service is resumed after the interruption. The initial customer's service may complete after a number of *prs* interruptions and before the arrival of any *prt* interruption. Otherwise it is preempted and repeated after the *prt* interrption. After a *pri* interruption the same initial customer's service is restarted. After a *pr* interruption a different customer's service ($\neq x$) is repeated. The new customer's service may be completed before any *prt* interruption, otherwise it is repeated, and so on.

Let $C_1(x)$ be the total time spent in servicing the customer and the *prs* interruptions until x units of service time are completed and before the arrival of any *prt* interruption. Note that $C_1(x)$ has incomplete distribution, since $C_1(x) = \infty$, if any *prt* interruption arrives before completing x units of service time.

Let $C_2(x)$ be the total time spent in servicing the customer and the *prs* interruptions until the arrival of any *prt* interruption and before the completion of x units of service time. Note that $C_2(x)$ has incomplete distribution, since $C_2(x) = \infty$, if x units of service time are comleted before the arrival of any *prt* interruption. Furthermore it is clear that

$$P(C_1(x) = \infty) + P(C_2(x) = \infty) = 1,$$

since the two events are exhaustive and mutually exclusive.

Define the following *LST*s

$$\tilde{C}_1(s,x) = E(e^{-sC_1(x)}), \tag{3.10}$$

and

$$\tilde{C}_2(s,x) = E(e^{-sC_2(x)}) \ . \tag{3.11}$$

The following two lemmas determine the above *LST*s which are useful for determining $\tilde{C}(s)$ defined in equation (3.3) as will be shown in theorem 3.1.

Lemma 3.1 : The *LST* $\tilde{C}_1(s,x)$ as defined in equation (3.10) is given by

$$\tilde{C}_1(s,x) = e^{-(s+v-v_s\tilde{\Gamma}_s(s))x}, \tag{3.12}$$

where v and v_s are the total and the *prs* interruption rates, respectively.

Proof : Conditioning on H, the holding time until the first interruption, we have

$$\tilde{C}_1(s,x|H=h) = E(e^{-sC_1(x)}|H=h)$$

$$= \begin{cases} e^{-sx}, & \text{if } h \geq x \\ \dfrac{v_s}{v} e^{-sh}\,\tilde{\Gamma}_s(s)\,\tilde{C}_1(s,x-h)\,, & \text{if } h < x \end{cases}$$

Unconditioning with respect to H, we get

$$\tilde{C}_1(s,x) = e^{-(s+v)x} + \int_0^x v_s\, e^{-(s+v)h}\,\tilde{\Gamma}_s(s)\,\tilde{C}_1(s,x-h)\,dh$$

Define the double transform

$$\tilde{C}_1^*(s,u) = \int_0^\infty e^{-ux}\, \tilde{C}_1(s,x)\, dx$$

After changing the order of integration and rearranging, it follows that

$$\tilde{C}_1^*(s,u) = \frac{1}{s+v-v_s\tilde{\Gamma}_s(s)+u}$$

Inverting with respect to u yields equation (3.12). *Q.E.D.*

Lemma 3.2 : The *LST* $\tilde{C}_2(s,x)$ as defined in equation (3.11) is given by

$$\tilde{C}_2(s,x) = \frac{v-v_s}{s+v-v_s\tilde{\Gamma}_s(s)}[1-e^{-(s+v-v_s\tilde{\Gamma}_s(s))x}], \tag{3.13}$$

where v and v_s are the total and the *prs* interruption rates, respectively.

Proof : Conditioning on H, the holding time until the first interruption, we have

$$\tilde{C}_2(s,x|H=h) = E(e^{-sC_2(x)}|H=h)$$
$$= \begin{cases} e^{-sh}, & \text{if } prt \text{ preemption} \\ e^{-sh}\,\tilde{\Gamma}_s(s)\,\tilde{C}_2(s,x-h), & \text{if } prs \text{ preemption} \end{cases}$$

Unconditioning with respect to H, we get

$$\tilde{C}_2(s,x) = \int_0^\infty (v-v_s)\, e^{-(s+v)h}\, dh + \int_0^\infty v_s\, e^{-(s+v)h}\, \tilde{\Gamma}_s(s)\, \tilde{C}_2(s,x-h)\, dh$$

Define the double transform

$$\tilde{C}_2^*(s,u) = \int_0^\infty e^{-ux}\, \tilde{C}_2(s,x)\, dx$$

After changing the order of integration and rearranging, it follows that

$$\tilde{C}_2^*(s,u) = \frac{v-v_s}{s+v-v_s\tilde{\Gamma}_s(s)}\left[\frac{1}{u} - \frac{1}{s+v-v_s\tilde{\Gamma}_s(s)}\right]$$

Inverting with respect to u yields equation (3.13). *Q.E.D.*

Now we can proceed to determine $C^*(s)$, the *LST* of the completion time, in the following main theorem.

Theorem 3.1 : The *LST* of the completion time associated with a customer's service is given by

$$C^*(s)=\frac{\displaystyle\int_0^\infty \frac{e^{-(s+v-v_3\Gamma_3(s))x}}{1-\dfrac{v_1\Gamma_1(s)}{s+v-v_3\Gamma_3(s)}[1-e^{-(s+v-v_3\Gamma_3(s))x}]}\,dG(x)}{1-\displaystyle\int_0^\infty \frac{\dfrac{v_2\Gamma_2(s)}{s+v-v_3\Gamma_3(s)}[1-e^{-(s+v-v_3\Gamma_3(s))x}]}{1-\dfrac{v_1\Gamma_1(s)}{s+v-v_3\Gamma_3(s)}[1-e^{-(s+v-v_3\Gamma_3(s))x}]}\,dG(x)}, \tag{3.14}$$

where v and v_3 are the total and the *prs* interruption rates, respectively.

Proof : Let $C(x)$ be the total time to completion of a customer's service given that its initial service time $S_0=x$. The corresponding *LST* is given by

$$C^*(s,x) = E(e^{-sC(x)})$$

$$= \begin{cases} \dfrac{C_1^*(s,x)}{P(C_1(x)<\infty)}, & \textit{if completion before prt interruption} \\ \dfrac{C_2^*(s,x)}{P(C_2(x)<\infty)}\left[\dfrac{v_1\Gamma_1(s)}{v-v_3}C^*(s,x)+\dfrac{v_2\Gamma_2(s)}{v-v_3}C^*(s)\right], & \textit{otherwise} \end{cases}$$

It follows that

$$C^*(s,x) = C_1^*(s,x) + \frac{v_1\Gamma_1(s)}{v-v_3}C_2^*(s,x)\,C^*(s,x) + \frac{v_2\Gamma_2(s)}{v-v_3}C_2^*(s,x)\,C^*(s)$$

Substituting for $C_1^*(s,x)$ and $C_2^*(s,x)$ from lemmas 3.1 and 3.2 and rearranging, we get

$$C^*(s,x)=\frac{e^{-(s+v-v_3\Gamma_3(s))x}+\dfrac{v_2\Gamma_2(s)}{s+v-v_3\Gamma_3(s)}C^*(s)\,[1-e^{-(s+v-v_3\Gamma_3(s))x}]}{1-\dfrac{v_1\Gamma_1(s)}{s+v-v_3\Gamma_3(s)}[1-e^{-(s+v-v_3\Gamma_3(s))x}]} \tag{3.15}$$

Unconditioning with respect to the initial customer's service time S_0 and rearranging yields equation (3.14). *Q.E.D.*

The following corollary determines $\tilde{S}_e(s)$, the *LST* of the effective service time.

Corollary 3.1 : The *LST* of the effective service time associated with a customer's service is given by

$$\tilde{S}_e(s) = \frac{\displaystyle\int_0^\infty \frac{e^{-(s+u_i+u_d)x}}{s+u_d+u_i\, e^{-(s+u_i+u_d)x}}\, dG(x)}{1-\displaystyle\int_0^\infty \frac{u_d\left[1-e^{-(s+u_i+u_d)x}\right]}{s+u_d+u_i\, e^{-(s+u_i+u_d)x}}\, dG(x)} \tag{3.16}$$

where u_i and u_d are the *pri* and the *prd* interruption rates, respectively.

Proof : Let $S_e(x)$ be the effective service time associated with a customer's service given that its initial service time $S_0=x$. The corresponding *LST* is denoted by $\tilde{S}_e(s,x)$. From earlier definitions in section 2, it should be clear that the effective service time follows from the completion time by setting the duration of all interruptions to zero. Thus substituting for $\tilde{\Gamma}_s(s) = \tilde{\Gamma}_i(s) = \tilde{\Gamma}_d(s) = 1$ in equation (3.15) yields the following

$$\tilde{S}_e(s,x) = E(e^{-sS_e(x)})$$
$$= \frac{(s+u_d+u_i)e^{-(s+u_i+u_d)x} + u_d\, \tilde{S}_e(s)\left[1-e^{-(s+u_i+u_d)x}\right]}{s+u_d+u_i\, e^{-(s+u_i+u_d)x}}, \tag{3.17}$$

where we made use of $v - u_s = u_d + u_i$.

Equation (3.16) follow from equation (3.17) or by similar substitution in equation (3.14). *Q.E.D.*

It is interesting to note that the effective service time depends only on the rates of the *pri* and *prd* interruptions and on the distribution of the customer's service time.

3.2. The First Moment and the Expected Residual Time

Clearly the i-th moment of the completion time can be determined from the *LST* $\tilde{C}(s)$ by equation (2.3), though this is a lengthy and error-prone task. A more elegant technique follows the same steps as in the derivation of $\tilde{C}(s)$.

We give an alternative derivation of the first moment and the expected residual time. This derivation is based on probabilistic arguments; it is presented as an intuitive confirmation of results which are proved rigorously. The derived expressions relate the moments of the completion time to the moments of the effective service time , the rates and the moments of the interruptions. This is advantageuos since it is esier to evaluate the moments of the effective service time.

First we introduce some quantities that will be used in the following discussions. Let A be the expected fraction of completion time spent by the system in actually servicing the customer. From the Poisson property of interruptions it follows that the expected fraction of completion time spent by the system in servicing source i interruptions, A_i, is given by $A\nu_i E(I_i)$. From the normalizing condition we get

$$A = (1 + \sum_{i \in T} \nu_i\, E(I_i))^{-1}\,, \tag{3.18}$$

and

$$A_i = \nu_i\, E(I_i)\,(1 + \sum_{i \in T} \nu_i\, E(I_i))^{-1}\,. \tag{3.19}$$

We state the main results in the following theorem.

Theorem 3.2 : In the presence of all types of interruptions; namely, *prs*, *pri*, and *prd*, the first moment and the expected residual time associated with a customer's service are given by the following relations

$$E(C) = \frac{E(S_0)}{A} , \tag{3.20}$$

and

$$\begin{aligned} R(C) = R(S_0) &+ A_0 \left[R(I_0) + \frac{R(S_0)}{A}\right] \\ &+ A_2 \left[R(I_2) + E(C)\right] \\ &+ A_1 \left[R(I_1) + \frac{1}{A}\int_0^\infty \frac{(E(S_0|S_2=x))^2}{E(S_0)} dG(x)\right] , \end{aligned} \tag{3.21}$$

where A, A_0, A_1, A_2 are given by equations (3.18) and (3.19),

$$\begin{aligned} E(S_0|S_2=x) &= -\left[\frac{d\, S^*_0(s,x)}{ds}\right]_{s=0} \\ &= (1 + u_2\, E(S_0))\ \frac{1-e^{-(u_1+u_2)x}}{u_2+u_1\, e^{-(u_1+u_2)x}} , \end{aligned}$$

$$E(S_0) = \frac{\displaystyle\int_0^\infty \frac{1-e^{-(u_1+u_2)x}}{u_2+u_1\, e^{-(u_1+u_2)x}} dG(x)}{\displaystyle 1 - \int_0^\infty \frac{u_2(1-e^{-(u_1+u_2)x})}{u_2+u_1\, e^{-(u_1+u_2)x}} dG(x)} ,$$

and

$$R(S_0) = (1 + u_2\, E(S_0))\ \frac{\displaystyle\int_0^\infty \frac{1-e^{-(u_1+u_2)x} - (u_1+u_2)\, x\, e^{-(u_1+u_2)x}}{(u_2+u_1\, e^{-(u_1+u_2)x})^2} dG(x)}{\displaystyle\int_0^\infty \frac{1-e^{-(u_1+u_2)x}}{u_2+u_1\, e^{-(u_1+u_2)x}} dG(x)} .$$

Proof : $E(S_0|S_2=x)$ and $E(S_0^2|S_2=x)$ follow by differentiating $S^*_0(s,x)$ from equation (3.17) with respect to s and using equation (2.3). Unconditioning with respect to S_2 and rearranging yields $E(S_0)$ and $E(S_0^2)$. Similarly, equation (3.16) can be used to obtain $E(C|S_2=x)$ and $E(C^2|S_2=x)$. Unconditioning with respect to S_2 and rearranging yields $E(C)$ and $E(C^2)$. $R(S_0)$ and $R(C)$ are obtained from equations (2.6) and (2.7), respectively. *Q.E.D.*

In the following remark we present a probabilistic argument to the derivation of theorem 3.2.

Remark : (probabilistic argument)

From the Poisson property of interruptions, it follows that the total expected duration of type t interruptions during the customer's service is given by $v_t E(S_e)E(I_t)$, $t \in T$. This property holds inspite of the dependency of the effective service time on the stream of *prt* interruptions.

Consider the completion time given that the initial customer's service $S_0 = x$. The expected completion time is the addition of the expected effective service time and the expected time spent in all interruptions during the customer's service. Hence, we have

$$E(C|S_0=x) = E(S_e|S_0=x)\,[1 + v_s E(I_s) + v_t E(I_t) + v_d E(I_d)] \qquad (3.22)$$

Unconditioning with respect to S_0 yields equation (3.20).

A random observer will find the server actually servicing the customer with probability A, and will find an interruption of type t with probability A_t, $t \in T$. Furthermore if the expected residual time of S_e is $R(S_e)$ then its contribution to the expected residual time of C is expanded by a factor A^{-1} due to interruptions that arrive during the residual effective service time. Therefore, the expected residual time of C given that $S_0=x$, may be written as follows

$$\begin{aligned} R(C|S_0=x) = {} & A\,[\frac{R(S_e|S_0=x)}{A}] \\ & + A_s\,[R(I_s) + \frac{R(S_e|S_0=x)}{A}] \\ & + A_d\,[R(I_d) + E(C)] \\ & + A_t\,[R(I_t) + E(C|S_0=x)] \end{aligned} \qquad (3.23)$$

Substituting for the expected residual time from the relation in (2.4) and using equation (3.22), we get

$$E(C^2|S_0=x) = \frac{E(S_e^2|S_0=x)}{A}$$

$$+ 2\, v_e\, E(I_e)\,[R(I_e)\, E(S_e|S_0=x) + \frac{E(S_e^2|S_0=x)}{A}]$$

$$+ 2\, v_d\, E(I_d)\, E(S_e|S_0=x)\,[R(I_d) + E(C)]$$

$$+ 2\, v_i\, E(I_i)\,[R(I_i)\, E(S_e|S_0=x) + \frac{(E(S_e|S_0=x))^2}{A}] \tag{3.24}$$

Unconditioning with respect to S_0 and deviding by $E(C)$ from equation (3.20) yields equation (3.21).

The following two corollaries specialize to the cases where either *prd* or *pri* interruptions are present, but not both simultaneously.

Corollary 3.2 : If no *pri* interruptions are present then the first moment and the expected residual time of C are given by

$$E(C) = \frac{E(S_e)}{A}\,, \tag{3.25}$$

and

$$R(C) = R(S_e) + A_e\,[R(I_e) + \frac{R(S_e)}{A}] + A_d\,[R(I_d) + E(C)]\,. \tag{3.26}$$

where

$$E(S_e) = \frac{(1 - \tilde{S}(v_d))}{v_d\, \tilde{S}(v_d)}\,,$$

and

$$R(S_e) = E(S_e) - \frac{1}{E(S_e)}\frac{dE(S_e)}{dv_d}\,.$$

Proof : Follows directly from theorem 3.2 by letting v_i=0 and evaluating $E(S_e)$ and $R(S_e)$ accordingly. *Q.E.D.*

Corollary 3.3 : If no *prd* interruptions are present then the first moment and the expected residual time of C are given by

$$E(C) = \frac{E(S_e)}{A} , \tag{3.27}$$

$$R(C) = R(S_e) + A_p \left[R(I_p) + \frac{R(S_e)}{A}\right] + A_i \left[R(I_i) + E(C)\right] , \tag{3.28}$$

where

$$E(S_e) = \frac{(\tilde{S}(-u_i) - 1)}{u_i} ,$$

and

$$R(S_e) = E(S_e) - \frac{1}{E(S_e)}\left[\frac{dE(S_e)}{du_i} - \frac{\tilde{S}(-2u_i) - (\tilde{S}(-u_i))^2}{u_i^2}\right] .$$

Proof : Follows from theorem 3.2 by coditioning on the customer's service time $S(=S_e)=x$. We get

$$E(C|S=x) = \frac{E(S_e|S=x)}{A} ,$$

and

$$R(C|S=x) = R(S_e|S=x) + A_p \left[R(I_p) + \frac{R(S_e|S=x)}{A}\right] + A_i \left[R(I_i) + E(C|S=x)\right]$$

Unconditioning with respect to the customer's service time S and evaluating $E(S_e)$ and $R(S_e)$, using equations (3.17) and (3.5), yields equations (3.27) and (3.28). *Q.E.D.*

So far we have been able to derive the *LST* of the effective service time and the completion time associated with a customer's service. We also obtained useful relations for the first moment and the expected residual time.

It should be noted that although we have considered one source of interruptions of each type, the same results hold in the case where there are more than one source of interruptions of each type. For example, the terms $u_q \tilde{\Gamma}_q(s)$ are replaced by $\sum_{i \in Q_q} u_i \tilde{\Gamma}_i(s)$, $q \in \{p,i,d\}$, in equation (3.14) for $\tilde{C}(s)$.

It remains to note that the completion times associated with successive customers are independent and identically distributed random variables. This is an important remark for the analysis of the steady-state average number of customers in the system which is considered in the next section.

4. The steady-state average number of customers

In [4] Gaver derived the steady-state distribution of the number of customers in the system in the case where only a single type of interruption sources are present. Similar derivation holds in the case with mixed types of interruption sources. In this section we show that if we are only interested in the steady-state average number of customers then it can be derived by rather simple probabilistic arguments [7,8,14].

Unlike the analysis of the completion time, it is important to distinguish active and independent interruptions. We consider the mixture of all types of interruptions, active and independent, with one or more sources of each type. We show that when all interruptions are of the active-preemptive type, the steady-state average number of customers is determined by the Pollaczek-Khintchine formula.

It is convenient at this point to introduce the concept of a virtual customer associated with each real customer; its service time is identical to the completion time that we studied in section 3. This implies that a virtual customer leaves the system only when its "own" postponable interruptions, if any, are serviced, while a real customer leaves the system just before servicing the postponable interruptions.

Denote by P_I the expected fraction of time the system is idle (i.e. neither customers nor interruptions are in the system). From the Poisson property of interruptions, it follows that the expected fraction of time the system is servicing independent interruptions from source i that start busy periods is given by $P_i = P_I v_i E(I_i)$, $i \in ind$. This is also the expected number of independent interruptions from source i, that start busy periods, in service. P_I can be determined from the normalizing relation

$$\lambda E(C) + P_I + \sum_{i \in ind} P_i = 1$$

It follows that

$$P_I = (1 - \lambda E(C))\,(1 + \sum_{i \in ind} v_i\, E(I_i))^{-1}\,, \tag{4.1}$$

and

$$P_i = v_i\, E(I_i)\,(1 - \lambda E(C))\,(1 + \sum_{i \in ind} v_i\, E(I_i))^{-1} \tag{4.2}$$

From equation (4.1) it is obviuos that the system is stable if $\lambda E(C) < 1$.

The following theorem gives the general result.

Theorem 4.1 : The steady-state average number of customers in the M/G/1 queue with mixed types of interruptions is given by

$$\begin{aligned} N = {} & \lambda\,(1 + \sum_{i \in ind} v_i\, E(I_i))^{-1} \sum_{i \in ind} v_i\, E(I_i)\, R(I_i) \\ & + (1 - \lambda E(C))^{-1}\, \lambda^2\, E(C)\, R(C) + \lambda E(C) \\ & - \lambda E(S_o) \sum_{i \in dep} v_i\, E(I_i) \end{aligned} \tag{4.3}$$

with $E(C)$ and $R(C)$ as given by equations (3.20) and (3.21), respectively.

Proof : Let W' be the mean response time of a virtual customer and denote by N' the steady-state average number of virtual customers seen by an arrival; it is identical to the time average since arrivals are Poisson [16]. The expected number of virtual customers in service seen by an arrival is $\lambda E(C)$ and hence the expected number of waiting virtual customers is $(N' - \lambda E(C))$.

The mean response time W' of a virtual customer is made up of the following terms

$$W' = W'_1 + W'_2 + W'_3 + W'_4 \ .$$

W'_1 is the expected remaining service of independent interruptions, that start busy periods, found in service

$$W'_1 = \sum_{i \in ind} P_i \, R(I_i)$$

with P_i from equation (4.2).

W'_2 is the expected remaining service of virtual customers found in service

$$W'_2 = \lambda E(C) \, R(C)$$

with $E(C)$ and $R(C)$ from equations (3.20) and (3.21).

W'_3 is the expected time spent in servicing the waiting virtual customers found in the system

$$W'_3 = (N' - \lambda E(C)) \, E(C) \ .$$

W'_4 is the expected service time of the arriving virtual customer

$$W'_4 = E(C) \ .$$

Substituting for N' from Little's formula ($N' = \lambda W'$) yields an explicit expression for W'

$$W' = (1 + \sum_{i \in ind} v_i \, E(I_i))^{-1} \sum_{i \in ind} v_i \, E(I_i) \, R(I_i)$$
$$+ (1 - \lambda E(C))^{-1} \lambda E(C) \, R(C) + E(C) \ .$$

The mean response time W of a real customer is determined by subtracting the expected duration of the postponable interruptions, if any

$$W = W' - E(S_a) \sum_{1 \le p \le P} v_i E(I_i)$$

The steady-state average number of real customers N follows from Little's formula ($N = \lambda W$). *Q.E.D.*

The following corollary specializes the result to the case where all interruptions are of the active-preemptive type.

Corollary 4.1 : In the special case whre all interruptions are of the active-preemptive type, the steady-state average number of customers is given by the Pollaczek-Khintchine formula

$$N = \lambda E(C) + (1 - \lambda E(C))^{-1} \lambda^2 E(C)\, R(C) \qquad (4.4)$$

with $E(C)$ and $R(C)$ as determined by equations (3.20) and (3.21), respectively.

Proof : Follows directly from theorem 4.1. *Q.E.D.*

This result could be anticipated since when all interruptions are active-preemptive, the system can be viewed as an M/G/1 queue with the real customer's service time replaced by the virtual customer's service time (or equivalently, the completion time).

5. Application to the Modeling of Checkpointing and Recovery

In this section we consider the application of the developed theory to the modeling of checkpointing and recovery in a transactional system.

5.1. Introduction

Periodical checkpointing is a common technique for maintaining the integrity of information in database systems subject to failures. During a checkpoint a copy of the system files is saved in a secondary storage device. When a failure occurs, a recovery action is initiated. It starts with reloading a copy of the system files that were saved at the last checkpoint into primary memory. This is followed by reprocessing all those transactions that have been processed since the last checkpoint. The recovery action brings the system to its correct status as just before the failure. The system is unavailable for processing new transactions during checkpointing and recovery operations. Too frequent checkpoints cost much time in making unnecessary copies, and too distant checkpoints cost much time in reprocessing after failures. Therefore it is of much interest to determine the checkpointing frequency that optimizes certain performance measure such as the system availability (the fraction of time the system is available for processing new transactions), or the mean response time of a transaction.

Although many authors have studied models to determine the system availability, only a few of them considered the queueing aspects in order to compute the mean response time [1,2,3,5,6,12,13].* In most of these models the mean recovery period is assumed to be proportional to the mean available time between checkpoints. Gelenbe and Derochette [5] considered an M/M/1 system with two sources of independent Poisson interruptions; namely, checkpointing and failure-recovery (indeed, this is a special case of Gaver's model). Nicola and Kylstra [12] extended the model to include state-dependent parameters and finite waiting room. In [13] they considered a model in which checkpoints are performed after a specified number of completed transactions (load-dependent

* we restrict our references to those that include queueing analysis

strategy). In [6] Gelenbe relaxed the exponential assumption and considered a general distribution of available time between checkpoints. Baccelli [1] continued the work of Gelenbe and derived numerical algorithm for the computation of the mean response time. Duda [3] further generalized the model to the GI/G/1 system. His analysis is based on a diffusion approximation approach, and it implies a preemptive-resume type of interruptions. In [2] Baccelli and Znati considered an M/G/1 system with two types of independent Poisson interruptions; namely, preemptive-resume (for checkpointing) and preemptive- repeat-different (for recovery).

In this section we consider an M/G/1 system. Checkpoints may occur when the system is idle or when it is processing. If the system is processing then the checkpoint operation is postponed until the end of the transaction being processed. Therefore checkpoints are modelled as independent-postponable (*ipsp*) Poisson interruptions. Checkpoint durations are independent and of identical general distribution. Failures may occur only when the system is processing. A recovery operation preempts the transaction being processed. When recovery is completed the preempted transaction is reprocessed. Therefore recoveries are modelled as active-preemptive-repeat-identical (*aprt*) Poisson interruptions.

We propose a more accurate recovery model than those considered in previous queueing models. It is assumed that a random number of transactions should be reprocessed in a recovery operation. The distribution of this number is identical to that of the random number of processed transactions between failure occurrence and the last checkpoint. This yields independent recovery durations of identical distribution. The mean and variance of this distribution can be expressed as functions of the checkpointing frequency, as will be shown.

5.2. Performance Measures

First we define the parameters and the random variables associated with the model described in section 5.1. Consider the M/G/1 system in which transactions arrive at rate λ. They are processed according to the FCFS discipline. The processing time of a transaction, S, is a random variable of general distribution; its LST is $S^*(s)$. Checkpoints are *tpsp* Poisson interruptions. They are perfomed at rate α. Checkpoint duration, B, is a random variable of general distribution. Failures are *apri* Poisson interruptions. They occur at rate γ. Recovery duration, Q, is a random variable of general distribution.

We proceed to compute the first moment and the expected residual time of the effective service time, S_e, and the completion time, C, as defined in section 2. Since there are no *prd* interruptions, we can use the results of corollary 3.3. It follows that

$$E(S_e) = \frac{(S^*(-\gamma) - 1)}{\gamma} \tag{5.1}$$

$$R(S_e) = \frac{\frac{(S^*(-2\gamma) - S^*(-\gamma))}{\gamma} - \frac{dS^*(-\gamma)}{d\gamma}}{(S^*(-\gamma) - 1)} \tag{5.2}$$

Let A, A_B, A_Q, be the expected fraction of completion time spent by the system in processing the transaction, in checkpoints and in recoveries, respectively. Then from equations (3.18) and (3.19) we have

$$A = (1 + \alpha\, E(B) + \gamma\, E(Q))^{-1} \ . \tag{5.3}$$

$$A_B = \alpha\, E(B)\, A \ . \tag{5.4}$$

and

$$A_Q = \gamma\, E(Q)\, A \ . \tag{5.5}$$

Equations (3.27) and (3.28) give for $E(C)$ and $R(C)$ the following

$$E(C) = \frac{E(S_e)}{A} \, . \qquad (5.6)$$

and

$$R(C) = R(S_e) + A_B \left[R(B) + \frac{R(S_e)}{A}\right] + A_Q \left[R(Q) + \frac{E(S_e)}{A}\right] . \qquad (5.7)$$

Let P_I be the probability that the system is idle; it is determined from equation (4.1). The system availability, A^* is given by

$$\begin{aligned} A^* &= \lambda\, E(S_e) + P_I \\ &= \frac{1 - \lambda E(S_e)\, \gamma E(Q)}{(1 + \alpha E(B))} \, . \end{aligned} \qquad (5.8)$$

The mean response time of a transaction, W, is determined from theorem 4.1 and Little's formula

$$W = \frac{\alpha E(B)\, R(B)}{(1 + \alpha E(B))} + \frac{\lambda E(C)\, R(C)}{(1 - \lambda E(C))} + E(C) - \alpha E(B)\, E(S_e) \, . \qquad (5.9)$$

The first term is the contribution of checkpoints that start busy periods and the last term is due to the postponement of checkpoints. The middle term correspond to the Pollaczek-Khintchine formula (see equation (4.4)).

For the optimization of performance measures we need to establish a model for the dependence of the recovery duration, Q, on the checkpointing rate, α. It can be shown [6], for exponential available time interval between checkpoints and Poisson failure occurrences, that the available time interval, F, between failure occurrence and the last checkpoint is exponentially distributed with a mean α^{-1}. We assume that the completion process of transactions (in the available time) is Poisson with rate $\frac{\lambda}{A}$. Let NF be the random number of completed transactions between failure occurrence and the last checkpoint; its mean and variance can easily be determined, and are given by

$$E(NF) = \frac{\lambda}{A} E(F)$$

$$= \frac{\lambda}{\alpha A^*} \tag{5.10}$$

and

$$var(NF) = (\frac{\lambda}{A^*})^2 var(F) + \frac{\lambda}{A^*} E(F)$$

$$= \frac{\lambda}{\alpha A^*}(1 + \frac{\lambda}{\alpha A^*}) \tag{5.11}$$

where var(.) denotes the variance of a random variable.

Further, we assume that a random number corresponding to NF and of identical distribution is to be reprocessed in a recovery operation. This yields a recovery duration, Q, of mean and variance given by [15]

$$E(Q) = E(NF)\, E(S)$$

$$= \frac{\lambda}{\alpha A^*} E(S) \ . \tag{5.12}$$

$$var(Q) = E(NF)\, var(S) + var(NF)\, (E(S))^2$$

$$= \frac{\lambda}{\alpha A^*}[E(S^2) + \frac{\lambda}{\alpha A^*}(E(S))^2] \ . \tag{5.13}$$

The expected residual time $R(Q)$ follows

$$R(Q) = R(S) + \frac{\lambda\, E(S)}{2\, \alpha\, A^*} \ . \tag{5.14}$$

Substituting for A^* from equation (5.8) in equation (5.12), we can solve for $E(Q)$ and A^*.

The optimization of performance measures with respect to the checkpointing rate, α, can be carried out analytically or numerically after substituting for $E(Q)$ and $R(Q)$ from equations (5.12) and (5.14). In general the maximization of the system availability, A^*, and the minimization of the mean response time of a transaction, $\bar{W}$, yield different values for the optimum checkpointing rate [5,12].

6. Conclusions

We have defined the effective service time and related it to the completion time associated with a cuctomer's service in a single server with mixed types of Poisson interruptions. A derivation of the *LST* of the completion time is presented. As an intuitive alternative, we have demonstrated a probabilistic argument to express the first moment and the expected residual time of the completion time in terms of those of the effective service time and the interruptions. Rigorous proofs are straightforward, though very lengthy and uninteresting for presentation. The moments of the completion time are used to obtain the steady-state average number of customers in an M/G/1 system with mixed types of interruptions. When all interruptions are active-preemptive, the average number of customers is given by the Pollaczek-Khintchine formula with the customer's service time replaced by the completion time.

The theory developed is relevant in many systems modeling applications. One such application; namely, the modeling of checkpointing and recovery in a transactional database system is considered. The theory enables us to model the interaction between transaction-processing and the two types of interruptions more realistically than in previous work.

ACKNOWLEDGEMENT

Special thanks go to Professor F. J. Kylstra and to Dr. E. A. van Doorn for useful comments on earlier drafts of this paper that included the basic results based on intuitive probabilistic arguments. Fruitful cooperation with Professor K. S. Trivedi and Professor V. G. Kulkarni has supported me to argument the paper with rigorous proofs. I indeed value their cooperation.

REFERENCES

[1] Baccelli, F., Analysis of a service facility with periodic checkpointing, Acta Informatica, vol. 15, no. 1, 1981, pp. 67-81.

[2] Baccelli, F. and T. Znati, Queueing algorithms with breakdowns in database mdelling, in "Performance 81", Ed. by F. J. Kylstra, North Holland, 1981, pp. 213-231.

[3] Duda, A., Performance analysis of the checkpoint- rollback-recovery system via diffusion approximation, in "Proceedings of the International Workshop on Applied Mathematics and Performance/Reliability Models of Computer/Communication Systems", Ed. by G. Iazeolla and S. Tucci, North Holland, 1983, pp. 387-399.

[4] Gaver, D. P., A waiting line with interrupted service, including priorities, J. Royal Stat. Soc., series B24, 1962, pp. 73-90.

[5] Gelenbe, E. and D. Derochette, Performance of rollback recovery systems under intermittent failures, Commun. ACM, vol. 21, 1978, pp. 493-499.

[6] Gelenbe, E., On the optimum checkpoint interval, J. ACM, vol. 26, 1979, pp. 259-270.

[7] Gelenbe, E. and I. Mitrani, Analysis and synthesis of computer systems, Academic Press, 1980.

[8] Green, L., A limit theorem on subintervals of interrenewal times, Operations Research, vol. 30, 1982, pp. 210-216.

[9] Jaiswal, N. K., Priority queues, Academic Press, 1968.

[10] Kulkarni, V. G., V. F. Nicola and K. S. Trivedi, On modeling the performance and reliability of multi-mode computer systems, Submitted, 1984.

[11] Kulkarni, V. G., V. F. Nicola and K. S. Trivedi, A unified model for performance and reliability of fault-tolerant systems, Submitted, 1984.

[12] Nicola, V. F. and F. J. Kylstra, A markovian model, with state-dependent parameters, of a transactional system supported by checkpointing and recovery strategies, in "Messung, Modellierung und Bewertung von Rechensystemen, Ed. by P. J. Kuhn and K. M. Schulz, IFB 61, Springer Verlag, 1983, pp. 189-206.

[13] Nicola, V. F. and F. J. Kylstra, A model of checkpointing and recovery with a specified number of transactions between checkpoints, in "Performance 83", Ed. by A. K. Agrawala and S. K. Tripathi, North Holland, 1983, pp. 83-100.

[14] Oliver, R. M., An alternate derivation of the Pollaczek-Khintchine formula, Operations Research, vol. 12, 1964, pp. 158-159.

[15] Trivedi, K. S., Probability and statistics with reliability, queueing and computer science applications, Prentice Hall, Englewood Cliffs, N.J., 1982.

[16] Wolff, R. W., Poisson arrivals see time averages, Operations research, vol. 30, 1982, pp. 223-231.