Contribution to the 6th International Conference on Computer Communication (ICCC), September 1982, London, GB.

Clocked Event Transfer Protocol in Distributed Processing Systems – a Performance Analysis

P Tran-Gia, H Jans
University of Siegen, Federal Republic of Germany

In modern real-time processing systems the intelligence is distributed among a number of individual processors operating in modes of functional or load sharing. Communication between distributed processors is often organised in the form of message interchanging according to an event transfer protocol. In this paper, a performance analysis is given for the commonly used clocked event transfer scheme, where a two-level queueing system is investigated. For the analysis a dimension-reduction of the two-dimensional description of the imbedded Markov chain is developed. Numerical results for dimensioning purposes are discussed, especially for event delay characteristics under different traffic conditions, clock intervals, and buffer sizes. Finally, the distribution function of the prequeueing delay is presented for the two event transfer disciplines first-in first-out and random.

1. INTRODUCTION

In distributed processing systems, especially in communication applications, the rates of real-time events, which have to be handled by several processors, are very high. These events are generated by peripheral devices or users, or are caused by inter-processor communications.

Fig. 1 illustrates a basic control structure of a multi-processor system where events are preprocessed by peripheral controllers (process level A) and stored in event buffers as valid events in a logical sense. They have to be transferred to another processor (process level B) according to an *event transfer protocol* for further processing. An optimal transfer protocol for this high event rate allows to increase the throughput and to optimize delay characteristics. The most commonly used event transfer protocol is the clocked scheme, whereby events are transmitted between different processors in a batch-wise manner at a scheduled time, initialised by a real-time clock. The event transfer protocol includes the initialisation of transfer, transmission control, acknowledgment etc..

There are a number of studies [1,2,5,6,7] which investigate the performance of event sampling and transfer schemes by means of *one-level* basic queueing models with scheduled batch arrivals. Some of them [1,4] deal with models having infinite waiting capacity. [5] considers dimensioning aspects for models with batch input and finite queue capacity, whereby the blocking probability is calculated for events and batches, which are partly or fully rejected according to the number of free waiting places. Several event transfer schemes with batch arrivals and overhead are discussed in [6], in which a number of events in a batch is considered to be lost when the actual batch is larger in size than the actual number of free waiting places.

In order to analyse the system depicted in Fig.1, a *two-level* queueing system is investigated in this paper, where a clocked event transfer protocol is considered.

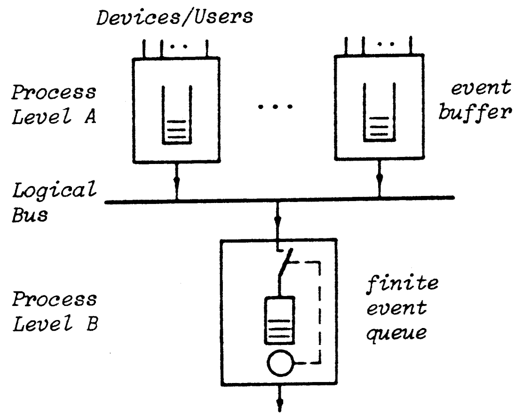

Fig.1 Basic Control Structure for a Distributed Processing System

In principle an approximation of this model is possible using two separate one-level queues. This method implies an independence assumption for the two queues, the *primary* queue and the *secondary* queue for events to be transferred (c.f. Fig. 2). Because of this assumption the accuracy of the one-dimensional approximation depends very strongly on system parameters. Therefore, an exact analysis requires a *two-dimensional* model which will be presented in the next section. Fortunately, computational efforts can be reduced by the analytical method described in section 3.

2. MODELLING OF CLOCKED EVENT TRANSFER BETWEEN TWO PROCESSORS

The queueing model considered in this paper has the structure shown in Fig. 2. Event arrivals constitute a Poisson process with rate λ. This assumption is based on the observation that the incoming event stream is the superposition of offered traffic from a large number of different devices and users connected to the processor. Taking into account the different types of events and the tasks and programs they may activate, the service time T_{SER} for events is assumed to be negative exponentially distributed with mean $1/\mu$.

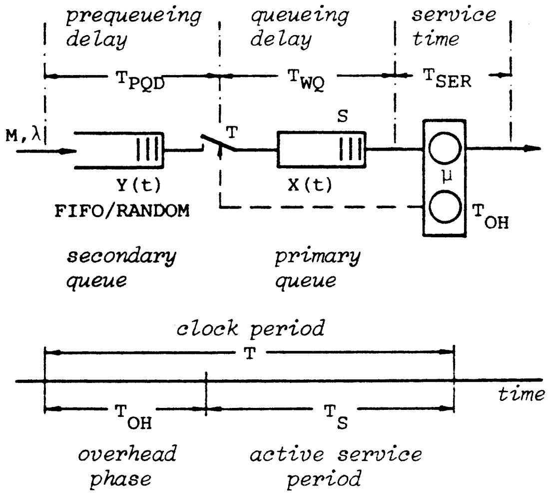

Fig.2 The two-level queueing system

Every event transfer activity which is controlled by the processor is usually performed by the same I/O task and has approximately the same run-time during which the processor is not available for event processing. Therefore, the whole clock period T consists of two parts: the overhead time T_{OH} and the active service period T_S (c.f. Fig. 2). It should be noted here that during the active service period the server is available for event processing but not necessarily busy. In this paper the clock period is chosen to be constant. This is often the case in real systems where the I/O phases are activated by a real-time clock. Another reason for this choice of T can be found in [6]. It is shown there that a well dimensioned clocked scheme is relatively robust with respect to event traffic intensity. The primary queue is considered to have the *finite capacity* S. At a clock instant when the actual batch (i.e. events existing in the secondary queue prior to the clock instant) is larger in size than the number of free waiting places in the primary queue, all free positions will be filled and the remaining events must wait for retrial until the next clock instant.

Each batch consists of two parts, the fresh part and the reattempt part. All arrivals during the clock period form the fresh part; the reattempt part contains events which have been rejected at the previous clock instant.

The total sojourn time of an event in the whole system is composed of three components:

- The prequeueing delay T_{PQD}, i.e. the waiting time in the secondary queue for events to be sent.
- The queueing delay T_{WQ}, i.e. the waiting time in the primary queue, including overhead periods.
- The service time T_{SER}

T_{PQD}, T_{WQ} and T_D will be considered as random variables in the next section.

3. ANALYSIS OF THE TWO-LEVEL QUEUEING SYSTEM

In this section performance measures of the above described two-level queueuing system are presented. Subsection 3.1 discusses queue stability conditions and subsection 3.2 deals with the system state probabilities. Subsequently, system characteristics will be derived in 3.3.

3.1 Queue Stability

Since we consider the overhead time T_{OH} and a finite primary queue capacity S, the system is only stable under certain conditions derived below. For given values of T_{OH}, S, and *offered traffic intensity* ρ ($\rho = \lambda/\mu$) we can calculate a lower and an upper limit for the clock period T, for which the system is stable. The lower limit T_{MIN} is found using the fact that the active service time T_S in which the server is available must be long enough to serve all arriving events on average:

$$T_{MIN} = T_{OH}/(1 - \rho) \tag{3.1}$$

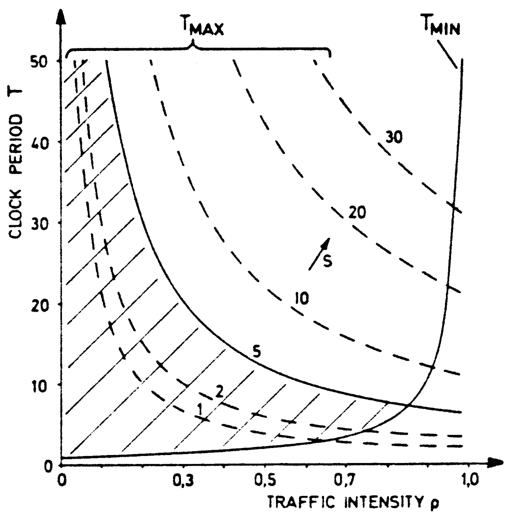

Fig.3 Queue Stability Conditions ($T_{OH}=1/\mu$)

On the other hand, if T becomes too long, the batch sizes are also very large on average and the finite queue is likely to be filled completely at each clock instant and tends to be empty before the next clock. The upper limit T_{MAX} is given by:

$$T_{MAX} = N/\lambda \tag{3.2}$$

where N = S + 1 is the *primary system size.*

The queueing system is stable for

$$T_{MIN} < T < T_{MAX} \tag{3.3}$$

The queue stability condition (3.3) is illustrated in Fig. 3. For S = 5 the system is stable in the hatched field. The dashed lines show the upper limit for other values of S.

3.2 The Imbedded Markov Chain and System State Probabilities

In Fig. 4 we consider the two-dimensional random process $\{X(t), Y(t), t \in (0,T)\}$ more closely

$$\Pi(x,y,t) = \Pr\{X(t)=x,\ Y(t)=y,\ t \in (0,T)\} \qquad (3.4)$$

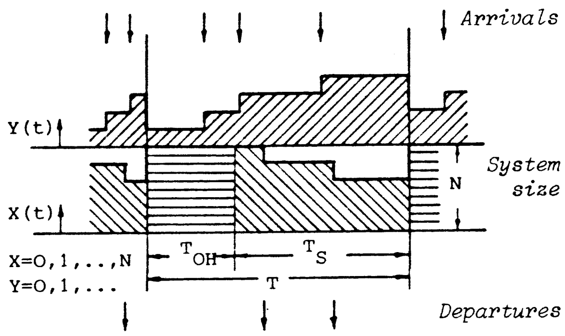

Fig.4 The Two-dimensional Random Process

The system can be analysed using the well-known technique of the imbedded Markov chain. It is here convenient to choose the time-epochs just after the event transfer instants as regeneration points ($t=0^+$), at which the two queues can be considered as connected. This argument allows us to reduce the two-dimensional description of the process into an one-dimensional description at regeneration points. The relationship between the two- and the one-dimensional description of state probabilities is given by

$$P^+(x+y)=\Pi(x,y,0^+) \qquad (3.5)$$

where $\quad \Pi(x,y,0^+)=0$ for $x \neq S+1$ and $y \neq 0$.

Observing two consecutive event transfer periods n and n+1 with state probabilities $\Pi_n(x,y,t)$ and $\Pi_{n+1}(x,y,t)$, $t \in (0,T)$, the main steps of the analysis can shortly be resumed as follows:

1. $\Pi_n(x,y,T^-)=f\{\Pi_n(x,y,0^+)\}$

$$x \neq 0:\ \Pi_n(x,y,T^-)= \sum_{i=x}^{N} \sum_{j=0}^{y} \Pi_n(i,j,0^+)\ d_{i-x}\ g_{y-j}$$

$$x = 0:\ \Pi_n(0,y,T^-)= \sum_{i=0}^{N} \sum_{j=0}^{y} \Pi_n(i,j,0^+) \sum_{\ell=i}^{\infty} d_\ell g_{y-j}$$

with $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ...(3.6)

$$d_\ell = \frac{(\mu T_S)^\ell}{\ell!} e^{-\mu T_S}\ ;\ g_m = \frac{(\lambda T)^m}{m!} e^{-\lambda T},$$

$$\ell, m = 0,1,...$$

2. $\Pi_{n+1}(x,y,0^+) = f\{\Pi_n(x,y,T^-)\}$

$$x=0,..,N-1;y=0:\ \Pi_{n+1}(x,0,0^+)=\sum_{i=0}^{x}\Pi_n(i,x-i,T^-)$$

$$x=N;y=0,1,..:\quad \Pi_{n+1}(N,y,0^+)=\sum_{i=0}^{N}\Pi_n(i,N+y-i,T^-)$$

...(3.7)

Eqns. (3.6) and (3.7) give us a relationship between state probabilities for two consecutive regeneration points. Using this consideration, the state probabilities at an arbitrary time can easily be derived. Therefore, in general, from a given starting probability vector, transient behaviour of the system can be investigated by means of the power method.

Under stationary conditions

$$\Pi_{n+1}(x,y,0^+) = \Pi_n(x,y,0^+) = P^+(x+y)$$

and after simple algebraic manipulations using (3.6) and (3.7) the system of difference equations for state probabilities can be obtained

$$P^+(k)= \sum_{i=0}^{N-1} P^+(i) \left[g_k \sum_{\ell=i}^{\infty} d_\ell + \sum_{x=1}^{\min(i,k)} d_{i-x}\ g_{k-x} \right]$$

$$+ \sum_{i=N}^{N+k} P^+(i) \left[g_{k+N-i} \sum_{\ell=N}^{\infty} d_\ell + \sum_{x=1}^{\min(N+k-i,N)} d_{N-x}\ g_{k+N-i-x} \right]$$

with $\sum_{x=a}^{b} (.) = 0$ for $b < a$. $\qquad$ ...(3.8)

Using eqn. (3.8), the state probabilities of the imbedded Markov chain can be obtained by means of an iteration method with over-relaxation, whereby a proper adaptive truncation of the state space must be provided.

3.3 Time-dependent State Probabilities and System Characteristics

In order to calculate the mean waiting time in the primary queue $E[T_{WQ}]$ and the mean prequeueing delay $E[T_{PQD}]$, it is convenient to use Little's law[8], for which it is necessary to know the mean queue lengths at an arbitrary point in time. Under stationary conditions, we only have to observe the system during *one* clock period ($t \in (0,T)$).

Using the terminology

$$P_k^X(t) = \Pr\{X(t)=k\}$$
$$P_k^Y(t) = \Pr\{Y(t)=k\} \qquad ...(3.9)$$

we obtain

$$\begin{cases} P_k^X(0) = P^+(k) & k=0,1,..,N-1 \\ P_N^X(0) = \sum_{i=N}^{\infty} P^+(i) & k=N \end{cases} \qquad (3.10a)$$

and

$$\begin{cases} P_0^Y(0) = \sum_{i=0}^{N} P^+(i) & k=0 \\ P_k^Y(0) = P^+(k+N) & k=1,2,..,\infty. \end{cases} \quad (3.10b)$$

Between two transfer instants, X(t) follows a pure death process ($T_{OH} \leq t < T$) and Y(t) follows a pure birth process ($0 \leq t < T$).

The mean numbers of events in the primary system E[X] and in the secondary queue E[Y] are given by

$$E[X] = \frac{1}{T} \int_0^T \sum_{k=1}^{N} k\, P_k^x(t)\, dt$$

$$= \frac{T_{OH}}{T} \sum_{i=1}^{N} i P_i^x(0)$$

$$+ \frac{1}{\mu T} \sum_{i=1}^{N} P_i^x(0) \sum_{k=1}^{i} k \sum_{j=i-k+1}^{\infty} d_j \quad (3.11)$$

$$E[Y] = \frac{1}{T} \int_0^T \sum_{k=1}^{\infty} k\, P_k^Y(t)\, dt$$

$$= \sum_{k=1}^{\infty} k\, P_k^Y(0) + \frac{\lambda T}{2} = E[Y(0)] + \frac{\lambda T}{2}. \quad (3.12)$$

Using Little's law, the mean waiting time in the primary queue $E[T_{WQ}]$ and the mean prequeueing delay $E[T_{PQD}]$ can be given as follows

$$E[T_{WQ}] = \frac{E[X]}{\lambda} - \frac{1}{\mu} \quad (3.13)$$

and $$E[T_{PQD}] = \frac{E[Y]}{\lambda}. \quad (3.14)$$

The mean total delay of events in the system is

$$E[T_D] = E[T_{WQ}] + E[T_{PQD}]. \quad (3.15)$$

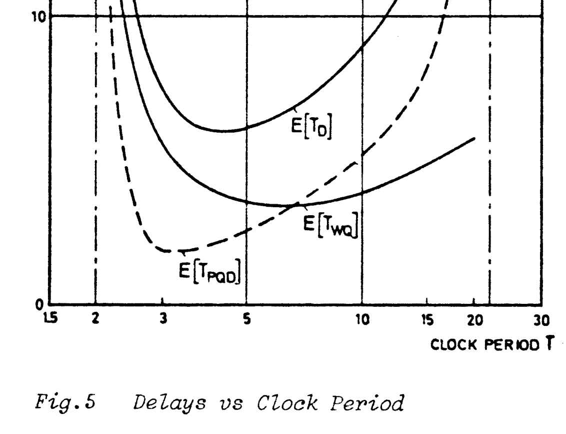

Fig.5 Delays vs Clock Period

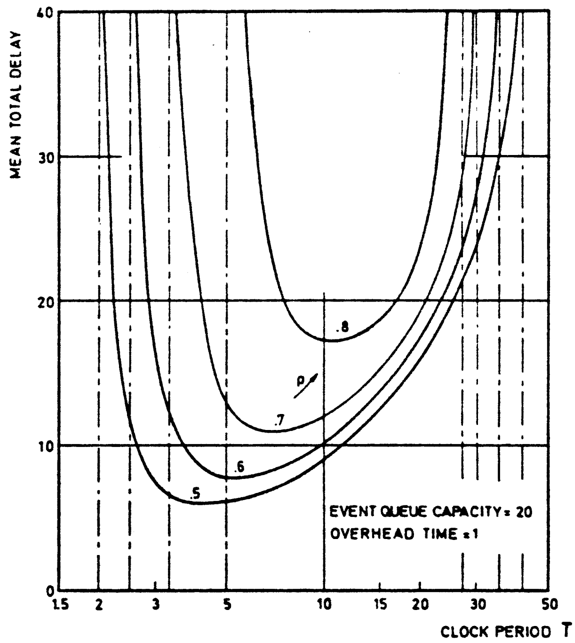

Fig.6 Total Delay vs Clock Period

4. RESULTS AND DISCUSSION

In this section numerical results are presented concerning the mean delays for events in the whole system using a clocked event protocol. All values are normalised by $E[T_{SER}] = 1/\mu = 1$. Fig. 5 depicts the different mean value of delays as a function of the clock period T. By the chosen parameters ($\rho = .5$, S = 10, $T_{OH} = 1$) the system is stable for $2 < T < 22$ (c.f.eqn. 3.3). It is clearly shown that a minimum of the total delay for events exists as expected.

In Fig. 6 the total delay is plotted as a function of clock period T and an optimum choice for T can be defined for a given level of offered traffic intensity. The sensitivity of these optimum values has to be taken into account for dimensioning purposes. The best choice T = 5 for $\rho = 0.6$ will make the system unstable for $\rho \geq 0.8$, which occurs in overload situations.

The delay characteristics discussed here can be used for dimensioning purposes where the clock period T and the capacity S of the event queue have to be chosen for a given traffic range.

5. PREQUEUEING DELAY DISTRIBUTION FUNCTION

5.1 General Relations

For the following derivation, we consider the two different disciplines FIFO and RANDOM for event transfer from secondary into primary queue. These two strategies can be considered as marginal cases for real systems, whereby FIFO strategy is a more optimistic and RANDOM strategy a more pessimistic case. In order to investigate the distribution function of the prequeueing delay $F_{PQD}(t)=Pr\{T_{PQD} \leq t\}$ in both cases the fate of a *test customer (t.c.)* in a fresh batch is considered (Fig. 7).

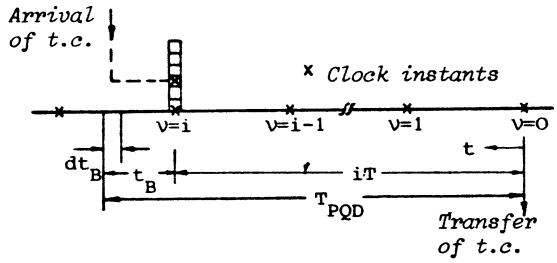

Fig.7 The Prequeueing Delay

As shown in Fig. 7 the prequeueing delay T_{PQD} consists of two components : the delay from arrival until the next clock instant (t_B) and a number i of clock periods (i.T), i=0, 1,2,.. The probability that the prequeueing delay varies between $iT+t_B-dt_B$ and $iT+t_B$ is

$$Pr\{iT+t_B-dt_B < T_{PQD} \leq iT+t_B\} = w(i,t_B)\cdot\frac{dt_B}{T}$$

with $\qquad 0 \leq t_B \leq T,\ i \geq 0 \qquad (5.1)$

dt_B/T :probability that the *t.c.* arrives in dt_B

$w(i,t_B)$:weighted sum over all possible positions in system the *t.c.* can take, who
- arrives t_B before the clock instant
- will be delayed i clock periods.

By integration we obtain from eqn. (5.1)

$$Pr\{iT<T_{PQD}\leq iT+t_B\} = \frac{1}{T}\int_0^{t_B} w(i,\tau)\cdot d\tau \qquad (5.2)$$

and for the delay distribution

$$F_{PQD}(t) = Pr\{T_{PQD}\leq iT+t_B\}$$

$$= Pr\{iT<T_{PQD}\leq iT+t_B\} + Pr\{T_{PQD}\leq iT\}. \qquad (5.3)$$

Eqn. (5.3) can be solved recursively and after some algebraic manipulations we have finally

$$F_{PQD}(t) = Pr\{T_{PQD}\leq iT+t_B\}$$

$$= \frac{1}{T}\left[\sum_{\nu=0}^{i-1}\int_0^T w(\nu,\tau)\cdot d\tau + \int_0^{t_B} w(i,\tau)\cdot d\tau\right] \qquad (5.4)$$

where $i=[t/T]^-$, is the largest integer less than t/T.
$i=0,1,2...,\ 0\leq t_B\leq T,\ \sum_i^j(\cdot)=0$ for $j<i$.

5.2 RANDOM Event Transfer Discipline

In RANDOM case the events are chosen randomly for transfer at clock instants. Based on the stationary state distribution of the imbedded Markov chain we can calculate the probability B_O that the *t.c.* is in the rejected batch part at the first clock instant after his arrival. In this case he has to compete with all customers in the secondary queue at transfer instants (including arrivals during his waiting time). The probability that he will be delayed at those clock instants is denoted by B_1.

From B_O and B_1, which are determined using combinatorial arguments, we obtain the probabilities $w(i,t_B)$ as follows

$$w(O,t_B)=1-B_O\ ;\ w(i,t_B)=B_O(1-B_1)B_1^{i-1} \qquad (5.5)$$

The delay distribution function is found by eqn. (5.4) to

$$F_{PQD}=Pr\{T_{PQD}\leq iT+t_B\} = \sum_{\nu=0}^{i-1} w(\nu,t_B)+w(i,t_B)\cdot\left(\frac{t}{T}-i\right) \qquad (5.6)$$

The mean pre-queueing delay $E[T_{PQD}]$ and its coefficient of variation c_{PQD} can be found from eqn. (5.6) and can be written in terms of B_O and B_1

$$E[T_{PQD}] = T\left(\frac{1}{2}+\frac{B_O}{1-B_1}\right) \qquad (5.7)$$

$$c_{PQD} = \sqrt{\left(\frac{T}{E[T_{PQD}]}\right)^2\cdot\left(\frac{1}{3}+\frac{2B_O}{(1-B_1)^2}\right)-1}$$

5.3 FIFO Event Transfer Discipline

In FIFO case the events are transferred between the two queues in order of arrival. Based on the stationary state distribution of the imbedded Markov chain, the prequeueing delay depends *only* on the service process of those customers in front of the *t.c.*. Due to the clocked scheme and the finiteness of the primary queue, idle periods of the server can exist although there are customers still waiting for transfer.

The prequeueing delay distribution function for the FIFO case has been developed analogously to the RANDOM case. However, the extensive derivation of the formula should not be discussed here in more detail.

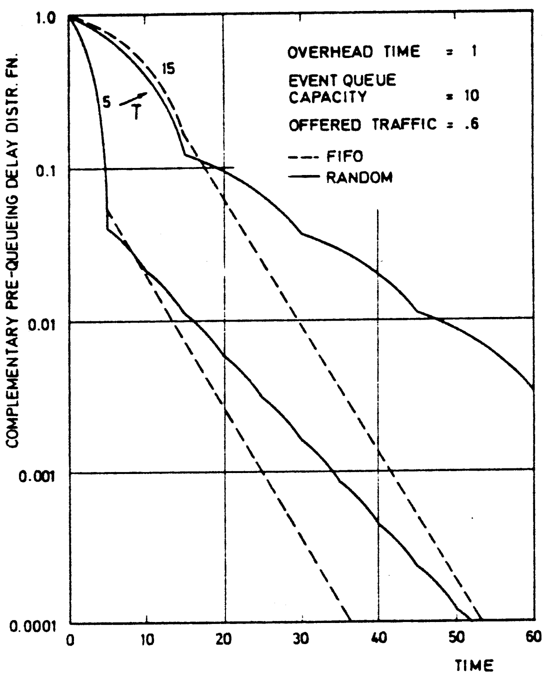

Fig.8 Comparison of $\overline{F}_{PQD}$ for FIFO and RANDOM Event Transfer Disciplines

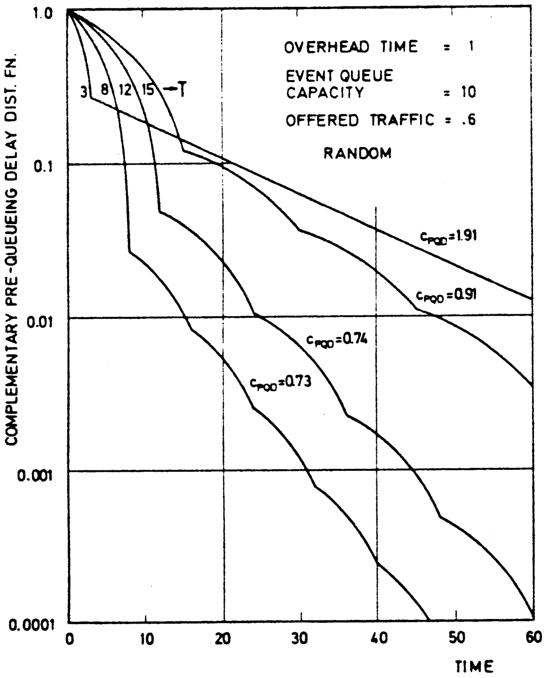

Fig.9 Complementary Prequeueing Delay Distribution Function for RANDOM case

5.4 Results and Comparison

In the following the complementary prequeueing delay distribution function $\bar{F}_{PQD}(t)$ will be discussed for the two considered event transfer disciplines FIFO and RANDOM. As expected the variance of T_{PQD} is higher for RANDOM case as shown by the gradient of the curves in Fig. 8, where $\bar{F}_{PQD}(t)$ is depicted for both disciplines and different values of T. Fig. 9 shows $\bar{F}_{PQD}(t)$ more closely in case of RANDOM for different values of the clock period T. Between T_{MIN} and T_{MAX}, a value of T can be found to optimize the coefficient of variation c_{PQD} for T_{PQD}. This argument can be taken into account together with the optimum choice of the mean total delay for different dimensioning conditions.

6. CONCLUSION

In this paper, a two-level queueing system has been developed and investigated, which models a clocked event transfer protocol between distributed real-time processing systems. In order to analyse the system, an exact one-dimensional description for the two-dimensional process is presented. From the steady state distribution system characteristics as event delays and mean values are derived and discussed. Finally, the prequeueing delay distribution function for events with FIFO and RANDOM transfer disciplines is investigated. The model is applicable for a wide range of distributed computing systems where high rate of real-time events have to be interchanged between processors and the response time is a critical factor. It is shown here that for dimensioning purposes, the queue stability and delay characteristics in overload situations must be taken into account to guarantee proper system performance.

ACKNOWLEDGEMENT

The authors would like to express their thanks to Prof. P.J. Kuehn for fruitful discussions during the course of this work.

REFERENCES

[1] Burke P.J.,"Delays in Single Server Queues with Batch Input".B.S.T.J.,23,830-833(1975).

[2] Chu W.W.,"Buffer Behaviour for Batch Poisson Arrivals and Single Constant Output" IEEE Trans.Comm.,18,613-618 (1970).

[3] Gross D. and Harris C.M.,"Fundamentals of Queueing Theory". Wiley, 1974.

[4] Langenbach-Belz M.,"Two-Stage Queueing System with Sampled Parallel Input Queues". Proc.7th ITC, Stockholm, 1973.

[5] Manfield D.R. and Tran-Gia P.,"Analysis of a Storage System with Batch Input Arising out of Message Packetisation". IEEE Trans. Comm., 30, 456-463(1982).

[6] Manfield D.R. and Tran-Gia P.,"Queueing Analysis of Scheduled Communications Phases in Distributed Processing Systems". Proc. 8th Symp.on Comp.Perf.Modelling,Meas. and Eval.,Amsterdam, 233-250(1981).

[7] Schwaertzel H.G.,"Serving Strategies of Batch Arrivals in Common Control Switching Systems". Proc.7th ITC, Stockholm, 1973.

[8] Little J.D.C.,"A Proof of the Queueing Formula L= λW". Operations Research, 9, 383-387(1961).

[9] Tran-Gia P. and Jans H.,"Delay Analysis of Clocked Event Transfer in Distributed Processing Systems". ORSA/TIMS Meeting, Detroit, April 1982.

P. Tran-Gia was born in Viet Nam and received the M.S. degree (Dipl.-Ing.) in EE from Stuttgart University, West-Germany, in 1977. In 1977 he joint SEL(ITT) in Stuttgart, where he was working in software development of digital switching systems. Since 1979 he has been Assistant Professor at the University of Siegen, Dept. of Communications. His current research activities are in the field of queueing theory and its application in performance analysis for telecommunication systems.

H. Jans was born in Niederkruechten, Germany. He received the Dipl.-Ing. degree in EE from Technical Univ. of Berlin, Germany, in 1975. Since then he has been Assistant Professor at the University of Siegen, Faculty of EE. His current research activities are in the field of queueing theory and its application in computer and communication systems.