Performance Analysis of Polling Mechanisms with Receiver Blocking

Thomas Raith, Phuoc Tran-Gia
Institute of Communications Switching and Data Technics
University of Stuttgart, Fed. Rep. of Germany

In distributed systems, loosely coupled units often communicate with each other by message passing through a communication subsystem, which consists of transmit and receive buffers interconnected by means of an interconection network with ring- or bus-oriented structure. The subject of this paper is the performance analysis of message interchanging mechanisms via an interconnection network, which operates according to a cyclic scheduling strategy. The investigation is based on a polling system with cyclic, nonexhaustive service, where the output process is further offered to the receiver having finite capacity buffers. An approximate analysis using a decomposition approach is developed. The method uses an alternating iteration of decomposed system parts, which combines the polling system analysis of the sending part with a feedback caused by the receiver blocking, whereby different bus scheduling modes in the case of receive buffer blocking are taken into account. Numerical results for performance characteristics like waiting time, cycle time, receive buffer blocking probability, etc. are presented, where the approximation accuracy and its dependency on system parameters are discussed by means of simulations.

1. INTRODUCTION

Autonomous loosely coupled processing units in systems with distributed control often communicate via message passing through an interconnection network. The units consist of, e.g., processing devices, local memories and communication buffers. They operate according to principles of load and function sharing and communicate with each other by addressed messages via their transmit and receive buffers, which are connected to a bus or ring architecture (c.f. Fig. 1). The allocation of the communication medium to the processing units is performed by a centralized or decentralized scheduling scheme, e.g. by setting a flag indicating the blocking condition.

The subject of this paper is the performance analysis of message interchanging mechamisms via a bus- or ring-oriented interconnection network. In general, the performance investigation presented here is based on an extended version of polling systems, where system constraints caused by the limitation of the receive buffer capacity are taken into account, in order to estimate its influence to the degradation of the system performance occuring in those systems. In the case of receive blocking, two scheduling schemes are investigated : either the scheduler proceeds to the transmit buffer of the succeeding unit and the message to be transmitted will wait further on within its transmit buffer until the next polling epoch, or the communication medium remains seized and the message waits in the transmit buffer until the blocked receive buffer becomes available.

Section 2 describes the queueing model, which is based on a multiqueue system with cyclic, nonexhaustive service, in which the output process will further be offered to the receiving part having finite capacity receive buffers. The approximate analysis will be dealt with in section 3. Numerical results will be presented in section 4, whereby the approximation accuracy and its dependency on system parameters will be discussed in comparison with simulations.

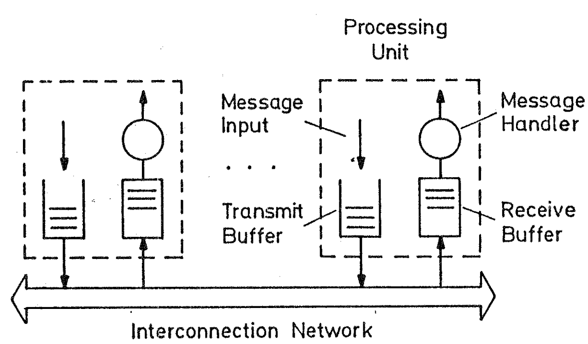


Fig. 1 : Communication in distributed systems via interconnection networks

2. MODELLING

2.1 General

In the literature, polling mechanisms with various service disciplines (e.g., exhaustive, nonexhaustive or gating) and interqueue schedules (e.g., cyclic or prioritized) have been the subject of numerous investigations. A survey on the analysis of multiqueue systems operating under different polling mechanisms was provided by Takagi and Kleinrock [3]. In most of the investigations the analysis was done by means of the imbedded Markov chain technique, where receiver constraints are not considered in the performance analysis. An approximation technique for cyclic queues with nonexhaustive service and general switchover time was developed by Kuehn [1]. This method has been extended in [2] where the output process of the cyclic server was calculated to investigate the performance of hierachical polling systems with feedback. Further investigations of some specific cyclic queueing systems were performed by Manfield [4] for systems with two-way traffic and priority cyclic service and by Morris and Wang [5] for multiqueue systems with multiple cyclic servers.

New Communication Services: A Challenge to Computer Technology
P. Kühn (ed.), Elsevier Science Publishers B.V. (North-Holland)

Raith [6] provided an analytical approach to deal with systems of multiple cyclic servers, where message blocking effects at receive buffer ports are taken into account. In the context of the analysis in [6], the blocking effect due to memory limitation is not investigated. In Tran-Gia and Raith [7,8] the effects of finite sending queue capacity in a nonexhaustive cyclic service polling system are investigated.

2.2 Performance Model of the Communication Mechanism

Fig. 2 depicts the queueing model of the communication structure described in section 1, which consists of a polling system where the receiver blocking effect is taken into account. In the model, a number g of transmit and receive buffers is considered. In order to investigate the receiver blocking, the receive buffers are assumed to have finite capacity S. The allocation of the communication medium to each particular connected processing unit is organized by means of a cyclic scheduling strategy, i.e., the server is cyclically allocated to one of the transmit buffers at a time. The transmit buffers are served in a nonexhaustive manner, according to which only one message can be transmitted per polling instant, if there are messages waiting for transfer. Furthermore, the processing units are assumed to operate in a full duplex mode, i.e., each of them can transmit and receive messages simultaneously.

The arrival process of messages at each processing unit is assumed to be Poissonian with rate λ. Receive buffers are served by the message handler according to a negative exponentially distributed service time T_p. The traffic between processing units in the system is assumed to be equally distributed. The transmission time T_h for messages via the interconnection network - from the ogirinating transmit buffer to the terminating receive buffer - is considered to be generally distributed. After finishing service in a unit, corresponding to the polling scheme assumed, the server will be allocated by the scheduler to the succeeding unit. If its transmit buffer is empty, the observed interscan period will be denoted as the switchover time T_u. This switchover time, which models all overheads spent and procedures performed by the scheduler to allocate the server cyclically, is assumed to have a general distribution function. In case of receiver blocking the server will proceed according to the following optional scheduling modes :

i) "Repeat" mode :
the scheduler proceeds to the following transmit buffer cyclically and the message waits further on within its transmit buffer until the next polling instant

ii) "Wait" mode :
the server remains occupied and the message has to wait until the blocked receive buffer gets idle and the destinating unit is able to accept the message.

It should be noted here that, without the receive buffer consideration, the model depicted in Fig. 2 corresponds to a single server polling system with nonexhaustive cyclic service [1].

The following symbols and random variables (r.v.) are used :

g	number of processing units
λ	arrival rate of messages offered to the transmit buffer
T_c	r.v. for the server cycle time
T_h	r.v. for the message transmission time in the sending part
T_u	r.v. for the switchover time
T_p	r.v. for the message processing time in the receiving part
S	capacity of a receive buffer
B	receive buffer blocking probability

The following notations are used for a r.v. T :

F(t)	probability distribution function
f(t)	probability density function
$\Phi(s)$	Laplace-Stieltjes-Transform of F(t)
E[T]	mean value of T
c	coefficient of variation of T

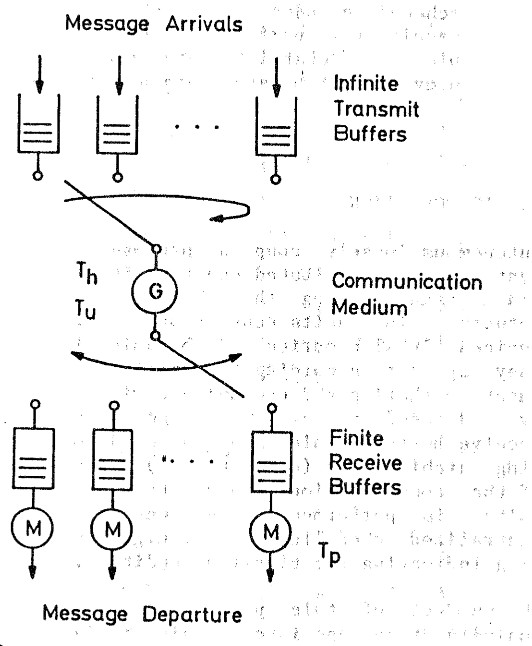


Fig. 2 : Performance model of the communication mechanism

3. ANALYSIS

3.1 Overview

In this subsection, an overview of an algorithm for an approximate analysis of a single server polling system is given, taking into account the feedback effect of the blocking due to finite capacity receive buffers. In general, the analysis is done in accordance with the method presented in [1]. However, some modifications must be provided in order to take into account the receive buffer blocking effect. The main idea in the calculation method presented in this paper is to combine the decomposition of the sending and receiving parts with an iteration approach. The analysis of the sending part will be described in subsection 3.2, followed by the receiving part analysis in subsection 3.3 and finally, the overall algorithm in subsection 3.4. System characteristics of interest are given in subsection 3.5.

3.2 Sending Part

3.2.1 Cycle Time Segment Analysis

The server cycle time is thought of to be divided into cycle time segments (c.f.[7,8]). The cycle time segment corresponding to an arbitrary transmit buffer is described by the random variable T_e, which is the time interval between the scanning epochs of two successive transmit buffers. The first two moments of the cycle time segment distribution function will be calculated. The moments depend on the probability of message blocking at the receive buffer and on the bus scheduling modes discussed above. The calculation method of the receive buffer blocking probability will be presented in subsection 3.3.

i) Scheduling mode "Repeat" :

According to the scheduling mode "Repeat" and the receive buffer blocking probability B, the Laplace-Stieltjes-Transform (LST) of the cycle time segment T_e corresponding to a transmit buffer can be given as :

$$\Phi_e(s) = \Phi_u(s)\{(1-P_0)\ ((1-B)\ \Phi_h(s) + B) + P_0\}, \tag{3.1}$$

where P_0 denotes the probability that no message is waiting for transfer in the transmit buffer, given by

$$P_0 = 1 - \lambda E[T_c]. \tag{3.2}$$

ii) Scheduling mode "Wait" :

Analogously, the Laplace-Stieltjes-Transform (LST) of T_e according to the scheduling mode "Wait" can be given as :

$$\Phi_e(s) = \Phi_u(s)\{(1-P_0)(B\ \Phi_p^r(s)+(1-B))\ \Phi_h(s) + P_0\}, \tag{3.3}$$

where $\Phi_p^r(s)$ corresponds to the LST of the message processing forward recurrence time in the receiving part.

3.2.2 Conditional Cycle Time

Under the assumption of independence between successive cycle segments, the LST of the server cycle time can be written as a product of all $\Phi_e(s)$. In order to reduce the independence assumption above, we use the consideration of conditional cycle times (c.f. [1]) and define the following random variables (r.v.) for the time between two succeeding server visits at a particular transmit buffer :

- $T_{c'}$ r.v. for the cycle time, conditioning on an empty transmit buffer at the previous scanning instant.
- $T_{c''}$ r.v. for the cycle time, conditioning on a non-empty transmit buffer at the previous scanning instant.

Thus, we obtain the LST of the conditional cycle times for the scheduling modes "Repeat" and "Wait" as follows :

i) Scheduling mode "Repeat" :

$$\begin{aligned}\Phi_{c'}(s) &= \Phi_u(s)\ \Phi_e(s)^{(g-1)} \\ \Phi_{c''}(s) &= \Phi_u(s)((1-B)\ \Phi_h(s) + B)\ \Phi_e(s)^{(g-1)}.\end{aligned} \tag{3.4}$$

ii) Scheduling mode "Wait" :

$$\begin{aligned}\Phi_{c'}(s) &= \Phi_u(s)\ \Phi_e(s)^{(g-1)} \\ \Phi_{c''}(s) &= \Phi_u(s)(B\ \Phi_p^r(s) + (1-B))\ \Phi_h(s))\ \Phi_e(s)^{(g-1)}.\end{aligned} \tag{3.5}$$

Based on equations (3.4) and (3.5), the first two moments of the conditional cycle times can easily be calculated for the scheduling modes [6].

3.2.3 Transfer Request Process

In order to determine the blocking probability for messages at the receive buffer, we consider the transfer request process, i.e. the point process consisting of instants, at which requests for transfer to receive buffers are observed. Assuming the transfer request process to be a renewal process and denoting T_r to be the r.v. for the interrequest time, i.e. the time between two successive transmission requests, we obtain the LST $\Phi_r(s)$ according to the considered scheduling modes "Repeat" and "Wait":

i) Scheduling mode "Repeat" :

$$\Phi_r(s) = \frac{(1-P_0)\Phi_u(s)}{1 - P_0\Phi_u(s)}\quad \{(1-B)\ \Phi_h(s) + B\}. \tag{3.6}$$

ii) Scheduling mode "Wait" :

$$\Phi_r(s) = \frac{(1-P_0)\Phi_u(s)}{1 - P_0\Phi_u(s)}\quad \{B\ \Phi_r^r(s) + (1-B)\}\ \Phi_h(s). \tag{3.7}$$

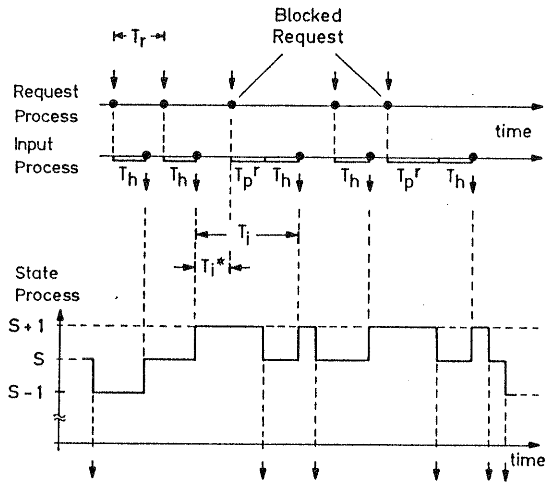
In order to calculate the input process to a particular receive buffer, the first two moments of the message transfer request process is decomposed into g input streams, which are offered to the receive buffers. Subsequently, using the decomposition method provided in [9], we obtain the moments of the distribution of the request process seen from a particular receiver.

3.3 Receiving Part

Using the decomposition approach discussed above, the model of the receiving part of a particular station is just a GI/M/1-S queue. Considering the equally distributed branching probability, the input process of this queue is given by a decomposed process out of the output process of the entire polling system in the sending part. Thus, the receiving part analysis reduces to the state analysis of standard type GI/M/1-S queueing model, in conjunction with a Markov chain analysis. In order to estimate the receive buffer blocking probability B at the transmission request instants, which is used for the feedback in the iteration discussed below, distinction is made between two processes concerning the receiving part (c.f. Figs. 3,4) :

- the process of transmission requests described by the r.v. T_r
- the input process of successsfully transmitted messages described by the r.v. T_i (i.e. the decomposed output process of the polling system).

By means of the state analysis of the GI/M/1-S model, however, we obtain the receive buffer state probabilities $\{P(k),\ k = 0,1,\ldots,S+1\}$ at regeneration points of the Markov chain, which is imbedded immediately before the events of the input process. As depicted in Figs. 3 and 4, the



T_r : Transmission interrequest time

T_h : Message transmission time

T_p : Message processing time at the receive buffer

T_p^r : Forward recurrence message processing time

T_i : Message interarrival time at the receive buffer (Regeneration points of the imbedded Markov chain)

T_i^* : Time between last message arrival and next transmission request

Fig. 3 : Sample path of the receive buffer state process (Scheduling mode "Wait")

instants of successfully transmitted messages are defered by the message transmission time compared to their request instants, while the blocked transmission requests (Fig. 3) do not affect the receive buffer state process. The input process, which is in general nonrecurrent, is assumed to be renewal for further analysis. Based on the transmission request process as well as the input process and out of Markov chain state probabilities, approximate formulae for the blocking probability B for messages at request instants will be derived in the following.

i) Scheduling mode "Repeat" :

Denote T_i^* the r.v. whose distribution is that of the difference between the interrequest time and the transmission time, we obtain

$$B = \{P(S) + P(S+1)\} \frac{e_0[T_i^*]}{1 - e_0[T_r]} , \qquad (3.8)$$

where $e_0[T]$ = Pr{no departures in the message handler during T}

$$= \int_{t=0}^{\infty} e^{-t/E[T_p]} \, dF(t). \qquad (3.9)$$

ii) Scheduling mode "Wait" :

Analogously, the blocking probability B can be derived for this scheduling mode (c.f. Fig. 4). The extensive derivation of the formula will not be discussed here in more detail.

$$B = \{P(S) + P(S+1)\} \; e_0[T_i^*]. \qquad (3.10)$$

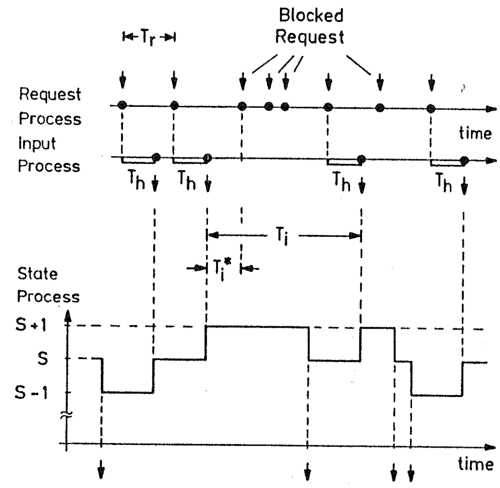


Fig. 4 : Sample path of the receive buffer state process (Scheduling mode "Repeat")

3.4 Calculation Algorithm

Based on the expressions derived in section 3.2 for the sending part and the receive buffer blocking probability derived in section 3.3 a numerical algorithm is developed. The main steps during an iteration cycle are :

i) sending part analysis :
- Calculation of the first two moments of a cycle time segment and the conditional cycle times
- Calculation of the server output and the transmission request processes (c.f. [2])
- Decomposition of the output and the request process into input and the request processes seen from the receivers.

ii) receiving part analysis :
- Analysis of the GI/M/1-S receiving part model to obtain the receive buffer state probability at the message arrival instants
- The GI-Messages interarrival time probability density function is estimated by means of a two-moment description according to [9] in conjunction with a substitute process
- Calculation of the blocking probability for messages at the request instants.

iii) use of the obtained blocking probability to update the conditional cycle times for the next iteration cycle.

iv) iteration until convergence and calculation of system characteristics according to [1].

3.5 System Characteristics

Based on eqns. (3.4) and (3.5) the transmit buffer state analysis shown in [1] can be performed. Thus, we obtain performance measures for the different scheduling modes, e.g. mean waiting time of messages in the transmit buffer, mean number of messages waiting in the transmit buffer or mean and coefficient of variation of the cycle time. All performance measures depend on the blocking probability of messages at their destinating receive buffer. It should be noted here that for the different scheduling modes the well-known formula for the mean cycle time (c.f. [1]) can be modified for the case of receive buffer constraints yielding

i) Scheduling mode "Repeat" :

$$E[T_c] = \frac{g\ E[T_u]}{1 - (1 - B)\rho} \tag{3.11}$$

ii) Scheduling mode "Wait" :

$$E[T_c] = \frac{g\ E[T_u]}{1 - (1 + g\ B)\rho}\ , \tag{3.12}$$

where the mean message processing time is chosen as

$$E[T_p] = g.E[T_h]\ . \tag{3.13}$$

4. NUMERICAL EXAMPLES

In the following, numerically obtained results will be presented and discussed for the case of a symmetrically loaded system in order to illustrate the accuracy of the derived algorithm. The system consists of g = 5 units. Each receive buffer has the capacity S = 5. For the results presented, the time variables are normalized to $E[T_h] = 1$. The switchover time is chosen to be constant with $E[T_u] = 0.5\ E[T_h]$. In order to validate the approximation, computer simulations are provided, whereby the simulation results will be depicted with their 95 % confidence intervals. The diagrams are drawn for performance measures as functions of the total offered traffic $\rho = g.\ \lambda E[T_h]$ to the communication medium.

Figs. 5 and 6 show performance comparisons of the two scheduling modes "Wait" and "Repeat", where attentions are devoted to the mean cycle time and the mean waiting time of messages. It can be seen clearly in Fig. 5 that at lower load conditions the cycle times are limited to the sum of all switchover times (empty cycle $E[T_c] = 2.5$) and no receiver blocking occurs, i.e. both scheduling modes lead to the same mean cycle time. Considering the scheduling mode "Repeat", only a switchover time is observed in the case of receive buffer blocking, where in the scheduling mode "Wait" the server waits until the blocking receive buffer is able to store the message, which has to be transmitted. This leads to a higher mean cycle time and consequently to higher mean waiting time, when the scheduler operates under the scheduling mode "Wait". Furthermore, Fig. 5 shows that the scheduling mode "Repeat" leads to smaller mean cycle times for transmission times with higher variances. This is caused by a higher blocking probability for larger c_h.

The cycle time coefficient of variation is depicted in Figs. 7 as function of the offered traffic intensity for the scheduling mode "Repeat", assuming Markovian ($c_h = 1$) and hyperexponential transmission time ($c_h = 5$). As expected, for disappearing message traffic, the cycle time coefficient of variation starts at zero, due to the chosen constant switchover time. With

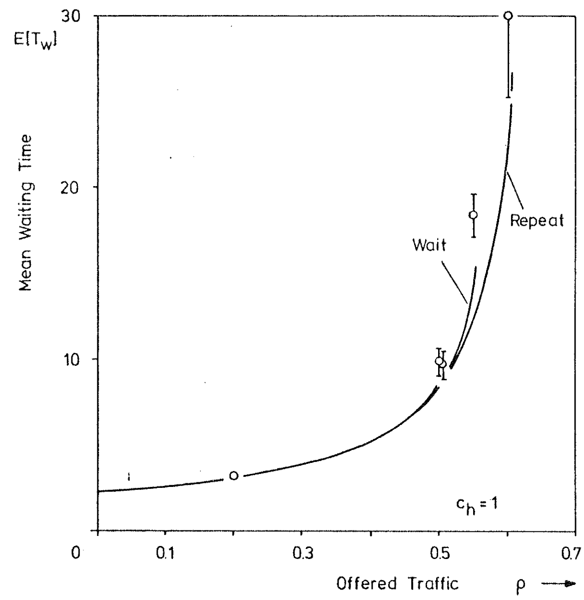


Fig. 6 : Mean message waiting time versus offered traffic intensity

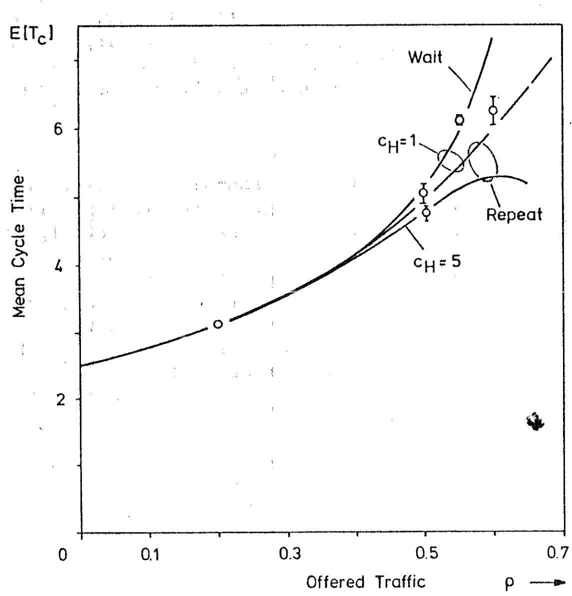


Fig. 5 : Mean cycle time versus offered traffic intensity

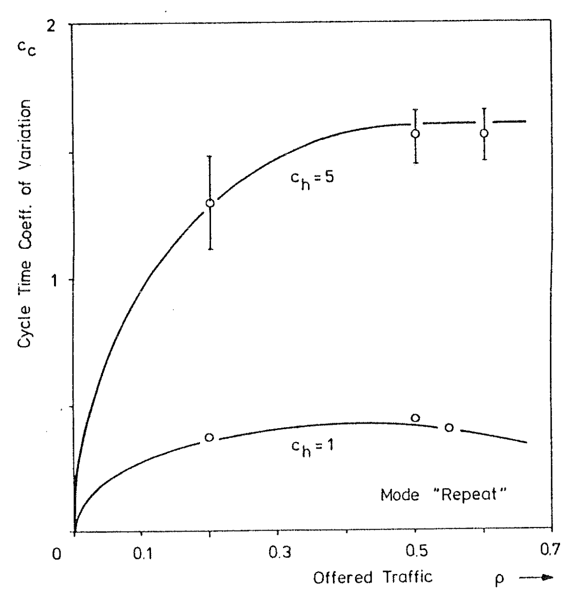


Fig. 7 : Mean cycle time coefficient of variation versus offered traffic intensity

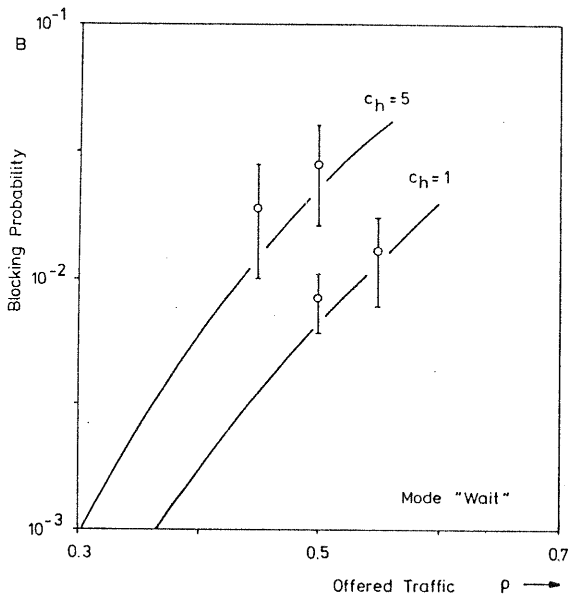


Fig. 8 : Message blocking probability versus offered traffic intensity

increasing values of the transmission time coefficient of variation and the offered traffic intensity, an increase of the cycle time coefficient of variation is also shown in this Figure.

The influence of the transmission time coefficient of variation on the blocking probability for scheduling mode "Wait" is shown in Fig. 8. It can be seen that the approximation accuracy is higher for smaller values of c_h. In general, results obtained by the presented method always show the same tendencies and phenomena as delivered by computer simulations.

5. CONCLUSION

In this paper, the traffic performance of message interchanging mechamisms via an interconnection network is investigated, which have, e.g., a bus- or ring-oriented structure and operates according to a cyclic scheduling strategy. The analysis is based on a polling system with cyclic, nonexhaustive service, whose output process is further offered to the receiver having finite capacity buffers. An approximate analysis approach is developed, where different bus scheduling modes in the case of receive buffer blocking are taken into account. Numerical results showing system characteristics are presented, whereby the approximation accuracy and its dependency on system parameters are discussed in accordance with validations by simulations. The modelling approach and the analysis method presented can be applied for a wide range of models in computer and communication systems, such as local area networks or stored program controlled switching systems with distributed structures.

ACKNOWLEDGEMENT

The authors would like to thank Mr. M. Bosch for helpful discussions and for his programming efforts.

REFERENCES

[1] P.J. Kuehn,"Multiqueue Systems with Nonexhaustive Cyclic Service", Bell Syst. Tech. J. 58 (1979) 671-699.

[2] P.J. Kuehn, "Performance of ARQ-Protocols for HDX-Transmission in Hierarchical Polling Systems", Perf. Eval. 1 (1981) 19-30.

[3] H. Takagi, L. Kleinrock, "Analysis of Polling Systems", Japan Science Inst. Research Report, 1985.

[4] D. R. Manfield, "Analysis of a Polling System with Priorities", Proc. Globecom'83, San Diego 1983, paper 43.4.

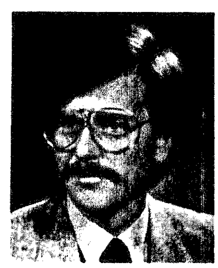
[5] R.J.T. Morris, Y.T. Wang, "Some Results for Multiqueue Systems with Multiple Cyclic Servers", Proc. 2nd Int. Symp. on the Performance of Comp. Comm. Systems, Zurich 1984, 245- 258.

[6] T. Raith, "Performance Analysis of Multibus Interconnection Networks", Proceedings of the 11th International Teletraffic Congress, Kyoto 1985, paper 4.2A-5.

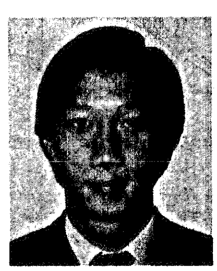
[7] P. Tran-Gia, T.Raith, "Multiqueue Systems with Finite Capacity and Nonexhaustive Cyclic Service", Int. Seminar on Computer Networking and Perf. Eval., Tokyo 1985, paper 5.3.1.

[8] P. Tran-Gia, T. Raith, "Approximation for Finite Capacity Multiqueue Systems", GI/NTG Conference on Measurement, Modelling and Eval. of Computer Systems, Oct. 1985, Dortmund, F.R.G.

[9] P.J. Kuehn, "Approximate Analysis of General Queueing Networks by Decomposition",IEEE Trans. Comm. COM-27(1979) 113-126.



Thomas Raith received the M.S. degree (Dipl.-Ing.) in Electrical Engineering from the Univ. of Stuttgart, Germany, in 1981. Since 1982 he has been a scientific staff member at the Institute of Communications Switching and Data Technics, Stuttgart Univ., where he is currently working on the performance investigation of computer communication systems.



Phuoc Tran-Gia received the M.S. degree (Dipl.-Ing.) from the Stuttgart Univ., Stuttgart, in 1977 and the Ph.D. degree (Dr.-Ing.) from the Univ. of Siegen, Siegen, in 1982, both in Electrical Engineering. In 1977, he joined Standard Elektrik Lorenz (ITT), Stuttgart, where he was working in software development of digital switching systems. From 1979 until 1982 he was working as assistant professor in the Dept. of Communications, Univ. of Siegen. Since 1983, Dr. Tran-Gia has been Head of a Research Group in the Institute of Comm. Switching and Data Technics, Stuttgart University, Germany. His current research activities are in the field of test methods and performance analysis of computer and communication systems.