ABSTRACT

In distributed systems loosely coupled units (e.g. microprocessor-based control devices, peripheral processors etc.) often communicate with each other through a communication subsystem. The communication subsystem consists of a transmit and a receive buffer per unit which are connected via an interconnection network through an individual port. In the paper the performance evaluation of a communication subsystem with an interconnection network consisting of one or several high-speed busses is considered. The approximate analytical solution is based on approaches as decomposition methods, two moment matching and imbedded Markov chains. The approximation is validated by means of computer simulations. Numerical results were found to be in good agreement over a wide parameter range. The class of models considered in this paper arise in performance evaluation of switching systems with distributed control, token-ring local area networks, and distributed systems of multiple interconnected computers.

1. INTRODUCTION

Autonomous loosely coupled units in systems with distributed control often communicate with each other by message passing through a communication subsystem. Each of the units is considered as a processor with its local memory. The units 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 the interconnection network through an individual port, see Fig. 1. The ports are assumed to operate in a full duplex mode, i.e. a port can transmit and receive messages simultaneously. In the paper an interconnection network consisting of one or several high-speed busses is considered. In the case of multibus systems the busses are assumed to work in parallel and independently, i.e. overtaking of bus grants can occur. The bus allocation to each particular unit is organized by means of a cyclic schedule. The transmit buffers are served in a nonexhaustive manner, that means per polling instant only one message will be transmitted.

The main subject of this paper is the performance investigation of the communication subsystem. The subsystem consists of the interconnection network, the transmit and receive buffers as well as the ports and may cause a performance degradation which depends mainly on the following blocking effects:

- transmit blocking due to port limitation, i.e. the transmit buffer can only be served by one bus at a time, therefore overtaking of bus grants may occur and the units are not able to use the full transmission capacity provided by the multibus interconnection network.

- receive blocking due to port limitation, i.e. the occupation of the receiver port is not possible due to the fact that the receiver port has been already occupied by another bus of the interconnection network.

- receive blocking due to buffer limitation, i.e. the transmission between units fails due to finite capacity of the receive buffer of the receiving unit.

All communication buffers are of finite capacity. Therefore, messages of a unit may be blocked in the transmit buffer due to its finite length. Blocking according to port limitation arises in interconnection networks where several busses are working in parallel and interfere with each other. In the case of receive blocking, either the scheduler proceeds to the transmit buffer of the following unit and the message waits further on within its transmit buffer until the next bus grant occurs, or the bus remains occupied and the message waits in the transmit buffer until the blocked receive buffer gets idle. Thus, throughput and delay performance of the communication subsystem depend on its

Fig. 1 Structure of a communication subsystem with a multibus interconnection network
structure and bus scheduling mode as well as on parameters like arrival rates of messages, the physical transmission rates of the buses, time out and receive buffer capacities, and the rate of emptying the receive buffer of the receiving units.

The paper aims at the modelling of a communication subsystem with a multibus interconnection network operating under several scheduling modes, the evaluation of its performance, and a comparison between a single- and multibus interconnection network.

2. MODELLING

Modelling of the communication subsystem with a multibus interconnection network leads to a queueing model depicted in Fig. 2, which consists of a number n of finite transmit and receive buffers with the capacity S_i and R_i respectively, and a number m of high-speed buses with transmission rate r_i. Each bus is allocated to one of the transmit buffers at a time. The allocation is done by a cyclic schedule with non-blocking service. Due to the parallel transmission capability and finite buffers, the blocking effects discussed above may occur. The arrival process of messages from the sending units are assumed to be Poissonian with queue-individual rates λ_i, i = 1,...,n at the transmit buffers. Receive buffers are emptied at rates μ_i, i = 1,...,m, with individual Markovian service times. The bus service time, i.e. the time to transmit a message from transmit buffer to receive buffer is considered to be generally distributed. After service of a transmit buffer the bus will be allocated by the scheduler to the succeeding unit to serve its transmit buffer, if there is at least one message waiting for service. If the transmit buffer is empty, the observed interarrival period will be denoted as switchover time. This switchover time, which models all overheads spent and procedures performed by the scheduler to allocate the busses in a cyclic manner, is assumed to have an unit-individual general distribution function. In case of receive blocking two bus scheduling modes are considered:

1) the scheduler proceeds to the following transmit buffer and the message waits further on within its transmit buffer until the next bus grant occurs (bus repeat mode).

2) the bus remains occupied and the message has to wait until the blocked receive buffer gets idle and the receiving unit is able to accept the message (bus wait mode).

Without receive buffer consideration the model in Fig. 2 corresponds to a multiserver polling system with nonexhaustive cyclic service and finite queue capacity. In the literature, multicue systems served by a single server have been subject of numerous investigations. An approximation technique introducing the method of conditional cycle times for cyclic queues with nonexhaustive service and general switchover time has been developed by Kuehn [3]. A survey on single server polling system analysis, where various system classes are considered, was provided by Takagi and Kleinrock [4]. In [2] Tran-Gia and Raitt propose an approximative analysis method for a single server polling system with finite buffer capacity, i.e. an investigation on the effect of message blocking due to transmit buffer limitation. A previous study on the problem of multiqueue systems with multiple cyclic servers was performed by Morris and Wang [1]. They give a simple formulae for the mean sojourn time in the multiqueue system and consider several service disciplines of the queues but they do not consider any blocking effects as well as additional overhead (switchover times) caused by server overloading.

Arrival of messages

λ_1 λ_2 λ_3 ... λ_n

Transmit buffers

Ports

Cyclic schedule

R_1 ... R_n

Busses

Receive buffers

Ports

Departure of messages

μ_1 μ_2 ... μ_m

Fig. 2 Model of the communication subsystem

3. PERFORMANCE ANALYSIS

In this section, a numerical algorithm for an approximate analysis of a multiserver polling system with finite buffer capacity, i.e. the communication subsystem of Fig. 1, will be derived. Basically, the analysis follows the method presented in [2]. However, some modifications must be provided in order to take into account the port and memory blocking effect of the transmit and receive buffers. The main idea in the calculation method presented in this paper is to develop an alternating calculation algorithm to obtain values for the Markov chain state probabilities and the server intervisit times of a considered transmit buffer. The probability density function (pdf) of the intervisit time was approximated by a two moment technique proposed by Kuehn [5].

3.1 Markov Chain State Probabilities

A particular transmit buffer j is considered in the following, which is observed at scanning instants, i.e. instants where a bus grant occurs.

Let t_i be the time of the n-th scanning epoch and let X(n)(0^t_i) be the number of messages in this buffer at time t_i, just prior to the n-th scanning epoch, then the Markov chain state probabilities are defined as

P_k = P[X(n)(0^t_i) = k], k = 0,1,...,S_j

For ease of reading, the subscript j indicating the observed transmit buffer will be suppressed, e.g., the notation P_k will be used instead of P_k,j.

4.2A-5-2
In order to calculate the transition probabilities of the Markov chain
\[ p_{jk} = \Pr( X^{(n+1)}(t) = k | X^{(n)}(t) = j ) \] (3.2)
the state \( X^{(n)}(t) \) of the transmit buffer at time \( t_n \) is observed. Considering the pure birth process in the buffer between two consecutive visits of an arbitrary server, the state probabilities at time \( t_n \) can be obtained as follows

\[ p_k(n)(t) = \Pr( X^{(n)}(t) = k | \lambda = 1, k = 0, 1, \ldots, S_j, \) (3.3)

According to the consideration of conditional cycle times \( [3] \), the following random variables (r.v.) for the time between two succeeding server visits to buffer \( j \) are defined:

- \( T_{iv} \) r.v. for the intervisit time with respect to the observed transmit buffer \( j \).
- \( T_{iv}' \) r.v. for the intervisit time, conditioning on an empty buffer at the previous scanning instant, that means without service of buffer \( j \) during the cycle.
- \( T_{iv}'' \) r.v. for the intervisit time, conditioning on a non-empty buffer at the previous scanning instant (i.e., with service of buffer \( j \) during the intervisit time).

Some algebraic manipulations, see \([2]\), yield to the set of Markov chain state equations (3.4), which are useful for the numerical calculation of the steady state probabilities \( \{ p_k \} \)

\[ p_k(n+1) = \sum_{i=1}^{S_j} \sum_{k=0}^{S_j} p_i(n) b_i' \] (3.4)

\[ p_k(n+1) = \sum_{i=1}^{S_j} \sum_{k=0}^{S_j} p_i(n) b_i'' \]

where the arrival probabilities, i.e., the probabilities for \( m \) arrivals during a conditional intervisit time of type \( T_{iv} \) or \( T_{iv}' \) are

\[ b_i' = \int_0^\infty a_m(t) f_{iv}(t) dt \]

\[ b_i'' = \int_0^\infty a_m(t) f_{iv}'(t) dt \] (3.5)

where \( a_m(t) \) corresponds to the probability that \( m \) Poisson arrivals occur during the time interval \( t \). In order to calculate the arrival probabilities, the pdf of the conditional intervisit times have to be determined.

3.2 Cycle Time Segment Analysis

In this subsection some considerations where made to obtain the probability of blocking effects with respect to the bus scheduling modes discussed above. Based on these blocking effects a random variable \( T_{iv} \) for the time interval between the scanning epochs of transmit buffer \( j \) and \((j+1)\) is defined, i.e., the cycle time segment corresponding to buffer \( j \) with respect to an arbitrary server.

The port blocking probability of a transmit or receive buffer \( j \) \( \{ B_{PT,j}, B_{PR,j} \} \) corresponds to the probability that the servers interfere with each other at the port of the considered buffers (transmit- or receive buffer). Under the assumption that messages are equally distributed to the receive buffers, they are obtained as

\[ P_{PT} = B_{PT} = \frac{(n-1)}{n} \lambda (1-B_{MT})E[T_h]. \] (3.6)

The subscript \( j \) is again suppressed for ease of reading. The memory blocking probability of the receive buffer \( (B_{MR,j}) \) corresponds to the probability that a bus transmission will be blocked due to finite capacity of the receive buffer and will be approximated by the blocking probability of a M/M/1-R system. The memory blocking probability of a transmit buffer \( (B_{MT,j}) \) corresponds to the probability that a message will be blocked due to finite capacity of the transmit buffer and will be obtained by the arbitrary time state probability \( B_{PT,j} \) of the considered transmit buffer. The probability of service of a transmit buffer where messages are waiting depends on the bus scheduling mode and is defined as \( B_{PT,j} \), i.e., the probability of additional overhead (switchover time) caused by server overtaking or receive buffer blocking.

3.2.1 Repeat Mode

According to bus scheduling mode "repeat" the probability of service of transmit buffer \( j \) \( (q_{s,j}) \), is obtained as \( (j \) being omitted) :

\[ q_s = 1 - (B_{PT} + B_{PR} + B_{MR}). \] (3.7)

The Laplace-Stieltjes-Transform (LST) of the cycle time segment corresponding to buffer \( j \) of an arbitrary server \( (T_{iv,j}) \) can be given as

\[ \phi_E(s) = \phi_u(s)((1-P_0) q_s \phi_h(s) + P_0). \] (3.8)

Thus, mean and variance of the cycle time segment are

\[ E[T_{iv}] = E[T_{iv}'] + (1-P_0) q_s E[T_h]. \]

\[ \text{VAR}[T_{iv}] = \text{VAR}[T_{iv}'] + (1-P_0) q_s (\text{VAR}[T_h] + E[T_h]^2(1-q_s(1-P_0))). \] (3.9)

3.2.2 Wait Mode

According to bus scheduling mode "wait" the probability of service transmit buffer \( j \) \( (q_{s,j}) \), is obtained as

\[ q_s = 1 - (B_{PT} + B_{PR} + B_{MR}). \] (3.10)

In this subsection the Laplace-Stieltjes-Transform (LST) of the cycle time segment \( (T_{iv,j}) \) includes the forward recurrence time of bus service time and receive buffer service time. Based on the LST, again mean and variance of the cycle time segment can be calculated.

3.3 Conditional Intervisit Time Approximation

Under the assumption of independence between \( T_{iv,j} \), \( j=1,2,\ldots,\beta \), the LST of the conditional cycle times of an arbitrary server can be given as follows
\[ \phi_{C',j}(s) = \Phi_{u_j}(s) \cdot \prod_{k=1}^{j-1} \Phi_{E,k}(s) \quad k \neq j \]  
\[ \phi_{C'',j}(s) = \Phi_{u_j}(s) \cdot \Phi_{h_j}(s) \cdot \prod_{k=1}^{j-1} \Phi_{E,k}(s) \quad k \neq j \]  
(Eqn. (3.11) yield the first two moments of the conditional cycle times, thus

\[ E[T'_{C,j}] = E[T_{u_j}] + \sum_{k=1}^{j} E[T_{E,k}] \]  
\[ \text{VAR}[T'_{C,j}] = \text{VAR}[T_{u_j}] + \sum_{k=1}^{j} \text{VAR}[T_{E,k}] \]  

\[ E[T''_{C,j}] = E[T_{u_j}] + E[T_{h_j}] + \sum_{k=1}^{j} E[T_{E,k}] \]  
\[ \text{VAR}[T''_{C,j}] = \text{VAR}[T_{u_j}] + \text{VAR}[T_{h_j}] + \sum_{k=1}^{j} \text{VAR}[T_{E,k}] \]  

To obtain the resulting conditional intervisit times, the bus-individual intervisit times are superimposed under the assumption of independent renewal processes [5]. Since only one bus can serve a queue at a time, overtaking has to be considered. Assuming a geometrical distribution for overtaking a transmitted buffer by an arbitrary server, the following intervisit times can be obtained. Based on these first two moments the calculation of the arrival probabilities (3.5) can be performed according to [2].

3.4 Calculation Algorithm for Markov Chain State Probabilities

Using the expressions for the Markov chain state probabilities and the conditional intervisit times obtained by composition of server cycle time processes where the server overtaking effect, i.e., transmit buffer blocking due to port limitation, are taken into account, a numerical algorithm is developed. Details of the algorithm are given in [2]. The main elements of the algorithm are:

- Iteration of the Markov chain state probabilities and the intervisit time.
- During an iteration cycle the state probabilities of all buffers are determined in a cyclic manner; the calculation for each buffer is done according to eqn. (3.4).
- During an iteration cycle, depending on the actual state probabilities the conditional intervisit times are updated; these values will be used in the next iteration cycle.
- Calculation of the arrival probabilities by means of a two-moment approximation of the intervisit time probability density function according to [5] in conjunction with a substitute process description.

3.5 Arbitrary Time State Probabilities

In order to calculate system characteristics, e.g., the blocking probability for messages or mean waiting time in a buffer, it is useful to obtain first the arbitrary time state probabiliti-
4. RESULTS

In the following, numerically obtained results will be presented and discussed for the case of a symmetrically loaded communication subsystem with a single- or double bus interconnection network, in order to illustrate the accuracy of the derived algorithm. The system consists of \( g = 8 \) units communicating over the communication subsystem. Each of the transmit and receive buffers have the same capacity \( S_j = R_j = 10 \). For the results presented, the time variables are standardized by \( E[T_{h_j}] = 1 \), \( j = 1, 2, \ldots, g \), i.e., the mean bus service time at transmit buffer \( j \). The receive buffers are assumed to be emptied according to a Markovian service time with mean \( g/n \cdot E[T_{h_j}] \). The switch-over time is chosen to be constant where \( E[T_{s_j}]/E[T_{h_j}] = 0.5 \).

![Graph showing Mean Waiting Time vs Offered Traffic Intensity](image1)

**Fig. 3** Mean waiting time vs offered traffic

In order to validate the approximation, computer simulations are provided. The simulation results will be depicted with their 95 percent confidence intervals, where the circular symbol will be used for Markovian bus service times \( (c[T_{h_j}] = 1) \) and the triangle symbol for hyperexponential bus service times \( (c[T_{h_j}] = 5) \). The graphs will be drawn as function of the offered traffic per bus

\[
\rho_o = \sum_{j=1}^{g} \frac{\lambda_j}{E[T_{h_j}]} \quad (4.1)
\]

The overall approximation accuracy for the given system parameters is in general less than 15 percent and depends strongly on the value of the service time coefficient of variation and the mean switchover time. The accuracy of the algorithm increases with increasing values of switchover time and decreasing values of the service time coefficient of variation. Results delivered by the presented method always show the same tendencies and phenomena as they are obtained by computer simulations.

4.1 Communication Subsystem with Double Bus Interconnection Network

In this subsection, an interconnection network of two identical buses is taken into account. The transmit buffer mean waiting time as well as its blocking probability for messages are shown as functions of the offered traffic intensity in Figs. 3 and 4, respectively, for different coefficients of variation of the service time and according to the various bus scheduling modes "wait" and "repeat".

In Fig. 3 a crossover effect of the waiting time characteristics can be observed for the service time coefficient of variation and the scheduling mode "wait" lead to higher waiting times and blocking probabilities than mode "repeat" which is caused by a large forward recurrence time in case of waiting for the port to become free. For small values of the service time coefficient of variation (cf. \( c[T_{h_j}] = 1 \)) there is only a small difference between modes "wait" and "repeat", therefore, "repeat" seems to be the best strategy.

![Graph showing Blocking Probability vs Offered Traffic Intensity](image2)

**Fig. 4** Blocking probability vs offered traffic

4.2 Comparison Between Single and Double Bus Interconnection Networks

In a single bus interconnection network no server interference occurs, thus only memory blocking effects arise, but the mean intervisit times of the single server to the transmit buffers is much greater compared to the double bus interconnection network intervisit times. Therefore, for higher load a significant difference between the mean waiting time is obtained (cf. Fig. 5). In general, it can be observed, that the receive buffer memory blocking effect is very small, because the total bus service capacity is equal to the total service capacity of all receive buffers (i.e. sum of all receive buffer empty rates).
Fig. 6 shows that the double bus network compared to the single bus leads to higher load condition to worse blocking probabilities caused by additional overhead (switchover times) in case of blocking a bus transmission.

The effect that service times with higher variances lead to shorter mean intervisit times in the range 0.5 < \( \rho_o \) < 0.75 of traffic intensity as depicted in Fig. 7 can be explained considering the higher blocking probability by large \( c[T_H] \) (cf. Fig. 6). Fig. 7 shows also the two limiting cases of the mean intervisit time which correspond to low and overload traffic levels. The limiting intervisit times of the single bus interconnection network, see Fig. 7, are given either by the sum of all switchover times (minimal intervisit \( E[T_{IV}] = 4 \)) or the sum of switchover and service times of all transmit buffers (maximal intervisit \( E[T_{IV}] = 12 \)). In principle, multibus intervisit times are obtained by the single bus intervisit times divided by the number of buses considered. However, for higher traffic the probability of service of a transmit buffer decreases; therefore, the mean intervisit time decreases, too.

As depicted in Fig. 8, the cycle time coefficient of variation increases by increasing service time variation. If the traffic intensity approaches zero the intervisit time coefficient of variation of the single bus network starts from zero because the empty bus cycle is determined by the constant switchover times. The
composition of two bus cycle time processes leads to the intervisit time coefficient of variation shown in Fig. 8. The simulation results show, that in case of higher service time variation the assumption of independece provide a higher calculated intervisit time coefficient of variation, i.e. the algorithm is less accurate.

Finally, the double bus network is considered under the assumption that one of the busses fails and the offered traffic stays equal, that means the residual bus is heavily overloaded. Figs. 9 and 10 show, that for total traffic load larger than 0.25 the mean waiting time and blocking probability of the communication subsystem strongly increase.

5. CONCLUSION

In the paper an approximative performance analysis for distributed systems with a communication subsystem consisting of transmit and receive buffers per unit and a multibus interconnection network is provided. The communication subsystem is modelled by means of a multi server polling system with finite buffer capacity and nonexhaustive cyclic service. An effective numerical algorithm is developed where different blocking effects are taken into account. Under consideration of two bus scheduling modes results for mean waiting time, blocking probability for messages etc. are derived. The accuracy of the presented algorithm is good over a wide range of parameters. This class of models can be applied to performance investigations of computer and communication systems, such as token-ring local area networks or systems of multiple interconnected computers with a distributed structure. Thus, the influence of structure, scheduling mode and parameters of the distributed system with a multibus interconnection network is analyzed and may be a basis for the decision process in the development and engineering of such systems.

ACKNOWLEDGMENT

The author would like to express his thanks to Prof. F. Kuehn and to Dr. P. Tran-Gia for helpful discussions.

REFERENCES


