Experimental research of a shared memory subsystem with limited queue length for specialized reconfigurable multiprocessor systems

Alexey I. Martyshkin
Penza State Technological University, Russian
E-mail: alexy.im@yahoo.com

Dmitry S. Martens-Atyushev
Penza State Technological University, Russian
E-mail: novoselich93@mail.ru

ABSTRACT

Recently, reconfigurable systems based on field programmable logic devices (FPLDs) have been widely used in high-performance computing. The paper discusses issues related to the experimental research of a shared memory subsystem with a limited queue length of specialized reconfigurable multiprocessor systems using the developed mathematical modelling method. The paper presents the results of the method proposed by the authors for modelling multiprocessor systems based on open queuing networks with limited queue lengths. Based on these conditions, as well as the architectural features of the investigated processor-memory subsystem, expressions are calculated to estimate the exchange time and the resulting delays at each exchange stage. During the research, the main attention was paid to the dependence of the increase in the number of processor nodes in the processor-memory subsystem. As a result, the data obtained showed that the processor growth significantly affects the exchange time, creating a significant load on the common bus, as well as increasing delays at the stages when request transfer operation from the processor to the memory is performed. At the same time, the inadequate behaviour of experimental results and inaccuracy of their values when using the basic modelling method are explicitly tracked, which is reflected in the obtained graphs. Computational experiments were carried out to calculate the probabilistic-temporal characteristics of the "processor-memory" subsystem using the developed mathematical modelling methods. Based on the experimental results, it was determined that the delays occurring in subsystem's nodes and the time of exchange between the processor and memory modules depend on the query parameters and the processor-memory subsystem's architectural characteristics.

s415

Keywords: Multiprocessor system, mathematical modelling, computational experiment, processor-memory subsystem, memory architecture.

1. INTRODUCTION

Some multiprocessor systems (MPS) are implemented directly using FPLDs (Ali et al., 2021). This is due to the fact that the reconfigurable hardware platform is fast enough, and most importantly, such multiprocessor systems become universal and flexibile in the process of computing a particular task (Reis & Fröhlich, 2020).

However, problems arising in the construction of traditional multiprocessor systems also appear in the design of reconfigurable multiprocessor systems (RMS). In particular, the task on increasing the speed in a "processor-memory" subsystem is still relevant and significant (Cudennec & Trabelsi, 2020).

An analysis of the sources for the problem under consideration showed that the queuing theory apparatus is mainly used for the study of such problems, as well as other problems associated with increasing the multiprocessor system speed. The advantages of this method are simplicity of implementation and minimal resource costs, since almost any node of a computing system can be represented as a queuing system (QS), and a multiprocessor system can be represented as a set of queuing systems, i.e. in the form of a queuing network (QN). The results of an analytical review on research methods are highlighted in the previous works of the authors (Ghose et al., 2019; Sinha et al., 2021).

Based on the above review, a number of disadvantages of the applied methods were revealed, which we will call basic methods. A basic method is usually characterized by an exponential input flow and service time; the queues in such research methods are, as a rule, not limited (otherwise, it could contradict the concept of limited resources of a computing system), which ultimately leads to not entirely accurate results for studying the probabilistic-temporal characteristics of the multiprocessor systems under study. Also, unlimited queues in systems do not allow studying the delays arising in the processor-memory subsystem devices.

The paper presents the study results for the proposed method on modelling multiprocessor systems based on open queuing networks with limited queues. Proceeding from these conditions, as well as from the architectural features of the investigated "processor-memory" subsystem, expressions were obtained to estimate the exchange time and arising delays at each exchange stage, which were described in detail in (Lakhani, 2020).

s416

The paper's object under study is a mathematical model of the "processor-memory" subsystem of reconfigurable multiprocessor system (RMS) with UMA (Uniform Memory Access) memory architecture type.

This architecture type is often used in the multiprocessor system design (EU4Business, 2021; EUR-Lex, 2003). There is a significant drawback: with the increase in the number of processor nodes (PN), the number of conflicts on the common bus (OS) will increase, since all processors access memory via a single trunk. However, there is some solution to this problem in this case using the function of splitting the read and write transaction, as well as the introduction of a specialized memory controller (MC) into the common bus, which implements the above function. The description of the memory controller functioning is described in sufficient detail in (Lant et al., 2019).

2. METHODOLOG

The model under study is an open queuing network (Figure 1), where S_0 is the source of requests, or processor nodes, S_1 is a common bus, S_2 is memory controller write buffer, S_3 is a memory controller read buffer, S_4 is a shared memory.
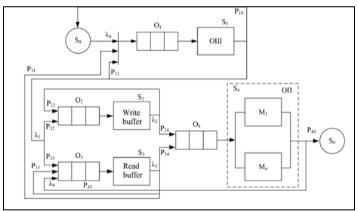

Figure 1: The investigated model of the UMA subsystem

To obtain the results for the probabilistic-temporal exchange characteristics in the "processor-memory" subsystem close to real, it is necessary to take the parameters for the object under study from the existing reconfigurable multiprocessor systems. The input flow rate varied depending on the number of processor nodes (from 2 to 16) and ranged from 0.0114 to 0.0912 requests / ns. In this case, the intensity values are based on the description data of the NIOS II processor cores operating at a frequency of 50 MHz (Eurostat, 2021a).

3. RESULTS AND DISCUSSION

Average service times are calculated according to expressions previously presented in papers (Lopes et al., 2019). Thus, for the basic model $\vartheta_{AT} = 140$ ns, for the model under study $\vartheta_{AT} = 20$ ns (according to the specification for the Avalon bus (Martyshkin et al., 2020; Martyshkin, 2021). The values of the average service time in the memory controller are the following: $\vartheta_{RB} = \vartheta_{WB} = 10$ ns. The queue lengths in a common bus and RAM for the traditional research method are assumed to be unlimited; for the proposed method, the read buffer capacity is 20 requests, the write buffer is 10 requests, and the queue length in the RAM is taken equal to 2 requests.

In the course of the computational experiment performed, the dependence of the increase in the number of processor nodes in the processor-memory subsystem was investigated. As a result, the data obtained showed that an increase in the number of processors significantly affects the exchange time creating a significant load on the common bus, as well as an increase in delays at the operation stages of transferring a request from processor to memory. At the same time, inappropriate behaviour and inaccuracy of values are monitored when using the basic modelling method, which is reflected in the diagrams obtained during the study.

The diagram of dependence between the average waiting times in the queue of the UMA subsystem queuing network and the number of processor nodes is built according to the study results presented in Table 1 and is presented in Figure 2. The following designations are adopted in Table 1, B is the basic modelling method, D is the developed modelling method with limited queue, I - Simulation model.

Table 1: Mean waiting time in UMA subsystem queues

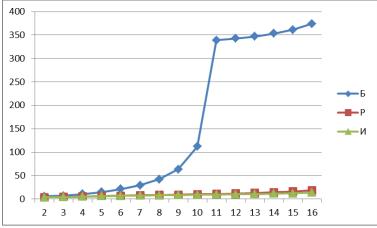| Processor node N	λ_0	Average queue time		
		B	D	S
1	0,0057			
2	0,0114	5,793681531	3,8245799	3,816938049
3	0,0171	7,215464022	4,375691874	4,361333922
4	0,0228	10,7416564	5,407073098	5,366686224
5	0,0285	15,23414843	6,295701471	6,207485422
6	0,0342	21,18351945	7,084174497	6,91950412
7	0,0399	29,49358307	7,812157418	7,535604248
8	0,0456	42,06089196	8,516546153	8,085546468
9	0,0513	63,73079709	9,232560144	8,596537355
10	0,057	111,9792963	9,995486293	9,09424073
11	0,0627	339,1166808	10,84304315	9,604155601
12	0,0684	342,630598	11,81789634	10,15273816
13	0,0741	347,164854	12,99687335	10,79452819
14	0,0798	353,1998716	14,44536547	11,56323519
15	0,0855	361,5945364	16,25848053	12,50732049
16	0,0912	374,0728058	18,58092047	13,69994378

s418

Figure 2: Dependence of the average waiting time in the UMA subsystem queues on the processor node number

It can be seen from the presented diagram that if we increase the number of processors, then the average waiting time in the queue increases, so we can characterize this as waiting for processor nodes to release the common bus from being served by another processor, as waiting time for a response to a request or reading RAM in actual SRMS. The average waiting time values for the basic modelling method have a sharp surge on the diagram with an increase in processor number from 5, and then the obtained indicators have too high values relative to the values of simulation. At the same time, the results obtained on the basis of the developed modelling method correlate with the values of simulation modelling, which confirms the adequacy of the developed method.

Let us present the results of the average time spent in the queuing network. As can be seen from the graph in Figure 3, the behaviour of the curves corresponds to the curves representing the average waiting time in a queue.
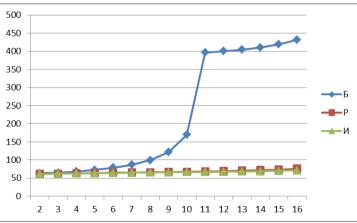

Figure 3: Dependence of the average residence time in the UMA subsystem on the number of processor nodes

s419

Table 2: The average residence time value for a request in the UMA subsystem

Processor node N	λ_0	Average residence time in a subsystem		
		B	D	S
1	**0,0057**			
2	0,0114	63,28546756	61,31636593	61,30872408
3	0,0171	64,70725006	61,86747791	61,85311995
4	0,0228	68,23344243	62,89885913	62,85847226
5	0,0285	72,72593447	63,7874875	63,69927145
6	0,0342	78,67530548	64,57596053	64,41129015
7	0,0399	86,9853691	65,30394345	65,02739028
8	0,0456	99,55267799	66,00833219	65,5773325
9	0,0513	121,2225831	66,72434618	66,08832339
10	0,057	169,4710823	67,48727233	66,58602676
11	0,0627	396,6084668	68,33482918	67,09594163
12	0,0684	400,122384	69,30968237	67,64452419
13	0,0741	404,65664	70,48865938	68,28631422
14	0,0798	410,6916576	71,9371515	69,05502122
15	0,0855	419,0863224	73,75026656	69,99910653
16	0,0912	431,5645918	76,0727065	71,19172982

Having obtained the probabilistic-temporal characteristics of the investigated subsystem "processor-memory", we determine the delays arising at the stages of processing a processor request to the memory. The developed modelling method makes it possible to obtain and evaluate the dependence of delays on the architectural parameters of a processor-memory subsystem. The delays were calculated according to the expressions (Derzhavna sluzhba statystyky Ukrainy, 2021e):

For common bus

$$q_{O\Sh} = \frac{\vartheta_{O\Sh}(\lambda_1\vartheta_{B3})^2(1-\lambda_1\vartheta_{B3})+(\lambda_1\vartheta_{B\Ch})^2(1-\lambda_1\vartheta_{B\Ch})}{(1-\lambda_1\vartheta_{O\Sh})},$$ (1)

For write buffer

$$q_{B3} = \frac{(\lambda_1 p_{12}\vartheta_{B3})^{k+1}\varphi_{0B3}\vartheta_{B3}^{\ 2}\lambda_1 p_{12}(1-(\lambda_1 p_{12}\vartheta_{B3})^k)[k(1-\lambda_1 p_{12}\vartheta_{B3})+1]}{(1-\lambda_1 p_{12}\vartheta_{B3})^2}\varphi_{0B3}+\vartheta_{B3}(\lambda_1 p_{12}\vartheta_{B3})^{k+1}\varphi_{0B3},$$ (2)

For read buffer

$$q_{B\Ch} = \frac{(\lambda_1 p_{13}\vartheta_{B\Ch})^{k+1}\varphi_{0B\Ch}\vartheta_{B\Ch}^{\ 2}\lambda_1 p_{13}(1-(\lambda_1 p_{13}\vartheta_{B\Ch})^k)[k(1-\lambda_1 p_{13}\vartheta_{B\Ch})+1]}{(1-\lambda_1 p_{13}\vartheta_{B\Ch})^2}\varphi_{0B\Ch}+\vartheta_{B\Ch}(\lambda_1 p_{13}\vartheta_{B\Ch})^{k+1}\varphi_{0B\Ch}.$$ (3)
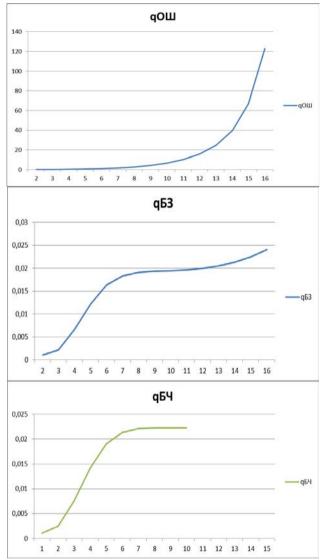
БЗ – WB

БЧ – RB

s420

ОШ - СВ

Figure 4: Dependence of delays in the queuing network of the UMA subsystem on the number of processor nodes

As we can see from the diagrams, the delay value on the common bus with 16 processors was 140 ns, which is quite satisfactory for the CPMS performance. These indicators characterize the increase in requests to shared memory modules, i.e. load on the common bus, but as we can see, this load is not significant, based on the test results of real multiprocessor systems (Martyshkin, 2019). We can say that there are no delays arising at the stages of processing in the memory controller. They were only tens of picoseconds (from 1 to 24 ps).

On the basis of probabilistic-temporal characteristics and values of delays, it is possible to calculate the exchange time in the "processor-memory" subsystem according to the following formula (Nguyen & Sanchez, 2021):

$$t_{o6} = \frac{3(\tau + q_{OW} + u_{OW} + \omega_{O\Pi}) + (u_{E3} + q_{E3})p_{12} + (\frac{(u_{E4} + q_{E4})p_{O\Pi}}{p_{E4}})p_{13}}{N_{cpu}},$$

(4)

БЗ – WB

БЧ – RB

ОШ – CB

Об - Ex

Where τ is the address \ data releasing time on the shared bus by the processor, p_{SM} is the probability that the read data is in the shared memory, p_{RB} is the probability that the read data is in the read buffer, N_{cpu} is the number of processors.

The results are shown in Table 3 and in the diaram (Figure 6)

Table 3: Values of exchange time in the UMA subsystem

Processor node N	λ_0	Exchange time in the "processor-memory" subsystem of the UMA type
1	**0,0057**	t_{ex}
2	0,0114	17,38369887
3	0,0171	11,81574481
4	0,0228	9,212100756
5	0,0285	7,649177355
6	0,0342	6,614648833
7	0,0399	5,889147527
8	0,0456	5,363246718
9	0,0513	4,976461746
10	0,057	4,693165813
11	0,0627	4,502388461
12	0,0684	4,380594168
13	0,0741	4,323551861
14	0,0798	4,320648443
15	0,0855	4,390322802
16	0,0912	4,535961367

s422

Figure 6: Dependence of the UMA subsystem exchange time on the number of processor nodes

The exchange time curve shows that with an increase in the number of processors, the data exchange time between processor nodes and shared memory modules smoothly decreases. Thus, we see that the values are at the 4.5 ns level with a 16-processor system. It can be judged that, given the characteristics of the CPMS, the increase in processor node number will lead to an increase in performance, which is the goal of achieving this study. Moreover, if we apply the basic modelling method already at the stage of obtaining probabilistic-temporal characteristics, we would get approximate values, on the basis of which it would be difficult to assess the research results.

4. CONCLUSION

Computational experiments were carried out to calculate the probabilistic-temporal characteristics of the "processor-memory" subsystem using the developed mathematical modelling methods. Based on the results of the experiments, it was determined that the delays occurring for the subsystem's nodes and the time of exchange between the processor and the memory modules depend on the query parameters and the architectural characteristics of the processor-memory subsystem.

REFERENCES

Ali, H., Tariq, U. U., Hardy, J., Zhai, X., Lu, L., Zheng, Y., ... & Antonopoulos, N. (2021). A survey on system level energy optimisation for MPSoCs in IoT and consumer electronics. **Computer Science Review,** 41, 100416.

Cudennec, L., & Trabelsi, K. (2020, August). Experiments Using a Software-Distributed Shared Memory, MPI and 0MQ over Heterogeneous Computing Resources**. In European Conference on Parallel Processing** (237-248). Springer, Cham.

Ghose, S., Hsieh, K., Boroumand, A., Ausavarungnirun, R., & Mutlu, O. (2019). The processing-in-memory paradigm: Mechanisms to enable adoption. In Beyond-CMOS Technologies for Next Generation Computer Design (133-194). **Springer**, Cham.

s423

Lakhani, K. J. (2020). **Using GPUDirect RDMA and IPC Shared Memory for Direct DMA in a Client-server (FPGA-GPU) System**: One Step Closer to a Fast and Robust DNA Sequencer (Doctoral dissertation, University of California, Davis).

Lant, J., Concatto, C., Attwood, A., Pascual, J. A., Ashworth, M., Navaridas, J., ... & Goodacre, J. (2019). Enabling shared memory communication in networks of MPSoCs. Concurrency and Computation: **Practice and Experience,** 31(21), e4774.

Lopes, A. S., Brandalero, M., Beck, A. C., & Pereira, M. M. (2019, November). Generating optimized multicore accelerator architectures. **In 2019 IX Brazilian Symposium on Computing Systems Engineering (SBESC)** (1-8). IEEE.

Martyshkin, A. I. (2021). Pilot Model of the Embedded Reconfigurable Real Time Computing System. **International Journal of Engineering Research and Technology**, 13(12), 4635-4645.

Martyshkin, A. I., Pashchenko, D. V., Trokoz, D. A., Sinev, M. P., & Svistunov, B. L. (2020). Using queuing theory to describe adaptive mathematical models of computing systems with resource virtualization and its verification using a virtual server with a configuration similar to the configuration of a given model. **Bulletin of Electrical Engineering and Informatics,** 9(3), 1106-1120.

Martyshkin, A. (2019). Software package for determining characteristics of task managers of reconfigurable computer systems using priority queueing networks. **Revista Inclusiones**, 463-474.

Nguyen, Q. M., & Sanchez, D. (2021, October). Fifer: Practical Acceleration of Irregular Applications on Reconfigurable Architectures. **In MICRO-54: 54th Annual IEEE/ACM International Symposium on Microarchitecture** (1064-1077).

Reis, J. G., & Fröhlich, A. A. (2020). Towards deterministic FPGA reconfiguration. **International Journal of Embedded Systems**, 13(2), 236-253.

Sinha, M., Harsha, G. S., Bhattacharyya, P., & Deb, S. (2021). Design space optimization of shared memory architecture in accelerator-rich systems. **ACM Transactions on Design Automation of Electronic Systems (TODAES),** 26(4), 1-31.