

A NOVEL 3D NoC SCHEME FOR HIGH THROUGHPUT UNICAST AND MULTICAST ROUTING PROTOCOLS

N. Ashokkumar, A. Kavitha

Original scientific article

Novel 3D-NoC architecture has been designed by expanding the impression of lossless compression of data. The proposed design shows remarkable results in terms of power efficiency and network throughput. In this scheme, proposed for 3D-NoC, the data to be transmitted is compressed on the transmitting side, so that the data packet is reduced before transmitting. And at the receiver, the original data is restored by decompressing the encoded data. Golomb-Rice algorithm is utilized to develop and implement the hardware encoder and decoder (Codec). The result is evaluated for both unicast and multicast routing. The improvement in power efficiency and throughput for proposed 3D-NoC (unicast) is 9,25 % and 17,61 % respectively. In similar, the improvement in power efficiency and throughput for proposed 3D-NoC (multicast) is 8,65 % and 14,66 % respectively. Further, from the result we observed that the improvement variation is higher for smaller bandwidth, which means the proposed 3D-NoC works well in case of narrow bandwidth.

Keywords: *codec; multicast; SoC; 3D-NoC; unicast*

Nova 3D NoC shema za unicast i multicast usmjerivačke protokole velike propusne moći

Izvorni znanstveni članak

Nova 3D-NoC arhitektura dizajnirana je s većom mogućnošću prihvatanja komprimiranih podataka. Predložena shema daje značajne rezultate u odnosu na učinkovitost i propusnu moć mreže. U toj shemi, predloženoj za 3D-NoC, podaci koje je potrebno prenijeti, komprimirani su već prije prijenosa tako da je veličina paketa s podacima smanjena već prije prenošenja. Pri prijemu, originalni podaci se obnavljaju dekompresijom kodiranih podataka. Golomb-Rice algoritam se primjenjuje za razvoj i primjenu koda i dekodera hardvera (kodek dekodek). Dobiven je rezultat i za unicast i multicast usmjeravanje. Povećanje učinkovitosti i propusne moći za predloženi 3D-NoC (unicast) je 9,25 % odnosno 17,61 %. Slično tome, poboljšanje učinkovitosti i propusne moći za predloženi 3D-NoC (multicast) je 8,65 % odnosno 14,66 %. Nadalje, rezultati pokazuju da je poboljšanje veće kod manjih širina pojasa, a to znači da predloženi 3D-NoC dobro funkcionira u slučaju uskih širina pojasa.

Ključne riječi: *kodek dekodek; multicast; SoC; 3D-NoC; unicast*

1 Introduction

The long term research in VLSI has led to a technology called System on Chip (SoC) [1]. The vast amount of integration makes SoC a perfect candidate for Multimedia Processor, consumer electronics, biological applications. From [2] we can have the knowledge of various components present in SoC. The functionality and performance of the SoC (system on Chip) are increased by utilizing the benefit of 3D IC (Three Dimensional Integrated Circuit). The conventional bus architecture is utilized to communicate among the various components in SoC. On the other hand, Network on Chip (NoC) is emerging as a promising technology [3] to overcome the restricted access of bus architecture in System on Chip (SoC).

The latency in 3D NoC is kept low by stacking the active layer vertically. High resource constraint is achieved, since single chip is used to implement all the hardware utilized in NoC. High scalability, reduced power consumption and low latency are some of the encouraging areas of NoC. Reliable switching model, routing scheme with reduced power along with specific topologies is the protocol followed in Network on Chip (NoC).

2 Related work

The research of interest has been turned towards the Network on Chip (NoC), since the performance of the NoC is promising over the conventional bus architecture. Many works are contributed towards the NoC based on routing, switching and topology. The high speed and low

power performance of the NoC is dominated by switching model and effective wormhole switching model is proposed in [4]. The routing scheme without the occurrence of deadlock is proposed in [5]. The concern of data buffering is overcome by using complex dynamic routing protocol [6]. [7] addresses the crisis faced by Task routing. Deflection routing techniques [8] are anticipated for output contention problems. The proficient layout and simplicity of 3D-mesh topology [9] make them popular among all the topologies for NoC. The network performance in 3D-mesh topology is increased using through-silicon vias (TSVs) [10]. The power consumption during communication is reduced by using three step routing algorithm [7]. So far, many works have been proposed for optimized NoC based on routing scheme, switching model, low power, contention less, dead-lock free and topology. But there is a minor contribution towards the power reduction strategy to improve the performance of NoC architecture. This paper presents a data compression (lossless) technique based on Golomb-Rice encoder and decoder to reduce the power consumption and increase the performance. To complicate the evaluation process of the proposed system, both unicast and multicast routing is incorporated in the design. In sender, the data to be transferred originally is converted to compressed packet by encoder and at the receiver end the compressed packet is converted to original form by the decoder. The lossless compression and decompression is achieved by hardware codec developed using Golomb-Rice algorithm [11]. The proposed codec for 3D-NoC router is developed using Verilog HDL and Design Vision from Synopsys is used to obtain the synthesized results. By analyzing the result, the

efficiency in power consumption and performance is improved for both unicast and multicast routing.

3 Proposed 3D-NoC scheme

A simple 4×4×3 mesh architecture for the proposed NoC scheme is illustrated in Fig. 1. The proposed scheme is practiced in both unicast and multicast routing strategy. The hardware encoder, present, on-chip will compress the data, as soon as transaction of data is initiated at the sender. Then the Network Interface (NI) is used to

packetize the compressed data before transmission. The delivered packets reach the receiver after the packet travels all the on-chip network. The NI in the receiver side will de-packetize the received packets. Then the hardware decoder, present, on-chip is used to bring back the original data. Hence the throughput of the network is improved by reducing the data packet generated by the Network Interface (NI). The feat of lossless compression of data is achieved by incorporating Golomb-Rice algorithm.

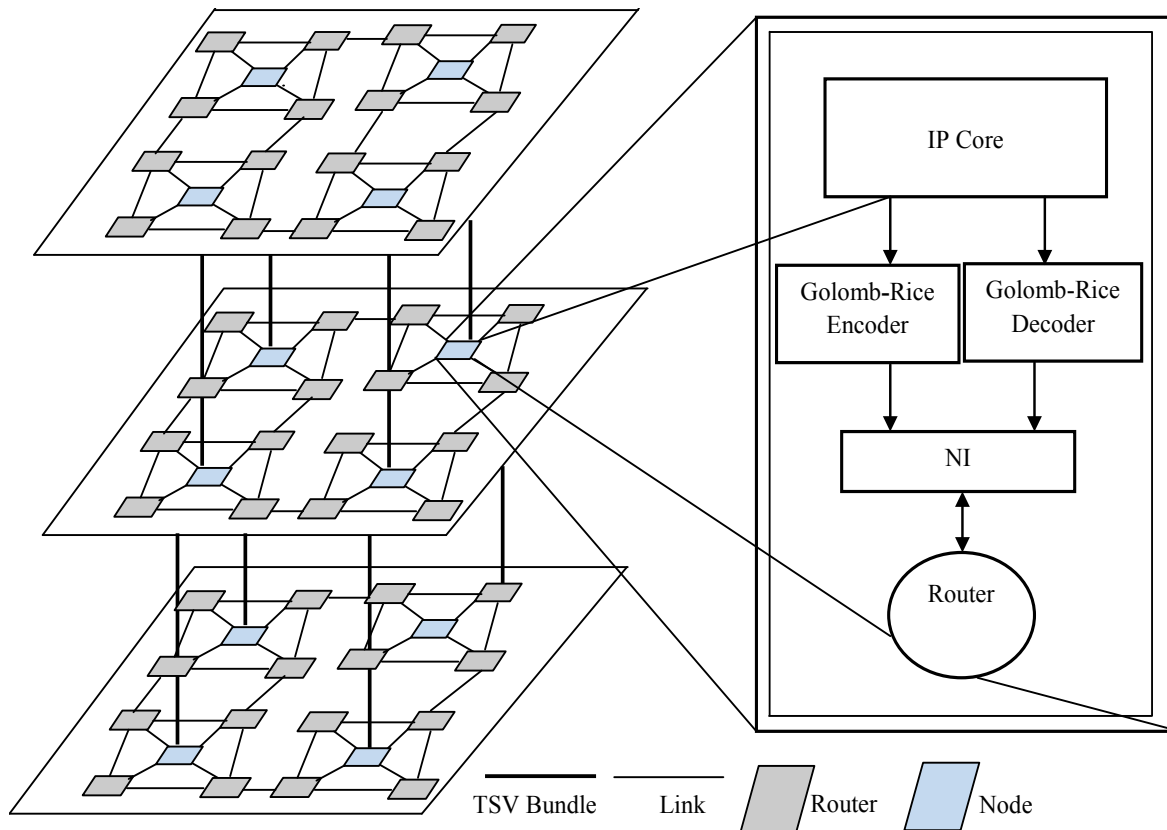


Figure 1 Proposed 3D-NoC Architecture

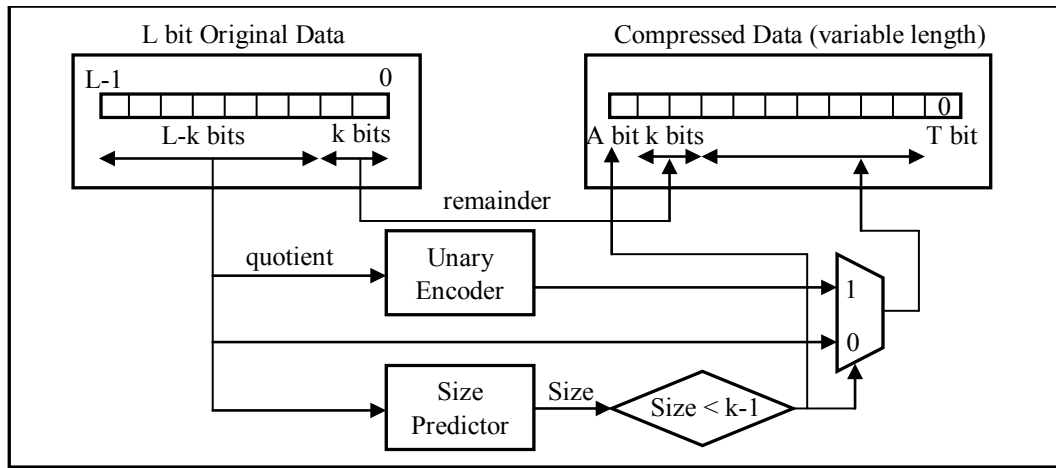
Low complexity of the algorithm made them a perfect match for the lossless compression. Golomb-Rice algorithm [11], for a ‘n’ input data, and 2^k dividend value, binary encode the residual value and unary encode the quotient values. The proposed codec based on a Golomb-Rice algorithm dominate entire lossless compression technique because the determination of the residual and quotient values are easy. The hardware complexity in finding the residual and quotient values is reduced because the dividend value is always a power of two (2^k).

The proposed encoder and decoder based on a Golomb-Rice algorithm are shown in Fig. 2a and 2b. Here 2^k dividend for original data ‘L’ is assumed. The length of the original data may be smaller than the encoded data length in the worst case scenario. It is due to the fact, that in the Golomb-Rice algorithm residual value is converted to fixed length data by binary encoding, whereas, quotient value is converted to variable length by unary encoding. The original data is transmitted as it is, if the

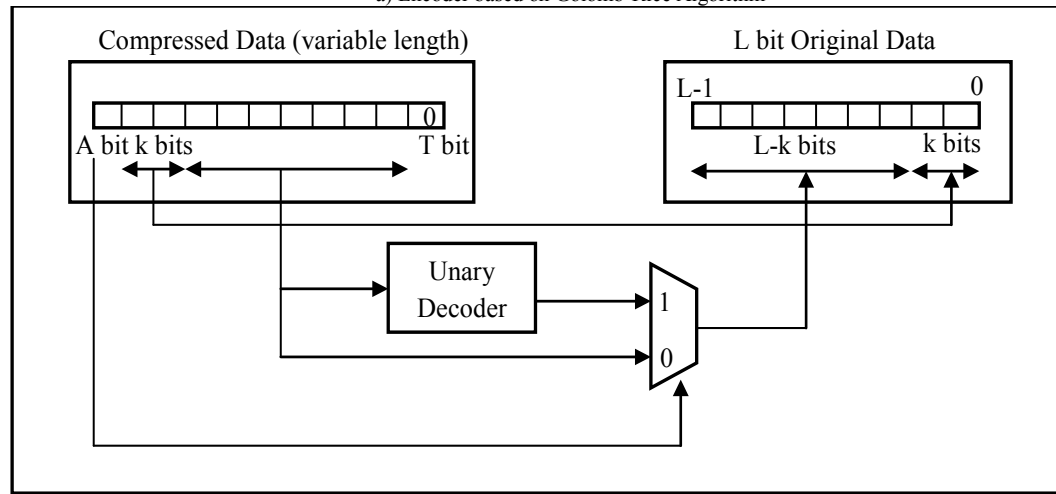
length of original data is less than the length of the encoded data.

The length is calculated by size prediction logic in prior. ‘A’ denotes, whether received encoded data is compressed or not. Due to the variable length of encoded quotient value, the ‘T’ bit is put together to the data encoded to indicate the termination.

Then in receiver side (i.e. Golomb-Rice decoding stage), to check the compression state of the input data, first ‘A’ bit is checked. If the condition $A=1$ is satisfied, the unary decoding of the data before the ‘T’ bit is performed. The process is followed by binary decoding of the k bits. The original data is generated by concatenating the two decode data. On the other hand, if condition $A=0$ is satisfied, the original data is generated by directly concatenating the k bit data after the ‘A’ bit and ‘L-k’ bit data before the ‘T’ bit. The flow of data packets in unicast routing and multicast routing is illustrated in Fig. 3.



a) Encoder based on Golomb-Rice Algorithm



b) Decoder based on Golomb-Rice Algorithm

Figure 2 Proposed Hardware Codec

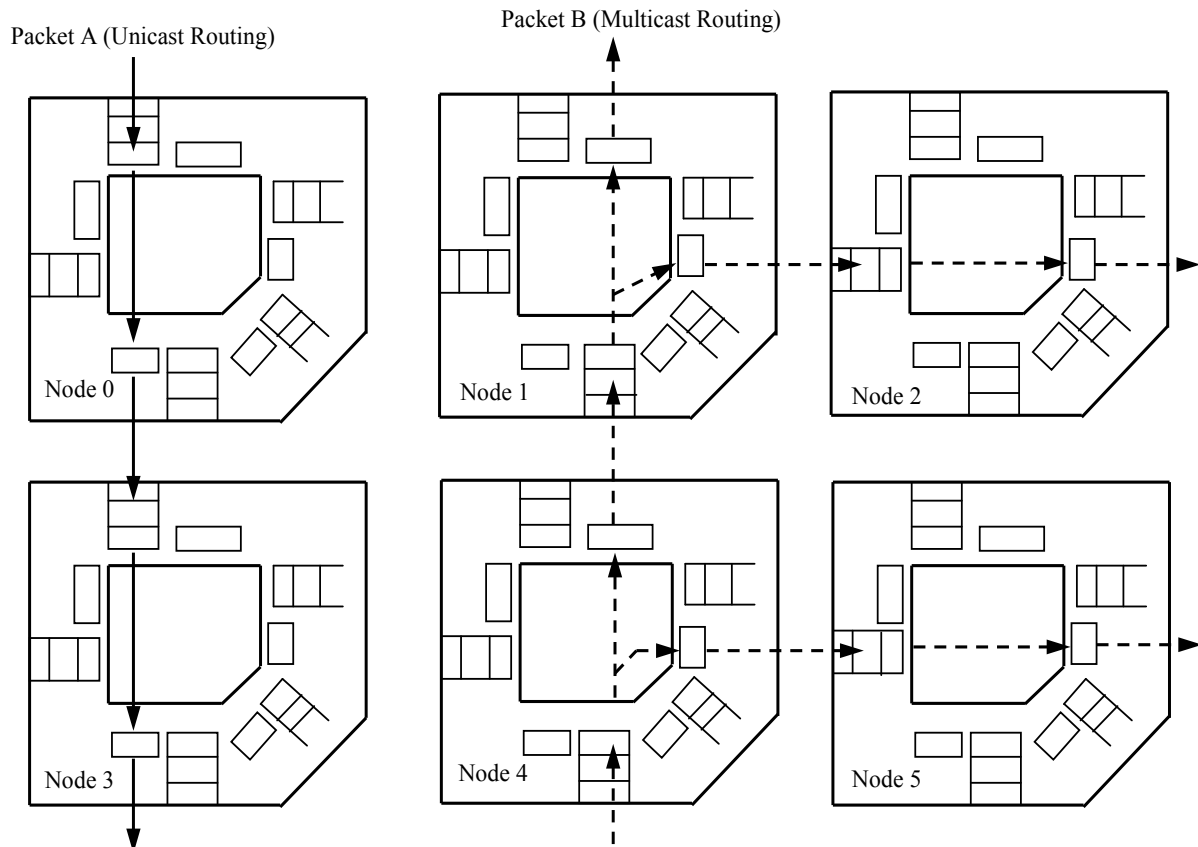


Figure 3 Employed Routing Scheme

4 Experimental results

The ‘C’ programming language is employed to develop the 3D-NoC simulator, which is based on cycle accurate technique. The implemented 3D-NoC simulator utilizes hardware codec (encoder/decoder) based on Golomb-Rice algorithm, power aware routing scheme, interconnect channels, router and Network Interfaces (NI’s). To improve the performance of the proposed 3D-NoC simulator several techniques such as virtual and pipelined channel along with effective routing techniques are engaged in the target platform. The Data generation Ratio (DR) is used by the IP-core present internally to generate the data randomly. On the other hand, the data size is selected between 4 ÷ 1024 bytes randomly. The encoder developed based on Golomb-Rice algorithm will compress the data generated randomly followed by packetization of the compressed data by the Network Interface (NI).

The receiver (or) node at the destination will receive the compressed packet after the packet travel through the networks present on-chip. And at the receive side, the original data are obtained by reversing the process with

the help of decoder based on Golomb-Rice algorithm. The calculation of cycle time between the sender (time from the data generated) and the receiver (restoration of the original data) is used to evaluate the throughput of the network. In search of accurate result the process is continued until the data generation and diminishing value reach 5 000 000. The impact of the performance is evaluated for various bandwidths namely; 8-bit, 16-bit and 32-bit. In synthesizing scenario, Synopsys Design Vision is used. The 3D-NoC router along with Network Interface is developed and designed using Verilog HDL. The CMOS technology of 0,15 um is employed to synthesize the design. The delay of critical path, dynamic power consumed by hardware and area (gate count) of the hardware for the proposed 3D-NoC simulator with compression scheme is computed as 3,57 ns, 701,19 uW and 11588 respectively. The value varies for conventional 3D-NoC architecture (without compression scheme) as 3,56 ns, 697,81 uW and 10096 respectively. The power efficiency and performance of the proposed 3D-NoC simulator is estimated from data of critical path delay and consumed power.

Table1 Evaluation of throughput performance

DR	CB (bits)	4×4×3 mesh 3D-NoC (Unicast Routing)			4×4×3 mesh 3D-NoC (Multicast Routing)		
		TP (Mpps)		(u)-(v)/(u) %	TP (Mpps)		
		Without ENC/DEC(u)	With ENC/DEC(v)		Without ENC/DEC(x)	With ENC/DEC(y)	(x)-(y)/(x) %
0,1	8	3,38	4,08	20,71	3,92	4,63	18,11
	16	4,22	4,86	15,17	4,63	5,29	14,25
	32	7,96	9,10	14,32	8,21	9,22	12,30
0,2	8	3,32	4,06	22,29	3,55	4,22	18,87
	16	4,02	4,73	17,66	4,31	4,98	15,54
	32	7,55	8,85	17,22	7,84	8,85	12,88
0,3	8	3,02	3,68	21,85	3,39	4,03	18,88
	16	3,96	4,71	18,94	4,10	4,66	13,66
	32	7,06	8,30	17,56	7,46	8,50	13,94
0,4	8	1,22	1,50	22,95	1,89	2,22	17,46
	16	2,01	2,29	13,93	2,49	2,85	14,46
	32	2,72	3,16	16,18	3,05	3,47	13,77
Average		4,20	4,94	18,23	4,57	5,24	15,34

The throughput of the proposed 4×4×3 mesh NoC (with encoding and decoding) system for both unicast and multicast routing is compared with the conventional 4×4×3 mesh NoC (without encoding and decoding) system for performance evaluation is shown in Tab. 1. The improvement in the power efficiency of the proposed 4×4×3 mesh NoC (with encoding and decoding) for both unicast and multicast routing is shown in Fig. 4. Here DR stands for the data generation ratio, which is defined as the ratio of the random data generated by number of IP cores to IP core totally present in one clock cycle. The throughput of the NoC system is represented by TP and CB denotes the bandwidth of the physical channel. The energy consumption of the proposed scheme (both unicast and multicast) over the conventional architecture is calculated for evaluating the power efficiency as shown in Fig. 4.

The power efficiency and throughput of the network is improved in entire cases (i.e. DR= 0,1; 0,2; 0,3 and 0,4) as shown in Fig. 4. And Tab. 1. The average improvement

in power efficiency and throughput of the network in unicast routing is estimated as 9,25 % and 17,61 % respectively. In similar, the average improvement in power efficiency and throughput of the network for multicast routing is determined as 8,65 % and 14,66 %. An account for this, throughput improvement of the proposed system shows greater variation as the bandwidth size decreases as shown in Tab. 1.

Two additional clock cycles are needed for compression and decompression (i.e. one cycle for compression and one cycle for decompression). In account to this the End-to End packet latency is tabulated in Tab. 2. Hence, the proposed system can be effectively used in case of narrow bandwidth. The utilization of the encoder and decoder in proposed scheme has led to the trivial increase in power consumption and area. The overhead due to area and power in the proposed scheme is negligible if we consider the improvement in power efficiency and throughput of the networks.

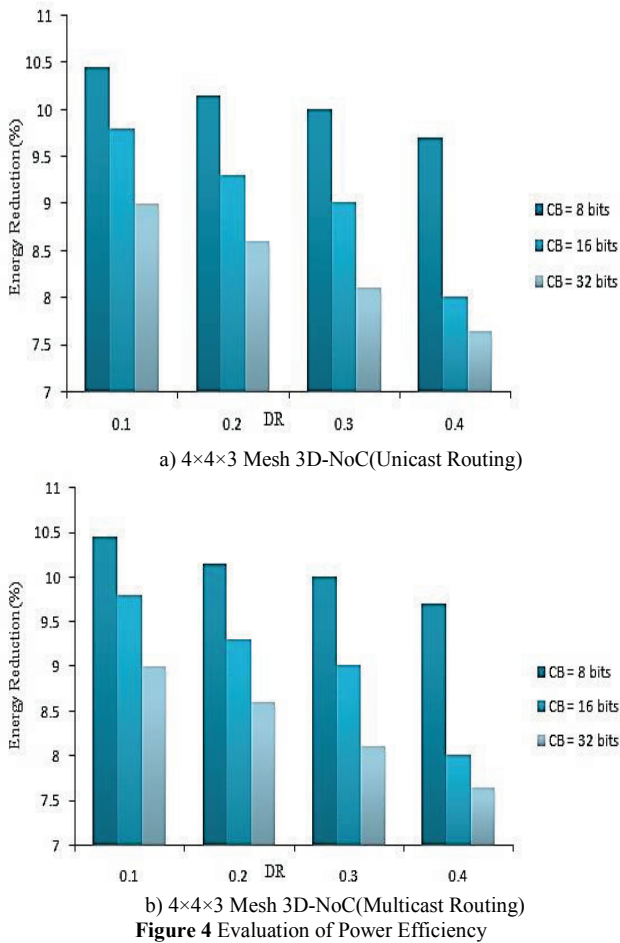


Figure 4 Evaluation of Power Efficiency

Table 2 End-to-End Packet Latency

DR	CB (Bits)	End-to-End Packet Latency (# of Cycles)	
		With Compression/ Decompression	Without Compression/ Decompression
0,1	8	17	16
	16	22	20
	32	30	28
0,2	8	24	23
	16	36	33
	32	43	41
0,3	8	61	59
	16	69	67
	32	76	73
0,4	8	87	85
	16	93	90
	32	107	105

6 Conclusion

A novel 3-Dimensional Network on Chip (3D-NoC) is designed to improve the throughput of the on-chip network. It is accomplished by utilizing the encoder and decoder for compressing and decompressing the data based on Golomb-Rice algorithm. The proposed system shows an improvement in power efficiency and throughput of the network with the small overhead in area compared to the conventional NoC (without encoding and decoding). As of now, the proposed scheme is evaluated for the 4x4x3 mesh topology; in the future the idea can be extended to different topologies for performance evaluation.

7 References

- [1] Benini, L.; Bertozzi, D. Network-on-chip architectures and design methods. // IEEE Computers and Digital Techniques. 152, (2005), pp. 261-272. DOI: 10.1049/ip-cdt:20045100
- [2] Pratim Pende, P.; Grecu, C.; Jones, M.; Ivanov, A.; Saleh, R. Performance Evaluation and Design Trade-Offs for Network-on-Chip Interconnect Architectures. // IEEE Transactions on Computers. 54, (2005), pp. 1025-1040. DOI: 10.1109/TC.2005.134
- [3] Rosinger, P.; Al-Hashimi, B. M.; Nicolici, N. Scan architecture with mutually exclusive scan segment activation for shift- and capture-power reduction. // IEEE Trans. Computer Aided Design Integr. Circuits Syst. 23, (2004), pp. 1142-1153. DOI: 10.1109/TCAD.2004.829797
- [4] Ogras, Y. U.; Marculescu, R. Analytical router modeling for networks-on-chip performance analysis. // Design, Automation and Test in Europe Conference, Acropolis, (2007), pp. 1096-1101.
- [5] Yeong Seob Jeong; Seung Eun Lee. Deadlock-free XY-YX router for on-chip interconnection network. // IEICE. 10, 20(2013), pp. 1-5. DOI: 10.1587/ielex.10.20130699
- [6] Moussa, H.; Baghdadi, A.; Jezequel, M. Binary de Bruijn interconnection network for a flexible LDPC/turbo decoder. // IEEE International Symposium on Circuits and Systems (ISCAS), (2008), pp. 97-100. DOI: 10.1109/iscas.2008.4541363
- [7] Saeidi, S.; Khademzadeh, A. Mehran, A. SMAP: An Intelligent Mapping Tool for network on Chip. // ISSCS 2007, PID360581, Iasi, Romania, (2007).
- [8] Kao, Y. H.; Chao, H. J. BLOCON: a bufferless photonic clos network-on-chip architecture. // 5th Conf. IEEE/ACM International Symposium on Networks on Chip (NoCS), (2011), pp. 81-88.
- [9] Feero, B. S.; Pande, P. P. Networks-on-Chip in a Three-Dimensional Environment: A Performance Evaluation. // IEEE Trans. Comput. 58, (2009), pp. 32-45. DOI: 10.1109/TC.2008.142
- [10] Lau, J. H. TSV Manufacturing Yield and Hidden Costs for 3D IC Integration. // 60th Conf. Electronic Components and Technology, Las Vegas, USA, (2010), pp. 1031-1042. DOI: 10.1109/ectc.2010.5490828
- [11] Rice, R. F. Practical Universal Noiseless Coding Techniques. // Technical Report 79-22, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, (1979).

Authors' addresses

N. Ashokkumar, *Research Scholar*,
 Faculty of Electronics and Communication Engineering,
 R. V. S. College of Engineering & Technology,
 R. V. S. Nagar, Dindigul - 624 005, Tamil Nadu, India
 E-mail: drashokkumarn@gmail.com

A. Kavitha, *Professor, Dr.*
 Faculty of Electronics and Communication Engineering,
 Chettinad College of Engineering & Technology,
 NH-67, Karur - Trichy Highways,
 Puliur CF, Karur, Tamil Nadu - 639114 India