A Novel Resource Scheduling Approach to Improve the Reliability of Shuffle-Exchange Networks

Farshad Mashhadi and Abu Asaduzzaman
EECS Department
Wichita State University
Wichita, Kansas, USA
{fxmashhadi@shockers. and abu.asaduzzaman@}wichita.edu

Muhammad F. Mridha
Department of Computer Science and Engineering
University of Asia Pacific
Dhaka, Bangladesh
Email: firoz@uap-bd.edu

***Abstract*—Approaches such as increasing the number of intermediate stages are introduced to increase the reliability and throughput of Multistage Interconnection Networks (MINs). However, they mainly try to change the network architecture to achieve the goal of having more reliable network. When multiple sources in such a network try to send data, collision of packets and blocking problems are inevitable. Using existing networks, they cant be prevented completely and a multiple access protocol must be used to that end. Time division multiple access (TDMA) protocol can be used to overcome these problems. To improve the performance of this protocol, we propose an adaptive slot allocation approach using Monte Carlo random sampling method. This approach is applied to Shuffle-exchange network (SEN) and Shuffle-exchange network with one additional stage (SEN+). Results for 4000 simulation cycles using Network Simulator 2 (NS2) show that the new SENs perform better in terms of reliability and throughput compared to their regular types.**

***Index Terms*—Confidence interval, Monte Carlo, Multistage interconnection network, Shuffle-exchange network, Time division multiple access.**

I. Introduction

High computational power have always been demanding due to its key contribution in applications like long-range weather forecasting, image processing, human genome, condensed matter theory, chemical reaction simulation, speech recognition, and automated reasoning [1]. Increasing the performance of processing elements solely cannot lead to increase the computational capability, and parallel processing techniques also should be taken into consideration. By taking advantage of multiple processors working in parallel, we can improve the speed and computational power that is needed to overcome the aforementioned problems [2]. To provide the interconnection among processors, memory modules, and other peripherals, a parallel computer also needs sub-system communications. Interconnection networks provide these communication needs in a parallel computer [3]. Figure 1 shows a generic model of interconnection networks. Although dual core processors can operate using a shared bus, but more advanced interconnection networks are needed in future for thousands of cores working on a single chip [4,5]. Many topologies are proposed to improve these so called networks-on-chips (NoCs). Multistage interconnection networks (MINs) is one of them that uses crossbar switches to provide communication needs between processors and memory modules [1,2].

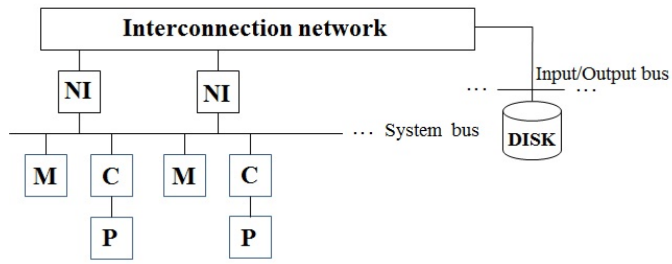

Fig. 1: Generic multiprocessor system with distributed memory.

Generally, in MINs inputs (source nodes) are connected to outputs (destination nodes) using one or multiple stages of crossbar switches and numerous inter-state links. Therefore, any communications between inputs and outpus are carried out through these switching elements [6]. These networks offer a good performance along being cost-effectiveness, and are mostly used in multiprocessor systems, communication networks, and modern embedded systems [1].

Crossbar switches of size 2×2 in MINs are mostly used in two possible states: Straight, in which upper input is connected to upper output, and lower input is connected to lower output and Exchange, that upper input is connected to lower output and lower input to upper output [2,3].

Different states of switching elements are managed by a centralized controller or by the individual switching element. The first strategy is being called centralized control and the second strategy is called distributed control. MINs are used in both circuit switching and packet switching networks with the introduction of buffered switches. Packet switching is efficient when messages are short and there are always packets to be sent. In contrast to circuit switching networks where the communication channels may be idle for a period of time, they can be fully utilized when inputs have packets for sending in a packet switching network [6].

According to their architecture, MINs provide a different number of distinct paths between any source/destination pairs. If there is only one path between a specific source/destination pair, the network is single-path and if there is more than one path the network is multi-path [7,8]. In designing the

978-1-5090-6004-7/17/$31.00 ©2017 IEEE

architecture of MINs, there are some important factors that should to be taken into consideration such as: performance, throughput, reliability, high availability, and lower latency to do multiple communications in parallel [1].

The rest of the paper is organized as follows. In Section II, motivation and study of related works are discussed. In Section III structure of SEN, routing model of the network, and a brief introduction to Monte Carlo method are covered. In section IV, the proposed approach is presented and in Section V, it is evaluated. Section VI summarizes and gives conclusions.

II. Background and Motivation

In this paper, our main objective is to use an adaptive TDMA protocol to allocate optimal time slots to source nodes, in order to make an existing type of MIN, SEN, more reliable.

Reliability of MINs is concerned with the ability of the network to accomplish the assigned tasks successfully, even when one or more failures happen [3,4]. When there is at least one available path between any source/destination pairs for packet transmission, the network is reliable [4,8,9]. Three approaches which are mostly used to increase the reliability of MINs are: increasing the number of switching stages, using a connection of multiple network in parallel, and replicating the network [8,9,10].

By adding more switching stages to MINs, we can increase the reliability and fault-tolerance capacity of MINs to some extent, by providing extra distinct paths between the source and destination nodes. However, increase in cost and network complexity is also inevitable in this approach [3,9,11].

Using multiplexers and de-multiplexers, multiple MINs can be connected in parallel arrangements with each other. Although using this approach it is possible to have a more reliable MIN, increase in cost and failure rate due to higher complexity in switching elements are also expected [2,4,10,12].

Another approach to improve the reliability of MINs is to replicate the network a number of times. Replicated MINs provide multiple paths between any source/destination pairs and clearly is more reliable than single-path MINs. Increase in costs and path length in these networks due to higher number of stages are inevitable [2,4,10,12].

Previous researches have shown that adding one stage to the network leads to reliability improvement in MINs. For instance, SEN+ has one additional more stage and is more reliable compared to SEN, but SEN+2 which has two additional stages has the least reliability among the others [8]. However, reliability analysis of SEN using reliability block diagram has shown more reliability parameters for SEN+2 than SEN, but SEN+ still has the highest reliability [3].

Our main purpose in this paper is to reevaluate the reliability of SEN and SEN+, while a new adaptive TDMA protocol is used in these networks. The maxmimum number of packet delivery in a unit of time, and delay/throughput tradeoff help characterizing the performance of a multiple access protocol in a distributed system. This tradeoff reflects the effect of an increasing arrival rate on the average message delay, defined as the amount of time between the message transmission at the sending station and its successful reception at a destination station [13]. Chiou and Shyi in [14], have analyzed the average waiting time of an optical network while TDMA was used in this network. Using TDMA as a media access protocol, experimented optical network have shown an ideal fair behavior.

Based on the above discussion we can conclude that there is no definite solution to have a perfect MIN without changing the network architecture. However, there are other ways to improve the reliability of these networks which have not been tested yet. In this paper, we suggest a new approach to improve the reliability, throughput, and performance of SENs. To achieve this goal, we use TDMA protocol to utilize the available resources in the most efficient way. Using TDMA in a network, multiple sources can share the same channel to send packets in a same time span, by dividing the time span into a number of time slots. The users transmit in rapid succession, one after the other, each using its own time slot [15]. Various approaches of TDMA try to divide time into slots and only allow nodes that do not interfere with each other to transmit data in a given slot. In most current TDMA networks a coordinator pre-computes the transmission schedule [16]. No changes are made to network architecture since this approach tries to improve the performance of current SENs.

III. Structures of SEN and SEN+

SENs are a class of interconnection networks with unidirectional switches, in which channels and switches are unidirectional. Figures 2 and 3 show the topology of two 16×16 class of unidirectional SENs: SEN and SEN+.

A. SEN

In this sub-section, we discuss more about the structure of SEN andSEN+. SENs of size $N \times N$ have $(\log_2 N)$ stages, in which every stage has $(N/2)$ number of switches of size 2×2. A SEN of size 16×16 is shown in Figure 2. The network complexity of an $N \times N$ SEN is $[N/2\ (\log_2 N)]$ [3]. In this network, only one path exists between every source and destination pair.

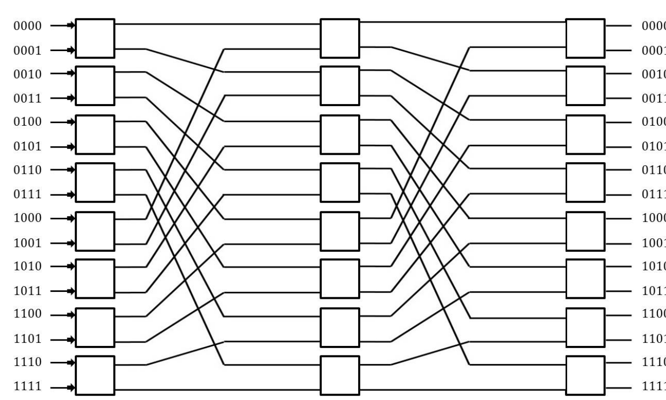

Fig. 2: SEN of size 16×16.

B. SEN+

An (SEN+), or Shuffle-exchange network with one additional stage of size $N \times N$ has $((\log_2 N)\ +1)$ stages, and each stage consists of $(N/2)$ switching elements of size 2×2. A SEN+2 of size 16×16 is shown in Figure 3. This network, has the complexity of $[N/2\ ((\log_2 N)\ +1)]$. In this network, there are two distinct paths between each source/destination pair [3].

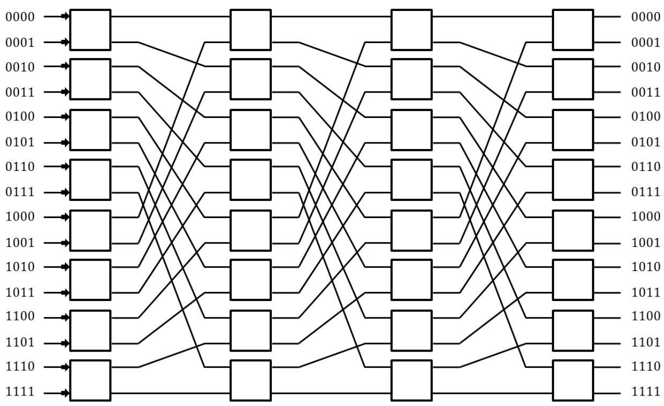

Fig. 3: SEN+ of size 16×16.

C. Network and routing model

Physical, switching and routing are three working layers in MINs. Packet transfer in these networks is done through via link-level and in some cases using routers. Switching layer has the control over the physical layer provide to help the packets reach the destination nodes. Routing the packet to the desired destination is also the duty of routing layer [6].

D. The basic Monte Carlo estimator procedure

We use Monte Carlo integration method for random sampling of a function to numerically compute an estimate of its integral. Consider the case that we need to integrate the one-dimensional function $f(x)$ from a to b:

$$F = \int_a^b f(x)dx \tag{1}$$

Finding the approximation of this integral is possible by averaging samples of the function f at uniform random points within the interval, which is used as the lengths of time slots. Given a set of N uniform random variables $X_i \in (a, b)$ with a corresponding probability density function or (PDF) of $1/(b-a)$, we can calculate F^N, which is the Monte Carlo estimator for computing F.

It should be noted that F^N is a function of X_i and it must be itself a random variable. This is an expression to say that F^N is the approximate of F, using N samples. In Equation (1), the estimator of Monte Carlo can compute the mean value of the function $f(n)$ over the interval a to b, by multiplying this mean with the length of the interval $(b-a)$. After moving $(b-a)$ into the summation, the estimator can be thought of as choosing a height at a random evaluation of the function and averaging a set of rectangular areas computed by multiplying this height by the interval length $(b-a)$ [17].

Quality of Monte Carlo integration should be improved by reducing the overall error and variance of the estimator. One powerful variance-reduction technique that addresses this problem is called stratified sampling. Using stratified sampling to divide $[a, b]$ into N number of sub-domains (or strata), a random sample is placed within the range of these intervals [18].

Predicting the length of the time slots needed for every node, especially when the number of them increases, is an NP-hard problem. In MINs with so many inputs, communication overheads between source nodes and the network controller can cause tremendous delay and effect the network performance negatively. The proposed approach uses Monte Carlo method to predict the time slot length for any source nodes and creates a new optimized TDMA frame. Numbers generated from Monte Carlo method, using confidence intervals are based on random sampling of time slot lengths and are evaluated for being accurate. This process is repeated for a specific number of times until it reaches the level in which it can generate more realistic samples.

IV. Proposed Resource Scheduling Algorithm

Monte Carlo method is able to predict almost exact numbers for time slot lengths needed for each of the source nodes. Exact prediction values are updated to the confidence interval. Every source node has its own confidence interval that shows the range of time slot length for that specific node. If the predicted time slot length is not enough, or the source node does not have packets to be sent, they send an acknowledgement to the network controller informing this. So we can say that after every cycle that these confidence intervals are updated, Monte Carlo method can generate more realistic numbers. Results from 4000 replications indicated that the computed confidence interval is becoming more reliable than a small number of replications.

The whole of TDMA frame includes information frame slot, assignment slot, and time frame slot, as shown in Figure 4 [19]:

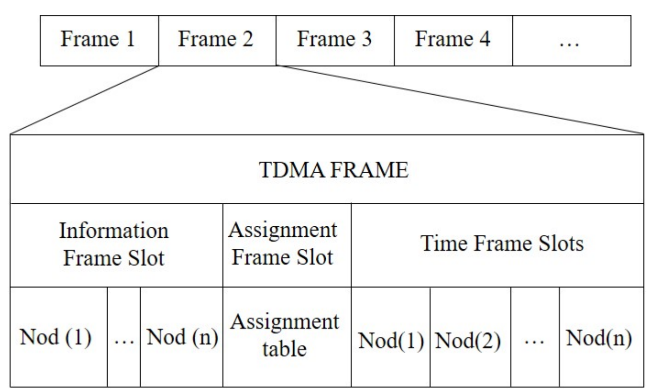

Fig. 4: Optimized TDMA structure.

There are of N number of information slots in every frame, and each of nodes occupies its own time slot. The information slot is generated by each source node, which contains the transmission information need to be sent in the MAC queue.

Assignment slot is used for sending time slot allocation table which generates based on the information slot generated from source nodes, and time slot length generated from Monte Carlo method.

After assignment slot, time frame slots begin immediately. Source nodes read slot allocation table, determine the transmit time length they are assigned in this cycle. The length of the time slot can be adjusted according to the allocation table. A node can send their packets in each time slot.

When the network is initializing, the network controller is fully responsible for the allocation of time slots to the source nodes. Network controller adds the information slot N at the beginning of the next cycle. Each of the source nodes generates information statistical package according to their transmission information (source/destination address) in each information slot. In the time frame slot, a node can occupy a plurality of time slots considering the number of data packets to be transmitted, so in the data frame using a dynamic TDMA. This cycle repeats after the algorithm converges and it begins to generate more reliable and realistic numbers. The whole slot allocation process is shown in Figure 5.

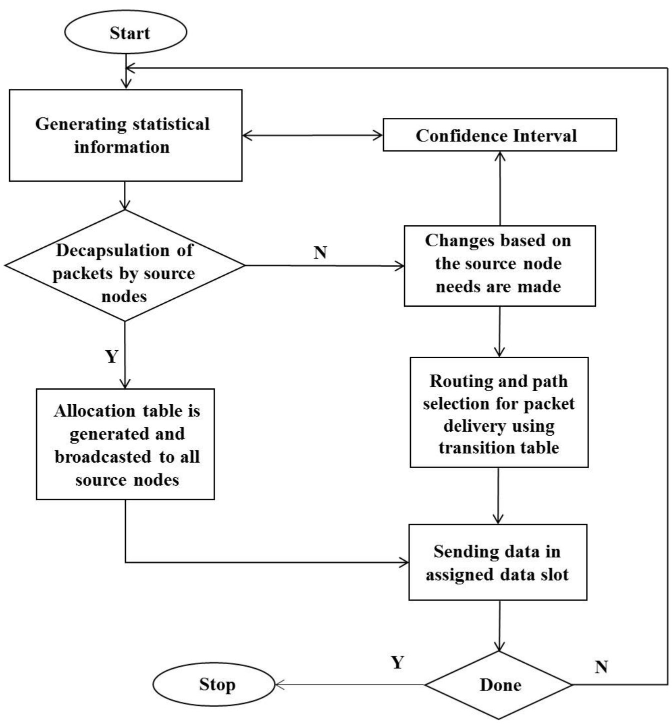

Fig. 5: The proposed approach flow.

V. Evaluation

In this section we evaluate our proposed approach (resource scheduling in SENs using adaptive TDMA) and compare it to the regular type of SENs. A suitable simulation platform for computer networks, NS2, is used as the simulation tool. In the following subsections, we discuss in detail about assumptions, simulation cases, and experimental results.

A. Assumptions

The SENs of size 16×16, with identical switching elements of size 2×2 is used in this study and the central controller establishes the path from inputs to outputs. The wormhole method is used for packet delivery at the flit level with the packet size of 64bit, where the routing information is in the first flit. In our simulation model, we consider no time restriction for a packet to leave the router, if both buffer and also the output are free [6]. Following assumptions are also made to facilitate the simulation process:

- Nodes in these networks are not completely reliable (except source and destination nodes).
- Node failures are random and independent.
- All of the switching elements have the same reliability.
- Traffic demand between 1 to 1024 kbps is randomly generated at every simulation cycle (to imitate the unbalanced of multiprocessor systems).

B. Simulation scenario

We use 16 different cases at every cycle of simulation, to find results for network throughput, delay, and terminal reliability. Every source node sends packet to its corresponding destination node at every simulation cycle, as it is described in Table I. It should be noted that desired destination node changes randomly for every source node at the beginning of every cycle. This procedure applies to all of the SENs.

TABLE I: Source/destination selection process.

Source Node#	Destination Node	Case 1	Case n
0000	RANDBETWEEN(0000, 1111)	1110	TBD
0001	RANDBETWEEN(0000, 1111)	0101	TBD
0010	RANDBETWEEN(0000, 1111)	1100	TBD
0011	RANDBETWEEN(0000, 1111)	0011	TBD
0100	RANDBETWEEN(0000, 1111)	1001	TBD
0101	RANDBETWEEN(0000, 1111)	0111	TBD
0110	RANDBETWEEN(0000, 1111)	0110	TBD
0111	RANDBETWEEN(0000, 1111)	1101	TBD
1000	RANDBETWEEN(0000, 1111)	0001	TBD
1001	RANDBETWEEN(0000, 1111)	0010	TBD
1010	RANDBETWEEN(0000, 1111)	1011	TBD
1011	RANDBETWEEN(0000, 1111)	1111	TBD
1100	RANDBETWEEN(0000, 1111)	0000	TBD
1101	RANDBETWEEN(0000, 1111)	1010	TBD
1110	RANDBETWEEN(0000, 1111)	1000	TBD
1111	RANDBETWEEN(0000, 1111)	0100	TBD

Since communication speed in multiprocessor systems are high, we assume a TDMA overall frame of 1 second, and since we have 16 source nodes, a fixed slot time of 62.5 mS to support 16 source nodes every 1 second. The Frame will repeat every 1 second and assuming the TDMA fixed time slot is used, every second each of the 16 nodes has 62.5mS of channel time available to send data packets. At the beginning of the sampling process, numbers generated are close to 62.5. These slot lengths begins to change for different source nodes and after running the simulation for more than 1000 cycles, more realistic numbers (lengths) are assigned to the source nodes, using the confidence interval.

C. *Experimental results*

As discussed in previous section, SEN and SEN+ are investigated to find out how the proposed resource scheduling approach affect the network performance. Network throughput and terminal reliability of SEN and SEN+ are experimented and results are compared to the regular type of the networks.

Terminal reliability is defined as the existence of at least one available path between every source/destination pair at every simulation cycle. It is also helps us to measure the accuracy of predicted time slot length and efficiency of proposed resource scheduling approach used in these networks.

In order to evaluate terminal reliability of these networks, for different reliability values of switching elements, 4000 simulation cycles are conducted. Reliability of switching elements in each of the 4000 simulation cycles is considered identical, ranging from 0.92 to 0.99. Figure 6 shows the results for terminal reliability experiment. We can see that SEN+ with TDMA is slightly better than the other types of network in terms of terminal reliability, for every reliability values. Both SEN+ and SEN show better values for terminal reliability, better than regular SEN+ and SEN.

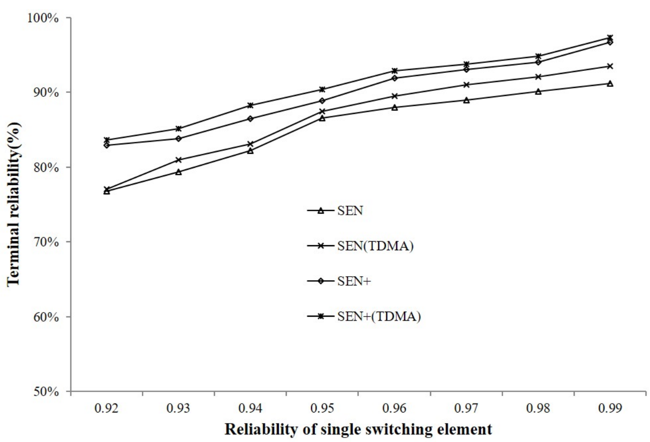

Fig. 6: Terminal reliability analysis with respect to single switches different reliability.

As discussed earlier, we can evaluate the efficiency of the proposed resource scheduling approach using values from terminal reliability experiment. As an example, for reliability value of 0.99 in SEN+ (TDMA), the average terminal reliability is 98.91% for 4000 simulation cycles. It means that 98.91% packets have the available resources to reach their destinations in that particular simulation cycles.

Figure 7 shows the effects of proposed resource scheduling approach on the throughput of the networks. Network throughput is defined as the average amount of data packets that are delivered to the destinations in a unit of time [20]. Results from simulation of the proposed approach show that the new SENs performs better in terms if throughput, above regular SENs. This gradual increase in network throughput begins after the confidence intervals convergs and more realistic time slot lengths are generated, as we have discussed in Section IV.

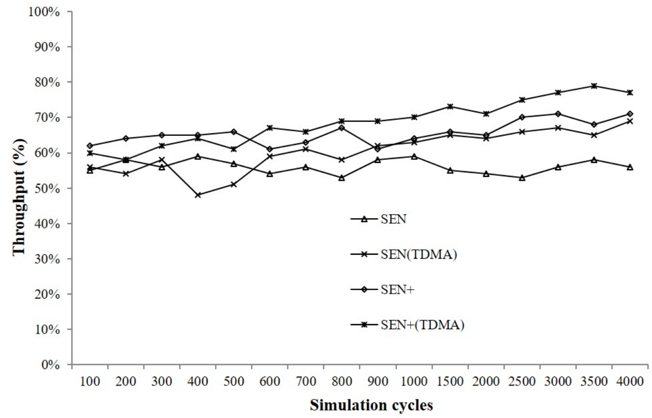

Fig. 7: Network throughput with respect to number of simulation cycles.

In this experiment, we can see the improvement of network throughput on the new SENs. According to the Figure7, SEN+ with TDMA has the network throughput of 77% when the simulation process ends at the 4000th cycle. Compared to the regular SEN+ which shows network throughput of 71%, SEN+ (TDMA) performs better, where regular SEN shows the least percentage of successful packet delivery. Although SEN+ provides 2 distinct paths between every source/destination nodes and is expected to have more throughput, our proposed approach could also increase the throughput of this network by 6%.

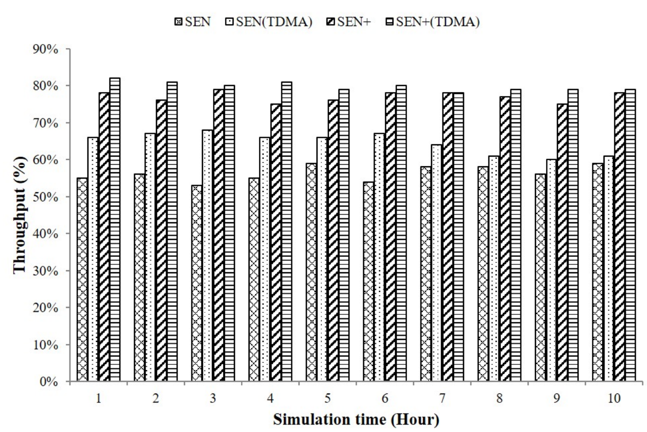

Fig. 8: Effects of synchronization problem on network throughput.

In Figure 8, effects of the synchronization problem associated with TDMA protocol is shown. One problem of TDMA protocol is synchronization of packets and nodes, which causes delay and less throughput in long-term simulation cycles [21]. In TDMA based communications, nodes must be synchronized in order to know when to send and receive data. In cases like a transient fault in which nodes lose their synchronization, they no longer know their position in the TDMA communication. Therefore, it would be impossible for the node to ascertain its message sending time. As it has shown in Figure 8, this synchronization problem effects on network throughput of SENs with TDMA to decrease slightly after a few hours of simulation time, but it is still better than the regular networks.

VI. Conclusion

In this paper, a new adaptive TDMA optimized frame structure based on Monte Carlo random sampling method is proposed. Monte Carlo helps us to predict an estimate of time slot lengths for every source node in every simulation cycle. As the Monte Carlo method is based on randomized sampling using confidence interval, results become more realistic after every simulation cycles. To solve the blocking problems caused by lack of resource scheduling, we use an optimized TDMA frame to allocate time slots to source nodes in a well-known type of MINs, Shuffle-exchange network.

Results show that by using this approach, network reliability and throughput of SENs with TDMA increase compared to regular networks. SEN+ with TDMA shows better results compared to the other networks, in both reliability and throughput measurements. Therefore, we can conclude that using the adaptive slot allocation approach, we can create more efficient SENs with better resource scheduling strategy. Furthermore, since the proposed approach does not change network topology by adding switching elements, it can be considered cost-effective.

However, synchronization of packets and nodes in the network is one of the problems associated with TDMA protocol. Since this problem can affect the network performance in long-term, our future study must be focused on solving the synchronization problem.

References

[1] I. Gunawan, Fundamental of Reliability Engineering: Application in Multistage Interconnection Networks, pp. 37-85, New Jersey, USA: John Wiley and Sons, Inc, 2014.

[2] F. Bistouni, and M. Jahanshahi, Pars network: A multistage interconnection network with fault-tolerance capability, J. Parappel Distrib. Comput., vol. 75, pp. 168-183, 2014.

[3] F. Bistouni, and M. Jahanshahi, Analyzing the reliability of shuffle-exchange networks using reliability block diagrams, Reliability Engineering and System Safety, vol. 132, pp. 97-106, 2014.

[4] J. T. Blake, and K. S. Trivedi, Reliability analysis of interconnection networks using hierarchical decomposition, Reliab IEEE Tran, vol. 38, pp. 111120, 1989.

[5] D. P. Agrawal, Graph theoretical analysis and design of multistage interconnection networks, IEEE Transactions on Computers, vol. 100, pp. 637-648, 1983.

[6] J. Duato, S. Yalamanchili, and L. NI, Interconnection Networks: An Engineering Approach, pp. 68-107, San Francisco, USA: Morgan Kaufmann, 2003.

[7] P. Newman, Fast packet switching for integrated services, University of Cambridge, UK, 1989.

[8] I. Gunawan, Reliability analysis of shuffle-exchange network systems, Reliability Engineering and System Safety, vol. 93, pp. 271-276, 2008.

[9] J. T. Blake, and K. S. Trivedi, Multistage interconnection network reliability, IEEE Transactions on Computers, pp. 1600-1604, 1989.

[10] F. Bistouni, and M. Jahanshahi, Improved extra group network: a new fault-tolerant multistage interconnection network, Journal of Supercomputing, vol. 69, pp. 161-199, 2014.

[11] G. Adams, and H. Siegel, The extra stage cube: A fault-tolerant interconnection network for supersystems, IEEE Transactions on Computers, vol. 100.

[12] P. K. Bansal, K. Singh, and R. C. Joshi, Reliability and performance analysis of a modular multistage interconection network, Microelectronics reliability, vol. 33, pp. 529-534, 1993.

[13] A. Boukerche, and T. Dash, Performance evaluation of a generalized hybrid TDMA/CDMA protocol for wireless multimedia with QoS adaptions, Computer Communications, vol. 28, no. 12, pp. 1468-1480, 2005.

[14] R. N. Chiou, and C. N. Shyi, The performance of optical networks with TDMA, in International Conference on Telecommunications, Marrakech, Morocco, 2009.

[15] G. Miao, J. Zander, K. W. Sung, and S. B. Slimane, "Fundamentals of Mobile Data Networks," 2016.

[16] R. Ouni, Dynamic slot assignment protocol for QoS support on TDMA-based mobile networks, Computer Standards and Interfaces, vol. 34, no.1, pp. 146-155, 2012.

[17] C. Rober, G. Castella, and I. Olkin, "Monte Carlo Statistical Methods, Springer," pp. 35-77, New York, USA: Springer, 2004.

[18] G. Fishman, "Monte Carlo: concepts, algorithms and ", pp. 145-254, New York, USA: Springer, 1996.

[19] G. Jia, Y. Wang, F. Meng, G. Y, G. Wang, "An adaptive TDMA slot assignment algorithm for scalable wireless mesh networks," Applied Mechanics and Materials, vol.392, pp. 872-875, 2013.

[20] R. Kaur, R. K. Bansal, and S. Bansal, Performance and reliability analysis of irregular multistage interconnection networks, in International Conference on Next Generation Computing Technologies, Dehradun, India, 2015.

[21] T. Nandagopal, T. E. Kim, X. Gao, and V. Bharghavan, Achieving mac layer fairness in wireless packet networks, in MobiCom, Boston, Massachusetts, 2000.