AN ESTIMATION OF TIME REQUIRED FOR MODELING OF AN ALGORITHM CALCULATE A NON-CONFLICT SCHEDULE FOR CROSSBAR SWITCH NODE BY MEANS OF GRID-STRUCTURE

Tasho Tashev

Abstract: The problem of calculating a non-conflict schedule by packets commutation in crossbar switch node is one of the foremost problems at the stage of node design. From a mathematical point of view this task is NP-complete. Constantly rising levels of traffic communication require developing of new algorithms. These algorithms must be correctly compared with known algorithms. In this paper we presented the investigations on the time execution of known PIM-algorithm for crossbar switch by means of CERN's grid-structure. By computer simulation of switching and using synthesized Generalized Nets (GN) model of the PIM-algorithm the execution time is obtained. Its assessment is based on the modeling of the throughput in the presence of uniform incoming traffic. It is shown that direct receiving of the characteristics of simulation time with the required accuracy would require the introduction of strictly defined control points.

Keywords*: Modeling, Generalized Nets, Communication Node, Crossbar Switch, Algorithm, Simulation.*

ACM Classification Keywords*: B.4.4 Performance Analysis and Design Aids, C.2.1 Network Architecture and Design, C.4 Performance of Systems*

Introduction

In information exchange networks the essential nodes are commutation nodes called switches and routers. Crossbar packet switches route traffic from input to output where a message packet is transmitted from the source to the destination.

The randomly incoming traffic must be controlled and scheduled to eliminate conflict at the crossbar switch. The goal of the traffic-scheduling for the crossbar switches is to minimize packet blocking probability and packet waiting time and to maximize the throughput of packet through a switch [Elhanany, 2007]. So achieving a maximum throughput of the switch depends on the calculation of non-conflict plan for switching incoming packets.

The problem for calculating of non-conflict schedule is NP-complete [Chen et al, 1990]. Algorithms are suggested which solve the problem partially. The origin of a series of parallel algorithms is the PIM-algorithm (Parallel Iterative Matching) [Anderson et al, 1993]. One of the research directions is working on modifications to PIM-algorithm, relying on input buffering with virtual output queuing (VOQ) [Guannan Qu et al., 2010]. Other studies are directed to the use of inputs and intermediate buffering (CICQ) [Shunyuan Ye et al., 2010].

Cellular automata, neural networks, etc. are used as formal means to describe and study the characteristics of crossbar switch nodes. In this investigation the apparatus of Generalized Nets (GN) are used as a powerful modern tool for formal modeling of parallel processes. Generalized nets (GN) [Atanassov, 1991, Atanassov, 1997] are a contemporary formal tool created to make detailed representation of connections between the structure and temporal dependencies in parallel processes. They are used in different fields of application, telecommunication is one of them [Gochev, 2008], [Tashev, 2010]. The apparatus of GN in this research is applied to synthesize a model of one algorithm for computing of non-conflict schedule in the crossbar switch node.

In this paper we presented the investigations on the time execution of PIM-algorithm for crossbar switch. By computer simulation of switching and using synthesized GN-model of the PIM-algorithm the execution time is obtained. Its assessment is based on the modeling of the throughput in the presence of uniform distributed incoming traffic. For this purpose, two templates are used to simulate uniform traffic.

Generalized Net Model of PIM-Algorithm

The requests for transmission through switching n x n line switch node is presented by an $n \times n$ matrix T, named traffic matrix (n is integer). Every element t_{ij} ($t_{ij} \in$ [0, 1, 2, …]) of the traffic matrix represents a request for packet from input i to output j. For example $t_{ij} = 4$ means that four packets from the i^{th} input line have to be send to j^{th} output line of the switch node, etc.

It is assumed that a conflict situation is formed when in any row of the T matrix the number of requests is more than 1 – this corresponds to the case when one source declares connection with more than one receiver. If a column of the matrix T hosts more than one digit 1, it indicates a conflict situation. Avoiding conflicts is related to the switch node efficiency [Elhanany, 2007].

The PIM-algorithm is based on the principle of distributed-random choice. Its informal description has three phases. 1) **Request**: Every input sends request to every output for which it has a packet for transmission; 2) **Grant**: Every output chooses randomly one of the received requests and grants permission for sending to the corresponding input; 3) **Accept**: Every input received grants chooses randomly one of them. This packet will be accepted for commutation.

Inputs execute in parallel the first phase. Outputs execute in parallel second phase. Inputs are working in parallel in the third phase [Anderson et al, 1993]. This parallelism is suitable for applying of GN

The PIM-algorithm can be described formally by the means of Generalized Nets. Based on a previous work [Tashev, 2010], here, we give an explicit form of the VOQ The model is developed for switch node with n inputs and n outputs. Its graphic form is shown on Figure 1.

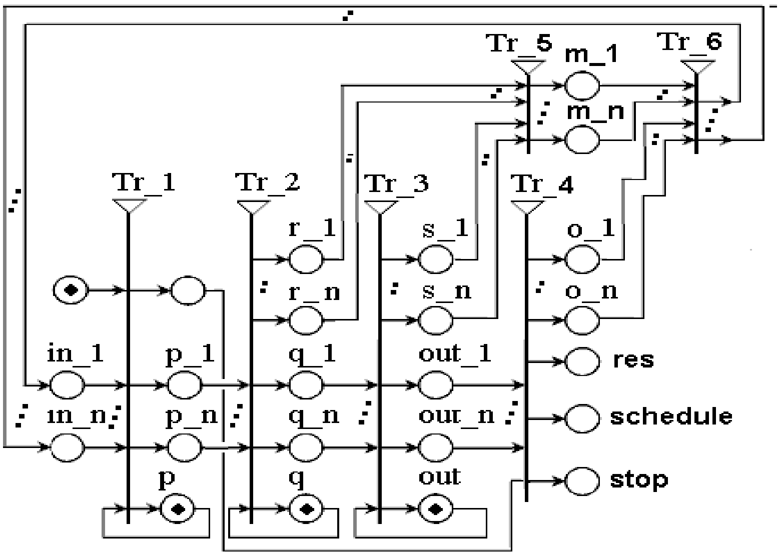

Figure 1: Graphical form of GN-model of PIM-algorithm.

The model has possibilities to provide information about the number of switching in crossbar matrix, as well as about the average number of packets transmitted by one switch. Analysis of the model proves receiving a non-conflict schedule. Calculation complexity of the solution depends on the power of three of the dimension *n* of the matrix *T* $(O(n^2))$.

Computer Simulation

The transition from GN-model to executive program is performed as in [Tashev and Vorobiov, 2007]. The program package Vfort of Institute of mathematical modeling of Russian Academy of Sciences is used [Vfort]. The source code has been tested by the IBM PC-compatible computer and then compiled by the means of the grid-structure of CERN (lxplus.cern.ch). The resulting executive code is executed in the grid under Scientific Linux SLC release 5.7. Main restriction for the choice of parameters in simulation (dimension *n* and type of load traffic) is the time for execution of the program.

Achieving maximal throughput of crossbar switch node depends on creation of non-conflict schedule for packet commutation. The first step while checking their efficiency is throughput modeling of the switch by uniform demand traffic. The matrix *T* defines a uniform traffic demand matrix if the total number of packets in each element in rows and columns are equal [Gupta and McKeown, 1999]. The uniform demand traffic matrix is called in the investigation as $Pattern_i$. The index *i* shows values of element in the traffic matrix. All elements in the traffic matrix are equal and in this case an optimal solution is known. The throughput is computed by dividing the result of optimal solution on the result of the simulated solution. The result of algorithm is a number of non-conflict matrices. Their sum is equal to *T*, as number of matrices shows number of commutations.

Figure 2 presents the used input data – *uniform matrix T*, defined by us. The first type of the matrix is called $Pattern_1$. Its specification is shown on the left of the figure 2. The optimal schedule requires *n* switching of crossbar matrix for *n* x *n* switch. The second type of the matrix is called $Pattern_i$. Its specification is shown on figure 2 (right). The optimal schedule requires (*i* . *n*) switching of crossbar matrix for *n* x *n* switch.

$$T=\underset{2\times 2}{\begin{bmatrix}1&1\\1&1\end{bmatrix}}\ \underset{3\times 3}{\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}}\cdots\underset{k\times k}{\begin{bmatrix}1&\cdots&1\\\vdots&\ddots&\vdots\\1&\cdots&1\end{bmatrix}}\cdots\qquad T=\underset{2\times 2}{\begin{bmatrix}i&i\\i&i\end{bmatrix}}\ \underset{3\times 3}{\begin{bmatrix}i&i&i\\i&i&i\\i&i&i\end{bmatrix}}\cdots\underset{k\times k}{\begin{bmatrix}i&\cdots&i\\\vdots&\ddots&\vdots\\i&\cdots&i\end{bmatrix}}\cdots$$

Figure 2: Types of the uniform traffic matrix T

The results from the computer simulation of the PIM-algorithm with input data $Pattern_1$ and $Pattern_{50}$ are displayed on figure 3 and 4.

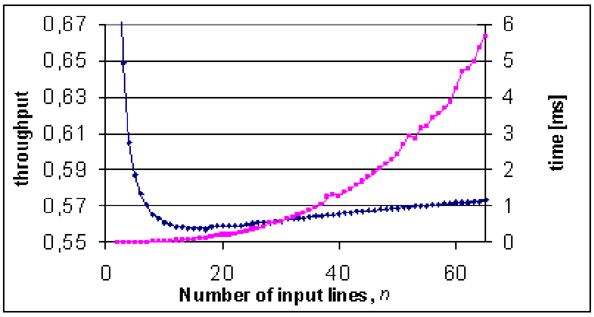

Figure 3: Throughput and time with $Pattern_1$

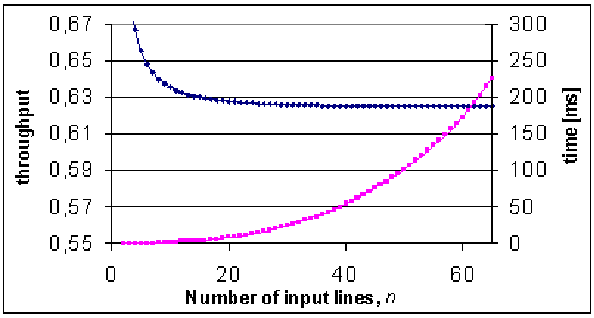

Figure 4:Throughput and time with $Pattern_{50}$

The crossbar matrixes of the size 2x2 up to 65x65 are simulated. It can be seen that when the size of *Pattern* and dimension of the switch field increases, the throughput asymptotically tends to the known boundary (~63,2 % [Gupta and McKeown, 1999]). Therefore the simulation is correct. Evaluation of the simulation results illustrates that PIM algorithm has high sensitivity to the increasing of input buffer.

The results of the approximated time with input data $Pattern_1$ and $Pattern_{50}$ are demonstrated on figure 5 and 6. We expect that the time coefficient tend to 3 since the simulation execution is not parallel.

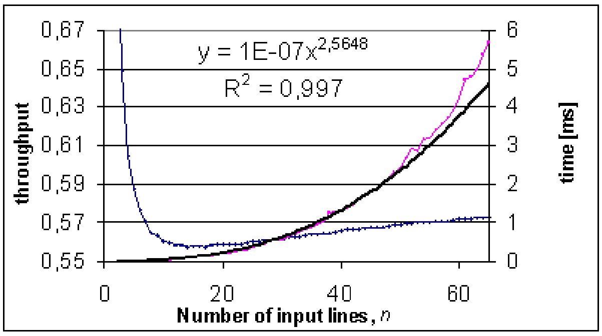

Figure 5: Approximation for time with $Pattern_1$

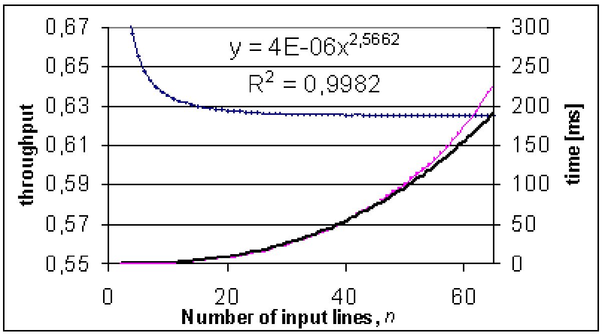

Figure 6: Approximation for time with $Pattern_{50}$

The figures give evidence of a difference between throughput and time in cases of $Pattern_1$ and $Pattern_{50}$ – the input buffer increases 50 times. The difference in time in the cases of $Pattern_{50}$ and $Pattern_1$ is increasing 40 times (from 1.10^{-7} sec to 4.10^{-6} sec) as opposed to the expected 50 times. The time coefficient is ~2.6 as opposed to the expected 3. The approximation is not very good.

If we reject the results for n = 2,..., 9, the difference between time for $Pattern_{50}$ and $Pattern_1$ is increasing 50 times (from 6.10^{-8} sec to 3.10^{-6} sec) - exactly as it should. The figures 7 and 8 shows the time approximation for this case. The time coefficient is increasing to ~2.7.

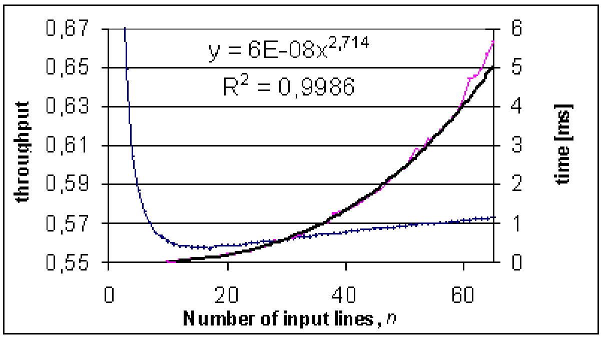

Figure 7: Correction for approximation of P_1

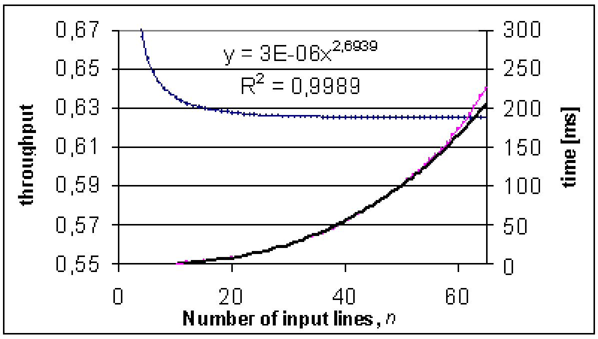

Figure 8: Correction for approximation of P_{50}

Let us increase the dimensionality of the simulation to 130. The results from the computer simulation of the PIM-algorithm with input data $Pattern_1$ for size *n*=2x2 up to 130x130 are displayed on figure 9.

The time coefficient is increasing to ~2.8. But the real data for time execution are still up from approximated data.

Let us reject the results for n = 2,..., 19. In this case y=3E-08x 2.8911. The time coefficient is ~2.9. Let us reject the results for n = 2,..., 29. In this case y=3E-08x 2.9343. The time coefficient is ~2.93.

It is shown that direct receipt of the characteristics of simulation time with the required accuracy would require the introduction of strictly defined intervals of size of simulation.

The promising results from the simulation on the PIM algorithm lead us to an idea of conducting of large-scale simulations for non-uniform demand traffic that can be direction for the future work.

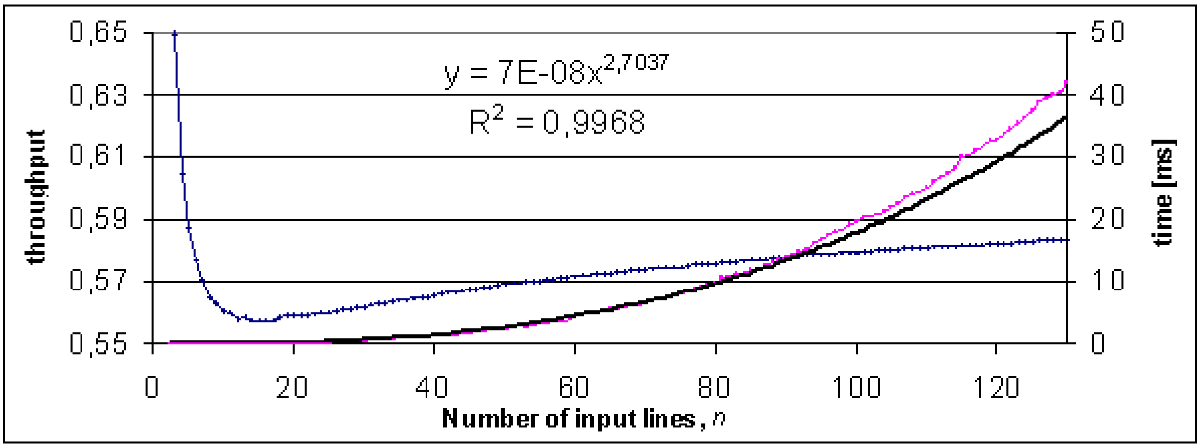

Figure 9: Throughput and time approximation for n=130x130 with Pattern₁

Let us reject the results for n = 2,..., 9. Figure 10 show the time approximation for this case.

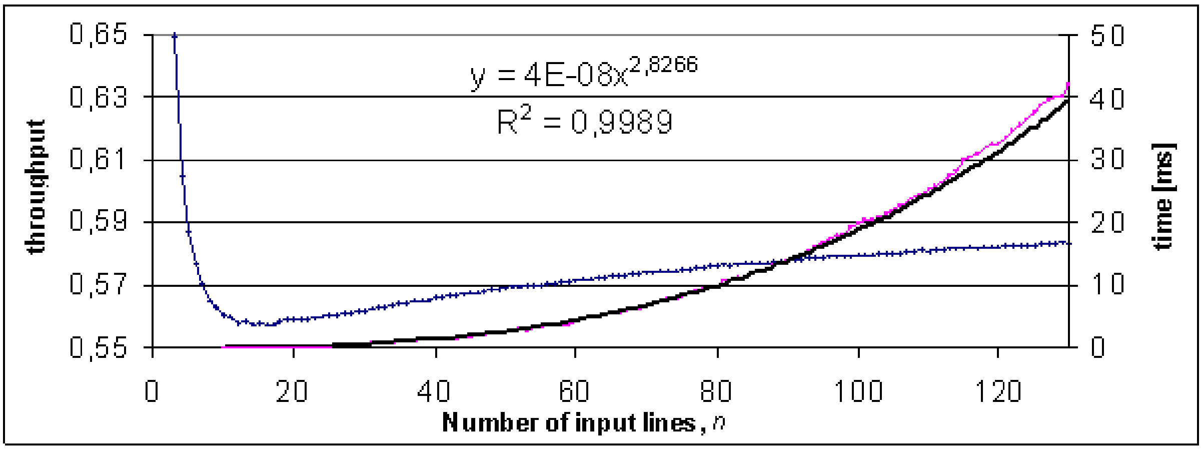

Figure 10: Coorection for approximation for time n=130x130 with Pattern₁

Conclusion

In this paper the investigations on PIM-algorithm for calculating a non-conflict schedule for crossbar switch node are presented. Computer simulations of a Generalized Nets-based model of PIM-algorithm performing uniform load traffic have been carried out. The results of simulation such as throughput and execution time are evaluated. It is shown that direct receipt of the characteristics of simulation time with the required accuracy would require the introduction of strictly defined control points.

Future work should be directed towards carrying out large-scaled computer simulation to study the throughput and execution time for a wide range of incoming demand traffic.

Acknowledgement

The author thank Assoc.Prof. PhD. Leandar Litov, chair of Department of Atomic physics, Faculty of Physics, Sofia University "St. Kliment Ohridski", Bulgaria, for providing necessary computer resource for CERN-grid.

Bibliography

[Elhanany, 2007] Elhanany, I., Hamdi, M. High-performance packet switching architectures. Springer-Verlag London, 2007.

[Chen et al, 1990], W., Mavor, J., Denyer, P., and Renshaw, D. Traffic routing algorithm for serial superchip system customisation. IEE Proc 137:[E]1, 1990.

[Anderson et al, 1993] Anderson T., Owicki S., Saxe J., and Thacker C. High speed switch scheduling for local area networks. ACM Trans. Comput. Syst., vol. 11, no.4, 1993, pp.319-352.

[Guannan Qu et al., 2010] Guannan Qu, Hyung Jae Chang, Jianping Wang, Zhiyi Fang, Si-Qing Zheng, Designing Fully Distributed Scheduling Algorithms for Contention-Tolerant Crossbar Switches, Proceedings of IEEE Conference on High-Performance Switching and Routing (HPSR 2010), pp. 69 - 74

[Shunyuan Ye et al., 2010] Shunyuan Ye, Yanming Shen, Shivendra Panwar, DISQUO: A Distributed 100% Throughput Algorithm for a Buffered Crossbar Switch, Proceedings of IEEE Conference on High-Performance Switching and Routing (HPSR 2010) , pp. 63 - 68

[Atanassov, 1991] Atanassov K. Generalized Nets. World Scientific, Sing., N.J., London, 1991

[Atanassov, 1997] Atanassov K. Generalized Nets and System Theory. Acad. Press “Prof.M.Drinov”, Sofia, Bulgaria, 1997.

[Gochev, 2008] Gochev V., Modeling of telecommunication traffic with Generalized nets. Proc. of National conference with international participation “Elektronika’2008”, 29-30 May 2008, Sofia, Bulgaria. pp. 88-93 (in Bulgarian).

[Tashev, 2010] Tashev T. Computering Simulation of Schedule Algorithm for High Performance Packet Switch Node Modelled by the Apparatus of Generalized Nets. Proceedings of 11th International Conference on Computer Systems and Technologies “ CompSysTech’10 ”, June 17-18 2010, Sofia, Bulgaria. ACM Press, ACM ICPS Vol. 471, pp.240-245. .

[Tashev and Vorobiov, 2007] Tashev, T., Vorobiov, V.: Generalized Net Model for Non-Conflict Switch in Communication Node. Proc. of Int. Workshop "DCCN'2007" Sept 10-12, 2007, Moscow, Russia. IPPI Publ., Moskow, 2007. Pp.158-163.

[Vfort] rhttp://www.imamod.ru

[Gupta and McKeown, 1999] Gupta, P., McKeown, N. Designing and Implementing a Fast Crossbar Scheduler. IEEE Micro, Jan-Feb 1999, pp. 20-28.

Authors' Information

Tasho Tashev, *Department "Optimization and Decision Making", Institute of Information and Communication Technologies – Bulgarian Academy of Sciences, “Acad. G. Bonchev” bl. 2 Sofia 1113, Bulgaria; e-mail: ttashev@iit.bas.bg*

Major Fields of Scientific Research: Distributed Information Systems Design, Methods and tools for net models researches