Performance of Switching Fabrics Used for the Scalable Router

Zbigniew Hulicki
Department of Telecommunication, AGH University of Science and Technology, ul. Czarnowiejska 78, Krakow, Poland

Keywords: Switching Fabric, Multistage Interconnection Network (MIN), Scalable High-Performance Router.

Abstract: The intention of this paper is to examine the performance aspects of switching fabrics which can be used for scalable high-performance routers. Topology and capabilities of switching fabrics are discussed, followed by an examination of their performance measures in diverse scenarios for possible operation of switching elements (SEs). It has been shown that the switching fabric based on the cube type multistage interconnection network (MIN) outperforms those based on the PM2I type MIN under both any traffic scenario and SE operation. The simulation results presented should be helpful in predicting the router performance before actual fabrication of the switching fabric.

1 INTRODUCTION

Routers and switches are viewed as the most critical parts of the current communication infrastructure and will be also needed to provide fast and efficient communication in next-generation networks. Recently there have been various efforts to design an efficient optical switch for use in a high-speed router (Chao, 2007). Besides, there is an immense interest in designing a simple and high performance switch which would satisfy the demands of an entirely new scenario for emerging broadband services.

In the past few decades, a number of switch fabrics have been proposed in literature and used in practical implementations to interconnect key components in routers, such as routing engines and line cards (Lien, 2010). Many of the proposed switch designs have been based on multistage interconnection networks (MINs). Various solutions, i.e. single or multiple panel (replicated), can be used in design of MIN based switches (Tzeng, 2004).

The purpose of this paper is to examine the issues dealing with the performance of single plane MINs used as switch fabrics for scalable high-performance routers and/or optical cross-connects OXCs. Topology and capabilities of switching fabric will be discussed first, followed by a specification of the switch model. Then, the performance and dependability of MIN based switch fabrics will be examined, taking into account diverse scenarios for possible operation of SEs. Lastly, there are some conclusions and remarks regarding the trade-offs between switch characteristics.

2 TOPOLOGY OF SWITCHING FABRIC

In selecting the architecture of switching fabric, four design decisions can be identified (Chao, 2007): operation mode, control strategy, switching method, and topology of MIN. The topology of a switching network is a key factor in determining a suitable fabric architecture (Skalis, 2010). MINs proposed for the switch fabrics are usually constructed using small crossbar switches organized in stages and may have a uniform or non-uniform connection pattern between stages. SEs can be buffered or non-buffered, but the use of self-routing MINs are favored because small delays can be achieved (Rongsen, 2007). Generally, the topologies of MINs tend to be regular and can be grouped into two categories: static or dynamic. The static MINs are simple to build, but fail even in the presence of a single fault. On the other hand, a dynamic MIN is able to reroute the data through a fault free path if the regular path is faulty or busy. Because the reliable operation of a MIN is an important factor in overall router performance, it is important to design a switching network that combines full connection capability – in spite of faults – with a slightly lower performance.

A number of diverse fabric solutions which offer redundant paths between the input and output ports

Hulicki Z..
Performance of Switching Fabrics Used for the Scalable Router.
DOI: 10.5220/0005033005670571
In *Proceedings of the 4th International Conference on Simulation and Modeling Methodologies, Technologies and Applications* (SIMULTECH-2014), pages 567-571
ISBN: 978-989-758-038-3

have been proposed to provide fast, efficient and reliable communications at a reasonable cost (cf. Chao, 2007). However, no single solution network is generally preferred because the cost-effectiveness of a particular design varies with such factors as its application, the required speed of data transfer, the actual hardware implementation of the switching network, the number of input/output ports, and the construction cost.

As far as MINs are concerned, they can be built using a variety of structures. Two significant examples of topological structures based on the single plane design are switching networks based on the cube $_i$ and PM2I interconnection functions (Lien, 2010). Therefore, there is an open question about the performance vs.-reliability trade-off that a multiple-path structure of the PM2I network might offer, for various distributed control algorithms and decision rules, when compared to a cube-type network.

3 THE SWITCH MODEL

To evaluate the properties of a fabric based on the given switching network (MIN), the following assumptions are made about the operation and the environment of the interconnection network: the single plane design of switching fabric is taken into account; because, in this contribution, we are focused primarily on the performance, the slotted traffic source model is used with a uniform random distribution of packet destinations; the switching network is operated synchronously, meaning that the packets are transmitted only at the beginning of a time slot given by the packet clock and each input link is offered the same traffic load; the buffering is external to the switch fabric, i.e. non-buffered switch fabrics formed from non-buffered SEs are under consideration. Therefore, the queuing effects are unaccounted in this model. All output ports perform the function of a perfect sync and a conflict is said to occur if more than one packet arrives at the same output link at the same time.

Two classes of packet switched MINs, one with redundant interconnection paths and one without are under consideration (i.e. the cube-type class of MINs known as unique-path networks and the PM2I class of MINs which have multiple paths between a given network input-output pair). It is clear that faults in the switching network often result in degraded performance of a fabric. Because failures occur at random and the combinations of the type, number and location of faults can also vary, the influence of failures on the performance of switching fabric is usually examined by simulation. These aspects have been already investigated (Hulicki, 2013) and hence they are not included in the model of switching fabric used in this contribution.

4 PERFORMANCE EVALUATION

Although recently a new (analytic) method for the performance analysis of multistage switching fabrics has been proposed (Hulicki, 2013a), in this paper, the performance of switching fabric, based on both of the aforementioned classes of packet switched MINs, has been studied using simulated experiments. A discrete time event-driven simulator has been designed to carry out simulations. Such a tool enables one to simulate the real operation of switch fabric as well as to estimate its performance measures. The simulator uses the switch model and the most commonly used performance measures. It enables performance evaluation of the fabric under different traffic and/or interconnection patterns, including both uniform and non-uniform traffic patterns. Moreover, in order to stop packets from entering the switching fabric once its resources are exhausted, the switching network of a router often employs a backpressure mechanism (Lien, 2010). Therefore, except for non-buffered MINs, the simulator also allows for a modification to the switching fabric, i.e. an implementation of buffering (introduced at the input/output of the switching network or to the SEs) or using different types of SEs (that allow for diverse switching functions), as well as an evaluation of the fabric’s dependability under diverse fault models, including combinations of the type, number and location of faults.

As it has been already mentioned, this contribution presents only selected results from the simulation experiments because the results point to a similar trend and the utilization of fabric, efficiency of switching and packet loss (or packet drop rate) can serve as good indicators for the performance of switching fabric. The simulation results have been shown and compared in a few figures (Figs. 2, 3, 4, 5, 6 and 7). Each data value given in these figures is the result after n_r = 100,000 clock cycles in the simulation (cf. Fig.1), where this number of cycles is found to yield steady-state outcomes (95% conf. lev.). If the number of cycles is less than 10,000 , the steady-state is not reached, so the simulation results cannot be valid (cf. lowest curve in Fig.1).

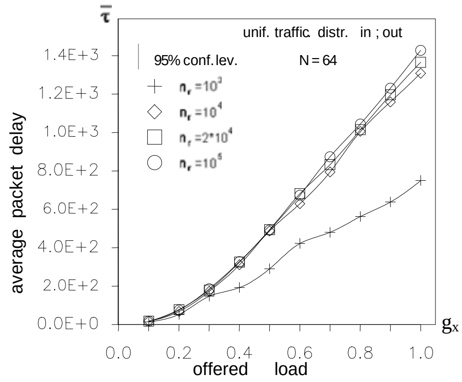

Figure 1: Comparison of packet delay for various simulation clock cycles.

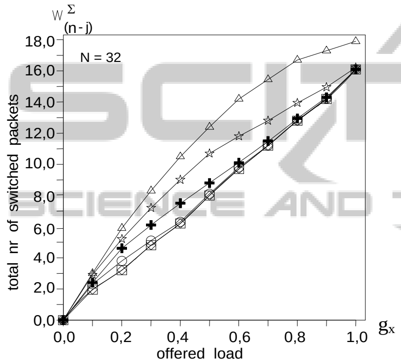

Figure 2: Total number of switched packets vs. traffic load.

The total number $W^{\Sigma}_{(n-j)}$ of switched packets at the outputs of the given ((*n-j*)–th) fabric stage versus traffic load offered to the input ports is demonstrated in Fig. 2. It is observed that increasing traffic load results in nonlinear increase of $W^{\Sigma}_{(n-j)}$, however along with that increase the inflexion of characteristic is also moved towards greater values of both parameters. The reason for that phenomenon is the cumulative effect of events occurring during a packet cross over the switching fabric.

Influence of SE operation mode on the performance of the cube type (MIN) fabric is shown in Fig. 3. One can see that implementation of switching function with the broadcast capability decreases fabric performance, even though it could potentially increase switching capabilities of SE. Implementation of such switching function increases packet drop rate at all stages in the fabric what deteriorates its performance measures.

As it has been already mentioned, the buffering is external to the switch fabric, i.e. non-buffered switch fabrics formed from non-buffered SEs are under consideration in this performance study. Hence, it was interesting to investigate the queuing effects at the inputs of switching fabric for its different size.

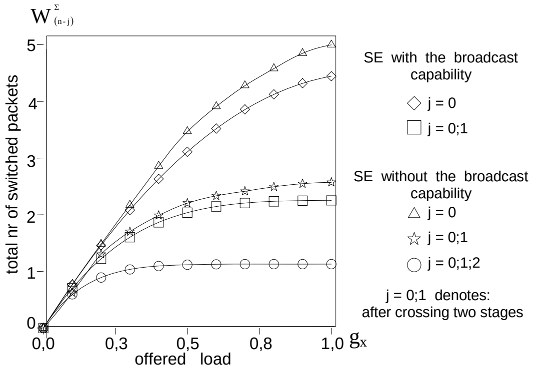

Figure 3: Influence of SE operation mode on the fabric performance.

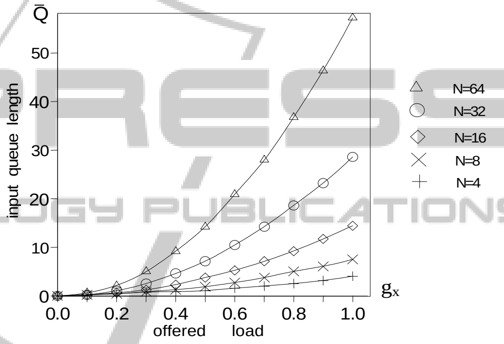

Figure 4: Queuing of packets at the input of switching fabric.

The input queue length versus offered load has been depicted in Fig. 4. One can observe that the queue length at the input of switching fabric increases nonlinearly along with that increase of traffic load and size of a fabric (MIN), but that increase intensifies for greater values of both parameters. This is result of packet collisions and lower probability to cross over a fabric (i.e. to be switched at the fabric output).

Usefulness of switching fabric in various scenario of its operation mode is illustrated in Fig. 5, i.e. the operation of switching fabric has been examined in both a circuit and a packet mode. In both cases diverse control rules were used for SEs of the fabric, i.e. *distributed* and two versions versions (*shift* and *flip*) of centralized control as well. Shift control means a group of SEs governed by a single control signal as opposite to the flip control whereby a single signal controls the whole stage of the cube type MIN.

It is clear (cf. Fig. 5) that usefulness of centralized control is questionable, because in both operation modes of the fabric such a control mechanism decreases fabric performance (i.e. for the

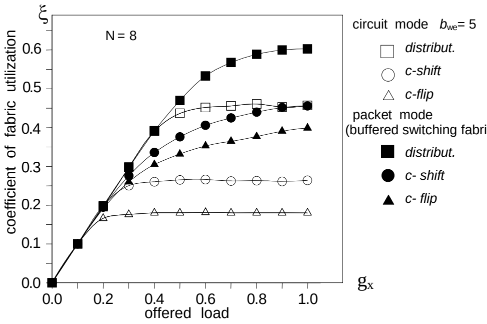

Figure 5: Usefulness of switching fabric in various scenario of its operation mode.

shift control, ~41% reduction in a circuit mode and ~25% for a packet mode; even worse performance is noticed for the flip mechanism). Hence, one can claim that the usage of MIN with centralized control can result in the congestion in switching fabric, i.e. such control mechanism can have a detrimental effect on the performance measures.

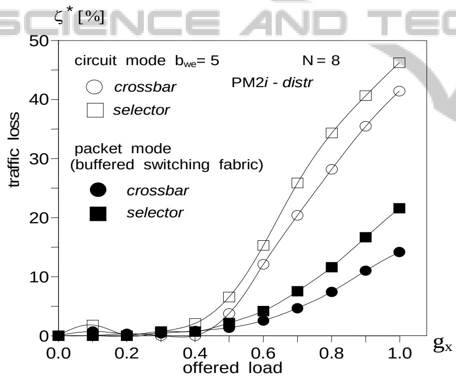

Figure 6: Packet loss vs. traffic load in the PM2I type MIN for various modes of SEs' operation.

On the other hand, there are architectures (the PM2I type MINs) which already have built-in redundancy (multiple paths between a given network input-output pair) that provides fault tolerance. Therefore, it was interesting to know what would be the effects on fabric performance when using different types of SEs (that allow for diverse switching functions). In consecutive tests, the impact of different SE type has been evaluated (cf. Fig. 6). One can observe a qualitatively similar nature of the performance measures (e.g. traffic loss) for both operation modes of the fabric. The traffic loss is higher if SEs can perform only one (*selector*) switching function, and the loss is lesser if all switching functions can be performed by SE simultaneously. This phenomenon is more clear when traffic load increases and it is result of the cumulative effect of events occurring during a packet cross over the switching fabric. Furthermore, it was also interesting to know what is the efficiency of switching when a fabric is based on various MINs. This aspect of performance evaluation has been depicted in Fig. 7.

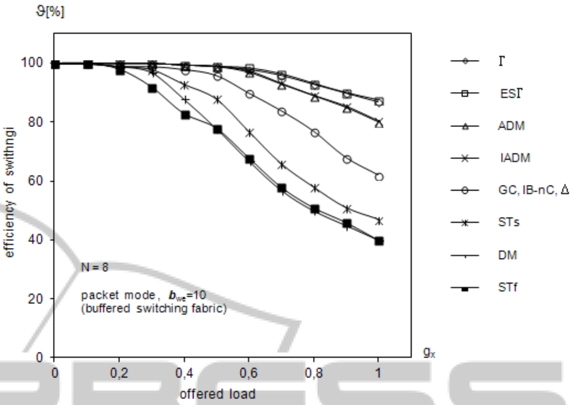

Figure 7: Efficiency of switching vs. traffic load for a fabric based on various MINs.

It is clear that capabilities of SEs as well as the architecture of switching network (MIN) can have a crucial impact on the fabric performance. According to expectations, the efficiency of switching deteriorates along with that increase of traffic load offered to the fabric, however the decrease depends on the switching capabilities and features of MINs.

5 CONCLUSIONS

This paper presents the results of performance evaluation of the single panel multistage switching fabrics. Two classes of packet switched MINs, those with redundant interconnection paths (the PM2I type MINs) and those without (the cube type MINs), have been examined by simulated experiments in diverse traffic conditions. Simulations under different scenarios (including the SE model and operation mode of switching fabric) have shown that the performance of multistage switching fabrics deteriorates after an increase of load. However, the switch based on the PM2I type MIN outperforms that based on the cube type MIN under any traffic scenario. Moreover, presented simulation results should be also helpful in predicting the performance vs. SE capability trade-off before actual fabrication of the switching fabric. The simulation also revealed that fault tolerance of the PM2I type multistage switching fabrics can be very useful for scalable routers, i.e. such single panel multistage switching fabrics seem to be an attractive alternative to the multiple panel architecture of switching networks.

ACKNOWLEDGEMENTS

This work has been funded by the AGH University of Science and Technology, Kraków, within the statutory research grant.

REFERENCES

Chao, H. J., Liu, B., 2007. *High Performance Switches and Routers*. J. Wiley.

Hulicki, Z., 2013. ‘Performance aspects of some switching networks’. *Proc. SIMULTECH 2013*, Jul. 2013, Reykjavik, Iceland, pp. 90-93.

Hulicki, Z., 2013a. ‘Performance analysis of multistage switching networks’. *Proc. IEEE ATNAC 2013*, Nov. 2013, Christchurch, New Zealand, pp. 52-56.

Lien, Ch.-M., Chang, Ch.-S., Cheng, J., Lee D.-S., Liao, J.-T., 2010. ‘Using banyan networks for load-balanced switches with incremental update’, *GLOBECOM Workshops*, 2010 IEEE, pp. 1–6.

Rongsen, H., Delgado-Frias, J.G., 2007. ‘Fault tolerant interleaved switching fabrics for scalable high-performance routers’, *IEEE Trans. Paral. Distrib. Syst.*, vol. 18, no. 12, Dec. 2007, pp. 1727-1738.

Skalis, N., Mhamd, L., 2010. ‘Performance guarantees in partially buffered crossbar switches’, *GLOBECOM Workshops*, 2010 IEEE, pp. 1–6.

Tzeng, N.-F., 2004. ‘Multistage-based switching fabrics for scalable routers’, *IEEE Trans. Parallel and Distributed Systems*, vol. 15, no. 4, pp. 304-318.