

Area Efficient Asynchronous SDM Routers Using 2-Stage Clos Switches

Wei Song, Doug Edwards, Jim Garside, and William J. Bainbridge

School of Computer Science, the University of Manchester

Manchester M13 9PL United Kingdom

Email: {songw, doug, jdg}@cs.man.ac.uk; jrbainbridge@ieee.org

Abstract—Asynchronous on-chip networks are good candidates for multi-core applications requiring low-power consumption. Asynchronous spatial division multiplexing (SDM) routers provide better throughput with lower area overhead than asynchronous virtual channel routers; however, the area overhead of SDM routers is still significant due to their high-radix central switches.

A new 2-stage Clos switch is proposed to reduce the area overhead of asynchronous SDM routers. It is shown that replacing the crossbars with the 2-stage Clos switches can significantly reduce the area overhead of SDM routers when more than two virtual circuits are used. The saturation throughput is slightly reduced but the area to throughput efficiency is improved. Using Clos switches increases the energy consumption of switches but the energy of buffers is reduced.

I. INTRODUCTION

Asynchronous on-chip networks provide promising communication structures for future on-chip multi-core systems [1]. They consume less dynamic power than synchronous on-chip networks under low network load. The handshake protocols used in asynchronous networks are naturally tolerant to delay variations. The unified sync/async interfaces between cores and the asynchronous network ease the chip-level timing closure. They also allow the clock frequency and the supply voltage of local cores to be tuned independently [2]. It is expected that half of the global signalling will be driven asynchronously by year 2024 [3].

‘Virtual channel’ (VC) flow control [4] has been extensively utilised in asynchronous on-chip networks [5]–[8], where they are usually used to support different qualities of service (QoS) rather than improve throughput. Recently, it has been found that spatial division multiplexing (SDM) can significantly improve the best-effort throughput of asynchronous networks with less area overhead than VC [9]. It is shown that, in synchronous on-chip networks, SDM can also be used to support guaranteed throughput [10].

In this paper, the area overhead of asynchronous SDM routers will be significantly reduced by replacing the internal crossbar with a 2-stage Clos switch. The major area overhead of SDM routers is the enlarged crossbar with increased radix. It is well known that the multi-stage Clos switches

are theoretically the most area efficient structure for high-radix switches. In all configurations of Clos switches, it is found that the 2-stage Clos switches, which are used in optical networks [11], can be adopted in asynchronous on-chip networks, introducing much smaller area overhead than crossbars. Similar to a crossbar, a 2-stage Clos switch must be reconfigured dynamically. A Clos scheduler, simplified from the one used in an asynchronous 3-stage Clos switch [12], is implemented in each router.

The remainder of the paper is organised as follows: Section II reviews SDM and its area overhead in asynchronous routers. Section III provides a short introduction to Clos switches and reveals the structure of the 2-stage Clos switch. Later an SDM router is implemented in Section IV using the 2-stage Clos switch. The performance of the new SDM router is analysed in Section V. Finally, the paper is concluded in Section VI.

II. SPATIAL DIVISION MULTIPLEXING

Spatial division multiplexing (SDM) was originally used in wireless communication systems to transmit multiple data streams on different antennas simultaneously for the best bandwidth efficiency. Similar concepts are utilised in synchronous on-chip networks [10,13] where multiple data streams can use the same network link by occupying a subset of the link wires, namely a *virtual circuit* [10].

The first asynchronous SDM router [9] was compared with asynchronous VC routers using various router configurations and traffic patterns. It was shown that both SDM and VC improve network throughput but SDM outperforms VC. SDM also consumes much less area than VC for the same saturation throughput.

The internal structure of an SDM router is shown in Fig. 1. It has P input ports (IPs) and P output ports (OPs). Each IP/OP is connected to an input/output buffer. These buffers are dynamically connected by the crossbar. Each input/output buffer is physically divided into several input/output virtual circuits. To ease the connection between different virtual circuits, all of them have the same data width. Assuming the data width of a port is W bits and every port has V virtual circuits, the data width of a virtual circuit is fixed to W/V bits and the crossbar connects VP input virtual circuits to VP output virtual circuits. A packet is transmitted on only one virtual circuit in a serialised way. Since a pausing packet

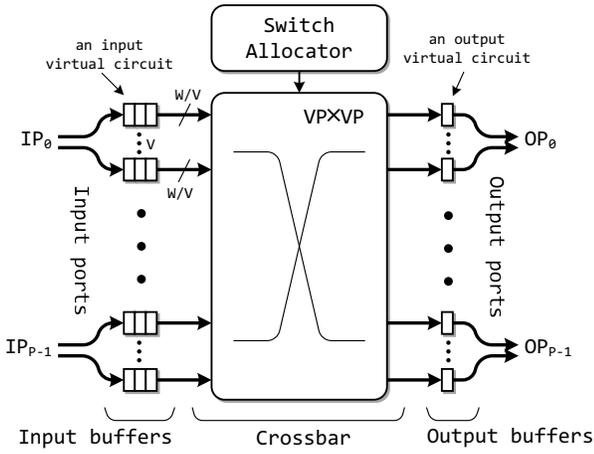


Fig. 1: An SDM router

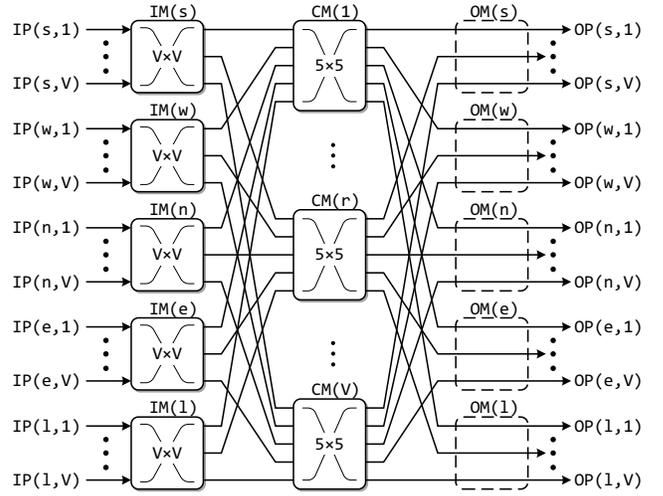


Fig. 2: A 2-stage Clos switch

blocks only one virtual circuit, a portion of the total bandwidth rather than all of it, the head-of-line (HOL) blockage is alleviated.

A quantitative analysis will show the area overhead of SDM routers. Assuming the buffer depth in a baseline router (a packet switched wormhole router using no VCs or virtual circuits) is D , the area of the buffer can be estimated as:

$$A_{\text{buf, baseline}} \approx P(WDA_b + A_{\text{ctl}}) \quad (1)$$

where A_b and A_{ctl} represent the area of a 1-bit buffer and the extra area of the control circuit respectively. In the same way, an estimation of the buffer area in an SDM router can be produced as follows:

$$A_{\text{buf, SDM}} \approx PV[(W/V)DA_b + A_{\text{ctl}}] \quad (2)$$

$$\approx P(WDA_b + VA_{\text{ctl}}) \quad (3)$$

It can be deduced from Equation 1 and 3 that some area overhead is introduced in the buffers of SDM routers due to the extra control circuits of independent virtual circuits. Since buffer control circuits normally consume much less area compared with buffers, this area overhead is moderate.

Crossbars lead to the major area overhead in SDM routers. The area of any switch is generally proportional to the number of cross points inside it. Assuming crossbars are fully connected, the area of the central crossbar in a baseline router and an SDM router is described as follows:

$$A_{\text{crossbar, baseline}} \approx (P \times P)WA_{\text{cp}} \quad (4)$$

$$A_{\text{crossbar, SDM}} \approx (VP \times VP)(W/V)A_{\text{cp}} \quad (5)$$

$$\approx V(P \times P)WA_{\text{cp}} \quad (6)$$

where A_{cp} is the equivalent area of a single cross point. By comparing Equation 4 and 6, it is revealed that the area of the crossbar in an SDM router is V times of the area in a baseline router. The rest of this paper concentrates on reducing this area overhead.

III. 2-STAGE CLOS SWITCHES

A general Clos switch comprises three stages of crossbars formed symmetrically [14]. It is reduced to a 2-stage Clos switch when the last stage is statically configured and, therefore, removed. Fig. 2 demonstrates a 2-stage Clos switch which can be used in normal SDM routers with five directions (south, west, north, east and local).

The three stages are called input modules (IMs), central modules (CMs) and output modules (OMs) respectively. The total N input ports (IPs) or output ports (OPs) are separated into k groups². Each group includes $n = N/k$ IPs/OPs connected to one IM/OM. There are m central modules implemented in a Clos switch. CMs are connected to all IMs/OMs in a distributed pattern that every IM/OM is connected to every CM by only one link. When the number of CMs is larger than or equal to n ($m \geq n$), the Clos switch is theoretically non-blocking. For the minimum area overhead, most Clos switches are configured with the minimum number of CMs, in which case, all IMs/OMs are $n \times n$ crossbars and all CMs are $k \times k$ crossbars.

In the 2-stage Clos switch shown in Fig. 2, the number of IP groups is set to the port count of the router ($k = P = 5$). As $n = PV$ in an SDM router, the number of ports in an IM is V , the number of virtual circuits implemented. This configuration has two major benefits:

- All IPs/OPs in one IM/OM have the same input/output direction. Since OPs heading to the same output direction are equivalent, the configuration of any OM can be determined statically without compromising the functionality. In this way, the outputs of CMs are directly connected to OPs, effectively removing all OMs.
- As every CM is a $k \times k$ crossbar and $k = 5$, its function is equivalent to the crossbar in baseline routers. Thus, any area reduction techniques used in baseline routers, such

²When a 2-stage Clos is used in an SDM router, a port of the Clos switch is connected to one virtual circuit and a port of the SDM router is connected to V virtual circuits.

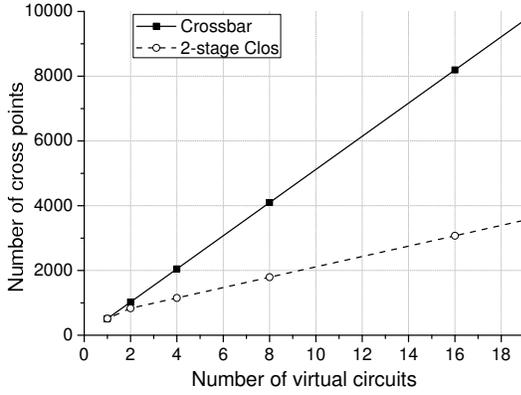


Fig. 3: Switching area of 5-port 32-bit routers

as removing the unused cross points [15], can be used in all CMs for further area reduction.

This configuration can be extended to support any router configurations as long as the three parameters that define a general Clos switch, (n, k, m) , are set as (V, P, V) . The area estimation is calculated as follows:

$$A_{\text{Clos,SDM}} \approx (PV^2 + VP^2)(W/V)A_{\text{cp}} \quad (7)$$

$$\approx (PV + P \times P)WA_{\text{cp}} \quad (8)$$

As shown in Equation 8, 4 and 6, using the 2-stage Clos switch reduces the extra switching area per virtual circuit from P^2WA_{cp} to PWA_{cp} . Fig. 3 further demonstrates the switching area of using crossbars or 2-stage Clos switches inside 5-port 32-bit routers. Assuming the XY routing algorithm is used, unused cross points are removed in switches and deducted from the area estimation. It is easy to see the area benefit of using Clos switches with a large number of virtual circuits.

The dynamic reconfiguration of Clos switches is an important research issue. Every I/O pair in a Clos switch has m different paths through the m CMs. The maximum number of paths that a CM can support is k , the number of IMs, but the total number of IPs/OPs of a Clos switch is $k \times n$. Therefore, it is important for the reconfiguration algorithm to choose an appropriate CM for each requesting IP without congesting CMs. The reconfiguration algorithms are normally classified into two categories [16]: *optimal algorithms*, which provide guaranteed results for all matches but with high complexity in time or implementation, and *heuristic algorithms*, which provide all or partial matches in low time and implementation complexity. The time and implementation overhead of optimal algorithms is normally intolerable for hardware reconfigured Clos switches. In this paper, the asynchronous heuristic algorithm proposed in [12] is used to dynamically reconfigure the 2-stage Clos switch.

Replacing the crossbar with the 2-stage Clos switch in an SDM router slightly compromises throughput. The reason lies in the Clos switch itself. The 2-stage Clos switch is theoretically non-blocking: it is non-blocking only if the established paths can be relocated when such operation is needed to

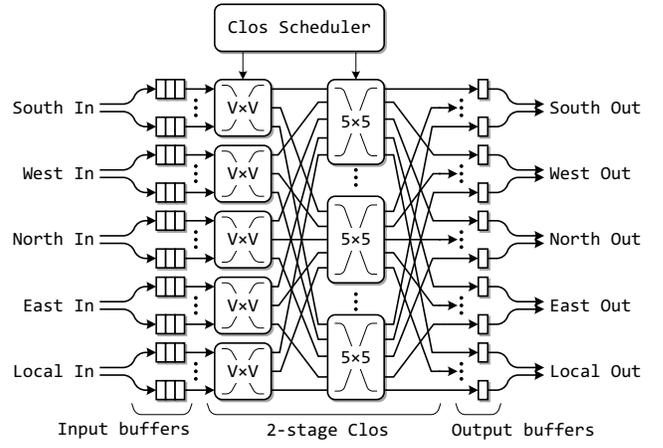


Fig. 4: An SDM router using the 2-stage Clos switch

connect a new I/O pair [17]. However, path relocation is normally prohibited in practical Clos switches and thus the 2-stage Clos switch is actually blocking.

Although the 2-stage Clos switch compromises throughput, the throughput improvement from baseline routers is still significant. This will be verified in the simulations presented in Section V.

IV. ROUTER IMPLEMENTATION

In this section, an asynchronous 5-port SDM router using the 2-stage Clos switch (SDM-Clos) is presented. As shown in Fig. 4, this router is compatible with normal mesh and torus networks. It is connected to a local processing element and four neighbouring routers. Every input port is connected to an input buffer which is divided into several virtual circuits. A 1-stage output buffer is connected to each output port to decouple the timing paths of inter-router links and intra-router switches. The data width of each port, the number of virtual circuits implemented in each port and the depth of buffers are configurable at design time.

Packets are delivered in a sequence of flits: a header flit, several data flits and a tail flit. The XY routing algorithm is used in this implementation but any routing algorithm can be supported. The address of the target processing element of each packet is stored in the header flit. Since the data width of one virtual circuit is W/V , multiple header flits may be required when the target address needs more than W/V bits. This would prolong the packet transmission latency and the minimal buffer depth. In this paper, the minimum data width of a virtual circuit is limited to 8 bits; therefore the maximum network size is 16×16 (using the XY routing algorithm).

All buffers are implemented using 4-phase 1-of-4 multi-rail pipelines [18], which are tolerant to delay variations. All 4-phase asynchronous pipelines [19], such as dual-rail pipelines, m-of-n pipelines [20] and self-timed bundled-data pipelines, can be used for data buffers without significant changes in control circuits. The 1-of-4 pipeline is chosen for its tolerance to variation, moderate area overhead and small energy consumption.

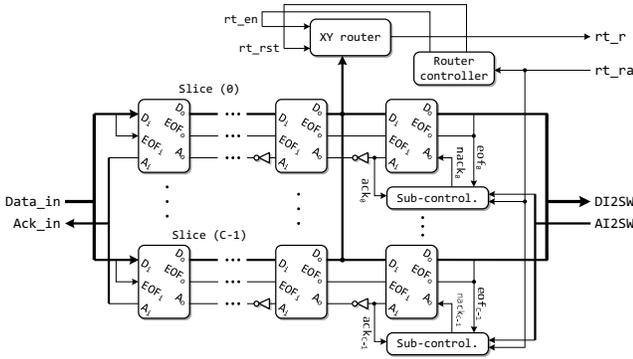


Fig. 5: Internal structures of an input virtual circuit

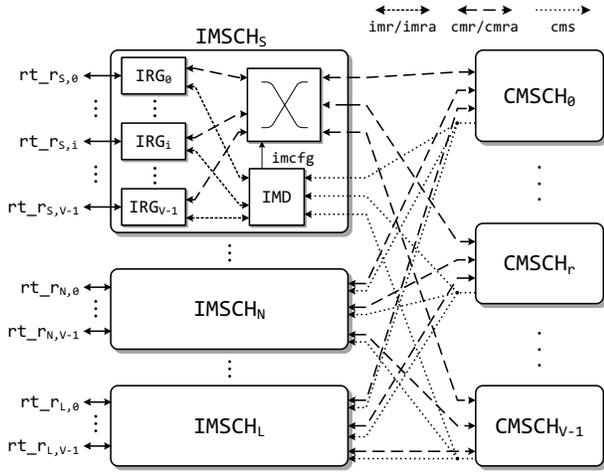


Fig. 6: Clos scheduler

As described in Section II, every virtual circuit transmits a packet independently. Thus the input buffer of every virtual circuit is equivalently controlled as an input port of the baseline router. Fig. 5 demonstrates the internal structure of an input virtual circuit. Buffer stages are further separated into fully decoupled slices to improve performance [21]. The extra control circuit in each slice ensures that the XY router is fed with the header flit for address decoding. These control circuits are reproduced from the first asynchronous SDM router [9] which uses a crossbar as the central switch.

Fig. 6 demonstrates the structure of the Clos scheduler, which is reduced from the scheduler [12] proposed for asynchronous 3-stage Clos switches. It receives the routing requests (rt_r) from all input virtual circuits, allocates a path for each active request and produces the ack (rt_{ra}) to the input virtual circuit once the path is configured. Running a heuristic dispatching algorithm [12], the scheduler adopts a distributed structure where each switching module has its own scheduler: IM schedulers (IMSCHs) and CM schedulers (CMSCHs).

An IM scheduler comprises of a group of input request generators (IRG), an IM dispatcher (IMD) and a crossbar delivering the requests to CMs. The request from each input virtual circuit is connected with one input request generator. It transfers the incoming request to two individual requests:

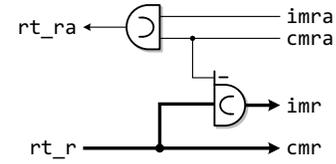


Fig. 7: Input request generator

imr for the IM dispatcher and cmr for the CM scheduler. It also controls the timing of the path allocation process. The implementation of an input request generator is shown in Fig. 7. Both imr and cmr are produced from rt_r at around the same time but imr is released after cmr is de-asserted. This ensures that the path in CMs are released before the IM path is released, otherwise the path in CMs may cause a false allocation. The final ack signal, rt_a , is set when the whole path is configured and is reset when the whole path is released.

All imr signals in one IM are sent to the IM dispatcher, which allocates available IM outputs to active requests. The allocation has taken the path availability inside CMs (denoted by the CM state feedback signal cms) into consideration to avoid most of the contention in CMs. Inside the IM dispatcher, there are two levels of arbiters: output arbiters and input arbiters. An active request is first sent to all of the output arbiters heading to the available CMs. Depending on their own status, these arbiters may or may not grant this request but the grant results are sent back to the input arbiter. Multiple outputs may grant the same input request; therefore, the input arbiter randomly chooses one output and releases the requests to other output arbiters. If all CMs are currently unavailable for a certain request, no output arbiter would grant it and it must wait until a CM returns available.

The allocation results from the IM dispatcher reconfigure the IM and the crossbar inside the IM scheduler which forwards the corresponding cmr to CM schedulers. Since a CM is equivalent to the crossbar in a baseline router, a CM scheduler has the same structure of the switch allocator of a baseline router. The state feedback signals, cms , are produced from the allocation results in order to identify the available output ports in each CM.

V. PERFORMANCE EVALUATION

Three routers were implemented for performance comparison: baseline routers (packet switched wormhole routers using no VCs or virtual circuits), SDM routers using crossbars and SDM-Clos routers using the 2-stage Clos switches. Described in gate-level SystemVerilog, all routers are laid out using the Synopsys design flow and a $0.13 \mu\text{m}$ standard cell library. Since most control circuits use speed independent timing assumptions [19] and data paths are implemented using multi-rail data encoding methods³, no timing constraint is required

³The arbiters in IM dispatchers are optimised using a delay assumption for fast arbitration [12]. In addition, lookahead pipelines [22] are used in the sliced buffers of input buffers, which also use some delay assumptions. Detailed analyses of these delay assumptions [12,21,22] show that they are already satisfied by practical gate delays and are robust enough to remain valid without special treatment in the automatic placement & routing process.

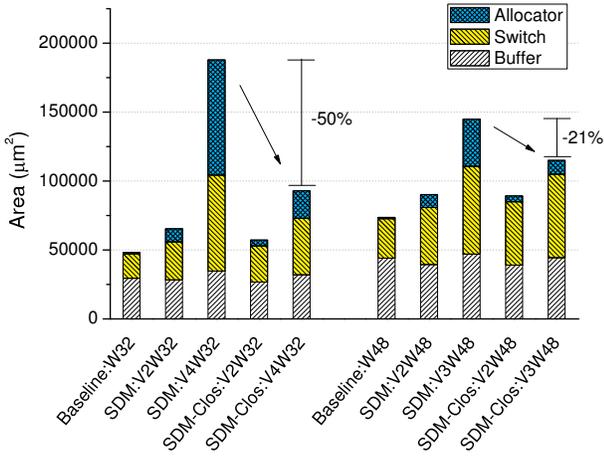


Fig. 8: Router area

to ensure correct circuit function in the back-end process. However, timing constraints are used, purely for the purpose of speed optimisation. All implementations are iterated multiple times for the best speed performance without any timing or design rule violation.

Fig. 8 reveals the area breakdown of routers with different numbers of virtual circuits (denoted by V) and port data widths (denoted by W). For the minimum area overhead, the input buffer in all routers has only one stage of buffers. It is shown that the differences in buffer area among routers with the same data width are marginal compared to the differences due to switches and allocators. The switch area in SDM routers is approximately proportional to the number of virtual circuits as described in Equation 6. The results also show that significant area overhead is introduced by the allocators in SDM routers. Although various allocators lead to different area overhead, it is generally true that the area increases faster than proportional to the number of requests and resources. In the case of SDM:V4W32, which has the largest number of virtual circuits, the allocator is larger than the crossbar.

Using the 2-stage Clos switches significantly reduces the area of both switches and allocators when more than two virtual circuits are implemented. Compared with SDM:V4W32, SDM-Clos:V4W32 saves around 41% area in switches and 76% area in allocators (50% of the total area). When two virtual circuits are used, the area reduction of Clos switches is marginal.

Several 8×8 mesh networks have been built. Random uniform traffic is injected into the network by the processing element (written in SystemC) connected to each router. Packets, containing 64-byte payloads, are generated by processing elements in a Poisson process. Fig. 9 shows the average packet transmission latency with various load injection rates. It is shown that virtual circuits improve the saturation throughput significantly at a cost of long transmission latency. Since a virtual circuit takes only a portion of the total data width, packets are serialised during transmission, which causes the extra latency. It is also shown that the throughput reduction

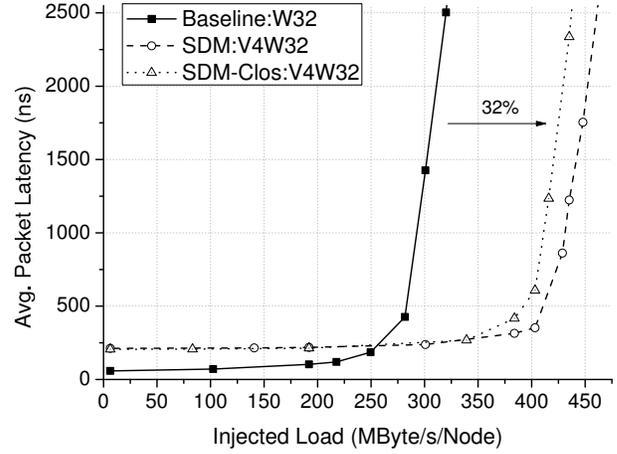


Fig. 9: Packet transmission latency

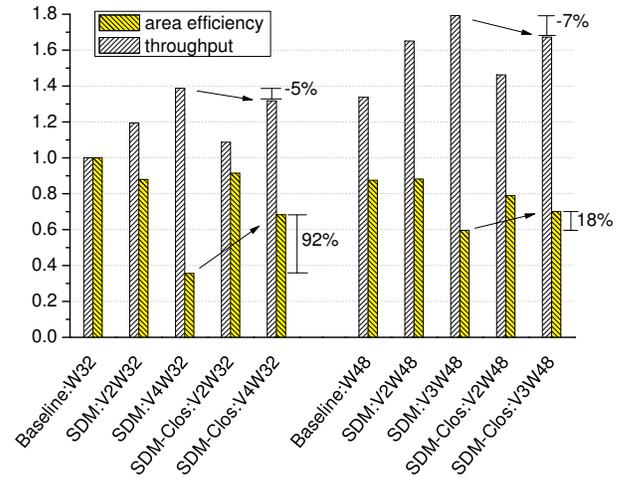


Fig. 10: Network saturation throughput and area efficiency

led by the 2-stage Clos switch is not significant. The period of Baseline:W32, SDM:V4W32 and SDM-Clos:V4W32 routers at typical corner are around 2.2 ns, 2.8 ns and 2.8 ns respectively. The extra 0.6 ns in period is introduced by the central switches, which are large and complicated in SDM and SDM-Clos routers.

Fig. 10 reveals the normalised saturation throughput and area efficiency of routers with data widths of 32 and 48 bits. The performance of Baseline:W32 is used as the baseline case for the normalisation. The area efficiency is defined as *saturation throughput divided by router area*. Greater area efficiency means less area overhead for the same saturation throughput. As shown in the figure, all SDM routers improve the saturation throughput at a cost of low area efficiency. Using the 2-stage Clos switch compromises the throughput but the area efficiency is boosted significantly in all the cases using more than two virtual circuits. If only two virtual circuits are used, the SDM routers using crossbars are the best choices.

The switching activities of the router at position (3,3) are recorded for power analysis. Fig. 11 shows the energy breakdown of the routers when networks are saturated. The

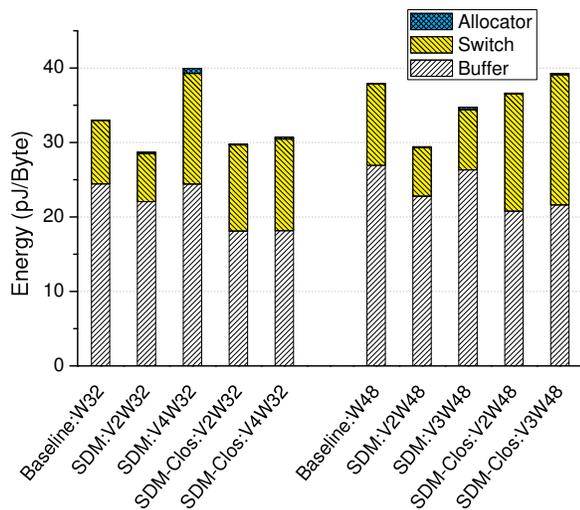


Fig. 11: Energy consumption

energy figures have been divided by the saturation throughput to directly reveal the energy consumption for every byte transmitted in the network. Most SDM routers show better energy efficiencies than baseline routers except for SDM:V4W32 and SDM-Clos:V3W48, whose switches consume too much energy. The energy breakdown shows that using the 2-stage Clos switch increases the energy consumption of the switches but the buffers consume less. This can be a benefit for the SDM-Clos routers using deep buffers as they may consume less energy than the SDM routers using crossbars.

VI. CONCLUSION

A new asynchronous spatial division multiplexing router using a 2-stage Clos switch as the central switch is proposed for asynchronous on-chip networks. Using the 2-stage Clos switch (SDM-Clos) rather than crossbars (SDM) significantly reduces the area overhead of asynchronous SDM routers when more than two virtual circuits are used. As shown in the comparison with the 32-bit SDM router using four virtual circuits, the SDM-Clos router saves 41% area in switches and 76% area in allocators. The saturation throughput is slightly reduced due to the internal blocking of the 2-stage Clos switch but the area efficiency is improved. Using the 2-stage Clos switch increases the energy consumption of the switches but the energy overhead of buffers is reduced. SDM-Clos routers may consume less energy than SDM routers when deep buffers are required.

All router implementations, including the router designs described in Verilog HDL, SystemC test benches, and the script files for the back-end design flow are available from the Asynchronous SDM NoC OpenCore project [23].

REFERENCES

[1] M. Krstić, E. Grass, F. K. Gürkaynak, and P. Vivet, "Globally asynchronous, locally synchronous circuits: overview and outlook," *IEEE Design and Test of Computers*, vol. 24, no. 5, pp. 430–441, 2007.

[2] E. Beigné, F. Clermidy, H. Lhermet, S. Miermont, Y. Thonnart, X.-T. Tran, A. Valentian, D. Varreau, P. Vivet, X. Popon, and H. Lebreton, "An asynchronous power aware and adaptive NoC based circuit," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 4, pp. 1167–1177, 2009.

[3] *International Technology Roadmap for Semiconductors*, 2009, ch. Design, pp. 12–13.

[4] W. J. Dally, "Virtual-channel flow control," *IEEE Transactions on Parallel and Distributed Systems*, vol. 3, no. 2, pp. 194–205, 1992.

[5] T. Felicijan and S. B. Furber, "An asynchronous on-chip network router with quality-of-service (QoS) support," in *Proc. of IEEE International SOC Conference*, 2004, pp. 274–277.

[6] T. Bjerregaard and J. Sparsø, "A router architecture for connection-oriented service guarantees in the MANGO clockless network-on-chip," in *Proc. of Design, Automation & Test in Europe Conference & Exhibition*, 2005, pp. 1226–1231.

[7] E. Beigné, F. Clermidy, P. Vivet, A. Clouard, and M. Renaudin, "An asynchronous NOC architecture providing low latency service and its multi-level design framework," in *Proc. of IEEE International Symposium on Asynchronous Circuits and Systems*, 2005, pp. 54–63.

[8] R. R. Dobkin, R. Ginosar, and A. Kolodny, "QNoC asynchronous router," *Integration, the VLSI Journal*, vol. 42, no. 2, pp. 103–115, 2009.

[9] W. Song and D. Edwards, "Asynchronous spatial division multiplexing router," *Microprocessors and Microsystems*, vol. 35, no. 2, pp. 85–97, 2011.

[10] A. Leroy, D. Milojevic, D. Verkest, F. Robert, and F. Catthoor, "Concepts and implementation of spatial division multiplexing for guaranteed throughput in networks-on-chip," *IEEE Transactions on Computers*, vol. 57, no. 9, pp. 1182–1195, 2008.

[11] J. Cheyens, C. Devellder, E. V. Breusegem, D. Colle, F. D. Turck, P. Lagasse, M. Pickavet, and P. Demeester, "Clos lives on in optical packet switching," *IEEE Communications Magazine*, vol. 42, no. 2, pp. 114–121, 2004.

[12] W. Song, D. Edwards, Z. Liu, and S. Dasgupta, "Routing of asynchronous Clos networks," *IET Computers & Digital Techniques*, vol. 5, no. 6, pp. 452–467, 2011.

[13] P. T. Wolkotte, G. J. Smit, G. K. Rauwerda, and L. T. Smit, "An energy-efficient reconfigurable circuit-switched network-on-chip," in *Proc. of IEEE International Parallel and Distributed Processing Symposium*, 2005.

[14] C. Clos, "A study of non-blocking switching networks," *Bell System Technical Journal*, vol. 32, no. 5, pp. 406–424, 1953.

[15] A. Adriahtantaina, H. Charlery, A. Greiner, L. Mortiez, and C. A. Zeferino, "SPIN: A scalable, packet switched, on-chip micro-network," in *Proc. of Design, Automation & Test in Europe Conference & Exhibition*, 2003, pp. 70–73.

[16] H. J. Chao, Z. Jing, and S. Y. Liew, "Matching algorithms for three-stage bufferless Clos network switches," *IEEE Communications Magazine*, vol. 41, no. 10, pp. 46–54, 2003.

[17] H. J. Chao, C. H. Lam, and E. Oki, *Broadband Packet Switching Technologies: A Practical Guide to ATM Switches and IP Routers*. John Wiley & Sons, Inc., 2001.

[18] J. Bainbridge and S. Furber, "Chain: a delay-insensitive chip area interconnect," *IEEE Micro*, vol. 22, pp. 16–23, 2002.

[19] J. Sparsø and S. Furber, *Principles of Asynchronous Circuit Design — A Systems Perspective*. Kluwer Academic Publishers, 2001.

[20] J. Bainbridge, W. Toms, D. Edwards, and S. Furber, "Delay-insensitive, point-to-point interconnect using m-of-n codes," in *Proc. of IEEE International Symposium on Asynchronous Circuits and Systems*, 2003, pp. 132–140.

[21] W. Song and D. Edwards, "A low latency wormhole router for asynchronous on-chip networks," in *Proc. of Asia and South Pacific Design Automation Conference*, 2010, pp. 437–443.

[22] M. Singh and S. M. Nowick, "The design of high-performance dynamic asynchronous pipelines: lookahead style," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 15, no. 11, pp. 1256–1269, 2007.

[23] W. Song and D. Edwards. (2011) Asynchronous SDM NoC. [Online]. Available: http://opencores.org/project,async_sdm_noc,Overview