

LETTER

A 1-Cycle 1.25 GHz Bufferless Router for 3D Network-on-Chip

Chaochao FENG^{†a)}, Student Member, Zhonghai LU^{††}, Axel JANTSCH^{††}, and Minxuan ZHANG[†], Nonmembers

SUMMARY In this paper, we propose a 1-cycle high-performance 3D bufferless router with a 3-stage permutation network. The proposed router utilizes the 3-stage permutation network instead of the serialized switch allocator and 7×7 crossbar to achieve the frequency of 1.25 GHz in TSMC 65 nm technology. Compared with the other two 3D bufferless routers, the proposed router occupies less area and consumes less power consumption. Simulation results under both synthetic and application workloads illustrate that the proposed router achieves less average packet latency than the other two 3D bufferless routers.

key words: 3D NoC, bufferless router, permutation network

1. Introduction

Recently, three-dimensional Network-on-Chip (3D NoC) is emerging as a promising network architecture that takes advantage of the short vertical interconnects of 3D-ICs [1]. As the number of elements integrated on a single chip increases significantly, power, area and design complexity become key factors for the on-chip network. Bufferless router, which provides a low-cost solution for NoC, has become a hot research issue [2], [3]. However, the serialized switch allocator which lies on the critical path degrades the performance of the traditional bufferless router [4]. A low-complexity bufferless router (called CHIPPER) has been proposed in [4] for 2D mesh NoC. A partial permutation network is used to replace the switch allocator and crossbar in the router. Although the router achieves smaller area and higher performance than the router in [3], it still has two pipeline stages and adopts an extra rule to avoid livelock. To the best of our knowledge, few previous works focus on the bufferless router for 3D NoC.

In this paper, we propose a 1-cycle high-performance 3D bufferless router with a 3-stage permutation network (called 3D_PERM). The serialized switch allocator and 7×7 crossbar are replaced with a 3-stage permutation network, which can reduce the number of logic levels of the router significantly. The 3D_PERM router is implemented in TSMC 65 nm technology, and is shown to achieve the frequency of 1.25 GHz. Compared with the baseline 3D bufferless router (called 3D_BASE) and the 3D CHIPPER bufferless router (called 3D_CHIPPER, extended from [4]), the 3D_PERM router achieves higher performance, occupies

less area and consumes less power consumption. Simulation results under both synthetic and application workloads illustrate that the 3D_PERM router achieves less average packet latency than the other two 3D bufferless routers.

The rest of paper is organized as follows. A baseline 3D bufferless router is introduced in Sect. 2. Section 3 proposes the 1-cycle 1.25 GHz 3D bufferless router. In Sect. 4, simulation experimental results are presented and analyzed, followed by the conclusion in Sect. 5.

2. Baseline 3D Bufferless Router

We extend the Nostrum bufferless router [5], which is a 5-port router for 2D mesh NoC, to a 7-port router for 3D mesh NoC as the baseline bufferless router. Besides the input registers for each input port, there are no extra buffers in the router. Deflection routing is suitable for routing packets in bufferless NoC to reduce the hardware overhead of the router. The structure of the 3D_BASE router is shown in Fig. 1. Each flit in a packet contains a header and is routed independently. Flits can be arrived at the destination out of order. So a reassembly buffer is needed in the network interface to reassemble the packet. In order to limit the number of misroutings for avoiding livelock, arriving flits are prioritized based on their "age", which is the number of hops the flit has gone through in the network, by the *Input_Priority_Sort* module. In the case of contention for one output port, the flit with the highest priority will be routed through this port, while other flits will be deflected to the non-productive ports.

A proximity congestion awareness technique [6] is used for output port arbitration to achieve load balance. Each router calculates the load information, which is the number of flits processed in the last 4 cycles, through the *Load_Computation* module and sends the load information to its 6 neighbors through the dedicated signal *Load_out*. After receiving the load information from neighbors with the signal *Load_in*, the router prioritizes the output ports according to the load information through the *Output_Priority_Sort* module. The port with the smallest loads has the highest priority. If there is more than one productive port to route the flit, the port with the smallest loads will be chosen. In addition, the router will always choose the port with the smallest loads to deflect the low-priority flit. The *Switch_Allocator* module generates the selecting signal for the 7×7 crossbar based on the results of the routing computation, the input and output priority sort. Relative addressing

Manuscript received November 4, 2011.

[†]The authors are with School of Computer, National University of Defense Technology, China.

^{††}The authors are with the Department of Electronic Systems, Royal Institute of Technology, Sweden.

a) E-mail: fengchaochao@nudt.edu.cn

DOI: 10.1587/transinf.E95.D.1519

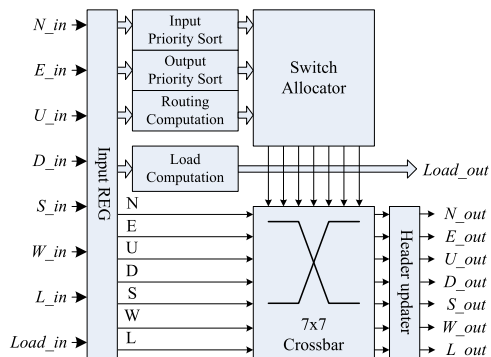


Fig. 1 Baseline 3D bufferless router.

is used for the source and destination address fields of the flit. The *Header_Updater* module following after the crossbar is used for updating the address fields and the hop count field of the flit. Because the router makes switch allocation for each arriving flit one by one from the highest priority to the lowest, the serialized switch allocation becomes a long critical path in the baseline 3D bufferless router, which lowers the frequency the router can achieve.

3. 1-Cycle 1.25 GHz 3D Bufferless Router

3.1 3D Bufferless Router with a Permutation Network

The serialized switch allocation slows down the performance of the baseline 3D bufferless router. Although pipelining the router can enhance the clock frequency of the baseline 3D bufferless router, the pipeline also increases the single hop latency. We propose a 1-cycle high-performance 3D bufferless router with a 3-stage permutation network instead of the switch allocator and 7×7 crossbar.

The structure of the 3D bufferless router with a permutation network is shown in Fig. 2. The flit ejector first checks whether a flit has reached the destination or not and generates the selecting signal of the 6-to-1 multiplexer. If one or more flits have reached the destination, the highest-priority flit can be routed to the local port through the 6-to-1 multiplexer. The flit ejector also decides whether the local node can inject a flit into the network or not. If the number of incoming flits is less than 6, the local input flit can be injected through one free input port of the permutation network. The router adopts a 3-stage permutation network instead of the switch allocator and crossbar. Each stage contains three 2×2 permutation cells. Permutation network is often used in indirect network such as butterfly topology [7]. The difference from the indirect network is the function of the 2×2 permutation cell. In the case of two flits contending for the same output of the permutation cell, the permutation cell in our design misroutes one flit to the other output rather than blocks. The input and output priority sort modules are completely removed from the router to reduce the number of logic levels further.

Although the CHIPPER router [4] can also be extended for 3D NoC, the differences between the CHIPPER router

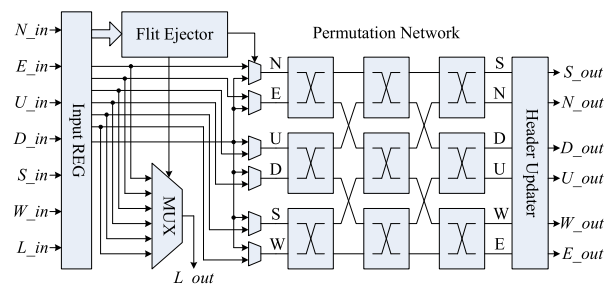


Fig. 2 3D bufferless router with a permutation network.

and the 3D_PERM router are the function of the 2×2 permutation cell which is the key element of the two routers and the method how to avoid livelock. In [4], an extra rule (called Golden Packet) has been used to avoid livelock. The Golden Packet can always win the arbitration in the permutation cell, while other packets are arbitrated pseudo-randomly. Each packet in the network has a chance to be treated as Golden Packet which will never be misrouted. The packet can be denoted as (sender, transaction ID) tuples uniquely. Each router consists of two counters to rotate all possible packet IDs as Golden Packet. The disadvantage of the rule has two aspects: (1) Rotate all possible packet IDs will increase the hardware overhead and take a long time; (2) Most packets are not golden, and arbitrated pseudo-randomly to pass through the permutation cell, which will increase the packet latency. The permutation cell in the 3D_PERM router performs the permutation based on the hop counts of the flit simply, which overcomes the disadvantage of the Golden Packet rule. The basic idea is the flit with larger hop counts wins the arbitration. The 3D_PERM router uses this simple rule for arbitration which can reduce hardware overhead and achieve livelock-free at the same time.

To illustrate the function of the 2×2 permutation cell in our design, we use the permutation cell at stage 1 for the north and east input flits as an example. The function of the permutation cell is shown in Fig. 3. The permutation cell first gets the productive direction(s) of the two incoming flits according to the relative address to the destination (d_addr), and then permutes the flits according to the hop counts (HC) of the flits. The flit with larger hop counts (higher priority) can be routed to its desired port, while the other flit takes the other port. For example, if the hop counts of *flit0* are larger than those of *flit1* and the desired output port for *flit0* is one of four directions (*North*, *South*, *Up* and *Down*), *flit0* will be routed through the *Up* port of the permutation cell. After passing through the stage-1 and stage-2, *flit0* can be routed to the output port through the up or middle cell of the stage-3. The flit with the highest priority can always be routed to its desired port at each stage, thus can be routed to the productive output port of the router finally. The 3-stage permutation network guarantees the highest-priority flit advances towards its destination finally, which can achieve freedom from livelock. This is a key different feature from [4] which uses the Golden Packet rule to avoid livelock. The order of

the output port (S, N, D, U, W, E) for the permutation network corresponding to the order of the input port (N, E, U, D, S, W) can avoid the input flit reflecting back to the input direction.

3.2 Hardware Implementation

In order to make a comparison, we also extend the CHIPPER router [4] from 2D to 3D and optimize the 3D_CHIPPER router from 2-cycle to 1-cycle. Three 3D bufferless routers (3D_BASE, 3D_CHIPPER and 3D_PERM) are developed at register transfer level (RTL-level) with VHDL. We synthesize three routers using Synopsys Design Compiler with TSMC 65 nm technology at

```

Permutation_cell
Input: in0, in1 (Flit type)
Output: out0, out1 (Flit type)

1: {d_productive0} ← get_productive_dir(in0.d_addr)
2: {d_productive1} ← get_productive_dir(in1.d_addr)
3: if in0.HC >= in1.HC then
4:   if (d_productive0 = North or South or Up or Down) then
5:     out0 ← in0; out1 ← in1;
6:   else
7:     out0 ← in1; out1 ← in0;
8:   end if
9: else
10:  if (d_productive1 = North or South or Up or Down) then
11:    out0 ← in1; out1 ← in0;
12:  else
13:    out0 ← in0; out1 ← in1;
14:  end if
15: end if

```

Fig. 3 Function of 2×2 permutation cell.

Table 1 Hardware cost comparison for three 3D bufferless routers.

	Timing (ns)	Area (μm^2)	Power (mW/MHz)
3D_BASE	4	73760	0.031
3D_CHIPPER	1	35639	0.016
3D_PERM	1	33621	0.011
	0.8	45026	0.018

the typical condition (25°C and 1.0v). The data path of the router has 128 bits, which contains a 48-bit flit header and an 80-bit flit payload. The critical path timing, area and power consumption of three 3D bufferless routers are show in Table 1. Power consumption is measured in power per frequency (mW/MHz). The 3D_PERM router can achieve the frequency of 1.25 GHz, which is 4 and 0.25 times more than that of the 3D_BASE and 3D_CHIPPER routers respectively. To make a fair comparison with the 3D_CHIPPER router, we also list the synthesize results of the 3D_PERM router with 1 ns timing constraint. Compared with the 3D_BASE router, the 3D_PERM router can save 39% and 42% less area and power consumption respectively. Under 1 ns timing constraint, the 3D_PERM router can achieve 6% and 31% less area and power consumption than the 3D_CHIPPER router.

4. Performance Evaluation

We evaluate the performance of three 3D bufferless routers using a cycle-accurate NoC simulator developed in VHDL under both synthetic and application workloads. The simulations are performed on a $4 \times 4 \times 4$ 3D mesh. Each packet contains 4 flits. For synthetic workloads, we use three traffic patterns: uniform random, transpose and local. In uniform random traffic, each resource node sends packets randomly to other nodes with an equal probability. For transpose traffic, resource node positioned at (x, y, z) sends packets to destination node $(N_x - 1 - x, N_y - 1 - y, N_z - 1 - z)$ for all $x \in \{0, \dots, N_x - 1\}, y \in \{0, \dots, N_y - 1\}, z \in \{0, \dots, N_z - 1\}$, where

Table 2 Full-system configuration for trace generation.

Number of Processors	64
ISA	SPARC
L1 Cache	32 K-I/D, 4-way associative, 64 B/line
L2 Cache	fully shared S-NUCA, 512 KB/bank 64 banks, 64 B/line, 8-way associative
Cache coherence protocol	MOESI_CMP_directory
Memory	8 on-chip memory controllers
Splash applications	barnes, cholesky, fft, fmm lu, radix, raytrace, water

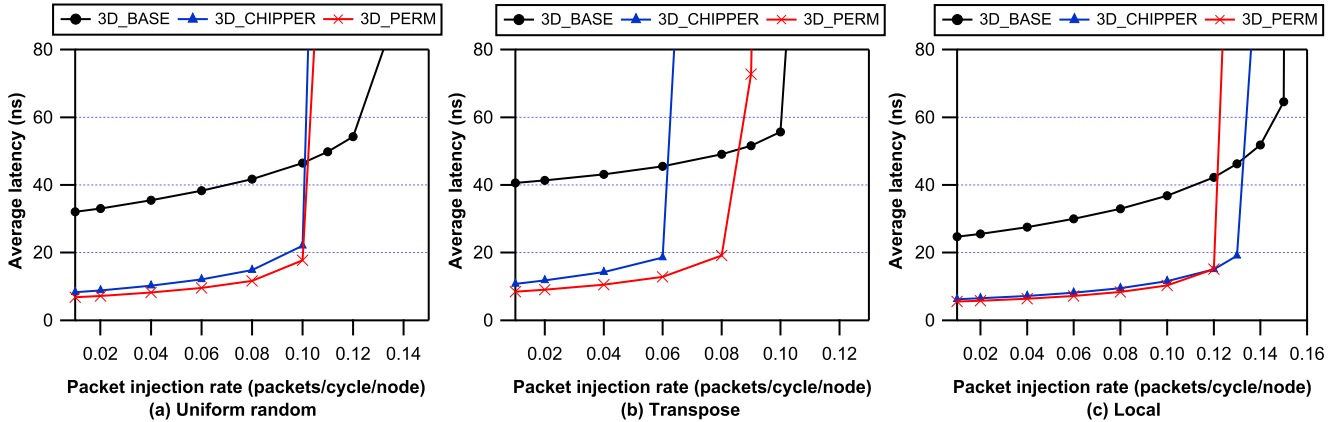


Fig. 4 Average packet latency with synthetic workloads.

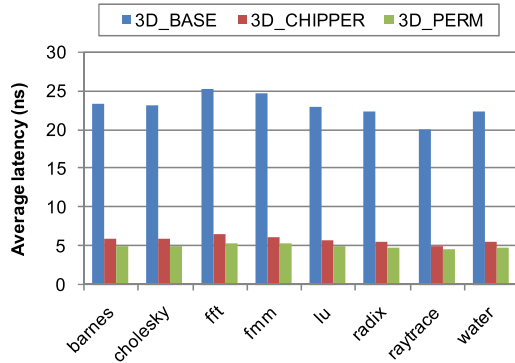


Fig. 5 Average packet latency with application workloads.

N_x , N_y and N_z are the number of nodes along each dimension. For local traffic, the resource node sends packets to the near neighbors with a higher probability than the remote nodes. The probability depends on the source-destination manhattan distance [8]. Real application traces are generated by executing the SPLASH-2 benchmark suites [9] on a full-system simulator Simics [10]. The full-system configurations are shown in Table 2. The multicore system contains 64 processors connected by a $4 \times 4 \times 4$ 3D mesh network. A 32 KB L1 I-Cache and D-Cache and a 512 KB L2 Cache are attached on each processor. The Cache coherence protocol is MOESI_CMP_directory protocol. 8 on-chip memory controllers are attached on 4 processors of the top and bottom layers respectively.

Figure 4(a)–(c) shows the average packet latency of three routers with three synthetic workloads respectively. To make a fair comparison, the clock cycle time of the three routers implemented in TSMC 65 nm technology is considered. The average packet latency, measured in *ns*, is calculated by multiplying the clock cycle time of the router and the number of cycles between the generation time of the first flit and the arrival time of the last flit. From the figure we can see that the network with the 3D_CHIPPER router and the 3D_PERM router reaches saturation point earlier than the network with the 3D_BASE router. This is due to the fact that the two routers remove the *Load_Computation* and *Output_Priority_Sort* modules, which cannot use the load-aware technique to avoid congestion. However, before the network reaches the saturation point, the average packet latency of the 3D_CHIPPER and 3D_PERM routers is much less than that of the 3D_BASE router. Compare with the 3D_CHIPPER router, the 3D_PERM router achieves 19%, 13% and 9% less average packet latency under three synthetic workloads respectively. Figure 5 shows the average packet latency of three routers with 8 Splash-2 applications. The 3D_PERM router achieves 78% and 14% less average packet latency than the 3D_BASE and 3D_CHIPPER routers respectively.

5. Conclusion

In this paper, we propose a 1-cycle high-performance bufferless router for 3D NoC. The proposed router adopts a 3-stage permutation network to replace the serialized switch allocator and 7×7 crossbar to achieve the frequency of 1.25 GHz in TSMC 65 nm technology. Compared with the other two 3D bufferless routers, the proposed router can reduce area requirement and power consumption. Simulation results under both synthetic and application workloads show that the proposed router outperforms the other two 3D bufferless routers in terms of average packet latency.

Acknowledgements

The research is partially supported by the National Natural Science Foundation of China, under Grant No.60970036, No.60873212 and No.61003301.

References

- [1] V.F. Pavlidis and E.G. Friedman, "3-D topologies for networks-on-chip," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol.15, no.10, pp.1081–1090, 2007.
- [2] M. Hayenga, N. Jerger, and M. Lipasti, "Scarab: A single cycle adaptive routing and bufferless network," *Proc. 42nd Annual IEEE/ACM Int. Symposium on Microarchitecture*, pp.244–254, 2009.
- [3] T. Moscibroda and O. Mutlu, "A case for bufferless routing in on-chip networks," *Proc. 36th Annual Int. Symposium on Computer Architecture*, pp.196–207, 2009.
- [4] C. Fallin, C. Craik, and O. Mutlu, "Chipper: A low-complexity bufferless deflection router," *Proc. Int. Symposium on High Performance Computer Architecture*, pp.144–155, 2011.
- [5] M. Millberg, E. Nilsson, R. Thid, S. Kumar, and A. Jantsch, "The nostrum backbone—a communication protocol stack for networks on chip," *Proc. IEEE Computer Society, Int. Conference on VLSI Design*, pp.693–696, 2004.
- [6] E. Nilsson, M. Millberg, J. Oberg, and A. Jantsch, "Load distribution with the proximity congestion awareness in a network on chip," *Proc. Design, Automation and Test in Europe Conference and Exhibition*, pp.1126–1127, 2003.
- [7] W.J. Dally and B. Towles, *Principles and Practices of Interconnection Networks*, Morgan Kaufmann, 2004.
- [8] Z. Lu, A. Jantsch, E. Salminen, and C. Grecu, "Network-on-chip benchmarking specification part 2: Micro-benchmark specification," *OCP-IP*, pp.7–8, May 2008.
- [9] S.C. Woo, M. Ohara, E. Torrie, J.P. Singh, and A. Gupta, "The splash-2 programs: Characterization and methodological considerations," *Proc. 22nd Annual Int. Symposium on Computer Architecture*, pp.24–36, 1995.
- [10] P.S. Magnusson, M. Christensson, J. Eskilson, D. Forsgren, G. Hallberg, J. Hogberg, F. Larsson, A. Moestedt, and B. Werner, "Simics: A full system simulation platform," *Computer*, vol.35, no.2, pp.50–58, 2002.