Partitioning Methods for Multicast in Bufferless 3D Network on Chip

Chaoyun Yao[1(✉)], Chaochao Feng[1], Minxuan Zhang[1], Wei Guo[1], Shouzhong Zhu[1], and Shaojun Wei[2]

[1] College of Computer, National University of Defense Technology, Changsha 410073, People's Republic of China
ycy021417@163.com

[2] The Department of Electronic System, Tsinghua University, Beijing 100083, People's Republic of China

Abstract. In this paper, we proposed two region partition multicast routing algorithms for the 3D mesh Interconnection Network to enhance the overall system performance. The proposed two algorithms shorten the network long path latency. Compared to the based multicast routing algorithm, our simulations with six different synthetic workloads reveal that our architecture acquires high system performance.

Keywords: NoC · 3D · Multicast

1 Introduction

The technology trends to scale with Moore's law, a single die can incorporate several hundreds of cores. When the number of core increases, interconnect becomes more and more complicated in System-on-chip design. The number of core on a single chip continues to increase to efficient utilize the process elements. Networks on chip were shown to be feasible and easy to scale for supporting a large number of cores and memory rather than shared buses or point-to-point interconnect wires [1]. However, with the increasing cores number in the 2D plane is not an efficient way due to its long wire interconnects [2,3]. The emerging of the 3D integrated circuit that stacks several dies to reduce the long wire delay, it results in lower power consumption and higher performance [4,5]. The 3D integrated circuit is also an attractive way for Network on chip design. The 3D NoC architecture is widely studied in the network topology [4,5], router architecture [6,7], and routing algorithms [8,9].

In order to run many kinds of communication applications in Multi-Processors System-on-Chip (MPSoC), collective communications must be supported. In an MPSoC system, the cache-coherent shared memory protocols (such as directory-based or token-based) require one-to-many or broadcast communications to

C. Yao—The research is supported by National Natural Science Foundation of China with Grant No. 61303066 and 61373032 and by Specialized Research Fund for the Doctor Program of Higher Education of China with Grant No. 20124307110016.

W. Xu et al. (Eds.): NCCET 2015, CCIS 592, pp. 13–22, 2016.
DOI: 10.1007/978-3-662-49283-3_2

obtain shared data or invalidate shared data on different cache blocks [10]. It has been proved that only percent 2–5 multicast traffic in the total network traffic will have a serious impact on system performance [10]. If multicast communication is supported in NoC can advance system performance significantly. Several recent research and industry products trend to support multicasting in hardware implementation.

There are three multicast mechanisms that can be classified as unicast-based, tree-based and path-based. In unicast-based multicast scheme [11], the multicast packet is partitioned into multiple unicast packets at the source node and transmits to the destinations separately. The advantage of this multicast scheme is which can be executed on the unicast routers, and the unicast routers did not require make any change. However, such a scheme is very inefficiency owe to the multiple copies of the same packet is injected into the network in turn. In path-based multicast scheme, the multicast packet selects a path to route the multicast packet to each destination sequentially until it arrives to the last node. Hence, path-based multicast routing does not block message along the path. So in path-based method is reducing the overall path length in the multi-packet [12] show that path-based routing is more beneficial in the wormhole switched network, yet in networks employing tree-based, store and forward routing is advantageous.

In this paper, we work on shorten multicast path length algorithm that is for 3D bufferless mesh architectures to enhance the overall system performance. The rest of paper is organized as follows. The related work is reviewed in Sect. 2. Section 3 proposes the general multicast schemes in bufferless 3D mesh architecture. In Sect. 4, Partition methods for is proposed. In Sect. 5, the simulation and experimental results are presented and analyzed, followed by the conclusion in Sect. 6.

2 Related Work

The fact is that the multicast communication is used to support several collective communication operations. There has been several multicast communication research in 2D NoCs. Low Distance multicast (LD) [13] algorithm is a path-based multicast communication method which optimized destination nodes order and utilized adaptive routing for multicast packets through the network. The Virtual Circuit Tree Multicasting (VCTM) [14] is a typical tree-based multicast routing method. This routing method first sends a setup packet to build a multicast tree, then sends multicast packet. Similar as VCTM multicast scheme, two phase multicast tree construct method is proposed in [15] which consume less power than VCTM algorithm. The Recursive Partitioning Multicast (RPM) [16] which supposes the network is divided into several partitions, the multicast packet selects intermediate nodes to replicate, minimizes the packet replication time. This scheme performs better than VCTM due to its more path diversities. However, it is not implemented in hardware. Feng [17] proposed bufferless multicast routing algorithms for 2D NoC architecture.

Recently, researches have considered on evaluating the performance metric of 3D NoCs. Feero [18] showed that 3D-NoC can reduce latency and the energy per packet by 40 % by decreasing the number of hopes. [5] is proved that 3D-NoC can decrease 62 % and 58 % power consumption compare to a traditional 2D-NoC topology for a network size of N = 128 and N = 256 nodes, respectively. The fact is that 3D mesh-based NoC architecture can reduction in the average wire length and wire delay resulting in better performance and lower power consumption [4].

In the open literature, there has been no work addressing bufferless multicasting for 3D NoCs systems. In this work, we first put forward a new idea of balanced partitioning in 3D bufferless mesh network for unicast/multicast traffic.

3 Preliminaries

The general path-based routing algorithm for buffer 3D mesh architecture is based on Hamiltonian path strategy. It is constructed two directed Hamiltonian Paths (high channel sub-network and low channel sub-network) [19]. However, this routing algorithm does not suit to the bufferless 3D mesh architecture. The algorithm in bufferless 3D mesh architecture is a non-deterministic path-based multicast scheme. The multicast packet is routed to every destination along a non-deterministic path. When a multicast packet arrives at a midway router, the router always selects a destination with the minimum manhattan distance to the current router from the destination address nodes. Owe to the packet would be deflected away from the shortest path to the destination, During the routing process the best destination nodes in the multicast will change dynamically. Table 1 is shown the pseudo code of the path-based bufferless 3D multicast routing algorithm. The algorithm first gains multicast destination node IDs in the packet to an array dst_id_array (Steps 1–7), then sets the first element of dst_id_array as the best candidate node ID and calculates the manhattan distance between the current node and the best candidate node (Steps 8–10). Third, acquires the minimum manhattan distance of the best candidate node to the current node (Steps 11–17). Finally, the optimal direction is based on the position of the best candidate node to the current node (Step 18).

4 Partitioning Methods for 3D Mesh Architecture

In this section, we put forward two partition methods for bufferless 3D mesh architecture. They are two block partition (TBP) and four block partition (FBP). This partition method is first applying for bufferless architecture.

4.1 Partitioning Methods

Two block partition, the network is partitioned into up block sub-network and down block sub-network. In up block sub-network contains all destination node IDs in this block. In down block sub-network contain all destination node IDs

Table 1. Path-based algorithm for bufferless 3D mesh architecture

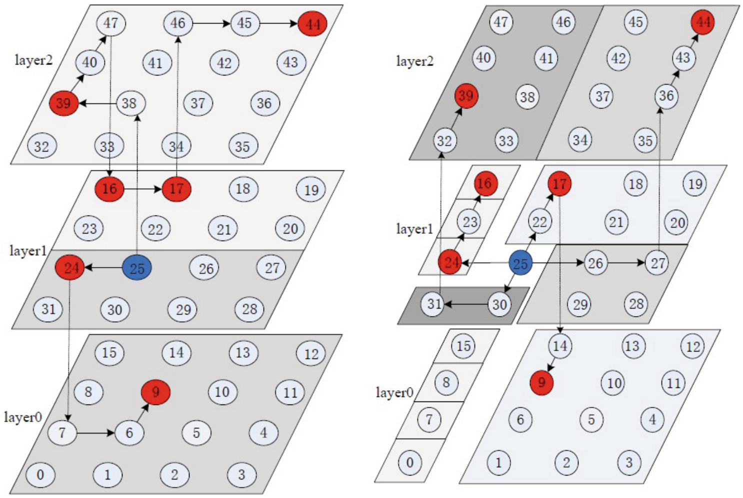
Routing computation function *for path-based bufferless 3D multicast routing algorithm*
Input*: Multicast destination address dst_add Coordinate of current node* (x_c, y_c, z_c)
Output*: Optimal direction set (doptimal)*

```
i←0
for j in 0 to N-1 loop
  if dst_add[j] = 1 then
    dst_id_array[i]←j
    i←i+1
  end if
end loop
best_candidate_id← dst_id_array[0]
(x_d, y_d, z_d)← get_dst_coordinate(dst_id_array[0])
distance← get_manhattan_distance(x_d, y_d, z_d, x_c, y_c, z_c)
for j in 1 to i-1 loop
  (x_d, y_d, z_d) ← get_dst_coordinate( dst_id_array[j] )
  if distance ¿ get_manhattan_distance(x_d, y_d, z_d, x_c, y_c, z_c) then
    distance ← get_manhattan_distance(x_d, y_d, z_d, x_c, y_c, z_c)
    best_candidate_id ← dst_id_array[j]
  end if
end loop
doptimal ← get_optimal_direction (best_candidate_id, x_c, y_c, z_c )
```

in it. All destination nodes are split into two disjoint groups use this method. They are up group and down group.

Figure 1 shows an example of this partitioning policy and the portions of each partition which depends on the source node position. If the multicast packet can inject into the network, the multicast will deflection routing in the network. When the coordinate (x, y, z) of the current node is equal to the destination node coordinate (x, y, z), the multicast arrives to the current destination and modifies the multicast destination address. In the example, it is assumed that for 3D mesh NoC architecture, the source node is at node 25 where the destination set is D = 9, 16, 17, 24, 39, 44 and for FBP multicast algorithm using our proposed multicasting strategy, the source node and the destination set is the same as the TBP algorithm. For the TBP multicast routing scheme is shown in Fig. 1(a), First Node 25 is copied into two node and separate sends to different partitions if the source node have more than two empty ports. Node 24 and 39 are chosen at the first best candidate since it has the minimum manhattan distance to the source node 25. After the packet is sent to node 24 and 39, node 16 and 9 are chosen as the second best candidate. Without contention, the multicast minimum latency is equal to 9 hops. The path shown in Fig. 1(a) is not the only one path since the packet may be deflected due to contention.

For the FBP multicast routing strategy, there is a great improvement in terms of packet latency for the same source node and destination set. As illustrated in Fig. 1(b), the destination set is partition into four parts. In this case, the latency is equal to 5 hops. The reason for such a significant improvement is that multiple region partition way is shorten the long path.



(a) 3D mesh NoC architecture using TBP multicast routing (b) 3D NoC mesh architecture using FBP multicast routing

Fig. 1. Two multicast routing where the source is at node 25

4.2 Packet Format

The router supports two packet types: unicast and multicast. The format of multicast packet type is shown in Fig. 2. The packet fields are explained as follows:

Type field (2 bits): indicate the type of the packet (10: multicast packet; 0: unicast packet; 00/11: invalid packet).
Dst_addr field: it use bit string encoding. A bit of 1 in the string means the corresponding nodes is one of destinations.
Hop_counter field (11 bits): record the number of hops the packet has been routed and used it as packet priority to avoid livelock.
Src_addr field (18 bits): denote the relative address to the source node (6 bits for row addresses, 6 bits for column addresses and the other 6 bits for layer addresses).
Payload field: the payload can be extended to contain more bits for different application requirements.

In our multicast routing, When a multicast packet arrive to a destination, the bit in the destination address corresponding to that destination node reset to 0, the message is copied and sent with its header to the next node in accordance with our proposed routing algorithm which will be shown in the following subsection. The hop counter field of the packet will be added 1.

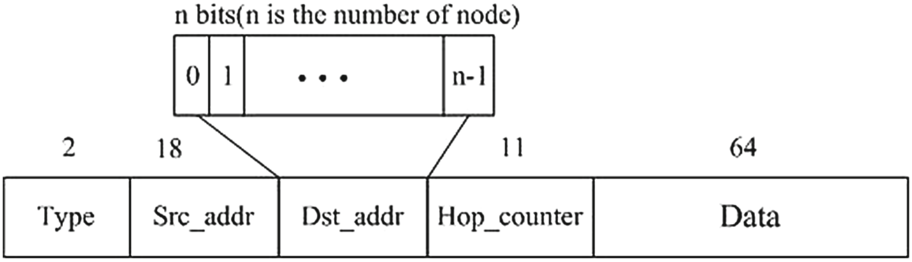


Fig. 2. Multicast packets format

5 Multicast Routing Algorithm

The TBP multicast routing of packets that takes place in the 3D bufferless NoC mesh architecture is a non-deterministic deflection routing based multicast algorithm. The pseudo code of the two block partition function for the routing scheme is shown in Table 2. The function judges the coordinate of current node and destination node, then collects multicast destination node IDs into array G1 and G2.

The FBP multicast routing of packets take place in the 3D bufferless NoC mesh architecture. The pseudo code of the algorithm for the routing scheme is shown in Table 3. The function judges the coordinate of current node and destination node, then collects multicast destination node IDs into array G1, G2, G3 and G4.

Table 2. The TBP algorithm for bufferless 3D mesh architecture

Algorithm: *Two block partitioning (TBP)*
Input: *Multicast destination address dst_add Coordinate of current node* (x_d, y_d, z_d)
Multicast current node Coordinate is (x_c, y_c, z_c)
Output: *group1 (G1) and group2 (G2)*
begin
for j in 0 to N-1 loop
if dst_add[j] = 1 then
if $z_d > z_c or (z_d = z_c and x_d < x_c) then$
G1←dst_add[i]
else
G2←dst_add[i]
end if
end if
end loop
end TBP

5.1 Deadlock and Livelock Avoidance

Deflection routing is deadlock-free, because of the packets never have to wait in a router. However, when a packet does not acquire the optimal port, it will deflect to another port away from destination node. Thus, livelock must be avoided by limiting the number of deflection. In our multicast routing algorithm, the multicast packets are prioritized based on its age that the number of hops already routed in the network. The age-based priority mechanism guarantees the oldest packet will first win the link arbitration and direct to its destination. Once the oldest packet reaches its destination, another packet becomes the oldest. Thus livelock can be avoided.

6 Performance Evaluation

We evaluate the performance of the proposed multicasting mechanism for the 3D mesh interconnection network, a cycle-accurate NoC simulator developed in VHDL. The based multicast routing scheme is use the path-based multicast routing algorithm. There multicast routing algorithms were analyzed for synthetic traffic patterns. Nostrum [19] router is as a baseline router structure. For the 3D mesh architecture, routers have 7 input/output ports. The arbitration scheme for the switch allocator is age-based.

The performance of the network was evaluated using latency curves and the packet injection rate function. The packet latency is the time duration from when the packet is created at the source node to the packet is delivered to the destination node. To perform the simulations, a packet generator is attached to

Table 3. The FBP algorithm for bufferless 3D mesh architecture

Algorithm: *Four block partitioning (FBP)*
Input: *Multicast destination address dst_add Coordinate of current node* (x_d, y_d, z_d)
Multicast current node Coordinate is (x_c, y_c, z_c)
Output: *group1 (G1), group2 (G2), group3(G3), group4(G4)*

begin
for j in 0 to N-1 loop
  if dst_add[j] = 1 then
   if $(z_d > z_c and y_d > y_c) or (z_d = z_c and x_d >= x_c and y_d > y_c) then$
    G1←dst_add[i]
   elsif$(z_d > z_c and y_d < y_c) or (z_d = z_c and x_d > x_c and y_d <= y_c) then$
    G2←dst_add[i]
   elsif$(z_d < z_c and y_d >= y_c) or (z_d = z_c and x_d < x_c and y_d >= y_c) then$
    G3←dst_add[i]
   else
    G4←dst_add[i]
   end if
  end if
end loop
end TBP

each router and uses a FIFO to buffer the packets which cannot be injected into the network immediately due to the fact that there is no free output port to route the packet. It is use A combination of unicast (80 %) and multicast (20 %) traffic. For the unicast portion of the traffic, we use six traffic patterns: uniform random, transpose, bit reverse, bit complement, shuffle, tornado. In uniform random traffic, each resource node sends packet randomly to other nodes with an equal probability. For transpose traffic, resource node coordinate is (x, y, z)(!(x = y = z)) sends package to the destination node coordinate is (z, y, x). If the four-bit source address is s3, s2, s1, s0 the destination address for bit reverse traffic is s0, s1, s2, s3 and for shuffle traffic is s2, s1, s0, s3. For bit complement traffic. The four-bit source node ID $\{si|i \in [0,3]\}$ sends packets to destination $\{\sim si|i \in [0,3]\}$. For tornado traffic, each (radix-k) digit of the destination address Dx is a function of a digit Sx of the source address, which is $D_x = S_x$ + (k/2-1) mod k. For multicast packet, the destination nodes are uniformly distributed. The simulations are performed on a $4 \times 4 \times 4$ 3D mesh architecture.

In the experiment set, the number of destination nodes has been set to 8. The average packet latency curves for uniform random (i.e. 80 % uniform random unicast 20 % uniform multicast), transpose, bit reverse, bit complement, shuffle, tornado are shown in Fig. 3. It can be observed for all six traffic patterns, that FBP and TBP multicast scheme achieves less average latency than the Path-based multicast scheme. The main reason is that the two proposed multicasting scheme shorten the long path latency.

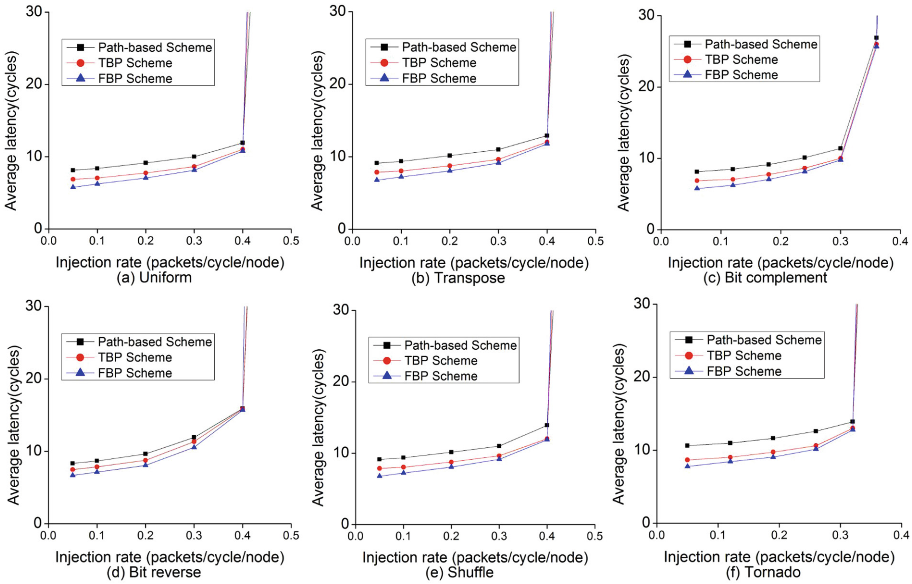


Fig. 3. Latency versus average packet inject rate with 8 multicast nodes

7 Conclusion and Future Work

In this paper, we propose two region patition multicast routing strategy for the 3D NoC mesh architecture, improving the overall NoC performance. The proposed architecture exploits region partition to shorten long path latency to provide high-performance hardware multicast support. Our simulations with different traffic profiles showed that the two multicast routing algorithms can achieve significant performance improvements over the pat-based multicast routing algorithm. In the future, our work will be extended by introducing faulty links for proposed architecture and simulating it using a set of realistic workloads. We will also measure power consumption and area for this architecture.

References

1. Dally, W.J., Towles, B.: Route packets, not wires: on-chip interconnection networks. In: Design Automation Conference (DAC), pp. 684–689 (2001)
2. Carloni, L.P., Pande, P., Xie, Y.: Networks-on-chip in emerging interconnect paradigms: advantages and challenges. In: Proceedings of the 2009 3rd ACM/IEEE International Symposium on Networks-on-Chip, pp. 93–102 (2009)
3. Rahmani, A.-M., Latif, K., Vaddina, K.R., Liljeberg, P., Plosila, J., Tenhunen, H.: Congestion aware, fault tolerant and thermally efficient inter-layer communication scheme for hybrid NoC-bus 3D architectures. In: Proceedings of the 5th ACM/IEEE International Symposium on Networks-on-Chip, pp. 65–2 (2011)
4. Feero, B.S., Pande, P.P.: Networks-on-chip in a three-dimensional environment: a performance evaluation. IEEE Trans. Comput. **58**(1), 32–45 (2009)
5. Pavlidis, V.F., Friedman, E.G.: 3-D topologies for networks-on-chip. IEEE Trans. Very Large Scale Integr. Syst. **15**(10), 1081–1090 (2007)
6. Kim, J., Nicopoulos, C., Park, D., Das, R., Xie, Y., Vijaykrishnan, N., Yousif, M., Das, C.: A novel dimensionally-decomposed router for on-chip communication in 3D architectures. In: Proceedings of the International Symposium on Computer Architecture (ISCA 2007), pp. 138–149 (2007)
7. Li, F., Nicopoulos, C., Richardson, T., Xie, Y., Narayanan, V., Kandemir, M.: Design and management of 3D chip multiprocessors using network-in-memory. In: Proceedings of the International Symposium on Computer Architecture, pp. 130–141, June 2006
8. Ramanujam, R.S., Lin, B.: Randomized partially-minimal routing on three-dimensional mesh networks. IEEE Comput. Archit. Lett. **7**(2), 37–40 (2008)
9. Moosavi, S.R., Rahmani, A.M., Liljeberg, P., et al.: Enhancing performance of 3D interconnection networks using efficient multicast communication protocol. In: 21st Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP), pp. 294–301. IEEE (2013)
10. Rahmani, A.-M., Vaddina, K.R., Latif, K., Liljeberg, P., Plosila, J., Tenhunen, H.: Generic monitoring and management infrastructure for 3D NoC-Bus hybrid architectures. In: Proceedings of the 6th ACM/IEEE International Symposium on Networks-on-Chip, pp. 177–184 (2012)
11. McKinley, P.K., Xu, H., Ni, L.M., Esfahanian, A.H.: Unicast-based multicast communication in wormhole-routed networks. IEEE Trans. Parallel Distrib. Syst. **5**(12), 1252–1265 (1994)

12. Lin, X., Ni, L.M.: Multicast communication in multicomputer networks. IEEE Trans. Parallel Distrib. Syst. **4**, 1105–1117 (1993)
13. Daneshtalab, M., Ebrahimi, M., Mohammadi, S., Afzali-Kusha, A.: Low distance path-based multicast algorithm in NOCs. IET (IEE) -Comput. Digital Tech. Spec. Issue NoC **3**(5), 430–442 (2009)
14. Jerger, N.E., Peh, L.-S., Lipasti, M.: Virtual circuit tree multicasting: a case for on-chip hardware multicast support. In: Proceedings of the 35th Annual International Symposium on Computer Architecture, pp. 229–240 (2008)
15. Hu, W., Lu, Z., Jantsch, A., Liu, H.: Power-efficient tree-based multicast support for networks-on-chip. In: Proceedings of 16th Asia and South Pacific Design Automation Conference, Piscataway, NJ, USA, pp. 363–368 (2011)
16. Wang, L., Kim, H., Kim, E.J.: Recursive partitioning multicast: a bandwidth-efficient routing for networks-on-chip. In: International Symposium on Networks-on-Chip (NOCS), San Diego, CA, May 2009
17. Chaochao, F., Zhonghai, L.U., Jantsch, A., et al.: Support efficient and fault-tolerant multicast in bufferless network-on-chip. IEICE Trans. Inf. Syst. **95**(4), 1052–1061 (2012)
18. Feero, B., Pande, P.P.: Performance evaluation for three-dimensional networks-on-chip. In: Proceedings of IEEE Computer Society Annual Symposium On VLSI (ISVLSI), pp. 305–310, 9th-11th May 2007
19. Millberg, M., Nilsson, E., Thid, R., Kumar, S., Jantsch, A.: The nostrum backbone-a communication protocols stack for networks on chip. In: Proceedings of IEEE Computer Society, International Conference on VLSI Design, pp. 693–696 (2004)

http://www.springer.com/978-3-662-49282-6

Computer Engineering and Technology
19th CCF Conference, NCCET 2015, Hefei, China,
October 18-20, 2015, Revised Selected Papers
Xu, W.; Xiao, L.; Li, J.; Zhang, C. (Eds.)
2016, XII, 191 p. 91 illus., Softcover
ISBN: 978-3-662-49282-6