

# Turn Prohibition Routing Investigation for Irregular, 2D-Mesh and 3D-Mesh Based Network on Chip

Naveen Choudhary

**Abstract**— *Network on Chip (NoC) has established itself as an alternative to the on chip bus to meet the increasing requirements of complex communication needs of system on Chip (SoC). A popular choice of topology for generic Network on Chip has been 2D Meshes. Similarly for application specific Network on Chip irregular topologies customized to application needs is preferred. However as the feature size continue to shrink and integration densities continue to increase, the interconnect delay is emerging as the critical bottleneck for the performance of 2D NoC. The advances in technology such as over the cell routing and Through-Silicon-Vias (TSV) has made possible performance conscious and scalable Network on Chip with more than 2 dimension. As was the case with 2D Mesh NoC, the 3D Mesh NoC is proving to be a preferred choice for the NoC designers due to its simple and scalable design.*

*The communication over the Network on Chip is required to be deadlock and livelock free. Turn prohibition based routing function are a popular choice for NoC communication as it provides deadlock free communication over the NoC without the requirement of additional physical or virtual channels. Moreover turn prohibition based routing is capable of providing deadlock free, livelock free, minimal or nonminimal and maximally adaptive communication over NoCs. Turn prohibition routing is based on analyzing the directions in which packets can turn in the network and the cycles that the turns can form. Prohibiting just enough turns to break all the resource dependence cycles in the network can help researchers design an effective and efficient deadlock and livelock free routing functions for the NoCs. This paper presents an investigation of the various popular turn prohibition based routing algorithms presented in the NoC research literature for 2D mesh, 3D mesh and irregular topology based on chip networks.*

**Index Terms**— *Turn Model, Routing, Network-on-Chip, Livelock, Deadlock.*

## I. INTRODUCTION

Shrinking feature size has allowed Systems-on-Chip (SoCs) [1] systems to grow continuously in scale and complexity. The conventional bus based systems were no more suitable as communication infrastructure for these complex SoC, due to their lack of communication efficiency, scalability, parallelism and power dissipation. Network-on-Chip [2] were introduced to address these problems of bus based systems. NoC connects processors, memories and other processing elements together using switching and routing packets on a hop-by-hop basis to offer higher bandwidth and higher performance.

The design of the topology or interconnection network

plays an important role in the communication performance of the complex systems on chip. Generally NoC with standard topologies such as meshes, tori, k-ary n-cubes or fat trees are favored as the the channels can be well structured in such topologies and moreover the generic architecture of such topologies allow mass production of the components reducing design efforts and time to market. The most popular regular topology proved to be meshes due to its simple and scalable design. Although regular NoCs were appropriate for general purpose systems where traffic characteristics cannot be predicted statically in advance, as in homogeneous chip multiprocessors. However, SoCs can also be heterogeneous, with each core having different size, functionality and communication requirements. In such cases regular or standard topology will not be the appropriate choice as it will poorly matches the application traffic leading to large wiring complexity after floorplanning, as well as significant power and area overhead. For such application, an irregular NoC custom designed according to the communication demands of the application will be more appropriate.

However future applications may include hundreds of processing elements and are predicted to be even more complex requiring improved communication infrastructure than 2D NoCs. The basic limitation of 2D NoC is large diameter as the number of processing elements increases exponentially. The advances in technology such as over the cell routing and Through-Silicon-Vias (TSV) has provided a solutions in the form of performance conscious 3D-NoC [3] making it feasible to keep the diameter of NoC in check even for hundreds of processing elements and a potential solution to resolve the interconnect bottleneck. Moreover 3D integrated circuit can achieve higher packaging density due to the availability of the third dimension. A 3D integrated circuit is a stack of multiple component layers with direct vertical interconnects tunneling through them [4]. The most popular 3D-NoC happens to be 3D meshes due to its simple and scalable design.

NoC generally require routing functions, which exhibit low packet latency, high throughput, and ease of implementation. To achieve the above, routing functions should be deadlock and livelock free. This paper presents a survey of turn Prohibition based routing proposed in the NoC research literature for 2D Mesh NoCs, 3D Mesh NoCs and irregular NoCs. The turn prohibition based routing is capable of providing deadlock free, livelock free, minimal or nonminimal and maximally adaptive communication over NoCs. Section 2 investigates the characteristics of deadlock and how turn prohibition based routing can address this issue.

**Manuscript received on May, 2013.**

Naveen Choudhary, Department of Computer Science and Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, Udaipur, Rajasthan, India.

Section 3 presents various turn model based routing proposed in the NoC literature for 2D Mesh NoCs. Section 4 presents the finding in the research literature of NoC regarding applicability of turn model based routing in irregular NoC. Section 4 presents various extension proposed in the literature of turn model based routing function from 2D NoC to 3D NoC. In section 5 a brief conclusion is presented.

II. NOC ROUTING AND DEADLOCKS

In NoC packets usually traverse several switches/routers before reaching the destinations. However, it may happen that some packets are not able to reach their destination, even if there exists a fault-free path connecting the source and destination for every packet. Assuming that the routing algorithm is able to use those paths, there are several situations that may prevent packet delivery like livelock, starvation and deadlock to name the few. Among these deadlock is the most significant. Deadlock [5] can be defined as a situation where each packet in the network whose header has not already arrived at its destination requests some buffers while keeping the buffers currently storing the packet, a deadlock may arise. A deadlock occurs when some packets cannot advance toward their destination because the buffers requested by them are full. Starvation may be defined as a situation when a packet may be permanently stopped if traffic is intense and the resources requested by it are always granted to other packets. Similarly livelock may be defined as a situation when some packets are not able to reach their destination, because the channels required to do so are always occupied by other packets. Livelock can be avoided by always using only the minimal path or by limiting the number of misrouting operation.

To handle deadlock we can use the strategy of deadlock prevention, deadlock avoidance or deadlock recovery. Most of the proposed strategy of routing in NoCs use deadlock avoidance [5] as it is reasonably efficient and has manageable overhead in terms of required resources. In Deadlock avoidance resources are requested as a packet advances through the network. However a resource is granted to a packet only if the resulting global state is safe. The deadlock can occur when there are cyclic channel dependencies among the channels of the network. When a packet is holding a channel and then it requests the use of another channel then there is channel dependency between these channels. If the network is as shown in fig. 1(a) and the corresponding routing function is defined as : if the current node  $n_i$  is equal to the destination node  $n_j$ , store the packet, otherwise use  $C_i$ , for all  $j \neq i$  then there is a cyclic dependency among the channels as the flits can be holding the channel  $C_i$  and requesting  $C_{i+1}$  as shown in fig. 1(b). Therefore if routing algorithm is designed in such a way that there are no cyclic channel dependencies in the network then the occurrence of deadlock can be avoided [5].

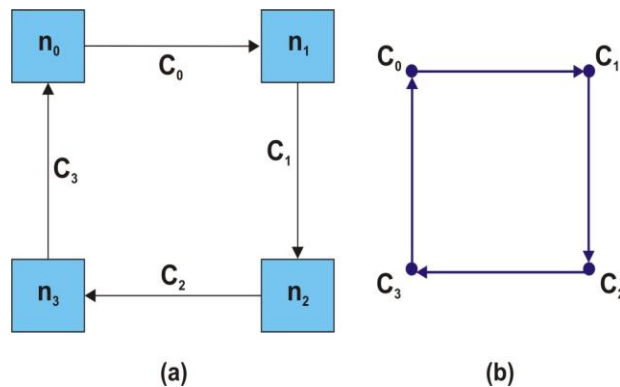


Fig. 1 : (a) NoC network (b) Channel dependency graph according to routing function 1.

The turn model proposed in [6] can be used to prohibit a set of turns in the network in such a way that network remains connected but there is no cyclic channel dependency in equivalent channel classes of the network. For example for XY routing, the turns shown with dashed line in fig. 2 can be prohibited to achieve deadlock free communication in 2D mesh NoCs.

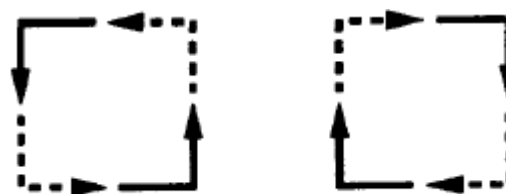


Fig. 2: Four allowed turns (solid lines) in XY routing for 2D Mesh NoCs

The generic steps for designing turn prohibition based routing for various topologies of NoC can be enumerated as follows [6].

1. Partition the channels in the network into sets according to the directions in which they route packets. If each node has  $u$  channels in a physical direction, treat these channels as being in  $u$  distinct virtual directions and divide them into  $u$  distinct sets accordingly. Put any wraparound channels in a separate set to be incorporated during Step 5.
2. Identify the possible turns from one virtual direction to another, ignoring 180-degree and 0-degree turns. A 0 degree turn is only possible when there are multiple channels in one direction. It represents a transition from one set of channels to another when the two sets route packets in the same physical direction, but different virtual directions.
3. Identify the cycles that these abstract turns can form. Generally, identifying the simplest cycles in each plane of the topology is adequate.
4. Prohibit one turn in each abstract cycle so as to prevent deadlock. The turns must be chosen carefully in order to break every possible cycle, including complex cycles not identified in Step 3. A useful approach is first to break the cycles in each plane and then to check whether this allows more complex cycles.

5. Incorporate as many turns as possible from the set of wraparound channels, without reintroducing cycles. At least one turn for each wraparound channel can always be incorporated.
6. Incorporate as many 180-degree and 0-degree turns as possible, without reintroducing cycles.

### III. TURN BASED ROUTING & 2D MESH NOCS

In [6] various turn based routing for 2D Mesh NoC are proposed. There can be four directions (North, South, East, West) and therefore eight 90-degree turns in 2D Mesh NoCs. We need to prohibit at least 1 turn from each possible cycles to achieve deadlock free routings. Based on this various deadlock free routing as shown below are proposed in [6].

Let us denote the communication source node coordinates as  $(X_s, Y_s)$  and Destination node coordinates as  $(X_d, Y_d)$ .

In **XY routing** (fig. 3(a)) the packets are first routed in horizontal direction, till the packet reaches the  $Y_d$  coordinates. After reaching the  $Y_d$ , the packet vertically moves to its destination in Y axis direction.

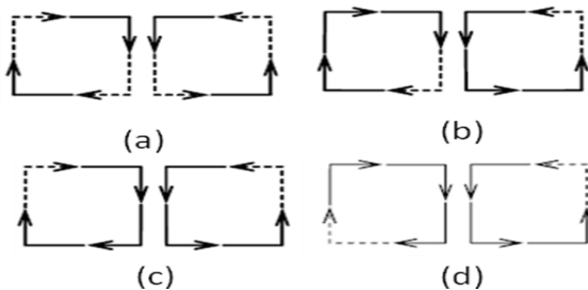


Fig. 3: Turn Model with dashed lines as prohibited turns (a) XY routing, (b) West first routing, (c) North Last routing, (d) Negative first routing

In **West-First routing** (fig 3(b)) if the X coordinates of the destination node is less than or equal to the X coordinates of the source node then packets are routed similar to XY routing algorithm otherwise the packets can be routed adaptively in all the directions (N, S, E) other than the west (W) direction.

In **North-Last routing** (fig. 3(c)) a packet should travel north when that is the last direction it needs to travel i.e. route a packet first adaptively west, south and east, and at last when there are no more turns required to be taken then travel north if required.

In **Negative first routing** (fig. 3(d)), a packet should be routed first adaptively to the west and south direction first if required and then adaptively to the east and north directions.

### IV. TURN BASED ROUTING & 2D IRREGULAR NOCS

The popular routing algorithms with irregular topologies such as prefix routing [7], up\*/down\* routing [8], L-turn routing [9], DOWN/UP routing [10] basically use the turn model to prohibit the certain turns among the equivalent channel classes of the irregular NoCs to avoid deadlocks. The equivalent channel classes in these algorithms are designed based on the direction and dimension of the channels. These routing algorithms do not necessarily rely on the availability of the virtual channels. We elaborate a popular turn model based routing for irregular NoCs and that is up\*/down\* routing.

**Up\*/down\* Routing:** The interconnection network between switches can be modeled by a multigraph  $G(N, C)$ ,

where  $N$  is the set of switches and  $C$  is the set of bidirectional links between the switches as shown in fig. 4.

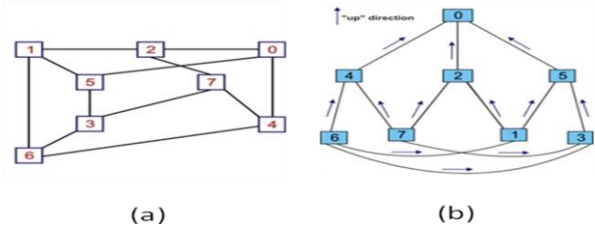


Fig. 4: (a) Irregular NoC, (b) Tree representation of (a) with appropriate direction(up/down)

Up\*/down\* distributed table based routing provides partially adaptive communication between the nodes of an irregular network. The routing table in each switch must be filled before messages can be routed. To do so, a breadth-first spanning tree (BFS) on the graph  $G$  is computed first. Routing is based on an assignment of direction to the channels, including the ones that do not belong to the tree. In particular, the "up" end of each link is defined as: 1) the end whose switch is closer to the root in the spanning tree; 2) the end whose switch has the lower ID, if both ends are switches at the same tree level. Channels looped back to the same switch are omitted from the configuration. The result of this assignment is that each cycle in the network has at least one link in the "up" direction and one link in the "down" direction. To eliminate deadlocks while still allowing all links to be used, the up\*/down\* routing algorithm uses the following rule: a legal route must traverse zero or more links in the "up" direction followed by zero or more links in the "down" direction. Thus, cyclic dependencies between channels are avoided because a message cannot traverse a link along the "up" direction after having traversed one in the "down" direction. Such routing not only guarantees deadlock-freedom, but also provides some adaptivity. The lookup tables can be constructed to support both minimal and nonminimal adaptive routing. However, in some cases, up\*/down\* routing is not able to supply any minimal path between some pairs of switches, as shown in the following example.

Fig. 4(b) shows the example tree representation of the irregular network shown in fig. 4(a). Switches are arranged in such a way that all the switches at the same tree level are at the same vertical position in the figure. The root switch for the corresponding BFS spanning tree is switch 0. The assignment of "up" direction to the links in the network is illustrated. The "down" direction is along the reverse direction of the up link.

Note that every cycle has at least one link in the "up" direction and one link in the "down" direction. It can be observed that all the alternative minimal paths are allowed in some cases. This is the case when the destination host is connected to the root switch. For example, a message transmitted from switch 7 to switch 0 can be routed either through switch 4 or switch 2. In some other cases, however, only some minimal paths are allowed. For example, a message transmitted from switch 2 to switch 5 can be routed through switch 0 but it cannot be routed through switch 1. It should be noted that any transmission between adjacent switches is always allowed to use the link(s) connecting them, regardless of the

direction assigned to that link. However, when two switches are located two or more links away, it may happen that all the minimal paths are forbidden. This is the case for messages transmitted from switch 4 to switch 1. The only minimal path (through switch 6) is not allowed, because it requires one transition from “down” to “up” direction. All the allowed paths (through switches 0, 2, and through switches 0, 5) are nonminimal since they require three hops, while the illegal path through switch 6 requires only two hops.

However although the up/down routing can not be considered minimal, it generally provides adaptivity to a good extent and so is a popular choice for application specific irregular NoCs.

V. TURN BASED ROUTING & 3D NOCS

C. J. Glass, L. M. Ni in [11] proposed an extension of 2D mesh turn prohibition based routing model for 3D mesh NoC. It was suggested in [11] that there are 24 possible 90-degree turns, four for each of the six directions in 3D Mesh leading to six abstract cycles i.e. two cycles in each of the three planes, as shown in Fig. 5. We need to prohibit at least one appropriate turn in each cycle to avoid communication deadlock situations.

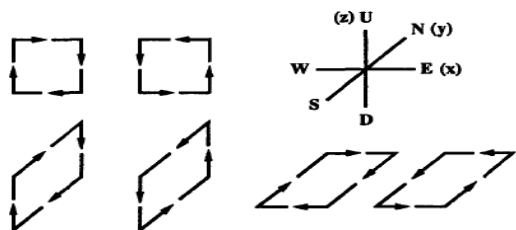


Fig. 5: Possible communication resource dependence cycles in 3D mesh.

The following turn model based routing for 3D mesh was proposed in [11].

**West South First:** Route a packet first adaptively west and south and then adaptively down, east, north, and up.

**North Up Last:** Route a packet first adaptively west, south, down, and east and then adaptively north and up.

**Negative First:** Route a packet first adaptively west, south, and down and then adaptively east, north, and up.

The turn model for the above mentioned turn model based 3D NoC routing is shown in fig. 6.

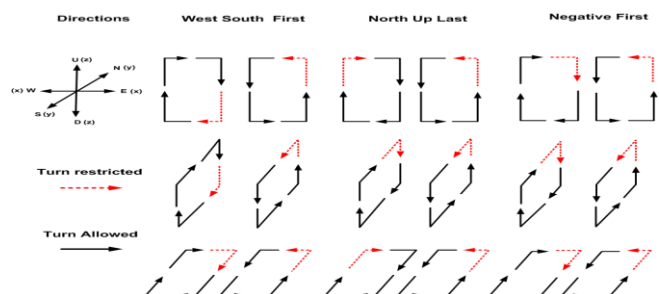


Fig. 6: Popular 3D NoC Turn model based routing with possible cycles in each plane and prohibited turns (dashed lines)

Similarly XYZ routing for 3D NoC can be explained as to route the packets first in X Direction, then in the Y direction and lastly in the Z direction towards its destination.

The algorithms for 3D meshes are more likely to be able to route any particular packet adaptively than are the algorithms for 2D meshes [11]. Overall, a 3D mesh with the same number of nodes as a 2D mesh has a lower average degree of

adaptivity but distribution of the adaptivity more uniformly distributed across the paths of source-destination pairs [11].

VI. CONCLUSION

The paper surveys the applicability and usefulness of the turn prohibition based routing in the popular 2D Mesh, application specific Irregular and 3D Mesh based NOCs. As presented in the paper it is obvious that prohibiting the minimum number of turns that break all the cycles can help us generate routing functions which are not only deadlock and livelock free but can also be minimal or non minimal and can also be maximally adaptive. Moreover the turn model, unlike some other routing function do not have the necessity of additional physical or virtual channels to avoid deadlock or livelock. However for adaptive turn model based routing the router logic can be complex for selecting the outgoing port leading to increased delay in communication.

ACKNOWLEDGMENT

This work is supported by Department of Science and Technology, Jaipur, Rajasthan, India under the research project “Network-on-Chip simulation framework for regular, irregular and 3D-Mesh Interconnection Architecture”

REFERENCES

- [1] L. Benini and G. De Micheli, “Networks on Chips: A New SoC Paradigm” in IEEE Computer, pp. 70-78, Jan. 2002.
- [2] W. J. Dally and B. Towles, “Route packets, not wires: on-chip interconnection networks,” in Proceedings of the 38th conference on Design automation, June 2001, pp. 684–689.
- [3] G. Philip, B. Christopher, and P. Ramm, Handbook of 3D Integration: Technology and Applications of 3D Integrated Circuits, Wiley-VCH, 2008.
- [4] P. Morrow, M. Kobrinsky, S. Ramanathan, C.-M. Park, M. Harnes, V. Ramachandrarao, H. Park, G. Kloster, S. List, and S. Kim, “Wafer-Level 3D Interconnects Via Cu Bonding”, in Proceedings of the 21st Advanced Metallization Conference, Oct. 2004.
- [5] J. Duato, S. Yalamanchili, L. Ni, Interconnection Networks : An Engineering Approach, Elsevier, 2003.
- [6] C. J. Glass, L. M. Ni, “The Turn Model for Adaptive Routing”, In the Journal on ACM, Vol. 41, No. 5, pp. 874-902, Sept. 1994.
- [7] J. Wu, L. Sheng, “Deadlock-Free Routing in Irregular Networks Using Prefix Routing”, DIMACS Tech. Rep. 99-19, Apr. 1999.
- [8] e. a. M. D. Schroeder, “Autonet: A High-Speed Self-Configuring Local Area Network Using Point-to-Point Links”, In Journal on Selected Areas in Communications, vol. 9, Oct. 1991.
- [9] A. Jouraku, A. Funahashi, H. Amano, M. Koibuchi, “L-turn routing: An Adaptive Routing in Irregular Networks”, In the International Conference on Parallel Processing, pp. 374-383, Sep. 2001.
- [10] Y.M. Sun, C.H. Yang, Y.C. Chung, T.Y. Hang, “An Efficient Deadlock-Free Tree-Based Routing Algorithm for Irregular Wormhole-Routed Networks Based on Turn Model”, In International Conference on Parallel Processing, vol. 1, pp. 343-352, Aug. 2004.
- [11] C. J. Glass, L. M. Ni, “Adaptive Routing in Mesh Connected Networks”, in proceedings of 12<sup>th</sup> international conference on Distributed Computing Systems, pp. 12-19, 1992.



**Dr. Naveen Choudhary** received his B.E, M.Tech and PhD degree in Computer Science & Engineering. He completed his M.Tech from Indian Institute of Technology, Guwahati, India and PhD from Malviya National Institute of technology, Jaipur, India in 2002 and 2011 respectively. Currently he is working as Professor and Head, Department of Computer Science and Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, Udaipur, India. His research interest includes Interconnection Networks, Network on Chip, Distributed System and Information Security. He is a life member The Indian Society of Technical Education, Computer Society of India and The Institution of Engineers, India. E-mail: naveenc121@yahoo.com

