# EbDa: A New Theory on Interconnection Networks

## Abstract

*Dally's theory has been used during the last three decades, saying that a design is deadlock-free if there is no cycle in the channel dependency graph. However, to prove such acyclic graph, rigorous analysis are needed using turn models. We introduce a unified theory, called EbDa, that directly allows designing an acyclic channel dependency graph and verifying algorithms on their freedom from deadlock. EbDa is composed of three theorems that enable extracting all allowable turns without dealing with turn models any more. Theorem1 introduces a new perspective on the formation of cycles. Theorem2 suggests a structural way of enabling U- and I-turns. Theorem3 traces the group of channels in an ascending order that permits new turns.*

## 1 Motivation

An interconnection network consists of a set of routers and links where a topology determines the arrangement of these network elements. Deadlock may occur in the network due to a cyclic dependency between channels such that each packet holds a channel needed by another packet [1, 4]. The two predominant theories in the area of on-chip/off-chip interconnection networks are Dally's theory (introduced in 1987 [2]) and Duato's theory (introduced in 1993 [3]). Dally proposed a methodology and applied it to deterministic routing, showing that a necessary and sufficient condition to design *a deadlock-free routing algorithm* is to remove cyclic dependencies on the channel dependency graph [2]. The turn model concept, defined by Glass and Ni, has enabled Dally's theory to be adapted for adaptive routing [5]. Since then, turn models have been extensively utilized in designing both deterministic and adaptive routing algorithms. However, Dally's theory does not scale well to large network sizes where it is not feasible to check all possible channel dependencies. We solve the scalability limitations of Dally's theory to networks with arbitrary large dimensions. Our theory show how to directly design an acyclic channel dependency graph given the available channels in the network. On the other hand, based on Duato's theory a necessary and sufficient condition to design *a deadlock-free fully adaptive routing* is the existence of a cycle-free subset of channels. All the other channels can be used with no restrictions. In more details, two types of channels are used as adaptive and escape. The header flit should be always in the head of the queue so that packets can always be transferred to the escape channels that are cycle-free [6] and thus avoiding deadlock. This assumption poses a strong limitation on wormhole switching as an input buffer must be emptied before receiving a new packet.

EbDa does not introduce a new routing algorithm but instead it shows a roadmap to design deadlock-free routing algorithms in a wormhole switching network. EbDa is a combination of three theorems that altogether remove all cyclic dependencies on the channel dependency graph and directly suggest deadlock-free routing options. We design an acyclic channel dependency graph in the network by an inspiration from a mathematical observation on the formation of cycles in a geometrical space. Based on this observation, the necessary condition to form a cycle is the availability of channels in both positive and negative directions. For example, a square or rectangular shape cannot be formed if any of the $X^+$, $X^-$, $Y^+$, and $Y^-$ channels is missing. In short, the whole process of designing a deadlock-free routing algorithm is as simple as dividing channels into disjoint cycle-free partitions and then tracing the partitions in an ascending or descending order. Based on these theorems different deadlock-free routing algorithms can be designed or algorithms can be verified on their freedom from deadlock.

## 2 The EbDa Theory

The EbDa theory introduces a new perspective on designing deadlock-free routing algorithms in networks which places no restrictions on the number of virtual channels (VCs), network topology, size, (ir)regularity, dimension, etc. We define some notations as *partition*, *D-pair*, and *channel dependence* to better describe the theorems. A *partition* covers a set of channels in an $n$-dimensional network where packets can take any channel inside the partition arbitrarily and repeatedly. A *D-pair* is formed if a partition covers both positive and negative directions of the $D$ dimension: $\mathbf{P} = \{D^+, D^-\}$. Two partitions are called disjoint with no *dependency* if they do not share any common channel with each other. In the finest level of granularity, each router-to-router link is disjoint from any other one in the network unless a connection is made by the means of 90-degree, U-turns (i.e. formed by a transition between channels in opposite directions), or I-turns (i.e. formed by a transition between channels in the same direction). For the sake of simplicity we increase the level of granularity from router-to-router links to channels (i.e. composed of a set of router-to-router links along the same dimension and with the same VC number). With this definition, two partitions are called disjoint if their channels, among the others, 1) are located in different layers or in different rows or columns; 2) have different VC numbers; 3) are located in opposite directions or belong to different dimensions.

EbDa is built based on three theorems as follows:

---

[†] *EbDa: A New Theory on Design and Verification of Deadlock-free Interconnection Networks*. M. Ebrahimi, M. Daneshtalab, In proceedings of the 44th International Symposium on Computer Architecture (ISCA), pages 703–715, 2017.

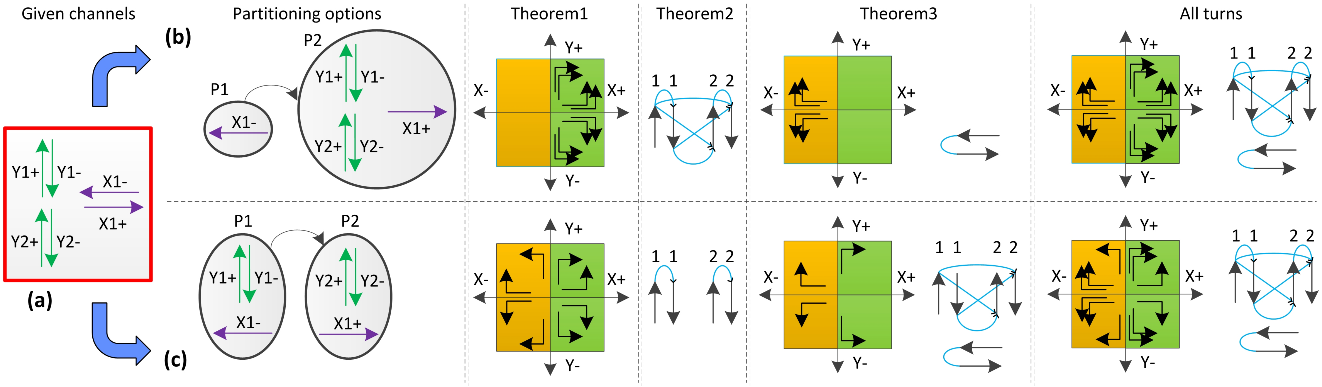

**Figure 1. An example of applying Theorem1, 2, and 3.**

- **Theorem1:** *A partition is cycle-free if it covers at most one D-pair in an n-dimensional network where* $\mathbf{D} = \{X, Y, \cdots, D_n\}$ *as long as no U- and I-turns are concerned.* A partition may cover several D-pairs as long as they represent the same dimension.
- **Theorem2:** *A partition is cycle-free if one U-turn is allowed per D-pair, taken in an ascending order. I-turns should be taken in an ascending order if a D-pair is presented along the D dimension. Otherwise, all I-turns within a partition are allowed.*
- **Theorem3:** *Transitions between disjoint acyclic partitions in a consecutive order do not form a cycle.*

**Example:** Let us assume that a 2D network is given with 1 and 2 VCs along the X and Y dimensions, respectively, as shown in Figure 1(a). The objective is to find a deadlock-free routing algorithm with maximum adaptiveness. For simplicity, $X^+$, $X^-$, $Y1^+$, $Y1^-$, $Y2^+$, and $Y2^-$ are replaced with $E$, $W$, $N1$, $S1$, $N2$, and $S2$, respectively, so that for instance a turn from $Y2^+$ to $X^-$ is represented by $N2W$.

If all channels are placed inside one partition (Figure 1(a)), Theorem1 is violated and there is a possibility of a complete cycle as two D-pairs (X-pair and Y-pair) are available inside the partition. Figures 1(b) and (c) show two possible ways of cycle-free partitioning. In the partitioning option of Figure 1(b), P1 covers $X^-$ and P2 covers the remaining channels. Based on Theorem1, to form a cycle in an $XY$ plane, it is necessary to take all four directions as $X^+$, $X^-$, $Y^+$, and $Y^-$. If any of these direction is not taken by packets, deadlock cannot be formed. Thereby, P2 is deadlock-free as the west channel is missing and at most one pair is completed (Y-pair). As the result, Theorem1 enables eight 90-degree turns as $EN1$, $ES1$, $EN2$, $ES2$, $N1E$, $S1E$, $N2E$, and $S2E$ by taking channels in P2 in any order (Figure 1(b)-Theorem1). Theorem2 allows some U- and I- turns by tracing channels of the D-pair channels in an ascending order. This enables four U-turns as $N1S1$, $N1S2$, $S1N1$, $N2S2$ and two I-turns as $N1N2$ and $S1S2$ (Figure 1(b)-Theorem2). While I-turns are useful to improve the performance, enabling U-turns is important in fault-tolerant designs or where rerouting brings an advantage. In addition, topologies with wrap-around channels such as Torus take benefits of U-turns. Theorem3 allows transitions from one partition (P1) to another (P2) in an ascending order. This transition enables four 90-degree turns as $WN1$, $WS1$, $WN2$, and $WS2$ and one U-turn as $WE$ (Figure 1(b)-Theorem3). Finally all allowable turns are illustrated in the last column of Figure 1(b). With a similar procedure, Figure 1(c) leads to a new set of turns.

## 3 Implementation Aspects

Using EbDa the whole process of designing a deadlock-free routing algorithm will be as simple as dividing channels into disjoint partitions and allowing transitions between partitions in a consecutive order. By minimizing the number of partitions and maximizing the number of channels inside a partition, maximum adaptiveness can be obtained. One way of forming the optimal partitioning is to sort dimensions based on the number of channels they cover and then form the first partition by selecting two channels (a D-pair) from the first partition and one channel from any other dimensions, if existed. The selected channels are removed and the procedure is repeated to form new partitions.

EbDa helps in finding compatible turns that otherwise might be found by heuristic methods or trial and error. So it does not impose any special hardware overhead as all connections between input and output ports are already supported by the current router architectures. Regarding the implementation and complexity aspects of the routing algorithm, turns can be expressed by some *if-else* statements [4]. Adding turns may simplify (combining *if-else* statements) or complicate (adding *if-else* statements) the routing logic.

## 4 Significance and Long-Term Impact

An interconnection network is in the heart of systems that are composed of several cores, either on-chip or off-chip. It determines the way the processing elements are connected and communicate with each other. Interconnection network is one of the main indicators of reliability and performance in different systems. However, designing an optimal deadlock-free routing algorithm is very challenging considering the increased number of cores, the expected performance level, the choice of topology, and the irregularity involved in connecting the cores. Most of the current interconnection networks are designed based on Duato's and Dally's theory. While Duato's theory relies on the scape channels and limits the number of packets to one at each buffer, Dally's theory is not scalable to large network sizes and irregular networks. Therefore, many of the current algorithms are limited to dimension-order routing or heuristic methods. EbDa is scalable and can offer deadlock-free algorithms from deterministic routing to maximally adaptive routing for the networks with any number of virtual/physical channels, dimensions, and sizes. This opens a new direction in the design of routing algorithms which was so far limited to simple approaches due to the complexity of finding an acyclic channel dependency graph. To sum, we anticipate that EbDa can have its impact in three directions as:

- **Industry:** The theorems by Duato and Dally have had huge impact on commercial systems such as *on-chip parallel systems* (e.g. GPUs and many-cores like Tilera) and *supercomputers* (e.g. Cray T3E and IBM BlueGene/L supercomputer). We expect that EbDa will be utilized in the network design of future commercial products or the network enhancement of current products. The usage of EbDa is motivated by a better system performance and a higher system reliability at no compromise such as power or area. EbDa is a new finding that can explain the basic concept behind the design of deadlock-free algorithms. Thereby, we look forward to see a revolution on designing routing algorithms from simple deterministic ones to more advanced algorithms that are capable of taking advantage of the full capacity of network resources. In addition, EbDa can be used for the verification of heuristic algorithms on their freedom from deadlock. The importance of EbDa is more obvious when considering 1) new trends such as many-core systems that require efficient routing for hundreds to thousands of cores; 2) technology limits such as dark silicon era that pose topology changes by turning the components on and off at run time; and 3) new enabled technologies such as 3D IC designs that demand advanced algorithms to take advantage of available resources [7].

- **Research:** During the last three decades the interconnection network area of research has been built based on the turn model which could have been built based on the channel model instead, like EbDa. Since the concept of the turn model lacks scalability, the improvement in this direction has reached its ever increased level of complexity which can be observed by thousands of published articles, each dealing with a new routing algorithm for a specific network requirement (e.g. number of channels, size, dimension, missing links and routers). We believe that EbDa will shift the research in this area from solving single isolated problems to more fundamental works such as advancing the theory to more radical topologies or increasing the adaptiveness of the algorithms even further by modifying the router architecture.

- **Education:** We predict that EbDa will be soon enter the educational system for the courses concerning the interconnection networks. This is due to the fact that EbDa is a fundamental and general approach that can explain not only the concept behind the turn models but also more complex designs that are beyond the capability of turn models.

Considering the number of citations on Dally's theory [2] (2594 citations in about 30 years) and Duato's theory [3] (946 citations in about 14 years), we estimate that EbDa will receive around 700 citations in ten years.

## References

[1] W. Dally and B. Towles. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003.

[2] W. J. Dally and C. L. Seitz. Deadlock-free message routing in multiprocessor interconnection networks. *Computers, IEEE Transactions on*, C-36(5):547–553, May 1987.

[3] J. Duato. A new theory of deadlock-free adaptive routing in wormhole networks. *IEEE Trans. Parallel Distrib. Syst.*, 4(12):1320–1331, Dec. 1993.

[4] J. Duato, S. Yalamanchili, and L. Ni. *Interconnection Networks: An Engineering Approach*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002.

[5] C. J. Glass and L. M. Ni. The turn model for adaptive routing. In *Proceedings of the 19th Annual International Symposium on Computer Architecture*, ISCA '92, pages 278–287, New York, NY, USA, 1992. ACM.

[6] P. Lopez, J. M. Martinez, and J. Duato. A very efficient distributed deadlock detection mechanism for wormhole networks. In *High-Performance Computer Architecture, 1998. Proceedings., 1998 Fourth International Symposium on*, pages 57–66, Feb 1998.

[7] R. Salamat, M. Khayambashi, M. Ebrahimi, and N. Bagherzadeh. A resilient routing algorithm with formal reliability analysis for partially connected 3d-nocs. *IEEE Transactions on Computers*, 65(11):3265–3279, Nov 2016.