A 3d stacked mesh NoC for reliable inter-layer communication and congestion reduction

K.A. Karthigeyan[1] and S. Sudhakar[2]

[1]Department of Electronics and Communication Engineering.

[2]Veltech Multitech Dr.Rangarajan Dr, Sakunthala Engineering College Chennai, India.

ABSTRACT

The increasing viability of 3D silicon integration technology has opened new opportunities for chip architecture innovations. One direction is in the extension of 2D mesh based chip multiprocessor architecture into three Dimensions. We present an efficient architecture to optimize system performance, power consumption, and reliability of stacked mesh 3D NoC is proposed. Stacked mesh is a feasible architecture which takes advantage of the short inter-layer wiring delays, while suffering from inefficient intermediate buffers. To cope with this, an inter-layer communication mechanism is developed to enhance the buffer utilization, load balancing, and system fault-tolerance. The mechanism benefits from a congestion-aware and bus failure tolerant routing algorithm for vertical communication.

Introduction

Ever increasing requirements on electronic systems are one of the key factors for evolution of the integrated circuit technology. Continuous technology scaling has made it possible to integrate billions of transistors on a single chip [1]. Thus the entire system with hundreds of components can be integrated on a single chip, which is known as a Multiprocessor System-on-Chip [2]. At such integration levels, communication plays a major role in the design and performance.

One outcome of higher integration levels is that interconnection platforms are replacing the shared buses.Networks-on-Chip (NoCs) are proposed to be used in complex SoCs for inter-core communication because of scalability, better throughput and reduced power consumption [3].

On the other hand, increasing the number of cores over a 2D plane is not efficient due to long interconnects. With the emergence of viable 3D integration technologies opportunities exist for chip architecture innovations to enhance system power/performance characteristics [4]. In 3D integration technologies, multiple layers of active devices are stacked above each other and vertically interconnected using through-silicon vias (TSVs) [5]. As compared to 2D designs, 3D ICs allow for performance enhancements even in the absence of scaling because of the reduced interconnect lengths [6]. In addition to this clear benefit, package density is increased significantly, power is reduced due to shorter wires, and system is more immune tonoise [7].

One of the well-known 2D NoC architectures is the 2D Mesh. This architecture consists of an $m \times n$ mesh of switches interconnecting IP blocks placed along with them. The straightforward extension of this popular planar structure is 3D Symmetric NoC by simply adding two additional physical ports to each router; one for Up and one for Down [8]. Despite simplicity, this architecture has two major inherent drawbacks. Firstly, it does not exploit the beneficial attribute of a negligible inter-wafer distance in 3D chips, because in this architecture, inter-layer and intra-layer hops are indistinguishable. Secondly, a considerably larger crossbar is required as a result of two extra ports [9].

The Stacked (Hybrid NoC-Bus) mesh architecture presented in [10] is a hybrid architecture between the packet switched network and the bus architecture to overcome the mentioned 3D Symmetric NoC challenges. It takes advantage of the short inter-layer distances, around 20μm, that are characteristics of 3D ICs [4]. It integrates the multiple layers of 2D mesh networks by connecting them with a bus spanning the entire vertical distance of the chip. As the inter-layer distance for 3D ICs is small, the bus length will also be smaller; approximately $(n-1) * 20\mu m$, where n is the number of layers.

This makes the bus suitable for interlayer communication in vertical direction. By using the stacked mesh architecture, six-port router is required instead of seven ports for typical 3D NoC router and vertical communication is just one hop away to any destination layer.

In this paper, we address the routing issues and buffer utilization to enhance the overall system power and performance of existing stacked mesh . In addition, the proposed architecture increases tolerance against single bus failure architectures.

Related Work

Three (3D) dimensional integrated circuits evolved to deal with the limitations of interconnect scaling by stacking active silicon layers. The detailed description of the challenges faced to manufacture the 3D ICs is provided in [4]. The authors have shown that 3D ICs are power and performance efficient, but when the 3D NoC is taken under consideration, the statistics are quite different. The 3D NoCs are extension to the 2D NoC

architecture. For each NoC router of mesh topology, two extra ports are needed resulting a 7×7 crossbar instead of 5×5 crossbar for the 2D mesh architecture. Since, the crossbar power increases quadratically with number of ports, the power consumption for a 3D router is much higher than a 2D router [11].

The solution to the power consumption for a 3D router has been proposed by Li *et al.* [10]. The proposed architecture is stacked mesh architecture. Its basic purpose was cache coherence. The dynamic Time-Division Multiple Access (dTDMA) bus was used as a communication pillar. Due to one hop vertical communication and 6×6 routers, proposed architecture is efficient enough in terms of power consumption and latency. The issue with this architecture is that each packet is traversed through two buffers: the source output buffer and destination input buffer as shown in Fig. 1. As we will discuss later, the output buffer hinders implementing congestion-aware inter-layer communications.

We improve the architecture to further enhance the throughput by using the available communication resources. We complement the architecture due to two major aspects. First is to traverse packets through two stage crossbar, which makes it less power efficient. Second is to use two hops per packet for vertical communication, which degrades system throughput. Our proposed architecture deals with both issues. Single hop vertical communication without using two-stage crossbar improves the system power and performance. Adaptive vertical routing mechanism deals with load balancing. In [12], an enhanced decision making routing algorithm to avoid congestion in 2D NoC architectures was proposed. In addition, the proposed dynamic routing approach can tolerate a single link failure.

Utilizing the available communication resources, we enhance our approach to deal with fault tolerance for inter-layer communication of stacked mesh architecture. Symmetrical 3D NoC architecture is high-throughput but not power efficient. On the other hand, architectures like True NoC or stacked mesh architectures are power efficient at the expense of reduced throughput. We reinforce the stacked mesh architecture without adding extra communication resources and enhance system throughput, fault-tolerance, and power efficiency.
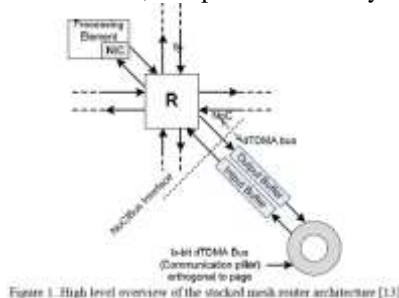

Figure 1. High level overview of the stacked mesh router architecture [13]

Motivation

In this architecture, routers connected to pillar nodes are different, as an interface between the dTDMA pillar (vertical link) and the NoC router must be provided to enable seamless integration of the vertical links with the 2D network within the layers. An extra physical channel is added to the router for the vertical communication. The extra channel has its own dedicated buffers, and is indistinguishable from the other channels.

In the static XYZ routing algorithm, consider that *RXYZ* is the source, which needs to send a data packet to the destination *RXYZ+1* as shown in Fig. 2. This particular architecture is neither power nor performance efficient because of inefficient

bus utilization. This can be justified by considering different communication scenarios as follows.
1. Bus is busy but input buffer of RXYZ+1 is free:
2. Bus is free but input buffer of RXYZ+1 is full:
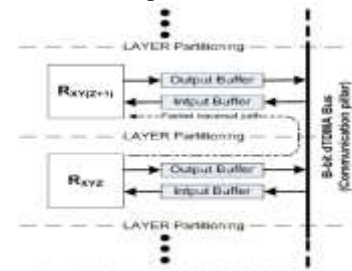3. Bus is busy and also the input buffer of RXYZ+1 is full:

Figure 2. Side view of the 3D NoC with the dTDMA bus

Proposed Architecture

It is not efficient to connect a NoC router to a vertical bus without considering their characteristics. We showed that the extra output buffer even can obstruct routing adaptivity and load balancing. In this section, we approach the problem in different phases. Initially, the buffer on output port is removed to reduce power consumption and to enable implementing an adaptive routing algorithm. Then, an adaptive inter-layer routing algorithm called Adaptive Z is applied.

For the sake of simplicity, it is assumed that for intra-layer communication a static XY routing algorithm is used. Note that the proposed inter-layer routing algorithm is not dependent to intra-layer routing policy. Therefore, in ZXY routing algorithm, the XY routing is projected in different layers identically as shown in Fig. 3 with the adaptive inter-layer communication. The Adaptive Z routing algorithm is elaborated in Algorithm 1.

As can be seen from Fig. 3, for inter-layer communication, the first bus pillar available on the way for the vertical communication is used.
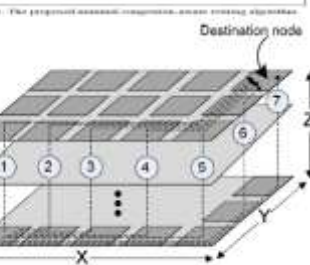

Figure 3. Example of ZXY routing

Deadlock Avoidance

To deal with deadlock, typical virtual channel architecture is used as shown in Fig. 4. Consider that the input buffer of router R001 contains a packet, which needs to be transmitted to the neighboring node R101. Router R101 is already waiting for the availability of dTDMA bus to deliver a packet at the node R100. Also, consider that the node R000 and R100 have the same situations. In case of typical stacked mesh architecture with adaptive vertical routing and without virtual channel, there will be a deadlock. In the proposed architecture, with the same number of buffers as compared to stacked mesh with dTDMA

bus architecture [10], we modify the inter-layer communication scheme to support the VC architecture.

The output buffer from the router for bus communication was removed. An extra input buffer is used to receive the data packets and support the VC concept. There is very small area overhead of one 2×1 multiplexer, one 1×2 de-multiplexer and few signaling wires. Reduction in dynamic power consumption due to removal of the buffer from the traversal path of packet and also static power optimization due to routing adaptivity dominates the power consumption of this small extra logic. In addition, the added VC not only avoids deadlock, but it also improves the throughput. Now, the buffer is used in a right way and makes the system power and performance efficient.
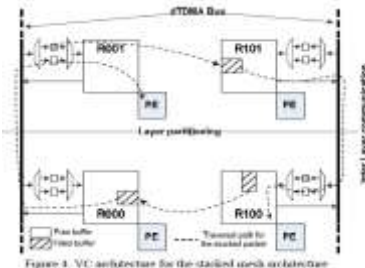

Figure 4. VC architecture for the stacked mesh architecture.

Fault Tolerance

Through Silicon Via, which provides communication links for dies in the vertical direction, is a critical design issue in 3D integration. Like other physical components, the fabrication and bonding of TSVs can fail. A failed TSV may cause a number of stacked known-gooddies to be discarded. As the number of dies to be stacked increases, the failed TSVs increase the cost and decrease the yield [13]. A reliable inter-layer communication scheme can considerably mitigate these issues. In this section, we explore how the available signals can enable the routing algorithm to avoid these paths (faulty buses) when there are other paths between the source and destination pair.

For fault tolerance, the existing signaling used foradaptive inter-layer communication is used without adding any extra wire. As explained in section IV, '*wait*' signal acts as a congestion flag in normal conditions. In case of any fault on the dTDMA bus, '*wait*' signal is permanently asserted '0' and bus arbiter does not serve any request raised by '*req*' signal. Thus, routing adaptivity is not affected by introducing the fault tolerance in existing architecture.

It is assumed that the dTDMA bus is equipped with enough resources (test units) to signal that either the bus link is broken or the bus is not able to provide services due to thermal problem or any other fault. When the bus arbiter receives any of such signals, which indicate that there is a fault on the bus link and bus is not able to provide the communication temporarily or permanently, it asserts '0' on the '*wait*' signal accordingly. When router checks the '*wait*' signal, it will reroute the packets within the layer and in future, will not derive any traffic for inter-layer communication to the corresponding dTDMA bus. Normally, the minimal path routing will be used even in case of faulty links without any modification in routing algorithm discussed in section IV. But there are few cases, when routing mechanism needs to be modified.

Consider that the node '000' in Fig. 6 needs to send a packet to node '112'. The vertical links on two paths are faulty as sown in Fig. 6. According to the proposed routing algorithm in section IV, the vertical dTDMA bus connecting the nodes '000', '001' and '002' should be tried but that bus link is faulty in current situation. So, the packet will be routed to the either of

nodes '010' or '100' according to the routing algorithm. Then the packet will be routed normally according to the proposed algorithm.

Now consider another situation, if the packet is exactly below/above its destination node and by inter-layer communication, it can be delivered to the destination but the bus link is broken, the packet will be rerouted to the neighboring node within the layer, where the packet is currently residing. Then the packet will be delivered to the destination. This is the situation, when the node '120' contains a packet for node '122'. For proposed routing algorithm, the packet can only be routed through the vertical dTDMA bus connecting the nodes '120', '121' and '122', which is faulty. In this situation, the routing algorithm requires modification. The packet will be routed to the nodes '110' or '020' following the non-minimal path.
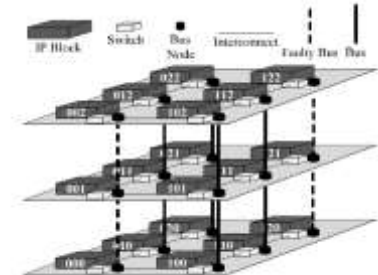

Figure 6. Example of modifications to the proposed routing algorithm to guarantee single bus-link failure tolerance

Conclusion

In this paper, an enhanced architecture for 3D stacked mesh NoC was proposed to enhance system performance, reduce power consumption, and improve the system reliability. To this end, a congestion aware adaptive interlayer communication algorithm was introduced. To deal with deadlock, an appropriate VC architecture was used with same number of buffers as compared to the existing stacked mesh architectures. In addition, the congestion signal triggered by the bus arbiter was used to deal with fault tolerance. This enabled avoiding a faulty vertical bus in a possible path for a number of source-destination pairs in the event of a bus failure.

References

[1] S. Borkar, "Designing reliable systems from unreliable components: The challenges of transistor variability and degradation," IEEE Micro, Vol. 25, No. 6, 2005, pp. 10-16.

[2] A. Jerraya and Wayne Wolf, "Multiprocessor Systems-on-Chips," Morgan Kaufmann, 1st edition, October 12, 2004.

[3] A. Jantsch, and H. Tenhunen, Network on Chip, Kluwer Academic Publishers, 2003.

[4] A. W. Topol et al., "Three-Dimensional Integrated Circuits," IBM J. Research and Development, Vol. 50, No. 4/5, July-Sept. 2006, pp. 491-506.

[5] V. F. Pavlidis and E. G. Friedman, "3-D Topologies for Networks-on- Chip, " in Proc. of IEEE Trans. Of VLSI Systems., Vol. 15, No. 10, 2007, pp.1081-1090.

[6] B. S. Feero and P. P. Pande, "Networks-on-Chip in a Three-Dimensional Environment: A Performance Evaluation," IEEE Transactions on Computers, Vol. 58, No. 1, 2009, pp. 32-45.

[7] R. S. Patti, "Three-Dimensional Integrated Circuits and the Future of System-on-Chip Designs", in Proc of IEEE, Vol. 94, No. 6, 2006, pp. 1214-1224

[8] L. P. Carloni et al., "Networks-on-Chip in Emerging Interconnect Paradigms: Advantages and Challenges," in Proc. of International Symposium on Networks-on-Chip, 2009, pp. 93-109.

[9] J. Kim et al., "A novel dimensionaly-decomposed router for on-chip communication in 3D architectures," in Proc. of International Symposium on Computer Architectures, 2007. pp. 138-149.

[10] F. Li et al., "Design and Management of 3D Chip Multiprocessors Using Network-in-Memory," in Proc. of International Symposium on Computer Architecture, 2006, pp. 130-141.

[11] R. S. Ramanujam and B. Lin, "A Layer-Multiplexed 3D On-Chip Network Architecture," IEEE Embedded Systems Letters, Vol. 1, No. 2, 2009, pp. 50-55.

[12] P. Lotfi-Kamran et al., "EDXY – A low cost congestion-aware routing algorithm for network-on-chips," Journal of Systems Architecture, Vol. 56, No. 7, 2010, pp. 256-264.

[13] A.-C. Hsieh et al., "TSV Redundancy: Architecture and Design Issues in 3D IC," in Proc. of International Conference on Design, Automation, and Test in Eroupe, 2010, pp. 166-171.