ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Performance Analysis of an Efficient NOC Router System Using Data Encoding Techniques

Y. Geetha
PG Scholar,
Electronics and Communication Engineering,
P.V.K.K Engineering College, AP, India

K. Ravikumar
Associate Professor,
Electronics and Communication Engineering,
P.V.K.K Engineering College, AP, India

ABSTRACT

As technology shrinks, the power dissipated by the links of a network-on-chip (NoC) starts to compete with thepower dissipated by the other elements of the communicationsubsystem, namely, the routers and the network interfaces (NIs).In this paper, we present a set of data encoding schemes aimedat reducing the power dissipated by the links of an NoC. Theproposed schemes are general and transparent with respect to theunderlying NoC fabric (i.e., their application does not require anymodification of the routers and link architecture). Experimentscarried out on both synthetic and real traffic scenarios show theeffectiveness of the proposed schemes, which allow to save upto 51% of power dissipation and 14% of energy consumptionwithout any significant performance

INTRODUCTION:

Shifting from a silicon technology node to the next one results in faster and more power efficient gates but slower and more power hungry wires. In fact, more than 50% of the total dynamic power is dissipated in interconnects in current processors, and this is expected to rise to 65%–80% over the next several years. Global interconnect length does not scale with smaller transistors and local wires. Chip size remains relatively constant because the chip function continues to increase and RC delay increases exponentially. At 32/28 nm, for instance, the RC delay in a 1-mm global wire at the minimum pitch is 25× higher than the intrinsic delay of a two-input NAND fan out. If the raw computation horsepower seems to be unlimited, thanks to the ability of instancing more and more cores in a single silicon die, scalability issues, due to the need of making efficient and reliable communication between the increasing number of cores, become the real problem . The network on- chip (NoC) design paradigm is recognized as the most viable way to tackle with scalability and variability issues that characterize the ultra deep submicron meter era. Nowadays, the on-chip communication issues are as relevant as, and in some cases more relevant than, the computation related issues. In fact, the communication subsystem increasingly impacts the traditional design objectives, including cost (i.e., silicon area), performance, power dissipation, energy consumption, reliability, etc. As technology shrinks, an ever more significant fraction of the total power budget of a complex many-core system-on-chip (SoC) is due to the communication subsystem.

An ever more significant fraction of the overall power dissipation of a network-on-chip (NoC) based system-on-chip (SoC) is due to the interconnection system. In fact, as technology shrinks, the power contributes of NoC links starts to compete with that of NoC routers. In this paper, we propose the use of data encoding techniques as a viable way to reduce both power dissipation and energy consumption of NoC links. The proposed encoding scheme exploits the wormhole switching techniques and works on an end-to-end basis. That is, flits are encoded by the network interface (NI) before they are injected in the network and are decoded by the destination NI. This makes the scheme transparent to the underlying network since the encoder and decoder logic is integrated in the NI and no modification of the routers architecture is required. We assess the proposed encoding scheme on a set of representative data streams (both synthetic and extracted from real applications) showing that it is

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

possible to reduce the power contribution of both the self-switching activity and the coupling switching activity in inter-routers links. As results, we obtain a reduction in total power dissipation and energy consumption up to 37% and 18%, respectively, without any significant degradation in terms of both performance and silicon area

NOC ROUTER

2.1 Introduction

A billion transistors, one million gates, thousands of circuits, hundreds of designs on a single IC chip; such intricate designs pose innumerable challenges to IC designers. The most successful IC designers overcome all such challenges to provide functionally correct and reliable operation of the IC's. As the integration increases the cost effectiveness is also a major area of concern in IC designs.

Reduced cost is one of the big attractions of integrated electronics, the cost advantage continues to increase with the evolution of technology toward the production of larger and larger circuit functions on a single semiconductor substrate; proposed by Gordon E Moore in his paper Cramming more components onto integrated circuits, 1965, the paper in which he proposed the well-known Moore's law. By 2025 the physical dimension of CMOS transistors are expected to cross the 10 nm threshold, according to a 2012 report of International Technology Roadmap for Semiconductors. The graph below shows the transistor integration on a single chip over the past two decades. Now two billion transistors are integr onto a single chip.

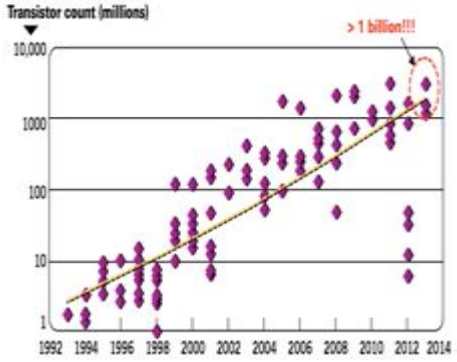

Figure 2.1. Evolution of transistor integration on a chip3

It is to keep pace with such intricate levels of integration that the design engineers have come up with a new design methodology called System-on-Chip (SoC) .

The SoC is a technology where maximum technology is crammed into the smallest possible space. The design of system on a chip is impacted strongly by the so called intellectual property (IP) core. An integrated circuit core is a predesigned, preverified silicon circuit block. The core usually contains at least 5,000 gates that can be used in building a larger or more complex application on a semiconductor chip. Examples of cores are memory controllers, processors, or peripheral devices such as MAC Ethernet or PCI bus controllers. In the semiconductor industry IP core is the property of any single vendor. The ip cores are the building blocks of various systems on chip designs for implementing larger and complex embedded system applications. The system on chip can hold hardware's like processors, memories peripherals, controllers, digital signal processors and various custom logic blocks and software's for controlling the hard wares.

The main advantage of system on chip is low power consumption, lower cost and higher reliability than the multi-chip systems it has replaced. But the transition to system on chip technology was faced with many challenges. Firstly, scalability of the system. It is really an enormous task to scale down large computer systems to the size of silicon die. The physical dimensions of various components, their inductive and capacitive effects on other components need to be taken care of. Secondly, it is difficult to maintain global synchronization as different systems will be using different clock signals. include an interconnection architecture and interfaces to peripheral devices. The interconnection architecture consists of physical interfaces and communication mechanisms. This allows the communication between SoC components to take place.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

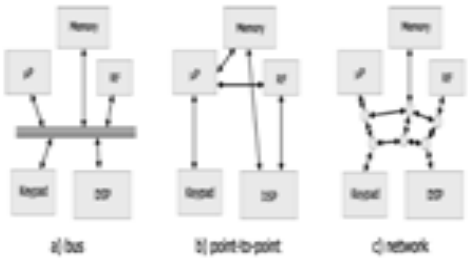

Fig. 2.2- Communication structures in SoC a) traditional bus based communication,

b) Dedicated point to point links c) network on a chip

Usually, the interconnection architecture is based on dedicated wires or shared busses. If a system has a limited number of cores then dedicated wire architecture is effective. As the system complexity grows the number of wires around the core also increases. Therefore, dedicated wires have poor reusability and flexibility. A shared bus is a set of wires which is common to multiple cores. The approach of shared bus is more flexible and is totally reusable, but it allows only one communication transaction at a time, all cores share the same communication bandwidth in the system and its scalability is limited to few dozen IP cores. Thus scalability is a major problem with buses. It is the issues in interconnection that paved the way for new paradigm in communication called the Network-on-chip (NoC). NoC architecture has been proposed as a high performance, scalable and power efficient alternative to the bus based architecture. It solves the scalability problem by supporting multiple concurrent connections with various systems. As system becomes more complex, more and more integration is possible to the 5 existing system with ease without any constraints. It can reduce the wire routing congestion to a great extent. The systems that are interconnected with a network on chip can be easily interchanged with other systems with any ip cores of any vendor available in the market. The NoC separates the communication part from the computation part for system simplicity and is ideally suited for integrated systems. NoC can take care of the communication part with utmost ease without any interference in the computation part. After the configuration and interconnection the other major area of concern in the SoC design is the implementation. Nowadays the implementation is done using Field Programmable Gate Arrays (FPGAs). The advantages of FPGAs are the lower time to market, lower development cost, less manufacturing steps and highly suitable for research activities.

2.2 Network on chip (NoC) :

NoC is a technology that is intended to solve the short coming of buses. It is an approach to design the communication subsystem between intellectual property cores in a SoC design. The communication strategy in system on chip uses dedicated buses between communicating resources. This will not give any flexibility for the needs of the communication, in each case, have to be thought of every time a design is made. Another possibility is the use of common buses, which have the problem that it does not scale very well, as the number of resources grows. NoC is intended to solve the shortcomings of these, by implementing a communication network of switches/micro routers and resources. The NoC design paradigm has been proposed as the future of ASIC design .The major driving force behind the transition to NoC based solutions is the inadequacy of current day VLSI inter-chip communication design methodology for the deep sub-micron chip manufacturing technology. The negative effect of technology scaling on global interconnects, increased dependence on fault-tolerant mechanisms as feature 6 size reduces, increasing use of parallel architectures are the reasons why NoC is becoming popular. The NoC based system on chips imposes various design issues on the fabrication of such integrated chips. Future network on chips will become more sensitive and prone to errors and faults. Fault tolerance is becoming critical for on chip communications. Today's SoCs need a network on chip IP interconnect fabric to reduce wire routing congestion, to ease timing closure, for higher operating frequencies and to change IP easily. Network on chips are a critical technology that will enable the success of future system on chips for embedded applications. This technology of

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

network on chip is expected to dominate computing platforms in the near future.

NoC design space is larger when compared to a bus based solution, as different routing and arbitration strategies can be implemented as different organizations of the communication infrastructure. The NoC paradigm is highly suited to provide SoC platforms scalable and adaptable over several technology generations. NoC platforms may allow the design productivity to grow as fast as technology capabilities and may eventually close the

2.3 NoC Architecture:

Fig 2.3. General Architecture ofNoC

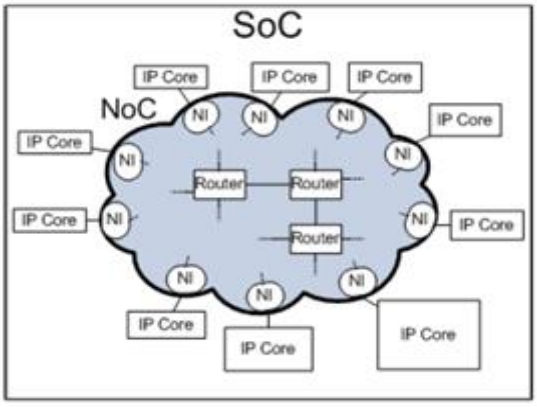

figure 2.3 shows the general architecture of NoC.

Routers, network interface (NI) and links are the maincomponents of NoC architecture. A router is responsible for routing information from a source port to its destination port. The network interface separates the computation part from the communication part and acts as a mediator between the router and the processing element. A link connects different routers in the network according to the chosen topology.

2.4 Link:

Links are used to transmit packet between routers. It physically connects the nodes and enables the communication in the network. It consists of a set of wires that connect the routers in the network. A NoC link has two physical channels making a full-duplex connection between the routers. The number of wires per channel is uniform throughout the network and it is known as channel bit width. But if the links become too long then it can cause wiring delay in the network. To overcome such problem the NoC pipelines long wires in interconnects by partitioning the wires into smaller segments.

2.5 Network interface

The network interface or network adapter makes the logical connection between the IP core and the network. It can be divided into two parts: a front end and back end. The requests from the IP core are handled by the front portion and it is unaware of the existence of the network. The back end part is connected directly to the network which handles the network protocol, ordering and reordering the packets, buffers and helps the router in terms of storage.

2.6 Router:

A router is the most important component in a NoC. It is the communication backbone of a NoC system. So it should be designed for maximum efficiency and throughput. A router is used in a network for directing the traffic from source to destination. It coordinates the data flow which is very crucial in communication networks. The architecture of a router consists of an input port, an output port, a switching matrix to connect the input port to the output port, and a local port that connects the router to the corresponding IP core. Routers are intelligent devices that receive incoming data packets, inspect their destination and figure out the best path for the data to move from source to destination.

A router's architecture determines its critical path delay which affects delay and network latency. So the design of the router should be such that it meets the required latency and throughput requirements among tight area and power constraints. The design efficiency of the router determines the performance of the network. A router decodes the information provided by the incoming message based on the routing function and destination of the message. A router is built according to the OSI model of NoC. Each layer has its own specific functions to perform.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

2.7 Network topology:

The physical layout and connections between nodes and channels are in the network is determined by the on chip network topology. A topology determines the number of hops (or routers) a message must traverse as well as the interconnect lengths, thus 10 influencing network latency significantly. As traversing routers and links incurs energy, a topology's effect on hop count also directly affects network energy consumption. The topology dictates the total number of alternate paths between nodes, affecting how well the network can spread out traffic and hence support bandwidth requirements. The first decision designers have to make when building of an on-chip network is, frequently, the choice of the topology.

2.7.1 Mesh:

Mesh topology is favored by many research groups because of its layout efficiency. It has good electrical property and can address the on-chip resources in a simple manner. A mesh-shaped network consists of m columns and n rows. The routers are situated in the intersections of the two wires and the computational resources are near routers. Addresses of routers and resources can be easily defined as x-y coordinates in mesh. The regular mesh network is also called as Manhattan Street network. A mesh-shaped network consists of m columns and n rows. The routers are situated in the intersections of the two wires and the computational resources are near routers. Addresses of routers and resources can be easily defined as x-y coordinates in mesh.

2.7.2 Tree:

In a tree topology nodes are routers and leaves are computational resources. The routers above a leaf are called as leaf's ancestors and correspondingly the leaves below the ancestor are its children. In a fat tree topology each node has replicated ancestors which mean that there are many alternative routes between nodes.

2.7.3 Star:

A star network consists of a central router in the middle of the star, and computational resources or sub networks in the spikes of the star. The capacity requirements of the central router are quite large, because all the traffic between the spikes goes through the central router. That causes a remarkable possibility of congestion in the middle of the star.

2.8 Switching:

A network consists of many switching devices. In order to connect multiple devices, one solution could be to have a point to point connection in between pair of devices. But this increases the number of connection. The other solution could be to have a central device and connect every device to each other the central device which is generally known as Star Topology. Both these methods are wasteful and impractical for very large network. The other topology also cannot be used at this stage. Hence a better solution for this situation is SWITCHING. A switched network is made up of a series of interconnected nodes called switches.

2.8.1 Types of Switching Techniques:

There are basically three types of switching methods are made available. Out of three methods, circuit switching and packet switching are commonly used but the message switching have been opposed out in the general communication but is still used in the networking application.

2.8.2 Virtual circuit Switching:

A virtual circuit (VC) is transferring data over a packet switched computer network in such a way that it appears as though there is a dedicated physical layer link between the source and destination. Before a connection or virtual circuit may be used, it has to be established, between two or more nodes or software applications, by configuring the relevant parts of the interconnecting network. After which, a bit stream or byte stream may be delivered between the nodes; the term virtual circuit is also synonymous with virtual connection and virtual channel. Hence, a virtual circuit protocol allows higher level protocols to avoid dealing with the division of data into segments, packets, or frames. Virtual circuit communication resembles circuit switching, since both are connection oriented,

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

meaning that in both cases data is delivered in correct order, and signaling overhead is required during a connection establishment phase. However, circuit switching provides a constant bit rate and latency, while these may vary in a virtual circuit service due to factors such as:

varying packet queue lengths in the network nodes,varying bit rate generated by the application,Varying load from other users sharing the same network resources by means of statistical multiplexing, etc.Many virtual circuit protocols, but not all, provide reliable communication service through the use of data retransmissions because of error detection and automatic repeat request (ARQ).

2.8.3Advantages:

The advantages to Message Switching are:Data channels are shared among communication devices improving the use of bandwidth.

- Messages can be stored temporarily at message switches, when network congestion becomes a problem.
- Priorities may be used to manage network traffic.
- Broadcast addressing uses bandwidth more efficiently because messages are delivered to multiple destinations.

REFERENCES

[1] *International Technology Roadmap for Semiconductors*. (2011) [Online]. Available: http://www.itrs.net

[2] M. S. Rahaman and M. H. Chowdhury, "Crosstalk avoidance and errorcorrection coding for coupled RLC interconnects," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2009, pp. 141–144.

[3] W. Wolf, A. A. Jerraya, and G. Martin, "Multiprocessor system-on-chip MPSoC technology," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 10, pp. 1701–1713, Oct. 2008.

[4] L. Benini and G. De Micheli, "Networks on chips: A new SoC paradigm," *Computer*, vol. 35, no. 1, pp. 70–78, Jan. 2002.

[5] S. E. Lee and N. Bagherzadeh, "A variable frequency link for a poweraware network-on-chip (NoC)," *Integr. VLSI J.*, vol. 42, no. 4, pp. 479–485, Sep. 2009.

[6] D. Yeh, L. S. Peh, S. Borkar, J. Darringer, A. Agarwal, andW. M. Hwu, "Thousand-core chips roundtable," *IEEE Design Test Comput.*, vol. 25, no. 3, pp. 272–278, May–Jun. 2008.

[7] A. Vittal and M. Marek-Sadowska, "Crosstalk reduction for VLSI," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 16, no. 3, pp. 290–298, Mar. 1997.

[8] M. Ghoneima, Y. I. Ismail, M. M. Khellah, J. W. Tschanz, and V. De, "Formal derivation of optimal active shielding for low-power on-chip buses," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 5, pp. 821–836, May 2006.

[9] L. Macchiarulo, E. Macii, and M. Poncino, "Wire placement for crosstalk energy minimization in address buses," in *Proc. Design Autom. Test Eur. Conf. Exhibit.*, Mar. 2002, pp. 158–162.

[10] R. Ayoub and A. Orailoglu, "A unified transformational approach for reductions in fault vulnerability, power, and crosstalk noise and delay on processor buses," in *Proc. Design Autom. Conf. Asia South Pacific*, vol. 2. Jan. 2005, pp. 729–734.

[11] K. Banerjee and A. Mehrotra, "A power-optimal repeater insertion methodology for global interconnects in nanometer designs," *IEEE Trans. Electron Devices*, vol. 49, no. 11, pp. 2001–2007, Nov. 2002.

[12] M. R. Stan and W. P. Burleson, "Bus-invert coding for low-power I/O," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 3, no. 1, pp. 49–58, Mar. 1995.

[13] S. Ramprasad, N. R. Shanbhag, and I. N. Hajj, "A coding framework for low-power address and data busses," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 7, no. 2, pp. 212–221, Jun. 1999.

[14] C. L. Su, C. Y. Tsui, and A. M. Despain, "Saving power in the control path of embedded processors," *IEEE Design Test Comput.*, vol. 11, no. 4, pp. 24–31, Oct.–Dec. 1994.

[15] L. Benini, G. De Micheli, E. Macii, D. Sciuto, and C. Silvano, "Asymptotic zero-transition activity encoding for address busses in low-power microprocessor-based systems," in *Proc. 7th Great Lakes Symp. VLSI*, Mar. 1997, pp. 77–82.

[16] E. Musoll, T. Lang, and J. Cortadella, "Working-zone encoding for reducing the energy in microprocessor address buses," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 6, no. 4, pp. 568–572, Dec. 1998.

[17] W. Fornaciari, M. Polentarutti, D. Sciuto, and C. Silvano, "Power optimization of system-level address buses based on software profiling," in *Proc. 8th Int. Workshop Hardw. Softw. Codesign*, May 2000, pp. 29–33.

[18] L. Benini, G. De Micheli, E. Macii, M. Poncino, and S. Quer, "Power optimization of core-based systems by address bus encoding," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 6, no. 4, pp. 554–562, Dec. 1998.

[19] L. Benini, A. Macii, M. Poncino, and R. Scarsi, "Architectures and synthesis algorithms for power-efficient bus interfaces," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 19, no. 9, pp. 969–980, Sep. 2000.

[20] G. Ascia, V. Catania, M. Palesi, and A. Parlato, "Switching activity reduction in embedded systems: A genetic bus encoding approach