Network on Chip - Design Aspects

Rakhi Mutha
Dept.of IT
JIET Group of Institutions
Jodhpur, Rajasthan, India

ABSTRACT
The Network-on-Chip is a packet switched platform for single chip systems which scales well to an arbitrary number of processor like resources. The network-on-chip (NoC) design paradigm is seen as a way of enabling the integration of an exceedingly high number of computational and storage blocks in a single chip. This provides vertical integration of physical and architectural levels in system design. A chip consists of contiguous areas called regions, which are physically isolated from each other but have special mechanism for communication among each other. The NOC architecture essentially is the on chip communication infrastructure comprising the physical layer, the data link layer and the network layer of the OSI protocol stack. NOC architecture, a general purpose on-chip interconnection network replaces the traditional design- specific global on-chip wiring, by the use of switching fabric or routers to connect IP cores or processing elements (PEs).

General Terms
Network on Chip, System on chip, Packet switching network.

Keywords
Packet switching, NOC, On chip communication, NOC design..

1. INTRODUCTION
On chip communications are improved using method of bit packet switching by employing either circuit switching or packet switching techniques. This technique is of most interest when taking into account the increased utilization of multi-core processor architectures, and the need to economically pipe information between cores. But besides the growing use of parallel architectures, this technology is also important for effective management of increased system architecture complexity, as well as efficient use of on-chip wiring. Furthermore, while not all approaches to modern day computer networking are applicable to NoC design.

The NOC design is typical to make dedicated architectures for applications. The possible solutions must be searched from platform based design and computer system design, which rely on the reuse of components, architectures, applications and implementations. The essential issue is the trade-off between generality and performance. Generality provides reusability of hardware, operating systems and development practices, while performance is achieved by using application specific structures. NOC platform is consisting of architecture and design methodology, which expandable from a numbers of more sources which can be a processor core, a dedicated HW block, a mixed signal block, or a memory block. There are various key points affecting the thought like Moore's law discussion [21], drawback of single processors which is not able to utilize the transistors of an entire chip. [2, 3], and A large number of communicating different tasks may have different characteristics

So the discussed aspect can make a collection of various implementations with different type of resources for different reasons and the most cost effective solution. On the basis of these points it can be conclude that a large number of different kinds of blocks, each of the size of a few hundred thousand gates, will constitute the computational resources.

The Network on chip is an effective platform to separate the specification of inter-task communication and the implementation. To make a high-performance, cost-effective products for improving design productivity are the consequent separation of different concerns.

To achieve it the device for processing should be work on multiple high data rate streams of data and the devices should be capable for multiple functionality, and devices should have high-capacity wire line.

NOC based design will not always be the preferred solution for all kinds of applications. It can be expect that NOC based designs will provide good solutions for flexible products that should be reconfigurable and programmable.

The design methodology must therefore support product family management. Tolerance of incomplete specifications, management of configurations and modifications, support for multiple languages and methods, and capability to handle different abstraction levels simultaneously are desirable characteristics.

Verification and testing are ever increasing challenges in today's design routines. The NOC platform effectively addresses these challenges by separating the computation resources from each other and from the communication network for all issues of design, verification and testing.

2. RELATED WORK
There are very big community of researchers are involved in this area of designing better NOC system, This system is very much useful to reutilize of predesigned components. The researcher like Gajski [5] who have prepare a system follows IP-centric design methodology and Keutzer [7] who have followed the principles of platform based design and expand new areas by including a layer of software on top of the hardware platform to help application development. The major challenges in the IP centric methodologies are the interface synthesis among various IP blocks and system verification. The idea of this researcher utilizes design for development of application not only speeds up application design but also reduces its verification time. This layer is called Software Platform. The combination of hardware and software platforms is referred as System Platform.

It has also been observed that the key to reuse and integration of IP components is the communication from the physical to the system and conceptual level, and consequently communication

centric architectures, platforms and methodologies have been developed [8, 9, 10]. The future system on a chip, incorporating many different types of processing and memory elements, has to operate using Globally Asynchronous Locally Synchronous paradigm [16], at least at the hardware level. This paradigm not only avoids the problem of clock skew but also leads to lower power consumption.

3. NOC ARCHITECTURE AND NETWORK

Architecture of network on chip provides the infrastructure for the communication in resources. The main objectives are to make possible to develop the independent hardware of resources and connect these resources to create the NOC, and make this network scalable and flexible to make this platform acceptable to different workloads. At this place the main topology is mesh, which is a simplest layout and local connectivity, also routing in a two-dimensional mesh is easy so it can make small switches, which have high capacity, short clock cycle, and overall scalability.

A NOC consists of resources and switches that are connected using channels as a mesh so that they are able to communicate with each other by sending messages. As shown in figure: 3 resources R is a computation or storage unit or their combination and switch S routes and buffers messages between resources. Each switch is connected to four other neighboring switches through input and output channels and channel C consists of two one-directional point-to-point buses between two switches or a resource and a switch. Switches may have internal queues to handle congestion.

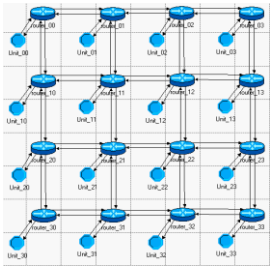

Figure: 1 A NOC with 16 resources.

The NOC design depends on the frequently changing technology and the area of a resource, which is the maximal synchronous region in a given technology. It varies with technology. The number of resources will grow, the switch-to-switch and the switch to resource bandwidth will grow, but the network wide communication protocols will be unaffected.

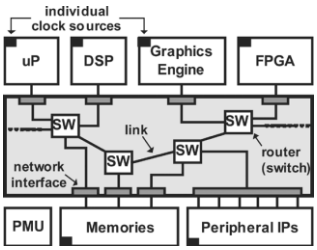

Figure: 2 -Heterogeneous NoC architecture.

4. NETWORK DESIGN

The network design task consists of implementations for upper layers of the protocol stack, namely presentation, session, transport, and network layers. The presentation layer is responsible for data formatting. It converts abstract data types in the application to blocks of ordered bytes as defined by the canonical byte layout requirements of the lower layers. The session layer implements end-to-end synchronization to provide synchronous communication as required between system components in the application. Furthermore, it is responsible for multiplexing messages of different channels into a number of end-to-end sequential message streams. The transport layer splits messages into smaller packets (e.g. to reduce required intermediate buffer sizes) and implements end-to-end flow control and error correction to guarantee reliable transmission.

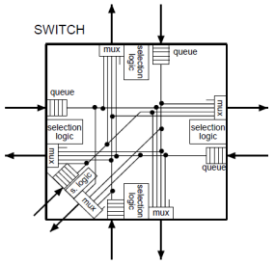

Figure: 3 Block diagram of a switch.

5. NOC RESOURCES

The NOC is a collection of different types of resources with personal caches and local memories, dedicated hardware resources, and configurable hardware resources. The resource can be a combination various resource with one synchronous clock. The internal communication inside a resource is synchronous. The model of computation is a heterogeneous network of resources executing local computation.

Communication between the resources is implemented by passing messages over the mesh network. Resources operate asynchronously with respect to each other. Synchronization is provided by synchronization primitives, which are implemented by passing messages around the network. Even a non-local memory is accessed through message passing.

In Figure:4 RNI=resource network interface, P=processor core, D=DSP core, c=cache, M=memory and re=reconfigurable block.

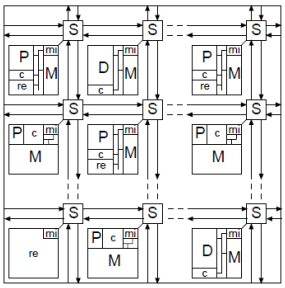

Figure: 4 Various types of resources in NOC.

The I/O could be of various kinds; they could glue many NOC chips together, interface with external memory or implement a

TCP/IP interface. Interface modules also handle data buffering and packet reordering.

6. COMMUNICATION

In NOC resource has particular specific address by which it can connect to a network via a switch. It communicates with the switch through a RNI. So that, any resource can be embedded in the network if its identification suitable to an available slot and if it is equipped with an RNI. The NOC defines four protocol layers, physical layer determines the number and length of wires connecting resources and switches, data-link layer defines the protocol to transmit a cell between a resource and a switch and between two switches, The network layer defines how a packet is transmitted over the network from an arbitrary sender to an arbitrary receiver directed by the receiver's network address and transport layer is technology independent. The transport layer message size can be variable. The RNI interface has to pack transport layer messages into network layer packets.

The basic communication mechanism envisioned among computing resources is message passing. However, it is possible to add additional protocols on top of the transport layer to provide for instance a virtual shared memory abstraction, which will help the programmers in development of data and computation intensive application.

7. REGIONS & WRAPERS

A 2-D mesh topology provides access to all resources of the NOC, it is scalable and it has a simple structure. A region G is an area inside the NOC, which is insulated from the network and which may have different internal topology and communication mechanisms.

The concept of region allows for resources of larger size than the atomic slots in the mesh. Regions are connected to the NOC by special communication arrangements called wrappers W, which route packets so that regions are insulated from external traffic. Specific IO wrappers allow communication between the region and its environment. The shape of regions cannot be arbitrary but their boundaries must be convex.

8. BACK BONE DESIGN

The NOC backbone design consist of various issues related to the topological and communication aspects issues such as channels, switches, and network interfaces. The backbone is the development platform for all NOC based systems, so it is important that every system follows the basic operation principles defined in the backbone. During the backbone design the focus is the network communication resources, e.g. switches and interfaces, and NOC system services and performance of different region topologies.

From the definition of resource area follows that the connections between neighboring switches and the switch design are issues where physical design has an important role.

The system-level communication challenges the technological limits. The amount of wires, wire lengths, synchronization, and buffering are all problems were physical layout and characteristics sets constraints. Customized region topology enables NOC based systems were the quality of the application mapping is optimized in the beginning.

Definition of region requires that potential applications are analyzed and modeled. Mathematical and performance analyses and even performance simulations are the main tools to be used.

9. COMMUNICATION DESIGN FLOW

The refinement-based communication design flow which is divided into two tasks: network design and link design. During the network design, the topology of communication architecture is defined and abstract message passing channels between system components are mapped into communication between adjacent communication stations (e.g. processing elements and communication elements) of the system architecture.

The network topology of communication stations connected by logical link channels is defined, bridges and other communication elements are allocated as necessary and abstract message passing channels are routed over sets of logical link channels. The result of the network design step is a refined link model of the system. The link model represents the topology of communication architecture where components and additional communication stations communicate via logical link channels.

10. SYSTEM DESIGN

In the application mapping the functionality of application is mapped to the resources. The NOC concept should ultimately support both dynamic and static mapping of applications [23], but the main problems with both are the resource allocation, optimization of network usage and verification of performance and correctness.

Static approaches are suitable only for specific platforms, not allowing the insertion of new applications into the system at run-time. As these approaches are performed during the design time, the algorithms may use a more thorough amount of information about the system to take decisions.

In opposition to static mapping, in dynamic scenarios the time taken to map each task is relevant, since it influences the overall application execution time. To reduce mapping overhead, greedy algorithms may be used, since these trade search space exploration quality by fast results.

Basically these issues are rather similar to what distributed and parallel system designers have to face.

The NOC platform is very heterogeneous. The resources can vary from configurable hardware to multiprocessor computers of almost every type. Therefore, several modeling languages should be supported by NOC application development environment making it easy to integrate different tools into the design flow. As with platform design, the decision support and quality validation needs special attention and new approaches.

11. CONCLUSION

The paper is a study on NOC architecture. NOC are useful to create a complex system on single chip. It supports integration of resources physically and logically. Communication between resources implements by packet switching and message passing through switches. NOC architecture works on four protocols for communication layers. Which are adapted from OSI standard. These protocols must be implemented in the resource to network interface for every resource in NOC. Before NOC architectural template can be used to develop applications, one need to work

out the details of architecture, communication, design flow, and system services.

12. REFERENCES

[1] Rakhi Mutha "Packet Switched Networks for Communications within Large Multi-core Systems on Chip " Proc. of the IEEE Int. Conference on simulation and modelling , Jan. 7-9 2011, Mumbai, India VI-97-101.

[2] D. Sylvester and K. Keutzer, "Getting to the Bottom of Deep Submicron", *Proc. of the Int. Conference on Computer-Aided Design*, 1998, 203-211.

[3] D. Sylvester and K. Keutzer, "Getting to the Bottom of Deep Submicron II: a global wiring paradigm", *Proc. of the 1999 Int. Symp. on Physical Design*, 1999, 193-200.

[4] C. Szyperski, *Component Software: Beyond Object Oriented Software*, Reading, MA, ACM/Addison Weseley, 1998.

[5] D. Gajski, R. Dömer and J. Zhu, "IP-Centric Methodology and Design with the SpecC Language" in *System Level Design*, Edited by Ahmed A. Jerraya and Jean Mermet, Nato Science Series 357, 1999.

[6] F. Vahid and T. Givargis, "Platform Tuning for Embedded Systems Design", *IEEE Computer* 34, 3.

[7] K. Keutzer, S. Malik, A. Newton, J. Rabaey and A. Sangiovanni-Vincentelli, "System Level Design: Orthogonolization of Concerns and Platform-Based Design", *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems* 19, 12 (Dec. 2000).

[8] W. Dally and B. Towles, "Route Packets, Not Wires: On-Chip Interconnection Networks", *Proc. of the Design Automation Conference*, Jun. 2001.

[9] D. Wingard, "MicroNetwork-Based Integration of SOCs", *Proc. of the 38th Design Automation Conference*, Jun. 2001

[10] M. Sgroi et. al., "Addressing the System-on-a-Chip Interconnect Woes Through Communication-Based Design", *Proc. of the 38th Design Automation Conference*, Jun. 2001.

[11] A. Hemani, A. Jantsch, S. Kumar, A. Postula, J. Oberg, M. Millberg, and D. Lindqvist, "Network on Chip:An architecture for billion transistor era", *Proc. Of the IEEE NorChip Conference*, Nov. 2000.

[12] A. Jantsch, J. Soininen, M. Forsell, L. Zheng, S.Kumar, M. Millberg, and J. Öberg, "Networks on Chip", *Workshop at the European Solid State Circuits Conference*, Sep. 2001.

[13] L. Benini and G. DeMicheli, "Powering Networks on Chip", *Proc. of the 14th Int. Symp. on System Synthesis*, 33-38, Oct. 2001.

[14] F. Catthoor, D. Verkest, and E. Brockmeyer, "Proposal for unified system design meta flow in task-level and instruction –level design technology research for multi-media applications", *Proc. 11th Int. Symp. on System Synthesis*, 1998, 89-95.

[15] P. Panda, N. D. Dutt, and A. Nicolau,, "Local Memory Exploration and Optimization in Embedded Systems", *IEEE Trans. on Computer Aided Design of Integrated Circuits and Systems* 18, 1 (1999), 3-13.

[16] A. Hemani et. al. , "Lowering power consumption in clock by using Globally Asynchronous Locally Synchronous Design style", *Proc. of Design Automation Conference*, 1999, USA. [17] Yi-Ran Sun, "Simulation and Performance Evaluation for Network on Chip", *MSc thesis*, Dept. of Microelectronics and Information Technology, Royal Institute of Technology, Stockholm.

[18] Dinesh Pamunuwa et. al., " A study of Physical Issuesin the design of an on-chip regular communication network", *submitted to DAC 2002*.

[19] V. Leppänen, Studies on the realization of PRAM, *Dissertation 3*, TUCS, University of Turku, 1996.

[20] M. Forsell and S. Kumar, Virtual Distributed Shared Memory for Network on Chip, *Proc. of the 19th IEEE NORCHIP Conference*, Nov. 12-13, 2001, Kista.

[21] Semiconductor Industry Association, *International Technology Roadmap for Semiconductors*, World Semiconductor Council, Edition 1999, 1999.

[22] Shashi Kumar, Axel Jantsch, Mikael Millberg, Johny Öberg, Juha-Pekka Soininen, Martti Forsell, Kari Tiensyrjä, Ahmed Hemani, "A Network on Chip Architecture and Design Methodology," isvlsi, pp.0117, IEEE Computer Society Annual Symposium on VLSI (ISVLSI'02), 2002

[23] Evaluation of Static and Dynamic Task Mapping Algorithms in NoC-Based MPSoCs Ewerson Carvalho, César Marcon, Ney Calazans and Fernando Moraes Pontifícia Universidade Católica do Rio Grande do Sul (FACIN-PUCRS) Av. Ipiranga, 6681 – P32 – 90619-900 – Porto Alegre – RS – Brasil