Virtual Circuit Switching Based Smart NOC

P. Jashuva 1, Sk. Rahil Hussain 2

P.G. Student, Department of ECE, DVR & HS MIC College of Technology, Kanchikacherla, AP, India 1
Assistant Professor, Department of ECE, DVR & HS MIC College of Technology, Kanchikacherla, AP, India 2

ABSTRACT: NoC solutions needs special kind of CPUs that consumes less power, is cheaper, smaller, but still has high performance requirements. To fulfil all these demands, they are getting more and more complex as the number of transistors is rapidly growing which has led to the emerging of multiprocessors systems-on-a-chip. This is developed on Xilinx 14.2 version by using vertex 4 tool kit to show the performance on the design of FPGA unit.

KEYWORDS: NoC, CPU, Power, Cheaper, Smaller, Performance.

I. INTRODUCTION

This State of the Art Report covers computer architecture topics with emphasis on Network on Chip (NoC). The reader is introduced to the basic ideas behind NoC and general computer architecture concepts before presenting an in-depth analysis of three important NoC components: CPU, Interconnect and Memory architecture.

NoC stands for Network-on-Chip and is a term for putting a complete system on a single piece of silicon. NoC has become a very popular word in the computer-industry, but very few agree on a general definition of NoC [19]. There are several alternative names for putting a system on a chip, such as system on silicon, system-on-a-chip, system-LSI, system-ASIC, and system-level integration (SLI) device. Some might say a large design automatically makes it a NoC, but that would probably include every existing design today. A better approach would be to say that a NoC should include different structures such as a CPU-core, embedded memory and peripheral cores. This still is a wide definition which could imply that any modern processor with an on-chip cache should be included into the NoC-community. Therefore a more suitable definition of NoC would be a complete system on a single piece of silicon, consisting of several types of modules including at least one processing unit designated for software execution, where the system depends on no or very few external components in order to execute its task.

A repetition is a task expressing how a single sub-task is repeated; each instance of the repeated task operates with sub-arrays of the inputs and outputs of the repetition. Making the repetitions independent and therefore parallel by construction, it allows expressing data parallelism. For a given input or output, all the sub-array instances have the same shape, are composed of regularly spaced elements and are regularly placed in the array. The RSM constructions can be used to compactly model repetitive architectures. It is especially useful for architectures with a large number of identical components, like a repetition of routers, where a compact representation both simplifies the modelling stage and concretely identifies the components with identical properties. The available modelling mechanisms in RSM are directed towards two aspects. The first aspect is that RSM helps specify the shape of a repetition using a multidimensional form and allows the representation of the potential instances as a multidimensional array. This repetition can be specified for an instance or a port of a component. RSM also provides a mechanism for expressing the potentially complex topologies of the links between these arrays of ports or instances. Complex, regular and repetitive structures such as cubes and grids can be modelled easily in a compact manner via the RSM package.
As NoC grew in quantity of IP cores, bus architectures and crossbars were exposed of their deficiencies. Shared bus architectures led to resource contention and hierarchical bus architectures and crossbar designs generated complexity. Interconnection networks offer an attractive solution to this communication predicament and are becoming persistent in NoC systems. A well-designed interconnection network makes a well-organized use of limited communication means while providing low latency, high bandwidth communication amongst different IPs with a minimum cost and low energy-dissipation. Undeniably, as system density and integration continued increasing, quite many designers discovered that it is more efficient to route packets, not wires. Utilizing an interconnection network in a NoC than a dedicated wiring permits limited bandwidth to be shared so that it can be used resourcefully. In contrast, dedicated wiring is idle mostly. Using a network also administrs regular, organized use of communication resources, making NoC easier to design, repair, and optimize.

II. ARBITER

The Arbiter unit shares resources among several potential bus masters in a round-robin fashion. The implementation is done generically, so it can be configured to handle arbitrary number of nodes. It is implemented as an asynchronous unit. A round-robin token passing bus or switch arbiter guarantees fairness (no starvation) among masters and allows any unused timeslot to be allocated to a master whose round-robin turn is later but who is ready now. A reliable prediction of the worst-case wait time is another advantage of the round-robin protocol. The worst-case wait time is proportional to number of requestors minus one. The protocol of a round-robin token passing bus or switch arbiter works as follows. In each cycle, one of the masters (in round-robin order) has the highest priority (i.e., owns the token) for access to a shared resource. If the token-holding master does not need the resource in this cycle, the master with the next highest priority who sends a request can be granted the resource, and the highest priority master then passes the token to the next master in round-robin order.
Multi processor Architecture

- Moore’s law is making the multiprocessor a commodity part
  1. 500M transistors on a chip, what to do with all of them?
  2. Not enough ILP to justify a huge uniprocessor.
  3. Really big caches that increases, diminishing miss returns.

- Chip multiprocessors (CMPs)
  1. Multiple full processors on a single chip
  2. Example: IBM POWER4: two 1GHz processors, 1MB L2, L3 tags

- Multiprocessors a huge part of computer architecture
  1. Multiprocessor equivalent of H+P is 50% bigger than H+P
  2. Another entire cores on multiprocessor architecture
     ❖ Hopefully every other year or so (not this year)

![Network Architecture](image)

- Multiprocessing can be very power efficient
- Recall: dynamic voltage and frequency scaling
- Performance vs. power is NOT linear
- Example: Intel’s Xscale
- 1 GHz " 200 MHz reduces energy used by 30x
- Impact of parallel execution
- What if we used 5 Xscale at 200 MHz?
- Similar performance as a 1Ghz Xscale, but 1/6th the energy
- 5 cores * 1/30th = 1/6th
- Assumes parallel speedup (a difficult task)
- Remember Ahmdal’s law

The memory on each node is a dual-ported, zero wait-state Xilinx block RAM. Memory accesses are made transparent through a wrapper module, which makes it easier to change between different target technologies. Both internal and external communication is managed by a Network Interface (NI). The NI encapsulates the complexity of communication and address translation, acting as a simple MMU.

In distinct usage in distributed computing, the term “Network Architecture” often describes the structure and classification of distributed application architecture, as the participating nodes in a distributed application are often referred to as a "network". For example, the applications architecture of the public switched telephone network (PSTN) has been termed the Advanced Intelligent Network. There are any number of specific classifications but all lie on a continuum between the dumb network (e.g. Internet) and the intelligent computer network (e.g. the telephone network).
Other networks contain various elements of these two classical types to make them suitable for various types of applications. Recently the context aware network, which is a synthesis of two, has gained much interest with its ability to combine the best elements of both. A popular example of such usage of the term in distributed applications, as well as PVCs (permanent virtual circuits), is the organization of nodes in peer-to-peer (P2P) services and networks. P2P networks usually implement overlay networks running over an underlying physical or logical network. These overlay networks may implement certain organizational structures of the nodes according to several distinct models, the network architecture of the system. Network architecture is a broad plan that specifies everything necessary for two application programs on different networks on an Internet to be able to work together effectively.

Network architecture is the design of a communication network. It is a framework for the specification of a network's physical components and their functional organization and configuration, its operational principles and procedures, as well as data formats used in its operation. In telecommunication, the specification of a network architecture may also include a detailed description of products and services delivered via a communications network, as well as detailed rate and billing structures under which services are compensated. The network architecture of the Internet is predominantly expressed by its use of the Internet Protocol Suite, rather than a specific model for interconnecting networks or nodes in the network, or the usage of specific types of hardware links. The Open Systems Interconnection model (OSI model) is a product of the Open Systems Interconnection effort at the International Organization for Standardization (ISO). It is a way of sub-dividing a communications system into smaller parts called layers.

Network architecture is the design of a communication network. It is a framework for the specification of a network's physical components and their functional organization and configuration, its operational principles and procedures, as well as data formats used in its operation. In telecommunication, the specification of a network architecture may also include a detailed description of products and services delivered via a communications network, as well as detailed rate and billing structures under which services are compensated. The network architecture of the Internet is predominantly expressed by its use of the Internet Protocol Suite, rather than a specific model for interconnecting networks or nodes in the network, or the usage of specific types of hardware links. The Open Systems Interconnection model (OSI model) is a product of the Open Systems Interconnection effort at the International Organization for Standardization (ISO). It is a way of sub-dividing a communications system into smaller parts called layers.
Each processor has switch to memory bus horizontally and processor-to-switch links vertically. A switch $S$ having four I/O paths ($0, 1, 2, 3$) provides the following twelve paths (assume even paths horizontal, odd paths vertical): $0-2$, $0-1$, $0-3$, $1-0$, $1-2$, $1-3$, $2-0$, $2-1$, $2-3$, $3-0$, $3-1$, $3-2$.

**Figure 6:** Examples of network architecture

III. PERFORMANCE CHART
IV. EXPERIMENTAL RESULTS

Figure 7: Simulation results for the proposed NoC multiprocessor

INPUT WAVEFORMS
OUTPUT WAVEFORMS

Table 1: Device Utilization

V. CONCLUSION

This project shows that it is possible to develop (including a whole CPU core) a multiprocessor NoC that fits on a single FPGA in very short time. A platform concept gives rapid design time and keeps the designers focused on debugging the application and not the platform itself. An “all synthesizable” approach is far from feasible in most projects but it is a convenient way of getting rapid results. A system view is crucial when designing NoC. All parts interact and focusing on narrow HW or SW issues does not give a satisfying solution. The platform concept will be dominating until IP companies have solved today’s problems with reusability and interoperability.

VI. FUTURE WORK

Design for Test is assumed to be implemented with standard tools and design flows. This can be accomplished due to our “all synthesizable” approach. Our scalability relies on a “perfect” interconnect. Today, a shared bus is used, which must be improved in the future with point-to-point/crossbar options in the generator. Software targeting is an important area and was solved by letting the user partition the application at thread level. This can be done automatically with respect to thread communication and code/data size. Inter-processors communication HW support will be integrated from existing research projects at MRTC. With automatic generation of synthesis scripts, a pushbutton design flow seems within reach. Virtual memory and dynamic memory allocation HW support will be added.
REFERENCES


BIOGRAphy

Mr. SK. Rahul Hussain was born in Guntakal, Ananthapur (Dist.), A.P, India. He received Bachelor of Technology in Electronics and Communication from ANNA University, Chennai, T.N, India. He has received Master of Technology (VLSI) from K L University, Andhra Pradesh, India. He is a Secretary of International Faculty and Students Multi Research Consortium. His research interests include Analog VLSI Design, Wireless Communication and Sensor Networks, Digital VLSI Design and Low Power Memory Design and Fault Diagnosis. He is working as Assistant Professor in Department of Electronics & Communication Engineering and Coordinator for Polytechnic Program, DVR and Dr. HS MIC College of Technology, Kanchikacherla, Krishn (Dist.), A.P. India. He has Published Nine International Journals and One International Conference.

Mr. P. Jashuva was born in Andhra Pradesh, India in 1991. He received Bachelor of Technology in Electronics and Communication Engineering (ECE) from DJR College of Engineering and Technology in 2013. Presently he is pursuing Master of Technology in VLSI & Embedded Systems Engineering in DVR and Dr. HS MIC College of Technology, Kanchikacherla, Krishna (Dist.), A.P. India. His research interests include field-programmable gate array architecture design, reconfigurable systems and NoC, high performance communicating circuitry and VLSI design.