Next Generation On-chip Communication Networks

Seongmoo Heo, Jason Kim, and Albert Ma
6.893 Project Report (the first checkpoint)
{heomoo,jasonkim,ama}@cag.lcs.mit.edu

Abstract

Due to the constraints of VLSI scaling, future processor and system-on-chip designs will by necessity incorporate on-chip communication networks. In this paper, protocols and signaling technologies are explored in the context of future on-chip multiprocessors in the 100nm regime. Scaling trends for devices and wires are predicted and based on these models, protocols and circuits are designed.

Because future multiprocessor chips require different network functionalities, tasks, and data transfer properties, design space for the new chip architecture is explored, including communication paradigm, data type, topology, switching technique, routing protocol, and node organization. A working communication protocol and a well-defined network architecture are designed to serve the need of next generation multi-processor chips.

At the 100nm regime, interconnect delay becomes a major challenge and needs to be taken into account at all levels. High-speed low-power CMOS drivers, receivers, and repeaters for global on-chip interconnection are designed and evaluated in terms of energy and delay. Our goal is 2.7GHz speed, which is about 10 FO4 delays, under 1.2V supply.

1 Related Work

A summary of scaling trends and issues are presented in [13, 6]. Based on these scaling trends, many researchers have concluded the only scalable architectures in the billion transistor era must consist of an array of processing nodes within a chip.[9, 7, 15]

As MIMD architectures became popular in the 80's, the network that provided the communication channel between processing elements became an important design focus. Since the first generation of multicomputers, interconnect networks have been designed, specifically tailored to the distinct features of each machine's communication style. One obvious interconnect characteristic is topology, rang-ing from the most popular 2D mesh to the exotic 3D cube. For example, MIT Alewife and Stanford Dash both had a 2D mesh network, while Cray T3D and MIT J-machine had a cube topology. There are other important characteristics as well such as switching technique, routing protocol, and node organization [2]. Also, some characteristics of the switch or the machine feature have been explored to give rise to new designs. For example, MIT Alewife machine attempted to integrate shared memory with message-passing communication [8].

The big difference between previous interconnect networks and the one we propose to build is that we deal with on-chip communication networks. Until now, process technologies allowed only a few processors on a chip, giving opportunity to build a bus-based network and leaving little incentive to build a switch-based network. However, with process technology nearing 0.13μm, designers began to have billions of transistors at their disposal and multiple tens of processors will be fabricated on chip. As a result, a switch-based network will rise to popularity. Then, the question is, what kind of a switch-based network will be appropriate? This question will be answered differently than in the past because the on-chip network should allow close coupling to the processing element at very high switching speeds and ,at the same time, deal with slow global interconnect delay at the future process. A very recent example of such an on-chip network is the communication network on RAW [15].

Low latency and low-power switching network also involves encoding and decoding of data at both ends of a point-to-point network. The weight encoding [10] and phase modulation [11] technique are such examples.

Simple inverters have been sufficient as a driver/receiver and a repeater for on-chip global interconnection previously. Therefore, most of the earlier works on interface circuit designs focused on sizing problems of multi-stage buffering to optimize for power [4].

However, recently low-power design has become one of the most important design criterion in VLSI circuits and it was found that long global interconnection can spend

big chunks of total power. Therefore, There have been many research efforts on the field of low-power on-chip global signaling scheme. There are two main techniques developed. The first one focuses on reducing the voltage swing on wire [5, 1]. This scheme needs a special low-swing driver and a level converting receiver. The second technique reduces power consumption by utilizing charge-sharing between bit-lines. Hui Zhang's paper shows excellent reviews on current low-power drivers/receiver circuit designs [16].

Until now, wire speed was not a main concern for on-chip interconnection designers. Even long wire delay was significantly faster than gate delay. However, many research works revealed that in the near future, global wire cannot keep up with the ever-growing speed of logic gate. It is expected that the RC delay of on-chip global wire will increase by 2 times per generation due to the increase of die size and the difficulty of metal-line scaling [6]. Many works on off-chip high-speed electrical signaling have been done [3, 12], but there is little work done in on-chip high-speed for the future process.

2 Methodology

2.1 Process

Based on the projections in [13], we generated spice device decks and Space[14] parameter decks. The spice decks model the transistors, while the Space decks are used in the calculation of wiring capacitances from layout. The spice decks were generated through BPTM, which is provided by the Device Group at UC Berkeley. The Space parameter decks were generated by hand.

2.2 Protocol

As a first step, design space for building network protocol and architecture will be explored. Since the sole user of the network is a programmer, the network must be able to handle the needs of the programmer and the surrounding chip environment. This means that we need to support message passing, both expected and unexpected, and shared memory. In addition, streaming data and I/O interrupts and transfers must also be supported. These systems affect the type of data that will travel on the network and most importantly, the behavior of the network. As for topology and routing scheme, the design is set to use a 2D topology and a point-to-point crossbar interconnect. This is because 2D is most suited for scaling in VLSI on-chip and a point-to-point crossbar allows fast multicast/broadcast messages on the network. Two other design spaces will be explored. For switching techniques, there
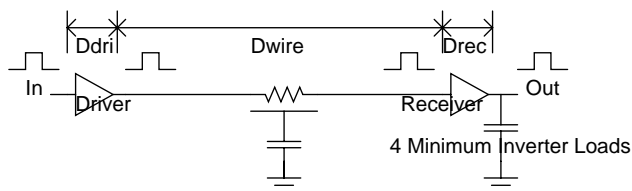

Figure 1: Test bench. The top metal layer is used for wire and its length is the predicted size of a tile.

are packet-switching (store-and-forward), circuit switching, and worm-hole routing. For routing schemes, there are deterministic and adaptive routing.

Further down the line, there are issues regarding the amount of buffering along the node and at the destination to the processor, network-to-processor interface, and handshake between communicating nodes. Another important issues in designing networks involve deadlock and livelock.

When all the design space have been explored, we will build a working protocol and a well-defined architecture that will be ready to be implemented in future on-chip multiprocessor.

2.3 Circuit

We focus on CMOS voltage-mode serial-link low-swing differential signaling circuits. Current-mode is currently popular in off-chip interconnection because it is robust to voltage supply noise and usually low-power [12]. But RC characteristics of on-chip wire makes the application of this technique hard to on-chip communication. We may work on it after finishing works on voltage-mode. Bus architecture is prohibited because it is almost impossible to get fast speed at the future process with bus signaling scheme. Low-swing differential signaling is robust to noise, low-power and fast. Differential signaling is believed to give almost 2 times faster speed.

First of all, we will choose the ones which are appropriate for the future process among various conventional interconnection circuits based on the predicted technology models. We build a standard test bench for fair comparison and compare candidates in terms of energy and delay. The test bench is shown in Figure 1. Besides, based on the comparison results, we will try to build new circuit designs. Also, we are planning to do research on optimal length for inserting repeaters and optimal width of wire. Xcircuit is used for the schematic design, Magic is used as the layout editor, and SPACE [14] extracts the layouts. The extracted netlists are simulated with HSpice simulator to measure energy consumption and delay.

References

[1] A. Bellaouar, I. S. Abu-Khater, and M. I. Elmasry. Low-power cmos/bicmos drivers and receivers for on-chip interconnects. *IEEE Journal of Solid-State Circuits*, 30(6):696–700, June 1995.

[2] Peter Kacsuk Dezso Sima, Terrence Fountain. *Advanced Computer Architecture: A Design Space Approach*. Addison-Wesley, 1997.

[3] M. Horowitz et al. High-speed electrical signaling: Overview and limitations. *IEEE Micro*, pages 12–24, Jan-Feb 1998.

[4] N. Hedenstierna et al. Cmos circuit speed and buffer optimization. *IEEE Trans. on Computer-Aided Design*, 6(2):270–281, March 1987.

[5] Y. Nakagome et al. Sub-1-v swing internal bus architecture for future low-power ulsi's. *IEEE Journal of Solid-State Circuits*, 28(4):414–419, April 1993.

[6] R. Ho, K. Mai, and M. Horowitz. The future of wires. In *IEEE Special Proceedings*, December 2000.

[7] Christoforos E. Kozyrakis et al. Scalable processors in the billion-transistor era: Iram. *IEEE Computer*, pages 75–78, September 1997.

[8] John Kubiatowicz. *Integrated Shred-Memory and Message-Passing Communication in the Alewife Multiprocessor*. PhD thesis, M.I.T., Feb 1998.

[9] Ken Mai et al. Smart memories: A modular reconfigurable architecture. In *Proc. 27th Intl. Symp. Computer Architecture*, pages 161–171, June 2000.

[10] Kazuyuki Nakamura and Mark A. Horowitz. A 50% noise reduction interface using low-weight coding. *Symposium on VLSI Circuits Digest of Technical Papers*, pages 144–145, June 1996.

[11] Kasutaka Nogami and Abbas El Gamal. A cmos 160mb/s phase modulation i/o interface circuit. *ISSCC Digest of Technical Papers*, pages 108–109, June 1994.

[12] J. Poulton. Signalling in high-performance memory systems. In *Proceedings ISSCC*, 1999.

[13] SEMATECH. *International Technology Roadmap for Semiconductors*, 1999 edition.

[14] N. P. van der Meijs and A. J. van Genderen. Space user's manual. Technical Report ET-NT 92.21, Delft University of Technology, Dept of EE, Delft, The Netherlands, April 1992,1997.

[15] Elliott Waingold et al. Baring it all to software: Raw machines. *IEEE Computer*, pages 86–93, September 1997.

[16] H. Zhang and J. Rabaey. Low-swing interconnect interface circuits. In *Proceedings ISLPED*, pages 161–166, Monterey, CA, 1998.