Bicephaly: Maximizing Bandwidth by Duplexing Power and Data

Eric Fontaine, School of Electrical and Computer Engineering, Georgia Tech
Hsien-Hsin S. Lee, School of Electrical and Computer Engineering, Georgia Tech
http://arch.ece.gatech.edu

1. INTRODUCTION

As the number of processor cores integrated on a single die continues to grow, so will the demand for power and data. However, ITRS [1] predicts the number of pads for high-performance processors to remain unchanged. With this trend, it will be a great challenge to sustain scalable performance by using a many-core processor to crunch data, particularly when workloads are memory- or bandwidth-bound. This *bandwidth* limitation will be a bottleneck for transporting sufficient data into the processor, jeopardizing the motivation of integrating a massive number of cores onto a single package. Furthermore, emerging 3D integration technology only exacerbates this situation, since the area footprint for I/O and power supply pads is shrunk and will be shared by microfluid channels for cooling. Although the industry has been successful in the last decade by elevating the front-side bus (FSB) frequency from 66MHz to today's 1333MHz, more creative techniques are needed. Data compression is one such technique, but it is only useful if the compression rate provides larger benefits over the overheads of the decompression/compression procedure during read and dirty eviction.

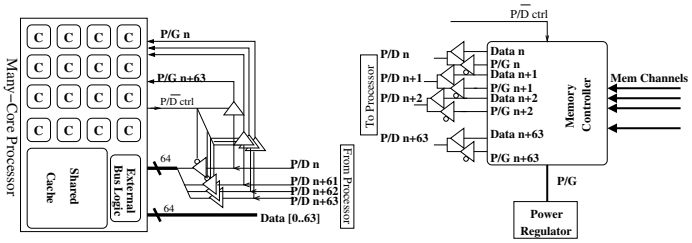
In this abstract, we introduce a drastically alternate approach to improving bandwidth by exploiting the following phenomenon: when a processor becomes bandwidth-bound, it does not require as much power. When the data supply is disrupted due to last level cache misses, cores and/or functional units will idle waiting for data to arrive. Since packages have a limited number of pads, our basic idea is to share the existing scare resources by converting some of the *power/ground* lines to *data* lines when the processor becomes bandwidth-bound.

2. OUR SOLUTION: BICEPHALY

Although two thirds of the I/O pads on a high-performance processor today are dedicated to power delivery to satisfy the worse-case scenario when large currents are demanded for driving all functional units, in common cases these power/ground lines will not be so stressed. Under these premises, similar to the concept of sharing resources in many part of a computing system, we ask the following question — can we address the bandwidth limitation issue by sharing and reconfiguring certain amount of the power/ground lines for data transmission?

Figure 1 shows a design instance to demonstrate our idea. The illustrations include changes in the interface between the processor and memory controller. In Figure 1(a), the duplexable power/ground lines (in a group of 64 bits) can be reconfigured into an expanded data bus when the processor switches into high-bandwidth mode. On the other hand, the power/ground signals generated by the power regulator are routed into the memory controller as illustrated in Figure 1(b). These signals can be displaced and reconfigured into the expanded data bus (denoted by Data n to Data n+63) with the same control bit that control the duplex of incoming power or data signals in Figure 1(a). Note that Figure 1 simply shows an example of duplexing 64 lines between power/ground and data, however any number of duplexable pins could be used. In fact, if the entire working set fits into the on-chip cache, all data lines could potentially be utilized for providing extra power.

During the high bandwidth mode, any power saving technique can be performed to lower the demand for power, including (1) dynamic voltage and frequency scaling (DVFS) of the core(s), (2) disabling some cores, (3) disabling functional units (clock gating or power gating), (4) disabling cache lines (for data streaming workloads). Although these techniques by themselves could reduce performance, during bandwidth-bound situations, the additional off-chip bandwidth using our technique allows the processor to increase throughput instead of sitting idle. Once the workload is no longer bandwidth-bound, the processor will revert back to the normal operation mode, taking the full advantage of its peak computing power.



(a) The Processor Side (a) The North Bridge Side

Figure 1: Duplexing Power/Ground and Data Signals

3. CHALLENGES

There are several challenges to be resolved in our proposed technique. First, we need to implement a dynamic reconfiguration mechanism to automatically duplex the power/ground and data lines. Furthermore, the mechanism needs to precisely determine which particular power/ground lines (out of some hundreds) can be duplexed without jeopardizing the power delivery network of the processor. Second, such duplexing could lead to large IR drop and simultaneous switching noise (L di/dt noise) if the displaced power/ground lines are the main power suppliers to nearby functional units. This reliability issue can be resolved by either choosing the duplexed power/ground pin candidates more cautiously or by implementing other signal integrity guarantees. Last but not least is the control logic used to monitor processor performance and determine when to switch bandwidth modes. This criterion could be in the form of some expression consisting of available performance counters such as IPC and FSB utilization. To avoid frequent detrimental switching patterns, and upper and lower threshold should be used.

Note that the overall goal is to improve the overall performance of a many-core system. Keeping all the cores busy to unleash the maximal computation throughput requires a system to balance data delivery and data consumption. Therefore, the performance compromised by reducing the power in the high-bandwidth mode needs to be compensated by the additional throughput enabled by the extra bandwidth. Naturally, there is an optimal reflection point when making such a trade-off.

4. REFERENCES

[1] ITRS Executive Summary, 2006 Update. Technical report, International Technology Roadmap for Semiconductors, 2006.