Full System Simulation of Optically Interconnected Chip Multiprocessors Using gem5

Anouk Van Laer†, Timothy Jones§, Philip M. Watts†

†Dept. of Electronic and Electrical Engineering, University College London, London, UK
§Computer Laboratory, University of Cambridge, Cambridge, UK
a.laer@ee.ucl.ac.uk, timothy.jones@cl.cam.ac.uk, philip.watts@ucl.ac.uk

Abstract: By extending the cycle accurate gem5 computer system simulator with optical network models, we demonstrate chip multiprocessor performance improvements of up to 18% using a wavelength striped optical crossbar interconnect and determine optimum optical parameters.
© 2013 Optical Society of America
OCIS codes: (060.4250) Networks (200.4650) Optical Interconnects

1. Introduction

It has been shown that optical networks-on-chip can reduce power consumption of chip multiprocessors (CMP), popularly known as multicore processors [1, 2, 3, 4]. Typically, these networks are evaluated for latency and throughput using simulation of either synthetic traffic patterns or trace data, the latter coming from either real systems or computer architecture simulations. However, this methodology does not allow the characteristics of the network to have any influence on the performance of the programs. Given the differences between optical and electronic networks, it is important to be able to evaluate the performance of real applications that utilise the network, which requires full system simulation. The full system simulations presented in [4] use the Graphite simulator which, while able to simulate a large number of cores relatively rapidly, cannot accurately model contention for network and other system resources. In this work, we propose a framework to fully investigate the performance of optically interconnected CMPs by extending the open source and cycle accurate gem5 simulator [5] with optical network models. gem5 can simulate a complete CMP system from the microarchitectural level up to the operating system level. This system is then capable of booting Linux and running parallel applications, for example, the PARSEC benchmark suite [6].

The gem5 optical extensions, which we intend to make available to the research community, allow us to investigate the consequences of choices made in the architecture of the optical network on the run-time and other performance metrics of real CMP systems running real applications. In this paper, we use the gem5 optical network extensions to demonstrate the performance effect of using a wavelength striped time division multiplexed optical crossbar on CMP performance and determine the optimum optical parameters.

2. Simulation set-up

In this work, we focus on CMPs with shared memory because this enables efficient parallel programming. On-chip communications in shared memory CMPs take place between the L1 caches, L2 caches and the main memory controllers. Messages consist of memory block with fixed length typically in the range of 64–$256B$ ($128B$ in this work) and cache coherence control messages ($8B$). We model CMPs based on a compute tile (Figure 1a) consisting of an x86 core with private L1 cache (16 kB for data, 16 kB for instructions), proportion of the distributed L2 cache (1MB total regardless of the number of cores) and a network interface which can be either electronic or optical, both of which operate at 1 GHz clock frequency. Although we use 8–16 cores in this work to reduce simulation time, gem5 is capable of simulating much larger systems.

The short messages of shared memory systems are problematic for a switched optical networks because the switching and arbitration overhead is large compared with the message length. Various network architectures have been proposed to overcome these issues including wavelength routed all-to-all networks [9, 4], hybrid networks which reserve the optical network for large messages [1] or for higher layers of hierarchy [3]. By contrast, we model a simple silicon photonic crossbar interconnecting all cores (Figure 1c) with reconfiguration time $\sim 1\,ns$ [10] carrying WDM wavelength striped data [11] to maximise bandwidth per port and hence minimise serialisation latency. All messages, both data and control, are routed onto the crossbar.

Arbitration of the crossbar is based on time division multiplexing (TDM) with a fixed slot length, as shown in Figure 1d. Tiles that wish to transmit a message, make a request to the arbiter which uses round robin arbitration to
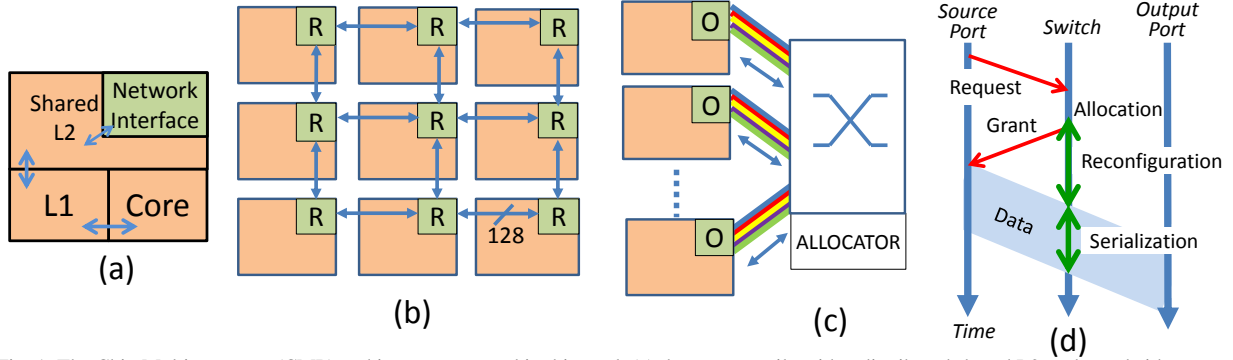
Fig. 1. The Chip Multiprocessor (CMP) architecture assumed in this work (a) the compute tile with a distributed shared L2 cache and either an electronic (R) or optical (O) network interface (b) tiles interconnected with an electronic mesh network of 5-port routers (c) tiles interconnected with an optical wavelength striped TDM crossbar. (d) the length of TDM slot mainly consists of the serialization latency and reconfiguration time. It should be noted however that the length of the arrows in this figure is not representative for actual timings.

determine the optical crossbar configuration for the next time slot. The slot length is given by the sum of the *serialization latency* and the switch *reconfiguration time* rounded up to an integer number of clock cycles. The serialization latency is determined by the number payload bits per slot, the number of wavelengths in the stripe and the modulation frequency. In this work, the number of payload bits per slot is fixed at four times the length of the cache block plus control overhead ($4 \times [8B + 128B]$) in order to cope with bursts and the modulation frequency is also fixed at 10 Gb/s. We then investigate the performance effect of varying the number of wavelengths and the switch reconfiguration time. Future work will also investigate the optimum slot length and the effect of varying it at run-time. When the arbiter has determined the crossbar configuration, it sends out grants to all nodes that are allowed to access the crossbar. These nodes then send as many messages as they have buffered (both cache block and control messages) up to the maximum which can fit within the slot, thus using the full data bandwidth available.

Current high performance shared memory servers use an electronic crossbar for on-chip interconnect [7], but area and power consumption does not scale well with increased number of cores. To compare with the optical TDM network, we use electronic mesh networks (Figure 1b), which have been widely proposed for larger core counts, e.g. [8]. We use a flit width (the parallel links between each router) of 128-bits for reduced latency. However, arguably this favours the mesh network, leading to approximately an order of magnitude greater power and area compared to the 32-bit width used in [8].

To investigate the influence of TDM arbitration on overall performance, four benchmarks from the PARSEC benchmark suite [6], `blackscholes`, `fluidanimate`, `swaptions` and X264 were used, representing typical CMP applications with different sharing patterns, parallelization models, size of the working set etc. Only the parallel part of the benchmark was simulated because network traffic outside this time is negligible.

3. Results

3.1 8 core simulations

Figure 2a (inset) shows the influence of reconfiguration time on the overall performance using 8 cores running `x264` with 4 wavelengths. Increasing reconfiguration time increases the slot size, resulting in a slight decrease in performance. However, across all benchmarks and numbers of wavelengths, the change in speed-up was no greater than 2% changing the reconfiguration time from 0.5 to 2.5 ns.

Figure 2a (main) shows the speed-up when using an optical crossbar with 1 ns reconfiguration time compared to an electrical mesh for the 4 benchmarks tested as the number of wavelengths in the stripe is varied. 4 or more wavelengths are required to consistently outperform the mesh network, but there is no significant improvement for more than 8 wavelengths. For 8 or more wavelengths, the speed-up that can be obtained varies greatly from 15.72% for `x264` to 1.67% for `fluidanimate`. X264 is more prone to changes in the interconnection network as the total message count is more than twice that of the other benchmarks. The effect is increased because of the parallization model used in X264. In `swaptions`, `blackscholes` and `fluidanimate`, the work is divided among threads that each work on a data parallel part of the program. The maximal speed-up is thus determined by the slowest thread. In `x264`, once one thread finishes, the results are then used by the next thread (pipelining), making the overall performance the combination of the speed-up of all threads. The speed-up in `fluidanimate` is small because it has a large working set that does not fit in the on-chip memory. As such, the frequent requests to the off-chip main memory
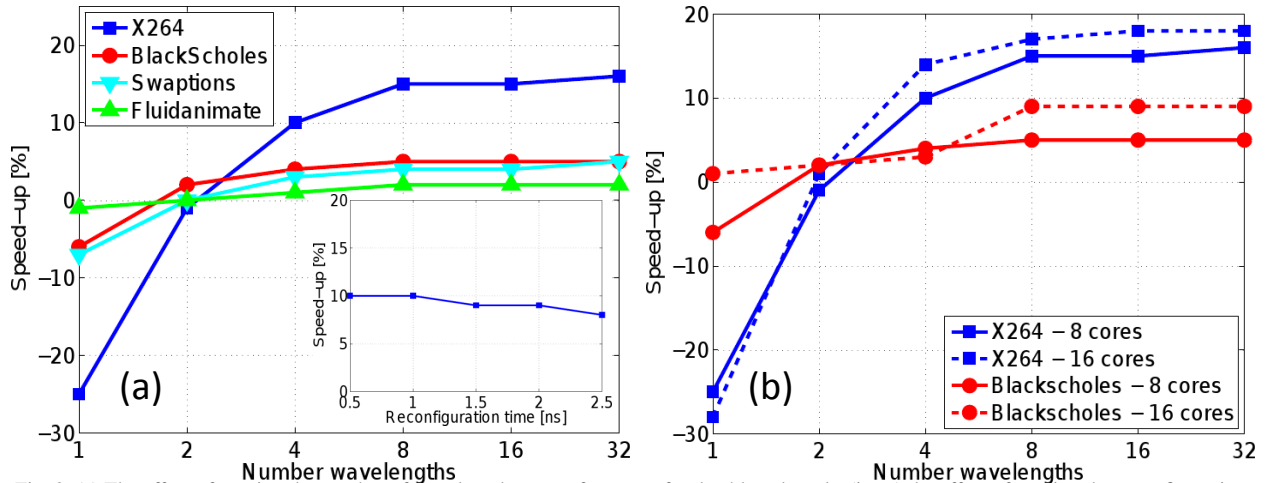
Fig. 2. (a) The effect of varying the number of wavelengths on performance for the 4 benchmarks (inset) the effect of varying the reconfiguration time on x264 with 4 wavelengths. (b) The effect of increasing the core count from 8 cores to 16 cores

dominate latency rather than the on-chip interconnection network.

3.2   16 core simulations

We repeated the simulations for x264 and blackscholes with a 16 core simulations to make a case for using optical networks for higher core counts. Only these 2 benchmarks were simulated as the behaviour of fluidanimate and swaptions is quite similar to blackscholes. Figure 2b shows the speed-up for both benchmarks when compared to a 16 core CMP with an E-mesh. The performance of both the E-mesh and the crossbar increase because of the benefits of increased parallelization. However, the speed-up of the optical crossbar relative to the mesh is increased for 16 cores due the increased hop count of the mesh network.

4.   Conclusions

Using full system cycle accurate simulation in gem5, we have demonstrated that a TDM optical crossbar increases performance of four parallel algorithms from the PARSEC benchmark suite compared with an electronic mesh network provided that wavelength striping with 4 or more wavelengths is used. This is despite the reconfiguration time required by the optical network and the favourable mesh network parameters used. We have shown that the benefits of the optical crossbar increase for larger core count. The gem5 simulator can be used to model performance of many different optical network architectures and we plan to make the optical extensions available to the research community.

References

1. G. Hendry, K. Bergman, L. Carloni, J. Shalf, "Analysis of Photonic Networks for a Chip Multiprocessor Using Scientific Applications", *Networks on Chips Symposium*, 2009
2. D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R. Beausoleil, J. Ahn, "Corona: System Implications of Emerging Nanophotonic Technology Computer Architecture", *Int. Symp. Computer Arch (ISCA)*, 2008
3. Y. Pan, P. Kumar, K. Kim, G. Memik, Y. Zhang, A. Choudhary, "Firefly: Illuminating future network-on-chip with nanophotonics", *Int. Symp. Computer Arch (ISCA)*, 2008
4. G. Kurian, J. E. Miller, J. Psota, J. Eastep, J. Liu, J. Michel, L. C. Kimerling, A. Agarwal,"ATAC: a 1000-core cache-coherent processor with on-chip optical network" *Proc. Int. Conf. on Parallel Architectures and Compilation Techniques*, 2010
5. N. Binkert, B. Beckmann, G. Black, S. K. Reinhardt, A. Saidi; A. Basu, J. Hestness; D. R. Hower, T. Krishna, S. Sardashti, R. Sen, K. Sewell, M. Shoaib, N. Vaish, M. D. Hill, D. A. Wood, "The gem5 simulator",*SIGARCH Comput. Archit. News* , 2011
6. C. Bienia,S. Kumar, J. P. Singh, K. Li, "The PARSEC Benchmark Suite: Characterization and Architectural Implications" *Proc. Int. Conf. on Parallel Architectures and Compilation Techniques*, 2008
7. J. L. Shin, H. Dawei,B. Petrick, H. Changku, K. W. Tam, A. Smith, Ha Pham, Hongping Li, T. Johnson, F. Schumacher, A. S. Leon, A. Strong, "A 40 nm 16-Core 128-Thread SPARC SoC Processor" *IEEE J. of Solid State Circuits*, Vol. 46, 2011
8. D. Wentzlaff, P. Griffin, H. Hoffmann, L. Bao, B. Edwards, C. Ramey, M. Mattina, C. Miao, Chyi-Chang, J. F. Brown III, A. Agarwal, "On-Chip Interconnection Architecture of the Tile Processor", *IEEE Micro*, 2007, 27
9. A. Krishnamoorthy, R. Ho, X. Zheng, H. Schwetman, J. Lexau, P. Koka, G. Li, I. Shubin, J. Cunningham, "Computer Systems Based on Silicon Photonic Interconnects" *Proceedings of the IEEE*, Vol. 97, 2009
10. A. W. Poon, X. S. Luo, F. Xu, H. Chen, "Cascaded Microresonator-Based Matrix Switch for Silicon On-Chip Optical Interconnection", *Proceedings of the IEEE*, Vol. 97, 2009
11. B. G.Lee, A. Lee, X. Qianfan, M. Lipson, K. Bergman."Characterization of a 4 4 Gb/s Parallel ElectronicBus to WDM Optical Link Silicon Photonic Translator", *IEEE Photonics Technology Letters*, Vol. 19, 2007