Alleviating Thermal Constraints While Maintaining Performance Via Silicon-Based On-Chip Optical Interconnects *

Nicholas Nelson, Gregory Briggs, Mikhail Haurylau, Guoqing Chen, Hui Chen
David H. Albonesi†, Eby G. Friedman, and Philippe M. Fauchet

Department of Electrical and Computer Engineering
University of Rochester, Rochester, New York 14627

†Computer Systems Laboratory
Cornell University, Ithaca, New York 14853

Abstract

The relentless pursuit of Moore's Law by the semiconductor industry has yielded significant increases in performance, but at the cost of greater power dissipation. As CMOS technology continues to scale, increasing power densities, or "hot spots," particularly in dense logic structures, may limit frequencies below projected targets in order to avoid circuit malfunction. A solution to this problem is to separate the hot spots by interleaving these units with cooler cache banks. This approach, however, increases the distance among processing functions, which can significantly degrade performance. While effort is made to localize communication as much as possible, global communication cannot be completely avoided, particularly in parallel applications.

In this paper, the use of silicon-based on-chip optical interconnects is investigated for minimizing the performance gap created by separating processing functions due to thermal constraints. Models of optical components are presented, and used to connect the common front-end with the distributed back-ends of a large-scale Clustered Multi-Threaded (CMT) processor. A significant reduction in thermal constraints (translated into an increase in clock frequency), combined with improved instructions per cycle (IPC), is demonstrated over a conventional all-electrical system.

*This research was supported by National Science Foundation grant CCR-0304574.

1. Introduction

Growing transistor densities, less than ideal scaling of global wires, and increasing clock frequencies, have led to excessive interconnect wire delay and significant heat dissipation in general purpose microprocessors. The industry move to multi-core chips creates the quandary of how to balance the need for high speed, high bandwidth communication and reasonable power density levels. These two criteria are often at odds as the former calls for functionality to be tightly packed, while the latter requires separation. This paper demonstrates that silicon-based on-chip optical interconnect technology is a promising solution to this growing problem.

In addition to interconnect delay, delay uncertainty has grown significantly. Greater delay uncertainty necessitates the introduction of registers along long distance lines, reducing the amount of useful work that can be accomplished within a clock cycle. Delay uncertainty is further increased by local and global temperature swings.

Increased power dissipation is a critical concern in microprocessors. The heat generated by localized high power dissipation leads to on-chip hot spots, producing potentially unstable circuit operation and local electromigration concerns. A solution to the problem of hot spots is physically separating the high power density components [12]. This strategy, however, exacerbates the problem of long lines and delay uncertainty. The temperature of a block is dependent on the amount of power dissipated in that block, and the temperature of the surrounding blocks. Highly active blocks inter-

spersed with blocks containing low activity will reduce the maximum temperature, although the overall power dissipation will remain the same.

This separation of microprocessor functions to alleviate thermal constraints has the undesirable effect of longer cycle times or deeper pipelines. A *clustered processor microarchitecture* separates processing units into *clusters*, with a dedicated interconnection network used for inter-cluster communication. Steering algorithms are used to limit inter-core forwarding, thereby limiting the increase in delay. A possible solution to long interconnect delay in *distributed microarchitectures* is the use of transmission-line connections [6]. While transmission-line connections provide fast communication, these structures are highly bandwidth limited. Wide thick lines also consume a significant amount of the upper metal layer area, limiting the number of possible connections.

Optical interconnects have previously been suggested as a potential solution to the global wire delay problem [23]. Traditionally, the use of on-chip optical interconnections require the integration of new materials, a prohibitively costly change, or bonding the optical components to a silicon CMOS circuit, also an expensive option. Accordingly, it was believed that optical interconnections are inappropriate for intra-chip communication [22]. Recent advances in silicon-based optical devices have solved many of the issues associated with CMOS-based optical devices. These proposed devices are constructed using traditional CMOS processing and materials, and significant progress has been made in electrical/optical conversion [7]. By 2010, for a 1 cm on-chip interconnect length, the propagation delay of an optical link is expected to be half that of an optimal electrical link with repeaters [8].

Although on-chip optical interconnects have recently been evaluated from device and circuit-level perspectives, similar work has yet to be performed at the architectural level. Thus, it is unclear from a systems perspective whether the use of optical interconnects to replace global on-chip wires is an attractive solution. In this paper, silicon-based optics for on-chip interconnects are investigated for a large-scale Clustered Multi-Threaded (CMT) processor microarchitecture [13]. Projections for optical and electrical interconnects for 45 nm CMOS are presented based on prior work [7, 8]. One potential benefit of optical interconnects is explored. Specifically, the processing elements are separated and interleaved with L2 cache banks to alleviate heat constraints, while low-latency optical connections from the centralized front-end to these back-end elements prevent undue performance loss. The resulting architecture exhibits a significant reduction in heat dissipation (translating into an increase in clock speed and improved reliability) for the same total power level with higher IPC. Although these results are obtained for a large-scale CMT organization, similar benefits can be achieved in a Chip Multi-Processor microarchitecture.

2. Optical System

The successful introduction of optical interconnects onto a microprocessor requires overcoming a number of barriers, the most significant being compatibility with a monolithic (silicon) microelectronic device technology. Due to the poor light emitting properties of crystalline silicon, the most viable option is to use an external light source (VCSEL laser, etc.) for optical signal generation. An external light source allows more compact and energy efficient electro-optical modulators as optical information transmitters. Furthermore, low-refractive index polymer waveguides for light propagation and SiGe detectors as receivers are potentially satisfactory candidates.

2.1. Modulator

An important example of an ultrafast silicon-based modulator has been demonstrated by Liu et al. [21]. The authors herein indicate that the physical device structure (without considering the driver delay) can operate at speeds in excess of 8 GHz. Moreover, Liu et al. mention that by thinning the gate oxide and using an epitaxial overgrowth technique, it is possible to enhance the phase modulation efficiency. Through additional device geometric optimization, it is also possible to increase the optical mode/active medium interaction volume. Thus, it is reasonable to assume that with technology improvements, the modulator speed will operate in the 30–40 GHz range by 2015. However, because the chosen device structure is a Mach-Zehnder interferometer, this type of modulator has a large footprint, resulting in excessive power consumption and increased driver delay. Simulations and initial experiments performed by Barrios et al. [2, 3] show that an alternative modulator topology—an optical microcavity—can drastically decrease the modulator area to 10–30 μm while maintaining the same operating speed. Based on these considerations, the capacitance of the modulator structure is estimated to be 1.36 pF.

A block diagram of a driver circuit is shown in Figure 1. The microcavity-based optical mod-

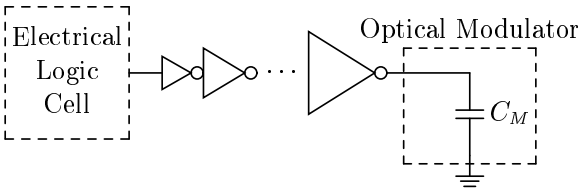

Figure 1. Circuit model of an optical transmitter.

ulator is assumed to be a purely capacitive load. A series of tapered inverters is used to drive the capacitor [10].

2.2. Receiver

The role of an optical receiver is to convert an optical signal into an electrical signal, thereby recovering the data transmitted through the lightwave system. The optical receiver has two primary components: a photodetector that converts light into electricity, and receiver circuits that amplify and digitize the electrical signal. A simplified equivalent circuit model is shown in Figure 2. In the context of on-chip optical interconnects, only those technologies that are fully compatible with silicon microelectronics are considered. A practical solution is a SiGe photodetector operating at a 1.3 μm wavelength.

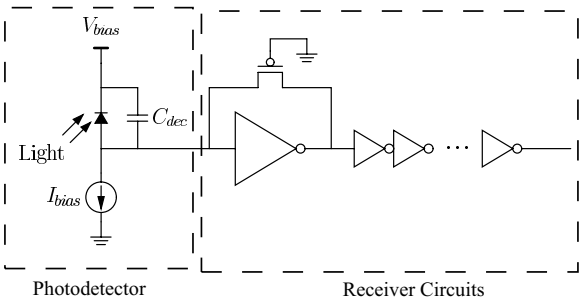

Figure 2. Circuit model of an optical receiver.

Many types of photodetectors exist due to the many different device structures and operating principles. Interdigitated SiGe p-i-n photodiodes and SiGe Metal-Semiconductor-Metal (MSM) detectors are considered here because these detectors tend to respond faster with the same quantum efficiency. In 2002, an interdigitated SiGe p-i-n detector fabricated on a Si substrate with a 3 dB bandwidth of 3.8 GHz at a 1.3 μm wavelength was demonstrated [24].

A summary of the delays of the individual elements along the optical data path is listed in Table 1. Note the significant delay advantage over optimal electrical interconnects with repeaters for a target length of 1 cm. More details describing the device/circuit aspects of the optical technology can be found in [7, 8].

Table 1. Delay (ps) in a 1 cm optical data path as compared with the electrical interconnect delay [8].

Modulator driver	25.8
Modulator	30.4
Waveguide	46.7
Photo-detector	0.3
Receiver amplifier	10.4
Total optical	113.6
Electrical	200.0

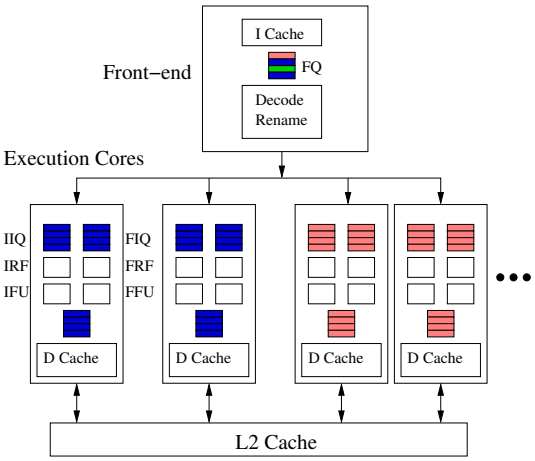

Figure 3. Clustered multi-threaded architecture with two cores per thread.

3. Architectural Design

The baseline processor is a clustered multi-threaded (CMT) machine [13] with a unified front-end, and 16 cores containing functional units, register files, and data caches for a back-end, as shown in Figure 3. The simulator is based on Simplescalar-3.0 [5] for the Alpha AXP instruction set with the Wattch [4] and HotSpot [16] extensions. Processor parameters are listed in Table 2.

3.1. Core Layout

A floorplan of the processing core (back-end) is illustrated in Figure 4. Each back-end is linearly scaled from the Alpha 21264 floorplan [17] to the 2010 (45 nm) technology node. Units whose parameters differ from the 21264 (*i.e.*, there are 64 integer registers rather than 80) are also linearly scaled.

The layout of the processor requires that each core has a level one data cache. The cache is as-

Table 2. Processor parameters.

Cluster	
L1 Data Cache	16 KB per core 2 way, 2 cycles
Load/Store Queue	64 entries
Register File	64 Int, 64 FP
Issue Queue	64 Int, 64 FP
Integer Units	2 ALU, 1 Mult
Floating Point	1 ALU, 1 Mult
Front end	
Combined Branch Predictor	2048 entry BTB
Return Address Stack	32 entries
Branch Mispredict Penalty	12
Fetch Queue Size	64 shared
Fetch Width	32 instructions from 2 threads
Dispatch	16 shared
Commit	12 per thread
Reorder Buffer	256 per thread
L1 Instruction Cache	32 kB 2 way
Unified L2 Cache	64 MB 32 way
TLB (each, I and D)	128 entries, 8KB fully associative per thread
Memory Latency	200 cycles

sumed to use a simplified coherence scheme. The mesh interconnect network is inherently unordered, and the delay from one point to another point is non-uniform. The cache coherence actions are performed in the order seen by the simulator. The level two data cache is clearly a non-uniform access time structure; for simplicity, however, it is simulated as a uniform access time structure. This approximation is accurate if the cache allows frequently accessed blocks to be moved closer to the utilizing cores [11].

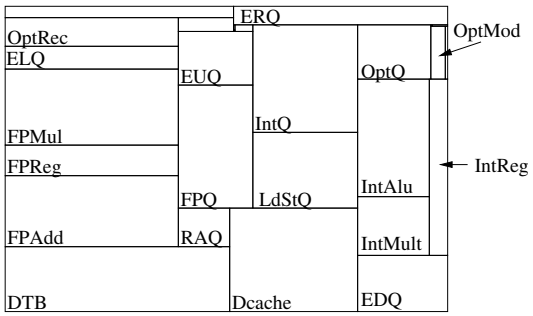

Figure 4. Core floorplan.

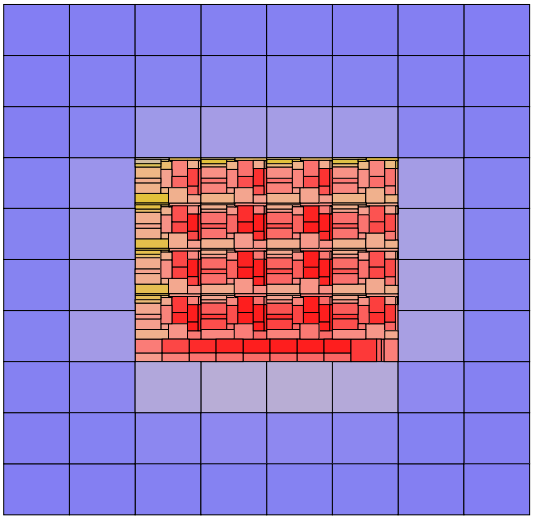

Figure 5. Grid floorplan. The back-end cores are in the center, above the common front-end, completely surrounded by a 64 MB unified L2 cache.

3.2. Processor Layout

Two layout strategies are compared to demonstrate the advantages of on-chip optical interconnects. The grid floorplan, as shown in Figure 5, is the baseline configuration, in which the cores are closely packed to minimize inter-cluster delay. This floorplan consists of 16 replicated cores surrounded by 64 banks of a unified level 2 cache. The second floorplan, shown in Figure 6, is proposed to reduce the maximum temperature while maintaining IPC performance. This floorplan has the advantage of spreading out the hot cores, thereby allowing the cool cache to reduce the temperature. Each of the 16 cores are surrounded by four banks of a unified level 2 cache.

A mesh Manhattan interconnection scheme is simulated; each core can communicate via electrical links with neighbors at a cost of one cycle. Communication between distant cores requires multiple hops, and congestion is considered. All of the electrical links are capable of serving two 64 bit values (2 registers) per cycle for each layout configuration. The shared front-end is located along the bottom of the core elements. In this study, optical links are only used for direct communication between the front-end (shown at the bottom) and each core. Communication over these optical links requires two cycles, compared to a worst case of seven cycles for wire interconnects.

4. Methodology

In this analysis, the maximum transient temperature of any functional unit limits the clock fre-

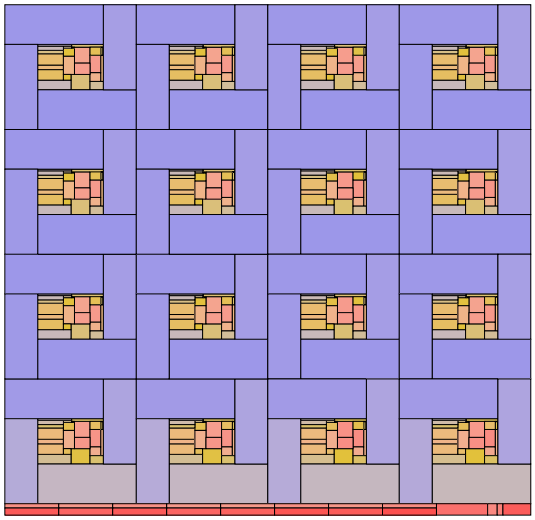

Figure 6. Checkers floorplan. Each core is surrounded by four unified L2 cache banks. The front-end is along the bottom edge of the layout.

quency. The maximum temperature is determined by executing the workload on a checkers layout (see Fig. 6) without an optical front-end communication network for a mix of benchmarks. To obtain the frequency for a grid layout (see Fig. 5) with the same maximum temperature, three different clock frequencies are simulated and interpolated. (In the region of interest, the temperature is approximately linear with clock frequency.)

To measure the effect of the impact on performance by spreading out the processing cores, the IPC performance of a microarchitecture with optical links between the front-end and back-ends is compared with a system with only electrical interconnects. In future work, the use of optical interconnects to reduce long distance inter-back-end communication latencies will also be investigated.

4.1. Power Model

Wattch version 1.02 [4] is used to compute the dynamic power of the units. Parameters for the 45 nm technology node are derived from the ITRS Roadmap [28]. The wire resistance and capacitance scaling factors are determined by log-log extrapolation from the technology nodes supplied with Wattch. Similarly, the sense voltage factor is determined by linear extrapolation from earlier technology nodes.

A simple temperature-dependent computation of leakage power is applied. Gate oxide leakage is assumed to not be significant (as a result of the adoption of a high-k dielectric technology) [18]. Therefore, only subthreshold leakage is considered. The units are divided into logic and SRAM groups, due to differences in ITRS predictions [28] for these two groups. The power is determined from the ITRS-predicted transistor density, static power per transistor width, and several additional assumptions: an average W/L of 3 for the SRAM circuitry and 3.6 for the logic, each PMOS transistor leaks twice as much as an NMOS transistor, and the NMOS and PMOS transistors are each on 50% of the time. The ITRS value for leakage power at room temperature provides a reference, and the BSIM3 model [33] is used to correlate leakage power with temperature. Equation (1) is used to adjust the leakage power of each unit based on the temperature of that individual unit, continually recalculated as the temperature changes.

$$P = \frac{P_{static}(W/L)L_{gate}Q_{density} * area}{T^2_{ITRS}} T^2 \; Watts \tag{1}$$

where

$$P_{static} = \frac{\tau_{N,leak}P_{N,static} + 2\tau_{P,leak}P_{N,static}}{2} \tag{2}$$

Equation (2) is given in terms of watts per meter of the transistor gate width with $\tau_{leak,N}$ and $\tau_{leak,P}$ referring to the fraction of the time that N and P transistors, respectively, dissipate leakage (rather than dynamic) power. L_{gate} is the printed length of the gate, $Q_{density}$ is the density of transistors and *area* is the actual die area of the device (box). T refers to the absolute temperature of the unit and is a function of time.

4.2. Temperature Model

Chip temperatures are derived from the power numbers using the HotSpot (version 2) [16] simulation tool. HotSpot determines the transient temperatures, so maximum transient temperatures are used. (Steady-state temperatures are not used because potential short-period hot spots are ignored.)

The HotSpot parameters are listed in Table 3. High end cooling technologies are assumed, since cooling will be more important in future processors. For the heat sink, the resistance of a "folded-fin" heat sink is used [20], as well as a thermal interface material with a resistivity of 0.14 mK/W [1] and a thickness of 30 μm. This thickness is about half of the coverage thickness used as a default in HotSpot or assumed by the Arctic Silver specifications [1]. Since the thermal interface material may play an important role in dissipating heat from the hot spots, it is assumed that by 2010 the thickness

will be reduced from the current 70 μm. Parameters not explicitly listed are the same as the default values specified in the HotSpot software.

Table 3. HotSpot parameters.

Heat Sink	
Convection resistance	0.02 K/W
Convection capacitance	140.4 J/K
Thermal Interface Material	
Thickness	30 μm
Thermal resistivity	0.14 mK/W

4.3. Benchmarks

Two classes of workloads are considered, mixes of SPEC2000 CPU benchmarks (groupA) and SPLASH-2 benchmarks operating in multi-threaded mode (groupB). Using the same classification system as [13], two communication bound workloads and an instruction level parallelism (ILP) bound workload are examined. The mixes are listed in Table 4.

GroupA benchmarks are mixes of independent threads. These benchmarks do not share virtual memory address space and therefore there is no inter-thread communication. Each SPEC benchmark in this group is run with the reference input set. The benchmarks are individually fast forwarded as suggested in [27], and run simultaneously until each thread reaches 100 million instructions. The geometric mean of the speedup of all of the threads is used as the performance metric.

GroupB benchmarks are parallel programs from the SPLASH-2 benchmark suite [34]. The relevant parameters are listed in Table 5. The threads share virtual address space and communicate with one another by means of shared memory facilitated by cache coherence. Each benchmark in groupB is run to completion. Speedup is calculated as the ratio of the execution times in cycles.

Each individual thread has exclusive access to two adjacent cores. Prior research has shown that the communication delays involved with additional cores negate any performance gain from the increase in the number of functional units [13, 19].

Table 4. Single-threaded mixes.

Load	Benchmarks included	Bound
Mix 1	bzip, parser, art, galgel	communication
Mix 2	bzip, vpr, gzip, parser, perlbmk, lucas, art, galgel	communication
Mix 3	gcc, mcf, twolf, applu, mgrid, swim, equake, mesa	ILP

Table 5. Parallel programs.

Program	Command Line Arguments
FFT	-m18 -p8 -n1024 -l6 -t
Jacobi	-p8 -v -s512 -i10
LU	-n512 -p8 -b16 -t
Radix	-p8 -n131072 -r16 -m524288 -t

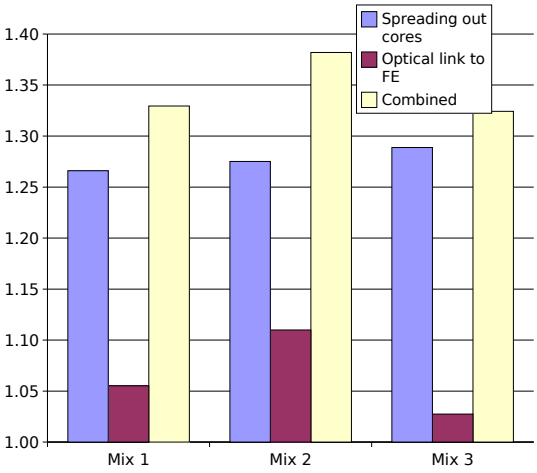

Figure 7. Speedup resulting for GroupA.

5. Results

The results are relative to a benchmark run with a grid layout (see Fig. 5) with no optical communication lines. Mixes of independent threads are first presented followed by parallel programs.

5.1. GroupA

The left bars in each group shown in Figure 7 quantify the change in the clock frequency (and therefore the performance) achieved by using the spread out checkers layout (Fig. 6). The middle bars include the optical communication lines from the shared front-end to each of the cores. The direct communication lines allow for faster dispatch of instructions to the cores and a shorter branch mispredict penalty (the recovery is started earlier). This modest application of optical interconnect leads to an increase in performance of up to 10% for multi-threaded workloads of independent applications.

The right bars combine the two techniques. The average speedup for these benchmark mixes is 35% with a maximum of 38%. The two enhancements are not completely orthogonal. The faster communication with the front-end leads to enhanced utilization of the functional units which in turn increases the baseline temperature. The increase in clock speed is therefore partly reduced.

5.2. GroupB

The multi-threaded benchmarks produce greater improvements. The left bars shown in

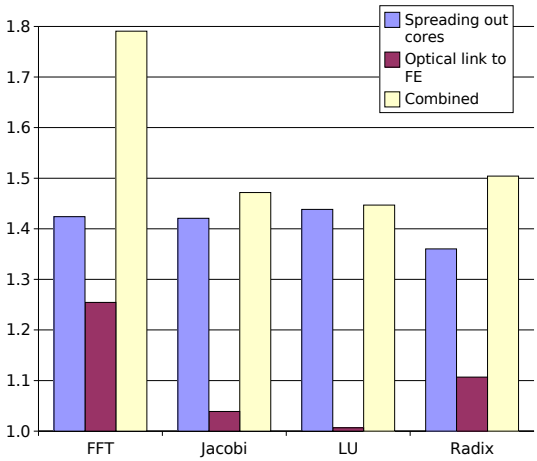

Figure 8. Speedup resulting for GroupB.

Figure 8 describe improvements from spreading out the cores. The speedup is roughly 40% across all of the benchmarks.

The middle bars are obtained by adding the high speed optical links from the front-end to each core. The improvement varies depending on the nature of each benchmark, but reaches 25% for FFT.

The right bars present the results of combining the two techniques. The average speedup for these benchmarks is 55% with a maximum of 78%.

6. Related Work

Modeling the effects of leakage current on power dissipation and temperature at the architectural level was first studied by Butts and Sohi [30] and later by Zhang et al. [35], the former based on the BSIM3 transistor leakage model [33].

Others have investigated dynamic temperature management schemes, such as frequency, voltage, and fetch rate control [29], software scheduling behavior [26], asymmetric dual core designs [14], and a combination of these techniques [15].

Additional researchers have also considered the impact of circuit layout on temperature, such as Cheng and Kang with their iTAS simulator [9]. Investigations have also been promoted by other VLSI-based simulation research, such as Rencz et al. [25], the SISSI package [31], and others [32].

Donald and Martonosi investigated thermal issues in SMT and CMP architectures [12], although these authors only consider steady-state temperatures and do not translate the temperature results into the effect on application performance.

In contrast to these previous research results, this work is the first to investigate the use of on-chip optical interconnects to reduce the performance gap created by increasing the physical distances between the front and back ends of the processor in order to alleviate thermal constraints.

7. Conclusions

With recent advances in silicon photonics, on-chip optical interconnects have become a prime candidate to alleviate a number of global communication challenges in future highly integrated microprocessors. In this paper, the use of optical interconnects to ameliorate the increased global wire delay due to intermixing hot and cold processing units is investigated. It is shown that the selective introduction of a few optical connections can significantly enhance overall processor performance. This study has also shown that intermingling the cluster cores with the on-chip cache reduces the maximum on-chip temperature. Since the maximum temperature limits the clock speed, spreading the cores can lead to increased clock frequencies. This technique does not reduce overall power dissipation (other than the decreased leakage current due to lower on-chip temperatures) but more uniformly redistributes the dissipated power. The use of optical interconnect for long distance communication makes spreading the cores a more viable proposition in terms of maintaining high performance levels.

In future work, the use of optical interconnect will be investigated to reduce inter-back-end communication for parallel workloads, increase link bandwidth through the use of Wave Division Multiplexing (WDM), and reduce the worst case latencies of large cache and main memory RAMs.

References

[1] Arctic Silver Incorporated. The Arctic Silver 5 Specifications. http://www.arcticsilver.com/as5.htm, 2004.

[2] C. A. Barrios, V. R. d. Almeida, and M. Lipson. Electrooptic modulation of silicon-on-insulator submicrometer-size waveguide devices. *Journal of Lightwave Technology*, 21(10):2332, Oct. 2003.

[3] C. A. Barrios, V. R. d. Almeida, and M. Lipson. Compact silicon tunable fabry-perot resonator with low power consumption. *IEEE Photonics Technology Letters*, 16(2):506, Feb. 2004.

[4] D. Brooks, M. Martonosi, and V. Tiwari. Wattch: A framework for architectural-level power analysis and optimizations. In *Proceedings of the 27th Annual International Symposium on Computer Architecture*, pages 83 – 94, June 2000.

[5] D. Burger and T. Austin. The simplescalar toolset, version 2.0. Technical Report TR-97–1342, University of Wisconsin-Madison, June 1997.

[6] R. T. Chang, N. Talwalkar, C. P. Yue, and S. S. Wong. Near speed-of-light signaling over on-chip electrical interconnects. *IEEE Journal of Solid-State Circuits*, 38(5):834–838, May 2003.

[7] G. Chen, H. Chen, M. Haurylau, N. Nelson, D. H. Albonesi, P. M. Fauchet, and E. G. Friedman. Electrical and optical on-chip interconnects in future microprocessors. In *IEEE International Symposium on Circuits and Systems*, May 2005.

[8] G. Chen, H. Chen, M. Haurylau, N. Nelson, P. M. Fauchet, E. G. Friedman, and D. H. Albonesi. Predictions of CMOS compatible on-chip optical interconnect. In *Proceedings of the IEEE/ACM International Workshop on System Level Interconnect Prediction*, Apr. 2005.

[9] Y.-K. Cheng and S.-M. Kang. A temperature-aware simulation environment for reliable ULSI chip design. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 19(10):1211–1220, Oct. 2000.

[10] B. S. Cherkauer and E. G. Friedman. A unified design methodology for CMOS tapered buffers. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 3(1):99–111, Mar. 1995.

[11] Z. Chishti, M. D. Powell, and T. N. Vijaykumar. Distance associativity for high-performance energy-efficient non-uniform cache architectures. In *Proceedings of the 36th International Symposium on Microarchitecture*, pages 55–66, Dec. 2003.

[12] J. Donald and M. Martonosi. Temperature-aware design issues for SMT and CMP architectures. In *Fifth Workshop on Complexity-Effective Design*, 2004 June.

[13] A. El-Moursy, R. Garg, S. Dwarkadas, and D. H. Albonesi. Partitioning multi-threaded processors with large number of threads. In *IEEE International Symposium on Performance Analysis of Systems and Software*, Austin, Texas, Mar. 2005.

[14] S. Ghiasi and D. Grunwald. Thermal management with asymmetric dual core designs. Technical Report CU-CS-965-03, Department of Computer Science, University of Colorado, 2003.

[15] M. Huang, J. Renau, S. Yoo, and J. Torrellas. The Design of DEETM: A Framework for Dynamic Energy Efficiency and Temperature Management. *Journal of Instruction-Level Parallelism*, 3, Oct. 2001.

[16] W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan, S. Ghosh, and S. Velusamy. Compact thermal modeling for temperature-aware design. In *Proceedings of the 41st IEEE/ACM Design Automation Conference*, June 2004.

[17] R. E. Kessler. The Alpha 21264 microprocessor. *IEEE Micro*, pages 24–36, Mar./Apr. 1999.

[18] N. S. Kim, T. Austin, D. Blaauw, T. Mudge, K. Flautner, J. S. Hu, M. J. Irwin, M. Kandemir, and V. Narayanan. Leakage current: Moore's law meets static power. *IEEE Computer*, 36(12):68–75, Dec. 2003.

[19] F. Latorre, J. González, and A. González. Back-end assignment schemes for clustered multithreaded processors. In *Proceedings of the 18th Annual ACM International Conference on Supercomputing*, pages 316–325, June 2004.

[20] S. Lee. How to select a heat sink. *Electronics Cooling*, 1(1), June 1995.

[21] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia. A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor. *Nature*, 427:615–618, Feb. 2004.

[22] R. Lytel, H. L. Davidson, N. Nettleton, and T. Sze. Optical interconnections within modern high-performance computing systems. *Proceedings of the IEEE*, 88(6):758–763, June 2000.

[23] D. A. B. Miller. Rationale and challenges for optical interconnects to electronic chips. *Proceedings of the IEEE*, 88(6):728–749, June 2000.

[24] J. Oh, J. Campbell, S. G. Thomas, S. Bharatan, R. Thoma, C. Jasper, R. E. Jones, and T. E. Zirkle. Interdigitated Ge p-i-n photodetectors fabricated on a Si substrate using graded SiGe buffer layers. *IEEE Journal of Quantum Electronics*, 38(9):1238–1241, Sept. 2002.

[25] M. Rencz, V. Szekely, A. Poppe, and B. Courtois. Friendly tools for the thermal simulation of power packages. *International Workshop on Integrated Power Packaging, 2000*, pages 51–54.

[26] E. Rohou and M. D. Smith. Dynamically managing processor temperature and power. In *Proceedings of the 2nd Workshop on Feedback-Directed Optimization*, Nov. 1999.

[27] S. Sair and M. Charney. Memory behavior of the SPEC2000 benchmark suite. Technical report, IBM T. J. Watson Research Center, Oct. 2000.

[28] Semiconductor Industry Association. *The International Technology Roadmap for Semiconductors*. 2003.

[29] K. Skadron, M. R. Stan, K. Sankaranarayanan, W. Huang, S. Velusamy, and D. Tarjan. Temperature-aware microarchitecture: Modeling and implementation. *ACM Transactions on Architecture and Code Optimization*, 1(1):94 – 125, Mar. 2004.

[30] G. S. Sohi and J. A. Butts. A static power model for architects. In *Proceedings of the 33rd annual ACM/IEEE International Symposium on Microarchitecture*, pages 191–201, Dec. 2000.

[31] V. Szekely, A. Poppe, A. Pahi, A. Csendes, G. Hajas, and M. Rencz. Electro-thermal and logi-thermal simulation of VLSI designs. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 5(3):258–269, Sept. 1997.

[32] K. Torki and F. Ciontu. IC thermal map from digital and thermal simulations. In *Proceedings of the 2002 International Workshop in THERMal Investigations of ICs and Systems*, pages 303–308, Oct. 2002.

[33] University of California, Berkeley. *BSIM3v3.2.2 Manual*, 1999.

[34] S. Woo, M. Ohara, E. Torrie, J. Singh, and A. Gupta. The SPLASH-2 Programs: Characterization and Methodological Considerations. In *International Symposium on Computer Architecture*, pages 24–36, Santa Margherita Ligure, Italy, June 1995.

[35] Y. Zhang, D. Parikh, K. Sankaranarayanan, K. Skadron, and M. R. Stan. Hotleakage: A temperature-aware model of subthreshold and gate leakage for architects. Technical Report CS-2003-05, Department of Computer Science, University of Virginia, Mar. 2003.