Computing With an All-Optical Cache Hierarchy Using Optical Phase Change Memory as Last Level Cache

Haiyang Han(1), Theoni Alexoudi(2), Chris Vagionas(2), Nikos Pleros(2), Nikos Hardavellas(3),

(1) Department of ECE, Northwestern University, Evanston, IL, USA haiyang.han@u.northwestern.edu
(2) Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece
(3) Department of CS & ECE, Northwestern University, Evanston, IL, USA

Abstract *We discuss the architecture of an all-optical cache hierarchy that extends existing optical cache designs with an optical PCM LLC. We design and analyze methods to mitigate PCM's slow write speed and limited lifetime for 20% execution time reduction and non-volatility.* ©*2022 The Author(s)*

Introduction

Recent experimental demonstrations of optical SRAM cells[1],[2] have stimulated a new research landscape towards overcoming the "memory wall" problem of modern computers, aiming to transfer the low latency, high bandwidth, and high energy efficiency credentials of optical technology into the memory domain. Optical memory layouts have started to migrate from simple memory cells to more complex memory architectures, demonstrating experimental optical RAM rows[3] and multi-bit optical RAM banks up to the recently reported first optical cache memory prototype[4]. All these indicate that the time is ripe to proceed towards an all-optical cache hierarchy to utilize the benefits of light-enabled caching. Within this frame, Pho$ has been a recently proposed opto-electronic memory architecture[5],[6] where a fast, large, and shared optical L1 cache with an optically connected main memory and a novel optical network-on-chip were successfully incorporated, demonstrating significant performance and energy benefits over electronic processors, although its last level cache (LLC) remains electronic. The use of an electronic LLC means that L1–LLC and LLC–DRAM traffic have to go through unnecessary OE/EO conversions, consuming additional latency and energy, while requiring almost 19.8% of the total chip area[6].

The employment of an all-optical cache hierarchy where the LLC relies also on the use of optical memory technologies can eliminate much of these complexities and inefficiencies. However, the simple approach of directly using optical SRAM bitcells[1],[2] in the LLC is impractical since InP photonic crystal-based memories, which are utilized as SRAM cells due to their high-speed read/write functionality, are volatile and consume a rather large area[7]. Small-footprint requirements together with non-volatile properties that minimize idle-time energy consumption can be offered only through optical Phase Change Memories (O-PCM)[8],[9], which are heavily researched towards neuromorphic photonic computing[10]. O-PCMs necessitate, however, rather long writing times, implying that the successful transfer of their size and non-volatility advantages into the LLC domain can be only realized if their slow writes can be absorbed by higher level caches within a properly architected cache hierarchy.

In this paper, we extend the Pho$ design[5] and replace its electronic LLC with an O-PCM LLC for all-optical communication between the processor and memory. We use write queues and a "no allocate" write policy to alleviate the slow write speed and low write endurance of PCM. We determine the optimal queue size and show that an 8-entry queue between the L1 cache and LLC can achieve a 20% reduction in execution time over the electronic baseline (2% reduction over Pho$). We also demonstrate that the write policy improves the average lifetime of the O-PCM LLC in a two-level optical cache hierarchy by 13× compared with just one level of O-PCM cache. To the best of our knowledge, we are the first to propose and analyze the design of an all-optical cache hierarchy by leveraging PhC SRAM cells[1], O-PCM cells[8],[9], and on-chip optical interconnects.

System Architecture

Figure 1a illustrates the architecture of the proposed design. The processor cores, 2.5D-stacked optical L1 cache banks, and O-PCM LLC all sit on the interposer with photonic links. The shared L1 optical cache is built using PhC SRAM cells[1]. The LLC is built using O-PCM cells[8]. The off-chip laser sources power the optical network, PhC L1 dies, and O-PCM LLC die. The entire cache hierarchy is in the optical domain.

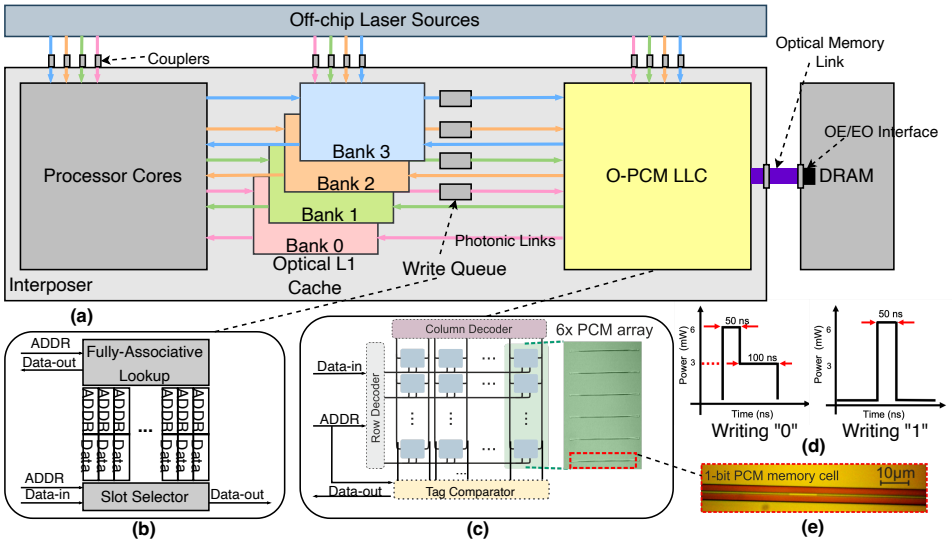

Fig. 1: System architecture: (a) PhC + O-PCM cache hierarchy (b) structure of a write queue (c) architecture of an O-PCM cache bank (d) latency and power of writing an O-PCM cell (e) photo of verified O-PCM cell.

Writing a logical "0" to an O-PCM cell requires 150 ns and 600 pJ. Writing a logical "1" requires 50 ns and 300 pJ. Figure 1d shows the latencies and power requirements for O-PCM write operations. The latency of a read operation is only attributed to the latency of light propagation through the waveguide, or time-of-flight. There is also no extra power required other than the Tx and Rx powers for the optical interconnects. Because the O-PCM material is directly integrated on an optical waveguide, the on-chip photonic links can directly access the O-PCM cache.

Tab. 1: Simulated system parameters.

Component	Details
Cores	16 cores, x86 ISA, 3.2 GHz, OoO, 4 wide dispatch/commit, 224-entry ROB, 72-entry load queue, 56-entry store queue
L1 ICache	**Baseline**: electronic, private, 64 B line,32 kB/core, 8-way, 4 cycles **One Level**: N/A **Pho$/Pho$OPCM**: optical, shared, 64 B line, 1 MB direct-mapped, 2-cycle read, 23-cycle write
L1 DCache	**Baseline**: electronic, private, 64 B line, 32 kB/core, 8-way, 4 cycles **One Level**: N/A **Pho$/Pho$OPCM**: optical, shared, 4 banks, 64 B line 4 MB direct-mapped, 2-cycle read, 23-cycle write
L2	**Baseline**: electronic, private, 64 B line, 256 kB/core, 4-way, 14 cycles **One Level/Pho$/Pho$OPCM**: N/A
LLC	**Baseline/Pho$**: electronic, shared, non-inclusive, 64 B line, 32 MB 16-way, 50 cycles **One Level/Pho$OPCM**: optical-PCM, shared, exclusive, 64 B line 32 MB, direct-mapped, 1-cycle read, 480-cycle write
Write Queue	**Baseline/Pho$**: N/A **One Level/Pho$OPCM**: 1 queue per L1D bank Sizes: 0, 4, 8, 16, 32, 64, 128, 256, 512, 1024 entries Latencies: 0, 1, 1, 1, 2, 2, 2, 2, 2, 3 cycles

Write Queue and O-PCM Cache Bank

To partially hide PCM's slow writes, all write operations to the LLC go through a fully-associative FIFO write queue (Figure 1b). Each entry in the write queue contains the address and data of an in-flight write to the LLC. When the L1 cache initiates a write, it writes to the next empty slot in the write queue. Instead of waiting for the block to be written, which incurs 150 ns, the L1 cache can return to its next operations immediately. In the background, the write queue is constantly writing the in-flight cache lines, in FIFO order to the O-PCM LLC. Only when the write queue is full will the L1 be forced to stall for a maximum wait time of 150 ns. The slot selector keeps two pointers: the next available entry for an incoming write and the next entry to be drained via the photonic link to the LLC. Cache misses in the L1 will check the write queue for a matching address first before the request is forwarded to the LLC.

Figure 1c shows the architecture of an O-PCM cache bank. The optical row and column decoders[11] interpret the incoming address and activate the corresponding O-PCM cell (photo in Figure 1e). The tag comparator[12] determines cache hit/miss status. We expect O-PCM caches have a larger capacity per unit area compared to electronic caches due to their high density. Our design uses a 32 MB O-PCM LLC with single-cycle read latency and 150 ns write latency. The read latency of the O-PCM cache is determined by the time-of-flight of the optical signal passing through the cache, and is related to the physical dimensions of the cache. The novelty of the technology, however, means that architecting the exact layout of the cells to model the area is left for future work. All the optical components (row/column decoders[11], tag comparator[12], O-PCM cell[8]) and an all-optical cache prototype[4] have been experimentally verified and demonstrated.

"No Allocate" Policy

To decrease the frequency of writing to the O-PCM LLC (for higher performance and longer lifetime), we propose a "no allocate" policy. We only allocate in the O-PCM LLC on L1 evictions (clean

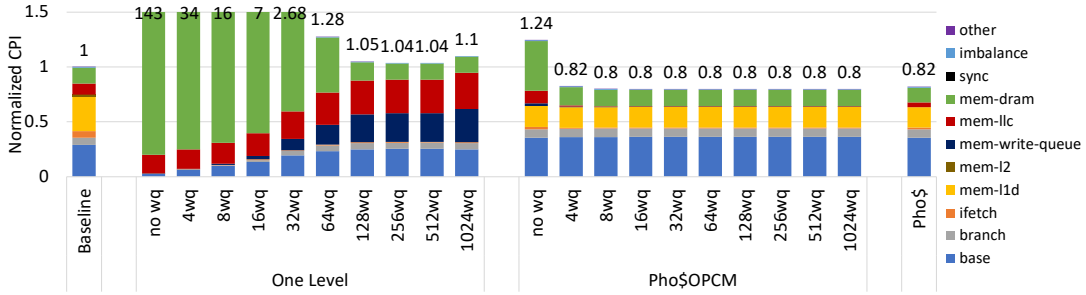

Fig. 2: Average CPU2017 CPI Stacks normalized to *baseline* (electronic multicore).

and dirty), which will go through the write queue first. All hits in the LLC will have the cache line moved to the L1. We never allocate in the O-PCM LLC if both L1 and the LLC miss to avoid write-after-write scenarios when the cache line is soon evicted from L1 and the LLC is charged with two writes. The policy can also be tuned for only one level of O-PCM cache between the processors and DRAM. In this case, the O-PCM cache also only contains clean lines. Writing in the O-PCM cache is avoided and the block is moved to be written in the DRAM for faster writes. Cache writes are only performed during a read miss.

Experimental Methodology and Results

We sweep through write queue sizes of 0–1024 to find the optimal size for "Pho$OPCM" and "One Level" of O-PCM. A larger write queue buffers more entries concurrently but suffers from higher access latency from fully-associative lookups. We modified the Sniper simulator[13],[14] extensively to model asymmetrical read/write latencies, the write queue, and the "no allocate" write policy. The configurations are compared against a baseline electronic multicore (16-core Intel Skylake) and Pho$[5] with an electronic LLC (Table 1).

Figure 2 shows the average CPI stacks[15] of the baseline, One Level and Pho$OPCM with various write queue sizes, and Pho$ running SPEC CPU2017[16]. For One Level, the O-PCM cache is under heavy traffic. Without a write queue, One Level's CPI is 143× that of baseline. As the size of the write queue increases, One Level's CPI decreases quickly. For write queue sizes of 256 and 512, we hit an optimal performance at 1.04× slower than baseline. Nonetheless, One Level is incapable of performing better than the baseline, no matter the queue size. It seems imperative that we add an L1 cache before the O-PCM cache. For Pho$OPCM, the optical shared L1 cache absorbs the majority of the traffic to the O-PCM LLC. An 8-entry write queue results in a 20% CPI reduction over baseline and 2% over Pho$. We can determine that from a performance aspect,

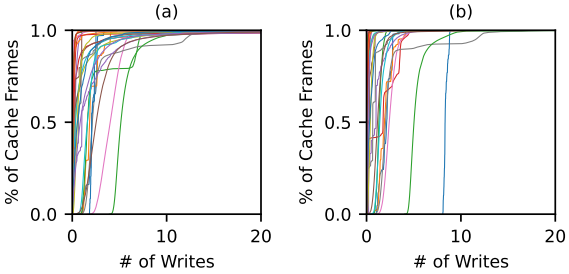

Fig. 3: Cache Frames Writes CDF: (a) One Level (b) Pho$OPCM.

the optimal write queue sizes for One Level and Pho$OPCM are 256 and 8, respectively.

Figure 3 shows the CDF of the percentage of physical cache frames and their total number of writes running CPU2017 for One Level and Pho$OPCM. For most benchmarks, Pho$OPCM shows a steeper plot and the CDF reaches 100% more quickly. This means that physical cache frames experience fewer writes for each run, indicating an overall longer lifetime for the O-PCM cache. The only exception is *lbm* where a stream of writes results in evictions of L1 and allocations in the LLC for Pho$OPCM but only DRAM writes for One Level. Our results show that when employing the "no allocate" policy, Pho$OPCM has a significantly longer average lifetime than One Level (13×).

Conclusions

We extended existing optical cache designs with an O-PCM LLC for an all-optical cache hierarchy. We studied the optimal write queue sizes and the effects of a "no allocate" write policy to increase PCM lifetime. By employing an 8-entry write queue and the write policy, we achieve similar performance to Pho$ despite O-PCM's long write latency and 13× cache lifetime vs One Level, while providing LLC-level non-volatility.

Acknowledgements

This work was partially funded by NSF award CCF-1453853, and HFRI and GSRT through the ORION (grant 585) and CAM-UP (grant 230) projects.

References

[1] K. Nozaki, A. Shinya, S. Matsuo, Y. Suzaki, T. Segawa, T. Sato, Y. Kawaguchi, R. Takahashi, and M. Notomi, "Ultralow-power all-optical RAM based on nanocavities", *Nature Photonics*, vol. 6, no. 4, pp. 248–252, 2012. DOI: 10.1038/nphoton.2012.2.

[2] T. Alexoudi, D. Fitsios, A. Bazin, P. Monnier, R. Raj, A. Miliou, G. T. Kanellos, N. Pleros, and F. Raineri, "III-V-on-Si photonic crystal nanocavity laser technology for optical static random access memories", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 22, no. 6, pp. 295–304, 2016. DOI: 10.1109/JSTQE.2016.2593636.

[3] T. Alexoudi, C. Pappas, T. Moschos, K. Fotiadis, G. Mourgias-Alexandris, N. Pleros, and C. Vagionas, "Optical RAM row with 20 gb/s optical word read/write", *Journal of Lightwave Technology*, vol. 39, no. 22, pp. 7061–7069, 2021. DOI: 10.1109/JLT.2021.3112913.

[4] C. Pappas, T. Moschos, T. Alexoudi, C. Vagionas, and N. Pleros, "Caching with light: First demonstration of an optical cache memory prototype", in *Optical Fiber Communication Conference (OFC) 2022*, Optica Publishing Group, 2022, Th4B.3. DOI: 10.1364/OFC.2022.Th4B.3.

[5] H. Han, T. Alexoudi, C. Vagionas, N. Pleros, and N. Hardavellas, "Pho$: A case for shared optical cache hierarchies", in *IEEE/ACM International Symposium on Low Power Electronics and Design, ISLPED 2021, Boston, MA, USA, July 26-28, 2021*, IEEE, 2021, pp. 1–6. DOI: 10.1109/ISLPED52811.2021.9502487.

[6] ——, "A practical shared optical cache with hybrid mwsr/r-swmr noc for multicore processors", *Journal on Emerging Technologies in Computing Systems*, Apr. 2022, ISSN: 1550-4832. DOI: 10.1145/3531012.

[7] T. Alexoudi, G. T. Kanellos, and N. Pleros, "Optical RAM and integrated optical memories: A survey", *Light: Science & Applications*, vol. 9, no. 1, pp. 1–16, 2020. DOI: 10.1038/s41377-020-0325-9.

[8] A. Manolis, J. Faneca, T. D. Bucio, A. Baldycheva, A. Miliou, F. Gardes, N. Pleros, and C. Vagionas, "Non-volatile integrated photonic memory using gst phase change material on a fully etched si3n4/sio2 waveguide", in *Conference on Lasers and Electro-Optics: Science and Innovations*, Optical Society of America, 2020, STh3R–4. DOI: 10.1364/CLEO_SI.2020.STh3R.4.

[9] C. Ríos, M. Stegmaier, P. Hosseini, D. Wang, T. Scherer, C. D. Wright, H. Bhaskaran, and W. H. Pernice, "Integrated all-photonic non-volatile multi-level memory", *Nature photonics*, vol. 9, no. 11, pp. 725–732, 2015. DOI: 10.1038/nphoton.2015.182.

[10] M. Miscuglio and V. J. Sorger, "Photonic tensor cores for machine learning", *Applied Physics Reviews*, vol. 7, no. 3, p. 031 404, 2020. DOI: 10.1063/5.0001942.

[11] T. Alexoudi, S. Papaioannou, G. T. Kanellos, A. Miliou, and N. Pleros, "Optical cache memory peripheral circuitry: Row and column address selectors for optical static RAM banks", *IEEE Journal of Lightwave Technology*, vol. 31, no. 24, pp. 4098–4110, 2013. DOI: 10.1109/JLT.2013.2286529.

[12] C. Vagionas, S. Pitris, C. Mitsolidou, J. Bos, P. Maniotis, D. Tsiokos, and N. Pleros, "All-optical tag comparison for hit/miss decision in optical cache memories", *IEEE Photonics Technology Letters*, vol. 28, no. 7, pp. 713–716, 2015. DOI: 10.1109/LPT.2015.2505500.

[13] T. E. Carlson, W. Heirman, and L. Eeckhout, "Sniper: Exploring the level of abstraction for scalable and accurate parallel multi-core simulation", in *Conference on High Performance Computing Networking, Storage and Analysis, SC 2011, Seattle, WA, USA, November 12-18, 2011*, ACM, 2011, 52:1–52:12. DOI: 10.1145/2063384.2063454.

[14] T. E. Carlson, W. Heirman, S. Eyerman, I. Hur, and L. Eeckhout, "An evaluation of high-level mechanistic core models", *ACM Transactions on Architecture and Code Optimization*, vol. 11, no. 3, 28:1–28:25, 2014. DOI: 10.1145/2629677.

[15] W. Heirman, T. E. Carlson, S. Che, K. Skadron, and L. Eeckhout, "Using cycle stacks to understand scaling bottlenecks in multi-threaded workloads", in *2011 IEEE International Symposium on Workload Characterization, IISWC 2011, Austin, TX, USA, November 6-8, 2011*, IEEE Computer Society, 2011, pp. 38–49. DOI: 10.1109/IISWC.2011.6114195.

[16] J. Bucek, K. Lange, and J. von Kistowski, "SPEC CPU2017: next-generation compute benchmark", in *Companion of the 2018 ACM/SPEC International Conference on Performance Engineering, ICPE 2018, Berlin, Germany, April 09-13, 2018*, ACM, 2018, pp. 41–42. DOI: 10.1145/3185768.3185771.