A Practical Shared Optical Cache With Hybrid MWSR/R-SWMR NoC for Multicore Processors

HAIYANG HAN, Northwestern University, USA
THEONI ALEXOUDI, Aristotle University of Thessaloniki, Greece
CHRIS VAGIONAS, Aristotle University of Thessaloniki, Greece
NIKOS PLEROS, Aristotle University of Thessaloniki, Greece
NIKOS HARDAVELLAS, Northwestern University, USA

Conventional electronic memory hierarchies are intrinsically limited in their ability to overcome the memory wall due to scaling constraints. Optical caches and interconnects can mitigate these constraints, and enable processors to reach performance and energy efficiency unattainable by purely electronic means. However, the promised benefits cannot be realized through a simple replacement process; to reach its full potential, the architecture needs to be holistically redesigned. This article proposes Pho$, an opto-electronic memory hierarchy architecture for multicores. Pho$ replaces conventional core-private electronic caches with a large shared optical L1 built with optical SRAMs. The shared optical cache is supported by Pho$Net, a novel hybrid MWSR/R-SWMR optical NoC that provides low-latency and high-bandwidth communication between the electronic cores and the shared optical L1 at low optical loss. Pho$Net's unique network arbitration protocol seamlessly co-arbitrates the request and reply sub-networks and facilitates cache requests and replies that optimize for the common case of cache hits. Through Pho$ we solve the problems that render previous designs impractical. Our results show that Pho$ achieves on average 1.41× performance speedup (3.89× max) and 31% lower energy-delay product (90% max) against conventional designs. Moreover, the Pho$Net optical NoC for core-cache communication consumes 70% less power compared to directly applying previously proposed optical NoC architectures.

CCS Concepts: • **Computer systems organization → Architectures**; • **Hardware → Emerging optical and photonic technologies**; **Memory and dense storage**; **Photonic and optical interconnect**.

Additional Key Words and Phrases: nanophotonic interconnects, cache hierarchy, network-on-chip, energy efficiency, optical caches

ACM Reference Format:
Haiyang Han, Theoni Alexoudi, Chris Vagionas, Nikos Pleros, and Nikos Hardavellas. 2022. A Practical Shared Optical Cache With Hybrid MWSR/R-SWMR NoC for Multicore Processors. *ACM J. Emerg. Technol. Comput. Syst.* 1, 1, Article 1 (January 2022), 28 pages. https://doi.org/10.1145/3531012

Preliminary portions of this article appeared in the Proceedings of the IEEE/ACM International Symposium on Low Power Electronics and Design, ISLPED 2021, Boston, MA, USA, July 26-28, 2021 [34].
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.
Authors' addresses: H. Han, and N. Hardavellas, Northwestern University, 2233 Tech Dr, Evanston, IL 60208, USA; emails: haiyang.han@u.northwestern.edu, nikos@northwestern.edu; T. Alexoudi, C. Vagionas, and N. Pleros, Aristotle University of Thessaloniki, University Campus 54124, Thessaloniki, Greece; emails: {theonial, chvagion, npleros}@csd.auth.gr.

1550-4832/2022/1-ART1 $15.00
https://doi.org/10.1145/3531012

1 INTRODUCTION

It has been nearly 25 years since the performance gap between CPUs and main memory, or the "Memory Wall", was identified as the main obstacle in increasing the performance of computer systems [88]. To mitigate the memory wall, stemming from the high latency of electronic memories and the limited bandwidth of electronic off-chip memory interconnects, modern chip multiprocessors (CMPs) have resorted to deep cache hierarchies. However, on-chip caches can occupy as much as 40% of the die area [9] and 32% of the processor's power [68]. As a result, multiple efforts on the device and architecture levels have focused on mitigating these issues, including caches based on STT-RAM [30, 37, 72, 75, 87], Phase Change Memory [38, 48, 86], and 3D-die stacking [6, 36, 45–47, 74, 89].

Alternatively, optical interconnects and nanophotonic technologies have emerged as promising yet underdeveloped solutions to tackle the disparity between processor and memory speeds. Today, we appear to have all the ingredients necessary to design novel optical cache architectures supported by optical interconnects. Optical Networks on Chip (NoCs) demonstrate higher bandwidth and energy efficiency than the traditional electronic NoCs used in CMPs [21–23, 25, 33, 41, 52, 58, 59, 81–83]. Optically connected memory (OCM) raises the possibility to switch much of the data transports between the processor and DRAM chips to the optical domain [5, 31, 32]. Silicon photonic IC optical interfaces have been integrated with an electronic IC using a 65 nm DRAM, providing a fast 4× wavelength 10 Gbit/s optical interface for HPC applications [11]. Optical Flip-Flops (FFs) in photonic crystal nanocavities (PhC) [1, 57] can form the building blocks of all-optical memory cells [2], which have demonstrated both speed and energy benefits over their electronic counterparts by boasting read/write speeds up to 40 Gbps [57, 78]. Several optical Flip-Flops (FFs) have been developed with materials like coupled semiconductor optical amplifiers (SOAs) [77], III-V-on-SOI microdisk lasers [43], polarization bistable vertical cavity surface emitting lasers (VCSELs) [67], coupled semiconductor optical amplifier-based Mach-Zehnder Interferometers (SOA-MZIs) [44], and photonic crystal nanocavities (PhC) [1, 57]. The confluence of these technologies seems to be all we need to develop an optical cache hierarchy.

However, the application of an optical cache is not a simple plug-and-play replacement of its conventional electronic counterpart. While prior works [49, 50] have tried to explore this topic, the proposed designs are infeasible for capacities larger than a few kB due to unrealistically high power consumption, and do not consider the challenges of interconnecting the electronic and optical domains. They also lack analysis for whole system power and energy, and their performance is compared against unrealistically weak baselines. In this paper, we address the issues that arise with the introduction of such optical cache devices, and bridge the gap between device- and architecture-level designs. More specifically, our contributions are:

- For the first time to our knowledge, we make optical caches practical. We employ a cascaded two-level row decoder to reduce laser power, active rather than passive components to reduce off-ring optical losses, and use a new technology for the optical bit cells that dramatically lowers the static power consumption.
- We propose Pho$[1], an opto-electronic memory hierarchy for CMPs. Pho$ replaces all the core-private levels of a conventional electronic cache hierarchy with a single-level shared L1 optical cache (split I/D) that utilizes PhC-based optical memory cells [57] operating at 20 GHz. Pho$ enables for the first time L1 caches to be **high capacity** (multiple MB), **fast** (2-processor-cycle access time at 3.2 GHz), and **shared** (obviating cache coherence).

[1]Pronounced "*phos*", a word play between the Greek "$\phi\omega\varsigma$", meaning light, the word "photonic", and "$", the symbol often used to denote a cache.

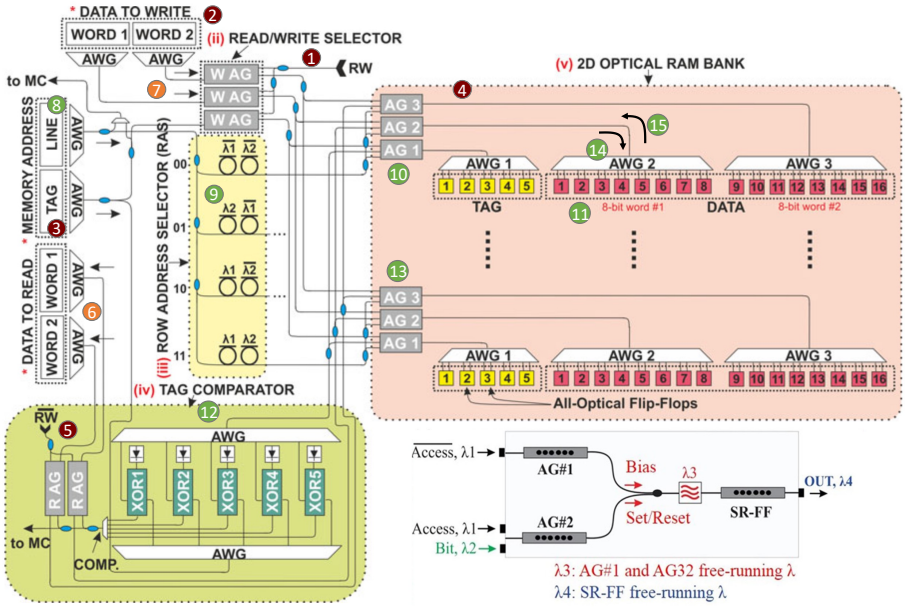


Fig. 1. 8 B optical cache [49] and PhC nanocavity optical SRAM cell [1].

- We propose Pho$Net, a novel hybrid MWSR/R-SWMR optical NoC to connect processor cores with optical cache banks in Pho$. Pho$Net disaggregates the request/reply paths to reduce laser power, and co-arbitrates both subnets simultaneously through a novel arbitration protocol. The optical network extends to the electronic LLC and main memory.
- We perform comprehensive modeling and evaluation of Pho$'s performance, power, and energy characteristics. Pho$ is up to 3.89× faster (1.41× on average) over a traditional electronic cache hierarchy, while achieving up to 90% lower energy-delay product (31% on average). Under realistic assumptions, the Pho$Net optical NoC achieves up to 70% power savings compared to directly applying previously available optical NoC architectures.

In the following section, we provide the necessary background on optical NoCs and the optical cache technologies on which Pho$ is based.

2 BACKGROUND

2.1 Optical Cache Operation

Figure 1 shows the layout of an 8 B direct-mapped optical cache with a 2 B cache line, 2-bit index, and 5-bit tags [49]. Each bit is encoded with two wavelengths.

Read/write operations are controlled by the RW and $\overline{RW}$ signals. During a write to the cache, a RW signal representing a logical "0" activates the Write Access Gates (WAG) ❶ and allows the incoming $data$ bits ❷, the tag bits ❸, and their complements $\overline{data}$ and $\overline{tag}$ to enter the optical RAM bank ❹. At the same time, $\overline{RW}$ represents a logical "1", blocking the Read Access Gates (RAG) ❺ and preventing a read operation. In the case of a read, the RW and $\overline{RW}$ signals are set to logical "1" and "0", respectively. This allows the data from the RAM bank to propagate onto the data reply channel ❻ and blocks the WAG to prevent any data from being overwritten ❼.

The cache line to read or write is designated by the incoming $index$ ❽ and $\overline{index}$ bits which drive the passive Row Address Selector (RAS) ❾. In Figure 1's example, the RAS consists of 4 rows of two

Table 1. Index-RAS truth table.

λ_1	$\overline{\lambda_1}$	λ_2	$\overline{\lambda_2}$	Row 00	Row 01	Row 10	Row 11
0	0	1	1	**0**	$\overline{\lambda_2}$	$\overline{\lambda_1}$	$\overline{\lambda_1\lambda_2}$
0	1	1	0	λ_2	**0**	$\lambda_2\overline{\lambda_1}$	$\overline{\lambda_1}$
1	0	0	1	λ_1	$\lambda_1\overline{\lambda_2}$	**0**	$\overline{\lambda_2}$
1	1	0	0	$\lambda_1\lambda_2$	λ_1	λ_2	**0**

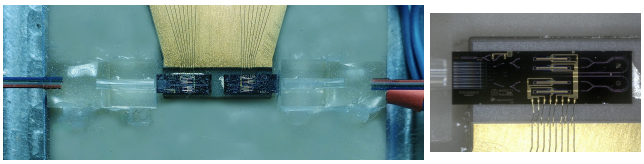


Fig. 2. (a) Packaged optical memory. (b) Monolithic InP Flip-Flop.

micro-rings (MRs) each. Each MR is tuned to a specific wavelength such that a pair of wavelengths λ_i and $\overline{\lambda_i}$ encode the logical "1" and "0" of the i-th bit of the *index*. The 2-bit *index* is encoded with 4 wavelengths: λ_1, $\overline{\lambda_1}$, λ_2, and $\overline{\lambda_2}$. As a result, only one of the four rows will have a logical "0" after the *index* bits pass through the RAS. For example, when the *index* bits are "10", meaning to select the third line, the corresponding logical values for the wavelengths are: $\lambda_1 = 1$, $\overline{\lambda_1} = 0$, $\lambda_2 = 0$, and $\overline{\lambda_2} = 1$. Only the third set of MRs is capable of absorbing both λ_1 and $\overline{\lambda_2}$, creating a logical "0". For each cache line, two access gates (AGs) (10) are responsible for the data words and a third AG is responsible for the tag bits. The AGs of the selected cache line now have a control signal of "0", which allows either incoming data-to-write and tags to pass through to the optical Flip-Flops (FFs) for writing (11), or the contents of the FFs pass through to the tag comparator for reading (12). Table 1 shows the possible combinations of the 2-bit *index* and the corresponding rows they activate. All other lines will have some wavelengths still propagating to their corresponding AGs, not activating them and blocking any data (13).

When the data and tag bits enter the optical RAM bank and propagate through the AGs in the row denoted by the index (10), the wavelengths are distributed to their corresponding optical FFs through Arrayed Waveguide Gratings (AWGs) (14). AWGs act as optical demultiplexers that retrieve individual wavelengths from Dense Wavelength Division Multiplexing (DWDM) optical channels [79]. Each pair of wavelengths λ_i and $\overline{\lambda_i}$ drive the optical FF at the i-th bit in each 8-bit optical word. For a read, the AWGs multiplex the bits from the FFs into a single waveguide in the reverse direction (15).

2.2 Optical SRAM Cells

We experimentally verified and characterized in our lab integrated photonic RAMs and optical FFs (Figure 2) which adopt the cross-coupled circuit-layout RAM cell architecture presented in Figure 1, and use technologies of optical gain elements integrated hybridly with InP PhC-on-SOI [1].

The optical SRAM cell can be built using emerging PhC technologies to reap the speed, energy, and footprint benefits they offer, as previously validated using a hybrid InP/Si PhC laser [1], and InGaAsP Buried Heterostructure cavities [40, 57]. Figure 1 shows a possible principle of operation of an optical SRAM cell when energy-efficient and compact hybrid PhC lasers are employed [1]. Two PhC-based nanocavity lasers act as AGs and another laser acts as the optical FF. The AGs are

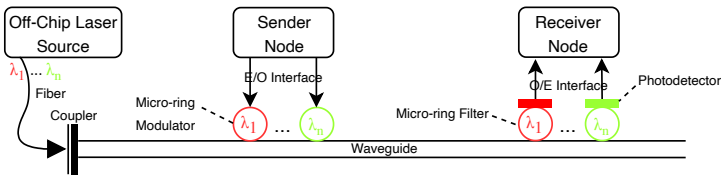


Fig. 3. Basic nanophotonic components.

controlled by a pair of *Access* and $\overline{Access}$ signals. When *Access* is a logical 1, AG#1 outputs the Bias pulse, and AG#2's output is suppressed. Thus, only the Bias pulse enters the FF, enabling the FF to output its previous stored value, successfully reading the content of the FF. When the *Access* is a logical 0, AG#1's output is suppressed, and AG#2 outputs the Bit signal. The value of the Bit signal facilitates a Set/Reset operation, thereby performing a write.

During a read operation, the data and tag bits of the selected cache line pass through the AGs and propagate to the tag comparator. The tag bits are demultiplexed through an AWG and each bit is XOR-ed with the corresponding bit of the tag array that the processor sent. The results of all XOR gates are then multiplexed to form a COMP signal, which is 0 if the tags match and 1 otherwise. The COMP signal is then used to drive the RAGs along with the $\overline{RW}$ signal, and allow the data to be replied to the processor only if there is a read operation and the tags match, i.e., a cache hit.

Both types of optical SRAM cells demonstrate extremely fast switching speeds. PhC nanolasers [1] exhibit 50 ps switching latency for fast memory operations. Alternatively, the InGaAsP/InP buried heterostructure PhC [57] has a 44 ps switch-on latency but requires 7 ns to switch off.

2.3 Optical Network-on-Chip

Recent breakthroughs in silicon photonics have propelled researchers to consider optical interconnects for on-chip communications. Optical NoCs provide low latency due to the fast propagation of light in silicon waveguides, and high bandwidth data transmission through dense wavelength-division multiplexing (DWDM), making them strong candidates to replace or partially replace traditional electronic NoCs. We first review the components for building optical NoCs, then briefly discuss two types of existing optical network interconnects.

2.3.1 Nanophotonic Building Blocks. Figure 3 shows the optical components to perform data transmission between a sender and a receiver on a chip. An off-chip laser source emits light with wavelengths $\lambda_1 \ldots \lambda_n$, which travels through an optical fiber and is brought onto the chip through a coupler. A single waveguide is capable of carrying multiple wavelengths in parallel by employing DWDM. The sender converts electrical signals into optical signals of specific wavelengths and modulates them onto the waveguide through micro-ring resonant modulators. MRs are placed next to waveguides and are tuned to modulate a specific wavelength by controlling their radius and temperature. The modulated wavelengths travel along the waveguide until they arrive at the receiver. MRs are also used on the receiver's side as filters to extract individual wavelengths from the waveguide. Then each wavelength is directed to a photodetector to convert the signal back to electrical currents, which subsequently go through amplifiers to be strong enough to drive electrical logic circuits. A DWDM density of n wavelength requires n modulator/filter pairs.

2.3.2 Corona. Corona [82] implements an optical crossbar to interconnect 64 four-core clusters. The crossbar is formed by 64 Multiple-Writer Single-Reader (MWSR) buses laid out in a serpentine fashion to connect all clusters. For each MWSR bus, 63 of the total 64 nodes can transmit on the

waveguide while the remaining one can receive from all others. An arbitration protocol is needed as multiple source nodes cannot transfer data to the same destination simultaneously. Token-based optical arbitration protocols [81] employ additional waveguides, in which receivers inject optical tokens for senders to acquire. A node can only transmit data to a destination when it has consumed the token on the waveguide corresponding to that destination node, meanwhile blocking other nodes from writing to the data bus. When the sender finishes transmitting data, it injects a new token onto the arbitration bus.

2.3.3 Firefly. Firefly [59] introduces an opto-electronic NoC using reservation-assisted Single-Writer Multiple-Reader optical crossbars (R-SWMR). A single R-SWMR bus involves one sender and multiple receivers. Arbitration, or reservation, is performed by the sender broadcasting a small optical reservation packet to all receivers. Upon receiving this packet, all nodes except the destination receiver turn off their corresponding receiving MRs, allowing only the destination node to receive the data from the sender. This saves power compared to a broadcast-based SWMR bus because all MRs on the optical path between the sender and receiver are off and induce minimal optical losses. An R-SWMR optical crossbar with a total of N nodes has N data channels, each with a data width of w bits, and N reservation channels of $\log N$ bits each.

3 THE PHO$ ARCHITECTURE

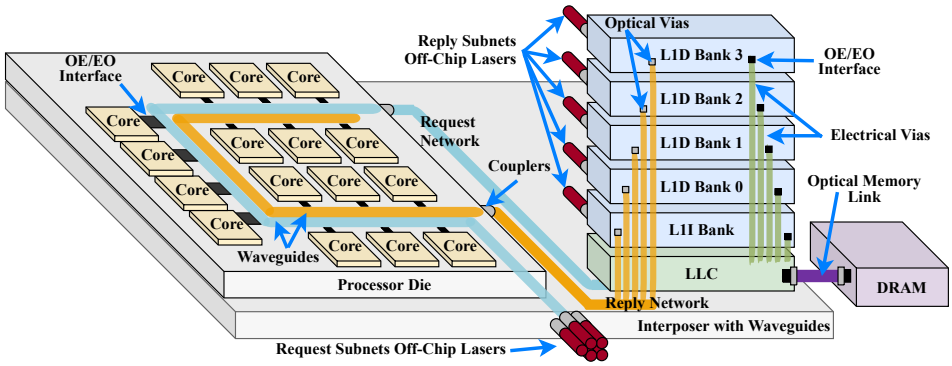
The optical cache prototype presented in Section 2 achieves very low latency. The optical SRAM cells can perform reads and writes in under 50 ps, and the outside decoding processing time is 100 ps, resulting in 150 ps cache read and write latencies. As long as the core-to-cache optical bus takes no more than 50 ps, such an optical cache can perform single-cycle cache accesses for core frequencies up to 5 GHz. However, while the InP/Si PhC laser-based optical SRAM cells have fast on/off switching speeds, each cell requires a pump power of 103.5 μW for storage operations [1]. Considering the number of components needed for a reasonably sized cache, static power quickly reaches hundreds of Watts, which is unrealistic. Thus, prior designs in this space [49, 50] are not implementable above 8 kB despite only taking up a physical footprint of 7.89 mm^2.

To avoid the additional pump power needed for biasing, Pho$ instead utilizes the InGaAsP-based optical SRAM cells demonstrated by Nozaki *et al.* [57]. These cells require a static power of only 30 nW, and their switch-on latency of 44 ps is on par with the 50 ps latency of the InP/Si PhC laser, allowing cache reads to still be completed within one cycle at 5 GHz. Cache writes are slow at 7 ns, but this can be mostly mitigated by memory-level parallelism (MLP) and a modern core's store queue. MLP allows for multiple concurrent memory requests, and store queues allow arithmetic operations and loads to bypass pending older writes. Thus, both MLP and store queues allow a core to overlap long write latencies with other work.

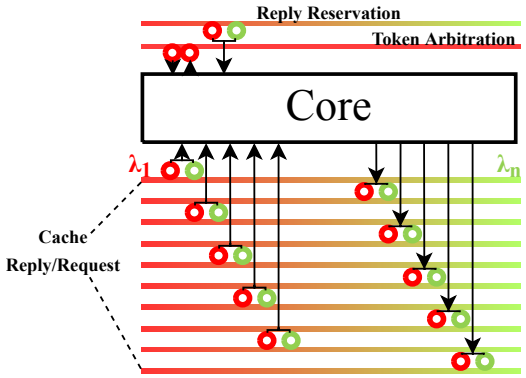
We propose Pho$, an opto-electronic cache hierarchy architecture that replaces all the electronic L1D, L1I, and L2 caches in a traditional CMP with a single, shared, high-capacity all-optical cache. Due to its high capacity and disaggregation from the cores, it is natural for the optical cache to be shared among all cores. We envision a shared optical L1D that employs 4 banks to provide high capacity and parallelism, and a shared optical L1I with one bank. The optical cache banks are fabricated on separate optical dies, while the processor cores remain on their original electronic die. The cores and optical caches are 2.5D-integrated on the same package and interconnected by an optical NoC, which handles arbitration and data transmission between them.

3.1 Pho$Net Network Topology

Figure 4a shows a high-level view of Pho$'s optical network topology, Pho$Net. The electronic processor die on the left houses the cores (16 cores in a 4×4 mesh layout) and sits atop an interposer



(a) Pho$Net optical network topology.



(b) Core optical connections.

Fig. 4. Pho$ and Pho$Net architecture.

with photonic waveguides. The dies on the right are 3D-stacked. The L1D and L1I banks are on optical dies, while the Last Level Cache (LLC) is a traditional electronic cache with its own die. Each optical cache bank has one input and one output port. We model single-ported cache banks throughout the manuscript, with the exception of the power investigation shown in Figure 9 where we analyze the impact of multi-ported caches on power consumption.

Communication between the cores and caches is entirely in the optical domain. Two sets of optical waveguides are laid between the processor and L1 cache dies. Each waveguide line in the figure is abstracted to represent multiple sub-networks, each comprising a bundle of waveguides with DWDM. The blue line depicts the subnets that carry requests from the cores to the cache banks (one subnet per bank). Within each request subnet, the cores are the writers and only one of the optical cache banks is the reader. Thus, each request subnet forms a Multiple-Writer Single-Reader (MWSR) crossbar [82] and uses token-based arbitration [81]. For each individual MWSR link, the cores are the writers and one of the optical cache banks is the reader. The orange line represents the reply subnets used by the cache banks to send data to the cores. For each reply subnet, one of the cache banks is the writer and the cores are the readers. Thus, the reply subnets are designed as Reservation-assisted Single-Writer Multiple-Reader (R-SWMR) crossbars [59]. Apart from waveguides for carrying data packets, additional waveguides are needed for both the request

networks' token arbitration channels and the reply networks' reservation channels. In essence, Pho$Net is a hybrid MWSR/R-SWMR optical network.

For a 16-core processor with 5 optical cache banks (as in Figure 4a), and assuming single-port cache banks, there are in total 5 hybrid subnets, each comprising an MWSR request and an R-SWMR reply crossbar with arbitration and reservation channels, respectively. If multi-port optical caches become possible in the future, we can further break the sub-networks to include only a subset of the processor cores. Note that with 16 cores, the reservation channel needs 4 wavelengths to represent the core ID.

The black squares in Figure 4a represent the Electrical-Optical (EO) and Optical-Electrical (OE) conversion interfaces for the cores to interact with the optical network, including modulators, filters, detectors, etc. More details are shown in Figure 4b. All cores have a pair of send/receive interfaces to interact with each cache bank and its corresponding set of waveguides, thus any core can send and receive on any sub-network. For diagram simplicity, not all connections between waveguides and the interfaces are shown.

Core-private caches, as employed by traditional multicores, require core-to-core communication to maintain coherence, which in turn requires full-blown MWSR or R-SWMR crossbars with all-to-all connectivity. By employing an L1 cache that is shared among all cores, Pho$ physically decouples the cores from the caches and removes the need for cache coherency and inter-core traffic. Thus, it is no longer necessary to build physical links between cores. It suffices to implement separate networks for carrying either requests or reply packets directly to and from caches, and optimize each for their purpose. The hybrid Pho$Net network capitalizes on this observation to shrink the network by avoiding full connectivity among all nodes, saving power, area, and cost.

The request and reply subnets are powered by separate off-chip lasers to minimize laser power (Section 3.4). Finally, the LLC can be connected to the DRAM through an optical interconnect [5] for low latency, high bandwidth DRAM accesses. The adoption of DWDM enables OCM to send out the entire cache line in one burst, which decreases the time needed to transmit and receive data on the memory interconnect. This, along with a higher propagation speed, reduces memory access latencies and increases bandwidth.

3.2 Pho$Net Arbitration Protocol

For each cache bank, all cores on the same request (or reply) subnet share the same channel, thus it is important to ensure that requests from (or replies to) different cores do not conflict. As Pho$Net is half MWSR and half R-SWMR (Section 3.1), it requires a new way to arbitrate packets.

Arbitration in Pho$Net is achieved through a protocol similar to optical token channel arbitration [81]. A single optical token circulates through each bank's request-reply subnets. When a core needs to send a cache request, it turns on its MRs on the arbitration channel to try to consume the token. If the token reaches the core's receiving MRs the core absorbs the token, acquiring the exclusive access to the request channel. The core then is able to send request packets to the optical cache. After sending a request, the core turns on its receiving MRs and starts listening on the reply channel without the need for further arbitration.

The cache, upon receiving the request, it processes it in the optical domain. Upon an L1 hit, the cache injects the data to the reply network followed by a new token. The reservation channel is not used at this moment because all cores except the original requester have their receiving MRs turned off (they are not expecting a data packet from the cache). If a cache miss is detected, the optical cache forwards the request to the electronic LLC after an OE conversion. It also sends out a Negative-Acknowledgement (NACK) packet on the reply waveguide. The requester core, who is still listening on the reply channel gets the NACK packet, realizes there is a cache miss, and turns off its receiving MRs. This mechanism ensures that during cache hits, the cache does not

need to arbitrate for the reply channel as the requester core and cache have exclusive permissions to transmit on the request and reply channels, respectively. Regardless of a cache hit or miss, a new token is injected in the arbitration channel. This new token can be grabbed by any core who wants to send a cache request. The requester, after consuming the reply packet (be it data from the cache or a NACK) also turns on its MRs on the arbitration channel so that it can grab the new token. In both cases, the total latency for the read and reply packets should be constant as there is always exactly one round trip made from core to the cache and back to the core again: $t_{total} = t_{arbitration} + t_{request} + t_{reply} = t_{arbitration} + t_{RTT} + t_{SERDES}$ (our design does not need SERDES, but we include it in the equation for completeness).

In the case of an L1 cache miss, the electronic LLC eventually responds with the data requested by the optical L1. The L1 first tries to arbitrate for the data channel like any other core would; this ensures that no other data packets are transferring on the data channel and no additional request can be made to the bank. When it has successfully grabbed the channel, it first broadcasts on the reservation channel so that the core that sent the original request for this cache access can turn on its receiving MRs. Then the cache transmits the reply data on the data channel, followed by a new token on the arbitration channel that all cores can grab to start a new cache request. The data packet needs to trail the reservation packet by a fixed delay to allow the receiving core enough time to turn on its data channel MRs. It is also important to note that the core ID needs to be included in the original request packet so that in case of a cache miss, the L1 knows which core to send the reply to once it gets the data from the LLC.

Figure 5 shows an example arbitration in a simplified 3-core 1-cache-bank setup. For simplicity, we have also combined the request and reply data channels in the figure.

3.3 Pho$ Optical Cache Architecture

This section describes the architecture design of the 1 MB optical cache banks employed by Pho$, as well as the components' optical losses for calculating the optical power budget. In a nutshell, Pho$ employs a cascaded two-level row decoder to reduce laser power, active rather than passive components to reduce off-ring optical losses, and uses PhC for the optical SRAM bit cells to dramatically lower the static power consumption. These design innovations allow Pho$ to be implementable within a reasonable power budget even for multi-MB cache capacities, in contrast to prior designs in this space [49, 50] that are not implementable above 8 kB.

Assuming a 64 B cache line, each of Pho$'s five 1 MB direct-mapped cache banks has 16384 lines. Row decoding with an MR-based matrix, as in prior work [49], is impractical: the number of MRs needed for each line increases as the matrix scales up, consuming inordinate amounts of power. Instead, Pho$ uses a two-level cascaded row decoding process (Figure 6a). The first-stage demultiplexing uses an active 9-to-512 tree global row selector, implemented with PhC nanocavity-based resonant switches [56], which activates only one of the 512 5-to-32 passive MR-based row decoders in the second stage. The second-stage row decoder then selects one line to activate, allowing a read/write operation to perform on the correct cache line. In this way, we build a 16384-line row selector with only 5 MRs per line instead of 14, drastically lowering laser power.

Active optical devices such as an PhC active tree are estimated to have optical losses of 0 dB. This is because in order to function, the active components within the device also provide some small amplification, sufficient to compensate its own losses. As a result, we can lower the laser power required for light to reach the optical FFs.

For the column decoding optical circuit (Figure 6b), we use 8 1-to-128 AWGs to demultiplex the wavelengths in the incoming light into their respective optical FFs. Each 1-to-128 AWG serves 64 bits, with 2 complementary channels per FF, so a total of 8 AWG-based column decoders are needed for a 64 B cache line. For each AWG, an AG controls the direction of data when switching between

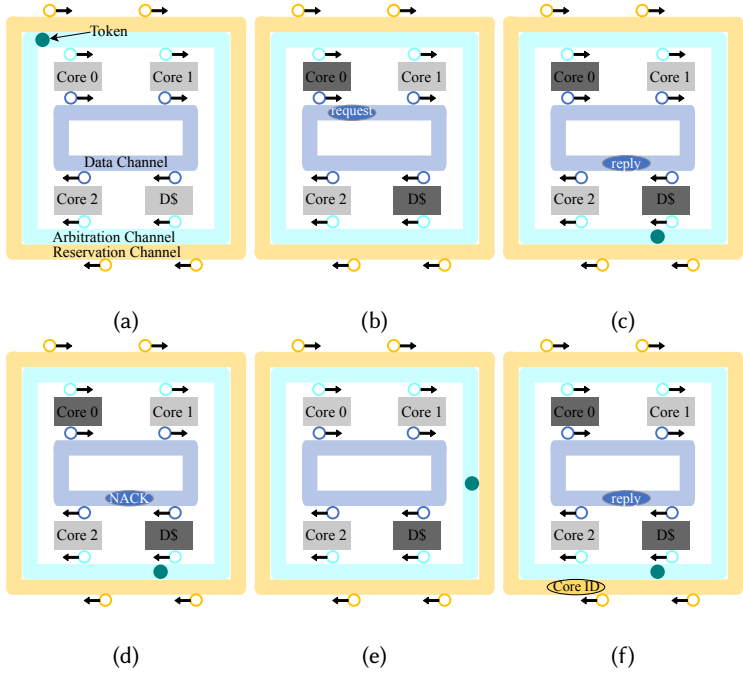


Fig. 5. Arbitration protocol. (a) Token circles the arbitration channel waiting to be grabbed. (b) Core 0 grabs the token and sends a request packet on the data channel. (c) Cache hit: reply packet sent, followed by a new token. (d) Cache miss: NACK sent, followed by a new token. (e) Cache has data following the miss, tries to grab token first. (f) Cache notifies core 0 with reservation channel, sends reply packet followed by new token.

(a) Row decoder (b) Column decoder

Fig. 6. Optical cache peripheral circuit.

Table 2. Optical cache components. See Section 3.3 for an explanation of 0 dB losses.

Component	Count	Passive/ Active	Optical Loss
Tag Comparator AWGs	2 AWGs	Passive	3 dB
XOR gates for Tag	52 PhCs	Active	0 dB
WAG	9 PhCs	Active	0 dB
RAG	8 PhCs	Active	0 dB
AG	147456 PhCs	Active	0 dB
Global Row Selector	1024 PhCs	Active	0 dB
Row Decoder	81920 MRs	Passive	#MRs × filter drop
AWG Column Decoder	147456 AWGs	Passive	3 dB
Optical RAM Bank	8814592 PhCs	Active	0 dB
Total PhCs: 8963141	Total MRs: 81920		
Total AWGs: 147458	Min power at optical FF: −14 dBm		

Table 3. Optical power source responsibilities.

Laser Source	Optical NoC Components	Optical Cache Components
Core network	Request network Token arbitration channel Reservation channel	Read/write selector (WAGs) Row decoder, Column decoder AGs, RAGs
Cache network	Reply network	Optical FFs, RAGS Tag comparator

writing and reading the FFs. The AGs are controlled by 8 WAGs acting as read/write selectors. Data are fed into the reply waveguides through 8 RAGs (Section 2).

For a 48-bit physical address with 14 bits used for the index, 6 bits used for the offset, and 2 bits used for bank selection, the tag field requires 26 bits (28 for L1I). Because our optical cache is direct-mapped, one 1-to-128 AWG is enough to demultiplex the incoming tag signal into separate wavelengths. As with the data cells in the above paragraph, an extra AG per line and one global WAG are needed to control the reading and writing of the tag cells. The tag comparator is built with 26 XOR gates and two 1×26 AWGs. The XOR gates and AGs are also implemented using active PhC resonant switches. The data and tag output of all 16384 cache lines need to be multiplexed onto one waveguide before the tag bits are passed into the tag comparator and the data bits into the RAGs. This can be done with active PhC trees acting as multiplexers.

Table 2 summarizes the component counts and the optical loss parameters for a 1 MB optical cache bank in Pho$. They will be used in calculating the laser powers required for the core and cache optical networks.

3.4 Laser Power Sources and Optical NoC Parameters

The request subnet and the optical cells and reply subnet are powered by separate laser sources. The laser used to power the request subnet also powers the row decoders, column decoders, read/write selectors, and AGs before the optical FFs, because additional lasers along the path can overwrite any data already traveling on the waveguide. The token arbitration and reservation channels are also powered by the same laser. On the other hand, the FFs in the optical cache cells need a continuous power source to store data using photons, and the same laser can be used to power the tag comparators as well as the reply network. Table 3 details the specific components of the optical system that each of the two lasers is in charge of powering. We call them the "core network" and "cache network", respectively.

Pho$ uses off-chip lasers because on-chip lasers may generate a lot of power and heat. The lasers are brought onto the chip using optical fibers and couplers, and then the light is distributed onto waveguides using splitters. There is only one splitter per waveguide in our design. We have a total of 105 waveguides combined across all subnets for the single-port design. We consider a comprehensive range of parameters for optical components by grouping the parameters of several seminal optical NoC designs from recent years [24, 25, 33, 39, 41, 52, 58, 76, 83] into two groups, conservative and aggressive (Table 4), which represent the worst-case and best-case parameters among these works, respectively. Showing both conservative and aggressive parameters highlights the spread of possible values for each nanophotonic parameter. We expect each device to exhibit losses between these two values.

The waveguides of the data channel in the request network feed directly into the input ports of the optical cache and continue onto the optical FFs after passing various optical components, so we use the minimum power needed at optical FFs in Table 2 in the calculation of the laser power

Table 4. Nanophotonic parameters for Pho$Net.

Component	Conservative	Aggressive	Component	Conservative	Aggressive
Waveguide	1 dB/cm	0.05 dB/cm	Waveguide bending	0.005 dB	0 dB
Coupler	2 dB	1 dB	Waveguide crossing	0.12 dB	0.05 dB
Nonlinearity	1 dB	1 dB	Photodetector	0.1 dB	0.1 dB
Ring-through	0.01 dB	0.001 dB	Modulator insertion	1 dB	0.001 dB
Filter drop	1.5 dB	0.5 dB	Detector sensitivity	−16 dBm	−28 dBm
Splitter	0.2 dB	0.1 dB	Laser Efficiency	30%	30%
Trimming	20 µW/ring	5 µW/ring	Modulation / Demod.	150 fJ/bit	20 fJ/bit

Table 5. Simulated system parameters.

	Baseline	Pho$	Pho$_OCM
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	electronic, private, 64 B line, 32 kB/core, 8-way, 4 cycles	optical, shared, 64 B line, 1 MB direct-mapped, 2-cycle read, 23-cycle write	
L1 DCache	electronic, private, 64 B line, 32 kB/core, 8-way, 4 cycles	optical, shared, 4 banks, 64 B line, 4 MB direct-mapped, 2-cycle read, 23-cycle write	
L2	electronic, private, 64 B line, 256 kB/core, 4-way, 14 cycles	N/A	
LLC	electronic, shared, non-inclusive, 64 B line, 22 MB, 11-way, 50 cycles		
Core-L1 Netw.	electronic, point-to-point	hybrid optical	
LLC Network	electronic, 4×4 mesh (NUCA)		
Memory	electrically connected, 49.37 ns		optically connected, 41.61 ns

for those data waveguides. For all other waveguides, including those in the reply network data channel, token arbitration and reservation channels, the detector sensitivity is used instead.

4 EXPERIMENTAL METHODOLOGY

4.1 Performance Simulations

We evaluate Pho$ using the Sniper simulator [12, 13] running workloads from SPEC CPU2017 [10] (SPECspeed, ref inputs) and Parsec 3.0 [7] (simlarge inputs) benchmark suites. For CPU2017, we used Pinpoints [62] to collect representative regions. We compare our results with a baseline electronic multicore whose configuration is similar to a 16-core Intel Skylake [19, 27, 53, 65, 84, 85]. For Pho$, we perform experiments with a per-bank capacity of 1 MB. We model both a conventional DRAM for Pho$, as well as an optically connected one (Pho$_OCM). DRAM bandwidth is modeled, but not the internals of DDR circuitry. Other non-cache and non-network related parameters are consistent across all the configurations. Table 5 summarizes the detailed configurations for our experiments. We run multi-threaded workloads in Parsec by pinning threads to individual cores. We also use methods introduced by Heirman *et al.* [35] to construct cycle stacks for better analysis of experiment results.

We perform physical measurements on a Dell PowerEdge R710 server [20] and estimate a 15 cm average distance between the LLC and DRAM DIMMs. We estimate the latency for DRAM accesses over that distance to be 46.7 ps/cm for light propagation in optical waveguides [15], and 50.4 ps/cm for electrical pulse propagation speed in electronic links[51].

4.2 Modeling Power, Energy, and Area

To get an insight into the optical NoC's power consumption, we compare our hybrid optical NoC, Pho$Net, against three network configurations. The first is a fully connected MWSR crossbar with 21 21-to-1 MWSR links (16 cores and 5 cache banks, a total of 21 nodes) with a token arbitration protocol. The second is a fully connected R-SWMR crossbar with 21 1-to-21 reservation-assisted SWMR links. Finally, we also compare against a "one channel" network where requests and replies share the waveguides as a single data channel, while all other characteristics are the same as in Pho$Net. For this comparison, we ignore the static power needed for optical FFs to operate as this depends on the number of cache components and not the network configuration.

Each data packet contains 512 bits of data, 42 bits of address, and 4 bits dedicated to the core ID (for the full R-SWMR configuration, 5 bits are used because there are 21 nodes in each R-SWMR link). Each bit is encoded with complementary wavelengths λ_i and $\overline{\lambda_i}$ to drive the optical cache circuits. We model a 64-λ DWDM. Because we stay in the optical domain for L1 cache accesses, we do not employ SERDES. All optical packets are sent in one burst.

To calculate the cores' die size, we use McPAT [42] to estimate the area of processor cores under the 14 nm technology node. Parameters are adapted from the International Technology Roadmap for Semiconductors 2015 Edition [17] and Fincacti [69]. We estimate the core die size to be 59 mm^2 and assume the distance between the core and cache dies is 1 mm. Using the scaling methodology presented by Maniotis *et al.* [49], we analyze the area footprint of a 1 MB direct-mapped optical cache by considering different component alignments and determine the optimal area to be 89 mm^2, where the distance traveled by the data and tag bits is 34.7 mm and that traveled by the index bits is 24.4 mm. We calculated the round-trip time for a cache access by considering EO/OE conversion latencies of 14.3 ps and 0.2 ps [15], the total distance traveled by the request and reply optical packets (175 ps roundtrip on the optical network), and the latency to access the optical cache bank itself (44 ps bitcell latency and 100 ps row/column decoder). This ensures 2-cycle access for up to 5 GHz. We calculated the area overhead for EO/EO interfaces of Pho$ to be 11.8% per core tile and 8% for optical dies by scaling numbers from Sun *et al.* [73]. To calculate the total area consumed by the NoC on the electronic die we use a waveguide pitch of 3 µm [91] and a micro-ring pitch of 5 µm [54]. The total area overhead is 0.17 mm^2 if the waveguides are stacked vertically and 3.32 mm^2 if laid out on the same plane (0.3% and 5.7% of total die area, respectively).

For every network configuration, we also explore the possibility of multi-port caches. For an N-port optical cache, each port serves $\frac{16}{N}$ cores. For each sub-network, the original 16-to-1 request MWSR link becomes a $\frac{16}{N}$-to-1 link and the 1-to-16 reply R-SWMR link likewise. Individual links can become shorter and need fewer optical components, but more links are needed. In this paper, we consider 1-port, 2-port and 4-port optical caches in our power analysis.

We estimate the energy consumption of cores, electronic caches, electronic on-chip interconnects, and DRAM using McPAT [42]. The energy consumption of the optical caches and Pho$Net are calculated analytically. We used detailed simulation results such as the number of cache loads, stores, misses, and evictions and calculated each operation's cache access, network arbitration, reservation, and data transfer energy. As the request and reply subnet lasers power the passive optical cache components and there is no need for additional modulation/demodulation within the optical domain, the optical cache dynamic energy is categorized as part of the NoC. Thus, to avoid double-counting, we do not include it in the energy of the optical cache, as it has already been included in the overall energy calculation as part of the NoC. The overall optical cache static power is calculated by multiplying the number of active components with the static power of each component. We use the 30 nW reported by Nozaki *et al.* [57] as the static power needed for every optical FF. For Pho$Net we model the best configuration determined by our design-space

exploration (Figure 9). The NoC dynamic power accounts for the modulation/demodulation during the EO/OE conversions at the cores and LLC.

Finally, we compare Pho$'s performance and energy-efficiency against prior works [49]. For fair comparison, we scaled important metrics like data cache capacity and processor frequency of Maniotis *et al.*'s implementation to the same level as Pho$.

5 EXPERIMENTAL RESULTS

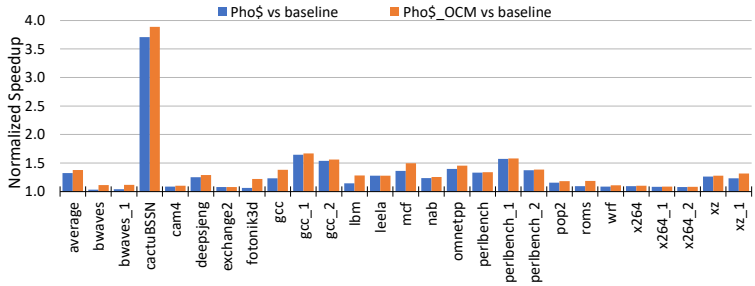
5.1 Benchmark Performance

Figures 7a and 7b summarize the speedup of Pho$ and Pho$_OCM over the baseline running SPEC CPU2017 and Parsec 3.0. Figures 7c and 7d show the normalized CPI stacks [35], respectively. Each bar shows the relative values of cycles per instruction that are spent waiting for a particular component in the system. The "busy" sub-bar denotes the fraction of time spent within the core itself. For each application, the left, middle, and right bars represent the normalized CPI stacks of baseline, Pho$, and Pho$_OCM, respectively. Pho$ achieves an average speedup (we use arithmetic average for all averages in this paper) of 1.34× and 1.41× without and with OCM, respectively. For CPU2017, we see an improved execution time across all applications, with *cactuBSSN* having a maximum of 3.89× speedup. Pho$ is able to significantly decrease instruction fetch delays because of its fast L1 read latency and large L1I capacity. Similarly, most applications enjoy a decrease in total L1D and L2 delay, like *leela* and *gcc_1*. The increased L1 capacity also means there are fewer misses that must visit the much slower LLC, and this is indicated by a reduced CPI for *mem-llc* in applications like *gcc*, *mcf*, and *xz*. The slow 7 ns L1 write time does not seem to have much adverse effect. OCM-enabled Pho$ makes an impact in applications like *fotonik3d* and *lbm*, providing on average an additional 5% speedup across the suite.

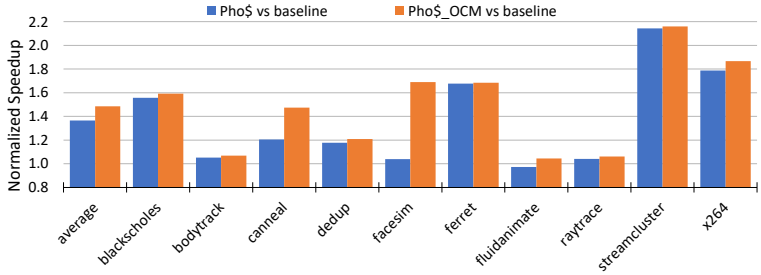
For the multi-threaded workloads in Parsec, Pho$ is able to speed up the execution of most applications, obtaining on average 1.37× speedup. Instruction fetch delays are greatly reduced, which is most prominent in *bodytrack* and *x264*. We find that Pho$ does not suffer from high contention from a shared L1I cache. This is due to Pho$ combining the aggregate capacity of the individual L1Is in baseline into a larger shared L1I, allowing more of the instruction stream to be L1-resident. Each fetched cache line also includes multiple instructions, eliminating the need for fetching on every cycle. The CPI component for L1D in Pho$ and Pho$_OCM is 44% lower on average than the CPI contribution of L1D+L2 in the baseline. The benefits of a low read latency and large capacity outweigh the disadvantage of a high write latency. Like in CPU2017, the large capacity of Pho$'s L1 cache also results in fewer visits to the LLC and thus fewer stalls. For example, Pho$ in *blackscholes* almost eliminates the CPI contribution of LLC and in *streamcluster* reduces it by about 4×. On average, Pho$ decreases LLC delays by 2.5×. Adding OCM to Pho$ reduces the average CPI spent waiting for DRAM by 2× and increases the overall speedup to 1.48×.

Pho$ shows a slight performance slowdown in *fluidanimate*. This is caused by serialization instructions, which force the processor to flush all pending writes in its store buffer before executing the next instruction [18], and the long write latency of the optical cache stalls the processor for a prolonged period of time. This shows up as a significant increase of the "other" component in the CPI stacks for *canneal* and *fluidanimate*. However, OCM helps to outweigh this scenario and allows Pho$ to attain speedups in all of Parsec's applications, even in *fluidanimate* which experiences slowdown without OCM.

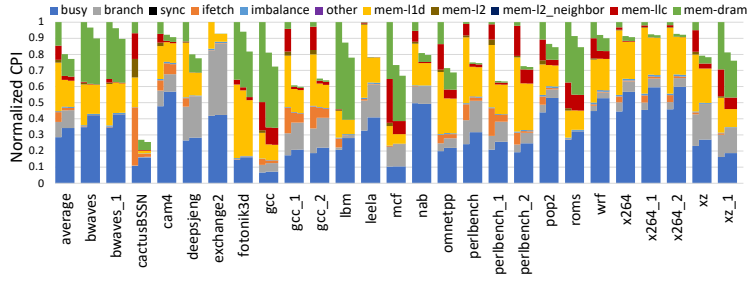
To isolate the source of the performance gains (capacity, latency, sharing), we examine three additional configurations derived from the electronic baseline (①): a hypothetical private L1 cache with increased size of 256 kB (8-way set-associative) but latency of a 32 kB cache (②), a 4 MB direct-mapped shared L1 but also with the latency of a 32 kB cache (③), and a 4 MB direct-mapped shared



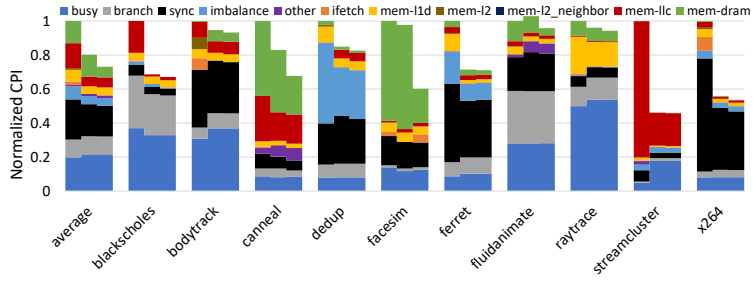
(a) CPU2017 Speedup over *baseline* (electronic multicore).



(b) Parsec Speedup over *baseline* (electronic multicore).

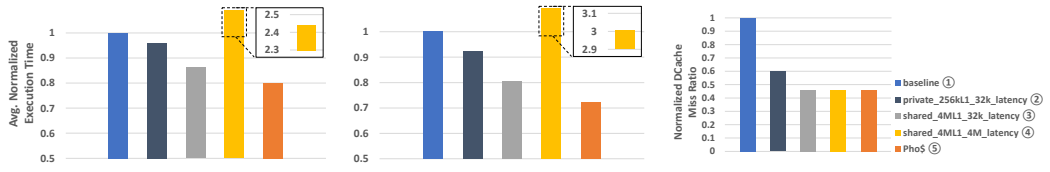


(c) CPU2017 CPI Stack normalized to *baseline* (electronic multicore).



(d) Parsec CPI Stack normalized to *baseline* (electronic multicore).

Fig. 7. Speedup and CPI Stacks for CPU2017 and Parsec. The three bars per benchmark in the CPI stack correspond to baseline, Pho$, and Pho$_OCM.



(a) All benchmarks (b) Cache-intensive benchmarks (c) L1 DCache Miss Ratio

Fig. 8. Sources of Pho$'s speedup. (a) presents the average normalized execution time of the electronic baseline (blue) and hypothetical electronic caches with zero-cost higher capacity (dark blue), zero-cost higher capacity plus sharing (grey), and a realistic high-capacity shared electronic cache (yellow) vs. Pho$ (orange). (b) presents the above results but only for cache-intensive benchmarks ($\frac{cacheCPI}{totalCPI} > 40\%$). (c) shows the average normalized L1 DCache miss ratio of all benchmarks.

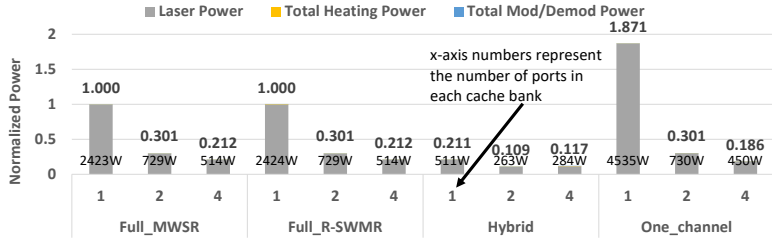
L1 with its respective real-world latency (④). Figure 8a shows the normalized average execution time of the baseline, the three hypothetical configurations, and Pho$. Pho$ gains 4% performance from increased capacity, 10% from sharing the L1, and 6% from latency. Designs ② and ③ are unrealistic but help us isolate the source of gains. Design ④ is realistic but impractical (2.4× slower than baseline; Pho$ beats it by 3.1×). Figure 8b shows the normalized average execution time of the same configurations, but only considering cache intensive benchmarks where the percentage of CPI spent on the cache hierarchy in total CPI is greater than 40%. The performance gain from more capacity, sharing, and latency increased to 8%, 12%, and 8%, respectively. Figure 8c shows the normalized L1 DCache miss ratio of all five configurations across all benchmarks. Comparing ② and ③, while an 8-way set-associative cache may lower miss rates compared to a direct-mapped one, increasing capacity by 16× lowers misses even more. When running single-threaded workloads (SPEC) the entirety of the 4 MB cache is available to the running thread, far surpassing the performance of a 256 kB 8-way cache, even with the unreasonably fast access of a 32 kB one. In multithreaded workloads (PARSEC), when this 16× larger L1 cache is shared, the threads act as prefetchers for one another, both for data and instructions, and also avoid cache-to-cache coherence traffic. As a result, ③ has a 14% lower average L1Dcache miss ratio than ②.

To further validate the need for a larger L1 DCache, we collected the working set sizes of SPEC CPU2017 and Parsec [8, 71] and found that most applications have working set sizes larger than 32 MB, and would not fit within Pho$'s L1D and LLC combined capacity (26 MB). Some benchmarks like *canneal* and *dedup* might even be considered to need unbounded cache and memory sizes. This implies the potential performance benefits that larger and faster cache hierarchies like Pho$ can bring.

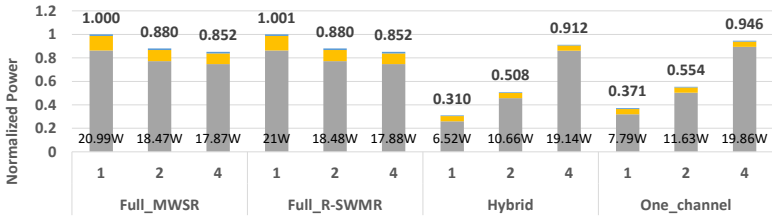
5.2 Optical NoC Power Analysis

Figure 9 shows the normalized optical power consumption of Full MWSR, Full R-SWMR, Pho$Net, and One Channel normalized to the Full MWSR configuration (normalized separately for the conservative and aggressive nanophotonic technologies) with 1-, 2-, and 4-port optical banks (x-axis). Our estimates include the power consumption of the off-chip laser, heating for MRs, and modulation/demodulation. Table 6 summarizes the different subnet and MR counts for the four optical NoC configurations with 1-, 2-, and 4-port optical caches.

Under conservative nanophotonic parameters, Pho$Net shows the lowest power consumption among alternatives and for all port numbers. Laser power constitutes over 99% of optical power for all configurations. This is due to the high optical loss accumulated along the data path. For each waveguide with a DWDM of 64 wavelengths, 64 MRs need to be placed at each node as



(a) Conservative nanophotonic parameters



(b) Aggressive nanophotonic parameters

Fig. 9. Optical NoC power for a range of nanophotonic parameters. All results are normalized to the 1-port Full MWSR design.

Table 6. Configuration comparison of different optical networks with 1-, 2-, and 4-port caches.

N_{subnet}: Number of sub-networks;
N_{ring}: Total number of MRs in the NoC;
$N_{ring.data}$: Number of MRs for each wavelength on the *data* channel;
$N_{ring.arb}$: Number of MRs for each wavelength on the *arbitration* channel;
$N_{ring.res}$: Number of MRs for each wavelength on the *reservation* channel.

	Network	Full MWSR	Full R-SWMR	One Channel	Pho$Net
1 port	N_{subnet}	1	1	5	5
	N_{ring}	515970	517293	187390	187390
	$N_{ring.data}$	21	21	32	16
	$N_{ring.arb}$	42	NA	34	34
	$N_{ring.res}$	NA	21	17	17
2 port	N_{subnet}	2	2	10	10
	N_{ring}	395460	396136	187330	187330
	$N_{ring.data}$	13	13	16	8
	$N_{ring.arb}$	26	NA	18	18
	$N_{ring.res}$	NA	13	9	9
4 port	N_{subnet}	4	4	20	20
	N_{ring}	379080	379728	187280	187280
	$N_{ring.data}$	9	9	8	4
	$N_{ring.arb}$	18	NA	10	10
	$N_{ring.res}$	NA	9	5	5

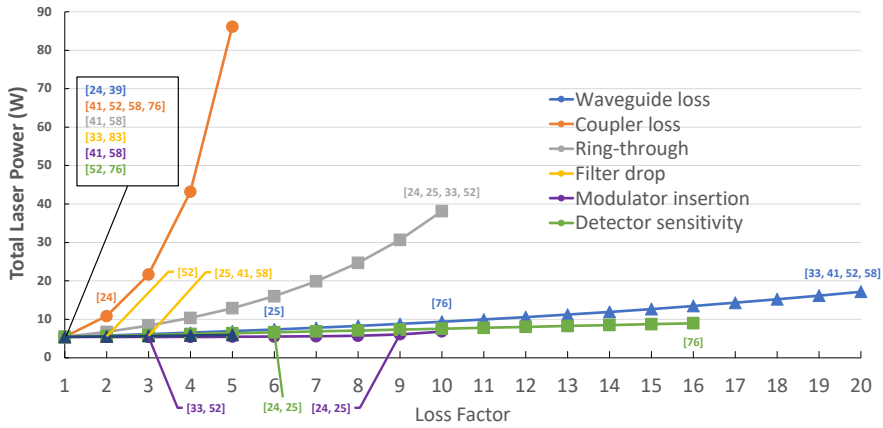


Fig. 10. Laser power sensitivity to nanophotonic parameters. The sources for each parameter value are noted in the figure, following the same color coding scheme shown at the legend.

either modulators or demodulators for Full MWSR, Full R-SWMR, and Pho$Net topologies (128 for One Channel). This causes the optical loss incurred by all MRs along one data waveguide to be high with a conservative ring-through loss of 0.01 dB. Pho$Net gains an advantage over the other three topologies because it does not need to keep all nodes fully connected, requiring the fewest MRs along each datapath as well as the fewest data channels, thus reducing its total off-ring losses. One Channel has the worst optical loss because for each waveguide twice as many MRs are needed. However, the high optical loss per device under conservative technology parameters still results in unrealistically high power requirements. For single port optical caches, even the most power-efficient Pho$Net configuration under the highly conservative nanophotonic parameters consumes 511 W for the network, requiring a 506 W laser power.

When we increase the number of ports of our optical cache to 2 and 4, the number of cores in each sub-network is halved and quartered, respectively, reducing the number of MRs that need to be placed along each waveguide and the total optical loss. At the same time, more sub-networks increase the number of waveguides, potentially offsetting the benefit above. All four configurations obtain lower laser power. However, Pho$Net still consumes the least power. Compared to the other topologies, Pho$Net saves 64% of total power with 2-port caches and 37–45% with 4-port caches.

We perform the same analysis using the aggressive nanophotonic parameters. The optical loss for off-resonance rings decreases from 0.01 dB to 0.001 dB. As a result, the total laser power can be lowered to a reasonable level. For single-port caches, Pho$Net achieves the lowest optical power of 6.52 W, requiring 5.43 W for the laser, 0.94 W for ring heating, and 0.15 W for modulation/demodulation. Compared to the other designs, Pho$Net still benefits from removing unnecessary links from the network and employing fewer MRs per waveguide. Having fewer MRs also reduces the MR heating and modulation/demodulation power. As a result, Pho$Net saves 70% of power compared to the two fully connected topologies and 16% compared to One Channel.

Increasing the number of cache ports under aggressive parameters increases power consumption for both Pho$Net and One Channel. As technology scales, off-ring losses have smaller weights in the overall loss. The total loss for a wavelength and even one waveguide does not decrease by much even if we can halve the number of off-rings. For example, the per-wavelength laser power required for Pho$Net's request network under conservative parameters decreases from 16.2 mW to 4.0 mW when we compare 1-port to 2-port caches; however, that number only decreases from 0.27 mW to

0.24 mW under aggressive parameters. Optical loss is now more sensitive to the number of parallel waveguides. For Pho$Net and One Channel, 2-port and 4-port cache designs result in twice and four times the number of waveguides. As a result, their power consumptions increase when we have multi-port caches. On the other hand, because the MWSR and R-SWMR topologies are fully connected no matter the number of ports, they benefit from shorter individual links and fewer off-rings per link. With 4-port caches, Pho$Net has almost the same total power consumption as the other configurations. However, multi-port caches in theory should provide more performance benefits by being able to serve multiple requests simultaneously. The optimal performance-power choice is beyond the scope of this paper.

Overall, the study using aggressive nanophotonic parameters gives us a very promising power consumption outlook with the lowest power consumption being under 7 W.

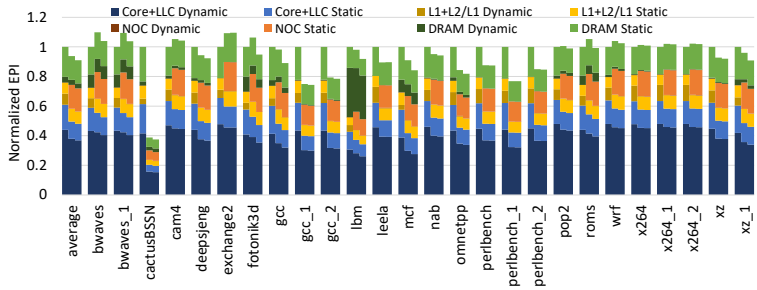
Figure 10 shows the sensitivity of Pho$Net's laser power to changes the scaling of optical loss for each nanophotonic parameter. Each parameter is scaled from its aggressive number up to its conservative counterpart (for modulator insertion loss, the maximum scaling factor is 1000, so we plot using the log_2 of the scaling factor on the x-axis). Pho$Net's laser power is most sensitive to coupler loss. It is also relatively sensitive to ring-through loss due to the large number of MRs required. It is relatively insensitive to all other nanophotonic parameters. This demonstrates the robustness of Pho$Net's laser power consumption under a wide range of nanophotonic technologies.

5.3 Energy Evaluation

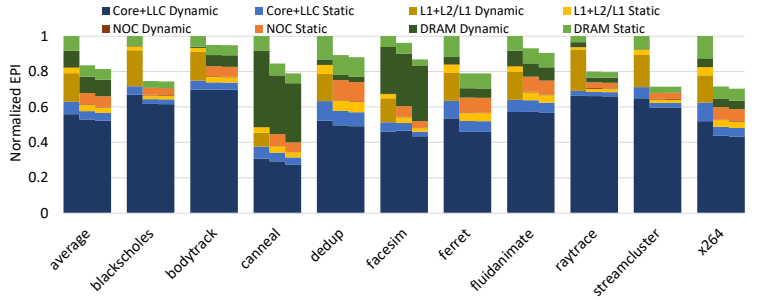
Figure 11 shows Pho$'s normalized energy per instruction (EPI, J/*insn*) and normalized energy × delay product (EDP, J×s). The three bars for each workload represent baseline, Pho$, and Pho$_OCM. By replacing conventional processors' electrical L1, L2, and mesh network with Pho$'s optical architecture, the original components' energy consumptions now become the energy consumed by the optical L1 cache and optical NoC. Pho$'s L1 static energy is considered to be the total pump energy needed for optical FF operations to be stable, and it is mostly on the same level with the combined L1 and L2 static energy in the baseline. We do not consider Pho$'s L1 dynamic energy as these dynamic energy consuming operations are considered as part of the NoC's operations. Pho$ also has lower core and LLC energy consumption as there are less frequent core stalls and fewer LLC accesses. The EO/OE conversion energy overhead for Pho$ is minimal, which is represented by NOC Dynamic. Overall, Pho$_OCM saves on average 12% EPI and 31% EDP, and is most energy-efficient in applications such as *blackscholes*, *streamcluster*, and *cactuBSSN*.

5.4 Comparison with Previous Optical Cache Designs

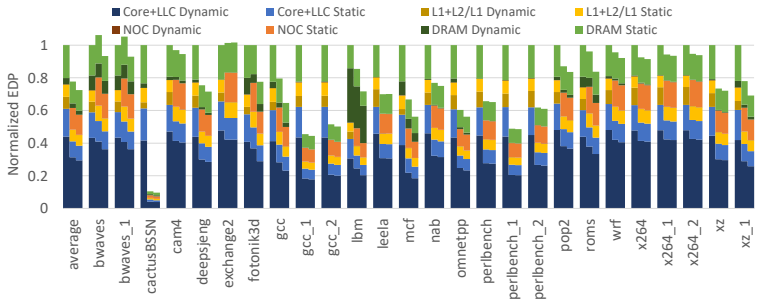
When comparing against Pho$, the advantages of previous optical cache design from Maniotis *et al.* [49] are its 2-way associative cache design, fast write latency at 2-cycles, and a Time-Division Multiplexed (TDM) optical bus. The TDM optical bus is a less complex design compared to Pho$Net's hybrid optical network: 1) it requires just a single set of waveguides that link all nodes in order; 2) employing a TDM-based bus eliminates the need for network arbitration. However, a number of practical problems exist in this design. First, it relies on set-associative optical caches, but no optical cache designs are capable of set-associative replacement due to the lack of a replacement algorithm in the optical domain that optical set-associativity relies on. Second, its high static power due to all-passive decoder and power-inefficient PhC cells [1] makes it impractical. Finally, to avoid data collision, its TDM-based optical bus requires the entire optical system to operate at 50 −80 GHz, as 1 CPU cycle needs to correspond to 16 optical cycles. To the best of our knowledge this is currently unattainable for optical interconnects and optical memory [2, 83]. Figure 12 shows the performance (speedup) and energy comparison (*log* scale) between Pho$ and Maniotis *et al.* [49], even under the assumption that the associativity and TDM challenges are resolved. Pho$ is able to achieve a



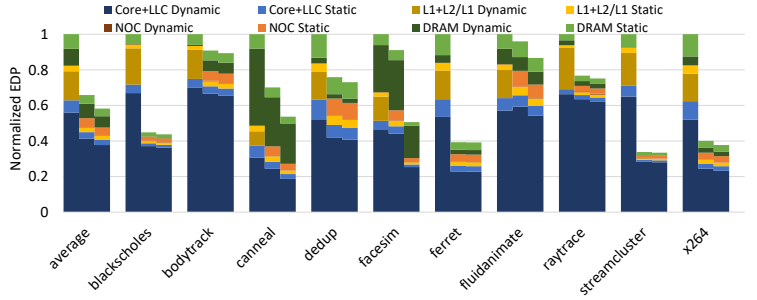
(a) CPU2017 energy per instruction normalized to *baseline*.



(b) Parsec energy per instruction normalized to *baseline*.



(c) CPU2017 energy×delay product normalized to *baseline*.



(d) Parsec energy×delay product normalized to *baseline*.

Fig. 11. Normalized energy per instruction and energy×delay product for CPU2017 and Parsec. For each benchmark, the three bars from left to right correspond to baseline (electronic multicore), Pho$, and Pho$_OCM, respectively.

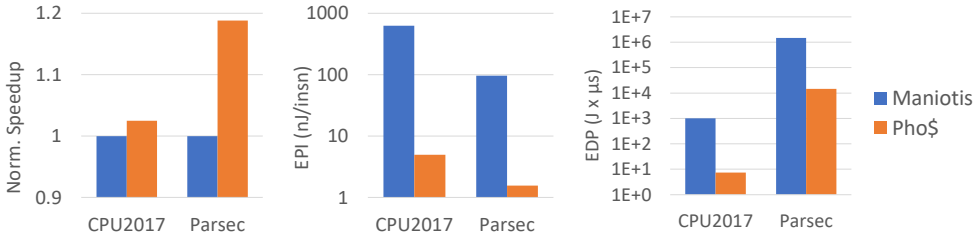


Fig. 12. Normalized speedup, energy per instruction (nJ/insn), and energy × delay product (J×µs) of Maniotis *et al.*'s optical cache compared with Pho$.

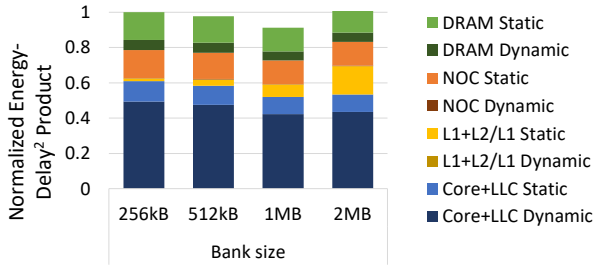


Fig. 13. Design space exploration of Pho$'s per-bank capacity.

performance increase despite a slower writing speed, while maintaining a two orders-of-magnitude lower energy consumption.

5.5 Capacity-Power Inflection Point

Pho$'s static power is dominated by the number of optical bit cells. To explore the capacity-power tradeoff, we compared the average EDP and energy delay squared product (ED^2P) of Pho$ with 256 kB, 512 kB, 1 MB, and 2 MB cache banks, shown in Figure 13. To strike a balance between capacity and power, 1 MB is currently the best design.

5.6 Iso-Area Comparison

We performed iso-area comparisons by giving electronic designs the same area as Pho$. We estimate that with the additional area each core tile can employ 3 MB more cache capacity. This extra cache capacity can be used either for L2 or for the L3 slice at each tile. We explore this design space and simulate L2 + NUCA L3 slice configurations of 1+4.375 MB, 2+3.275 MB, 3+2.375 MB and 4+1.375 MB, respectively, and adjust latencies. Pho$ achieves 1.27× speedup and 37% lower EDP over the best iso-area electronic configuration.

6 DISCUSSION

6.1 Cache Contention

While contention for the optical cache does happen, its detrimental effect (increasing the average cache access latency beyond 2 cycles) is relieved in Pho$. In our system, with 2 cycles per L1 cache access, even when 3 cores contend for the same bank at exactly the same time and serialize, one will observe a 2-cycle delay, one will observe a 4-cycle delay, and the third core will observe a 6-cycle delay, bringing the average effective cache access latency to 4 cycles per access. This exactly

matches the 4–5 cycle delay of modern electronic private L1s [53, 85]. Moreover, each bank has its own private optical subnetwork in Pho$Net, so requests to different banks do not need to arbitrate with each other, effectively cutting contention by a factor of 4. This makes severe contention a much rarer event: for Pho$ to have a 4-cycle average cache delay (i.e., double the Pho$ cache latency), all 4 subnets need to have 3 requests each arriving at the same time (i.e., 12 requests contenting for the cache each time). Even the most memory-intensive application in Parsec (*streamcluster*) does not generate that amount of traffic to the cache, as less than 43% of its instructions are load/stores [8] and the baseline IPC is a mere 0.579. Thus, cache contention rarely rises to a level that presents a problem.

Even in the case when the contention is so high as for the average access latency to exceed 4 cycles, the load and store queues on the core act as buffering for cache accesses, which allow the out-of-order core to continue executing speculatively past these contended memory operations, and hence most of the time the contention delay will be overlapped with useful computation and will not increase execution time. This is exactly how modern cores can absorb most of the delay of L1, L2 and L3 electronic caches, and Pho$ can take advantage of the same mechanism to hide the latency of severe cache contention, when it occurs.

6.2 Future Technology and Scalability

Higher core counts will require Pho$ to scale capacity and avoid contention. While a proper scaling study requires physical-level details that are beyond the scope of this paper, we can make educated guesses by drawing from prior work. Over the last 20 years optical memory cell footprints decreased by 12 orders of magnitude, compared to 3 for SRAM [2], and are fast converging to their electronic counterparts. The steep improvement slope shows little signs of a slowdown, and as this is still a nascent technology, it holds the potential to exceed them in the future as it matures. Scaling the capacity with a small area footprint can be further addressed by 3D-die stacking: the optical banks do not need to communicate with each other, only with waveguides, which can be facilitated by optical TSVs (connected to waveguides through micro-mirrors), stacked multilayer waveguides or 3D opto-electronic interconnects, which are being developed globally and have been demonstrated [14, 26, 55, 60, 61, 66, 70, 90, 92]. Optical TSVs are vital in supporting the integration of 3D stacked photonic chips, and optical TSVs that can achieve loss lower than 0.1 dB while carrying error-free operations at a bandwidth of up to 40 Gbit/s have been demonstrated and validated for fabrication [14].

If Pho$ is to be scaled to a 64-core design, there are several aspects of the design to consider. By increasing total core count from 16 to 64, there will need to be four times the number of cache banks, with each bank requiring its own optical subnet. Pho$Net's optical waveguides will also need to make four more turns on the core die to ensure all cores are connected. This ultimately will lead to longer waveguides, increasing both cache access latency and laser power. For example, with 16 cores in a 4×4 layout, the length of waveguides on the core die is approximately 10× the length of one core (Figure 4a). With 64 cores in an 8×8 layout, the total length of waveguides on the core die is 38× the length of one core, an almost 4× increase. Scaling up the total number of cores might also require a more complicated arbitration protocol, which will need additional optical components to function. It is also worth noting that a chiplet-based design [25] may fit well with a high-performance target, while leaving individual chiplets relatively smaller in size.

Other designs are also in principle compatible with Pho$. Pho$'s interconnect can easily connect chiplets on the same network or separate sub-networks per chiplet, and use multiported optical caches (Figure 9) to reduce contention or optical losses. Pho$ on SMPs may also be implemented individually within each socket.

6.3 Cost and CMOS Compatibility

Estimating the cost of Pho$ is very challenging at this point because, while these devices have been manufactured and characterized in research lab settings, they have not been manufactured at volume, so the economy-of-scale benefits and mature yield numbers are unknown. To the best of our knowledge, PhC cells are research devices that are not commercially available. It is important to note that in the design of Pho$ we assume separate electronic and photonic dies, which simplifies the design and reduces the associated costs. We emphasize that photonic/CMOS integration has been shown in the integration of high-speed optical modulators, optical waveguides, resonators, and sensitive avalanche photodetectors in bulk CMOS chips [4] and the manufacturing of a photonic-electronic processor [73]. The latter work adopted a "zero-change" approach to the integration of photonics. Instead of developing a custom process to enable the fabrication of photonics, which would complicate or eliminate the possibility of integration with state-of-the-art transistors at large scale and at high yield, the authors designed optical devices using a standard microelectronics foundry process that is used for modern microprocessors. Thus, there is proof-of-concept work showing that the photonic devices required on the logic die can be integrated with CMOS. The devices needed for the optical cache banks can be developed and optimized separately, as they are on a separate photonic die.

7 RELATED WORK

In the 1990s, Guilfoyle *et al.* [29] first introduced photonic random optical memory for faster and less power consuming accesses of the main memory and Chiarulli *et al.* [16] proposed an opto-electronic memory hierarchy for similar purposes. Pleros *et al.* [63], Alexoudi *et al.* [1], and Nozaki *et al.* [57] proposed using different materials to build optical SRAM cells with memory operations. Alexoudi *et al.* [3], Vagionas *et al.* [80], Maniotis *et al.* [49, 50], and Pleros *et al.* [64] present a series of work that propose physical-level optical cache architectures and integrating them into processors with simple interconnects. Fotouhi *et al.* [28] exploits silicon-photonic interconnects in chiplet-based systems to build uniform memory architectures. Pho$, in contrast, takes a more all-round perspective in that we not only re-engineer the configuration and architecture of the optical cache itself, but also leverage on-chip as well as off-chip optical interconnects to take as much advantage of optical cache's low latency, high capacity, and high energy-efficiency as possible.

Several optical NoC works try to integrate photonics into on-chip communication. Corona [82] implements an MWSR optical crossbar where nodes contend for an optical token before they are allowed to transmit data, allowing the arbitration of a shared channel. FireFly [59] partitions optical R-SWMR crossbars to connect clusters of electrical mesh networks. ProLaser [23] segregates the data channel and the control channel and manages them separately in order to save laser power. Pho$Net takes a leap by partitioning the network into separate subnets and segregating the request and reply data channels, which have their dedicated laser source. This makes for a more optimized and efficient distribution of power.

8 CONCLUSION

Recent discoveries of new materials and research on optical SRAM cells enable us to build fast, low-power optical cache architectures. In this paper we propose Pho$, an opto-electronic memory hierarchy architecture for multicores. Pho$ replaces private electronic L1 and L2 caches with a large shared optical cache, and on-chip electronic mesh networks with a novel optical NoC that uses a unique network arbitration protocol. We estimate that Pho$ is on average 1.41× faster and 31% more energy-efficient (in terms of EDP) over purely electronic designs with similar configurations. Assuming aggressive technology projections, Pho$'s network design, Pho$Net,

consumes 70% less power than previously proposed optical NoCs. We also solve a number of problems that make previous optical cache designs impractical, achieving a performance lead and two orders-of-magnitude lower energy consumption.

REFERENCES

[1] Theonitsa Alexoudi, Dimitrios Fitsios, Alexandre Bazin, Paul Monnier, Rama Raj, Amalia Miliou, George T. Kanellos, Nikos Pleros, and Fabrice Raineri. 2016. III-V-on-Si Photonic Crystal Nanocavity Laser Technology for Optical Static Random Access Memories. *IEEE J. Sel. Topics Quantum Electron.* 22, 6 (2016), 295–304. https://doi.org/10.1109/JSTQE.2016.2593636

[2] Theonitsa Alexoudi, George T. Kanellos, and Nikos Pleros. 2020. Optical RAM and integrated optical memories: a survey. *Light: Science & Applications* 9, 1 (2020), 1–16. https://doi.org/10.1038/s41377-020-0325-9

[3] Theonitsa Alexoudi, Sotiris Papaioannou, George T. Kanellos, Amalia Miliou, and Nikos Pleros. 2013. Optical Cache Memory Peripheral Circuitry: Row and Column Address Selectors for Optical Static RAM Banks. *J. Lightw. Tech.* 31, 24 (2013), 4098–4110. https://doi.org/10.1109/JLT.2013.2286529

[4] Amir H. Atabaki, Sajjad Moazeni, Fabio Pavanello, Hayk Gevorgyan, Jelena Notaros, Luca Alloatti, Mark T. Wade, Chen Sun, Seth A. Kruger, Huaiyu Meng, Kenaish Al Qubaisi, Imbert Wang, Bohan Zhang, Anatol Khilo, Christopher V. Baiocco, Miloš A. Popović, Vladimir M. Stojanović, and Rajeev J. Ram. 2018. Integrating photonics with silicon nanoelectronics for the next generation of systems on a chip. *Nature* 556, 7701 (2018), 349–354. https://doi.org/10.1038/s41586-018-0028-z

[5] Christopher Batten, Ajay Joshi, Jason Orcutt, Anatol Khilo, Benjamin Moss, Charles W. Holzwarth, Miloš A. Popovic, Hanqing Li, Henry I. Smith, Judy L. Hoyt, et al. 2009. Building Many-Core Processor-to-DRAM Networks with Monolithic CMOS Silicon Photonics. *IEEE Micro* 29, 4 (2009), 8–21. https://doi.org/10.1109/MM.2009.60

[6] Majed Valad Beigi and Gokhan Memik. 2016. TAPAS: Temperature-aware Adaptive Placement for 3D Stacked Hybrid Caches. In *2nd International Symposium on Memory Systems, MEMSYS 2016, Alexandria, VA, USA, October 3-6, 2016.* ACM, 415–426. https://doi.org/10.1145/2989081.2989085

[7] Christian Bienia. 2011. *Benchmarking Modern Multiprocessors.* Ph. D. Dissertation. Princeton University. https://parsec.cs.princeton.edu/publications/bienia11benchmarking.pdf

[8] Christian Bienia, Sanjeev Kumar, Jaswinder Pal Singh, and Kai Li. 2008. The PARSEC benchmark suite: characterization and architectural implications. In *17th International Conference on Parallel Architectures and Compilation Techniques, PACT 2008, Toronto, Ontario, Canada, October 25-29, 2008.* ACM, 72–81. https://doi.org/10.1145/1454115.1454128

[9] Shekhar Borkar and Andrew A. Chien. 2011. The future of microprocessors. *Comm. ACM* 54, 5 (2011), 67–77. https://doi.org/10.1145/1941487.1941507

[10] James Bucek, Klaus-Dieter Lange, and Jóakim von Kistowski. 2018. 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, 41–42. https://doi.org/10.1145/3185768.3185771

[11] Hyunil Byun, Jinkwon Bok, Kwansik Cho, Keunyeong Cho, Hanmei Choi, Jinyong Choi, Sanghun Choi, Sangdeuk Han, Seokyong Hong, Seokhun Hyun, et al. 2014. Bulk-Si photonics technology for DRAM interface. *Photonics Research* 2, 3 (2014), A25–A33. https://doi.org/10.1364/PRJ.2.000A25

[12] Trevor E. Carlson, Wim Heirman, and Lieven Eeckhout. 2011. 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, 52:1–52:12. https://doi.org/10.1145/2063384.2063454

[13] Trevor E. Carlson, Wim Heirman, Stijn Eyerman, Ibrahim Hur, and Lieven Eeckhout. 2014. An Evaluation of High-Level Mechanistic Core Models. *ACM Trans. Arch. Code Opt.* 11, 3 (2014), 28:1–28:25. https://doi.org/10.1145/2629677

[14] Sujay Charania, Niels Neumann, Sebastian Killge, Felix Winkler, Zaid Al-Husseini, Laszlo Szilagyi, Ronny Henker, Frank Ellinger, Dirk Plettemeier, and Johann W Bartha. 2020. Design, Fabrication, and Comparison of 3D Multimode Optical Interconnects on Silicon Interposer. *Journal of Lightwave Technology* 38, 13 (2020), 3454–3460.

[15] Guoqing Chen, Hui Chen, Mikhail Haurylau, Nicholas A. Nelson, David H. Albonesi, Philippe M. Fauchet, and Eby G. Friedman. 2007. Predictions of CMOS compatible on-chip optical interconnect. *Integration* 40, 4 (2007), 434–446. https://doi.org/10.1016/j.vlsi.2006.10.001

[16] Donald M. Chiarulli and Steven P. Levitan. 1996. Optoelectronic-cache memory system architecture. *Applied Optics* 35, 14 (1996), 2449–2456. https://doi.org/10.1364/AO.35.002449

[17] International Roadmap Committee. 2015. *International Technology Roadmap For Semiconductors 2.0 2015 Edition.* Technical Report. European and Japan and Korean and Taiwan and United States Semiconductor Industry Associations. https://www.semiconductors.org/wp-content/uploads/2018/06/0_2015-ITRS-2.0-Executive-Report-1.pdf

[18] Intel Corporation. 2020. *Intel 64 and IA-32 Architectures Software Developer's Manual Volume 3(3A, 3B, 3C & 3D): System Programming Guide.* Intel Corporation. https://software.intel.com/content/dam/develop/public/us/en/documents/

325384-sdm-vol-3abcd.pdf

[19] Johan De Gelas and Ian Cutress. 2017. Sizing Up Servers: Intel's Skylake-SP Xeon versus AMD's EPYC 7000 - The Server CPU Battle of the Decade? https://www.anandtech.com/show/11544/intel-skylake-ep-vs-amd-epyc-7000-cpu-battle-of-the-decade/13

[20] Dell Inc. 2012. *PowerEdge R710 Technical Guide.* Dell Inc. https://i.dell.com/sites/doccontent/business/solutions/engineering-docs/en/Documents/server-poweredge-r710-tech-guidebook.pdf Ver. 4.0.

[21] Yigit Demir and Nikos Hardavellas. 2014. EcoLaser: An Adaptive Laser Control for Energy Efficient On-Chip Photonic Interconnects. In *IEEE/ACM International Symposium on Low-Power Electronics and Design, ISLPED 2014, La Jolla, CA, USA, August 11-13, 2014.* ACM, 3–8. https://doi.org/10.1145/2627369.2627620

[22] Yigit Demir and Nikos Hardavellas. 2015. Parka: Thermally Insulated Nanophotonic Interconnects. In *9th International Symposium on Networks-on-Chip, NOCS 2015, Vancouver, BC, Canada, September 28-30, 2015.* ACM, 1–8. https://doi.org/10.1145/2786572.2786597

[23] Yigit Demir and Nikos Hardavellas. 2016. Energy-proportional photonic interconnects. *ACM Trans. Arch. Code Opt.* 13, 4 (2016), 54:1–54:26. https://doi.org/10.1145/3018110

[24] Yigit Demir and Nikos Hardavellas. 2016. SLaC: Stage laser control for a flattened butterfly network. In *2016 IEEE International Symposium on High Performance Computer Architecture, HPCA 2016, Barcelona, Spain, March 12-16, 2016.* IEEE Computer Society, 321–332. https://doi.org/10.1109/HPCA.2016.7446075

[25] Yigit Demir, Yan Pan, Seukwoo Song, Nikos Hardavellas, John Kim, and Gokhan Memik. 2014. Galaxy: a high-performance energy-efficient multi-chip architecture using photonic interconnects. In *2014 International Conference on Supercomputing, ICS'14, Muenchen, Germany, June 10-13, 2014.* ACM, 303–312. https://doi.org/10.1145/2597652.2597664

[26] Fuad E. Doany, Benjamin G. Lee, Daniel M. Kuchta, Alexander V. Rylyakov, Christian Baks, Christopher Jahnes, Frank Libsch, and Clint L. Schow. 2012. Terabit/Sec VCSEL-Based 48-Channel Optical Module Based on Holey CMOS Transceiver IC. *J. Lightw. Tech.* 31, 4 (2012), 672–680. https://doi.org/10.1109/JLT.2012.2217938

[27] Agner Fog. 2016. The microarchitecture of Intel, AMD and VIA CPUs: An optimization guide for assembly programmers and compiler makers, 2016. https://www.agner.org/optimize/microarchitecture.pdf

[28] Pouya Fotouhi, Sebastian Werner, Jason Lowe-Power, and SJ Ben Yoo. 2019. Enabling scalable chiplet-based uniform memory architectures with silicon photonics. In *International Symposium on Memory Systems, MEMSYS 2019, Washington, DC, USA, September 30 - October 03, 2019.* ACM, 222–334. https://doi.org/10.1145/3357526.3357564

[29] Peter S. Guilfoyle and Richard V. Stone. 1992. Photonic random optical memory access cache. In *Image Storage and Retrieval Systems*, Vol. 1662. International Society for Optics and Photonics, SPIE, 218–223. https://doi.org/10.1117/12.58506

[30] Xiaochen Guo, Engin Ipek, and Tolga Soyata. 2010. Resistive computation: avoiding the power wall with low-leakage, STT-MRAM based computing. In *37th International Symposium on Computer Architecture (ISCA 2010), June 19-23, 2010, Saint-Malo, France.* ACM, 371–382. https://doi.org/10.1145/1815961.1816012

[31] Amit Hadke, Tony Benavides, Rajeevan Amirtharajah, Matthew Farrens, and Venkatesh Akella. 2008. Design and evaluation of an optical CPU-DRAM interconnect. In *26th International Conference on Computer Design, ICCD 2008, 12-15 October 2008, Lake Tahoe, CA, USA.* IEEE Computer Society, 492–497. https://doi.org/10.1109/ICCD.2008.4751906

[32] Amit Hadke, Tony Benavides, S.J. Ben Yoo, Rajeevan Amirtharajah, and Venkatesh Akella. 2008. OCDIMM: Scaling the DRAM Memory Wall Using WDM Based Optical Interconnects. In *16th Annual IEEE Symposium on High Performance Interconnects (HOTI 2008), 26-28 August 2008, Stanford, CA, USA.* IEEE Computer Society, 57–63. https://doi.org/10.1109/HOTI.2008.25

[33] Parisa Khadem Hamedani, Natalie Enright Jerger, and Shaahin Hessabi. 2014. QuT: A low-power optical Network-on-Chip. In *Eighth IEEE/ACM International Symposium on Networks-on-Chip, NOCS 2014, Ferrara, Italy, September 17-19, 2014.* IEEE, 80–87. https://doi.org/10.1109/NOCS.2014.7008765

[34] Haiyang Han, Theoni Alexoudi, Chris Vagionas, Nikos Pleros, and Nikos Hardavellas. 2021. 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, 1–6. https://doi.org/10.1109/ISLPED52811.2021.9502487

[35] Wim Heirman, Trevor E. Carlson, Shuai Che, Kevin Skadron, and Lieven Eeckhout. 2011. 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, 38–49. https://doi.org/10.1109/IISWC.2011.6114195

[36] Djordje Jevdjic, Gabriel H. Loh, Cansu Kaynak, and Babak Falsafi. 2014. Unison Cache: A Scalable and Effective Die-Stacked DRAM Cache. In *47th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO 2014, Cambridge, United Kingdom, December 13-17, 2014.* IEEE Computer Society, 25–37. https://doi.org/10.1109/MICRO.2014.51

[37] Adwait Jog, Asit K. Mishra, Cong Xu, Yuan Xie, Vijaykrishnan Narayanan, Ravishankar Iyer, and Chita R. Das. 2012. Cache revive: Architecting volatile STT-RAM caches for enhanced performance in CMPs. In *The 49th Annual Design Automation Conference 2012, DAC '12, San Francisco, CA, USA, June 3-7, 2012.* ACM, 243–252. https://doi.org/10.1145/

2228360.2228406

[38] Yongsoo Joo, Dimin Niu, Xiangyu Dong, Guangyu Sun, Naehyuck Chang, and Yuan Xie. 2010. Energy- and endurance-aware design of phase change memory caches. In *Design, Automation and Test in Europe, DATE 2010, Dresden, Germany, March 8-12, 2010*. IEEE Computer Society, 136–141. https://doi.org/10.1109/DATE.2010.5457221

[39] Pranay Koka, Michael O McCracken, Herb Schwetman, Xuezhe Zheng, Ron Ho, and Ashok V Krishnamoorthy. 2010. Silicon-photonic network architectures for scalable, power-efficient multi-chip systems. In *37th International Symposium on Computer Architecture (ISCA 2010), June 19-23, 2010, Saint-Malo, France*. ACM, 117–128. https://doi.org/10.1145/1815961.1815977

[40] Eiichi Kuramochi, Kengo Nozaki, Akihiko Shinya, Koji Takeda, Tomonari Sato, Shinji Matsuo, Hideaki Taniyama, Hisashi Sumikura, and Masaya Notomi. 2014. Large-scale integration of wavelength-addressable all-optical memories on a photonic crystal chip. *Nature Photonics* 8, 6 (2014), 474–481. https://doi.org/10.1038/nphoton.2014.93

[41] Cheng Li, Mark Browning, Paul V. Gratz, and Samuel Palermo. 2014. LumiNOC: A Power-Efficient, High-Performance, Photonic Network-on-Chip. *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.* 33, 6 (2014), 826–838. https://doi.org/10.1109/TCAD.2014.2320510

[42] Sheng Li, Jung Ho Ahn, Richard D. Strong, Jay B. Brockman, Dean M. Tullsen, and Norman P. Jouppi. 2009. McPAT: an integrated power, area, and timing modeling framework for multicore and manycore architectures. In *42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-42 2009), December 12-16, 2009, New York, New York, USA*. ACM, 469–480. https://doi.org/10.1145/1669112.1669172

[43] Liu Liu, Rajesh Kumar, Koen Huybrechts, Thijs Spuesens, Günther Roelkens, Erik-Jan Geluk, Tjibbe De Vries, Philippe Regreny, Dries Van Thourhout, Roel Baets, and Geert Morthier. 2010. An ultra-small, low-power, all-optical flip-flop memory on a silicon chip. *Nature Photonics* 4, 3 (2010), 182–187. https://doi.org/10.1038/nphoton.2009.268

[44] Y. Liu, R. McDougall, J. Seoane, E. Kehayas, M. T. Hill, G. Maxwell, S. Zhang, R. Harmon, F.M. Huijskens, L. Rivers, et al. 2006. Characterization of Hybrid Integrated All-Optical Flip-Flop. In *LEOS 2006 - 19th Annual Meeting of the IEEE Lasers and Electro-Optics Society*. IEEE, 943–944. https://doi.org/10.1109/LEOS.2006.279158

[45] Gabriel H. Loh. 2009. Extending the effectiveness of 3D-stacked DRAM caches with an adaptive multi-queue policy. In *42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-42 2009), December 12-16, 2009, New York, New York, USA*. ACM, 201–212. https://doi.org/10.1145/1669112.1669139

[46] Gabriel H. Loh and Mark D. Hill. 2011. Efficiently enabling conventional block sizes for very large die-stacked DRAM caches. In *44th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO 2011, Porto Alegre, Brazil, December 3-7, 2011*. ACM, 454–464. https://doi.org/10.1145/2155620.2155673

[47] Niti Madan, Li Zhao, Naveen Muralimanohar, Aniruddha Udipi, Rajeev Balasubramonian, Ravishankar Iyer, Srihari Makineni, and Donald Newell. 2009. Optimizing communication and capacity in a 3D stacked reconfigurable cache hierarchy. In *15th International Conference on High-Performance Computer Architecture (HPCA-15 2009), 14-18 February 2009, Raleigh, North Carolina, USA*. IEEE Computer Society, 262–274. https://doi.org/10.1109/HPCA.2009.4798261

[48] Prasanth Mangalagiri, Karthik Sarpatwari, Aditya Yanamandra, VijayKrishnan Narayanan, Yuan Xie, Mary Jane Irwin, and Osama Awadel Karim. 2008. A low-power phase change memory based hybrid cache architecture. In *18th ACM Great Lakes Symposium on VLSI 2008, Orlando, Florida, USA, May 4-6, 2008*. ACM, 395–398. https://doi.org/10.1145/1366110.1366204

[49] Pavlos Maniotis, Dimitrios Fitsios, George T. Kanellos, and Nikos Pleros. 2013. Optical Buffering for Chip Multiprocessors: A 16GHz Optical Cache Memory Architecture. *J. Lightw. Tech.* 31, 24 (2013), 4175–4191. https://doi.org/10.1109/JLT.2013.2290741

[50] Pavlos Maniotis, Savvas Gitzenis, Leandros Tassiulas, and Nikos Pleros. 2016. An optically-enabled chip-multiprocessor architecture using a single-level shared optical cache memory. *Optical Switching and Networking* 22 (2016), 54–68. https://doi.org/10.1016/j.osn.2016.05.001

[51] Atar Mittal. 2020. What is a PCB Transmission Line? https://www.protoexpress.com/blog/pcb-transmission-line/

[52] Vikram K. Narayana, Shuai Sun, Abdel-Hameed A. Badawy, Volker J. Sorger, and Tarek A. El-Ghazawi. 2017. MorphoNoC: Exploring the design space of a configurable hybrid NoC using nanophotonics. *Microprocess. Microsystems* 50 (2017), 113–126. https://doi.org/10.1016/j.micpro.2017.03.006

[53] NASA. 2021. Skylake Processors. https://www.nas.nasa.gov/hecc/support/kb/skylake-processors_550.html

[54] Christopher J. Nitta, Matthew Farrens, and Venkatesh Akella. 2013. On-Chip Photonic Interconnects: A Computer Architect's Perspective. In *Synthesis Lectures on Computer Architecture*. Morgan & Claypool Publishers, 1–111. https://doi.org/10.2200/S00537ED1V01Y201309CAC027

[55] Akihiro Noriki, Isao Tamai, Yasuhiro Ibusuki, Akio Ukita, Satoshi Suda, Daisuke Shimura, Yosuke Onawa, Hiroki Yaegashi, and Takeru Amano. 2019. Optical TSV Using Si-Photonics Integrated Curved Micro-Mirror. In *2019 International 3D Systems Integration Conference (3DIC), Sendai, Japan, October 8-10, 2019*. IEEE, 1–4. https://doi.org/10.1109/3DIC48104.2019.9058779

[56] Kengo Nozaki, Akihiko Shinya, Shinji Matsuo, Tomonari Sato, Eiichi Kuramochi, and Masaya Notomi. 2013. Ultralow-energy and high-contrast all-optical switch involving Fano resonance based on coupled photonic crystal nanocavities. *Optics Express* 21, 10 (2013), 11877–11888. https://doi.org/10.1364/OE.21.011877

[57] Kengo Nozaki, Akihiko Shinya, Shinji Matsuo, Yasumasa Suzaki, Toru Segawa, Tomonari Sato, Yoshihiro Kawaguchi, Ryo Takahashi, and Masaya Notomi. 2012. Ultralow-power all-optical RAM based on nanocavities. *Nature Photonics* 6, 4 (2012), 248–252. https://doi.org/10.1038/nphoton.2012.2

[58] Yan Pan, John Kim, and Gokhan Memik. 2011. FeatherWeight: low-cost optical arbitration with QoS support. In *44th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO 2011, Porto Alegre, Brazil, December 3-7, 2011*. ACM, 105–116. https://doi.org/10.1145/2155620.2155633

[59] Yan Pan, Prabhat Kumar, John Kim, Gokhan Memik, Yu Zhang, and Alok N. Choudhary. 2009. Firefly: illuminating future network-on-chip with nanophotonics. In *36th International Symposium on Computer Architecture (ISCA 2009), June 20-24, 2009, Austin, TX, USA*. ACM, 429–440. https://doi.org/10.1145/1555754.1555808

[60] Mahavir S. Parekh, Paragkumar A. Thadesar, and Muhannad S. Bakir. 2011. Electrical, optical and fluidic through-silicon vias for silicon interposer applications. In *2011 IEEE 61st Electronic Components and Technology Conference (ECTC)*. IEEE, 1992–1998. https://doi.org/10.1109/ECTC.2011.5898790

[61] Chirag S. Patel, Cornelia K. Tsang, Christian Schuster, Fuad E. Doany, Harold Nyikal, Christian W. Baks, Russell Budd, L. Paivikki Buchwalter, Paul S. Andry, Donald F. Canaperi, et al. 2005. Silicon Carrier with Deep Through-Vias, Fine Pitch Wiring and Through Cavity for Parallel Optical Transceiver. In *Electronic Components and Technology, 2005. ECTC'05.* IEEE, 1318–1324. https://doi.org/10.1109/ECTC.2005.1441439

[62] Harish Patil, Robert Cohn, Mark Charney, Rajiv Kapoor, Andrew Sun, and Anand Karunanidhi. 2004. Pinpointing Representative Portions of Large Intel® Itanium® Programs with Dynamic Instrumentation. In *37th Annual International Symposium on Microarchitecture (MICRO-37 2004), 4-8 December 2004, Portland, OR, USA*. IEEE Computer Society, 81–92. https://doi.org/10.1109/MICRO.2004.28

[63] Nikos Pleros, Dimitrios Apostolopoulos, Dimitrios Petrantonakis, Christos Stamatiadis, and Hercules Avramopoulos. 2009. Optical Static RAM Cell. *IEEE Photon. Tech. Lett.* 21, 2 (2009), 73–75. https://doi.org/10.1109/LPT.2008.2008444

[64] Nikos Pleros, Pavlos Maniotis, Theonitsa Alexoudi, Dimitris Fitsios, Christos Vagionas, Sotiris Papaioannou, Konstantinos Vyrsokinos, and George T. Kanellos. 2014. Optical RAM-enabled cache memory and optical routing for chip multiprocessors: technologies and architectures. In *Optical Interconnects XIV*, Vol. 8991. International Society for Optics and Photonics, SPIE, 206–213. https://doi.org/10.1117/12.2042732

[65] Timothy Prickett Morgan. 2017. Drilling Down Into The Xeon Skylake Architecture. https://www.nextplatform.com/2017/08/04/drilling-xeon-skylake-architecture/

[66] Wesley D. Sacher, Jared C. Mikkelsen, Ying Huang, Jason C.C. Mak, Zheng Yong, Xianshu Luo, Yu Li, Patrick Dumais, Jia Jiang, Dominic J. Goodwill, Eric Bernier, Patrick Guo-Qiang Lo, and Joyce K. S. Poon. 2018. Monolithically Integrated Multilayer Silicon Nitride-on-Silicon Waveguide Platforms for 3-D Photonic Circuits and Devices. *Proc. IEEE* 106, 12 (2018), 2232–2245. https://doi.org/10.1109/JPROC.2018.2860994

[67] Jun Sakaguchi, Takeo Katayama, and Hitoshi Kawaguchi. 2010. High Switching-Speed Operation of Optical Memory Based on Polarization Bistable Vertical-Cavity Surface-Emitting Laser. *IEEE J. Quantum Electron.* 46, 11 (2010), 1526–1534. https://doi.org/10.1109/JQE.2010.2052590

[68] Rathijit Sen and David A. Wood. 2013. *Cache power budgeting for performance.* Technical Report TR1791. Univ. of Wisconsin-Madison, Computer Science Dept. http://digital.library.wisc.edu/1793/65385

[69] Alireza Shafaei, Yanzhi Wang, Xue Lin, and Massoud Pedram. 2014. FinCACTI: Architectural Analysis and Modeling of Caches with Deeply-Scaled FinFET Devices. In *IEEE Computer Society Annual Symposium on VLSI, ISVLSI 2014, Tampa, FL, USA, July 9-11, 2014*. IEEE Computer Society, 290–295. https://doi.org/10.1109/ISVLSI.2014.94

[70] Nicolás Sherwood-Droz and Michal Lipson. 2011. Scalable 3D dense integration of photonics on bulk silicon. *Optics Express* 19, 18 (2011), 17758–17765. https://doi.org/10.1364/OE.19.017758

[71] Sarabjeet Singh and Manu Awasthi. 2019. Memory Centric Characterization and Analysis of SPEC CPU2017 Suite. In *Proceedings of the 2019 ACM/SPEC International Conference on Performance Engineering, ICPE 2019, Mumbai, India, April 7-11, 2019*. ACM, 285–292. https://doi.org/10.1145/3297663.3310311

[72] Clinton W. Smullen, Vidyabhushan Mohan, Anurag Nigam, Sudhanva Gurumurthi, and Mircea R. Stan. 2011. Relaxing non-volatility for fast and energy-efficient STT-RAM caches. In *17th International Conference on High-Performance Computer Architecture, HPCA-17 2011, February 12-16, 2011, San Antonio, Texas, USA*. IEEE Computer Society, 50–61. https://doi.org/10.1109/HPCA.2011.5749716

[73] Chen Sun, Mark T. Wade, Yunsup Lee, Jason S. Orcutt, Luca Alloatti, Michael S. Georgas, Andrew S. Waterman, Jeffrey M. Shainline, Rimas R. Avizienis, Sen Lin, et al. 2015. Single-chip microprocessor that communicates directly using light. *Nature* 528, 7581 (2015), 534–538. https://doi.org/10.1038/nature16454

[74] Guangyu Sun, Xiangyu Dong, Yuan Xie, Jian Li, and Yiran Chen. 2009. A novel architecture of the 3D stacked MRAM L2 cache for CMPs. In *15th International Conference on High-Performance Computer Architecture (HPCA-15 2009), 14-18*

February 2009, Raleigh, North Carolina, USA. IEEE Computer Society, 239–249. https://doi.org/10.1109/HPCA.2009.4798259

[75] Zhenyu Sun, Xiuyuan Bi, Hai Li, Weng-Fai Wong, Zhong-Liang Ong, Xiaochun Zhu, and Wenqing Wu. 2011. Multi retention level STT-RAM cache designs with a dynamic refresh scheme. In *44th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO 2011, Porto Alegre, Brazil, December 3-7, 2011*. ACM, 329–338. https://doi.org/10.1145/2155620.2155659

[76] Christos A. Thraskias, Eythimios N. Lallas, Niels Neumann, Laurent Schares, Bert J. Offrein, Ronny Henker, Dirk Plettemeier, Frank Ellinger, Juerg Leuthold, and Ioannis Tomkos. 2018. Survey of Photonic and Plasmonic Interconnect Technologies for Intra-Datacenter and High-Performance Computing Communications. *IEEE Commun. Surveys Tuts.* 20, 4 (2018), 2758–2783. https://doi.org/10.1109/COMST.2018.2839672

[77] Christos Vagionas, Dimitris Fitsios, George T. Kanellos, Nikos Pleros, and Amalia Miliou. 2012. All Optical Flip Flop with two Coupled Travelling Waveguide SOA-XGM Switches. In *Conference on Lasers and Electro-Optics 2012*. Optical Society of America, JW4A.2. https://doi.org/10.1364/CLEO_AT.2012.JW4A.2

[78] Christos Vagionas, Dimitrios Fitsios, Konstantinos Vyrsokinos, George T. Kanellos, Amalia Miliou, and Nikos Pleros. 2014. XPM- and XGM-Based Optical RAM Memories: Frequency and Time Domain Theoretical Analysis. *IEEE J. Quantum Electron.* 50, 8 (2014), 1–15. https://doi.org/10.1109/JQE.2014.2330068

[79] Christos Vagionas, Stella Markou, George Dabos, Theonitsa Alexoudi, Dimitris Tsiokos, Amalia Miliou, Nikos Pleros, and George T. Kanellos. 2013. Optical RAM Row Access and Column Decoding for WDM-formatted optical words. In *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*. Optical Society of America, JW2A.56. https://doi.org/10.1364/NFOEC.2013.JW2A.56

[80] Christos Vagionas, Stelios Pitris, Charoula Mitsolidou, Jan Bos, Pavlos Maniotis, Dimitris Tsiokos, and Nikos Pleros. 2015. All-Optical Tag Comparison for Hit/Miss Decision in Optical Cache Memories. *IEEE Photon. Tech. Lett.* 28, 7 (2015), 713–716. https://doi.org/10.1109/LPT.2015.2505500

[81] Dana Vantrease, Nathan Binkert, Robert Schreiber, and Mikko H. Lipasti. 2009. Light speed arbitration and flow control for nanophotonic interconnects. In *42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-42 2009), December 12-16, 2009, New York, New York, USA*. ACM, 304–315. https://doi.org/10.1145/1669112.1669152

[82] Dana Vantrease, Robert Schreiber, Matteo Monchiero, Moray McLaren, Norman P. Jouppi, Marco Fiorentino, Al Davis, Nathan Binkert, Raymond G. Beausoleil, and Jung Ho Ahn. 2008. Corona: System Implications of Emerging Nanophotonic Technology. In *35th International Symposium on Computer Architecture (ISCA 2008), June 21-25, 2008, Beijing, China*. IEEE Computer Society, 153–164. https://doi.org/10.1109/ISCA.2008.35

[83] Sebastian Werner, Javier Navaridas, and Mikel Luján. 2018. A Survey on Optical Network-on-Chip Architectures. *Comput. Surveys* 50, 6 (2018), 89:1–89:37. https://doi.org/10.1145/3131346

[84] WikiChip. 2020. Skylake (client) - Microarchitectures - Intel. https://en.wikichip.org/wiki/intel/microarchitectures/skylake_(client)

[85] WikiChip. 2020. Skylake (server) - Microarchitectures - Intel. https://en.wikichip.org/wiki/intel/microarchitectures/skylake_(server)

[86] H.-S. Philip Wong, Simone Raoux, SangBum Kim, Jiale Liang, John P. Reifenberg, Bipin Rajendran, Mehdi Asheghi, and Kenneth E. Goodson. 2010. Phase Change Memory. *Proc. IEEE* 98, 12 (2010), 2201–2227. https://doi.org/10.1109/JPROC.2010.2070050

[87] Xiaoxia Wu, Jian Li, Lixin Zhang, Evan Speight, Ram Rajamony, and Yuan Xie. 2009. Hybrid cache architecture with disparate memory technologies. In *36th International Symposium on Computer Architecture (ISCA 2009), June 20-24, 2009, Austin, TX, USA*. ACM, 34–45. https://doi.org/10.1145/1555754.1555761

[88] William A. Wulf and Sally A. McKee. 1995. Hitting the memory wall: implications of the obvious. *ACM SIGARCH Comp. Arch. News* 23, 1 (1995), 20–24. https://doi.org/10.1145/216585.216588

[89] Lei Yi, Guangbao Shan, Song Liu, and Chengmin Xie. 2016. High-performance processor design based on 3D on-chip cache. *Microprocess. Microsystems* 47 (2016), 486–490. https://doi.org/10.1016/j.micpro.2016.07.009

[90] S.J. Ben Yoo, Binbin Guan, and Ryan P. Scott. 2016. Heterogeneous 2D/3D photonic integrated microsystems. *Microsystems & Nanoengineering* 2, 1 (2016), 1–9. https://doi.org/10.1038/micronano.2016.30

[91] Yang Zhang, Amir Hosseini, Xiaochuan Xu, David Kwong, and Ray T. Chen. 2013. Ultralow-loss silicon waveguide crossing using Bloch modes in index-engineered cascaded multimode-interference couplers. *Optics Letters* 38, 18 (2013), 3608–3611. https://doi.org/10.1364/OL.38.003608

[92] Yu Zhang, Anirban Samanta, Kuanping Shang, and S.J. Ben Yoo. 2020. Scalable 3D Silicon Photonic Electronic Integrated Circuits and Their Applications. *IEEE J. Sel. Topics Quantum Electron.* 26, 2 (2020), 1–10. https://doi.org/10.1109/JSTQE.2020.2975656