Silicon Photonics in Post Moore’s Law Era: Technological and Architectural Implications

Ke Wen*, Sébastien Rumley, Payman Samadi, Christine P. Chen, Keren Bergman

Department of Electrical Engineering, Columbia University, New York, *kw2501@columbia.edu

1 A RIPPLE EFFECT FROM CACHE REDUCTION

Moore’s law is ending as we enter the last years of shrinking transistors. Chip designers will thus have to use the available transistors more effectively. One may interpret a few already signs of this trend. As shown in Figure 1, area devoted to cache on a CPU chip is decreasing, both in terms of (a) MB per FLOPS and (b) normalized chip area -- cache size (MB) × features size (nm²) / die size (mm²). Especially, a sharp fall (Fig. 1a) is clear as the industry gets into the many-core era (around 2013). Interestingly, this cache cliff matches the time when Moore’s Law was said to be dead in the economic sense -- starting from 2013, the number of transistors bought per dollar has stayed stagnant [1]. The fact that chipmakers are unwillingly trading the cache area for more FLOPS, along with the rise of data-centric throughput computing [2], calls for significantly higher off-chip memory bandwidth. Fig. 1c shows this trend: the sharp increase of the off-chip memory bandwidth matches the cache cliff of Fig. 1a. This increase, however, is still not enough to balance the FLOPS increase as the bytes per flop ratio continues to drift away from the ideal point.

There is more to this grim description. The memory bandwidth increase is also rapidly stressing the pin count limit of the processor package. For example, KNL requires 3647 pins in the socket, plus 1024 pins in the interposer for each of the eight on-package memory stacks. The pin density of standard chip package, however, cannot scale indefinitely. The ever-increasing bandwidth demand thus requires a more efficient chip I/O technology for processors beyond the Moore’s Law.

2 SILICON PHOTONICS FOR CHIP I/O

Compatible with CMOS lithography fabrication, Silicon photonics (SiP) has become one of the leading solutions to the aforementioned chip I/O issue. An example of “extending the power of silicon to new arenas” [3], SiP leverages the transparency of silicon to light with 1.2−5 μm wavelength for high-speed transmission. Each SiP waveguide can support terabit/s bandwidth, orders of magnitude higher than what can be achieved with conventional electrical I/O. For example, while an 8-channel (4-layer) High Bandwidth Memory (HBM) cube requires a 1024-bit bus for 100 GB/s, a single SiP waveguide can provide the same bandwidth with 32 wavelengths each at 25 Gb/s.

Silicon photonics is compatible with silicon interposers used to carry processor and memory chips, forming a high-bandwidth chip-to-chip interconnect on package (Fig. 1d). Components such as waveguides, modulators, photodectors and switches can be directly fabricated on the silicon interposer with low cost. The SiP switch, controllable by the processor, can provide flexible and transparent connection between any memory stack and any processor interface. A SiP interposer fabricated by PEST of Japan was reported to achieve bandwidth density of 6.6 Tb/s/cm² [4].

Another important aspect of SiP is extending high-bandwidth I/O off package, enabled by efficient coupling
between waveguides and fibers. This is a much-needed capability, as the interposer area (about 700 mm2) will limit the capacity of on-package memory (OPM). The current solution is to pair the fast OPM with slow, off-package, DRAM. Such small-fast, large-slow exclusiveness may significantly complicate application programming and memory management. The distance-independent transmission of photons can solve this problem, enabling a uniform, high-capacity HBM architecture as shown in Fig. 1e. With 1 Tbit/s bandwidth per fiber, four fibers can supply the 256 GB/s bandwidth needed by a HBM2 cube. With 24 fibers per coupling assembly and four such assemblies, an interposer hosting processors can connect to a total of 24 HBM cubes, accounting for 192 GB memory capacity and 6 Tbit/s aggregate bandwidth. SiP technologies can thus enable a flat, easy-to-manage memory hierarchy.

3 ARCHITECTURAL IMPLICATIONS

3.1 Node Level: Optimizing Memory Locality

The benefit of silicon photonics is not limited to sheer bandwidth growth. As mentioned earlier, the reconfigurable SiP switch can provide connection between any memory cube and any processor interface. This functionality can help precisely deliver memory data to the consumer cores, without traversing the network on chip (NoC), effectively mitigating the NUMA problem faced by the many-core era [5] [6]. As shown in Fig. 1f, a reconfiguration of the SiP switch can reduce the NoC hop count from 10 (dashed yellow, as in native connection) to 1 (solid yellow). This hop decrease immediately translates into a few tens of nanoseconds less latency and a significant drop in energy dissipation. The routing of high-speed memory data out of the NoC plane may also save the NoC bandwidth for more core-to-core communication, a trend as “MPI everywhere” (assigning each core with a MPI process) emerges [7]. SiP waveguides with ultra-low loss of 1.2 dB/m has been demonstrated [8], meaning nearly distance-independent energy consumption for chip scale, as compared to 25 pJ per 64-bits per mm in case of moving data electrically on chip [2].

Another possibility is to use the SiP switch to alleviate the hotspot effect on the NoC when hotspot memory access happens (Fig. 1g). In this scenario, the SiP switch can TDM select the memory interface to inject data stream from the hotspot memory, thus distributing the traffic to different NoC sections [6].

3.2 System Level: Flexible Topology

SiP switching can be also utilized to form flexible system-level topology [9]. The need for flexible topology roots from the diverse spectrum of applications that run in a supercomputer. The clear difference in their communication characteristics, in terms of neighboring relationship, traffic volume, etc., makes it very difficult to find a “best-for-all” topology. SiP switching, in contrast, is capable of dynamically “rewiring” the connections among a set of electronic endpoints. These electronic endpoints can be either compute nodes or electrical routers. The benefit is directing bandwidth to where it is needed without over-provisioning it [10]. Recently, a reconfigurable Dragonfly architecture utilizing small-radix SiP switches has been demonstrated [11]. The architecture, called Flexfly, is capable of concentrating the fully-connected group-to-group links of Dragonfly into, for example, a thick ring-like topology (Fig. 1h). It is shown to help applications like GTC to achieve 1.8x speedup over conventional adaptive UGAL routing.

4 PHOTONIC-ELECTRONIC INTEGRATION

There are three methods for integrating SiP and electronic devices: front-end, back-end and hybrid integrations.

In front-end integration, electronic and photonic devices are formed on the same layer. The advantage is that nanophotonics can piggyback on the same mask. However, the challenge remains to guide light with sufficient isolation, especially, separation between the waveguide and the silicon substrate. While CMOS-SOI uses a thin buried oxide (BOX) of 200 nm, photonics SOI requires a BOX of 1 µm. Approaches that utilize thicker-BOX have been proposed [12, 13]. This may, however, reduce the heat dissipation capability of electronics [14]. Methods that do not modify the standard CMOS have thus been proposed [14-17], which locally remove the underlying Si substrate to mitigate losses.

Back-end integration is another monolithic method [18, 19]. It allows deposition of sufficient isolation oxide on top of the existing CMOS-SOI. However, this method introduces additional back-end steps and thus extra cost. The back-end method may also face a stricter thermal budget in order to prevent damage to electronic CMOS. As a result, engineers have to look at using other materials for the photonic layer. Yet, to date, silicon nitride [18], amorphous silicon [20] and laser-annealed polysilicon [18] have been proven as feasible material.

The hybrid integration method forms photonic and electronic circuits on separate chips and bonds them through flip-chip bonding. As such, the photonic and electronic chips can be each optimized using different process flows. Hybrid integration is to date the majority choice of SiP research and development parties. Signaling speeds of 25 Gb/s [4] and 50 Gb/s [21] have recently been demonstrated using flip-chip bonding.

7 CONCLUSION

The end of Moore’s Law comes at a time when efficient allocation of transistor real estate has become imperative for computing. The resulting cache reduction and the rise of data-centric throughput computing calls for efficient off-chip, off-package data movement. Silicon photonics could potentially be one of the promising solutions the computing world is looking for to continue performance growth. Yet, industry-level electronic-photonic integration, and system co-design are yet to be realized, along with reducing manufacturing costs. New architectural implications of silicon photonics at both node level and system level also require further investigation into how to enable new dimensions of performance improvement.
ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy Lawrence Berkeley National Laboratory under subcontract 7257488 and Sandia National Laboratories under contract PO 1319001.

REFERENCES