

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

Single-chip microprocessor that communicates directly using light.

### Permalink

<https://escholarship.org/uc/item/4dh1v4px>

### Journal

Nature, 528(7583)

### ISSN

0028-0836

### Authors

Sun, Chen  
Wade, Mark T  
Lee, Yunsup  
[et al.](#)

### Publication Date

2015-12-01

### DOI

10.1038/nature16454

### License

<https://creativecommons.org/licenses/by-nc-nd/4.0/> 4.0

Peer reviewed

# Single-Chip Microprocessor with Integrated Photonic I/O

Chen Sun<sup>\*1,2</sup>, Mark T. Wade<sup>\*3</sup>, Yunsup Lee<sup>\*1</sup>, Jason S. Orcutt<sup>\*2,4</sup>, Luca Alloatti<sup>2</sup>, Michael S. Georgas<sup>2</sup>, Andrew S. Waterman<sup>1</sup>, Jeffrey M. Shainline<sup>3,5</sup>, Rimas R. Avizienis<sup>1</sup>, Sen Lin<sup>1</sup>, Benjamin R. Moss<sup>2</sup>, Rajesh Kumar<sup>3</sup>, Fabio Pavanello<sup>3</sup>, Amir H. Atabaki<sup>2</sup>, Henry M. Cook<sup>1</sup>, Albert J. Ou<sup>1</sup>, Jonathan C. Leu<sup>2</sup>, Yu-Hsin Chen<sup>2</sup>, Krste Asanović<sup>1</sup>, Rajeev J. Ram<sup>2</sup>, Miloš A. Popović<sup>3</sup>, Vladimir M. Stojanović<sup>1</sup>

*\*Contributed equally to this work*

<sup>1</sup>*University of California, Berkeley, Berkeley, CA*

<sup>2</sup>*Massachusetts Institute of Technology, Cambridge, MA*

<sup>3</sup>*University of Colorado, Boulder, Boulder, CO*

<sup>4</sup>*Now at IBM T.J. Watson Research Center, Yorktown Heights, NY*

<sup>5</sup>*Now at National Institute of Science and Technology, Boulder, CO*

**Data transport across short-reach electrical wires is both bandwidth-density and power-density-limited, creating a performance bottleneck for semiconductor microchips in modern computer systems, from mobile phones to large-scale datacenters. Optical communications based on chip-scale electronic-photonic systems<sup>1-4</sup> enabled by silicon-based nanophotonic devices<sup>5</sup> can overcome these limitations<sup>6-8</sup>. However, combining electronics and photonics on the same chip has proved challenging due to microchip manufacturing conflicts between electronics and photonics. Consequently, current electronic-photonic chips<sup>9-11</sup> are limited to niche manufacturing processes and integrate only a few optical devices alongside simple**

**circuits. Here we report an electronic-photonic system-on-chip (SoC) integrating over 70 million transistors and 850 photonic components that work in concert to provide logic, memory, and interconnect functions, realizing a microprocessor chip that can optically communicate directly to the outside world for the first time. To integrate electronics and photonics at this scale, we adopt a *zero-change* approach to the integration of photonics. Instead of developing a custom process to enable the fabrication of photonics<sup>12</sup>, which complicates or eliminates the possibility of integration with state-of-the-art transistors at large scale and at high yield, we design optical devices directly within a standard microelectronics foundry process used for modern microprocessors<sup>13</sup> (Cell<sup>14</sup>, BlueGene/Q<sup>15</sup>, Power7<sup>16</sup>, etc.). We expect that this demonstration signals the beginning of an era of electronic-photonic SoCs with potential for transformative impact on computing system architecture, enabling a leap to new kinds of more powerful computers, from network infrastructure to datacenters and supercomputers.**

The electro-optic SoC (Figure 1) contains a dual-core RISC-V instruction set architecture<sup>17</sup> (ISA) microprocessor and an independent 1 MB bank of static random access memory used for memory. The on-chip electro-optic transceivers for data input/output (I/O) enable both the microprocessor and the memory to communicate directly to off-chip components using light, without the need for separate chips or components to host the optical devices. The chip was fabricated in a commercial high-performance 45 nm complementary metal-oxide semiconductor (CMOS) silicon-on-insulator (SOI) process<sup>18</sup>. No changes to the foundry process were necessary to accommodate photonics and all optical devices were designed to comply with the native process manufacturing rules. This *zero-change* integration enables high-performance transistors on the same chip as optics, reuse of all existing designs in the process, compatibility with electronics design tools, and manufacturing

in an existing high-volume foundry.

The process includes a crystalline Si (c-Si) layer which is patterned to form both the body of electronic transistors and the core of optical waveguides. A thin buried oxide (BOX) layer separates the c-Si layer from the silicon handle wafer (Extended Data Figure 1). As the BOX is <200 nm thick, light propagating in c-Si waveguides will evanescently leak into the silicon handle wafer, resulting in high waveguide loss. To resolve this, we perform selective substrate removal on the chips after electrical packaging to etch away the silicon handle under regions with optical devices (Extended Data Figure 2). We leave the silicon handle intact under the microprocessor and memory (which dissipate the most power) to allow a heat sink to be contacted, if necessary. Substrate removal has a negligible impact on the electronics<sup>13</sup> and the processor is completely functional even with a fully-removed substrate.

Silicon-germanium (SiGe) is present, though in low germanium mole fractions, in advanced CMOS processes for enhancing hole mobility and transistor performance via compressive strain engineering of p-channel transistors<sup>18</sup>. Selecting an 1180 nm wavelength band for the optical channel enables use of photodetectors (PDs) built using this SiGe<sup>19</sup>. Silicon is transparent at 1180 nm and no adverse effects are observed. At these wavelengths, the optical propagation loss in silicon strip waveguides is 4.3 dB/cm (losses at industry-standard wavelengths of 1300 nm and 1550 nm are 3.7 dB/cm and 4.6 dB/cm, respectively<sup>13</sup>). The receiver circuit<sup>20</sup> resolves photocurrent produced by the illuminated PD into digital ones and zeros. The receiver sensitivity in optical modulation amplitude (OMA) is -5 dBm for a better than  $10^{-12}$  bit-error-rate.

The electro-optic transmitter consists of an electro-optic modulator and its electronic driver. The

modulator is a 10  $\mu\text{m}$  diameter silicon microring resonator, coupled to a waveguide. We dope the structure with the n-well and p-well implants used for transistors to form radially extending p-n junctions, interleaved along the azimuthal dimension<sup>21, 22</sup>, taking the form of a “spoked ring”. The ring exhibits a sharp, notched filter optical transmission response, with a stop-band at the ring’s resonant wavelength ( $\lambda_0$ ). Applying a negative voltage across the junctions depletes the ring of free carriers (electron and hole concentrations), while a small positive voltage refills the carriers. A change in carrier concentration influences the ring waveguide’s index of refraction through the carrier plasma dispersion effect<sup>23</sup> which, in turn, shifts  $\lambda_0$ . Electro-optic modulation (on-off keying) is achieved by changing the voltage applied across the junction to move the  $\lambda_0$  stop-band in and out of the laser wavelength ( $\lambda_L$ ). The modulator has a loaded quality factor of approximately 10,000, and a voltage swing of only 1  $V_{pp}$  across the modulator achieves 6 dB on-to-off ratios at a 3 dB insertion loss for non-return-to-zero (NRZ) binary data. The low voltage, near-zero quiescent current, and low capacitance (15 fF, including wiring capacitance) result in an energy-efficient modulator driven by a standard CMOS logic inverter at gigabit data rates using the same 1 V nominal supply that powers digital electronics.

As a resonant device, the modulator is highly sensitive to c-Si layer thickness variations within and across SOI wafers<sup>24</sup> as well as spatially and rapidly temporally varying thermal environments created by the electrical components on the chip<sup>25, 26</sup>. Both effects cause  $\lambda_0$  to deviate from the design value, necessitating tuning circuitry. We embedded a 400  $\Omega$  resistive microheater inside the ring to efficiently tune  $\lambda_0$  and added a monitoring PD weakly coupled to the modulator drop port. When light resonates in the modulator ring, a small fraction of the light couples to and illuminates the PD. This generates photocurrent proportional to the amount of resonating light, which is

maximized when resonance  $\lambda_0$  is equal to laser wavelength  $\lambda_L$  (modulator is directly on resonance). Taking advantage of the densely integrated electronics, we designed a digital controller which monitors the photocurrent and controls the power to the microheater to keep  $\lambda_0$  locked to  $\lambda_L$  under thermal variations<sup>20</sup>. In the case where  $\lambda_0$  has a large offset to  $\lambda_L$ , such as during chip power-up, and no photocurrent feedback is available, the controller steps the power output of the heater to sweep  $\lambda_0$  to perform an initial alignment with  $\lambda_L$  to reach a state where there is sufficient photocurrent to begin the main feedback loop. The controller achieves initial lock within 7 ms and has a tracking time constant of 13  $\mu$ s after lock-on. This system provides up to 3 nm of  $\lambda_0$  change and can compensate temperature swings of 60 K<sup>20</sup>, aided by the superior thermal isolation afforded by selective substrate removal.

We use the direct chip-to-chip optical connectivity of the microprocessor chip to build a photonically-connected main memory system for the microprocessor (Figure 2). The microprocessor chip optically communicates to the 1 MB memory array located remotely on a second identical chip an arbitrary distance away. The microprocessor sends requests (a *read* or *write*), the memory address (location in memory to *read* or *write*), and write data (for *write* requests) via the microprocessor to memory (P2M) link. The memory to microprocessor (M2P) link returns read data for *read* requests. A field programmable gate array (FPGA) provides the peripheral functionality of a motherboard, completing a user controllable computer.

For both P2M and M2P links, the laser light first couples into an electro-optic transmitter; laser light arriving in a single-mode (SM) fiber couples into an on-chip waveguide through a vertical grating coupler (VGC). The optical modulator, driven by circuits, modulates light in the waveguide and imprints it with on-off keyed binary data from the source. The light then exits the

chip through a second vertical grating into an SM fiber bound for the other chip. Once there, the light couples into the receive site through a VGC, illuminates a receive PD, and is resolved back by the receiver circuit into binary data for the destination. The communication between the microprocessor and memory is full duplex. Both P2M and M2P links run at 2.5 Gb/s, providing an aggregate 5 Gb/s of memory bandwidth. The shown demonstration uses only one wavelength of light; each additional wavelength increases the memory bandwidth by 5 Gb/s for a total potential aggregate bandwidth of 55 Gb/s without the need to use additional fibers.

A single 1183 nm continuous wave (CW) off-chip solid-state laser acts as the light source, with output power split 50/50 to share it across both the P2M and M2P links. To overcome the 4 dB to 6 dB coupling losses through each VGC due to unoptimized grating couplers, we insert an optical amplifier, which provides about 9 dB of gain, to obtain sufficient optical power at the receiver to resolve the signal. Using the optimized VGCs with losses of 1.2 dB<sup>27</sup> that exist as standalone test devices elsewhere on the same chip would eliminate the need for optical amplifiers in future design iterations.

To verify functionality of the photonicly-connected memory in the computer, we ran a combination of terminal-based and graphical programs (excerpted in Figure 3). To run a program, the control FPGA first performs direct memory access (DMA) through the memory controller to write all of the program's instructions into memory. Once the program is fully loaded, the FPGA issues a *reset* signal to the processor and the processor begins execution of the program by fetching the first program instruction from memory (from address 0x00002000). During program execution, the processor writes and reads program data to and from memory, in addition to reading the instructions from the memory. The control FPGA handles the printing of terminal outputs and acts

as a display driver that reads from the frame buffer residing in memory to display a screen to the user. In all cases, the P2M and M2P optical links handle all communications to and from memory (which holds all the program instructions and data). We note that the processor clock frequency is locked to a 1-to-80 ratio of the aggregate P2M link bit rate (corresponding to a clock frequency of 31.25 MHz at 2.5 Gb/s) when demonstrating the processor using the optical link, the result of a decision which simplified engineering efforts during chip design. When operating in non-optical mode – by electrically communicating to the 1 MB bank of memory local to the same chip, or memory connected to the control FPGA by time-multiplexing memory data over the control interface – the processor can run at a maximum speed of 1.65 GHz. A demonstration of the system running these programs can be found in the supplementary video.

To evaluate the robustness of the optical links and ring tuning control against thermal perturbations, we create a synthetic processor power trace by changing the processor's voltage and frequency operating points (Figure 4) over a 1000 s period. The changes in processor power are representative of a processor's behavior as it runs different loads, affecting the chip temperature. The difference in temperature between the highest temperature and lowest temperature (processor at maximum and minimum power, respectively) is approximately 8 K. The thermal tuning circuitry controls the output of the microheater integrated with the ring modulator to keep the resonant device locked to the laser wavelength, keeping the link free of bit-errors despite changes in temperature produced by the processor. With the tuning circuitry disabled, the same link experiences a number of bit-errors depending on the processor power draw. The impact of thermal perturbations on the system during the execution of a program is shown in the supplementary video.



The demonstration of the first electronic-photonics microprocessor chip opens a transformative path for advances in very-large scale integrated circuit (VLSI) technology, by adding nanophotonics as a new design dimension. Tailoring photonic devices to be integrated directly with electronics in an advanced-node CMOS process enabled a fully-functioning electronic-photonics SoC to be produced in a high-volume electronics foundry. The level of integration allowed on-chip thermal tuning control systems to guarantee robust operation of compact and energy efficient, but also thermally sensitive, optical resonator devices, addressing one of the key remaining challenges for nanophotonic circuits adoption in VLSI technology.

## References

1. Vantrease, D. *et al.* Corona: System implications of emerging nanophotonic technology. In *Proceedings of the 35th Annual International Symposium on Computer Architecture, ISCA '08*, 153–164 (IEEE Computer Society, 2008).
2. Shacham, A., Bergman, K. & Carloni, L. P. Photonic networks-on-chip for future generations of chip multiprocessors. *Computers, IEEE Transactions on* **57**, 1246–1260 (2008).
3. Batten, C. *et al.* Building manycore processor-to-DRAM networks with monolithic CMOS silicon photonics. *Micro, IEEE* **PP**, 1 (2009).
4. Beamer, S. *et al.* Re-architecting DRAM memory systems with monolithically integrated silicon photonics. In *International Symposium on Computer Architecture*, 129–140 (ACM, New York, NY, USA, 2010).
5. Xu, Q., Schmidt, B., Pradhan, S. & Lipson, M. Micrometre-scale silicon electro-optic modulator. *Nature* **435**, 325–327 (2005).
6. Goodman, J. W., Leonberger, F. J., Kung, S.-Y. & Athale, R. A. Optical interconnections for VLSI systems. *Proceedings of the IEEE* **72**, 850–866 (1984).
7. Miller, D. A. Rationale and challenges for optical interconnects to electronic chips. *Proceedings of the IEEE* **88**, 728–749 (2000).
8. Young, I. *et al.* Optical I/O technology for tera-scale computing. *Solid-State Circuits, IEEE Journal of* **45**, 235–248 (2010).
9. Narasimha, A. *et al.* A 40-Gb/s QSFP optoelectronic transceiver in a 0.13  $\mu\text{m}$  CMOS silicon-on-insulator technology. In *Optical Fiber Communication Conference, OMK7* (Optical Society of America, 2008).
10. Assefa, S. *et al.* CMOS integrated nanophotonics: Enabling technology for exascale computing systems. In *Optical Fiber Communication Conference, OMM6* (Optical Society

- of America, 2011).
11. Buckwalter, J., Zheng, X., Li, G., Raj, K. & Krishnamoorthy, A. A monolithic 25-Gb/s transceiver with photonic ring modulators and Ge detectors in a 130-nm CMOS SOI process. *Solid-State Circuits, IEEE Journal of* **47**, 1309–1322 (2012).
  12. Dupuis, N. *et al.* 30Gbps optical link utilizing heterogeneously integrated III-V/Si photonics and CMOS circuits. In *Optical Fiber Communications Conference and Exhibition (OFC), 2014*, 1–3 (2014).
  13. Orcutt, J. S. *et al.* Open foundry platform for high-performance electronic-photonic integration. *Opt. Express* **20**, 12222–12232 (2012).
  14. Takahashi, O. *et al.* Migration of Cell broadband engine from 65nm SOI to 45nm SOI. In *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, 86–597 (IEEE, 2008).
  15. IBM Blue Gene team. Design of the IBM Blue Gene/Q Compute chip. *IBM Journal of Research and Development* **57**, 1:1–1:13 (2013).
  16. Wendel, D. *et al.* The implementation of POWER7 (TM): A highly parallel and scalable multi-core high-end server processor. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International*, 102–103 (IEEE, 2010).
  17. Waterman, A., Lee, Y., Patterson, D. A. & Asanović, K. The RISC-V Instruction Set Manual, Volume I: User-Level ISA, Version 2.0. *EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2014-54* (2014).
  18. Narasimha, S. *et al.* High performance 45-nm SOI technology with enhanced strain, porous low-k BEOL, and immersion lithography. In *Electron Devices Meeting, 2006. IEDM '06. International*, 1–4 (2006).
  19. Alloatti, L., Srinivasan, S., Orcutt, J. & Ram, R. Waveguide-coupled detector in zero-change

- complementary metal–oxide–semiconductor. *Applied Physics Letters* **107**, 041104 (2015).
20. Sun, C. *et al.* A 45nm SOI monolithic photonics chip-to-chip link with bit-statistics-based resonant microring thermal tuning. In *VLSI Circuits Digest of Technical Papers, 2015 Symposium on* (2015).
  21. Shainline, J. M. *et al.* Depletion-mode carrier-plasma optical modulator in zero-change advanced CMOS. *Optics letters* **38**, 2657–2659 (2013).
  22. Wade, M. T. *et al.* Energy-efficient active photonics in a zero-change, state-of-the-art CMOS process. In *Optical Fiber Communication Conference, Tu2E.7* (Optical Society of America, 2014).
  23. Soref, R. A. & Bennett, B. Electrooptical effects in silicon. *Quantum Electronics, IEEE Journal of* **23**, 123–129 (1987).
  24. Selvaraja, S., Bogaerts, W., Thourhout, D. V. & R.Baets. Fabrication of uniform photonic devices using 193nm optical lithography in silicon-on-insulator. *Proc. 14th Eur. Conf. Integr. Opt. (ECIO)* (2008).
  25. Padmaraju, K., Chan, J., Chen, L., Lipson, M. & Bergman, K. Thermal stabilization of a microring modulator using feedback control. *Opt. Express* **20**, 27999–28008 (2012).
  26. Sun, C. *et al.* A monolithically-integrated chip-to-chip optical link in bulk CMOS. *Solid-State Circuits, IEEE Journal of* **50**, 828–844 (2015).
  27. Wade, M. *et al.* 75% efficient wide bandwidth grating couplers in a 45 nm microelectronics CMOS process. In *Optical Interconnects Conference, 2015 IEEE* (2015).

## Methods

**Chip Implementation.** The key chip characteristics are summarized in Extended Data Table 1. Photonic devices were prepared in Cadence Virtuoso (an industry-standard design tool for frontend electronics) in conjunction with mixed-signal electronics<sup>28</sup>. Digital electronics were implemented using a combination of digital synthesis and place and route tools from Synopsys and Cadence. All photonic and electronic designs conform to the CMOS manufacturing rules (more than 5000 rules) of IBM's commercial 45 nm thin-BOX SOI process (12SOI), with physical verification performed using Mentor Graphics Calibre.

**Chip Fabrication.** The chips are fabricated through the standard 12SOI process flow. We submit our design for mask aggregation through the Trusted Access Program Office (TAPO) shuttle run, with the chip mask set treated as if it were an ordinary electronics design. We note that physical design dimensions, including the cross-sectional layer type and thickness information explicitly not reported in this work, are provided as part of the standard electronic design kit that is made available to IBM foundry customers under a non-disclosure agreement. A subset of process and performance information regarding this process can be found in various official IBM publications on electronic CMOS process development<sup>18, 29, 30</sup>.

**Electrical Packaging.** The chips from the foundry are bumped with controlled-collapse chip connection (C4) solder balls. The chips are then flip-chip mounted (the chip's substrate is exposed on top) to an 8-layer FR4 printed circuit board (PCB) through C4 solder reflow. This forms all 249 electrical connections (including power and ground) from the chip to the PCB. Epoxy encapsulation is added to the mounted chips for additional mechanical support and to protect the

mounted chips. These steps are typical for an electrical chip package and were performed by CVInc.

**Patterned Substrate Removal of a Packaged Chip.** The electrically packaged samples are first backside-ground to thin the chip substrate down to 100  $\mu\text{m}$  to 150  $\mu\text{m}$  (performed by Aptek Industries). We then clean the backside surface with isopropyl alcohol and an  $\text{N}_2$  air gun. We next apply Kapton tape over the substrate regions that we do not wish to remove (over the processor and the DRAM emulator bank). Afterwards, the chips are placed in a chamber which supplies  $\text{XeF}_2$  gas to isotropically etch the silicon substrate, removing it as the volatile product  $\text{SiF}_4$ . We use a pulsed-etch technique, where etch steps of 120 s were interleaved with 60 s periods where we pump out the reaction products. The pressure used in the chamber is 3.4 Torr. As electronics are unaffected by the substrate removal, the very coarse feature definition provided by tape and hand-alignment is sufficient. On average, the substrate removal process takes 10-30 cycles (depending on the thickness after the backside-grind) with a success yield of 80% (defined as having a working processor after substrate removal). We stop the etch when the substrate over the desired etch region has disappeared when inspected by eye. The steps above are easily implementable at wafer-scale in high-volume manufacturing using standard photolithographic techniques<sup>31</sup>, which can also improve uniformity and yield of the post-processing as well as the resolution and alignment of the etch regions.

**Optical Testing.** The 1183 nm laser is a quantum dot DFB laser available from QDLaser. We use lensed fibers available from Oz Optics with a spot size of 5  $\mu\text{m}$  and a working distance of 26  $\mu\text{m}$  to couple light into the vertical grating couplers through the chip backside (after substrate removal). The spot size is matched to the the 5  $\mu\text{m}$  mode-field diameter of the vertical grating

couplers. We use 3-axis positioner stages (Thorlabs NanoMax) to position and align fibers over the grating couplers of the test sites. The shown demonstrations require a total of 3 fibers coupled to each chip. Minimum fiber-to-coupler insertion loss was achieved by angling the fibers at  $19^\circ$  off-normal from the chip's surface. To adjust the polarization of the input light, we use 3-paddle manual polarization controllers from Thorlabs (though these can be avoided if using polarization-maintaining fibers). For this first demonstration, we chose the manual fiber alignment approach to freely couple into any of the hundreds of optical test sites located throughout the chip. To make a permanent fiber-attach, we can leverage commercial optical packaging techniques for vertical grating couplers, such as through horizontal fiber array blocks with angle-cleaved fibers<sup>32</sup> or through vertical fiber array pigtailed<sup>9,33</sup>.

**Processor Testing.** The control FPGA is a Zedboard FPGA, providing an intermediate hardware interface between the processor's electrical links and an ethernet connection to the laboratory control computer. The individual cores incorporate a 64-bit scalar core, floating-point unit, vector accelerator, and private caches<sup>34</sup>. Programs are compiled from C source code using a *gcc*-based C compiler targeted for the RISC-V ISA. The implementations of the RISC-V processor and the software compilation stack are available at *bar.eecs.berkeley.edu*. Details of the RISC-V ISA standard are found at *www.riscv.org*. The full system is stable and can execute an arbitrary number of programs. We list a representative set of programs tested on this processor as follows:

- *memory test* – the control FPGA writes to and reads from every location in memory through direct memory access (DMA) to verify that the memory interface is fully functional and that all bits are correct. The processor is idled for this test.

- *hello world!* – program which asks the processor to print out a single line of text to the terminal, which is sent to the control FPGA to be displayed to the user.
- *STREAM* – a popular memory benchmarking application<sup>35</sup>, the program's outputs are printed to the terminal and displayed to the user.
- *teapot renderer* – a program which pixel shades a 3D teapot using the Blinn-Phong shading model and outputs the rendered image. The location and color of the light source illuminating the teapot in the rendered image can be controlled by the user using keyboard. The processor performs all calculations and writes the image to the frame buffer in memory using the optical links. It then reads the content of the frame buffer over the optical link and sends it to the control FPGA to display it as an image to the user.
- *linux* – a full Linux operating system. Once linux boots, the user is free to run any program, including running *python*, *top*, or file system operations (the file system behavior is coordinated by the control FPGA). This test uses memory connected to the control FPGA and not the optically-connected memory as the memory footprint of the Linux kernel is too big to fit in the 1 MB memory bank.

The 1-to-80 ratio between processor clock frequency and the P2M link throughput was chosen to keep processor frequency reasonable if the links operated at higher data rates than anticipated at design time and when all wavelengths in the P2M link are active. For example, if the P2M link supported an 80 Gb/s aggregate data rate, the processor needs to operate at 1.0 GHz, which is well within its abilities. Alternatively, if the ratio was 1-to-10, the processor would need to operate at 8



GHz, which is impractical.

**Transmitter and Receiver Circuit Specifications.** At the 2.5 Gb/s operating point used in the demonstration, the transmitter uses the 1 V digital supply, corresponding to a transmitter energy of 20 fJ/bit and achieving 3 dB insertion loss at 6 dB on-off ratio for non-return-to-zero (NRZ) binary data. The modulator is effectively “driverless” insofar as no analog driver electronics are needed to bridge between digital logic and the optical modulator due to the efficiency of the latter. The thermal tuning for the modulator ring consumes a fixed 192 fJ/bit for the control circuit and 0 mW to 2.5 mW for the heater power, dependent on the tuned range (for the 1.5 mW heater output power in the supplemental video, this corresponds to 600 fJ/bit). More detailed transmitter and thermal tuner descriptions have been previously reported<sup>20</sup>. The receiver has a  $10^{-12}$  bit-error-rate sensitivity (OMA) of  $-5$  dBm up to 5 Gb/s, degrading to  $-3.8$  dBm at 8 Gb/s, and  $-0.8$  dBm at 10 Gb/s. At 2.5 Gb/s, the receiver energy efficiency is 496 fJ/b, improving to 297 fJ/b at 10 Gb/s. Summing up, we report a total circuit energy efficiency of 1.3 pJ/b at 2.5 Gb/s, a power consumption of 3.25 mW. The bandwidth density of the transceivers is approximately 300 Gb/s/mm<sup>2</sup> of chip area. The key specifications are summarized in Extended Data Table 2.

**Link Specifications.** In the 2.5 Gb/s P2M and M2P links used in the demo, the transmitter input VGC, the transmitter output VGC, and receiver input VGC contribute 4 dB, 4 dB, and 6 dB of link insertion loss, respectively. The 1183 nm laser outputs 9.2 dBm such that 5.2 dBm (50/50 split, with an approximate 1 dB excess loss of the splitter) is incident upon each of the input transmit VGC of each link (P2M and M2P). At this laser power level, the OMA of each transmitter is  $-7$  dBm, with an average optical power of  $-9$  dBm. Each amplifier adds 9 dB of optical gain, completing the P2M and M2P links each with an extra 1 dB link margin. A chip iteration

incorporating 1.2 dB loss VGCs<sup>27</sup> into the P2M and M2P link would remove 10.4 dB of excess insertion loss. These devices were high-risk test structures on the current chip and so were not placed in the P2M and M2P transceivers. Using these couplers at the same input laser power level as before, both links could complete, without an amplifier, with an extra 2.4 dB of link margin. The 1183 nm laser was made by QDLaser and uses 55 mA of pump current at a laser diode bias of 1.3 V to output 9.2 dBm (8.3 mW) of power. This corresponds to a power use of 71.5 mW and a wall-plug efficiency of 11.6 %. The laser is shared across both P2M and M2P links and the total wall-plug energy-efficiency (laser and circuit) is 15.6 pJ/bit. The laser has a threshold of 29 mA and a slope efficiency of 0.32 mW/mA.

**Potential for Improved Performance.** It is important to note that the chip is a first working research prototype, and the current achieved performance is by no means representative of the absolute performance limits of this technology. We describe a few known ways to improve performance as follows:

- The current modulator design used a mid-level p-implant ( $10^{17} \text{ cm}^{-3}$  to  $10^{18} \text{ cm}^{-3}$ ) for p-contacts as opposed to a p+ implant, creating high series contact resistance which limited its bandwidth. Future design iterations will use the p+ implant to improve device bandwidth. Moreover, the modulators used only two out of a number of different doping implants available in the process for different transistors and transistor thresholds. Substantial improvements may be possible with other available implants.
- The current detector is absorption-length limited<sup>19</sup> and resonating the detector can improve sensitivity without a device size increase. Resonant detectors, implemented as a spoked-ring cavity in a manner similar to the modulator, exist on the same chip as standalone

devices in the independent device and transceiver regions. If incorporated with processor and memory transceivers in a future chip, they would improve sensitivity by  $\approx 6$  dB (to  $-11$  dBm OMA sensitivity), which would be competitive with state-of-the-art receivers. In addition, the current receiver circuit design is very conservative and could be optimized to further improve the sensitivity by 6 dB. The circuit could also be placed closer to the photo-detector to minimize wiring capacitance.

- The demonstration uses the laser at a power level far below that for peak efficiency, which is 16 % at 30 mW. Operation of the current laser at the peak-efficiency power level and sharing the output power across multiple links on the chip or, alternatively, usage of a laser optimized for the given output power are both techniques for improving energy-efficiency of the link, even without any device improvements.

**Applicability to CMOS Processes with Bulk Silicon Substrates.** CMOS processes utilizing a bulk silicon substrate lack a patternable crystalline silicon layer, motivating alternative devices in polysilicon and a small number of process changes<sup>36</sup>. However, some guiding principles of zero-change integration, such as reuse of existing transistor mask levels, repurposing of transistor materials for optics, and compact integration through silicon microrings, can be applied to minimize changes to the process frontend, which are the most harmful to process-native electronics. These concepts have been applied successfully in practice to enable functional photonics in bulk<sup>26, 36</sup>, though at far smaller scale.

## References for Methods

28. Orcutt, J. S. & Ram, R. J. Photonic device layout within the foundry CMOS design environment. *Photonics Technology Letters, IEEE* **22**, 544–546.
29. Kalla, R., Sinharoy, B., Starke, W. J. & Floyd, M. Power7: IBM’s next-generation server processor. *IEEE micro* **30**, 7–15 (2010).
30. Lee, S. *et al.* Record RF performance of 45-nm SOI CMOS technology. In *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, 255–258 (IEEE, 2007).
31. Roger, A. Breaking a new sound barrier; it’s a mic-on-a-chip. *Electronic Design* **54** (2006).
32. Pavarelli, N., Lee, J. S. & O’Brien, P. A. Packaging challenges for integrated silicon photonic circuits. In *SPIE Photonics Europe*, 91330F–91330F (International Society for Optics and Photonics, 2014).
33. Kopp, C. *et al.* Silicon photonic circuits: on-CMOS integration, fiber optical coupling, and packaging. *Selected Topics in Quantum Electronics, IEEE Journal of* **17**, 498–509 (2011).
34. Lee, Y. *et al.* A 45nm 1.3GHz 16.7 Double-Precision GFLOPS/W RISC-V Processor with Vector Accelerators. In *European Solid State Circuits Conference, 2104. ESSCIRC 2014*, 199–202 (2014).
35. McCalpin, J. D. Stream: Sustainable memory bandwidth in high performance computers (1995).
36. Meade, R. *et al.* Integration of silicon photonics in bulk CMOS. In *VLSI Technology, 2014. Digest of Technical Papers. 2014 Symposium on*, 228–229 (2014).

**Figure 1: The electro-optic SoC.** **a.** Die photo of the 3 mm × 6 mm chip showing locations and relative sizes of the processor, memory, and transceiver banks, imaged from the chip’s backside. **b.** The processor transmitter and receiver banks (the memory transmitter and receiver banks are identical) with zoom-ins of individual transmitters and receivers sites. **c.** Micrographs of the grating coupler, photodetector, and resonant microring modulators.

**Figure 2: Optical memory system block diagram.** The system uses one chip acting as the processor and the other acting as *memory*, connected together by a full duplex optical link with a round-trip distance of 20 m by fiber.

**Figure 3: Processor optical demonstration.** **a.** Program loading and execution **b.** Successful execution of the *Hello World!* basic functionality test and the *STREAM*<sup>35</sup> memory benchmark, two examples of terminal-based programs. **c.** Screen capture of the output of a 3D teapot rendering application running on the processor.

**Figure 4: Thermal tuning stress test of the P2M link.** **a.** Modulator heater output power with tuning switched on overlaid on top of the power trace for the processor. The thermal tuning controller changes the heater power output to adapt to the changes in temperature created by the changes in processor power. **b.** Measured bit errors per second vs. time with the thermal tuning controller switched *on* and *off* overlaid on top of the power trace for the processor. The link with the tuning controller *on* has no bit errors over the entire interval (a total of 2.5 Terabits transmitted and received).

**Extended Data Figure 1: Chip cross-section.** **a.** Full chip cross-section (not drawn to scale) from the silicon substrate to C4 solder balls, showing the structures of electrical transistors, waveguides, and contacted optical devices. The minimum separation between transistors and waveguides is  $<1$   $\mu\text{m}$ , set only by the distance at which evanescent light from the waveguide begins to interact with the transistor's structures. **b.** TEM cross-section micrograph of an optical waveguide, prior to substrate removal.

**Extended Data Figure 2: Selective substrate removal.** **a.** Selective substrate removal steps for the flip-chip packaged chip, using tape as a coarse mask for defining areas that retain the substrate. **b.** Photo of a selective-substrate-removed fully electrically packaged electronic-photonic processor chip.

**Extended Data Table 1:** Chip characteristics summary

**Extended Data Table 2:** Transceiver performance summary

## **Supplementary Information**

**Supplementary Video** The supplementary video (7 minutes 31 seconds) provides an animated overview of the chip (starts at 0:05) and description of the test setup for the optical memory (starts at 0:59). The demonstration of the processor running programs starts at 2:36, with temperature changing events applied starting at 5:46, with and without the thermal tuning circuit enabled.

**Acknowledgements** We thank Stephen Twigg, Quan Nguyen, and Miquel Moreto Planas for help with processor infrastructure as well as Ashwyn Srinivasan for help with photodetector characterization. We would also like to thank Sangyoon Han for helping with chip photos. This work was supported by DARPA POEM award HR0011-11-C-0100, led by Dr. Jagdeep Shah, and DARPA PERFECT award HR0011-12-2-0016, led by Dr. Joseph Cross. We also extend our sincere appreciation to Matt Casper, Jerry Torneden and the team at the Kansas City Plant for their support of our design submissions over the many years leading up to this work. Support is also acknowledged from the Berkeley Wireless Research Center, UC Berkeley ASPIRE Lab, MIT CICS, National Science Foundation, FCRP IFC, Trusted Foundry, Intel, Santec, and NSERC. The views expressed are those of the authors and do not reflect the official policy or position of the DoD or the U.S. Government.

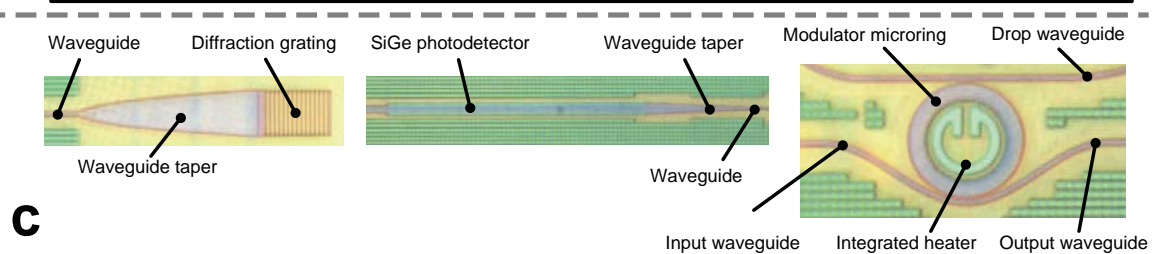
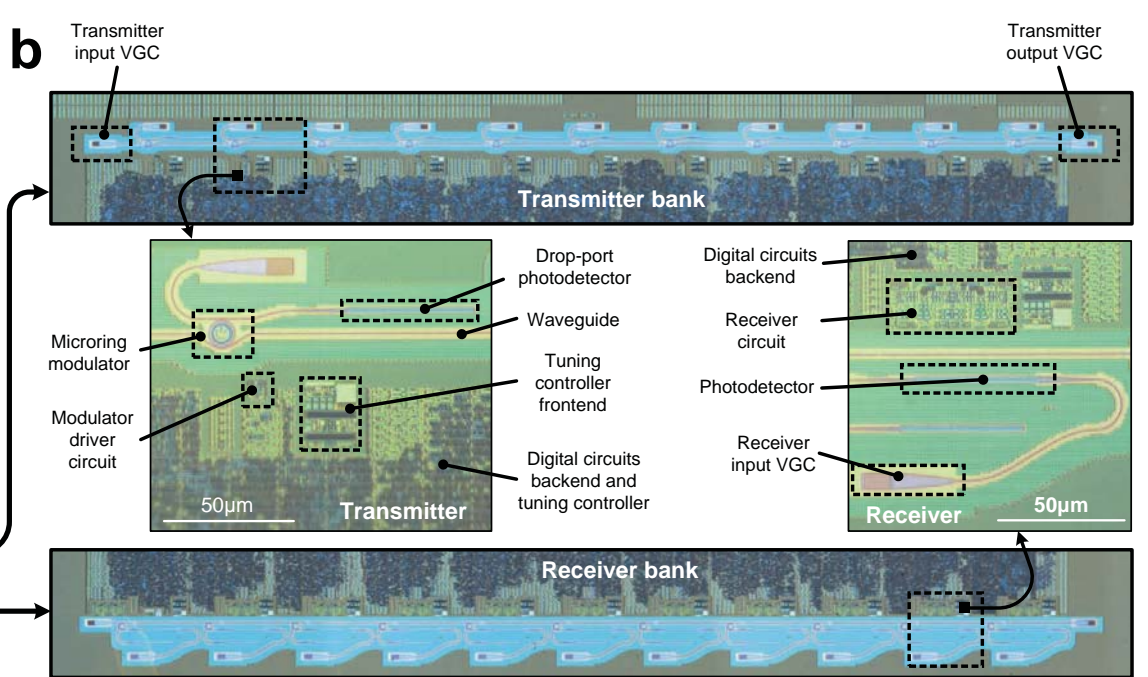
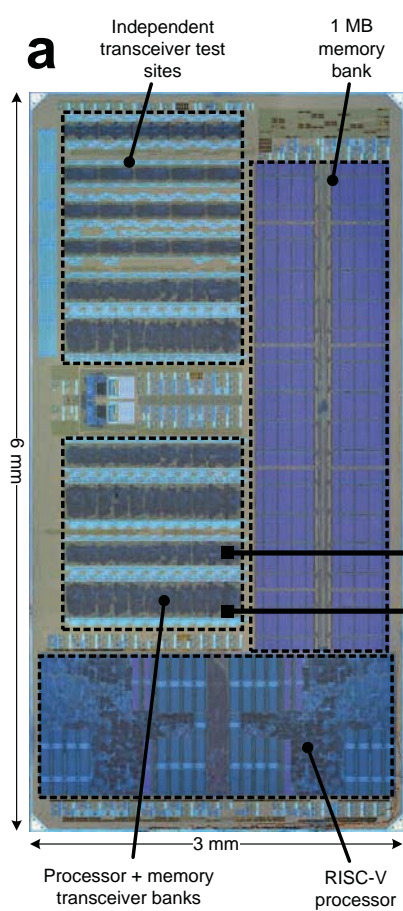
**Author Contributions** C. Sun, M. Wade, Y. Lee, and J. Orcutt contributed equally to this work. C. Sun developed the thermal tuning circuitry, designed the memory bank, implemented the “glue-logic” between various electronic components, and performed top-level assembly of electronics and photonics on the chip. M. Wade contributed modulator designs optimized for thermal tuning, designed the grating couplers, and performed top-level assembly of photonics regions used in the processor demonstration. C. Sun and Y. Lee designed the system-level architecture, tested the chip, and demonstrated the processor with photonic I/O. Y. Lee wrote and/or adapted the test programs for use in the processor demonstration. Y. Lee and A. Waterman developed the RISC-V ISA and the processor implementation. J. Orcutt created the CAD infrastructure for photonic layouts, designed the photodetector used in the demonstration, assembled earlier versions of photonic layouts, and contributed to passive devices. L. Alloatti improved the CAD infrastructure, developed new rules for design rule checking, and contributed new photodetector designs. C. Sun,

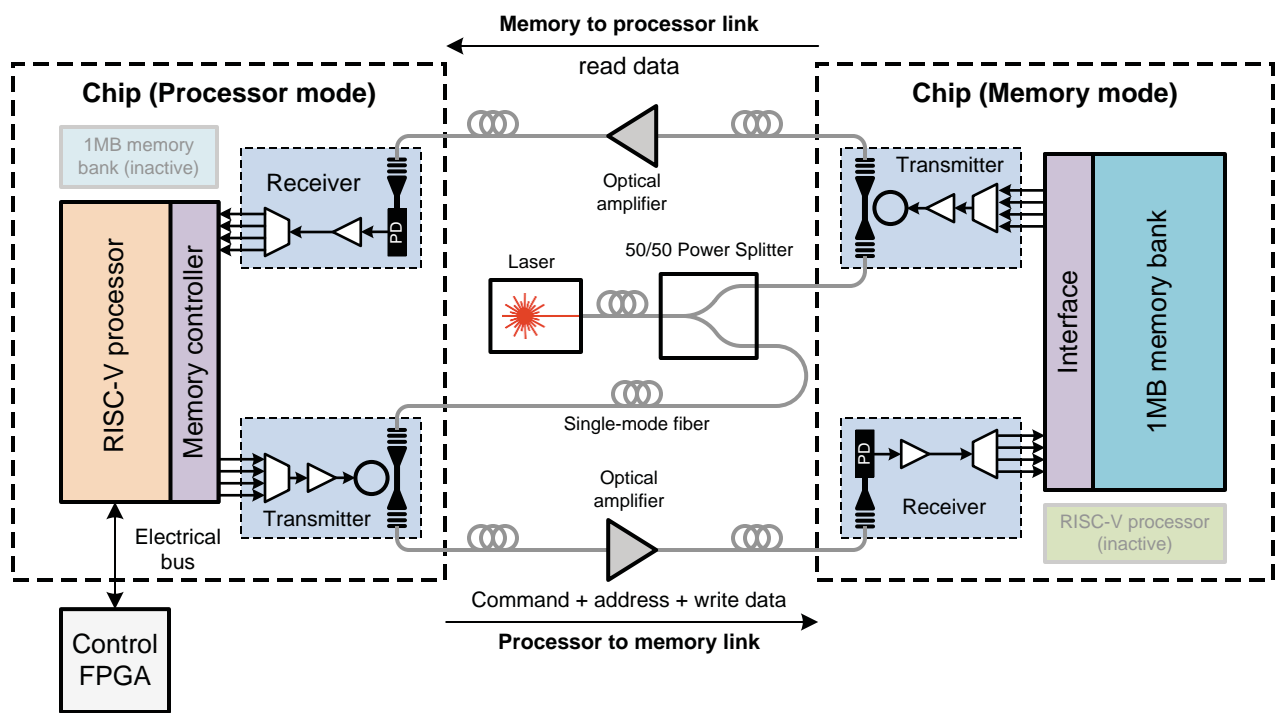


M. Wade, Y. Lee, and L. Alloatti all contributed to the chip verification. M. Georgas designed and implemented the receiver circuit. J. Shainline designed, implemented, and tested the original version of the modulator device. R. Avizienis did the physical implementation of the processor and designed the chip and adapter PCBs. S. Lin developed the selective substrate removal process and contributed to the thermal tuning method. B. Moss assisted with the chip implementation and performed initial substrate removal experiments. R. Kumar assisted in the rework of new grating coupler designs. F. Pavanello contributed to layout and data interpretation for couplers and modulators. A. Atabaki created new photodetector designs. H. Cook and A. Ou assisted with the design of the processor. J. Leu and Y.-H. Chen designed component pieces in the transceiver regions. V. Stojanović, M. Popović, R. Ram, and K. Asanović supervised the project.

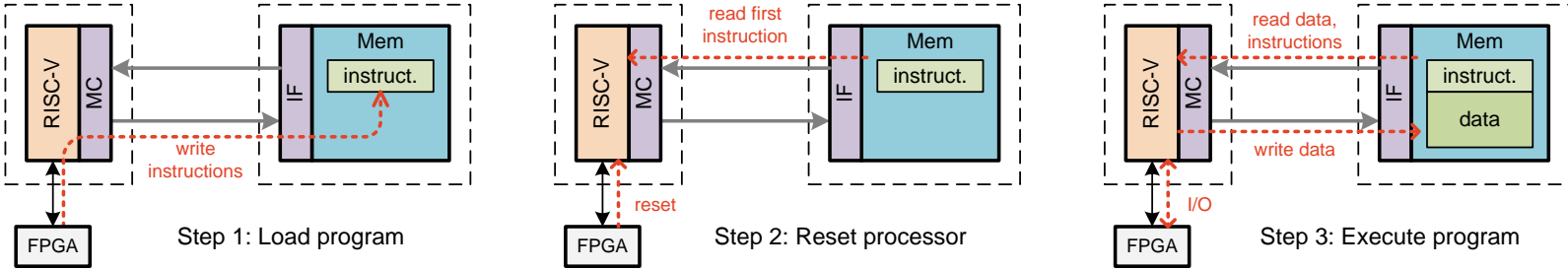
**Competing Interests** Competing financial interests: Chen Sun, Mark Wade, Rajeev Ram, Miloš Popović, and Vladimir Stojanović are involved in developing silicon photonic technologies at Ayar Labs, Inc. Yunsup Lee, Andrew Waterman, and Krste Asanović are working on platforms based on the RISC-V ISA at SiFive Inc. Jason Orcutt is now employed at IBM developing silicon photonics technologies.

**Correspondence** Correspondence and requests for materials should be addressed to Vladimir Stojanović (email: vlada@berkeley.edu), Miloš Popović (email: milos.popovic@colorado.edu), Rajeev Ram (email: rajeev@mit.edu), or Krste Asanović (email: krste@berkeley.edu).





**a**



```
>> ./fesvr-zedboard-head.1MB +divisor=1 +hold=1 ./pk -p ./hello.photonics
```

```
CPU reset complete
uncore slowio divisor=1, hold=1
host_clk frequency = 15.64 MHz
cpu_clk frequency = 31.27 MHz
hello world with photonics!
```

```
>> ./fesvr-zedboard-head.1MB +divisor=1 +hold=1 ./pk -p ./stream.256KB
```

```
CPU reset complete
uncore slowio divisor=1, hold=1
host_clk frequency = 15.64 MHz
cpu_clk frequency = 31.27 MHz
```

```
STREAM version $Revision: 5.10 $
```

```
This system uses 8 bytes per array element.
```

```
Array size = 32768 (elements), Offset = 0 (elements)
```

```
Memory per array = 0.2 MiB (= 0.0 GiB).
```

```
Total memory required = 0.8 MiB (= 0.0 GiB).
```

```
Each kernel will be executed 10 times.
```

```
The *best* time for each kernel (excluding the first iteration)
```

```
will be used to compute the reported bandwidth.
```

```
Your clock granularity/precision appears to be 1 microseconds.
```

```
Each test below will take on the order of 799 microseconds.
```

```
(= 799 clock ticks)
```

```
Increase the size of the arrays if this shows that
```

```
you are not getting at least 20 clock ticks per test.
```

```
WARNING -- The above is only a rough guideline.
```

```
For best results, please be sure you know the
```

```
precision of your system timer.
```

Function	Best Rate MB/s	Avg time	Min time	Max time
Copy:	640.3	0.000823	0.000819	0.000831
Scale:	551.3	0.000954	0.000951	0.000964
Add:	584.8	0.001350	0.001345	0.001364
Triad:	585.9	0.001351	0.001342	0.001364

```
Solution validates: avg error less than 1.000000e-13 on all three arrays
```

**b**

**c**



