Let There Be Light!

The Future of Memory Systems is Photonics and 3D Stacking

Keren Bergman Gilbert Hendry
Columbia University
{bergman,gilbert}@ee.columbia.edu

Paul Hargrove John Shalf
Lawrence Berkeley National Lab
{phhargrove,jshalf}@lbl.gov

Bruce Jacob
University of Maryland
~~bji~~@umd.edu
blj

K. Scott Hemmert Arun Rodrigues David Resnick
Sandia National Labs
{kshemme,afrodri,drresni}@sandia.gov

Abstract

Energy consumption is the fundamental barrier to exascale supercomputing and it is dominated by the cost of moving data from one point to another, not computation. Similarly, performance is dominated by data movement, not computation. The solution to this problem requires three critical technologies: 3D integration, optical chip-to-chip communication, and a new communication model. A memory system based on these technologies has the potential to lower the cost of local memory accesses by orders of magnitude and provide substantially more bandwidth. To reach the goals of exascale computing with a manageable power budget, the industry will have to adopt these technologies. Doing so will enable exascale computing, and will have a major worldwide economic impact.

Categories and Subject Descriptors B.3.1 [*Hardware*]: Memory Structures

General Terms Design, Economics

Keywords Memory Systems, DRAM, Photonics

1. Introduction

Energy consumption is the fundamental barrier to exascale supercomputing, and it is dominated not by the cost of computation, but rather the cost of moving data from one point to another. As an example, a 25MW supercomputer that performs only DRAM accesses at 100 pJ/bit would allow an aggregate memory bandwidth of 31 PB/sec. If the same machine supported an exaflop of computation at zero power cost, this would be 0.031 byte/flop of memory bandwidth, compared to the 1 byte/flop available in the original Cray XT architecture. Similarly, performance is dominated by memory movement, not computation. The solution to this problem requires two critical technologies: 3D integration and optical chip-to-chip communication. A memory system based on these technologies has the potential to lower the cost of memory accesses by orders of magnitude and provide substantially more bandwidth. Early versions of stacking are already seen in low power commodity devices, particularly cell phones, however it is only recently that fabrication techniques to provide high bandwidth connections between stacked chips (with Thru-Silicon Vias) have been developed and applied at commercial scale. Similarly, Silicon photonic communication has been demonstrated before, but it is only when coupled with 3D stacking that it provides low power communication at a reasonable fabrication cost. To reach the goal of an exascale computer with a manageable power budget, the industry will have to adopt these technologies.

Adoption of these technologies will not only enable exascale computing, but will also have a major worldwide economic impact of several billions of dollars per year.

2. Data Movement Dominates

The diverging performance of DRAM and processing, referred to as the Memory Wall [16], is a well known phenomena in computer architecture. Within a node, application performance is dominated by the memory system, especially for scientific simulation codes. Analysis of several scientific applications shows that even in ostensibly "floating point intensive" codes, less than 10% of instructions executed are floating point computation [17]. Instead, memory access operations (loads and stores) dominate. The large

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MSPC'11, June 5, 2011, San Jose, California, USA.

Copyright © 2011 ACM 978-1-4503-0794-9/11/06...$10.00

datasets and irregular access patterns of these codes mean that L1 miss rates of 50% or higher is not uncommon. With substantial portions of an application's instructions requiring the L2 (10s of ns) or main memory (100s of ns), memory clearly dominates the single node performance.

For large parallel jobs, the internode network is the dominant factor in determining how well an application will scale. Even at small scale, applications can spend 10-30% of their time waiting for messages to arrive[20] and reductions in bandwidth can have serious effects on overall performance [19].

2.1 Energy is Dominated by Moving Data

Accessing data from a DRAM today requires over 100 pJ per bit. By 2018, this is expected to drop to about 30 pJ/bit[13], making DRAM access one of the largest power consumers in an exascale-class machine. This is because modern DDR DRAM must access much more data than is actually required to fill a cacheline. For example, each memory access will activate several thousand memory cells in each of the 8-18 DRAM parts on a DIMM before transferring only a few hundred bits several centimeters back to the processor cache.

Current internode networks are based on high speed serial electrical and electrical/optical links. It is believed that this technology can scale to about 2 pJ/bit for short distances (<1m). For longer links, signaling speeds will require optical signaling. These optical links typically consist of an electrical link connected to an optical transceiver at each end. The result is a link which incurs both the electrical and optical energy. These links would consume around 4.6 pJ/bit (2 pJ each for electrical links on each end and 0.6 pJ for optical link). Routing requires roughly an additional 1 pJ/bit. With an even split between short and long-distance cables, 4.3 pJ/bit total is a reasonable average.

Current processors require several hundred to several thousand pJ for each operation performed. Aggressively power optimized general purpose processors and expected process scaling improvements may push this down to around 11 pJ per operation by 2018[11]. A floating point operation is expected to consume an additional 10 pJ per flop by 2018, and a L2 cache access perhaps 20 pJ. Assuming 10% of our instructions are floating point, and 50% require L2 cache access, this would mean an average instruction would require 22pJ.

If we theorize a 2018 exascale machine capable of 1 exaop (10^{18} operations per second), 1 exabyte/second of bandwidth to memory, 1 exabyte/second bandwidth to the nodes, 3 exabytes/second bandwidth in the router-to-router links, and 200 petabytes of memory capacity, and use the power and energy estimates above, such a machine would require 74.8 MW of power[1]. Of this, 57.2 MW (76%) is required for the data transfer (from memory, or between nodes) or masking data transfer latency (caches). This is another indication that data movement dominates system power consumption.

2.2 World Energy Impact

In addition to having a critical effect on machines at the exascale, DRAM power consumption has a major global impact. As detailed in Table 1, total DRAM power consumption is in the tens of *tera*watts-hours. Assuming the US holds one third to one fifth of the world's computers[2], this would mean billions of dollars spent on DRAM power. Even a small reduction in DRAM power could save hundreds of millions to billions of dollars.

Note, that this is probably a conservative estimate, as it neglects the DRAM found in portable electronics or embedded systems. Also, it is probable that the percentage of power used by computers has grown from 3-4% in 2005.

3741 Billion	Total US Power Consumption (KW-Hrs)[7]
× 3-4%	Used by computers (server & household) [5, 8, 9]
× 15-35%	DRAM power consumption [13]
=16.8-22.5 Billion	US DRAM Power (KW-Hrs)
× \$0.1 \$ / KW-Hrs	Retail Cost of Power, US Average [6]
=\$1.7-\$2.3 Billion	USD in Memory Power
× 3-5	Computers outside of US (Estimate)
=\$6.7-\$9.0 Billion	USD in worldwide memory power

Table 1. Worldwide DRAM power consumption

3. Solution

The solution to these problems is to combine 3D stacked memory, processing, and photonic interconnect. Photonic interconnect allows significant power reduction, especially over long distances, and huge increases in bandwidth. 3D stacking allows close integration between memory and processing elements (again reducing the power and increasing bandwidth) and allows close integration with a photonic interconnect.

3.1 3D Integration

3D packaging with Thru Silicon Vias (TSV)[10] is the key to low-power, high performance memory and is the solution to the data movement problem. 3D integration enables high bandwidth, low-power communication over short distances between heterogeneously fabricated devices. This is critical because logic devices (such as processors) are fabricated on different processes than DRAM or photonic devices. 3D integration provides a path for fabricating optics and DRAM in simpler processes and tightly binding them to logic devices. This technology will be the cornerstone of future memory technologies, and enable in-memory processing and new communication models.

[1] DRAM refresh should require about 28 mW per GB in this timeframe.

[2] The US is roughly one quarter of the world's GDP

3.2 Photonic Communication

In recent years, photonics has become increasingly attractive as a solution to chip-chip communications, as it is not clear that electronics can provide the energy-efficient solutions to high-bandwidth IO for Exascale computing [10]. Photonics can achieve extremely high bandwidth density by using wavelength division multiplexing (WDM), or transmitting many signals on different wavelengths in the same waveguide. Considering that a waveguide-fiber coupling from a chip out through the packaging is transparent to energy-per-bit, latency, and bandwidth, photonics provides the possibility of changing the game in chip IO.

Device research in silicon photonics has yielded a promising path towards realizing the full benefits of optical transmission at the chip- and board-scale. One of the basic devices for optical transmission, the silicon Mach-Zehnder interferometer modulator[14], is already commercially viable and available[1]. When used with an arrayed waveguide grating (AWG) for (de)multiplexing [3], a full link can be constructed.

Photonic devices based on the versatile ring-resonator structure are also attractive for their lower area and energy-per-bit. WDM modulation using ring-modulators has been demonstrated, and a full on-chip link has been shown with real data [18].

3.3 In-memory Computation

Close connection of logic and memory has long been a goal of computer architecture. Low latency, high bandwidth connections between processing and memory would eliminate the von Neumann bottleneck. Unfortunately, attempts to integrate memory and logic onto the same die have met with limited success due to differences in the fabrication processes for DRAM and high performance logic. Hybrid approaches such as eDRAM typically sacrifice memory density and have a high cost.

The use of 3D stacking will be the most fundamental change to main memory systems since the invention of DRAM. Its most important feature is the ability to integrate dense DRAM memory with high performance CMOS logic parts in the same package. By connecting logic and memory together in this way it is possible to move processing, data handling, and other tasks closer to the memory, reducing power and latency.

3.4 Communication Models

These technologies can enable new communication models that will have a major impact on system performance.

Unlike traditional communication models, message-driven computation moves the work to the location of the data, rather than moving or copying data. Many problem domains, such as those which use a work queue model or graph-based informatics, naturally fit into the message-driven computation model. Moving the work to the data can reduce bandwidth requirements and eliminate many network transactions. This can dramatically reduce power consumption, especially if it is coupled with hardware support. For example, the hardware could provide an "enqueue" instruction to manage remote shared queue structures, and alleviate the software overhead of message processing.

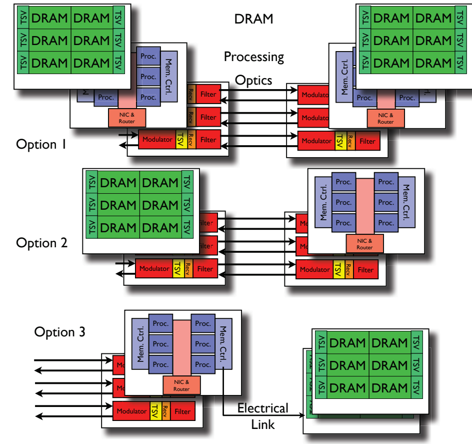

Figure 1. Architectures

Currently, communication protocols use the memory system for communication; that is, inter-process communication is accomplished by moving data through DRAM, even when the ultimate goal of many communications is simply to pass data from one thread's register file to another thread's register file. This can be avoided by dividing the memory system into its two separate functions: naming and storage, and treating them separately. Current communication models use both to facilitate transmission of data; a datum is *stored* in the memory system, and the recipient reads from the storage, using the expensive coherence fabric tied to the datum's *name* to ensure that the proper data value is transmitted. Instead, the storage portion can be avoided, but the naming mechanism (i.e. address) can be used.

3.5 Possible Architectures

With the fundamental building blocks of memory, processing, and photonic layers, held together by the mortar of 3D integration with Thru-Silicon Vias, it is possible to create a number of interesting architectures (Figure 1):

1. **Full Processing In Memory**: The most aggressive design, and potentially most efficient, would be to eliminate separate processing chips altogether and perform all processing in layers stacked with the memory parts. By replacing the expensive board-level connection between the processors and memory with TSVs, it is possible to boost bandwidth, reduce latency, and decrease power. The latency reduction may be significant enough that the processor itself can be redesigned to increase efficiency even more, such as by reducing or removing the L2 cache.

2. **Separate Processing w/ Photonics**: A less aggressive design would be to keep memory and processing in separate packages, but to use silicon photonics to communicate between them and across the system. This would reduce power required substantially, and improve bandwidth, but not as much as the full processing in memory approach.
3. **Stacked Photonics**: A simpler conservative alternative would be to utilize a conventional electrical link between processor and memory, but to stack photonic components with the processor for system-level communication. This would reduce internode communication power, especially for long-distance communication, but would not fully address the memory system power.

4. Analysis

4.1 Stacked Memory Performance Analysis

The performance of memory interconnected with TSVs depends on a number of implementation choices. Some areas of optimization and exploration include:

- **The number and use of TSVs** Each data TSV can contribute directly to the total bandwidth. Some TSVs will need to be used for such things as addressing, power delivery, or other functional uses. Another design choice is if TSVs should be unidirectional or bidirectional.
- **The TSV data rate** The energy efficient signaling rate for data lines in a DRAM is significantly less than in a logic-optimized CMOS part. Essentially, DRAMs are optimized for memory density and retention, not for the smallest logic delay. Still, TSV rate of 1-2 Gbits per second are achievable with an energy cost of 1-11 femto-Joules per bit.
- **The layout and organization of the memory dies containing the TSVs** To ensure the greatest energy reduction and highest data rates, the distances from the memory bank to the TSV will have to be carefully managed. Additionally, it will be necessary to avoid hot-spots, both by having enough memory banks, and by making internal routing as conflict free as is reasonable.
- **Performance of the DRAM banks** There is little reason to change the basic architecture of a DRAM cell or bank. Thus, the performance of a a DRAM bank with TSVs will be very similar to that of current DRAM parts, and may be a limitation.
- **External connections** In the most aggressive architectures, DRAM is stacked directly on the processors. If the processing is packaged separately, even with high-bandwidth optical IO, then care must be taken to avoid bottlenecks and reduce latency in the communication protocol.

It is reasonable to envision memory parts with TSVs that have bandwidths from tens to low hundreds of Gbytes per second, however, there will be multiple cost/complexity and manufacturing issues that must be overcome to bring such parts to market. The memory that results will provide much higher bandwidths than current parts, while using much less energy than current parts.

The external memory latency of parts with TSVs will be similar to current parts. However, if processing elements are contained within the memory stack, they will be able to take advantage of slightly improved latencies.

Parts with TSVs will offer a range of design and optimization possibilities. Many of these optimization are the same as current parts, but start from a base of significantly higher bandwidths and significantly lower energy per bit than current parts.

4.2 Energy Analysis

4.2.1 Stacked Memory Energy Analysis

As mentioned, modern DDR DRAM is energy inefficient because it accesses more data than it needs, and because it must transfer that data long distances. These inefficiencies are caused by the relatively small number of pins between a memory part and the processor and the geometries of separate packages.

3D Stacking effectively eliminates these bottlenecks, enabling much more efficient access of DRAM over much shorter distances. The dynamic dissipation of DRAM is primarily capacitance based, and can be estimated by (adapted from[13]):

$$E_{energyPerBit} = \quad (V^2 * (C_{bl} + C_{cell} + \frac{C_{wl}}{bA}) * O_{cmd} + E_{TSV}) * O_{ECC} \tag{1}$$

Using estimates for 2018 DRAM (Table 2), we estimate about 255 fJ to move one bit from the DRAM cell to a processor on a stacked die. Including a memory controller, this may rise to 500 fJ per bit – dramatically less than the 30pJ/bit for conventional DRAM.

Table 2. Stacked DRAM Energy Components (2018)

Symbol	Meaning	Used
$E_{energyPerBit}$	Avg. energy (fJ) per bit	248.5
V	DRAM Core Voltage, in Volts	1.0
C_{bl}	Capacitance of the bitline (fF)	80
C_{cell}	DRAM cell Capacitance (fF)	25
C_{wl}	Wordline Capacitance (fF)	3000
bA	Avg. bits read from a wordline.	512
O_{cmd}	Command & Address Overhead	2.0
E_{TSV}	TSV transfer energy, per bit (fJ)	5
O_{ECC}	ECC Overhead.	1.125

4.2.2 Photonic Communication Energy

The most significant sources of energy consumption in a photonic link are laser power and driver/receiver energy-per-bit, both of which can be computed on a per-wavelength basis. Laser power per wavelength is a function of detector

Table 3. Insertion Loss Parameters

Symbol	Parameter	Value
$\zeta_{coupler}$	Off-chip coupler	0.5 dB [2]
ζ_{prop}	Waveguide propagation	1.5 dB/cm [23]
ζ_{bend}	Waveguide bend	0.005 dB [23]
ζ_{mod}	Modulator thru loss	0.05 dB
$\zeta_{de/mul}$	(De)multiplexing loss	2 dB
S_{det}	Detector sensitivity	-15 dBm
η_{laser}	Laser efficiency	50%

sensitivity, insertion loss, and laser efficiency as follows (in W):

$$P_{laser} = \frac{10^{(S_{det}+\zeta_{link}-30)/10}}{\eta_{laser}} \quad (2)$$

where η_{laser} is the laser quantum efficiency, S_{det} is the detector sensitivity (in dBm), and ζ_{link} is the worst-case insertion loss (in dB). We can estimate the insertion loss for each link with the following equation:

$$\zeta_{link} = 2\zeta_{coupler} + \zeta_{wg} + \zeta_{mod} + 2\zeta_{de/mul} \quad (3)$$

where $\zeta_{coupler}$ is the loss from coupling the laser into the chip, ζ_{mod} is the loss experienced through a modulator, $\zeta_{de/mul}$ is the loss from wavelength multiplexing and demultimplexing, and ζ_{wg} is the insertion loss of traveling in the routed waveguides including propagation (ζ_{prop}) and bending (ζ_{bend}). Table 3 shows the parameters used for estimating link insertion loss and laser power. This results in a conservative prediction for laser power of around 0.29mW for a 1cm distance traveled on-chip per wavelength per link, regardless of signaling rate. As off-chip fibers are much less lossy, (a few dV per kilometer), photonics is highly suited to long distance chip-to-chip interconnect.

Dynamic power is likely to come mainly from the driver, receiver, and SerDes circuits associated with each wavelength in a link. One source shows these circuits taking up approximately 1pJ/bit[22]. By 2018, 0.5pJ is obtainable [13].

4.3 Cost Analysis

Photonics and 3D stacking are still areas of active development, and the cost of these technologies is not fully understood. However, there is reason to believe it will not be prohibitive.

4.3.1 Photonics Costs

Because of the availability of implementing photonic devices with CMOS-compatible deposited materials such as silicon nitride and polysilicon crystalline, it is likely that the actual fabrication of the photonic links will not incur any additional significant costs over the chips themselves. Packaging the optics, namely in fiber-coupling, is likely to add some cost, though this will become cheaper as optical packaging reaches maturity. Integrated or discrete lasers will add to the total cost of a single computing node, though using comb-lasers which can produce multiple wavelengths of light with one device [12] can significantly reduce this cost.

4.3.2 Stacked Memory Costs

Stacking memory will add cost due to the expense of forming the TSVs, extra area required for the TSVs, and yield loss due to stacking.

Creation of the TSVs is performed at the wafer level, using a reactive ion etch to excavate the hole, and then various deposition processes to fill it. This cost is largely independent of the number of TSVs per wafer, and is estimated to add 10% per wafer.

TSVs can be made quite small, with a 10 micron diameter being a very conservative estimate. With driver circuitry and additional overhead, we assume a single TSV requires a 20 by 20 micron space. With TSVs clocked at 1-2 Ghz, and accounting for overheads in addressing, redundancy, ECC, power and ground signals, and bidirectional transfer, 1,000 TSVs would have about 20 GByte/sec effective bandwidth, or about 9 times the estimated bandwidth of a DDR4 DRAM [21]. Assuming a DRAM chip is about 100 mm^2, and using the ICKnowledge chip costing model[15], 1000 TSVs would increase the silicon area by about 0.4% and the cost by about 0.5%. Ten thousand TSVs (90 times DDR4 bandwidth) would increase area by 4% and cost by 5%. Note, this neglects the positive effect that removing off-chip driver circuitry would have on chip area.

Studies[4] suggests a 99% yield per stacking operation. With an eight deep stack, this would increase cost by 8.4%. In total, this would mean 1,000 TSVs would increase total cost by about 20%, and 10,000 TSVs would add about 25%.

A major cost impact is the number of package pins on a component. Both 3D stacking and Photonics have the potential to decrease the total number of pins on a package by using more efficient TSVs or optical fiber to move data. Also, roughly half of the pins on a package are for power delivery. As the power requirements for a part are reduced, it is possible that some of the power delivery pins can be eliminated.

5. Conclusion

The solutions proposed here will have a dramatic impact on future exascale computers. If we revisit the sample exascale machine from Section 2.1 with the proposed architectures in Section 3.5 we can see dramatic power savings, summarized in Table 4. The aggressive Option 1 architecture, combining 3D stacking, in-memory computation, and photonic communication would reduce the power budget for an exascale machine by almost 70%.

The combination of 3D integration with thru silicon vias, silicon photonic interconnect, and new communication models will revolutionize how data is transfered, how machines are architected, and how power is consumed. In addition to making exascale computing affordable, it has the potential to save hundreds of millions to billions of dollars in world electrical costs.

Component	Base (MW)	Option 1 (MW)	Option 2 (MW)	Option 3 (MW)
Memory BW	30	0.5	1.1	30
Memory Cap.	5.6	5.6	5.6	5.6
Network	17.2	6.4	6.4	6.4
Processor	22	12	22	22
Total	74.8	24.5	35.1	64

Table 4. Exascale Power Consumption

Acknowledgments

This work was supported by the ASCR Office in the DOE Office of Science under contract number DE-AC02-05CH11231.

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin company, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

[1] A. Alduino et al. Demonstration of a high speed 4-channel integrated silicon photonics WDM link with hybrid silicon lasers. In *Proceedings of Hot Chips*, Aug. 2010.

[2] V. R. Almeida, R. R. Panepucci, and M. Lipson. Nanotaper for compact mode conversion. *Optics Letters*, 28(15):1302–1304, August 2003.

[3] P. Cheben and et al. A high-resolution silicon-on-insulator arrayed waveguide grating microspectrometer with sub-micrometer aperture waveguides. *Opt. Express*, 15(5): 2299–2306, Mar 2007. doi: 10.1364/OE.15.002299. URL `http://www.opticsexpress.org/abstract.cfm?URI=oe-15-5-2299`.

[4] X. Dong and Y. Xie. System-level cost analysis and design exploration for three-dimensional integrated circuits (3d ics). In *Proceedings of the 2009 Asia and South Pacific Design Automation Conference*, ASP-DAC '09, pages 234–241, Piscataway, NJ, USA, 2009. IEEE Press. ISBN 978-1-4244-2748-2. URL `http://portal.acm.org/citation.cfm?id=1509633.1509700`.

[5] Energy Information Administration. Commercial buildings energy consumption survey. Technical report, Department of Energy, 2007.

[6] Energy Information Administration. Electric power monthly. Technical report, Department of Energy, March 2011. URL `http://www.eia.doe.gov/cneaf/electricity/epm/`.

[7] Energy Information Administration. Annual energy review 2007. Technical Report DOE/EIA-0384(2007), Department of Energy, 2007.

[8] Energy Information Administration. U.s. household electricity report. Technical report, Department of Energy, July 2005. URL `http://www.eia.doe.gov/emeu/reps/enduse/er01_us.html`.

[9] ENERGY STAR Program. Report to congress on server and data center energy efficiency public law 109-431. Technical report, Environmental Protection Agency, August 2007. URL `http://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf`.

[10] ITRS Committee. International Technology Roadmap for Semiconductors 2008 Update, 2008. http://www.itrs.net/Links/2008ITRS/Home2008.htm.

[11] D. Jensen and A. Rodrigues. Embedded systems and exascale computing. *Computing in Science And Engineering*, 2010. Accepted for Publication.

[12] B. Koch, A. Fang, R. Jones, O. Cohen, M. Paniccia, D. Blumenthal, and J. Bowers. Silicon evanescent optical frequency comb generator. In *Group IV Photonics, 2008 5th IEEE International Conference on*, pages 64 –66, Sept 2008. doi: 10.1109/GROUP4.2008.4638098.

[13] P. M. e. Kogge. ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems. Technical report, University of Notre Dame CSE Department Technical Report, TR-2008-13, September 28, 2008.

[14] L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, U. Keil, and T. Franck. High speed silicon mach-zehnder modulator. *Opt. Express*, 13(8):3129–3135, Apr 2005. URL `http://www.opticsexpress.org/abstract.cfm?URI=oe-13-8-3129`.

[15] I. K. LLC. Ic cost model 0909b. Software, November 2009.

[16] S. A. McKee. Reflections on the memory wall. In *CF '04: Proceedings of the 1st conference on Computing frontiers*, page 162, New York, NY, USA, 2004. ACM. ISBN 1-58113-741-9. doi: http://doi.acm.org/10.1145/977091.977115.

[17] R. C. Murphy, A. Rodrigues, P. Kogge, and K. Underwood. The implications of working set analysis on supercomputing memory hierarchy design. In *The 2005 International Conference on Supercomputing*, June 20-22, 2005.

[18] N. Ophir et al. First demonstration of error-free operation of a full silicon on-chip photonic link. In *OFC*, Mar. 2011.

[19] K. Pedretti, C. Vaughan, K. S. Hemmert, and B. Barrett. Application sensitivity to link and injection bandwidth on a cray xt4 system. In *Proceedings of the 2008 Cray User Group Annual Technical Conference*, May 2008.

[20] A. Rodrigues. *Dusty Decks, Memory Walls, and the Speed of Light*. PhD thesis, University of Notre Dame, 2006.

[21] Samsung. Samsung develops industry's first ddr4 dram, using 30nm class technology. News Release, January 2011. URL `http://www.samsung.com/us/business/semiconductor/newsView.do?news_id=1202`.

[22] H. D. Thacker and others. Flip-chip integrated silicon photonic bridge chips for sub-picojoule per bit optical links. In *Electronic Components and Technology Conference (ECTC), 2010 Proceedings 60th*, pages 240 –246, June 2010.

[23] F. Xia, L. Sekaric, and Y. Vlasov. Ultracompact optical buffers on a silicon chip. *Nature Photonics*, 1:65–71, Jan. 2007. doi: 10.1038/nphoton.2006.42.