Guest Editors' Introduction: Special Section on Emerging Memory Technologies in Very Large Scale Computing and Storage Systems

Hossein Asadi, Paolo Ienne, and Hamid Sarbazi-Azad

◆

1 INTRODUCTION

THE overwhelmingly increasing demand for both storage and computation necessitates revisiting the traditional memory subsystems used in processors and storage systems to take advantage of emerging memory technologies. *Dynamic Random Access Memory* (DRAM) and *Static Random Access Memory* (SRAM) have been ubiquitously used for decades as main memory and as on-chip cache, respectively. Scaling and power issues of these traditional memory technologies have led to significant investment in emerging memory technologies. On the other hand, fundamental limitations of mechanical disk drives have brought great attention to further explore the design space of *Solid-State Drives* (SSDs). The promising features of emerging memory technologies such as low power consumption, increased performance, lower susceptibility to particle strikes, and higher bit density have prompted researchers to seek new organizations in the different memory-hierarchy levels, to propose new circuitry and algorithms to improve performance and reduce power consumption, and to develop new schemes to enhance system reliability.

While performance and power are the major motivations to revisit memory organization in the cache hierarchy of modern processors, several contributing factors such as endurance, reliability, power, throughput, and cost are the major concerns in storage subsystems and SSDs. Design refinement techniques in both processors and storage systems often span different abstraction levels in computer systems, ranging from circuit design to micro-architectural techniques, from system architecture to operating systems.

To foster the research momentum on emerging memory technologies in high-performance processors and mainstream and enterprise storage systems, *IEEE Transactions on Computers* has scheduled a special section on both system- and circuit-level aspects of emerging memory

and storage technologies. For this special section, we have received 50 submissions from 17 countries, out of which 18 submissions have been selected in the first round of review process. Finally, after detailed review of the revised submissions, 11 papers have been accepted for publication in this section. More than 200 reviewers have helped us to make publishing decisions for this special section. The reviewers have been carefully selected among distinguished researchers from both industry (IBM, HP, Microsoft, EMC, NetApps, Intel, Data Storage Institute, and Fusion IO) and academia. Most of them have served as members of the organization or program committees of top conferences, including the IEEE/ACM International Symposium on Computer Architecture (ISCA), the IEEE/ACM International Symposium on Microarchitecture (MICRO), the USENIX Conference on File and Storage Technologies (FAST), the ACM/IEEE Design Automation Conference (DAC), the IEEE/ACM International Conference on Computer-Aided Design (ICCAD), the IEEE/ACM Design, Automation, and Test in Europe (DATE), the IEEE/ACM Asia and South Pacific Design Automation Conference (ASP-DAC), and the IEEE Non-Volatile Memory Technology Symposium (NVMTS).

2 CONTRIBUTIONS

This special section brings the great opportunity of sharing innovative ideas and research developments in two major tracks:

a) Emerging uses of memory technologies in processor architecture.
b) Emerging uses of memory technologies in storage systems.

The first track tries to address limitations of emerging memory technologies including both volatile and non-volatile memories. Most of the articles in this track focus on circumventing the intrinsic limitations of non-volatile memory technologies such as *Phase-Change Memory* (PCM) and *Spin-Transfer Torque Magnetic RAM* (STT-MRAM). Nonetheless, there is also work trying to elude shortcomings of standard DRAM technology for *Embedded DRAM* (eDRAM) and suggesting the use of FinFETs. The second track mainly focuses on different aspects of storage subsystems such as performance, reliability, and endurance.

- *H. Asadi is with the Sharif University of Technology.*
  E-mail: asadi@sharif.edu.
- *P. Ienne is with Ecole Polytechnique Fédérale de Lausanne (EFPL).*
  E-mail: paolo.ienne@epfl.ch.
- *H. Sarbazi-Azad is with the Sharif University of Technology and Institute for Research in Fundamental Sciences (IPM).*
  E-mail: azad@sharif.edu, azad@ipm.ir.

The papers of these two tracks are summarized in the following subsections.

2.1 Emerging Uses of Memory Technologies in Processor Architectures

The first paper in this track is entitled "Cache Design with Domain Wall Memory" by Rangharajan Venkatesan, Vivek J. Kozhikkottu, Mrigank Sharad, Charles Augustine, Arijit Raychowdhury, Kaushik Roy, and Anand Raghunathan. It discusses the use of *Domain Wall Memory* (DWM), a recently developed spin-based memory technology, for caches. This work first provides a comprehensive overview of DWM and two main latency- and capacity-optimized cell designs. The authors then focus on different design choices for using DWM, such as cell type, organization, and head management. Of note is the use of a hybrid array comprised of both types of cells and a thorough study regarding head positioning. A two-level array is then designed and tested using simulations. The authors demonstrate massive gains in area and energy efficiency.

The second paper, "WOM-Code Solutions for Low Latency and High Endurance in Phase Change Memory" by Poovaiah Palangappa, Jiayin Li, and Kartik Mohanram, presents a *Write-Once-Memory* (WOM) code; it addresses the long latency of PCM technology by integrating the WOM code at both the memory architecture end and in the memory controller. Since PCM RESET operation is much faster than SET operation, the paper proposes to use an inverted WOM-code that transforms PCM writes to comprise of low latency RESET operations. To minimize the memory overhead of the WOM-code encoded data, the authors present two memory organizations: wide-column and hidden-page. To verify the proposed methods, the DRAMSim2 simulator has been extended to evaluate the WOM-code PCM architecture across different benchmarks.

The third paper, "Array Organization and Data Management Exploration in Racetrack Memory" by Zhenyu Sun, Xiuyuan Bi, Wenqing Wu, Sungjoo Yoo, and Hai Li, presents an ultra-dense memory architecture based on a racetrack memory—a descendant of STT-RAM which was formerly introduced by IBM. In the proposed architecture, proper data management scheme and cross-layer optimization techniques are employed to enhance both performance and energy efficiency. Efficient cell and array designs are also presented to eliminate the area constraint of the access transistor size and to enable the uniform access ports for read and write operations.

A CMOS-compatible non-volatile SRAM-based hybrid memory scheme entitled "cNV SRAM: CMOS Technology Compatible Non-Volatile SRAM-Based Ultra-low Leakage Energy Hybrid Memory System" is presented by Jinhui Wang, Lina Wang, Haibin Yin, Zikui Wei, Zezhong Yang, and Na Gong. The proposed non-volatile SRAM, called c-NV SRAM, is implemented using the Fowler-Nordheim tunneling mechanism. Efficient power gating techniques are also employed to achieve ultra-low leakage energy. Conventional 8T-SRAM is used for fast operation while data is backed up with non-volatile parts in sleep state and the power supply is switched off to eliminate leakage energy.

Another interesting study, entitled "Feasibility of Embedded DRAM Cells on FinFET Technology" by E. Amat, A. Calomarde, F. Moll, R. Canal, and A. Rubio, characterizes several eDRAM designs which are based on FinFET technology. In this study, several important parameters such as process variation, voltage fluctuation, environment temperature variation, yield estimate, soft error tolerance due to ion strikes, speed performance, different topologies, and layout comparison are investigated in various eDRAM technologies. The impact of device fluctuation on the parameters of eDRAMs has been evaluated by Monte-Carlo simulations. The variability is also modeled as a change in the device threshold voltage. Numerous findings have been presented in this study; for instance, it is shown that the cells based on multiple-gate devices exhibit a factor of two improvement in retention time over the implementation with planar transistors.

2.2 Emerging Uses of Memory Technologies in Storage Systems

The first paper in this track is entitled "Effective Lifetime-Aware Dynamic Throttling for NAND Flash-Based SSDs" and is authored by Sungjin Lee and Jihong Kim. It presents a dynamic throttling technique to improve the lifetime of SSDs. In this technique, the number of data blocks written over each time interval is controlled by adjusting the write speed of SSDs and by taking into account the effective write endurance of NAND flash memory. To have minimal performance degradation, the dynamic throttling is performed by predicting future write traffic and adopting appropriate write speed. This also prevents significant fluctuations in write response times.

A coding scheme for flash-based SSDs is presented in "Improving Read Performance of NAND Flash SSDs by Exploiting Error Locality" by Ren-Shuo Liu, Meng-Yen Chuang, Chia-Lin Yang, Cheng-Hsuan Li, Kin-Chu Ho, and Hsiang-Pang Li. The main aim of the proposed encoding scheme is to speed up decoding latency of *Low-Density Parity-Check* (LDPC) in read operations. Flash memory has a high error locality and error correction calculations can be reused to avoid costly LDPC decode operations. This scheme adaptively partitions DRAM for use as a regular cache and a cache for storing error correction data. These partitions are periodically updated using analysis based on stack distance. This scheme was evaluated using real hardware and a significant read performance gain has been achieved.

The study presented in "Exploring Design Challenges in Getting Solid State Drives Closer to CPU" by Myoungsoo Jung explores challenges of two representative PCI-Express SSD architectures, namely, from-scratch SSD and bridge-based SSD. This study reveals that the former architecture, which typically employs *Field-Programmable Gate Array* (FPGA)- or *Application-Specific Integrated Circuit* (ASIC)-based native PCIe controllers, offers considerable performance improvement at the cost of significant host-side memory and computational resources. The quantitative results also show that the latter architecture, which leverages conventional high-performance SSD controllers, imposes less memory and computation power while

delivering longer latency and less throughput as compared to the from-scratch SSD architecture.

An intra-SSD level *Redundant-Array of Independent Disks* (RAID) scheme which allows flexible stripe sizes and parity placements is presented in "Chip-Level RAID with Flexible Stripe Size and Parity Placement for Enhanced SSD Reliability" by Jaeho Kim, Eunjae Lee, Jongmoo Choi, Donghee Lee, and Sam H. Noh. The proposed RAID scheme constructs stripes based on the arrival order of write requests and tries to place parity data in the stripes such that parity update overhead is minimized. By modeling performance and reliability of both conventional RAID-5 and the proposed RAID schemes, it is shown that reducing the number of parity writes committed to flash enhances both performance and reliability as compared to the conventional RAID5 scheme.

An analytical model to quantify the reliability of SSDs is presented in "Analysis of Reliability Dynamics of SSD RAID" by Yongkun Li, Patrick P.C. Lee, and John C.S. Lui. The proposed model takes into account the flash wear-outs due to limited program/erase cycles of flash chips as well as bit errors due to read disturbs, program disturbs, and retention errors. In particular, the paper explores the reliability of even parity distribution as opposed to uneven parity distribution across multiple SSDs. Due to time-varying bit error rates of SSDs and with the purpose of modeling the reliability of an SSD-based RAID array over time, the authors use non-homogeneous continuous time Markov chains and conduct transient analysis to derive reliability metrics.

The last but not the least paper in this track is entitled "Eliminating Periodic Flush Overhead of File I/O with Non-volatile Buffer Cache" by Eunji Lee, Hyojung Kang, Hyokyung Bahn, and Kang G. Shin. It addresses the load of synchronous periodic writes performed by the operating system: to prevent data loss, dirty data blocks are periodically saved to preserve data in case of power loss, resulting in synchronous writes to disk. The authors propose using *Non-Volatile Memory* (NVM) to store the dirty blocks as, unlike DRAM, NVM preserves data upon a power loss. This scheme reduces writes and space requirements by only storing the modified parts of blocks in NVM. Also, when blocks are flushed from main memory, the corresponding data blocks are also flushed from NVM to free up NVM space. This solution was implemented on the Linux Kernel: the authors demonstrated reduced writes (by nearly half) and improved performance.

3   CONCLUDING REMARKS

In conclusion, as evidenced by the numerous submissions to this special issue, we believe that emerging memory technologies have great potential for further research and applications in computing and storage systems. The guest editors would like to thank all the authors who have submitted papers to this special issue, including those authors whose contributions could not be selected for inclusion in this issue. They would also like to offer our sincere thanks to the reviewers, without whom this issue would not have been possible and who provided very constructive comments and suggestions to the authors. They would also like to express their gratitude to the retired editor-in-chief, Prof. Albert Zomaya, for his support and help in making this special issue a reality. Last but not least, they thank the staff of the editorial office of the *IEEE Transactions on Computers* for their cooperation during this project. But, of course, most of all, we hope you, the readers, will enjoy the papers in this special issue.

<div align="right">

Hossein Asadi
Paolo Ienne
Hamid Sarbazi-Azad
Guest Editors

</div>

Hossein Asadi (M'08-SM'14) received the BSc and MSc degrees in computer engineering from the Sharif University of Technology (SUT), Tehran, Iran, in 2000 and 2002, respectively, and the PhD degree in electrical and computer engineering from Northeastern University, Boston, MA, in 2007. He was with EMC Corporation, Hopkinton, MA, as a research scientist and senior hardware engineer, from 2006 to 2009. From 2002 to 2003, he was a member in the Dependable Systems Laboratory, SUT, where he researched hardware verification techniques. From 2001 to 2002, he was a member in the Sharif Rescue Robots Group. He has been with the Department of Computer Engineering, SUT, since 2009, where he is currently a tenured associate professor. He is the founder and director in the Data Storage Systems Laboratory at SUT. He has authored and coauthored more than 50 technical papers in reputable journals and conference proceedings. His current research interests include data storage systems and networks, solid-state drives, operating system support for I/O and memory management, and reconfigurable and dependable computing. He received the Technical Award for the Best Robot Design from the International RoboCup Rescue Competition, organized by AAAI and RoboCup, received the Best Paper Award at the 15th CSI Internation Symposium on Computer Architecture and Digital Systems (CADS) in 2010, and the Distinguished Lecturer Award from SUT in 2010, one of the most prestigious awards in the university. He also received an "*Extraordinary Ability in Science*" visa from US Citizenship and Immigration Services in 2008. He is a senior member of the IEEE.

Paolo Ienne (M'90-SM'13) has been a professor at the EPFL since 2000 and heads the Processor Architecture Laboratory (LAP). Prior to that, from 1990 to 1991, he was an undergraduate researcher with Brunel University, Uxbridge, United Kingdom. From 1992 to 1996, he was a research assistant at the Microcomputing Laboratory (LAMI) and at the MANTRA Center for Neuro-Mimetic Systems of the EPFL. In December 1996, he joined the Semiconductors Group of Siemens AG, Munich, Germany (which later became Infineon Technologies AG). After working on datapath generation tools, he became the head in the embedded memory unit in the Design Libraries division. His research interests include various aspects of computer and processor architecture, electronic design automation, computer arithmetic, FPGAs and reconfigurable computing, and multiprocessor systems-on-chip. He received the Best Paper Award at the 40th Design Automation Conference (DAC) in 2003, at the International Conference on Compilers, Architectures, and Synthesis for Embedded Systems (CASES) in 2007, at the 19th International Conference on Field-Programmable Logic and Applications (FPL) in 2009, and at the 20th ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA) in 2012. In 2008, he was general co-chair of the 6th IEEE Symposium on Application Specific Processors (SASP) and a guest editor of a Special Section on Application Specific Processors which appeared in October 2008 on the *IEEE Transactions on Very Large Scale Integration Systems*. In 2010, he was the program subcommittee chair of the Design Automation Conference (DAC) on High-Level and Logic Synthesis. From 2010 to 2012, he was a topic co-chair of Design Automation and Test in Europe (DATE) for Architectural and High-Level Synthesis topic. In 2011, he was a program co-chair of the 20th IEEE Symposium on Computer Arithmetic (ARITH) and a Program Co-Chair of the 22nd IEEE International Conference on Application-specific Systems, Architectures and Processors (ASAP). In 2012, he was a guest editor of the Special Section on Computer Arithmetic of the *IEEE Transactions on Computers*. In 2014, he was a chair in the Program Committee of the 23rd International Workshop on Logic & Synthesis (IWLS) and a guest editor in the *IEEE Micro* Special Issue on Reconfigurable Computing. He has been a member of some 50 program committees of international workshops and conferences in the areas of design automation, computer architecture, embedded systems, compilers, FPGAs, and asynchronous design. He has been an associate editor of *ACM Transactions on Design Automation of Electronic Systems* (TODAES), since 2011, and of *ACM Computing Surveys* (CSUR), since 2014. He is a senior member of the IEEE.

Hamid Sarbazi-Azad received the BSc degree in electrical and computer engineering from Shahid-Beheshti University, Tehran, Iran, in 1992, the MSc degree in computer engineering from Sharif University of Technology, Tehran, Iran, in 1994, and the PhD degree in computing science from the University of Glasgow, Glasgow, United Kingdom, in 2002. He is currently a professor in the Department of Computer Engineering, Sharif University of Technology, and heads the School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran. His research interests include high-performance computer/memory architectures, NoCs and SoCs, parallel and distributed systems, performance modelling/evaluation, and storage systems, on which he has published more than 300 refereed conference and journal papers. He received Khwarizmi International Award in 2006, TWAS Young Scientist Award in engineering sciences in 2007, and Sharif University Distinguished Researcher awards in years 2004, 2007, 2008, 2010, and 2013. He is a member of the managing board of the Computer Society of Iran (CSI), and has served as the editor-in-chief for the *CSI Journal on Computer Science and Engineering*, an associate editor for *IEEE Transactions on Computers* and *ACM Computing Surveys*, and editorial board member for Elsevier's *Computers & Electrical Engineering* journal, *International Journal of Computers & Their Applications*, and *Journal of Parallel and Distributed Computing and Networks*. He has also served as a guest-editor for several special issues on high-performance computing architectures and networks in related journals, and as a chair/program chair for several international events.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.