

# High Performance Computing Co-Design Strategies

James A. Ang  
Sandia National Laboratories  
Center for Computing Research  
Albuquerque, NM 87185-1319  
+1-505-844-0068  
jaang@sandia.gov

## ABSTRACT

The MEMSYS Call for Papers contains this passage: *Many of the problems we see in the memory system are cross-disciplinary in nature – their solution would likely require work at all levels, from applications to circuits. Thus, while the scope of the problem is memory, the scope of the solutions will be much wider.*

The Department of Energy's (DOE) high performance computing (HPC) community is thinking about how to define, support and execute *work at all levels* for the development of future supercomputers to run our portfolio of mission applications. Borrowing a concept from embedded computing, the DOE HPC community is calling our work at all levels *co-design* [1]. Co-design for embedded computing is focused on hardware/software partitioning of activities to execute a well-defined task within specific constraints. Co-design for general-purpose HPC has many dimensions for both the work to be performed and the constraints, e.g. hardware designs, runtime software, applications and algorithms. The subject of this extended abstract is a description of two alternative DOE HPC co-design strategies. While DOE co-design efforts include more than the memory system, as noted in the MEMSYS call, the memory system impacts applications, circuits and all levels between.

## Categories and Subject Descriptors

• Computer systems organization~architectures • Computing methodologies~Massively parallel and high-performance simulations.

## Keywords

High Performance Computing; Co-design; Exascale Computing Initiative;

## 1. BACKGROUND

In the 1990's the DOE high performance computing (HPC) community shifted from the use of custom vector processors and memory, e.g. Cray vector supercomputers, to the use of systems based on the integration of commodity computing components into large-scale massively parallel processors (MPPs.) This was a very effective strategy because it rode the dual benefits of

Moore's Law and Dennard scaling. There have been opportunities for DOE to invest in technologies that improve the scalability of MPP systems, for example to develop lightweight kernel operating systems [2], or to improve the performance of interconnection networks [3]. But the majority of the components in MPPs are commodity off the shelf (COTS) computing components.

Since the end of Dennard Scaling over a decade ago, and the subsequent introduction of multi-core processors and many-core accelerators we have seen the commodity computing ecosystem depart further and further from the DOE's needs for HPC. To a large degree this is because multi-core and many core-processors exacerbate the memory wall [4]. The HPC community is on the precipice of a new era in supercomputing. Unfortunately we do not yet know what will replace the MPP era. This is why there is an international race underway to establish major research and development programs in exascale computing. With active programs underway in China, Europe, and Japan, the Department of Energy is working to establish the U.S. Exascale Computing Initiative (ECI) [5].

## 2. DOE CO-DESIGN STRATEGIES

The DOE has defined a co-design approach for the development of HPC capabilities and for several years has invested in the development of a key portfolio of co-design capabilities [6]. These include: proxy applications, e.g. Mantevo mini-applications [7], architectural simulation frameworks, e.g. the Structural Simulation Toolkit [8], and advanced architecture testbeds. The ECI supports the ability for the DOE to pursue two distinct co-design strategies where one is application-centric and the other computer architecture-centric.

### 2.1 Application-centric Co-design

*Co-design with hardware and system architectures largely predetermined using a clean sheet approach to the application development.* A concrete example of this Co-design strategy was set in motion last year when the DOE's Advanced Simulation and Computing (ASC) program awarded the *Trinity* platform to Cray [9], for a system that will use Intel's Xeon Phi Knights Landing (KNL) processors [10]. A key architectural change in KNL is the integration of Micron's Multi-Channel-DRAM, which provides a high bandwidth scratchpad memory albeit of limited capacity. In response to this pending architectural change, a Sandia and University team collaborated on an algorithmic and architectural analysis of how to refactor a sorting algorithm to leverage the capabilities of the KNL's two-level memory system [11]. With DOE support, this type of analysis will expand to cover more applications and algorithms.

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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM First International Symposium on Memory Systems, October 5–9, 2015, Washington, DC, USA.

Copyright 2015 ACM 1-58113-000-0/00/0010 ...\$15.00.

The recent announcement by Intel and Micron on their 3D-XPoint technology [12] and previous announcements from HP Labs on memristor devices for universal memory [13], means this strategy will grow. New application-centric co-design efforts are needed to understand how these new memory designs can address performance limits for DOE multi-physics applications with very large sparse linear systems. The ASC program calls this strategy Advanced Technology Development and Mitigation (ATDM). The DOE ECI program will allow this application-centric co-design to expand beyond the initial efforts with one multi-physics application per lab. But there will probably not be enough ECI budget to scale this strategy to the entire portfolio of ASC legacy applications, and furthermore, ASC does not have enough application and algorithm developers to rely solely on this clean sheet application strategy. In short, DOE and ASC need a complimentary co-design strategy.

## 2.2 Architecture-centric Co-design

*Co-design with applications and algorithms largely predetermined using a clean sheet approach to the hardware/system architecture development.* Given our portfolio of legacy application codes, our architecture-centric approach pursues clean-sheet development of revolutionary hardware and system architectures including associated system software, which is required to bridge to the DOE application code base. This strategy will support efforts such as the development of modules of chains of stacked DRAM to increase capacity and resilience of the memory system on what may be a single tier of main memory [14]. A research and development investment in this type of capability will have synergy with a large base of legacy scientific and engineering applications that exist within DOE and in a broad range of commercial and industry sectors. The DOE ECI provides both the funding and the longer time frame to pursue this strategy that maps to architecture-centric co-design with a requirement to “bridge” to the ASC portfolio of legacy applications. To create a foundation for ECI, the DOE has funded Industry-led architectural research and development efforts since 2012 [15]. Under ECI, this architecture-centric strategy complements the application-centric strategy by focusing a new set of research and development efforts with the U.S. computer industry to reduce the workload and effort that will be required of DOE application and algorithm developers.

## 3. CONCLUSIONS

Application-centric and architecture-centric co-design strategies, while distinct are not independent. A fundamental principle of co-design is that the multi-disciplinary process requires design space exploration with multiple iterations. While these distinct co-design strategies start with different assumptions, progress in each approach can inform the other. For example, *application-centric* co-design while focused on rewriting applications, can also inform hardware and system architecture design alternatives. Conversely, *architecture-centric* co-design can also inform changes to application and system software that help bridge to the DOE application portfolio.

Our strategy of creating supercomputers from the integration of commodity computing components may still be valid, but we need to see *if* and *how* we can influence future commodity computing components. The forthcoming ECI provides the DOE with the opportunity to extend the strategy of integrating commodity components into future supercomputers. But the last decade of limited MPP performance efficiency has demonstrated that current commodity component technology roadmaps will be

unable to support future DOE HPC requirements and constraints. Co-design is required for future COTS computing components to be useful to HPC.

## 4. ACKNOWLEDGMENTS

I thank my colleagues at Sandia’s Center for Computing Research. This is a special place that fosters multi-disciplinary discussions which are the foundation for holistic co-design. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

## 5. REFERENCES

- [1] DOE Advanced Scientific Computing Advisory Committee (ASCAC) Subcommittee Report 2014. *Top Ten Exascale Research Challenges* (February 10, 2014). <http://science.energy.gov/~media/ascr/ascac/pdf/meetings/20140210/Top10reportFEB14.pdf>.
- [2] Kelly, S. M., Brightwell, R. 2005. Software Architecture of the Light Weight Kernel, Catamount. *Sandia Technical Report, SAND2005-2780C*.
- [3] Brightwell, R., Pedretti, K. T., Underwood, K. D., Hudson, T. 2006. SeaStar Interconnect: Balanced Bandwidth for Scalable Performance. *IEEE Micro*, Vol. 3, Issue 3, pp. 41-57 (May-June 2006). <http://doi.ieeecomputersociety.org/10.1109/MM.2006.65>
- [4] McKee, S. A. 2004. Reflections on the Memory Wall. In *ACM Proceedings Computing Frontiers 2004*. (April 14-15, 2004).
- [5] <http://www.hpcwire.com/2015/07/28/doe-exascale-plan-gets-support-with-caveats/>.
- [6] Ang, J. A., Henning, P. J., Hoang, T. T., and Neely, R. *Advanced Simulation and Computing Program: Computing Strategy* (May 2013), pp.10-12. <http://nnsa.energy.gov/sites/default/files/nnsa/05-13-inlinefiles/2013-05-23%20ASC-StratV9.pdf>.
- [7] <http://mantevo.org>
- [8] <http://sst-simulator.org>
- [9] <http://insidehpc.com/2014/07/cray-wins-174-million-contract-trinity-supercomputer-based-knights-landing>.
- [10] <http://www.hpcwire.com/2014/06/24/micron-intel-reveal-memory-slice-knights-landing/>.
- [11] Bender, M. A., Berry, J., Hammond, S. D., Hemmert, K. S., McCauley, S., Moore, B., Moseley, B., Phillips, C. A., Resnick, D., Rodrigues, A. 2015. Two-Level Main Memory Co-Design: Multi-threaded Algorithmic Primitives, Analysis, and Simulation. In *Proceedings of the IEEE International Symposium on Parallel & Distributed Processing (IPDPS)* (2015), pp. 835-846.
- [12] <http://www.micron.com/about/innovations/3d-xpoint-technology>
- [13] <http://www.hpl.hp.com/news/2008/apr-jun/memristor.html>
- [14] Resnick, D. 2015. Opportunities to Upgrade Main Memory. In *Proceedings of the ACM International Symposium on Memory Systems* (2015).
- [15] [http://www.hpcwire.com/2012/07/12/doe\\_primes\\_pump\\_for\\_exascale\\_supercomputer](http://www.hpcwire.com/2012/07/12/doe_primes_pump_for_exascale_supercomputer)

