Co-design Breakout

Sudip Dosanjh and Stephane Requena

Co-design has been proposed as a key strategy for achieving Exascale computing. In this decade, industry roadmaps will not result an Exascale system that can solve a broad range of science and engineering problems within an affordable power budget. Architectures will need to change significantly to reach Exascale. At the same time, algorithms and applications will need to adapt to work effectively on these new systems. Two key elements of co-design are to (1) enable algorithms R&D for supercomputers that don't yet exist, are much different from today and are not well-defined and (2) help prioritize hardware/architectures research and development investments. In particular, we need to quantify the benefit of architectural choices relative to costs in energy, silicon area and R&D. The complexity is that these tradeoff studies must be done for algorithms and applications that are yet to be developed and defined.

The co-design breakout session at the Cologne IESP meeting was very well attended and included researchers from the U.S., Europe, Japan and China. Representatives from several major computer companies including Bull, Cray, IBM, Intel and Nvidia attended. The outline was to discuss:

(1) Co-design centers
(2) Whether the high performance computing community can really influence microprocessors, memory and architectures
(3) Co-design methodology
(4) Co-design and the software stack
(5) International collaborations

Co-design Centers

United States

The U.S. Department of Energy's Office of Advanced Scientific Computing Research has established three co-design centers on advanced reactors, combustion, extreme materials and nuclear energy. Several computer companies (AMD, Cray, HP, IBM, Intel and Nvidia) are participating in these centers. All three will be developing mini-applications that are representative of their computer codes to aid in the co-design process.

The advanced reactors co-design center is led by Argonne National Laboratory. Partners include Lawrence Livermore, Los Alamos, Oak Ridge, Pacific Northwest, Rice, Texas A&N, AREVA, General Atomics, and TerraPower. The goal of the center is to enable complete and detailed nuclear power system simulations to compress the nuclear regulatory timeline by guiding expensive experimental efforts, to prototype new designs, to study accident scenarios, and to narrow design margins.

The Combustion co-design center is led by Sandia National Laboratories. Partners include NREL, Los Alamos, Lawrence Livermore, Lawrence Berkeley, Oak Ridge, Stanford, Georgia Tech, University of Utah and University of Texas. The goal of the center is to (1) redesign all aspects of the combustion simulation process from core simulation and analysis methodology to programming models and languages to enable high-fidelity combustion simulations on exascale architectures and (2) identify key hardware features that have a critical impact on combustion simulation performance

The extreme materials co-design center is led by Los Alamos National Laboratory. Partners include Lawrence Livermore, Oak Ridge, Sandia, Stanford and Caltech. The goal of the center is to establish the interrelationship between hardware, middleware (software stack), programming models, and algorithms required to enable a productive exascale environment for multiphysics simulations of materials in extreme mechanical and radiation environments.

<u>Europe</u>

One of the major recommendations of the FP7 EESI project (European Exascale Software Initiative[1]) disclosed during the Barcelona conference on October 2011 was to setup in Europe between 5 to 10 specific Co-design centers, starting with Life Sciences & Health, Climate and Earth Sciences and Engineering.

However there are already existing initiatives/projects in Europe which implement co-design methodologies on various domains:

- The 3 European Intel Exascale Labs established in Belgium[2], France[3] and Germany are targeting co-design of specific applications in the field of space weather, Oil & Gas, Life Sciences and Particle Physics.
  The ECR (Exascale Computing Research) joint laboratory in France with Intel, CEA, GENCI and UVSQ is also implementing a full co-design methodology based on tools for automatic generation of codelets (or kernels/skeletons), performance characterization (based on static and dynamic analysis) and performance prediction using micro-benchmarks and simulators.
- CERFACS[4] is a French research organization that aims to develop advanced methods for the numerical simulation and the algorithmic solution of large scientific and technological problems of interest for research as well as

[1] http://www.eesi-project.eu/pages/menu/homepage.php
[2] http://www.exascience.com/
[3] http://www.exascale-computing.eu/
[4] Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (www.cerfacs.fr)

industry. Like the US Co-esign center for combustion, CERFACS is very known for co designing with IFP and industrial partners (Snecma, AirLiquide, Alstom, …) massive parallel CFD codes for hi-fidelity LES simulation applied to the combustion process (turbines or piston engines) or aerodynamics.

- The Swiss Platform for High-Performance and High-Productivity Computing (HP2C5) consists of domain science projects that are lead by research groups at Swiss universities and Institutes of the ETH Domain, and supported by a core group of scientific computing experts in the Lugano area.
  HP2C is jointly operated by the Swiss National Supercomputer Center (CSCS) and the Institute for Computational Sciences of the University of Lugano (USI). Current projects include DFT simulations, cosmology, fusion, climate or high-resolution earthquake modelisation.
- The JSC Simulation Laboratories6 at Juelich (Germany) are performing domain-specific research and support on scientific fields like Computational Biology, Molecular Systems, Plasma Physics and climate modeling.
- The Collaborative Research into Exascale Systemware, Tools and Applications (CRESTA[7]), a FP7 project leaded by EPCC with 12 more academic and industrial partners. CRESTA has two integrated strands: one focused on enabling a key set of co-design applications for exascale, the other focused on building and exploring appropriate systemware for exascale platforms.
  The six co-design vehicles represent a group of applications used by European academia and industry to solve critical grand challenge issues, including: biomolecular systems, fusion energy, the virtual physiological human, numerical weather prediction and engineering

<u>Japan</u>

Japan will be launching a co-design effort that will be described at the next IESP meeting. Approximately three co-design subject areas are being identified.

<u>China</u>

China has embraced co-design as a key element of its strategy. This is partially based on the observation that applications are only achieving a small fraction of peak on current Petascale systems and this situation will likely only become worse in the future. Several application areas are being considered:

- Aircraft Design
- Spaceship

---

[5] http://www.hp2c.ch
[6] http://www2.fz-juelich.de/jsc/simlabs
[7] http://www.cresta-project.eu/

- Drug Design
- Animation
- Mechanics of Giant Engineering Equipment
- Electromagnetic Environment Simulation
- New Type Material Design

Can the high performance computing community really influence microprocessors, memory and architectures?

All the companies present (Bull, Cray, IBM, Intel and Nvidia) said that HPC is an important customer and that there are many opportunities to have influence. The longest lead time item is microelectronics. Networks can be influenced later in time. The shortest lead time is software (but key R&D must still be started early). The HPC community must understand and leverage industry roadmaps. We need the ability to do a deep dive and to develop abstractions. Frequent communication is needed and computer companies should be engaged early. There was considerable discussion about what is realistic. There was a strong sentiment that a realistic view must be developed between the systems and applications community and that, in particular, 128 PB of memory, was not realistic. The concern was that unrealistic targets might send the computer companies in a fundamentally wrong direction. It was also noted that the research for 2018 processors will be done during the next two years.

Co-design methodology

Kernels (at the core level), compact applications (at the socket level and up to a few nodes), skeleton applications (at the system level) and full applications all have an important role. We must represent the breadth of applications and capture both current and future needs. However, we will still need permanent communication between these communities (vendors/hpc experts and end users). Compact and skeleton applications must evolve through co-design and their behavior needs to be validated against full applications. Kernels are most useful at an early stage in conjunction with simulators . Due to code complexity and the size of datasets, full applications are typically not used until prototype hardware is available.

Performance models, simulation and emulation are all critical. Also useful would be tools for automatically extracting kernels/skeletons and constructing performance models. Open simulation tools are needed that can interface to proprietary models and these tools should be

- Modular so that a variety of processor, network and memory models can be used

- Multilevel so that some components can be modeled in greater detail (otherwise simulations can take too long to perform)
- Holistic so that system attributes like power and reliability can be calculated in addition to performance

It is also useful if these simulation tools can be coupled to emulation to provide faster turn-around times and to study more complex applications. Validation of these models will be a key concern.

Co-design and the software stack

Key observations from the breakout session were that

- A reduced number of software stacks would be useful
- Software should be open source whenever possible
- A mechanism for sharing and coordination across the co-design centers is needed and the establishment of Exascale Software Centers for managing the entire software stack would be very useful.

International collaborations

Key conclusions were:

- Applications/software/architectures communities need a forum to openly exchange information, lessons learned
- There was general agreement that we should continue co-design methodology discussions within IESP. A standing breakout on co-design methodologies in the US, Europe and Asia would be useful and it was recommended that we hold a deep dive at next IESP meeting.