Update

# High-Performance Computing in Practice: HPC User Forum, September 15–17, 2014, Seattle, Washington


Earl C. Joseph, Ph.D.    Steve Conway
Robert Sorensen    Chirag Dekate, Ph.D.


## IN THIS UPDATE

This IDC update captures part of the proceedings at the 54th HPC User Forum held in Seattle, Washington. Organizations worldwide are taking decisive action to develop high-performance computing capability through regional, national, and international initiatives. Key leaders in the high-performance computing (HPC) industry spoke at the HPC User Forum in Seattle. Presentations made by Stephen Bique, Naval Research Laboratory; Jan van Lunteren, IBM Research Labs, Zurich, Switzerland; Michael Brown, Intel; Igor Bolotnov, North Carolina State University; and Jack Deslippe, NERSC, are captured in this document. Key vendor updates from Intel and HP have also been captured. Advances in liquid cooling technologies presented by Steven Hammond (NREL) and Jon Summers (University of Leeds) are also captured in this document. Bique from the Naval Research Laboratory presented initial experiences in programming using Xilinx Virtex-6 FPGAs inside Convey's HC-1ex system. Lunteren from IBM Research Labs presented details about memory-driven near-data acceleration and its application to DOME/SKA program. Brown from Intel presented advances in alternative energy and bioengineering simulations driven by many-core processors and accelerators. Deslippe from NERSC highlighted the use of many-core processors with BerkeleyGW Code. Bolotnov from North Carolina State University presented advances in multiphase flow modeling and simulation. Hammond from NREL showcased utilization of warm water cooling in NREL supercomputing ecosystems. Summers from the University of Leeds shared results of a study comparing liquid cooling methods.

High-performance computing innovation will be driven by disruptive developments in several of the enabling technologies. Many of the HPC industries will be key in driving the innovation to the market and packaging it for broader consumption. IDC believes that innovations and investments in high-performance computing will be needed on a sustained basis to catalyze regional and national high-performance computing efforts.

## "Initial Experiences Programming Xilinx Virtex-6 FPGAs Inside Convey's HC-1ex": Stephen Bique, Naval Research Laboratory

Bluespec SystemVerilog is a modern language used in the design of electronic systems. An FPGA is a large array of configurable logic blocks.

Xilinx has been an FPGA vendor for three decades. Xilinx Virtex-6 has slices for doing 18 x 25 multiplies. The Convey HC-1ex has a Xeon quad-core processor and a custom board with four Virtex-6 LX760s. There's tight integration with the x86 server. Coprocessor memory looks like standard memory to the x86 — it is cache coherent with the x86's own memory. FPGAs avoid the instruction/fetch/decode bottleneck of Von Neumann architectures. Convey's memory system was designed to avoid bottlenecks.

Bluespec SystemVerilog — I chose this because you can design a circuit with it; yet it's still a high-level language so I don't have to manage a lot of details. You can verify correctness using atomic rules. You can do computations on the FPGA using usual block matrix multiplication. Our initial experience tells us to use brute force by programming with 64-bit integers. Synthesis would not succeed due to 64-bit multiples. Virtex-6 has only an 18 x 23 multiplier. No runtime support for 32-bit numbers, so this needed a custom implementation. Optimizations reduced complexity.

Current work:

- Forget about alignment — vectors can be anywhere in memory
- Computations happen independently of reading and writing
- Don't interleave reads and writes
- Use multiple time-multiplexed MACs

Future work:

- Implement double-precision version
- Design a sparse matrix-matrix solution
- Find other applications

Comment: Why did you pick FPGAs instead of GPUs?

Speaker: I like FPGAs. They're so much more powerful. GPUs have an advantage because they've always been part of systems.

## "Memory-Driven Near-Data Acceleration and Its Application to DOME/SKA": Jan van Lunteren, IBM Research Labs

Project goal is to develop technology for the Square Kilometer Array (SKA) telescope that will be deployed in Australia and South Africa. SKA will have thousands of antennas that will deliver an exascale per day of data.

Extrapolating the current HPC trends would result in 37GF/W in 2022 (20% efficiency). We need to improve this. CPU, GPU, FPGA, and DSP do not meet performance or energy efficiency targets. ASIC can meet performance and power-efficiency targets but do not provide enough flexibility and have high development costs. Memory system bandwidth/latency/power depends on complex factors. Operation is typically fixed and can't be adapted to workloads.

Our approach is to utilize programmable near-memory processing. This approach keeps data close to memory to reduce data movement and makes memory system operation programmable. Integration could be done through on die with eDRAM or with stacked memory modules.

Near-data acceleration involves a decoupled access/execute architecture. Access processor handles all memory access. Execution pipelines are configured by the access processor as is the tag-based data transfer. All operand values in the execution pipeline input buffer/register trigger execution.

New programmable near-memory accelerator made feasible by novel-state machine and address mapping technologies. Goal is to minimize energy overhead beyond basic storage and processing needs. Proof of concept happened for selected workloads using FPGA and initial compiler stacks.

## "Advancing Science in Alternative Energy and Bioengineering with Many-Core Processors and Accelerators": Michael Brown, Intel

This presentation includes results from my tenure at ORNL. Liquid crystal biosensors commonly used for optical switching and biosensing applications can potentially perform label-free diagnostics. Alignment changes induced by interactions with biomolecules can be detected with optical microscopy. Before 2013, there were no molecular simulations at experimentally relevant scales.

Another problem: icephobic surfaces for wind power. Ice formation can reduce turbine efficiency and is hard to follow experimentally. Organic solar cells are promising renewable energy sources: low cost, flexible, lightweight. This is a multiscale problem. Potential benefits of many core and GPUs for multiscale simulations include power consumption and heat dissipation.

The goal of CAAR at ORNL is to explore potential performance gains with the heterogeneous Cray Titan supercomputer. This includes study of phenomena like spinoidal de-wetting (prevent defects from forming in liquid crystal layers). Simulations like these are not possible to study on Jaguar without accessing the entire system. Resultant simulations demonstrate a 7x performance gain (1 CPU plus GPU versus 2 CPUs).

Icephobic surfaces: Probe for water droplet freezing on a surface with molecular detail. Result: Simulation is now typical with hundreds of nodes, thanks to 2-3x speedup. Issues with heterogeneous machines. Programming model: Using CUDA and C++ introduces multiple languages, codes paths, etc. Optimizations for GPUs don't necessarily improve performance of the code on the CPUs. Future Intel Phis will be bootable.

Current Intel optimizations:

- Data alignment
- Support for mixed and single-precision modes
- SIMD directives for compiler vectorization for routines with data dependencies
- Modifications for conditional branches
- Neighbor-list padding

- Offload directives to manage data allocation, transfer, and concurrent computation on the code processor

Main advantage of Phi versus GPU is that you can use the same code on both devices. Future improvements:

- Compiler vectorization
- MPI performance in symmetric codes

### "Use of Many-Core Processors with the BerkeleyGW Code": Jack Deslippe, NERSC

BerkeleyGW is a material science code that came out of the Berkeley physics department. It's an example of big data coming into HPC. Density functional theory (DFT) GW fixes some of the failings of DFT. DFT is not good at predicting electron properties in the excited state as opposed to ground state. Systems excited by light are very important today for complex materials used for photovoltaics, LEDs, nanomaterials (e.g., nanotubes), and more.

BerkeleyGW started in the 1980s and spawned lots of versions; some having "basic" parallelization with MPI. The code was prohibitively slow for large systems and way more expensive than DFT. I and others moved the code into the MPP era with minor improvements such as version control, profiling, and more. We rewrote the code to address computational bottlenecks.

NERSC-8 will begin the transition to more energy-efficient architectures. We're working with 20 teams to get users ready for this architecture. BerkeleyGW is one of the selected codes. The MPI-only programming model is not well suited for BerkeleyGW. Up to 90% of the computing capability can be lost to overhead.

## Key Vendor Updates

### *Mike Lafferty, Intel Update*

Yesterday we repeatedly heard that software needs to be better addressed. Modernizing code: What you do for Xeon Phi will likely benefit Xeon too. Knights Landing is a good place to do this because optimizations will carry forward to future Intel processors. We believe, based on benchmarks, that from a performance standpoint, you're just fine with Xeon Phi compared with NVIDIA, and from a programming standpoint Phi is a better choice.

People can apply online for Intel grant money to modernize codes. To get started with Phi, you can buy a starter kit from various vendors for under $5,000, or you can ask Intel for a free sample, or you can request a loaner from a vendor and we'll help you with tools. There are many resources on our MIC developer Web site. We also take requests to do a one-day workshop at your site if you have enough people.

## *Ed Turkel, HP Update*

Customers in the HPC markets and segments we're in are telling us what they need next. Also important for us is making accessibility to HPC easier with our products and services and the cloud. There's increasing demand for our HPC services, since customers are saying they don't want to manage their next big system themselves or they want to pay based more on performance. We have designed systems more and more for specific workloads.

### ARM

Our customers will tell us where this best fits. People today are trying out Moonshot. We'll see a whole new set of applications for this rather than the traditional set of HPC applications. HPC (density-optimized servers) part of HP has strong margins and is expected to become half of all our server business in not that many years, thanks to its growth and not because other parts of the business will shrink.

## "Multiphase Flow Modeling and Simulation: HPC-Enabled Capabilities Today and Tomorrow": Igor Bolotnov, North Carolina State University

North Carolina State College of Engineering; 9,000 students. Nuclear engineering department has been growing quickly. CFD-DNA code: PHASTA is a parallel, hierarchic, high-order accurate, adaptive code for multiphase flows, biology, nuclear fuel flow, etc. Using ANL's BG/Q "Mira," complex wing design on 11 billion element mesh, on up to 3 million parts. We have demonstrated strong scaling with 5 billion elements up to 294,912 cores on JUGENE and up to 98,304 cores on Kraken at the University of Tennessee.

One simple example involves lift force. Four factors govern it:

- Relative velocity
- Shear rate
- Bubble rotational speed
- Bubble surface boundary condition

Also use the code to model lift and drag coefficient versus shear rate, for applications such as what happens in nuclear reactors' cores. We also simulated realistic reactor spacer grids and mixing vanes for turbulent flow.

Future virtual experiments:

- Increasing fidelity of bubbly flow simulations
- Incorporate complex geometry analysis for nuclear applications
- Phase change
- Boiling flows with challenging conditions

PWR-relevant problems sizes:

- **Current runs:** Mesh size 190 million elements, Reynolds number 180
- **2015:** 2 billion elements, Reynolds number 500; 512,000 BG cores at 4,000 elements/core
- **2020:** 355 billion elements, Reynolds number 5,000; 90 million cores at 4,000 elements/core

Larger domain problem sizes:

- **2060:** Simulation of an entire reactor core could use 40 trillion elements with 320 billion processor cores at 128,000 elements/core.

## "Bits, Bytes, and BTUs: Warm Water Liquid Cooling at NREL": Steven Hammond, NREL

We are a DOE lab outside of Denver. Datacenters are energy intensive — 10–100x more than a typical office. Power and cooling constrains existing facilities. EPA estimate is that 3% of U.S. electricity goes to datacenters. Demand is surging. Densities are increasing. Silicon photonics is promising but also has thermal issues.

Why care?

- Carbon footprint
- Water use
- Limited utility power
- Cost

We pay about $1 million per kilowatt per year. Most sites have a relatively fixed electrical supply. It can be costly and disruptive to add more.

Move to liquid cooling:

- Fans are inefficient and noisy
- Rising densities require component-level liquid cooling
- Air cooling — humidity, fan failures, hot spots, more
- Liquid issues — pH, bacteria, types of pipes, corrosion inhibitors

NREL datacenter:

- 10MW, 10,000 sq ft
- Leverage climate
- No chillers

- Waste heat heats labs and offices (35C/95F)
- PUE 1.06

Planned to accommodate 5–6 HPC system generations. Warm water cooling: 24C/75F, 480VAC. Estimated 10-year savings is $24 million plus $200,000/year savings in heating labs and offices. ASHRAE TC9.9 liquid standards are an excellent guide. We run our disks as warm as 75C.

## "Results of Study Comparing Liquid Cooling Methods": Jon Summers, University of Leeds

We have 15 grad students in our Institute of Thermofluids. Liquid cooling (LC) of electronics has been around for a long time, at least since the 1960s — 1973 paper on "liquid immersion cooling of small electronic devices" from Bell Laboratories. Traditional liquid cooling brings liquid to the edge of the datacenter. Better is to bring them closer to the computers/components.

- **Indirect LC:** 100% air, with liquid to the rack
- **Direct LC:** Some air, with liquid to the rack
- **Total LC:** No air, with liquid to the rack (We found each air-cooled server is 230W, while liquid cooled is 45W.)

Total LC by full immersion: Vendor Iceotope and 3M, University of Leeds, won the 2013 Green Enterprise IT Award awarded by Uptime Institute. All electronics immersed in dielectric liquid. Heat transfer by natural convection into the lab; then from the rack to a secondary heating water circuit. We use a 20W central heating pump. Partial PUE 1.05. ILC: 80–91% heat capture in liquid at 18C; DLC: Up to 78% capture in liquid at 25C; and TLC: 89% capture in liquid at 45C.

# LEARN MORE

## Related Research

Additional research from IDC in the technical computing hardware program includes the following

- *Worldwide Broader HPC 2014–2018 Forecast: Servers, Storage, Software, Middleware, and Services* (IDC #248835, June 2014)
- *When Massive Data Never Becomes Big Data* (IDC #lcUS24922014, June 2014)
- *Worldwide Technical Computing Server 2014–2018 Forecast* (IDC #248779, May 2014)
- *Perspectives on High-Performance Data Analysis: The Life Sciences* (IDC #248348, May 2014)
- *Global HPC Market Dynamics in 2013* (IDC #248137, April 2014)
- *Industrial Partnership Programs and High-Performance Computing: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico* (IDC #248113, April 2014)

- *Disruptive Technologies in High-Performance Computing: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico* (IDC #248112, April 2014)
- *Advances in Processors, Coprocessors, and Accelerators in High-Performance Computing: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico* (IDC #248111, April 2014)
- *International Perspectives on Industrial High-Performance Computing Partnerships: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico* (IDC #248122, April 2014)
- *Worldwide HPC Public Cloud Computing 2014-2017 Forecast* (IDC #247846, April 2014)
- *Summary of IDC's 2014 Research in the Use of HPC by Oil and Gas Organizations* (IDC #247704, March 2014)
- *IBM Sale to Lenovo Opens Opportunity for Other HPC Vendors* (IDC #lcUS24694314, February 2014)
- *IDC's Worldwide High-Performance Computing Predictions 2014* (IDC #WC20140211, February 2014)
- *Seagate Looking for the X Factor in Its Acquisition of Xyratex* (IDC #lcUS24555413, December 2013)
- *Micron Demonstrates Technologies to Address Emerging Challenges in Big Data Applications* (IDC #244843, December 2013)
- *Market Analysis Perspective: Worldwide HPC, 2013 — Directions, Trends, and Customer Requirements* (IDC #244742, December 2013)
- *HPDA Pulse: 2013 Software and Consulting Market Analysis* (IDC #244513, November 2013)
- *HPDA Pulse Results: 2013 Hardware and Storage Market Analysis* (IDC #244493, November 2013)
- *HP FY13: Revenue Declines Abate on Stronger Core Business* (IDC #lcUS24466413, November 2013)
- *Catalyst Supercomputer Heralds Shift to More Balanced Architectures* (IDC #lcUS24437513, November 2013)
- *China Eyes 10,000-Fold Data Reduction for Internet of Things* (IDC #lcUS24392513, October 2013)
- *HPC User Forum, October 2013, Seoul, Korea* (IDC #243786, October 2013)
- *Tools and Techniques for Technical Computing in Life Sciences: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243778, October 2013)
- *Perspectives on Quantum Computing: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243777, October 2013)
- *National and International Initiatives: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243776, October 2013)
- *Issues in High-Performance Computing: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243775, October 2013)
- *High-Performance Data Analysis in the Life Sciences: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243774, October 2013)

- *Chinese Research in Processor Designs for High-Performance Computing and Other Uses* (IDC #243502, October 2013)
- *World's Fastest Supercomputer Set to Reach Customer in October 2013* (IDC #lcUS24300913, September 2013)
- *The Broader HPC Market 2012-2017 Forecast: Servers, Storage, Software, Middleware, and Services* (IDC #242742, August 2013)
- *IDC's Worldwide Technical Server Taxonomy, 2013* (IDC #242725, August 2013)
- *China Regains Top Supercomputer Title* (IDC #lcUS24190613, June 2013)
- *10 Things CIOs Should Know About High-Performance Computing* (IDC #241565, June 2013)
- *Worldwide High-Performance Data Analysis 2013-2017 Forecast* (IDC #241315, June 2013)
- *Top Issues for HPC Sites: HPC User Forum, April 29-May 1, 2013, Tucson, Arizona* (IDC #241463, June 2013)documents:

## About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

## Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com