

HPC User Forum Update

Interviews with Steering Committee Members: Merle Giles, Moonshot Research

Steve Conway and Thomas Gerard
July 2020

IN THIS UPDATE

After the global pandemic forced Hyperion Research to cancel the April 2020 HPC User Forum planned for Princeton, New Jersey, we decided to reach out to the HPC community in another way – by publishing a series of interviews with members of the HPC User Forum Steering Committee. Our hope is that these seasoned leaders' perspectives on HPC's past, present and future will be interesting and beneficial to others. To conduct the interviews, Hyperion Research engaged insideHPC Media. We welcome comments and questions addressed to Steve Conway, sconway@hyperionres.com or Earl Joseph, ejoseph@hyperionres.com.

This interview is with Merle Giles. Giles is a tireless advocate for HPC's role as an accelerator to industrial innovation. While he was at NCSA, his team partnered with nearly 60% of the manufacturers in the U.S. FORTUNE100®, as well as with bio-medical, chemical, tech, oil and gas, and agriculture companies. He and his corporate partners were founding members of two key digital manufacturing consortia: NDEMC (National Digital Engineering and Manufacturing Consortium), a \$5 million public-private partnership pilot serving small and medium manufacturers in the American Midwest, and DMDII (Digital Manufacturing and Design Innovation Institute), a \$320 million partnership announced in 2014 as one of the USA's National Network for Manufacturing Innovation (NNMI) institutes. He is a co-founder, along with Germany's HLRS and South Korea's KiSTi, of the International Industrial Supercomputing Workshop. He earned undergraduate degrees in accounting and business and holds a CPA and an MBA from University of Illinois. Giles is also co-editor of the book, *Industrial Applications of High Performance Computing: Best Global Practices* (CRC Press, 2015).

The HPC User Forum was established in 1999 to promote the health of the global HPC industry and address issues of common concern to users. More than 75 HPC User Forum meetings have been held in the Americas, Europe and the Asia-Pacific region since the organization's founding in 2000.

MERLE GILES INTERVIEWED BY DAN OLDS, insideHPC

Olds: Today we have Merle Giles, who we are going to be talking to about his experience in HPC. So, how did you become involved in HPC, Merle?

Giles: Sometimes I like to say I couldn't even spell HPC when I got into it. I had been at the University of Illinois running the executive MBA program for the business school. I'm a business guy myself, had an MBA from that very program, ended up running it for a while, and left because we had moved it to Chicago. But anyway, I fell into the position at NCSA, at the university, as director of what they called at that time the private sector program. I built an industry-facing HPC services operation right here in Illinois.

Olds: Can you summarize your background in HPC?

Giles: We learned that there were barriers to access and education and all sorts of the classical things that we've talked about for years. As a student of the business models, I learned that the research machines are wonderful but not always accessible. We had as partners in our industry program a fair number of manufacturers, just being here in Illinois and dealing with other U.S. companies. The manufacturing business is heavily dependent on commercial codes that do simulation, and the commercial code developers didn't get access often enough to HPC systems. We actually began to focus on how to help the commercial side access systems designed for what manufacturers care about, so we built a system ourselves with different kinds of access models and ran it for industry-only, and began to get very deeply into the problem sets with these companies, many of which were Fortune 100. We learned a ton and like to think we helped a lot.

Olds: And this is before the Fortune 100 started building out their own internal HPC capabilities?

Giles: Oh, no. The Boeings of the world, the large ones, the big boys have been doing this for 50 years in one form or another. They all had high interest in HPC. In fact, they were running clusters of their own. But as an HPC center our attitude was, if we can't run HPC better than the average manufacturer, then we shouldn't be in the business. So, that was our goal - let's run this better and let's work on some things to break down the barriers in this space.

Olds: So, what are the biggest changes that you've seen in HPC during your career?

Giles: For decades, even from the early 1980s when the NSF systems came about, HPC has had this high-level, lab-level focus on some of the world's toughest problems. And those are research problems of their own. But, if you step back just a bit, the companies – manufacturers, biotech firms and so-forth – also have exceedingly difficult challenges, but not always at the scale that those large machines are built for solving. So, the access becomes kludge code maybe. We did a lot of work with Fluent, because we had all sorts of manufacturing communities that cared about Ansys Fluent, for instance. But they didn't have people to hold their hands to actually make HPC systems work better for their problems. So, we did research projects, we tried to build systems that were very good for running codes like this. The main change I've seen in HPC has been what we learned working with in the second tier of HPC users, just below the big research labs. That is exceedingly important in the commercial space. It's not research HPC. So, that next tier actually brought us some problems that were pretty difficult. And then that tier, of build-the-cluster-on-premise, has now moved a little bit and is shifting to the cloud. So, that's some of the biggest changes I've seen. The supply chain in manufacturing, for instance, can access cloud HPC in ways that would beat what they could do on-

prem. An example is the whole HPC emphasis that is really special at Microsoft Azure. Amazon does it a different way. Azure's got InfiniBand, for instance. So, the business model shift has been to greater access and, in theory, and I think in practice, there's more access further down the supply chain in a commercial domain that was once dominated by large companies.

Olds: That makes a lot of sense. So, given that, where do you see HPC going in the future, and is there anything you're particularly excited about or concerned about?

Giles: HPC as a narrow slice of computing, as one of the pillars of computing, is truly high-performant. But I once walked into a manufacturing company, months into my job, and said, hey, if we could help you speed up your simulation by 100 times, wouldn't you just love us? And the answer was no. And I'm thinking, oh my goodness, what did I just say? So, it got us into a conversation about workflow. Building the models at the manufacturing company actually takes more time than running the simulation. So, if we could help with the model building, actually it'd save more time overall, which is more expensive on the personnel side than the cost of running those simulations. The simulations are crucial because they are often seeking out parameters and trying to fix your product design.

The other thing I learned is that the engineers who use HPC are sometimes given just two weeks out of the entire design cycle inside the company to add their special sauce to the design. The designers are moving on in two weeks whether you've done your stuff or not. And it turns out that if you want to simulate oil in a differential, since it's an incompressible fluid, the math is exceedingly difficult. It would take a month to run a single job because it would only run on 32 cores. But gee whiz, the airplane engine guys are running on 1,000 cores with the same code. What's the difference? And my guys would say, "it's simple, it's the difference between compressible and incompressible fluids." Well, who knew? So, that's the difficulty of the physics. And if that oil in a differential simulation takes a month but you're only given two weeks to add something to the design, HPC stays in the R&D basement forever and never gets out. It's a long-term process. We think, okay, we're going to out-compute to out-compete, but in reality, if you can't do it in two weeks you have to do it another way. So, the old-school manufacturers will do it another way.

So, part of the challenge is to help the user spend more of their time doing engineering instead of wrangling access to the code and trying to do things HPC-like. So, there are lots of things that actually changed that rubric, but what I learned, and I think is a change in this industry, is that as we move HPC beyond research and into operations, it drives ROI because operations is where the money is. R&D is a fixed budget which everybody wants to squeeze. So, this workflow matters. The TRL [technology readiness level] system, for instance, when you go from research to different kinds of research, to the POCs [proofs of concept] and demonstrations and one day you're going to the moon, as we move up that chain HPC really does matter, but we don't often deploy it on the operational side of a company.

Olds: So, it's not deployed in the right place yet with a lot of companies?

Giles: Tom Lang at P&G, for instance, would claim that HPC was never a boardroom conversation because it was so focused on R&D. In today's world, AI *is* a boardroom conversation. Well, AI needs high performance computing, just of a different kind. So, if AI drives operations in a business and we can have computing to speed that up, everybody wins. In a Procter and Gamble situation, for instance, Tom would say very publicly that their challenge is in the last mile, delivery. You can make the product, you can make the product better, but distribution is a big deal. So, part of what I'm seeing today is a desire to get information as to what is going on under the hood of delivery vehicles and where those vehicles are going.

Olds: And to optimize.

Giles: And optimize, that's right. So, it's a big data problem. It's really a high performance data analytics problem. But the challenge is how to get that done. AI has brought computing to the boardroom conversation in a way that HPC historically had not.

Olds: In the example you were using, it's a travelling salesman problem on steroids.

Giles: It is. And real-time data analytics now is coming about in ways that can give some insight beyond batch analytics, for instance. So, I'm a fan of the movie *Ford vs. Ferrari*, with one of the most memorable parts of that movie being when they put Post-it notes on the car and drove it around the track pretty hard to find out how the wind was affecting the suspension. Real-time analytics offers the Post-it notes in ways that few other things can do. You come back, you park the car, you take apart things and look at it and say, well what went wrong? And the Ford driver is saying, it's doing something to me here and I've got to fix it. The classical mechanics cannot fix it, ergo the Post-it notes. They came up with something.

So those are the biggest changes in this space that I see. How do we offer Post-it notes to people if we actually use operations as the target? And now we are talking edge computing, we're talking 5G doing some things in IoT that have never been done before. But do we go into the R&D basement and do batch analytics to try to figure out what's happening in real-time on a manufacturing floor? No. Same story as classical HPC. If that's what happens, that's what you have to do to gain insight but you actually have to put it into rows and columns to figure out what's happening and know what to ask for. It's back to an R&D project. So, if it's R&D, if it's a research approach to an operational problem, we've got a gap.

Olds: And we have to bring it down to that on-the-street operational need. That's the key.

Giles: That's where the money is. So, let's think about an HPC deployment. At some point the opex [operational expenditures] become larger than the capex [capital expenditures]. The cost to run those CPUs at some point, let's just call it a year, and maybe you bought it to hold for three years, is going to eclipse how much you paid for them in the first place., So, do we have HPC systems instrumented to be able to optimize the costs, the operating costs? Not typically. We don't watch what goes under the hood and don't have all the data to do that. And if you do it in a classical batch analytics approach you might not be asking all the right questions, because you don't know what's happening under the hood. So, if you have real-time analytics in the form of Post-it notes you may have more insight and can lower your opex costs and affect operations.

Olds: You know, I love how you brought it back to the Post-it notes, and I think that's a great point to end on and keep in people's mind. You've got to find those Post-it note places in your operations.

Giles: That's where the biggest money is. And if you save money on the operational side, you can fund research in better ways. It pays for itself.

Olds: Great, well thank you so much Merle I really appreciate your time and I know that our listeners and watchers are really going to enjoy this. Thank you, again.

Giles: Dan, it's a pleasure, thanks for having me on.

About Hyperion Research, LLC

Hyperion Research provides data-driven research, analysis and recommendations for technologies, applications, and markets in high performance computing and emerging technology areas to help organizations worldwide make effective decisions and seize growth opportunities. Research includes market sizing and forecasting, share tracking, segmentation, technology and related trend analysis, and both user & vendor analysis for multi-user technical server technology used for HPC and HPDA (high performance data analysis). We provide thought leadership and practical guidance for users, vendors and other members of the HPC community by focusing on key market and technology trends across government, industry, commerce, and academia.

Headquarters

365 Summit Avenue

St. Paul, MN 55102

USA

612.812.5798

www.HyperionResearch.com and www.hpcuserforum.com

Copyright Notice

Copyright 2020 Hyperion Research LLC. Reproduction is forbidden unless authorized. All rights reserved. Visit www.HyperionResearch.com or www.hpcuserforum.com to learn more. Please contact 612.812.5798 and/or email info@hyperionres.com for information on reprints, additional copies, web rights, or quoting permission.