

# HPC User Forum Update

# Interviews with Steering Committee Members: Suzy Tichenor, Oak Ridge National Laboratory

Steve Conway and Thomas Gerard August 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, <u>sconway@hyperionres.com</u> or Earl Joseph, <u>ejoseph@hyperionres.com</u>.

This interview is with Suzy Tichenor, director, Industrial Partnerships Program for the Computing and Computational Sciences Directorate at Oak Ridge National Laboratory. The ACCEL Industrial Partnerships Program (Accelerating Competitiveness through Computational ExceLlence) provides companies access to the laboratory's world class computational science expertise and Summit, the nation's most powerful supercomputer for open science. Tichenor has more than 20 years' experience creating partnerships and programs at all levels of the government, private sector, and not-for-profit organizations. Prior to joining Oak Ridge, she was vice president of the Council on Competitiveness and directed its High Performance Computing Initiative.

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.

### SUZY TICHENOR INTERVIEWED BY DAN OLDS, insideHPC

**Olds**: Hello. Dan Olds here on behalf of Hyperion Research and InsideHPC and we have a great treat today. We're going to be talking to Suzy Tichenor and we're going to be talking about where HPC and industry come together. How you doing, Suzy?

Tichenor: Hi, Dan, I'm doing great. How're you?

Olds: I'm good, I'm just going through these virus-riddled times and still shut in.

Tichenor: You're healthy, though?

Olds: Yes, like a zoo animal.

Tichenor: Good. Consider it a good day. You're healthy, you're blessed.

Olds: Yes, and I assume you are, too?

Tichenor: Yes, I consider myself very blessed, thank you.

Olds: Very much so. So, tell me, how did you get involved with HPC?

**Tichenor:** Well, I began, really, in the Washington, DC office of Cray and I was handling a lot of their international trade issues. That really got me started and I did that for a number of years - working for Cray. I ended up at the Council on Competitiveness, which is a policy organization in Washington, and they had a very interesting project. They had funding from DoD and Department of Energy and the National Science Foundation, looking at high performance computing as a driver for economic competitiveness. And they were interested in understanding whether companies, industry, was using HPC as aggressively as they could and should. And if not, why not? What where the barriers? And then, what were the opportunities for public/private partnerships to address those barriers? What would be the role of the public sector and what would be the role of the private sector?

#### Olds: That's a big project!

**Tichenor:** Well, it was. I was there for four years running this and it was very interesting. Some of the key findings were that for the companies - we did a number of surveys, national surveys, conferences, and we had a wonderful advisory committee of heads-of-HPC from the national labs, heads-of-computing and R&D from a number of companies - so, in these surveys we found that for the companies that were using high performance computing it was really essential to their business survival. But there were barriers to using it for them and even for newcomers. Those certainly still exist, the same things. You would also find today that for companies that use HPC it's really important for their business survival. But the same barriers are still around: lack of access to talent, or maybe not being able to afford to find and hire the talent; ease-of-use issues, of course as architectures change that's always a challenge. Some of those barriers still exist.

And then one of the really interesting findings that came out was kind of an 'aha' moment. I remember it in a meeting talking about industrial use of HPC. This advisory committee included public and private sector people. The public sector participants on the advisory committee had the impression that companies didn't have large-scale HPC systems because they didn't have big problems, because the feeling was companies are for-profit entities. They invest where they need it, and if they had big problems they'd go out and buy big computers.

So, what the companies at the advisory committee said was, "No, you have a misconception. First off, we all have big problems sitting on the shelf, but our business and financial model doesn't allow us to invest tens of millions of dollars in a computer that's going to be out-of-date in a couple years. Our CEO's won't allow that, our boards of directors won't permit that. So, we have to wait for that performance to drop down into a lower price point so we can jump on, which is why we are always drafting behind you. But we have big problems. And, by the way, you've been talking about these great supercomputers you have, and DoE gets access to them and universities get access to them, why can't companies have access to them?" That was a big outcome of that project.

Olds: That's a huge 'aha' moment.

**Tichenor:** Well, it was! And DoE, in particular, took it very seriously. The first step was they opened up the INCITE program [Innovative and Novel Computational Impact on Theory and Experiment], which is one of the pathways to get time on a DoE supercomputer. The first year they opened it up three companies applied from our advisory committee, actually, and they were selected. Of course, it's all peer review. We had a conference, a user conference, that summer and the director for science for DoE came and gave a talk and he said that the highest-rated application that year in the INCITE program came from a company. It came from Pratt and Whitney. He said, "I have to be honest and tell you I was surprised that it was the highest rated, but as I read it, there was no question why it was the highest rated. These are complex problems that need leadership-scale computing." So, that opened up the idea that the DoE supercomputer centers should be more accessible to companies.

Olds: So, a huge success case right off the bat?

**Tichenor:** Well, this was at year four. Not right off the bat, it took a long time. There were a lot of discussions that came out, all these findings, we had surveys and conferences and meetings, this doesn't just pop out at you. It takes time. Then, you had to work with DoE, and DoE had to hit the cycle where they would open up INCITE. It took a while for this to happen. Then, once that happened the DoE labs started to say, "how can we be a part of this?" So, Oak Ridge, actually, was kind of the first there and stepped forward. They just made their own decision that they wanted to start an industrial partnership program for their computing center. That's when I joined Oak Ridge.

Olds: That's what brought you there?

Tichenor: That's what brought me to Oak Ridge, to help to launch that and then to manage that. And it's been, I think, pretty successful.

**Olds:** That's fantastic. So, looking back over the years, what are the biggest changes that you've seen throughout your career when it comes to HPC and industry?

**Tichenor:** Well, a number of things. First off, I think right now it's a wonderful time because there's a lot of opportunity to access high performance computing through different pathways. First off, there are a lot of opportunities at much lower price points than there were a number of years ago. There are smaller systems that you can have and bring in that are very powerful. There is cloud computing now, there is HPC on demand. That's hugely important. And now, at the top end, at the leadership computing end, you have the NSF supercomputer centers, they're open to industry, and you have the

DoE user facilities at the very highest end. They've made their supercomputers much more accessible to industry and I think that's been really helpful to a lot of companies.

Olds: And it's certainly been a successful partnership, right?

**Tichenor:** I would I think so. We're still doing it after ten years here. What's exciting is that we've seen companies start - I'm just speaking for our center; I'm sure the other supercomputing centers have their own stories. At Oak Ridge we made a conscious decision when we started this that we would not have a set-aside for industry - that industry was going to have to compete in each pathway. There were three different ways you could apply for time, but industry was going to have to compete in each one with everyone else. We weren't going to say, "okay, ten percent of the system is going to be for industry," or two percent or whatever it may be. They'd have to compete, and it was kind of hard in the beginning. They weren't used to writing the kinds of proposals that the DoE was looking for. There was a lot to learn. And the DoE side had to get reviewers that were used to reviewing industrial proposals because they were different kinds of problems. There were little stumbles along the way but everybody worked with each other and we saw companies come in and start with very small allocations through what we call our Director's Discretionary program, which is where companies, anyone, can get their feet wet and kind of prepare themselves for some of the larger calls for proposals.

We've seen companies move successfully now from those smaller allocations over several years, then to the ALCC [ASCR Leadership Computing Challenge] program and then to compete successfully for INCITE. It's been really exciting because companies are growing in their experience to apply high performance computing to real world problems and they're growing in that experience through access to the DoE systems. That's a wonderful contribution to competitiveness, it's great. And then, as those companies are progressing, they're also upgrading their own internal systems because they realize they can't go to DoE for everything. But as they get their feet wet then they go back and now they've built a real ROI case for the CFO.

Olds: And that's key to convince business management to invest the money.

**Tichenor:** That's right. Because they'll outsource that test case to a larger system at the DoE center and it'll be a proof of concept and they'll show what they could do if they had access to more. And then they can go back and use that to justify an upgrade. Now, they're probably not going to upgrade to what DoE has, but they could do a big upgrade, anyway, and do more internally. So, that's been really helpful because. Ultimately, it's helped to seed the market, too.

**Olds:** So, Suzy, artificial intelligence, AI, is all the rage right now. How can that work out for a private company that's looking to get their feet wet?

**Tichenor:** Well, this is a very, very exciting time in high performance computing with the new systems that are coming out. I can tell you that at Oak Ridge the Summit system that we have is really uniquely architected to do not only the traditional modelling and simulation that you think about - car crashes and airplane design - but also very large-scale data analytics, deep learning and machine learning. Summit is just wonderful for that. So we see companies now applying for time with us not to do the traditional mod/sim, but to really start to learn about and perfect their Al models and how to do that at scale. And we have, interestingly, a lot of small companies and startups in that area that are very capable of using a very large system like this and who would benefit from access.

You know, one thing we learned, Dan, from our industry program, is that large-scale computing is not just the purview of large companies. A lot of large companies are at the beginning of their HPC journey

and there are a lot of small and nimble startups who are very capable of using large systems, but they just don't have access to them. They just can't afford them. So, we're agnostic as to the size of the company that wants to come and use these resources. We're always looking at what's the science and what's their ability to use these systems. Now, especially in the AI and machine learning area, we're seeing there are a lot of small companies out there that could very much benefit from coming to a center like OLCF and use a system like Summit, and I think that's going to be a big trend in all the DoE supercomputing centers: seeing a lot more machine learning and artificial intelligence and data analytics being done on these big systems and then being done, also, in combination with data from physical experiments. So, it's really a very exciting and broadening time for the use of high-performance computing with the use of these new architectures.

**Olds:** It's interesting. I did an HPC road trip last year. I drove from Portland, Oregon, down to Dallas, Texas for SC. But I stopped along the way at national labs on the way down and on the way back up and did interviews, and almost all of them said that they were going to be, maybe heavily, using AI to inform their simulations. So, definitely on the cutting edge, and that was nearly two years ago.

**Tichenor:** And now we have systems that can do that at scale. Because when you think about the simulation data it's enormous. So, you need enormous systems like a Summit, like the exascale systems that are coming in order to make sense of all of that in a reasonable amount of time.

Olds: That's fantastic - giving them a window into AI.

Tichenor: It's really a crystal ball look.

Olds: Yes, absolutely. And who couldn't use that now.

Tichenor: Ah, I don't think anyone would turn that down if they had the chance.

**Olds:** Now, we've talked about what has you excited and it's obvious you were excited about this and you still are.

**Tichenor:** Working with these companies is awesome. Their problems are exciting, the people are excited to be working on them, and they're real-world problems, they're now problems. Whereas a lot of the research is fundamental research that goes on at user facilities, so you might not see the practical application of some of that for years down the road. Whereas companies, of course, have a different business model. Their research isn't just for curiosity - it is very much aligned with a business objective.

Olds: It's going to have an impact right away in a lot of cases?

Tichenor: It needs to or they can't justify it.

Olds: Yes. Now, is there anything that has you concerned looking forward and looking down the road?

**Tichenor:** That's a really good question. I tend to take a longer view and say that these things work themselves out over time. But in the near term, we don't see companies jumping on to GPUs as quickly as DoE has.

Olds: As they should?

**Tichenor:** You know, that's a business question. There is a cost to making that move. There is a cost if you have your own internal software, there is a cost to porting it. And, then, a lot of companies are still dependent on ISVs. ISVs are going to port their software as there is demand. There tends to be more demand at the higher end, which is the peak of the industrial pyramid, which is smaller than the base which may not be demanding it quite yet.

**Olds:** That'll also lead to higher ISV costs if they're doing it for a small number. They'll charge whatever they can get.

**Tichenor:** There's that. What will the market bear? There's a lot of that. For companies that are still writing their own internal software, they're moving along a little more quickly and that's gratifying, and we're starting to see that now. It's important for companies that have been using the DoE centers as a springboard to an early attack on large-scale problems that they can't solve on their internal systems, it's really important for them, now, to seriously think about GPUs. Because you've seen the announcement of the exascale systems. They're all going to be heterogenous architectures. So, if a company has been using DoE systems for their larger-scale problems and they haven't been able to port their software yet to GPUs or find a substitute open code, they're not going to be able use those DoE systems until they do. In most cases the project just won't be accepted. It won't be approved, because the horse-power is coming from the GPU.

**Olds:** About 90% or more of the horse-power is coming from the GPUs. So, they need to be on the train.

**Tichenor:** Now, one thing that the DoE labs have done that has been very helpful is they started, a few years ago, hackathons, where organizations can bring their code to a week-long intensive workshop. It's all free, you just have to pay for yourself to get there - and then we have experts from the labs and the vendors there to work with you to get that code ported over and get it started. These have been very, very successful and we have found ways to work with proprietary code, too. Companies can bring mini-apps and it's been successful. We've had some companies that have come in with their own codes and that's worked well, and that's given them a jump. So, there are tools out there and training and workshops that can help. But as far a concern, I guess my concern would be that more companies see the importance of moving of GPUs if they want to scale up because that's where the big systems are.

**Olds:** Yes, and that's a little bit of a surprise to me because I know that, as an industry analyst, I've been talking about and promoting the use of GPUs since 2008.

**Tichenor:** I'm sympathetic to the companies. Even when Titan was installed at Oak Ridge a lot of people thought Oak Ridge had sort of wandered a bit by putting in a hybrid architecture. It took a couple of years before people realized, "Wow, there was a real benefit to that." Now you see all the exascale systems that will be with GPUs. So, I can understand companies sort of holding back a little bit to say, "before we take the plunge, an expensive plunge, let's see if it's really going to pay off." That's one of the important roles that DoE plays. It's the early adopter. It's where you find serial no. 1 in the hardware. DoE helps to push and prove out the technology, so I can see why companies have waited to see, "Well, let's see what happens. Is DoE really going to do this and stick with it?" I think the answer is a resounding yes. And, so, it's really time, if you haven't don't it already, that you need to do that or you're going to be locked out of certain things.

**Olds:** One thing, and I imagine these tools are out there but I'm not sure specifically what they are, but do you have something that you can give to a company that says that these applications will have x-type of acceleration using accelerators, like GPUs?

**Tichenor:** Well, we don't have a fact-sheet that has that but, certainly, there is information on websites and you can find out. I mean, if they have a code that they're interested in, an open code, we can certainly tell them whether it has been ported over to GPUs yet. They can send us an email and we can tell them that. If they're working with an open scientific code already, they can just go to the community themselves and find that out. And DoE has made a lot of investments in porting a lot of community codes that are important to DoE at least, and that are used by other researchers that are porting them to GPUs. So, some of that work's been done. Not every company uses open codes. A lot of them use commercial codes because there's a lot more support there. A lot of them still have kept the secret sauce internally and use their own.

**Olds:** Well, this has been great Suzy. Thank you so much for your time. It's been a real education for me and I'm sure our audience is going to listen with rapt attention.

**Tichenor:** Well, I don't know about that, but let me just close by saying if there is a company listening and you have some large problems that really exceed what you can do internally don't hesitate to get in touch with us at Oak Ridge, or one of the other DoE labs that has an industry partnership program, and see if there might be an opportunity for you to bring that problem to a DoE center and get an advance chance to start working on it before you run it internally. It could be a great opportunity.

**Olds:** That's great advice and I hope everybody out there follows up on it. I'll see if I can come up with a problem here and send it to you. Thank you again Suzy.

Tichenor: Alrighty, thank you very much.

## 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 <u>www.HyperionResearch.com</u> of <u>www.hpcuserforum.com</u> 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.