Episode 19: On a Mission of Disaster Management and Scientific Discoveries
Host: Nicole Huesman, Intel
Guests: Paul Navratil, Texas Advanced Computing Center; Jim Jeffers, Intel

---

Nicole Huesman: Welcome to *Code Together*, an interview series exploring the possibilities of cross-architecture development with those who live it. I'm your host, Nicole Huesman.
High-fidelity visualization is helping spur advancements across a broad spectrum—from our favorite movies to breakthroughs in medical research, geophysical exploration, cosmology, and much more.

Today's guests live at this forefront, and I'm very excited to welcome them.

Paul Navrátil, director of visualization at the Texas Advanced Computing Center at the University of Texas at Austin, is an expert in high-performance visualization technologies, accelerator-based computing and advanced rendering techniques. His recent work includes algorithms for large-scale, distributed-memory ray tracing, which enables photorealistic rendering of the largest data sets produced on supercomputers today. Paul, thanks so much for joining the program.

Paul Navrátil: Thanks Nicole. I'm thrilled to be here.

Nicole Huesman: Jim Jeffers, senior principal engineer and senior director of Intel's Advanced Rendering and Visualization team, leads the design and development of the open source rendering library family known as the Intel oneAPI Rendering Toolkit. This toolkit includes the Embree Ray Tracing library, which was just recognized with a Technical Achievement Award by the Academy of Motion Picture Arts and Sciences. Jim, so great to have you with us today.

Jim Jeffers: Thank you, Nicole. Happy to be here.

Nicole Huesman: So, Paul, given that Frontera, the world's largest academic supercomputer is within your purview, and given the state of our world these days, I'll bet there's no shortage of projects that are keeping you busy. I'll let the two of you dive in.

Paul Navrátil: Thanks Nicole.

Jim Jeffers: Hey Paul, first off, we're recording this just a little after a little bit of ice and other stuff happening down in Austin. How you doing, bud?

Paul Navrátil: Doing fine, Jim. Thanks for asking. It was five degrees this time last week and it's 75 degrees today. So, you know, you can't beat Texas for a variety.

Jim Jeffers: That's great. Well, I was actually just being nice! I'm really more interested in how Frontera is doing and how that was impacted. So, what's the situation with Stampede II, and Frontera? I know that occasionally you get involved in disaster management and computing. Was there anything done with that here?

Paul Navrátil: Yeah. A great question, Jim. I'm happy to say that all systems are back up and running at full capacity and, actually, the way we helped out in this particular crisis, because there was a severe energy shortage across Texas' grid, which is unique in the U.S. for having its own grid. There's basically Eastern U.S., Western U.S. and Texas. And so we voluntarily powered down our system. No jobs were harmed in this process. Basically, we paused the queues whenever a job finished. We then did not allow a new job to schedule on those. And we powered off all the compute nodes as they freed up. And so we kept the log-ins on. People were still able to execute. I think we still kept the development queues open just so folks could get some science done, but the lion's share of our consumption was taken offline, which returned several megawatts of energy back to the grid. And then as normal operations resumed, we powered everything back up. And I think we were running back at full operations sometime on Friday of that week. There's a considerable backlog in the queues now that we're still working through, but we're happy to say we're able to serve the state and still serve our science mission both to our best capacity during the disaster.

In general, we're very involved in events like this, particularly where supercomputing capabilities can come to the fore. One of the most regular ones, particularly in Texas, is hurricane forecasting. And so we have a partnership with NOAA to run the hurricane models on our systems and they do that on a number of other systems as well. But one of the primary centers is in Louisiana, which is also sometimes in a hurricane's path. And so we provide them general capability, but also we'll expedite jobs when there's a storm eminent. When hurricane Harvey parked over Houston and caused so much flooding, we were actually part of the mobilization of the state of Texas to respond to that crisis and the governor's office mobilized the university's capabilities and the university mobilized us, which we're all too happy to do to run predictive flood modeling, to help understand what areas would be most impacted and what remedies might be possible. And that's of course, in partnership with some of the great researchers at the university that have those computational models available.

And then just as a final note, particularly with Frontera, we’re part of the national HPC mobilization effort to fight COVID-19. And so we've been making a number of resources available primarily on Frontera to the national Open Science community, running both viral models like Rommie Amaro out in San Diego on studying the spike protein and other interventions for the virus itself, and then also, epidemiological models like Lauren Meyers here at UT and her team to understand potential spread and transmission scenarios, and those have been a really vital part of our response effort going on a year now.

Jim Jeffers: Wow, that's awesome. And I like that juxtaposition of, in this case, when TACC and Frontera were in the midst of, you came up with a silver lining solution to shut the thing off because of the power situation in your local area. So I'm very, very impressed with the way you guys manage the systems instantaneously to take care of us. So that's fantastic. I just want to delve a little further related to the area of viz and everything in this scope and ask you to relay the questions, you know, the disaster stuff and the weather modeling and even COVID. So, two-part question is, what is the role of visualization, specifically? You know, what are the top things that vis analysis itself provides, in addition to just the compute, like you said, the predictive flood modeling. And then the additional question is, since we both are leaning forward in ray tracing, in these circumstances or in general circumstances in HPC, how is ray tracing being utilized to do better visualization than maybe traditional, you know, rasterized OpenGL. What is your view of those two things, you know, with the specific domain of disaster analysis?

Paul Navrátil: Yeah, absolutely. So, for the first part, visualization and visual analytics play a critical role in the entire discovery lifecycle for scientists. There are three phases that I would classify that into. First is the individual or the research team discovery process, where they're just trying to see the data initially and understand what's coming off of this simulation or out of their data. And that plays an important role in terms of directing the next investigative question for the team. The second part is the communication among peers. And so for instance, in Dr. Meyer's group for COVID fighting, there's her research team here in Austin, but then there's a broader network of folks across the state and really internationally that they're engaging with. And so they will have a discussion of ideas about the latest results and having visualizations and analytics of their various models and simulations are key in terms of advancing those discussions. It's very true that a picture communicates a thousand words and it can communicate a thousand simulation runs. We're visual creatures, and so, [visualizations] just help advance that understanding that much more. And then really importantly, in the COVID discussion in particular, but for any of these types of disaster interventions is, communicating to the decision makers and to the broader public. Decision makers need good data in order to make good actions and policy decisions. And the general public needs good data to be informed about the situation and having a visualization of that that is easy to consume and straightforwardly shows the data and the insight contained in the data is key. And so having good means to perform each of those phases because sometimes the visualizations might be different for each of those phases, either more polished or more focused in particular ways to communicate the particular insight in question, the scientists operate best if they have a single tool and a single environment in which to perform those analyses so that they can refine them in a familiar workbench and a familiar sandbox, if you will, and they don't have to stop their personal analysis and then jump to another completely different environment that they may not be familiar with as much to do some polished final results to send to a government agency or to a funding agency or to the press, the general public. And so, ray tracing really helps in a number of ways in that avenue. I know that we've been involved through the Intel Graphics and Visualization Institutes (GVIs) of Excellence in incorporating ray tracing capabilities into frontline scientific discovery tools like ParaView, models and engines like Embree that are used by film companies to create computer generated imagery like from the Avengers movies, for instance, or like from other Pixar-type films. And so the ability to have that level of rendering capability in those tools allows a scientist to have a single environment to do their initial discovery and then refine and improve that imagery for final publication, final dissemination, without having to change the entire way that the data is read in and represented. And that's a really powerful time-saving device for them and just enhances the initial quality of the imagery and analysis that they can do, which can lead to better science. A great recent example is that ParaView plus Intel OSPRay was featured on the January '21 cover of Science magazine. There's a research team from Iowa State that one of my team members here at TACC was working with and analyzed their data within ParaView using the OSPRay rendering module, and that went from their initial discussions around the visualization to a magazine-ready cover all within ParaView, which is amazingly powerful for both us as visualization people and then also for the domain scientists themselves.

Jim Jeffers: Yeah. That's fantastic. Gonna switch gears a little and just get your thoughts on a combination of the overall oneAPI initiative led by Intel, particularly the open-standards-based part of it, and your thoughts, your feelings about what the benefits that that program can bring to HPC visualization.

Paul Navrátil: Absolutely. So oneAPI is a really powerful application metaphor and implementation metaphor that would allow for folks to have a unified code base for their entire scientific workflow. And one of the additional things that we've been working on is in-situ analysis. In other words, analyzing data as it is being generated by simulation without writing it to the disc in the meantime. And there's a number of ways to do that, but oftentimes it requires some level of integration with the simulation application. And so if you have applications and analysis tools that are all able to be run wherever they need to be run, and that's typically wherever the simulation is being run, that's a really powerful expression to enable scientists to run where they have the cycles and to not have to adjust their workflows or to minimize tweaks to their workflows based on which system they're on. It also eases the life for developers so that they can continue to advance great new features of simulations and of analysis capabilities without having to then port those existing capabilities to a plethora of different architecture types in order to support the various architectures that they may be asked to run on.

Jim Jeffers: Yeah. In situ, one, is a great term because it's totally, what, six letters, but it's so powerful. I think the plethora of outcomes and what you would call the virtuous cycle that comes through it, I think a lot of people don't fully grasp. So you, you know, a second ago, were talking about data movement and we've talked about this a lot, but you know, petabytes, exabytes, as we're heading into exascale, it just doesn't run as fast as compute. It's just the way that physics are. And so moving data around, as soon as you have to copy the data somewhere, you almost have hit a huge wall. And so the industry, us at Intel, Paul, the Department of Energy, we're doing a lot of work in this arena, in the Exascale Computing Project, as well as work of our Intel GVIs with Paul and others to enable this, you know. So, that's one thing to, you know, basically better utilize the system that you have and the memory and all the key expensive stuff on the system it enables and improvement of that flow. But then my favorite term, I'd love for you to confirm with me that I'm kind of thinking of this correctly is, is what I call real-time science, which is you're literally doing the traditional five years ago, or even many codes now are run, run, run, run, run, run, run for, you know, hours, days, weeks, get all your data together, store it on a disc, get back on the computer again, do another job to do the visualization, which you even just outlined before is often three or four different x styles of viz, and what in-situ provides is just all the tools for you to do it as you compute, then, essentially at the end of your compute run, you know, you might linger a little longer and hang on to the data, but do your vis and all those preliminary things, and this is happening in close to real time. So, it is to me, a true virtuous cycle that I just think is going to be critical. You know, you could potentially call it the next big thing or the current big thing. So, what do you think about that?

Paul Navrátil: Absolutely, Jim. The push to move from reactive science to proactive science, I think particularly in the realm of disaster relief, it's an opportunity that requires the full workflow pathway to be as optimal as possible. And it's been a great partnership with Intel in terms of advancing the variety of capabilities for that. You know, ray tracing plays a key role in terms of both high-fidelity rendering, but also rendering efficiency. There are a number of ways that ray tracing can be an accelerant, if you will, in terms of overall processing, in terms of the memory that might be required to hold the visualization that's being produced. And then another initiative that we've been working on is utilizing these hardware-optimized ray tracing components to do the simulation itself. So, it turns out that there's whole classes of simulations that do things that look like ray tracing, at least from a visualization or a ray tracing expert's perspective, vector mathematics, spatial sorting of data in which ray tracing has had a decades of research in how to do that efficiently in a variety of ways. And so having those capabilities enable simulations to potentially utilize these new hardware pathways or these new software libraries like Embree, like Open VKL, like OSPRay, to effectively use the vector units in Intel CPU to improve the vector math or the spatial sorting, or what have you, or in the new X^e processors, utilizing the ray tracing cores to optimize what the ray tracing core considers ray traversal, but it could be particle advection, it could be heat and radiative transfer, it again, could be […] search. And that's another exciting opportunity that's really been unlocked by these capabilities.

Jim Jeffers: Yeah. I want to extract two things you said there that I think are everyday for you and me, but the audience might not realize. One was what you were calling ray tracing efficiency. I think there is a reasonable belief that, 'Oh my God, ray tracing is used in the movies, and a lot of people who, at least, get the slightest into 3D digital movies and how that works go, 'Oh my God, it's, you know, 24 hours per frame. It's not 60 frames per second. It's 24 hours per frame.' That is an option of ray tracing when you really want to dial it up. But I think a lot of people are unaware. There's a cost of entry, right? Ray tracing starts with a cost of entry that's generally higher than the cost of entry for rasterization. But once you pay that cost for entry, essentially everything is almost free after that. And we've been bearing this out in some of the benchmarking and stuff my team has done. By the way, the ray tracing community has actually presented this for years and years, for rasterization, just the way it's structured, it's not bad or good, it's just the way it's structured is, the more stuff you throw at it, the more you have to compute. If you have a fixed cost of flops and compute capability, if you want higher complexity, you have to throw more triangles or objects at it, and the frame rate begins to drop. It turns out that ray tracing is not quite the opposite, but is a flat curve. Once you cross over the point and a certain complexity level, you can throw all the complexity you want and the flops remain about the same. Ray tracing has this entry cost, but once you pay it, you know, it's free. It's like paying for your cable TV once in the beginning, you know, maybe you pay a little bit more, but they never come back and bill you again, which will never happen in cable. That's one thing that I wanted to reiterate.

And then the other thing, which is really, really cool, and it's related to something I'll try to enhance here is the notion that ray tracing is actually a physics algorithm. It's really physics, it's wave particle physics.

Paul Navrátil: Yep.

Jim Jeffers: And also line-of-sight physics. It kind of fits in there. And so it can be utilized to do physics. You're basically utilizing tools that intercept objects or particles or other elements, and it can help track those things or identify where they are. So again, ray tracing has this physics of light or wave particle physics element that can map very often to HPC compute codes.

I'll just use another realm where people might be shocked. Although I think occasionally Paul has … you usually don't play actual games like Doom or something on a supercomputer …

Paul Navrátil: Purely system testing, purely system testing!

Jim Jeffers: But an interesting use is Valve's Steam gaming engine. It uses Embree, but right now, not for rendering at all. It uses it for surround sound 3D audio.

Paul Navrátil: Cool.

Jim Jeffers: And so, you know, it's computing positional data, which is what ray tracing is great at. So just then map that to things like particle advection, heat transfer and everything you were talking about, and you see what the potential of this algorithm is across HPC. Do you want to just quickly talk about the SOLAR (Ray Tracing) Consortium that's related to this area?

Paul Navrátil: Yeah, absolutely. So, the SOLAR Consortium is dealing exactly with this, Jim, how to use these ray tracing components from rendering in simulation, like the Steam game engine that you just mentioned, for auditory positioning and to do 3D sound, which is outstanding. And there's a company, SURVICE Engineering, out of the DC area, and Christiaan Gribble has been a great partner in leading that effort for other types of radio frequency computation, and for multi-hit ballistics, a variety of ray-like components in simulation modules.

Jim Jeffers: That's like microwave and cell towers and things like that, which also have line-of-sight kind of mechanisms, right?

Paul Navrátil: Yep, exactly. And so, the SOLAR Consortium grew out of the IXPUG Institute hackathon. The Intel Extreme Performance User Group has been a great sponsor since 2017 to really get the simulation folks and the visualization folks in the same room to not just talk about in-situ, but to actually advance it and develop on the simulation codes because it really takes that expertise of both simulation and visualization folks in order to preserve the science and optimize the analysis. And so, that has been a successful series.

And then we realized that there's this additional opportunity to utilize the ray tracing engines for science simulation. And so, we created a second group to bring in more folks to that community. We've had a successful birds of a feather (session) at SC19, and then we had another exhibitors forum during the virtual SC20. And also, from this effort, we engaged—we, being the community—engaged with the Khronos Group to create, first, an exploratory group, and now, a working group, around analytic rendering. And so, combining both ray tracing and rasterization, but as an effort to provide a single API for analytic codes so that folks doing this type of development, both purely at the analysis level, but then also incorporating them into simulations, can write once, run virtually anywhere, kind of in the spirit of the oneAPI work as well.

I'm excited to say that we have a special issue of IEEE Computing in Science and Engineering. The call for papers has just come out and they'll be due in September of 2021 submissions. For folks doing work in this area, we'd love to have your submissions for that. And just this week, we're also putting in a workshop proposal for SC2021, which is the workshop on analytic rendering and render-accelerated simulation, or WARRAS.

Jim Jeffers: Oh, okay, that's a new one for me! Thank you! I love making acronyms, by the way, so that's great!

Paul Navrátil: You know, I understand every good workshop needs an acronym, so we're hopeful that this will get accepted.

Jim Jeffers: You gotta make it WALRUS, man!

Paul Navrátil: Yeah, exactly, exactly! We'll work eggman in there as well!

But this would be a workshop co-hosted at SC, or combined with SC, that really combines both the ANARI community and the SOLAR community to really enable a venue for further discussion, publishing of early work, and to really share best practices, and just really allow for the community to grow and expand. We've had great uptake, you know, on the order of 40 to 50 participants at the IXPUG Institute hackathons and at the SOLAR events, and so I'm really excited that this might take things to the next level. We won't know at the time this pod gets posted whether or not that's actually happening, but if you go to WARRAS.SOLARRT.org, you can find a little more information about the workshop and hopefully you'll see soon that it'll be open for submissions.

Jim Jeffers: Yes. In one sense, it seems like this SOLAR and ANARI, you know, people are listening and thinking, this is out there in the future, like three years or whatever. No, this stuff can be done, like today, you know, you're doing active work. This is just kind of pie in the sky. We're talking about it. These consortium talks have the world's leading compute and visualization specialists coming together to bring these worlds to you out there. So, I really do encourage it.

Paul Navrátil: Yeah. I mean, the whole reason we pursued a special issue of Science is that from the exhibitors forum at SC20, the community said, 'Hey, we're doing really good work in this area that we want to share with people, where can we publish stuff?' And so, we wanted to really have a targeted venue for this level, the Science special issue we're thrilled about. And then hopefully this workshop will also be a more continuous and perennial opportunity because of its focused nature. I think there are other conferences that may be receptive to these types of papers, but I think that there's enough interest that we can have a focused discussion and really get some exciting early work put out there.

Jim Jeffers: I just want to chime in and just a shout-out to Jim Ahrens of the DoE for being also a partner with us. Paul usually has more of an NSF academic hat on, and Jim has the DoE hat and Jim leads the ECP programs for data analysis. So, they are two powerhouses in it. So just again, shout out to Jim, you know, check with Jim and Paul on some of these things.

Paul Navrátil: Great point, Jim. Jim Ahrens at Los Alamos has been a great proponent for this and really an engine encouraging the community to advance to these levels. And so I really appreciate you bringing it up. We wouldn't be here without Jim's support, so it's very appropriate to recognize him.

Nicole Huesman: I knew when we asked the two of you to come on the program, we would have a fantastic discussion. So I would love to invite you both back on the program in the future, because there's so much more to talk about. With that, unfortunately, that's about all the time that we have today. As we wrap up, any other places you'd like to point folks to learn more?

Paul Navrátil: Thanks, Nicole. It's been a great discussion. I always enjoy having a spirited talk with Jim. We'd love for folks, if you're interested in the SOLAR community, just SOLARRT.org. And of course, TACC, www.tacc.utexas.edu/. You can find out about all the interesting and important research that we're really fortunate to help support in terms of the Open Science community and read about Stampede II and Frontera and these great systems that are being powered by Intel and powered by the Intel oneAPI rendering framework that gives us the analysis capability on those machines.

Nicole Huesman: And Jim, how about you? Any pointers you'd like to give folks before we close today?

Jim Jeffers: Yeah, so, I think two things. I'm a random tweeter and I also work with our social media folks at Intel. So, my Twitter handle is at @jamesljeffers, even though everybody calls me Jim. And then, you can also learn a lot about what we're doing using @IntelHPC @IntelDevTools and @IntelGraphics, follow all of those guys and you'll get a bunch of cool stuff. And then, finally, to learn, and really we have a great landing page that I'm really pleased with, lots of things for you to peruse, you certainly can Google Intel oneAPI Rendering Toolkit, that'll probably take you there, but if you want to just type it quickly, it's http://bit.ly/renderkit, and that'll take you to a landing page that tells you all about the rendering kit tools under oneAPI and our direction and all that kind of stuff. So yeah, I think that would be helpful.

Nicole Huesman: Great, thank you. Paul, it's been so exciting to hear all about the work happening at TACC. Thanks so much for being here today.

Paul Navrátil: My pleasure, Nicole, and we really appreciate the invitation.

Nicole Huesman: Absolutely. And Jim, I have always loved our discussions. Again, thanks so much for your insights.

Jim Jeffers: Love to do it. And I can't wait 'til we all actually see each other live. Maybe, maybe at SC, we'll see.

Nicole Huesman: Yeah. Oh gosh. Here's hoping, right? Yeah. And a big thank you to all of our listeners for joining us. Let's continue the conversation at oneapi.com.