

THIS WEEK IN HPC: IBM Edge Conference in Las Vegas

Addison Snell

Michael Feldman

Podcast excerpt

May 2014

PODCAST EXCERPT

The following is an excerpt from the weekly Intersect360 Research podcast, "This Week in HPC," available on iTunes, Stitcher, and through our media partnership with insideHPC. The full podcast can be found at <http://www.intersect360.com/industry/podcasts.php> and is hosted at <http://insidehpc.com/2014/05/23/week-hpc-ibm-edge-conference-new-federated-cloud-europe/>.

In this podcast segment, recorded on May 22, 2014, host analysts Addison Snell and Michael Feldman discuss the IBM Edge Conference.

Addison Snell: I've been having a good time in Las Vegas this week with our friends at IBM.... This is an evolution of the original IBM StorEDGE Conference — StorEDGE, get it? It used to be a little more storage-focused, but now really gets into not only executive tracks in Executive Edge, which is really CxO-designed materials, but then goes deep-dive into a lot of technical tracks, from IT storage to networking, Flash, the different IBM technologies like System x. I was there in particular for a lot of the Big Data and analytics track stuff in their Technical Edge areas. It really goes enterprise-wide.

Michael Feldman: Right, it looks like there were quite a few announcements, like you said ...sort of focused on storage, but also now of course on Big Data and the analytics community that's surrounding that. It looks like a pretty interesting selection of sessions and talks there.

AS: Very interesting. I always like seeing the end-user perspective on things. I got to attend a presentation by Jason Bucholtz of DigitalGlobe — always a great talker — looking at how they can use geospatial imaging in a lot of different application areas and the amount of data they're bringing down by satellite. It's really a certain perspective on that aspect of Big Data. It really is these very huge files. [I] went from that right into how to build a multi-tenant Big Data infrastructure using Hadoop by Robert Ghavidel of USAA, looking at their fraud detection applications and how that went over into Hadoop. [He was] very clear on the types of technical computing or High Performance Computing that goes on at USAA and how that overlaps with Big Data. Good presentations all around.

MF: One of the most interesting ones that I was reading about... I guess it was a presentation on analytics applications that was predicting — or actually suggesting — how to build the NFL TV schedule, which seems pretty evident to me but there's a lot that goes into it. It requires quite a bit of decision-making and compute power these days.

AS: Really interesting presentation — and probably the one that people were talking about the most at IBM Edge — this presentation by Mike North of the NFL. He was saying as recently as five years ago the NFL schedule was done by hand. You'd get the big board and you'd pick out of the hundreds of games that go on over the course of just the regular season.... You're saying well, all right, these are the fifty highest-value games and we want to do them, and then you'd try to get the schedule to work the rest around it.

You're dealing not only with the stadium and scheduling constraints — and you don't want one team to be on the road eight weeks in a row, or mundane features like that — but they're dealing with, what? five different network contracts. (They're) balancing out which games are on CBS or FOX or NBC or ESPN or the NFL Network, and trying to satisfy all of the different viewer demands so that when you've got a high-value game, like Peyton Manning's going to be playing against Tom Brady when the Broncos face the Patriots.... All the networks want that game and only one of them's going to get it.

MF: Right! So I guess the impetus behind this is sort of to maximize the exposure of these games to a certain extent, but also to balance that with all these other variables. Like you said, not having teams play on the road for a certain number of weeks in a row, and it's probably tough to schedule three-day weeks after a Monday night game, things like that. There's a lot of decision-making that has to go on, and it's not arbitrary. They really do want to maximize the exposure into money for themselves and the networks in the big-name games but also be fair to the smaller teams as well.

“As recently as five years ago the NFL schedule was done by hand. They're getting more hardware involved every year, moving into a midrange HPC cluster to plan this out.”

—Addison Snell

AS: That's right. Between the teams as stakeholders and the network stakeholders... You know, the teams care, “Am I going to be playing in Green Bay or in Miami when it's January?” The timing of when you go to Green Bay can matter quite a great deal in the outcome of that game. It's hard to make everybody perfectly happy. What they're really trying to do is balance the value of all of those games. It was interesting that Mike North mentioned in his presentation and in subsequent interviews — you can find these interviews, by the way, by searching IBM Edge and something called The Cube on their YouTube channels; you can see an interview with Mike North on this — using software from a Canadian company called Optimal Planning Solutions, that really helps do a lot of the scheduling. He said they're getting more and more hardware involved in this every year, going to multi-core and then parallel solutions with 10, 20, 40 different boxes. So, really, (they're) moving up into what we would consider a midrange cluster for High Performance Computing to plan this out.

MF: That's very interesting. I would not have guessed that it was an HPC-type application behind the scheduling of this. It just goes to show you it's getting into a lot of different application areas that heretofore you would not have suspected.

AS: Yeah, it impressed me, the level of scale that it's getting to as well. Now I did note that Mike North did say that IBM lends them the equipment for this. He never said that they buy it, and I wonder if his testimony at IBM Edge is part of the deal.

MF: Yeah, there you go.

AS: That wasn't the only interesting announcement coming out of IBM Edge, though. On the technology side, another interesting technology I heard a fair amount of, was tape. And that was interesting because last year at IBM Edge, everything was all flash, all the time. And here we certainly continued to hear about flash, but tape was making a splash at IBM Edge as well.

MF: Right. I caught the announcement from IBM about this, basically from IBM Research; it's still in the lab, but they're collaborating with Fuji Film and building a tape system—

AS: — a tape cartridge.

MF: Yeah, a tape cartridge — around that technology. It's very dense now. This new tape technology that Fuji Film has come out with, it's like 62 times as dense as the current library tape cartridges out there now. You can stick something like 154 terabytes — that's *terabytes* — of data on a single cartridge now.

AS: That's very impressive. And, first of all, it bears mentioning that there have been a lot of important developments in tape libraries and tape technologies over the course of the last five years, starting with LTFS and the ability to do a directory of what's on your tape. That leads you into active archives and all of the other developments that we've heard around tape. Now this idea coming out of IBM Research and Fuji Film, I think it essentially takes a lot of the mechanisms or procedures that you've done in disk drive manufacture and bring those same concepts into tape manufacture. They've found that they're able to dramatically increase the density of storage on tape. That's going to help, because obviously if it's on tape you're talking about pretty big volumes of data, and if you can reduce the amount of distance that your robot has to travel, then that can improve access times.

MF: Yeah, definitely. Well, it's still the technology of choice for cold data and archives and things like that, and though people will occasionally...they'll talk about, like tape is dead, I think —

AS: Bah, tape is not dead! Tape is growing, man.

MF: When you see improvements like this it seems like the lifetime is into the ever-foreseeable future.

AS: Tape is, and is going to continue to be, an active part of enterprise infrastructures, HPC infrastructures, Big Data infrastructures, and is certainly not going away, and in fact is likely growing in its footprint in these kinds of areas. I think we'll continue to see that.

MF: Yeah, I think you're absolutely right.

“With this new tape technology, you can stick 154 terabytes — that’s terabytes — of data on a single cartridge now.”

—Michael Feldman