



BBBT Podcast Transcript



About the BBT

The Boulder Business Intelligence Brain Trust, or BBT, was founded in 2006 by Claudia Imhoff. Its mission is to leverage business intelligence for industry vendors, for its members, who are independent analysts and experts, and for its subscribers, who are practitioners. To accomplish this mission, the BBT provides a variety of services, centered around vendor presentations.

For more, see: www.bbbt.us.

Vendor:	Action - ParAccel
Date recorded:	July 26, 2013
Host:	Claudia Imhoff , President, BBT
Guest(s):	John Santaferraro , VP, Marketing, Paracel Platform Group
Run time:	00:12:52
Audio link:	<u>Action - ParAccel Podcast</u>
Transcript:	[See next page]
Transcribed By:	CastingWords



Claudia Imhoff: Welcome to this special edition of the Boulder BI Brain Trust podcast. I'm Claudia Imhoff, and I'm here in Grants Pass, Oregon, at the Humphrey Strategic Communications annual Pacific Northwest BI Summit. I'm pleased to have the opportunity to interview each of the vendors attending this event.

CI: With me is John Santaferaro. He's the Vice President of Marketing for the ParAccel Platform Group of Actian. Now, that's a whole new title for you and it's a mouthful, so why don't you tell me a little bit about this new Actian acquisition here?

John Santaferaro: It's very exciting, Claudia. Actian has purchased ParAccel. ParAccel is now part of Actian, but the brand and name and product lives on. They've created what is called the ParAccel Platform Group. They've put into that all of their assets that they acquired from ParAccel, the whole analytic platform, all of the assets that were formerly VectorWise, and all of the assets from Pervasive that were known as DataRush and RushAnalytics.

It really expands the set of capabilities that we can take to market and offer customers around an analytic platform. Suddenly, you have an analytic platform that is able to run high-performance analytics on massive amounts of data.

You have 600 analytic functions embedded in the database. You have the concept of on-demand integration, where you can now connect to over 200 data sources. Added to that, this whole idea of data provisioning, where you can create these data flows, where you're grabbing data and applying operators to transform the data, to enrich it, and then to actually do the analytics.

Those analytics can be done either parallel-ized on our data-flow engine servers or actually run right on the HDFS cluster -- so everything from high-performance, complex, sophisticated analytics to data provisioning and creation of data flows to analytics running right on the Hadoop cluster.



CI: Very exciting, because it has so expanded the breadth of ParAccel. You can do things now that you couldn't do before.

JS: Exactly. I always look at it from the standpoint of the analyst. When an analyst comes to his computer, to his workbench, what are the things that that person needs to do to do their job? Well, guess what? They've got to connect to data, they've got to prepare and enrich that data to ready it for the analytics, and then they've got to run the analytics against it.

We're now providing everything that that analyst needs to do their job. What we're really doing is providing unconstrained analytics for that analyst, the idea that he can get any data, run any analytics, and expect a response at high speed.

CI: Action's done a very good job of picking the right pieces, acquiring the right pieces, and then grouping them and putting them together in a very smart fashion.

Now, let me change the topic on you just a little bit, because you've got an article coming out. It's going to be out in the Data Warehousing Institute's "BI Journal," and it's all about a new breed of analyst that's emerging.

We talked this morning here all about big data and who's using it. The whole emphasis was, for heaven's sake, let's stop focusing on the technology. Let's focus on what we do with this data, not so much the fact that we can capture it. Yes, that's always very useful, but if you don't do something with it, then it loses its value. That's where this new breed of analyst comes in. Why don't you tell me about that?

JS: I think everybody agrees that the data scientist has shown up on the scene, highly valued, highly prized, highly paid now...

CI: And highly technical.

JS: Yeah, but in very short...

CI: Supply.



JS: Right. There aren't enough of them, and there never will be enough of these PhDs, these quants, these mathematicians that build algorithms and know the data and write the programs to run the algorithms.

You could go and try and train people to be PhDs, but a better approach is to take the business analyst and the data analyst and the BI analyst and just grow them up into a bigdata analyst. You can train them on a couple areas. One, you teach them more about the data. Where is it? What kind of data is it? What do you do with different kinds of data? How do you prepare and enrich it to make sense out of it? How do you take a Web log, for example, and session-ize it so that it means something to the person who's doing the analytic?

You begin to teach them more about the data, and then you take them through Analytics 101. They don't have to be the person who creates the algorithm and a grand mathematician. All they have to do is know how to use the analytic function and apply it to the data. I can teach them what can be done with pattern match. I can teach them what can be done with a K-means cluster algorithm.

CI: They may not have to create the cluster algorithm, but they need to be able to interpret the results of the algorithm. I think that's what you're getting at is that, OK, we can now give them the algorithms, but they still need to interpret the results. They need to be able to understand the business well enough to say, "Here's what's going on in this scenario. Here's what this analytic is telling me, this analysis, and the results of that analysis are telling me, here's what's going on in my business." Would that be a correct way of putting that?

JS: Exactly. They become users of the analytics, not creators of them. I think it's the right community of people to go after to bring their capabilities up to the analytic level, because, when you think about it, most of them were Computer Science majors. Because of that, they took a healthy dose of math, and so they understand math. They understand statistics. They understand a lot of the kind of things that they're going to need to do with analytics. They get that already.



CI: Well, and they're already sort of data engineers as well. They have to understand the underlying data structure to a certain extent. They're certainly not DBAs, but they need to at least be able to engineer the data to the point where they can now run it through the algorithm.

JS: Exactly. I love the idea of the data engineer as well. The ParAccel platform supports all of this, right? Because you can pour massive amounts of data into this massively parallel system with N number of nodes, and because we embed these analytic functions in the database, the big-data analyst then basically picks the function, picks the data, and gets a response very quickly.

CI: Yeah. Put something together, and boom, out comes the result.

JS: The data scientist, meanwhile, can be creating these functions, and as he creates them, at a point when he's got the algorithm correct, he embeds it in the database and it becomes a part of that library. Then this army of big-data analysts can suddenly have access to that and have at it and start discovering things in the data by the hundreds, right, instead of just the one data scientist.

CI: It certainly does expand the audience. It's what we in business intelligence have been dreaming about is, you don't have to be so technologically savvy or so "PhD in statistics" and that sort of thing. A mortal human could actually create some very intelligent analytics and interpretations of the analyses -- and be able to then act on it, which I think is the most exciting part.

The data scientist may create the algorithms, may create the analyses, but they're not necessarily the people that are going to act on it. It's the business analyst, it's the person in the trenches that's going to take those results, interpret them according to the situation they find themselves in, and then act on it, do something about it.

JS: Exactly. I think there's a great opportunity as well for this new breed of analyst. They'll have the data and the analytics expertise, but there's also a huge opportunity for them to become the storytellers. If they understand the data, if they understand how to use the analytic, they



know the result and they can tell the story of how that applies to the business. They need to be quick thinkers, innovative thinkers, and storytellers, to be able to put all of these pieces together, to drive them through that iterative, analytic process, and then to capture the story and be able to go tell that story to the business or to the customer or to the supplier, whoever they're interacting with.

Cl: When you say "storytelling", why don't you explain a little bit about what you mean by that? Because what you're saying is that there's a context behind the content itself, right?

JS: Data always means something, based on its context. Application data is a great example. You carry around a smart phone. Every application in there accesses your GPS. They know where you're at. Every click that you make or move that you make in the application, they have a record of that. After a while, the context there is a session -- somebody picked up the phone, somebody opened that app, they spent three minutes on the app, they did these seven things. I can now categorize them. Then there's an event at the end, right? The event is either I buy something, or maybe it's an event that I want to prevent, that they get to a certain point and they get rid of the app. It didn't work for them. All of that is incredible knowledge to the people who are building that application and trying to sell them something or trying to get them to upgrade to a paid subscription.

The context is about the event that they want to create or avoid and then the pattern that comes right before it. What's the pattern of events? What kinds of things happen before that? Then what's the trigger? What's the trigger that gets them to either leave or to buy or whatever that event is? These new big-data analysts, that's the story that they can articulate to the business.

Cl: It used to be "every picture tells a story." Now it's "every data result tells a story."

JS: Exactly. I love that. I love that.



CI: That's what we're trying to get to, is that story. Every analytic result tells a story. If it's going up or down or sideways or staying the same, there's some reason behind that, some story that goes along with it to explain what it means.

JS: Exactly. My car produces a massive amount of data that is shipped back to the manufacturer. Every day, that manufacturer is learning more about how their engineering works, what works in different kinds of conditions, in different locations, at different slopes and elevations, and they're capturing all that telemetry data to be able to figure out how to build better cars and keep them out of the shop and not have to recall things and all that.

CI: Make you a happier customer. Excellent. All right. Very good, John, I really appreciate your time today. Again, it's John Santaferaro. He's the vice president of marketing for the ParAccel Platform Group, a newly formed group for Actian. Thanks so much, John.

JS: Thank you, Claudia.

CI: Thank you for listening to this special edition of the BBBT podcast, and thanks to Scott Humphrey for giving me this opportunity and for hosting the Pacific Northwest BI Summit.