

## **Data Warehouse Podcast Transcription**

Andrew Verdesca: Hello everyone. My name is Andrew Verdesca, and I'd like to welcome you to another IBM podcast. Joining us today to discuss what to consider when shopping for data warehouse appliances is Glen Sheffield, Technical Evangelist from the IBM Canada Lab. Glen has been with IBM for almost 30 years and currently helps customers understand IBM technology for data warehouse and how it compares to other products in the marketplace.

And so, without further ado, let's get started.

So Glen, welcome.

Glen Sheffield: Yes, hi Andrew, how are you today?

Andrew Verdesca: Very good. Thanks for joining us. Now, there's been a lot happening in the data warehouse marketplace recently with a definite trend toward what are being called data warehouse appliances. What's driving this?

Glen Sheffield: Well, what's happening is that customers are looking for simplicity and they want to get value from their purchase of the data warehouse right away. They don't want to spend, you know months installing and configuring a system from scratch. Customers realize that data warehouse systems need to be configured differently than traditional computer systems that run the business, for example for traditional transaction processing. Data warehouse systems need to have an abundance of I/O bandwidth, for example, to efficiently scan and analyze vast amounts of data. So pre-configured data warehouse appliances take the guesswork out of choosing and configuring the necessary system components, but in a very cost-effective way.

Andrew Verdesca: I see. Well, that certainly makes sense. So what does IBM have to offer customers in this regard?

Glen Sheffield: Well, IBM's been in the data warehouse marketplace for quite some time. But so, what we have to offer now is something called the Smart Analytics System. And we recently completed the acquisition of a company called Netezza, which offers what's called the Twin Fin data warehouse appliance. So both of these IBM offerings are designed specifically for high-speed analytics. Netezza actually entered the marketplace in 2003 and they've been a huge success with customers. In many ways, Netezza actually spearheaded the appliance trend with a tremendous focus on simplicity and low cost. Smart Analytics systems from IBM also have that focus, but they

## Data Warehouse Podcast Transcription

provided a more flexibility but offered solutions that are based on DB2 that are pre-configured on both the IBM Power Systems as well as z/OS architecture.

Andrew Verdesca: I see. So in other words, it sounds like IBM offers some options here, some flexibility for our end-users. What about other vendors? What other choices do customers have today?

Glen Sheffield: Well Andrew, you know, it's been a very interesting marketplace for the past few years. There's been a number of new entries. Some have come and gone. And there's been some consolidation.

And also, for example, SAP recently acquired Sybase IQ. Microsoft acquired a company called DATAlegro. EMC just acquired a company called Greenplumb this year. Terradata, a long-term player in this marketplace, acquired a company called Kickfire. HP, you know just a couple of years ago, introduced an offering called Neoview, but now, it looks like they're about to withdraw it. And Oracle has introduced something called Exadata. So there's certainly is definitely a lot of choices for customers.

Andrew Verdesca: Wow, there sure are. So, given that multitude of choices out there, what factors would you say customers should consider when evaluating the different product offerings out there?

Glen Sheffield: Well, there's traditional factors of course, that would include things like performance, price, price performance, scalability, ease of use...but another factor that customers should be concerned about, I think, is product maturity.

Customers are deploying these solutions now for bet-your-business, mission-critical applications. And relatively new products introduce risk. And I don't necessarily mean the viability of the vendor, but risk that the product itself hasn't matured enough for enterprise-class reliability and stability.

IBM Smart Analytics system is on the fifth generation of its data warehouse appliance based on DB2, with a 15-year heritage. And Netezza, our latest offering and acquisition I spoke about earlier, is on its fourth-generation data warehouse appliance and is proven across more than 350 customers with data volumes into the petabytes. So, in this regard, I do believe that IBM solutions reduce risk.

## Data Warehouse Podcast Transcription

Andrew Verdesca: That's a great point. Especially since these systems represent a significant investment for customers. One last question. We've been hearing a lot from Oracle regarding Exadata with bold claims about performance vs. the competition. Any comments on Exadata and where they fit in all this?

Glen Sheffield: Yes. Sure. First of all, the proven architecture for data warehouse solutions is what we call share nothing for MPP, which stands for Massively Parallel Processing. This is because, as systems scale to larger sizes to support the larger data volumes, the share nothing architecture ensures that that necessary bandwidth I/O that I mentioned earlier, also scales. In other words, each server that's added to this share nothing cluster has its own dedicated storage and bandwidth to enable this.

In the marketplace and all the vendors I spoke about, Oracle is the only vendor without this architecture and recognized that they were therefore limited in terms of storage bandwidth. So Oracle introduced Exadata to address this, and what Exadata does is it adds another layer of servers below the Oracle database to manage the I/O operations and provide that scalable bandwidth.

I think it's an interesting idea, but it's still not a true share nothing architecture such as that of Smart Analytics or Netezza. But more importantly, Exadata is relatively new. So in particular, the software that runs on the Exadata storage servers themselves was first introduced just two years ago on HP. And then one year ago, Oracle dropped HP and moved the Exadata to Sun.

So while some customer testimonials are starting to emerge, there's still a lack of widespread factual information about Exadata installations. So I think customers need to assess the risk. In particular the risk associated with product maturity of Exadata in the context that I mentioned earlier.

Andrew Verdesca: Yeah, that's a great point. I see what you mean. Well OK, with that, I want to thank you Glen, for participating in today's podcast. And for our listeners, thank you for joining us today as well. You can always visit [ibm.com/smarterystems](http://ibm.com/smarterystems) for more information.

And of course, stay tuned for more podcasts in this series. Until then, for Glen Sheffield, I'm Andrew Verdesca wishing you good listening and good luck.