

INTERVIEW WITH LARRY WEBER

Eric Green: Hello and welcome to a new podcast series from IBM software that explores the challenges IT managers and business professionals are facing today. I'm Eric Green and I'll be talking with a range of experts to discover new perspectives, approaches and examples that can help meet these challenges and introduce you to the capabilities of smarter software from IBM. So let's get started.

Welcome back to our next episode. As we know, the preponderance of data in organizations has just been growing and growing, and over the years, there's been a lot of changes in how we deal with that data. But today, to talk about data warehousing, we have Larry Weber, who is the team leader with IBM Data Warehousing Product Marketing and Strategy. Thanks for joining us, Larry.

Larry Weber: Hey, thanks for having me.

Eric Green: So to start with, what exactly is data warehousing and can you talk about some of the challenges you're seeing organizations face today?

Larry Weber: Yeah sure. Now data warehousing means a lot of different things to different people, but, you know, let's net it out, keep it simple. Data warehousing is really used to make decisions, make business decisions. This is going to be just reporting, you know, if you hear business intelligence, predictive analytics, being able to predict the future from your own data, these are types of situations where you're going to want and use a data warehouse.

Some of the challenges around data warehouses is, you know, let's be real honest here. It's not easy. What software are you going to pick? What hardware? How's it going to run? How are you going to optimize it? At this point, we even have the appliance in our – in IBM's purchasing of IBM Netiza now. Then okay, let's assume you've got the data warehouse running on the hardware, ready to go, and how do you load it? You know, how do you bring in the disparate data sources and then how do you store that data, and then ultimately how do you actually leverage that data for other decisions? It's difficult. You know, I'm not even going to go into the point now about talking with quality of data. There's a lot of things that can go wrong in data warehousing situations, and it takes a lot of expertise. The two key areas to focus on, right, some other issues that people are having with data warehouses once they

get up and started are going to be issues around scalability, or building on databases rather than data warehouses that don't scale and don't grow, once the rest of their organization starts utilizing and leveraging it. And the second one is going to be around performance, if they didn't configure it correctly, if they didn't put together the correct allocation of hardware, storage, server, etc.

Eric Green: So just in all of this, what has been changing over the last several years that has really pushed organizations to have to think so much about data warehousing? I mean what dynamic has sort of caused this shift, or has there not been a shift, this has been happening for longer than the last few years?

Larry Weber: Well, if you think about what data warehousing is used for, right? Making business decisions. If you think about the economy, right? And think about companies – more and more companies, they're competing. They're getting aggressive, there's acquisitions, there's, you know, mergers. And we're fighting these mano a mano battles in – across all industries. Business intelligence and analytics runs on the data warehouse. You need good, quality data to make decisions with. If you're pulling from bad data, you're making bad decisions, they should be irrelevant.

So organizations that want to be able to leverage existing information, through maybe the past 100 years of sales or even the past five years, whatever it may be, if you want to use that to predict the future, to predict your customer's behavior, to understand hmm, okay, this customer purchased Product A. What's the likelihood that customer's going to purchase Product B, C, D, E, F? This is all information that organizations can have at their fingertips. These are essentially the results of what data warehousing brings. Data warehousing is the core of this. It's the foundation to allow these kinds of questions from organizations.

Eric Green: So to add to that, how can our listeners know they need one. Or, in other words, can you talk about how important organizing data is?

Larry Weber: So ask yourself how you're using your own data. You know, I mentioned, you know, a basic case of market basket analysis. You know, where you buy an item – who is more likely to buy another group of items associated there? Do you want to understand questions like this? I mean, would you like to ask them? Would you like to understand which of your customers are more profitable? Do you want informed decisions, analytic decisions, predictive decisions, based on your own data? Right then and

there, you should say gosh, I need a data warehouse. Situations where you have your own database in store. Ah, I've got it already. I've collected all this data. That's what I'm using in my spreadsheet or my reporting tool. Well make sure you're not pulling it from your basic transactional data. Because many times, you know, there's an inherent difference, right, between transactional data, and analytical data processing. You know, transactions can be short, sweet – think like an ATM machine, right? And a bank that's around the world making many transactions. It's very, very important to have these transactions, have them available, have them stored, and your system behind the scenes has to make sure that no matter what, you're always able to process them.

You think about someone ordering let's say a pair of shoes. Right? The company has to keep that customer profile, keep that data. You know, there could be a separate database on the returns of the shoes. Analytical processing is going to be completely different. Analytical process is the core of data warehousing. Right? And in the analytical processing side of things, you're going to look at sales for those shoes for the past ten years. All right? By different customers, by third party products, whatever it is associated with that, and then potentially take that transactional data that we just talked about, like premier of shoe purchases, bringing it together into a view, into the data warehouse that I now as, you know, either an executive or as a business analyst or as a marketer, I can look at that data in one single view, data mine, dig through it, drill down through it, to find those golden nuggets of information that are going to tell me more information about my customers and my business and how to perform better in my own marketplace.

Eric Green: Hey, so does one size fit all when it comes to data warehousing or is that just not necessarily the case?

Larry Weber: It's definitely not the case. You know, there are probably, Eric, a good number of companies out there that will say hey, you know, we have the magic solution that will solve all of your problems. It's really just not true. There are many solutions that hey, they do a great job across the board, but different companies utilize data in different ways. One example is this is, you know, you might have a company that just one line of business, you know, really wants to focus on – gosh, we have a problem, we just want some simple churning data warehouse, and we just want to get it out the door, don't want to spend a lot of money, and I need it yesterday. There's a place for that, and there are solutions for that offering.

There are also different areas in which an organization could be on many different operating systems. They could be using different processors, different types, there's different requirements. Even an example of real-time, you know, business analytics – the whole gamut's going to be run across there. And there are a number of solutions from, you know, appliances that basically, you know, open up the box, plug it in. It's a black box, right? Throw data at it, that thing hums. It runs super fast, it crunches through analytic questions, queries, super fast, and gets up and running, you know, basically in 24 hours.

You also have workable optimized systems, right? And from an IBM perspective, it's our Smart Analytic system, etc. in our portfolio there. And which, you know, these are modular units that can process both transactional and analytical queries, together in real time. And we also have those people out there in organizations that might be standardizing one piece of hardware, one piece of whatever it may be, they want to run their own software. They want to see it through. They have the teams to be able to support it, and they're going to want a software only solution. So to make that short answer super long, no, there are many, many different ways and sizes for data warehouses for many different organizations.

Eric Green: Well fair enough. So about some examples, maybe, of how organizations are using data warehouses.

Larry Weber: Sure. One example would be a travel agency who we've worked with in Germany. In this situation, they're bringing together real time data from both transactions and combining that with the data in their existing data warehouse. This enables them, basically on demand, travel bundle that they can offer to their customers. Also online bookings based on their customer preferences, etc. The net is here is, en masse, they have an ability to tailor their offerings to their market demand, in real time.

Another example, banking – we have a leading bank in Brazil we're working with. Here, they're leveraging their data warehouse solution to enable, you know, all their different branch banks with the ability to offer real time offers to their own customers. If you think if you walk into a bank, and oh, you know, it's Larry Weber, here's his customer ID. That bank then in real time can calculate and look through a pre-developed data mine and say, ah, Larry is maybe in the 80% of our top customers. But we can put him in the 100% mark if we upsell him or cross sell him on A, B, C, or D.

These are kind of the ideas that these banks – that this bank specifically – is implementing, real time offers.

And even third one out there is in the government space – the New York Police Department. They're using our business analytics and they're utilizing it with mapping and visualization software to show crimes as they are happening in real time, so that they can proactively attack them. So it's not dependent on your industry, or, you know, whether you're in government or you're in, you know, profitable business. You know, business analytics and data warehousing have a need across the gamut.

Eric Green: Very interesting. So how is IBM innovating in this space?

Larry Weber: Well, you know, you've seen one - recently we purchased IBM Netiza and really looking at, you know, the appliance market and adding that and rounding out our portfolio with Netiza. To tag onto that, you know, a lot of simplification, ease of use. You know, one of the things I talked about early on is the difficulty of putting these data warehouses together. Right? There are so many failed data warehousing projects that are in the marketplace.

One of the areas is warehouse model packs, in which we're, you know, basically taking goodness from industry models, and taking all of the data models and hardening down to the database, data warehouse level, you know, both for Netiza and for Infosphere warehouse software and Smart Analytics systems. Here, we actually – organizations can come out of the box running and seek immediate answers to their questions. It can integrate a _____ reports and all of the, you know, kit and caboodle that you're going to need to run a data warehouse out of the box.

That's part of it. Another angle is looking at, you know, ease of purchase, ease of understanding, you know, the costs associated. You know, recently we have announced terabyte pricing as an option for data warehouse, an Infosphere warehouse. Basically, you pay for what you use. You know, in the past it's been hey, I have to buy this big old crazy bumble and I hope I use. And I'm not sure when I'm going to get to a certain point and when I'm going to scale out of it. And it becomes a big, confusing mess, specifically when you're looking to plan for costs. Terabyte processing allows you to plan ahead of time, and allows you to grow as you need.

And a third point here would be virtualization. Right? The idea of deploying warehouses, you know, now we have offering in the VMware image. You know, the virtual machine of a warehouse that now allows business partners and customers to download the warehouse, pre-tuned and configured in that VMware space, that they can throw in pretty much any hardware they want. So we're making it easier, we're making it simpler, and really a quicker time to value.

Eric Green: So Larry, earlier you mentioned Netiza, which I believe is one of the IBM products. Could you explain what that is and does, please?

Larry Weber: Yeah, absolutely. So IBM Netiza is our data warehousing appliance. What is the appliance compared to the rest of the data warehousing portfolio? Well what Netiza brings to the table is essentially a black box solution. It is an out of the box data warehouse machine. You basically plug the thing in, turn it on, and throw data at it. This is going to be pretty much your best friend for super fast analytical queries. Netiza itself really fits nicely into our entire portfolio.

Historically, you know, we have our Infosphere warehouse software based on DBII which we're using with many enterprise customers around the globe. We also have smart analytics systems and workload optimized systems, which are, you know, basically warehouse software that's been hard-tuned and coded for hardware modules, right, that organizations can just add and scale as they need. Pre-tuned, pre-optimized, all that goodness. Where Netiza really fits in here and really, really rounds out the portfolio is the idea of speed and simplicity. The Netiza box actually comes out ready and willing just to take on your analytical queries and at a super fast speed. They're quite affordable as well.

Eric Green: That is great stuff. Thank you so much for your insight. And it seems we're actually out of time for this podcast, but Larry, thanks so much for joining us today.

Larry Weber: Hey it was a pleasure. Thank you so much for having me.

Eric Green: Thanks for listening. Please do visit IBM.com/software to connect with our experts, continue the conversation, and to learn more about smarter software from IBM. Let's build a smarter planet.