Big Data Challenges on System z

Enterprise Tech Journal recently had the pleasure of speaking with Kevin Goulet about overcoming the challenges inherent in managing Big Data on System z. As vice president of Product Management for the CA Technologies Database Management portfolio with a long history of product leadership roles, Kevin is in a prime position to observe the Big Data market and the issues companies face in establishing effective Big Data management solutions. Since a sizeable portion of the

By Denny Yost
The audience of Enterprise Tech Journal is DBAs who are often charged with the day-to-day handling of Big Data solutions, we focused our questions on issues that will help them be more effective in that role. Let’s see what this expert in the field had to say.

**Enterprise Tech Journal:** Thanks for joining us today, Kevin. Let’s start with your observations about the current state of the Big Data market.

**Kevin Goulet:** Glad to be here! As you know, a majority of companies have massive data stockpiles on their mainframes culled from the cloud, social media, mobile devices, email, the Internet of Things, relational and non-relational databases, spreadsheets, video and countless other sources. This store of data is known as Big Data because it’s grown so large that traditional data analysis and management solutions are too slow, too small and too expensive to handle it. Despite (or perhaps because of) the exponential growth of data, and as with any technology in the early stages of adoption and development, most companies are in the discovery stage of evaluating the best means of extracting value from it.

**ETJ:** Before we talk about how to deal with the challenges of Big Data management, can you define those challenges for us?

**Goulet:** The challenges facing Big Data administrators, the hands-on users of any business intelligence solutions that companies have launched, are critical to the success of any such initiative. First, you need an effective process for moving a cornucopia of structured and unstructured data from the mainframe to the Hadoop environment and potentially back again. (Hadoop is the technology for map reducing software, the engine behind extracting value from Big Data analysis.) Those processes need to be secure, close to real-time and performed at regular intervals. Hadoop clusters need to be up and running all the time, and the data moving back into the mainframe must be clean and in sync with the original database schemas, so that it can be used productively. It’s also important to automate management so that DBAs don’t spend all their time doing manual scheduling. So I would say that the overall concerns of Big Data administrators are processes and reining in the amount of time they commit to Big Data management.

**ETJ:** Just how widespread are these challenges?

**Goulet:** We speak with administrators across the gamut of industries and all face these challenges. If misery loves company, then DBAs can take solace in the fact they aren’t alone in dealing with these issues.

**ETJ:** What about the challenges to the enterprise in creating a Big Data management solution?

**Goulet:** Funny you should ask, as I wrote a column for this issue of Enterprise Tech Journal (see page 68) that addresses that very topic.

**ETJ:** Who would the Big Data administrator be in the typical enterprise? (Editor’s note: Big Data administrators may want to suggest that their managers read Kevin’s column.)

**Goulet:** Big Data administrators, the people we envision being responsible for business intelligence and Big Data management, will likely be plucked from the ranks of DBAs.

**ETJ:** What would be their role, beyond what you’ve already talked about?

**Goulet:** They would interface with many teams across the enterprise, such as supply chain management, lead generation, customer behavior, development, operations and other business users. The ongoing goal would be to
ensure the Big Data solution provides intelligence (i.e., deeper insight) that the teams—and the enterprise as a whole—need to be successful. That will call for someone with highly developed diplomatic skills who can partner with “competing” interests to prioritize what’s best for the business.

And there’s more to it than that. When data returns from analytics to the system of record, the Big Data admin can provide Ops with info about greater efficiencies they could gain or Marketing with info about customer buying patterns or how they respond to marketing efforts. They could supply Development with customer responses to surveys or a more streamlined way of approaching feature sets or bug fixes. They could give other business users intelligence about customer needs and how to improve supply chain management to make it more efficient and effective. In fact, the value of properly analyzed Big Data is practically endless.

**ETJ:** What other roles are essential to an effective Big Data solution?

**Goulet:** Someone needs to manage the Hadoop environment. An application developer (and often a solution architect) would be responsible for creating the infrastructure and processes that the Big Data admin uses. The Big Data scientist/engineer applies the data to the business.

**ETJ:** What do current Big Data solutions look like?

**Goulet:** Today, there’s no unified solution that addresses all these challenges. Current solutions, and I use that term loosely, are homegrown “science projects” cobbled together from individual products from a variety of vendors to meet short-term, tactical, siloed requirements (rather than strategically addressing the entire enterprise). These “solutions” operate on separate platforms with varying degrees of interoperability. Simply put, the current state is far from ideal.

**ETJ:** What can CA Technologies do to help customers get on the Big Data bandwagon?

**Goulet:** CA Chorus is designed to manage multiple solutions and Workload Automation solutions take care of that problem. Products in the CA Database Administration Suite for DB2 for z/OS can support aspects of data management. But what companies really need is a solution to help them consolidate the management of their entire Big Data environment—a Big Data management platform that makes it easier and more efficient to transform Big Data into real value to the business.

**ETJ:** In your view, what will the ideal solution look like?

**Goulet:** When talking about Big Data, building the ideal solution is the ultimate challenge. The ideal solution for companies seeking the advantage of powerful business intelligence will be a unified Big Data management product that eliminates the complexity of today’s piecemeal solutions. It will move data from the mainframe to the analytics environment and potentially back into the system of record in an efficient, repeatable, automated and secure way. This new approach will allow for the most sophisticated analytics and automate scheduling. To accomplish this, the solution will incorporate data integration that probably needs to include both extract, transform and load (ETL) and data virtualization to support the physical movement of data between environments as well as federation capability, which allows queries that span both environments without physically moving data. Also necessary is a full management console that spans the two environments and includes collaborative capabilities to accommodate two distinct sets of users: the mainframe database
“Despite (or perhaps because of) the exponential growth of data, and as with any technology in the early stages of adoption and development, most companies are in the discovery stage of evaluating the best means of extracting value from it.”

~ Kevin Goulet

administrators and administrators managing the capabilities that make up the Big Data environment. I could go on, but the bottom line is that the ability to schedule and manage the flow of data and eliminate complexity will be key to a useful, cost-effective solution.

ETJ: What’s your best advice to someone who wants to learn more?

Goulet: At the risk of sounding self-serving, one of the best actions readers could take is to visit the CA Technologies Big Data resource page at www.ca.com/bigdata. Here they will find all the current thinking on this fascinating approach to gathering business intelligence, which is rapidly becoming recognized as a competitive advantage in nearly every industry. Also, reach out to your most trusted software solutions provider and let them know what you’re seeking from Big Data. Their development team may already be working on a solution and would welcome your input. And if your input is incorporated into the solution, you could become a hero to your business by giving it a leg up on the Big Data challenge once the product is available for use. ETJ

Denny Yost has more than 30 years of mainframe and IT experience. He’s the associate publisher and editor-in-chief of Enterprise Executive and associate publisher of Enterprise Tech Journal.

Email: Denny@ESMPubs.com