

## DB2 News from IDUG 2013

*By Craig S. Mullins*

This past month, from April 29th - May 3rd, 2013, IDUG held its annual North American DB2 Tech Conference in Orlando, Florida. It was a significant year in IDUG's history as the conference marked the 25th anniversary for IDUG... and the 30<sup>th</sup> anniversary for DB2!

I was lucky enough to attend and deliver several presentations as part of the conference. Regardless of whether you, too, were lucky enough to attend (or not), this article will cover the highlights of the conference.

Looking back on the event, it is clear that the focus of the conference was squarely on Big Data. However, this was not the only theme. From my perspective, the conference highlighted three major themes:

- Big Data
- BLU Acceleration
- DB2 11 for z/OS

We'll take a look at each of these areas and examine what was uncovered and discussed at IDUG 2013.

### **Big Data**

Given the excitement in the market regarding Big Data, it should come as no surprise that IBM is embracing Big Data and extending DB2 to embrace and participate in Big Data and analytics-related projects. Big Data support and development is fast becoming a key basis for competition and growth across the industry... which comes with the promise of driving a new wave of productivity.

The Tuesday keynote session at IDUG was titled "Database Software for the Era of Big Data" and it focused on the Big Data capabilities required to capture, organize and analyze a variety of data types within the context of your organization's information. The session was delivered by Bob Picciano, General Manager for Information Management at IBM. Bob began auspiciously as he stated that "Big data is the next natural resource."

It is obvious that IBM has heavily invested in DB2 to deliver advanced functionality for enabling Big Data and analytical capabilities. Another of the conference's highlighted themes, BLU Acceleration, is a Big Data advancement. Without getting ahead of ourselves, BLU Acceleration delivers rapid access to large amounts of data using a combination of technologies, which we will discuss shortly.

IBM also talked about its Big Data focused PureData System for Hadoop. Announced in early April, IBM PureData System for Hadoop is built to optimize Apache Hadoop data services for big data analytics and online archive with appliance simplicity. In other words, it delivers a pre-packaged system for Hadoop implementation.

IBM PureData System for Hadoop combines IBM InfoSphere BigInsights and IBM System x hardware for an integrated Hadoop system. The combination of hardware and software enables the IBM PureData System for Hadoop to provide enterprise Hadoop capabilities with easy-to-use analytic tools and visualization capabilities. It includes developer tools, analytic functions, and administration and management capabilities, along with the latest versions of Hadoop and associated projects. Furthermore, it integrates with other IBM technologies including DB2, Netezza, PureData System for Analytics and InfoSphere Guardium.

At the keynote IBM also highlighted its Big Data Platform. IBM claims to be unique in having developed an enterprise class big data platform that allows its customers to address the full spectrum of Big Data business challenges. The primary benefit of the platform is the ability to start with one capability and easily add others as customers add Big Data requirements.

The keynote session also spent some time touting the extreme performance boost that can be achieved using IDAA – IBM DB2 Analytics Accelerator for z/OS. IDAA is a workload optimized appliance that blends System z and Netezza technologies to deliver, mixed workload performance for complex analytic needs. It runs complex queries up to 2000x faster while retaining single record lookup speed and eliminates costly query tuning while offloading query processing. Yes, you read that correctly – up to two thousand times faster!

### **BLU Acceleration**

But let's get back to BLU Acceleration. Basically, BLU Acceleration adds a column store capability to DB2 10.5 for LUW (with the promise that it will come to z/OS soon). A column store physically stores data as sections of columns rather than as rows of data. By doing so, data warehouse queries, customer relationship management (CRM) systems, and other types of ad-hoc queries where aggregates are computed over large numbers of similar data items can be optimized.

But BLU Acceleration is not just a column store. IBM had delivered three additional capabilities and improvements with BLU Acceleration. The first is called “actionable compression,” which can deliver up to 10x storage space savings. Some of the beta customers are getting 90-95% data compression for their large data warehouse database tables. But why is it called “actionable?” Well, there are two key ingredients that make the compression actionable. There are (1) new algorithms enabling many predicates to be evaluated without having to decompress and (2) the most frequently

occurring values are compressed the most, thereby saving the greatest level of storage space.

The second new feature of BLU Acceleration comes via the exploitation of the SIMD (Single Instruction Multiple Data) capabilities of modern CPUs. The basic idea behind SIMD is the ability for a single instruction to be able to act upon multiple items at the same time, which obviously can speed up processing.

And finally, BLU Acceleration adds data skipping technology. You can probably guess what this does, but let's explain it a little bit anyway. The basic idea is to skip over data that is not required in order to deliver an answer set for a query. Metadata is stored for sets of data records that can be accessed by DB2 to determine whether that particular set of data holds anything of interest. If not, it can be skipped over.

BLU Acceleration is being delivered first in DB2 10.5 for LUW. And best of all, it is simple to use. All that is necessary is to specify ORGANIZE BY COLUMN in the DDL for a table to make it BLU.

So what, you may ask? In my opinion, BLU Acceleration is a very significant milestone in the history of DB2. It brings a column store capability that can be implemented right inside of DB2, without any additional product or technology. So you can implement a multi-workload database implementation for the Big Data era using nothing more than DB2 software. BLU Acceleration provides blazing speed and can act upon large amounts of analytical data. And that is something we all should consider when embarking on our Big Data projects.

## **DB2 11 for z/OS**

Given my mainframe background, I was excited to hear some of the details of DB2 11 for z/OS sneak out at the IDUG DB2 Tech Conference. Of course, the highest level details were already available given that IBM had announced the Early Support Program (ESP) in the Fall of 2012. If you are interested in reading the details of the ESP it is in the web at <http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?subtype=ca&infotype=an&appname=iSource&supplier=897&letternum=ENUS212-364>.

As usual, IBM offered up the very high-level goals of the new release, which are shown here:

- CPU reductions and performance improvements for certain OLTP, heavy INSERT, select query workloads, and when running queries against compressed tables.
- Improved data sharing performance and efficiency.
- Improved utility performance and additional zIIP eligible workload.
- Cost-effective archiving of warm and cold data with easy access to both within a single query.

- Intelligent statistics gathering and advanced optimization technology for efficient query execution in dynamic workloads.
- Additional online schema changes that simplify management, reduce the need for planned outages, and minimize the need for REORG.
- Productivity improvements for DBAs, application developers and system administrators.
- Efficient real-time scoring within your existing transaction environment.
- Enhanced analysis, forecasting, reporting and presentation capabilities, as well as improved storage management, in QMF.
- Expanded SQL, SQL PL, temporal and XML function for better application performance.
- Faster migration with application protection from incompatible SQL and XML changes and simpler catalog migration.

But additional details were unveiled at IDUG. To me, the most significant one was at the DB2 Trends and Direction session when the speaker let it leak that DB2 11 for z/OS was about six months away... and then added after a brief pause “unofficially.” This places it squarely within the fourth quarter of 2013. And that makes sense. After all, DB2 10 has been generally available now for three years and that marks the time span between Version 7 and 8, Versions 8 and 9, and Versions 9 and 10.

Another interesting tidbit was that 25 percent of DB2 10 for z/OS customers migrated directly from V8 using skip-level migration support. But there will not be skip-level support in DB2 11. That means that you will have to be on DB2 10 in order to get to DB2 11. Although it was not spoken as a promise, it was noted that skip-level support comes along approximately every 10 years, whereas new versions are released on a three year cycle.

But what are some of the tidbits that were revealed about DB2 11? Here are a few:

- DB2 11 will not support packages bound before V9
- Transparent archiving support will be delivered leveraging the temporal support added in V10
- The open data set limit will increase to 200K
- RBAs and LRSNs will increase from 6 bytes to 10 bytes
- DROP COLUMN support is being added
- Performance improvements for DPSIs (Data Partitioned Secondary Indexes)

All-in-all, it sounds like it is a release to look forward to... and I'm sure more details will be forthcoming over the next six months.

## **Planning Ahead**

You can learn an awful lot by attending an IDUG DB2 Tech Conference event. That should be clear just if you skim this article. But keep in mind that this article only hits the highlights. There was a lot more covered at this year's conference. Four days of technical session, one day of full-day seminars, the opportunity to take certification exams, complimentary workshops sponsored by IBM, and the chance to network with vendors and your peers. All in one convenient location.

So be sure to circle the week of May 12 thru 16, 2014. The North American IDUG DB2 Tech Conference will be held that week in Phoenix, Arizona at the Sheraton Phoenix.

See you all there!