

# Do GPU optimized databases threaten the hegemony of Oracle, Splunk and Hadoop?

By  Kurt Marko April 11, 2016

## SUMMARY:

GPU optimized databases are rapidly moving from science project to business reality. Here's what it means.



6 Comments



I don't use the R word too often so sit tight and bear with me.

There's a revolution brewing in how applications are designed that is fueled by the desire to exploit non-traditional processing engines instead of traditional general purpose CPUs and it's rapidly moving from academia and science projects to business. As I recently explained in a [column on the looming specialization of cloud hardware](#),

*"The days of homogeneous server farms with racks and racks of largely identical systems are over.*

The past week's events (major announcements by Intel, NVIDIA and IBM) highlight a continuation of the trend towards greater hardware and software specialization for specific workloads. Just as the era of mainframe hegemony, where IT was built upon a one-size-fits-all platform, was largely replaced by cookie-cutter x86 systems powering client-server and Web applications, today a new generation of hardware accelerators designed for various categories of applications and workloads are displacing the do-everything, plain-vanilla x86 system.

The catalyst for hardware-optimized software has been deep learning algorithms using brain-inspired neural networks for image recognition, categorization and positional localization. The applications are many, not just mundane auto-tagging of Facebook pictures, but extending to manufacturing, such as allowing robots to 'see' and respond for example by tightening a bolt or selecting a package, and autonomous vehicles by enabling cars to accurately navigate through crowded, urban terrain in foul weather.

The high-profile robotics applications of hardware-accelerated algorithms are just the start. The most pervasive and versatile of these accelerators, the GPU, can also tackle big data analysis by breaking apart and parallelizing the data set, then filtering, recombining and presenting the results in rich graphic visualizations.

Several companies at the NVIDIA Developers Conference (GTC) had products on display that used GPUs to achieve database query and visualization performance traditional software could only dream about. For example, try real time, interactive filtering, clustering and display of billion-row datasets as easy as creating an Excel graph using a standard RDMS and visualization add-on like Tableau.

Two pioneers of GPU-optimized relational and column-store databases are GPUdb and MapD. Each can tackle many of the same problems that enterprises currently solve with a traditional RDMS like Oracle and SQL Server, or log analytics software like Splunk and Sumo Logic. The killer difference is that they process data in near real time, including traditional SQL functions like selects (filters) and joins (merges), but add accelerated visualizations to the output. Both vendors show numerous tests in which their product is one to two *orders of magnitude faster* than an SQL database.

To get a bit geeky for a moment, it's all made possible by loading the data set into very fast GPU RAM, breaking it into chunks usable by an individual GPU core, of which **there are thousands in today's processors**, and independently processing the elements before recombining the results. Since the data is already on a powerful graphics rendering engine, the software can then display the resulting using all manner of 2- and 3-D visualizations at lightning speed. Both

GPU accelerated databases are superficially analogous to in-memory DBs, however the availability of many more GPU cores, much faster GDDR5 memory and instruction sets optimized for combining, multiplying, summing and filtering complex data sets makes GPUs particularly adept at accelerating many database operations.

For example, the U.S. Postal Service uses a rack full of GPU servers running GPUdb to analyze each carrier's location data to calculate delivery time estimates based on aggregated results on a particular route, provide real-time notifications to supervisors about abnormal carrier activity (e.g. outside the assigned route, unusually slow deliveries) and optimize ad hoc routes when a carrier is out sick and no substitute is available. With 154 million addresses spread over more than 200,000 delivery routes, the USPS has a sizable database, yet GPUdb is able to process complex queries and display 2D visualizations in the time it takes to load a Web page.

Although the typical demonstrations used by GPUdb and MapD, filtering and mapping queries against a database of domestic **flight records** or **Tweets**, are gimmicky, there are serious business applications for this kind of speed including analysis of retail transactions, security and IT operations logs or geo-located energy usage trends. Plus, both products are SQL compatible, GPUdb already fully supports the SQL92 standard and MapD expects to fill its remaining gaps later this year.

Yet relational and column store databases aren't the only type amenable to GPU acceleration. **Blazegraph**, also on display at GTC, does the same for graph databases, supporting up to 50 billion edges on a single system. Like GPUdb, its origins lie in defense research, and it likewise boasts workload acceleration metrics measured in orders of magnitude. Indeed, the company cites a 300x performance increase on a standard **SPARQL benchmark**. Although graph databases are commonly used for social network analysis, they are also useful for security forensics and fraud detection, recommendation engines and network and IT operations analysis.

The key benefits of GPU-powered databases include very fast data ingestion rates and near real-time data exploration and display with complex visualizations. And thanks to GPU-enabled servers on **AWS**, **Azure** and **Softlayer**, each can be deployed in the cloud, avoiding the expense of building dedicated systems.

## My Take

A year or two ago it was easy to dismiss GPU-optimized databases as niche products designed for data scientists

operating on relatively modest data sets. However, with vast improvements in GPU performance and memory capacity, the ceiling is rapidly rising. With a performance curve exceeding that of CPUs and incorporation of a much faster data interconnect, **NVLink, into just-announced Pascal latest GPUs**, which IBM is also integrating onto its **next generation POWER CPU**, enables dramatically expanding the size of GPU workloads by scaling across multiple systems with little loss in performance.

At the GTC opening keynote, NVIDIA CEO Jen-Hsun Huang discussed the accelerating rate of GPU performance. In speaking of deep learning applications, **he highlighted** a 12x speedup in performance, far faster than the 20-25% annual improvements seen with conventional CPUs. The slope of this performance curve will not be lost on enterprise application developers where databases represent the next frontier in GPU penetration.

GPU-optimized databases have already gained the attention of the defense-intelligence complex, some large enterprises and influential technology executives. To wit, as I stopped by the GPUdb booth at GTC, Huang himself was engaged with its representatives, questioning them about the product, getting a demo and closing with an offer, “Just let me know what I can do to help.” Indeed, Huang encouraged the GPUdb team to take on both traditional enterprise staples like Oracle and emergent alternatives like Hadoop based on performance *and* licensing cost advantages.

As my last column detailed, hardware-optimized software is primed to disrupt enterprise IT, just as it’s already done in many areas of academia and HPC. Databases are at the forefront of a trend that will extend to many applications in realtime analysis, correlation and prediction of events from all manner of systems and applications. While incumbent RDMS providers like Oracle and big data platforms like Hadoop don’t face an immediate threat, they ignore this technological trend at their peril.

*Image credit - Story image via VM World, featured image*

*Disclosure - The author’s T&E were mostly covered for the events referred to in this story by the vendors mentioned.*

Read more on:

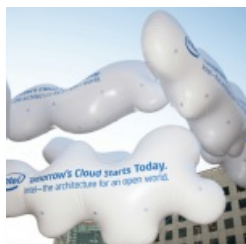
Analytics planning and data analysis



Would you like to subscribe for more content?

Click for options

## Related Stories



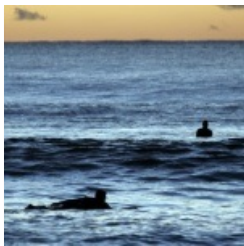
Intel's software evangelist talks cloud architectures and the Internet of Things



## Cloudera CTO on Big Data analytics and security risks



## Augmented reality for business - present and future use cases from Metaio



## Where are all the Salesforce Wave analytics apps?

Powered by **Contextly**



Ashin bhatt says: April 11, 2016 at 11:48 pm

GPU based development could struggle a bit due to lack of Linux (open source) support.  
([Reference](#))



Alex C. says: April 11, 2016 at 11:56 pm

I'm actually impressed with the GPU scaling techniques that they developed. Sadly no product is available to download, so it will still take some time until someone will release software for the public.



Georges says: April 12, 2016 at 11:23 am

Alex, Ashin,

There are two PostgreSQL-based GPU acceleration projects:

- PgStorm: <https://github.com/pg-strom/devel> , <https://wiki.postgresql.org/wiki/PGStrom>
- PgOpenCL: <https://wiki.postgresql.org/images/6/65/Pgopencl.pdf>



Evelyn Kotler says: April 12, 2016 at 1:23 am

Adding SQream Technologies to the list of GPU-powered big data analytics databases.



Arnon says: April 12, 2016 at 2:05 am

SQream DB predates both!  
([Reference](#))



Milosz Tanski says: April 12, 2016 at 6:05 am

I wouldn't count out Intel's Xeon-Phi hardware when it comes to databases. There's quite a few SQL operations who's implementation requires branchy code or working with variable length data, etc... Lots of those those general purpose

cores of Xeon PHI (plus AVX512) would be a good use here.

## Leave a Reply

Your email address will not be published. Required fields are marked \*

Comment

Name \*

Email \*

Website

Post Comment

## Premier Partner Newsfeed



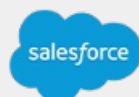
SysAdmins: Transform your IT dept into a center of innovation with the Salesforce platform



6 Priorities CEOs Care Most About

**Rimini Street**

Rimini Street Recognized for Excellence in Customer Service with Three Prestigious Awards



Introducing Salesforce for Messenger



Fueling A Transportation Revolution



May 4: Macy's—Oracle HCM Cloud Customer Forum



NetSuite Supports the 2016 International ICT Awards with LightCMS



Top-Level Parts Profitability Must Be Measured In The Field



It's not all about the code: Crack the soft skills employers value and hack your career



Clark County CIO Heads to .NEXT



Green Flash's Growth Brewed in the Cloud



New Video Playlist for Infor CloudSuite Industrial



HSN is Crushing Retail with 3 Big Innovations



SAP HANA 4 IoT Post-Event Updates

## Partner Newsfeeds

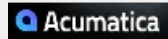
---



Supplier enablement: An idea whose time has come, and gone - Gabe Perez



Your B2B Starter Kit for Cross-Channel Marketing



Technical Tuesday: Building a Project Overview Report in Acumatica Cloud ERP Software, Part 3

## Our value

---

## Recent Conversations

---

Marcus Mossberger on [Time for HR to get our heads in the cloud](#)

Georges on [Do GPU optimized databases threaten the hegemony of Oracle, Splunk and Hadoop?](#)

Tom Harmon on [HR- equal partner or weakest link?](#)

Which tech news site has the sorriest ad tech UX? – Diginomica | Everyday News Update on [Terms and privacy](#)

Tim Janisch on [Time for HR to get our heads in the cloud](#)

## Affiliate Partners

---



TUESDAY 10TH FEBRUARY 2015





Go to Topics

Select Category

Premier Partner Newsfeed



SysAdmins: Transform your IT dept into a center of innovation with the Salesforce platform



6 Priorities CEOs Care Most About

Rimini Street

Rimini Street Recognized for Excellence in Customer Service with Three Prestigious Awards





Introducing Salesforce for Messenger



Fueling A Transportation Revolution



May 4: Macy's—Oracle HCM Cloud Customer Forum



NetSuite Supports the 2016 International ICT Awards with LightCMS



Top-Level Parts Profitability Must Be Measured In The Field



It's not all about the code: Crack the soft skills employers value and hack your career



Clark County CIO Heads to .NEXT



Green Flash's Growth Brewed in the Cloud



New Video Playlist for Infor CloudSuite Industrial



SAP HANA 4 IoT Post-Event Updates

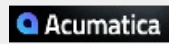
Partner Newsfeeds



Supplier enablement: An idea whose time has come, and gone - Gabe Perez



Your B2B Starter Kit for Cross-Channel Marketing



Technical Tuesday: Building a Project Overview Report in Acumatica Cloud ERP Software, Part 3

Go to Topics

Select Category

C O P Y R I G H T   D I G I N O M I C A   L T D   2 0 1 3 - 1  
C R E A T E D   U N D E R   C R E A T I V E   C O M M O N S   L I C E N

Developed by DevriX