

## REPORT REPRINT

# Kinetica delves into in-database analytics with user-defined functions powered by GPUs

**JAMES CURTIS**

**3 FEB 2017**

The GPU-powered database vendor has delivered its version 6.0 release with in-database analytics, an enhanced front-end visualization tool and the ability to 'pin' data in video memory for certain workloads.

---

THIS REPORT, LICENSED EXCLUSIVELY TO KINETICA, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR REPOSTED, IN WHOLE OR IN PART, BY THE RECIPIENT, WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



©2017 451 Research, LLC | [WWW.451RESEARCH.COM](http://WWW.451RESEARCH.COM)

Kinetica continues to enhance its GPU-powered database with the launch of version 6.0. Headlining the release, the company has added in-database analytics with a tie-in with user-defined functions (UDFs), enabling organizations to drive BI and machine learning powered by GPUs. Kinetica has also announced its Reveal product, a front-end visualization framework tool, as well as the ability to leverage video memory for certain workloads.

---

## THE 451 TAKE

While the GPU database market is fairly new, Kinetica, one of the sector's early players, has shown a propensity to continue to add relevant capabilities to its GPU-powered database offering. Updates in the company's latest version 6.0 include in-database analytics, a visualization front-end tool and the ability to drive some workloads in GPU memory. Machine learning, deep learning and similar advanced analytical workloads are known use cases for GPUs, so pairing Kinetica's GPU database with these analytical workloads appears to be a logical next step. The concept of data democratization is a goal of many organizations, so we are anxious to see how this strategy plays out with the company's customers. What we are intrigued about, however, is the broadening appeal of GPUs - encroaching, for instance, on the capabilities of data-warehousing wares.

---

## CONTEXT

In October 2016, we initiated coverage of Kinetica, noting its early start doing consultancy work for the US government by the company's founders. Their early work led the pair to launch GIS Federal with its GPUdb product. The company and its GPU database offering were later renamed as Kinetica in 2016. Certainly, over these past seven years, the company and its product have evolved. But as we recently wrote about, the GPU database space in general has been quite active and growing as a whole.

As for Kinetica's growth, the company is based in San Francisco and also has offices in Arlington, Virginia. In 2015, it entered Europe and currently has a managing partner there to drive its business. The company claims more than 10 paying customers and over 50 employees, and reports solid revenue growth, although it does not disclose revenue. Funding totals \$13m.

## PRODUCTS

Kinetica positions its GPU columnar database as being ideal for OLAP workloads, and particularly so for high-cardinality workloads that often require a good deal of compute power, given the complexity of the data. For version 6.0, the company has extended the database's OLAP capability by enabling in-database analytics. It notes that it encounters many potential customers that are doing deep learning and more advanced analytics on HPC systems that leverage GPU processors. Kinetica cites a few challenges these organizations face. One is that data scientists or other technical specialists often need to pull data from a system of record and load it into an HPC system to perform the analytics leveraging certain algorithms. While HPC systems are well equipped to handle advanced analytics because they leverage GPUs, there is also a price to be paid as it requires moving the data from one system to the other.

With its in-database analytics, Kinetica is addressing these challenges. In order to enable in-database analytics, the company developed an orchestration layer that leverages UDFs along with facilitated integration to the GPUs. UDFs are user-created functions that can accept customer algorithms or third-party algorithms, which can also access known machine-learning libraries such as TensorFlow, Spark ML and Caffe. The UDFs have direct access to the CUDA APIs, which can then access the GPUs for parallel processing. Initially, the company supports UDFs with C++ and Java, with Python in the works - to run on the GPUs, however, the functions need to be converted to CUDA.

Kinetica is promoting its in-database capabilities for a few reasons. One is that it effectively opens up the door for machine learning, deep learning and other advanced analytical workloads to be run alongside BI workloads, all within the same environment. There is also the benefit of less data movement because, theoretically, a data scientist can develop algorithms and models and then make those available to business analysts, for instance. As such, the company views this function as broadening its user profile – a type of data-democratization strategy.

In addition to in-database analytics, Kinetica has unveiled a front-end visualization tool called Reveal. Visualization tools make a lot of sense, especially when paired with GPUs, primarily since such pairing enables interactive or drill-down analytics, something traditional BI visualization tools often struggle with, particularly on extremely large multi-hundred, million-row tables. Spatial or location-based data is a particularly favorable use case with GPUs. As such, Kinetica processes the spatial data on the GPUs and then sends a small PNG file to Reveal, which allows users to interactively query a map. Reveal also enables users to create dashboards by dragging and dropping widgets, as well as integrate with several map providers such as Google, Bing, ESRI and others.

Other enhancements include VRAM Boost, which allows users to ‘pin’ data in GPU memory (VRAM) for some workloads and use cases that require low latencies. Kinetica runs in a hybrid CPU in-memory architecture that leverage GPUs, so with VRAM Boost, customers have the option to have certain workloads run in GPU memory (VRAM), which offers even greater performance.

## COMPETITION

Kinetica is part of a small but growing collection of GPU-powered database vendors. And although there may not be as many players in the market currently, there does exist a certain level of competition among them. The company differentiates itself with a hybrid in-memory CPU- and GPU-enabled database, so the addition of VRAM Boost further enhances its hybrid story. That being the case, MapD is likely Kinetica’s closest rival, given that it runs primarily in VRAM, but can leverage CPUs as well.

Further, Kinetica’s Reveal visualization tool is comparable to MapD’s Immerse visualization offering. And while both firms enable third-party BI tool integration, each of their visualization tools is particularly suited to work with GPUs – especially well with drill-down analytics. BlazingDB is another contender, although its GPU database runs on disk. SQream Technologies deploys a NoSQL database under its covers, while brytlyt leverage PostgreSQL.

While we noted that Kinetica’s GPU-powered database can be competition for the data-warehouse suppliers, with the company’s newly released in-database analytics functionality, we believe it is even more so now. As such, we see Kinetica encroaching on a number of data-warehouse vendors, including Oracle, IBM, Teradata, Microsoft, SAP, HPE and Amazon Web Services, among others.

Lastly, some in-memory and NoSQL database providers also deserve a mention. These include SAP HANA, MemSQL and VoltDB, as well as NoSQL specialists such as Redis Labs, Cassandra, MapR-DB and Aerospike.

## SWOT ANALYSIS

### STRENGTHS

By enabling in-database analytics, Kinetica is demonstrating that GPUs can have a broader appeal for organizations, particularly for more advanced analytics.

### WEAKNESSES

The GPU market is maturing in many ways as organizations continue to understand and leverage the technology. Consumption can also be challenging, particularly on the cloud front, where instances can be limited.

### OPPORTUNITIES

GPUs shine for drill-down analytics, and when combined with the company's Reveal tool, this should give Kinetica additional opportunities. However, the notion of driving both BI and more advanced analytics such as machine learning within a single environment should, at a minimum, spur new customer conversations.

### THREATS

While the GPU market is growing and does have an expanding roster of players, the overall sector is not as large as some of the other database segments we track - data warehousing, for instance. The biggest challenge lies in educating organizations on what GPUs can do, given that most of them still have a CPU-first mentality.