In our quest for disruptive big data technology, we crossed the Atlantic to find German-based rasdaman, which truly has the potential for becoming a game changer right tomorrow. Turns out the Caribbean sounding name actually stands for “raster data manager.”

Peter Baumann, inventor CEO of the company, early on has recognized a gap in big data analytics: missing support for massive matrices and datacubes, also known as multi-dimensional arrays. We find these arrays everywhere: in business, such as stock risk analysis and OLAP; in Life Science; in exploration data; and in industrial simulation. Even analyzing large graphs, like the Facebook One, can be done through array operations. In fact, datacubes are recognized as a key paradigm that is simple to grasp and easy to query. As we see the potential for array engines, we wonder: How could that go undiscovered?

“SQL has missed a train,” reveals Baumann. “Its designers didn’t like what didn’t fit world. NoSQL ultimately is a response to this, and array engines are part of it.” Left alone, data centers and companies created homegrown solutions which typically ended up in silo solutions and whole architecture is crafted from scratch.

“rasdaman fast. Adaptive data partitioning and parallelization is done individually for each incoming query, as opposed to static parallelization like in Spark,” describes Baumann. And confronts us with a demo where we see a Terabyte analyzed in less than 100 milliseconds.

With rasdaman’s unique adaptive data partitioning and parallelization, data cubes are analyzed and combined in a straightforward and ultrafast manner.

How rasdaman does this job? “The whole architecture is crafted from scratch, optimized for array handling,” outlines Baumann. There is a series of strong optimizations which altogether make rasdaman fast. Adaptive data partitioning and distribution is one element, augmented with effective compression on dense and sparse data cubes; intelligent processing utilizes all silicon it can get hold of, within and across nodes and even data centers while respecting security. “Query optimization and parallelization is done individually for each incoming query, as opposed to static parallelization like in Spark,”

And how do data get in? Ingestion pipelines are configured quickly – or you avoid copying at all, as rasdaman can be adjusted to tap directly into any preexisting archive structure. “This allows us to be uniquely customer oriented and fast while accommodating any change request swiftly,” says Baumann. “Customers deploy rasdaman as a cost-effective solution for value-adding services, essentially enabling users to build their own data product on the go,” he adds.

This lead got rasdaman on the radar of standardization bodies. ISO is finalizing an SQL extension called MDA (for Multi-Dimensional Arrays) – using rasdaman as its blueprint. “In 2017, ISO-SQL-MDA will become the standard for large-scale array analytics, and rasdaman will be the first fully scalable implementation,” notes Baumann.

“The net effect is a tremendous boost in productivity of analysts, scientists, and engineers, achieved on commodity hardware and clouds,” concludes Baumann. We follow this, and wholeheartedly include this pick in our list of 100 Most Promising Big Data technologies.