

Tokutek promises performance boost for MySQL with new storage engine

Analyst: Matt Aslett

Tokutek is a new startup preparing for the general release of its first product, TokuDB for **MySQL**, which is designed to boost the performance and scalability of the open source database by adding Fractal Tree indexing capabilities. TokuDB is a new storage engine for MySQL that is designed to improve its ability to capture and query information simultaneously, and it will be targeted at the needs of e-commerce, advertising, marketing, vertical search and social networking Web applications.

The 451 Take

Tokutek has chosen a demanding market niche to target with TokuDB for MySQL, although it could be a lucrative one given that the majority of MySQL's adoption has been for Web applications, but it lacks the performance characteristics required for the most demanding applications. With an experienced management team and the brains behind Fractal Tree indexing behind it, the company also boasts a strong executive team. The big question surrounding Tokutek is whether Oracle, assuming it closes its acquisition of Sun, will continue to encourage third-party storage engines for the MySQL database. Tokutek's strategy does not depend on Oracle's approval, but it would be much easier with it than without it.

Context

Tokutek was founded in 2006 to commercialize the research work of its founders into Fractal Tree indexing, an alternative method of indexing data within a database that is designed to improve the performance of the database by overcoming disk input/output (I/O) limitations. The Fractal Tree indexing approach was originally developed at **MIT**, **Stony Brook University** and **Rutgers** by researchers such as cofounders Bradley Kuszmaul, Martin Farach-Colton and Michael Bender, who are respectively the company's chief architect, chief technology officer and vice president of engineering. The fourth cofounder is chief operating officer David Wells, previously of **Giganet** and **Thinking Machines**, while president and chief executive officer John Partridge recently joined the team from complex event processing vendor **StreamBase Systems**, which he cofounded with Michael Stonebraker.

The company has a dozen employees in offices in New York and Lexington, Massachusetts, and has one paying customer up and running, although TokuDB for MySQL will not reach

general availability until later in May. It also claims to have a further dozen customers in its early-release program. An undisclosed amount of seed funding was provided in late 2008 by **General Catalyst Partners** and **Atlas Venture**, and the company is currently exploring its options for raising its series A funding round.

Technology

The key to Tokutek's business strategy is the Fractal Tree indexing algorithm, which is specifically designed to overcome disk I/O limitations that impact B-tree indexes, the indexing method used by the majority of existing database products. Essentially, Fractal Trees convert random I/O into sequential I/O to improve data indexing, insertion and deletion rates, as well as improving query performance.

Fractal Tree indexing is based on the theory of cache-oblivious algorithms (COAs), which are designed to remove the dependency on block sizes in system memory. This has enabled the Fractal Tree indexing algorithm itself, as well as improving data-compression capabilities, which also improves query performance. The company also claims that COAs remove the need for tuning, since Fractal Trees are oblivious to factors such as data characteristics, workload and hardware configuration.

Products

The Fractal Tree indexing technology has been productized in the form of TokuDB for MySQL. The first commercial release is numbered version 2.0, and will be generally available in May under annual subscription or perpetual licensing.

It acts as a storage engine for the MySQL Enterprise 5.1 database and improves performance by 10-50 times compared to existing MySQL storage engines InnoDB and MyISAM, according to Tokutek's benchmark tests. It supports features such as compression, row-level locking and ACID (atomicity, consistency, isolation, durability)-compliant transactions and offers compatibility with existing MySQL-based applications and SQL code. The product is available for both 32- and 64-bit Linux and Windows, initially on a single server, and the company is working on multi-server support.

Strategy

Tokutek's business strategy is focused on delivering additional performance capabilities to the estimated 11 million installations of MySQL worldwide, providing a product for demanding Web applications such as social-networking sites and transactional Web applications that will enable users to query live data without the need for offline batch processing. The company has a commercial relationship with **Sun Microsystems**, and TokuDB is currently licensed to work with the MySQL Enterprise subscription version of the database. This should enable the company to encourage MySQL Community users to become MySQL Enterprise paying customers and enable a symbiotic relationship with Sun Microsystems to develop, such as has been established by **Infobright**, which has a similar strategy in the data warehousing market. Tokutek is initially taking a direct approach to selling TokuDB for MySQL since it seeks to get early adopters up and running and to

benefit from the experience of live customer deployments. In the long run, the strategy is to exploit indirect sales channels in order to reach as wide an audience as possible.

Competition

Since TokuDB offers a new approach to improving the performance of MySQL, there is some justification for the company's claim that its biggest competition is from MySQL users attempting to improve performance themselves by working around the limitations of the MyISAM and InnoDB storage engines. InnoDB is owned by **Oracle**, which will soon also own the MySQL database itself when it closes its acquisition of Sun Microsystems. While we believe Oracle will continue to develop MySQL and strengthen its position as a database for Web applications, we are not convinced that it will continue to support commercial relationships with third-parties that better enable MySQL to compete with the Oracle Database, so Tokutek's current deal with Sun may be called into question.

Tokutek believes that its strategy of targeting high-end Web applications might work in its favor – especially if it enables MySQL to better compete with **Microsoft's** SQL Server – and that the acquisition by Oracle will be a positive event in the long run. In any case, much depends on Oracle's attitude.

A number of other vendors have created storage engines for MySQL that target data warehousing (**KickFire** and Infobright) and clustering (**ScaleDB**), which we do not believe are directly competitive. The data-caching appliances recently launched by **Gear6**, **Schooner Information Technology** and **Virident Systems** provide a hardware-based approach to the problem of improving MySQL performance.

SWOT analysis

Strengths	Weaknesses
Tokutek boasts the brains trust when it comes to Fractal Tree indexing and cache-oblivious algorithms.	The company will always have to explain to potential customers why a new approach to database indexing is required – an opportunity and a potential barrier to adoption.
Opportunities	Threats
There are an estimated 11 million MySQL users to be targeted, many of which are going to be tempted by the claimed performance gains.	It remains to be seen whether Oracle will invest in the MySQL database, let alone support the use of third-party storage engines.

Reproduced by permission of The 451 Group; copyright 2009. This report was originally published within The 451 Group's Market Insight Service.

For additional information on The 451 Group or to apply for trial access, go to: www.the451group.com