



# Analytics

Analytics applications are rapidly becoming the key applications for Big Data workloads. Analytics applications address the large data-sets that are generated by transactional processing to find the patterns in the data that can be leveraged to take decisive action in a fast-moving marketplace.

*“Analytics is rapidly becoming the key application for Big Data workloads”*

Jean S. Bozman

Analytics is an umbrella term used to describe a number of specific workloads that are widely deployed within financial services companies. These workloads are needed to cope with the data tsunami that is hitting financial services firms—generated from a variety of data sources. Customers must quickly find the patterns in the data in order to make accurate and timely business decisions.

MongoDB is a leading NoSQL database that allows customers to identify the key data-points in extremely large multi-terabyte (TB) datasets. SanDisk® evaluated a scalable MongoDB workload, to determine the impact of running these workloads on flash-enabled servers based on SanDisk® solid-state drives (SSDs).

and videos can be viewed and analyzed as easily as more-structured, traditional employee records and customer records.

MongoDB, with its dynamic query language and its built-in Aggregation Framework, are software tools that allow customers to group, shape and analyze data that ranges up to 1TB, or more, of storage capacity. Its methods are different from those of the SQL-based relational databases, and so its database engine is adaptable to changing business conditions – and to input that takes varied forms.

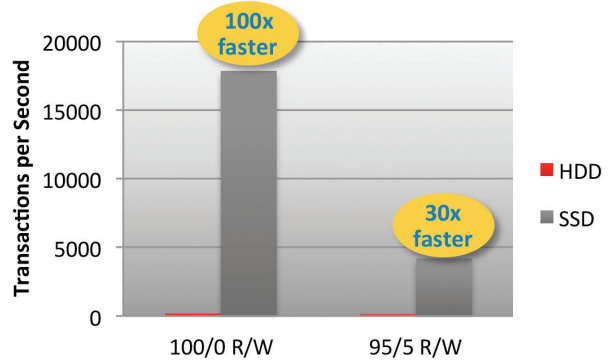
Customers use MongoDB for many aspects of financial applications, including risk-analytics and reporting; distribution and synchronization of data across

## MongoDB

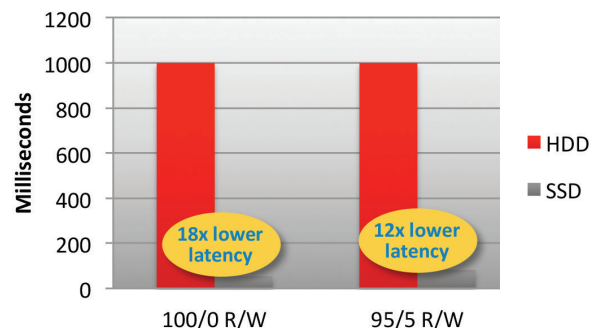
MongoDB is an open source NoSQL database. Unlike traditional relational databases, MongoDB enables unstructured or schema-less databases. As applications evolve, many types of data - including photos, videos, audio files and other types of data - can be analyzed more efficiently compared to traditional SQL databases.

The business benefits of this approach are clear: NoSQL databases are highly scalable - and can handle large data files. Importantly, they can address multiple data-types, including structured, semi-structured and unstructured data—with all data stored in a single data repository. This means that images, photos

MongoDB Throughput



MongoDB Latency (99%)



geographically separated sites; market data management; and aggregation and storing of trading information. "Auto-sharding" of data makes it possible to distribute the data under MongoDB management across many hundreds of physical servers. Those servers can reside within traditional enterprise data centers—or they can leverage cloud-style scale-out compute nodes.

## Testing MongoDB

SanDisk tested MongoDB using the Yahoo! Cloud Serving Benchmark (YCSB) on a flash-enabled server. When running in a 100% read mode, performance for the flash-enabled servers with SSDs was 100 times faster than hard-disk drives (HDDs), with latency that was 18 times lower than the HDDs tested. When running in a mode that was 95% read, and 5% write, the performance was 30 times faster, with 12 times lower latency, than was seen for the HDDs tested.

In this test, one server node was used, allowing all data to be loaded onto a single server. The benchmark used was the YCSB (Yahoo! Cloud Serving Benchmark), running against a MongoDB database with 1TB of data.

## Advantages of Flash

Flash technology accelerates the performance of servers running MongoDB. Object-oriented NoSQL databases like MongoDB, when deployed for financial workloads, handle financial data in new ways, combining multiple data types in the storage and retrieval of market data. Customers who acquire flash-enabled servers will see performance benefits, with dramatically reduced latency for I/O – improving the time-to-results and speeding time-to-decision.

Using SSDs brings a number of advantages to customers in terms of CapEx and OpEx costs. First, fewer SSDs will be needed to deliver the same

outcome as HDDs. The performance characteristics of SSDs make them much less subject to the response time issues that affect HDDs. Operational expenditures are less than for HDDs, because the number of servers required to run the same workload within the data center is less.

With fewer drives and fewer systems required, power and cooling costs are lower than for an HDD-based server solution. SSDs save time and money, because they reduce latency, while improving quality of service (QoS). And with no moving parts, SSDs don't experience failures due to mechanical parts wearing out. In terms of high availability for mission-critical data, SSDs' non-volatile memory preserves data, reducing time to recovery from outages.

## Summary

The digital universe is expanding – creating new demands on those who must analyze it, and take actions based on the analytics results. SanDisk SSDs can be put into use through in-place swap-outs for HDDs or as built-in devices in systems vendor products.

For technology refresh, SanDisk SSDs plug into standardized interfaces for SAS, SATA, and PCIe directly—so they fit into existing data center systems with no disruption of the infrastructure. New deployments bring the benefits of flash technology, as well. Solid-state drives are built into the servers being acquired from major systems vendors worldwide. SanDisk SSDs are being shipped by 6 of the top 7 server and storage OEMs worldwide.

Fast-paced financial markets value technology that allows them to analyze transactional data—and to predict where the market is heading. Solid-state drives provide rapid processing, and shorter time-frames to meet customers' quick decision horizons.

---

### Contact information

ESS.channelsales@sandisk.com  
ESS.OEMsales@sandisk.com  
businesspartners@sandisk.co

### Western Digital Technologies, Inc.

951 SanDisk Drive  
Milpitas, CA 95035-7933, USA  
T: 1-866-744-2165

Western Digital Technologies, Inc. is the seller of record and licensee in the Americas of SanDisk® products.

For more information, please visit:

[www.sandisk.com/enterprise](http://www.sandisk.com/enterprise)

---

# SanDisk®

a Western Digital brand

At SanDisk, we're expanding the possibilities of data storage. For more than 25 years, SanDisk's ideas have helped transform the industry, delivering next generation storage solutions for consumers and businesses around the globe.