CASSANDRA

SOLUTION BRIEF

The SanDisk® Solution for Scaling Cassandra

SanDisk's heritage of innovation in flash storage technologies has led to the development of our flash-optimized software technology that enables applications to capitalize on the benefits of flash scalability and performance. Multi-core processors allow parallel execution of multiple threads. But Cassandra was not written to exploit multi-core processors or flash memory. Cores go unused. Flash IOPS go unexploited. Precious IT budget is inefficiently used.

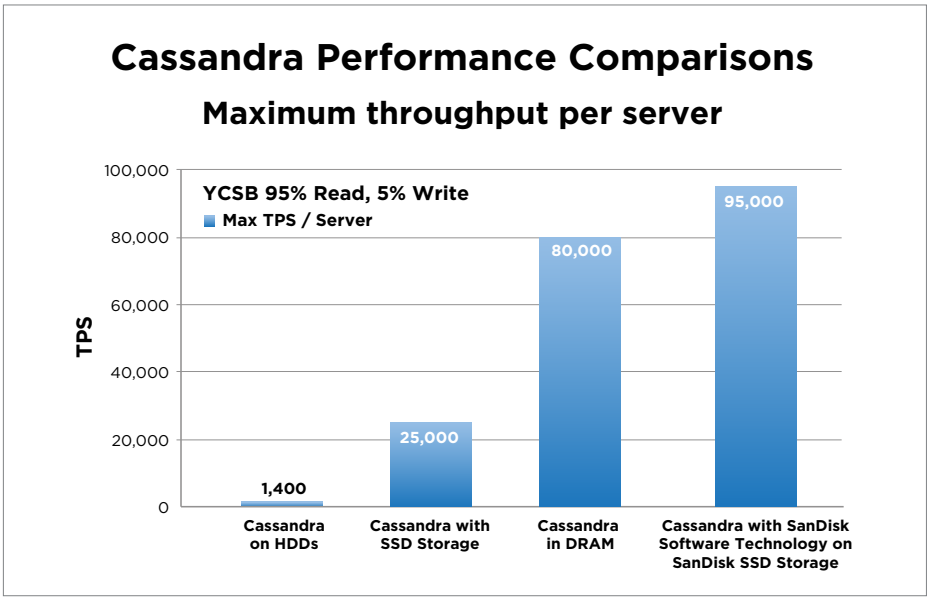
The SanDisk Solution is unique in its ability to fully exploit flash memory and multi-core servers to provide optimal performance and server consolidation.

SanDisk Software Technology delivers cost-effective Cassandra scalability

SanDisk Software Technology is a software development kit (SDK) that offers cost-effective application scaling by enabling flash memory to exploit the capability of multi-core processors parallel execution of multiple threads providing DRAM-like speed at a lower cost. Greater overall performance can be achieved while minimizing infrastructure investment, eliminating the need for rack after rack of servers, or the need to keep data in DRAM to get great performance. SanDisk Software Technology delivers higher availability for mission-critical workloads with synchronous replication at very high throughput.

The SanDisk Solution—SanDisk Software Technology and Cassandra on SanDisk SSDs

To demonstrate the multiple benefits of SanDisk Software Technology's ability to scale Cassandra, we conducted an internal test using Yahoo! Cloud Serving Benchmark (YCSB). The benchmark configuration used consisted of transactions which were 95% reads and 5% writes. Each row of the database had 10 fields comprising 1kByte rows. Write transactions updated one field, while read transactions read all fields, both with uniform distribution. The system test was set-up with a server that ran the Cassandra database and that was driven by client servers running the YCSB workload generator. What varied were



SanDisk®
a Western Digital brand

the storage configurations (HDDs, SSDs, DRAM and the SanDisk Solution) and the Cassandra software between the runs.

Maximize cost savings with server consolidation and lower cost per transaction

In addition to TPS results, TCO calculations were conducted to show server consolidation savings and $/TPS using operating and capital cost required to execute 80,000 TPS workload and 3.1 TB database for 3 years. The 3-year operating expense calculations are based on standard server provider rates for racks per month, power per month, pipe per month and the annual software license cost for SanDisk Software Technology. Note that the $/TPS = (3 year CapEx and 3 year OpEx) / 80,000 TPS. The HDDs were throughput limited (2.4 TPS/server) and require sharding to achieve the 80K TPS requirements. The DRAM was capacity limited at 600GB/server and required several servers to achieve the 3.1 TB database requirement. The competitive SSDs were performance throughput constrained requiring numerous servers to achieve the 80K TPS. The SanDisk Solution was able to achieve throughput and capacity with just one server. The SanDisk Solution proved to be 16x cheaper than deploying HDDs, 4.8x cheaper than deploying large DRAM servers, and 3.3x cheaper than just replacing HDDs with SSDs.

Cassandra Demo Configuration	Servers Required	Cost per Transaction
Cassandra on HDDs	34	$8.44
Cassandra in DRAM	6	$2.49
Cassandra on competitive SSDs	4	$1.70
The SanDisk Solution	1	$0.51

The SanDisk Solution delivers optimal Cassandra performance

This benchmark demonstrates how the SanDisk Solution provides superior performance per server and the most cost-effective data center deployment alternative for a Cassandra database relative to servers with HDDs, replacing HDDs with SSD storage or using large DRAM in servers. The SanDisk Solution is unique in its ability to fully exploit flash memory and multi-core servers to provide optimal performance and server consolidation. SanDisk Software Technology enables multi-threaded parallel data access to flash memory that can deliver performance comparable to sequential access from DRAM. SanDisk Software Technology is designed to allow applications to access data in flash at DRAM-like performance maximizing Cassandra's ability to deliver cost-effective and persistent performance, scalability, and availability for mission-critical workloads. SanDisk is committed to providing innovative flash technology solutions that scale to meet your business needs.

SanDisk is committed to enterprise data center solutions

SanDisk has been transforming flash memory storage for 25 years. Being a vertically integrated producer of flash, from fabrication to our broad portfolio of products, we maintain a consistent commitment to advancing data storage technologies. As a global leader in flash memory storage solutions, we invest over $0.7 billion annually in research and development to provide scalable enterprise hardware and software solutions that are accessible and compatible with existing infrastructures. These solutions empower customers with options best suited for their specific workloads. Businesses can create more efficient and reliable storage that saves money, space and power. SanDisk provides flash technology solutions that help businesses anticipate and respond rapidly to changing needs to maximize competitive advantages and IT investments.

Contact Information

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

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

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.