Increasing Oracle NoSQL Database Performance Using SanDisk Solid State Drives (SSDs)

April 2015
Table of Contents

Executive Summary .............................................................................................................................................. 3
SanDisk’s CloudSpeed SSD .................................................................................................................................... 3
Oracle NoSQL Database ...................................................................................................................................... 4
YCSB benchmark .................................................................................................................................................. 5
Test Environment ................................................................................................................................................... 6
  Technical Component Specifications .................................................................................................................... 6
  Test Workloads .................................................................................................................................................... 7
Oracle NoSQL Database Configuration .................................................................................................................. 7
Results Summary ................................................................................................................................................... 8
  Workload A: Update Heavy .................................................................................................................................. 8
  Workload B: Read Heavy ...................................................................................................................................... 9
  Workload C: Read Only ....................................................................................................................................... 9
Conclusion ............................................................................................................................................................ 10
References .............................................................................................................................................................. 10
Executive Summary

The rapid growth of Big Data, spurred by business-generated operational data, and new data from cameras, smartphones, sensors and the Internet of Things (IoT), threatens to swamp organizations struggling to analyze it, and to leverage it for business advantages. The size and scope of this data explosion—the amount of data generated in an “internet minute”—is challenging traditional data storage practices, in terms of scale and cost. This is the opportunity that created a new type of “NoSQL” database system that can store, manage and analyze huge amounts of data—including new data types such as video and images—on a distributed system architecture.

The Oracle NoSQL database belongs to the same family of NoSQL database systems. It is based on a distributed key-value architecture that offers enterprise-grade reliability and availability features. As organizations deploy NoSQL databases to deal with vast amounts of data, performance remains the key criteria to drive business objective scenarios. These include high-throughput event processing, real time data access, and support for high-volume, web-based commerce.

SanDisk has a rich portfolio of SSD products that support these types of data processing, both for traditional Relational Database Management Systems (RDBMS) and new-generation NoSQL databases.

SanDisk tested NoSQL databases running on SSDs, to explore the performance advantages of SSDs, when compared with spinning hard disk drives (HDDs) for NoSQL databases. To do this, SanDisk used the standard Yahoo! Cloud Serving Benchmark (YCSB), working with an Oracle NoSQL database on a single server node. The testing compared the results when running on HDDs compared to SanDisk's CloudSpeed™ SATA (Serial ATA Interface) SSDs.

Benchmark results discussed in this white paper show that the Oracle NoSQL database, running on servers using SanDisk’s CloudSpeed SSDs, delivered much higher performance compared to HDDs. Clearly, this performance advantage could be leveraged to provide business benefits to customers.

SanDisk’s CloudSpeed SSD

SanDisk is a leader in flash storage solutions. Its solid state drives (SSDs) support the megatrends in the industry that are driving new deployments, including Cloud, Big Data/Analytics, Mobility and Social Media. NoSQL database applications drive many of these new workloads—and they must provide optimized performance to support timely business results.

The CloudSpeed SATA SSD product family offers a full portfolio of options, ranging from entry-level to enterprise-grade high performance requirements. SanDisk’s portfolio of CloudSpeed SATA SSDs is optimized for read-intensive, mixed-use, and write-intensive application workloads in enterprise and cloud computing environments. The CloudSpeed drives, leveraging a 6Gb/s SATA interface, provide data transfer rates up to 450/400 MB/s for sequential read/write operations—and they support performance up to 80K/25K IOPS random read/writes with an average latency of less than 2 milliseconds.

CloudSpeed SATA SSDs are available in storage capacities ranging from 100GB to 960GB. These drives come with a warranty of 3 to 5 years, and support high reliability with a MTBF (Mean Time Between Failure) of 2 to 2.5 million hours.

For additional product details and SanDisk’s portfolio of enterprise products, please refer to the SanDisk website in the References section.
**Oracle NoSQL Database**

The Oracle NoSQL database employs a simple key-value data model, and a non-SQL method of accessing and querying the database. It is used to store non-relational data, such as images, video, and web-based content that is not typically stored in SQL-based relational databases. Importantly, customers can leverage the Oracle database operating concepts they already know to understand Oracle NoSQL database operations.

SQL-based Oracle relational databases serve data requests by using an index key of the stored tables. In an Oracle NoSQL database, data is stored as key-value pair. The Oracle NoSQL database serves up data requests by using a primary key that is associated with key-value pair. The data store where the key-value is stored is referred to as KVStore. It is a collection of storage nodes that host a set of replication nodes.

Replication nodes are organized into shards. A shard contains a single replication node, called the master node, which is responsible for performing database writes, as well as one or more read-only replicas of the database. The master node copies all writes to the replicas. These replicas are used to serve read-only operations. A pictorial representation of Oracle NoSQL storage topology is shown in Figure 1.

Oracle NoSQL server offers full create, read, update and delete operations. It provides flexibility in selecting the consistency models, based on the business requirements. It supports four different consistency models, starting with “no consistency”, “time-based”, “version-based” and finally “absolute consistency.” This flexibility is achieved while still meeting the application latency and scalability requirements.

Oracle NoSQL provides an “administration service” that can be accessed both from either the command line interface (CLI) and or a web-based console. This administration service can be used to administer, manage, and monitor NoSQL operations.

Some of the important Oracle NoSQL features are:

- **Data Model:** Key-value, Java Script Object Notification (JSON), SQL-like tables
- **ACID (Atomicity, Consistency, Isolation and Durability) and BASE (Basically Available, Soft State, Eventually Consistent) transactions**
- **Management:** Easier cluster management and online rolling upgrade option
- **Core server features:** Secondary indices, large objects, availability zone (aka data centers), security
- **Integration:** Oracle database external tables, Oracle Coherence, Oracle event processing, Oracle Enterprise Manager, Oracle Golden Gate, Apache Hadoop/MapReduce, Apache Spark, Apache Hive
The Yahoo! Cloud Serving Benchmark (YCSB) is a standard benchmark framework for evaluating performance of cloud data serving systems like MongoDB, Cassandra, and Oracle NoSQL. The framework consists of a workload generating client and a package of standard workloads.

The performance section of the YCSB benchmark focuses on measuring the throughput of the system for defined latency. The scalability section of the benchmark focuses on ability to scale elastically so that these systems can handle more load, as demands grow, as applications add more features, and as more business users demand data services.

The YCSB benchmark also provides workload distribution options based on how real-world applications run, over time, such as insert/update/scan operations acting on random set of data.

There are two major YCSB workload distribution options:

- **Uniform:** This benchmark option is based on the assumption that all the records in the database are uniformly accessed.
- **Zipfian:** This benchmark option is based on the assumption that the popularity of the record being accessed (e.g., the results of the World Soccer finals) is being accessed more often than other records in the database.
Table 1 shows the workload types that are used for this benchmark:

<table>
<thead>
<tr>
<th>Workload</th>
<th>Operations</th>
<th>Record Selection/Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workload A: Update Heavy</td>
<td>Read: 50%</td>
<td>Uniform/Zipfian</td>
</tr>
<tr>
<td></td>
<td>Update: 50%</td>
<td></td>
</tr>
<tr>
<td>Workload B: Read Heavy</td>
<td>Read: 95%</td>
<td>Uniform/Zipfian</td>
</tr>
<tr>
<td></td>
<td>Update: 5%</td>
<td></td>
</tr>
<tr>
<td>Workload C: Read Only</td>
<td>Read: 100%</td>
<td>Uniform/Zipfian</td>
</tr>
</tbody>
</table>

Table 1: YCSB Workload Options Used for This Benchmark

Test Environment

The benchmark test environment consists of one Dell PowerEdge R720 two-socket server with a total of 12 Intel® Xeon® E5-2620 processor cores—or 24 logical cores via hyperthreading. The environment also had 96GB RAM for hosting the Oracle NoSQL server and one Dell PowerEdge R720, which serves as a client for the YCSB benchmark tool. A 10GbE network interconnect is used to link the Oracle NoSQL server and the YCSB client. The local storage is varied between HDDs and SSDs. Table 2 below provides complete hardware and software components used for this test environment.

Technical Component Specifications

<table>
<thead>
<tr>
<th>Hardware</th>
<th>Software if applicable</th>
<th>Purpose</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dell PowerEdge R720 Intel® Xeon E5-2620 processor, 2 Socket, 12 physical cores, 96GB RAM</td>
<td>OS: Centos: 5.10</td>
<td>Oracle NoSQL server</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Oracle NOSQL Database: 12CR1: kv-ce-3.0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dell PowerEdge R720 Intel Xeon E5-2620 processor, 2 Socket, 12 physical cores, 16GB RAM</td>
<td>OS: Centos: 5.10</td>
<td>YCSB Client</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>YCSB: 0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500GB 7.2K RPM Dell SATA HDDs</td>
<td>Used as Just a Bunch of Disks (JBODs)</td>
<td>Data node drives</td>
<td>6</td>
</tr>
<tr>
<td>480GB CloudSpeed SATA SSDs</td>
<td>Used as Just a Bunch of Disks (JBODs)</td>
<td>Data node drives</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 2: Benchmark Hardware and Software components

Figure 2: Testing Environment
Test Workloads

The primary objective of these benchmark tests is to explore the advantage of using SanDisk SSDs for an Oracle NoSQL database store, and to provide performance data points for SSDs and HDDs. This benchmark consists of a single-node Oracle NoSQL database, with the standard YCSB benchmark workload types A, B and C and data set is varied from 32GB to 128GB.

- The YCSB workload types are based on the percentage of reads and writes, as follows:
  - **Workload A**: Update Heavy: 50% update / 50% read
  - **Workload B**: Read Heavy: 5% update / 95% read
  - **Workload C**: Read Only: 100% read-only

- The YCSB default data size: 1 KB records (1 field, 1000 bytes each, plus key)

- The data set sizes are as follows:
  - **Data Set 1**: 32GB
  - **Data Set 2**: 128GB

- The YCSB workload distribution types are as follows:
  - **Uniform**: All database records are uniformly accessed
  - **Zipfian**: Few records in the database are accessed more often than other records

- All the above configurations is switched between
  - Hard drives
  - SanDisk CloudSpeed SATA SSDs

All of these workload variations resulted in 24 different tests for this benchmark.

Oracle NoSQL Database Configuration

Oracle NoSQL Database default configurations that were used during the test, and all of the workloads, are all running on the same Oracle NoSQL configurations. Because this testing is about performance on a single node server, one storage node was created, containing one data shard with a replication factor set to one.
Results Summary
The YCSB benchmark was conducted on the 24 configuration testing types, and its throughput and latency test results were collected for detailed analysis and reporting.

Workload A: Update Heavy
Workload A consists of an equal mix of 50% reads and 50% updates. It is a write-intensive workload and we tested the data storage systems to determine the workload throughput and latency the drive write throughput and latency. Results of this test are shown in Figure 3.

As expected, when a 32GB data set being tested can completely reside in system memory (DRAM) there is minimal throughput advantage to using SSDs over HDDs for write-intensive workloads. However, when the size of the data set size increases beyond the capacity of available DRAM, a large performance benefit in throughput and latency is realized—as shown by the results for the 128GB data set—for the Uniform and Zipfian workload types. For the Zipfian workload, a 20x performance increase was measured, while for the Uniform workload a 30x performance boost was observed over results using HDDs.

As the dataset was increased to 128GB, HDD performance was 1,537 operations per second, while SSD performance was 48,975 operations per second. As the dataset was increased from 32GB to 128GB, the SSD latency increased marginally from 9 to 15 milliseconds, while on HDDs it increased from 237 to 460 milliseconds. These high throughput numbers are an important factor for large data sets, especially for those involving high update operations, because more business operations can be accomplished in a given period of time.
Workload B: Read Heavy

Workload B involves 5% updates and 95% read operations, representing mixed-workload scenarios. Results of this test, as shown in Figure 4, are that SSDs provide exceptional performance for mixed workloads compared to HDDs.

Consistent with the results from the previous workload, when the data set being tested is entirely resident in system memory, little throughput advantage is seen when using SSDs over HDDs. Only when the data set becomes larger than available DRAM do the HDD results lag far behind at 1,709 operations per second, while the SSDs deliver 58,386 operations per second. Overall, in this test, SSDs provided over a 30x throughput performance advantage compared to HDDs. This will directly translate into efficient utilization of enterprise applications, delivering business results.

Workload C: Read Only

In our third test, throughput results similar to the previous two—when the data set resides in system memory—are observed. As expected, when the data set size increases beyond what will completely reside in system memory, significant advantages in throughput emerge. In this testing of 128GB databases, SSDs provided a 25x advantage in throughput and a 13x to 20x advantage in latency.
Conclusion

The YCSB benchmark test results discussed in this paper demonstrate that running the Oracle NoSQL database on SanDisk SATA CloudSpeed SSDs provided significantly higher performance levels than running the same workload on HDDs—when the database being analyzed exceeds the available system DRAM for it, an extremely common occurrence.

In these situations, the SSD-based solution provided, on average, a 20x to 30x performance benefit for read-intensive, mixed and write-intensive workloads. This means that the performance advantage for SSDs can be leveraged for a wide range of enterprise and cloud applications involving Oracle NoSQL databases.

References

1. What Happens in an Internet Minute:

2. Oracle NoSQL Database Architecture:

SanDisk website: www.sandisk.com/enterprise


Yahoo LABS: http://research.yahoo.com/Web_Information_Management/YCSB/
Products, samples and prototypes are subject to update and change for technological and manufacturing purposes.

SanDisk Corporation general policy does not recommend the use of its products in life support applications wherein a failure or malfunction of the product may directly threaten life or injury. Without limitation to the foregoing, SanDisk shall not be liable for any loss, injury or damage caused by use of its products in any of the following applications:

- Special applications such as military related equipment, nuclear reactor control, and aerospace
- Control devices for automotive vehicles, train, ship and traffic equipment
- Safety system for disaster prevention and crime prevention
- Medical-related equipment including medical measurement device

Accordingly, in any use of SanDisk products in life support systems or other applications where failure could cause damage, injury or loss of life, the products should only be incorporated in systems designed with appropriate redundancy, fault tolerant or back-up features. Per SanDisk Terms and Conditions of Sale, the user of SanDisk products in life support or other such applications assumes all risk of such use and agrees to indemnify, defend and hold harmless SanDisk Corporation and its affiliates against all damages.

Security safeguards, by their nature, are capable of circumvention. SanDisk cannot, and does not, guarantee that data will not be accessed by unauthorized persons, and SanDisk disclaims any warranties to that effect to the fullest extent permitted by law.

This document and related material is for information use only and is subject to change without prior notice. SanDisk Corporation assumes no responsibility for any errors that may appear in this document or related material, nor for any damages or claims resulting from the furnishing, performance or use of this document or related material. SanDisk Corporation explicitly disclaims any express and implied warranties and indemnities of any kind that may or could be associated with this document and related material, and any user of this document or related material agrees to such disclaimer as a precondition to receipt and usage hereof. EACH USER OF THIS DOCUMENT EXPRESSLY WAIVES ALL GUARANTIES AND WARRANTIES OF ANY KIND ASSOCIATED WITH THIS DOCUMENT AND/OR RELATED MATERIALS, WHETHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR INFRINGEMENT, TOGETHER WITH ANY LIABILITY OF SANDISK CORPORATION AND ITS AFFILIATES UNDER ANY CONTRACT, NEGLIGENCE, STRICT LIABILITY OR OTHER LEGAL OR EQUITABLE THEORY FOR LOSS OF USE, REVENUE, OR PROFIT OR OTHER INCIDENTAL, PUNITIVE, INDIRECT, SPECIAL OR CONSEQUENTIAL DAMAGES, INCLUDING WITHOUT LIMITATION PHYSICAL INJURY OR DEATH, PROPERTY DAMAGE, LOST DATA, OR COSTS OF PROCUREMENT OF SUBSTITUTE GOODS, TECHNOLOGY OR SERVICES.

No part of this document may be reproduced, transmitted, transcribed, stored in a retrievable manner or translated into any language or computer language, in any form or by any means, electronic, mechanical, magnetic, optical, chemical, manual or otherwise, without the prior written consent of an officer of SanDisk Corporation.

All parts of the SanDisk documentation are protected by copyright law and all rights are reserved.
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.

For more information, please visit:
www.sandisk.com/enterprise