

SOLID STATE STORAGE FOR DATA WAREHOUSING

By Richard Winter

Summary

Solid state devices (SSD) provide online data storage for computer systems without the mechanical moving parts of hard disk drives (HDD). SSDs have recently become practical for enterprise use in data warehouses, delivering large performance advantages at the device level. While individual solid state devices cost more than hard disk drives with the same storage capacity, the high performance and higher throughput of SSDs can make it practical to replace multiple HDDs with one SSD in heavily used data warehouses. In addition, SSDs have now increased in capacity and price/performance.

SSDs vary considerably in their design, operating characteristics and operating lifespan. In addition, the customer advantages actually realized with SSDs depend heavily on the design of the data warehouse software, storage enclosures, controllers and servers with which they are used.

This WinterCorp Executive Report discusses key issues in the design and use of SSDs in enterprise scale data warehousing. This report focuses on particular products which combine to deliver the desired advantages: the SSDs, manufactured and marketed as enterprise flash drives by [SanDisk](#); storage enclosures and controllers provided by [NetApp Inc.](#); and, the data warehouse platforms provided by [Teradata Corporation](#) that incorporate these storage products.

Enterprise data warehouse solutions from Teradata that employ SSDs are able to replace up to 18 hard disk drives with one solid state drive, while delivering equivalent performance on a typical data warehouse workload. Depending on the configuration and the type of data warehouse platform, the benefit to the customer of using the SSD's is then either much higher performance overall; higher performance on hot data; overall system energy and data center space savings; lower total cost of operation; or a combination of these.

The development of this report by the independent expert WinterCorp has been sponsored by NetApp Inc., which integrates the SanDisk drives into storage systems and provides them to Teradata. A sidebar on page 3 describes the research methodology used.

Hard Disk Drives (HDDs)

The standard solution for decades for online data storage in enterprise systems has been the hard disk drive (HDD). In the remarkable evolution of HDDs since their introduction in the nineteen fifties, the standard enterprise product has shrunk from the size of a washing machine to a box that is 4 inches by 5.75 inches by 1 inch—this is the



size of the standard enclosure for a 3.5 inch diameter disk. Storage capacities per drive have increased by a factor of roughly one million over the same period.

HDD enterprise drive storage capacities now range up to 2 TB and typically double every 12 to 18 months, while maintaining a stable or slowly declining price per drive.

But HDDs are mechanical devices with two motors. They store the data on platters that spin at up to 15,000 rpm. In addition, HDDs reading and writing data from random locations must typically move an electro-mechanical arm to a different physical position each time—an operation that requires an average of 2 milliseconds (ms) in the highest performance devices. Neither the rotational speed nor the average time to access data on disk has improved much in the last few decades. Neither is expected to increase rapidly in the next several years.



Figure 1: Inner Workings of a Hard Disk Drive (HDD)

The result is that HDDs are wonderfully cost efficient for the *storage* of data, but—relative to modern processors—are now too slow for *access* to data in many cases. Today’s highest performance enterprise quality HDDs perform about 200 data reads or writes per second when accessing data at random locations in the device.

In many data warehouses, the speed at which the storage system can read and write data—particularly the speed at which it can read data from random locations—has become the limiting factor in system performance and throughput. Every 12 to 18 months, processor speeds double; main memory sizes double; data volumes increase by 100% or more; HDD capacities double; BUT, data access times on HDDs stay pretty much the same. So, not only is HDD data access time often the limiting factor on system performance—every year the problem gets worse relative to other factors.

Data warehouse operators have dealt with this problem by buying more HDDs every year. In order to obtain more data input/output performance, they buy more drives. If they fill this ever increasing number of drives with data, in many cases, they find that they have insufficient read/write capacity to service the query needs of their users across this expanded data volume. So, each year, they leave more and more empty space on the drives. This often results in inefficient use of space, electric power, storage capacity and capital investment.

Solid State Storage Devices (SSDs)

Enterprise SSDs inside high performance storage arrays have emerged to address the limitations of HDDs in high-performance applications.

ABOUT WINTERCORP

WinterCorp is an independent consulting firm that specializes in the performance and scalability of terabyte- and petabyte-scale data management systems throughout their lifecycle.

Since our inception in 1992, we have architected many of the world’s largest and most challenging databases in production today. Our consulting services help organizations define business-critical database solutions, select their platforms, engineer their implementations, and manage their growth to optimize business value.

With decades of experience in large-scale database implementations and in-depth knowledge of database products, we deliver unmatched insight into the issues that impede performance and the technologies that enable success.



WinterCorp

245 FIRST STREET,
SUITE 1800
CAMBRIDGE, MA 02145
617-695-1800

visit us at
www.wintercorp.com

Purpose and Methodology for this Report

This WinterCorp Executive Report reviews the SanDisk enterprise SSDs and their use in NetApp E-Series Storage Arrays and Teradata data warehouse platforms.

This report is sponsored by NetApp.

In developing this paper, WinterCorp operated as an independent industry expert, interviewing NetApp, SanDisk and Teradata employees; reviewing product documentation; and critically reviewing product design, measurements and evidence in order to arrive at the descriptions and conclusions presented here.

NetApp was provided an opportunity to comment on the paper with respect to facts.

WinterCorp has final editorial control over the content of this publication.

SSDs store the data on devices with no moving parts, typically using flash memory instead. Flash memory has been in use for many years in consumer devices such as phones, music players and thumb drives. However, the flash memories in these devices are not designed for the rigors of enterprise use.

Enterprise SSDs use a different type of flash memory; they employ sophisticated controllers and algorithms; and, they are designed for higher performance and reliability.

Within the class of drives known as enterprise SSDs, there is considerable variation in the characteristics needed for enterprise class data warehousing. This report focuses on the SSDs of one manufacturer: SanDisk.

SanDisk SSDs



Figure 2:
Interior of a SanDisk Solid State Device (SSD)

For several years, SanDisk has been pursuing a long term strategy to produce a particular type of SSD that is distinctively suited to enterprise data management applications. In particular, the SanDisk drives inside NetApp storage arrays use:

- SLC flash (Single-Level Cell) rather than the MLC (Multi-Level Cell) flash used in consumer devices. This increases the stability of data in storage by a factor of ten;
- A rigorous process of testing and selection of chips to be used, resulting in much lower hardware level error and wear rates than are typically accepted, even in “enterprise SSDs”;
- Intensive error correction strategies designed specifically for NAND flash media. Most non-SanDisk SSDs simply employ error correction codes that were designed for HDDs, which have a different hardware level error detection pattern; and,
- Proprietary algorithms and software in the flash controller to optimize performance and reliability in enterprise applications. In particular, these algorithms enable the device to perform more reads per second and to minimize wear resulting from repeated writes to the same chip.

As a result of these techniques, and others distinctive to SanDisk, the SSD drives they provide are particularly well suited to the active enterprise data warehouse platforms. SSD reliability and life span have increased greatly over the SSD’s previously available in the industry.

NetApp E-Series Storage Arrays

NetApp Corporation has incorporated the SanDisk SSDs into storage arrays designed to leverage their high performance. Thus, NetApp has designed its array controllers to respond effectively to the rapid response times—a few microseconds—and higher data transfer rates of SSDs. In addition, NetApp supports RAID on SSDs. For example, in RAID1, the NetApp array controller automatically maintains a twin of each drive. Then, if a drive fails, the same data is available on its twin, and the data warehouse continues operating normally, using only the twin. When the failed drive is replaced, data on the twin is automatically copied in the background to the new drive, restoring the replicated status of the drive.

Implementing RAID1 on SSDs requires a high performance architecture for the storage array—an architecture that NetApp provides in its E-Series line of storage systems. In contrast to storage architectures that use SSDs but do not support RAID, the E-Series systems provide much higher reliability at the storage system level, complementing the high reliability of the SanDisk technology at the drive level.

Use of NetApp Storage arrays with SanDisk SSDs within Teradata Data Warehouse Platforms

Teradata has approached the development of its data warehouse platforms with a distinctive philosophy concerning data storage, integrated with its massively parallel (MPP) architecture.

First, the Teradata system relieves the customer of virtually all decisions with respect to the placement and management of data in storage. Teradata automatically determines where to place data within the storage system. Teradata makes this decision partly on the basis of data “temperature.” That is hot data (data which is accessed frequently) is placed by the system on faster storage devices, and cool data is placed on slower storage devices. In a system that contains both SSDs and HDDs, hot data will migrate to the SSDs and cool data will migrate to the HDDs. Within HDDs, some locations on the disk (e.g., the outer cylinders) transfer data at higher rates and this difference, also, is automatically exploited by the automated storage management of Teradata.

Second, the basic Teradata architecture employs a sophisticated hash based storage scheme for all data objects. This means that individual data objects (e.g., a row or record) can be located very quickly in storage. But, it also means that Teradata reads data more selectively than most other systems and does a large proportion of random disk reads. As a result, the principally random disk read pattern of Teradata benefits disproportionately from SSDs. The SanDisk SSDs used by Teradata produce about 22 times more random reads per second than a high-performance HDD.

Most other data warehouse products are optimized for sequential reads. In sequential read performance, SSDs are only about three times faster than HDDs. As a result, most other data warehouse products will benefit less from the use of SSDs than Teradata.

These two factors have a large impact on the benefit to the customer of Teradata systems that provide a mix of HDDs and SSDs. First, the customer benefits from the performance of the SSDs without any additional administrative burden: hot data automatically migrates to the SSDs. Second, because of the predominately random pattern of accessing data, the benefit from SSDs is typically much larger with Teradata than with other data warehouse platforms.

Teradata’s automatic data placement contrasts with the approach employed on other data warehouse products in which either: (a) the customer must specifically manage the placement of data on SSDs and HDDs; this is a burden, not only because of the difficulty of figuring this out, but also because the temperature of data naturally changes over time; thus, the customer never escapes the ongoing burden of allocating and reallocating data to the SSDs; or (b) only temporary data is placed on the SSDs. The customer’s hottest retained data (for example, the sales of the last 30 days) will be on HDDs where it takes many times longer to retrieve.

Teradata's automatic placement of data according to temperature—branded as Teradata Virtual Storage also differs from another industry approach known as “tiered storage.” In tiered storage, the units migrated are typically gigabyte-scale granules. Data migration occurs at scheduled times, typically daily or weekly, and is driven by complex rules or policies that must be maintained by human storage or system administrators.

In Teradata Virtual Storage, the data is placed automatically by the system on the basis of data temperature, and the granules of data moved by the system are a thousand times smaller (megabyte scale). Further, migration occurs continuously in response to actual user behavior. Teradata Virtual Storage is thus much better suited to the dynamic, online environment of a shared, active data warehouse.

Summary

Enterprise SSDs provide dramatically higher performance in data retrieval and update, when compared to HDDs. Retrieval of a randomly located block of data from storage, in particular, is many times faster with an SSD.

However, SSDs have not been widely used in data warehousing until recently. First, earlier SSDs lacked the storage capacity to have a large impact on enterprise scale data warehousing. Second, they were expensive. Third, most early SSDs, including those used in consumer products, were not as reliable for enterprise use as initially expected or promoted.

Some enterprise SSDs have now matured to the point where they can play a significant role in data warehousing. Such products can be more than ten times as reliable and much higher performance than the “commercial” SSDs used in consumer products, smart phones, tablets and notebook computers. One company that delivers an enterprise class product with suitable, proven characteristics for enterprise data warehousing is SanDisk.

In a data warehouse system, it is not enough to have individual devices—such as the SanDisk SSDs—that provide superior performance and reliability. For the SSDs to actually deliver the increased throughput and reliability to the data warehouse requires an overall storage architecture that supports RAID without bottlenecks and that leverages the high data rates and low access times of the SSDs. The E-Series Storage Arrays by NetApp Corporation are designed and built to address these requirements. When configured with SanDisk SSDs, the E-Series arrays provide increased performance, increased reliability and reduced footprint for equivalent I/O capacity.

Teradata platforms are designed to exploit the E-Series Storage Arrays (containing SanDisk SSDs) effectively. Teradata's system managed storage automatically places hot data on SSDs and cooler data on HDDs. Thus, customers benefit from SSD performance with no added administrative burden. In addition—both Teradata's random placement of data via hashing and Teradata's emphasis on selective (rather than sequential) reading of data for retrieval—result in a large benefit from the use of SSDs. So, while products which favor sequential scanning may not benefit greatly from SSDs, Teradata does.

Teradata now offers a variety of data warehouse platforms, including models that use SSDs entirely; models that use a mix of SSDs and HDDs; and, models that use all HDD storage. This provides users with a range of price/performance tradeoffs. In addition to providing better performance on hot data, the models that include SSDs all provide savings in data center space and energy.

A customer with the average Teradata workload can use up to 18 fewer HDDs for each SSD installed. The relatively higher device level price of the SSDs can thus be recovered via reduced total cost of operation. In addition, replacing multiple HDDs with one SSD greatly reduces the number of moving parts in the system, thus substantially increasing system level reliability.