Spring 2013

Ssd Flash Drives Used to Improve Performance with Clarity Data Warehouse

Douglas Ervin
Regis University

SSD FLASH DRIVES USED TO IMPROVE PERFORMANCE WITH CLARITY DATA WAREHOUSE

A THESIS

SUBMITTED ON 23 OF January, 2013

TO THE DEPARTMENT OF INFORMATION TECHNOLOGY

OF THE SCHOOL OF COMPUTER & INFORMATION SCIENCES

OF REGIS UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF MASTER OF SCIENCE IN DATABASE TECHNOLGY

BY

Douglas Ervin

Douglas Ervin

APPROVALS

Robert T. Mason

Robert Mason, Thesis Advisor

D. M. Likarish

Daniel Likarish

Nancy Birkenheuer

Nancy Birkenheuer

Abstract

Since the introduction of solid-state devices (SSD), both storage area network (SAN) administrators and database administrators (DBA) have imagined the performance gains promised by replacing hard disk drives (HDD). The initial testing in the laboratory did not promise those gains in the real world. The SSD vendors worked between 2007 and 2010 to improve performance, which in industry standard tests showed steady progress. Despite the gains in the laboratory, there were few examples of real world usage particularly in the field of data warehousing. The process of extracting, transforming and loading (ETL) places extreme loads on the ability of the storage device to update data. This paper studies the effect on one such data warehouse.

Acknowledgements

The author wishes to thank Scott Alderman, Director – Sisters of Charity of Leavenworth Health Systems, allowing the use of Clarity Data Warehouse Console and his encouragement of this research paper. My wife, Debbie Ervin, deserves appreciation and acknowledgement due to the supporting my efforts over the two and half years this program has taken along with our son, Chris Norris, and his wife, Kelly Norris.

Table of Contents

SSD FLASH DRIVES USED TO IMPROVE PERFORMANCE WITH CLARITY

Chapter 1 – Introduction

Problem Statement

Solid-state device (SSD) drives promise improved read and write performance over magnetic disk drives. Limitations to performance of magnetic disk drives are the speed at which the disks can spin, and how fast the mechanical read/write heads can move. To overcome the physical limitations, administrators have started using large number of spindles, using only half of the drives capacity, and striping everywhere.

SSD drives started being used in cell phones with the advent of the smart phone, in portable music devices (iPod), document readers (Kindle), and finally as a means of increasing the performance of high end laptops (MacBook Air). System administrators and database administrators have imagined what performance benefits SSD drives would provide on servers.

Finally, several vendors are marketing enterprise level SSD drives to the large database community. For the last ten years, the SSD drive has promised to make the magnetic disk drive obsolete, but it has always been the next great step.

Purpose Statement

This study will explore the performance differences between magnetic disk drives and SSD drives during the extract-transform-load (ETL) process, a complex mixture of sequential reads and writes, and random writes. The goal is to determine if SSD drives will offer improved performance in a normal business situation.

Rational

Agrawal et al. (2008) found that SSD drives have the potential to revolutionize storage systems, particularly with large databases. However, the authors determined the need for further study regarding the internal organization of SSD drives and design choices affecting optimal performance. Agrawal et al. (2008) demonstrated the advantages of SSD drives for writes and reads, showing sequential write latency of 327 µsec and a random write latency of 433 µsec and both sequential reads and random reads at 130 µsec. Narayanan, Thereska, Donnelly, Elnikety, & Rowstron (2009) looked at the problem from the viewpoint of MB per second and in-out operations per second (IOPS). They compared the Memoright MR 25.2 SSD, Seagate Cheetah 10K, Seagate Cheetah 15K, and Seagate Momentus 7200. The Seagate Momentus 7200 served as a reference point, being a high-end consumer drive.

Drive	Read MB/s	Write MB/s	Read IOPS	Write IOPS
Memoright MR 25.2 SSD	121	126	6450	351
Seagate Cheetah 15K	88	85	384	269
Seagate Cheetah 10K	85	84	277	256
Seagate Momentus 7200	64	54	102	118

Table 1 – Partial Storage Device Characteristics

The researchers’ conclusion was that it is not cost effective to replace magnetic disk drives with SSD flash drives at the time of the study (Narayanan et al., 2009).

Schmidt, Ou, & Härder (2009) looked at SSD drives, their favorable IO model, and the total cost of ownership (TCO). The advantages cited were transactional throughput, power consumption, reliability, and scaling. In large storage arrays, the power consumption can be a significant factor in mitigating the TCO. The difference in energy used is almost a factor of 10 in favor of the SSD drives. The authors reliably confirm the superiority of SSD read and sequential write capability but also confirm the problem with random writes. In their experiment, the magnetic disk drive out performed the SSD drive in random writes by a factor of three. Their conclusion is SSD drives are not a solution for all database application but should be considered where performance/$ is not a major consideration and pure performance is.

Lee, Moon, & Park (2009), prototypes of SSD drives manufactured by Samsung Electronics were tested against an enterprise magnetic disk drive: Personal Class SSD Models A and B, and Enterprise Class SSD. The authors found that a single Enterprise Class SSD could outperform a level 0 RAID array with eight-enterprise class 15k-RPM disk drives in terms of transaction throughput, cost effectiveness and energy consumption. The prototype Enterprise Class SSD was able to outperform the magnetic disk drive by a factor of eight for random writes. The authors concluded the SSD drive had advanced to the point where consideration as a storage solution for database applications is reasonable.

Research Question

In the proposed study, the research question is, will the enterprise class SSD drives will improve the average time to perform the extract-transform-load process of the Epic System's Clarity Reporting Repository?

Chapter 2 -Literary Review

The ETL process is a complex mixture of sequential and random writes and reads from the disk array. There are question concerning the suitability of SSD drives in conjunction with relational database management system (RDBMS). The organization of the literature review is from the oldest to the latest available literature emphasizing points of advancement in the technology and areas of continued concern.

Microsoft Research – An early study

Microsoft Research produced one of the earliest papers, which was submitted for publishing in the Proceedings of the USENIX Technical Conference, June 2008 (Agrawal et al., 2008). The study showed SSD drive performance was dependent on data placement, parallelism, write ordering, and workload management. Proper placement of data on the SSD drives was critical, since it affected both load balancing and wear leveling, which is the technique to prevent the tendency of the chip to wear out due to constant writing to the same sector. The earliest SSD drives did not have enough parallel paths for the data to provide optimal performance, and the study recommended additional work in this area. The random write (also called an in-place update) proved to be a matter of serious concern. The logic of write ordering, due to the time needed to read the sector into memory, change the contents in memory, erase the original sector then write the new sector is a major drawback to SSD drives. The study saw some random writes taking 200 µsec for a write to a SSD drive as opposed to 12 µsec for a magnetic disk drive. Workload management concerns how the SSD drive balances the sequential writes (0.4 µsec) and reads (0.2 µsec) with the 200 µsec writes and overall improve performance. Agrawal et al. (2008) determined SSD drives showed promise but were not ready for wholesale adoption.

ACM Queue – Show the Potential Gains

Moshayedi & Wilkison (2008) described the potential benefits of SSD drives over 15,000 RPM magnetic disk drives (HDD). On a one-to-one comparison of a SSD to a HDD, the study determined in terms of in/out operations per second (IOPS) that 30 properly configured HDDs were required to match the performance of a single SSD. In the case study, the article focused on a data center reducing from 9000 HDDs to 300 SSDs and saving 90% of the energy costs due to the conversion. They found SSDs have very fast sequential read and write performance, very fast random read, which helps in fast application load and boot-up times, very poor random write performance (less than one tenth the speed of existing HDDs), and horrible performance under workloads with intermixed reads and writes (Moshayedi & Wilkison, 2008, p. 36). Referring back to Agrawal et al. (2008) the last two points are expressions of write ordering and workload management. The authors concluded while potential gains are possible, a buyer will have to do some research, as not all SSDs are truly enterprise ready.

SIGMOD 2009

Lee, Moon & Park (2009) researched the problems from the aspect of Access Density or the ration of IOPS to the disk capacity. For HDD a drive, as the capacity has gone up the access density has actually dropped, one of the few areas in computers where a measure of performance has decreased over time. One large scale TPC-C (a testing protocol for transactional databases) used 170 HDDs per each of its 64 cores to maximize performance. The study went on to test three prototype Samsung SSD drives, two personal-class drives (32 and 64 GB respectively) and a 128 GB enterprise class drive. The researchers concentrated on the pattern of a write followed by a read, which was determined to be a problem by Moshayedi & Wilkison (2008). The enterprise class SSD performed better than a comparable HDD due to the use of eight-channel parallelism to allow reads to be processed along with the slower write. This coupled with a buffer to queue up requests allowed the SDD surpass the HDD in random write and mixed read-write performance. Cost per GB remained a major barrier to adoption.

Lee, Moon, Park, Kim, & Kim (2009) evaluated different functions performed by a database in terms of their suitability to using SSD drives. The areas studied were transaction log, rollback segments, and temporary table spaces. When a transaction commits, the transaction contents append to the end of the transaction log, a sequential write. Sequential writes are favorable to both HDDs and SSDs, with only an average of one quarter of a spindle turn for the HDD and about 4 msec latency. The SSD is even more impressive at 0.4 msec, issuing an update statement causes the pre-update contents of the row to be written to a rollback segment. The rollback segment allows both for read consistency of the data prior to the commit and for the update statement to roll back to its original value. The rollback segment is a hot spot for concentrated mixed reads and writes, one of the SSDs problem areas. In their study, the researchers found no advantage of the SDD over the HDD. The temporary tablespace is used both for external sorts of data and merges of data from joins of table, another form of sort. In both of these cases, the database writes the data sequentially to disk and then reads the sorted data back to be presented as the result set. The conclusion was a hybrid database (both SSD and HDD), which placed the transaction logs and temporary tablespace on SSD drives would give a measurable performance increase, while there was no gain by moving the rollback segment from a HDD to a SSD.

European Conference on Computer Systems 2009

Narayanan et al. (2009) looked at the wholesale and tiered approach of replacing HDD drives with SSD drives using an improved performance versus a cost/capacity model. The evaluated the Memoright MR-25.2 SSD versus the Seagate Cheetah 10K and Seagate Cheetah 15K. The Memoright MR-25,2 won the IOPS/$ comparison and the Seagates won the GB/$ comparison. The interesting one was the throughput test, MB/s/$, which favored the Seagates due to their cost but in absolute performance the Memoright was superior. Overall, they recommended to not replacing HDDs with SSDs with one exception. If the database reduced the buffer cache in size and used SSD for the buffer cache, it would prove nearly as fast but cheaper since static RAM in the form of the SSD is less expensive than dynamic RAM. The authors expressed the opinion the total cost of ownership (TCO) needed to weight against the need for superior performance, and there were times when TCO did not adequately balance with application needs.

Texas Memory Systems White Papers

One significant difference of the Hutsell & Ault (2009) study is they choose to include the TPC-H test, which simulates a data warehouse and is applicable to this study. The test using HDDs took 5 hours 24 minutes 7 seconds to complete while the SSD test took 1 hour 7minutes 42 seconds. The memory component of Oracle, the SGA, varied from as large as the server could support to as little as was necessary for Oracle to run. Hutsell & Ault (2009) also discussed moving one database component at a time from HDD drives to SSD drives. The results paralleled those found in Narayanan et al. (2009) with the exception of rollback segments where the results were superior for the SSD over the HDD. Ault (2010) showed the SSD array outperformed the HDD array by a factor of 5.01 for the single streaming – maximum bandwidth test. The major source of wait in both tests was the direct path read, which for the HDD array were 137,651 seconds and 11,633 seconds for the SSD array (Ault, 2010, p. 8). The average overall data and index latency was 1.05 milliseconds for the SSD array and 27.47 milliseconds for the HDD array. Moving the temporary tablespace from the SSD to a HHD changed the "direct path read temp wait event has increased from an average of 1 millisecond per wait to 96 milliseconds per wait and the direct path write temp increased from 1 millisecond to 63 milliseconds (Ault, 2010, p. 14)." This result directly supports the findings of Lee, Moon, Park, Kim, & Kim (2009) that moving the temporary tablespace from HDD to SSD would be beneficial.

Cost, Energy Savings and Performance

Schmidt et al. (2009) evaluated SSD drives in conjunction with a database serving up XML records. Two criteria govern the management of the data center; providing ever-increasing capacity to store the amount of data and increased bandwidth to serve up the data when demanded. As the data center grow the energy consumption costs becomes a percentage of the operational costs. SSD proved to be superior to HDD in performance but failed to provide a superior IOPS/$ or MB/s/$ advantage. The authors proposed using a hybrid system where the most accessed XML existed on the SSD drives and the least accessed XML on the HDD drives with an aging process to move stagnant XML to the HDD drives. In a situation where IOPS are the primary factor, the authors did recommend full replacement of HDD drives with SDD drives. In terms of energy consumption, the report recommends switching HDD off and only switching them on when required to access the data particularly with seldom-used data would be more cost effective than going to SSD drives.

Problems with One-To-One Replacement

Stoica, Athanassoulis, Johnson, & Ailamaki (2009) concluded SSD should not be modeled for use with databases the same as magnetic disk drives but need their own specific abstraction to optimally operate. The authors suggest databases be flash aware and use an append and pack algorithm instead of the traditional method of read the block, change the block in memory, erase the block, write the new block. “The Append and Pack algorithm transforms temporal locality as provided by the overlying application, to spatial locality when data is placed on persistent storage, which allows to improve write performance” (Stoica et al., 2009, p. 14). The append and pack algorithm creates a layer of abstraction between the logical data organization and physical data placement. The pages written by the database append to new pages in the SSD in a sequential manner. Once a new page on the SSD is full, the append and pack algorithm erases the old and invalid page and returns it to the available pool of pages. The manufacturers of enterprise SSD drives add an abstraction, Flash Translation Layer, behind a controller to ensure backwards compatibility with HHD drives. Databases, in particular, are highly optimized to deal with the long access time of HDD drives as such simply replacing a HDD with a SSD does not take full advantage of the SSD, which is why they developed the Append and Pack algorithm.

Increased Energy Efficiency

One promise of SSD drives is a payback of the additional cost by increased energy efficiency but as Schall, Hudlet, & Härder (2010) points out without changes to the way a database interfaces with the SSD drives there is not enough gain to justify the cost. The authors propose changes to the storage algorithm in the database, which will net increased energy efficiency, but this will require database vendors to build in SSD aware algorithms. Stoica et al. (2009) determined the flash translation layer (FTL) may interfere with optimal performance of the new media but Schall et al. (2010) looked at the wear leveling characteristics of FTL. FTL implements a shadow block mechanism, writing updates to a new block instead of trying to erase and rewrite the original block. This improves the in-place update time and eliminates any wear endurance concerns. Agrawal et al. (2008) gave wear leveling as one of Microsoft Research's primary concerns with adopting SSD drives. Mirroring one technique advised by Schmidt et al. (2009) to switch off unused HDD drives this study disagrees based on the difference in startup times for the drives; 30 seconds for HDD and 1 second for SDD. Their recommendation is to place SDD drives in standby when not being actively accessed.

Using SSD Drives Instead Of Increasing the Buffer Cache

Taking the middle path, Canim, Mihaila, Bhattacharjee, Ross, & Lan (2010) proposes instead of a either or solution, a hierarchical mixture of SSD and HDD drives. With this proposal all data initially resides on the HDD drives and as data is accessed, an evaluation on how often it is accessed is performed. Least used data remains on the HHD drives and most often access data moves to the SSD drives. Therefore, the data most commonly accessed resides on the fastest drives improving overall database performance. In this configuration, the SSD drives function as a write-through cache. Working with the buffer pool, there needs to be rules for caching the data in the buffer spool, SSD write-through cache or left on disk. The very highest value data will reside in memory in the buffer pool with high value data in the next layer. The database queries the buffer pool first to determine if the data is there then queries the SSD drives and finally the HDD drives. Admittedly if the data is on the HHD drives the access time is slightly extended but for higher value data access time should be reduced. For the TPC-H test, the authors found when the size of the SSD buffer pool was 10% of the data set it produced a factor of three times in performance. Several vendors have implemented data storage systems using a hierarchical mixture: Oracle's Exadata Smart Flash Cache, Teradata Virtual Storage, and finally Sun's ZFS enterprise file system.

Chapter 3 – Methodology

The goal of this study is to determine if replacing magnetic disk drives with solid-state device drives will improve the extract-transform-load performance on the Exempla Healthcare Clarity Reporting Data Warehouse.

Background

In the field of large data warehouses, one of the problems is the time it takes to perform the extraction, transformation, and load (ETL). Epic Systems is a vendor of enterprise medical record systems (EMR) and has the Clarity reporting data warehouse as part of its solution. The ETL in this auxiliary application consists of using SQL to query the EMR for all data from the previous day. The outputs of these queries are spooled into SQL Loader files consisting of a series of sequential writes on the data warehouse. The next step is the sequential reading of these files and using SQL Loader appending the records into the staging tables, using both sequential reads and writes. The final step is a series of MERGE statements into the reporting tables. A merge statement inserts a row if the data does not exist, or updates the row if it already exists. The Clarity reporting data warehouse uses EMC 15K magnetic disk drives in its current format and averages 2 hours and 40 minutes to complete the ETL process. This is with three hospitals and two clinics on the current system. Due to the acquisition of Exempla Healthcare by the Sisters of Charity of Leavenworth Health Systems (SCLHS), the three-year expectation is to expand to eleven hospitals and eight hundred clinics. A simple extrapolation just based on the hospitals puts the ETL time at 10 hours, which would be unacceptable since pre-shifts reports need to run at 0600 daily. The decision was to build a new server using EMC enterprise flash drives (EFD) in the hope the superior read/write latency would provide acceptable performance.

Place

The study was performed at the Exempla Healthcare Lutheran Medical Center Data Center where the servers reside.

Exempla Lutheran Medical Center
8300 W. 38th Avenue - Wheat Ridge, CO 80033

Participants

Douglas Ervin, Clarity ETL Administrator and author of this study, gathered the data and Gordana Nichols, Clarity ETL Backup Administrator, verified the collected data.

Materials

There are no materials needed beyond the normal monitoring tools and a Microsoft Excel© spreadsheet. The Exempla Healthcare owns and operates the database servers and has granted permission to perform this study. The existing server is an IBM Power 6 570 with five P6 cores allotted to the Clarity Reporting data warehouse. The operating system is AIX 6.2 and the disk array is an EMC^2 DMX using 15K rpm magnetic disk drives and set up as RAID 1+0. The future server will be an IBM Power 7 750 with eight P7 cores allotted to the data warehouse. The operating system will also be AIX 6.1. The disk array will be an EMC^2 VMAX with enterprise flash drives using a RAID 5 configuration.

Procedure

The application server, henceforth referred to as the Clarity Console, controls the ETL process for the data warehouse. The Clarity Console measures the time each step of the ETL for each table plus the overall time for each table and the composite time for the entire evolution. Each table execution has seventeen discrete steps so very fine-grained measurements are possible. If it is determined, a composite time does not give meaningful information as to the effect of solid-state device drives on the ETL process a more detailed investigation is possible.

Initially each day the study will record the individual reading for a single table as well as the composite time for the entire execution. Due to fact, weekend processing times are approximately one hour quicker due to less patient encounters; separation of the weekday and weekend times eliminates a false low average. The only patient encounters on the weekend are in-patient and emergency departments, while the weekdays include outpatient encounters.

Data Analysis

Using formulas presented in Leedy & Ormond (2010), calculations of the mean, range, and standard deviation for each system and summarized. Student T test will be used to determine if the two means are significantly different. Weekday and weekend times, both the individual table steps and composite times, are analyzed separately to give a truer measure of differences in the performance of the two systems.

The study cannot be 100% replicated due to following problems. The new system uses a more advanced processor than the original system. By looking at the individual steps for a single table, the author hopes to minimize the effect of the processor on ETL performance. Due to being a production system, some natural fluctuation of patient census will result in some skewing of the normal curve of the data. As an example, a sudden outbreak of flu would dramatically change the patient population and affect the ETL times. The final problem with replication is no hospital systems are identical in population or volume. For this reason, a similar study performed at another facility should have similar but not identical results.

Chapter 4 Results

The following tables represent the data collected using the 15K RPM hard disk drives for the execution of the ETL script as a whole. Since the data so clearly demonstrate and advantage for the solid-state devices, the data for an individual table is not included.

Table 1: Execution Times Using the Hard Disk

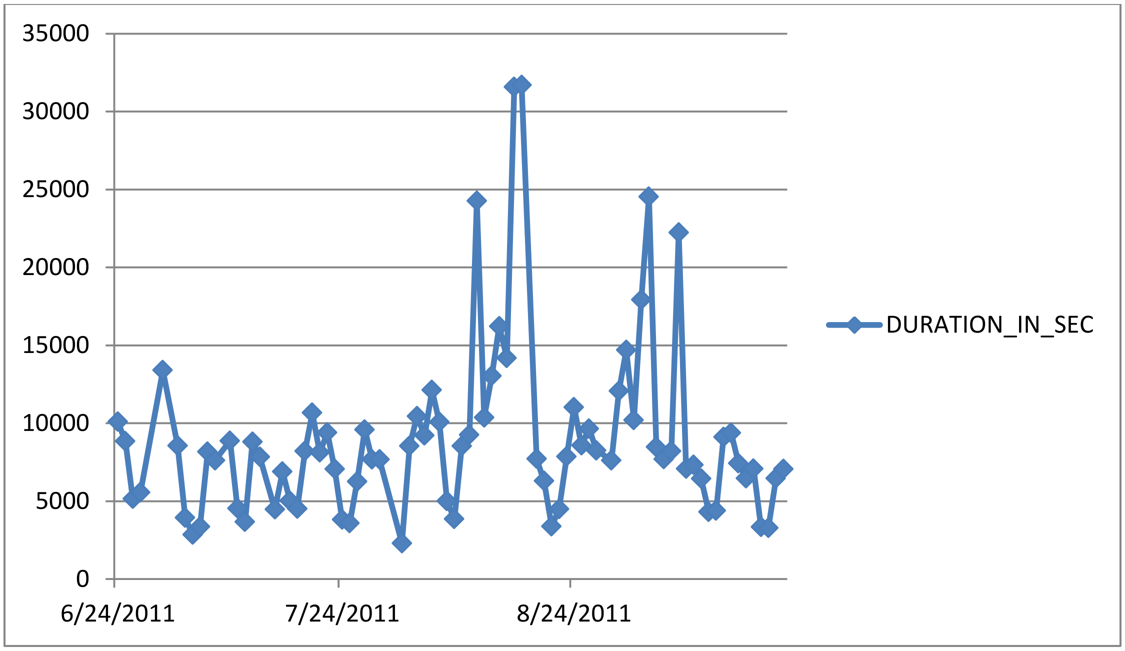

Table 2: Analysis of Hard Disk Execution Times - Weekdays

Mean	Minimum	Maximum	Standard Deviation
9,287.54	3,373.00	24,526.00	3,708.00

Table 3: Analysis of Hard Disk Execution Times - Weekends

Mean	Minimum	Maximum	Standard Deviation
4,584.05	2,304.00	8,480.00	1,574.80

A comparison of Tables 2 and 6, and Tables 3 and 7, show a considerable difference in execution times for the ELT script.

Table 5: Execution Times Using the Solid State Devices

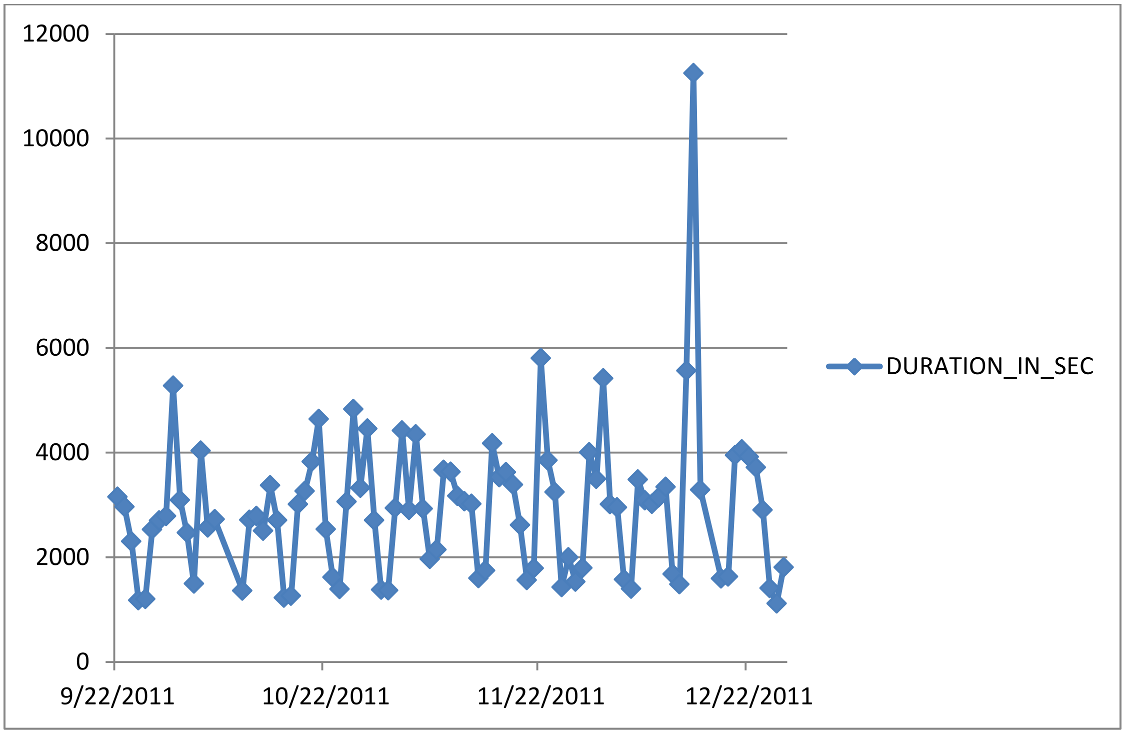

Table 6: Analysis of Solid-State Device Execution Times - Weekdays

Mean	Minimum	Maximum	Standard Deviation
3,491.91	1,429.00	11,250.00	1,297.41

Table 7: Analysis of Solid-State Device Execution Times - Weekend

Mean	Minimum	Maximum	Standard Deviation
1,553.74	1,115.00	2,468.00	301.81

The Student T test is one method of determining if two populations of data are significantly different or if the two sets of data are just groupings of the same populations (Leedy & Ormond, 2010).

The T-test probability as calculated using Microsoft Excel 2010 is as shown in Table 8: T-test Results below and the analysis used the parameters: =T.TEST(B2:B82,D2:D93,2,3).

B2:B82 denotes the duration in seconds for the executions using the hard disk drives, D2:D93 denotes the duration in seconds for the execution using the solid-state devices, 2 specifies a two distribution, and 3 specifies two-sample with unequal variance (heteroscedastic). The result of 6.45 E-15 indicates there is an extremely low probability the two data sets are the same so a comparison of the means can be used to correlate a conclusion as to validity of the hypotheses.

Table 8: T-test Results

T-test	6.45701E-15

Chapter 5 Discussion and Conclusion

Discussion

This study originally proposed gathering data for the entire ETL execution and the data for a single table out of that execution. Due to the clear difference of the data from the ETL execution as a whole, this study omits the discussion of the individual table for brevity. All further discussion shall only involve the entire ETL execution.

Comparing the time for the HDD ETL execution time 9,287.54 ± 3,708.00 seconds to the SSD ETL execution time 3,491.91 ± 1,297.41 seconds shows an increase in overall performance of 2.66 for the weekday ETL. The results for the weekend ETL execution are 4,584.05 ± 1,574.80 seconds versus 1,553.74 ± 301.81 seconds, with a performance increase of 2.98. Comparing the mean minus the standard deviation of the HDD ETL weekday time versus mean plus the standard deviation SSD ETL times intuitively correlates with the Student T test performed as part of the analysis; 9,287.54 - 3,708.00 = 5,579.51 seconds as compared to 3,491.91 + 1,297.41 = 4,789.32 seconds. A difference of 790.19 seconds or 13.17 minutes between two results as a worst-case scenario satisfied upper management of the success of the upgrade.

During 2012, Sisters of Charity of Leavenworth Health Systems (SCLHS) added seven more hospitals and several hundred community-based physicians, thus bringing the number of patients being monitored to over 200, 000. The number of discrete data elements being gathered and aggregated has risen to over 2 million rows of data a day. The SSD drives proved instrumental in allowing this rapid growth in the EMR.

Conclusion

In terms of absolute performance, the SSD devices have proved superior to HHD devices in all but a few cases. The data gathered from monitoring ETL executions shows the marked drop in execution time and decreased variance due to switching from HHD to SSD devices. The change in enterprise architecture for SCLHS was a success and expanded the capabilities of the data warehouse. For SCLHS, the increased cost of SSD over HHD was justified.

References

Agrawal, N., Prabhakaran, V., Wobber, T., Davis, J. D., Manasse, M., & Panigrahy, R. (2008). *Design Tradeoffs for SSD Performance*. Manuscript submitted for publication.

Ault, M. (2010, August). *Oracle, TPC-H and RamSan* (White Paper). Retrieved from Texas Memory Systems: http://www.ramsan.com/

Canim, M., Mihaila, G. A., Bhattacharjee, B., Ross, K. A., & Lan, C. A. (2010). SSD Bufferpool Extensions for Database Systems. In (Ed.), *Proceedings of the VLDB Endowment, Vol. 3, No. 2* (pp. 1435-1446). Singapore: VLDB Endowment.

Hutsell, W., & Ault, M. (2009, November). *Faster Oracle Performance with Solid State Disks* (White Paper). Retrieved from Texas Memory System: http://www.superSSD.com

Lee, S., Moon, B., & Park, C. (2009). Advances in Flash Memory SSD Technologyfor Enterprise Database Applications. In (Ed.), *The Proceedings of the ACM Special Interest Group on Management of Data 2009* (pp. 863-870). Providence, Rhode Island, U.S.A.: Association of Computing Machines.

Lee, S., Moon, B., Park, C., Kim, J., & Kim, S. (2009). A Case for Flash Memory SSD in Enterprise DatabaseApplications. In (Ed.), *The Proceedings of the ACM Special Interest Group on Management of Data 2009* (pp. 1075-1086). Vancouver, BC, Canada: Association of Computing Machinery.

Leedy, P. D., & Ormond, J. E. (2010). *Practical Research* (9th ed.). UpperSaddle River, New Jersey, USA: Merrill.

Moshayedi, M., & Wilkison, P. (2008, July/August). Enterprise SSDs [Magazine]. *ACM QUEUE*, 32-39.

Narayanan, D., Thereska, E., Donnelly, A., Elnikety, S., & Rowstron, A. (2009). Migrating Server Storage to SSDs: Analysis of Tradeoffs. In (Ed.), *Proceedings of the 4th ACM European conference on Computer systems* (pp. 145-158). Nuremberg, Germany: Association of Computing Machinery.

Schall, D., Hudlet, V., & Härder, T. (2010). Enhancing Energy Efficiency of Database Applications Using SSDs. In (Ed.), *Proceedings of the Conference on Computer Science & Software Engineering 2010* (pp. 1-9). Montreal, Canada: Association of Computing Machinery.

Schmidt, K., Ou, Y., & Härder, T. (2009). The Promise of Solid State Disks Increasing efficiency and reducing cost of DBMS processing. In (Ed.), *Proceedings of the Conference on Computer Science & Software Engineering 2009* (pp. 35-41). Montreal, Canada: Association of Computing Machinery.

Stoica, R., Athanassoulis, M., Johnson, R., & Ailamaki, A. (2009). Evaluating and Repairing Write Performance on FlashDevices. In (Ed.), *Proceedings of the Fifth International Workshop on Data Management on New Hardware* (pp. 9-14). Providence, Rhode-Island, U.S.A.: Association of Computing Machinery.

Appendix A

Raw Data

Table 9: Raw Data Used in Analysis

Date HDD	Duration In Seconds	Date SSD	Duration In Seconds
6/24/2011	10103	9/22/2011	3151
6/25/2011	8844	9/23/2011	2962
6/26/2011	5147	9/24/2011	2301
6/27/2011	5570	9/25/2011	1176
6/30/2011	13403	9/26/2011	1203
7/2/2011	8564	9/27/2011	2530
7/3/2011	3932	9/28/2011	2697
7/4/2011	2857	9/29/2011	2786
7/5/2011	3373	9/30/2011	5273
7/6/2011	8161	10/1/2011	3093
7/7/2011	7621	10/2/2011	2468
7/9/2011	8878	10/3/2011	1492
7/10/2011	4540	10/4/2011	4031
7/11/2011	3676	10/5/2011	2565
7/12/2011	8808	10/6/2011	2722
7/13/2011	7840	10/10/2011	1361
7/15/2011	4467	10/11/2011	2714
7/16/2011	6904	10/12/2011	2786
7/17/2011	5032	10/13/2011	2503
7/18/2011	4513	10/14/2011	3371
7/19/2011	8205	10/15/2011	2706
7/20/2011	10665	10/16/2011	1227
7/21/2011	8123	10/17/2011	1261
7/22/2011	9395	10/18/2011	3009
7/23/2011	7058	10/19/2011	3263
7/24/2011	3821	10/20/2011	3823
7/25/2011	3582	10/21/2011	4639
7/26/2011	6268	10/22/2011	2535
7/27/2011	9595	10/23/2011	1618
7/28/2011	7674	10/24/2011	1392
7/29/2011	7670	10/25/2011	3062
8/1/2011	2304	10/26/2011	4829
8/2/2011	8537	10/27/2011	3325
8/3/2011	10453	10/28/2011	4458
8/4/2011	9216	10/29/2011	2707
8/5/2011	12128	10/30/2011	1380
8/6/2011	10076	10/31/2011	1368
8/7/2011	5006	11/1/2011	2941

8/8/2011	3856	11/2/2011	4422
8/9/2011	8546	11/3/2011	2900
8/10/2011	9254	11/4/2011	4345
8/11/2011	24261	11/5/2011	2925
8/12/2011	10373	11/6/2011	1966
8/13/2011	13036	11/7/2011	2144
8/14/2011	16220	11/8/2011	3665
8/15/2011	14186	11/9/2011	3631
8/16/2011	31579	11/10/2011	3172
8/17/2011	31691	11/11/2011	3065
8/19/2011	7709	11/12/2011	3016
8/20/2011	6300	11/13/2011	1600
8/21/2011	3393	11/14/2011	1748
8/22/2011	4490	11/15/2011	4172
8/23/2011	7859	11/16/2011	3527
8/24/2011	11027	11/17/2011	3625
8/25/2011	8590	11/18/2011	3384
8/26/2011	9658	11/19/2011	2617
8/27/2011	8259	11/20/2011	1561
8/29/2011	7622	11/21/2011	1786
8/30/2011	12068	11/22/2011	5801
8/31/2011	14697	11/23/2011	3852
9/1/2011	10209	11/24/2011	3247
9/2/2011	17917	11/25/2011	1429
9/3/2011	24526	11/26/2011	1994
9/4/2011	8480	11/27/2011	1530
9/5/2011	7695	11/28/2011	1796
9/6/2011	8207	11/29/2011	4004
9/7/2011	22234	11/30/2011	3496
9/8/2011	7080	12/1/2011	5414
9/9/2011	7337	12/2/2011	3012
9/10/2011	6446	12/3/2011	2950
9/11/2011	4306	12/4/2011	1572
9/12/2011	4401	12/5/2011	1398
9/13/2011	9114	12/6/2011	3485
9/14/2011	9379	12/7/2011	3088
9/15/2011	7431	12/8/2011	3017
9/16/2011	6471	12/9/2011	3143
9/17/2011	7089	12/10/2011	3342
9/18/2011	3348	12/11/2011	1677
9/19/2011	3278	12/12/2011	1480
9/20/2011	6460	12/13/2011	5563
9/21/2011	7056	12/14/2011	11250
		12/15/2011	3285
		12/18/2011	1594
		12/19/2011	1631

		12/20/2011	3951
T-test	6.45701E-15	12/21/2011	4061
		12/22/2011	3919
		12/23/2011	3714
		12/24/2011	2902
		12/25/2011	1407
		12/26/2011	1115
		12/27/2011	1807

Annotated Bibliography

Agrawal, N., Prabhakaran, V., Wobber, T., Davis, J. D., Manasse, M., & Panigrahy, R. (2008). Design Tradeoffs for SSD Performance. *USENIX Technical Conference* (pp. 1-14). Madison WI USA: Microsoft Research.

Microsoft research produced one of the first studies involving the performance of solid-state device drives when used with a database. The researchers found the new media has superior read characteristics, both sequential and random, and superior sequential write characteristics. Random writes were inferior to magnetic disk drives. The study showed SSD drive performance was dependent on data placement, parallelism, write ordering, and workload management. This was the study, which prompted the need for the current study.

Canim, M., Mihaila, G. A., Bhattacharjee, B., Ross, K. A., & Lang, C. A. (2010). SSD Bufferpool Extensions for Database Systems. *The 36th International Conference on Very Large Data Bases* (pp. 1435-1446). Singapore: VLDB Endowment.

Canim, Mihaila, Bhattacharjee, Ross, & Lan (2010) proposes instead of a either or solution, a hierarchical mixture of solid-state device and magnetic disk drives. In an effort to improve performance of a database they proposed using the solid-state device drives as replacement for the buffer cache in RAM. The authors pointed out several commercially available solutions. This is a good study showing the worth of only replacing part of the storage media with solid-state device drives.

He, Z., & Veeraraghavan, P. (2009). Fine-grained updates in database management systems for flash memory. *Information Sciences 179* , 3162–3181.

The study deals with the shortcomings of solid-state device drives in buffering the data to be written. Algorithms were shown to be in need of redesign and added to the inefficiency of random writes. The study did not add any new information and was deselected from this study.

Huang, M., Serres, O., Narayana, V. K., El-Ghazawi, T., & Newby, G. (2010). Efficient Cache Design for Solid-State Drives. *CF'10* (pp. 41-50). Bertinoro, Italy: ACM.

The authors focused on proper design of the flash translation layer in regards to how they perform and how to make them more efficient. A good paper in terms of recommendations to electronic engineers but not as useful to the database or system administrator trying to determine if solid-state device drives will improve performance.

Hutsell, W., & Ault, M. (2009, November ND). *Faster Oracle Performance with Solid State Disks.* Retrieved May 13, 2011, from Texas Memory Systems: http://www.ramsan.com/solutions/applications/oracle-solutions

This whitepaper specifically addresses the performance gains possible with partial and whole sale replacement of magnetic disk drives with solid-state device drives. The author explore the effect of placing each database subsystem on each type of drive and the change in performance on various industry performance tests. While not peer reviewed, this study is pivotal in predicting performance gains.

Lee, S.-W., Moon, B., & Park, C. (2009). Advances in Flash Memory SSD Technology for Enterprise Database Applications. *SIGMOD'09* (pp. 863-870). Providence, Rhode Island, U.S.A.: ACM.

The authors researched the problems from the aspect of Access Density or the ration of IOPS to the disk capacity. For HDD a drive, as the capacity has gone up the access density has actually dropped, one of the few areas in computers where a measure of performance has decreased over time. One large scale TPC-C (a testing protocol for transactional databases) used 170 HDDs per each of its 64 cores to maximize performance. The researchers concentrated on the pattern of a write followed by a read, which was determined to be a problem in other studies.

Lee, S.-W., Moon, B., Park, C., Kim, J.-M., & Kim, S.-W. (2008). A Case for Flash Memory SSD in Enterprise Database Applications. *SIGMOD'08* (pp. 1075-1086). Vancouver, BC, Canada: ACM.

The authors researched performance gains associated with placing the transaction log, rollback segments, and temporary tablespaces on solid-state device drives. Both the transaction logs and temporary tablespaces showed marked improvement but the rollback segments times were much slower than having them on magnetic disk drives. A very good study for partially replacing existing drives.

Li, Y., He, B., Yang, R. J., Luo, Q., & Yi, K. (2010). Tree Indexing on Solid State Drives. *The 36th International Conference on Very Large Data Bases* (pp. 1195-1209). Singapore: VLDB Endowment.

The authors proposed an alternative to B-Tree file structure on the solid-state device drives. The alternative is a FD-tree index using a logarithmic structure. This proposal promises increased performance in future drives. While an interesting design paper, it is not applicable to the proposed study.

Moshayedi, M., & Wilkison, P. (2008, July/August ND). Enterprise SSDs. *ACM Queue*, pp. 32-39.

This article is a case study of how a company dealt with the issue of magnetic disk drives not being able to meet their input/output needs. The company replaced their existing drives with a smaller number of solid-state device drives, at a ratio of 30 to 1. The new drives meet their requirements but there were multiple difficulties to be addressed. Excellent article showing the promise of the new media.

Narayanan, D., Thereska, E., Donnelly, A., Elnikety, S., & Rowstron, A. (2009). Migrating Server Storage to SSDs: Analysis of Tradeoffs. *EuroSys'09* (pp. 145-158). Nuremberg, Germany: ACM.

Recently, flash-based solid-state drives (SSDs) have become standard options for laptop and desktop storage, but their impact on enterprise server storage has not been studied. Provisioning server storage is challenging. It requires optimizing for the performance, capacity, power and reliability needs of the expected workload, all while minimizing financial costs.

Park, H., & Shin, D. (2010). Buffer flush and address mapping scheme for flash memory solid-state disk. *Journal of Systems Architecture 56*, 208–220.

The authors studied the write performance of the internal write buffer and parallel architecture of solid-state device drives. The study focused on the coarse-grained address mapping called superblock mapping, which the parallel architecture uses. The study demonstrated the poor performance of random writes. The study proposes a multi-level address mapping scheme to use small-sized writes and achieved a 64% increase. The study does not support the research question.

Schall, D., Hudlet, V., & Härder, T. (2010). Enhancing Energy Efficiency of Database Applications Using SSDs. *C3S2E − 10* (pp. 1-9). Montreal: ACM.

One promise of SSD drives is a payback of the additional cost by increased energy efficiency but as this study points out without changes to the way a database interfaces with the SSD drives there is not enough gain to justify the cost. The authors propose changes to the storage algorithm in the database, which will net increased energy efficiency, but this will require database vendors to build in SSD aware algorithms.

Schmidt, K., Ou, Y., & Härder, T. (2009). The Promise of Solid State Disks. *C3S2E-09* (pp. 35-41). Montreal: ACM.

This study evaluated SSD drives in conjunction with a database serving up XML records. Two criteria govern the management of the data center; providing ever-increasing capacity to store the amount of data and increased bandwidth to serve up the data when demanded. As the data center grow the energy consumption costs becomes a percentage of the operational costs. SSD proved to be superior to HDD in performance but failed to provide a superior IOPS/$ or MB/s/$ advantage. This study will be used.

Stoica, R., Athanassoulis, M., Johnson, R., & Ailamaki, A. (2009). Evaluating and Repairing Write Performance on Flash Devices. *Proceedings of the Fifth International Workshop on Data Management on New Hardware* (pp. 6-14). Providence, Rhode-Island: ACM.

The authors concluded SSD should not be modeled for use with databases the same as magnetic disk drives but need their own specific abstraction to optimally operate. The authors suggest databases be flash aware and use an append and pack algorithm instead of the traditional method of read the block, change the block in memory, erase the block, write the new block. “The Append and Pack algorithm transforms temporal locality as provided by the overlying application, to spatial locality when data is placed on persistent storage, which allows to improve write performance” (Stoica et al., 2009, p. 14)

Tsirogiannis, D., Harizopoulos, S., Shah, M. A., Wiener, J. L., & Graefe, G. (2009). Query Processing Techniques for Solid State Drives. *SIGMOD'09* (pp. 59-71). Providence, Rhode Island: ACM.

One of the problems this study investigated was the ineffeciencies of solid-state device drives in modern relational databases. Relational databases have been optimized to overcome the problems with magnetic disk dives and the new media presents different challenges to the database vendors. The conclusion was gains can be achieved by using the new drives but they may not be as great as could be achived if the databases were flash drive aware.