

## Various Factor Comparison of DRAM-SSD and HDD using TPC-H Benchmark

Hyun-Ju Song and Young-Hun Lee\*

*Dept. of Electronic Eng., Hannam University, Ojeong -dong,  
Daedeok-gu, Daejeon 306-791, Korea  
iisong86@nate.com, \*Corresponding Author: yhlee00@hnu.kr*

### **Abstract**

*Recently, Users needed storage for processing high-capacity data. Meanwhile, the HDD was used primarily as storage device, but SSD was developed as a fast access device. So, the use of SSD increased a lot of capacity-process. Using a SAN Switch, we tested in order to resolve data process faster. So in this paper, DRAM-SSD and HDD's performance ability in the data process will be confirmed using TPC-H Benchmark on MYSQL at SAN environment. From the performance analysis results, performance difference of HDD and DRAM-SSD is little when database size is low at san switch connecting with storage. But ad-hoc query process ability difference between DRAM-SSD and HDD increased. Based on these results in a SAN-based, DRAM-SSD has a better performance than the HDD. Therefore San is judged to be more effective when dealing with the large amounts of data using a SAN to manage data. And CO2 & power consumption was tested. CO2 emissions analysis result during analysis of performance of each storage using Postmark and TPC-H, as shown in Figure 4 HDD storage can be considered less CO2 emissions due to the power consumption when non-load. But, increasing the load to DRAM-SSD storage generated less CO2 emission.*

**Keywords:** *TPC-H, postmark benchmark, power consumption, CO2*

### **1. Introduction**

Recently, Users needed storage for processing high-capacity data. Meanwhile, the HDD was used primarily as storage device, but the development of SSD is faster access as storage devices and a lot of research is being accelerated. From the research of HDD and SSD, the difference of data I/O processing performance was progressed by comparing performance of storage device of each [1-3]. Basic storage device's manage data and use of DBMS which provided efficient and convenient way increased. So in this paper, we will compare DRAM-SSD to HDD and confirm showing good performance in data processing, using TPC-H Benchmark through SAN Switch on MYSQL.

The process of this paper is as follow. In Chapter 2 we will introduce used technology that evaluates performance of HDD and SSD, in Chapter 3 we will introduce analysis environment and condition by thinking each storage device and using tool. And in Chapter 4 analyzes test results will be shown with conditions described in Chapter 3, and Chapter 5 is concluded.

.....  
\* Corresponding Author

## 2. Related Study

### 2.1 MYSQL

MYSQL is the relative database management system of open source that uses SQL which is standard database quality language, which is very fast, flexible, and easy to use. MYSQL offers client/server environment, and a server installed MYSQL has MYSQL daemon called mysqld, so client program connects via network by this daemon so that it can control data [5].

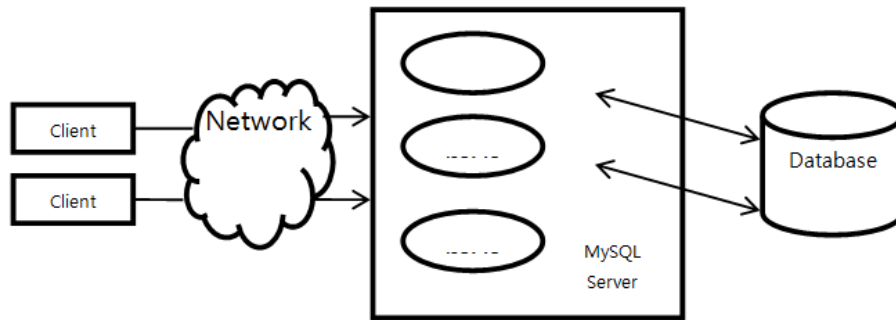


Figure 1. Performance Structure of MYSQL

### 2.2 SAN

SAN stands for Storage Area network and say move available high-speed network for large-capacity between storage equipment which unrelated distributed sort of Host. i.e., say all together configured network to communication storage component and system component at fiber channel networks. SAN integrate and share storage, and supply high-performance link to data device, and add overlapping-link to storage, and speeding up data backup, and support available-high-clustered systems. SAN have advantage by building implementing a highly available data access, integration resources and management of storage, required window to backup and traffic reduce, occupation solution of host CPU cycle, data retention function through disaster Accepted techniques [6].

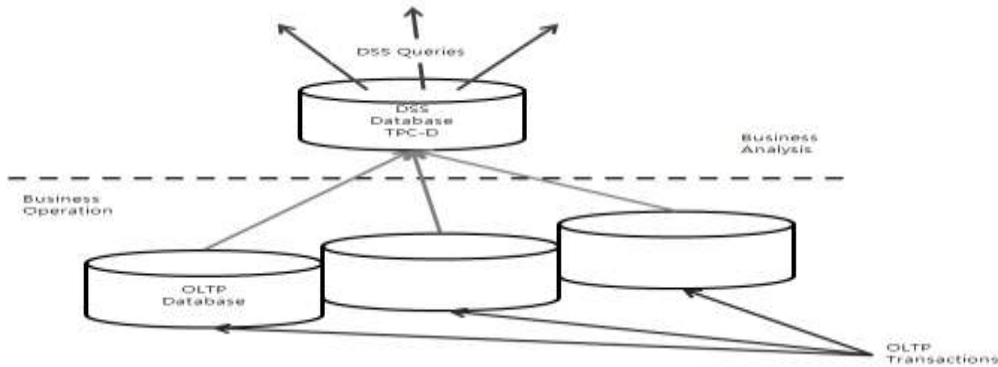


Figure 2. General Storage Server VS SAN Structure

### 2.3 TPC-H

TPC stands for Transaction processing performance and is non-profit organization that has multiple hardware and software companies and small number of user organizations. Usually TPC called transaction processing performance evaluation committee, but TPC-alphabet notation tells the benchmark model. TPC becomes standard evaluating processing performance of on-line transaction processing systems. TPC defines transaction processing

and benchmarks of database and is used to performance measure of total system including disk I/O and s/w. As a benchmarking tool to measure that how fast it can handle complex SQL, and defines 22 SQL statements and DB schema, and set of data about 1GB. TPC-H benchmark is public performance test that is used SQL that Business-oriented ad-hoc Query and concurrent data modifications made by the combination about large data. Figure 3 showing business environment of TPC-H, impromptu Query and modification transactions was performed at table from multiple users, model situation entered data from database of decision support system from OLTP system [7].

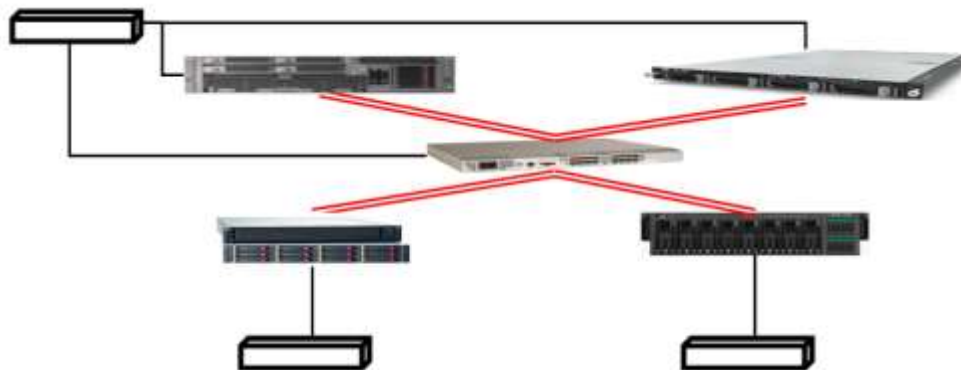


**Figure 3. TPC-H's Business Environment**

### 3. Performance Analysis Environment and Conditions

#### 3.1 Performance Test Environment

This paper comparative analyzed the performance of each, which the HDD as a storage device in modern mainly used and the SSD as a storage device in modern society used increasingly. HDD Storage and DRAM-SSD Storage were constructed in conjunction with a SAN switch for test Environment. Test Environment constituted as Figure 4, and test subject server formed as Linux CentOS 5.3 version, performance measure tool used TPC-H Benchmark, to use a tool as the database was installed MYSQL 5.0.90 version.



**Figure 4. Test Environment Configuration**

### 3.2 Test Procedure and Condition

Performance of each storage device comparative analyzed to configure SAN environment using TPC-H Benchmark. Test using the TPC-H Benchmark compared to performance analysis through total three step procedure with the results. Load Test is the first step, which make up database and store to generate data. Next step is power test, which was analyzed by measuring the ability when Single active user is run the Query. Final step is Throughput Test, which was analyzed by measuring the ability when Multi active user is run Query at same time. The performance of each storage device analyzed through Power TEST and Throughput Test combined results in SAN environment. Data generation and total of eight tables produced to store generated data in data storage of Mysql. Read data in the created file store data which get by running stored-code in the DBMS, confirm insert play time of stored data. A total of 22 Query given in the TPC-H benchmark, which the play-time check by running through the Query Browser, total of 22 Query stored for multi active user test through vim command in TPC-H installed the directory. After run total of 22 Query a script run to delete set by checking the result. Result of insert execution time and 22Query execution time and delete execution time calculate value of Power Test apply in below expression 1.

$$\frac{1}{24} \times \frac{3600 \times SF}{\sqrt{\prod_{i=1}^{22} QI(i,0) \times \prod_{j=1}^2 RI(j,0)}} \quad (1)$$

QI (i, 0) is the execution time of I second query, RI (j, 0) as the running time of insert and delete unit second the data generated by the size of SF, i.e., the value of the database. When each of Query accordingly practice ability of multiple users at the same time make amount of data as table 1 number of users to differ.

**Table 1. Conditions for Throughput Test**

Database capacity	User number
1GB	2
5GB	2
10GB	3

The fixed users perform as shown in above table1 parallel 22 Query of each. And time check was expressed in units of sec from first user starts the first Query to last user finish last Query. By substituting in Equation 2 is to evaluate performance time to check on the recipe.

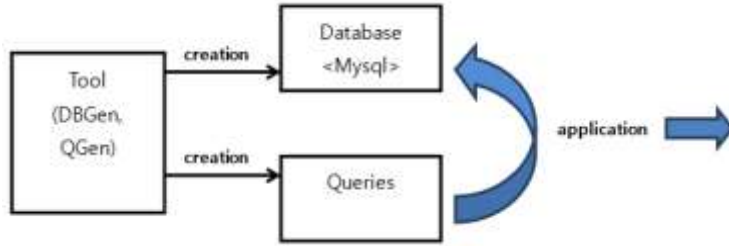
$$(2) \quad \frac{S \times 22 \times 3600}{T_s} \times SF$$

In the above expression 2,  $T_s$  is performance time, S is number of user, SF show amount of data. Through the result of expression 1 and expression 2, expression 3 has produced result to get final result to analysis performance of each storage device.

$$\sqrt{Power@Size \times Throughput@Size} \quad (3)$$

QphH@Size which is a calculated value from Equation 3 is reflected value of the many features of Query processing system, these features are performed Query, Query processing

ability when performed Query, Query performed by multiple users simultaneously including all the processing ability. Thus, QphH@Size shows ad-hoc query capabilities that to handle hourly capacities of the database.



**Figure 5. TPC-H Applies to Get the Results the Block Diagram**

## 4. Test Results

### 4.1 Performance Measurements along the Environment

Test results using TPC-H Benchmark comparative analyzed performance at SAN environment in according to described result calculation method in Section 3.2. Some of total 22 Query for test when Database capacity 1GB, 5GB, 10GB, by modifying comparative analysis performance for each size Shorten the time the test case.

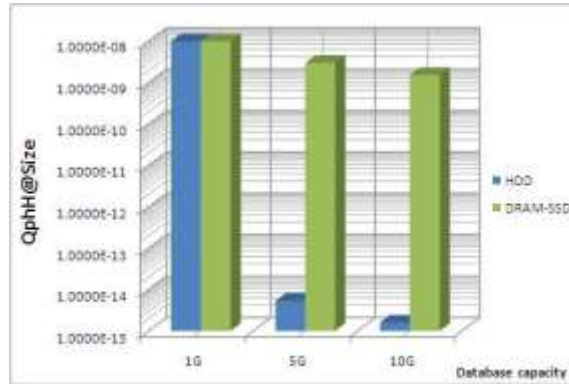
**Table 2. SAN Environment HDD Performance Measurements**

HDD	Database capacity		
	1GB	5GB	10GB
Test result	1GB	5GB	10GB
Power@size	1.0891E-09	3.8770E-30	3.6293E-31
Throughput@size	15.7283	7.372	7.0681
QphH@size	1.3088E-04	5.3461E-15	1.6016E-15

**Table 3. Local Environment DRAM-SSD Performance Measurements**

DRAM-SSD	Database capacity		
	1GB	5GB	10GB
Test result	1GB	5GB	10GB
Power@size	9.7911E-09	4.6299E-20	8.8710E-22
Throughput@size	19.3928	34.7933	80.1214
QphH@size	4.3575E-04	1.2692E-09	2.6660E-10

As can be seen through above Table 2 and Table 3, the practical value for comparing the performance of DRAM-SSD Storage and HDD Storage in SAN environment is QphH@size, compared result to only database-capacity of a storage device of each same below picture.



**Figure 6. SAN Environment Database Capacities of Each Storage Device QphH@size Value**

QphH@size value that can handle ad-hoc queries, the little difference find out between ad-hoc query handle-capabilities of HDD Storage and DRAM-SSD Storage in low carrying a load 1G, but DRAM-SSD storagere can be seen that much higher which per hour to handle ad-hoc query capabilities more than HDD storage when amounts of data is increasing.

#### 4.2 Analysis and Comparison of Power Consumption

At this section analyzed power consumption having a test results by attaching the measuring instrument.

##### 4.2.1 Using a Postmark Benchmark

**Table 4. Power Consumption Measurement Results Table**

Postmark	Test result(W)		
	HDD	DRAM-SSD	Rate
Test Level			
Non-load	263.8	475.7	0.55
Low	264.18	476.55	0.55
Medium	264.63	484.74	0.55
High	281.26	508.01	0.55

**Table 5. Total Power Consumption Comparison**

Postmark Test Level	Test result(W)		
	HDD	DRAM-SSD	Rate
Non-load	263.8	475.7	0.55
Low	264.18	476.55	0.55
Medium	264.63	484.74	0.55
High	281.26	508.01	0.55

Based on the results of Table 4 and Table 5, case of low data I/O occurs, power consumption of HDD storage evaluate less occurs by difference of non-load power consumption, but, power consumption of DRAM-SSD is less than HDD storage to difference of run time by increasing database capacity.

### 4.3 CO2 Emission Comparison

When power spend of each storage device comparative analyzed emission of CO2.

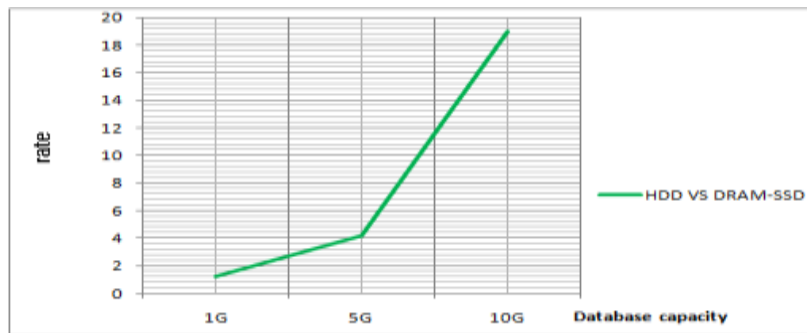


Figure 7. CO2 Occur Ratio of HDD and DRAM-SSD

CO2 emissions analysis result during analysis of performance of each storage using Postmark and TPC-H, as shown in Figure 7 HDD storage can be considered less CO2 emissions due to the power consumption when non-load. But, increasing the load to DRAM-SSD storage generated less CO2 emission.

## 5. Conclusion

In this paper, we analyzed the data processing performance of a DRAM-SSD and HDD as a data processing storage using the TPC-H Benchmark on mysql in a san environment.

From the performance analysis results, when database Size increased performance compared analysis value QphH@size was increased with TPC-H, but difference of the ad-hoc query processing ability of DRAM-SSD and HDD are increased when compared of each database capacity.

As a results of a postmark benchmark analyzed about the HDD storage and DRAM-SSD storage for performance analysis, the performance difference of HDD and DRAM-SSD was little in low data I/O. But, the DRAM-SSD had a better performance than the HDD in large amounts of data I/O.

And we analyzed the data processing performance of a DRAM-SSD and HDD as a data processing storage using the TPC-H Benchmark on mysql in a san environment.

From the performance analysis results, when database Size increased performance compared analysis value QphH@size was increased with TPC-H, but difference of the ad-hoc query processing ability of DRAM-SSD and HDD are increased when compared of each database capacity.

Based on these results, using by connecting SAN Switch and DRAM-SSD storage is judged to be effective at large amounts of data I/O required field or applications. Also mentioned in the introduction, SSD has benefits if SSD price is stabled. Using SSD is thought to be effective in case of industry side which required large amounts of data I/O or Media Server, using storage device of computer and notebook computer. In the future, we will be analyzed the power consumption from the TPC-H Benchmark at SAN Environment.

In addition, the database become larger than 10G, DRAM-SSD had less power consumption and emissions about 92% than the HDD in analysis of power consumption and CO2 emissions.

In the future, the new direction is needed through the more reliable comparison of performance with SSD storage system for to implement the testing.

## References

- [1] K. -H. Park, J. -k. Choe, Y. -H. Lee, S. -K. Cheong and Y. -S. Kim, "Performance Analysis for DRAM-based SSD system", Korean information technical academic society dissertation, vol. 7, Unit 4, (2009) August, pp. 41-47.
- [2] Y. -H. Kang, J. -H. Yoo and S. -K. Cheong, "Performance Evaluation of the SSD based on DRAM Storage System IOPS", Korean information technical academic society dissertation, vol. 7, Unit 1, (2009), pp. 265-272.
- [3] S. -K. Cheong, Y. -W. Jeong, Y. -J. Jeong and J. -j. Jeong, "Input-Output Performance Analysis of HDD and DDR-SSD Storage under the Streaming Workload", Korean information technical academic society dissertation, (2010) May, pp. 322-325.
- [4] J. Judd and C. Beaucamp, "Building SANs Witch Brocade", syngess, (1996) November, pp. 9-16.
- [5] H. Kim, "Learn to MYSQL Database Programming", Young-jin.com.
- [6] <http://www.tpc.org/tpch/spec/tpch2.13.0.pdf>.
- [7] H. -j. song and Y. -h. Lee, "CO2 & power consumption comperision Analysis of HDD and DRAM-SSD", UCMA 2012 proc., (2012) June, pp. 147-151.

## Authors



### Hyun-Ju Song

2010: Electronics Engineering, Hannam University, Daejon, Korea

2012: M.S. Electronics Engineering, Hannam University, Daejon, Korea

Present: Clerk of Strategic Business Development in DB Communication & Systems

\*major Field: SSD, Server



### Young-Hun Lee

1973: Electronics Engineering, Donga University, Korea

1980: M.S. Electronics Engineering, Soongsil University, Korea

1984: Ph. D. Electronics Engineering, Kyoung Hee University, Korea

Present: Professor in Electronics Engineering, Hannam University, Daejon, Korea

\*major Field: Hybrid Circuit Design