INTERNATIONAL CONFERENCE
INFORMATION AND COMMUNICATION TECHNOLOGIES IN BUSINESS AND EDUCATION

CONFERENCE PROCEEDINGS

Publishing house "Science and economics"
University of Economics - Varna

International Conference

INFORMATION AND COMMUNICATION TECHNOLOGIES IN BUSINESS AND EDUCATION

Conference proceedings

INFORMATION AND COMMUNICATION TECHNOLOGIES IN BUSINESS AND EDUCATION

Proceedings of the International Conference dedicated to the 50th anniversary of the Department of Informatics

2019

University publishing house "Science and economics"
University of Economics – Varna

Published reports have not been edited or adjusted. The authors are solely responsible for the content, originality and errors of their own fault.

With the support of ADASTRA, 411 Marketing Business Services.

ISBN 978-954-21-1004-0

PROGRAMME COMMITTEE

Abdel-Badeeh M. Salem, Egypt
Aijaz A. Shaikh, Finland
Aleksandar Dimov, Bulgaria
Alexander Feldman, USA
Alexander Sergeev, Russia
Antoanela Naaji, Romania
Avram Eskenazi, Bulgaria
Dariusz Mrozek, Poland
Dimitrina Polimirova, Bulgaria
Dmitriy Berg, Russia
Elena Nechita, Romania
Emil Denchev, Bulgaria
Georgi Dimitrov, Bulgaria
Ivan Popchev, Bulgaria
Jeanne Schreurs, Belgium
Julian Vasilev, Bulgaria
Kameliya Stefanova, Bulgaria
Krasimir Shishmanov, Bulgaria
Maciej Czaplewski, Poland
Maciej Pondel, Poland
Maksim Medvedev, Russia
Maria Hristova, Bulgaria
Marian Cristescu, Romania
Marina Medvedeva, Russia
Michael Voskoglou, Greece
Nadezhda Filipova, Bulgaria
Nikola Nikolov, Ireland
Niv Ahituv, Israel
Pavel Petrov, Bulgaria
Rami Malkawi, Jordan
Sara Paiva, Portugal
Silvia Parusheva, Bulgaria
Snezhana Sulova, Bulgaria
Todorka Atanasova, Bulgaria

Valentin Kisimov, Bulgaria
Valery Trofimov, Russia
Veselin Popov, Bulgaria
Violeta Kraeva, Bulgaria
Vladimir Sulov, Bulgaria
Vladimir Zanev, USA
Zdzislaw Polkowski, Poland

ORGANIZING COMMITTEE

Chairman: Prof. Vladimir Sulov, PhD

Members:
Prof. Avram Eskenazi, PhD
Prof. Julian Vasilev, PhD
Assoc. Prof. Todorka Atanasova, PhD
Assoc. Prof. Silvia Parusheva, PhD
Assoc. Prof. Snezhana Sulova, PhD
Assoc. Prof. Pavel Petrov, PhD

Secretary: Tijen Talib

FOREWORD

These proceedings contain the papers of the International Conference "Information and Communication Technologies in Business and Education" which took place at the University of Economics – Varna, Bulgaria, 18 October 2019.

The international scientific conference is dedicated to the **50th anniversary of the Department of Informatics at the University of Economics – Varna.** The conference is also dedicated to the 100th anniversary of the University. The included papers describe recent scientific and practical developments in the field of information and communication technologies, information systems, and their applications in business and education.

The papers in the Proceedings are peer reviewed and are checked for plagiarism.

C O N T E N T S

CHOOSING STORAGE DEVICES WITH CONSIDERATION OF BACKUP AND RESTORE PERFORMANCE IN MS SQL SERVER

Ivan Kuyumdzhiev[1], Radka Nacheva[2]
[1] University of Economics – Varna/Informatics Department, Varna, Bulgaria, ivan_ognyanov@ue-varna.bg
[2] University of Economics – Varna/Informatics Department, Varna, Bulgaria, r.nacheva@ue-varna.bg

Abstract

Importance of the process of restoring a database requires precise examination of all factors that affects it. Considering the decline of price of SSD and it's better read and write speeds it seems like logical choice when reducing downtime is primary concern for the organization. This paper categorizes storage devices in use as well as main backup types in MS SQL Server. Tests are conducted to evaluate the impact of the storage device on backup and restore efficiency. In this regard, **the purpose of the research** is to examine problems associated with selection of a storage device for efficient backup and restore in MS SQL Server. Conducted tests showed some surprising results regarding SSD's efficiency. The reason these findings are important is that they could be used as a guideline for choosing proper storage device for organization's needs.

Keywords: *backup and restore performance, storage media, databases performance tests, MS SQL server.*

INTRODUCTION

The volume of data worldwide is growing enormously, hence the need for storage capacity is increasing. Over five years, the generated data has more than tripled - from 12 ZB for 2014 to the expected 40 ZB by the end of 2019 (Statista, 2019c). Global data center IP traffic for 2018 is 952 EB and until the end of 2019 it is predicting it will be 997 EB only in traditional data centers but the traffic from the cloud data centers is over 10 times larger (Statista, 2019d). Especially considering the fact that more and more businesses are getting oriented towards online presence (Stanoeva, 2017), the data is increasing rapidly, it is important to choose a storage data device for better performance of backup and restore processes.

MS SQL Server is one of the most popular commercial relational database management system. It offers built-in backup and restore functionality, that can mitigate the risk of disastrous events. This functionality includes couple backup types that could be combined in many different strategies. Characteristics of each backup type is outlined by Microsoft regarding factors as database size and amount of changed data but there is no specific guidance on which storage device to be chosen.

In this regard, **the purpose of the research** is to examine problems associated with selection of a storage device for efficient backup and restore in MS SQL Server. On this basis the paper could be used as a guideline for choosing proper storage device for organization's needs.

1. COMPARISION OF DIGITAL DATA STORAGE MEDIA

As an integral element of the information technology sector, the global demand and supply of data storages by the end of 2019 will be around 31,000 EB and 20,000 EB compared to a period of five years ago, when they were approximately 6 times smaller. (Statista.com, 2019a). Worldwide spending on data storage has exceed 51 billion U.S. dollars in 2019 (Horst, 2018). According to other statistics global data will grow 61% to 175 zettabytes by 2025, with as much of the data residing in the cloud as in data centers (Patrizio, 2018). This shows that the need for using storage devices is increasing, including for private users.

Data storage refers to the devices which are used for retaining and archiving digital data that will be used by computer or another device. It could be divided to primary (main), secondary (auxiliary) or tertiary memory. The first type is random access memory (RAM) and its derivatives. The second type are the commonly used hard disks and solid-state drives. The third type are networked or cloud storages. Other bearers of digital information could be magnetic tapes, floppy disks, optical disks (CDs, DVDs, Blu-ray disks), flash drives and flash memory cards. Considering the purpose of this study and the capabilities of each of the digital data media, we are focusing on researching capabilities of secondary memory devices - hard disk drives (HDDs) and solid-state drives (SSDs).

Currently, many of the computers (home or enterprise machines) use both main types of secondary memory media. Computer configura-

tions equipped with both types of devices are often encountered, as home and corporate users prefer to take advantage of both technologies. Statistics show that around 360 million HDDs and 280 million SSDs are expected to sell by the end of 2019, compared to the previous three or four years, when SSDs sales were approximately 4 times less than HDDs (Statista.com, 2019b).

The answer to the question "Which device is better?" cannot be given unequivocally because it depends mainly on the purposes for which each of them will be used. The main difference between them is the technology they are designed for - hard disk drives are magnetic data carriers and solid-state drives are based on flash technology. Through Table 1, we compare the characteristics of the two main media types.

Table 1

Digital Data Storage Media[1]

Characteristic	Hard Disk Drive	Solid-State Drive
Cost	buying a cheap 4TB model - around €0.02 / GB	buying a cheap 1TB model - around €0.09 / GB
Capacity[2]	2 ÷ 10 TB	128 GB ÷ 4 TB
Power Consumption	7 Watt average	2 Watt average
Read Speed	approx. 150 MB/s	over 450 MB/s
Write Speed	approx. 100 MB/s	over 500 MB/s
Time to Access	approx. 2 sec. before it can read/write	read/write immediately
Annualized Failure Rate	1,7%	0,4%
Encryption	Full Disk Encryption	Full Disk Encryption
CPU Power	7%	0,7%

Still, the digital data storage capacity of SSDs is much smaller than HDDs'. However, the price ratio is inversely proportional - SSDs have a

[1] The data in Table 1 are averaged based on tests performed by (UserBenchmark, 2019a), (UserBenchmark, 2019b), (Xu, 2019), (Safford, 2019) and (Klein, 2019). The specifications vary from manufacturer to manufacturer.

[2] Depending if it is a laptop or desktop machine.

significantly higher price than HDDs. The flash based SSDs' technology, as seen in Table 1, enables faster reading and writing of data, faster access and a correspondingly lower annualized failure rate. CPU power consumption on SSDs is also less than on HDDs.

2. BACKUP AND RESTORE IN MS SQL SERVER

MS SQL Server is one of the most used commercial databases systems up to date. It offers a wide variety of options for continuity. To keep the business processes running, the organization can implement solutions such as a clustering, replication, mirroring the database, and log file transfer, so that even if the main server stops working, the requests are sent to failover server (Jorgensen, 2012). Main issue with these strategies is that they require extra financing – including software, hardware and professionals with high level of expertise in the area. This factor is more than enough for majority of the organizations to look for cheaper alternatives – one of which is using the built-in backup and restore functionality in MS SQL Server Express.

Main database backup types can be considered depending on whether clients have access to the data during the process. With offline (cold) backup, the database management system suspends access to the archived object and all applications that use it must wait for the process to complete. Online (hot) backup enables the database to remain operational and allow clients to perform the necessary operations without knowing that backup is currently underway. Considering the fact that organizations prefer to keep systems running, hot backup types are examined in the following tests.

Since MS SQL Server version 7 there are four primary database backup types (Desai, 2000):

- Full Backup – backups all data and objects part of the database.
- Differential Backup – it only backs up database changes that have occurred since the last full backup to date, and therefore runs in less time than a full backup would run at the same point in time.
- Log Backup – copies all operations recorded in the transaction log performed since the last log backup.
- Partial Backup - Backing up file groups or individual files is used when the database is so large that the full backup goes beyond ac-

ceptable organization timeframes. This option is added in MS SQL Server version 2005 and backups all file groups except marked as read-only.

Backup plan determines the intensity, sequence, and types of backup that apply to the database. It depends on various factors: what is the amount of information available, at what periods and how it enters the database, what part of it could be recovered from other sources, etc. Recovery plan depends on the backup strategy implemented. Time needed to execute a restore command is similar to time needed for corresponding backup type – fastest when using log backup and slowest with full backup.

Considering wide variety of backup types and strategies test should be conducted to measure their performance with different settings for used storage device.

3. COMPARISON OF THE PERFORMANCE OF DIFFERENT STORAGE MEDIA WITH MS SQL SERVER

To conduct the tests, we chose to use two of the fastest and most reliable storage devices examined in the first section – HDD and SSD. To properly compare their efficiency, we created a web-based application that executes commands to MS SQL Server instance and records the results of their execution. Database in use is a variant of Northwind sample database. To observe if the characteristics of each backup and restore type is preserved with different size of the database, the application executes insert queries with every iteration. The algorithm in use is increasing number of execution of steps 2 and 3 at every iteration. This means that at the last step there are 10 log backups, containing database incremented 10 times. After the end of this the experiment is repeated using an HDD instead of SSD storage device. The program logic takes the following steps:

1) create full database backup;

2) increasing database size by adding 3 000 records in tables production.TransactionHistory, person.person and sales.SalesOrderDetail ;

3) create log backup;

4) create differential backup;

5) create full backup;

6) restore full backup executed in step 1;

7) restore log backups executed in step 3;
8) restore full backup executed in step 1;
9) restore differential backup from step 4;
10) restore full backup executed in step 5;

Hardware testing environment is close to the one that a low budget organization would be using: CPU AMD FX 6100 3.30GHz, RAM 4GB DD3-1333, Samsung SSD 850 EVO and SATA III Hard Disk. Software in use is as follows: Windows 10 64bit, PHP 5.4.7, Apache Server 2.4.3.

Below are the results.

Table 2

HDD Log Backup and Restore results

Iteration	Log backup files	Log Backup			Log Restore	
		Time (s)	Backup Size	Speed (MB/s)	Time (s)	Speed (MB/s)
1	1	0,70	63,37	62,76	4,43	9,94
2	2	1,06	89,72	83,36	5,19	16,95
3	3	1,81	133,25	73,03	6,97	18,94
4	4	2,30	170,79	76,67	8,63	20,39
5	5	3,02	212,70	72,93	9,84	22,35
6	6	3,29	254,74	80,32	11,47	23,02
7	7	3,89	298,29	79,20	13,35	23,07
8	8	4,39	348,46	80,12	15,32	22,98
9	9	5,55	425,52	71,38	18,06	21,93
10	10	5,91	469,55	74,47	20,43	21,54

Table 3

SSD Log Backup and Restore results

Iteration	Log backup files	Log Backup			Log Restore	
		Time (s)	Backup Size	Speed (MB/s)	Time (s)	Speed (MB/s)
1	1	0,55	63,06	80,62	4,37	10,06
2	2	0,78	89,80	112,45	5,10	17,25

3	3	1,18	131,97	111,96	6,68	19,77
4	4	1,48	170,51	119,30	8,62	20,43
5	5	2,54	212,93	86,55	10,32	21,31
6	6	3,09	254,75	85,45	12,59	20,97
7	7	3,43	297,22	89,91	14,67	20,99
8	8	4,26	347,56	82,65	17,02	20,68
9	9	5,05	424,55	78,49	20,19	19,62
10	10	5,61	468,91	78,49	22,92	19,19

Table 4

HDD Differential Backup and Restore results

Iteration	Differential Backup			Differential Backup Restore	
	Time (s)	Backup Size	Speed (MB/s)	Time (s)	Speed (MB/s)
1	1,55	140,13	13,53	1,84	11,42
2	1,61	164,75	26,03	2,16	19,46
3	2,01	195,82	31,30	2,32	27,20
4	2,35	224,38	35,81	2,65	31,70
5	2,63	254,82	39,88	2,91	36,08
6	3,01	287,63	41,84	3,25	38,74
7	3,28	324,88	44,81	3,60	40,84
8	3,56	365,82	47,19	4,30	39,06
9	4,35	420,00	43,47	4,54	41,64
10	4,84	467,25	43,39	5,40	38,90

Table 5

SSD Differential Backup and Restore results

Iteration	Differential Backup			Differential Backup Restore	
	Time (s)	Backup Size	Speed (MB/s)	Time (s)	Speed (MB/s)
1	1,43	139,04	14,68	1,73	12,14

2	1,54	164,23	27,36	1,98	21,24
3	1,82	196,17	34,54	2,20	28,58
4	1,97	223,04	42,59	2,48	33,92
5	3,13	254,04	33,50	2,73	38,43
6	3,49	287,29	36,07	3,53	35,71
7	3,81	323,92	38,61	4,28	34,35
8	3,82	365,23	43,94	4,74	35,44
9	4,61	419,98	41,04	5,07	37,26
10	5,61	466,10	37,41	6,22	33,78

Table 6

HDD Full Backup and Restore results

Iteration	Full Database Backup from step 1			Full Database Backup Restore	
	Time (s)	Backup Size	Speed (MB/s)	Time (s)	Speed (MB/s)
1	1,85	190,23	102,63	1,92	121,73
2	2,19	234,13	106,78	2,49	109,77
3	2,54	273,63	107,61	2,72	121,47
4	3,20	330,94	103,42	3,34	120,27
5	3,77	401,57	106,61	3,94	124,17
6	4,51	489,32	108,57	4,51	131,48
7	5,45	593,25	108,90	5,64	127,06
8	6,57	717,13	109,19	6,92	124,08
9	7,71	858,25	111,29	7,80	132,23
10	9,46	1031,69	109,02	9,56	127,73

Table 7

SSD Full Backup and Restore results

Itera-tion	Full Database Backup from step 1			Full Database Backup Restore	
	Time (s)	Backup Size	Speed (MB/s)	Time (s)	Speed (MB/s)
1	1,32	190,21	144,28	1,59	146,81
2	1,84	233,23	126,60	2,02	133,45
3	2,08	273,17	131,13	2,43	135,91
4	2,59	330,98	127,68	2,93	136,96
5	2,96	400,98	135,58	3,51	139,12
6	3,63	488,67	134,45	4,03	147,03
7	5,02	593,10	118,03	5,55	128,98
8	5,61	716,54	127,68	7,40	115,95
9	6,92	857,85	123,88	9,06	113,86
10	8,89	1031,23	115,94	10,82	112,72

Calculating average speeds leads to interesting results. Table 8 shows that surprisingly speed of differential backup and restore is better when HDD storage type is used. Speed of both full backup and restore is slightly better with SSD but once again log backup and restore speeds are not undeniably faster with the more expensive configuration.

Table 8

Average results

Data storage	Operation Type	Log backup/ restore	Differential backup / restore	Full backup / restore
HDD	Backup	75,424 MB/s	**36,725 MB/s**	107,402 MB/s
HDD	Restore	**20,111 MB/s**	**32,504 MB/s**	123,999 MB/s
SSD	Backup	**92,587 MB/s**	34,974 MB/s	**128,525 MB/s**
SSD	Restore	19,027 MB/s	31,085 MB/s	**131,079 MB/s**

CONCLUSION

The publication researches the different types of storage media and their combination with backup and restore in MS SQL Server. The analysis of the theoretical capabilities and qualities of the carriers shows that the SSDs has more advantages than HDDs. For example, better read, write and access performance, lower failure rate and CPU power consumption. On the other hand, we found that the actual speed of different backup types is not as better as it should be expected with SSD technology.

That is why HDD storage type should be used when the database is as large as several gigabytes or backup strategy relies on usage of differential backups. SSD should be used with databases as large of several thousand gigabytes with backup strategy using many full and log backups.

REFERENCES

1. DESAI, A. (2000). *SQL Server 7 Backup & Recovery*. Mcgraw-Hill Osborne Media.
2. HOLST, A. (2018). Data Storage - Statistics & Facts. [Online] Available from: https://www.statista.com/topics/3150/data-storage/ [Accessed 08/08/2019].
3. JORGENSEN, A., Wort, S., LoForte. R. and Knight, B. (2012). *Professional Microsoft SQL Server 2012 Administration*. Wrox.
4. KLEIN, A. (2019). Backblaze Hard Drive Stats Q2 2019. [Online] Available from: https://www.backblaze.com/blog/hard-drive-stats-q2-2019/ [Accessed 08/08/2019].
5. PATRIZIO, A. (2018). IDC: Expect 175 zettabytes of data worldwide by 2025. [Online] Available from: https://www.networkworld.com /article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html [Accessed 08/08/2019].
6. SAFFORD, M. (2019). How to Buy the Right SSD: A Guide for 2019. [Online] Available from: https://www.tomshardware.com/ reviews/ssd-buying-guide,5602.html [Accessed 08/08/2019].
7. STANOEVA, I. (2017) Creating "Digital Marketing" Curriculum For "Art Management" Programmes. *Proceedings of The Second Conference on Innovative Teaching Methods (ITM 2017)*, 1(1), pp. 5-13.

8. STATISTA.COM. (2019a). Data storage supply and demand worldwide, from 2009 to 2020 (in exabytes). [Online] Available from: https://www.statista.com/statistics/751749/worldwide-data-storage-capacity-and-demand/ [Accessed 08/08/2019].
9. STATISTA.COM. (2019b). Shipments of hard and solid state disk (HDD/SSD) drives worldwide from 2015 to 2021 (in millions). [Online] Available from: https://www.statista.com/statistics/285474/hdds-and-ssds-in-pcs-global-shipments-2012-2017/ [Accessed 08/08/2019].
10. STATISTA.COM. (2019c). Volume of data/information created worldwide from 2010 to 2025 (in zetabytes). [Online] Available from: https://www.statista.com/statistics/871513/worldwide-data-created/ [Accessed 08/08/2019].
11. STATISTA.COM. (2019d). Global data center IP traffic from 2012 to 2021, by type (in exabytes per year). [Online] Available from: https://www.statista.com/statistics/227259/forecast-of-global-data-center-ip-traffic-by-type/ [Accessed 08/08/2019].
12. USERBENCHMARK. (2019a). Speed test your SSD in less than a minute. [Online] Available from: https://ssd.userbenchmark.com/ [Accessed 08/08/2019].
13. USERBENCHMARK. (2019b). Speed test your HDD in less than a minute. [Online] Available from: https://hdd.userbenchmark.com/ [Accessed 08/08/2019].
14. XU, E. et. al. (2019). Lessons and Actions: What We Learned from 10K SSD-Related Storage System Failures. *Proceedings of the 2019 USENIX Annual Technical Conference*. pp. 961-975.