ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2022. № 1 (19) ISSN 2524-2296 (online)

UDC 004.75 DOI: https://doi.org/10.30837/ITSSI.2022.19.027

S. MINUKHIN

PERFORMANCE STUDY OF THE DTU MODEL FOR RELATIONAL DATABASES ON THE AZURE PLATFORM

When solving problems of working with relational databases on cloud platforms, the problem arises of choosing a specific model to ensure the performance of executing queries of varying complexity. The **object** of research is the processes of implementing various types of queries to relational databases within the framework of the DTU **purchase** model of the MS Azure platform. The **subject** is methods for evaluating the performance of work with relational databases based on the timing of query execution and indicators of the load on the resources of the cloud platform. The **aim** of the study is to develop a system of indicators for monitoring the current state of work with the database for reasonable decision-making on the choice of a certain price category of the DTU model of the MS Azure cloud service, which will optimize the results of working with the database. platforms Achieving the set goals involves the following **tasks**: to analyze modern tools and services for working with databases, in particular relational databases, on Azure and AWS cloud platforms, the features of their application and implementation; develop software for generating test relational databases of different sizes; test the generated databases on a local resource; taking into account the characteristics of the levels of the Azure DTU model, develop a new system of performance indicators, which includes 2 groups - time indicators and indicators of the load on existing platform resources; develop and implement queries of varying complexity for the generated test database for different levels of the DTU model and analyze the results. **Methods.** The following methods were used in the research: methods of relational database design; methods of creating queries in SQL-oriented databases with any number of tables; methods of creating and migrating data to cloud platforms; methods of monitoring the results of queries based on time and resource indicators; methods of generating test data for relational databases; system approach for complex assessment and analysis of productivity of work with relational databases. **Results.** On the basis of the developed scorecard used for the current analysis of the processes of working with relational databases of the MS Azure platform, numerous experiments were carried out for different levels of the model for simple and complex queries to a database with a total volume of 20 GB: loading of DTU model levels when executing various queries, the influence of model levels DTU Azure SQL database on the performance of simple and complex queries, the dependence of the execution time of various queries on the load of the CPU and the speed of write/read operations for different levels of the model. **Conclusions.** The results of the experiments allow us to conclude that the levels of the DTU model - S3 and S7 - are used to generate test data of various sizes (up to 20 GB) and execute database queries. The practical use of the proposed indicators to evaluate the results of applying the DTU model will improve the efficiency of decision-making on choosing the model level when implementing various queries and generating test data on the Azure cloud platform. The developed set of indicators for working with relational databases on the Azure cloud platform expands the basis of the methodological framework for evaluating the performance of working with relational databases on cloud platforms by analyzing the results of executing the simple and complex database queries on the resources involved.

Keywords: cloud platform; relational database; DTU purchase model; indicators of time and workload; data generator; test data; request complexity.

Relevance

One of the most powerful and relevant trends in the use of modern information and communication systems of different spheres of application is the creation and implementation of database technologies. The use of these technologies allows a significant increase in the efficiency of information systems such as the Internet of Things, cloud platform services, distributed data processing systems – in distributed computing systems (e.g., grid systems for intensive data processing), clusters with the ability to connect data processing centers (DPCs) using service-oriented archetypes, etc. One of the tasks here is to create and test relational databases based on different tools. Relational databases are now not only one of the most popular technologies in terms of their versatility in solving a significant number of tasks of various purposes, but also a sufficiently large variety of their proposals in the market of information services. DB development is an important stage of its use – the creation of logical and physical models is accompanied by constant monitoring of the status of the existing database for optimization of their work in conditions of scale and dynamic changes in the number of users. This raises the problem of locating and storing the database on remote servers, providing access to them, which requires compliance with the requirements for communication systems (the capabilities of Internet resources), data center capacity, as well as the associated financial costs. Therefore, there is a need for a general solution to this problem and related partial tasks. One of the possible ways is to use the services of the data storage platforms designed to work with databases and data warehouses

A brief overview of cloud platform services

Cloud computing is one of the trends of modern information technology, which is rapidly developing and is increasingly used at different levels of management. Cloud computing is the provision of computing capacities, warehouses, databases, resources and other resources of cloud service platforms via the Internet [1]. Cloud computing has the following advantages [2]:

1. Capital expenditures are being converted into changes. There is no need to invest large sums of money in data processing centers (DPC) and servers, not knowing in advance what capacities will be needed in the future. You can pay only for the actual use of computing resources.

2. Significant savings in large volumes. When using cloud computing, you can achieve a lower variable cost than when creating your own computing infrastructure.

3.There is no need to predict what amount of infrastructure resources will be needed. When making a

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2022. No. 1 (19)

decision on the volume of resources for program implementation, often in the final time you can run into downtime of expensive resources or lack of capacity.

4. Increasing the speed and flexibility of developing the features.

5. Lack of costs for launching and maintaining DPC resources.

The global market for cloud solutions and services is growing quite rapidly, so it is difficult to predict the rate of its increase in practice. However, the existing data record the same trends: the rapid growth rate of cloud computing costs, as well as the associated market for services, data centers and data traffic in these systems [1, 2]. In April 2018, Gartner analysts [3] published the results of a study of the global public cloud market. The cost of these solutions is increased (table 1) due to infrastructure services (IaaS model – DigitalOcean, Linode, Rackspace, Amazon Web Services (AWS), Cisco Metapod, Microsoft Azure, Google Compute Engine (GCE).) And software provided as a service (SaaS model - AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos, OpenShift) [1, 2, 4]. In 2021, the cost of cloud services is projected to increase almost 2 times compared to 2017. The largest segment of the public cloud market remains SaaS, which, according to Gartner, by 2021 will account for 45% of software costs in the world. The fastest growing segment of analytics is the IaaS service model, the volume of which in 2018 increased by almost 40%.

Table 1. *World forecast of cloud services (billion dollars)*

Cloud services	2017	2018	2019	2020	2021
Business Process as a Service (BPaaS)	42,6	46,4	50,1	54,1	58,4
Platform as a Service (PaaS)	11,9	15,0	18,6	22,7	27,3
Software as a Service (SaaS)	60,2	73,6	87,2	101,9	117,1
Cloud management and security services	8,7	10,5	12,3	14,1	16,1
Infrastructure as a Service (IaaS)	30,0	40,8	52,9	67,4	83,5
Total market	153,5	186,4	221,1	260,2	302,5

In recent years, cloud services such as DBaaS (Database as a Service, AWS) and MWaaS (Middleware as a Service) have appeared, which are a type of PaaS. Using DBaaS, the user can access any type of database on request and quickly deploy databases on any hardware in the environment of the selected software platform (operating system). MWaaS is a specialized cloud solution for companies. This service provides access to a comprehensive platform with the appropriate infrastructure to service enterprise programs and security tools. Using MWaaS, the user can quickly prepare a specific software environment to perform the appropriate tasks.

The highest rates of DBaaS use are based on the Azure SQL database [5, 6, 9] and RDS AWS [7, 8]. Tools and technologies related to the use of the following models have been developed: deployment models, migration models, and purchase models. At present, these models are in fact the stages of creation, migration and use of a certain price category of models used for relational DB on cloud platforms [10-12].

These models of services for working with relational DB on cloud platforms to justify the choice of resources when working with certain DB do not fully take into account the nature of work with DB, namely, the volume of DB (number of records in tables), its filling method (location) migration from a local resource or direct deployment to available Azure resources), DB complexity (number and volume of tables), DB query types - simple and complex queries – that significantly affect the time and efficiency of working with DB. This is especially important when scaling DB and when using so-called "cold" and operational data in the process of analytical data processing of enterprises and institutions. Therefore, it is key and promising to develop indicators for quantitative (qualitative) evaluation of the results of various types of queries to DB, taking into account practical issues related to the scalability of systems (data volumes) based on existing price categories of cloud platform models.

Analysis of publications and setting the task

Let's analyze the state of the problem on the basis of DB servers on the Azure and AWS cloud platforms [5 – 8].

Microsoft Azure cloud platform: characteristics and models of work with relational DB [5, 6].

Azure database SQL resides in the Azure cloud and is part of the platform-as-a-service (PaaS) server model. The SQL Azure database allows you to easily purchase a fully managed DBMS PaaS core that meets your productivity and cost requirements. Depending on the model of SQL Azure data base deployment, one can choose the model that meets the user's needs [5, 6]:

1. Based on DTU units (Database Transaction Units). This model offers logical servers in SQL Azure database.

2. On the basis of virtual cores (vCore) – this model offers logical servers in SQL Azure database and kerned instances in SQL Azure database.

Сучасний стан наукових досліджень та технологій в промисловості. 2022. № 1 (19) ISSN 2522-9818 (print) ISSN 2524-2296 (online)

Table 2. *Comparison of types of purchase models on the Microsoft Azure cloud platform*

Purchase model	Description	Optimal
Purchase model based on DTU units	It is based on complex evaluation of computing resources, storage resources and resources for input-output operations. The productivity levels are expressed in data base transaction units (DTU) for individual databases, and for elastic pools - in elastic data base transaction units (eDTU).	For customers who need simple, pre-configured resource options.
Purchase model based on virtual cores	This model allows to independently choose computing resources and storages. In addition, it allows you to save costs due to the "Advantage of Azure Hybrid Usage for SQL Server" program.	For customers who value flexibility, control and transparency.

The DTU-based purchase model offers a number of pre-configured computing resource packages and included storage capacity to ensure different levels of productivity of add-ons. Data base transaction unit (DTU) is an integral indicator of the capacity (productivity) of the CPU, RAM, read and write operations and memory on a hard disk drive (HDD, SSD). The physical characteristics (CPU, memory, input/output operations) associated with each DTU measurement are calibrated using the performance test which has the operating load of the real data base. The productivity test consists of a set of transactions of different types, which are executed in a data model containing a number of tables and data types. When using a model based on DTU units, users can choose between different service levels (Basic, Standard and Premium) for certain databases and elastic pools [5] (tables 3, 4).

Table 3. *Levels of DTU-based purchase models*

Indicator	Basic	Standard	Premium
Target workload	Development and application in the working environment.	Development and application in the working environment.	Development and application in the working environment.
Service level agreement with a guarantee of continuous operation time	99.99%	99.99%	99.99%
Backup storage	7 days	35 days	35 days
CPU	Low	Low, medium, high	Medium, high
I / O bandwidth (approx.)	2.5 I/O operations on the DTU	2.5 I/O operations on the DTU	48 I/O operations on the DTU
I / O delay (approximately)	5 ms (read), 10 ms (write)	5 ms (read), 10 ms (write)	2 ms (read and write)
Indexing columnstore	Not available	S3 and above	Supported
In-memory OLTP	Not available	Not available	Supported
Maximum storage size	2 GB	1 TB	4 TB
Maximum number of DTUs	5	3,000	4,000

Table 4. *Characteristics of levels of the purchase model based on DTU units*

Level, model		DTU	Storage volume included, GB	Max. storage volume, GB	Max. OLTP volume in memory, GB	Max. number of queries, that are performed simultaneously, in 1 second.	Max. number of simultaneous sessions
1		2	3	4	5	6	7
Basic	B	5	2	2	N/A	30	300
Standard	S0	10	250	250	N/A	60	600
	S1	20	250	250	N/A	90	900
	S2	50	250	250	N/A	120	1200
	S3	100	250	250, 500, 750, 1024	N/A	200	2400
	S4	200	250	250, 500, 750, 1024	N/A	400	4800
	S6	400	250	250, 500, 750, 1024	N/A	800	9600
	S7	800	250	250, 500, 750, 1024	N/A	1600	19200
	S9	1600	250	250, 500, 750, 1024	N/A	3200	30000
	S12	3000	250	250, 500, 750, 1024	N/A	6000	30000

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2022. No. 1 (19)

The virtual core purchase model (vCore) provides flexibility, control, and transparency in the consumption of individual resources. A virtual core is a logical central processor (CPU) that can select hardware generation and physical hardware characteristics (such as number of cores, memory capacity, and storage size). The model allows you to **independently** choose computing and storage resources, ensure performance at the local environment level and optimize costs and uses two levels of services: general purpose and critical for business [5]. Service levels differ in the range of performance levels, the structure to ensure a high level of availability, the method of providing isolation from failures, types of storage and the range of input- output operations (table 5).

Table 5. *Levels of vCore-based purchase models*

Characteristics	General purpose	Critically important for business
1	2	3
Optimal for	Most workloads. Offers budget balanced and scalable variants of computing and storage resources.	Business applications with high demands on input-output operations. Offers the highest resistance to failures due to the use of several isolated replicas.
Computing services	- 4th generation: from 1 to 24 virtual cores;	- 4th generation: from 1 to 24 virtual cores;
Memory	- 5th generation: from 1 to 80 virtual cores.	- 5th generation: from 1 to 80 virtual cores.
Input/output bandwidth (speed)	- 4th generation: 7 GB per core;	- 4th generation: 7 GB per core;
Storage service	- 5th generation: 5.1 GB per core.	- 5th generation: 5.1 GB per core
Availability	500 input/output operations per second on the virtual core with a maximum of 7,000 operations per second.	5,000 input/output operations per second per core with a maximum of 200,000 operations per second
Backups	- Premium level remote control storage;	- Local storage on SSD storage devices;
In memory	- individual database: 5 GB - 4 TB;	- separate database: 5 GB - 1 TB;

Amazon Web Services cloud platform: characteristics and models of working with relational DB [7, 8].

Amazon Relational Database Service (Amazon RDS) is a web service that allows you to set up, use, and scale relational databases in Amazon Web Services cloud. Amazon RDS is available in the form of database instances of several types: optimized for working with memory, for a high productivity or implementation of the write/read operations (table 6) and offers a choice of six database cores, including Microsoft SQL Server. When working with the SQL Server database on the AWS platform, the deployment model, the migration model and the purchase model are also used.

Table 6. *SQL Server database instance classes of the Amazon RDS service*

Instance type	vCPU processors	Memory, GB	Optimized for IOPS (write/read speed)	Instance type
Standard - the last generation				
db.m4.large	2	8	Yes	Medium
db.m4.xlarge	4	16	Yes	High
db.m4.2xlarge	8	32	Yes	High
db.m4.4xlarge	16	64	Yes	High
db.m4.10xlarge	40	160	Yes	10 gigabits
Standard - the previous generation				
db.m3.medium	1	3,75	–	Medium
db.m3.large	2	7,5	–	Medium
db.m3.xlarge	4	15	Yes	High
db.m3.2xlarge	8	30	Yes	High
Memory optimization - current generation				
db.r3.large	2	15	–	Medium
db.r3.xlarge	4	30,5	Yes	Medium
db.r3.2xlarge	8	61	Yes	High
db.r3.4xlarge	16	122	Yes	High
db.r3.8xlarge	32	244	–	10 gigabits
Microinstances				
db.t2.micro	1	1	–	Low
db.t2.small	1	2	–	Low
db.t2.medium	2	4	–	Medium
db.t2.large	2	8	–	Medium

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2022. № 1 (19) ISSN 2524-2296 (online)

You do not need to purchase Microsoft SQL Server licenses separately for service models that include a license. The cost of licensed models includes the cost of software, basic hardware resources and Amazon RDS control capabilities. This model supports: Express Edition, Web Edition, Standard Edition and Enterprise Edition [7]. The free use of AWS for Amazon RDS allows you to use micro DB instances for free in one availability zone running SQL Server Express Edition. By default, users can create up to 40 instances of the Amazon RDS database, including up to 10 instances of the SQL Server DB under the licensing model. Therefore, the functionality of using such a service is quite limited from the point of view of the average user of cloud platforms.

Thus, the analysis and literary sources revealed the advantages of the Microsoft Azure platform over the Amazon Web Services platform:

1. Azure gives 200 dollars for 30 days as a free subscription, which allows you to use a wide range of services during this period. This allows us to investigate the impact of increasing the price category of the database on its productivity. AWS does not have such a possibility.
2. Access to the Azure database can be obtained using the built-in editor of records in Azure Portal, Microsoft SQL Server Management Studio, and Microsoft Visual Studio. Access to the SQL Server database, located on the AWS platform, is available only through Microsoft SQL Server Management Studio.
3. With more regions available than any other service provider, the Azure platform offers scale and regions to keep additions as close as possible to users.
4. Microsoft Azure has a more intuitive interface for creating and working with databases.
5. Microsoft Azure provides databases with extensive real-time monitoring capabilities, automated performance tuning based on inverse access models and requests performance analysis.
6. Microsoft Azure has a more flexible interface for working with relational databases and provides the ability to easily set up and smoothly switch from one level (price category) to another in the DTU model (within S0-S12) (see table 3, table 4) as compared to the vCore model.

The purpose of this article is to develop indicators for assessing and analyzing the performance of the relational database service and to investigate the use of price categories of existing models for different types of queries (on the basis of the Azure cloud platform).

Materials and methods

Algorithm of the research conducting.

The algorithm was developed according to the given sequence of stages of work with the database on the cloud platform **with such modification**: designing and deploying a relational database, developing queries of varying complexity to the database, generating test data of slow database volume and conducting experiments for the DTU purchase model at the S0-S12 levels of the SQL database service.

The algorithm is implemented by the following steps.

Step 1 Creation of SQL Server database (projecting).

Step 2 Deployment of SQL Server database on local resource.

Step 3. Filling SQL Server database with test data using random number generator. The generator allows to dynamically change the number of records in the created database tables for evaluation of the influence of their volume on the performance indicators of the model DTU.

Step 4: Development of simple and complex types of queries for further evaluation of their results for some level of the DTU Azure SQL database model.

Step 5: Testing the created database on a local resource.

Step 6. Development of relational database productivity indicators for DTU Azure SQL database model taking into account characteristics of the purchase model levels on the basis of DTU units (see tables 3 and 4).

Step 7: Analysis of the results obtained by performing different types of queries to the database, and formulating recommendations for choosing the price category of the purchase model DTU Azure SQL database for different types of queries.

Database design.

2 tables with a size of 1,000,000 records each (10 GB) were designed for testing (tables 7, 8).

Table 7. *Customer*

Field	Data type
Customer ID	unique identifier, Primary key
Email	nvarchar(100)
Password	nvarchar(50)
Creation Date	Date time

Table 8. *Permission*

Field	Data type
Permission ID	unique identifier, Primary key
Customer ID	unique identifier, Foreign key
CanRead	Bit
CanWrite	Bit

Performance indicators for relational databases for the DTU Azure SQL database purchase model.

To monitor the performance of the Azure SQL database and the managed instance of Azure SQL, a system for monitoring CPU usage and CPU usage is used to control a certain level of the database model on an ongoing basis. To do this, the Azure SQL database and the managed instance of Azure SQL in the process can provide resource metrics that can be viewed on the Azure portal or using Azure Data Studio or SQL Server Management Studio (SSMS). However, these indicators are not enough for a comprehensive analysis of the use of certain levels of models of database work on the Azure platform, for which it is proposed to expand their composition taking into account the characteristics of the DTU model.

Based on the analysis of characteristics [9, 10, 13] and indicators of levels (price categories) of the purchase model DTU Azure SQL database [14, 15] the following

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2022. No. 1 (19)

indicators are proposed for evaluation and analysis of its performance (table 9).

Table 9. *Database performance indicators for the DTU purchase model*

№	Characteristic	Description	Comment
1	Request type	Determined by the level of complexity of the query (simple, complex)	Determined by the number of tables and poles in the query
2	Query execution time	Query type execution time at DTU model level, sec	It is determined by the execution time of a certain type of query at the model level
3	The number of selected records in the query	Number of selected records from fixed-volume database tables for simple and complex queries	The result of the query,
4	The ratio of the number of selected records to the total number of records in the tables,%	The effectiveness of the query	absolute
5	DTU load,%	DTU model level load	The result of the query,
6	CPU load,%	CPU usage for DTU model level	relative
7	Load OP, GB	RAM load capacity, GB	Determined by the load level of the DTU model for the query type
8	Number of operations per 1 second	Number of query execution operations per 1 sec. (input / read speed)	The CPU load for the DTU model level for the request type is determined
9	Number of DTUs on request	Must not exceed 80% for all levels of the DTU model	Determined by the amount of OP used to execute the query

The first 4 indicators are time indicators and performance of requests; the rest are indicators of resource utilization in the execution of requests.

Assessing the impact of DTU Azure SQL database model levels on the performance of simple and complex queries to relational databases.

3 price categories (levels) of the model were selected to conduct research to assess the impact of price categories of the DTU purchase model: S0, S3 and S7.

The results of execution of different types of queries for performance indicators of the DTU model "Execution time", "Number of selected records per query" and "Ratio of the number of selected records to their total number" are summarized in tables 10 – 12 and in fig. 1.

Table 10. *Results of execution of queries to the database for the price category S0*

№ of the query	Description of the query	Table	Execution time, sec.	Number of selected records on the query	The ratio of the number of selected records to their total number,%
	Simple queries (SQ)				
SQ 1	SELECT * FROM dbo.Customers	dbo.Customer	79	1001002	50,03
SQ 2	SELECT * FROM dbo.Customers WHERE Email LIKE 'A1%'	dbo.Customer	2	706	0,04
SQ 3	SELECT * FROM dbo.Customers WHERE CreationDate < GETDATE()	dbo.Customer	93	201002	10,05
	Complex queries (CQ)				
CQ 1	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId	dbo.Customer, dbo.Permission	105	1001002	50,03
CQ 2	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.Password LIKE 'A%' AND per.CanRead = 1	dbo.Customer, dbo.Permission	54	13827	0,69
CQ 3	SELECT * FROM dbo.Customers AS cus LEFT JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.CreationDate < GETDATE()	dbo.Customer, dbo.Permission	103	201002	10,05

Сучасний стан наукових досліджень та технологій в промисловості. 2022. № 1 (19) ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The results of queries for the S3 category of the model are given in table 11.

Table 11. *Results of database queries for price category S3*

№ of the query	Description of the query	Table	Execution time, sec.	Number of selected records on the query	The ratio of the number of selected records to their total number,%
Simple queries (SQ)					
SQ 1	SELECT * FROM dbo.Customers	dbo.Customer	73	1001002	50,03
SQ 2	SELECT * FROM dbo.Customers WHERE Email LIKE 'A1%'	dbo.Customer	1	706	0,04
SQ 3	SELECT * FROM dbo.Customers WHERE CreationDate < GETDATE()	dbo.Customer	99	201002	10,05
Complex queries (CQ)					
CQ 1	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId	dbo.Customer, dbo.Permission	97	1001002	50,03
CQ 2	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.Password LIKE 'A%' AND per.CanRead = 1	dbo.Customer, dbo.Permission	43	13827	0,69
CQ 3	SELECT * FROM dbo.Customers AS cus LEFT JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.CreationDate < GETDATE()	dbo.Customer, dbo.Permission	54	201002	10,05

The results of queries for the S7 model category are given in table 12.

Table 12. *Results of database queries for price category S7*

№ of the query	Description of the query	Table	Execution time, sec.	Number of selected records on the query	The ratio of the number of selected records to their total number,%
Simple queries (SQ)					
SQ 1	SELECT * FROM dbo.Customers	dbo.Customer	53	1001002	50,03
SQ 2	SELECT * FROM dbo.Customers WHERE Email LIKE 'A1%'	dbo.Customer	1	706	0,04
SQ 3	SELECT * FROM dbo.Customers WHERE CreationDate < GETDATE()	dbo.Customer	74	201002	10,05
Complex queries (CQ)					
CQ 1	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId	dbo.Customer, dbo.Permission	84	1001002	50,03
CQ 2	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.Password LIKE 'A%' AND per.CanRead = 1	dbo.Customer, dbo.Permission	37	13827	0,69
CQ 3	SELECT * FROM dbo.Customers AS cus LEFT JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.CreationDate < GETDATE()	dbo.Customer, dbo.Permission	42	201002	10,05

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2022. No. 1 (19)

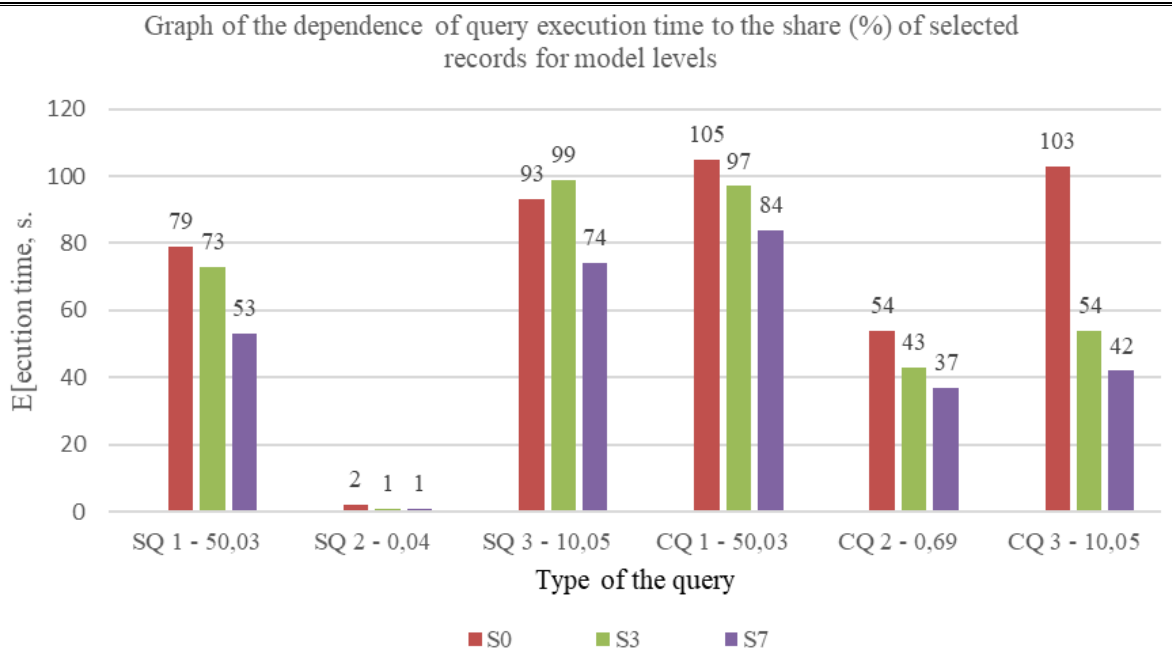

Fig. 1. Graph of query execution time dependence on the share of selected records for different queries, %, for DTU model levels

From the graph we can conclude that the execution time of different types of queries for their performance for different levels of the model varies in the range: for **simple queries** – 1.3 - 1.5 times, for **complex** queries - 1.25 - 2.5 times. Thus, increasing the level of the model allows to increase the productivity of the database in the range of 130% - 250% (for databases of 20 GB), and the complexity of the query is determined by the relative indicator – **the number of received records for queries** to the total database volume **during the time interval of their implementation.**

Further analysis of the model's performance will be based on the **indicators of resource load**: "DTU load", "CPU load", "RAM load" and "Number of operations per 1 sec." (see table 9).

The load levels of the DTU model when generating database test data are given in table 13.

Table 13. *Indicators of the load levels of the DTU model when generating test data of the database*

Model level	DTU load,%	CPU load, %
S0	100	15,57
S3	100	3,36
S7	55	0,72

The load levels of the DTU model when performing various queries are given in table 14.

Table 14. *Load rate levels of the DTU model when executing different types of queries*

Model level	DTU load,%	CPU load, %	RAM load, GB.	Number of operations per 1 sec.
1	2	3	4	5
Query SQ1				
S0	44	4,2	1,17	12671
S3	19	1,82	1,17	13712
S7	1,7	0,08	1,17	18887
Query SQ2				
S0	52	3,8	0,9	353
S3	23	3,85	0,9	706
S7	2	0,31	0,9	706
Query SQ3				
S0	68	5,2	1,05	2161
S3	22	1,54	1,05	2030
S7	1,8	0,9	1,05	2716

Сучасний стан наукових досліджень та технологій в промисловості. 2022. № 1 (19) ISSN 2522-9818 (print) ISSN 2524-2296 (online)

The end **Table 14**

1	2	3	4	5
Query CQ1				
S0	94	4,95	1,43	9533
S3	37	1,32	1,43	10320
S7	3,7	0,35	1,43	11917
Query CQ2				
S0	83	5,74	1,22	256
S3	32	3,57	1,22	322
S7	1,4	0,21	1,22	374
Query CQ3				
S0	81	4,3	1,14	1951
S3	31,7	2,69	1,14	3722
S7	2,1	0,31	1,14	4786

Thus, for different types of queries, the value of the key indicator -load of DTU varies: for simple queries – in the range of 1.8% - 68%, complex queries - in the range of 1.4% - 94%, and therefore the optimal level of the S7 model.

Table 15. *Number of operations per second (speed) when executing queries for level S0*

Query type	**Number of selected records**	**Query execution time, sec.**	**Number of operations per second**
SQ 1	1001002	79	12671
SQ 2	706	2	353
SQ 3	201002	93	2161
CQ 1	1001002	105	9533
CQ 2	13827	54	256
CQ 3	201002	103	1951

Table 16. *Number of operations per second (speed) when executing queries for level S3*

Query type	**Number of selected records**	**Query execution time, sec.**	**Number of operations per second**
SQ 1	1001002	73	13712
SQ 2	706	1	706
SQ 3	201002	99	2030
CQ 1	1001002	97	10320
CQ 2	13827	43	322
CQ 3	201002	54	3722

Table 17. *Number of operations per second (speed) when executing queries for level S7*

Query type	**Number of selected records**	**Query execution time, sec.**	**Number of operations per second**
SQ 1	1001002	53	18887
SQ 2	706	1	706
SQ 3	201002	74	2716
CQ 1	1001002	84	11917
CQ 2	13827	37	374
CQ 3	201002	42	4786

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) *Innovative technologies and scientific solutions for industries. 2022. No. 1 (19)*

Thus, the speed of input/read operations when performing various queries to the database varies in the ranges: for simple queries 12671 (S0) – 18887 (S7), for complex queries – 256 (S0) – 374 (S7). Therefore, the choice of the S7 model level is optimal.

Figures 2 - 4 show the results of visualization to analyze the relationships between query execution time and resource indicators used in the model. They show that it is desirable to use S2 level for simple queries and at least S3 level for complex queries.

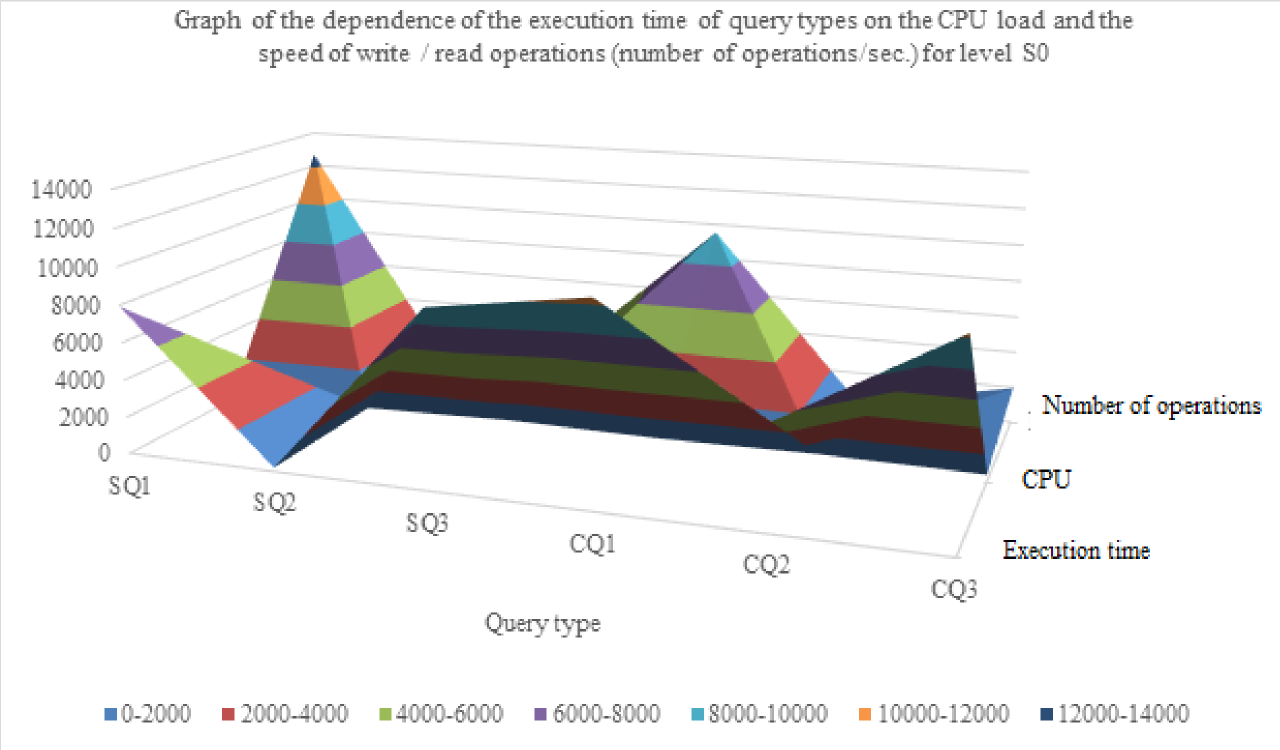

Fig. 2. Graph of the dependence of query execution time on CPU load and write / read operation speed (number of operations/sec.) for level S0

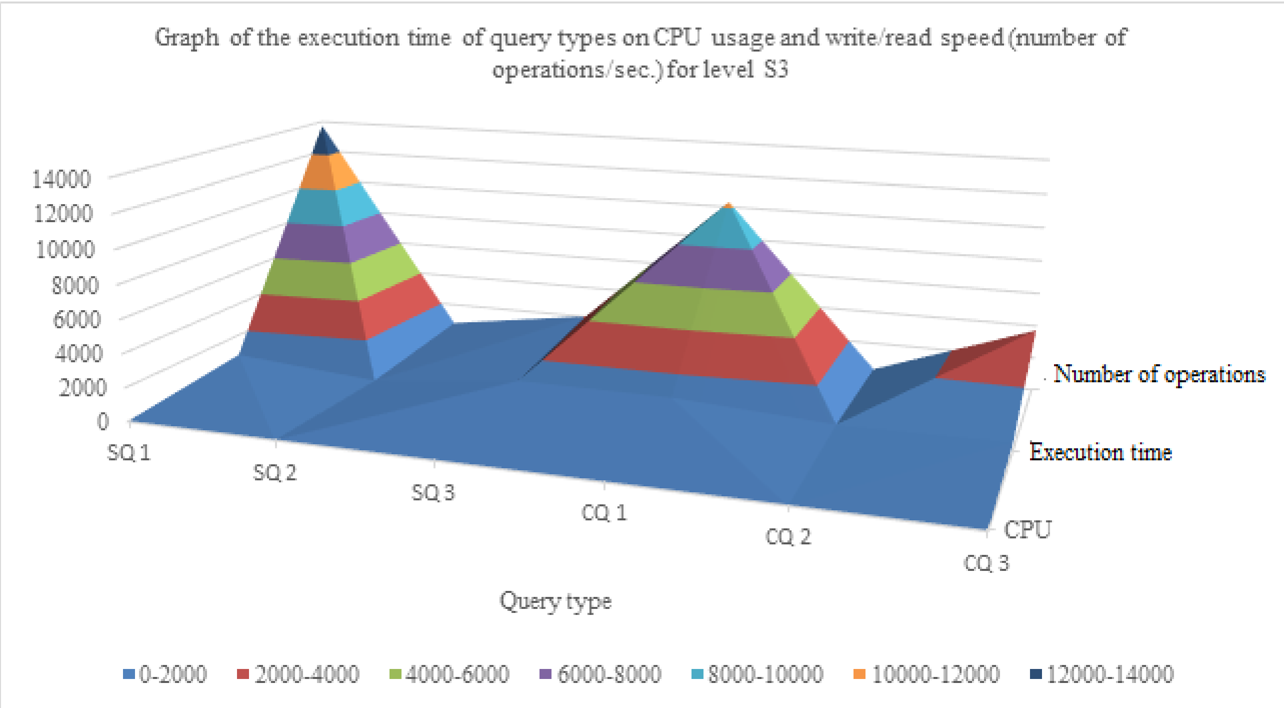

Fig. 3. Graph of query execution time dependence on CPU load and write/read operation speed (number of operations/sec.) for level S3

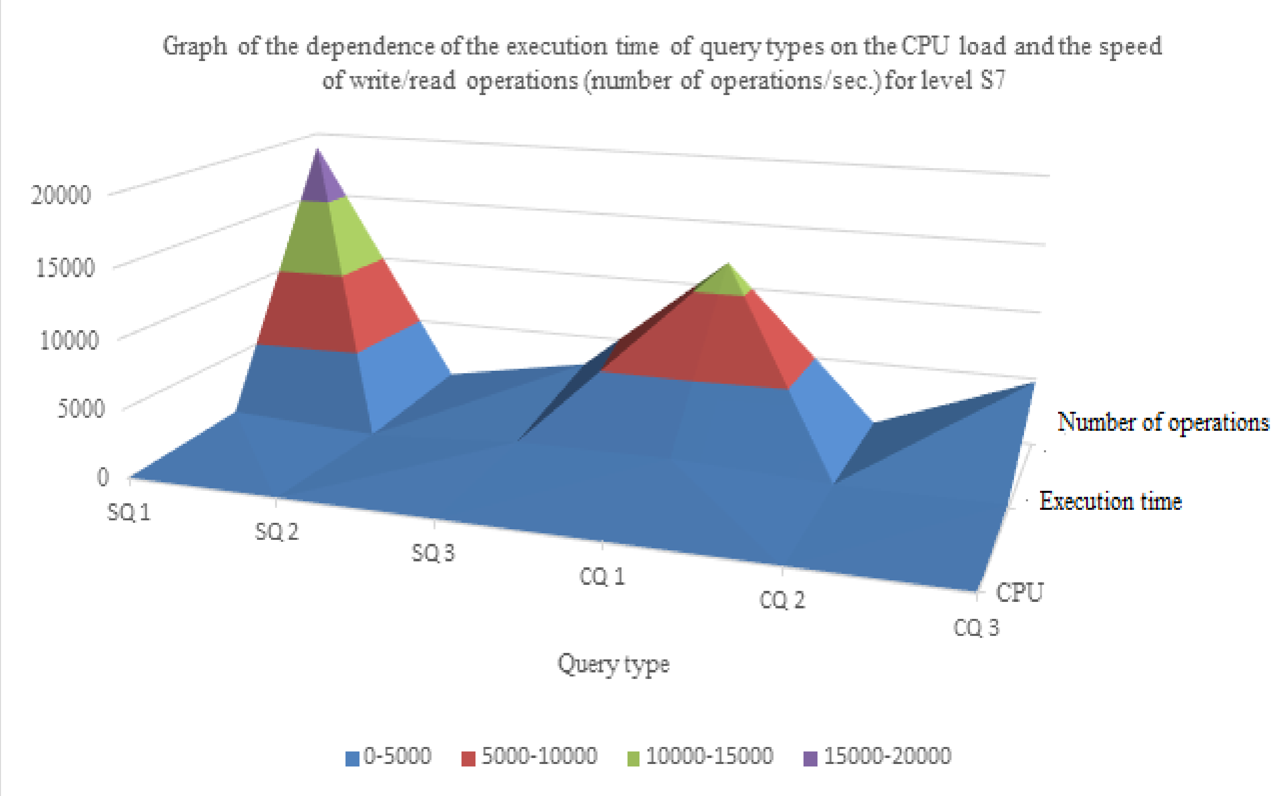

Fig. 4. Graph of dependence of time of execution of requests on CPU loading and speed of write/read operations (number of operations/sec.) for level S7

Conclusions.

The analysis of services and models of database work using cloud platforms showed that in the current global trends of their use the most acceptable from a practical point of view are the models of MS Azure platform – DTU and vCore. To assess the effectiveness of their use in working with relational databases, tools are used to monitor indicators - query execution time and resources involved (CPU and RAM load). It is proved that for a detailed assessment and further analysis of productivity it is necessary to expand the composition of these indicators, for which 2 groups of indicators of database work were proposed for the first time: indicators of execution time and performance of simple and complex queries the key indicator is the level of the DTU model. An additional factor influencing the performance of the selected DTU model is the data generation time – the results show that the generation of large data (10 GB or more) requires the use of the S7 level. Thus, based on the results of the experiments, we can draw a general conclusion – the optimal choice is the level of the model within S3 and S7. The developed set of database performance indicators on the Azure cloud platform expands the basis of methodological principles of using and evaluating the performance of relational databases on cloud platforms [16 - 19] by analyzing the results of simple and complex queries for the DTU model and allows final selection of a model key factors - the type of queries and the amount of test data.

Practical use of the proposed indicators for the DTU model will increase the efficiency of decision-making on the choice of model level in the implementation of different types of queries and database data generation on cloud platforms.

References

1. TOP 10 cloud trends in 2018, available at: https://ecm-journal.ru/post/TOP-10-oblachnykh-tendencii-v-2018-godu.aspx.
2. What is cloud computing?, available at: https://aws.amazon.com/ru/what-is-cloud-computing/.
3. Gartner., available at: https://www.gartner.com/en.
4. Marinescu, D., C. (2018), "Cloud Service Providers and the Cloud Ecosystem", *Cloud Computing*, No. 2, P. 13 – 49. DOI: https://doi.org/10.1016/B978-0-12-812810-7.00002-9
5. Database prices SQL Azure, available at: https://azure.microsoft.com/ru-ru/pricing/details/sql-database/single/.
6. Database SQL Azure, available at: https://azure.microsoft.com/ru-ru/services/sql-database/.
7. "How to migrate to Amazon RDS for SQL Server using transactional replication", available at: https://aws.amazon.com/ru/blogs/database/how-to-migrate-to-amazon-rds-for-sql-server-using-transactional-replication/.
8. Amazon RDS pricing for SQL Server, available at: https://aws.amazon.com/ru/rds/sqlserver/pricing/.
9. Delaney, K. (2018), *Microsoft Azure Options for SQL Server Relational Databases*, Microsoft Corporation, 22 p.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2022. No. 1 (19)

10. Marozzo, F. (2019), "Infrastructures for High-Performance Computing: Cloud Infrastructures", *Reference Module in Life Sciences*, No. 1, P. 240 – 246. DOI: https://doi.org/10.1016/B978-0-12-809633-8.20374-9
11. Boukhelef, D., Boukhobza, J., Boukhalfa, K., Ouarnoughi, H., Lemarchand, L (2019),"Optimizing the cost of DBaaS object placement in hybrid storage systems", *Future Generation Computer Systems*, Vol. 93, P. 176 – 187. DOI: https://doi.org/10.1016/j.future.2018.10.030
12. Saif, S., Wazir, I. (2018), "Performance Analysis of Big Data and Cloud Computing Techniques: A Survey", *Procedia Computer Science*, No. 132, P. 118 – 127. DOI: https://doi.org/10.1016/j.procs.2018.05.172
13. Importing and Exporting SQL Server Databases, available at:https://docs.aws.amazon.com/en_us/AmazonRDS/latest/UserGuide/SQLServer.Procedural.Importing.html.
14. Resource limits for single databases using the DTU-based purchasing model, available at: https://docs.microsoft.com/en-us/azure/sql-database/sql-database-dtu-resource-limits-single-databases.
15. Azure SQL Database DTU Calculator, available at: http://dtucalculator.azurewebsites.net/.
16. Passos da Costa, P. J., Rosado da Cruz, A. M. (2012), "Migration to Windows Azure – Analysis and Comparison", *Procedia Technology*, Vol. 5, P. 93–102. DOI: https://doi.org/10.1016/j.protcy.2012.09.011
17. Han, J., Kamber, M., Pei, J. (2012), "Data Warehousing and Online Analytical Processing". DOI: https://doi.org/10.1016/B978-0-12-381479-1.00004-6
18. Dabbechi, H,. Nabli, A., and Bouzguenda, L. (2016), "Towards Cloud-Based Data Warehouse as a Service for Big Data Analytics", *Lecture Notes in Computer Science*, 9876, P.180–189. DOI: https://doi.org/10.1007/978-3-319-45246-317
18. Zolotariov, D. (2020), "The system of automated billing on the basis of bad infrastructure has been developed" ["Rozpodilena sistema avtomatizovanih obchislen' na bazi hmarnoї infrastrukturi"], *Innovative technologies and scientific solutions for industry*, No. 4 (14), P.47-56. DOI: https://doi.org/10.30837/ITSSI.2020.14.047

Received 14.01.2022

Відомості про авторів / Сведения об авторах / About the Authors

Мінухін Сергій Володимирович – доктор технічних наук, професор, професор кафедри інформаційних систем ХНЕУ ім. С.Кузнеця, Харків, Україна; e-mail: minukhin.sv@gmail.com; ORCID: https://orcid.org/ 0000-0002-9314-3750.

Минухин Сергей Владимирович – доктор технических наук, профессор, профессор кафедры информационных систем ХНЭУ им. С.Кузнеца, Харків, Україна.

Minukhin Serhii – Doctor of Sciences (Engineering), Professor, Professor of the Department of Information Systems, Simon Kuznets Kharkiv National University of Economics Kharkiv, Ukraine.

ДОСЛІДЖЕННЯ ПРОДУКТИВНОСТІ МОДЕЛІ DTU ДЛЯ РЕЛЯЦІЙНИХ БАЗ ДАНИХ НА ПЛАТФОРМІ AZURE

При рішенні завдань роботи з реляційними БД на хмарних платформах виникає проблема вибору певної моделі для забезпечення продуктивності виконання запитів різної складності. **Об’єктом** дослідження є процеси реалізації різних типів запитів до реляційних БД в рамках моделі придбання DTU платформи MS Azure. **Предметом** є методи оцінювання продуктивності роботи з реляційними БД на основі часових показників виконання запитів та показників завантаженості ресурсів хмарної платформи. **Мета** дослідження полягає в розробленні системи показників для моніторингу поточного стану роботи з БД для обгрунтованого прийняття рішень щодо вибору певної цінової категорії моделі DTU хмарного сервісу MS Azure, що дозволить оптимізувати результати роботи з БД платформи Досягнення поставленої мети передбачає виконання таких **завдань**: проаналізувати сучасні засоби та сервіси роботи з БД, зокрема реляційними БД, на хмарних платформах Azure та AWS, особливості їх застосування та реалізації; розробити програмне забезпечення для генерація тестових реляційних БД різних об’ємів; провести тестування згенерованих БД на локальному ресурсі; з врахуванням характеристик рівнів моделі DTU Azure розробити нову систему показників продуктивності , яка включає 2 групи - часові показники та показники завантаженості існуючих ресурсів платформи; розробити та реалізувати запити різної складності для згенерованої тестової БД для різних рівнів моделі DTU та провести аналіз отриманих результатів. **Методи**. В дослідженні використано такі методи: методи проектування реляційних баз даних; методи створення запитів у SQL-орієнтованих базах даних з довільною кількістю таблиць; методи створення та міграції даних у хмарні платформи; методи моніторингу результатів виконання запитів на основі часових та ресурсних показників; методи генерації тестових даних для реляційних БД; системний підхід для комплексного оцінювання та аналізу продуктивності роботи з реляційними БД. **Результати**. На основі розробленої системи показників, що використовується для поточного аналізу процесів роботи з реляційними БД платформи MS Azure, проведено чисельні експерименти для різних рівнів моделі для простих та складних запитів до БД загальним об’ємом 20 ГБ: навантаження рівнів моделі DTU при виконанні різних запитів, вплив рівнів моделі DTU Azure SQL database на показники виконання простих та складних запитів, залежність часу виконання різних запитів від завантаженості CPU (ЦП) та швидкості операцій запису/читання для різних рівнів моделі. **Висновки**. Отримані результати експериментів дозволяють зробити висновок щодо використання рівнів моделі DTU - S3 та S7 - для генерації тестових даних різного об’єму (до 20 ГБ) та виконання запитів до БД. Практичне використання запропонованих показників для оцінювання результатів застосування моделі DTU дозволить підвищити ефективність прийняття рішень щодо вибору рівня моделі при реалізації різних запитів та генерації тестових даних БД на хмарній платформі Azure. Розроблений набір показників роботи з реляційними БД на хмарній платформі Azure розширює базис методологічних засад оцінювання продуктивності роботи з реляційними БД на хмарних платформах шляхом аналізу результатів виконання простих та складних запитів до БД на задіяних ресурсах.

Ключові слова: хмарна платформа; реляційна база даних; модель придбання DTU; показники часу та завантаженості; генератор даних; тестові дані; складність запиту.

ИССЛЕДОВАНИЕ ПРОИЗВОДИТЕЛЬНОСТИ МОДЕЛИ DTU ДЛЯ РЕЛЯЦИОННЫХ БАЗ ДАННЫХ НА ПЛАТФОРМЕ AZURE

При решении задач работы с реляционными БД на облачных платформах возникает проблема выбора определенной модели для обеспечения производительности выполнения запросов разной сложности. **Объектом** исследования являются процессы реализации различных типов запросов к реляционным базам данных в рамках модели приобретения DTU платформы MS Azure. **Предметом** является методы оценки производительности работы с реляционными базами данных на основе временных показателей выполнения запросов и показателей загруженности ресурсов облачной платформы. **Цель** исследования заключается в разработке системы показателей для мониторинга текущего состояния работы с БД для обоснованного принятия решений выбора определенной ценовой категории модели DTU облачного сервиса MS Azure, что позволит оптимизировать результаты работы с БД. платформы Достижение поставленных целей предполагает выполнение следующих **задач**: проанализировать современные средства и сервисы работы с БД, в частности реляционными БД, на облачных платформах Azure и AWS, особенности их применения и реализации; разработать программное обеспечение для генерации тестовых реляционных БД разных объемов; провести тестирование сгенерированных БД на локальном ресурсе; с учетом характеристик уровней модели DTU Azure разработать новую систему показателей производительности, которая включает 2 группы – временные показатели и показатели загруженности существующих ресурсов платформы; разработать и реализовать запросы разной сложности для сгенерированной тестовой БД для разных уровней модели DTU и провести анализ полученных результатов. **Методы**. В исследовании использованы следующие методы: методы проектирования реляционных баз данных; методы создания запросов в SQL-ориентированных базах данных с произвольным количеством таблиц; методы создания и миграции данных в облачные платформы; методы мониторинга результатов выполнения запросов на основе временных и ресурсных показателей; методы генерации тестовых данных для реляционных БД; системный подход для комплексной оценки и анализа производительности работы с реляционными БД. **Результаты**. На основе разработанной системы показателей, используемой для текущего анализа процессов работы с реляционными БД платформы MS Azure, проведены многочисленные эксперименты для разных уровней модели для простых и сложных запросов к БД общим объемом 20 ГБ: нагрузка уровней модели DTU при выполнении различных запросов, влияние уровней модели DTU Azure SQL database на показатели выполнения простых и сложных запросов, зависимость времени выполнения различных запросов от загруженности CPU (ЦП) и скорости операций записи/чтения для разных уровней модели. **Выводы**. Полученные результаты экспериментов позволяют сделать вывод об использовании уровней модели DTU - S3 и S7 - для генерации тестовых данных разного объема (до 20 ГБ) и выполнения запросов к БД. Практическое использование предложенных показателей для оценки результатов применения модели DTU позволит повысить эффективность принятия решений по выбору уровня модели при реализации различных запросов и генерации тестовых данных на облачной платформе Azure. Разработанный набор показателей работы с реляционными БД на облачной платформе Azure расширяет базис методологических основ для оценки производительности работы с реляционными БД на облачных платформах путем анализа результатов выполнения простых и сложных запросов к БД на задействованных ресурсах.

Ключевые слова: облачная платформа; реляционная база данных; модель приобретения DTU; показатели времени и загруженности; генератор данных; тестовые данные; сложность запроса.

Бібліографічні описи / Bibliographic descriptions

Мінухін С. В. Дослідження продуктивності моделі DTU для реляційних баз даних на платформі Azure. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 27–39. DOI: https://doi.org/10.30837/ITSSI.2022.19.027

Minukhin, S. (2022), "Performance study of the DTU model for relational databases on the Azure platform", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 27–39. DOI: https://doi.org/10.30837/ITSSI.2022.19.027