Beltyukov Mikahil Romanovich
Student
Ural Federal University named after the first
President of Russia B.N. Yeltsin
Russia, Ekaterinburg
Academic supervisor: Kovaleva Alexandra Georgievna

FAULT TOLERANCE OF DATABASE SYSTEMS WITH REPLICATION AND MIRRORING TECHNOLOGIES

Abstract. *The paper describes approaches to organizing a fault tolerant database. To ensure security, the technologies of database replication and reflection are used. The methodology of applying these technologies is considered. The study proposes schemes for applying such approaches to ensure the required level of fault tolerance of a database.*

Keywords: *availability, databases, fault tolerant systems, database replication, database mirroring.*

Бельтюков Михаил Романович
Студент
Уральский федеральный университет имени первого
Президента России Б.Н. Ельцина
Россия, г. Екатеринбург
Научный руководитель: Ковалева Александра Георгиевна

ОТКАЗОУСТОЙЧИВОСТЬ БАЗ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ ОТРАЖЕНИЯ И РЕПЛИКАЦИИ

Аннотация. *В статье рассматриваются подходы к организации отказоустойчивой базы данных. Для обеспечения безопасности используются*

технологии репликации и отражения баз данных. Рассматривается методология применения данных технологий. Предлагаются схемы применения таких подходов для обеспечения требуемого уровня отказоустойчивости базы данных.

Ключевые слова: *доступность, базы данных, отказоустойчивые системы, репликация баз данных, отражение баз данных.*

Today, many information systems operate with large and complex data sets organized into databases. For this reason, there should be an access to this data at any given time. Technological support for the continuity of interaction with stored data is one of the priorities of the information system. This may be achieved by organizing a fault-tolerant database. The database server and the information system may stop working on many factors, such as temperature rise, equipment failure, software error and human factor. These situations bring direct and indirect damage, despite their possible short duration.

For example, fines from suppliers, buyers, government and other organizations for the late submission of information, loss of profit from unexecuted transactions, lack of internal operational information during the key decisions of the company's management, etc. These problems also lead to a negative image of the company in the market and may reduce the desire of partners to work with this an unstable partner.

The rapid pace of penetration of digital technologies in all spheres of life of modern society entails a change in the familiar models of economic and social mode of states. The Russian government, in the context of some national projects aimed at the digitalization of many areas of the economy and the public sector, should provide enough conditions for the organization of modern information systems 0. The development of this direction is declared as one of the priority directions of the state policy and is considered as an obligatory condition of competitiveness in the modern global market and a strategic component of economic independence of the state.

Probably the main and most popular information system among the people of the Russian Federation is the State Services Portal of the Russian Federation –

«Gosuslugi». It provides access to information about all available state and municipal services. This information system must be as fault tolerant as possible, since at any given time it can be accessed by a huge number of users. As an example, after the President of the Russian Federation appealed for financial support to the citizens of those affected by the COVID-19 crisis, the site was inaccessible. Thousands of users simultaneously tried to send a request for money. This event should have become the reason to increase the fault tolerance of this information system.

In addition, the Russian government plans to switch to the electronic form of certification of human labor activity from 2021. For this purpose, a database will be organized where all necessary information about the employee will be stored. Such information system and the database shall be fault tolerant 0.

One of the approaches to organizing such a database is to use database replication and mirroring technology. It is a relation database management technique which allows maintaining data consistency in spite of its high availability by creating redundant copies of a data set.

Fault tolerance - property of the technical system to maintain its operability after failure of one or more of its components Fault tolerance is determined by the number of single failures of components (elements) of the system, after which the performance of the system is maintained. The basic level of fault tolerance implies protection from failure of one of any elements.

The main way to increase fault tolerance of the system is redundancy of elements of this system. Under redundancy we mean functionality in which there is no need if the system works without faults. Approaches to implementing redundancy can be implemented in several ways: hardware (additional components), software (special programs), and time (repetition of operations) 0.

Replication is the process of copying data from a central database to one or more databases. It is a process of storing data in more than one site or node. Database replication can be used on many database management systems (DBMS), usually with a master-slave relationship between the original and the copies. When replicating with

multiple main servers, updates can be sent to any database node and then transferred to other servers.

The result is a distributed database in which users can quickly access data relevant to their tasks without interfering with the work of others. Numerous elements contribute to the overall process of creating and managing database replication.

In general, distributed database management systems (DBMS) work to ensure that changes, additions and deletions to data performed at any given location are automatically reflected in the data stored at all other locations. DBMS is, in fact, the name of the infrastructure that allows or performs database replication - the system that manages the distributed database, which is a product of database replication.

Data replication involves constant duplication of transactions, so that the replication is constantly updated and synchronized with the source. The replication may be complete, in which the entire database is stored on each site. There may also be partial replication, in which some frequently used database fragments are replicated while others are not.

There are two schemes of database replication. The first of them is a scheme of full data replication. Replication of the entire database on each server in a distributed system. This will improve the availability of the system as the system can continue to operate while at least one server is running. This scheme provides high availability of data. It ensures that database queries can be executed on separate servers. The disadvantages are slow updating of the data, as they must be added to each server to be consistent.

The number of copies is jealous of weaving data. This approach will allow to take advantage of replication for important data, those segments to which queries are performed less frequently will not be replicated. This will increase the speed of data update, while ensuring enough data availability and fault tolerance of the system. By analyzing the database and requests to it, we can build an optimal scheme of data replication, the most loaded nodes of which will be replicated

Partial replication may be considered as a compromise solution. In this case some fragments of the database may be replicated whereas others may not. The number

of copies of the fragment may range from one to the total number of sites in the distributed system. The description of replication of fragments is sometimes called the replication schema.

Database mirroring is the process of creating multiple copies of data and database. Usually, when mirroring, the database is copied to another place from its main database. In contrast to replication, the database mirroring technology does not allow creating a distributed system, since in this case it creates a complete copy of the original database, which starts to be used only if the main one fails. Only one copy may be obtained at any time.

The close link between the Mirror Database and the Primary Database is established by using the cause-and-effect blocks of the operation log with the reflected database. Transactions and changes to the primary database are transferred directly to the mirror and processed immediately so the mirror is always up-to-date and available as a standby.

In case of correct operation of the main server, the interaction of users with the database is performed directly with the main server. Each successfully completed transaction will be reflected in a mirrored database located on a separate server. When the main server is unavailable, the reflected database may be included in the operation. Since this database contains all source data from the main database, the user may not notice any difference in the system operation

By combining these two technologies, an optimal database scheme for a specific task may be created. The core factors are the necessary level of fault tolerance as well as the financial costs to provide redundancy and payment to the specialists who create and operate the database.

There are six different ways to use mirroring and replication. The first is not to use these technologies at all. This approach ensures the maximum ease of organization and operation of the database. The lack of redundancy significantly reduces the cost of such a database, but in the event of a critical failure, users lose access to it until it is fully restored. This is an undesirable approach to database organization, it makes sense to apply it only for training in specific DBMS, or the creation and organization of

databases. A simple database allows students not to get distracted from general concepts.

As opposed to this, mirroring and full replication in the database may be implemented. In addition to the fact that mirroring is limited to only two servers, it is possible to mirror each server that hosts the replicated database. This ensures maximum fault tolerance using these technologies. As the replication of database is a distributed database, in case of failure of one of the mirroring servers, the mirroring server becomes active (Figure 1). Therefore, the performance of the entire system does not change.

In this case, the system is too redundant and the road to operate. This scheme finds its application in systems for which the performance of servers when accessing databases is critical, as well as providing the highest level of fault tolerance of the database.

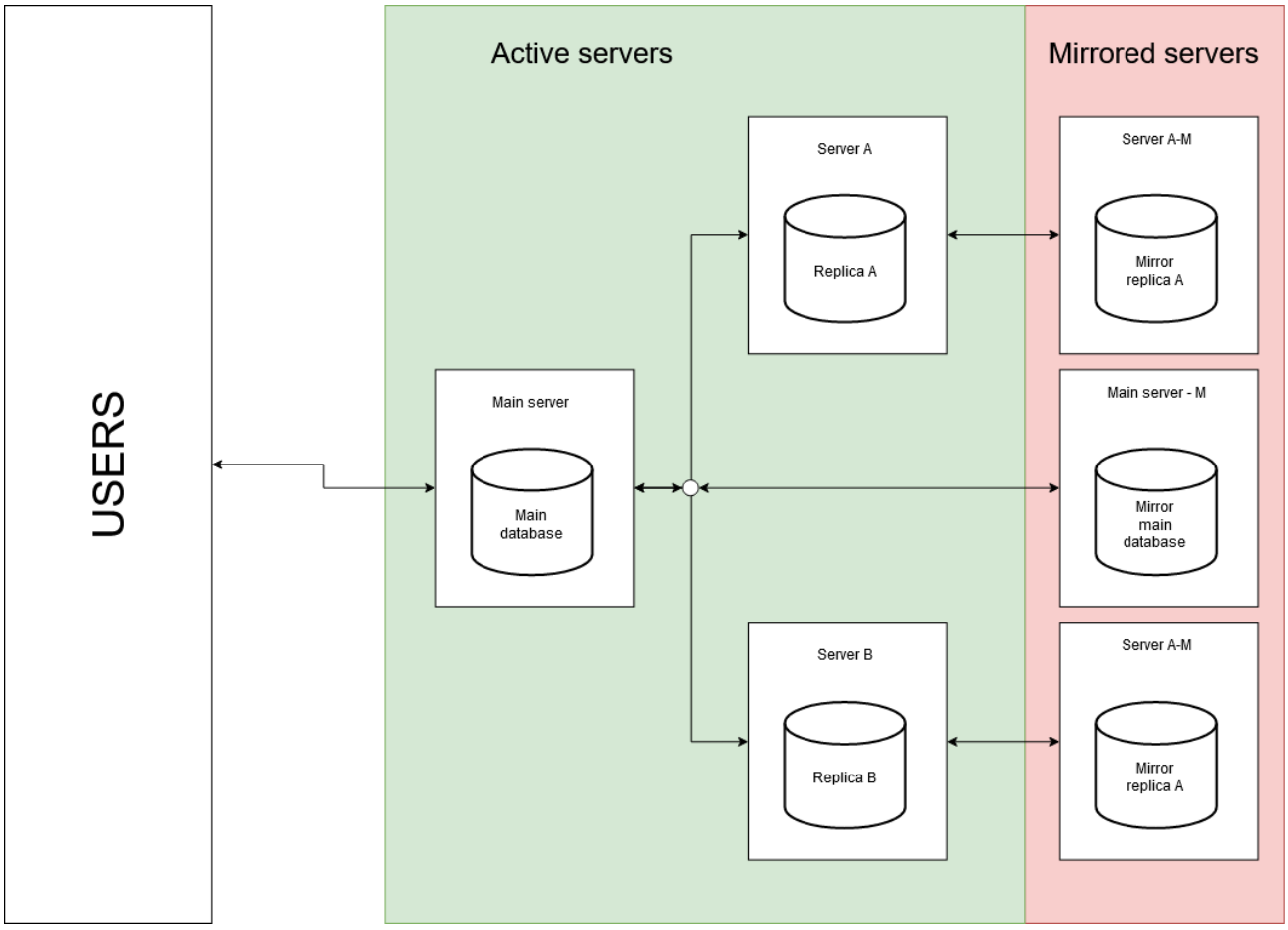

Figure 1. – Full replication scheme with mirroring

To reduce the cost of creating and operating such a database, it is necessary to reduce the redundancy of such a system by identifying those parts of the database that are not expected to be heavily loaded. For this purpose, the replication is carried out in accordance with the partial scheme (**Ошибка! Источник ссылки не найден.**). It is possible to produce mirroring in different ways:

1) Mirroring replicated database parts
2) Mirroring non-replicated database parts
3) Combine the two approaches

The most balanced option is to use the second option. This ensures high fault tolerance of the database, as all important parts of the database are replicated to individual servers that work synchronously with each other. Less loaded parts are mirrored to a separate server and in the case of a failure, this server becomes active. In order to do this, a mirror copy of the main database is created and the server may work with servers for heavily loaded parts.

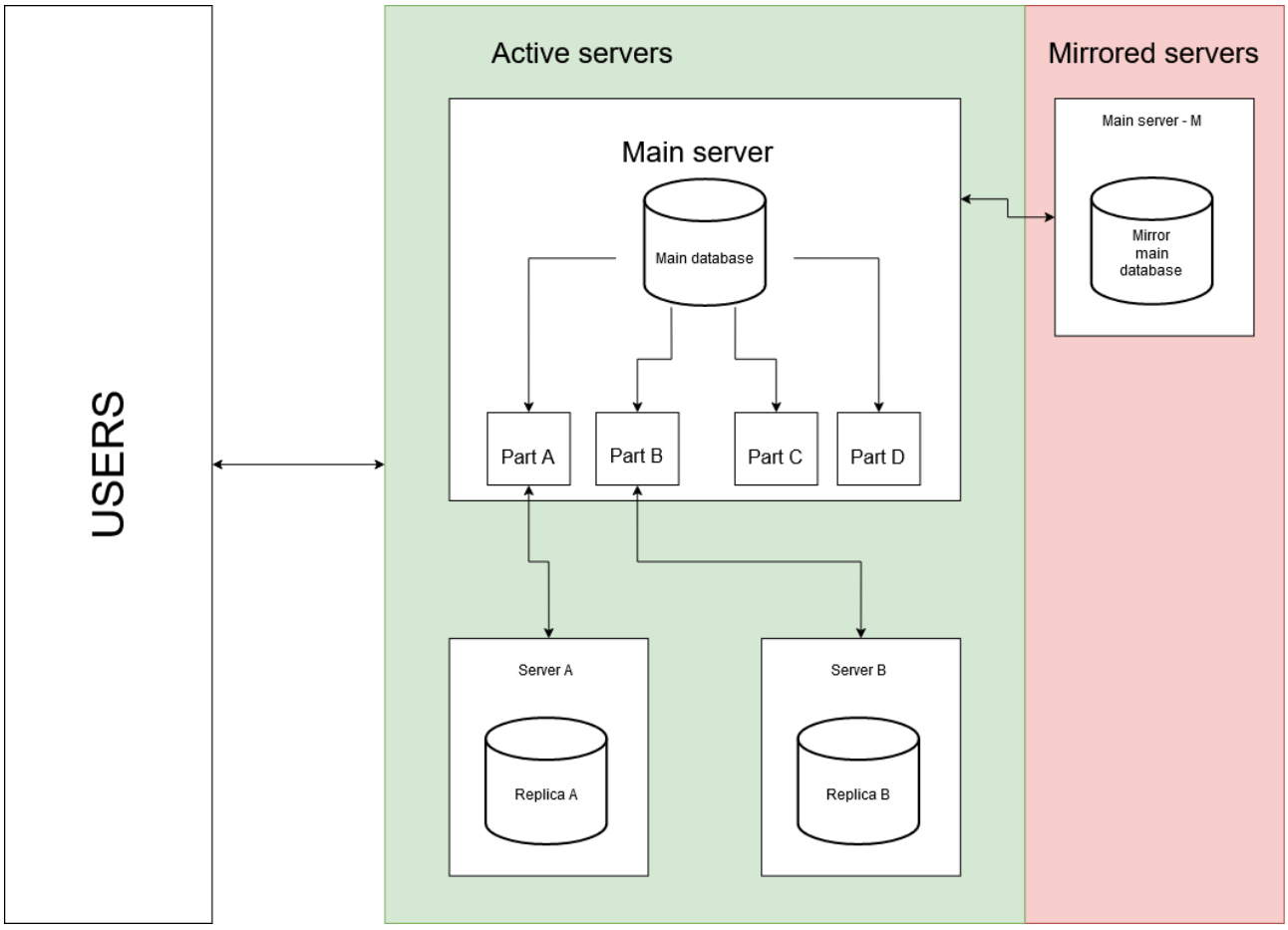

Figure 2. – Mirroring main database server with partial replication scheme

In addition, if there is a need to reduce system redundancy, mirroring may be abandoned. The database is replicated by full or partial scheme. Then, in the case of a main server failure, access to the database may be disrupted.

For small databases, whose owners do not have a lot of resources, it is relevant to use mirroring as the main tool to improve fault tolerance.

It is possible to conclude that the choice of fault-tolerant technologies should be based on the availability of financial and human resources. According to these criteria, the necessary level of redundancy of information system elements, including the database, is determined.

REFERENCES

1. Nacionalnaya programma «Cifrovaya ekonomika Rossijskoj Federacii»: PASPORT nacionalnogo proekta ot 04.06.2019 – URL: https://digital.gov.ru/ru/activity/directions/858/

2. O vnesenii izmenenij v Trudovoj kodeks Rossijskoj Federacii v chasti formirovaniya svedenij o trudovoj deyatelnosti v elektronnom vide: fed. Zalon ot 16.12.2019 N 439-FZ – URL: http://www.consultant.ru/document/cons_doc_LAW_340241/

3. Avizienis, A. - Fault-Tolerant Systems // IEEE Transactions on Computers / Vol. 25, No. 12, Dec. 1976, pp. 1304—1312 / URL: https://www.computer.org/csdl/journal/tc/1976/12/01674598/13rRUwInv3f

4. Watt A., N. Eng. Database Design – 2nd Edition / Victoria, B.C.: BCcampus. Retrieved from https://opentextbc.ca/dbdesign01/.

5. SearchDataManagement.com: a knowledge guide for data management professionals and business intelligence leaders. – URL: https://searchdatamanagement.techtarget.com/

6. Microsoft SQL documentation / URL: https://docs.microsoft.com/en-us/sql