ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫНЫҢ
Қ. И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті

ХАБАРЛАРЫ

ИЗВЕСТИЯ

НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК
РЕСПУБЛИКИ КАЗАХСТАН
Казахский национальный исследовательский
технический университет им. К. И. Сатпаева

NEWS

OF THE ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN
Kazakh national research technical university
named after K. I. Satpayev

SERIES
OF GEOLOGY AND TECHNICAL SCIENCES

5 (437)

SEPTEMBER – OCTOBER 2019

THE JOURNAL WAS FOUNDED IN 1940

PUBLISHED 6 TIMES A YEAR

ALMATY, NAS RK

NAS RK is pleased to announce that News of NAS RK. Series of geology and technical sciences scientific journal has been accepted for indexing in the Emerging Sources Citation Index, a new edition of Web of Science. Content in this index is under consideration by Clarivate Analytics to be accepted in the Science Citation Index Expanded, the Social Sciences Citation Index, and the Arts & Humanities Citation Index. The quality and depth of content Web of Science offers to researchers, authors, publishers, and institutions sets it apart from other research databases. The inclusion of News of NAS RK. Series of geology and technical sciences in the Emerging Sources Citation Index demonstrates our dedication to providing the most relevant and influential content of geology and engineering sciences to our community.

Қазақстан Республикасы Ұлттық ғылым академиясы "ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы" ғылыми журналының Web of Science-тің жаңаланған нұсқасы Emerging Sources Citation Index-те индекстелуге қабылданғанын хабарлайды. Бұл индекстелу барысында Clarivate Analytics компаниясы журналды одан әрі the Science Citation Index Expanded, the Social Sciences Citation Index және the Arts & Humanities Citation Index-ке қабылдау мәселесін қарастыруда. Webof Science зерттеушілер, авторлар, баспашылар мен мекемелерге контент тереңдігі мен сапасын ұсынады. ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы Emerging Sources Citation Index-ке енуі біздің қоғамдастық үшін ең өзекті және беделді геология және техникалық ғылымдар бойынша контентке адалдығымызды білдіреді.

НАН РК сообщает, что научный журнал «Известия НАН РК. Серия геологии и технических наук» был принят для индексирования в Emerging Sources Citation Index, обновленной версии Web of Science. Содержание в этом индексировании находится в стадии рассмотрения компанией Clarivate Analytics для дальнейшего принятия журнала в the Science Citation Index Expanded, the Social Sciences Citation Index и the Arts & Humanities Citation Index. Web of Science предлагает качество и глубину контента для исследователей, авторов, издателей и учреждений. Включение Известия НАН РК. Серия геологии и технических наук в Emerging Sources Citation Index демонстрирует нашу приверженность к наиболее актуальному и влиятельному контенту по геологии и техническим наукам для нашего сообщества.

«ҚР ҰҒА Хабарлары. Геология мен техникалық ғылымдар сериясы».
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)
Меншіктенуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.).
Қазақстан республикасының Мәдениет пен ақпарат министрлігінің Ақпарат және мұрағат комитетінде 30.04.2010 ж. берілген **№10892-Ж** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Мерзімділігі: жылына 6 рет.
Тиражы: 300 дана.

Редакцияның мекенжайы: 050010, Алматы қ., Шевченко көш., 28, 219 бөл., 220, тел.: 272-13-19, 272-13-18,
http://www.geolog-technical.kz/index.php/en/

Редакцияның Қазақстан, 050010, Алматы қ., Қабанбай батыра көш., 69а.
мекенжайы: Қ. И. Сәтбаев атындағы геология ғылымдар институты, 334 бөлме. Тел.: 291-59-38.

Типографияның мекенжайы: «Аруна» ЖК, Алматы қ., Муратбаева көш., 75.

«Известия НАН РК. Серия геологии и технических наук».
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)

Собственник: Республиканское общественное объединение «Национальная академия наук Республики Казахстан (г. Алматы)

Свидетельство о постановке на учет периодического печатного издания в Комитете информации и архивов Министерства культуры и информации Республики Казахстан **№10892-Ж,** выданное 30.04.2010 г.

Периодичность: 6 раз в год
Тираж: 300 экземпляров

Адрес редакции: 050010, г. Алматы, ул. Шевченко, 28, ком. 219, 220, тел.: 272-13-19, 272-13-18,
http://nauka-nanrk.kz /geology-technical.kz

Адрес редакции: Казахстан, 050010, г. Алматы, ул. Кабанбай батыра, 69а.
Институт геологических наук им. К. И. Сатпаева, комната 334. Тел.: 291-59-38.

Адрес типографии: ИП «Аруна», г. Алматы, ул. Муратбаева, 75

News of the National Academy of Sciences of the Republic of Kazakhstan. Series of geology and technology sciences.
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)
Owner: RPA "National Academy of Sciences of the Republic of Kazakhstan" (Almaty)
The certificate of registration of a periodic printed publication in the Committee of information and archives of the Ministry of culture and information of the Republic of Kazakhstan N 10892-Ж, issued 30.04.2010

Periodicity: 6 times a year
Circulation: 300 copies

Editorial address: 28, Shevchenko str., of. 219, 220, Almaty, 050010, tel. 272-13-19, 272-13-18,
http://nauka-nanrk.kz/geology-technical.kz

Editorial address: Institute of Geological Sciences named after K.I. Satpayev
69a, Kabanbai batyr str., of. 334, Almaty, 050010, Kazakhstan, tel.: 291-59-38.

Address of printing house: ST "Aruna", 75, Muratbayev str, Almaty

NEWS
OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
SERIES OF GEOLOGY AND TECHNICAL SCIENCES
ISSN 2224-5278
Volume 5, Number 437 (2019), 82 – 90 **https://doi.org/10.32014/2019.2518-170X.128**

UDC 004.056
МРНТИ 81.93.29

A. Zh. Karipzhanova[1], K. M. Sagindykov[1], A.M. Gudov[2], Kalin Dimitrov[3]

[1]L. N. Gumilyov Eurasian National University, Astana, Kazakhstan,
[2]Kemerovo State University, Kemerovo, Russia,
[3]Technical university of Sofia, Sofia, Bulgaria.
E-mail: kamilakz2001@mail.ru, ksagin@mail.ru, good@kemsu.ru, kld@tu-sofia.bg

PARAMETERS OF THE DISTRIBUTED DATABASES OF INFORMATION SYSTEMS WHEN SPLITTING DATA WITH APPLICATION OF ALGORITHMS OF MULTIDIMENSIONAL PARITY

Abstract. So far the standard method of prevention of loss of information is a repeated reservation that leads to huge material inputs. Big Data security system developed by authors of the article with an application of the algorithms of multidimensional parity steady against partial losses of places of storage, showed the increased safety level, in particular when using in a cloud computing. In this algorithm data, it is split on a large number of files, each of which does not may contain even one bit of the initial information dispersed in a cloud. The system realized by authors does not demand additional expensive infrastructure of reservation, easily is scaled, extends, and in process of increase in the sizes of infrastructure from the addition of units of storage safety and reliability of data storage automatically increases. Results of testing of parameters of the safety of the main basic subsystems of a technology of the distributed storage with a splitting of data are received. Results showed compliance of parameters of safety to modern requirements and, respectively, a possibility of reduction of domination of expensive infrastructure of reservation and Backup. The technology of the distributed storage with splitting of data personifies the new paradigm of safety opening a possibility of effective counteraction to numerous threats to the stored information and big calculations.

Key words: safety of information, cloud computing, the distributed storage, splitting of data.

Introduction. Safety of big data and calculations in clouds represents a problem which value only increases in the process of increase of external threats over time [1]. According to Brookings Institution, safety is the main obstacle for the federal U.S. Government in plans to transfer more functions to cloud platforms. The government of Kazakhstan, in general, was forced to issue Resolution No. 965 of September 14, 2004 "About some measures for ensuring information security in the Republic of Kazakhstan" according to which "processing and storage of the data making the state secrets, office information of limited distribution are carried out on the computer aids which are not connected to the international (global) data transmission networks, the Internet and/or to the information networks, communication networks having an exit in the international (global) data transmission networks, the Internet". Moreover, the possibilities of practical realization of optical neural networks with a small number of elements is considered [2, 3].

How in such a situation to strengthen the reputation of public clouds?

Only by ensuring reliable safety of the stored information. Badly operated infrastructures of storage of cloudy data threaten any business, any governmental activity in case of catastrophic failure. Therefore, the main objective is the creation of cost-effective architecture with extremely safe storage, scalability and seamless integration between local and cloudy ways of data storage.

We offered for the solution of security of cloudy systems technology of the distributed storage of information with splitting of data [4]. Originally this technology was approved in local networks [5], then established and passed tests in cloudy systems [6].

The essence of the offered technology consists in what the stored data is split on considerable number of files, each of which does not may contain even one bit of initial information. The divided files are distributed on a set of the servers programmed on self-recovery and self-preservation that ensures constant safety of data and safety [4].

The algorithm is not enciphering [7] therefore it cannot be deciphered. The complex structure of safety and resistance to damages of places of storage provides increased protection against the hacker attacks. Thanks to the innovative hierarchical protocol of access, the system prevents unauthorized access to data and plunder of information by insiders. At last, the system guarantees 100% of safe data transmission through open Internet channels and private networks.

It is possible to tell that in a sense our technology of the distributed storage of information with splitting of data is entered in the theory of universal objects in a class of the positive preorders considered a rather computable reducibility [8].

In the real work results of testing of parameters of safety of the main basic subsystems of technology of the distributed storage with splitting of data with application of algorithms of multidimensional parity which have to show the high level of protection and recovery of data in technology of the distributed storage of information with splitting of data are presented. The preliminary tests which are carried out on prototypes guarantee safety level which does not have now analogs in the world.

Method. The method of the distributed storage of information applied by us with splitting of data is intended for ensuring unique protection of databases and large-scale calculations in cloud systems. The method delivers new hybrid the object/block architecture which easily integrates optimal solutions for storage of the structured and unstructured data. In principle, the formulation and solution of the problem of optimal allocation of program modules and database arrays to the nodes of computing systems of a given topology is considered in a number of works [9-11]. But as a result of application of our method the new class of codes splits Big Data in a large number of files, each of which does not may contain even one bit of initial information. The divided files are distributed on a set of the servers programmed on self-recovery and self-preservation that ensures constant safety of data and safety. The separate split data in itself do not bear intelligent information. Another aspect is connected with the fact that the configuration of splitting/restoration can be made so that recovery of data could be executed with application only of a part of the split data, that is it is possible to provide resistance to loss of data.

The frontend of a system is realized on web technologies and uses the simple protocol of hypertext references of HTTP together with interactive technologies on AJAX/PHP (WEB 2.0). The basic control system of content of a web part is realized on CMS Word Press. The website is located on dedicated the server of Internet PS Company LLP, under the Linux Ubuntu Server operating system and the Apache 2.25 web server with PHP 7.0. As the database is used MySQL 5.0.

Technology. The technology offers seamless integration between local and cloud storages. Guarantees unprecedented safety of data and safety and also high scalability. The multilayered infrastructure of safety the system of the false IP addresses provides protection against the attacks of external hackers. Flexible hierarchical protocols of access convince that system administrators, service providers and other strangers have no access to the stored data that excludes need of personal safety, execution of protocols and promotes cost reduction. The technology also guarantees 100% of safe data transmission through open Internet channels and private networks.

In the developed technology patent algorithms of parity with resistance to multiple refusals are applied. The target platform on which the system functions, – MS Windows XP/Vista/7/8 with .NET Framework v.4/4.5.

Structurally the system of the distributed storage of information consists of knots with the installed software of NODE connected among themselves by a confidential communication channel and the software of CLIENT working with these knots. Such multi-agent systems are well known [12]. Many such systems are the basis of decentralized systems with autonomous components [13]. The platforms used are as follows: Java Agent Development Framework [14], JACK Intelligent Agents [15], Multi-Agent Development Kit [16], Agent Builder [17], Cognitive Agent Architecture [18], Agent Building Tool-kit [19], etc.

The main modules of a system, the user interface, server a component, knots of storage are subordinated to a program complex according to the card of interaction of components (figure 1).

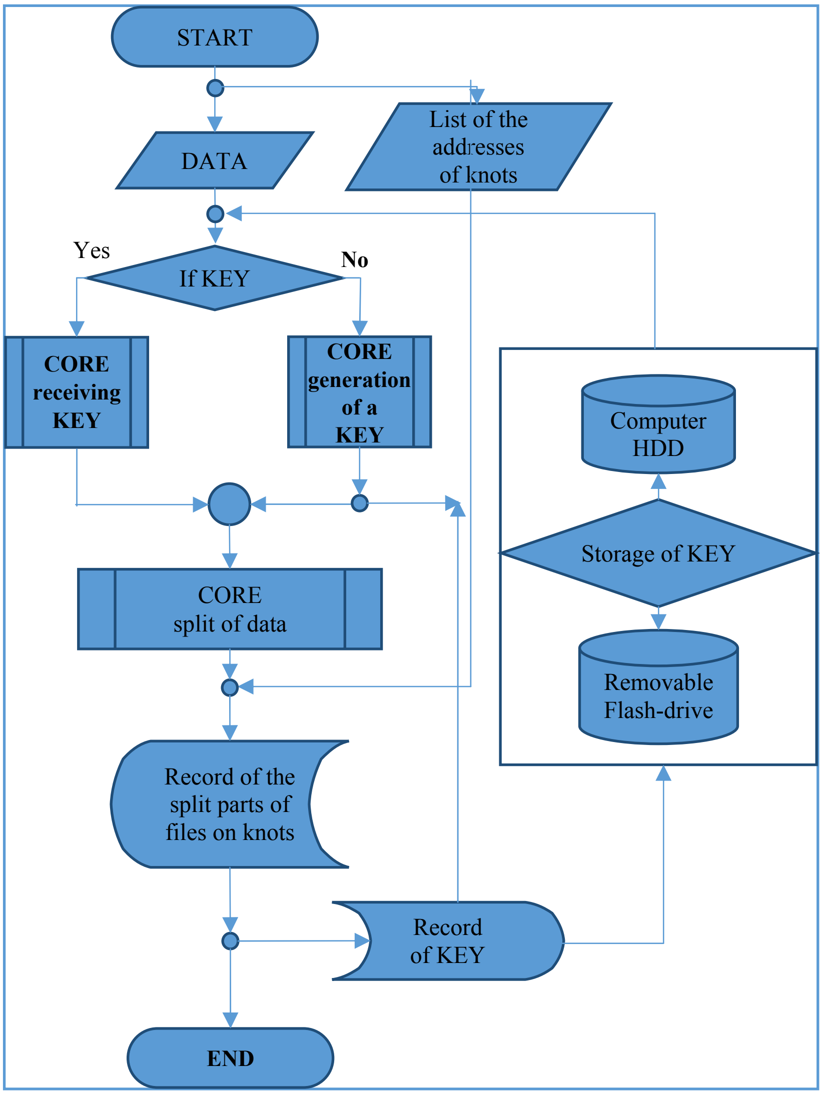

Figure 1 – The card of interaction of components of a system of the distributed storage of information

The client application is responsible for the breakdown of the file on blocks of data and distribution of blocks of data between knots. For storage of the file system and the distributed block, the client application is used by the metafile. Thanks to a tree-like system, the metafile describes file structure and arrangement of blocks of data. The metafile at completion of work is divided into blocks and is also distributed between knots. Only a unique key of the client of CID together with data of authorization (the account, the password) can collect the metafile.

The user establishes a client part of the application and will become authorized in the Distributed Cloud System system, receiving at the same time a unique key of CID which serves for formation of the metafile. Having chosen the file necessary to it, the client starts the Put in a Cloud function. At the same time, this file is blocked, both in the background divided into blocks and distributed on system knots. In need of obtaining the file, the client starts the Receive from a Cloud function, and to it from a cloud blocks of data of its file arrive and by means of an algorithm are packed in the file. The client can also give access to the directory to other users.

The server part of the application (figure 2) serves for storage and processing of information arriving from the client. Storage is carried out in data files which are broken into clusters. The cluster is in turn

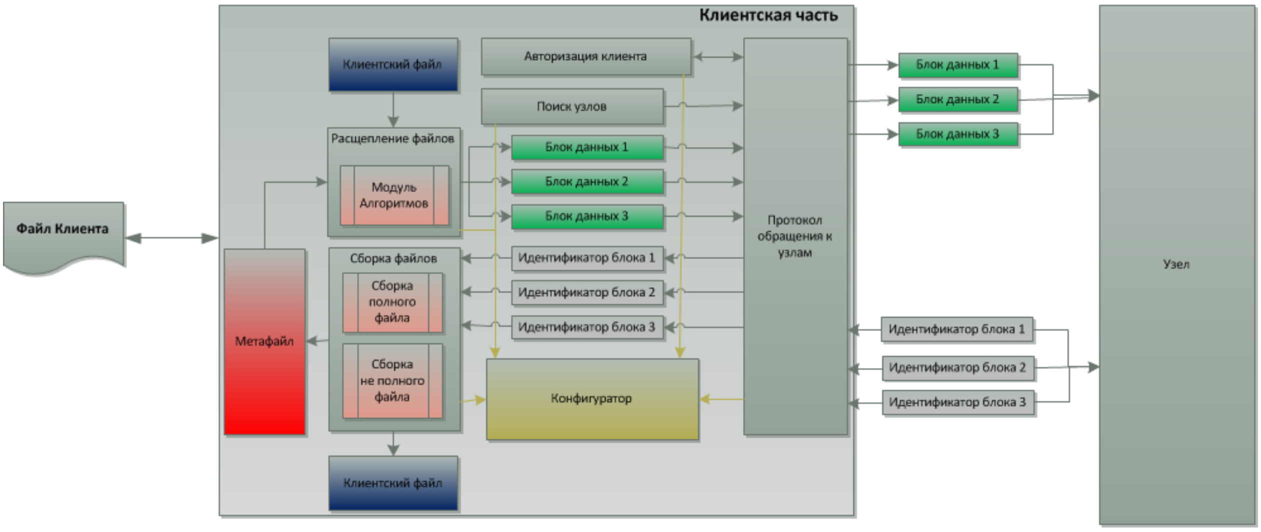

Figure 2 – Scheme of work of a server part of a system

broken into pages which are ranged on a 2-fold system. The size of a cluster is configured in settings of a system, the cluster contains multiple two the number of pages. For example, if the cluster consists of 32 megabytes, then it contains 256 pages of 128 kilobytes in size. In the process of receipt of new information the system creates new clusters for data storage with various dimension of pages.

All arriving data register in turn which monitors all processes of data recording. After successful completion of receiving the client block of data the system writes down this block in the data file and sends to the client the notice of successful receiving the block of data in the form of UID.

"Client" is an ordinary computer device in the form of the personal computer, the server or any other intelligent device on which the software of "CLIENT" is installed. The client serves as a lock for login.

"Knot" is a server with the installed software of "NODE".

"Client" works with a system as with an abstract cloudy subsystem. Interaction of "Client" with the massif of knots is carried out under the one-to-many protocol (one to many).

Any act of recording information into the system is pre-processed by a special algorithm that: a) splits information into unreadable components; b) adds dynamically generated redundant data to increase resistance to partial loss of split parts; c) generates a service metafile describing the created array.

The act of record of set of the created data is implemented by their distribution on system Knots. Any act of reading information by "Client" from "System" is possible only by means of processing of data array received from "Knots" by an algorithm which can restore it, only knowing its metadata. The system knots making the distributed massif know only the limited number of the neighbors. No knot can know all system and knots making it. Knots can dynamically be connected and be disconnected in a system that does not affect operability of all system. Knots can exchange data and automatically update the gone parts of the stored information on the teams of the client.

The client can become authorized on any knot of a system for work with all system.

All information exchange between components of a system is carried out by means of channels in the form of virtual tunnels.

The system is capable to sustain mass shutdowns and damages of Knots, up to 50% and more, depending on the size of a system and configuration parameters of an algorithm.

The system does not demand certification regarding use crypto - algorithms as it does not use enciphering for protection of the stored data. It is possible to read data, having only restored them from the parts of information "smeared" in a system on "Client". Only the owner of a system (creator) or the one to whom the rights were delegated by him can restore data. Delegation of the rights does not mean transfer of rights to possession. The owner always has complete control over any changes in its files.

Node is the software of knot installed on servers in local network of Datacenter I form network of storage of SAN. Knots interact with each other, updating information on the user database of a cloud and notifying neighbors on changes in structure of knots. Thus the transparent scalability of a system is provided.

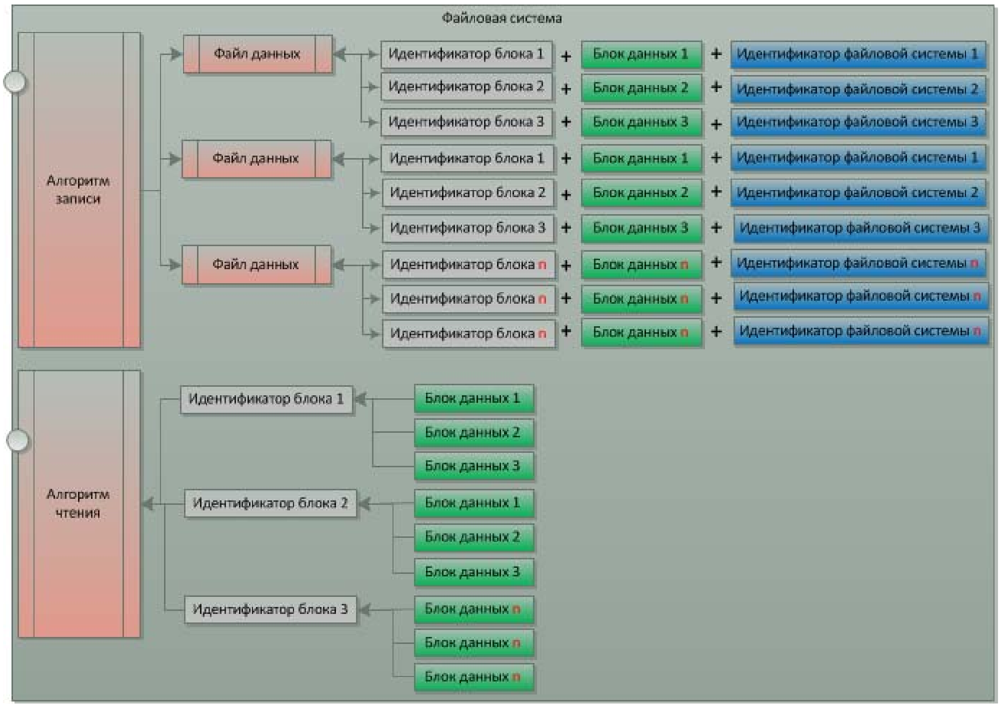
Файловая система
Алгоритм записи
Файл данных
Идентификатор блока 1 + Блок данных 1 + Идентификатор файловой системы 1
Идентификатор блока 2 + Блок данных 2 + Идентификатор файловой системы 2
Идентификатор блока 3 + Блок данных 3 + Идентификатор файловой системы 3
Файл данных
Идентификатор блока 1 + Блок данных 1 + Идентификатор файловой системы 1
Идентификатор блока 2 + Блок данных 2 + Идентификатор файловой системы 2
Идентификатор блока 3 + Блок данных 3 + Идентификатор файловой системы 3
Файл данных
Идентификатор блока n + Блок данных n + Идентификатор файловой системы n
Идентификатор блока n + Блок данных n + Идентификатор файловой системы n
Идентификатор блока n + Блок данных n + Идентификатор файловой системы n
Алгоритм чтения
Идентификатор блока 1
Блок данных 1
Блок данных 2
Блок данных 3
Идентификатор блока 2
Блок данных 1
Блок данных 2
Блок данных 3
Идентификатор блока 3
Блок данных n
Блок данных n
Блок данных n

Figure 3 – Scheme of work of the file system

The system is designed in such a way that all main innovative part of the functionality providing privacy of the stored data and resistance to mass losses of files on knots is implemented only in the client software (figure 3).

In the current version, the mode of access to data on the individual keys developed by the splitting module used in the only place namely at the time of reception/data transmission by the client to knots is realized. Thus, outside the client application with initialized by the user key access to data is impossible, and in a cloud, they are stored in the split unreadable look.

Testing of a security system. The considered security technology in cloud systems is installed by the author on the servers of Kazakhstan hosting company LLP "Internet company PS". 3 servers are used, each of which runs three services in SaaS (Software as a Service) mode. As a result of the use of such a configuration, the goal of a full-fledged layout of the system of 9 nodes has been achieved [20].

The configuration of 9 nodes corresponds to the second level of splitting, which allows achieving resistance to losses of parts from 3 to 5. As shown by the testing of the security system under consideration, conducted by the British company Locked Space Technology, Ltd (figure 4), the loss of 3 pieces of

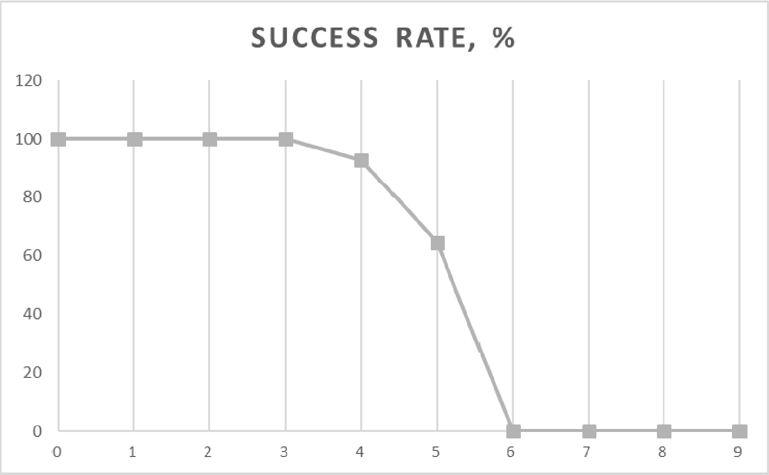

Figure 4 – The results of testing the security system (Locked Space Technology, Ltd)

split files allows you to restore the original file with a probability of 100%. That is, the system has an absolute resistance to file losses up to 33%. Loss of 4 files out of 9 (44.4%) allows you to recover a file with a probability of 92.8%. Even the loss of 5 parts of the split file (55.5%) still has a fairly high probability of 64.3% recovery.

The servers are joined by an internal network of 100 MB/sec. As shown by the preliminary performance tests of the system conducted at the stage of research and development, the speed of splitting/Assembly processes in this implementation depends mainly on the performance of the disk subsystem and the level of splitting. For the second level of splitting, the download and read speeds are between 80-100 MB/sec. Taking into account the existing network capacity of 70-80%, the utilization of the channel can be considered satisfactory. Given the speed of external access from the Internet to the cluster of nodes on average no more than 2-5 MB/s, we can assume that the speed of data access will not be limited to the splitting/Assembly subsystem. The utilization rate of the external access channel to the internal SAN:

$$K_{eff} = \frac{V_{IO}}{\left(\frac{V_{SAN}}{9}\right)} \cong 45\%,$$

where V_{IO} – the speed of Internet access and V_{SAN} – speed of access to nodes in the internal network.

External testing. Modern security methods are implemented by encrypting data, providing security around the perimeter and at the end points, verification protocols and control of employees. A system that stores data in a split form does not need additional protection from unauthorized access since split files do not carry any meaningful content.

The Swiss company Equivalence AG, which is interested in our results of studies on the security of distributed systems with data splitting, has independently tested the security parameters of the main basic subsystems of our technology. In particular, various types of cryptographic attacks on data stored in a split form and on traffic circulating in the system were simulated. The test results are shown in table 1.

Table 1 – The testing results of the distributed storage of information
with a breakdown of the data conducted by an Equivalence AG

No	Name of test element	Test parameter	Amount	Results
1	Traffic	Attempt to determine the type of document: docx/odt, jpg, pdf, mp3, txt.	~15GB	Not recognized
2	Traffic	Search in the stream of recognizable symbols	~15GB	Not recognized
3	Login Process	Attempt to intercept the password	10 000 tests	No success
4	Login Process	Trying to Client software spoofing (MITM attack)	10 000 tests	No success
5	Split Files/Metafiles	Brute Force	8760 hours	No success

Of particular interest is the Brute Force attack on split files in the conditions of a specially created on the servers of the company "sandbox", in the form of an isolated virtual machine with a processor pool in 48x10 cores, with a clock frequency of 3.6 GHz. For the organization of attack, the split files of the 3rd level of splitting were selected. As can be seen from the above table of the test Protocol (5th line), the attack lasted 8760 hours, i.e. 1 year of continuous search of bits in 480 parallel streams, which did not give any results.

As you know, the number of search combinations **n** for each level will be equal to the number **k** of ordered sets from the array **j** elements, i.e. it will be equal to the placement A_j^k:

$$n = A_j^k = \frac{j!}{(j-k)!}$$

The number of possible bit permutations is: $s = P_\tau = \tau!$, where τ – the length of the bit sequence.

If you need to choose a sequence of bits in the i files from which you want to collect the original file, we get:

$$m = s1 \cdot s2 \cdot s3 \ldots \cdot si,$$

where *m* the resulting number of combinations and permutations of the bits in the file.

In our algorithm, the split parts are equal to each other, so

$$m = s1 \cdot s2 \cdot s3 \cdot \ldots si = s^{i}$$

And with the minimum required number of *k* parts for assembly of the file obtained the dependence of the level splitting *λ*:

$$k = 2^{\lambda}, m = (\tau!)^{k}$$

Get the General formula for calculating the number of combinations:

$$n \cdot m = \left\{\frac{j!}{(j-k)!}\right\} \cdot \{(\tau!)^{k}\}$$

Here, the total number of files and the minimum required number of file parts are related to the split level *λ* relations: $j = 3^{\lambda}$ and $k = 2^{\lambda}$ – when the length of the sequence of bits to be rearranged = 16 (τ = 16, τ! = 20 922 789 888 000).

The number of search combinations at the 3rd level is for a file with the length of 16 byte $1.35 \cdot 10^{213}$ (table 2). For example, in AES 256 the number of combinations of key search is 10^{77}.

τ = 16 (2 bytes)	*k*	*m*
1 level	2	$\sim 4.38 \cdot 10^{26}$
2-level	4	$\sim 1.92 \cdot 10^{53}$
3-level	16	$\sim 1.35 \cdot 10^{213}$
4-level	32	$\sim 1.82 \cdot 10^{426}$

Table 2 – The number of search combinations

Discussion. The deployed distributed storage and information exchange system is easily scalable by simply adding servers with NODE software installed. Infrastructure expansion is easy and inexpensive. In addition, as the size of the infrastructure increases with the addition of storage units, the security and reliability of data storage automatically increase. And as you know, reliability should always be taken into account when deciding on operation and maintenance: "Reliability is the ability of technical devices to perform certain functions, maintaining their operation within the specified limits for the required period of time or the required operating time in certain operating conditions" [21].

Thus, it can be concluded that the parameters of information security in the cloud with distributed storage using the method of splitting data meet modern requirements for the security of cloud technologies.

To date, the most common method of preventing data loss is multiple backup and replication (real-time copying), i.e. creating multiple copies of the source file. Implemented the system requires no additional costly infrastructure of redundancy and Backup, and can be used, in particular, in the field of processing of large amounts of data on the Mapreduce model [22].

А. Ж. Карипжанова[1], К. М. Сагиндыков[1], А. М. Гудов[2], Kalin Dimitrov[3]

[1]Л. Н. Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан, Қазақстан,
[2]Кемеровоның ұлттық университеті, Кемерово, Ресей,
[3]Софияның техникалық университеті, Болгария

КӨП ӨЛШЕМДІ ЖҰПТЫҚ АЛГОРИТМДЕРІН ҚОЛДАНА ОТЫРЫП, ДЕРЕКТЕРДІ ЫДЫРАТУ КЕЗІНДЕ АҚПАРАТТЫҚ ЖҮЙЕЛЕРДІҢ ТАРАТЫЛҒАН ДЕРЕКТЕР ҚОРЫНЫҢ ПАРАМЕТРЛЕРІ

Аннотация. Қазіргі уақытта ақпараттың жоғалуын болдырмаудың жалпы қабылданған әдісі көп мәрте резервтеу болып табылады, бұл үлкен материалдық шығындарға әкеледі. Мақала авторлары әзірлеген Big Data қауіпсіздік жүйесі сақтау орындарының ішінара жоғалуына төзімді көп өлшемді айқындық алгоритмдерін қолдана отырып, жоғары қауіпсіздік деңгейін көрсетті, атап айтқанда бұлтты технологияларда пайдалану кезінде. Бұл алгоритмде деректер файлдардың көп санына бөлінеді, олардың әрқайсысы бұлтта шашыраған бастапқы ақпараттың бір битін де қамтуы мүмкін емес. Авторлар іске асырған жүйе қосымша қымбат резервтеу инфрақұрылымын талап етпейді, оңай масштабталады, кеңейтіледі, әрі сақтау бірліктерін қоса отырып, инфрақұрылым көлемінің ұлғаюына қарай деректерді сақтаудың қауіпсіздігі мен сенімділігі автоматты түрде артады. Деректерді ыдыратумен үлестірілген сақтау технологиясының негізгі базалық кіші жүйелерінің қауіпсіздік параметрлерін тестілеу нәтижелері алынды. Нәтижелер қауіпсіздік параметрлерінің қазіргі заманғы талаптарға сәйкестігін және тиісінше қымбат резервтеу инфрақұрылымы мен Backup басымдығын азайту мүмкіндігін көрсетті. Деректерді ажыратумен үлестірілген сақтау технологиясы сақтаудағы ақпараттың көптеген қатерлеріне және үлкен есептеулерге тиімді қарсы әрекет ету мүмкіндігін ашатын қауіпсіздіктің жаңа парадигмасын өзіне көрсетеді.

Түйін сөздер: ақпараттық қауіпсіздік, бұлыңғыр технологиялар, деректерді үлестірілген сақтау, ажырату.

А. Ж. Карипжанова[1], **К. М. Сагиндыков**[1], **А. М. Гудов**[2], **Kalin Dimitrov**[3]

[1]Евразийский национальный университет им. Л. Н. Гумилева, Нур-Султан, Казахстан,
[2]Кемеровский государственный университет, Кемерово, Россия,
[3]Технический университет Софии, Болгария

ПАРАМЕТРЫ РАСПРЕДЕЛЕННЫХ БАЗ ДАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ ПРИ РАСЩЕПЛЕНИИ ДАННЫХ С ПРИМЕНЕНИЕМ АЛГОРИТМОВ МНОГОМЕРНОЙ ЧЕТНОСТИ

Аннотация. К настоящему времени общепринятым методом предотвращения потери информации является многократное резервирование, что приводит к огромным материальным затратам. Разработанная авторами статьи система безопасности Big Data с применением алгоритмов многомерной четности, устойчивых к частичным потерям мест хранения, показала повышенный уровень безопасности, в частности при использовании в облачных технологиях. В этом алгоритме данные расщепляются на большое число файлов, каждый из которых не может содержать даже одного бита рассредоточенной в облаке исходной информации. Реализованная авторами система не требует дополнительной дорогой инфраструктуры резервирования, легко масштабируется, расширяется, причем по мере увеличения размеров инфраструктуры с добавления единиц хранения автоматически увеличивается безопасность и надежность хранения данных. Получены результаты тестирования параметров безопасности основных базовых подсистем технологии распределенного хранения с расщеплением данных. Результаты показали соответствие параметров безопасности современным требованиям и, соответственно, возможность уменьшения доминирования дорогой инфраструктуры резервирования и Backup. Технология распределенного хранения с расщеплением данных воплощает в себе новую парадигму безопасности, открывающую возможность эффективного противодействия многочисленным угрозам хранимой информации и большим вычислениям.

Ключевые слова: безопасность информации, облачные технологии, распределенное хранение, расщепление данных.

Information about authors:

Karipzhanova Ardak Zhumagazievna, doctoral candidate of ENU named after L. N. Gumilev, by specialty "Information systems", Eurasian national university. L. N. Gumilyova, Astana, Kazakhstan; kamilakz2001@mail.ru; https://orcid.org/0000-0002-0113-6132

Sagindykov Kakim Moldabekovich, associate professor, head of the Department of “Computer science and information security”, the candidate of Technical Sciences, Eurasian national university. L. N. Gumilyova, Astana, Kazakhstan; ksagin@mail.ru; https://orcid.org/0000-0003-3315-798X

Gudov Alexandr Mikhailovich, director of the Institute of fundamental sciences, associate professor, doctor of technical sciences, Kemerovo state university, Russia; good@kemsu.ru; https://orcid.org/0000-0002-4775-071X

Dimitrov Kalin, associate professor department of radiocommunications and videotechnologies, faculty of Telecommunications, Technical university of Sofia, Sofia, Republic of Bulgaria; kld@tu-sofia.bg; https://orcid.org/0000-0003-1104-4685

REFERENCES

[1] Sagindikov K.M. Algorithm for calculation of parameters of the bearing elements of oil heating installations // International journal of chemical sciences, 14: 1 (**2016**), 355-362.

[2] Kalimoldayev M.N., Suleimenov E.I., Pak I.T., Vitulyova E.S., Tasbulatova Z.S., Yevstifeyev V.N., Mun G.A. To the question of physical implementation of optical neural networks // News of the national academy of sciences of the Republic of Kazakhstan. Series of geology and technical sciences. Vol. 2, N 434 (**2019**), 217-224. ISSN 2518-170X (Online), ISSN 2224-5278 (Print). https://doi.org/10.32014/2019.2518-170X.57

[3] Kalimoldayev M.N., Pak I.T., Baipakbayeva S.T., Mun G.A., Shaltykova D.B., Suleimenov I.E. (2018). Methodological basis for the development strategy of artificial intelligence systems in the Republic of Kazakhstan in the message of the President of the Republic of Kazakhstan dated october 5, 2018 // News of the National academy of sciences of the Republic of Kazakhstan. Series of geology and technical sciences. 2018. Vol. 5, N 431. P. 62-68. ISSN 2518-170X (Online). ISSN 2224-5278 (Print). https://doi.org/10.32014/2018.2518-170X.34

[4] Kurmanbaev E.A., Syrgabekov I. N., Zadauly E. Karipzhanova A.Zh., Urazbaeva K.T. Information Security System on the Basis of the Distributed Storage with Splitting of Data // International Journal of Applied Engineering Research, 12 (**2017**), N 8, 1703-1711.

[5] Syrgabekov I., Zadauly E., Kurmanbaev E. Zashhita informacionnyh baz po metodu raspredelennogo hranenija // Doklady Nacional'noj akademii nauk Respubliki Kazahstan, 5 **(2014)**, 141-153 (in Rus.).

[6] Zadauly E., Kurmanbaev E., Syrgabekov I. Innovacionnaja sistema bezopasnosti na baze raspredelennogo hranenija informacii s rasshhepleniem dannyh // Patriot Engineering, 2: 7 **(2015)**, 111-119 (in Rus.).

[7] Tussupov J., Johnson J., Knight J.F., Ocasio V., Van Den Driessche S. Preserving Categoricity and Complexity of Relations // Algebra and Logic, 54 (**2015**), N 2, May, 140-154.

[8] Badaev S.A., Kalmurzayev B.S., Kabylzhanova D.K., Abeshev K.Sh. (2018) Universal positive preorders // News of the National academy of sciences of the Republic of Kazakhstan. Physical-mathematical series. Vol. 6, N 322 (**2018**), 49-53. ISSN 2518-1726 (Online), ISSN 1991-346X (Print). https://doi.org/10.32014/2018.2518-1726.17

[9] Kaziyev G.Z., Markosiyan M.B., Taurbekova A.A. Methods of distribution of data processing systems to the nodes of computing systems // News of the National academy of sciences of the Republic of Kazakhstan. Series of geology and technical sciences. Vol. 4, N 430 (**2018**), 124-131. ISSN 2518-170X (Online), ISSN 2224-5278 (Print).

[10] Sigal I.Kh. Parametrizasia priblizhennyh algoritmov reshenie nekotoryh klassov zadach diskretnoi optimizasii bolshoi razmernosti // Izvestiya RAN. Teoriya i sistemy upravlenia. **(2002)**. N 6. P. 63-72 (in Rus.).

[11] Sigal I.Kh., Ivanova A.P. Vvedenie v prikladnoe diskretnoe programmirovanie: modeli i vychislitelnye algoritmy. 2nd ed., rev., and add. M.: Fizmatlit, **(2007)**. 304 p. (in Rus.).

[12] Samigulina G.A., Nyusupov A.T., Shayakhmetova A.S. Analytical review of software for multi-agent systems and their applications // News of the National academy of sciences of the Republic of Kazakhstan. Series of geology and technical sciences. Vol. 3, N 429 (**2018**), 173-181. ISSN 2518-170X (Online), ISSN 2224-5278 (Print).

[13] Müller J.P., Fischer K. (**2014**). Application impact of multi-agent systems and technologies: A survey, Agent-oriented software engineering, 27-53. DOI: https://doi.org/10.1007/978-3-642-54432-3_3

[14] Flores-Mendez R.A. **(1999)**. Towards a standardization of multi-agent system framework, Crossroads. P. 18-24.

[15] Multi-Agent Development Kit. **(2017)**. URL: http://www.madkit.net/madkit/madkit.php (date of the application: 20.11.2017).

[16] Mourabit E.L., Toumanari A., Zougagh H.A. **(2014)** Mobile Agent Approach for IDS in Mobile Ad Hoc Network // JCSI International Journal of Computer Science, 2: 148-152. https://DOI:10.3844/jcssp.2014.970.975 .

[17] Agent Builder. **(2017)**. URL: http://www.agentbuilder.com /Documentation/product.html (date of the application: 20.11.2017).

[18] Cognitive Agent Architecture. **(2017)**. URL: cougaar.org (date of the application: 21.11.2017).

[19] Zeus Agent Toolkit. URL: labs.bt.com/projects/agents/zeus (date of the application: 21.11.2017).

[20] Goudov F.V., Perminov V.A. Mathematical simulation of contaminant flow in the square reservoir // International Journal of Geomate. 11(**2016**), 2558-2562.

[21] Akhmetov J.W., Seitova S.M., Toibazarov D.B., Kadyrbayeva G.T., Dauletkulova A.U., Issayeva G.B. Verification of reliability technical devices through resolving probability of failure and failure // News of the National academy of sciences of the Republic of Kazakhstan. Physical-mathematical series. Vol. 5, N 321 (**2018**), 49-61. ISSN 2518-1726 (Online), ISSN 1991-346X (Print). https://doi.org/10.32014/2018.2518-1726.7

[22] Shomanov A.S., Akhmed-Zaki D.Zh., Amirgaliyev E.N., Mansurova M.E. About the problem of key distribution in Mapreduce model // News of the National academy of sciences of the Republic of Kazakhstan. Physical-mathematical series. Vol. 3, N 313 (**2017**), p. 167-172. ISSN 2518-1726 (Online), ISSN 1991-346X (Print).

Publication Ethics and Publication Malpractice in the journals of the National Academy of Sciences of the Republic of Kazakhstan

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайте:

www:nauka-nanrk.kz

ISSN 2518-170X (Online), ISSN 2224-5278 (Print)

http://www.geolog-technical.kz/index.php/en/

Верстка *Д. Н. Калкабековой*

Подписано в печать 14.10.2019.
Формат 70х881/8. Бумага офсетная. Печать – ризограф.
15,0 п.л. Тираж 300. Заказ 5.

Национальная академия наук РК
050010, Алматы, ул. Шевченко 28, т. 272-13-19, 272-13-18