

UOT 004.75

Ələkbərov R.Q.¹, Paşayev F.H.², Həşimov M.A.³

^{1,3}AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan

²AMEA İdarəetmə Sistemləri İnstitutu, Bakı, Azərbaycan

¹rashid@iit.ab.az, ²pasha.farhad@gmail.com, ³mamedhashimov@gmail.com

VERİLƏNLƏRİN EMAL MƏRKƏZLƏRİNDƏ İSTİFADƏÇİLƏRİN MƏLUMATLARININ YADDAŞ QURĞULARINDA DİNAMİKİ PAYLANMASI METODUNUN İŞLƏNMƏSİ

Məqalədə verilənlərin emal mərkəzlərində (VEM-də) məlumatların bulud texnologiyaları əsasında yaddaş resurslarında dinamik paylanması məsələsinə baxılmışdır. İstifadəçilərin məlumatlarının istifadə intesivliyindən asılı olaraq müxtəlif xarakteristikali yaddaş qurğuları arasında dinamiki paylanması metodu təklif olunmuşdur. Metod yaddaş resurslarından istifadəçilərin daha səmərəli istifadə etməsinə imkan verir.

Açar sözlər: *verilənlərin emal mərkəzi, bulud texnologiyası, bulud xidmətləri, yaddaşın həcmi, məlumatlara müraciət tezliyi, virtual resurs, yaddaş resursu, sadə və çəkili sürüşən orta qiymət metodu.*

Giriş

Hal-hazırda bulud texnologiyalarının (*ing. cloud computing*) köməyi ilə VEM-in hesablama və yaddaş resurslarından səmərəli istifadə etmək istiqamətində intensiv tədqiqat işləri aparılır [1]. Hal-hazırda kompüter şəbəkələri əsasında böyük hesablama və yaddaş resursları tələb edən mürəkkəb məsələlərin həlli üçün paylanmış hesablama sistemlərinin yaradılmasında bulud texnologiyalarından istifadə edilir. Bulud texnologiyası imkan verir ki, təşkilatların emal mərkəzlərinin istifadəsiz qalan hesablama və yaddaş resurslarından daha səmərəli istifadə edilsin. Bu texnologiyanın köməyi ilə istifadəçinin məlumatları hesablama buludlarının serverlərində saxlanılır, emal edilir və eyni zamanda brauzerlərin köməyi ilə nəticələrə baxılması təmin edilir. Bulud texnologiyası emal mərkəzlərinin hesablama və yaddaş resurslarının klasterləşməsi və virtuallaşdırılmasını təmin edir [2].

Bulud texnologiyası fiziki resursları (məsələn: prosessor və disk yaddaş fəzasını) miqyaslaşdırmaq, İnternet vasitəsi ilə bu resurslardan istifadə etmək imkanları verir. Bu halda, məlumatların emalı və yadda saxlanması prosesinə bir xidmət növü kimi baxılır. Qeyd edilən texnologiyanın köməyi ilə VEM-in yaddaş resurslarını istifadəçilər arasında optimal paylamaqla ondan daha səmərəli istifadə etmək olar [3].

VEM-də yaddaş resurslarının istifadəçilər arasında dinamiki paylanması və məsələnin qoyuluşu

Məlumdur ki, VEM-də virtual yaddaş resurslarının yaradılmasında SAN (*ing. Storage Area Network*) texnologiyasından geniş istifadə edilir. Ənənəvi SAN texnologiyalarının tətbiqi zamanı istifadəçi üçün ayrılmış resurs tam olaraq onun istifadəsinə verilir. İstifadəçi bu resursdan tam istifadə etmədikdə bu resursdan digər istifadəçilər istifadə edə bilmir. Nəticədə əgər istifadəçi ona ayrılan resursun yarısından istifadə edərsə digər yarısı boş qalır [4].

Virtual yaddaş resurslarının istifadəçilər arasında optimal paylanması və səmərəli istifadə edilməsi üçün hipervizordan (*ing. hypervisor*) istifadə olunur [5]. Bu sistem ənənəvi SAN texnologiyasının çatışmazlıqlarının aradan qaldırılmasına kömək edir.

Hipervizorun tətbiq edilməsi ilə virtual yaddaş resurslarının idarə edilməsində yeni funksiyalardan istifadə olunur. Bu yeni funksiyalara proseslərin avtomatlaşdırılması, qəzalar zamanı avtomatik bərpa mexanizmi, rezervlərin yaradılması, yaddaş resurslarının dinamik paylanması (*ing. thin provisioning*) və faylların (məlumatların) avtomatik olaraq səviyyələr (müxtəlif xarakteristikali yaddaş avadanlıqları) üzrə paylanması (*ing. auto-tiering*) və s.

daxildir [6]. *Thin provisioning* və *auto-tiering* texnologiyalarının birlikdə istifadə olunması virtual yaddaş resurslarının idarə olunmasında və paylanmasında effektivliyin artırılmasına böyük təsir göstərir.

Konkret olaraq, axırını iki texnologiyanın birləşməsi bulud texnologiyalarında yaddaş resurslarının effektiv istifadəsinə böyük təsir göstərir.

SAN sistemlərdə *thin provisioning* texnologiyasının tətbiqi imkan verir ki, istifadəçi ona ayrılan resursun yalnız onun hal-hazırda istifadə etdiyi həcmi əlində saxlasın [7]. Beləliklə, resurslar daha səmərəli istifadə olunur. İstifadəçi ona ayrılmış yaddaş resursunu doldurduqca, ümumi ehtiyat resurslardan onun tələbləri bərpa olunur. *Thin provisioning* texnologiyasından yaddaş resurslarının virtuallaşdırılmasında geniş istifadə olunur və yaddaş resurslarının israf olunmadan istifadə olunmasını təmin edir.

Bu texnologiya imkan verir ki, hər hansı bir məqsəd üçün ayrılmış resurs yaddaş qurğusunda tam olaraq istifadə olunduğu qədər yer tutsun. Ayrılan resurs istifadə olunmadığı müddətdə bron edilmir və sərt disklərdə yer tutmur. Bu texnologiya mərkəzləşdirilmiş böyük yaddaş resurslarını planlı şəkildə istifadə etməyə imkan verir.

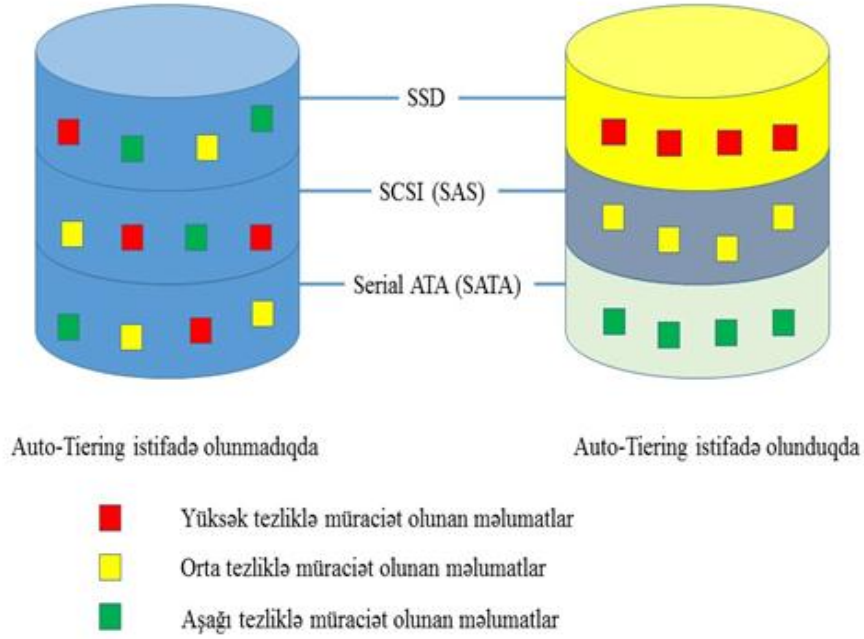
Beləliklə, resurslar yalnız istifadə olunduğu qədər ayrılır və israf olunmur. İşin kommertiya tərəfi həm bulud provayder, həm də istifadəçi üçün sərfəlidir. Belə ki, istifadəçi bron etdiyi resursun deyil, sadəcə faktiki istifadə etdiyi resursun pulunu ödəyir. Provayder isə həm əlavə avadanlığın alınması, qurulması, yerləşdirilməsi kimi xərclərdən azad olur, həm də eyni xidmət müqabilində daha aşağı qiymət təklif edə bilər ki, bu da daha çoxlu sayda istifadəçinin bu tip xidmətlərə cəlb olunmasına gətirib çıxarır [3].

Məlumdur ki, VEM-də yaddaş resursları müxtəlif xarakteristikalara malik olan yaddaş qurğuları (SSD (*ing. Solid State Disks*), maqnit disk qurğuları və s.) əsasında yaradılır. Qeyd edilən yaddaş qurğularının qiymətləri də müxtəlif olur.

İstifadəçilərin məlumatlarının bu yaddaş qurğularının hansında yerləşdirilməsindən asılı olaraq icarəyə götürülmüş resursa görə ödəniş pulu müxtəlif olur. İstifadəçi VEM-də yerləşdirdiyi məlumatlardan il ərzində çox az istifadə edərsə, onda onun yaddaş resursu kimi maqnit disk qurğusunun resurslarından istifadə etməsi daha yaxşı olar. Bu halda sərf edilən xərcin həcmi də az olacaqdır. Yüksək sürətli qurğuların (*ing. SSD*) qiymətləri baha olduğundan, istifadəçi bu qurğulardan istifadə etdikdə, təbii ki, ödəmə xərcləri də çox olur.

Virtual yaddaş resurslarının optimal paylanması və onlardan istifadənin səmərəliliyinin təmin olması üçün vacib komponentlərdən biri də faylların (məlumatların) səviyyələrə (yaddaş qurğularının tipləri üzrə) görə avtomatik paylanması funksiyasıdır. Onun köməyi ilə pik yüklənmələrin qarşısını almaq və faylları istifadə olunma xüsusiyyətlərinə görə səviyyələrə bölmək mümkün olur. Bu funksiya *auto-tiering* adlanır. Bu funksiyanın köməyi ilə VEM-in virtual yaddaş resurslarının birinci səviyyəsi yüksək sürətli SSD yaddaş qurğularından təşkil olunur.

Sonrakı səviyyədə SSD yaddaş qurğularından daha az sürətlə işləyən və qiyməti ucuz olan serial Ataachid SCSI (SAS) maqnit disk qurğularından istifadə olunur. Üçüncü səviyyədə qeyd edilən qurğulara nəzərən aşağı sürətlə işləyən və onlarla müqayisədə daha da ucuz olan serial ATA (*ing. SATA*) qurğularından istifadə olunur (şəkil 1) [8]. Yuxarıda deyilənləri nəzərə alsaq, bulud texnologiyaları əsasında istifadəçilərin məlumatlarının hansı yaddaş qurğularında yerləşdirilməsi aktual məsələlərdəndir.



Şəkil 1. Yaddaş qurğularının səviyyələr üzrə yerləşdirilməsi

İstifadəçilərin məlumatlarının yaddaş resurslarının səviyyələri arasında dinamik paylanması metodunun işlənməsi

Məqalədə faylların (məlumatların) sonlu sayda səviyyələr (yaddaş qurğuları) arasında dinamik paylanması məsələsinin bir həlli göstərilmişdir.

Faylların sonlu sayda səviyyələr arasında dinamik paylanması məsələsinin həlli üçün aşağıdakıları qəbul edək:

- səviyyələrin sayı - M ;
- səviyyələrin ümumi yaddaş həcmi - V ;
- m -ci səviyyənin yaddaş həcmi: - V_m ;
- faylların sayı - I ;
- i -ci faylın həcmi - F_i ;
- Müşahidələrin sayı - N ;
- i -ci fayla n -ci müşahidə anı - T_{in} ;
- i -ci fayla n -ci müşahidə anında müraciətlərin cari tezliyi - f_{in} ;
- i -ci fayla müraciətlərin orta tezliyi - f_i ;
- i -ci fayla n -ci müşahidə anında müraciətlərin cari periodu - h_{in} ;
- i -ci fayla müraciətlərin orta periodu - h_i .

Səviyyələrin malik olduqları yaddaş həcmələri və faylların həcmələri aşağıdakı ilkin şərtləri ödəməlidirlər:

1-ci şərt - səviyyələrin yaddaş həcmələrinin cəmi ümumi yaddaş həcmi olan V -yə bərabər olmalıdır:

$$\sum_{m=1}^M V_m = V, \quad m \in [1, M]; \quad (1)$$

2-ci şərt – istifadəçilər tərəfindən müraciət edilən faylların ümumi həcmi ümumi yaddaş həcmnin p hissəsindən (məsələn, 70 faizindən) çox olmamalıdır:

$$\sum_{i=1}^I F_i \leq 0.7 * V; \quad (2)$$

Qeyd etmək lazımdır ki, i -ci fayla n -ci müşahidə anında müraciətlərin cari intervalı

$$h_{in} = T_{in} - T_{i(n-1)} \quad (3)$$

kimi hesablanır. i -ci fayla n -ci müşahidə anında müraciətlərin cari tezliyi - f_{in} ilə cari periodu - h_{in} arasında $f_{in} = \frac{1}{h_{in}}$ münasibəti vardır.

Fayllara müraciət tezliyi təyin edilərkən faylların atributlarından və müraciət üsullarından istifadə edilir [9].

Məsələnin həllinə hazırlıq proseduraları hər hansı müşahidə anında fayllara müraciətlərin tezliklərini nizamlı şəkildə düzməkdən ibarətdir. Ona görə də onları səviyyələr arasında bölüşdürmək olar. Tutaq ki, səviyyələrə bölüşdürmə nəticəsində səviyyələrə k_1, k_2, \dots, k_M sayda fayllar yerləşdirilmişdir. Bu bölüşdürmə (1), (2) başlanğıc şərtlərə uyğun olaraq aşağıdakı şərtləri ödəməlidir:

$$\sum_{j=1}^M k_j = I; \quad (4)$$

m -ci səviyyəyə qədər olan $m-1$ sayda təbəqədə $M_{m-1} = \sum_{j=1}^{m-1} k_j$ sayda fayl yerləşdirilmiş olar. m -ci səviyyədə yerləşmiş k_m sayda fayl ikinci başlanğıc şərtinə (2) uyğun olaraq

$$\sum_{i=1}^{m_i} F_{M_{i-1}+i} \leq 0.7 V_m \quad (5)$$

şərtini ödəyər.

Nəticə etibarlı ilə bu bölgü faylların səviyyələr üzrə klasterizasiya məsələsinin həllini verir. Bölgünü həyata keçirən proqram təminatı hər hansı m -ci səviyyədə faylları yerləşdirərkən yerləşməni sıra nömrəsi $M_{m-1} + 1$ olan fayldan başlayaraq (5) şərti pozulana qədər davam etdirir.

Göstərilən metod sadə olsa da, burada fayllara müraciətlərin tarixi nəzərə alınmır. Cari müraciət zamanı formalaşan tezliklər təsadüfi xarakter daşıya bilər. Ona görə də əvvəlki müşahidələr zamanı formalaşan tezlikləri də nəzərə almaq lazımdır. Bu məsələni sürüşən orta qiymət metodlarının müxtəlif variantlarının tətbiqi ilə həll etmək olar [10–12].

Sadə sürüşən orta qiymət metodu: i -ci fayla n -ci müşahidə anında müraciətin tezliyi f_{in} olduğundan orta qiyməti k sayda elementdən hesablasaq

$$SMA_k = \frac{f_{in} + f_{i(n-1)} + \dots + f_{i(n-(k-1))}}{k} \quad (6)$$

Burada SMA_k (ing.simple moving average) - sadə sürüşən orta qiymətdir. Klasterizasiya üçün f_{in} əvəzinə SMA_k götürmək lazımdır.

Çəkili sürüşən orta qiymət metodu: sadə sürüşən orta qiymətlər tətbiq edildikdə müxtəlif müşahidə anlarında hesablanmış k sayda tezlik qiymətlərinin hər biri eynigüclü hesab edilir. Bəzən sonuncu qiymətlərə üstünlük vermək əhəmiyyətli ola bilər. Ona görə də çəkili sürüşən orta qiymətin aşağıdakı düsturunu tətbiq etmək olar:

$$WMA_k = \frac{k * f_{in} + (k-1) * f_{i(n-1)} + \dots + f_{i(n-(k-1))}}{k + (k-1) + \dots + 1}, \quad (7)$$

Burada WMA_k (ing.weighted moving average) - çəkili sürüşən orta qiymətdir və tezliklərin axırıncı qiymətləri daha böyük çəkiyə malikdirlər. Analoji olaraq klasterizasiya üçün f_{in} əvəzinə VMA_k götürmək lazımdır.

Ekspensial hamaralama metodu: ekspensial hamaralama zamanı hər hansı α ($0 < \alpha < 1$) hamaralama parametri götürərək axtarılan parametri aşağıdakı kimi hesablamaq olar:

$$AR_k = a * f_{in} + a(1 - a)f_{i(n-1)} + a(1 - a)^2 f_{i(n-2)} + \dots \quad (8)$$

(8) düsturundan görünür ki, α parametrinin qiymətləri 1-ə yaxınlaşdıqca, sonuncu qiymətlərə daha çox üstünlük verilmiş olur. Praktiki tətbiqlər zamanı α parametri 0,01-dən 0,30-a kimi qiymətlər ala bilər.

Beləliklə, biz fayllara müraciət tezliklərinin aşağıda qeyd olunan 4 metoddan biri ilə alınmış qiymətlərindən istifadə edə bilərik:

- Cari qiymətlər. Bu zaman cari qiymətlərin təsadüfi dəyişmələri hər hansı bir faylın səviyyələri arasında yerdəyişməsinə və yeni qiymətin özünü təsdiq etməyəcəyi hallarda prosesin balansının pozulmasına səbəb ola bilər.
- Sadə sürüşən orta qiymət metodu ilə təyin edilmiş qiymətlər. Bu zaman cari qiymətlərin təsadüfi dəyişmələri ciddi balans pozğunluğuna səbəb ola bilmir. Tezliyin yeni qiyməti özünü təsdiq etdiyi halda fayl yavaş-yavaş öz tezliyinə uyğun təbəqədə yer alır.
- Çəkili sürüşən orta qiymət metodu ilə təyin edilmiş qiymətlər. Bu zaman cari qiymətlərin təsadüfi dəyişmələri ciddi balans pozğunluğuna səbəb ola bilmir. Tezliyin yeni qiyməti özünü təsdiq etdiyi halda fayl öz tezliyinə uyğun təbəqədə sadə sürüşən orta qiymət metodu ilə müqayisədə daha tez yer alır.
- Eksponensial hamarlaşdırma metodu ilə təyin edilmiş qiymətlər. Bu zaman cari qiymətlər hamarlanır və onların təsadüfi səbəblərdən qiymət dəyişmələri yekun nəticəyə təsir edə bilmir. Tezliyin yeni qiyməti özünü təsdiq etdiyi halda uyğun səviyyədə kifayət qədər tez yer tuta bilər.

Yekun olaraq, faylların sonlu sayda səviyyələri arasında dinamik paylanması məsələsinin həlli üçün aşağıdakılar həyata keçirilməlidir:

1. Müşahidə anlarında eksperimental verilənlərdən fayllara müraciətlərin tezlikləri və periodları formalaşdırılmalıdır;
2. Cari qiymətlər, sadə sürüşən orta qiymətlər və ya çəkili sürüşən orta qiymətlərdən biri seçilməklə klasterizasiya üçün orta tezliklər formalaşdırılmalıdır;
3. Orta tezliklər və onlara uyğun olan fayllar qiymətlərinin artması ardıcılığı ilə nizamlı düzüləlidir;
4. (5) şərti ödənilməklə klasterizasiya və səviyyələrə görə bölgü həyata keçirilməlidir.

Nəticə

Məqalədə istifadəçilərin yaddaş resurslarında yerləşdirdikləri məlumatlara müraciət intensivliyi təhlil edilmiş və bunun əsasında istifadəçiyə müraciət tezliyinə uyğun hansı yaddaş qurğusundan istifadə edilməsi təklif olunmuşdur. Eyni zamanda faylların (məlumatların) sonlu sayda səviyyələri (yaddaş qurğuları) arasında dinamik paylanmasının həlli məsələsi qarşıya qoyulmuşdur. Bu zaman fayllara müraciətlərin müşahidə anlarında formalaşan cari qiymətlərindən, sadə sürüşən orta, çəkili sürüşən orta və ya eksponensial hamarlaşdırma metodu ilə tapılmış qiymətlərdən istifadə etməklə səviyyələri arasında klasterizasiya metodu verilmişdir. Məqalədə təklif olunan sürüşən orta qiymət və ya çəkili sürüşən orta qiymət metodu analoji məsələlərin həllində istifadə edilə bilər.

Minnətdarlıq

Bu iş Azərbaycan Respublikasının Prezidenti yanında Elmin İnkişafı Fondunun maliyyə yardımını ilə yerinə yetirilmişdir – **Qrant № EIF-2014-9(24)-KETPL-14/02/1.**

Ədəbiyyat

1. Alguliev R.M., Alakbarov R.G. Cloud Computing: Modern State, Problems and Prospects / Telecommunications and Radio Engineering, USA, 2013, vol.72, no.3, pp.255–266.
2. Marios D. Dikaiakos, George Pallis, Dimitrios Katsaros, Pankaj Mehra, Athena Vakali. Cloud Computing -Distributed Internet Computing for IT and Scientific Research / IEEE Internet Computing, 2009, № 9, pp.10–13.

3. Alakbarov R.G., Pashaev F.H., Hashimov M.A. Development of the Model of Dynamic Storage Distribution in Data Processing Centers / Information Technology and Computer Science, 2015, no.5, pp.18–24.
4. An introduction to thin provisioning. <http://searchstorage.techtarget.com>
5. Storage hypervisor tutorial. <http://searchvirtualstorage.techtarget.com>
6. Jeremy LeBlanc, Adam Mendoza, Mike McNamara, NetApp Ashish Yajnik, Rishi Manocha, Symantec. Maximize Storage Efficiency with NetApp Thin Provisioning and Symantec Thin Reclamation. September 2010, WP-7111.
7. Floyer David. Thin provisioning. [http://wikibon.org/wiki/Thin provisioning](http://wikibon.org/wiki/Thin_provisioning)
8. Automated Storage Tiering. www.datacore.com
9. Tanenbaum A.S. Modern Operating Systems. Second edition. Piter, 2007, 1038 p.
10. Anatoly Homonenko and others. Delphi 7. St. Petersburg, "BHV Petersburg", 2008, p.1200.
11. Thomopoulos N.T. Applied Forecasting Methods. Prentice-Hall, New Jersey, 1980, 369 p.
12. Hyndman R.J., Athanasopoulos G. Forecasting: principles and practice. www.otextst.org/book/fpp

УДК 004.75

Алекперов Рашид Г.¹, Пашаев Фахрад Г.², Гашимов Мамед А.³

^{1,3}Институт Информационных Технологий НАНА, Баку, Азербайджан

²Институт Систем Управления НАНА, Баку, Азербайджан

¹rashid@iit.ab.az, ²pasha.farhad@gmail.com, ³mamedhashimov@gmail.com

Разработка метода динамического распределения данных о пользователях в устройствах памяти в центрах обработки данных

В статье рассматривается вопрос динамического распределения данных в ресурсах памяти пользователей центрах обработки данных на основе облачных вычислений. Предложен метод динамического распределения между устройствами памяти с различными характеристиками в зависимости от интенсивности использования данных пользователями. Метод обеспечивает более эффективное использование ресурсов памяти пользователями.

Ключевые слова: центр обработки данных, облачные технологии, облачные услуги, размер памяти, частота обращения к данным, виртуальные ресурсы, ресурсы памяти, метод простой величины, метод средних скользящих величины.

Rashid G. Alakbarov¹, Fahrhad H. Pashayev², Mammad A. Hashimov³

^{1,3}Institute of Information Technology of ANAS, Baku, Azerbaijan

²Institute of Control Systems of ANAS, Baku, Azerbaijan

¹rashid@iit.ab.az, ²pasha.farhad@gmail.com, ³mamedhashimov@gmail.com

Development of the method of dynamic distribution of users' data in storage devices at data processing centers

The paper reviews a dynamic data distribution in storage resources at data processing centers based on cloud technologies. A method is proposed for the dynamic distribution of users' data among storage devices with various characteristics depending on the intensity of the data usage. The method allows the users using storage resources more efficiently.

Keywords: data processing center, cloud technology, cloud services, storage capacity, data call frequency, virtual resource, storage resource, simple and weighted moving average method.