

МЕТОД РЕЗЕРВУВАННЯ ТА ПРИСКОРЕНОГО ВІДНОВЛЕННЯ ДАНИХ В СИСТЕМАХ ЇХ ВІДДАЛЕНОГО ЗБЕРІГАННЯ

В статті запропоновано метод резервування та відновлення даних. Ці дані зберігаються на різних віддалених носіях, до p з яких постійно або тимчасово втрачено доступ. Висока ефективність запропонованого методу забезпечується тим, що використовується теоретично мінімальна кількість резервних носіїв, яка дорівнює кількості носіїв, до яких втрачено доступ. Детально викладені математична ідея та процедури відновлення даних, доступ до яких втрачено. Розроблений метод ілюстровано прикладами. Наведено теоретичні та експериментальні оцінки запропонованого методу. Розроблений метод має на меті забезпечення високого рівня надійності систем розподіленого зберігання даних.

In this work, a method for backup and restoring data is proposed. The data is stored on different remote disks and temporary or permanent loss of access to any n of them appears. The high efficiency of the proposed method is achieved by using a theoretically proved minimum number of backup copies of the data, which is equal to the number of the disks to which access has been lost. The mathematical idea of proposed method and procedure for recovering of data from access lost storage unit are described in details. A numerical example for developed recovering procedure are given. The theoretical and experimental effectiveness evaluation of the proposed method is demonstrated as well. The proposed method is aimed to ensure a high level of survivability of system of distributed data storage.

Вступ

Швидкий прогрес засобів телекомунікацій та технології комп'ютерних мереж на початку нового тисячоліття призвів до динамічного розвитку систем віддаленого зберігання та обробки даних. Основною перевагою віддаленої обробки інформації в рамках хмарних технологій є можливість надання користувачам значних за обсягом обчислювальних ресурсів та пам'яті [1].

Динамічний розвиток розподілених систем обробки та зберігання даних неможливий без ефективного вирішення ряду наукових задач, чільне місце серед яких посідає забезпечення надійності та безперебійності доступу до віддаленої інформації користувачів. Вирішення цих важливих задач вимагає застосування спеціальних засобів резервування та відновлення даних, до яких втрачено доступ. За умови швидкого зростання об'ємів такої інформації і розширення кола користувачів технологій віддаленого зберігання даних набуває ваги задача підвищення ефективності методів та засобів резервування та відновлення інформації в рамках таких технологій. Це вимагає постійного пошуку нових підходів, методів та розробки засобів забезпечення неперервності доступу до даних шляхом їх відновлення з використанням резервної інформації.

Таким чином, задача підвищення ефективності та розширення функціональних можливостей засобів резервування та відновлення даних при їх зберіганні на віддалених та розподілених носіях є актуальною та практично важливою з огляду на сучасний стан розвитку технологій комп'ютерної обробки інформації.

Аналіз існуючих систем резервування та відновлення даних

При віддаленому зберіганні даних користувача вони розподіляються по окремим вузлам зберігання інформації. В рамках окремого вузла організовано розподілення даних по носіях, доступ користувачів до даних, їх захист, резервування в разі втрати доступу до одного або декількох носіїв.

Для забезпечення високого рівня неперервності доступу кожному користувачеві до своєї інформації, що зберігається на віддалених носіях, найчастіше використовується їх резервування.

В якості критеріїв ефективності систем резервування найчастіше виступають:

- обчислювальна складність процедури відновлення даних з використанням резервних носіїв;
- кількість носіїв (K_B), дані з яких можуть бути відновлені в разі втрати доступу до них;

– відношення кількості резервних носіїв (K_p) до числа носіїв, дані з яких можуть бути відновлені ($\alpha = K_p / K_B$).

Проведений аналіз літературних джерел [2] показав, що найчастіше причинами втрати доступу до даних віддалених користувачів стає вихід з ладу окремих носіїв (або стирання інформації на них), а також тимчасова нездатність вузлу обслужити запит користувача.

При втраті даних на окремих носіях в результаті виходу їх з ладу, або стирання (помилкового чи цілеспрямованого) даних на них, кількість q таких носіїв залежить від часу t , що пройшов від моменту останнього звертання до них. Тимчасова нездатність вузлу обслужити запит користувача здебільшого пов'язана з перевантаженістю вузла зберігання даних, виходом з ладу його апаратних чи програмних компонентів, економічними чи природними катаклізмами.

Проблема забезпечення надійного доступу до даних, що містяться на віддалених від користувачів вузлах зберігання інформації спонукала до створення ряду технологій резервування.

Найбільш простою схемою резервування є використання простого дублювання даних на двох носіях. До такого типу відносяться системи InterMemory та RAID-1. Використання простого дублювання пов'язане зі значними затратами об'єму пам'яті. При цьому воно не гарантує відновлення даних при втраті доступу до обох носіїв, на яких зберігаються копії даних.

Значно меншого об'єму пам'яті потребує схема резервування, що передбачає для групи носіїв використання одного контрольного, на якому зберігається сума за модулем 2 відповідних даних всіх носіїв групи. Ця схема дозволяє доволі просто відновити дані при втраті доступу до одного з носіїв групи. Найбільш відомим застосуванням описаної схеми резервування є система RAID-2. Проте ця схема не дозволяє відновлювати дані при втраті доступу до більш як одного носія.

Найбільшого поширення на практиці набули технології відновлення даних на основі корегуючих та відновлюючих (erasure) кодів. При відновленні даних з носіїв, до яких втрачено доступ, як правило, не має потреби в їх локалізації. Класичні корегуючі коди, такі, як коди Хемінга, Ріда-Соломона орієнтовані на послідовне виконання двох процедур: локалізації спотвореної частини даних та їх виправлення. З цієї причини при використанні зазначених вище

корегуючих кодів для відновлення даних з носіїв, до яких втрачено доступ, потрібна їх модифікація. Модифіковані коди Каучі - Ріда-Соломона, зокрема, використовуються в системі відновлення даних з носіїв RAID-6 [3].

Більш ефективно використання для цієї цілі спеціальних відновлюючих кодів. Більшість таких кодів мають за основу лінійні перетворення, і це зумовлює швидке зростання кількості резервних носіїв, при збільшенні числа носіїв до яких втрачено доступ.

Загальною рисою відомих технологій відновлення даних з носіїв, до яких втрачено доступ є те, що вони реалізовані в рамках окремого вузла зберігання інформації. Це означає, що в разі втрати доступу до вузла в результаті тимчасового виходу його з ладу, перевантаження, вірусної атаки, відключення від мережі, техногенних або природних катаклізмів, відомі механізми відновлення даних або доступу до них для конкретного користувача не спрацьовують.

Таким чином, існуючі методи відновлення доступу до даних в системах їх віддаленого зберігання не гарантують вирішення цієї задачі в разі втрати доступу до вузла зберігання інформації.

Доведено також, що існуючі методи резервування даних мають високий рівень надлишковості і що число резервних носіїв може бути зменшено без втрати здатності відновлювати дані.

В роботах [4-7] запропоновано організацію резервування та відновлення даних не більше ніж з трьох носіїв на основі діагональних контрольних сум. Ця організація забезпечує можливість відновлення даних з використанням не більше трьох резервних носіїв, тобто забезпечує теоретичний мінімум інформаційної надлишковості резервування. В роботі [7] запропоновано варіант згаданої організації, який дозволяє відновлювати окремі фрагменти не більше ніж 3-х носіїв, до яких втрачено доступ.

Недоліком організації є те, що вона не забезпечує можливості відновлення даних більш з 3-х носіїв. Аналіз використання систем розподіленого зберігання даних показує, що в ситуаціях, коли носії, на яких зберігаються дані користувача фізично знаходяться в рамках одного вузла, частими є ситуації відмови доступу до більш ніж 3-х носіїв.

Ціллю досліджень є підвищення ефективності резервування та відновлення даних в системах їх віддаленого зберігання за рахунок

зняття обмежень на кількість носіїв, доступ до яких втрачено.

Метод відновлення даних з довільної заданої кількості носіїв

Інформація користувача зберігається на m віддалених носіях. Інформація на кожному з них може розглядатися як така, що складається з n фрагментів. Поняття фрагменту не прив'язано до певної фізичної природи і фрагментом може бути сектор жорсткого диску або блок інформації довжина якого співпадає з розрядністю процесора. Відповідно кожен j -ий фрагмент, $j \in \{1, \dots, n\}$, i -го віддаленого носія може бути позначений як $s_{i,j}$. Тоді інформація користувача, що зберігається на віддалених носіях може бути представлена у вигляді матриці S :

$$S = \begin{pmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,n} \\ s_{2,1} & s_{2,2} & \dots & s_{2,n} \\ \dots & \dots & \dots & \dots \\ s_{m,1} & s_{m,2} & \dots & s_{m,n} \end{pmatrix}$$

В такій матриці кожен стовпець містить одинойменні фрагменти всіх носіїв, а рядок – фрагменти одного носія.

Пропонується метод для відновлення інформації з будь якої непарної кількості віддалених носіїв до яких втрачено доступ. При використанні запропонованого методу розробник такої системи має обрати граничне значення кількості p віддалених носіїв, інформація з яких може бути відновлена. Особливістю методу є те, що кількість резервних носіїв, що використовуються для відновлення даних з p носіїв також становить p . Це значить, що метод забезпечує теоретичний мінімум інформаційної надлишковості резервування. Тоді кожен j -ий фрагмент, $j \in \{1, \dots, n\}$, i -го резервного носія, $i \in \{1, \dots, p\}$, може бути позначений як $d_{i,j}$, а інформація, що зберігається на резервних носіях може бути представлена у вигляді матриці D :

$$D = \begin{pmatrix} d_{1,1} & d_{1,2} & \dots & d_{1,n} \\ d_{2,1} & d_{2,2} & \dots & d_{2,n} \\ \dots & \dots & \dots & \dots \\ d_{p,1} & d_{p,2} & \dots & d_{p,n} \end{pmatrix}$$

Для відновлення даних використовуються p резервних носія, які пропонується розділити на три групи:

- 1) група, яка складається з носіїв з індексами від 1 до $(p-1)/2$ включно;
- 2) група, яка складається з єдиного носія з

індексом $(p+1)/2$;

3) група, яка складається з носіїв з індексами від $(p+3)/2$ до p включно.

Дані для кожної групи резервних носіїв пропонується формувати наступним чином.

Фрагменти для першої групи резервних носіїв пропонується формувати з фрагментів, що належать висхідним діагоналям матриці під різними кутами, а формально як суму за модулем два наступних фрагментів:

$$\forall i \in \{1, \dots, (p-1)/2\}, \forall j \in \{1, \dots, m \cdot i\}: \quad (1)$$

$$d_{(p-1)/2-i+1,j} = \bigoplus_{q=1}^{\lceil j/i \rceil} s_{\lceil j/i \rceil - q + 1, (j-1) \bmod i + 1 + i(q-1)}$$

$$\forall i \in \{1, \dots, (p-1)/2\}, \forall j \in \{m \cdot i + 1, \dots, n\}: \quad (2)$$

$$d_{(p-1)/2-i+1,j} = \bigoplus_{q=1}^m s_{m-q+1, i \cdot q + j - i \cdot m}$$

Фрагменти резервного носія з другої групи пропонується формувати з фрагментів, що належать стовпцям матриці, а формально як суму за модулем два наступних фрагментів:

$$\forall j \in \{1, \dots, n\}: d_{(p+1)/2,j} = \bigoplus_{q=1}^m s_{q,j} \quad (3)$$

Фрагменти для третьої групи резервних носіїв пропонується формувати з фрагментів, що належать низхідним діагоналям матриці під різними кутами, а формально як суму за модулем два наступних фрагментів:

$$\forall i \in \{1, \dots, (p-1)/2\}, \forall j \in \{1, \dots, m \cdot i\}: \quad (4)$$

$$d_{i+(p+1)/2,j} = \bigoplus_{q=1}^{\lceil j/i \rceil} s_{m - \lceil j/i \rceil + q, (j-1) \bmod i + 1 + i(q-1)}$$

$$\forall i \in \{1, \dots, (p-1)/2\}, \forall j \in \{m \cdot i + 1, \dots, n\}: \quad (5)$$

$$d_{i+(p+1)/2,j} = \bigoplus_{q=1}^m s_{q, j - i \cdot m + i \cdot q}$$

Запропонований порядок формування фрагментів резервних носіїв може бути ілюстровано прикладом, який зображено на рис. 1-5. В рамках цього прикладу вважається, що загальна кількість віддалених носіїв дорівнює семи ($m=7$), а граничне значення кількості носіїв до яких втрачено доступ – п'яти ($p=5$).

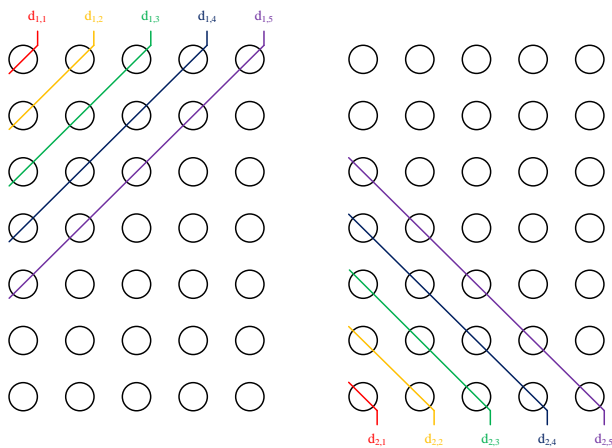


Рис.1 – Формування фрагментів першого резервного носія

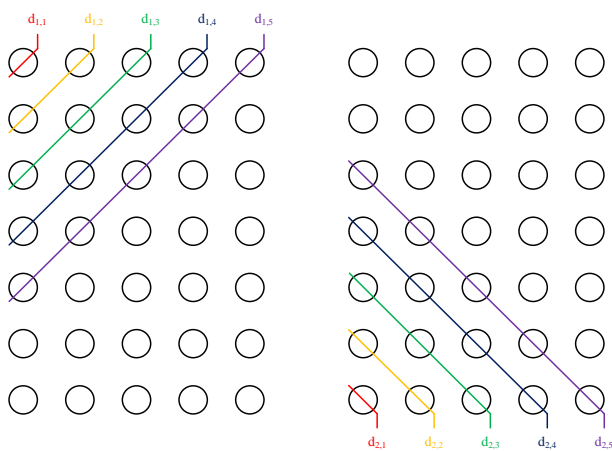


Рис. 2 – Формування фрагментів другого резервного носія

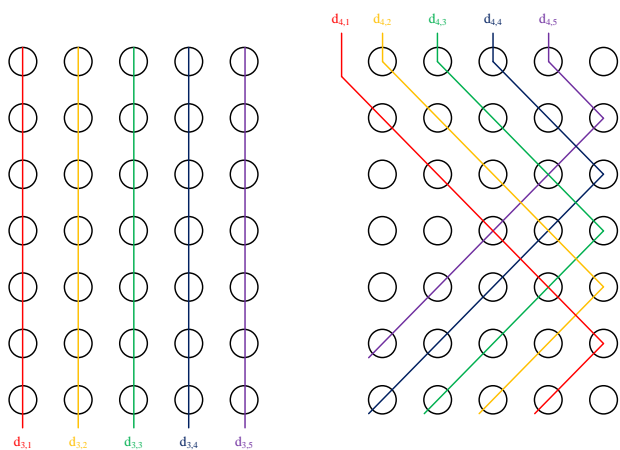


Рис. 3 – Формування фрагментів третього резервного носія

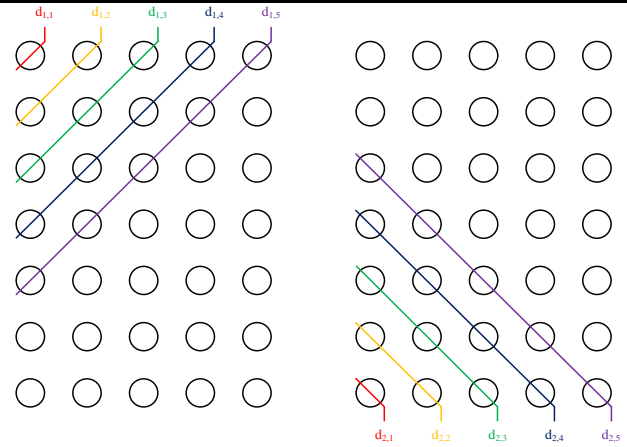


Рис. 4 – Формування фрагментів четвертого резервного носія

Запропонований метод передбачає відновлення майже всіх фрагментів носіїв, до яких втрачено доступ. В найкращому випадку метод надає можливість відновити всі фрагменти, в найгіршому випадку метод не дає змоги відновити останні $(m-p)/2$ однойменних фрагментів носіїв. Зважаючи на те, що кількість фрагментів n на порядки більша за m та p , на цей недолік можна не звертати уваги.

Запропонований метод передбачає впорядковане відновлення фрагментів носіїв.

Індекси носіїв, до яких втрачено доступ, в порядку зростання позначимо як $X = (x_1, x_2, \dots, x_p)$, де $x_1 < x_2 < \dots < x_p$.

Процес відновлення фрагментів складається двох наступних етапів:

1) відновлення першого фрагменту з носія $x_{(p+1)/2}$, до якого втрачено доступ, і всіх потрібних для цього фрагментів з інших недоступних носіїв;

2) Відновлення по одному фрагменту з кожного носія, до якого втрачено доступ доки не будуть відновлені всі фрагменти.

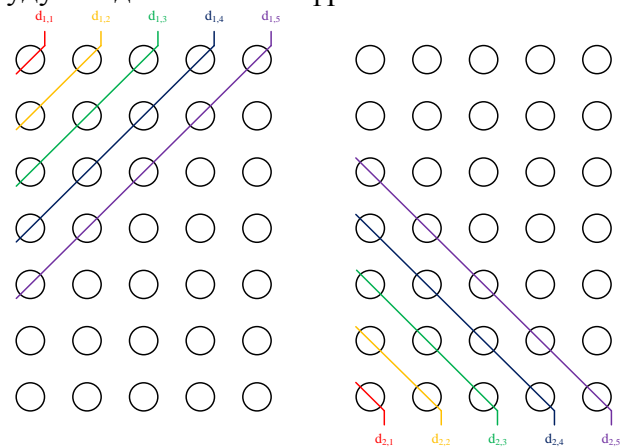


Рис. 5 – Формування фрагментів п'ятого резервного носія

Визначимо кількість фрагментів $B = (b_1 \ b_2 \ \dots \ b_p)$, яку потрібно відновити з кожного недоступного носія для реконструювання першого фрагменту носія $x_{(p+1)/2}$. Для цього визначимо різниці індексів сусідніх носіїв, до яких втрачено доступ. Позначимо різниці індексів сусідніх носіїв як

$$A = (a_1 \ a_2 \ \dots \ a_{p-1}), \text{ де} \\ \forall i \in \{1, \dots, p-1\}: a_i = x_{i+1} - x_i \quad (6)$$

Тоді маємо наступну рівність:

$$\forall i \in \{1, \dots, p\}: \\ b_i = \begin{cases} a_i \cdot ((p-1)/2 - i) + b_{i+1}, & i < (p-1)/2 \\ 1, & (p-1)/2 \leq i \leq (p+3)/2 \\ a_{i-1} \cdot (i - (p+3)/2) + b_{i-1}, & i > (p+3)/2 \end{cases} \quad (7)$$

Введемо поняття останнього відновленого фрагмента кожного недоступного носія $C = (c_1 \ c_2 \ \dots \ c_p)$. Цілком очевидним є факт того, що на початку відновлення:

$$\forall i \in \{1, \dots, p\}: c_i = 0 \quad (8)$$

Алгоритм відновлення для першого етапу складається з наступної послідовності дій:

1. Встановити номер кроку $i = b_1$.
2. Встановити номер відновлюваного носія $j = 1$.
3. Якщо кількість фрагментів, яку потрібно відновити для j -го носія, менша за номер кроку, тобто умова $b_j \geq i$ не виконується перейти до кроку під номером 6.
4. Збільшити індекс останнього відновленого фрагменту j -го носія c_j на одиницю.
5. Відновити фрагмент s_{x_j, c_j} .
6. Збільшити j на 1.
7. Якщо умова $j < (p+1)/2$ виконується перейти до кроку під номером 3.
8. Зменшити i на одиницю.
9. Якщо умова $i \geq 1$ виконується перейти до кроку 2.
10. Встановити номер кроку $i = b_p$.
11. Встановити номер відновлюваного носія $j = p$.
12. Якщо кількість фрагментів, яку потрібно відновити для j -го носія, менша за номер кроку, тобто умова $b_j \geq i$ не виконується перейти до кроку під номером 15.
13. Збільшити індекс останнього відновлено-

го фрагменту j -го носія c_j на одиницю.

14. Відновити фрагмент s_{x_j, c_j} .
15. Зменшити j на 1.
16. Якщо умова $j > (p+1)/2$ виконується перейти до кроку під номером 12.
17. Зменшити i на одиницю.
18. Якщо умова $i \geq 1$ виконується перейти до кроку 11.
19. Збільшити індекс останнього відновленого фрагменту $(p+1)/2$ -го носія $c_{(p+1)/2}$ на одиницю.
20. Відновити фрагмент $s_{x_{(p+1)/2}, c_{(p+1)/2}}$.

Другий етап відновлення складається з p кроків, по одному фрагменту з кожного носія, до яких втрачено доступ. На всіх не парних кроках, окрім останнього, реконструюється фрагмент носія з найменшим порядковим номером за допомогою резервних носіїв першої групи. На всіх парних кроках реконструюється фрагмент носія з найбільшим порядковим номером за допомогою резервних носіїв третьої групи. На останньому кроці реконструюється останній фрагмент за допомогою резервного носія другої групи.

Фрагменти, які відновляються на непарних кроках, окрім останнього, належать висхідним діагоналям, сума яких зберігається в резервних носіях першої групи. Такий фрагмент може бути реконструйований, так як фрагменти, що належать діагоналі і лежать на ній вище мають порядковий номер носія менший за j , тобто вже відновленні на даному кроці або зберігаються на непошкоджених носіях з меншими порядковими номерами. Фрагменти, що лежать на діагоналі нижче мають номери менші за c_j , тобто відновлені на одному з минулих кроків або належать доступним носіям.

Фрагменти, які відновляються на парних кроках, належать низхідним діагоналям, сума яких зберігається в резервних носіях третьої групи. Такий фрагмент може бути реконструйований, оскільки фрагменти, що належать діагоналі і лежать на ній нижче мають порядковий номер носія більший за j , тобто вже відновленні на даному кроці або зберігаються на непошкоджених носіях. Фрагменти, що лежать на діагоналі вище мають номери менші за c_j , тобто відновлені на одному з минулих кроків або належать доступним носіям.

Фрагмент, який відновлюється на останньому кроці належить вертикальній прямій, сума

якої зберігається в резервному носіїв другої групи. Фрагмент може бути реконструйований так, як всі інші однойменні c_j -ті фрагменти вже відновлені на даному або минулих кроках, а інші фрагменти відомі.

Алгоритм відновлення для другого етапу:

1. Встановити номер кроку $i = c_j$.
2. Встановити номер відновлюваного носія $j = 1$.
3. Збільшити індекс останнього відновленого фрагменту j -го носія c_j на одиницю.
4. Відновити фрагмент s_{x_j, c_j} .
5. Збільшити j на 1.
6. Якщо умова $j < (p+1)/2$ виконується перейти до кроку під номером 3.
7. Встановити номер відновлюваного носія $j = p$.
8. Збільшити індекс останнього відновленого фрагменту j -го носія c_j на одиницю.
9. Відновити фрагмент s_{x_j, c_j} .
10. Зменшити j на 1.
11. Якщо умова $j > (p+1)/2$ виконується перейти до кроку під номером 8.
12. Збільшити індекс останнього відновленого фрагменту $(p+1)/2$ -го носія $c_{(p+1)/2}$ на одиницю.
13. Відновити фрагмент $s_{x_{(p+1)/2}, c_{(p+1)/2}}$.

Для відновлення конкретного фрагменту s_{x_j, c_j} віддаленого носія, до якого втрачено доступ, будемо посилатись на групи описані на етапі резервування. Тоді для відновлення фрагменту з індексом резервного носія j , що належить носіям з першої групи достатньо використати одну з наступних формул:

$$\forall j \in \{1, \dots, (p-1)/2\}, \forall c_j \in \{1, \dots, m \cdot i\}:$$

$$i = (p-1)/2 - j + 1$$

$$s_{x_j, c_j} = d_{j, c_j} \oplus \oplus_{\substack{q=1 \\ q \neq x_j}}^{m} s_{\lceil c_j/i \rceil} \quad (9)$$

$$\oplus_{\substack{q=1 \\ q \neq \lceil c_j/i \rceil - x_j + 1}}^{m} s_{\lceil c_j/i \rceil - q + 1, (c_j - 1) \bmod i + 1 + i(q-1)}$$

$$\forall j \in \{1, \dots, (p-1)/2\}, \forall c_j \in \{m \cdot i + 1, \dots, n\}:$$

$$i = (p-1)/2 - j + 1$$

$$s_{x_j, c_j} = d_{j, c_j} \oplus \oplus_{\substack{q=1 \\ q \neq m - x_j + 1}}^m s_{m - q + 1, i \cdot q + c_j - i \cdot m} \quad (10)$$

Для відновлення фрагменту з індексом резервного носія j , що належить носію з другої групи $j = (p+1)/2, \forall c_j \in \{1, \dots, n\}$:

$$s_{x_j, c_j} = d_{(p+1)/2, c_j} \oplus \oplus_{\substack{q=1 \\ q \neq x_j}}^m s_{q, c_j} \quad (11)$$

Для відновлення фрагменту з індексом резервного носія j , що належить носіям з третьої групи достатньо використати одну з наступних формул:

$$\forall j \in \{(p+3)/2, \dots, p\}, \forall c_j \in \{1, \dots, m \cdot i\}:$$

$$i = j - (p+1)/2$$

$$s_{x_j, c_j} = d_{j, c_j} \oplus \oplus_{\substack{q=1 \\ q \neq x_j}}^{m} s_{\lceil c_j/i \rceil} \quad (12)$$

$$\oplus_{\substack{q=1 \\ q \neq x_j - m + \lceil c_j/i \rceil}}^{m} s_{m - \lceil c_j/i \rceil + q, (c_j - 1) \bmod i + 1 + i(q-1)}$$

$$\forall j \in \{(p+3)/2, \dots, p\}, \forall c_j \in \{m \cdot i + 1, \dots, n\}:$$

$$i = j - (p+1)/2$$

$$s_{x_j, c_j} = d_{j, c_j} \oplus \oplus_{\substack{q=1 \\ q \neq x_j}}^m s_{q, c_j - i \cdot m + i \cdot q} \quad (13)$$

Зважаючи на алгоритм відновлення можна зробити висновок, що запропонований метод відновлює не всі фрагменти носіїв, до яких втрачено доступ. В середньому кількість не відновлювальних фрагментів дорівнює $\max\{c_l, c_p\}$ після першого етапу відновлення.

Оцінка ефективності

Зважаючи на те, що кожен віддалений носій складається більш як з тисячі фрагментів, при підрахунку обчислювальної складності процедури резервування даних можна знехтувати формулами (1) та (4) при не великих значеннях p , та вважати, що всі фрагменти резервних носіїв першої групи формуються за формулою (2), а третьої – за формулою (5). Тоді очевидно, що обчислювальна складність резервування даних складає $O(p \cdot m \cdot n)$. Звертаючи увагу на той самий факт, при підрахунку обчислювальної складності відновлення даних можна знехтувати алгоритмом відновлення першого етапу. Та вважати, що всі n однойменних фрагментів відновлюються за алгоритмом другого етапу. Тоді очевидним є факт того, що обчислювальна складність відновлення даних складає $O(p \cdot m \cdot n)$.

Таким чином, показано, що обчислювальна складність процесів формування резервного коду та відновлення будь-яких p з m носіїв, до

яких втрачено доступ, лінійно залежно від основних параметрів – кількості носіїв, кількості фрагментів та кількості носіїв, дані з яких відновлюються.

Основною перевагою запропонованого методу є те, що він не накладає обмежень на кількість носіїв, дані з яких можуть бути відновлені на відміну від відомих методів [4-7].

Висновки

В результаті проведених досліджень запропоновано метод резервування та відновлення даних при їх зберіганні на віддалених від користувача носіях на основі діагональних контрольних сум. Основною відмінністю запропонованого методу від відомих є те, що він не на-

кладає обмежень на кількість носіїв, дані з яких можуть бути відновлені.

Для досягнення можливості відновлення даних з довільної кількості носіїв, доступ до яких втрачено, запропоновано використання контрольних сум фрагментів, що належать діагоналям, розташованими під різними кутами. Оскільки для відновлення даних з довільних p носіїв використано таку ж кількість резервних носіїв, то метод забезпечує теоретичний мінімум інформаційної надлишковості резервування даних.

Розроблений метод орієнтовано для використання в системах з великою кількістю користувачів та віддалених від них рознесених вузлах зберігання інформації.

Список посилань

1. Patterson D. Case of Redundant Array of Inexpensive Disk (RAID) / Patterson D., Gibson G., Katz R.- Berkeley: University of California. – 1987. – 238 p.
2. Plank J. S., Thomasson M.G., On Practical Use of LDPC Erasure Codes for Distribute Storage Application: Technical Report UT-CS-03-510. – Department of Computer Science, University of Tennessee. – 2004.
3. Berger C.R. Optimizing Joint Erasure – and Error-Correction Coding for Wireless Packet Transmissions // Willett P., Pattirati K. / IEEE Transactions on wireless communications. – Vol.7 – No. 11. – 2008. – Pp. 4586-4595.
4. Иванов Д. Г. Организация резервирования в системах распределенного хранения данных // Вісник Національного технічного університету України "КПІ" Інформатика, управління та обчислювальна техніка, – Київ: ВЕК+ – 2012 – № 56. – С.160-164.
5. Иванов Д.Г. Метод восстановления данных на основе скошенных матриц в системах распределенного хранения / Иванов Д.Г. Луцкий Г.М. // Известия высших учебных заведений. Проблемы полиграфии и издательского дела. Информационные технологии. – М.: УПИПК МГУП им. И.Федорова. – 2013. – № 2. – С.47-52
6. Марковський О.П. Організація резервування та відновлення даних при їх віддаленому зберіганні / Марковський О.П., Иванов Д.Г., Ванчугов Б.Ю. // Вісник Національного технічного університету України "КПІ" Інформатика, управління та обчислювальна техніка, – Київ: ВЕК+ – 2013. – № 59. – С. 50-55.
7. Марковський О.П. Метод резервування та прискореного відновлення даних в системах їх віддаленого зберігання / Марковський О.П., Иванов Д. Г., Великий М. М., Невдащенко М. В. // Вісник Національного технічного університету України "КПІ" Інформатика, управління та обчислювальна техніка, – Київ: ВЕК+ – 2014 – № 60. – С.46-54.