

УДК 519.816

DOI: 10.15587/2312-8372.2019.182622

РОЗРОБКА МЕТОДИКИ ОЦІНКИ НАДІЙНОСТІ БАЗ ДАНИХ, ЯКІ ВИКОРИСТОВУЮТЬСЯ В УПРАВЛІННІ РАДІОМЕРЕЖАМИ

Мищенко В. О.

РАЗРАБОТКА МЕТОДИКИ ОЦЕНКИ НАДЕЖНОСТИ БАЗ ДАННЫХ, ПРИМЕНЯЕМЫХ В УПРАВЛЕНИИ РАДИОСЕТЯМИ

Мищенко В. А.

DEVELOPMENT OF A METHODOLOGY FOR ASSESSING RELIABILITY OF DATABASES USED IN RADIO NETWORK MANAGEMENT

Mishchenko V.

У роботі проаналізовано задачі, які виконуються під час обробки розподілених даних. Об'єктом дослідження є процес автоматизації управління у розподілених системах та організація обробки розподілених даних. У ході дослідження використовувався метод реплікації даних, який передбачає їх дублювання в різних вузлах інформаційно-аналітичної системи. Це пов'язано з тим, що зазначений метод дозволяє привести декілька баз даних з однаковою структурою в один і той же несуперечливий стан, який супроводжується взаємним внесенням змін. Завдяки цьому забезпечується ефективне інформаційне та інтелектуальне (аналітичне) забезпечення територіально розподілених, але функціонально взаємозалежних груп посадових осіб в процесі вироблення ними управлінських рішень. Аналіз етапів елементарного циклу управління надскладними системами – збір даних обстановки, вироблення рішення, генерація управляючих команд компонент мережі – виявляє загальну для усіх етапів особливість. А саме – усі вони є складними інформаційними технологічними процесами (перші три – обмін інформацією в системі передачі даних та її обробка в системах управління даними, останній – зміна параметрів мережі генерації управляючих команд компонент мережі). Висока швидкість протікання зазначених процесів і логічна складність їх структури призвели до необхідності автоматизації управління даними процесами, коли значна частина інформаційних функцій центру управління передається технічними засобами – інформаційно-логічними пристроями автоматизованої системи управління. У результаті проведення дослідження отримані експериментальні результати, на підставі яких можливо значно підвищити рівень функціональної надійності баз даних інформаційно-аналітичних систем. Також стало можливим зменшити час на обробку та синхронізацію розподілених даних інформаційно-аналітичних систем. Результати дослідження мають використовуватися під час розрахунків функціональної

надійності розподілених баз даних інформаційно-аналітичних систем, інформаційних та телекомунікаційних мереж. Планується використовувати зазначену методику під час побудови великих масивів розподілених баз даних у інформаційних мережах, побудованих на сучасних безпроводових радіотехнологіях.

Ключові слова: інформаційна мережа, технології реплікації даних, система управління базами даних, обробка розподілених даних.

В работе проанализированы задачи, которые выполняются во время обработки распределенных данных. Объектом исследования является процесс автоматизации управления в распределенных системах и организация обработки распределенных данных. В ходе исследования использовался метод репликации данных, предусматривающий их дублирование в разных узлах информационно-аналитической системы. Это связано с тем, что указанный метод позволяет привести несколько баз данных с одинаковой структурой в одно и то же непротиворечивое состояние, которое сопровождается взаимным внесением изменений. Благодаря этому обеспечивается эффективное информационное и интеллектуальное (аналитическое) обеспечение территориально распределенных, но функционально взаимосвязанных групп должностных лиц в процессе выработки ими управленческих решений. Анализ этапов элементарного цикла управления сложнейшими системами – сбор данных обстановки, выработка решения, генерация управляющих команд компонент сети – обнаруживает общую для всех этапов особенность. А именно – все они являются сложными информационными процессами (первые три – обмен информацией в системе передачи данных и ее обработка в системах управления данными, последний – изменение параметров сети генерации управляющих команд компонент сети). Высокая скорость протекания указанных процессов и логическая сложность их структуры привели к необходимости автоматизации управления данными процессами, когда значительная часть информационных функций центра управления передается техническими средствами – информационно-логическими устройствами автоматизированной системы управления. В результате проведения исследования получены экспериментальные результаты, на основании которых возможно значительно повысить уровень функциональной надежности баз данных информационно-аналитических систем. Также стало возможным уменьшить время на обработку и синхронизацию распределенных данных информационно-аналитических систем. Результаты исследования должны использоваться при расчетах функциональной надежности распределенных баз данных информационно-аналитических систем, информационных и телекоммуникационных сетей. Планируется использовать указанную методику при построении больших массивов распределенных баз данных в информационных сетях, построенных на современных беспроводных радиотехнологиях.

Ключевые слова: информационная сеть, технологии репликации данных, система управления базами данных, обработка распределенных данных.

1. Вступ

В даний час інформаційна інфраструктура сучасних підприємств і організацій має все більш розподілений характер. Рівень прийнятих управлінських рішень, контроль та управління інформаційними ресурсами вимагає все більшої децентралізації [1–3]. Структура сучасних організацій являє собою розгалужену схему, до складу якої входять десятки територіально розділених підрозділів, функціонування яких один без одного, часто, взагалі не можливо. При цьому якість функціонування такої складної, корпоративної системи багато в чому залежить від застосування в її діяльності новітніх інформаційних технологій.

Основною проблемою під час автоматизації розподілених систем є організація обробки розподілених даних. Застосування технології реплікації даних [4–6], є важливим етапом в успішній роботі організації. Впровадження зазначеної технології забезпечує гарантовану доставку, своєчасність і цілісність переданих даних.

На даний час актуальним є розробка нових підходів до оцінки надійності баз даних. Оскільки цього вимагає:

- стрімке впровадження автоматизованих систем управління на початку ХХ століття [7–9];

- поява, розвиток та розповсюдження складних систем з такими компонентами, як технічні, програмні, комунікаційні засоби та люди [10];

- засоби боротьби у кібернетичному просторі [11].

Таким чином, *об'єктом дослідження* є процес автоматизації управління у розподілених системах та організація обробки розподілених даних. *А мета роботи* полягає у розробці методики підвищення функціональної надійності баз даних інформаційно-аналітичних систем методом реплікації даних, який передбачає їх дублювання в різних вузлах інформаційно-аналітичної системи.

2. Методика проведення дослідження

Під час проведення досліджень використовувалися такі наукові методи:

- метод наукового аналізу;

- метод узагальнення;

- метод порівняння;

- метод моделювання;

- метод прогнозування.

Основною метою застосування зазначених методів у дослідженні є забезпечення можливості обробки даних будь-якої структури та складності.

Зазначений підхід дозволяє обробляти та аналізувати дані без прив'язки до їхнього подання та територіального розташування. Оперативного виведення даних для вироблення управлінських рішень відповідними посадовими особами.

Задачею прогнозування є знаходження залежності між датою та параметрами ресурсів, отриманих з бази даних вказаної структури, на основі аналізу попередньо отриманих даних. Динаміка зміни виробленої потужності може бути охарактеризована стосовно якогось базисного (зазвичай першого) спостереження і

величиною зміни сусідніх рівнів. У якості статистичних характеристик часового ряду $Y_i, i=1...n$ використовується середнеарифметичне Y :

$$Y = \frac{Y_n - Y_1}{N - 1},$$

де N – число рівнів ряду; Y_i – рівні ряду [12].

Відповідно до методу перевірки істотності різниці середніх показників, початковий часовий ряд розкладається на дві однакові (або майже однакові) частини, після чого перевіряється гіпотеза про істотність різниці середніх показників для зазначених частин. Недолік методу полягає в неможливості правильно визначити наявність тренда у тому випадку, коли часовий ряд містить точку зміни тенденції в районі середини ряду [12].

Гіпотеза про відсутність тренда перевіряється за допомогою допоміжних функцій:

$$L = \sum_{t=2}^N I_t, I_t = U_t - V_t, u_t = \begin{cases} 1, & < Y_{t-1} \\ Y & 0 \end{cases}, \dots, v_t = \begin{cases} 1, & > Y_{t-1} \\ Y & 0 \end{cases}, \dots$$

Гіпотеза про відсутність тенденції не приймається, якщо розрахункове t -значення більше табличного на вибраному рівні значущості 0,95.

Перевірка однорідності даних здійснюється на основі критерію Ірвіна, який заснован на порівнянні сусідніх значень ряду. Відповідно до нього розраховується характеристика t :

$$t = \frac{Y_t - Y_{t-1}}{Y}.$$

Оцінка властивостей зводиться до дослідження автокореляційної функції вихідного і різницевого рядів. Аналіз автокореляції виконується за допомогою графіка і критичних значень коефіцієнтів, встановлених експертно.

3. Результати дослідження та обговорення

Оскільки під інформаційно-аналітичною системою слід розуміти взаємопов'язану сукупність інформаційних ресурсів, людського інтелекту, сучасних засобів інформатизації та інформаційно-телекомунікаційних технологій, зазначені ресурси об'єднані у телекомунікаційні та комп'ютерні мережі. А також засновані на поглибленому аналізі, широкому застосуванні інформаційно-аналітичних систем підтримки рішень, електронному документообізі та діловодстві та призначені об'єднати передові інформаційні технології [13].

Відповідно під функціональною надійністю інформаційно-аналітичних систем розуміється їх спроможність забезпечити визначений рівень працездатності при виконанні визначених їм функцій. Ефективне інформаційне та інтелектуальне (аналітичне) забезпечення територіально розподілених, але

функціонально взаємозалежних груп посадових осіб в процесі вироблення ними управлінських рішень.

Реплікація даних передбачає їх дублювання в різних вузлах інформаційно-аналітичної системи. При цьому будь-яка база даних (БД) є локальною (як для системи управління базами даних (СУБД), так і для користувача, що працює з нею). Тобто дані завжди розташовуються локально на визначеному вузлі мережі, де вони обробляються, і всі транзакції в системі відпрацьовуються та завершуються локально. Ефективність технології реплікації даних в розподіленій системі залежить від того, які операції мають виконуватися в системі.

Будь-яка інформаційно-аналітична система (ІАС) S_{IAC} описується відношенням типу:

$$S_{IAC} \leq T + D + P + U, \quad (1)$$

де T – множина технічних засобів інформаційно-аналітичної системи;

D – множина баз даних системи;

P – множина програмних засобів системи;

U – множина користувачів системи.

БД є найважливішим компонентом ІАС і від її надійного функціонування залежить ефективне виконання практично всіх функцій системи.

Застосування реплікації БД дозволяє привести декілька БД з однаковою структурою в один і той же несуперечливий стан, який супроводжується взаємним внесенням змін.

В [14, 15] реплікацію визначають як процес генерації та відтворення декількох копій даних, розташовуваних на одному або декількох сайтах. В [1] під реплікацією даних СУБД розуміють приведення баз даних, які функціонують у розподіленому середовищі в актуальний стан, за рахунок виявлення змінених даних, передавання і застосування цих змін у базі даних отримувача.

Найпростішу модель для оцінки надійності роботи реплікованої бази даних ІАС можна представити як відновлювану систему з обмеженим обсягом запасних частин (у наведеному випадку – реплік БД).

Розглянемо систему БД із «гарячим» резервуванням, яка піддається пуасонівському потоку відмов з інтенсивністю λ .

У зазначеній системі є n локальних копій, тобто реплік БД, а також файлів внесених змін (локальних журналів транзакцій).

У найпростішому випадку відновлення реплік здійснюється за допомогою еталонної версії БД, а їхня подальша актуалізація генерується на основі відповідного електронного локального журналу змін.

Вважатимемо, що витрати часу на відновлення деякої репліки БД за допомогою еталонної версії набагато вищі розглянутого інтервалу функціонування. Оскільки вимагається виконати пересилання достатньо великих обсягів даних (повної копії еталонної БД), а також зробити їхню актуалізацію.

У цьому випадку надійність БД визначається ймовірністю того, що наявного числа її реплік вистачить для забезпечення сталої роботи системи на

інтервалі часу $(0,t)$. Оскільки, при відмові будь-якої репліки БД, запити будуть переадресовані до однієї з реплік, які залишились в робочому стані.

Надійність бази даних ІАС доцільно оцінювати як ймовірність того, що на інтервалі $(0,t)$ число відмов реплік БД буде не більше $(n-1)$, що обчислюється за формулою:

$$P_{\text{БД}}(t) = \sum_{k=0}^{n-1} P_k(t) = e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}. \quad (2)$$

Зазначену величину можемо прийняти за оцінку ймовірності безвідмовної роботи системи з «гарячим» резервуванням, яка містить n реплікованих БД ІАС, на момент часу t .

Позначивши через $P_{\text{БД}}^{\text{вим}}$ рівень надійності БД, що вимагається у складі ІАС, можна визначити кількість $n^{\text{вим}}$ реплік БД, що вимагається для її роботи з рівнем надійності $P_{\text{БД}}^{\text{вим}}$ за час t , з урахуванням виразу (2) за формулою:

$$n^{\text{вим}} = \arg \min_{n=1,2,\dots} \left\{ \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} \geq P_{\text{БД}}^{\text{вим}} e^{\lambda t} \right\}. \quad (3)$$

Розглянемо набагато складну, але досить реальну ситуацію, коли в територіально-розподілених вузлах обробки даних ІАС здійснюється локальне резервне копіювання станів реплік БД через деякі апріорно задані інтервали часу.

В цей час передбачається, що послідовність зняття резервних копій (тобто контрольних точок) у цілому для всієї системи синхронізована.

У зазначеному випадку кожна з реплік БД відновлюється зі своєї останньої локальної резервної копії (РК), а далі здійснюється її актуалізація відповідно до локального журналу змін.

Подальша побудова моделей, призначених для оцінки характеристик надійності розглянутої системи, буде здійснюватись на основі формалізації кінцевих ланцюгів Маркова [16].

Розглянемо наступні стани системи:

S_0 – у системі всі репліки БД функціонують та доступні;

S_1 – одна з реплік перебуває в стані відмови (непрацездатна), здійснюється її відновлення за допомогою відповідної локальної РК;

S_2 – дві репліки в стані відмови;

S_n – усі наявні репліки БД ІАС перебувають у стані відмови, здійснюється їхнє відновлення. Саме в цьому випадку можна говорити про відмову всієї БД ІАС, оскільки жодна з реплік не буде працездатною та доступною для роботи користувачів.

Як і раніше, будемо вважати, що система піддається пуасонівському потоку відмов з інтенсивністю λ , а відновлення кожної з реплік здійснюється з інтенсивністю μ .

Зазначений процес переходів станів складної системи відомий як марківський ланцюг «загибелі-розмноження» [17].

Також розглянемо приведену інтенсивність потоку відмов, яка визначається:

$$\rho = \frac{\lambda}{\mu}.$$

Базову модель розрахунку надійності БД ІАС для графа, визначимо через так звані кінцеві ймовірності станів системи за формулою Ерланга [18], враховуючи при цьому, що $\mu_1 = \mu, \mu_2 = 2\mu, \mu_3 = 3\mu, \dots, \mu_n = n\mu$:

$$\begin{aligned} p_0 &= \left(1 + \rho + \frac{\rho^2}{2!} + \frac{\rho^3}{3!} + \dots + \frac{\rho^n}{n!}\right)^{-1}; \\ p_1 &= \rho p_0; \\ p_2 &= \frac{\rho^2}{2!} p_0; \\ p_3 &= \frac{\rho^3}{3!} p_0; \\ &\dots; \\ p_n &= \frac{\rho^n}{n!} p_0; \\ \sum_{i=0}^n p_i &= 1. \end{aligned} \tag{4}$$

Ймовірність безвідмовної роботи розглянутого алгоритма функціонування бази даних ІАС обчислюється за допомогою системи рівнянь (4), яка описує кінцеві ймовірності всіх можливих станів системи в рамках даної моделі:

$$Q_{\text{БД}} = 1 - p_n = 1 - \frac{\rho^n}{n!} p_0. \tag{5}$$

Таким чином, зазначена методика дозволяє:

- привести декілька баз даних з однаковою структурою в один і той же несуперечливий стан, який супроводжується взаємним внесенням змін;
- одержати досить адекватні оцінки таких важливих показників надійності бази даних ІАС;
- забезпечити високу ймовірність безвідмовної роботи та ймовірність роботи із продуктивністю не нижче заданої розподілених баз даних ІАС.

4. Висновки

Під час проведення дослідження отримані експериментальні результати, на підставі яких можливо значно підвищити рівень функціональної надійності баз даних ІАС. Також стало можливим зменшити час на обробку та синхронізацію розподілених даних ІАС.

Результати дослідження стануть у нагоді під час розрахунків функціональної надійності розподілених баз даних ІАС, інформаційних та телекомунікаційних мереж.

Запропоновано використовувати зазначену методику під час побудови великих масивів розподілених баз даних у інформаційних мережах, побудованих на сучасних безпроводових радіотехнологіях.

Література

1. Date, C. J. (2003). *An Introduction to Database Systems*. Pearson, 1040.
2. Третяк, В. Ф., Голубничий, Д. Ю., Челенко, Ю. В. (2005). Тиражирование данных в системе управления базами данных. *Управління розвитком*, 3, 94–95.
3. Третяк, В. Ф., Приходько, В. М., Голубничий, Д. Ю. (2005). Актуальність побудови grid систем. *Перша науково-технічна конференція Харківського університету Повітряних Сил*. Харків: ХУ ПС, 240–241.
4. Третяк, В. Ф. (2004). Технология репликации в распределенных системах управления базами данных. *Інформаційно-керуючі системи на залізничному транспорті*, 1, 7–10.
5. Третяк, В. Ф., Голубничий, Д. Ю., Золотарьова, І. О. (2007). Проблеми розвитку програмного забезпечення середовища розподілених обчислень GRID. *Матеріали третьої наукової конференції Харківського університету Повітряних Сил імені Івана Кожедуба*. Харків: ХУПС, 7.
6. Голубятников, И. В., Матвеев, Ю. В., Сергеев, И. В. (2001). Создание программного комплекса репликации СУБД-независимых данных. *Системный анализ, информатика и оптимизация*. Москва: РЗИТЛП, 106–116.
7. Ляшенко, І. О. (2009). Мережоцентризм у військовій справі. *Сучасні інформаційні технології у сфері безпеки та оборони*, 2 (5), 78–81.
8. Ляшенко, І. О. (2009). Еволюція розвитку концепцій ведення збройної боротьби. *Сучасні інформаційні технології у сфері безпеки та оборони*, 3 (6), 23–25.
9. Ляшенко, І. О. (2010). Застосування перспективних інформаційних технологій у збройній боротьбі. *Труди університету*, 1 (94), 144–150.
10. Ляшенко, І. О. (2010). Кібернетичний простір – як еволюція розвитку інформаційних технологій. *Сучасні інформаційні технології у сфері безпеки та оборони*, 7, 18–21.
11. Ляшенко, І. О., Кириленко, В. А. (2010). Кібернетичні операції – майбутня форма збройної боротьби. *Збірник наукових праць. Серія: військові та технічні науки*, 53, 162.
12. Болюбаш, Ю. Я. (2017). *Методи та засоби опрацювання інформаційних ресурсів великих даних в системах територіального управління*. Львів, 178.

13. Бройдо, В. Л. (2003). *Вычислительные системы, сет и телекоммуникации*. Санкт-Петербург: Питер, 688.
14. Ciciani, B., Dias, D. M., Yu, P. S. (1992). Analysis of Concurrency-Coherency Control Protocols for Distributed Transaction Processing Systems with Regional Locality. *IEEE Transactions on Software Engineering*, 18 (10), 899–914. doi: <http://doi.org/10.1109/32.163606>
15. Connolly, T., Begg, C. (2000). *Database Systems: A Practical Approach to Design, Implementation, and Management*. Addison Wesley, 1236.
16. Kemeny, J. G., Snell, J. L. (1983). *Finite Markov Chains: With a New Appendix «Generalization of a Fundamental Matrix» (Undergraduate Texts in Mathematics)*. Springer, 226.
17. Wentzel, E. S. (1983). *Operations Research*. Mir Publishers, 256.
18. Бейз, Г., Кук, Д. (1990). *Компьютерная математика*. Москва: Наука, 383.

The paper analyzes the tasks that are performed during the processing of distributed data. The object of research is the process of automation of control in distributed systems and the organization of processing distributed data. During the study, a data replication method was used, providing for their duplication in different nodes of the information-analytical system. This is due to the fact that this method allows to bring several databases with the same structure into the same consistent state, which is accompanied by a mutual introduction of changes. This ensures effective informational and intellectual (analytical) support for geographically distributed, but functionally interconnected groups of officials in the process of developing managerial decisions. The analysis of the stages of the elementary cycle of managing complex systems – the collection of situation data, the development of solutions, the generation of control commands for the network component – reveals a common feature for all stages. Namely, all of them are complex information processes (the first three are the exchange of information in the data transmission system and its processing in data management systems, the last is the change in the network parameters for generating control commands of the network component). The high speed of these processes and the logical complexity of their structure led to the need to automate the management of these processes, when a significant part of the information functions of the control center is transferred by technical means – information and logic devices of an automated control system. As a result of the study, experimental results are obtained, on the basis of which it is possible to significantly increase the level of functional reliability of databases of information-analytical systems. It also became possible to reduce the time for processing and synchronization of distributed data of information-analytical systems. The research results should be used in calculating the functional reliability of distributed databases of information-analytical systems, information and telecommunication networks. It is planned to use this methodology when building large arrays of distributed databases in information networks built on modern wireless radio technologies.

Keywords: *information network, data replication technology, database management system, distributed data processing.*