

УДК 004.658.2 + 051

ТЕСТУВАННЯ БАЗ ДАНИХ КОРПОРАТИВНИХ ІНФОРМАЦІЙНИХ СИСТЕМ*Левченко О.Ю.**Одеський національний політехнічний університет*

У статті показано, що для підвищення продуктивності КІС методом налаштування параметрів конфігурації СУБД необхідно виконувати навантажне тестування. Представлено імітаційний підхід до побудови моделі тестування БД. Розписана методика побудови тесту та тестування БД із застосуванням імітаційної моделі. Для перевірки достовірності отриманих результатів тестування запропоновано використовувати довірчий інтервал та довірку ймовірність.

Вступ

З метою підвищити продуктивність корпоративних інформаційних систем (КІС) методом налаштування параметрів конфігурації систем керування базами даних (СКБД), що обслуговують ці КІС, необхідно досліджувати вплив різних конфігурацій параметрів на продуктивність БД досліджуваної КІС [1]. Навантажне тестування СКБД дозволяє фіксувати показники продуктивності із заданим робочим навантаженням в різних діапазонах зміни конфігураційних параметрів СКБД. Аналіз результатів тестування дозволить підібрати такі значення параметрів конфігурації СКБД, що забезпечать підвищення продуктивності БД КІС в заданих реальних умовах функціонування системи.

1 Імітаційна модель тестування БД

Одним з методів побудови моделі тестування БД КІС є імітаційне моделювання [2]. Модель описує реальну БД, включаючи структуру таблиць, дані, запити користувачів з урахуванням послідовності і частоти їх виникнення.

Методика тестування СКБД включає 3 етапи:

1. Побудова імітаційної моделі тестування (ІМТ) БД, що є репрезентативною відносно реальної БД досліджуваної КІС.
2. Моделювання із використанням ІМТ та заданим набором параметрів, збір статистики і аналіз отриманих показників продуктивності.
3. Формування рекомендованих наборів параметрів конфігурації СКБД і їх значень P_{opt} для досліджуваного класу КІС.

Етап формування моделі включає збір інформації про стан БД на момент часу P_0 (структура таблиць, зв'язки між ними і наповнення таблиць). При цьому, з метою скорочення ресурсів в процесі тестування, таблиці з однаковою структурою об'єднуються під новими іменами. Потім виконується збір інформації щодо здійснених запитів $Q = \{q_1, q_2, \dots, q_n\}$ у досліджуваній період часу. Синтаксичний розбір (парсинг) журналу транзакцій дозволяє сформулювати шаблони запитів та класифікувати їх, що дозволить абстрагуватися від смислового значення імен таблиць та поєднати запити з однаковою структурою реляційних операцій. Класифікація включає віднесення шаблонів запитів з безлічі Q до одного з класів, що описуються еталонами шаблонів. На основі зібраної інформації про структуру таблиць і їх наповнення створюється ІМТ БД.

2 Моделювання із використанням ІМТ БД

На етапі моделювання досліджується вплив різних значень параметрів конфігурації СКБД з множини:

$$P = \langle \{p_1, v_1\} \{p_2, v_2\} \dots \{p_n, v_n\} \rangle, \quad (1)$$

де $p_i, i = 1, n$ – найменування параметру конфігурації, v_i – поточне значення параметру p_i . За початковим

набором параметрів конфігурації P_0 фіксуються характеристики загальної продуктивності БД: для усієї множини запитів Q – показник H^{P_0} , для окремих запитів – $H_{qi}^{P_0}$. Потім виконується їх порівняння з новими характеристиками H^{P_i} та $H_{qi}^{P_i}$, що було отримано у ході моделювання із набором параметрів P_i . Моделювання виконується доки $H_{qi}^{P_i} > H_{qi}^{P_{i-1}}$. Щоб виключити ситуацію, коли ефект від застосування найкращого набору параметрів $H^{P_{opt}}$ (отриманого в ході ітерацій моделювання) може статися меншим, ніж при заміні значення окремого параметру, виконується порівняння показників продуктивності: сумарного $H^{P_{opt}}$ та за окремими параметрами H^{P_i} . Отримані таким чином погоджені значення параметрів конфігурації зберігаються у базі знань у вигляді рекомендацій щодо зміни значень параметрів для КІС заданого класу. Алгоритм тестування наведено на рис. 1.

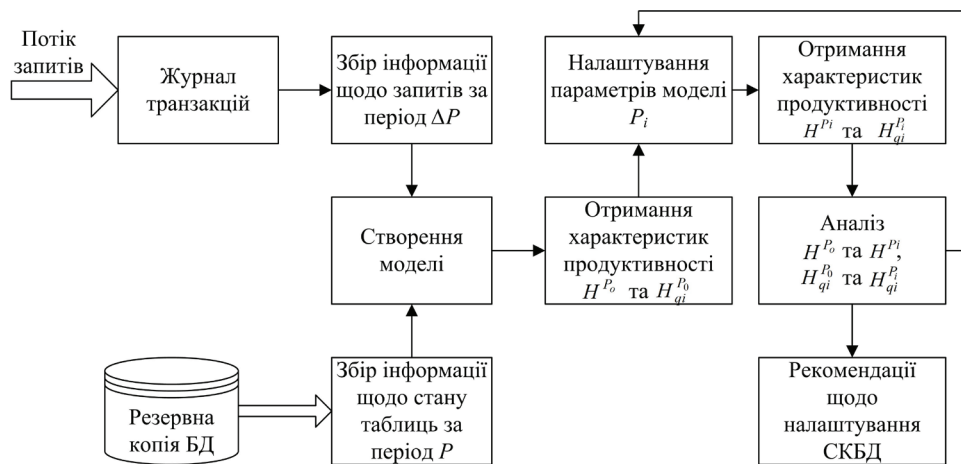


Рисунок 1. Алгоритм навантаженого тестування БД

3 Статистична обробка результатів тестування

Характеристики, що отримано в процесі тестування, необхідно оцінювати з погляду достовірності. Важливим також є питання щодо кількості експериментів, що виконуються: їх недостатня кількість призводить до зниження достовірності отриманих характеристик, тоді як збільшення кількості експериментів вимагає більшої кількості ресурсів, що призводить до зниження продуктивності. Для оцінки точності й надійності отриманої характеристики пропонується використовувати інструментарій математичної статистики: довірчий інтервал та довірчу вірогідність. У такому разі достовірною вважається така характеристика, яка потрапляє в інтервал l_β заздалегідь розрахований виходячи з довірчої вірогідності β , що задається дослідником:

$$l_\beta = (m - t_\beta \sigma_m; m + t_\beta \sigma_m), \quad (2)$$

де m – математичне очікування, σ – середньоквадратичне відхилення оцінки m ; t_β – (t - критерій) статистичний показник, визначається значенням довірчої вірогідності β та числом ступенів свободи за таблицями розподілів Стюдента.

4 Висновок

Запропонована методика дозволяє проаналізувати міру впливу кожного окремого параметру конфігурації БД на загальну продуктивність КІС. Аналіз впливу на продуктивність окремих груп запитів (шаблонів) дозволить додатково зробити висновок про необхідність подальшого їх використання в імітаційній моделі БД.

Для оцінки достовірності отриманих характеристик пропонується розраховувати довірчий інтервал по заданій довірчій вірогідності.

Для реалізації запропонованої методики розроблено комплекс програмних засобів, що реалізовує алгоритми тестування параметрів налаштування СКБД, побудовані із застосуванням

наведеної методики тестування. Розроблений модуль тестування дозволяє змінювати значення вибраних параметрів конфігурації у заданому діапазоні із заданим кроком та виконувати тестування для обраного робочого навантаження (кількість клієнтів, що симулюються, коефіцієнт масштабування, кількість транзакцій та ін.). В ході експерименту з використанням цього комплексу було протестоване 14 параметрів налаштування СКБД PostgreSQL. Було отримано рекомендації щодо налаштування параметрів конфігурації СКБД, що обслуговує КІС деканату студентів.

Література

- [1] Блажко А.А. Модели для автоматизированной оптимизации производительности СУБД / А.А. Блажко, А.Ю. Левченко, А.С. Пригожев // Радиоэлектронні і комп'ютерні системи. – ХАІ, Харків, 2010. – №7(48). – С. 24-29.
- [2] Зиноватная С.Л., Задачи имитационного моделирования реляционных баз данных/ С.Л. Зиноватная, А.Б. Кунгурцев, А.А. Мунзер // Труды XI международной научно-практической конференции «Сучасні інформаційні та електронні технології» (СИЭТ-2010). – Одеса, 2010. – Том1. – С. 118.