



## ДОСЛІДЖЕННЯ ПІДХОДІВ ТА МЕТОДІВ НАВАНТАЖЕНОГО ТЕСТУВАННЯ (LOAD TESTING) ДЛЯ ПЕРЕВІРКИ ШВИДКОСТІ БД ТА ЗАПИТІВ ДО НЕЇ

*Бабій А.С., доцент, кафедра ПІ, ХНУРЕ*  
*Шуміла А.І., студент, кафедри ПІ ХНУРЕ*

Мета дослідження полягає в аналізі існуючих підходів та методів навантажувального тестування для виявлення найоптимальнішого рішення оцінки швидкодії запитів до баз даних.

Основна проблема, яку має вирішити дослідження, полягає у визначенні ефективності обробки запитів в реальних умовах експлуатації систем. Важливість цієї проблеми виходить із зростання обсягів даних і збільшення вимог до швидкості їх обробки в сучасних інформаційних системах. Вирішення цієї проблеми включає в себе розробку і аплікацію методів тестування, які дозволяють ефективно симулювати реальні навантаження, визначати потенційні "вузькі місця" і оптимізувати роботу системи для забезпечення високої продуктивності.

Також важливою є задача оптимізації запитів до баз даних, яка передбачає розробку методів для покращення часу реакції на запити, що включає аналіз існуючих запитів, оптимізацію SQL-команд і структур даних, а також вдосконалення архітектури баз даних.

Вибір метода оцінки продуктивності баз даних у нашому дослідженні був обґрунтований необхідністю забезпечення всебічного аналізу швидкодії запитів, що включає виявлення не тільки загальних трендів, але й відхилень та потенційних проблем у відповідях системи. Далі розглянемо підходи.

Процентилі (90-й, 95-й, та 99-й). Процентилі використовуються для визначення часів відповіді, які перевищують звичайні, але не є екстремальними, а також для аналізу "хвостів" розподілу, які містять потенційні проблемні зони. Це дозволяє оцінити реальну продуктивність під час пікових навантажень, а не тільки при середніх умовах.

Середній час відгуку. Цей показник використовується для оцінки загальної продуктивності системи. Він дає змогу визначити базовий рівень ефективності обробки запитів, що є корисним для порівняльного аналізу та планування масштабування.

Медіана. Медіанний час відгуку вказує на те, який час обробки є типовим для половини запитів, ігноруючи аномальні випадки. Це корисно для виявлення стандартної продуктивності системи без упереджень через виключні значення.

Аналіз нормального розподілу: цей аналіз дозволяє зрозуміти, як розподіляються часи відгуку, виявляючи відхилення від норми, що можуть вказувати на наявність технічних чи програмних проблем.

Застосування методу нормального розподілу також дає змогу виконати прогнозування та моніторинг ефективності. Враховуючи динаміку розподілу в часі, можна спрогнозувати, як система поведе себе під впливом різних умов навантаження та оптимізації.



Також є можливість виявлення аномалій чи викидів в розподілі часів відповіді. Поєднуючи цю інформацію зі статистикою викидів, можна легко визначити та вивчати незвичайні патерни, які можуть вказувати на проблеми в системі або неочікувані зміни.

Застосування методу нормального розподілу також дає змогу виконати прогнозування та моніторинг ефективності. Враховуючи динаміку розподілу в часі, можна спрогнозувати, як система поведе себе під впливом різних умов навантаження та оптимізації.

Виходячи із зазначеного – до нашого методу нормального розподілу варто додати метод міжквартального розмаху. Міжквартальний розмах (IQR) – це міра розмаху, яка визначається як різниця між верхнім (Q3) та нижнім (Q1) квантилями набору даних.

Цей показник дозволяє виявити різницю між центральними 50% значень в наборі даних, і він широко використовується для виявлення викидів та визначення ступеня варіабельності в даних [1].

Застосування IQR в нашій вибірці, може бути вкрай корисним. IQR дозволяє визначити «типовий» діапазон значень, які лежать в межах між Q1 та Q3. Це дозволяє оцінити ступінь стабільності та неперервності швидкісного режиму системи. IQR може бути використаний для визначення асиметрії розподілу швидкості, але важливо також розглядати відхилення від середнього за допомогою методу нормального розподілу. Загальна ідея полягає в тому, що IQR дозволяє визначити та відокремити центральні 50% значень у наборі даних, що може бути корисним при виявленні аномалій та змін в швидкодії системи. Поєднання цього методу з методом нормального розподілу надає більш комплексний підхід для розуміння та оптимізації швидкодії запитів.

Хоча метод наведені вище є доволі самостійними, але нам необхідно буде бачити усю картину, усі параметри часу нашої вибірки, тому варто застосувати мультимодальний підхід який буде реалізовано програмно а візуальна складова буде відображена у консольному режимі.

#### Список літератури

1. Statistics By Jim. (2018). Normal Distribution in Statistics. <https://statisticsbyjim.com/basics/normal-distribution/>.