

ДИНАМІЧНЕ КЕРУВАННЯ ЧЕРГАМИ

Гольдінер Д. І.

Харківський національний університет радіоелектроніки, м. Харків

Попит є базовим елементом розвитку технологій і послуг, оскільки саме він надає змісту та формує мотивацію для подальших удосконалень. Чим вищим стає попит, тим активніше це стимулює еволюцію пропозиції, що, у свою чергу, визначає критичну важливість здатності послуги до масштабування.

Системи масового обслуговування описують поведінку процесів надходження, очікування та обробки запитів у ситуаціях, коли зростаючий попит перевищує можливості пропозиції забезпечити своєчасне обслуговування.

Завжди вирашним є шлях оптимізації процесів для підвищення пропускної здатності системи – це обов'язковий етап побудови продуктивних сервісів. Проте, на жаль, у певний момент складність застосування оптимізаційних рішень починає зростати експоненційно, що призводить до значного збільшення витрат на підтримку та подальший розвиток системи.

Протягом певного періоду часу потреба може задовольнятися шляхом переходу до багатоканальних систем і збільшення кількості серверів обслуговування. Однак у більшості випадків таке рішення має обмеження, зокрема: лімітований фізичний простір для масштабування, високі витрати або тривалий час, необхідний для розширення виробничих потужностей.

У момент, коли всі інструменти підвищення продуктивності системи досягають своїх меж, неминуче починають накопичуватися необроблені заявки, що зумовлює виникнення черги. Така черга, залежно від обставин, може мати обмеження на свою довжину або бути нескінченною. Водночас стан переповнення для обмеженої черги або нескінченно довге очікування у випадку відсутності обмежень на кількість місць у черзі неминуче виникають тоді, коли інтенсивність надходження заявок перевищує пропускну здатність системи.

Останнім бастіоном у боротьбі за підвищення ефективності системи може виступати складніша організація пріоритетів і структури самої черги. Для досягнення бажаного ефекту важливо мати глибоке розуміння прикладної області та якісні аналітичні дані щодо ймовірного розподілу вхідного потоку заявок і часу їх обслуговування. При впровадженні пріоритезації запитів у черзі можна поєднувати різні політики, щоб досягти максимальної продуктивності системи [1].

Ефективність поєднання різних стратегій динамічного балансування черг у системах масового обслуговування було досліджено з використанням комп'ютерної програми, написаної мовою Go, у якій реалізовано модель взаємодіючого співвиконання для паралельних обчислень.

Література:

1. Гольдінер Д. І., Матвієнко О. І. Зменшення ймовірності відмови в системах масового обслуговування з обмеженою чергою із застосуванням пріоритезації за розміром та штучного інтелекту // Біоніка інтелекту, 2024. № 1 (100). с. 36-42.