

О. О. Ковалюк, к. т. н., доц.

АНАЛІЗ ПРОГРАМНИХ ЗАСОБІВ КЕШУВАННЯ ДАНИХ У ВИСОКОНАВАНТАЖЕНИХ СИСТЕМАХ

У статті досліджено програмне забезпечення кешування даних для оптимізації роботи високонавантажених веб-систем та обґрунтовано вибір засобу кешування залежно від обсягів даних.

Ключові слова: кешування, високонавантажені системи, веб-сайти, оптимізація.

Вступ

Сьогодні значну кількість комп'ютеризованих систем створюють у вигляді веб-додатків, коли клієнтом виступає веб-браузер, що взаємодіє із веб-сервером. Таку архітектуру реалізують численні комп'ютеризовані системи управління, системи управління бізнес-процесами, веб-сайти. Головною перевагою цього підходу є використання "тонкого клієнта" (браузера), що усуває необхідність інсталяції додаткового програмного забезпечення в користувача. Одним із головних завдань, що постають під час роботи веб-систем, є забезпечення її швидкодії. На сьогодні найрозповсюдженіший шлях розв'язання цього завдання полягає в кешуванні даних — розміщенні даних у проміжному буфері з метою більш швидкого доступу до них.

Існують різні підходи щодо кешування даних та засоби їхньої реалізації [1].

За механікою роботи виділяють такі види кешування [2]:

- lazy cache (лінійний кеш) – зберігає дані й віддає їх, поки кеш не застаріє;
- synchronized cache (синхронізований кеш) – клієнт разом із даними отримує мітку часу й під час наступного звернення запитує, чи не змінилися дані, щоб повторно їх не отримувати. Такий підхід реалізує http-протокол;
- write-through cache (кеш наскрізного запису) – будь-яка зміна даних одночасно відбувається як у джерелі даних, так і в кеші.

За типом даних можна виділити кешування:

- даних, які відображаються користувачеві (front-end);
- даних, які використовують для формування html-розмітки (back-end).

За місцем розташування кешу можна виділити такі види кешування (рис. 1):

- кешування на стороні клієнта (локальний кеш браузера);
- локальний кеш веб-сервера (використовують для кешування статичного контенту);
- кешування сервером додатків.

Незважаючи на широке використання кешування, загальних правил вибору системи кешування не існує [3, 4]. Вибір засобів кешування як правило здійснюється розробником на основі власного досвіду та вподобань, тому актуальною є задача дослідження ефективності засобів кешування для різних обсягів даних.

Метою статті є отримання залежностей часу запису/зчитування від кількості даних у кеші та визначення оптимального засобу кешування на основі експериментальних досліджень.

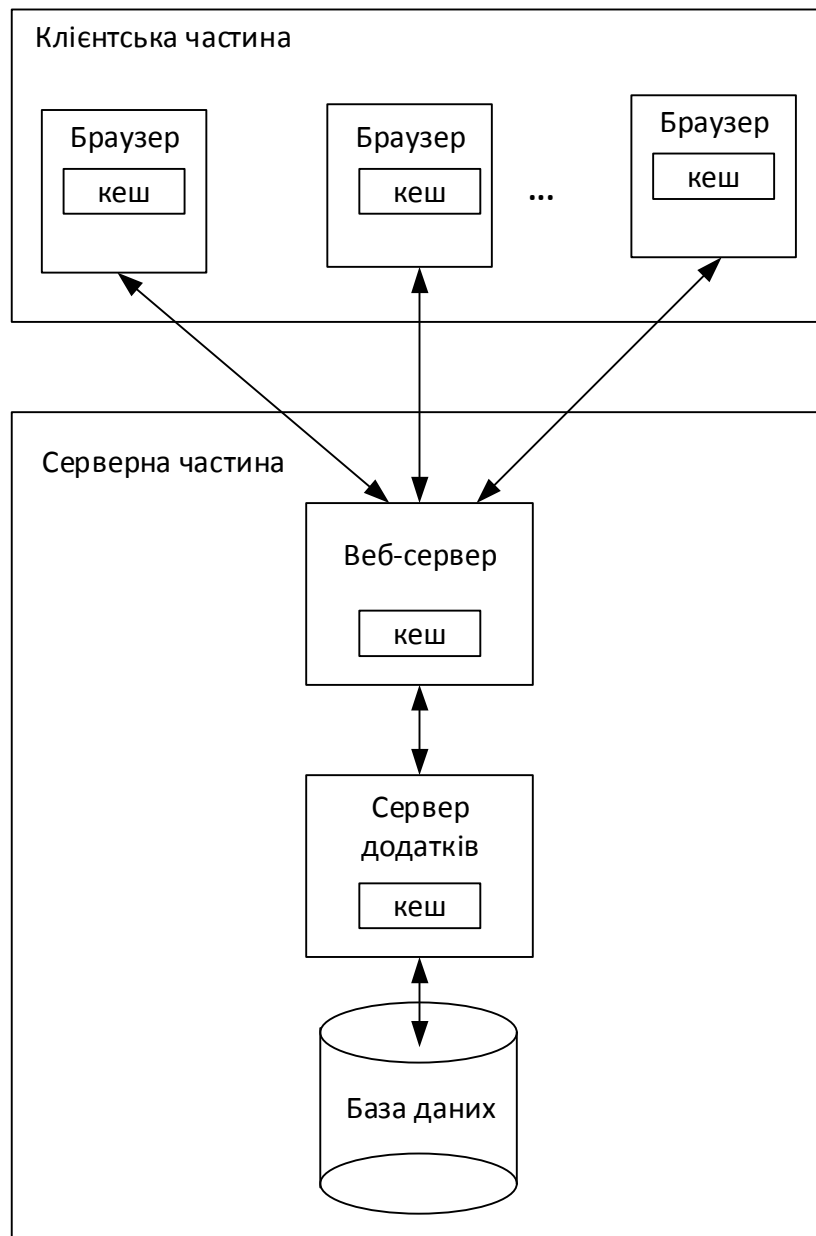


Рис. 1. Можливе розташування засобів кешування

Методика проведення експериментальних досліджень

У дослідженні проведено аналіз кешування результатів виконання SQL-запитів на прикладі інтернет-сайту www.zabudovnyk.com.ua — високонавантаженого порталу нерухомості. Кешування результатів виконання SQL-запитів обґрунтоване тим, що звернення до бази даних є “вузьким місцем” багатьох веб-систем. Портал створено з використанням бібліотеки Zend Framework 1, тому в якості засобів кешування тестуватимемо програмне забезпечення, підтримка якого реалізована в Zend Framework:

- APC;
- Memcached;
- Файловий кеш.

Методика проведення експерименту складається з таких кроків:

1. Отримання експериментальних даних щодо швидкості *запису* даних у кеш для різних обсягів даних. На цьому кроці дані вибирають з бази даних, формують ключ для

ідентифікації даних у кеші, дані розміщують у кеші. Під часом розміщення даних у кеші розглядають час виконання функції, яка зберігає підготовлені дані в кеш.

2. Отримання експериментальних даних щодо швидкості зчитування даних з кешу для даних, розміщених у кеші на кроці 1.
3. Візуалізація даних.

Умови проведення експерименту:

- дослідження проводять на тестовому сервері (Linux Fedora, SSD-диски);
- час видалення даних з кешу не досліджують;
- вважають, що дані в кеші існують (потрапляння в кеш).
- для отримання більш точних результатів використовують кілька ітерацій моделювання й визначають усереднені результати.

З метою гнучкого налаштування системи та полегшення дослідження проведено автоматизацію вибору засобів кешування. Тип засобу кешування вказують у конфігураційному файлі системи, і перемикання між засобами кешування не потребує зміни програмного коду.

Результати експериментальних досліджень

Результати запису даних у кеш наведено в табл. 1 та на рис. 2.

Таблиця 1

Час запису даних у кеш, с

Кількість даних/ Кеш	500	5000	50000	500000
APC	0,046	0,46	4,5	45
File	1,8853	70,0000	-	-
Memcached	0,053	0,53	5,3	53

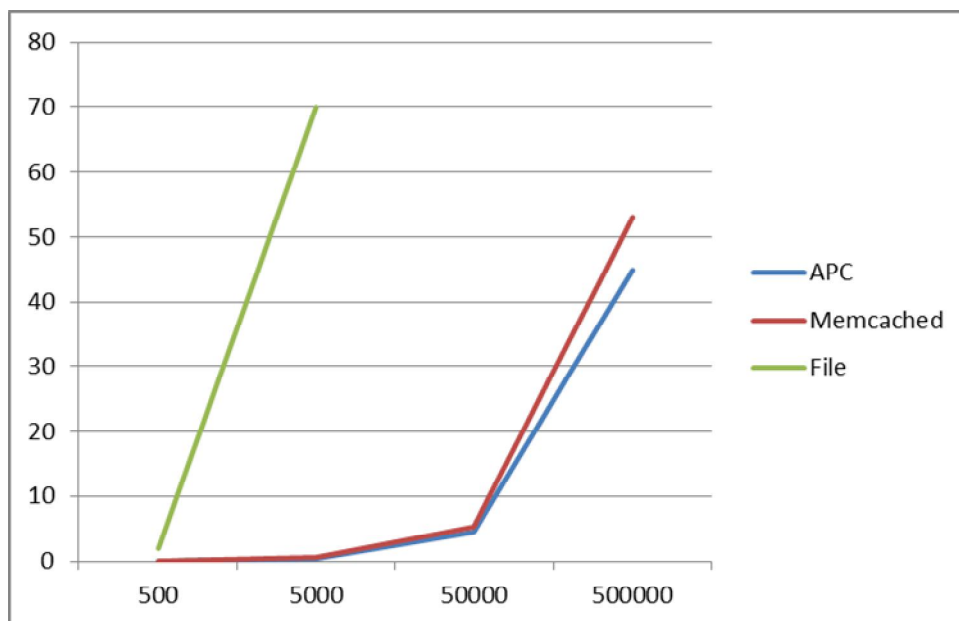


Рис. 2. Результати запису даних у кеш

Результати зчитування даних з кешу наведено в табл. 2 та на рис. 3.

Таблиця 2

Результати зчитування даних з кешу, с

Кількість даних/ Кеш	500	5000	50000	500000
APC	0,0108	0,1089	0,6809	13,3850
File	0,0932	1,0519	-	-
Memcache	0,0178	0,1760	1,7935	18,0930

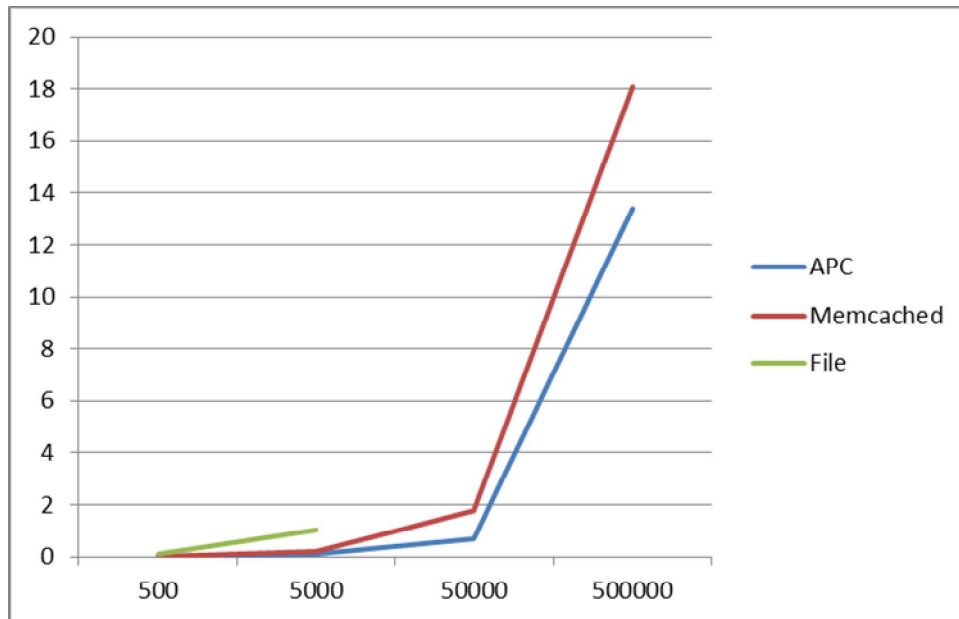


Рис. 3. Результати зчитування даних з кешу

Висновки

Отже, проведено дослідження засобів кешування, на основі якого можна зробити такі висновки:

- Засоби кешування даних в оперативній пам'яті (APC, Memcached) значно швидші, ніж файловий кеш.
- Для великої кількості даних використання файлового кешу недоцільне, оскільки кожне значення зберігається в окремому файлі. З цієї причини тестування файлового кешу для кількості ітерацій понад 5000 було перерване через тривалий час виконання.
- Час запису перевищує час зчитування. Для APC і Memcached у кілька разів. Для файлового кешу в кілька десятків разів.
- Для APC і Memcached час запису практично не залежить від кількості даних, що вже містяться в кеші.
- Для APC і Memcached час зчитування незначно збільшується під час зростання кількості даних, що містяться в кеші.

Отже, методику вибору засобів кешування можна сформулювати у вигляді таких кроків:

1. У якості основного кешу обрати програмний засіб, що використовує для зберігання даних оперативну пам'ять. Перевага надають APC.
2. Якщо ресурси оперативної пам'яті обмежені, то можна додатково використати файловий кеш для кешування до 5000 сутностей.

У разі застосування APC доречним буде використання його вбудованого веб-інтерфейсу для моніторингу основних показників використання кешу (рис. 4).

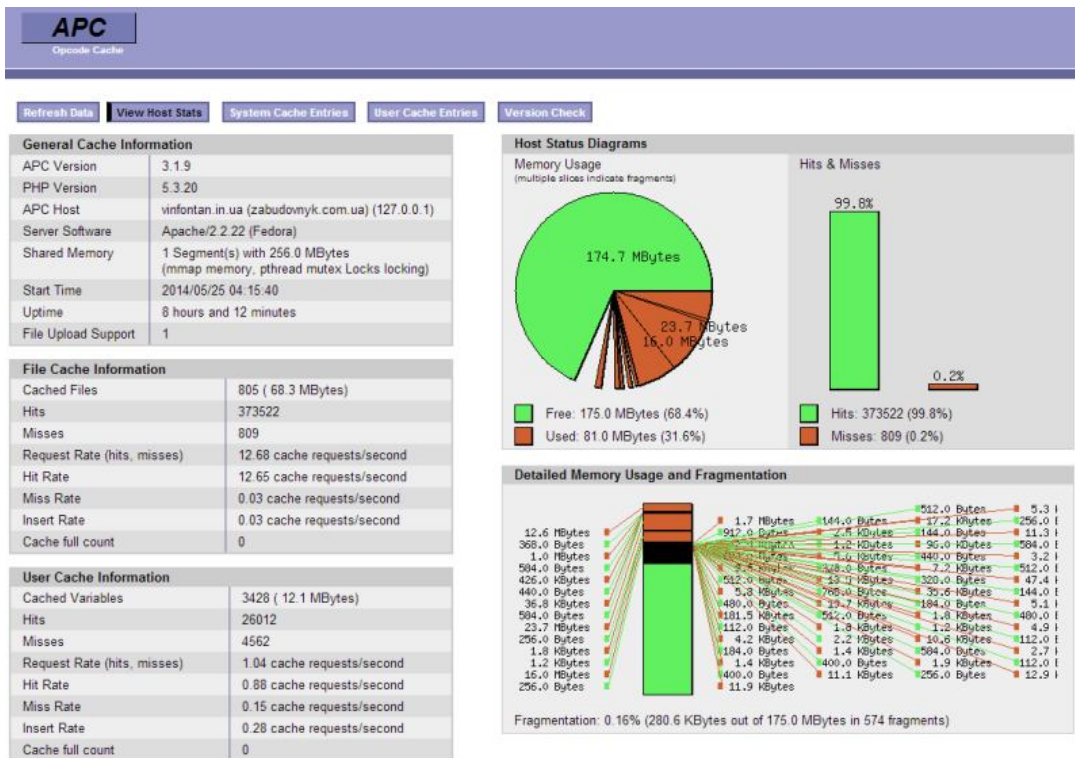


Рис. 4. Веб-інтерфейс APC

СПИСОК ЛІТЕРАТУРИ

1. Ботыгин И. А. Исследование методов увеличения производительности веб-приложений // И. А. Ботыгин, К. А. Каликин / Известия Томского политехнического университета. – 2008. – Т. 312, № 5. – С. 109 – 114.
2. Стас Выщепан Стратегия кеширования в приложении [Электронный ресурс] // Режим доступа : <http://habrahabr.ru/post/168725>.
3. Теория кэша (часть вторая, практическая, дополненная) [Электронный ресурс] // Режим доступа : <http://habrahabr.ru/post/38911/>.
4. Michal Špaček Caching Strategies [Электронный ресурс] // Режим доступа : <http://www.slideshare.net/spaace/caching-strategies>.

Ковалюк Олег Александрович – к. т. н., доцент кафедри комп'ютерних систем управління. Вінницький національний технічний університет.