

УДК 004.652.5

ЗАСТОСУВАННЯ «ТЕОРЕМИ БРЮЕРА» В РОЗПОДІЛЕНИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ

Черніков М.М., Смітюк О. Ю.

к.т.н., доцент каф. ІС Болтъонков В. А.

Одеський національний політехнічний університет, Україна

АННОТАЦІЯ. Проведено аналіз теореми Брюера для розподілених інформаційних систем. Показано, як реалізуються компроміси у виборі умов CAP-теореми на прикладі сучасних розподілених баз даних.

Вступ. Сучасні комп'ютерні системи повинні витримувати велику кількість постійних користувачів. Навіть дуже потужні апаратні рішення не витримують інтенсивного навантаження та коштують дуже дорого. Такі проблеми вирішуються реалізацією систем, які мають не один, а декілька різних серверів з'єднаних між собою за допомогою мережі. Складність такої системи існує в програмному з'єднанні усіх компонентів в єдине ціле.

Мета роботи. Методом дослідження є аналіз використання теореми Брюера для реалізації розподілених систем.

Основна частина роботи. Теорема Брюера, яка також відома як CAP-теорема (від англ. абревіатури Consistency – узгодженість даних, Availability – доступність, Partition tolerance – стійкість до розділення) використовується для усіх сучасних розподілених систем. Твердження теореми свідчить про те, що неможливо одночасно забезпечити виконання усіх наступних вимог: узгодженість даних (усі компоненти системи бачать однакові дані на будь-який момент часу), доступність (гарантія того, що кожен запит отримає конкретну відповідь), стійкість до розділення (гарантія того, що система при ізолюваності декількох секцій або їх втраті не втрачає стабільності і здатна правильно відповідати на запити). У 2002 році Сет Джилберт та Ненсі Ліч підібрали формальні моделі синхронних та асинхронних розподілених систем за допомогою яких показали можливість компромісу в тоді ще «правилі Брюера». Цей експеримент вважається доказом CAP-теореми. [1]

Оскільки терміни теореми дуже розпливчасті, після початку її використання для NoSQL баз даних було добавлено деякі уточнення:

Від самого початку теорема передбачала, що можуть траплятися розділення мережі, які можуть тривати нескінченно довго. Тоді виконати умову узгодженості буде неможливо. Джилберт та Лінч не розглядали такі випадки, тому було вирішено вважати, що в межах теореми розглядають лише тимчасові розділення з узгодженістю у довгостроковій перспективі. Розподілені системи найчастіше з'єднані між собою локальною мережею в якій її розділення майже неможливі.

Узгодженість у довгостроковій перспективі може бути виконана лише при умові відсутності оновлення даних до отримання цієї узгодженості. Якщо система потребує N часу для отримання узгодженості, оновлення даних цієї системи раніше N часу зробить її постійно неузгодженою.

Якщо система гарантує лише неатомарну узгодженість, тобто потребує певного часу N для отримання узгодженості, неможливо гарантувати коректність її роботи. [2]

Основною задачею при розробці розподілених систем є правильне використання компромісу для отримання найкращих результатів. Відхилення в одну чи іншу сторону неможливо використовувати для будь-якої системи, це залежить лише від задач, які до неї ставлять.

На рис. 1 зображено приклади баз даних, які реалізують компроміс між умовами теореми Брюера.

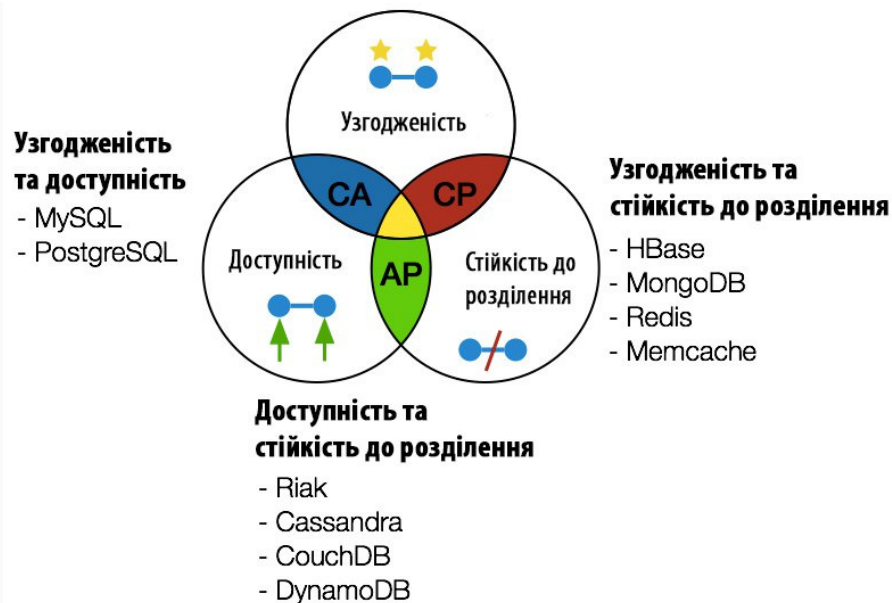


Рисунок 1 – Приклад деяких розподілених баз даних.

Розглянемо декілька з цих систем:

PostgreSQL. Типова SQL база даних, яка не має стійкості до розділення. Як система вона працює в рамках лише одного комп'ютера та не може підтримувати горизонтальної масштабованості. При запиті усі дані завжди доступні та узгоджені (оскільки не має розділення даних), але неможливо отримати перевагу розділення навантаження.

Riak. NoSQL база даних, дані якої завжди доступні та можуть бути розподілені на декілька компонентів у системі, але не завжди узгоджені між ними. Riak передбачає випадок, що дані з однаковим індексом будуть різні у різних вузлах та потребують часу на отримання повної узгодженості. При роз'єднанні одного з компонентів системи, його дані завжди будуть доступні в інших компонентах, але не завжди будуть мати фінальний стан.

MongoDB. Ця база даних реалізує умову постійної узгодженості та можливості розподілення, але не реалізує доступність. При збереженні даних можливо вибрати вузол для роботи. Це дає можливість розділити дані між компонентами системи для зменшення навантаження, але кожен з компонентів буде мати свій набір і відповідати за нього. При роз'єднанні вузла від системи його дані перестають бути доступними на час роз'єднання.

Висновки. В результаті аналізу сучасних розподілених інформаційних систем на прикладі баз даних було виявлено переваги та недоліки різних типів компромісу у виборі умов теореми Брюера. Для отримання цілісності та повної доступності даних використовуються нерозподілені системи, які не мають переваг в потужності. Розподілені системи можуть бути або узгоджені, але не завжди доступними для використання, або завжди доступними та не завжди узгодженими. Але оскільки сучасні мережі досить рідко дозволяють роз'єднання вузлів в системі, розподілені системи вважають найкращим рішенням для витримування великих навантажень.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Gilbert S., Lynch N. Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services. // ACM SIGACT News. – 2002 – Vol. 33, № 2 – p.51-59.
2. Brewer, E. (2012). CAP Twelve Years Later: How the "Rules" Have Changed. // IEEE Computer. 2012 – Vol. 45, № 2 –23-29.