

УДК 004.6

¹С. Теленик, ²О. Карнаухов, ¹Р. Плотник

¹(Національний технічний університет України «КПІ»)

²(Тернопільський національний технічний університет імені Івана Пулюя)

ПЕРСПЕКТИВИ ВИКОРИСТАННЯ ВІРТУАЛІЗАЦІЇ СХОВИЩ ДАНИХ

Сьогодні технології віртуалізації широко використовуються для підвищення ефективності використання ресурсів зберігання даних у розподілених системах. Сутність віртуалізації сховищ даних становить агрегування фізичних пристроїв зберігання даних із різними протоколами інтерфейсів в єдиний віртуальний пул. В цьому пулі можна створювати віртуальні томи, що розглядаються хостовим сервером як локально підключені логічні пристрої. Перевагами віртуалізації сховищ даних є:

- вирішення проблем роботи з гетерогенними архітектурами зберігання даних (диски, контролери сховищ та інше обладнання тепер працюють «за лаштунками»);
- забезпечення важливих технологічних переваг, насамперед високої доступності, масштабованості і безпеки, швидкого резервування і відновлення даних;
- створення економічно вигідного середовища даних потрібного об'єму і легке його розширення за необхідності;
- простота управління та ін.

Але досягнення наведених вище переваг можливе лише за умови використання ефективної системи управління сховищами даних, побудованими на основі технології віртуалізації. На сьогоднішній день існує ряд готових рішень, наприклад VMWare View, Citrix XenDesktop, Microsoft Virtual Desktop Infrastructure, Parallels Virtual Desktop Infrastructure. На жаль, покладені в основу їх функціонування моделі невідомі. До того ж завжди залишається потреба підвищити їх ефективність за рахунок оптимізації параметрів і режимів функціонування сховища даних. У доповіді розглядається проблема оптимального розподілу пам'яті пулу між томами. Користувачам виділяються віртуальні машини, які можуть мігрувати між серверами. Уся пам'ять, що об'єднує окремі сховища даних та окремі дискові масиви, віртуалізується і складає один віртуальний пул. Віртуальний пул розподіляється на віртуальні томи. Кожний користувач працює з відповідним віртуальним томом, на якому зберігаються його дані. Віртуальні томи складаються з виділених для них елементів пам'яті різних рівнів. Рівні сховища даних характеризуються набором технічних параметрів, а дані і, отже, віртуальні томи користувачів – набором вимог до віртуальних томів. Основними параметрами рівнів сховищ даних є об'єм (Gb), час доступу до 1 блоку збереження (ms), вартість збереження 1 блоку (\$/Mb). Дані зберігаються розділеними на блоки, які реплікуються на декілька рівнів. Запропонована відповідна формальна модель розподілу фізичних ресурсів зберігання даних між віртуальними томами з урахуванням обмежень на час доступу, надійність збереження даних та технологічних обмежень. В якості критерію оптимального розподілу фізичних ресурсів запропоновано використовувати мінімізацію витрат на збереження даних.

Модель розподілу становить собою задачу цілочисельного програмування. Для її розв'язання запропоновано варіант генетичного, керований генетичний та евристичний алгоритми. Наведені результати експериментальних досліджень підтверджують, що застосування реалізації зазначених моделі і методу дозволяє підвищити ефективність розподілу пам'яті сховища між користувачами.

Подальші дослідження пов'язані з розробленням моделей, які підвищують ефективність використання технології MapReduce для управління сховищами даних.