

УДК 004.33

М.А. Сибиряков

Способы повышения производительности систем хранения данных

Аннотация. В статье рассматриваются основные способы повышения производительности систем хранения данных (СХД).

Определяется направление совершенствования высокопроизводительных СХД (ВСХД).

Ключевые слова и фразы: системы хранения данных, кэш-память.

Введение

Система хранения данных представляет собой сложный аппаратно-программный комплекс, реализующий функции надежного хранения информации ([рис.1](#)). Первые архитектуры таких систем стали появляться в начале 90-х годов прошлого века. Тогда элементная база уже была достаточно развита для возможности создания подобных систем.

Требования к производительности СХД постоянно растут в связи с быстрым увеличением объемов хранимой информации и потребностью в обеспечении быстрого доступа к ней [1, 2, 3]. Создается потребность в новых решениях: в разработке новых архитектур, компонентов и более совершенных методов алгоритмов обработки данных в системах хранения. Поэтому основной целью данной публикации является анализ основных способов повышения производительности СХД для определения наиболее актуального направления исследований в этой области.

Производительность систем хранения данных

Производительность любой СХД определяется параметром IOPS (числом операций ввода вывода в секунду, которое способна обработать СХД) и временем отклика системы [6, 8, 10, 11]. Данные параметры зависят от архитектурных особенностей системы и от характеристик ее компонент.

Существует два основных способа повышения производительности СХД [2]:

© М. А. Сибиряков, 2015

© Поволжский государственный технологический университет, 2015

© ПРОГРАММНЫЕ СИСТЕМЫ: ТЕОРИЯ И ПРИЛОЖЕНИЯ, 2015

- применение более быстрых систем и компонентов;
- разработка улучшенных алгоритмов обработки.

В основном производительность подобных систем повышается первым способом, за счет использования более быстрых и производительных компонентов (таких как процессоры, коммутационные структуры и др.), за счет увеличения их числа и наращивания объемов кэш-памяти [1, 2, 3].

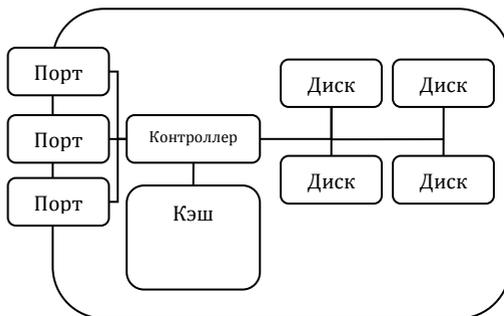


Рис. 1. Базовая структура СХД

Причиной этого является получение быстрой коммерческой выгоды компаниями-разработчиками за счет модернизации и масштабирования уже существующих систем. Это самый простой путь, в отличие от исследований в области разработки алгоритмов обработки, которые требуют больших затрат и большего времени на коммерческую отдачу.

Кэш-память в СХД

В любых вычислительных системах память всегда являлась узким местом в плане производительности. Также и в системах хранения данных кэш-память является основным компонентом, влияющим на производительность всей системы.

Рассмотрим основные параметры, влияющие на производительность подсистемы кэш-памяти. Производительность кэш-памяти определяется следующим образом [5].

$T_{average} = T_{hit} + MissRate * T_{miss}$, где $T_{average}$ – среднее время доступа к памяти, T_{hit} – время доступа при кэш-попадании, $MissRate$ – доля промахов при обращении к кэш-памяти, T_{miss} – время выборки данных из основной памяти (RAM, HDD)

Соответственно, чем меньше значение параметра $T_{average}$, тем выше производительность и эффективность работы кэш-памяти.

Для повышения скорости доступа к данным (т.е. снижения среднего времени доступа) необходимо:

- Уменьшение параметра *MissRate*. Данный параметр зависит от организации кэш-памяти (ассоциативность, размер блока), от ее объемов, алгоритмов вытеснения данных.
- Уменьшение параметров T_{hit} и T_{miss} . Они зависят от используемой аппаратной логики. Параметр T_{hit} в большей степени зависит от алгоритмов поиска данных в кэш-памяти.

При обработке больших объемов информации, именно кэширование позволяет значительно увеличить скорость доступа к данным, а, следовательно, повысить производительность кэш-памяти и системы хранения данных в целом [9].

В первую очередь это касается высокопроизводительных СХД. Они обеспечивают очень большие объемы кэш-памяти. Однако, несмотря на то,

что существует тенденция наращивания объемов кэш-памяти (в современных ВСХД кэш-память достигает объема в 4 Тб, рис. 2), существует проблема эффективного использования ее ресурсов [1, 2, 7].

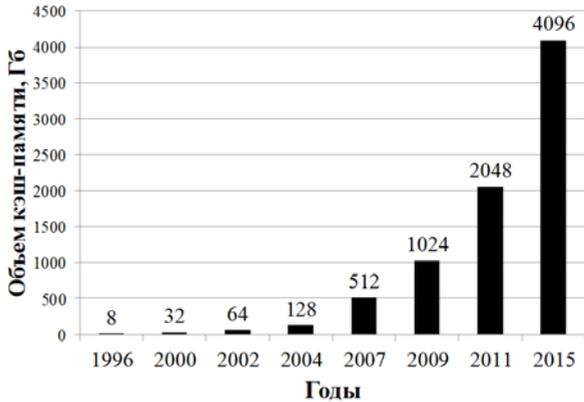


Рис. 2. Тенденция роста объема кэша в ВСХД

Сравнение с кэш-памятью процессора

Кэш-память СХД значительно отличается от кэш-памяти процессора. Главными отличиями являются: размер кэш-памяти, тип памяти и размер блока (таблица 1).

Для эффективного использования кэш-памяти процессора достаточно использование алгоритмов вытеснения данных (в основном используются LRU, FIFO и псевдо-LRU) [4, 7, 12]. В то время как для систем хранения данных необходимы более сложные методы

управления кэш-памятью, поскольку кэш-память разделяется несколькими внешними хост-узлами. Требуется разграничение кэша между ними, приоритетный доступ и правильное использование структур данных для эффективной работы алгоритмов вытеснения данных. Все это реализуется с использованием индексных таблиц, эффективных алгоритмов поиска данных.

Таблица 1

Сравнение кэш-памяти процессора и СХД

Параметры	Кэш-память процессора	Кэш-память СХД
Объем	8 Кб < L1 < 128 Кб 128 Кб < L2 < 1 Мб 1 Мб < L3 < 10 Мб	До 4096 Гб
Тип памяти	Регистровая, SDRAM	DRAM
Размер строки (блока)	16, 32, 64 байта Кэш_строка ← → ОП_блок	4 Кб Кэш_сегмент ← → Диск_трек

Заключение

Постоянное наращивание и масштабирование компонент не является единственным способом повышения производительности СХД. Необходимо также развивать и алгоритмы обработки данных.

С точки зрения развития архитектур СХД важным является совершенствование подсистемы кэш-памяти. Повышение эффективности работы кэш-памяти позволит увеличить производительность СХД. Наиболее перспективными в плане исследования и разработок в этой области являются: методы управления кэш-памятью, алгоритмы замещения и поиска данных в ней.

Список литературы

- [1] Васяева Е.С., Сибиряков М.А. Особенности развития высокопроизводительных систем хранения данных // сборник тезисов VIII Международной молодежной научной конференции «Научному прогрессу творчество молодых» (2013).
- [2] Сибиряков М.А., Васяева Е.С. Сравнительный анализ основных моделей интеллектуальных систем хранения данных в процессе их эволюции // Информационные технологии в профессиональной деятельности и научной работе. 2012. Ч1. С. 99-105.
- [3] Farley Marc. Building Storage Networks, second edition. The McGraw-Hill Companies, Inc., 2001. 596 p.
- [4] Henessy John L., Patterson David A. Computer Architecture Quantitative Approach, fourth edition. Morgan Kaufmann Publishers, 2007. 676 p.

- [5] Jacob Bruce, Ng Spencer W., Wang David T. Memory Systems Cache, Dram, Disk. Morgan Kaufmann Publishers, 2008. 982 p.
- [6] White paper: Storage is Still Not a Commodity: an Updated Comparison of High End Storage Subsystems of 9 August 2009 by Josh Krischer // Josh Krischer & Associates GmbH. 24p.

Ссылки на электронные ресурсы

- [7] Крис Касперски, Подсистема кэш-памяти как она есть URL: http://kpnc.opennet.ru/hw_cach.pdf (дата обращения 19.06.2014).
- [8] Расчет параметра IOPS URL: http://wikibon.org/wiki/v/Calculating_the_max_IOPS (дата обращения 18.06.2014).
- [9] Статья о производительности систем хранения данных URL: http://www.remmag.ru/admin/upload_data/remmag/10-3/Verysell.pdf
- [10] Статья о производительности систем хранения данных URL: <http://pc-hard.ru/hardarticles/87-what-is-iops.html> (дата обращения 18.06.2014).
- [11] Статья про IOPS URL: <http://denserov.com/2012/09/29/> (дата обращения 18.06.2014).
- [12] Технические характеристики современных процессоров IOPS URL: <http://www.intel.com/content/dam/doc/manual/64-ia-32-architectures-optimization-manual.pdf> (дата обращения 19.06.2014).

Об авторе:



Сибиряков Максим Андреевич

Аспирант ПГТУ, 25 лет. Разработки: алгоритмы обработки данных в кэш-памяти систем хранения данных.

e-mail:

maxover777@bk.ru

Образец ссылки на публикацию:

М.А. Сибиряков. Способы повышения производительности систем хранения данных // Программные системы: теория и приложения: электрон. научн. журн. 2015. Т.?, №? (?), с.??-??.URL: <http://psta.psriras.ru/read/???>

M.A. Sibiryakov. Ways of increasing productivity storage systems. ABSTRACT. The article discusses the main ways to improve the performance of storage systems.

The direction of improvement of high-end storage systems is defined.
Key Words and Phrases: storage system, cache-memory.