

УДК 004.94

## ОЦЕНКА ЭФФЕКТИВНОСТИ АЛГОРИТМОВ РАБОТЫ КЭШ-ПАМЯТИ МУЛЬТИПРОЦЕССОРА С ПОМОЩЬЮ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

Сорока Т.Е., Фельдман Л.П., Михайлова Т.В.  
Донецкий национальный технический университет

*Рассматриваются различные подходы к моделированию вычислительных систем. Приводится схема имитационной модели кэш-памяти мультипроцессора и описание ее реализации. Представлены полученные в результате моделирования зависимости производительности системы от организации кэш-памяти.*

**Введение.** Для многопроцессорных вычислительных систем задача правильной организации памяти является одной из важнейших. Применение многоуровневой кэш-памяти (наряду с прогнозированием ветвлений) является эффективным средством повышения производительности современных многоядерных процессоров [1].

Проектирование высокопроизводительных процессоров является сложной задачей и требует предварительной верификации и анализа на уровне моделей, для чего обычно используется аналитическое и имитационное моделирование [2, 3]. Достоинством аналитических моделей является их высокая гибкость, но аналитическое моделирование параллельных вычислительных систем является довольно сложным, так как представление подобных систем в виде цепей Маркова – очень громоздкая задача, а модели на основе сетей СМО обычно не достигают желаемой точности.

Решением задачи может служить разработка и анализ имитационных моделей различной сложности в сочетании с аналитической оценкой отдельных алгоритмов [2, 3]. Данная работа посвящена проектированию многопараметрической имитационной модели мультипроцессора для оценки эффективности работы алгоритмов кэш-памяти и оптимальности ее структуры. Оптимизация структуры и алгоритмов работы кэш-памяти позволяет ускорить взаимодействие процессора с памятью и, таким образом, повысить производительность всей системы.

**Постановка задачі.** В качестве первоначального объекта исследований рассмотрим простейшую топологию: все процессоры располагаются на общей системной шине. Общая структура моделируемой системы представлена на рис. 1.

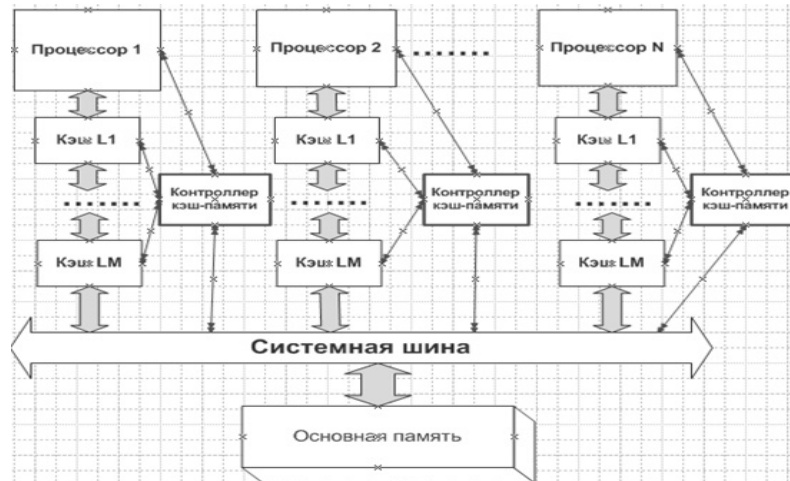


Рисунок 1 – Общая структура мультипроцессора

Зависимость между организацией кэш-памяти и производительностью системы целесообразно представить в виде некоторой функции  $G=F(Y)$ , где  $Y$  – вектор входных параметров модели (характеристики мультипроцессора),  $G$  – вектор выходных параметров (показателей эффективности системы). В свою очередь вектор  $Y$  можно представить как объединение двух множеств параметров:  $Y = Y_1 \cup Y_2, (Y_1 \cap Y_2 = \emptyset)$ , где  $Y_1$  – множество аппаратных характеристик системы, а  $Y_2$  – множество случайных величин – параметров моделирования.

В качестве аппаратных характеристик рассмотрим вектор из 11 параметров:

$$Y_1 = \{n, k, v_{кэш}, v_{оп}, C, R, M, A, t_{оп}, t_{кэш}, t_{ук}\}, \text{ где}$$

$n$  – число процессоров,  $k$  – число уровней-кэш-памяти,  $v_{кэш}$  – емкость кэш-памяти (L1),  $v_{оп}$  – емкость оперативной памяти,  $C$  – используемый протокол когерентности кэш-памяти (Write once, Synapse, MSI, MESI, MOESI),  $R$  – используемый алгоритм замещения (LRU, LFU, Random),  $M$  – используемый алгоритм согласования содержимого основной и кэш-памяти (алгоритмы сквозной записи, обратной записи и обратной записи с флагом),  $A$  – алгоритм отображения основной памяти на кэш (характеризуется ассоциативностью кэша),  $t_{оп}$  – время доступа к оперативной памяти (в

пересчете на такты процессора),  $t_{кэш}$  - время чтения/записи в кэш-память,  $t_{ук}$  - время поиска указателя в кэш-памяти.

В качестве параметров моделирования будем рассматривать вектор из 4 элементов:

$$Y_2 = \{P_{запроса}, P_{чтения}, \delta, t_{мод}\}, \text{ где}$$

$P_{запроса}$  - вероятность того, что в очередной такт процессор сгенерирует запрос к памяти,  $P_{чтения}$  - вероятность того, что это будет запрос на чтение,  $\delta$  - отклонение следующего адреса запроса по отношению к предыдущему (с целью соблюдения принципа локальности используем нормальное распределение),  $t_{мод}$  - время моделирования (число тактов, которые должен выполнить процессор).

В качестве показателей эффективности системы будем рассматривать вектор из 3 параметров:

$$G = \{t_{цикла}, \rho, h\}, \text{ где}$$

$t_{цикла}$  - среднее время обмена с памятью,  $\rho$  - загрузка системы,  $h$  - коэффициент кэш-попаданий.

Проектируемая система должна моделировать поведение мультипроцессора с характеристиками  $Y_1$ , в среде с параметрами  $Y_2$ , и как результат моделирования получать вектор характеристик эффективности  $G$ .

Система считается наиболее эффективной, если  $t_{цикла} \rightarrow (t_{кэш} + t_{ук})$ ,  $\rho \rightarrow 1$ ,  $h \rightarrow 1$ .

**Общая структура имитационной модели.** Моделирующая система реализована как многопоточное приложение (см. рис. 2) и легко может быть перенесена на распределенную архитектуру.

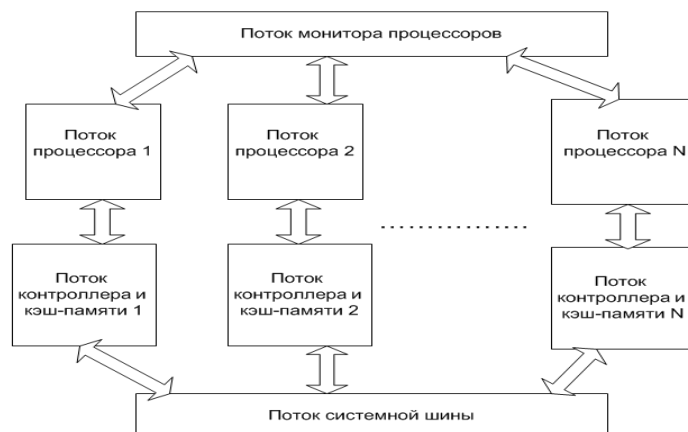


Рисунок 2 – Структура имитационной модели

Сама модель состоит из  $N$  потоков, моделирующих процессоры,  $N$  потоков для моделирования их кэш-контроллеров и кэш-памяти, единственного потока для системной шины и потока «монитора процессоров». Таким образом, для моделирования мультипроцессора с  $N$  ядрами требуется  $2N+2$  потоков; отдельный поток реализует интерфейс пользователя.

Поток процессора раз в несколько тактов генерирует запрос к памяти, после чего помещает запрос в очередь запросов кэш-контроллера и приостанавливает свою работу. Поток кэш-контроллера обрабатывает запросы процессора и после получения запрошенных данных возвращает их процессору, позволяя ему продолжить работу. Системная шина служит связующим звеном для связи кэш-контроллеров между собой и с основной памятью. Монитор процессоров не является частью мультипроцессора, он периодически опрашивает состояние процессоров и предназначен для определения коэффициента загрузки системы.

**Результаты моделирования.** Система моделирования позволяет исследовать влияние любой комбинации входных параметров на эффективность системы. В качестве примера рассмотрим зависимость времени среднего цикла памяти от алгоритмов свопинга и доли операций чтения/записи (см. рис. 3).

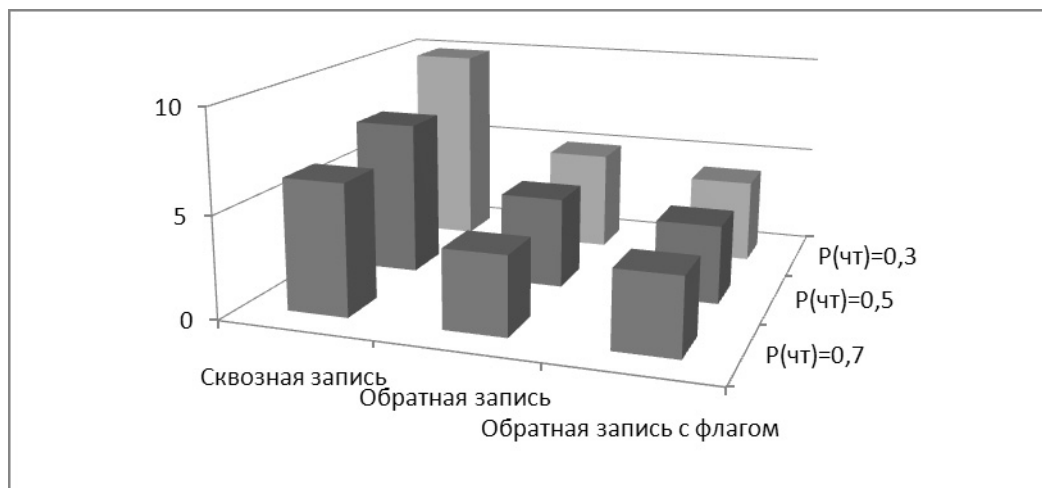


Рисунок 3 – Среднее время доступа к памяти ( $t_{\text{цикла\_оп}}=10$ )

Наилучший результат дает алгоритм обратной записи с флагом, кроме того, он малочувствителен к увеличению доли операций записи. Худший результат дает алгоритм сквозной записи, который практически неэффективен при большом количестве операций записи.

В качестве второго примера рассмотрим зависимость коэффициента кэш-попаданий от ассоциативности кэша (см. рис. 4).

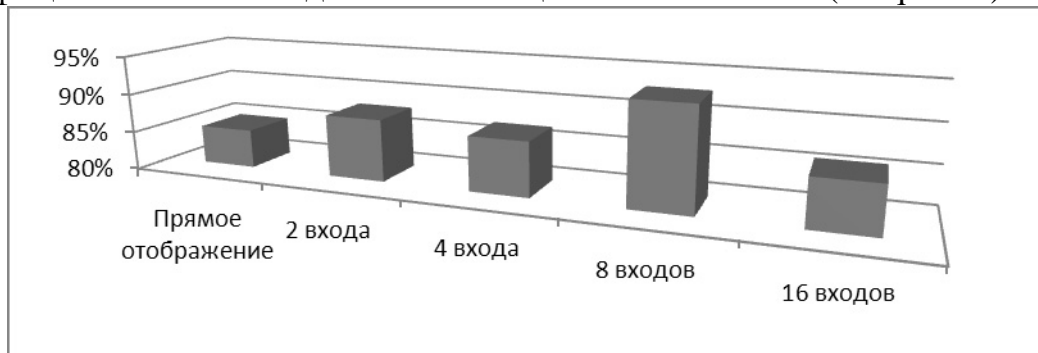


Рисунок 4 – Зависимость частоты попаданий от числа входов

Наиболее эффективно применения кэша с 8 входами, дальнейшее увеличение ассоциативности нецелесообразно.

Полученные относительные результаты практически совпадают с результатами более ранних аналитических исследований и могут служить доказательством корректности разработанной модели [4,5].

**Выводы.** Разработанная модель может применяться для выявления зависимостей между структурой кэш-памяти мультипроцессора и его производительностью. Так как некоторые входные характеристики модели являются не числовыми величинами, а алгоритмами, использование имитационной модели является удачным решением, поскольку аналитический подход потребовал бы составления отдельной модели для каждого возможного алгоритма. Модель позволяет оценить эффективность алгоритмов работы памяти без учета физических особенностей аппаратной реализации.

#### Список литературы

1. Таненбаум Э. Архитектура компьютера. 5-е изд. – СПб.: Питер, 2007. – 844 с.
2. G. E. Suh, S. Devadas, and L. Rudolph, "Analytical Cache Models with Application to Cache Partitioning." Proceedings of the 15th International Conference on Supercomputing, June 2001.
3. K. Skadron, P.S. Ahuja, M. Martonosi, and D.W. Clark. "Branch Prediction, Instruction-Window Size, and Cache Size: Performance Tradeoffs and Simulation Techniques." IEEE Transactions on Computers, 48(11):1260-81, Nov. 1999.
4. Фельдман Л.П., Дедищев В.О. Математическое обеспечение САПР: Моделирование вычислительных и управляющих систем. – К.: УМК ВО, 1992. – 256 с.
5. Пом А., Агравал О. Быстродействующие системы памяти: Пер. с англ. – М.: Мир, 1987. – 264 с.

Получено 08.09.2011