

Контроллер памяти с дополнительными признаками данных и его реализация на ПЛИС

Монахов Александр Михайлович

Студент

Физический факультет МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: monahov.aleksandr@physics.msu.ru

Последние десятилетия отмечены мощным развитием вычислительной техники. В ведущих странах мира полным ходом идут работы по созданию до 2020 года вычислительных комплексов мощностью порядка экзафлопса. Рост производительности обеспечивается в первую очередь за счет параллелизма и увеличения количества вычислительных блоков, требует соответствующего роста объемов и скорости оперативной памяти. ОЗУ экзафлопных систем, по современным расчетам, будет составлять экзабайт и более.

Важнейшими научными и инженерно-техническими проблемами развития вычислительной техники сегодня является отставание производительности ОЗУ от производительности микропроцессоров, так называемый эффект "бутылочного горлышка"[1] и сложность обеспечения на программном и аппаратном уровнях взаимодействия сотен и тысяч вычислительных потоков с общими данными в памяти, когерентность данных, так называемая проблема "синхронизации вычислительных потоков"[2]. Элегантное решение проблемы совместного использования данных в памяти несколькими потоками (тредами) и синхронизации вычислительных потоков было предложено группой Питера Кегге [3] и заключалось в переносе функционала арбитража доступа к данным из программы пользователя и/или операционной системы (семафоры (semaphore), мьютексы (mutex) и прочее) прямо в контроллер памяти и созданию т.н. «умной памяти». В простейшей интерпретации «умной памяти» к данным в ОЗУ прилагается дополнительный признак, т. н. Full-Empty(далее FE) бит. Его значение, равное единице сигнализирует, что ячейка пуста и данные недоступны для чтения (например ожидается запись результата работы другого вычислительного потока), ноль - ячейка полна и данные доступны для считывания. Считывание возможно двумя способами: обычный, когда состояние ячейки не изменяется и "с изъятием данных когда FE-бит аппаратно переводится в состояние "пусто". В этом случае любой запрос на повторное чтение будет задержан до появления данных в ячейке. Таким образом реализуется-

ся простой и понятный механизм синхронизации любого количества потоков без использования громоздких глобальных барьеров, аппаратно, без накладных расходов и без программирования.

В данной работе разработан контроллер памяти с дополнительными признаками, построенный по указанным выше принципам. Контроллер позволяет работать с ОЗУ типа DDR2 SDRAM, DDR3 SDRAM, RLDRAMIII, ZBT SRAM. Контроллер отлажен на ПЛИС Xilinx Virtex5 50T, Virtex6 240T. Рабочая частота 100-400 MHz. Контроллер поддерживает одновременную работу с 1024 вычислительными потоками, максимальная длина очереди запросов - 1024, латентность выполнения операций без дополнительных признаков данных зависит только от загруженности входной очереди и равняется количеству занятых в ней строк. Латентность контроллера памяти с дополнительными признаками $m-n+1$, где m - кол-во отложенных запросов на чтение, n - кол-во "заполняющих"запросов на запись в очереди контроллера, $m>n$.

Контроллер памяти с дополнительными признаками, разработан и применяется в рамках проекта MALT - Multicore Architecture with Lightweight Threads [4], открытый проект создания прото-типа масштабируемого сервера, предназначенного для применения в хостинг-центрах и энергоэффективных ЦОД, содержащего тысячи или десятки тысяч предельно простых процессорных ядер с разрешением конфликтов одновременного доступа к данным на уровне контроллера общей памяти и аппаратными механизмами планирования и исполнения легких потоков, работающего под управлением единой POSIX-совместимой операционной системы.

Литература

1. Bing Xie, Jeffrey Chase Characterizing Output Bottlenecks in a Supercomputer. Intel Corporation
2. Корнеев В.В. Параллельные вычислительные системы. М.: "Нолидж 1999. 320 с.
3. P.M. Kogge et al., Computer Systems with Lightweight Multithreaded Architectures, U.S. Patent 7,584,332
4. Создание прототипа масштабируемого сервера на базе ПЛИС, <http://maltsystem.ru/ru/>