

Суперкомпьютеры для графовых задач

Интерес к большим графовым задачам и технологиям их решения непрерывно растет. Какие сегодня имеются подходы к суперкомпьютерным системам для решения графовых задач и что делается в этом направлении в России?

Александр Семенов, Александр Фролов, Анатолий Никитин, Владимир Кабыкин

Обработка графов — интенсивно развивающееся сегодня направление, что неудивительно, ведь речь идет о задачах анализа социальных сетей, биоинформатики и бизнес-аналитики, о семантических базах данных и системах обеспечения национальной безопасности. Например, в 2011 году количество активных пользователей сети Facebook составило 750 млн, что на 50% больше, чем в 2010 году, и для обработки такого массива данных требуются все более мощные компьютерные системы. Считается, что приложения, использующие графовые алгоритмы, потенциально самые «тяжелые» как по объему обрабатываемых данных, так по режиму доступа (плохая пространственно-временная локализация), при этом растет интерес к ним и увеличивается их число.

Графовые задачи характеризуются нерегулярной структурой графа, низкой локализацией данных и превалированием доступа к данным над вычислениями. Все это делает их крайне неудобными для решения на наиболее распространенных сегодня суперкомпьютерах с кластерной архитектурой, ориентированных на программную модель передачи сообщений. А главное, в ориентации на применение кэшей для обеспечения толерантности к задержкам по памяти, которая на графовых задачах не срабатывает из-за плохой локализации. Для оценки производительности суперкомпьютеров на графовых задачах был специально создан рейтинговый список Graph500, в качестве вычислительного ядра реализации которого используется тест «поиск вширь» в графе. Рейтинг суперкомпьютеров Graph500 публикуется два раза в год, а места в списке распределяются с учетом размера графа и полученной производительности суперкомпьютера при его обработке (единица измерения — TEPS, то есть количество пройденных дуг в секунду), причем размер имеет более приоритетное значение, чем производительность.

Graph500: адекватный рейтинг

В 2010 году был предложен список Graph500, претендующий на более адекватное, чем в Top500, отражение рейтинга суперкомпьютеров, способных обрабатывать большие массивы данных. Что побудило создать очередной тест и в чем его суть?

Леонид Эйсымонт, Александр Фролов, Александр Семенов

В настоящее время опубликована вторая редакция списка [Graph500](http://www.graph500.org/june2011.html) (<http://www.graph500.org/june2011.html>), в которой представлены 29 суперкомпьютеров (таблица 1). Первые два места занимает BlueGene/P, однако на третьем месте находится отечественная система «Ломоносов». При этом на данной машине была достигнута рекордная производительность на Graph500 благодаря качественной агрегации коротких сообщений, использованной в библиотеке DISLIB, разработанной в компании «Т-Платформы». Для сравнения, не представленный официально в рейтинге компьютер «Ангара» по результатам имитационного моделирования показал 56380 MTEPS при обработке графа с количеством вершин 12,8 млн.

Таблица 1. Наиболее интересные результаты из Graph500 (июнь 2011)

№	Суперкомпьютер	Размер графа ¹	Производительность
1	Intrepid (IBM BlueGene/P, 32786 узлов, 131072 ядер)	38	18508 MTEPS
3	Lomonosov (MPP, 4096 узлов, 8192 ядер)	37	43471 MTEPS
4	Hopper (Cray XE6, 1800 узлов, 43200 ядер)	37	25075 MTEPS
9	Endeavor (Westmere X5670, 256 процессоров/3072 ядер)	33	6860 MTEPS
15	DAS-4/MU (Supermicro, 64 узла, 512 ядер)	31	4642 MTEPS
16	Matterhorn (Cray XMT2, 64 узла)	31	885 MTEPS
21	Cougarxmt (Cray XMT, 128 узлов)	29	1223 MTEPS

26	Vortex (Convey HC-1ex, 1 узел, 4 ядра, 4 ПЛИС)	27	773 MTEPS
27	Westmere E7-4870 2,4 ГГц (один узел, 40 ядер)	27	705 MTEPS
¹ Тест Graph500 имеет два параметра: SCALE и edgfactor. SCALE задает общее число вершин графа, которое равно 2SCALE. Параметр edgfactor определяет количество дуг в графе, равное edgfactor*(число вершин). По умолчанию edgfactor = 16. Размер графа в таблице — это SCALE.			

Повышение интереса к графовым задачам сопровождается появлением новых технологий, ориентированных на их эффективное решение; например, в феврале 2011 года компания Cray анонсировала суперкомпьютер Cray XMT2 из линейки XMT [1] мультитредовых суперкомпьютеров с глобально адресуемой памятью. В Cray XMT2 применяется обновленный 128-тредовый микропроцессор Threadstorm и коммуникационная сеть SeaStar2. От Cray не отстают и производители ускорителей на программируемых логических матрицах (FPGA). Компания Convey Computer продемонстрировала результаты выполнения Graph500 на своей системе Convey HC-1ex на базе FPGA с уникальной подсистемой памяти, оптимизированной под нерегулярный доступ [2].

Интересна программная система SWARM (SWift Adaptive Runtime Machine), разрабатываемая компанией ET International. Мультитредово-потокковая вычислительная модель, на которой основана SWARM, позволяет резко поднять эффективность решения нерегулярных задач по сравнению с моделью передачи сообщений, применяемой в MPI.

Российский суперкомпьютер с глобально адресуемой памятью

В пылу массового увлечения вычислительными кластерами часто забывают, что эти системы непригодны для решения задач, требующих эффективной работы с глобально адресуемой памятью большого объема.

Анатолий Слущкин, Леонид Эйсымонт

В июле 2011 года Национальная лаборатория Сандия и Университет Нью-Мексико начали долгосрочный совместный проект по разработке принципиально новой вычислительной системы для решения графовых задач, которая в перспективе будет использоваться в экзафлопсных суперкомпьютерах. В качестве требования предъявляется не только повышение производительности, но и снижение количества потребляемой энергии.

В России, судя по открытым источникам, компьютерной обработке графов уделяется пока мало внимания, что отчасти обусловлено мнением об отсутствии приложений для обработки больших графов, и, как следствие, слабо развиваются технологии, ориентированные на ускорение выполнения графовых задач.

В ОАО «НИЦЭВТ» проводятся исследования по созданию мультитредового суперкомпьютера с глобально адресуемой памятью [3]. На данный момент разработан макет конвейера команд ядра мультитредового процессора J7

на FPGA, разрабатывается новый макет на FPGA, который будет включать конвейер команд и подсистему памяти.

Далее приведен обзор технологий, ориентированных на повышение эффективности решения графовых задач, а также текущие результаты по макетированию мультитредового процессора J7 и планы по его развитию.

Cray XMT2

Суперкомпьютер Cray XMT2 был представлен в феврале 2011 года спустя четыре года после появления первой версии Cray XMT, хотя архитектура Cray XMT2 была разработана еще в конце 80-х годов Бартоном Смитом и впервые реализована в системе Tera MTA, а затем — в Cray MTA-2. В целом с момента создания Tera MTA архитектура не сильно изменилась, за исключением того что в Tera MTA и Cray MTA-2 использовалась глобально адресуемая память с равномерным доступом (UMA), а в Cray XMT и Cray XMT2 — с неравномерным доступом (NUMA).

Особенность Cray XMT2 — 128-тредовый микропроцессор Threadstorm, поддерживающий аппаратно глобально адресуемую память. Предполагается, что большое количество тредовых устройств позволит скрыть задержки выполнения команд обращений к памяти на фоне выполнения других тредов (совмещение обращений к памяти с вычислениями), что позволит работать не на задержке, а на темпе выдачи обращений.

В Threadstorm реализована сегментная глобально адресуемая память. Виртуальные сегменты могут быть распределены между узлами, поддерживается блочно-циклическое распределение, а также скремблирование — зашумление виртуального адреса, позволяющее реализовать UMA для отдельных сегментов. Threadstorm совместим с сокетом AMD 940 (технология Torrenza), что позволило в XMT и XMT2 полностью использовать инфраструктуру Cray XT4, в частности коммуникационную сеть SeaStar2 с топологией 3D-тор.

Каждое тредовое устройство обладает своим набором регистров (управляющие регистры, 8 регистров адресов команд, 32 регистра общего назначения) и разделяет три конвейера команд: М-конвейер, А-конвейер, С-конвейер, которые отвечают за выполнение команд — обращений к памяти, за арифметические операции и за команды передачи управления соответственно. На каждом такте выдается до трех команд из одного треда, но не более одной для каждого конвейера. В Threadstorm поддерживается совмещение выполнения команд — обращений к памяти от одного тредового устройства. Каждое тредовое устройство может выдать до восьми обращений к памяти, что в сумме для всех тредов дает 1024 отложенных обращения.

Отличия между Cray XMT и Cray XMT2 касаются производительности подсистемы памяти в части контроллеров памяти. Кроме того, в XMT2 появились суперсегменты, позволяющие одним содержимым буфера быстрой переадресации (TLB) перекрывать всю виртуальную память, что дало возможность для задач, оперирующих большими массивами данных, избавиться от промахов в TLB. К сожалению, более подробная информация недоступна.

Продвижением суперкомпьютеров серии XMT занимается Custom Engineering Group, и сегодня известно о двух поставленных системах Cray XMT2: 64-процессорном суперкомпьютере в Швейцарском национальном суперкомпьютерном центре и в Центре прикладных высокопроизводительных вычислений в Денвилле (США). Последний организован некоммерческой группой Noblis, одним из направлений деятельности которой является продвижение проектов по национальной безопасности.

Cray XMT на 128 узлах и Cray XMT2 на 64 узлах показывают производительность в 1223 и 885 МТЕPS соответственно, что меньше, чем у кластера DAS-4/U, построенного на основе коммерческих компонентов. Удивительного в этом ничего нет — инфраструктура Cray XT4, используемая в Cray XMT2, отстает на два поколения сетевых маршрутизаторов от Cray XE6. С другой стороны, в тесте для Graph500 поиска вширь в графе для продолжения счета при посылке данных на удаленные узлы нет необходимости получения обработанной информации с этих узлов — природа обменов односторонняя, что позволяет агрегировать пакеты и лучше использовать пропускную способность сети. Однако в реальных графовых задачах, напротив, требуется послать сообщения на удаленный узел, получить сообщение обратно и обработать его, поэтому Cray XMT продолжает позиционироваться как идеальная машина для решения реальных графовых задач. Среди преимуществ Cray XMT/XMT2 стоит отметить глобально адресуемую память, а также богатые аппаратные возможности взаимодействия и синхронизации большого количества аппаратных тредов. Параллельное программирование для Cray XMT отличается простотой и похоже на OpenMP, а продуктивность программирования является важнейшим показателем современных высокопроизводительных технологий.

Convey HC-1ex

Многофункциональная реконфигурируемая система Convey HC-1ex — второе поколение встраиваемых в сокет Intel LGA 775 сопроцессоров на базе FPGA, выпускаемых компанией Convey Computers. На плате Convey HC-1ex две микросхемы FPGA (Xilinx Virtex5 LX110) используются для интерфейса с центральным процессором и декодирования команд, четыре (Virtex6 LX760) — для выполнения пользовательской программы (эти платы также называют Application Engines, AE), остальные восемь (Virtex5 LX155) реализуют контроллеры памяти. Для перепрограммирования пользователю доступны четыре AE, в которые можно загружать либо собственную прошивку, либо одну из предоставляемых производителем, специализированных под различные области приложений (финансовый анализ, биоинформатика, распознавание речи, графовые задачи и др.).

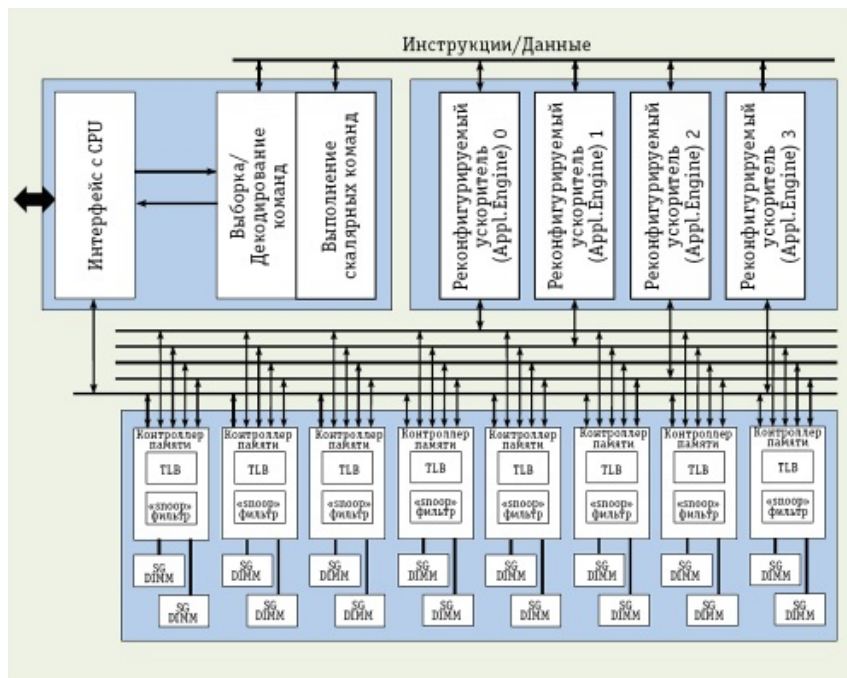


Рис. 1. Блок-схема сопроцессора Convey HC-1ex

В Convey HC-1ex (рис. 1) реализована уникальная подсистема памяти с общей пропускной способностью 80 Гбайт/с. Подсистема памяти Convey HC-1ex состоит из восьми контроллеров, каждый из которых размещен на отдельной FPGA. Каждый контроллер соединен с двумя модулями памяти SG-DIMM производства Convey Computers. Модуль SG-DIMM состоит из 64 банков и встроенной FPGA Stratix-2. В отличие от обычных модулей DIMM, ориентированных на подкачку сразу восьми слов (одной кэш-строки), модули SG-DIMM поддерживают восемь независимых 8-байтовых каналов данных, что позволяет достигать максимальной пропускной способности памяти при обращении в одно 8-байтовое слово с шагом большим, чем восемь слов. Пропускная способность одного модуля SG-DIMM составляет 5 Гбайт/с. Подсистема памяти Convey HC-1ex позволяет использовать как обычные модули DIMM, так и специализированные SG-DIMM. В первом случае максимальный объем памяти составляет 128 Гбайт, во втором — 64 Гбайт.

Именно нестандартная подсистема памяти, приспособленная к нерегулярному доступу с низкой пространственной и временной локализацией, позволила Convey HC-1ex достигнуть показателя в 773 MTEPS, что неплохо для одного узла. Для сравнения заметим, что для такого же графа на четырехсокетном узле с 10-ядерными процессорами Intel Westmere-EX было получено 705 MTEPS. Это выглядит не в пользу Convey HC-1ex, однако следует принять во внимание, что четырехсокетный узел с Westmere-EX на сегодняшний день — это новейший SMP-узел с процессорами Intel Xeon, в то время как Convey HC-1ex появился в 2010 году и в настоящий момент ведется разработка новой версии системы, использующей QPI (Quick Path Interconnect) в качестве интерфейса с центральным процессором.

SWARM

Система обеспечения параллельного выполнения программ SWARM предназначена для поддержки времени выполнения в больших гетерогенных вычислительных системах (runtime). SWARM разрабатывается компанией ET International, основанной профессором Гуанг Гао из Университета Делавэра. Во многом SWARM есть реинкарнация проекта EARTH, разрабатывавшегося в Университете Делавэра под руководством Гао в первой половине 90-х, адаптированная к суперкомпьютерам с гетерогенной архитектурой на базе многоядерных процессоров и различного рода ускорителей (GPGPU, FPGA, MIC). Целью создателей SWARM было создание унифицированной runtime-системы для высокопроизводительных систем экзафлопсного уровня, разрабатываемых в рамках программы DARPA UNPC.

В основе SWARM лежит двухуровневая мультитредовая вычислительная модель (рис. 2) с элементами управления потоком данных. В вычислительной модели SWARM выделяются два уровня параллельных программных потоков: процедуры и треды (codelet в терминах SWARM). Процедура — набор асинхронно выполняющихся тредов, разделяющих контекст процедуры. В каждой процедуре должен быть объявлен один тред, автоматически запускаемый при вызове процедуры.

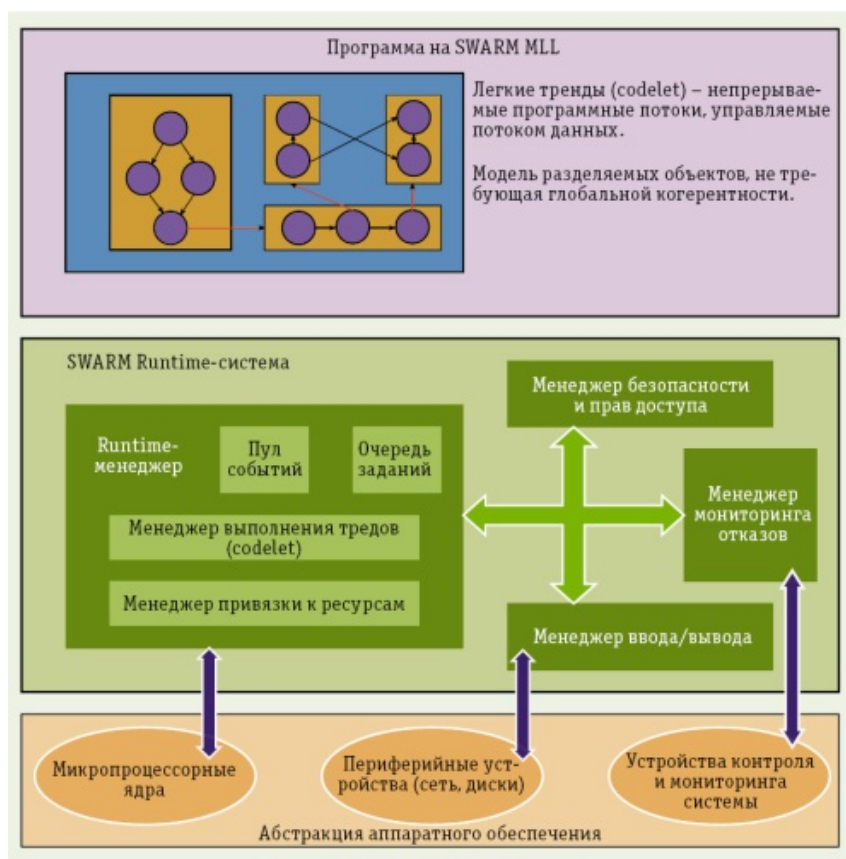


Рис. 2. Вычислительная модель SWARM

Планирование выполнения тредов осуществляется runtime-системой, при этом треды могут быть запущены только после того, как все зависимости по данным и ограничения будут разрешены. Запущенные треды не прерываются до завершения их выполнения. Также треды (в отличие от процедур) не могут возвращать значения. Коммуникации между тредями осуществляются с помощью асинхронных операций put/get.

Программа для SWARM представляется на языке среднего уровня SWARM MLL. Предполагается, что в дальнейшем будут написаны трансляторы с параллельных языков высокого уровня, таких как UPC, X10, Chapel в SWARM MLL. Программа на SWARM MLL транслируется в язык Си с вызовами runtime-системы SWARM посредством транслятора swarm2c.

В списке Graph500 6-е, 8-е и 9-е места принадлежат кластерным системам, на которых запуск теста осуществлялся с использованием SWARM. Для кластера Intel Endeavor ускорение, полученное на SWARM относительно реализации на MPI, составило 11 раз, это подтверждает, что мультитредово-поточковая вычислительная модель даже с учетом программной реализации позволяет повысить эффективность использования вычислительных ресурсов. Дальнейшее развитие SWARM предполагает реализацию системы на новейших многоядерных процессорах типа Intel Knight-Ferry и Adapteva Epihany и суперкомпьютеров на их основе.

J7

В ОАО «НИЦЭВТ» с 2005 года ведутся исследования по суперкомпьютерам стратегического назначения (СКЧН) на базе мультитредово-поточкового микропроцессора J7. За это время были разработаны архитектура суперкомпьютера, имитационная модель, инструментальные средства программирования, включая два компилятора Си.

В СКЧН «Ангара» в максимальной конфигурации из 32 768 узлов доступно 8 Пбайт глобально адресуемой памяти, причем вся эта память доступна для виртуальной адресации из одного узла без промахов в TLB. Организация памяти — сегментно-страничная с разделением вопросов отображения виртуального сегмента на памяти узлов и управления размещением данных на каждом из узлов. При этом возможно гибкое управление отображением виртуального сегмента на памяти узлов при помощи блочного, блочно-циклического распределения и двухуровневого скремблирования.

Основой вычислительного узла СКЧН является многоядерный мультитредово-поточковый процессор J7 (два ядра по 64 тредовых устройства, 500 МГц) с аппаратной поддержкой трансляции адресов глобально адресуемой памяти и передачи реализующих такие обращения коротких системных пакетов. Каждое тредовое

устройство микропроцессора J7 содержит регистры управления выполнением потока команд загруженного на него треда, наборы 64-разрядных архитектурных регистров для хранения чисел с фиксированной и плавающей запятой, адресуемых однобитовых регистров-признаков и регистров адресов передачи управления.

Мультитредовое ядро содержит четыре конвейера, по 16 тредовых устройств в каждом. Каждый конвейер выдает с какого-либо тредового устройства одну команду за такт, причем с одного тредового устройства это происходит не чаще чем раз в четыре такта. Информационная зависимость выполняемых команд с одного тредового устройства контролируется аппаратно посредством таблиц признаков занятости регистров. Переключение выполнения команд с одного тредового устройства на выполнение команд с другого происходит без перезагрузки регистров, за один такт процессора. Каждый конвейер имеет собственное устройство для выполнения целочисленных и логических операций, каждая пара конвейеров разделяет устройство выполнения операций с плавающей запятой.

Данные по тесту Graph500 были получены при помощи моделирования процессора J7 и кристалла маршрутизатора коммуникационной сети на 4D-тор, также разрабатываемой в НИЦЭВТ. Для кристалла маршрутизатора сети планируется использовать интерфейс PCIe gen2 x16 (80 Гбит/с) и каналы с пропускной способностью 75 Гбит/с. В реализации на J7 для сетевых взаимодействий используется удаленный вызов процедуры с передачей коротких пакетов. Производительность может быть поднята еще выше посредством агрегации и длинных сетевых пакетов.

Для отработки схемотехнических решений, принятых при проектировании микроархитектуры J7, было запланировано создание серии макетов на FPGA с поэтапным введением блоков процессора в FPGA. В настоящее время разработан макет J7-M2 для отработки конвейера команд и ведется работа над созданием макета J7-M3 для отработки подсистемы памяти. Макет представляет собой плату расширения с Xilinx Virtex XC5VFX100T, подключенную через PCIe к хост-машине. Макет содержит ядро J7 в конфигурации: тактовая частота ядра — 25 МГц; кэш команд (ICACHE) — 16 Кбайт, поддерживается только режим физической адресации; конвейер команд (блоки FETCH и ISSUE) с 16 тредовыми устройствами, у каждого из которых тридцать два 64-разрядных регистра общего назначения, восемь однобитовых регистров-признаков, восемь 34-разрядных регистров адресов команд; блок целочисленной арифметики ALU над 64-разрядными числами; блок устройства умножения-деления MDU (умножение 64-разрядных целых со знаком и без знака); блок запросов в память LSU (обрабатывает скалярные команды записи/чтения, поддерживается до двух незавершенных команд от одного треда); блок BUI для взаимодействия с управляющей системой (хост-системой), в том числе для доступа к оперативной памяти хост-системы через PCIe.

Для управления макетом разработан драйвер для ОС Linux. Обращения к памяти, выдаваемые блоком LSU, передаются в блок BUI, который передает их дальше через PCIe контроллеру памяти хост-системы. Таким образом эмулируется работа с локальной памятью мультитредового процессора. Макет J7-M2 позволил верифицировать скалярные команды обращения в память, команды целочисленной арифметики, команды переходов, команды управления тредовыми устройствами.

Следующую версию макета J7-M3 планируется реализовать с использованием Xilinx Virtex6 XC6VLX240T и 512 Мбайт памяти DDR3 (рис. 3). По сравнению с предыдущим пакетом в J7-M3 будут реализованы следующие дополнительные функциональные возможности:

- блок запросов в память LSU: векторные команды записи/чтения, поддержка до восьми незавершенных команд от одного треда;
- блок трансляции виртуальных адресов MMU: поддержка виртуальных сегментов, кэш дескрипторов сегментов;
- блок интерфейса работы с памятью EDI: наборно-ассоциативный кэш данных L2 — 4 Мбайт, длина кэш-строки 64 байт, поддержка атомарных операций;
- блок BUI: взаимодействие с драйвером макета, загрузка теста в память макета, получение результата выполнения теста.

Дальнейшее развитие макетов предполагает разработку макета, в котором будет обрабатываться взаимодействие мультитредового ядра с маршрутизатором коммуникационной сети.

Развитие архитектуры J7

Микропроцессор J7 для целей исследования предполагается встроить в виде коммуникационного сопроцессора в кристалл маршрутизатора сети 4D-тор. При этом основные архитектурные принципы — мультитредовость и глобально адресуемая память — будут сохранены, вместе с тем будет внесен ряд изменений в сторону упрощения конвейера и поддержки функций сопроцессора.

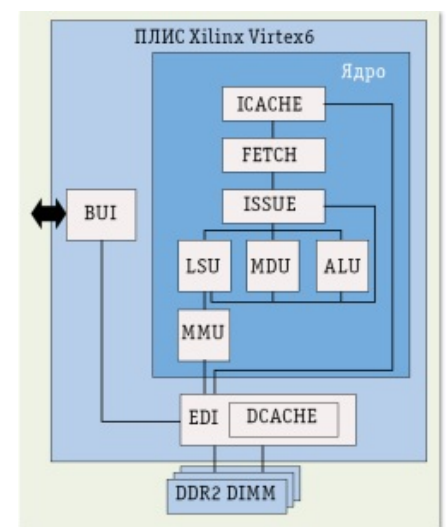


Рис. 3. Структура макета J7-M3

На рис. 4 представлена блок-схема контроллера коммуникационной сети 4D-тор с встроенным сопроцессором J7-C. Коммуникационный сопроцессор J7-C состоит из мультитредового ядра и контроллера памяти EDI, отвечающего за работу с локальной памятью DRAM.

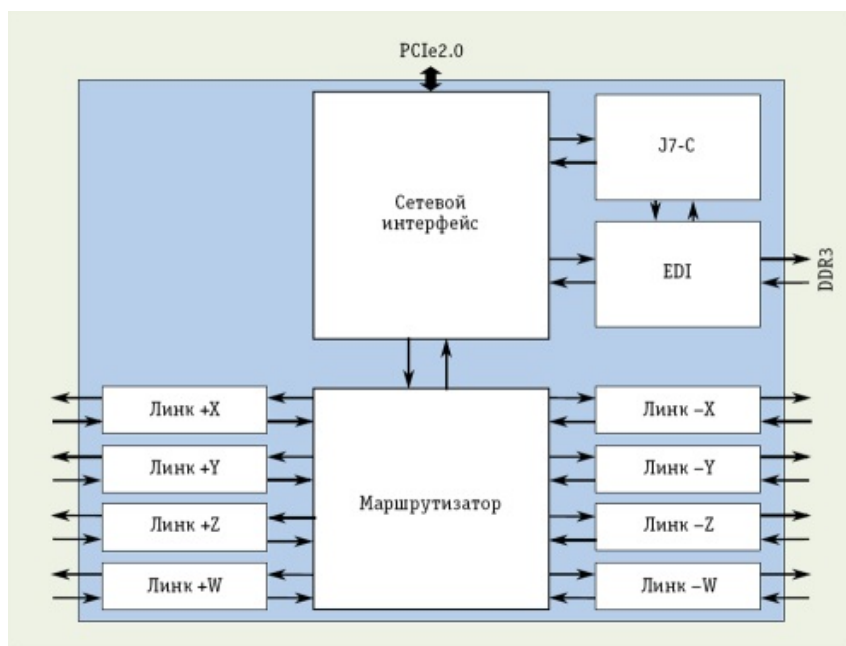


Рис. 4. Блок-схема контроллера сети 4D-тор с встроенным коммуникационным сопроцессором J7-C

В блоке EDI находится наборно-ассоциативный кэш данных L2 размером 4 Мбайт и контроллер памяти DDR3. В отличие от J7, в J7-C нет поддержки битов f/e, которыми в J7 снабжена каждая 64-разрядная ячейка памяти для эффективной синхронизации тредов, что значительно упрощает структуру блока, однако остается поддержка широкого набора атомарных операций.

Ядро J7-C состоит из четырех конвейеров команд, каждый из которых поддерживает одновременное выполнение до 16 тредовых устройств. Каждый конвейер оснащен собственным ALU для выполнения сложения, вычитания и логических операций. Также конвейеры разделяют два устройства MDU для деления и умножения целых чисел, одно MDU на два конвейера. В первых вариантах J7-C не поддерживаются операции с плавающей запятой ввиду отсутствия их необходимости для большинства графовых задач, что делает ядро более компактным с точки зрения места на кристалле и снижает энергопотребление.

В мультитредовое ядро J7-C введен блок Control Unit для управления работой тредов из центрального процессора. Основные функции Control Unit — загрузка программы и данных в локальную память J7-C из памяти центрального процессора, запуск и останов тредовых устройств, диагностика состояния J7-C. Блок Control Unit обменивается пакетами с центральным процессором через интерфейс PCIe, для чего в блоке PCIe выделяется отдельный базовый адресный регистр (BAR).

Работа с памятью осуществляется посредством команд записи, чтения и атомарных операций. Поддерживаются скалярные и векторные (до восьми 64-разрядных слов) обращения к памяти. Размер слова — 64 разряда. Адресация данных осуществляется с точностью до одного слова. Поддерживается глобальное виртуальное адресное пространство с сегментной организацией.

Тредовые устройства всегда работают на одном уровне привилегий — в пространстве пользователя. Управление тредовыми устройствами возможно только через Control Unit, то есть исключены команды резервирования и создания тредов. При этом поддерживается механизм удаленного запуска тредовых устройств для реализации активных сообщений.

Принципиальное отличие графовых задач от традиционных приложений способствует появлению новых вычислительных моделей и парадигм программирования. Программные технологии типа SWARM ориентируются на повышение утилизации аппаратных ресурсов системы за счет асинхронности и высокого параллелизма программы, управления потоком данных и агрегации асинхронных коммуникационных операций.

Аппаратные решения в первую очередь нацелены на достижение высокой реальной пропускной способности

памяти и сети. В Cray XMT и в проекте «Ангара» для этого используется мультитредовый процессор, в Convey HC-1ex — многоканальная подсистема памяти, оптимизированная для случайного доступа с грануляцией в одно слово. Большое количество потоков в графическом процессоре также используют большую пропускную способность памяти, причем даже на традиционно неудобных для графических систем графовых задачах начали получать высокие результаты [4]. Однако Cray XMT и «Ангара» имеют качественное отличие от GPGPU и Convey HC-1ex — поддержка глобально адресуемой памяти, что повышает продуктивность программирования и позволяет работать с большими объемами данных.

Умеренные результаты Cray XMT в списке Graph500 не должны вводить в заблуждение, поиск вширь в графе, лежащий в основе Graph500, не в полной мере соответствует реальным графовым задачам. В настоящее время создатели Graph500 работают над добавлением новых ядер, учитывающих природу реальных графовых задач.

Одним из возможных направлений дальнейшего развития отечественных мультитредовых процессоров является разработка коммуникационного сопроцессора J7-C, встроенного в кристалл контроллера коммуникационной сети. В данный момент осуществляется проработка этого решения с учетом планов по реализации СБИС маршрутизатора коммуникационной сети «Ангара».

Литература

1. Feo J., Harper D., Kahan S., Konecny P. Eldorado. Proceedings of the 2nd Conference on Computing Frontiers. New York, NY, USA: ACM, 2005.
2. Jason Bakos, High-Performance Heterogeneous Computing with the Convey HC-1. Computing in Science and Engineering, 2010, Vol. 12, No. 6.
3. Семенов А. С., Соколов А. А., Эйсымонт Л. К. Архитектура глобально адресуемой памяти мультитредово-потокового суперкомпьютера. //Электроника: Наука, Технология, Бизнес. – 2009. – №1.
4. Hong S., Oguntebi T., Olukotun K. Efficient Parallel Graph Exploration on Multi-Core CPU and GPU//Proceedings of the PACT '11: 20th International Conference on Parallel Architectures and Compilation Techniques, October 2011.

Александр Фролов (frolov@nicevt.ru), **Александр Семенов** (semenov@nicevt.ru), **Денис Мошкин** (den_m@nicevt.ru), **Владимир Кабыкин** (vk@nicevt.ru), **Анатолий Никитин** — сотрудники ОАО «НИЦЭВТ» (Москва). Авторы благодарят **Леонида Эйсымонта**, заместителя главного конструктора по архитектуре и программному обеспечению суперкомпьютера стратегического назначения «Ангара» в период 2005-2010 гг.: подходы в данной статье базируются на идеях, заложенных в проекте «Ангара».

Постоянный URL статьи: <http://www.osp.ru/os/2011/07/13010498/>

© «Открытые системы», 1992-2011. Все права защищены.