

Стратегические ИТ: китайский сюрприз №863

Дмитрий Волков

Пять лет назад Китай приступил к реализации проекта, сопоставимого с аналогичными американскими и японскими суперкомпьютерными программами. Каковы шансы Поднебесной оставить всех позади в области стратегических ИТ?



В рамках реализуемой министерством науки и технологий национальной программы 863, начатой еще в марте 1986 года (отсюда и название), руководство КНР приняло в 2006 году пятилетний проект **China National Initiative of High Productivity Computer and Grid Service Environment** с бюджетом около 400 млн долл. [1,2]. Цель этого проекта 863/ИТ – развитие специальных, или заказных, стратегических ИТ и создание на их основе перспективных стационарных и бортовых суперкомпьютеров, предназначенных для обеспечения национальной

безопасности и решения важнейших научно-технических задач, стоящих перед государством. На начальных этапах проекта допускалось применение новейших зарубежных технологий, но основной замысел – добиться полной технологической независимости в области стратегических ИТ.

Одним из результатов проекта 863/ИТ стало создание суперкластера «Млечный путь-1» (Tianhe-1) петафлопного уровня производительности и двух суперкомпьютеров такого же типа с производительностью около 100 TFLOPS. Наряду с применением импортных коммерчески доступных микропроцессоров (что типично для многих, в том числе и российских разработчиков кластерных систем), в Китае не забыли и про собственные микропроцессоры Loongson (другое название Godson). Но все это лишь вершина айсберга – в рамках проекта 863/ИТ была развернута разработка высокопродуктивного гиперпараллельного компьютера с глобально адресуемой памятью, в котором должна быть, в частности, решена проблема «стены памяти». Сегодня именно такие цели ставят перед собой правительства США и Японии в программах создания суперкомпьютеров стратегического назначения (СКСН).

Китайская платформа

Бытует мнение, что в Китае, как и в других странах, строятся лишь кластерные суперкомпьютеры и копируются устаревшие по архитектуре американские микропроцессоры. Такое представление слишком упрощено, но даже если рассматривать только работы по кластерам и копируемым микропроцессорам, то можно обнаружить ряд весьма перспективных решений и высокий профессионализм местных разработчиков.

Суперкомпьютер Tianhe-1, созданный в Национальном университете оборонных технологий, содержит 6144 микропроцессора Intel Xeon E54505540, 5120 графических процессоров AMDATI Radeon HD 4870 и имеет память 98 Тбайт. Вроде бы, ничего особенного, однако даже в этой разработке имеются интересные находки.

Xeon E5450/3 ГГц имеет четыре ядра, кэш-память второго уровня 12 Мбайт, но серверная плата с этими микропроцессорами построена по обычной схеме с общей физической памятью, доступной через общую шину. Это оборудование эффективно при решении задач с хорошей пространственно-временной локализацией обращений к памяти. Xeon E5540/2,5 ГГц также имеет четыре ядра, но в каждом поддерживаются по два SMT-треда, и, по сути, это восьмитредовый микропроцессор. Он еще имеет межкристальные каналы QuickPath, что позволяет на плате реализовать сеть и распределенную по сокетам

процессоров физическую память, доступную через логически единое адресное пространство. Однако это более медленный микропроцессор, чем E5450, да и кэш второго уровня у него меньше – 8 Мбайт. Зачем же была выбрана именно эта многосокетная плата?

Оказывается, она обладает свойством толерантности к задержкам обращений к памяти – эффективность работы приложения при определенных условиях на ней определяется уже не задержками, а темпом выполнения операций с памятью, за счет чего при распараллеливании растет реальная производительность даже для задач с нерегулярным доступом к памяти, например таких, как: расчеты на нерегулярных адаптивных сетках, боевые и гражданские информационно-управляющие комплексы, системы предотвращения террористических операций, обработка научной и разведывательной информации.

Свойство толерантности многосокетных плат и возможность подключения к ним графических ускорителей позволяет, в свою очередь, эффективно реализовать DAE-модель организации вычислений, когда в программе между асинхронными параллельными процессами разделяется доступ к данным и вычисления. При таком подходе E5540 может успешно выполнять работу с данными (передачу их в графическую память и обратно), а графические ускорители – выполнять роль вычислителя над удобно расположенными в своей быстрой памяти данными. Такая идея параллельных гетерогенных систем сегодня просматривается во многих перспективных суперкомпьютерах. Более того, кроме DAE-модели, толерантность позволяет на множестве многосокетных плат эмулировать работу с глобально адресуемой памятью, для чего можно применить библиотеки доступа к памяти удаленных узлов ARMCi (Aggregate Remote Memory Copy Interface) или GASNet (Global-Address Space Networking). Можно пойти дальше – начать опытное использование PGAS-языков (Partitioned global address space) UPC (Unified Parallel C) и CAF (Co-Array Fortran).

Итак, даже такой вроде бы небольшой нюанс в выборе микропроцессоров для Tianhe-1 открывает возможности для решения задач с нерегулярным доступом к памяти (Data-Intensive Computing, DIC), требующих использования больших объемов общей памяти, для повышения продуктивности разработки параллельных программ за счет применения специальных библиотек и PGAS-языков.

Микропроцессор Godson-1 появился в 2002 году в Институте компьютерных технологий академии наук Китая. Это 32-разрядный микропроцессор MIPS, но с собственной суперскалярной микроархитектурой. Его характеристики: тактовая частота 200-266 МГц, технология 180 нм, производительность на SPECintfp2000 – 19/25.

Первый микропроцессор семейства Godson-2 появился в 2003 году, и до 2007 года вышло еще четыре его 64-разрядных представителя: 2B, 2C, STLS2E и STLS2F. Последние два выпущены по технологии 90 нм компанией STMicroelectronics, соглашение о сотрудничестве с которой было заключено в 2004 году. В 2006 году объявлено о готовности к массовому выпуску модели STLS2E (Godson 2E): суперскалярная микроархитектура – четыре команды за такт; два функциональных устройства операций над числами с плавающей запятой (одно из них для SIMD-команд), два устройства целочисленной арифметики и одно устройство выполнения обращений к памяти; 47 млн транзисторов; тактовая частота 0,7-1 ГГц; рекордно низкая в индустрии потребляемая энергия – 4 Вт (для сравнения, IBM Cell BE, SPE потребляют свыше 110 Вт, Intel Xeon 7400 – 50-130 Вт); размер кэша первого уровня команд и данных – 64 Кбайт, а кэш второго уровня – 512 Кбайт; производительность на SPECintfp2000 – 503/503.

Первый микропроцессор семейства Godson-3 (2008 год) имел уже четыре ядра, а в 2009 году последовал восьмиядерный микропроцессор. Оба созданы по технологии 65 нм, имеют тактовую частоту 1-1,2 ГГц, 400 и 600 млн транзисторов, потребляемая энергия – 10 и 20 Вт. В этих изделиях введено дополнительно более 200 команд для аппаратно-программной эмуляции архитектуры x86. Имеется встроенный контроллер DDR2-памяти

и ведутся разработки вариантов с разнородными ядрами (графические ускорители, специальные функции), варианта введения в ядра мультитредовости, добавления встроенного сетевого интерфейса.

Микропроцессоры Godson применяются в персональных компьютерах и ноутбуках, а также суперкомпьютерах петафлопной производительности – узлы с микропроцессорами Godson будут и в суперкомпьютере петафлопного уровня Dawning 6000A.

Таким образом, за очень короткий срок Китай создал плацдарм для развития своих стратегических ИТ – вышел на мировой уровень в разработке микропроцессоров, подключив к проекту около 300 специалистов и эффективно используя возможности работы с западноевропейскими партнерами. Развитие в сторону увеличения количества ядер и их специализации, введения мультитредовости, улучшения внешних интерфейсов – достаточно общие тенденции, однако низкое потребление энергии очень выгодно отличает Godson (сравнимые показатели имеют лишь специализированные микропроцессоры ARM Cortex-A9 MPCore и Tiler TILE 64). Это может быть использовано для создания на его базе мультиядерных конфигураций (32-64 ядра) методами 3D-сборки, что будет уникальным достижением.

Год 2006: перелом

Суперкомпьютер Tianhe-1 и микропроцессор Godson можно отнести к области обычных ИТ, они важны, но их аналоги общедоступны и ограничены по возможностям, поэтому не обеспечивают стратегического превосходства в ИТ над другими странами. Начиная с 2006 года в Китае стали развиваться специальные стратегические ИТ, позволяющие получить реальное преимущество как в научно-технической области, так и в обеспечении национальной безопасности – запущен проект 863/ИТ. В работе [2], подготовленной в Национальном университете оборонных технологий, была оценена общая ситуация в мире, научный потенциал страны в ИТ, очерчены цели и задачи проекта 863/ИТ, явно противопоставляемого аналогичным проектам США по программе DARPA HPCS, предусматривающей создание к 2010 году перспективных высокопродуктивных СКСН с реальной производительностью 1 PFLOPS, причем развиваемой на широком классе задач, а не только на специфическом Linpack, высокие показатели на котором позволяют попасть в престижный рейтинг Top 500. В работе [2] для противопоставления упоминается и японский проект NGSP создания к 2012 году СКСН с пиковой производительностью 10 PFLOPS на базе коммерческих микропроцессоров и быстрой заказной сети, но по глубине замысла и новаторству этот проект явно уступает американскому и китайскому.

До недавнего времени процесс разработки суперкомпьютеров во всем мире напоминал гонку по созданию конкретных образцов высокопроизводительных систем с преодолением очередных уровней производительности, однако в современных условиях этого мало – сегодня преимущество имеет интеграция национальных стратегических ИТ, инфраструктуры и суперкомпьютерной индустрии, включающих науку и промышленность.

В 1976 году векторный суперкомпьютер Cray-1 первым преодолел уровень производительности 100 MFLOPS, а уже в начале 80-х годов был создан китайский суперкомпьютер Galaxy-1 с той же производительностью. После появления суперкомпьютера с разделяемой общей памятью Cray-XMP в Китае в конце 80-х годов был создан аналогичный суперкомпьютер Galaxy-2. Последовавшая с наступлением 90-х годов эра массивно-параллельных полужаказных суперкомпьютеров на базе коммерчески доступных микропроцессоров началась с создания решений, содержащих до 1 тыс. процессоров, и системы такого типа были созданы в Китае: Dawning 1000, SHENWEI I, Galaxy-3.

В США терафлопный барьер был преодолен в 1996 году на системе ASCI RED, а в 2000 году был создан китайский массово-параллельный суперкомпьютер с аналогичной

производительностью. Затем ИТ-индустрия увлеклась кластерными конфигурациями на базе коммерчески доступных микропроцессоров, коммуникационных сетей и системного программного обеспечения. Знаковым достижением этого периода стало создание в июне 2006 года монстра от компании Dell – суперкомпьютера ThunderBird, который содержал до 9 тыс. процессоров и показал на тесте Linpack производительность 39 TFLOPS. Работы по этому направлению были подхвачены и китайскими фирмами Dawning Information Industry, Lenova и др., важным событием стало создание еще в 2004 году системы Dawning 10 TFLOPS, которая оказалась тогда на десятом месте в списке Top 500.

Гонка за пиковую производительность на тесте Linpack закончилась в 2003 году для США драматично – задуматься о новой системе ценностей и концепциях стратегических ИТ, нацеленных на поддержку национальных интересов, а не бизнеса конкретных компаний, заставило появление созданного по заказным технологиям японского векторного суперкомпьютера Earth Simulator. Этот суперкомпьютер превосходил американские кластерные монстры на тесте Linpack, но еще в большей степени он превосходил их при решении стратегически важных задач, именно это шокировало американскую ИТ-общественность и истеблишмент.

Создание Earth Simulator, во-первых, ознаменовало начало эры возрождения специальных стратегических ИТ – заказные технологии вновь стали востребованными и спустя три года появился американский массово-параллельный заказной суперкомпьютер IBM BlueGene/L с производительностью 280 TFLOPS. Во-вторых, изменились способы оценки качества суперкомпьютеров, вместо теста Linpack перешли к применению комплексного набора тестов HPC Challenge и другим методикам (например, GUPS – Giga Updates Per Second), в которых использовалась многопараметрическая оценка, при этом особое внимание уделялось исследованию поведения суперкомпьютера в разных режимах пространственно-временной локализации обращений к памяти, особенно в режимах плохой локализации.

Общий вывод китайских специалистов состоит в том, что вызовом на этот раз стало не создание конкретного суперкомпьютера с очередной рекордной производительностью, а формирование новых концепций разработки перспективных СКСН с применением специальных стратегических ИТ, для появления которых нужны фундаментальные исследования. Разработка СКСН вновь стала наукой, и в этом основная причина появления проекта 863/ИТ.

В чем стратегия и тактика проекта 863/ИТ? Авторы проекта выделяют шесть основных достижений, составляющих платформу проекта: закон Густафсона о теоретической масштабируемости производительности параллельных компьютеров; создание библиотеки MPI; разработки в области микропроцессоров общего назначения; опыт применения кластерных технологий, позволивший расширить круг пользователей высокопроизводительных средств; создание средств визуализации результатов вычислений; технологии построения глобальных сетей и grid-технологии.

Однако эти достижения – явления недавнего прошлого, а будущее, по мнению авторов проекта, будет определено другими факторами:

- создание больших заказных вычислительных систем с параллельной архитектурой на базе оригинальных вариантов всех компонентов, от функциональных устройств процессора до межузловой сети;
- разработка разных вариантов оригинальных масштабируемых операционных систем, в том числе и параллельных, которые для пользователя выглядят как единое целое (single system image);
- разработка распараллеливающих и высокооптимизирующих компиляторов, эффективно использующих методы машинно-зависимой оптимизации программ;
- разработка оригинальных средств отладки и профилирования параллельных программ;
- работы в области grid.

Кроме программы 863 в Китае действует еще программа фундаментальных исследований 973, ведутся программы развития ключевых технологий, развития информационной инфраструктуры, мегапроекты, курируемые Министерством науки и технологий КНР, Национальным университетом оборонных технологий и Институтом компьютерных технологий. Ставится задача вовлечения в эти работы как можно большего числа исследователей и разработчиков. Кстати, до наступления эры микропроцессоров в области стратегических ИТ в США работало около 50 исследовательских коллективов, но потом их количество резко сократилось. Затем Министерству энергетики США было предписано активизировать исследования и разработки по стратегическим ИТ в университетской среде, результаты таких действий в настоящее время уже известны.

Еще одной особенностью китайского перелома 2006 года является активное использование возможностей международного сотрудничества в области ИТ с США, Европой, Японией и странами Юго-Восточной Азии:

- предоставление на своей территории льготных условий работы филиалов зарубежных ИТ-компаний (Intel, Cray, NEC, Fujitsu и др.);
- обмен специалистами, участие китайских ученых в проектах ведущих американских центров (IBM Cyclops 64, eXMT PRAM, CASS-MT);
- финансирование совместных проектов (STMicroelectronics и т.п.);
- организация международных конференций на территории Китая в области важнейших направлений ИТ и прикладных проблем (International Conference on Theory and Applications of Satisfiability Testing, графовые базы данных).

Первые результаты

Направления работ проекта 863/ИТ охватывают создание заказных процессоров и коммуникационных сетей, а также системного программного обеспечения.

Исторически сложилось так, что большее внимание всегда уделялось оптимизации выполнения операций над числами в формате с плавающей запятой, особенно это касалось вычислений по обработке сигналов в бортовых суперкомпьютерах. По этой причине популярны работы по потоковым (stream based) суперкомпьютерам, ориентированным на модели вычислений, представимые статическими графами потоков данных. Потоковые архитектуры – это один из методов решения проблемы «стены памяти», поскольку при вычислениях данные обычно обмениваются через регистровые ресурсы или быстрые блоки небольшой памяти, минуя общую память с медленным доступом. Высокие требования к производительности потоковых суперкомпьютеров выдвигаются прежде всего создателями радаров космического базирования, предназначенных для контроля за поверхностью Земли, а также разработчиками авиационных беспилотных средств. Неудивительно, что значительное число публикаций по работам, проводимым в Китае, связано именно с этим направлением, а точнее, с процессорами FT64 [3] и MASA (Multiple-morphs Adaptive Stream Architecture) [4], которые также рассматриваются и как процессоры для вычислительно емких научных приложений. Потоковые процессоры FT64 и MASA иллюстрируют возможности специальных стратегических ИТ при создании бортовых суперкомпьютеров и суперкомпьютеров для вычислительно емких алгоритмов, однако имеется и другой способ преодоления стены памяти – посредством искусственного обеспечения толерантности процессоров к задержкам.

Китайские работы по массово-мультитредовым микропроцессорам пока не встречались среди доступных источников, что странно, поскольку для Китая характерна систематизация и полнота охвата в разрабатываемой области, какой бы она ни была. Это тем более удивительно при наличии китайских работ по приложениям, для которых наиболее эффективно применение массово-мультитредовых процессоров: решение SAT-проблемы с использованием графового алгоритма Survey Propagation и графовые базы

нерегулярных данных, решение задач на графах в области биологии и нанотехнологий, управления сетевыми структурами, социологии. Кроме того, разработчики из КНР наверняка знают об успешных попытках решения таких задач на Cray XMT, а также о направлениях работ Центра CASS-MT по освоению такой мультитредовой техники и созданию для нее программного обеспечения.

Технологии создания адаптивных сетей с топологией N-тор в Китае хорошо освоены, о чем говорит хотя бы факт разработки внутриплатной сети в FT64. По-видимому, в перспективных системах ставка будет сделана на многосвязные реконфигурируемые сети с оптическими терабитными оптическими WDM-линками, в которых передача сообщений производится одновременно на разных длинах волн. Здесь следует ожидать расширения сотрудничества китайских ученых с японскими специалистами.

Уникальным событием можно считать создание в рамках программы 863 микроядерной распределенной операционной системы Kylin, призванной обеспечить надежную защиту компьютерной инфраструктуры на уровне всей страны от атак извне. Как отметил в докладе Конгрессу США Джон Колеман, ведущий американский эксперт в области информационной безопасности, средства кибернаступления, разрабатываемые американскими программистами для сетей компьютеров на платформах Linux, Unix и Windows, могут оказаться бесполезны против ОС Kylin. Комментируя инициативы КНР по подготовке к кибервойнам, Колеман отметил высокий уровень китайских специалистов: «Эта страна играет в шахматы, в то время как США разыгрывают партию в шашки».

Внедрение собственной ОС Kylin в правительственных организациях и военных ведомствах КНР началось еще в 2007 году, а сама программная платформа находится в разработке с 2001 года. ОС Kylin окажется эффективна в сочетании с новым китайским процессором, который будет безопаснее и надежнее, чем зарубежные аналоги. Некоторые специалисты считают, что этим микропроцессором может быть предполагаемый китайский массивно-мультитредовый микропроцессор, обеспечивающий аппаратную поддержку работы с огромной глобально адресуемой физической памятью и мощную многоуровневую аппаратную защиту программ и данных.

Второе десятилетие XXI века

Согласно [1,2], следующим шагом КНР будет, возможно, гетерогенный, гиперпараллельный суперкомпьютер на базе массивно-мультитредовых и потоковых микропроцессоров собственного производства, толерантный к задержкам выполнения операций с огромной глобально адресуемой пространственно распределенной памятью.

Сегодня исследователи из развитых стран приступили к концептуальной проработке систем экзафлопного уровня, концентрируясь на основных проблемах [5]. Характерно, что при этом явно осознается тот факт, что путем увеличения количества однородных ядер в коммерчески доступных микропроцессорах экзафлопный барьер преодолеть не удастся, поскольку еще сильнее обостряются проблемы «стены памяти», плохой масштабируемости и чрезмерного энергопотребления. Если двигаться по проторенному пути использования коммерческих микропроцессоров с увеличивающимся количеством ядер, то экзафлопная система-монстр будет потреблять около 200 МВт, что практически неприемлемо.

Надо полагать, что, как и зарубежные коллеги, из множества вариантов создания экзафлопных систем китайские специалисты выберут два взаимодополняющих решения: многоядерные гетерогенные мультитредово-потоковые архитектуры и 3D-сборку (трехмерные СБИС) [6]. Такие стратегические ИТ, по-видимому, и позволят создать экзафлопную китайскую СКСН по проекту 863/ИТ.

* * *

Системность государственной организации работ в области стратегических ИТ на базе серьезных экспертных проработок, высокий профессионализм специалистов,

продемонстрированные темпы роста и масштабность позволяют предположить, что к середине следующего десятилетия КНР вполне может стать лидером в этой области.

Литература

1. D.K.Karper, N.Hirose, D.Chen, Asian HPC Update (Japan, China, India), April 2008.
2. Yang X.J., Dou Y., Hu Q.F. Progress and Challenges in High Performance Computer Technology, Journal of Computer Science & Technology, Sept.2006, Vol.21, N5.
3. Wen M., Wu N., Zhang C., Wu.W., Yang Q., Xun C. FT64: Scientific Computing with Streams. 14th IEEE HIPS , 2007.
4. Wu N., Yang Q., Wen M., He Y., Ren J., Guang M., Zhang C. Multiple Macro-Tile Stream Architecture, Workshop on SHCMP08 in conjunction with 35th ISCA, 2008.
5. Geist A. Paving the Roadmap to EXASCALE. SciDAG Review, 2010.
6. Huang W. et al. Interaction of Scaling Trends in Processor Architecture and Cooling. To be appear in Proceedings of the 26 Semi-Therm Symposium, 2010.

Дмитрий Волков (vlk@keldysh.ru) – старший научный сотрудник ИПМ им. М.В. Келдыша РАН (Москва).

Программа создания перспективных суперкомпьютеров

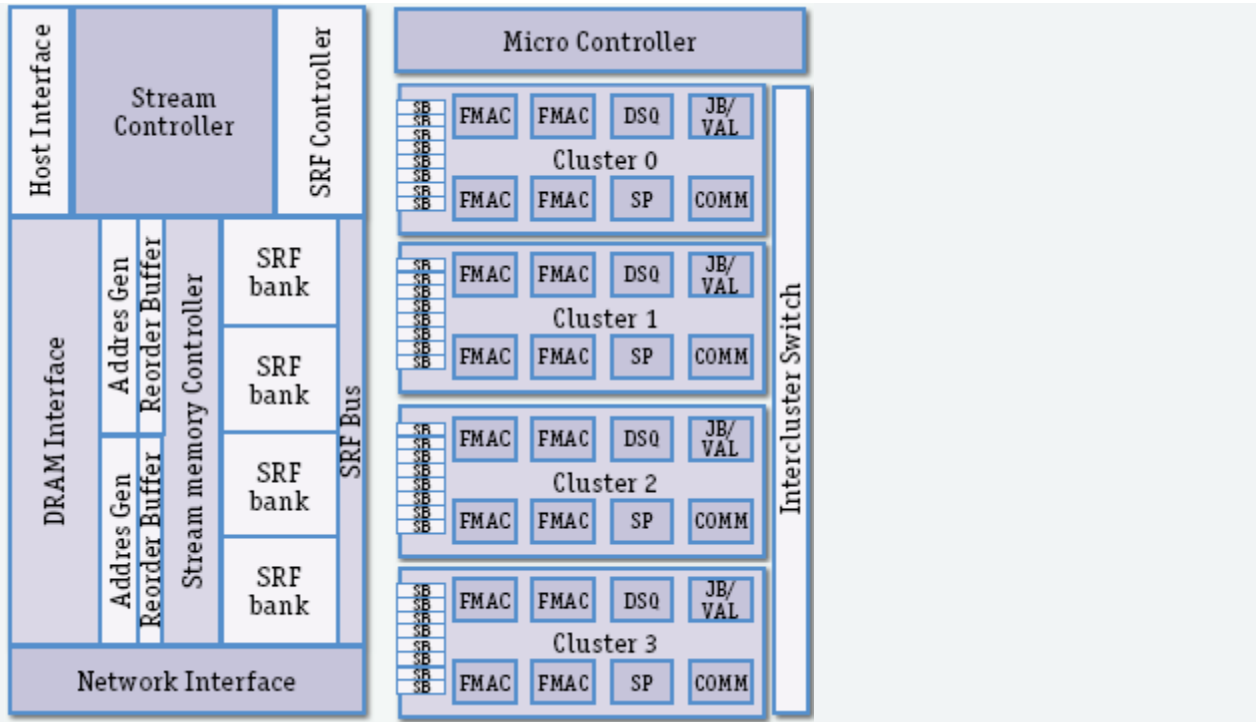
Министерство обороны США реализует программу создания суперкомпьютеров с перспективной архитектурой для решения стратегически важных государственных задач обеспечения национальной безопасности. Эта программа уже оказала влияние на отрасль и определила дальнейшее развитие индустрии суперкомпьютеров.

Суперкластеры – между прошлым и будущим

Успехи в области микропроцессоров и сетей, а также наличие разнообразного открытого программного обеспечения создали предпосылки для разработки высокопроизводительных систем на базе коммерчески доступных компонентов. К новому поколению таких компонентов относятся многосокетные платы, представляющие собой коммерчески приемлемые варианты мощных многопроцессорных узлов с логически общей и достаточно эффективной в разных режимах использования памятью большого объема.

Процессоры: made in China

Процессор FT64 (рис. 1) базируется на 32-разрядном процессоре Imagine Стэнфордского университета, предназначенном для мультимедийной обработки, но в отличие от него FT64 ориентирован на научные вычисления, работает с 64-разрядными данными и подключается к Itanium 2 как процессор-ускоритель (рис. 2).

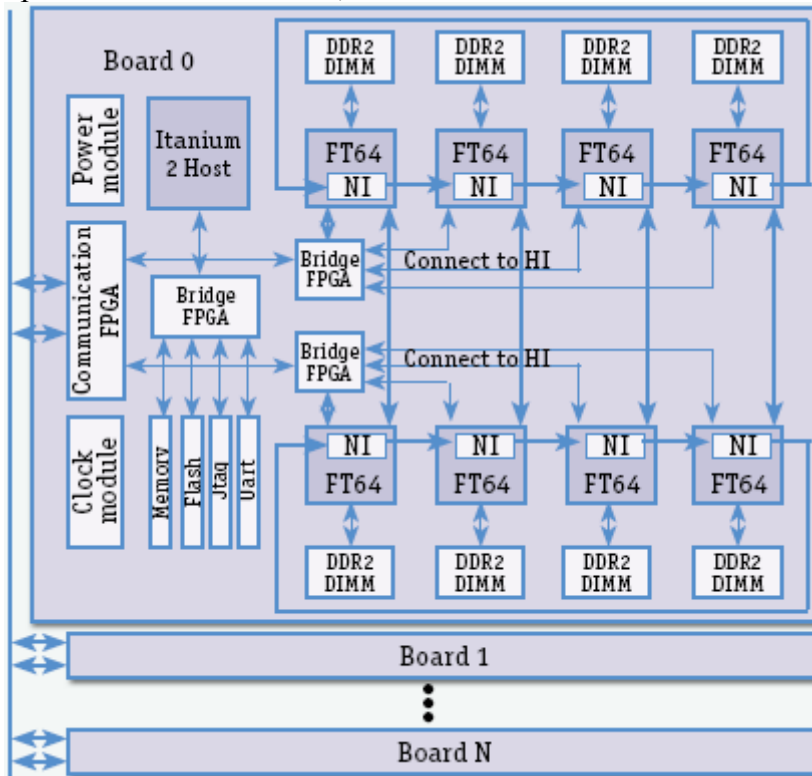


Host Interface – интерфейс с управляющим микропроцессором, в качестве него используется Itanium 2; DRAM Interface – интерфейс с внекристальной динамической памятью, у каждого FT64 своя такая внешняя память (см. DDR2 DIMM на рис. 2); Network Interface – интерфейс с внутрисетевой сетью, могут быть разные топологии; Address Gen – адресный генератор; Reorder Buffer – буфер переупорядочения; Stream memory Controller – контроллер потоковой памяти; SRF bank – банк файла потоковых регистров; SRF Bus – общая шина банков файла потоковых регистров; Stream Controller – контроллер потоков данных, получает указания от управляющего микропроцессора, передает их в контроллер потоковой памяти и микроконтроллер арифметических кластеров; SRF Controller – контроллер файла потоковых регистров; Cluster 0,1,2,3 – арифметические кластеры, содержат множество функциональных устройств и настраиваемый коммутатор для передачи данных между ними, управляются одной широкой командой, выдаваемой микроконтроллером; Micro Controller – микроконтроллер арифметических кластеров; Intercluster Switch – межкластерный коммутатор, позволяющий соединять функциональные устройства разных арифметических кластеров; SB – потоковый буфер для сборки/разборки потоков данных, удобных по организации для работы конвейерных алгоритмов обработки, реализуемых в арифметических кластерах; FMAC – устройство выполнения операции умножения и сложения над 64-разрядными числами в формате с плавающей запятой; DSQ – устройство выполнения деления и извлечения квадратного корня; SP – регистровая блокнотная память; COMM – блок связи с межкластерным коммутатором; JB/VAL (jump bit/check value) – блок проверки «на лету» по заданным условиям битов пакетов сообщений, преобразование их и при невыполнении условий – отбраковка.

Потоковая модель вычислений – это вычислительный граф, в узлах которого находятся вычислительные ядра, а по дугам передаются данные в виде наборов записей однородных данных. В FT64 вычислительные ядра реализуются на четырех арифметических кластерах, в каждом из которых по четыре 64-разрядных конвейерных устройства сложения-умножения с локальными регистровыми файлами LRF на входах. Вычислительные ядра отображаются на арифметические кластеры, передача данных между ядрами происходит через потоковые регистровые файлы SRF, имеющие объем 256 Кбайт, а для более сложных случаев – через внешнюю DRAM-память и блок интерфейса с внешней сетью. Соотношение пропускной способности DRAM, LRF и SRF – 1:10:85, что принципиально для процессоров такого типа. Управление арифметическими кластерами и передачами данных осуществляется внутрикристалльным контроллером и микроконтроллером, который содержит память 2Кх688 бит команд. FT64 был разработан за один год с использованием технологии 130 нм на кристалле 12х12 мм с тактовой частотой 500 МГц и потреблением 8,6 Вт. Пиковая производительность такого кристалла

– 16 GFLOPS. На оценочных тестах научных приложений один FT64 развивает реальную производительность в 4,2 раза выше, чем Itanium 2/1,6 ГГц, а плата из восьми процессоров дает почти линейный прирост в 6,8 раза. Ускоритель из восьми процессоров по энергетической эффективности превосходит Itanium 2 почти в 100 раз.

Проект MASA сравним с проектами Merrimac и TRIPS (США), ориентированными на бортовые приложения, но процессор MASA можно считать обобщенной реализацией на кристалле одной платы с несколькими FT64. Если в FT64 используется параллелизм уровня машинных команд и обработки данных (ILP- и DLP-параллелизм), то в MASA используется еще и тредовый (TLP) параллелизм. Один процессор MASA содержит два MIPS-процессора, управляющих 2D-сетью четырехъядерных блоков (тайлов), каждое ядро которых можно сопоставить с одним FT64. Ядро содержит 16 конвейерных АЛУ обработки вещественных чисел, файл потоковых регистров (SRF), потоковые буферы (SB) и локальные регистровые файлы (LRF). На периферии 2D-сети тайлов процессор MASA имеет интерфейсы с памятью и внешними устройствами с программируемым локальным управлением. Вариант процессора MASA с 256 АЛУ, реализованный по норме 45 нм, может развивать реальную производительность в 100-350 GFLOPS на большинстве оценочных тестов. Например, на задаче трехмерного преобразования Фурье процессор MASA 1 ГГц с пиковой производительностью 512 GFLOPS развивает реальную производительность 100,7 GFLOPS.



Board 0,1...N – платы с процессорами FT64; DDR2 DIMM – внекристалльная динамическая память, у каждого процессора FT64 своя; NI – блоки сетевых интерфейсов FT64; Bridge FPGA – схемы-мосты на программируемых логических матрицах; Memory – оперативная DRAM-память платы; Flash – внешняя флэш-память; Jtag? – специализированный аппаратный интерфейс тестирования; Uart – универсальный асинхронный приемопередатчик, производит преобразование параллельного кода в последовательный при выдаче и обратно – при приеме; Power module – модуль источника питания; Clock module – модуль выдачи тактовых сигналов; Communication FPGA – коммуникационная схема на программируемой логической матрице; Itanium 2 Host – управляющий процессор платы в виде микропроцессора Itanium 2; Connect to HI – подключение к интерфейсу FT64 с управляющим процессором, по нему осуществляется управление процессорами FT64 и их контроль.

Толерантность работы с памятью лучшим образом обеспечивают мультитредовые процессоры. Исторически сложилось, что тут выделяется направление с малым количеством тредов в процессорном ядре (единицы) и массово-мультитредовые

процессоры с большим количеством тредов (сотни) в ядре. В Китае ведутся работы по мультитредовым микропроцессорам первого направления, причем для VLIW- или EPIC-архитектур реализуется наиболее сложная – SMT-мультитредовость. Мультитредовые процессоры второго типа с запуском VLIW-команды за такт от одного тредра применяются в Cray XMT, а в более обобщенном виде, с запуском за такт нескольких RISC-команд, но от разных тредов ядра, – в российском массово-мультитредовом микропроцессоре СКСН «Ангара».

Андрей Моляков (andrei_molyakov@mail.ru) – сотрудник ЗАО «ВТ-Консалтинг» (Москва).

Проблемы современных суперкомпьютеров

Для создания суперкомпьютеров стратегического назначения необходимо решить следующие задачи:

- устранить «стену памяти», не позволяющую «снимать» с современных суперкомпьютеров реальную производительность выше 5-10% пиковой на задачах с плохой пространственно-временной локализацией обращений к памяти;
- добиться масштабируемости производительности и отказоустойчивости при увеличении количества процессоров;
- снизить энергопотребление.

Решение всех этих проблем и было определено в числе главных целей проекта 863/ИТ, который по мнению его разработчиков, как и аналогичные проекты США и Японии, будет промежуточным (кремниевым) этапом при переходе к квантовым компьютерам.