

Принципы построения многопроцессорных вычислительных систем на основе ПЛИС

И.А. Каляев¹, И.И. Левин¹, Е.А. Семерников²

- 1) НИИ многопроцессорных вычислительных систем имени академика А.В. Каляева Южного федерального университета, г. Таганрог, Россия
- 2) Южный научный центр Российской академии наук, г. Ростов-на-Дону, Россия

Аннотация

В статье приводятся принципы построения высокопроизводительных вычислительных систем на основе реконфигурируемой элементной базы. В качестве основного вычислительного элемента в них используются не универсальные микропроцессоры, а программируемые логические интегральные схемы сверхбольшой интеграции. Такие системы предназначены для решения вычислительно трудоемких задач различных предметных областей, обеспечивают реальную производительность не ниже 50% от пиковой производительности на широком классе задач и имеют практически линейный рост производительности при наращивании аппаратного ресурса.

Principles of multiprocessor computer systems design on basis of FPGA

I.A. Kaliaev¹, I.I. Levin¹, E.A. Semernikov²

- 1) SRI of multiprocessor computer systems of academician A.V. Kaliaev of Southern federal university (Taganrog)
- 2) Southern scientific center of Russian academy of sciences (Rostov-on-Don)

Abstract

In the article are given principles of high-performance computer systems design on basis of reconfigurable element base. As basic computing element there are used not general purpose microprocessors, but programmable logical integrated super large scale integration circuits. Such systems are considered to be used for solving tasks of high computational complexity from various subject fields. They provide real performance not less than 50% from the peak one for wide class of tasks and have practically linear performance growth at hardware resource growth.

Введение

Наиболее распространенные в мире кластерные МВС имеют высокие значения пиковой производительности. В то же время пользователи отмечают принципиальные недостатки кластерных МВС, которые связаны с относительно низкой скоростью процедур межпроцессорного обмена, ограниченной пропускной способностью сети передачи данных, необходимостью синхронизации множества взаимосвязанных последовательных процессов, каждый из которых выполняется на отдельном процессоре, и т.д. Все это приводит к тому, что высокую реальную производительность кластерные суперЭВМ демонстрируют, в основном, только при решении класса слабосвязанных задач, не требующих большого числа

информационных обменов, в то время как при решении задач других классов их реальная производительность существенно снижается и не превышает 5-10% от декларируемой пиковой производительности системы. Это является следствием неадекватности данной конкретной архитектуры суперкомпьютера информационной структуре решаемой задачи [1, 2, 3, 4].

Недостатки кластерных систем, связанные с их неизменной «жесткой» архитектурой, позволяет преодолеть концепция создания МВС с «гибкой», динамически перестраиваемой (программируемой) архитектурой на основе реконфигурируемой элементной базы или иначе – реконфигурируемые вычислительные системы (РВС) [3, 4]. Концепция создания РВС разработана и развивается в Научно-

исследовательском институте многопроцессорных вычислительных систем имени академика А.В. Каляева Южного федерального университета (НИИ МВС ЮФУ) уже более 20 лет. В отличие от многопроцессорных вычислительных систем с «жесткой» архитектурой архитектура РВС может динамически изменяться в процессе их функционирования. В результате у пользователя появляется возможность адаптации архитектуры вычислительной системы под структуру решаемой им задачи. Иными словами, пользователь может, оставаясь в рамках базовой архитектуры системы, создавать проблемно-ориентированные вычислители, структура которых соответствует структуре решаемой задачи. Проведенные исследования и многочисленные практические разработки, выполненные в НИИ МВС ЮФУ, показали [5,6,7], что реализация данной концепции обеспечивает высокую реальную производительность РВС, близкую к пиковой на широком классе задач, в том числе при решении «сильносвязанных» задач. В качестве основного вычислительного элемента в РВС используются не универсальные микропроцессоры, а программируемые логические интегральные схемы (ПЛИС) сверхбольшой интеграции. На основе предложенной концепции возможно без привлечения больших финансовых затрат создавать высокопроизводительные вычислительные системы, существенно опережающие зарубежные суперЭВМ по таким характеристикам как соотношения «реальная производительность/пиковая производительность» (эффективность) и «реальная производительность/объем» (компактность). Еще одной важной характеристикой РВС является практически линейный рост производительности при наращивании аппаратного ресурса [4], что является следствием исключения из вычислительного процесса множества вспомогательных действий, которые неизбежно сопровождают вычисления в универсальных процессорах, на базе которых построены суперЭВМ традиционной архитектуры.

Архитектурные особенности и принципы организации вычислений в РВС

В отличие от традиционных методов организации параллельных вычислений РВС ориентируются на абсолютно параллельную форму алгоритма задачи – ее информационный граф. Под информационным графом [2] понимается граф, вершины которого соответствуют арифметико-логическим операциям над операндами или ячейками (каналами) памяти, в которых расположены информационные массивы. Дуги информационного графа соответствуют информационной зависимости между вершинами. Если существует направленная дуга, соединяющая вершину a с вершиной b , то это означает, что результат операции, соответствующий вершине информационного графа a , является входным операндом операции, соответствующей вершине b . Вершины, соответствующие каналам памяти, в которых располагается входная и выходная информация, являются входными и выходными вершинами информационного графа. Информационные графы больших и сложных задач могут содержать миллионы вершин.

Идея концепции построения РВС заключается в аппаратной реализации всех операций, предписанных вершинами информационного графа задачи, всех каналов передачи данных между вершинами, соответствующих дугам графа, и всех информационных каналов, соответствующих входным и выходным вершинам. В этом случае задача, определенная информационным графом, будет выполнена максимально быстро, поскольку обеспечивается максимально возможное распараллеливание вычислений. Такое решение задачи принято называть *структурным* [4].

Основными вычислительными блоками в РВС являются *макропроцессоры* (МАП). Макропроцессор позволяет реализовывать крупные операции, которые предписаны вершинами информационного графа. Макропроцессор представляет собой некоторый набор элементарных процессоров (ЭП), объединяемых в единый

программно-неделимый вычислительный ресурс с помощью локального пространственного коммутатора (K_1) (рис. 1, а). Внутренний коммутатор макропроцессора, как правило, соединяет по полному графу все информационные входы и выходы элементарных процессоров, внешние информационные входы X , внешние информационные выходы Z макропроцессора. Настройка макропроцессора на крупную функционально законченную операцию (макрооперацию) производится с помощью блока макроопераций (БМ). В каждый момент времени макропроцессор может реализовывать только одну макрооперацию. Незадействованные в макрооперации элементарные процессоры будут простаивать.

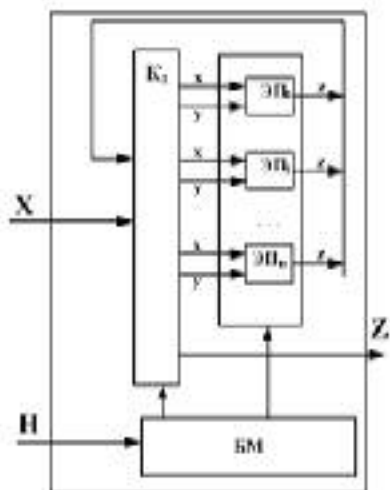


Рис.1. Структурная схема макропроцессора

Следует отметить, что в отличие от стандартного микропроцессора элементарный процессор макропроцессора не управляет процессом обработки информации, а лишь реализует соответствующую ему операцию над операндами, поступающими на его входы.

Для реализации того или иного информационного графа решаемой задачи макропроцессоры должны иметь возможности соединения в вычислительные параллельно-конвейерные структуры. С этой целью в состав РВС включен коммутатор второго уровня или системный коммутатор, обеспечивающий различные варианты соединения макропроцессоров друг с другом. Кроме того, в состав РВС входит распределенная память, обеспечивающая возможности

параллельной выдачи массива входных данных на входы макропроцессоров и записи результатов вычислений с их выходов.

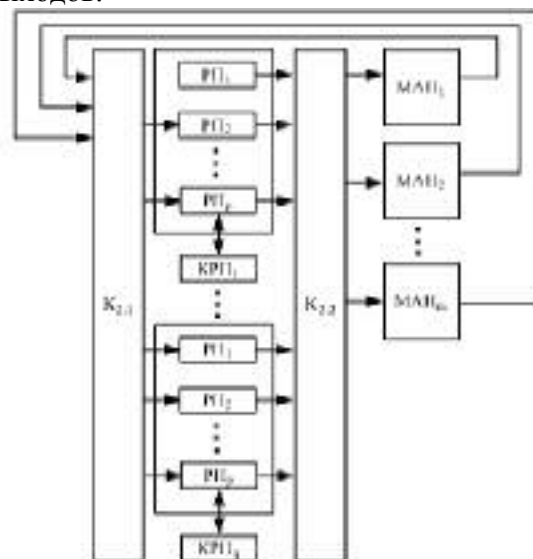


Рис.2. Обобщенная схема РВС

Обобщенная схема РВС приведена на рис. 2. В ее состав входят множество макропроцессоров, состоящих из элементарных процессоров, соединяемых между собой с помощью локального коммутатора; распределенной памяти (РП), включающей в свой состав контроллер распределенной памяти (КРП) и множество каналов РП, а также системный коммутатор K_2 , предназначенный как для соединения макропроцессоров с друг с другом в вычислительные структуры, так и для их параллельного доступа к каналам распределенной памяти. Множество входных и выходных дуг информационного графа реализуется системным коммутатором K_2 , обеспечивающим подключение каналов распределенной памяти к тем или иным входам и выходам макропроцессоров. При этом массивы входных данных отображаются в соответствующие каналы распределенной памяти, процедура обращения к которым реализуется с помощью КРП.

Основные положения концепции были сформулированы академиком А.В. Каляевым в середине 80-х годов прошлого века и развиваются в НИИ МВС ЮФУ уже более двадцати лет. Опираясь на сформулированную концепцию, в 80-е-90-е годы в НИИ МВС ЮФУ был создан целый ряд вычислительных систем с

реконфигурируемой архитектурой. Созданные системы представляли собой первые шаги на пути практического становления данного направления вычислительной техники.

Несмотря на большой объем проведенных 80–90-е годы исследований в области высокопроизводительных вычислительных систем с реконфигурируемой архитектурой и создание ряда опытных образцов, дело до серийного выпуска систем такого типа не дошло. Это было вызвано тем, что реконфигурируемые вычислительные системы, созданные на основе элементной базы того времени, не позволяли в полной мере реализовать преимущества концепции. Перестройка архитектуры затрагивала только связи между процессорами, в то время как структура самих процессоров перестраивалась в узких пределах и была достаточно жесткой, поскольку процессоры были выполнены на серийных микропроцессорах с неизменной структурой. Вычислительные системы получались громоздкими, малонадежными, сложными в эксплуатации.

Таким образом, реализация и широкое внедрение в практику идеи высокопроизводительных вычислительных систем с реконфигурируемой архитектурой сдерживалась отсутствием в нашей стране соответствующей этой концепции элементной базы. Полномасштабная же реализация PBC требует наличия принципиально иной элементной базы, удовлетворяющей следующим основным критериям:

- высокой степени интеграции, позволяющей реализовать крупные фрагменты вычислений;
- сравнительно низкой стоимости при мелкосерийном производстве вычислительных систем;
- возможности аппаратной реализации специализированных вычислительных структур;
- возможности реализации различных схем распараллеливания вычислений;
- возможности реконфигурации сформированных вычислительных структур при переходе от одних алгоритмов и схем распараллеливания к другим;

- поддержки разработки схемотехнических решений системами автоматизированного проектирования.

Всем этим требованиям отвечают появившиеся в конце XX-го и начале XXI-го века программируемые логические интегральные схемы (ПЛИС) со сверхвысокой степенью интеграции (в английской аббревиатуре FPGA – Filed Programmable Gates Array). В ПЛИС изначально заложены возможности реконфигурирования их внутренней архитектуры, и поэтому они наилучшим образом соответствуют концепции реконфигурируемых вычислительных систем.

Основными преимуществами программируемых логических интегральных схем являются: возможность аппаратной реализации сложных параллельных алгоритмов; наличие средств САПР, позволяющих провести полное моделирование системы; возможность программирования или изменения конфигурации непосредственно в системе; совместимость различных проектов при их переводе на языки описания аппаратуры - VHDL, AHDL, Verilog и др.

ПЛИС выпускаются несколькими фирмами: Xilinx, Altera, Atmel, Vantis, Lucent и др. Ведущим производителем ПЛИС является фирма Xilinx, занимающая более 40% мирового рынка. При этом фирма каждые 1,5-2 года выпускает на рынок все новые поколения ПЛИС, превосходящие предыдущие поколения по таким показателям, как потребляемая мощность, тактовая частота и функциональные возможности.

На рис. 3 показаны графики, отражающие основные тенденции развития высокопроизводительных ПЛИС фирмы Xilinx.

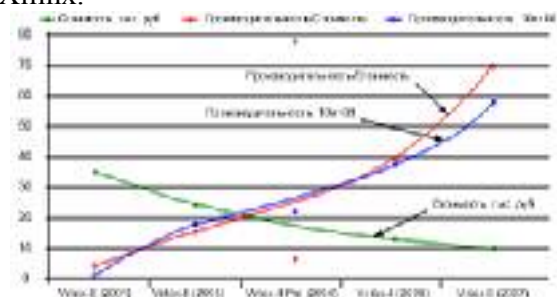


Рис. 3. Основные тенденции развития высокопроизводительных ПЛИС фирмы Xilinx

Как видно, стоимость ПЛИС имеет тенденцию к снижению в абсолютном выражении. Производительность (синий график) оценивалась путем размещения в ПЛИС максимально возможного количества 32-разрядных устройств с плавающей запятой стандарта IEEE-754, функционирующих на частоте работы микросхемы. Красный график показывает экспоненциальный рост отношения «Производительность/Стоимость», что, в свою очередь, свидетельствует о росте экономической эффективности элементной базы на основе ПЛИС.

Таким образом, наличие широких возможностей реконfigurирования, поддержанных удобными САПР, и высокая экономическая эффективность современных ПЛИС позволяют сделать однозначный вывод о том, что они являются наиболее перспективной элементной базой для построения РВС.

Применение ПЛИС большой степени интеграции дало возможность перейти к созданию больших решающих полей, объединяющих вычислительные ресурсы множества ПЛИС в единый вычислительный ресурс. Это позволило отказаться от прежнего подхода, в котором отдельным микросхемам отводилось определенное постоянное функциональное назначение, и перейти к более прогрессивной архитектуре, когда в любой микросхеме может быть реализовано любое функциональное устройство, которое необходимо в данном месте решающего поля. В этом смысле решающее поле является как бы аналогом сверхбольшой ПЛИС, объединяющей в себе все их аппаратные ресурсы.

Однако построение больших решающих полей на ПЛИС выявило новые проблемы. Первая проблема - это негативный эффект границ, возникающих на стыках отдельных ПЛИС при их объединении в решающее поле. Эффекты границ сказываются при реализации больших фрагментов вычислительных структур, которые не могут быть реализованы в пределах отдельной микросхемы, а требуют ресурса нескольких ПЛИС. Вторая проблема - это конструктивно-технологические ограничения, которые заключается в том, что невозможно разместить неограниченное количество ПЛИС на

печатной плате приемлемого размера. Эта проблема решается путем модульного построения аппаратных средств РВС и модульной наращиваемости на основе унифицированных базовых модулей. Базовые модули не только позволяют путем комплексирования создавать РВС необходимой производительности, но и сами по себе являются реконfigurируемой вычислительной системой, способной совместно с персональным компьютером (в качестве интеллектуального терминала) решать пользовательские задачи.

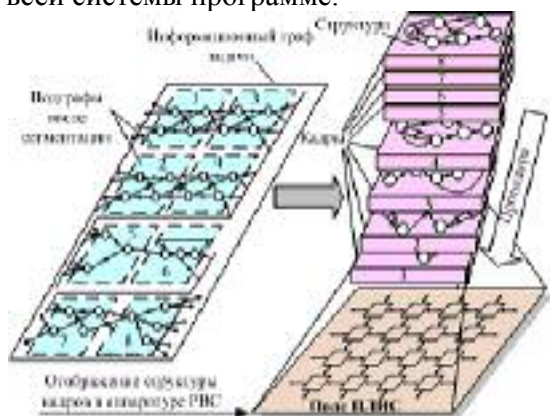
Однако введение модульного принципа построения аппаратных средств обострило первую проблему, породив еще один тип границ – межмодульный. Суть решения проблемы границ заключается в создании единого типа интерфейса как между ПЛИС в пределах базового модуля, так и между ПЛИС разных базовых модулей.

Ограниченность аппаратного ресурса приводит к тому, что для больших задач весь информационный граф не может быть отображен в имеющемся аппаратном ресурсе РВС. В этом случае информационный граф сегментируется на фрагменты – непересекающиеся базовые подграфы, физически реализуемые в аппаратуре РВС, а решение большой задачи выполняется *структурно-процедурным* способом, при котором на решающее поле ПЛИС поочередно отображаются базовые подграфы информационного графа, и вычисления в соответствии с отображенным подграфом выполняются структурно, а смена подграфов выполняется процедурно [4, 8].

В простейшем случае структурно-процедурный вычислительный процесс в РВС организуется следующим образом. В аппаратном ресурсе РВС реализуется один из базовых подграфов сегментированного информационного графа задачи и на его входы из блоков памяти подаются входные данные этого фрагмента. Процесс вычисления реализуется вычислительной структурой, соответствующей данному подграфу, и на выходе формируется результат вычислений, который запоминается в соответствующих блоках памяти. Далее в аппаратном ресурсе РВС реализуется следующий базовый подграф сегментированного информационного

графа задачи, и процесс повторяется. Этот процесс показан на рис. 4.

При реализации структурно-процедурного метода информационный граф задачи представляется в виде кортежа изоморфных базовых подграфов, которые являются информационно-независимыми или непосредственно зависящими друг от друга. Кортеж информационных подграфов преобразуется в специальную вычислительную конструкцию - *кадр*. Можно сказать, что кадру соответствует подграф задачи, реализованный аппаратно, через который следует поток операндов. При этом каждая группа операндов (результатов) соответствует входным (выходным) вершинам определенного подграфа кортежа. Смена кадров в РВС осуществляется процедурно по единой для всей системы программе.



Р

ис. 4. Процесс решения задачи в РВС

Механизм последовательного обхода подграфов информационного графа задачи кадрами принято называть *структурно-процедурной организацией вычислений*.

РВС со структурно-процедурной организацией вычислений является гибридом фон-неймановской архитектуры и архитектуры потока данных. Такая организация вычислений обеспечивает детерминизм выполнения программы, что в общем случае недостижимо в многопроцессорных системах, построенных по традиционной мультипроцедурной архитектуре (кластерные МВС). При этом обеспечивается также и высокая эффективность параллельных вычислений на широком классе задач.

Особенность архитектуры РВС заключается в том, что в процессе

конструирования она не формируется окончательно, а остается в определенном смысле незавершенной и открытой. Окончательное программирование архитектуры РВС включает создание функциональных узлов (элементарных процессоров) для выполнения вычислений, настройку прямых информационных каналов в коммутаторах K_1 и K_2 между элементарными процессорами, настройку блоков распределенной памяти на реализацию процедур чтения и записи информационных массивов. Аппаратно-программные средства РВС позволяют синтезировать произвольные соединения между компонентами системы, создавая необходимые вычислительные структуры. Совокупность вычислительных структур, созданных в рамках базовой архитектуры РВС, образуют виртуальный проблемно-ориентированный вычислитель, структура которого адекватна информационному графу (подграфу) решаемой задачи [4].

Организация аппаратных средств РВС на основе ПЛИС

Высокопроизводительные реконфигурируемые вычислительные системы целесообразно строить по принципу модульной наращиваемости из однотипных базовых модулей. Базовые модули, с одной стороны, несут на себе конструктивные ограничения, а с другой стороны, должны в полной мере сохранить все концептуальные особенности архитектуры реконфигурируемых систем. Поэтому требования и ограничения, оказывающие влияние на характеристики базового модуля РВС, можно условно разделить на архитектурные и конструктивно-технологические.

К архитектурным требованиям относятся параметры системы, которые должны найти свое отражение в основных характеристиках базового модуля таким образом, чтобы при построении системы реализовать все отличительные особенности РВС. Исходя из архитектурных требований, базовый модуль должен включать некоторое сбалансированное количество элементарных процессоров и блоков памяти, объединенных с помощью полнодоступной системы быстрых каналов

и способных реализовать, желательного целиком, базовые подграфы различных задач. С точки зрения операционной системы базовый модуль должен представлять собой программно-неделимый аппаратный ресурс при организации вычислительного процесса в РВС.

К конструктивно-технологическим требованиям и ограничениям относятся возможности конструктивной реализации всех архитектурных особенностей базового модуля на современном уровне технологии. Параметры базового модуля зависят от многих факторов: элементной базы, которая определяет вычислительные возможности базового модуля; размеров и технологических норм печатной платы, характеристик разъемов и соединителей для организации внешних связей, подсистемы электропитания и отвода тепла и т.п. Все это накладывает определенные ограничения на архитектурные возможности базовых модулей.

Сформулированные архитектурные и конструктивные требования и ограничения являются определяющими при создании РВС, которые предназначены для решения широкого круга задач. Для проблемно-ориентированных РВС определяющим может стать требование обязательной реализации в пределах базового модуля базовых подграфов данной проблемной области. При этом базовый модуль может быть конструктивно реализован на нескольких печатных платах.

Укрупненная структура базового модуля показана на рис. 5.

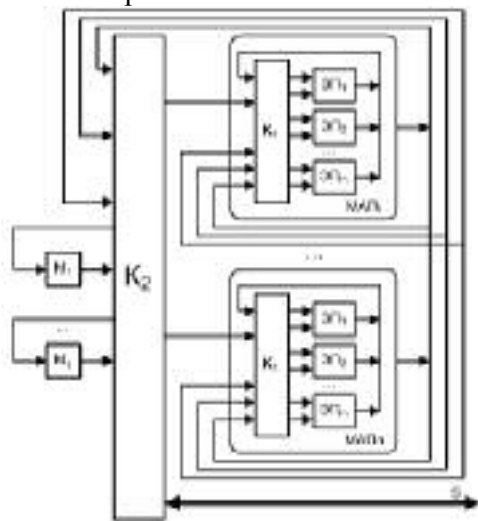


Рис. 5. Укрупненная структура базового модуля

В состав базового модуля (БМ) входит некоторое множество макропроцессоров (МАП) и блоков памяти (М), объединенных с помощью пространственной коммутационной системы K_2 . Каждый макропроцессор, в свою очередь, содержит t элементарных процессоров (ЭП), объединенных коммутатором K_1 . Таким образом, в базовом модуле, показанном на рис. 5, имеет место двухуровневая коммутационная система. Первый уровень устанавливает связи между элементарными процессорами внутри макропроцессора, а второй уровень – связи между макропроцессорами внутри базового модуля.

Наращивание производительности путем объединения в вычислительный блок нескольких базовых модулей возможно несколькими путями. Первый из них заключается во введении еще одного уровня коммутации для создания связей между базовыми модулями. Дальнейшее объединение модулей и создание более производительных структур возможно путем введения дополнительных уровней коммутации между все более крупными объединениями базовых модулей.

Другой путь заключается в объединении базовых модулей посредством использования ресурсов внутримодульного коммутатора K_2 , как это показано на рис. 6. В этом случае можно обойтись одним типом базового модуля. Однако удаленные друг от друга базовые модули будут соединяться по довольно длинным цепочкам и по все более узким каналам. Возможны также любые сочетания рассмотренных вариантов.

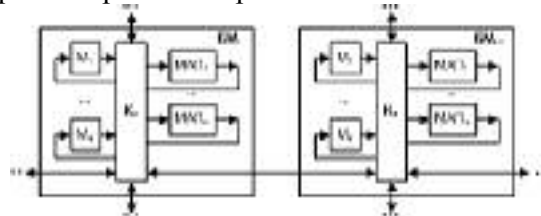


Рис. 6. Обобщенная структура модульно-наращиваемой РВС на базе внутримодульного коммутатора

Базовый модуль на основе ПЛИС несет в себе все характерные признаки завершённой реконфигурируемой системы

[4,6,7]. Типовая компоновка базового модуля на основе ПЛИС показана на рис. 7.

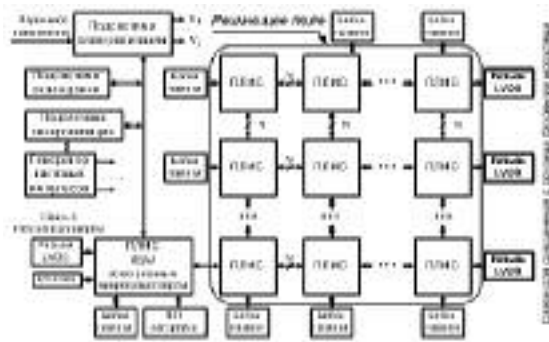


Рис. 7. Структура базового модуля реконфигурируемой системы на основе ПЛИС

Основные вычислительные возможности базового модуля сосредоточены в решающем поле, которое содержит некоторое множество ПЛИС большой степени интеграции. В решающем поле создаются вычислительные и коммутационные структуры: макропроцессоры $МАП_1 - МАП_n$, каждый из которых, в свою очередь, содержит несколько элементарных процессоров $ЭП_1 - ЭП_m$, объединенных коммутаторами K_1 , внутримодульный коммутатор K_2 , контроллеры распределенной памяти для управления блоками распределенной памяти. При этом сами блоки распределенной памяти выполняются на типовых микросхемах ОЗУ SRAM или SDRAM необходимого объема и быстродействия.

Контроллер базового модуля (КБМ) выполняет функции управления и контроля всех систем базового модуля.

На базовом модуле располагаются также и вспомогательные подсистемы: синхронизации, электропитания и охлаждения.

Центральное место в организации вычислительного процесса отводится контроллерам распределенной памяти. Контроллеры памяти работают с фрагментами параллельной программы, которые загружены в их блоки памяти. Исполняя параллельную программу, контроллеры распределенной памяти участвуют в настройке элементарных процессоров на выполнение необходимых операций и создают с помощью коммутаторов K_1 и K_2 необходимые каналы

связи между ними, тем самым реализуя в пределах базового модуля мультиконвейерную вычислительную структуру, соответствующую базовому подграфу задачи.

Одна часть контроллеров распределенной памяти, исполняя фрагменты параллельной программы, организует и синхронизирует потоки данных, подаваемые в вычислительные структуры. Другая часть контроллеров выполняет функции приема результатов вычислений. Базовый модуль с точки зрения организации потоков данных через аппаратно реализованный в нем базовый подграф показан на рис. 8.

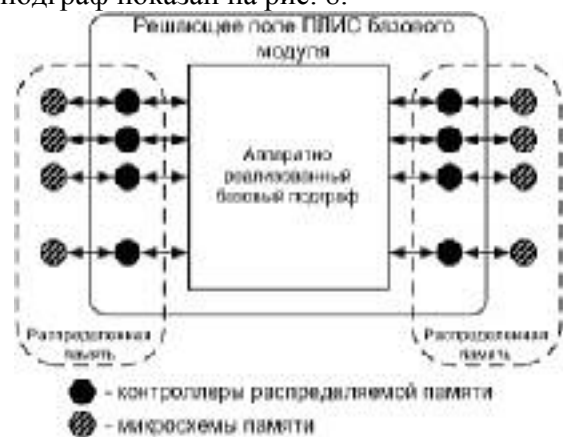


Рис. 8. Организации потоков данных через аппаратно реализованный базовый подграф

Для создания эффективных вычислительных структур в пределах базового модуля необходимо оптимальное соотношение количества ПЛИС, количества блоков распределенной памяти и их объема. Для различных задач, решаемых на РВС, это соотношение различно. В то же время для структурной реализации вычислений не требуется запоминание множества промежуточных данных, так как они передаются для дальнейшей обработки в последующие ступени конвейерного вычислителя без промежуточного запоминания. Это снижает требования к общему объему памяти на базовом модуле. Для структурной реализации вычислений более критичным является количество блоков распределенной памяти, а не объем каждого блока или общий объем памяти. Помимо распределенной памяти, реализованной на типовых микросхемах ОЗУ, при создании вычислительных

структур широко используется внутрикристалльная память ПЛИС.

Поскольку в решающем поле реализуются вычислительные структуры, которые располагаются в ресурсах нескольких или даже всех микросхем, то необходимо стремиться к тому, чтобы максимально снизить негативный эффект границ между кристаллами ПЛИС. Снижение межкристалльного эффекта границ достигается двумя путями. Первый из них предполагает использование ПЛИС максимальной интеграции, которые позволяют размещать все более крупные фрагменты функциональных узлов. Второй путь заключается в создании регулярных и однородных связей между всеми ПЛИС базового модуля да и всей РВС в целом.

С этой целью связи между ПЛИС решающего поля базового модуля целесообразно выполнять на основе стандарта LVDS (Low Voltage Differential Signaling – дифференциальные сигналы низкого напряжения). Преимуществами стандарта LVDS являются: низкая потребляемая мощность выходных каскадов, низкий уровень создаваемых электромагнитных излучений, невосприимчивость к синфазным электромагнитным помехам и наличие в микросхемах современных ПЛИС аппаратной поддержки для организации высокоскоростных передач данных на основе стандарта LVDS. Темп передачи данных по каждой двухпроводной линии в зависимости от реализации может составлять несколько гигабит в секунду.

ПЛИС решающего поля располагаются в узлах двумерной решетки и соединяются между собой ортогональной системой связей по близкодействию. Такая система связей позволяет существенно упростить печатную плату и улучшить ее частотные характеристики, поскольку связи между соседними микросхемами не превышают единицы сантиметров. Данные между дальними микросхемами передаются по транзитным каналам через промежуточные микросхемы, используя систему ортогональных связей.

Следующим конструктивным уровнем, где необходимо принимать специальные меры для снижения эффекта границ, являются базовые модули.

Поскольку при реализации вычислительных структур больших базовых подграфов может не хватить ресурсов одного базового модуля, то возникает необходимость реализации таких структур в пределах двух и более базовых модулей. При этом важно сохранить темп передачи данных из микросхем решающего поля одного базового модуля непосредственно в ПЛИС другого базового модуля. Для сохранения темпа передачи в межмодульных обменах целесообразно также использовать стандарт LVDS. Единый тип интерфейса между ПЛИС как в пределах одного базового модуля, так и между различными базовыми модулями, обеспечивает одинаковые скорости обмена и снижает эффект границ в пределах общего схемотехнического ресурса РВС.

Принцип модульной наращиваемости позволяет увеличить производительность РВС при увеличении количества базовых модулей. При этом обеспечивается возможность организации ресурсонезависимого и отказоустойчивого программирования, параметризуемого относительно вычислительного ресурса (количества базовых модулей), выделяемого для решения задачи. Задача может быть решена на любой конфигурации вычислительной структуры. Выход из строя одного или нескольких базовых модулей не влечет за собой прекращение решения задачи и необходимость ее полной перетрансляции, а лишь несколько замедляет процесс решения [4,7].

На основе описанной выше концепции в НИИ МВС ЮФУ за последние 5-6 лет был создан ряд РВС различной производительности и назначения, из которых две системы в настоящее время выпускаются серийно.

Внешний вид одной из созданных РВС производительностью 200 Гфлопс и ее базового модуля 16V4-50 показан на рис. 9 и рис 10.



ис.9. PBC «Медведь»



Рис. 10. Универсальный базовый модуль 16V4-50

Показатели «компактности» и «эффективности» созданных систем составили 5-6 Гфлопс/дм³ и менее 18000-20000 руб/Гфлопс. Для задач символьной обработки эти же показатели составили 115x10⁹ операций/дм³ и 2000 руб/млрд.оп.

В настоящее время по заданию Федерального агентства по науке и инновациям ведется выполнение Государственного контракта № 02.524.12.4002 по теме «Создание семейства высокопроизводительных многопроцессорных вычислительных систем с динамически перестраиваемой архитектурой на основе реконфигурируемой элементной базы и их математического обеспечения для решения вычислительно трудоемких задач» в рамках Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития

научно-технологического комплекса России на 2007-2012 годы».

Целью разработки является создание на единых архитектурных принципах семейства программно-совместимых PBC производительностью от 0,025 Тфлопс до 6 Тфлопс. Создаваемое семейство PBC включает:

- PBC-5 – высокопроизводительная систему производительностью 6 Тфлопс;
- PBC-1P и PBC-1K – системы производительностью более 1 Тфлопс;
- PBC-0.2-PC – рабочую станцию производительностью 300 ГФлопс;
- РУПК-50 и РУПК-25 – ускорители персональных компьютеров производительностью 50 и 25 ГФлопс.

Старшие представители семейства создаются на принципах модульной наращиваемости и будут обладать почти линейным ростом реальной производительности в зависимости от увеличения аппаратного ресурса [7].

Для создания семейства PBC, системного и прикладного программного обеспечения его представителей, информационной и программной инфраструктуры для обучения и поддержки потенциальных пользователей семейства образован консорциум из ряда научных и научно-производственных организаций, в который вошли: НИИ МВС ЮФУ (г. Таганрог) – головной исполнитель проекта; ФГУП «НИИ «Квант» (г. Москва); «Южный научный центр РАН» (г. Ростов-на-Дону); «Специальная астрофизическая обсерватория РАН» (пос. Нижний Архыз); НИЦ «СуперЭВМ и нейрокомпьютеров» (г. Таганрог); НИВЦ МГУ (г. Москва).

Реальная производительность всех представителей семейства PBC планируется не ниже 50% от указанной пиковой производительности на широком классе вычислительных задач. При этом они будут обладать существенно меньшими габаритами, энергопотреблением и стоимостью по сравнению с суперЭВМ кластерного типа аналогичной производительности.

Системное программное обеспечение

Программирование PBC отличается от программирования МВС традиционной архитектуры, и его можно условно

Р

разделить на две составляющие: программирование структурное, которое создает необходимые вычислительные структуры в поле логических ячеек ПЛИС, и программирование процедурное – программирование в традиционном смысле, заключающиеся в организации вычислительного процесса в РВС. При этом программирование вычислительных структур вызывает у пользователей наибольшие трудности [10]. Это связано с тем, что традиционно, пользователи привыкли программировать только организацию вычислительного процесса, опираясь на неизменяемую аппаратную поддержку средств вычислительной техники, в то время как для программирования вычислительных структур РВС требуется совершенно другая квалификация, а именно - квалификация схемотехника.

При программировании пользовательской задачи структура РВС приобретает черты специализированной многопроцессорной ЭВМ, которая оптимально соответствует структуре решаемой задачи из предметной области. Это обеспечивает высокую реальную производительность системы, близкую к пиковой производительности на широком классе задач, и позволяет достичь практически линейного роста производительности при наращивании аппаратного ресурса. Эффективность вычислительного процесса при реконфигурации архитектуры РВС на низком (схемотехническом) уровне может быть повышена от 10 до 100 раз по сравнению с вычислительными системами, архитектура которых не может быть изменена. Это делает, с одной стороны, чрезвычайно привлекательной реконфигурируемые на низком уровне системы, а с другой стороны, их программирование становится по сложности сопоставимым с созданием новой вычислительной системы. Такой подход требует новых методов и средств организации параллельных вычислительных процессов.

Создаваемое в рамках проекта системное и прикладное программное обеспечение обеспечивает потенциальным пользователям удобство программирования сложных практических задач на РВС и

включает: программный комплекс средств разработки прикладных программ, средства администрирования вычислительных ресурсов РВС и служебные программы и драйверы.

Целью создания программного комплекса средств разработки прикладных программ для РВС является предоставление пользователю возможностей, которые позволяют создавать программы без привлечения специальных знаний в области схемотехники ПЛИС и по сложности будут приближены к обычному программированию для многопроцессорных ЭВМ. Основными задачами программного комплекса средств разработки являются эффективная реализация вычислительно трудоемких фрагментов задач различных проблемных областей на произвольном количестве взаимосвязанных кристаллов ПЛИС и произвольном количестве базовых модулей, а также поддержка разработки и отладки прикладных программ на языках структурно-процедурных вычислений, в том числе с вызовом библиотечных функций настройки архитектуры системы и реализации необходимых вычислительных структур на множестве ПЛИС.

Программный комплекс средств разработки прикладных программ для РВС содержит: язык ассемблера Argus v.3.0; язык структурно-процедурного программирования высокого уровня COLAMO v.2.0; интегрированную среду разработки аппаратно-программных решений прикладных задач Argus IDE v.3.0 и поддерживающую языки программирования Argus v.3.0 и COLAMO v.2.0; отладчик параллельных программ на базовых модулях РВС, поддерживающий межмодульные связи; программный интерфейс доступа к вычислительным ресурсам РВС из различных сред программирования; среду разработки вычислительных структур для синтеза масштабируемых параллельно-конвейерных процедур, оперирующую библиотекой схемных решений (IP-ядер).

Язык структурно-процедурного программирования высокого уровня COLAMO [4, 8] обеспечивает синтаксическую поддержку реконфигурации аппаратной платформы РВС и возможность использования

элементов библиотеки масштабируемых IP-ядер. Транслятор COLAMO v.2.0 обеспечивает трансляцию исходного кода программы в язык ассемблера Argus v.3.0 и в VHDL посредством среды разработки масштабируемых параллельно-конвейерных процедур Fire!Konstruktor, создавая тем самым конфигурационные файлы для ПЛИС.

Средства администрирования вычислительных ресурсов PBC, необходимые для обеспечения доступа и управления вычислительными ресурсами базовых модулей, состоят из сервера удаленного доступа к вычислительным ресурсам, обеспечивающего обработку удаленных заявок на использование вычислительных ресурсов и поддерживающего очередь заявок с учетом приоритетов пользователей и клиентской части удаленного доступа, обеспечивающей прозрачный с точки зрения пользователя режим доступа и функционирования удаленных базовых модулей системы.

Заключение

Описанные в статье принципы построения высокопроизводительных систем на основе реконфигурируемой элементной базы и создаваемое на этих принципах семейство PBC являются принципиально новым направлением развития высокопроизводительной вычислительной техники. В качестве основного вычислительного элемента в PBC используются не универсальные микропроцессоры, а ПЛИС. Это дает возможность пользователям создавать в базовой архитектуре PBC виртуальные специализированные вычислители, структура которых адекватна структуре решаемой задачи, что, в свою очередь, обеспечивает высокую эффективность вычислений и близкий к линейному рост производительности при наращивании вычислительного ресурса.

Проблема снижения производительности традиционных МВС на сильносвязанных задачах может быть разрешена только путем предоставления пользователю больших возможностей в части архитектурного программирования аппаратного параллелизма вычислительной

системы. В этом контексте PBC, построенные на принципах, описанных в статье, обладают рядом преимуществ перед многопроцессорными системами традиционной организации, основанными на использовании микропроцессоров в качестве основного вычислительного элемента.

1 Аладышев О.С., Дикарев Н.И., Овсянников А.П. и др. СуперЭВМ: области применения и требования к производительности - Известия ВУЗов. Электроника, 2004. - №1. - С.13-17.

2 Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. - С.-Петербург: «БХВ-Петербург», 2002. - 599 с..

3. Каляев А.В. Многопроцессорные вычислительные системы с программируемой архитектурой. - М.: Радио и Связь, 1984. - 240 с.

4. Каляев А.В., Левин И.И. Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. - М.: Янус-К, 2003. - 380 с.

5. Каляев И.А., Левин И.И. Многопроцессорные вычислительные системы (суперкомпьютеры): состояние и перспективы. // Вестник компьютерных и информационных технологий. - М.: Машиностроение, 2004. - №5. - С. 25-44.

6. Беседин И.В., Дмитренко Н.Н., Каляев И.А., Левин И.И., Семерников Е.А. Семейство базовых модулей для построения реконфигурируемых многопроцессорных вычислительных систем со структурно-процедурной организацией вычислений. // Материалы Всероссийской научной конференции «Научный сервис в сети Интернет: технологии распределенных вычислений», Новороссийск. - М.: Издательство Московского университета, 2006. - С 47-49.

7. Каляев И.А., Левин И.И. Высокопроизводительные модульно-наращиваемые многопроцессорные системы на основе реконфигурируемой элементной базы // Вычислительные методы и программирование. - М.:Изд-во Московского Университета, 2007. - Т.8. - №1. - С. 181-190.

8. Каляев А.В., Каляев И.А., Левин И.И. Многопроцессорные вычислительные системы с программируемой архитектурой

на основе ПЛИС - Вестник ЮНЦ РАН, 2004. – С.24-33.

9. Дордопуло А.И., Каляев И.А., Левин И.И., Семерников Е.А. Семейство многопроцессорных вычислительных систем с динамически перестраиваемой архитектурой // Материалы Четвертой Международной научной молодежной школы «Высокопроизводительные вычислительные системы». - Таганрог: Изд-во ТТИ ЮФУ, 2007, 68-74.

10. Левин И.И. Язык параллельного программирования высокого уровня для структурно-процедурной организации вычислений // Труды Всероссийской научной конференции. - М.: Изд-во МГУ, 2000, 108-112.