

УДК 004.272.43

И.И. Левин

СЕМЕЙСТВО ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ РЕКОНФИГУРИРУЕМЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Рассматриваются вопросы создания многопроцессорных высокопроизводительных вычислительных систем (МВС), в которых в качестве основного вычислительного элемента используются не универсальные микропроцессоры, а программируемые логические интегральные схемы (ПЛИС). Созданные аппаратно-программные средства поддерживают структурно-процедурную организацию вычислительного процесса в системе и позволяют довести реальную производительность до 60% от пиковой производительности на широком классе задач, а также добиться линейного роста производительности при наращивании аппаратного ресурса.

Ключевые слова: *высокопроизводительные вычислительные системы, программируемые логические интегральные схемы, аппаратно-программные средства, структурно-процедурная организация вычислительного процесса*

Если десять лет назад компьютерному сообществу казалось, что вычислительная техника может неограниченно развиваться, совершенствуя микропроцессорные схмотехнические и технологические решения, то сейчас ясно, что без координатной перестройки архитектуры вычислительной техники сохранить темп роста производительности невозможно.

Возможности роста производительности вычислительных систем с помощью дальнейшего уменьшения технологических норм оказались практически исчерпанными. Кроме того, при увеличении плотности размещения транзисторов на кристалле значительно возрастает сложность процессоров, поэтому прирост производительности не пропорционален затрачиваемым аппаратным ресурсам и энергии. Не помогает дальнейшее наращивание объемов кэш-памяти микропроцессоров, а также числа одновременного выполнения команд. Заявленные производителями пиковые характеристики процессоров практически не достижимы без низкоуровневого программирования. Большинство вычислительных задач выполняется на массовых микропроцессорах с эффективностью не более 10 – 20% [1].

Одним из способов повышения производительности вычислительных систем является распараллеливание вычислительных процессов. В настоящее время многопроцессорные вычислительные системы (МВС) используются в различных областях науки и техники, а также в промышленности и экономике. В то же время реальная производительность многопроцессорных вычислительных систем, которые ориентированы на традиционные методы организации параллельных вычислений и представляют собой механистически соединенные традиционные микропроцессоры, зачастую не превышает 10 – 15% от заявляемой пиковой производительности вследствие необходимости реализации множества процедур межпроцессорного обмена, а также синхронизации последовательных процессов, выполняемых в процессорах системы [2]. Более того, для задач, требующих интенсивных, но нерегулярных межпроцессорных обменов и обращений к системе распределенной памяти, реальная производительность МВС снижается до 1%.

Основная причина – это несоответствие между «жесткой» архитектурой МВС и информационной структурой широкого класса решаемых задач, из-за чего существующие суперкомпьютеры и работают столь неэффективно. Данный недостаток позволяет устранить способ построения МВС с «гибкой», динамически перестраиваемой (программируемой) архитектурой, подстраиваемой под информационную структуру каждой конкретной задачи, решаемой в текущий момент времени на таком компьютере [3, 4].

Практическое внедрение данной концепции сдерживалось отсутствием необходимой для ее реализации реконфигурируемой элементной базы. В последние годы такая элементная база на рынке появилась – это программируемые логические интегральные схемы (ПЛИС) высокой степени интеграции. На основе ПЛИС возможно без привлечения больших финансовых затрат создавать МВС с программируемой архитектурой, существенно опережающие зарубежные аналоги по таким характеристикам, как соотношение «реальная производительность/пиковая производительность» (эффективность) и «реальная производительность/объем» (компактность).

По этому пути уже идут ведущие мировые производители: компания Сray предлагает снабдить процессоры Orteron на двухпроцессорной платформе помощником в виде дополнительного модуля на базе ПЛИС Virtex 4 (DRC Coprocessor Module), включаемого во второй сокет Orteron-процессора. На DRC Coprocessor Module возлагается решение трудоемких для универсального процессора задач. Структура ПЛИС DRC Coprocessor Module при этом перестраивается под конкретную решаемую задачу.

В то же время западные производители используют ПЛИС, как правило, только в качестве сопроцессоров к стандартным вычислительным узлам – универсальным микропроцессорам. Концепция же построения МВС с программируемой архитектурой предполагает использовать ПЛИС в качестве базы для создания динамически реконфигурируемых вычислительных систем (РВС), адаптируемых под структуру решаемой задачи. В качестве основного вычислительного элемента в РВС используются вычислительные структуры, созданные в поле логических блоков ПЛИС, в то время как небольшое количество универсальных микропроцессоров выполняют вспомогательные функции: загрузку конфигураций ПЛИС, тестирование, распределение вычислительной нагрузки, загрузку исходных данных, визуализацию результатов и т.п.

Теоретические положения данной концепции уже подтверждены созданием в НИИ многопроцессорных вычислительных систем имени академика А.В. Каляева Южного федерального университета (НИИ МВС ЮФУ) целого ряда экспериментальных образцов РВС различных конфигураций на базе ПЛИС – от малогабаритных одноплатных устройств до систем, содержащих тысячи процессоров. В частности, в рамках ФЦНТП «Исследования и разработки по приоритетным направлениям науки и техники» на 2002 – 2006 годы была создана модульно-наращиваемая РВС с производительностью 200 Гфлопс в объеме всего 40 дм³. Данная система показала исключительно высокую реальную производительность на широком классе задач линейной алгебры, цифровой обработки сигналов, математической физики, символьной обработки, криптографии и других, работая с эффективностью не менее 60%.

В основу построения семейства РВС положены архитектурные принципы модульно-наращиваемых многопроцессорных вычислительных систем с программируемой архитектурой и структурно-процедурной обработкой информации [4].

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 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]. Входные и выходные вершины графа соответствуют блокам распределенной памяти, внутренние (операционные) вершины графа соответствуют процессорам, настроенным на выполнение определенной операции. Дуги информационного графа соответствуют каналам передачи данных. Если ресурс РВС позволяет полностью реализовать информационный граф задачи, то такая организация вычислений называется структурной. При структурной реализации вычислений обеспечивается максимальная скорость обработки информации.

Разумеется, прямое отображение информационного графа в архитектуру системы нецелесообразно, да и в большинстве случаев невозможно, поскольку число функциональных устройств, необходимых для покрытия информационного графа, может достигать $10^{12} - 10^{15}$. В то же время задачи, которые должны решаться на многопроцессорных системах, как правило, характеризуются цикличностью обработки данных, т.е. содержат повторяющиеся фрагменты – базовые подграфы информационного графа задачи. Поэтому информационный граф задачи может быть сегментирован на совокупность базовых подграфов, каждый из которых может быть реализован в РВС. В этом случае задача может быть решена *структурно-процедурными* методами [4]. В идеале необходимо стремиться к тому, чтобы информационный граф задачи был сегментирован на одинаковые по своей структуре – *изоморфные* подграфы. Однако в общем случае информационный граф задачи разбивается на несколько непересекающихся областей, каждая из которых, в свою очередь, сегментируется на изоморфные в пределах области подграфы.

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 подграф сегментированного информационного графа задачи, и процесс повторяется. Этот процесс показан на рис. 1.

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.

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

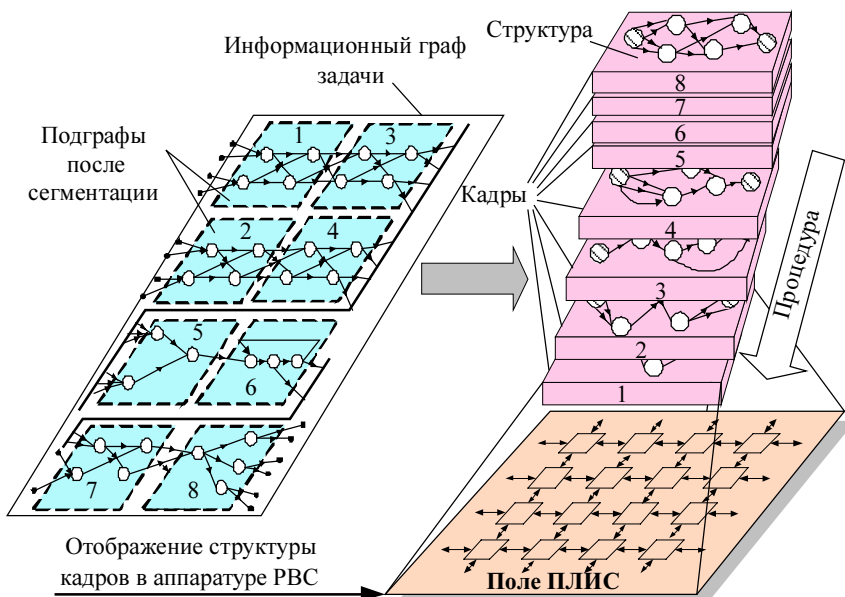


Рис. 1. Процесс решения задачи в РВС

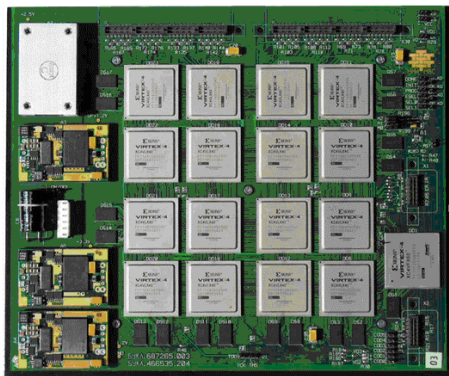
РВС со структурно-процедурной организацией вычислений является гибридом фоннеймановской архитектуры и архитектуры потока данных. Такая организация вычислений обеспечивает детерминизм выполнения программы, что в общем случае недостижимо в многопроцессорных системах, построенных по традиционной мультипроцедурной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

В течение 10 лет в НИИ МВС ЮФУ разрабатывались и создавались различные реконфигурируемые вычислительные системы, построенные по принципу модульного наращивания.

Фотографии базового модуля и модульно-наращиваемой РВС показаны на рис. 2.

Результаты теоретических и прикладных исследований позволяют сделать вывод об эффективности использования концепции динамической перестройки архитектуры системы для создания целого семейства высокопроизводительных модульно-наращиваемых многопроцессорных суперкомпьютеров широкого применения на основе ПЛИС-технологий (от суперЭВМ и рабочих станций до ускорителей персональных компьютеров), значительно опережающих по своим техническим характеристикам зарубежные аналоги, и в то же время отличающихся существенно более низкой стоимостью и сохранением преемственности методов и средств программирования.

В качестве прототипа реконфигурируемого ускорителя персонального компьютера может выступать изделие «Мангуст», внешний вид которого показан на рис. 3.



Базовый модуль 16M50



Модульно-наращиваемая PBC

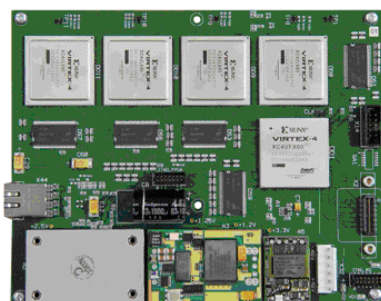
Характеристики базового модуля

Число ПЛИС Virtex IV, шт.	16
Объем памяти, ГБайт	1
Производительность, Гфлопс	50
Рабочая частота, МГц	160
Количество каналов LVDS, шт.	112
Суммарная пропускная способность LVDS каналов, Мбит/сек	22400
Потребляемая мощность, КВт	0,16
Габариты, мм	233,35 × 280

Характеристики изделия

Число унифицированных вычислительных БМ, шт.	4
Число вычислительно реконфигурируемых блоков, шт.	64
Объем памяти, Гбайт	4
Производительность, ГФлопс	200
Потребляемая мощность, КВт	0,9
Объем, л	39,9

Рис. 2. Базовый модуль 16M50 и модульно-наращиваемая PBC



Базовый модуль «МАНГУСТ»



Ускоритель персонального компьютера «МАНГУСТ»

Характеристики ускорителя персонального компьютера

Число реконфигурируемых вычислительных блоков, шт.	4
Число эквивалентных вентиляей, млн.шт.	32
Объем динамической памяти, Мбайт	576
Производительность, Гфлопс	25
Частота базового модуля, МГц	160
Скорость обмена данными по каналу Ethernet, Гбит/с	1
Потребляемая мощность, КВт	0,145
Габариты, мм	300 × 263 × 55

Рис. 3. Базовый модуль «МАНГУСТ» и ускоритель персонального компьютера «МАНГУСТ»

В настоящее время в НИИ МВС ЮФУ выполняется госконтракт 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 ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007 – 2012 годы».

В рамках госконтракта будет создано семейство РВС: высокопроизводительные системы отраслевого уровня производительностью 5 Тфлопс (РВС-5), уровня крупных организаций или научных центров производительностью 1 Тфлопс (РВС-1), рабочая станция уровня средних организаций производительностью 200 ГФлопс (РВС-0.2), а также ускорители для персональных компьютеров производительностью 50 и 25 ГФлопс (РУПК-50 и РУПК-25). Структура семейства РВС показана на рис. 4.

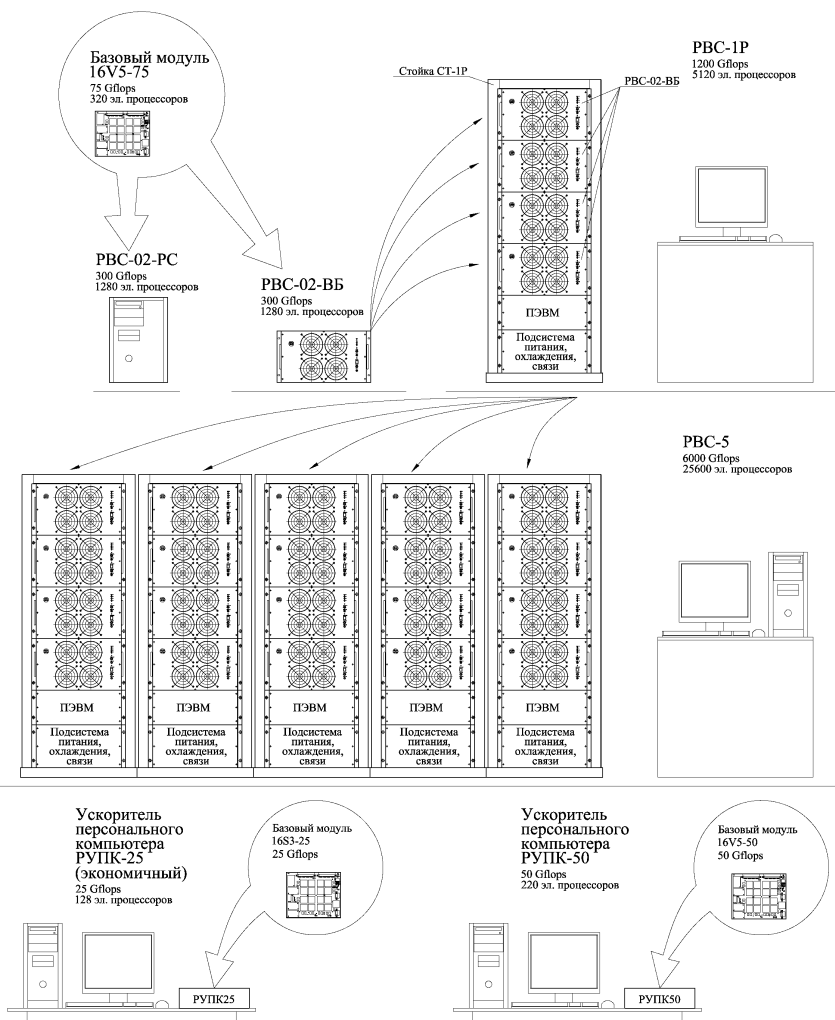


Рис. 4. Структура семейства РВС

Целесообразность проекта также обусловлена экономической выгодностью разработки данных технологий в Российской Федерации по сравнению с приобретением отдельных компонентов за рубежом, поскольку стоимость лицензий на их использование может составлять от нескольких сот тысяч до десятков миллионов долларов.

1. Базовые модули семейства PBC

Базовые модули, из которых строятся PBC, в полной мере сохраняют все концептуальные особенности архитектуры реконфигурируемых систем. Семейство PBC строится всего из трех типов базовых модулей: 16V5-75, 16V5-50 и 16S3-25. Базовый модуль 16V5-75, как самый производительный, используется в моделях PBC-5, PBC-1P и PBC-0.2-PC. Базовые модули 16V5-50 и 16S3-25 входят в состав ускорителей персональных компьютеров РУПК-50 и РУПК-25. Общие принципы построения базовых модулей одинаковы, и поэтому рассмотрим более подробно базовый модуль 16V5-75.

Базовый модуль 16V5-75 обладает характеристиками, представленными в таблице.

Характеристики базового модуля 16V5-75

Параметр	Значение
Производительность (64 разряда), Гфлопс	75
Производительность (32 разряда), Гфлопс	140
Потребляемая мощность, ВА	200
Объем оперативной распределенной памяти, Гбайт	1,25
ПЛИС решающего поля Virtex 5 XC5VLX110, шт.	16
Количество эквивалентных вентилях в ПЛИС, шт.	$8 \cdot 10^6$
Тактовая частота, МГц	250
Количество внешних LVDS каналов, шт.	224
Скорость межмодульного обмена, Мбит/с	22000

Решающее поле базового модуля 16V5-75 выполнено на 16-ти ПЛИС Virtex 5 XC5VLX110-2FF1153 фирмы Xilinx, каждая из которых содержат около $8 \cdot 10^6$ эквивалентных вентилях.

Микросхемы расположены в узлах двумерной решетки 4×4 и соединены между собой ортогональной системой связей по близкодействию. Такая система связей позволяет существенно упростить печатную плату и улучшить ее частотные характеристики, поскольку связи между соседними ПЛИС не превышают четырех сантиметров. Данные между несмежными микросхемами передаются по транзитным каналам через промежуточные ПЛИС с использованием системы ортогональных связей.

Связи между ПЛИС решающего поля базового модуля реализованы на основе стандарта LVDS. Преимуществами стандарта LVDS являются низкая потребляемая мощность выходных каскадов, низкий уровень создаваемых электромагнитных излучений, невосприимчивость к синфазным электромагнитным помехам и наличие в микросхемах семейства Vertex 5 аппаратной поддержки для организации высокоскоростных передач данных на основе стандарта LVDS. Физически шины связи представляют собой набор пар дифференциальных полосковых передающих линий, с обоих концов подключенных к определенным выводам микросхем. Для надежной передачи данных по LVDS в 16V5-75 задействованы специальные ресурсы семейства микросхем Vertex 5, поддерживающие алгоритм опти-

мальной битной и кадровой синхронизации и позволяющие учесть все нюансы линии передачи.

Базовый модуль имеет подсистему обмена информацией по быстрым LVDS-каналам с другими базовыми модулями, которая предназначена для передачи промежуточных результатов вычислений непосредственно из микросхем данного модуля непосредственно в микросхемы решающего поля других базовых модулей. Передача осуществляется посредством специальных кабелей, подключаемых к соединителям типа QTE-056 (рис. 5). 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 схмотехнического ресурса РВС.

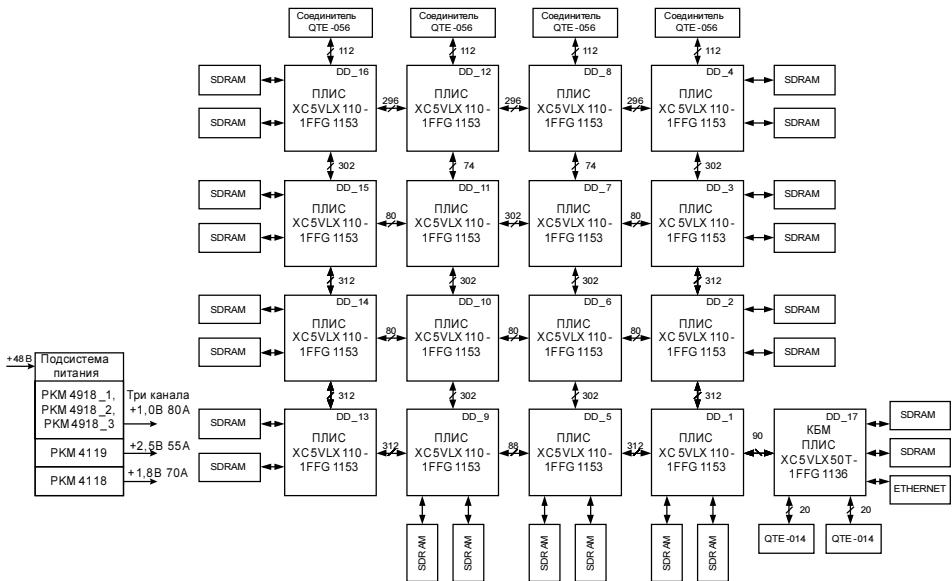


Рис. 5. Структура базового модуля 16V5-75

К периферийным ПЛИС решающего поля, расположенных на краях двумерной решетки 4×4 , подключены микросхемы динамических ОЗУ SDRAM типа DDR2 (Double Data Rate Two Synchronous Dynamic Random Access Memory), которые образуют распределенную память базового модуля 16V5-75. Для организации распределенной памяти используется 20 микросхем MT47N32M16BN фирмы Micron Technology Inc. Каждая микросхема имеет структуру 2^{25} 16-разрядных слов. Общий объем распределенной памяти составляет 1,25 Гбайт. Кроме того, на базовом модуле установлено еще 128 Мбайт памяти, подключенной к контроллеру базового модуля – КБМ. Распределенная память предназначена для хранения входных данных, промежуточных и конечных результатов обработки информации, а также фрагментов параллельных программ.

Основным назначением КБМ являются функции управления подсистемами базового модуля, а также передача информации между базовым модулем и host-машиной, в том числе загрузка конфигурации ПЛИС. Помимо этого, КБМ участвует в диагностике устройств базового модуля, настраивает подсистему синхро-

низации, собирает и передает информацию о текущем состоянии базового модуля в управляющий контроллер. КБМ выполнен на ПЛИС Virtex 5 XC5VLX50T-1FFG1136 фирмы Xilinx. Настройка конфигурации этой микросхемы осуществляется по включению питания из энергонезависимой памяти, в которой хранится конфигурационный файл КБМ. Связь КБМ с управляющим контроллером осуществляется посредством LVDS-каналов через два соединителя QTE-014, а также с помощью канала Ethernet. Функция контроля температуры ПЛИС решающего поля осуществляется КБМ путем опроса аппаратных блоков System Monitor, изменяющих температуру кристаллов, которые встроены в ПЛИС семейства Virtex 5.

Подсистема электропитания расположена непосредственно на базовом модуле и содержит портативные преобразователи напряжения серии РКМ фирмы Ericsson, на которые подается напряжение +48 В от источника первичного питания. Для отвода тепла и поддержания необходимых температурных режимов в микросхемах базового модуля предусмотрена комбинированная система охлаждения, которая включает радиаторы на ПЛИС решающего поля и вентиляторы для их обдува. На рис. 6 представлен внешний вид печатной платы базового модуля РВС нового поколения.

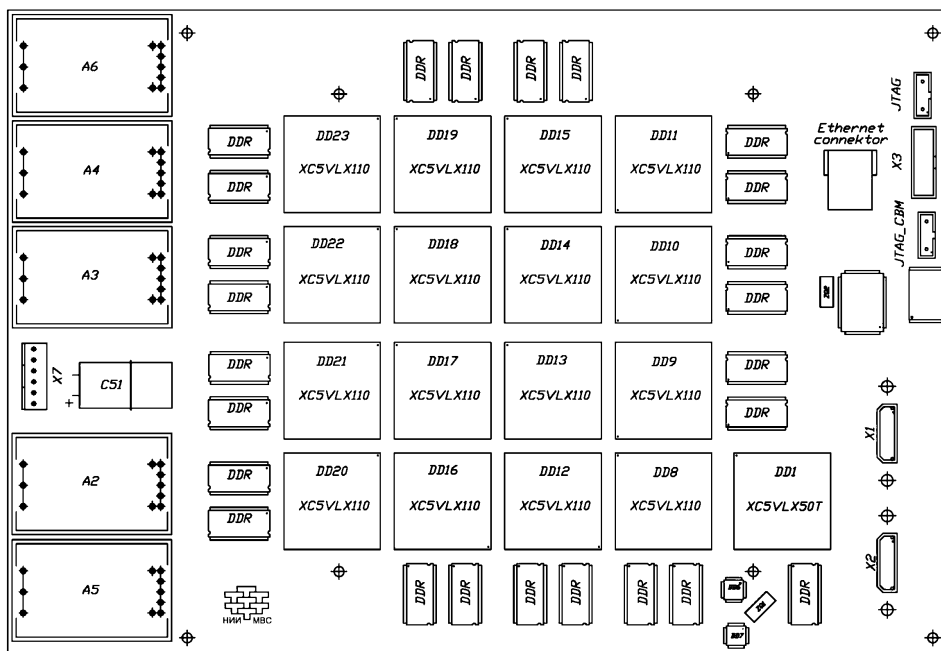


Рис. 6. Внешний вид печатной платы базового модуля РВС нового поколения

Базовый модуль 16V5-75 представляет собой мощный вычислительный узел производительностью свыше 75 Гфлопс. На его основе могут строиться вычислительные блоки, содержащие от одного до восьми базовых модулей производительностью от 75 до 600 Гфлопс. В то же время базовый модуль обладает достаточной автономностью и может легко комплексоваться с персональным компьютером типа IBM PC в качестве ускорителя при решении различных задач.

2. Структура PBC-0.2-PC и PBC-0.2-VB

На основе базового модуля 16V5-75 в рамках проекта разработаны рабочая станция PBC-0.2-PC и вычислительный блок PBC-0.2-VB производительностью 300 Гфлопс. Основу этих изделий составляет многопроцессорный вычислительный блок (МВБ), который включает четыре базовых модуля, соединенных между собой в единый вычислительный ресурс быстрыми каналами LVDS и посредством коммутатора Ethernet.

Рабочая станция PBC-0.2-PC является представителем семейства PBC и предназначена для решения прикладных задач проектирования изделий микроэлектроники, управления в реальном времени сложными объектами, моделирования сложных технических и природных объектов и процессов, построения систем мониторинга, дистанционного зондирования и томографии и др.

Вычислительный блок PBC-0.2-VB практически полностью повторяет архитектуру рабочей станции PBC-0.2-PC, однако конструкции этих изделий значительно отличаются. Конструктивные отличия определяются назначением этих изделий: рабочая станция – это настольный вариант вычислительной системы, предназначенный для автономного использования, а вычислительный блок – это встраиваемый вариант, предназначенный для комплектования стоек СТ-1P в составе представителей семейства PBC-1P и PBC-5 и для создания суперЭВМ различных конфигураций. Вычислительный блок PBC-0.2-VB обладает теми же техническими параметрами, что и рабочая станция PBC-0.2-PC, и предназначен для решения перечисленных выше задач в составе PBC-1P и PBC-5.

Реальная производительность PBC-0.2-PC и PBC-0.2-VB на задачах линейной алгебры, математической физики, цифровой обработки сигналов, символьной обработки составляет не менее 60% от пиковой производительности. Аппаратно-программные средства PBC-0.2-PC и PBC-0.2-VB позволяют динамически перестраивать архитектуру в процессе решения задачи на двух уровнях: программном (на уровне элементарных процессоров и каналов распределенной памяти), обеспечивающем высокую скорость реконфигурации системы на задачи из данного класса; схемотехническом (на уровне логических ячеек ПЛИС), обеспечивающем модернизацию системы команд элементарных процессоров и высокую удельную производительность системы при переходе на задачи различных классов.

PBC-0.2-PC и PBC-0.2-VB обладают высокой надежностью и готовностью к работе.

Конструктивные отличия PBC-0.2-VB от PBC-0.2-PC заключаются в особенностях соединения базовых модулей в составе МВБ. На рис. 7, а и б показана укрупненная структура рабочей станции PBC-0.2-PC и МВБ вычислительного блока PBC-0.2-VB. В МВБ рабочей станции базовые модули соединяются в кольцо, а в МВБ вычислительного блока базовые модули БМ0 и БМ3 имеют выходы за пределы PBC-0.2-VB с целью комплексирования нескольких вычислительных блоков PBC-0.2-VB в единую вычислительную структуру в составе стойки СТ-1P. Во всем остальном архитектура PBC-0.2-VB совпадает с архитектурой PBC-0.2-PC.

Управляющий контроллер, входящий в состав обоих изделий, представляет собой персональный компьютер типа IBM PC, встроенный в корпус PBC-0.2-PC или PBC-0.2-VB и функционирующий под управлением операционной системы Windows XP. Управляющий контроллер предназначен для управления всеми компонентами PBC-0.2-PC или PBC-0.2-VB; хранения информации (исходной, промежуточной и результатов); трансляции параллельных программ с языков высоко-

кого уровня в исполняемые коды; тестирования и диагностики; подключения периферийного оборудования; подключения к локальным и глобальным информационным сетям посредством Ethernet.

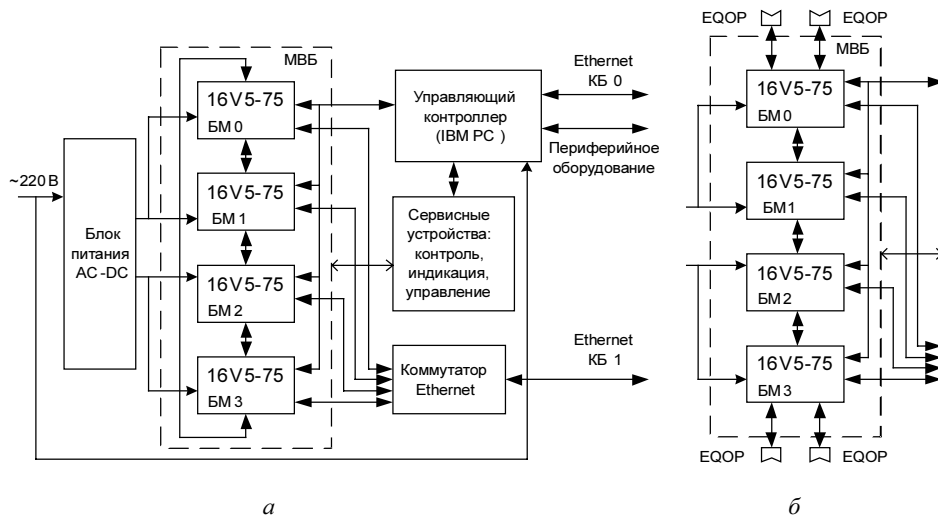


Рис. 7. Структура рабочей станции РВС-0.2-РС (а) и МВБ вычислительного блока РВС-0.2-ВБ (б)

Системы охлаждения РВС-0.2-РС и РВС-0.2-ВБ имеют некоторые отличия, связанные с особенностями их назначения и эксплуатации. Автономная работа рабочей станции позволяет использовать для охлаждения ПЛИС базовых модулей медные штыревые радиаторы и проточно-вытяжную вентиляцию корпуса РВС-0.2-РС, что, в свою очередь, приводит к уменьшению габаритов рабочей станции по сравнению с РВС-0.2-ВБ, а также к уменьшению потребляемой мощности и шумности. Вычислительный блок работает в более жестких условиях по сравнению с рабочей станцией из-за наличия фонового перегрева, создаваемого другими блоками, и необходимо прогонять воздушный поток не только через корпус блока, но и через корпус стойки, где он встречает дополнительное сопротивление. Поэтому для РВС-0.2-ВБ используется проточная вентиляция корпуса, дополненная вентиляторами прямого обдува, установленными непосредственно на медных штыревых радиаторах ПЛИС базовых модулей.

3. Реконфигурируемая вычислительная система РВС-1Р

Реконфигурируемая вычислительная система РВС-1Р предназначена: для оснащения научных центров с целью проведения исследований в области физики, химии, биологии, космоса, построения информационно-управляющих систем для управления потенциально опасными производствами, решения задач аэрокосмической, автомобильной промышленности и энергетики. Пиковая производительность РВС-1Р составляет 1200 Гфлопс. Основу РВС-1Р составляют вычислительная стойка СТ-1Р, в которую монтируются четыре вычислительных блока РВС-0.2-ВБ, ПЭВМ, коммутатор Ethernet, система питания и система охлаждения. Структурная схема РВС-1Р показана на рис. 8.

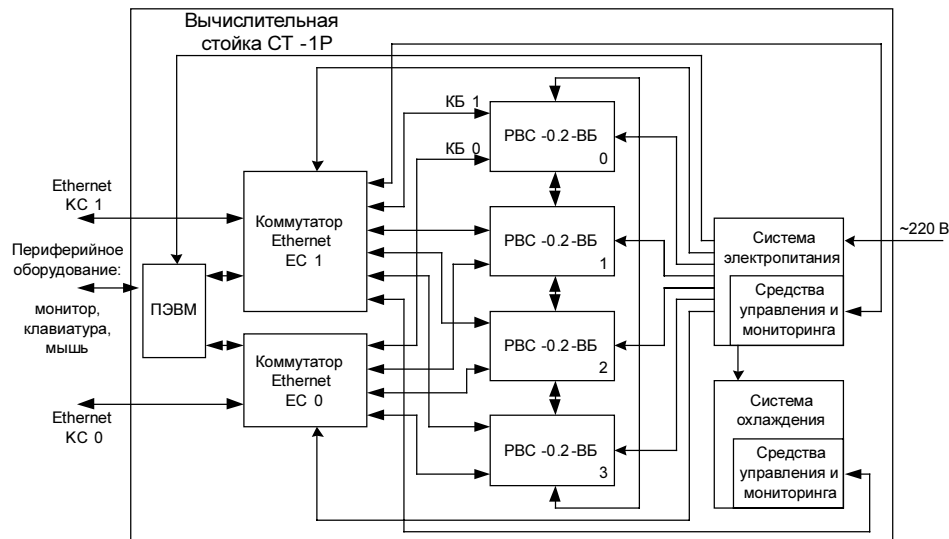


Рис. 8. Структурная схема PBC-1P

Четыре вычислительных блока PBC-0.2-BB объединяются с помощью быстрых LVDS-каналов в единый вычислительный ресурс, содержащий до 16-ти базовых модулей 16V5-75 с общей пиковой производительностью 1,2 Тфлопс. Межблочные связи являются продолжением межмодульных и, в свою очередь, продолжением связей между ПЛИС решающих полей базовых модулей. В целом подобная организация логических связей реализует в составе стойки СТ-1P глобальный LVDS-канал передачи данных с единым темпом продвижения информации.

Для осуществления функций управления и мониторинга управляющие контроллеры вычислительных блоков по выходам KB1 соединяются между собой посредством коммутатора EC1 под общим управлением ПЭВМ. Сетевые выходы KB0 блока PBC-0.2-BB объединяются сетевым коммутатором стойки EC0, посредством которого можно установить прямые связи с любым из базовых модулей в составе стойки, минуя управляющие контроллеры блоков. Такое соединение вычислительных ресурсов позволит максимально эффективно использовать возможности реконфигурируемой элементной базы вычислительной системы PBC-1P. ПЭВМ предназначена для управления работой четырех вычислительных блоков PBC-0.2-BB, управления работой коммутаторов Ethernet, управления работой системы питания и системы охлаждения, подключения периферийного оборудования, а также соединения реконфигурируемой вычислительной системы к локальным и глобальным информационным сетям.

4. Реконфигурируемая вычислительная система PBC-5

Старшим представителем семейства PBC является изделие PBC-5 с пиковой производительностью 6000 Гфлопс. Реконфигурируемая вычислительная система PBC-5 предназначена для использования в научно-исследовательских центрах с целью проведения исследований в области молекулярного моделирования, создания принципиально новых лекарственных препаратов и материалов нового поколения, построения информационно-моделирующих систем для долгосрочного

прогнозирования погоды, моделирования задач сейсмоактивности, томографических исследований приповерхностных слоев Земли акустическими и электромагнитными волнами. Упрощенная структурная схема реконфигурируемой вычислительной системы РВС-5 показана на рис. 9. РВС-5 содержит пять стоек СТ-1Р, управляющую ЭВМ (УЭВМ), систему питания, систему охлаждения.

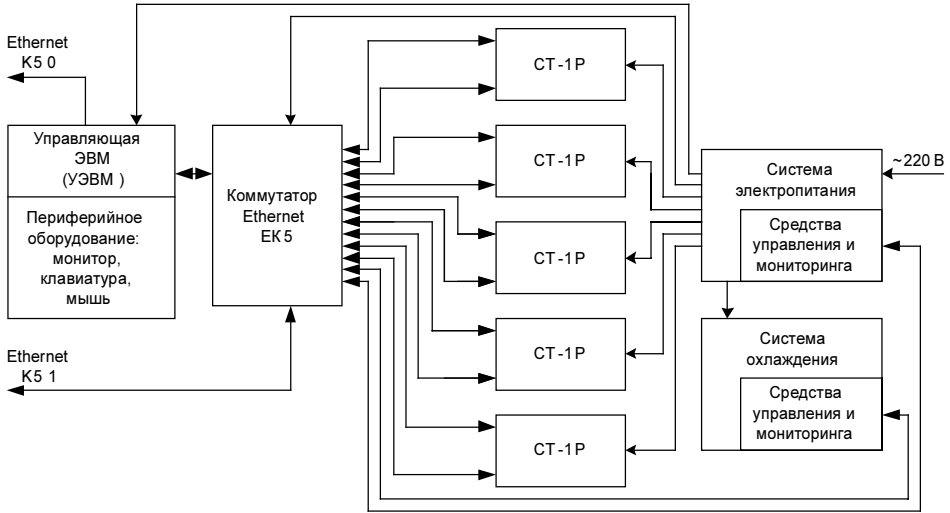


Рис. 9. Упрощенная структурная схема РВС-5

Основным вычислительным ресурсом РВС-5 являются пять стоек СТ-1Р, которые соединяются между собой посредством коммутаторов Ethernet с использованием сетевых технологий под общим управлением УЭВМ. Такое построение

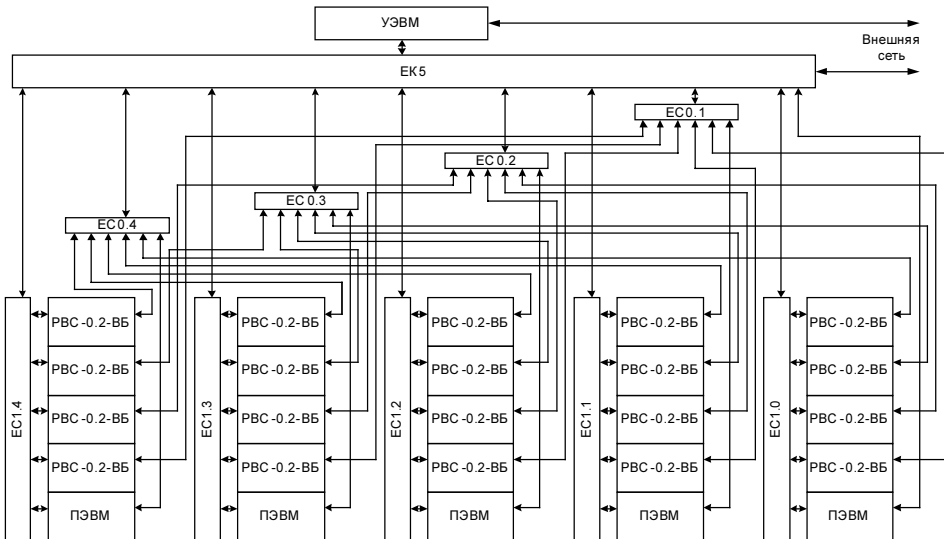


Рис. 10. Сетевое соединение вычислительных ресурсов РВС-5

аппаратных средств позволяет, с одной стороны, связать воедино все вычислительные ресурсы стоек СТ-1Р, а с другой – максимально унифицировать оборудование, поскольку вычислительная часть РВС-5 строится из типовых узлов, уже использовавшихся для построения других представителей семейства РВС – базовых модулей 16V5-75, вычислительных блоков РВС-0.2-ВБ, стоек СТ-1Р. Один из возможных вариантов соединения вычислительных ресурсов РВС-5 показан на рис. 10. В этом случае коммутатор Ethernet EC0 в составе стоек СТ-1Р подключается так, чтобы создать «горизонтальные» связи между блоками РВС-0.2-ВБ различных стоек системы РВС-5.

5. Реконфигурируемые ускорители персонального компьютера РУПК-50 и РУПК-25

Основным назначением реконфигурируемых ускорителей персонального компьютера РУПК-25 и РУПК-50 является наращивание возможностей персональных компьютеров типа IBM PC при решении вычислительно трудоемких задач, проектировании изделий микроэлектроники, математического моделирования сложных технических и природных объектов и процессов, символьной обработки информации, оптимизации при эксплуатации нефтяных месторождений, дистанционного зондирования и томографии и др.

РУПК-50 и РУПК-25 предоставляют пользователю дополнительный вычислительный ресурс с пиковой производительностью свыше 50 и 25 Гфлопс соответственно. Реальная производительность РУПК-50 и РУПК-25 на задачах линейной алгебры, математической физики, цифровой обработки сигналов, символьной обработки превышает 60% от пиковой производительности. РУПК-50 и РУПК-25 имеют одинаковую структуру (рис. 11) и различаются только типом используемого базового модуля.

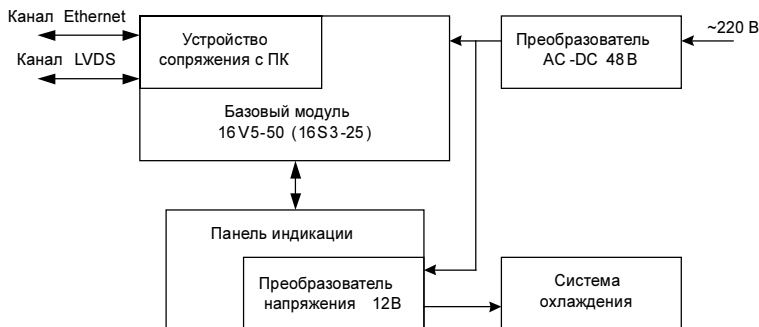


Рис. 11. Структурная схема РУПК-50 (РУПК-25)

РУПК-50 (РУПК-25) содержит: базовый модуль 16V5-50 (16S3-25) с устройством сопряжения с персональным компьютером; блок питания АС-DC; панель индикации; систему охлаждения.

Базовые модули с динамически перестраиваемой архитектурой на основе ПЛИС для реконфигурируемых ускорителей персональных компьютеров строятся на тех же принципах, что и базовый модуль 16V5-75 старших представителей семейства РВС-0.2-РС, РВС-1Р и РВС-5. Основным отличием от базового модуля 16V5-75 является более низкая стоимость и меньшая потребляемая мощность при

сохранении высокой производительности. Поэтому базовые модули для РУПК-50 и РУПК-25 строятся на основе более дешевых электронных компонентов.

Решающее поле базового модуля 16V5-50 выполнено на 16-ти ПЛИС Virtex 5 XC5VLX110-1FFG1153 фирмы Xilinx, контроллер базового модуля – на ПЛИС XC5VLX50T-1FFG1136. Распределенная память реализована на 24 микросхемах SDRAM типа DDR2 и имеет общий объем 1,5 Гбайта. В отличие от базового модуля 16V5-75 базовый модуль 16V5-50 не содержит каналов связи с другими базовыми модулями. Производительность базового модуля 16V5-50 составляет 50 Гфлопс.

Решающее поле базового модуля 16S3-25 выполнено на 16-ти ПЛИС Spartan 3 XC3S4000-5FG1156 фирмы Xilinx, каждая из которых содержат около $4 \cdot 10^6$ эквивалентных вентиляей. Контроллер базового модуля выполнен на ПЛИС Virtex 4 XC4VFX60-10FF1152C. Распределенная память реализована на 24 микросхемах SDRAM типа DDR2 и имеет общий объем 1,5 Гбайта. Производительность базового модуля 16S3-25 составляет 25 Гфлопс. Обмен информацией РУПК-50 и РУПК-25 с персональным компьютером по данным и управлению осуществляется с помощью ПЛИС КБМ с использованием LVDS- или Ethernet-каналов.

Система охлаждения поддерживает необходимый температурный режим электронных компонентов РУПК-50 (РУПК-25) и включает радиаторы и вентиляторы на температурно-напряженных микросхемах базового модуля, а также вентиляторы для принудительной прокачки воздуха в пределах всего объема корпуса.

6. Системное программное обеспечение

Программирование РВС отличается от программирования МВС традиционной архитектуры и его можно условно разделить на две составляющие: программирование структурное, которое создает необходимые вычислительные структуры в поле логических ячеек ПЛИС, и программирование процедурное – программирование в традиционном смысле, заключающееся в организации вычислительного процесса в РВС. При этом программирование вычислительных структур вызывает у пользователей наибольшие трудности. 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 привлекательными реконфигурируемые на низком уровне системы, а с другой – их программирование становится по сложности сопоставимым с созданием новой вычислительной системы. Такой подход требует новых методов и средств организации параллельных вычислительных процессов.

Целью создания программного комплекса средств разработки прикладных программ семейства PBC 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 микропроцессоров и микропроцессорных ЭВМ. Основными задачами программного комплекса средств разработки является 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 реализацией необходимых вычислительных структур на множестве ПЛИС.

Программный комплекс средств разработки прикладных программ семейства PBC 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, единую для всех представителей семейства PBC и поддерживающую языки программирования Argus v.3.0 и COLAMO v.2.0; отладчик параллельных программ на базовых модулях семейства PBC, поддерживающий межмодульные связи; программный интерфейс доступа к вычислительным ресурсам всех представителей PBC 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, 5] обеспечивает синтаксическую поддержку реконфигурации аппаратной платформы PBC и возможность использования элементов библиотеки масштабируемых IP-ядер. Транслятор COLAMO v.2.0 для всех представителей семейства PBC 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 всех представителей семейства PBC, 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.

Одновременно с разработкой семейства PBC создается и будет развиваться инфраструктура для обучения потенциальных пользователей и организации доступа к библиотекам масштабируемых IP-ядер для различных предметных областей с целью их широкого продвижения на рынке высокопроизводительной вычислительной техники. За счет использования возможностей сервисов глобальной компьютерной сети Internet планируется обеспечить широкий доступ потенциальных пользователей к информационным ресурсам и программным продуктам, созданным в результате выполнения проекта.

Заключение

Описанное в статье семейство РВС является принципиально новым направлением в создании высокопроизводительных вычислительных систем, в которых в качестве основного вычислительного элемента используются не универсальные микропроцессоры, а ПЛИС-микросхемы. Это дает возможность пользователям создавать в базовой архитектуре РВС 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. Семейство РВС с динамически перестраиваемой архитектурой на основе ПЛИС предназначено для решения вычислительно трудоемких задач, решение которых на МВС традиционной архитектуры либо требует недопустимо больших временных затрат, либо вообще невозможно.

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 семейство РВС, описание которого посвящена данная статья, обладает рядом преимуществ перед многопроцессорными системами традиционной организации, основанными на использовании микропроцессоров в качестве основного вычислительного элемента. Реконфигурируемые системы с высокой реальной производительностью позволяют решить важнейшие научно-технические задачи и обеспечить прорывы в таких наукоемких технологических областях, как молекулярная фармакология, наноэлектроника, создание энергетических комплексов нового поколения, а также проведение фундаментальных научных исследований в астрофизике, микробиологии, физике твердого тела, томографическом исследовании приповерхностных слоев Земли.

ЛИТЕРАТУРА

1. Волков Д. Реальность и фантазии // Открытые системы. 2006. № 5.
2. Аладышев О.С., Дикарев Н.И., Овсянников А.П. и др. СуперЭВМ: области применения и требования к производительности // Изв. вузов. Электроника. 2004. № 1. С. 13 – 17.
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. Левин И.И. Язык параллельного программирования высокого уровня для структурно-процедурной организации вычислений // Тр. Всерос. науч. конф. М.: Изд-во МГУ, 2000. С. 108 – 112.

Статья представлена кафедрой информационных технологий в исследовании дискретных структур радиофизического факультета Томского государственного университета и оргкомитетом 7 Российской конференции с международным участием «Новые информационные технологии в исследовании сложных структур», поступила в научную редакцию 10 мая 2008 г.