

СЕМЕЙСТВО МНОГОПРОЦЕССОРНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ НА ОСНОВЕ ПЛИС

А.И. Дордопуло¹, И.А. Каляев², И.И. Левин², Е.А. Семерников¹

¹Южный научный центр РАН, г. Ростов-на-Дону;

²НИИ многопроцессорных вычислительных систем имени академика А.В. Каляева
Южного федерального университета, г. Таганрог

В статье приводится описание семейства многопроцессорных вычислительных систем с динамически перестраиваемой архитектурой на основе реконфигурируемой элементной базы производительностью до 6 Тфлопс, создаваемого в рамках Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2012 годы». В качестве основного вычислительного элемента в них используются не универсальные микропроцессоры, а программируемые логические интегральные схемы сверхбольшой интеграции.

Введение

В настоящее время наибольшее распространение в мире получили суперЭВМ кластерного типа. Это связано с тем, что они строятся из серийно выпускаемых узлов и основываются на массовом использовании универсальных микропроцессоров традиционной архитектуры в качестве основного вычислительного элемента. В то же время пользователи отмечают принципиальные недостатки кластерных суперЭВМ, которые заключаются в существенном снижении их производительности при решении многих практических задач. Так, например, 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].

Недостатки кластерных систем, связанные с их неизменной «жесткой» архитектурой, позволяет преодолеть концепция создания реконфигурируемых вычислительных систем (РВС) с «гибкой», динамически перестраиваемой (программируемой) архитектурой [2, 3]. 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 функционирования. В результате у пользователя появляется возможность адаптации архитектуры вычислительной системы под структуру решаемой 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 [2, 3], что реализация данной концепции обеспечивает высокую реальную производительность РВС, близкую к пиковой, на широком классе задач, в том числе при решении «сильносвязанных» задач.

Состав и характеристики семейства РВС

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 реконфигурируемых высокопроизводительных вычислительных систем производительностью от 0,025 Тфлопс до 6 Тфлопс.

В рамках государственного контракта будет создано семейство РВС:

- РВС-5 – высокопроизводительная система производительностью 6 Тфлопс;
- РВС-1Р и РВС-1К – системы производительностью более 1 Тфлопс;
- РВС-0.2-РС – рабочая станция производительностью 300 ГФлопс;
- РУПК-50 и РУПК-25 – ускорители персональных компьютеров

производительностью 50 и 25 ГФлопс.

Для создания семейства РВС, системного и прикладного программного обеспечения, 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).

Старшие представители семейства создаются на принципах модульной наращиваемости и будут обладать почти линейным ростом реальной производительности в зависимости от увеличения аппаратного ресурса [2]. Реальная производительность всех представителей семейства РВС 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.

Базовые модули семейства РВС

Семейство РВС строится всего из трех типов базовых модулей: 16V5-75, 16V5-50 и 16S3-25. Базовый модуль 16V5-75, как самый производительный, используется в моделях РВС-5, РВС-1Р и РВС-0.2-РС. 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 (см. рис. 1). Базовый модуль 16V5-75 обладает следующими характеристиками: производительность (64 разряда) – 75 Гфлопс; производительность (32 разряда) – 140 Гфлопс; потребляемая мощность – 200 ВА; объем оперативной распределенной памяти – 1,25 Гбайт; количество внешних LVDS каналов – 224 шт.; скорость межмодульного обмена – 22000 Мбит/сек.

Шестнадцать ПЛИС Virtex 5 XC5VLX110-2FF1153, образующие решающее поле, расположены в узлах двумерной решетки 4 x 4 и соединены между собой ортогональной системой связей по близкодействию посредством каналов LVDS. Такая система связей позволяет существенно упростить печатную плату и улучшить ее

частотные характеристики. Данные между несмежными микросхемами передаются по транзитным каналам через промежуточные ПЛИС.

Базовый модуль имеет подсистему обмена информацией с другими базовыми модулями по LVDS-каналам, обеспечивающим передачу промежуточных результатов вычислений непосредственно из микросхем данного модуля в микросхемы решающего поля других базовых модулей в пределах общего схемотехнического ресурса РВС.

К периферийным ПЛИС решающего поля подключены микросхемы динамических ОЗУ SDRAM типа DDR2, которые образуют распределенную память базового модуля 16V5-75. Распределенная память предназначена для хранения входных данных, промежуточных и конечных результатов обработки информации, а также фрагментов параллельных программ. На базовом модуле имеется еще 128 Мбайт памяти, подключенной к контроллеру базового модуля – КБМ.

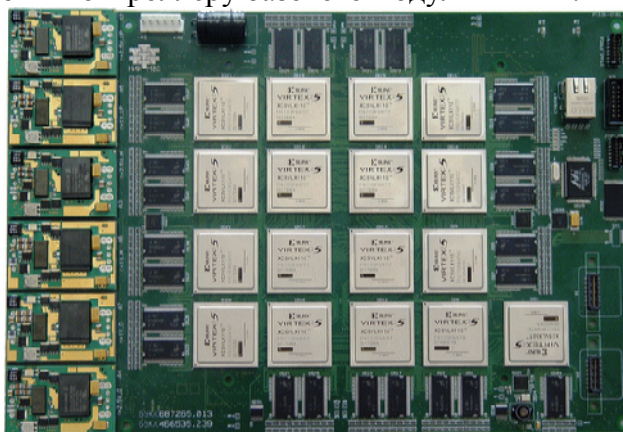


Рис. 1. Базовый модуль 16V5-75

Основным назначением КБМ являются функции управления подсистемами базового модуля, а также функции связи базового модуля с управляющим контроллером (ЭВМ типа IBM PC) посредством LVDS-каналов и канала Ethernet.

Подсистема электропитания расположена непосредственно на базовом модуле и содержит портативные преобразователи напряжения серии РКМ фирмы Ericsson.

Таким образом, базовый модуль 16V5-75 представляет собой мощный вычислительный ресурс производительностью свыше 75 Гфлопс. На основе БМ 16V5-75 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 и использоваться при решении различных задач.

Рабочая станция РВС-0.2-РС и вычислительный блок РВС-0.2-ВБ

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

Рабочая станция РВС-0.2-РС является представителем семейства РВС и предназначена для решения прикладных задач различных предметных областей.

Вычислительный блок РВС-0.2-ВБ практически полностью повторяет архитектуру рабочей станции РВС-0.2-РС, однако конструкции этих изделий значительно отличаются, поскольку рабочая станция – это настольный вариант

вычислительной системы, предназначенный для автономного использования, а вычислительный блок – это встраиваемый вариант, предназначенный для комплектования стоек СТ-1Р в составе представителей семейства РВС-1Р и РВС-5 и для создания суперЭВМ различных конфигураций.

Аппаратно-программные средства РВС-0.2-РС и РВС-0.2-ВБ позволяют динамически перестраивать их архитектуру в процессе решения задачи на двух уровнях: программном (на уровне команд 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.

Архитектурные отличия РВС-0.2-ВБ от РВС-0.2-РС заключаются в особенностях соединения базовых модулей. В РВС-0.2-РС базовые модули соединяются в кольцо, а в РВС-0.2-ВБ базовые модули имеют выходы за пределы корпуса блока с целью комплексирования нескольких вычислительных блоков РВС-0.2-ВБ в единую вычислительную структуру в составе стойки СТ-1Р.

Управляющий контроллер, входящий в состав обоих изделий, представляет собой персональный компьютер типа IBM PC, встроенный в корпус РВС-0.2-РС или РВС-0.2-ВБ и функционирующий под управлением операционной системы Windows XP, предназначен для управления всеми компонентами РВС-0.2-РС или РВС-0.2-ВБ, хранения информации, трансляции параллельных программ, тестирования и диагностики, подключения периферийного оборудования, подключения к локальным и глобальным информационным сетям посредством Ethernet.

Реконфигурируемые вычислительные системы РВС-1Р и РВС-5

Представители семейства РВС-1Р и РВС-5 предназначены для оснащения научных центров и имеют пиковую производительность 1,2 и 6,0 Тфлопс соответственно.

Основу РВС-1Р и РВС-5 составляет вычислительная стойка СТ-1Р, в которую монтируются четыре вычислительных блока РВС-0.2-ВБ, ПЭВМ, коммутатор Ethernet, системы питания и охлаждения. Четыре вычислительных блока РВС-0.2-ВБ 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-1Р глобальный LVDS-канал передачи данных с единым темпом продвижения информации. Для осуществления функций управления и мониторинга вычислительные блоки соединяются между собой посредством коммутаторов Ethernet.

В состав РВС-1Р входят одна стойка СТ-1Р и периферийное оборудование.

РВС-5 с пиковой производительностью 6000 Гфлопс является старшим представителем семейства РВС и содержит пять стоек СТ-1Р. В состав РВС-5 помимо этого входит управляющая ЭВМ. Управляющая ЭВМ представляет собой компьютер типа IBM PC и предназначена для управления работой ресурсов вычислительной системы РВС-5, а также для ее подключения к локальным и глобальным информационным сетям посредством Ethernet.

Ускорители персонального компьютера РУПК-50 и РУПК-25

Основным назначением реконфигурируемых ускорителей РУПК-25 и РУПК-50 с пиковой производительностью 25 и 50 Гфлопс является наращивание возможностей

персональных компьютеров типа IBM PC при решении вычислительно трудоемких задач. Аппаратно-программные средства РУПК-25 и РУПК-50 позволяют динамически перестраивать их архитектуру в процессе решения задачи на программном и схемотехническом уровнях. РУПК-50 и РУПК-25 имеют одинаковую структуру, содержат одинаковые узлы и различаются только типом базового модуля. РУПК-50 содержит: базовый модуль 16V5-50, выполненный на 16-ти ПЛИС Virtex 5 XC5VLX110-1FFG1153, а базовый модуль 16S3-25 – на 16-ти ПЛИС Spartan 3 XC3S4000-5FG1156 фирмы Xilinx. Обмен информацией РУПК-50 и РУПК-25 с персональным компьютером по данным и управлению осуществляется с помощью ПЛИС КБМ, используя LVDS или Ethernet-каналы.

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, заключающиеся в организации вычислительного процесса в РВС. При этом программирование вычислительных структур вызывает у пользователей наибольшие трудности [2, 3].

В настоящее время ведется создание комплекса средств разработки, который обеспечит потенциальным пользователям удобство программирования сложных практических задач на РВС. Создаваемый комплекс предоставит пользователю РВС возможности, которые позволят создавать программы без привлечения специальных знаний в области схемотехники ПЛИС и по сложности будут приближены к обычному программированию для многопроцессорных ЭВМ.

Заключение

Описанное семейство РВС является принципиально новым направлением развития высокопроизводительной вычислительной техники. В качестве основного вычислительного элемента в РВС используются не универсальные микропроцессоры, а ПЛИС. 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 предназначено для решения вычислительно трудоемких задач, решение которых на МВС традиционной архитектуры либо требует недопустимо больших временных затрат, либо вообще невозможно.

Список литературы

- 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. Каляев А.В., Левин И.И. Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. - М.: Янус-К, 2003. – 380 с.
3. Каляев И.А., Левин И.И., Семерников Е.А., Шмойлов В.И. Реконфигурируемые мультиконвейерные вычислительные структуры / Под общ. ред. И.А.Каляева. – Ростов/Д: Издательство ЮНЦ РАН, 2008. - 320 с.