

ВЫСОКОПРОИЗВОДИТЕЛЬНЫЕ РЕКОНФИГУРИРУЕМЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ НОВОГО ПОКОЛЕНИЯ

А.И. Дордопуло, И.А. Каляев, И.И. Левин, Е.А. Семерников

В последние годы во всем мире идет поиск новых архитектурных решений с целью достижения пиковых значений производительности многопроцессорных систем. Одним из наиболее распространенных решений является использование программируемых логических интегральных схем (ПЛИС) для выполнения вычислений. На второй строчке списка TOP-500 за ноябрь 2010 года значится суперЭВМ Jaguar - Cray XT5-HE, произведенная фирмой Cray Inc., с пиковой производительностью 2331.00 Тфлопс (в предыдущем списке TOP-500 за июнь 2010 года эта суперЭВМ находилась на первом месте), в составе которой в качестве сопроцессоров используются ПЛИС большой интеграции. В большинстве содержащих ПЛИС вычислительных систем, так же как и в Jaguar - Cray XT5-HE, кристаллы ПЛИС используются как дополнение к микропроцессорам, выполняющее трудно- или неэффективно реализуемые на универсальных микропроцессорах фрагменты вычислений.

Однако, в ряде работ [1,2,3] показано, что ПЛИС обладают значительно большим вычислительным потенциалом, который в полной мере может быть реализован в реконфигурируемых вычислительных системах (РВС), содержащих множество кристаллов ПЛИС, используемых как основной вычислительный элемент. Успешно развивающаяся более 20 лет в НИИ многопроцессорных вычислительных систем Южного федерального университета (г. Таганрог) концепция построения многопроцессорных вычислительных систем с программируемой архитектурой позволила создать целый ряд РВС различных архитектур и конфигураций, предназначенных для решения вычислительно трудоемких задач различных предметных областей, успешно эксплуатируемых организациями и ведомствами Российской Федерации. В качестве элементной базы для построения таких РВС используются ПЛИС Xilinx семейства Virtex большой интеграции, соединенные в единый вычислительный ресурс быстрыми каналами передачи данных – LVDS и Rocket GTX.

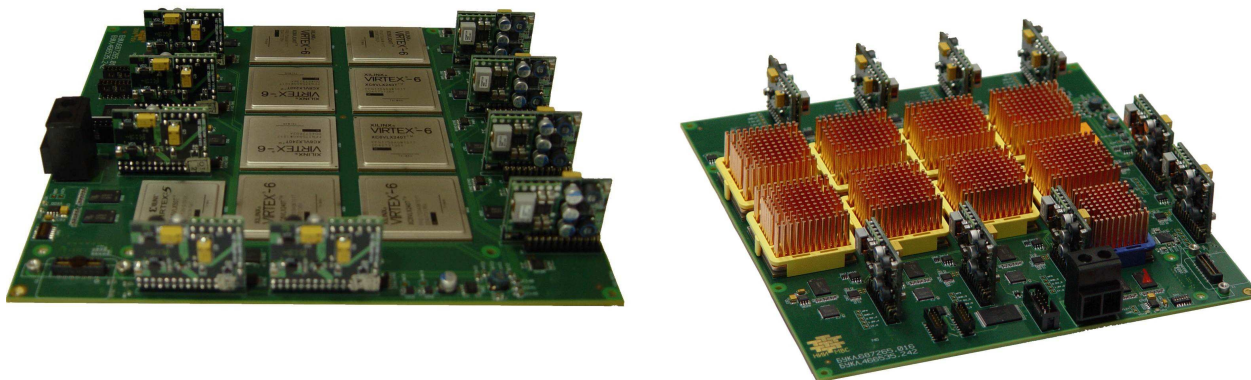
В НИИ многопроцессорных вычислительных систем Южного федерального университета серийно выпускались реконфигурируемые вычислительные системы на основе ПЛИС семейства Virtex 5, описанные в [3,4], разработанные по государственному контракту №02.524.12.4002 «Создание семейства высокопроизводительных многопроцессорных вычислительных систем с динамически перестраиваемой архитектурой на основе реконфигурируемой элементной базы и их математического обеспечения для решения вычислительно трудоемких задач», выполняемого по заданию Федерального агентства по науке и инновациям в рамках Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2012 годы».

Переход к принципам открытой масштабируемой архитектуры [1] в области разработки РВС положил начало новому семейству вычислительных систем под названием «Орион» и привел к созданию в 2010 году платы модифицированного вычислительного модуля с новой компоновкой и конструктивными решениями на основе ПЛИС семейства Virtex 5, принципы построения и технические характеристики которого описаны в [5].

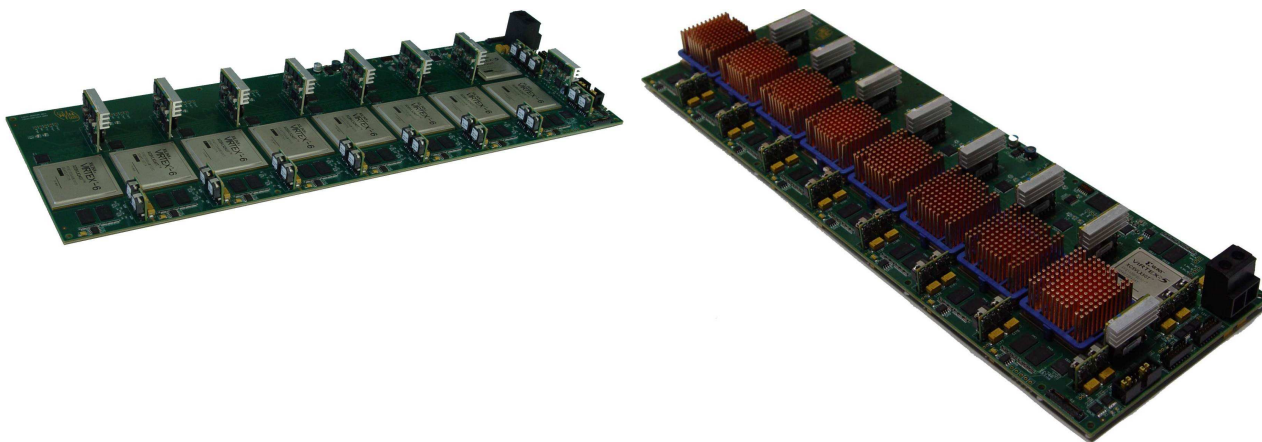
В настоящее время коллектив разработчиков НИИ многопроцессорных вычислительных систем Южного федерального университета приступил к выпуску РВС нового поколения на основе разработанных вычислительных модулей с использованием ПЛИС семейства Virtex-6. Разработаны и созданы платы нового поколения на основе ПЛИС семейства Virtex-6, построенные на основе открытой масштабируемой архитектуры [1] для вычислительных модулей перспективных конструктивных исполнений – «Саиф» и «Ригель», названных именами звезд из астрономического созвездия «Орион». Фотографии платы модифицированного вычислительного модуля на основе ПЛИС семейства Virtex 5 и плат нового поколения представлены на рисунке 1а-в.



а) плата вычислительного модуля на основе ПЛИС семейства Virtex 5;



б) плата вычислительного модуля «Саиф»;



в) плата вычислительного модуля «Ригель»

Рис. 1. Платы вычислительных модулей

Еще одним направлением применения высокопроизводительных РВС являются телекоммуникационные технологии. С целью освоения этой перспективной высокотехнологичной области в НИИ МВС ЮФУ на базе открытой масштабируемой архитектуры разработана плата вычислительного модуля «Орфей», предназначенная для построения высокопроизводительных систем мониторинга телекоммуникационных сетей и обработки сигналов. Серийное производство плат и вычислительных модулей на их основе намечено на третий квартал 2011 года. Плата вычислительного модуля «Орфей» обладает следующими характеристиками:

тип ПЛИС Virtex-6 вычислительного поля – XC6VSX475T-1FFG1759C;

количество ПЛИС вычислительного поля – 8;

количество элементарных процессоров – 16384;

пиковая производительность, операций в секунду – $16,2 \cdot 10^{12}$;

пиковая производительность, флопс – $1 \cdot 10^{12}$;

количество вх/вых LVDS (1,2 ГГц) – 312;

количество вх/вых RocketGTX (5,0 ГГц) – 512;

суммарная скорость обмена, бит в секунду – $2,4 \cdot 10^{12}$.

Вычислительные модули на основе плат «Орфей» имеют высоту 2U и устанавливаются в стандартную стойку шириной 19 дюймов. Основное назначение систем на основе плат «Орфей» – решение сильносвязанных задач, задач обработки сигналов и мониторинга телекоммуникационных сетей.

Одновременно ведется проработка вариантов изделий на ПЛИС семейства Virtex-7. Ожидаемые характеристики платы вычислительного модуля MM777 с вычислительным полем из ПЛИС семейства Virtex-7 XC7VX1140T-1FFG1928 и изделий на ее основе приведены в таблицах 1-4.

В таблице 1 приведены технические характеристики рассматриваемых плат вычислительных модулей. Вычислительные модули на основе этих плат «Орион-5», «Саиф», «Ригель», «Орфей» и MM777 имеют высоту 1U, 6U, 1U 2U и 1U соответственно и предназначены для установки в стандартную 19-дюймовую вычислительную стойку, которая является базовым компонентом для создания сверхвысокопроизводительных

комплексов на основе ПЛИС. Фотографии вычислительных модулей «Орион-5», «Саиф» и «Ригель» представлены на рисунке 2.

Таблица 1. Технические характеристики плат вычислительных модулей

Плата вычислительного модуля	Число ПЛИС	Тип и наименование ПЛИС	Количество эквивалентных вентилях в 1 ПЛИС, млн. шт.	Интерфейс и скорость межмодульного обмена, Гбит/сек	Потребляемая мощность, ВА
«Орион-5»	16	Virtex 5	11	LVDS, 1,2	250
«Саиф»	8	Virtex 6	24	Gigabit Ethernet, 1	300
«Ригель»	8	Virtex 6	24	Gigabit Ethernet, 1	300
«Орфей»	8	Virtex 6	47,5	LVDS, 1,2, RocketGTX, 5,0	350
MM777*	6	Virtex 7	113	LVDS, 1,2, RocketGTX, 6,0	350*

Примечание – Здесь и далее символом * отмечены ожидаемые характеристики.



а) вычислительный модуль «Орион-5»;

б) вычислительный модуль «Саиф»;

в) вычислительный модуль «Ригель»

Рис. 2. Вычислительные модули нового поколения

Применение ПЛИС семейства Virtex 6 в качестве элементной базы для построения вычислительных модулей «Саиф» и «Ригель» позволяет при сохранении стоимости поставки вычислительного модуля увеличить производительность в 1,5-2 раза по сравнению с аналогичным решением на основе ПЛИС семейства Virtex 5 для вычислительного модуля «Орион-5». Этот факт позволяет рассматривать созданные вычислительные модули нового поколения как наиболее перспективные варианты для построения РВС различных архитектур и конфигураций и обеспечивает им существенное конкурентное преимущество по большинству технико-экономических параметров: удельной производительности, энергоэффективности и др.

В таблице 2 представлены пиковые производительности рассматриваемых вычислительных модулей и вычислительных стоек на их основе. Производительность соответствует обработке данных с одинарной (Pi_{32}) и двойной (Pi_{64}) точностью в соответствии со стандартом IEEE-754 для вычислительных модулей и стоек описанных изделий. Технические характеристики вычислительных модулей представлены в таблице 2.

Таблица 2. Производительность вычислительных модулей и стоек

Наименование вычислительного модуля	Производительность вычислительного модуля Pi_{32}/Pi_{64} (Гфлопс)	Число вычислительных модулей в 19"стойке	Производительность стойки Pi_{32}/Pi_{64} (Тфлопс)
«Орион-5»	1000/340	24	24/8,1
«Саиф»	1600/500	6	9/3
«Ригель»	1600/500	32	51,2/16,0
«Орфей»	3000/1000	22	66,0/22,0

ММ777	4800*/1600*	32	105,6*/35,2*
-------	-------------	----	--------------

В таблице 3 приведены производительности вычислительных модулей на задачах символьной обработки данных, использующих битовые преобразования, и задачах математической физики на основе арифметики с плавающей запятой одинарной точности.

Таблица 3. Производительность вычислительных модулей

Вычислительный модуль	Символьная обработка данных (Топ/с)	Математическая физика, арифметика с плавающей запятой (Тфлопс)
«Орион-5»	116	1/0,34
«Саиф»	199,6	1,6/0,5
«Ригель»	199,6	1,6/0,5
«Орфей»	360,0	3,0/1,0
ММ777	216,0*	4,8*/1,6*

В таблице 4 приведены суммарные скорости передачи данных между кристаллами ПЛИС и блоками распределенной памяти, между ПЛИС в пределах одного вычислительного модуля и других вычислительных модулей.

Таблица 4. Скорость передачи данных

Вычислительный модуль	С блоками распределенной памяти (Гбит/с)	Между ПЛИС вычислительного поля (Тбит/с)	С другими вычислительными модулями (Тбит/с)
«Орион-5»	12,8	1,2	1,2
«Саиф»	12,8	1,0	1,0
«Ригель»	12,8	1,0	1,0
«Орфей»	12,8	2,1	2,4
ММ777	15,4*	2,0*	1,6*

Таким образом, вычислительные модули нового поколения «Саиф», «Ригель» и «Орфей» на основе ПЛИС семейства Virtex 6, а также модуль ММ777 на основе ПЛИС семейства Virtex 7 открывают перспективы для построения вычислительных систем более высокой производительности при сохранении стоимости системы по сравнению с РВС на основе вычислительного модуля «Орион-5». В то же время вычислительные модули обладают достаточной автономностью и могут легко комплексоваться с персональным компьютером типа IBM PC в качестве ускорителей и использоваться при решении различных задач.

Для вычислительных модулей нового поколения «Саиф», «Ригель», «Орфей» и ММ777 сохраняется преемственность принципов программирования РВС. Программирование всех рассмотренных вычислительных модулей и систем на их основе осуществляется с помощью единого комплекса системного программного обеспечения, поддерживающего структурно-процедурные методы организации вычислений. Программирование РВС отличается от программирования суперЭВМ традиционной архитектуры, поскольку включает организацию не только параллельных процессов и потоков данных, но и программирование структуры вычислительной системы в поле логических ячеек ПЛИС. Комплекс программного обеспечения вычислительных модулей предоставляет прикладному программисту следующие возможности:

- программирование как структурной, так и процедурной составляющих на языке высокого уровня без участия высококвалифицированного схемотехника;
- реконфигурацию прикладных программ при перераспределении вычислительного ресурса РВС;
- обеспечение совместимости и переносимости проектов между РВС разных архитектур;
- масштабирование прикладной задачи при увеличении ресурса;
- удаленное использование вычислительных ресурсов РВС.

Созданный комплекс программного обеспечения [3] по функциональному назначению разделяется на комплекс средств разработки прикладных программ и комплекс средств управления и администрирования ресурсов РВС.

Средства разработки прикладных программ содержат: транслятор языка ассемблера; транслятор языка

программирования PBC высокого уровня COLAMO; интегрированную среду разработки прикладных задач IDE, поддерживающую языки ассемблера и COLAMO; синтезатор масштабируемых параллельно-конвейерных решений, оперирующий библиотекой IP-ядер и интерфейсов.

Язык программирования высокого уровня COLAMO [3,4,5] обеспечивает поддержку создания как структурной, так и процедурной составляющих прикладной программы, реконфигурацию прикладных задач без участия высококвалифицированного схемотехника за счет неявного описания параллелизма и переносимость прикладных задач между PBC разных архитектур за счет использования файла описания архитектуры PBC и элементов библиотеки масштабируемых IP-ядер. Транслятор COLAMO v.2.0 осуществляет трансляцию процедурной составляющей программы, организующей потоки данных, в язык ассемблера Argus v.3.0 и создание структурной составляющей в объектном представлении, которая автоматически передается в среду разработки масштабируемых параллельно-конвейерных процедур Fire!Constructor для синтеза конфигурационных файлов ПЛИС на языке VHDL.

Фундаментальным типом вычислительной структуры в языке COLAMO является конструкция "кадр". Кадром является программно-неделимый компонент, представляющий собой совокупность арифметико-логических команд, выполняемых на различных элементарных процессорах, обладающих распределенной памятью и соединенных между собой в соответствии с информационной структурой алгоритма таким образом, что вычисления производятся с максимально возможным параллелизмом и асинхронностью.

Кадр фактически определяет вычислительную структуру и потоки данных в PBC в данный момент времени. При этом все операции в теле кадра выполняются асинхронно с максимальным параллелизмом, а последовательность смены кадров однозначно определяется программистом.

В языке отсутствуют явные формы описания параллелизма. Распараллеливание достигается с помощью объявления типов доступа к переменным и индексации элементов массивов. Для исключения конфликтов одновременного чтения и записи ячеек памяти в пределах текущего кадра используется широко распространенное в языках потока данных правило единственной подстановки: переменная, хранящаяся в памяти, может получить значение в кадре только один раз.

Для обращения к данным используются два основных метода доступа: параллельный доступ (задаваемый типом Vector) и последовательный доступ (задаваемый типом Stream). На рисунке 3 представлены программы, являющиеся граничными примерами извлечения параллелизма, и графы синтезируемых вычислительных структур.

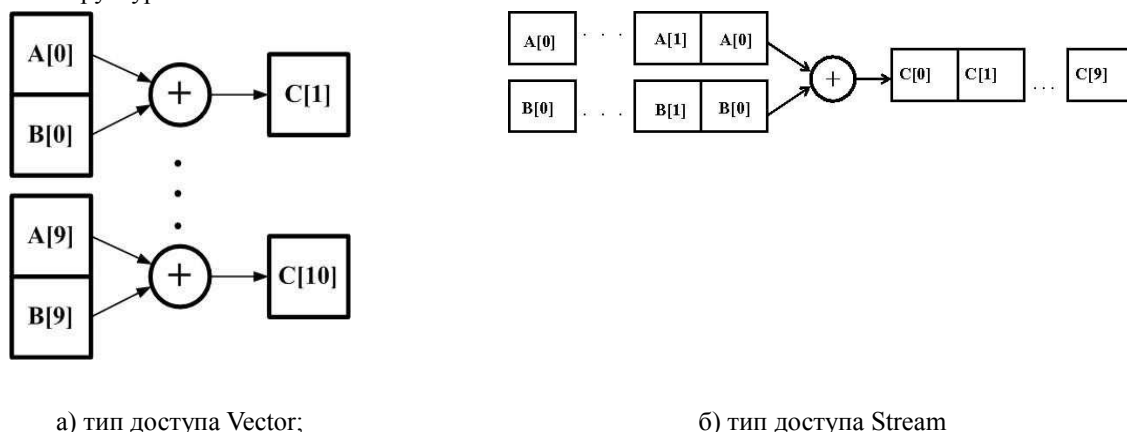


Рис. 3. Параллельное и последовательное сложение массивов

Тип доступа Stream указывает на последовательную обработку элементов одномерного массива, а тип Vector позволяет обрабатывать элементы одномерного массива одновременно.

Многомерные массивы состоят из множества измерений, каждое из которых может иметь последовательный или параллельный тип доступа, задаваемый ключевым словом Stream или Vector соответственно. Применение неявного описания параллелизма за счет задания типа доступа позволяет достаточно просто управлять степенью распараллеливания программы на уровне описания структур данных и дает возможность программисту максимально просто описывать различные виды параллелизма в достаточно сжатом виде.

Трансляция программы на языке высокого уровня COLAMO состоит в создании схемотехнической конфигурации вычислительной системы (структурной составляющей) и параллельной программы, управляющей потоками данных (поточковой и процедурной составляющих).

Операторы и функции языка (сумматоры, умножители, функции сравнения, тригонометрические функции и др.), используемые в тексте параллельной программы, имеют готовые схемотехнические решения. Данные решения разрабатываются специалистами-схемотехниками в интегрированной среде разработки цифровых устройств ISE фирмы XILINX или с ней совместимых и включаются в библиотеку транслятора языка COLAMO и библиотеку стандартных примитивов среды Fire!Constructor.

В процессе работы транслятора языка COLAMO формируется информационный граф прикладной задачи из текста параллельной программы, где операторы и функции языка по определённым правилам заменяются соответствующими блоками или группами блоков из библиотеки стандартных примитивов.

Синтезированный вычислительный граф задачи передается в среду разработки вычислительных структур Fire!Constructor для укладки на множество ПЛИС PBC и обеспечения синхронизации между ПЛИС [5]. Одной из задач среды является формирование разбиения информационного графа прикладной задачи на непересекающиеся подграфы, каждый из которых будет структурно реализован в кристаллах ПЛИС выбранной PBC.

Процесс синтеза результата разбиения информационных графов прикладных задач состоит из следующих этапов:

- решения задачи разбиения (компоновки) узлов информационного графа прикладной задачи на непересекающиеся подграфы, каждый из которых будет размещён в соответствующем БМ;
- решения задачи размещения и трассировки для узлов информационного графа в каждом БМ в отдельности и задачи трассировки связей между БМ;
- синтеза файлов VHDL-описаний и файлов временных и топологических ограничений для каждой ПЛИС, каждого БМ выбранной PBC.

Среда Fire!Constructor упрощает создание масштабируемых структурных решений и сокращает время разработки за счет автоматизированного выполнения следующих трудоемких процедур:

- согласования входов и выходов совместно работающих ПЛИС (ucf-файлов);
- автоматической синхронизации информационных потоков при размещении функциональных устройств в едином вычислительном контуре, расположенном в различных кристаллах ПЛИС;
- автоматического обеспечения сбалансированного размещения функциональных устройств по различным ПЛИС.

Технология создания прикладных программ для PBC и общая взаимосвязь транслятора языка COLAMO, среды Fire!Constructor и синтезатора конфигурации ПЛИС в рамках комплекса системного программного обеспечения при создании многокристального схемотехнического решения для PBC представлена на рисунке 4.

Такой подход к программированию реконфигурируемых вычислительных систем позволяет освободить программиста от построения графа задачи в виде функциональных библиотек в среде Fire!Constructor и синхронизации потоков данных в PBC, сократив время создания параллельных программ для PBC в 3-10 раз, и исключить участие специалиста-схемотехника при разработке параллельных прикладных программ.

Язык структурно-процедурного программирования Argus представляет собой низкоуровневый язык (ассемблер), предназначенный для описания процедурной составляющей прикладной параллельной программы PBC [3,4]. Программа на языке Argus организует потоки данных на уровне команд контроллеров распределенной памяти, обеспечивая их синхронизацию.

Интегрированная среда разработки Argus IDE предназначена для интерактивной разработки параллельных программ на языках высокого уровня COLAMO и языке ассемблера Argus в едином языковом пространстве. Среда Argus IDE, объединяя в своем составе трансляторы языков COLAMO и Argus, обеспечивает эффективную разработку масштабируемых параллельных программ для PBC.



Рис. 4. Технология создания прикладных программ для РВС

Созданное параллельное решение прикладной задачи в виде загрузочного модуля РВС с помощью драйвера загружается в вычислительный модуль РВС. Драйвер вычислительных модулей обеспечивает программную поддержку функций непосредственного доступа к высокоскоростному аппаратному интерфейсу, поддерживающему пакетные режимы работы и обеспечивающему механизмы прямого доступа к физической памяти управляющего компьютера.

Для удаленного доступа и управления вычислительными ресурсами РВС разработана система удаленного доступа, которая состоит из сервера, обрабатывающего удаленные заявки на использование вычислительных ресурсов и поддерживающего очередь заявок, и клиента, формирующего заявки на основе команд пользователя. К функциям системы удаленного доступа относятся функции включения, выключения, остановки и запуска как отдельных вычислительных модулей, так и стоек и РВС в целом.

Созданный комплекс программного обеспечения позволяет создавать эффективные прикладные программы для РВС при решении задач различных предметных областей, обеспечивает удобство программирования и сокращает время разработки прикладного решения в 3-5 раз, обеспечивая при этом автоматизированный перенос структурного решения с одной архитектуры РВС на другую.

РВС являются перспективным направлением развития высокопроизводительной вычислительной техники, которые, в отличие от кластерных суперЭВМ, предоставляют пользователю возможность создавать в базовой архитектуре виртуальные специализированные вычислители, структура которых адекватна структуре решаемой задачи. Это, в свою очередь, обеспечивает высокую эффективность вычислений и близкий к линейному рост производительности при наращивании вычислительного ресурса.

Следует отметить, что переход на новую компоновку модулей, позволившую сосредоточить в пределах вычислительных модулей «Орион-5» и «Ригель» высотой 1U мощный вычислительный ресурс на основе ПЛИС, обеспечивает удельную производительность РВС на уровне лучших мировых показателей для суперЭВМ с кластерной архитектурой. Учитывая тот факт, что РВС обладают до 10 раз превосходящей удельной производительностью на широком классе задач, чем кластерные суперЭВМ, можно сделать вывод о том, что РВС являются средством создания суперЭВМ нового поколения.

ЛИТЕРАТУРА:

1. И.И. Левин Реконфигурируемые вычислительные системы с открытой масштабируемой архитектурой // Труды Пятой Международной конференции «Параллельные вычисления и задачи управления» РАСО'2010. - М.: Учреждение Российской академии наук Институт проблем управления им. В.А. Трапезникова РАН, 2010. - С.83-95.
2. А.В. Каляев, И.И. Левин Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. - М.: Янус-К, 2003. – 380 с.
3. И.А. Каляев, И.И. Левин, Е.А. Семерников, В.И. Шмойлов Реконфигурируемые мультимодульные вычислительные структуры /Изд. 2-е, перераб. и доп. / Под общ. ред. И.А. Каляева. - Ростов-на-Дону: Изд-во ЮНЦ РАН, 2009. – 344 с.
4. И.А. Каляев, И.И. Левин Семейство реконфигурируемых вычислительных системы с высокой

- реальной производительностью // Труды международной научной конференции «Параллельные вычислительные технологии» (ПАВТ'2009). – Нижний Новгород: электронное издание НГУ имени Н.И. Лобачевского, 2009. – С.186-196.
5. Н.Н. Дмитренко, И.А. Каляев, И.И. Левин, Е.А. Семерников Развитие аппаратной платформы реконфигурируемых вычислительных систем // Труды Международной суперкомпьютерной конференции «Научный сервис в сети Интернет: суперкомпьютерные центры и задачи. – М.: Изд-во МГУ, 2010. – С. 315-320.