

Ю.И. Доронченко¹, И.А. Каляев¹, И.И. Левин¹, Е.А. Семерников²

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ С РЕКОНФИГУРИРУЕМОЙ АРХИТЕКТУРОЙ НА ОСНОВЕ СОВРЕМЕННЫХ И ПЕРСПЕКТИВНЫХ ПЛИС

ООО «Научно-исследовательский центр СуперЭВМ и
нейрокомпьютеров», г. Таганрог, Россия¹,

Южный научный центр Российской академии наук, г. Ростов-на-Дону,
Россия²

¹doronchenko@superevm.ru, ¹kaliaev@superevm.ru, ¹levin@superevm.ru,
²semernikov@superevm.ru

1. Введение

В настоящее время отечественные [1] и зарубежные производители вычислительной техники все чаще включают в их состав программируемые логические интегральные схемы (ПЛИС) для ускорения работы при реализации вычислительно трудоемких алгоритмов. Создаются как отдельные ускорители с одним-двумя кристаллами ПЛИС, так и целые вычислительные комплексы. Такие фирмы как Nallatech [2] и Pico Computing [3] выпускают ряд ускорителей и несущих плат с небольшим числом (до четырёх) кристаллов ПЛИС, которые используются в создании серверов и гибридных кластерных систем фирмами HP и IBM. Компании Convey [4] и Maxeler Technologies [5] создают гибридные суперкомпьютеры на основе собственных гетерогенных кластерных узлов, каждый из которых может содержать от одного до четырех кристаллов ПЛИС и несколько универсальных процессоров. Похожее решение используется и компанией SRC [6], которая выпускает узлы, названные MAP processor, для стойки (MAPstation) форм-фактором 1U, 2U и 4U. MAPstation 1U содержит один MAP processor. MAPstation 2U содержит до трёх MAP processor. MAPstation 4U может содержать до 10 различных модулей – MAP processor, модуль с универсальным микропроцессором или модуль памяти.

В ООО «Научно-исследовательский центр СуперЭВМ и нейрокомпьютеров» (ООО «НИЦ СЭ и НК», г. Таганрог) разрабатываются и производятся реконфигурируемые вычислительные системы (PVC) класса суперЭВМ, где основным вычислительным ресурсом являются не микропроцессоры, а множество кристаллов ПЛИС, объединенных в вычислительные поля посредством высокоскоростных каналов передачи данных. PVC такой архитектуры

находят применение для решения вычислительно трудоемких задач в различных областях науки и техники, поскольку обладают рядом существенных преимуществ по сравнению с многопроцессорными вычислительными системами кластерной архитектуры: высокими реальной и удельной производительностями, высокой энергоэффективностью и др.

2. PBC на основе ПЛИС Xilinx седьмого семейства

2.1 PBC на основе вычислительных модулей «Тайгета»

На основе ПЛИС Virtex-7 разработан вычислительный 19" модуль (BM) «Тайгета» высотой 2U, предназначенный для создания высокопроизводительных многостоечных PBC.

BM «Тайгета», представленный на рис. 1а, содержит четыре платы 8V7-200, встроенную управляющую ЭВМ, систему питания, систему управления, систему охлаждения и другие подсистемы. Платы в составе BM «Тайгета» соединяются LVDS-каналами, работающими на частотах до 1000 МГц. Внешний вид платы 8V7-200 показан на рис. 1б.

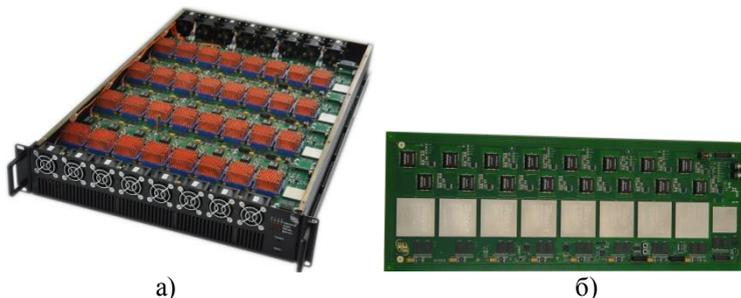


Рис. 1. Фотография BM «Тайгета» (а – фотография BM «Тайгета» со снятой верхней крышкой, б – фотография ПВМ 8V7-200)

Плата вычислительного модуля (ПВМ) 8V7-200 представляет собой 20-слойную печатную плату с двухсторонним монтажом элементов, на которой располагаются 8 ПЛИС типа XC7VX485T-1FFG1761, содержащих по 48,5 миллионов эквивалентных вентилей, 16 микросхем распределенной памяти SDRAM типа DDR2 общим объемом 2 Гбайт, интерфейсы LVDS и Ethernet и другие компоненты.

Производительность одной ПВМ 8V7-200 составляет 667 Гфлопс при обработке 32-разрядных данных с плавающей запятой, а

производительность ВМ «Тайгета» составляет соответственно 2,66 Тфлопс. Производительность одной стойки РВС при комплектации 18 ВМ «Тайгета» составляет 48 Тфлопс при обработке данных с плавающей запятой одинарной точности и 23 Тфлопс при обработке 64-разрядных данных с плавающей запятой. Внешний вид стандартной 19” стойки РВС на основе ВМ «Тайгета» показан на рис. 2.

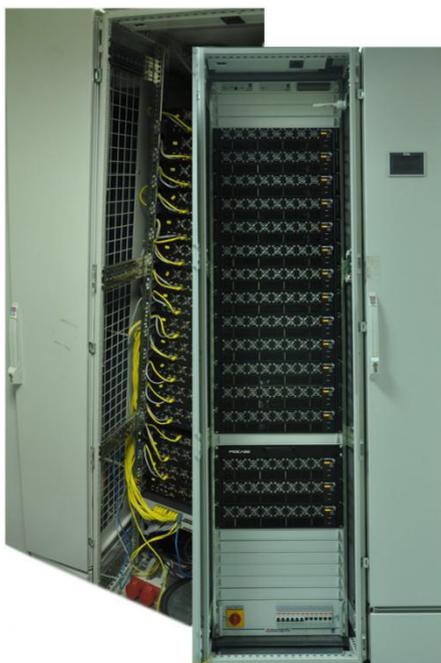


Рис. 2. Внешний вид стойки РВС на основе ВМ «Тайгета»

Высокопроизводительные РВС на основе ВМ «Тайгета» ориентированы на решение вычислительно трудоемких задач науки и промышленности, задач синтеза лекарств и задач символьной обработки и обеспечивают при этом существенное конкурентное преимущество по большинству технико-экономических параметров: удельной производительности, энергоэффективности и другим по сравнению с суперЭВМ традиционной кластерной архитектуры.

2.2 Реконфигурируемый вычислительный блок «Калеано»

На основе ПЛИС Xilinx седьмой серии разработан новый реконфигурируемый вычислительный блок (РВБ) «Калеано» в конструктивном исполнении высотой 1U, предназначенный для обработки данных, поступающих по каналу Gigabit Ethernet без поддержки IP-протоколов. Возможные области применения: символьная обработка, математическая физика, моделирование и вычислительный эксперимент, цифровая обработка сигналов, линейная алгебра и т.п.

На рис. 3а представлена фотография РВБ «Калеано», на рис. 3б - блок со снятой верхней крышкой и печатная плата.

РВБ «Калеано» содержит вычислительное поле из шести ПЛИС, встроенную управляющую ЭВМ, систему питания, систему управления, систему охлаждения и другие подсистемы. Все ПЛИС вычислительного поля соединены между собой по принципу близкодействия с помощью LVDS-каналов, ко всем ПЛИС РВБ подключены модули динамической памяти емкостью 256 Мбайт каждый.



Рис. 3. Фотографии РВБ «Калеано» и его составных частей

Для управления и конфигурирования вычислительного поля РВБ используется ЭВМ семейства COM-Express фирмы Kontron, установленная непосредственно на плате вычислительного модуля. С ее помощью подключается периферийное оборудование, осуществляются подготовка и отладка параллельной программы для решения вычислительно трудоемких задач, формируются файлы исходных данных, которые вместе с исполняемым модулем задачи через шину PCI-Express загружаются в вычислительное поле по LVDS-каналу. После выполнения задачи результаты решения пересылаются в процессорный модуль COM-Express.

РВБ «Калеано» выпускается в двух модификациях «Калеано-К» на ПЛИС Kintex-7 XC7K160T и «Калеано-V» на ПЛИС Virtex-7

XC7VX485T. Технические характеристики РВБ «Калеано» для этих модификаций приведены в таблице 1.

Таблица 1 - Технические характеристики РВБ «Калеано» для модификаций «Калеано-К» и «Калеано-V»

Технический параметр	Калеано-К	Калеано-V
Число и тип ПЛИС	6	6
Общее число эквивалентных вентилях в ПЛИС вычислительного поля, млн.	96	288
Объем оперативной памяти, Гб	3	3
Производительность вычислительного модуля P_{i32}/P_{i64} , Гфлопс	150/75	440/220
Частота работы, МГц	330	400
Скорость обмена данными по каналу Ethernet, Гбит/с	1	1
Частота передачи данных по LVDS между ПЛИС вычислительного поля, МГц	900	1200
Потребляемая мощность, Вт	200	320
Габаритные размеры, мм	480x270x70	480x270x70
Стоимость, млн. руб.	1,3	2,0

2.3 РВС на основе ПЛИС UltraScale

Развитием направления реконфигурируемых компьютеров «Калеано» является использование в данных изделиях новой элементной базы – ПЛИС Xilinx нового поколения семейства UltraScale. ПЛИС данного семейства, выполненные по технологии 20 нм, обладают сниженным энергопотреблением и повышенным быстродействием по сравнению с ПЛИС седьмого семейства. Эскиз компоновки ПВМ показан на рис. 4.

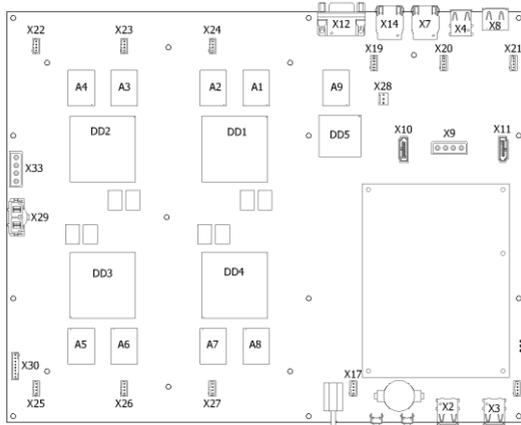


Рис. 4. Эскиз компоновки ПБМ «Калеано-U»

На рис. 4 обозначены:

- DD1-DD4 – вычислительные ПЛИС Xilinx UltraScale XCVU095;
- DD5 – ПЛИС контроллера ПБМ Xilinx UltraScale XCKU040;
- A1-A9 – модули распределенной памяти;
- X2-X4, X7-X12 – разъемы различных типов интерфейсов.

По сравнению с предыдущей версией ПРК «Калеано-V» производительность ПРК «Калеано-U» возрастет в 1,7–1,8 раза, при увеличении потребляемой мощности - не более чем в 1,3 раза.

3. Реконфигурируемые системы на основе жидкостного охлаждения

Использование воздушных систем охлаждения современных вычислителей и построенных на их основе суперкомпьютеров, в том числе реконфигурируемых, практически достигло своего предела. В современных суперЭВМ на основе новых поколений ПЛИС целесообразно использовать жидкостное охлаждение, в частности, непосредственное погружение плат вычислительных модулей в жидкостный хладагент на основе минерального масла.

В ООО «НИЦ СЭ и НК» (г. Таганрог) активно развивается направление по созданию РВС нового поколения на основе жидкостного охлаждения. Разработаны новые конструкции печатных плат и вычислительных модулей, характеризующиеся высокой плотностью компоновки. В частности, в настоящее время ведется разработка

перспективных вычислительных модулей «Скат-8» для многостоечных PBC сверхвысокой производительности.

Плата перспективного вычислительного модуля содержит восемь ПЛИС семейства Virtex UltraScale логической емкостью 100 млн. эквивалентных вентилей каждая. Вычислительный модуль состоит из двух секций. В первой секции размещается 16 плат вычислительных модулей с потребляемой мощностью до 800 Вт каждая. Платы полностью погружены в электрически нейтральный жидкостный хладагент. Во второй секции располагаются насосная группа и теплообменник, обеспечивающие проток и охлаждение хладагента. Высота конструктива ВМ составляет 3U. Эскиз ВМ показан на рис. 5.

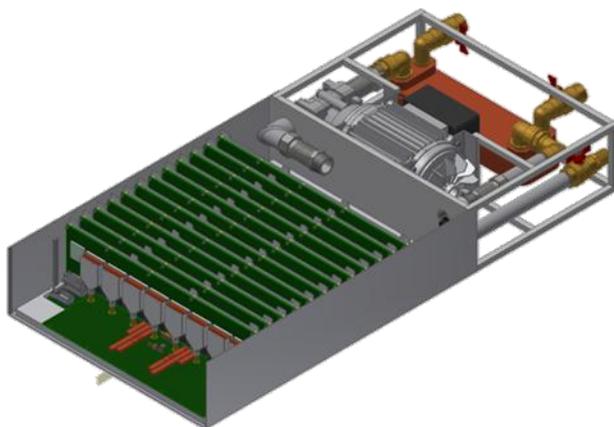


Рис. 5. Эскиз ВМ «Скат-8»

На основе ВМ «Скат-8» обеспечивается построение реконфигурируемых суперкомпьютеров петафлопсной производительности. Вычислительная 19” стойка суперкомпьютера может содержать до 12 ВМ «Скат-8» с жидкостным охлаждением. Эскиз вычислительной стойки показан на рис. 6.

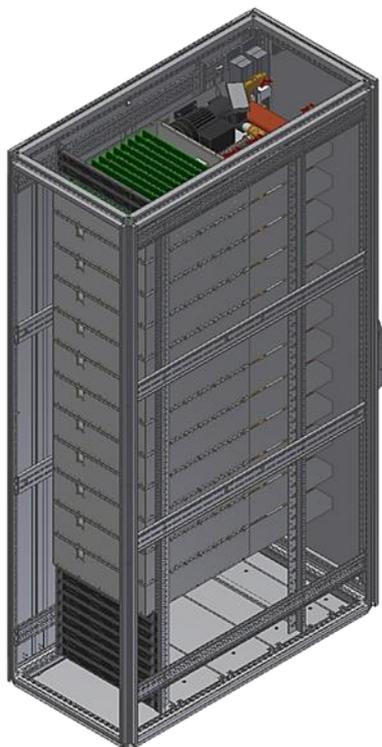


Рис. 6. Эскиз вычислительной стойки на основе VM «Скат-8»

Производительность и потребляемая мощность перспективной PBC приведены в таблице 2.

Таблица 2 – Производительность и потребляемая мощность перспективной PBC

Характеристика	Значение
Производительность VM «Скат-8»	105 Тфлопс
Производительность вычислительной стойки на основе VM «Скат-8»	1,25 Пфлопс
Потребляемая мощность VM «Скат-8»	13 кВт
Потребляемая мощность вычислительной стойки на основе VM «Скат-8»	150 кВт

На основе данных конструкций будут созданы сверхпроизводительные вычислительные комплексы, в которых обеспечивается эффективное охлаждение вычислительных ПЛИС как семейства UltraScale, так и следующего прогнозируемого семейства ПЛИС.

4. Программное обеспечение РВС на основе ПЛИС

Для вычислительных систем на основе ПЛИС седьмого семейства и UltraScale сохраняется преемственность принципов программирования [7-10]: программирование всех рассмотренных РВС и их вычислительных модулей осуществляется с помощью единого комплекса системного программного обеспечения, поддерживающего структурно-процедурные методы организации вычислений и определяющего не только организацию параллельных процессов и потоков данных, но и структуру вычислительной системы в поле логических ячеек всех ПЛИС. Наиболее характерной отличительной особенностью комплекса программного обеспечения РВС на основе ПЛИС Xilinx является поддержка проблемно-ориентированных софт-архитектур, позволяющих создавать и программировать макрообъекты, представляющие собой совокупность вычислительных устройств, выполняющих определенную группу команд и соединенных между собой коммутационной системой. Это обеспечивает при тех же принципах программирования и использовании языка высокого уровня для программирования РВС возможность простой адаптации программных компонентов средств разработки для РВС при переходе на новые аппаратные платформы без внесения существенных изменений в код программных компонентов комплекса, а также позволяет сократить время решения прикладных задач.

Программирование РВС осуществляется в два этапа: на первом этапе схемотехник создает вычислительную структуру для решения прикладной задачи, а на втором этапе прикладной программист создает параллельную программу, определяющую потоки данных в созданной схемотехником вычислительной структуре.

Большинство существующих зарубежных коммерческих систем проектирования (Xilinx ISE, Altium Designer и другие) обеспечивает в рамках одного проекта работу только с одним кристаллом ПЛИС. Поэтому при разработке конфигурации для нескольких ПЛИС инженеру-схемотехнику приходится самому распределять элементы вычислительной структуры между микросхемами и учитывать топологию связей между ними. Учет особенностей внутренней

архитектуры, топологии и элементной базы существенно усложняет специалисту-схемотехнику разработку конфигураций вычислительной структуры прикладной задачи для многокристалльных решений и практически исключает возможность переноса (портации) готового решения на платформы другой конфигурации или архитектуры. Это приводит к тому, что сроки разработки прикладной программы для решения прикладной задачи на реконфигурируемой системе достаточно велики и составляют 4-9 месяцев.

Для программирования РВС, созданных в ООО «НИЦ СЭ и НК», используется разработанный комплекс программного обеспечения [1, 7, 9, 11], поддерживающий структурно-процедурные методы организации вычислений и определяющий как структуру вычислительной системы в поле логических ячеек ПЛИС, так и организацию параллельных процессов и потоков данных. Отличительными особенностями комплекса программного обеспечения на основе языка программирования COLAMO по сравнению с известными средствами разработки MitrionC [12] и CatapultC [13], являются автоматическое размещение, синхронизация и создание конфигурации для многокристалльных РВС, высокий процент заполнения кристалла (60-90%) и высокие частоты работы (250 – 350 МГц).

Средства программирования прикладных задач на языке высокого уровня COLAMO содержат следующие основные компоненты:

- транслятор языка программирования COLAMO, осуществляющий трансляцию исходного кода на COLAMO в информационный граф параллельной прикладной программы;

- синтезатор масштабируемых схемотехнических решений на уровне логических ячеек ПЛИС Fire!Constructor, осуществляющий отображение полученного от транслятора языка программирования COLAMO информационного графа на архитектуру РВС, размещение отображенного решения по кристаллам ПЛИС и автоматическую синхронизацию фрагментов информационного графа в разных кристаллах ПЛИС;

- библиотеку IP-ядер, соответствующих операторам языка COLAMO (функционально-законченных структурно-реализованных аппаратных устройств), для различных предметных областей и интерфейсов для согласования скорости обработки информации и связи в единую вычислительную структуру.

5. Заключение

На основе производимых в ООО «НИЦ СЭ и НК» реконфигурируемых систем можно проследить темпы роста производительности PBC при переходе от семейства к семейству ПЛИС. Как видно из таблицы 3, ПЛИС как элементная база реконфигурируемых суперЭВМ обеспечивают их устойчивый, близкий к линейному рост производительности, открывая новые перспективы по созданию суперкомпьютеров петафлопсной производительности.

Таблица 3 – Производительность реконфигурируемых суперЭВМ

Изделие	ПВМ $P_{i_{32}}/P_{i_{64}}$, Гфлопс	ВМ $P_{i_{32}}/P_{i_{64}}$, Гфлопс	Стойка $P_{i_{64}}$, Тфлопс
«Орион-5» 2009 год Virtex-5	250/85	1000/340	19,2 - 28,8
«Ригель» 2010/2012 год Virtex-6	400/125	1600/500	34,5 – 51,8
«Тайгета» 2012/2013 год Virtex-7	900/300	3600/1200	68 – 100
«Скат» 2015/2016 год UltraScale	7250/2500	82500/30000	1000 – 1250

Таким образом, конструктивные решения, положенные в основу перспективных вычислительных модулей на основе ПЛИС Xilinx Virtex-7 и UltraScale, позволяют сосредоточить в пределах одной вычислительной стойки высотой 47U мощный вычислительный ресурс и обеспечивают удельную производительность PBC на основе ПЛИС на уровне лучших мировых показателей для суперЭВМ с кластерной архитектурой. Это позволяет рассматривать PBC на основе ПЛИС Xilinx как основу для создания высокопроизводительных вычислительных комплексов нового поколения, обеспечивающих высокую эффективность вычислений и близкий к линейному рост производительности при наращивании вычислительного ресурса.

Литература

1. Каляев А.В., Левин И.И. Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. М.: Янус-К, 2003. 380 с.
2. <http://www.nallatech.com/>
3. <http://picocomputing.com/>
4. <http://www.conveycomputer.com>
5. <http://www.maxeler.com/>
6. <http://www.srccomp.com/>
7. Каляев И.А., Левин И.И., Семерников Е.А., Шмойлов В.И. Реконфигурируемые мультиконвейерные вычислительные структуры /Изд. 2-е, перераб. и доп. / Под общ. ред. И.А. Каляева. Ростов-на-Дону: Изд-во ЮНЦ РАН, 2009. 344 с.
8. Каляев И.А., Левин И.И. 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.
9. Дордопуло А.И., Каляев И.А., Левин И.И., Семерников Е.А. Высокопроизводительные реконфигурируемые вычислительные системы нового поколения // Труды Международной суперкомпьютерной конференции с элементами научной школы для молодежи «Научный сервис в сети Интернет: экзафлопсное будущее». М.: Изд-во МГУ, 2011. С. 42-49.
10. Каляев И.А., Левин И.И., Семерников Е.А., Дордопуло А.И. Реконфигурируемые вычислительные системы на основе ПЛИС семейства Virtex-6 // Сборник трудов Международной научной конференции «Параллельные вычислительные технологии 2011» (ПАВТ 2011). Челябинск-М.: Издательский центр ЮУрГУ [Электронный ресурс], 2011. С. 203–210.
11. I.A. Kalyaev, I.I. Levin, E.A. Semernikov, V.I. Shmoilov. (2012). Reconfigurable multipipeline computing structures. Nova Science Publishers, New York.