

И.А. КАЛЯЕВ, А.И. ДОРДОПУЛО, И.И. ЛЕВИН, А.М. ФЕДОРОВ
**РАЗВИТИЕ ОТЕЧЕСТВЕННЫХ МНОГОКРИСТАЛЬНЫХ
РЕКОНФИГУРИРУЕМЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ: ОТ
ВОЗДУШНОГО К ЖИДКОСТНОМУ ОХЛАЖДЕНИЮ**

Каляев И.А., Дордопуло А.И., Левин И.И., Федоров А.М. Развитие отечественных многокристалльных реконфигурируемых вычислительных систем: от воздушного к жидкостному охлаждению.

Аннотация. В статье рассматривается история развития технологий построения реконфигурируемых вычислительных систем на основе кристаллов ПЛИС различных семейств. Описано пять поколений реконфигурируемых вычислительных систем с высокой плотностью компоновки, созданных на основе ПЛИС, начиная с Xilinx Virtex-E и заканчивая современными ПЛИС Virtex UltraScale. Представлены последние достижения в области проектирования энергоэффективных реконфигурируемых вычислительных систем с высокой реальной производительностью — разработанная система жидкостного охлаждения для кристаллов Virtex UltraScale, которая обеспечивает автономную циркуляцию охлаждающей жидкости в вычислительном модуле с размерами 3U на 19" для охлаждения 96-128 кристаллов ПЛИС с выделяемой суммарной тепловой мощностью 9,6-12,8 кВт. Отличительными характеристиками разработанной погружной системы жидкостного охлаждения являются высокая эффективность охлаждения с резервом мощности для проектируемых перспективных семейств ПЛИС, нечувствительностью к протечкам и их последствиям и совместимость с традиционными системами водяного охлаждения на базе промышленных чиллеров.

Ключевые слова: реконфигурируемые вычислительные системы, ПЛИС, энергоэффективность, жидкостное охлаждение, промышленные чиллеры

1. Введение. Одним из перспективных способов достижения высокой реальной производительности вычислительной системы является адаптация ее архитектуры под структуру решаемой задачи, а также создание специализированного вычислительного устройства, аппаратно реализующего все вычислительные операции информационного графа задачи с наименьшими задержками. Естественным требованием к современной вычислительной системе является аппаратная поддержка возможности изменения как алгоритма решаемой задачи, так и самой задачи, что обуславливает применение программируемых логических интегральных схем (ПЛИС) в качестве основного вычислительного ресурса реконфигурируемых вычислительных систем [1].

Основными преимуществами ПЛИС являются: возможность реализации сложных параллельных алгоритмов, возможность адаптации конфигурации вычислительной системы под структуру решаемой задачи, высокая производительность получаемого решения, совместимость различных проектов на языках описания аппаратуры — VHDL, AHDL, Verilog и др. Следует учитывать, что вычислительные системы на основе ПЛИС не лишены и недостатков: сложность проектирования, изго-

товления и программирования, необходимость эффективного охлаждения, приобретающая особое значение для многокристалльных вычислительных систем, поскольку рабочие температуры ПЛИС, как правило, на 30-40% ниже, чем у микропроцессоров. Поэтому в настоящее время разработчикам реконфигурируемых вычислительных систем необходимо обеспечить эффективное охлаждение вычислительных модулей, содержащих не менее восьми кристаллов ПЛИС с высокой плотностью компоновки, поскольку воздушные системы охлаждения, успешно использовавшиеся для ПЛИС с момента их появления, практически достигли своего предела для охлаждения суперкомпьютеров.

2. История развития ПЛИС как элементной базы для РВС. История развития архитектур ПЛИС берет свое начало с конца 1970-х годов, когда на рынке появились ПЛИС с программируемыми матрицами «И» и «ИЛИ» — архитектуры FPLA (Field Programmable Logic Array) и FPLS (Field Programmable Logic Sequencers) [2]. Их недостатком было слабое использование ресурсов программируемой матрицы «ИЛИ».

Идя по пути совершенствования архитектуры, в конце 1980-х годов разработчики ПЛИС предложили более простую и изящную архитектуру программируемой матричной логики PAL (Programmable Array Logic) и GAL (Gate Array Logic) — ПЛИС фирм Intel, Altera, AMD, Lattice и других с небольшой степенью интеграции, имеющую программируемую матрицу «И» и фиксированную матрицу «ИЛИ» [3]. Другим подходом к уменьшению избыточности программируемой матрицы «ИЛИ» являлась программируемая макрологика. Построенные на основе данной архитектуры микросхемы содержали единственную программируемую матрицу «И-НЕ» или «ИЛИ-НЕ», но за счет многочисленных инверсных обратных связей были способны формировать сложные логические функции.

В начале 1980-х годов на мировой рынок микроэлектронных изделий выходят три ведущие фирмы-производители ПЛИС. В июне 1983 года основана фирма Altera Corporation [4], в феврале 1984 — компания Xilinx Inc. [5], в 1985 году — Actel Corporation [6]. Эти три компании в настоящее время занимают до 80% всего рынка ПЛИС и являются основными идеологами их применения. Если ранее ПЛИС являлись одним из множества продуктов, выпускаемых такими гигантами как Intel, AMD и другие, то начиная с середины 1980-х годов на рынке ПЛИС происходит специализация, и законодателями мод становятся фирмы, специализирующиеся только на разработке и производстве ПЛИС.

С появлением новых производителей возникли и новые архитектуры. ПЛИС CPLD (Complex Programmable Logic Devices) [7] имеют довольно высокую степень интеграции и содержат несколько логиче-

ских блоков, объединенных коммутационной матрицей. Каждый логический блок представляет собой программируемую матрицу «И» и фиксированную матрицу «ИЛИ». К этому классу относятся ПЛИС семейств MAX7000 фирмы Altera, XC9500 фирмы Xilinx, а также большое число микросхем других производителей (Atmel, Vantis, Lucent и др.).

Настоящая революция в устройствах программируемой логики произошла с появлением ПЛИС, имеющих архитектуру Field Programmable Gate Array (FPGA) [8]. К FPGA относятся ПЛИС XC2000, XC3000, XC4000, Spartan фирмы Xilinx, АСТ1, АСТ2 фирмы Actel, а также семейства FLEX8000 фирмы Altera, некоторые ПЛИС Atmel и Vantis.

В ПЛИС FPGA множество конфигурируемых логических блоков (Configurable Logic Blocks (CLBs)) объединяется с помощью матрицы соединений. Логические блоки таких ПЛИС состоят из одного или нескольких относительно простых логических элементов, в основе которых лежат: четырехходовая таблица преобразования (Look-up table, LUT), программируемый мультиплексор, D-триггер, а также цепи управления. Характерными для FPGA-архитектур являются элементы ввода-вывода, позволяющие реализовать двунаправленный ввод/вывод, третье состояние и т.п. Особенностью современных ПЛИС FPGA является возможность тестирования узлов с помощью порта JTAG (B-scan), а также наличие внутреннего генератора (Osc) и схем управления последовательной конфигурацией.

Дальнейшее развитие архитектуры FPGA привело к появлению ПЛИС, имеющих встроенные модули памяти, позволяющие использовать ПЛИС без применения внешних запоминающих устройств. Благодаря большой логической емкости, удобной архитектуре, достаточно высокой надежности и удачному соотношению цена/логическая емкость данные ПЛИС удовлетворяют разнообразным требованиям, возникающим у разработчика.

В последние годы, наряду с заказными SoC, ПЛИС также становятся основой для «систем-на-кристалле». IP-ядра этих систем разрабатываются отдельно и могут многократно использоваться в различных проектах. Окончательная структура SoC-микросхемы на базе ПЛИС выполняется на основе IP-ядер с помощью программ автоматизации проектирования электронных устройств.

Идеология построения SoC подстегнула ведущих производителей ПЛИС к выпуску в конце 1998 – начале 1999 года изделий с эквивалентной емкостью миллион эквивалентных вентилях и более. Примером новых семейств ПЛИС, пригодных для реализации SoC, является семейство ApEX20K фирмы Altera.

Сходной архитектурой обладает и семейство Virtex фирмы Xilinx [9], которое характеризуется широким разнообразием высокоскоростных трассировочных ресурсов, наличием выделенного блочно-го ОЗУ, развитой логикой ускоренного переноса. ПЛИС данной серии обеспечивают высокие скорости межкристального обмена — до 200 МГц (стандарт HSTL IV). Кристаллы серии Virtex за счет развитой технологии производства и усовершенствованного процесса верификации имеют достаточно низкую стоимость (до 40 % от эквивалентной стоимости серии XC4000XL).

Увеличение эквивалентной логической емкости ПЛИС привело к тому, что в 1998–1999-х годах началось изменение отношения к программному обеспечению САПР ПЛИС со стороны как разработчиков программного обеспечения, так и пользователей. Если до конца 1990-х годов основным средством описания проекта являлся ввод схемы с помощью графических редакторов с использованием библиотек стандартных логических примитивов — логических элементов, простейших комбинационных и последовательностных функциональных узлов, аналогов стандартных интегральных схем малой и средней степени интеграции, то в настоящее время разработчиками широко используются языки описания аппаратуры (Hardware Description Languages) для реализации алгоритмов на ПЛИС. Причем в современных САПР поддерживаются как стандартизированные языки описания аппаратуры типа VHDL, Verilog, HDL, так и языки описания аппаратуры, разработанные компаниями-производителями ПЛИС специально для использования только в своих САПР и учитывающие архитектурные особенности конкретных семейств ПЛИС. Примером может служить AHDL (Altera Hardware Description Languages), поддерживаемый САПР MAX PLUS II и Quartus компании Altera. HDL-языки являются очень удобным средством описания различного рода интерфейсов, однако при реализации сложных вычислений разработчик не имеет возможности влиять на процедуру отображения написанного на HDL кода на физические ресурсы кристалла, что отрицательно сказывается на эффективности реализации. Кроме того, очевидна наглядность схемотехнического описания, поэтому при реализации сложного вычислительного алгоритма представляется предпочтительным использование графического редактора.

Фирма Xilinx предлагает IP-ядра для часто используемых элементов, включая блоки цифровой обработки сигналов, шинные интерфейсы, процессоры и периферийные устройства. Использование пакета LogiCORE™ фирмы Xilinx и аналогичных программ сторонних производителей для создания IP-ядер сокращает время разработки

проектов, уменьшает риски и позволяет получать наивысшую производительность. В дополнение к этому система генерации IP-ядер CORE Generator™ позволяет разработчикам реализовать собственные IP-ядра с предсказуемыми и воспроизводимыми временными характеристиками. Данный продукт имеет простой пользовательский интерфейс для генерации параметризованных IP-ядер, оптимизированных для FPGA Xilinx.

3. Построение вычислительных систем на основе ПЛИС. В то же время появление на рынке высокопроизводительных кристаллов ПЛИС с FPGA-архитектурой открывает широкие перспективы по их использованию в качестве элементной базы для построения высокопроизводительных вычислительных систем (суперкомпьютеров).

В настоящее время высокопроизводительные вычислительные системы, использующие ПЛИС в качестве элементной базы, принято разбивать на два типа. К первому типу относятся так называемые гибридные вычислительные системы, представляющие собой классические кластерные вычислители, в микропроцессорных узлах которых ПЛИС используются в качестве ускорителей вычислений. В качестве примера гибридных суперкомпьютеров можно привести XT4 компании Cray и RASC компании SiliconGraphics. В этих системах блоки программируемых сопроцессоров реализованы на базе ПЛИС, связанных высокоскоростными магистралями с основными процессорами и между собой.

Второй тип вычислительных систем, использующих ПЛИС в качестве элементной базы, представляют реконфигурируемые вычислительные системы (PBC). В PBC ПЛИС используются в качестве основного вычислительного элемента, в то время как процессорам общего назначения отводится второстепенная функция сопровождения и контроля над работой реконфигурируемой части системы.

Признанным лидером в области создания высокопроизводительных реконфигурируемых систем на основе вычислительных полей ПЛИС является таганрогская школа, основанная академиком А.В. Каляевым и представленная в настоящее время разработками НИИ многопроцессорных вычислительных систем Таганрогского радиотехнического университета (ныне Южного федерального университета) и Научно-исследовательского центра супер-ЭВМ и нейрокомпьютеров. Научным коллективом разрабатываются и производятся PBC класса супер-ЭВМ, в которых основным вычислительным ресурсом являются не микропроцессоры, а множество кристаллов ПЛИС, объединенных в вычислительные поля высокоскоростными каналами передачи данных. Спектр выпускаемых и проектируемых изделий до-

статочно широк: от полностью автономных малогабаритных реконфигурируемых ускорителей (вычислительных блоков), вычислительных модулей в настольном или стоечном конструктивном исполнении (на основе Xilinx Virtex-6, Virtex-7 и Virtex UltraScale) до вычислительных систем, состоящих из нескольких вычислительных шкафов, размещаемых в специально оборудованном машинном зале[10].

Начиная с 2001 года произошла смена четырех поколений реконфигурируемых вычислительных систем, построенных на базе ПЛИС, которая обусловлена выходом новых семейств ПЛИС, повышением вычислительной сложности задач, требующих постоянного роста производительности РВС. Первым поколением РВС были РВС с макропроцессорной архитектурой (РВС МПА), которая включала в свой состав группу базовых модулей, реализованных с помощью ПЛИС, и персональную ЭВМ (ПЭВМ). Каждый базовый модуль (БМ) представлял собой реконфигурируемое вычислительное устройство, построенное по тем же архитектурным принципам, что и система в целом. Подобный подход обеспечивал естественную реализацию структурно-процедурных параллельных программ для различных степеней параллелизма и конвейеризации вычислений.

В 2001 году был разработан и создан одноплатный базовый модуль РВС МПА на ПЛИС Xilinx Virtex E. Плата была изготовлена по 12-слойной технологии с двухсторонним монтажом элементов, при этом шесть слоев — сигнальных и шесть слоев — потенциальных: два слоя земли и по два слоя каждого напряжения питания 1.8 В и 3.3 В. На рисунке 1 показаны лицевая и тыльная стороны платы базового модуля РВС МПА.

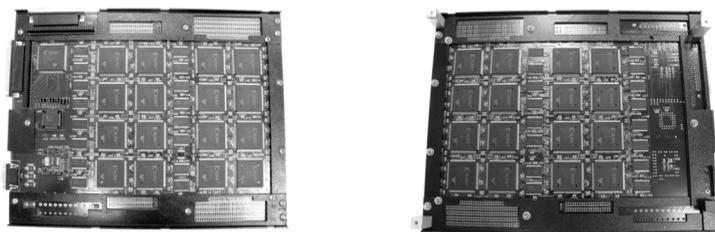


Рис. 1. Лицевая и тыльная стороны платы базового модуля РВС МПА

Плата БМ содержала 32 ПЛИС и 32 микросхемы оперативной памяти, конструктивно расположенные на обеих ее сторонах. Элементы на плате были размещены для обеспечения минимальной длины связей между элементами. Производительность базового модуля РВС МПА составляла $2,5 \cdot 10^{10}$ оп/сек, количество элементарных процессоров обработки данных — 64, а потребляемая мощность 30 Вт. На ос-

нове платы базового модуля РВС МПА был разработан и создан ряд модульно-наращиваемых многопроцессорных систем [1]. На смену РВС МПА пришли реконфигурируемые вычислительные системы второго поколения, к которым относят РВС с макрообъектной архитектурой. Программирование РВС с макрообъектной архитектурой обеспечивает пользователю возможность двухуровневого программирования архитектуры РВС. Такая организация программирования архитектуры РВС позволяет на порядок повысить эффективность вычислительного процесса по сравнению с РВС МПА. Первым представителем семейства РВС второго поколения с макрообъектной архитектурой была модульно-наращиваемая РВС «Медведь», базовый модуль которой представлен на рисунке 2.

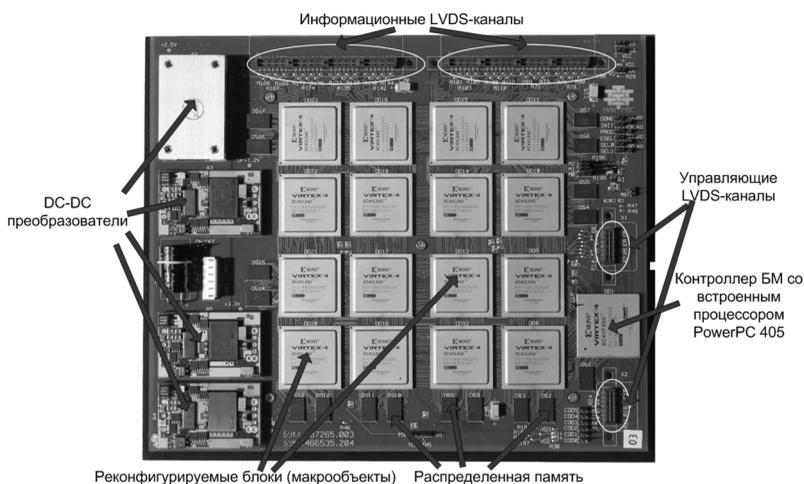


Рис. 2. Общий вид базового модуля РВС «Медведь»

Основу базового модуля составляют: 20-слойная печатная плата с двухсторонним монтажом элементов, вычислительное поле из 16 ПЛИС Xilinx XC4VLX40-10FF1148, контроллер обращений, 17 микросхем динамической памяти SDRAM, программируемый тактовый генератор, четыре малогабаритных DC-DC преобразователя напряжения, LVDS-разъемы для соединения базовых модулей по информационным каналам и другие элементы поверхностного монтажа. Производительность базового модуля РВС «Медведь» составляла 50 Гфлопс, частота базового модуля — 160 МГц, а потребляемая мощность не превышала 150 Вт.

Принципы макрообъектной архитектуры также были использованы при создании малогабаритного реконфигурируемого ускорителя персонального компьютера, предназначенного для решения вычисли-

тельно трудоемких фрагментов задач различных предметных областей. На рисунке 3 представлен внешний вид малогабаритного реконфигурируемого ускорителя персонального компьютера.

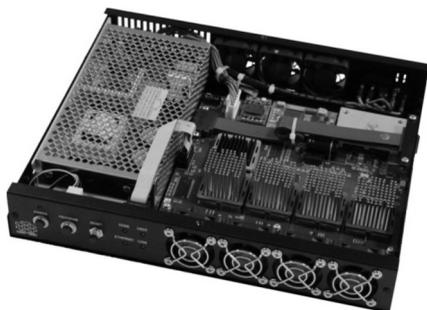


Рис. 3. Малогабаритный реконфигурируемый ускоритель персонального компьютера

В основе ускорителя лежал базовый модуль 4V4-25, выполненный на 18-слойной печатной плате размером 150×190 мм с двухсторонним расположением элементов. Производительность базового модуля 4V4-25 составляла 25 Гфлопс, частота базового модуля — 160 МГц, а потребляемая мощность не превышала 145 Вт.

Накопленный при создании РВС МПА и РВС «Медведь» опыт позволил перейти к реализации РВС третьего поколения, первым представителем которого является семейство РВС «Большая Медведица», которое строилось на основе трех типов базовых модулей: 16V5-75, 16V5-50 и 16S3-25. Базовый модуль 16V5-75 (как самый производительный) использовался в старших моделях семейства: системах РВС-5, РВС-1Р и РВС-0.2-РС. Базовые модули 16V5-50 и 16S3-25 входили в состав ускорителей персональных компьютеров РУПК-50 и РУПК-25.

Решающее поле базового модуля 16V5-75 выполнено на 16-ти ПЛИС Virtex-5 XC5VLX110-2FF1153 фирмы Xilinx, содержащих 11 миллионов эквивалентных вентилей. ПЛИС 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 микросхемами не превышают четыре сантиметра. Данные между удаленными ПЛИС передаются по транзитным каналам через промежуточные микросхемы, используя систему ортогональных связей. К периферийным ПЛИС решающего поля, расположенным на краях двумерной решетки 4×4, подключено 20 микросхем динамических SDRAM DDR2,

которые образуют распределенную память общим объемом 1,25 Гбайт. На рисунке 4 показан внешний вид платы базового модуля 16V5-75 с установленными электронными элементами и в сборе с системой охлаждения.

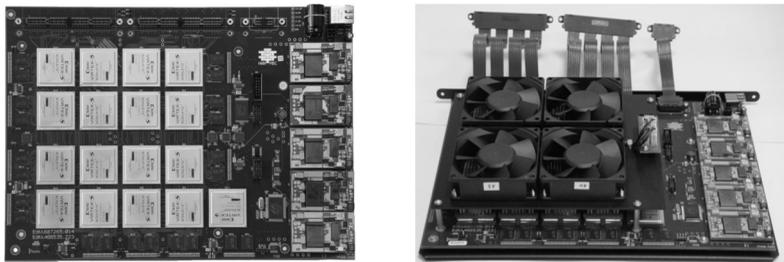


Рис. 4. Плата базового модуля 16V5-75

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. В целом базовый модуль 16V5-75 представляет собой мощный вычислительный узел производительностью свыше 75 (140) Гфлопс. На основе базового модуля 16V5-75 были разработаны рабочая станция PBC-0.2-PC и вычислительный блок PBC-0.2-ВБ производительностью 300 Гфлопс. На рисунке 5 показан общий вид рабочей станции PBC-0.2-ВБ со снятой верхней крышкой и вычислительный блок PBC-0.2-В.



Рис. 5. Рабочая станция PBC-0.2-PC и вычислительный блок PBC-0.2-ВБ

На основе вычислительных блоков PBC-0.2-ВБ была создана реконфигурируемая вычислительная система PBC-1P с пиковой производительностью 1200 Гфлопс, предназначенная для оснащения научных центров, проводящих исследования в области физики, химии, биологии,

космоса, построения информационно-управляющих систем для управления потенциально опасными производствами, решения задач аэрокосмической и автомобильной промышленности, энергетики и т.п.

Старшим представителем РВС третьего поколения являлась система РВС-5 (рисунок 6) с пиковой производительностью 6000 Гфлопс, содержащая пять стоек РВС-1Р, соединенных между собой коммутаторами Ethernet с общим управлением.

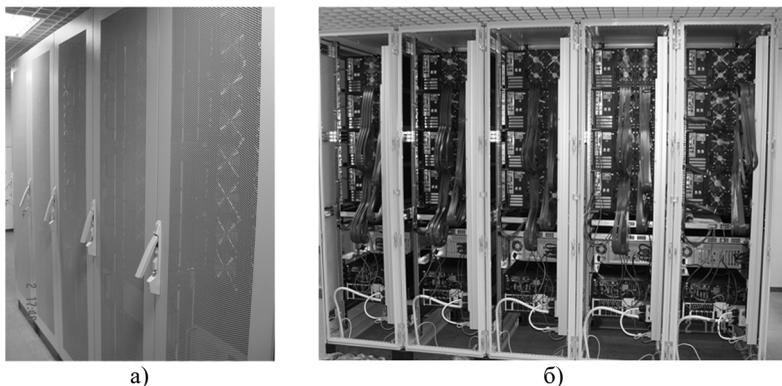


Рис. 6. РВС-5: а) вид спереди; б) вид сзади

РВС-5, установленная в НИВЦ МГУ им. М.В. Ломоносова, содержит 20 вычислительных блоков РВС-0.2-ВБ, 80 базовых модулей 16V5-75, 16384 элементарных процессора (IEEE-754), осуществляющих обработку 64-разрядных данных по стандарту IEEE-754 на тактовой частоте 330 МГц с производительностью более 6000 Гфлопс. Обмен между блоками осуществляется по внешним интерфейсам LVDS, Gigabit Ethernet на частоте 640 МГц.

Растущие требования к производительности РВС обусловили четырехкратное повышение плотности компоновки вычислительных систем и создание в 2012-2014 годах на основе ПЛИС Xilinx Virtex-7 РВС четвертого поколения на базе вычислительных модулей (ВМ) 24V7-750 («Плеяда») и «Тайгета» [11].

Вычислительный модуль 24V7-750, представленный на рисунке 7, содержит четыре платы 6V7-180; управляющий модуль УМ-7; подсистему питания; подсистему охлаждения и другие подсистемы. Производительность ВМ 24V7-750 составляет 2,58 Тфлопс при обработке 32-разрядных данных с плавающей запятой.



Рис. 7. Вычислительный модуль 24V7-750: а) со снятой верхней крышкой; б) с установленной верхней крышкой)

ВМ 24V7-750 использовался для создания по государственному контракту №14.527.12.0004 от 03.10.2011 реконфигурируемой вычислительной системы PBC-7, которая содержит 24 ВМ 24V7-750 с вычислительным полем из 576 микросхем ПЛИС Virtex-7 XC7V585T-FFG1761 объемом 58 миллионов эквивалентных вентилях каждая, конструктивно объединенных в один вычислительный шкаф высотой 47U с пиковой производительностью 10^{15} операций с фиксированной запятой в секунду. Производительность PBC-7 при комплектации от 24 до 36 ВМ 24V7-750 составит от 62 до 93 Тфлопс при обработке 32-разрядных данных с плавающей запятой и 19,4–29,4 Тфлопс при обработке 64-разрядных данных с плавающей запятой. Области применения PBC-7 и вычислительных комплексов на ее основе являются решение задач цифровой обработки сигналов и многоканальная цифровая фильтрация.

На основе ПЛИС Virtex-7 также был разработан ВМ «Тайгета» в конструктивном исполнении высотой 2U, предназначенный для установки в стандартную 19" вычислительную стойку. ВМ «Тайгета», представленный на рисунке 8а, содержит четыре ПВМ 8V7-200, представленных на рисунке 8б, соединенных быстрыми LVDS-каналами; встроенную управляющую ЭВМ; систему питания; систему управления; систему охлаждения и другие подсистемы.

ПВМ 8V7-200, лежащая в основе ВМ «Тайгета», представляет собой 20-слойную печатную плату с двухсторонним монтажом элементов, на которой располагаются 8 ПЛИС типа XC7VX485T-1FFG1761, содержащих 48,5 миллионов эквивалентных вентилях, 16 микросхем распределенной памяти SDRAM типа DDR2 общим объемом 2 Гбайт, интерфейсы LVDS и Ethernet и другие компоненты. Производительность ВМ «Тайгета» составляет 2,66 Тфлопс при обработке 32-разрядных данных с плавающей запятой. Производительность

ность PBC на основе ВМ «Тайгета» при комплектации 18 ВМ «Тайгета» составляет 48 Тфлопс при обработке 32-разрядных данных с плавающей запятой и 23 Тфлопс при обработке 64-разрядных данных с плавающей запятой.

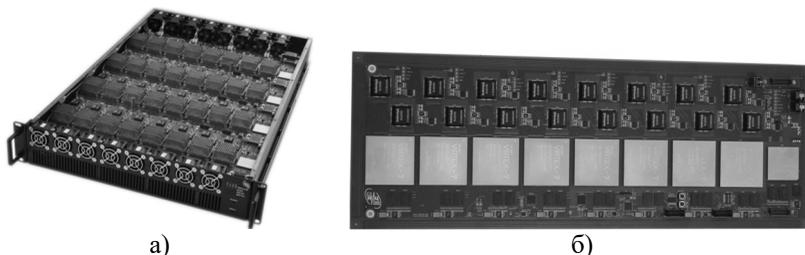


Рис. 8. ВМ «Тайгета»: а) со снятой верхней крышкой, б) плата 8V7-200

Переход к четвертому поколению PBC был обусловлен проблемами системы охлаждения. Согласно полученным экспериментальным данным переход от семейства ПЛИС Virtex-6 к следующему семейству ПЛИС Virtex-7 ведет к росту максимальной температуры ПЛИС на 11...15°C. Поэтому дальнейшее развитие технологий изготовления ПЛИС и переход на следующее семейство ПЛИС Virtex Ultra Scale приведет к росту величины перегрева ПЛИС на дополнительные 10...15°C, что сместит зону их рабочей температуры до 80...85°C с выходом за допустимые пределы рабочей температуры ПЛИС (65...70°C), что негативно скажется на их надежности.

Практический опыт эксплуатации больших вычислительных комплексов на основе ВМ PBC показывает, что воздушные системы охлаждения подошли к своему тепловому пределу. Поэтому использование воздушного охлаждения при переходе к следующему поколению ПЛИС Virtex UltraScale, содержащих порядка 100 млн. эквивалентных вентиляей и обладающих энергопотреблением не менее 100 Вт на кристалл ПЛИС, не позволяет обеспечить стабильную и надежную работу PBC при заполнении кристаллов на 85-95% от доступного аппаратного ресурса.

Это обстоятельство потребовало перехода к принципиально иному способу охлаждения, позволившему сохранить темпы роста производительности PBC для перспективных проектируемых семейств ПЛИС фирмы Xilinx: Virtex UltraScale, Virtex UltraScale+, Virtex UltraScale2 и др.

Большинство разработчиков вычислительной техники видит выход в применении систем жидкостного охлаждения, которые на сегодняшний день представляются наиболее перспективной областью

разработки для охлаждения современных высоконагруженных электронных компонентов вычислительных систем.

Существенными преимуществами всех жидкостных схем охлаждения являются лучшая теплоемкость жидкостей по сравнению с воздухом (от 1500 до 4000 раз) и больший коэффициент теплоотдачи (увеличение до 100 раз). Для охлаждения одного современного кристалла ПЛИС необходимо наличие 1 м^3 воздуха или $0,00025 \text{ м}^3$ (250 мл) воды в минуту. На перекачивание 250 мл воды тратится намного меньше электроэнергии, чем для перекачивания 1 м^3 воздуха.

4. Жидкостное охлаждение для реконфигурируемых вычислительных систем. Развитие вычислительных технологий приводит к разработке все более производительной в плане вычислительной мощности, а значит, и более тепловыделяющей вычислительной техники. Утилизация выделяемого тепла осуществляется системой охлаждения электронных компонентов, которая реализует перенос тепла от более горячего тела (охлаждаемый объект) к менее горячему (система охлаждения). При постоянном нагреве охлаждаемого объекта температура системы охлаждения повышается и со временем сравняется с температурой охлаждаемого объекта, передача тепла прекратится, что вызовет его перегрев. Система охлаждения защищается от перегрева с помощью холодного вещества — хладагента (теплоносителя), характеристиками эффективности охлаждения которого являются теплоемкость и теплоотдача. Перенос тепла, как правило, осуществляется либо с помощью механизма теплопроводности, требующего физического контакта теплоносителя с охлаждаемым объектом, либо с помощью механизма конвективного теплообмена с хладагентом, который связан с физическим переносом свободно циркулирующего охлаждающего вещества.

Для организации переноса тепла к хладагенту необходимо организовать тепловой контакт системы охлаждения с хладагентом. Для этого применяют различные *радиаторы* — устройства для рассеивания тепла в хладагенте, которые устанавливают на наиболее нагреваемые компоненты вычислительных систем. Для повышения эффективности переноса тепла от электронного компонента к радиатору между ними устанавливается *термоинтерфейс* — слой теплопроводящего состава (обычно многокомпонентного) между охлаждаемой поверхностью и отводящим тепло устройством, применяемый для уменьшения термического сопротивления между двумя соприкасающимися поверхностями. Современные процессоры и ПЛИС нуждаются в охлаждающих устройствах с как можно более низким термическим сопротивлением, так как в настоящий момент даже самые продвинутые ра-

диаторы и термоинтерфейсы не справляются с этой задачей при воздушном охлаждении.

Системы воздушного охлаждения достаточно успешно использовались для охлаждения суперкомпьютеров до 2012 года. С ростом мощности и степени интеграции применяемых кристаллов микропроцессоров и ПЛИС использование воздушных систем охлаждения для разрабатываемых перспективных суперкомпьютеров, в том числе вычислительных систем гибридного типа, практически достигло своего предела. Поэтому большинство разработчиков вычислительной техники видит выход в применении систем жидкостного охлаждения, которые на сегодняшний день представляются наиболее перспективной областью разработки для охлаждения современных высоконагруженных электронных компонентов вычислительных систем.

Существенными преимуществами всех жидкостных схем охлаждения являются лучшая теплоемкость жидкостей по сравнению с воздухом (от 1500 до 4000 раз) и больший коэффициент теплоотдачи (увеличение до 100 раз). Для охлаждения одного современного кристалла ПЛИС необходимо наличие 1 м^3 воздуха или $0,00025 \text{ м}^3$ (250 мл) воды в минуту. На перекачивание 250 мл воды тратится намного меньше электроэнергии, чем для перекачивания 1 м^3 воздуха. Тепловой поток, передаваемый одинаковыми поверхностями при традиционных скоростях теплоносителя, становится при жидкостном охлаждении в 70 раз большим, чем при воздушном охлаждении. Дополнительным преимуществом является использование традиционных, достаточно надежных и недорогих рабочих элементов (насосы, теплообменники, клапаны, механизмы управления и т.д.). Для организаций, которые работают с высокой плотностью оборудования при высоких тепловых режимах, жидкостное охлаждение, по существу, оказывается единственным решением проблемы организации охлаждения современных вычислительных систем. Дополнительным резервом повышения эффективности жидкостного охлаждения является улучшение исходных параметров теплоносителя: увеличение скорости, снижение температуры, создание турбулентного режима течения, повышение теплоемкости, снижение вязкости.

В жидкостных системах охлаждения вычислительной техники роль хладагента выполняет жидкость (вода либо диэлектрические жидкости). Нагретые электронные компоненты отдают тепло постоянно циркулирующему хладагенту — жидкости, которая после ее охлаждения во внешнем теплообменнике вновь поступает для охлаждения нагретых электронных компонентов. Жидкостные системы охлаждения можно разделить на системы закрытого типа, где нет прямого контакта

между жидкостью и электронными компонентами на печатных платах [12, 13], и системы открытого типа (погружные), в которых жидкость непосредственно омывает электронные компоненты. Каждый тип систем жидкостного охлаждения имеет свои достоинства и недостатки.

В системах закрытого типа все тепловыделяющие элементы печатной платы накрываются одним или несколькими плоскими пластинами с каналом для прокачки жидкости [14, 15, 16]. Так, например, в суперкомпьютере «СКИФ-Аврора» использовался принцип «одна охлаждающая пластина на одну печатную плату». Соответственно пластина имела сложный рельеф, чтобы иметь плотный тепловой контакт с каждой микросхемой. В суперкомпьютере IBM Aquasar используется принцип «одна охлаждающая пластина на одну (горячую) микросхему». В любом случае каналы пластин объединяются коллекторами в единый контур, соединенный с общим радиатором (или другим теплообменником), обычно расположенным за пределами корпуса и/или шкафа или даже машинного зала. С помощью насоса через пластины прокачивается теплоноситель, который отводит тепло от элементов вычислителя и рассеивает его через теплообменник. В такой системе требуется обеспечить доступ теплоносителя к каждому тепловыделяющему элементу вычислителя, что означает довольно сложную «трубопроводную систему», большое количество герметичных соединений. При этом требования к возможности обслуживания печатных плат без серьезного демонтажа в системе охлаждения неизбежно означают применение специальных жидкостных разъемов, которые обеспечивают герметичность соединений с одновременной легкостью операций по сборке/разборке системы.

В системах жидкостного охлаждения закрытого типа можно в качестве хладагента использовать обычную воду или растворы гликолей. Однако утечка теплоносителя может привести к возможному попаданию электропроводной жидкости на незащищенные контакты рабочих плат охлаждаемого компьютера, что, в свою очередь, может быть фатальным как для отдельных электронных компонентов, так и для вычислительной системы в целом. Для устранения неисправности требуются остановка всего комплекса, проверка системы электропитания и ее просушка. В системах управления и мониторинга таких компьютеров всегда предусматривают многочисленные внутренние датчики влажности и протечек. Для решения проблемы протечек часто используется метод, основанный на использовании в системе охлаждения отрицательного давления воды, когда вода не накачивается под давлением, а высасывается, что практически исключает утечку влаги. При нарушении герметичности в систему охлаждения попадет

воздух, а вода оттуда не вытечет. Для детектирования утечек используются специальные датчики, а модульная конструкция позволяет осуществлять ремонт без прерывания работы системы в целом. Однако все эти мероприятия значительно усложняют конструкцию гидравлической системы.

Еще одна проблема систем жидкостного охлаждения закрытого типа — проблема «точки росы». Воздух центра обработки данных (ЦОД) входит в контакт с пластинами охлаждения. Это значит, что если какие-то участки этих пластин слишком холодные, а воздух в ЦОД более теплый и не слишком сухой, то на пластинах может сконденсироваться влага. Последствия этого процесса аналогичны последствиям протечек. Проблема решается либо переходом на охлаждение горячей водой, что неэффективно, либо контролем и поддержанием параметров температуры и влажности воздуха в ЦОД, что сложно и дорого.

Конструкция еще более усложняется, когда необходимо охладить несколько компонентов потоком воды, пропорциональным их тепловыделению. Не считая разветвленных трубок, приходится применять сложные регулирующие приборы (простыми тройниками и крестовинами не обойдешься). Альтернативный вариант — использовать конструкцию с заводскими регулировками потоков; но в этом случае пользователь лишен возможности существенно изменить конфигурацию охлаждаемых вычислительных модулей.

К достоинствам систем жидкостного охлаждения закрытого типа относятся:

- возможность использования в качестве хладагента воды или водных растворов, преимуществами которых являются доступность, прекрасные теплотехнические свойства (теплопроводность, теплоемкость, вязкость), простота и сравнительная безопасность эксплуатации;
- наличие большого числа унифицированных механизмов, узлов и деталей для систем водоснабжения, которые можно использовать;
- большой опыт эксплуатации систем с водяным охлаждением в промышленности.

При этом системы жидкостного охлаждения закрытого типа обладают целым рядом существенных недостатков, сдерживающих их массовое применение:

- сложностью определения места утечки воды;
- зачастую катастрофическими последствиями не обнаруженных своевременно протечек;
- технологическими сложностями ликвидации протечек (необходимость отключения всего вычислительного шкафа, что не всегда возможно и удобно);

- необходимостью поддержания микроклимата помещения вычислительного зала (проблема «точки росы»);
- проблемой охлаждения остальных элементов на плате вычислительного модуля PBC. Даже при небольшом изменении конфигурации PBC необходимо изготавливать новый теплообменник;
- проблемой электрохимической коррозии при использовании алюминиевых теплообменников либо массогабаритными ограничениями при использовании более устойчивых медных теплообменников (алюминий в три раза легче меди);
- необходимостью удаления воздуха из системы охлаждения как при пуско-наладке, так и в процессе эксплуатации;
- сложностью компоновки вычислительных модулей в шкафу при большом количестве фитингов, необходимых для подключения каждого вычислительного модуля;
- необходимостью использования специализированного вычислительного шкафа со значительными массогабаритными характеристиками.

В системах жидкостного охлаждения открытого типа основным компонентом является хладагент, представляющий собой диэлектрическую жидкость, как правило, на основе белого минерального масла, благодаря которому теплоаккумулирующая способность хладагента гораздо выше, чем у воздуха при том же объеме. Конструктивно такие системы представляют собой заполненную хладагентом ванну (в том числе устанавливаемую в вычислительном шкафу), в которой размещаются печатные платы и серверы вычислительного оборудования. Выделяемое электронными компонентами тепло рассеивается циркулирующим в пределах всего объема ванны хладагентом. Достоинствами погружной системы охлаждения являются простота конструкции и адаптация под меняющуюся геометрию печатных плат, отсутствие сложной системы коллекторов и жидкостных разъемов, отсутствие проблем управления потоками жидкости и проблемы «точки росы», повышение надежности и снижение стоимости изделия.

Основная проблема систем жидкостного охлаждения открытого типа заключается в химическом составе используемого хладагента, который должен удовлетворять жестким требованиям по теплопроводности, электропроводности, вязкости, токсичности, пожаробезопасности, стабильности основных параметров, и при этом стоимость жидкости должна быть разумной.

Системы жидкостного охлаждения открытого типа обладают следующими преимуществами:

- нечувствительностью к протечкам и их последствиям, возможностью эксплуатации системы даже при наличии местных протечек хладагента;
- нечувствительностью к климатическим характеристикам помещения вычислительного зала;
- решением проблемы охлаждения остальных элементов РВС, так как плата вычислительного модуля погружена в хладагент;
- возможностью изменения конфигурации платы вычислительного модуля без изменения системы охлаждения;
- простотой гидравлической балансировки системы благодаря отсутствию сложной системы коллекторов;
- возможностью использования как унифицированных механизмов, узлов и деталей для гидравлических систем общего машиностроения, так и накопленного опыта эксплуатации электрооборудования с использованием трансформаторных масел;
- повышением общей надежности системы жидкостного охлаждения.

Недостатками систем жидкостного охлаждения открытого типа являются:

- необходимость дополнительного насосного и теплообменного оборудования для улучшения теплотехнических свойств (теплопроводность, теплоемкость, вязкость) хладагента, в роли которого используются специальные диэлектрические органические жидкости;
- необходимость обучения обслуживающего персонала и соблюдение повышенных мер безопасности при работе с хладагентом;
- необходимость более частой очистки помещения машинного зала из-за высокой проникающей способности хладагента, особенно в случае протечки;
- необходимость специальной оснастки для регламентных и аварийных эксплуатационных операций (монтаж/демонтаж вычислительного модуля, залив/слив хладагента и т.д.);
- повышение стоимости эксплуатации из-за необходимости регулярной замены хладагента по истечении срока службы и необходимость организации учета (транспортирование, прием, учет, хранение, выдача, утилизация хладагента и т.д.) хладагента в организации.

Оценивая приведенные достоинства и недостатки двух систем жидкостного охлаждения, можно с достаточной степенью уверенности отметить более весомые преимущества систем жидкостного охлаждения электронных компонентов вычислительных систем открытого типа. Поэтому для вычислительных модулей РВС, проектируемых на основе перспективных семейств ПЛИС, целесообразно использовать жидкостное

охлаждение, в частности непосредственное погружение плат вычислительных модулей в жидкостный хладагент на основе минерального масла.

Особенностями выпускаемых в Научно-исследовательском центре супер-ЭВМ и нейрокомпьютеров РВС являются наличие не менее 6-8 кристаллов ПЛИС на одной печатной плате и высокая плотность компоновки платы, что существенно увеличивает число тепловыделяющих элементов по сравнению с микропроцессорными модулями. Это затрудняет применение существующих готовых решений погружных систем [17-19] и требует дополнительных технических и конструктивных решений для эффективного непосредственного жидкостного охлаждения вычислительных модулей РВС.

5. Реконфигурируемая вычислительная система на основе ПЛИС Xilinx UltraScale. Поэтому с 2013 года в НИЦ СЭ и НК активно развивается направление по созданию РВС нового поколения на основе собственной системы жидкостного охлаждения для печатных плат с высокой степенью компоновки и большим количеством тепловыделяющих электронных компонентов. В основу критериев проектирования вычислительного модуля (ВМ) нового поколения для РВС с жидкостным охлаждением открытого типа были положены следующие принципы:

- основой конфигурации вычислительного шкафа является вычислительный модуль с размерами 3U на 19" и автономной циркуляцией охлаждающей жидкости;

- в одном вычислительном шкафу стандартной высоты 47U должно размещаться не менее 12 вычислительных модулей с жидкостным охлаждением;

- в одном вычислительном модуле должно размещаться 12-16 плат с установленными на них кристаллами ПЛИС;

- на каждой печатной плате должно быть размещено до восьми ПЛИС с выделяемым тепловым потоком порядка 100 Вт от каждой ПЛИС;

- для охлаждения жидкости должна использоваться традиционная система водяного охлаждения на базе промышленных чиллеров.

Основным элементом модульной реализации погружной системы жидкостного охлаждения электронных компонентов вычислительных систем открытого типа является реконфигурируемый вычислительный модуль нового поколения, технологический образец которого представлен на рисунке 9. ВМ нового поколения состоит из вычислительной секции, теплообменной секции, корпуса, насоса, теплообменника и штуцера. В корпусе, образующем основу вычислительной секции, размещены герметичный контейнер с диэлектрической охлаждающей жидкостью, содержащий электронные устройства с компонен-

тами, которые выделяют тепло в процессе работы. В качестве электронных устройств могут быть вычислительные модули (не менее 12-16 шт.), платы управления, ОЗУ, блоки питания, накопители, дочерние платы и так далее. Сверху вычислительная секция закрыта крышкой.

К вычислительной секции примыкает теплообменная секция, содержащая насос и теплообменник. Насос обеспечивает циркуляцию хладагента в ВМ по следующему замкнутому контуру: из вычислительного модуля нагретый хладагент поступает в теплообменник, охлаждается там, из теплообменника охлажденный хладагент вновь поступает в вычислительный модуль, охлаждает там нагретые электронные компоненты, там же нагревается и затем вновь поступает в теплообменник и так далее. Теплообменник подключен к внешнему теплообменному контуру через штуцеры и предназначен для охлаждения хладагента с помощью вторичной охлаждающей жидкости. В качестве теплообменника можно использовать пластинчатый теплообменник, в котором первичный и вторичный контуры разделены между собой. Тогда в качестве вторичной охлаждающей жидкости можно использовать обыкновенную воду, охлаждаемую промышленным чиллером. При этом чиллер может располагаться вне здания серверной и соединяться с реконфигурируемыми вычислительными модулями через стационарную систему инженерных коммуникаций.

Вычислительная и теплообменная секции механически соединены между собой, образуя единый реконфигурируемый вычислительный модуль. Для эксплуатации реконфигурируемого вычислительного модуля требуется подключение к источнику вторичной охлаждающей жидкости (через краны), к источнику энергоснабжения и к сетевому концентратору (через электрические разъемы).

В корпусе вычислительного шкафа ВМ размещены друг над другом. Их количество ограничивается линейными размерами шкафа, техническими возможностями помещения вычислительного зала и подводимых инженерных коммуникаций. Каждый ВМ вычислительного шкафа подключается к источнику вторичной охлаждающей жидкости с помощью подающих обратных коллекторов через штуцеры (или балансировочные клапаны) и гибкие трубопроводы; подключение к источнику энергоснабжения и сетевому концентратору осуществляется через электрические разъемы. Через штуцеры (или балансировочные клапаны) подают холодную и отводят нагретую вторичную охлаждающую жидкость в стационарную систему инженерных коммуникаций, подключенную к шкафу. Множество вычислительных шкафов, размещенных в одном или нескольких помещениях, образует вычислительный комплекс. Для эксплуатации вычислительный ком-

плекс также подключается к источнику вторичной охлаждающей жидкости, к источнику энергоснабжения, модулю управления вычислительным комплексом.

Рассмотренная модульная реализация системы жидкостного охлаждения электронных компонентов вычислительных систем открытого типа, помимо свойственных системам открытого типа достоинств, дополнительно обладает рядом преимуществ:

- платы вычислительных модулей и реконфигурируемые вычислительные модули идентичны, относительно автономны и взаимозаменяемы. При выходе из строя одного ВМ или при необходимости проведения профилактических работ нет необходимости полностью отключать вычислительный шкаф и останавливать рабочую задачу;

- достигается повышенная плотность компоновки ПЛИС в ВМ;

- предложенное техническое решение позволяет при необходимости наращивать производительность реконфигурируемых вычислительных модулей без кардинального увеличения габаритных размеров (более производительные насос и теплообменник позволяют вписаться в выбранные габариты). Увеличение количества плат вычислительных модулей незначительно увеличит размер реконфигурируемого вычислительного модуля по глубине при неизменной плотности компоновки.

В силу простоты конструкции теплообменной секции реконфигурируемого вычислительного модуля значительно повышается его надежность.

Вычислительный 19” шкаф суперкомпьютера имеет следующие технические характеристики:

- стандартный вычислительный шкаф высотой 47U;

- 12 вычислительных модулей высотой 3U с жидкостным охлаждением;

- каждый вычислительный модуль содержит 12 плат мощностью 800 Вт каждая;

- каждая плата содержит восемь кристаллов ПЛИС Kintex UltraScale - XCKU095-1FFVB2104C по 95 млн. эквивалентных вентилей (134 400 логических блоков) каждый;

- производительность ВМ нового поколения — 105 Тфлопс;

- производительность вычислительного шкафа с 12 ВМ — 1 Пфлопс;

- потребляемая мощность вычислительного шкафа с 12 ВМ — 124 кВт.

Производительность одного вычислительного шкафа с системой жидкостного охлаждения, содержащего 12 ВМ с 12 платами каждый, в 6,55 раз превышает производительность аналогичного шкафа с ВМ

«Тайгета». При этом производительность одного ВМ нового поколения по сравнению с ВМ «Тайгета» повышена в 8,74 раза. Такое качественное повышение удельной производительности системы обеспечено более чем трехкратным повышением плотности компоновки системы за счет оригинальных конструктивных решений, а также повышением тактовой частоты и увеличением числа вентиляторов в кристалле.

Для проверки технических, технологических решений и определения ожидаемых технико-экономических и эксплуатационных характеристик разрабатываемой высокопроизводительной реконфигурируемой вычислительной системы с жидкостным охлаждением был создан ряд макетов, опытных и технологических образцов. На рисунке 9 представлен технологический образец ВМ нового поколения, для которого были разработаны новые конструкции печатных плат и вычислительных модулей, характеризующиеся высокой плотностью компоновки.

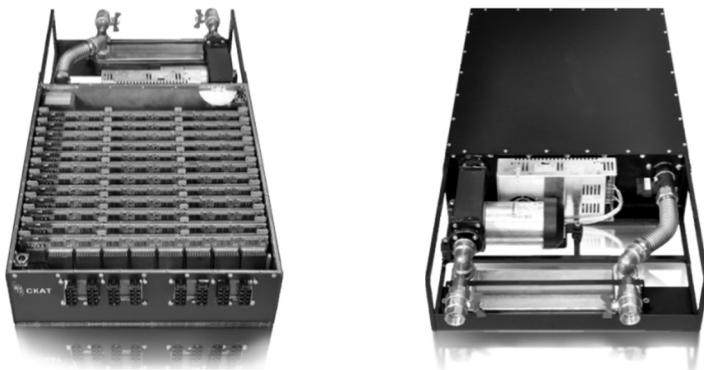


Рис. 9. Технологический образец ВМ нового поколения

Для создания эффективной погружной системы охлаждения был разработан диэлектрический хладагент, обладающий наилучшей электрической прочностью, высокой теплопроводностью, максимально возможной теплоемкостью при низкой вязкости. На основе трансформаторного масла ГК методом вакуумной перегонки было получено масло с пониженной вязкостью «Масло маловязкое диэлектрик МД-4,5 для охлаждения электронных компонентов ЭВМ», на него разработаны технические условия ТУ 38.401-58-421-2015 и получены рекомендации для его применения. Масло МД-4,5 прошло всестороннее исследование в теплотехнической лаборатории НИЦ СЭ и НК на технологическом образце вычислительного модуля с погружной системой жидкостного охлаждения открытого типа. Цикл проведенных лабораторных и эксплуатационных испытаний подтвердил целесообразность использования масла МД-4,5 для охлаждения электронных компонен-

тов ЭВМ и применение маломощных насосов для его циркуляции (вследствие пониженной вязкости).

При разработке ВМ нового поколения получен ряд прорывных технических решений: разработаны погружной блок питания для напряжения 380В и преобразователь DC/DC 380/12 В, обеспечена минимальная высота платы ВМ 100 мм, разработана погружная плата управления собственного производства. Для подсистемы охлаждения ВМ нового поколения определены необходимые элементы системы охлаждения: выбран оригинальный термоинтерфейс, разработана собственная конструкция радиатора ПЛИС малой высоты для конвективного теплообмена, выбраны оптимальный для используемого хладагента насос и теплообменник. Также определены конструкция компенсатора объемного расширения хладагента и элементы контроля подсистемы охлаждения: оптические датчики уровня и датчик протока. Разработанные решения по конструкции системы охлаждения и циркуляции хладагента позволяют эффективно решать проблему теплоотвода с наиболее нагретых компонентов ВМ.

Комплекс разработанных решений погружной системы жидкостного охлаждения позволил обеспечить в рабочем режиме ВМ температуру хладагента не более 33 °С при потребляемой мощности 91 Вт для каждой ПЛИС (8736 Вт для ВМ). При этом максимальная температура ПЛИС не превышает 55°С, что свидетельствует о резерве мощности разработанной погружной системы жидкостного охлаждения, позволяющей обеспечить эффективное охлаждение для проектируемых перспективных семейств ПЛИС фирмы Xilinx (UltraScale+, UltraScale 2 и т.д.).

6. Заключение. Использование воздушных систем охлаждения для разрабатываемых суперкомпьютеров практически достигло своего предела, что обусловлено снижением эффективности охлаждения с ростом потребляемой (и отдаваемой) мощности, вызванным увеличением степени интеграции кристаллов микропроцессоров и других микросхем. Поэтому использование жидкостного охлаждения для организации работы современных вычислительных систем является приоритетным направлением совершенствования систем охлаждения с широкими перспективами дальнейшего развития. Жидкостное охлаждение вычислительных модулей PBC, содержащих не менее восьми кристаллов ПЛИС с высокой плотностью компоновки, обладает спецификой по сравнению с охлаждением кристаллов микропроцессоров, что требует разработки специализированной погружной системы охлаждения. Разработанная оригинальная жидкостная система охлаждения для вычислительного модуля PBC нового поколения позволяет обеспечить высокие эксплуатационные показатели системы — макси-

мальную температуру ПЛИС, не превышающую 55 °С при температуре хладагента не более 33 °С в рабочем режиме. Полученные прорывные решения погружной системы жидкостного охлаждения позволяют в пределах одного вычислительного шкафа высотой 47U разместить не менее 12 ВМ нового поколения с суммарной производительностью свыше 1 Пфлопс. Резерв мощности системы жидкостного охлаждения ВМ нового поколения позволяет обеспечить эффективное охлаждение не только для существующих, но и для проектируемых перспективных семейств ПЛИС фирмы Xilinx UltraScale+ и UltraScale 2.

Литература

1. *Kalyaev I.A., Levin I.I., Semernikov E.A., Shmoilov V.I.* Reconfigurable multipipeline computing structures // New York: Nova Science Publishers. USA. 2012. 330 p.
2. *Carter W.S.* The evolution of programmable logic // VLSI Circuits. Digest of Technical Papers. 1991 Symposium on IEEE. 1991. pp. 43–46.
3. *Wilson R.* In the Beginning. URL: https://www.altera.com/solutions/technology/system-design/articles/_2013/in-the-beginning.html (дата обращения: 04.04.2016).
4. *Chartrand L.* Advanced Digital Systems: Experiments and Concepts with CPLDs // Cengage Learning. 2005. 268 p.
5. *Power Methodology Guide.* UG786 (v14.5). 2013. 54 p. URL: http://www.xilinx.com/support/documentation/sw_manuals/xilinx13_1/ug786_PowerMethodology.pdf (дата обращения 04.04.2016).
6. *User guide Microsemi.* 158 p. URL: http://www.actel.ru/files/uploads/Novosti/Download/Libero_UG_rus.pdf (дата обращения 04.04.2016).
7. *Barkalov A., Titarenko L.* Evolution of programmable logic // Logic Synthesis for FSM-Based Control Units. Springer Berlin Heidelberg. 2009. LNEE 53. pp. 53–75.
8. *Hutton M.* Architecture and CAD for FPGAs // Proceedings of the 17th Symposium on Integrated Cuits and Systems Design (SBCCI2004). 2004. pp. 3.
9. *Kuon I., Tessier R., Rose J.* FPGA Architecture: Survey and Challenges // Foundations and Trends in Electronic Design Automation. 2008. vol. 2. no. 2. pp. 135–253.
10. *Levin I.I., Dordopulo A.I., Doronchenko Y.I., Raskladkin M.K.* Reconfigurable computer system on the base of Virtex UltraScale FPGAs with liquid cooling // Proceedings of international scientific conference “Parallel computer technologies” (PaCT’2016). 2016. pp. 221–230.
11. *Kalyaev I.A., Levin I.I., Dordopulo A.I., Slasten L.M.* Reconfigurable Computer Systems Based on Virtex-6 and Virtex-7 FPGAs // IFAC Proceedings Volumes. 2013. vol. 12. no. 1. pp. 210–214.
12. *Li L., Zheng W., Wang X., Wang X.* Data center power minimization with placement optimization of liquid-cooled servers and free air cooling // Sustainable Computing: Informatics and Systems. 2016. vol. 11. no. 1. pp. 3–15.
13. *Sun L. et al.* Corrosion investigation of the inlet section of REAC pipes in the refinery // Engineering Failure Analysis. 2016. vol. 66. no. 1. pp. 468–478.
14. *Абрамов С.М. и др.* Суперкомпьютеры ряда 4 семейства “СКИФ”: Штурм вершины суперкомпьютерных технологий // Вестник Нижегородского университета им. Н.Н. Лобачевского. 2009. № 5. С. 200–210.
15. *Левшин И., Орлов Д., Суомин А., Челищев А.* Не революция, но переворот // М.: Суперкомпьютеры. 2012. № 2(10). С. 6–21. URL: <http://admin.clus-ter.sfu>

kras.ru/data/uploads/2015/02/12/ Supercomputers_10-2012.pdf. (дата обращения: 04.04.2016).

16. Новое суперкомпьютерное решение «Т-Платформы». А-Класс. Пресс-релиз от 25.06.2014. URL: https://www.pcweek.ru/infrastructure/news-company/detail_print.php?ID=164557&print=Y. (дата обращения 04.04.2016).
17. *Shah J.M., Eiland R., Siddarth, A. Agonafer D.* Effects of mineral oil immersion cooling on IT equipment reliability and reliability enhancements to data center operations // Proceedings of the 15th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm 2016). 2016. pp. 316–325.
18. *Gess J.L., Bhavnani S.H., Johnson R.W.* Experimental Investigation of a Direct Liquid Immersion Cooled Prototype for High Performance Electronic Systems // IEEE Transactions on Components, Packaging and Manufacturing Technology. 2015. vol. 5. no. 10. pp. 1451–1464.
19. *Gess J., Dreher T., Bhavnani S., Johnson W.* Effect of flow guide integration on the thermal performance of high performance liquid cooled immersion server modules // ASME 2015 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems collocated with the ASME 2015 13th International Conference on Nanochannels, Microchannels, and Minichannels. American Society of Mechanical Engineers. 2015. pp. V001T09A047–V001T09A047.

Каляев Игорь Анатольевич — д-р техн. наук, профессор, академик РАН, главный научный сотрудник научно-исследовательского института многопроцессорных вычислительных систем имени академика А.В. Каляева, Южный федеральный университет (НИИ МВС ЮФУ). Область научных интересов: многопроцессорные вычислительные и информационно-управляющие системы. Число научных публикаций — 330. kaliaev@mvs.sfedu.ru; ул. Чехова, 2, ГСП-28, Таганрог, Ростовская область, 347928; р.т.: +7(8634)360-757, Факс: +7(8634)360-376.

Дордопуло Алексей Игоревич — к-т техн. наук, начальник отдела математического и алгоритмического обеспечения, ООО «Научно-исследовательский центр супер-ЭВМ и нейрокомпьютеров» (ООО «НИЦ СЭ и НК»). Область научных интересов: языки программирования высокого уровня, программное обеспечение многопроцессорных вычислительных систем. Число научных публикаций — 153. scorpio@mvs.tsure.ru, <http://www.superevm.ru>; пер. Итальянский, 106, Таганрог, 347900; р.т.: +7(8634)612-111, Факс: +7(8634)477-415.

Левин Илья Израилевич — д-р техн. наук, профессор, директор, ООО «Научно-исследовательский центр супер-ЭВМ и нейрокомпьютеров» (ООО «НИЦ СЭ и НК»). Область научных интересов: высокопроизводительные многопроцессорные вычислительные системы, языки программирования высокого уровня, программное обеспечение. Число научных публикаций — 313. levin@superevm.ru, <http://www.superevm.ru>; пер. Итальянский, 106, Таганрог, 347900; р.т.: +7(8634)612-111, Факс: +7(8634)477-415.

Федоров Александр Михайлович — главный теплотехник, ООО «Научно-исследовательский центр супер-ЭВМ и нейрокомпьютеров» (ООО «НИЦ СЭ и НК»). Область научных интересов: системы охлаждения высокопроизводительных многопроцессорных вычислительных систем. Число научных публикаций — 3. ss24@mail.ru; пер. Итальянский, 106, Таганрог, 347900; р.т.: +7(8634)612-111, Факс: +7(8634)477-415.

Поддержка исследований. Работа выполнена при частичной финансовой поддержке Министерства образования и науки РФ по Соглашению о предоставлении субсидии № 14.5578.21.0006 от 05.06.2014, уникальный идентификатор RFMEFI57814X0006.

I.A. KALYAEV, A.I. DORDOPULO, I.I. LEVIN, A.M. FEDOROV
**EVOLUTION OF DOMESTIC MULTICHIP RECONFIGURABLE
 COMPUTER SYSTEMS: FROM AIR TO LIQUID COOLING**

Kalyaev I.A., Dordopulo A.I., Levin I.I., Fedorov A.M. Evolution Domestic of Multichip Reconfigurable Computer Systems: from Air to Liquid Cooling.

Abstract. In the paper we consider the evolution of design technologies of reconfigurable computer systems based on FPGAs of various families. Five FPGA-based generations of reconfigurable computer systems with high placement density, from Xilinx Virtex-E to modern Virtex UltraScale, are described. We show results of design of high real performance energy-efficient reconfigurable computer systems. The main contribution is a liquid cooling system designed for Virtex UltraScale FPGAs. It provides independent circulation of the cooling liquid in the 19" 3U computational module for cooling of 96-128 FPGA chips that generate 9.6-12.8 kWatt of heat in total. The distinctive features of the designed immersion liquid cooling system are high cooling efficiency with power reserve for the designed perspective FPGA families, resistance to leaks and their consequences, and compatibility with traditional water cooling systems based on industrial chillers.

Keywords: reconfigurable computer systems, FPGA, immersion liquid cooling system, industrial chillers.

Kalyaev Igor Anatolievich — Ph.D., Dr. Sci., professor, academician of RAS, chief researcher of Scientific Research Institute of Multiprocessor Computer Systems, Southern Federal University (SFedU). Research interests: multiprocessor computer and data control systems. The number of publications — 330. kaliaev@mvs.sfedu.ru; 2, Chehov st., GSP-284, Taganrog, 347928, Russia; office phone: +7(8634)360-757, Fax: +7(8634)360-376.

Dordopulo Alexey Igorevich — Ph.D., head of the mathematical and algorithmic support department, "Scientific Research Centre of Supercomputers and Neurocomputers" Co Ltd ("SRC SC & NC" Co Ltd). Research interests: high level programming languages, high-performance multiprocessor computer systems software. The number of publications — 153. scorpio@mvs.tsure.ru, <http://www.superevm.ru>; 106, Italyansky lane, Taganrog, 347900, Russia; office phone: +7(8634)612-111, Fax: +7(8634)477-415.

Levin Ilya Izrailevich — Ph.D., Dr. Sci., professor, director, "Scientific Research Centre of Supercomputers and Neurocomputers" Co Ltd ("SRC SC & NC" Co Ltd). Research interests: high-performance multiprocessor computer systems, high level programming languages, software. The number of publications — 313. levin@superevm.ru, <http://www.superevm.ru>; 106, Italyansky lane, Taganrog, 347900, Russia; office phone: +7(8634)612-111, Fax: +7(8634)477-415.

Fedorov Alexander Michailovich — chief heat engineer, "Scientific Research Centre of Supercomputers and Neurocomputers" Co Ltd ("SRC SC & NC" Co Ltd). Research interests: cooling systems for multiprocessor computer systems. The number of publications — 3. ss24@mail.ru; 106, Italyansky lane, Taganrog, 347900, Russia; office phone: +7(8634)612-111, Fax: +7(8634)477-415.

Acknowledgements. This paper was financially supported in part by the Ministry of Education and Science of the Russian Federation under Grant № 14.578.21.0006 from 05.06.2014, ID RFMEFI57814X0006.

References

1. Kalyaev I.A., Levin I.I., Semernikov E.A., Shmoilov V.I. Reconfigurable multipipeline computing structures. New York: Nova Science Publishers. USA. 2012. 330 p.

2. Carter W.S. The evolution of programmable logic. VLSI Circuits. Digest of Technical Papers. 1991 Symposium on IEEE. 1991. pp. 43–46.
3. Wilson R. In the Beginning. Available at: https://www.altera.com/solutions/technology/system-design/articles/_2013/in-the-beginning.html. Available at: https://www.altera.com/solutions/technology/system-design/articles/_2013/in-the-beginning.html (accessed 04.04.2016).
4. Chartrand L. Advanced Digital Systems: Experiments and Concepts with CPLDs. Cengage Learning. 2005. 268 p.
5. Power Methodology Guide. UG786 (v14.5). 2013. 54 p. Available at: http://www.xilinx.com/support/documentation/sw_manuals/xilinx13_1/ug786_PowerMethodology.pdf (accessed 04.04.2016).
6. User guide Microsemi. 158 p. Available at: http://www.actel.ru/files/uploads/Novosti/Download/Libero_UG_rus.pdf (accessed 04.04.2016).
7. Barkalov A., Titarenko L. Evolution of programmable logic. Logic Synthesis for FSM-Based Control Units. Springer Berlin Heidelberg. 2009. LNEE 53. pp. 53–75.
8. Hutton M. Architecture and CAD for FPGAs. Proceedings of the 17th Symposium on Integrated Circuits and Systems Design (SBCCI2004). 2004. pp. 3.
9. Kuon I., Tessier R., Rose J. FPGA Architecture: Survey and Challenges. Foundations and Trends in Electronic Design Automation. 2008. vol. 2. no. 2. pp. 135–253.
10. Levin I.I., Dordopulo A.I., Doronchenko Y.I., Raskladkin M.K. Reconfigurable computer system on the base of Virtex UltraScale FPGAs with liquid cooling. Proceedings of international scientific conference “Parallel computer technologies” (PaCT’2016). 2016. pp. 221–230.
11. Kalyaev I.A., Levin I.I., Dordopulo A.I., Slasten L.M. Reconfigurable Computer Systems Based on Virtex-6 and Virtex-7 FPGAs. IFAC Proceedings Volumes. Programmable Devices and Embedded Systems. 2013. vol. 12. no. 1. pp. 210–214.
12. Li L., Zheng W., Wang X., Wang X. Data center power minimization with placement optimization of liquid-cooled servers and free air cooling. *Sustainable Computing: Informatics and Systems*. 2016. vol. 11. no. 1. pp. 3–15.
13. Sun L. et al. Corrosion investigation of the inlet section of REAC pipes in the refinery. *Engineering Failure Analysis*. 2016. vol. 66. no. 1. pp. 468–478.
14. Abramov S.M. et al. [Supercomputers series 4 family SKIF: summit supercomputer technologies]. *Vestnik Nizhegorodskogo universiteta im. N. N. Lobachevsky – Vestnik of Lobachevsky University of Nizhni Novgorod*. 2009. vol. 5. pp. 200–210. (In Russ.).
15. Lyovshin I., Orlov D., Syomin A., Chelyshev A.J. [Not a revolution, but a coup]. *Superkomp'yutery – Supercomputers*. 2012. vol. 2(10). pp. 6–21. Available at: http://admin.clus-ter.sfu-kras.ru/data/uploads/2015/02/12/Supercom-puters_10-2012.pdf. (accessed 04.04.2016). (In Russ.).
16. Novoe superkomp'yuternoe reshenie «T-Platformy» A-Class [A new supercomputing solution "T-Platforms" A-Class]. Press-reliz ot 25.06.2014. https://www.pcweek.ru/infrastructure/news-company/detail_print.php?ID=164557&print=Y. (accessed 04.04.2016). (In Russ.).
17. Shah J.M., Eiland R., Siddarth, A., Agonafer D. Effects of mineral oil immersion cooling on IT equipment reliability and reliability enhancements to data center operations. Proceedings of the 15th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm 2016). 2016. pp. 316–325.
18. Gess J.L., Bhavnani S.H., Johnson R.W. Experimental Investigation of a Direct Liquid Immersion Cooled Prototype for High Performance Electronic Systems IEEE Transactions on Components, Packaging and Manufacturing Technology. 2015. vol. 5. no. 10. pp. 1451–1464.
19. Gess J., Dreher T., Bhavnani S., Johnson W. Effect of flow guide integration on the thermal performance of high performance liquid cooled immersion server modules. ASME 2015 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems collocated with the ASME 2015 13th International Conference on Nanochannels, Microchannels, and Minichannels. American Society of Mechanical Engineers. 2015. pp. V001T09A047–V001T09A047.