

## **ВЫСОКОПРОИЗВОДИТЕЛЬНАЯ РЕКОНФИГУРИРУЕМАЯ ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА НА ОСНОВЕ ПЛИС XILINX ULTRASCALE+**

*И. И. Левин, Ю. И. Доронченко, А. И. Дордопуло, М. Г. Левина*

ООО «Научно-исследовательский центр супер-ЭВМ и нейрокомпьютеров», Таганрог

В статье рассматривается перспективная реконфигурируемая компьютерная система на базе современных ПЛИС с погружной подсистемой охлаждения электронных компонентов. Система состоит из автономных вычислительных блоков. Мы представляем архитектуру базового компонента системы, такого как вычислительный модуль. Благодаря коммутационной структуре между вычислительными модулями и блоками, можно решать широкий класс ресурсоемких задач с многократными информационными обменами. Мы представляем новые подходы к питанию ПЛИС, которые потребляют до 180 Вт каждая. Благодаря уникальной конструкции вычислительного блока, можно достичь вычислительной плотности до 128 ПЛИС в конструкции 3U. Мы демонстрируем преимущества подсистемы погружного охлаждения. Кроме того, мы представляем описания технических решений для улучшения подсистемы охлаждения. Благодаря такому усовершенствованию, можно отводить до 20 кВт тепла в целом от одного вычислительного блока. Кроме того, мы рассматриваем оригинальный комплекс средств программирования на языке высокого уровня для многокристалльных PBC.

*Ключевые слова:* реконфигурируемые вычислительные системы, производительность вычислений, ПЛИС, системы погружного охлаждения, вычислительная плотность, средства программирования, язык высокого уровня.

### **1. Введение**

Научно-технический прогресс требует неуклонного повышения производительности суперкомпьютеров для решения вычислительно трудоемких задач. В то же время рост производительности традиционных многопроцессорных систем замедляется, повышение тактовых частот процессоров остановилось. Для реконфигурируемых вычислительных систем (PBC), построенных на основе программируемых логических интегральных схем (ПЛИС), удается поддерживать рост производительности вычислений в 2,5–3 раза с выходом каждого нового поколения ПЛИС (приблизительно раз в два года).

В отличие от вычислительных систем с традиционной архитектурой концепция PBC позволяет создавать вычислительные комплексы, обеспечивающие высокую реальную производительность вычислений для решения широкого круга задач различных проблемных областей за счет нивелирования противоречий между реализацией проблемно ориентированного устройства (обладающего максимальной производительностью) и необходимостью его универсального применения [1]. PBC достигают существенно более высокой реальной производительности и энергоэффективности за счет адаптации своей архитектуры к структуре решаемой задачи: пользователь может в PBC для каждой решаемой задачи создать специализированное вычислительное устройство, реализующее вычислительные операции информационного графа задачи с наименьшими задержками [2].

Современный уровень интеграции ПЛИС и достигаемые значения рабочей тактовой частоты реализаций алгоритмов на ПЛИС очень высоки – сотни миллионов эквивалентных вентилях способны работать на частотах до 700 МГц. Использование этих возможностей и близкое к максимальному заполнение ресурсов микросхемы делают невозможным отвод тепла от корпуса ПЛИС с использованием воздушных систем охлаждения при ограниченных габаритах изделия. Одной из труднейших проблем построения эффективных вычислителей является обеспечение приемлемых тепловых режимов работы как отдельных компонентов с высоким тепловыделением, так и функционально законченного вычислительного устройства в целом. Важным является выбор системы охлаждения. Проведенные в Научно-исследовательском центре супер-ЭВМ и нейрокомпьютеров (НИЦ СЭ и НК) исследования показали, что из различных вариантов жидкостного охлаждения наиболее эффективной является иммерсионная технология охлаждения [3].

## 2. РВБ «Арктур» на основе ПЛИС Xilinx UltraScale+

Каждый выпуск нового семейства ПЛИС сопровождается увеличением аппаратного вычислительного ресурса, ростом тактовой частоты вследствие уменьшения проектных норм производства, изменениями в технологии передачи данных, что подталкивает разработчиков реконфигурируемых суперкомпьютеров к поиску новых технических решений, позволяющих использовать эти новые преимущества.

Для построения РВБ нового поколения на основе ПЛИС Xilinx семейства UltraScale+ потребовался ряд изменений как в системе охлаждения, так и в конструкции изделия и его составных частей.

Корпус ПЛИС семейства UltraScale+ на 5 мм больше корпуса ПЛИС семейства UltraScale, используемых в предыдущих РВБ, например РВБ «Нектар» [4]. В этой связи ширина печатной платы вычислительного модуля не позволяет разместить восемь вычислительных ПЛИС и контроллер. Для сохранения вычислительной плотности изделия из состава вычислительного модуля исключена ПЛИС контроллера. Функции контроллера (обеспечение процесса загрузки конфигураций ПЛИС, реализация интерфейса доступа к вычислительным ресурсам и мониторинг параметров) возложены на вычислительную ПЛИС, что потребует не более 5 % ее ресурса. Эскиз компоновки вычислительного модуля показан на рис. 1.

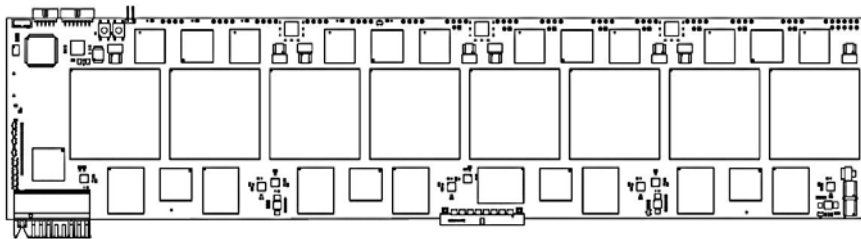


Рис. 1. Эскиз компоновки вычислительного модуля «Арктур»

Вычислительный ресурс новых микросхем UltraScale+ примерно в два раза превышает ресурс используемых ПЛИС предыдущего поколения. Следовательно, с учетом роста энергоэффективности потребление одной ПЛИС может вырасти до 170–200 Вт, что потребовало кардинальной переработки подсистемы вторичного электропитания вычислительного модуля.

В РВБ «Арктур» система питания спроектирована, исходя из расчетного потребления ПЛИС не менее 200 Вт. При таком потреблении вычислительный модуль с восемью ПЛИС может потреблять до 2 кВт, что делает обеспечение его питания от 12 В практически невозможным, так как ток в этом случае будет до 170 А. В этой связи в качестве основной шины питания выбрано напряжение 48 В, что позволяет снизить ток на плате вычислительного модуля в 4 раза до приемлемых значений 32 – 42 А. Нестабилизированные гальванические преобразователи 400/48 В могут быть реализованы распределенным образом, в том числе непосредственно на кросс-плате. Соответственно на вычислительном модуле выполняются преобразования из 48 В во все требуемые напряжения со стабилизацией.

Рост потребляемой мощности по ядру ПЛИС приводит к необходимости модернизации вторичного питания. При одноточечном размещении источника питания вследствие омического сопротивления возникает градиент напряжения, при этом падение при большом токе потребления становится существенным – до нескольких десятков милливольт. Корректировка топологии не приводит к существенному улучшению, поэтому необходимо использовать несколько источников напряжения ядра, что позволит сократить длины проводников и снизить падение напряжения в несколько раз.

Применение ПЛИС фирмы Xilinx семейства UltraScale+, выполненных по 16-нм технологии, позволит до трех раз повысить производительность вычислений за счет роста тактовой частоты и степени интеграции ПЛИС, причем без увеличения объема вычислительной системы. По заявлениям производителя в связи с новыми технологическими нормами производства снижение потребляемой мощности по сравнению с UltraScale может составить до 50 %. Однако при этом, несмотря на некоторый имеющийся резерв мощности разработанной системы иммерсионного жидкостного охлаждения, можно ожидать очередное стремление рабочих температур ПЛИС к своим критическим значениям.

РВБ «Арктур» обладает поддержкой широких возможностей информационного обмена. На рис. 2 показана структурная схема блока. МЗУ обеспечивает функциональную законченность изделия и выполнен в виде отдельной платы. Основными узлами МЗУ являются процессор Intel Core i5–6300U и ПЛИС семейства Kintex Ultrascale. ПЛИС выполняет функцию контроллера и обеспечивает взаимодействие между процессором и ПЛИС в первом вычислительном модуле. Связь между МЗУ и вычислительным модулем осуществляется через кросс-плату. МЗУ также обеспечивает аппаратный контроль датчиков уровня жидкости, датчика протока и параметров насоса с помощью микроконтроллера. На плату индикации выведены кнопки управления, индикаторы, а также интерфейсы для подключения клавиатуры, мыши, дисплея, Ethernet и оптический интерфейс.

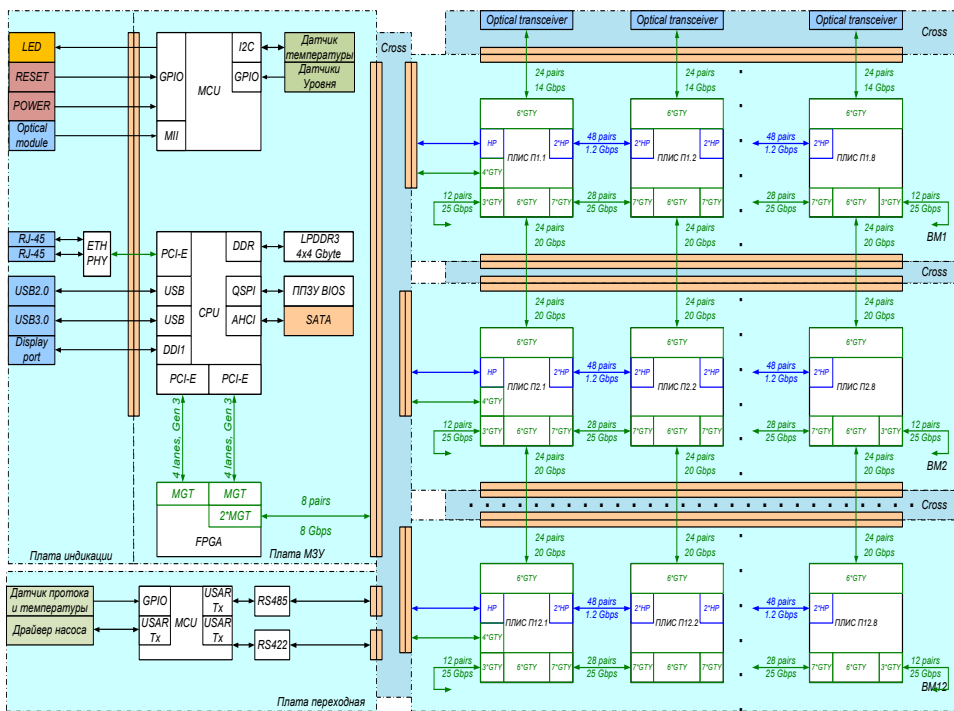


Рис. 2. Структурная схема РВБ «Арктур»

На рис. 3 показана структура информационных связей. В качестве основных связей между ПЛИС предполагается использование дифференциальных линий с подключенными к ним мульти-гигабитными трансиверами (MGT), вспомогательными связями являются дифференциальные линии, подключенные к НР – банкам.

На вычислительном модуле ПЛИС связаны горизонтальными и вертикальными линиями связей. Горизонтальные связи между ПЛИС представлены 28 и 12-ю дифференциальными парами

в каждом направлении со скоростью передачи данных до 25 Гбит/с по одной паре, что обеспечивает пропускную способность 700 и 300 Гбит/с независимо в каждом направлении. К вертикальным связям относятся 192 дифференциальные пары в каждом направлении с допустимой скоростью передачи данных до 20 Гбит/с по одной паре, что обеспечивает пропускную способность до 3840 Гбит/с независимо в каждом направлении.

Дополнительные линии связи представлены горизонтальной сетью от П1.1 до П1.8 на каждой плате (не показана на рис. 2) и содержит 48 дифференциальных линий с максимальной скоростью до 1,2 Гбит/с, каждая из которых может работать в любом направлении.

Кроме того, между платами существуют связи, предназначенные для служебных целей, таких как загрузка конфигурации ПЛИС, передача команд и данных от МЗУ к любому вычислительному модулю. Пропускная способность данного канала – 64 Гбит/с для обоих направлений независимо.

В настоящее время рассматривается два варианта производства РВБ «Арктур» – с использованием ПЛИС XCVU9P для вычислительно трудоемких задач и с использованием ПЛИС XCVU37P для задач с особыми требованиями к объему памяти. ПЛИС XCVU37P является одной из четырех в особой линейке ПЛИС UltraScale+, обладающей новым аппаратным ресурсом – встроенным модулем DDR памяти HBM2. Память HBM2 обладает объемом до 8 Гбайт и может обеспечивать многоканальный доступ 16×64 бит.

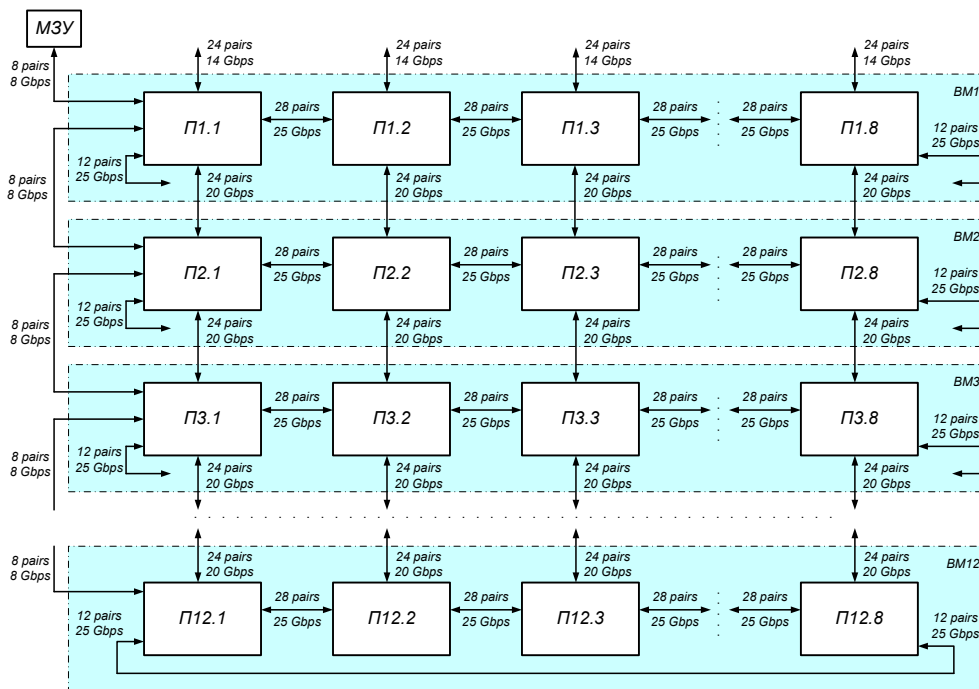


Рис. 3. Структура информационных связей РВБ «Арктур»

Все технические решения по созданию вычислительного модуля отмакетированы.

При построении вычислительных комплексов информационное взаимодействие между РВБ осуществляется через оптические каналы (рис. 4). Оптические приемопередатчики установлены на отдельной кросс-плате и подключаются к первому вычислительному модулю через разъемное соединение. Скорость приема и передачи данных в каждом из оптических каналов достигает 14 Гбит/с. Полное количество оптических каналов составляет 384 и 192 канала на прием и 192 канала на передачу, это позволяет получить пропускную способность до 2688 Гбит/с независимо в каждом направлении.

ПЛИС вычислительного блока можно условно представить в виде строк и столбцов. Тогда строками будут являться ПЛИС, связанные между собой в пределах каждого вычислительного модуля РВБ, а столбцами – ПЛИС, связанные соответствующими межмодульными соединениями. Таким образом, обеспечивается информационная связь между соответствующими столбцами всех РВБ, объединенных оптическими каналами. Первый и последний РВБ в вычислительной стойке комплекса

могут быть как соединены между собой, образуя кольцевое соединение по столбцам, так и подключаться к РВБ соседних вычислительных стоек комплекса, наращивая объединенный в единое пространство ресурс. Для обеспечения изоморфности топологии информационных связей во всех РВБ первый канал столбца предыдущего РВБ всегда подключается ко второму каналу столбца последующего РВБ.

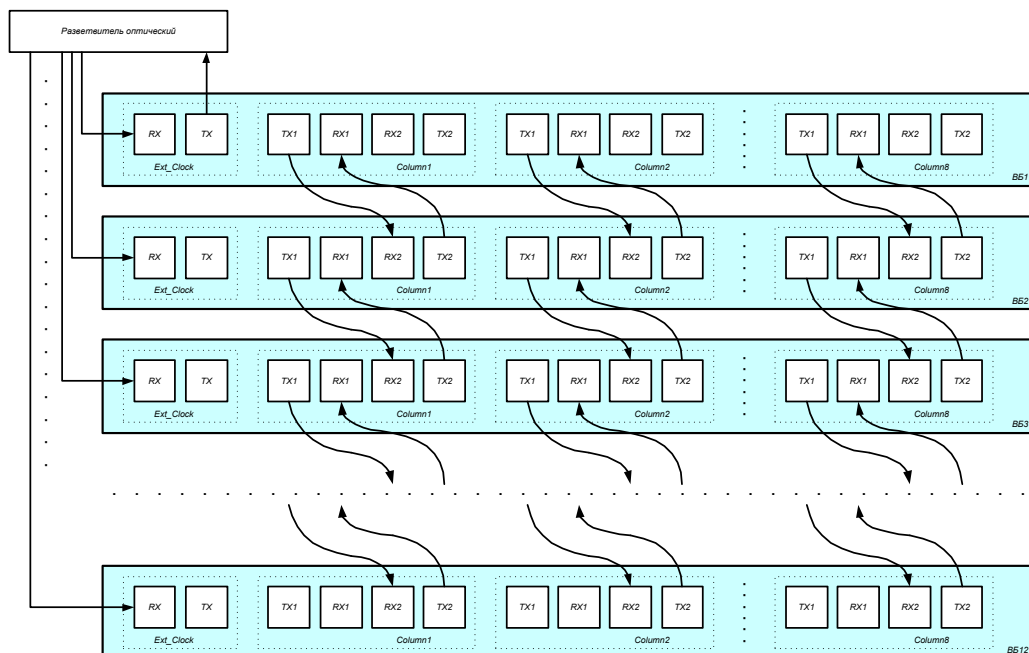


Рис. 4. Структура информационных связей между РВБ

Выполнена оценка производительности РВБ «Арктур» при использовании ПЛИС UltraScale+ XCVU9P на 12-ти вычислительных модулях.

Для оценки производительности использовалась конвейерная вычислительная структура, реализующая функцию фильтра с конечной импульсной характеристикой (КИХ) с симметричными коэффициентами. Вычислительный конвейер представляет собой последовательно соединенные каскады фильтрации. При реализации данной тестовой задачи цифровой обработки данных может быть достигнута производительность 240 Tflops (single precision). Таким образом, построение РВС из 16 РВБ «Арктур» обеспечит производительность при решении данной задачи 3,8 Pflops (single precision).

Для оценки производительности также использовалась задача решения СЛАУ с помощью конвейерной вычислительной структуры, реализующей функцию LU – разложения, поскольку данная функция является алгоритмической основой теста производительности компьютеров LINPACK. При реализации данной тестовой задачи может быть достигнута производительность 31,5 Tflops (double precision). При построении РВС на основе 16 РВБ «Арктур» при решении задачи LU – разложения производительность составит 500 Tflops (double precision).

### 3. Система охлаждения РВБ «Арктур»

Особенностью выпускаемых в НИЦ СЭ и НК реконфигурируемых вычислительных блоков (РВБ) для построения РВС является высокая плотность компоновки вычислительных компонентов. Каждый РВБ состоит из 12–16 базовых вычислительных элементов – вычислительных модулей. На одном вычислительном модуле устанавливается до восьми ПЛИС, каждая из которых требует обеспечения электропитанием с помощью микросхем модулей питания, которые также являются активными источниками тепла. Данное обстоятельство существенно увеличивает число тепловыделяющих элементов по сравнению с микропроцессорными модулями, состоящими из одного–двух процессоров.

В ступенчатом переносе тепла в иммерсионной (открытой) системе охлаждения принимают участие несколько компонентов. Эффективность системы жидкостного охлаждения обуславливается перспективными техническими решениями при реализации каждого компонента: применяемого хладагента, конструкции используемых радиаторов ПЛИС, насосного оборудования, теплообменников. Хладагент должен обладать наилучшей электрической прочностью, высокой теплопроводностью, максимально возможной теплоемкостью при низкой вязкости. Радиатор должен обеспечивать максимальную поверхность теплосъема, возможность организации циркуляции хладагента через радиатор, турбулентность потока хладагента в радиаторе, технологичность изготовления. Применяемый термоинтерфейс должен не деградировать и не вымываться хладагентом, иметь стабильно высокий коэффициент теплопроводности. Теплообменник должен обеспечивать высокий коэффициент теплопередачи между основным и вторичным контуром охлаждения. Для обеспечения циркуляции хладагента в объеме РВБ используется насос, который должен обладать необходимой производительностью.

Проведенные в НИЦ СЭ и НК работы позволили получить и апробировать технические решения по обеспечению эффективного теплопереноса каждым из компонентов системы охлаждения. В настоящее время в НИЦ СЭ и НК серийно производится РВБ «Неккар» (рис. 5). Рассмотрим основные составные части и особенности технологии производства данного изделия для общности понимания перспектив развития направления РВС с иммерсионной системой охлаждения.

Корпус РВБ «Неккар» [5] в конструктиве 3U 19" состоит из вычислительной секции и теплообменной секции. Вычислительная секция представляет собой герметичный контейнер с хладагентом, в которой размещаются вычислительные модули (12 штук), модули питания, плата управления, кросс-платы. Сверху вычислительная секция закрыта крышкой.

К вычислительной секции примыкает теплообменная секция, основными компонентами которой являются насос и теплообменник. Насос обеспечивает циркуляцию хладагента в РВБ по следующему замкнутому контуру: из вычислительной секции нагретый хладагент поступает в пластинчатый теплообменник, там охлаждается, затем вновь поступает в вычислительную секцию, охлаждает там нагретые электронные компоненты, отчего нагревается сам и вновь поступает в теплообменник. Теплообменник подключен к вторичному контуру охлаждения через штуцеры и предназначен для охлаждения хладагента, как правило, с помощью воды. Также в теплообменной секции располагается расширительный бачок для компенсации объемного расширения хладагента.

Каждый вычислительный модуль РВБ «Неккар», представленный на рис. 6, содержит восемь ПЛИС семейства Kintex UltraScale XCKU095T, а также контроллер вычислительного модуля, выполненный на ПЛИС Xilinx Kintex – 7. Для охлаждения ПЛИС используются радиаторы малой высоты оригинальной конструкции, установленные по особой технологии с помощью специально разработанного термоинтерфейса.



Рис. 5. РВБ «Неккар»



Рис. 6. Вычислительный модуль

Специалистами НИЦ СЭ и НК были проведены теплотехнические расчеты и предложена принципиально новая конструкция радиатора с оригинальными ламелями, которые позволяют осуществлять местное закручивание потока хладагента, переводя его из ламинарного в условно турбулентный поток.

Для осуществления загрузки ПЛИС и управления процессом вычислений в вычислительных модулях РВБ разработана оригинальная материнская плата на основе процессора Intel Skylake® Core I5–6300U – модуль загрузки и управления (МЗУ). Для МЗУ разработана оригинальная базовая система ввода-вывода (BIOS), позволяющая использовать все возможности системы на кристалле Intel Skylake и внешнего периферийного оборудования. Для теплоотвода от процессора используется радиатор оригинальной конструкции.

РВБ «Неккар» использует электропитание 400 В постоянного тока. Для электропитания вычислительных модулей разработан модуль питания, обеспечивающий четыре канала преобразования DC/DC 400/12 В мощностью до 4 кВт для электропитания четырех ПВМ. Максимальная потребляемая мощность РВБ – 12 кВт.

Для охлаждения электронных компонентов РВБ был разработан диэлектрический хладагент – масло маловязкое, диэлектрик МД–4,4 Gazpromneft для охлаждения электронных компонентов ЭВМ. Хладагент обладает высокой электрической прочностью и теплопроводностью, а также максимально возможной теплоемкостью при низкой вязкости.

Для эксплуатации РВБ и РВС на их основе требуется подключение с помощью подающих и обратных коллекторов через штуцеры (или балансировочные клапаны) и гибкие трубопроводы к вторичному контуру охлаждения, а также к источнику энергоснабжения и сетевому концентратору.

Уникальная вычислительная плотность РВБ «Неккар» позволяет достигать производительности при решении тестовой задачи цифровой обработки данных до 60 Tflops (single precision). Таким образом, построение РВС из 16 РВБ обеспечит производительность 1 Pflops. Производительность РВС на основе 16 РВБ «Неккар» при решении задачи LU – разложения (тест LinPack) составляет 160 Tflops (double precision).

Комплекс технических решений разработанной погружной системы жидкостного охлаждения позволил обеспечить в рабочем режиме температуру хладагента не более 33 °С при потребляемой мощности до 90 Вт для каждой ПЛИС. При этом максимальная температура ПЛИС не превышает 57–60 °С, что свидетельствует о резерве мощности, позволяющем обеспечить эффективное охлаждение для проектируемых РВС на основе перспективного семейства ПЛИС Xilinx UltraScale+.

При проектировании РВБ «Арктур» в рамках модернизации системы охлаждения решены следующие задачи:

- разработка оптимальной конструкции радиаторов за счет увеличения эффективной площади поверхности теплообмена между радиатором ПЛИС и хладагентом;
- увеличение производительности насоса подачи хладагента и повышение надежности системы жидкостного охлаждения посредством применения погружных насосов;
- повышение производительности теплообменника.

Для отвода большего количества тепла от ПЛИС необходимо увеличивать площадь теплосъема у радиаторов, монтируемых на ПЛИС, исключив при этом увеличение высоты радиатора. В рамках проведенных исследований рассматривалось множество вариантов конструкции радиаторов, часть из которых были эффективными, но очень дорогостоящими в технологии изготовления. Для радиатора удалось увеличить площадь теплосъема на 17 и 43 % для различных исполнений радиаторов, а если учесть, что с увеличением габаритов ПЛИС можно увеличить и габариты (длину и ширину) радиатора, площадь теплосъема возрастает более чем в два раза.

Циркуляцию хладагента в системе жидкостного охлаждения обеспечивает насос с приводом от электродвигателя постоянного тока. Электродвигатель является единственным элементом в РВБ с жидкостным охлаждением, требующим отдельного охлаждения, для чего в вычислительных стойках предусматривается принудительная циркуляция холодного воздуха. Предложено охлаждать электродвигатель хладагентом, циркулирующим в РВБ. Для этих целей разработаны и успешно тестируются два насоса, оснащенных погружными электродвигателями мощностью 0,5 и 1 кВт. Использование этого новшества позволяет снизить требования к вычислительным стойкам и поме-

щениям, где эксплуатируются РВС с иммерсионным охлаждением. Кроме того, значительно упрощается техническое обслуживание, а также повышается надежность за счет уменьшения количества открытых гидравлических соединений при эксплуатации.

Увеличение выделяемой тепловой мощности ПЛИС влечет увеличение мощности теплообменника не менее чем в 1,5 раза, что потребовало нарастить количество рабочих пластин с 24 до 36 штук, что неизменно повлияло на его высоту. При существующей схеме расположения гидравлических элементов сохранение габарита изделия 3U стало невозможным. Проблему могло решить расположение теплообменника на боковой поверхности, но в этом случае теплообменник переставал выполнять свою основную функцию – отвод тепла. Дело в том, что у современных теплообменников штуцеры для подключения теплоносителей одной среды, как правило, расположены в одной плоскости, параллельной боковой поверхности, а не по диагонали. Поэтому при установке теплообменника на боковой поверхности невозможно удалить воздух из теплообменника для среды, у которой штуцеры расположены внизу.

Для конструкции РВБ идеальным виделось диагональное расположение штуцеров сред на противоположных сторонах. Это позволило бы разделить подключение хладагента и воды на разных сторонах теплообменника, сократить количество гидравлической арматуры и минимизировать последствия нарушения герметичности подключения изделия. Кроме того, необходимо, чтобы теплообменник эффективно работал с более вязкой средой, чем вода, – хладагентом МД – 4,4 Gazpromneft. Это требовало применения пластин специальной конфигурации. В сотрудничестве с немецким производителем пластинчатых теплообменников фирмой «Funke» данная проблема была решена. За основу были взяты теплообменники серии TPL, которые разрабатывались специально для охлаждения гидравлических и моторных масел водой. Ширина канала теплообменников TPL до 80 % больше, чем у классических паяных водо-водяных теплообменников. Специальными встроенными турбулизаторами и эффективными диагональными направляющими создается максимальная эффективность. Теплообменники TPL обладают повышенными требованиями к технологии пайки, исключая возможность перемешивания двух сред.

Претерпела изменения и система компенсации объемного расширения хладагента вследствие изменения его температуры. В РВБ «Неккар» для этих целей использовался расширительный бачок, сообщающийся с атмосферой, где изменялся уровень хладагента в зависимости от его объемного расширения. Такая система имела ряд недостатков: сложность конструкции расширительного бачка и наличие соединительных гидравлических узлов требовало тщательной герметизации изделия и регулярного контроля уровня хладагента. Кроме того, хладагент, находящийся в расширительном бачке, контактировал с атмосферным воздухом, что приводило к его дополнительному окислению и неизбежному сокращению срока службы.

В предложенном техническом решении роль компенсатора объемного расширения хладагента играют гофрированные резиновые рукава. При этом полностью исключен контакт хладагента с атмосферным воздухом и минимизированы места потенциальных протечек.

Эскиз конструкции РВБ «Арктур» с модернизированной погружной системой охлаждения представлен на рис. 7.

Все технические решения по модернизации системы охлаждения для РВБ «Арктур» проверены на различных стендах и макетах и показали возможность эффективного функционирования изделия.



Рис. 7. Эскиз конструкции РВБ «Арктур» с модернизированной системой охлаждения



#### 4. Комплекс средств программирования реконфигурируемых вычислительных систем нового поколения

Оригинальный комплекс средств программирования PBC [2, 6] в течение 17 лет разрабатывается коллективом Научно-исследовательского центра супер-ЭВМ и нейрокомпьютеров (г. Таганрог), ведущего российского разработчика и производителя высокопроизводительных PBC. Отличительной характеристикой созданного комплекса является программирование многокристалльных PBC на языке программирования высокого уровня с автоматической синхронизацией размещенных в каждой ПЛИС фрагментов задачи.

С 2017 года в НИЦ супер-ЭВМ и нейрокомпьютеров разрабатывается новый комплекс средств разработки прикладных программ для трансляции исходной программы на языке C в программу на языке COLAMO с последующей адаптацией под архитектуру PBC и синтезом конфигурационных файлов ПЛИС.

В отличие от автоматических распараллеливающих компиляторов [7] (Polaris, SUIF, VAST/Parallel, Intel/OpenMP и др. для вычислительных систем с общей памятью и ParaWise, Parallel Studio, САПФОР [7, 8], OPC и ДВОР [9] и др. для систем с распределенной памятью), для создания параллельной программы в PBC используется альтернативный распараллеливанию последовательной программы подход. Для PBC задача исходно представляется в абсолютно параллельной форме в виде отражающего естественный параллелизм и конвейеризацию операций информационного графа [1, 2], который адаптируется к текущей конфигурации вычислительной системы с помощью методов масштабирования (индукции и редукции) производительности вычислительной структуры. При этом информационный граф может масштабироваться не только по данным, но и по итерациям, что сокращает число используемых каналов памяти (или ввода-вывода).

Структура комплекса и схема его функционирования представлены на рис. 8.

Комплекс средств программирования PBC содержит:

- конвертор «Ангел» для преобразования программы на языке программирования «C» в стандарте ISO/IEC 9899:1999 в абсолютно-параллельную программу на языке Colamo;
- конвертор «Русалка», преобразующий абсолютно-параллельную программу на языке Colamo в ресурсонезависимую (масштабируемую) программу на языке Colamo;
- программу «Прокруст», осуществляющую адаптацию параметров ресурсонезависимой программы на языке Colamo под архитектуру PBC;
- конвертор «Щелкунчик», осуществляющий редукцию производительности ресурсонезависимой программы при нехватке аппаратного ресурса;
- интегрированную среду разработки;
- транслятор языка программирования высокого уровня Colamo;
- синтезатор масштабируемых вычислительных структур Fire!Constructor;
- описания (паспорта) PBC и библиотеку элементов;
- средства диагностики и мониторинга;
- средства описания софт-архитектур и синтезатор Steam!Constructor;
- управляющую программу на языке C.

В отличие от большинства известных средств трансляции языка C в конфигурационные файлы ПЛИС (трансляторы C-to-RTL, MitrionC, CatapultC и др.) рассматриваемый комплекс средств программирования не синтезирует специализированный вычислитель на основе автоматной модели или процессорной парадигмы, а отображает последовательную программу в абсолютно параллельную форму, которую адаптирует под доступный аппаратный ресурс PBC или гибридной PBC с учетом информационных зависимостей данных.

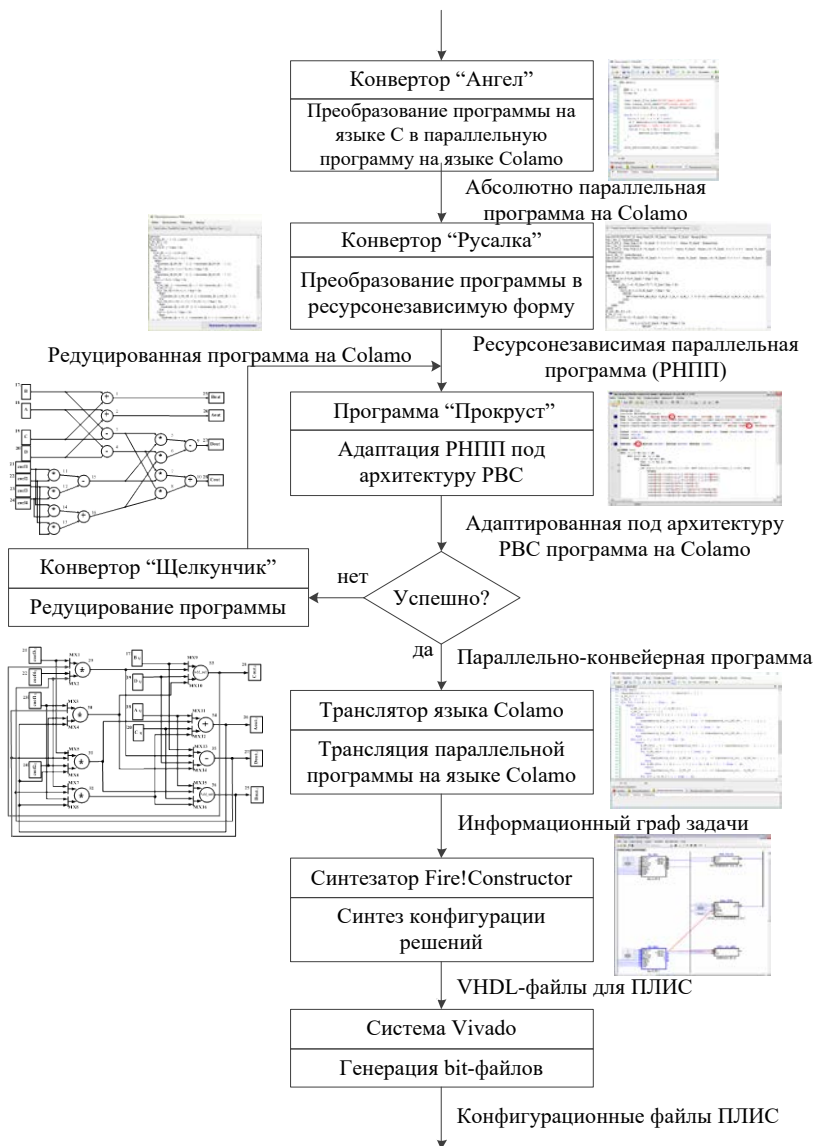


Рис. 8. Структура и схема функционирования комплекса средств программирования нового поколения

От наиболее близкого аналога Xilinx Vivado HLS (High-Level Synthesis) при схожей функциональности комплекс отличается поддержкой многокристальных решений с автоматической синхронизацией информационных и управляющих сигналов.

Входной программой является последовательная программа на языке программирования C в наиболее распространенном стандарте ISO/IEC 9899:1999 с GNU-расширением для компилятора gcc. Методика перевода последовательной программы на языке программирования C в программу на языке программирования COLAMO содержит 5 основных этапов.

Первый этап: перевод программы на языке программирования C в абсолютно-параллельную форму на языке программирования COLAMO. Результатом первого этапа является абсолютно параллельная программа в синтаксисе COLAMO, но требующая для своей реализации большое число каналов памяти.

Второй этап: анализ информационных зависимостей исходной программы, расщепление скалярных переменных и растягивание массивов по итерациям выполняется транслятором «Ангел». При сохранении информационных связей исходной последовательной программы обеспечивается выполнение правил единственной подстановки и однократного присваивания в параллельной программе

с помощью анализа информационных зависимостей переменных и массивов данных исходной программы. Если для адресации массива данных с явно выраженной информационной зависимостью используются не все переменные циклов, то те циклы, которые не используются для адресации элементов массива (потока) данных, являются для него итерационными (как правило, итерации действуют для самого нижнего вложенного цикла). На каждой итерации итерационного цикла будет осуществляться запись в одни и те же ячейки коммутационного массива, что приведет к нарушению правила однократного присваивания. Поэтому необходимо осуществить растягивание коммутационного массива по итерациям (исходный мемориальный массив остается без изменений).

Третий этап: переход в параллельно-конвейерную форму, разбиение массивов и циклов на векторную и потоковую составляющие выполняются процессором «Русалка».

Каждое измерение любого массива разбивается на 2 составляющие – векторную и потоковую с параллельным и последовательным типами доступа соответственно. Для этого используется следующая схема разбиения: длина потоковой составляющей задается равной  $R$ , а длина векторной определяется как  $(T + R - 1) / R$ , где  $T$  – исходная длина измерения массива). Векторная составляющая должна быть целочисленной, т.е. исходная длина измерения массива должна нацело делиться на  $R$ .

Четвертый этап: перевод в масштабируемую параллельно-конвейерную форму выполняется процессором «Прокруст». Если каналов памяти недостаточно для реализации базового подграфа, то необходимо выполнять редукцию программы [10]. Полученный коэффициент масштабирования  $K$  определяет параллелизм по каждому измерению массива. Полученный коэффициент  $K$  не всегда является итоговым, так как он характеризует максимальную степень масштабирования и на него нацело должен делиться исходный размер задачи  $N$ . Аналогичным образом определяются необходимый для реализации базового подграфа аппаратный ресурс (по числу LUT, DSP, Flip-Flop и других ресурсов ПЛИС) и коэффициент масштабирования по ресурсу  $Q$ .

Пятый этап: создание вычислительной структуры выполняется процессором «Прокруст».

Выполнение описанной методики позволяет синтезировать ресурснезависимую параллельную прикладную программу, которая адаптируется под доступный вычислительный ресурс изменением значений нескольких констант в автоматическом режиме без существенной корректировки исходного текста программы. В результате выполнения всех 5-ти этапов создается масштабируемая прикладная программа на языке программирования COLAMO, которая с помощью существующих транслятора COLAMO, синтезатора Fire! Constructor и других средств разработки преобразуется в конфигурационные файлы ПЛИС многокристальных PBC.

В настоящее время создана экспериментальная версия комплекса средств разработки пятого поколения, включающая все программы и конверторы, с помощью которой осуществлена успешная трансляция ряда последовательных программ на языке C в программы на языке программирования высокого уровня COLAMO и далее в конфигурационные файлы ПЛИС для различных предметных областей. Успешно транслированы следующие прикладные программы математической физики: решение систем линейных алгебраических уравнений (СЛАУ) методом Гаусса, решение СЛАУ методом Якоби, решения СЛАУ с помощью верхней треугольной и нижней треугольной матриц (LU-разложение). В области символьной обработки решены 5 прикладных задач, при этом программы, транслируемые с помощью комплекса средств нового поколения, показывают высокую эффективность: не менее 85 % от созданных прикладными программистами на языке COLAMO и не менее 70 % от прикладных решений, разработанных специалистами-схемотехниками. Время трансляции последовательных программ на языке C в программы на языке программирования высокого уровня COLAMO и в конфигурационные файлы ПЛИС (без синтеза загрузочных конфигурационных bit-файлов) не превышает 30 минут, основное время трансляции составляет работа среды проектирования Xilinx Vivado.

## 5. Заключение

ПЛИС как элементная база реконфигурируемых супер-ЭВМ обеспечивают устойчивый, близкий к линейному, рост производительности РВС, открывая новые перспективы по созданию суперкомпьютеров петафлопсной производительности. В отличие от элементной базы многопроцессорных вычислительных систем, ПЛИС поддерживают развитие РВС за счет сохранения темпов роста аппаратного ресурса и рабочей тактовой частоты. Построение РВС из 16 РВБ «Арктур» (одна вычислительная стойка) обеспечит пиковую производительность вычислений до 4 Pflops (single precision), позволив при этом за счет мощной системы информационных связей решать широкий класс задач, в том числе сильносвязанных.

Эффективность применения иммерсионной системы жидкостного охлаждения в перспективных системах обуславливается техническими решениями при реализации каждого из компонентов системы – хладагента, радиаторов ПЛИС, насосного оборудования, теплообменников. При переходе к новой элементной базе требования к этим компонентам усиливаются, необходима модернизация готовых и поиск новых решений. Для построения нового РВБ «Арктур» на основе ПЛИС Xilinx UltraScale+ в ООО «НИЦ СЭ и НК» проведены масштабные работы по доработке компонентов системы охлаждения, что позволит осуществить эффективный отвод тепла, оставив при этом конструктив изделия 3U 19” без изменений.

Продолжение активного развития технологии построения РВС открывает возможности для создания вычислительных систем эксафлопсной производительности. Одним из шагов в этом направлении является создание перспективного РВБ «Арктур» на основе ПЛИС Xilinx UltraScale+, серийное производство которого намечено ООО «НИЦ СЭ и НК» в 2021 г.

Для эффективного программирования РВС и гибридных реконфигурируемых вычислительных систем разработана методика трансляции последовательной программы на языке С в программу на языке программирования высокого уровня COLAMO и далее, с помощью разработанных средств трансляции, – в конфигурационные файлы ПЛИС. Создаваемый комплекс средств разработки прикладных программ позволит существенно сократить время программирования многокристалльных РВС с помощью трансляции исходной последовательной программы на языке программирования С в загрузочные конфигурационные файлы ПЛИС.

Совокупность разработанных аппаратно-программных средств на основе ПЛИС новых поколений открывает возможность создания высокопроизводительных вычислительных комплексов мирового уровня.

## Литература

1. Каляев А. В., Левин И. И. Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. М.: Янус-К, 2003.
2. Гузик В. Ф., Каляев И. А., Левин И. И. Реконфигурируемые вычислительные системы / Учебное пособие под общей редакцией Каляева. И. А. 2016. – Ростов-на-Дону: Изд-во ЮФУ, 2016.
3. Levin I. I., Dordopulo A. I., Fedorov A. M., Gulenok A. A. Reconfigurable computer based on Virtex UltraScale+ FPGAs with immersion cooling system // XI международная научная конференция "Параллельные вычислительные технологии (ПаВТ) 2017 (Parallel computational technologies (PCT'2017) agora.guru.ru/pavt)", Казань, Республика Татарстан, Россия, 3–7 апреля 2017 г. – Челябинск: Издательский центр ЮУрГУ, 2017. С. 27–41. ISBN 978-5-696-04880-2 URL: [https://link.springer.com/chapter/10.1007/978-3-319-67035-5\\_3](https://link.springer.com/chapter/10.1007/978-3-319-67035-5_3) (дата обращения: 03.10.2019).
4. Каляев И. А., Дордопуло А. И., Левин И. И., Федоров А. М. Развитие отечественных многокристалльных реконфигурируемых вычислительных систем: от воздушного к жидкостному охлаждению // Труды СПИИРАН. – Санкт-Петербург: Изд-во СПИИРАН ФГБУН Санкт-Петербургский институт информатики и автоматизации РАН, 2017. № 1(50). С. 5–31. DOI 10.15622/sp.50.1 <http://dx.doi.org/10.15622/sp.50>. (дата обращения: 03.10.2019).

5. Levin Ilya, Dordopulo Alexey, Fedorov Alexander, Doronchenko Yuriy. High-Performance Reconfigurable Computer Systems with Immersion Cooling // Parallel Computational Technologies (12th International Conference, PCT 2018, Rostov-on-Don, Russia, April 2–6, 2018. P. 62–76. [https://link.springer.com/content/pdf/10.1007%2F978-3-319-99673-8\\_5.pdf](https://link.springer.com/content/pdf/10.1007%2F978-3-319-99673-8_5.pdf) (дата обращения: 03.10.2019).

6. Gulenok A. A., Dordopulo A. I., Levin I. I., Gudkov V. A. Hybrid computer system programming technology with adaptation and scaling of calculations // Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering. 2017. Vol. 6. № 1. P. 73–86. DOI: 10.14529/cmse170105.

7. Бахтин В. А., Жукова О. Ф., Катаев Н. А., Колганов А. С., Крюков В. А., Поддерюгина Н. В., Притула М. Н., Савицкая О. А., Смирнов А. А. Автоматизация распараллеливания программных комплексов // Научный сервис в сети Интернет: труды XVIII Всероссийской научной конференции (19-24 сентября 2016 г., г. Новороссийск). – М.: ИПМ им. Келдыша М. В, 2016. С. 76–85. – doi: 10.20948/abrau-2016 – 31

8. Система САПФОР. URL:<http://www.keldysh.ru/dvm/SAPFOR/> (дата обращения 03.10.2019)

9. Открытая распараллеливающая система. URL: <http://www.ops.rsu.ru> (дата обращения 03.10.2019).

10. Сорокин Д. А., Дордопуло А. И., Левин И. И., Мельников А. К. Решение задач с существенно-переменной интенсивностью потоков данных на реконфигурируемых вычислительных системах // Вестник компьютерных и информационных технологий. 2012. № 2. С.49–56. DOI: 10.14489/issn.1810 – 7206

## **HIGH-PERFORMANCE RECONFIGURABLE COMPUTER SYSTEM BASED ON XILINX ULTRASCALE+ FPGAs**

*I. I. Levin, Yu. I. Doronchenko, A. I. Dordopulo, M. G. Levina*

ООО “Scientific research centre of supercomputers and neurocomputers”, Taganrog

In the paper we review an advanced reconfigurable computer system on the base of modern FPGAs and with an immersion subsystem for electronic components cooling. The system consists of stand-alone computational blocks. We present the architecture of the system basic component such as a computational module. Owing to the structure of data connections among computational modules and among blocks, it is possible to solve a wide class of resource-intensive problems with multiple data exchange. We show new approaches to power supply of FPGAs which consume up to 180 Watt each. Owing to the unique design of the computational block it is possible to achieve the computational density of up to 128 FPGAs in a 3U construction. We show advantages of an immersion cooling subsystem. Besides, we describe technical solutions for improvement of the cooling subsystem. Owing to such improvement, it is possible to dissipate up to 20 KWatt of total heat from one computational block. Also, we review an original complex of high-level language programming tools for multichip RCSs.

*Key words:* reconfigurable computer systems, performance of calculations, FPGA, immersion cooling systems, computational density, programming tools, high-level language.